A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity

Claudia A. Blindauer*, Mark D. Harrison†, John A. Parkinson*, Andrea K. Robinson†, Jennifer S. Cavet†, Nigel J. Robinson†‡, and Peter J. Sadler*‡

*Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, United Kingdom; and †Department of Biochemistry and Genetics, University of Newcastle, Newcastle NE2 4HH, United Kingdom

Edited by Richard D. Palmiter, University of Washington School of Medicine, Seattle, WA, and approved June 19, 2001 (received for review March 12, 2001)

Zinc is essential for many cellular processes, including DNA synthesis, transcription, and translation, but excess can be toxic. A zinc-induced gene, *smtA*, is required for normal zinc-tolerance in the cyanobacterium *Synechococcus* PCC 7942. Here we report that the protein SmtA contains a cleft lined with Cys-sulfur and His-imidazole ligands that binds four zinc ions in a $Zn_4Cys_9His_2$ cluster. The thiolate sulfurs of five Cys ligands provide bridges between the two $ZnCys_4$ and two $ZnCys_3His$ sites, giving two fused six-membered rings with distorted boat conformations. The inorganic core strongly resembles the Zn_4Cys_{11} cluster of mammalian metallothionein, despite different amino acid sequences, a different linear order of the ligands, and presence of histidine ligands. Also, SmtA contains elements of secondary structure not found in metallothioneins. One of the two Cys_4-coordinated zinc ions in SmtA readily exchanges with exogenous metal (^{111}Cd), whereas the other is inert. The thiolate sulfur ligands bound to zinc in this site are buried within the protein. Regions of β-strand and α-helix surround the inert site to form a zinc finger resembling the zinc fingers in GATA and LIM-domain proteins. Eukaryotic zinc fingers interact specifically with other proteins or DNA and an analogous interaction can therefore be anticipated for prokaryotic zinc fingers. SmtA now provides structural proof for the existence of zinc fingers in prokaryotes, and sequences related to the zinc finger motif can be identified in several bacterial genomes.

The cyanobacterium *Synechococcus* PCC 7942 synthesizes a 56-aa cysteine-rich protein, SmtA, on exposure to elevated concentrations of metals (Zn, Cd, and Cu; refs. 1 and 2). Zinc is the most potent inducer of transcription from the *smtA* operator promoter, and repression of *smtA* in the absence of elevated zinc is mediated by the zinc-sensing, DNA-binding product of the divergently transcribed gene, *smtB* (3). Mutants lacking the *smtA* locus are hypersensitive to elevated zinc (5-fold), but otherwise viable (4). In addition to sequestering excess, it has been postulated that SmtA may be involved in controlling the intracellular distribution of zinc (5), but a specific partner (protein or DNA) for interaction with SmtA has not yet been identified.

SmtA has been classified as a bacterial metallothionein (MT), although its sequence is very different from eukaryotic MTs. It is unclear whether these proteins are in any way functionally analogous, because the function of neither has been fully defined. Metallothioneins are ubiquitous in eukaryotes and are small proteins containing about 60 amino acids (6). The solution structures of several vertebrate and invertebrate MTs have been determined by multidimensional NMR techniques (7). Such studies have provided the basis for understanding the dynamic behavior of MTs, especially of their metal ion uptake and release reactions. Mammalian isoforms MT-1 and -2 are commonly isolated with seven Zn^{2+} ions bound exclusively to the thiolate groups of cysteine residues in two clusters (8). In view of the cellular importance of zinc, there is considerable interest in establishing the conditions in which, and locations where, MT acquires or releases zinc (5, 9–13).

SmtA contains a smaller proportion of Cys than mammalian MT (16% vs. 33% of residues), and, unlike mammalian MTs, contains His residues, which have been implicated in metal coordination (14). The Cys coordination pattern for the MT clusters cannot be mapped onto the sequence of SmtA. Genome sequences have recently revealed *smtA*-related genes in several bacteria, including pathogens, and we previously detected an MT-like protein in *Pseudomonas putida* from a metal-polluted site (15). To investigate the analogy between SmtA and MT, and to understand better how bacterial MTs buffer excess heavy metals, we have determined the solution structure of SmtA and investigated its metal uptake and exchange behavior.

BIOCHEMISTRY

Materials and Methods

Protein Expression and Purification. To express SmtA in *E. coli*, an *Nde*I restriction site was introduced at the 5′ end of the SmtA coding region in pMDNR1.1 (14) by using Quik-change mutagenesis (Stratagene). The resulting *Nde*I/*Bam*HI fragment was subcloned into pET29a (Novagen), creating pMHNR1.1. Aliquots (5 ml) of cleared lysate from cells carrying pMHNR1.1 were applied to a Sephadex G-50 column (2 × 60 cm; Amersham Pharmacia) equilibrated with 25 mM Tris (pH 8.8). Fractions containing SmtA were pooled and applied to a HiTrapQ column (1 ml; Amersham Pharmacia) equilibrated with 25 mM Tris (pH 8.8). SmtA was eluted with a 40-ml 0–100 mM NaCl gradient. The purity of SmtA was confirmed by SDS/PAGE and HPLC. N-terminal amino acid sequencing showed that the protein began with TSTTL, indicating excision of the terminal Met in *E. coli*. ^{15}N-labeled SmtA was prepared from cells grown in minimal medium supplemented with $^{15}NH_4Cl$-containing M9 salts.

Preparation of Cd_4SmtA. Homogeneous Cd^{2+}-loaded SmtA was prepared by acidification of Zn_4SmtA to pH 1 under nitrogen, separation of the apoprotein on a Sephadex G-25 column (prepacked Amersham Pharmacia PD-10, equilibrated with 0.01 M HCl), and reconstitution in the presence of a slight excess of $^{111}Cd^{2+}$ by raising the pH to 7.0 with Tris base (16). Excess Cd^{2+} was removed with Chelex-100 resin (Bio-Rad).

Zn–Cd Exchange. Zn_4SmtA was incubated in 50 mM Tris·HCl buffer containing 50 mM NaCl (all reagents Fisher Scientific, analytical grade) for periods of 6 h to 7 days and at pH values from 5 to 8, at

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ESI, electrospray ionization; ICP-AES, inductively coupled plasma–atomic emission spectrometry; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; rmsd, rms deviation; amu, atomic mass unit; MT, metallothionein; HSQC, heteronuclear single quantum coherence.

