JOURNAL OF CHEMICAL INFORMATION AND MODELING

LETTER

pubs.acs.org/jcim

Model Structures of *Helicobacter pylori* UreD(H) Domains: A Putative Molecular Recognition Platform

Francesco Musiani,[†] Matteo Bellucci,[†,§] and Stefano Ciurli*[,†,‡]

[†]Laboratory of Bioinorganic Chemistry, University of Bologna, Viale G. Fanin 40, 40127 Bologna, Italy

[‡]Center for Magnetic Resonance (CERM), University of Florence, Florence, Italy

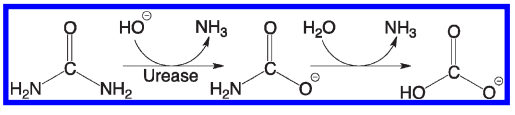
Supporting Information

ABSTRACT: The analysis of the sequence of *Helicobacter pylori* UreD(H), an accessory protein involved in the activation of urease through the assembly of the Ni^{2+}-containing active site, revealed the presence of two domains. The structure of these domains was calculated using threading and modeling algorithms. A search for putative binding sites on the protein surface was carried out using dedicated algorithms sensitive to either sequence conservation or structural similarity based on geometry and physicochemical properties. The results suggest that UreD(H) acts as a multifunctional molecular recognition platform facilitating the interaction between apo-urease and the ancillary proteins UreG, UreF, and UreE, responsible for nickel trafficking and delivering.

INTRODUCTION

The essentiality of transition metals for living organisms, coupled with their limited environmental availability and toxicity, has prompted life to develop strict mechanisms for specific and selective metal ion sensing, utilization, and trafficking.[1,2] The Ni^{2+}-dependent urease enzymatic system represents an emblematic case of the metal ions essentiality/toxicity paradigm and is useful to unravel the strategies used by cells to handle nickel, a rare and potentially harmful element, and to turn it into a micronutrient for essential biological functions. Urease, produced by eukaryotes and prokaryotes, catalyzes urea hydrolysis to give ammonia and carbamate, which spontaneously decomposes to give a second molecule of ammonia and hydrogen carbonate (Scheme 1).[3,4]

Scheme 1



This reaction is responsible for the decomposition of urea, produced by the majority of vertebrates as the catabolic product of nitrogen-containing compounds. The hydrolysis of the reaction products induces an overall pH increase that has negative implications for human health as well as in the environmental ecosphere. Urease efficiency and essentiality are exemplified by the role played in the survival of the human pathogen *Helicobacter pylori* (*Hp*). This bacterium is able to colonize the gastric mucosa through the activity of urease, which creates a local pH environment compatible with the survival of this microorganism in such a specific niche that cycles from nearly neutral to very acidic conditions.[5]

Several crystal structures of urease from bacteria[6−8] and plants[9] are available. The active site of the enzyme contains two Ni^{2+} ions bridged by a fully conserved, post-translationally carbamylated lysine residue and by a hydroxide ion, which acts as the nucleophile in the catalytic mechanism.[4,10] Urease is synthesized in vivo as an inactive apoenzyme and undergoes a maturation process that involves Ni^{2+} incorporation and lysine carbamylation, guided by four accessory proteins: UreG, UreF, UreE, and UreD (named UreH in *H. pylori*, hereafter termed UreD(H)).[11] The current proposed activation mechanism proceeds via a series of protein−protein associations between the apoenzyme and these protein chaperones, followed by the delivery of Ni^{2+} concomitantly with the guanosine 5′-triphosphate (GTP)-dependent transfer of a CO_2 molecule necessary for lysine carbamylation.[11,12]

UreG is responsible for the GTP hydrolysis associated to the transfer of CO_2 to the active site lysine; the protein is intrinsically disordered, specifically binds Zn^{2+}, and features very low or absent hydrolyzing activity in vitro, probably needing the interaction with other protein partners to achieve its fully folded and functional state.[13] UreF has been suggested to act as an activator of the GTPase activity of UreG on the basis of structural biomodeling studies,[14] and the recent structure of a truncated form of *Hp*UreF appears to support this hypothesis.[15] Finally, UreE is believed to act as a Ni^{2+} carrier based on the evidence that the levels of Ni^{2+} necessary for the in vitro assembly of the urease active site are significantly reduced, and a much larger portion of enzyme molecules is activated if UreE is present.[16−18] In *H. pylori*, UreE and UreG directly interact to achieve the urease maturation process, as demonstrated by yeast two-hybrid analysis[19,20] as well as coimmunoprecipitation assays[20] and isothermal titration calorimetry.[21] A calculated model structure of the UreE−UreG complex has been proposed.[21]

Among all urease accessory proteins, the structural properties and the functional role of UreD(H) remain largely obscure. Most of the evidence has been collected on the *Klebsiella aerogenes* (*Ka*) system. Deletion of the *KaureD* gene blocks in vivo[22] and in vitro[23] *K. aerogenes* urease activation. The presence of protein complexes containing *Ka*UreD and apo-urease in 1:1, 2:1, and

Published: May 29, 2011

3:1 stoichiometric ratios has been detected; notably, the assemblies dissociate in the presence of Ni^{2+}, suggesting that UreD stabilizes a conformation of urease that is competent for in vivo nickel incorporation.[23] The UreE, UreF, and UreG proteins are not required for formation of the *Ka*UreD−urease complex, suggesting a direct interaction between UreD and apo-urease, even though the presence of UreF appears to alter the conformation of the UreD−urease complex.[24] Consistently, complexes of apo-urease/UreD/UreF[25] and apo-urease/UreD/UreF/UreG[24] have been reported. Moreover, an insoluble UreD/UreF/UreG complex has been detected[26] and isolated in a urease-activation competent complex with apo-urease,[27] further activated in the presence of UreE.[18]

Overall, these results consistently highlight a role of UreD in the formation of a multiprotein complex with apo-urease and the other accessory proteins that is necessary for urease activation through lysine carbamylation and concomitant insertion of Ni^{2+} ions in the active site. However, very little is known about the interaction between UreD and the other urease accessory components. Two-hybrid analyses and immuno-precipitation studies identified UreD−urease and UreD−UreF complexes in *Proteus mirabilis*[28] and *Helicobacter pylori*.[19,20] Chemical cross-linking experiments coupled with mass spectrometry allowed the identification of a superficial region on the *K. aerogenes* urease surface that directly interacts with *Ka*UreD.[29] This result was confirmed using flexibility analysis, mutagenesis, and small-angle X-ray scattering, which also suggested the occurrence of a conformational change induced on urease by interaction with UreD.[30] However, no information on the UreD side of the interaction has ever been obtained by any means.

