Zinc Binding by the Methylation Signaling Domain of the *Escherichia coli* Ada Protein†

Lawrence C. Myers,‡ Michael P. Terranova,§ Huw M. Nash,§ Michelle A. Markus,‡ and Gregory L. Verdine*,‡,§
Department of Chemistry and Program for Higher Degrees in Biophysics, Harvard University, Cambridge, Massachusetts 02138
Received January 28, 1992; Revised Manuscript Received March 26, 1992

ABSTRACT: The *Escherichia coli* Ada protein repairs O^6-methylguanine residues and methyl phosphotriesters in DNA by direct transfer of the methyl group to a cysteine residue located in its C- or N-terminal domain, respectively. Methyl transfer to the N-terminal domain causes it to acquire a sequence-specific DNA binding activity, which directs binding to the regulatory region of several methylation-resistance genes. In this paper we show that the N-terminal domain of Ada contains a high-affinity binding site for a single zinc atom, whereas the C-terminal domain is free of zinc. The metal-binding domain is apparently located within the first 92 amino acids of Ada, which contains four conserved cysteine residues. We propose that these four cysteines serve as the zinc ligand residues, coordinating the metal in a tetrahedral arrangement. One of the putative ligand residues, namely, Cys_{69}, also serves as the acceptor site for a phosphotriester-derived methyl group. This raises the possibility that methylation-dependent ligand reorganization about the metal plays a role in the conformational switching mechanism that converts Ada from a non-sequence-specific to a sequence-specific DNA-binding protein.

Nonenzymatic DNA methylation gives rise to the base adduct O^6-methylguanine (O^6-meG,[1] Figure 1), which causes mutations by pairing with thymine during replication (Hall & Saffhill, 1983; Loechler et al., 1984). This lesion poses such a serious threat that virtually all organisms possess mechanisms to eradicate it from their genome. The mechanism for repair of O^6-meG is both conserved and unusual: a repair protein, O^6-methylguanine methyltransferase (MGMtase), transfers the offending methyl group directly from DNA to one of its own cysteine residues (Teo et al., 1984; Nakabeppu et al., 1985). Since there appears to be no pathway for regeneration of the active site cysteine, the MGMtase protein functions as a stoichiometric reagent rather than a catalyst; its deliberate self-inactivation has been referred to as a suicidal reaction (Demple, 1988).

Escherichia coli elaborates two MGMtases, one of which (the *ogt* gene product) is expressed constitutively, while the other (the *ada* gene product) is induced upon challenge with methylating agents. In addition to *ada*, several other genes that encode DNA repair functions are induced by methylating agents; the most well understood of these, *alkA*, encodes a protein that catalyzes glycosidic bond cleavage of the methyl-bearing adducts N^7-methylguanine and N^3-methyladenine (Nakabeppu et al., 1984). These genes together form a small regulon that is under the positive transcriptional control of *ada* itself (Nakabeppu & Sekiguchi, 1986; Teo et al., 1986). Thus, the Ada protein acts not only as a specific repair protein for O^6-meG but also as a positive regulator for transcription of the *ada* regulon.

The molecular basis for the transcriptional regulatory function of Ada has been described in general terms. In addition to the active site for repair of O^6-meG (Cys_{321}), Ada possesses a second active site (Cys_{69}), which carries out an analogous suicidal methyl transfer from the S_p diastereomer of methyl phosphotriesters (Figure 1) (Sedgwick et al., 1988). Methylation of Cys_{69}, which can also occur via direct attack of methylating reagents such as methyl methanesulfonate (Takahashi et al., 1988), reveals in Ada a sequence-specific DNA binding activity that recognizes a promoter element—the "Ada box"—present in *ada* regulon genes (Teo et al., 1986; Sakumi & Sekiguchi, 1989). Thus, methylation of Cys_{69} in Ada is the triggering mechanism for a genetic switch that regulates resistance to methylating agents in *E. coli*; the peculiar feature of this triggering mechanism is that it is not directly reversible.

The structural details of how methylation at Cys_{69} brings about such a profound change in the biological activity of Ada, and how DNA-bound Ada activates transcription, remain poorly understood. Proteolysis studies have indicated that the 39-kDa Ada protein possesses two distinct domains. Each domain bears an independently functional active site, attached by a flexible connector. Cleavage takes place predominantly after Lys_{178} to furnish a 20-kDa N-terminal domain (N-Ada20), which demethylates (S_p)-methyl phosphotriesters, and a 19-kDa C-terminal domain (C-Ada19), which demethylates O^6-meG (Figure 1) (Demple et al., 1982, 1985; Yoskikai et al., 1988). The sequence-specific DNA binding activity of Ada appears to reside entirely within its N-terminal domain, since the methylated N-terminal fragment (me_{69}N-Ada20) binds the Ada box with an affinity similar to that of me_{69}Ada, whereas N-Ada20 binds weakly and the methylated C-terminal domain (me_{321}C-Ada19) not at all (Akimaru et al., 1990). The individual contributions of the N- and C-terminal

† This work was supported by grants (to G.L.V.) from the Chicago Community Trust (Searle Scholars Program), the National Science Foundation (Presidential Young Investigator Award), the Alfred P. Sloan Foundation, Hoffmann-La Roche, Eli Lilly, Bristol-Myers Squibb, and Sandoz. L.C.M. is a Howard Hughes Medical Institute Predoctoral Fellow. M.A.M. was supported by a National Science Foundation Predoctoral Fellowship.
‡ Program for Higher Degrees in Biophysics.
§ Department of Chemistry.

[1] Abbreviations: MGMtase, O^6-methylguanine methyltransferase; PMBS, *p*-(hydroxymercuri)benzenesulfonate; PAR, 4-(2-pyridylazo)-resorcinol; oligo,oligo-2′-deoxynucleotide; IPTG, isopropyl β-D-thiogalactopyranoside; MCS, multiple cloning site; PCR, polymerase chain reaction; PEI, poly(ethylenimine); O^6-meG, O^6-methylguanine; N-Ada20, Ada fragment containing residues 1–178; N-Ada10, Ada fragment containing residues 1–92; C-Ada19, Ada fragment containing residues 179–354; AAS, atomic absorption spectroscopy; M9-Glc, M9 minimal media supplemented with glucose.

