local scaling. The anomalous component of the structure factor was calculated by XPREP, and Patterson correlation analysis was done by the program SHELXD[28] at a resolution of 4.5 Å. Because there was no clear occupancy drop for the sites, phase refinement was carried out with SHARP[29], which refined coordinates and occupancy but not temperature factors. By this approach, 25–27 selenium sites were confirmed out of an expected total of 39, leading to a map with visible helical density.

Phased molecular replacement using BRUTEPTF (http://russel.bioc.aecom.yu.edu/server/NYSGRC.html) located the intact Vt (Protein Data Bank (PDB) accession code 1QKR) and α-catenin dimerization (PDB 1DOV) domains, as well as one half of the α-catenin M-fragment (PDB 1H6G). Chain tracing for the remainder of the molecule was guided by the selenium positions. Refinement was done with the CNS package[30] against the maximum-likelihood target using amplitudes and phase probability distribution to 3.1 Å resolution. No σ cut-off was used in refinement. Convergence was reached at an R_{work} of 31.6% and an R_{free} of 35.7%. The current model includes residues 1–855 and 875–1,065. Of these, 87.4%, 10.9% and 1.5% are in the most favoured, additionally and generously allowed regions, respectively, of the Ramachandran plot. The structure was used as a molecular replacement search model for the other crystal forms. Refinement proceeded for each to R_{free} values of 0.35–0.40. Minor differences were observed in interhelical loops and the proline-rich region at sites of crystal contacts. Coordinates for the model of α-catenin are available from the authors on request.

Calorimetry

Differential scanning calorimetry (DSC) experiments were done at a scanning rate of 1 K min^{-1} under 3.0 atm of pressure on an N-DSC II differential scanning calorimeter (Calorimetry Sciences). Before measurements, all protein samples were dialysed against PBS buffer, which was used as the reference solution. All thermal transitions were irreversible under the conditions used. See Supplementary Methods for a further description of the DSC method. ITC was carried out on a VP-ITC calorimeter (Microcal). We injected 8-μl aliquots of solution containing either 940 μM α-catenin D3 or 1.0 mM talin peptide into a cell containing 70–100 μM vinculin D1–D3 construct (wild type or the Ala50Ile mutant). In each experiment 35 injections were made. The experiments were done at 23°C. Before ITC titrations, all protein samples were dialysed against PBS buffer. Experimental data were analysed with Microcal Origin software provided by the ITC manufacturer (Microcal).

Received 1 April; accepted 29 April 2004; doi:10.1038/nature02610.
Published online 13 June 2004.

