NMR Study Suggests a Major Role for Arg111 in Maintaining the Structure and Dynamical Properties of Type II Human Cellular Retinoic Acid Binding Protein[†]

Lincong Wang and Honggao Yan*

Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824

Received May 4, 1998; Revised Manuscript Received July 21, 1998

ABSTRACT: The solution structure of a site-directed mutant of type-II human cellular retinoic acid binding protein (CRABPII) with Arg111 replaced by methionine (R111M) has been determined by NMR spectroscopy. The sequential assignments of the ^{1}H and ^{15}N resonances of apo-R111M were established by multinuclear multidimensional NMR. The solution structure was calculated from 2302 distance restraints and 77 ϕ dihedral restraints derived from the NMR data. The root-mean-square deviation of the ensemble of 28 refined conformers that represent the structure from the mean coordinate set derived from them was 0.54 ± 0.26 and 0.98 ± 0.23 Å for the backbone atoms and all heavy atoms, respectively. The solution structure of apo-R111M is similar to that of wild-type apo-CRABPII. However, there are significant conformational differences between the two proteins, localized mainly to three segments (Leu19−Ala36, Glu73−Cys81, and Leu99−Pro105) clustered around the ligand entrance more than 17 Å away from the point mutation. In apo-R111M, all the three segments move toward the center of the ligand entrance so that the opening of the ligand-binding pocket in apo-R111M is much smaller than that in wild-type apo-CRABPII. Furthermore, the ligand-binding pocket of apo-R111M, especially the ligand entrance, is much less flexible than that of apo-CRABPII. Surprisingly, apo-R111M is more similar to holo-CRABPII than to apo-CRABPII in both structure and dynamical properties. The conformational and dynamical changes caused by the mutation are similar to those induced by binding of RA, although the magnitudes of the changes caused by the mutation are smaller than those induced by binding of RA. The results suggest that Arg111 plays a critical role in determining the structure and dynamical properties of CRABPII.

Retinoic acid (RA),[1] a hormonally active metabolite of vitamin A, has profound effects on cell growth, differentiation and morphogenesis (*1*). Either RA deficiency or excess results in developmental defects. The plethora of biological effects of RA and other retinoids are believed to be mediated by various nuclear RA receptors (RARs and RXRs), all of which are RA-activated transcriptional factors that regulate the expression of target genes (*2*).

In addition to nuclear RA receptors, two types of cellular retinoic acid binding proteins (CRABPI and CRABPII) are also implicated in RA signal transduction (*3*). However, their physiological functions have not been established at present. It is thought that CRABPs may protect cells from the detrimental effects of RA by preventing it from being incorporated into membranes. Indeed, in the tissues that express CRABPs, RA is found predominately in the protein-bound forms. More recently, a model for RA metabolism with CRABPI playing more active roles has been proposed (*4*). It appears that the two proteins may play distinct cellular functions because they differ in spatial and temporal expression patterns. It has been shown that production of RA is correlated with the expression of CRABPII but not CRABPI (*5*). Overexpression of CRABPII enhances cellular response to RA in breast cancer cells (*6*), but overexpression of CRABPI decreases the biological potency of RA in F9 teratocarcinoma cells (*7*).

The crystal structures of holo-CRABPI and holo-CRABPII (both in complex with *all-trans*-RA) have been determined (*8*). Both structures consist of a helix-turn-helix motif and two nearly orthogonal five-stranded β-sheets. A flattened β-barrel formed by the two β-sheets serves as a deep ligand-binding pocket with the helix-turn-helix motif at the entrance. RA is buried in the binding pocket with its carboxyl group interacting with Arg111, Arg132, and Tyr134 (CRABPII numbering) at the bottom of the pocket. Arg132 and Tyr134 interact directly with the carboxyl group of RA, whereas the interaction between Arg111 and the carboxyl group of RA is mediated by a water molecule. The β-ionone ring is

[†] This work was supported by funds from the REF Center of Protein Structure and Design and the Cancer Center at Michigan State University and NIH (GM51901). [†]This study made use of a Bruker DMX-600 NMR spectrometer at The Ohio State University funded by NIH Grant RR08299 and NSF Grant BIR-9221639 and the National Magnetic Resonance Facility at Madison with operating costs subsidized by NIH Grant RR02301 and equipment funded by the University of Wisconsin, NSF, NIH, and USDA.

[‡] The NMR structures and constraints have been deposited at the Protein Data Bank under accession number 1bm5.

* To whom correspondence should be addressed. Tel: 517-353-8786. Fax: 517-353-9334. E-mail: yanh@pilot.msu.edu.

[1] Abbreviations: 2D, two-dimensional; 3D, three-dimensional; CRABP, cellular retinoic acid binding protein; CRABPI, type-I cellular retinoic acid binding protein; CRABPII, type-II cellular retinoic acid binding protein; DQF-COSY, double quantum filtered correlation spectroscopy; HSQC, heteronuclear single quantum coherence; IPTG, isopropyl-1-thio-β-D-galactopyranoside; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; RA, *all-trans*-retinoic acid; RAR, retinoic acid receptor; RMSD, root-mean-square deviation; RXR, retinoid X receptor; TOCSY, total correlation spectroscopy.

S0006-2960(98)01021-6 CCC: $15.00
Published on Web 08/26/1998

snugly fixed at the entrance of the RA-binding pocket with only one edge of the ring accessible to the solvent. It appears that RA cannot enter or exit the deep binding pocket in the absence of major conformational changes in the protein.

The crystal structure of apo-CRABPI has also been determined (*9*). The ligand entrance of apo-CRABPI is slightly more open than that of holo-CRABPI and thus a bit more accessible to RA. We have determined recently the solution structure of apo-CRABPII [Wang, L., Li, Y., Abildgaard, F., Markley, J. L., and Yan, H. (1998) *Biochemistry* (in press)]. The ligand entrance of apo-CRABPII is greatly enlarged relative to that of holo-CRABPII and readily accessible to RA. The enlargement of the ligand entrance is mainly due to a concerted conformational change in three structural elements, namely the second helix, the βC-βD loop and the βE-βF loop. Furthermore, the ligand-binding pocket of apo-CRABPII exhibits dynamic disorder in solution.

To study the roles of Arg111 and Arg132 in binding of RA, we have made two site-directed mutants of CRABPII with the arginine residues replaced by methionine (R111M and R132M) (*10*). Competitive binding studies of the mutants have shown that both Arg111 and Arg132 are important for CRABPII to bind RA but Arg111 contributes more to the binding energy than Arg132. In this paper, we report the structure determination of the CRABPII mutant R111M in solution in the absence of RA by NMR spectroscopy. The results suggest that Arg111 plays a critical role in determining the structure and dynamical properties of CRABPII.

EXPERIMENTAL PROCEDURES

Sample Preparation. Site-directed mutagenesis, expression, and purification of the CRABPII mutant R111M have been previously described (*10*). Uniformly ^{15}N-labled R111M was obtained by growing the expression strain BL21-(DE3)/pLysS containing the plasmid pET-17b/R111M in M9 media with $^{15}NH_4Cl$ as the sole nitrogen source. For ^{15}N-labeling of leucine, isoleucine, and phenylalanine residues, a rich medium containing [^{15}N]leucine, [^{15}N]isoleucine, or [^{15}N]phenylalanine was used (*11*). The *Escherichia coli* strain DL49PS pLysS (kindly provided by Dr. David M. LeMaster) was the host for the selective isotopic labeling. For ^{15}N-labeling of lysine, the expression strain BL21(DE3)/pLysS was used. The strain was grown in M9 media supplemented with 110 mg/L of [^{15}N]lysine added 10 min before the induction by IPTG. D_2O samples were prepared by dissolving lyophilized unlabeled R111M in PBS buffer (20 mM sodium phosphate, 150 mM sodium chloride in 99.9% D_2O, pD 7.5, uncorrected). The protein solution was incubated for 6 h at 37 °C and lyophilized again. The lyophilized protein was then dissolved in 99.96% D_2O. H_2O samples were prepared by directly concentrating R111M fractions eluted with the PBS buffer (pH 7.3) from a gel-filtration column at the final step of the purification procedure (*10*). About 10% D_2O was added to the concentrated protein solutions. The protein concentrations of the NMR samples were ~2 mM.

