Nitrosative stress sensing in *Porphyromonas gingivalis*: structure of and heme binding by the transcriptional regulator HcpR

B. Ross Belvin,^{a,b} Faik N. Musayev,^{c,d} John Burgner,^b J. Neel Scarsdale,^{a,d} Carlos R. Escalante^e and Janina P. Lewis^{a,b,f,*}

Received 4 October 2018

Accepted 21 February 2019

Edited by C. S. Bond, University of Western Australia, Crawley, Australia

Keywords: HcpR; *Porphyromonas gingivalis*; microbiology; transcriptional regulators; nitric oxide; nitrosative stress; heme proteins; anaerobes.

PDB references: sensor domain of HcpR from *Porphyromonas gingivalis*, P4₁22 form, 5v30; C22₁ form, 6np6

Supporting information: this article has supporting information at journals.iucr.org/d

^aDepartment of Biochemistry and Molecular Biology, School of Medicine, Virginia Commonwealth University, Richmond, VA 23298, USA, ^bThe Philips Institute for Oral Health Research, School of Dentistry, Virginia Commonwealth University, Richmond, VA 23298, USA, ^cDepartment of Medicinal Chemistry, School of Pharmacy, Virginia Commonwealth University, Richmond, VA 23298, USA, ^dThe Institute for Structural Biology, Drug Discovery, and Development, Virginia Commonwealth University, Richmond, VA 23298, USA, ^eDepartment of Physiology and Biophysics, Virginia Commonwealth University, Richmond, VA 23298, USA, and ^fDepartment of Microbiology and Immunology, School of Medicine, Virginia Commonwealth University, Richmond, VA 23298, USA. *Correspondence e-mail: jplewis@vcu.edu

Although the HcpR regulator plays a vital step in initiation of the nitrosative stress response in many Gram-negative anaerobic bacteria, the molecular mechanisms that it uses to mediate gas sensing are not well understood. Here, a 2.6 Å resolution crystal structure of the N-terminal sensing domain of the anaerobic periodontopathogen *Porphyromonas gingivalis* HcpR is presented. The protein has classical features of the regulators belonging to the FNR-CRP family and contains a hydrophobic pocket in its N-terminal sensing domain. It is shown that heme bound to HcpR exhibits heme iron as a hexacoordinate system in the absence of nitric oxide (NO) and that upon nitrosylation it transitions to a pentacoordinate system. Finally, small-angle X-ray scattering experiments on full-length HcpR reveal that the C-terminal DNA-binding domain of HcpR has a high degree of interdomain flexibility.

1. Introduction

HcpR, which is a member of the fumarate–nitrate regulator/cyclic AMP receptor protein (FNR-CRP) family of regulators, is found in many anaerobic bacteria such as *Desulfovibrio*, *Thermatoga*, *Prevotella* and *Porphyromonas*, where it regulates genes coding for proteins mediating the nitrosative stress response. In the periodontal pathogen *Porphyromonas gingivalis*, HcpR regulates the transcription of the hybrid cluster protein (Hcp; also known as prismane protein), a redox enzyme that plays a role in the response of bacteria to nitrosative stress as a putative high-affinity nitric oxide reductase (Rodionov *et al.*, 2004, 2005; Cadby *et al.*, 2011; Boutrín *et al.*, 2012). Transcriptomic and gene-regulation studies have shown that *hcp* is upregulated in response to NO and nitrite at the transcript level and that this upregulation is dependent on HcpR (Lewis *et al.*, 2012). A strain deficient in the *hcpR* gene (PG1053) does not grow in the presence of physiological concentrations of nitrite (1–2 mM) and NO (0.7 µM *S*-nitroso-glutathione) and shows decreased survival with host cells (Lewis *et al.*, 2012). A transposon-insertion library also identified an HcpR-deficient strain as having a reduced ability to colonize host cells as well as survive in mice, thus indicating that the regulator plays an important role in the virulence of the bacterium and its survival in the host (Miller *et al.*, 2017). Furthermore, it is revealed that *P. gingivalis* HcpR requires heme to bind to DNA of the *hcp* promoter, implicating heme

as a potential cofactor (Lewis *et al.*, 2012). Such data justified further characterization of the molecular mechanism of activation of HcpR that will not only give insight into our ability to interfere with the function of the regulator but will also shed light on the mechanisms of other heme-based regulators belonging to the FNR-CRP family.

FNR-CRP regulators are a class of bacterial transcriptional regulators that play essential roles in metabolism, stress response and virulence (Körner *et al.*, 2003). In all cases, these proteins are capable of sensing an intracellular stimulus and inducing transcription of the appropriate response. Despite the diversity of stimuli and function in bacteria, all members of the FNR-CRP family have a similar structural composition: (i) they contain a β -barrel fold in the N-terminal sensing domain, (ii) they form homodimers through the use of a long helical domain and (iii) they are capable of binding to DNA using a C-terminal helix–turn–helix domain (Schultz *et al.*, 1991; Kuchinskas *et al.*, 2006; Giardina *et al.*, 2008). In general, the binding of a stimulant leads to a conformational change in the protein that results in an increase in affinity to the promoter DNA sites. Beyond these shared attributes, there are significant differences in sequence identity between the members, indicating potential differences in the allosteric mechanisms of activation. Most variation between members of the family occurs in the N-terminal sensing domain, where changes in the sequence accommodate different stimulants or the binding of cofactors.

The gas-sensing members of the FNR-CRP family sense changes in the environment through their cofactors, which are either iron–sulfur clusters (FNR) or heme (CooA/DNR) (Körner *et al.*, 2003; Kiley & Beinert, 2003). In heme-based gas sensors, the heme iron complex functions as the binding site for gaseous molecules such as NO, CO and O₂. Although not all HcpRs utilize heme as a cofactor, a growing subset have been characterized as heme-binding proteins (Cadby *et al.*, 2017). The hypothesis that HcpR may use the heme cofactor is supported by studies that have been performed with other sensor regulators and homologs of HcpR: DNR (NO sensing in *P. aeruginosa*) and CooA (CO sensing in *Rhodospirillum rubrum*) (Liebl *et al.*, 2013). CooA utilizes the heme cofactor to bind to CO and to induce transcription at operons coding for genes necessary for the oxidation of CO to CO₂ (Lobato *et al.*, 2014). Although it is a CO sensor, CooA has been shown to bind to NO under saturating conditions (Reynolds *et al.*, 2000). DNR is a heme-binding transcriptional regulator that senses and binds NO, controlling the expression of denitrification gene clusters (Arai *et al.*, 1995; Giardina *et al.*, 2008). The full-length and sensing-domain structures of both CooA and DNR in the apo form and the heme-bound form of CooA have been determined by X-ray crystallography (Giardina *et al.*, 2008, 2009; Lanzilotta *et al.*, 2000; Kuchinskas *et al.*, 2006). Both proteins form a hexacoordinate heme system utilizing two axial bonds from the protein (Vogel, Spiro *et al.*, 1999; Giardina *et al.*, 2008). Of the diatomic gas molecules that are most relevant to biology (NO, CO and O₂), NO has the highest affinity for heme, thus making it a prime scheme for the selective detection of NO. However, there are significant

differences in the sequences of the known structures of bacterial heme-binding sensor proteins (CooA and DNR): the residues involved in the coordination of the heme iron are not conserved and the location of the heme-binding site in DNR (the closest relative of HcpR) is not known. Furthermore, the most well studied heme-based gas sensors are derived from facultative anaerobic organisms; HcpR is found in obligate anaerobes that grow under low oxygenated conditions. This creates a very different oxidative environment and could play a significant role in the mechanism of protein detection and activation.

