Structure, Dynamics, and Stability Variation in Bacterial Albumin Binding Modules: Implications for Species Specificity[†,‡]

Yanan He, David A. Rozak, Nese Sari, Yihong Chen, Philip Bryan, and John Orban*

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received March 1, 2006; Revised Manuscript Received June 16, 2006

ABSTRACT: Protein G-related albumin-binding (GA) modules are frequently expressed on the surfaces of bacterial cells. The limited amino acid sequence variation among GA modules results in structural and functional differences with possible implications for bacterial pathogenesis and host specificity. In particular, the streptococcal G148-GA3 and *F. magna* ALB8-GA albumin-binding domains exhibit a degree of structural and dynamic diversity that may account for their varied affinities for different species of albumin. To explore the impact of GA module polymorphisms on albumin binding and specificity, we recently used offset recombinant PCR to shuffle seven artificially constructed representatives of the GA sequence space and scan the phage-displayed recombinant domains for mutations that supported binding to the phylogenetically distinct human and guinea pig serum albumins (HSA and GPSA) (Rozak et al. (2006) *Biochemistry 45*, 3263−3271). Surprisingly, phage selection revealed an overwhelming preference for a single recombinant domain (PSD-1, phage-selected *d*omain-1) regardless of whether the phages were enriched for their abilities to bind one or both of these albumins. We describe here the NMR-derived structure, dynamics, and stability of unbound PSD-1. Our results demonstrate that increased flexibility is not a requirement for broadened specificity, as had been suggested earlier (Johansson et al. (2002) *J. Mol. Biol.* 316, 1083−1099), because PSD-1 binds the phylogenetically diverse HSA and GPSA even more tightly than G148-GA3 but is less flexible. The structural basis for albumin-binding specificity is analyzed in light of these new results.

In 1994, de Château and Björck described a broad class of three-helix albumin-binding domains, representing the first example of a shuffled bacterial module (*1*). Since then, 16 Protein G-related albumin-binding (GA[1]) modules with a high degree of sequence identity have been found in multiple bacterial species (*2*). The domains are thought to support pathogenesis by scavenging albumin-bound nutrients (*3*) and possibly camouflaging the bacteria. The limited amino acid sequence variation among GA modules results in both structural and functional differences with possible implications for bacterial pathogenesis and host specificity (*4*). Indeed, structural and competitive binding studies have shown that streptococcal G148-GA3 and *Finegoldia magna* ALB8-GA albumin-binding domains exhibit a degree of structural diversity that may account for their varied affinities for different species of albumin (*4*, *5*). Specifically, G148-GA3 has been shown to have a more flexible backbone than ALB8-GA (*5*) and atypical folding thermodynamics (*6*) that might enhance the domain's affinity for diverse albumins.

To explore the impact of GA module polymorphisms on albumin-binding specificity, we recently used offset recombinant polymerase chain reaction (OR-PCR) (*7*) to shuffle seven artificially constructed representatives of the GA sequence space and scan the phage-displayed recombinant domains for mutations that supported binding to the phylogenetically distinct human and guinea pig serum albumins (HSA and GPSA) (*8*). Surprisingly, phage selection revealed an overwhelming preference for a single recombinant domain (PSD-1, phage selected domain-1), regardless of whether the phages were enriched for their abilities to bind one or both of the phylogenetically diverse albumins.

Thermodynamic analysis of PSD-1 showed that the phage-selected domain was more stable than G148-GA3 at high temperatures and exhibited 4- and 10-fold increases in HSA- and GPSA-binding constants, respectively, compared to those of the wild-type streptococcal domain (*8*). A review of PSD-1 polymorphisms with respect to the structurally defined G148-GA3 and ALB8-GA albumin-binding domains suggests that a single mutation in the protein's core may largely be responsible for simultaneous gains in PSD-1 stability and binding by stabilizing the domain in a conformation that is more conducive to albumin interaction.

This article describes a series of NMR experiments aimed at resolving the PSD-1 structure and assessing whether the domain's enhanced affinities for human and guinea pig serum albumins are attributable to domain-stabilizing mutations. Our results support the hypothesis that a core mutation

[†] This work was supported by NIH Grants GM62154 and 1S10RR15744 and the W. M. Keck Foundation.

[‡] Structure coordinates for PSD-1 are deposited in the PDB (accession code 2fs1). NMR assignments for PSD-1 are deposited in BioMagResBank under BMRB accession code 6945.

* Corresponding author. Tel: 240-314-6221. Fax: 240-314-6255. E-mail: orban@umbi.umd.edu.

[1] Abbreviations: GA, protein G-related albumin-binding; HSA, human serum albumin; GPSA, guinea pig serum albumin; PSD-1, phage selected domain-1; OR−PCR, offset recombinant polymerase chain reaction.

10.1021/bi060409m CCC: $33.50
Published on Web 07/27/2006

substantially stabilizes the PSD-1 backbone in a conformation consistent with the albumin-bound ALB8-GA domain. We have not evaluated PSD-1 interactions with the full range of albumin species. However, the NMR data presented here show that increased flexibility is not a requirement for broadened specificity because PSD-1 binds the phylogenetically diverse HSA and GPSA even more tightly than G148-GA3 but has a more rigid backbone. Rather, PSD-1's structural and thermodynamic features appear to support the proposition made by Lejon et al. that broad albumin specificity is largely attributable to a common tyrosine near the *N*-terminus of the second alpha helix of the GA module (*9*). Previously, it was difficult to assess the relative contributions of backbone dynamics and the tyrosine binding interface to broad species specificity because both factors were present in G148-GA3 and absent from ALB8-GA.

MATERIALS AND METHODS

Sample Preparation. PSD-1 was cloned into a pG58 vector as described previously (*8*). In this approach, the protein is expressed as a fusion with subtilisin propeptide and, subsequently, purified and cleaved on an immobilized subtilisin column (*10*). *E. coli* BL21(DE3) competent cells were transformed with the pG58/PSD-1 plasmid and grown at 37 °C in 1 L of ^{15}N-labeled or ^{15}N/^{13}C-labeled minimal media containing 100 mg/mL of ampicilin. At an OD_{600} of approximately 0.8, the cells were induced with 1 mM IPTG for 4 h. The cells were harvested by centrifugation, and the cell pellet was resuspended in 100 mM potassium phosphate at pH 7.0, 30 mg/mL of DNaseI, and 0.1 mM PMSF for sonication on ice. After centrifugation for 1 h at 100 000*g*, the cleared lysate was loaded onto a S189 column (5 mL bed volume) and purified using the method of Ruan et al. (2004). The eluent containing purified PSD-1 was dialyzed against 2 mM ammonium bicarbonate buffer (pH 7.0) at 4 °C overnight and then lyophilized. NMR samples of PSD-1 were prepared at a concentration of 0.2−0.4 mM in 50 mM potassium phosphate buffer at pH 7.0 with 0.1 mM EDTA, and 10% D_2O was added.

