# Identification of the Two Zinc-Bound Cysteines in the Ferric Uptake Regulation Protein from *Escherichia coli*: Chemical Modification and Mass Spectrometry Analysis†

Anne Gonzalez de Peredo,‡ Christine Saint-Pierre,‡ Annie Adrait,§ Lilian Jacquamet,§ Jean-Marc Latour,§ Isabelle Michaud-Soret,*,§ and Eric Forest*,‡

*Laboratoire de Spectrométrie de Masse des Protéines, Institut de Biologie Structurale, 38027 Grenoble Cedex 1, France, and Laboratoire de Chimie de Coordination (Unité de Recherche Associée au CNRS 1194), Service de Chimie Inorganique et Biologique, Département de Recherche Fondamentale sur la Matière Condensée, CEA-Grenoble, 38054 Grenoble Cedex 9, France*

*Received February 1, 1999; Revised Manuscript Received April 20, 1999*

ABSTRACT: Selective chemical modification of thiol groups combined with mass spectrometry analysis was used to characterize cysteine ligands in the zinc-binding site of the Fur protein. Fur is a metalloregulatory protein involved in the regulation of almost all bacterial genes related to iron uptake in Gram-negative bacteria such as *Escherichia coli*. In addition to the iron site, Fur also possesses a tight-binding zinc site that likely comprises two cysteines. Using a new procedure, we confirm the involvement of two cysteines in zinc binding and identify them within the two pairs of cysteines present in the protein. The protein was treated under nondenaturing conditions with iodoacetamide, and the progressive alkylation of the thiol groups monitored by quenching the reaction at different times and measuring the extent of alkylation by mass spectrometry. Complementary experiments were carried out in the absence or presence of EDTA, a strong zinc chelator, to determine which of the cysteines were protected from alkylation by the zinc atom. Enzymatic digestion of the modified protein and analysis of the peptide mixture by mass spectrometry enabled fast identification of reactive and protected thiol groups. Two cysteines, Cys92 and Cys95, were thus assigned as zinc ligands. Examination of the sequence comprising the zinc site indicates that it may belong to a new type of structural zinc site. Furthermore, Cys132 was shown to be the fastest reacting cysteine, implying it is a surface-exposed residue.

Iron plays a major role in almost all living organisms, as it is involved in many biological functions (*1*, *2*). The iron uptake system must however be very carefully regulated since ferrous iron is potentially toxic and ferric iron is highly insoluble ([soluble Fe(III)] < $10^{-17}$ M) (*2*). The control of the intracellular iron concentration in bacteria such as *Escherichia coli* occurs at the level of the iron uptake inside the cell (*3*, *4*), and Fur[1] is the key protein for this regulation. Fur is a metalloregulatory protein involved in the regulation of almost all bacterial genes related to the iron uptake in Gram-negative bacteria such as *E. coli* (*5*). It has been proposed that Fur utilizes iron in vivo as a corepressor and then acts as a negative regulator via sequence-specific protein−DNA interactions at the promoter regions of Fur-regulated genes (*6*). There is a 19 bp consensus palindromic DNA sequence called the "iron box" which is specifically recognized by the iron-bound Fur (*7*, *8*). In *fur* mutant strains, poor control of the iron concentration results in iron overload and the appearance of an oxidative stress and DNA damage (*9*).

Fur consists of two identical subunits of 148 amino acids. Results of chromatographic experiments (*10*) as well as mass spectrometry data (*11*) suggest that the protein exists as a noncovalent dimer in solution, even in the absence of metal. The metal binding site and the dimerization site are proposed to be in the C-terminal region of the protein and the DNA binding site in the N-terminal domain (*12*, *13*). Bagg et al. (*6*) found that all first-row divalent metal ions can, with varying degrees of efficiency, cause Fur to bind the operator. In vivo, Mn, Fe, and Co, but not Zn, are able to activate the protein, and in vitro Fe, Mn, Co, Cd, Cu, and Zn can activate the protein (*6*, *8*). When the metal is bound, a conformational change is proposed to allow the binding of the N-terminal part of the protein to the DNA. Although several spectroscopic studies of Fur using NMR (*14*−*16*), EPR (*17*), and UV−visible spectroscopy (*18*) have previously been reported, the environment of the iron binding site responsible for the activation of Fur is still not known precisely. Recent spectroscopic data indicate that the metal is penta- or hexacoordinated with N or O donor ligands, but not with sulfur ligands (*19*). In addition to this metallic site, Fur also

† This work was supported by Grant E 90010 00 00 from the "Région Rhône-Alpes".

* To whom correspondence should be addressed. Fax: (33) 4 76 88 54 94. E-mail: forest@ibs.fr and michaud@drfmc.ceng.cea.fr.

‡ Institut de Biologie Structurale.

§ CEA-Grenoble.

[1] Abbreviations: Fur, ferric uptake regulation; EXAFS, extended X-ray absorption fine structure; ICP-AES, inductive coupling plasma-atomic emission spectroscopy; EDTA, ethylenediaminotetraacetic acid; PMSF, phenylmethanesulfonyl fluoride; MOPS, morpholinopropanesulfonic acid; Tris, tris(hydroxymethyl)aminomethane; MALDI, matrix-assisted laser desorption/ionization; HPLC, high-pressure liquid chromatography.

