Intraresidue ^{1}H-^{15}N-^{13}C′ and ^{1}Hα-^{13}Cα-^{13}C′ dipole-CSA relaxation interference as a source of constraints for structural refinement of metal-binding sites in zinc-finger proteins

Karin Kloiber, Wolfgang Schüler & Robert Konrat*
Institute of Organic Chemistry, University of Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria

Abstract

^{1}H(i)-^{15}N(i)-^{13}C′(i) dipole-chemical shift anisotropy (CSA) relaxation interference was quantified for the ^{13}C,^{15}N labeled zinc-finger protein qCRP2(LIM2). The cross-correlation rates obtained for residues located in the metal coordination sites of qCRP2(LIM2) show a high degree of correlation with the peptide plane torsion angles φ and ψ taken from the solution structure. ^{1}H(i)-^{15}N(i)-^{13}C′(i) as well as ^{13}Cα(i)-^{1}Hα(i)-^{13}C′(i) dipole-CSA cross-correlation rates were subsequently used to improve the geometry of the metal binding site. The optimized dihedral angles of the two zinc-binding sites in qCRP2(LIM2) are in better agreement with values obtained from crystal structures of other zinc-finger proteins and thus establish the utility of this approach to improve the metal-binding site geometry of zinc-finger proteins studied by NMR spectroscopy in solution.

Introduction

Zinc fingers are among the most abundant protein domains. For example, in the genome of *Caenorhabditis elegans*, more than 1000 zinc-binding domains were identified (Clarke and Berg, 1998). These autonomously folding and functional protein domains or 'fingers' constitute an important protein family and play an important role in gene regulation and transcriptional control due to their DNA-binding capacity (Schwabe and Klug, 1994; Berg and Shi, 1996). The DNA-binding selectivity of zinc-finger proteins is greatly enhanced by the fact that they can be present in proteins as sequential arrays of variable length (Berg and Shi, 1996). In addition to their importance as a DNA-binding motif, recent reports suggest that zinc-finger proteins can also serve as a protein–protein interaction scaffold (Mackay and Crossley, 1998; Hata et al., 2000).

The LIM domain is composed of a cysteine-rich motif that was first observed in the protein products of three regulatory genes (Dawid et al., 1998) and contains two tetrahedral Zn(II)-coordinating sites of the CCHC and CCCC type, capable of binding two zinc ions (Michelsen et al., 1993,1994; Kosa et al., 1994). The *CSRP* genes encode a specific class of LIM-only proteins, termed cysteine-rich proteins (CRP) and comprising two LIM domains (Weiskirchen et al., 1995). The family members found to date are involved in the regulation of cell differentiation and proliferation, particularly muscle differentiation (Arber et al., 1994; Louis et al., 1997). To date a large number of three-dimensional structures of LIM domains has been determined (Pérez-Alvarado et al., 1994, 1996; Konrat et al., 1997, 1998a,b; Kontaxis et al., 1998; Kloiber et al., 1999; Yao et al., 1999). From these studies it followed that all LIM domains share a unique globular fold and consist of CCHC and CCCC zinc-binding subdomains tightly packed via a hydrophobic core region with distinct aliphatic and aromatic side-chain interactions.

*To whom correspondence should be addressed. E-mail: robert.konrat@uibk.ac.at

It is now well established that cross-correlated NMR spin relaxation provides unique structural and dynamical information which can be beneficially applied to solution structural studies of proteins. The methods rely on quantification of cross-correlated fluctuations of different dipolar couplings (Reif et al., 1997, 1998, 2000; Carlomagno et al., 1999; Chiarparin et al., 1999; Pelupessy et al., 1999; Junker et al., 2000), dipolar couplings and anisotropic chemical shifts (Yang et al., 1997, 1998; Yang and Kay, 1998; Blommers et al., 1999; Kloiber and Konrat, 2000a,b; Sprangers et al., 2000) or different chemical shift tensors (Skrynnikov et al., 2000). In a recent paper (Kloiber and Konrat, 2000b), we showed that large ^{1}H(i)-^{15}N(i)-^{13}C′(i) dipole-CSA cross-correlation rates are observed for residues outside α-helical or β-sheet regions of Ramachandran space and proposed this relaxation interference as a simple criterion to discriminate between type I and type II β-turns in proteins. Here we demonstrate that this relaxation interference can also be applied to the refinement of the geometry of metal coordinating sites in zinc-finger proteins.

Materials and methods

Uniformly ^{13}C,^{15}N-labeled carboxyl-terminal LIM domain qCRP2(LIM2) of quail cysteine- and glycine-rich protein CRP2 was prepared and purified as described previously (Konrat et al., 1997). Signal assignment and structural data of qCRP2(LIM2) have already been published (Konrat et al., 1997).

