Unraveling the *Helicobacter pylori* UreG zinc binding site using X-ray absorption spectroscopy (XAS) and structural modeling

Vlad Martin-Diaconescu,
Department of Chemistry, University of Massachusetts, Amherst 01003, MA, USA

Matteo Bellucci,
Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Bologna, Italy

Francesco Musiani,
Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Bologna, Italy

Stefano Ciurli, and
Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Bologna, Italy; Center for Magnetic Resonance (CERM), University of Florence, Florence, Italy

Michael J. Maroney
Department of Chemistry, University of Massachusetts, Amherst 01003, MA, USA

Abstract

The pathogenicity of *Helicobacter pylori* depends on the activity of urease for pH modification. Urease activity requires assembly of a dinickel active site that is facilitated in part by GTP hydrolysis by UreG. The proper functioning of *Helicobacter pylori* UreG (*Hp*UreG) is dependent on Zn(II) binding and dimerization. X-ray absorption spectroscopy and structural modeling were used to elucidate the structure of the Zn(II) site in *Hp*UreG. These studies independently indicated a site at the dimer interface that has trigonal bipyramidal geometry and is composed of two axial cysteines at 2.29(2)Å, two equatorial histidines at 1.99(1)Å, and a solvent-accessible coordination site. The final model for the Zn(II) site structure was determined by refining multiple-scattering extended X-ray absorption fine structure fits using the geometry predicted by homology modeling and ab initio calculations.

Keywords

X-ray absorption spectroscopy; *Helicobacter pylori* UreG; Nickel trafficking; Homology modeling; Urease

Introduction

Helicobacter pylori is a human pathogen affecting almost half the world's population [1, 2]. It can lead to severe conditions such as chronic gastritis and duodenal and peptic ulcers and

mmaroney@chemistry.umass.edu, stefano.ciurli@unibo.it.

may lead to gastric carcinoma, which accounts for 10% of cancer deaths in the world [1-3]. The ability of the bacterium to colonize the gastric mucosa despite the acidic conditions present in the stomach is dependent on both active and passive acid resistance mechanisms. These include membrane proteins with high isoelectric points that reduce membrane proton permeability as well as mechanisms that actively neutralize pH through the production of ammonia [4]. Ammonia formation is catalyzed by the enzyme urease, which hydrolyzes urea to yield ammonia and bicarbonate [5]. *H. pylori* urease is a nickel-dependent enzyme whose maturation involves the formation of a supramolecular complex [(UreAB)$_3$]$_4$ [6] that is facilitated by a host of accessory proteins. These include UreD and UreF, involved in protein-protein interactions and protein conformational changes [7, 8], and a metallochaperone (UreE) [9, 10] that delivers nickel and appears to stimulate the nucleotide triphosphate hydrolase activity performed by UreG [11]. Urease activation involves the incorporation of two Ni(II) ions and the posttranslational modification of a conserved active site lysine to form a carbamate bridge between the metals [12]. Formation of the bridging lysine carbamate is accomplished by guanosine triphosphate (GTP)-dependent transfer of a CO_2 molecule. The protein responsible for GTP hydrolysis is *H. pylori* UreG (*Hp*UreG) [13, 14].

Several factors are believed to be important for *Hp*UreG function, including formation of a multiprotein complex, GTP binding and GDP release mechanisms, and monomer/dimer equilibria that depend on metal binding [11]. Specifically, it has been found that Zn(II) induces *Hp*UreG dimerization, with a stoichiometry of one Zn(II) per dimer [15]. Moreover, Zn(II) binding has been demonstrated to strengthen the interaction between *Hp*UreG and *H. pylori* UreE (HpUreE), the metallochaperone that appears to provide nickel for urease maturation [10]. This suggests a structural role for Zn(II), which may facilitate the folding of *Hp*UreG, an enzyme belonging to what can be described as a class of intrinsically disordered proteins [16, 17]. In vitro site-directed mutagenesis, supported by homology modeling of the apoprotein [15] and by in vivo assays [18], has identified a set of amino acid residues that are essential for binding of zinc to *Hp*UreG. The residues Cys-66 and His-68 are part of a conserved Cys-Pro-His motif in a region of the amino acid sequence that has a high propensity to be non-structured according to disorder predictors [15, 16]. X-ray absorption spectroscopy (XAS) was applied at the zinc *K*-edge to structurally characterize the zinc binding site, leading to a better understanding of the coordination number, ligand speciation, and metal–ligand bond distances. The models derived from the extended X-ray absorption fine structure (EXAFS) analysis were compared with UreG models independently obtained from molecular modeling of the holoprotein. The coordination geometry at the zinc binding site was further refined by building an EXAFS fitting model that combines molecular modeled coordination geometries with the coordination number, bond distances, and ligands determined by EXAFS analysis.