In this study, we aimed to shed light on the molecular mechanisms by which *P. gingivalis* HcpR is capable of sensing changes in the levels of reactive nitrogen species in the environment and how this mechanism differs or agrees with those of known heme-sensor proteins. Towards this goal, we have solved the structure of the N-terminal sensing domain of HcpR and characterized the heme- and NO-binding properties of the protein. As HcpR plays an important regulatory step in the nitrosative stress response in many Gram-negative anaerobes, our findings shed light on the understanding of the mechanism not only in the periodontal pathogen *P. gingivalis* but also in other anaerobic bacteria. Furthermore, this work contributes to our knowledge of the mechanism of action of the growing family of FNR-CRP heme-based regulators that are present in a variety of organisms.

2. Materials and methods

2.1. Generation of recombinant proteins

The *hcpR* gene (PG1053) was cloned into a modified pET-21d (m-pET-21d) vector (a TEV cleavage site was added downstream of the N-terminal 6×His tag for removal via TEV protease digestion). Using gene-specific primers, BamHI and XhoI sites were added to the 5' and 3' ends of the gene, respectively. The PCR fragment containing the *hcpR* gene was digested and cloned into the m-pET-21d vector using these restriction sites following a standard T4-ligase cloning procedure. As an alternative purification route, the *hcpR* gene was also cloned into the pFC20K vector (Promega) using the SgfI and EcoICRI restriction sites to add a C-terminal HaloTag. The recombinant plasmids were transformed and screened in *Escherichia coli* DH5 α cells and positive clones were subsequently sequenced to confirm and transformed into *E. coli* BL21 (DE3) cells (Bioline) for protein expression.

The N-terminal sensing domain of HcpR (referred to here as HcpR-SD) was produced using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). Primers were designed to change Leu156 to a stop codon on the HcpR-pET-21 plasmid. Subsequent PCR-generated plasmids were screened and sequenced. Plasmids that were positive for stop-codon insertion were transformed into *E. coli* BL21 (DE3) cells for protein expression.

Cells for the expression of HcpR (full length and HcpR-SD) were grown overnight with antibiotics in auto-induction medium or LB broth to an OD₆₆₀ of 0.5–0.7, and were induced using 1 mM IPTG. HcpR expressed from the m-pET-21d

plasmid was purified using Ni-NTA agarose (Qiagen) following the manufacturer's protocol and was eluted in 50 mM sodium phosphate, 300 mM NaCl, 250 mM imidazole pH 8.0 buffer. HcpR expressed from the pFC20K plasmid was purified using Halo affinity agarose (Promega) and was removed from the agarose via acTEV (Thermo Fisher Scientific) protease digestion as per the manufacturer's protocol in 25 mM HEPES, 150 mM NaCl, 1 mM TCEP buffer. Recombinant HcpR from the m-pET-21d vector purified from the Ni-NTA column was digested with TEV protease to remove the 6×His tag from HcpR; after digestion, the sample was passed through an Ni-NTA column to remove free tag, uncut protein and TEV protease.

To reconstitute it with heme, an ~2.5 molar excess of heme was added to the purified and digested HcpR. The sample was then dialyzed overnight in 1 l 20 mM sodium phosphate pH 7.5, 150 mM NaCl, 1 mM TCEP, 5% glycerol buffer to remove excess heme or passed through a HiTrap Desalting column (GE Healthcare) using the same buffer.

2.2. Crystallization and structure determination

His-tagged purified HcpR-SD protein was dialyzed against 25 mM Tris-HCl pH 7.0, 0.1 M NaCl, 1 mM TCEP after TEV digestion and concentrated to 15 mg ml⁻¹. Crystallization experiments for both HcpR-SD alone and an HcpR-SD-heme complex (1:2 ratio) were carried out using a Crystal Gryphon robot (Art Robbins Instruments) at 20°C. A wide range of commercially available crystallization conditions were screened; 58 µl reservoir solution and 400 nl crystallization drops were dispensed onto 96-well INTELLI-PLATES by a Gryphon dispenser. Unfortunately, thus far we have not been able to obtain crystals of the HcpR-SD-heme complex. Small crystals of HcpR-SD were obtained in one week with several different precipitants, and further refinement of the conditions resulted in diffracting crystals using ammonium sulfate as a precipitant. Attempts to improve the quality and the size of the crystals using the sitting-drop vapor-diffusion method were performed up to the microlitre range in 24-well VDX crystallization plates (Hampton Research). Two crystal forms were obtained using 1.2–1.4 M ammonium sulfate, 0.2 M NaCl, 0.1 M sodium acetate pH 4.5 as the reservoir solution. However, both crystal forms diffracted to 4–5 Å resolution. Crystal dehydration was applied to improve the diffraction quality of the crystals (Heras & Martin, 2005). Prior to data collection, the crystals were first washed in 2.5 µl cryoprotectant solution consisting of 0.1 M sodium acetate pH 4.5, 1.3 M ammonium sulfate, 0.2 M NaCl and were then transferred stepwise to similar solutions containing 5%, 10%, 12.5%, 15%, 17.5% and 20% glycerol. All steps were exposed to air and the soaking time for each step was about 3–4 min. Two additional annealing steps in cryoprotectant solution (25% glycerol) also improved the diffraction quality of the crystals. X-ray data sets for the tetragonal crystal form were obtained at 100 K on an R-Axis IV++ image-plate detector using Cu Kα X-rays ($\lambda = 1.5418$ Å) from a Rigaku MicroMax-007 X-ray source operating at 40 kV and 20 mA equipped with VariMax confocal

optics (Rigaku, The Woodlands, Texas, USA). The crystals diffracted to 3.15 Å resolution and belonged to space group $P4_122$, with typical unit-cell parameters $a = b = 145.30$, $c = 77.93$ Å and two monomers in the asymmetric unit. Intensity data were integrated, scaled and merged using *d*TREK* and were converted to amplitudes with *TRUNCATE* from the *CCP4* suite (Winn *et al.*, 2011). A heavy-atom derivative of HcpR was prepared by soaking a tetragonal crystal in 0.1 M sodium acetate pH 4.5, 1.3 M ammonium sulfate, 0.2 M NaCl solution containing 0.5 mM potassium tetrachloroplatinate(II) (K_2PtCl_4) for 2 h. The second crystal form of HcpR belonged to space group $C222_1$, with unit-cell parameters $a = 133.47$, $b = 138.85$, $c = 44.55$ Å, and diffracted to 2.24 Å resolution. A diffraction data set was collected at the Stanford Synchrotron Radiation Light Source (SSRL). The data were processed with *iMosflm* (*CCP4* suite). Diffraction data statistics for all three data sets are shown in Table 1.

2.3. Phasing, model building and refinement

Phases were calculated to 3.5 Å resolution based on isomorphous and anomalous differences for the platinum derivative. Heavy-atom sites were determined, and phasing was performed by *SOLVE* following density modification in *RESOLVE* (Adams *et al.*, 2010). The optimal solution had a BAYES-CC of 35.1, an FOM of 0.39 and a map skew of 0.12, with two Pt sites identified for the K_2PtCl_4 derivative. Auto model building resulted in an initial model of 238 amino-acid residues built with an R_{work} of 4.1% and an R_{free} of 4.7%. The phases were transferred to the isomorphous native data and extended to 3.15 Å resolution. The model was refined using *phenix.refine* to a final R_{work} of 19.60% and R_{free} of 23.90%. The crystal structure of the orthorhombic form was determined by the molecular-replacement method using *Phaser* as implemented in *PHENIX* with the structure of the tetragonal crystal form as a starting model (McCoy *et al.*, 2007; Adams *et al.*, 2010). The model was refined using *phenix.refine* and refinement was accomplished in *PHENIX* followed by manual rebuilding into $2mF_o - DF_c$ maps in *Coot* (Emsley *et al.*, 2010; Adams *et al.*, 2010). The structure was refined to 2.6 Å resolution with a final R_{work} of 23.50% and R_{free} of 30.47%. The final model consists of two monomers, while the biological dimer is formed with other symmetry-related molecules.