NMR Spectroscopy. All NMR spectra were recorded at 27 °C. Homonuclear 2D spectra of unlabeled R111M in D_2O and heteronuclear 2D ^{1}H-^{15}N HMQC (*12*) spectra of selectively ^{15}N-labeled proteins in H_2O were acquired on a Varian VXR-500 NMR spectrometer. All 2D experiments were recorded in the phase-sensitive mode with F1 quadrature detection achieved by the hypercomplex method (*13*). The spectral width was 7238 Hz for ^{1}H and 2000 Hz for ^{15}N. The relaxation delay was 1.5 s, and the residual solvent resonance was suppressed by low-power presaturation. The following 2D homonuclear spectra were recorded: one DQF-COSY (*14*, *15*) spectrum, one E.COSY (*16*) spectrum, three NOESY (*17*, *18*) spectra with mixing times of 50, 100, and 150 ms, and three clean-TOCSY (*19*−*21*) spectra with mixing times of 20, 45, and 75 ms. The time domain data were composed of 2048 × 320 complex points for the E.COSY and DQF-COSY experiments and 2048 × 256 complex points for all other homonuclear 2D experiments. The 2D ^{1}H-^{15}N HMQC spectra were recorded with 1024 (^{1}H) × 64 (^{15}N) complex points. The data were processed with the program VNMR v. 5.1 (Varian Associates). A ~30°-shifted sine bell was applied in both dimensions for processing the DQF-COSY and E.COSY data, and a ~45°-shifted sine bell for processing the ^{1}H-^{15}N HMQC data. A Gaussian and minus LB combination was applied in F2 dimension and a ~75°-shifted sine bell in F1 dimension for processing the NOESY and TOCSY data. The time domain data were then zero-filled to 2048 × 512, 4096 × 2048, 4096 × 2048, and 8192 × 4096 real points for the 2D HMQC, TOCSY, NOESY, DQF-COSY, and E.COSY, respectively. A five-order polynomial was applied for baseline correction in the F2 dimension after Fourier transformation.

Heteronuclear NMR experiments of the uniformly ^{15}N-labeled protein were carried out on a Bruker DMX600 spectrometer equipped with a *z*-shielded gradient triple probe at The Ohio State University. 2D ^{1}H-^{15}N HSQC (*22*, *23*) spectrum was recorded with coherent selection by pulsed field gradients and sensitivity enhancement. The acquisition times and complex data points were ^{1}H (F2) 113.9 ms, 1024, and ^{15}N (F1) 56.9 ms, 256 (32 scans). 3D ^{1}H-^{15}N TOCSY-HMQC (*24*) and NOESY-HMQC (*24*, *25*) spectra were recorded with the following acquisition times and complex data points: ^{1}H (F3) 113.9 ms, 1024; ^{15}N (F2) 11.9 ms, 32; ^{1}H (F1) 14.3 ms, 128. The mixing time was 45 ms for the TOCSY-HMQC experiment and 150 ms for the NOESY-HMQC experiment. The ^{1}H isotropic mixing in the TOCSY-HMQC experiment was achieved by a DIPSI-3 sequence (*26*). In both 3D experiments, the solvent resonance was suppressed by low power presaturation. Quadrature detection in the indirectly detected dimensions was achieved by the States-TPPI method (*27*). A low power GARP-I sequence (*28*) was applied for ^{15}N broad-band decoupling during the data acquisition. The carriers for ^{1}H and ^{15}N were set at 4.70 and 115.7 ppm, respectively. The 2D ^{1}H-^{15}N HMQC-J (*29*) spectrum was recorded on a Bruker DMX750 spectrometer at the National Magnetic Resonance Facility at Madison with the following acquisition times and complex data points: ^{1}H 136.4 ms, 2048, and ^{15}N 232.7 ms, 512. The NMR data were processed with the program NMRPipe (*30*). A ~45°-shifted sine bell was applied in both F2 (^{1}H) and F1 (^{15}N) dimensions of the 2D data. They were zero-filled to 2048 (F2) × 512 (F1) and 2048 (F2) × 1024 (F1) real points for the ^{1}H-^{15}N HSQC and HMQC-J, respectively. A ~75°-shifted sine bell was applied in both F1 (^{1}H) and F3 (^{1}H) dimensions and a cosine bell in the F2 (^{15}N) dimension for processing the 3D spectra. The first points in the F1 and F2 dimensions of the 3D data were calculated by linear

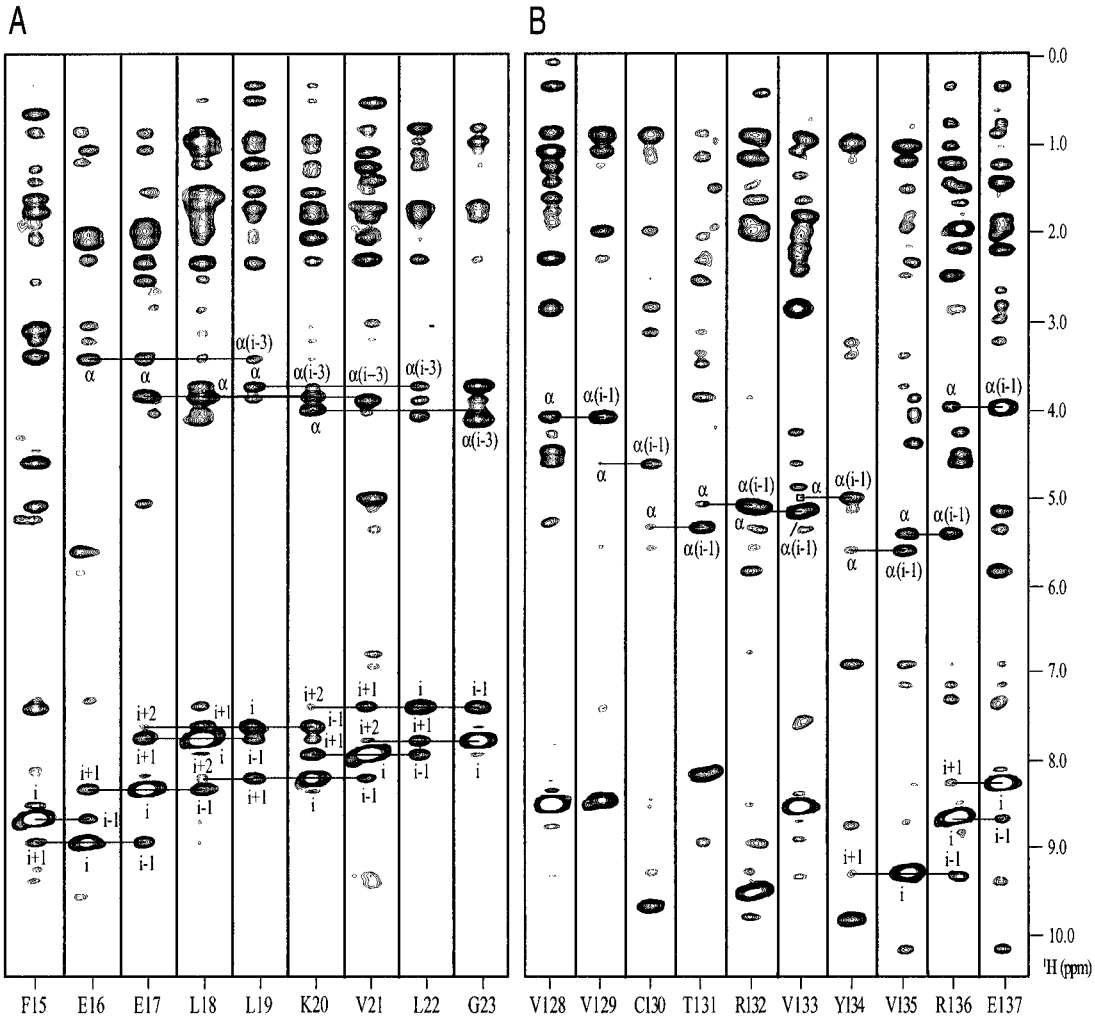

FIGURE 1: Strip plots extracted from the 3D ^{15}N-edited NOESY-HMQC spectrum of R111M showing the sequential and short-range NOE connectivities of α-helix A (A) and β-strand J (B).

prediction. Both 3D data were zero-filled to 2048 (F3) × 128 (F2) × 512 (F1) real points. The aliphatic part in the F3 dimension was removed after Fourier transformation.

Resonance Assignment. The sequential assignments were carried out following the standard procedure developed by Wüthrich and others (*31*). The spin systems were identified by the analysis of the homonuclear 2D TOCSY and DQF-COSY spectra and the 3D ^{1}H-^{15}N TOCSY-HMQC spectrum or by selective labeling. Sequential linking was obtained from NOE connectivities observed in 3D ^{1}H-^{15}N NOESY-HMQC experiment. The stereospecific assignments of β-mythylene protons were obtained by the joint analyses of the qualitative $^3J_{\alpha\beta}$ coupling constants derived from the E.COSY and DQF-COSY spectra and the relative volumes of the NOE cross-peaks between ^{1}H$_\alpha$-^{1}H$_\beta$ and ^{1}H$_N$-^{1}H$_\beta$ from the 2D homonuclear NOESY and 3D ^{1}H-^{15}N NOESY-HMQC spectra (*32*). The methyl protons of valine residue were stereospecifically assigned in a similar manner except that the relative volumes of the NOE cross-peaks between ^{1}H$_\alpha$-^{1}H$_\beta$ and ^{1}H$_N$-^{1}H$_\beta$ were used instead. Methyl protons of leucine residue were stereospecifically assigned after two rounds of structure refinement.