NMR Spectroscopy. Spectra for structural studies were acquired on a Bruker DRX-600 equipped with a *z*-axis gradient triple resonance cryoprobe. Spectra for dynamics studies were acquired on a Bruker DMX-600 with a sample concentration of about 1 mM. All experiments on PSD-1 were collected at 25 °C. Data were processed on a Linux workstation using nmrPipe (*11*) and analyzed using Sparky (Goddard, T. D., and Kneller, D. G., UCSF). Standard triple resonance experiments were used to make NMR assignments (*12*−*16*). Backbone resonance assignments were obtained from HNCACB, CBCA(CO)NH, and HNCO spectra, whereas side chain assignments were made from HBHA(CO)NH, (H)C(CO)NH-TOCSY, H(CCO)NH-TOCSY, 3D ^{15}N-TOCSY, 3D ^{15}N-NOESY, and 3D ^{13}C-NOESY spectra. Distance restraints were obtained from a 3D ^{15}N-edited NOESY spectrum in H_2O (*17*) and 3D ^{13}C-edited NOESY (*18*) spectra in H_2O and D_2O with mixing times of 100 and 150 ms.

^{15}N-T_1 and -T_2 measurements were performed using 2D ^{15}N-HSQC experiments (*19*) with water flip-back (*20*) and Watergate (*21*) modifications for solvent suppression. For the T_1 and T_2 experiments, a recycle delay of 1.5 s was used between acquisitions to ensure sufficient recovery of ^{1}H magnetization. For T_1 experiments, 12 variable delays were used (11.9, 51.9, 101.9, 151.9, 201.9, 301.9, 401.9, 501.9, 601.9 1001.9, 1201.9, 1501.9 ms), and for T_2 experiments, 12 variable delays were used (8, 16, 24, 32, 40, 48, 64, 80, 96, 120, 160, 200 ms). Duplicate spectra were recorded for the T_1 measurement at delays of 11.9, 51.9, and 201.9 ms and the T_2 experiment at delays of 8, 16, and 40 ms, and these were used to estimate noise levels and fitting errors. T_1 and T_2 values were calculated by a nonlinear least-squares fitting of the cross-peak heights to a single-exponential decay with error estimates for *T* using Sparky.

Steady-state {^{1}H}-^{15}N NOEs were measured using ^{15}N-HSQC spectra with a gradient-selected, sensitivity-enhanced pulse sequence (*22*). A relaxation delay of 2 s prior to a 3 s proton presaturation period was employed for the experiments with NOE, and presaturation was applied at 4 MHz off-resonance with the same recycling delay for the experiments without NOE. The NOEs were calculated as the ratios of peak heights in the spectrum recorded with proton saturation to those in the spectrum recorded without saturation. The standard deviation of the NOE value was estimated on the basis of measured background noise levels (*22*). An analysis of ^{15}N relaxation data was carried out using the Modelfree program (*23*). The global correlation time (τ_m) was obtained from a trimmed mean value of T_1/T_2 utilizing standard procedures.

Structure Calculations. Structures were calculated using CNS 1.1 (*24*) and standard simulated annealing and torsion angle dynamics protocols. Initial NOE assignments and an initial set of structural models were generated by the automated CANDID iterative refinements module of Cyana (*25*). After this initial assignment step using Cyana, additional NOE assignments were added and erroneous ones corrected through a manual examination of NMR spectra using Sparky prior to recalculation of the structure by CNS1.1. Distance restraints based on peak intensities were categorized as strong (1.8−3.0 Å), medium (1.8−4.0 Å), weak (1.8−5.5 Å), and very weak (2.8−6.0 Å). Backbone dihedral angle restraints were obtained on the basis of chemical shift values using TALOS (*26*). Two restraints were used to define each hydrogen bond, 1.5−2.5 Å for r_{HN-O} and 2.3−3.2 Å for r_{N-O}, and these restraints were used only in the later stages of refinement. Nonbonded contacts were represented by a quartic van der Waals repulsion term. Final values for force constants were 1000 kcal mol^{-1} Å^{-2} for bond lengths, 500 kcal mol^{-1} rad^{-2} for angles and improper torsions, 40 kcal mol^{-1} Å^{-2} for experimental distance restraints, 200 kcal mol^{-1} rad^{-2} for dihedral angle restraints, and 4.0 kcal mol^{-1} Å^{-4} for the van der Waals repulsion term. The twenty best structures for each protein were chosen on the basis of low total energy, no NOE distance violations greater than 0.30 Å, no dihedral restraint violations greater than 5°, and standard indicators of structure quality as shown in Table 1. The structures were analyzed using PROCHECK-NMR (*27*) and QUANTA (Molecular Simulations, Inc.) and displayed using MOLMOL (*28*).

Hydrogen−*Deuterium (H*−*D) Exchange Measurements.* Exchange experiments were carried out at 25 °C at pH 7.0. Lyophilized ^{15}N-PSD-1 was dissolved in D_2O, and 30 ^{15}N-HSQC spectra were recorded in automated mode, each with 16 scans and 64 t_1 increments. Most of the amide exchange was complete at this point. Individual amide proton peak

Table 1: Statistics for the PSD-1 Ensemble of 20 Structures

A. Experimental Restraints	
NOE restraints	
all NOE	1063
intraresidue	423
sequential ($\|i-j\|=1$)	273
medium-range ($1<\|i-j\|\leq 5$)	263
long-range ($\|i-j\|>5$)	104
hydrogen bond restraints	58
dihedral angle restraints	64
total restraints	1185
B. rmsd's to the Mean Structure (Å)	
over all residues[a]	
backbone atoms	0.31 ± 0.08
heavy atoms	1.10 ± 0.12
secondary structures[b]	
backbone atoms	0.27 ± 0.08
heavy atoms	1.05 ± 0.13
C. Measures of Structure Quality	
Ramachandran distribution	
most favored (%)	84.4 ± 2.4
additionally allowed (%)	13.8 ± 2.3
generously allowed (%)	1.4 ± 1.6
disallowed (%)	0.4 ± 0.82
bad contacts/100 residues	3.8 ± 0.89
overall dihedral *G* factor	0.05 ± 0.02

[a] Residues 6−52. [b] The secondary elements used were as follows: residues 6−23, 27−35, and 39−51.

intensity data were fitted to a first-order exponential decay of the form $A_t = A_0\exp(-k_{ex}t)$, where A_t is the peak volume at time t and k_{ex} is the first-order rate constant.

