THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 12, Issue of March 22, pp. 10187–10193, 2002
Printed in U.S.A.

The Crystal Structure of *Helicobacter pylori* Cysteine-rich Protein B Reveals a Novel Fold for a Penicillin-binding Protein*

Received for publication, September 18, 2001, and in revised form, December 28, 2001
Published, JBC Papers in Press, January 2, 2002, DOI 10.1074/jbc.M108993200

Lucas Lüthy, Markus G. Grütter, and Peer R. E. Mittl‡

From the Biochemisches Institut, Universität Zürich, Winterthurer Strasse 190, 8057 Zürich, Switzerland

Colonization of the gastric mucosa with the spiral-shaped Gram-negative proteobacterium *Helicobacter pylori* is probably the most common chronic infection in humans. The genomes of *H. pylori* strains J99 and 26695 have been completely sequenced. Functional and three-dimensional structural information is available for less than one third of all open reading frames. We investigated the function and three-dimensional structure of a member from a family of cysteine-rich hypothetical proteins that are unique to *H. pylori* and *Campylobacter jejuni*. The structure of *H. pylori* cysteine-rich protein (Hcp) B possesses a modular architecture consisting of four α/α-motifs that are cross-linked by disulfide bridges. The Hcp repeat is similar to the tetratricopeptide repeat, which is frequently found in protein/protein interactions. In contrast to the tetratricopeptide repeat, the Hcp repeat is 36 amino acids long. HcpB is capable of binding and hydrolyzing 6-amino penicillinic acid and 7-amino cephalosporanic acid derivatives. The HcpB fold is distinct from the fold of any known penicillin-binding protein, indicating that the Hcp proteins comprise a new family of penicillin-binding proteins. The putative penicillin binding site is located in an amphipathic groove on the concave side of the molecule.

The large number of protein sequences that have been derived by more than 80 genome sequencing projects of archaea, bacteria, and eukaryotes (www.cbs.dtu.dk/services/GenomeAtlas/) has provided the scientific community with sequences where neither a function nor a three-dimensional structure is available. These sequences, which are annotated as "hypothetical proteins," will become a rich source of information, provided that their structures and biological functions are investigated. Here we present the structure and function analysis of a hypothetical protein from the pathogenic microorganism *Helicobacter pylori*. Several implications of *H. pylori* on human health have been established since its discovery in 1983 by Warren and Marshal (1). It is generally accepted that gastric diseases such as duodenal ulcers, gastric ulcers, adenocarcinoma of the distal stomach, and gastric mucosa-associated lymphoid tissue lymphoma are caused by *H. pylori*, and its implication in extradigestive diseases is under discussion. Infection by *H. pylori* has also been linked to dyspepsia and to a multitude of non-gastric diseases including cardiovascular, autoimmune, dermatological, and liver diseases. Implications of *H. pylori* on human health have been reviewed in several articles (2–5). In addition, it has also been reported that *H. pylori* infection may be beneficial and protect against gastric esophageal reflux disease (6).

The *H. pylori* genomes of strains 26695 and J99 have been completely sequenced, facilitating a detailed genome analysis (7, 8). For approximately two-thirds of all *H. pylori* ORFs,[1] functions were assigned by sequence comparison methods, and for approximately one-third, the three-dimensional structure of a homologous protein is available. Among the ORFs without a functional annotation, there is a group of hypothetical proteins that are rich in cysteine residues. Therefore the corresponding gene products are designated *Helicobacter* cysteine-rich proteins (Hcp) (9, 10). The Hcps, which are so far unique to microorganisms from the *Helicobacter* and *Campylobacter* genera, possess molecular sizes in the range between 15 and 40 kDa and show a stringent pattern of cysteine pairs. Two cysteine residues are separated by 7 amino acids, and there are 36 amino acids between adjacent cysteine pairs, suggesting that the Hcp proteins possess modular architectures of repetitive α/β-motifs. Sequence conservation among this family varies between 22 and 66% sequence identity (Fig. 1). It was shown recently that the *Helicobacter* cysteine-rich protein A (HcpA) possesses a β-lactamase activity, although there was no detectable sequence homology to known β-lactamases (10). To work toward a functional and structural characterization of the Hcp family, we expressed and characterized the HP0336 gene product, designated HcpB, and determined its crystal structure. The HcpB structure possesses a fold that is related to the structures of tetratricopeptide repeat proteins. This fold has so far never been observed for a penicillin-binding protein.

MATERIALS AND METHODS

Molecular Biology and Protein Expression—The plasmid GHPDN49 harboring the ORF HP0336 was obtained from the American Tissue and Culture Collection, and the ORF was amplified by PCR. The sequences of the sense and antisense primers were 5′-GCACCCCATGGTAGGGGGTGGAACGGTAAA-3′ and 5′-TACGCTCCCGGGTTA**GTGGTGGTGGTGGTGGTG**GTAGTTGTTTAAAATACCACATGC-3′, respectively. The PCR reaction amplified the entire HP0336 gene sequence and included additional *Nco*I and *Xma*I restriction sites (underlined) at the 5′- and 3′-ends, respectively. In addition, the PCR reaction introduced a stop codon and six codons for histidine residues (bold characters) at the 3′-end of the HP0336 gene.

The PCR products were inserted into pTFT74 expression vectors using the *Nco*I and *Xma*I restriction sites. After sequencing the inserted

* This work was supported by the Hartmann-Müller Foundation (Zürich/CH), the Baugarten Stiftung (Zürich/CH), and Grant 3100-063794.001 from the Swiss National Science Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1KLX) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ To whom correspondence should be addressed. Tel.: 41-1-635-6559; Fax: 41-1-635-6834; E-mail: mittl@bioc.unizh.ch.

[1] The abbreviations used are: ORF, open reading frame; GdmHCl, guanidinium hydrochloride; Hcp, *Helicobacter* cysteine-rich protein; NAM, *N*-acetylmuramic acid; PP5, human protein phosphatase 5; TPR tetratricopeptide repeat; PBP, penicillin-binding proteins; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid.

