Bispheric Coordinative Structuring in a Zinc Finger Protein: NMR Analysis of a Point Mutant of the Carboxy-Terminal LIM Domain of Quail Cysteine- and Glycine-Rich Protein CRP2

Robert Konrat,*,† Ralf Weiskirchen,‡ Klaus Bister,‡ and Bernhard Kräutler*,†

Institute of Organic Chemistry, Innrain 52a
Institute of Biochemistry, Peter-Mayr-Strasse 1a
University of Innsbruck, A-6020 Innsbruck, Austria

Received October 16, 1997

Tetrahedral zinc-binding domains ("zinc fingers") are important structural elements in a wide variety of proteins, and more than 10 different classes of such Zn^{2+}-binding motifs have been identified and biochemically characterized.[1,2] In these proteins the side chains of histidine and cysteine are the characteristic ligands for the Zn^{2+} ion, and for several of them specific interactions with nucleic acids have been established.[1,2] Proteins of the CRP (cysteine- and glycine-rich protein) family (CRP1, CRP2, CRP3) contain two domains (LIM domains), each composed of two zinc-binding sites of the CCHC and CCCC types (three cysteines and one histidine, and four cysteines, respectively).[3] They are implicated in diverse processes linked to cellular differentiation and growth control.[4] We have recently determined the solution structure of the carboxy-terminal LIM domain (LIM2) from recombinant quail CRP2 by heteronuclear magnetic resonance spectroscopy.[5,6] The solution structure of this zinc-binding module of CRP2 gave evidence for the existence of an H-bond and/or salt bridge between the side chain functionalities of a glutamate (E155) and an arginine residue (R122) as a possible link between the CCHC and CCCC structural units.[5]

We have synthesized [7] the mutant protein CRP2(LIM2)E155G in which the relevant glutamic acid E155 was substituted by glycine, and report here on the comparison of the solution structures of the wild type (wt) and the mutant LIM2 domains.[8,12] In addition, this work allowed the full characterization by NMR of the *extended H-bonding networks* in the coordination polyhedron around the Zn^{2+} ions in the two zinc finger units of CRP2-(LIM2) and the *first experimental determination of the protonation states* of the zinc-coordinating ligands in a natural zinc finger protein in neutral aqueous solution, a problem of considerable interest.[13,14] Our data demonstrate that *indirect* or *outer-sphere* coordination is important in a zinc finger protein, in which the Zn^{2+} centers are assigned to have a structural role.[15] In catalytically active zinc centers, *indirect* metal coordination has been recognized to be crucial for their catalytic activity and to modulate the basicity or nucleophilicity of the coordinating ligands.[15,16]

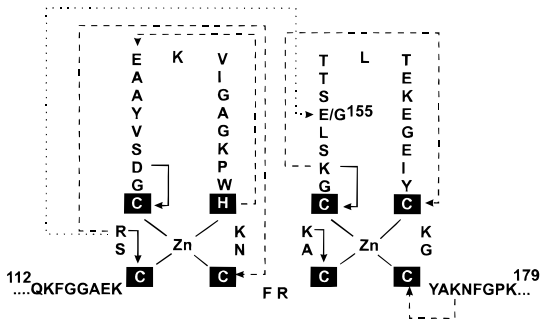

Figure 1. Amino acid sequence and zinc binding sites of wt-CRP2-(LIM2) and of the mutant CRP2(LIM2)E155G and H-bonding networks (solid arrow, H-donating backbone amide groups; dashed and dotted arrows, H-donating side chain functionality).

The uniformly ^{15}N-labeled point mutant CRP2(LIM2)E155G (see Figure 1) was synthesized using oligonucleotide-directed mutagenesis and expression in a growth medium supplemented with [^{15}N]ammonium chloride.[7] The solution structure of CRP2-(LIM2)E155G was determined by multidimensional NMR spectroscopy.[8] The folding topologies (arrangement of secondary structure elements, conformation, and the relative orientation of the two zinc-binding sites) were identical in wt-CRP2(LIM2) and

* To whom correspondence should be addressed. E-mail: Robert.Konrat@uibk.ac.at or Bernhard.Kraeutler@uibk.ac.at.
† Institute of Organic Chemistry.
‡ Institute of Biochemistry.

