Dissecting the structural and functional roles of a putative metal entry site in encapsulated ferritins

Dissecting the structural and functional roles of a putative metal entry site in encapsulated ferritins

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

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

Email: Dave.clarke@ed.ac.uk

Tel: +44(0)131 650 4808

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

Tel: +44(0)191 2084855

Running title: Roles of a secondary iron site in encapsulated ferritins

Keywords: *Ferritin, Crystal structure, Mass spectrometry (MS)*

Encapsulated ferritin, Encapsulin, Ferroxidase, Native Mass Spectrometry

Abstract
Encapsulated ferritins belong to the universally distributed ferritin superfamily, which function as iron detoxification and storage systems. Encapsulated ferritins have a distinct annular structure and must associate with an encapsulin nanocage to form a competent iron store that is 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 organization of the ferroxidase catalytic site and neighboring secondary metal binding sites. We have previously identified a putative metal binding site on the inner surface of the *Rhodospirillum rubrum* encapsulated ferritin at the interface between the two-helix subunits and proximal to the ferroxidase center. 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 center and suggest a dual role for the secondary site, which both serves to attract metal ions to the ferroxidase center and acts as a flow-restricting valve to limit the activity of the ferroxidase center. Moreover, confinement of encapsulated ferritins within the encapsulin nanocage, while enhancing the ability of the encapsulated ferritin to undergo catalysis, does not influence the function of the secondary site. Our study demonstrates a novel molecular mechanism by which substrate flux to the ferroxidase center is controlled, potentially to ensure that iron oxidation is productively coupled to mineralization.

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 which forms an iron-binding di-iron ferroxidase active site (the ferroxidase center, FOC) (1). The quaternary structure of these protein families shows a high degree of variation: from monomeric rubrerythrin proteins (2, 3); dimeric iron-mineralizing encapsulin-associated firmicute (IMEF) proteins (4); dodecameric DNA Protection in Starved (DPS) cells proteins and DPS-like (DPSL) proteins (5–9); the 24-meric classical ferritins (Ftn) and bacterioferritins (Bfr) (10–13); and the decameric encapsulated ferritins (EncFtn) (14, 15). The rubrerythrin and DPS proteins play key roles in the protection of cells from oxidative stress, while DPS and the other members of the ferritin superfamily are vital iron stores (7, 16–22). These latter proteins oxidize and convert free Fe(II) to inert Fe(III) oxyhydroxide or Fe(III) phosphate mineral forms that provide a reservoir of bioavailable iron (23). The nanocages formed by the quaternary structure of DPS, DPSL, Ftn, and Bfr allow iron oxidation and mineralization to occur within a biochemically privileged environment within the cell, which protects the host cell from iron-induced oxidative stress. Interestingly, the recently described IMEF and EncFtn proteins both have ferroxidase activity, but they do not form nanocage structures on their own, instead they must be localized to the lumen of encapsulin nanocages to form a competent iron-store (4, 14, 15). EncFtn and IMEF are directed into encapsulins by a terminal localization sequence that binds to the interior wall of encapsulin. The EncFtn proteins form a distinct homodecameric annular assembly, comprising a pentamer of dimers with D5 symmetry (Figure 1A). Each EncFtn subunit is formed of two long anti-parallel helices (α1 and α2) followed by a shorter helix (α3) (Figure 1B); this is in contrast to the classical ferritins, which have 4-helix subunits. The annular assembly has two classes of 2-fold symmetry forming two distinct dimer interfaces. One of the dimer assemblies contains the FOC across the dimer interface, with the two iron ions bound within the interface (Figure 1C, FOC dimer). The other dimer forms a back-to-back conformation, exposing the metal ion binding residues and is known as the non-FOC dimer (Figure 1D) (24).

A model for the engagement of EncFtn decamers with the encapsulin nanocage has been proposed based on the crystallographic model of the *Thermotoga maritima* encapsulin complex (15, 25). In this model the C-terminal

encapsulation sequence of the EncFtn interacts with a specific binding site on the encapsulin monomer, which is oriented to face the lumen of the nanocage (Figure S1). This model places the central hole of the EncFtn decamer directly under the pores formed at the icosahedral 5-fold axis of the encapsulin nanocage, which would form a pathway for iron to transit across the encapsulin shell towards the putative metal entry site. The residues lining the five-fold pore in the encapsulin nanocage are not well conserved between species; and based on crystallographic evidence, this pore is primarily made up of main-chain atoms, implying that this pore does not have a role as a selectivity filter.

