Dissecting the structural and functional roles of a vicinal iron-binding site in encapsulated ferritins

Cecilia Piergentili[1,♣], Didi He[2,#,♣], Jennifer Ross[3,♣], Will A. Stanley[1], Laurène Adam[1], C. Logan Mackay[3], Kevin J. Waldron[4], David J. Clarke[3*], Jon Marles-Wright[1*]

[1]School of Natural and Environmental Sciences, Newcastle University. Newcastle upon Tyne, NE1 7RU; [2]Institute of Quantitative Biology, Biochemistry and Biotechnology, School of Biological Sciences, The University of Edinburgh, Max Born Crescent, Edinburgh, EH9 3BF; [3]EaStCHEM School of Chemistry, The University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, Scotland, EH9 3FJ; [4]Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, NE2 4HH.

[#]*Current address: Structural Genomics Consortium, University of Oxford, Old Road Campus Research Building, Roosvelt Drive, Oxford, OX3 7DQ*

♣ These authors contributed equally to this work

* Corresponding authors

Contact details:

Email: Dave.clarke@ed.ac.uk

Tel: +44(0)131 650 4808

Email: Jon.marles-wright1@ncl.ac.uk

Tel: +44(0)191 2084855

Abstract

The Encapsulated ferritin-like proteins belong to the universally distributed ferritin superfamily, which function as iron detoxification and storage systems. The encapsulated ferritins have a distinct annular structure and must associate with an encapsulin nanocage to form a competent iron store capable of holding significantly more iron than classical ferritins. The catalytic mechanism of iron oxidation in the ferritin family is still an open question due to differences in organisation of the ferroxidase catalytic site and secondary metal binding sites vicinal to this. We have previously identified a putative metal binding site on the inner surface of *Rhodospirillum rubrum* encapsulated ferritin at the interface between the two-helix subunits and proximal to the ferroxidase centre. Here we present a comprehensive structural and functional study to investigate the functional relevance of this putative iron entry site by means of enzymatic assays, mass-spectrometry, and X-ray crystallography. We show that catalysis occurs in the ferroxidase centre and suggest a dual role for the secondary site as an electrostatic trap guiding ferrous ions toward the ferroxidase centre and, at the same time, acting as a barrier protecting the ferroxidase site against non-cognate inhibiting species. Moreover, confinement of encapsulated ferritins within the encapsulin nanocage, while enhancing the ferritin ability to undergo catalysis, does not influence the specific function of the secondary site.

Introduction

The ferritin superfamily consists of a number of structurally and functionally distinct members, each of which is built around a four-helix bundle scaffold for an iron-binding ferroxidase active site (Andrews, 2010). The quaternary structure of these protein families show a high degree of plasticity: from the monomeric rubrerythrin proteins (Dillard et al., 2011; Cardenas et al., 2016); dimeric IMEF proteins (Giessen et al. 2019); dodecameric DNA protection in Starved cells (DPS) (Grant et al., 1998; Pesek et al., 2011; Martinez et al., 1997); the 24-meric classical (Ftn) and bacterioferritins (Bfr) (Andrews, 1998, Pfaffen et al., 2013; Frolow et al., 1994; Harrison et al., 1996); and the decameric encapsulated ferritins (EncFtn) (He et al. 2016, 2019). The rubrerythrin and DPS proteins play key roles in the protection of cells from oxidative stress (Ratnayake et al., 2000; Sztukowska et al., 2002; Almirón et al., 1992; Martinez and Kolter, 1997; Choi et al., 2000; Nair et al., 2004; Chiancone and Ceci, 2010; De Martino et al., 2016); while DPS and the other members of the ferritin superfamily are vital iron stores. These latter proteins oxidise and convert free Fe(II) to inert Fe(III) oxyhydroxide or Fe(III) phosphate mineral forms, to provide a reservoir of bioavailable iron (Chasteen and Harrison, 1999). The nanocages formed by the quaternary structure of DPS, Ftn, and Bfr, allow iron oxidation and mineralisation to occur within a biochemically privileged environment within the cell, which protecting the host cell from iron-induced oxidative stress. Interestingly, the recently described IMEF and EncFtn proteins have ferroxidase activity, but they do not form nanocage structures on their own, instead they must be localised to the lumen of encapsulin nanocages to form a competent iron-store (Giessen et al., 2019; He et al., 2016, 2019).

The catalytic route to iron mineralisation in the different ferritin iron-stores is still a subject of vigorous debate (Xu and Chasteen, 1991; Proulx-Curry and Chasteen, 1995; Yang et al., 1998; Chasteen and Harrison, 1999; Baaghil et al., 2003; Ebrahimi et al., 2009) and given the lack of absolute conservation in the ferroxidase centre, it is likely that there is no single conserved mechanism for their activity ((Xu and Chasteen, 1991; Proulx-Curry and Chasteen, 1995; Yang et al., 1998; Chasteen and Harrison, 1999; Baaghil et al., 2003; Ebrahimi et al., 2009). All of the ferritin family proteins that have been structurally characterised have additional iron-binding sites to the ferroxidase centre. In the case of sites proximal to the ferroxidase centre in DPS and ferritins, these have been suggested to be iron-entry, transit, and exit points from the ferroxidase centre (Theil et al., 2013; Masuda et al., 2010; Behera et al., 2015), or the sites of electron transfer from Fe(II) substrates to the FOC (Yao et al., 2012). The subunit boundaries in the DPS and ferritin nanocages form a variety of symmetrical pores, which have variously been suggested and shown to be entry and exit sites to the lumen for iron, water and phosphate substrates (Treffry and Harrison, 1978; Andrews, 1998; Liu et al., 2007; Bellapadrona et al., 2009; Giessen et al., 2019). Whatever their function, these secondary metal sites are essential to the correct functioning of the ferritin iron-stores (Treffry et al., 1998; Ebrahimi et al., 2016; Hagen et al., 20refs).

In our previous work on the EncFtn family we identified a number of conserved secondary metal-binding sites in their structures (He et al., 2016; 2019). A partially occupied 'weak' site on the outer surface of the EncFtn decamer (Figure 1A in He et al., 2016) and a highly occupied 'entry' site on the inner face, 17; which is proximal to the ferroxidase centre (Figure 1B in He et al., 2016). Both of these sites are found at dimer interfaces between the two-helix EncFtn subunits and are formed when the four-helical ferritin fold is reconstituted by the iron-mediated assembly of the EncFtn ferroxidase centre (Ross et al., *submitted*). In our structure of the *Rhodospirillum rubrum* EncFtn protein, the entry site coordinates a calcium ion, from the crystallisation solution, in a symmetrical arrangement of glutamic acid residues (E31/E34) and a tryptophan residue (W38) contributed by each chain from the ferroxidase centre dimer (Ross et al., *submitted*) (Figure 1). The metal coordination distances between the side-chain carboxylic acid oxygens and calcium ion are 2.1 Å Ca – Glu31-Ox, 2.7 Å Glu34-Ox, which is similar to the Fe-Glu-Ox coordination distances seen in the FOC of 2.2-2.6 Å.

We hypothesise that this site acts as an entry site for iron ions approaching the ferroxidase centre, and acts as a filter to control the access of metal ions to the active site of the protein. The glutamic acid residues in this site are arranged in such a way as to capture and coordinate an Fe(II) ion in an octahedral geometry before passing it to the FOC. The positioning of W38 adjacent to this site suggests that the indole ring sterically constrains the position of the E31 side chain to an optimal position for metal coordination. We have considered several hypotheses regarding the functional role of this additional metal-binding site, including a gate prior to the channel leading to FOC; an electrostatic trap to direct metal ions to the FOC; a discriminatory filter to avoid mismetallation and consequent inhibition of the active site (Figure 2); or for protein stability and assembly. These functions may overlap, and the site may play a number of key roles in the stability and activity of the Encapsulated ferritins.

In order to understand the role of this site a number of single-point variants of the *R. rubrum* EncFtn protein were produced by site-directed mutagenesis to replace the residue of interest with non-metal binding residues. These variants were produced in the isolated *R. rubrum* EncFtn protein and an *R. rubrum* Encapsulin:EncFtn (Enc:EncFtn) co-expression system. We produced and purified recombinant E31A, E34A and W38G/A variants of EncFtn and the Enc:EncFtn complex and analysed the solution and gas phase behaviour of these proteins to understand the influence of these variants on the stability of the EncFtn complex. We obtained X-ray crystal structures of the E31A and E34A variants of the EncFtn protein and their structures display altered entry site geometry, with the abrogation of metal binding in this site in the E34A variant. The functional role of the secondary metal binding site was explored biochemically, we show that changes of metal coordination in this site enhances the ferroxidase activity of the EncFtn protein, while increasing its susceptibility to inhibition by competing zinc ions. Taken together, these results indicate that this metal binding site plays a minor role in stabilising the oligomeric state of the EncFtn decamer, and primarily acts as an electrostatic filter to control metal access to the ferroxidase centre.

Methods

Cloning

Expression constructs for the *R. rubrum* EncFtn protein (pACYCDuet-1-Rru_EncFtn) and Enc:EncFtn protein complex (pACYCDuet-1-Enc:EncFtn) were produced in the previous work (He et al., 2016). Site-directed mutagenesis was performed using Stratagene QuikChange kit following manufacturer instructions. Protein nomenclature used throughout the text is shown in Supplementary Table S1. Primers used in this work for this purpose are listed in Supplementary Table S2.

Protein production and purification

Expression plasmids for the *R. rubrum* EncFtn protein and Enc:EncFtn complex were transformed into *Escherichia coli* BL21(DE3) cells and grown in LB medium and expressed as described previously (He et al., 2016; 2019).

EncFtn and its variants, were purified as described in He et al., 2019; concisely, the proteins were purified by combining anion exchange chromatography (AXC) and Size exclusion chromatography (SEC). Purity of protein samples was analysed by SDA-PAGE.

Enc:EncFtn wild-type and variant complex cell pellets were resuspended in 10-times (v/w) of Lysis Buffer (20 mM HEPES pH 8.0, 2 mM $MgCl_2$, 1 mg/ml lysozyme, 20 units/ml benzonase). Sonication on ice (5 minutes at 30 seconds on/off cycles, 35 %) of cell sample was followed by incubation for 30 minutes at 37 °C using a water bath. Cell lysate was clarified by centrifugation (35,000 x g, 30 minutes, 4 °C) and the supernatant was collected and incubated at 80 °C for 10 minutes. After cooling the sample was centrifuged again (30 minutes, 35,000 x g, 4 °C), and the supernatant was collected and filtered using a 0.22 um syringe filter (Millipore, UK). Enc:EncFtn protein complex sample was loaded onto a 20 ml anion exchange column (Hi-Trap Q-sepharose FF, GE Healthcare) pre-

