The crystal structure of the pentahaem c-type cytochrome NrfB and characterisation of its solution-state interaction with the pentahaem nitrite reductase NrfA

Thomas A Clarke, Jeffrey A Cole, David J Richardson, Andrew M Hemmings

▶ **To cite this version:**

Thomas A Clarke, Jeffrey A Cole, David J Richardson, Andrew M Hemmings. The crystal structure of the pentahaem c-type cytochrome NrfB and characterisation of its solution-state interaction with the pentahaem nitrite reductase NrfA. Biochemical Journal, 2007, 406 (1), pp.19-30. 10.1042/BJ20070321 . hal-00478769

HAL Id: hal-00478769
https://hal.science/hal-00478769

Submitted on 30 Apr 2010

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.

Biochemical Journal Immediate Publication. Published on 23 May 2007 as manuscript BJ20070321

The crystal structure of the pentahaem *c*-type cytochrome NrfB and characterisation of its solution-state interaction with the pentahaem nitrite reductase NrfA

Thomas A. Clarke[1], Jeffrey A. Cole[3], David J. Richardson[1*] and Andrew M. Hemmings[1,2*]

[1]Centre for Metallospectroscopy and Biology, [1]School of Biological Sciences and [2]School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich, NR4 7TJ, United Kingdom; [3]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom.

Short title: Crystal structure of NrfB and characterisation of its interaction with NrfA

NrfB is a small pentahaem electron transfer protein widely involved in the respiratory reduction of nitrite or nitric oxide to ammonia, processes that provide energy for anaerobic metabolism in many enteric bacteria and also serve to detoxify these reactive nitrogen species. The X-ray crystal structure of *Escherichia coli* NrfB is presented at 1.74 Å resolution. The architecture of the protein is that of a 40 Å 'nanowire' in which the five haems are positioned within 6 Å of each other along a polypeptide scaffold. During nitrite reduction the physiological role of NrfB is to mediate electron transfer to another pentahaem protein, NrfA, the enzyme that catalyses periplasmic nitrite or nitric oxide reduction. Protein-protein interaction studies suggest NrfA and NrfB can form a twenty haem $NrfA_2B_2$ hetero-tetrameric complex.

Key words: nitrite reductase, ammonification, cytochrome, nitric oxide, electron transport, protein-protein interaction.

Abbreviations used: AUC, Analytical Ultracentrifugation; DLS, dynamic light scattering; HAO, hydroxylamine oxidoreductase; NrfA, cytochrome *c* nitrite reductase; NrfB, penta-haem electron donor to NrfA; Stc, small tetra-haem cytochrome *c*.

*Address correspondence to: David Richardson or Andrew Hemmings at the School of Biological Sciences, University of East Anglia, Norwich, NR4 7TJ, U.K., d.richardson@uea.ac.uk or a.hemmings@uea.ac.uk

© 2007 The Authors Journal compilation © 2007 Biochemical Society

Biochemical Journal Immediate Publication. Published on 23 May 2007 as manuscript BJ20070321

INTRODUCTION

A key reaction in the metabolism of *Escherichia coli* and related enteric pathogens, including *Salmonella enteritica*, is the reduction of nitrite to ammonium that is catalysed by cytochrome *c* nitrite reductase (NrfA)[1]. This anaerobic respiratory reaction is coupled to the conservation of energy and so contributes to the survival of enteric bacteria in low-oxygen environments, such as those of the gut. In addition to the six e^- reduction of nitrite, NrfA can catalyze the five e^- reduction of nitric oxide to ammonium and a physiological role for the reduction of NO in the detoxification of this cytotoxin has been suggested for *E. coli* [1]. The crystal structure of *E. coli* NrfA has been solved [2], along with those of the homologous enzymes from *Sulphospirillium deleyianum* [3], *Wolinella succinogenes* [4] and *Desulfovibrio desulfuricans* [5]. In all cases the structure of NrfA is that of a ~110 kDa homodimer with a total of 5 closely packed haems per monomer. By contrast to the biochemical conservation of NrfA, there are two biochemically distinct systems by which electrons required for nitrite and nitric oxide reduction are transferred through the respiratory electron transfer system to NrfA [6]. In delta and epsilon proteobacteria, such as *W. succinogenes* and *S. deleyianum*, the electron donor to NrfA is the membrane-bound tetrahaem NrfH [6,7]. This forms a membrane anchored NrfHA complex on the periplasmic face of the cytoplasmic membrane that has recently been resolved structurally from *Desulfovibrio vulgaris, w*here it was shown to form a complex comprising 2 NrfH and 4 NrfA molecules [8]. By contrast, in enteric gamma proteobacteria, such as *E. coli* and *S. enterica,* the direct electron donor to NrfA is NrfB, a water soluble 21 kDa pentahaem protein [9]. NrfB does not co-purify with NrfA and so, unlike the stable $NrfA_4H_2$ complex, the NrfAB complex is an example of a transient, rather than a stable electron transfer complex. In this work we present the first structure of NrfB in conjunction with data from solution state studies on the interaction with NrfA, demonstrating that it forms a $NrfA_2B_2$ complex which must be biochemically distinct from the $NrfA_4H_2$ complex.

EXPERIMENTAL PROCEDURES

Purification of NrfB and NrfA

NrfB was prepared using a method modified from that of [9]. The plasmid pNRFB6 was transformed into *E. coli* JM109 cells containing the pEC89 plasmid and grown overnight at 37 °C on LB agar plates containing 100 μg ml^{-1} carbenicillin and 30 μg ml^{-1} chloramphenicol. Colonies were used to inoculate 5 ml of LB medium containing 100 μg ml^{-1} carbenicillin, 30 μg ml^{-1} chloramphenicol and grown for 8 hr at 37 °C with shaking. 1 litre LB cultures containing 100 μg ml^{-1} carbenicillin and 30 μg ml^{-1}

© 2007 The Authors Journal compilation © 2007 Biochemical Society

chloramphenicol were inoculated with the 5 ml cultures and grown overnight at 37 °C with shaking. Cells were harvested by centrifugation, resuspended in 50 mM Tris-HCl pH 7.0, 1 mM EDTA, broken by the French press method and centrifuged to remove membrane fragments. NrfB in the supernatant was precipitated in 30 % (w/v) ammonium sulfate at 4 ºC and then resuspended in and dialyzed against 50 mM Tris-HCl pH 7.0. The dialysed NrfB was loaded onto a 2.5 × 20 cm Q-Sepharose Fast Flow column (Amersham) equilibrated with 50 mM Tris-HCl pH 7.0 and eluted using a 250 ml gradient from 0 - 1 M NaCl. Fractions containing NrfB were identified spectrophotometrically as having an A_{410}/A_{280} ratio of greater than 1. NrfB-containing fractions were loaded onto a 5/50 Fast Flow Phenyl Sepharose column (Amersham) equilibrated with 50 mM Tris-HCl pH 7.0, 100 mM NaCl and equilibrated with a 100 ml gradient of 0 – 30 % (v/v) isopropanol, 10 mM Tris-HCl pH 7.0. 3 ml fractions containing NrfB with an $A_{410}/A_{280} > 5$ were pooled and dialyzed into 50 mM Na-HEPES pH 7.5. As a final purification step, NrfB was passed through a Superdex 75 16/60 gel filtration column (Amersham) equilibrated with 50 mM Na-HEPES pH 7.5. NrfB eluting from the gel filtration column was judged to be pure by Coomassie stained SDS-PAGE gels and a spectrophotometric A_{410}/A_{280} ratio greater than 9. NrfA was prepared from *E. coli* K-12 strain LCB2048 [10] and purified according to established procedures [2]. The turnover number of purified NrfA was 800 NO_2^- s^{-1} at 25 °C and the NrfA concentration[1] was determined using an ε_{410} of 497 mM^{-1} cm^{-1}, based on a monomeric molecular mass of 53 kDa.

Crystallisation of NrfB, data collection and processing

Purified NrfB was concentrated to 8 mg ml^{-1} and centrifuged at 16,000 g for 10 min at 4 °C before crystallization. Crystals were obtained by the vapour diffusion method under anaerobic conditions using 250 mM tri-sodium citrate, 10 mM $Na_2S_2O_4$ and 20 % (v/v) isopropanol in 100 mM sodium Na-HEPES pH 7.5. Oxidised crystals were prepared by moving trays containing fully formed crystals to an aerobic environment at 4 ºC for 24 hours before harvesting. Crystals were judged to be oxidised on the basis of the colour change that occurs to cytochromes when becoming oxidised or reduced. For data collection, crystals were soaked with a solution of reservoir buffer containing 20 % (v/v) ethylene glycol as a cryoprotectant. The oxidised absorbance spectrum of NrfB has peaks at 410 and 530 nm that are typical of *c*-type cytochromes [9]. On reduction, the 410 nm peak decreases and is replaced by peak a 420 nm, while the 530 nm peak is split into two peaks at 520 and 550 nm. This spectral change can be visualised as a colour change from orange-red to pink-red. NrfB crystals typically had approximate dimensions of 10 x 10 x 100 μM and were too small to be analysed spectrophotometrically. However, crystals of reduced NrfB

© 2007 The Authors Journal compilation © 2007 Biochemical Society

were prepared by soaking the oxidised crystals for 1 minute in oxygen-free reservoir solution with 20 % (v/v) ethylene glycol and 50 mM $Na_2S_2O_4$ until the colour change was observed, indicating reduction, and then immediately plunged into liquid nitrogen. To ensure that reducing conditions were maintained, the reducing solution was tested after each crystal was harvested using strips of filter paper soaked in 100 mM methyl viologen.

