NIH Public Access
Author Manuscript
Proteins. Author manuscript; available in PMC 2009 October 1.

Published in final edited form as:
Proteins. 2008 October ; 73(1): 265–268. doi:10.1002/prot.22159.

Structure of HP0564 from *Helicobacter pylori* identifies it as a new transcriptional regulator

Brendan N. Borin and **Andrzej M. Krezel**
Department of Biological Sciences, Center for Structural Biology, Vanderbilt University, Nashville, TN

Keywords
NMR; repressor; structural genomics; DNA-binding; ribbon-helix-helix; ulcer; gastric cancer

Introduction

Helicobacter pylori (H. pylori) is a gram-negative, pathogenic bacterium that infects half the world's population and is responsible for the majority of cases of gastric and duodenal ulcers1. Uniquely adapted to survive the low pH conditions, it is the only organism that can establish a permanent infection of the human stomach. In the most severe cases, long-term infection can lead to gastric cancer.

With approximately 1,500 genes, *H. pylori* has a relatively small genome and very few transcriptional regulators described to date2›3. This has been explained by the fact that it has only one environmental niche and must respond to only a small number of stimuli compared to free living bacteria. It has been reported that the number of transcriptional regulators in bacteria increases proportionally to the square of the number of genes in the genome4. Nevertheless, our understanding of the regulation of metabolic processes and environmental responses in *H. pylori* is far from complete.

Here, we report the solution structure of HP0564 (JHP0511 in the sequenced strain J993), a protein with no assigned function. Although it has no sequence homologs outside of *H. pylori*, our structural analysis indicates that it is a member of the ribbon-helix-helix superfamily (Pfam protein domain family PF01402) of transcriptional regulators. These proteins bind to specific DNA sequences with high affinity and usually act as repressors.

Materials and Methods

Protein Expression and Purification

For structural work, a shortened construct of HP0564 was created that lacks the flexible N-terminal 20 residues as well as the C-terminal 7 residues, leaving only the stably folded region. The corresponding sequence was PCR amplified from genomic DNA of *H. pylori* strain J99 and cloned into a modified pET vector with an N-terminal, 12 residue His_6 tag (MRGSHHHHHHGS). Transformed *Escherichia coli* BL21 (DE3) cells were grown in LB media to the OD_{600}=1 and induced with 0.4 mM IPTG for 3 h. Cells were spun down, resuspended in binding buffer (20 mM Tris, 0.5 M NaCl, 5 mM imidazole, 8 M urea, pH 7.9), and disrupted by sonication (6×30 seconds). Filtered (1 μm) cell extract was loaded on

Corresponding author: Andrzej M. Krezel, Department of Biological Sciences, Vanderbilt University, 465 21st Avenue S, MRBIII, 5270A, (615) 322-2073, email: andrzej.m.krezel@vanderbilt.edu, phone: 615 322 2073, fax: 615 343 6707 .

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

a Ni-NTA column, followed by a 100 mL wash (20 mM Tris, 0.5 M NaCl, 30 mM imidazole, 8 M urea, pH 7.9) and elution (20 mM Tris, 0.5 M NaCl, 0.1 M EDTA, 8 M urea, pH 7.9). Refolding was achieved by dialysis against distilled water. No additional protein bands could be detected by tricine, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Isotope-labeled samples were prepared by growing cells in M9 minimal media supplemented with $^{15}NH_4Cl$ and/or ^{13}C-u-glucose (CIL, Andover, MA). All other aspects of the expression and purification of labeled samples were identical to those used for natural abundance protein.

Crosslinking Experiments

Crosslinking experiments were performed with BS3 (Pierce, Rockford, IL). Reaction buffer was 20 mM NaH_2PO_4, pH 7.0. Crosslinker was dissolved in reaction buffer to 10 mM stock concentration immediately prior to setting up reactions. All reactions were in 20 μL, consisting of 17 μL reaction buffer, 2 μL HP0564 (10 μg/μL), and 1 μL of an appropriate dilution of BS3. Final concentrations of 0, 0.005, 0.05, and 0.5 mM BS3 were used. Reactions were allowed to proceed for 3 minutes before being quenched with 5 μL of 1 M Tris, pH 7.5. All reactions were run on a 10% SDS-PAGE gel and stained with Coomassie Blue.

Gel Filtration Experiments

Size exclusion separations were performed on a Superdex 75 10/30 FPLC column (Pharmacia, Piscataway, NJ) at 4 °C in 50 mM KH_2PO_4, pH 4.0. Elution was followed by UV absorption at 214 nm. The calibration curve used to calculate the molecular weight was prepared with ubiquitin, thioredoxin, and ovalbumin run under identical conditions.

