Transient Structure of the Amyloid Precursor Protein Cytoplasmic Tail Indicates Preordering of Structure for Binding to Cytosolic Factors†

Theresa A. Ramelot, Lisa N. Gentile,‡ and Linda K. Nicholson*

Department of Molecular Biology and Genetics, 239 Biotechnology Building, Cornell University, Ithaca, New York 14853

Received November 8, 1999

ABSTRACT: The cytoplasmic tail of the amyloid precursor protein (APP) appears to play two important roles in the cell through participation in intracellular signaling and proteolytic processing of APP. Hence, knowledge of the structure of the 47 residue cytoplasmic tail of APP is important for understanding the molecular interactions involved in normal cell function as well as in the pathogenesis of Alzheimer's disease. Multidimensional solution NMR spectroscopy has been applied to examine the structural features of a 49-residue peptide (APP-C) containing two N-terminal residues (GS) and the APP cytoplasmic tail, over the pH range of 4.2−7.1. Although the peptide does not adopt a stable folded structure, regions of unstable structure exist over the pH range examined and have been characterized by a combination of H^{α} chemical shifts, NOE analysis, and $^{3}J_{HNH\alpha}$ coupling constants and by identification of transient hydrogen bonds between amide protons and titrating carboxylate groups. These studies extend the work of others [Kroenke et al. (1997) *Biochemistry 36*, 8145−8152] by identifying an additional nascent helix and a hydrophobic cluster within the N-terminal 20 amino acid residues and by further characterizing the TPEE turn as a helix capping box. The transient structure of APP-C provides insight into the importance of preordering of this cytoplasmic tail in governing specificity and affinity for cytosolic binding partners.

The amyloid precursor protein (APP)[1] is an integral transmembrane glycoprotein that is the precursor to the 39−43 residue amyloid β peptide (AβP), the major component of amyloid plaques that characterize Alzheimer's disease. The predominant APP isoform in the brain, APP_{695}, is composed of a large glycosylated extracellular component, a single membrane-spanning region, and a short 47 residue cytoplasmic tail that is highly conserved (Figure 1A). The cytoplasmic tail participates in the important cellular processes of intracellular trafficking of APP and signal transduction via interactions with adaptor and signaling proteins, respectively. Cleavage of APP to yield amyloidogenic AβP involves a complicated metabolic pathway that is mediated by intracellular targeting sequences located in the cytoplasmic tail. Processing of APP takes place in different intracellular compartments and can result in either cleavage within the AβP sequence, precluding AβP formation, or cleavage that yields C-terminal fragments containing the AβP sequence (for a review, see ref *1*). Subsequent cleavage of these C-terminal fragments results in release of the AβP. Knowledge of the intermolecular interactions that govern the trafficking and processing of APP, the production and secretion of AβP, and amyloid plaque formation is essential for understanding the pathogenesis of Alzheimer's disease.

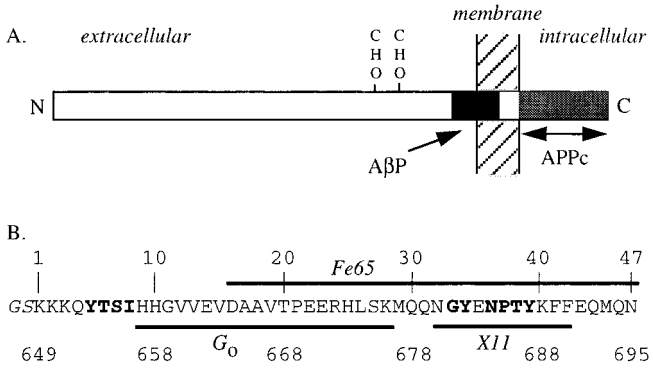

FIGURE 1: (A) Schematic diagram of APP_{695}. Main features of the 695 amino acid residue protein include a large extracellular region (white box), glycosylation sites (CHO), the amyloid β-peptide (AβP, black box), and the 47-residue cytoplasmic tail (APPc, gray box). (B) Amino acid sequence of the 49-residue APP-C peptide containing the 47-residue APPc and two nonnative N-terminal residues (italic type) remaining from the thrombin cut site. Intracellular targeting sequences are highlighted in boldface type. The G_o, Fe65, and X11 binding regions are indicated by heavy lines. Numbering schemes are indicated relative to either the APP-C peptide (top) or the full-length APP_{695} isoform (bottom).

† This work was supported by funds provided by National Science Foundation Grants MCB-9507144 (L.K.N.) and BIR-9512501 (L.K.N.) and by a National Institutes of Health Biophysics Training Grant (T.A.R.).

* To whom correspondence should be addressed: Phone (607) 255-7208; Fax (607) 255-2428; email lkn2@cornell.edu.

‡ Present address: Department of Chemistry and Physics, University of New England, 11 Hills Beach Rd., Biddeford, ME 04005.

[1] Abbreviations: APP, amyloid precursor protein; APP-C, the 49-residue peptide containing two N-terminal residues (GS) and the cytoplasmic tail of the amyloid precursor protein; APPc, the 47-residue cytoplasmic tail of the amyloid precursor protein; AβP, the 39−43 residue amyloid β peptide; APP_{695}, the 695-residue predominant amyloid precursor protein isoform in the brain; FAD, familial Alzheimer's disease; pY, phosphotyrosine; NMR, nuclear magnetic resonance spectroscopy; PID, phosphotyrosine interaction domain; 2D, two-dimensional; 3D, three-dimensional; HSQC, heteronuclear single quantum spectroscopy; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; HMQC-*J*, heteronuclear multiple quantum *J*-correlation; HNHA, three-dimensional ^{15}N-separated $H^{N}-H^{\alpha}$ homonuclear *J*-correlation experiment; FID, free induction decay; RMSD, root-mean-square deviation.

10.1021/bi992580m CCC: $19.00
Published on Web 02/17/2000

The cytoplasmic tail of APP (APPc) contains important signal sequences for intracellular trafficking, which include signals sufficient for endocytosis, for recycling to the plasma membrane, and for sorting to the basolateral surface in polarized cells (*2*, *3*). Deletion of APP residues 682−687, which contains the widely recognized endocytosis signal NPXY (Figure 1B), was shown to prevent endocytosis of APP (*2*) and to alter AβP production (*4*, *5*). Adjacent to this signal is the conserved GY motif (Figure 1B), present in many lysosomal glycoproteins that are endocytosed and targeted to the lysosomes (*6*, *7*). An additional sequence, YTSI (APP residues 653−656, Figure 1B), mediated endocytosis when substituted into the cytoplasmic tail of the transferrin receptor (*2*). Furthermore, Y_{653} is required for sorting of APP to the basolateral membrane surface in polarized MDCK cells (*8*). Hence, APPc contains two signal sequences that govern the trafficking and processing of this important protein.

Additional regions of APPc are important for interactions with intracellular factors. The overall organization of full-length APP displays features similar to cell surface receptors (*9*). Apoptosis associated with DNA fragmentation is observed in neuronal cell lines with V_{642} familial Alzheimer's disease (FAD) mutants in an AβP-independent process (*10*). This process is believed to be mediated by the 20 amino acid residue G_o binding region of APPc (APP residues 657−676, Figure 1B) (*11*, *12*). The abnormal function of the APP FAD mutants appears to result from the constitutive activation of G_o (*13*). Other cytosolic factors that are known to interact with the cytoplasmic tail may also be involved in signal transduction events. A recently discovered APP binding protein, APP-BP1, may act as an effector molecule in transducing APP signals into the cell (*14*). Other neuronal APP binding proteins, X11, FE65, and the FE65-like protein, bind in vitro and in vivo through a phosphotyrosine interaction domain (PID) (*15*−*17*). However, unlike the well-characterized homologous PID domain of Shc, FE65 and X11 binding to target sequences is independent of Tyr phosphorylation (*15*). Furthermore, the X-ray crystal structure of X11 bound to a peptide derived from APPc reveals an interaction region spanning residues 680−690 (Figure 1B) (*18*), and biochemical studies have shown that more than residues 675−695 are required for interaction with FE65 (Figure 1B) (*16*). Hence, the structure of the entire APPc region is of great interest for understanding interactions between this cytoplasmic tail and intracellular signal transduction machinery.

A single NMR structural study at pH 6.3 and 4 °C has been reported for a 49-residue peptide (APP-C) containing two N-terminal residues (GS) fused to APPc, which characterized two type I β-turns for the sequences TPEE and NPTY and found nascent helical character primarily in the C-terminal half and random coil conformations for the N-terminal 20 amino acid residues (*19*). We report here additional NMR studies of APP-C performed at 25 °C and at a range of pH values that reveal transient hydrogen bonds and considerable structure in the N-terminal half as well. The H^{α} chemical shifts were used to characterize the structure of this peptide at three different pH values: 4.2, 5.6, and 7.1. In addition, pH titration in the range of 2.0−6.1 allowed identification of transient hydrogen bonds between amide protons and the ionizable side chains of the Glu and Asp residues in APP-C. Saturation transfer is used to indicate the level of protection of amide protons from exchange with solvent. A detailed analysis of NOE cross-peaks and measurement of $^3J_{HNH\alpha}$ coupling constants provided structural constraints that, along with the hydrogen bond information, were used to further characterize the transient structure of APP-C in aqueous solution. Comparison of this solution structure with the X11-bound form (*18*) demonstrates how transiently structured regions in APP-C are either stabilized or rearranged upon binding. The unbound structure is of great importance for understanding the biophysical basis for transient interactions between the APP cytoplasmic tail and intracellular protein machinery that facilitate intracellular trafficking and signal transduction.

MATERIALS AND METHODS

Plasmid Construction. DNA containing the carboxy-terminal 47 residues of APP was PCR-amplified from clone λAm4 (American Type Culture Collection No. 40305). The 5′ primer contained a *Bam*HI cut site along with DNA encoding an amino terminal 6×His tag, SGSG spacer, thrombin cut site, and the first 6 residues of APPc. The 3′ primer contained the complementary DNA for the carboxy-terminal 20 amino acids of APPc, a stop codon, and a *Sal*I cut site. To convert residue 34 from serine to tyrosine, this mutation was included in the 3′ primer, along with the complementary DNA for 6 residues upstream from this mutation. This was necessary to convert the purchased sequence (GenBank accession no. M15533) (*20*) to that obtained by other researchers for human APP (*9*, *21*).