(1) Schwabe, J. W. R.; Klug, A. *Nat. Struct. Biol.* **1994**, 345−349.
(2) Berg, J. M.; Shi, Y. *Science* **1996**, *271*, 1081−1085.
(3) Sanchez-Garcia, I.; Rabbits, T. H. *Trends Genet.* **1994**, *10*, 315−320.
(4) Weiskirchen, R.; Pino, J. D.; Macalma, T.; Bister, K.; Beckerle, M. C. *J. Biol. Chem.* **1995**, *270*, 28946−28954.
(5) Konrat, R.; Weiskirchen, R.; Kräutler, B.; Bister, K. *J. Biol. Chem.* **1997**, *272*, 12001−12007.
(6) The solution structure of the LIM2 domain of the related but functionally distinct CRP family member chicken CRP1 has also been determined: Perez-Alvarado, G. C.; Miles, C.; Michelsen, J. W.; Louis, H. A.; Winge, D. R.; Beckerle, M. C.; Summers, M. F. *Nat. Struct. Biol.* **1994**, *1*, 388−398.
(7) *Expression plasmid encoding recombinant quail CRP2(LIM2)E155G*: Using DNA from the expression plasmid pET3d-qCRP2(LIM2) as template,[5] mutagenesis primer (E155G) 5′-d(GTTAGGGTTGTAGATCCTAGACTTTTC-CCAC)-3′ and pET3d (*Afl*III → *Bgl*II) selection primer 5′-d(CAGGAAA-GAAGATCTGAGCAAAAG)-3′ in a transformer site-directed mutagenesis reaction (Clontech, Vienna, Austria), E155 was substituted by glycine; the entire coding sequence for quail CRP2(LIM2)E155G was verified by dideoxynucleotide chain termination method with the T7 sequencing kit (Pharmacia, Vienna, Austria) and pET-specific primers. *Preparation of ^{15}N-labeled recombinant quail CRP2(LIM2)E155G*: expression and ^{15}N-labeling of CRP2-(LIM2)E155G was performed essentially as described[5] for quail wt-CRP2-(LIM2); the yield of ^{15}N-labeled CRP2(LIM2)E155G was 20 mg/L of bacterial culture. The amino acid substitution was verified mass spectrometrically (F. Lottspeich, Max-Planck-Institute of Biochemistry, Martinsried, Germany).