Materials and methods

*Hp*UreG expression and purification

Large-scale expression of *Hp*UreG in *Escherichia coli* BL21(DE3) host cells (Novagen) was conducted as previously described [15]. Cells were grown in 2 L batches of M9 autoinduction medium at 28 °C for 48 h, starting from 50 mL of preinoculate cell culture that was grown for 16 h. Cells were harvested by centrifugation at 8,000g for 20 min at 4 °C. The pellet was resuspended in 30 mL of 50 mM tris(hydroxymethyl)aminomethane, Tris–HCl, pH 7.0, containing 2 mM ethylenediaminetetraacetic acid (EDTA), 2 mM dithiothreitol (DTT), 10 mM MgCl$_2$, and 20 μg mL^{-1} DNAse I. Cells were disrupted by two passages through a French pressure cell (SLM-Aminco) operating at 20,000 psi. Cell debris was separated from the supernatant by centrifugation at 15,000g for 30 min at 4 °C.

The protein was isolated using a combination of anion-exchange and size-exclusion chromatographic separations. The soluble fraction obtained from the cell lysate was loaded onto a Q-Sepharose XK 26/10 column (GE Healthcare) that had been preequilibrated with 20 mM Tris–HCl buffer, pH 7.0, containing 2 mM EDTA. The column was washed with the starting buffer until the baseline was stable. The protein was eluted from the column with a 500-mL linear gradient of NaCl (0–1 M). The fractions containing *Hp*UreG were collected, and the protein was concentrated using 5 kDa molecular mass cutoff Amicon and Centricon ultrafiltration units (Millipore) after addition of 2 mM DTT. The sample obtained was loaded onto a Superdex 75 XK 16/60 column (GE Healthcare) previously conditioned with 20 mM Tris–HCl buffer, at pH 7.0, containing 150 mM NaCl and 10 mM EDTA. The protein fractions were collected, stored in 2 mM DTT, and further purified by another step of size-exclusion chromatography with Superdex 75 HR 10/30 (GE Healthcare), using 20 mM Tris–HCl, pH 7.0, containing 150 mM NaCl. Purified *Hp*UreG samples were concentrated using 5 kDa molecular mass cutoff Centricon ultrafiltration units (Millipore), until a final volume of 0.5–1 mL had been reached. For XAS studies, the protein buffer was 20 mM Tris–HCl, pH 7.0, containing 150 mM NaBr.

In each purification step, the purity of *Hp*UreG, as well as its molecular mass under denaturing conditions, was estimated by sodium dodecyl sulfate polyacrylamide gel electrophoresis using a NuPAGE Novex precast gel system (Invitrogen). Proteins were separated on NuPAGE 4–12% [bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane] gels (Invitrogen) and stained using SimplyBlue Safestain (Invitrogen). Protein concentration was determined using the theoretical value of the molar extinction coefficient at 280 nm ($10,095 M^{-1} cm^{-1}$), calculated from the protein sequence using the ProtParam webpage (http://au.expasy.org/tools/protparam.html). This value was confirmed by absolute protein quantitation, performed by measuring the amount of sulfur in a protein sample with known absorbance at 280 nm by inductively coupled plasma emission spectroscopy, as previously described [15]. The absence of any metal bound to the purified protein was confirmed by inductively coupled plasma emission spectroscopy using a procedure described elsewhere [19].

EXAFS sample preparation, data collection, and analysis

To a stock solution of UreG monomer (470 μM) 0.45 equiv of zinc(II) sulfate was added. UreG is expected to fully bind zinc with a K_d of 0.33 μM. The protein sample was incubated for 5 min upon metal ion addition, loaded into sample cells consisting of a polycarbonate sample holder with a Kapton window, and frozen in liquid nitrogen. The sample was then stored at −80 °C and transported in liquid nitrogen before exposure to synchrotron radiation. The samples were run at the Stanford Synchrotron Radiation Lightsource (3-GeV ring) on beam line 9-3 equipped with a Si(220) double-crystal monochromator, a liquidhelium cryostat, and a 100-element X-ray fluorescence detector array (Canberra). Scattering was reduced by placing a Söller slit with a Z-1 element filter.

The SixPack software program was used to remove bad detector elements, to average the data, and to perform energy calibration [20]. An internal energy calibration was performed using the first inflection point in the X-ray absorption near edge structure (XANES) spectrum of a zinc foil (E_{cal} of 9,660.7 eV). Data reduction and normalization was performed with the Athena software program using the AUTOBK algorithm [21]. R_{bkg} was set to 0.9Å and a spline between $k = 1 Å^{-1}$ and $k = 14.5 Å^{-1}$ was used for data extraction. E_0 was set to 9,670 eV for the zinc K-edge data. Edge normalization and background subtraction was performed using a linear pre-edge function (between −200 and −30 eV with respect to E_0) and a quadratic polynomial for the postedge region (between 300 and 800 eV with respect to E_0) followed by normalization of the edge jump to 1.

The Artemis software program was used for EXAFS analysis with parameters generated using FEFF6, and scattering paths were fit using the IFEFFIT engine [22, 23]. The k^3-weighted data were fit in r-space over a k range of 2–13.5Å^{-1}, with an S_0 value of 0.9, and with a Kaiser–Bessel window with a windowsill of 2. The spectra were not phase-corrected during fitting and are presented as such in the text with the r-space axis being labeled "$R + \Delta$" in the figures to emphasize this point. Separate sets of Δr_{eff} and σ^2 for the sulfur, nitrogen, and bromide ligands were used with initial values of 0.0Å and 0.003Å^2, respectively, with a universal E_0 initially set to 9,670.0 eV and ΔE_0 of 0 eV. Initial input metal–ligand distances were 2.0Å for Zn–N, 2.3Å for Zn–S, and 2.4Å for Zn–Br. Single-scatter fits were performed over an r-space range of 1–2.5Å, and multiple-scattering contributions from histidines were explored as described in the text using a 1–4.5-Å r-space range. The final model using the geometry predicted from homology-based computational studies also included multiple-scattering components expected for transaxial sulfur donors.

