



HAL
open science

Revealing the structural origin of the redox-Bohr effect: the first solution structure of a cytochrome from *Geobacter sulfurreducens*

Leonor Morgado, Vítor B. Paixão, Marianne Schiffer, P Raj Pokkuluri, Marta Bruix, Carlos A. Salgueiro

► To cite this version:

Leonor Morgado, Vítor B. Paixão, Marianne Schiffer, P Raj Pokkuluri, Marta Bruix, et al.. Revealing the structural origin of the redox-Bohr effect: the first solution structure of a cytochrome from *Geobacter sulfurreducens*. *Biochemical Journal*, 2011, 441 (1), pp.179-187. 10.1042/BJ20111103 . hal-00658161

HAL Id: hal-00658161

<https://hal.science/hal-00658161>

Submitted on 10 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Revealing the structural origin of the redox-Bohr effect: the first solution structure of a cytochrome from *Geobacter sulfurreducens*

Leonor Morgado^{*}, Vitor B. Paixão[†], Marianne Schiffer[‡], P. Raj Pokkuluri[‡], Marta Bruix^{§1}, Carlos A. Salgueiro^{*1}

^{*} Requite-CQFB, Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Campus Caparica, 2829-516 Caparica, Portugal

[†] Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Rua da Quinta Grande 6, 2780-156 Oeiras, Portugal

[‡] Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

[§] Departamento de Química Física Biológica, Instituto de Química-Física "Rocasolano", CSIC, Serrano 119, 28006 Madrid, Spain

¹Corresponding authors: mbruix@iqfr.csic.es and csalgueiro@dq.fct.unl.pt

SYNOPSIS

Geobacter sulfurreducens (*Gs*) can transfer electrons to the exterior of its cells, a property that makes it a preferential candidate for the development of biotechnological applications. Its genome encodes for over one hundred cytochromes and despite their abundance and key functional roles, to date there is no structural information for these proteins in solution. The trihaem cytochrome PpcA has a crucial role in the conversion of electronic energy into proton motive force, a fundamental step for ATP synthesis in the presence of extracellular electron acceptors. In this work, ¹⁵N-labelled PpcA was produced and NMR spectroscopy was used to determine its solution structure in the fully reduced state, its backbone dynamics and the pH-dependent conformational changes. The structure obtained is well defined, with an average pairwise root-mean-square deviation of 0.25 Å for the backbone atoms and 0.99 Å for all heavy atoms, and constitutes the first solution structure of a *Gs* cytochrome. The redox-Bohr centre responsible for controlling the electron/proton transfer was identified, as well as the putative interacting regions between PpcA and its redox partners. The solution structure of PpcA will constitute the foundation for studies aimed at mapping out in detail these interacting regions.

Heading title: PpcA solution structure

Keywords: Multihem cytochromes, Redox proteins, NMR, *Geobacter*, structure-function relationship

Abbreviations: PpcA – *Geobacter sulfurreducens* trihaem cytochrome encoded by gene GSU0612, *Gsc*₇ – *Geobacter sulfurreducens* trihaem cytochromes, rmsd – root-mean-square deviation.

INTRODUCTION

The bacterium *Geobacter sulfurreducens* (*Gs*) is able to reduce extracellular radioactive and toxic metal ions, as well as transfer electrons to electrode surfaces. These properties make this microorganism an important candidate for developing biotechnological applications in bioremediation and microbial electricity production [1, 2]. The genome of *Gs* was fully sequenced and encodes for 111 *c*-type cytochromes, of which 73 are multihaemic [3]. Among these, a family composed by five periplasmic trihaem cytochromes *c*₇ has been identified. These cytochromes contain approximately 70 amino acids, three haem groups with bis-histidinyl axial coordination, and share between 41 and 77% sequence identities. They were designated as PpcA (encoded by gene GSU0612), PpcB (GSU0364), PpcC (GSU0365), PpcD (GSU1024) and PpcE (GSU1760) [4]. Genomic and proteomic studies showed that these proteins affect differently the extracellular reduction of iron and uranium by *Gs* [5, 6].

The trihaem cytochrome PpcA is involved in extracellular metal ion reduction *in vivo*, and *in vitro* studies showed that it displays Fe^{III} reductase activity [7]. The detailed thermodynamic properties of PpcA redox centres have been determined, showing that the haem groups have negative reduction potentials, which are modulated by haem-haem redox interactions and by redox-Bohr interactions [8]. The functional rationalization of these properties at physiological pH showed that PpcA can couple electron and proton transfer. Thus, PpcA might contribute to the proton electrochemical potential gradient across the periplasmic membrane that drives ATP synthesis, constituting an additional mechanism to sustain cellular growth in presence of extracellular electron acceptors [9]. Given the importance of PpcA on metal ions respiration, understanding the function of this protein is required for the study of the biogeochemical activities of this bacterium. Studies of this type could lead to design of improved biotechnological applications based on *Gs* metabolic activities.

The modulation of multihaem cytochromes reduction potentials by redox and redox-Bohr interactions requires control of both redox state and the protonation state of the protein. Within this scope, structure determination of the protein in solution is the preferred method. The crystal structure of PpcA was previously determined [4]. However, in contrast to the structures obtained for the other four *Gs* cytochromes *c*₇ (*Gsc*₇), PpcA crystals could only be obtained in the presence of the additive deoxycholate [4]. Preliminary NMR studies carried out for PpcA samples without deoxycholate showed that the haem core architecture in solution is different from the PpcA crystal structure, but is similar to that observed in the crystal structures of the other four *Gsc*₇ [10, 11]. These studies showed that the binding of deoxycholate to PpcA in the crystals altered its structure [11]. Attempts to crystallize PpcA in absence of deoxycholate were unsuccessful and, thus, the structural information for this protein in an unaltered form is not available.

In this work, we used our recent developed methodology to obtain isotopically labelled recombinant multihaem cytochromes to produce ¹⁵N-labelled PpcA in a cost-effective manner [12]. The yields obtained were sufficient to determine its high-resolution solution structure. The structure presented is the first solution structure reported for a cytochrome from *Gs*. The structural information obtained was used to explore the backbone dynamics and to fingerprint the residues involved in the redox-Bohr effect, a crucial property that allows PpcA to couple electron and proton transfer at physiological pH.

EXPERIMENTAL

Bacterial growth and protein purification

Uniformly ^{15}N labelled and unlabelled PpcA were expressed in *Escherichia coli* as previously described [12]. Briefly, PpcA was expressed in plasmid pCK32 in *E. coli* BL21(DE3) by co-expression with plasmid pEC86 that encodes for the cytochrome *c* maturation gene cluster [13]. After reaching an OD_{600} of ~ 1.5 cultures were processed in either of two ways: (a) addition of $10\ \mu\text{M}$ isopropyl β -D-thiogalactoside (IPTG) and growing overnight at $30\ ^\circ\text{C}$ to express unlabelled protein, followed by harvesting by centrifugation; (b) cells were collected by centrifugation after reaching an OD_{600} of 1.0, washed twice with 250 ml M9 medium, resuspended in minimal media (in a ratio of 250 ml of minimal medium for each litre of 2xYT medium) supplied with $1\ \text{g/l}$ $^{15}\text{NH}_4\text{Cl}$ as nitrogen source (together with $1\ \text{mM}$ of the haem precursor α -aminolevulinic acid, trace amounts of metal ion salts, biotin and thiamine), grown overnight at $30\ ^\circ\text{C}$ in the presence of $0.8\ \text{mM}$ IPTG and harvested by centrifugation. Isolation of the periplasmic fraction and purification of the protein was done as previously described [12, 14].