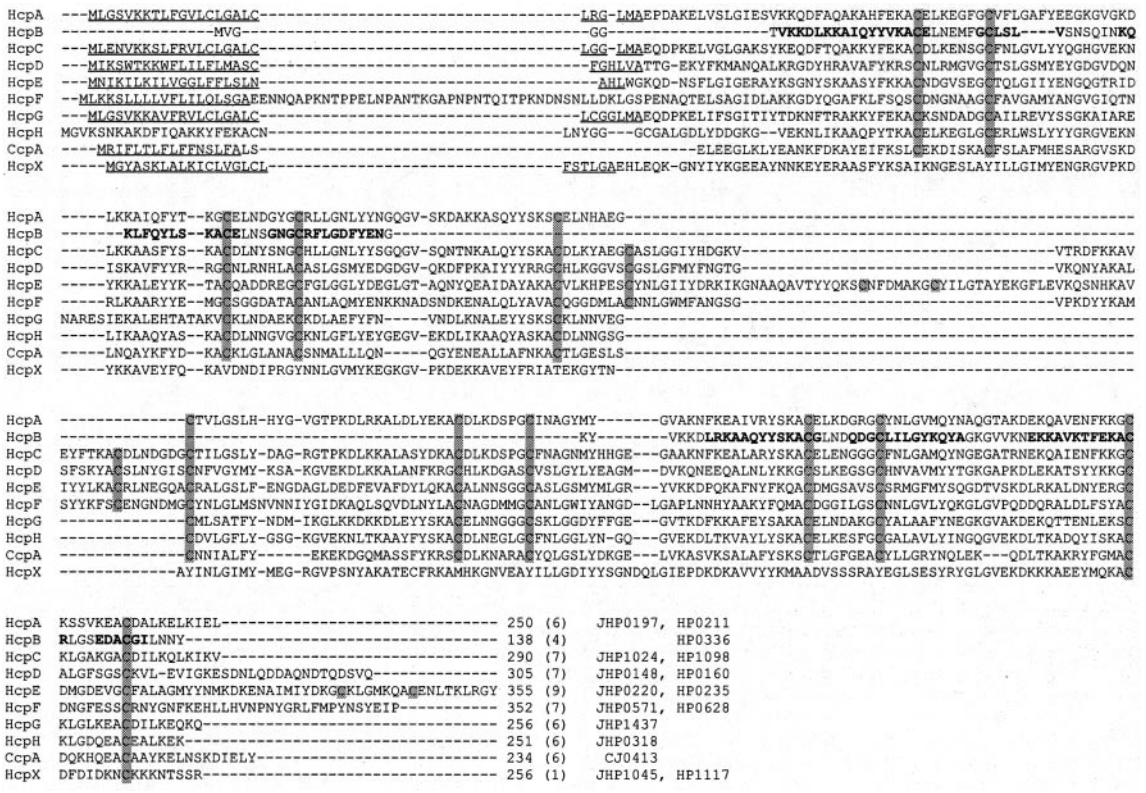

FIG. 1. **Multiple sequence alignment of the Hcp family generated with the program CLUSTALX (33).** Residues that are predicted by the SignalP server (34) to form leader peptides are *underlined*, and residues that form α-helices in the HcpB structure are shown in *bold* characters. Cysteine pairs are highlighted by *gray bars*. The protein nomenclature suggested earlier (9,10) is expanded to the entire Hcp family. At the end of each sequence, the *number* of residues, the *number* of cysteine pairs (in *parentheses*), and the *names* of the ORFs are given. ORFs starting with *JHP*, *HP*, and *CJ* refer to the *H. pylori* strain J99 (8), strain 26695 (7), and *Campylobacter jejuni* strain NCTC11168 genomes (35), respectively. In cases where orthologues in both *H. pylori* strains exist, only the sequence from the J99 strain is given.

ORF, the pTFT74/HP0336 plasmid was used to transform competent *Escherichia coli* BL21(DE3) cells. For the expression of native HcpB, protein cells were grown in LB medium at 37 °C with constant agitation (280 rpm). When an OD_{600} of 0.6 was reached, the expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside, and the culture was grown for an additional 3 h.

Selenomethionine-labeled HcpB was overexpressed in the same strain using M9 salt medium containing 1 mg/liter biotin and 1 mg/liter thiamin. 20 min before induction, additional L-selenomethionine (Sigma, 50 mg/liter), lysine hydrochloride (100 mg/liter), threonine (100 mg/liter), phenylalanine (100 mg/liter), leucine (50 mg/liter), isoleucine (50 mg/liter), and valine (50 mg/liter) were added as solid salts, and the culture was grown for an additional 13 h after induction.

Isolation of Inclusion Bodies—HcpB protein was refolded in a similar way to HcpA (10). Cells were harvested by centrifugation (30 min, 2000 × *g*, 4 °C), and the pellet was suspended in 10–20 ml of ice-cold lysis buffer (10 mM Tris/HCl, 2 mM magnesium chloride, pH 6.8). After passing the suspension two times through a French pressure cell, 50 μg/ml DNase and 65 μg/ml RNase were added, and the solution was incubated at 37 °C for 30 min. After adding EDTA and CHAPS to final concentrations of 25 mM and 0.25%, the solution was kept on ice for an additional 30 min. Inclusion bodies were collected by centrifugation (15 min, 20,000 × *g*, 4 °C), and the soluble fraction was discarded. The pellet was washed two times with buffer A (0.1 M Tris/HCl, 20 mM EDTA, pH 6.8) and subsequently buffer B (0.5 M GdmHCl in buffer A). Inclusion bodies were solubilized in buffer C (5 M GdmHCl, 0.2 M Tris/HCl, 0.1 M dithiothreitol, 10 mM EDTA, pH 8.0), and insoluble material was removed by centrifugation. Solubilized inclusion bodies were dialyzed overnight against buffer D (5 M GdmHCl, 0.1 M acetic acid). Protein concentration was determined by measuring the absorption at λ = 280 nm (ϵ_{280} = 14,800 M^{-1} cm^{-1}).