*Ka*UreD has been repeatedly reported to be completely insoluble when synthesized in the absence of the other accessory proteins or of the urease enzyme itself.[16,23,26,31−33] The complete insolubility of recombinant isolated *Hp*UreD(H) has also been observed experimentally in our laboratory (data not shown). Therefore, UreD has never been overexpressed alone, leading to difficulty in characterizing this protein in solution. The only reported case in which UreD has been studied in solution refers to the translational fusion product between *Ka*UreD and the maltose binding protein (MBP-UreD).[33] This chimerical protein binds Ni^{2+} and Zn^{2+} with micromolar dissociation constants and forms strong and weak complexes with UreF and UreG, respectively, and surprisingly it does not bind to apo-urease. Furthermore, UreD maintains its insolubility when separated from MBP, and MBP-UreD forms large multimers in solution (>670 KDa) that can only partially complement the urease activation network. These observations suggest that MBP-UreD is present as an agglomerate of misfolded protein, maintained in solution by the folded and highly soluble MBP moieties, as observed in other cases,[34] challenging the physiological relevance of this chimerical product. Finally, the study did not report any structural characterization of the UreD component of the fusion protein.

The present work is an attempt to unravel the structural properties of *Hp*UreD(H) and its putative interaction surface using a biomodeling approach, in the actual unfeasibility of experimental approaches. The sequence of *Hp*UreD(H) was analyzed, and two distinct domains were identified. The structures of these domains were calculated using algorithms for threading and modeling in order to investigate and possibly clarify the function of *Hp*UreD(H) in the urease activation process. The presence of surface regions likely involved in protein−protein interactions was investigated using dedicated algorithms. The results provide new bases for possible strategies to clone and express the protein and to reveal some unexpected features that suggest a possible role of UreD(H) as a molecular recognition platform, acting as protein hub for the recruitment of other urease accessory components.

MATERIALS AND METHODS

The DomPred server[35] was used to investigate the presence of domains in the sequence of UreD(H) from *H. pylori* 26695. The BioInfoBank Meta Server/3D-Jury[36] was utilized to search for proteins compatible with the sequence of the two identified *Hp*UreD(H) domains. Multiple sequence alignments of the *Hp*UreD(H) domains sequences with the corresponding selected templates were performed using ClustalW,[37] while manual optimization of the alignments was carried out using information derived from the secondary structure predictions calculated by PSIPRED.[38] The obtained alignments were used to calculate 50 structural models for each *Hp*UreD(H) domain using the program Modeller 9v8.[39] The best model for each domain was selected on the basis of the lowest value of the DOPE score in Modeller[40] and was subjected to a refining step of loop optimization using Modeller. The stereochemical quality of the structures was evaluated using PROCHECK,[41] and the distribution of residual energy was evaluated in ProSA.[42] The results of this analysis, reported in Table 1SI, Supporting Information, confirm the high reliability of the model structures.

The molecular (solvent-excluded) surfaces and graphics of the domains of *Hp*UreD(H) were calculated using the Chimera package.[43] The electrostatic color coding was generated using the DelPhi software,[44] using the default atomic radii and charges parameters, a box fill of 60%, and a grid spacing of 1 Å. All histidine residues were considered neutral. The protein internal dielectric constant was set to 4 in all calculations, and the solvent dielectric constant was 80. The salt concentration was set to 0.15 M NaCl, which corresponds to the physiological ionic strength.

Detection of putative binding sites on the surface of the *Hp*UreD(H) domains was explored using two different computational tools: (i) the ConSurf server[45] was used to calculate surface residue conservation, and (ii) the ProBiS server[46,59] was used to search for surface regions bearing structural similarity with known protein binding sites based on geometric and physicochemical parameters. In the case of ConSurf, the sequence and the structure of the two domains were used as input to search for homologues in the NR nonredundant protein database by using the program PSI-BLAST, an *E* value cutoff of 0.0001, a maximal percentage of sequence identity of 95%, and a minimal percentage for homologues of 35%. Conservation scores were mapped onto the protein structure using a representative coloring scheme. In the case of ProBiS, the results were filtered using the default "neutral" settings, which entailed 5 fingerprint residues, a minimum surface patch size of 10 vertices, a surface vector angle of 1.5708, a surface patch root-mean-square deviation (RMSD) of 2 Å, and an *E* value of 0.0001.

RESULTS

Structure Modeling. The search for structural analogues of *Hp*UreD(H) was initially conducted using the full-length protein sequence as input for a large number of databases and predictors. These initial attempts did not provide reliable results and prompted