The PCR-generated DNA was first cloned into the pCRII vector (Invitrogen) for amplification and purification and verified by DNA sequencing. After restriction enzyme digestion, the desired fragment was gel-purified and ligated into the codigested pTrxFus vector from the ThioFusion expression system (Invitrogen), resulting in the thioredoxin-APP-C fusion protein expression vector pTrxAPPc47.

Protein Expression and Purification. The thioredoxin fusion protein was expressed in *Escherichia coli* strain GI724 cells. Cells were grown at 30 °C in an air bath incubator/shaker in 1 L of M9 minimal medium containing ampicillin (100 μg/mL) in a 2.8 L Fernbach flask. Expression was induced when cells reached an OD_{600nm} of 0.30−0.35 by the addition of 100 mg of L-tryptophan. The cells were then grown for 4 h at 37 °C to an OD_{600nm} of 0.7−0.9. The cells were harvested, washed with PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4), and lysed in a French pressure cell in sonication buffer (50 mM Na_2HPO_4 and 300 mM NaCl, pH 8.0) containing 10 μg/mL each pepstatin A, aprotinin, and leupeptin. The lysate was centrifuged at 16000*g* for 20 min, and the supernatant was applied to an 8 mL Ni−NTA agarose column (Qiagen) equilibrated with the same buffer. The column was washed with 10 column volumes each of sonication buffer, wash buffer (50 mM Na_2HPO_4, 300 mM NaCl, and 10% glycerol, pH 6.0), 5 mM imidazole in wash buffer, and again with sonication buffer, leaving 8 mL above the column.

The desired peptide was cleaved from the thioredoxin fusion protein by addition of 2 units of human plasma thrombin (Boehringer Mannheim) to the Ni−NTA column while shaking for 10 h at room temperature, yielding two

fragments: a 49 residue peptide (APP-C) containing the 47 residue APP cytoplasmic tail and two additional amino terminal residues (GS) that remain from the thrombin cut site, and the 6×His tagged thioredoxin fragment. APP-C was then eluted with sonication buffer adjusted to pH 6 and was further purified on a preparative HPLC Novapac C-18 column, 2.5 × 10 cm (Waters Inc.) using an elution gradient of 0−50% acetonitrile with 0.1% trifluoroacetic acid in water after first adjusting the pH of the peptide solution to 3.5 and filtering. Purified APP-C was lyophilized to dryness, yielding 3−4 mg of peptide (>99% pure as judged by analytical HPLC), and was further characterized by mass spectroscopy and amino acid analysis, verifying the molecular weight (5.7 kDa) and the amino acid composition of the desired product. Protein enriched with ^{15}N was obtained as described above by substituting $^{15}NH_4Cl$ in the M9 media.

NMR Sample Preparation and Experiments. NMR samples were prepared from lyophilized APP-C powder by dissolving in either 50 mM Sorenson's phosphate buffer at pH 5.6 (63 mM KH_2PO_4 and 4 mM Na_2HPO_4) or 50 mM potassium phosphate buffer with 10% D_2O and 5 mM sodium azide. The pH of the samples was adjusted by adding small amounts of 0.1 N NaOH or 0.1 N HCl as needed. Protein concentrations ranged from 1 to 3 mM APP-C.

All NMR spectra were recorded at 25 °C on a Varian Inova 600 MHz spectrometer using the States−Haberkorn hypercomplex method of frequency discrimination (*22*). Two-dimensional $^1H-^{15}N$ HSQC spectra were recorded with spectra widths of 2 kHz in *t*1 (512 complex data points) and 8 kHz in *t*2 (1024 complex data points) and processed with zero-filling to a final data size of 2048 by 1024 data points. Two-dimensional total correlation spectroscopy (TOCSY) (*23*) and nuclear Overhauser enhancement spectroscopy (NOESY) (*24*) spectra were recorded with spectral widths of 8 kHz in *t*2 (2048 complex data points) and *t*1 (1024 complex data points), unless noted otherwise. Two-dimensional TOCSY spectra were acquired with a DIPSI-2 (*25*) isotropic mixing period of 50 or 70 ms and processed with a single zero-fill to a final data size of 2048 by 1024 data points. Two-dimensional NOESY spectra were processed with zero-filling to a final data size of 2048 by 2048 data points. Two-dimensional TOCSY spectra of ^{15}N-labeled APP-C were recorded for the pH titration study with spectral widths of 8 kHz (1024 complex data points) in *t*2 and 5.1 kHz (668 complex data points) in *t*1 with a DIPSI-2 mixing time of 70 ms, with a 180° ^{15}N pulse centered on the *t*1 evolution period to refocus $^1J_{HN}$ coupling and with WALTZ-16 ^{15}N decoupling during *t*2. These spectra were processed with zero-filling to a final data size of 2048 by 2048 data points. Three-dimensional ^{15}N-TOCSY−HSQC and ^{15}N-NOESY−HSQC spectra were recorded with spectral widths of 5.7 kHz (440 complex data points) in *t*1, 1.1 kHz (62 complex data points) in *t*2, and 6 kHz (1024 complex data points) in *t*3. Data were processed to a final data size of 2048 by 1024 by 128, using linear prediction to double the data size in the *t*2 dimension (^{15}N) and by zero-filling in all dimensions to the final data size.

The $^3J_{HNH\alpha}$ coupling constants were determined for ^{15}N labeled APP-C at pH 6.0 and 25 °C using both the 2D $^1H-^{15}N$ HMQC-*J* (*26*) and 3D ^{15}N-separated H^N-H^α homonuclear *J*-correlation experiment, HNHA (*27*) experiments. For the 2D $^1H-^{15}N$ HMQC-*J* experiment, spectral widths were 8 kHz in *t*2 (1024 complex data points) and 1.5 kHz in *t*1 (512 complex data points), and the data were processed with zero-filling to a final data size of 2048 by 2048 data points. A Lorentz-to-Gaussian window function was applied to the FIDs in both dimensions prior to zero-filling and Fourier transformation to remove 12 and 5 Hz (*t*2 and *t*1, respectively) Lorentzian line width (based on the non-apodized line widths) and to apply a Gaussian apodization function in each dimension using the same line widths. Peak splitting in the ^{15}N dimension was determined by deconvolution of the Gaussian line shapes using the nonlinear least-squares spectral line shape modeling program, nlinLS (Frank Delaglio, NIH/NIDDK). The values obtained in this manner agreed well with those determined from the 3D HNHA experiment by use of the diagonal to cross-peak intensity ratio with a 12.7 ms dephasing and rephasing delay (*27*). For the 3D experiment, data were collected and processed as described for ^{15}N-separated TOCSY−HSQC spectra.

Saturation transfer experiments, which measure the rate of hydrogen exchange between H_2O and amide protons, were performed on an ^{15}N-labeled APP-C sample at pH 6.0. Two $^1H-^{15}N$ HSQC spectra, each with a presaturation pulse either on- or off-resonance with water, were recorded as described above. Peak intensities were determined by the peak-picking routine PIPP, which uses parabolic interpolation in all dimensions. Intensities from the HSQC spectrum recorded with the 3 s presaturation pulse at 66 Hz applied on-resonance with water are denoted by $I(t)$, with $t = 3$. Peaks intensities determined from the HSQC spectrum with the off-resonance 3 s presaturation pulse are denoted by $I(0)$. The value T_1k_{ex} was calculated for the amide proton of each residue by use of eq 1 (*28*), where T_1 is the longitudinal relaxation time constant of the amide proton:

$$k_{ex} = \{[I(0)/I(t)] - 1\}/T_1 \quad (1)$$

Predicted intrinsic exchange rates, k_{int}, were estimated by the method described by Bai et al. (*29*), which takes into account the intrinsic influence on amide proton lability by the neighboring side chains. The amide/H_2O exchange parameters at 293 K from Connelly et al. (*30*) were used in determination of k_{int}: $k_A = 1.39$ min^{-1}, $k_B = 9.95$ min^{-1}, and $k_W = 0$ min^{-1}. Exchange rates at 298 K and pH 6.0 were obtained by correcting for temperature dependence as described in Bai et al. (*29*) for the base-catalyzed reaction ($E_a = 17$ kcal mol^{-1}). The protection of an amide proton from exchange is often described by the protection factor, *P*:

$$P = k_{int}/k_{ex} = T_1k_{int}/\{[I(0)/I(t)] - 1\} \quad (2)$$

An estimate of amide proton T_1 values across the APP-C backbone was obtained by recording HSQC spectra with recycle delay times of $t = 0.6$, 0.7, 0.8, and 1.5 s and fitting the peak intensity build-up curves to

$$y = A[1 - \exp(-t/T_1)] \quad (3)$$

Protection factors for amide protons in close proximity to H^α protons with resonance frequencies close to water may be attenuated by NOE effects resulting from saturation of

the H^{α} resonances by the presaturation pulse. In this case, protection factors for these residues would reflect a lower bound.

The dependence of ^1H chemical shifts on pH was determined by recording a series of TOCSY (t_{mix} = 70 ms) spectra at seven pH values: 2.0, 3.1, 3.5, 4.1, 4.6, 5.1, and 6.1 as described above. Chemical shifts were obtained for all observable amide protons, for Asp H^{β} protons, for Glu H^{γ} protons, and for the H^{α} proton of N_{47}. The dependence of a given proton chemical shift on pH was fit to the single-proton titration curve described by eq 4:

$$\delta(\mathrm{pH}) = \frac{\delta_{HA} + \delta_{A^-}\, 10^{(\mathrm{pH}-\mathrm{p}K_a)}}{1 + 10^{(\mathrm{pH}-\mathrm{p}K_a)}} \quad (4)$$

where δ_{HA} is the chemical shift at the acidic limit and δ_{A^-} is the chemical shift at the basic limit. The pK_a value reflected by a given proton was obtained from the best fit of the experimental data, and the corresponding error was obtained by fitting 100 synthetic data sets in which δ(pH) values are randomly varied over their range of uncertainty (σ = 0.01 ppm) to obtain the standard deviation in pK_a using a Monte Carlo approach.