Derivation of Structural Restraints. Three types of structural restraints were derived from the experimental NMR data: interproton distance restraints, torsion angle restraints, and hydrogen bond restraints. Approximate interproton distance restraints were derived from the homonuclear 2D NOESY with 100 and 150 ms mixing times and 3D ^{15}N-edited NOESY-HMQC with 150 ms mixing time. The NOE cross-peaks of the homonuclear 2D NOESY and the 3D ^{15}N-edited NOESY-HMQC were picked by the programs VNMR (Varian Associates) and PIPP (*33*), respectively. The NOE intensities were converted to approximate interproton distances by normalization against the NOE cross-peaks between the backbone protons of the identified β-sheets. The NOEs were then classified as strong, medium, weak, and very weak, corresponding to interproton distances 1.8−2.7 Å (1.8−2.9 Å for NOE distances involving NH protons), 1.8−3.3 Å (1.8−3.5 Å for NOE distances involving NH protons), 1.8−5.0 Å, and 2.5−5.0 Å, respectively. Pseudoatom corrections were made for nonstereospecifically assigned methylene protons and methyl protons (*34*). An additional 0.5 Å was added to the upper bounds for methyl protons. The backbone torsion angle ϕ was restrained according to the values of $^3J_{NH\alpha}$ obtained from the 2D HMQC-J spectrum. It was restrained to $-120 \pm 40°$ if $^3J_{NH\alpha} \geq 7.5$ Hz and to $-60 \pm 20°$ for $^3J_{NH\alpha} \leq 5.0$ Hz and when the NOE data indicated the presence of α-helix. Backbone hydrogen bonds were derived from patterns of cross-peaks in the fingerprint regions of the homonuclear 2D TOCSY and NOESY spectra recorded in D_2O. Hydrogen bond restraints were set to 2.0 ± 0.5 Å for H^N···O and 3.0 ± 0.3 Å for N···O distances.

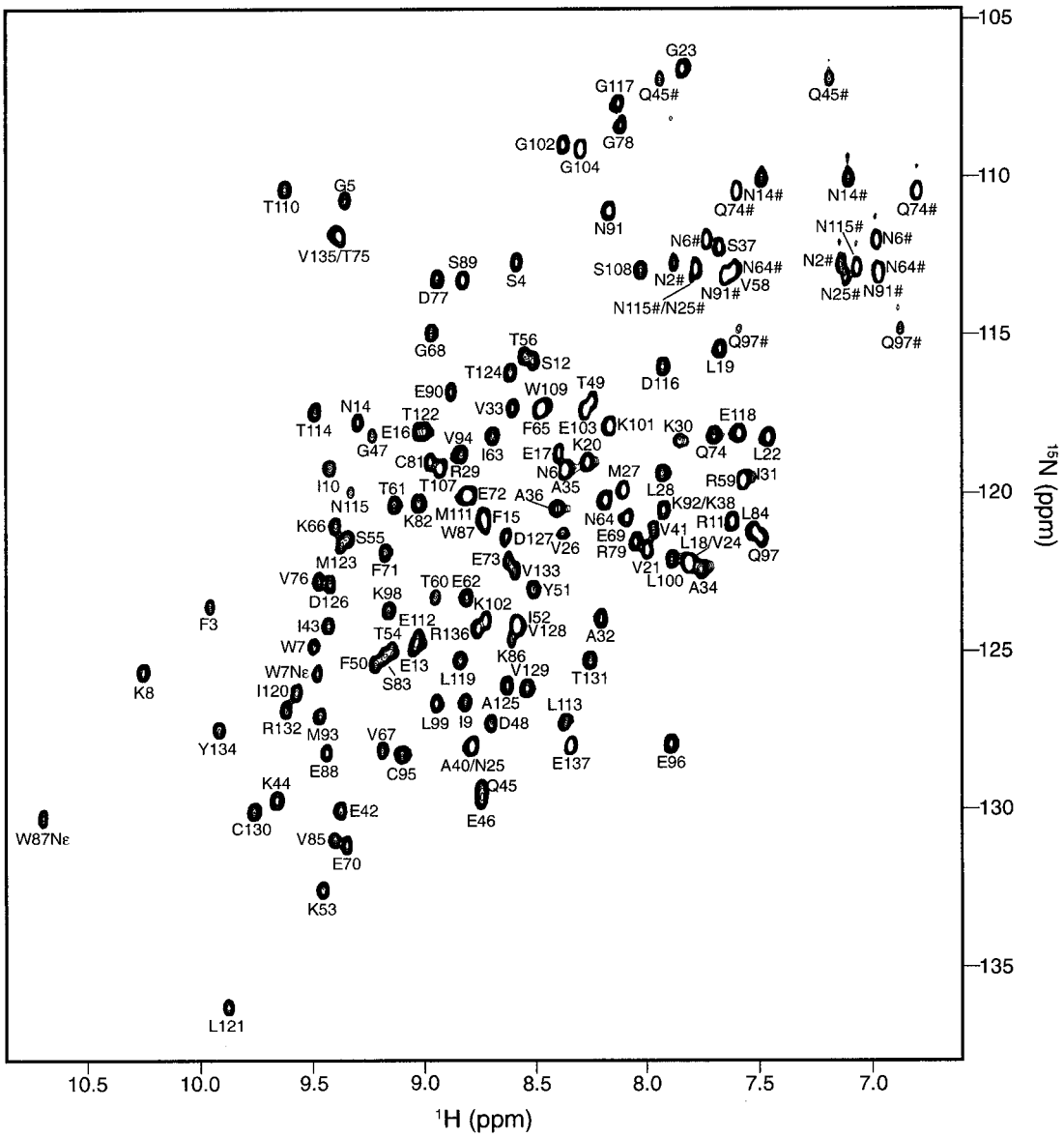

FIGURE 2: 2D ^{1}H-^{15}N HSQC spectrum of uniformly ^{15}N-labeled apo-R111M recorded at 600 MHz ^{1}H frequency with coherence selection by pulsed field gradients and sensitivity enhancement. Sequential assignments are indicated with one-letter amino acid code and residue number. The side-chain amides of Asn and Gln are indicated (#).

Structure Calculation. The structures were calculated with a hybrid distance geometry-simulated annealing (DGSA) protocol (*35*) using the program X-PLOR (version 3.1) (*36*) on a SGI Indigo II workstation. A square-well potential function with a force constant of 50 kcal mol^{-1} Å^{-2} or 200 kcal mol^{-1} rad^{-2} was applied for the NMR-drived restraints. The X-PLOR f_{repel} function was used to simulate van der Waals interactions with atomic radii set to 0.80 times their CHARMM values (*37*). Hydrogen bond restraints within the regions of regular secondary structures were introduced at a later stage of structural refinement. A total of 70 structures were generated using this protocol. The structures were inspected by the programs QUANTA96 (Molecular Simulations), InsightII (Biosym Technologies) and analyzed by PROCHECK-NMR (version 3.4.4) (*38*, *39*). An iterative strategy was used for the structure refinement. In each round of structure refinement, newly computed NMR structures were employed to assign more NOE restraints, to correct wrong assignments and to loosen the NOE distance bounds if spectral overlapping was deduced. Then another round of structure refinement was carried out with the modified NMR restraints. After several rounds of such refinement, an ensemble of 28 structures was selected according to their best fit to the experimental NMR restraints and the low values of their total energies.

RESULTS AND DISCUSSION

Spin System Identification and Sequential Assignment. Spin systems were first identified by the analysis of the homonuclear 2D DQF-COSY and TOCSY spectra. Most of the threonine, alanine, valine, and glycine residues could be identified easily. The majority of the residues belonging to AMX or AM(PT)X spin system could also be identified. Except for a few, the residues with long side chains (lysine, leucine, isoleucine, arginine and proline) could not be distinguished from each other because of the severe spectral overlapping at the upfield region (1−2 ppm) and the inefficiency of TOCSY experiments for long-relayed coherence transfers. The aliphatic spin systems were then extended to amide resonances by the 3D ^{1}H-^{15}N TOCSY-

