NMR Identification and Characterization of the Flexible Regions in the 160 kDa Molten Globule-Like Aggregate of Barstar at Low pH[†]

Juhi Juneja,[‡] Neel S. Bhavesh,[§] Jayant B. Udgaonkar,[*,‡] and Ramakrishna V. Hosur[*,§]

National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bangalore 560 065, India, and Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India

ABSTRACT: Barstar is known to form a molten globule-like A form below pH 4. This form exists as a soluble aggregate of 16 monomeric subunits, and appears to remain homogeneous in solution for at least two weeks. Here, structural characterization by NMR of the flexible regions in the A form of barstar has been carried out at pH 2.7 and 25 °C. Significantly, the A form appears to be a symmetrical aggregate. Using the recently described fast assignment strategy from HNN and HN(C)N spectra, along with the standard triple resonance and three-dimensional NMR experiments, the flexible segment of the aggregate has been identified to belong largely to the N-terminal end of the polypeptide chain; sequential connectivities were obtained for the first 20 residues (except two) from these experiments. This segment is free in each of the monomeric subunits, and does not form a part of the aggregated core of the A form. The secondary chemical shifts of these residues suggest propensity toward an extended structure. Their $^3J_{HN,H\alpha}$ coupling constants have values corresponding to those in a random coil structure. However, a few medium-range NOEs, some of them involving side chain atoms, are observed between some residues in this segment. The lowered temperature coefficients of the H^N chemical shifts compared to random coil values indicate possibilities of some hydrogen bonding in this region. Analysis of the ^{15}N relaxation parameters and reduced spectral density functions, in particular the negative values of heteronuclear NOEs, indicates large-amplitude high-frequency motions in the N-terminal segments; the first three residues show more negative NOEs than the others. The ^{15}N transverse relaxation rates and the $J(0)$ spectral density values for residues Ser12 and Ser69 are significantly larger than for the rest, indicating some microsecond to millisecond time scale conformational exchange contributions to the relaxation of these residues. Taken all together, the data suggest that the A form of barstar is an aggregate with a rigid core, but with the N-terminal 20 residues of each of the monomeric subunits, in a highly dynamic random coil conformation which shows transient local ordering of structure. The N-terminal segment, anchored to the aggregated core, exhibits free-flight motion.

Protein aggregation is well-known as a side-reaction that accompanies the folding reactions of many proteins, but the mechanisms by which proteins aggregate are poorly understood. The formation of aggregated protein instead of the viable native state may be merely the result of kinetic competition between intra- and intermolecular interactions during protein folding (*1, 2*). Such competition may arise because partially folded intermediates that populate protein folding pathways are very often molten globule-like in

possessing exposed hydrophobic patches, which facilitate aggregation (*3, 4*). Aggregation-prone intermediates may accumulate either early (*1*) or late (*5, 6*) on folding pathways, aggregation may occur only transiently during folding (*6−8*), and aggregated forms may transform to native protein without first undergoing unfolding (*6*).

In vivo aggregation is usually evident in the form of inclusion bodies (*9*) or amyloid fibrils (*9, 10*). Thus, not unexpectedly, partially folded intermediates of the tailspike protein of phage P22 that aggregate in vitro, are also responsible for inclusion body formation (*11, 12*). Conditions that lead to the accumulation of partially folded intermediates are likely to lead to aggregation, and point mutations in a number of proteins have been shown to affect the amount of inclusion body formation (*13, 14*). Some of the mutations that lead to the formation of amyloid fibrils in transthyretin (TTR) (*15*), lysozyme (*16*), and immunoglobulin variable light chain domains (*17, 18*) have been shown to destabilize the native states of these proteins relative to the partially folded intermediates. Similarly, temperature and solution conditions such as pH, ionic strength, and denaturant concentration that lead to the formation of different partially

* To whom correspondence should be addressed. Prof. Ramakrishna V. Hosur, Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai- 400 005, India. E-mail: hosur@tifr.res.in; tel: +91-22-2152971, fax: +91-22-2152110. Prof. Jayant B. Udgaonkar, National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bangalore-560 065, India. E-mail: jayant@ncbs.res.in; tel: +91-80-3636422, fax: +91-80-3636662.

[‡] National Centre for Biological Sciences, Tata Institute of Fundamental Research.

[§] Department of Chemical Sciences, Tata Institute of Fundamental Research.

(un)folded intermediates with varying stabilities affect the degree of aggregation (*19, 20*) and may also control aggregate morphology (*21−23*).

The observations that inclusion bodies are essentially homogeneous amorphous aggregates (*24*), and that amyloid fibrils are highly ordered β-sheet structures (*25*), point to some degree of specificity in protein aggregation. The involvement of specific interactions in protein aggregation has been suggested, for example, by the ability of a specific peptide fragment of human growth hormone to inhibit aggregation during refolding of the protein (*26*), and by the observation that the early multimeric intermediates identified along the aggregation pathway for P22 tailspike and P22 coat proteins (*27*), do not aggregate with each other but only among themselves in a mixture of these proteins folding in vitro (*12*). A recent in vivo study using aggregation-prone proteins biosynthetically tagged with fluorophores has shown that nonspecific aggregation between hydrophobic or hydrophobic and hydrophilic proteins does not occur even in mammalian cells. Apparently, specific coaggregation can only occur between proteins that share a common aggregation-promoting motif (*28*). These observations also strongly indicate that protein aggregation is a very specific process because, unlike in bacterial cells, inclusion bodies and aggresomes in mammalian cells are complex structures that contain many proteins including molecular chaperones, components of the ubiquitin-proteasome system, and cytoskeletal proteins (*29*).

Many proteins unfold to partially folded, molten globule-like conformations at low pH (*30*). Not surprisingly, such acid forms of proteins have a strong tendency to aggregate (*30*). A partially folded, molten globule-like conformation, in dynamic equilibrium with the native state of the protein, self-associates and initiates fibril formation in human lysozyme (*16*). Amyloid fibril formation of TTR results, in vitro, from the self-assembly of an intermediate formed during partial acid denaturation of the protein (*31*). Recombinant human prion protein (PrP) can reversibly switch, in vitro, between the native α-helical conformation at high pH, and a highly soluble monomeric form rich in β-structure at low pH and with reduction of the disulfide bond (*32*). The soluble β-form is a direct precursor of fibrillar structures of the protein associated with the disease. More recently, the transition of human PrP to an oligomeric β-sheet structure was observed in the presence of moderate concentrations of urea and small amount of NaCl at acid pH, suggesting the involvement of partially unfolded intermediates in the process (*33*). Indeed, aggregation leading to extracellular amyloid fibril deposition in neurodegenerative diseases, most likely involves unfolding of a normally folded protein under low pH conditions (*4*).

One protein that unfolds reversibly to a molten globule-like A form at acidic pH is the 89-amino acid residue single-domain protein, barstar, which is a natural inhibitor of barnase, an extracellular endoribonuclease in *Bacillus amyloliquefaciens* (*34*). The transition between the soluble A form at low pH and the native (N) state at pH 7 is completely reversible at protein concentrations in the micromolar range (*35*). Native barstar at pH 6.6 is monomeric (10.1 kDa) with three parallel α-helices packed against a three-stranded parallel β-sheet, and with a fourth more poorly defined helix connecting the second β-strand and the third major helix (*36*). The A form of barstar at pH 3 appears to be devoid of well-defined tertiary interactions, has 60% of the native-state secondary structure, has exposed hydrophobic surfaces (*34*), and sedimentation velocity measurements indicate that it is a soluble aggregated form with an apparent molecular mass of about 150 kDa (*35*). The rotational correlation time of the A form determined from time-resolved fluorescence anisotropy decay measurements of IAEDANS[1]-labeled protein was observed to be 90 ns (*37*), as expected for a protein of molecular mass about 160 kDa. The aggregated nature of the A form is also evident in higher concentrations of chemical denaturant being needed to unfold it than are required to unfold the native state (*34*).

The formation of a soluble protein aggregate must precede final formation of an insoluble aggregate, and structural characterization of a soluble protein aggregate, such as the A form of barstar, is expected to be useful in the study of the initial structural changes that eventually lead to irreversible protein aggregation. Overall conformational changes occurring during the aggregation process can be studied by traditional optical techniques, Fourier transform infrared (IR) spectroscopy, light scattering, and small-angle X-ray scatterring measurements. It is known for example, from IR spectroscopy (*38*), X-ray fiber diffraction (*39, 40*), electron microscopy (*41*), and solid-state NMR studies (*42*) that amyloid fibrils form an ordered β-helix structure. The best way to characterize a soluble protein aggregate would be by solution NMR methods that yield structural information at atomic resolution. Moreover, measurements of protein dynamics by NMR is expected to provide an insight into the initial stages of aggregation processes. For instance, the dependence of transverse relaxation rates of backbone ^{15}N nuclei has been used to study the aggregation (polymerization) process of cold-shock protein A (*43*).

Structural characterization by NMR of partially folded and unfolded forms of proteins is often hampered by the poor chemical shift dispersion of amide and carbon resonances in these forms. In some favorable situations, this problem has been overcome by using three-dimensional NOESY experiments and a set of triple resonance experiments such as HNCA, HN(CO)CA, CBCANH, and CBCA(CO)NH on a ^{15}N- and ^{13}C-labeled protein sample. For example, almost complete backbone assignments have been obtained for acid-unfolded apomyoglobin (*44*) and its partially folded state formed at low pH (*45*), but these experiments were not successful in obtaining sequential assignments for HIV-1 protease (*46*). Moreover, NMR structural studies of large molecular weight proteins (including soluble aggregated forms) are not straightforward because of low sensitivity and line broadening due to rapid transverse spin relaxation and extensive signal overlap in the highly complex spectra (*47, 48*).

The recently developed, novel three-dimensional (3D) NMR experimental procedures of HNN and HN(C)N provide a new protocol for the sequential assignment of both folded and unfolded forms of proteins (*49*). The successful applica-

---

[1] Abbreviations: NMR, nuclear magnetic resonance; ANS, 1-anilino-8-naphthalenesulfonate; 1,5-IAEDANS, (5-((((2-iodoacetyl)amino)-ethyl)amino)naphthalene-1-sulfonic acid); HSQC, heteronuclear single-quantum coherence; TOCSY, total correlation spectroscopy; TROSY, transverse relaxation-optimized spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; HX, hydrogen exchange; pH*, pH of D_2O solutions uncorrected for isotope effects, DLS, dynamic light scattering.

tion of this approach to assigning the unfolded form of the HIV-1 protease-tethered dimer, where the standard triple resonance experiments were unsuccessful, has been demonstrated (*46*). The 3D spectra display direct correlations between resonances of the amides and ^{15}N nuclei of the *i*, *i* − 1, and *i* + 1 residues in the ^{15}N plane of residue *i*. An important application of the HNN and HN(C)N based approach is the possibility of studying large proteins including aggregated states of proteins: large and/or aggregated proteins may have regions that are flexible and show sharp resonances, and a few characteristic residues in these regions can then serve as triplet fixed points in the HNN and HN(C)N spectra which enable unambiguous sequential assignments when the primary sequence is known.