1. Volberg, T. *et al.* Focal adhesion formation by F9 embryonal carcinoma cells after vinculin gene disruption. *J. Cell Sci.* **108,** 2253–2260 (1995).
2. Xu, W. M., Baribault, H. & Adamson, E. D. Vinculin knockout results in heart and brain defects during embryonic development. *Development* **125,** 327–337 (1998).
3. DeMali, K. A., Barlow, C. A. & Burridge, K. Recruitment of the Arp2/3 complex to vinculin: coupling membrane protrusion to matrix adhesion. *J. Cell Biol.* **159,** 881–891 (2002).
4. Johnson, R. P. & Craig, S. W. An Intramolecular association between the head and tail domains of vinculin modulates talin binding. *J. Biol. Chem.* **269,** 12611–12619 (1994).
5. Johnson, R. P. & Craig, S. W. F-actin binding site masked by the intramolecular association of vinculin head and tail domains. *Nature* **373,** 261–264 (1995).
6. Winkler, J., Lunsdorf, H. & Jockusch, B. M. The ultrastructure of chicken gizzard vinculin as visualised by high-resolution electron microscopy. *J. Struct. Biol.* **116,** 270–277 (1996).
7. Bakolitsa, C., de Pereda, J. M., Bagshaw, C. R., Critchley, D. R. & Liddington, R. C. Crystal structure of the vinculin tail suggests a pathway for activation. *Cell* **99,** 603–613 (1999).
8. Yang, J., Dokurno, P., Tonks, N. K. & Barford, D. Crystal structure of the M-fragment of α-catenin: implications for modulation of cell adhesion. *EMBO J.* **20,** 3645–3656 (2001).
9. Izard, T. *et al.* Vinculin activation by talin through helical bundle conversion. *Nature* **427,** 171–175 (2004).
10. Pokutta, S., Drees, F., Takai, Y., Nelson, W. J. & Weis, W. I. Biochemical and structural definition of the l-afadin- and actin-binding sites of α-catenin. *J. Biol. Chem.* **277,** 18868–18874 (2002).
11. Pokutta, S. & Weis, W. I. Structure of the dimerization and β-catenin-binding region of α-catenin. *Mol. Cell* **5,** 533–543 (2000).
12. Miller, G. J. & Ball, E. H. Conformational change in the vinculin C-terminal depends on a critical histidine residue (His-906). *J. Biol. Chem.* **276,** 28829–28834 (2001).
13. Brindle, N. P. J., Holt, M. R., Davies, J. E., Price, C. J. & Critchley, D. R. The focal adhesion vasodilator-stimulated phosphoprotein (VASP) binds to the proline-rich domain in vinculin. *Biochem. J.* **318,** 753–757 (1996).
14. Kioka, N. *et al.* Vinexin: a novel vinculin-binding protein with multiple SH3 domains enhances actin cytoskeletal organisation. *J. Cell Biol.* **144,** 59–69 (1999).
15. Huttelmaier, S. *et al.* The interaction of the cell-contact proteins VASP and vinculin is regulated by phosphatidylinositol-4,5-bisphosphate. *Curr. Biol.* **8,** 479–488 (1998).
16. Price, G. J. *et al.* Primary sequence and domain-structure of chicken vinculin. *Biochem. J.* **259,** 453–461 (1989).
17. Bass, M. D. *et al.* Further characterization of the interaction between the cytoskeletal proteins talin and vinculin. *Biochem. J.* **362,** 761–768 (2002).
18. Imamura, Y., Itoh, M., Maeno, Y., Tsukita, S. & Nagafuchi, A. Functional domains of α-catenin required for the strong state of cadherin-based cell adhesion. *J. Cell Biol.* **144,** 1311–1322 (1999).
19. Adey, N. B. & Kay, B. K. Isolation of peptides from phage-displayed random peptide libraries that interact with the talin-binding domain of vinculin. *Biochem. J.* **324,** 523–528 (1997).
20. Steimle, P. A., Hoffert, J. D., Adey, N. B. & Craig, S. W. Polyphosphoinositides inhibit the interaction of vinculin with actin filaments. *J. Biol. Chem.* **274,** 18414–18420 (1999).
21. Johnson, R. P., Niggli, V., Durrer, P. & Craig, S. W. A conserved motif in the tail domain of vinculin mediates association with and insertion into acidic phospholipid bilayers. *Biochemistry* **37,** 10211–10222 (1998).
22. Gilmore, A. P. & Burridge, K. Regulation of vinculin binding to talin and actin by phosphatidylinositol-4-5-bisphosphate. *Nature* **381,** 531–535 (1996).
23. Janssen, M. E. W. *et al.* Three-dimensional structure of vinculin bound to actin filaments. *J. Cell Biol.* (submitted).
24. Liddington, R. C. & Bankston, L. A. The structural basis of dynamic cell adhesion: heads, tails and allostery. *Exp. Cell Res.* **261,** 37–43 (2000).
25. Lim, W. A. The modular logic of signaling proteins: building allosteric switches from simple binding domains. *Curr. Opin. Struct. Biol.* **12,** 61–68 (2002).
26. Xing, B., Jedsadayanmata, A. & Lam, S. C. Localization of an integrin binding site to the C terminus of talin. *J. Biol. Chem.* **276,** 44373–44378 (2001).
27. O'Halloran, T., Molony, L. & Burridge, K. Purification and assay of vinculin, metavinculin, and talin. *Methods Enzymol.* **134,** 69–77 (1986).
28. Uson, I. & Sheldrick, G. M. Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **9,** 642–648 (1999).
29. de la Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276,** 472–494 (1997).
30. Brunger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank the Cell Migration Consortium, the NIH, the BBSRC and the Wellcome Trust for financial support; the DOE and staff at SSRL for synchrotron access and support; S. Lam for the talin rod expression vector; M. Zhang for talin expression; and B. Patel for vinculin purification. D.M.C. is supported by a Howard Hughes predoctoral fellowship.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.C.L. (rlidding@burnham.org). The atomic coordinates have been deposited in the PDB under accession codes 1ST6 (coordinates) and r1ST6sf (structure factors).

Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR

Dmitry M. Korzhnev[1,2]*, Xavier Salvatella[3]*, Michele Vendruscolo[3], Ariel A. Di Nardo[1], Alan R. Davidson[1], Christopher M. Dobson[3] & Lewis E. Kay[1,2]

[1]*Departments of Molecular and Medical Genetics and Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada*
[2]*Department of Chemistry and Protein Engineering Network Center of Excellence, University of Toronto, Toronto, Ontario M5S 1A8, Canada*
[3]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK*

* These authors contributed equally to this work

Many biochemical processes proceed through the formation of functionally significant intermediates[1,2]. Although the identification and characterization of such species can provide vital clues about the mechanisms of the reactions involved, it is challenging to obtain information of this type in cases where the intermediates are transient or present only at low population[1–4]. One important example of such a situation involves the folding behaviour of small proteins that represents a model for the acquisition of functional structure in biology[1]. Here we use relaxation dispersion nuclear magnetic resonance (NMR) spectroscopy to identify, for two mutational variants of one such protein, the SH3 domain from Fyn tyrosine kinase[5], a low-population folding intermediate in equilibrium with its unfolded and fully folded states. By performing the NMR experiments at different temperatures, this approach has enabled characterization of the kinetics and energetics of the folding process as well as providing structures of the intermediates. A general strategy emerges for an experimental determination of the energy landscape of a protein by applying this methodology to a series of

©2004 Nature Publishing Group

mutants whose intermediates have differing degrees of native-like structure.

