Structure and Intramodular Dynamics of the Amino-Terminal LIM Domain from Quail Cysteine- and Glycine-Rich Protein CRP2[†,‡]

Georg Kontaxis,[§] Robert Konrat,*[,§] Bernhard Kräutler,[§] Ralf Weiskirchen,[||] and Klaus Bister[||]

Institutes of Organic Chemistry and Biochemistry, University of Innsbruck, Austria

Received December 12, 1997; Revised Manuscript Received February 20, 1998

ABSTRACT: Members of the cysteine and glycine-rich protein (CRP) family (CRP1, CRP2, and CRP3) contain two zinc-binding LIM domains, LIM1 and LIM2, and are implicated in diverse cellular processes linked to differentiation, growth control and pathogenesis. The solution structure of an 81-amino acid recombinant peptide encompassing the amino-terminal LIM1 domain of quail CRP2 has been determined by 2D and 3D homo- and heteronuclear NMR spectroscopy. The LIM1 domain consists of two zinc binding sites of the CCHC and the CCCC type, respectively, which both contain two orthogonally arranged antiparallel β-sheets and which are packed together by a hydrophobic core composed of residues from the zinc finger loop regions. The CCCC zinc finger is followed by a short α-helical stretch. The structural analysis revealed that the global fold of LIM1 closely resembles the recently determined solution structures of the carboxyl-terminal LIM2 domains of quail CRP2 and chicken CRP1, and that LIM1 and LIM2 are independently folded structural and presumably functional domains of CRP proteins. To explore the dynamical properties of CRP proteins, we have used ^{15}N relaxation values (T_1, T_2, and nuclear Overhauser effect (NOE) to describe the dynamical behavior of a LIM domain. A model-free analysis revealed local variations in mobility along the backbone of the quail CRP2 LIM1 motif. Slow motions are evident in turn regions located between the various antiparallel β-sheets or between their strands. By use of an extended motional model, fast backbone motions were detected for backbone amide NH groups of hydrophobic residues located in the core region of the LIM1 domain. These findings point to a flexible hydrophobic core in the LIM1 domain allowing residual relative mobility of the two zinc fingers, which might be important to optimize the LIM1 interface for interaction with its physiological target molecule(s) and to compensate enthalpically for the entropy loss upon binding.

Cysteine- and glycine-rich proteins (CRP)[1] are encoded by a gene family (*CSRP*) that so far includes three distinct but closely related genes: the *CSRP1* gene originally identified in human (*1*), the *CSRP2* gene first isolated from quail (*2*), and the *CSRP3* gene originally cloned from rat and chicken (*3*). *CSRP1* homologues from chicken, rat, and quail (*4*−*7*), the chicken and human *CSRP2* homologues (*7, 8*), and the human *CSRP3* homologue (*9, 10*) were subsequently isolated. The definition of the *CSRP* gene family is based on the unique structural properties shared between the CRP protein products (*7*). The CRP1, CRP2, and CRP3 proteins contain 192−194 amino acid residues and display two LIM motifs in their amino acid sequences. The LIM motif is defined by two cysteine-rich zinc finger structures separated by a two amino acid spacer and conforms to the general consensus sequence $CX_2CX_{16-23}HX_2CX_2CX_2CX_{16-21}$-$CX_2$(C/H/D) (*11*). It has been identified in many proteins of diverse regulatory functions, either alone (in single or multiple copies) or in combination with defined functional domains, like homeodomains or protein kinase domains (*12, 13*). Their structural and functional properties suggest that CRP proteins have important regulatory roles in cell differentiation and proliferation (*14*). The *CSRP1* gene was shown to have properties typical for a primary response gene (*15*), and its protein product was found to be associated with specific components of the cytoskeleton including α-actinin and the LIM protein zyxin (*4, 5, 16*). The *CSRP2* gene was discovered on the basis of its strong suppression in avian fibroblasts transformed by retroviral oncogenes or chemical carcinogens (*2*), and the suppression of *CSRP2* gene expression was directly linked to the transformed phenotype of these cells (*7*). The *CSRP3* gene and its protein product (CRP3, or MLP for muscle LIM protein) were identified as positive regulators of myogenesis (*3*). *CSRP3/MLP*-deficient mice exhibit a disruption of cardiac cytoarchitectural organization and heart failure (*17*).

[†] This research was supported by Grants P11600 (to B.K.) and SFB-F002/211 (to K.B.) from the Austrian Science Foundation (FWF).

[‡] The atomic coordinates and structure factors (code 1A7I) have been deposited in the Brookhaven Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

* To whom correspondence should be addressed at the Institute of Organic Chemistry, University of Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria. Fax +43-512-507-2892; E-mail robert.konrat@uibk.ac.at.

[§] Institute of Organic Chemistry.

[||] Institute of Biochemistry.

[1] Abbreviations: CRP, cysteine- and glycine-rich protein; LIM, specific double zinc finger domain; LIM1 and LIM2, amino- and carboxyl-terminal zinc finger domains of cysteine- and glycine-rich protein; *CSRP*, gene encoding cysteine- and glycine-rich protein; X, variable amino acid; CRIP, cysteine-rich intestinal protein; PCR, polymerase chain reaction; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single-quantum correlation spectroscopy; T_1, longitudinal relaxation time; T_2, transverse relaxation time; rms, root-mean-square.

