Biochemical Studies on *Mycobacterium tuberculosis* UreG and Comparative Modeling Reveal Structural and Functional Conservation among the Bacterial UreG Family†

Barbara Zambelli,‡ Francesco Musiani,‡ Matteo Savini,‡ Paul Tucker,§ and Stefano Ciurli*,‡,||

Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Viale Giuseppe Fanin 40, 40127 Bologna, Italy, EMBL (European Molecular Biology Laboratory) Hamburg c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany, and CERM (Center for Magnetic Resonance), Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Firenze, Italy

Received November 30, 2006; Revised Manuscript Received January 4, 2007

ABSTRACT: Nickel is a fundamental micronutrient for cellular life, but it is toxic in soluble form at nonphysiological concentrations. Such potentially contradictory features required living organisms to develop efficient systems for nickel utilization and homeostasis. This is the case for incorporation of nickel into the active site of urease, a multistep, tightly regulated process, requiring the interplay of various accessory proteins. The understanding of this activation mechanism may find medical applications against ureolytic bacteria, among which *Mycobacterium tuberculosis* is a deadly pathogen for humans. The topic of this study is UreG, an essential chaperone in the in vivo activation of urease upon insertion of Ni^{2+} into the active site. The protein was examined using both experimental and computational approaches. In particular, the soluble *M. tuberculosis* UreG (*Mt*UreG) was overexpressed in *Escherichia coli* and purified to homogeneity. The identity of the isolated protein was established by mass spectrometry. On-line size-exclusion chromatography and light scattering indicated that *Mt*UreG exists as a dimeric form in solution. Determination of the free thiol concentration revealed that a disulfide bond is present in the dimer. The isolated *Mt*UreG shows low GTPase activity under native conditions, with a k_{cat} of 0.01 min^{-1}. Circular dichroism spectroscopy demonstrated the presence of a well-defined secondary structure (8% α-helices, 29% β-strands) in *Mt*UreG, whereas NMR spectroscopy indicated that this protein does not behave as a rigid three-dimensional fold and thus can be assigned to the class of intrinsically unstructured polypeptides. The molecular model of *Mt*UreG in the fully folded and functional form was built using fold recognition algorithms. An extensive similarity search was performed to determine conservation patterns in all known bacterial UreG sequences. The generation of a multiple-sequence alignment and the related phylogenetic tree allowed us to recognize key residues and motifs that are likely important for protein function. A structural database containing the homology-built models of the most representative UreG proteins was created, confirming the structural analogies among the UreG family. A flexible region, likely to be important for protein function, is identified. The structural conservation among this class of GTPases is discussed on the basis of their function in the urease assembly process.

Transition metals are fundamental micronutrients for cellular life, being cofactors of several biological macromolecules and playing crucial functions for the catalysis of many biological reactions. Accordingly, cells concentrate metal ions, even several orders of magnitude relative to their abundance in a typical growth medium (*1*), through the activity of specific membrane transporters. On the other hand, the toxicity of free transition metals for cellular life prompted living organisms to develop sophisticated mechanisms for their detoxification, and efficient intracellular carriers for delivering them into the correct subcellular location (*2*−*4*).

Nickel possesses this dual nature. Its toxicity and carcinogenesis are well-documented for animals and humans (*5*), while conversely, the importance of nickel in many fundamental biological processes is well-known (*6*). This metal is an essential cofactor in the active site of important

† This work was supported by Grants PRIN-2003 and PRIN-2005 from the Ministero Italiano dell'Università e della Ricerca (MIUR). B.Z. was awarded an EMBO Short Term Fellowship in 2005 to visit the EMBL of Hamburg and is now the recipient of a grant from the University of Bologna and Magnetic Resonance Center (CERM) of Firenze. F.M. is a recipient of a fellowship from Consorzio Interuniversitario di Risonanze Magnetiche di Metalloproteine Paramagnetiche (CIRMMP). P.T. acknowledges infrastructural support from German Ministry of Science Education Consortium Grant BIO/0312992 and from the NIH *Mycobacterium tuberculosis* Structural Genomics Consortium. We thank Arie Geerlof (EMBL Hamburg, Hamburg, Germany) for useful advice in the use of auto-induction medium for protein expression.

* To whom correspondence should be addressed: Laboratory of Bioinorganic Chemistry, Department of Agro-Environmental Science and Technology, University of Bologna, Viale Giuseppe Fanin 40, 40127 Bologna, Italy. Phone: +39-051-209-6204. Fax: +39-051-209-6203. E-mail: stefano.ciurli@unibo.it.

‡ University of Bologna.

§ EMBL (European Molecular Biology Laboratory) Hamburg c/o DESY.

|| CERM (Center for Magnetic Resonance).

10.1021/bi6024676 CCC: $37.00
Published on Web 02/20/2007

enzymes, such as CO-dehydrogenase and acetyl-CoA synthase from photosynthetic and methanogenic bacteria, methyl-CoM reductase from methanogenic bacteria, bacterial [NiFe]-hydrogenase and acireductone dioxygenase, superoxide dismutase from actinomyces, nickel-dependent glyoxylase from *Escherichia coli*, and urease from plants and bacteria (*6*). Nickel is also a fundamental element for several medically relevant microorganisms, being important for their growth and pathogenesis. Therefore, a strategy against nickel-dependent pathogens could pass through the preclusion of regular intracellular nickel trafficking. In this regard, the understanding of nickel handling and detoxification mechanisms is a prerequisite for pinpointing specific antibacterial targets.

Among nickel-dependent microorganisms, *Mycobacterium tuberculosis*, the etiologic agent of tuberculosis disease, is one of the most representative and important. In humans, this intracellular bacterium infects macrophages, living inside their phagosomes. In this environment, its survival depends on the activity of nickel-dependent urease. In particular, urea hydrolysis is essential for bacterial survival, since it contributes to nitrogen availability and environmental pH modulation (*7*). Moreover, ammonia derived from this reaction can block the phagosome−lysosome fusion, being an important defensive mechanism against the immune system of the host (*8*). Therefore, urease, one of the best-characterized nickel-dependent enzymes (*9*, *10*), can be considered an ideal target for the development of specific strategies against ureolytic pathogens, such as *M. tuberculosis*. The atomic structure of the hetero-multimeric urease enzyme has been elucidated for *Bacillus pasteurii* (*Bp*)[1] (*11*), *Klebsiella aerogenes* (*Ka*) (*12*), and *Helicobacter pylori* (*Hp*) (*13*). In addition to the structural subunits that compose the apoenzyme, the urease system includes four conserved helper proteins, named UreD, UreE, UreF, and UreG, which ensure the correct sequence of events for the assembly of the catalytic site (*14*). Among them, UreE and UreG are the best-characterized at the structural and molecular level. UreE is responsible for delivery of nickel into the apourease precursor through a metal site able to bind two nickel ions (*15*), and the crystallographic structures of *Bp*UreE and *Ka*UreE have been determined (*16*, *17*). In many organisms, this protein likely plays a further role in nickel storage, binding several nickel ions in a histidine-rich C-terminal sequence (*18*). This histidine-rich trait is also found in other chaperones of nickel-dependent enzymatic systems, for instance, in the N-terminal part of some HypB GTPases (*19*) and in the C-termini of the nickel transporter CooJ (*20*), involved in the activation of nickel-dependent hydrogenase and carbon monoxide dehydrogenase, respectively. Eu3, the orthologue of UreG in the urease system of soybean, also has an N-terminal histidine-rich tag (*21*).

UreG was identified as the essential chaperone responsible for the hydrolysis of GTP in the process of urease active site assembly (*22*). A well-characterized ortholog is UreG from *B. pasteurii* (*Bp*UreG), a dimeric protein in solution that exhibits both stoichiometric zinc binding and low GTPase activities (*23*). NMR spectroscopy revealed that it does not possess a rigid tertiary structure and exists in solution in fast equilibrium among different conformations, with large portions of unfolded backbone (*23*). However, *Bp*UreG contains a certain degree of secondary and tertiary structure, as revealed by CD and fluorescence spectroscopy, and can be classified as an intrinsically disordered protein (*24*). The structural model of *Bp*UreG, developed using threading algorithms, indicated that, in the fully folded state, this protein probably assumes an overall fold typical of a GTPase (*23*).

This study extends the previously available information about the essential chaperone UreG using both experimental and computational approaches. The hydrodynamic, functional, and structural properties of UreG from *M. tuberculosis* (*Mt*UreG) are investigated and compared to the previously available information for other UreG proteins. The molecular model of fully folded *Mt*UreG is reported, and the structure of the protein is discussed on the basis of its function as an intrinsically disordered GTPase involved in intracellular nickel trafficking and urease activation. Moreover, the emerging structural homology between different UreG proteins is investigated by building a multiple-sequence alignment, a phylogenetic tree, and a structural database containing the most significant UreG proteins, retrieved in an extensive similarity search. The distinctive structural tracts of this protein family are identified and discussed relative to protein functionality and to the available experimental data.