To assess the goodness of fit from different fitting models, the fit parameters χ^2, reduced χ^2 ($\chi_\nu{}^2{}_{\text{m}}$) and the R factor were minimized. Increasing the number of adjustable parameters is generally expected to improve the R factor; however $\chi_\nu{}^2$ may go through a minimum then increase, indicating the model is overfitting the data. These parameters are defined as follows:

$$\chi^2 = \frac{N_{\text{idp}}}{N_{\text{pts}}\varepsilon^2} \sum_{t=1}^{N} \left\{ \left[\text{Re}\left(\chi_{\text{data}}(R_i) - \chi_{\text{theory}}(R_i)\right) \right]^2 + \left[\text{Im}\left(\chi_{\text{data}}(R_i) - \chi_{\text{theory}}(R_i)\right) \right]^2 \right\},$$

and

$$\chi_\nu^2 = \frac{\chi^2}{\left(N_{\text{idp}} - N_{\text{var}}\right)},$$

where N_{idp} is the number of independent data points defined as $N_{\text{idp}} = (2\Delta r \Delta k)/\pi$, Δr is the fitting range in r-space, Δk the fitting range in k-space, N_{pts} is the number of points in the fitting range, N_{var} is the number of variables floating during the fit, ε is the measurement uncertainty, Re() is the real part of the EXAFS Fourier-transformed data and theory functions, Im() is the imaginary part of the EXAFS Fourier-transformed data and theory functions, $\chi(R_i)$ is the Fourier-transformed data or theory function, and

$$R = \frac{\sum_{i=1}^{n} \left\{ \left[\text{Re}\left(\chi_{\text{data}}(R_i) - \chi_{\text{theory}}(R_i)\right) \right]^2 + \left[\text{Im}\left(\chi_{\text{data}}(R_i) - \chi_{\text{theory}}(R_i)\right) \right]^2 \right\}}{\sum_{i=1}^{N} \left\{ \left[\text{Re}\left(\chi_{\text{data}_i}(R_i)\right) \right]^2 + \left[\text{Im}\left(\chi_{\text{data}}(R_i)\right) \right]^2 \right\}}.$$

Structure prediction and homology modeling

An initial alignment of *Methanocaldococcus jannaschii* HypB (MjHypB) and *Hp*UreG sequences was produced using the program MODELLER version 9.8 [24] and this alignment was manually adjusted to match up the primary and the secondary structure of the proteins, as previously described [15]. In the resulting alignment, the cysteine and histidine residues involved in the binding of Zn(II) in the structure of MjHypB were aligned with the corresponding residues in the fully conserved Cys-Pro-His motif found in all UreG proteins [16]. This alignment was then used to calculate 100 structural models of the dimer of *Hp*UreG (which included all hydrogen atoms as well as two molecules of a GTPγS

analogue and associated Mg^{2+} ions in the positions found in the MjHypB structure) using MODELLER version 9.8 and the MjHypB dimer structure (Protein Data Bank code 2HF8) as a template. A twofold axis of symmetry passing through the center of the two monomers was imposed in the calculation. The Zn(II) binding site was modeled with one Zn(II) ion bound to Cys-66 S_c and His-68 N_e at 2.2(1) and 2.0(1)Å, respectively. These values were taken as the lower values for the ranges of the Zn–S_{Cys} and Zn–N_{His} bond lengths statistically derived using the Cambridge Structural Database [25, 26]. In all modeling calculations, angle and torsional constraints were included to correctly position the metal ion with respect to the bound residues. In particular, for the Zn–S_{Cys} bond, the Zn–S_γ–C_β angle was constrained to 109° ± 5°, whereas for the Zn-N_{His} bond the angles between Zn–N_e–C_e and Zn–N_e–C_δ were constrained to 120° ± 10° and the torsional angles between Zn–N_e–C_e–N_δ and Zn–C_e–C_δ–C_γ were constrained to 180° ± 10°. No restraints were added to constrain the coordination geometry around the Zn(II) ion.

The best model was selected on the basis of the lowest value of the MODELLER objective function. A loop optimization routine was used to refine the regions that showed higher than average energy as calculated using the DOPE [27] potential function built into MODELLER. The results of the PROCHECK analysis [28] for the final model were fully satisfactory. The entire modeling procedure was repeated using a water molecule [at 2.0(1)Å] bound to the Zn(II) ion in the vacant coordination site. This distance was taken as the lower value for the range of the Zn–O_{water} bond length statistically derived using the Cambridge Structural Database [25, 26].

The resulting metal binding site for the aqua complex was used as the starting structure for ab initio geometry optimizations, performed using the program ORCA 2.8.0 [29]. Histidine residues were modeled as methylimidazole rings and cysteines were modeled as methylthiolates. The optimization was performed at the B3LYP/6-31G(d) level. The parameters obtained from the ab initio geometry optimization are provided as electronic supplementary material. A model of *Hp*UreG was then calculated using MODELLER, which included the ORCA-based metal–ligand coordination geometric parameters.