X-ray diffraction datasets were measured using an ADSC detector on beamline ID29 at the ESRF (Grenoble). The NrfB crystals used for data collection were all of space group $P2_12_12_1$ and essentially isomorphous with typical cell dimensions a = 49.7 Å, b = 59.7 Å, c = 65.2 Å. Assuming one molecule of NrfB in the crystallographic asymmetric unit the solvent content of the crystals was approximately 45 % (v/v). Due to the small size of the NrfB crystals we were unable to measure an iron K-edge absorption spectrum and so instead collected a single-wavelength anomalous dispersion (SAD) dataset near the presumed iron K-edge at a wavelength of 1.722 Å to a resolution of 3.0 Å. Further diffraction datasets for both oxidised and dithionite-reduced NrfB were obtained from single crystals to a resolution of 1.74 Å using an X-ray wavelength of 0.976 Å. All datasets were processed using MOSFLM [11] and SCALA [12] as part of the CCP4 package [13]. Data collection statistics are summarized in Table 1.

Structure determination and refinement

The structure of NrfB was determined by SAD phasing methods using the anomalous signal originating from the five haem iron atoms present in the single molecular copy of NrfB present in the asymmetric unit. Identification of the iron atom sites and subsequent phasing at 3.5 Å resolution was performed using SHELX [14]. An electron density map calculated with these phases allowed manual building of haems 1, 3, 4 and 5 and 45 % of the amino acid backbone using COOT [15]. Phases were improved using the partial model and extended to 3.0 Å resolution with SHARP [16], allowing 66 % of the amino acid backbone to be built. The phases from the SAD dataset were extended to 2.5 Å using the oxidised NrfB dataset. Following four rounds of simulated annealing refinement at 2.5 Å resolutions against the oxidised NrfB data using CNS [17] an improved map was obtained which allowed haem 2, the amino acid side-chains and an additional 9 % building of the amino acid chain to be added to the structural model. Subsequent rounds of simulated annealing and manual intervention using COOT gave a model containing all five haems and 141 amino acids (86 % of the total) corresponding to NrfB residues 19 to 160[2]. This interim structural model had an R_{cryst} (R_{free}) of 28.4 % (32.5 %) using data in the resolution range 65 – 2.5 Å. In all cases, the R_{free} -value provided refers to an R-value calculated using 5 % of the data randomly-selected and otherwise

© 2007 The Authors Journal compilation © 2007 Biochemical Society

excluded from refinement [18]. The final refined structure of oxidised NrfB was obtained by alternating rounds of manual model building and automatic refinement using REFMAC [19] against all data to 1.74 Å resolution. This gave a structural model containing all five haems and residues 19-160 having an R_{cryst} (R_{free}) of 22.1 % (24.8 %) for data in the resolution range 60 - 1.74 Å. Addition of 198 water molecules using ARP [20] gave a final structure with an R_{cryst} (R_{free}) of 16.9 % (20.4 %) for data in the same resolution range. When analysed for stereochemical quality using PROCHECK [21] the structure has 90.9 % of residues in the most favoured regions of the Ramachandran plot with the remainder falling into additional allowed regions.

The structure of reduced NrfB was determined by molecular replacement using as a search model the final oxidised NrfB structure, but lacking water molecules [22]. This was corrected by cycles of manual model building and automatic refinement in the same fashion as used for the structure of the oxidised protein. After addition of 184 water molecules the refined structure had an R_{cryst} (R_{free}) of 15.9 % (19.2 %) for data in the resolution range 65 - 1.74 Å. The structure has 90.9 % of the residues in the most favoured regions with the remainder in additional favoured regions. As with oxidised NrfB, the final model for the reduced protein also lacks the N-terminal 18 residues and four residues from the C-terminus. N-terminal sequencing (PNAC facility, University of Cambridge) of the soluble expressed protein showed that the NrfB signal peptide is cleaved after position 25 of the protein sequence predicted from the DNA sequencing (Supplementary figure 2). The residues of the N-terminus point towards the solvent cavity in the crystal lattice, suggesting that these residues are flexible and disordered in the NrfB crystal structure. The structure coordinates have been deposited to the protein database as accession numbers 2OZY and 2POB for the oxidised and reduced proteins, respectively.

Estimation of the hydrodynamic radii of NrfA and NrfB

The hydrodynamic radius of NrfA in solution was measured using dynamic light scattering (DLS). 20 µl samples of 5 – 40 µM NrfA in 50 mM Na-HEPES pH 7.0, 2 mM $CaCl_2$ were centrifuged for 10 min at 10,000 *g* to remove dust and large aggregates. 12 µl samples were then pipetted into a sample cell and analysed at 20 °C using a Dynapro-MSX-TC equipped with a 50 mW, 843 nm solid state laser (Protein solutions, Charlottesburg, VA). The hydrodynamic radius of NrfB in 50 mM Na-HEPES pH 7.5, 1 mM CHAPS was measured in an identical manner. Two sets of data, each with more than 20 readings, were taken at various concentrations, analysed, and the hydrodynamic radii obtained using DYNAMICS version

© 2007 The Authors Journal compilation © 2007 Biochemical Society

5.0 software. The molecular weight at each concentration was determined from the hydrodynamic radius by reference to the standard curve method in the software.

Titration of the NrfA-NrfB complex using UV-visible absorbance and fluorescence emission spectroscopy

Absorbance measurements were performed on a Hitachi V-3310 dual beam spectrophotometer; the sample cuvette contained 50 mM Na-HEPES pH 7.0, 2 mM $CaCl_2$ and 1.7 μM NrfA, while the reference cuvette contained an equal volume of 50 mM Na-HEPES pH 7.0, 2 mM $CaCl_2$. After allowing the cuvettes to equilibrate to 20 °C for 5 min 15 scans were recorded of the NrfA spectrum in the wavelength range of 250 – 600 nm. Aliquots of a 200 μM NrfB solution were added to sample and reference cuvettes. The cuvettes were then left to equilibrate for 15 min before a further 15 scans were recorded. Each set of scans were averaged and the NrfB spectrum was subtracted from the NrfA plus NrfB spectrum to obtain a difference spectrum caused by a change in the environment of a haem due to complex formation. The dissociation constant of the NrfA-NrfB complex was determined by fitting the obtained data to Equation 1 using the program TABLECURVE (Systat software).

$$A = A_0\left(\frac{\left([NrfA] + [NrfB] + K_d{}^{AB}\right)}{2} - \sqrt{\frac{\left([NrfA] + [NrfB] + K_d{}^{AB}\right)^2}{4} + [NrfA][NrfB]}\right)$$

(Equation 1)

A is the observed spectral change, A_0 is the maximum theoretical spectral change, $[NrfA]$ is the concentration of monomeric NrfA, $[NrfB]$ is the concentration of the monomeric NrfB and $K_d{}^{AB}$ is the dissociation constant of the NrfA-NrfB complex.

A Perkin Elmer LS50B fluorimeter equipped with a temperature controlled water bath at 20 °C was used for fluorescence emission spectroscopy measurements. The excitation wavelength was 295 nm with a 5 nm bandwidth, and the emission wavelength measured was 348 nm with a 12.5 nm bandwidth. For each measurement the average of 30 scans were taken. A fluorescence cuvette containing 2 ml of 50 mM Na-HEPES pH 7.0, 2 mM $CaCl_2$ was used to measure a baseline. A sample of either NrfA or NrfB was then added and the change in fluorescence was measured. Experiments were performed in triplicate.

© 2007 The Authors Journal compilation © 2007 Biochemical Society

For experiments analyzing complex formation, 1.8 μM NrfB was added to the cuvette and the baseline measured, sequential additions of NrfA were then added to give a final concentration between 0.045 and 0.32 μM. The overall change in fluorescence was recorded between each addition.

Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) experiments were performed using a Beckman XLI analytical ultracentrifuge equipped with absorbance optics. The partial specific volumes of NrfA and NrfB were calculated from the amino acid sequence using SEDNTERP [23] and had values of 0.735 and 0.715 ml g^{-1} respectively. All experiments were performed in the presence of 50 mM Na-HEPES pH 7.0, 2 mM $CaCl_2$ using either NrfA or an equimolar mixture of NrfA and NrfB at a range of different concentrations. Sedimentation equilibrium experiments were performed at 20 °C, 9000 rpm, using 2, 10 or 20 μM oxidised NrfA to measure the dissociation constant of the NrfA monomer/dimer equilibrium. The concentration profiles of NrfA during analytical ultracentrifugation were measured at 410, 440 and 530 nm for 2, 10 or 20 μM NrfA respectively. Scans were recorded every four hours to determine when protein samples had reached equilibrium in the centrifuge. When equilibrium had been achieved, 10 scans were recorded for each sample. The program Ultrascan II [24] was used to fit the obtained sedimentation equilibrium profiles to both single species, as well as monomer-dimer, equilibrium models where the NrfA monomer and dimeric states are in a dynamic equilibrium and the NrfA monomer has a M_w of 53 kDa.

To measure the dissociation constant of the NrfA-NrfB complex, a stock equimolar mixture of NrfA and NrfB at 20 μM was prepared using the extinction coefficients of the haem peaks of both NrfA and NrfB [2,9]. Equilibrium experiments were executed at 7500 and 9000 rpm at a temperature of 20 °C, using concentrations of 1, 5 and 20 μM NrfA and NrfB. The concentration profiles of the NrfAB complex were measured at 410, 530 and 600 nm for 1, 5 and 20 μM NrfA respectively. When equilibrium was achieved, 10 scans were recorded for each sample. When fitting all 6 datasets simultaneously, only 3 scans for each dataset were used. The data was fitted as before to obtain sedimentation equilibrium profiles to both single species models and as monomer-dimer equilibrium models where the NrfA-NrfB heterodimer and $NrfA_2$-$NrfB_2$ heterotetramer are in a dynamic equilibrium and the NrfA-NrfB dimer has a M_w of 74 kDa. The standard deviations shown are obtained by measuring the standard deviation of the equilibrium constants measured using only one scan for each dataset.