0006-2960/92/0431-4541$03.00/0

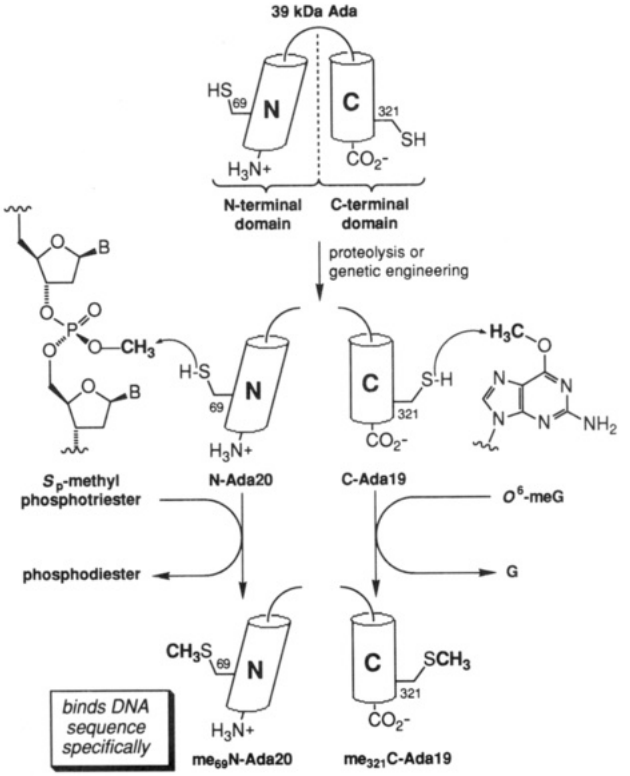

FIGURE 1: Repair of methylated DNA by Ada and its domains. Ada consists of two domains connected by a flexible linker. Proteolytic cleavage generates a 20-kDa N-terminal domain (N-Ada20) and a 19-kDa C-terminal domain (C-Ada19). Ada and N-Ada20 repair (S_p)-methyl phosphotriesters by direct transfer of the methyl group to Cys_{69}. Ada and C-Ada19 repair O^6-meG by direct transfer of the methyl group to Cys_{321}. Methylation on Cys_{69} converts Ada and N-Ada20 into a sequence-specific DNA-binding protein.

domains to transcriptional activation are less well understood, in part because the mechanism of transcriptional activation at the *ada* promoter is probably different from that at the *alkA* promoter. Evidence has suggested that the methylation-dependent switch in the Ada DNA-binding domain involves a conformational change rather than a local change in the protein–DNA contact repertoire, since $me_{69}me_{321}$Ada and Ada have markedly different proteolytic susceptibilities in the N-terminal region (Sedgwick et al., 1988).

To gain insight into the structural basis for methylation-dependent DNA binding by Ada, we have overproduced the full-length protein and its N-terminal domain at high levels in *E. coli*. During attempts to prepare isotopically labeled N-Ada20, we discovered that Zn^{2+} must be added to the growth medium in order for the protein to assume a correctly folded structure in vivo. Here we present conclusive evidence that the DNA-binding domain of Ada binds tightly 1 equiv of Zn^{2+}. Analysis of the *E. coli* Ada sequence in comparison with those of Ada proteins from other organisms suggests that the metal may coordinate the motif Cys-X_3-Cys-X_{26}-Cys-X_2-Cys, which is reminiscent of $(Cys)_4$ zinc-binding elements shown or proposed to be present in numerous DNA-binding proteins. Within this motif in Ada, one of the four putative ligand residues—Cys_{69}—is the active site for methylation, which raises the possibility that Zn^{2+} participates directly in the conformational switching mechanism.

Experimental Procedures

Materials

[*methyl*-^{3}H]MNU was purchased from Moravek Biochemicals Inc. (Brea, CA). PMBS and PAR were purchased from Sigma (St. Louis MO). S Sepharose Fast Flow and Sephacryl S-200 chromatographic media were from Pharmacia (Milwaukee, WI). IPTG was purchased from Bachem (Torrance, CA). The plasmid pBAR (McCarthy & Lindahl, 1985), on which is borne the full-length *E. coli* B *ada* gene, was a gift from Prof. Bruce Demple (Harvard School of Public Health). pKK223-3 (Brosius & Holy, 1984)) was purchased from Pharmacia, and pBS$^+$ (Short et al., 1988) was from Stratagene (La Jolla, CA). *Taq* polymerase was from Promega (Madison, WI) and T4 DNA ligase from BRL (Gaithersburg, MD). Restriction enzymes were from New England Biolabs (Beverly, MA). Oligos were synthesized on an Applied Biosystems 380A DNA synthesizer and purified by OPS cartridge (ABI, Foster City, CA). H_2O was twice distilled. Ultrafiltration was performed using Omegacells (Pharmacia).

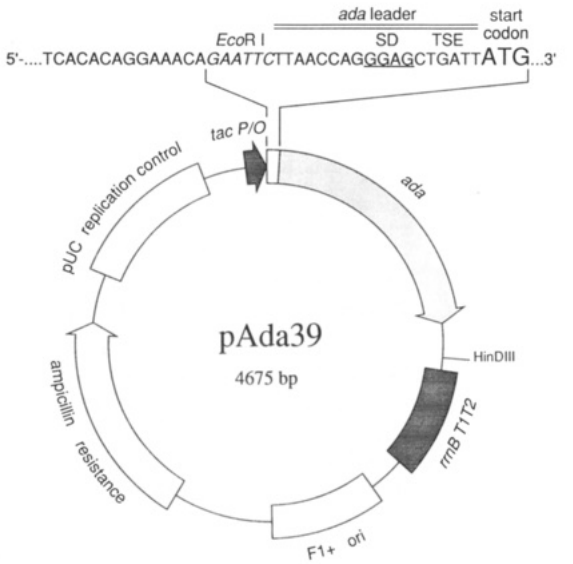

FIGURE 2: Ada-overproducing phagemid pAda39. In the expansion above the circular map is shown the sequence (coding strand) of the region from the 3′-end of the *lac* operator to the initiator methionine (start) codon. Within this region is the *ada* translational leader, which directs high-level expression of Ada. pAda39 was derived from pHN1$^+$ by replacement of the pUC9 MCS (between the *Eco*RI and *Hin*dIII) with the *ada* gene. pN-Ada20 and pN-Ada10 were derived from pAda39 by modifications within the *ada* gene.

