A Novel Type of Zinc Finger DNA Binding Domain in the *Agrobacterium tumefaciens* Transcriptional Regulator Ros[†]

Sabrina Esposito,[‡] Ilaria Baglivo,[‡] Gaetano Malgieri,[‡] Luigi Russo,[‡] Laura Zaccaro,[§] Luca D. D'Andrea,[§] Marco Mammucari,[‡] Benedetto Di Blasio,[‡] Carla Isernia,[‡] Roberto Fattorusso,[‡] and Paolo V. Pedone*[,‡]

Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, via Vivaldi 43, 81100 Caserta, Italy, and Istituto di Biostrutture e Bioimmagini, CNR, via Mezzocannone, 16, 80134 Napoli, Italy

Received April 11, 2006; Revised Manuscript Received July 4, 2006

ABSTRACT: Transcriptional factors bearing a Cys_2His_2 zinc finger were thought to be confined to eukaryotes, but recent studies have suggested their presence also in prokaryotes. In this paper, we report the first complete functional characterization of the DNA binding domain present in the putative Cys_2His_2 zinc finger-containing prokaryotic transcriptional regulator Ros from *Agrobacterium tumefaciens*. We demonstrate that in the single zinc binding motif present in the Ros protein the metal ion is coordinated by two cysteines (Cys79 and Cys82) and two histidines (His92 and His97), separated by a shorter spacer with respect to the eukaryotic classical Cys_2His_2 domains. The Cys_2His_2 zinc finger motif is essential for Ros DNA binding and is part of a larger DNA binding domain which includes four basic regions located on either side of the finger, one at the N-terminus and three at the C-terminus. The one described here is a novel type of DNA binding domain containing a noncanonical Cys_2His_2 zinc finger motif which, by sequence alignment, seems to be conserved in all the bacterial putative zinc finger proteins identified so far. Interestingly, basic amino acids have been shown to be important in stabilizing the DNA binding of eukaryotic single Cys_2His_2 zinc finger domains, confirming that the modality of DNA binding using a single zinc finger motif flanked by basic residues is widespread throughout the living kingdom from eukaryotic, both animal and plant, to prokaryotic, even if in each kingdom it presents its peculiarity.

Proteins containing classical type (or Cys_2His_2) zinc finger domains represent the largest group of eukaryotic DNA binding proteins known to date (*1*−*3*). The classical zinc finger domain, originally described in *Xenopus laevis* transcription factor IIIA (TFIIIA) (*4*), contains the amino acid sequence $(F/Y)XCX_{2-5}CX_3(F/Y)X_5\Psi X_2HX_{3-5}H$, where X represents any amino acid and Ψ is a hydrophobic residue; the two cysteines and the two histidines tetrahedrally coordinate a zinc ion to form a compact structure containing a β-hairpin and an α-helix ($\beta\beta\alpha$) (*1*, *4*, *5*). Structural studies carried out with animal protein classical zinc finger−DNA complexes have revealed that sequence specific recognition is achieved by contacts between the α-helix of the zinc finger and bases in the major groove of the DNA. The key residues of the helix involved in base specific DNA contacts are highly variable, depending on the recognized target sequence, and located at positions 6, 3, 2, and −1 relative to the N-terminus of the helix (*1*, *6*−*12*).

Many proteins containing Cys_2His_2 zinc finger domains have recently been identified in plants, and several data suggest that these proteins are involved in important biological processes (*13*−*15*). A structural feature common to the plant TFIIIA type zinc finger proteins, which distinguishes them from their counterparts in other eukaryotes, is the high degree of conservation of a unique amino acid sequence (QALGGH) present in their zinc finger domains. This sequence is localized at the N-terminus of the α-helix (from position 2 to 7), wherein animal zinc finger protein residues that are responsible for recognizing the DNA bases are located (*16*). Considering that in the QALGGH zinc finger two of the canonical DNA recognizing positions of the α-helix are occupied by amino acid residues (Ala in position 3 and Gly in position 6) which hardly could be involved in base specific contacts, it has been proposed that the plant QALGGH-containing zinc finger motif may adopt a peculiar DNA recognition code, different from that adopted by the animal counterparts (*16*).

The number of zinc finger domains present in proteins varies from 1 to 37, and a single zinc finger domain in itself is not sufficient for high-affinity binding to a specific DNA target sequence. In fact, proteins containing multiple zinc finger domains usually require a minimum of two zinc fingers for high-affinity DNA binding (*1*). Nevertheless, it has been demonstrated that the single zinc finger domain present in the *Drosophila* GAGA transcription factor (*10*, *17*) as well as the QALGGH single zinc finger domain of the *Arabidopsis thaliana* SUPERMAN protein (*18*) are capable of sequence specific DNA binding. In both cases,

[†] This work was funded, in part, by grants from MIUR, PRIN 2003, to R.F. and FIRB 2003 to P.V.P., and from Regione Campania, LR N.5 2003, to P.V.P.

* To whom correspondence should be addressed: Dipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, Via Vivaldi 43, Caserta, 81100 Italy. Phone: 39-0823-274598. Fax: 39-0823-274605. E-mail: PaoloV.Pedone@unina2.it.

[‡] Seconda Università degli Studi di Napoli.

[§] CNR.

10.1021/bi060697m CCC: $33.50
Published on Web 08/04/2006

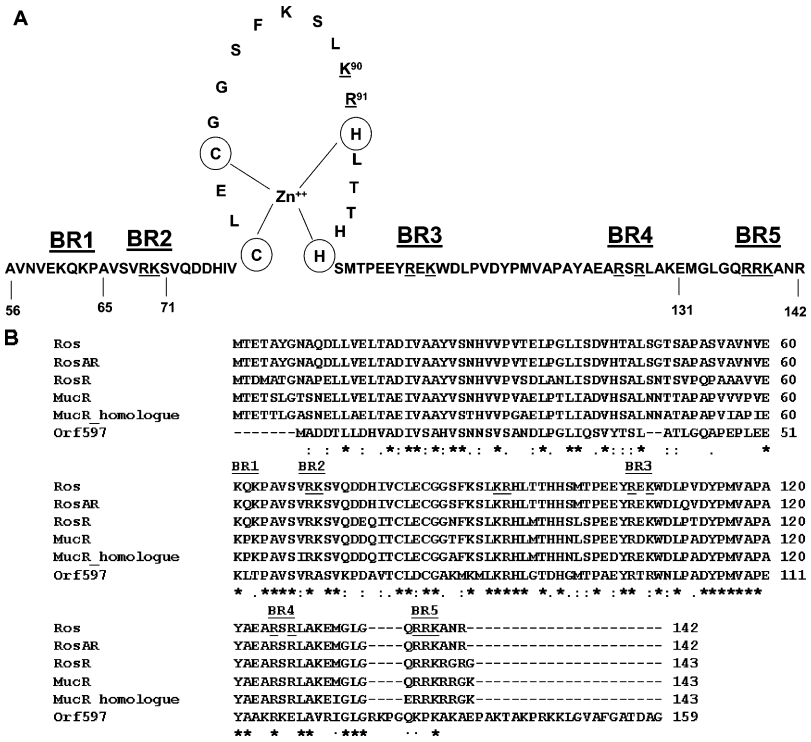

FIGURE 1: (A) Schematic representation of the amino acid sequence of the Ros protein, including the Cys_2His_2 zinc finger motif (residues 56−142). Multiple basic regions, located on either side of the finger, are indicated (BR). Residues which have been mutated in the different mutant proteins are underlined. The terminal residues of the Ros protein deletion mutants utilized in this study and the basic residues which have been mutated inside the zinc finger motif are numbered with reference to the full-length protein. (B) Sequence alignment of the Ros from *A. tumefaciens* with its homologous proteins: RosAR from *A. radiobacter*, RosR from *Rh. etli*, MucR from *S. meliloti*, MucR homologue from *Rhizobium* sp. NGR234, and Orf597 from *N. aromaticivorans* F199 (*21*). The basic residues shown to be important for Ros DNA binding capability are underlined. The basic regions (BR) flanking the zinc finger domain are indicated. An asterisk means that residues in that column are identical in all sequences in the alignment. A colon means that conserved substitutions are observed. A period means that semiconserved substitutions are observed.

the DNA binding of the finger domain is stabilized by flanking basic residues providing further contacts with the target sequences.

Transcriptional factors bearing a Cys_2His_2 zinc finger were thought to be confined to eukaryotes; however, recent studies have suggested their presence also in prokaryotes (*19*). The first putative prokaryotic Cys_2His_2 type zinc finger transcriptional regulator, the Ros protein, has been discovered in *Agrobacterium tumefaciens* (*20*). Ros homologues have been identified in different bacteria, primarily among members of the Rhizobiaceae, including RosAR in *Agrobacterium radiobacter*, RosR in *Rhizobium etli*, MucR and Orf2 in *Sinorhizobium meliloti*, the MucR homologue in *Rhizobium* sp. NGR234, and a coding region called ORF597 in *Novoshingobium aromaticivorum* F199 (Figure 1) (*21*−*27*).

Ros is a 15.5 kDa protein, encoded by the *A. tumefaciens* chromosomal gene *ros*. It regulates the expression of the virulence genes of the *virC* and *virD* operons, the products of which are involved in processing the oncogene-bearing T-DNA region of the Ti plasmid for horizontal gene transfer from *A. tumefaciens* to plant cells (*21*). Ros also regulates the expression of the *ipt* oncogene, located on the T-DNA region, which, in the host plant, is recognized by the plant transcriptional machinery and encodes the isopentenyl transferase required for the biosynthesis of cytokinins (*20*). Mutations in the *ros* gene result in upregulation of the *virC* and *virD* operons, which leads to the appearance of T-DNA intermediates in *A. tumefaciens* cells, and also derepression of the *ipt* oncogene in the T-DNA, which activates the production of cytokinin in the bacterial cells (*21*). Ros regulates the transcription by acting as a transcription factor which binds a 40 bp AT-rich DNA sequence, named Ros box, originally identified in the promoter region of virC and virD operons (*28*); a similar Ros box has also been identified in the promoter of the ipt gene (*20*).