RESULTS

Structure and Backbone Dynamics of PSD-1. The oligomeric state of the 56 amino acid PSD-1 protein was determined to be monomeric on the basis of size exclusion chromatography (Supporting Information, Figure 1) and concentration-dependent thermal-melting characteristics. Specifically, a calorimetric study of 123 mM PSD-1 buffered in 50 mM potassium phosphate at pH 7.0 yielded an identical melting point because circular dichroic measurements on 3 mM PSD-1 melted under identical conditions (*8*). Near complete chemical shift assignments were made at 25 °C for all backbone amide resonances (Supporting Information, Figure 2), more than 90% of all side chain protons, and most side chain ^{13}C and ^{15}N resonances. Chemical shift and NOE data analysis indicated that the structure contains three amphipathic α-helices, α1 (residues 6−23), α2 (residues 27−35), and α3 (residues 39−51), arranged in a compact bundle. The three α-helices are well ordered with the two loop regions, α1−α2 and α2−α3, being less well defined (Figure 1A). Residues 1−5 have few interresidue NOEs and form a disordered *N*-terminal tail, whereas residues 53−56 at the *C*-terminus are also poorly defined. Figure 2 shows an overlay of the 20 final structures with an average backbone RMSD for residues 6−52 of 0.31 ± 0.08 Å. Complete structure statistics are shown in Table 1.

The first turn of the α1-helix includes the three residues A6, N7, and S8, which are artifacts from the cloning procedure but do not contribute to the hydrophobic core of the protein. The core residues include K13, A16, and L20 from α1, I33 from α2, and V42, K46, and I49 from α3. Both K13 and K46 have extensive contacts from their hydrophobic

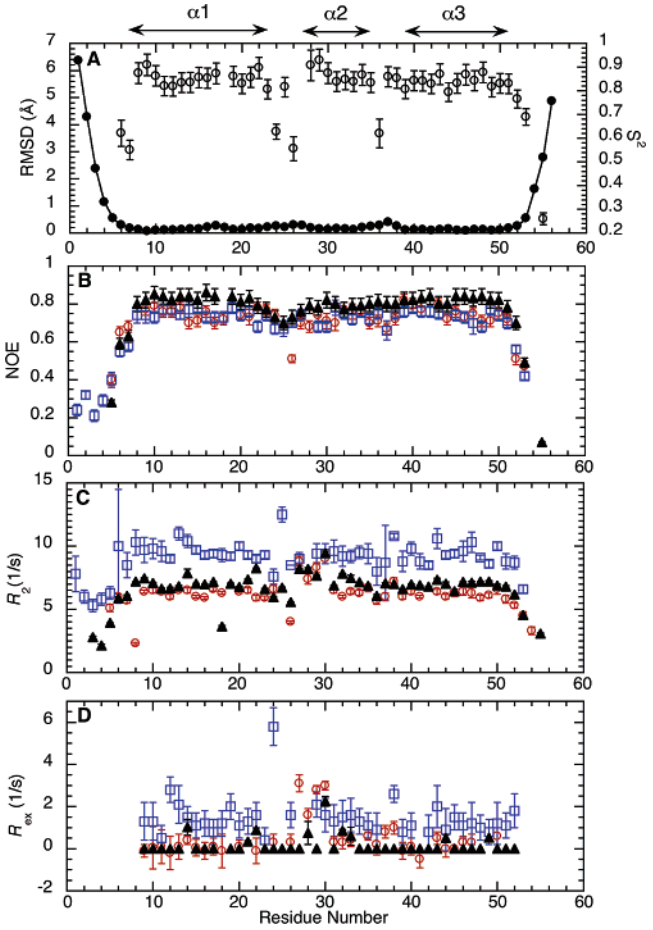

FIGURE 1: (A) Average rmsd per residue for the backbone N, C_α, and C′ atoms of PSD-1 (●). The secondary structure elements are shown at the top of the plot. Order parameters for PSD-1 are shown as open circles (○). (B) 600 MHz steady state {^{1}H}-^{15}N nuclear Overhauser effect (NOE) values per residue for PSD-1 (triangles), ALB8-GA (red circles), and G148-GA3 (blue squares). (C) 600 MHz spin−spin (R_2) relaxation rates (s^{-1}) as a function of residue number for PSD-1 (triangles), ALB8-GA (red circles), and G148-GA3 (blue squares). (D) 600 MHz R_{ex} data for PSD-1 (triangles), ALB8-GA (red circles), and G148-GA3 (blue squares). The data for ALB8-GA and G148-GA3 in (B), (C), and (D) are from ref 5 (*5*). The data for all three proteins were recorded under similar conditions of protein concentration, temperature and pH. PSD-1 numbering is used, and the values for T24 in the α1−α2 loop of ALB8-GA are omitted so that the data can be aligned correctly.

methylene groups to the core, whereas their ϵ-ammonium moieties are solvent accessible. Furthermore, L53 near the *C*-terminus of α3 packs against the hydrophobic side chains of I25 (α1-α2 loop) and L50 (α3) forming a stabilizing cap on the three-helix bundle.