150 equilibrated with Buffer QA (50 mM Tris, pH 8.0). Un-bound protein was washed with Buffer QA.
151 A step gradient elution was performed by mixing Buffers QA and QB (50 mM Tris, pH 8.0, 1 M
152 NaCl). Since Enc:EncFtn does not bind the Q-sepharose column, proteins of interest were found in
153 flow-through sample, as confirmed by 15 % (w/v) SDS-PAGE. Protein was concentrated using a
154 Vivaspin centrifugal concentrator (MWCO 10 kDa) following the manufacturer's instructions.
155 Concentrated protein was then applied to a Size exclusion column (HiLoad 16/600 Superdex 200, GE
156 Healthcare) pre-equilibrated with Buffer GF (50 mM Tris, pH 8.0, 150 mM NaCl). Fractions of
157 interest were run on a 15% (w/v) SDS-PAGE to assess protein purity and oligomerisation state, and
158 further concentrated using a Vivaspin centrifugal concentrator as before. After concentration, protein
159 samples were analysed by 10% Tricine-SDS-PAGE (Schägger, 2006) in order to resolve bands
160 around 30 kDa. Elution volumes of EnFtn and Enc:EncFtn complexes from the size exclusion column
161 are listed in Supplementary Table S3.
162
163 **His-tagged protein production and purification**
164 Hexahistidine tagged EncFtn E31A and E34A variants (EncFtn_sH) were produced in *E. coli*
165 BL21(DE3) as described above and purified following the same protocols as previously described
166 (He et al., 2016). Briefly, a 5 ml HisTrap column (GE Healthcare, UK) was used for immobilised
167 metal affinity chromatography (IMAC), followed by size exclusion chromatography (HiLoad 16/600
168 Superdex 200, GE Healthcare) using the protocol described above.
169
170 **Protein quantification**
171 Both untagged and His-tagged EncFtn proteins were quantified using the Pierce BCA Assay kit
172 (Thermo Fisher Scientific) following the manufacturer's instructions for the Test-tube procedure.
173 Concentration of Enc:EncFtn proteins was determined using the Beer-Lambert equation, measuring
174 absorbance at 280 nm with a Nanodrop Spectrophotometer (Thermo Scientific) and using an
175 extinction coefficient calculated by ProtParam tool on the Expasy platform entering a sequence
176 composed of 3 x Enc and 2 x Enc-Ftn sequences. This ratio was determined by examination of SDS-
177 PAGE gels of Enc:EncFtn proteins and densitometry of bands.
178
179 **Transmission electron microscopy**
180 TEM imaging was performed as previously described (He et al., 2016) on purified Enc:EncFtn
181 proteins. Proteins were diluted in Buffer GF to a final concentration of 0.1 mg/ml before being spotted
182 on glow-discharged 300 mesh carbon-coated copper grids. Excess liquid was removed by blotting
183 with filter paper (Whatman, UK) before sample was stained with 0.2 % uranyl acetate, blotted again,
184 and air-dried. Grids were imaged using a Hitachi HT7800 transmission electron microscope (LaB6
185 filament) and images were collected with an EMSIS Xarosa camera (Radius software). A
186 representative set of images are presented here.
187
188 **Mass spectrometry analysis**
189 Sample preparation, instrumentation, native MS conditions, ion mobility-MS conditions and the
190 method of collision cross section (CCS) determination have all been reported in previous works by
191 our group (He et al., 2016; 2019). For collision induced unfolding (CIU) experiments, protein samples
192 were prepared and the Synapt G2 was set up as for IM-MS. Samples were analysed with incremental
193 increasing trap voltages between 10 V to 60 V, using 2 V increments. Data was processed using
194 Spyder CIUSuite (Dixit et al., 2018).
195
196 **Protein crystallization and X-ray data collection**
197 EncFtn_sH variants (E31A or E34A) were concentrated to 10 mg/ml (based on extinction coefficient
198 calculation) and subjected to crystallisation under similar conditions to the wild-type protein
199 (WT)（Supplementary Table S4）. Crystallization drops were set up in 24−well Linbro plates using
200 the hanging drop vapor diffusion method at 292 K. Glass coverslips were set up with 1–2 µl protein
201 mixed with 1 µl well solution and sealed over 1 ml of well solution. Crystals appeared after one week

202 to two months and were mounted using a LithoLoop (Molecular Dimensions Limited, UK),
203 transferred briefly to a cryoprotection solution containing well solution supplemented with 1 mM
204 $FeSO_4$ (in 0.1% (v/v) HCl) and 20% (v/v) PEG 200, which were subsequently flash cooled in liquid
205 nitrogen.
206
207 All crystallographic datasets were collected at 10–100 eV above the experimentally determined Fe-
208 *K*α edge on the macromolecular crystallography beamlines at Diamond Light Source (Didcot, UK)
209 at 100 K using Pilatus 6M detectors. Diffraction data were integrated and scaled using XDS (Kabsch,
210 2010) and symmetry related reflections were merged with Aimless (Evans, 2011). Data collection
211 statistics are shown in Table 3. The resolution cut-off used for structure determination and refinement
212 was determined based on the $CC_{1/2}$ criterion proposed by Karplus and Diederichs, 2012.
213 The structures of the EncFtn_sH variants were determined by molecular replacement using decameric
214 wild-type protein (PDB ID: 5DA5) as the search model (He et al., 2016). A single solution comprising
215 three decamers in the asymmetric unit was found by molecular replacement using Phaser (McCoy et
216 al., 2007). The initial model was rebuilt using Phenix.autobuild (Adams et al., 2010) followed by
217 cycles of refinement with Phenix.refine (Afonine et al., 2012), with manual rebuilding and model
218 inspection in Coot (Emsley et al., 2010). The final model was refined with isotropic B-factors,
219 torsional NCS restraints, and with anomalous group refinement to correctly model the iron ions in
220 the ferroxidase centre. The model was validated using MolProbity (Chen et al., 2010). Structural
221 superimpositions were calculated using Coot (Emsley et al., 2010). Crystallographic figures were
222 generated with PyMOL. Model refinement statistics are shown in Table 3. The final models and
223 experimental data are deposited in the PDB.
224
225 **Ferroxidase assay of EncFtn proteins**
226 Enzymatic activity of EncFtn proteins was tested by ferroxidase assay, as previously described (He
227 et al., 2019). Fe(II) aliquots were prepared under anaerobic conditions by dissolving FeSO4·7H2O in
228 0.1% (v/v) HCl. Purified protein was diluted anaerobically in Buffer H (10 mM HEPES, pH 8.0, 150
229 mM NaCl), previously purged with gaseous nitrogen, to a final concentration of 10 µM (EncFtn).
230 Protein and ferrous ion samples were added to a quartz cuvette (Hellma) under aerobic conditions at
231 a final concentration of 10 µM and 50 µM, respectively, corresponding to a FOC:Fe ratio of 1:10.
232 Absorbance at 315 nm was recorded every second for 1500 s by a UV-visible spectrophotometer
233 (PerkinElmer Lambda 35), using the TimeDrive software. The same experiment was performed in
234 the absence of the enzyme to determine the oxidation of ferrous salts by atmospheric oxygen.
235 The same setup was adopted for recording the activity of EncFtn_E34A and EncFtn_E31A/E34A in
236 a separate assay at higher concentration (20 µM EncFtn and 100 µM Fe(II)) for comparison purposes.
237 Data presented here are the mean of three technical replicates of time zero-subtracted progress curves
238 with standard deviations calculated from the mean.
239 Calculation of the EncFtn enzymatic reaction initial rate (v_0) was made by applying the Linear
240 Regression tool in GraphPad (Prism8) on the absorbance at 315 nm measured for the first 10 s or 200
241 s, when curves are still linear and following an order zero kinetics (Figure 10, Supplementary Tables
242 S5, S6, S8). Slopes obtained from these curves correspond to initial reaction rate. $v_0^{variant}/ v_0^{wild\text{-}type}$
243 factors were calculated by dividing the variants initial rates by the initial rate of the wild-type protein
244 (Supplementary Tables S7).
245
246 **Ferroxidase assay of Enc:EncFtn protein complexes**
247 Ferroxidase assay of Enc:EncFtn protein complexes was performed in the same conditions as those
248 used for EncFtn, but with different protein and iron salt concentration. Enc:EncFtn was diluted to a
249 final concentration of 25 µM which corresponds to 10 µM EncFtn based on a 3:2 Enc:EncFtn ratio,
250 as previously observed by SDS-PAGE by our group (Figure 1A in He et al., 2016). Fe(II) sample was
251 diluted to 50 µM in the reaction system to maintain a final ratio of 1:10 FOC:Fe(II).
252 Calculation of the Enc-EncFtn enzymatic reaction initial rate (v_0) and $v_0^{variant}/ v_0^{wild\text{-}type}$ factors was
253 carried out as described in the above section (Figure 10, Supplementary Tables S5, S6, S7, S8).

254 **Zinc inhibition of ferroxidase activity of EncFtn proteins**
255 Ferroxidase assay was performed as above with EncFtn (10 μM) and Fe(II) samples (50 μM) in the
256 presence of 34 μM $ZnSO_4 \cdot 7H_2O$. The chosen concentration corresponds to the IC50 value previously
257 determined for EncFtn_StrepII, under the same experimental conditions, by our group (He et al.,
258 2019). Oxygen-free metal samples were added to the quartz cuvette under aerobic conditions,
259 followed by the protein sample. Data were replicated three times and mean and standard deviation of
260 time zero-subtracted progress curves were calculated. Negative control was performed by monitoring
261 A_{315} of Zn(II) and Fe(II) salts mixed in the absence of enzyme.
262
263 **Accession codes and datasets**
264 Data sets supporting this paper have been deposited in appropriate public data repositories. Please
265 see figure legends and tables for links to these.
266
267 **Results**
268 *Purification of encapsulated ferritins from R. rubrum*
269 In order to explore the functional roles of the residues in the putative entry site of EncFtn, a number
270 of single-point variants were produced by site-directed mutagenesis, selectively substituting each
271 metal-binding residue in the site with alanine or glycine. The resulting set of variants were also
272 generated in association with the encapsulin component to produce Enc:EncFtn complexes. All the
273 proteins were produced by heterologous expression in *Escherichia coli.*
274
275 To investigate the influence of these changes to the proteins in solution, all proteins were subjected
276 to size-exclusion chromatography using an S200 column calibrated with standards of known
277 molecular mass (Figure 3A/B). The resulting fractions were analysed by SDS-PAGE (Figure S1A)
278 and the average molecular masses of the EncFtn proteins were confirmed by LC-MS (Table 1).
279 Each of the variant EncFtn proteins eluted with a similar profile, with a major peak at retention
280 volume around 60 ml, a second peak at retention volume of 75 ml, and the final peak at 82 ml (Figure
281 3A, Supplementary Table S3). The E34A variant had significantly more protein in the 82 ml peak
282 when compared to the wild type and E31A variant, suggesting that this change affects the stability of
283 the protein complex.
284
285 Based on our prior work on the assembly pathway of the *Haliangium ochraceum* EncFtn protein
286 (Ross et al., *submitted*) and native mass spectrometry of other representative members of the EncFtn
287 family (He et al., 2019) these peaks would correspond to decamer, tetramer, and dimer species. The
288 anomalous size estimation based on calibration standards can be attributed to the fact that the EncFtn
289 proteins used in this study were full length proteins with the flexible localisation sequence, which
290 would increase the hydrodynamic radius of the oligomers considerably.
291
292 The Enc:EncFtn protein complexes all eluted as single peaks starting in the void volume of the S200
293 column, indicating the formation of large protein complexes not able to enter the matrix pores (Figure
294 3B, Supplementary Table S3). SDS-PAGE analysis of peak fractions confirmed the presence of both
295 the EncFtn and encapsulin proteins (Figure S1) To determine if these complexes had formed correctly
296 assembled encapsulin nanocages we performed transmission electron microscopy on uranyl acetate
297 stained samples (Figure 4). The encapsulins purified with wild type and variant EncFtn all showed
298 characteristic 25 nm encapsulin cages with visible internal density for the EncFtn proteins, consistent
299 with what previously observed for Rru_Enc-EncFtn wild-type (He et al., 2016).
300
301 *Native mass spectrometry of EncFtn proteins*
302 To interrogate the stability and oligomeric states of the EncFtn variants we performed native MS and
303 ion mobility MS experiments on the samples purified by size-exclusion chromatography (Figure 5A)
304 (He et al., 2019). In native MS analysis, EncFtn_E31A behaves similarly to EncFtn and adopts a
305 decameric charge state distribution (Figure 5A, pink circles). In contrast, under the same instrumental

conditions, EncFtn_E34A and EncFtn_W38G display significant differences – with both displaying decameric and monomeric charge state distributions (Figure 5, stressed with pink circles and blue triangles, respectively). The presence of monomeric with relatively high charge states (8+ to 16+), suggests a degree of gas phase dissociation from the decameric complex during native MS analysis. Even with careful control of instrument conditions, this monomeric dissociation product could not be completely eliminated, suggesting that the EncFtn_E34A and EncFtn_W38G variants are less stable than the wild type protein. The difference in gas phase stability of the EncFtn variants was further highlighted when increasing the sampling cone voltage and/or activating the trap voltage of the mass spectrometer, which activates the complex and causes dissociation (Figure S2). EncFtn_W38G fully dissociates into its monomer charge state distribution at lower activation conditions than the other variants, further highlighting its instability (Figure S2D).