The catalytic route to iron mineralization in the different ferritin iron-stores is still a subject of vigorous debate (23, 26–30). Given the lack of absolute conservation in structure of the ferroxidase center, it is likely that there is no single conserved mechanism for their activity (23, 26–30). All of the ferritin family proteins that have been structurally characterized have additional iron-binding sites proximal to the ferroxidase center. In the case of the sites proximal to the ferroxidase center identified in DPS and ferritins, these have been suggested to be iron-entry, transit, and exit points from the ferroxidase center (31–33), or the sites of electron transfer from Fe(II) substrates to the FOC (34). The subunit boundaries in the DPS and ferritin nanocages form a variety of symmetrical pores, which have variously been suggested, or shown, to be entry and exit sites to the lumen for iron, water and phosphate substrates (4, 10, 35–37). The proposed iron entry site in DPSL is distinct to other ferritin family proteins, with an unusual arrangement of cysteine residues and a disulfide bond (8). Furthermore, an additional metal-binding site located on the inner surface of the dodecamer has been identified in members of the DPSL proteins, because of its location this site is reminiscent of the third iron site in *E. coli* ferritins and bacterioferritins and it is proposed to be involved in the iron translocation through the protein shell (9). Whatever their function, these secondary metal sites are essential to the correct functioning of the ferritin iron-stores (38–40).

In our previous work we identified a number of conserved secondary metal-binding sites in EncFtn proteins (14, 15). A partially occupied metal ion site was present on the outer surface (proposed exit site) of the EncFtn decamer, and a highly occupied metal ion site was present on the inner face (proposed entry site), which is proximal to the ferroxidase center (Figure 2A and B). 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 center (24). A comparison of these sites to secondary metal binding sites in other ferritin family members highlights the fact that 'C' site, first identified in the *E. coli* ferritin (41, 42), is topologically similar to the proposed iron-exit site identified in our previous work, whereas the entry site is unique to the EncFtn family (15). In our structure of the *R. rubrum* EncFtn protein, the entry site coordinates a calcium ion from the crystallization solution in a symmetrical arrangement of glutamic acid residues (E31/E34) and a tryptophan residue (W38) contributed by each chain from the ferroxidase center dimer (24) (Figure 2B). 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 are similar to the Fe – Glu – Ox coordination distances seen in the FOC of 2.2-2.6 Å.

We hypothesize that this site acts as an entry site for iron ions approaching the ferroxidase center, and acts in some way to control the access of metal ions to the active site of the protein (Figure 2C). The glutamic acid residues in the entry 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 the entry site, including

acting as a flow restrictor valve to limit the flow of metal ions to the FOC ("valve" function); an electrostatic funnel to direct metal ions to the FOC ("attraction" function); a discriminatory filter to avoid mis-metalation and consequent inhibition of the active site ("filter" function) (Figure 2D/E/F); or for protein stability and assembly ("stabilizing" function). It is also possible that these functions may overlap, and the entry site may play a number of key roles in the stability and activity of the encapsulated ferritins.

In order to understand the role of the entry site, a number of single-point amino acid variants of the *R. rubrum* EncFtn protein were produced to replace the residue of interest with non-metal binding residues. These variants were produced in the isolated *R. rubrum* EncFtn protein alone and in 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 analyzed the solution and gas phase behavior 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, finding that their structures display altered entry site geometry and the abrogation of metal binding in this site. The functional role of the entry site was explored biochemically; we show that changes to the metal coordinating amino acids in this site enhance the ferroxidase activity of the EncFtn protein, while increasing its susceptibility to inhibition by competing zinc ions. Taken together, these results indicate that the entry site plays a minor role in stabilizing the oligomeric state of the EncFtn decamer, and has a dual role as a site of metal ion attraction and as a flow restrictor valve to limit metal access to the ferroxidase center and thus slow the oxidation of ferrous ions to ferric ions.

Results

Changes to the metal ion entry site of Encapsulated ferritins affect their oligomeric state in solution

In order to explore the functional roles of the residues in the entry site of the *R. rubrum* EncFtn, single point variants were produced by site-directed mutagenesis, removing each metal binding residue in turn and substituting each with alanine or glycine. The average molecular masses of the EncFtn proteins were confirmed by liquid chromatography-mass spectrometry (LC-MS, Table S1). The resulting variants were also generated in association with the encapsulin (Enc) to produce Enc:EncFtn complexes.

Changes to the ferroxidase center affect both the assembly and activity of the *R. rubrum* EncFtn (EncFtn-WT) (15). To investigate whether changes to the entry site influence the quaternary structure of the EncFtn protein, each variant was subjected to size-exclusion chromatography using a HiLoad 16/60 Superdex 200 (S200) column calibrated with standards of known molecular mass (Figure 3A and B). Each of the variant EncFtn proteins eluted with a similar profile, with a major peak at 60 ml retention volume, a second peak at 75 ml, and the final peak at 82 ml (Figure 3A, Figure S2A, Table S2). The E34A variant had significantly more protein in the 82 ml peak when compared to the EncFtn-WT and E31A variant, suggesting that this change affects the stability of the protein complex. Based on our prior work, these peaks correspond to decamer, tetramer, and dimer species, respectively (14, 24).