S0006-2960(97)03055-9 CCC: $15.00
Published on Web 04/29/1998

The only defined sequence motifs found in CRP proteins are the two LIM domains linked to short glycine-rich regions (*7*). Each CRP LIM motif contains two tetrahedral Zn(II)-coordinating sites of the CCHC and CCCC type, respectively (*18*−*20*). The three-dimensional structures of the carboxyl-terminal LIM2 domains of chicken CRP1 and quail CRP2 were determined by nuclear magnetic resonance spectroscopy (*21, 22*), and the protein fold of the CCCC subdomain was found to be strikingly similar to that observed for the DNA-interactive CCCC zinc fingers of the GATA-1 and steroid hormone receptor DNA binding domains (*23, 24*). CRP LIM domains are involved in specific protein−protein interactions, presumably mediated by LIM−LIM contacts (*25*−*27*), and a molecular dissection of the zyxin LIM1 domain by mutagenesis revealed that the amino-terminal CCHC zinc finger of this LIM domain is essential for the interaction with CRP1 (*28*). Also, protein−protein interactions involving LIM domains and other protein interaction domains have been described, like the interaction between the LIM protein Enigma and tyrosine-containing motifs within the endocytic code of the insulin receptor (*29, 30*), the interaction of LIM protein ENH with protein kinase C (*31*), the interactions of LIM-only (LMO) proteins or LIM homeodomain proteins with the LIM domain binding factor Ldb1/NLI (*32, 33*), or the interactions of the LIM protein LMO2(RBTN2) with the basic helix−loop−helix domain (bHLH) of transcription factor TAL1 (*34, 35*). Intriguingly, it was recently demonstrated that LMO2 and the LIM binding factor Ldb1/NLI are components of an erythroid DNA-binding complex containing the DNA-contacting transcription factors TAL1-E47 and GATA-1 (*36*). Also, specific interactions of the LIM1 domain of CRP3/MLP with the bHLH domain of myogenic transcriptional regulator MyoD were reported to enhance the DNA-binding activity of the MyoD−E47 transcription factor complex (*37*), and LIM domain-mediated protein−protein interactions are implicated in transcriptional synergism between homeodomain and bHLH proteins (*38, 39*). Although most of the available evidence points to a function of LIM domains in protein−protein interactions both in the cytoplasm and in the nucleus, direct interactions with nucleic acids cannot be ruled out (*40*).

Three-dimensional structures of LIM domains have been determined for the carboxyl-terminal LIM2 domains of chicken CRP1 (*21*) and quail CRP2 (*22*), for the single LIM domain of the cysteine-rich intestinal protein (CRIP) (*41*), and for the CCHC subdomain of the single LIM domain from Lasp-1 (*42*). Here, we have determined the solution structure of the amino-terminal LIM1 domain of quail CRP2, compared it to the structure of the carboxyl-terminal LIM2 domain in the same protein, and analyzed its dynamic properties.

EXPERIMENTAL PROCEDURES

Construction of a pET Expression Plasmid Encoding Recombinant Quail CRP2(LIM1). A polymerase chain reaction (PCR) was performed with DNA from plasmid clone W15 containing quail *CSRP2* (q*CSRP2*) cDNA (*2*) as a template and the oligonucleotides 5′-d(CTAAAGGCCTAACTGGGGAGGTGGCAAC)-3′ and 5′-d(CTTGAATTCAGCGTGCCCGCCCC)-3′ as 5′ and 3′ primers, respectively. The 5′ primer corresponds to nucleotides 4−31 of the published q*CSRP2* cDNA sequence (*2*) with nucleotide substitutions (underlined) at its 5′ end introducing a novel *Stu*I site (AGG′CCT). The 3′ primer corresponds to the sequence complementary to nucleotides 233−255 of the q*CSRP2* cDNA sequence with nucleotide substitutions (underlined) at its 5′ end introducing a novel *Eco*RI site (G′AATTC). The PCR product was digested with *Stu*I and *Eco*RI, and the *Eco*RI overhangs were filled in with Klenow DNA polymerase. The blunt-ended 241-nucleotide fragment was ligated into expression plasmid pET3d (*43*) that had been cut by *Nco*I (C′CATGG) and *Bam*HI (G′GATCC) and then blunt-ended by filling in overhangs with Klenow DNA polymerase. In the course of this cloning strategy, the codon for Pro2 was changed from CCG to CCT and the codon for Asn81 was changed from AAC to AAT, the ATG start codon was provided by the pET3d vector, and a TGA stop codon was generated at the 3′ insert−vector junction. To rule out PCR-induced mutations and to verify the integrity of the *CSRP2* coding region, the entire nucleotide sequence of the inserted PCR fragment was determined by the dideoxynucleotide chain-termination method using a T7 sequencing kit (Pharmacia, Vienna, Austria) and pET-specific primers. The expression plasmid pET3d-qCRP2(LIM1) encodes an 81 amino acid peptide encompassing amino acids 1−81 of qCRP2 including the amino-terminal LIM domain (LIM1) (*2, 7*).

Purification of Recombinant Quail CRP2(LIM1). The bacterial expression and purification of ^{15}N-labeled qCRP2(LIM1) protein was performed essentially as described for recombinant protein qCRP2(LIM2) (*22*), except that buffer A was modified to contain 50 mM sodium phosphate, pH 5.6, 10 mM NaCl, and 0.1% (v/v) β-mercaptoethanol, and that a gel-filtration step was included. Pooled qCRP2(LIM1)-containing fractions of the CM-52 eluate were dialyzed against buffer C (*22*) and concentrated by centrifugation through Centriprep 3 ultrafiltration filters (Amicon, Witten, Germany), and 1.8-ml portions of concentrated protein solutions containing approximately 8 mg of recombinant qCRP2(LIM1) were subsequently subjected to Superdex 75 HiLoad 16/60 prep grade (Pharmacia, Vienna, Austria) gel-filtration chromatography in buffer C at a flow rate of 0.25 mL/min. Protein containing fractions of the eluate were analyzed by a photometric assay (*44*) and by SDS−polyacrylamide (17% w/v) gel electrophoresis to determine protein concentration and purity. The final yield of homogeneous qCRP2(LIM1) was approximately 10 mg/L of bacterial culture. The structural integrity and purity of the protein preparation was verified by amino-terminal sequencing (F. Lottspeich, Max Planck Institute of Biochemistry, Martinsried, Germany). For NMR analysis, protein samples were concentrated by centrifugation through Centriprep 3 ultrafiltration filters.

NMR Measurements. All NMR experiments were performed on a Varian Unity Plus 500 MHz spectrometer equipped with a pulsed field gradient (PFG) unit using a triple-resonance probe with actively shielded *z* gradients. The sample contained 1.5 mM qCRP2(LIM1), 20 mM potassium phosphate, pH 7.2, 50 mM potassium chloride, and 0.5 mM dithiothreitol in 90% H_2O/10% D_2O. All spectra were recorded at 26 °C. Spectra recorded for sequential assignment include 2D TOCSY, 2D NOESY, 2D X-filtered

NOESY (*45, 46*), 2D ^{15}N HSQC, 3D ^{15}N TOCSY−HSQC, and ^{15}N NOESY−HSQC (*47*). Isotropic mixing was performed with DIPSI-2 (*48*). Water suppression was achieved by using either a WATERGATE (*49*) spin−echo or shifted laminar pulses (*50*). Heteronuclear spectra were acquired as described elsewhere (*51, 52*) with ^{15}N GARP (*53*) decoupling during acquisition. The 3D TOCSY−, and NOESY−HSQC spectra were performed including a water flipback pulse (*54, 55*). The measure of the amide attenuation factors are described in ref 22. A 3D HNHA experiment was performed to measure $^3J_{HNH\alpha}$ (*56*). All NMR spectra were processed and analyzed using NMRPipe (*57*), ANSIG (*58, 59*), and FELIX software (60). For 3D spectra the number of data points was doubled by mirror image linear prediction (LP) (*61*) in the indirect dimensions prior to zero-filling and Fourier transformation.