^{15}N-relaxation data were obtained for 53 out of 56 residues in PSD-1. No data were obtained for M1 or E2, and residue 54 is a proline. Standard Model Free analysis of the R_1, R_2, and NOE data gave the generalized order parameters (S^2) shown in Figure 1A. The dynamics results are consistent with the structural data, indicating that PSD-1 has flexible *N*- and *C*-terminal tails, whereas residues 6−51 form a well-ordered domain. Within this domain, the S^2 values are high for residues in all three helices, indicating a relatively rigid backbone. A number of residues in the loop region between α1 and α2 and, to a lesser extent, between α2 and α3 have lower order parameters suggesting that these regions are more flexible on the ns to ps time scale. Estimates of R_{ex} values

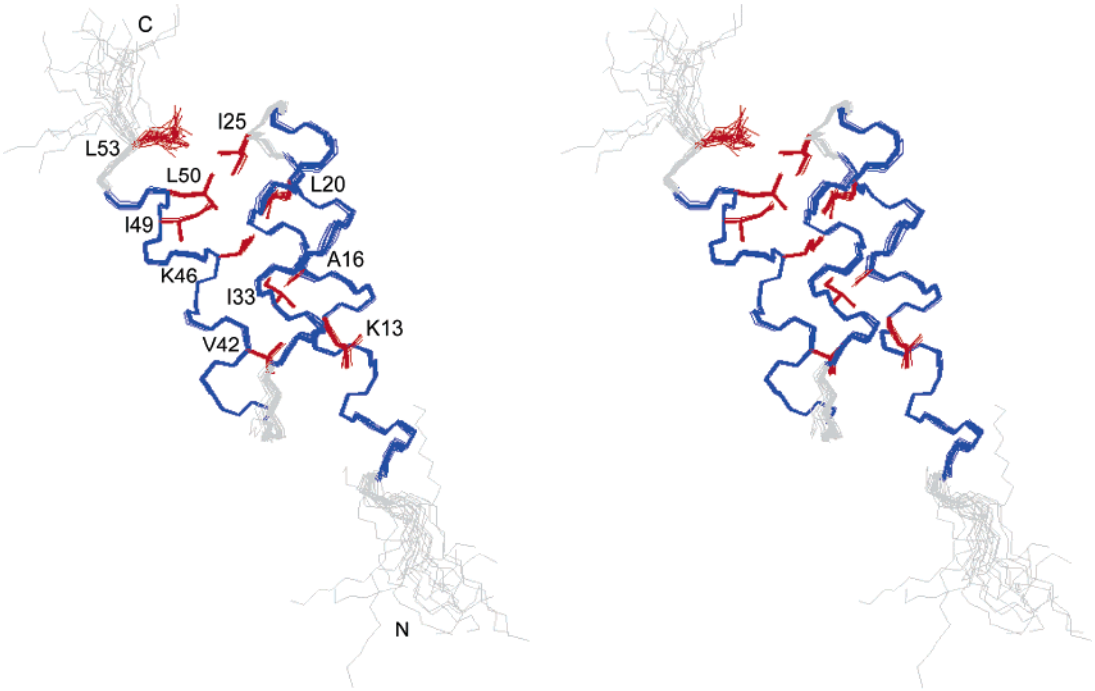

FIGURE 2: Stereoview of the PSD-1 ensemble containing the 20 lowest energy structures. The α-helical backbone regions are highlighted in blue, whereas buried and partly buried side chains are shown in red and labeled. Unstructured and loop regions are in gray.

for PSD-1 were also obtained, and these are displayed in Figure 1D. These results show that only a limited number of residues have relatively small R_{ex} terms. ^{15}N-relaxation data for ALB8-GA and G148-GA3 were also available (*5*). These are compared with the PSD-1 results in Figure 1B, C, and D and discussed below.

Derivation of the Thermodynamic State Functions for PSD-1 Unfolding. Slow-exchanging amides were identified in the α1-, α2-, and α3-helices, consistent with their buried positions in the NOE-derived structure of PSD-1. The H−D exchange values for slow-exchanging backbone amides were used to compute PSD-1 $\Delta G_{unfolding}$ at room temperature. The H−D exchange for each backbone amide can be described by the reaction path

$$F(H) \underset{k_f}{\overset{k_u}{\rightleftharpoons}} U(H) \xrightarrow[D_2O]{k_c} U(D) \underset{k_u}{\overset{k_f}{\rightleftharpoons}} F(D) \qquad (1)$$

where k_u, k_f, and k_c are the respective unfolding, folding, and intrinsic exchange rates, F is the folded state, and U is the unfolded state (*29*). Because the exchange experiment was conducted at temperatures well below T_m ($k_f \gg k_u$) with an intrinsic exchange rate significantly less than the rate of protein folding ($k_f \gg k_c$), the measured exchange rate k_{ex} can be given by

$$k_{ex} = \frac{k_u k_c}{k_f} = K_{op} k_c \qquad (2)$$

The free energy required for transient opening can then be expressed as

$$\Delta G_{op} = -RT \ln K_{op} = -RT \ln(k_{ex}/k_c) \qquad (3)$$

where K_{op} is the equilibrium constant for transient opening, and ΔG_{op} is the free energy difference between folded and unfolded states. The values for k_c were calculated using the SPHERE algorithm, which employs sequence-dependent correction factors and amide exchange rates measured for model peptides (*30*, *31*). The histograms in Figure 3 were produced by applying this analysis to H−D exchange data obtained here for PSD-1 and that reported elsewhere for ALB8-GA and G148-GA3 (*5*). Because the slow-exchanging

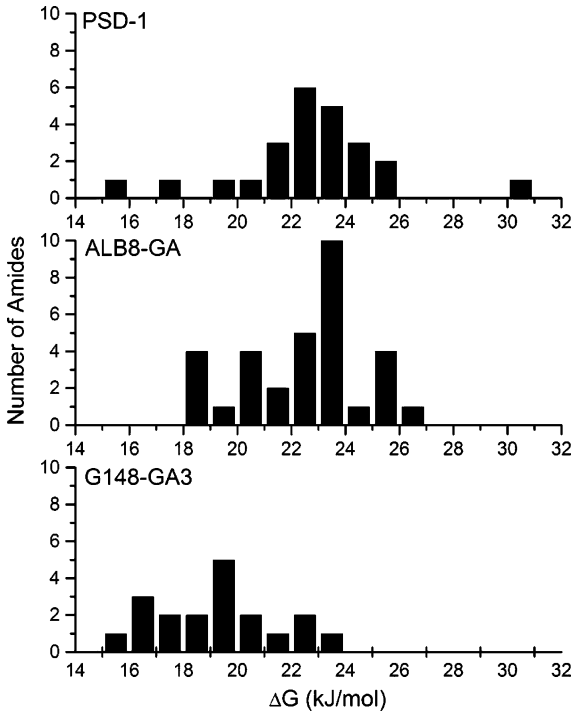

FIGURE 3: Histograms comparing the free energies of opening (ΔG_{op}) for slow H−D exchanging amides in PSD-1, ALB8-GA, and G148-GA3 suggest that PSD-1 and ALB8-GA similarly exhibit a $\Delta G_{unfolding}$ value of around 23 kJ/mol near 25 °C compared to an estimated 19 kJ/mol for G148-GA3. ΔG_{op} was computed from H−D exchange data reported here for PSD-1 and elsewhere for ALB8-GA and G148-GA3 (*5*).