NMR sample preparation

For NMR experiments, samples were lyophilized and dissolved in $45\ \text{mM}$ sodium phosphate buffer ($100\ \text{mM}$ ionic strength) pH 7.1 (direct reading with a glass micro electrode without correction for isotope effects) and with 0.04% sodium azide to avoid bacterial growth. ^{15}N labelled sample was prepared in 92% $\text{H}_2\text{O}/8\%$ D_2O and unlabelled samples prepared in 92% $\text{H}_2\text{O}/8\%$ D_2O or in pure D_2O . Final protein concentration was about $1\ \text{mM}$. In order to avoid oxidation of the samples, the NMR tubes were sealed with a gas-tight serum cap and the air was flushed out from the sample. The samples were then reduced directly in the NMR tube with gaseous hydrogen in the presence of catalytic amounts of hydrogenase from *Desulfovibrio vulgaris* (Hildenborough), as previously described [10].

NMR spectroscopy

NMR experiments were acquired on a Bruker Avance 800 MHz or on a Bruker Avance 600 MHz spectrometers equipped with triple-resonance cryoprobes. All experiments were acquired at 298K and ^1H and ^{15}N chemical shifts were respectively calibrated using the water signal as internal reference and through indirect referencing [15]. Spectra were processed using TOPSPIN (Bruker Biospin, Karlsruhe, Germany) and analyzed with Sparky [16]. The following set of experiments was acquired: ^{15}N labelled sample: 2D- ^1H - ^{15}N -HSQC, 3D- ^1H - ^{15}N -TOCSY (60 ms) and 3D- ^1H - ^{15}N -NOESY (80 ms); unlabelled sample in 92% $\text{H}_2\text{O}/8\%$ D_2O : 2D- ^1H -COSY, 2D- ^1H -TOCSY (60 ms) and 2D- ^1H -NOESY (50 ms); unlabelled sample in D_2O : 2D- ^1H -COSY, 2D- ^1H -TOCSY (45 ms) and 2D- ^1H -NOESY (100 ms). 1D- ^1H NMR spectra were obtained before and after each multidimensional spectrum to confirm protein integrity and fully reduction.

A series of ^1H - ^{15}N correlation spectra were measured at 800 MHz and at 600 MHz to obtain the ^{15}N longitudinal (T_1) and transverse (T_2) relaxation times in the same conditions used for structure calculation. For the longitudinal relaxation rate, a set of eight experiments was acquired with relaxation delays from 5 to 1200 ms, and for the transverse relaxation rate, eight experiments were recorded with relaxation delays between 16 to 130 ms. Heteronuclear $^{15}\text{N}\{^1\text{H}\}$ NOE were determined from the ratio of two experiments with and without saturation. Data from the different spectra were treated as peak heights.

The effect of pH titration on the chemical shifts was determined by analyzing a series of 2D- ^1H - ^{15}N -HSQC spectra acquired in the pH range 5.5 to 9.5. The pH of the sample was adjusted by addition of small amounts of NaOD or DCl inside an anaerobic chamber (MBraun LABstar) to avoid sample oxidation. The weighted average chemical shift ($\Delta\delta_{\text{avg}}$) of each backbone and side chain amide was calculated as: $\Delta\delta_{\text{avg}} = \sqrt{[(\Delta\delta^2\text{N}/25 + \Delta\delta^2\text{H})/2]}$, where $\Delta\delta\text{H}$ and $\Delta\delta\text{N}$ are the differences in the ^1H and ^{15}N chemical shifts, respectively [17].

Assignment and determination of restraints

Protein assignment, including backbone, side chain and haem resonances, was previously described [18] and data was deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) under BMRB

accession number 16842. For structure determination, the cross-peaks assigned in the 50 ms 2D-¹H-NOESY were integrated using the program Sparky. A Gaussian function was used for the integration of isolated peaks and sum data in a box surrounding the peak for overlapped ones. Cross-peaks due to protons separated by fixed distances and all intra-haem cross-peaks, except those involving the propionate groups, were not included. Volume for integrated cross-peaks were converted into volume restraints and used as input for the program PARADYANA [19], as described previously [20]. Initial structures were calculated with a preliminary set of NOE data and the resulting conformers were then analyzed and used to assign additional peaks in the NOESY spectra. The program GLOMSA [21] was used for stereospecific assignment during the process of structure calculation. Three non-standard residues were used for structure calculations: fast-flipping aromatic residues with pseudo-atoms to limit the orientations of the planes, flexible haem groups and proline residues with fixed upper limit distances for ring closure. A set of 69 fixed upper limit distances associated with this type of residues was also used as input for PARADYANA. In the final stages of structure refinement, the calculated structures were checked for short (less than 2.5 Å) distances between assigned protons that should give rise to significant NOEs.

Structure calculation and analysis

Structure calculation was performed with the program PARADYANA, a version of DYANA that takes peak volumes as input and pseudocontact shifts in the case of paramagnetic proteins [19]. In each calculation 200 conformers were determined, from which the 10 structures with lowest target function value were selected for further analysis. In the final calculation, 400 conformers were determined and the 20 with lowest target function value were selected. The program CHIMERA [22] was used for visual analysis during preliminary calculations and the program MOLMOL [23] was used for superimposition and identification of elements of secondary structure in the final set of conformers. Quality of the structures was analysed with PROCHECK-NMR [24].

Data Bank accession number

The structure of PpcA has been deposited in the Protein Data Bank with accession number 2LDO.

RESULTS

^{15}N -labelled PpcA was produced as previously described [12] with a final yield of 4 mg of pure protein per litre of culture. We used the expression system developed by Londer and co-workers [14], which uses the *lac* promoter in the expression plasmid with co-expression of the cytochrome *c* maturation gene cluster on a separate plasmid. The use of this system allows obtaining high protein yields with the correct fold and post-translational modification of the haem groups in a cost-effective manner and, concomitantly to overcome the traditional difficulties associated with the determination of solution structures using natural abundance samples of multihaem proteins.

Sequential assignment and structure calculations

PpcA has a MW of 9.57 kDa, and contains three low spin *c*-type haem groups with bis-histidinyl axial coordination and an unusually high content of lysine residues (24%). The haems are numbered I, III and IV, a designation that derives from the superimposition of the haems in cytochromes *c*₇ with those of the structurally homologous tetrahaem cytochromes *c*₃. The assignment of all haem protons, the protein backbone and side chains resonances, except for the first two residues, was previously reported [18]. The 60 haem proton signals (20 per haem) were unambiguously assigned, according to the strategy described by Turner and co-workers [25]. After that, individual spin-systems were first identified according to their type and then sequence-specific assignment made. A summary of the sequential connectivities between NH, H α and H β protons is shown in supplementary Figure S1. The total extent of the assignment is 93%, excluding carboxyl, amino and hydroxyl groups.

Assigned cross-peaks in the H₂O spectra were integrated and converted into volume restraints, resulting in 1115 lower limits for volumes (lov) and 1434 upper limits (upv) (Table 1). These were used as input for the program PARADYANA [19] together with a set of 69 fixed upper limit distances. The preliminary structures were analysed using the program GLOMSA [21], modified to take NOE volumes as input, and 32 stereospecific assignments were made for diastereotopic pairs of protons or methyl groups.