In this paper, the novel assignment strategy based on HNN and HN(C)N triple resonance experiments has been used to identify flexible regions in the aggregated A form of barstar at pH 2.7 and 25 °C. These assignments then enabled a rapid analysis of the standard triple resonance spectra in a straightforward manner to obtain the C^{α}, C^{β}, and CO chemical shifts. Backbone assignments and sequential connectivities could be obtained for 21 residues corresponding to 20 N-terminal residues (except Ile13 and Ser14), and a triplet sequence that occurs in helix 4 of the monomeric N state of barstar. The data suggests that at pH 2.7, the A form of barstar exists as a symmetrical, specific aggregated state: only a single set of peaks is observed in the ^{1}H-^{15}N HSQC spectrum. The N-terminal segments of all individual monomeric subunits appear to be hanging free in solution, and the oligomer is so arranged that they are in identical average environments. Secondary shifts, $^{3}J_{HN,H\alpha}$ coupling constants, and the nature of sequential NOE connectivities of these residues show that they have a propensity toward an extended conformation. Medium-range NOE connectivities between a few residues (belonging to helix 1 in native barstar), however, indicate transient formation of medium-range order in the structure. While the N-terminal residues do not show any protection from hydrogen exchange, the dependence on temperature of their H^{N} chemical shifts shows the existence of hydrogen bonds in the free N-terminal segments of the monomeric subunits in the aggregate. ^{15}N relaxation dynamics measurements indicate a highly flexible protein backbone in the free N-terminal segments, as expected for an essentially unfolded polypeptide chain.

MATERIALS AND METHODS

Protein Purification. The pMT316 plasmid containing the barstar gene was expressed in MM294 *Escherichia coli* cells and purified as described earlier (*34*). For uniform labeling of the protein either singly with ^{15}N or doubly with ^{15}N and ^{13}C, the cells were grown for 30 h in M9 minimal medium prepared with $^{15}NH_4Cl$ and ^{13}C-glucose as the only sources of nitrogen and carbon, respectively. The barstar gene expression was induced by IPTG about 6 h after inoculation, and the protein was then purified by the standardized protocol. The yields of labeled protein were 12−15 mg/L culture. The protein used for all experiments was >95% pure.

NMR Sample Preparation. Lyophilized protein was dissolved in 90% H_2O/10% D_2O to a final pH of 2.7 to form the A form of barstar. The protein concentration in all NMR samples was about 1.2 mM. The samples were equilibrated at pH 2.7 for at least 10 h before collecting spectra. For NMR experiments with native barstar, the sample was prepared by dissolving lyophilized protein in 20 mM sodium phosphate buffer/10% D_2O, pH 6.8.

NMR Spectroscopy. The NMR experiments were carried out on a Varian Unity Plus 600 MHz spectrometer at 25 °C. An external DSS sample was used to determine reference frequencies for all nuclei. Typical spectral widths for all two- and three-dimensional experiments recorded were 7002.8 Hz for ^{1}H, 1460 Hz for ^{15}N, 6000 Hz for $^{13}C^{\alpha}$, 13500 Hz for $^{13}C^{\alpha\beta}$, and 2500 Hz for ^{13}CO. Two-dimensional (2D) HSQC spectra were recorded with 512 complex t_1 increments, 8192 t_2 points, and four scans for each fid. A three-dimensional (3D) HNN spectrum was recorded with the following parameters: 32 complex points along t_1 (^{15}N) and t_2 (^{15}N) and 1024 complex points along t_3 (H^{N}), 16 scans for each fid, and $T_N = T_C = 14$ ms. A 3D HN(C)N spectrum of the same sample was recorded using parameters identical to those used for the HNN spectrum. The T_{CC} delay was set to 4.5 ms. The acquisition time for the two experiments was about 26 h each. The following triple resonance experiments were also recorded on the ^{15}N-^{13}C-double-labeled sample: HNCA and HN(CO)CA spectra were acquired with 1024 (H^{N}), 32 (^{15}N), and 96 ($^{13}C^{\alpha}$) complex points; CBCANH and CBCA-(CO)NH spectra were acquired with 1024 (H^{N}), 32 (^{15}N), and 80 (^{13}C) complex points; and an HNCO spectrum was acquired with 1024 (H^{N}), 32 (^{15}N), and 50 (^{13}CO) complex points.

2D ^{1}H TOCSY and ^{1}H-^{15}N HSQC-TOCSY experiments were recorded on a ^{15}N-labeled barstar sample, both with 256 complex t_1 points and 2048 complex t_2 points. A mixing time of 80 ms was used for both experiments. A 3D ^{15}N NOESY−HSQC spectrum was recorded on the ^{15}N-labeled sample with 1024, 32, and 128 complex points along t_3 (H^{N}), t_2 (^{15}N), and t_1 (^{1}H) dimensions, respectively, and with a mixing time of 150 ms. A 2D ^{1}H NOESY spectrum was recorded on an unlabeled protein sample with 256 complex t_1 points and 2048 complex t_2 points, and a mixing time of 150 ms. A 3D HNHA experiment was carried out on the ^{15}N sample with 2048, 48, and 96 complex points along t_3 (H^{N}), t_2 (^{15}N), and t_1 (^{1}H) dimensions, respectively.

Amide Hydrogen Exchange and Temperature Coefficients. A ^{15}N-labeled barstar sample prepared in H_2O at pH 2.7 was lyophilized, and H → D exchange was initiated by dissolving this protein in D_2O, pH 2.7. A series of 2D ^{1}H-^{15}N HSQC spectra, with 70 complex t_1 points (acquisition time 20 min), were collected immediately after starting the HX. The dead time of the experiment was ∼13 min. Amide proton temperature coefficients were determined from a series of ^{1}H-^{15}N HSQC spectra recorded at pH 2.7 over the temperature range 20−30 °C. The temperature coefficients were obtained from a linear regression to all the temperature points for various residues.

Measurement of Relaxation Parameters. All relaxation experiments were performed at 600 MHz at a temperature of 25 °C. R_1 (longitudinal relaxation rate) measurements, R_2 (transverse relaxation rate) measurements, and the steady-state ^{1}H-^{15}N NOE measurements were performed using the pulse sequences described by Farrow et al. (*50*), which use pulse field gradients for coherence transfer pathway selection combined with sensitivity enhancement. The R_1 and R_2 experiments were collected with 512 complex t_1 increments,

2048 t_2 points, and four scans for each fid. Typically for R_1 (= $1/T_1$) measurement, spectra were recorded with eight inversion recovery delays of 10.034, 60.082, 120.165, 240.330, 360.495, 530.729, 761.045, and 1151.581 ms and spectra duplicated at 60.082 and 240.330 ms. For R_2 (= $1/T_2$) measurement, spectra were recorded at eight Carr-Purcell-Meiboom-Gill (CPMG) delays of 8.352, 25.056, 41.760, 66.816, 83.520, 116.928, 141.984, and 192.096 ms and spectra duplicated at 25.056 and 116.928 ms. ^{1}H-^{15}N NOE spectra with 512 complex t_1 increments, 2048 t_2 points, and eight scans for each fid were recorded with and without proton saturation during relaxation delay. NOE experiments were performed in duplicate to estimate the error in measurement. A recycle delay of 5 s was used for the spectrum recorded in the absence of proton saturation, and a 2 s delay was used in the NOE experiment in which protons were saturated. The ^{1}H saturation was achieved by the application of 120° ^{1}H pulses separated by 5 ms, for a period of 3 s.

Data Analysis. All NMR spectra were processed using the Felix 97 software (Molecular Simulations Inc.). All triple resonance spectra were typically apodized using a 90°-shifted square sinebell function before zero-filling and Fourier transformation. R_1, R_2, and steady-state NOE spectra were processed so as to achieve maximum peak heights, and a 60°-shifted square sinebell function was used to improve resolution in the data. Intensities or peak heights (in arbitrary units) for the ^{15}N-^{1}H cross-peaks in these spectra were measured using the Felix software. The uncertainty in peak heights was determined from the spectra collected in duplicate. The intensities of the cross-peaks corresponding to individual residues were then fit to a single-exponential decay function,

$$I(t) = A + Be^{-R_1, R_2 t}$$

to get the values of R_1 or R_2. $I(t)$ is the intensity at recovery delay t (ms) used in the experiments for R_1 and R_2 measurements. $A + B$ is the intensity at time $t = 0$, and A is the steady-state value, that is, the intensity at $t = \infty$. The errors in R_1 and R_2 were estimated as standard errors from the fitting routine.

The {^{1}H}-^{15}N heteronuclear NOEs were calculated from the following equation:

$$\text{NOE} = \frac{I_{\text{sat}}}{I_{\text{eq}}}$$

I_{sat} and I_{eq} are the intensities of a peak in the spectra collected with and without proton saturation, respectively. The errors in the measurement were determined by similarly analyzing the spectra collected in duplicate.

Spectral density functions $J(0)$, $J(\omega_N)$, and $J(\omega_H)$ were calculated as described by Lefèvre et al., by reduced spectral density mapping (*51*). The linear correlation between $J(0)$ and $J(\omega_N)$, and between $J(0)$ and $J(\omega_H)$ was then examined (*51*). Reduced spectral density mapping uses only three ^{15}N relaxation parameters, with the assumption that at high frequencies the spectral density functions $J(\omega_H) \approx J(\omega_H + \omega_N) \approx J(\omega_H - \omega_N)$. By this approach, the spectral density functions are expressed as follows:

$$J(0) = \frac{3}{2(3d' + c')}\left[-\frac{1}{2}R_1 + R_2 - \frac{3}{5}R_{\text{noe}}\right] \quad (1)$$

$$J(\omega_N) = \frac{1}{3d' + c'}\left[R_1 - \frac{7}{5}R_{\text{noe}}\right] \quad (2)$$

$$J(\omega_H) = \frac{1}{5d'}R_{\text{noe}} \quad (3)$$

where,

$$R_{\text{noe}} = (\{^1\text{H} - {}^{15}\text{N}\} \text{ NOE} - 1)R_1(\gamma_N/\gamma_H)$$

The constants c' and d' are approximately equal to 1.25 × 10⁹ (rad/s)² and 1.35 × 10⁹ (rad/s)², respectively, at 14.1 T (*52*). The errors in the spectral density functions were calculated by solving eqs 1–3 for the errors estimated in the ^{15}N relaxation parameters.

Dynamic Light Scattering Measurements. Dynamic light scattering experiments were carried out on a DynaPro-99 machine (Protein Solutions Ltd.) to check if the A form of barstar exists as a homogeneous aggregated population at a concentration of ~1 mM, which is similar to the concentration used in the NMR samples.

RESULTS AND DISCUSSION

Barstar has been observed to form a soluble 160 kDa aggregate in the 1 μM to 1 mM protein concentration range, indicating a low dissociation constant for the aggregate (*35*). This suggested that the A form of barstar exists as a stable oligomer of 16 monomeric subunits in this concentration range. Dynamic light scattering measurements, at the protein concentrations of 1.2–1.4 mM used for NMR, yielded an anomalously high hydrodynamic radius (R_h) of 8 nm (the R_h of monomeric barstar is ≈2 nm). The size distribution obtained by DLS did not change for at least two weeks, indicating that the state of aggregation was stable for this period. After two weeks in solution, the DLS measurements indicated further growth of the aggregate to a heterogeneous population of different sizes of aggregated molecules, finally leading to visible precipitation. This is not surprising since the degree and rate of aggregation are expected to depend on protein concentration because aggregation minimally involves a second-order kinetic process. It is similar to what has normally been observed for proteins in which partially folded intermediates form dimers and soluble oligomers, ultimately leading to the formation of larger insoluble aggregates (*19, 26, 27*). These aggregation processes are therefore effectively irreversible under nativelike conditions. In the case of barstar, the time course of formation of larger aggregates from the 16-mer appears to be an extremely slow process and is being investigated in greater detail. The A form of barstar provides a good model to study aggregated states of proteins and the aggregation process, because it remains stable in its soluble form for a long time and at high enough concentrations to characterize it by solution NMR.