(8) NMR experiments: Varian UNITYPlus 500 MHz spectrometer, pulse-field gradient unit, triple resonance probes with actively shielded *z* gradients; Varian Vnmr and NMRPipe software systems;[9] experimental parameters;[5] CRP2 (LIM2)E155G 1−2 mm, 20 mM potassium phosphate pH 7.2, 50 mM KCl, 0.5 mM dithiothreitol in 90% H_2O/10% ^2H_2O. Signal assignment of CRP2(LIM2)E155G from 2D and 3D NMR experiments (3D ^{15}N TOCSY−HSQC, 3D ^{15}N NOESY−HSQC, 2D NOESY, 2D TOCSY).[10] ^{15}N attenuation factors from two sets of PFG-SE 2D ^{15}N HSQC spectra with and without presaturation of the water signal[11] (attenuation factor: ratio of the signal intensities in these two experiments). The 3D HNHA spectra[12] resulted from a 64 × 32 × 512 complex data set. The transfer delay was set to 26.8 ms.
(9) Delaglio, F. *NMRPipe System of Software*; NIH, Bethesda, MD, 1993.
(10) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J. *Protein NMR Spectroscopy*; Academic Press: New York, 1996.
(11) Spera, S.; Ikura, M.; Bax, A. *J. Biomol. NMR* **1991**, *1*, 155−165.
(12) A comparison of 2D ^{15}N HSQC spectra of wt-CRP2(LIM2) and mutant CRP2(LIM2)E155G proteins revealed no significant chemical shift differences: average chemical shift difference $\Delta\delta = 0.3$ ppm; $\Delta\delta = [(\Delta\omega_H)^2 + (\Delta\omega_N)^2]^{1/2}$; $\Delta\omega_H = \delta(^1H)$(wt-CRP2(LIM2)) − $\delta(^1H)$(CRP2(LIM2)E155G); $\Delta\omega_N = \delta(^{15}N)$(wt-CRP2(LIM2)) − $\delta(^{15}N)$(CRP2(LIM2)E155G). Maximum chemical shift difference (1.1 ppm) for Val127, located at the end of the first β-sheet; identical NOE cross-peak pattern for CRP2(LIM2)E155G; significant NOE cross-peak pattern (i.e., strong $H^N(i)-H^N(i+1)$ and $H^N(i)-H^N(i+2)$ NOEs) were found for residues C120−D125 and C147−K152 in the CCHC and CCCC units, respectively; strong sequential NOE connectivities ($H^N(i)-H^\alpha(i-1)$), indicative of extended antiparallel β-sheet conformations, were found for the sections Lys119−Cys120, Ser126−Tyr128, Val133−Ala136, Lys138−His141 (CCHC site) and Phe145−Ala148, Lys152−Leu154, Thr160−Lys162, Glu165−Lys169 (CCCC site); a C-terminal helix was detected by means of strong sequential $H^N(i)-H^N(i+1)$ NOEs. Backbone dihedral angle information stemming from HNHA-derived $^3J(H^NH^\alpha)$ scalar coupling data supported the NOE-derived arrangement of secondary structure elements. Aliphatic protons located in the hydrophobic core appeared at identical resonance positions and displayed the same long-range NOE connectivities as wt-CRP2(LIM2); Pdb-code: 1QLI (Protein Data Bank, Brookhaven National Laboratory, Upton, NY); rmsd values from the mean structure for ordered backbone atoms in the two CCHC and CCCC subdomains were 0.87 ± 0.22 and 0.81 ± 0.17 Å.

(13) Fabris, D.; Zaia, J.; Hathout, Y.; Fenselau, C. *J. Am. Chem. Soc.* **1996**, *118*, 12242−12243.
(14) (a) Wilker, J. J.; Lippard, S. J. *J. Am. Chem. Soc.* **1995**, *117*, 8682−8683. (b) Myers, L. C.; Wagner, G.; Verdine, G. L. *J. Am. Chem. Soc.* **1995**, *117*, 10749−10750.
(15) Lipscomb, W. N.; Sträter, N. *Chem. Rev.* **1996**, *96*, 2375−2433.
(16) Christianson, D. W.; Fierke, C. A. *Acc. Chem. Res.* **1996**, *29*, 331−339.

S0002-7863(97)03599-3 CCC: $15.00
Published on Web 07/03/1998

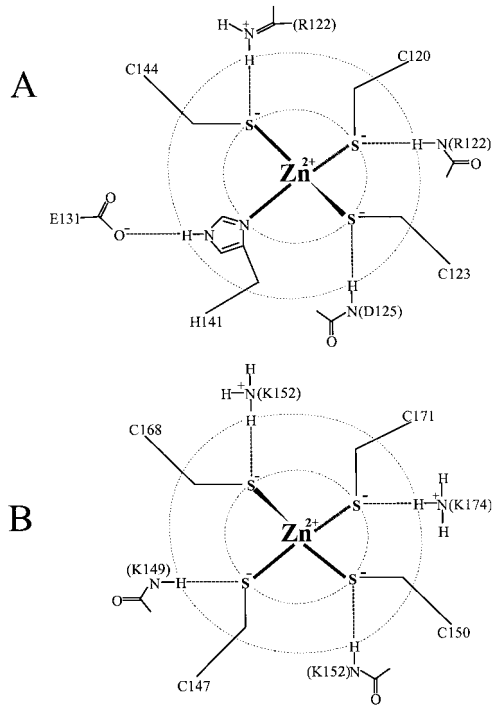

Figure 2. Schematic representation of the inner and outer coordination spheres and their associated H-bonding networks around the two zinc centers in the CCHC (A) and the CCCC (B) subdomains of CRP2(LIM2).