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 of ^{13}C,^{15}N-labeled qCRP2(LIM2), as well as 20 mM potassium phosphate buffer at pH 7.2, 50 mM KCl and 0.5 mM dithiothreitol in 90% H$_2$O/10% D$_2$O. All spectra were recorded at 26 °C. Experimental ^{1}H(i)-^{15}N(i)-^{13}C′(i) dipole-CSA cross-correlation rates were obtained using the sequence described elsewhere (Kloiber and Konrat, 2000b). However, in contrast to the published version, only one 3D data set was recorded. This leads to an HNCO-type cross peak (see Figure 1), centered at the intra-residue carbonyl frequency and split by the one-bond ^{15}N-^{1}H^N scalar coupling. In case of spectral overlap, the resolution can be increased twofold by recording two 3D data sets as described (Kloiber and
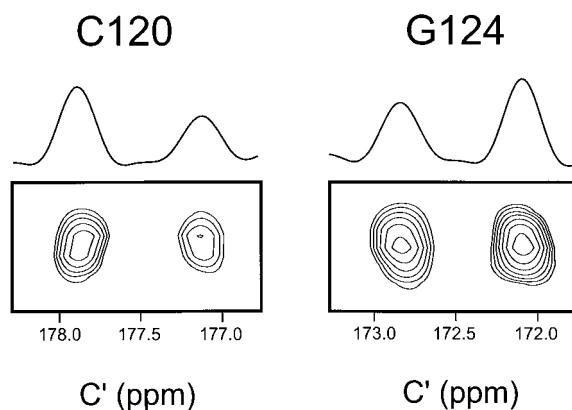

Figure 1. Experimental $\Gamma_{N(i),H(i),C'(i)}$ data for qCRP2(LIM2). Cross sections and F_1 slices through residues located in the N-terminal CCHC zinc-binding site (C120: $\Gamma_{N(i),H(i),C'(i)} = -6.70$ s^{-1} and G124: $\Gamma_{N(i),H(i),C'(i)} = +4.30$ s^{-1}) are illustrated.

Konrat, 2000b). Constraints for the protein backbone dihedral angle ψ were obtained from C^α(i)-H^α(i)-C′(i) dipole/CSA cross-correlation rates according to the procedure described by Kay and co-workers (Yang et al., 1998).

The precision limits of the experimental ^{1}H(i)-^{15}N(i)-^{13}C′(i) and C^α(i)-H^α(i)-C′(i) dipole-CSA cross-correlation rates were derived from Monte Carlo simulations of the distributions of the experimental differential peak intensities. The root-mean-square baseline noise in the experiment was taken as a measure of the standard deviations of the peak heights in the cross-correlated relaxation experiment. Data were processed and analyzed using the programs NMRPipe (Delaglio et al., 1995) and PIPP/CAPP (Garrett et al., 1991).

Values for φ and ψ were extracted from a combination of $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ cross-correlation rates by employing a 'Z-surface' approach (Le et al., 1995) and by maximizing the function

$$Z = \exp\left(-\chi^2_{\Gamma_1}\right)\exp\left(-\chi^2_{\Gamma_2}\right)$$

$$\chi^2_{\Gamma_1} = \left(\Gamma^{calc.}_{N(i)H(i),C'(i)} - \Gamma^{exp.}_{N(i)H(i),C'(i)}\right)^2 / \sigma^2_{(i)}$$

$$\chi^2_{\Gamma_2} = \left(\Gamma^{calc.}_{C\alpha(i)H\alpha(i),C'(i)} - \Gamma^{exp.}_{C\alpha(i)H\alpha(i),C'(i)}\right)^2 / \sigma^2_{(i)}$$

independently for each residue i, where σ$_{(i)}$ denotes the experimental error of the cross-correlation rates. For the calculation of $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$, standard bond lengths and angles, planar peptide bond geometry, and uniform values for both the components of the ^{13}C′ CSA tensor and the orientation of the tensor with respect to the

Table 1. Comparison between φ/ψ values from the ^{15}N-solution structure (Konrat et al., 1997) (NMR), cross-correlation refined (refined) and X-ray values obtained for residues of the metal-binding turns in the zinc-finger proteins rubredoxin from *Pyrococcus furiosus* (Day et al., 1992), Tramtrack (Fairall et al., 1993), GATA-1 DNA-binding domain (DBD) (Omichinski et al., 1993) and Gal4 (Marmorstein et al., 1992)

Position	CCHC[a]		CCCC[b]		Rubredoxin	Tramtrack	GATA-1	Gal4
	NMR[c]	Refined[d]	NMR[c]	Refined[d]				
C (i)	−90/102	−69/106	−58/145	−105/121	−75/107	−66/116	−52/113	−75/162
X (i+2)	−106/−64	−109/−44	n.a.	n.a.	−87/−50	−69/−48	−98/−29	−77/−49
C (i+3)	−80/−46	−109/−5	−85/−3	−124/−3	−111/−18	−126/15	−142/17	−63/−31
X (i+4)	120/24	119/0	77/−72	58/−19	103/−4	66/−6	52/16	−61/−55
H/C (i)	−67/169	−113/136	−130/153	−72/170	−63/131	−63/138	−49/134	−70/157
X (i+1)	n.a.	n.a.	−69/−27	−41/−26	−68/−16	−62/−32	−37/−46	−41/−52
C (i+3)	−88/26	−120/2	−95/−86	−81/−28	−116/−1	−76/−11	−65/−43	−69/−41
X (i+4)	−66/103	−101/75	−134/37	−58/−26	80/7	61/29	−63/−37	−54/−60