All NMR data were processed and analyzed using the *nmrPipe* and *nmrDraw* software tools (*31*) and with the program PIPP (*32*). Solvent subtraction was obtained by subtraction of a fourth-order polynomial fit in the time domain. Typically, the free induction decay (FID) was multiplied by a 80° shifted squared sine-bell function (^1H dimensions) or a 90° shifted sine-bell function (^{15}N dimensions) before zero-filling and Fourier transformation. Linear prediction was used to back-predict one point in the *t*1 dimension of the ^{15}N-TOCSY−HSQC experiment prior to Fourier transformation. Linear baseline correction was applied in the acquisition dimension of all spectra.

Molecular Modeling of the VTPEER Capping Box. Solution structures of the $_{19}VTPEER_{24}$ region (numbering relative to APP-C, Figure 1B) were calculated by the hybrid distance geometry/simulated annealing approach (*33*) implemented by the program X-PLOR 3.8 (*34*). The 8-residue fragment $_{18}AVTPEERH_{25}$ was used for these calculations to minimize end effects resulting from the N- and C-termini. By use of the standard X-PLOR protocol, an extended structure was generated by the *generate* and the *generate_template* routines, followed by embedding using the *dg_sub_embed* routine. Structures are generated after application of the *dgsa* routine and refinement by the *refine* routine. The structures were improved in an iterative process with addition of NOEs and resolution of stereochemical proton assignments. Final structures were analyzed by the superposition of the backbone atoms of residues $_{20}TPEE_{23}$ of the turn.

RESULTS

Resonance Assignments. Full proton and ^{15}N resonance assignments were first determined for APP-C at pH 5.6. These assignments were used for a detailed structural analysis as well as for evaluating chemical shift changes as a function of pH. Figure 2 shows the $^1H-^{15}N$ HSQC spectrum of APP-C at pH 5.6, with residue assignments denoted by the numbering scheme relative to the APP-C peptide (Figure 1B). Hereafter, this numbering scheme will be employed. The

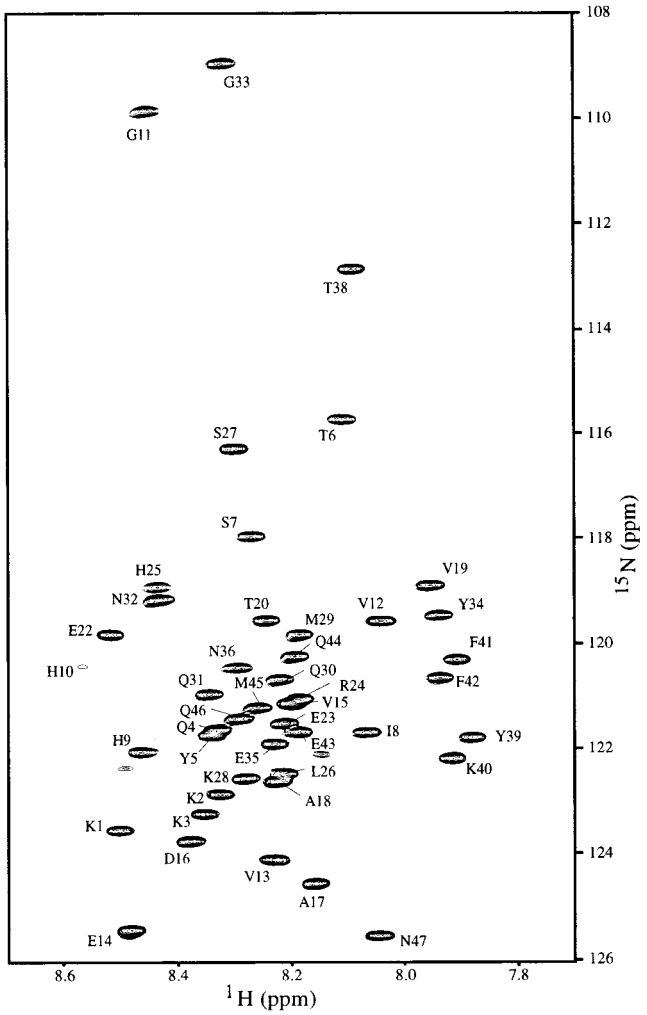

FIGURE 2: $^1H-^{15}N$ HSQC spectrum of APP-C at pH 5.6 and 25 °C. Residue assignments are indicated with numbering relative to APP-C.

narrow range of the amide proton chemical shifts (<1 ppm) is typical of random coil proteins. However, due to the wide dispersion in the ^{15}N dimension only two pairs of peaks, corresponding to residues Q_4/Y_5 and V_{15}/R_{24}, overlap within their line widths at half-height. Peaks are observed for all backbone amide protons from K_1 through N_{47} (44 peaks), while those for the first two N-terminal residues (GS) that remain from the thrombin cut site are not observed due to rapid amide proton exchange. Essentially complete ^1H and ^{15}N assignments at pH 5.6 were obtained for residues K1 through N47 using homonuclear TOCSY (t_{mix} = 70 ms) and NOESY (t_{mix} = 200 ms) spectra along with ^{15}N-TOCSY−HSQC (t_{mix} = 70 ms) and ^{15}N-NOESY−HSQC (t_{mix} = 200 ms) spectra to resolve those spin systems with spectral overlap in the amide proton dimension (Supporting Information).

To examine the effects of pH on the structure of APP-C, full ^1H assignments were also obtained at pH 4.2 and 7.1 (Supporting Information). Assignments were obtained at pH 4.2 from homonuclear TOCSY (t_{mix} = 70 ms) and NOESY (t_{mix} = 200 ms) spectra and at pH 7.1 from a single TOCSY (t_{mix} = 50 ms) spectrum. A comparison of the structure of APP-C at the three pH values (pH 4.2, 5.6, and 7.1) was made by comparing the H^{α} chemical shifts of each of the residues as described below. In addition, the chemical shifts

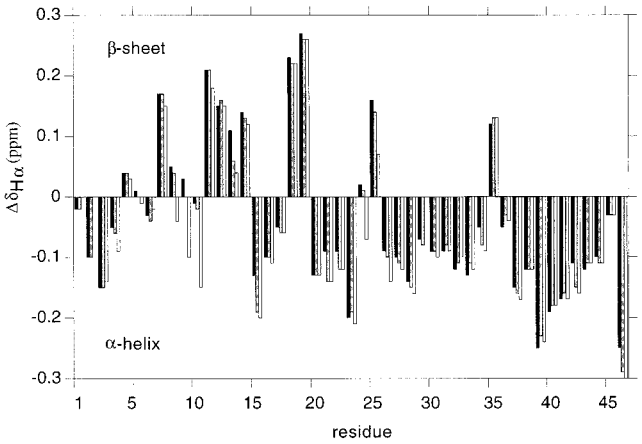

FIGURE 3: Plot of the difference in H^{α} chemical shifts ($\Delta\delta_{H\alpha}$) from random coil values for APP-C at pH 4.2 (black bars), pH 5.6 (gray bars), and pH 7.1 (white bars). A series of $\Delta\delta_{H\alpha}$ values less than -0.1 ppm indicates helical structure, and a series of values greater than 0.1 indicates β-strand. Error in H^{α} chemical shift values is ± 0.01 ppm.

of all amide protons, the H^{β} protons of Asp, H^{γ} protons of Glu residues, and the H^{α} proton of the C-terminal Asn were determined at seven pH values ranging from 2.0 to 7.1 by homonuclear TOCSY spectra ($t_{mix} = 70$ ms). The pH dependence of these chemical shifts was used to identify transient hydrogen bonds between titrating carboxylate groups and amide protons as described in detail below.

pH Dependence of Secondary Structure from H_{α} Chemical Shifts. The difference between an observed H^{α} chemical shift and the corresponding random coil value for a particular type of amino acid residue yields an accurate measure of the secondary structure of peptides and proteins (*35*). This difference, denoted by $\Delta\delta_{H\alpha}$, was determined for APP-C at three pH values (4.2, 5.6, and 7.1) and the results are plotted in Figure 3. The largest changes in $\Delta\delta_{H\alpha}$ between pH 4.2 and 5.6 are observed for residues that have titrating groups with pK_a values near pH 4. These correspond to the five Glu residues, the single Asp, and the C-terminal Asn, which display changes consistent with the intrinsic effect due to titration. Upon increasing pH from 5.6 to 7.1, large upfield chemical shifts are observed for His residues (H_9, H_{10}, and H_{25}) that are expected to have pK_a values near this range. Upfield shifts were also observed for residues G_{11}, V_{12}, L_{26}, and S_{27}, consistent with intrinsic effects due to titration of the neighboring His residues. Hence, the largest changes in H^{α} chemical shift result from intrinsic effects of titration rather than from conformational changes. For the remaining residues, the similarity of the H^{α} chemical shifts at the different pH values indicates that the average backbone conformation of APP-C is generally invariant to pH. This is also supported by the invariance of methyl proton frequencies with pH, which indicates that the level of solvent exposure of these hydrophobic groups is not altered by titration (Supporting Information). The stability of structural features indicates that, unlike the AβP, APP-C does not have a pH-dependent phase transition in the pH range examined. This justifies the use of pH titration studies to identify transient hydrogen bonds, as discussed below.