10.1021/bi9902283 CCC: $18.00 
Published on Web 06/11/1999

| *organisms* | Ref | C-terminal sequences |
|---|---|---|
| Number (from *E. coli*) | | 80 92 95 124 132 137 |
| *Escherichia coli* | (*36*) | ELTQQH**HHDH**LICLDCGK(..)RLTN**H**SLYLYGHCAEGD**C**REDEHAHEGK - |
| *Neisseria gonorrhoeae* | (*37*) | ELDKGD**HHDH**IVCVK**C**GE(..)RIVD**H**ALYMYGVCSDCQAKGKR |
| *Yersinia pestis* | (*38*) | ELTQQH**HHDH**LICLDCGK(..)KLTN**H**SLYLYGHCETGN**C**REDESAHSKR |
| *Vibrio cholerae* | (*39*) | ELSTQH**HHDH**LVCLDCGE(..)QLTN**H**SLYLYGK**C**GSDGS**C**KDNPNAHKPKK |
| *Campylobacter jejuni* | (*40*) | ELANKP**HHDH**MICKN**C**GK(..)KLTG**H**LMQLYGVCGD**C**NNQKAKVKI |
| *Erwinia chrysanthemi* | (*41*) | ELTQQH**HHDH**LICLDCGR(..)KLTN**H**SLYLYGHCSSGD**C**REDDNAHNER |
| *Pseudomonas aeruginosa* | (*23*) | ELADSG**HHDH**MV**C**VDTGE(..)ELVD**H**NLVLYVRKKK |

FIGURE 1: Sequence alignments of parts of C-terminal sequences of Fur from several organisms containing the conserved cysteines.

possesses a tight-binding zinc site. The presence of zinc in the protein after overexpression and purification was shown by ICP analysis, which gave a stoichiometry of 0.5−0.8 Zn atom per monomer (*11*). An X-ray absorption spectroscopy (XAS) study unambiguously demonstrated the existence of a specific zinc site (*20*). This zinc site is tetrahedral and comprises two sulfur donor ligands, probably two cysteines, at 2.3 Å, and two N or O donor ligands at 2.1 Å, one of them at least being a histidine (*20*). Many histidines, glutamates, aspartates, tyrosines, and cysteines, usually involved in metal sites, are well-conserved in the sequences of the Fur proteins reported in the literature (*13*) and found in the GenBank database. It is proposed that they are potential ligands of the metal ions. The four cysteines present in *E. coli* Fur are well-conserved in almost all the Fur sequences except in the Fur from some bacteria such as *Rhizobium leguminosarum* (*21*), *Pseudomonas putida* (*22*), and *Pseudomonas aeruginosa* (*23*). Most of the histidines and the four cysteines are contained in the C-terminal region of the protein. Figure 1 shows sequence alignments of part of the C-terminal sequences of Fur from several organisms. The motif HHDHXXCXXC is highly conserved, as is the position of three of the four cysteines. The fourth cysteine is located two to five amino acids from the third cysteine. The four cysteines can be grouped into two pairs (C92-$X_2$-C95 and C132-$X_4$-C137 in *E. coli* Fur) with a spacing reminiscent of that observed in the sequences of zinc finger-containing proteins (*24*). Mutation studies with *E. coli* Fur involving the mutation of the 12 histidines present in the sequence into leucines gave a partial loss of activity. Furthermore, mutations of the four cysteines to serines showed that cysteines 92 and 95 were essential to the activity (*25*). The cysteine mutations resulted in inactivation of the protein, due either to destabilization of the protein structure or to impaired metal binding. The inactivation could not be due to the simple loss of a disulfide bond since it has been shown that Fur does not contain any (*26*). Furthermore, chemical modification reagents specific for thiol groups have been used to demonstrate that free cysteines are essential to the activity (*26*).

Using a new procedure, we here confirm the involvement of two cysteines in zinc binding and identify them within the two pairs of cysteines present in the protein. The protein was treated under nondenaturing conditions with iodoacetamide, and the progressive alkylation of the thiol groups was monitored by quenching the reaction at different times and measuring the extent of alkylation by electrospray ionization mass spectrometry. Complementary experiments were carried out in the absence or presence of EDTA, a strong zinc chelator, to verify that some of the cysteines were protected from alkylation by the zinc atom. Enzymatic digestion of the modified protein and analysis of the peptide mixture by MALDI mass spectrometry enabled fast identification of reactive and protected thiol groups. The two cysteines 92 and 95, essential for the activity, were thus assigned as zinc ligands. The sequence of the zinc site in Fur is different from the consensus sequences of other known classes of stuctural zinc sites. This indicates that the two cysteines may belong to a new type of structural zinc site.

## EXPERIMENTAL PROCEDURES

*Materials*. Trizma buffer, ethylenediaminetetraacetic acid (EDTA), and iodoacetamide were purchased from Sigma. Endoproteinase Lys-C and endoproteinase Asp-N were purchased from Boehringer Mannheim. Poly(propylene glycol) for electrospray mass spectrometer calibration was purchased from Aldrich.

*Overproduction and Purification of Fur*. The T7 RNA polymerase/promoter system was used to overproduce Fur from *E. coli* as previously described (*27*). Fur was purified as previously described (*28*) but with some modifications (*11*, *20*). The samples collected from the gel filtration column were concentrated in 0.1 M Tris-HCl (pH 8) containing 0.1 M KCl and 10% v/v glycerol.

*Protein Concentration Measurements.* Proteins concentrations were determined spectrophotometrically using an

absorption coefficient at 275 nm of 0.4 $mg^{-1}$ mL $cm^{-1}$ for one monomer of pure apo-Fur (*29*).

*Activity Assay.* The capacity of metal-substituted Fur to bind DNA was checked using a gel retardation assay. These experiments were performed by incubating a 25 bp $^{32}$P-end-labeled oligomer containing the iron box consensus sequence (sequence of one strand (5′-GGGGATAATGATAATCATTATCGGG-3′) at a nanomolar concentration with Fur samples (micromolar concentration) in the presence of 100 $\mu$M $MnCl_2$ (unpublished results).