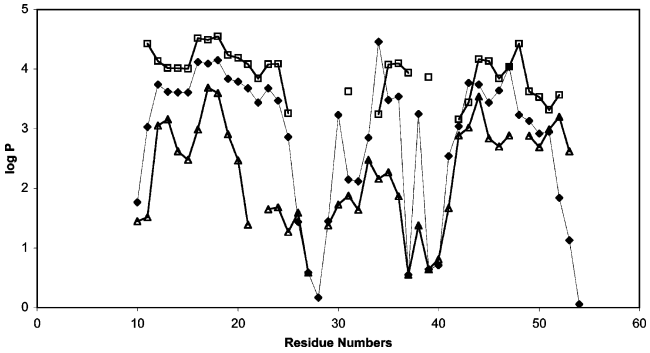

FIGURE 4: Plot of log(protection factor) per residue for PSD-1 (□), ALB8-GA (♦), and G148-GA3 (△). The PSD-1 results are from data obtained in the current study at 25 °C at pH 7.0, whereas the protection factors for ALB8-GA and G148-GA3 are from ref 5 (*5*), where exchange was done at 27 °C at pH 7.2. The protection factor P is defined as k_c/k_{ex}, where k_c is the calculated intrinsic exchange rate, and k_{ex} is the measured amide exchange rate. The values for PSD-1 only include slow-exchanging amide protons.

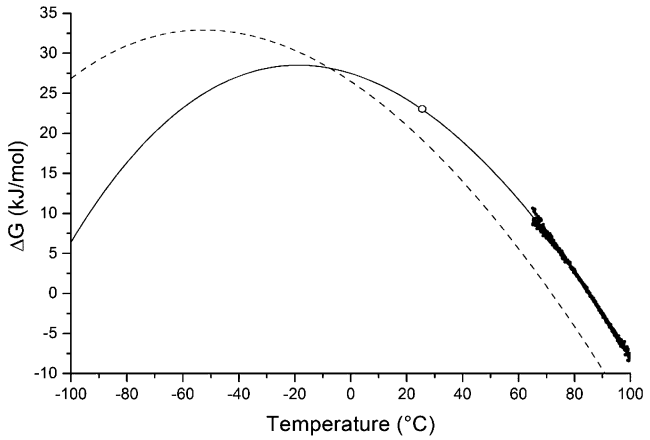

FIGURE 5: Comparison of PSD-1 (—) and G148-GA3 (---) free energy plots under comparable pH 7.0 buffer conditions. The graph shows a tight fit for eq 4 to NMR (○) and circular dichroic derived (dots) free energies at different temperatures. Circular dichroic-derived PSD-1 data and the G148-GA3 free energy plot were obtained from earlier publications (*6*, *8*).

amides are expected to swap hydrogen isotopes only when the molecules are fully unfolded, ΔG_{op} values for these amides are centered around the $\Delta G_{unfolding}$ for the entire domain. The slow-exchanging amides of PSD-1 and ALB8-GA are similarly clustered around a value of 23 kJ/mol. G148-GA3 slow exchangers, however, have an average ΔG_{op} of 19 kJ/mol, which is consistent with an earlier calorimetrically derived estimate of the protein's free energy under similar conditions (*6*). The protection factors per residue obtained here for PSD-1 are shown in Figure 4 and compared with previously obtained results for G148-GA3 and ALB8-GA (*5*).

The NMR-derived value for PSD-1 $\Delta G_{unfolding}$ at room temperature can be combined with high-temperature circular dichroism-derived values for $\Delta G_{unfolding}$ reported earlier (*8*) to compute the remaining thermodynamic state functions for this reaction. Fitting the equation

$$\Delta G = \Delta H_0 - T\Delta S_0 + \Delta C_p[T - T_0 - T\ln(T/T_0)] \quad (4)$$

to these two data sets in Figure 5 yields the values for ΔH, ΔS, and ΔC_P reported in Table 2 alongside those obtained previously for G148-GA3 (*6*). Consistent with the NMR structure and backbone dynamics observed for both domains, the thermodynamic data in Table 2 suggests that PSD-1 achieves greater stability at room temperature by doing a better job of decreasing its hydrophobic surface area upon folding. This conclusion is based on the lower change in entropy and broader shift in heat capacity associated with PSD-1 unfolding.

Table 2: Thermodynamic State Functions for PSD-1 and G148-GA3 Unfolding at 25 °C

protein	ΔG (kJ/mol)	ΔH (kJ/mol)	ΔS (J/deg mol)	ΔCp (kJ/deg mol)
PSD-1[a]	23 ± 1	95 ± 3	240 ± 9	1.51 ± 0.05
G148-GA3[b]	19	118	332	1.1 ± 0.1

[a] PSD-1 state functions were obtained by fitting eq 4 to the NMR and circular dichroic data displayed in Figure 5. [b] G148-GA3 state functions were derived from calorimetric data presented in an earlier article (*6*).

DISCUSSION

Comparative Analysis of GA Modules. Three other GA module structures have been described in the literature, an NMR structure of ALB8-GA (pdb code 1prb (*2*)), a 2.7 Å X-ray structure of ALB8-GA in complex with HSA (pdb code 1tf0 (*9*)), and an NMR structure of G148-GA3 (pdb code 1gjt (*4*)). A sequence alignment of PSD-1 with ALB8-GA and G148-GA3 is shown in Figure 6A. On the basis of sequence homology alone, PSD-1 is more similar to G148-GA3 (68.5%, 37 identities over 54 amino acids) than ALB8-GA (56.6%, 30 identities over 53 amino acids). However, although all four structures are three-helix bundles, PSD-1 most closely resembles HSA-bound ALB8-GA in its overall 3D structure, particularly with regard to helix lengths, helix orientations, and packing interactions. The main difference between free and bound ALB8-GA is that the α1-helix is one turn longer at the *N*-terminus in the HSA-bound form just as in PSD-1. A backbone structural alignment of PSD-1, HSA-bound ALB8-GA, and G148-GA3 shows that PSD-1 superimposes better with G148-GA3 in α1, about evenly well with the other two structures in α2, and better with ALB8-GA in α3 (Figures 6B and 7). The relatively large displacement in the α1−α2 loop of ALB8-GA is presumably due to the fact that this loop contains an extra residue (T24) compared with the number of residues in PSD-1 and G148-GA3. In contrast, the α2−α3 loop is in good agreement in all three structures. The consistently larger rmsd of displacement of the residues in the α3-helix of G148-GA3 relative to that of the other structures is also evident when the ensembles of NMR structures for G148-GA3, ALB8-GA, and PSD-1 are compared (Supporting Information, Figure 3).