In contrast to the EncFtn variants, the Enc:EncFtn protein complexes all eluted as a single peak within the void volume of the S200 column, indicating the formation of protein complexes larger than 600,000 Da (separation range M_r for this column is 10,000-600,000 Da), hence not able to enter the matrix pores (Figure 3B, Figure S2B, Table S2). SDS-PAGE analysis of peak fractions confirmed the presence of both the EncFtn and encapsulin proteins (Figure S3A). We performed transmission electron microscopy to determine if these complexes assembled into encapsulin nanocages (Figure S2D/E and Figure S3B). All samples showed characteristic 25 nm encapsulin cages, with visible internal density

for the EncFtn proteins. This is consistent with our previous observations of the *R. rubrum* encapsulin:EncFtn complex (15).

X-ray crystal structures of EncFtn variants show altered metal binding within the entry site

Given the observation that the E31A and E34A EncFtn variants display some differences in terms of oligomeric state and stability in solution, we determined the X-ray crystal structures of C-terminal hexa-histidine tagged variants of these proteins (EncFtn-sH). Both the E31A and E34A variants crystallized in similar conditions to EncFtn-WT; however, we were unable to obtain crystals of either the W38G or W38A variants of the EncFtn protein, suggesting that removing this tryptophan destabilizes the quaternary structure of the protein sufficiently to present a barrier to crystallization.

The structures of the *R. rubrum* EncFtn-sH E31A and E34A variants were determined by X-ray crystallography to 2.66 and 2.19 Å resolution respectively. They adopt the same overall quaternary fold as EncFtn-WT, with root-mean-square deviation (RMSD) Cα differences of 0.2 A^2 over 91 Cα and 0.17 $Å^2$ over 90 Cα for the E31A and E34A variants respectively, and an RMSD Cα of 0.23 $Å^2$ over 90 Cα between the two variants. This indicates that these amino acid substitutions do not significantly alter the structure of the polypeptide backbone in the E31A and E34A variants. The position of side chains in the FOC and their metal coordination geometry are conserved for both variants, indicating that changes to the entry site do not fundamentally change the structure of the FOC.

Analysis of the entry site of the E31A variant shows that octahedral metal ion coordination seen for the wild-type protein is disrupted (Figure 4), with a weak electron density feature at approximately 4 Å distance from the carboxyl oxygens of Glu34 (Figure S4A). It was not possible to unambiguously model this feature, but it may be interpreted as a partially occupied divalent cation, such as calcium, which is present at 140 mM in the crystallization solution. Loss of the coordinating Glu31 side chain ligand causes Trp38 to move when compared to the wild-type protein; this residue is found in two predominant conformations in this site, where it moves closer to the Glu34 side chain (Supplementary Movie 1). The change in metal ion binding when compared to the wild type protein can be ascribed to the change in the coordination environment around this site. Similarly, the entry site of the E34A variant does not appear to coordinate any metal ions and the side chain of Glu31 is flipped away by 120° from this site (Figure 4, Figure S4B), and the Trp38 indole ring is also moved away from this site when compared to EncFtn-WT (Supplementary Movie 2). Together, the local structural changes seen for both entry site variants disrupt the metal coordination in this site when compared to the wild-type protein.

Native mass spectrometry reveals conformational heterogeneity in the EncFtn variants.

To interrogate the stability and oligomeric states of the EncFtn variants, we performed native MS and ion mobility MS experiments on samples purified by size-exclusion chromatography (Figure 5) (14). In our initial studies we found that EncFtn-W38A was not amenable to native mass spectrometry analysis due to poor solubility in native MS ammonium acetate buffer; consequently, native MS was performed on EncFtn-WT, EncFtn-E31A, EncFtn-E34A, and EncFtn-W38G.