2.4. Small-angle X-ray scattering experiments

Recombinant HcpR was purified and sent to the SIBYLS beamline at the Lawrence Berkeley National Laboratory. A concentration of 2 mg ml⁻¹ apo HcpR was exchanged into 25 mM Tris, 100 mM NaCl, 2 mM TCEP pH 7.5 buffer. Each form of the protein was dialyzed against 2 l buffer overnight before being sent off. The matching buffer was sent with it. All SAXS data were collected at the Advanced Light Source (ALS), a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program supported by the DOE Office of Biological and

Table 1

Data-collection and refinement statistics.

	Platinum derivative	Native	Native
Data-collection statistics			
Radiation source	Rigaku MicroMax-007	Rigaku MicroMax-007	SSRL
Wavelength (Å)	1.5418	1.5418	1.18076
Space group	$P4_122$	$P4_122$	$C222_1$
Unit-cell parameters (Å)	$a = b = 145.30$, $c = 77.93$	$a = b = 144.97$, $c = 77.97$	$a = 133.47$, $b = 138.85$, $c = 44.55$
Monomers in asymmetric unit	2	2	2
Resolution (Å)	39.58–3.50 (3.63–3.50)	29.93–3.15 (3.26–3.15)	43.73–2.60 (2.70–2.60)
Measured reflections	91938	92849	93652
Unique reflections	10962 (1067)	14801 (1429)	13182 (1544)
Multiplicity	8.4 (8.2)	6.2 (6.1)	7.3 (7.5)
$\langle I/\sigma(I) \rangle$	11.7 (5.2)	13.9 (4.2)	16.2 (5.5)
Completeness (%)	99.7 (100.0)	99.5 (99.3)	99.8 (99.4)
$R_{\text{merge}}^{\dagger}$ (%)	13.5 (37.5)	7.7 (36.7)	6.6 (26.4)
R_{meas} (%)	14.4 (40.1)	8.4 (40.1)	7.6 (30.5)
$CC_{1/2}$ (%)	99.0 (88.0)	99.8 (92.5)	99.9 (96.6)
Structure refinement			
Resolution limit (Å)		28.99–3.15 (3.39–3.15)	43.73–2.60 (2.70–2.60)
No. of reflections		14778 (756)	13165 (1294)
R_{work} (%)		19.6 (28.8)	25.47 (30.63)
$R_{\text{free}}^{\ddagger}$ (%)		23.9 (37.4)	29.22 (39.29)
R.m.s.d. from standard geometry			
Bond lengths (Å)		0.010	0.003
Bond angles (°)		1.25	0.75
Ramachandran statistics (%)			
Most favored regions		96.0	96.30
Allowed regions		3.33	3.70
Outliers		0.67	0.00
Average B factors (Å ²)			
All atoms		67.1	57.39
Protein alone		67.1	55.90
Water		60.0	53.50

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$. $\ddagger R_{\text{free}}$ was calculated with 5% of reflections that were excluded from the refinement.

Environmental Research. Data were collected at a sample-to-detector distance of approximately 3.5 m at a wavelength of ~ 0.103 nm [$I(s)$ versus s , where $s = 4\pi \sin \theta / \lambda$ and 2θ is the scattering angle] at a temperature of 21°C. The data were normalized to the intensity of the transmitted beam and radially averaged; the scattering of the solvent blank was subtracted. The *ATSAS* software package was used to analyze the scattering profiles and to create the graphs (Franke *et al.*, 2017). *DAMMIF/DAMMIN* were used to create the *ab initio* space model (Franke & Svergun, 2009). *EOM* was used to evaluate the conformational flexibility of full-length HcpR (Tria *et al.*, 2015). All SAXS data were deposited in the SASBDB under accession number SASDEN9 (Valentini *et al.*, 2015).

2.5. Sedimentation-velocity experiments

Recombinant Halo-purified HcpR was dialyzed into 50 mM HEPES, 150 mM NaCl, 1 mM DTT pH 7 buffer and diluted to 0.1, 0.2, 0.5 and 1.0 mg ml⁻¹ with dialysis buffer. A Beckman Optima XL-1 analytical ultracentrifuge was used to analyze the samples. The sedimentation-velocity experiment was run at 30 000 rev min⁻¹ (~ 69 000g) in an eight-hole An-60 Ti rotor at 20°C using carbon-filled Epon double-sector cells capped with sapphire windows. The samples were incubated in the centrifuge for 1 h prior to starting the run. Boundary sedimentation was recorded for each cell using UV absorption (280 nm) and collected every 5 min.

2.6. Luminol assay

HcpR-SD was reconstituted with heme and dialyzed into 10 mM phosphate, 100 mM NaCl, 1 mM TCEP pH 7.5 buffer. Approximately 10 μ g was added to a native 4–16% polyacrylamide gel and the protein was resolved under native conditions. *P. gingivalis* OxyR reconstituted with heme was used as a negative control in the heme-binding study. The heme-binding proteins were detected by soaking the gel in luminol (PerkinElmer) for 5–10 min and then activating the heme with hydrogen peroxide (3%). The luminescence signal marking the presence of heme was detected using film.

2.7. UV–Vis spectrum studies

The spectrum of heme-reconstituted HcpR or HcpR-SD was recorded using a Biomate 3S UV–Vis spectrophotometer (Thermo Fisher Scientific) in a gas-tight 1 cm quartz cuvette in 25 mM Tris 100 mM NaCl, 1 mM TCEP pH 7.5 buffer or 10 mM phosphate, 100 mM NaCl, 1 mM TCEP pH 7.5 buffer. The ferrous form of heme was obtained by adding an excess of sodium dithionite (1 mM). An excess of *S*-nitrosoglutathione (GSNO) or diethylamine NONOate sodium salt hydrate (NONOate; Sigma) was used to obtain the nitrosylated form of the protein.

2.8. Resonance Raman spectroscopy

Recombinant HcpR was dialyzed into 25 mM Tris, 100 mM NaCl, 1 mM TCEP pH 7 buffer and concentrated to

20 mg ml⁻¹. An equimolar amount of heme was added to the sample with 5 mM dithionite to achieve ferrous heme. The sample was then desalted through a PD-10 desalting column in an anaerobic chamber to remove excess unbound heme and was diluted to 10 mg ml⁻¹. The nitrosylated form of HcpR was obtained by the addition of 100 µM NONOate. The anaerobic samples were added to glass melting-point capillaries and sealed. Resonance Raman spectra were obtained using a krypton ion laser at 406.7 nm (Spectra-Physics, model 171-01; Mountain View, California, USA). The detection system used was a liquid-nitrogen-cooled 400 × 1340 CCD detector (Princeton Instruments, Roper Scientific, Trenton, New Jersey, USA) and a 0.5 m spectrograph (Spex model 1870; Horiba/Jobin-Yvon, Edison, New Jersey, USA). *GRAMS/AI* v.7.0 was used to perform spectral baseline leveling by a fifth-order polynomial routine. The mathematical peak-fitting module of *OriginPro* v.7.5 was used to deconvolute band shapes and to generate the spectral graph.

2.9. Bioinformatics analysis

All figures showing structures in this article were generated using *UCSF Chimera* (Pettersen *et al.*, 2004). The buried solvent-excluded area calculations and comparison overlays of homologs were performed using *UCSF Chimera*. The volume of the hydrophobic pocket was calculated using the *CASTp* server (Dundas *et al.*, 2006). The alignments were made using *Clustal Omega* (Sievers *et al.*, 2011). *MODELLER* was used to create the chimeric HcpR model (Webb & Sali, 2016). The C-terminal domain was modeled using the DNA-binding domain of the crystal structure of full-length DNR (PDB entry 3dkw; Giardina *et al.*, 2009) as a template.

