Structural Comparison of the Unstable drkN SH3 Domain and a Stable Mutant[†,‡]

Irina Bezsonova,[§,||] Alex Singer,[§,#] Wing-Yiu Choy,[||,⊥,¶] Martin Tollinger,[§,£] and Julie D. Forman-Kay*[,§,⊥]

The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada, M5G 1X8, and Departments of Chemistry and Biochemistry, University of Toronto, Toronto, Ontario, Canada, M5S 1A8

Received July 4, 2005; Revised Manuscript Received October 3, 2005

ABSTRACT: The N-terminal SH3 domain of the *Drosophila* adapter protein Drk (drkN SH3 domain) is marginally stable ($\Delta G_U = 1$ kcal/mol) and exists in equilibrium between folded and highly populated unfolded states. The single substitution T22G, however, completely stabilizes the protein ($\Delta G_U = 4.0$ kcal/mol). To probe the causes of instability of the wild-type (WT) protein and the dramatic stabilization of the mutant, we determined and compared nuclear magnetic resonance structures of the folded WT and mutant drkN SH3 domains. Residual dipolar coupling (RDC) and carbonyl chemical-shift anisotropy (C′-CSA) restraints measured for the WT and T22G domains were used for calculating the structures. The structures for the WT and mutant are highly similar. Thr22 of the WT and Gly22 of the mutant are at the $i + 2$ position of the diverging, type-II β-turn. Interestingly, not only Gly22 but also Thr22 successfully adopt an α_L conformation, required at this position of the turn, despite the fact that positive ϕ values are energetically unfavorable and normally disallowed for threonine residues. Forcing the Thr22 residue into this unnatural conformation increases the free energy of the folded state of the WT domain relative to its T22G mutant. Evidence for residual helix formation in the diverging turn region has been previously reported for the unfolded state of the WT drkN SH3 domain, and this, in addition to other residual structure, has been proposed to play a role in decreasing the free energy of the unfolded state of the protein. Together these data provide evidence that both increasing the free energy of the folded state and decreasing the free energy of the unfolded state of the protein contribute to instability of the WT drkN SH3 domain.

The *Drosophila* adapter protein Drk is a 23-kDa protein containing a central Src homology 2 (SH2) domain surrounded by two Src homology 3 (SH3)[1] domains. Drk mediates signaling between receptor tyrosine kinases and *ras* G proteins by the binding of its SH3 domains to proline-rich, hydrophobic regions of the guanine nucleotide exchange factor Sos (*1*−*3*). SH3 domains, which are 50−70 residues in length, are found in a large number of proteins involved in signal transduction and cellular localization (*4*). The isolated N-terminal SH3 domain of Drk (drkN SH3), unlike other known SH3 domains, is marginally stable and exists in equilibrium between folded (F_{exch}) and highly populated unfolded (U_{exch}) states under non-denaturing buffer conditions (*5*), with a ratio of folded/unfolded close to 1:1. The interconversion between the two states is slow on the nuclear magnetic resonance (NMR) chemical-shift time scale, giving rise to distinct sets of resonances, making the drkN SH3 domain an excellent model for the study of disordered states, protein folding, and stability (*5*−*11*).

Sequential alignment of a large number of SH3 domains revealed a highly conserved glycine in the diverging β-turn (residues 20−23 in the drkN SH3 domain), which links unpaired β-strands (*12*). However, the drkN SH3 domain has a threonine at this position (Thr22). Thermodynamic and kinetic studies of the drkN SH3 domain and various mutants have been performed (*13*) to understand the nature of the instability of the drkN SH3 domain. A single substitution of threonine in position 22 to glycine leads to dramatic stabilization of the domain, resulting in an increase of the melting temperature by 20 °C with respect to the WT protein. This mutant folds 7 times faster and unfolds 15 times slower than the WT. The difference in unfolding free energy between the T22G mutant and WT under salt-stabilizing conditions is 2.4 kcal/mol on the basis of GdmCl denaturation data. Evidence for non-native structure in the unfolded state of the drkN SH3 domain based on experimental NOE data (*14*, *15*) and on predictions using the algorithm

[†] This research was supported by grants from the Canadian Institutes of Health Research (to J.D.F.-K.). I.B. is the recipient of funding from the Canadian Institutes of Health Research Strategic Training Program in Protein Folding: Principles and Diseases.

[‡] The coordinates of the WT and T22G structures have been deposited in the Protein Data Bank, accession codes: 2A36 and 2A37 (with NOEs) and 2AZS and 2AZV (without NOEs), respectively.

* To whom correspondence should be addressed. Telephone: 1 (416) 813-5358. Fax: 1 (416) 813-5022. E-mail: forman@sickkids.ca.

[§] The Hospital for Sick Children.

[||] Department of Chemistry, University of Toronto.

[⊥] Department of Biochemistry, University of Toronto.

[#] Present address: Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, NC, 25799.

[¶] Present address: Department of Biochemistry, University of Western Ontario, London, Ontario, Canada, N6A 5C1.

[£] Present address: Institute for Theoretical Chemistry and Structural Biology, University of Vienna, Dr.-Bohr-Gasse 9, A-130 Vienna, Austria.

[1] Abbreviations: SH3 domain, Src homology 3 domain; drkN SH3 domain, N-terminal SH3 domain of the *Drosophila* adapter protein Drk; WT, wild type; F_{exch}, folded state under low-salt conditions that allow exchange with the highly populated unfolded state; U_{exch}, unfolded state under non-denaturing conditions that allow exchange with the highly populated folded state; NMR, nuclear magnetic resonance; RDC, residual dipolar coupling; C′-CSA, carbonyl chemical-shift anisotropy; rmsd, root-mean-square deviation.

10.1021/bi0512795 CCC: $30.25
Published on Web 11/05/2005

AGADIR (*16*, *17*), in conjunction with this kinetic data, led to the proposal that stabilization of the T22G mutant is, at least in part, due to the disruption of non-native helical structure in the wild-type (WT) unfolded state in residues corresponding to the diverging turn region.

A detailed understanding of the energetic components of protein stability is a major challenge. While it is clear from theoretical principles that both the folded and unfolded states must contribute to the overall energetics, much experimental work has focused only on the folded state. Specific conclusions, particularly when comparing WT and mutant proteins, have often been drawn based on the structure of the folded state alone (*18*, *19*). Increasingly, however, studies in which both states have been probed to understand the causes of protein stability are being reported (*20*−*24*). We have investigated detailed structural properties of both folded and unfolded states of the drkN SH3 domain. In fact, much of our previous work has focused on the unfolded state, and we have found that at least part of the instability of the drkN SH3 domain is a result of decreasing the free energy of the unfolded, U_{exch}, state of the domain. The structure of the folded drkN SH3 domain, however, has not been published before. In the present work, we probe the role of the folded state in the stability of the domain, by determining and comparing the structures of the WT and T22G mutant drkN SH3 domains. Five sets of residual dipolar coupling (RDC) restraints (in contrast to the 1−3 sets often utilized) and a set of carbonyl chemical-shift anisotropy (C′-CSA) restraints (*25*), which provide structural information analogous to that given by RDC restraints, have been chosen as a rich source of data for WT and T22G folded structure determinations. The large number of RDCs combined with C′-CSA data is particularly useful for determining local backbone geometry, appropriate given our focus on the structural differences at position 22. Our results suggest that both increasing the free energy of the folded state and decreasing the free energy of the unfolded state contribute to the instability of the WT drkN SH3 domain.

MATERIALS AND METHODS

Sample Preparation. A plasmid coding for the isolated WT drkN SH3 domain, residues 1−59 of the protein Drk (*5*), under the control of the T7 promoter was transfected into *Escherichia coli* HMS 174 cells. The expression of the double-labeled ^{15}N, ^{13}C protein was induced for 4 h at an OD_{600} of 0.6 by addition of 250 mg/L IPTG to bacterial growths at 37 °C in M9 minimal medium, supplemented with 0.3% ^{13}C-labeled glucose, 0.1% $^{15}NH_4Cl$, 100 mg/L ampicillin, 10 mg/L thiamine, 10 mg/L biotin, and 1 mM of each of $MgSO_4$ and $CaCl_2$. Cells were lysed by sonication in 50 mM Tris, 2 mM EDTA, 5 mM benzamidine HCl, and 7 mM β-mercaptoethanol. The drkN SH3 domain was purified on a DE 52 ion-exchange column with a linear gradient of NaCl (0−1 M) followed by Superdex 75 gel-filtration column in 0.15 M NaCl, 50 mM Tris, 2 mM EDTA, 5 mM benzamidine HCl, and 7 mM β-mercaptoethanol, followed by a Mono Q ion-exchange column with a linear gradient of NaCl (0−0.3 M). The T22G mutant (*13*) was expressed and purified in a manner similar to WT. The identity and purity of the proteins were confirmed by mass spectrometry.

The NMR sample of the WT drkN SH3 domain contained 1.0 mM protein in 50 mM sodium phosphate at pH 6.0, 90% H_2O/10% D_2O, 3 mM DSS, and 0.5 M Na_2SO_4 for stabilization of the folded state of the domain. The primary NMR sample of the T22G mutant contained 1.0 mM protein in 50 mM sodium phosphate at pH 6.0, 90% H_2O/10% D_2O, and 3 mM DSS. An additional sample of the T22G mutant also contained 0.5 M Na_2SO_4. Samples were oriented using pf1 phage prepared as described by Hansen and co-workers (*26*).