MATERIALS AND METHODS

Expression and Purification of MtUreG. The *pET*22b-*MtureG* construct was kindly provided by J. Perry (UCLA-DOE, Los Angeles, CA), and it was used to transform the *E. coli* BL21(DE3) expression host (Novagen). On the basis of the T7 expression system (*25*), large-scale expression of His-tagged *Mt*UreG was achieved in 2 L batches of LB or auto-induction (*26*) medium. The ^{15}N-enriched protein was obtained using M9 medium [6 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1.25 g/L $(NH_4)_2SO_4$ (pH 7.5), 1 mM $MgSO_4$, and 4 g/L glucose], containing $(^{15}NH_4)_2SO_4$. In LB and M9 media, cells were grown at 37 °C. When the OD_{600} reached 0.5−0.6, expression was induced by addition of IPTG (isopropyl β-thiogalactopyranoside) to a final concentration of 0.5 mM and the temperature was decreased to 28 °C after induction. The cells were harvested 4 h after induction by centrifugation at 8000*g* for 15 min, at 4 °C. In auto-induction medium, cells were grown at 28 °C for 48 h. The cellular pellet was resuspended in 20 mL of 50 mM Tris-HCl (pH 8), and cells were disrupted by two passages through a French pressure cell (SLM-Aminco) operating at 20 000 psi. The supernatant after pellet separation was loaded onto a 5 mL Hi-Trap chelating column (Amersham Pharmacia Biotech), pre-equilibrated with 40 mL of 50 mM Tris-HCl (pH 8.0). The column was subsequently washed with 50 mL of 20 mM Tris-HCl (pH 8.0) containing 300 mM NaCl and with 50 mL of the same buffer containing 20 mM imidazole, until the baseline was stable. The protein was

[1] Abbreviations: GTPase, guanosine 5′-triphosphate hydrolase; *Bp*, *Bacillus pasteurii*; *Ka*, *Klebsiella aerogenes*; *Hp*, *Helicobacter pylori*; *Mt*, *Mycobacterium tuberculosis*; LB, Luria-Bertani; OD_{600}, optical density at 600 nm; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; EMBL, European Molecular Biology Laboratory; MALS, multiple-angle light scattering; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); DTT, dithiothreitol or *threo*-1,4-dimercaptobutane-2,3-diol; CD, circular dichroism; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; IPTG, isopropyl β-D-1-thiogalactopyranoside; ESI-Q-TOF, electrospray ionization quadrupole time-of-flight; SEC, size-exclusion chromatography.

eluted with a 50 mL linear imidazole gradient from 0.02 to 0.5 M. The fractions containing *Mt*UreG were combined, concentrated using 5 kDa Centricon ultrafiltration units (Millipore) to a final concentration of 5 mg/mL, and centrifuged (15 min at 14000*g*) to remove the precipitated material. The solution was loaded onto a Superdex 75 XK 16/60 column conditioned with 50 mM Tris-HCl buffer (pH 8) containing 0.15 M NaCl. *Mt*UreG was eluted at a flow rate of 1 mL/min. The purified protein was concentrated and stored at −80 °C.

Protein purity, as well as the apparent molecular mass of *Mt*UreG under denaturing conditions, was estimated by SDS−PAGE according to the method of Laemmli (*27*), by using a BioRad Mini-Protean II apparatus. Proteins were separated on 15% (w/v) acrylamide−bisacrylamide separating gels, stained using Coomassie Brilliant Blue R-250.

Protein concentrations were measured using a JASCO 7800 spectrophotometer and a value for the extinction coefficient ($\epsilon_{280} = 8370\ M^{-1}\ cm^{-1}$ for *Mt*UreG) calculated from the amino acid sequence using the ProtParam website (http://au.expasy.org/tools/protparam.html). This value is in good agreement with that obtained by using the Bio-Rad assay that is based on the Bradford colorimetric method (*28*).

The purified protein, isolated from a denaturing SDS−PAGE gel, was identified by tryptic digestion and ESI Q-TOF mass spectrometry, performed at the EMBL Proteomic Core Facility (Heidelberg, Germany).

Hydrodynamic Properties of MtUreG. An estimate of the absolute molecular mass of *Mt*UreG under native conditions was determined using a combination of size-exclusion chromatography and MALS (multiple-angle light scattering). *Mt*UreG (100 μL, 2.5 mg/mL) in 20 mM Tris-HCl (pH 8.0) containing 150 mM NaCl was loaded onto a Superdex-75 HR10/30 column (Amersham), pre-equilibrated with the same buffer, and eluted at room temperature at a flow rate of 0.5 mL/min. The column was connected downstream to a multiple-angle laser light (690.0 nm) scattering DAWN EOS photometer (Wyatt Technology). Values of 0.185 for the refractive index increment (d*n*/d*c*) and 1.330 for the solvent refractive index were used. Molecular weights were determined from a Zimm plot. Data were analyzed using Astra version 5.1.7 (Wyatt Technology), following the manufacturer's instructions.

Determination of the Concentration of Free Cysteine Thiols in MtUreG. A solution of *Mt*UreG [0.96 mL in 20 mM Tris-HCl (pH 8.0), 15 μM] was incubated with 40 μL of 10 mM DTNB at room temperature for 15 min. The absorbance at 412 nm was measured and was stable after this incubation time. The concentration of free cysteine thiols in *Mt*UreG was calculated using a standardization curve, obtained with standard cysteine concentrations ranging from 0 to 50 μM, under identical buffer conditions.

Measurement of the GTPase Activity of MtUreG. GTP hydrolyzing activity was measured using a colorimetric method. The reaction mixture, containing 20 mM Tris-HCl (pH 8.0), 0.15 M NaCl, 5 mM $MgCl_2$, 2 mM GTP, and 5 μM *Mt*UreG, in the presence and absence of 2 mM DTT, was incubated at 37 °C. Aliquots (50 μL) were removed at different incubation times and added to 50 μL of a 17.5% trichloroacetic acid/water solution. The phosphate concentration was determined by the malachite green assay (*29*).

Circular Dichroism Spectroscopy. The CD spectrum of *Mt*UreG in 10 mM phosphate buffer (pH 7.5) was recorded using a Jasco J-810 spectropolarimeter, flushed with N_2, and a cuvette with a path length of 0.1 cm. The spectra were registered from 185 to 250 nm at 0.2 nm intervals. Ten spectra were accumulated at room temperature and averaged to achieve an appropriate signal-to-noise ratio. The spectrum of the buffer was subtracted. The secondary structure composition of *Mt*UreG was evaluated with the tool available on the Dichroweb server of the Centre for Protein and Membrane Structure and Dynamics [http://www.cryst.bbk.ac.uk/cdweb/html/home.html (*30*)] using reference sets 3, 4, 6, and 7.

$^{15}N-^{1}H$ HSQC NMR Spectroscopy. NMR spectra of 0.5 mM ^{15}N-enriched *Mt*UreG were recorded at pH 8.0 and 298 K on a Varian Inova 600 MHz spectrometer. The $^{1}H-^{15}N$ HSQC spectrum consisted of 64 scans, spectral windows of 20 ppm in the proton dimension and 80 ppm in the nitrogen dimension, with the carrier set at the water frequency and 118 ppm, respectively. Matrices of 1024 × 512 points were acquired and transformed into 2048 × 1024 points and processed using nmrPipe (*31*).

Sequence Search and Alignment. Sequences of UreG were searched using sequence similarity criteria and the primary structure of *Bp*UreG as a template. FASTA3 (*32*, *33*) available at http://www.ebi.ac.uk/fasta33 and BLAST (basic local alignment search tool) (*34*, *35*) available at http://www.ncbi.nlm.nih.gov/BLAST were utilized for the search. UreG sequences were retrieved from a nonredundant sum of different databases (SwissProt, TrEMBL, TrEMBLNew, GenBank CDS, PDB, PIR, and PRF). Sequences not belonging to bacteria or not annotated as UreG's were discarded. Multiple-sequence alignments were performed on the remaining sequences using ClustalW (*36*), available at http://www.ebi.ac.uk/clustalw, and alignment optimization was carried out using information deriving from secondary structure predictions provided by JPRED (*37*), available at http://www.compbio.dundee.ac.uk/~www-jpred.

The obtained final alignment was employed to derive an unrooted phylogenetic tree as provided by PHYLIP (*38*), available at http://evolution.genetics.washington.edu/phylip.html. Such a tree was then divided into portions on the basis of taxonomic order of the sequences. To reduce the number of modeled structures, from every portion of the phylogenetic tree the most representative sequence was identified as the sequence having the higher degree of homology with all the others of the same group.

Structure Prediction and Homology Modeling. The selected sequences were used to search for template structures using the 3D-Jury predictor meta-server (*39*, *40*) available at http://bioinfo.pl/Meta/. Only the templates with a 3D-Jury score higher than 80% of the best score were selected. Multiple-sequence alignment of the sequence of the target sequence with the selected templates was performed using ClustalW, while manual optimization of the alignment was carried out using information deriving from the secondary structure prediction provided by JPRED.

