

Running title: Reassessment of NOEs in an unfolded SH3 domain

Title: Reassessment of NOE data for the unfolded state of
an SH3 domain highlights residual β -structure

Karin A. Crowhurst^{a,b}, Wing-Yiu Choy^c, Yu-Keung Mok^d and Julie D. Forman-Kay^{a,b}

^a Department of Biochemistry, University of Toronto, Toronto, Ontario, M5S 1A8,
Canada

^b Structural Biology and Biochemistry Program, The Hospital for Sick Children,
555 University Avenue, Toronto, Ontario, M5G 1X8, Canada

^c Department of Medical Genetics and Microbiology, University of Toronto, Toronto,
Ontario, M5S 1A8, Canada

^d Department of Biological Sciences, National University of Singapore,
14 Science Drive 4, Singapore, 117543

Corresponding author: Julie D. Forman-Kay <forman@sickkids.ca>

Tel. 416-813-5358 Fax 416-813-5022

Total number of manuscript pages: 23 (including 3 pages of tables and figures)

Total number of figures: 2

Total number of tables: 1

Abstract

Preferential interactions in disordered states of proteins are important in folding initiation, the stability of the folded state and the roles disordered proteins play *in vivo*.

The N-terminal SH3 domain of the *Drosophila* protein drk (drkN SH3) exists in equilibrium between a folded (F_{exch}) and a significantly populated unfolded (U_{exch}) state under non-denaturing conditions, facilitating studies of preferential structure in the U_{exch} state by NMR methods. Previously described HN-HN NOE data were recently reassessed following experiments recorded at a higher spectrometer field. In the reanalysis described here, we have identified a large number of native-like NOEs in the U_{exch} state that cluster in regions corresponding to the RT and r-src loops, as well as between β -strands in the central β -sheet in the folded state. A group of non-native NOEs has also been observed, involving Trp 36, Tyr 37 and residues corresponding to the β 3 strand and the diverging turn (residues 20 – 23) in the folded state. Residual native-like structure in the loops suggests possible locations for initiation of protein folding within this domain and non-native interactions help explain previous evidence for the greater burial of the Trp 36 indole in the U_{exch} than the F_{exch} state.

Keywords: unfolded state, SH3 domain, residual structure, NOE, folding initiation

Disordered states of proteins were originally described as random coils with no residual structure and freedom of rotation around every bond (Tanford, 1968). Advances in experimental techniques for studying disordered proteins, particularly the increasing use of multinuclear NMR, have shown that even in the presence of high concentrations of strong denaturants polypeptide chains contain at least some degree of residual structure (Dill & Shortle, 1991; Schulman *et al.*, 1997; Wang & Shortle, 1995). In addition, many intrinsically unstructured proteins that exist in the absence of denaturants are being described, with a number of studies highlighting proteins that are unstructured in the absence of their binding partners (Wright & Dyson, 1999). It is also becoming clear that disordered states must be studied in order to fully understand the stability of the folded state. It has even been estimated that 30 % of eukaryotic proteins have unstructured regions of greater than 50 consecutive residues (Dunker *et al.*, 2001). Together, these underscore the importance and relevance of studying disordered states, specifically in identifying native-like and non-native interactions, as well as the functional role of disordered proteins *in vivo*.

Our group has focussed on the disordered states of the N-terminal SH3 domain of the *Drosophila* protein drk (drkN SH3 domain). Under non-denaturing *in vitro* conditions, the drkN SH3 domain exists in equilibrium between a folded (F_{exch}) and a highly populated unfolded (U_{exch}) state. These states exchange slowly on an NMR chemical shift timescale, such that resonances for the two states can be seen separately and therefore studied exclusive of the other. A large number of experiments have been performed on the U_{exch} state; results demonstrate a fairly compact state (Choy *et al.*, 2002; Farrow *et al.*, 1995; Mok *et al.*, 1999) with burial of the single Trp indole (Trp 36) (Crowhurst *et al.*, 2002; Mok *et al.*, 1999) and cooperative interactions within the region corresponding to the central β -sheet in the folded state (Crowhurst *et al.*, 2002).

In one previous study, we presented HN-HN NOE data recorded on a highly deuterated sample of the drkN SH3 domain (Mok et al., 1999). Many long-range NOEs were reported, providing evidence for a compact unfolded state ensemble that contained a significant amount of non-native structure. Recently the same NOE experiment has been recorded at 800 MHz. By comparing the two data sets it has become apparent that a portion of the previous data was misassigned. Due to the intrinsic challenges of analysing NOE data from an unfolded protein, with very low intensity or highly overlapped crosspeaks, a number of NOEs previously reported have now been attributed to spectral artifacts, while others were misassigned as long range backbone-backbone NOEs when in fact they were NOEs from backbone to nearby sidechain NH₂ groups (whose resonance positions were only recently identified). The data presented here will therefore constitute a revision of the previous report (Mok et al., 1999) and a summary of all current HN-HN NOE data observed for the U_{exch} state of the drkN SH3 domain.