The effect of spin diffusion introduces an uncertainty into the conversion of experimental data to distance constraints. This effect was simulated by complete relaxation matrix calculations based on the initial protein structures and, accordingly, a parameter was set in the program PARADYANA to loosen all distance restraints by 5%. An average of 36 NOE restraints per amino acid residue (16 lovs and 20 upvs) and 170 per haem residue (74 lovs and 96 upvs) was used for the final calculation (Figure 1). The distribution of the number of constraints is not uniform along the protein sequence, as haem groups attached to positions 30, 54 and 68 show many long distance contacts.

Quality and analysis of the structures

The twenty structures with the lowest target function values (from 1.94 Å² to 2.16 Å², average value 2.07 Å², 11 % range from the lowest value) were selected as being representative of the solution structure of the protein. The structures superimpose with an average pairwise backbone (N-C α -C β) rmsd of 0.25 Å and a heavy atom rmsd of 0.99 Å (Supplementary Figure S2). Thus, the backbone is very well defined and the amino acid side chains showing larger conformational variability correspond to regions with higher solvent exposure. In particular, the N- and the C-terminals are disordered. The statistics for this family of structures is shown in Table 1. The Ramachandran plot shows 70.4% of the residues in the most favoured regions and 29% in the additionally allowed. A total of 35 hydrogen bonds were identified in the family of 20 structures with the program MOLMOL [23], 23 of which were present in at least 50% of the structures. In order to avoid different levels of protonation of the groups involved in the redox-Bohr effect, and a concomitant broadening of the signals, the PpcA solution structure (Figure 2) was determined at pH 7.1, which is more than one pH unit lower than the redox-Bohr centre pK_a in the reduced protein [8].

^{15}N relaxation measurements

^{15}N NMR relaxation has been used to characterize the dynamic properties of reduced PpcA in solution. The relaxation parameters T_1 , T_2 and NOE were determined at 600 and 800 MHz (Figure 3). The relaxation data were obtained for all residues except for proline residues (positions 16, 25, 35 and

62) and the two N-terminal amino acids Ala¹ and Asp². The average values for T_1 are 500 and 655 ms and for T_2 values are 119 and 104 ms, at 600 and 800 MHz, respectively. The average NOE value is 0.76 at 600 MHz and 0.81 at 800 MHz, respectively. Good correlation was observed between structure and experimental relaxation data. Most residues in regular secondary structure elements exhibit heteronuclear NOE values close to the theoretical maximum indicating high rigidity in these regions (Figure 3). The overall differences in the T_1 values are not significant. Higher variability was clearly observed in the NOE and T_2 data. High T_2 values correlate with a decrease in the NOE ratio in positions Asp²⁶, Gly³⁴, Gly³⁶ and the three residues at the C-terminus (positions His⁶⁹, Lys⁷⁰ and Lys⁷¹), which is in agreement with motions in the ns-ps time scale. Other regions of PpcA with low or average NOE values present lower T_2 values with respect to the mean. These correspond to residues His¹⁷, Val²⁴ and Cys²⁷, indicating that these residues are affected by conformational exchange processes in the μ s-ms time scale.

pH titration

The pH titration of PpcA was carried out by ¹H-¹⁵N HSQC NMR in the pH range 5.5-9.5 and all ¹H and ¹⁵N chemical shifts of the polypeptide backbone (except for residues 1 and 2) and side chains were measured. To estimate the effects of pH changes on the PpcA solution structure, the average chemical shift differences ($\Delta\delta_{\text{avg}}$) of each amide were calculated as described by Garret and co-workers [17]. The residues whose NH signals indicated larger differences ($\Delta\delta_{\text{avg}} > 0.2$ ppm) were from Lys⁷, Ala⁸, Asn¹⁰, Ile³⁸ backbone and also from Asn¹⁰ side chain (Figure 4A). These residues were mapped on the structure indicated in Figure 4B. The pH titration of the amide signals of Lys⁷, Ala⁸, Asn¹⁰ and Ile³⁸ are indicated in Figure 4C. The NH signals of Lys⁷, Ala⁸, Asn¹⁰ have basic pK_a values of 7.0, 7.9 and 7.5 respectively, whereas that of Ile³⁸ has a much more acidic pK_a (5.5). These pK_a values obtained from the pH dependence of the NH chemical shifts show that the pH-linked conformational changes have different origins.

DISCUSSION

Comparison of PpcA solution structure with *Gsc*₇ crystal structures

In solution, the structure of the trihaem cytochrome PpcA folds in a two-strand β -sheet at the N-terminus formed by Asp³-Leu⁶ and Val¹³-Pro¹⁶ strands, followed by three α -helices between residues Ala¹⁹-Lys²², Lys⁴³-His⁴⁷, and Lys⁵²-Met⁵⁸. The three bis-histidinyll haem groups are arranged in a triangular way with haems I and IV almost parallel with each other and both nearly perpendicular to haem III (Figure 2). This structure was compared with the crystal structures available for the *Gsc*₇ [4, 10, 11]. The parameters describing the haem geometry of PpcA in solution along with equivalent values for *Gsc*₇ crystal structures [4, 10, 11] and cytochrome *c*₇ from *Desulfuromonas acetoxidans* (*Dac*-7) [26] are presented in Table 2. With exception of the crystal structure of PpcA, for which the Fe-Fe distances and angles between the haem planes in the PpcA crystal structure are different due to the binding of the additive deoxycholate [11], the general fold and the relative positions of the three haems are similar, with a good agreement between the consensus secondary structure elements identified in the NMR and crystal structures. The average iron-iron distances found among the crystal structures of PpcB, PpcC, PpcD and PpcE differ by less than 4% compared to those of PpcA solution structure (Table 2). Differences between the haem core in PpcA solution and crystal structures were previously anticipated from the analysis of NOE interhaem connectivities between haems III and IV. Indeed, several unexpected connectivities were observed on basis of the distances taken from PpcA crystal structure [10]. The differences between the solution and crystal structures of PpcA are not just limited to the haem core but extend beyond. In particular, significant differences are observed in the polypeptide region between His¹⁷ and Glu³⁹ that surrounds haem I.

On the other hand, the patterns of interhaem NOE connectivities were identical to those observed for PpcB, which has the highest sequence identity with PpcA (77%), suggesting the same haem core arrangement for both proteins in solution. The comparison between the solution structure of PpcA and PpcB crystal structures (Figure 5) showed that the global rmsd values of the lowest-energy NMR structure to molecule A and molecule B of PpcB in the unit cell are 1.22 Å and 1.12 Å for backbone atoms, respectively. Among the four PpcA homologue crystal structures, significant structural differences were noted by Pokkuluri and co-workers [11], in particular, on the axial ligand orientations of haem I and the polypeptide region surrounding haem I (residues 16-40). The region around haem I is also where the sequences of *Gsc*₇ homologues differ the most suggesting a functional significance. In this region, the PpcA solution structure showed a non-regular helix forming a bulge between the two helical regions comprised of residues Lys⁴³-His⁴⁷ and Lys⁵²-Met⁵⁸. This structural feature is similar to the crystal structures of PpcB and PpcC but not to those of PpcD and PpcE where a regular helix is formed in this region.