The structural differences between PSD-1 and G148-GA3 in particular must be due to a subset of the 17 nonidentical residues in these two sequences. An examination of the distinct residues in each structure indicates that 13 are solvent-accessible surface residues (PSD-1 residue−G148-GA3 residue: M1-D12, Q11-E22, E14-V25, A15-L26, K18-R29, K21-D32, Q22-K33, G26-S37, K31-N42, E43-K54, S44-A55, N47-D58, and K51-A62), three are partly buried (I17-N28, I25-V36, and I30-K41), and one is mostly buried (K46-I57). The solvent-accessible residues do not affect the

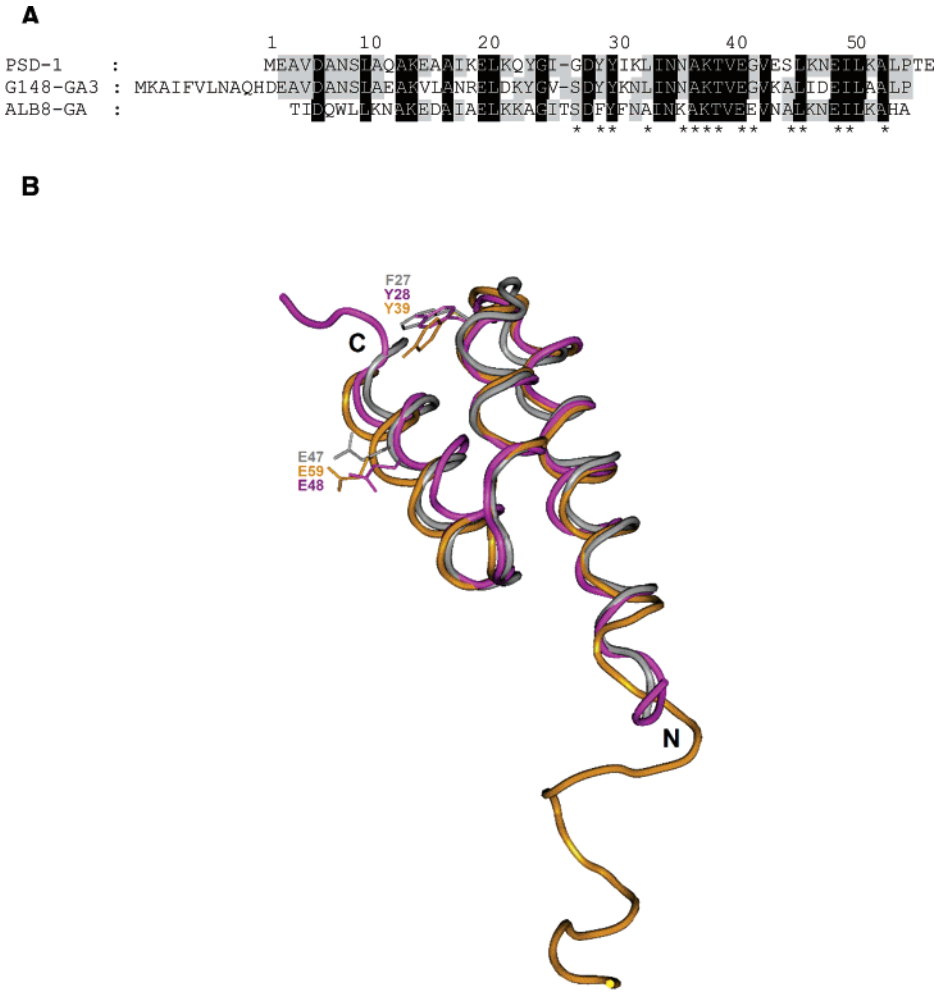

FIGURE 6: (A) Sequence alignment for PSD-1, streptococcal G148-GA3, and *Finegoldia magna* ALB8-GA albumin-binding domains. The numbers indicate PSD-1 residue positions. Identical residues are highlighted in gray or black, depending on whether they are conserved across two or three of the sequences, respectively. Amino acids implicated in the albumin-binding interface by the ALB8-GA/HSA crystal structure or a VAST (*32*) alignment with this structure are indicated with asterisks. (B) VAST alignment (*32*) of unbound PSD-1 (purple trace), unbound G148-GA3 (orange trace), and HSA-bound ALB8-GA (gray trace) shows G148-GA3 deviates significantly from PSD-1 and ALB8-GA with respect to the backbone alignment of a glutamic acid (rendered side chain) involved in the albumin-binding interface. PSD-1's closer alignment with HSA-bound ALB8-GA at this point likely contributes to the domain's enhanced albumin-binding ability compared to that of G148-GA3. The tyrosine (rendered side chain) at the start of PSD-1 α2 likely explains why the domain does not exhibit ALB8-GA's 1000-fold preference for HSA over GPSA, despite other structural and thermodynamic similarities. Consistent with a theory proposed by Lejon et al. (*9*), the tyrosine found in G148-GA3 and PSD-1 may give the domains a significant advantage in binding several nonprimate albumins, including GPSA.

packing of the buried residues, and therefore, some or all of the four remaining amino acids must be responsible for the structural variations, likely through tighter hydrophobic interactions with physically adjacent residues. A mutation from a β-branched amino acid (I57) in G148-GA3 to a lysine in PSD-1 may be part of the reason why PSD-1 has better packing of the α3-helix against the other two helices. It is interesting to note that this position is lysine in 9 of the 16 naturally occurring GA variants discovered to date. Other interactions may also contribute to the improved packing of hydrophobic residues in PSD-1 relative to that in G148-GA3. Residues I17 and I30 in PSD-1 form stabilizing hydrophobic contacts between α1 and α2. In contrast, the corresponding residues in G148-GA3, N28, and K41 do not have as many stabilizing close contacts. Furthermore, the hydrophobic capping interaction between L53 and I25 in PSD-1 is not present between the analogous residues, L64 and V36, in G148-GA3. However, both the free and bound forms of