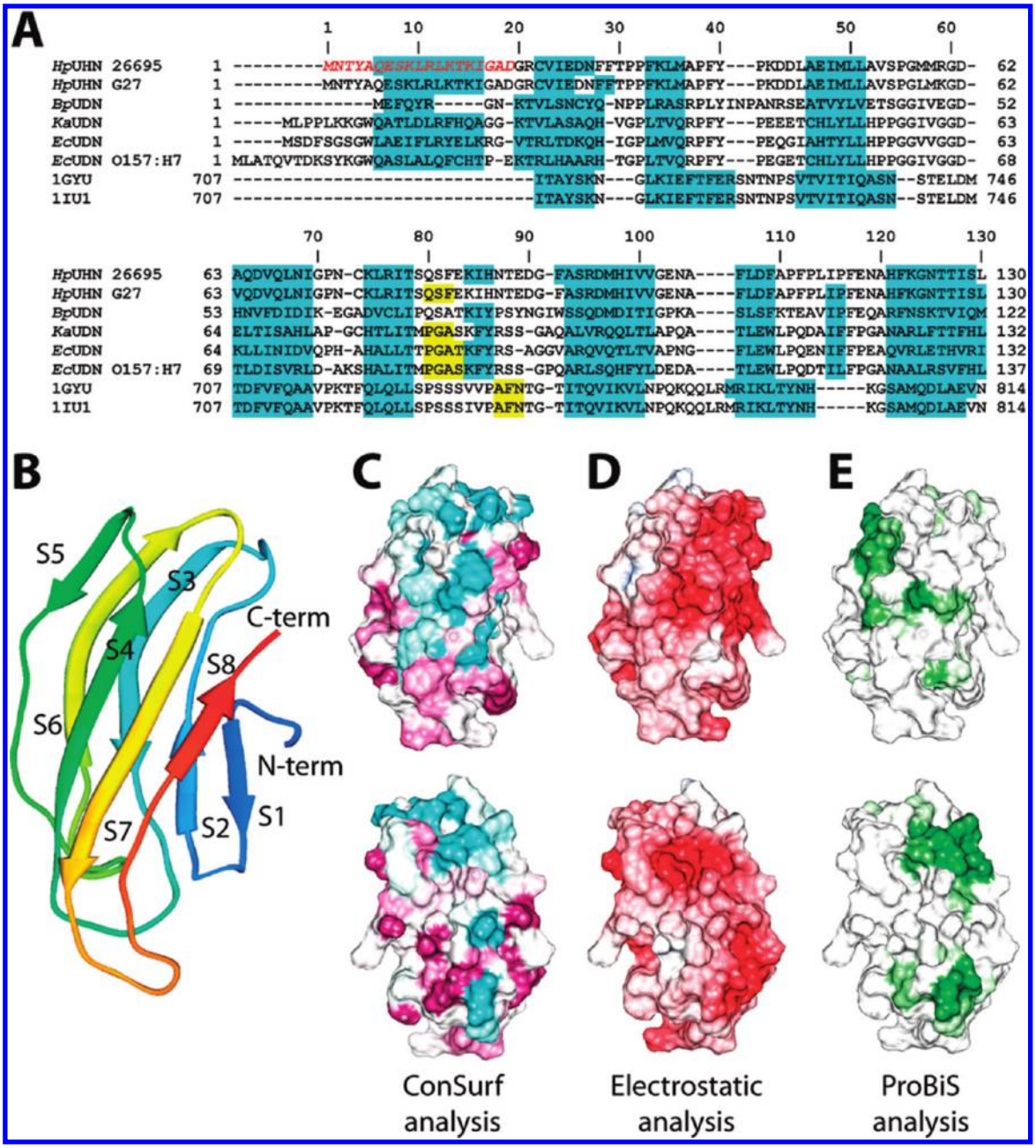


Figure 1. (A) Alignment of the sequences of UreD(H) N-terminal domain from *H. pylori* 26695 (*Hp*UHN 26696), *H. pylori* G27 (*Hp*UHN G27), *Bacillus pasteurii* (*Bp*UDN), *K. aerogenes* (*Ka*UDN), *E. coli* (*Ec*UDN), *E. coli* O157:H7 (*Ec*UHN O157:H7), and γ-adaptin appendage domain from mouse clathrin adaptor AP-1 (1GYU) and to ear domain of human γ1-adaptin (1IU1). The secondary structure derived from PSIPRED prediction for UreD(H) sequences and from the PDB structure for the templates is indicated (α-helix, yellow; β-strand, cyan). Residues in red and italics were not modeled due to the absence of available structural template structures. (B) Ribbon scheme of the N-terminal domain of the *Hp*UreD model structure colored from blue in the proximity of the N-terminal to red at the C-terminus. (C−D) Solvent excluded and surface colored according to the residue conservation calculated with ConSurf (C: maroon to white to turquoise as the extent of conservation decreases), to the electrostatic potential (D: red, negative; blue, positive), and to the degree of structural conservation calculated with ProBiS (E: green to white as the extent of conservation decreases). The surface in the bottom panels of C−D are rotated by 180° around the vertical axis respect to the surfaces in the top panels.

us to investigate in depth *Hp*UreD(H)'s protein composition and modularity. In particular, the DomPred server, which predicts domain boundaries in target sequences using a combined homology and fold recognition-based approach, revealed the presence of two distinct domains, located at the N-terminal (residues 1−130) and at the C-terminal (residues 131−265). On this basis, the search for putative template structures was thus repeated using the sequences of the two identified modules.

The adaptin C-terminal domain family (PFAM ID: PF02883), and specifically the structurally elucidated γ-adaptin appendage domain from mouse clathrin adaptor AP-1 (PDB code: 1GYU),[47] and the ear domain of human γ1-adaptin (PDB: 1IU1)[48] were detected as feasible templates for the N-terminal portion of *Hp*UreD(H), encompassing residues 20−130 (Figure 1A). The γ-adaptins are heterotrimeric adaptor proteins involved in the interaction of clathrin with cargo proteins and lipids at vesicle