Thermodynamic Stability. The equilibrium denaturation experiment on the T22G mutant was performed on an Aviv ATF 105 ratio spectrofluorimeter at 25 °C in a 1 cm pathlength quartz cell. The sample contained 5 μM T22G drkN SH3 domain in 50 mM phosphate buffer. Denaturation was performed in steps of 0.2 M GdmCl from 0 to 6.0 M GdmCl. Renaturation started at 6.8 M and ended at 0.6 M GdmCl. The samples were equilibrated at the appropriate GdmCl concentrations for 5 min prior to measurements. A wavelength of 280 nm was used for excitation, and a fluorescence emission intensity at 360 nm was measured as a function of the denaturant concentration. The denaturation and renaturation curves could be well-fit to the equation for a two-state transition

$$F_{obs} = F_F - (F_F - F_U)\frac{\exp(m[\mathrm{GdnCl}] - \Delta G_{H_2O})/RT}{1 + \exp(m[\mathrm{GdnCl}] - \Delta G_{H_2O})/RT}$$

where F_{obs} is the observed fluorescence, F_F and F_U are the fluorescence intensities of the folded and unfolded states extrapolated to a 0 M concentration of GdmCl, respectively, and ΔG_{H_2O} is the free energy of unfolding in water (or in the phosphate buffer) (*27*). The *m* value is related to the increase in the degree of the surface exposure upon denaturation.

NMR Spectroscopy. Spectra were recorded at 20 °C on four-channel Varian Inova 500 and 600 MHz spectrometers equipped with *z*-axis pulsed field gradient units and actively shielded triple resonance probes. Data sets were processed with NMRPipe software (*28*) and analyzed with PIPP (*29*) and NMRView software (*30*). Backbone ^1H, ^{15}N, and ^{13}C resonance assignments for the WT protein and the T22G mutant and side-chain resonance assignments for the WT domain were obtained from standard double- and triple-resonance experiments (*31*, *32*).

Interproton distance restraints were obtained from multidimensional NOE spectra recorded with mixing times ranging from 50 to 200 ms. NOE experiments (*33*) included a 2D 200 ms NOESY and a 3D 150 ms ^{15}N-edited NOESY recorded in H_2O and a 2D 150 ms NOESY and a 2D 50 ms NOESY recorded in D_2O. The 50 ms NOESY experiment in D_2O has been used for stereospecific assignment of β-methylene protons. NOE experiments were recorded only for the WT protein.

Hydrogen-bond restraints for the WT drkN SH3 domain were obtained from a $^1H/D$ amide exchange experiment. Corresponding restraints for the T22G mutant were obtained from a CLEANEX experiment (*34*). A (CLEANEX-PM)−HSQC spectrum was recorded with a mixing time of 24.6 ms, and an HSQC spectrum was used as a reference spectrum. Peaks that were present in the reference spectrum but absent in the (CLEANEX-PM)−HSQC spectrum were considered to be hydrogen-bond donors. Hydrogen-bond acceptors were determined using the previously calculated

WT structure (A. U. Singer, O. Zhang, and J. D. Forman-Kay, unpublished results).

HNCO-based 3D experiments (non-TROSY) described in detail by Yang and co-workers (*35*) have been used for measuring ^{1}HN-^{15}N, ^{15}N-^{13}C′, and ^{13}C′-^{13}C$^{\alpha}$ RDCs. ^{13}C$^{\alpha}$-^{13}C$^{\beta}$ and ^{13}C$^{\alpha}$-^{1}H$^{\alpha}$ RDCs (*36*, *37*) were also measured. An HNCO experiment has been used as previously described (*25*) for measuring C′-CSA. All RDC and C′-CSA experiments were carried out in the absence and presence of pf1 phage (*26*). For sample alignment, ~6 mg of phage were added to the T22G mutant drkN SH3 domain in low salt (deuterium splitting, 10.6 Hz) and ~16 mg of phage were added to the WT protein (deuterium splitting, 12.1 Hz) and T22G sample in high-salt conditions for similar deuterium splitting. The difference in the amount of phage required for similar net alignments is explained by the difference in ionic strengths between the low- and high-salt (0.5 M Na_2SO_4) conditions. The effect of pf1 phage as an alignment agent is partly electrostatic, and therefore, a high concentration of salt in the sample lowers the efficiency of alignment by pf1 phage.

Three bond H^{N}-H$^{\alpha}$ *J*-coupling constants were measured for both WT and T22G samples using 3D ^{15}N-separated *J* correlation spectra as a function of the diagonal-peak/cross-peak intensity ratio as previously described (*38*). H$^{\alpha}$ spin flip rates were estimated using 2D ^{13}C CT-HSQC experiments and used to correct the H^{N}-H$^{\alpha}$ *J* couplings (*39*). Three bond H$^{\alpha}$-C′ *J*-coupling constants were measured only for the WT domain to distinguish unambiguously between positive and negative ϕ angle values for the Thr22 residue after inconclusive H^{N}-H$^{\alpha}$ *J*-coupling results. The values of $^{3}J_{\mathrm{H}\alpha\mathrm{C}'}$ were obtained by recording a 3D HN(CO)HB quantitative *J* correlation spectrum (*40*, *41*).

Estimation of Helical Content of the U_{exch} State. The population of conformers within the U_{exch} ensemble with a helical structure from residues 18−26 (α) was estimated on the basis of the comparison of the U_{exch} H$^{\alpha}$ chemical shifts of residues 18−26 to the residue-specific chemical shifts of a random coil and α helix (*42*) as

$$\alpha = \frac{1}{9}\left(\sum_{i=18}^{26} \frac{\delta_{\mathrm{rc}(i)} - \delta_{\mathrm{u}(i)}}{\delta_{\mathrm{rc}(i)} - \delta_{\mathrm{h}(i)}}\right)$$

where α is the fraction of conformers of the U_{exch} state in a helical conformation and δ_u, δ_h, and δ_{rc} are the chemical shifts in the U_{exch} state, helix, and random coil, respectively (*43*).

Structure Calculations. Initial structure calculations for the WT and T22G mutant of the drkN SH3 domain were based on restraints from five sets of RDC, C′-CSA, hydrogen-bonding, backbone dihedral angle, and ^{3}J-coupling constant experiments. NOE-derived distance restraints determined for the WT protein were used later to refine both WT and T22G structures. The upper bound of NOE-derived distance restraints was increased by 0.5 Å for both WT and T22G calculations. NOE restraints used for the T22G structure calculation excluded data corresponding to the diverging turn region (residues 19−24).

Backbone dihedral angle restraints were derived from ^{15}N, ^{13}C$^{\alpha}$, ^{13}C$^{\beta}$, and ^{13}C′ chemical shifts using TALOS (*44*). Restraints consisting of the average ϕ and ψ values ±2 standard deviations or at least ±20° from the average predicted values were employed for 28 and 22 residues of the WT and T22G mutant of the drkN SH3 domain, respectively. The ϕ and ψ TALOS restraints for residues in the loops and turns were not included into the structure calculations for WT or T22G (excluded residues were 6−8, 11−17, 20−24, 30−35, and 42−43). In addition, experimentally measured (*38*, *40*, *41*) ϕ angle restraints were introduced into the structure calculations, in the form of direct refinement against $^{3}J_{\mathrm{HNH}\alpha}$ and $^{3}J_{\mathrm{H}\alpha\mathrm{C}'}$.

Five sets of RDCs were used as restraints, including ^{1}HN-^{15}N, ^{15}N-^{13}C′, ^{13}C$^{\alpha}$-^{13}C′, ^{13}C$^{\alpha}$-^{1}H$^{\alpha}$, and ^{13}C$^{\alpha}$-^{13}C$^{\beta}$. The ^{15}N-^{13}C′, ^{13}C$^{\alpha}$-^{13}C′, ^{13}C$^{\alpha}$-^{1}H$^{\alpha}$, and ^{13}C$^{\alpha}$-^{13}C$^{\beta}$ couplings were normalized to the magnitude of ^{1}HN-^{15}N couplings so that similar deviations between each class of measured and calculated couplings contribute equally to the RDC pseudo-energy during the structure refinement. The parameters D_a and R, the axial and rhombic components of the alignment tensor, respectively, were obtained directly from the experimental data by examining the distribution of RDCs (*45*), yielding values of D_a and R of 9.65 Hz and 0.59 for the WT and −15.32 Hz and 0.33 for the mutant, respectively. ^{13}C′ chemical shifts were measured using HNCO experiments, and changes in ^{13}C′ chemical shifts upon alignment (Δ_C) were calculated from the difference in the carbonyl chemical shifts between aligned and unaligned samples. Because, in the case of aligned samples, the spectrometer is locked on one of the two ^{2}H lines (frequencies $\omega_D \pm D_D$), while in the case of unaligned samples, it is locked on the ^{2}H singlet resonating at ω_D, all ^{13}C shifts are offset by a constant amount in addition to the CSA-dependent offset. This constant offset was estimated by minimizing the difference between experimental and predicted Δ_C values using lowest energy structures calculated on the basis of NOE, dihedral angle, hydrogen-bond, and RDC restraints. An offset of −83.9 ppb for WT and −76.1 ppb for the mutant protein was subtracted from all measured Δ_C values prior to their use in structure refinement.