In native MS analysis, EncFtn-E31A behaves similarly to EncFtn-WT and presents solely as a decameric charge state distribution centered around the 25+ charge state (Figure 5, pink triangles). In contrast, under the same instrument conditions, EncFtn-E34A and EncFtn-W38G displayed both decameric and monomeric charge state distributions (Figure 5, highlighted with pink triangles and blue circles, respectively). The presence of monomeric high charge states (8+ to 16+ for EncFtn-E34A, and 8+ to 15+ for EncFtn-W38G) suggests a degree of gas phase dissociation of 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 complexes of the EncFtn-E34A and EncFtn-W38G are less stable than EncFtn-WT. Thus, these initial observations suggest that loss of either E34 or W38 results in gas phase destabilization of the decamer assembly, although we cannot rule out the observed difference in stability being due to different iron content within the FOC in each of these protein preparations. We also note that the decameric charge state distributions of EncFtn-WT, EncFtn-E34A and EncFtn-W38G are extended and include several additional higher charge, low abundance charge states (Figure 5, highlighted with *) when compared to our previous EncFtn native MS analysis of a truncated EncFtn-WT protein (15). We attribute their presence to the ability of the solvent-exposed encapsulation sequence to readily protonate in solution. The difference in gas phase stability of the EncFtn variants was also evident when increasing the sampling cone voltage and/or activating the trap voltage of the mass spectrometer, which activates the complex and causes dissociation (Figure S5).

To further investigate the gas phase structure and stability of the variants, ion mobility collision induced unfolding (CIU) experiments were performed (43, 44). Each gas phase decamer protein complex was activated in a stepwise manner by incremental changes of the trap voltage prior to analysis of the conformation of the native protein ions using ion mobility mass spectrometry. The resulting ion mobility profile (arrival time distributions) 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 S6 A-D). It is clear from these experiments that two major discrete conformations of each EncFtn variant are observed during CIU; a conformation with a compact structure and a second conformation with a more extended, unfolded structure (Figure S6 A-D labelled ‘C’ and ‘E’ respectively, and Figure S6 E i and ii). 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 (characterized by a drift time of 12.5 ms). A second more complex transition to higher drift time is also observed at high activation (around 45V) before dissociation of the complex to release monomeric subunits. The transition from the C to E states for each variant was quantified using CIU50 values (the collision voltage required to convert 50 percent of the compact protein form into the adjacent extended state) to give values of 35.2 V, 26.5 V, and 35.1 V for EncFtn-WT, EncFtn-E31A and EncFtn-E34A respectively. The close similarity of these CIU profiles strongly suggests similar gas phase structures for the EncFtn-WT and the E31A and E34A variants (Figure S6) with a slight destabilizing effect resulting from the mutation of E31 residue. In contrast, EncFtn-W38G occurs in the more extended conformation (12.5 ms) throughout the CIU experiment (Figure S6 D and Figure S6 E iii), with the compact conformation only observed as a minor species at low activations. As there is already less than 50 percent of compact conformation present in the lowest activating conditions, no CIU50 can be established for the transition from compact to extended state for W38G. EncFtn-W38G also undergoes a more complex transition from the extended conformation but with a lower CIU50 of 35.2 V compared to the other variants of approximately 45 V. This is in accord with the native MS studies and suggests that although the W38G assembly shares similar overall structure to the WT and glutamate variants, the tryptophan substitution has a significant destabilizing effect on the decameric assembly. This increased structural/conformational heterogeneity in the tryptophan variant of EncFtn may account for our inability to obtain crystals in standard crystal screens.

The EncFtn metal entry site acts to attract metals to the FOC and restrict its catalytic activity

The changes in metal ion coordination observed in the entry sites in the crystal structures of the E31A and E34A variants suggest that these residues play a direct role in controlling the accessibility of the ferroxidase center to metal ions. To test our hypotheses about the functional role of this site (Figure 2 D,E,F), we investigated the ferroxidase activity of all of the EncFtn variants produced, both in isolation, and in complex with the encapsulin protein.

The ferroxidase activity of the wild-type EncFtn protein was confirmed to be comparable to that previously determined for the EncFtn-strepII variant (14). Surprisingly, all three variants displayed higher catalytic activity than the wild-type protein (Figure 6). The EncFtn-E34A variant had a 2-fold higher activity, while both the W38A and E31A mutations oxidize Fe(II) to Fe(III) around 5-fold faster, and their progress curves present a distinct shape in contrast to the wild-type protein (Figure 6). In order to test the contribution of Trp38 to the EncFtn catalytic activity we tested the W38A variant, allowing for a direct comparison with the other variants. We also verified that the W38G and W38A variants had comparable activities in complex with the Encapsulin protein (Figure S2C).

The increase in activity observed by all of these variants suggest that the entry site plays a role in restricting the flow of substrate Fe(II) ions to the FOC. When the differences in enzymatic activity are considered in the context of the X-ray crystal structures, it is clear that the disruption of metal coordination in the entry site has a significant impact on the catalytic activity of the EncFtn protein. The EncFtn-E31A structure shows that, as a consequence of changes in metal ion coordination, the restriction in metal ion flow provided by this variant is partially removed when compared to EncFtn-WT. However, the residual glutamic acid, Glu34, may still be able to coordinate metal ions, albeit with lower affinity than the wild-type enzyme, given the increased distance between these residues and lack of additional coordination from Glu31 (Figure 4).