Refolding and Purification—HcpB was refolded by immobilizing the solubilized inclusion bodies on a nickel nitrilotriacetic acid-agarose (Qiagen) and removing the guanidinium hydrochloride from the buffer. To bind the unfolded inclusion bodies to the resin, 20 mg of unfolded HcpB was added to 5–10 ml of nickel nitrilotriacetic acid-agarose in buffer D. After adjusting the pH to 8.0, the slurry was filled into a column. The column was washed with 50 ml of buffer E (5 M GdmHCl, 0.1 M Tris, pH 8.0). HcpB was refolded by replacing buffer E immediately with buffer F (50 mM Tris/HCl, 150 mM sodium chloride, 5 mM glutathione, pH 8.0) and washing the column with 50 ml of buffer F at a flow rate of 1 ml/min. The protein was eluted with buffer G (250 mM imidazole, 50 mM Tris/HCl, 150 mM sodium chloride, 5 mM glutathione, pH 7.0). Protein containing fractions were pooled and dialyzed against 1000 ml of buffer H (40 mM sodium acetate, 1 mM EDTA, pH 5.5). Buffer H was also used for gel-permeation chromatography. After concentrating the protein in a Centriprep (Millipore), 0.4 ml of refolded HcpB (1 mg) was loaded onto a Superdex 75 HR 10/30 column (Amersham Biosciences, Inc.) at a flow rate of 0.5 ml/min. Purified HcpB eluted as a single peak at a volume of 13.47 ml. The comparison with the calibration profile (blue dextran (2 MDa), 8.63 ml; bovine serum albumin (67 kDa), 9.97 ml; ovalbumin (43 kDa), 10.90 ml; chymotrypsinogen A (25 kDa), 13.17 ml; ribonuclease A (13.7 kDa), 14.17 ml) revealed that HcpB eluted as a monomer.

Folding Characterization—The folding/unfolding behavior of HcpB was investigated by CD spectroscopy. Spectra were recorded at a protein concentration of 10 μM in 0–4 M GdmHCl, 5 mM sodium phosphate, pH 6.9 on a Jasco J-751 CD spectrometer. The temperature was maintained at 22 °C, and the data were fitted against Eq. 1 (11). Y_{obs} is the observed CD signal; a and b and c and d are the intercepts and the slopes at low and high GdmHCl concentrations, respectively. $[GdmHCl]_{½}$ is the GdmHCl concentration where half of the protein is unfolded, and m is the cooperativity of the unfolding reaction. R is the ideal gas constant, and T is the absolute temperature. The theoretical value for the cooperativity of the unfolding reaction was calculated according to the literature (12).

$$Y_{\text{obs}} = ((a \cdot [GdmHcl] + b)/(1 + k)) + (k \cdot (c \cdot [GdmHcl] + d)/(1 + k)) \quad \text{(Eq. 1)}$$

$$k = exp\ (-\Delta G^{GdmHcl}/RT) = exp\ (m \cdot ([GdmHcl]_{1/2} - [GdmHcl])/RT) \quad \text{(Eq. 2)}$$

TABLE I
Kinetic constants for β-lactam hydrolysis

Antibiotic	Δε	λ	[Antibiotic]	HcpA[a]			HcpB		
				K_m	k_{cat}	IC_{50}	K_m	k_{cat}	IC_{50}
	$M^{-}cm^{-}$	*nm*	μM	μM	min^{-1}	μM	μM	min^{-1}	μM
6-Aminopenillinic acid derivatives									
Ampicillin	−820	235	0.1–800	145	1.06 ± 0.04	ND[b]	174	0.76 ± 0.08	ND
Benzylpenicillin	−775	235	0.1–400	48	0.74 ± 0.02	ND	126	0.68 ± 0.14	ND
Amoxicillin	−800	235	0.1–400	155	0.55 ± 0.17	ND	47	0.45 ± 0.04	ND
Carbenicillin	−800	235	0.1–30,000	ND	0.04	63	ND	0.12	38
Cloxacillin	−800	235	0.1–30,000	ND	< 0.01	19	ND	< 0.01	17
Oxacillin	−800	235	0.1–30,000	ND	< 0.01	6.7	ND	< 0.01	10
7-Aminocephalosporanic acid derivatives									
Nitrocefin	16000	486	0.1–300	47	0.28 ± 0.08	ND	53	0.45 ± 0.10	ND
Cephaloridin	−10000	260	0.1–1000	406	0.30 ± 0.07	ND	>1000	ND	ND
Cefotaxim	−7500	260	0.1–30,000	ND	< 0.01	> 30,000	ND	< 0.01	> 30,000
Cefoxitin	−7700	260	0.1–30,000	ND	0.03	4300	ND	< 0.01	1100
Cephalothin	−6500	260	0.1–30,000	ND	0.04	1000	ND	< 0.02	640

[a] Taken from reference 10.
[b] ND, not determined.

Kinetic Parameters—The hydrolysis of antibiotics by HcpB was monitored by following the absorption variation resulting from the opening of the β-lactam ring. Absorption maxima and molar absorption coefficients are given in Table I. Ampicillin, amoxicillin, cefotaxim, cloxacillin, and benzylpenicillin were from Fluka; carbenicillin, cefalotin, cefoxitin, cephaloridin, and oxacillin were from Sigma; and nitrocefin was from Becton Dickinson (Franklin Lakes, New York). All reactions were performed in 20 mM sodium acetate, 150 mM sodium chloride, pH 6.0, at 25 °C on a Cary 300 UV-spectrophotometer. The steady-state rate constants (K_m and k_{cat}) were determined by fitting all data to the Michaelis-Menten equation using the KALEIDOGRAPH software. IC_{50} values were determined by inhibiting nitocefin hydrolysis at substrate and protein concentrations of 200 and 2 μM, respectively. Protein concentration was determined by amino acid analysis.

Crystallization and Data Collection—Crystallization trials using the sitting drop vapor diffusion method of native and selenomethionine-labeled HcpB were set up exactly the same way. Droplets consisted of 2 μl of reservoir buffer and 2 μl of refolded HcpB (4.4 mg/ml protein in 40 mM sodium acetate, 1 mM EDTA, pH 5.5). The droplets were equilibrated against 500 μl of reservoir solution (25% polyethylene glycol 8000, 0.1 M sodium citrate, pH 3.0). Pencil-shaped crystals were obtained within 14 days at 20 °C. They belonged to space group $P6_522$ with unit cell dimensions $a = b = 51.07$ Å, $c = 206.39$ Å, and a Matthew's parameter of 2.40 $Å^3$/Da with one molecule per asymmetric unit.

Single crystals were transferred into a cryo-buffer (25% polyethylene glycol 8000, 0.1 M citrate, 20% ethylene glycol, pH 3.0) and flash-frozen in a stream of liquid nitrogen at a temperature of 110 K. For phasing by multiple wavelength anomalous dispersion, three data sets were collected up to a 2.5-Å resolution from a single crystal at the BM14 beamline (European Synchrotron Radiation Facility, Grenoble). Later, a further high resolution native data set was collected at a 1.95-Å resolution on station ID14–3. Data were scaled and integrated using the DENZO/SCALEPACK package (13). Statistics on data collection and refinement are given in Table II.