Structures were calculated using CNS software version 1.0 (*46*) starting from extended conformations and using a combination of torsion angle dynamics and Cartesian dynamics, with C′-CSA restraints incorporated into the calculation as previously described (*25*). All parameters for structure calculations were used as previously described (*25*), except the values of force constants for RDC and C′-CSA restraints. The force constants for RDC restraints were scaled from an initial value of 0.015 to a final value of 0.30 kcal Hz^{-2}, while the force constants for C′-CSA values were increased from the initial value of 0.00015 to the final value of 0.003 kcal ppb^{-2} during the second torsion angle dynamics stage in the structure calculations. Typically, a set of 100 structures was generated, and the 10 lowest energy structures were selected for structural analysis.

RESULTS AND DISCUSSION

Thermodynamic Stabilities of the WT and T22G drkN SH3 Domains. To investigate the change in stability between the WT drkN SH3 domain and its T22G mutant, we determined the free-energy difference between folded and unfolded states (ΔG_U) for the T22G mutant at 20 °C. Denaturation by GdmCl was monitored using intrinsic tryptophan fluorescence, and the resulting denaturation and renaturation curves

could be well-fit to the equation for a two-state transition (see the Materials and Methods). From denaturing and renaturing experiments, nearly identical values of $\Delta G_U = 4.0 \pm 0.2$ kcal/mol (with an m value of 2.2 ± 0.5 kcal mol^{-1} M^{-1}) are obtained. It was shown earlier (*9*) that folding and unfolding rate constants of the WT drkN SH3 domain at 20 °C are 1.5 ± 0.1 and 0.7 ± 0.1 s^{-1}, respectively, corresponding to a ΔG_U of 0.4 ± 0.1 kcal/mol. Thus, the difference in unfolding free energy, $\Delta\Delta G_U$, between WT and T22G domains is 3.6 ± 0.2 kcal/mol, a remarkable stabilization for a single mutation. We have previously determined (*13*) ΔG_U values for the WT and T22G domains of 1.8 ± 0.5 and 4.2 ± 0.2 kcal/mol, respectively, in the presence of 0.4 M Na_2SO_4, which stabilizes the folded state of the WT.

NOE and RDC Measurements. Assignments of backbone and side-chain resonances of the WT drkN SH3 domain were made and NOE-derived distance restraints were measured on the WT sample stabilized with 0.5 M Na_2SO_4 (F_s state) as described in the Materials and Methods. These restraints were used for preliminary WT structure calculations (A. U. Singer, O. Zhang, and J. D. Forman-Kay, unpublished results). To determine the WT and T22G mutant structures, we measured five sets of RDCs, including ^{1}HN-^{15}N, ^{15}N-^{13}C′, ^{13}C′-^{13}C$^\alpha$, ^{13}C$^\alpha$-^{1}H$^\alpha$, and ^{13}C$^\alpha$-^{13}C$^\beta$ dipolar couplings and the C′-CSA for both WT and T22G drkN SH3 domain samples (see the Materials and Methods).

The NMR sample used to determine the structure of the stabilized folded state of the WT domain contained 0.5 M Na_2SO_4, yielding a single set of resonances, while the T22G NMR sample did not require salt. To ensure that addition of 0.5 M Na_2SO_4 does not affect the structure of the folded state of the domain and to facilitate a comparison with the T22G mutant in the absence of salt, we have also measured ^{1}HN-^{15}N RDCs for the T22G mutant in the presence of 0.5 M Na_2SO_4. Because a direct comparison of the two sets of RDC/C′-CSA restraints is not possible due to the different molecular alignment frames of the two proteins, we compared RDC restraints measured for the T22G mutant in the presence of 0.5 M Na_2SO_4 with RDCs predicted on the basis of the final T22G mutant structure in the absence of Na_2SO_4 (see below). The results are shown in Figure 1, where red circles represent the correlation between ^{1}HN-^{15}N RDCs measured in high-salt buffer and RDCs back-calculated from the final T22G mutant structure in low-salt buffer. Filled circles represent residues of the diverging turn. Black circles on the plot represent, as a control, the correlation between experimental RDCs recorded in low-salt buffer with RDC values that were back-calculated from the structure obtained from low-salt experimental data [root-mean-square deviation (rmsd) = 2.9 ± 0.3 Hz]. If the structure of the protein in high- and low-salt buffer conditions remains the same, the black circles and red circles should show a similar degree of deviation between experimental and predicted values of RDCs. Figure 1 reveals that RDC restraints measured in 0.5 M Na_2SO_4 are in good agreement with restraints measured in the absence of high salt (rmsd = 3.12 ± 0.18 Hz). This suggests that the addition of 0.5 M Na_2SO_4 to the sample does not affect the structure of the protein.

Structure Determination. Classical NMR structure determination requires extensive assignment of backbone and side-chain resonances followed by unambiguous identification of

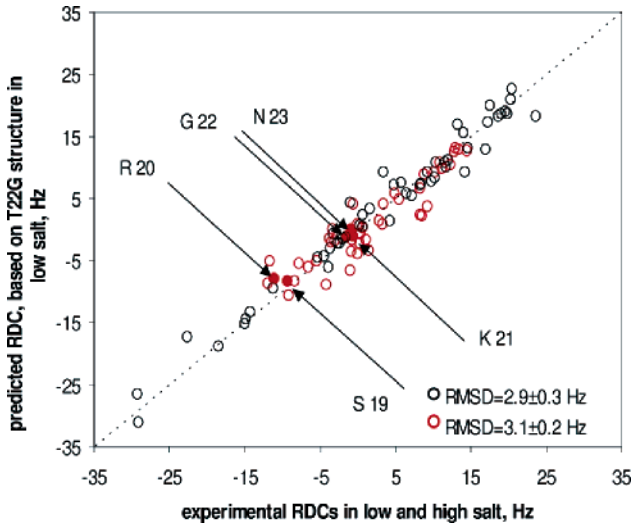

FIGURE 1: Comparison of ^{1}HN-^{15}N RDC data recorded for the T22G drkN SH3 domain in the presence and absence of 0.5 M Na_2SO_4. Red circles represent the correlation between RDC data recorded in the presence of 0.5 M Na_2SO_4 (experimental RDCs) and those predicted on the basis of the 10 lowest energy structures calculated using RDC data recorded without Na_2SO_4 (predicted RDCs). Filled circles represent residues of the diverging turn. Black circles represent the control correlation between ^{1}HN-^{15}N RDCs that are observed (experimental) and those back-calculated from the 10 lowest energy structures (predicted), using the same RDC data recorded in the absence of Na_2SO_4.

nuclear Overhauser effect (NOE) correlations between these assigned frequencies to yield semiquantitative short-range (<6 Å) distance restraints. This process is often very time-consuming. RDC/C′-CSA restraints, on the other hand, can be measured routinely, immediately following the backbone resonance assignment. They provide distance-independent structural data related to the angle between an internuclear vector and the magnetic field and serve as an ideal tool for molecular structure refinement and comparison of homologues (*47*).

Initial structure calculations were performed using as structural restraints five sets of RDC (^{1}HN-^{15}N, ^{15}N-^{13}C′, ^{13}C′-^{13}C$^\alpha$, ^{13}C$^\alpha$-^{1}H$^\alpha$, and ^{13}C$^\alpha$-^{13}C$^\beta$), C′-CSA, hydrogen-bond, and dihedral angle measurements recorded independently for each of the WT and T22G domains. The quality of structures calculated with RDC restraints depends upon the force constant value (k_{dip}) for the dipolar restraints used in the calculations. To find the optimal force constant, we performed several structure calculations varying the value of k_{dip} from 0.1 to 0.6 kcal Hz^{-2}. The resulting structures showed that increasing the force constant to greater than 0.3 kcal Hz^{-2} led to significant deviations from standard covalent geometry in the structures for both WT and mutant drkN SH3 domains, as previously reported (*48*−*51*). Therefore, a force constant of 0.3 kcal Hz^{-2} has been chosen for the final structure calculations of both proteins.