To further investigate the gas phase stability of the variants, ion mobility collision induced unfolding (CIU) experiments were performed (Eschweiler et al. 2015). The complexes were unfolded by gradually increasing the trap voltage prior to the gas phase conformation of the native protein ions is monitored by ion mobility mass spectrometry. The resulting ion mobility profiles for an individual protein ion is then plotted as a function of activation voltage to produce a heat map to reveal discrete conformations and gas phase unfolding transitions (Figure 6). It is clear from these experiments, that two discrete conformations of each EncFtn variant are observed during CIU; a major conformation with a compact structure and a minor conformation with a more extended, unfolded structure (Figure 6, Figure S3). At minimal activation, EncFtn_WT, EncFtn_E34A and EncFtn_E31A all exist in a single, compact decameric conformation with a drift time of between 10 and 11 ms, suggesting the same overall structure. As activation is increased all three variants undergo a discrete transition to a more extended conformation (characterised by a drift time of 12.5 ms) at around 30 V activation, followed by a more complex transition to higher drift time conformations with activation above ca. 40V. The close similarity of these CIU profiles strongly suggest have similar gas phase structures and stability (Figure 6, Figure S3). In contrast, EncFtn_W38G occurs in a more extended conformation (12.5 ms) throughout the CIU experiment, with the compact conformation only observed as a minor species at low activations. This is in accord with the native MS studies and suggests that the W38G substitution has a significant structural effect and destabilises the decameric EncFtn assembly.

By calibration of the ion mobility data, the collision cross sections (CCS) of both the discrete conformations of the EncFtn decamer were determined (Table 2). The compact conformation of EncFtn_WT has a CCS of 71.40 nm^2; whereas the more extended conformation, observed at higher activation has a CCS of 80.62 nm^2 (Table 2). We note that the CCs of the compact conformer is significantly larger than the CCS calculated from the EncFtn_ sH is crystal structure of 58.7 nm^2 and 58.2 nm^2 observed in previously reported IM-MS experiments (He et al., 2016). However, prior studies were performed on truncated EncFtn_WT which has the C-terminal localisation sequence region removed. The higher observed CCS for EncFtn_WT suggests that the localisation sequence is on the surface of the decamer as it has an increased cross-sectional area. This is in accord with our observations for the size-exclusion experiments, where the EncFtn variants elute with an apparently higher hydrodynamic radius than the expected oligomeric states.

X-ray crystal structures of R. rubrum EncFtn variants

Given the observation E31A and E34A variants do not display any significant differences in terms of oligomeric state and stability both in solution and the gas phase, we subjected truncated his-tagged versions of these proteins to crystallisation screening to determine any changes at the metal binding sites. We obtained crystals in similar conditions to the wild type sequence protein. We were unable to obtain crystals of the W38G, or W38A variants of the EncFtn protein, indicating that the conformational differences seen in the mass spectrometry experiments destabilise the quaternary structure of the protein sufficiently to present a barrier to stable crystallisation.

357 The structures of the *R. rubrum* EncFtn_sH E31A and E34A variants were determined by X-ray
358 crystallography and confirmed to adopt the same overall quaternary fold as the wild type sequence
359 EncFtn, with RMSD Cα differences of 0.2 A^2 over 91 Cα and 0.17 $Å^2$ over 90 Cα for the E31A and
360 E34A variants respectively, and an RMSD Ca of 0.23 $Å^2$ over 90 Cα between the two variants. This
361 indicates that these amino acid substitutions do not significantly alter the structure of the protein in
362 the E31A and E34A variants.
363
364 Inspection of anomalous difference maps calculated for the X-ray data collected near to the Fe-*K*α
365 edge indicate the presence of anomalously scattering metal ions in the ferroxidase centre for both the
366 E31A and E34A variants. We have assigned these as iron ions based on the size of the peaks and the
367 nature of this site (Figure 7A/B, 8A/B). The ferroxidase centre of the E31A variant has an additional
368 peak of electron density assigned as a water molecule, bridging the di-iron centre.
369
370 The entry site of the E31A variant appears to coordinate a metal ion via the Glu34 side-chain,
371 inspection of anomalous maps allowed us to assign this as a calcium ion; however, the coordination
372 distances are much longer than the wild type, at 3.8 Å compared to 2.8 Å (Figure 7C.). Loss of the
373 coordinating Glu31 side chain ligand does not appear to cause gross structural rearrangements in this
374 site, and the change in metal ion binding can be ascribed to the change in the coordination
375 environment around this site.
376
377 The entry site of the E34A variant does not appear to coordinate any metal ions and the side chain of
378 Glu31 is flipped away from this site towards Arg42 (Figure 8C), the Trp38 indole ring is also moved
379 away from the site when compared to the wild type protein. These changes greatly disrupt the metal
380 coordination in this site.
381
382 *Ferroxidase activity of EncFtn and Enc:EncFtn proteins*
383 While the MS and crystallographic data were in agreement with respect to the quaternary structure
384 of the E31A and E34A variants, our inability to produce crystals of the W38G protein and reduced
385 stability in our MS experiments highlight the importance of this residue for the oligomerisation of the
386 EncFtn decamer. The changes in metal ion coordination seen in the E31A and E34A crystal structures
387 suggest these may play a role in controlling entry of iron ions to the ferroxidase centre. To test this
388 hypothesis, we investigated the activity of all of the variants produced, both in isolation, and in
389 complex with the encapsulin protein. The rationale behind producing single point mutants was to
390 discern between the different actions that the E31-E34-W38 residues could exert on the enzyme
391 activity. At the start of our work we delineated a number of hypothesis involving this site, including
392 it being an access gate at the entry of the channel, and providing attractive forces which, as a
393 consequence, guide the metal ions to the FOC, buried at the dimer interface.
394
395 The ferroxidase activity of the wild-type EncFtn protein was confirmed to be comparable to that
396 previously determined for the EncFtn_StrepII variant (He et al., 2019). Both the W38A and E31A
397 mutations clearly affect the activity of the protein, as oxidation of Fe(II) to Fe(III) occurs faster and
398 both curves present a distinct shape to the wild-type protein (Figure 9). The EncFtn_E34A variant
399 shows a smaller increase in catalytic activity when compared to the other variants (Figure 9).
400
401 Considering these differences in activity in the context of the X-ray crystal structures, it is clear that
402 disruption of the metal coordination in this entry site has a significant impact on the catalytic activity
403 of the EncFtn protein. The EncFtn_E31A structure shows that the gating provided by this mutant is
404 partially diminished, as is its binding ability. However, the residual glutamic acid, Glu34, is still able
405 to coordinate metal ions, albeit less tightly than the wild type given the increased distance between
406 these residues and lack of additional coordination from Glu31 (Figure 7). This variant possesses the
407 highest ferroxidase activity, along with EncFtn_W38A.
408

The crystal structure of the EncFtn_E34A variant shows that the mutation of Glu34 into Ala results in a reorganisation of the site, with the side chain of Glu3 flipped by 117°, and is no longer correctly positioned for metal coordination, essentially removing both the electrostatic attraction and gating of this site (Figure 8). A ferroxidase assay carried out with an EncFtn_E31A/E34Avariant shows that its activity is comparable to EncFtn_E34A (Figure 11). This data corroborates our hypothesis that in EncFtn_E34A, the gating function is completely abolished, and no attraction is conveyed by the residual Glu31. Hence, in Rru_EncFtn_E34A metal ions slowly reach the FOC by diffusion only. As no gating is provided at this level, we would expect to observe a higher activity, however the absence of attracting forces results in modest increase when compared to wild type. The difference in activity between EncFtn_E31A and EncFtn_E34A shows how critical Glu34 is and more broadly the importance of this electrostatic attraction for the activity of the protein.

As we were unable to produce crystals of either the W38G or W38A variants, we could not determine the effect of removing this tryptophan residue on the structure of the entry site. However, analysis of the crystal structure of the wild type EncFtn shows that the tryptophan indole ring appears to sterically constrain the Glu31 side chain in the correct position for metal-coordination in WT (Figure 1, He et al., 2016), we propose that the removal of Trp38 results in the Glu31 side chain moving away from the site. Our ferroxidase assay data presents this variant as having a comparable activity profile to the E31A variant, we therefore propose that the two variants possess similar entry site structures. The broader destabilising effects of the loss of the tryptophan side chain are harder to model, but the loss of a side chain of the size of the tryptophan indole may lead to further structural rearrangements around the internal dimer interfaces engaged by this residue.

The same ferroxidase assay was performed with the Enc:EncFtn complex, with wild-type and variants carrying the same mutations introduced in Rru_EncFtn protein (Figure 9). These deviations from wild-type behaviour were compared to what observed in the absence of the encapsulin nanocompartment (Figure 9A). Enzymatic reaction initial rates (v_0) were calculated for each of the variants (in the first 200 s of the assay) and then divided by the initial rate of the corresponding wild-type protein (EncFtn or Enc:EncFtn) (Supplementary Table S7). This allowed comparison across the two experimental setups (EncFtn versus Enc:EncFtn proteins). Data show that E31A and W38X (where X is Gly or Ala) mutations have a greater effect on activity within both systems (about 4 folds faster when compared to wild-type initial rate) while removal of Glu34 results in a 1.5-2-folds increase in the initial velocity v_0 (Figure 10, Supplementary Table S7). It appears that the EncFtn variants and wild-type retain the same relative differences in their enzymatic activities, whether they are associated to the encapsulin shell or not. This outcome suggests that the gating and the attraction properties are actually exerted by the entry site instead of the pore level.

However, if we examine Enc:EncFtn_E31A curve in more details, we notice that it is not completely linear in the first 200 s of reaction (Figure 10B). Hence, we chose to directly compare the initial velocity (v_0) of EncFtn_E31A and Enc:EncFtn_E31A within this shorter period of time (Figure S5). Although EncFtn_E31A data are far too noisy to allow an acceptable fit of the curve resulting from the Linear Regression tool (Supplementary Table S8), a qualitative comparison can still be made (Figure S5). When encapsulated, EncFtn_E31A appears to catalyse the ferroxidase reaction much faster than other variants, including EncFtn_E31A when not compartmentalised within A0974 (Supplementary Table S8). After the first 20-30 s the reaction starts to slow down, reaching a rate comparable to Enc:EncFtn_W38G. It is plausible to believe that the nanocompartment improves the stabilisation of the inner protein, resulting in a different reaction profile.

Zinc inhibition of ferroxidase activity of Rru_EncFtn

In order to test the hypothesis that the entry site acts as a selectivity filter, and is able to discriminate between metal ions approaching the FOC, ferroxidase assays were performed with the addition of zinc as a competitive inhibitor of ferroxidase activity. The activity of EncFtn wild-type and variant

proteins (10 μM, monomer) with 50 μM Fe(II) was assayed in the presence of 34 μM Zn(II) (Figure 12A). The selected zinc concentration was previously determined to be the IC50 for wild type EncFtn_StrepII (He et al., 2019). This concentration was chosen to allow identification of the impact of inhibition across the set of proteins. In line with our previous results the EncFtn wild type sequence protein showed a 54% inhibition (He et al., 2019) (Figures 12A/B). The activity of the three variant proteins (E31, E34, W38) showed a markedly increased relative inhibition when compared to their enhanced activity in the absence of zinc (66%, 68%, and 88%, respectively) (Figures 12A/B).

The increased inhibition seen in the variants can be explained by the changes in metal coordination at this site. With all the variants, the change in metal coordination at the putative entry site permits greater access to the FOC for competing metal ions, as they are no longer effectively captured in the entry site. Of note is the fact that in the presence of zinc, the activity of each variant is the same as the wild type protein.

We have shown that substitution of residues in the entry site triad (E31, E34, W38) with non-metal binding residues results in the enhancement of enzymatic activity (Figure 9A). Our crystal structures support the conclusion that this is a consequence of the creation of a wider and more loosely coordinating site, which allows the faster passage of metal ions to the FOC due to a reduced energy barrier at this site (Figure 7, Figure 8). Furthermore, when these key residues are removed the protein is more susceptible to inhibition by Zn(II). As wild type EncFtn presents an extra metal binding site with which competing metals would interact, this presents a barrier to the FOC, which would trap metal ligands before they are able to reach the FOC itself. In the absence of this site both iron and competing metals can approach the FOC, either productively in the case of iron, or unproductively in the case of zinc.

Discussion

Our work presents a comprehensive structure/function study investigating the role of the putative metal ion entry site proximal to the ferroxidase centre FOC in encapsulated ferritin-like proteins. To investigate the role of this site we produced single point mutations in residues in this site to replace metal coordinating residues in the *R. rubrum* EncFtn protein with alanine (E31A, E34A) and to remove the large tryptophan side chain proximal to this site (W38G/A).