[a]N-terminal zinc-binding site CCHC (**C**XX**C**X$_{17}$**H**XX**C**X) comprising the residues C120-S121-R122-C123-G124 X$_{16}$ H141-K142-N143-C144-F145. No cross-correlation rates are available for S121, K142 and N143.
[b]C-terminal zinc-binding site CCCC (**C**XX**C**X$_{17}$**C**XX**C**X) comprising the residues C147-A148-K149-C150-G151 X$_{16}$ C168-K169-G170-C171-Y172. No cross-correlation rates are available for A148, K149 and G170.
[c]φ/ψ values from the ^{15}N-solution structure of CRP2(LIM2) (Konrat et al., 1997).
[d]Refined φ/ψ values obtained from simultaneous minimization of $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ cross-correlation rates, respectively (see text).

molecular peptide frame were assumed (Teng et al., 1992). χ^2 was determined within the (φ, ψ) interval [−180°,180°]. From the set of minima, the nearest minimum to the (φ, ψ) values of the solution structure was taken. In cases of a residual ambiguity, $^3J_{HNH\alpha}$ values were used to define the best solution. Dynamic information was obtained from published [15]N T_1 and T_2 relaxation (Konrat et al., 1997). For an estimation of the overall correlation time, τ_C, a grid search in τ_C was performed while optimizing the order parameter (S^2) (Lipari and Szabo, 1982a,b) for each residue.

Results and discussion

Figure 1 shows cross sections and F_1 slices through residues located in the N-terminal CCHC zinc-binding site (C120 and G124). A detailed description of the sequence and experimental parameters can be found elsewhere (Kloiber and Konrat, 2000b). The desired $\Gamma_{N(i)H(i),C'(i)}$ cross-correlation rate can be obtained from the intensity ratios of the downfield (I_{df}) and upfield (I_{uf}) multiplet components, respectively (Kloiber and Konrat, 2000b). Extracted $^1H(i)$-$^{15}N(i)$-$^{13}C'(i)$ dipole-CSA cross-correlation rates $\Gamma_{N(i)H(i),C'(i)}$ were compared with calculated values based on the solution structure of qCRP2(LIM2) (Konrat et al., 1997) and using an isotropic τ_C value of 6.2 ns. Figure 2A shows the correlation between experimental and cal-

culated $\Gamma_{N(i)H(i),C'(i)}$ cross-correlation rates obtained for residues located in the zinc-coordinating turns of qCRP2(LIM2). The agreement between experimental and theoretical $\Gamma_{N(i)H(i),C'(i)}$ cross-correlation rates is worse compared to data obtained on ubiquitin (Kloiber and Konrat, 2000b). Significant deviations occur for residues C120, C123, G124, H141, F145, C147, C150, G151, C168, K169, C171, and Y172. We have also measured $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ cross-correlation rates (Yang et al., 1998) and compared these to calculated values based on the solution structure of qCRP2(LIM2) (Figure 2B). Again, the agreement between experimental and calculated values is, although better than $\Gamma_{N(i)H(i),C'(i)}$ cross-correlation rates, significantly worse compared to other proteins which have been investigated by cross-correlated NMR relaxation (Yang et al., 1997, 1998; Chiarparin et al., 1999; Pelupessy et al., 1999; Kloiber and Konrat, 2000a). The data thus suggests a possible improvement in the accuracy and precision of the structural data of qCRP2(LIM2) when employing these cross-correlation rates for structural refinement. Based on these findings, we set out to test whether the local backbone geometry of the zinc-binding site of qCRP2(LIM2) can be optimized using intra-residue $^1H(i)$-$^{15}N(i)$-$^{13}C'(i)$ dipole-CSA cross-correlation rates. It is important to note that the dependence of $\Gamma_{N(i)H(i),C'(i)}$ on the backbone dihedral angles φ and ψ is rather complex and many (φ,ψ) val-
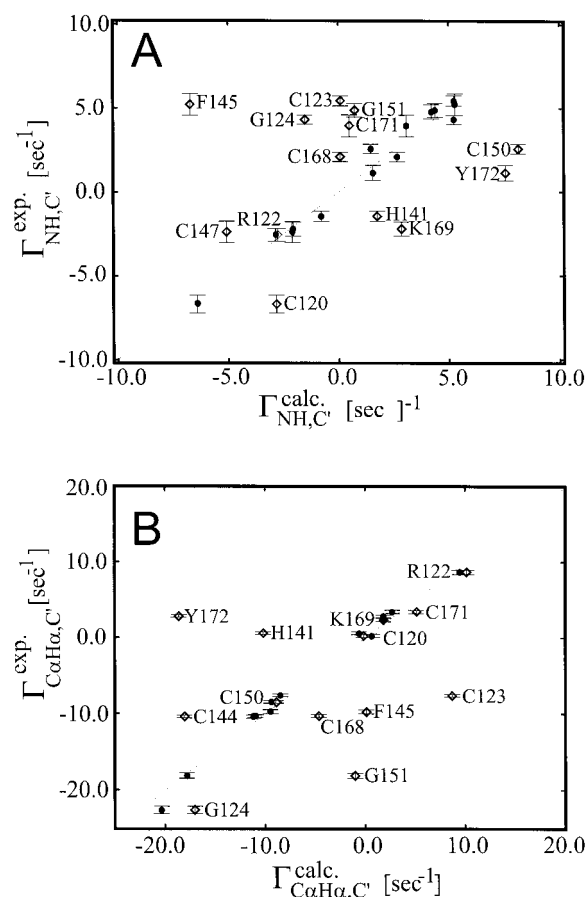
Figure 2. Correlation between experimental and calculated values of $\Gamma_{N(i),H(i),C'(i)}$ (A) and $\Gamma_{C\alpha(i),H\alpha(i),C'(i)}$ (B) for qCRP2(LIM2) before (open diamonds) and after (filled circles) dihedral angle optimization. Calculated values before optimization (open diamonds) were obtained from backbone dihedral angles taken from the solution structure of ^{15}N-labeled qCRP2(LIM2) (Konrat et al., 1997). Dihedral angles were determined by simultaneous minimization of $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ cross-correlation rate differences. Values of 244, 178 and 90 ppm were used for σ_{xx}, σ_{yy} and σ_{zz} (Teng et al., 1992). An overall correlation time τ_C of 6.2 ns was used.