© 2007 The Authors Journal compilation © 2007 Biochemical Society

Biochemical Journal Immediate Publication. Published on 23 May 2007 as manuscript BJ20070321

RESULTS

The 1.7 Å crystal structure of NrfB

The overall structure of the NrfB polypeptide is completely unique amongst multihaem *c*-type cytochromes. The approximate overall dimensions are 40 x 30 x 20 Å, with a chain of five haems distributed through the structure (Fig 1A, 1C). The maximum edge-to-edge distance of the haem porphyrin rings is 40 Å and the propionate side chains of all 5 haems are solvent exposed. The first observed residue at the N-terminus of the crystal structure is Asn19 and the polypeptide chain is continuous for 142 residues until Lys160 near the C-terminus of the chain of the processed protein. A haem group is covalently attached approximately every 20-25 amino acids via thiol ester linkages to the cysteines of five classical CXXCH attachment sites (Fig. 2A), so that all five of the minimum inter-porphyrin ring distances between neighbouring haems lie within 6 Å. 35% of the protein chain is folded in an α-helical manner and the remainder is extended loop regions that serve to connect the chain of haem groups (Fig 1A). The haems are arranged in a series of alternative parallel and perpendicular pairs (Fig. 1B) such that there are two parallel pairs (haems 1,2 and haems 3,4) and two perpendicular pairs (haems 2,3 and haems 4,5). Due to the close packing of the haems, the porphyrin rings of haem 2 and haem 4 are only separated by a distance of 11 Å, making it possible, in principle, for electron transfer to by-pass haem 3 (Fig. 1A). It is notable that, in making this arrangement, each set of adjacent parallel and perpendicular haem pairs can be superimposed on top of each other so that the haem pair 1, 2 can be superimposed onto haem pair 3, 4 with an r.m.s.d of 0.39 Å. The two haem elbow motifs comprising haems 2, 3 and 4, 5 can also be superposed with an r.m.s.d of 1.4 Å, where the Fe atoms of each haem are directly superimposable and the increase in r.m.s.d is caused by slight changes in the relative angle of the haems. This reflects the conservation of these haem-pairing arrangements, which is maintained despite quite different protein loops connecting the haem pairs in the two halves of the protein. (Fig. 1A). It is also possible to superimpose haem triads, such that haems 1,2,3 can be superimposed onto haems 3,4,5 with an r.m.s.d of 1.9 Å (Fig. 1D). The other two potential haem triad combinations, superimposing haems 1,2,3 with haems 2,3,4, or haems 2,3,4 with 3,4,5, have much higher r.m.s.d (5.3 Å and 3.8 Å, respectively). This is due to the difference in the position of the parallel haem pairs compared to the position of the haem elbow.

There are only a few significant changes between the structure derived from untreated air-oxidised crystals and that derived from crystals pre-reduced with dithionite immediately before data collection (Fig. 1C). The alpha-carbons of the main chain backbones of NrfB deviate with an r.m.s.d of only 0.20 Å.

© 2007 The Authors Journal compilation © 2007 Biochemical Society

However, two loops containing residues 21 to 30 and 106 to 112 deviate by 0.46 Å and 0.70 Å r.m.s.d., respectively. Both loops are on the molecular surface close to haems 2 and 3 (Fig. 1C) and the change in the NrfB crystal structure is best described as a slight contraction in the surface of the protein near haems 2 and 3. The propionate groups of the haems are also changed slightly in position, most significantly in haem 5 where one propionate group moves closer to the haem iron.

All five haem iron ions are axially coordinated by two histidine ligands. The resolution of the electron density map is sufficient to identify the orientation of the imidazole rings with respect to the haem-iron ligand (Fig. 1E). The imidazole planes of the two ligands are either near-parallel or twisted into a more perpendicular arrangement (Fig. 1B). Previous studies showed that the EPR spectrum of purified oxidised NrfB exhibited two main features: a rhombic signal at 2.99, 2.27 and 1.5 that was attributed to low spin haems where the histidine imidazole rings were near-parallel, and a weak 'large g_{max}' signal with a g-value of 3.57 attributed to haems where the imidazole rings were in a near perpendicular arrangement [9]. The 'large gmax' EPR signal is well documented [25] and results from the near-degenerate state of the d_{xz} and d_{yz} orbitals, which is caused by the perpendicular arrangement of the planar imidazole rings. In contrast, the near-parallel arrangement of imidazole rings results in a difference in energy between the d_{xz} and d_{yz} orbitals which results in a rhombic EPR signal. Quantification of the rhombic and large g_{max} signals gave values of 2 spin $NrfB^{-1}$ and 3 spin $NrfB^{-1}$ respectively, in complete agreement with the orientation of the imidazole ring planes observed in the crystal structure. Each proximal ligand is contributed by the CXXCH motif anchoring it to the protein and each distal ligand arises from one of the five histidines that are conserved throughout the NrfB family of proteins (Fig. 2A). The histidine axial ligands provide important determinants to the final folding with the following pairings: His64 - His28; His57 – His120; His92 – His39; His117 – His104 and His142 – His107 (Fig. 2). These histidines are completely conserved throughout the NrfB family, suggesting that NrfB proteins from different bacteria share a common structure. It is possible that the distal histidine ligands are important for protein stability, as they are clustered on one side of the protein which is predominantly comprised of loops.

The C-terminal haem 5 is the most solvent- exposed of all the haems, contributing 300 $Å^2$ of the surface of the protein (Fig. 2 B1). The next most exposed is the N-terminal haem 1, contributing 216 $Å^2$. The surface exposures of the middle haems 2, 3 and 4 are 132, 138 and 140 $Å^2$, respectively. These figures suggest routes of electron entry/exit to the NrfB haem-wire via haems 1 and 5. However, the non-negligible solvent exposures of haems 2, 3 and 4 opens up the possibility for electron input or egress at multiple sites in the protein. By mapping the most conserved residues of the NrfB amino acid sequence onto the surface

© 2007 The Authors Journal compilation © 2007 Biochemical Society

of the *E. coli* NrfB structure, structurally important surface regions can be identified (Fig. 2B). The area around haem 5 is predominantly conserved, suggestive of a potential protein - protein binding site. This region includes an exposed conserved tryptophan residue, Trp102 (see below). The next most conserved region is that around the surface of haem 1, suggesting a potential second binding site. This supports the hypothesis that haems 1 and 5 are the likely sites of electron input to or output from NrfB.

The solution state interaction of NrfB with NrfA

E. coli NrfB contains only one tryptophan residue and this is conserved in NrfB from other bacteria, including *Pasteurella multocida*; *Haemophilus influenzae*; *Shigella flexneri*; *Salmonella enteritis*; *Erwinia carotovora; Vibrio alginolyticus* (Fig. 2A). This tryptophan is positioned on the surface near haem 5 in the refined crystal structure (Figs. 1A, 2B). The fluorescence of this tryptophan can be used to reveal changes in the vicinity of haem 5. The fluorescence of NrfA and NrfB was monitored using an excitation wavelength of 295 nm and the corresponding emission at 348 nm. In control experiments, as the concentration of NrfB increases from 0.05 to 0.28 μM, the fluorescence intensity increases linearly As the concentration of NrfA increases from 0.05 to 0.31 μM, the fluorescence also increases linearly, although the observed fluorescence is only 20 % that of the NrfB fluorescence at the same concentration. However, when NrfA is added to a fluorescing solution of NrfB, the total observed fluorescence initially decreases (Fig. 3B). This decrease continues until NrfA and NrfB are present at equimolar concentrations (0.18 μM in Fig. 3B). Further additions of NrfA then lead to an increase of fluorescence in a manner expected for the additive effects of free NrfA. This suggests that the two proteins form a complex that leads to a decrease in the fluorescence of the NrfB tryptophan. This decrease in fluorescence continues until a NrfA:NrfB ratio of 1.0 is achieved, the fluorescence then increases for a situation in which there is a NrfA-NrfB complex and excess free NrfA, showing that the change in the tryptophan environment is consistent with the formation of an equimolar NrfA-NrfB complex.

UV-visible spectroscopy was also used to study this complex formation. Although NrfA and NrfB have strongly overlapping UV-visible spectra, a small change occurs when increasing amounts of NrfB are added to NrfA and this can be resolved through use of difference spectra (Fig. 3C). It is impossible to attribute this change to a particular protein or haem group as the change is caused by mixing two pentahaem proteins. Control experiments in which NrfB was added to buffer, or buffer added to NrfA, did not cause the observed spectral shift. As the amount of NrfB added increases, the amplitude of the change between 410 and 420 nm increases until the NrfA concentration is equivalent to the NrfB concentration and

© 2007 The Authors Journal compilation © 2007 Biochemical Society

the amplitude of the spectral change remains constant (Fig. 3). This plateau is reached at equimolar concentrations of 0.17 μM NrfA and NrfB in Fig 3D and the data were fitted using equation 1 to give a dissociation constant (K_d^{AB}) of 37 ± 23 nM for the dissociation of a NrfA-NrfB complex with a stoichiometry of 1: 1. We recognise that this represents an upper limit as it is necessary to use concentrations much higher than the dissociation constant to observe the subtle change in absorbance. Taken together the fluorescence and UV-visible studies indicate that NrfB forms a tight stable complex with NrfA with a likely stoichiometry of either 1:1 or, as the crystal structure shows NrfA to be dimeric, 2:2.