Backbone dynamics

The backbone dynamics obtained for PpcA corroborates the ordered nature of the protein, and confirms that the less defined regions of the molecule are the result of their intrinsic flexibility. The ¹⁵N NMR relaxation experiments showed that the polypeptide segment between His¹⁷ and Cys²⁷ that surrounds haem I has a dynamic behaviour that differs from rest of the protein. Interestingly, *T*₂ values in this region (below the mean value) point to the participation of exchange processes in the μ s-ms time regime (Figure 3). The dynamic behaviour of this region suggested that this could be an interaction surface of PpcA with other proteins, as dynamics in this time scale have been postulated to be essential for the recognition and interaction between protein molecules [27]. In fact, the detailed thermodynamic characterization of PpcA hinted that haem I and IV might have a role in recognition of PpcA redox partners [8]. The proposed functional redox-cycle for PpcA involves two dominant microstates: a microstate in the one-electron oxidised protein with haem I oxidised while keeping the redox-Bohr centre protonated (*P*_{IH}) and another in the two-electron oxidised protein with haems I and IV oxidised and deprotonated redox-Bohr centre (*P*_{IA}). Since haem I is oxidised in the two functional microstates, it is most likely that the haem I region of PpcA interacts with its electron acceptor. Haem IV, which alters its oxidation state between the one-and two-electron oxidised protein, is very likely the interaction region with the electron donor. However, in contrast with the region around haem I, no backbone dynamics effects were detected around haem IV. This behaviour is not completely unexpected as there are some examples of rigid protein recognition sites [28, 29]. The lack of

dynamics has been suggested as a requisite for the selective sequence recognition. Structurally, haem IV is stabilized by hydrogen bonds to both propionate groups (see below), restricting the backbone dynamics around this haem. This is in line with the fact that, in the crystal structures of the entire family of *Gsc7* [11], the region around haem I varies significantly but that around haem IV is highly conserved, probably reflecting the higher structural rigidity in this region.

Haem reduction potentials and redox interactions

The detailed thermodynamic properties of PpcA redox centres were previously reported [8] and are summarized in Table 3. The solution structure determined in the presented work allowed addressing the structural basis for the PpcA redox properties. The haem negative reduction potentials are expected for relatively exposed *c*-type haems with bis-histidine axial coordination. The haem exposures are 231.7, 215.8, and 171.3 Å² for haems I, III and IV respectively, and the reduction potential values of the haem groups correlate with their solvent exposure. In fact, the more exposed haems (haem I and III) have more negative reduction potentials (-154 and -138 mV, respectively) whereas the least exposed one (haem IV) has the least negative reduction potential value (-125 mV). The higher reduction potential of the latter is also reinforced by its substantial positive environment due to the presence of several neighbouring lysine residues that stabilize the reduced state of haem IV. The haem reduction potentials are also modulated by haem-haem redox interactions [8]. The positive values obtained for these interactions suggest that they are dominated by electrostatic effects, rather than conformational changes between redox stages. This is confirmed by the fact that the highest (41 mV) and the lowest (16 mV) redox interaction are observed between the closest (III-IV) and furthest (I-IV) pair of haem groups.

Structural close-up on PpcA redox-Bohr centre: pH-linked conformational changes

In addition to the haem-haem redox interactions, the haem reduction potentials are modulated by the solution pH [8]. This is designated redox-Bohr effect by analogy with the Bohr effect in the haemoglobin [30]. The redox-Bohr interaction for haem IV is twice as high as compared to that of the other two haems (Table 3), suggesting that the redox-Bohr centre is located in the vicinity of haem IV. In fact, the redox-Bohr centre was previously assigned to haem IV propionate 13 (P₁₃^{IV}, according to the IUPAC Nomenclature [31]) [10, 32]. The redox-Bohr effect, as observed for PpcA, is crucial for coupling electron and proton transfer at physiological pH. The detailed thermodynamic characterization of this trihaem cytochrome indicated that such coupling involves the two microstates *P_{IH}* and *P_{I4}* [8]. Thus, a fraction of the energy associated with electrons received from the donor by microstate *P_{IH}* is used by the protein to lower the *pK_a* value of the redox-Bohr centre so that protons can be released in the periplasm at physiological pH. This mechanism would imply structural conformational changes in the neighbourhood of the redox-Bohr centre.

The solution structure determined in this work confirms P₁₃^{IV} as the redox-Bohr centre and allowed for the first time mapping the pH-linked conformational changes associated with the protonation/deprotonation of this centre. The backbone NH signals of Lys⁷, Ala⁸, Asn¹⁰, Ile³⁸ and from Asn¹⁰ side chain are the most affected by the protonation/deprotonation of the redox-Bohr centre (Figure 4). Residues 7, 8 and 10 are part of the β-turn segment connecting the two-strand β-sheet near haem IV, whereas Ile³⁸ is a fully conserved residue within the family of *Gsc7* and protects haem I and axial His³¹ from solvent exposure.

The amide proton of Ile³⁸ forms a conserved hydrogen bond with the carboxyl oxygen of P₁₃^I in the crystal structures of the entire family of *Gsc7* [11]. This hydrogen bond is not seen in PpcA solution structure probably due to the flexibility of the propionate groups of haem I. However, at lower pH values, the proton chemical shift of the Ile³⁸ NH signal decreases significantly suggesting the disruption of a hydrogen bond formed between the amide proton of Ile³⁸ and the carboxyl oxygen of P₁₃^I upon protonation of the latter (Figure 4C).

The *pK_a* values obtained for the NH signals of Lys⁷, Ala⁸, Asn¹⁰ in this work correlate with the *pK_a* value of the redox-Bohr centre (8.6) previously determined for the fully reduced and protonated protein [8], and is likely to be caused by the protonation/deprotonation of the redox-Bohr centre (P₁₃^{IV}). The differences of the *pK_a* values obtained for the redox-Bohr centre and those described from the analysis of the NMR signals of Lys⁷, Ala⁸, Asn¹⁰ are expected since the former value was derived

from the global thermodynamic analysis of the properties of the redox-centres and may reflect the combined contribution of fractional protonation of several acid-base groups [33].

Although being also part of the β -turn segment connecting the two-strand β -sheet, no significant changes on the NH backbone signal were observed for residue Lys⁹. The most likely explanation for this resides is the conserved hydrogen bond between the Lys⁹ backbone NH and the O2A atom of P₁₇^{IV}. Indeed, P₁₇^{IV} is exposed to surface (44.1 Å² for carboxyl oxygens) and would have nearly the same pK_a of a free propionate group (~ 4.0). Therefore, in the pH range studied it remains deprotonated and forms the hydrogen bond with Lys⁹ backbone NH.

The pH-linked conformational changes observed in the neighbourhood of P₁₃^{IV} and the structural features around this group help to rationalize the redox-Bohr effect observed in PpcA. P₁₃^{IV} is considerably less exposed to solvent than P₁₇^{IV} (16.3 Å² and 44.1 Å², respectively) and thus would be very sensitive to the conformation of the protein surroundings depending on its protonation state. Being shielded from the solvent, the protonated form is stabilized which explains the pK_a of the redox-Bohr centre well above its expected pK_a in free solution (~4.0). Furthermore, in the solution structure of PpcA, carbonyl oxygen of Lys⁷ in the β -turn (residues 7-12) is within hydrogen bonding distance from P₁₃^{IV} (2.6 Å between O7 and haem IV O1D in the lowest-energy structure). Thus, deprotonation of the propionate P₁₃^{IV} disrupts the hydrogen bond with Lys⁷ leading to structural changes in the neighbourhood as a function of pH. This is likely to be the origin of pH-linked conformational changes in the β -turn as the above mentioned contact is not favoured when this buried propionate is deprotonated.