^{1}H-^{15}N HSQC Spectrum of the A form of Barstar. Figure 1A shows the 2D ^{1}H-^{15}N HSQC spectrum of native barstar at pH 6.8 and 25 °C. The ^{1}H-^{15}N HSQC spectrum of the A form of barstar at pH 2.7 and 25 °C is shown in Figure 1B. Only 32 cross-peaks are observed in the A form spectrum. The chemical shifts of H^N are within a range of 0.7 ppm
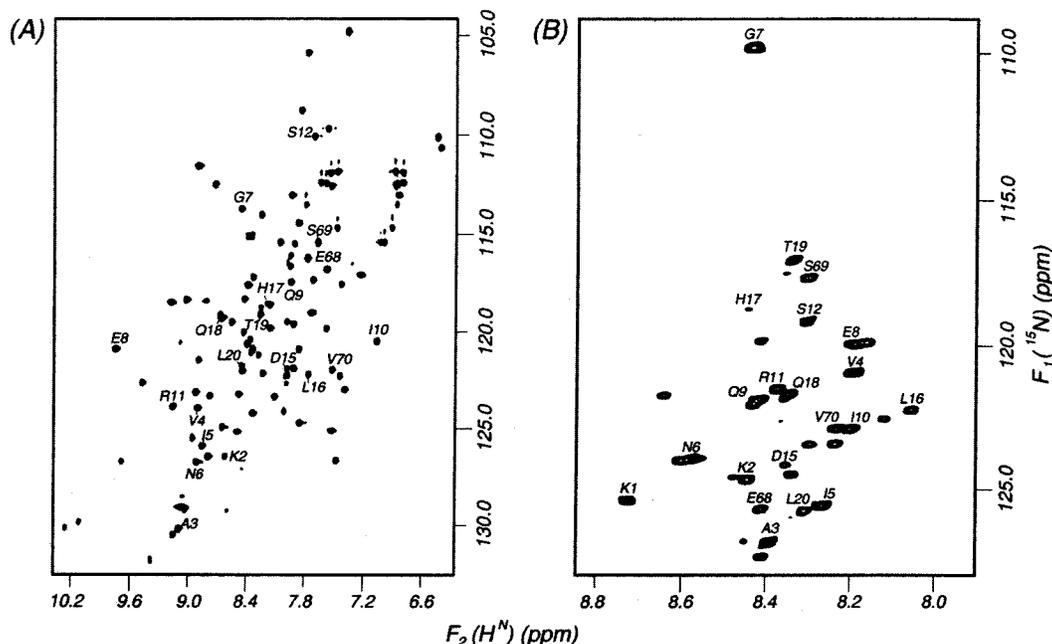
FIGURE 1: ^{1}H-^{15}N HSQC spectra of (A) native barstar at pH 6.8, and (B) A form of barstar at pH 2.7, and 25 °C. In panels A and B, cross-peaks corresponding to only those residues have been labeled for which spin systems could be unambiguously assigned at pH 2.7, and sequential connectivities obtained from the HNN and HN(C)N spectra, the HNCA and HN(CO)CA spectra, and the NOESY spectra. The residues in the ^{1}H-^{15}N HSQC spectrum of native barstar in panel A have been marked comparing peak positions with ref *81*.

and those of ^{15}N are within a range of 20 ppm. The small chemical shift dispersion is characteristic of unfolded proteins. Moreover, the A form of barstar is an aggregate. It is, therefore, surprising that about 30 well-resolved resonances could be observed at all. In contrast, the H^N and ^{15}N dispersion in the HSQC spectrum of folded barstar is about 4 and 28 ppm, respectively. It should be mentioned that a TROSY pulse sequence did not resolve any more peaks than were observed in the HSQC spectrum of the A form of barstar, although the TROSY technique has been successfully used to improve spectral resolution and sensitivity enabling structural studies on high molecular weight proteins (*53*).

Sequential Connectivities in the A form of Barstar. The recently developed strategy for sequential assignments of proteins based on HNN and HN(C)N triple resonance NMR experiments has been exploited here to obtain structural information about the aggregated A form of barstar formed at low pH. The HNN experiment generates correlations from residue i to both the $i - 1$ and $i + 1$ residues in the protein primary sequence, while the HN(C)N experiment generates only i to $i - 1$ correlations. The two experiments together are useful in obtaining rapid sequential assignments of the H^N and ^{15}N chemical shifts of individual residues in a protein, as described earlier (*49*). In the HNN spectrum, the F_1-F_3 plane at the F_2 chemical shift of residue i displays self correlations and sequential correlations to ^{15}N chemical shifts of residues $i - 1$ and $i + 1$ at the amide position of i; the F_2-F_3 plane at the F_1 chemical shift of i displays all three correlations at their respective amide positions. In the HN(C)N spectrum, the F_1-F_3 plane at the F_2 chemical shift of residue i displays self and sequential correlations to the $i + 1$ residue. Thus, in these spectra, unlike the conventional 3D experiments, repeated scanning through the ^{15}N planes to locate peaks at the desired chemical shifts is not required. On the basis of the positive and negative peak patterns obtained in the HNN and HN(C)N spectra, proline and

glycine neighbors can be easily identified in the primary sequence of the protein. Furthermore, glycines are distinguishable as showing a distinctly upfield ^{15}N chemical shift in both HNN and HN(C)N spectra. When glycine is the central residue in a triplet in the HNN spectrum, its diagonal peak is given a negative sign; the $i - 1$ peak is also negative and the $i + 1$ peak is always positive (product operator formalism described in ref *49*). Again, the glycine diagonal peak is given a negative sign at the $i - 1$ position in a HN(C)N spectrum (*49*). Thus, glycine residues can serve as starting fixed points in the sequential walk through the HNN spectrum of a protein whose primary sequence is known.

In the HNN spectrum of the A form of barstar, one glycine cross-peak was observed, which served as a starting fixed point to obtain sequential connectivities. In conjunction with spin system identification from 2D TOCSY spectra (described below), this glycine was identified to be Gly7 in the barstar primary sequence. Figure 2A shows the sequential F_1-F_3 planes at the F_2 chemical shifts of residues Lys1 to Ile10 in the N-terminal segment of barstar. In any given F_1-F_3 plane, distinction between the peaks corresponding to the $i - 1$ and $i + 1$ residues was confirmed by comparison with the identical plane in the HN(C)N spectrum. There was a break in the sequential walk at the Ile5−Asn6 segment, as seen in Figure 2A. This was bridged, however, by sequential connectivities obtained from the triple resonance HNCA and HN(CO)CA experiments. The sequential walk through the HNCA spectrum for the first 10 residues in barstar at pH 2.7 is shown in Figure 2B.

Of the 30 resonances observed in the HSQC spectrum, sequential connectivities for 18 residues could be unambiguously established from the HNN, HN(C)N, HNCA, and HN(CO)CA spectra. These include the first 12 residues in the N-terminal segment of barstar (Figure 2), indicating that the N-terminal segments of all the monomeric subunits in the A form are free and not a part of the aggregated core. A
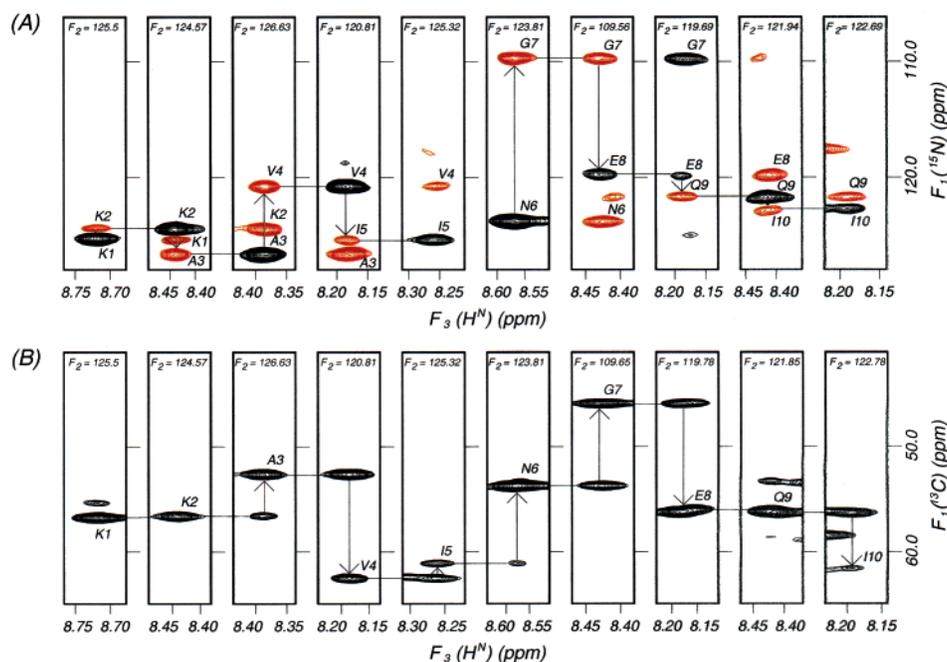
FIGURE 2: Illustrative sequential walks through F_1-F_3 planes in the (A) HNN and (B) HNCA spectra of the A form of barstar at pH 2.7. Sequential connectivities are indicated for the Lys1 to Ile10 stretch in the protein primary sequence. Strips from the spectra at the appropriate H^N chemical shifts are shown. Black and red contours are positive and negative peaks, respectively. The distinct Gly plane in the HNN spectrum served as a starting triplet fixed point to obtain sequential connectivities. The F_2 (^{15}N) chemical shift is indicated at the top of each strip of the HNN and HNCA spectra.

few other resonances in the protein could be identified suggesting that these must be part of flexible regions in the aggregate. Sequential connectivities for two more triplet sequences in the protein could be found by these experiments (data not shown). That these peaks correspond to Gln18−Thr19−Leu20 and Glu68−Ser69−Val70 in the barstar primary sequence was determined after spin system assignments from TOCSY experiments (described below). The former triplet of amino acids occurs in helix 1 in native barstar and lies close to the N-terminal which does not appear to be aggregated in the A form. The other triplet sequence forms part of helix 4 in the native protein. The assigned residues in the A form of barstar are indicated in the $^1H-^{15}N$ HSQC spectrum of barstar at pH 2.7 in Figure 1B. No sequential connectivities were discernible for the other peaks visible in the HSQC spectrum, although the self-peaks for a few of them were present in the appropriate F_1-F_3 planes in the HNN spectrum. It is possible that the resonances of their sequential $i − 1$ and $i + 1$ partners are broadened out and, therefore, not observed in the HNN spectrum. No more sequential connectivities could be identified in the other triple resonance experiments as well.