As was previously found for APP-C at pH 6.3 and 4 °C (*19*), nascent helix is shown here to be present in this cytoplasmic tail over a broad range of pH values. Nascent helix is defined as a series of turnlike structures in equilibrium with unfolded states (*36*). Helical structure, indicated by a series of $\Delta\delta_{H\alpha}$ values less than -0.1 ppm, is observed for residues 21−24 and 38−45 at pH values of 4.2, 5.6, and 7.1 (Figure 3). In addition, although not all less than -0.1 ppm, sequential negative values are also consistently observed for residues 2−4, 16−18, and 27−35. Values greater than $+0.1$ ppm are observed for residues 12−15 at pH 4.2, indicating β-strand conformation for this short segment. This trend persists at pH 5.6 and 7.1 as indicated by positive values for these residues. Hence, based solely on H^{α} chemical shifts, a short extended segment (residues 12−15) and several helical segments (residues 2−4, 16−18, 21−24, 27−35 and 37−45) are predicted.

β-Turns and Nascent Helix Persist over a Range of pH Values. In agreement with the structure previously characterized at pH 6.3 and 4 °C (*19*), β-turns are present at residues $_{20}$TPEE$_{23}$ and $_{36}$NPTY$_{39}$ over the pH range examined. The TPEE turn contains NOEs consistent with a type I β-turn but also displays features that more fully characterize it as an N-terminal helix capping box structure. In the standard nomenclature for N-terminal helix capping box residues, -N′-Ncap-N1-N2-N3-N4-, the Ncap, N1, N2, and N3 residues share many structural features with a type I β-turn. The capping box structure is distinguished by a hydrogen bond between the Ncap side chain and the backbone of N3 and a reciprocal hydrogen bond between the N3 side chain and the backbone of Ncap (*37*, *38*). A type Ib capping box structure also has a hydrophobic interaction between the side chains of N′ and N4 (*39*). Reciprocal hydrogen bonding in TPEE is supported by NOEs observed between the Ncap side chain (T_{20} H^{γ}) and the N3 backbone (E_{23} H^{N}), between the N3 side chain (E_{23} H^{β}, H^{γ}) and the Ncap backbone (T_{20} H^{N}), and by additional NOEs that further define the Ncap side-chain position (T_{20} H^{γ} to E_{22} H^{N} and to P_{21} H^{δ}s). A N′−N4 hydrophobic interaction is supported by weak NOEs between the N′ and N4 side chains (V_{19} H^{γ} and R_{24} H^{δ}s, respectively) and between the N4 side chain (R_{24} H^{δ}s) and the Ncap residue (T_{20} H^{N}). In addition, the small coupling constant observed (5.8 Hz) for the third residue of this turn, E_{22}, indicates that φ for this residue is more helical than the standard type I β-turn. The helical nature of this region is further supported by observation of weak NOEs between the N4 backbone (R_{24} H^{N}) and the Ncap residue (T_{20} H^{α}, H^{β}, H^{γ}) as well as a $d_{\alpha N}(i, i + 3)$ between R_{24} H^{α} and S_{27} H^{N}, which indicates that helical character is extended beyond the TPEE region. This is also in agreement with the H^{α} chemical shifts observed for residues 21−24. Side-chain NOEs between T_{20} and P_{21} (T_{20} H^{α} to P_{21} H^{δ}s and T_{20} H^{β} to P_{21} H^{δ}s) identify the trans conformation of P_{21}. Taken together, these chemical shift, NOE, and $^3J_{HNH\alpha}$ coupling data demonstrate that this turn is more accurately described as an N-terminal helix capping box (*40*). This conclusion is further supported by identification of transient hydrogen bonds between T_{20} H^{N} and E_{23} COO^{-} and between E_{22} H^{N} and its own COO^{-} by pH titration studies and by molecular modeling as described below.

The sequence NPTY also possesses the pattern of local NOEs that identifies a type I β-turn, including the interresidue side-chain NOEs between N_{36} and P_{37} that characterize the trans Pro. Furthermore, the large coupling constant observed for T_{38} (8.2 Hz) is typical for the third residue in a type I

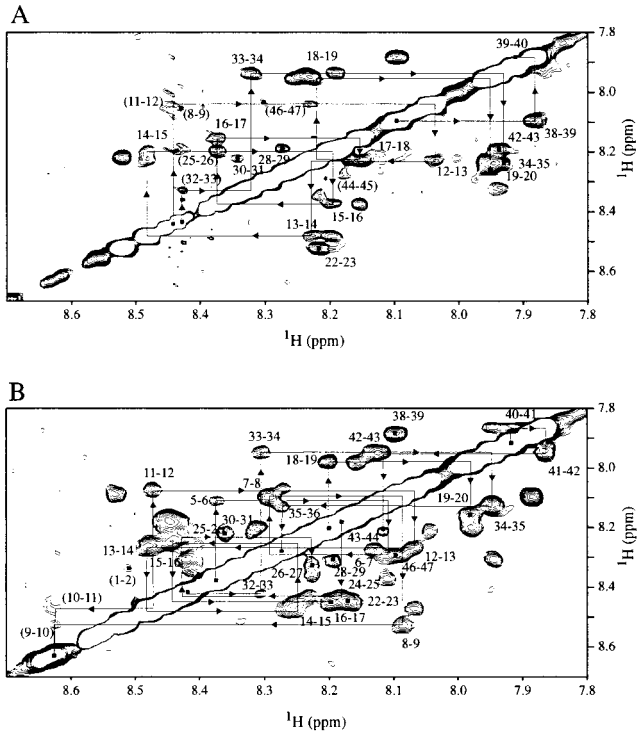

FIGURE 4: H^N-H^N region of the homonuclear NOESY spectrum (t_{mix} = 200 ms) of APP-C at 25 °C and (A) pH 5.6 or (B) pH 4.2. The positive sign of both the diagonal and cross-peaks indicates that motions governing cross-relaxation are in the slow-motion limit ($\tau \gg \omega_o^{-1}$) (*65*). Cross-peaks are labeled with APP-C residue number, and arrows indicate sequential connectivities. The slower amide proton exchange rate at pH 4.2 allowed observation of several additional weak d_{NN} cross-peaks that were not observed at pH 5.6.

β-turn. However, long-range NOEs between residues on either side of the turn were not observed. Hence, in comparison, this turn is less stable than the TPEE structure described above. This is in agreement with the lower temperature coefficient for E_{23} as compared with that of Y_{39}, measured by Kroenke et al. (*19*).

The identification of nascent helix on each side of the helix capping box and type I β-turn is supported by NOEs, H^α chemical shifts, and $^3J_{HNH\alpha}$ coupling constants. NOEs characteristic of nascent helix [i.e., sequential $d_{NN}(i, i + 1)$, and medium-range $d_{\alpha N}(i, i + 2)$ and $d_{\beta N}(i, i + 2)$ NOEs (*36*)], were observed for many residues from G_{11} to Q_{44}. Sequential $d_{NN}(i, i + 1)$ NOEs observed at pH 5.6 (Figure 4A) and pH 4.2 (Figure 4B), and short- and medium-range NOEs to amide protons observed at pH 5.6 (Figure 5A) clearly demonstrate that, in addition to the more stable capping box and type I β-turn structures identified for the $_{20}$TPEE$_{23}$ and $_{36}$NPTY$_{39}$ sequences, transient turnlike structures (i.e., nascent helix) exist in flanking regions. Good agreement between these observed NOEs and secondary structure predictions based on observed H^α chemical shifts is obtained for residues 16−18, 27−35, and 40−45, characterizing these regions that surround the TPEE capping box and the NPTY type I β-turn as nascent helix. Furthermore, the $^3J_{HNH\alpha}$ coupling constants obtained from both the 2D HMQC-*J* and 3D HNHA spectra for almost all residues in these regions were between 6.5 and 7.5, consistent with nascent helix.

Additional support for the presence of nascent helix flanking the TPEE capping box and the NPTY type I β-turn is provided by hydrogen exchange measurements. The fast amide proton exchange rates expected for the transiently structured APP-C were best studied by saturation transfer experiments, which can measure rates on the order of T_1, the longitudinal relaxation time of the amide proton (0.1−10 s^{-1}) (*28*). A qualitative measure of the protection from exchange for amide protons of APP-C, *P*, was calculated as described in the Materials and Methods section. The approximate amide proton T_1 values ranged between 494 and 802 ms, with an average value of 633 ms and standard deviation of 68 ms. A plot of the resulting *P* values versus residue is shown in Figure 6. From this plot, it is evident that amide protons in the nascent helical regions described above are protected from exchange relative to an unstructured peptide. Residues G_{11}, V_{13}, and E_{14} also display elevated protection values, a region included in a hydrophobic cluster as discussed below.

The N-Terminal Region Contains a Hydrophobic Cluster. A hydropathy plot for APP-C was constructed by the method of Kyte and Doolittle (*41*) and is shown in Figure 5B. This plot reveals a single hydrophobic region delineated by positive hydropathy values that extends from residue H_9 to A_{17}. Residues I_8, V_{12}, V_{13}, V_{15}, A_{17}, A_{18}, and V_{19}, which contain 12 of the 19 methyl groups in APP-C, contribute significantly to the positive hydropathy values obtained for this region. Of these 12 methyl groups, only the I_8 H^γ and H^δ resonances are well resolved (0.72 and 0.78 ppm). The Val methyl groups all fall between 0.88 and 0.90 ppm, and the Ala methyl groups both resonate at 1.35 ppm. Although limited in frequency resolution, weak NOE cross-peaks observed between these methyl groups as well as other protons reveal transient population of a previously undetected hydrophobic cluster within the N-terminal half of APP-C.