Structure Calculations. Three-dimensional structures were generated according to standard distance geometry/simulated annealing/refinement and energy minimization protocols using X-PLOR software (*62*) running on a cluster of Silicon Graphics workstations. NOE distance constraints were derived from 3D NOESY−HSQC, 2D NOESY, and 2D X-filtered NOESY spectra and classified as strong (1.8−2.7 Å), medium (1.8−3.3 Å), weak (1.8−5 Å), and very weak (1.8−6 Å). $^3J_{HNH\alpha}$ coupling constants were derived from the 3D HNHA experiment (*56*). Dihedral restraints for the backbone dihedral angle ϕ could be extracted from the 3D HNHA experiment by use of the appropriate Karplus relation (*63, 64*). The structures were generated from a covalent template with randomized dihedral angles by use of a distance geometry protocol followed by restrained simulated annealing (*65*), the zinc ions were covalently attached to the coordinating residues using standard force field parameters (*66*), and the resulting structures were finally minimized until convergence by use of the CHARMM force field (*67*). The sense of chirality of the zinc centers could unambiguously be determined, as only these structures exhibited minimal energies and resulted in the fewest constraint violations, in which the peptide chain winds itself around both zinc ions in the same sense as found in all earlier work on LIM proteins, and resulting in a Λ topology (*68*) for both zinc-binding sites of qCRP2(LIM1). However, following the stereochemical nomenclature proposed by Berg (*69*), the configuration of the two zinc centers can be assigned as *S* for the CCCC subdomain and as *R* for the CCHC subdomain.

^{15}N Backbone Relaxation. Dynamic information was obtained from measuring ^{15}N T_1 and T_2, and heteronuclear ^{1}H−^{15}N NOE, using proton-detected sensitivity-enhanced pulse schemes (*70*). To interpret the dynamic information contained in the ^{15}N relaxation rates, either the model-free approach (*71, 72*) or the extended treatment including an additional Lorentzian term for the expression of the spectral density function (*73, 74*) was employed. To determine the overall correlation time (τ_m), a grid search in τ_m was performed while optimizing the order parameter (S^2) and the effective correlation time (τ_e) for each residue. With this optimum τ_m as a starting point, the correlation time was refined in an iterative manner by accounting for more sophisticated motional models and exchange terms (R_{ex}), respectively.

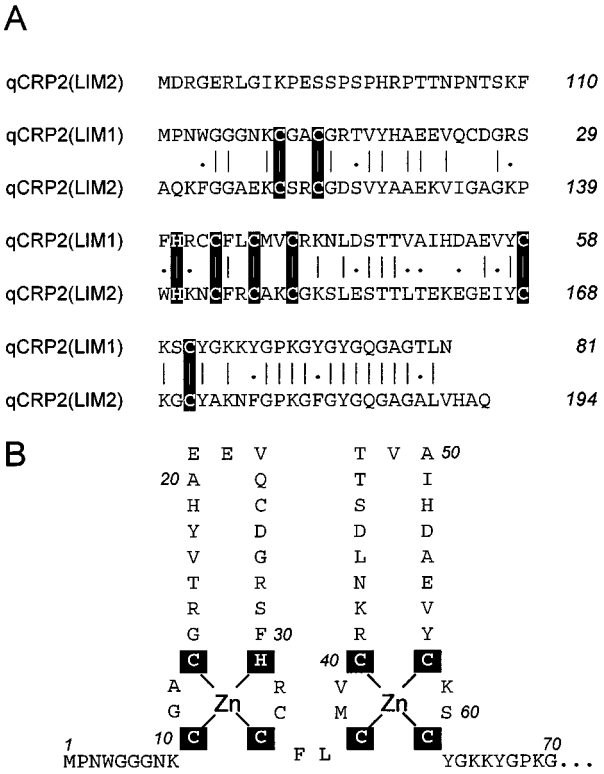

FIGURE 1: Amino acid sequence and zinc-binding sites of recombinant protein qCRP2(LIM1) containing the amino-terminal LIM domain (LIM1) of quail CRP2. (A) Alignment of the amino acid sequences of recombinant protein qCRP2(LIM1) and of recombinant protein qCRP2(LIM2) containing the carboxyl-terminal LIM domain (LIM2) of quail CRP2 (*22*). The numbering in both proteins (1−81 and 82−194, respectively) corresponds to amino acid positions in the native quail CRP2 protein (*2, 7*). Identical residues are indicated by bars, chemically similar residues by dots. Similarity groups were Ala, Ser, and Thr; Asp and Glu; Asn and Gln; Arg and Lys; Ile, Leu, Met, and Val; and Phe, Tyr, and Trp. The zinc-coordinating residues of the LIM domains in both sequences are shown in reversed type. (B) Schematic diagram of the CCHC and CCCC zinc-binding sites of the LIM double zinc finger motif in qCRP2(LIM1). The zinc-coordinating residues are shown in reversed type.