Structure Solution and Refinement—The HcpB structure was solved by multiple wavelength anomalous dispersion phasing using the selenium absorption edge. Several dispersive and difference Patterson maps were calculated among the selenomethionine derivative data sets. To improve the signal to noise ratio, the maps were merged, and the selenium site was identified by the automated Patterson search method implemented into the program CNS (14). Heavy atom parameters were refined using the program SHARP (15). Initial phases were calculated using data between 25- and 3.8-Å resolution. Solvent flattening using the program SOLOMON (16) revealed an electron density map that was suitable to build an initial poly-alanine model using the display software O (17). Subsequently, phases were calculated to a 2.5-Å resolution, and side chains became visible, allowing the sequence to be fitted into the electron density. The refinement was performed using the programs CNS and REFMAC (18). The free *R*-factor was calculated with a test set containing 10% of the data. When the 1.95-Å data set became available, refinement was finalized using the program ArpWarp (19). Amino acids Met-1, Val-2, Asn-136, Asn-137, and Tyr-138 as well as the six C-terminal histidine residues were not modeled due to the lack of interpretable electron density. Fold analysis was performed using the Dali internet service (20). Figures within this publication were prepared using the programs MOLSCRIPT (21) and BOBSCRIPT (22). Helix packing angles were calculated using the program INTERHELIX.

RESULTS

The HcpB Structure—The crystal structure analysis of HcpB revealed, in contrast to the sequence-based secondary structure prediction, an essentially α-helical fold. The 133 residues of HcpB fold into eight α-helices that pack into a right-handed superhelix with overall dimensions of 63 × 35 × 25 Å (Fig. 2*a*). Four disulfide bridges are observed between cysteine pairs Cys-22/Cys-30, Cys-52/Cys-60, Cys-88/Cys-96, and Cys-124/Cys-132. The disulfide bridges subdivide the structure into four (*1*, *2*, *3*, *4*) pairs (*A*, *B*) of α-helices confirming the proposed modular architecture. Helices A and B are 14 and 10 residues long, respectively. The two cysteine residues forming a disulfide bridge are located at the C terminus of helix A and four residues behind the N terminus of helix B. However, there are three exceptions. Helix 1A has a three-residue-long α-helical extension at the N terminus, and helix 4B is two residues shorter. In addition, two residues at the N terminus of helix 1B are not in an α-helical conformation. The packing angle of helices A and B belonging to the same α/α-motif (*e.g.* 1A/1B) is 42°, whereas the angle between helices B and A of adjacent motifs (*e.g.* 1B/2A) is 14°. The helix packing creates a fan-like structure with an angle between the first and the last α-helix of 130° (Fig. 2*a*). The convex surface of the molecule is formed by helices 1A, 2A, 3A, and 4A. This surface area is predominately positively charged. On the opposite side of the molecule, helices 1B, 2B, 3B, and 4B create an amphipathic groove. Polar side chains of helix 2B form the bottom of the grove that is flanked on both sides by hydrophobic side chains coming from helices 1B, 3B, and 4B.

The four α-helix pairs possess very similar conformations (Fig. 2*b*). The sequence identity for the pairwise alignments varies between 33 and 58%, and the root mean square deviation (r.m.s.d.) varies between 0.33 and 1.35 Å (Table III). Although the overall sequence composition of motif 1 is similar to motifs 2–4, the conformation of motif 1 is different from motifs 2–4. The r.m.s.d. between motif 1 and motifs 2–4 is well above 1 Å, whereas the r.m.s.d. among motifs 2–4 is much smaller (Table III). The increased r.m.s.d. is due to a different conformation of the loop that connects helices 1A and 1B. In loop 1, the amino acid at position 26 is in the left-handed helix conformation ($\phi/\varphi_{Phe\text{-}28}$ = 70°/3°), whereas the corresponding residues in loops 2–4 are all in right-handed helix conforma-

TABLE II
X-ray data collection and refinement statistics

Data collection	Native	Se λ1 (peak)	Seλ2 (inflection)	Seλ3(remote)
Beamline	ID14–3	BM14	BM14	BM14
λ (Å)	0.9330	0.9791	0.9794	0.8856
Maximum resolution (Å)	1.95	2.5	2.5	2.5
Completeness	99.8%	99.9%	99.9%	99.9%
Average I/σ	12.5	17.0	18.4	12.6
R^{a}_{sym}	5.4% (49.8%)	4.9% (11.8%)	4.9% (10.4%)	6.3% (17%)
Redundancy	9.3	5.4	5.4	5.7
Mean figure of merit[b]		0.606		
Refinement				
Resolution (Å)	30–1.95			
R-factor (%)	18.67			
R_{free} (%)	23.84			
No. of protein atoms	1030			
No. of water molecules	190			
Ramachandran plot				
Most favored (%)	96			
Allowed (%)	4			
Disallowed (%)	0			

[a] Last shell (2.00–1.95 Å) data in parentheses.
[b] Data from 25 to 3.8 Å resolution.

tions ($\phi/\varphi_{Asn\text{-}58} = -56°/-40°$, $\phi/\varphi_{Asp\text{-}94} = -83°/-33°$, and $\phi/\varphi_{Asp\text{-}130} = -54°/-38°$).

The structure-based sequence alignment of the four motifs reveals that the sequence pattern extends beyond the conserved cysteine pairs (Fig. 2*c*). The cysteine residues at positions 20 and 28, alanine at position 19, and glycine at position 27 are conserved for structural reasons. The disulfide bridge fixes helices A and B in a defined orientation and restrains the conformation of the loop. The covalent disulfide bond brings the helices very close together in space. Therefore the side chains of residues preceding the cysteines (*e.g.* alanine at position 19 and glycine at position 27) are at van der Waals distances. Throughout the whole Hcp family, residues preceding the cysteines are always glycine, alanine, or serine residues because these residue types possess sufficiently small side chains. Residues with larger side chains would prevent helices A and B from adopting the proper packing angles. Leucine at position 31 is also conserved because its side chain fits like a knob into a hole on the surface of the preceding helix A. The leucine at position 31 in motif 1 (Leu-33) is completely buried in the center of a hydrophobic core formed by helices 1A, 1B, and 2A, whereas leucine residues in motifs 2–4 (Leu-63, Leu-99, and Leu-135) are solvent accessible. In addition, lysine residues at positions 11 and 18, leucine residues at position 22, and asparagine residues at position 14 are also conserved (Fig. 2, *b* and *c*). Since these amino acids occur in subsequent turns on the solvent-exposed side of helix A, they form rims of identical residues on the convex side of the molecule.