Analysis of the Ros primary structure revealed the presence of the $CX_2CX_3FX_2LX_2HX_3HH$ amino acid sequence (from residue 79 to 97; see ref *20* and Figure 1A), which bears a significant resemblance to the eukaryotic Cys_2His_2 zinc finger motif. This motif differs from the classical Cys_2His_2 motif in that there are 9 amino acids instead of 12 in the loop that separates the second cysteine from the first histidine. By showing that the substitution of either the second cysteine or the first histidine residue in this putative zinc finger motif resulted in the release of zinc and the loss of Ros box binding activity, Chou and colleagues (*20*) demonstrated that this motif was indeed able to bind zinc and was involved in DNA binding. Whether the other zinc coordinating residues, and especially both the second and the third histidine residues, were important for DNA binding activity remained to be determined. More recently, an amino acid substitution close to the carboxy terminus of Ros which causes loss of the protein DNA binding activity has been identified, suggesting that residues outside of the zinc finger also interact with DNA or are important in maintaining the conformational integrity of the Ros DNA binding domain (*29*).

In this paper, we present the functional characterization of the Ros protein DNA binding domain. We demonstrate

that the single Cys_2His_2 zinc finger motif present in this protein is part of a larger DNA binding domain which includes multiple basic regions located on either side of the finger. Therefore, this study represents the first complete functional characterization of a novel type of prokaryotic DNA binding domain containing a noncanonical Cys_2His_2 zinc finger motif.

EXPERIMENTAL PROCEDURES

Cloning and Purification of the Proteins. DNA fragments encoding the different segments of the Ros protein to be expressed as gluthatione *S*-transferase (GST) fusion proteins were generated by PCR from *A. tumefaciens* genomic DNA. Oligonucleotides were synthesized on the basis of the published sequence (M65201). The following oligonucleotides were used as primers: primer 1, 5′-CGGGATCCGCGGTCAATGTTGAAAAGCA-3′, and primer 2, 5′-CGGAATTCTCAACGGTTCGCCTTGCGG-3′, for $_{gst}$Ros56−142; primer 3, 5′-CGGGATCCGCTGTGTCGGTTCGCAAGTC-3′, and primer 2bis, 5′-CGGAATTCTCAACGGTTCGCCTTGCGG-3′, for $_{gst}$Ros65−142; primer 4, 5′-CGGGATCCTCGGTTCAGGACGATC-3′, and primer 2bis for $_{gst}$Ros71−142; and primer 1 and primer 5, 5′-CGGAATTCTCATTCCTTGGCGAGCCGCGAAC-3′, for $_{gst}$Ros56−131. All the point mutants were generated by PCR-mediated mutagenesis according to the method of White et al. (*30*) using a single-step or double-step reaction. The following primers were utilized: primer 1, C79AREV 5′-CGAGCCACCACATTCCAAAGCGAC-3′, C79AFOR 5′-GTCGCTTTGGAATGTGGTGGCTCG-3′, and primer 2bis for C79A; primer 1, C82AREV 5′-CGAGCCACCAGCTTCCAAACAGACGATATG-3′, C82AFOR 5′-GGAAGCTGGTGGCTCGTTCAAGTCG-3′, and primer 2bis for C82A; primer 1, H96AREV 5′-CTTCCGGCGTCATGCTGTGAGCCGTCGTCAGGTGGCGTTT-3′, H96AFOR 5′-AAACGCCACCTGACGACGGCTCACAGCATGACGCCGGAAG-3′, and primer 2bis for H96A; primer 1, H97AREV 5′-ATTCTTCCGGCGTCATGCTGGCATGCGTCGTCAGGTGGCG-3′, H97AFOR 5′-CGCCACCTGACGACGCATGCCAGCATGACGCCGGAAGAAT-3′, and primer 2bis for H97A; primer 1, H96AH97AREV 5′-CTTCCGGCGTCATGCTGGCAGCCGTCGTCAGGTGGCG-3′, H96AH97AFOR 5′-CGCCACCTGACGACGGCTGCCAGCATGACGCCGGAAG-3′, and primer 2bis for H96AH97A; primer 1, H92AREV 5′-CGTCAGGGCGCGTTTGAGCGACTTGAAC-3′, H92AFOR 5′-CAAACGCGCCCTGACGACGCATCACAG-3′, and primer 2bis for H92A; primer 3, R69AK70AREV 5′-GAACCGACGCGGCAACCGACACAGCAGGCTTC-3′, R69AK70AFOR 5′-CGGTTGCCGCGTCGGTTCAGGACGATC-3′, and primer 2bis for $_{gst}$Ros65−142BR2mut; primer 1 and R125AREV 5′-CGGAATTCTCAACGGTTCGCCTTGCGGCGCTGACCGAGACCCATTTCCTTGGCGAGCCGCGAAGCGGCTTCGGCATAGGC-3′ for R125A; primer 1, R127AREV 5′-CGAGCGCCGAACGGGCTTCGGCATAG-3′, R127AFOR 5′-GTTCGGCGCTCGCCAAGGAAATGGG-3′, and primer 2bis for R127A; primer 1, R105AK107AREV 5′-CCCATGCTTCGGCATATTCTTCCGGCGTCATG-3′, R105AK107AFOR 5′-GAATATGCCGAAGCATGGGATCTGCCGGTCG-3′, and primer 2bis for BR3mut; primer 1 and R125AR127AREV 5′-CGGAATTCTCAACGGTTCGCCTTGCGGCGCTGACCGAGACCCATTTCCTTGGCGAGCGCCGAAGCGGCTTCGGCATAGGCGGGAG-3′ for BR4mut; primer 1 and R137AR138AK139AREV 5′-CGGAATTCTCAACGGTTCGCCGCGGCGGCCTGACCGAGACCC-3′ for BR5mut; and primer 1, K90AR91AREV 5′-GGTGGGCTGCGAGCGACTTGAACGAGCC-3′, K90AR91AFOR 5′-CGCTCGCAGCCCACCTGACGACGCATCAC-3′, and primer 2bis for K90AR91A.

All the PCR products were digested with the restriction enzymes *Bam*HI (for $_{gst}$Ros56−142) or *Bam*HI and *Eco*RI (all the others) and cloned into a *Bam*HI-digested or a *Bam*HI- and *Eco*RI-digested pGEX-6P-1 (Amersham Biosciences) bacterial expression vector. All the plasmids that were obtained were sequenced to confirm that there were no mutations in the coding sequences. The fusion proteins were expressed in the *Escherichia coli* BL21 host strain, induced for 2 h in the presence of 1 mM IPTG at 37 °C, sonicated in 1× PBS (pH 7.4) (*31*), 1 mM PMSF, 1 μM leupeptin, 1 μM aprotinin, and 10 μg/mL lysozyme, adjusted to 1% Triton X-100, and centrifuged for 20 min at 29 000 rcf. The supernatant was then loaded on a glutathione−Sepharose resin (Amersham Biosciences) according to the manufacturer's protocol. Following washes with 1× PBS, purified fractions were eluted in glutathione elution buffer [10 mM glutathione, 100 mM Tris (pH 8.0), and 100 mM NaCl] and used for band-shift assays and ICP-Ms analysis.

To remove the GST from the $_{gst}$Ros56−142 fusion protein, according to the manufacturer's advice, a cleavage reaction mix containing 50 mM Tris (pH 8.0), 150 mM NaCl, 1 mM EDTA, 1 mM DTT, and 1 unit/100 μg of GST fusion protein of Prescission Protease (Amersham Biosciences) was loaded on the resin where the fusion protein was bound, incubated for 12 h at 4 °C, and centrifuged for 2 min at 500 rcf, and the supernatant containing the protein without the GST tag was then recovered.

The GST protein was expressed by transforming the *E. coli* BL21 host strain with the empty pGex-6P-1 plasmid and purified as described for the other fusion proteins.

The coding sequences for the proteins $_{pet}$Ros56−142 and $_{pet}$Ros56−142H97A, not expressed as GST fusions, were generated by PCR using as templates the plasmids carrying the coding sequence for $_{gst}$Ros56−142 and $_{pet}$Ros56−142-H97A, respectively. The primers utilized were Rosdel56NcoI, 5′-ACATGCCATGGCGGTCAATGTTGAAAAGCA-3′, and primer 2. The obtained DNA fragments were then digested with the restriction enzymes *Nco*I and *Bam*HI, cloned into the pET-11d vector, and expressed in the *E. coli* BL21(DE3) host strain. For NMR spectral analysis, the $_{pet}$Ros56−142 and $_{pet}$Ros56−142H97A proteins were produced as follows. Uniform (>95%) ^{15}N and ^{13}C labeling was achieved by growing the cells in a modified minimal medium containing 0.5 g/L $^{15}NH_4Cl$ and/or 0.9 g/L [$^{13}C_6$]glucose as the sole nitrogen and carbon sources, respectively. The cells were grown at 37 °C until the OD_{600} equaled 0.6, after which the protein expression was induced for 2.5 h with 0.7 mM IPTG. The cells were then harvested, resuspended in buffer A [20 mM Na_2HPO_4 (pH 6.8)], and lysed by sonication. The crude cell extracts were clarified by centrifugation at 29 000 rcf for 40 min, and the supernatant was applied to a Mono S HR 5/5 cation exchange chromatography column (Amersham Biosciences) equilibrated with buffer A. The fractions containing the $_{pet}$Ros56−142 or $_{pet}$Ros56−142H97A protein were eluted using a gradient (from 0 to 100% over 40 mL) of buffer B [20 mM Na_2HPO_4 (pH 6.8) and 1 M NaCl]. The

pooled fractions containing the proteins were applied to a HiLoad 26/60 Superdex 75 (Amersham Biosciences) gel filtration chromatography column (320 mL bed volume) equilibrated with buffer C [20 mM Na_2HPO_4 (pH 6.8) and 0.2 M NaCl]. The purified products were then concentrated using a Centriplus YM-3 (Amicon) concentrator to give a final concentration of ~1.0 mM. For MALDI analysis, the PAR-PCMB assay, and circular dichroism experiments, the proteins were expressed in LB medium and then purified as described above.