ues are consistent with a single cross-correlation rate (Figure 3). We thus decided to simultaneously minimize differences between experimental and calculated $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ cross-correlation rate differences (see Materials and methods), as the combined use of both cross-correlation rates considerably reduces the possible (φ,ψ) areas. At this point, we only demonstrate the possibility to obtain more accurate (φ,ψ) angles. A more detailed description of the refined solution structure and the intramolecular dynamics of $^{13}C,^{15}N$-labeled qCRP2(LIM2) obtained from cross-correlated NMR spin relaxation will be
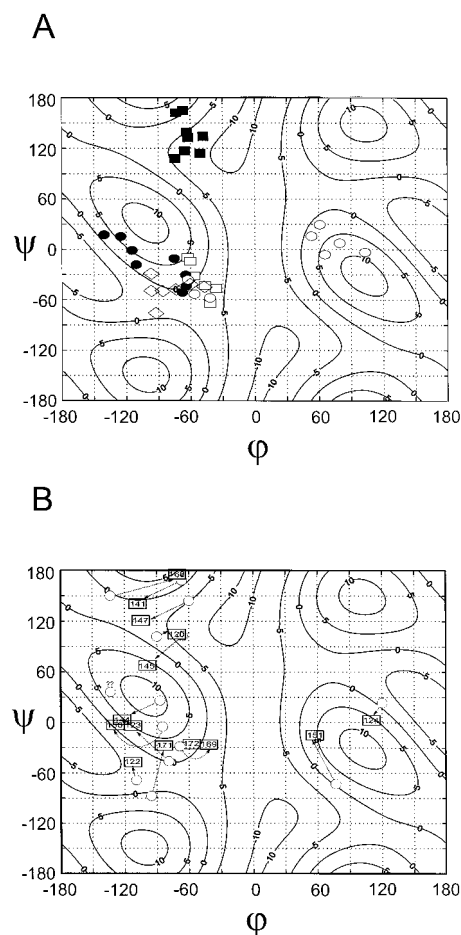
Figure 3. $\Gamma_{N(i)H(i),C'(i)}$ as a function of the intervening dihedral angles φ and ψ. (A) Typical metal-binding geometries (φ,ψ) of zinc-finger proteins, illustrated by turns from the classical zinc-finger Tramtrack (Fairall et al., 1993), rubredoxin from Pyrococcus furiosus (Day et al., 1992), GATA-1 DNA-binding domain (DBD) (Omichinski et al., 1993), and Gal4 (Marmorstein et al., 1992). Residues in the zinc-coordinating turns C(i)/H(i)-X(i+1)-X(i+2)-C(i+3)-X(i+4) (X is a variable amino acid) are depicted as filled squares, open squares, open diamonds, filled circles, and open circles. (B) Illustration of the (φ,ψ) dihedral angle changes in the two zinc-binding sites of qCRP2(LIM2) obtained from simultaneous optimization of $\Gamma_{N(i),H(i),C'(i)}$ and $\Gamma_{C\alpha(i),H\alpha(i),C'(i)}$ cross-correlation rate differences. (φ,ψ) pairs for the CCHC and CCCC zinc-binding sites of qCRP2(LIM2) before (gray circles) and after (open squares with residue number) cross-correlation rate refinement are connected by dotted arrows. Only residues of qCRP2(LIM2) are shown for which both $\Gamma_{N(i),H(i),C'(i)}$ and $\Gamma_{C\alpha(i),H\alpha(i),C'(i)}$ cross-correlation rates were observed. The residues of the consensus sequence C/H-X-X-C-X are as follows: **CCHC**: C120-S121-R122-C123-G124 and H141-K142-N143-C144-F145; **CCCC**: C147-A148-K149-C150-G151 and C168-K169-G170-C171-Y172. The gray circle labeled with ?? denotes the φ,ψ location of residue Y172 in the solution structure, for which a substantial change of the backbone dihedral angles was observed. For the calculation of the surface, CSA and τ_C parameters were as in Figure 2.