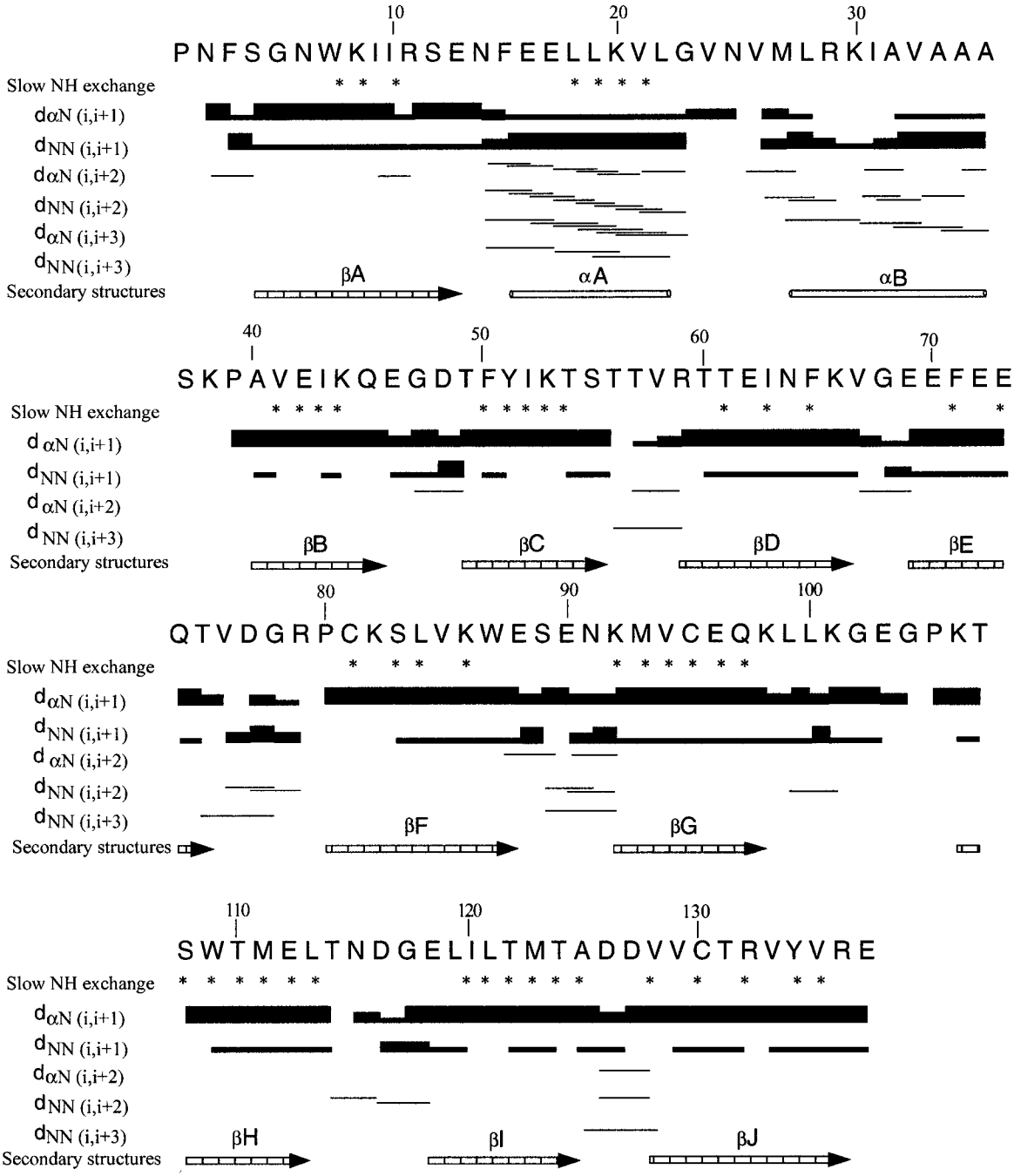

FIGURE 3: Summary of the sequential and medium-range NOEs involving backbone ^1H_N and $^1H_\alpha$ atoms, slow-exchange backbone amide protons and the deduced secondary structures of apo-R111M. Line thickness for the $d_{\alpha N}$ and d_{NN} sequential NOE distances reflects the intensities of the cross-peaks. Asterisks indicate the residues whose amide protons remained after 24 h of exchange with D_2O in PBS buffer at room temperature and pD 7.5 as determined by 2D homonuclear NMR experiments.

HMQC experiment. Leucine, isoleucine, lysine, and phenylalanine residues were identified by selective ^{15}N labeling. The aromatic ring protons were identified in the 2D DQF-COSY and TOCSY spectra and linked to the aliphatic protons through intraresidue NOEs.

The sequential assignments were primarily based on the sequential NOE connectivities such as $d_{\alpha N}$, d_{NN} and $d_{\beta N}$ (Figure 1). The assignments were initiated after over 60% of the total 137 residues had been identified. Most residues from the 10 β-strands and the first α-helix could be easily assigned (Figures 2 and 3). The sequential assignments of Val26−Lys38 posed a challenge. Val26, Arg29, Ile31, and Lys38 were very weak in the 2D ^1H-^{15}N HSQC spectrum. No correlation between their amide and aliphatic protons could be observed in the 3D ^1H-^{15}N TOCSY-HMQC spectrum. Only a few NOEs useful for their sequential assignments could be obtained from the 3D ^1H-^{15}N NOESY-HMQC spectrum. Moreover, the ^1H chemical shifts of the four alanines (Ala32, Ala34, Ala35, and Ala36) were very similar. The ^1H and ^{15}N resonances of Val26 and Arg29 were tentatively assigned based on a few weak sequential NOEs and the comparisons with the sequential assignments of wild type holo-CRABPII (L.W. et al., unpublished material). Both residues showed stronger cross-peaks after RA binding. The assignments of Ile31 and Lys38 were helped greatly by selective labeling and the sequential

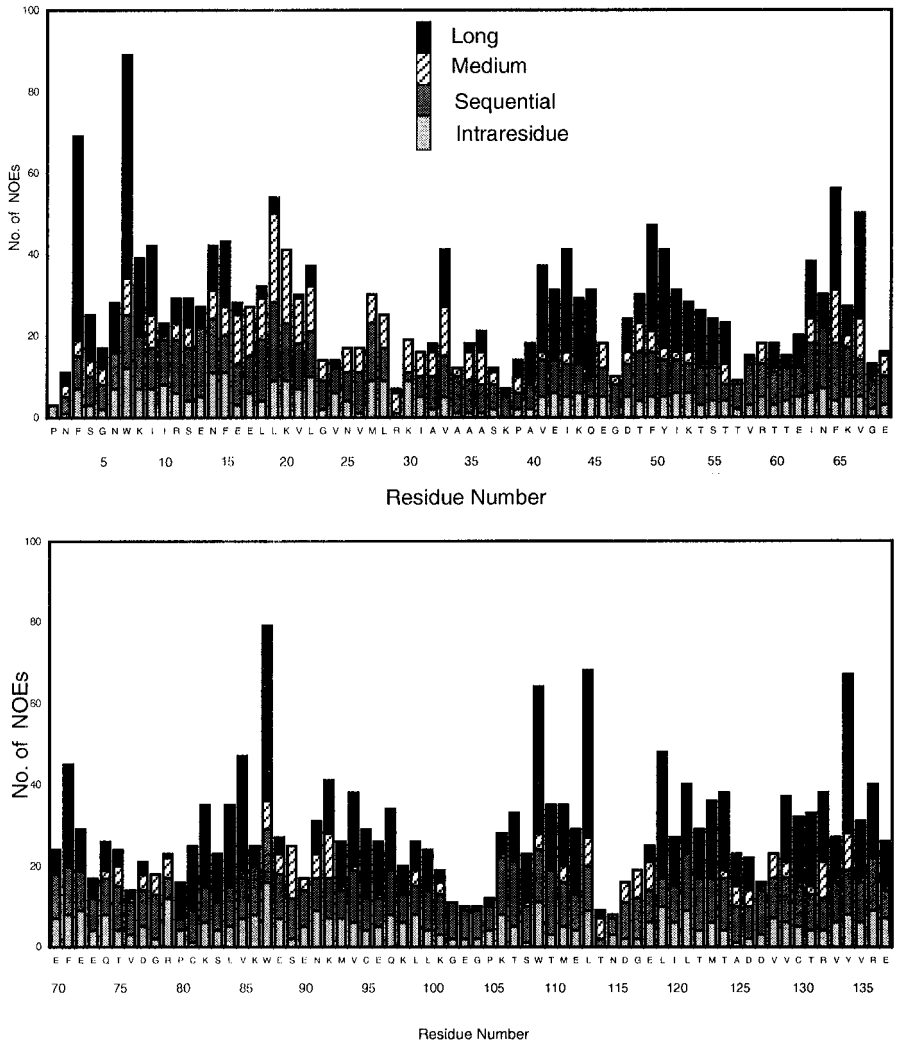

FIGURE 4: The distribution of NOEs along the amino acid sequence of R111M. For intraresidue NOEs, only those that are structurally useful are included.

assignments of the wild type holo-CRABPII. The spin system of Asn115 could be easily identified by the homonuclear NMR experiments, but its sequential assignment was achieved only at a later stage because it had weak peaks in all ^{15}N-edited spectra. The chemical shifts of the α- and β-protons of Thr114 were very close to the water signal and they were not observed in the 3D ^{15}N-edited spectra because of presaturation. Thr114 was assigned by the analysis of the 2D homonuclear spectra recorded in D_2O and the resonance assignments of the wild-type holo-CRABPII. Asn2 and Thr57 were the only two residues that were not observed in the 2D ^{1}H-^{15}N HSQC spectrum. The aliphatic protons of the two residues were assigned by spin system identification and NOEs at a later stage of the sequential assignments. The side-chain assignments were completed except for Arg11, Leu22, Arg29, Ile31, Arg59, and eight lysine residues. Stereospecific assignments were made for the β-methylene protons of 51 residues and the methyl protons of 10 valine and six leucine residues. The chemical shifts of all the assigned ^{1}H and^{15}N resonances are listed in Table S1 of the Supporting Information.