Methods

Construction of the Phagemid Overexpression Vector pHN1$^+$. A 620-bp *Ssp*I fragment of pKK223-2 containing the *tac* promoter, pUC9 MCS, and *rrnBT*$_1$*T*$_2$ transcription terminator was ligated into the *Pvu*II-digested phagemid vector pBS$^+$. The orientation of the insert was determined by restriction mapping, and its sequence was verified by dideoxy sequencing. For constructs containing gene inserts in the MCS of pHN1$^+$, the following sequencing primers were used: *tac* primer (promoter-proximal), 5′-TATAATGTGTGGAATT-3′; TT primer (terminator-proximal), 5′-TCTCTCATCCGCCAAAACAG-3′. Competent XA 90(*F′lacI*^{Q1}) was prepared by the method of Hanahan (1985).

Construction of the 39-kDa Ada Overproducer. An *Eco*RI site was introduced immediately upstream of the *ada* Shine–Dalgarno (SD) sequence by site-directed mutagenesis in M13mp18. Cleavage at this site and the downstream *Hin*dIII furnished a 1259-bp *ada* gene cassette, which was inserted into the MCS of pHN1$^+$ to yield the Ada overexpression vector pAda39 (Figure 2). The *ada* overproducing construct and those described below were sequenced as reported (Sanger et al., 1977).

Construction of the N-Ada20 Overproducer. The 20-kDa

N-terminal domain of Ada (N-Ada20) was overproduced by a PCR-based approach, using the *tac* primer (see above) in combination with a 3′-primer that encoded a stop codon and a *Hin*dIII restriction site following Lys_{178}. The PCR product was ligated into pHN1$^+$ and transformed into XA 90 to yield the overproducing phagemid pN-Ada20.

Construction of the N-Ada10 Overproducer. The *ada* gene in pAda39 possesses a unique *Sal*I restriction site within codons 91–93. Into this restriction site was inserted a duplex oligonucleotide, 5′-TCGATAATTA-3′, to afford the overproducing phagemid pN-Ada10, in which a stop codon follows the codon for Arg_{92}.

Expression and Purification of Ada. *E. coli* XA 90-(pAda39) was grown at 37 °C with shaking in 1 L of LB–A medium (LB plus 50 μg/mL ampicillin) to an OD_{550} of 0.5–0.6. Protein overproduction was then induced by addition of IPTG to a final concentration of 0.25 mM. At the time of induction, $ZnCl_2$ was added to a final concentration of 0.1 mM. After 4 h, the culture was centrifuged at 5000*g* for 30 min, after which time the supernatant was discarded to yield ~3 g of wet cells. All subsequent steps were carried out at 4 °C. The cells were resuspended in 10 mL of lysis buffer [100 mM Tris-HCl (pH 8.0), 500 mM NaCl, 1 mM EDTA, 5% (w/v) glycerol, 2 mM $CaCl_2$, 10 mM $MgCl_2$, and 10 mM 2-mercaptoethanol] and lysed in a French pressure cell. The lysate was centrifuged at 22000*g* for 30 min, and the supernatant was then transferred to a fresh tube. A 5% (v/v, pH 8.0) solution of PEI·HCl was added to the supernatant to a final concentration of 0.5%. The PEI slurry was centrifuged at 22000*g* for 30 min, the supernatant was transferred to a fresh tube, and to it was added saturated aqueous $(NH_4)_2SO_4$ to 15% saturation. Following centrifugation at 22000*g* for 30 min, the pellet was discarded, and saturated aqueous $(NH_4)_2SO_4$ was added to 45% saturation. The $(NH_4)_2SO_4$ slurry was centrifuged at 22000*g* for 30 min, and the precipitate was resuspended in 10 mL of resuspension buffer [30 mM sodium phosphate (pH 7.8), 200 mM NaCl, 1 mM EDTA, 10 mM 2-mercaptoethanol, and 5% glycerol (v/v)]. The protein solution was then loaded onto a 2.6 × 33 cm S Sepharose column and eluted with an 800-mL linear gradient from 200 mM to 1 M NaCl in buffer A [20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 10 mM 2-mercaptoethanol, and 5% glycerol (v/v)]. The fractions were analyzed by SDS–PAGE; the Ada-containing fractions, eluting from 575 to 680 mM NaCl were pooled and then concentrated in a ultrafiltration cell (10-kDa cutoff). The concentrated protein sample was applied to a 2.5 × 108 cm Sephacryl S-200 column and eluted in buffer B [20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 10 mM 2-mercaptoethanol, and 100 mM NaCl]. SDS–PAGE analysis showed Ada to be present in fractions eluting at 260–290 mL; these were pooled, concentrated in an ultrafiltration cell, and assayed for methyltransferase activity. Active fractions were either used directly or flash frozen and stored at −70 °C. SDS–PAGE analysis of Ada prepared by this procedure showed it to be virtually homogeneous and to correspond in size to the expected molecular size (39.3 kDa).

Expression and Purification of N-Ada20. The expression and purification of N-Ada20 protein followed the same procedure as that used for Ada, with the exception of the following. Induced cells were lysed in 10 mL of buffer A plus 200 mM NaCl. The final concentration of PEI was 0.25%. Ammonium sulfate precipitation was carried out in a single step at 52% saturation, and the ammonium sulfate pellet was resuspended in buffer A plus 100 mM NaCl. N-Ada20 eluted from S Sepharose at 585–690 mM NaCl and in the Sephacryl S-200 column from 310 to 340 mL.

Expression and Purification of N-Ada10. The expression and purification of the N-Ada10 protein followed the same procedure as that used for Ada, with the exception of the following. Induced cells were lysed in 10 mL of buffer A. Ammonium sulfate precipitation was carried out in a single step at 70% saturation, and the ammonium sulfate pellet was resuspended in buffer A. N-Ada10 was eluted from the S Sepharose column with an 800-mL linear gradient from 0 mM to 1 M NaCl in buffer A. N-Ada10 eluted from S Sepharose at 535–665 mM NaCl and in the Sephacryl S-200 column from 340 to 360 mL.