published elsewhere. Interestingly, on average only small dihedral angle changes were necessary to minimize the differences between experimental and calculated cross-correlation rates (see below). This is due to the fact that there is a very steep dependence of $\Gamma_{N(i)H(i),C'(i)}$ on both backbone dihedral angles φ and ψ (Kloiber and Konrat, 2000b), as well as of $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ on the protein backbone dihedral angle ψ (Yang et al., 1997, 1998). A comparison between dihedral angles from the ^{15}N-solution structure of qCRP2(LIM2) (Konrat et al., 1997) and the refined φ, ψ values is given in Table 1. The largest φ changes occur for residues C123, H141, C144, F145, C147, C150, C168, and Y172. Specifically, the φ angle changes observed for residues C123, C144 and C150 are in good agreement with the large $^3J(H^N H^\alpha)$ values observed in the HMQC-J experiment (Konrat et al., 1997). The φ changes towards a more α-helical conformation for residues C168 and Y172 are consistent with the experimental fact that only small $^3J(H^N H^\alpha)$ values were observed for these residues (Konrat et al., 1997). The agreement between refined φ values and $^3J(H^N H^\alpha)$ coupling constants is worse for residues H141, F145 and C147, presumably due to the presence of conformational flexibility in the two zinc-binding sites (Konrat et al., 1997). For example, residues K142 and C147 show differential multiple-quantum relaxation, $\Gamma_{C',N}$ and $\Gamma_{N,H}$, significantly different from average values, and thus suggesting the presence of conformational exchange on the ms–µs time scale at these sites (Kloiber and Konrat, 2000c). To rule out that internal dynamics may compromise the reliability of the derived angles, we analyzed ^{15}N relaxation data (Konrat et al., 1997) in terms of the model-free approach (Lipari and Szabo, 1982a,b). The average order parameter was determined to be $S^2 = 0.826 \pm 0.056$. For the residues in the metal-binding sites, the following order parameters were observed: C120: 0.829, R122: 0.833, C123: 0.852, G124: 0.870, H141: 0.864, C144: 0.827, F145: 0.794, C147: 0.870, C150: 0.799, G151: 0.893, C168: 0.876, K169: 0.859, C171: 0.762, Y172: 0.890. Apart from residue C171, no indications for internal dynamics were observed (C150 and F145 are still within one standard deviation). Finally, due to the steep dependence of $\Gamma_{N(i)H(i),C'(i)}$ on the dihedral angles φ and ψ anisotropic rotational diffusion was not taken into account. Figures 2A and 2B show that the agreements between theoretical rates and experimental values for both cross-correlations have been significantly improved.

We also compared the experimental $\Gamma_{N(i)H(i),C'(i)}$ cross-correlation rates with calculated $\Gamma_{N(i)H(i),C'(i)}$ data based on the crystal structures of zinc finger proteins with similar local folding and NH-S hydrogen bonding patterns to the coordinating cysteine thiolates (Figure 3A). The metal-binding turns in these zinc-finger proteins contained either the so-called rubredoxin full 'knuckle' (Day et al., 1992; Fairall et al., 1993; Omichinski et al., 1993) or half 'knuckle' (Marmorstein et al., 1992). The full 'knuckle' requires that the backbone displays an extended conformation before and after the zinc-binding site, whereas the half 'knuckle' is realized in cases where the zinc-binding site is followed by, for example, an α-helix. From Figure 3A it can be seen that the turn residues (C(i)/H(i)-X(i+1)-X(i+2)-C(i+3)-X(i+4)) of the various zinc-finger proteins cluster in well-defined regions of φ, ψ-Ramachandran space. Specifically, the first three coordinating residues of the metal-binding site occupy well-defined areas: (C(i)/H(i): $-80° \leq \varphi \leq -40°$/ $100° \leq \psi \leq 170°$ (filled squares); X(i+1): $-60° \leq \varphi \leq -30°$/$-60° \leq \psi \leq 0°$ (open squares); X(i+2): $-100° \leq \varphi \leq -50°$/$-80° \leq \psi \leq -30°$ (open diamonds). Using the τ_C value (6.2 ns) of qCRP2(LIM2), the corresponding theoretical $\Gamma_{N(i)H(i),C'(i)}$ cross-correlation rates are: C(i)/H(i): -9.0 to -5.0 s^{-1} (the only exception being Gal4 $\Gamma_{N(i)H(i),C'(i)} \approx 0.0$ s^{-1}); X(i+1): -4.0 to $+5.0$ s^{-1}; X(i+2): -1.5 to $+1.5$ s^{-1}. The last two coordinating residues, however, fall into two classes depending on the structure of the knuckle. A 'full' knuckle appears to occur in rubredoxin (Day et al., 1992), Tramtrack (Fairall et al., 1993) and the first turn of GATA-1 (Omichinski et al., 1993), whereas a 'half' knuckle structure is observed in the second turn of GATA-1 (Omichinski et al., 1993) and in both turns of Gal4 (Marmorstein et al., 1992). In case of a 'full' knuckle, the turn-terminating residues are as follows: C(i+3): $-150° \leq \varphi \leq -110°$/$-30° \leq \psi \leq 20°$ (filled circles); $\Gamma_{N(i)H(i),C'(i)} \approx +1.5$ to $+6.5$ s^{-1}; (the second knuckle in Tramtrack ($\varphi \approx -75°$, $\psi \approx -15°$) is slightly different and shows $\Gamma_{N(i)H(i),C'(i)} \approx +6.0$ s^{-1}); X(i+4): $50° \leq \varphi \leq 100°$/$-10° \leq \psi \leq 30°$ (open circles); $\Gamma_{N(i)H(i),C'(i)} \approx +1.5$ to $+9.0$ s^{-1}. In contrast, a 'half' knuckle is characterized by: C(i+3): $-70° \leq \varphi \leq -60°$/$-60° \leq \psi \leq -30°$ (filled circles); $\Gamma_{N(i)H(i),C'(i)} \approx +0.0$ to $+1.5$ s^{-1}; X(i+4): $-60° \leq \varphi \leq -30°$/$-60° \leq \psi \leq -30°$ (open circles); $\Gamma_{N(i)H(i),C'(i)} \approx -2.0$ to -1.0 s^{-1}. Figure 3A thus indicates that $\Gamma_{N(i)H(i),C'(i)}$ cross-correlation rates observed for successive residues (C/HXXCX) can be