Gel Mobility Shift Analysis. Unless otherwise specified, 500 ng of the purified proteins was incubated for 10 min on ice with 50 fmol of the labeled duplex oligonucleotide VirC, 5′-GATTTTATATTTCAATTTTATTGTAATATAATTTC-AATTG-3′, in the presence of 25 mM HEPES (pH 7.9), 50 mM KCl, 6.25 mM $MgCl_2$, 1% NP-40, 5% glycerol, and 200 ng of double-stranded poly(dI/dC-dI/dC) (Roche). After incubation, the mixture was loaded on a 5% polyacrylamide gel (29:1 acrylamide:bisacrylamide ratio) and run in 0.5× TBE (*31*) at 4 °C (200 V for 135 min). As a nonspecific competitor for competition experiments, the oligonucleotide NS, 5′-TGGCCAGGGCCGCGCCGTGGCGGGGCCAGGG-CGCGGGGCT-3′, was used. The affinity of $_{gst}$Ros56−142 and $_{pet}$Ros56−142 for the VirC oligonucleotide was measured by a gel mobility shift assay by performing a titration of the proteins with the oligonucleotide. In the case of $_{gst}$Ros56−142, in a volume of 20 μL, 5 pmol of the protein was incubated with 0.25, 0.5, 0.75, 1.25, and 2 pmol of the duplex oligonucleotide; in the case of $_{pet}$Ros56−142 in a volume of 20 μL, 15 pmol of the protein was incubated with 2, 3, 4, 5, and 6 pmol of the duplex VirC oligonucleotide in the presence of 10 pmol of the NS oligonucleotide as a nonspecific competitor. The nonspecific competitor was necessary to obtain a cleaner data set. Scatchard analysis of the gel shift binding data was performed to obtain the K_d values. All numerical values were obtained by computer quantification of the image using a Amersham Biosciences Typhoon Trio + apparatus.

MALDI Analysis. To observe the $_{pet}$Ros56−142−zinc complex at neutral pH and the dissociation of the Zn^{2+}−$_{pet}$Ros56−142 complex at acidic pH, MALDI spectra were acquired on a Voyager Perseptive Biosystem spectrometer using 3,5-dimethoxy-4-hydroxycinnamic acid as a matrix. To change the solution pH, the matrix (2.5 mg) was dissolved in a water/acetonitrile mixture (7:3, v/v) and in a water (0.1% TFA)/acetonitrile (0.1% TFA) mixture (7:3, v/v) to yield pH values of 7.0 and 1.9, respectively. For both sample preparations, 2 μL of a $_{pet}$Ros56−142 solution (150 ng/μL) and 4 μL of the matrix were mixed and 2 μL of the resulting solution was deposited on the plate.

PAR-PCMB Assay. Zinc assessment and analysis of the number of zinc-coordinating cysteines were performed by using a modified version of the spectrophotometric assay originally reported by Boyer (*32*). The assay is based on the formation of a complex of free zinc and the zinc-complexing dye 4-(2-pyridylazo)resorcinol (PAR), which produces an intense red dye (ϵ_{500} = 66 000 M^{-1} cm^{-1}). To analyze the amount of free zinc in a wild-type $_{pet}$Ros56−142 solution, 2 nmol of protein was incubated in 500 μL of a 100 μM PAR, 40 mM metal-free HEPES-KOH (pH 7.5) solution, and the A_{500} was monitored using a Hewlett-Packard 8453 spectrophotometer; the $A_{500\ nm}$ obtained value was then subtracted in the following measurements. To determine the number of cysteines involved in the zinc coordination, aliquots of 2 nmol of a 4-(chloromercuric)benzoic acid (PCMB) titration solution (2 μL of a 1 mM solution) were added, and the changes in A_{500} were recorded. The 1 mM PCMB stock solution was obtained by suspending the PCMB powder in water and treating the solution with sufficient 1 M NaOH to dissolve the material. PCMB, a strongly sulfhydryl-dissociating reagent, was used to remove Zn^{2+} from $_{pet}$Ros56−142 by forming mercaptide bonds with thiols, leading to the release of zinc into the solution where it is immediately complexed by PAR, thereby turning the solution red (*32*, *33*). The PCMB titration solution was added until all zinc was released from the protein, as indicated by a constant A_{500} signal.

ICP-MS Analysis. The amount of zinc contained in $_{gst}$Ros proteins was measured by ICP-MS. The recombinant proteins (15.3 nmol) were eluted in glutathione elution buffer, and the dried samples were resuspended in 600 μL of nitric acid and heated at 200 °C for 30 min to hydrolyze the protein. The hydrolyzed samples were filled up to 10 mL with 1% nitric acid and analyzed using ICP-MS (Perkin-Elmer). As a blank, a sample without the protein was prepared in the same way.

NMR Experiments. The NMR experiments were recorded with 0.8 mM samples of uniformly ^{15}N-labeled $_{pet}$Ros56−142 and $_{pet}$Ros56−142H97A, containing 20 mM phosphate buffer, at pH 6.8 and 25 °C, in a 90:10 $H_2O/^2H_2O$ mixture. $^1H-^{15}N$ HSQC spectra were acquired using as transfer delays a τ_m [1/(4*J*)] of 13.8 ms, to obtain the coherence transfer from the $H_{\epsilon 1}$ and $H_{\delta 2}$ histidine side chain protons to $N_{\epsilon 2}$ and $N_{\delta 1}$ through the $^2J_{HN}$ coupling constant. The spectral width values were 6000 Hz and carrier at 4.75 ppm (t_2) and 8610 Hz and carrier at 175 ppm (t_1). The experimental matrix was 1024 × 256 complex data points and was transformed to yield a final matrix of 4096 × 1024 data points.

Circular Dichroism Experiments. A JASCO 700 spectropolarimeter was used to acquire CD spectra at pH 7.0 and 2.6. The spectrum of $_{pet}$Ros56−142 (0.3 μM) in 20 mM phosphate buffer containing 40 mM NaCl (pH 7.0) was recorded in a 0.1 cm path length cell from 260 to 195 nm. The pH of the sample was adjusted to pH 2.6 with a concentrated HCl solution. The spectrum of the protein at pH 2.6 was acquired from 260 to 200 nm using a 0.1 path length cell. All spectra represent the average of three scans. CD intensities are expressed as mean residue ellipticities (decimoles square centimeter per mole). The CD spectrum of $_{pet}$Ros56−142 at pH 7.0 was analyzed for secondary structure content via a neural network approach (*34*).

RESULTS

The Ros56−142 Fragment Contains the DNA Binding Domain. To define the minimal region of the Ros protein that is able to bind DNA, fragments encoding the full-length protein (Ros1−142) and different N-terminal deletion mutants (Ros29−142, Ros48−142, and Ros56−142), including the putative zinc finger motif, were cloned into the pGEX-6P1 expression vector. The proteins were expressed in *E. coli* as a glutathione *S*-transferase (GST) fusion and purified. A gel mobility shift analysis was performed on the purified proteins to determine their ability to bind DNA. As shown