The crystal structure of the EncFtn-E34A variant shows that the mutation of Glu34 to Ala results in a reorganization of the entry site, with the side chain of Glu31 reoriented in such a way that it is no longer correctly positioned for metal coordination, essentially removing any electrostatic attraction and flow restriction for this site (Figure 4). A ferroxidase assay carried out with an EncFtn-E31A/E34A variant shows that its activity is comparable to EncFtn-E34A (Figure S7). These data corroborate our hypothesis that in EncFtn-E34A, the flow restricting function is completely abolished, and no attraction is conveyed by the residual Glu31. As no flow restriction is provided at this level, we would expect to observe a higher activity, however the restructuring of this site results in a loss of attractive forces, and consequently only a modest increase in activity is observed when compared to EncFtn-WT. The difference in activity between EncFtn-E31A and EncFtn-E34A demonstrates that the electrostatic attraction provided by Glu34 plays an important role in the attraction of metal ions to the FOC.

As we were unable to produce crystals of the W38A variant, we could not determine the effect of removing this tryptophan residue on the structure of the entry site. Analysis of our crystal structures shows that the indole ring of the tryptophan may act to sterically constrain the Glu31 side chain in the correct position for metal-coordination. Consequently, it is possible that the removal of Trp38 results in the Glu31 side chain moving away from this site. Our ferroxidase assay data show that the Trp38 variants have a comparable activity profiles to the E31A variant; we therefore propose that the two variants possess similar entry site structures. The broader destabilizing 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 EncFtn variants (Figure 6B). The deviations from wild type behavior were compared to those observed in the absence of the encapsulin nanocompartment (Figure 6A). 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-WT or Enc:EncFtn-WT) (Figure S8, Tables S3, S4 and S5). This allowed comparison of the assay data across the two contexts for the EncFtn enzyme (encapsulated vs non-encapsulated). Our data show that the E31A and W38A variants have a greater effect on activity within both systems (about 4-fold faster when compared to wild type initial rate) while the E34A variant displays a 2-fold increase in the initial velocity v_0 (Figure S8, Table S5). Therefore, the EncFtn variants and wild type retain the same relative differences in their activities whether they are associated to the encapsulin shell, or not. This observation suggests that the attraction and flow restricting properties are exerted by the entry site and are not changed by the sequestration of the EncFtn within the encapsulin cage.

The ferroxidase activity of EncFtn is inhibited by zinc

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 (45). 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 7A). The selected zinc concentration was previously determined to be the IC_{50} for wild-type EncFtn-strepII (14). This concentration was chosen to allow identification of the impact of inhibition across the set of variant proteins. In line with our previous observations, EncFtn-WT showed 54% inhibition (14) (Figure 7A and B). The activity of the three variant proteins (E31, E34, W38) showed a markedly increased relative inhibition when compared to their activity in the absence of zinc (88%, 66%, and 69%, respectively) (Figures 7A and B). The increased inhibition seen in the variants can be explained by the changes in metal coordination at the entry site. With all the variants, the abrogation of metal coordination at this site permits greater access to the FOC for competing metal ions, as they are no longer effectively captured in the entry site. While these data show that the entry site restricts the general accessibility of the FOC to competing metal ions, they do not directly address the question of whether the site is acting as a selectivity filter.

Steady state fluorescence emission reveals zinc ion binding at the entry site

To determine if zinc ions interact with the entry site and thus compete with iron for access to the FOC, intrinsic tryptophan fluorescence was exploited to study the Trp38 microenvironment in the presence and absence of zinc sulfate. In addition to Trp38, EncFtn possesses two additional tryptophan residues (W72 and W80), one phenylalanine (F89), and two tyrosine residues (Y39 and Y87). Tyr39 is located in close proximity to the FOC and Glu31/Glu34, while the other aromatic residues are > 10 Å from both metal-coordination sites. Therefore, any observed change in the protein fluorescence emission should be due to changes in the entry site, and hence to metal binding at this site. This hypothesis was tested by monitoring steady state fluorescence emission of EncFtn proteins over time and upon addition of increasing sub-stoichiometric concentrations of Zn(II). The control experiment with EncFtn-W38A showed an hypsochromic shift of 8 nm in the emission maximum (λ_{EM} ~ 322 nm compared to λ_{EM} ~ 330 nm of EncFtn-WT) (Figure S9A), suggesting a contribution to intrinsic fluorescence from more buried residues in the absence of Trp38. Although this variant still possesses both glutamic acid residues in the entry site, no specific trend was observed upon zinc addition (Figure 8).