The resulting ensembles of 10 lowest energy structures of 100 calculated are presented in parts a and b of Figure 2, where the WT structural ensemble is shown in red and the T22G structural ensemble is shown in blue. Structural statistics for the ensemble of 10 lowest energy structures of the WT and T22G drkN SH3 domains are given in Table 1. The quality of the fit of the experimental RDC data (D_{obs}) with that predicted on the basis of the protein structure (D_{calc})

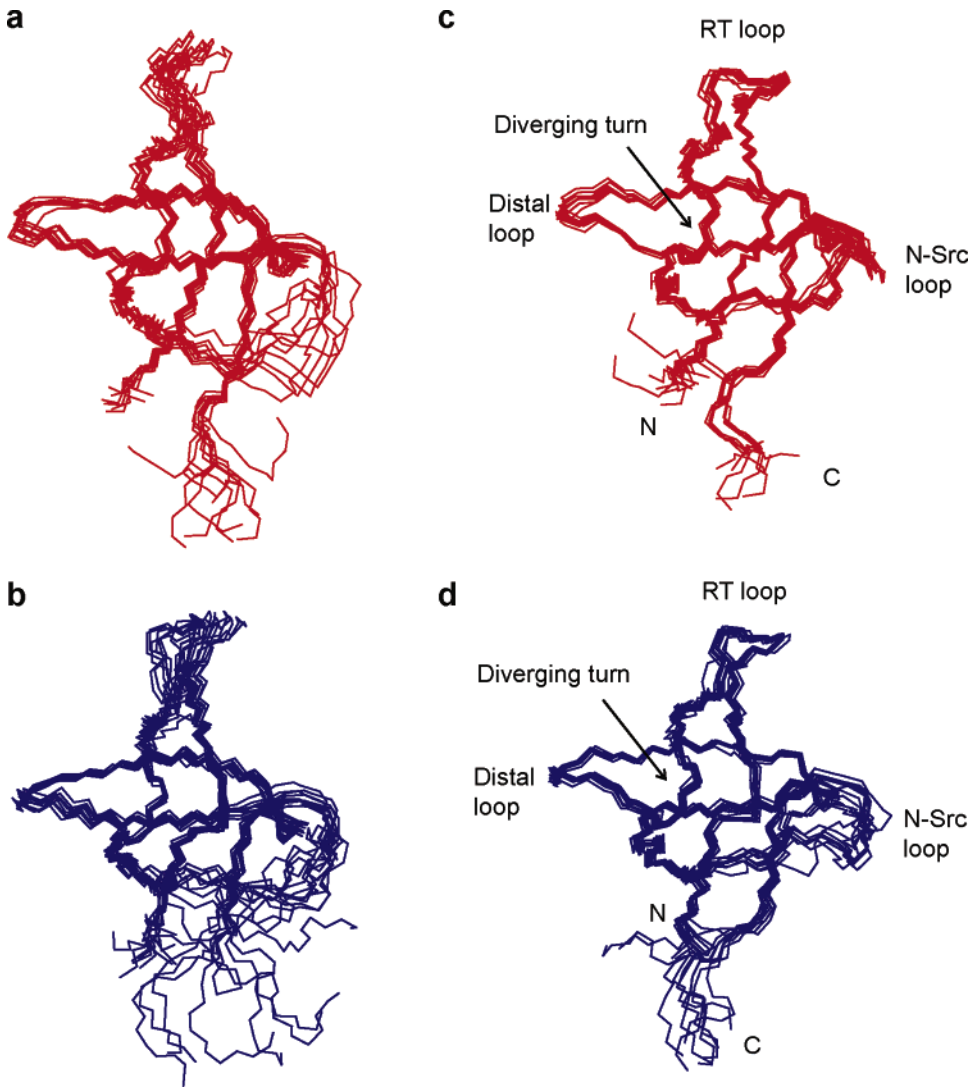

FIGURE 2: Superposition of the 10 lowest energy structures of the WT (a) and T22G (b) drkN SH3 domains calculated on the basis of RDC, C′-CSA, dihedral angle, and hydrogen-bond experimental restraints and WT (c) and T22G (d) structures refined with NOE restraints. The pairwise backbone rmsd values (excluding disordered loops) for the structural ensembles are WT (a) 0.73 ± 0.12 Å, T22G (b) 1.09 ± 0.24 Å, refined WT (c) 0.51 ± 0.10 Å, and refined T22G (d) 0.52 ± 0.09 Å.

was characterized by means of dipolar coupling R factors (*48*), given as

$$R_{\mathrm{dip}} = \sqrt{\frac{5\langle (D_{\mathrm{obs}} - D_{\mathrm{calc}})^2 \rangle}{2(D_{\mathrm{a}})^2(4 + 3R^2)}},$$

where D_{a} is the magnitude of the axial component of the alignment tensor and R is the rhombicity. R_{dip} is a comparison of the rms difference between observed and calculated values with that expected for a random distribution of vector directions, and it varies from 0 to 1, where $R_{\mathrm{dip}} = 0$ reflects the perfect agreement between observed and calculated RDCs and $R_{\mathrm{dip}} = 1$ corresponds to a completely random structure.

Despite the fact that no NOE-derived structural restraints were used in the calculations, both structures are well-defined with the exception of the RT loop and the large flexible N-Src loop and C terminus. The overall pairwise backbone rmsd for the WT structures is 1.36 ± 0.29 Å (using residues 2−56 for fitting) and 0.73 ± 0.12 Å, excluding disordered loops (using residues 2−10, 17−27, and 37−56). For the T22G mutant structure, the corresponding rmsd values are 1.51 ± 0.28 and 1.09 ± 0.24 Å, respectively. The higher rmsd values for the T22G structure reflect the fact that there are fewer RDC restraints available for the T22G structure calculations (216) than for the WT (253). The reduction in data is a result of the doubling of several peaks in the T22G spectra caused by partial formylation of the N-terminal methionine of the T22G domain, determined by mass spectrometry analysis. The diverging turn (residues 20−23) carrying the mutation is also well-defined in both WT and T22G structures with rmsd values of 0.18 ± 0.07 and 0.15 ± 0.10 Å for the WT and T22G, respectively.

The majority of the residues display backbone dihedral angles within the most favored and additionally allowed regions of the Ramachandran plot, 92.2 ± 6.0% and 93.7 ± 5.8% for the WT and T22G, respectively, with 3.4 ± 1.5% (WT) and 2.5 ± 1.6% (T22G) in the generously allowed, and 4.5 ± 2.0% (WT) and 4.0 ± 2.1% (T22G) in the disallowed regions of the plot. The residues in the disallowed regions are either located within the disordered C terminus or N-Src loop of the molecule (residues Glu31, Ser34, and Asn57) or have known positive ϕ angles (Gly/Thr22 and Asp42).

Table 1: Structural Statistics for the 10 Lowest Energy Structures of the WT and T22G Mutant drkN SH3 Domain without and with NOE Restraints

	WT (without NOEs)	T22G (without NOEs)	WT (with NOEs)	T22G (with NOEs)
	Deviations from Experimental Restraints[a]			
rmsd values from hydrogen bond and NOE restraints (Å)	0.018 ± 0.020 (40)	0.009 ± 0.011 (40)	0.025 ± 0.003 (949)	0.017 ± 0.002 (777)
rmsd values from TALOS dihedral angle restraints (deg)	0.86 ± 0.63 (56)	0.92 ± 0.73 (44)	3.45 ± 0.22 (56)	1.40 ± 0.36 (44)
rmsd values from $^3J_{HNH\alpha}$ and $^3J_{H\alpha C'}$ *J*-coupling restraints (Hz)	1.09 ± 0.10 (89)	0.66 ± 0.58 (58)	1.53 ± 0.08 (89)	1.73 ± 0.11 (58)
overall RDC *R* factors[b]				
$^1D_{HN}$	0.21 ± 0.01 (52)	0.17 ± 0.04 (49)	0.21 ± 0.01 (52)	0.16 ± 0.02 (49)
$^1D_{NC'}$	0.13 ± 0.13 (52)	0.08 ± 0.09 (47)	0.12 ± 0.13 (52)	0.07 ± 0.08 (47)
$^1D_{C'C\alpha}$	0.36 ± 0.02 (55)	0.16 ± 0.01 (42)	0.36 ± 0.01 (55)	0.16 ± 0.01 (42)
$^1D_{C\alpha H\alpha}$	0.16 ± 0.02 (46)	0.10 ± 0.02 (41)	0.17 ± 0.01 (46)	0.10 ± 0.00 (41)
$^1D_{C\alpha C\beta}$	0.26 ± 0.02 (48)	0.15 ± 0.04 (37)	0.22 ± 0.01 (48)	0.20 ± 0.01 (37)
rmsd values from C′-CSA (ppb)	23.0 ± 1.3 (54)	22.5 ± 1.5 (45)	20.1 ± 0.5 (54)	20.0 ± 0.7 (45)
	Structural Precision			
rmsd (Å)				
backbone	1.36 ± 0.29	1.51 ± 0.28	0.60 ± 0.15	0.72 ± 0.17
heavy atoms	2.57 ± 0.29	2.85 ± 0.27	1.56 ± 0.18	1.89 ± 0.25
backbone, excluding N-Src and RT loops	0.73 ± 0.12	1.09 ± 0.24	0.51 ± 0.10	0.52 ± 0.09
heavy atoms, excluding N-Src and RT loops	2.09 ± 0.22	2.43 ± 0.21	1.50 ± 0.14	1.77 ± 0.16
	Deviations from Standard Covalent Geometry			
bonds (Å)	0.004 ± 0.0003	0.004 ± 0.0004	0.004 ± 0.0003	0.004 ± 0.0002
angles (deg)	0.94 ± 0.05	0.82 ± 0.06	0.99 ± 0.02	0.80 ± 0.03
impropers (deg)	1.16 ± 0.07	1.03 ± 0.04	1.28 ± 0.04	1.00 ± 0.05
dihedrals (deg)	34.2 ± 1.7	32.7 ± 1.3	32.5 ± 1.4	31.4 ± 1.5
	Percentage of Residues in Different Regions of the Ramachandran Diagram			
most favored regions	68.7 ± 4.4	76.3 ± 3.3	67.0 ± 3.8	83.0 ± 4.0
additional allowed regions	23.5 ± 4.0	17.4 ± 4.8	28.3 ± 3.4	10.5 ± 2.2
generously allowed regions	3.4 ± 1.5	2.5 ± 1.6	2.6 ± 1.3	4.0 ± 1.9
disallowed regions	4.5 ± 2.0	4.0 ± 2.1	2.1 ± 1.6	2.5 ± 2.4

[a] Number of restraints used is given in parentheses. All calculations included RDC, C′-CSA, hydrogen bond, TALOS-derived dihedral angle, and $^3J_{HNH\alpha}$ and $^3J_{H\alpha C'}$ coupling restraints. [b] The dipolar coupling *R* factor is defined as the ratio of the rmsd between observed and calculated values to the expected rmsd if the vectors were randomly oriented. The latter is given by $\{2D_a^2[4 + 3R^2]/5\}^{1/2}$, where D_a is a magnitude of the axial component of the alignment tensor and *R* is the rhombicity (*48*).