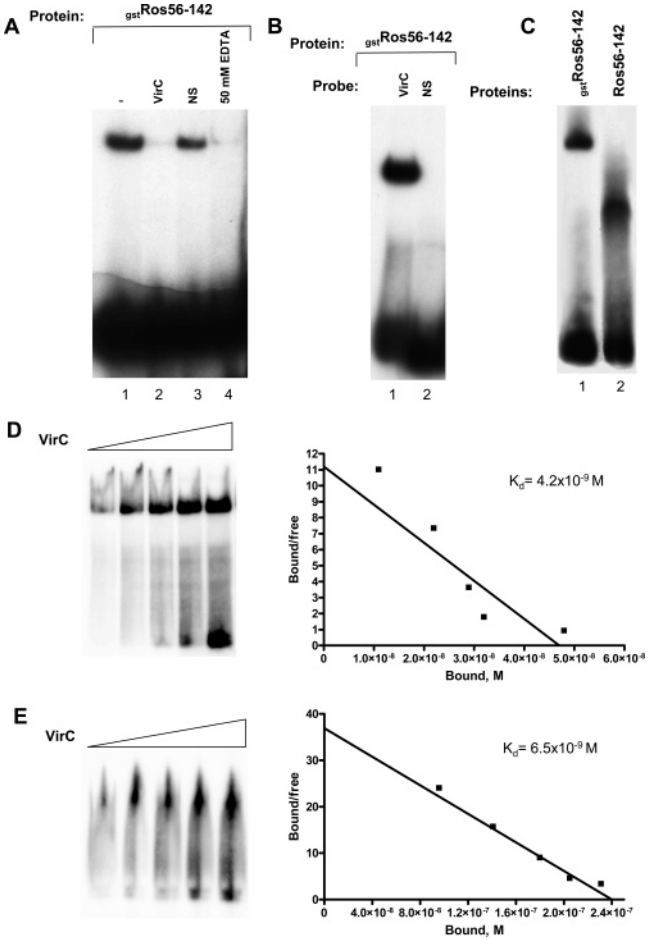

FIGURE 2: Analysis of Ros fragment 56−142 DNA binding. (A) Gel mobility shift analysis of $_{gst}$Ros56−142 DNA binding. The DNA binding specificity and zinc requirement for binding have been investigated. Purified $_{gst}$Ros56−142 protein (75 ng) was incubated with 1 pmol of the labeled VirC oligonucleotide in the absence (lane 1) or presence of a 100-fold excess of unlabeled specific oligonucleotide VirC (lane 2), a 100-fold excess of an unlabeled oligonucleotide with a nonspecific sequence (NS, lane 3), or 50 mM EDTA (lane 4) and then subjected to gel shift analysis. (B) $_{gst}$Ros56−142 protein was incubated with the labeled VirC (lane 1) and NS (lane 2) oligonucleotides and then subjected to gel shift analysis. (C) Gel mobility shift analysis of the Ros56−142 fragment after removal of the GST tag. The purified $_{gst}$Ros56−142 protein (1 μg, lane 1) and the purified Ros56−142 fragment after the proteolytic cleavage of the GST (1 μg, lane 2) were incubated with 10 pmol of labeled VirC oligonucleotide and then subjected to gel shift analysis. (D) Gel mobility shift titrations of $_{gst}$Ros56−142 with the VirC oligonucleotide (see Experimental Procedures) (left) and Scatchard analysis of the gel shift binding data (right). The ratio of bound to free DNA is plotted vs the molar concentration of bound DNA in the reaction mixture. (E) Gel mobility shift titrations of $_{pet}$Ros56−142 with the VirC oligonucleotide (see Experimental Procedures) (left) and Scatchard analysis of the gel shift binding data (right).

in Figure 2A (lane 1), $_{gst}$Ros56−142 protein, as well as the other proteins that were tested (data not shown), binds the VirC oligonucleotide, which is derived from the Ros box present in the virC/D operator (*20*, *28*), producing a single complex. This result indicates that the N-terminal region of the protein (from residue 1 to 55) is not required for DNA binding. The binding specificity of the purified $_{gst}$Ros56−142 protein was demonstrated by competition experiments with unlabeled oligonucleotides; the complex is challenged by addition of a 100-fold excess of unlabeled VirC oligonucleotide (Figure 2A, lane 2), but not by the same amount of an unrelated oligonucleotide sequence (NS, Figure 2A, lane 3). Similarly, $_{gst}$Ros56−142 is not able to bind the NS-labeled oligonucleotide when it is used as a probe (Figure 2B, lane 2). Addition of 50 mM EDTA to the binding reaction (Figure 2A, lane 4) abolishes the DNA binding activity, suggesting that the binding of the protein to DNA is metal-dependent. To exclude the possibility that the GST moiety could influence the DNA capability of the fusion protein, we proved by gel shift analysis that the Ros56−142 fragment is able to bind DNA after the GST tag had been removed (Figure 2C, lane 2). The affinity of $_{gst}$Ros56−142 for the VirC oligonucleotide was measured by a gel mobility shift assay by performing a titration of the protein with the oligonucleotide. Scatchard analysis of the data leads to an apparent dissociation constant of 4.2×10^{-9} M (Figure 2D). Similar results ($K_d = 6.5 \times 10^{-9}$ M) were obtained with the Ros56−142 fragment expressed without the GST moiety ($_{pet}$Ros56−142) (Figure 2E).

FIGURE 3: Analysis of the role of zinc finger residues K90 and R91 in Ros DNA binding. Gel shift analysis of the K90AR91A mutant protein DNA binding (lanes 2 and 3); where indicated (4×, lane 3) a 4-fold higher K90AR91A protein concentration was used. The $_{gst}$Ros56−142 protein was used as a control (lane 1).

To further confirm that the Ros putative zinc finger motif was involved in DNA binding, we produced a $_{gst}$Ros56−142 double point mutant in which residues Lys90 and Arg91 were substituted with alanine (K90AR91A). These two amino acids, highly conserved in the bacterial Ros homologues (Figure 1) (*19*), are located in the region corresponding to the DNA recognition α-helix in the classical eukaryotic Cys_2His_2 $\beta\beta\alpha$ fold. As shown in Figure 3, the K90AR91A mutant does not bind DNA even when used in a 4-fold molar excess. This result indicates that also residues of the putative zinc finger motif not involved in zinc coordination are important for high-affinity DNA binding.

Definition of the Ros Amino Acid Residues Involved in Zinc Coordination. The sequence analysis of the Ros putative zinc finger clearly suggests that it should represent a novel type of Cys_2His_2 zinc finger domain considering that the loop spacing the second cysteine from the first histidine is 9 amino acids long as opposed to the invariant 12 amino acids in the classical Cys_2His_2 motif present in the eukaryotic proteins. Chou and colleagues (*20*) already reported that the Ros protein is able to bind zinc and that residues Cys82 and His92 are involved in binding a single zinc ion and are essential for DNA interaction. The role of residue Cys79 as the first coordination position of the zinc finger domain remained to

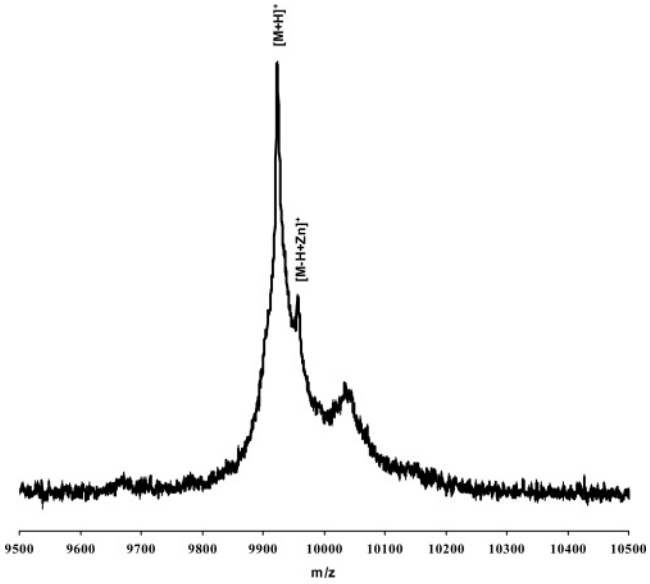

FIGURE 4: MALDI mass spectrum of $_{pet}$Ros56−142. The matrix was 3,5-dimethoxy-4-hydroxycinnamic acid. The solvent was water and acetonitrile (7:3, v/v) at pH 7.0. The spectrum, normalized to the $[M + H]^+$ signal, exhibited both the protonated molecular ion peak, $[M + H]^+$ (*m*/*z* 9934), and the corresponding Zn^{2+} complex, $[M - H + Zn]^+$ (*m*/*z* 9998).

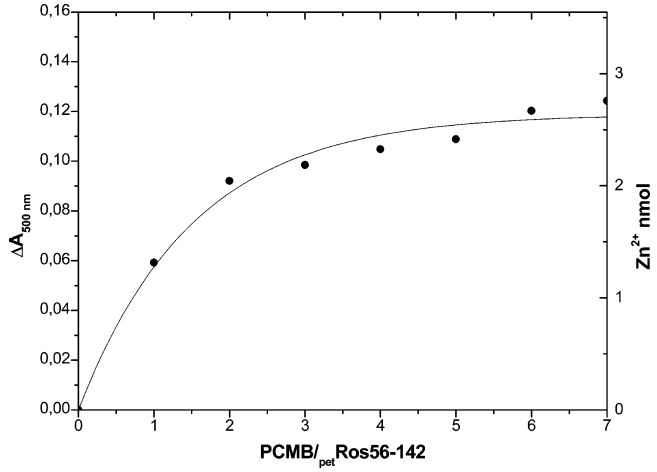

FIGURE 5: PCMB titration of wild-type $_{pet}$Ros56−142 used to determine the number of cysteine residues involved in zinc coordination. $_{pet}$Ros56−142 (2.0 nmol) was incubated in 0.5 mL of a reaction solution containing 40 mM metal-free HEPES-KOH (pH 7.5) and 100 μM PAR. PCMB was added with stepwise addition of 1 equiv of protein. The ratio of PCMB molecules to $_{pet}$Ros56−142 molecules at the different point values is reported along the abscissa. The Zn^{2+} release was monitored by following the increase in absorbance at 500 nm. The right ordinate indicates the number of nanomoles of Zn^{2+} released according to the data of a titration curve obtained using a $ZnCl_2$ standard solution.

be determined and, more importantly, the fourth coordinating position needed to be defined, considering that residues His96 and His97 both occupy a position compatible with the zinc binding role.

Starting from these observations, we decided to perform a complete analysis of the zinc coordination in this protein domain. We first confirmed the presence of the Ros56−142−zinc complex by recording MALDI mass spectra at pH 7.0 and 1.9. In this experiment, the recombinant protein $_{pet}$Ros56−142 was used. The spectrum at neutral pH exhibited both the protonated molecular ion peak, $[M + H]^+$ (*m*/*z* 9934), and the corresponding Zn^{2+} complex, $[M - H + Zn]^+$ (*m*/*z* 9998) (Figure 4), whereas at acidic pH, at which protein side chains are no longer able to bind zinc ion, the mass spectrum exhibited only the protonated molecular ion peak, $[M + H]^+$ (data not shown).