Model structures were calculated using MODELLER 6.2 (*41*) with the model-default options and the iterative procedure developed for the modeling of the urease accessory protein UreE (*18*). PROSA II (version 3.0) (*42*) was used for selecting the best models provided by MODELLER and

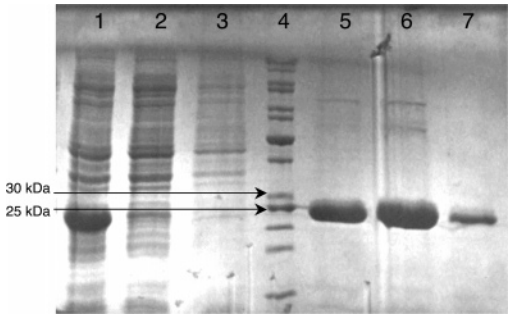


FIGURE 1: Expression and purification of *Mt*UreG. SDS−PAGE of cell extracts of BL21(DE3) *E. coli* cells harboring *pET22b-MtureG* after induction for 16 h with IPTG (lane 1) and prior to induction of expression (lane 2), the flow-through of the Ni affinity column (lane 3), molecular mass markers (lane 4), and fractions eluted with an imidazole gradient, as described in the text (lanes 5 and 6), purified *Mt*UreG (lane 7).

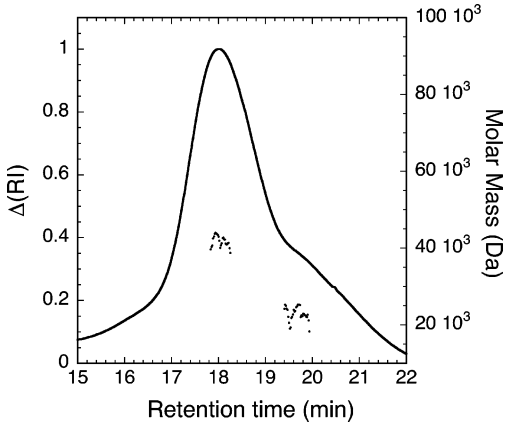


FIGURE 2: Molar mass distribution plot for *Mt*UreG as detected by size-exclusion chromatography and light scattering. The solid line indicates the trace from the refractive index detector, and the dots are the weight-average molecular weights for each slice (i.e., measured every second), measured by multiple-angle light scattering.

for protein structure analysis to test the coherency and validity of the model structures. The *Z* score reported in this work is derived through the standard "hide and seek" procedure of the program, by which the score is correlated to the difference in potential energy, calculated using mean field potentials, between the input structure and other randomly assigned folds for its amino acid sequence. A lower *Z* score corresponds to a more favorable potential energy associated with the structure that is being examined.

Structure validation was performed using PROCHECK (*43*) and WHATIF (*44*). The calculated final structures were deposited in the http://www.postgenomicnmr.net site. The molecular surface and the electrostatic color coding were generated with GRASP (*45*) using a probe radius of 1.4 Å. The electrostatic potential was calculated using a simple version of a Poisson−Boltzmann solver with the GRASP full charge set. All the histidine residues were considered neutral, and the N- and C-terminal residues were charged. Dielectric constants of 80 and 2 were used for the solvent and protein interior, respectively.

RESULTS

Expression and Purification of MtUreG. The *pET*22b-*MtureG* construct contained the gene *MtureG* cloned in frame with a C-terminal His tag between the *Nde*I and *Xho*I restriction sites of the *pET*22b vector. This plasmid, the sequence of which was confirmed by double-strand DNA sequencing in the cloned gene, was employed to overproduce the His-tagged *Mt*UreG (25.09 kDa) in the *E. coli* BL21-(DE3) strain by induction with IPTG. This protocol provided an abundant polypeptide with an apparent molecular mass in agreement with the theoretical molecular weight. This protein was absent from the non-induced cells, as demonstrated by SDS−PAGE (in Figure 1, compare lanes 1 and 2). Fractionation of cells into soluble and insoluble extracts showed that the overproduced protein completely accumulated in the soluble fraction. The protein was purified by affinity chromatography, which specifically retained the His-tagged protein, as indicated by SDS−PAGE (Figure 1, lane 3). The protein was eluted using an imidazole gradient and presented a very good level of purity (Figure 1, lanes 5 and 6), which was further improved with another chromatographic passage in size exclusion. The final protein yield (ca. 100 mg/L of initial culture) was achieved by using auto-induction medium (*26*), which increased the amount of purified protein by 1 order of magnitude, with respect to induction in LB medium (16 mg/L of initial culture). The final protein purity was checked by SDS−PAGE (Figure 1, lane 7).

The identity of the recombinant protein was demonstrated by tryptic digestion of the protein, recovered from a SDS−PAGE gel. The correspondence between the protein fragments and the expected fragments was confirmed by ESI Q-TOF tandem mass spectrometry. Three digested fragments were sequenced as TDLAALVGADLAVMAR (1585.8 Da), DGRPTVLQSLTEDPAASDVVAWVR (2581.3 Da), and SQLAADGV (759.4 Da), in complete agreement with the theoretical sequence of *Mt*UreG.

Hydrodynamic Properties and Oligomerization of MtUreG. A combination of size-exclusion chromatography (SEC) and multiple-angle light scattering (MALS) (*46*, *47*) (Figure 2) indicated that *Mt*UreG exists in solution as a dimer (estimated $M_w = 41.7 \pm 0.2$ kDa), as also found for *Bp*UreG (*23*). A very minor fraction is also present, constituted by the monomeric form (estimated $M_w = 23.1 \pm 0.2$ kDa).

Determination of the Concentration of Free Cysteine Thiols in MtUreG. Analysis of the *Mt*UreG sequence indicates that this protein contains two cysteine residues per protein monomer (Cys^{46} and Cys^{90} in *Mt*UreG numeration). While Cys^{90} is fully conserved among all UreG proteins (*23*) and is involved in a disulfide bridge in the *Bp*UreG dimer (*24*), Cys^{46} is not conserved and is often substituted with threonine, valine, or, occasionally, serine (see the alignment in Figure 6A). To monitor the nature of the interaction responsible for protein dimerization of *Mt*UreG, the oxidation state of the two cysteine residues was evaluated using Ellman's reagent under native conditions (*48*). For a solution containing 15 μM *Mt*UreG monomer, the concentration of free cysteine thiols is ca. 50% of the expected value if both cysteines were reduced, suggesting that each protein monomer contains an oxidized cysteine. This observation, coupled to the reported dimeric state of *Mt*UreG, denotes the presence of a disulfide bridge involved in protein dimerization.

Measurement of the GTPase Activity of MtUreG. The GTPase activity of UreG, initially postulated on the basis of its primary structure, was recently demonstrated for *Bp*UreG

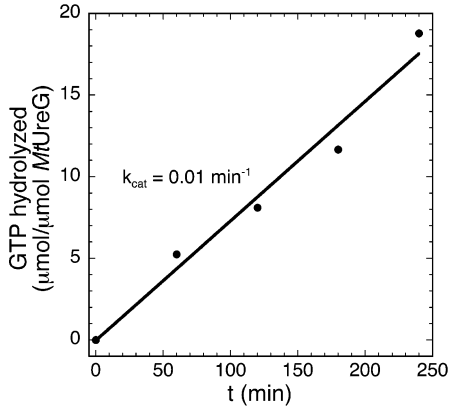


FIGURE 3: Enzymatic activity of *Mt*UreG. Time course of GTPase activity of *Mt*UreG (●), fitted using linear regression (—). The calculated k_{cat} is given.

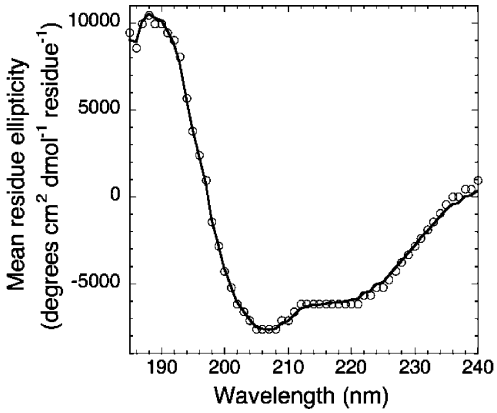


FIGURE 4: Far-UV circular dichroism spectrum of *Mt*UreG. The experimental points are shown as empty circles, and the solid line represents the best fit calculated using the Dichroweb server, which was used to calculate the secondary structure, as described in the text.

(*23*). To test the functionality of the purified *Mt*UreG, the GTPase capability of this protein was measured using the colorimetric method previously employed for *Bp*UreG.

In the assay mixture, the concentration of substrate GTP is 2 orders of magnitude larger than the concentration of enzyme, a condition that allows the use of the time course data of the reaction for the derivation of the value of k_{cat} of 0.01 min^{-1}. The GTPase activity was not influenced by the addition of 2 mM DTT to the reaction mixture.

Circular Dichroism Spectroscopy on MtUreG. The secondary structure composition of *Mt*UreG in solution was evaluated using CD spectroscopy. The far-UV CD spectrum (Figure 4) features negative ellipticity at ca. 206 and 220 nm and a pronounced shoulder at ca. 193 nm. This shape indicates the presence of both α-helices and β-strands.