It is important to note that only one peak for every assigned residue is observed in the HSQC spectrum of the A form of barstar, indicating that all these residues are equivalent and that the aggregate is symmetric. The residues that could be assigned in the low pH state have also been indicated in the $^1H-^{15}N$ HSQC spectrum of native barstar shown in Figure 1A. Comparing the H^N and ^{15}N chemical shifts of these peaks with those in the HSQC spectrum collected at pH 2.7 and 25 °C shows that the chemical shifts are very different in the aggregate. This indicates that the environment of these residues in the A form is definitely different from that in native barstar.

Analogous to this is the example of the cartilage matrix protein C-terminal domain (CMPcc) structure (*54, 55*). CMPcc corresponds to 43 residues of the full-length protein and self-assembles into a homotrimer of 16.1 kDa. NMR spectra of CMPcc show only resonances corresponding to the primary sequence of the monomers, indicating an identical magnetic environment for each of the three polypeptide chains and thus a parallel coiled-coil structure with 3-fold symmetry (*55*).

Homonuclear Sequential NOE Connectivities. The spin systems of the resolved peaks were determined from two-dimensional ^1H TOCSY and $^1H-^{15}N$ HSQC-TOCSY experiments. For the 18 peaks that showed sequential connectivities in the HNN and HNCA spectra, spin systems could be unambiguously assigned from the TOCSY experiments. Homonuclear sequential NOE connectivities for these residues were also detected in the 3D NOESY−HSQC experiment, though there is partial overlap in some of the self and sequential peaks due to insufficient dispersion among the H^α proton resonances. Figure 3A shows the $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ NOE connectivites from Lys1 to Ser12 in the N-terminal segment of barstar. Strips of the 3D NOESY spectrum in appropriate F_1-F_3 planes are shown at the respective F_2 chemical shifts of residues from Lys1 through Ser12.

A few more spin systems in the TOCSY spectra were tentatively identified initially, based on random coil chemical shifts. Two of these belonged to either Asn or Asp residues, one to Glu or Gln, one to Leu, and one to His. There is a unique His residue in barstar at position 17 in helix 1, and the His spin system identified in the TOCSY spectra must correspond to this. His17 is preceded by a Leu residue and followed by a Gln residue in the primary sequence of the protein. Although only the diagonal peak of His17 was
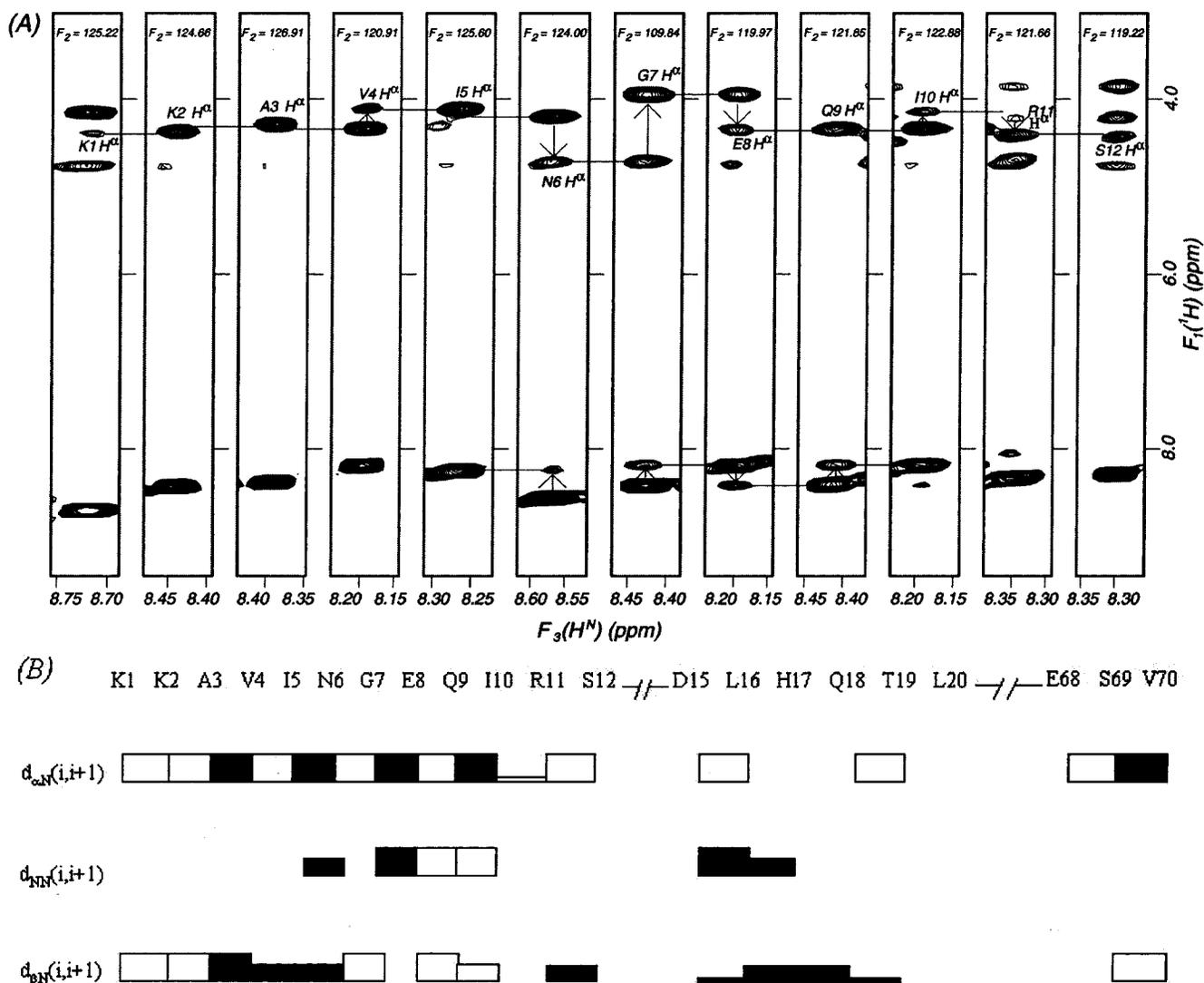
FIGURE 3: Homonuclear sequential NOE connectivities in barstar at pH 2.7. (A) Strips of different F_1-F_3 planes in the 3D NOESY-HSQC spectrum of the A form of barstar. The observed $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ sequential connectivities in the N-terminal region of barstar from Lys1 through Ser12 are shown here. The F_2 (^{15}N) chemical shift is indicated at the top of each strip. (B) Summary of homonuclear sequential NOE connectivities observed in the A form of barstar at pH 2.7. The thickness of the bars indicates the relative intensities of the NOE cross-peaks observed in the 3D NOESY−HSQC spectrum. Unfilled bars signify either overlapped NOE cross-peaks with the self-peaks or unresolved $C^\beta H$ resonances for a residue. Absence of a bar indicates that no NOE was observed in the particular segment of the protein.

visible at the appropriate chemical shifts for this spin system in the 3D NOESY−HSQC spectrum, it clearly showed $d_{NN}(i,i-1)$ and $d_{\beta N}(i-1,i)$ NOE connectivites to the Leu16 spin system and $d_{\beta N}(i,i+1)$ NOE connectivities to the Gln18 in the 2D ^{1}H NOESY spectrum of the A form of barstar. This Leu spin system, identified in the TOCSY experiments, and present in the 3D NOESY spectrum at the H^N chemical shift and ^{15}N plane corresponding to an unassigned peak in the HSQC spectrum, showed sequential NOE cross-peaks, in both the 3D and 2D NOESY experiments, to an Asp or Asn spin system. It must, therefore, correspond to the Asp15−Leu16 pair in the primary sequence of barstar. No sequential connectivities were, however, observed in the HNN or HNCA spectra for the Asp15−Leu16−His17 triplet. This may be attributed to the low sensitivity of these spectra.

A summary of the sequential NOEs observed in the A form of barstar for all the assigned residues is shown in Figure 3B. The figure shows that $d_{\alpha N}(i,i+1)$ NOE connec-

tivities are observed between most of the assigned residues, indicating that all the identifiable flexible segments in the A form sample the β-region of (ϕ, ψ) space. The Asn6−Ile10 and Asp15−His17 segments also show $d_{NN}(i,i+1)$ NOE connectivities, indicating that these regions of the polypeptide backbone fluctuate over both α and β regions of (ϕ, ψ) space. The relative populations of α and β backbone dihedral angles could not be assessed from the relative intensities of the sequential NOEs, since many of the peaks were partially overlapped with self-peaks (Figure 3B).

The H^N positions of the other spin systems identified in the TOCSY experiments matched with some of the unassigned peaks in the HSQC spectrum. The diagonal peaks for most of these unassigned residues were observed in the 3D NOESY spectrum, but no NOE cross-peaks were detected in the F_1-F_3 planes at their respective F_2 chemical shifts. Since no sequential connectivities were found for these residues even in the HNN and other triple resonance

experiments, it would mean that they lie in regions that are flexible in the aggregated core of the A form; the resonance frequencies of their sequential partners, however, are too broadened to be observed above the noise level.

Homonuclear Medium Range NOE Connectivities. A few medium range NOE connectivities between backbone atoms as well as with some side-chain atoms were observed in the 3D NOESY−HSQC spectrum of the A form of barstar (data not shown). Their origin could be identified on the basis of the sequential assignments described above and the knowledge of the spin system assignments. In a few cases where ambiguities arose owing to poor dispersion of chemical shifts, a judgment was made assuming that some native propensities would be observable even in unfolded proteins; this has been seen in many proteins (see for example refs 44−46). A brief description of these medium range NOEs is given below.

The Gly7 F_1−F_3 plane of the 3D NOESY−HSQC spectrum showed an NOE connectivity from Gly7 H^N to a γCH_3 pseudoatom, which was attributed to Ile5, based on the 5.06 Å distance between them in native barstar. Ile5 lies in the N-terminal β-strand 1 of native barstar while Gly7 is the first residue of the loop following it and extending up to Ser12 (36). (The proton−proton distances are from the three-dimensional structure of barstar, 1BTA, submitted by Lubienski et al. in the PDB). The Glu8 and Gln9 F_1−F_3 planes also show NOE peaks that may arise because of medium range connectivities from the respective H^N to the γCH_3 protons of Ile5 and Ile10, respectively. The proton−proton distance between the H^N of Glu8 and the γCH_3 pseudoatom of Ile5 is 7.43 Å, and that between the H^N of Gln9 and the γCH_3 pseudoatom of Ile10 is 6.6 Å in native barstar. Although these distances are large to give rise to NOE peaks in folded monomeric barstar, these residues possibly come close to each other transiently in the free N-terminal of the aggregated A form of the protein to give rise to NOE effects. There appears to be a weak medium range $d_{\alpha N}(i,i+2)$ connectivity between Ile10 and Ser12 in the NOESY spectrum. The proton−proton distance between these atoms in native barstar is 3.52 Å. This peak may be partially overlapped with a $d_{\alpha N}(i,i-4)$ NOE between the Ser12 H^N and the Leu16 $C^{\alpha H}$; the distance between which in native barstar is 5.76 Å (36). Arg11 shows a $d_{\beta N}(i,i-1)$ NOE connectivity to Ser12; the distances between the H^N proton of Arg11 and $C^{\beta H}1$ and $C^{\beta H}2$ protons of Ser12 are, respectively, 5.25 and 5.54 Å in native barstar. (The two Ser12 $C^{\beta H}$ protons are not resolved as separate peaks in the NOESY spectrum). Ile10, Arg11, and Ser12 occur in the loop connecting β-strand 1 to helix 1 in native barstar (36). These medium range NOE peaks, though weak, in the segment Gly7−Ser12 in the A form, seem to indicate structure formation which must, however, be transient since the N-terminal segment appears to have a highly dynamic backbone (discussed below).