Results

XAS analysis

XAS at the metal K edge was used to investigate the *Hp*UreG zinc binding site. The XANES region of the spectrum (Fig. 1) has two features at 9,665.7 and 9,670.8 eV. These well-resolved features are associated with $4p \leftarrow 1s$ transitions and are consistent with a Zn(II) center coordinated by a combination of S- and N/O-donor ligands. Previous studies showed the correlation between increased intensity of the zinc XANES features with increasing coordination number by analyzing a series of both inorganic and biological samples [30]. The XANES intensity of Zn–*Hp*UreG bound transitions was between 1.2 and 1.5, indicating a four- or five-coordinate zinc site.

At higher energies, electrons are excited to the continuum and their scattering interactions with nearby atoms give rise to the EXAFS component of the XAS spectrum (Fig. 2b). A second shell of scatterers is indicated by the splitting observed in the first EXAFS oscillations in the k-space spectrum between 3 and 5Å⁻¹ and is suggestive of histidine ligands [31]. The presence of histidines is rein-forced by the features in the 3–4-Å region of the Fourier-transformed EXAFS spectrum (*r*-space, Fig. 2a) [32]. The *r*-space spectrum also exhibits two intense peaks between 1 and 2.5Å that are consistent with at least two shells of scattering atoms in the primary coordination sphere. In biological samples, the first peak would generally be attributed to N/O-donor ligands, whereas the second peak usually consists of sulfur/halogen ligands.

The EXAFS data were first analyzed using single-scattering models and r-space data over a 1–2.5-Å range, and employing a systematic series of N/S ligand combinations (see the electronic supplementary material). These fits became approximately tenfold better based on R and χ_ν^2 values once contributions from bromide, present in the buffer, were included. Because chloride ligands are not easily distinguished by EXAFS from sulfur donors, such as cysteines, and aqua ligands are not easily distinguished from other N/O donors, the sample was purposely prepared in bromide-containing buffer to assess the presence of labile ligands and/or open coordination positions. The best single-scattering models (Table 1) had coordination numbers of 4–5 and had a shell of one or two sulfur donors, a Br⁻, and two N/ O donors. These fits were further analyzed using multiple scattering (see the electronic supplementary material). Fits generated with independent shells of sulfur scattering atoms gave unimproved or increased values of R and χ_ν^2, and σ^2 values with large errors. A six-coordinate geometry for the zinc site was ruled out on the basis of the XANES analysis (vide supra) and generally poorer EXAFS fits (see Table 1 and the electronic supplementary material).

The best single-scattering models were then analyzed for the presence of histidine imidazole ligation by multiple-scattering analysis using an extended fitting range with an upper limit in r-space of 4.5Å. Histidine ligands were fit with geometrically rigid imidazoles constructed by using average values for bond lengths and angles in the imidazole ring that were gathered from crystallographic data [33]. The radial distance of imidazole atoms from the zinc metal center was expressed in terms of the Zn–N bond and the rotation angle a, and all paths with an amplitude of more than 10% were included (see the electronic supplementary material) [34, 35]. The best fit was obtained using a combination of three shells containing two histidines at 1.99(1)Å with an α of 5°, two sulfurs at 2.29(2)Å, and a bromide at 2.40(1)Å (Table 2; Fig. 2). The fits obtained using two sulfur scatterers were approximately 30% better than the corresponding fit using only one sulfur scattering atom. The distances obtained from the EXAFS analysis are consistent with available statistical analyses of Zn(II) coordination geometries observed in near-atomic-resolution structures in the Protein Data Bank as well as in the structures of coordination complexes found in the Cambridge Structural Database [25, 26]. Therefore, the EXAFS analysis indicates that Zn(II) forms a five-coordinate complex with *Hp*UreG involving two cysteine and two histidine ligands, as well as a solvent-accessible coordination site that is occupied by bromide in the sample studied.

The ligand environment and geometry of the Zn(II) metal binding site was determined independently using molecular modeling techniques (see "Materials and methods"). The metal binding site resulting from the homology-based modeling procedure features a Zn(II) ion lying at the center of a trigonal bipyramidal ligand geometry, with the two axial Cys-66 S_γ atoms and two His-68 N_e atoms in the equatorial plane. An empty coordination site, in this case occupied by a water molecule, completes the five-coordinate geometry. This model is consistent with the pentacoordinated Zn(II) site observed by XAS (Fig. 3), where the aqua site is occupied by bromide. The average Zn(II)–ligand bond distances from the homology-based computational model are 2.25Å for Zn–S_{Cys}, 1.95Å for Zn–N_{His}, and 2.0Å for Zn– O_{water}. These values are in agreement with the accurate metal–ligand distances experimentally determined using EXAFS analysis (see Table 2) and with available statistical analyses [25, 26]. Geometry optimization of the metal binding site based on ab initio calculations in the gas phase, which included a water molecule in the empty coordination site, supports the structural stability of the coordination environment of this chemical moiety. The observed differences in the metal-ligand distances between the ab initio calculation on one hand and the homology-based modeling or the fits of the experimental EXAFS data on the other hand are likely due, at least in part, to the bromide versus water ligation of the Zn(II) ion in the two studies, and/or to the limitations of the theoretical

approach, which cannot fully account for the presence of the protein framework around the metal binding site and solvent effects.