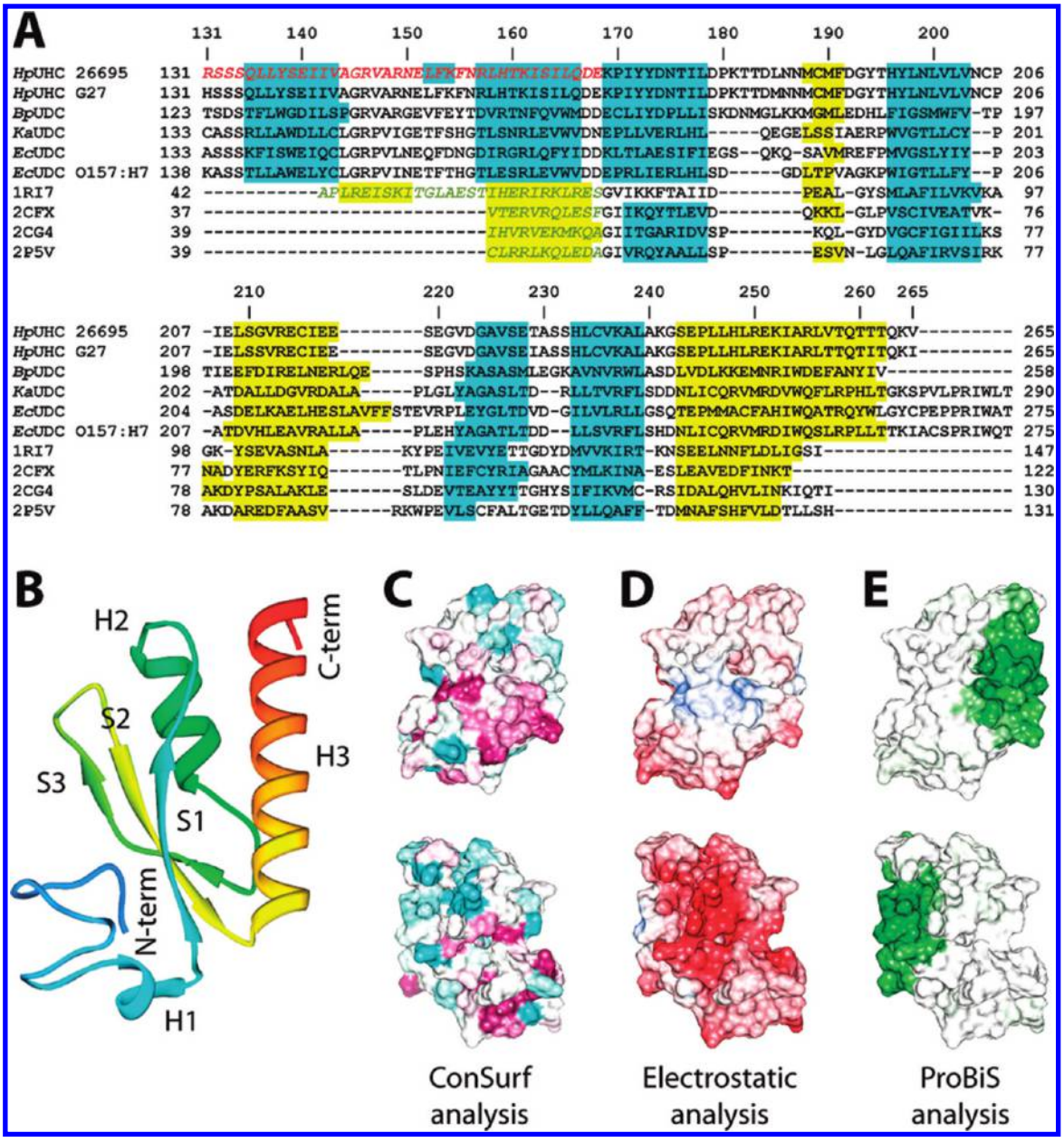


Figure 2. (A) Alignment of the sequences of UreD(H) C-terminal domain from *H. pylori* 26695 (*Hp*UHN 26696), *H. pylori* G27 (*Hp*UHN G27), *Bacillus pasteurii* (*Bp*UDN), *K. aerogenes* (*Ka*UDN), *E. coli* (*Ec*UDN), *E. coli* O157:H7 (*Ec*UHN O157:H7), and C-terminal domain of the AsnC/Lrp family of bacterial transcription regulators from *Pyrococcus* sp. (1RI7), *Bacillus subtilis* (2CFX), *E. coli* (2CG4) and *Neisseria meningitides* (2P5V). The secondary structure derived from PSIPRED prediction for UreD(H) sequences and from the PDB structure for the templates is indicated (α-helix, yellow; β-strand, cyan). Residues in red and italics were not modeled due to the absence of available structural template structures. Residues in gray in the template structures were not employed in the modeling due to the predicted structural differences respect to the target sequence. (B) Ribbon scheme of the C-terminal domain of *Hp*UreD model structure colored from blue in the proximity of the N-terminal to red at the C-terminus. (C−E) Solvent excluded surface colored according to the residue conservation calculated with ConSurf (C: maroon to white to turquoise as the extent of conservation decreases), to the electrostatic potential (D: red, negative; blue, positive), and to the degree of structural conservation calculated with ProBiS (E: green to white as the extent of conservation decreases). The surface in the bottom panels of C−E are rotated by 180° around the vertical axis respect to the surfaces in the top panels.

budding sites as well as binding accessory proteins that regulate coat assembly and disassembly.[49] On the basis of homology modeling calculations based on these two template structures, the *Hp*UreD(H) N-terminal domain appears to be an eight-stranded β sandwich (Figure 1B) containing one five-stranded antiparallel β-sheet I (made of strands S1−S3 and S5 and S6) and one three-stranded antiparallel β-sheet II (made of strands S4, S7, and S8) in a Greek key topology. This architecture properly reflects the immunoglobulin-like fold found in α, β, and γ adaptins,[50,51] stabilized by an hydrophobic core formed by residues Phe[33], Leu[35], Met[36], and Ala[37] from strand S2, Leu[50] and Ala[52] from S3, Ile[85] located on the loop between S5 and S6, Ala[93] from S6, and Leu[107] and Phe[109] from S7. In this domain, Pro[32] was found to be highly conserved among UreD proteins;[52]

dx.doi.org/10.1021/ci200183n |*J. Chem. Inf. Model.* 2011, 51, 1513–1520

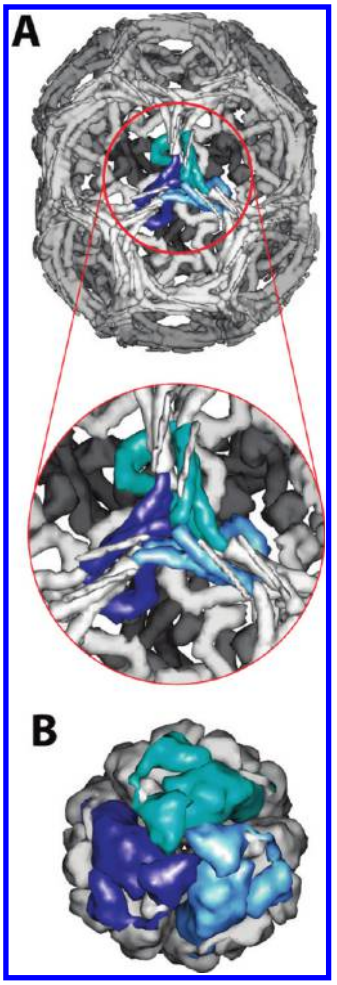


Figure 3. Surface of *Bos taurus* clathrin (panel A, PDB code 1XI4) and *H. pylori* urease (panel B, PDB code 1E9Z) showing the trigonal arrangement of subunits displayed in different shades of blue. The lower picture in panel A shows a zoom of the clathrin surface having the same dimensions as urease.

this residue is located in the loop connecting strands S1 and S2, in a solvent-exposed position.