RESULTS AND DISCUSSION

Preparation of Recombinant qCRP2(LIM1). A constructed derivative of the pET3d expression vector directed the synthesis of an 81-amino acid peptide with a calculated M_r of 8824 and an estimated isoelectric point of 8.22, encompassing amino acids 1−81 of quail CRP2 comprising the amino-terminal LIM domain (LIM1). The highly soluble recombinant protein was purified to homogeneity in a two-step strategy. An alignment of the amino acid sequence of the qCRP2(LIM1) peptide used in this study with the sequence of the qCRP2(LIM2) peptide (*22*), corresponding to amino acids 82−194 of quail CRP2 including the carboxyl-terminal LIM domain (LIM2), and a schematic topology of the quail CRP2 LIM1 motif are shown in Figure 1. The sequence identity between the LIM1 (residues 10−61) and the LIM2 (residues 120−171) domains is 46.2%, and 17.3% of the residues in equivalent positions are chemically similar. The zinc coordinating residues within the LIM1 and LIM2 domains and the spacing between them are absolutely conserved. Remarkably, the sequence identities between the LIM1 and LIM2 domains of any given CRP protein range from 46% to 56% and hence are far less

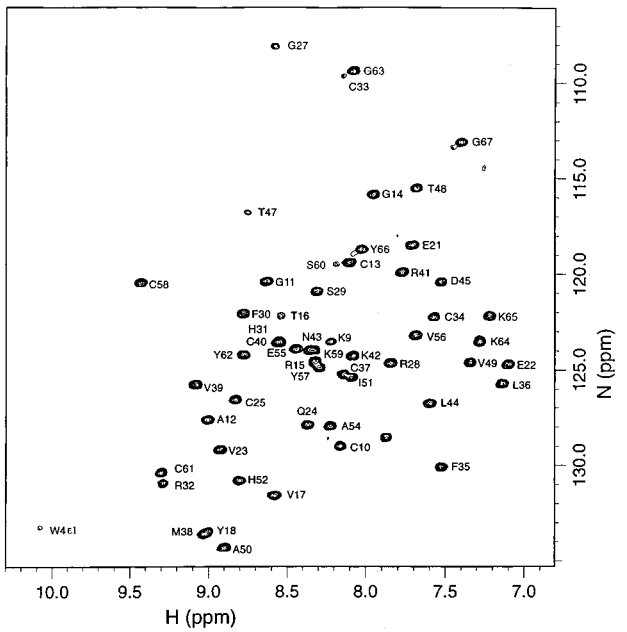

FIGURE 2: $^{15}N-^{1}H$ HSQC of qCRP2(LIM1). Cross-peaks have been labeled with their sequential assignments.

pronounced than the 71−84% sequence identities observed when either the LIM1 or the LIM2 domains of different members of the CRP protein family from a single species are aligned (*7, 8*).

NMR Analyses. From the $^{15}N-^{1}H$ HSQC analysis shown in Figure 2 it can be inferred that qCRP2(LIM1) adopts a well-defined structure in aqueous solution. However, presumably due to rapid proton exchange with the bulk solvent and/or rapid conformational exchange, only 55 out of 81 backbone amide NH groups can be observed under the current experimental conditions. Sequential signal assignment could be carried out by following well-established strategies (*75*). Spin systems were identified in the 3D TOCSY−HSQC spectrum and sequentially assigned by combining this information with the 2D NOESY or the 3D NOESY−HSQC spectrum. Regions with defined secondary structure were identified primarily by their characteristic sequential NOESY cross-peak pattern. Hydrogen bonds between the β-strands could be identified on the basis of amide proton attenuation factors and characteristic interstrand NOE patterns. This secondary structure information is supported by secondary H^{α} shifts. The β-strands generally exhibit a characteristic downfield shift, whereas the helical region is clearly shifted toward higher field (*76*). Amide proton attenuation factors support the identification of regions with defined secondary structure. Residues located in loop regions or regions with weakly defined secondary structures show low attenuation factors due to easy exchange with bulk solvent.

Tertiary Structure of qCRP2(LIM1) and Structural Relationship between LIM1 and LIM2. The structural statistics of 15 final structures of qCRP2(LIM1) are summarized in Table 1, and an overlay of the 15 calculated X-PLOR structures is shown in Figure 3. The average rms deviation is 1.30 ± 0.27 Å for the backbone atoms (N, C^{α}, C′) of the structured region (residues 9−67) of qCRP2(LIM1) (Figure 3A). Better matches within the zinc binding units were obtained when the CCHC and CCCC zinc fingers are

Table 1: Distance Geometry and Structural Statistics of the Final 15 Structures of qCRP2(LIM1)

parameter	value
distance restraints (total)	468
intraresidue	158
interresidue sequential ($\|i-j\|=1$)	143
interresidue medium-range ($1<\|i-j\|\le 4$)	55
interresidue long-range ($\|i-j\|>4$)	112
hydrogen bonds	20
dihedral angle restraints	42
average rms deviation from idealized covalent geometry	
bonds (Å)	0.013
angles (deg)	3.31
impropers (deg)	2.37
E_{L-J}[a] (kcal/mol)	−314

	atomic rms deviation (Å)	
	backbone[b]	all atoms[c]
residues 9−67[d]	1.30 ± 0.27	2.02 ± 0.31
residues 9−35	0.72 ± 0.13	1.78 ± 0.26
residues 36−67	0.83 ± 0.21	1.78 ± 0.30

[a] E_{L-J} is the Lennard Jones van der Waals energy calculated with X-PLOR using the CHARMM force field parameters. [b] Pairwise rms deviations calculated from the superposition of backbone heavy atoms (C′, C^{a}, and N) of the 15 X-PLOR structures. [c] Pairwise rms deviations calculated from the superposition of all non-hydrogen atoms of the 15 X-PLOR structures. [d] Numbering of residues is as in Figure 1. Residues 9−35 and 36−67 encompass the CCHC and CCCC zinc finger subdomains, respectively, and residues 9−67 represent the entire LIM1 domain.

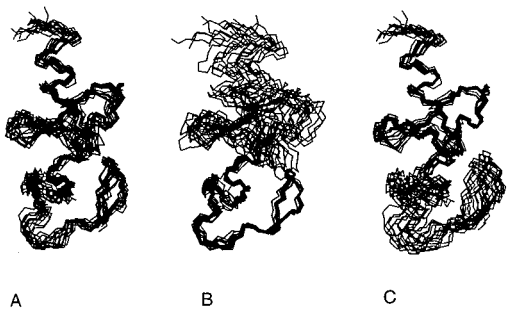

FIGURE 3: Solution structure of qCRP2(LIM1) (residues 9−67) depicted as a backbone overlay of the best 15 X-PLOR structures. (A) The whole structured region including both zinc fingers was matched. (B, C) Separate matches for each of the two zinc fingers encompassing residues 9−35 and 36−67, respectively, indicate residual mobility between the two zinc finger subdomains.