A final fit of the EXAFS data was performed by using the geometry and ligand arrangement, predicted by computational modeling. This EXAFS model included multiple scattering components from the transaxial sulfur donors and resulted in an improved fit to the EXAFS data (approximately 20% in R and χ_ν^2), as discussed in the next section.

Discussion

Urease is a nickel metalloenzyme that catalyzes the breakdown of urea to ammonia and bicarbonate, and is thus involved in one of the primary mechanisms for pH homeostasis in *H. pylori* [4]. Hydrolysis of GTP is an essential step in urease maturation. In *H. pylori*, the GTP hydrolysis is catalyzed by UreG [8, 14]. The activity of UreG may be affected by metal binding and protein-protein interactions as well as monomer to dimer formation [11, 15]. In particular, it is known that zinc binding causes UreG monomer subunits to form homodimers. However, zinc binding alone does not promote enzymatic GTP hydrolysis, and additional protein-protein interactions are believed to be necessary for activity [15]. It has been observed that less GTP is required for urease activation in vitro in the presence of UreE, suggesting a possible role for a UreE–UreG interaction in GTP hydrolysis [36]. In fact, immunoprecipitation assays have shown a direct interaction between UreE and UreG [37], which is stabilized by zinc binding as illustrated by size exclusion chromatography and light scattering experiments [10]. Therefore, a thorough characterization of the UreG zinc binding site may lead to a better understanding of not only its GTPase activity but also the protein–protein interactions essential for its function.

Homology modeling of the UreG apoprotein and studies employing site-directed mutagenesis have proposed that Cys-66 and His-68 are possible ligands for zinc binding [15]. We performed a comprehensive characterization of the zinc binding site both in terms of the ligands involved and bond distances, and the coordination geometry by XAS and complemented by molecular modeling of the holoprotein. Inspection of the zinc K-edge XANES spectra of holo-*Hp*UreG points to a five- or four-coordinate zinc site. However, both single-scattering and multiple-scattering EXAFS fits are significantly better for a five-coordinate site over the four-coordinate models (Tables 1, 2, see also the electronic supplementary material). A poorer fit is achieved both in terms of goodness of fit (R) and χ_ν^2 values for both four- or six-coordinate models, and an N_2S_2Br ligand donor atom set is proposed for the zinc site. In the case of the multiple-scattering model incorporating histidine imidazole ligands, the six-coordinate $N_2^{(His)}S_3Br$ model has a σ^2 component of the Debye–Waller factor that not only is large, but also has a significantly larger error, allowing us to differentiate it from the $N_2^{(His)}S_2Br$ five-coordinate model despite the fact that both models have similar goodness of fit (Table 2). Furthermore, the addition of multiple-scattering paths from histidines is justified not only by the presence of features attributable to the presence of histidines in the EXAFS spectra of holo-*Hp*UreG (Fig. 2), but also by the threefold decrease in χ_ν^2 values of multiple-scattering fits when compared with the single-scattering analysis (see the electronic supplementary material). This improvement in χ_ν^2 values further supports the model including the addition of paths necessary to describe histidine scattering.

The XAS experiment also allowed the identification of a coordinating bromide ligand, indicating the presence of a solvent-accessible coordination site. The final coordination environment obtained from the EXAFS analysis contains two histidines at 1.99(1)Å with an α of 5°, two sulfurs at 2.29(2)Å, and a bromide at 2.40(1)Å (Table 2; Fig. 2). The resulting

bond distances are consistent with previously reported zinc–ligand distances found in proteins [25, 26].

The five-coordinate *Hp*UreG zinc binding site consisting of two histidines, two sulfur donors, and a solvent-accessible site is also independently supported by molecular modeling of holo-*Hp*UreG. The model shows the presence of a trigonal bipyramidal zinc binding site at the UreG homodimer interface, with one (axial) cysteine and one (equatorial) histidine ligand from each monomer subunit and a solvent-accessible coordination site in the third equatorial position (Fig. 3). The Zn(II)–ligand bond distances obtained from the homology–structural modeling are 2.25Å for Zn–S_{Cys}, 1.95Å for Zn–N_{His}, and 2.0Å for Zn–O_{water}, and are consistent with both the EXAFS data and available statistical analyses of zinc–ligand bond distances found from crystallographic data [25, 26]. XAS provided experimental confirmation of the theoretical prediction of the ligands involved in, and the geometry of, the zinc site, as well as accurate bond distances. Combining EXAFS analysis with the homology-based model allowed us to refine the multiple-scattering EXAFS model with the predicted coordination geometry. This resulted in an improved fit in terms of both goodness of fit (R) and χ_ν^2 (Table 3; Fig. 4). The bond distances determined are similar to those obtained in the initial EXAFS analysis, the largest difference being for the sulfur bond distance, which shifted from 2.29(2) to 2.25(2)Å, a difference that is not significant given the errors. The value of σ^2 for the shell of sulfur scattering atoms (0.013Å^2) is larger than that of the other shells, but falls within accepted limits for EXAFS fits [38]. The larger value might be attributed to a distance difference for the two Zn–S vectors that is below the resolution limit of the data (approximately 0.1Å) [39-41]. Attempts to split the shell into two separate shells gave poorer fits (see the electronic supplementary material). It is not clear whether this represents a real difference or is the result of a small amount of sample heterogeneity or small correlations between parameters in fitting the data [40-42]. However, geometrically unrestricted gas-phase ab initio calculations on the water-bound zinc model suggest that asymmetric binding of the water molecule may lead to distinct Zn–S bond distances (see the electronic supplementary material). This would lead to the conclusion that binding of solvent molecules may affect the Zn–S bond distances. This is expected to be minimal in the case of the symmetric bromide ligand. The resulting five-coordinate trigonal bipyramidal His$_2$Cys$_2$X (where X is a solvent molecule or anion) Zn(II) site with axial cysteine ligation is fully consistent with both the XAS data and the homology-based and ab initio structural modeling results. In addition, the predicted Zn(II) binding site supports the available model structure of the complex between *Hp*UreG and HpUreE, which shows His-68 and Cys-66 of *Hp*UreG forming part of a binding site within the HpUreE–*Hp*UreG complex [10].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