Solution and gas-phase analyses of variant proteins showed that removal of the metal coordinating ligands did not have a significant effect on protein oligomerisation, however the removal of the Trp38 side chain destabilised the protein considerably. X-ray crystal structures of the E31A and E34A variants highlighted changes in metal coordination at this site.

The effect of these changes on the activity of the protein was tested using ferroxidase assays and zinc inhibition experiments. Removal of the Glu31 and Trp38 side chains significantly enhanced the activity of the enzyme. While removal of metal coordinating residues increased the susceptibility of the protein to inhibition with zinc.

Taken together these results support the hypothesis that this metal ion binding site has a dual function in modulating the rate of ferroxidase activity of the EncFtn protein and protecting the ferroxidase centre from competing metal ions by acting as an electrostatic trap prior to the active site. These results also confirm that the ferroxidase centre is indeed the active site of this enzyme, as the loss of the metal ion entry site enhances enzyme activity rather than abrogates it. Why the EncFtn protein has evolved a mechanism to effectively slow its activity is still an open question. There may be a trade-off between enzyme activity and the availability of a sink for the electrons liberated from the iron in the ferroxidase centre. The catalytic mechanism and electron sink for ferritins is still a subject of vigorous debate and the EncFtn family presents a unique structural arrangement of the active site, which will no doubt present a distinct mechanism when probed further. The mineralisation step is

513 also likely to be a rate limiting step and the entry site may provide a brake to ensure productive
514 coupling of iron oxidation and mineralisation.
515
516 Encapsulation of the EncFtn protein within the encapsulin nanocage enhances the activity of all of
517 the EncFtn variants but does not change the way in which the variants affect the activity, indicating
518 that the function of the entry site is not changed through association with the encapsulin shell. These
519 results suggest that the encapsulin shell acts to guide productive iron ion interactions with the EncFtn
520 protein, indeed, our previous work showed that the encapsulin shell interacts with a significant
521 amount of iron in the absence of the EncFtn protein. This warrants future investigation into the ability
522 of the encapsulin shell to conduct metal ions to the EncFtn active site and the role of the encapsulin
523 shell pores in the access of substrates to the encapsulated enzyme.
524
525
526
527
528
529
530

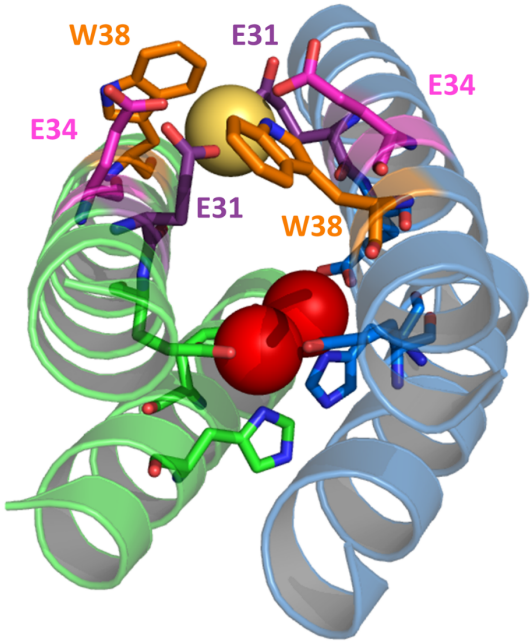
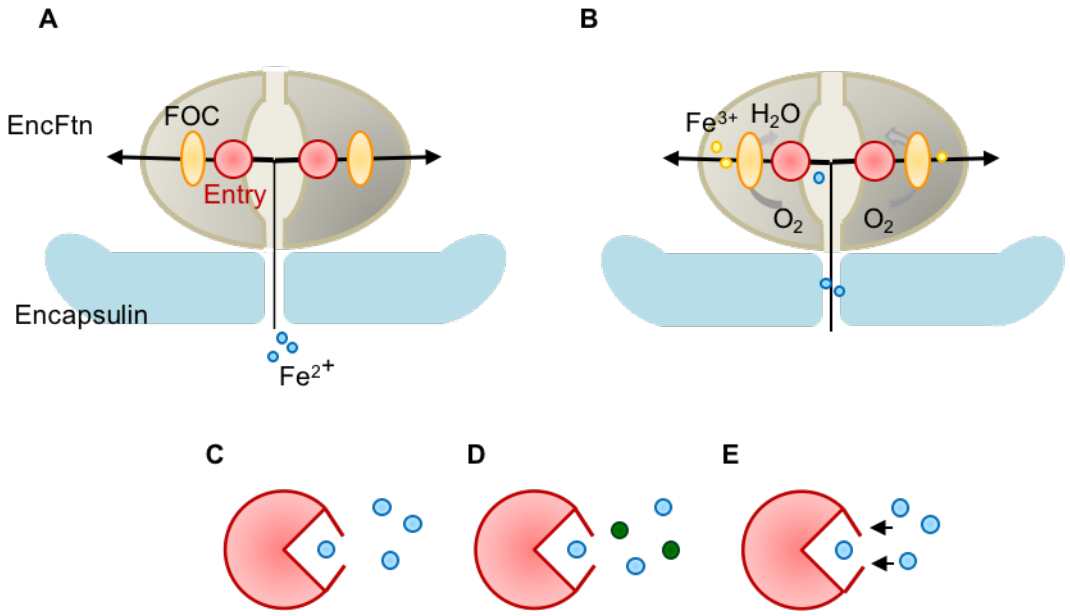
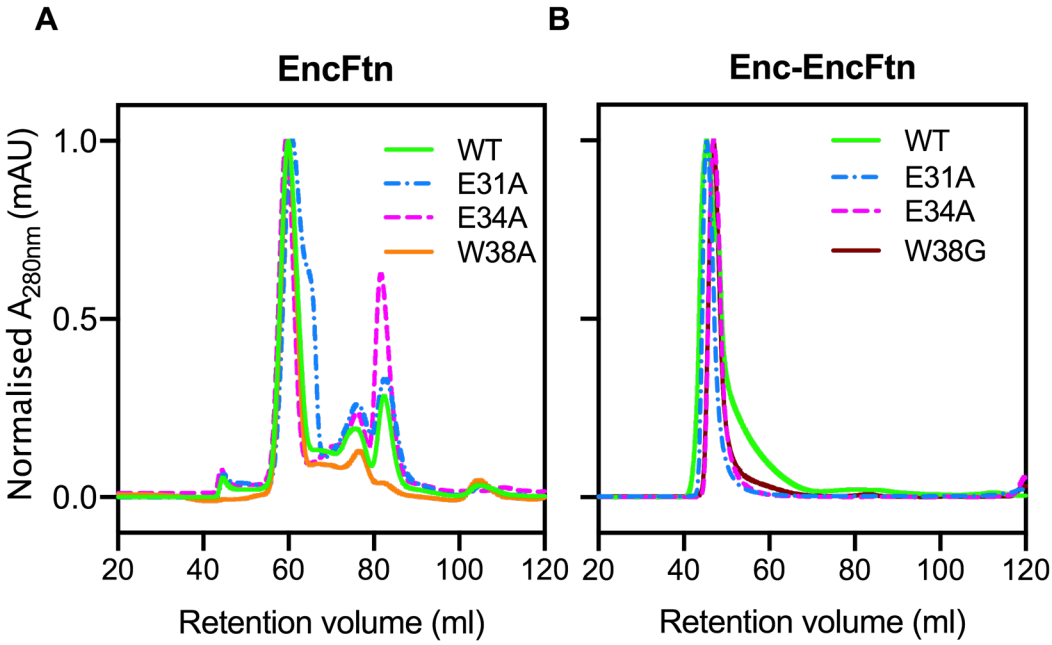


Figure 1. Putative iron ion entry site in *Rhodospirillum rubrum* EncFtn.
Coordination of calcium (yellow sphere) within the dimer interface by Glu31 (purple) and Glu34 (pink) from the two different subunits. Iron ions coordinated in the ferroxidase centre (FOC) are depicted as red spheres. Trp38 side chains (orange) from both chains are also shown.

539



540
541
542 **Figure 2. Illustration of the hypotheses for the functional role of the putative metal ion entry**
543 **site in *Rhodospirillum rubrum* EncFtn.**
544 ***A, B,*** Schematic view of a putative pathway for incoming ferrous ions moving towards the active
545 site (FOC). EncFtn is represented by a grey oval, where the putative entry sites and ferroxidase
546 centres are shown in red and yellow, respectively. A portion of the encapsulin is displayed in pale
547 blue, with emphasis given to the 5-fold pore used by ferrous ions to reach the compartmentalised
548 ferritin protein. Ferrous ions (blue circles) are shown to be converted into ferric ions (yellow
549 circles) in the ferroxidase centre with the concomitant reduction of molecular oxygen (***B***). Several
550 hypotheses (gating (***C***), selectivity filter (***D***), and attraction (***E***)) on the role played by the residues
551 contributing to the putative entry site are shown in ***C-E***. Ferrous ions are depicted as blue circles,
552 while non-cognate species are in dark green.
553



554
555 **Figure 3. Purification of recombinant *R. rubrum* EncFtn and Enc:EncFtn protein complexes.**
556 Recombinant EncFtn (***A***) and Enc:EncFtn (***B***) proteins were purified by anion exchange column
557 (Hi-Trap Q-sepharose FF, GE Healthcare) and then subjected to analytical size-exclusion
558 chromatography using a Superdex 200 16/60 column (GE Healthcare) previously equilibrated with
559 50 mM Tris-HCl, pH 8.0, 150 mM NaCl. Investigation of the elution profiles in Figure 3A reveal
560 that all EncFtn proteins show the main peak at around 60 ml, diagnostic of oligomerisation states
561 close to 20-mer, and smaller peaks at ~ 76 ml and 82 ml (indicative of smaller assembly states,
562 Supplementary Table S3). Enc:EncFtn proteins elute in a single peak around 46 ml, suggesting that
563 EncFtn is compartmentalised within the Enc shell. doi. 10.6084/m9.figshare.9885557
564
565