Having gained an insight into the stoichiometry of the NrfA-NrfB electron transfer complex, the effect of NrfB on NrfA self-association was investigated. First, the self-association of air-oxidised NrfA was assessed by measuring the hydrodynamic radius using dynamic light scattering (DLS). The range of measured hydrodynamic radius broadens and the percentage polydispersity decreases as the concentration of NrfA increases (Table 2). A low polydispersity value indicates a homogenous sample. Thus, the data suggests that oxidised NrfA is in a dynamic equilibrium of monomer and dimer over the concentration range 5-38 μM and that, as the concentration of NrfA increases, so does the proportion of dimeric NrfA at equilibrium. Estimating the molecular weight (M_w) of the protein in solution from the hydrodynamic radius measured yielded a value of 102 kDa at NrfA concentrations ≥ 9 μM (Table 2). This is similar to the calculated molecular weight of the $(NrfA)_2$ complex (106 kDa). However, as the concentration of NrfA decreases below 9 μM, the apparent M_w changes to 62 kDa, close to the molecular weight calculated for the NrfA monomer. The sudden decrease in molecular weight observed between 9 and 7.5 uM is probably most likely to small changes in the concentration of monomeric and dimeric species in solution causing artificially large changes in the measured molecular weight distribution and this is reflected in the increasing polydispersity. DLS analysis of samples of NrfB over a similar concentration range showed that NrfB did not dimerise. The dissociation constant of the NrfA homodimer into its component monomers (K_d^{AA}) was measured using sedimentation equilibrium experiments on air-oxidised NrfA at 2, 10 and 20 μM concentration. The absorbance of the NrfA haem groups was used to measure the concentration of NrfA across the ultracentrifuge cell (note that the haem : protein ratio remains constant regardless of the oligomeric state of the protein and we established that the absorbance properties were not changed through the monomer-dimer transitions). When NrfA was subjected to centrifugation at 9000 rpm a stable concentration gradient across the ultracentrifuge cell was formed after approximately 12 hrs (Fig. 4). This gradient is directly dependent on the averaged molecular weight ($M_{w(av)}$) of NrfA in solution and this value

© 2007 The Authors Journal compilation © 2007 Biochemical Society

increased with increasing concentration in a similar manner to that observed in the DLS experiments, confirming that dissociation of the $(NrfA)_2$ complex occurs over the concentration range of 2 to 20 μM. Ultrascan II [24] was used to simultaneously fit the data at all three concentrations and the best fit to the data (Fig. 4A) yielded a K_d of 4.0 ± 0.7 μM for the dissociation of the $(NrfA)_2$ deca-haem complex into pentahaem monomers, this value was consistent with the data obtained from DLS experiments (Table 2).

Sedimentation equilibrium experiments on solutions containing equal amounts of NrfA (53 kDa) and NrfB (21 kDa) at 2, 5 or 20 μM showed that the $M_{w(av)}$ of the NrfA-NrfB sample increased significantly when compared with NrfA alone. The $M_{w(av)}$ in 20 μM mix of each protein (147 ± 9 kDa) was similar to that expected of a 148 kDa $NrfA_2B_2$ heterotetramer (Table 3). The sedimentation equilibrium profiles of the NrfAB complex (Fig. 4B) were simultaneously fitted to a monomer-dimer equilibrium with a NrfAB heterodimer molecular weight of 74 kDa to yield a dissociation constant of 4.1 ± 0.2 μM for the 2(NrfAB) ↔ $NrfA_2B_2$ equilibrium. This is similar to the K_d of 4.0 μM obtained for the 2NrfA ↔ $NrfA_2$ equilibrium. The data implies that formation of the $NrfA_2B_2$) complex does not affect the dissociation at the $NrfA_2$ dimer, suggesting that NrfB binds at a site that is remote from this NrfA interface.

DISCUSSION

In this paper we have presented the first crystal structure of NrfB, a periplasmic pentahaem electron transfer protein. NrfB is involved in shuttling electrons from an integral membrane quinol-oxidising protein complex, NrfC-NrfD, to a pentahaem nitrite reductase (NrfA) and is conserved in many enteric pathogens. We have presented two crystal structures of NrfB, one derived from data collected from crystals that were in an 'air-oxidised' state immediately prior to exposure to X-rays and another derived from data collected from crystals that were treated with dithionite reductant immediately prior to exposure to X-ray irradiation. Whilst the colour of the crystals indicated that the two treatments yielded NrfB haems in different redox states, we recognise that the level of photoreduction that occurs to the 'air-oxidised' NrfB through exposure to X-rays during data collection is not known so that it can not be excluded that both structures represent reduced forms of the protein. Importantly, however, the crystallographically-derived coordination structure of the five NrfB haems is exactly as predicted from solution and frozen-solution MCD and EPR spectroscopy [9]. These two methods together show that all five ferric haems are low-spin hexacoordinate species with *bis*-histidinyl axial coordinate. This is also the case in both crystal structures. Furthermore the EPR analyses showed that in three of the NrfB haems the imidazole rings of the histidine ligands are near-

© 2007 The Authors Journal compilation © 2007 Biochemical Society

perpendicular, whilst in the remaining two they are near-parallel. This is also the case in both NrfB crystal structures. This then gives confidence that the coordination environment of the NrfB haems in the crystal structures reflects that in solution.

A number of lines of evidence are also presented that suggest NrfA and NrfB form a 1:1 complex in solution. The mass of this complex (~150 kDa), determined by AUC indicate that it is a twenty haem $NrfA_2B_2$ heterotetramer. We have previously determined the structure of NrfA, which crystallizes as a 106 kDa homodimer with a total of 10 closely packed haems [2] (Supplementary Fig. 1). Four of the haems in each monomer are *bis*-His ligated, while the active site haem is ligated by a proximal lysine ligand and, depending whether a substrate such as nitrite or hydroxylamine is included in the crystallization conditions, hydroxide or substrate on the distal side. Much of the homodimer interface is formed from a pair of long, interacting helices, at the end of which, partially exposed to solvent, is haem 5. The distance between the haems 5, one from each NrfA monomer, is only 4 Å [edge-edge]. This distance of separation shows that electrons could easily travel from one monomer to another by exchange through these haems [2]. In the present work we have established the equilibrium dissociation constant for 2(NrfA) ↔ $NrfA_2$ and show that this is the same as for the 2(NrfAB) ↔ $NrfA_2B_2$ equilibrium. This suggests that the binding site on NrfA for NrfB does not overlap with the NrfA dimer interface. Such a site could be provided around NrfA haem 2, which is the most solvent exposed haem [2]. Thus we feel that any model for electron transfer between NrfB and NrfA would have to bring NrfB haem 5 (the putative electron egress haem) to within 14 Å of NrfA haem 2.

Whilst the present work was in progress the structure of a $NrfA_4H_2$ complex for *D. vulgaris* was published [8]. In contrast to NrfAB, NrfH and NrfA form a very tight 'hard-wired' interaction, indeed NrfH has never been purified in a native form independent from its cognate NrfA. The crystal structure provides some insight into this tight interaction since two of the NrfA proteins provide lysine ligands to one of the haems on each NrfH protein in the $NrfA_4H_2$ complex. We have not been able to purify a NrfAB complex, or co-crystallize such a complex but, as discussed above, our solution-state protein-protein interaction studies strongly suggest the NrfA and NrfB form a 2:2 complex that will be structurally distinct from the $NrfA_4H_2$ complex. However, our structure of NrfB also reveals that NrfH and NrfB are structurally distinct in many ways so that there is no reason to expect *a priori* that the NrfAB and NrfAH stoichiometries should be the same. For example: (i) the sequence identity between NrfB and NrfH proteins is very low (~15%) and this is largely contributed to by alignment of four CXXCH motifs (Supplementary Fig. S2); (ii) NrfH has a transmembrane helix located towards the N-terminus, whereas NrfB is a soluble

© 2007 The Authors Journal compilation © 2007 Biochemical Society

periplasmic protein in which the N-terminal signal peptide has been cleaved; and (iii) the polypeptide chain of NrfB and the soluble domain of NrfH can not be superimposed. Despite this, four of the NrfB haems (haems 1-4) adopt similar haem-haem packing motifs to the four NrfH haems (Fig. 5) and can be superimposed with an r.m.s.d of 1.93 Å. However, two of these four superimposable haems have different haem iron ligation. In NrfB all five haem irons are low spin hexacoordinate with *bis*-histidinyl axial ligation. However, in the $NrfA_4H_2$ complex, NrfH haem 1 is pentacoordinate (and therefore likely to be high spin) with a Met distal ligand and an aspartate occupying the proximal position, but not being within bonding distance. NrfH haem 4 is ligated by His 140 provided by NrfH and Lys 331 from NrfA. Whilst we are unable to provide a structure of NrfB in complex with NrfA, it should be noted that the EPR spectrum of a $NrfA_2B_2$ complex is simply the sum of the NrfA and NrfB spectra and thus does not indicate that any substantial change occurs in haems visible by EPR on complex formation (data not shown). A change to the EPR spectra would be indicative of either changes in haem coordination, such as a change from His-His to His-Met, or ligand orientation such as movement of the planar imidazole rings from a perpendicular to parallel arrangement. As no such changes are observed in the EPR spectrum of the $NrfA_2B_2$ complex it seems likely that the environment of the NrfA and NrfB haems remains constant during complex formation.