overlaid separately, yielding rms deviations of 0.72 ± 0.13 and 0.83 ± 0.21 Å for the first (residues 9−35) (Figure 3B) and second (residues 36−67) (Figure 3C) zinc finger subdomains, respectively. Figure 4 shows a schematic ribbon drawing of the tertiary structure of qCRP2(LIM1) (residues 9−67) containing rather short regions of regular secondary structure connected by flexible loop regions. Both zinc-coordinating subdomains display a similar fold containing two antiparallel β-sheets. The two zinc fingers pack together, forming a hydrophobic core region. The CCHC zinc finger starts out with a loosely defined antiparallel β-sheet (β-strands 1 and 2; residues Lys9−Cys10 and Thr16−Tyr18) connected by a rubredoxin-type turn (*77*) that accommodates the two coordinating residues Cys10 and Cys13 of this zinc finger. Hydrogen bonds are observed between the amide NH groups of Ala12 or Arg15 and Cys10 S^{γ} or Cys13 S^{γ}, respectively, a characteristic feature of the CCHC site in LIM domains (*21, 22*). A second antiparallel β-sheet (β-strands

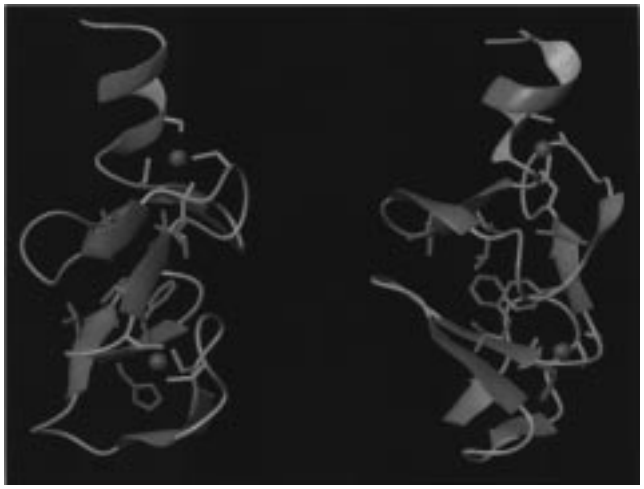

FIGURE 4: Ribbon drawing of qCRP2(LIM1) (residues 9−67). A representative structure selected from the calculated set shown in Figure 3 is depicted in two different views, including the zinc ions (magenta), the coordinating Cys side chains (yellow), the coordinating His side chain (green), and the side chains of the hydrophobic core residues (brown). The plot was generated by the program MolMol (*82*).

3 and 4; residues Val23−Cys25 and Arg28−His31) extends to the zinc-coordinating His31. It is connected at an orthogonal angle to the first β-sheet by a flexible and highly dynamic loop region. A turn accommodating the coordinating Cys34 terminates the CCHC zinc finger subdomain. ^{15}N relaxation data (see below) and protection factors indicate that the second β-sheet is significantly more rigid and therefore better defined than the first β-sheet.

The CCCC zinc finger displays a tertiary structure similar to that of the CCHC zinc finger. The two strands of a third, again loosely defined, antiparallel β-sheet (β-strands 5 and 6; Phe35−Cys37 and Lys42−Leu44) are also connected by a rubredoxin-type turn that includes the zinc-coordinating Cys40. Again, the characteristic hydrogen-bonding pattern can be observed between Val39 NH and Cys37 S$^{\gamma}$ and between Lys42 NH and Cys40 S$^{\gamma}$. A fourth antiparallel β-sheet (β-strands 7 and 8; Ala50−His52 and Glu55−Cys58) is oriented orthogonal to the third one and connected to it by a highly dynamic loop region, as indicated by ^{15}N relaxation (see below) and protection factors. Again, this β-sheet is better defined and more rigid than the preceding one. Following the fourth β-sheet, a helical stretch starts out at the coordinating Cys61 and extends to Gly67, with a kink between Lys64 and Lys65. Interresidue NOEs suggest an H-bond between S$^{\gamma}$ of Cys61 and the side chain of Lys64, and there is evidence that S$^{\gamma}$ of Cys58 is hydrogen-bonded to the side-chain amino group of Lys42. Such an extended hydrogen-bonding pattern defining a bispheric coordinative structuring of the zinc-binding sites was also found for qCRP2(LIM2) (unpublished data).

In general, the overall tertiary structure of the amino-terminal LIM1 domain of quail CRP2 is very similar to that of the carboxyl-terminal LIM2 domains of chicken CRP1 (*21*) and quail CRP2 (*22*) or to the single LIM motif of CRIP (*41*). Despite this highly similar global fold, there are distinct structural and in particular dynamic differences between the LIM1 and the LIM2 domains of qCRP2. In general, the structure of qCRP2(LIM1) appears to be less compact than that of qCRP2(LIM2). Fewer and weaker NOE contacts

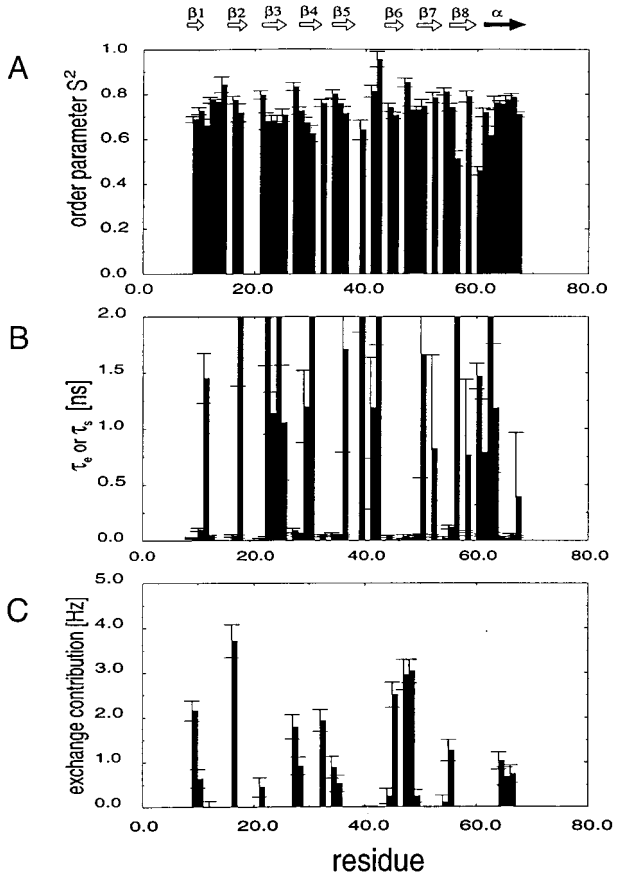

FIGURE 5: Characteristics of the qCRP2(LIM1) backbone dynamics as determined by the analysis of ^{15}N relaxation measurements (T_1, T_2, ^{1}H−^{15}N NOE) of the amide nitrogens. (A) Order parameters S^2, (B) correlation times τ_e or τ_s for internal mobility (in nanoseconds), and (C) exchange contributions (in reciprocal seconds) are plotted as a function of the residue position. For residues whose relaxation could not be measured accurately (due to signal overlap), no values are shown in the diagrams. Residues with τ_e out of the range of the scale in panel B have high uncertainties in τ_e but have nevertheless a significant contribution of slow internal motion.