A few medium range NOE connectivities are also observed in the F_1−F_3 plane of Gln18 in the 3D NOESY−HSQC spectrum of the A form. Gln18 shows $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+2)$, and $d_{\beta N}(i,i+2)$ connectivities to Leu16. It also possibly shows $d_{\alpha N}(i,i+3)$ and $d_{\beta N}(i,i+3)$ connectivities to Asp15, but these peaks have partial overlap with self and sequential peaks of Arg11 whose H^N and ^{15}N chemical shifts lie very close to those of Gln18. The proton−proton distances between each of these pairs of atoms lies in the range of 3−5 Å in native

barstar. The occurrence of medium range $d_{\alpha N}(i,i+2)$ and $d_{\alpha N}(i,i+3)$ NOE connectivities is consistent with the observation of sequential $d_{NN}(i,i+1)$ NOE connectivities (discussed above) between these residues, and seems to indicate that the Asp15−Gln18 segment also populates the α region of (ϕ, ψ) space in the A form. These residues form part of helix 1 (Ile13−Glu23) in native barstar (36). Since no medium range NOEs were observed for the rest of the assigned residues in helix 1, viz, Asp15, Leu16, His17, Thr19, and Leu20, this region probably forms a transient structure with some medium range order, and not a well-formed helix. A bend formed in the protein backbone, for example, could also give rise to such localized medium range NOEs.

Deviation of Chemical Shifts from Random Coil Values. The deviations of specific chemical shifts (secondary shifts) in proteins from their random coil values are highly sensitive to conformational preferences of the protein backbone (56). The resonance assignments for $^{13}C^\alpha$, $^{13}C^\beta$, and ^{13}CO for the assigned peaks in the A form of barstar were obtained from HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH, and HNCO experiments. The resonance assignments for $^1H^\alpha$ were obtained from the TOCSY spectra. Since the chemical shifts of certain nuclei are influenced both by the neighboring amino acids and by the local backbone structure, it was important to correct these for contributions from the local amino acid sequence. The random coil values (57) for all the residues used in the secondary shift analysis were corrected using sequence-dependent correction factors determined for a set of Ac-GGXGG-NH$_2$ peptides in 8 M urea at pH 2.3 (58). Deviations in specific chemical shifts were then calculated by subtracting the corrected random coil values from the measured chemical shifts for the assigned residues in the A form. Positive secondary shifts for $^{13}C^\alpha$ and ^{13}CO indicate a preference for ϕ, ψ angles characteristic of helical conformations, while negative secondary shifts indicate a preference for ϕ, ψ angles in the β-sheet region. The trend is opposite for $^1H^\alpha$ and $^{13}C^\beta$ chemical shifts; positive and negative secondary shifts show a propensity for extended and helical conformations, respectively.

Secondary shifts in the A form at pH 2.7 are shown in Figure 4. For the N-terminal residues, $^{13}C^\alpha$ shows negative secondary shifts, indicating a propensity of the polypeptide backbone toward a β-strand conformation in this region. This is confirmed by the negative secondary shifts observed for ^{13}CO as well. The residues belonging to helices 1 and 4 in native barstar also show negative $^{13}C^\alpha$ and ^{13}CO secondary shifts, indicating a preference for backbone dihedral angles in the β-region of (ϕ, ψ) space. This region of the Ramachandran plot is also most populated in random coil polypeptides (59). Thus, the negative $^{13}C^\alpha$ and ^{13}CO secondary shifts may not differentiate between a conformational propensity toward β-structure or a random coil. The $^1H^\alpha$ chemical shifts show essentially no deviation from the random coil values and, therefore, do not indicate backbone conformational propensities for barstar at pH 2.7. The $^{13}C^\beta$ secondary shifts seem to show a negative trend. These have, however, not been corrected for contributions from the local amino acid sequence, and may not be diagnostic of the conformational preferences of ϕ, ψ angles in the N-terminal and other regions of the A form.

Backbone $^3J_{HN,H\alpha}$ Coupling Constants. The $^3J_{HN,H\alpha}$ coupling constant is sensitive to the dihedral angle ϕ, and thus
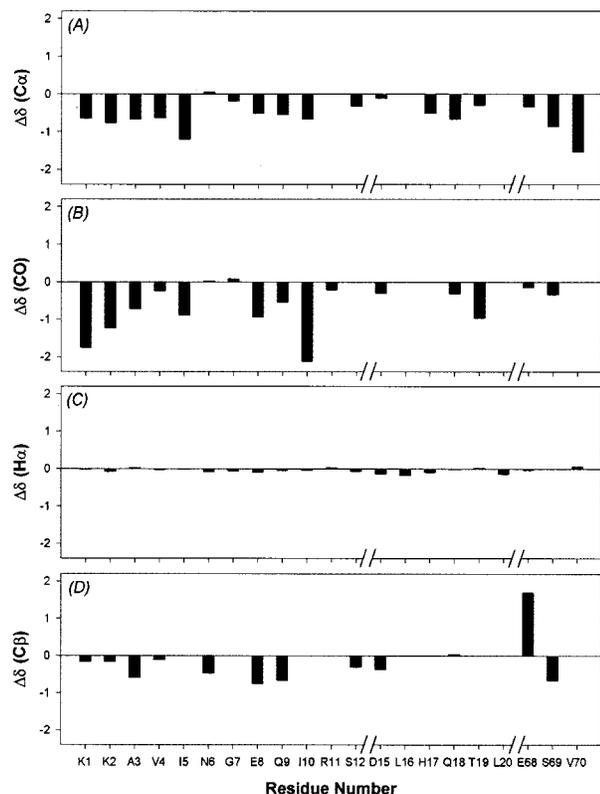
FIGURE 4: Secondary chemical shifts for (A) $^{13}C^{\alpha}$, (B) ^{13}CO, (C) $^{1}H^{\alpha}$, and (D) $^{13}C^{\beta}$ resonances of the assigned residues in the A form of barstar at pH 2.7. The chemical shifts have been corrected for contributions from the local amino acid sequence (58). Reference random coil shifts used are those determined for the peptide Ac-GGXGG-NH$_2$ in 8 M urea at pH 2.3 (57).

Table 1: $^3J_{HN,H\alpha}$ Coupling Constants for Non-Glycine Residues in the Flexible Regions of the A Form of Barstar at pH 2.7 and 25 °C

residue	measured $^3J_{HN,H\alpha}$ coupling constants (Hz)[a]	random coil $^3J_{HN,H\alpha}$ coupling constants (Hz)[b]
K1	6.5	
K2	6.4	6.5
A3	5.6	5.7
V4	8.1	7.3
I5	7.9	7.6
N6	6.2	7.3
E8	6.4	
Q9	6.8	6.1
I10	7.0	7.0
R11	6.0	6.6
S12	6.0	6.4
D15		6.5
L16	6.2	6.5
Q18		6.6
T19	6.7	7.4
L20	5.7	6.6
E68	8.0	5.9
S69	6.5	6.4
V70	6.7	7.3

[a] $^3J_{HN,H\alpha}$ coupling constants (uncorrected for relaxation) measured for barstar at pH 2.7 from a 3D HNHA spectrum. [b] Random coil J values corrected for sequence context (62). There is no value for E8 since coupling constant was not measured for G7.

Random coil temperature coefficients were subtracted from the measured H^N shifts. Figure 5B shows the deviation from random coil values of the H^N temperature coefficients in the A form at pH 2.7. Lowered temperature coefficients are evident for many of the residues analyzed, with Lys1, Glu8, Gln9, Arg11, Leu16, His17, Gln18, Leu20, and Val70 showing a deviation of $\geq$ 1 ppb/K, indicating that these residues are involved, at least transiently, in hydrogen bonding. The H^N temperature coefficients cannot distinguish between intra- and intermolecular hydrogen bonds. Thus, the hydrogen bonding may be between residues in the same monomer or between neighboring monomeric subunits in the aggregated A form. The hydrogen bonds, however, must be rapidly fluctuating since the ^{15}N relaxation measurements indicate a highly flexible polypeptide backbone in this segment (discussed below).

Hydrogen Exchange. An H $\rightarrow$ D HX experiment was carried out to check if any of the assigned amide protons show protection from hydrogen exchange in the A form of barstar. While most residues exchanged with the solvent deuterons within the dead time of the experiment (about 13 min), 7 H^N protons exchanged more slowly. These include the amide protons of Ala3, Val4, Ile5, Gln9, Ile10, Leu20, and Val70. The ^{1}H-^{15}N HSQC spectrum collected 13 min after initiating hydrogen exchange in D$_2$O is shown in Figure 6. Val4 and Ile10 exchanged out after 40 min, and the cross-peak corresponding to Ile5 disappeared only after 60 min of starting the HX data collection. The observed time constants of HX for all these amide protons are similar to their expected intrinsic exchange rates at pH 2.7 and 25 °C in barstar (66). Thus, all residues in the free N-terminals of the A form, and also the three assigned residues in helix 4, are completely exposed to HX, with only Gln9 showing a marginal protection factor of ~5, suggesting essentially random coil-like conformations for these residues.

Backbone Dynamics. The longitudinal and transverse relaxation rates (R_1 and R_2) of backbone ^{15}N nuclei as well

provides a probe for backbone conformational preferences (60). The $^3J_{HN,H\alpha}$ coupling constants for the A form at pH 2.7 were measured from a 3D HNHA spectrum, as described in ref 61, for all the well-resolved strong peaks in the HSQC spectrum, and are listed in Table 1 along with the values of statistical random coil coupling constants (62). These random coil J values were corrected for residue type based on the nature of the preceding amino acid residue (62). The measured $^3J_{HN,H\alpha}$ coupling constants show only small deviations, < 1 Hz, from the expected random coil values for most of the residues, and are, thus, poor indicators of major backbone conformational preferences (63). Glu68, however, shows a large $^3J_{HN,H\alpha}$ coupling constant, suggesting a strong bias toward the β-region of (ϕ, ψ) space.

Hydrogen Bonding. The temperature dependence of the H^N chemical shift, that is temperature coefficient, provides an estimate of the involvement of the amide proton in hydrogen bonding (63, 64). Random coil temperature coefficients determined for residue X in a series of GGXGG unstructured peptide models at pH 5 over the temperature range 278 to 318 K are around -8 ppb/K (65). ^{1}H-^{15}N HSQC spectra were collected for the A form at pH 2.7 over the temperature range 293–303 K. Temperature coefficients were calculated for the H^N chemical shifts of all the assigned residues in the A form over this temperature range. The data were highly linear, as is evident from Figure 5A. The H^N temperature coefficients calculated by linear regression analysis for the 21 assigned peaks in the A form of barstar, along with those in a random coil are given in Table 2.