The C-terminal domain of *Hp*UreD(H), covering residues 158−265, appears to be similar to the C-terminal domain of the AsnC/Lrp family of bacterial transcription regulators (PFAM ID: 01037)[53] and specifically to protein domains from *Pyrococcus* sp. (PDB: 1RI7),[54] *Bacillus subtilis* (PDB: 2CFX),[55] *Escherichia coli* (PDB: 2CG4),[55] and *Neisseria meningitides* (PDB: 2P5V)[56] (Figure 2A). These proteins are multimers of dimers (1RI7, dimer of dimers; 2CFX, 2CG4, 2P5V tetramer of dimers) in which the C-terminal domain is responsible for oligomerization. The N-terminal domain of these transcription factors, which does not feature any sequence or structural similarity with *Hp*UreD(H), contains a HTH (helix−turn−helix) motif, and is associated with DNA binding. The model structure of the *Hp*UreD(H) C-terminal domain thus appears to be formed by a three-stranded antiparallel β-sheet comprised by a small α-helix on one side (H1) and by two α-helices (H2 and H3) on the other side (Figure 2B), reflecting the typical $\alpha\beta$-sandwich fold of the AsnC/Lrp family. This domain, also designated as regulation of amino acid metabolism (RAM), is found as a stand-alone module in many prokaryotic genomes,[57] in fusion with DNA-binding domains encoding Lrp-homologues, or fused to a catalytic domain in different enzyme classes.[58] As such, its main role resides in protein oligomerization mechanisms. As in the case of the N-terminal domain, the fold is stabilized by a hydrophobic core formed by Val^{212} and Ile^{216} from helix H2, Val^{222} and Ala^{225} from strand S1, Val^{236} and Leu^{239} from S2, and Leu^{249}, Ile^{253}, and Val^{257} from H3. In this domain, Pro^{206} is highly conserved,[52] is located in the loop connecting strand S1 and helix H2, and is positioned in a solvent-exposed spot possibly involved in molecular interactions.

Protein Surface Analysis. An understanding of the mechanism of protein−protein interactions, a function that appears, as experimentally suggested, to be related to UreD(H) proteins, requires a detailed analysis of the protein surface to identify putative binding sites. In the present study, the search for regions on the surface of the two modeled domains of *Hp*UreD(H) that could be involved in protein−protein interactions was carried out using two different approaches that involved, respectively, either sequence or structural conservation. This was compared with the analysis of the surface electrostatics carried out using DelPhi,[44] a software that calculates electrostatic potentials by incorporating the effects of ionic strength through the nonlinear Poisson−Boltzmann equation. The ConSurf server[45] carries out a residue conservation search using the three-dimensional protein structure and a phylogenetic relation among its close sequence homologues. The server extracts the protein sequence from the PDB file used as input and automatically carries out a search for close homologous sequences. The multiple sequence alignment that results is used to build a phylogenetic tree and to calculate the conservation scores. The protein surface can finally be visualized with the sequence conservation score color-coded onto its surface. On the other hand, the ProBiS server[46,59] uses the input PDB file to identify the solvent accessible surface atoms, which are then represented as vertices replacing the functional groups within the residues on the protein surface. Each vertex is labeled according to a set of five physicochemical properties associated to the functional group that it represents (hydrogen-bond acceptor or donor, mixed hydrogen-bond acceptor and donor, aliphatic, or aromatic residue). Then, the query protein structure is compared in a pairwise manner with each of the ca. 27 000 nonredundant protein structures in the ProBiS database, all of them represented, in analogy with the query protein, as ensembles of vertices. Finally, an alignment is carried out, a scoring rank is produced, and a structural conservation score is visualized as a color-coded surface.

The electrostatic surface potential map of the *Hp*UreD(H) N-terminal domain (Figure 1D) reveals the presence of localized patches of negative electric potential, surrounded by large hydrophobic neutral areas. In particular, clusters of Glu and Asp residues are found at the N-terminal (Glu^{25} and Asp^{26}), in the loops connecting S2−S3 and S6−S7 (Asp^{43}, Asp^{44}, Glu^{103}) and S5−S6 (Glu^{83}, Glu^{89}, Asp^{90}) as well as on strands S3 and S6 (Glu^{47} and Asp^{96}, respectively). According to the analysis conducted with ConSurf[45] (Figure 1C), residues 20, 22, 29−32, 34, 38, 39, 54−56, 61, 65, 80, 81, 83−85, 88, 91, 93, 105, 107, 109, 113, 115, 118, 119, 122−124, 130, and 132 show conservation scores in the 8−9 range, which is generally considered a good indication for reliable binding site predictions. All these residues are hydrophobic (with the exception of Asp^{61}, Lys^{84}, and Glu^{83}, involved in an alternate chain of negative and positive charge interactions stabilizing the loops connecting S3−S4 and S5−S6) and compose either the protein core or a large neutral patch

centered on Phe[122]. An analysis performed with ProBiS[46] identified the three strands S3, S5, and S6 of β-sheet I and the entire β-sheet II, as surface regions putatively involved in protein−protein interactions (Figure 1E), with a distribution of electrostatic charge on these regions ranging from neutral to negative. These regions have, as a common feature, their flatness and little residue conservancy.