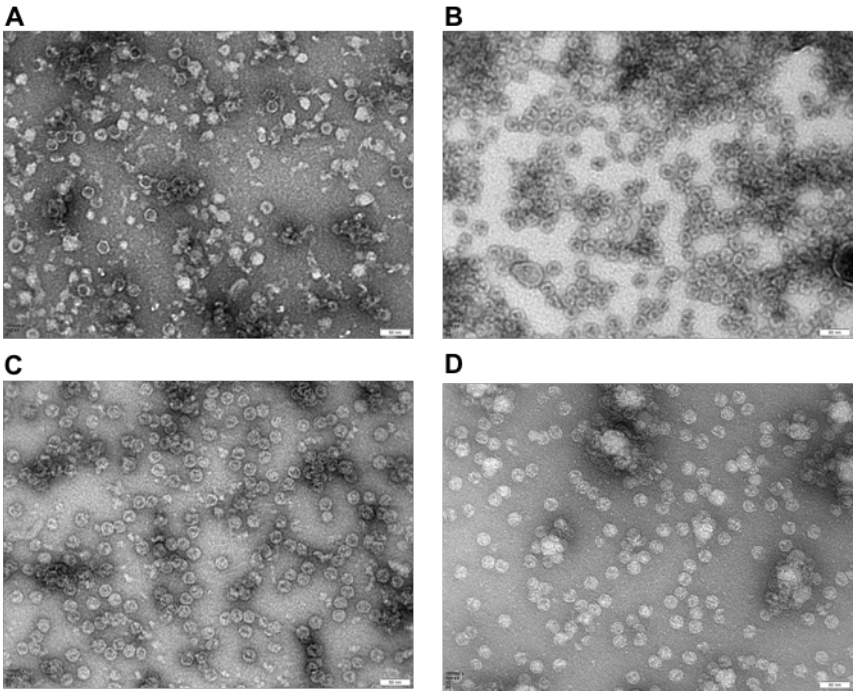
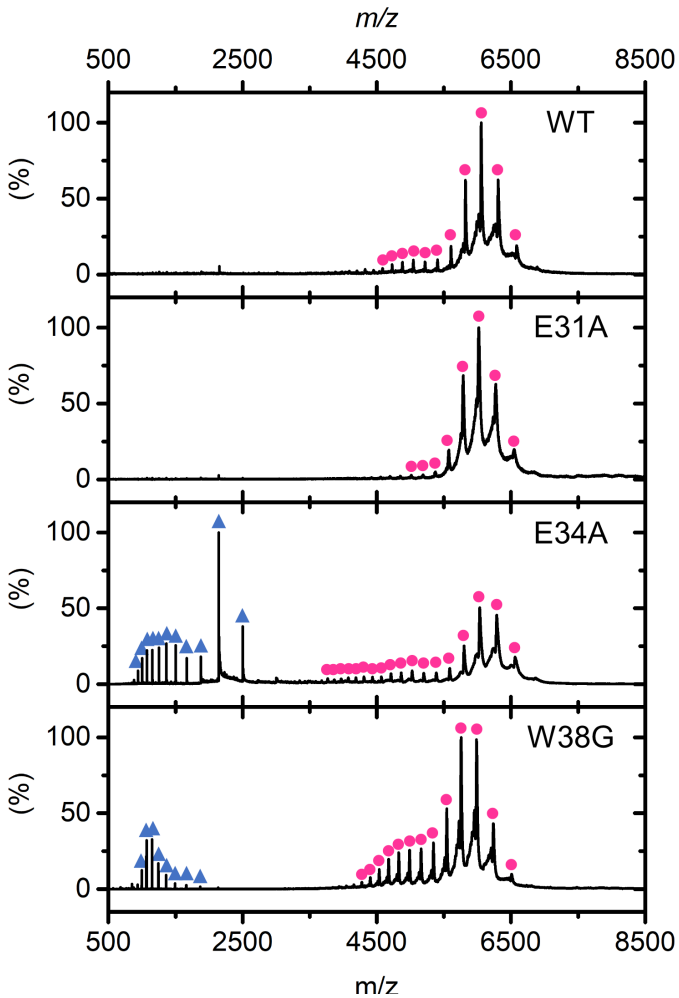
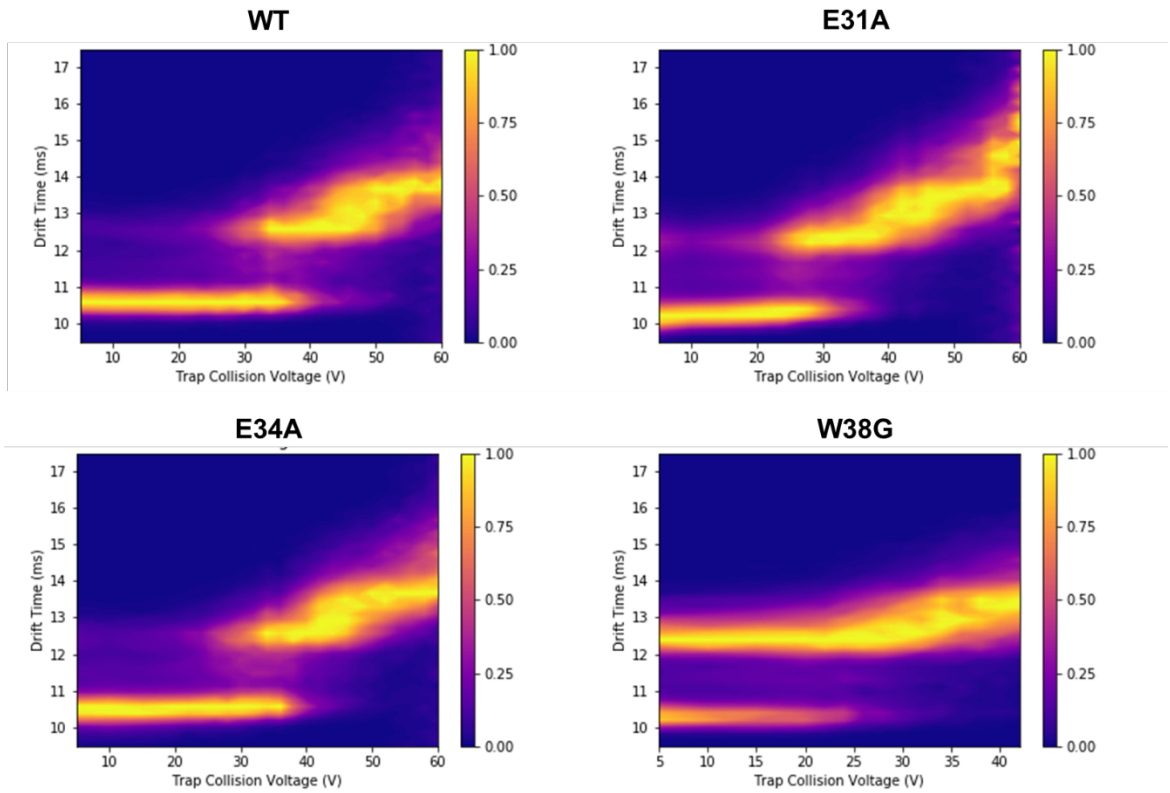


Figure 4. Transmission electron micrographs of *Rhodospirillum rubrum* Enc:EncFtn complexes.
Negative stain TEM images of *R. rubrum* Enc:EncFtn nanocompartments stained with uranyl acetate. Enc:EncFtn_WT (***A***), Enc:EncFtn_E34A (***B***), Enc:EncFtn_W38G (***C***), and Enc:EncFtn_E31A (***D***). Samples were imaged at a magnification of 100,000, scale bars correspond to 50 nm. Proteins were diluted in Buffer GF to 0.1 mg/ml.



576
577
578 **Figure 5. Native MS Analysis of *Rhodospirillum rubrum* EncFtn variants.**
579 All spectra were acquired in 100 mM ammonium acetate, pH 8.0. nESI spectra of EncFtn variants
580 displaying decamer (pink circles) and monomer (blue triangle) charge state distributions.
581
582
583
584
585



586
587
588 **Figure 6. Collision Induced unfolding (CIU) of the *Rhodospirillum rubrum* EncFtn variants**
589 **determined by ion mobility mass spectrometry.**
590 The ion mobility drift time of the 26+ charge state of each EncFtn variant is displayed as a function
591 of increasing collision voltage. All variants display similar CIU profiles, with a compact structural
592 conformation (10.5 ms) present at low activation which transitions to a more extended
593 conformation (drift time 12.5-13.5 ms) as activation increases. The onset of this transition occurs at
594 around 32 V in WT, E31A, and E34A; However, the transition occurs at significantly lower
595 activation energy in W38G and the extended structural conformation (12.5 ms) dominates the
596 profile, even at low activation voltages.
597
598

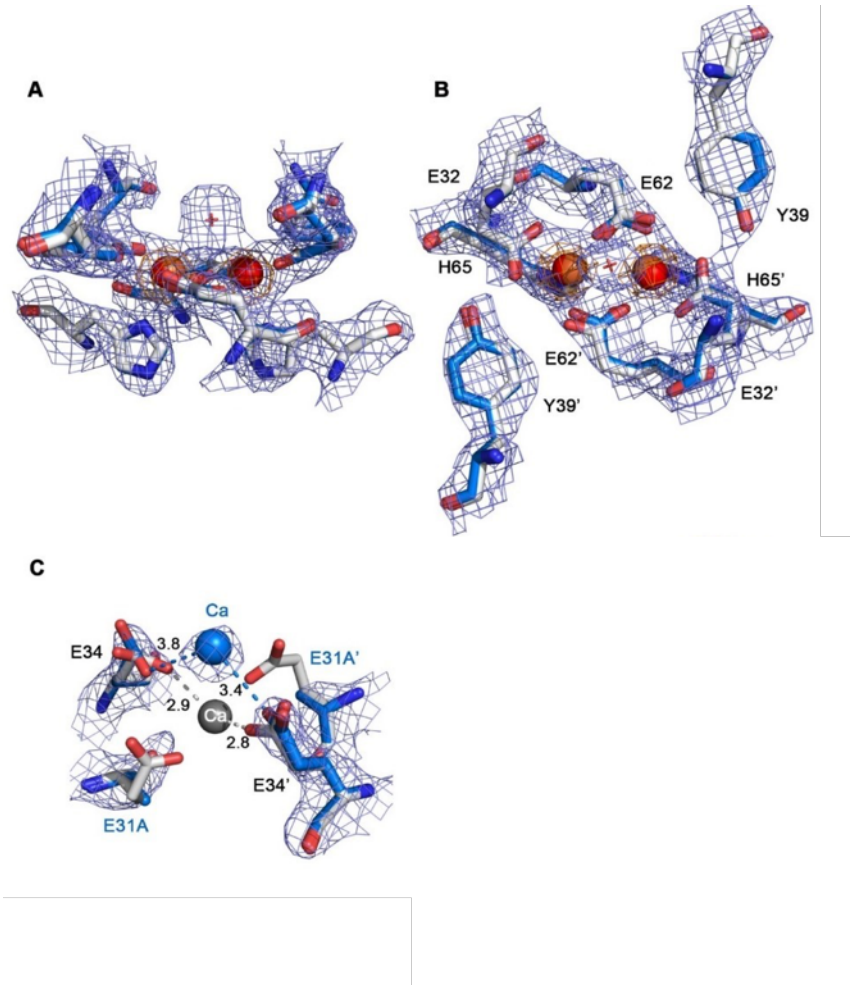


Figure 7. Comparison of metal binding sites in EncFtn_sH wild type and E31A structures. Metal binding sites in the EncFtn E31A variant and WT are shown as stick representations with Cα coloured in marine and white, respectively. The final experimental 2mFo-DFc density map is shown contoured at 1.5 σ around the metal binding sites of the E31A variant as a blue mesh. An NCS-averaged anomalous difference map is superimposed on the FOC of E31A variant as an orange mesh and contoured at 20 σ. Residues from dimers are differentiated with a prime symbol in their labels. The distance between metal ions and coordinating residues is marked by dashed lines with distances shown in Å. ***A***, Aligned FOCs with di-iron bound as orange spheres in E31A mutant and red spheres in WT. A water molecule shown as a red cross is coordinated above the di-iron of E31A mutant. ***B***, Top view of ***A*** with residues labelled. ***C***, Metal ion entry site comparison between E31A mutant and WT with both Ca(II) bound in grey and brown spheres, respectively. In the E31A mutant the Ca(II) moves further from FOC and is coordinated by E34 from two chains in a distance of 3.4 Å/3.8 Å by contrast to the 2.8 Å/2.9 Å in the WT.

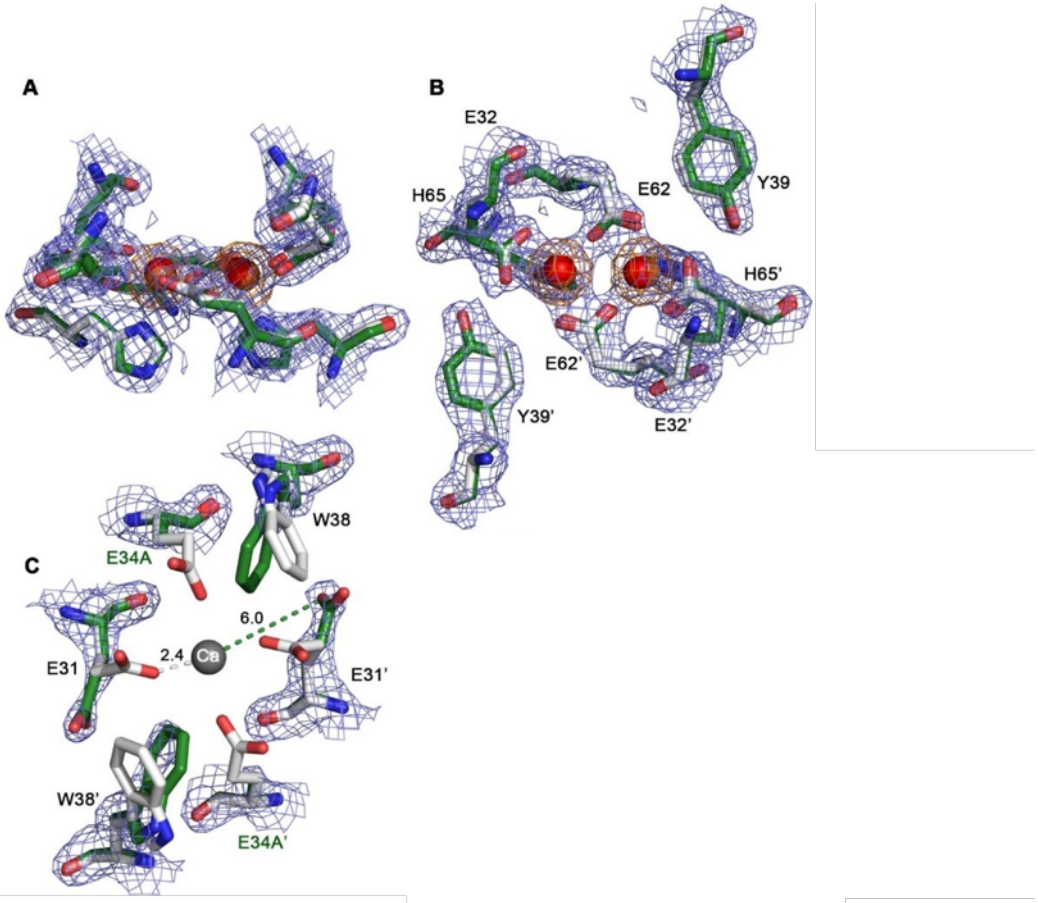


Figure 8. Comparison of metal binding sites in the E34A and WT *Rhodospirillum rubrum* EncFtn_sH structures.