Investigating the folding of small single-domain proteins has been crucial in understanding how biological molecules have evolved to self-assemble into complex functional structures[6–8]. SH3 domains are protein modules consisting of about 60 residues that fold into five-stranded β-sandwiches composed of two orthogonal β-sheets[5] (see Fig. 1a). The folding behaviour of several of these domains has been characterized in detail and found to be consistent with a two-state process; that is, there is no measurable population of partly structured intermediates[9,10]. Indeed, studies of this behaviour have provided important insights into the detailed mechanism of protein folding[11,12].

We have recently shown that mutations of the highly conserved glycine residue at position 48 of the Fyn SH3 domain both destabilize the protein and markedly increase the folding rate[13], k_f. We now use ^{15}N relaxation dispersion NMR spectroscopy[14] to probe the folding/unfolding reaction of two Fyn SH3 mutants, G48M and G48V, under native conditions. In these experiments, contributions to line widths of cross peaks in ^{1}H– ^{15}N correlation spectra are interpreted in terms of the rates of interconversion between states (kinetics), their chemical-shift differences (structure) and their populations (thermodynamics)[14,15]. For exchange rates varying from a few hundred to a few thousand per second and, for populations of the higher-energy (unfolded or intermediate) state greater than about 0.5%, this approach provides a sensitive measure of the exchange dynamics. Distinct advantages of the NMR experiments over global measurements such as those provided by circular dichroism or fluorescence are that site-specific folding information is available from each probe (in this case individual backbone amides) and that the experiments can be performed in the absence of denaturant.

Figure 1b, c (insets) illustrates typical ^{15}N Carr–Purcell–Meiboom–Gill (CPMG) dispersion profiles for Ala 12, a representative residue of Fyn SH3, analysed under the assumption of a two-state exchange process (solid lines). The kinetic parameters obtained are very similar to those extrapolated from unfolding/folding experiments in urea by stopped-flow fluorescence[13] (see Supplementary Table 1S), strongly supporting our interpretation that the increased line widths in NMR spectra arise from the folding reaction.

The high quality of the fits and the associated statistical analysis show that more complex models of exchange are not required to explain the experimental data when dispersion profiles are considered independently for the different residues. However, when k_f and unfolding rate (k_u) values measured for different probes (namely, different amides) are compared (Fig. 1b, c), it becomes clear that the exchange process is more complex for both mutants, particularly at temperatures below 20 °C. In the case of two-state folding, all of the folding/unfolding rates obtained should be the same within experimental error, because all probes are reporting on the same physical process. By contrast, if the folding mechanism is more complex, fits to a two-state exchange model are expected to produce residue-specific variations in the parameters describing the exchange process. The latter is true for many residues of the SH3 mutants at lower temperatures (Fig. 1b, c) with the differences in observed rates significantly larger than the error bounds. However, it is interesting to note that k_f (k_u) values for different residues converge with increasing temperature, so that the unfolding or folding process approximates to a two-state model at temperatures above about 25 °C for both G48M and G48V.

Recognizing that the exchange kinetics are more complex than two-state under at least some conditions, the dispersion data at all temperatures and for a significant set of residues with large dispersions were fitted simultaneously to a three-state exchange model that includes U, I and F states (see Methods). In this approach we assume that the chemical-shift differences between states are invariant with temperature, an assumption supported by previous work of Palmer *et al*[14], that the temperature dependence of the rates obeys standard transition-state theory[1] and that over the range of temperatures examined ΔH and ΔS are essentially constant (see Methods). Figure 2a, b shows examples of fits of the dispersions, here for Ser 41 and Thr 44 of the G48M mutant, based on the global analysis.

The temperature dependence of the populations of U, I and F, along with their rates of interconversion, are shown in Fig. 2c, d. The population of the I state is relatively independent of temperature, whereas that of the U state increases from 3% to 5% and from 1% to