*Alkylation of the Fur Protein.* Alkylation of the thiol groups was performed in 0.1 M Tris-HCl/0.1 M KCl (pH 8) or in 0.1 M Tris-HCl/0.1 M KCl/50 mM EDTA (pH 8) at room temperature (25 °C). The Fur protein was taken from a stock solution at 437 $\mu$M (7.33 $\mu$g/$\mu$L) in 0.1 M Tris-HCl (pH 8) containing 0.1 M KCl and 10% v/v glycerol, and diluted in the appropriate buffer (with or without EDTA) to a final concentration of 50 $\mu$M (0.84 $\mu$g/$\mu$L). A solution of idoacetamide was freshly prepared in the same buffer, and added to the protein to a final concentration of 6 mM. The reaction was carried out in the dark. These conditions of alkylation were adapted from ref *29*. Several concentrations of iodoacetamide had been previously tested (6, 2, and 0.5 mM, corresponding to 120, 40, and 10 equiv of alkylating reagent against the protein). Using lower concentrations of iodoacetamide resulted in the same pattern of alkylating kinetics, but with lower rates. Aliquots of 20 $\mu$L (1 nmol of Fur) were taken at various times, and immediately frozen in liquid nitrogen to stop the reaction. The samples were then purified by reverse phase HPLC on a C4 Brownlee column (2.1 mm × 150 mm, Applied Biosystems), using a model 130A syringe pump system (Applied Biosystems). Solvent A was a solution of 0.1% trifluoroacetic acid in water, and solvent B was a 90/10/0.08 acetonitrile/water/trifluoroacetic acid mixture. The column was equilibrated in 10% solvent B, and the protein was then eluted with a fast gradient of solvent B (10 to 80% over the course of 20 min). Fractions containing the purified protein were freezed-dried, and the product was then dissolved in 20 $\mu$L of water.

*Electrospray Mass Spectrometry Analysis of Alkylated Fur.* Electrospray mass spectrometry was performed using a Sciex API III+ triple-quadrupole mass spectrometer (Perkin-Elmer Sciex) equipped with a nebulizer-assisted electrospray source. Calibration was performed in positive mode using poly(propylene glycol) ions. Mass spectra were analyzed using a Quadra 950 data system (Apple). The ion spray voltage was set at 5 kV and the orifice voltage at 80 V. The interface temperature was set at 55 °C. Sample volumes of 1 $\mu$L were injected into the loop of a Valco C6W valve and infused into the spectrometer using a syringe pump (Harvard 22), at a flow rate of 5 $\mu$L/min, with 74/25/1 water/methanol/acetic acid as the vector solvent. Spectra were acquired in multichannel acquisition mode, with a scan step of 0.1 *m*/*z* and a dwell time of 2 ms, from *m*/*z* 920 to 1100, a range including three charge states of the denatured protein.

*Enzymatic Digestion of Alkylated Fur with Endoproteinase Lys-C.* After HPLC purification, alkylated samples of Fur were diluted in 0.1 M ammonium bicarbonate buffer (pH 7.9) to a concentration of ~25 $\mu$M protein. About 1% w/w endoproteinase Lys-C was added, and the mixture was incubated at 37 °C for 6 h.

*Separation of the Peptides and Enzymatic Digestion with Endoproteinase Asp-N.* To identify the fast-reacting cysteine, the peptides obtained by cleavage of the sample alkylated for the shortest time period (15 min) with endoproteinase Lys-C were separated by reverse phase HPLC on a C18 Brownlee column (1 mm × 150 mm, Applied Biosystems), using a model 130A syringe pump system (Applied Biosystems). Solvent A and solvent B were the same as those previously described. The gradient used was as follows: (1) 2 to 32% solvent B over the course of 90 min and (2) 32 to 82% solvent B over the course of 30 min. Fractions containing a mixture of peptide H117−K147 free and singly alkylated were dried with a Speed-vac concentrator, and the peptides were then redissolved in 20 mM ammonium bicarbonate buffer (pH 7.9). About 0.05 $\mu$g of endoproteinase Asp-N was added, and the mixture was incubated at 37 °C for 4 h.

*MALDI Mass Spectrometry Analysis.* MALDI mass spectrometry was performed on a Perseptive Voyager XL (Perseptive Biosystems) time-of-flight mass spectrometer. A solution of 2,5-dihydrobenzoic acid in 50/50 water/acetonitrile was used as a matrix. A volume of 1 $\mu$L of crude peptide mixture was mixed on the target with an equal volume of matrix, and the spot was allowed to air-dry. Ions were accelerated with an extraction voltage of 20 kV. Spectra were obtained by accumulation of 256 laser shots (Nd:YAG laser, $\lambda$ = 337 nm).

*Specificity of Iodoacetamide*. To check the specificity of iodoacetamide toward the cysteines, the fully alkylated Fur was cleaved with endoproteinase Lys-C as described above. After HPLC separation, the peptide Ser77−Lys97 was cleaved with endoproteinase Asp-N as described above. The peptide Asp88−Leu93 was submitted to tandem mass spectrometry using the Sciex API III+ instrument: parent ion, *m*/*z* 770; collision gas, argon; collision energy, 60 eV; and target gas thickness, $1.2 \times 10^{15}$ atoms/$cm^2$. The peptide His117−Lys147 was cleaved with chymotrypsin; it was redissolved in 100 mM ammonium acetate buffer (pH 7.6), about 0.6 $\mu$g of chymotrypsin added, and the mixture incubated at 37 °C for 6 h. After HPLC separation, the peptide Gly130−Lys147 was entirely sequenced using Edman degradation with an Applied Biosystems model 477A protein sequencer; amino acid phenylthiohydantoin derivatives were identified and quantitated on-line with a model 120A HPLC system (Applied Biosystems) as recommended by the manufacturer.

## RESULTS

It has previously been proposed that two cysteines of the Fur protein from *E. coli* are involved in zinc binding (*20*). Our objectives in this study were to confirm that only two cysteines are bound to the zinc ion and to identify them within the four cysteines of the Fur monomer. The time course of alkylation of the Fur protein was monitored in the presence and absence of EDTA. Iodoacetamide was used to probe the accessibility of cysteines in Fur, through modification of thiol groups. The specificity of the reaction toward the cysteines was checked, as discussed below. The degree of modification of the protein was assessed by electrospray ionization mass spectrometry. As the ionizable side chains (from lysine, arginine, or histidine residues) are generally