NOEs to side-chain methyl groups from nonsequential residues along with other amide−side-chain and side-chain−side-chain NOEs support transient clustering of hydrophobic side chains. As mentioned previously, this region has several NOEs characteristic of nascent helical structure. In addition, several amide proton NOEs denoted $dx\mathrm{N}(i, i - n)$, where n = 1, 2, 3, or 4 (Figure 5A), are observed. These NOEs could indicate loops or turn structures in this region. Each of the resolved I_8 methyl groups displayed NOE cross-peak intensity in the Val methyl frequency range, indicating hydrophobic clustering of I_8 with at least one of the Val residues. In addition, weak NOE cross-peaks are observed between V_{12} H^N and both methyl groups of I_8 and between V_{12} H^N and S_7 H^α, demonstrating close proximity of I_8 and V_{12}. The closeness of these two regions of sequence is further supported by NOE cross-peaks observed between G_{11} H^α and the I_8 H^γ methyl group, and between the resolved T_6 H^γ methyl group (1.15 ppm) and V_{13} methyl group (0.88 ppm). Confirmation of clustering between the four Val side chains of V_{12}, V_{13}, V_{15}, and V_{19} would not be observable due to resonance overlap of the H^γ methyl groups. However, two NOEs were observed that indicate clustering of Val and Ala side chains: one between the resolved H^β of V_{15} and the methyl group of either A_{17} or A_{18} and the other between a Val H^γ methyl group and the methyl group of either A_{17} or A_{18}. Together, this collection of NOEs suggests that a hydrophobic cluster extends from I_8 through V_{19}, which most likely represents multiple conformations in dynamic equi-

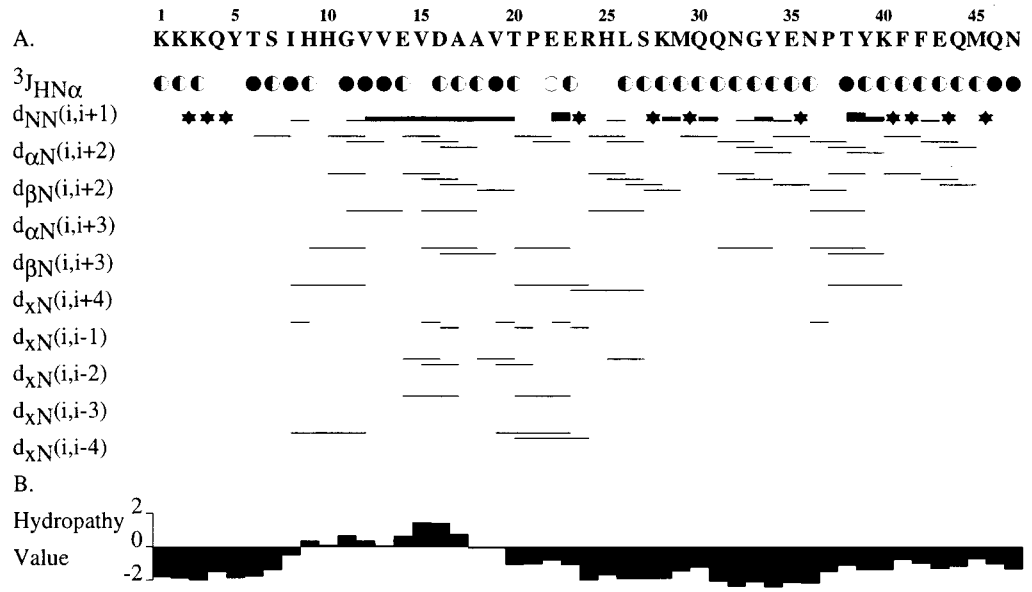

FIGURE 5: (A) Diagram of $^3J_{HNH\alpha}$ coupling constant values and short- and medium-range NOE connectivities observed for APP-C at pH 5.6 and 25 °C. Coupling constant values are indicated by open circles ($^3J_{HNH\alpha} < 6$ Hz), half circles (6 Hz $\leq$ $^3J_{HNH\alpha} \leq$ 8 Hz), and solid circles ($^3J_{HNH\alpha} > 8$ Hz). Stars indicate that the d_{NN} NOE cross-peak was not observable due to spectral overlap. Standard nomenclature for NOE connectivities is employed, and *x* denotes any proton of residue *i*. (B) Hydropathy plot of APP-C. The plot was constructed by the method of Kyte and Doolittle (*41*) with a 9-residue scanning window. The average hydropathy value over the 9-residue window is plotted versus the central residue in the window. Positive values indicate hydrophobic regions.

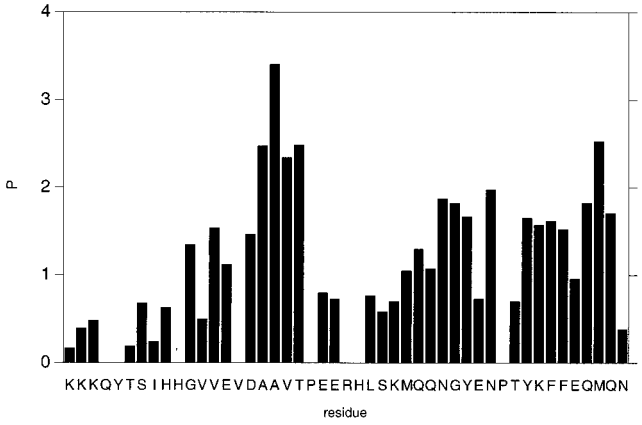

FIGURE 6: P/T_1 versus residue for amide protons of APP-C calculated from saturation transfer experiments at pH 6.0 at 298 K. Missing values correspond to overlapped peaks (Q_4, Y_5, V_{15}, and R_{24}), peaks too weak to characterize (H_{10} and H_{25}), or Pro residues (P_{21} and P_{37}).

librium. In this situation, NOE-derived distance constraints cannot be used to determine one single meaningful structure.

Identification of Transient Hydrogen Bonds. Measurement of downfield amide proton shifts upon increasing pH is a sensitive method for detecting transient hydrogen bonds between amide protons and titrating carboxylate groups (*42*). To investigate such interactions in APP-C, pH titration studies were performed to determine the pK_a for each of the seven carboxylate groups in the protein and apparent pK_a for amide protons that shift in response to titration. Matching of a carboxylate pK_a value with the apparent pK_a of a downfield-shifted amide proton identifies through-space interactions between the two groups. The chemical shifts of amide protons, the H^β protons of Asp, H^γ protons of Glu, and the H^α proton of the C-terminal Asn were obtained from TOCSY spectra recorded at seven pH values ranging from 2.0 to 6.1. The pK_a values were determined by plotting the chemical shifts of these protons as a function of pH and fitting this curve to the single proton titration curve (*42*) as described in Materials and Methods.

Table 1: pK_a and $\Delta\delta$ Values of APP-C Amide Protons, Asp and Glu Residues, and C-Terminus

residue	proton[a]	pK_a[b] ± error	$\Delta\delta^{b,c}$
E14	H^γs	4.3 ± 0.1	−0.17
D16	H^N	4.2 ± 0.1	−0.16
D16	H^βs	3.9 ± 0.1	−0.22
		4.0 ± 0.1	−0.23
A18	H^N	3.8 ± 0.1	0.12
T20	H^N	4.5 ± 0.1	0.19
E22	H^N	4.0 ± 0.1	0.24
E22	H^γs	4.1 ± 0.1	−0.21
E23	H^N	4.4 ± 0.1	0.11
E23	H^γs	4.3 ± 0.1	−0.22
E35	H^N	4.3 ± 0.1	0.22
E35	H^γs	4.5 ± 0.1	−0.14
E43	H^N	4.4 ± 0.1	0.14
E43	H^γs	4.5 ± 0.1	−0.15
N47	H^N	3.4 ± 0.1	−0.28
N47	H^α	3.4 ± 0.1	−0.24

[a] Apparent pK_a values are given for those H_N for which $|\delta_{HN}{}^{pH5.6} - \delta_{HN}{}^{pH2.0}| > 0.10$. [b] Obtained from nonlinear least-squares fit of the data to the single proton titration curve. [c] The net titration shift, $\Delta\delta = (\delta_{A^-} - \delta_{HA})$, where δ_{A^-} and δ_{HA} are the chemical shifts corresponding to the fully deprotonated and fully protonated states, respectively.

All carboxylate groups in APP-C display pK_a values that are consistent with a solvent-exposed environment (Table 1). The pK_a values obtained for the five Glu residues in APP-C (E_{14}, E_{22}, E_{23}, E_{35}, and E_{43}) range from 4.1 to 4.5 based on titration curves of their H^γ protons. As expected, a slightly lower pK_a value of 3.9 was determined from the resolved H^β proton chemical shifts from the single Asp residue, D_{16}. The calculated pK_a of the C-terminus was determined to be 3.4 by following the titration of the H^α proton of N_{47}. The sigmoidal shape and small standard deviation of data points from the best-fit theoretical curve

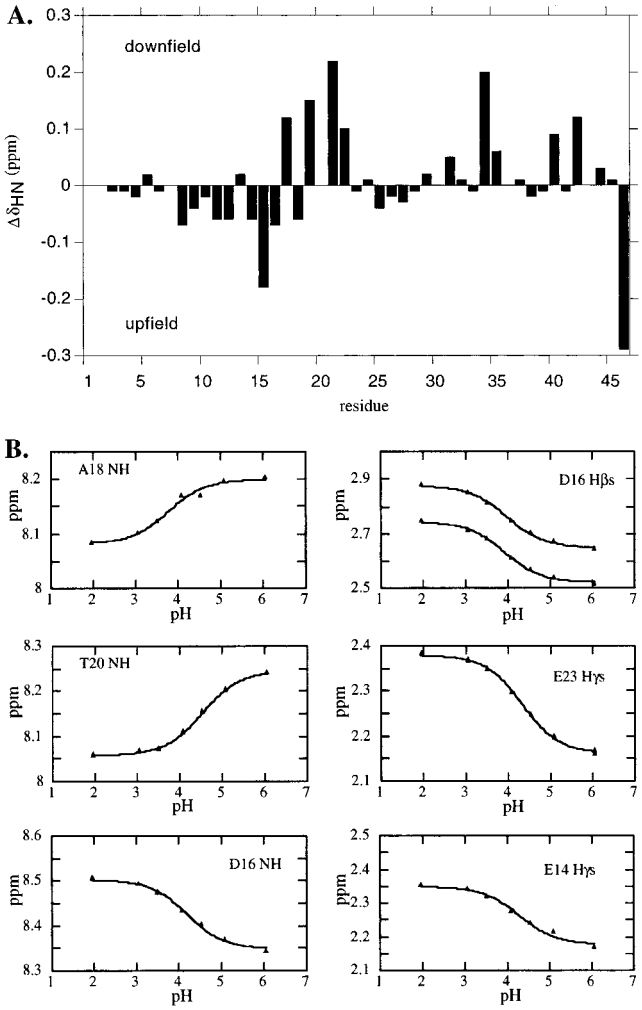

FIGURE 7: (A) Plot of the amide proton chemical shift differences observed for APP-C between pH 5.6 and 2.0, $\Delta\delta_{HN} = \delta_{HN}^{pH5.6} - \delta_{HN}^{pH2.0}$ (ppm), for each amide proton in APP-C. (B) Representative plots of chemical shift vs pH for APP-C protons. Plots are shown for the H^N protons of A_{18}, T_{20}, and D_{16}, along with the H^β protons of D_{16} and the H^γ protons of E_{23} and E_{14}. Curve fits were obtained by a nonlinear least-squares fit of the single proton titration curve.

for each carboxylate group demonstrate that the data is fit well by a single proton titration.