The resulting structure of the WT drkN SH3 domain, like other SH3 domains (*12*), is a β-barrel or β-sandwich (Figure 3). The β-strands are arranged into two sheets and an irregular β-hairpin, with strands β1, β7, and part of β4 forming the first sheet and strands β4, β5, and β6 forming the second sheet. Strands β2 and β3 form an irregular twisted β-hairpin with two main-chain hydrogen bonds (9HN−19C and 19HN−9C) connecting two strands. This hairpin is frequently referred to as the RT loop. The long flexible N-Src and the distal loops connect strands β4 and β5 and β5 and β6, respectively. Unpaired β3 and β4 strands are connected by the diverging turn. Strands β6 and β7 are connected by a short three-residue 3_{10} helix typically seen in other SH3 domains. The structure of the T22G drkN SH3 domain has the same structural characteristics.

Comparison of the Structures of the WT and T22G Mutant of the drkN SH3 Domain. A superposition of the WT and T22G drkN SH3 domain structures is shown in Figure 4a. The overall structures of the two proteins are very similar with a backbone rmsd of 1.65 ± 0.40 or 1.27 ± 0.27 Å, excluding N-Src and RT loops. These values, within error, are nearly the same as the rmsd values within the T22G structural ensembles (1.51 ± 0.28 or 1.09 ± 0.24 Å). A more detailed comparison of the structure of the two domains is presented in Figure 5, showing the difference between the WT and T22G dihedral angles ϕ and ψ as a function of a residue number. The plot reveals that the backbone structure of the domains is quite similar, with the most significant differences in the backbone conformation of residues 5 and 6, situated in close proximity of the mutation point, and residues 42 and 43 in the distal hairpin. These differences are evident in the $^3J_{HNH\alpha}$- and $^3J_{H\alpha C'}$-coupling data. Poorly defined regions such as the RT loop, N-Src loop, and C terminus are shaded on the plot.

Despite the similarities of the two structures, the stabilities of the two molecules are very different, with the increase in the stability of T22G triggered by the substitution of Thr22 with a glycine in the diverging turn. The vast majority of SH3 domains have a glycine residue in the position corresponding to position 22 in the drkN SH3 domain, with the diverging turns adopting a type-II β-turn conformation. The type-II β-turn is a tight turn with a hydrogen bond between the carbonyl of residue *i* and the amide proton of residue *i* + 3. This type of turn requires ϕ and ψ angles for residue *i* + 2 in the α_L region (with a positive ϕ), generally not allowed in L-amino acids. Glycine, asparagine, or aspartate are most frequently found in this position of the turn (85% of 405 type-II turns investigated) (*52*). Glycine (76%) has a unique structure lacking a side-chain and, therefore, can adopt ϕ and ψ angles in all four quadrants of the Ramachandran plot. Asparagine (6%) and aspartate (3%), on the other hand, are capable of adopting α_L conformations, owing to formation of a hydrogen bond between their side-chain and main chain, which stabilizes otherwise unfavorable conformations.

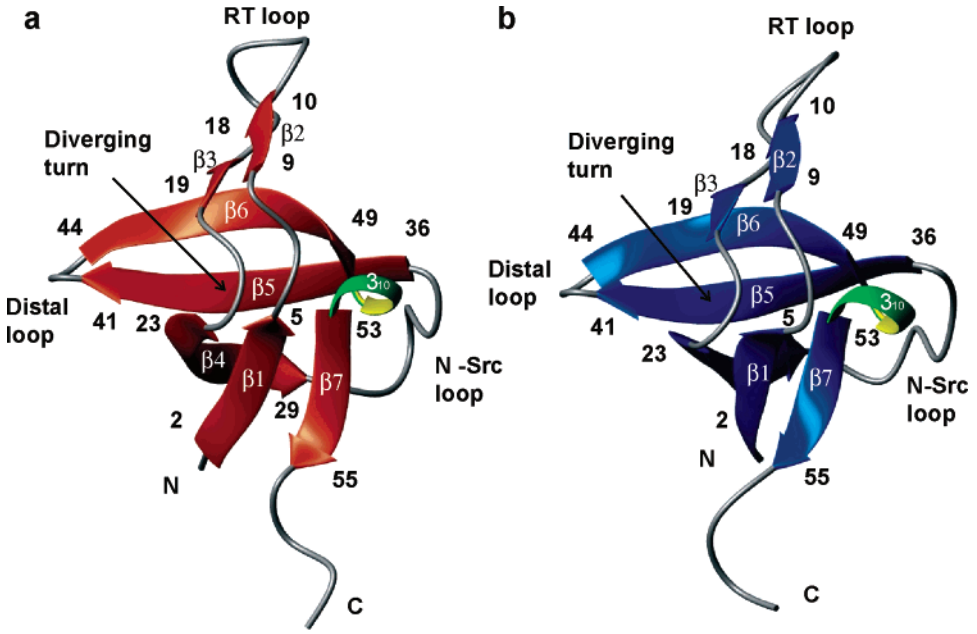

FIGURE 3: Ribbon diagrams of the lowest energy structures of the (a) WT and (b) T22G drkN SH3 domain. The β-strands of the WT and T22G mutant drkN SH3 domains are shown in red and blue, respectively, and the 3_{10} helices are shown in green. All loops and secondary structure elements are labeled.

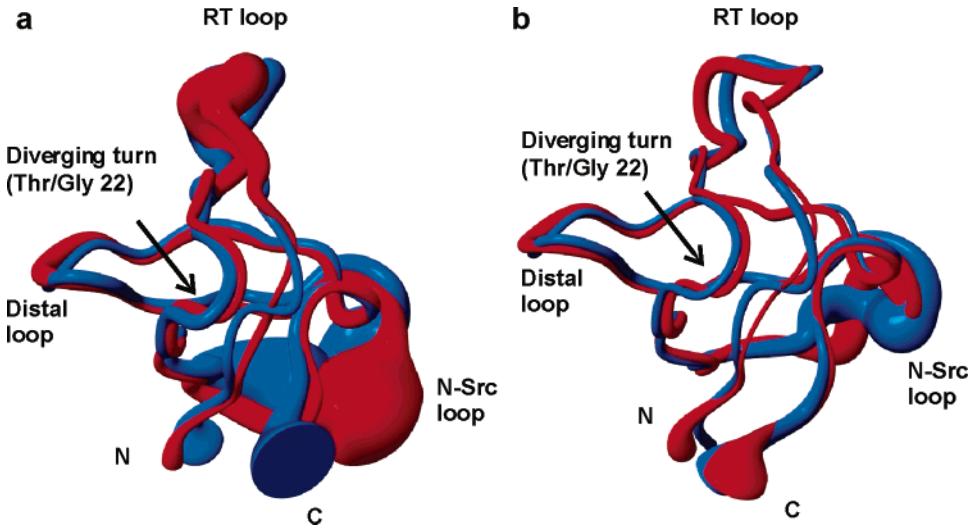

FIGURE 4: Superposition of the WT (red) and the T22G (blue) structural ensembles represented as tubes with radii proportional to the rmsd values within the ensembles. All of the loops and the mutation point are labeled. (a) Structures calculated on the basis of RDC, C′-CSA, dihedral angle, and hydrogen-bond experimental restraints. The rmsd between the two ensembles is 1.65 ± 0.40 Å. (b) Structures refined with NOE-based restraints. The rmsd between the two ensembles is 1.02 ± 0.27 Å.

In contrast, threonine is extremely rare in this position of type-II turns. It has been found in only 0.7% of all type-II turns investigated (*52*). Therefore, we analyzed in detail the conformations of the diverging turn and, specifically, Thr22 and Gly22.