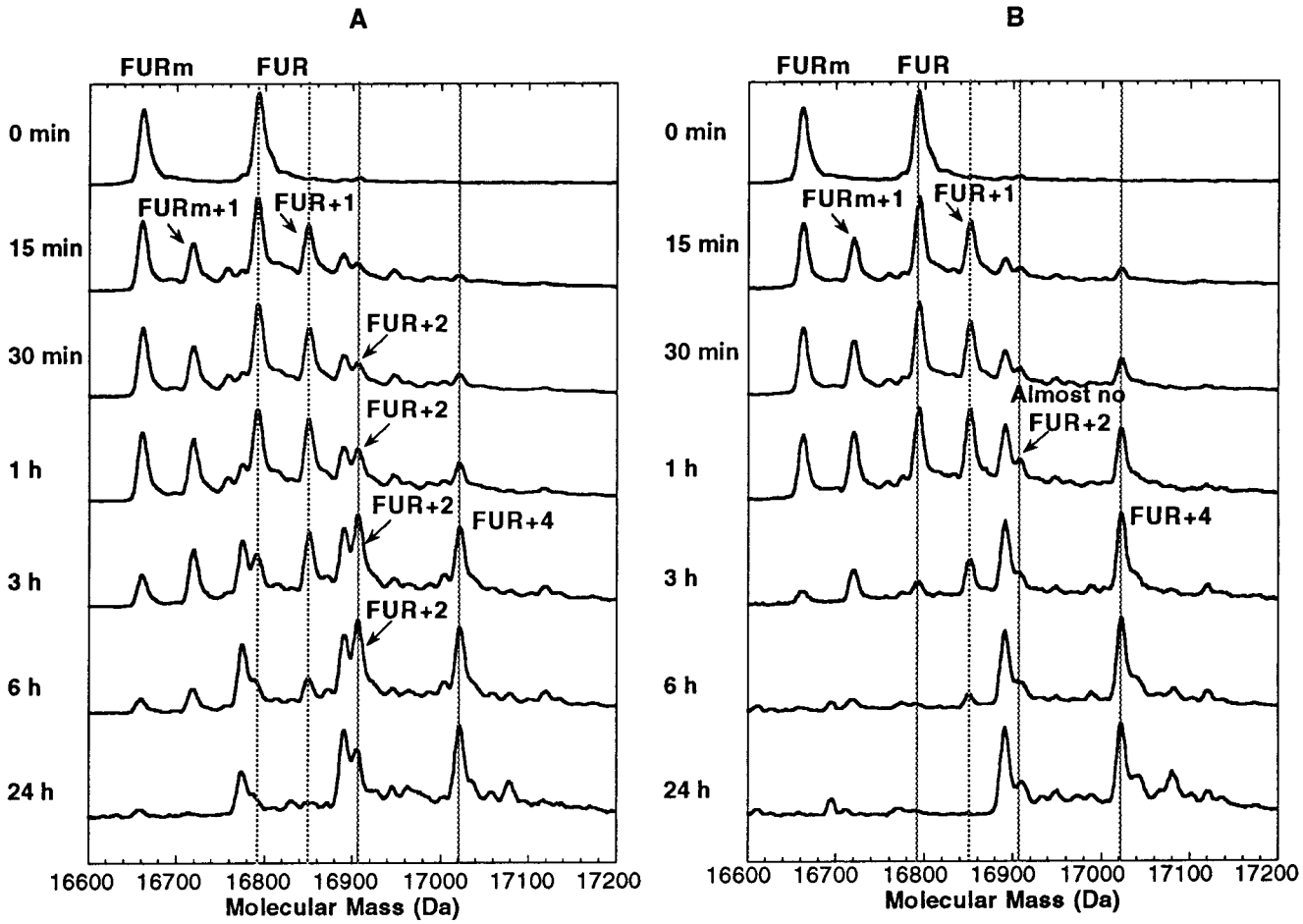



FIGURE 2: Kinetics of alkylation of FUR by iodoacetamide monitored by electrospray ionization mass spectrometry (A) in 0.1 M Tris-HCl/0.1 M KCl at pH 8 and 25 °C and (B) in 0.1 M Tris-HCl/0.1 M KCl/50 mM EDTA at pH 8 and 25 °C. The mass spectrum of the unmodified protein ($t = 0$ min) shows the presence of two species corresponding to the entire protein (16 794 Da) and to the protein in which the N-terminal methionine has been excised (16 663 Da). Addition of one alkylating group to a cysteine results in a mass increase of 57 Da. Peaks labeled FUR+1, FUR+2, and FUR+4 correspond to the entire protein having one, two, and four fixed alkylating groups. The others peaks in the spectra correspond to the same adducts arising from the protein without N-terminal methionine.

not involved in these modifications, we considered the addition of alkylating groups to have only a minor influence on the ionization efficiency of the Fur protein. On the basis of this consideration, the relative intensities of the peaks in the mass spectra, representing the relative proportion of gas-phase ions, were also indicative of the relative proportion of the different species in solution.

*Alkylation of Fur in the Absence of EDTA.* Figure 2A shows reconstructed mass spectra of Fur obtained at various times during alkylation. The spectrum obtained before the beginning of the reaction ($t = 0$ min) shows the presence of two species with a mass difference of 131 Da. The highest mass at 16794 Da (termed FUR) precisely corresponds to the mass of the peptidic chain calculated from the amino acid sequence. The absence of the dimeric form or of zinc atom can be explained by the denaturing conditions used to obtain the mass spectra with a good sensitivity. Only very soft conditions using nondenaturing solvent allow the observation of the dimeric form of the protein as previously described (*11*). The second species (16 663 Da) corresponds to the N-terminal methionine-excised form (termed FURm) previously shown to exist due to the overproduction conditions (*11*). After alkylation for 15 min, mainly two new peaks appear with 57 Da mass increases. They correspond to the addition of one alkylating group to both Fur (labeled FUR+1) and excised Fur (labeled FURm+1), thus indicating that one of the cysteines is more reactive. Its identification is described below. As the reaction continues, peaks corresponding to the addition of two (FUR+2) and four (FUR+4) alkylating groups steadily increase whereas FUR and FUR+1 progressively decrease and have practically disappeared after alkylation for 24 h. After alkylation for 6 h, the almost exclusive and simultaneous presence of equal intensities of FUR+2 and FUR+4 peaks clearly indicates that two of the cysteines reacted more slowly. It is noteworthy that FUR+3 species did not appear. After alkylation for 24 h, a small peak at 17 079 Da indicates that a fifth alkylation occurs. This is due to a nonspecific reaction, probably with one of the histidines, as sometimes occurs, especially with long reaction times (*29*). The different species which appear during the experiment and their respective masses are given in Table 1.