Results and Discussion

HSQC-NOESY-HSQC experiments (Zhang *et al.*, 1997) were recorded at 5 °C at a field strength of 800 MHz on a highly deuterated sample of the drkN SH3 domain (Figure 1a and c). As described previously (Mok et al., 1999), a high level of deuteration was used to enhance the efficiency of NOE transfer for the remaining exchangeable NH protons in the molecule and to improve lineshape by removing proton-mediated dipolar relaxation pathways (LeMaster, 1989; Mok et al., 1999). Additionally, a 600 ms mixing time and low temperature (5 °C) were utilised to further facilitate the observation of weak, longer range NOEs in specific conformations within the unfolded state

ensemble (Pachter *et al.*, 1992). Refer to the Materials and Methods section for sample and experimental details.

Comparison of the 800 MHz and 600 MHz HSQC-NOESY-HSQC spectra

As is apparent from Figure 1, the new HSQC-NOESY-HSQC spectrum recorded at 800 MHz is significantly better resolved than the previously described data recorded at 600 MHz (Mok *et al.*, 1999), enabling much more comprehensive assignment. In the original analysis of the HN-HN NOE data, Mok *et al.* (Mok *et al.*, 1999) were forced to exclude data from highly overlapped regions, such as the area surrounding Phe 19 shown in Figure 1b. This enhanced peak definition in the 800 MHz experiment is due to two crucial factors: first, the higher spectrometer field, and second, a different method of processing the raw data. The F_1 nitrogen dimension was linear predicted for the 600 MHz data but not the 800 MHz data, while both were linear predicted in the F_2 nitrogen dimension. Successful linear prediction relies on a high signal-to-noise ratio (S/N), therefore in regions of a NOESY spectrum where weak peaks exist (due to long-range or weak interactions between two atoms), linear prediction can, in some cases, make weak peaks even less intense, and conversely, make some noise look like real peaks (Stern *et al.*, 2002). Close analysis of the 600 MHz data reveals that the additional linear prediction adds significant noise (as seen in Figure 1b), as well as “false peaks”. This helps in part to explain why the lower field experiment displays some peaks that are not observable in the other experiment (an example of this is not shown). After taking this into account, however, there are still some extra peaks in the 600 MHz experiment, which could possibly be explained by a higher level of deuteration or a higher concentration than reported of the NMR sample. Our reanalysis of the data considered all of these differences between the two experiments. The variations in resolution and

data processing often allowed us to assess more accurately the assignment of particular peaks, and to discard data that were deemed to arise from spectral artifacts.

Use of Pf1 phage to minimise exchange-NOEs

In recent years, Pf1 filamentous phage has become popular for use as orienting media to partially align molecules for measurement of dipolar couplings (Hansen *et al.*, 1998). However, when phage is added to a solution sample of the drkN SH3 domain under non-denaturing buffer conditions (50 mM sodium phosphate, pH 6), the ratio of folded to unfolded conformations is significantly shifted towards the unfolded state. Since there are no measurable changes in the chemical shifts of resonances in the unfolded state conformation, we have concluded that the phage has likely destabilised the F_{exch} state conformation while having little effect on the overall ensemble of structures in the U_{exch} state. Since filamentous phage has an overall negative charge, it is possible that this repels the cluster of negative surface charges in the F_{exch} state, thereby destabilising the F_{exch} state and causing the equilibrium shift.

Although the folded and unfolded states exchange slowly on an NMR chemical shift timescale allowing peaks for each state to be resolved separately, there are applications for which the selective stabilisation of one state over the other is desired. It has been possible to stabilise the folded state by adding 0.4 M sodium sulphate to the sample (Zhang *et al.*, 1994), however only recently have we found a means of stabilising the U_{exch} state in order to record NMR experiments in the absence of the folded state. The advantage of removing folded state peaks for NOE experiments is twofold: (i) simplification of spectra from decreased overlap with U_{exch} peaks and (ii) elimination of the “exchange-NOE” peaks arising when the transition from one conformational state to another occurs during NOE transfer. Although the folded state is essentially unobservable in samples of uniformly protonated drkN SH3 domain in phage, the

sharpness and intensity of the peaks in spectra recorded on deuterated samples prevents the complete elimination of these peaks when phage is added, although they are significantly weaker and broader than the unfolded state peaks (see Figure 1c). However, the unfolded state peaks also show line broadening in the HN dimension. This may be due to some alignment of the protein in the unfolded state, which would then cause broadening in the experiment due to dipolar coupling interactions.