Data deposition: The atomic coordinates for Zn_4SmtA have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 1JJD).

‡To whom reprint requests should be addressed. E-mail: n.j.robinson@ncl.ac.uk or p.j.sadler@ed.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

ambient temperature with various mol equivalents of $^{111}CdCl_2$ [prepared by dissolution of ^{111}CdO (95.29%; Oak Ridge National Laboratory, TN) in 1 M HCl]. Excess metal was removed on a Sephadex G-25 column (prepacked Amersham Pharmacia PD-10) by elution with 50 mM Tris·HCl buffer (50 mM NaCl, pH 7.0). Typically, the ^{111}Cd NMR spectrum was recorded *ca.* 30 min. after the addition of Cd^{2+}. The incubation time had little effect on the NMR spectra.

Metal Determinations. Samples of SmtA were diluted to 5 ml with 50 mM Tris·HCl buffer and 50 mM NaCl. A Thermo Jarrell Ash IRIS inductively coupled plasma–atomic emission spectrometry (ICP-AES) instrument was calibrated by using five sulfur, zinc, or cadmium standard solutions in the same buffer, prepared from 1,000 ppm standards (Fisher Scientific), and the metal ion and sulfur contents were determined simultaneously, allowing the direct determination of the protein:metal ratio.

Mass Spectrometry. Samples were prepared by buffer exchange into 50 mM NH_4Ac (Fisher Scientific) by using Amicon YM-3 centrifugation filters. Positive ion electrospray ionization (ESI)-MS was performed on a Micromass Platform II instrument on samples diluted to 40 pmol μl^{-1} in 10% methanol. Samples at pH 2, 5, 7, and 10 were studied. MASS-LYNX (Version 2.3) software was used for deconvolution of the spectra.

Liquid Chromatography. The homogeneity of the protein preparations was determined on a Hewlett Packard 1100 Series LC System, employing a C_{18} reverse-phase column (250 × 4.6 mm, Lichrosorb, 5 μm diameter (BAS Technical, Congleton, UK); detection at 220 nm). A 10-μl aliquot of 15 μM Zn_4SmtA in 10 mM Tris·HCl (pH 7.5) was eluted over 25 min with buffer A and the gradients: 0–3 min, 0–12% buffer B; 3–15 min, 12–18% buffer B; 15–25 min, 18–20% buffer B. [buffer A, aqueous 10 mM Tris·HCl; buffer B, 60% acetonitrile (Fisher Scientific, HPLC grade)/10 mM Tris·HCl].

NMR Spectroscopy. Samples of SmtA (typically 3–4 mM) were prepared by buffer exchange into 50 mM [D_{11}]Tris·HCl and 50 mM NaCl in 90% (vol/vol) H_2O/D_2O. Three-dimensional [^1H, ^1H, ^{15}N] nuclear Overhauser enhancement spectroscopy–heteronuclear single quantum coherence (NOESY-HSQC), total correlation spectroscopy (TOCSY)–HSQC, and HNHA data were acquired by using a uniformly ^{15}N-labeled sample (0.4 mM) in 90% H_2O/D_2O. All NMR spectra were acquired at pH 7.0 and 308 K on Varian UNITY INOVA (operating at 599.82 MHz for ^1H) or Bruker DMX500 (operating at 500.13 MHz for ^1H) NMR spectrometers by using triple resonance (^1H, ^{13}C, ^{15}N) z-gradient probe heads. Two-dimensional (2D) [^1H,^1H] double quantum filtered (DQF)–correlated spectroscopy (COSY; ref. 17), TOCSY (65 ms spin-lock time; ref. 18), and NOESY (τ_m = 70, 100, and 130 ms; ref. 19) data were acquired with 64, 32, and 32 transients, respectively, for each of 2 × 512 t_1 increments [hypercomplex acquisition (20)] using TPPI (time proportional phase incrementation) into 4 k complex data points. The spectral width was 8 kHz, and the real Fourier transform (FT) was carried out on 2 k × 2 k data points. Water suppression was achieved by double-pulsed field-gradient spin-echo (DPFGSE; ref. 21).

Three-dimensional [^1H, ^1H, ^{15}N] TOCSY-HSQC (65 ms spin-lock time) and NOESY-HSQC (τ_m = 100 ms; ref. 22) NMR data were acquired with eight transients for each of 2 × 128 t_1 increments and 2 × 40 t_2 increments into 1 k complex points. Spectral widths were 5 kHz in F1, 8 kHz in F3, and 2.2 kHz in F2. The data were Fourier transformed into 1 k × 512 × 64 data points.

$^3J_{HN,HA}$ coupling constants were estimated by using a ^{15}N-HNHA experiment (23). Sixteen transients were acquired for each of 2 × 128 t_1 increments and 2 × 32 t_2 increments into 2 k complex data points, using spectral widths of 8, 8, and 2.3 kHz, respectively. The data were transformed into 1 k × 512 × 64 data points.

One-dimensional ^{111}Cd-{^1H} NMR spectra [106.04 MHz, Bruker (Ettlingen, Germany) DMX500, 4 k complex points] were acquired on a BBO (direct observe) probe head with 64 k transients (acquisition time 0.06 s, recycle delay 0.34 s). Longer acquisition times or delays did not alter the spectra. Metal-ion-to-ligand connectivities were determined by 2D [^1H,^{111}Cd] HSQC and 2D [^1H,^{111}Cd] HSQC-TOCSY NMR experiments: 256 transients were acquired for each of 128 t_1 increments into 2 k complex points, over a frequency width of 6.5 kHz in F2 and 28.6 kHz in F1. $^3J_{^1H,^{111}Cd}$ was varied over the range from 5 Hz to 55 Hz. Connectivities to His residues were best observed at $^3J_{^1H,^{111}Cd}$ = 5–15 Hz, and connectivities to Cys β protons at $^3J_{^1H,^{111}Cd}$ = 25–55 Hz.