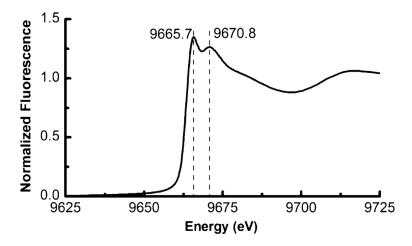|	
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EXAFS	Extended X-ray absorption fine structure
GTP	Guanosine triphosphate
HpUreE	Helicobacter pylori UreE
*Hp*UreG	Helicobacter pylori UreG
MjHypB	*Methanocaldococcus jannaschii* HypB
Tris	Tris(hydroxymethyl)aminomethane
XANES	X-ray absorption near edge structure
XAS	X-ray absorption spectroscopy

References

1. Bruce MG, Maaroos HI. Helicobacter. 2008; 13:1–6. [PubMed: 18783514]

2. Ferreira AC, Isomoto H, Moriyama M, Fujioka T, Machado JC, Yamaoka Y. Helicobacter. 2008; 13:28–34. [PubMed: 18783519]

3. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. CA Cancer J Clin. 2011; 61:69–90. doi: 10.3322/caac.20107. [PubMed: 21296855]

4. Sachs G, Weeks DL, Melchers K, Scott DR. Annu Rev Physiol. 2003; 65:349–369. doi: 10.1146/annurev.physiol.65.092101.142156. [PubMed: 12471160]

5. Zambelli B, Musiani F, Benini S, Ciurli S. Acc Chem Res. 2011; 44(7):520–530. doi: 10.1021/ar200041k. [PubMed: 21542631]

6. Ha NC, Oh ST, Sung JY, Cha KA, Lee MH, Oh BH. Nat Struct Biol. 2001; 8:505–509. [PubMed: 11373617]

7. Lam R, Romanov V, Johns K, Battaile KP, Wu-Brown J, Guthrie JL, Hausinger RP, Pai EF, Chirgadze NY. Proteins Struct Funct Bioinform. 2010; 78:2839–2848. doi: 10.1002/prot.22802.

8. Moncrief MBC, Hausinger RP. J Bacteriol. 1996; 178:5417–5421. [PubMed: 8808930]

9. Benoit S, Maier RJ. J Bacteriol. 2003; 185:4787–4795. doi: 10.1128/jb.185.16.4787-4795.2003. [PubMed: 12896998]

10. Bellucci M, Zambelli B, Musiani F, Turano P, Ciurli S. Biochem J. 2009; 422:91–100. doi: 10.1042/bj20090434. [PubMed: 19476442]

11. Zambelli B, Musiani F, Savini M, Tucker P, Ciurli S. Biochemistry. 2007; 46:3171–3182. doi: 10.1021/bi6024676. [PubMed: 17309280]

12. Ciurli S, Benini S, Rypniewski WR, Wilson KS, Miletti S, Mangani S. Coord Chem Rev. 1999; 192:331–355.

13. Moncrief MBC, Hausinger RP. J Bacteriol. 1997; 179:4081–4086. [PubMed: 9209019]

14. Soriano A, Hausinger RP. Proc Natl Acad Sci USA. 1999; 96:11140–11144. [PubMed: 10500143]

15. Zambelli B, Turano P, Musiani F, Neyroz P, Ciurli S. Proteins Struct Funct Bioinform. 2009; 74:222–239. doi: 10.1002/prot.22205.