in the mutant.[12] This indicates that the relative orientation of the two zinc fingers in CRP2(LIM2) domains is mainly determined by hydrophobic interactions in the core region. Accordingly, H-bonding (and electrostatic) interaction between the two units, as present by the presumed link between the side chain from functionalization of glutamate E155 and arginine R122, is less important. However, H-bonding does play a pivotal role in defining the conformation[17] of the "rubredoxin knuckles" (Rd knuckles) in CCHC and CCCC zinc-binding sites.[5,6] Within the CCHC sites of both the wt-CRP2(LIM2) and the mutant CRP2-(LIM2)E155G, characteristic H-bonding occurs between S^{γ}-(Cys120) and H^{N}(Arg122) as well as between S^{γ}(Cys123) and H^{N}(Asp125) (backbone amides) (see Figures 1 and 2A). Similarly, in the CCCC unit, H-bonding exists between the backbone amide groups H^{N}(Lys149) to S^{γ}(Cys147) and H^{N}(Lys152) to S^{γ}-(Cys150) (see Figures 1 and 2B).

Complete H-bonding networks in the CCHC and CCCC zinc-binding sites could be inferred from further inspection of the final three-dimensional structures and amide attenuation factors (i.e., retardation of intermolecular exchange of amide protons with bulk water). Strong NOEs between $H^{\delta 2}$(His141) and $H^{\beta,\gamma}$(Glu131) suggest H-bonding between $H^{N\epsilon 2}$(His141) of the CCHC unit and the carboxyl group of Glu131 in both LIM2 proteins (see Figure 2A). This interaction was also found for chicken CRP1(LIM2)[6] and the single LIM-domain protein CRIP,[18] and it was suggested to be important for defining the conformation of the CCHC structural unit. The H-bond donor of the remaining thiolate of this site at Cys144 was deduced from ^{1}H and ^{15}N chemical shifts and NOE-derived distance constraints to be the side chain function of Arg122. The $H^{N\epsilon}$(Arg122) appeared at remarkably low fields, at 8.60 ppm in wt-CRP2(LIM2) and at 8.36 ppm in the mutant CRP2(LIM2)E155G. Two separate $H^{N\eta}$(Arg122) resonances (6.93 and 7.33 ppm) were observed in both proteins, indicating the H-bonding pattern of the side chain guanidinium group of Arg122 not to be changed significantly by the presence or absence of a carboxyl group in residue 155 (E155 in wt-CRP2(LIM2), G155 in the mutant). H-bonding of the protons $H^{N\epsilon}$ and $H^{N\eta}$ in the guanidinium group of an Arg residue leads to a significant downfield shift in the ^{1}H NMR spectrum.[19] Backbone to side chain NOEs were detected between H^{N}(C123) and H^{β}(R122) and H^{γ}(R122), respectively, as well as NOEs between H^{N}(C123) and the primary guanidinium protons of $H^{N\eta 1,2}$(R122). However, no NOE connectivities were found for $H^{N\epsilon}$(R122) but an intraresidue NOE with H^{δ}(R122). These results and the low attenuation factor (0.3) suggest that $H^{N\epsilon}$(R122) is exposed to the solvent and that $H^{N\eta}$(R122) forms a H-bond to the cysteine thiolate S^{γ}(C144).