CONCLUSIONS

The structure of PpcA determined in this work constitutes the first solution structure of a cytochrome from *G. sulfurreducens*. The polypeptide fragment formed by residues 17-27 located near haem I of PpcA was identified as the most dynamic segment by analysis of the ¹⁵N relaxation parameters obtained for the backbone. This segment may be involved in interaction with other molecules. The analysis of the chemical shift variation of the backbone and side chain amide signals with pH allowed mapping the pH-linked conformational changes caused by protonation/deprotonation of the redox-Bohr centre, which is now confirmed to be the haem propionate P₁₃^{IV}. The redox-Bohr effect of PpcA is a crucial property that allows it to perform a concerted e⁻/H⁺ transfer and to contribute to the proton electrochemical gradient across the bacterial cytoplasmic membrane that drives ATP synthesis. The solution structure of PpcA will be used to assist in the rational design of mutants and in the mapping of interaction sites with redox partners, an essential step towards engineering variants of *Gs* with increased respiratory rates and the concomitant improvement in their biotechnological applications.

ACKNOWLEDGMENTS

We would like to acknowledge to Prof. David L. Turner for the upgraded versions of structure calculation software and Dr. Daniel F. Jana for technical support. We would also like to thank Dr. David M. Tiede (Argonne National Laboratory) for many useful discussions. We acknowledge LabRMN at FCT-UNL and Rede Nacional de RMN for access to the facilities.

FUNDING

This work was supported by Fundação para a Ciência e Tecnologia (Portugal) research grant [PTDC/QUI/70182/2006] and by Ministerio de Ciencia e Innovación (Spain) research grant [CTQ2008-0080/BQU]. L.M. was supported by Fundação para a Ciência e Tecnologia [SFRH/BD/37415/2007].

P.R.P and M.S (USA) are supported in part by the US Department of Energy's Office of Science, Biological and Environmental Research GTL program under contract [DE-AC02-06CH11357] and by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under contract [DE-AC02-06CH11357].

Rede Nacional de RMN is supported with funds from Fundação para a Ciência e Tecnologia, Projecto de Re-Equipamento Científico [REDE/1517/RMN/2005], Portugal.

REFERENCES

- 1 Lovley, D. R. (2003) Cleaning up with genomics: applying molecular biology to bioremediation. *Nat Rev Microbiol* **1**, 35-44
- 2 Lovley, D. R. (2008) The microbe electric: conversion of organic matter to electricity. *Curr Opin Biotechnol* **19**, 564-571
- 3 Methé, B. A., Nelson, K. E., Eisen, J. A., Paulsen, I. T., Nelson, W., Heidelberg, J. F., Wu, D., Wu, M., Ward, N., Beanan, M. J., Dodson, R. J., Madupu, R., Brinkac, L. M., Daugherty, S. C., DeBoy, R. T., Durkin, A. S., Gwinn, M., Kolonay, J. F., Sullivan, S. A., Haft, D. H., Selengut, J., Davidsen, T. M., Zafar, N., White, O., Tran, B., Romero, C., Forberger, H. A., Weidman, J., Khouri, H., Feldblyum, T. V., Utterback, T. R., Van Aken, S. E., Lovley, D. R. and Fraser, C. M. (2003) Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments. *Science* **302**, 1967-1969
- 4 Pokkuluri, P. R., Londer, Y. Y., Duke, N. E., Long, W. C. and Schiffer, M. (2004) Family of cytochrome *c*₇-type proteins from *Geobacter sulfurreducens*: structure of one cytochrome *c*₇ at 1.45 Å resolution. *Biochemistry* **43**, 849-859
- 5 Ding, Y. H., Hixson, K. K., Giometti, C. S., Stanley, A., Esteve-Núñez, A., Khare, T., Tollaksen, S. L., Zhu, W., Adkins, J. N., Lipton, M. S., Smith, R. D., Mester, T. and Lovley, D. R. (2006) The proteome of dissimilatory metal-reducing microorganism *Geobacter sulfurreducens* under various growth conditions. *Biochim Biophys Acta* **1764**, 1198-1206
- 6 Shelobolina, E. S., Coppi, M. V., Korenevsky, A. A., Didonato, L. N., Sullivan, S. A., Konishi, H., Xu, H., Leang, C., Butler, J. E., Kim, B. C. and Lovley, D. R. (2007) Importance of *c*-type cytochromes for U(VI) reduction by *Geobacter sulfurreducens*. *BMC Microbiol* **7**, 16
- 7 Lloyd, J. R., Leang, C., Hodges Myerson, A. L., Coppi, M. V., Cuifo, S., Methe, B., Sandler, S. J. and Lovley, D. R. (2003) Biochemical and genetic characterization of PpcA, a periplasmic *c*-type cytochrome in *Geobacter sulfurreducens*. *Biochem J* **369**, 153-161
- 8 Morgado, L., Bruix, M., Pessanha, M., Londer, Y. Y. and Salgueiro, C. A. (2010) Thermodynamic characterization of a triheme cytochrome family from *Geobacter sulfurreducens* reveals mechanistic and functional diversity. *Biophys J* **99**, 293-301
- 9 Mahadevan, R., Bond, D. R., Butler, J. E., Esteve-Núñez, A., Coppi, M. V., Palsson, B. O., Schilling, C. H. and Lovley, D. R. (2006) Characterization of metabolism in the Fe(III)-reducing organism *Geobacter sulfurreducens* by constraint-based modeling. *Appl Environ Microbiol* **72**, 1558-1568
- 10 Morgado, L., Bruix, M., Orshonsky, V., Londer, Y. Y., Duke, N. E., Yang, X., Pokkuluri, P. R., Schiffer, M. and Salgueiro, C. A. (2008) Structural insights into the modulation of the redox properties of two *Geobacter sulfurreducens* homologous triheme cytochromes. *Biochim Biophys Acta* **1777**, 1157-1165
- 11 Pokkuluri, P. R., Londer, Y. Y., Yang, X., Duke, N. E., Erickson, J., Orshonsky, V., Johnson, G. and Schiffer, M. (2010) Structural characterization of a family of cytochromes *c*₇ involved in Fe(III) respiration by *Geobacter sulfurreducens*. *Biochim Biophys Acta* **1797**, 222-232
- 12 Fernandes, A. P., Couto, I., Morgado, L., Londer, Y. Y. and Salgueiro, C. A. (2008) Isotopic labeling of *c*-type multiheme cytochromes overexpressed in *E. coli*. *Prot Expr Purif* **59**, 182-188