To directly demonstrate the involvement of the cysteine residues in zinc coordination, we utilized a modified version of the protocol previously described by Boyer et al. (*32*) and induced zinc release from the Ros protein by treating the protein with PCMB, a reagent reacting with sulfhydryl groups. Again the $_{pet}$Ros56−142 protein was used so that the only cysteines reacting with PCMB were the two located in the Ros zinc finger motif (see Figure 1A). The zinc release accompanying the reaction of PCMB with the cysteine residues was detected by including the Zn^{2+} complexing dye 4(2-pyridylazo)resorcinol (PAR) in the reaction mixture; Zn(II)PAR_2 has an absorption maximum at 500 nm (*33*). The change in A_{500} during the titration of $_{pet}$Ros56−142 with PCMB in the presence of PAR is shown in Figure 5. The increase in A_{500} almost completely ceases after the addition of 2 equiv of PCMB, indicating that two thiol groups per protein molecule must be titrated by PCMB to induce maximum zinc release and so confirming that two cysteine residues are involved in Zn^{2+} coordination. If compared with the data of a titration curve obtained by using a $ZnCl_2$ standard solution, the total increase in A_{500}, obtained after the addition of 2 equiv of PCMB, corresponds to the release of 1.05 mol of Zn^{2+}/mol of $_{pet}$Ros56−142 (Figure 5) and confirms the binding of one Zn^{2+} per Ros molecule.

To further confirm the role of the zinc coordinating residues in specifying the correct fold of the Ros zinc finger motif, we produced different mutant proteins in which these residues were mutagenized to alanines. First, we produced point mutants of $_{gst}$Ros56−142 in which the putative coordinating residues Cys79, Cys82, and His92 were substituted with alanines (C79A, C82A, and H92A, respectively). The purified proteins were analyzed for their DNA binding activity in a gel mobility shift assay (Figure 6A). As expected, the mutant proteins were not able to bind DNA even when used in a 4-fold molar excess, indicating that these three residues are essential for high-affinity DNA binding and likely represent the first three zinc coordinating positions, all necessary for maintenance of the correct fold of the domain. To determine whether His96 or His97 was the fourth metal coordinating residue, we produced two $_{gst}$Ros56−142 point mutants in which we replaced either of these residues with alanine (H96A and H97A). Surprisingly, when we tested the mutant proteins for their DNA binding ability, both H96A and H97A mutants were able to bind DNA (Figure 6B, lanes 2 and 3). Reasoning that each of the single mutants still had a histidine residue available to coordinate the zinc ion in the fourth position, we produced a double mutant in which both histidines were substituted with alanines (H96AH97A). Interestingly, this mutant was not able to bind the VirC oligonucleotide with high affinity when tested in a gel shift assay (Figure 6B, lanes 4 and 5). These results suggest that either one of the histidine residues may function as the fourth coordinating position alternatively, and only upon elimination of both residues is the protein structural integrity and therefore its DNA binding activity compromised.

To determine how much the different mutations of the coordinating residues affected the ability of the Ros protein

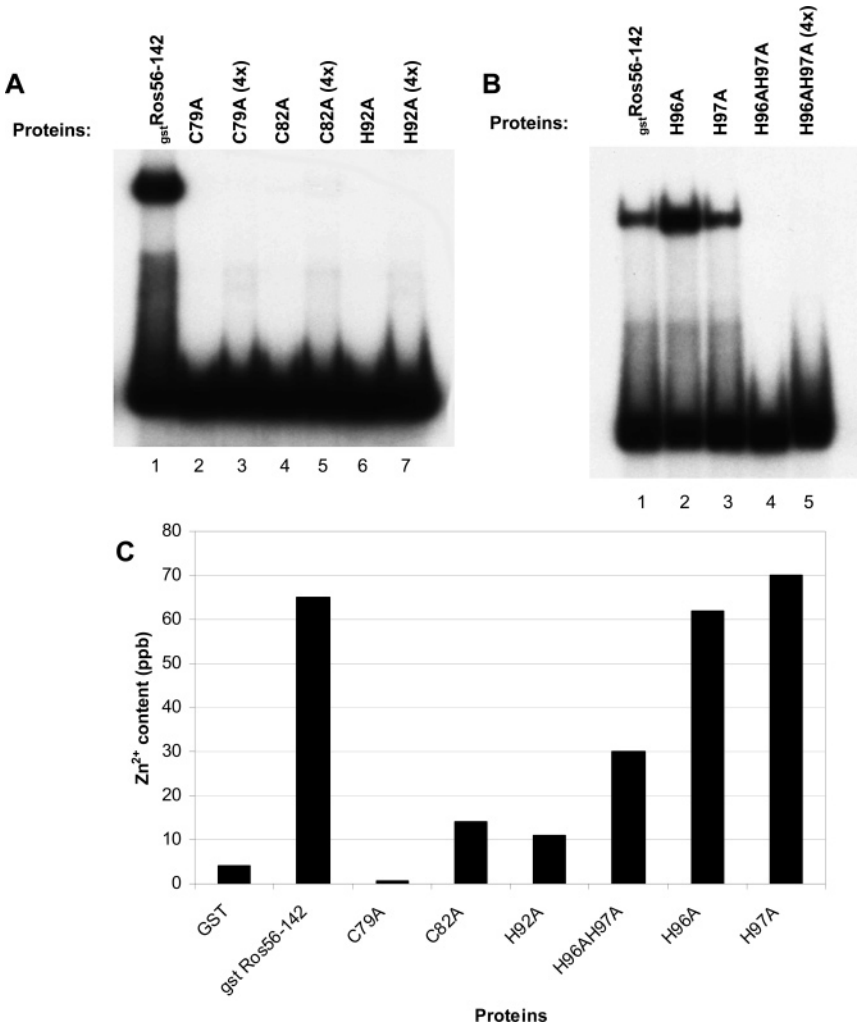

FIGURE 6: Analysis of the role of the potential zinc coordinating residues of the Ros zinc finger domain for protein DNA binding. (A) Gel mobility shift analysis of DNA binding of the C79A, C82A, and H92A mutant proteins (lanes 2−7). The $_{gst}$Ros56−142 protein was used as a control (lane 1), and where indicated (lanes 3, 5, and 7), a 4-fold higher concentration of the different proteins was used. (B) Gel mobility shift analysis of DNA binding of the H96A, H97A, and H96AH97A mutant proteins (lanes 2−5); where indicated (lane 5), a 4-fold higher H96AH97A protein concentration was used. The $_{gst}$Ros56−142 protein was used as a control (lane 1). (C) Zinc content of the $_{gst}$Ros56−142 protein and the different protein mutants as determined by ICP-MS analysis. Considering that all the proteins were expressed as GST fusion proteins, the glutathione *S*-transferase protein was used as a control. The blank value was subtracted for each sample (see Experimental Procedures).

to bind Zn^{2+}, we measured by inductively coupled argon plasma mass spectrometry (ICP-MS) the amount of metal ion in the wild-type $_{gst}$Ros56−142 protein and in all the mutant Ros proteins in which the coordinating residues were substituted with alanine. As shown in Figure 6C, the content of zinc ion was significantly reduced in the C79A, C82A, and H92A mutants and to a lesser extent in the double mutant, H96AH97A, while single mutants H96A and H97A were both still capable of binding zinc. These results confirm the hypothesis that in the Ros zinc finger domain residues Cys79, Cys82, and His92 represent the first three coordinating positions and that the domain can fold using either His96 or His97 as the fourth coordinating residue. The observation that double mutant H96AH97A still binds a certain amount of zinc, even if the overall structure does not allow it to bind DNA and likely is not correctly preserved, is consistent with the previously published results showing that substitution mutants at the final histidine in the Cys_2His_2 classical zinc finger domain retain some ability to fold in the presence of zinc ion (*35*).

To understand in the wild-type Ros protein whether His96 or His97 is indeed involved in the coordination of the zinc ion, we performed a NMR analysis on the $_{pet}$Ros56−142 protein (Figure 7). $^1H-^{15}N$ HSQC experiments of wild-type $_{pet}$Ros56−142 and of a mutant version in which His97 was mutated to alanine ($_{pet}$Ros56−142H97A) were carried out to derive the tautomeric and zinc binding states of the histidine side chains of the two proteins (Figure 7). This experiment allows the coherence transfer from $H_{\epsilon 1}$ and $H_{\delta 2}$ to $N_{\epsilon 2}$ and $N_{\delta 1}$ through the $^2J_{HN}$ coupling constants. The analysis of the two $^1H-^{15}N$ HSQC spectra was based on $_{pet}$Ros56−142 and $_{pet}$Ros56−142H97A chemical shift assignments obtained with standard triple-resonance NMR experiments (R. Fattorusso et al., manuscript in preparation). Four histidines are found in $_{pet}$Ros56−142, three located inside the classical zinc finger consensus sequence and possibly involved in zinc coordination (His92, His96, and His97) and one outside (His76) (see Figure 1A). The pattern of cross-peaks present in the $^1H-^{15}N$ HSQC spectrum (Figure 7, left panel) clearly indicates that in $_{pet}$Ros56−142

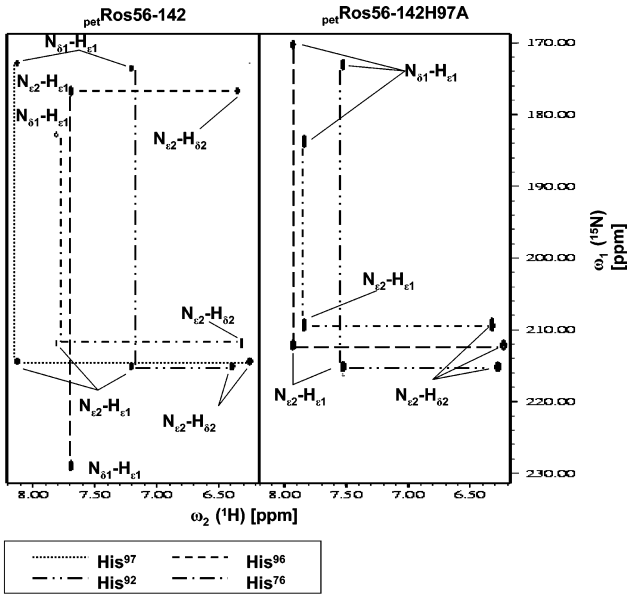

FIGURE 7: Definition of the role of the His residues in zinc coordination by NMR analysis. Portions of the $^{1}H-^{15}N$ HSQC spectra of $_{pet}$Ros56−142 and $_{pet}$Ros56−142H97A acquired at pH 6.8. The assignments of the cross-peaks reveal that His96 is a $N_{\epsilon 2}-H$ tautomer in $_{pet}$Ros56−142 and becomes a $N_{\delta 1}-H$ tautomer in $_{pet}$Ros56−142H97A. Differently, His76 and His92 are $N_{\delta 1}-H$ tautomers in both the proteins as well as His97 in $_{pet}$Ros56−142. Nitrogen chemical shifts indicate also that zinc ion is coordinated by His92 and His97 in $_{pet}$Ros56−142 and by His92 and His96 in $_{pet}$Ros56−142H97A.