Data base searches revealed that the structure of HcpB is most similar to the tetratricopeptide repeat (TPR) domain of the human protein phosphatase 5 (PP5, Protein Data Bank accession number 1a17) (23). The isolated PP5 TPR repeats superimpose well onto the HcpB structure (Fig. 2*d*). However, the relative orientation of repeats in HcpB and PP5 are different.

Characterization of Folding—Since HcpB was refolded from inclusion bodies, proper refolding was verified by CD spectroscopy. The CD spectrum shown in Fig. 3*a* reveals a pronounced minimum at 222 nm. Based on the CD spectrum, the α-helix content was predicted to be 73%, which is in perfect agreement with the crystal structure. Upon the addition of GdmHCl, the minimum at 222 nm vanishes from the spectrum. By plotting the CD signal at 222 nm over the GdmHCl concentration, the free energy of unfolding and the cooperativity parameter (*m*) were determined from the intercepts and the slopes of the titration curve at the transition phase. From the titration curve shown in Fig. 3*b*, we derived $[GdmHCl]_{½}$ and *m* values of 1.93 ± 0.02 M and 11.24 ± 0.99 kJ/(mol·M), respectively, yielding a free energy of unfolding of −22 kJ/mol. The theoretical cooperativity of unfolding calculated from the amino acid sequence is 12 kJ/(mol·M).

β-Lactam Hydrolysis—It was shown recently that HcpA has β-lactamase and penicillin binding activities (10). Kinetic data summarized in Table I reveal that HcpB possesses similar activities. Generally, 6-aminopenicillinic acid compounds are better substrates or inhibitors than 7-aminocephalosporanic acid derivatives. With the exception of nitrocefin, 6-aminopenicillinic acid derivatives show K_m and IC_{50} values in the micromolar range, whereas the kinetic parameters for 7-aminocephalosporanic acid derivatives are in the millimolar range.

The Binding Site—Attempts to detect the nitrocefin binding site in HcpB failed because the crystals disintegrated upon soaking nitrocefin into the HcpB crystals. However, the crystal color turned dark red, indicating that the nitrocefin β-lactam ring was cleaved by HcpB. Electron density maps calculated between the refined HcpB structure and x-ray diffraction data collected on HcpB co-crystallized with oxacillin (data not shown) revealed significant difference electron density in the amphipathic grove, but the maps were not sufficiently clear to fit oxacillin precisely into the HcpB structure.

Upon refinement of the HcpB crystal structure, we observed strong electron density at the putative penicillin binding site. This density was refined as a cluster of densely packed water molecules as shown in Fig. 4*a*. However, the close distances of water molecules and the continuous electron density suggest that this density might represent a copurified ligand rather than a cluster of isolated water molecules. Mass spectrometric analysis of HcpB revealed two peaks with molecular masses of 16,159.2 and 16,450.8 Da (data not shown). The two peaks account for a mixture of free HcpB and a complex between HcpB and a compound with a molecular weight of ~292 Da. *N*-acetylmuramic acid (NAM) is a compound that is found in the peptidoglycan of all Gram-negative bacteria. NAM has the right molecular size (molecular size = 293.3 Da) and fits the observed electron density as indicated in Fig. 4*a*. The proposed binding site is located in the amphipathic grove close to the N termini of helices 1B, 2B, and 3B (Fig. 4*b*). Modeling NAM into the proposed binding site revealed that NAM would be recognized by a number of hydrogen bonds. Residues that could