The data were quantitatively analyzed in the range of 185−240 nm, on a per peptide basis calculated from protein content and sequence, using all the different fitting programs available at the Dichroweb server (*30*) and all the possible reference sets. The best fit was selected on the basis of the normalized root-mean-square deviation (nrmsd = 0.036) between the experimental and calculated data, obtained using the variable selection method program CDSSTR (*49*) and reference set 6. From this analysis, a secondary structure composition of 8% α-helix, 29% β-strand, 19% turns, and 45% random coil was estimated for *Mt*UreG.

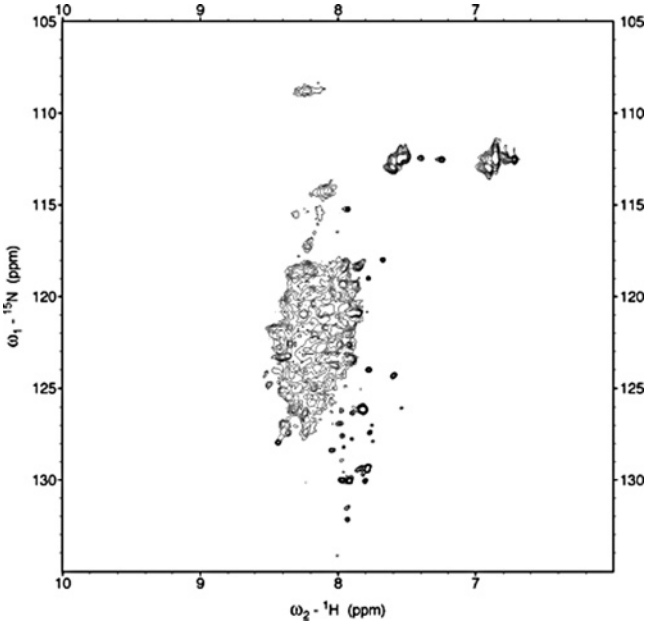
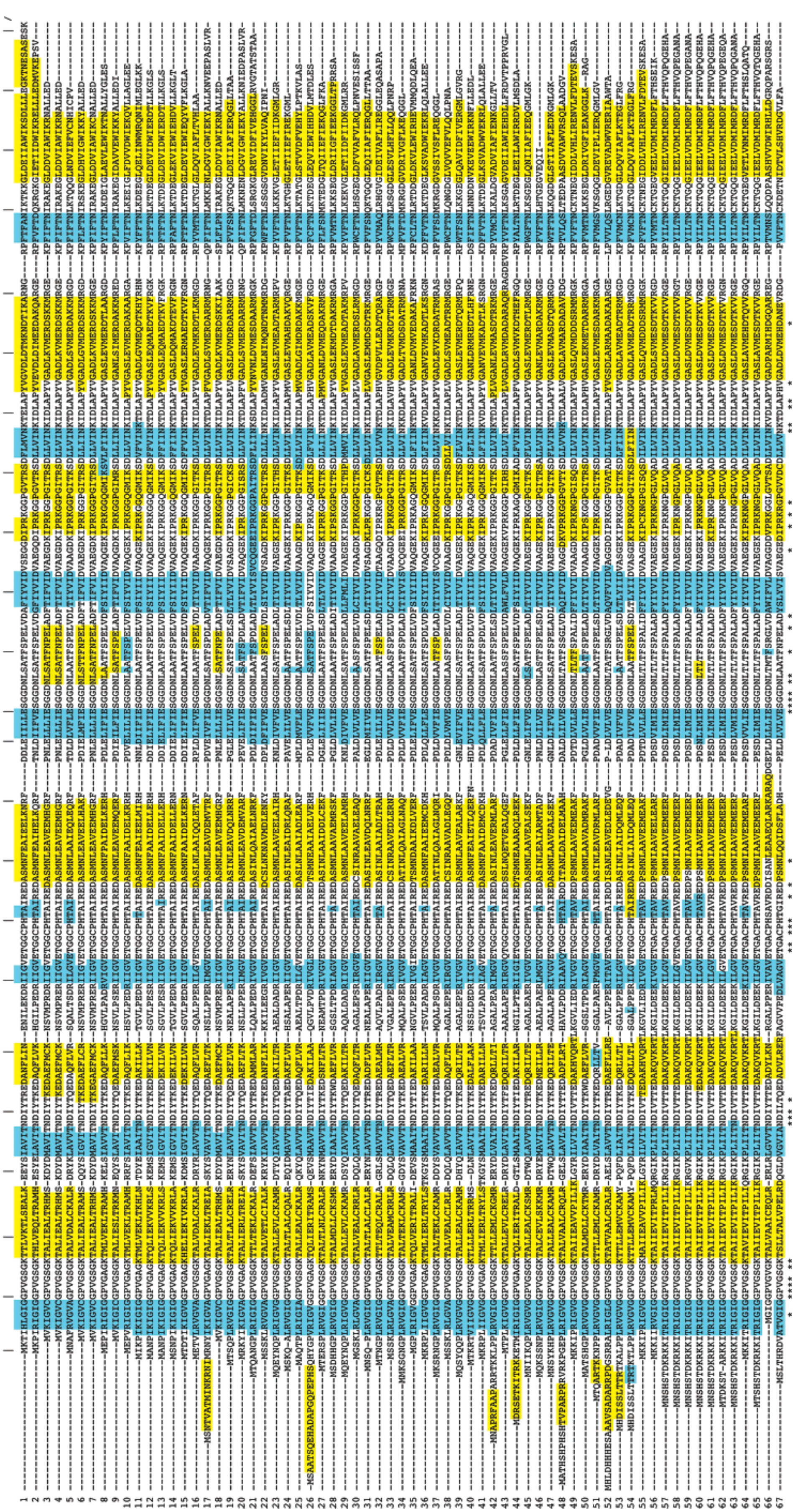


FIGURE 5: NMR spectroscopic properties of *Mt*UreG. ^{1}H−^{15}N HSQC spectrum of *Mt*UreG. The experimental conditions are described in the text.

^{1}H−^{15}N HSQC NMR Spectroscopy. To monitor the three-dimensional structural properties of *Mt*UreG, NMR spectroscopy was used. As in the case of *Bp*UreG and *Ka*UreG, the NMR spectrum of *Mt*UreG shows poorly resolved resonances with little dispersion in the ^{1}H dimension of the backbone amides, indicating that large portions of the protein backbone are very flexible in solution and lack a well-defined tertiary structure (Figure 5). The signals of the side chain NH_2 groups of the four asparagine residues present in the protein are resolved and give rise to two broad envelopes, centered at 112.5 ppm (^{15}N) and 7.6−6.8 ppm (^{1}H), confirming the fluxional behavior for *Mt*UreG in solution. No effect on the NMR spectrum was observed in the presence of 1 equiv of GTPγS in solution.

UreG Multiple-Sequence Alignment Construction and Taxonomic Analysis. A search for UreG sequences using *Bp*UreG as seed resulted in 67 hits. Their multiple-sequence alignment was optimized by hand on the basis of the secondary structure predictions provided by JPRED (Figure 6A). In the final alignment, the sequences feature a level of identity between 69.2 and 47.8% with respect to the sequence of *Bp*UreG and contain 39 fully conserved residues. Some conserved characteristics that confirm that UreG can be classified as a small G-protein are identifiable (*50*): (i) the P-loop (GXXXXGK[S/T]) sequence in the N-terminal region, (ii) a threonine residue putatively involved in Mg^{2+} binding (position 64 of the *Mt*UreG consensus sequence), corresponding to the so-called "Walker A box" and required for GTP hydrolysis, (iii) the sequence [D/E/Q]XXG (residues 86−89 in *Mt*UreG numeration) corresponding to the so-called "Walker B box", and (iv) the sequence NKX[D/E] (residues 169−172 in *Mt*UreG numeration) involved in the binding of the GTP guanidine ring. In addition, the sequences belonging to *M. tuberculosis* (sequence number 48) and *Streptomyces coelicolor* (sequence number 52) present a characteristic N-terminal His-rich tag, with the presence of four histidines in the first 10 N-terminal residues.