FIGURE 5: (A) Linear regression analysis of backbone H^N chemical shifts in the A form of barstar over the temperature range 293–303 K. Representative data are shown for Lys2 (●), Ile5 (□), Glu8 (△), and Gln18 (▼). The solid lines are the linear fits through the data, all with R^2 values ≥ 0.99. (B) Deviation of the temperature coefficients from their random coil values (65) for backbone H^N chemical shifts in the A form of barstar at pH 2.7. The temperature dependence of the H^N chemical shifts has been determined for only the unambiguously assigned residues.

as the ^1H-^{15}N steady-state heteronuclear NOE are useful probes of protein backbone dynamics and overall molecular tumbling motions (67, 68). While all the three relaxation parameters are sensitive to motions on a picosecond to nanosecond time scale, the ^1H-^{15}N NOE is most sensitive to high-frequency motions of the protein backbone. Likewise, transverse relaxation is quite sensitive to slow time scale (micro- to millisecond) motions and conformational exchange. The ^{15}N longitudinal (T_1) and transverse (T_2) relaxation data for a few of the assigned residues in the A form of barstar is shown in Figure 7. The relaxation rates, R_1 and R_2, and heteronuclear NOEs for 20 assigned residues in the A form were measured at pH 2.7, as described in the Materials and Methods section, and are plotted as bar graphs in Figure 8. The ^{15}N relaxation parameters of the N-terminal segments and the Glu68–Ser69–Val70 triplet in helix 4 of the monomeric subunits in the aggregated A form indicate a highly flexible polypeptide chain in these regions. This is evident from the negative values of heteronuclear NOEs (Figure 8C), which indicate occurrence of large-amplitude motions on a subnanosecond time scale. The residues toward the N-terminal of the protein show more negative NOEs. Interestingly, the R_2 values for Ser12 and Ser69 are significantly larger than those for the rest of the residues, which

Table 2: H^N Temperature Coefficients in the A Form of Barstar at pH 2.7 Measured over the temperature range 293–303 K

residue	measured H^N temperature coefficients (ppb/K)	random coil H^N temperature coefficients (ppb/K)[a]
K1	−5.44	−7.87
K2	−7.34	−7.87
A3	−7.46	−8.20
V4	−8.63	−8.35
I5	−8.40	−8.35
N6	−8.03	−7.02
G7	−6.63	−7.02
E8	−4.83	−7.01
Q9	−6.07	−7.65
I10	−8.03	−8.35
R11	−6.29	−7.64
S12	−6.36	−7.02
D15	−7.89	−6.43
L16	−5.40	−8.42
H17	−4.53	−7.49
Q18	−6.24	−7.65
T19	−7.24	−7.40
L20	−7.44	−8.42
E68	−7.19	−7.01
S69	−7.34	−7.02
V70	−7.49	−8.35

[a] Determined for residue X in a series of GGXGG unstructured peptide models at pH 5 over the temperature range 278–318 K (65).
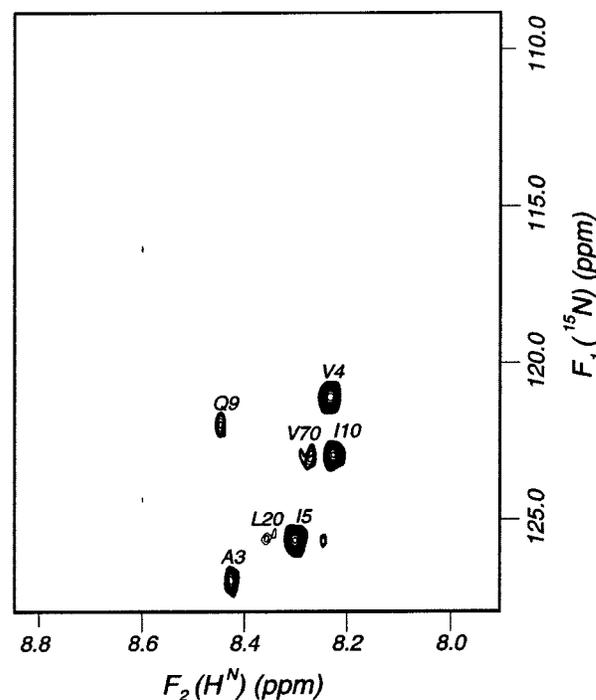
FIGURE 6: ^1H-^{15}N HSQC spectrum of the A form of barstar at pH 2.7 after H → D exchange was initiated in D_2O. The dead time of the hydrogen exchange experiment was 13 min. The relatively slow-exchanging residues have been indicated in the spectrum.

suggests the presence of significant conformational exchange contributions for these two residues. It is important to note that Ser12 is the last residue of the loop connecting β-strand 1 to helix 1 in the native barstar structure (36) and loop residues are often found to exhibit conformational dynamism. It may be mentioned that, even in folded monomeric barstar, the relaxation data indicated significant conformational dynamism in the residues Ile13, Glu68, and Ser69 among others (69).
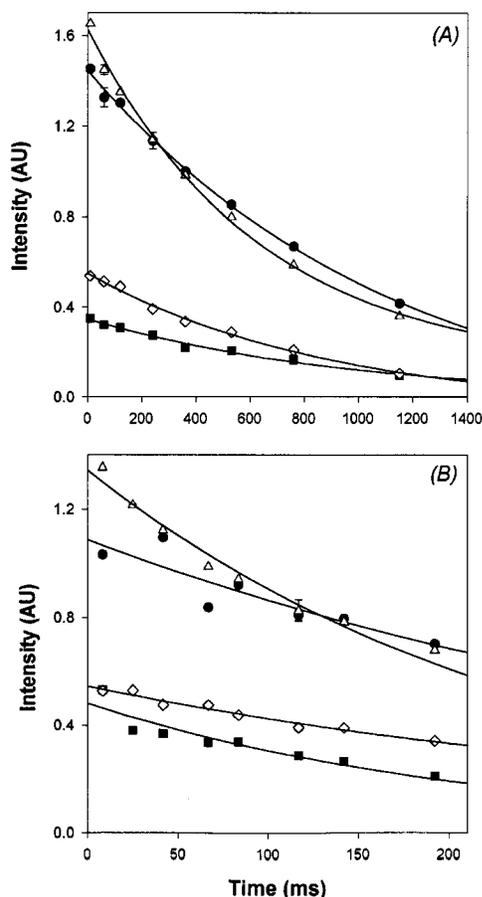
FIGURE 7: ^{15}N relaxation data for the measurement of R_1 and R_2. The figure depicts representative plots of intensities in arbitrary units vs relaxation delays showing (A) T_1 relaxation and (B) T_2 relaxation of Lys2 (●), Glu8 (△), Gln18 (■), and Val70 (◇). The solid lines represent fits of the data to a single-exponential function, as described in Materials and Methods.

FIGURE 8: Relaxation parameters for the A form of barstar at pH 2.7 and 25 °C. (A) R_1 relaxation rates, (B) R_2 relaxation rates, and (C) {^{1}H}-^{15}N heteronuclear NOEs are plotted as a function of residue number in the protein sequence. The horizontal lines in panels A and B indicate the mean R_1 and R_2 values of 1.2 and 3.3 s^{-1}, respectively. The mean value of R_2 was calculated omitting the R_2 values of Ser12 and Ser69 residues, which have exchange contributions.

The spectral density function, $J(\omega)$, of the protein backbone was calculated from the ^{15}N relaxation parameters at three specific frequencies, viz., $\omega = 0$, ω_N, and ω_H. The spectral density functions were determined using the reduced spectral density mapping approach (*51*) from eqs 1−3 described in the Materials and Methods section. $J(\omega_H)$ is largely determined by heteronuclear NOEs and is thus most sensitive to higher frequency motions of the protein backbone. $J(\omega_N)$ is dominated by R_1 and $J(0)$ is dominated by both R_1 and R_2. Thus, $J(0)$ is sensitive to both nanosecond time scale motions and also contributions from slower micro- to millisecond exchange processes. The bar graphs in Figure 9 show the spectral densities of the A form as a function of residue number. In a protein, the area under the spectral density curve, a Lorentzian function of frequency, is a constant and does not vary from one NH vector to another (*70*). Smaller values of $J(0)$ are compensated, therefore, by larger values of spectral densities at higher frequencies, suggesting fast internal motions at frequencies approaching ω_N and ω_H. It is evident from Figure 9A that most of the assigned residues in the A form show a value of the $J(0)$ spectral density lower than what is seen in native barstar (*69*) (discussed below). The low $J(0)$ values are compensated by higher values of the $J(\omega_N)$ and $J(\omega_H)$ spectral density functions, as seen in Figure 9B,C. The spectral density functions, again, indicate high frequency motions of a highly

flexible protein backbone in the N-terminal segments as well as the identifiable region of helix 4 of the monomeric subunits in the A form. Here again, it is observed that, for Ser12 and Ser69, the $J(0)$ value is significantly larger than for the rest of the residues, indicating contributions from slow time scale conformational exchanges.

The ^{15}N relaxation parameters and spectral density functions have been calculated earlier for 69 residues in native monomeric barstar (*69*). The R_1 and R_2 values in native barstar are 1.6−1.9-fold higher than those observed for the flexible N-terminal segment in the A form of the protein. Lower values of R_2 reflect the relatively faster internal motions in the essentially unstructured N-terminal segment of barstar at low pH. Moreover, positive NOEs are observed in the monomer, as is expected for a folded protein, in contrast to all the assigned residues showing negative NOEs in the A form. Negative NOEs have also been observed for the unfolded states of apomyoglobin (*44*) and apo-plastocyanin (*71*). A striking feature of the spectral densities in the flexible regions of the A form are the low values of $J(0)$, and the large values of $J(\omega_H)$, relative to those typically found for folded globular proteins. The $J(0)$ values for all 69 residues in native barstar are at least 2-fold higher than the average $J(0)$ value of 0.86 ns/rad (0.72 ns/rad if Ser12 and Ser69 are excluded) in the A form, while the $J(\omega_H)$ values are 3−4-fold lower than the average $J(\omega_H)$ value of 0.03 ns/rad in the A form. Low $J(0)$ and high $J(\omega_H)$ values of
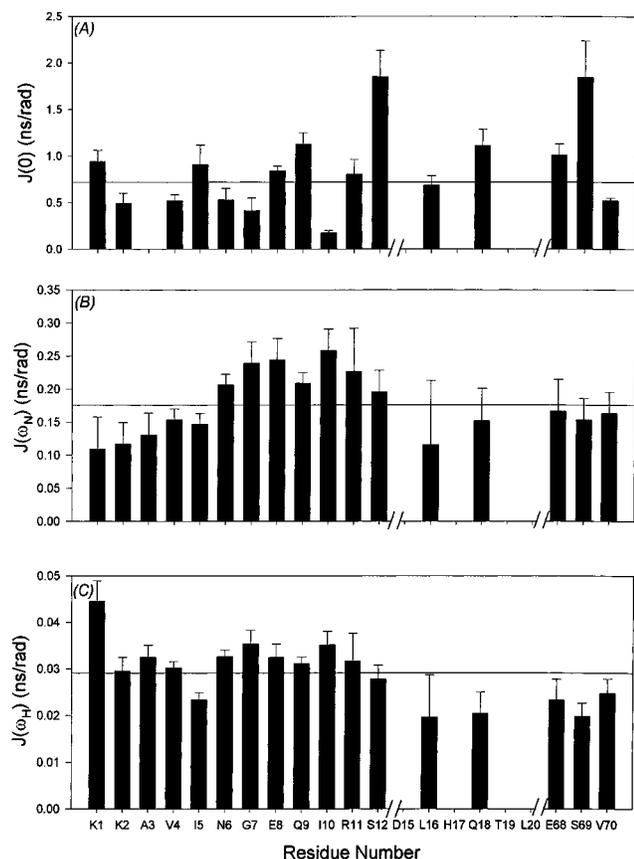
FIGURE 9: Spectral density functions obtained from reduced spectral density mapping (*51*) for aggregated barstar at pH 2.7. Calculated values of (A) $J(0)$, (B) $J(\omega_N)$, and (C) $J(\omega_H)$ are plotted versus residue number in the protein sequence. The horizontal lines in all panels indicate the mean $J(0)$, $J(\omega_N)$, and $J(\omega_H)$ values of 0.72, 0.18, and 0.03 ns rad^{-1}, respectively. The mean $J(0)$ value was calculated omitting the $J(0)$ values of Ser12 and Ser69 residues, which have exchange contributions.