NMR Experiments

NMR experiments were performed on Bruker Avance 600 and 800 MHz spectrometers at 25 °C. Samples were prepared at 1 mM monomer concentration in 50 mM KH_2PO_4, pH 4.0. Natural abundance protein was used to acquire ^1H 2D NOESY spectra using mixing times of 25, 50, and 100 ms. Singly-labeled ^{15}N and ^{13}C samples were used to acquire 2D HSQCs. Doubly-labeled $^{15}N,^{13}C$ samples were used to acquire 3D HNCO, CBCANH, CCCONH, and HCCCONH experiments used for backbone and sidechain assignments5. Over 98% of backbone resonances (H^N, N, C^α, H^α, C') and 85% of commonly assignable carbon and proton sidechain resonances were assigned.

Structure Calculations

NMR data were processed using XWINNMR (Bruker, Billerica, MA) and analyzed using SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). Structure calculations were performed using CYANA6 version 2.1 with 25,000 steps for each structure. NOE crosspeaks corresponding to both intramolecular and intermolecular interactions were assigned manually, and intensities were automatically converted to distance restraints using built-in CYANA routines. Given the small size of the protein (7.8 kDa monomer), the 2D NOESY was sufficiently resolved to assign all crosspeaks. 3D heteronuclear-resolved NOESY spectra were recorded, but offered no additional distance information and were not used in structure calculations. In the initial stages of calculations, only NOE-derived restraints were used. Hydrogen bond restraints were added in later stages when they could be identified in a majority of calculated structures. In the last stage, out of 1,000 initial structures, the 50 with the lowest target function values were minimized in AMBER7 version 9 using 10,000 steps of conjugate

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

gradient energy minimization (Table I). Of these 50 energy-minimized structures, the 20 with the lowest nonbonded backbone energies were used in the final ensemble, which was analyzed using AQUA and PROCHECK-NMR 8. The PDB entry, including the structural ensemble as well as the restraints used in structure calculations, has the PDB accession code 2k1o. BMRB entry 15761 contains ^{1}H, ^{13}C, and ^{15}N chemical shift assignments. Chimera9 was used for interactive analysis and figure production.

Results and Discussion

A Genbank search with the DNA or protein sequence of *Helicobacter pylori* HP0564 (Uniprot Q9ZLR7_HELPJ) yields no orthologs and only one paralog, HP022210 (Uniprot Q9ZML0_HELPJ), which has 17 identical residues out of 40 in the stably folded region consisting of residues 23-62 (Fig. 1b). The structure of HP0564 shows it to be a member of the ribbon-helix-helix (RHH) superfamily of transcriptional regulators (Fig. 1a). A DALI11 search yielded the Arc repressor (PDB accession code 1baz), CopG (PDB accession code 1ea4), and HP0222 (PDB accession code 1×93) as its closest structural relatives, all with Z-scores greater than 5.0. These proteins are always found in solution as dimers12. Dimerization creates an antiparallel double-stranded β-sheet with several sidechains exposed to solvent that are used in making sequence-specific contacts with DNA13. Upon binding DNA, proteins in this superfamily form tetramers or higher order oligomers, where each dimer binds several base pairs of DNA.

Chemical crosslinking was performed to confirm that HP0564 could form dimers (Supplementary Figure S1). The amount of dimer and species corresponding to higher-order oligomers increased with increasing BS3 concentration. Without crosslinking, traces of noncovalent dimers were present on SDS-PAGE gels. In experiments with HP022210, we did not observe the higher-order, cross-linked forms. Gel filtration experiments showed only stable dimers in solution (Supplementary Figure S2).

The DNA binding residues are not conserved between the β-sheets of HP0222 and HP0564. The presence of intact *HP0564* does not complement deleted *HP0222*. HP0222 null mutants are viable, but show significantly slower growth than parent wild-type strains. This suggests that they are not functionally redundant and will bind different DNA sequences and regulate different genes. Structurally, the two proteins are very similar, with a backbone RMSD of 1.24 Å (Fig. 1c). Superimposing HP0222 and HP0564, one can see that the β-sheet in HP0564 packs more closely to the helices than in HP0222, possibly due to the less bulky valine at position 53 compared to isoleucine in HP0222. Although there are no absolutely conserved amino acids in the RHH family14, the HP0564 sequence agrees with the sequence motifs featured in all RHH proteins, including the alternating hydrophilic and hydrophobic residues within the β-sheet and the hydrophobic core residues - F24, V26, F28 from the β-sheet, L38 from α-helix 1, and V53, I57, I61 from α-helix 2. All of these residues are involved in making contacts with residues from the other subunit in the dimer.

Because so few transcriptional regulators have been identified in *Helicobacter*, it is exciting to discover a new one. We expect HP0564 to play an important role in transcriptional regulation. We are working on determining its cognate DNA-binding sequence and its function in the cell.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Acknowledgments

This work was supported by NIH grant AI67628 to AMK. BNB was supported by NIH Molecular Biophysics Training Program grant GM 008320.