16. Zambelli B, Stola M, Musiani F, De Vriendt K, Samyn B, Devreese B, Van Beeumen J, Turano P, Dikiy A, Bryant DA, Ciurli S. J Biol Chem. 2005; 280:4684–4695. doi: 10.1074/jbc.M408483200. [PubMed: 15542602]

17. Dyson HJ, Wright PE. Nat Rev Mol Cell Biol. 2005; 6:197–208. doi: 10.1038/nrm1589. [PubMed: 15738986]

18. Boer JL, Quiroz-Valenzuela S, Anderson KL, Hausinger RP. Biochemistry. 2010; 49:5859–5869. doi: 10.1021/bi1004987. [PubMed: 20533838]

19. Stola M, Musiani F, Mangani S, Turano P, Safarov N, Zambelli B, Ciurli S. Biochemistry. 2006; 45:6495–6509. doi: 10.1021/bi0601003. [PubMed: 16700560]

20. Webb SM. Phys Scr. 2005; T115:1011–1014.

21. Ravel B, Newville M. J Synchrotron Radiat. 2005; 12:537–541. doi: 10.1107/s0909049505012719.

22. Newville M. J Synchrotron Radiat. 2001; 8:96–100. [PubMed: 11512993]

23. Zabinsky SI, Rehr JJ, Ankudinov A, Albers RC, Eller MJ. Phys Rev B. 1995; 52:2995–3009.

24. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Annu Rev Biophys Biomol Struct. 2000; 29:291–325. [PubMed: 10940251]

25. Harding MM. Acta Crystallogr. 1999; 55:1432–1443.

26. Harding MM. Acta Crystallogr. 2006; 62:678–682. doi: 10.1107/s0907444906014594.

27. Shen MY, Sali A. Protein Sci. 2006; 15:2507–2524. doi:10.1110/ps.062416606. [PubMed: 17075131]

28. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. J Appl Crystallogr. 1993; 26:283–291.

29. Neese, F. ORCAv2.8.0, an ab initio, density functional and semiempirical program package. Institute for Physical and Theoretical Chemistry, Universitat Bonn; Bonn: 2010. http://www.thch.uni-bonn.de/tc/orca/

30. Jacquamet L, Aberdam D, Adrait A, Hazemann JL, Latour JM, Michaud-Soret I. Biochemistry. 1998; 37:2564–2571. [PubMed: 9485406]

31. Diakun GP, Fairall L, Klug A. Nature. 1986; 324:698–699. [PubMed: 3796733]

32. Strange RW, Blackburn NJ, Knowles PF, Hasnain SS. J Am Chem Soc. 1987; 109:7157–7162. doi: 10.1021/ja00257a042.

33. Engh RA, Huber R. Acta Crystallogr. 1991; 47:392–400.

34. Ferreira GC, Franco R, Mangravita A, George GN. Biochemistry. 2002; 41:4809–4818. doi: 10.1021/bi015814m. [PubMed: 11939775]

35. Blackburn NJ, Hasnain SS, Pettingill TM, Strange RW. J Biol Chem. 1991; 266:23120–23127. [PubMed: 1744110]

36. Soriano A, Colpas GJ, Hausinger RP. Biochemistry. 2000; 39:12435–12440. [PubMed: 11015224]

37. Voland P, Weeks DL, Marcus EA, Prinz C, Sachs G, Scott D. Am J Physiol Gastrointest Liver Physiol. 2003; 284:G96–G106. doi: 10.1152/ajpgi.00160.2002.

38. Levina A, Armstrong RS, Lay PA. Coord Chem Rev. 2005; 249:141–160. doi: 10.1016/j.ccr.2004.10.008.

39. Robert, AS. Methods in Enzymology. Hirs, CHW.; Timasheff, SN., editors. Academic Press; New York: 1985. p. 414-459.

40. Teo, BK. EXAFS: basic principles and data analysis. Springer; Berlin: 1986.

41. Bunker, G. Introduction to XAFS: a practical guide to X-ray absorption fine structure spectroscopy. Cambridge University Press; London: 2010.

42. Clark-Baldwin K, Tierney DL, Govindaswamy N, Gruff ES, Kim C, Berg J, Koch SA, Penner-Hahn JE. J Am Chem Soc. 1998; 120:8401–8409. doi: 10.1021/ja980580o.

Fig. 1.
Helicobacter pylori UreG (*Hp*UreG) zinc K-edge X-ray absorption near edge structure spectra [0.2 mM holo-*Hp*UreG–Zn dimer in 20 mM Tris–HCl buffer, pH 7.0, with 150 mM NaBr]
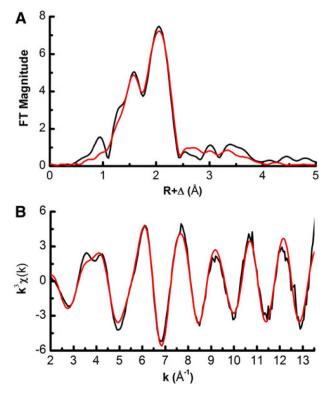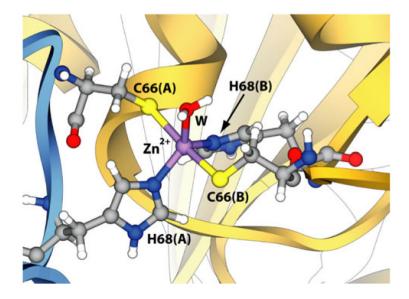
Fig. 2.
*Hp*UreG zinc K-edge extended X-ray absorption fine structure (EXAFS) spectra (0.2 mM holo-*Hp*UreG–Zn dimer in 20 mM Tris–HCl buffer, pH 7.0, with 150 mM NaBr). **a** Fourier-transformed EXAFS spectra [no phase correction, Fourier transform (*FT*) window 2-13.5Å^{-1}]: *black line* data, *red line* best fit. **b** k^3-weighted unfiltered EXAFS spectra: *black line* data, *red line* best fit. Fits were performed in *r*-space ($\Delta k = 2–13.5$Å^{-1}; $\Delta r = 1–4.5$Å) with a Kaiser–Bessel window (dk = 2), k^3-weighted data, and $S_0 = 0.9$