Comparison of ^{111}Cd spectra at 106.04 and 76.37 MHz processed with optimum apodization functions, together with homonuclear decoupling experiments, showed that ^{111}Cd–^{111}Cd couplings are present. A reasonable interpretation of the splittings gave $^2J_{Cd(A),Cd(B)}$ = 45 Hz and $^2J_{Cd(B),Cd(C)}$ = 32 Hz. However, a full analysis of these couplings was complicated by the low signal-to-noise ratio and line-broadening.

All NMR data were processed with XWINNMR software (Version 2.0 or 2.1, Bruker), and analyzed with the aid of the NMR Triad module of SYBYL (Version 6.3, Tripos Associates, St. Louis).

Table 1. NMR and structure statistics for Zn_4SmtA

Restraints for structure calculations (residues 5–56)	
Total restraints	380
Total NOE restraints	291
Intraresidue	41
Sequential	79
Medium range	49
Long range	122
Dihedral angles (ϕ)	37
Zinc distance restraints	52
Structure statistics	
Number of starting structures	500
Number of final structures	20
Average of DIANA target function	0.67 ± 0.13
Number of violations > 0.25 Å	none
rmsd to the average structure (residues 5–56), Å	
Backbone	0.90 ± 0.11
All heavy atoms	1.48 ± 0.14

Structure Calculations. Distance restraints (see Table 1) were derived from a 100-ms 2D NOESY experiment on an unlabeled sample. Initially, about 1,000 crosspeaks were picked, and 600 thereof were unambiguously assigned. The intensities of the NOE crosspeaks were classified as strong (1.8–3.0 Å), medium (1.8–4.0 Å), or weak (1.8–5.0 Å). Pseudoatom corrections were added to the upper distance limits where appropriate. Initial DIANA (24) calculations classified 291 distance restraints as being relevant. The three-dimensional HNHA experiment yielded 37 ϕ dihedral angle restraints via the determination of $^3J_{HN,HA}$ coupling constants, which were divided into four groups ($^3J_{HN,HA}$ < 5 Hz: $-40° < \phi < -80°$; 5 Hz < $^3J_{HN,HA}$ < 7 Hz: $-50° < \phi < -100°$; 7 Hz < $^3J_{HN,HA}$ < 9 Hz: $-60° < \phi < -120°$; $^3J_{HN,HA}$ > 9 Hz: $-75 < \phi < -150°$. Stereospecific assignments for the β protons of a few residues (Asn-17, Tyr-31, Cys-47, and His-49) were determined by using the program HABAS (24), which was also used for the validation of the values of the dihedral angles ϕ. Structures were calculated by using the SYBYL built-in version of DIANA (Version 2.2.1), which makes use of a distance-geometry algorithm. Two cycles of REDAC (Redundant Dihedral Angle Constraint algorithm) were applied to optimize a total of 500 starting structures to yield the 20 best

structures. Initial structures were calculated without any assumptions about zinc coordination, but including modified Cys residues containing a dummy atom instead of H(γ) (25). The subsequent calculations included upper and lower distance limits for zinc and ligating sulfur and nitrogen atoms. The Zn-S and Zn-N distances were restrained according to published values (26, 27). Lower limits for S-S distances were used, but no angle restraints were applied. Both metal ion-derived and NOE-derived restraints were considered with equal weight. Other ligation schemes were tested, but none gave satisfactory results. The average structures were obtained by 200 steps of a force-field-based energy minimization procedure within the program SYBYL, using the Tripos force-field. Statistical analyses were performed within SYBYL and MOLMOL (Version 2k.1; ref. 28). Protein cartoons were also prepared by using MOLMOL.

Sequence and Structure Comparisons. The Swiss-Pdb Viewer (Version 3.51; http://expasy.cbr.nrc.ca/spdbv/) was used to construct the overlay of GATA-1 (PDB ID code 1gat), CRP1 LIM2 (PDB ID code 1b8t), and SmtA structures.

Results and Discussion

SmtA Contains Four Zinc Ions. Under acidic conditions, the most abundant peak in the positive ion electrospray mass spectrum (ESI-MS) of apo-SmtA was at 5610 atomic mass units (amu), corresponding to apo-SmtA lacking the N-terminal Met residue (calculated 5610.3). For pH values above 5, the most abundant peak (5863 amu) corresponded to Zn_4SmtA (calculated 5862.8 assuming that nine protons are released upon metal binding). N-terminal sequencing and one-dimensional NMR spectroscopy also confirmed the absence of the N-terminal Met residue. The zinc and sulfur contents of SmtA were determined by ICP-AES. On the basis of the presence of nine sulfurs per mol protein, the Zn content was determined to be 4.4 $\pm$ 0.4 Zn per mol. Other metals (Cu and Cd) were not detected. Analysis of SmtA by titration with the organomercurial *p*-(hydroxymercuri)phenylsulphonate (PMPS) and 4-(2)-(pyridylazo)resorcinol (PAR) also gave a zinc content of four Zn^{2+} ions per mol protein (A.K.R., M.D.H., C.A.B., J.A.P., P. Bowness, P.J.S., and N.J.R., manuscript in preparation), a greater zinc content than observed previously for SmtA produced as a glutathione *S*-transferase fusion protein (14).

Backbone and Side Chain Structure. A combination of standard 2D [^{1}H] and three-dimensional [^{15}N, ^{1}H, ^{1}H] NMR experiments on Zn_4SmtA (residues 2–56, Fig. 1*a*) allowed a complete assignment of peaks for residues 5–56 (see Table 2 and Fig. 5, which are published as supplemental data on the PNAS web site, www.pnas.org), with the exception of residues Gly-42, Thr-50, and Gly-51, for which no NH resonances were detected. Residues 2–4 at the N terminus were not detected in NOE experiments, probably due to the high flexibility of the N terminus, and were not considered in the structure calculations. The structure of Zn_4SmtA was initially determined without incorporation of Zn^{2+} ions in the calculation. The ensemble of the best 20 structures (Fig. 1*b*) shows that the backbone is well defined for a substantial part of the protein (residues 6–40), whereas some loop regions (e.g., 41–45) are more flexible and do not give rise to long-range NOE crosspeaks.

The structure derived from these NMR constraints, even with no specific attempt to incorporate Zn^{2+} ions, contains a cleft lined with the side chains of cysteine and histidine residues, potential ligands for zinc. The side chains of charged or polar residues (Arg, Lys, Asx, Glu) are located on the surface of the protein.