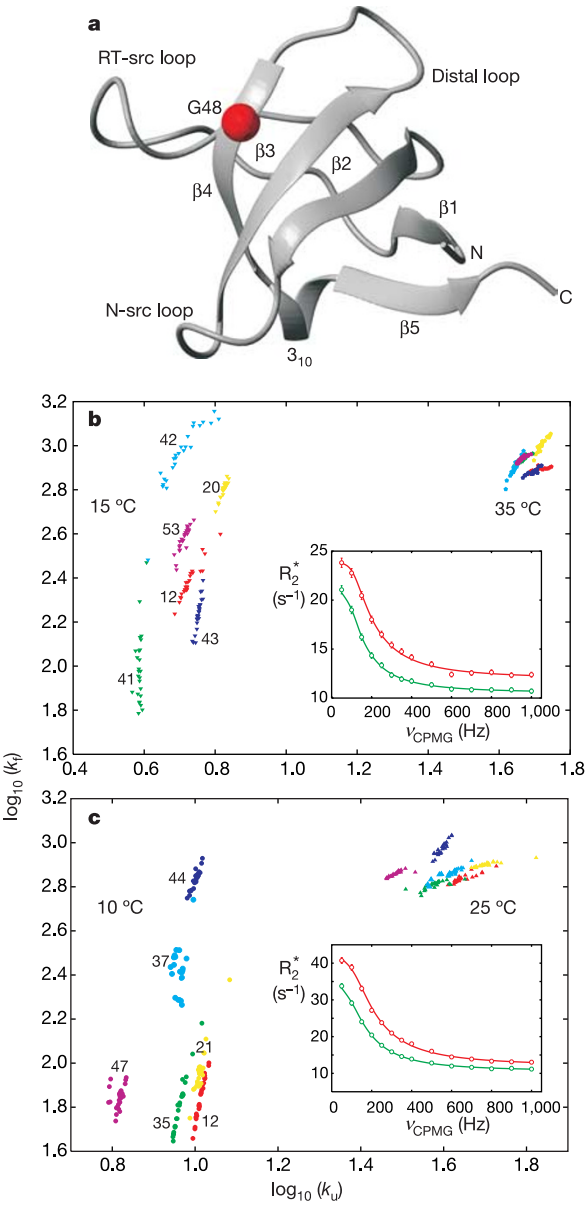

Figure 1 Measurement of k_f and k_u in the G48M and G48V Fyn SH3 domains. **a**, Ribbon diagram of the wild-type Fyn SH3 domain (pdb accession code 1SHF; ref. 27), highlighting the site of mutation, Gly 48. **b**, **c**, Values of k_f and k_u for selected residues (labelled with residue number) from fits of dispersions assuming a two-state model at 15 and 35 °C for G48M (**b**) and 10 and 25 °C (G48V) (**c**). Errors in rates were measured with a jacknife procedure by which 25% of the points in each dispersion were randomly removed; subsequent fits of the data produce the ranges of rates indicated. Insets show representative dispersion curves (for Ala 12) recorded at 800 MHz (red) and 600 MHz (green) at 25 °C, with R_2^* the measured relaxation rate and ν_{CPMG} the field strength employed.

©2004 Nature Publishing Group

5% for G48V and G48M, respectively, over the temperature range examined. In Fig. 2d we have plotted $k_{UI} + k_{IU}$ (labelled UI, where k_{ij} is the rate from state i to j) and $k_{IF} + k_{FI}$ (labelled IF) for both G48V and G48M. In particular, $k_{UI} + k_{IU}$ increases rapidly with temperature so that at higher temperatures the exchange between states U and I approaches the fast regime on the NMR chemical-shift timescale for both mutants, leading to a rapid equilibration between U and I.

This analysis allows relative values to be extracted for the thermodynamic parameters of each state for each mutant, yielding the results shown in Fig. 2c, where the U state is used as a reference (see Supplementary Table 3S). The G48M mutant is more stable than G48V, ΔG_{U-F} ($G_U - G_F$) = 2.3 kcal mol^{-1} (1 kcal ≈ 4.18 kJ) versus 1.8 kcal mol^{-1}, with the I state similar in free energy to the U state (within 0.5 kcal mol^{-1} in both cases). Notably, very similar ΔG_{U-F} values are obtained by fitting the data to a two-state model (within 0.2 kcal mol^{-1}), whereas the ΔH_{U-F} values extracted from the two models (two-state versus three-state) differ by less than about 25%. The potential interplay between H and S values[16], and the difficulties involved in deconvolving the contributions of the thermodynamic properties of the polypeptide chain from those of the solvent, means that their profiles must be interpreted more conservatively than those for G. Nevertheless, for both mutants $H_I \ll H_U$, an observation that is compatible with the formation of considerable structure in the I state. However, the larger enthalpy difference between states I and F for G48M ($\Delta H_{F-I} = -3.4$ kcal mol^{-1}) than for G48V ($\Delta H_{F-I} = 0.7$ kcal mol^{-1}) is particularly intriguing and indicates that the I state of G48V might be structurally more native-like than that of G48M (see below). In addition, this analysis indicates that the relative contribution of the entropic term to the free energy barrier for folding of Fyn SH3 is rather small, indicating that the expected decrease in chain entropy could be offset by the increased disorder of the solvent, particularly as the folding rates were measured at relatively low temperatures[1,17,18].

In addition to a kinetic and thermodynamic characterization of the three-site exchange process, structural information is also available from the dispersion experiments in the form of backbone ^{15}N chemical-shift differences between the various states. Figure 3a shows these differences, displayed in the form of a ratio,