- 13 Arslan, E., Schulz, H., Zufferey, R., Kunzler, P. and Thöny-Meyer, L. (1998) Overproduction of the *Bradyrhizobium japonicum* *c*-type cytochrome subunits of the *cbb*₃ oxidase in *Escherichia coli*. *Biochem Biophys Res Commun* **251**, 744-747
- 14 Londer, Y. Y., Pokkuluri, P. R., Tiede, D. M. and Schiffer, M. (2002) Production and preliminary characterization of a recombinant triheme cytochrome *c*₇ from *Geobacter sulfurreducens* in *Escherichia coli*. *Biochim Biophys Acta* **1554**, 202-211
- 15 Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. and Sykes, B. D. (1995) ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6**, 135-140
- 16 Goddard, T. D. and Kneller, D. G. (2007) Sparky 3.114. University of California, San Francisco, USA
- 17 Garrett, D. S., Seok, Y. J., Peterkofsky, A., Clore, G. M. and Gronenborn, A. M. (1997) Identification by NMR of the binding surface for the histidine-containing phosphocarrier protein HPr on the N-terminal domain of enzyme I of the *Escherichia coli* phosphotransferase system. *Biochemistry* **36**, 4393-4398
- 18 Morgado, L., Paixão, V. B., Salgueiro, C. A. and Bruix, M. (2011) Backbone, side chain and heme resonance assignments of the triheme cytochrome PpcA from *Geobacter sulfurreducens*. *Biomol NMR Assign* **5**, 113-116
- 19 Turner, D. L., Brennan, L., Chamberlin, S. G., Louro, R. O. and Xavier, A. V. (1998) Determination of solution structures of paramagnetic proteins by NMR. *Eur Biophys J* **27**, 367-375
- 20 Paixão, V. B., Salgueiro, C. A., Brennan, L., Reid, G. A., Chapman, S. K. and Turner, D. L. (2008) The solution structure of a tetraheme cytochrome from *Shewanella frigidimarina* reveals a novel family structural motif. *Biochemistry* **47**, 11973-11980
- 21 Güntert, P., Braun, W. and Wüthrich, K. (1991) Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J Mol Biol* **217**, 517-530
- 22 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612
- 23 Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **14**, 51-55, 29-32
- 24 Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-486
- 25 Turner, D. L., Salgueiro, C. A., LeGall, J. and Xavier, A. V. (1992) Structural studies of *Desulfovibrio vulgaris* ferrocyclochrome *c*₃ by two-dimensional NMR. *Eur J Biochem* **210**, 931-936
- 26 Assfalg, M., Banci, L., Bertini, I., Bruschi, M., Giudici-Ortoni, M. T. and Turano, P. (1999) A proton-NMR investigation of the fully reduced cytochrome *c*₇ from *Desulfuromonas acetoxidans*. Comparison between the reduced and the oxidized forms. *Eur J Biochem* **266**, 634-643

- 27 Mandel, A. M., Akke, M. and Palmer, A. G., 3rd. (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme. *J Mol Biol* **246**, 144-163
- 28 Pérez-Cañadillas, J. M., Guenneugues, M., Campos-Olivas, R., Santoro, J., Martínez del Pozo, A., Gavilanes, J. G., Rico, M. and Bruix, M. (2002) Backbone dynamics of the cytotoxic ribonuclease α -sarcin by ^{15}N NMR relaxation methods. *J Biomol NMR* **24**, 301-316
- 29 Fushman, D., Cahill, S. and Cowburn, D. (1997) The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of ^{15}N relaxation with monomer/dimer equilibration. *J Mol Biol* **266**, 173-194
- 30 Perutz, M. F., Wilkinson, A. J., Paoli, M. and Dodson, G. G. (1998) The stereochemical mechanism of the cooperative effects in hemoglobin revisited. *Annu Rev Biophys Biomol Struct* **27**, 1-34
- 31 Moss, G. P. (1988) Nomenclature of tetrapyrroles. Recommendations 1986 IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). *Eur J Biochem* **178**, 277-328
- 32 Dantas, J. M., Morgado, L., Londer, Y. Y., Fernandes, A. P., Louro, R. O., Pokkuluri, P. R., Schiffer, M. and Salgueiro, C. A. (2011) Pivotal role of the strictly conserved aromatic residue F15 in the cytochrome c_7 family. *J Biol Inorg Chem* *in press*, doi:10.1007/s00775-00011-00821-00778
- 33 Fonseca, B. M., Saraiva, I. H., Paquete, C. M., Soares, C. M., Pacheco, I., Salgueiro, C. A. and Louro, R. O. (2009) The tetraheme cytochrome from *Shewanella oneidensis* MR-1 shows thermodynamic bias for functional specificity of the hemes. *J Biol Inorg Chem* **14**, 375-385
- 34 Assfalg, M., Banci, L., Bertini, I., Bruschi, M. and Turano, P. (1998) 800 MHz ^1H NMR solution structure refinement of oxidized cytochrome c_7 from *Desulfuromonas acetoxidans*. *Eur J Biochem* **256**, 261-270
- 35 Czjzek, M., Arnoux, P., Haser, R. and Shepard, W. (2001) Structure of cytochrome c_7 from *Desulfuromonas acetoxidans* at 1.9 Å resolution. *Acta Crystallogr D Biol Crystallogr* **57**, 670-678

Accepted Manuscript

TABLES

Table 1. Summary of restraint violations and quality analysis for the final families of structures for PpcA.

Parameter	
<i>Type of distance restraint</i>	
Intra-residue	879
Sequential	550
Medium range ($2 \leq i - j < 5$)	501
Long range ($ i - j \geq 5$)	619
Total	2549
	(1115 lov + 1434 upv)
<i>Upper distance limit violations</i>	
Average maximum	0.19 ± 0.01
Number of consistent violations ($>0.2\text{\AA}$)	0
<i>Lower distance limit violations</i>	
Average maximum	0.20 ± 0.01
Number of consistent violations ($>0.2\text{\AA}$)	0
<i>Van de Waals violations</i>	
Average maximum	0.17 ± 0.01
Number of consistent violations ($>0.2\text{\AA}$)	0
<i>Ramachandran Plot (%)^a</i>	
Most favoured regions	70.4
Additionally allowed regions	29
Generously allowed regions	0.5
Disallowed regions	0.2
<i>Stereospecific Assignments^b</i>	
	32
<i>Precision</i>	
Average pairwise rmsd backbone (\AA)	0.25 ± 0.07
Average pairwise rmsd heavy atoms (\AA)	0.99 ± 0.10

^a Values obtained with PROCHECK-NMR^b Analysis with GLOMSA

Accepted Manuscript

Table 2. Haem geometry for *Gs* cytochromes *c*₇. The values for PpcA NMR reduced structure were obtained from the lowest-energy structure. The values for the NMR solution structure of *Dac*₇ [26] are also included for the sake of completeness and are not discussed.

	NMR		Crystal structures ^a				NMR
	PpcA	PpcA ^b	PpcB ^c	PpcC	PpcD ^c	PpcE	<i>Dac</i> ₇ ^d
Haem Fe-Fe distance (Å)							
I-III	11.7	11.2	11.7 (11.5)	11.2	11.2 (11.3)	11.8	11.9
I-IV	19.0	20.8	18.2 (18.8)	18.5	18.2 (18.2)	18.2	18.9
III-IV	12.6	12.6	12.6 (12.8)	12.6	12.6 (12.6)	12.4	12.5
Angle between haem planes (°)							
I-III	82	86	73 (72)	81	78 (76)	71	41
I-IV	27	35	20 (16)	18	25 (24)	22	11
III-IV	74	69	70 (71)	75	72 (70)	72	56
Angle between His planes (°)							
I	50	57	79 (88)	52	12 (3)	77	55
III	23	22	35 (36)	38	38 (22)	34	27
IV	80	89	75 (81)	86	85 (84)	89	53

^a The values were taken from [11].