Metal binding sites in the EncFtn E34A mutant and WT are shown as stick representations with Cα coloured in forest and white, respectively. The final calculated 2mFo-DFc density map is shown contoured at 1.5 σ around the metal binding sites of the E34A variant as a blue mesh. An NCS-averaged anomalous map is superimposed on FOC of E34A mutant as an orange mesh and contoured at 20 σ. Residues from dimers are differentiated with a prime symbol in their labels. The distance between metal ions and coordinating residues is marked by dashed lines and labelled with numbers nearby and Å as the unit. ***A***, Aligned FOCs with di-iron ions bound as orange spheres in E34A variant and red spheres in WT. ***B***, Top view of ***A*** with residues labelled. ***C***, Metal ion entry site comparison between E34A variant and WT. No metal is coordinated in this site of E34A mutant while WT binds a Ca(II) shown in grey sphere. The E31 in E34A mutant moves away by 117° and would be 6 Å away from the position of the Ca(II) of the WT. The indole ring of W38 rotates towards the central Ca(II) site of the WT by 38°.

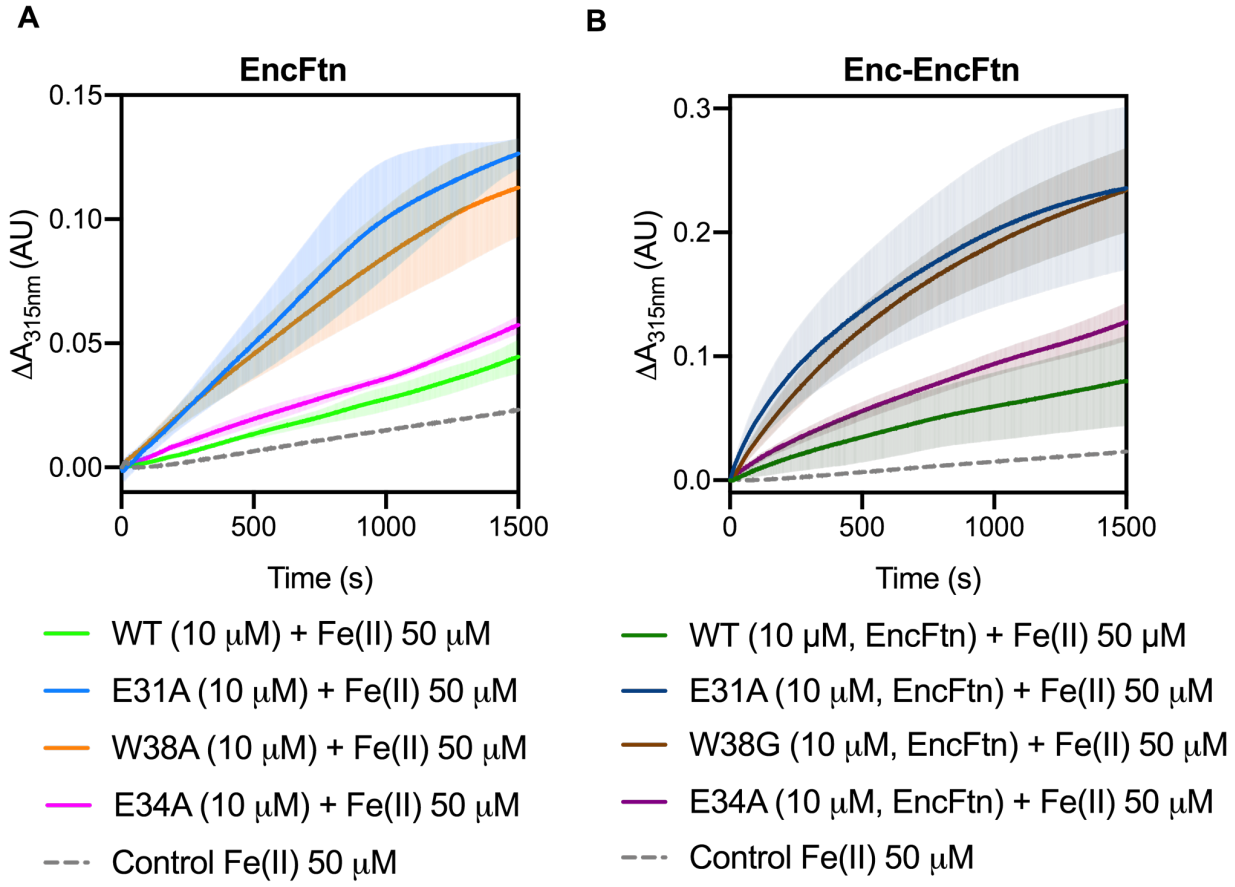


Figure 9. Ferroxidase activity of *Rhodospirillum rubrum* EncFtn and Enc:EncFtn complexes. ***A***, EncFtn wild-type (green line), EncFtn_E34A (pink line), EncFtn_E31A (blue line), and EncFtn_W38A (orange line), (10 μM, monomer) were incubated with 50 μM $FeSO_4.7H_2O$ (10 times molar equivalent Fe(II) per FOC) and progress curves of the oxidation of Fe(II) to Fe(III) was monitored at 315 nm at room temperature. The background oxidation of iron at 50 μM in enzyme-free control is shown for reference (dotted grey line). Solid lines represent the average (n = 3) of technical replicates, shaded areas represent standard deviation from the mean. Protein and iron samples were prepared anaerobically in Buffer H (10 mM HEPES pH 8.0, 150 mM NaCl), and 0.1 % (v/v) HCl, respectively. ***B***, Enc:EncFtn wild-type (dark green line), Enc:EncFtn_E34A (dark pink line), Enc:EncFtn_E31A (dark blue line), and Enc:EncFtn_W38G (chestnut line) (25 μM Enc:EncFtn, corresponding to 10 μM EncFtn monomer) were incubated with 50 μM $FeSO_4.7H_2O$ and progress curves of the oxidation of Fe(II) to Fe(III) was monitored at 315 nm at room temperature. The background oxidation of iron at 50 μM in enzyme-free control is shown for reference (dotted grey line). Solid lines represent the average (n = 3) of technical replicates, shaded areas represent standard deviation from the mean. Protein and iron samples were prepared anaerobically in Buffer H (10 mM HEPES pH 8.0, 150 mM NaCl), and 0.1 % (v/v) HCl, respectively. doi.10.6084/m9.figshare.9885575

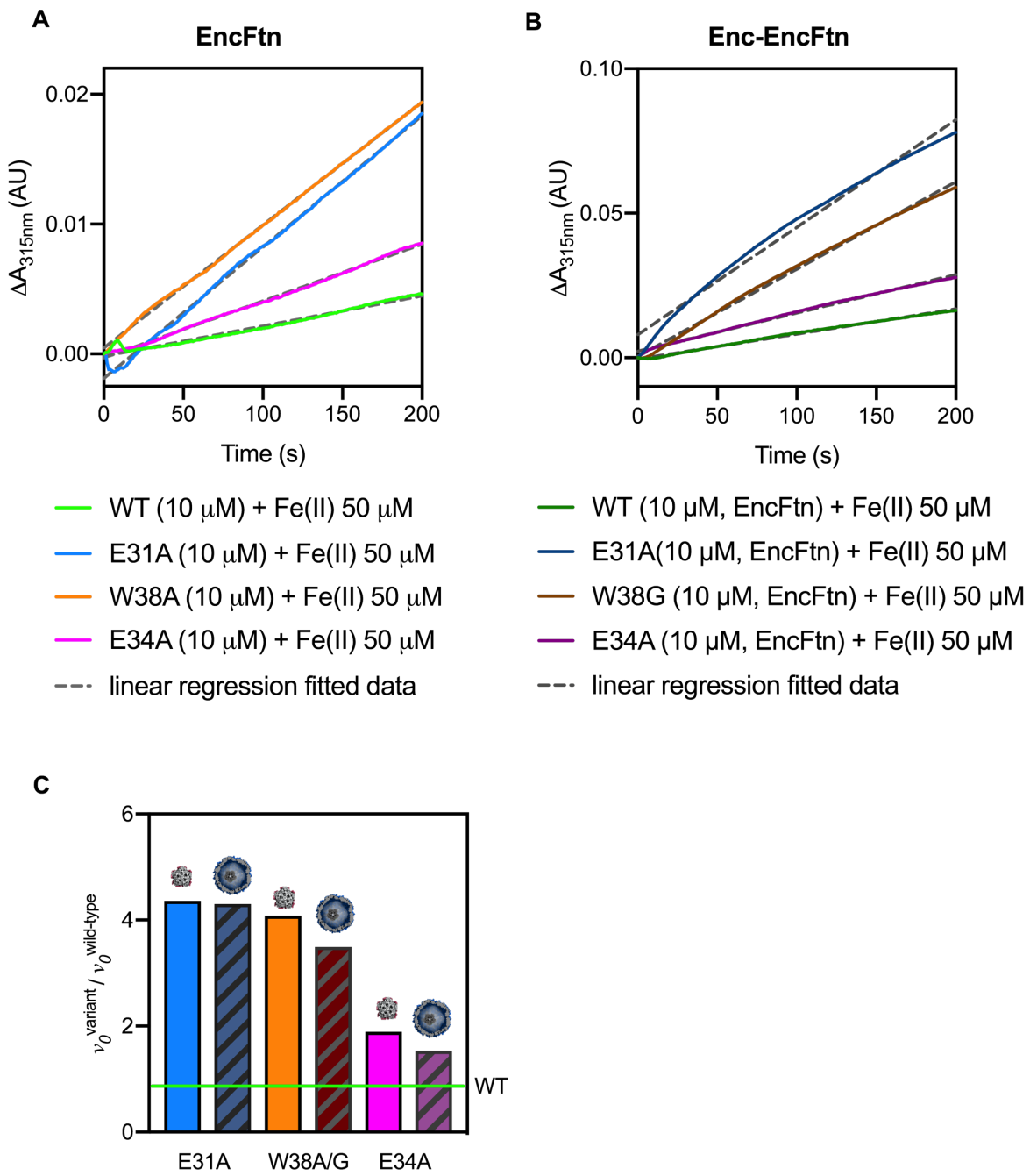


Figure 10. Linear Regression on first 200s of ferroxidase assays with EncFtn and Enc:EncFtn complexes.