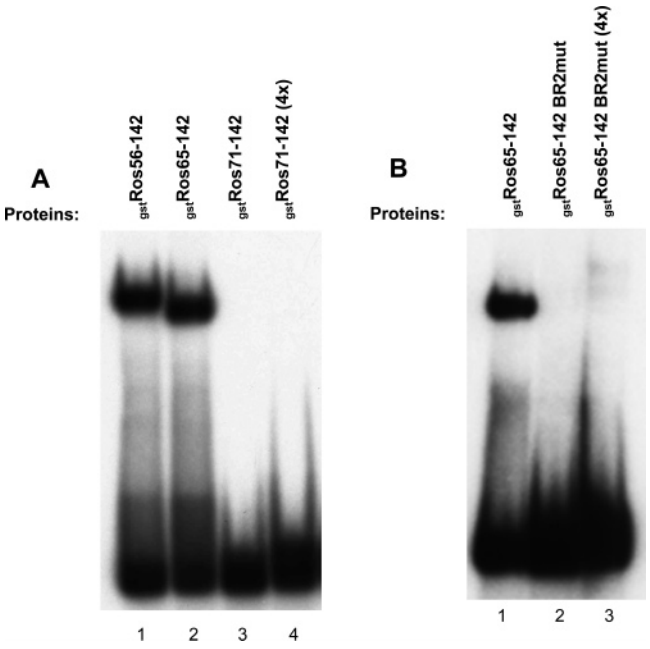

FIGURE 8: Analysis of the role of the basic residues at the N-terminus of the zinc finger domain for Ros DNA binding. (A) Gel mobility shift analysis of DNA binding of the N-terminal deletion mutants $_{gst}$Ros65−142 (lane 2) and $_{gst}$Ros71−142 (lanes 3 and 4); where indicated (lane 4), a 4-fold higher $_{gst}$Ros71−142 protein concentration was used. The $_{gst}$Ros56−142 protein was used as a control (lane 1). (B) Gel mobility shift analysis of DNA binding of $_{gst}$Ros65−142BR2mut (lanes 2 and 3); where indicated (lane 3), a 4-fold higher $_{gst}$Ros65−142BR2mut protein concentration was used. The $_{gst}$Ros65−142 protein was used as a control (lane 1).

His92 and His97 imidazoles are $N_{\delta 1}-H$ tautomers (i.e., with $N_{\delta 1}$ protonated and $N_{\epsilon 2}$ unprotonated), commonly found to bind Zn(II) in the zinc finger domain, with both $N_{\epsilon 2}$ and $N_{\delta 1}$ chemical shifts typical of zinc binding histidine side chains (*36*). Differently, His96 imidazole is a $N_{\epsilon 2}-H$ tautomer, having a $N_{\delta 1}$ chemical shift that is largely out of the range that is typical for unprotonated nitrogens of zinc binding histidine. The His76 side chain is a $N_{\delta 1}-H$ tautomer, but as expected, its $N_{\epsilon 2}$ chemical shift is not characteristic of the zinc coordinating histidine side chain. In the $_{pet}$Ros56−142H97A mutant, analysis of the $^{1}H-^{15}N$ HSQC spectrum (Figure 7, right panel) shows that His76 and His92 retain the same chemical shifts as in the wild-type protein, while His96 becomes a $N_{\delta 1}-H$ tautomer, with both $N_{\epsilon 2}$ and $N_{\delta 1}$ chemical shifts characteristic of zinc binding histidine side chains. These data therefore demonstrate that in $_{pet}$Ros56−142 His92 and His97 occupy the third and fourth positions, respectively, of the zinc coordination whereas in $_{pet}$Ros56−142H97A the fourth position is taken by His96, whose tautomeric state is changed after the zinc binding.

Zinc Finger Flanking Basic Regions Are Essential for Ros DNA Binding. The zinc finger motif present in the Ros protein is flanked on both termini by highly basic stretches of amino acids (BR1, BR2, BR3, BR4, and BR5; see Figure 1A). It has previously been demonstrated that in eukaryotic proteins, such as the SUPERMAN and GAGA proteins, a single Cys_2His_2 zinc finger is capable of high-affinity specific DNA binding but that this binding requires flanking basic amino acids to stabilize the interaction (*17*, *18*). Using a series of deletions and point mutations, the role of all the basic regions present in Ros was investigated. According to what has been observed in other single zinc finger-containing DNA binding proteins, the basic regions are here defined as two or more basic amino acids spaced by no more than one residue.

To determine the role of the amino acids flanking the Ros zinc finger domain on the N-terminal side in stabilizing the interaction with DNA, we analyzed a series of new N-terminal deletion mutants. $_{gst}$Ros65−142 was still able to bind DNA (Figure 8A; compare $_{gst}$Ros56−142, lane 1, with $_{gst}$Ros65−142, lane 2), implying that also the residues from Ala56 to Pro64, including K61 and K63 (BR1), are not necessary for the interaction with DNA. Further amino-terminal deletion resulted in a protein, $_{gst}$Ros71−142, that is not able to bind DNA with high affinity (Figure 8A, lanes 3 and 4). These results indicate that residues 66−70 of the Ros protein are important in stabilizing the DNA interaction. We then focused our attention on basic residues R69 and K70 (BR2), which have been removed in the Ros71−142 deletion mutant. To test if these amino acid residues were important for DNA binding, we prepared a mutant $_{gst}$Ros65−142 protein in which these residues were both substituted with alanine ($_{gst}$Ros65−142BR2mut) and tested its DNA binding activity. As shown in Figure 8B, the mutant protein did not exhibit high-affinity DNA binding. This result implies that basic residues of the Ros protein flanking the finger domain on the N-terminal side (BR2, Figure 1A) are important in stabilizing DNA binding.

We then focused our attention on the basic amino acids located at the C-terminus of the finger to evaluate if they also were part of the DNA binding domain and could play a role in stabilizing the interaction with DNA. First, we prepared a deletion mutant ($_{gst}$Ros56−131) in which all the basic residues located at the C-terminus of the protein (BR5)

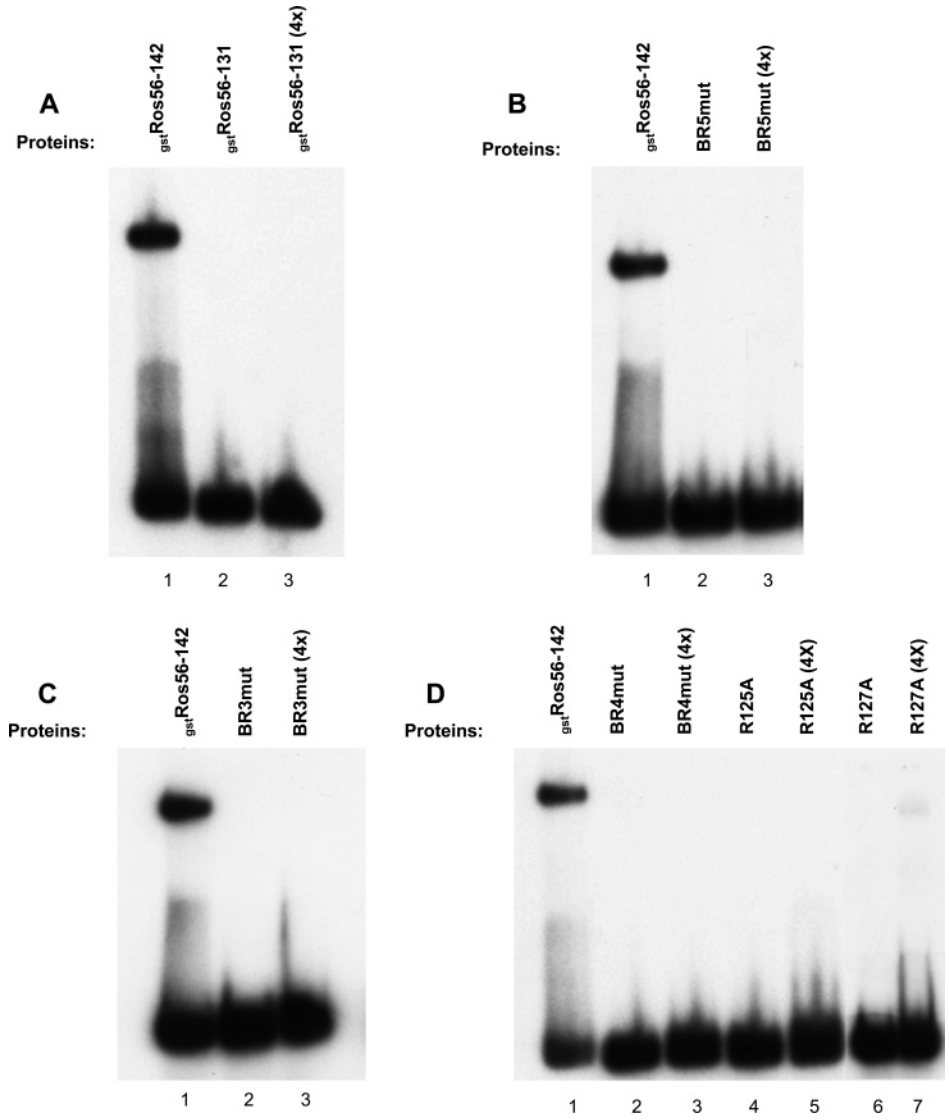

FIGURE 9: Analysis of the role of the basic residues at the C-terminus of the zinc finger domain in Ros DNA binding. (A) Gel mobility shift analysis of DNA binding of the C-terminal $_{gst}$Ros56−131 deletion mutant (lanes 2 and 3). (B) Gel mobility shift analysis of DNA binding of BR5mut (lanes 2 and 3). (C) Gel mobility shift analysis of DNA binding of BR3mut (lanes 2 and 3). (D) Gel mobility shift analysis of DNA binding of BR4mut (lanes 2 and 3), R125A (lanes 4 and 5), and R127A (lanes 6 and 7). In all the experiments, the $_{gst}$Ros56−142 protein was used as a control (lane 1 of all the panels) and where indicated a 4-fold higher protein concentration was used.