Secondary Structure Determination. The sequential resonance assignment led directly to the determination of the secondary structures of apo-R111M based on their characteristic NOE patterns (Figure 3). Except the N-terminal half (Leu28−Ile31) of the second α-helix, all the secondary structure elements had typical NOE patterns. The strength and number of the NOEs observed for residues Leu28−Ile31were well below what is expected for a regular α-helix. However, the chemical shifts of these residues indicated that they are predominately in a helical conformation. Long-range NOEs were observed between the backbone protons of all adjacent β-strands except between βD and βE. There were also NOEs between the backbone protons of the N-terminal half of βA and the C-terminal half of βJ.

Tertiary Structure Determination. To determine the solution structure of apo-R111M, a total of 2302 structurally useful distance restraints were obtained by the analysis of the homonuclear 2D NOESY and ^{15}N-edited 3D NOESY spectra. Of the 2302 distance restraints, 711 were derived from intraresidue NOEs, 635 sequential NOEs, 252 medium-range NOEs, and 704 long-range NOEs. The distance restraints were quite evenly distributed along the amino acid sequence except for the two N-terminal residues, the residues in several turns and in the second α-helix (Figure 4). The numbers of distance restraints for these residues were well below the average. A total of 77 ϕ dihedral angle restraints were obtained from the 2D HMQC-J data. A total of 98

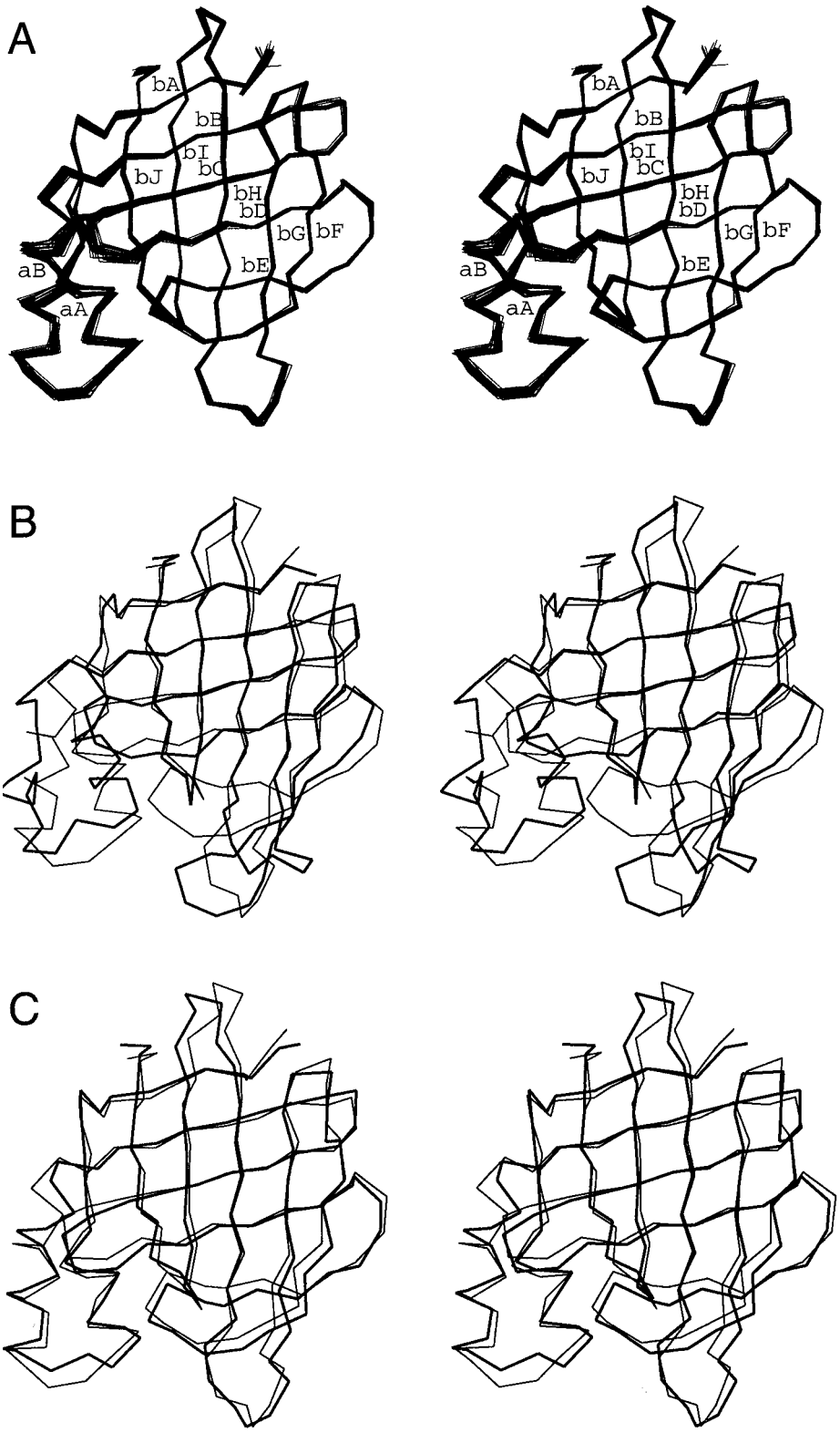

FIGURE 5: (A) Stereoview of the C_α traces of the superimposed 28 refined solution structures of apo-R111M. (B) Stereoview of the superimposed C_α traces of the restrained minimized mean structures of apo-R111M (thin line) and wild-type apo-CRABAPII (thick line). (C) Stereoview of the C_α trace of the restrained minimized mean solution structure of apo-R111M (thin line) superimposed with that of the crystal structure of wild-type holo-CRABAPII (thick line).

hydrogen bond restraints from 49 hydrogen bonds was included in the structural refinement. Figure 5A shows the superposition of 28 refined structures with no NOE violations exceeding 0.25 Å and no dihedral angle violation larger than 5.0°. The statistics of the NMR structures is shown in Table 1. The stereochemical qualities of the structures were examined by the program PROCHECK-NMR and were comparable to those X-ray structures with 2.5 Å resolution and *R*-factors less than 20%.

Recently, we have determined the crystal structure of apo-R111M (*40*). The crystal structure contains two independent molecules, Mol A and Mol B. The most striking difference

Table 1: Restraint and Structural Statistics of Apo-R111M

Restraint Statistics		
no. of experimental NOE restraints		
intraresidue	711	
sequential	635	
medium-range	252	
long-range	704	
total	2302	
no. of experimental ϕ dihedral restraints	77	
no. of hydrogen bonds	98	
Structural Statistics		
	$\langle SA\rangle^a$	$\langle SA\rangle_r^b$
RMSD from experimental distance restraints (Å) (2302)	0.032 ± 0.001	0.024
RMSD from experimental dihedral restraints (deg) (77)	0.926 ± 0.102	0.836
deviation from idealized covalent geometry		
bonds (Å)	0.004 ± 0.001	0.003
angles (deg)	0.621 ± 0.028	0.520
impropers (deg)	0.495 ± 0.004	0.411
Measures of Structure Quality (by Procheck)		
Ramachandran plot		
% residues in most favorable regions	71.8 ± 2.4	71.0
% residues in additional allowed regions	25.8 ± 0.8	25.8
no. of bad contacts	6 ± 2	6
hydrogen bond energy	0.40 ± 0.10	0.50
overall *G*-factor	−0.19 ± 0.01	−0.20
Coordinate Precision		
RMSD for C_α trace (Å)	0.43 ± 0.14	
RMSD for backbone atoms (Å)	0.54 ± 0.26	
RMSD for all heavy atoms (Å)	0.98 ± 0.23	

[a] $\langle SA\rangle$ are the final 28 simulated annealing structures. [b] $\langle SA\rangle_r$ is the restrained minimized mean structure obtained by restrained regularization of the mean structure, which is obtained by averaging the coordinates of the individual SA structures best fitted to each other.

between Mol A and Mol B occurs at the second α-helix, with the N-terminus of the second α-helix partially unwound in Mol A. The NMR solution structure is highly similar to the X-ray crystal structure, and the restrained minimized mean NMR structure can be superimposed with both Mol A and Mol B with a RMSD of 1.6 Å for all C_α atoms. The main difference between the NMR and crystal structures also occurs at the second α-helix. It is likely that the second α-helix undergoes conformational exchange as detailed in a late section and the NMR structure represents an average conformation.

Comparison of the Solution Structures of apo-R111M and Wild-Type apo-CRABPII. The solution structure of apo-R111M is similar to that of apo-CRABPII in backbone folding [Wang, L., Li. Y., Abildgaard, F., Markley, J. L., and Yan, H. (1998) *Biochemistry* (in press)]. The restrained minimized mean NMR structures of the two proteins could be superimposed with a RMSD of 3.2 Å for all C_α atoms (Figures 5B and 6B). However, there are significant conformational differences between the two structures, mainly localized to three segments (Leu19−Ala36, Glu73−Cys81, and Leu99−Pro105) clustered around the ligand entrance. Surprisingly, the conformational differences occur in the region far away from the site of the point mutation. If these segments are excluded, the C_α RMSD between the two structures is reduced to 1.6 Å. The first segment (Leu19−Ala36) encompasses the middle of the first helix to the end of second helix, the second segment (Glu73−Cys81) the end of βE and the beginning of the βF, and the third segment (Leu99−Pro105) the long βG-βH loop. In apo-R111M, all three segments move toward the center of the ligand entrance so that the opening of the ligand-binding pocket in apo-R111M is much smaller than that in wild-type apo-CRABPII. For example, the distance between the C_α atoms of Val58 in the βC-βD loop and Val24 in the turn connecting the two α-helices is 7.4 Å shorter in apo-R111M than in apo-CRABPII. The distance between the C_α atoms of Val58 and Asp77 in the βE-βF loop is 10.8 Å shorter in apo-R111M than in apo-CRABPII.