3. Results

3.1. Overview of the HcpR sensing domain

The C-terminal DNA-binding domain is highly dynamic, making determination of the full-length structure through crystallization difficult; thus, in order to gain insight into the structural characteristics of HcpR the N-terminal sensing domain was crystallized. The N-terminal sensing domain spans residues 1–156 and contains most of the dimerization helix (Fig. 1). The initial crystals belonging to space group *P*₄₁₂₂ diffracted to 3.5 Å resolution. A structure was solved using platinum single-wavelength anomalous diffraction (SAD) by soaking crystals in a solution of potassium tetrachloroplatinate. A model from the tetragonal crystal was refined to a final *R*_{work} of 19.6% and *R*_{free} of 23.9% (Table 1). This HcpR model was used for molecular replacement of a second crystal form that belonged to space group *C*222₁. This new model was refined to 2.6 Å resolution. There was no significant loss of electron density along the backbone of the tetragonal or orthorhombic crystal structures.

HcpR-SD forms a homodimer, as shown in Fig. 1(*a*). In the tetragonal crystal form there are two molecules in the asymmetric unit that represent the biological dimer. However, in the orthorhombic crystal form the biological dimer is formed

between monomers of symmetry-related molecules. In both crystal structures no electron density accounting for the cofactor was found; therefore, we conclude that this is the apo form of the protein. The ligand-sensing domain of HcpR consists of seven α-helices and seven β-sheets. The seven β-sheets form a β-barrel-like structure or ‘jelly-roll’ fold, which is common in the sensing domain of CRP-like regulators (Fig. 1*b*). According to the standard topology, the fold consists of eight β-sheets oriented in an antiparallel fashion. It is common for the sixth β-sheet to be shortened in the FNR-CRP family; however, in the case of HcpR the sixth β-sheet is completely distorted, in part owing to a proline residue (Pro89) in this region creating an α-helical structure. This also affects the loop region between this helix and the subsequent β-sheet 7, creating an extended loop region. To maintain a standard naming convention, we have numbered the β-sheets based on the eight-sheet topology of the jelly-roll fold (Fig. 1*b*).

A small subdomain composed of a helical bundle is located at the N-terminus. Additionally, the ‘flap’, a region that has been shown to be important in the allosteric activation of CRP and FNR-CRP proteins, is located between β-strands 4 and 5 (Passner *et al.*, 2000). The seventh α-helix (α-7) forms a dimerization interface that interacts with the corresponding helix of the opposing chain. The two dimerization helices interact with a buried solvent-excluded area of 258.4 Å³, which acts to stabilize the dimer. Superposition of chains *A* and *B* results in an r.m.s.d. of 0.413 Å for the backbone atoms, indicating no significant deviation between the two backbone chains of the homodimer.

A *CASTp* search resulted in two well defined pockets that could incorporate a heme molecule. The pockets are located in the space between the β-barrel motif and the dimerization helix of the opposite chain. Thus, the homodimer has two pockets on opposite faces of the protein (Fig. 2*a*). The core of the cavity is surrounded by hydrophobic residues derived from β-sheet 5, the helical region between β-5 and β-6, and the dimerization helices of both chains, with the opening lined by a ring of hydrophilic residues derived from the lower part of the dimerization helix and Arg85 (Fig. 2*a*). Together, these residues form a pocket that is approximately 520 Å³ in volume. Heme (protoporphyrin IX) has a volume of approximately 510 Å³, and thus it is possible for the pocket to dock heme. Aligning the sequences of members of the HcpR subfamily from related species, many of the conserved residues are located around or are in contact with this identified pocket (Supplementary Fig. S1). Of note, there is a relatively high sequence similarity in the region that corresponds to the distorted sixth β-sheet and the loop region before β-sheet 7 when compared with the rest of the sensing domain.

Coordination of the heme iron is essential to the function of heme-based gas sensors. Inside this pocket are several residues that could be implicated in the coordination of heme (Fig. 2*b*). His149 is oriented near the opening of the pocket. HcpR has three methionine residues extending into the pocket (Met68, Met135 and Met145). Methionine has been shown to be essential in heme coordination in cytochromes. Lys147 extends

into the opening of the pocket. Some of these residues (Met68, His124 and Lys147) are well conserved in the HcpR sequence alignments and could act as potential coordination sites in HcpR (Supplementary Fig. S1). The side chain of Arg85 also extends prominently into the pocket. Although arginine residues do not typically coordinate heme iron directly, arginine plays an important role in stabilizing the binding of NO to heme in NO synthase (Wang *et al.*, 1994).

3.2. Comparison of HcpR with other members of the FNR-CRP family

The best functionally and structurally characterized proteins with similarity to HcpR are DNR (from *P. aeruginosa*) and CoxA (from *R. rubrum*); we thus compared their structures with our structure of *P. gingivalis* HcpR. Despite low sequence identity between the three proteins, there is high similarity at the secondary and tertiary levels of protein folding (Figs. 3a,

3b and 3c). The CoxA and HcpR sensing domains superimpose with an r.m.s.d. of 0.958 Å and the HcpR and DNR sensing domains superimpose with an r.m.s.d. of 1.13 Å. The core β -barrel fold of the sensing domains, a key property of FNR-CRP family proteins, is maintained and is positioned similarly in all three structures. Likewise, all three proteins utilize the dimerization helix to stabilize the formation of a homodimer and it is positioned with the β -barrel analogously. Furthermore, the flap region of HcpR clearly aligns in the sequences and in the structural superposition of each protein.

In CoxA, the His N-terminal proline residue (Pro3) of one subunit in the dimer directly acts as an axial ligand for the iron in the heme cofactor of the opposite subunit. Binding of CO to the heme iron displaces this proline, leading to a substantial change in the structure. This displacement is the initial event that sends the allosteric signal through the protein, thereby activating it. Although HcpR has a proline located at its N-terminus (Pro3), its position is a large distance away from a

Figure 1

Overview of the N-terminal domain of HcpR. (a) Ribbon diagram of the N-terminal domain of HcpR shown from a side view (top) and a top view (bottom), showing the tertiary and quaternary orientation of the structure. Chain A is colored green and chain B is colored blue. The two chains form a homodimer through the interaction of the dimerization helix. (b) Secondary-structure orientation of HcpR shown from side and top views. HcpR contains seven different α -helical regions (α -1 through α -7) that are shown in red. α -7 is the dimerization helix. There are seven different β -sheet regions in HcpR (β -1 through β -8) that make up the β -barrel structure or 'jelly-roll' fold. In HcpR, the β -6 sheet is distorted. The flap region is labeled and is located in the loop region between β -4 and β -5. (c) Schematic representation of the domain organization of full-length HcpR (the region of the protein that was not crystallized is denoted by dashes).

potential heme-binding site. The other heme axial ligand, His77, is located in the sensing domain and is not conserved in HcpR (Supplementary Fig. S2*b*). Thus, the coordination of the heme iron in HcpR occurs through residues that are not analogous to those in CooA. Furthermore, the position and opening of the hydrophobic pocket that coordinates heme in CooA is different from that found in HcpR and other FNR-CRP members. When superimposed, the openings of the pockets are on opposite faces of the protein. This is most likely owing to differences in the lengths of β -sheets 4 and 5 in CooA (Arg56–Ala62 and Glu67–Glu74) and HcpR (Ile64–Met68 and Asp79–Leu81). It should be noted that the location and

opening of the pocket in HcpR is similar to those in CRP (and other members of the FNR-CRP family), although the size of the pocket is larger in HcpR (Seok *et al.*, 2014).