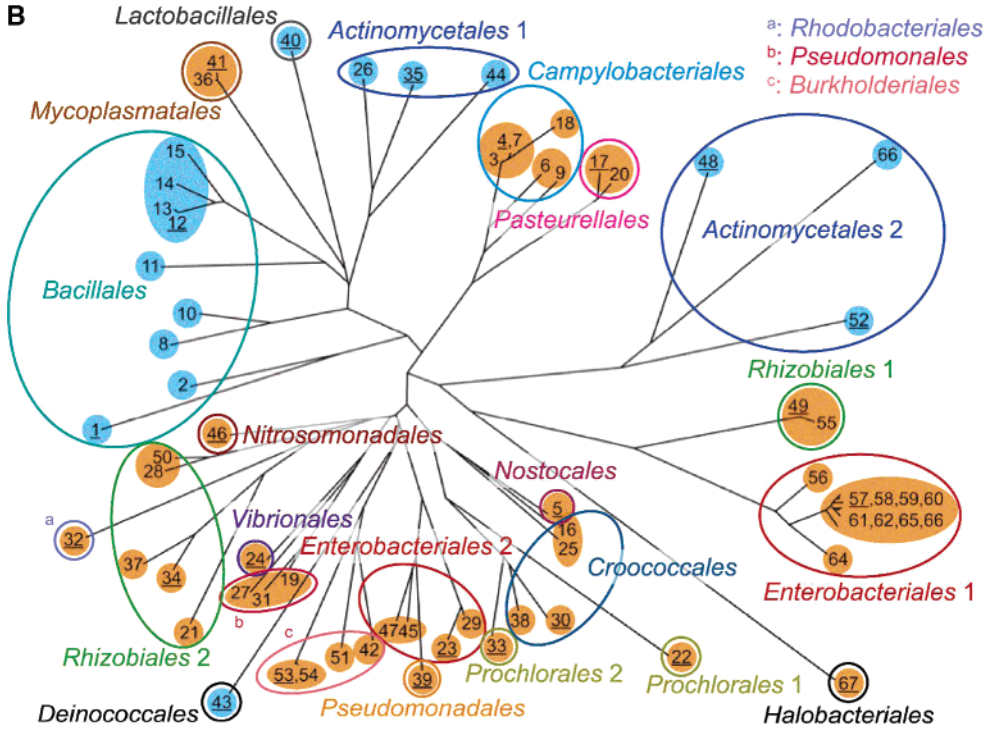


FIGURE 6: (A) Multiple-sequence alignment of UreG proteins. The alignment was obtained as described in the text. The predicted secondary structural elements are highlighted in yellow (helix) and cyan (strand). Fully conserved residues are denoted with an asterisk: (1) *B. pasteurii*, (2) *Bacillus halodurans* C-125, (3) *H. pylori* J99, (4) *H. pylori* 26695, (5) *Nostoc* sp. PCC 7120, (6) *Helicobacter bizzozeronii*, (7) *H. pylori* HPK5, (8) *Bacillus* sp. (strain TB-90), (9) *Helicobacter hepaticus* ATCC 51449, (10) *Geobacillus kaustophilus*, (11) *Bacillus cereus* ATCC 10987, (12) *Staphylococcus aureus* subsp. *aureus* MRSA252, (13) *S. aureus* subsp. *aureus* MSSA476, (14) *Staphylococcus epidermidis* ATCC 12228, (15) *Staphylococcus xylosus*, (16) *Thermosynechococcus elongatus* BP-1, (17) *Haemophilus influenzae* Rd KW20, (18) *H. pylori* 85P, (19) *Pseudomonas aeruginosa* PAO1, (20) *Actinobacillus pleuropneumoniae* CM5, (21) *Mesorhizobium loti* MAFF303099, (22) *Prochlorococcus marinus* subsp. *pastoris* strain CCMP1986, (23) *E. coli* 1440, (24) *V. parahaemolyticus* TH3996, (25) *Synechocystis* sp. PCC 6803, (26) *Actinomyces naeslundii* WVU45, (27) *Acinetobacter* sp. ADP1, (28) *Rhodopseudomonas palustris* CGA009, (29) *Proteus mirabilis*, (30) *Synechoccus* sp. WH 7805, (31) *Pseudomonas syringae* pv. tomato strain DC3000, (32) *Rhodobacter capsulatus*, (33) *Prochlorococcus marinus* strain MIT 9313, (34) *Agrobacterium tumefaciens* strain C58, (35) *Corynebacterium glutamicum* ATCC 13032, (36) *Ureaplasma parvum serovar* 3 strain ATCC 700970, (37) *Rhizobium leguminosarum* bv. *viciae*, (38) *Synechococcus* sp. WH 8102, (39) *Pseudomonas putida* KT2440, (40) *Streptococcus salivarius*, (41) *Ureaplasma urealyticum*, (42) *Burkholderia mallei* ATCC 23344, (43) *Deinococcus radiodurans* R1, (44) *Corynebacterium efficiens* YS-314, (45) *E. coli* O157:H7 EDL933, (46) *Nitrosospira* sp. NpAV, (47) *K. aerogenes*, (48) *M. tuberculosis*, (49) *Brucella melitensis* 16M, (50) *Bradyrhizobium japonicum* USDA 110, (51) *Ralstonia eutropha*, (52) *Streptomyces coelicolor* A3(2), (53) *Bordetella pertussis* Tohama I, (54) *Bordetella bronchiseptica* RB50, (55) *Brucella suis* 1330, (56) *Edwardsiella ictaluri*, (57) *Yersinia kristensenii*, (58) *Yersinia intermedia*, (59) *Yersinia mollaretii*, (60) *Yersinia enterocolitica*, (61) *Yersinia frederiksenii*, (62) *Yersinia pestis*, (63) *Yersinia aldovae*, (64) *Photorhabdus luminescens* ssp. *laumondii* TTO1, (65) *Yersinia rohdei*, (66) *Saccharopolyspora spinosa*, and (67) *Haloarcula marismortui*. (B) Unrooted phylogenetic tree of the UreG sequences. The division on the basis of taxonomic classification is reported. Gram-positive and Gram-negative bacteria are colored light blue and light orange, respectively. The sequence numeration is that indicated for panel A.

The unrooted phylogenetic tree, generated on the basis of the final multiple-sequence alignment, is shown in Figure 6B.

In general, the correlation between sequence homology and taxonomic distance is maintained (i.e., sequences belonging to the same taxonomic order are localized in the same region of the phylogenetic tree). Exceptions are some sequences belonging to the orders Actinomycetales, Enterobacteriales, Prochlorales, and Rhizobiales, which are localized in a different region of the phylogenetic tree respect to other sequences of the same order. In such cases, the classification of the order has been annotated sequentially (e.g., *Actinomycetales 1* and *Actinomycetales 2*). The Gram-positive and Gram-negative classification is also reported in Figure 6B.

Model Structure Calculation and Analysis. For every order identified in the phylogenetic tree, the most representative sequence was chosen for modeling (underlined sequences in Figure 6B). In the *Actinomycetales 2* group, the sequence belonging to *M. tuberculosis* was also modeled. To calculate model structures of the selected proteins, the meta-server 3D-Jury was employed for a template structure search. For each UreG sequence, the same procedure previously applied to choose the protein templates for *Bp*UreG modeling was used (*23*). The results of this search are summarized in Table SI-1. UreG's model structures were calculated following a consolidated procedure, developed to model the urease accessory proteins *Bp*UreE, *Bp*UreG, and *Bp*UreF (*18*, *23*, *51*). The sequence alignments of the found template proteins with the sequence of the corresponding UreG's, optimized by hand using information derived from secondary structure predictions and energy profile minimization of PROSA, are reported in Figure SI-1. Table 1 reports quality parameters for the modeled UreG structures.

All models present a good percentage of residues in the "most favored" and "additionally allowed" regions of the Ramachandran plot (at least 81.4 and 8.2%, respectively, with average values of 85.6 and 12.2%, respectively), while the few residues found in the "disallowed" region are localized in loops where the modeling process is knowingly more difficult. PROSA *Z* scores obtained for all UreG model

Table 1: Quality Parameters for the Calculated UreG Model Structures

order	sequence number	Ramachandran plot (%)				PROSA *Z* score
		favorite	allowed	generously allowed	disallowed	
Actinomycetales 1	35	87.7	9.0	3.2	0.0	−6.47
Actinomycetales 2	52	88.1	9.8	2.1	0.0	−6.54
Bacillales	12	83.3	14.6	1.4	0.7	−6.86
Burkholderiales	53	87.0	11.6	1.4	0.0	−6.33
Campylobacteriales	4	86.0	12.0	2.0	0.0	−5.95
Chroococcales	30	86.8	11.1	2.1	0.0	−5.60
Deinococcales	43	84.9	12.9	2.2	0.0	−6.17
Enterobacteriales 1	67	85.0	13.6	1.4	0.0	−5.57
Enterobacteriales 2	23	86.6	12.8	0.7	0.0	−6.39
Halobacteriales	67	86.5	11.5	2.0	0.0	−6.41
Lactobacillales	40	84.7	13.3	2.0	0.0	−5.62
M. tuberculosis	48	86.1	11.1	2.1	0.7	−5.69
Nitrosomonales	46	89.1	9.5	1.4	0.0	−6.08
Nostocales	5	85.0	14.3	0.7	0.0	−6.05
Pasturellales	17	81.9	16.0	2.1	0.0	−6.60
Prochlorales 1	22	88.4	9.5	2.0	0.0	−5.94
Prochlorales 2	33	88.4	8.2	3.4	0.0	−6.23
Pseudomonales	39	86.3	11.0	2.7	0.0	−6.42
Rhizobiales 1	49	81.4	14.5	4.1	0.0	−6.39
Rhizobiales 2	34	82.6	15.3	2.1	0.0	−6.24
Rhodobacteriales	32	83.9	13.3	2.8	0.0	−6.24
Vibrionales	24	82.9	14.4	2.7	0.0	−6.83

structures range from −5.57 to −6.86 (with an average value of −6.21), confirming the good quality of the model structures.