The change in APP-C amide proton chemical shifts between pH 2.0 and 5.6 are shown in Figure 7A. This pH range allows identification of amide protons whose chemical environments are affected by deprotonation of Glu, Asp, or C-terminal carboxylate groups. Intrinsic effects cause an amide proton of a residue containing a titrating carboxylate group to shift upfield upon deprotonation (*43*, *44*), whereas a downfield shift of an amide proton upon increasing pH is an indication of a hydrogen-bond interaction between the amide proton and a titrating carboxylate group. As illustrated in Figure 7A, several residues display significant changes in amide proton chemical shift ($\geq \pm 0.1$ ppm) over this pH range.

Upfield shifts for the D_{16} and N_{47} amide protons are observed, indicating that these N−H groups are not involved in hydrogen bonds with titrating groups. The large upfield shift of N_{47} H^N is as expected due to intrinsic effects, and the resulting p*K* value correlates well with that obtained from N_{47} H^α. The net upfield shift of D_{16} H^N is also as expected due to intrinsic effects. However, the resulting pK_a value of 4.2 ± 0.1 is not within experimental error of the pK_a value of 3.9 ± 0.1 determined for its own side chain (Table 1). Furthermore, the H^α of D_{16} titrates with a pK_a value of 4.4 ± 0.1, indicating that a conformational change in response to titration of a Glu residue occurs that influences the backbone conformation of D_{16}. The pK_a of the sequentially nearest Glu residue (E_{14}) is 4.3 and is a likely candidate for imparting possible indirect effects on the conformation of D_{16}. The correlated titration curves for the H^N of D_{16} and the H^γ protons of E14 are shown in Figure 7B.

For Glu residues, upfield shifts are observed in highly structured proteins due to intrinsic effects (*45*). However, downfield shifts are observed for Glu residues in flexible regions of proteins due to the formation of a transient cyclic intraresidue hydrogen bond between the amide proton and the side-chain carboxylate group (*43*). The downfield amide proton chemical shifts of E_{22}, E_{23}, E_{35}, and E_{43} indicate that transient intraresidue hydrogen-bond interactions are present for these residues. The amide proton of E_{14} does not show a significant dependence on pH (Figure 7A), with a total change in chemical shift of 0.04 ppm across the full pH range examined. An estimate of the extent of intraresidue hydrogen bonding in Glu residues can be approximated from the magnitude of the downfield H^N shift upon increasing pH (*42*). A magnitude of 0.41 ppm was estimated as the net downfield shift for an amide proton in a stable hydrogen bond with a titrating carboxylate group (*42*). With this as an approximate standard for 100% occupancy of the intraresidue hydrogen bond, the largest net downfield shift of 0.24 ppm observed for E_{22} H^N indicates that this cyclic intraresidue hydrogen bond is approximately 58% populated. Similarly, intraresidue hydrogen-bond populations are estimated as 10% for E_{14}, 27% for E_{23}, 54% for E_{35}, and 34% for E_{43}. Hence, the amide protons of E_{22} and E_{35} are less likely to be involved in interresidue hydrogen bonding, while E_{14}, E_{23}, and E_{43} are more likely to participate in such a structural hydrogen bond through either their amide proton or carboxylate group. For E_{23} and E_{43}, reduced occupancy of the intraresidue hydrogen bond is consistent with the capping box and nascent helix characterized for these regions, respectively. The reduced occupancy of the E_{14} intraresidue hydrogen bond could be accounted for by interresidue hydrogen bonding between either its amide proton or carboxylate group and polar group(s) of another side chain.

Other than titrating residues, the only additional amide protons that have significant (>0.10 ppm) downfield chemical shifts upon increasing pH are those of A_{18} and T_{20}. Correlation of an amide proton to its hydrogen-bonding partner is obtained by comparing its apparent pK_a value to the pK_a values of titrating side chains (*43*). The amide proton of A_{18} has a low pK_a value of 3.8, which clearly matches with the pK_a value of the carboxylate of D_{16} (Figure 7B). The hydrogen bond between the amide proton of A_{18} and the carboxylate of D_{16} presents a distance constraint which is very similar to a $d_{xN}(i, i + 2)$ NOE, where *x* is a proton in the side chain of residue *i*. This type of NOE is often found in nascent helical structures and is consistent with the $\Delta\delta_{H\alpha}$ values observed for residues D_{16}, A_{17}, and A_{18}. Additional support for this hydrogen bond is provided by saturation transfer measurements in which A_{18} showed the highest protection from exchange (Figure 6). The T_{20} amide proton has an apparent pK_a value of 4.5, which indicates

that it is hydrogen-bonded to the carboxylate of a Glu residue (either E_{14}, E_{23}, E_{35}, or E_{43}). The high protection factor for amide proton exchange observed for T_{20} (Figure 6) supports hydrogen bonding of this amide NH group. On the basis of numerous NOE contacts observed across the TPEE turn and the reduced population of the E_{23} intraresidue hydrogen bond, the carboxylate of E_{23} is assigned as the acceptor for the T_{20} H^N hydrogen-bond interaction. The corresponding titration curves are shown in Figure 7B. A Thr residue in the Ncap position of the N-terminal helix capping box structure is stabilized by such a hydrogen bond as well as by the reciprocal hydrogen bond between the Thr side chain and the H^N of the N3 residue, E_{23} in the case of APP-C (*46*). The T_{20} $O^{\gamma}-E_{23}$ H^N reciprocal hydrogen bond is supported by NOEs, and is consistent with the low-temperature coefficient for E_{23} observed by Kroenke et al. (*19*). Together, these reciprocal hydrogen bonds can explain the disruption of the E_{23} cyclic intraresidue hydrogen bond observed.

Molecular Modeling Confirms the VTPEER Capping Box. Residues displaying characteristics consistent with an N-terminal helix capping box structure, $_{19}$VTPEER$_{24}$, were modeled with X-PLOR 3.8 (*34*) as described in Materials and Methods. The three-dimensional structure of this region was calculated with 28 NOE restraints, three dihedral angle restraints, and two hydrogen-bond distance restraints. NOEs obtained from 2D and 3D NOESY experiments at pH 5.6 were classified as strong (1.8−2.7 Å), medium (1.8−3.3 Å), weak (1.8−4.3 Å), and very weak (1.8−5.0 Å) (*47*). The strong short-range $d_{\alpha N}$, $d_{\beta N}$, and $d_{xN}(i, i+1)$ NOEs were omitted from the analysis since these arise from unfolded states that are in equilibrium with the capping box structure. Torsion angle constraints for the φ angles of V_{19} and E_{22} were derived from measured $^3J_{HNH\alpha}$ values to be $-120° \pm 40°$ for V_{19} ($^3J_{HNH\alpha} > 8$ Hz) and $-65° \pm 30°$ for E_{22} ($^3J_{HNH\alpha} < 6$ Hz). The ω torsional angle for P_{21} was constrained to $-180° \pm 20°$ to specify the trans conformation. Two hydrogen bonds identified from the pH titration experiment were constrained (1.8−2.5 Å): the H^N of T_{20} to the carboxylate of E_{23} and the intraresidue hydrogen bond between the H^N and carboxylate of E_{22}. As described in Materials and Methods, an ensemble of 20 structures was generated in which no dihedral, NOE, or hydrogen-bond distance violations exist that are greater than 5°, 0.3 Å, or 0.3 Å, respectively. These structures had low energies (total X-PLOR energy < 30 kcal/mol). The backbone atomic RMSD of residues $_{20}$TPEE$_{23}$ for the 20 calculated structures was less than 0.80 Å, indicating that the structure is well-defined. As a control, an ensemble of 20 structures was calculated with no constraints, which yielded a significantly higher average backbone atomic RMSD of 1.55 Å.

A representative structure is displayed in Figure 8, illustrating the characteristic features of the N-terminal helix capping box structure. The defined capping box hydrogen bond between the H^N of the Ncap residue, T_{20}, and the side chain of the N3 residue, E_{23}, was observed, along with the reciprocal hydrogen bond between the N3 side chain and the Ncap H^N. These reciprocal hydrogen bonds, along with two backbone hydrogen bonds between the carbonyl oxygen of T_{20} and the H^N of E_{23} and R_{24}, are observed in all the calculated structures. These additional hydrogen bonds would not have been directly detected in our NMR measurements but are supported by NOE contacts discussed previously for

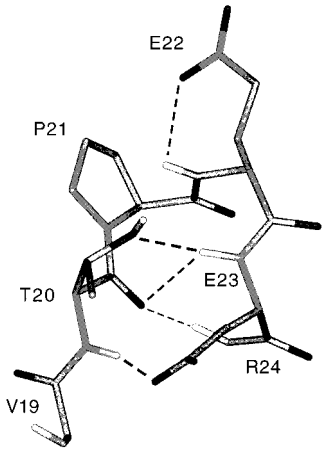

FIGURE 8: Representative structure of $_{19}$VTPEER$_{24}$ N-terminal capping box calculated by X-PLOR 3.8. Backbone and side-chain O atoms are black, H^N atoms and T_{20} side-chain hydroxyl H^{δ} atom are white, and C and N atoms are gray. Hydrogen bonds are indicated by dashed lines. The figure was generated by use of insightII (MSI) and Adobe Illustrator.