Despite the lack of a solved structure of a heme-bound form, DNR shares a slightly higher sequence similarity with HcpR than CooA. The opening of the hydrophobic pocket in DNR is in a similar position in the superposition of HcpR and DNR (Fig. 3*c*). Of the conserved residues found in the DNR subgroup, most are in contact with this cavity or form part of the cavity wall; however, these residues are not conserved in HcpR (Giardina *et al.*, 2008). This would imply that HcpR and DNR could bind to heme in a similar manner (the location of

Figure 2

Hydrophobic pocket of HcpR. (*a*) The locations of the hydrophobic pockets with respect to the full structure are marked by black boxes. The solid box indicates the pocket which is enlarged in the inset; the dotted box indicates the location of the pocket on the opposite side of HcpR. Inset: the surface area of the pocket region of HcpR is denoted using the Kyte–Doolittle hydrophobicity scale, where blue is more hydrophilic and red is more hydrophobic. An electron density consistent with glycerol is found in this pocket. (*b*) The location of potential heme-coordinating residues, denoted by stick figures, in and around the hydrophobic pocket.

the hydrophobic pocket is similar) but the exact mechanisms and coordination may differ. His187, the crucial residue that is displaced in NO-mediated activation of DNR, is not conserved in HcpR (Supplementary Fig. S2a). The other axial ligand believed to be involved in the coordination of heme in DNR, His139, is directly positioned over Met145 in the HcpR–DNR overlay. A single mutation in His139 is not capable of full inactivation of DNR, thus Met145 may serve an auxiliary role in heme binding and coordination in HcpR (Rinaldo *et al.*, 2012).

Just as in HcpR, β -sheet 6 is completely distorted in DNR, leading to an extended loop region before β -sheet 7. This appears to be an important distinction in the DNR and HcpR-like heme-binding regulators. Changes in this region can have large impacts on the size and shape of the hydrophobic pocket,

allowing the binding of diverse ligands across the FNR–CRP family. In the case of HcpR and DNR, increased flexibility in this loop region and loss of β -sheet 6 may allow incorporation of the heme cofactor.

3.3. Small-angle X-ray scattering modeling and conformational flexibility of full-length HcpR

To gain further insight into the mechanism of HcpR, we also characterized the full-length protein. Sedimentation-velocity experiments yielded a single, discrete species that sedimented at an average sedimentation coefficient of ~ 3.55 S, corresponding to a molecular weight of 54.2 kDa (Fig. 4, Supplementary Figs. S5 and S6). The frictional ratio calculated from the fits is 1.3. The *hcpR* gene encodes a protein with an

Figure 3

Comparison of HcpR with other members of the CRP–FNR family of bacterial transcription factors. (a) Sequence alignment of the sensing domains of HcpR, DNR and CooA. Secondary-structure features are shown in red (α -helices) and blue (β -sheets). The secondary-structural elements of HcpR are labeled above the sequence alignments. An asterisk (*) denotes a residue that is fully conserved, a colon (:) indicates a strongly similar residue and a period (.) indicates a weakly similar residue. The sequence identity between HcpR and CooA is 18%, that between HcpR and DNR is 20% and that between DNR and CooA is 16%. (b) Superposition of the HcpR (green) and CooA (blue; PDB entry 4k8f; Kuchinskas *et al.*, 2006) N-terminal sensing domain monomers, showing the orientation of the dimerization helices and the heme-binding domain. The r.m.s.d. on alignment of the backbone atoms (minus the first 12 residues of the N-terminus) is 0.958 Å. (c) Superposition of truncated DNR (red; PDB entry 2z69; Giardina *et al.*, 2008) and HcpR (green). The r.m.s.d. between the two structures is 1.31 Å.

Figure 4
HcpR sedimentation velocity: $c(s)$ plot of the sedimentation-coefficient distribution. Four concentrations of HcpR were used: $2.1 \mu\text{M}$ (0.1 mg ml^{-1}), $2.4 \mu\text{M}$ (0.2 mg ml^{-1}), $10.4 \mu\text{M}$ (0.5 mg ml^{-1}) and $20.8 \mu\text{M}$ (1.0 mg ml^{-1}). HcpR gives a single discrete species at $3.55 \pm 0.05 \text{ S}$, corresponding to a molecular weight of approximately $54.2 \pm 1 \text{ kDa}$.

Figure 5
SAXS profile and *ab initio* model of HcpR. (a) Scattering data shown in a $\log(I)$ versus q graph. Error bars are shown on the graph for each data point. (b) Kratky plot of the scattering data. This graph is characteristic of a partially disordered protein. (c) The distance distribution function [$P(r)$ profile] of HcpR indicates a maximum diameter of 96.2 Å . Inset: Guinier plot and the linear fit (solid line) of the Guinier region ($n_{\text{begin}} = 15$; $n_{\text{end}} = 51$).

Table 2
Structural parameters from SAXS experiments.

$I(0)$ [from $P(r)$]	283.1
R_g [from $P(r)$] (Å)	28.9
$I(0)$ (from Guinier)	282.9 ± 0.57
R_g (from Guinier) (Å)	29.3
D_{max} (Å)	104
Molecular mass (V_c) (kDa)	43
Molecular mass (MoW) (kDa)	44
Molecular mass (from dimer sequence) (kDa)	48

estimated molecular weight of 24.5 kDa ; thus, the observed 54.2 kDa product suggest that HcpR forms a dimer. Such data agree with the results obtained from our structural analysis, making it consistent with other members of the FNR-CRP family.

To gain insight into the full-length structure of HcpR, we performed small-angle X-ray scattering (SAXS) to model the missing domain. The SAXS structural parameters are presented in Table 2. The SAXS data confirm the homodimerization of the full-length protein, revealing an estimated molecular weight of 43 and 44 kDa using V_c and MoW calculations, respectively (Table 2, Fig. 5). Approximation of the Guinier region of the SAXS data indicates that HcpR has an R_g of approximately 29.2 Å and an $I(0)$ of 282.9 . The $p(r)$ function predicates a globular protein with a maximum diameter of 104 Å and an R_g of 28.9 Å . Using the DNA-binding domain of DNR (PDB entry 3dkw) as a template, a model of full-length HcpR was generated. The DNA-binding domain of all FNR-CRP proteins is a classical helix–turn–helix (HTH) motif and the full-length model is represented in Supplementary Fig. S3(a). However, this model did not agree with the scattering data. The theoretical scattering curve of the chimeric model fitted the HcpR SAXS data with an overall χ^2 of 38.09 according to *CRY SOL* (Supplementary Fig. S3b). The full-length model was refined using *SREFLEX* to a χ^2 of 14.75 .

(Supplementary Fig. S3c). However, this model only fitted the experimental scattering data in the low- s regions, with a significant variation in the $s > 0.2$ region. *Ab initio* models were created using *DAMMIN* from the SAXS data set. The generated SAXS molecular envelopes that best fitted the scattering data differed significantly from the chimeric model, with a best NSD of 3.1 (Supplementary Figs. S3d and S3e).

Observation of the Kratky plot reveals a form that differs from the typical bell shape of globular proteins and that is more consistent with a flexible or partially unfolded protein (Fig. 5b; Kikhney & Svergun, 2015; Receveur-Brechot & Durand, 2012). Thus, we used the ensemble-optimization method (*EOM*) to determine whether a mixture of multiple conformers fitted the data better. In this method, a random pool of 10 000 conformers of HcpR were generated where the N-terminal sensing domain dimer was kept fixed and the DBD model could assume multiple conformations. The distribution of R_g was calculated from the initial pool and were compared with 100 sub-ensembles. The best ensemble gave an average R_g of 28.87 Å and an average maximum diameter of 93.61 Å.