All UreG model structures and their surface representations calculated in this work are reported in Figures SI-2 and SI-3, respectively, and are available at http://www.postgenomicnmr.net. All models share almost the same overall tertiary structure, with a central open β-barrel, formed by six parallel strands and a variable number (from 0 to 2) of antiparallel strands, surrounded by five or, rarely, six α-helices connected with loops. On average, 36.3% of the modeled residues of UreG are involved in α-helices and 19.4% constitute the central β-barrel, whereas the remaining 44.3% comprises turns and coils (Table SI-2).

Figure 7A shows a backbone superimposition of all calculated UreG model structures. All models present a similar structure, with a good degree of superimposition. The only exception is the region located between β-strands S2 and S3 (generally modeled as α-helix H2). The position of α-helix H2 influences the accessibility of the GTP binding pocket, which appears to exist in an "open" or "closed" conformation. This aspect is evident by comparing the molecular surfaces of the model structures from *Prochlorococcus marinus* (*Pm*UreG, model number 22), *E. coli* (*Ec*UreG, number 23), *Vibrio parahaemolyticus* (*Vp*UreG, number 24), as reported in Figure 7B. In the *Ec*UreG model structure, the position of α-helix H2 allows the open conformation, while in the *Vp*UreG model, its position closes the hydrophilic pocket. In the *Pm*UreG model structure, helix H2 is absent and the pocket is partially open.

The region between α-helix H2 and β-strand S4 is the most difficult to model because the template structure sequences do not completely cover this part, but on the other hand, it is a specific feature of UreG proteins. It is noteworthy that, near this region, a fully conserved motif is present ([E/Q]TG[G/A]CPH, residues 86−92 in *Mt*UreG numeration) that may be involved in Zn^{2+} ion binding (*23*).

MtUreG Model Calculation. Figure 8 reports the optimized alignment and the model structure of *Mt*UreG. Like the other UreG proteins, *Mt*UreG consists of a central open β-barrel, with six parallel strand and one antiparallel strand, surrounded by six α-helices connected by loops. Forty percent of the modeled residues of *Mt*UreG are involved in α-helices, and 20% are part of the central β-barrel. The remaining 40% is modeled as turns and coils. The P-loop is located between strand S1 and helix H1 and resides on one side of a deep pocket defined by the loop located between strand S5 and helix H4, on the other side. The N-terminal and C-terminal portions of *Mt*UreG, comprising residues 1−21 and residues 192−225, were not modeled due to the absence of a structural template for these regions.

DISCUSSION

This study investigated the structural behavior of UreG, an essential GTPase in the urease system, using both experimental and computational analysis. Among all UreG sequences found in a database search, the one from *M. tuberculosis* contains a distinctive His-rich N-terminal region, which, in other urease systems, resides in the chaperone UreE and plays a role in intracellular nickel storage (*18*). This trait was also found in some HypB GTPases (*19*) and in CooJ (*20*), as discussed above. The occurrence of this characteristic feature in *Mt*UreG may be related to the lack of UreE in the genome of *M. tuberculosis* (*7*). These observations suggest a modular organization of the functions required for incorporation of metal in nickel-containing centers, with different functions residing in different domains of the same protein or, alternatively, located in different polypeptides, and prompted us to further investigate the properties of UreG in the peculiar case represented by *M. tuberculosis*.

The *pET22b* expression plasmid, containing the *MtureG* gene fused to a C-terminal His tag, was employed to overexpress the recombinant protein in *E. coli*. *Mt*UreG was totally expressed in the soluble fraction of the cellular extract (Figure 1), as reported in the past for *Ka*UreG (*52*) and *Hp*UreG (*53*). In contrast, *Bp*UreG has been isolated primarily from inclusion bodies (*23*). The elution profile and

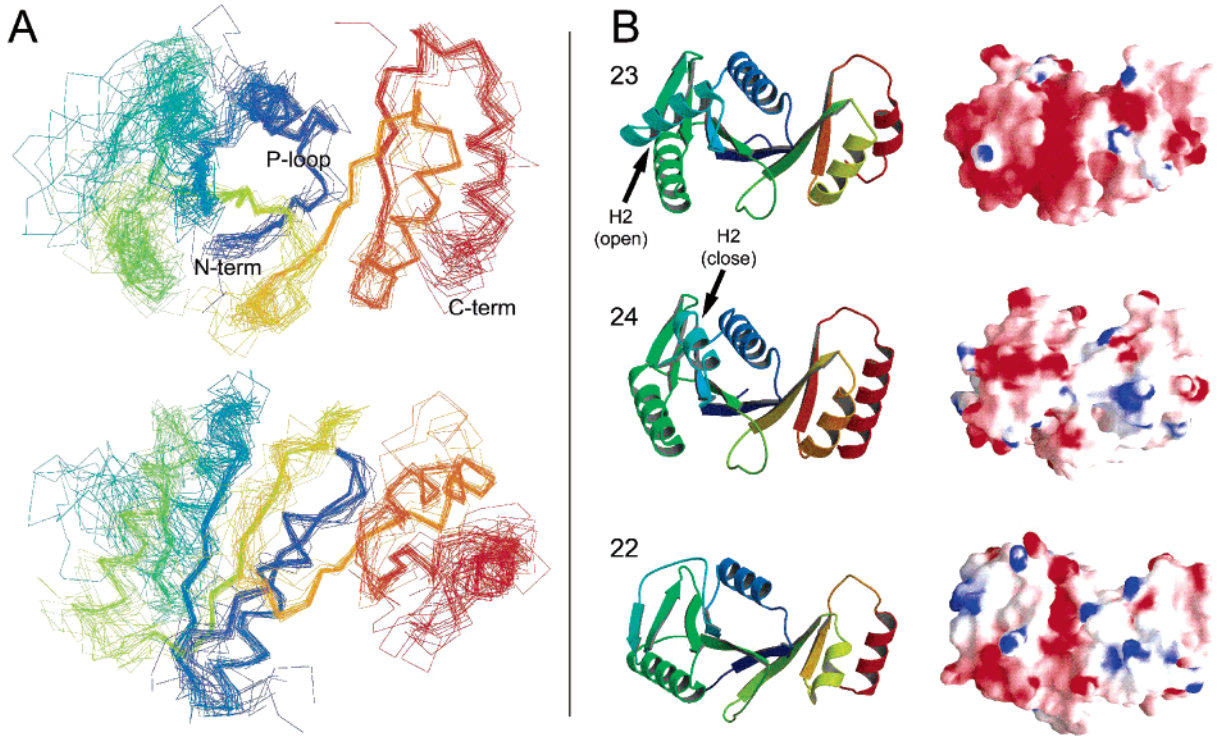


FIGURE 7: (A) Backbone superimposition of modeled structures of UreG. The proteins are shown with the P-loop toward the viewer (top) and rotated by 90° around the long horizontal axis (bottom). (B) Model structures of UreG from *E. coli* (top), *V. parahaemolyticus* (middle), and *P. marinus* (bottom). UreG is depicted as a ribbon diagram (left) or GRASP solid surface representation (right). The proteins are shown with the P-loop toward the viewer. The secondary structure elements range from deep blue in the proximity of the N-terminus to red at the C-terminus. Ribbon diagrams were made with MOLSCRIPT (*55*) and RASTER3D (*56*). The surface is colored according to the calculated electrostatic potential contoured from −10.0 (intense red) to +10.0 kT/e (intense blue) (where k is the Boltzmann constant, T the absolute temperature, and e the electron charge).

the hydrodynamic properties of *Mt*UreG under native conditions were investigated using a combination of size-exclusion chromatography and multiple-angle light scattering, showing that in solution the protein is homodimeric (Figure 2). This oligomeric form corresponds to that observed in the past for *Bp*UreG (*23*). In *Mt*UreG, as for *Bp*UreG, a disulfide bond is present in the dimer. Indeed, in spite of the presence of two cysteine residues in the protein sequence (Cys^{46} and Cys^{90} in *Mt*UreG numeration), the estimation of only one free thiol per monomer in a solution of *Mt*UreG indicates that one cysteine must be oxidized. Since Cys^{90} is fully conserved among all UreG sequences (*23*) and is involved in dimerization of *Bp*UreG (*24*), it is reasonable to presume that this residue is the one oxidized in the intermolecular disulfide bond also in *Mt*UreG.

In solution, *Mt*UreG contains a well-defined secondary structure, with 8% of the residues in an α-helix and 29% in a β-strand conformation (Figure 4); however, the protein does not show a rigid tertiary structure and displays a conformationally fluctional behavior as detected by NMR (Figure 5). The unstructured tract of native *Mt*UreG is further confirmed by the partial lack of secondary structure in the native protein when compared to the secondary structure composition derived from the molecular model of the fully folded protein (40% α-helices, 20% β-sheets). In *Mt*UreG, this variation mostly involves the α-helices that, in the model, surround the internal and hydrophobic β-barrel. A similar deficiency of α-helical content relative to the theoretical prediction was recently demonstrated for *E. coli* HypA, a protein involved in the synthesis of the [NiFe]-hydrogenase active site (*54*).