“Exchange-NOEs” can be observed in an NOE experiment with two highly populated conformational states in equilibrium. In this case, we observe not only NOE crosspeaks between amide protons in the U_{exch} state (U_1 ? U_2) and between amide protons in the F_{exch} state (F_1 ? F_2) if the protons are close in both states, but also crosspeaks between amide protons in each state (U_1 ? F_2 or F_1 ? U_2). These “exchange-NOE” crosspeaks are the result of either NOE transfer followed by conformational exchange, or the reverse. In other words, a U_1 ? F_2 crosspeak can be due to U_1 ? U_2 NOE transfer, followed by exchange (U_2 to F_2), or due to exchange first (U_1 to F_1), followed by F_1 ? F_2 NOE transfer. In the first case, the NOE is between amide protons in the unfolded state, while in the latter case, the NOE occurs within the folded state structure. A more detailed discussion of exchange-NOEs can be found in Mok *et al.* (Mok *et al.*, 1999). In the past, there was no way to differentiate between the two possibilities, thus requiring this data to be excluded. Despite the poor resolution and residual overlap of F_{exch} with U_{exch} peaks, analysis of the HSQC-NOESY-HSQC experiment recorded in phage can help discriminate, from the intensity and shape of the peak, whether or not a particular NOE is a result of chemical exchange to the F_{exch} state or a genuine through-space interaction within the U_{exch} state. Consequently, we were able to include several NOEs that would have otherwise been discarded (see items marked with “*” in Table 1).

Native-like short- and medium-range NOEs in the U_{exch} state

Table 1 summarises the NOEs in the U_{exch} state that could be unambiguously assigned, ranging from $(i, i+3)$ to $(i, i+18)$ distance restraints. Interactions that were listed in the original long-range HN-HN NOE paper by Mok *et al.* (Mok *et al.*, 1999) are indicated in Table 1 with “#”. In this table, we also report the shorter-range non-ambiguous $(i, i+3)$ NOEs because they are not seen uniformly throughout the sequence; in fact, in combination with the $(i, i+4)$ and $(i, i+5)$ NOEs they help to illustrate how many U_{exch} state contacts are clustered in areas that correspond to loop regions in the folded state structure, particularly around the n-src and RT loops (see Figure 2a). Short- and medium-range NOEs should not be disregarded when analysing interactions in disordered states, particularly if they cluster in a non-random manner, as in this case.

Numerous groups have also observed residual structure in the loop and turn regions of disordered proteins. For example, Shortle *et al.* (Wang & Shortle, 1995) have reported that in $\Delta 131\Delta$ (a large, partially folded fragment of staphylococcal nuclease) the most stable secondary structural element is a β -hairpin that is located between the $\beta 2$ and $\beta 3$ strands; this hairpin exists even in high concentrations of denaturant. They argue that its stability is due to the burial of a large amount of nonpolar surface area when this hairpin is formed. NMR experiments on the F30H mutant of the acid denatured state of protein G (F30H-G_{B1}) (Sari *et al.*, 2000) show that the most stable residual structure is the native-like $\beta 3$ - $\beta 4$ hairpin, and that residual structure is also present in the $\beta 1$ - $\beta 2$ turn, although some non-native turns are also detected. A recent study of mutants of bovine pancreatic trypsin inhibitor supported the hypothesis that the native-like structure observed in the unfolded state was important in folding, specifically an interaction in the 25-28 turn which is believed to initiate folding (Li *et al.*, 2002). Finally, in our own research, previous NOE experiments that probed some short-range HN-HN

interactions had already highlighted the potential of turn regions in the drkN SH3 domain to be involved in folding initiation (Zhang & Forman-Kay, 1997). The authors of these papers have suggested that there is a connection between stable residual structure in turn regions of the disordered states and its role in the initiation of protein folding, possibly as a nucleation site. Indeed, this concept is supported by Baker's hypothesis that in smaller protein domains folding is highly dependent on topology (Baker, 2000) and that local interactions, such as those within a turn region, are more entropically favoured at early stages in folding than interactions made between residues that are distant along the sequence.

Longer-range NOEs in the U_{exch} state

The longer-range NOEs in the U_{exch} state of the drkN SH3 domain that are not associated with interactions within turns cluster primarily in two areas. Native-like interactions are observed between residues corresponding to the central β -sheet, particularly between the $\beta 5$ and $\beta 6$ strands, in the folded state (see Figure 2b). The observation of NOEs in this region supports earlier evidence for interactions in the central β -sheet of the U_{exch} state that are cooperatively interrupted by the addition of denaturants (Crowhurst et al., 2002). Furthermore, molecular dynamics unfolding simulations indicate the central β -sheet to be one of the most stable regions of the domain (Philippopoulos, M., Forman-Kay, J.D. & Pomès, R., *manuscript in preparation*).