FIGURE 10: Plots of (A) $J(\omega_N)$ and (B) $J(\omega_H)$ as a function of $J(0)$. The solid lines are fits to eq 4 (see text). The α and β values (see text) for $J(\omega_N)$ vs $J(0)$ are -0.0191 and 0.1949 ns rad^{-1}, respectively, and the correlation coefficient for the fit is 35%. The α and β values for $J(\omega_H)$ vs $J(0)$ are -0.0054 and 0.0335 ns rad^{-1}, respectively, and the correlation coefficient for the fit is 14%.

similar magnitude as seen here, were also observed for the acid-unfolded state of apomyoglobin (*44*) and the unfolded state of apoplastocyanin (*71*). As for unfolded proteins that have little ordered structure (*71, 72*), the spectral density function for the free N-terminal regions in the A form of barstar is flat and extended (Figure 9), indicating that their relaxation is dominated by rapid backbone motions occurring on a subnanosecond time scale. No information about backbone dynamics in the aggregated part of the protein could be obtained since those residues were not resolved in the NMR spectra. The three residues belonging to helix 4 that could be resolved also appear to be highly flexible in the aggregate.

A linear correlation between $J(\omega_{N,H})$ and $J(0)$ values for individual NH bonds has been proposed by Lefèvre et al. (*51*). This is based on the assumption that the spectral density function is a linear combination of a few Lorentzians and these are similar for every residue along the sequence. All $J_i(\omega)$ components for each residue obey the following linear relationship (*51*):

$$J_i(\omega_{N,H}) = \alpha J_i(0) + \beta \qquad (4)$$

The linearity, however, is often very poor, which is possibly due to contributions from chemical/conformational exchange which corrupt the $J(0)$ values. Nevertheless, an analysis with

the linearity assumption would still help in identifying a few of the Lorentzian contributions and the corresponding correlation times. Such an analysis for the A form of barstar is shown in Figure 10A,B. The values of α and β obtained from a linear least-squares fit of $J(\omega_N)$ vs $J(0)$ were -0.0191 and 0.1949 ns/rad, respectively, and those obtained from a similar fit of $J(\omega_H)$ vs $J(0)$ were -0.0054 and 0.0335 ns/rad, respectively. These values were then put in the following cubic equation in τ to calculate the time constants characterizing various motions of the protein (*51*):

$$2\alpha\omega_N^2\tau^3 + 5\beta\omega_N^2\tau^2 + 2(\alpha - 1)\tau + 5\beta = 0 \qquad (5)$$

The equation was solved analytically using the software Mathematica (*73*), which yielded one real root and two complex roots. The value of τ obtained is 0.49 ns, which may be assigned to the high-frequency internal motions (τ_i) of the protein backbone. The same equation when solved for ω_H yielded three roots: 0.09, 0.80, and 14.73 ns. To avoid the influence of conformational exchange in residues Ser12 and Ser69 on these derived parameters, the above analysis was repeated by excluding these residues from the data. This yielded a value of 0.53 ns using the $J(\omega_N)$ vs $J(0)$ linear correlation parameters and values of 0.09, 0.79, and 13.28 ns using the $J(\omega_H)$ vs $J(0)$ linear correlation parameters. Among these, the 0.09 ns and 0.79 ns values must be

interpreted to reflect internal motions (τ_i) of the free N-terminal segments and the flexible region in helix 4 of the aggregated monomers of barstar in its A form. The 0.79 ns value of τ_i is similar to the 0.6 ns value calculated by a similar analysis for folded monomeric barstar (*69*). The third 13−15 ns correlation time cannot reflect the overall tumbling time of the aggregate, because it is too small for a molecule of mass 160 kDa. The overall molecular tumbling correlation time for native barstar has been determined to be 5.2 ns from the "model-free" analysis (*69*). Thus, for the 16-mer aggregate of barstar formed at low pH the overall rotational correlation time would be expected to be about 16-fold larger. Indeed, a rotational correlation time of 90 ns has been calculated from fluorescence anisotropy decay measurements carried out at pH 3 on a single tryptophan-containing mutant of barstar labeled with the fluorophore 1,5-IAEDANS at the sole cysteine residue in the mutant protein (*37*). Thus, the 13−15 ns correlation time probably describes a free-flight motion of the N-terminal segment of the protein chain anchored to the aggregated core. Here, the overall tumbling time for the aggregate, however, could not be obtained because residues other than Glu68−Val70 in the aggregated part of the A form were not resolved in the NMR spectra.

It may be mentioned that analysis of the dynamics of the A form of barstar was not done using the "model-free" approach (*74*) because such an analysis does not provide meaningful information in partly folded and random coil proteins due to the wide variation in the local correlation times.

Random Coil Structure in the N-terminal Segment of the A form of Barstar at pH 2.7. The segment consisting of the first 20 residues in each monomeric subunit in the A form appears to be hanging out in solution from the rest of the protein which is aggregated. The secondary $^{13}C^\alpha$ and ^{13}CO chemical shifts of these residues indicate that they have backbone conformational propensity toward an extended structure, β-strand, or random coil. Although less sensitive to conformational preferences than secondary chemical shifts in disordered protein structures, the $^3J_{HN,H\alpha}$ coupling constants of most of these residues have values expected for a random coil structure. Moreover, the existence of $d_{\alpha N}(i,i+1)$ sequential NOE connectivities between most of the residues from Lys1 to Leu20 provides a strong indication of preference toward dihedral angles in the β-region. The occurrence, however, of $d_{NN}(i,i+1)$ sequential NOE connectivities among a few residues in this region suggests that they also sample the α-region of (ϕ, ψ) space. From all these data, it appears that the flexible N-terminal region in the A form is in an essentially random coil conformation. The secondary shifts and presence of only $d_{\alpha N}(i,i+1)$ NOEs in the Lys1 to Ile5 segment, which constitutes the N-terminal region of the first β-strand in native barstar (*36*), show evidence of nativelike structural propensity. The pattern of medium-range NOE connectivities, some of them involving side chain atoms, in the Asp15−Gln18 stretch, and in the Gly7−Ser12 stretch, may be indicative of transient local structuring in this segment. Conformational dynamism in this region is also supported by relatively higher R_2 values for the residues Gln9, Gln18, and Thr19. In addition, the lowered temperature coefficients of NH chemical shifts for a few residues in the N-terminal segment (Lys1, Glu8, Gln9, Arg11, Leu16, His17, Gln18) suggest some transient hydrogen bonding.

Possible Structural Arrangement of the Aggregated A form of Barstar. The peak intensities of residues Lys2−Gln9 in the ^1H-^{15}N HSQC spectrum of native barstar (Lys1 could not be assigned in the native protein spectrum) and in the ^1H-^{15}N HSQC spectrum of the A form are comparable, within experimental error, at equivalent monomer concentrations. This suggests that these N-terminal residues in all 16 subunits in the aggregated state are completely free in solution, and all of them contribute to the peak intensities. The other assigned residues in the N-terminal, however, show lower than the expected peak intensities in the A form spectrum, indicating that the Ile10−Leu20 segment probably fluctuates slowly, on the NMR chemical shifts time scale, between being part of the A form aggregated core and hanging free in solution. Thus, these residues contribute to the magnetization only when they are free in solution, leading to their lower peak intensities. It may be mentioned that this is not in contradiction with high-frequency dynamic motions of the NH bonds, and the free-flight motion of the N-terminal chain inferred from the relaxation data.

From the data obtained here, no definitive information can be derived about how the 16 monomers are arranged in the aggregated A form of barstar, except that it appears to be a symmetrical arrangement so that the flexible N-terminal segments of all monomers are equivalent. The hydrodynamic radius of this 160 kDa aggregate is about 8 nm and that of the monomer is 2 nm. The molecular weight determination of the barstar A form by sedimentation analysis has been done assuming a sphere (*35*), and the DLS measurements also assume the protein to be spherical. It may, therefore, not be unreasonable to speculate a reverse micelle-like structure for the A form of barstar, with a hydrophobic aggregated core and the N-terminal segments lying free in solution. Micelle-like aggregates of the amyloid β-peptide, Aβ(1−40), have been observed to form above a critical concentration of monomers under acidic conditions, and appear to be centers of fibril nucleation (*75*). In a recent small angle neutron scattering study, these micellar intermediates have been proposed to have a spherocylinderal structure with Aβ monomers stretched along the cylinder axis (*76*).

Alternatively, the A form of barstar may be formed by two rings of eight monomers each, arranged one on top of the other forming an axially symmetric cylinder, such that the height of the cylinder is similar to its diameter. The N-terminal segments of the monomers in either ring of monomeric subunits must hang out from opposite ends forming an overall symmetric arrangement of the aggregated A form. The N-terminal segments may lie in the plane of this toroid-like arrangement of monomers or in a plane perpendicular to it.

CONCLUDING REMARKS

The work presented here is a useful first step toward understanding the process of aggregation of barstar at low pH, which proceeds via a rapidly formed soluble aggregate, the A form, that is stable for two weeks. The aggregation process continues over a period of several weeks, through larger, intermediate aggregates, and ultimately terminates in the formation of an insoluble aggregate. Aggregation to the A form is most likely due to intermolecular hydrophobic

interactions (*34*). There may, however, be positively charged residues on the protein surface because of which the segment of barstar comprising β-strand 1, helix 1, and the loop connecting the two is pushed out from the rest of the aggregate due to electrostatic repulsion. The region of the protein that appears to be hanging free in solution from these NMR experiments, and helices 3 and 4 that must be part of the aggregated structure, have Lys, Arg, Asn, and Gln residues that will be protonated and thus positively charged at pH 2.7.

More sensitive NMR experiments utilizing the TROSY and CRINEPT pulse sequences (*53*) on a protein sample triple-labeled with deuterium, ^{15}N and ^{13}C may be able to resolve resonances in the aggregated core of the A form of barstar and provide more information about the structure of the aggregate. Earlier circular dichroism experiments have shown that the A form has 60% of the native helical signal (*34*). Thus, the monomeric subunits in the aggregate must still retain some native or non-native secondary structure. Secondary chemical shifts of the three residues identified in helix 4 in native barstar, in this study, indicate their propensity toward an extended conformation in the A form. Recent ATR-FTIR studies on barstar at low pH show a shift of the peak corresponding to β-structure to a lower wavenumber as compared to the β-peak in the native state, indicating the formation of stronger inter-β-strand hydrogen bonds in the oligomeric A form (Khurana, R. et al., unpublished results).