Fig. 3.
MODELLER version 9.8 model of *Hp*UreG dimer Zn(II) binding site based on ab initio geometric parameters. A trigonal bipyramidal coordination sphere for zinc (*purple sphere*) is formed by pairs of Cys-66 and His-68 residues from each of the monomers, leaving a solvent-accessible coordination site, in this case occupied by water (W). The protein chains are reported as ribbons (*light blue* and *orange* for chain A and chain B, respectively). Residues involved in zinc binding are reported as balls and sticks and atoms are colored according to atom type: *blue* nitrogen, *red* oxygen, *yellow* sulfur, *gray* carbon, *white* hydrogen
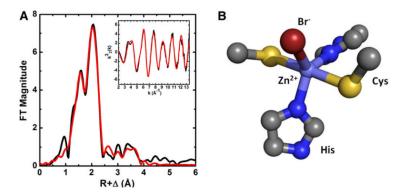
Fig. 4.
*Hp*UreG zinc *K*-edge EXAFS data (*black line*) and refined fit (*red line*) using the molecular model that accounts for coordination geometry as input. **a** Fourier-transformed EXAFS spectra (no phase correction, FT window 2–13.5Å^{-1}). *Inset:* k^3-weighted unfiltered EXAFS spectra and best fit. Fits were performed in *r*-space ($\Delta k = 2$–13.5Å^{-1}; $\Delta r = 1$–4.5Å) with a Kaiser–Bessel window (dk = 2), k^3-weighted data, and $S_0 = 0.9$. **b** Structure of the homology-modeled zinc site that was used as input for EXAFS refinement. Average metal ligand bond distances were 2.25Å for Zn–S$_{Cys}$, 1.95Å for Zn–N$_{His}$, and 2.40Å for Zn–Br$^-$

Table 1

Selected single-scattering extended X-ray absorption fine structure (EXAFS) fits showing zinc–ligand bond distances and coordination number for *Helicobacter pylori* UreG (*Hp*UreG)

Model	N	r(Å)	σ² (×10⁻³ Å²)	R (%)	χ_ν^2
N₂S₁Br₁	2 N	1.98(2)	4.5(2)	1.18	33.60
	1 S	2.24(7)	6(3)		
	1 Br	2.39(1)	3.0(8)		
N₂S₂Br₁	2 N	2.00(2)	6.1(3)	1.08	30.87
	2 S	2.21(8)	15(5)		
	1 Br	2.38(1)	2.7(4)		
N₂S₃Br₁	2 N	2.00(2)	5.4(2)	1.12	31.89
	3 S	2.19(8)	20(6)		
	1 Br	2.38(1)	2.6(4)		

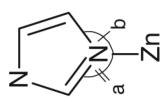Fits were performed in *r*-space ($\Delta k = 2$–13.5 Å^{-1}; $\Delta r = 1$–2.5 Å)

Table 2

Selected multiple-scattering EXAFS fits showing zinc–ligand bond distances and coordination number for *Hp*UreG

Model[a]	N	r(Å)	σ^2 (×10⁻³ Å²)	R (%)	$\chi_\nu{}^2$
N₂^(His)S₂Br₁ α = 0	2 N^(His)	1.99(1)	5.3(8)	2.01	12.24
	2 S	2.29(2)	13(2)		
	1 Br	2.40(1)	3.0(3)		
N₂^(His)S₂Br₁ α = 5	2 N^(His)	1.99(1)	5.0(9)	1.90	11.55
	2 S	2.29(2)	13(2)		
	1 Br	2.40(1)	3.0(3)		
N₂^(His)S₃Br₁ α = 5	2 N^(His)	2.0(1)	2.0(1)	1.96	11.96
	3 S	2.28(2)	19(19)		
	1 Br	2.40(1)	2.8(2)		

Fits were performed in r-space ($\Delta k = 2$–13.5 Å^{-1}; $\Delta r = 1$–4.5 Å)

[a] α is the imidazole angle of rotation, around an axis perpendicular to the plane of the ring and going through the coordinating nitrogen, where α = $(b - a)/2$, with a and b as the zinc–imidazole bond angles shown below:

Table 3

Refined multiple-scattering EXAFS fit of *Hp*UreG, using the computational model (Fig. 4) as initial input

Model	N		r(Å)	σ^2 ($\times 10^{-3}$ Å^2)	R (%)	χ_ν^2
N$_2^{(His)}$S$_2$Br$_1$	2 N$^{(His)}$	1.98(1)	5.2(9)	1.52	9.25
	2 S	2.25(2)	12(1)		
	1 Br	2.40(1)	2.7(2)		

Fits were performed in *r*-space ($\Delta k = 2$–13.5 Å^{-1}; $\Delta r = 1$–4.5 Å)