Longer-range NOEs in the U_{exch} state are also observed in non-native interactions between Trp 36 / Tyr 37 and residues Ser 18, Thr 22 and Gln 23, which are located in strand $\beta 3$ and the diverging turn in the folded state. These observations support previous data that indicate greater burial of the Trp 36 indole in the U_{exch} state

than in the folded state where it is located on the surface of the protein (Crowhurst et al., 2002; Mok et al., 1999; Yang *et al.*, 1999). In order to more fully explore hydrophobic clustering in the unfolded state, we are pursuing experiments to probe NOE interactions between aromatic sidechains in the U_{exch} state.

ENSEMBLE calculations

We have developed a software program called ENSEMBLE (Choy & Forman-Kay, 2001) that uses experimental data to generate ensembles of structures representing the unfolded state. The original HN-HN NOE data from Mok *et al.* (Mok et al., 1999) were used as one set of restraints in the first ENSEMBLE calculation performed on the drkN SH3 domain. A recalculation using the new HN-HN NOE data list yielded no observable change in the average structural properties including compactness of the unfolded state ensemble of conformations (*data not shown*). This was expected since the hydrodynamic radius provides the most influential restraint in both calculations, diminishing the impact of other experimental data included. Nonetheless, the presence of stabilised residual structure in the central β -sheet and turn regions of the most highly populated conformers of both ENSEMBLE runs demonstrates that the newly reassessed HN-HN NOE data are more compatible with the calculated ensemble than the previous data (Mok et al., 1999) where many of the long range NOEs were inconsistent with the major conformers predicted by ENSEMBLE.

Comparison to the denatured state ensemble (DSE) of the α -spectrin SH3 domain

Kortemme et al. (Kortemme *et al.*, 2000) have performed HSQC-NOESY-HSQC experiments on a uniformly ^2H , ^{15}N -labelled chicken α -spectrin SH3 domain that is denatured under mildly acidic conditions (pH 2.2). Medium- and long-range NOEs in this

denatured state ensemble (DSE) were found to cluster primarily in the loop and turn regions, particularly in the distal β -hairpin, with many NOEs from the two sequential Trp indoles at positions 41 and 42 (analogous to Trp 36 and Tyr 37 in the drkN SH3 domain) and within the central β -sheet. In comparing their DSE to the U_{exch} state of the drkN SH3 domain, the authors at that time had difficulties in reconciling the fact that, while the two conformational ensembles had a similar degree of compactness, they differed greatly in their numbers of long-range NOEs (Kortemme et al., 2000; Mok et al., 1999). In light of our reassessment, a comparison of the NOE data for the α -spectrin and drkN SH3 domains now indicates significant similarities between the two unfolded states. Thus, despite only 25 % sequence identity, these two proteins that share the same fold also share similar structural characteristics in their unfolded state ensembles. There are differences, however: i) non-native contacts between residues near the diverging turn and the aromatics in β 5 are not observed in the DSE of α -spectrin; ii) when focussing on contacts in turn regions, the vast majority of NOEs are in the distal hairpin of the DSE of α -spectrin, while they cluster more prominently in regions corresponding to the n-src and RT loops in the unfolded state of the drkN SH3 domain. When comparing the two domains, the differences in the number of contacts in the n-src and RT loop regions as well as the non-native long-range interactions that may stabilise helical structure near the diverging turn could help to explain why the unfolded state of the drkN SH3 domain is much more stable than that of α -spectrin; perhaps this combination of stabilised loops and non-native helical structure are more crucial to the stability of the unfolded structure of the drkN SH3 domain than the residual interactions that are common between the two unfolded ensembles.

Summary

Our reassessment of HN-HN NOE data has highlighted several areas of residual structure that are maintained in the unfolded state of the drkN SH3 domain. Residues associated with the loops and β -strands in the central β -sheet are involved in native-like interactions, agreeing with previous experiments that suggest the U_{exch} state maintains some residual structure in this region (Choy & Forman-Kay, 2001; Crowhurst et al., 2002; Zhang & Forman-Kay, 1997). Of the 51 U_{exch} state NOEs listed, about half of them reflect residual native-like turn structure in the U_{exch} state as they are found between residues within or adjacent to loops which may act as nucleation sites in domain folding initiation (Zhang & Forman-Kay, 1997). One specific cluster of non-native interactions was also observed, involving NOEs between Trp 36 / Tyr 37 and a region corresponding to strand β 3 and the diverging turn in the folded state. These interactions are consistent with the greater burial of the Trp 36 indole in the U_{exch} than the folded state within the drkN SH3 domain and possible stabilisation of non-native helical structure.