In most proteins for which structures have been determined in the monomeric and associated states that involve domain swapping, the monomer fold is observed to be substantially conserved (*77*). The formation of amyloid fibrils by some proteins, may, however, involve more extensive structural rearrangements such as conversion from essentially α-helical to β-sheet structure (*15*). The conversion of cellular prion protein (PrP^C) to the disease-specific scrapie form (PrP^{Sc}) also involves a shift from a predominantly α-helical monomeric protein to an oligomeric β-sheet structure (*78*). On the other hand, nativelike structure appears to be stabilized upon aggregation under acidic denaturing conditions for staphylococcal nuclease (*19*, *79*), cold shock protein A (*43*), and the N-terminal anticodon-binding domain of lysyl tRNA synthetase (*80*).

ACKNOWLEDGMENT

REFERENCES

1. Jaenicke, R., and Seckler, R. (1997) *Adv. Prot. Chem. 50*, 1−59.
2. Kiefhaber, T., Rudolph, R., Kohler, H. H., and Buchner, J. (1991) *Biotechnology (NY) 9*, 825−829.
3. Fink, A. L. (1998) *Folding Des. 3*, 9−15.
4. Wetzel, R. (1996) *Cell 86*, 699−702.
5. Finke, J. M., Gross, L. A., Ho, H. M., Sept, D., Zimm, B. H., and Jennings, P. A. (2000a) *Biochemistry 39*, 15633−15642.
6. Ganesh, C., Zaidi, F. N., Udgaonkar, J. B., and Raghavan, V. (2001) *Protein Sci. 10*, 1635−1644.
7. Silow, M., and Oliveberg, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 6084−6086.
8. Pecorari, F., Minard, P., Desmadril, M., and Yon, J. M. (1996) *J. Biol. Chem. 271*, 5270−5276.
9. Wetzel, R. (1992) in *Stability of Protein Pharmaceuticals, Part B.; In Vivo Pathways of Degradation and Strategies for Protein Stabilization*, Vol 3, Plenum Press, New York.
10. Sipe, J. D. (1992) *Annu. Rev. Biochem. 61*, 947−975.
11. Haase-Pattingell, C. A., and King, J. (1988) *J. Biol. Chem. 263*, 4977−4983.
12. Speed, M. A., Wang, D. I. C., and King, J. (1996) *Nat. Biotechnol. 14*, 1283−1287.
13. Mitraki, A., and King, J. (1992) *FEBS Lett. 307*, 20−25.
14. Chrunyk, B. A., and Wetzel, R. (1993) *Protein Eng. 6*, 733−738.
15. Kelly, J. W. (1996) *Curr. Opin. Struct. Biol. 6*, 11−17.
16. Booth, D. R., Sunde, M., Bellotti, V., Robinson, C. V., Hutchinson, W. L., Fraser, P. E., Hawkins, P. N., Dobson, C. M., Radford, S. E., Blake, C. C. F., and Pepys, M. B. (1997) *Nature 385*, 787−793.
17. Hurle, M. R., Helms, L. R., Li, L., Chan, W., and Wetzel, R. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5446−5450.
18. Helms, L. R., and Wetzel, R. (1996) *J. Mol. Biol. 257*, 77−86.
19. Uversky, V. N., Karnoup, A. S., Khurana, R., Segel, D. J., Doniach, S., and Fink, A. L. (1999) *Protein Sci. 8*, 161−173.
20. Quintas, A., Vaz, D. C., Cardoso, I., Saraiva, M. J. M., and Brito, R. M. M. (2001) *J. Biol. Chem. 276*, 27207−27213.
21. Khurana, R., Gillespie, J. R., Talapatra, A., Minert, L. J., Zanetti, C. I., Millett, I., and Fink, A. L. (2001) *Biochemistry 40*, 3525−3535.
22. Wood, S. J., Maleeff, B., Hart, T., and Wetzel, R. (1996) *J. Mol. Biol. 256*, 870−877.
23. Zurdo, J., Guijarro, J. I., Jimenez, J. L., Saibil, H. R., and Dobson, C. M. (2001) *J. Mol. Biol. 311*, 325−340.
24. Bowden, G. A., Paredes, A. M., and Georgiou, G. (1991) *Biotechnology (NY) 9*, 725−730.
25. Sunde, M., and Blake, C. (1997) *Adv. Prot. Chem. 50*, 123−159.
26. DeFelippis, M. R., Alter, L. A., Pekar, A. H., Havel, H. A., and Brems, D. N. (1993) *Biochemistry 32*, 1555−1562.
27. Speed, M. A., Wang, D. I. C., and King, J. (1995) *Protein Sci. 4*, 900−908.
28. Rajan, R. S., Illing, M. E., Bence, N. F., and Kopito, R. R. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 13060−13065.
29. Kopito, R. R. (2000) *Trends Cell Biol. 10*, 524−530.
30. Fink, A. L. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 495−522.
31. Lai, Z., Colon, W., and Kelly, J. W. (1996) *Biochemistry 35*, 6470−6482.
32. Jackson, G. S., Hosszu, L. L. P., Power, A., Hill, A. F., Kenney, J., Saibil, H., Craven, C. J., Waltho, J. P., Clarke, A. R., and Collinge, J. (1999) *Science 283*, 1935−1937.
33. Morillas, M., Vanik, D. L., and Surewicz, W. K. (2001) *Biochemistry 40*, 6982−6987.
34. Khurana, R., and Udgaonkar, J. B. (1994) *Biochemistry 33*, 106−115.
35. Khurana, R. (1995) Ph.D. Thesis: *pH Dependence of the Conformation of the Small Protein Barstar*, Mumbai University, Mumbai.
36. Lubienski, M. J., Bycroft, M., Freund, S. M. V., and Fersht, A. R. (1994) *Biochemistry 33*, 8866−8877.
37. Lakshmikanth, G. S. (2001) Ph.D. Thesis: *Protein Dynamics and Folding*, Mumbai University, Mumbai.
38. Fraser, P. E., Nguyen, J. T., Surewicz, W. K., and Kirschner, D. A. (1991) *Biophys. J. 60*, 1190−1201.
39. Inouye, H., Fraser, P. E., and Kirschner, D. A. (1993) *Biophys. J. 64*, 502−519.
40. Blake, C., and Serpell, L. (1996) *Structure 4*, 989−998.
41. Serpell, L. C., Sunde, M., Fraser, P. E., Luther, P. K., Morris, E. P., Sangren, O., Lundgren, E., and Blake, C. C. (1995) *J. Mol. Biol. 254*, 113−118.
42. Lansbury, P. T., Jr., Costa, P. R., Griffiths, J. M., Simon, E. J., Auger, M., Halverson, K. J., Kocisko, D. A., Hendsch, Z. S., Ashburn, T. T., Spencer, R. G. S., Tidor, B., and Griffin, R. G. (1995) *Nat. Struct. Biol. 2*, 990−998.
43. Alexandrescu, A. T., and Rathgeb-Szabo, K. (1999) *J. Mol. Biol. 291*, 1191−1206.
44. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) *Biochemistry 40*, 3561−3571.
45. Eliezer, D., Chung, J., Dyson, H. J., and Wright, P. E. (2000) *Biochemistry 39*, 2894−2901.
46. Bhavesh, N. S., Panchal, S. C., and Hosur, R. V. (2001) *Biochemistry 40*, 14727−14735.
47. Kay, L. E., and Gardner, K. H. (1997) *Curr. Opin. Struct. Biol. 7*, 722−731.

48. Clore, G. M., and Gronenborn, A. M. (1998) *Curr. Opin. Chem. Biol. 2*, 564−570.

49. Panchal, S. C., Bhavesh, N. S., and Hosur, R. V. (2001) *J. Biomol. NMR 20*, 135−147.

50. 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) *Biochemistry 33*, 5984−6003.

51. Lefèvre, J.-F., Dayie, K. T., Peng, J. W., and Wagner, G. (1996) *Biochemistry 35*, 2674−2686.

52. Peng, J. W., and Wagner, G. (1994) *Methods Enzymol. 239*, 563−596.

53. Wider G., and Wüthrich, K. (1999) *Curr. Opin. Struct. Biol. 9*, 594−601.

54. Wiltscheck, R., Kammerer, R. A., Dames, S. A., Schulthess, T., Blommers, M. J. J., Engel, J., and Alexandrescu, A. T. (1997) *Protein Sci. 6*, 1734−1745.

55. Dames, S. A., Kammerer, R. A., Wiltscheck, R., Engel, J., and Alexandrescu, A. T. (1998) *Nat. Struct. Biol. 5*, 687−691.

56. Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.

57. Schwarzinger, S., Kroon, G. J. A., Foss, T. R., Wright, P. E., and Dyson, H. J. (2000) *J. Biomol. NMR 18*, 43−48.

58. Schwarzinger, S., Kroon, G. J. A., Foss, T. R., Chung, J., Wright, P. E., and Dyson, H. J. (2001) *J. Am. Chem. Soc. 123*, 2970−2978.

59. Brant, D. A., Miller, W. G., and Flory, P. J. (1967) *J. Mol. Biol. 23*, 47−65.

60. Karplus, M. (1959) *J. Chem. Phys. 30*, 11−15.

61. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996) in *Protein NMR Spectroscopy Principles and Practice*, Academic Press, San Diego, CL.

62. Penkett, C. J., Redfield, C., Dodd, I., Hubbard, J., McBay, D. L., Mossakowska, D. E., Smith, R. A. G., Dobson, C. M., and Smith, L. J. (1997) *J. Mol. Biol. 274*, 152−159.

63. Dyson, H. J., and Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519−538.

64. Rose, G. D., Gierasch, L. M., and Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1−106.

65. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) *J. Biomol. NMR 5*, 14−24.

66. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.

67. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry 28*, 8972−8979.

68. Kay, L. E. (1998) *Nat. Struct. Biol. 5 (NMR Suppl.)*, 513−517.

69. Sahu, S. C., Bhuyan, A. K., Majumdar, A., and Udgaonkar, J. B. (2000) *Proteins: Struct., Funct., Genet. 41*, 460−474.

70. Slichter, C. P. (1990) in *Principles of Magnetic Resonance*, Springer-Verlag, New York.

71. Bai, Y., Chung, J., Dyson, J. H., and Wright, P. E. (2001) *Protein Sci. 10*, 1056−1066.

72. Meekhof, A. E., and Freund, S. M. V. (1999) *J. Mol. Biol. 286*, 579−592.

73. Wolfram, S. (1999) *The Mathematica Book*, 4th ed., Wolfram Media/ Cambridge University Press, Cambridge U. K.

74. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.

75. Lomakin, A., Chung, D. S., Benedek, G. B., Kirschner, D. A., and Teplow, D. B. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 1125−1129.

76. Yong, W., Lomakin, A., Kirkitadze, M. D., Teplow, D. B., Chen, S. H., Benedek, G. B. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 150−154.

77. Schlunegger, M. P., Bennet, M. J., and Eisenberg, D. (1997) *Adv. Prot. Chem. 50*, 61−122.

78. Prusiner, S. B., and DeArmond, S. J. (1995) *Amyloid 2*, 39−65.

79. Uversky, V. N., Segel, D. J., Doniach, S., and Fink, A. L. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 5480−5483.

80. Alexandrescu, A. T., Lamour, F. P., and Jarvine, V. A. (2000) *J. Mol. Biol. 295*, 239−255.

81. Lubienski, M. J., Bycroft, M., Jones, D. N. M., and Fersht, A. R. (1993) *FEBS Lett. 332*, 81−87.