were deleted. When tested in a gel mobility shift assay, the $_{gst}$Ros56−131 protein did not bind the VirC oligonucleotide with high affinity (Figure 9A, lanes 2 and 3); similarly, a triple point mutant of $_{gst}$Ros56−142 in which basic residues R137, R138, and K139, deleted in Ros56−131, were substituted with alanine (BR5mut) failed to exhibit specific high-affinity DNA binding (Figure 9B, lanes 2 and 3). These results indicate that the stretch of amino acids which are located at the C-terminus of the protein is important for stabilization of the interaction with the target sequence. To determine the role of the other basic regions located at the C-terminal side of the finger (BR3 and BR4), we prepared point mutants of the $_{gst}$Ros56−142 protein in which the basic amino acids constituting these basic regions were substituted. In BR3mut, the substitutions were R105A and K107A. As shown in Figure 9C, the BR3mut protein does not bind DNA even when tested at a 4-fold higher concentration, indicating that one or both of the mutagenized residues are important for high-affinity DNA binding.

The mutation in the *ros* gene present in the spontaneous mutant 4011R of *A. tumefaciens* has been shown to be a single substitution of one of the basic residues of BR4, Arg125, with a cysteine (*29*). Interestingly, this mutation abolishes the protein DNA binding capability, already suggesting a possible role in DNA binding for Arg125. To confirm the importance of BR4 in the interaction with the DNA target sequence, we prepared three different mutant of the $_{gst}$Ros56−142 protein: a double point mutant, BR4mut, in which alanines were substituted in place of both BR4 wild-type arginines in positions 125 and 127, and two single mutants, R125A and R127A, in which only residues Arg125 and Arg127, respectively, were substituted. As shown in Figure 9D, none of these mutants displayed any DNA binding activity, even when used in a 4-fold molar excess, suggesting that both R125 and R127 are essential for Ros high-affinity DNA binding. By mutating R125 to alanine instead of cysteine, we could exclude the possibility that the addition of an extra cysteine in the natural mutant

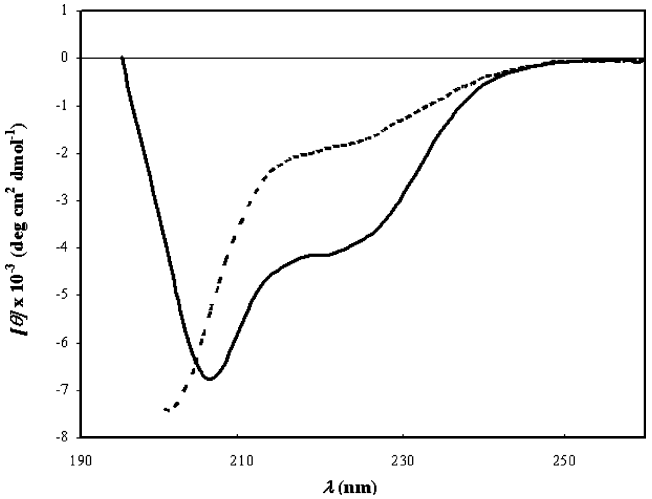

FIGURE 10: CD spectra of $_{pet}$Ros56−142 (0.30 μM) at pH 2.6 (- - -) and 7.0 (—).

4011R inhibited the Ros DNA binding simply by interfering with the correct protein fold around the zinc ion and confirm a more direct involvement of residue R125 in the DNA binding activity of the Ros protein. Taken together, these results imply that basic residues flanking either side of the Ros zinc finger domain are important in stabilizing the DNA binding interaction.

The DNA binding domain present in the $_{gst}$Ros56−142 protein, constituted by a noncanonical Cys_2His_2 zinc finger motif flanked by four basic regions that are important for stabilization of its interaction with DNA, represents a novel type of nucleic acids interaction domain.

To gain more structural information about the $_{pet}$Ros56−142 protein, CD spectra at pH 7.0 and 2.6 (Figure 10) were acquired. The spectrum at pH 7.0 displays two mimima at 207 and 223 nm indicating the presence of an α-helix structure. The deconvolution of this spectrum via a neural network approach indicates a total structural content (α-helix and β-structures) of 68% at pH 7.0. This relatively high content of secondary structure suggests that the Ros region, including residues 56−142, could consist of a domain larger than the classical $\beta\beta\alpha$ zinc finger domain. Addition of HCl to the $_{pet}$Ros56−142 solution up to pH 2.6 produces marked changes in the spectrum: a decrease in the magnitude of the 222 nm minimum and a shift of the 207 nm minimum to 200 nm indicating the loss of the secondary structure. All these changes can be interpreted as the dissociation of the Zn^{2+}−$_{pet}$Ros56−142 complex, which causes the loss of the domain structure integrity.

DISCUSSION

In this paper, we present the first complete functional characterization of the bacterial Ros transcriptional regulator DNA binding domain. This domain is constituted by a single Cys_2His_2 zinc binding motif flanked by four basic regions, one located at the N-terminus and three at the C-terminus, which all appear to be essential for DNA binding. Interestingly, basic amino acids have been shown to be important in stabilizing the DNA binding of other types of single zinc finger domains such as the classical Cys_2His_2 zinc finger present in the *Drosophila* GAGA factor and in the *A. thaliana* SUPERMAN protein and the Cys_4 zinc finger present in the GATA proteins (*10*, *17*, *18*, *37*, *38*). These results, taken together, confirm that the modality of DNA binding using a single zinc finger motif flanked by basic residues is widespread throughout the living kingdom from eukaryotic, both animal and plant, to prokaryotic, even if in each kingdom it presents its peculiarity.

The zinc binding motif present in the Ros DNA binding domain differs from the classical Cys_2His_2 motif in that there are 9 amino acids instead of 12 in the loop which separates the second cysteine from the first histidine (see Figure 1A). Moreover, the Ros zinc binding motif contains three histidine residues, two of which are adjacent and occupy a position compatible with the fourth zinc coordinating position. In this paper, we present an accurate analysis to define the residues involved in binding the zinc ion in this new type of zinc binding domain. By showing that chemical modification of Cys79 and Cys82 induced the release of zinc from the Ros protein and substitutions with alanines of Cys79, Cys82, and His92 residues resulted in the loss of zinc and DNA binding activity, we clarified that these residues are indeed involved in zinc coordination. Interestingly, single mutations of either His96 and His97 do not alter the zinc and DNA binding ability of the Ros protein, whereas mutation of both residues abolishes the DNA binding activity and strongly reduces the zinc coordination capability. These results suggest that either of the histidine residues, His96 or His97, may function as the fourth coordinating position alternatively, and only upon elimination of both residues is the protein DNA binding activity compromised. By NMR spectral analysis, we demonstrated that in the Ros wild-type protein, where both His96 and His97 residues are present, the histidines involved in zinc coordination are His92, as the third coordinating residue, and His97 as the fourth, and both bind the ion through their $N_{\epsilon 2}$ atom. Moreover, when His97 is mutated in Ala, His96 is able to occupy the fourth position of the zinc coordination, changing its tautomeric form from the $N_{\epsilon 2}$−H tautomer, observed in the wild-type protein, to the $N_{\delta 1}$−H tautomer (Figure 7). Therefore, it is not surprising that Ros homologues in which either the second or third histidine is absent have been identified (*19*), confirming that both can function in coordinating zinc, conferring a correct protein folding. Taken together, our results for the zinc coordination residues present in the Ros protein confirm that the one present in Ros is a Cys_2His_2 zinc binding motif. Just as in the eukaryotic Cys_2His_2 zinc finger domains, where the spacer between the first and second histidine can vary from three to five amino acids (*1*), also in the prokaryotic domains the correct folding is possible with a variable spacer between the third and the fourth zinc coordination position (three to four residues).