The major difference between NrfB and NrfH is the presence of haem 5 in NrfB that is absent in the tetrahaem NrfH (Supplemental Fig. S2). This haem is the most likely electron output site for NrfB and its presence makes the construction of a NrfAB complex model using a NrfAH structural template impossible. However, it is notable that the position of NrfB haem 5 has equivalents in some other multi-haem cytochrome systems. One of these is the 12 kDa 'small tetrahaem cytochrome' (STC) of *Shewanella species* that is involved in periplasmic electron transfer and respiration on soluble Fe (III) [26]. Haems 1-4 of STC overlay onto haems 2-5 of NrfB with an r.m.s.d of 1.20 Å (Fig. 5), which reveals clear similarities between these two small periplasmic electron transfer proteins. Perhaps more intriguing though, it has previously been noted that haems 1-5 of NrfA can be superimposed onto the haems 4-8 of the octahaem hydroxylamine oxidoreductase (HAO) subunit with an r.m.s.d of 1.1 Å [3, 27]. We can now show that the remaining haems 1-3 of HAO can be superimposed onto haems 3-5 of NrfB with an r.m.s.d of 1.46 Å (Fig. 5). This introduces the possibility that the octahaem HAO polypeptide, rather than the NrfHA complex, may provide a better conceptual template for a model for the NrfAB arrangements. Such a model would be consistent with the suggestion raised earlier that NrfB might dock onto NrfA such that NrfB haem 5 approaches NrfA haem 2 (Fig. 5). Previously, NrfA haem 2 was identified as the site of electron input into

© 2007 The Authors Journal compilation © 2007 Biochemical Society

NrfA [2] due to a conserved region in NrfA sequences which employ NrfB as an electron donor. This conserved region is absent in NrfA proteins that use NrfH as an electron donor. It should be noted that to achieve the HAO-type of haem arrangement for NrfAB, a rearrangement of the NrfA structure involving two flexible α-helices located at the surface near haem 2 (Supplemental Fig. S2) would be required. These structural elements have temperature factors that are amongst the highest for residues in the NrfA structure and are therefore predicted to be mobile. A number of significant structural changes also occur in the NrfH-dependent NrfA proteins on binding NrfH and so a degree of structural re-arrangement of NrfB-dependent NrfA on binding NrfB cannot be excluded.

In closing this discussion of the NrfB structure we move away from electron egress from haem 5 and the interaction with NrfA to consider electron input. Haem 1 of NrfB superimposes onto haem 1 of NrfH (Fig. 5). In NrfH this is the likely primary acceptor haem for electrons extracted from quinol through the quinol dehydrogenase activity of the protein. It is also then reasonable to argue that this is the most likely electron input site in NrfB. However, NrfB is not a quinol dehydrogenase. Unlike the $NrfA_4H_2$ electron transport complex that binds a complete 'quinol-to-nitrite' electron transport chain. The $NrfA_2B_2$ complex has to be interfaced to the quinol pool by yet another redox protein complex, the NrfCD complex that is predicted to bind four iron sulphur clusters. Thus the crystal structure of this complex is required before the full 'quinol-to-nitrite' electron transport chain of the NrfB-dependent cytochrome c nitrite reductase system can be fully understood.

ACKNOWLEDGEMENTS

We are grateful to Ann Reilly and Christine Moore for excellent technical support and to Drs Jörg Simon, Julea Butt, Myles Cheesman and James Gwyer for helpful discussions. This work was supported by: BBSRC grants B18695 to DJR, AMH, and JAC; Wellcome Trust JIF grant 0162178 to DJR and AMH; and the U.S. Dept. of Energy Office of Biological and Environmental Research under the Genomics-Genomes to Life Program and an EMSL Scientific Grand Challenge project.

REFERENCES

© 2007 The Authors Journal compilation © 2007 Biochemical Society

1. Poock, S.R., Leach, E.R., Moir, J.W., Cole, J.A. and Richardson D.J. (2002) Respiratory detoxification of nitric oxide by the cytochrome c nitrite reductase of Escherichia coli. J. Biol. Chem. **277,** 23664-23669.
2. Bamford, V.A., Angove, H.C., Seward. H.E., Thomson, A.J., Cole, J.A., Butt, J.N., Hemmings, A.M. and Richardson, D.J. (2002) Structure and Spectroscopy of the Periplasmic Cytochrome c Nitrite Reductase from Escherichia coli. Biochemistry **41,** 2921-2931.
3. Einsle, O., Messerschmidt, A., Stach, P., Bourenkov, G. P., Bartunik, H. D., Huber, R. and Kroneck, P. M. (1999) Structure of cytochrome c nitrite reductase. Nature **400**, 476-480.
4. Einsle, O., Stach, P., Messerschmidt, A., Simon, J., Kroger, A., Huber, R. and Kroneck, P. M. (2000) Cytochrome c nitrite reductase from Wolinella succinogenes. Structure at 1.6 A resolution, inhibitor binding, and heme-packing motifs. J. Biol. Chem. **275,** 39608-39616.
5. Cuhna, C. A., Macieira, S., Dias, J. M., Almeida, G., Goncalves, L. L., Costa, Lampreia, J., Huber, R., Moura, J. J. G., Moura, I. and Romao M. J. (2003) Cytochrome c nitrite reductase from Desulfovibrio desulfuricans ATCC 27774. The relevance of the two calcium sites in the structure of the catalytic subunit (NrfA). J. Biol. Chem. **278,** 17455-17465.
6. Simon, J. (2002) Enzymology and bioenergetics of respiratory nitrite ammonification. FEMS Microbiol. Rev. **26,** 285-309.
7. Schumacher, W., Hole, U. and Kroneck P.M.H (1994) Ammonia-forming cytochrome c nitrite reductase from Sulfurospirillum deleyianum is a tetraheme protein: new aspects of the molecular composition and spectroscopic properties. *Biochem. Biophy. Res. Commun.* **205,** 911-916.
8. Rodriques, M.L., Oliveira, T.F., Pereira I.A.C. and Archer, M. (2006) X-ray structure of the membrane-bound cytochrome c quinol dehydrogenase NrfH reveals novel haem coordination. EMBO J. **25**, 5951-5960
9. Clarke, T.A., Dennison, V., Seward, H., Burlat, B., Cole, J.A., Hemmings, A.M. and Richardson, D.J. (2004) Purification and spectropotentiometric characterization of Escherichia coli NrfB, a decaheme homodimer that transfers electrons to the decaheme periplasmic nitrite reductase complex. J. Biol. Chem. **279,** 41333-41339.
10. Potter, L. C. and Cole, J. A., (1999) Essential roles for the products of the napABCD genes, but not napFGH, in periplasmic nitrate reduction by Escherichia coli K-12. Biochem. J. **344,** 69-76.
11. Leslie, A.G.W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography. **26.**

© 2007 The Authors Journal compilation © 2007 Biochemical Society

12. Evans, P. (2006) Scaling and assessment of data quality. Acta Cryst. D. **62**, 72-82.

13. Collaborative Computational Project, Number 4. (1994) The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst. D **50,** 760-763.

14. Sheldrick, G. M. (1990). Phase annealing in *SHELX*-90: direct methods for larger structures. Acta Cryst. A **46**, 467-473.

15. Emsley, P. and Cowtan, K. (2004) Coot: Model-Building Tools for Molecular Graphics. Acta Cryst D**. 60**, 2126-2132.

16. Bricogne, G., C. Vonrhein, C. Flensburg, M. Schiltz. and W. Paciorek. (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Cryst. D **59,** 2023-30.

17. Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R.J., Rice, L. M., Simonson, T. and Warren, G. L., Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination (1998) Acta Cryst. D. **54,** 905-921.

18. Brunger A.T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature **355,** 472-475

19. Murshudov, G. N., Vagin A. A. and Dodson, E. J.(1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Cryst.D **53**, 240-255.

20. Cohen, S. X., Morris, R. J., Fernandez, F. J., Jelloul, M. B., Kakaris, M., Parthasarathy, V., Lamzin, V.S., Kleywegt, G.J. and Perrakis, A. (2004). Towards complete validated models in the next generation of *ARP*/*wARP.* Acta Cryst. D **60**, 2222-2229.

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

22. Vagin, A. and Teplyakov, A. (2000) An approach to multi-copy search in molecular replacement. Acta Cryst. D. **56**, 1622-1624.

23. Hayes, D. B., Laue, T. and Philo, J. (1995-2005) SEDNTERP- Sedimentation interpretation program. Version 1.05, NH, USA; University of New Hampshire.

24. Demeler, B. (2005) UltraScan: A Comprehensive Data Analysis Software Package for Analytical Ultracentrifugation Experiments. In Modern Analytical Ultracentrifugation: Techniques and Methods. (Scott, D. J., Harding, S. E. and Rowe, A.J. Eds.) pp210-229, Royal Society of Chemistry (UK)

© 2007 The Authors Journal compilation © 2007 Biochemical Society

Biochemical Journal Immediate Publication. Published on 23 May 2007 as manuscript BJ20070321

25. Walker, F. A. (1999) Magnetic spectroscopic (EPR, ESEEM, Mössbauer, MCD and NMR) studies of low-spin ferriheme centers and their corresponding heme proteins. Coord. Chem Rev. **186,** 471-534

26. Leys, D., Meyer, T. E., Tsapin, A. S., Nealson, K. H., Cusanovich, M. A. and Van Beeumen J. J. (2002) Crystal structures at atomic resolution reveal the novel concept of "electron-harvesting" as a role for the small tetraheme cytochrome c. J. Biol. Chem. **277,** 35703-35711.

27. Igarashi, N., Morigama, H., Fujiwara, T., Fukomon Y. and Tanaka, N. (1997) The 2.8 A structure of hydroxylamine oxidoreductase from a nitrifying chemoautotrophic bacterium, Nitrosomonas europaea. Nat. Struct. Biol. **4,** 276-289.

FOOTNOTES

[1]Unless otherwise stated, the concentration of NrfA refers to the total concentration of the 53 kDa NrfA subunit.

[2]Unless otherwise stated the amino acid numbering refers to the processed protein without the signal peptide, which is experimentally determined to start at position 26 on the pre-protein sequence.

FIGURE LEGENDS

Figure 1. Crystal structures of NrfB. All figures showing protein structures were prepared using PYMOL (DeLano Scientific) (A) The crystal structure of *E. coli* NrfB represented as the α-carbon trace with 5 haems and tryptophan 102 displayed in green. (B) Haem organization within NrfB, histidines sidechains are shown in blue. The distances (in Å) between haems are shown above; the haem numbers are shown underneath. (C) Superimposition of the oxidised and reduced haem structures. The α-carbon trace and haems of the oxidised and reduced enzymes are shown in red and blue, respectively. (D) Superimposition of haems 1, 2, 3 (blue) over haems 3, 4, 5 (green) with an r.m.s.d. of 1.9 Å. The haem numbering refers to the green haems. (E) Region of the electron density map contoured at 1.1 σ calculated from the final refined structure of oxidised NrfB in the region of haem 5, 4, 3 and trytophan 102. The figure demonstrates the quality of the fit of the structural model to the map.