The *Hp*UreD(H) C-terminal domain (Figure 2D) can be envisioned as divided in two zones featuring a negatively charged and a hydrophobic protein surface located on the opposite sides of the $\alpha\beta$-sandwich. The ConSurf[45] analysis (Figure 2C) identifies residues 174, 175, 178, 180, 190, 191, 195−197, 199, 217, 219, 227, 237−239, and 243−247 having high conservation scores (8 or 9). Most of these residues constitute the hydrophobic core of the protein, with some notable exception of the residues in the 243−247 range, found at the beginning of H3 and probably important for the initiation and stabilization of this helix. As in the case of the N-terminal domain model structure, the most conserved regions of the protein surface correspond to hydrophobic patches (residues Met[190], Phe[191], His[196], Tyr[197], and Ser[243]) potentially acting as putative docking positions for the formation of the quaternary structure of the full protein or, more in general, involved in mediating protein−protein interactions. Consistently, ProBiS[46,59] indicates that the entire helix H3 features a well-defined ensemble of fingerprinting residues possessing structural and chemical properties previously observed to be involved in intermolecular interactions in a wide range of biological settings.

DISCUSSION

The lack of experiment-based information on the structural properties of UreD proteins, due to their well-established insolubility, constitutes a severe obstacle for the comprehension of the role of this protein in the urease activation process. The identification of two domains within the protein sequence prompted us to overproduce, using recombinant DNA techniques, the two separate domains of *Hp*UreD(H), which, however, also resulted insoluble (data not shown). Therefore, a bioinformatics approach was used in the present study, which involved homology-based molecular modeling of the two separate domains and an analysis of superficial putative hot spots for protein−protein recognition. In general, this analysis, conducted using different algorithms that investigate different surface properties, has allowed the identification of protein regions possibly involved either in intra- or interdomain interactions/stabilization or in interactions between UreD(H) and the different accessory proteins in the urease activation process.

The fold resemblance of the N-terminal domain to adaptor proteins (AP),[50,51] also named adaptins, suggests the possibility that UreD(H) employs this module to directly interact with the apo-urease protein and/or with the other accessory proteins UreF and UreG, known to form a supercomplex with UreD.[24−27] Adaptins are a well-known class of macromolecules featuring a wide repertoire of protein-binding modules that interlink related partners in order to regulate the assembly of larger protein complexes. Historically, the term "adaptin" is referred to a group of proteins that copurify with clathrin upon isolation of clathrin-coated vesicles.[60] Clathrin is a triskelion-shaped trimer of heavy chain subunits, each connected to a single small chain, which constitutes the repetitive unit found in hexagonal lattice coat structure (Figure 3A).[61] This trimer-based architecture geometrically resembles the trigonal structure adopted by urease, in particular that from *H. pylori*, built according to a peculiar supramolecular organization made of a tetramer of trimer of $\alpha\beta$ heterodimers (Figure 3B). Therefore, the similarity between the N-terminal domain of UreD(H) with the eukaryotic AP-1 ear domain, known to interact with clathrin and to mediate the recruitment of other proteins such as γ-synergin, rabaptin-5, and cyclin G-associated kinase, is consistent with a role of UreD(H) as a protein interaction hub.

The coupling of an adaptin-like module (that is, the N-terminal domain), with an AsnC/Lrp-like module (the C-terminal domain) able to promote protein oligomerization, could represent an expedient to modulate the interaction of a multimeric target, such as the urease enzyme with the set of urease accessory proteins, UreF and UreG. A well-documented case in support of this hypothesis is the bric-a-brac tramtrack broad complex (BTB) protein−protein interaction motif,[62] a globular domain that generally contains a five α-helices cluster capped at one end by a short three-stranded β-sheet. Also in this case, a close similarity is found between BTB and AsnC/Lrp domains from a topological and a functional point of view, both being $\alpha\beta$-sandwich motifs involved in protein oligomerization. Following this reasoning, the occurrence of a domain similar to AsnC/Lrp in the *Hp*UreD(H) sequence could be rationalized with the high variability of protein−protein interaction modules, as in the case of BTB, which allows to evolve diverse but convergent functional capabilities of proteins.

In conclusion, our modeling studies allowed us to propose for UreD(H) a putative role as a molecular recognition platform,[63−65] providing an efficient docking module for the correct recruitment and functional positioning of multiple protein partners. In this framework, UreD(H) could operate as a protein hub for apo-urease and the accessory proteins UreF and UreG, having key roles in urease activation. The experimental validation of our models could possibly stem from a comparison of their features with the secondary structure content of an isolated UreDFG complex, provided the same information become available for the isolated UreF and UreG proteins, the former being studied so far only as a truncated form,[15] and the latter being known to exist in an intrinsically disorder form possibly undergoing a disorder-to-order transition upon binding to partner proteins.[13]

ASSOCIATED CONTENT

Supporting Information. Table 1SI containing the parameters of ProCheck and ProSA analysis on *Hp*UreD(H) template and model structures. This information is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: stefano.ciurli@unibo.it Telephone: (+39)-051-2096204.

Present Addresses

[§]Bioinformatics & Genomic Program, Centre for Genomic Regulation (CRG), C/Dr. Aiguader 88, 08003 Barcelona, Spain.

ACKNOWLEDGMENT

F.M. was a recipient of a postdoctoral fellowship from UniBO and of a fellowship from CERM-CIRMMP, and M.B. was

supported by a Ph.D. fellowship granted by UniBO. The structures of the two domains of *Hp*UreD(H) are freely available upon request.

REFERENCES

(1) Finney, L. A.; O'Halloran, T. V. Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **2003**, *300*, 931–936.

(2) Giedroc, D. P.; Arunkumar, A. I. Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Trans.* **2007**, 3107–3120.

(3) Ciurli, S. In *Biological Inorganic Chemistry: Structure & Reactivity*; Bertini, I., Gray, H. B., Valentine, J. S., Eds.; University Science Books: Sausalito, CA, 2007, p 198−208.

(4) Ciurli, S. In *Nickel and its surprising role in nature*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; Wiley & Sons, Ltd.: Chichester, U.K., 2007, p 241−278.

(5) Zanotti, G.; Cendron, L. Functional and structural aspects of *Helicobacter pylori* acidic stress response factors. *IUBMB Life* **2010**, *62*, 715–723.

(6) Jabri, E.; Carr, M. B.; Hausinger, R. P.; Karplus, P. A. The crystal structure of urease from Klebsiella aerogenes. *Science* **1995**, *268*, 998–1004.

(7) Benini, S.; Rypniewski, W. R.; Wilson, K. S.; Miletti, S.; Ciurli, S.; Mangani, S. 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. *Structure* **1999**, *7*, 205–216.