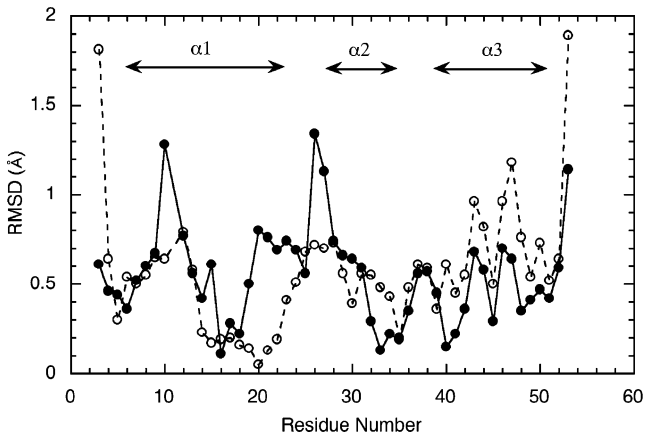

FIGURE 7: Plot of per residue backbone rmsd between PSD-1 and HSA-bound ALB8-GA (●), and PSD-1 and G148-GA3 (○). The average PSD-1 and G148-GA3 NMR structures were used in this analysis.

ALB8-GA have this capping interaction, where H52 at the end of the α3-helix packs against I23 in the α1−α2 loop. The structural results are, therefore, consistent with the thermodynamic data described above, which indicate that PSD-1 buries its hydrophobic residues more efficiently than G148-GA3.

Although these structural changes are small, they have a significant effect on the backbone dynamics, stability, and binding affinities of the individual domains. PSD-1 and ALB8-GA (free and HSA-bound) are structurally more compact than G148-GA3. Consistent with this, Johansson et al. (*5*) found that most residues in G148-GA3 have a notable R_{ex} contribution to R_2 and undergo extensive conformational exchange on the μs−ms time scale (Figure 1D). In contrast, ALB8-GA has only a few residues in the α1−α2 loop and the *N*-terminus of the α2-helix that exhibit such motions. Our results indicate that PSD-1 has ^{15}N-relaxation rates similar to the more rigid ALB8-GA module (Figures 1C, D) under comparable conditions. In agreement with these data are the hydrogen−deuterium exchange results, which show that PSD-1 and ALB8-GA have analogous protection factors (Figure 4) and ΔG_{op} values (Figure 3) under similar conditions, whereas G148-GA3 exhibits faster exchange characteristics and has lower stability. An earlier work suggested that G148-GA3 acquires its broader binding specificity for albumins from multiple species, partly through enhanced backbone flexibility compared with that of ALB8-GA (*5*). Although this may be true for G148-GA3, the work presented here demonstrates that increased flexibility is not a requirement for broadened specificity because PSD-1 binds the phylogenetically diverse HSA and GPSA even more tightly than G148-GA3 but is less flexible.

Analysis of the PSD-1 Albumin-Binding Interface: Implications for Binding Specificity. What then is the structural basis for these changes in affinity and specificity? We previously used isothermal titration calorimetry to show that, compared to G148-GA3, PSD-1 exhibited approximately 4- and 10-fold increases in its ability to bind HSA and GPSA, respectively (*8*). Improvements to PSD-1 HSA binding appear to be driven largely by more favorable changes in entropy upon binding when compared to the same G148-GA3 reaction. However, PSD-1's enhanced GPSA binding benefits from more balanced improvements to ΔH and ΔS when compared with that of G148-GA3. Despite sharing a number of ALB8-GA's structural and thermodynamic characteristics, PSD-1 resembles G148-GA3 in its affinity for phylogenetically diverse human and guinea pig albumins and its failure to exhibit ALB8-GA's 1000-fold preference for HSA (*4*). By comparing the structural data presented here for PSD-1 with that obtained elsewhere for G148-GA3 (*4*) and ALB8-GA (*9*), one can offer possible explanations for PSD-1's improved HSA and GPSA binding and broader range of affinities.

On the basis of their analysis of the HSA-bound ALB8-GA crystal structure, Lejon et al. identified an α3 glutamate (E47) that forms an important interaction with HSA K212 at the edge of a fatty-acid binding pocket (*9*). We note that this glutamate is similarly found in G148-GA3 (E59) and PSD-1 (E48). When all three domains are aligned in Figure 6B, the backbone trace of G148-GA3's α3-helix deviates most from the HSA-bound ALB8-GA structure at ALB8-GA E47, suggesting that the G148-GA3 E47 equivalent (E59) and surrounding helix may become distorted slightly upon albumin binding. However, the alignment in Figure 6B shows that this is not the case for the unbound PSD-1 structure, which closely approximates the HSA-bound ALB8-GA in its placement of the same glutamate (E48). Stabilization of this glutamate residue in the proper position for HSA and likely GPSA binding may partially account for PSD-1's enhanced ability to bind both albumins relative to that of G148-GA3.

Lejon et al. go on to argue that another residue, ALB8-GA F27, may contribute to the high species specificity exhibited by the domain because the phenylalanine appears to fill a hydrophobic pocket in HSA that is disrupted by a polar residue in many nonprimate albumins (*9*). The existence of a polar threonine in position 329 of GPSA (Accession No. AY294645) may account for the 1000-fold lower affinity ALB8-GA shows for this species of albumin compared to that for HSA (*4*). Consistent with this analysis, the substitution of a tyrosine for phenylalanine in PSD-1 (Y28) and G148-GA3 (Y39) may support the interaction with the GPSA T329 and account for the high affinities both domains show for GPSA when compared with that for ALB8-GA.

In addition to Y28, the sequence/structure alignment implicates six other PSD-1 polymorphisms in the albumin-binding interface (Figure 6A). Of note are G26 and G41, which appear to reduce the size of the polar interface compared with that of G148-GA3 or ALB8-GA and S44, which may mitigate this loss by supporting an additional polar interaction with the ligand. Moreover, a bulky leucine at position 32 appears to fill a hydrophobic pocket in the binding interface better than a similarly placed alanine in ALB8-GA. Further experiments will be needed to assess the relative impacts of these polymorphisms.