Figure 2. Amino acid sequence conservation in NrfB primary and tertiary structures. (A) An amino acid sequence alignment of *E. coli* NrfB (residues 14-163) with NrfBs from: *Pasteurella multocida*

© 2007 The Authors Journal compilation © 2007 Biochemical Society

*(*P. mult); *Haemophilus influenzae (*H. infl); *Shigella flexneri* (S. flex); *Salmonella enteritis* (S. ente); *Erwinia carotovora* (E. caro)*; Vibrio alginolyticus* (V. algi). Red-filled boxed amino acids are identical in the alignment, unfilled boxed amino acids are similar in the alignment. The secondary structure of *E. coli* NrfB is indicated above the sequence alignment. The CXXCH motifs are indicated by coloured triangles under the alignment according to: haem 1, red; haem 2, yellow; haem 3, green; haem 4, blue; haem 5, magenta. The corresponding distal histidines of each haem are also coloured accordingly. (B) Surface map of the side (panel right) and C-terminal end (panel left) of NrfB. Completely conserved residues are shown in red, haems 5, 2 and 1 are shown in green. All other residues are shown in white.

Figure 3. Complex formation between NrfA and NrfB measured by fluorescence and absorption spectroscopy. (A) Dependence of the relative fluorescence of NrfB (■) or NrfA (●) on protein concentration. 100% represents the maximum fluorescence obtained with 0.31 μM NrfB. (B) Titration of 0.18 μM NrfB with 0 - 0.31 μM NrfA. The percentage scale on the Y-axis represents the change in fluorescence obtained, where 100% is the maximum observed in the experiment (0.18 μM NrfB in the absence of NrfA) and 0% is the minimum observed in the experiment (0.18 μM NrfB in the presence of 0.18 μM NrfA). In all plots each data point is the average of three separate experiments with error bars derived from the standard deviations, (C) Shift in the UV-visible spectrum following addition of 0.4 μM (▬), 0.8 μM (---) or 2.4 μM (─) NrfB to a solution containing 1.7 μM NrfA. (D) The amplitude of the change in absorbance shown in 'C' plotted as a function of NrfB concentration; the line was fitted iteratively using Equation 1 as described in methods.

Figure 4. Sedimentation equilibrium analyses of NrfA and NrfB. (A) Dissociation of the NrfA dimer in solution as monitored by sedimentation equilibrium. NrfA in 50 mM Na-HEPES pH 7.0, 2 mM $CaCl_2$ was centrifuged at 9000 rpm at 20 °C for 24 hours until equilibrium had been achieved. *Lower panel:* The absorbance profiles of NrfA at 2 μM (○), 10 μM (□) and 20 μM (Δ) were all fitted to a model where the NrfA dimer dissociated into 53 kDa monomers with a K_d^{AA} of 4.0 ± 0.7 μM. *Upper panel*: residual difference between the experimental data and the fitted curves. (B) Dissociation of the NrfA-NrfB tetramer into NrfA-NrfB heterodimers. NrfA-NrfB at 1, 5 and 20 μM was centrifuged at 7,500 or 9000 rpm at 20 °C until equilibrium was achieved, *Lower panel*: Absorbance profiles of the NrfA-NrfB complex at 1 μM (○), 5 μM (□) and 20 μM (Δ) at both 7,500 and 9000 rpm were fitted to a model where the NrfA-NrfB tetramer dissociated into 74 kDa NrfA-NrfB heterodimers with a K_d of 4.1 ± 0.2 μM. *Upper panel* : residual difference between the experimental data and the fitted curves.

© 2007 The Authors Journal compilation © 2007 Biochemical Society

Figure 5. Comparison of the haem packing motifs of NrfB and NrfA (PDB entry 1GU6) with $NrfA_2H$ (PDB entry 2J7A), HAO (PDB entry 1FGJ) and Stc (PDB entry 1M1Q). In the case of the putative NrfAB complex the NrfB haems are shown in blue, the NrfA haems are shown in green and the numbers denote NrfA haems 2 and 5. Dimerisation of NrfA occurs at the haem 5 interface (see supplementary Figure S2)

© 2007 The Authors Journal compilation © 2007 Biochemical Society

Table 1. Data for collection and refinement statistics for oxidised and reduced NrfB

	SAD data	Oxidised	Reduced
Data collection [a]			
Wavelength (Å)	1.722 Å	0.976 Å	0.976 Å
Resolution (Å)	65 – 3.0 (3.16 - 3.0)	44 - 1.74 (1.83 - 1.74)	44 – 1.74 (1.83 - 1.74)
Unique reflections	4106 (551)	19836 (2915)	19070 (2846)
Completeness (%)	98.6 (93.9)	97.4 (99.3)	94.0 (97.0)
Anomalous Completeness (%)	98.4 (92.7)	-	-
R_{sym} (%)	12.6 (22.1)	12.2 (45.0)	13.1 (43.3)
$\langle I/\sigma \rangle$	13.9 (6.8)	4.8 (1.6)	4.7 (1.5)
Multiplicity	6.8 (6.0)	2.8 (2.7)	4.2 (4.0)
0Anomalous Multiplicity	3.8 (3.3)	-	-
Average atomic B-factor (A^2)	42.3	11.2	8.5
Refinement [a,b]			
R_{cryst}		16.9 (24.9)	15.9 (20.4)
R_{free}		20.4 (29.7)	19.2 (30.0)
Model			
Protein atoms		1159	1153
Water		198	184
Haem atoms		215	215
Bond length rmsd (Å)		0.013	0.012
Bond angle rmsd (°)		1.50	1.51
Average B-factor ($Å^2$)		13.7	10.9

[a] Values in parentheses indicate the highest resolution shell.

[b] $R = | F_o - F_c | / F_o$. R_{cryst} is calculated with the 95 % of data used during refinement.

R_{free} is calculated with a 5 % subset of data not used during refinement.

© 2007 The Authors Journal compilation © 2007 Biochemical Society

Table 2. Dynamic light scattering analyses of the dimerization of NrfA and NrfB. Experiments were performed at 20 ºC using solutions containing 5 to 38 μM NrfA or 11 to 45 μM NrfB in 50 mM Na-HEPES pH 7.0, 2 mM $CaCl_2$.

[NrfA] (μM)	R_H (nm)	M_w (kDa)	Polydispersity (%)
38	4.2 ± 0.2	102 ± 7	29 ± 11
9	4.2 ± 0.2	102 ± 7	51 ± 31
7.5	3.4 ± 0.2	63 ± 9	72 ± 19
5	3.7 ± 0.2	62 ± 13	91 ± 11
[NrfB] (μM)			
11	2.5 ± 0.1	30 ± 6	13 ± 10
21	2.5 ± 0.1	27 ± 6	26 ± 15
45	2.6 ± 0.1	30 ± 6	28 ± 19

© 2007 The Authors Journal compilation © 2007 Biochemical Society

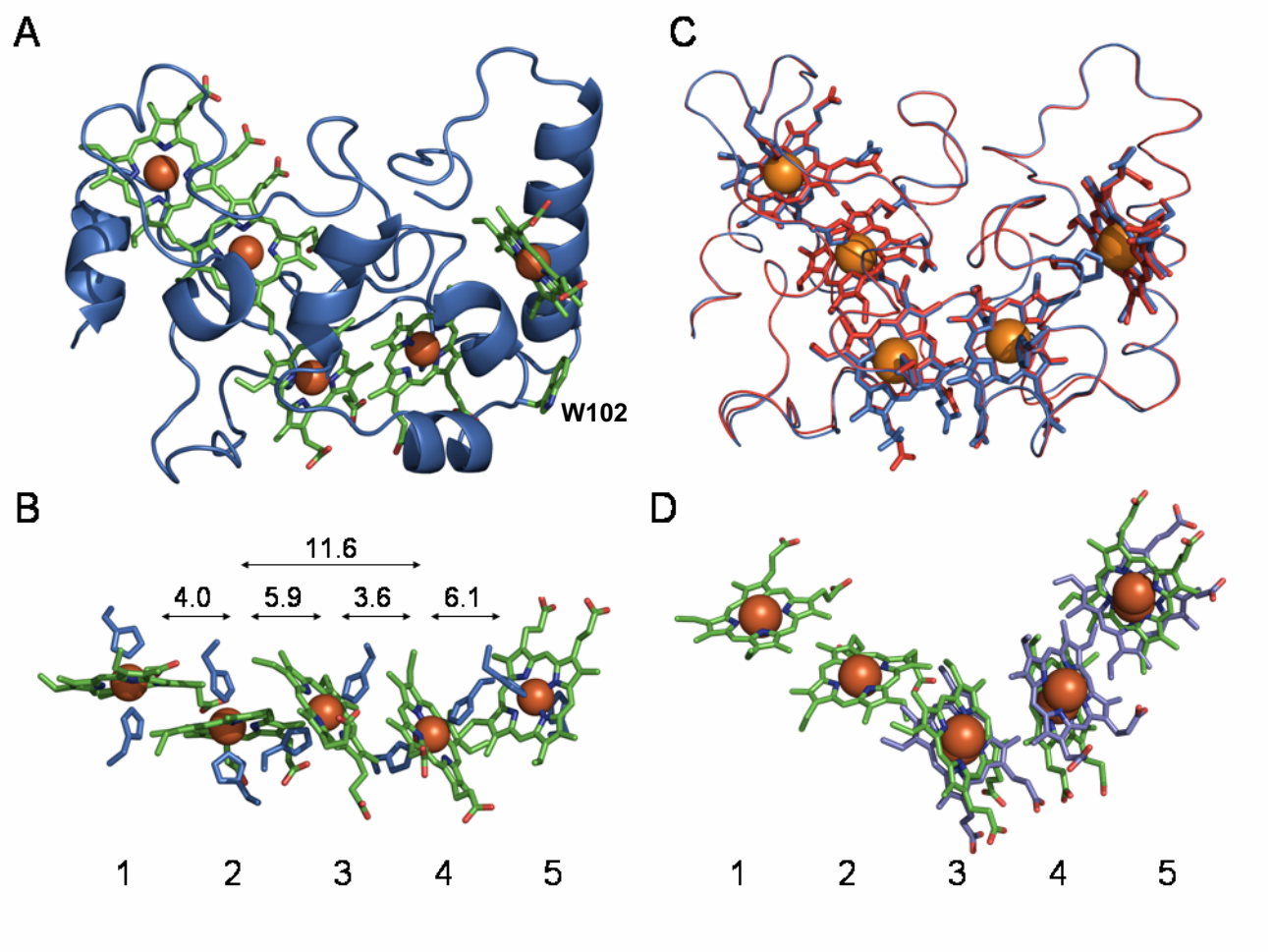
Table 3. Analysis of NrfA and NrfB interactions by analytical ultracentrifugation. Sedimentation experiments were performed on NrfA and mixtures of NrfA and NrfB over a range of concentrations using a Beckman XL-I analytical centrifuge. NrfA was centrifuged at 9000 r.p.m, while the NrfA-NrfB complex was centrifuged at 7,500 and 9,000 rpm. Scans were collected every four hours until equilibrium was achieved. The $M_{w(av)}$ was determined for each sample using Ultrascan II (Demeller, 2005). The K_d was determined for NrfA and the NrfA-NrfB complex assuming monomeric molecular weights of 53 kDa and 74 kDa respectively.