(8) Ha, N. C.; Oh, S. T.; Sung, J. Y.; Cha, K. A.; Lee, M. H.; Oh, B. H. Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. *Nat. Struct. Biol.* **2001**, *8*, 505–509.

(9) Balasubramanian, A.; Ponnuraj, K. Crystal structure of the first plant urease from jack bean: 83 years of journey from its first crystal to molecular structure. *J. Mol. Biol.* **2010**, *400*, 274–283.

(10) Musiani, F.; Arnofi, E.; Casadio, R.; Ciurli, S. Structure-based computational study of the catalytic and inhibition mechanisms of urease. *J. Biol. Inorg. Chem.* **2001**, *6*, 300–314.

(11) Quiroz, S.; Kim, J. K.; Mulrooney, S. B.; Hausinger, R. P. In *Nickel and its surprising role in nature*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; Wiley & Sons, Ltd.: Chichester, UK, 2007, p 519−544.

(12) Kaluarachchi, H.; Chan Chung, K. C.; Zamble, D. B. Microbial nickel proteins. *Nat. Prod. Rep.* **2010**, *27*, 681–694.

(13) Zambelli, B.; Turano, P.; Musiani, F.; Neyroz, P.; Ciurli, S. Zn2+-linked dimerization of UreG from *Helicobacter pylori*, a chaperone involved in nickel trafficking and urease activation. *Proteins* **2009**, *74*, 222–239.

(14) Salomone-Stagni, M.; Zambelli, B.; Musiani, F.; Ciurli, S. A model-based proposal for the role of UreF as a GTPase-activating protein in the urease active site biosynthesis. *Proteins* **2007**, *68*, 749–761.

(15) Lam, R.; Romanov, V.; Johns, K.; Battaile, K. P.; Wu-Brown, J.; Guthrie, J. L.; Hausinger, R. P.; Pai, E. F.; Chirgadze, N. Y. Crystal structure of a truncated urease accessory protein UreF from Helicobacter pylori. *Proteins* **2010**, *78*, 2839–2848.

(16) Park, I.-S.; Hausinger, R. P. Metal ion interactions with urease and UreD-urease apoproteins. *Biochemistry* **1996**, *35*, 5345–5352.

(17) Park, I. S.; Hausinger, R. P. Requirement of carbon dioxide for in vitro assembly of the urease nickel metallocenter. *Science* **1995**, *267*, 1156–1158.

(18) Soriano, A.; Colpas, G. J.; Hausinger, R. P. UreE stimulation of GTP-dependent urease activation in the UreD-UreF-UreG-urease apoprotein complex. *Biochemistry* **2000**, *39*, 12435–12440.

(19) Rain, J. C.; Selig, L.; De Reuse, H.; Battaglia, V.; Reverdy, C.; Simon, S.; Lenzen, G.; Petel, F.; Wojcik, J.; Schachter, V.; Chemama, Y.; Labigne, A.; Legrain, P. The protein-protein interaction map of Helicobacter pylori. *Nature* **2001**, *409*, 211–215.

(20) Voland, P.; Weeks, D. L.; Marcus, E. A.; Prinz, C.; Sachs, G.; Scott, D. Interactions among the seven *Helicobacter pylori* proteins encoded by the urease gene cluster. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2003**, *284*, G96–G106.

(21) Bellucci, M.; Zambelli, B.; Musiani, F.; Turano, P.; Ciurli, S. *Helicobacter pylori* UreE, a urease accessory protein: specific Ni2+- and Zn2+-binding properties and interaction with its cognate UreG. *Biochem. J.* **2009**, *422*, 91–100.

(22) Lee, M. H.; Mulrooney, S. B.; Renner, M. J.; Markowicz, Y.; Hausinger, R. P. *Klebsiella aerogenes* urease gene cluster: sequence of ureD and demonstration that four accessory genes (ureD, ureE, ureF, ureG) are involved in nickel metallocenter biosynthesis. *J. Bacteriol.* **1992**, *174*, 4324–4330.

(23) Park, I. S.; Carr, M. B.; Hausinger, R. P. *In vitro* activation of urease apoprotein and role of UreD as a chaperone required for nickel metallocenter assembly. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3233–3237.

(24) Park, I. S.; Hausinger, R. P. Evidence for the presence of urease apoprotein complexes containing UreD, UreF, and UreG in cells that are competent for in vivo enzyme activation. *J. Bacteriol.* **1995**, *177*, 1947–1951.

(25) Moncrief, M. C.; Hausinger, R. P. Purification and activation properties of UreD-UreF-urease apoprotein complexes. *J. Bacteriol.* **1996**, *178*, 5417–5421.

(26) Moncrief, M. C.; Hausinger, R. P. 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.* **1997**, *179*, 4081–4086.

(27) Soriano, A.; Hausinger, R. P. GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11140–11144.

(28) Heimer, S. R.; Mobley, H. L. Interaction of *Proteus mirabilis* urease apoenzyme and accessory proteins identified with yeast two-hybrid technology. *J. Bacteriol.* **2001**, *183*, 1423–1433.

(29) Chang, Z.; Kuchar, J.; Hausinger, R. P. Chemical cross-linking and mass spectrometric identification of sites of interaction for UreD, UreF, and urease. *J. Biol. Chem.* **2004**, *279*, 15305–15313.

(30) Quiroz-Valenzuela, S.; Sukuru, S. C.; Hausinger, R. P.; Kuhn, L. A.; Heller, W. T. The structure of urease activation complexes examined by flexibility analysis, mutagenesis, and small-angle X-ray scattering. *Arch. Biochem. Biophys.* **2008**, *480*, 51–57.

(31) Kim, J. K.; Mulrooney, S. B.; Hausinger, R. P. The UreEF fusion protein provides a soluble and functional form of the UreF urease accessory protein. *J. Bacteriol.* **2006**, *188*, 8413–20.

(32) Boer, J. L.; Quiroz-Valenzuela, S.; Anderson, K. L.; Hausinger, R. P. Mutagenesis of *Klebsiella aerogenes* UreG to probe nickel binding and interactions with other urease-related proteins. *Biochemistry* **2010**, *49*, 5859–69.