Only wild-type EncFtn and the E31A variant showed quenching of their fluorescence emission on addition of zinc, with an inflection

at ~1 molar equivalent of Zn(II) per entry site (Figure 8, Figure S9B,C). Variants with abrogated metal coordination (E34A and E31A/E34A) did not show any specific trend, or inflection upon zinc addition (Figure 8, Figure S9F). Moreover, tryptophan fluorescence is highly influenced by the solvent polarity of the surrounding environment and fluorescence emission spectra from both E34A and E31A/E34A variants exhibited a bathochromic shift of 25 nm (λ_{EM} ~ 355 nm), implying an increasing polarity in Trp38 surroundings (Figure S9A). These observations suggest that upon metal coordination, the Trp38 indole ring is restrained towards the gateway of the channel leading to FOC.

Thus, we propose that metal ions can indeed interact with the entry site when it is intact (wild type), or partially present (E31A). In the absence of the entry site, Zn(II) will likely reach the FOC, an event that is not detectable with this assay, but shown by ferroxidase analysis in the presence of Zn(II) where increased inhibition of all variants was observed.

Zn(II) displaces Fe(II) in the FOC and stabilizes the FOC dimer interface

Given the observation in the intrinsic tryptophan fluorescence experiments that zinc directly interacts with the entry site, we performed high-resolution native FT-ICR MS to further probe whether this site was acting as a selectivity filter to restrict the passage of competing ions to the FOC. The high mass resolution of FT-ICR MS affords isotopic resolution of native protein complexes and, when applied to the study of metalloproteins, allows assignment of coordinated metals with high confidence. To further aid this analysis, we also employed an isotope depletion strategy (46). The depletion of ^{13}C and ^{15}N stable isotopes from the EncFtn protein sample (referred to as *ID*-EncFtn) results in a dramatically reduced isotope distribution in mass spectrometry analysis. This simplified signal results in increased sensitivity and, most importantly for native MS analysis, reduces the occurrence of overlapping signals from proteoforms of similar mass, such as complete separation of sodiated and potassiated protein ions (a common observation in native protein MS). Thus, this strategy simplifies data interpretation and allows more confident and precise assignment of metal bound species. Native FT-ICR analysis of apo-*ID*-EncFtn reveals that it exists exclusively as a monomer (Figure 9A), as previously seen for EncFtn expressed in minimal media (15). Isotopic resolution confirms there is no metal bound within the purified *ID*-EncFtn monomer; which is as predicted, due to the limited bioavailability of metals in the minimal media required for the production of isotopically depleted proteins (Figure S10 A). Titration with Fe(II) prior to MS analysis results in the disappearance of the monomeric species and the appearance of a decameric charge state distribution (Figure 9B, pink triangles). This result is consistent with the Fe(II)-dependent assembly of EncFtn (15). Interestingly, titration with Zn(II) prior to native MS reveals a similar result (Figure 9C, pink triangles), supporting our previous solution-phase observations that metal-mediated higher order assembly is not specific for iron (15). Loading *ID*-EncFtn with Fe(II) followed by the addition of Zn(II) results in a decameric charge distribution similar to that of the Zn(II)-mediated decamer (Figure 9D). Isotopic resolution cannot be achieved for the decameric species due to heterogenous metal loading, but sufficient mass resolving power indicates metal loading occurs to the decameric complex in each case (Figure S10 B). We used collision-induced dissociation (CID) to perform gas phase disruption experiments to gain additional understanding of metal loading in *ID*-EncFtn. Iron-loaded *ID*-EncFtn dissociates predominantly into a high-charge monomer species, the so-called 'typical' dissociation product of protein complexes (47). However, upon closer inspection, evidence of higher order collision induced subcomplexes is present - a minor tetramer species is observed at *m/z* 3552 (charge 14+) (Figure S11 A i, blue circles and purple diamond respectively). This species has low signal-to-noise ratio (S/N); however, its depleted isotope distribution is similar to that

of four *ID*-EncFtn monomers with two Fe(III) ions bound (Figure S11 A ii). We assign this species as two non-FOC dimers bound together by a di-iron containing FOC, this dissociation pathway is consistent with the proposed assembly of EncFtn (24) (Figure S11 A iii).