Comparison of the Solution Structure of apo-R111M and the Crystal Structure of Wild-Type holo-CRABPII. Figure 5C shows the superposition of the restrained minimized mean NMR structure of apo-R111M and the crystal structure of wild-type holo-CRABPII. The RMSD between the two superimposed structures is 0.96 Å for the C_α atoms of all regular secondary structures and 1.38 Å for all C_α atoms (Figure 6A). It is obvious from examining Figure 6 that the solution structure of apo-R111M is more similar to the crystal structure of holo-CRABPII than to the solution structure of apo-CRABPII. The ligand entrance of apo-R111M is much smaller than that of apo-CRABPII, but it is still larger than that of holo-CRABPII. The distance between the C_α atoms of Val58 in the βC-βD loop and Val24 in the turn connecting the two α-helices is 15.0, 21.3, and 28.8 Å in holo-CRABPII, apo-R111M, and apo-CRABPII, respectively. The distance between the C_α atoms of Val58 and Asp77 in the βE-βF loop is 12.4, 13.4, and 24.2 Å in holo-CRABPII, apo-R111M, and apo-CRABPII, respectively. It appears that the mutation and the binding of RA causes similar conformational changes, although the magnitudes of the conformational changes caused by the mutation are not as larger as those induced by the binding of RA.

Dynamical Properties of apo-R111M. More than 50% backbone amide protons of apo-R111M remained after a day of exchange with D_2O (Figures 3 and 7), indicating that the half-lives of these amide protons with respect to exchange with D_2O are at least 10 h. These slow-exchanging amide protons were distributed in the first α-helix and all strands of the two β-sheets. The intrinsic exchange rates were calculated based on the R111M sequence using the formula and parameters developed and calibrated by Englander and colleagues (*41*) with random coil peptides, resulting in protection factors (the ratios of intrinsic and observed exchange rates) for these amide protons on the order of 10^6. Generally, hydrogen-exchange rates are correlated with hydrogen bonding and/or accessibility to solvent. In order for the exchange to occur, amide protons must be accessible to solvent. If the amide protons are buried, transient opening of solvent channels by structural fluctuation is required for the exchange. If the amide protons are involved in hydrogen bonding, transient breakup of the hydrogen bonds by structural fluctuation or local unfolding is also thought to be required for the exchange. Slow exchange usually indicates that the amide proton is buried and/or involved in hydrogen bonding. In addition, it also suggests that the amide group undergo less structural fluctuation. On examination of the NMR (this work) and crystal (*40*) structures of apo-R111M, we found that all the slow-exchanging amide protons, except those of Val85, Lys92, Gln97, and Val128, are involved in hydrogen bonding. Those slow-exchanging amide protons that are not involved in hydrogen bonding

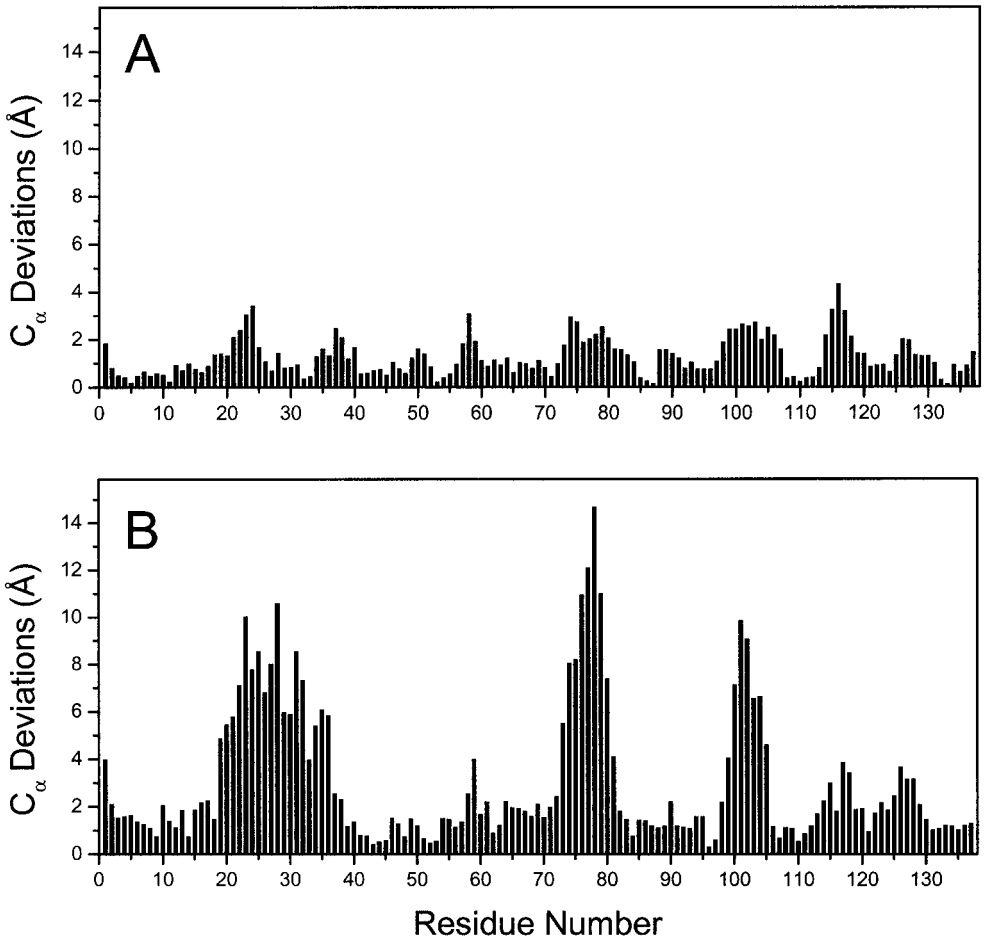

FIGURE 6: C_α atom deviations between apo-R111M and wild-type holo-CRABPII (A) and between apo-R111M and wild-type apo-CRABPII (B).

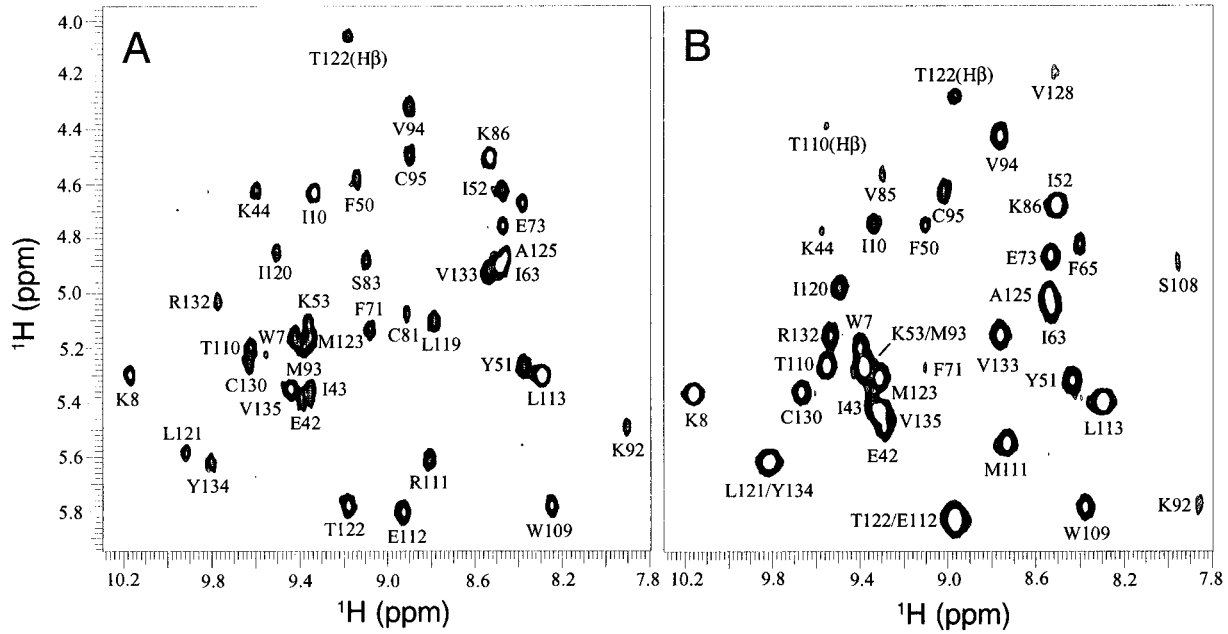

FIGURE 7: The finger regions of the homonuclear 2D TOCSY spectra of apo-CRABPII (A) and apo-R111M (B) recorded with 40 ms mixing time. The TOCSY experiment was initiated more than 24 h after dissolving a lyophilized protein sample in PBS buffer, pD 7.5, in D_2O.

are all inaccessible to solvent. The indole ϵ_1-proton of Trp7 was also detected in homonuclear 2D spectra recorded in D_2O. The hydroxyl protons of Thr107 and Ser83 were observed in 2D NMR spectra recorded in H_2O. All the observations suggested that R111M is for the most part fairly stable and rigid.