Identification of Zn Ligands Through Cd Replacement. It is not possible to determine from ^{1}H NMR data alone which protein side chains are bound to the four zinc ions, because no NMR connectivities can be made between protons and quadrupolar Zn^{2+} ions. Therefore, the native Zn^{2+} ions were replaced by $^{111}Cd^{2+}$ ions (nuclear spin $I = 1/2$), an approach used successfully for MTs (16).

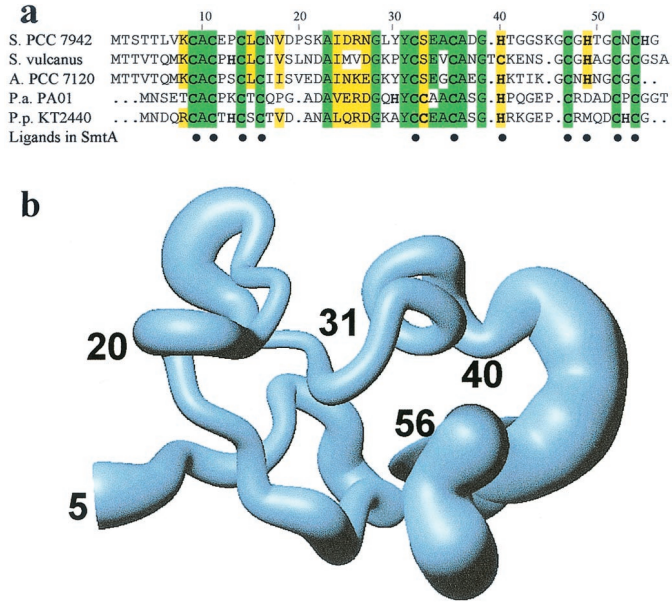

Fig. 1. Sequence and backbone structure of SmtA. (*a*) Sequence alignment of SmtA homologues from various strains of *Synechococcus* (S.), *Anabaena* (A.), *Pseudomonas aeruginosa* (P.a.), and *P. putida* (P.p.). Green, conserved residue; yellow, semiconserved residue. The sequences for the Pseudomonads are truncated at the C terminus. (*b*) ''Sausage'' representation of a superposition of the backbone traces (N, Cα, C) of the 20 best NMR-derived structures of holo-SmtA without specific incorporation of Zn^{2+}. The cylindrical rod represents a trace through the Cα atoms of the protein. Its radius is proportional to the mean global backbone displacement per residue. The rmsd about the mean coordinates for residues 5–56 is 1.36 $\pm$ 0.35 Å for a best fit of the backbone atoms.

$^{111}Cd_4$-SmtA, prepared by reconstitution of apo-SmtA with $^{111}Cd^{2+}$ ions, gave a single peak on reverse-phase C_{18} HPLC with a Cd^{2+} content of 4.3 $\pm$ 0.4 Cd per mol protein (ICP-AES, Zn^{2+} not detectable). The most abundant peak in the positive ion ESI-MS of the ^{111}Cd-reconstituted protein (6051 amu) corresponded to Cd_4SmtA (calculated 6050.6 amu). [^{1}H, ^{1}H] 2D NMR spectra for $^{111}Cd_4$SmtA confirmed the homogeneity of the preparation and showed that the reconstitution had yielded a fully structured protein, with the spectra of the Zn^{2+}- and Cd^{2+}-loaded proteins showing very similar features.

The one-dimensional ^{111}Cd NMR spectrum of Cd_4SmtA (Fig. 2*a*) contains four peaks at 712 (A), 660 (B), 596 (C), and 567 ppm (D) with equal integrated peak areas, and chemical shifts indicative of S_4 or S_3N coordination (29). Analysis of 2D [^{1}H, ^{111}Cd] HSQC NMR spectra (see Fig. 6, which is published as supplemental data) and 2-bond ^{111}Cd–^{111}Cd couplings allowed the assignment of the ligands in the four Cd^{2+} binding sites as shown in Fig. 2*c*. There are two $CdCys_4$ and two $CdCys_3His$ sites linked by five bridging Cys thiolate sulfurs.

The ^{1}H NMR chemical shifts of Cys $\beta(CH_2)$ and His CH(δ) and CH(ε) protons of coordinated ligands in Cd_4SmtA are similar to those of the same residues in Zn_4SmtA suggesting that the cluster connectivities are the same in the two proteins. There are differences in shift for several backbone NH and CH(α) resonances for residues linked to the cluster structure. These differences probably arise from small changes in the global backbone conformation due to the larger size of Cd^{2+} (tetrahedral ionic radius 0.92 Å) compared with Zn^{2+} (0.74 Å). Small differences in the structures of human Zn^{2+}-MT-2 and Cd^{2+}-MT-2 (30) have also been attributed to the difference in size between the bound metal ions. Average M-S bond lengths in MTs are about 2.32 Å for Zn^{2+} and 2.53 Å for Cd^{2+} (26), and this difference gives rise to an expansion in the volume of the coordination sphere.