between the zinc finger subdomains are observed in qCRP2-(LIM1). In particular, the hydrophobic core region involving the side chains of Val23, Phe30, Phe35, Leu44, Val49, Ile51, and Val56 appears to be less tightly packed in qCRP2(LIM1) than the corresponding region in qCRP2(LIM2). The relative orientation of the two zinc fingers is less rigid in LIM1 than in LIM2, apparently a consequence of weaker packing interactions in the hydrophobic core. In addition, the hydrogen-bonding and electrostatic interactions between Glu155 and Arg122 are missing in LIM1, where Ala12 corresponds to the H-bonding Arg122 in qCRP2(LIM2) (cf. Figure 1).

^{15}N Backbone Relaxation. ^{15}N relaxation data were analyzed in terms of the so-called model-free approach (*71*, *72*) and extensions thereof (*73*, *74*). From the 55 cross-peaks detected and assigned in the 2D ^{15}N-HSQC spectra out of 59 residues encompassing the structured part of qCRP2-(LIM1), 43 cross-peaks have been analyzed. The results of the ^{15}N relaxation measurements are illustrated in Figure 5, in which order parameters (S^2), the effective internal correlation time (either τ_e in the Lipari−Szabo model or τ_s in the extended Lipari−Szabo model), and exchange contributions (R_{ex}) are plotted as a function of residue position. An

overall correlation time τ_m of 5.6 ns was determined by fitting data from residues with T_1/T_2 ratios less than 1 standard deviation from the mean value. The various residues could be divided into three classes, according to the time scales of their internal mobilities. The first class comprises residues described by the classical Lipari−Szabo model showing order parameters (average $S^2 \sim 0.8$) and internal correlation times (average τ_e 50 ps) typical for residues found in a secondary structure element (Figure 5). The second class of residues display motions on a millisecond-to-microsecond time scale. These residues are located at the beginning of β-strands (Thr16, Glu55), at the end of β-strands (Arg32), or in turn regions connecting β-strands (Gly27, Asp45, and Thr47−Val49). Note that the backbone amide NH groups of His19, Asp26, Ser46, and Asp53 for which the correlation peaks could not be observed are adjacent to residues located in flexible loop regions, which provides additional evidence for these slow conformational exchange processes. Similar mobility can be detected for Lys9, which appears to be the first structured residue in qCRP2(LIM1), and, notably, the zinc-coordinating Cys34. This was also found for the corresponding residue Cys144 in qCRP2(LIM2) (*22*). Similar observations were also made for the residues involved in zinc binding within the zinc finger DNA binding domain of Xfin (*66*) and some of the ligand binding cysteines of *Escherichia coli* Ada (*78*). The third class comprises residues that cannot be described by the conventional Lipari−Szabo model. These residues show significant internal mobility in the nanosecond time range, superimposed on small angle fluctuations in the picosecond time scale. Residues of this class are Gly11, Val17, Glu22−Cys25, Ser29, Phe30, Leu36, Val39, Arg41, Lys42, Ala50, His52, Val56, Cys58, and Ser60−Gly63. Interestingly, most of these residues are located within well-defined secondary structure elements, where typically higher order parameters are found. Significantly, these local backbone motions are found for residues involved in the formation of the hydrophobic core, most likely reflecting the time-dependent orientation of the two zinc-binding sites. This is consistent with the fact that NOE connectivities within the hydrophobic core are sparse and confirms that the hydrophobic packing interactions between the zinc fingers are less pronounced in qCRP2(LIM1) than in qCRP2(LIM2).

Functional Aspects of LIM Domain Structure and Dynamics. Does the intramolecular long-range mobility in qCRP2(LIM1) have any functional impact? Although the precise biochemical and physiological functions of CRP proteins are unknown, there is experimental evidence that the LIM domain provides a multifunctional binding interface, capable of interacting with both nucleic acids and proteins. Interactions of LIM domains with nucleic acids have been demonstrated (*40*), and striking structural similarities between the CCCC zinc fingers within the LIM2 domains of chicken CRP1 and quail CRP2 and the DNA-interacting CCCC zinc fingers within the glucocorticoid receptor and the GATA-1 transcription factor have been reported (*21, 22*). Conformational flexibility in the loop regions connecting the third and fourth β-sheets in LIM1 or LIM2 domains [Ser46−Thr48 in qCRP2(LIM1) and Ser156−Thr158 in qCRP2(LIM2) (*22*)] may be relevant for the optimization of the binding interface interacting with a putative DNA target. It is well demonstrated that DNA-binding proteins display analogous intramolecular mobilities in the free form and get structured by induced fit upon binding to their respective nucleic acid targets (*79*). Biochemical studies have convincingly shown that LIM domains play a role also in protein−protein interactions (*25*−*39*). In many of these studies, it was unequivocally shown, that a solitary LIM domain defines a discrete functional unit that mediates protein−protein interactions. Analogously, our structural data and those from others demonstrate that a solitary LIM domain also defines an autonomously folding structural unit, and the uniformity of the protein fold based on the consensus amino acid sequence suggests that the protein binding activity may be a general feature of the LIM domain. Nevertheless, individual LIM domains can show distinct protein partner preferences. For instance, only one of the three LIM domains in zyxin is necessary and sufficient to bind to CRP1 (*26, 28*). The finding that a single LIM domain acts as a specific and distinct protein-binding interface suggests that proteins with multiple LIM domains may function as adaptor molecules, arranging two or more protein constituents into a macromolecular complex in a manner analogous to Src homology domains (SH2 and SH3) (*80*). The NMR structure (*81*) of an SH2 domain of phospholipase C-γ1 complexed with a high-affinity binding peptide revealed details about the binding or interaction mode between an SH2 domain and its partner(s). It was found that ligand specificity is determined not only by H-bond or salt bridge interactions to the phosphotyrosine moiety (pTyr) but mainly due to side-chain interactions between residues in the hydrophobic binding interface of the SH2 domain and predominantly hydrophobic residues carboxyl-terminal of pTyr. Surface potential maps (Figure 6) show that qCRP2(LIM1) contains a continuous hydrophobic patch surrounded by clusters of positively or negatively charged residues. It is formed by residues Val23, Val49, Ala50, and Ile51 and by hydrophobic residues in the core region of qCRP2(LIM1) that are partly solvent-accessible (Phe35, Leu44, and Val56). Hence, a putative protein−protein interaction site could encompass