By showing that the mutant protein K90AR91A, in which Lys90 and Arg91 of the zinc finger motif were substituted with alanines, does not bind DNA, we have demonstrated that not only the zinc coordinating residues but also amino acids located in the putative DNA recognition helix are essential for DNA binding. Moreover, in the case of the Ros protein, we have proved that amino acids included in a region of ∼70 amino acids, comprising the four basic stretches and the zinc binding motif, are part of the DNA binding domain. So far, in neither of the eukaryotic single zinc finger-containing DNA binding motif have so many flanking basic stretches been demonstrated to be necessary for high-affinity

DNA binding. Interestingly, the basic regions that we have identified as being important for Ros DNA binding are all completely or partially conserved in the other bacterial Ros homologues (ref *21* and Figure 1B), confirming a similar DNA binding modality between these proteins. The exact role of the basic residues as well as the role of the residues present in the zinc finger motif in stabilizing the interaction with the DNA target sequence remains to be defined. In fact, for all of these residues, whether they play a role in contacting the DNA or contribute in maintaining the conformational integrity of the domain must be determined. In this regard, our CD data suggest that a large part of the Ros DNA binding domain has a defined secondary structure. This would indicate not only that the zinc finger region is folded in the absence of DNA but also that other secondary structure elements might be present in the Ros56−142 region. Whether the zinc binding region present in the Ros protein and in the other bacterial homologues is indeed a domain, able to fold in the presence of zinc in a manner independent of the rest of the protein, remains to be determined. Considering all the differences between the bacterial Cys_2His_2 zinc binding motif and the eukaryotic counterpart evidenced by our biochemical and preliminary structural data, in fact, we can now hypothesize that the zinc finger present in these bacterial proteins is part of a larger domain that could be significantly different from the ones so far characterized. Structural characterization of the Ros DNA binding domain will be of fundamental interest in clarifying the differences between the prokaryotic and eukaryotic zinc finger binding modules and will help in elucidating the mechanism of the interaction with DNA.

ACKNOWLEDGMENT

We are grateful to Prof. Mauro Iuliano for his help in ICP-MS analysis and to Maurizio Muselli and Dr. Vincenzo Piscopo for their excellent technical assistance.

REFERENCES

1. Klug, A., and Schwabe, J. W. R. (1995) Protein motifs 5. Zinc finger, *FASEB J. 9*, 597−604.
2. Tupler, R., Perini, G., and Green, M. R. (2001) Expressing the human genome, *Nature 409*, 832−833.
3. Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratclie, O. J., Samaha, R. R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J. Z., Ghandehari, D., Sherman, B. K., and Yu, G. (2000) *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes, *Science 290*, 2105−2110.
4. Miller, J., McLachlan, A. D., and Klug, A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes, *EMBO J. 4*, 1609−1614.
5. Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000) DNA recognition by Cys2His2 zinc finger proteins, *Annu. Rev. Biophys. Biomol. Struct. 29*, 183−212.
6. Laity, J. H., Lee, B. M., and Wright, P. E. (2001) Zinc finger proteins: New insights into structural and functional diversity, *Curr. Opin. Struct. Biol. 11*, 39−46.
7. Pavletich, N. P., and Pabo, C. O. (1991) Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 Å, *Science 252*, 809−817.
8. Pavletich, N. P., and Pabo, C. O. (1993) Crystal structure of a five-finger GLI-DNA complex: New perspectives on zinc fingers, *Science 261*, 1701−1707.
9. Fairall, L., Schwabe, J. W. R., Chapman, L., Finch, J. T., and Rhodes, D. (1993) The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition, *Nature 366*, 483−487.
10. Omichinski, J. G., Pedone, P. V., Felsenfeld, G., Gronenborn, A. M., and Clore, G. M. (1997) The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode, *Nat. Struct. Biol. 4*, 122−130.
11. Houbaviy, H. B., Usheva, A., Shenk, T., and Burley, S. K. (1996) Cocrystal structure of YY1 bound to the adeno-associated virus P5 initiator, *Proc. Natl. Acad. Sci. U.S.A. 93*, 13577−13582.
12. Nolte, R. T., Conlin, R. M., Harrison, S. C., and Brown, R. S. (1998) Differing roles for zinc fingers in DNA recognition: Structure of a six-finger transcription factor IIIA complex, *Proc. Natl. Acad. Sci. U.S.A. 95*, 2938−2943.
13. Takatsuji, H. (1998) Zinc-finger transcription factors in plants, *Cell. Mol. Life Sci. 54*, 582−596.
14. Takatsuji, H. (1999) Zinc-finger proteins: The classical zinc finger emerges in contemporary plant science, *Plant Mol. Biol. 39*, 1073−1078.
15. Englbrecht, C. C., Schoof, H., and Bohm, S. (2004) Conservation, diversification and expansion of C2H2 zinc finger proteins in the *Arabidopsis thaliana* genome, *BMC Genomics 5*, 39.
16. Isernia, C., Bucci, E., Leone, M., Zaccaro, L., Di Lello, P., Digilio, G., Esposito, S., Saviano, M., Di Blasio, B., Pedone, C., Pedone, P. V., and Fattorusso, R. (2003) NMR structure of the single QALGGH zinc finger domain from the *Arabidopsis thaliana* SUPERMAN protein, *ChemBioChem 4*, 171−180.
17. Pedone, P. V., Ghirlando, R., Clore, G. M., Gronenborn, A. M., Felsenfeld, G., and Omichinski, J. G. (1996) The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high affinity specific DNA binding, *Proc. Natl. Acad. Sci. U.S.A. 93*, 2822−2826.
18. Dathan, N., Zaccaro, L., Esposito, S., Isernia, C., Omichinski, J. G., Riccio, A., Pedone, C., Di Blasio, B., Fattorusso, R., and Pedone, P. V. (2002) The *Arabidopsis* SUPERMAN protein is able to specifically bind DNA through its single Cys2-His2 zinc finger motif, *Nucleic Acids Res. 30*, 4945−4951.
19. Bouhouche, N., Syvanen, M., and Kado, C. I. (2000) The origin of prokaryotes C2H2 zinc finger regulators, *Trends Microbiol. 8*, 77−81.
20. Chou, A. Y., Archdeacon, J., and Kado, C. I. (1998) *Agrobacterium* transcriptional regulator Ros is prokaryotic zinc finger protein that regulates the plant oncogene *ipt*, *Proc. Natl. Acad. Sci. U.S.A. 95*, 5293−5298.
21. Kado, C. I. (2002) Negative transcriptional regulation of virulence and oncogenes of the Ti plasmid by bearing a consrved C_2H_2− zinc finger motif, *Plasmid 48*, 179−185.
22. Tiburtius, A., De Luca, N. G., Hussian, H., and Johnston, A. W. B. (1996) Expression of the *exoY* gene, required for exopolysaccharide synthesis in *Agrobacterium*, is activated by the regulatory *ros* gene, *Microbiology 142*, 2621−2629.
23. Bittinger, M. A., Milner, J. L., Saville, B. J., and Handelsman, J. (1997) *rosR*, a determinant of nodulation competitivness in *Rhizobium etli*, *Mol. Plant−Microbe Interact. 10*, 80−186.
24. Bertram-Dogratz, P. A., Quester, I., Becker, A., and Puhler, A. (1998) The *Sinorhizobium meliloti* MucR protein, which is essential for the production of high-molecular-weight succinoglycan exopolysaccharide, binds to short DANN regions upstream of *exoH* and *exoY*, *Mol. Gen. Genet. 257*, 433−441.
25. Keller, M., Roxlau, A., Wenig, W. M., Schmidt, M., Quandt, J., Niehaus, K., Jording, D., Arnold, W., and Puhler, A. (1995) Molecular analysis of the *Rhizobium meliloti mucR* gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan, *Mol. Plant−Microbe Interact. 8*, 267−277.
26. Freiberg, C., Fellay, R., Bairoch, A., Broughton, W. J., Rosenthal, A., and Perret, X. (1997) Molecular basis of symbiosis between *Rhizobium* and legumes, *Nature 387*, 394−401.
27. Romine, M. F., Stillwell, L. C., Wong, K. K., Thurston, S. J., Sisk, E. C., Sensen, C., Gaasterland, T., Fredrickson, J. K., and Saffer, J. D. (1999) Complete sequence of a 184 kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199, *J. Bacteriol. 181*, 1585−1602.
28. D'Souza-Ault, M. R., Cooley, M. B., and Kado, C. I. (1993) Analysis of the Ros repressor of *Agrobacterium virC* and *virD* operons: Molecular intercommunication between plasmid and chromosomal genes, *J. Bacteriol. 175*, 3486−3490.
29. Archdeacon, J., Bouhouche, N., O'Connel, F., and Kado, C. I. (2000) A single amino acid substitution beyond the C_2H_2-zinc finger in Ros derepresses virulence and T-DNA genes in *Agrobacterium tumefaciens*, *FEMS Microbiol. Lett. 187*, 175−178.
30. White, B. A. (1993) *Methods in Molecular Biology*, Vol. 15, Humana Press, Totowa, NJ.

31. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
32. Boyer, P. D. (1954) Spectrophotometric study of the reaction of protein sulfhydryl groups with organic mercurials, *J. Am. Chem. Soc. 76*, 4331−4337.
33. Tang, W., and Wang, C. C. (2001) Zinc finger and thiol-disulfide oxidoreductase activities of chaperone DnaJ, *Biochemistry 40*, 14985−14994.
34. Bohm, G., Muhr, R., and Jaenicke, R. (1992) Quantitative analysis of protein far-UV circular dichroism spectra by neurak networks, *Protein Eng. 5*, 191−195.
35. Simpson, R. J. Y., Cram, E. D., Czolij, R., Matthews, J. M., Crossley, M., and Mackay, J. P. (2003) CCHX zinc finger derivatives retain the ability to bind Zn(II) and mediate protein-DNA interaction, *J. Biol. Chem. 278*, 28011−28018.
36. Urbani, A., Bazzo, R., Nardi, M. C., Cicero, D. O., De Francesco, R., Steinkühler, C., and Barbato, G. (1998) The metal binding site of the hepatitis C virus NS3 protease. A spectroscopic investigation, *J. Biol. Chem. 273*, 18760−18769.
37. Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., and Gronenborn, A. M. (1993) NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1, *Science 261*, 438−446.
38. Pedone, P. V., Omichinski, J. G., Nony, P., Trainor, C., Gronenborn, A. M., Clore, G. M., and Felsenfeld, G. (1997) The N-terminal fingers of chicken GATA-2 and GATA-3 are independent sequence-specific DNA binding domains, *EMBO J. 16*, 2874−2882.

BI060697M