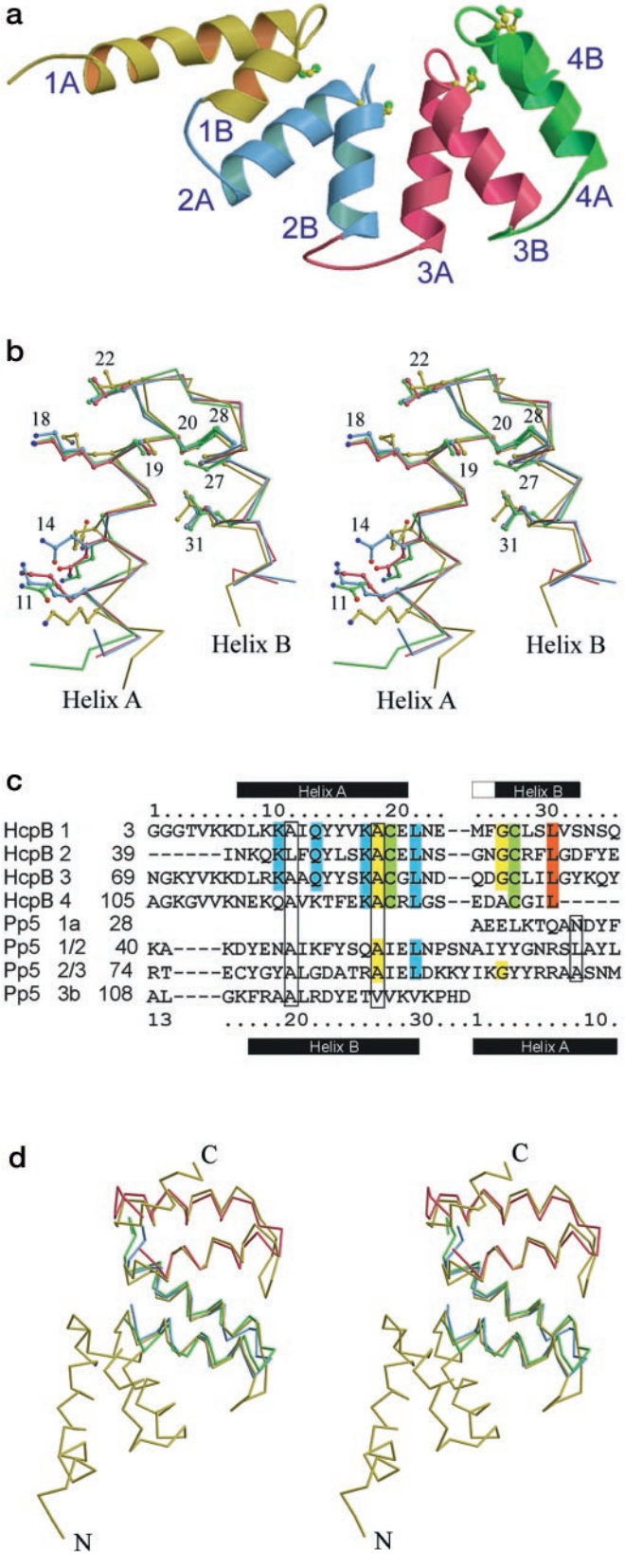

FIG. 2. *a*, ribbon diagram showing also the disulfide bridges in HcpB. The four α/α-motifs are *shaded* and *labeled*. *b*, stereo view of the superposition of four HcpB α/α-motifs. Motifs are *shaded* as in *panel a*. The superposition was calculated based on the residue selection given in Table III. The side chains of amino acids that are conserved in all four motifs are depicted. *Numbering* refers to the position in the motif as indicated in *panel c*. *c*, structure-based sequence alignment of HcpB motifs 1–4 and PP5 TPR repeats 1–3. Residues that are conserved are highlighted: surface residues (*blue*), cysteine residues (*green*), residues with small side chains (*yellow*), and hydrophobic side chains (*red*). Helices and *numbering* on the *top* refer to the HcpB motif. The TPR *numbering* and helix assignment is given at the *bottom*. Small hydrophobic residues that are conserved in TPRs are *boxed*. *d*, stereo view of the Cα traces of TPR repeats 1 (*blue*), 2 (*red*), and 3 (*green*) superimposed individually onto HcpB (*yellow*).

TABLE III
R.m.s.d. and sequence identity among α/α-motifs of HcpB

Rmsd [Å] and sequence identity[a]	M1	M2	M3	M4
M1		46%	58%	31%
M2	1.18		50%	27%
M3	1.11	0.33		35%
M4	1.35	0.66	0.64	

[a] Motifs M1, M2, M3, and M4 refer to residues 10–36, 40–66, 76–102, and 112–135, respectively. Sequences were aligned as shown in Fig. 2c.

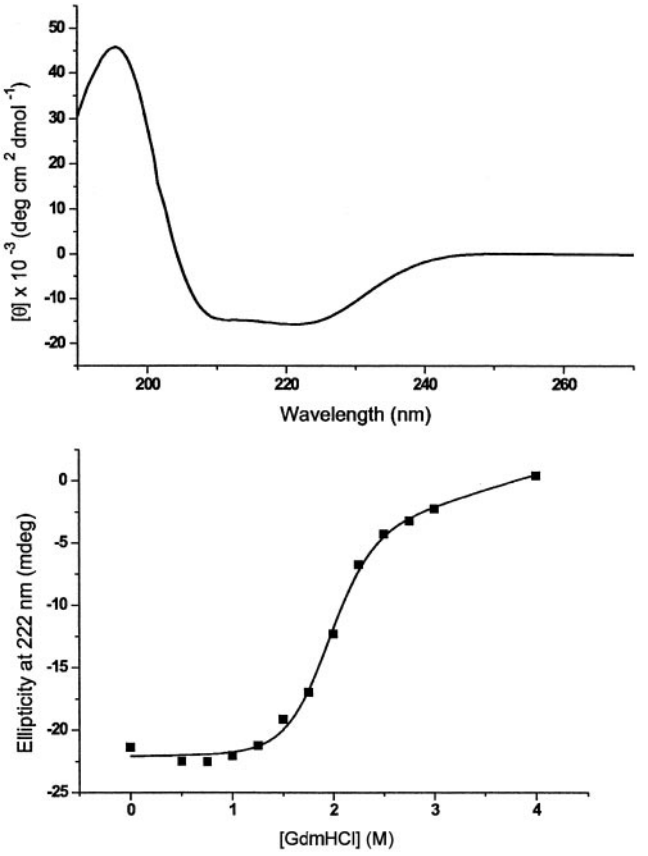

FIG. 3. *a*, CD spectrum of refolded HcpB. *b*, ellipticity at a wavelength of 222 nm as a function of GdmHCl concentration. *mdeg*, millidegrees.

interact with the putative ligand are Asn-58, Asp-92, Asp-94, and Ser-128.

DISCUSSION

The conservation of the sequence pattern among the Hcp family suggests that all family members are composed of the same α/α-motif. This motif is similar to the TPR repeat, although there are substantial differences. As the name implies, TPR proteins consist of repeats of 34 amino acids that fold into two α-helices and are frequently found in multidomain proteins where they serve as protein/protein interaction modules. The TPR sequences are very versatile, and there is no position characterized by an invariant residue. Small hydrophobic residues are observed at positions 8, 20, and 27 of the TPR motif. The sequence alignment deduced from the superposition of the HcpB motifs onto the three TPRs of PP5 reveals that this