A, Data from first 200 s of the assays shown in Figure 9*A* were processed using the Linear Regression tool in GraphPad (Prism8) to determine v_0, initial enzymatic rate, from slope of each of the linear curves (see Table S5). ***B***, Data from first 200 s of the assays shown in Figure 9*B* were processed using the Linear Regression tool in GraphPad (Prism8) to determine v_0, initial enzymatic rate, from slope of each of the linear curves (see Table S6). ***C***, Variants initial enzymatic rate ($v_0^{variant}$) calculated from linear regression curves in the first 200 s of ferroxidase assay experiments were divided by the slope calculated for the corresponding wild-type species ($v_0^{wild\text{-}type}$). Evaluation of the resulting $v_0^{variant}/ v_0^{wild\text{-}type}$ factors allow comparison between enzymatic kinetics of EncFtn (solid coloured columns) and Enc:EncFtn (striped columns) proteins. Colour coding is consistent with Figures 9*A,B* and Figures 10*A,B* and columns are labelled with corresponding mutation. Horizontal green line represents wild type's $v_0^{variant}/ v_0^{wild\text{-}type}$ ratio = 1. Representations of EncFtn or Enc:EncFtn systems (non-descriptive of true size ratios) are shown above each column for clarity. doi.10.6084/m9.figshare.9885575

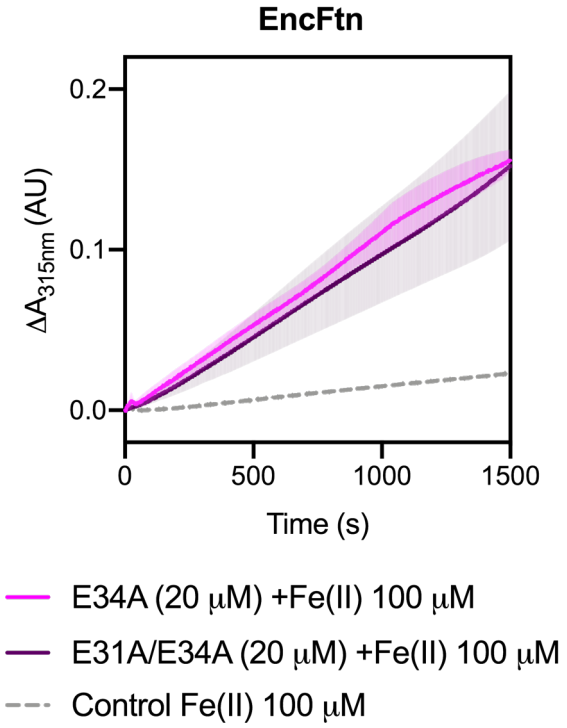


Figure 11. Comparison of ferroxidase activities of *Rhodospirillum rubrum*_EncFtn_E34A and *Rhodospirillum rubrum* EncFtn_E31A/E34A.
EncFtn_E34A (pink line) and EncFtn-E31A/E34A (dark purple line), (20 μM, monomer) were incubated with 100 μM $FeSO_4.7H_2O$ (10 times molar equivalent Fe(II) per FOC) and progress curves of the oxidation of Fe(II) to Fe(III) was monitored at 315 nm at room temperature. The background oxidation of iron at 100 μM in enzyme-free control is shown for reference (dotted grey line). Solid lines represent the average (n = 3) of technical replicates, shaded areas represent standard deviation from the mean. Protein and iron samples were prepared anaerobically in Buffer H (10 mM HEPES pH 8.0, 150 mM NaCl), and 0.1 % (v/v) HCl, respectively.
doi.10.6084/m9.figshare.9885575

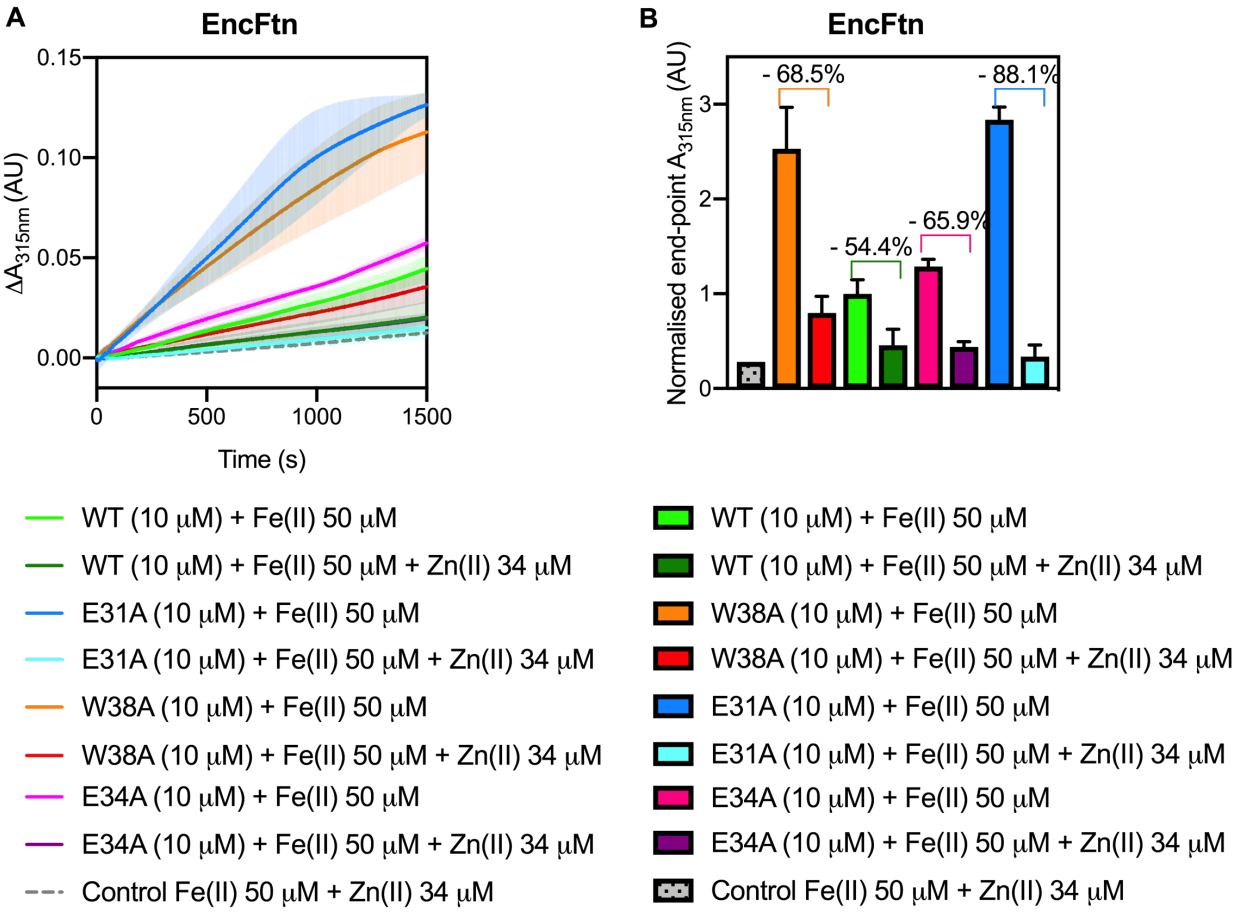


Figure 12. Ferroxidase activity of EncFtn wild-type and variants is inhibited by zinc. ***A***, Comparison between data shown in Figure 9A, collected in the absence of competing metals, are here plotted with curves from ferroxidase assays of EncFtn wild-type and variants carried out using [protein] = 10 μM and [Fe(II)] = 50 μM in the presence of [Zn(II)] = 34 μM. The zinc concentration corresponds to a response inhibited by 50% for Rru_EncFtn_Strep under the same experimental conditions (He et al., 2019). EncFtn wild-type is represented by a dark green line, EncFtn_E34A by a purple line, EncFtn_E31A by a pale blue line, and EncFtn_W38A by a red line. As a control, a mix of Fe(II) and Zn(II) salts (grey dotted line) were assayed in the absence of the enzyme. ***B***, End-point data presented in Figure 12A were plotted as columns to compare inhibition by zinc. Percentage of inhibition is shown above corresponding columns. Colour-coding is consistent with Figure 12A. doi.10.6084/m9.figshare.9885575

695 **Table 1. Average molecular masses of encapsulated ferritins obtained by LC-MS.**

Protein	Observed Mass (Da)	Assignment	Theoretical Mass (Da)
Rru_EncFtn_WT	15060.8 Da	monomer without Met	15061.4 Da
		dimer without Met	
Rru_EncFtn_E31A	15003.6 Da	monomer without Met	15003.4 Da
Rru_EncFtn_E34A	15003.3 Da	monomer without Met	15003.4 Da
Rru_EncFtn_W38G	14932.3 Da	monomer without Met	14932.2 Da

696 The starting Methionine residue is not always retained, and this has been indicated in "Assignment".
697
698
699
700
701
702 **Table 2. Collision cross sections (CCS) of the compact and extended conformations of the Rru-EncFtn**
703 **variants**

Protein	Compact CCS (nm^2)	Extended CCS (nm^2)
Rru_EncFtn_ WT	71.40	80.62
Rru_EncFtn_E31A	69.50	82.65
Rru_EncFtn_E34A	71.30	81.68
Rru_EncFtn_W38G	71.17	81.41

704 CCS values were calculated using Driftscope v2.5.
705
706
707
708

709 **Table 3. Data collection and refinement statistics.**

	*E31A**	*E34A**
Data collection		
Wavelength	1.74	1.74
Resolution range	49.40 - 2.66 (2.75 - 2.66)	48.73 - 2.19 (2.27 - 2.19)
Space group	P 1 2_1 1	P 1 2_1 1
Unit cell (Å) *a* *b* *c*	97.89 120.16 139.37	97.79 120.14 139.34
β (°)	95.33	95.20
Total reflections	310,119 (30,187)	906,663 (86,277)
Unique reflections	170,103 (16,471)	31,4834 (30,221)
Multiplicity	1.8 (1.8)	2.9 (2.9)
Completeness (%)	93.30 (90.35)	96.37 (92.66)
Mean I/sigma(I)	5.45 (1.40)	8.04 (1.43)
Wilson B-factor ($Å^2$)	39.14	33.05
R_{merge}	0.136 (0.657)	0.094 (0.688)
R_{meas}	0.184 (0.878)	0.116 (0.842)
R_{pim}	0.123 (0.578)	0.065 (0.478)
$CC_{1/2}$	0.978 (0.517)	0.995 (0.579)
CC*	0.994 (0.826)	0.999 (0.857)
Refinement		
Reflections used in refinement	170,031 (16,463)	314,739 (30,212)
Reflections used for R_{free}	8,346 (794)	15,590 (1548)
R_{work}	0.195 (0.297)	0.184 (0.282)
R_{free}	0.236 (0.327)	0.212 (0.315)
CC_{work}	0.946 (0.748)	0.959 (0.768)
CC_{free}	0.914 (0.684)	0.952 (0.689)
Model		
Number of non-hydrogen atoms	22,382	22,760
macromolecules	21,988	22,072
ligands	65	48
solvent	329	640
Protein residues	2680	2696
RMS(bonds) (Å)	0.007	0.005
RMS(angles) (°)	1.03	0.87
Ramachandran		
favoured (%)	99.69	99.77
allowed (%)	0.31	0.23
outliers (%)	0.00	0.00
Rotamer outliers (%)	0.26	1.07
Clashscore	4.79	3.01
Average B-factor ($Å^2$)	41.64	40.53
macromolecules	41.66	40.52
ligands	55.57	46.17
solvent	37.08	40.62
PDB ID	6SUW	6SV1

710 Statistics for the highest-resolution shell are shown in parentheses. *Friedel pairs are treated as
711 separate reflections for data analysis and refinement.
712