Expression of Ada and N-Ada20 in Zinc-Supplemented Minimal Media. Growth on zinc-supplemented minimal media and purification were as described above for expression in rich media, with the exception of the following. Cells were grown at 30 °C in M9-Glc media (Miller, 1972). When the cells reached an OD_{550} of 0.5–0.6, $ZnCl_2$ was added to 0.1 mM and IPTG to 0.25 mM, and induction was allowed to proceed for 6 h.

Expression of Ada and N-Ada20 in Zinc-Free Minimal Media and Purification of the Denatured Proteins. Cells were grown at 37 °C in M9-Glc media (Miller, 1972) until an OD_{550} of 0.5–0.6, IPTG was then added to 0.25 mM, and incubation was continued 3 h (Ada) or 4 h (N-Ada20). The cells were lysed and the inclusion bodies purified by the Parker method as described in Bowden et al. (1991). The purified inclusion bodies were washed three times in 10 mM Tris-HCl (pH 8.0) containing 5.0 mM EDTA.

Quantitation of Zinc in Ada. Zinc content was assayed by both a colorimetric titration method and atomic absorption spectroscopy. Soluble protein samples were diluted in TNG buffer [10 mM Tris-HCl (pH 8.0), 200 mM NaCl, and 5% (v/v) glycerol] and quantified by amino acid analysis. Buffer controls were generated for all soluble protein samples by diluting a protein blank into TNG. Insoluble protein samples (Ada and N-Ada20 expressed in M9-Glc minus added zinc) were suspended in 10 mM Tris-HCl (pH 8.0) containing 5.0 mM EDTA and compared alongside supernatants obtained by centrifugation. Nanopure water and ultrapure reagents were used to minimize metal contamination in the buffers.

(*a*) *Colorimetric Method.* The colorimetric assay was carried out according to the procedure of Geidroc et al. (1986). Ada in TNG buffer plus 6.6 μM DTT was titrated, in the presence of 0.1 mM PAR, with an excess of PMBS to release all cysteine-bound metal. After each addition of PMBS, the solution was centrifuged to remove precipitated Ada protein, and the concentration of the zinc–PAR complex was quantified by measurement of OD_{500}. Purified N-Ada20, in TNG buffer, was assayed similarly except that the protein does not precipitate upon addition of PMBS; hence OD_{500} was measured directly. The Ada and N-Ada20 inclusion body proteins were not assayed.

(*b*) *Atomic Absorption Spectroscopy.* Zinc content was also determined by the method of electrothermal atomic absorption spectrometry as previously described (Medina et al., 1991). Protein quantitations were based on amino acid analysis, using the PicoTag method (Waters) and the known content of Ile + Leu + Phe for each protein.

Activity Assays. Methyltransferase activity of Ada and N-Ada20 was quantified by the method of Takahashi et al. (1988).

Results

To support structural studies, we have generated *E. coli* strains that overproduce the Ada protein from *E. coli* B, an

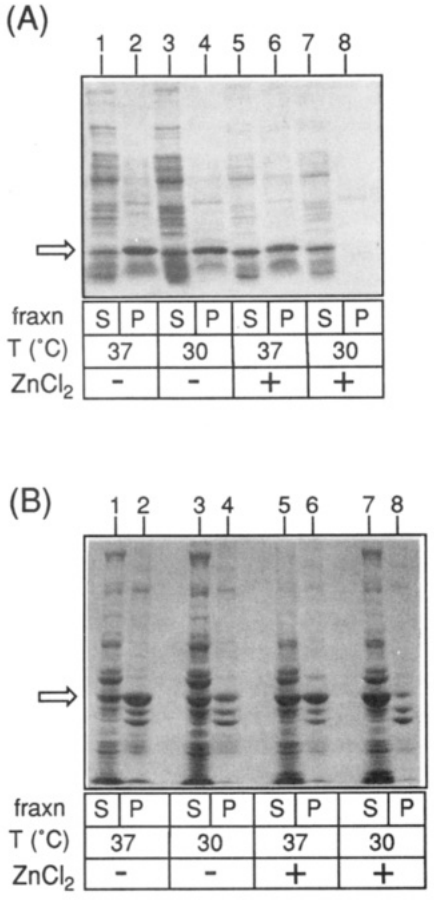

FIGURE 3: SDS–PAGE gel showing the solubility of N-Ada20 (A) and Ada (B) expressed in vivo as a function of temperature and added Zn^{2+}. Lanes marked S underneath represent the supernatant fractions (fraxn), and P the pellet fractions, from lysis of induced cells. An arrow marks the position of N-Ada20 (A) and Ada (B).

N-terminal fragment, N-Ada20, which contains its sequence-specific DNA-binding domain, and a smaller N-terminal fragment, N-Ada10. In these constructs, the native *ada* Shine–Dalgarno sequence and translational spacer element were used to direct translation initiation, whereas transcriptional control elements were provided by the strong *tac* promoter and *rrnBT*$_1$*T*$_2$ transcription terminator present in the vector pHN1$^+$ (Figure 2). The *ada*, *n-ada20*, and *n-ada10* genes thus equipped typically expressed their respective proteins at levels comprising greater than 40% of the total cell protein. Although the levels of expression were exceedingly high, we noted that sometimes as much as 50% of the recombinant protein was present in the insoluble fraction of induced cell lysates, depending on the conditions of growth. We later found that this problem of insolubility was greatly reduced by the addition of Zn^{2+} to the growth media (see below). Using these improved conditions, it was possible routinely to obtain ~20 mg of highly purified protein per liter of induced cell culture. The N-Ada20 overproducer yields 100-fold more protein than a previously reported construct that overproduces the corresponding gene product from *E. coli* K-12 (Akimaru et al., 1990). Given that the *ada* gene sequences from *E. coli* B and K-12 are almost identical (Demple et al., 1985; Nakabeppu et al., 1985), the greater expression in the pHN1$^+$ constructs probably results from their stronger promoter and abbreviated 5′-untranslated region.