The second α-helix, especially the N-terminal part of it, appeared to be flexible. While all other secondary structure

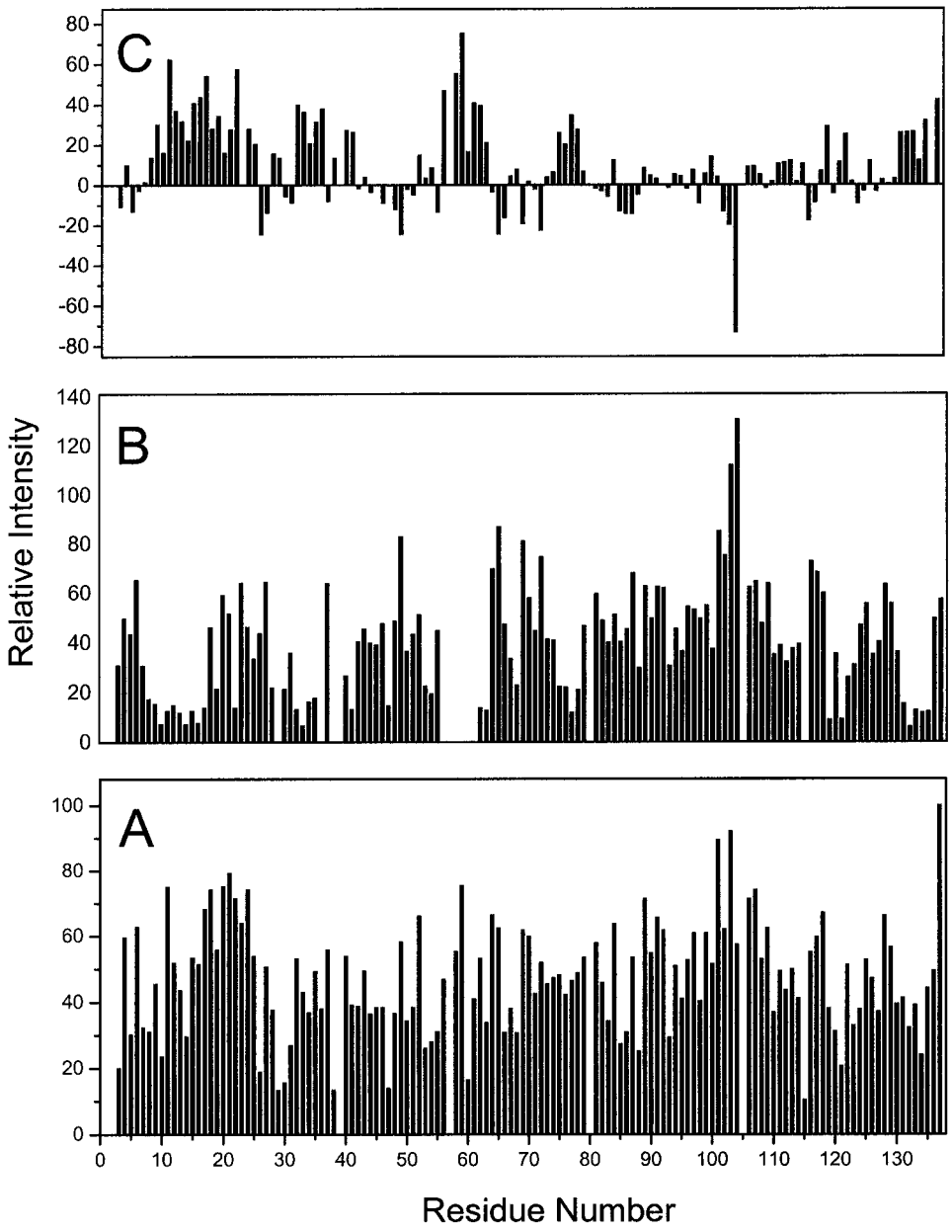

FIGURE 8: Relative peak intensities as measured by normalized peak heights of the 2D ^{1}H-^{15}N HSQC spectra of apo-R111M (A) and wild-type apo-CRABPII (B) as a function of residue number. Spectrum C was obtained by subtracting spectrum B from spectrum A.

elements had slow-exchanging amide protons, all amide protons of the second helix underwent fast exchange with D_2O. Furthermore, the N-terminal half (Leu28−Ile31) of the second α-helix had much less characteristic NOE indicative of a helical conformation. Residues Arg29−Ile31 also had very weak peaks in 2D ^{1}H-^{15}N HSQC spectrum (Figures 2 and 8A), indicating that the N-terminal part of the second helix experienced significant conformational exchange as discussed below.

Did the Mutation Change the Dynamical Properties of CRABPII? For the most part, apo-R111M and wild-type apo-CRABPII had similar dynamical properties. Thus, both proteins had about the same number of slow-exchanging backbone amide protons. Furthermore, the identities of most of the amide protons were the same for the two proteins (Figure 7). However, the ligand entrance of wild-type apo-CRABPII appeared to be much more flexible than that of apo-R111M. In particular, the second α-helix was apparently disordered in apo-CRABPII but fairly ordered, albeit still flexible, in apo-R111M.

To compare the dynamical properties of the two proteins in more detail, we examined the intensities of the peaks (as measured by peak heights) in their 2D ^{1}H-^{15}N HSQC spectra. In general, the peak intensities of the HSQC spectra are modulated by amide proton exchange and conformational exchange. But amide proton exchange appeared to have little effects on the peak intensities of the HSQC spectra of the two proteins, partly owing to a "water flip-back" pulse sequence for the HSQC experiments that minimized water saturation and dephasing (*42*, *43*). In fact, the average peak intensity of the slow-exchanging amides was the same as that of all amides. Thus, variation in the line widths of amide resonances caused by conformational exchange was the major factor in determining the peak intensities of the two HSQC spectra. For apo-R111M, the intensities of the peaks in the 2D ^{1}H-^{15}N HSQC spectrum (Figure 8A) were fairly

uniform. Of all secondary structure elements, only three residues (Arg29−Ile31) at the flexible N-terminal half of the second helix showed weak peaks. The rest of the weak peaks invariably belonged to the residues at the loops or turns.

Wild-type apo-CRABPII displayed large variations in the peak intensities (Figure 8B). Most residues at the RA-binding pocket and near the ligand entrance showed weak or no peaks in the 2D ^{1}H-^{15}N HSQC spectrum, indicating that the RA-binding pocket, especially the ligand entrance, is rather flexible. Most of these peaks became much higher after the mutation (Figure 8C). In fact, most of the residues with significant intensity differences between the two spectra were clustered in the ligand-binding pocket. Of the 21 residues that were in contact with RA as revealed by the crystal structure of holo-CRABPII (*8*), all but three residues (Ile31, Thr54, and Met123) showed significantly increased peak intensities in the HSQC spectrum of apo-R111M.

The changes in the dynamical properties caused by the mutation are very similar to those induced by binding of RA. Gierasch and co-workers compared the hydrogen/deuterium exchange rates of apo- and holo-CRABPI (*44*). The results showed that binding of RA reduces structural fluctuation at the ligand entrance of CRABPI, particularly the second α-helix and the βC-βD loop. Our NMR study of holo-CRABPII showed that binding of RA also reduces the flexibility of the ligand-binding pocket of CRABPII, especially the ligand entrance (L.W, and H.Y., unpublished material). Reduction of backbone mobility was also observed at the ligand entrance of a homologous fatty acid binding protein upon binding of the ligand (*45*).

Roles of Arg111. Arg111 interacts with the carboxyl group of RA via hydrogen bonds mediated by a water molecule (*8*). RA binds to CRABPII with a binding energy of 11.9 kcal/mol (K_d = 1.9 nM) (*10*). Substitution of Arg111 with methionine in CRABPII decreases the affinity for RA by ∼45-fold (2.2 kcal/mol), indicating that Arg111 is indeed important for binding of RA (*10*). X-ray crystallographic and molecular modeling studies suggest that Arg111 is one of four arginine residues that generate a positive potential that may facilitate the binding of RA (*40*, *46*). The NMR results presented in this paper suggest an additional role for Arg111 related to the structure and dynamics of CRABPII.

As mentioned earlier, RA binding induces significant but localized conformational and dynamical changes in both CRABPI and CRABPII (*8*, *9*, *40*, *44*). The ligand entrance, in particular, is much more open and flexible in the apo-forms than in the holo-forms. The large opening of the ligand entrance and its flexibility is probably essential for the unhindered entry of RA to the deep ligand-binding pocket and the subsequent closure of the ligand entrance. Because apo-CRABPI was crystallized as a dimeric protein, dimerization has been suggested to be the mechanism that allows RA to enter or exit the RA-binding pocket (*9*). However, biochemical and NMR evidence indicates that apo-CRABPII is predominately monomeric in solution [Wang, L., Li. Y., Abildgaard, F., Markley, J. L., and Yan, H. (1998) *Biochemistry* (in press)].