713 **References**
714
715 Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral
716 GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC,
717 Richardson JS, Terwilliger TC, Zwart PH. 2010. *PHENIX*: a comprehensive Python-based system
718 for macromolecular structure solution. *Acta Cryst*. **D66**:213-221. doi:10.1107/S0907444909052925
719
720 Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M,
721 Terwilliger TC, Urzhumtsev a, Zwart PH, Adams PD. 2012. Towards automated crystallographic
722 structure refinement with phenix.refine. *Acta Cryst.* **D68(4)**:352-67. doi:
723 10.1107/S0907444912001308
724
725 Almirón M, Link AJ, Furlong D, Kolter R. 1992. A novel DNA-binding protein with regulatory and
726 protective roles in starved Escherichia coli. *Genes Dev* **6**:2646–2654. doi:10.1101/gad.6.12b.2646
727
728 Andrews SC. 1998. Iron storage in bacteria. *Adv Microb Physiol* **40**:281-351
729
730 Andrews SC. 2010. The ferritin-like superfamily:evolution of the biological iron storeman from a
731 rubrerythrin like ancestor. *Biochim Biophys Acta – General Subjects* **1800**:691-705.
732 https://doi.org/10.1016/j.bbagen.2010.05.010
733
734 Baaghil S, Lewin A, Moore GR, Le Brun NE. 2003. Core formation in Escherichia coli
735 bacterioferritin requires a functional ferroxidase center. *Biochem* **42**(**47**):14047-56.
736 doi:10.1021/bi035253u
737
738 Behera RK, Torres R, Tosha T, Bradley JM, Goulding CW, Theil EC. 2015. Fe(II) substrate
739 transport through ferritin protein cage ion channels influences enzyme activity and
740 biomineralisation. *J Biol Inorg Chem.* **20(6)**:957-969. doi:10.1007/s00775-015-1279-x
741
742 Bellapadrona G, Stefanini S, Zamparelli C, Theil EC, Chiancone E. 2009. Iron translocation into
743 and out of Listeria innocua Dps and size distribution of the protein-enclosed nanomineral are
744 modulated by the electrostatic gradient at the 3-fold "ferritin-like" pores. *J Biol*
745 *Chem.* **284(28)**:19101-9. doi: 10.1074/jbc.M109.014670
746
747 Cardenas JP, Quatrini R, Holmes DS. 2016 Aerobic lineage of the oxidative stress response protein
748 rubrerythrin emerged in an ancient microaerobic, (hyper)thermophilic environment. *Front*
749 *Microbiol* **7**:1822 doi:10.3389/fmicb.2016.01822
750
751 Chasteen ND, Harrison PM. 1999. Mineralization in ferritin: an efficient means of iron storage. *J*
752 *Struct Biol.* **126**:182-194. doi:10.1006/jsbi.1999.4118
753
754 Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW,
755 Richardson JS, Richardson DC. 2010. *MolProbity*: all-atom structure validation for macromolecular
756 crystallography. *Acta Cryst.* **D66(1)**:12-21 doi: 10.1107/S0907444909042073
757
758 Chiancone E, Ceci P. 2010. The multifaceted capacity of Dps proteins to combat bacterial stress
759 conditions: Detoxification of iron and hydrogen peroxide and DNA binding. *Biochim Biophys Acta*.
760 **1800(8)**:798-805. doi: 10.1016/j.bbagen.2010.01.013
761
762 Choi SH, Baumler DJ, Kaspar CW. 2000. Contribution of *dps* to acid stress tolerance and oxidative
763 stress tolerance in Escherichia coli O157:H7. *Appl Environ Microbiol* **66**:3911–3916.
764 doi:10.1128/AEM.66.9.3911-3916.2000

765
766 Crow A, Lawson TL, Lewin A, Moore GR, Le Brun NE. 2009. Structural basis for iron
767 mineralization by bacterioferritin. *J Am Chem Soc* **131**:6808–6813 doi: 10.1021/ja8093444
768
769 De Martino M, Ershov D, van den Berg PJ, Tans SJ, Meyer AS. 2016. Single-cell analysis of the
770 Dps Response to oxidative stress. *J Bacteriol* **198(11)**:1662-1674. doi: 10.1128/JB.00239-16.
771
772 Dillard BD, Demick JM, Adams MW, Lanzilotta WN. 2011. A cryo-crystallographic time course
773 for peroxide reduction by rubrerythrin from Pyrococcus furiosus. *J Biol Inorg Chem.* **16(6)**:949-59.
774 doi:10.1007/s00775-011-0795-6
775
776 Dixit SM, Polasky DA, Ruotolo BT. 2018. Collision induced unfolding of isolated proteins in the
777 gas phase: past, present, and future. *Curr Opin Chem Biol*. **42**:93-100. doi:
778 10.1016/j.cbpa.2017.11.010
779
780 Ebrahimi HK, Hagedoorn PL, Jongejan JA, Hagen WR. 2009. Catalysis of iron core formation in
781 Pyrococcus furiosus ferritin. *J Biol Inorg Chem* **14(8)**:1265-74 doi:10.1007/s00775-009-0571-z
782
783 Ebrahimi KH, Bill E, Hagedoorn P-L, Hagen W. 2016. Spectroscopic evidence for the role of a site
784 of the di-iron catalytic center of ferritins in tuning the kinetics of Fe(II) oxidation. *Mol BioSyst.*
785 **12**:3576-3588. doi: 10.1039/C6MB00235H
786
787 Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta*
788 *Cryst* **D66**:486-501. doi:10.1107/S0907444910007493
789
790 Evans PR. 2011. An introduction to data resolution: space-group determination, scaling and
791 intensity statistics. *Acta Cryst* **D76(4)**:282-92. doi:10.1107/S090744491003982X
792
793 Frolow F, Kalb AJ, Yariv J. 1994. Structure of a unique twofold symmetric haem-binding site. *Nat*
794 *Struct Biol* 453-460.
795
796 Giessen TW, Orlando BJ, Verdegaal AA, Chambers MG, Gardener J, Bell DC, Birrane G, Liao M,
797 Silver PA, 2019. Large protein organelles form a new iron sequestration system with high storage
798 capacity. *eLife.* **8**:e46070 doi: 10.7554/eLife.46070
799
800 Grant RA, Filman DJ, Finkel SE, Kolter R, Hogle JM. 1998. The crystal structure of Dps, a ferritin
801 homolog that binds and protects DNA. *Nat Struct Biol.* **5**:294-303.
802
803 Hagen WR, Hagedoorn P-L, Ebrahimi KH. 2017. The working of ferritin: a crossroad of options.
804 *Metallomics*. **9**:595-605.
805
806 Harrison PM, Arosio P. 1996. The ferritins: molecular properties, iron storage function and cellular
807 regulation. *Biochim Biophys Acta* **1275**:161-203.
808
809 He D and Marles-Wright J. 2015. Ferritin family proteins and their use in biotechnology. *New*
810 *Biotech* **32**:651-657 doi:org/10.1016/j.nbt.2014.12.006
811
812 He D, Hughes S, Vanden-Hehir A, Georgiev K, Altenbach K, Tarrant E, Mackay CL, Waldron KJ,
813 Clarke DJ, Marles-Wright J. 2016. Structural characterisation of encapsulated ferritin provides
814 insight into iron storage in bacterial nanocompartments. *eLife* **5**. doi:10.7554/eLife.18972
815

He D, Piergentili C, Ross J, Tarrant E, Tuck LR, Mackay CL, McIver Z, Waldron KJ, Clarke DJ, Marles-Wright J. 2019. Conservation of the structural and functional architecture of encapsulated ferritins in bacteria and archaea. *Biochem J* **476(6)**:975-989. doi:10.1042/BCJ20180922

Kabsch W. 2010. Integration, scaling, space-group assignment and post-refinement. *Acta Cryst.* **D66(2)**:133-44. doi:10.1107/S0907444909047374

Karlplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. *Science* **336(6084)**:1030-33. doi:10.1126/science.1218231

Liu XS, Patterson LD, Miller MJ, Theil EC. 2007. Peptides selected for the protein nanocage pores change the rate of iron recovery from the ferritin mineral. *J Biol Chem*. **282(44)**:31821-5. doi:10.1074/jbc.C700153200

Martinez A, Kolter R. 1997. Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J Bacteriol*. **179**:5188-5194. doi:10.1128/jb.179.16.5188-5194.1997

Masuda T, Goto F, Yoshihara T, Mikami B. 2010. Crystal structure of plant ferritin reveals a novel metal binding site that functions as a transit site for metal transfer in ferritin. *J Biol Chem* **285**:4049-4059. doi:10.1074/jbc.M109.059790

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MS, Storoni LC, Read RJ. 2007. *Phaser* crystallographic software. *J Appl Cryst*. **40**:658-674. doi:10.1107/S0021889807021206

Nair S, Finkel SE. 2004. Dps protects cells against multiple stresses during stationary phase. *J Bacteriol* **186**:4192–4198. doi:10.1128/JB.186.13.4192-4198.2004

Pesek J, Büchler R, Albrecht R, Boland W, Zeth K. 2011. Structure and mechanism of iron translocation by a Dps protein from Microbacterium arborescens. *J Biol Inorg Chem*. **286**:34872-34882. doi:10.1074/jbc.M111.246108

Pfaffen S, Abdulqadir R, Le Brun NE, Murphy MEP. 2013. Mechanism of ferrous iron binding and oxidation by ferritin from a pennate diatom. *J Biol Inorg Chem*. **288**:14917-14925. doi:0.1074/jbc.M113.454496

Proulx-Curry PM, Chasteen ND. 1995. Molecular aspects of iron uptake and storage in ferritin. *Coord Chem Rev* **144**:347–368. doi:10.1016/0010-8545(95)01148-I

Ratnayake DB, Wai SN, Shi Y, Amako K, Nakayama H, Nakayama K. 2000. Ferritin from the obligate anaerobe *Porphyromonas gingivalis*: purification, gene cloning and mutant studies. *Microbiol* **146**:1119-1127. doi:10.1099/00221287-146-5-1119

Ross J, Lambert T, Piergentili C, He D, Waldron KJ, Mackay CL, Marles-Wright J, Clarke DJ. 2019. Mass spectrometry reveals the assembly pathway of encapsulated ferritins and highlights a dynamic ferroxidase interface. *Chem Comm* (submitted for review) doi xxxx

Schägger H. 2006. Tricine-SDS-PAGE. *Nat Protoc*. **1(1)**:16-22. doi:10.1038/nprot.2006.4

Sztukowska M, Bugno M, Potempa J, Travis J, Kurtz DM Jr. 2002. Role of rubrerythrin in the oxidative stress response of Porphyromonas gingivalis. *Mol Microbiol.* **44(2)**:479-88. doi:10.1046/j.1365-2958.2002.02892.x

868 Theil EC, Behera RK, Tosha T. 2013. Ferritins for chemistry and for life. *Coord Chem Rev*.
869 **257(2)**:579-586. doi:10.1016/j.ccr.2012.05.013
870
871 Treffry A, Harrison PM. 1978. Incorporation and release of inorganic phosphate in horse spleen
872 ferritin. *Biochem J*. **171**:313–320.
873
874 Treffry A, Zhao Z, Quail MA, Guest JR, Harrison PM. 1998. How the presence of three iron
875 binding sites affects the iron storage function of the ferritin (EcFtnA) of *Escherichia coli*. *FEBS*
876 **432**:1873-3468
877
878 Xu B, Chasteen ND. 1991. Iron oxidation chemistry in ferritin. Increasing Fe/O2 stoichiometry
879 during core formation. *J Biol Chem.* **266(30)**:19965-70.
880
881 Yang X, Chen-Barrett Y, Arosio P, Chasteen ND. 1998. Reaction paths of iron oxidation and
882 hydrolysis in horse spleen and recombinant human ferritins. *Biochem* **37(27)**:9743-50.
883
884 Yao H, Wang Y, Lovell S, Kumar R, Ruvinsky AM, Battaile KP, et al. 2012. The structure of the
885 BfrB-Bfd complex reveals protein-protein interactions enabling iron release from bacterioferritin.
886 *J Am Chem Soc.* **134**:13470-13481. doi:10.1021/ja305180n
887
888 Zhao G, Bou-Abdallah F, Arosio P, Levi S, Janus-Chandler C, Chasteen ND. 2003. Multiple
889 pathways for mineral core formation in mammalian apoferritin. The role of hydrogen peroxide.
890 *Biochem* **42(10)**:3142-50. doi:10.1021/bi027357v
891
892 Zheng H, Cooper DR, Porebski PJ, Shabalin IG, Handing KB, Minor W. 2017. CheckMyMetal: a
893 macromolecular metal-binding validation tool. *Acta Crystallogr D Struct Biol*. **73(3)**:223–233.
894 doi:10.1107/S2059798317001061
895

Competing interests

The authors declare that no competing interests exist.

Author contributions

CP, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article.

DH, Conception and design, Acquisition of data, Analysis and interpretation of data.

JR, Conception and design, Acquisition of data, Analysis and interpretation of data. Drafting or revising the article.

WAS, Acquisition of data.

LA, Acquisition of data.

CLM, Acquisition of data.

KJW, Acquisition of data, Interpretation of data.

DJC, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article.

JMW, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article.

Funding

This work was supported a Royal Society Research Grant awarded to JMW [RG130585] and a BBSRC New Investigator Grant to JMW and DJC [BB/N005570/1]. CP and WAS are funded by the BBSRC New Investigator Grant [BB/N005570/1]. JMW is funded by Newcastle University. DJC and JR are funded by the University of Edinburgh. JR is funded by a BBSRC EastBio DTP studentship [BB/M010996/1]. DH was funded by the China Scholarship Council. KJW was funded by the Wellcome Trust and Royal Society through a Sir Henry Dale Fellowship awarded to KJW [098375/Z/12/Z].

Abbreviations

Enc, encapsulin; EncFtn, encapsulated ferritin; FOC, ferroxidase centre; MS, mass spectrometry; nESI, Native nanoelectrospray ionization; CIU, collision induced unfolding.