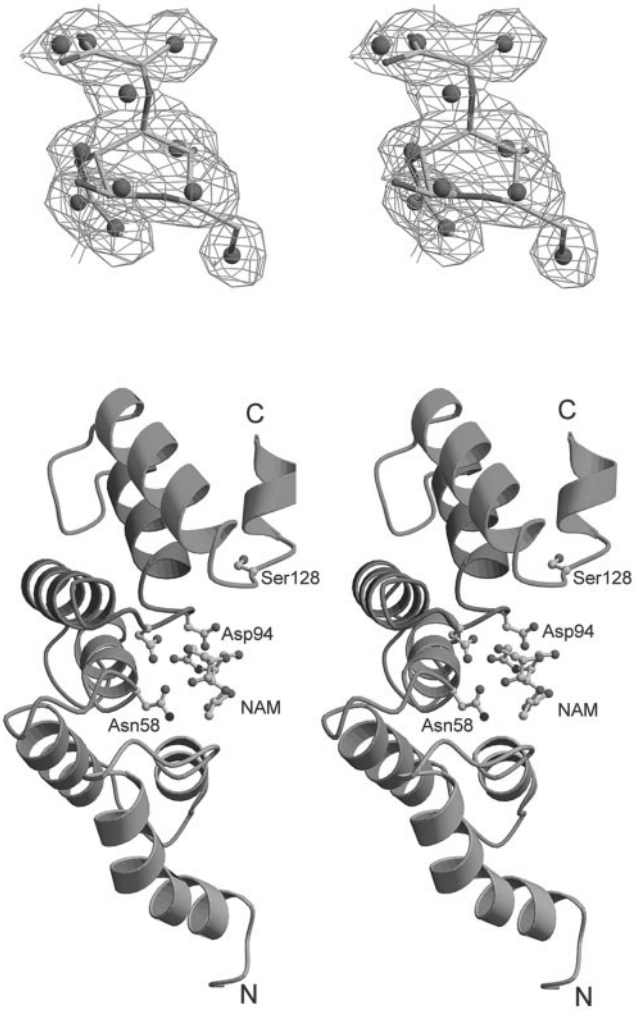

FIG. 4. *a*, water molecules and 2Fo-Fc electron density (contour level of 1.3 σ) in the putative ligand binding site. The density is explained by 11 water molecules. *N*-acetylmuramic acid was modeled into the electron density of the water molecule cluster. *b*, modeled *N*-acetylmuramic acid/HcpB complex. The ligand could form hydrogen bonds with residues in the loops between helices A and B of motifs 1, −2, and −3.

pattern is partially conserved in the HcpB structure (Fig. 2*c*). The alanine/leucine residues at positions 12 and 19 of the HcpB motif superimpose onto the alanine/valine residues at positions 20 and 27 of the TPR. In addition, leucine at position 22 is also conserved in TPR repeats 1 and 2, whereas the lysine and asparagine residues at positions 11, 14, and 18 that are located on the convex surface of HcpB are not. There might be a functional requirement for the conservation of these amino acids, particularly if HcpB interacts with proteins that also show a modular architecture. On the other hand, the conservation might be a remnant from the duplication of an ancestral α/α-motif sequence. Since many residues that participate in helix packing are not conserved among HcpB motifs, it seems that there is a selective pressure for the conservation of these surface residues.

However, the TPR and Hcp repeats consist of 34 and 36 amino acids, respectively. In HcpB, four amino acids are inserted between helices A and B of the TPR. The loop between cysteine pairs that is conserved throughout the Hcp family is two amino acids shorter in TPR proteins (Fig. 2, *c* and *d*). In HcpB, the inter- and intra-repeat helix packing angles are 42° and 14°, respectively, whereas in TPR proteins, these angles are always ~24°. Therefore the Hcp and TPR folds are related because they consist of similar pairs of antiparallel α-helices. However, the loops connecting the helices and the helix packing angles are different.

The biological functions of TPR proteins are very diverse. Many TPR proteins are involved in regulation of the cell cycle, in protein transport, and in chaperone-assisted protein folding (24), which makes it impossible to assign a possible biological function to HcpB based on the overall structural topology alone. Most members of the Hcp family have only been recognized on the genome level. *In vivo* expression was shown for the gene products of HP0211 and HP0160. HP0211 messenger RNA (designated orf2 in the literature (25)) was detected by slot-blot analysis, and the gene product (HcpA) was recognized in *H. pylori* culture broth supernatant, verifying that this gene was expressed and secreted into the medium (9). In another study, the HP0160 gene product was identified in *H. pylori* membrane fractions (26).

It was shown that HcpA had a moderate β-lactamase activity (10) and the HP0160 gene product (PBP4) is capable of binding penicillin derivatives (26). HcpB possesses a penicillin binding activity like other Hcp family members. The substrate profile shows that HcpB must be regarded as a penicillinase because most 7-aminocephalosporanic acid derivatives are neither good substrates nor tight binding inhibitors. The substrate profiles of HcpA and -B are similar, but there are also subtle differences that distinguish these family members. The K_m value for amoxicillin hydrolysis by HcpB is three times smaller than for HcpA, whereas for benzylpenicillin, this relationship is inverted. The turnover rates for β-lactam hydrolysis by HcpA and -B are 5 orders of magnitude lower than for typical β-lactamases such as the *Bacillus licheniformis* β-lactamase, but they are still 4 orders of magnitude higher than for typical penicillin-binding proteins (PBP) such as the *Streptomyces R61* DD-peptidase (27).

A possible explanation why the turnover rates are much lower than for known β-lactamases might be that the true activities are substantially higher, but only a small fraction of HcpB refolded into an active conformation. Although this hypothesis can ultimately be tested only by the analysis of HcpB that has been isolated from natural sources, there is little evidence to support this idea. If the measured β-lactamase activities would be exerted by a small fraction of correctly refolded protein, there should be considerable batch-to-batch variation, and the k_{cat} values of HcpA and -B should differ significantly. However, the measured turnover rates for HcpA and -B are very similar, and the k_{cat} error is just 0.1 min^{-1}. The fact that two different proteins that have been refolded under different conditions possess very similar k_{cat} values makes it unlikely that the natural activities are substantially higher than the measured activities. On the other hand, the true activities might just be slightly bigger than the measured activities, and the protein preparation may consist of equally sized fractions of active and inactive protein conformations. In this case, one would expect that the GdmHCl titration curve would show a multiphase transition, which is not the case. In fact the measured cooperativity of unfolding agrees very well with the theoretical value.

The kinetic data given in Table I characterize HcpA and -B as intermediates between classical PBPs and true β-lactamases. Although β-lactamases and PBP have evolved from a common ancestor, which is indicated by similar active site topologies and three-dimensional folds, their biological functions are different (27). Classical PBPs are involved in peptidoglycan biosynthesis where they catalyze the glycan chain elongation and cross-linking. Therefore high molecular weight PBPs are bifunctional. They contain a D-Ala-D-Ala-specific transpeptidase activity that can be inhibited by β-lactam antibiotics and

a transglycosylase activity (28). In contrast to PBPs, β-lactamases have evolved to combat treatment with β-lactam antibiotics. β-lactamases are very potent enzymes that rapidly hydrolyze β-lactam antibiotics to prevent inhibition of PBPs. None of the members of the Hcp family possess significant sequence or structural similarity with the currently known β-lactamases or PBPs, and the well known sequence motif that is ubiquitously found in active site serine PBPs (28) is also absent throughout the Hcp family. Therefore the *in vivo* functions of HcpA and -B still remain unclear. Due to their moderate turnover rates, it is unlikely that these enzymes confer significant resistance against antibiotics by β-lactam hydrolysis, which is also supported by the observation that the *H. pylori* strain 26695 is still sensitive to amoxicillin (29).