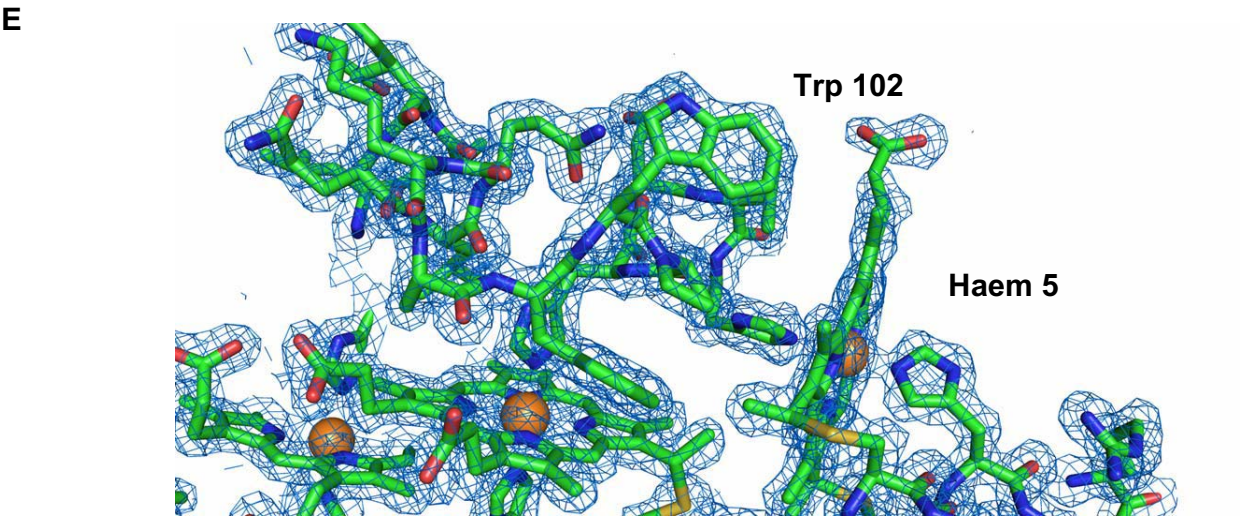
Sample (μM)	Λ (nm)	E (mM^{-1} cm^{-1})	M_w^{av} (kDa)[1]	K_d (μM)
NrfA (2.0)	410	497	72	
NrfA (10)	440	110	83	4.0
NrfA (20)	530	48	98	
NrfA-NrfB (1.0)	410	1000	97 ± 4	
NrfA-NrfB (5.0)	530	96	134 ± 1	4.1
NrfA-NrfB (20)	600	15	147 ± 9	

[1]Molecular weight shown for NrfA-NrfB samples is averaged molecular weight determined at two different speeds.

© 2007 The Authors Journal compilation © 2007 Biochemical Society

Figure 1

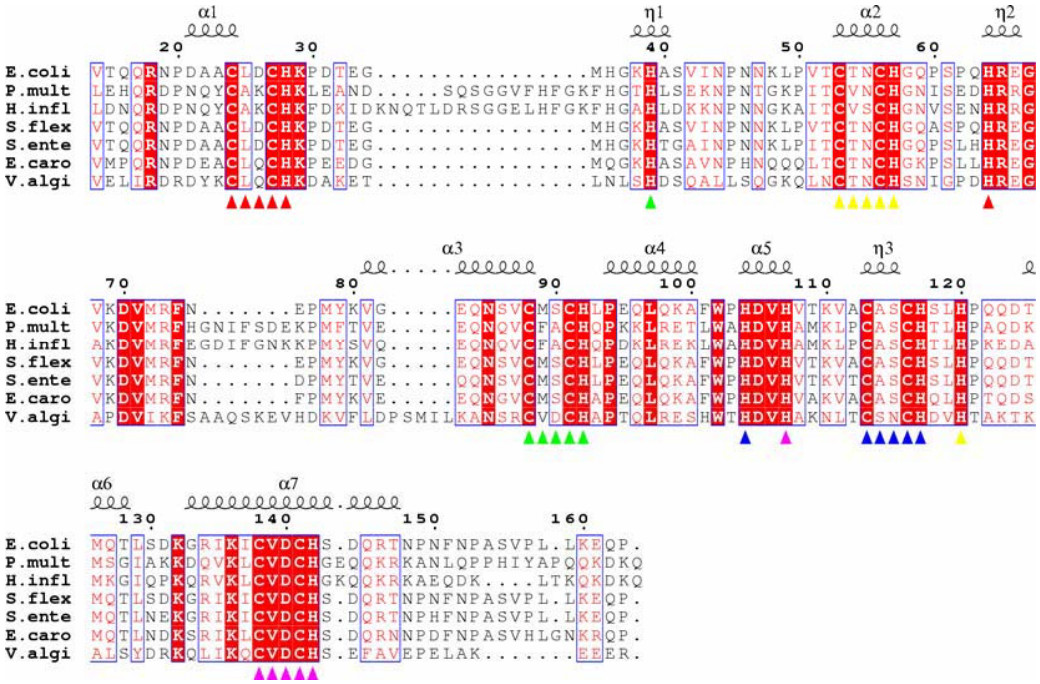




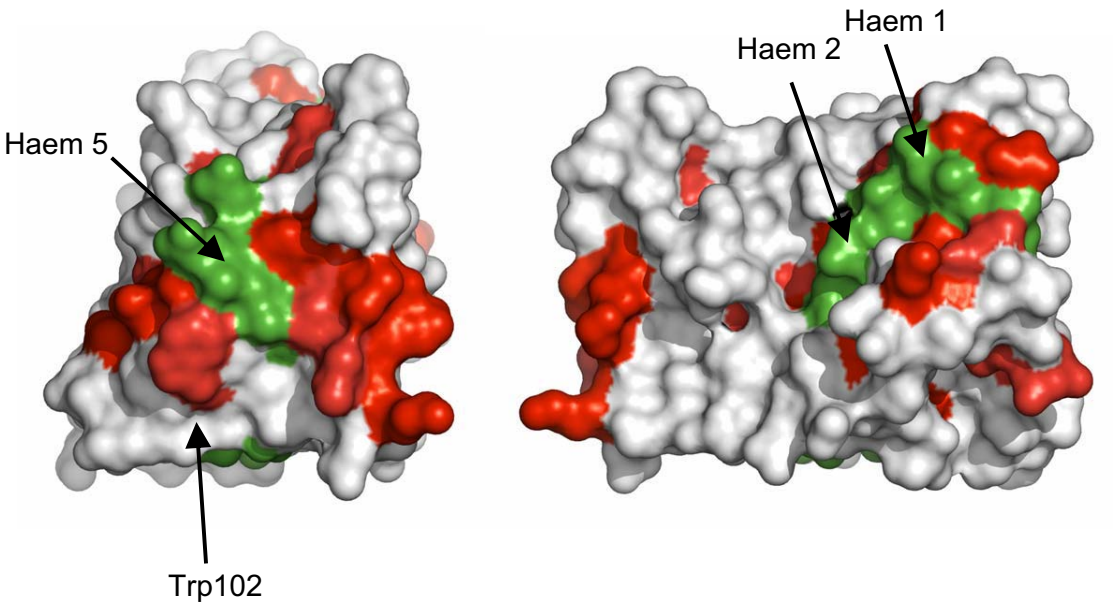
© 2007 The Authors Journal compilation © 2007 Biochemical Society