Similarly, in the CCCC zinc-binding site assigned NOE connectivities to the β- and γ-protons of Lys152 suggest an orientation which brings the protonated amino group of Lys152 in close proximity to the thiolate of Cys168 (severe signal overlap in the aliphatic region of the NOESY map impeded determination of the precise side chain conformation of Lys152). Likewise, the NOE connectivities for H^{γ}(Lys174) indicate an extended conformation of the side chain of Lys174. In this way, the protonated amino group of Lys174 acts as a H-bond donor to the cysteine thiolate of Cys171. In the NMR structure of the chicken erythroid transcription factor GATA-1 complexed with its cognate DNA site, an arginine guanidinium side chain might act as the H-bond donor for this thiolate.[20] As was described for the CCCC unit in GATA-1,[20] the tetrahedrally coordinated zinc ion serves as a scaffold to orient the helix with respect to the preceding antiparallel β-sheet. Our structural data show the H-bond formed between Lys174 and Cys171 to be a significant determinant of the orientation of the C-terminal helix and to point also to a possible functional relevance of Lys174, which is conserved in all avian and human family members of the CRP protein family.[4,21] In sum, all the directly ligating residues of the two Zn finger units of LIM2 are engaged as H-bonding partners in interactions with other residues in the same Zn finger unit, extending the structuring role of the Zn^{2+} ion via the directly bound ligands to a second, outer coordination sphere[22] (see Figure 2).

Our structural definition of the H-bond donors in the CCHC and CCCC zinc-binding sites now provides a means to address the question regarding the protonation state of thiol groups in zinc ligating cysteine residues. As all zinc ligating cysteine sulfurs act as H-bond acceptors for backbone amide and side chain functional H donors, they are coordinated as thiolates, and thiol protons are not retained at the metal center. Therefore, in the CCCC unit four thiolates and in the CCHC motif three thiolates are direct ligands of the Zn^{2+} ions. Electroneutrality[23] is closely preserved in these zinc-binding sites by H-bonding interactions of some metal-coordinated thiolates with positively charged basic amino acid residues.

The results presented here reveal the existence of an integral proteinic outer-coordination sphere in zinc finger proteins, which expands the structuring role of the metal centers via extended H-bonding networks. In addition the questions[14] concerning the protonation state(s) of the zinc-coordinated cysteine ligands in such a protein have found an exemplary, but unprecedented, answer. Our conclusions concerning the latter problem are in remarkable contrast to a recent mass spectrometric study, which claimed evidence for "retention of thiol protons in two classes of protein zinc ion coordination centers",[13] with CCHC and CCCC units.

Acknowledgment. This work was supported by the Austrian National Science Foundation with Grants P-11600 (to B.K.) and SFB-F002/211 (to K.B.).

JA973599K

(17) Adman, E.; Watenpaugh, E. D.; Jensen, L. H. *Proc. Nat. Acad. Sci. U.S.A.* **1975**, 72, 4854.

(18) Perez-Alvarado, G. C.; Kosa, J. L.; Louis, H. A.; Beckerle, M. C.; Winge, D. R.; Summers, M.F. *J. Mol. Biol.* **1996**, *257*, 153−174.

(19) Pascal, S. M.; Yamazaki, T.; Singer, A. U.; Kay, L. E.; Forman-Kay, J. D. *Biochemistry* **1995**, *34*, 11353−11362.

(20) Omichinski, J. G.; Clore, G. M.; Schaad, O.; Felsenfeld, G.; Trainor, C.; Appella, E.; Stahl, E.; Gronenborn, A. M. *Science* **1993**, *261*, 438−446.

(21) Weiskirchen, R.; Erdel, M.; Utermann, G.; Bister, K. *Genomics* **1997**, *44*, 83−93.

(22) Cowan, J. A. *Inorganic Biochemistry*; VCH: New York, 1993; p 16.

(23) Lippard, S. J.; Berg, J. M. *Principles of Bioinorganic Chemistry*; University Science Books: Mill Valley, CA, 1994; p 218.