Materials and Methods

Sample preparation

All isotopically labelled reagents were purchased from Cambridge Isotope Labs.

All 800 MHz NMR experiments employed a uniformly $^{13}\text{C} / ^{15}\text{N} / ^2\text{H}$ -labelled sample with ^1H -labelled δ 1 methyl isoleucines and ^1H -labelled δ methyl leucines. Please note that the selective labelling strategy described for the sample was utilised for a different set of experiments (Crowhurst, K.A. & Forman-Kay, J.D., *manuscript in preparation*) and does not have any bearing on this discussion. A 1.5 mM sample in sodium phosphate (50 mM, pH 6.0, 10 % D_2O) and a 2.5 mM sample in sodium phosphate (50 mM, pH 6.0, 10 % D_2O) were used for experiments recorded without and

with phage, respectively. The latter sample was dissolved in Pf1 phage liquid-crystalline medium (^2H splitting = 24 Hz). Experiments at 600 MHz were recorded on a uniformly ^{15}N / ^2H -labelled, 1.5 mM drkN SH3 domain in sodium phosphate (50 mM, pH 6, 10 % D_2O); details of sample preparation have been described previously (Mok et al., 1999).

HN-HN NOE experiments

NMR experiments recorded at 800 MHz. ^1H - ^{15}N HSQC-NOESY-HSQC experiments (Zhang et al., 1997) were recorded at 5°C on a Varian Inova 800 MHz spectrometer, equipped with a pulsed-field gradient triple resonance probe, utilised matrices of 64 x 32 x 814 complex points corresponding to an acquisition time of 64 ms, a mixing time of 600 ms and were acquired with spectral widths of 1783.0, 1783.0 and 12775.5 Hz (F_1 , F_2 and F_3). 24 scans were acquired for each FID. The ^1H and ^{15}N carrier frequencies were placed at 4.99 and 119.11 ppm, respectively, for experiments without phage, and at 5.00 and 119.22 ppm, respectively, for experiments with phage. Both ^1H and ^{15}N chemical shifts were referenced to 0.2 mM external DSS (directly for ^1H and indirectly for ^{15}N chemical shifts) (Wishart *et al.*, 1995). Data were processed and analysed on a Linux-based PC using NMRPipe / NMRDraw (Delaglio, 1995; Delaglio *et al.*, 1995) and NMRView (Johnson & Blevins, 1997; Johnson & Blevins, 1994) software, respectively. For the F_2 ^{15}N dimension the size of the time domain was doubled using forward-backward linear prediction (Zhu & Bax, 1992). A 65°-shifted sine-bell window function was applied to all three dimensions. After zero filling twice in each dimension and extraction of the data to retain 6.5 ppm to 10.5 ppm in the acquisition dimension, 3D data sets of 256 x 256 x 816 real points were obtained.

Summary of parameters for previously reported experiments recorded at 600 MHz. The ^1H - ^{15}N HSQC-NOESY-HSQC experiment (Zhang et al., 1997) was recorded at 5 °C on a

Varian Inova 600 MHz spectrometer equipped with a z-axis pulsed field gradient unit and an actively shielded triple resonance probe. Matrices of 64 x 32 x 576 complex points were acquired with an acquisition time of 64 ms, a mixing time of 600 ms and spectral widths of 1700.0, 1700.0 and 9009.9 Hz (F_1 , F_2 and F_3). 24 scans were acquired for each FID. For the ^{15}N dimensions (F_1 and F_2) the size of the time domain was doubled using forward-backward linear prediction (Zhu & Bax, 1992). A 65° -shifted sine-bell window function was applied to the F_1 and F_3 dimensions, while a 60° -shifted sine-bell window function was applied to F_2 . After zero filling and extraction of the data to retain 6.7 ppm to 10.3 ppm in the acquisition dimension, this 3D data set contained 256 x 128 x 496 real points. All other details of this experiment have been described previously (Mok et al., 1999).

Acknowledgements

The authors wish to sincerely thank Dr. Ranjith Muhandiram for NMR assistance and Dr. Lewis Kay for thoughtful input during manuscript preparation. K.A.C. and W.-Y.C. would like to acknowledge the Canadian Institutes of Health Research for a Doctoral Research Award and a Senior Research Fellowship, respectively. This work was supported by funds from the Medical Research Council of Canada / Canadian Institutes of Health Research (J.D.F.-K.).