(33) Carter, E. L.; Hausinger, R. P. Characterization of the *Klebsiella aerogenes* urease accessory protein UreD in fusion with the maltose binding protein. *J. Bacteriol.* **2010**, *192*, 2294–2304.

(34) Nomine, Y.; Ristriani, T.; Laurent, C.; Lefevre, J. F.; Weiss, E.; Trave, G. Formation of soluble inclusion bodies by hpv e6 oncoprotein fused to maltose-binding protein. *Protein Expression Purif.* **2001**, *23*, 22–32.

(35) Marsden, R. L.; McGuffin, L. J.; Jones, D. T. Rapid protein domain assignment from amino acid sequence using predicted secondary structure. *Protein Sci.* **2002**, *11*, 2814–2824.

(36) Ginalski, K.; Elofsson, A.; Fischer, D.; Rychlewski, L. 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics* **2003**, *19*, 1015–1018.

(37) Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **2003**, *31*, 3497–3500.

(38) Bryson, K.; McGuffin, L. J.; Marsden, R. L.; Ward, J. J.; Sodhi, J. S.; Jones, D. T. Protein structure prediction servers at University College London. *Nucleic Acids Res.* **2005**, *33*, W36–W38.

(39) Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.

(40) Shen, M. Y.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **2006**, *15*, 2507–2524.

(41) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.

(42) Wiederstein, M.; Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, W407–W410.

(43) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

(44) Honig, B.; Nicholls, A. Classical electrostatics in biology and chemistry. *Science* **1995**, *268*, 1144–1149.

(45) Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **2010**, *38*, W529–W533.

(46) Konc, J.; Janezic, D. ProBiS algorithm for detection of structurally similar protein binding sites by local structural alignment. *Bioinformatics* **2010**, *26*, 1160–1168.

(47) Kent, H. M.; McMahon, H. T.; Evans, P. R.; Benmerah, A.; Owen, D. J. Gamma-adaptin appendage domain: structure and binding site for Eps15 and gamma-synergin. *Structure* **2002**, *10*, 1139–1148.

(48) Nogi, T.; Shiba, Y.; Kawasaki, M.; Shiba, T.; Matsugaki, N.; Igarashi, N.; Suzuki, M.; Kato, R.; Takatsu, H.; Nakayama, K.; Wakatsuki, S. Structural basis for the accessory protein recruitment by the gamma-adaptin ear domain. *Nat. Struct. Biol.* **2002**, *9*, 527–531.

(49) Owen, D. J.; Collins, B. M.; Evans, P. R. Adaptors for clathrin coats: structure and function. *Annu. Rev. Cell Dev. Biol.* **2004**, *20*, 153–191.

(50) Boehm, M.; Bonifacino, J. S. Adaptins: the final recount. *Mol. Biol. Cell* **2001**, *12*, 2907–2920.

(51) Robinson, M. S. Adaptable adaptors for coated vesicles. *Trends Cell. Biol.* **2004**, *14*, 167–174.

(52) Steyert, S. R.; Rasko, D. A.; Kaper, J. B. Functional and phylogenetic analysis of ureD in Shiga toxin-producing Escherichia coli. *J. Bacteriol.* **2011**, *193*, 875–886.

(53) Brinkman, A. B.; Ettema, T. J.; de Vos, W. M.; van der Oost, J. The Lrp family of transcriptional regulators. *Mol. Microbiol.* **2003**, *48*, 287–294.

(54) Koike, H.; Ishijima, S. A.; Clowney, L.; Suzuki, M. The archaeal feast/famine regulatory protein: potential roles of its assembly forms for regulating transcription. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2840–2845.

(55) Thaw, P.; Sedelnikova, S. E.; Muranova, T.; Wiese, S.; Ayora, S.; Alonso, J. C.; Brinkman, A. B.; Akerboom, J.; van der Oost, J.; Rafferty, J. B. Structural insight into gene transcriptional regulation and effector binding by the Lrp/AsnC family. *Nucleic Acids Res.* **2006**, *34*, 1439–1449.

(56) Ren, J.; Sainsbury, S.; Combs, S. E.; Capper, R. G.; Jordan, P. W.; Berrow, N. S.; Stammers, D. K.; Saunders, N. J.; Owens, R. J. The structure and transcriptional analysis of a global regulator from *Neisseria meningitidis*. *J. Biol. Chem.* **2007**, *282*, 14655–14664.

(57) Nakano, N.; Okazaki, N.; Satoh, S.; Takio, K.; Kuramitsu, S.; Shinkai, A.; Yokoyama, S. Structure of the stand-alone RAM-domain protein from *Thermus thermophilus* HB8. *Acta Crystallogr.* **2006**, *F62*, 855–860.

(58) Ettema, T. J.; Brinkman, A. B.; Tani, T. H.; Rafferty, J. B.; Van Der Oost, J. A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes. *J. Biol. Chem.* **2002**, *277*, 37464–37468.

(59) Konc, J.; Janezic, D. ProBiS: a web server for detection of structurally similar protein binding sites. *Nucleic Acids Res.* **2010**, *38*, W436–W440.

(60) Pearse, B. M. Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1255–1259.

(61) Smith, C. J.; Pearse, B. M. Clathrin: anatomy of a coat protein. *Trends Cell. Biol.* **1999**, *9*, 335–338.

(62) Perez-Torrado, R.; Yamada, D.; Defossez, P. A. Born to bind: the BTB protein-protein interaction domain. *Bioessays* **2006**, *28*, 1194–1202.

(63) Dzhindzhev, N. S.; Yu, Q. D.; Weiskopf, K.; Tzolovsky, G.; Cunha-Ferreira, I.; Riparbelli, M.; Rodrigues-Martins, A.; Bettencourt-Dias, M.; Callaini, G.; Glover, D. M. Asterless is a scaffold for the onset of centriole assembly. *Nature* **2010**, *467*, 714–718.

(64) Seiradake, E.; Harlos, K.; Sutton, G.; Aricescu, A. R.; Jones, E. Y. An extracellular steric seeding mechanism for Eph-ephrin signaling platform assembly. *Nat. Struct. Mol. Biol.* **2010**, *17*, 398–402.

(65) Farré, J.-C.; Subramani, S. Rallying the exocyst as an autophagy scaffold. *Cell* **2011**, *144*, 172–174.