The theoretical *EOM* scattering curve fitted the HcpR experimental scattering data with a χ^2 of 0.655 (Fig. 6a). The best-fitting ensemble is represented by a flexibility of 88.5% [$R_{\text{flex}} \simeq 88.5\%$ (85.51%); $R_{\text{sigma}} = 1.50$]. The conformations represented are among the possible ones that can be attained, and should be interpreted as showing that full-length apo HcpR has a significant degree of interdomain flexibility (Fig. 6c). The ensemble is primarily represented by two main populations at $R_g \simeq 26.44$ Å and $R_g \simeq 31.13$ Å (Fig. 6b).

In the representative models of the *EOM* ensemble, the DNA-binding domain has a significantly different orientation to the chimeric model based on the structure of DNR. In all of the models the HTH domain moves via the hinge region connecting it to the dimerization helix, and the DNA-recognition helix of the HTH motif does not assume the correct position to bind to DNA. In each of the models in the ensemble the DNA-binding domain rotates around the dimerization helix of the sensing domain. This is expected, as the DNA-binding domains of FNR-CRP regulators can exhibit a remarkable degree of flexibility around the sensing

Figure 6

Flexibility of full-length HcpR. (a) Scattering curves of the HcpR SAXS data overlaid with the experimental scattering curve of the *EOM* model. The black dots represent the full-length HcpR SAXS scattering data and the red line indicates the best-fit theoretical scattering curve of the *EOM* model. The curves fit with a χ^2 of 0.622. (b) Distribution of the orientations of the C-terminal DNA-binding domain. The R_g of the best ensemble is 28.87 Å, with a D_{max} of 93.61 Å. (c) Representative models of the main subpopulations of the representative ensemble. The models are representative of the $R_g \simeq 26.4$ Å (i), $R_g \simeq 28.3$ Å (ii) and $R_g \simeq 31.1$ Å (iii) regions.

domain (Levy *et al.*, 2008; Lanzilotta *et al.*, 2000; Eiting *et al.*, 2005). This degree of flexibility also brings elements of the domain into the proximity of the hydrophobic pocket. In the $R_g \approx 26.44$ Å model (Fig. 6c, i), the helical region just below the hinge region is positioned near the opening of the pocket of the opposite monomer. Thus, the possibility that residues in this region (which includes the well conserved Lys160) may influence heme binding or coordination cannot be ruled out.

3.4. Heme- and nitric-oxide-binding properties of HcpR

It has been shown that heme is required for HcpR–DNA interaction, implicating heme as a necessary component in the

activation of HcpR (Lewis *et al.*, 2012). To confirm the heme-binding ability of HcpR, an equimolar amount of heme was added to purified HcpR–SD. The binding of heme to the protein was detected by resolving the protein on a native gel and visualizing the peroxidase activity of heme through luminescence in a gel that was soaked in luminol and exposed to peroxide. Excess heme ran down the gel and heme bound to HcpR migrated with the protein; no heme binding was observed in the negative-control sample (Fig. 7a).

Heme binding was verified by the UV–Vis spectrum of HcpR that was reconstituted with ferrous heme under anaerobic conditions (the reduced derivative was obtained by adding excess sodium dithionite to the sample; Fig. 7b). The absorption bands of the HcpR–heme spectrum also yield clues into the nature of heme binding. In the unligated form, a Soret band appears at 425 nm and α and β bands appear at 558 and 535 nm, respectively; this is consistent with a hexacoordinate, low-spin system. The ferrous NO derivative of HcpR yields a spectrum with a peak at 401 nm and a loss of the α and β bands (Fig. 7c). This result suggests a five-coordinate, NO-bound system, indicating that heme coordination in HcpR goes from a six-coordinate unbound state to a five-coordinate ligand-bound system when nitric oxide is added.

A series of resonance Raman spectra of heme-bound HcpR before and after the addition of NO is displayed in Fig. 8 and the results are summarized in Table 3. In the reduced HcpR sample, the most prominent band of the high-frequency region is the ν_4 band located at the 1362 cm^{-1} position of the ferrous sample, which is characteristic of a hexacoordinate ferrous heme. Other vibrational markers in the high-frequency region appear, ν_3 at 1492 cm^{-1} , ν_2 at 1586 cm^{-1} and ν_{10} at 1625 cm^{-1} , which have been shown to be sensitive to gas binding and confirm the hexacoordinate, low-spin nature of the un-nitrosylated heme (Fig. 8a; Spiro & Strekas, 1974; Andrew *et al.*, 2001). When compared with the prototypical NO sensor

Figure 7
Heme-binding and NO-binding properties of HcpR. (a) Luminol heme-binding assay of HcpR. All samples (lane 1, HcpR–SD; lane 2, negative control, *P. gingivalis* OxyR) were reconstituted with heme and run on a native gel. Luminol was added to the gel and chemiluminescence indicates the presence of heme after the addition of 3% H_2O_2 . Excess heme ran down the gel, whereas bound heme migrated with the protein. (b) Spectral properties of reconstituted heme-bound HcpR (1 mg ml^{-1}) under reduced, anaerobic conditions. Bound heme reveals a Soret peak at 425 nm which is indicative of a ferrous six-coordinate heme. Inset: $\alpha\beta$ region of the heme-bound HcpR. (c) Spectra of reconstituted HcpR ($10\text{ }\mu\text{M}$) before (bold line) and after (dotted line) the addition of NO under anaerobic conditions. After the addition of NONOate to the sample the Soret peak shifts to 401 nm, indicating a transition from a six-coordinate to a five-coordinate state.

Table 3
Heme vibrational markers and vibrational modes and nitric oxide binding.

Protein	Coordination No.	Spin†	$\nu_2^\ddagger$	ν_3	ν_4	ν_{10}	$\nu(\text{Fe-NO})$	Reference
Unligated								
HcpR	6	ls	1586	1492	1362	1625		This work
sGC	5	hs	1562	1471	1358	1606		Deinum <i>et al.</i> (1996)
CooA	6	ls	1580	1492	1361	1616		Vogel, Spiro <i>et al.</i> (1999)
NO bound								
HcpR	5	hs	1584	1508	1376	1646	535	This work
sGC	5	hs	1584	1509	1375	1646	525	Deinum <i>et al.</i> (1996)
CooA	5	hs	1582	1506	1376	1641	523	Reynolds <i>et al.</i> (2000)

† ls, low spin; hs, high spin. ‡ All values are given in cm^{-1} .

Figure 8
Resonance Raman spectra of reconstituted HcpR at 406.7 nm excitation. (a) Reduced heme-bound HcpR (10 mg ml^{-1}) under anaerobic conditions was excited at 406.7 nm at room temperature. The frequencies of the marker bands are representative of a six-coordinate system. (b) Reduced heme-bound HcpR (10 mg ml^{-1}) plus NO under anaerobic conditions was excited at 406.7 nm at room temperature. The frequencies of the marker bands are representative of a five-coordinate NO-bound system. The $\nu(\text{Fe-NO})$ frequency appears at 535 nm.

soluble guanylate cyclase (sGC), there are marked differences in the vibrational spectrum (Table 3). In the un-nitrosylated state, sGC exists as a pentacoordinate, high-spin system (Deinum *et al.*, 1996; Karow *et al.*, 2004). The Raman spectrum of HcpR in the un-nitrosylated state is more consistent with that of a hexacoordinate system as seen in CooA (Vogel, Spiro *et al.*, 1999). Although the large majority of the HcpR sample is in the hexacoordinate state, a small percentage of the sample appears to be in the pentacoordinate state, indicated by the shoulder at 1471 cm^{-1} . This is further confirmed by observing the polarizing component of the Raman spectrum (Supplementary Fig. S4). In the Raman spectrum the 1362 cm^{-1} peak (indicative of hexacoordinate system) dominates the ν_4 frequency; however, observing the polarizing component of the spectrum allows us

to see the hidden peak at 1376 cm^{-1} . This would imply that one of the axial side-chain ligands responsible for the sixth bond to the heme iron is a weak or moderate donor that forms a weak, transient bond to the iron.