*Alkylation of Fur in the Presence of EDTA.* To determine if the two slow-reacting cysteines were protected due to their zinc binding, we performed the alkylation in the presence of EDTA. It was supposed that this divalent cation scavenger could trap the $Zn^{2+}$ ion and suppress the protection effect. The addition of a large excess of EDTA dramatically changed the alkylation kinetics obtained through the different mass spectra presented in Figure 2B. Although approximately the

Table 1: Different Species Which Appeared during the Experiment and Their Respective Masses[a]

| name | mass (Da) | name | mass (Da) |
|---|---|---|---|
| FUR | 16 794 | FURm | 16 663 |
| FUR+1 | 16 851 | FURm+1 | 16 720 |
| FUR+2 | 16 908 | FURm+2 | 16 777 |
| FUR+4 | 17 022 | FURm+4 | 16 891 |
| FUR+5 | 17 079 | FURm+5 | 16 948 |

[a] FURm corresponds to the N-terminal methionine-excised form; FUR+*x* and FURm+*x* correspond to each form of the Fur protein with *x* alkylations.

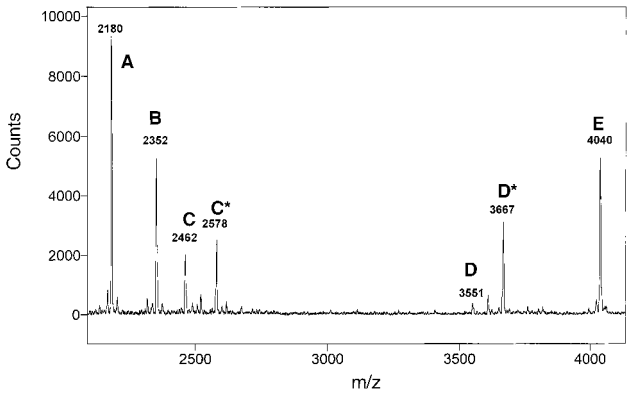



FIGURE 3: MALDI mass spectrum of the peptides obtained after enzymatic cleavage with endoproteinase Lys-C of a mixture of FUR molecules bearing two or four alkylating groups (corresponding to alkylation for 6 h in Figure 2A). The different peaks correspond to singly charged ions from several peptides of the protein: (A) V98−K116, (B) I21−K40, (C) S77−K97, and (C*) S77−K97 plus two alkylating groups, (D) H117−K147 and (D*) H117−K147 plus two alkylating groups, and (E) R41−K76.

same mass spectrum was obtained after alkylation for 15 min, with major peaks for FUR and FUR+1, the presence of FUR+2 was always very limited. In contrast with the experiment in the absence of EDTA, the peaks corresponding to FUR+4, already appearing at 15 min, progressively increase to become almost exclusive at 24 h. The almost complete disappearance of peaks corresponding to FUR+2, induced by the presence of EDTA, indicates that two cysteines involved in zinc binding became accessible to the alkylating reagent after removal of the zinc by EDTA treatment. As for the alkylation without EDTA, the minor species FUR+5 appears after alkylation for 24 h, indicating a nonspecific reaction.

*Identification of the Two Least Reactive Cysteines.* To identify the two cysteines which were less reactive in the absence of EDTA, specific proteolysis of the samples was carried out. The mixture of species FUR+2 and FUR+4, obtained after alkylation for 6 h (Figure 2A), was cleaved with endoproteinase Lys-C. It would have been obviously easier to work on the purified FUR+2 to characterize each cysteine. Unfortunately, it was not possible as we did not find chromatographic conditions that were suitable for the separation of the different forms of Fur obtained after alkylation. A MALDI mass spectrum of the crude mixture of the resulting peptides is presented in Figure 3. A peak at *m*/*z* 2462 precisely corresponds to the mono charged ion of peptide Ser77−Lys97 (C), possessing Cys92 and Cys95 which were not alkylated. Another peak at *m*/*z* 2578 (C*), with a mass increase of 116 Da, corresponds to the same peptide with two alkylating groups. The intensities of both

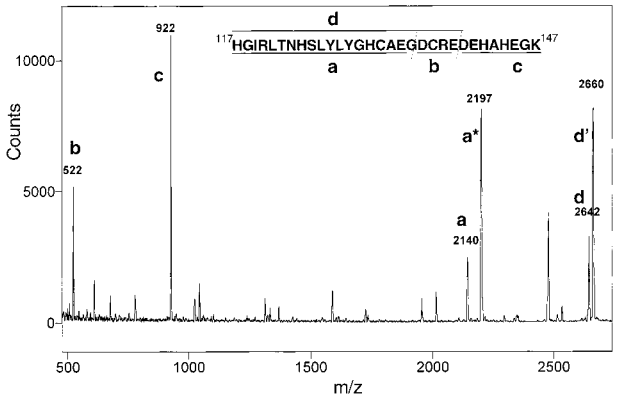



FIGURE 4: MALDI mass spectrum of the peptides obtained after enzymatic cleavage with endoproteinase Asp-N of a mixture containing peptide H117−K147 bearing one alkylating group and peptide H117−K147 free: (a) H117−G135, (a*) H117−G135 plus one alkylating group, (b) D136−E139, (c) D140−K147, (d) H117−E139, and (d′) H117−G135 linked to D136−E139 with a disulfide bridge.

peaks are similar. In contrast, the peptide His117−Lys147, containing Cys132 and Cys137, exhibited a very small peak at *m*/*z* 3551 (D) and a much more intense peak at *m*/*z* 3667 (D*), indicating that Cys132 and Cys137 are almost fully alkylated. From these results, it appears that a significant proportion of Cys92 and Cys95 was protected from alkylation whereas Cys132 and Cys137 were practically fully alkylated. The values of the three other major peaks precisely fit the calculated masses of three peptides expected from the endoproteinase Lys-C treatment.