As a prelude to preparation of isotopically labeled protein for multidimensional NMR experiments, we attempted to express N-Ada20 in minimal (M9-Glc) media, in which glucose and ammonium chloride, respectively, provide the sole sources of carbon and nitrogen. Although overproduction of N-Ada20 was induced to the same extent in M9-Glc as with rich media, with M9-Glc virtually all of the N-Ada20 was found in the pellet fraction after lysis of the cells (Figure 3A, lane 2), even when the pellet was washed with 1 M NaCl (data not shown). Whereas the supernatant fraction (Figure 3A, lane 1) contained detectable methyltransferase activity, the pellet did not, suggesting strongly that N-Ada20 was not folded in an active conformation. It was thus apparent that some factor present in rich media but absent in M9-Glc was required for the proper folding of N-Ada20 in vivo.

Table I: Zinc Content of Ada and Its Fragments[a]

sample	Zn(II) content (mol/mol of protein) colorimetric	AAS
intact Ada	1.35	0.91
N-Ada20 (M9-Glc)	1.56	1.33
N-Ada20 (LBA)		0.94
N-Ada10		1.59[b]
Ada inclusion bodies		0.11
N-Ada20 inclusion bodies		0.11

[a] All values are corrected for background metal content of the buffer and are the result of several duplicate runs. Notations in parentheses refer to the growth medium in which the proteins were overproduced. [b] The accuracy of the amino acid determination for this sample was possibly compromised by constituents present in the buffer.

Given that many DNA-binding proteins coordinate structural Zn^{2+} atoms and that an array of cysteines in N-Ada20 resembles that found in known or suspected Zn^{2+}-binding elements of many DNA-binding proteins, we decided to test whether addition of Zn^{2+} to the M9-Glc medium would have an effect on the folding of N-Ada20 in vivo. As shown in Figure 3A (lanes 5 and 6), with the addition of 100 μM Zn^{2+} to M9-Glc, roughly 50% of the N-Ada20 was soluble (active). Elsewhere it has been observed that reducing the growth temperature can have a beneficial effect on expression of soluble protein in vivo (Schein, 1989). We therefore tested the effect of reducing the growth temperature from 37 to 30 °C, in both the presence and absence of Zn^{2+}. In the absence of added Zn^{2+} (Figure 3A, lanes 3 and 4) lowering the temperature had a very modest effect, if any; however, in the presence of added Zn^{2+} (Figure 3A, lanes 7 and 8) lowering the temperature had a profound effect, with virtually all of the N-Ada20 being expressed in soluble form. Thus, the ability of Zn^{2+} to confer solubility on N-Ada20 is augmented by a reduction in the growth temperature. Similar effects of Zn^{2+} and temperature were observed for Ada (Figure 3B). The observed effect of added Zn^{2+} on N-Ada20 and Ada itself suggested either (i) that Zn^{2+} was required for some cellular machinery that accelerated the folding of Ada or (ii) that Ada directly ligates Zn^{2+} and therefore requires it for folding.

To determine whether Ada coordinates Zn^{2+}, and if so using which domains, we first purified soluble samples of Ada, N-Ada20, and N-Ada10 from cells grown in rich media (LB–A) and minimal media supplemented with Zn^{2+}. The buffer used for these purifications contained 1 mM EDTA, which was included to remove weakly bound or extrinsic metals. Aliquots were then titrated with PMBS—a sulfhydryl reagent that releases even strongly bound Zn^{2+}—in the presence of PAR, a Zn^{2+}-chelating chromogenic reagent (Hunt et al., 1985). The data for these titrations, shown in Table I, indicate that Ada, N-Ada20, and N-Ada10 each contain one atom of bound zinc. The lack of a significant difference between Ada and the N-terminal fragments allows one to conclude that the C-terminal domain of Ada does not coordinate Zn^{2+} tightly. A more rigorous measure of Zn^{2+} content