Substitution of Arg111 by methione causes significant but localized changes in the conformation and dynamical properties of CRABPII as shown by this NMR study. Strikingly, apo-R111M is more similar to wild-type holo-CRABPII than to wild-type apo-CRABPII in both structure and dynamical properties. The mutation also apparently increases the stability of the protein. Thus, the NMR samples of R111M last for several months without deterioration of the spectral properties. In contrast, the NMR samples of apo-CRABPII last only for about 2 weeks. This observation is consistent with the results of the site-directed mutagenesis study of CRABPI (*46*). Substitution of Arg111 of CRABPI with glutamine increases the stability of the protein by 3.0 kcal/mol, whereas binding of RA increases the stability by 5.7 kcal/mol. The changes are rather significant considering that the free energy change of unfolding apo-CRABPI is only 6.2 kcal/mol (*47*). These results indicate that RA binding and the Arg111 mutations have similar effects on the structure, dynamics, and stability of CRABPS, although the effects caused by RA binding are greater than those caused by the mutations. The results also suggest that Arg111 may play a major role in maintaining the ligand entrance in an open and dynamic state for the entry of RA.

The next question is how the mutation R111M ∼17 Å away from the ligand entrance affects its structure and dynamics. It has been suggested that the repulsive interaction between the positively charged Arg111 and Arg131 (corresponding to Arg132 in CRABPII) provide the driving force for the conformational flexibility of apo-CRABPI (*46*). The guanidinium groups of the two residues are ∼7.6 Å apart in apo-CRABPII [Wang, L., Li, Y., Abildgaard, F., Markley, J. L., and Yan, H. (1998) *Biochemistry* (in press)]. Our site-directed mutagenesis study has shown that substitution of Arg132 by methionine stabilizes CRABPII to some extent but much less than substitution of Arg111 by methionine (*10*). The NMR spectral quality of apo-R132M is better than that of apo-CRABPII but worse than that of apo-R111M, suggesting that apo-R132M is less flexible than apo-CRABPII but not as rigid as apo-R111M. Thus, the interaction between the two arginine residues may be only partially responsible for the conformational flexibility of apo-CRABPII.

The different dynamical behaviors of R111M and R132M may reflect different environments of the two arginine residues. The solution structure of apo-CRABPII reveals that Arg111 is situated in the small hydrophobic core of CRABPII with its guanidinium group surrounded by the side chains of mostly hydrophobic residues, Val41, Ile43, Ile52, Thr54, Ile63, Trp109, and Leu121. The only favorable interactions may be between the guanidinium group of Arg111 and the aromatic ring of Trp109 and between the guanidinium group of Arg111 and the hydroxyl group of Thr54. On the other hand, the guanidinium group of Arg132 is surrounded by the side chains of mostly hydrophilic residues, Ser12, Pro39, Thr54 (γ-methyl group), Cys130, and Tyr134. The hydroxyl groups of Ser12 and Tyr134 and the thiol group of Cys130 provide the favorable polar interactions with the guanidinium group of Arg132. In particular, the hydroxyl group of the phenolic ring of Tyr134 is within the hydrogen bond distance to the NE nitrogen of the guanidinium group of Arg132. Furthermore, the guanidinium group of Arg111 is almost completely buried, whereas that of Arg132 is partially exposed. Two water molecules are found within the hydrogen bond distance of the guanidinium group of Arg132 in the crystal structure of apo-R111M (*40*), although there is no evidence at present for water molecules

bound to either of the two guanidinium groups in apo-CRABPII. It is tempting to speculate that the charged guanidinium group of Arg111 is mobile because of its hydrophobic environment and the repulsive interaction with Arg132. The mobility of guanidinium group of Arg111 may propagate to the ligand entrance, resulting in a more open and flexible conformation. Binding of RA neutralizes, at least partially, the positive charge of Arg111 and, therefore, reduces the mobility of its guanidinium group. CRABPS are most stable in the holo-forms because there are additional hydrophobic interactions between the proteins and RA. It is noted that the guanidinium group of Arg111 is more than 17 Å away from the ligand entrance, the most flexible region of apo-CRABPII. It is not clear how the mobility of the guanidinium group propagates to the ligand entrance.

ACKNOWLEDGMENT

We thank Dr. Frits Abildgaard for recording the 2D ^{1}H-^{15}N HMQC-J spectrum and Dr. Yue Li for critical reading of the manuscript.

SUPPORTING INFORMATION AVAILABLE

A table listing the ^{1}H and^{15}N resonance assignments of apo-R111M at pH 7.3 and 27 °C (3 pages). Ordering information is given on any current masthead page.

REFERENCES

1. Sporn, M. B., Roberts, A. B., and Goodman, D. S. (1994) in *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., and Goodman, D. S., Eds.) Raven Press, New York.
2. Mangelsdorf, D. J., Umesono, K., and Evans, R. M. (1994) in *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., and Goodman, D. S., Eds.) pp 319−349, Raven Press, New York.
3. Ong, D. E., Newcomer, M. E., and Chytil, F. (1994) in *The Retinoids: Biology, Chemistry, and Medicine* (Sporn, M. B., Roberts, A. B., and Goodman, D. S., Eds.) Raven Press, New York.
4. Napoli, J. L. (1996) *FASEB J. 10*, 993−1001.
5. Bucco, R. A., Zheng, W. L., Davis, J. T., Sierra Rivera, E., Osteen, K. G., Chaudhary, A. K., and Ong, D. E. (1997) *Biochemistry 36*, 4009−4014.
6. Jing, Y., Waxman, S., and Mira-y-Lopez, R. (1997) *Cancer Res. 57*, 1668−1672.
7. Boylan, J. F., and Gudas, L. J. (1991) *J. Cell Biol. 112*, 965−979.
8. Kleywegt, G. J., Bergfors, T., Senn, H., Le Motte, P., Gsell, B., Shudo, K., and Jones, T. A. (1994) *Structure 2*, 1241−1258.
9. Thompson, J. R., Bratt, J. M., and Banaszak, L. J. (1995) *J. Mol. Biol. 252*, 433−446.
10. Wang, L., Li, Y., and Yan, H. (1997) *J. Biol. Chem. 272*, 1541−1547.
11. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahluist, F. W. (1989) *Methods Enzymol. 177*, 44−73.
12. Mueller, L. (1979) *J. Am. Soc. Chem. 101*, 4481−4484.
13. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
14. Piantini, U., Sørensen, O. W., and Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800−6801.
15. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
16. Griesinger, C., Søresen, O. W., and Ernst, R. R. (1985) *J. Am. Soc. Chem. 107*, 6394−6396.
17. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
18. Macura, S., and Ernst, R. R. (1980) *Mol. Phys. 41*, 95−117.
19. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
20. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
21. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. P. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
22. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−188.
23. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
24. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry 28*, 6150−6156.
25. Fesik, S. W., and Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 78*, 588−593.
26. Shaka, A. J., Lee, C. J., and Pines, A. (1988) *J. Magn. Reson. 77*, 274−293.
27. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
28. Shaka, A. J., Barker, P. B., and Freeman, R. (1985) *J. Magn. Reson. 64*, 547−552.
29. Kay, L. E., and Bax, A. (1990) *J. Magn. Reson. 86*, 120−126.
30. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−93.
31. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
32. Basus, V. J. (1989) *Methods Enzymol. 177*, 132−149.
33. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Reson. 95*, 214−220.
34. Wüthrich, K., Billeter, M., and Braun, W. (1983) *J. Mol. Biol. 169*, 949−61.
35. Nilges, M., Gronenborn, A. M., and Clore, G. M. (1988) *FEBS Lett. 229*, 317−324.
36. Brünger, A. T. (1992) *X-PLOR Version 3.1: A System for Crystallography and NMR*, Yale Univeristy, New Haven, CT.
37. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
38. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
39. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1997) *J. Biomol. NMR 8*, 477−486.
40. Chen, X., Tordova, M., Gilliland, G. L., Wang, L., Li, Y., Yan, H., and Ji, X. (1998) *J. Mol. Biol. 278*, 641−653.
41. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins 17*, 75−86.
42. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
43. Zhang, O., Kay, L. E., Olivier, J. P., and Forman Kay, J. D. (1994) *J Biomol NMR 4*, 845−858.
44. Rizo, J., Liu, Z. P., and Gierasch, L. M. (1994) *J. Biomol. NMR 4*, 741−760.
45. Hodsdon, M. E., and Cistola, D. P. (1997) *Biochemistry 36*, 2278−2290.
46. Zhang, J., Liu, Z. P., Jones, T. A., Gierasch, L. M., and Sambrook, J. F. (1992) *Proteins 13*, 87−99.
47. Liu, Z.-P. (1992) Ph.D. Dissertation, University of Texas Southwestern Medical Center at Dallas, Dallas, TX.

BI981021X