CCHC

G124

C120
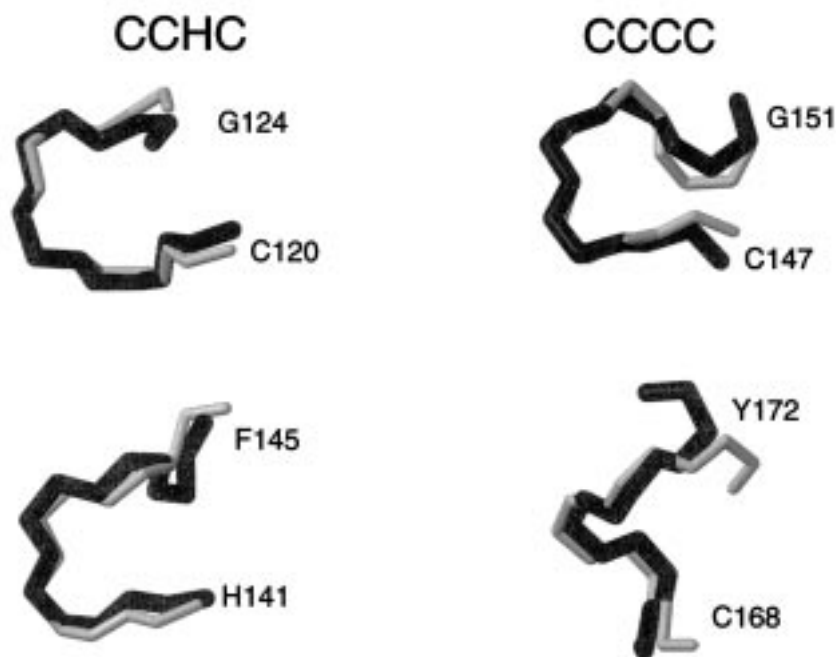
CCCC

G151

C147

F145

H141

Y172

C168

Figure 4. Cross-correlation optimized zinc-binding site geometry in qCRP2(LIM2). Best-fit superposition of the backbone atoms located in the N-terminal CCHC (left) and C-terminal CCCC (right) before and after dihedral angle optimization. The N-terminal (top) and C-terminal (bottom) knuckles of the two zinc-binding sites are shown as separate overlays. The residues at the beginning and at the end of the turns are labeled. The ^{15}N-labeled solution structure (Konrat et al., 1997) is shown in light gray, the optimized backbone structure is shown in black.

used to distinguish between the two knuckle conformations of metal-binding sites in zinc-finger proteins. For a 'full' knuckle, $\Gamma_{N(i+3)H(i+3),C'(i+3)}/\tau_C$ (ns) $\approx$ 0.25 to 1.0 s^{-1} and $\Gamma_{N(i+4)H(i+4),C'(i+4)}/\tau_C$ (ns) $\approx$ 0.25 to 1.5 s^{-1} are observed for residues C(i+3) and X(i+4). In contrast, in the case of a 'half' knuckle values on the order of $\approx$ 0.0 to 0.25 s^{-1} and $\approx$ −0.5 to 0.0 s^{-1} are obtained for $\Gamma_{N(i+3)H(i+3),C'(i+3)}/\tau_C$ (ns) and $\Gamma_{N(i+4)H(i+4),C'(i+4)}/\tau_C$ (ns), respectively.