Figure 2

A

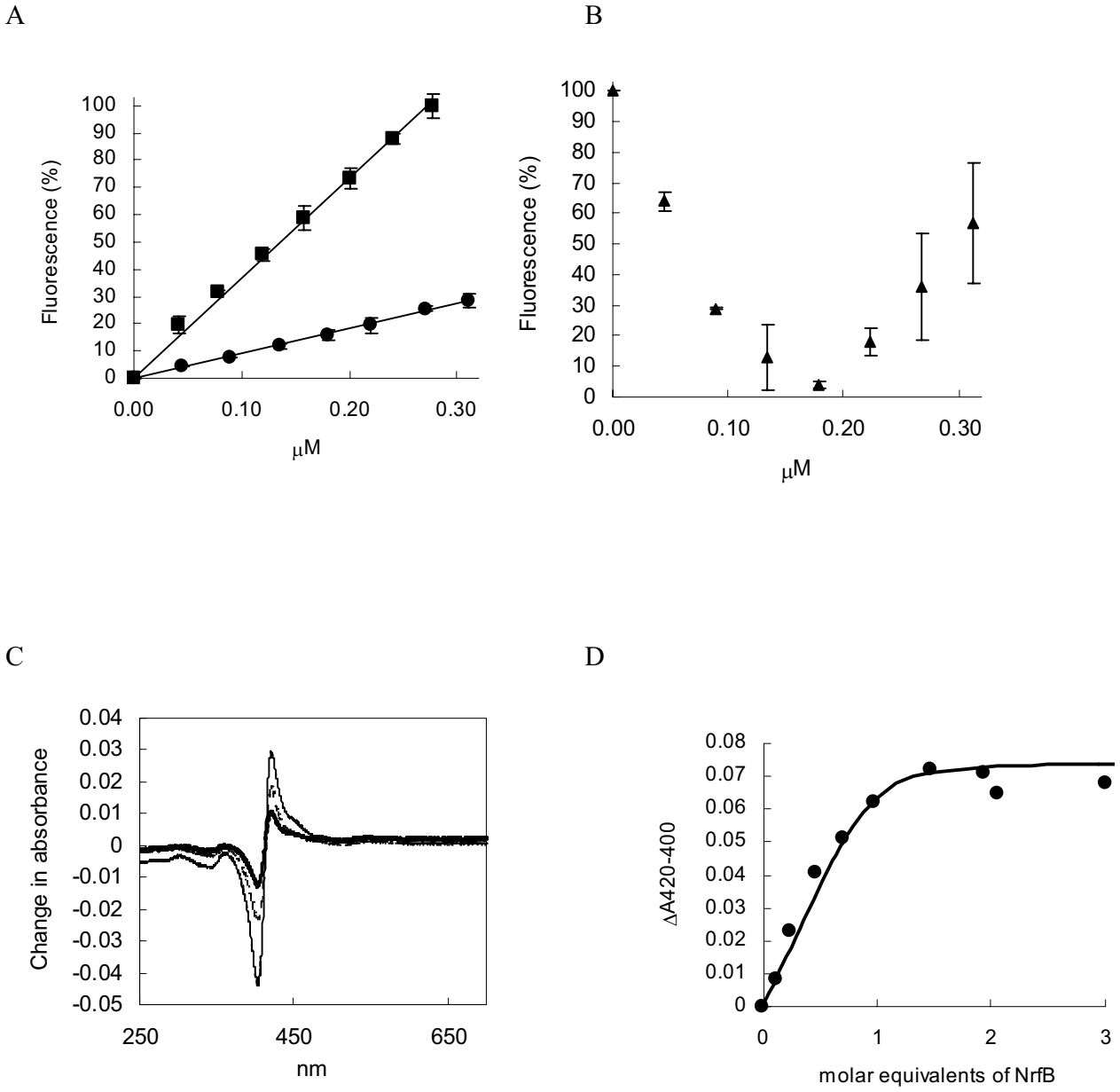


B



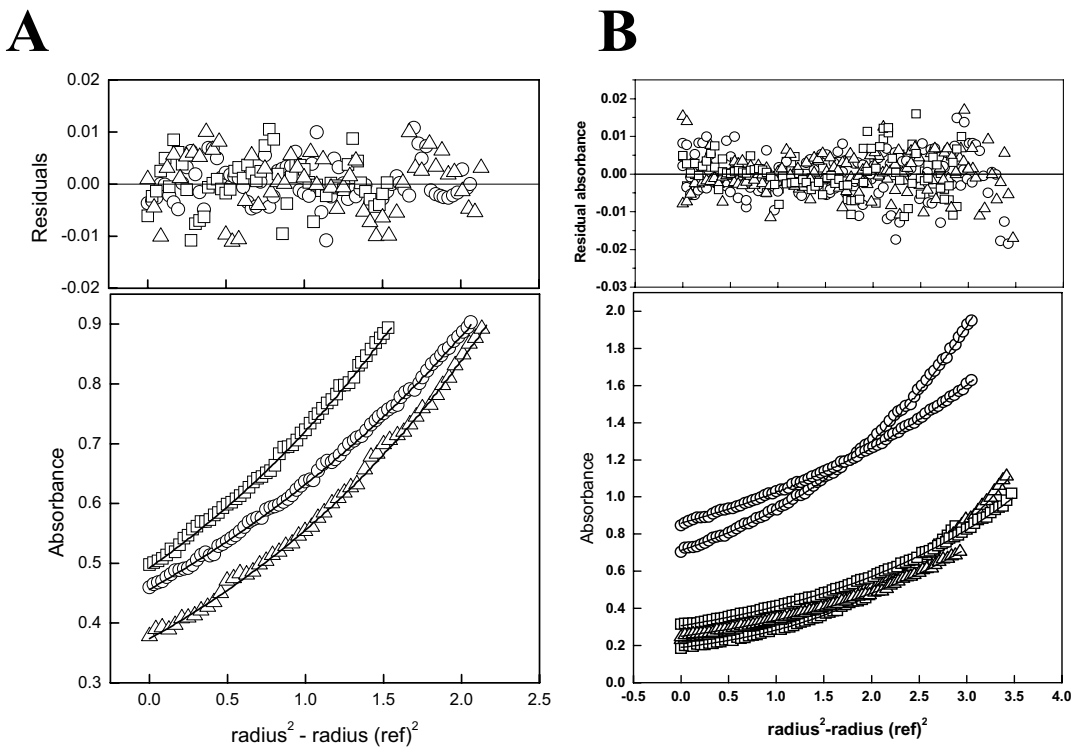
© 2007 The Authors Journal compilation © 2007 Biochemical Society

Figure 3



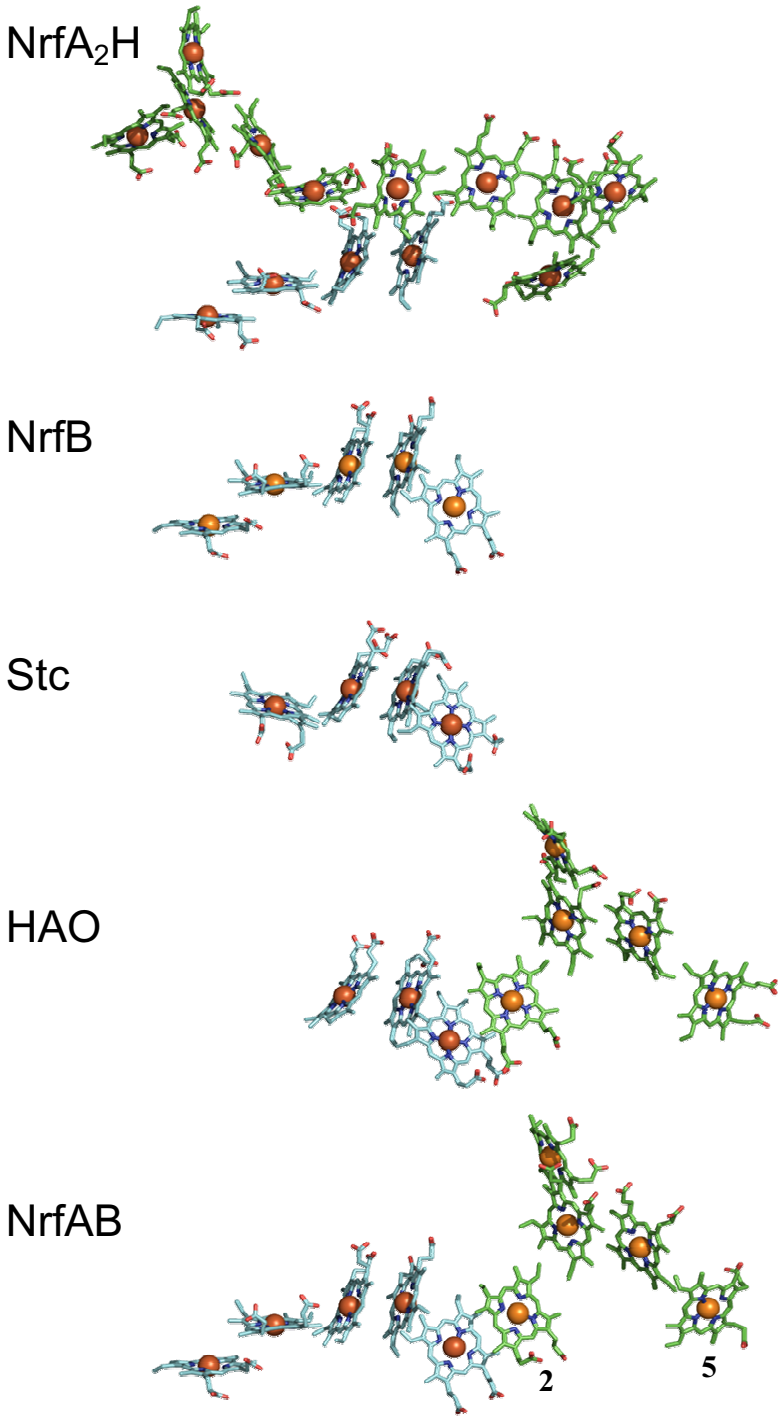
© 2007 The Authors Journal compilation © 2007 Biochemical Society

Figure 4



© 2007 The Authors Journal compilation © 2007 Biochemical Society

Figure 5



© 2007 The Authors Journal compilation © 2007 Biochemical Society