References

- Baker, D. 2000. A surprising simplicity to protein folding. *Nature* **405**: 39-42.
- Choy, W.-Y. & Forman-Kay, J. D. 2001. Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J. Mol. Biol.* **308**: 1011-1032.
- Choy, W.-Y., Mulder, F. A. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D. & Kay, L. E. 2002. Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. *J. Mol. Biol.* **316**: 101-112.
- Crowhurst, K. A., Tollinger, M. & Forman-Kay, J. D. 2002. Cooperative interactions and a non-native buried Trp in the unfolded state of an SH3 domain. *J. Mol. Biol.* **322**: 163-178.
- Delaglio, F. (1995). *NMRDraw*. v.1.8, NIH Laboratory of Chemical Physics, Bethesda, MD USA.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277-293.
- Dill, K. A. & Shortle, D. 1991. Denatured states of proteins. *Annu. Rev. Biochem.* **60**: 795-825.
- Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C. & Obradovic, Z. 2001. Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**: 26-59.

- Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. 1995. Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry-US* **34**: 868-878.
- Guex, N. & Peitsch, M. C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714-2723.
- Hansen, M. R., Mueller, L. & Pardi, A. 1998. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**: 1065-1074.
- Johnson, B. A. & Blevins, R. (1997). *NMRView*. v.4.1.2, Merck and Co., Whitehouse Station, NJ USA.
- Johnson, B. A. & Blevins, R. A. 1994. NMR View - a computer-program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**: 603-614.
- Kortemme, T., Kelly, M. J. S., Kay, L. E., Forman-Kay, J. & Serrano, L. 2000. Similarities between the spectrin SH3 domain denatured state and its folding transition state. *J. Mol. Biol.* **297**: 1217-1229.
- LeMaster, D. M. 1989. Deuteration in protein proton magnetic resonance. *Method. Enzymol.* **177**: 23-43.
- Li, R., Battiste, J. L. & Woodward, C. 2002. Native-like interactions favored in the unfolded bovine pancreatic trypsin inhibitor have different roles in folding. *Biochemistry-US* **41**: 2246-2253.
- Mok, Y.-K., Kay, C. M., Kay, L. E. & Forman-Kay, J. 1999. NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* **289**: 619-638.
- Pachter, R., Arrowsmith, C. H. & Jardetzky, O. 1992. The effect of selective deuteration on magnetization transfer in larger proteins. *J. Biomol. NMR* **2**: 183-194.

- Sari, N., Alexander, P., Bryan, P. N. & Orban, J. 2000. Structure and dynamics of an acid-denatured Protein G mutant. *Biochemistry-US* **39**: 965-977.
- Schulman, B. A., Kim, P. S., Dobson, C. M. & Redfield, C. 1997. A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nat. Struct. Biol.* **4**: 630-634.
- Stern, A. S., Li, K.-B. & Hoch, J. C. 2002. Modern spectrum analysis in multidimensional NMR spectroscopy: Comparison of linear-prediction extrapolation and maximum-entropy reconstruction. *J. Am. Chem. Soc.* **124**: 1982-1993.
- Tanford, C. 1968. Protein denaturation. *Adv. Protein Chem.* **23**: 121-282.
- Wang, Y. & Shortle, D. 1995. The equilibrium folding pathway of Staphylococcal Nuclease: Identification of the most stable chain-chain interactions by NMR and CD spectroscopy. *Biochemistry-US* **34**: 15895-15905.
- Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. & Sykes, B. D. 1995. ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**: 135-140.
- Wright, P. E. & Dyson, H. J. 1999. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**: 321-331.
- Yang, D., Mok, Y.-K., Muhandiram, D. R., Forman-Kay, J. D. & Kay, L. E. 1999. ^1H - ^{13}C Dipole-dipole cross-correlated spin relaxation as a probe of dynamics in unfolded proteins: Application to the drkN SH3 domain. *J. Am. Chem. Soc.* **121**: 3555-3556.
- Zhang, O. & Forman-Kay, J. D. 1997. NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry-US* **36**: 3959-3970.
- Zhang, O., Forman-Kay, J. D., Shortle, D. & Kay, L. E. 1997. Triple-resonance NOESY-based experiments with improved spectral resolution: Applications to structural

- characterization of unfolded, partially folded and folded proteins. *J. Biomol. NMR* **9**: 181-200.
- Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. 1994. Backbone ^1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**: 845-858.
- Zhu, G. & Bax, A. 1992. Two-dimensional linear prediction for signals truncated in both dimensions. *J. Magn. Reson.* **98**: 192-199.

Table 1. Unambiguous NH-NH NOEs ($=i, i-3$) observed in the U_{exch} state of the highly deuterated drkN SH3 domain using a mixing time of 600 ms.