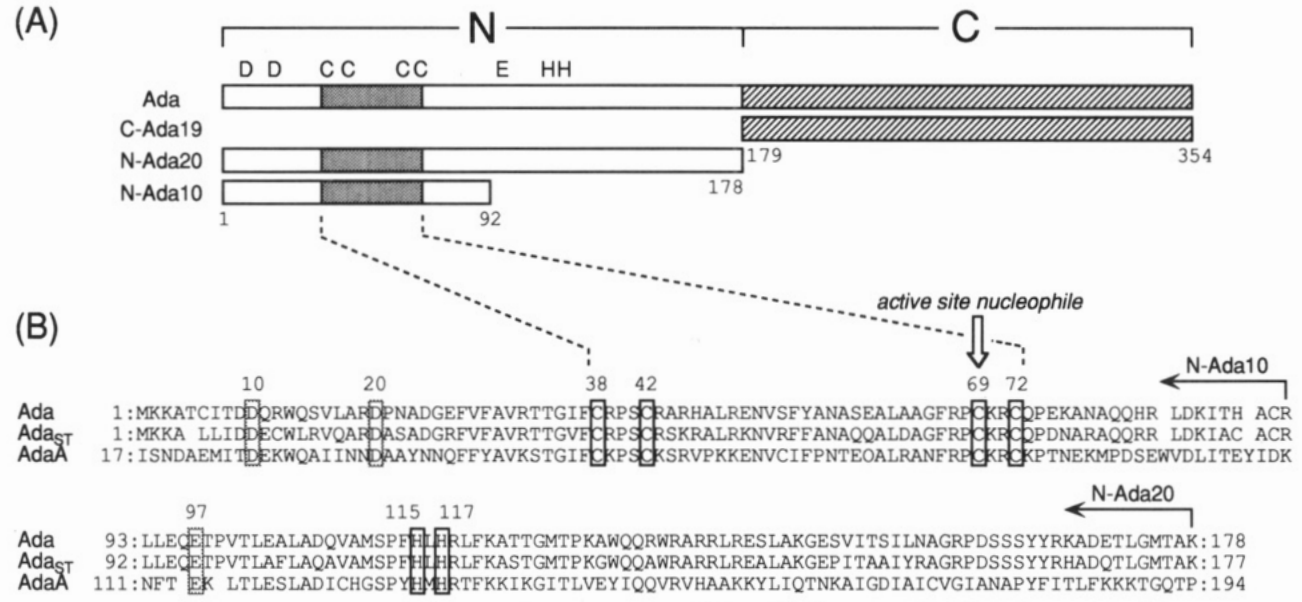

FIGURE 4: The zinc-binding domain of Ada. (A) Schematic representation of Ada and its fragments. Positions of potential Zn^{2+} ligand residues within the N-terminal domain are denoted above the diagram in one-letter code. The shaded region corresponds to the putative zinc-binding domain of Ada, which is of the Cys_4 type. (B) Sequence comparison of Ada protein from *E. coli* B (Demple et al., 1985), *S. typhimurium* (Hakura et al., 1991), and *B. subtilus* (Moroshi et al., 1990) within the N-terminal domain. Parts of the sequence corresponding to N-Ada10 and N-Ada20 are denoted by arrows. Carboxylate-type ligands are in lighter boxes, and Cys/His-type ligands are in darker boxes. The position of the active site nucleophile, Cys_{69}, is marked by an open arrow. The sequence of *E. coli* K12 *ada* (not shown) differs at three positions in the N-terminus: Leu_7, Thr_{151}, and Phe_{159} (Nakabeppu et al., 1985).

was gained from atomic absorption measurements. In these experiments, insoluble samples of Ada and N-Ada20 purified from M9-Glc (without added Zn^{2+}) were also included. These data (Table I) provide confirmatory evidence that the N-terminal domain contains a specific binding site for a Zn^{2+} atom. The insoluble proteins, on the other hand, lacked significant levels of Zn^{2+}. Incubation of the insoluble N-Ada20 with Zn^{2+} under nondenaturing conditions did not result in restoration of solubility; however, in rapid-dilution renaturation experiments on pellet samples of Ada and N-Ada20 taken up in 6 M guanidine hydrochloride, addition of Zn^{2+} increased markedly the yield of soluble protein.[2] These data suggest that the metal is required during folding for the Ada N-terminal to adopt a correct tertiary structure.

Discussion

The results of colorimetric and atomic absorption measurements demonstrate conclusively that *E. coli* Ada binds a single Zn^{2+} atom, thus documenting the first example, to our knowledge, of a prokaryotic transcription factor with a structural Zn^{2+}. Since Ada and its N-terminal frament, N-Ada20, each possess one Zn^{2+} atom, the metal binding site clearly resides in the N-terminal domain, which contains the protein's phosphotriester repair and methylation-dependent DNA binding activities. During high-level overproduction of Ada and N-Ada20 in *E. coli*, we observed that folding of the proteins to a soluble, active form in vivo is greatly facilitated by the addition of Zn^{2+} (or Cd^{2+}).[3] Ada and N-Ada20 also require Zn^{2+} (or Cd^{2+})[2] for efficient folding in vitro; moreover, removal of Zn^{2+} from Ada and N-Ada20 by treatment with PMBS results in rapid precipitation of the protein. Thus, the tightly bound Zn of Ada appears to be required not only for folding but also for maintenance of the folded state. Such Zn^{2+}-dependent folding effects, which have been previously observed both in vivo (Gao & Knipe, 1989) and in vitro (Frankel et al., 1987; Lee et al., 1991), may result from enthalpic stabilization of the folded protein by the presence of strong metal–ligand bonds (Berg, 1990).

In an attempt to identify putative ligand residues, we compared the sequence of N-Ada20 with those of functionally equivalent domains from two other proteins, the Ada protein from *Salmonella typhimurium* (Hakura et al., 1991) and the AdaA protein from *Bacillus subtilis* (Moroshi et al., 1990) (Figure 4B). Zn^{2+} exhibits very specific preferences for ligand moieties in proteins: the thiolate sulfur atom of Cys, the imidazole nitrogen atoms of His, and the carboxylate oxygens of Asp and Glu; of these, the carboxylate type appear to ligate only catalytic Zn^{2+}, whereas Cys and His ligate both catalytic and structural Zn^{2+} (Vallee & Auld, 1990a,b). Conserved among the N-terminal domains of Ada proteins are nine such residues, three of the carboxylate type and six of the Cys/His type (Figure 4). Of these, Glu_{97}, His_{115}, and His_{117} appear not to play a role, since N-Ada10 binds tightly one atom of Zn^{2+}, even though it lacks these residues (Figure 4A). Loss of even one ligand residue would be expected to labilize the Zn^{2+} and consequently the protein's tertiary structure, and yet N-Ada10 exhibits no such instabilities. Thus, we favor the hypothesis that the ligand residues are contained entirely within N-Ada10. It has previously been suggested, on the basis of sequence comparisons, that the N-terminal domain of Ada might possess a $(Cys)_2(His)_2$-type zinc-binding domain (Lindahl et al., 1988); however, our results do not support the presence of such a coordination scheme in Ada.

Of the six conserved ligand-type residues in N-Ada10 (Figure 4A), two are carboxylates (Asp_{10} and Asp_{20}) and the remaining four are thiols (Cys_{38}, Cys_{42}, Cys_{69}, and Cys_{72}). Since it is not clear whether the Zn^{2+} of Ada participates in the methyl-transfer reaction, it is not possible to rule out conclusively one or the other Asp residue as a ligand. It should be noted, however, that there is no known example of a Zn^{2+}-binding domain that uses only Glu/Asp and Cys, with no His (Vallee & Auld, 1990b). For this reason and because

[2] M. P. Terranova, unpublished results.
[3] L. C. Myers, unpublished results.

of the particular spacing of the conserved Cys residues in Ada (see below), we favor a model in which the four ligand residues are Cys_{38}, Cys_{42}, Cys_{69}, and Cys_{72}.

Zn^{2+}-binding domains containing only Cys residues have been shown or proposed to exist in numerous DNA-binding proteins (Berg, 1990; Kaptein, 1991). For two such proteins—GAL4 and the glucocorticoid hormone receptor—the presence and coordination environment of the metal have been determined rigorously. GAL4 binds two Zn^{2+} using a $(Cys)_6$ arrangement (Gardner et al., 1991; Povey et al., 1990). Ada binds only one Zn^{2+}, and although N-Ada10 possesses six Cys residues (6, 38, 42, 69, 72, 91), their spacing is unlike that in GAL4, and only four are conserved. The two tandemly arranged $(Cys)_4$ Zn-binding elements in the glucocorticoid hormone receptor (Härd et al., 1990) compare more favorably with the single array in Ada; however, Ada possesses only one $(Cys)_4$ element. Despite such vague similarities with these and other DNA-binding proteins in which the DNA-binding domain is organized about the metal, the differences are significant enough to suggest that the metal and DNA-binding domains of Ada are unique to this protein.