Upon NO binding, the heme skeletal markers exhibit a very similar spectrum to those of sGC and CooA, producing a five-coordinate NO-bound system with a high spin factor. The ν_4 mode is shifted completely to 1376 cm^{-1} upon the addition of NO. Furthermore, ν_3 shifts to 1508 cm^{-1} , ν_2 shifts to 1584 cm^{-1} and ν_{10} shifts to 1646 cm^{-1} (Fig. 8b). These changes in the spectrum are indicative of heme binding and are in the expected range for a five-coordinate, high-spin iron(II)-NO heme, as opposed to a six-coordinate iron(II)-NO heme. Furthermore, $\nu(\text{Fe-NO})$ is found at 535 cm^{-1} . Bands in this region are consistent with a five-coordinate Fe-NO system (Vogel, Kozlowski *et al.*, 1999).

4. Discussion

Here, we report the structure of the sensing domain of HcpR (HcpR-SD). The structure verifies the inclusion of HcpR in the FNR-CRP family of regulators: the N-terminal domain forms the characteristic β -barrel structure and the homodimer oligomerization is confirmed by SAXS and sedimentation-velocity experiments. The mechanism and the exact location of heme binding are not immediately evident upon observation of the structure. There are several hydrophobic residues that line the space between the dimerization helix (α -7) and two of the β -sheets (β -3 and β -5) forming the hydrophobic pocket. Of note, the hydrophobic pocket is in contact with the structural elements important in the allosteric activation of FNR-CRP regulators: the dimerization helix and the β -hairpin region known as the ‘flap’ between β -sheets 4 and 5 (located at residues 68–77 in HcpR). Studies of other FNR-CRP regulators have shown that the flap plays an important role in transmitting the signal to the C-terminal helix–turn–helix DNA-binding domain and has been shown to play a role in both the interdomain and intersubunit interaction necessary for the transmission of a binding signal (Levy *et al.*, 2008).

In heme-based sensor proteins, the ability to change the coordination state of the heme iron is utilized for signaling and allosteric activation. Thus, understanding and identifying the residues involved in coordination is imperative in order to fully understand the mechanisms of these proteins. It is not

immediately clear which residues are involved in the coordination of heme iron in HcpR. Several residues that extend prominently into the pocket are possible candidates for heme coordination. These include Met68, Arg85, Met135, Met145, Lys146 and His149. Furthermore, the flexible nature of the DNA-binding domain, as confirmed by the SAXS analysis, potentially brings elements of this domain into contact with the opening of the hydrophobic pocket. Thus, residues such as Lys160 could play a role in heme axial coordination. This is not improbable, as His187 in DNR, which is located on the DNA-binding domain, has been implicated in heme axial coordination (Cutruzzolà *et al.*, 2014).

The heme-bound protein is stable, and the spectral and resonance Raman data suggest that the ferrous form is primarily hexacoordinate. The primarily hexacoordinate nature of the heme-bound form of HcpR transitions to a pentacoordinate system on the addition of NO to form the holoprotein. This is not uncommon: a five-coordinate system is observed in many heme-binding proteins when NO is bound (as seen in DNR, CooA and sGC). The innate affinity of NO for heme is much higher than that of CO or O₂ owing in large part to its strong back-bonding with the iron. This increased affinity causes NO to exert a strong *trans* effect on iron(II), resulting in a long and weak bond to an axial ligand (Scheidt *et al.*, 2010; Spiro *et al.*, 2013). This unique property of the NO–iron(II) bond is used to the advantage of other heme-based NO sensors. The primary molecular event correlated with sGC activation is the dissociation of the heme-proximal histidine bond upon NO binding to the distal face of the heme. This event (loss of proximal His coordination) triggers the structural allosteric changes within sGC that activate it (Stone & Marletta, 1994). A similar mechanism is also hypothesized to take place in DNR, in which the dissociation of an axial histidine residue (His187) is the key regulatory step in the NO-mediated activation of DNR (Cutruzzolà *et al.*, 2014). In both cases, CO and O₂ do not exert a sufficiently strong *trans* effect on the iron to break the His–iron bond. This allows NO-selective activation in both sensors. It is probable that HcpR follows a similar mechanism. Hcp is not upregulated in response to oxygen or oxidative stress and the $\Delta hcpR$ mutant of *P. gingivalis* is not sensitive to oxidative stress; thus, HcpR is capable of selective activation in response to NO *in vivo* (Boutrín *et al.*, 2012). Furthermore, the transient nature of the sixth axial ligand in the un-nitrosylated state implies that the fifth axial ligand may be the key regulator in the NO activation of HcpR. However, discerning the location and the identity of this residue will require a more targeted approach and the development of an activity assay.

In CRP, the binding of cAMP leads to substantial structural change in the homodimer that is characteristic of FNR-CRP proteins (Popovych *et al.*, 2009). Likewise, in DNR it is hypothesized that the protein undergoes a similar large structural change on NO binding (Giardina *et al.*, 2009). It is believed that the sensor domain, dimerization helix and DNA-binding domain act as separate bodies, changing their orientation with respect to each other upon ligand binding. In the case of CRP, it is believed that cAMP binding in the space

between the dimerization helix and the β -barrel of the N-terminal domain causes the N-terminal domain to move around the dimerization helices (Passner *et al.*, 2000; Tzeng & Kalodimos, 2013). The movement of the N-terminal domain will influence the orientation of the C-terminal DNA-binding domain through direct interaction with the ‘flap’ region, and the lower portion of the dimerization helix transitions from a disordered coil to an ordered helix (Popovych *et al.*, 2009). For this to occur, cAMP must make interactions with residues located on the dimerization helix (Passner *et al.*, 2000). Since its overall structure is similar to those of other FNR-CRP proteins, it could be assumed that NO binding to HcpR allosterically promotes the adoption of similar structural changes to promote activation and DNA binding. In HcpR, the binding of NO to the heme cofactor displaces two of the side chains coordinating the heme iron. It is possible that this displacement would then yield an allosteric transition. Although the identities of these residues are currently unknown, for activation to occur in the same manner as CRP and DNR one of these residues is most likely to be located on the dimerization helix.

Understanding the complete structural basis for NO-induced allosteric activation in HcpR is most likely to require the structures of the full-length and heme-bound forms of the protein. Our attempts to crystallize full-length HcpR have been unsuccessful thus far, probably owing to the dynamic character of the DNA-binding domain, as shown by the SAXS analysis. A simple hypothesis based on the mechanism of heme-based sensors is a two-state model with an ‘on’ and ‘off’ state (Liebl *et al.*, 2013). From the spectroscopic data, the non-gas-bound state is a six-coordinate system and the NO-bound state is a five-coordinate system. Thus, the off state of the protein is most likely to be a six-coordinate system and the active on state is most likely to be a five-coordinate system.

The apo form of the protein is not dominated by a single conformer. SAXS reveals that the DNA-binding domain of apo HcpR expresses a degree of flexibility. This would imply that the inactive, apo form of the protein is an assemblage of different conformers, which may be the source of its lower affinity for DNA. A possible mechanism of allosteric activation involves bound NO influencing the folding equilibrium of the DNA-binding domain, causing one conformer with a high affinity for DNA to dominate. SAXS experiments using the heme-bound and NO-bound forms of the protein would shed light on this mechanism; however, our attempts to obtain solution scattering data for these forms of HcpR were unsuccessful owing to aggregation and precipitation.

Based on structural and biochemical data, we show that HcpR is a member of the FNR-CRP family of regulators and demonstrate that the N-terminal sensing domain is capable of binding heme. We also have identified a hydrophobic pocket which may accommodate heme binding and have confirmed the functionality of this pocket through mutagenesis studies, implicating two methionine residues located in the pocket (Met68 and Met145) as important in the activation of the protein. Finally, the heme-bound form of the protein is

primarily a hexacoordinate system, which changes to a pentacoordinate system after the addition of NO.