The RDC/C′-CSA and hydrogen-bond restraints alone were not enough to unambiguously confirm positive ϕ values at position 22 of the protein. Thus, we determined ϕ angles directly for both proteins using the 3D HNHA experiment for measuring three-bond $^3J_{HNH\alpha}$-coupling constants (*38*). In general, measured *J*-coupling constants showed a good agreement with structures calculated before introducing *J*-coupling restraints into structure calculations. $^3J_{HNH\alpha}$-coupling constants yield two potential values for ϕ, with the exception of coupling constants from stereoassigned glycine residues. The *J*-coupling value for residue Gly22 unambiguously confirmed a positive ϕ angle for the mutant; however, the *J*-coupling value for Thr22 in the WT was ambiguous, allowing, in principle, both positive and negative values of the ϕ angle. To verify the angle for Thr22 in the WT drkN SH3 structure, we performed an H^{α}-C′ *J*-coupling experiment. This experiment, described earlier by Bax et al. (*40*) and Wang et al. (*41*), is particularly useful because it allows one to distinguish between positive and negative values of ϕ, with high (>3 Hz) $^3J_{H\alpha C'}$ couplings representing positive ϕ angles and low $^3J_{H\alpha C'}$ values corresponding to negative ϕ angles. Measured $^3J_{H\alpha C'}$ values agreed well with ϕ angles of the calculated WT structure and, most importantly, unambiguously confirmed the presence of the positive ϕ angle for Thr22 with one of the highest values of $^3J_{H\alpha C'}$ for this residue (4.7 Hz). Gly43 and Asp42, which have positive ϕ angles in the calculated structure, also showed high values

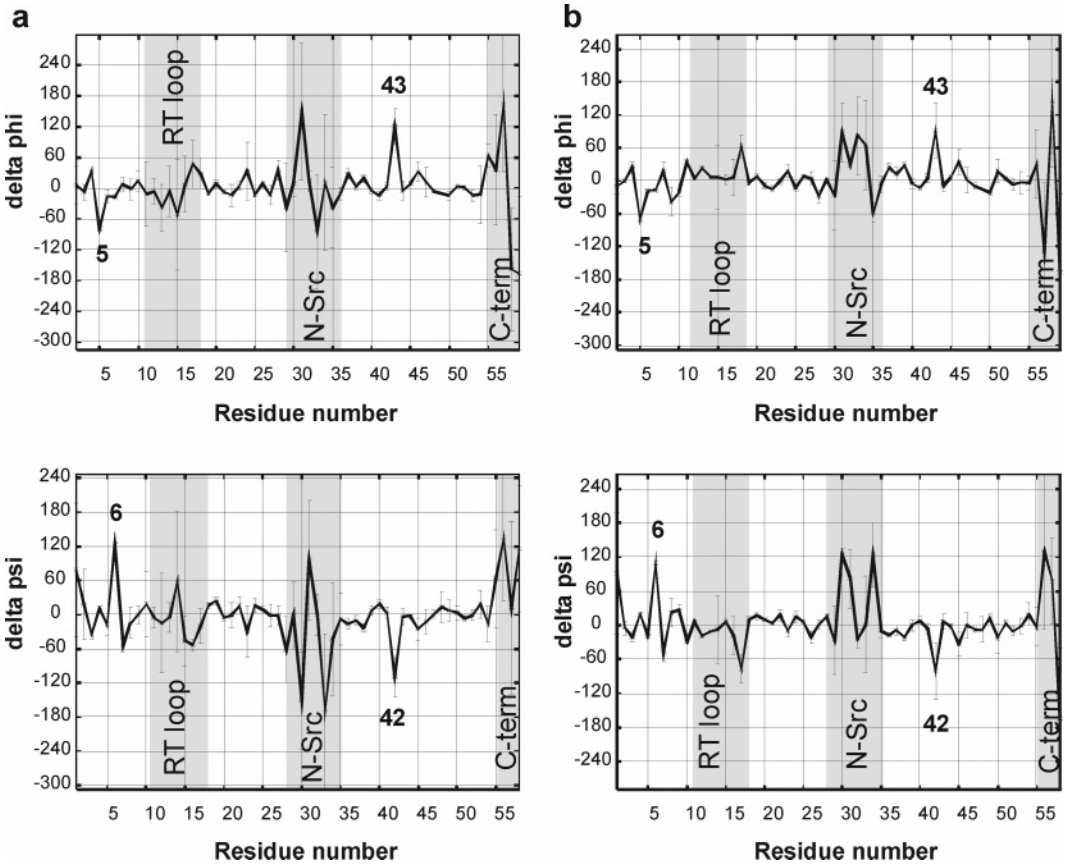

FIGURE 5: Difference between the backbone dihedral angles of the WT and T22G structures calculated without (a) and with (b) NOE restraints, shown as a function of the residue number. Poorly defined regions of the molecule are shaded on the plot. Residues outside the poorly defined regions displaying the largest change in dihedral angles are labeled.

Table 2: Comparison of the Diverging Turn Geometry of the drkN SH3 Domain and Its T22G Mutant

		without NOEs		with NOEs	
residue	angle (turn position)[a]	WT	T22G	WT	T22G
Arg 20	ϕ (i)	-114.5 ± 3.5	-123.0 ± 7.0	-114.0 ± 2.0	-121.6 ± 6.4
	ψ (i)	151.9 ± 4.3	157.5 ± 2.5	166.7 ± 1.8	156.9 ± 1.4
Lys 21	ϕ ($i+1$)	-61.6 ± 3.8	-53.8 ± 3.0	-62.6 ± 1.9	-53.6 ± 3.6
	ψ ($i+1$)	141.1 ± 3.7	142.3 ± 22.2	136.5 ± 2.6	133.2 ± 1.3
Thr/Gly 22	ϕ ($i+2$)	73.6 ± 8.7	87.1 ± 19.9	80.1 ± 2.4	95.4 ± 4.6
	ψ ($i+2$)	-1.7 ± 12.2	17.5 ± 7.5	4.1 ± 5.8	-13.8 ± 8.4
Gln 23	ϕ ($i+3$)	-75.0 ± 8.8	-79.2 ± 1.2	-81.5 ± 5.2	-80.4 ± 1.1
	ψ ($i+3$)	129.2 ± 46.0	159.3 ± 2.0	149.3 ± 9.9	159.3 ± 1.8

[a] ϕ and ψ angles of the diverging turn region are given as average ϕ and ψ angles in the ensembles of 10 lowest energy structures. For each residue, its corresponding position in the type-II β-turn is given in parentheses. Standard angles for the type-II β-turns are $i+1$ (-60, 120), $i+2$ (80, 0).

of $^3J_{H\alpha C'}$ in this experiment (4.4 and 5.5 Hz, respectively), providing a good control for the accuracy of the method. The $^3J_{HNH\alpha}$-coupling restraints were included into final T22G structure calculations, and $^3J_{H\alpha C'}$- and $^3J_{HNH\alpha}$-coupling restraints were included into final WT calculations (Table 1).

It is interesting to note that dihedral angle restraints derived from TALOS, on the basis of chemical-shift values and statistical data from the protein database, predict a negative value for the ϕ angle of Thr22. This value, however, is inconsistent with the bulk of the structural restraints, causing dihedral angle distortions around the Thr22 residue in the WT structure. We suggest that TALOS predictions be used with caution, particularly in turns or loop regions. Because TALOS is based on statistics, it potentially provides inaccurate predictions for residues with unusual dihedral angles. Therefore, we did not include TALOS-derived dihedral angle restraints for the diverging turn or any of the loops in the protein in the final structure calculations.

Calculated structures of the T22G mutant of the drkN SH3 domain with the conserved glycine residue in the $i+2$ position of the diverging turn demonstrate a type-II conformation of the turn, as expected, with ϕ and ψ angles of 87.1 $\pm$ 19.9 and -17.5 ± 7.5, respectively, in position $i+2$ (see Table 2). It is somewhat surprising, however, that T22, in structures of the WT drkN SH3 domain, has ϕ and ψ angles of 73.6 ± 8.7 and -1.7 ± 12.2, respectively, close to the ideal angles (80, 0) for the $i+2$ position in type-II β-turns. Therefore, a comparison of the two structures, WT and T22G, in topology and specific backbone conformation at position 22, shows that they are very similar and that the mutation does not lead to significant structural differences in the folded state.

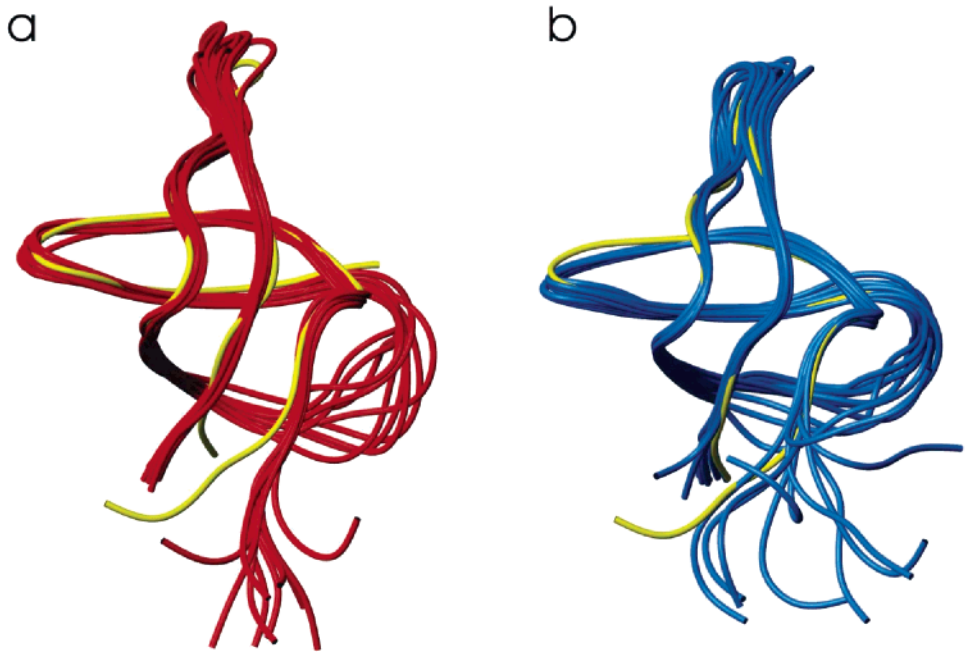

FIGURE 6: Structural comparison of the WT (a, red) and T22G (b, blue) drkN SH3 domains with the Grb-2 N-terminal SH3 domain (*54*), PDB ID 1GRI (yellow). Note that the Grb-2 structure lacks density for residues 28−33 of the N-Src loop.