Sequence comparisons have suggested the presence of a helix–turn–helix DNA-binding element in Ada (Dodd & Egan, 1990) that is distinct from the metal-binding element we propose. However, our results suggest a means for communication of the methylation event to the DNA-binding element. A necessary consequence of the $(Cys)_4$ ligation scheme proposed above is the requirement that Cys_{69} function not only as a ligand for the metal but also as the acceptor of a methyl group acceptor from the phosphotriester substrate—this appears to be the first case in which (i) a $(Cys)_4$-ligated Zn^{2+} does not fulfill an exclusively structural role and (ii) a Zn^{2+}-bound Cys acts as nucleophile. Upon methylation, Cys_{69} is converted from the strongly metal-bonding thiolate form to the more weakly bonding thioether form (Chakrabarti, 1989; Kuehn & Isied, 1990). Hence methylation might be accompanied by ligand reorganization about the metal and conversion to a protein conformation that is capable of binding DNA sequence specifically. The metal center in Ada may thus function as a methylation-dependent conformational switch.

ACKNOWLEDGMENTS

We are deeply grateful to Dr. Bert Vallee and his co-workers Dr. Bob Shapiro and Ms. Lisa Kunze for carrying out the atomic absorption analyses and for related discussions. We thank Dr. Bruce Demple for providing pBAR and for helpful advice and Dr. Gerhard Wagner and Mr. François Jackow for stimulating discussions.

REFERENCES

Akimaru, H., Sakumi, K., Yoshikai, T., Anai, M., & Sekiguchi, M. (1990) *J. Mol. Biol. 216*, 261–273.

Berg, J. M. (1990) *J. Biol. Chem. 265*, 6513–6516.

Bowden, G. A., Paredes, A. M., & Georgiou, G. (1991) *Bio/Technology 9*, 725–730.

Brosius, J., & Holy, A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6929–6933.

Chakrabarti, P. (1989) *Biochemistry 28*, 6081–6085.

Demple, B. (1988) in *Protein Methylation* (Paik, W. K., & Kim, S., Eds.) CRC Press, Boca Raton, FL.

Demple, B., Jacobsson, A., Olsson, M., Robins, P., & Lindahl, T. (1982) *J. Biol. Chem. 257*, 13776–13780.

Demple, B., Sedgwick, B., Robins, P., Totty, N., Waterfield, M. D., & Lindahl, T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2688–2692.

Dodd, I. B., & Egan, J. B. (1990) *Nucleic Acids Res. 18*, 5019–5026.

Frankel, A. D., Berg, J. M., & Pabo, C. O. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4841–4845.

Gao, M., & Knipe, D. M. (1989) *J. Virol. 63*, 5258–5267.

Gardner, K. H., Pan, T., Narula, S., Rivera, E., & Coleman, J. E. (1991) *Biochemistry 30*, 11292–11302.

Geidroc, D. P., Keating, K. M., Williams, K. R., Konigsberg, W. H., & Coleman, J. E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8452–8456.

Hakura, A., Morimoto, K., Sofuni, T., & Nohmi, T. (1991) *J. Bacteriol. 173*, 3663–3672.

Hall, J. A., & Saffhill, R. (1983) *Nucleic Acids Res. 11*, 4185–4193.

Hanahan, D. (1985) in *DNA Cloning: A Practical Approach* (Glover, D. M., Ed.) Vol. 1, pp 109–135, IRL Press, Washington, DC.

Härd, T., Kellenbach, E., Boelens, R., Maler, B. A., Dahlman, K., Freedman, L. P., Carlstedt-Duke, J., Yamamoto, K. R., Gustafsson, J.-Å., & Kaptein, R. (1990) *Science 249*, 157–160.

Hunt, J. B., Neece, S. H., & Ginsburg, A. (1985) *Anal. Biochem. 146*, 150–157.

Kaptein, R. (1991) *Curr. Opin. Struct. Biol. 1*, 63–70.

Kuehn, C. G., & Isied, S. S. (1980) *Progr. Inorg. Chem. 27*, 154–221.

Lee, M. S., Gottesfeld, J. M., & Wright, P. E. (1991) *FEBS Lett. 279*, 289–294.

Lindahl, T., Sedgwick, B., Sekiguchi, M., & Nakabeppu, Y. (1988) *Annu. Rev. Biochem. 57*, 133–157.

Loechler, E. L., Green, C. L., & Essigmann, J. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6271–6275.

McCarthy, T. V., & Lindahl, T. (1985) *Nucleic Acids Res. 13*, 2683–2698.

Medina, J. F., Wetterholm, A., Radmark, O., Shapiro, R., Haeggström, J. Z., Vallee, B. L., & Samuelsson, B. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 7620–7624.

Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Moroshi, F., Hayashi, K., & Munakata, N. (1990) *Nucleic Acids Res. 18*, 5473–5480.

Nakabeppu, Y., & Sekiguchi, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6297–6301.

Nakabeppu, Y., Kondo, H., & Sekiguchi, M. (1984) *J. Biol. Chem. 259*, 13723–13729.

Nakabeppu, Y., Kondo, H., Kawabata, S.-I., Iwanga, S., & Sekiguchi, M. (1985) *J. Biol. Chem. 260*, 7281–7288.

Povey, J. F., Diakun, G. P., Garner, C. D., Wilson, S. P., & Laue, E. D. (1990) *FEBS Lett. 266*, 142–146.

Sakumi, K., & Sekiguchi, M. (1989) *J. Mol. Biol. 205*, 373–385.

Sanger, F., Nicklen, S., & Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Schein, C. H. (1989) *Bio/Technology 7*, 1141–1149.

Sedgwick, B., Robins, P., Totty, N., & Lindahl, T. (1988) *J. Biol. Chem. 263*, 4430–4433.

Short, J. M., Fernandez, J. M., Sorge, J. A., & Huse, W. D. (1988) *Nucleic Acids Res. 16*, 7583–7599.