The isolated *Mt*UreG exhibits GTPase activity under native conditions, with a k_{cat} of 0.01 min^{-1}. This activity is comparable but lower than that reported for *Bp*UreG (0.04 min^{-1}) (*23*). This indicated that in vitro *Mt*UreG, even if isolated from the soluble fraction of the cellular extract, is less active than its homologue, which was isolated from inclusion bodies. Such lower functionality possibly reflects the presence of a smaller amount of tertiary structure in *Mt*UreG with respect to *Bp*UreG under native conditions, a hypothesis supported also by the presence of an α-helical content in *Mt*UreG smaller than that in *Bp*UreG, as shown by the analysis of CD spectra. In this regard, it is worth mentioning that previous studies have never detected any GTPase activity for isolated *Ka*UreG (*52*) and *Hp*UreG (*53*), which were expressed in the soluble cellular fraction as well.

The described results, coupled to the previously reported observations on *Bp*UreG and the NMR spectrum observed for *Ka*UreG (*23*), point out that the intrinsically unfolded character observed for UreG is probably a general property of this class of accessory proteins and is not directly related to a particular sequence, organism, and/or purification protocol.

The evidence of structural conservation among UreG proteins belonging to different organisms prompted us to perform an extensive similarity search, to identify all known bacterial UreG sequences, and to compare their primary and, if possible, secondary and tertiary structures. In total, 67 protein sequences, belonging to a nonredundant sum of different databases, were retrieved. The high degree of sequence similarity (from 69.2 to 47.8%), as well as the

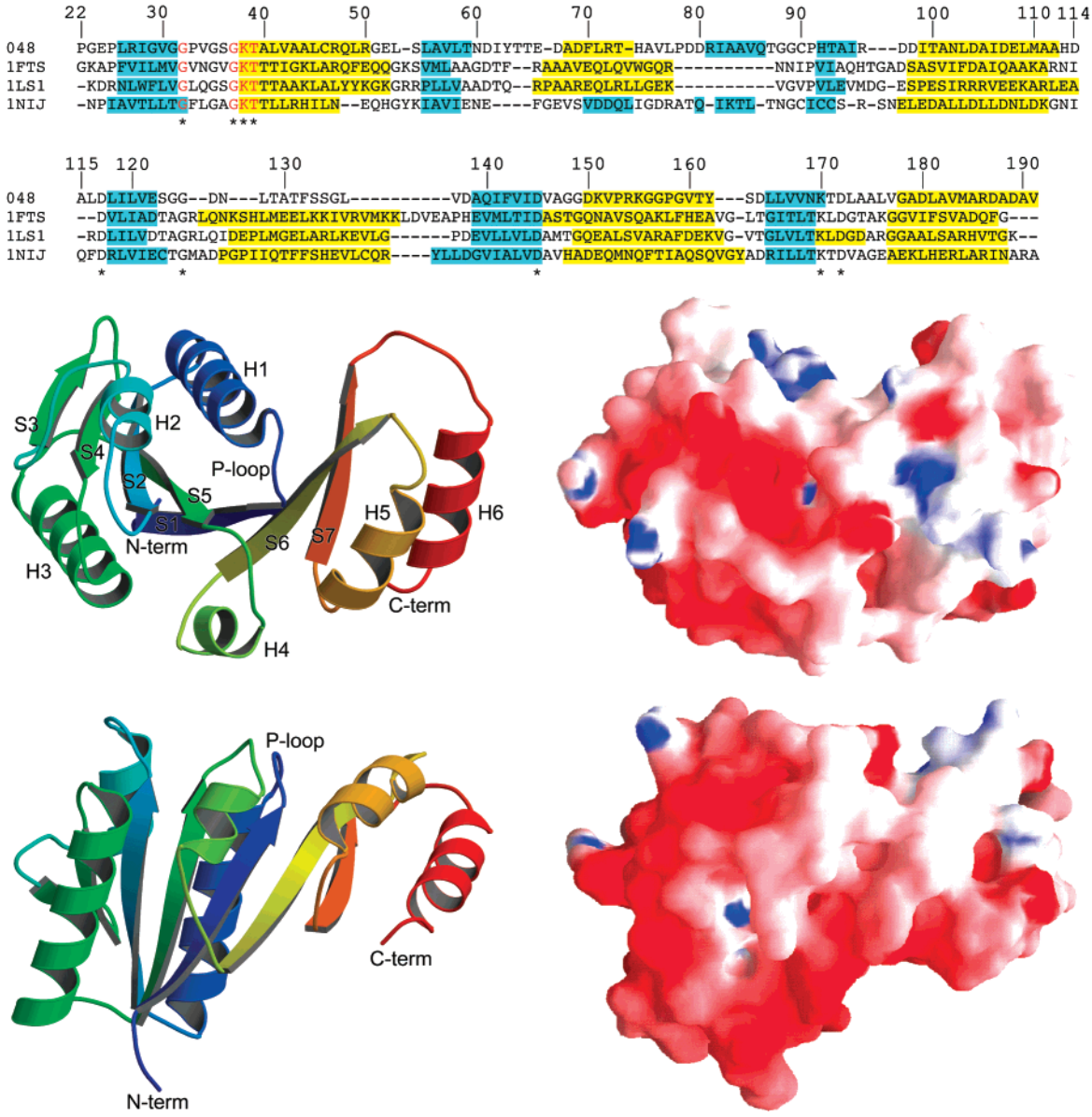


FIGURE 8: Structural model of *Mt*UreG. Multiple-sequence alignment (top) of *Mt*UreG (sequence 48 in the Figure 5 numeration) with the template structures as obtained from ClustalW and optimized using JPRED and PROSA energy profile. The selected templates are the G domain of signal recognition protein Ffh from *Thermus aquaticus* (PDB entry 1LS1), the signal recognition particle receptor from *E. coli* (PDB entry 1FTS), and the hypothetical protein Yjia from *E. coli* (PDB entry 1NIJ). The predicted secondary structure elements are highlighted in yellow (helices) and cyan (strands). The residues involved in the P-loop are red boldface type. The fully conserved residues are denoted with an asterisk. Middle and bottom panels report *Mt*UreG model structures shown as ribbon diagrams (left) and GRASP solid surface representations (right). The protein is shown with the P-loop toward the viewer (middle) and rotated by 90° around the long horizontal axis (bottom). The secondary structure elements range from deep blue in the proximity of the N-terminus to red at the C-terminus. Ribbon diagrams were made with MOLSCRIPT (*55*) and RASTER3D (*56*). The surface is colored according to the calculated electrostatic potential contoured from −10.0 (intense red) to +10.0 kT/e (intense blue) (where k is the Boltzmann constant, T the absolute temperature, and e the electron charge).

secondary structure conservation observed in the multiple-sequence alignment, supports the idea of a general structural homology among the UreG family, in agreement with experimental results for *Bp*UreG, *Mt*UreG, and *Ka*UreG. The presence, in all the retrieved sequences, of some conserved motifs, typical of GTPases, confirms the role of UreG in GTP hydrolysis. The putative zinc-binding site (Glu/Gln^{86}, Cys^{90}, Pro^{91}, and His^{92} in *Mt*UreG numeration) (*23*) is also fully conserved, indicating its possible functional importance.

Structural models were generated for *Mt*UreG and for the most representative UreG sequences belonging to different regions of the phylogenetic tree, using a threading procedure already employed for *Bp*UreG model calculation (*23*). This study further confirms the presence of an overall conserved structure for the UreG family, with every protein folded in an internal β-barrel surrounded by α-helices. The lack of structural agreement for the region that surrounds α-helix 2 suggests two possible alternative positions for this part of protein, which would result in an open or closed hydrophilic GTP binding pocket. The structural flexibility in this region, previously indicated by the disorder predictor algorithms (*23*), may be a peculiarity of UreG, and it is possibly related to the presence, in this region, of the putative zinc-binding site. It is plausible that, upon zinc binding, the region folds in a more ordered conformation, contributing to UreG activation or stabilization.

All the reported results clearly indicate the presence of a general structural conservation among the UreG family and an apparent functional importance of the particular behavior of this class of GTPases. These findings, coupled to the evidence of structural similarity previously demonstrated for many UreE proteins (*18*), support the idea of a common mechanism for delivery of nickel into urease enzymes in different organisms. Since nickel is fundamental for urease function, this result contributes to the identification of general targets, potentially able to control enzymatic activity in diverse ureolytic organisms, in some cases, as in *M. tuberculosis*, representing pathogenic bacteria. These findings are suitable for important biological and medical applications, such as the eradication of several important bacterial infections.