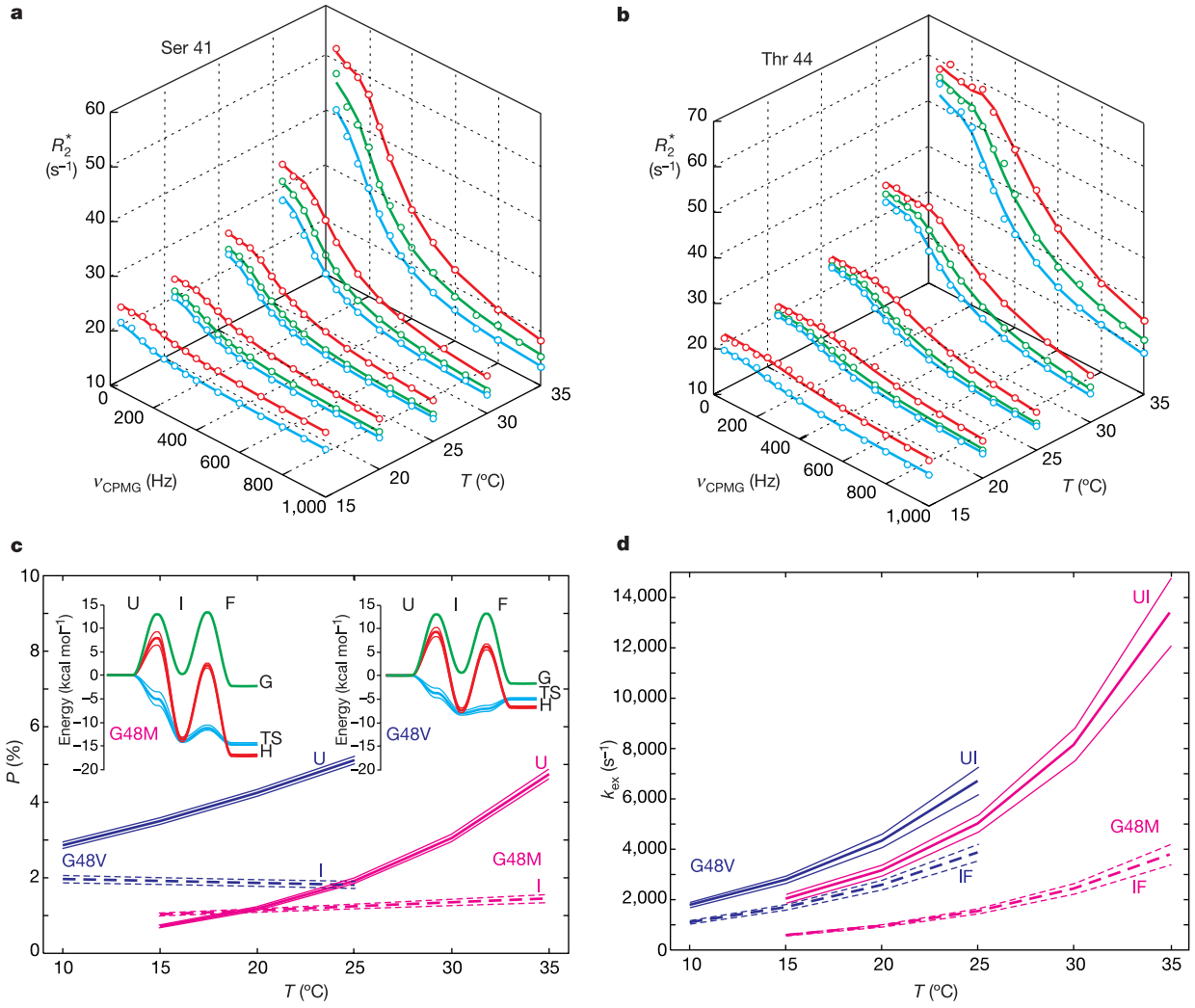

Figure 2 Three-site folding model for G48M and G48V Fyn SH3. **a**, **b**, Typical fits of dispersion data with a global three-state folding model for Ser 41 (**a**) and Thr 44 (**b**) of the G48M Fyn SH3 domain. Data from all residues, all temperatures and all spectrometer fields (blue, 500 MHz; green, 600 MHz; red, 800 MHz) were included in the analysis. **c**, **d**, Temperature dependence of populations (**c**) and rates of interconversion between states (**d**), along with energy diagrams for G48M (25 °C) and G48V (17.5 °C) are shown in the middle curve of each three, with the outer curves corresponding to errors of 1 standard deviation, estimated by the covariance matrix method[28]. The k_{ex} rates labelled UI and IF correspond to $k_{UI} + k_{IU}$ and $k_{IF} + k_{FI}$, respectively. Note that placement of the reference state (U) of each mutant in the energy-level diagrams is arbitrary; comparison of absolute parameters between mutants must therefore be made with caution.

©2004 Nature Publishing Group

$|\Delta^{exp}(k)| = |\delta_{FI}(k)/\delta_{FU}(k)|$, colour coded on the structure of the wild-type protein, where $\delta_{ij}(k)$ is the difference in ^{15}N shifts between states i and j for residue k. For a given residue, a ratio smaller than 0.5 indicates that the I state shift is closer to that of the folded state than to that of the unfolded state. An important validation of the analysis is the close similarity of the chemical-shift data for the U states of the two mutants (all chemical-shift data are provided in Supplementary Information). However, the chemical shifts of the I state in G48V seem on average to be significantly more native-like than the corresponding shifts in the G48M mutant ($\langle\Delta^{exp}\rangle_{G48M} = 0.61$, whereas $\langle\Delta^{exp}\rangle_{G48V} = 0.38$).