Takahashi, K., Kawazoe, Y., Sakumi, K., Nakabeppu, Y., & Sekiguchi, M. (1988) *J. Biol. Chem. 263*, 13490–13492.

Teo, I., Sedgwick, B., Demple, B., Li, B., & Lindahl, T. (1984) *EMBO J. 3*, 2151–2157.

Teo, I., Sedgwick, B., Kilpatrick, M. W., McCarthy, T. V., & Lindahl, T. (1986) *Cell 45*, 315–324.
Vallee, B. L., & Auld, D. S. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 220–224.
Vallee, B. L., & Auld, D. S. (1990b) *Biochemistry 29*, 5647–5659.
Yoshikai, T., Nakabeppu, Y., & Sekiguchi, M. (1988) *J. Biol. Chem. 263*, 19174–19180.

Articles

Purification, Characterization, and Fibrinogen Cleavage Sites of Three Fibrinolytic Enzymes from the Venom of *Crotalus basiliscus basiliscus*†

Anastassios D. Retzios and Francis S. Markland, Jr.*

Department of Biochemistry, School of Medicine, University of Southern California, 1303 North Mission Road, CRL 106, Los Angeles, California 90033

Received July 23, 1991; Revised Manuscript Received February 19, 1992

ABSTRACT: Three distinct fibrinolytic enzymes have been purified from the venom of *Crotalus basiliscus basiliscus* (the Mexican west coast rattlesnake). The high-performance liquid chromatography-based purification comprised the following steps: (a) hydrophobic interaction chromatography; (b) hydroxylapatite chromatography; (c) anion-exchange chromatography. Following hydrophobic interaction chromtography two fibrinolytic activity peaks were detected, Cbfib1 and Cbfib2. Cbfib2 was rendered homogeneous following hydroxylapatite chromatography. Upon hydroxylapatite chromatography Cbfib1 was shown to consist of two components, Cbfib1.1 and Cbfib1.2. Both Cbfib1.1 and Cbfib1.2 were purified to homogeneity using anion-exchange chromatography. SDS–polyacrylamide gel electrophoresis revealed that Cbfib1.1 and Cbfib1.2 had similar molecular weights (approximately 23 500), whereas Cbfib2 displayed a molecular weight of approximately 22 500. The enzymes do not appear to be glycosylated. Tryptic digests of all three enzymes, analyzed by high-performance reverse-phase chromatography, suggest that Cbfib1.1 and Cbfib1.2 are closely related and different from Cbfib2. The latter displayed more similarity with Cbfib1.2 than with Cbfib1.1. Specific fibrinolytic activity for all three enzymes was very similar, but general proteolytic activity varied substantially with Cbfib2 showing a 12-fold higher specific proteolytic activity when compared to Cbfib1.1 and Cbfib1.2. None of these enzymes exhibited hemorrhagic activity when injected (up to 100 μg) subcutaneously into mice. Cbfib1.1 and Cbfib1.2 action against fibrinogen was directed equally against both the Aα- and Bβ-chains. Against fibrin the rate of degradation of the α-chain was considerably higher than that of the β-chain. Cbfib2 showed mainly α-fibrin(ogen)ase activity with limited activity on the β-chain. Several fibrinogen cleavage sites on the Aα-chain have been identified: Cbfib1.1 and Cbfib1.2 cleave at Lys^{413}–Leu^{414}, Ser^{505}–Thr^{506}, and Tyr^{560}–Ser^{561}. Cbfib2 cleaves mainly at Gly^{254}–Ser^{255} and Pro^{516}–Met^{517}.

A number of enzymes that influence blood coagulation have been isolated from various snake venoms. These enzymes can either promote or inhibit coagulation (Markland & Damus, 1971; Ouyang & Teng, 1972; Holeman & Weiss, 1976; Teng & Seegers, 1981). Fibrinolytic activities have received special attention because of their possible therapeutic role for dissolution of blood clots (Didisheim & Lewis, 1956) and because they may serve as templates for the development of de novo agents.

Fibrinolytic enzymes have been identified in the venom of several *Agkistrodon* species such as *contortrix*, *acutus*, and *piscivorus* [Kornalik (1966), Ouyang and Teng (1972), and Moran and Geren (1981), respectively]. A fibrinolytic enzyme from the venom of *Agkistrodon contortrix contortrix* (southern copperhead snake) has been purified and characterized by this laboratory (Markland et al., 1988; Guan et al., 1991). This enzyme (fibrolase or Accfib[1]) is a zinc metalloproteinase (1 mol of zinc/mol of enzyme) with a molecular weight close to 23 800. It readily cleaves the α- and Aα-chains of fibrin and fibrinogen, respectively, between Lys^{413}–Leu^{414}, without activating or degrading plasminogen or protein C (Retzios & Markland, 1988). A similar enzyme, Apcfib, has been purified from the venom of *Agkstrodon piscivorus conanti* (Florida cottonmouth) (Retzios & Markland, 1990). In vivo Accfib has been shown to dissolve aged clots in the renal arteries and iliac veins of rabbits (Markland et al., 1989). It may, therefore, have significant clinical potential. A nonhemorrhagic fibrinolytic enzyme, called atroxase, has also been

†This study has been supported in part by NIH Grant HL 31389 from the USPHS and by Cortech, Inc., Denver, CO.

[1] Abbreviations: Accfib, fibrinolytic enzyme from *Agkistrodon contortrix contortrix* (southern copperhead) venom, also known as fibrolase; Apcfib, fibrinolytic enzyme from *Agkistrodon piscivorus conanti* (Florida cottonmouth) venom; Cbfib1.1, Cbfib1.2, and Cbfib2, fibrinolytic enzymes from *Crotalus basiliscus basiliscus* (Mexican west coast rattlesnake) venom; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; DTT, dithiothreitol; TPCK, *N*-tosyl-L-phenylalanine chloromethyl ketone; PIU, plasmin international units of activity; SBTI, soybean trypsin inhibitor; PMSF, phenylmethanesulfonyl fluoride; HPLC, high-performance liquid chromatography; Zincov, 2-(*N*-hydroxycarbamoyl)-4-methylpentanoyl-L-alanylglycine amide.

0006-2960/92/0431-4547$03.00/0