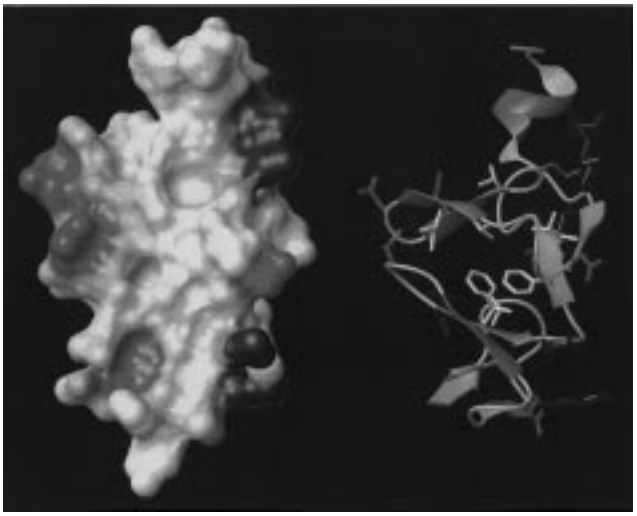

FIGURE 6: Surface structure of qCRP2(LIM1) showing the distribution of hydrophobic residues and of the electrostatic potential near the surface of the molecule. In the surface structure and in the corresponding ribbon drawing, hydrophobic residues are shown in white, positively charged residues in blue, and negatively charged residues in red. The electrostatic surface potential was calculated by the program MolMol (*82*).

part of the hydrophobic core region of the LIM domain. In this case, the intramodular flexibility would also contribute to the optimal embedding of the interaction site of the protein partner into the hydrophobic binding site and may compensate entropy losses due to decreased conformational entropy by favorable energetic interactions. From the structural and dynamical data presented here, it can be concluded that despite the well-defined global fold, the LIM domain nevertheless presents a hydrophobic core with inherent conformational flexibility. This conformational variability of the LIM domain may be of relevance for its physiological function, mediating associations with appropriate protein partners and/or nucleic acids.

ACKNOWLEDGMENT

We thank Friedrich Lottspeich (Max Planck Institute of Biochemistry, Martinsried, Germany) for protein sequencing and Sabine Weiskirchen and Gabi Langanger for excellent technical assistance. R.K. thanks N. D. Farrow (DuPont Merck, Pharmaceutical Company, Wilmington, DE) for making available numerical procedures to extract dynamical parameters from ^{15}N relaxation data.