Figure 3B shows the changes in (φ,ψ) dihedral angles obtained from the combined use of $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ cross-correlation rates for qCRP2(LIM2). Significant angle changes ($\Delta\varphi$ or $\Delta\psi > 30°$) were observed for residues C123, H141, C144, F145, C147, C150, G151, C168, C171 and, in particular, for Y172. In general, the optimized backbone dihedral angles for all of the residues in the zinc-binding sites of qCRP2(LIM2) are in better agreement with corresponding residues of other zinc-finger proteins (Figure 3A), which additionally corroborates the reliability of the approach and allows for an unambiguous assignment of the knuckle type for the zinc-binding sites. In the first turns of the N-terminal CCHC (C120 to G124) and the C-terminal

CCCC zinc finger (C147 to G151) only minor structural changes occur and the geometry is very similar to a typical rubredoxin full 'knuckle' (Konrat et al., 1997). In both cases the protein backbone adjacent to the zinc-binding site displays an extended conformation, which is the prerequisite for a full 'knuckle' structure. The geometries of the second turns in both zinc-binding sites are, however, considerably different. The second turn of the C-terminal CCCC zinc-binding module is part of the LIM domain terminating α-helix which precludes the formation of an extended structure at the end of the turn (Konrat et al., 1997). This zinc-binding site is thus structurally comparable to a half 'knuckle'. In contrast, the second turn of the N-terminal CCHC zinc-binding site does not resemble any of the two canonical knuckle types. This may be due to the specific steric demands of the two-residue linker between the two zinc-coordinating residues C144 (CCH**C**) and C147 (**C**CCC). Most importantly, the dihedral angles for F145 are very different compared to a classical knuckle structure (Figures 3A and 3B). This may be due to the fact that the side chain of F145 is involved in specific hydrophobic interactions which stabilize the global fold of the LIM domain

and determine the relative orientation of the two zinc-binding sites (Konrat et al., 1997). The importance of this interaction is evidenced by the fact that aromatic or hydrophobic residues are conserved at this position in the family of LIM domain proteins (Weiskirchen et al., 1995). It is also important to note that K142, N143 and C144 show conformational flexibility on the ms to µs time scale, presumably as a result of the steric strain caused by the specific local fold of this part of the polypeptide chain. Although these residues are part of the zinc coordinating site, similar observations have been made for the zinc-finger DNA binding domain of Xfin (Palmer et al., 1991) and some of the ligand binding cysteines of *E. coli* Ada (Habazettl et al., 1996). As a summary, Figure 4 shows a superposition of the backbone atoms located in the N-terminal CCHC and C-terminal CCCC zinc-finger of qCRP2(LIM2) before and after dihedral angle optimization, demonstrating that only small angle changes were necessary to account for the experimental $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ dipole/CSA cross-correlation rates. The geometries of the four knuckles in the two CCHC and CCCC zinc-binding sites were previously determined by characteristic $^3J(H^NH^\alpha)$ scalar couplings, and NOE and hydrogen bonding patterns (Pérez-Alvarado et al., 1994, 1996; Konrat et al., 1997, 1998a,b; Kontaxis et al., 1998; Kloiber et al., 1999; Yao et al., 1999). Due to the relatively large number of NMR constraints for the central residues of the various knuckles in qCRP2(LIM2), the geometries of the zinc-binding sites were reasonably well defined. The combined use of cross-correlation rates indicates, however, a significant improvement of the turn geometry, especially at the terminating residues. A striking example is the substantial change of backbone dihedral angles for Y172, which was only loosely defined in the ^{15}N-labeled sample due to a lack of characteristic long-range NOE connectivities. We are currently refining the solution structure of $^{13}C,^{15}N$-labeled qCRP2(LIM2) employing a significantly enlarged number of NOEs and (φ,ψ) values for the zinc-binding sites outlined in this report.

Conclusions

In summary, we have demonstrated that the combined use of intra-residue $\Gamma_{N(i)H(i),C'(i)}$ and $\Gamma_{C\alpha(i)H\alpha(i),C'(i)}$ dipole-CSA cross-correlated relaxation can be very effectively used to refine the geometry of metal binding sites in zinc-finger proteins. The steep dependence of

the cross-correlation rates on the backbone dihedral angles φ and ψ allows for a very high accuracy. The improved structural resolution may provide important insights into the factors which control the stability and geometry of the metal-binding sites in zinc-finger proteins.

References

Arber, S., Halder, G. and Caroni, P. (1994) *Cell*, **79**, 221–231.

Berg, J.M. and Shi, Y. (1996) *Science*, **271**, 1081–1085.

Blommers, M.J.J., Stark, W., Jones, C.E., Head, D., Owen, C.E. and Jahnke, W. (1999) *J. Am. Chem. Soc.*, **121**, 1949–1953.

Carlomagno, T., Felli, I.C., Czech, M., Fischer, R., Sprinzl, M. and Griesinger, C. (1999) *J. Am. Chem. Soc.*, **121**, 1945–1948.

Chiarparin, E., Pelupessy, P., Ghose, R. and Bodenhausen, G. (1999) *J. Am. Chem. Soc.*, **121**, 6876–6883.

Clarke, N.D. and Berg, J.M. (1998) *Science*, **282**, 2018–2022.

Dawid, I.B., Breen, J.J. and Toyama, R. (1998) *Trends Genet.*, **14**, 156–162.