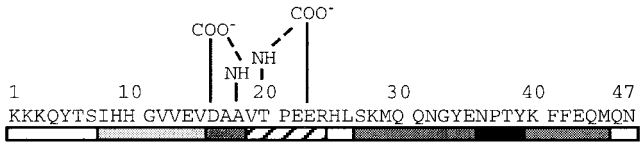

FIGURE 9: Schematic diagram of the transient structural features determined for APP-C. Hydrogen bonds between backbone N−H and side-chain carboxylate groups are depicted as dashed lines. Shown are the hydrophobic cluster (light gray), nascent helical regions (darker gray), the N-terminal helix capping box (hatched), and the type I β-turn (black).

the capping box. All structures displayed the hydrophobic interaction between the side chains of V_{19} and R_{24}, which classifies this capping box as type Ib as defined by Aurora and Rose (*39*). A striking sequence similarity with the helix capping box composed of residues $_{43}$LTPDER$_{48}$ in the Trp repressor (*40*, *48*) was identified through investigation of other characterized helix capping box structures. The main difference between the two structures is the degree of helicity of residues N4 and N5. While the capping box of the Trp repressor is followed by a stable α-helix, the APP-C capping box is followed by nascent helix, which results in larger $^3J_{HNH\alpha}$ and a lack of $i, i+3$ and $i, i+4$ NOE contacts between these and subsequent residues. Hence, the APP-C capping box structure is less stable than an N-terminal helix capping box in the context of a stable folded protein.

DISCUSSION

The Cytoplasmic Tail of APP Possesses Transient Structure over a Broad pH Range. The present study has shown that although APP-C does not have a stable folded structure over the pH range examined, regions of transient structure exist and have been characterized by a combination of H^{α} chemical shifts, NOE analysis, $^3J_{HNH\alpha}$ coupling constants, saturation transfer, and identification of hydrogen bonds between amide protons and titrating carboxylate groups. The transient structural features of APP-C are summarized in Figure 9. These features include a hydrophobic cluster involving residues I_8 through V_{19}, an N-terminal helix capping box formed by residues $_{19}$VTPEER$_{24}$, a type I β-turn at $_{36}$NPXY$_{39}$, and nascent helix for residues $_{16}$DAA$_{18}$,

$_{27}$SKMQQNGYE$_{35}$, and $_{40}$KFFEQM$_{45}$. Transient hydrogen bonds were identified between the amide proton of A_{18} and the carboxylate of D_{16} and between the amide proton of T_{20} and the carboxylate of E_{23}, in good agreement with the nascent helix and capping box structures determined for these regions, respectively.

A key question is whether the transient structural features characterized here for the APP-C peptide in aqueous solution accurately represent this region in the context of native full-length APP. In its native environment, the APP cytoplasmic tail is juxtaposed to the membrane. Hence, it is possible that the local environment of the membrane surface alters its conformation. In addition, since it is easier to continue a helix than to initiate one, it is also possible that the putative transmembrane helix could extend into the cytoplasmic tail of native APP. However, APP-C and several peptide fragments derived from it have displayed function, indicating that neither the adjacent transmembrane segment nor the membrane is required for interactions with cytosolic binding partners. The 32-residue $D_{16}-N_{47}$ fragment efficiently competes for binding of GST-Fe65 to all three APP isoforms in PC12 cell extracts (*16*). Similarly, the 20-residue H_9-K_{28} fragment shows specific G_o-activating function and prevents anti-APP antibody coimmunoprecipitation of native APP and G_o from bovine brain membranes (*11*). In addition, affinity chromatography employing an immobilized peptide comprising the APP cytoplasmic tail preceded by four residues from the transmembrane region was used to identify a 127-kDa APP-binding protein in rat brain cytosol, and this binding protein coimmunoprecipitated with full-length APP in PC12 cells (*49*). Hence, the retention of native function by peptide fragments derived from the cytoplasmic tail of APP suggests that the transient structural features described here for APP-C are likely to be adopted by the corresponding regions in native APP.

Binding of APP-C to X11 Requires Little Conformational Rearrangement. Recent crystal structures of the PID domain of the neuronal protein X11 bound to each of two peptide fragments derived from APP-C (10 and 14 residues long, respectively, and each spanning the NPTY sequence) show a highly specific interaction involving approximately 11 residues of APP-C (*18*). Although the function of X11 is currently not known, the high specificity and affinity of this protein in vivo with the known endocytosis target sequence of APP suggests a possible role in the processing of APP (*15*). In fact, binding of X11 to the cytoplasmic tail of APP has been shown to reduce cellular processing of APP in transfected cells (*50*).

It is informative to compare the free and bound conformations of this region of APP-C in order to gain insight into the energetics of the APP/X11 interaction. In the bound state, three structural motifs are found in the APP peptide: β-strand, type I β-turn, and 3_{10} helix. The β-strand is formed by four residues ($_{32}$NGYE$_{35}$) N-terminal to the NPTY turn that are incorportated into a β-sheet of the X11 PID domain, where the peptide backbone participates in antiparallel β-sheet hydrogen bonding and side chains participate in electrostatic and hydrophobic interactions with the X11 binding surface. As presented above, these four residues are characterized as nascent helix in free APP-C. Hence, binding to X11 requires a conformational rearrangement for these four residues. Since nascent helix is a transient structure, backbone and side-chain conformational entropy is lost upon binding. However, this loss is offset by favorable electrostatic and hydrophobic interactions involving these four residues in the bound complex, which contribute significantly to the high affinity and specificity of binding (*18*).

The type I β-turn and 3_{10} helix of the X11-bound APP peptide are stabilized forms of structural features found in free APP-C. The $_{36}$NPTY$_{39}$ sequence binds as a type I β-turn, retaining and stabilizing the conformation found for these residues in APP-C free in solution. Preordering of these residues into the conformation required for binding is entropically favorable, and additional hydrophobic and enthalpic interactions are gained upon binding. Similarly, three residues C-terminal to the turn ($_{40}$KFF$_{42}$) form a 3_{10} helix in the bound state, again retaining and stabilizing the helical nature of these residues observed in APP-C free in solution. Although binding would result in loss of entropy of this region since sampling of turnlike conformations would no longer be transient, hydrophobic interactions between the APP-C Phe side chains and a hydrophobic pocket in the X11 binding surface are gained, supported by the fact that these C-terminal residues contribute positively to binding affinity (*18*). Residues after the β-turn endocytosis signals in the low-density lipoprotein receptor and lysosomal acid phosphatase contribute to the endocytosis rate of these two proteins (*51*, *52*), indicating the importance of these residues in interactions with the endocytosis machinery. These residues in the low-density lipoprotein receptor are also nascent helix and have been suggested as part of the internalization motif for this protein (*53*). For the case of APP binding to X11, the native type I β-turn and a short segment of nascent helix are stabilized in the bound state, while a short segment of nascent helix is rearranged into β-strand upon binding.

Helix Capping Box May Play an Important Role in APP-C Interactions with G_o and Fe65. Other neuronal proteins that APP-C interacts with include the heterotrimeric protein G_o and one of the PID domains of the adapter protein Fe65. However, the regions of APP-C recognized by these proteins differ from the 11 residues recognized by X11. Unlike X11, the C-terminal PID domain of Fe65 does not bind to a short peptide containing the YENPTY sequence (a 20-residue peptide comprising APP-C residues S_{27} through Q_{46}) but does bind with high affinity to a peptide containing both the helix capping box and the β-turn (residues D_{16} through N_{47}) (*16*). Similarly, G_o binds to a 20-residue APP peptide comprising residues H_9-K_{28} (*54*), an interaction that activates G_o function in vivo (*55*). The two N-terminal His residues and the BBXXB (where B is a basic residue and X is any residue) sequence at the C-terminus of this peptide are essential for G_o binding (*11*). This region includes the capping box but not the NPTY β-turn important in recognition by X11. Hence, G_o and the PID domain of FE65 are expected to bind to APP-C in a manner different from X11, and it is reasonable to expect that the capping box structure and flanking nascent helix regions play important roles in recognition.

Transient Structure of APP-C Is Analogous to an Early Folding State. The functional relevance of the transient structure observed in APP-C is an important question to consider. The nascent helix, capping box, type I β-turn, and hydrophobic cluster characterized here for regions within APP-C are relatively unstable structures in rapid equilibrium

with unfolded states. This description of transient structure shares many characteristics used to describe early states in the protein folding process (*56*, *57*). Cytoplasmic domains of other membrane-spanning proteins have been shown to possess unstable structure in physiologically compatible buffers (*58*−*61*). Hence, transient structure appears to be common in cytoplasmic domains. In APP-C, no tertiary contacts occur, secondary structural elements are transiently populated, and a cluster of hydrophobic side chains is observed. As demonstrated by studies of the folding kinetics of barnase (*62*) and lysozyme (*63*), helix capping can be an early folding event. Similarly, relatively unstable nascent helix and β-turns characterized in peptide fragments of larger proteins have been proposed as important features of protein folding initiation states, since these transient structures would prevent the trapping of incorrectly folded conformations (*36*, *64*). The striking similarity between protein folding initiation states and the transient structure of APP-C suggests that binding of cytoplasmic tails to their intracellular binding partners resembles the final step(s) of protein folding, in which some structural features of the folding initiation state (unbound state) are retained and stabilized in the folded (bound) state and others are rearranged into different conformations upon folding (binding) without large energetic cost. However, the bound state must not be too stable, since this interaction must be transient. For APP-C and other cytoplasmic tails that lack stable folded structure, the higher entropy of the unliganded state would weaken binding interactions by lowering the free energy of the unbound state, while local preordering of regions into transient structure may govern specificity of binding to cellular factors.