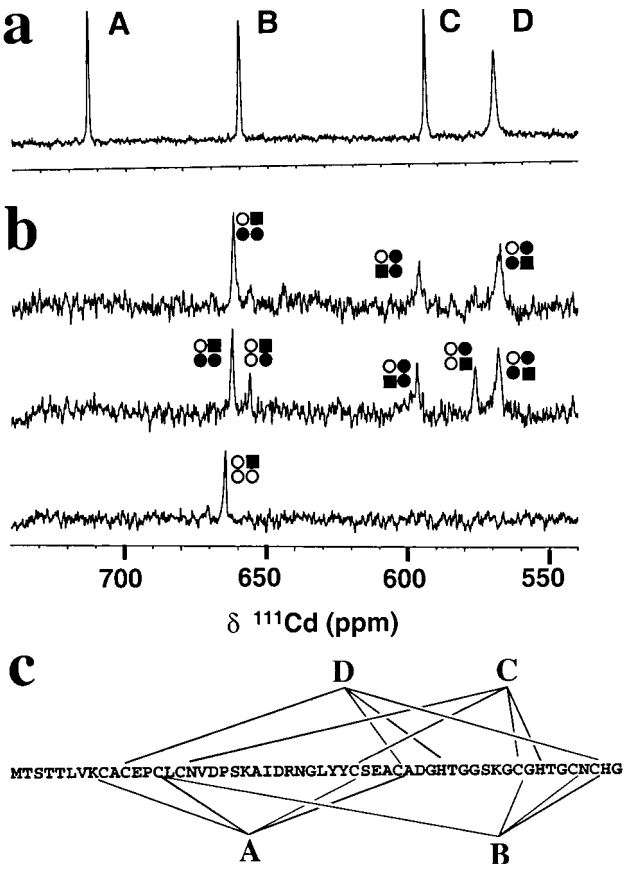

Fig. 2. Proton-decoupled one-dimensional ^{111}Cd NMR spectra of SmtA. (*a*) $^{111}Cd_4$SmtA prepared by reconstitution of apo-SmtA with $^{111}Cd^{2+}$ (106.04 MHz, pH 7.0, 318 K, 10% D_2O, 50 mM Tris·HCl, 50 mM NaCl) showing four peaks with equal integrals. (*b*) Substitution of Zn^{2+} by Cd^{2+}. Spectra recorded after mixing Zn_4SmtA with 1 (bottom), 4 (middle), and 8 (top) mol equivalents of $^{111}Cd^{2+}$ (pH 7.0, 308 K, 10% D_2O, 50 mM Tris·HCl, 50 mM NaCl). Comparison with *a* shows that the first mol equivalent of $^{111}Cd^{2+}$ selectively occupies binding site B, and that binding site A is not occupied at all. The various peaks reflect $^{111}Cd^{2+}$ in mixed Cd, Zn clusters. The occupation of the sites is depicted by squares and circles in the order ABDC, clockwise from top left. Filled square, Cd in the site which gives rise to the respective peak; filled circle, Cd in an adjacent site; open circle, Zn; e.g., the peak at 656 ppm corresponds to $^{111}Cd^{2+}$ in site B (■), with Zn^{2+} in sites A and C (○) and $^{111}Cd^{2+}$ in site D (●). (*c*) Metal-to-ligand connectivities for SmtA (as determined by 2D heteronuclear NMR experiments).

Refined Structure of SmtA. Inclusion of distance restraints for all four metal sites yielded the average structure shown in Fig. 3*a*. The best fit backbone rms deviation (rmsd) of the 20 best structures to the mean structure (0.90 ± 0.11 Å) was much improved by incorporating the metal ion restraints, without violating the previously used restraints. Comparison of the backbones of the averaged structures derived from the two different determinations yielded a rmsd value of 1.16 Å for all heavy atoms, showing that the ^{1}H NMR data alone are sufficient to define the overall conformation of the protein.

The structure of SmtA shows several intriguing features. Most noteworthy, and unusual for a metallothionein, is the presence of well defined elements of secondary structure. There is a short α-helix from Glu-34 to Gly-39, and two small antiparallel β-sheets formed by residues Val-7, Lys-8, Asn-17, and Val-18, and residues Ile-24 to Tyr-31. Residues Asn-27 and Gly-28 form a loop. Regular secondary structure elements, apart from short 3_{10}-helices and "half-turns," are rare in metallothioneins (25, 31), and it can be seen from the structure of the α-domain of rat MT-2 (32), shown in Fig. 3*b*, that the structure consists mainly of turns and loops. In Zn_4SmtA, a CH/π interaction (33) between H(α) of Ala-37, part of the α-helix, and the aromatic ring of Tyr-31, part of the β-sheet, provides direct communication between the two secondary structure elements (Fig. 4*a*), and accounts for the remarkable high-field shift of the H(α) proton of Ala-37 (1.73 ppm; induced by the ring current of Tyr-31). These two residues are conserved in related protein sequences (see Fig. 1*a*), suggesting that the interaction may be of functional significance. Notable too is the H-bond between the backbone amide proton of Cys-32 and the coordinated sulfur of Cys-9, both ligands for metal A (Fig. 3*c*). Backbone amide hydrogen bonds to coordinated side chain sulfurs have been observed previously, and the bond length (2.61 Å for the S $\cdots$ H distance) and S $\cdots$ HN angle (153°) are similar to values found in other published structures, such as high-potential iron-sulfur proteins (2.50 Å, 164°; ref. 34), and the Zn_2Cys_6 cluster of fungal transcription factor GAL4 (2.19–2.63 Å, 149–171°; ref. 35). The presence of the H-bond is corroborated by the extremely low-field ^{1}H chemical shift for the NH of Cys-32 (10.09 ppm). For Cd_4SmtA, this peak is shifted to high field by *ca.* 0.5 ppm (to 9.58 ppm), which is probably due to a lengthening of the H-bond brought about by the larger coordination sphere of Cd^{2+}.

The ligand content of SmtA (*ca.* 20%) is substantially lower than for other MTs (*ca.* 33% for mammalian MT) and SmtA has larger stretches of sequence devoid of metal binding residues. Although there is no correspondence between the position of the ligands in the sequences of the two proteins (Fig. 3 *a* and *b*), there is a remarkable overall structural similarity between the inorganic cores of the $M_4Cys_9His_2$ cluster in SmtA and the M_4Cys_{11} cluster in the α-domain of mammalian MTs (Fig. 4*a*). They are both four-metal clusters containing five bridging thiolates and six terminal ligands, with the metals and donor atoms arranged in two fused six-membered rings, which adopt distorted boat conformations. In SmtA, two terminal Cys residues are replaced by His ligands. This leads to a change in the overall charge on the cluster from -3 in MT-2 to -1 in SmtA. In mammalian MTs, second-shell Lys residues provide a charge balance for the negatively charged S(γ) atoms (36).