*Identification of the Fast-Reacting Cysteine.* As shown in panels A and B of Figure 2 after alkylation for 30 min, one of the cysteines was much more reactive, giving FUR+1 species. With the objective of identifying it, we cleaved the mixture obtained after alkylation for 30 min without EDTA, using endoproteinase Lys-C. The peptides were analyzed using MALDI mass spectrometry. The peptide Ser77−Lys97 was practically not alkylated, giving a peak at *m*/*z* 2461. The traces of alkylated forms were due to the small amount of FUR+4 present in the mixture. After HPLC separation, a fraction gave two major peaks at *m*/*z* 3551 and 3607 corresponding to the peptide His117−Lys147 with zero and one alkylating group, respectively (data not shown). This fraction was further digested using endoproteinase Asp-N. The resulting peptides were analyzed without separation using MALDI mass spectrometry, giving the mass spectrum in Figure 4. A peak at *m*/*z* 2660 likely corresponds to peptides a and b linked together by a disulfide bridge that was formed during the digestion of the nonalkylated form (peptide d′). It is likely that because of the disulfide bridge this digestion was not complete, producing the peptide d at *m*/*z* 2642. A peak at *m*/*z* 2197 can be attributed to peptide His117−Gly135 (a*) with one alkylating group. The presence of this modified peptide indicates that Cys132 was alkylated. In contrast, a peak at *m*/*z* 522, corresponding to the peptide Asp136−Glu139 (b), indicates that Cys137 was not alkylated. Cys132 is the fastest-reacting cysteine and therefore the most accessible.

*Specificity of Iodoacetamide*. Iodoacetamide may react with histidines, especially with long reaction times (*29*). We determined that the alkylations we observed occurred only with cysteines and not with histidines. FUR+4 obtained after

alkylation for 6 h in the presence of EDTA was cleaved with endoproteinase Lys-C. The masses of the resulting peptides were measured, indicating that the peptides without cysteines were not alkylated. After HPLC separation, the peptides Ser77−Lys97 and His117−Lys147 were cleaved with endoproteinase Asp-N and chymotrypsin, respectively. The two resulting peptides possessing both histidines and cysteines were sequenced using tandem mass spectrometry or Edman degradation. The others were mass analyzed. We unambiguously found that the four cysteines were the only alkylated amino acids.

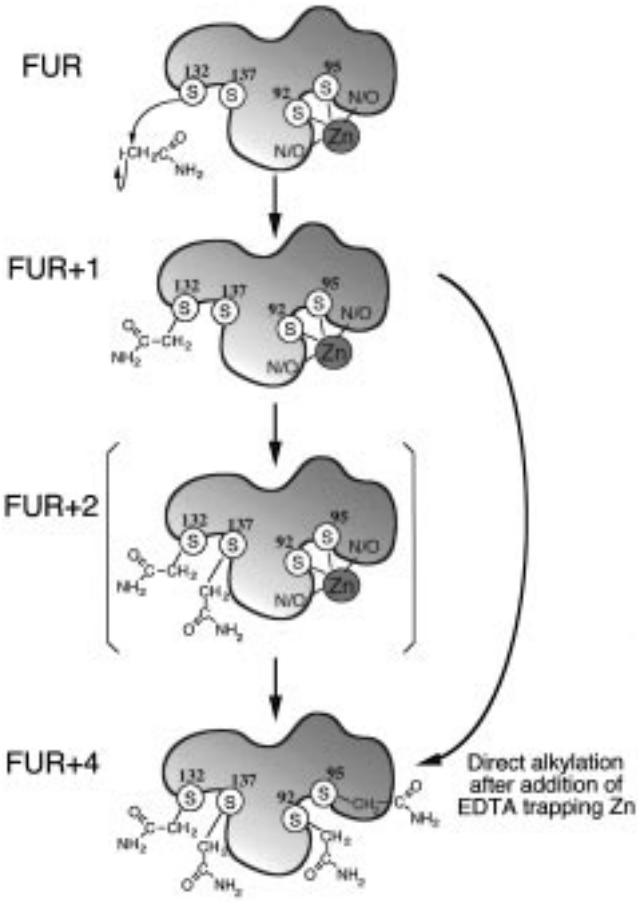
## DISCUSSION

The Fur protein, a metalloregulatory protein, has been proposed to utilize iron in vivo as a corepressor and then to act as a negative regulator via sequence-specific protein−DNA interactions. Fur also has a tight binding zinc site, and EXAFS analysis enabled the first structural characterization of this site (*20*). The chemical environment was proposed to consist of two sulfur and two light (N or O) atoms, at least one of them being a histidine. The sulfur ligands were expected to be cysteines since the Zn−S distances (2.3 Å) are in perfect agreement with Zn−S distances found in the literature for sulfur atoms from zinc-bound cysteine molecules (*20*). In this study, we implemented a differential alkylation study with and without EDTA, a strong zinc chelator, to characterize the behavior of each four cysteines. The results are summarized in Scheme 1.

*Fast-Reacting Cysteine.* Under all the experimental conditions that we tried, one of the cysteines was more reactive than the others, indicating that it is more accessible. After proteolysis experiments and mass spectrometry measurements, we identified it as Cys132. This cysteine thus likely belongs to an exposed portion of the Fur protein. This is in agreement with secondary structure predictions, based on primary sequence, assuming a turn between Gly130 and Asp136 (*14*−*16*). It gives a new partial element of the tertiary structure of Fur.

*Two Cysteines Are Protected from Alkylation by the Zinc Atom.* After alkylation for 6 h without EDTA, the simultaneous presence of forms with two and four alkyl groups clearly indicated that a portion of the molecules of the Fur protein that were examined had two protected cysteines and two accessible cysteines, whereas a portion of the molecules possessed four accessible cysteines. Because of the denaturing conditions used to obtain the mass spectra with a good sensitivity, we were unable to show the presence of zinc or of the dimeric form of the Fur protein. The stoichiometry of zinc has not yet been clearly determined as Fur has been shown to contain 0.5−0.8 zinc molecule per monomer from ICP analysis (*11*). These values do not allow one to determine if there is one zinc per monomer or one zinc per dimer. Furthermore, an underestimation of the zinc amount in Fur was possible since the protein samples were treated with EDTA after passage through the chelating zinc iminodiacetate column to remove the nonspecific zinc (*20*). The Fur molecules possessing four accessible cysteines might thus be formed without zinc. FUR+2 species, indicative of Fur molecules possessing a zinc, were more abundant than FUR+4 in the beginning of the alkylation reaction (up to 6 h). Afterward, they became less and less abundant from 24 to 36 h (data not shown). This evolution is indicative of a zinc displacement induced by the high concentration of the alkylating agent. The absence of FUR+3 indicates that the two least reactive cysteines react together, which is induced by a common gain of accessibility. The use of EDTA in the subsequent alkylation experiments clearly showed that the zinc protection of the two cysteines was eliminated since the FUR+2 species had almost disappeared. This disappearance also ruled out the presence of a disulfide bridge which could also have accounted for the FUR+2 species obtained without EDTA.