As stated above, positive ϕ values are not favored in non-glycine amino acid residues because of the steric clash of the β carbon and subsequent atoms of the side-chain, and aspartate and asparagine overcome this potential energy barrier by forming an intraresidue hydrogen bond. Threonine is likely to be even less favored because of additional steric clashes due to its branching at the β position. Therefore, there could be a potential for an intraresidue hydrogen bond involving the NH and O-γ atoms to stabilize the turn in the WT structure. Amide exchange protection data provide evidence for a hydrogen bond involving the backbone NH of Thr22 as well as HN of Gly22. On the basis of the structures, there could be only one acceptor of the hydrogen bond for T22G (carbonyl of Ala5) and potentially two for WT (carbonyl group of Ala5 and/or the hydroxyl group of Thr22). On the basis of the conformational differences between the WT and T22G around residues 5 and 6 (see Figures 4 and 5), it is most likely that the HN of Thr22 is engaged in an intraresidue hydrogen bond.

Structure Refinement with WT NOE Restraints. Because the structures of the WT and T22G calculated on the basis of RDCs and C′-CSA are highly similar and we had made complete backbone and side-chain resonance assignments and recorded and analyzed NOE spectra to determine the structure of the WT protein (A. U. Singer, O. Zhang, and J. D. Forman-Kay, unpublished results), we have used these WT NOE restraints to refine both WT and T22G structures. All 909 NOE restraints were added to the WT structural restraints, and a subset, 737, were added to the T22G restraints. NOE restraints used in T22G structure calculations excluded all of the structural information for the diverging turn region (residues 19−24).

Addition of the NOE restraints significantly improved the structural precision, decreasing pairwise backbone rmsd from 0.73 ± 0.12 to 0.51 ± 0.10 Å for the WT and from 1.09 ± 0.24 to 0.52 ± 0.09 Å for T22G (parts c and d of Figure 2). Residual dipolar coupling *R* factors and deviations from C′-CSA restraints have not changed significantly with the introduction of NOE restraints (Table 1), indicating that NOE-derived restraints are in good agreement with RDC and C′-CSA data. Refined structures of the domains also have diverging turns adopting type-II β-turn conformations in both proteins with dihedral angles in good agreement with those calculated without NOE restraints (Table 2). The superposition of the refined WT and T22G domains is presented in Figure 4b (backbone rmsd = 1.02 ± 0.27 Å).

Comparison to the Grb-2 N-Terminal SH3 Domain. The human homologue of Drk, the growth factor receptor-bound protein 2 (Grb-2), also contains a central SH2 domain surrounded by two (N- and C-terminal) SH3 domains and is involved in activation of *ras* G proteins in response to growth factors (*53*). The N-terminal SH3 domain of Grb-2 is highly similar to the drkN SH3 domain with a sequence identity of 71%. As for most SH3 domains, but unlike the drkN SH3 domain, this protein has a glycine residue at position $i + 2$ of the diverging turn. Figure 6 shows a superposition of the crystal structure of the Grb-2 N-terminal SH3 domain (PDB ID 1GRI, yellow) (*54*) and the drkN SH3 domain structures (WT, red; and T22G, blue). The structures coincide very well, with slight differences in the RT loop conformation. (Note that, because of the lack of density for the N-Src loop, this region is missing from the crystal structure of Grb-2.) The backbone rmsd of the Grb-2 N-terminal SH3 domain structure with respect to the WT ensemble is 1.05 ± 0.05 Å, and its rmsd with respect to the T22G structures is 0.97 ± 0.05 Å.

Effect of the T22G Mutation on the Stability of the drkN SH3 Domain. Recent aromatic and methyl NOE data on the unfolded (U_{exch}) state of the WT drkN SH3 domain (*14*) confirm α-helical structure formation for residues 18−26, as previously predicted by the algorithm AGADIR and suggested by thermodynamic and kinetic analysis of Thr22 mutants of the drkN SH3 domain (*13*). Contacts between the Trp36 indole and Ser18, Thr22, and Lys26, which are in *i*, $i + 4$, and $i + 8$ positions with respect to each other, were detected. It was also shown that residues Phe19 and

Gln23, in positions *i* and *i* + 4 with respect to each other, make contacts with Tyr37. This is highly suggestive of a significantly populated conformation or subset of conformations in the unfolded state ensemble in which the aromatic groups Trp36 and Tyr37 help stabilize non-native α-helical structure between residues 18−26. The comparison of H^{α} chemical-shift data for the 18−26 region of the U_{exch} state with corresponding values in random coil and helix conformations (*41*) estimates the relative population of the helical conformers in this region as 30%. In addition, NOE and other structural data provide evidence for the presence of native-like and non-native structure within the unfolded state ensemble (*7*, *15*). These data support the idea that the instability of the WT drkN SH3 domain is partly due to decreasing the free energy of the unfolded state. In particular, the mutation of Thr22 to glycine disrupts non-native helical interactions in this region of the unfolded state, increasing its free energy and therefore forcing the equilibrium between U_{exch} and F_{exch} states to be shifted toward the folded state. The contribution of conformational entropy in unfolded states to the stability in the case of glycine substitutions has been discussed (*55*). While this would oppose the enthalpic effect of destabilization of the helix in the unfolded ensemble, it is not a dominant effect.

In the present work, the effect of the stabilizing mutation T22G on the folded state structure of the protein has been investigated. Our structural data suggest that a significant contribution to the instability of the WT protein is the presence of a threonine residue in the diverging turn. This residue adopts a sterically and energetically highly unfavorable α_L conformation, which is required in this position to form a type-II β-turn, characteristic of the diverging turn in all SH3 domains. This likely causes an increase in the free energy of the folded state of the WT protein compared to the T22G mutant folded state.

The T22G substitution decreases the free energy of the folded state of the protein by placing the most favorable glycine residue in the diverging turn and therefore avoiding steric problems induced by a threonine residue in this position of the WT protein. In addition, the Gly22 residue prevents formation of the non-native helical structure in the unfolded state by disrupting interactions that involve Thr22 of the WT protein and, therefore, increases the free energy of the unfolded state of the protein. Thus, dramatic stabilization of the T22G drkN SH3 domain is a result of both positive (avoiding steric problems in the folded state and favorable propensity for the consensus β-turn structure) and negative (disruption of the non-native contacts in the unfolded state and prevention of the alternative non-native structure formation) design elements. We conclude that both decreasing the free energy of the unfolded state and increasing the free energy of the folded state of the WT drkN SH3 domain contribute to its low thermodynamic stability and suggest that these issues are features involved in the stability of all proteins.

NMR Data and Coordinates. NMR resonance assignments for the WT and T22G have been deposited in the BioMagResBank (accession codes 5925 and 5923, respectively). The coordinates and NMR-derived restraints have been deposited in the Protein Data Bank (structures calculated without NOEs: accession codes 2AZS and 2AXV, respectively; refined structures: accession codes 2A36 and 2A37, respectively).

ACKNOWLEDGMENT

We thank Dr. Lewis E. Kay for many useful discussions, Dr. Ranjith Muhandiram for help in setting up NMR experiments, Dr. Alan Davidson for the use of his fluorescence spectrophotometer and helpful discussions, and Arash Zarrine-Afsar for assistance with fluorescence experiments.