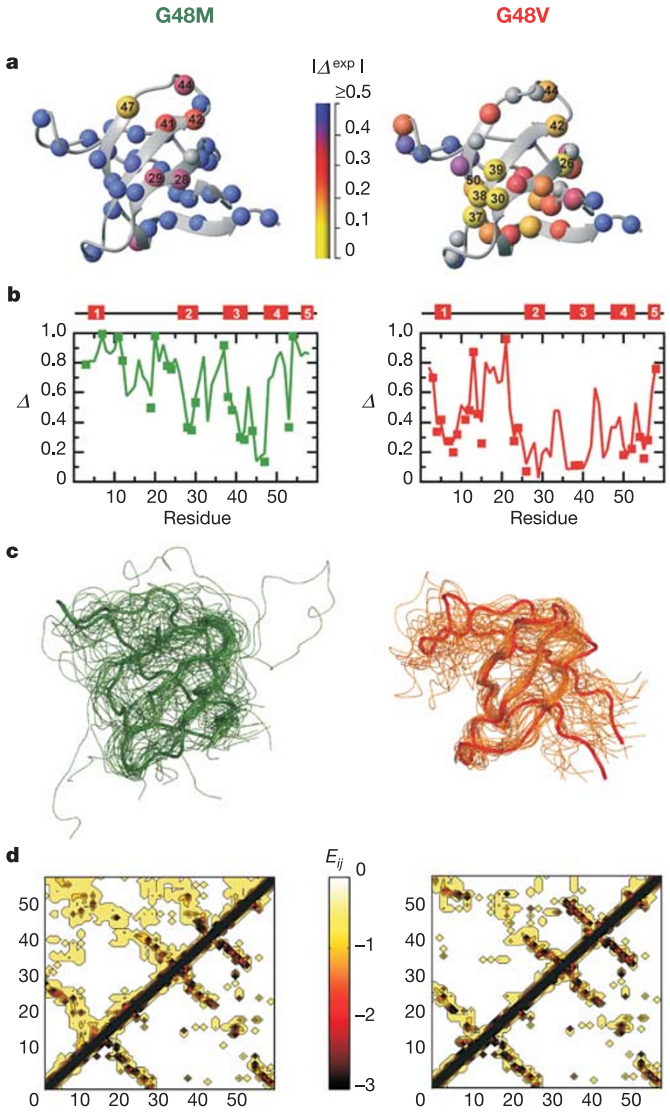

Figure 3 Structural analysis of the I state of mutants G48M and G48V of Fyn SH3 using chemical shifts. **a**, Ratio of ^{15}N chemical-shift differences, $|\Delta^{exp}(k)| = |\delta_{FI}(k)/\delta_{FU}(k)|$, plotted on the ribbon trace of the structure of the wild-type Fyn SH3 domain[27], with the colour code indicated. Residues with values of $\Delta^{exp} < -0.5$ or $\Delta^{exp} > 2.5$ are in grey. **b**, Comparison of Δ^{exp} (filled squares) with the Δ^{calc} values obtained in the structure determinations (lines); only residues with more than four native contacts have been represented. **c**, Ribbon representation of the ensemble of 25 structures determined for each mutant; the structure with the thicker ribbon is the more representative (lowest average pairwise r.m.s.d.) of the ensemble. **d**, Average effective energy contact map of the native state (bottom) and the ensembles obtained for the I state of each mutant (top); E_{ij} is the interaction energy (in kcal mol^{-1}) between the side chains of residues i and j according to the CHARMM19 force field. Residue numbers are indicated along each of the axes.

The measured chemical-shift differences were then used to calculate structural ensembles for the I state by using an approach similar to that originally developed to determine transition-state structures from protein engineering ϕ values[19]. Because most of the conformational factors that are known to influence ^{15}N shifts in native proteins can be related to the formation of inter-atomic interactions[20], in this approach we assume that the degree of unfolding of a given residue is related to Δ^{exp}, and computationally to the fraction of native contacts lost by the amide nitrogen of that residue in the I state, Δ^{calc}. A simulated annealing protocol developed to sample protein configurations for which $\Delta^{exp}(k) = \Delta^{calc}(k)$ (Fig. 3b) was used to produce an ensemble of 25 structures for the I state of each Fyn SH3 mutant.

The ensembles of conformations (Fig. 3c) and the ensemble-averaged energy maps (Fig. 3d) illustrate how in both intermediates the topology of the native state is already defined. Analysis of the calculated structures indicates that the most native-like region of both intermediates is the β-sheet formed by strands 2, 3 and 4 (the average root-mean-square deviation (r.m.s.d.) of residues in strand 4 between the intermediates and the native state is 3.7 and 1.9 Å for mutants G48M and G48V, respectively) and that their only significant structural differences with the native state are found in the connecting regions such as the RT loop and at the termini of the protein (details given in Supplementary Fig. 1S). Interestingly, several studies on the folding of Fyn SH3 domains have clearly established that the region of the native protein involving residues 45–50, which is part of strand 4, is formed almost completely in the transition state[9,11,12,21]. Given the fundamental differences between the two techniques used to probe the structure present in these states (ϕ values for transition states and ^{15}N chemical shifts for intermediates) it is remarkable that similar structural regions are identified in each case. However, the I states of the G48M and G48V mutants differ significantly in their structural heterogeneity (the average r.m.s.d. between two members of the ensemble is 5.6 Å in G48M and 3.3 Å in G48V) and in their degree of compaction. A recent study has shown a very strong correlation of folding kinetics with both β-sheet propensity and residue volume in a series of mutations at residue 48 (ref. 13); taken together, these results show that, although the acquisition of the native topology is the dominant structural event in the folding of Fyn SH3, the interactions established by the side chain at position 48 are very important in defining both the energetics and the degree of structure present at different stages along the folding reaction.