References

1. Marshall BJ, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancet. 1984; 1(8390):1311–1315. [PubMed: 6145023]
2. Tomb JF, White O, Kerlavage AR, Clayton RA, Sutton GG, Fleischmann RD, Ketchum KA, Klenk HP, Gill S, Dougherty BA, Nelson K, Quackenbush J, Zhou L, Kirkness EF, Peterson S, Loftus B, Richardson D, Dodson R, Khalak HG, Glodek A, McKenney K, Fitzegerald LM, Lee N, Adams MD, Hickey EK, Berg DE, Gocayne JD, Utterback TR, Peterson JD, Kelley JM, Cotton MD, Weidman JM, Fujii C, Bowman C, Watthey L, Wallin E, Hayes WS, Borodovsky M, Karp PD, Smith HO, Fraser CM, Venter JC. The complete genome sequence of the gastric pathogen Helicobacter pylori. Nature. 1997; 388(6642):539–547. [PubMed: 9252185]
3. Alm RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, Smith DR, Noonan B, Guild BC, deJonge BL, Carmel G, Tummino PJ, Caruso A, Uria-Nickelsen M, Mills DM, Ives C, Gibson R, Merberg D, Mills SD, Jiang Q, Taylor DE, Vovis GF, Trust TJ. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. Nature. 1999; 397(6715): 176–180. [PubMed: 9923682]
4. van Nimwegen E. Scaling laws in the functional content of genomes. Trends Genet. 2003; 19(9): 479–484. [PubMed: 12957540]
5. Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progress in Nuclear Magnetic Resonance Spectroscopy. 1999; 34:93–158.
6. Guntert P, Mumenthaler C, Wuthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. Journal of molecular biology. 1997; 273(1):283–298. [PubMed: 9367762]
7. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr. Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. Journal of computational chemistry. 2005; 26(16):1668–1688. [PubMed: 16200636]
8. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. Journal of biomolecular NMR. 1996; 8(4):477–486. [PubMed: 9008363]
9. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004; 25(13):1605–1612. [PubMed: 15264254]
10. Popescu A, Karpay A, Israel DA, Peek RM Jr. Krezel AM. Helicobacter pylori protein HP0222 belongs to Arc/MetJ family of transcriptional regulators. Proteins. 2005; 59(2):303–311. [PubMed: 15723352]
11. Holm L, Sander C. Mapping the protein universe. Science (New York, NY. 1996; 273(5275):595–603.
12. Breg JN, van Opheusden JH, Burgering MJ, Boelens R, Kaptein R. Structure of Arc repressor in solution: evidence for a family of beta-sheet DNA-binding proteins. Nature. 1990; 346(6284):586–589. [PubMed: 2377232]
13. Raumann BE, Rould MA, Pabo CO, Sauer RT. DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. Nature. 1994; 367(6465):754–757. [PubMed: 8107872]
14. Schreiter ER, Drennan CL. Ribbon-helix-helix transcription factors: variations on a theme. Nature reviews. 2007; 5(9):710–720.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

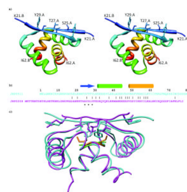

Figure 1.

a) Stereo view ribbon diagram of HP0564 showing residues 21-62 of each subunit. Side chain heavy atoms of β-sheet residues S25, T27, and Y29 that make up the DNA binding interface of ribbon-helix-helix proteins are labeled for one subunit. b) Sequence alignment of JHP0511 (HP0564) and JHP0208 (HP0222) from the J99 strain of *H. pylori.* Numbering is according to the JHP0511 sequence. Secondary structural elements are indicated for HP0564, with an arrow representing the β-strand and cylinders representing the α-helices. DNA-binding residues from the β-sheet are indicated by asterisks. c) Superposition of HP0564 (cyan) on HP0222 (magenta). The β-sheet of HP0564 packs more closely to the α-helices than in HP0222, possibly due to its less bulky valine (green) at position 53 compared to the isoleucine (orange) of HP0222.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table I

Structural statistics for ensemble of 20 structures of HP0564 (JHP0511)

NOE restraints	797
Intraresidue	222
Short	220
Medium	134
Long	221
Intramolecular	626
Intermolecular	171
Hydrogen bonds per dimer	44
Average CYANA target function	0.11
Number of violations > 0.2 Å	0
Average AMBER energies (± standard deviation)	
Input structures	-3894 (± 146)
Energy minimized structures	-5123 (± 13)
Average Ramachandran statistics from PROCHECK (residues 23-62)	
Most favored (%)	88.8
Additionally allowed (%)	10.9
Generously allowed (%)	0.2
Disallowed (%)	0
Average RMSD from mean structure (Å, residues 23-62)	
Backbone (N,C^{α},C',O)	0.59
Heavy atoms	1.08