Scheme 1: Schematic Representation of the Process of FUR Alkylation by Iodoacetamide[a]




[a] Only one monomer is represented for clarity. One cysteine reacts very fast, and then another one, in leading to an intermediate species denoted FUR+2. The other two cysteines are protected from alkylation by the zinc atom, and a longer reaction time is required to modify them and obtain the final species denoted FUR+4. Addition of EDTA during the reaction enables deprotection of these two cysteines, and the fully alkylated protein can be obtained directly, without any intermediate.

*Identification of These Two Protected Cysteines*. The identification of the two zinc-protected cysteines was achieved after a proteolytic cleavage of the mixture of species FUR+2 and FUR+4. The peptide bearing Cys92 and Cys95 was found in both alkylated (two alkylations) and nonalkylated forms. This result was consistent with the two cysteines being part of the zinc site. The essential role played by these cysteines, which are well-conserved in the sequences of the Fur proteins, has already been demonstrated by directed mutagenesis. Their replacement by serines caused a major loss of activity of Fur both in vivo and in vitro (*25*). In vivo,

each of these mutations depressed the activity of the proteins from 4- to 10-fold. Furthermore, the destabilization induced by these two mutations was confirmed by proteolysis experiments; both mutant proteins were 10-fold more sensitive to trypsin digestion than wild-type Fur (*25*). In contrast, in vitro, the mutations seemed to lower activity of the proteins about 100-fold. This difference suggested that the structure of these proteins was destabilized such that there was a loss of activity during the purification process (*25*). It would be interesting to understand how the Fur from bacteria such as *R. leguminosarum* (*21*), *P. putida* (*22*), and *P. aeruginosa* (*23*) which do not contain the four conserved cysteines are managing to conserve an active and stable conformation.

*A New Type of Structural Zinc Site?* The EXAFS data together with the present results strongly suggest the stoichiometry of one $Zn^{2+}$ ion per Fur monomer, each one ligated to cysteines 92 and 95. To our knowledge, it is the first time that a combination of chemical modification, the use of a zinc scavenger, and mass spectrometry enables the identification of ligands of structural zinc. In a recent study, mass spectrometry was also used in association with a chemical modification of cysteines to determine which ones were readily accessible in *E. coli* dihydroorotase (*30*). The authors assumed that the two accessible cysteines that they had identified were potential ligands of structural zinc. In our study, we identified the two cysteines of Fur that became accessible with the action of EDTA, a strong chelator of zinc. We thus assumed that these two cysteines were ligands of structural zinc, which gives new insight into the three-dimensional structure of the Fur protein.

Over the course of more than fifty years, zinc has been shown to be an essential element of the active site in many enzymes (*31*). Indeed, in addition to serving as a catalytic metal, Zn(II) often plays a significant and purely structural role in many metalloproteins such as in alcohol dehydrogenase (*32*). In this case, the metal is tetracoordinated and has two or four cysteine ligands (*24*, *32*). Karlin et al. (*33*) recently proposed a classification of the zinc sites which included structural zinc site-containing thiolate ligands in class III of their classification. Within the structural zinc sites, there is a subfamily of the DNA-binding zinc proteins which contains three distinct motifs: zinc fingers, zinc clusters, and zinc twists (*34*). In proteins such as metallothionein and GAL4 possessing zinc clusters, six cysteines are ligands of two zinc atoms (*34*). In proteins such as glucocorticoid receptor possessing a zinc twist, each zinc binds to four cysteines (*34*). In zinc finger domains, zinc binds four ligands from the side chains of cysteine, histidine, and occasionally aspartate and glutamate (*24*). Ten classes of zinc fingers domains have been discovered and biochemically characterized (*32*). Each class is defined by a consensus sequence involving two to eight cysteines and zero to two histidines with a frequent CysXXCys motif. The two cysteines that we identified in this study as zinc ligands (Cys92 and Cys95) are arranged in a CysXXCys motif. This motif is very well conserved within the Fur proteins. Although Fur is a DNA-binding protein, it does not possess in its sequence the full consensus sequence which could allow it to be classified into one of the families of zinc finger domains identified previously (*12*, *24*). Furthermore, in zinc finger domains, the specificity of the DNA−protein recognition is usually contained in the peptide loop of the zinc finger (*24*). In the Fur protein, the DNA-binding activity is proposed to be in the N-terminal region (*12*, *13*, *35*), devoid of cysteines and far from the zinc site in the C-terminal region (*20*).

The presence of a structural zinc binding site is not surprising since the interior of the cell is fairly strongly reducing, through relatively high concentrations of glutathione; therefore, disulfide bridges, which are so commonly exploited by extracellular proteins, are unstable and hence rare in the intracellular environment. Since metal-binding sites such as zinc sites are insensitive to a reducing environment, they can be used to stabilize intracellular proteins (*32*).

Furthermore, the EXAFS analysis gave a tetrahedral environment for the zinc atom of Fur with two cysteines and two N or O donor ligands, one of them at least being an histidine. Using a similar approach, we plan to identify the two other zinc ligands. A histidine-rich domain close to the cysteine motif (His85, -86, -87, and -89 in *E. coli* Fur) is highly conserved within the Fur proteins. His124 of the *E. coli* Fur is also conserved in other Fur proteins. We will determine whether one or two of these histidines are the other ligands of the zinc atom. Whereas none of the consensus sequence found in the known zinc finger motifs was present within the Fur amino acid sequence, our findings might enable us to characterize a zinc domain belonging to a new class. Whether this new zinc site is involved in DNA binding or in the dimerization of Fur is still a question to be answered.