Thus, we have introduced here a novel methodology for the study of the mechanisms of protein folding and have applied it to characterize the behaviour of two destabilized, but fast-folding, mutants of the Fyn SH3 domain. Despite the apparent two-state nature of the folding and unfolding of this protein, we have shown that an intermediate can be characterized even though it is present at a level of only about 1% of the total population of molecules; because the folding of Fyn SH3 seems to be typical of small proteins, we believe that this approach will be widely applicable. Our results indicate that similar regions of the overall structure are present in the G48M and G48V intermediates, but also that these states differ in their degree of structural homogeneity and compaction owing to the crucial role of position 48 in the folding of Fyn SH3 (ref. 13). The two mutants that we have studied have allowed different stages of the folding landscape to be explored, indicating that a more complete picture could emerge from the investigation of a wider range of mutants whose intermediates are likely to span a greater portion of conformational space. This study is therefore an important step forwards in characterizing the detailed energy landscape of a folding reaction through the generation of extensive kinetic, thermodynamic and structural information. □

©2004 Nature Publishing Group

Methods

Materials

^{15}N and ^{15}N, ^{13}C-labelled mutants, G48M and G48V, were prepared as described previously[22]. Samples were 1.0 mM in protein, 50 mM sodium phosphate, 0.05% NaN_3, 0.2 mM EDTA, pH 7.

NMR methods

Chemical shifts of the mutant proteins were assigned by using standard triple-resonance NMR techniques[23] and were subsequently used to establish that each mutant had the same secondary structure as the wild type, by using the chemical-shift index[24]. Relaxation dispersion spectra were recorded at 600 and 800 MHz for G48V at temperatures of 10–25 °C in 5 °C intervals by methods that have been described previously[15,25]. Dispersion spectra of G48M were obtained at 15–35 °C in 5 °C steps at 500, 600 and 800 MHz, except at 15 °C where fields of 600 and 800 MHz only were employed. Three duplicate points for each dispersion were recorded for error analysis. Dispersion profiles for each residue were considered for subsequent computations with a set of criteria described in Supplementary Information. Dispersion curves were fitted analytically with either two-site or three-site exchange models by using software written in house. For each mutant, all dispersion profiles were fitted together, enforcing the relation $k = (k_b T/h)\exp(-\Delta G^{\dagger}/RT)$, where k is a rate constant, $k_b(h)$ is the Boltzmann (Planck) constant and $\Delta G^{\dagger}$ is the activation free energy. Note that the values for ΔG, ΔH and ΔS are independent of the prefactor used in the expression for k. Extensive minimizations with different initial conditions were performed to ensure that minimum-energy solutions to the fits of dispersion data were obtained. Extracted parameters from fits to the model U ↔ I ↔ F are k_{UI}, k_{IU}, k_{IF} and k_{FI} (at each temperature, from which populations can be calculated), chemical-shift differences between each pair of states (assumed to be temperature invariant[14]) and values of R_2* extrapolated to infinite CPMG repetition rate for each amide ^{15}N, at each spectrometer field and temperature. For all initial fits (and results in Fig. 2) ΔC_p values were set to 0, so that the extrapolated thermodynamic parameters correspond approximately to those at the temperature midpoint of the studies. The ΔG values extracted in this way are essentially independent of ΔC_p for the range of temperatures studied owing to the compensation between ΔH and ΔS. Additional fits were performed assuming $\Delta C_p = 300\ \mathrm{cal\ mol^{-1}\ K^{-1}}$, estimated from the slope of the plot of ΔH against T_m obtained for a large number of G48 mutants[13]. Here we assumed that the transition state between U and I, IU, and all subsequent states on the folding pathway have similar levels of compactness so that ΔC_p is attributed to differences between U and IU. Qualitatively, the energy profiles of Fig. 2c are very similar (details are given in Supplementary Table 3S).

Concentration-dependent studies have been performed on the G48M mutant, establishing that the presence of the third state in the exchange processes is not due to oligomerization (see Supplementary Fig. 2S).

Structure calculations

^{15}N chemical shifts were used as restraints for the calculation of ensembles representing the I state for G48M and G48V, with the use of an approach developed for the calculation of transition-state structures from protein engineering ϕ values, described previously[19,26]. The exact definition of native contact, the expression used to compute $N_i(k)$ and that used for the restraint violation term can be found in Supplementary Information, as well as a detailed description of the protocol used for the structure determination.

Received 8 January; accepted 17 May 2004; doi:10.1038/nature02655.