One Zinc Ion Is Inert to Cd Substitution. Next we studied reactions of Zn_4SmtA with Cd^{2+} ions. After addition of 10 mol equivalents of Cd^{2+} to Zn_4SmtA and removal of excess metal, 0.9 ± 0.1 mol equivalents of Zn^{2+} were still bound to the protein together with 2.4 ± 0.3 mol equivalents Cd^{2+}, even after long incubation times, as determined by ICP-AES. This result was surprising, because the metal ions in MTs studied so far are kinetically labile, and the bound metal ions undergo exchange reactions on a second-to-minute time scale (29). Moreover, "all-cysteine" MTs bind Cd^{2+} ions 10^4 more strongly than Zn^{2+} (37), and stoichiometric amounts of Cd^{2+} are usually sufficient to displace all of the Zn^{2+} from MT under similar conditions to those used here (29). Cys-to-His mutations in MT-2 have been shown to enhance Zn^{2+} binding (38), consistent with predictions based on the hard–soft acids and bases concept. A 100-fold preference for Cd^{2+}over Zn^{2+} for a $ZnCys_4$ zinc finger site has been reported, whereas a CCHC site has comparable affinity for Cd^{2+} and Zn^{2+} (39). ^{111}Cd NMR experiments showed that the Zn^{2+} ion that is inert to Cd^{2+} substitution is, curiously, in a Cys_4 site.

Fig. 2*b* shows ^{111}Cd NMR spectra of Zn_4-SmtA in the presence of increasing amounts of $^{111}Cd^{2+}$. Under these conditions, Zn–Cd exchange appeared to be rapid and equilibrium was established within 30 min. Comparison with Fig. 2*a* shows that $^{111}Cd^{2+}$ preferentially displaces Zn^{2+} from the Cys_4 binding site B, followed by CCHC sites D and C. In contrast, Cd^{2+} does not readily displace Zn^{2+} from the Cys_4 site A. The retention of one Zn^{2+} by the protein in the presence of excess Cd^{2+} is in agreement with the ICP-MS data. The inability of excess Cd^{2+} to displace Zn^{2+} from site A of SmtA appears to have no parallel in mammalian MTs, although site-specific displacement of Zn^{2+} by Cd^{2+} has been observed (40).

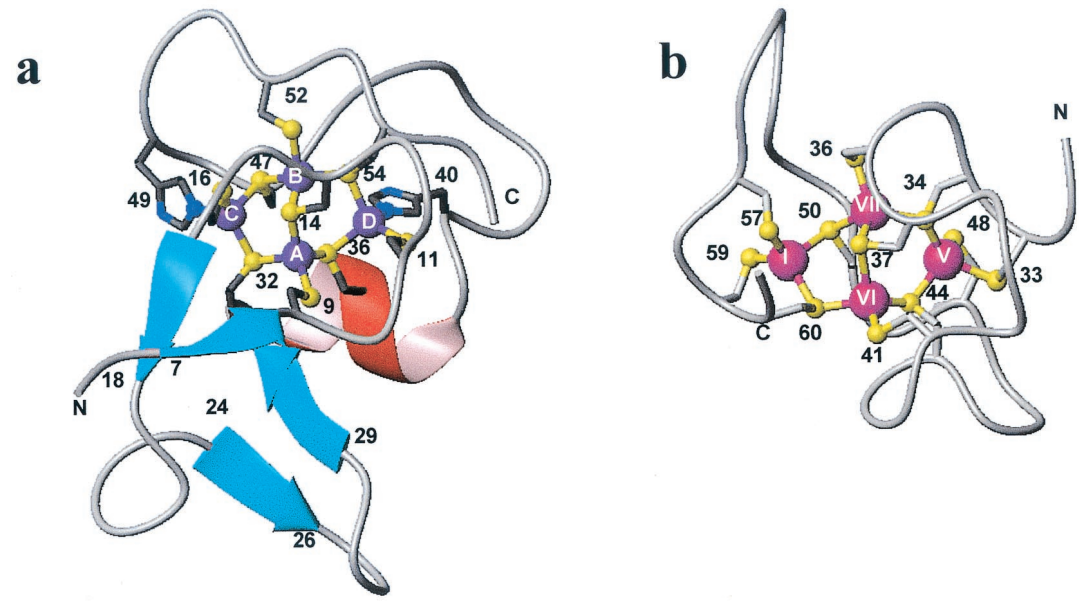

Fig. 3. Features of the NMR solution structure of bacterial Zn_4SmtA and comparison with mammalian MT-2. (*a*) Energy-minimized average structure of Zn_4SmtA including connectivities for the Zn^{2+} ions. The structure is derived from the 20 best structures. (*b*) Structure of the $[Cd_4]$ α-domain of rat MT-2 (32) showing the similarity between the inorganic cores of SmtA and the α-domain of mammalian MTs.

Cd^{2+} occupation of site A in SmtA was achievable only by reconstitution of the apoprotein.

Zinc Finger in the Cluster. While comparing the metal coordination sites of SmtA with those of reported zinc fingers, we discovered a remarkable similarity in the arrangement and orientation of site A and associated α-helix and antiparallel β-strands of Zn_4SmtA and the zinc site and secondary structure elements in the C-terminal zinc finger of the eukaryotic DNA binding protein GATA-1 (41), as well as in zinc finger 2 of LIM domains (e.g., CRP1 or PINCH LIM1; refs. 42 and 43), which are involved in protein–protein interactions. Fig. 4*b* shows an overlay of Zn(A)-SmtA (i.e., Zn^{2+} ions in sites B, C, and D omitted from the Figure) and the C-terminal zinc finger domain of GATA-1. Three of the four Cys ligands have very similar side chain conformations. The close overlay of SmtA with CRP1 LIM2 is shown in Fig. 7, which is published as supplemental data. Comparison of the sequences of SmtA, GATA-1 (residues 157–217, Fig. 4*c*), and CRP1 reveals that, despite the close structural resemblance, there is no sequence homology. The spacing patterns for the coordinated cysteines ($CX_4CX_{17}CX_3C$ for SmtA, and $CX_2CX_{17}CX_2C$ for GATA and CRP1 LIM2) differ, and only the stretch between the two pairs of Cys ligands in which the β-sheet lies is of identical length. The loop connecting the two β-strands in GATA is known to be essential for