Acknowledgements

The authors would like to thank Dr Darrell Peterson for providing the modified pET-21d vector and Dr James Turner for help with the Raman experiments. The authors declare that they have no conflicts of interest with the contents of this article. Author contributions were as follows. BRB, FNM and JB performed all experiments. All authors contributed to the experimental design and the analysis of data. BRB wrote the initial manuscript and all authors contributed to the writing and editing of the final manuscript.

Funding information

Funding for this work was provided by NIH grants 1R01DE023304 (JPL) and 1F31DE025158 (BRB).

References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* **D66**, 213–221.
- Andrew, C. R., Green, E. L., Lawson, D. M. & Eady, R. R. (2001). *Biochemistry*, **40**, 4115–4122.
- Arai, H., Igarashi, Y. & Kodama, T. (1995). *Biochim. Biophys. Acta*, **1261**, 279–284.
- Boutrin, M. C., Wang, C., Aruni, W., Li, X. & Fletcher, H. M. (2012). *J. Bacteriol.* **194**, 1582–1592.
- Cadby, I. T., Busby, S. J. & Cole, J. A. (2011). *Biochem. Soc. Trans.* **39**, 224–229.
- Cadby, I. T., Faulkner, M., Cheneby, J., Long, J., Van Helden, J., Dolla, A. & Cole, J. A. (2017). *Sci. Rep.* **7**, 16228.
- Cutruzzola, F., Arcovito, A., Giardina, G., della Longa, S., D'Angelo, P. & Rinaldo, S. (2014). *Biomaterials*, **27**, 763–773.
- Deinum, G., Stone, J. R., Babcock, G. T. & Marletta, M. A. (1996). *Biochemistry*, **35**, 1540–1547.
- Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acids Res.* **34**, W116–W118.
- Eiting, M., Hagelüken, G., Schubert, W.-D. & Heinz, D. W. (2005). *Mol. Microbiol.* **56**, 433–446.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Franke, D., Petoukhov, M. V., Konarev, P. V., Panjikovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M. & Svergun, D. I. (2017). *J. Appl. Cryst.* **50**, 1212–1225.
- Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.
- Giardina, G., Rinaldo, S., Castiglione, N., Caruso, M. & Cutruzzola, F. (2009). *Proteins*, **77**, 174–180.
- Giardina, G., Rinaldo, S., Johnson, K. A., Di Matteo, A., Brunori, M. & Cutruzzola, F. (2008). *J. Mol. Biol.* **378**, 1002–1015.
- Heras, B. & Martin, J. L. (2005). *Acta Cryst.* **D61**, 1173–1180.
- Karow, D. S., Pan, D., Tran, R., Pellicena, P., Presley, A., Mathies, R. A. & Marletta, M. A. (2004). *Biochemistry*, **43**, 10203–10211.
- Kikhney, A. G. & Svergun, D. I. (2015). *FEBS Lett.* **589**, 2570–2577.
- Kiley, P. J. & Beinert, H. (2003). *Curr. Opin. Microbiol.* **6**, 181–185.
- Körner, H., Sofia, H. J. & Zumft, W. G. (2003). *FEMS Microbiol. Rev.* **27**, 559–592.
- Kuchinkas, M., Li, H., Conrad, M., Roberts, G. & Poulos, T. L. (2006). *Biochemistry*, **45**, 7148–7153.
- Lanzilotta, W. N., Schuller, D. J., Thorsteinsson, M. V., Kerby, R. L., Roberts, G. P. & Poulos, T. L. (2000). *Nature Struct. Biol.* **7**, 876–880.
- Levy, C., Pike, K., Heyes, D. J., Joyce, M. G., Gabor, K., Smidt, H., van der Oost, J. & Leys, D. (2008). *Mol. Microbiol.* **70**, 151–167.
- Lewis, J. P., Yanamandra, S. S. & Anaya-Bergman, C. (2012). *Infect. Immun.* **80**, 3319–3331.
- Liebl, U., Lambry, J. C. & Vos, M. H. (2013). *Biochim. Biophys. Acta*, **1834**, 1684–1692.
- Lobato, L., Bouzahir-Sima, L., Yamashita, T., Wilson, M. T., Vos, M. H. & Liebl, U. (2014). *J. Biol. Chem.* **289**, 26514–26524.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Miller, D. P., Hutcherson, J. A., Wang, Y., Nowakowska, Z. M., Potempa, J., Yoder-Himes, D. R., Scott, D. A., Whiteley, M. & Lamont, R. J. (2017). *Front. Cell. Infect. Microbiol.* **7**, 378.
- Passner, J. M., Schultz, S. C. & Steitz, T. A. (2000). *J. Mol. Biol.* **304**, 847–859.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
- Popovych, N., Tzeng, S. R., Tonelli, M., Ebright, R. H. & Kalodimos, C. G. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 6927–6932.
- Receveur-Brechot, V. & Durand, D. (2012). *Curr. Protein Pept. Sci.* **13**, 55–75.
- Reynolds, M. F., Parks, R. B., Burstyn, J. N., Shelper, D., Thorsteinsson, M. V., Kerby, R. L., Roberts, G. P., Vogel, K. M. & Spiro, T. G. (2000). *Biochemistry*, **39**, 388–396.
- Rinaldo, S., Castiglione, N., Giardina, G., Caruso, M., Arcovito, A., Longa, S. D., D'Angelo, P. & Cutruzzola, F. (2012). *Antioxid. Redox Signal.* **17**, 1178–1189.
- Rodionov, D. A., Dubchak, I., Arkin, A., Alm, E. & Gelfand, M. S. (2004). *Genome Biol.* **5**, R90.
- Rodionov, D. A., Dubchak, I. L., Arkin, A. P., Alm, E. J. & Gelfand, M. S. (2005). *PLoS Comput. Biol.* **1**, e55.
- Scheidt, W. R., Barabanschikov, A., Pavlik, J. W., Silvernail, N. J. & Sage, J. T. (2010). *Inorg. Chem.* **49**, 6240–6252.
- Schultz, S. C., Shields, G. C. & Steitz, T. A. (1991). *Science*, **253**, 1001–1007.
- Seok, S.-H., Im, H., Won, H.-S., Seo, M.-D., Lee, Y.-S., Yoon, H.-J., Cha, M.-J., Park, J.-Y. & Lee, B.-J. (2014). *Acta Cryst.* **D70**, 1726–1742.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Spiro, T. G., Soldatova, A. V. & Balakrishnan, G. (2013). *Coord. Chem. Rev.* **257**, 511–527.
- Spiro, T. G. & Strekas, T. C. (1974). *J. Am. Chem. Soc.* **96**, 338–345.
- Stone, J. R. & Marletta, M. A. (1994). *Biochemistry*, **33**, 5636–5640.
- Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. (2015). *IUCrJ*, **2**, 207–217.
- Tzeng, S.-R. & Kalodimos, C. G. (2013). *Nature Chem. Biol.* **9**, 462–465.
- Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. (2015). *Nucleic Acids Res.* **43**, D357–D363.
- Vogel, K. M., Kozlowski, P. M., Zgierski, M. Z. & Spiro, T. G. (1999). *J. Am. Chem. Soc.* **121**, 9915–9921.
- Vogel, K. M., Spiro, T. G., Shelper, D., Thorsteinsson, M. V. & Roberts, G. P. (1999). *Biochemistry*, **38**, 2679–2687.
- Wang, J., Rousseau, D. L., Abu-Soud, H. M. & Stuehr, D. J. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 10512–10516.
- Webb, B. & Sali, A. (2016). *Curr. Protoc. Bioinformatics*, **54**, 5.6.1–5.6.37.
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Pottert, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* **D67**, 235–242.