In conclusion, the results presented here have revealed specific structural features of APP-C free in solution, known to contain regions recognized by adapter and signaling proteins. Interactions between APP-C and these proteins are required for processing of APP and for intracellular signaling. Alteration of these interactions may lead to abnormal processing or targeting that would contribute to the pathogenesis of Alzheimer's disease or to abnormal signaling that can result in apoptosis. Knowledge of the structure of APP-C in both free and bound conformations is required to understand the energetics of binding, and for gaining insight on the transient interactions that facilitate normal cellular function. The present study shows that APP-C resembles a partially folded protein and possesses structural features free in solution that are present when bound to the PID domain of the neuronal protein X11. Hence, it is likely that the transient structure of the APP cytoplasmic tail plays an important role in guiding recognition by intracellular machinery. This type of preordering of the unbound structure has important implications for our understanding of the mechanisms and energetics of formation of multiple-protein complexes involved in intracellular trafficking and signaling.

ACKNOWLEDGMENT

We thank T. Thanhauser and the Cornell BioResources Center for HPLC purification, mass spectrometry, and amino acid analysis, L. E. Kay (University of Toronto) for the use of his Varian NMR pulse sequence library, F. Delaglio (NIH/NIDDK) for use of NMRPipe and nmrDraw programs, D. Garrett (NIH/NIDDK) for use of PIPP, and D. A. Torchia (NIH/NIDCR) for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Three tables containing chemical shift assignments for APP-C at pH 4.2, 5.6, and 7.1. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Selkoe, D. J., Yamazaki, T., Citron, M., Podlisny, M. B., Koo, E. H., Teplow, D. B., and Haas, C. (1996) in *Annals of the New York Academy of Sciences* (Wurtman, R. J., Corkin, S., Growdon, J. H., and Nitsch, R. M., Eds.) pp 57−64, New York Academy of Sciences, New York.
2. Lai, A., Sisodia, S. S., and Trowbridge, I. S. (1995) *J. Biol. Chem. 270*, 3565−3573.
3. Lai, A., Gibson, A., Hopkins, C. R., and Trowbridge, I. S. (1998) *J. Biol. Chem. 273*, 3732−3739.
4. De Strooper, B., Umans, L., Van Leuven, F., and Van Den Berghe, H. (1993) *J. Cell Biol. 121*, 295−304.
5. Koo, E. H., and Squazzo, S. L. (1994) *J. Biol. Chem. 269*, 17386-17389.
6. Kornfeld, S., and Mellman, I. (1989) *Annu. Rev. Cell. Biol. 5*, 483−525.
7. Fukuda, M. (1991) *J. Biol. Chem. 266*, 21327−21330.
8. Haass, C., Koo, E. H., Capell, A., Teplow, D. B., and Selkoe, D. J. (1995) *J. Cell Biol. 128*, 537−547.
9. Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K.-H., Multhaup, G., Beyreuther, K., and Muller-Hill, B. (1987) *Nature 325*, 733−736.
10. Yamatsuji, T., Okamoto, T., Takeda, S., Murayama, Y., Tanaka, N., and Nishimoto, I. (1996) *EMBO J. 15*, 498−509.
11. Nishimoto, I., Okamoto, T., Matsuura, Y., Takahashi, S., Okamoto, T., Murayama, Y., and Ogata, E. (1993) *Nature 362*, 75−79.
12. Okamoto, T., Takeda, S., Murayama, Y., Ogata, E., and Nishimoto, I. (1995) *J. Biol. Chem. 270*, 4205−4208.
13. Yamatsuji, T., Matsui, T., Okamoto, T., Komatsuzaki, K., Takeda, S., Fukumoto, H., Iwatsubo, T., Suzuki, N., Asami-Odaka, A., Ireland, S., Kinane, T. B., Giambarella, U., and Nishimoto, I. (1996) *Science 272*, 1349−1352.
14. Chow, N., Korenberg, J. R., Chen, X.-N., and Neve, R. L. (1996) *J. Biol. Chem. 271*, 11339−11346.
15. Borg, J.-P., Ooi, J., Levy, E., and Margolis, B. (1996) *Mol. Cell. Biol. 16*, 6229−6241.
16. Zambrano, N., Buxbaum, J. D., Minopoli, G., Fiore, F., De Candia, P., De Renzis, S., Faraonio, R., Sabo, S., Cheetham, J., Sudol, M., and Russo, T. (1997) *J. Biol. Chem. 272*, 6399−6405.
17. Guenette, S. Y., Chen, J., Paul, J. D., and Tanzi, R. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10832−10837.
18. Zhang, Z., Lee, C.-H., Mandiyan, V., Borg, J.-P., Margolis, B., Schelssinger, J., and Kuriyan, J. (1997) *EMBO J. 16*, 6141−6150.
19. Kroenke, C. D., Ziemnicka-Kotula, D., Xu, J., and Palmer, A. G. (1997) *Biochemistry 36*, 8145−8152.
20. Goldgaber, D., Lerman, M. I., McBride, O. W., Saffiotti, U., and Gajdusek, D. C. (1987) *Science 235*, 877−880.
21. Robakis, N. K., Ramakrishna, N., Wolfe, G., and Misniewski, H. M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4190−4194.
22. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
23. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
24. Bodenhausen, G., Kogler, H., and Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370−388.
25. Shaka, A. J., Lee, C. J., and Pines, A. (1988) *J. Magn. Reson.* 274−293.
26. Kay, L. E., and Bax, A. (1990) *J. Magn. Reson. 86*, 100−126.
27. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
28. Skelton, N. J., Kordel, J., Akke, M., and Chazin, W. J. (1992) *J. Mol. Biol. 227*, 1100−17.

29. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75.
30. Connelly, G. P., Bai, Y., Mei-Fen, J., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 87−92.
31. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−93.
32. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson. 95*, 214−220.
33. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
34. Brünger, A. T. (1992) *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR*, Harvard University Press, New Haven, CT.
35. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
36. Dyson, J. H., Rance, M., Houghten, R. A., Wright, P. E., and Lerner, R. (1988) *J. Mol. Biol. 201*, 201−217.
37. Seale, J. W., Srinivasan, R., and Rose, G. D. (1994) *Protein Sci. 3*, 1741−1745.
38. Stickle, D. F., Presta, L. G., Dill, K. A., and Rose, G. D. (1992) *J. Mol. Biol. 226*, 1143−1159.
39. Aurora, R., and Rose, G. D. (1998) *Protein Sci. 7*, 21−38.
40. Harper, E. T., and Rose, G. D. (1993) *Biochemistry 32*, 7605−9.
41. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105.
42. Bundi, A., and Wüthrich, K. (1979) *Biopolymers 18*, 299−311.
43. Szyperski, T., Antuch, W., Schick, M., Betz, A., Stone, S. R., and Wüthrich, K. (1994) *Biochemistry 33*, 9303−9310.
44. Schaller, W., and Robertson, A. D. (1995) *Biochemistry 34*, 4714−4723.
45. Khoda, D., and Inagaki, F. (1992) *Biochemistry 31*, 11928−11939.
46. Doig, A., Macarthur, M. W., Stapley, B. J., and Thornton, J. M. (1997) *Protein Sci. 6*, 147−155.
47. Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., and Gronenborn, A. M. (1986) *EMBO J. 5*, 2729−2735.
48. Lawson, C. L., Zhang, R. G., Schevitz, R. W., Otwinowski, Z., Joachimiak, A., and Sigler, P. B. (1988) *Proteins: Struct., Funct., Genet. 3*, 18−31.
49. Watanabe, T., Sukegawa, J., Sukegawa, I., Tomita, S., Iijima, K., Oguchi, S., Suzuki, T., Nairn, A. C., and Greengard, P. (1999) *J. Neurochem. 72*, 549−556.
50. Borg, J. P., Yang, Y., De Taddeo-Borg, M., Margolis, B., and Turner, R. S. (1998) *J. Biol. Chem. 273*, 14761−14766.
51. Lehmann, L. E., Eberle, W., Krull, S., Prill, V., Schmidt, B., Sander, C., von Figura, K., and Peters, C. (1992) *EMBO J. 11*, 4391−4399.
52. Chen, W.-J., Goldstein, J. L., and Brown, M. S. (1990) *J. Biol. Chem. 265*, 3116−3123.
53. Bansal, A., and Gierasch, L. M. (1991) *Cell 67*, 1195−1201.
54. Giambarella, U., Yamatsuji, T., Okamoto, T., Matsui, T., Ikezu, T., Murayama, Y., Levine, M. A., Katz, A., Gautam, N., and Nishimoto, I. (1997) *EMBO J. 16*, 4897−4907.
55. Lang, J., Nishimoto, I., Okamoto, T., Sadoul, K., Regazzi, R., Kiraly, K., Willer, U., and Wollheim, C. B. (1995) *EMBO J. 14*, 3635−3644.
56. Dobson, C. M., and Karplus, M. (1999) *Curr. Opin. Struct. Biol. 9*, 92−101.
57. Wright, P. E., Dyson, H. J., and Lerner, R. A. (1988) *Biochemistry 27*, 7167−7175.
58. Hubbard, J. A., MacLachlan, L. K., Meenan, E., Salter, C. J., Reid, D. G., Lahouratate, P., Humphries, J., Stevens, N., Bell, D., Neville, W. A., Murray, K., and Darker, J. G. (1994) *Mol. Membr. Biol. 11*, 263−269.
59. Sandoval, I. V., Arredondo, J. J., Alcalde, J., Gonzalez Noriega, A., Vandekerckhove, J., Jimenez, M. A., and Rico, M. (1994) *J. Biol. Chem. 269*, 6622−6631.
60. Wilde, A., Dempsey, C., and Banting, G. (1994) *J. Biol. Chem. 10*, 7131−7136.
61. Wray, V., Mertins, D., Kiess, M., Henklein, P., Trowitzsch-Kienast, W., and Schubert, U. (1998) *Biochemistry 37*, 8527−8538.
62. Serrano, L., Sancho, J., Hirshberg, M., and Fersht, A. R. (1992) *J. Mol. Biol. 227*, 544−559.
63. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) *Nature 358*, 302−307.
64. Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., and Wright, P. E. (1988) *J. Mol. Biol. 201*, 161−200.
65. Ernst, R. R., Bodenhausen, G., and Wokaun, A. (1997) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, England.

BI992580M