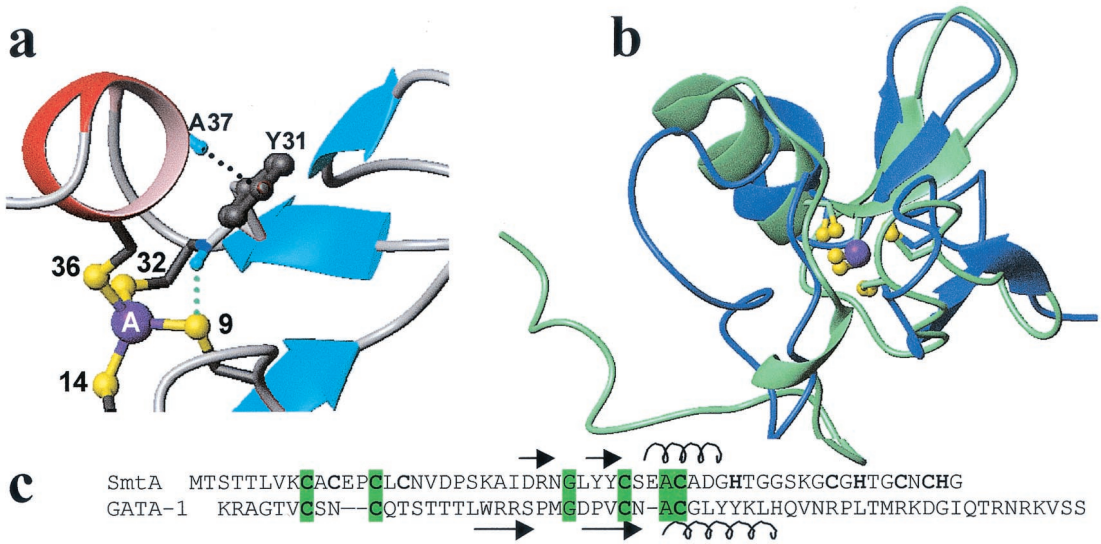

Fig. 4. (*a*) Weak interactions around binding site A: (*i*) CH/π interaction between the β-sheet and the α-helix. The distance between H(α) of Ala-37 and the aromatic ring of Tyr-31 is 2.47 Å. (*ii*) Hydrogen bond between NH of Cys-32 and the sulfur of Cys-9 (2.61 Å, 153°). (*b*) Overlay of site A in SmtA with the zinc site in a DNA-binding domain of the eukaryotic transcription factor GATA-1 (41). Overlay of all side chain atoms of the Cys residues results in a rmsd of 1.40 Å. (*c*) Sequence alignment of the molecules shown in *b*. Arrows and curls indicate elements of secondary structure. The only conserved residues are the binding site Cys, the Gly in the loop between the β-strands, and an Ala residue adjacent to Cys-36 in SmtA.

DNA binding (41). It is noteworthy that the Gly residue in this loop (Gly-28 in SmtA) is conserved in the sequences of the three proteins, as well as in those of the SmtA homologues. The potential biological importance of the zinc finger motif associated with site A is indicated by its apparent conservation in the sequences of gene products reported for 13 different bacteria (see Fig. 8, which is published as supplemental data; A.K.R., M.D.H., C.A.B., J.A.P., P. Bowness, P.J.S., and N.J.R., manuscript in preparation). Zinc fingers were thought to be confined to eukaryotes, but recent studies have indicated the presence of Krüppel-like zinc fingers in some prokaryotes (5, 44–46), although they remain to be confirmed by structure determination.

Zn^{2+} in site A not only appears to be inert toward Cd^{2+} substitution, but is also removed more slowly by chelating agents. The mass spectrum of Zn_4SmtA treated with a 125-fold excess of EDTA for 5 min showed a strong peak for Zn_1SmtA, and ^{1}H NMR spectroscopy gave evidence for a long-lived intermediate during Zn^{2+} removal. These unusual characteristics appear to arise from the presence of both the cluster and the secondary structure elements.

Accessibility of Cys Sulfurs from the Surface of the Protein. The mechanism of metal substitution in Zn_4SmtA may involve attack of Cd^{2+} on accessible Zn-bound terminal Cys sulfur atoms. Examination of the protein surface of Zn_4SmtA (see Fig. 7, which is published as supplemental data) shows that the coordinated terminal sulfur atoms of Cys-11, Cys-16, and Cys-52 in sites D, C, and B, respectively, are likely to be accessible, whereas the sulfur of Cys-9 in site A is completely buried. The negative electrostatic potential around Cys-11 may favor the entry of Cd^{2+} from this side of the protein (via site D). In contrast, the opposite side of the protein, which provides access to Cys-16 and Cys-52, is more hydrophobic. The differing environments of these potential entry and exit sites for metal ions could allow selective docking onto donor or acceptor partners.

Conclusions

The NMR structure of SmtA reveals a $Zn_4Cys_9His_2$ cluster. The four Zn^{2+} ions and five bridging Cys thiolate sulfurs form two fused six-membered rings with distorted boat conformations, a topology very similar to that of the Zn_4Cys_{11} cluster of the α-domain of mammalian metallothionein, despite the lack of sequence similarity. To date, Zn_4 clusters are unique to SmtA and to eukaryotic metallothioneins. However, while the two $ZnCys_3His$ sites and one of the $ZnCys_4$ sites in SmtA readily exchange Zn^{2+} for exogenous Cd^{2+}, the remaining $ZnCys_4$ site (A) is inert. The topology of this inert site and the surrounding β-sheet and α-helix closely resemble that of zinc finger portions of GATA and LIM proteins. The zinc sites in SmtA therefore appear to have evolved not only to allow zinc uptake and release for buffering roles, but also to endow specific protein and/or DNA recognition properties on SmtA. The next challenge is to identify recognition partner(s) for SmtA and to elucidate the nature of interaction processes, which may include metal transfer.

We thank Dr. Rhodri Thomas and Ms. Sally Shirran for assistance with ICP-AES and ESI-MS. This work was supported by fellowships to C.A.B. from the Swiss National Science Foundation, the Novartis Foundation (formerly Ciba–Geigy Jubilee Foundation), and the European Community (Marie Curie Fellowship under the Improving Human Potential program), and by the Wellcome Trust (Edinburgh Protein Interaction Centre), Engineering and Physical Sciences Research Council, and the Wolfson Foundation. N.J.R. and M.D.H. thank the Biotechnology and Biological Sciences Research Council for support.