1. Fersht, A. *Structure and Mechanism in Protein Science* (W.H. Freeman and Company, New York, 1999).
2. Kahn, F., Chuang, J. I., Gianni, S. & Fersht, A. R. The kinetic pathway of folding of barnase. *J. Mol. Biol.* **333,** 169–186 (2003).
3. Sanchez, I. E. & Keifhaber, T. Hammond behavior *versus* ground state effects in protein folding: Evidence for narrow free energy barriers and residual structure in unfolded states. *J. Mol. Biol.* **327,** 867–884 (2003).
4. Klein-Seetharaman, J. *et al.* Long-range interactions within a nonnative protein. *Science* **295,** 1719–1722 (2002).
5. Pawson, T. & Gish, G. D. SH2 and SH3 domains: from structure to function. *Cell* **71,** 359–362 (1992).
6. Jackson, S. E. How do small, single domain proteins fold? *Fold. Design* **3,** R81–R91 (1998).
7. Dobson, C. M. Protein folding and misfolding. *Nature* **426,** 884–890 (2003).
8. Daggett, V. & Fersht, A. R. The present view of the mechanism of protein folding. *Nature Rev. Mol. Cell Biol.* **4,** 497–502 (2003).
9. Northey, J. G. B., Di Nardo, A. A. & Davidson, A. R. Hydrophobic core packing in the SH3 domain folding transition state. *Nature Struct. Biol.* **9,** 126–130 (2002).
10. Plaxco, K. W. *et al.* The folding kinetics and thermodynamics of the Fyn-SH3 domain. *Biochemistry* **37,** 2529–2537 (1998).
11. Martinez, J. C. & Serrano, L. The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved. *Nature Struct. Biol.* **6,** 1010–1016 (1999).
12. Grantcharova, V. P., Riddle, D. S. & Baker, D. Long-range order in the src SH3 folding transition state. *Proc. Natl Acad. Sci. USA* **97,** 7084–7089 (2000).
13. Di Nardo, A. A., Korzhnev, D. M., Stogios, P. J., Zarrine-Afsar, A., Kay, L. E. & Davidson, A. R. Dramatic acceleration of protein folding by stabilization of a non-native backbone conformation. *Proc. Natl Acad. Sci. USA* **101,** 7954–7959 (2004).
14. Palmer, A. G., Kroenke, C. D. & Loria, J. P. NMR methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **339,** 204–238 (2001).
15. Mulder, F. A. A., Mittermaier, A., Hon, B., Dahlquist, F. W. & Kay, L. E. Studying excited states of protein by NMR spectroscopy. *Nature Struct. Biol.* **8,** 932–935 (2001).
16. Dunitz, J. D. Win some, lose some. Enthalpy–entropy compensation in weak intermolecular interactions. *Chem. Biol.* **2,** 709–712 (1995).
17. Oliveberg, M., Tan, Y. J. & Fersht, A. R. Negative activation enthalpies on the kinetics of protein folding. *Proc. Natl Acad. Sci. USA* **92,** 8926–8929 (1995).
18. Dobson, C. M., Sali, A. & Karplus, M. Protein folding: A perspective from theory and experiment. *Angew. Chem. Int. Edn Engl.* **37,** 867–893 (1998).
19. Vendruscolo, M., Paci, E., Dobson, C. M. & Karplus, M. Three key residues from a critical contact network in a protein folding transition state. *Nature* **409,** 641–645 (2001).
20. Xu, X. P. & Case, D. A. Probing multiple effects on ^{15}N, ^{13}Cα, ^{13}Cβ, ^{13}C′ chemical shifts in peptides using density functional theory. *Biopolymers* **65,** 408–423 (2002).
21. Northey, J. G. B., Maxwell, K. L. & Davidson, A. R. Probing folding kinetics beyond the Phi value: Using multiple amino acid substitutions to investigate the structure of the SH3 domain folding transition state. *J. Mol. Biol.* **320,** 389–402 (2002).
22. Maxwell, K. L. & Davidson, A. R. Mutagenesis of a buried polar interaction in an SH3 domain: sequence conservation provides the best prediction of stability effects. *Biochemistry* **37,** 16172–16182 (1998).
23. Bax, A. Multidimensional nuclear magnetic resonance methods for protein studies. *Curr. Opin. Struct. Biol.* **4,** 738–744 (1994).
24. Wishart, D. S. & Sykes, B. D. The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR* **4,** 171–180 (1994).
25. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D. & Kay, L. E. Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* **123,** 11341–11352 (2001).
26. Paci, E., Vendruscolo, M., Dobson, C. M. & Karplus, M. Determination of a transition state at atomic resolution from protein engineering data. *J. Mol. Biol.* **324,** 151–163 (2002).
27. Noble, M. E., Musacchio, A., Saraste, M., Courtneidge, S. A. & Wierenga, R. K. Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *EMBO J.* **12,** 2617–2624 (1993).
28. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical Recipes in C* (Cambridge Univ. Press, 1988).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements This work was supported through grants from the Canadian Institutes of Health Research (A.R.D, L.E.K and D.M.K.), the European Comission (X.S.), the Royal Society (M.V.), the Leverhulme Trust (M.V. and C.M.D.) and the Wellcome Trust (C.M.D.). L.E.K. holds a Canada Research Chair in Biochemistry.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to L.E.K. (kay@pound.med.utoronto.ca) or C.M.D. (cmd44@cam.ac.uk).

©2004 Nature Publishing Group