## ACKNOWLEDGMENT

We thank Jean Gagnon and Andrew Mundt for helpful discussions. Jean-Pierre Andrieu is also acknowledged for the sequencing analysis.

## REFERENCES

1. Crichton, R. R., and Charloteaux-Wauters, M. (1987) *Eur. J. Biochem. 164*, 485−506.
2. Neilands, J. B. (1995) *J. Biol. Chem. 270*, 26723−26726.
3. Braun, V., Schäffer, S., Hantke, K., and Tröger, W. (1990) in *41. Colloquium Mosbach 1990*, pp 164−179, Springer-Verlag, Berlin and Heidelberg.
4. Braun, V., Hantke, K., and Köster, W. (1998) in *Metal Ions in Biological Systems*, Vol. 35, pp 67−145, Marcel Dekker, Inc., New York.
5. Stojiljkovic, I., Bäumler, A. J., and Hantke, K. (1994) *J. Mol. Biol. 236*, 531−545.
6. Bagg, A., and Neilands, J. B. (1987) *Biochemistry 26*, 5471−5477.
7. Calderwood, S. B., and Mekalanos, J. J. (1988) *J. Bacteriol. 170*, 1015−1017.
8. De Lorenzo, V., Wee, S., Herrero, M., and Neilands, J. B. (1987) *J. Bacteriol. 169*, 2624−2630.
9. Touati, D., Jacques, M., Tardat, B., Bouchard, L., and Despied, S. (1995) *J. Bacteriol. 177*, 2305−2314.
10. Neilands, J. B., and Nakamura, K. (1991) in *Handbook of Microbial Iron Chelates* (Winkelmann, G., Ed.) pp 1−14, CRC Press, Boca Raton, FL.
11. Michaud-Soret, I., Adrait, A., Jaquinod, M., Forest, E., Touati, D., and Latour, J. M. (1997) *FEBS Lett. 413*, 473−476.
12. Coy, M. (1995) *Biochem. Biophys. Res. Commun. 212*, 784−792.
13. Stojiljkovic, I., and Hantke, K. (1995) *Mol. Gen. Genet. 247*, 199−205.
14. Saito, T., and Williams, R. P. J. (1991) *Eur. J. Biochem. 197*, 43−47.

15. Saito, T., Duly, D., and Williams, R. P. J. (1991) *Eur. J. Biochem. 197*, 39−42.
16. Saito, T., Wormald, M. R., and Williams, R. P. J. (1991) *Eur. J. Biochem. 197*, 29−38.
17. Hamed, M. Y., and Neilands, J. B. (1994) *J. Inorg. Biochem. 53*, 235−248.
18. Hamed, M. Y. (1993) *J. Inorg. Biochem. 15*, 193−210.
19. Adrait, A., Jacquamet, L., Le Pape, L., Gonzalez de Peredo, A., Aberdam, D., Hazemann, J. L., Latour, J. M., and Michaud-Soret, I. (1999) *Biochemistry 38*, 6248−6260.
20. Jacquamet, L., Aberdam, D., Adrait, A., Hazemann, J. L., Latour, J. M., and Michaud-Soret, I. (1998) *Biochemistry 37*, 2564−2571.
21. de Luca, N. G., Wexler, M., Pereira, M. J., Yeoman, K. H., and Johnston, A. W. B. (1998) *FEMS Microbiol. Lett. 168*, 289−295.
22. Venturi, V., Ottevanger, C., Bracke, M., and Weisbeek, P. (1995) *Mol. Microbiol. 15*, 1081−1093.
23. Prince, R. W., Cox, C. D., and Vasil, M. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 175*, 2589−2593.
24. Berg, J. M., and Shi, Y. (1996) *Science 271*, 1081−1085.
25. Coy, M., Doyle, C., Besser, J., and Neilands, J. B. (1994) *BioMetals 7*, 292−298.
26. del Cardayre, S., and Neilands, J. B. (1991) in *Iron Biominerals* (Frankel, R. E., Ed.) pp 387−396, Plenum Press, New York.
27. Tardat, B., and Touati, D. (1993) *Mol. Microbiol. 9*, 53−63.
28. Wee, S., Neilands, J. B., Bittner, M. L., Hemming, B. C., Haymore, B. L., and Seetharam, R. (1988) *Biol. Met. 1*, 62−68.
29. Gurd, F. (1972) *Methods Enzymol. 25*, 424−438.
30. Daniel, R., Caminade, E., Martel, A., Le Goffic, F., Canosa, D., Carrasca, M., and Abian, J. (1997) *J. Biol. Chem. 272*, 26934−26939.
31. Coleman, J. E. (1992) *Annu. Rev. Biochem. 61*, 897−946.
32. Schwabe, J. W. R., and Klug, A. (1994) *Struct. Biol. 1*, 345−349.
33. Karlin, S., and Zhu, Z.-Y. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 14231−14236.
34. Vallee, B. L., Coleman, J. E., and Auld, D. S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 999−1003.
35. Coy, M., and Neilands, J. B. (1991) *Biochemistry 30*, 8201−8210.
36. Schäffer, S., Hantke, K., and Braun, V. (1985) *Mol. Gen. Genet. 200*, 110−113.
37. Berish, S. A., Subbarao, S., Chen, C.-Y., Trees, D. L., and Morse, S. A. (1993) *Infect. Immun. 61*, 4599−4606.
38. Staggs, T. M., and Perry, R. D. (1992) *Mol. Microbiol. 6*, 2507−2516.
39. Litwin, C. M., Boyko, S. A., and Calderwood, S. B. (1992) *J. Bacteriol. 174*, 1897−1903.
40. Wooldridge, K. G., Williams, P. H., and Ketley, J. M. (1994) *J. Bacteriol. 176*, 5852−5856.
41. Franza, T., Sauvage, C., and Expert, D. (1999) *Mol. Plant-Microbe Interact.* (in press).

BI9902283