SUPPORTING INFORMATION AVAILABLE

3D-Jury meta-predictions for folding templates of modeled UreG, multiple-sequence alignments of target UreG's with the selected templates, all UreG model structures and their surface representations, and percentages of secondary structural elements calculated for all UreG structural models. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Finney, L. A., and O'Halloran, T. V. (2003) Transition metal speciation in the cell: Insights from the chemistry of metal ion receptors, *Science 300*, 931−936.
2. Tottey, S., Harvie, D. R., and Robinson, N. J. (2005) Understanding how cells allocate metals using metal sensors and metallochaperones, *Acc. Chem. Res. 38*, 775−783.
3. Rosenzweig, A. C. (2002) Metallochaperones: Bind and deliver, *Chem. Biol. 9*, 673−677.
4. O'Halloran, T. V., and Culotta, V. C. (2000) Metallochaperones, an intracellular shuttle service for metal ions, *J. Biol. Chem. 275*, 25057−25060.
5. Kasprzak, K. S., Sunderman, F. W., Jr., and Salnikow, K. (2003) Nickel carcinogenesis, *Mutat. Res. 533*, 67−97.
6. Mulrooney, S. B., and Hausinger, R. P. (2003) Nickel uptake and utilization by microorganisms, *FEMS Microbiol. Rev. 27*, 239−261.
7. Clemens, D. L., Lee, B. Y., and Horwitz, M. A. (1995) Purification, characterization, and genetic analysis of *Mycobacterium tuberculosis* urease, a potentially critical determinant of host-pathogen interaction, *J. Bacteriol. 177*, 5644−5652.
8. Gordon, A. H., Hart, P. D., and Young, M. R. (1980) Ammonia inhibits phagosome-lysosome fusion in macrophages, *Nature 286*, 79−80.
9. Hausinger, R. P., and Karplus, P. A. (2001) Urease, in *Handbook of Metalloproteins* (Messerschmidt, A., Huber, R., Poulos, T., and Wieghardt, K., Eds.) pp 867−879, John Wiley & Sons, Chichester, U.K.
10. Ciurli, S., and Mangani, S. (2001) Nickel-containing enzymes, in *Handbook on Metalloproteins* (Bertini, I., Sigel, A., and Sigel, H., Eds.) pp 669−708, Marcel Dekker, New York.
11. Benini, S., Rypniewski, W. R., Wilson, K. S., Miletti, S., Ciurli, S., and Mangani, S. (1999) A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: Why urea hydrolysis costs two nickels, *Struct. Folding Des. 7*, 205−216.
12. Jabri, E., Carr, M. B., Hausinger, R. P., and Karplus, P. A. (1995) The crystal structure of urease from *Klebsiella aerogenes*, *Science 268*, 998−1004.
13. Ha, N.-C., Oh, S.-T., Sung, J. Y., Cha, K. A., Lee, M. H., and Oh, B.-H. (2001) Supramolecular assembly and acid resistance of *Helicobacter pylori* urease, *Nat. Struct. Biol. 8*, 505−509.
14. Mobley, H. L. T., Island, M. D., and Hausinger, R. P. (1995) Molecular biology of microbial ureases, *Microbiol. Rev. 59*, 451−480.
15. Stola, M., Musiani, F., Mangani, S., Turano, P., Safarov, N., Zambelli, B., and Ciurli, S. (2006) The nickel site of *Bacillus pasteurii* UreE, a urease metallo-chaperone, as revealed by metal-binding studies and X-ray absorption spectroscopy, *Biochemistry 45*, 6495−6509.
16. Remaut, H., Safarov, N., Ciurli, S., and Van Beeumen, J. J. (2001) Structural basis for Ni transport and assembly of the urease active site by the metallo-chaperone UreE from *Bacillus pasteurii*, *J. Biol. Chem. 276*, 49365−49370.
17. Song, H.-K., Mulrooney, S. B., Huber, R., and Hausinger, R. P. (2001) Crystal structure of *Klebsiella aerogenes* UreE, a nickel-binding metallochaperone for urease activation, *J. Biol. Chem. 276*, 49359−49364.
18. Musiani, F., Zambelli, B., Stola, M., and Ciurli, S. (2004) Nickel trafficking: Insights into the fold and function of UreE, a urease metallochaperone, *J. Inorg. Biochem. 98*, 803−813.
19. Olson, J. W., and Maier, R. J. (2000) Dual roles of *Bradyrhizobium japonicum* nickel in protein in nickel storage and GTP-dependent Ni mobilization, *J. Bacteriol. 182*, 1702−1705.
20. Kerby, R. L., Ludden, P. W., and Roberts, G. P. (1997) In vivo nickel insertion into the carbon monoxide dehydrogenase of *Rhodospirillum rubrum*: Molecular and physiological characterization of cooCTJ, *J. Bacteriol. 179*, 2259−2266.
21. Freyermuth, S. K., Bacanamwo, M., and Polacco, J. C. (2000) The soybean Eu3 gene encodes an Ni-binding protein necessary for urease activity, *Plant J. 21*, 53−60.
22. Soriano, A., and Hausinger, R. P. (1999) GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins, *Proc. Natl. Acad. Sci. U.S.A. 96*, 11140−11144.
23. Zambelli, B., Stola, M., Musiani, F., De Vriendt, K., Samyn, B., Devreese, B., Van Beeumen, J., Turano, P., Dikiy, A., Bryant, D. A., and Ciurli, S. (2005) UreG, a chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds Zn^{2+}, *J. Biol. Chem. 280*, 4684−4695.
24. Neyroz, P., Zambelli, B., and Ciurli, S. (2006) Intrinsically disordered structure of *Bacillus pasteurii* UreG as revealed by steady-state and time-resolved fluorescence spectroscopy, *Biochemistry 45*, 8918−8930.
25. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct express of cloned genes, *Methods Enzymol. 185*, 66−89.
26. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures, *Protein Expression Purif. 41*, 207−234.
27. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680−685.
28. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248−254.
29. Lanzetta, P. A., Alvarez, L. J., Reinach, P. S., and Candia, O. A. (1979) An improved assay for nanomole amounts of inorganic phosphate, *Anal. Biochem. 100*, 95−97.
30. Lobley, A., Whitmore, L., and Wallace, B. A. (2002) DICHROWEB: An interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics 18*, 211−212.
31. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
32. Pearson, W. R., and Lipman, D. J. (1988) Improved Tools for Biological Sequence Comparison, *Proc. Natl. Acad. Sci. U.S.A. 85*, 2444−2448.
33. Pearson, W. R. (1990) Rapid and Sensitive Sequence Comparison with FASTP and FASTA, *Methods Enzymol. 183*, 63−98.
34. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *J. Mol. Biol. 215*, 403−410.
35. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
36. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.

37. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998) Jpred: A Consensus Secondary Structure Prediction Server, *Bioinformatics 14*, 892−893.
38. Felsenstein, J. (1993) *PHYLIP* (Phylogenetic Inference Package), 3.5c ed., Department of Genetics, University of Washington, Seattle.
39. Ginalski, K., Elofsson, A., Fischer, D., and Rychlewski, L. (2003) 3D-Jury: A simple approach to improve protein structure predictions, *Bioinformatics 19*, 1015−1018.
40. Ginalski, K., and Rychlewski, L. (2003) Detection of reliable and unexpected protein fold predictions using 3D-Jury, *Nucleic Acids Res. 31*, 3291−3292.
41. Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000) Comparative protein structure modeling of genes and genomes, *Annu. Rev. Biophys. Biomol. Struct. 29*, 291−325.
42. Sippl, M. J. (1993) Recognition of errors in three-dimensional structures of proteins, *Proteins: Struct., Funct., Genet. 17*, 355−362.
43. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
44. Vriend, G. (1990) WHATIF: A molecular modelling and drug design program, *J. Mol. Graphics 8*, 52−56.
45. Nicholls, A., Bharadwaj, R., and Honig, B. (1993) GRASP: Graphical representation and analysis of surface properties, *Biophys. J. 64*, A166.
46. Wen, J., Arakawa, T., and Philo, J. S. (1996) Size-exclusion chromatography with on-line light scattering, absorbance, and refractive index detectors for studying proteins and their interactions, *Anal. Biochem. 240*, 155−166.
47. Wyatt, P. J. (1993) Light scattering and the absolute characterization of macromolecules, *Anal. Chim. Acta 272*, 1−40.
48. Habeeb, A. F. S. A. (1972) Reaction of protein sulfhydryl groups with Ellman's reagent, *Methods Enzymol. 25*, 457−464.
49. Johnson, W. C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures, *Proteins 35*, 307−312.
50. Sprang, S. R. (1997) G protein mechanisms: Insights from structural analysis, *Annu. Rev. Biochem. 66*, 639−678.
51. Salomone-Stagni, M., Zambelli, B., Musiani, F., and Ciurli, S. (2007) A model-based proposal for the role of UreF as a GTPase-activating protein in the urease active site biosynthesis, *Proteins: Struct., Funct., Bioinf.*, in press.
52. Moncrief, M. C., and Hausinger, R. P. (1997) Characterization of UreG, identification of a UreD-UreF-UreG complex, and evidence suggesting that a nucleotide-binding site in UreG is required for in vivo metallocenter assembly of *Klebsiella aerogenes* urease, *J. Bacteriol. 179*, 4081−4086.
53. Mehta, N., Benoit, S., and Maier, R. J. (2003) Roles of conserved nucleotide-binding domains in accessory proteins, HypB and UreG, in the maturation of nickel-enzymes required for efficient *Helicobacter pylori* colonization, *Microb. Pathog. 35*, 229−234.
54. Atanassova, A., and Zamble, D. B. (2005) *Escherichia coli* HypA is a zinc metalloprotein with a weak affinity for nickel, *J. Bacteriol. 187*, 4689−4697.
55. Kraulis, P. J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 949−950.
56. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic molecular graphics, *Methods Enzymol. 277*, 505−524.

BI6024676