HcpA and -B could also be involved in the biosynthesis of peptidoglycan, which is supported by the chemical similarity between penicillins and the D-Ala-D-Ala dipeptide that is cleaved upon cross-linking of adjacent glycan strains. PBPs that catalyze this reaction play a crucial role in maintaining the cellular morphology (30). *H. pylori* possesses a characteristic spiral-shaped morphology, suggesting that the biosynthesis of the *H. pylori* peptidoglycan has some unique features (31). Genome analysis revealed that there are three PBP homologues but only one of them has a proposed transglycosylation activity (32). Since no additional monofunctional transglycosylases were found by sequence comparison methods, there is just one enzyme that can catalyze the glycan chain elongation and recycling of proteoglycan fragments. The spirulate morphology of *H. pylori* might either be attributed to the lack of transglycosylases or to alternative enzymes that participate in proteoglycan metabolism. Perhaps proteins from the Hcp family are responsible for the spiral-shaped morphology of *H. pylori*.

The HcpB crystal structure is the prototype structure for a protein family that is restricted to the *Helicobacter* and *Campylobacter* genera. So far, all investigated proteins from this family possess penicillin binding activities. However, the biological functions are still unclear. Therefore further *in vivo* experiments to elucidate the biological functions of these proteins in more detail are needed.

Acknowledgments—We thank Ragna Sack and Drs. Peter Gehrig and Peter Hunziker for mass spectrometric analysis. The support of data collection of Drs. Germaine Sainz and Gordon Leonard on station BM14 and Drs. Julien Lescar and Ed Mitchel on station ID14–3 (European Synchrotron Radiation Facility, Grenoble) is also gratefully acknowledged.

REFERENCES

1. Warren, J. R., and Marshal, B. (1983) *Lancet* **1,** 1273–1275
2. Blaser, M. J., Perez-Perez, G. I., Kleanthous, H., Cover, T. L., Peek, R. M., Chyou, P. H., Stemmermann, G. N., and Nomura, A. (1995) *Cancer Res.* **55,** 2111–2115
3. Graham, D. Y., and Yamaoka, Y. (1998) *Helicobacter* **3,** 145–151
4. McGee, D. J., and Mobley, H. L. (1999) *Curr. Top. Microbiol. Immunol.* **241,** 155–180
5. Covacci, A., Telford, J. L., Del Giudice, G., Parsonnet, J., and Rappuoli, R. (1999) *Science* **284,** 1328–1333
6. Powell, J., and McConkey, C. C. (1992) *Eur. J. of Cancer Prev.* **1,** 265–269
7. Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., and Venter, J. C. (1997) *Nature* **388,** 539–547
8. Alm, R. A., Ling, L. S., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., Smith, D. R., Noonan, B., Guild, B. C., deJonge, B. L., Carmel, G., Tummino, P. J., Caruso, A., Uria-Nickelsen, M., Mills, D. M., Ives, C., Gibson, R., Merberg, D., Mills, S. D., Jiang, Q., Taylor, D. E., Vovis, G. F., and Trust, T. J. (1999) *Nature* **397,** 176–180
9. Cao, P., McClain, M. S., Forsyth, M. H., and Cover, T. L. (1998) *Infect. Immun.* **66,** 2984–2986
10. Mittl, P. R. E., Lüthy, L., Hunziker, P., and Grütter, M. G. (2000) *J. Biol. Chem.* **275,** 17693–17699
11. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry* **27,** 8063–8068
12. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci.* **4,** 2138–2148
13. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
14. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** D905–D921
15. Brodersen, D. E., de La Fortelle, E., Vonrhein, C., Bricogne, G., Nyborg, J., and Kjeldgaard, M. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** D431–D441
16. Abrahams, J. P., and Leslie, A. G. W. (1996) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **52,** D40–D42
17. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr. Sect. A* **47,** A110–A119
18. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** D247–D255
19. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
20. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
21. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
22. Esnouf, R. M. (1997) *J. Mol. Graph. Model* **15,** 132–134
23. Das, A. K., Cohen, P. W., and Barford, D. (1998) *EMBO J.* **17,** 1192–1199
24. Blatch, G. L., and Lassle, M. (1999) *Bioessays* **21,** 932–939
25. Karita, M., Etterbeek, M. L., Forsyth, M. H., Tummuru, M. K., and Blaser, M. J. (1997) *Infect Immun.* **65,** 4158–4164
26. Krishnamurthy, P., Parlow, M. H., Schneider, J., Burroughs, S., Wickland, C., Vakil, N. B., Dunn, B. E., and Phadnis, S. H. (1999) *J. Bacteriol.* **181,** 5107–5110
27. Kelly, J. A., Dideberg, O., Charlier, P., Wery, J. P., Libert, M., Moews, P. C., Knox, J. R., Duez, C., Fraipont, C., and Joris, B. (1986) *Science* **231,** 1429–1431
28. Massova, I., and Mobashery, S. (1998) *Antimicrob. Agents Chemother.* **42,** 1–17
29. Simala-Grant, J. L., Zopf, D., and Taylor, D. E. (2001) *J. Antimicrob. Chemother.* **47,** 555–563
30. Nelson, D. E., and Young, K. D. (2001) *J. Bacteriol.* **183,** 3055–3064
31. Costa, K., Bacher, G., Allmaier, G., Dominguez-Bello, M. G., Engstrand, L., Falk, P., de Pedro, M. A., and Garcia-del Portillo, F. (1999) *J. Bacteriol.* **181,** 3710–3715
32. Doig, P., de Jonge, B. L., Alm, R. A., Brown, E. D., Uria-Nickelsen, M., Noonan, B., Mills, S. D., Tummino, P., Carmel, G., Guild, B. C., Moir, D. T., Vovis, G. F., and Trust, T. J. (1999) *Microbiol. Mol. Biol. Rev.* **63,** 675–707
33. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) *Nucleic Acids Res.* **25,** 4876–4882
34. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Int. J. Neural Syst.* **8,** 581–599
35. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, S., Jagels, K., Karlyshev, A. V., Moule, S., Pallen, M. J., Penn, C. W., Quail, M. A., Rajandream, M. A., Rutherford, K. M., van Vliet, A. H., Whitehead, S., and Barrell, B. G. (2000) *Nature* **403,** 665–668