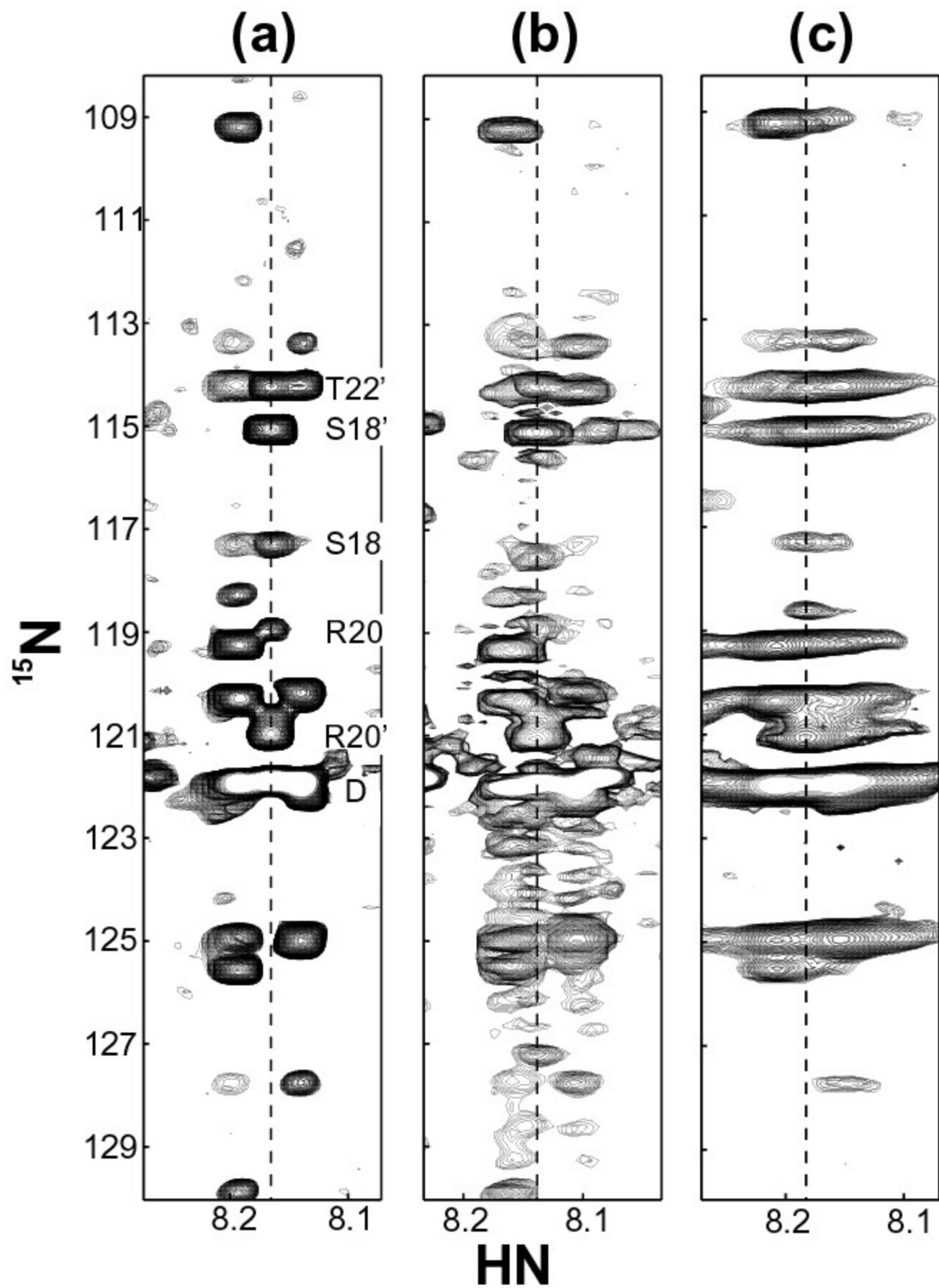
Atom 1	Atom 2	Range	Atom 1	Atom 2	Range
D8 #	A11	3	D14 #	S18	4
F9 #	T12	3	L25	N29 δ	4
T12 #	D15	3	N29 δ	D33	4
A13 #*	E16	3	E31	N35 δ	4
D14 #	L17	3	D32	W36 ϵ	4
D15 #	S18	3	S34 #*	R38	4
F19	T22	3	N35 δ	A39	4
T22	L25	3	W36 ϵ	E40	4
N29	D32	3	A39	G43	4
N29 δ	D32	3	D42 #	G46	4
E31 #	S34	3	S50 #	E54	4
D32	N35 δ	3	N51 δ	M55	4
D33 #	W36 ϵ	3	A13 #	S18	5
S34 #*	Y37	3	N29 δ	S34	5
N35	R38	3	M30	N35 δ	5
N35 δ	R38	3	E31	W36 ϵ	5
W36	A39	3	W36 ϵ	L41	5
W36 ϵ	A39	3	N29 δ	N35	6
E40 #	G43	3	N29	N35 δ	6
I48	N51 δ	3	N29	W36	7
S50 #	I53	3	A39 *	G46	7
N51 δ	E54	3	Y37 #	G46	9
E54	N57 δ	3	T22 #	W36 ϵ	14
M55 #	H58	3	Q23 ϵ	Y37	14
T12 #*	E16	4	S18 #	W36 ϵ	18
A13 #	L17	4			

NOEs are sorted in ascending order according to the range of their correlation. Items marked with # agree with previously reported data (from Mok *et al.* ⁹), and items marked with * correspond to NOEs verified using data from the sample in phage. Asn residues denoted with " δ " refer to NOEs to the H ^{δ 2} proton, and NOEs to the Gln H ^{ϵ 2} or Trp H ^{ϵ 1} proton are denoted with " ϵ ". All other NOEs are to backbone amide protons.