1. Robinson, N. J., Gupta, A., Fordham-Skelton, A. P., Croy, R. R. D., Whitton, B. A. & Huckle, J. W. (1990) *Proc. R. Soc. London Ser. B* **242,** 241–247.
2. Olafson, R. W., McCubbin, W. D. & Kay, C. M. (1988) *Biochem. J.* **251,** 691–699.
3. Huckle, J. W., Morby, A. P., Turner, J. S. & Robinson, N. J. (1993) *Mol. Microbiol.* **7,** 177–187.
4. Turner, J. S., Morby, A. P., Whitton, A. B., Gupta, A. & Robinson, N. J. (1993) *J. Biol. Chem.* **268,** 4494–4498.
5. Bird, A. J., Turner-Cavet, J. S., Lakey, J. S. & Robinson, N. J. (1998) *J. Biol. Chem.* **273,** 21246–21252.
6. Vašák, M. & Hasler, D. W. (2000) *Curr. Opin. Chem. Biol.* **4,** 177–183.
7. Wüthrich, K. (1991) *Methods Enzymol.* **205,** 502–529.
8. Kägi, J. H. R. (1991) *Methods Enzymol.* **205,** 613–626.
9. Palmiter, R. D. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 8428–8430.
10. Suhy, D. A., Simon, K. D., Linzer, D. I. H. & O'Halloran, T. V. (1999) *J. Biol. Chem.* **274,** 9183–9192.
11. Andrews, G. K. (2000) *Biochem. Pharmacol.* **59,** 95–104.
12. Jacob, C., Maret, W. & Vallee, B. L. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 3489–3494.
13. Cano-Gauci, D. F. & Sarkar, B. (1996) *FEBS Lett.* **386,** 1–4.
14. Daniels, M. J., Turner-Cavet, J. S., Selkirk, R., Sun, H., Parkinson, J. A., Sadler, P. J. & Robinson, N. J. (1998) *J. Biol. Chem.* **273,** 22957–22961.
15. Higham, D. P., Sadler, P. J. & Scawen, M. D. (1984) *Science* **225,** 1043–1046.
16. Vašák, M. (1991) *Methods Enzymol.* **205,** 450–458.
17. Piantini, U., Sørensen, O. & Ernst, R. R. (1982) *J. Am. Chem. Soc.* **104,** 6800–6801.
18. Bax, A. & Davis, D. G. (1985) *J. Magn. Reson.* **65,** 355–360.
19. Kumar, A., Wüthrich, K. & Ernst, R. R. (1980) *Biochem. Biophys. Res. Commun.* **95,** 1–6.
20. States, D. J., Haberkorn, R. A. & Ruben, D. J. (1982) *J. Magn. Reson.* **48,** 286–292.
21. Hwang, T. L. & Shaka, A. J. (1995) *J. Magn. Reson. Ser. A* **112,** 275–279.
22. Zhang, O. W., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. (1994) *J. Biomol. NMR* **4,** 845–858.
23. Kuboniwa, H., Grzesiek, S., Delaglio, F. & Bax, A. (1994) *J. Biomol. NMR* **4,** 871–878.
24. Güntert, P., Braun, W. & Wüthrich, K. (1991) *J. Mol. Biol.* **217,** 517–530.
25. Braun, W., Wagner, G., Wörgötter, E., Vašák, M., Kägi, J. H. R. & Wüthrich, K. (1986) *J. Mol. Biol.* **187,** 125–129.
26. Hasnain, S. S., Diakun, G. P., Abrahams, I., Ross, I., Garner, C. D., Bremner, I. & Vašák, M. (1985) *Experientia Suppl.* **52,** 227–236.
27. Karlin, S. & Zhu, Z.-Y. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 14231–14236.
28. Koradi, R., Billeter, M. & Wüthrich, K. (1996) *J. Mol. Graphics* **14,** 51–55.
29. Öz, G., Pountney, D. L. & Armitage, I. M. (1998) *Biochem. Cell Biol.* **76,** 223–234.
30. Messerle, B. A., Schäffer, A., Vašák, M., Kägi, J. H. R. & Wüthrich, K. (1992) *J. Mol. Biol.* **225,** 433–443.
31. Wagner, G., Neuhaus, D., Wörgötter, E., Vašák, M., Kägi, J. H. R. & Wüthrich, K. (1986) *J. Mol. Biol.* **187,** 131–135.
32. Schultze, P., Wörgötter, E., Vašák, M., Kägi, J. H. R. & Wüthrich, K. (1988) *J. Mol. Biol.* **203,** 251–268.
33. Umezawa, Y., Tsuboyama, S., Takahashi, H., Usawa, J. & Nishio, M. (1999) *Bioorg. Med. Chem.* **7,** 2021–2026.
34. Rayment, I., Wesenberg, G., Meyer, T. E., Cusanovich, M. A. & Holden, H. M. (1992) *J. Mol. Biol.* **228,** 672–686.
35. Marmorstein, R., Carey, M., Ptashne, M. & Harrison, S. C. (1992) *Nature (London)* **356,** 408–414.
36. Vašák, M., McClelland, C. E., Hill, H. A. O. & Kägi, J. H. R. (1985) *Experientia* **41,** 30–34.
37. Vašák, M. & Kägi, J. H. R. (1983) *Metal Ions Biol. Syst.* **15,** 213–273.
38. Zhou, Y., Zhang, N., Li, L. & Ru, B. (2000) *Protein Peptide Lett.* **7,** 9–16.
39. Krizek, B. A., Merkle, D. L. & Berg, J. M. (1993) *Inorg. Chem.* **32,** 937–940.
40. Nettesheim, D. G., Engeseth, H. R. & Otvos, J. D. (1985) *Biochemistry* **24,** 6744–6751.
41. Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J. & Gronenborn, A. M. (1993) *Science* **261,** 438–446.
42. Yao, X., Perez-Alvaro, G. C., Louis, H. A., Pomies, P., Hatt, C., Summers, M. F. & Beckerle, M. C. (1999) *Biochemistry* **38,** 5701–5713.
43. Velyvis, A., Yang, Y., Wu, C. & Qin, J. (2001) *J. Biol. Chem.* **276,** 4932–4939.
44. Chou, A. Y., Archdeacon, J. & Kado, C. I. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 5293–5298.
45. Bouhouche, N., Syvanen, M. & Kado, C. I. (2000) *Trends Microbiol.* **8,** 77–81.
46. Wolfe, S. A., Nekludova, L. & Pablo, C. O. (1999) *Annu. Rev. Biophys. Biomol. Struct.* **3,** 183–212.