In stark contrast, *ID*-EncFtn with coordinated Zn(II) ions dissociates into monomer and dimer species (Figure S11 B i, blue circles and green squares respectively). Isotopic resolution was achieved for the dimer species and reveals this species has a mass consistent with two *ID*-EncFtn monomers associated two Zn(II) ions (Figure S11 B ii). The presence of two metal ions bound within the dimer suggests that this species is the FOC dimer with the FOC occupied by two Zn(II) ions. There is no peak observed consistent with an apo-dimer, nor a dimer with one, or three, Zn(II) ions bound. This 'atypical' dissociation pathway was previously observed for EncFtn (15) (Figure S11 B iii). The differing dissociation pathways of zinc-coordinated *ID*-EncFtn and iron-coordinated *ID*-EncFtn suggests that zinc stabilizes the FOC interface and binds with a higher affinity than iron in the FOC catalytic center; this higher affinity is as predicted from the Irving-Williams series. This supports our observations from the ferroxidase assay that show zinc inhibiting the FOC. Analysis of *ID*-EncFtn first loaded with iron and then subsequently challenged with zinc reveals an identical dissociation pathway to that titrated solely with zinc (Figure S11 C i). Analysis of the isotopic distribution of the dimer species shows the same dimer species seen when titrated with zinc alone (Figure S11 C ii). No iron containing dissociated subcomplexes were observed in the mass spectrum, suggesting that in EncFtn-WT, zinc can efficiently displace iron in the FOC. Thus, taken together, our native MS and biochemical analysis suggests that zinc inhibition acts at the catalytic ferroxidase center, and that zinc is not held within, or discriminated against, at the entry site (Figure S11 C iii).

Discussion

We have shown that substitution of residues in the entry site triad (Glu31, Glu34, Trp38) with non-metal binding residues results in the enhancement of enzymatic activity (Figure 6A). Our crystal structures support the conclusion that this is a consequence of the creation of a wider and more loosely coordinating entry site, which allows the faster passage of metal ions to the FOC due to a reduced energy barrier at this site (Figure 4). Furthermore, when these key residues are removed, the protein is more susceptible to inhibition by Zn(II).

Solution and gas-phase analyses of variant proteins showed that, in contrast to the FOC, removal of the metal coordinating ligands (E31A/E34A) did not have a large effect on protein oligomerization; however, removal of the Trp38 side chain had a significant destabilizing effect on the quaternary structure of the protein. Investigation of the EncFtn intrinsic fluorescence emission properties upon zinc binding revealed coordination in the entry site when this is intact or partially formed (WT and E31A, respectively), confirming that this site acts as an additional metal binding site prior to the FOC.

High resolution MS analysis of *ID*-EncFtn metal binding reveals that *ID*-EncFtn decamer assembly can be induced by iron and zinc ions. Furthermore, zinc appears not to be trapped by the metal ion entry site but is bound within the FOC more tightly than iron and stabilizes the FOC dimer. When *ID*-EncFtn is loaded with iron and subsequently challenged with zinc, only zinc is observed bound in *ID*-EncFtn subcomplexes, indicating that zinc displaces iron from the FOC.

Taken together, these results imply that the entry site has a lower affinity for metal ions than the ferroxidase center and that in accord with the Irving-Williams series the FOC forms more stable complexes with zinc than iron. However, bacterial cells buffer metal concentrations via a polydisperse buffer in the order of the Irving-William series (48–51), resulting in highly competitive Zn(II) ions being buffered to lower available levels than

Fe(II), thus presumably preventing the mis-metalation of ferritin sites *in vivo*.

These observations support the hypothesis that the entry site has a dual function, in attracting metal ions to the catalytic site, and as a flow restrictor to limit the movement of metal ions to the FOC and thus slow the rate of iron oxidation (Figure 2 D and F). Our results confirm that the ferroxidase center is indeed the active site of this enzyme, as the loss of the 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 substrate turnover and the availability of a sink for the electrons liberated from the iron in the ferroxidase center. The catalytic mechanism and electron sink for ferritins is still a subject of vigorous debate. The EncFtn family presents a unique structural arrangement of the active site, which will no doubt present a distinct mechanism to other members of the ferritin family when probed further (52). The entry site may well provide a brake to ensure productive coupling of iron oxidation and mineralization.

Encapsulation of the EncFtn protein within the encapsulin nanocage enhances the activity of all of the EncFtn variants (Figure 6), but does not change the way in which modifications to the entry site affect the catalytic activity, indicating that the function of the entry site is not changed through association with the encapsulin shell. These results suggest that the encapsulin shell acts to guide productive iron ion interactions with the EncFtn protein; indeed, our previous work showed that the encapsulin shell interacts with a significant amount of iron in the absence of the EncFtn protein. This warrants future investigation into the ability of the encapsulin shell to conduct metal ions to the EncFtn active site and the role of the encapsulin shell pores in the access of substrates to the encapsulated enzyme.

Experimental procedures

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 (15).

Site-directed mutagenesis to generate the desired variants of EncFtn and Enc:EncFtn was conducted via the QuikChange® protocol (Stratagene) using primers listed in Table S6. pACYCDuet-1-EncFtn (15) or pACYCDuet-1-Enc:EncFtn (15) were used as template to produce single mutants, while pACYCDuet-1-EncFtn-E31A was used to produce EncFtn-E31A/E34A double mutant.