REFERENCES

1. Liebhaber, S. A., Emery, J. G., Urbanek, M., Wang, X., and Cooke, N. E. (1990) *Nucleic Acids Res. 18*, 3871−3879.
2. Weiskirchen, R., and Bister, K. (1993) *Oncogene 8*, 2317−2324.
3. Arber, S., Halder, G., and Caroni, P. (1994) *Cell 79*, 221−231.
4. Sadler, I., Crawford, A. W., Michelsen, J. W., and Beckerle, M. C. (1992) *J. Cell Biol. 119*, 1573−1587.
5. Crawford, A. W., Pino, J. D., and Beckerle, M. C. (1994) *J. Cell Biol. 124*, 117−127.
6. McLaughlin, C. R., Tao, Q. and Abood, M. E. (1994). *Nucleic Acids Res. 22*, 5477−5483.
7. Weiskirchen, R., Pino, J. D., Macalma, T., Bister, K., and Beckerle, M. C. (1995) *J. Biol. Chem. 270*, 28946−28954.
8. Weiskirchen, R., Erdel, M., Utermann, G., and Bister, K. (1997) *Genomics 44*, 83−93.
9. Fung, Y. W., Wang, R. X., Heng, H. H. Q., and Liew, C. C. (1995). *Genomics 28*, 602−603.
10. Fung, Y. W., Wang, R., and Liew, C. C. (1996) *J. Mol. Cell. Cardiol. 28*, 1203−1210.
11. Sanchez-Garcia, I., and Rabbitts, T. H. (1994) *Trends Genet. 10*, 315−320.
12. Dawid, I. B., Toyama, R., and Taira, M. (1995) *C. R. Acad. Sci. Paris 318*, 295−306.
13. Taira, M., Evrard J.-L., Steinmetz, A., and Dawid, I. B. (1995) *Trends Genet. 11*, 431−432.
14. Louis, H. A., Pino, J. D., Schmeichel, K. L., Pomiès, P., and Beckerle, M. C. (1997) *J. Biol. Chem. 272*, 27484−27491.
15. Wang, X., Lee, G., Liebhaber, S. A., and Cooke, N. E. (1992) *J. Biol. Chem. 267*, 9176−9184.
16. Pomiès, P., Louis, H. A., and Beckerle, M. C. (1997) *J. Cell Biol. 139*, 157−168.
17. Arber, S., Hunter, J. J., Ross, J., Jr., Hongo, M., Sansig, G., Borg, J., Perriard, J.-C., Chien, K. R., and Caroni, P. (1997). *Cell 88*, 393−403.
18. Michelsen, J. W., Schmeichel, K. L., Beckerle, M. C., and Winge, D. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4404−4408.
19. Michelsen, J. W., Sewell, A. K., Louis, H. A., Olsen, J. I., Davis, D. R., Winge, D. R., and Beckerle, M. C. (1994) *J. Biol. Chem. 269*, 11108−11113.
20. Kosa, J. L., Michelsen, J. W., Louis, H. A., Olsen, J. I., Davis, D. R., Beckerle, M. C., and Winge, D. R. (1994) *Biochemistry* 33, 468−477.
21. Perez-Alvarado, G. C., Miles, C., Michelsen, J. W., Louis, H. A., Winge, D. R., Beckerle, M. C., and Summers, M. F. (1994) *Nat. Struct. Biol. 1*, 388−398.
22. Konrat, R., Weiskirchen, R., Kräutler, B., and Bister, K. (1997) *J. Biol. Chem. 272*, 12001−12007.
23. Omichinski, J. G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., and Gronenborn, A. M. (1993) *Science 261*, 438−446.
24. Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., and Sigler, P. B. (1991) *Nature 352*, 497−505.
25. Feuerstein, R., Wang, X., Song, D., Cooke, N. E., and Liebhaber, S. A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10655−10659.
26. Schmeichel, K. L., and Beckerle, M. C. (1994) *Cell 79*, 211−219.
27. Arber, S., and Caroni, P. (1996) *Genes Dev. 10*, 289−300.
28. Schmeichel, K. L., and Beckerle, M. C. (1997) *Mol. Biol. Cell 8*, 219−230.
29. Wu, R. Y., and Gill, G. N. (1994) *J. Biol. Chem. 269*, 25085−25090.
30. Wu, R. Y., Durick, K., Songyang, Z., Cantley, L. C., Taylor, S. S., and Gill, G. N. (1996) *J. Biol. Chem. 271*, 15934−15941.
31. Kuroda, S., Tokunaga, C., Kiyohara, Y., Higuchi, O., Konishi, H., Mizuno, K., Gill, G. N., and Kikkawa, U. (1996) *J. Biol. Chem. 271*, 31029−31032.
32. Agulnick, A. D., Taira, M., Breen, J. J., Tanaka, T., Dawid, I. B., and Westphal, H. (1996) *Nature 384*, 270−272.
33. Jurata, L. W., Kenny, D. A., and Gill, G. N. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 11693−11698.
34. Valge-Archer, V. E., Osada, H., Warren, A. J., Forster, A., Li, J., Baer, R., and Rabbitts, T. H. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 8617−8621.
35. Wadman, I., Li, J., Bash, R. O., Forster, A., Osada, H., Rabbitts, T. H., and Baer, R. (1994) *EMBO J. 13*, 4831−4839.
36. Wadman, I. A., Osada, H., Grütz, G. G., Agulnick, A. D., Westphal, H., Forster, A., and Rabbitts, T. H. (1997) *EMBO J. 16*, 3145−3157.
37. Kong, Y., Flick, M. J., Kudla, A. J., and Konieczny, S. E. (1997) *Mol. Cell. Biol. 17*, 4750−4760.
38. Bach, I., Carrière, C., Ostendorff, H. P., Andersen, B., and Rosenfeld, M. G. (1997) *Genes Dev. 11*, 1370−1380.
39. Johnson, J. D., Zhang, W., Rudnick, A., Rutter, W. J., and German, M. S. (1997) *Mol. Cell. Biol. 17*, 3488−3496.
40. Shibanuma, M., Mochizuki, E., Maniwa, R., Mashimo, J. I., Nishiya, N., Imai, S. I., Takano, T., Oshimura, M., and Nose, K. (1997) *Mol. Cell. Biol. 17*, 1224−1235.
41. Perez-Alvarado, G. C., Kosa, J. L., Louis, H. A., Beckerle, M. C., Winge, D. R, and Summers, M. F. (1996) *J. Mol. Biol. 257*, 153−174.
42. Hammarström, A., Berndt, K. D., Sillard, R., Adermann, K., and Otting, G. (1996) *Biochemistry 35*, 12723−12732.
43. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.
44. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
45. Griffey, R. H., and Redfield, A. G. (1987) *Q. Rev. Biophys. 19*, 51−82.
46. Otting, G., and Wüthrich, K. (1990) *Q. Rev. Biophys. 23*, 39−96.
47. 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.
48. Rucker, S. P., and Shaka, A. J. (1989) *Mol. Phys. 68*, 509−517.
49. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
50. Smallcombe, S. H. (1993) *J. Am. Chem. Soc. 115*, 4776−4785.
51. Palmer, A. G., Cavanagh, J., Wright, P. E., and Rance, M. (1991) *J. Magn. Reson. 93*, 151−170.
52. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.

53. Shaka, A. J., Barker, P. B., and Freeman, R. (1985) *J. Magn. Reson. 64*, 547−552.
54. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
55. Stonehouse, J., Shaw, G. L., Keeler, J., and Laue, E. D. (1994) *J. Magn. Reson., Ser. A 107*, 178−184.
56. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR 4*, 871−878.
57. Delaglio, F. (1993) *NMRPipe System of Software*, National Institutes of Health, Bethesda, MD.
58. Kraulis, P. J. (1987) *J. Magn. Reson. 84*, 627−633.
59. Kraulis, P. J., Domaille, P. J., Campbell-Burk, S. L., van Aken, T., and Laue, E. D. (1994) *Biochemistry 33*, 3515−3531.
60. Biosym/MSI (1995) *Felix User Guide, Version 95.0*, San Diego, CA.
61. Zhu, G., and Bax, A. (1990) *J. Magn. Reson. 90*, 405−410.
62. Brünger, A. T. (1992) *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
63. Pardi, A., Billeter, M., and Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741−751.
64. Wang, A. C., and Bax, A. (1996) *J. Am. Chem. Soc. 118*, 2483−2494.
65. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 239*, 317−324.
66. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., and Wright, P. E. (1989) *Science 245*, 635−637.
67. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
68. Purcell, K. F., and Kotz, J. C. (1977) *Inorganic Chemistry*, W. B. Saunders Company, Philadelphia, London, and Toronto.
69. Berg, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 99−102.
70. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
71. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.
72. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4559−4570.
73. Clore, M. G., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) *J. Am. Chem. Soc. 112*, 4989−4991.
74. Clore, M. G., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) *Biochemistry 29*, 7387−7401.
75. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.
76. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
77. Blake, P. B., Lee, B., Park, J. B., and Summers M. F. (1994) *New J. Chem. 18*, 387−395.
78. Habazettl, J., Myers, L. C., Yuan, F., Verdine, G. L., and Wagner, G. (1996) *Biochemistry 35*, 9335−9348.
79. Slijper, M., Boelens, R., Davis, A. L., Konings, R. N., van der Marel, G. A., van Boom, J. H., and Kaptein, R. (1997) *Biochemistry 36*, 249−254.
80. Pawson, T., and Gish, G. D. (1992) *Cell 71*, 359−362.
81. Pascal, S. M., Singer, A. U., Gish, G., Yamazaki, T., Shoelson, S. E., Pawson, T., Kay, L. E., and Forman-Kay, J. D. (1994) *Cell 77*, 461−472.
82. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 29−32.

BI973055V