REFERENCES

1. Simon, M. A., Dodson, G. S., and Rubin, G. M. (1993) An SH3−SH2−SH3 protein is required for P21(Ras1) activation and binds to sevenless and Sos proteins *in vitro*, *Cell 73*, 169−177.
2. Olivier, J. P., Raabe, T., Henkemeyer, M., Dickson, B., Mbamalu, G., Margolis, B., Schlessinger, J., Hafen, E., and Pawson, T. (1993) A *Drosophila* SH2−SH3 adapter protein implicated in coupling the sevenless tyrosine kinase to an activator of *ras* guanine nucleotide exchange, Sos, *Cell 73*, 179−191.
3. Barsagi, D. (1994) The Sos (son of sevenless) protein, *Trends Endocrinol. Metab. 5*, 165−169.
4. McPherson, P. S. (1999) Regulatory role of SH3 domain-mediated protein−protein interactions in synaptic vesicle endocytosis, *Cell. Signalling 11*, 229−238.
5. Zhang, O., and Forman-Kay, J. D. (1995) Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer, *Biochemistry. 34*, 6784−6794.
6. Choy, W. Y., Mulder, F. A. A., Crowhurst, K. A., Muhandiram, D. R., Millett, I. S., Doniach, S., Forman-Kay, J. D., and 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.
7. Crowhurst, K. A., Tollinger, M., and 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.
8. Choy, W. Y., and Forman-Kay, J. D. (2001) Calculation of ensembles of structures representing the unfolded state of an SH3 domain, *J. Mol. Biol. 308*, 1011−1032.
9. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D., and Kay, L. E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain, *J. Am. Chem. Soc. 123*, 11341−11352.
10. Zhang, O. W., and Forman-Kay, J. D. (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions, *Biochemistry 36*, 3959−3970.
11. Farrow, N. A., Zhang, O. W., Forman-Kay, J. D., and Kay, L. E. (1995) Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer, *Biochemistry 34*, 868−878.
12. Larson, S. M., and Davidson, A. R. (2000) The identification of conserved interactions within the SH3 domain by alignment of sequences and structures, *Protein Sci. 9*, 2170−2180.
13. Mok, Y. K., Elisseeva, E. L., Davidson, A. R., and Forman-Kay, J. D. (2001) Dramatic stabilization of an SH3 domain by a single substitution: Roles of the folded and unfolded states, *J. Mol. Biol. 307*, 913−928.
14. Crowhurst, K. A., and Forman-Kay, J. D. (2003) Aromatic and methyl NOEs highlight hydrophobic clustering in the unfolded state of an SH3 domain, *Biochemistry 42*, 8687−8695.
15. Crowhurst, K. A. (2003) Corrigendum to the paper by Mok et al. (1999) NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions, *J. Mol. Biol. 329*, 185−187.
16. Munoz, V., Serrano, L. (1994) Elucidating the folding problem of helical peptides using empirical parameters, *Nat. Struct. Biol. 1*, 399−409.
17. Lacroix, E., Viguera, A. R., Serrano, L. (1998) Elucidating the folding problem of α-helices: Local motifs, long-range electrostatics, ionic-strength dependence, and prediction of NMR parameters, *J. Mol. Biol. 284*, 173−191.
18. Matthews, B. W. (1996) Structural and genetic analysis of the folding and function of T4 lysozyme, *FASEB J. 10*, 35−41.
19. Gassner, N. C., Baase, W. A., Mooers, B. H. M., Busam, R. D., Weaver, L. H., Lindstrom, J. D., Quillin, M. L., and Matthews, B. W. (2003) Multiple methionine substitutions are tolerated in

T4 lysozyme and have coupled effects on folding and stability, *Biophys. Chem. 100*, 325−340.
20. Shortle, D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability, *FASEB J. 10*, 27−34.
21. Wrabl, J. O., and Shortle, D. (1996) Perturbations of the denatured state ensemble: Modeling their effects on protein stability and folding kinetics, *Protein Sci. 5*, 2343−2352.
22. Wrabl, J., and Shortle, D. (1999) A model of the changes in denatured state structure underlying *m* value effects in staphylococcal nuclease, *Nat. Struct. Biol. 6*, 876−883.
23. Pace, C. N., Alston, R. W., and Shaw, K. L. (2000) Charge−charge interactions influence the denatured state ensemble and contribute to protein stability, *Protein Sci. 9*, 1395−1398.
24. Arai, M., Kataoka, M., Kuwajima, K., Matthews, C. R., and Iwakura, M. (2003) Effects of the difference in the unfolded-state ensemble on the folding of *Escherichia coli* dihydrofolate reductase, *J. Mol. Biol. 329*, 779−791.
25. Choy, W. Y., Tollinger, M., Mueller, G. A., and Kay, L. E. (2001) Direct structure refinement of high molecular weight proteins against residual dipolar couplings and carbonyl chemical shift changes upon alignment: An application to maltose binding protein, *J. Biomol. NMR 21*, 31−40.
26. Hansen, M. R., Mueller, L., and Pardi, A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions, *Nat. Struct. Biol. 5*, 1065−1074.
27. Jackson, S. E., and Fersht, A. R. (1991) Folding of chymotrypsin inhibitor-2. 2. Influence of proline isomerization on the folding kinetics and thermodynamic characterization of the transition-state of folding, *Biochemistry 30*, 10436−10443.
28. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe−A multidimensional spectral processing system based on Unix Pipes, *J. Biomol. NMR 6*, 277−293.
29. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) A common sense approach to peak picking in two-dimensional, three-dimensional, and four-dimensional spectra using automatic computer analysis of contour diagrams, *J. Magn. Reson. 95*, 214−220.
30. Johnson, B. A., and Blevins, R. A. (1994) NMR view−A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
31. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity, *J. Magn. Reson. Ser. B 103*, 203-216.
32. Kay, L. E. (1995) Pulsed field gradient multi-dimensional NMR methods for the study of protein structure and dynamics in solution, *Prog. Biophys. Mol. Biol. 63*, 277−299.
33. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
34. Hwang, T. L., van Zijl, P. C. M., and Mori, S. (1998) Accurate quantitation of water−amide proton exchange rates using the phase-modulated clean chemical exchange (CLEANEX-PM) approach with a fast-HSQC (FHSQC) detection scheme, *J. Biomol. NMR 11*, 221−226.
35. Yang, D. W., Venters, R. A., Mueller, G. A., Choy, W. Y., and Kay, L. E. (1999) TROSY-based HNCO pulse sequences for the measurement of (HN)-H-1-N-15, N-15-(CO)-C-13, (HN)-H-1-(CO)-C-13, (CO)-C-13-C-13(α), and (HN)-H-1-C-13(α) dipolar couplings in N-15, C-13, H-2-labeled proteins, *J. Biomol. NMR 14*, 333−343.
36. Evenas, J., Mittermaier, A., Yang, D. W., and Kay, L. E. (2001) Measurement of C-13(α)-C-13(*β*) dipolar couplings in N-15, C-13, H-2-labeled proteins: Application to domain orientation in maltose binding protein, *J. Am. Chem. Soc. 123*, 2858−2864.
37. Mittermaier, A., and Kay, L. E. (2001) *χ*(1) torsion angle dynamics in proteins from dipolar couplings, *J. Am. Chem. Soc. 123*, 6892−6903.
38. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) Measurement of H-N-H-α *J*-couplings in calcium-free calmodulin using new 2d and 3d water-flip-back methods, *J. Biomol. NMR 4*, 871−878.
39. Vuister, G. W., and Bax, A. (1993) Quantitative *J* correlation−A new approach for measuring homonuclear 3-bond J(H(N)H(α) coupling constants in N-15-enriched proteins, *J. Am. Chem. Soc. 115*, 7772−7777.
40. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994) Measurement of homonuclear and heteronuclear *J*-couplings from quantitative *J* correlation, in *Nuclear Magnetic Resonance*, part C, pp 79−105, Academic Press, Inc., San Diego, CA.
41. Wang, A. C., and Bax, A. (1996) Determination of the backbone dihedral angles *ϕ* in human ubiquitin from reparametrized empirical Karplus equations, *J. Am. Chem. Soc. 118*, 2483−2494.
42. Wang, Y. J., and Jardetzky, O. (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data, *Protein Sci. 11*, 852−861.
43. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding, *Biochemistry 40*, 3561−3571.
44. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
45. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information, *J. Magn. Reson. 133*, 216−221.
46. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
47. Prestegard, J. H., Al-Hashimi, H. M., and Tolman, J. R. (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings, *Q. Rev. Biophys. 33*, 371−424.
48. Clore, G. M., and Garrett, D. S. (1999) *R* factor, free *R*, and complete cross-validation for dipolar coupling refinement of NMR structures, *J. Am. Chem. Soc. 121*, 9008−9012.
49. Mueller, G. A., Choy, W. Y., Yang, D. W., Forman-Kay, J. D., Venters, R. A., and Kay, L. E. (2000) Global folds of proteins with low densities of NOEs using residual dipolar couplings: Application to the 370-residue maltodextrin-binding protein, *J. Mol. Biol. 300*, 197−212.
50. Schwalbe, H., Grimshaw, S. B., Spencer, A., Buck, M., Boyd, J., Dobson, C. M., Redfield, C., and Smith, L. J. (2001) A refined solution structure of hen lysozyme determined using residual dipolar coupling data, *Protein Sci. 10*, 677−688.
51. Demene, H., Ducat, T., Barthe, P., Delsuc, M. A., and Roumestand, C. (2002) Structure refinement of flexible proteins using dipolar couplings: Application to the protein p8MTCP1, *J. Biomol. NMR 22*, 47−56.
52. Hutchinson, E. G., and Thornton, J. M. (1994) A revised set of potentials for *β*-turn formation in proteins, *Protein Sci. 3*, 2207−2216.
53. Chardin, P., Cussac, D., Maignan, S., and Ducruix, A. (1995) The Grb2 adaptor, *FEBS Lett. 369*, 47−51.
54. Maignan, S., Guilloteau, J. P., Fromage, N., Arnoux, B., Becquart, J., and Ducruix, A. (1995) Crystal structure of the mammalian Grb2 adaptor, *Science 268*, 291−293.
55. Daquino, J. A., Gomez, J., Hilser, V. J., Lee, K. H., Amzel, L. M., and Freire, E. (1996) The magnitude of the backbone conformational entropy change in protein folding, *Proteins: Struct., Funct., Genet. 25*, 143−156.

BI0512795