^b The crystal structure of PpcA between haems I and III is altered in comparison with the other *Gs* cytochromes *c*₇ due to the presence of the additive deoxycholate used in crystallization. It should not be used for structural comparisons.

^c PpcB and PpcD cytochromes present two monomers in the crystal. The values corresponding to monomer B are given in parentheses.

^d *Dac*₇ shares 46% sequence identity with PpcA and was extensively studied by Assfalg and co-workers [26, 34] and Czjzek and co-workers [35].

Table 3. Haem redox potentials and pairwise interactions (mV) of the fully reduced and protonated forms of PpcA [8]. The standard errors are given in parentheses.

Haem redox potentials			Redox interactions			Redox-Bohr interactions		
I	III	IV	I-III	I-IV	III-IV	I-H	III-H	IV-H
-154(5)	-138(5)	-125(5)	27(2)	16(3)	41(3)	-32(4)	-31(4)	-58(4)

FIGURE LEGENDS

Figure 1. Number of constraints per residue used for the calculation of the structure of PpcA. Bars are white, light grey, dark grey and black for intra residue, sequential, medium and long range restraints, respectively. Residues 30, 54 and 68 also include restraints to haems I, III and IV, respectively.

Figure 2. PpcA solution structure. (A) Overlay of the 20 lowest energy NMR structures of PpcA at pH 7.1. Superimposition was performed using all the heavy-atoms. The peptide chain and the haems are colored grey and black, respectively. (B) Ribbon diagram of PpcA structure. Figures were produced using MOLMOL [23].

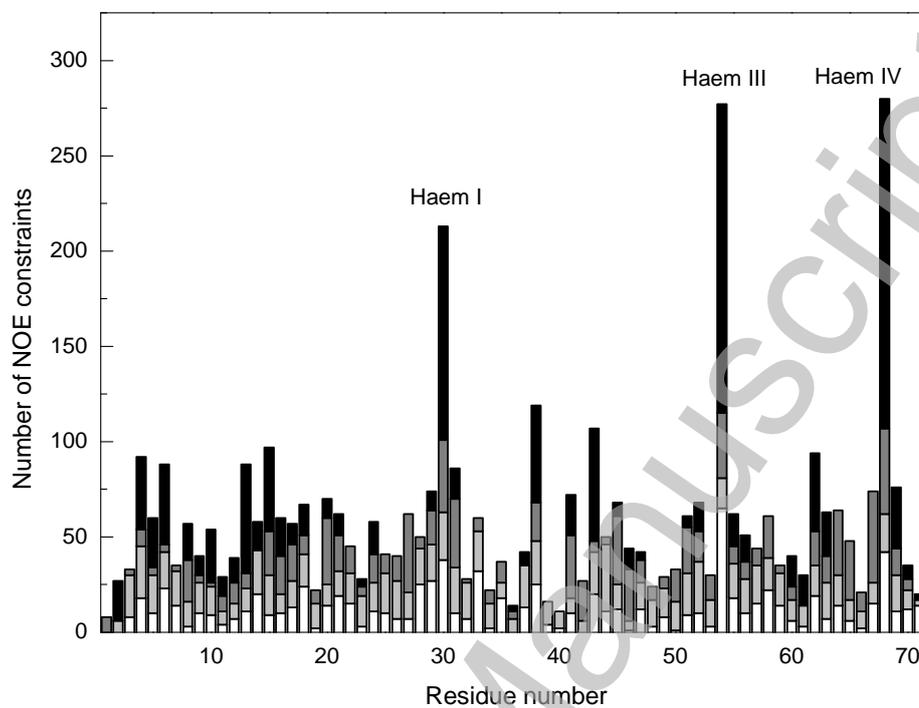
Figure 3. ^{15}N relaxation parameters for PpcA backbone in the reduced state. (A) longitudinal T_1 relaxation times, (B) transverse T_2 relaxation times, and (C) ^1H - ^{15}N heteronuclear NOEs. Solid straight lines represent the average T_1 , T_2 and NOE at the respective magnetic field; open circles 800 MHz and closed circles 600 MHz. Secondary structure elements are shown as arrows (β -strands) and cylinders (α -helices).

Figure 4. pH-linked conformational changes in PpcA. (A) Weighted average of ^1H and ^{15}N chemical shifts ($\Delta\delta_{\text{avg}}$) between pH 5.5 and 9.5. (B) Mapping of the residues showing large pH-dependent shifts on PpcA solution structure. Residues that are close to the haem IV are indicated in green and Ile³⁸ that forms a conserved hydrogen bond with P₁₃¹ carboxyl oxygen is indicated in blue. Haems groups are colored red. Dashed lines indicate hydrogen bonds. (C) pH titration data of the most affected PpcA amide signals. In the expansion of each ^1H - ^{15}N -HSQC NMR spectrum pH increases from violet to pink. Lys⁷ and Ala⁸ backbone amide signals are not shown since they appear in a very crowded region of the spectra.

Figure 5. Comparison of PpcA lowest energy solution structure with PpcB crystal structure (monomers A and B). Structures were superimposed in MOLMOL [23] using backbone atoms. (A) Average rmsd between each pair of structures. Closed and open symbols represent PpcB monomers A and B, respectively. (B) PpcA solution structure *versus* PpcB crystal structure (monomer B). PpcA solution structure is colored light grey, and PpcB crystal structure is black.

FIGURES

Figure 1



Accepted Manuscript

Figure 2

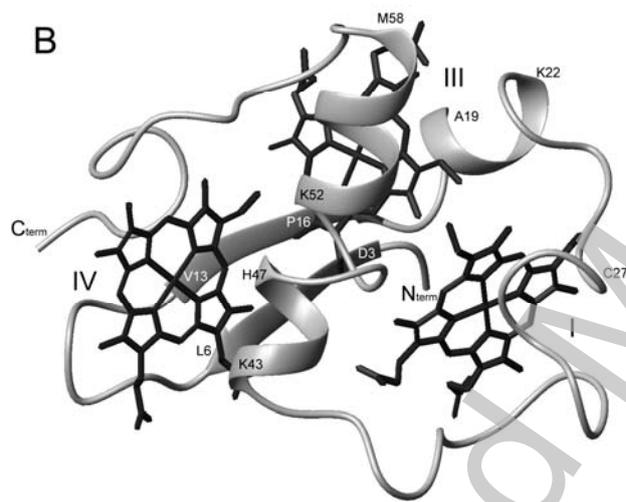
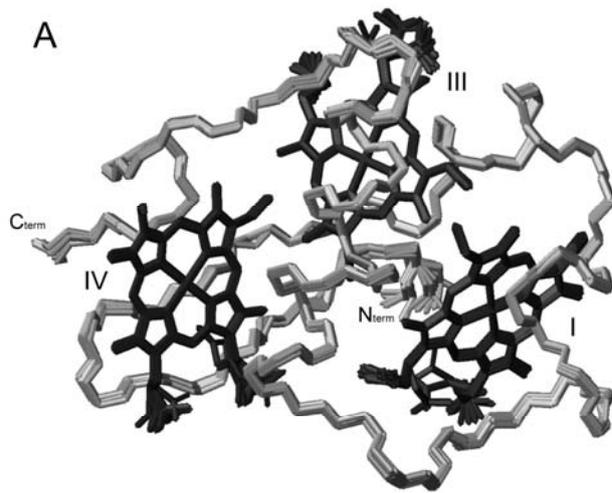