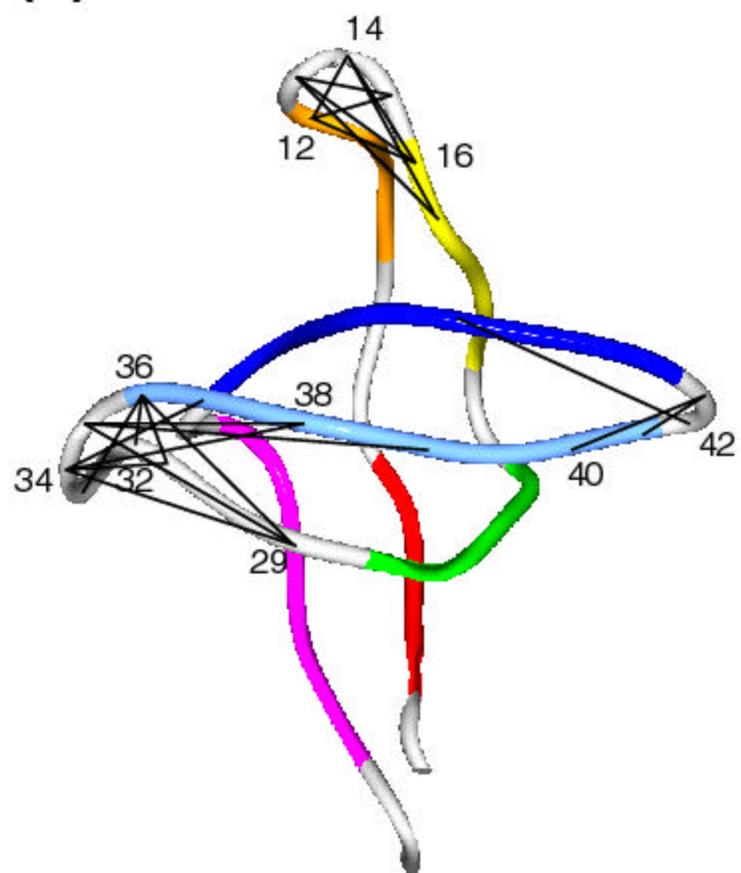
Figure Legends

Figure 1. Strips of ^1H - ^{15}N HSQC-NOESY-HSQC spectra (Zhang et al., 1997) taken at the ^1H and ^{15}N chemical shifts of residue Phe 19 in the U_{exch} state of the drkN SH3 domain. NOEs to Phe 19 lie along the dotted line. Peaks not lying on the line correspond to other NOE strips (for Leu 41 to the left of the line, and for Ile 24 to the right). A prime following a residue number denotes an NOE from the U_{exch} state, while an unprimed number indicates an NOE from the F_{exch} state. D denotes the diagonal peak. (a) Sample in sodium phosphate (50 mM, pH 6.0), 800 MHz; (b) Sample in sodium phosphate (50 mM, pH 6), 600 MHz sample (Mok et al., 1999); (c) Sample in sodium phosphate (50 mM, pH 6.0) and Pf1 phage liquid-crystalline medium, 800 MHz.

Figure 2. Selected HN-HN NOEs observed in the U_{exch} state of the drkN SH3 domain, depicted on a ribbon representation of the folded state structure (Singer, A.U. *et al.*, *unpublished results*). The residues corresponding to the seven β -strands of the folded state are coloured as follows: β 1, red; β 2, orange; β 3, yellow; β 4, green; β 5, light blue; β 6, indigo; β 7, magenta. (a) Primarily local HN-HN NOEs illustrating the predominance of NOEs in the RT and n-src loop, and, to a lesser extent, the distal hairpin regions. (b) Medium-and long-range HN-HN NOEs showing the maintenance of native-like interactions in the β 5 and β 6 strands in the central β -sheet, as well as non-native interactions primarily involving Trp 36 / Tyr 37, the β 3 strand, and the RT loop. The figure was generated using Swiss-PdbViewer v3.7b2 (Guex & Peitsch, 1997) and rendered using POV-Ray for Windows v3.1g.



(a)



(b)