CONCLUSION

Researchers have alternatively proposed that G148-GA3's flexible backbone structure (*5*) or its Y39 polar aromatic residue (*9*) might account for the domain's broader albumin specificity compared to that of ALB8-GA. However, these features were either jointly present in G148-GA3 or absent from ALB8-GA, making it difficult to assess the relative contributions of the factors on albumin specificity. A validated mechanism for albumin specificity is likely to have broad medical implications, given the expression of GA homologues in multiple bacterial pathogens and the possible impact for domain polymorphisms on host specificity. We recently used offset recombinant PCR and phage display to search the GA sequence space for molecules that exhibit broad albumin affinity via their abilities to efficiently bind both HSA and GPSA (*8*). NMR analysis of the dominant phage-selected mutant presented here shows that an α2 tyrosine at Y28 and not backbone dynamics offers the most likely explanation for the domain's comparable human and guinea pig albumin affinities. These findings lend weight to Lejon et al.'s proposed mechanism for albumin specificity (*9*) and support predictions on the species specificities of other albumin-binding domains.

SUPPORTING INFORMATION AVAILABLE

Gel filtration data, an HSQC spectrum, and structural comparisons of PSD-1 with G148-GA3 and ALB8-GA. This

material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. de Château, M., and Björck, L. (1994) Protein PAB, a mosaic albumin-binding bacterial protein representing the first contemporary example of module shuffling, *J. Biol. Chem. 269*, 12147−12151.
2. Johansson, M. U., de Château, M., Wikström, M., Forsén, S., Drakenberg, T., and Björck, L. (1997) Solution structure of the albumin-binding GA module: a versatile bacterial protein domain, *J. Mol. Biol. 266*, 859−865.
3. de Château, M., Holst, E., and Björck, L. (1996) Protein PAB, an albumin-binding bacterial surface protein promoting growth and virulence, *J. Biol. Chem. 271*, 26609−26615.
4. Johansson, M. U., Frick, I. M., Nilsson, H., Kraulis, P. J., Hober, S., Jonasson, P., Linhult, M., Nygren, P. A., Uhlén, M., Björck, L., Drakenberg, T., Forsén, S., and Wikström, M. (2002) Structure, specificity, and mode of interaction for bacterial albumin-binding modules, *J. Biol. Chem. 277*, 8114−8120.
5. Johansson, M. U., Nilsson, H., Evenäs, J., Forsén, S., Drakenberg, T., Björck, L., and Wikström, M. (2002) Differences in backbone dynamics of two homologous bacterial albumin-binding modules: implications for binding specificity and bacterial adaptation, *J. Mol. Biol. 316*, 1083−1099.
6. Rozak, D. A., Orban, J., and Bryan, P. N. (2005) G148-GA3: a streptococcal virulence module with atypical thermodynamics of folding optimally binds human serum albumin at physiological temperatures, *Biochim. Biophys. Acta 1753*, 226−233.
7. Rozak, D. A., and Bryan, P. N. (2005) Offset recombinant PCR: a simple but effective method for shuffling compact heterologous domains, *Nucleic Acids Res. 33*, e82.
8. Rozak, D. A., Alexander, P. A., He, Y., Chen, Y., Orban, J., and Bryan, P. N. (2006) Using offset recombinant PCR to identify functional determinants in a common family of bacterial albumin binding domains, *Biochemistry 45*, 3263−3271.
9. Lejon, S., Frick, I. M., Björck, L., Wikström, M., and Svensson, S. (2004) Crystal structure and biological implications of a bacterial albumin binding module in complex with human serum albumin, *J. Biol. Chem. 279*, 42924−42928.
10. Ruan, B., Fisher, K. E., Alexander, P. A., Doroshko, V., and Bryan, P. N. (2004) Engineering subtilisin into a fluoride-triggered processing protease useful for one-step protein purification, *Biochemistry 43*, 14539−14546.
11. 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.
12. Ikura, M., Kay, L. E., and Bax, A. (1990) A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin, *Biochemistry 29*, 4659−4667.
13. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *J. Am. Chem. Soc. 114*, 6291−6293.
14. Wittekind, M., and Mueller, L. (1993) HNCACB, a high sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha and beta carbon resonances in proteins, *J. Magn. Reson., Ser. B 101*, 201−205.
15. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of backbone amide and aliphatic side-chain resonances in C-13/N-15-enriched proteins by isotropic mixing of C-13 magnetization, *J. Magn. Reson., Ser. B 101*, 114−119.
16. Montelione, G. T., Lyons, B. A., Emerson, S. D., and Tashiro, M. (1992) An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically enriched proteins, *J. Am. Chem. Soc. 114*, 10974−10975.
17. Fesik, S. W., and Zuiderweg, E. R. P. (1988) Heteronuclear three-dimensional NMR spectroscopy: a strategy for the simplification of homonuclear two-dimensional NMR spectra, *J. Magn. Reson. 78*, 588−593.
18. Ikura, M., Kay, L. E., Tschudin, R., and Bax, A. (1990) Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C-labeled protein, *J. Magn. Reson. 86*, 204−209.
19. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease, *Biochemistry 28*, 8972−8979.
20. Grzesiek, S., and Bax, A. (1993) The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements, *J. Am. Chem. Soc. 115*, 12593−12594.
21. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR 2*, 661−665.
22. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation, *Biochemistry 33*, 5984−6003.
23. Mandel, A. M., Akke, M., and Palmer, A. G. (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme, *J. Mol. Biol. 246*, 144−163.
24. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse, K. R., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D 54*, 905−921.
25. Herrmann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *J. Mol. Biol. 319*, 209−227.
26. 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.
27. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
28. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
29. Hvidt, A., and Nielsen, S. O. (1966) Hydrogen exchange in proteins, *Adv. Protein Chem. 21*, 287−386.
30. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange, *Proteins 17*, 75−86.
31. Connelly, G. P., Bai, Y., Jeng, M. F., and Englander, S. W. (1993) Isotope effects in peptide group hydrogen exchange, *Proteins 17*, 87−92.
32. Gibrat, J. F., Madej, T., and Bryant, S. H. (1996) Surprising similarities in structure comparison, *Curr. Opin. Struct. Biol. 6*, 377−385.

BI060409M