Figure 3

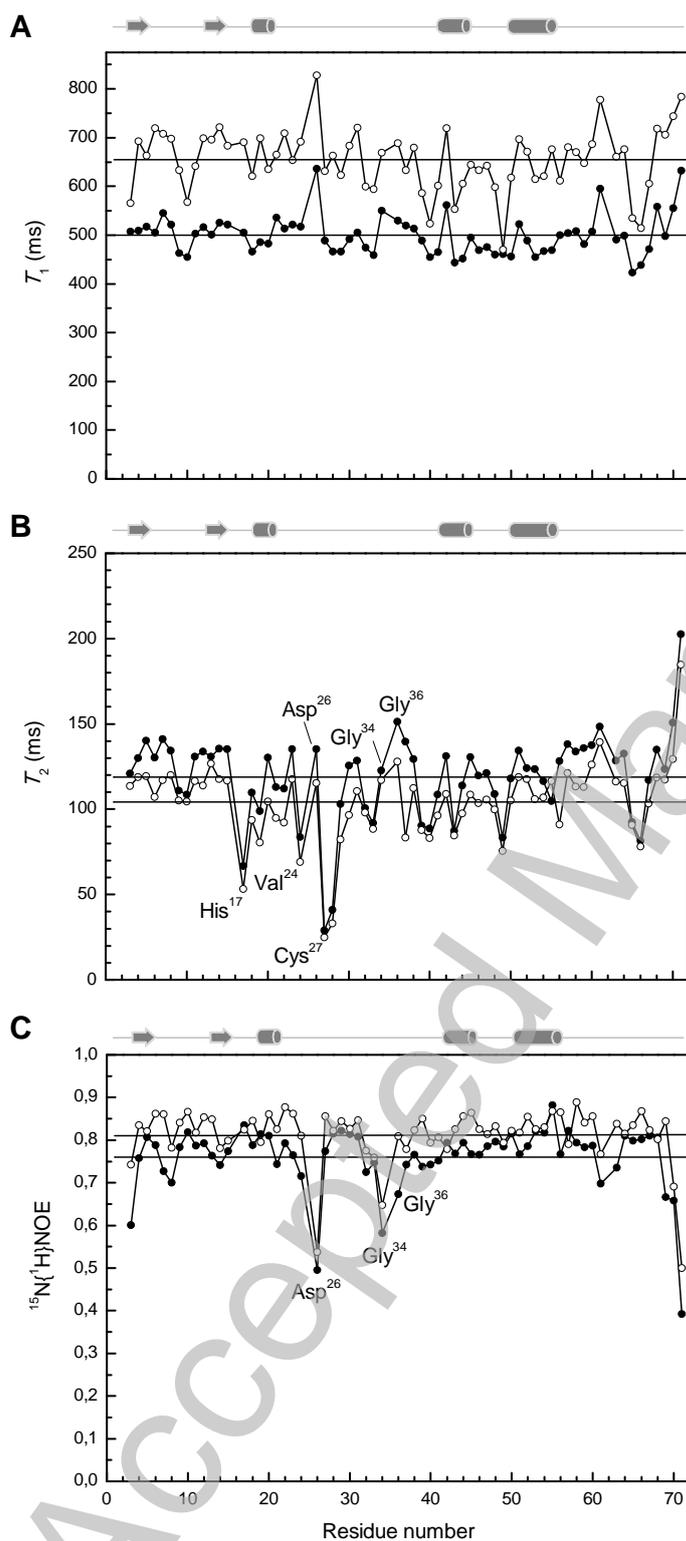


Figure 4

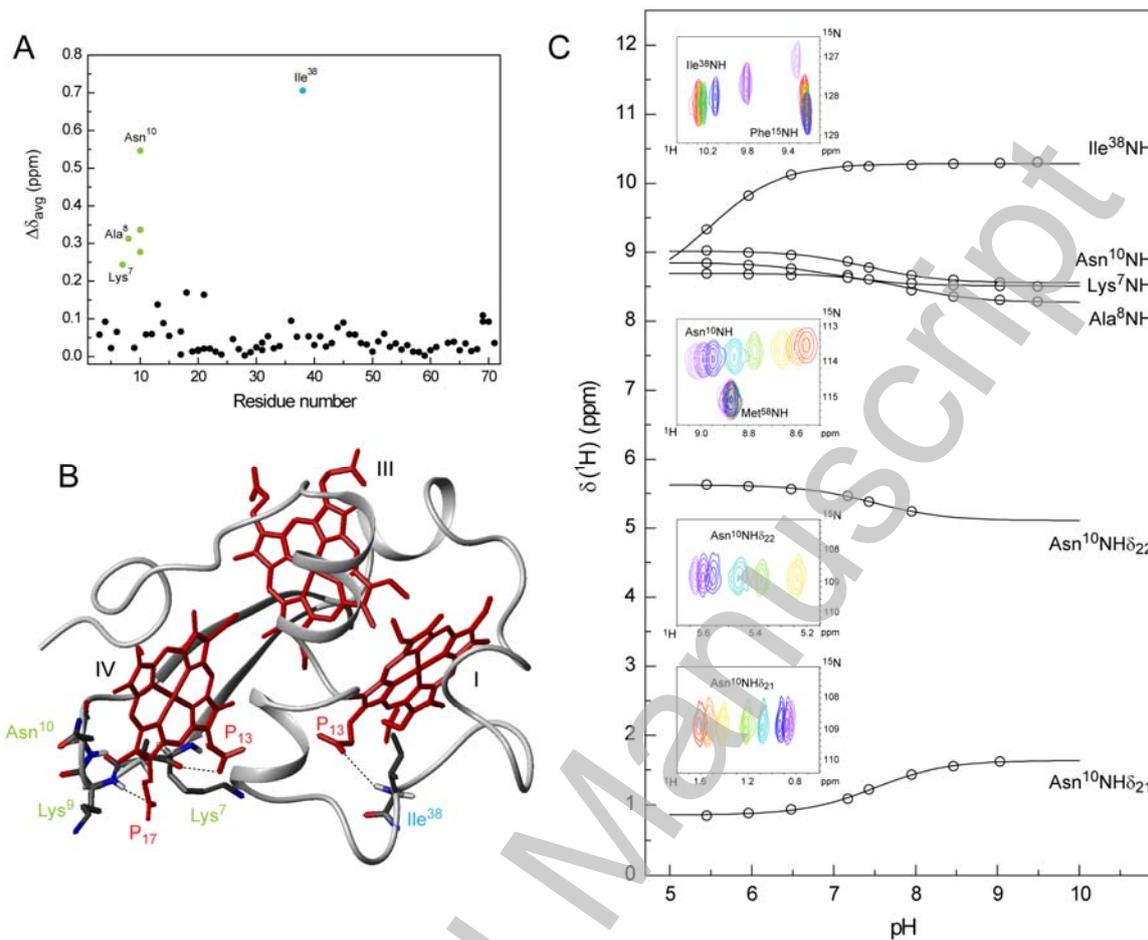


Figure 5