Day, M.W., Hsu, B.T., Joshua-Tor, L., Park, J.-B., Zhou, Z.H., Adams, M.W.W. and Rees, D.C. (1992) *Protein Sci.*, **1**, 1494–1507.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Fairall, L., Schwabe, J.W.R., Chapman, L., Finch, J.T. and Rhodes, D. (1993) *Nature*, **366**, 483–487.

Garrett, D.S., Powers, P., Gronenborn, A.M. and Clore, G.M. (1991) *J. Magn. Reson.*, **95**, 214–220.

Habazettl, J., Myers, L.C., Yuan, F., Verdine, G.L. and Wagner, G. (1996) *Biochemistry*, **35**, 9335–9348.

Hata, A., Seoane, J., Lagna, G., Montalvo, E., Hemmati-Brivanlou, A. and Massague, J. (2000) *Cell*, **100**, 229–240.

Junker, J., Reif, B., Steinhagen, H., Junker, B., Felli, I.C., Reggelin, M. and Griesinger, C. (2000) *Chem. Eur. J.*, **6**, 3281–3286.

Kloiber, K., Weiskirchen, R., Kräutler, B., Bister, K. and Konrat, R. (1999) *J. Mol. Biol.*, **292**, 893–908.

Kloiber, K. and Konrat, R. (2000a) *J. Biomol. NMR*, **17**, 265–268.

Kloiber, K. and Konrat, R. (2000b) *J. Am. Chem. Soc.*, **122**, 12033–12034.

Kloiber, K. and Konrat, R. (2000c) *J. Biomol. NMR*, **18**, 33–42.

Konrat, R., Weiskirchen, R., Kräutler, B. and Bister, K. (1997) *J. Biol. Chem.*, **272**, 12001–12007.

Konrat, R., Kräutler, B., Weiskirchen, R. and Bister, K. (1998a) *J. Biol. Chem.*, **273**, 23233–23240.

Konrat, R., Weiskirchen, R., Bister, K. and Kräutler, B. (1998b) *J. Am. Chem. Soc.*, **120**, 7127–7129.

354

Kontaxis, G., Konrat, R., Kräutler, B., Weiskirchen, R. and Bister, K. (1998) *Biochemistry*, **37**, 7127–7134.

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.

Le, H.-B., Pearson, J.G., de Dios, A.C. and Oldfield, E. (1995) *J. Am. Chem. Soc.*, **117**, 3800–3807.

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

Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**, 4559–4570.

Louis, H.A., Pino, J.D., Schmeichel, K.L., Pomiés, P. and Beckerle, M.C. (1997) *J. Biol. Chem.*, **272**, 27484–27491.

Mackay, J.P. and Crossley, M. (1998) *Trends Biochem. Sci.*, **265**, 1–4.

Marmorstein, R., Carey, M., Ptashne, M. and Harrison, S.C. (1992) *Nature*, **356**, 408–414.

Michelsen, J.W., Schmeichel, K.L., Beckerle, M.C. and Winge, D.R. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 4404–4408.

Michelsen, J.W., Sevell, 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.

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.

Palmer III, A.G., Rance, M. and Wright, P.E. (1991) *J. Am. Chem. Soc.*, **113**, 4371–4380.

Pelupessy, P., Chiarparin, E., Ghose, R. and Bodenhausen, G. (1999) *J. Biomol. NMR*, **14**, 277–280.

Pérez-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.

Pérez-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.

Reif, B., Hennig, M. and Griesinger, C. (1997) *Science*, **276**, 1230–1233.

Reif, B., Steinhagen, H., Junker, B., Reggelin, M. and Griesinger, C. (1998) *Angew. Chem.*, **110**, 2006–2009.

Reif, B., Diener, A., Hennig, M., Maurer, M. and Griesinger, C. (2000) *J. Magn. Reson.*, **143**, 45–68.

Schwabe, J.W.R. and Klug, A. (1994) *Nat. Struct. Biol.*, **1**, 345–349.

Skrynnikov, N.R., Konrat, R., Muhandiram, D.R. and Kay, L.E. (2000) *J. Am. Chem. Soc.*, **122**, 7059–7071.

Sprangers, R., Bottomley, M.J., Linge, J.P., Schuktz, J., Nilges, M. and Sattler, M. (2000) *J. Biomol. NMR*, **16**, 47–58.

Teng, Q., Iqbal, M. and Cross, T.A. (1992) *J. Am. Chem. Soc.*, **114**, 5312–5321.

Weiskirchen, R., Pino, J.D., Macalma, T., Bister, K. and Beckerle, M.C. (1995) *J. Biol. Chem.*, **270**, 28946–28954.

Yang, D., Konrat, R. and Kay, L.E. (1997) *J. Am. Chem. Soc.*, **119**, 11938–11940.

Yang, D., Gardner, K.H. and Kay, L.E. (1998) *J. Biomol. NMR*, **11**, 1–8.

Yang, D. and Kay, L.E. (1998) *J. Am. Chem. Soc.*, **120**, 9880–9887.

Yao, X., Pérez-Alvarado, C., Louis, H.A., Pomiès, P., Hatt, C., Summers, M.F. and Beckerle, M.C. (1999) *Biochemistry*, **38**, 5701–5713.