Protein nomenclature used throughout the text is shown in Table S7. Primers used in this work for this purpose are listed in Table S6.

Protein production and purification

Expression plasmids for the *R. rubrum* EncFtn protein and Enc:EncFtn complex were transformed into competent *Escherichia coli* BL21(DE3) cells and grown in LB medium supplemented with appropriate antibiotics, and expressed as described previously (14, 15). Protein sequences are listed in Table S8.

EncFtn and its variants, were expressed and purified as described in (14). A single colony of *E. coli* BL21(DE3) cells transformed with protein expression plasmid was grown overnight at 37 °C in 10 ml LB medium, supplemented with 35 µg/ml chloramphenicol or 50 µg/ml kanamycin, while shaking at 180 rpm.
The starting culture was then inoculated into 1 l of LB medium, grown until OD_{600} = 0.5, and protein expression was induced with 1 mM IPTG, while the temperature was reduced to 30 °C and the cells were incubated for a further 18 h. Cells were pelleted by centrifugation at 4000 *g* for 25 min at 4 °C and washed with PBS before an additional centrifugation step.

EncFtn wild-type and variants cell pellets were resuspended in 10-times (v/w) of Buffer QA (50 mM Tris, pH 8.0) and lysed by sonication on ice (5 minutes of 30-second burst of sonication at 35 % amplitude and 30-second cool down). Cell lysate was clarified by ultracentrifugation at 16,000 *g* for 25 min at 4

°C, followed by filtration using a 0.22 µM syringe filter (Millipore, UK). EncFtn protein sample was applied to a 20 ml anion exchange column (Hi-Prep Q HP, GE Healthcare) pre-equilibrated with Buffer QA. Un-bound protein was then washed off with 10 volumes of Buffer QA. Proteins were eluted with a linear gradient of 20 column volumes of Buffers QA and QB (50 mM Tris, pH 8.0, 1 M NaCl) to 100 % QB. Protein of interest was identified by SDS-PAGE and concentrated using a Vivaspin centrifugal concentrator (MWCO 10 kDa, Sartorius) following the manufacturer's instructions. Concentrated sample was loaded onto a size exclusion column (HiLoad 16/60 Superdex 200, GE Healthcare) pre-equilibrated with Buffer GF (50 mM Tris, pH 8.0, 150 mM NaCl). Fractions were analyzed by 15 % SDS-PAGE to assess purity and oligomerization state and those containing the protein of interest were pooled and further concentrated as described before.

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). Cells were sonicated on ice (5 minutes at 30 seconds on/off cycles, 35 % amplitude), followed by incubation for 30 minutes at 37 °C in a water bath. Cell lysate was clarified by centrifugation (35,000 x g, 30 minutes, 4 °C) and the supernatant was collected and incubated for 10 minutes at 80 °C in a water bath. 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 µm syringe filter (Millipore, UK). Enc:EncFtn protein complex sample was loaded onto a 20 ml anion exchange column (Hi-Prep Q HP, GE Healthcare) pre-equilibrated with Buffer QA (50 mM Tris, pH 8.0). Un-bound protein was washed with Buffer QA. A linear gradient elution of 20 column volumes was performed by mixing Buffers QA and QB (50 mM Tris, pH 8.0, 1 M NaCl) to 100 % QB. Since Enc:EncFtn does not bind the anion exchange column, proteins of interest were found in flow-through sample, as confirmed by 15 % SDS-PAGE. Protein was concentrated using a Vivaspin centrifugal concentrator and then applied to a size exclusion column (HiLoad 16/60 Superdex 200, GE Healthcare) pre-equilibrated with Buffer GF (50 mM Tris, pH 8.0, 150 mM NaCl). Fractions of interest were run on a 15% SDS-PAGE, and further concentrated using a Vivaspin centrifugal concentrator as before. After concentration, protein samples were analyzed by 10% Tricine-SDS-PAGE (53) in order to resolve bands around 30 kDa. Elution volumes of EncFtn and Enc:EncFtn complexes from the size exclusion column are listed in Table S2.

His-tagged protein production and purification

Hexahistidine tagged EncFtn-E31A and -E34A variants (EncFtn-sH) were produced in *E. coli* BL21(DE3) as described above and purified following the same protocols as previously described (15). A 5 ml HisTrap column (GE Healthcare, UK) was equilibrated with Buffer HisA (50 mM Tris, 500 ml NaCl and 50 mM imidazole, pH 8.0) and loaded with the clarified cell lysate. Unbound proteins were washed with Buffer HisA. A step gradient of 50 % and 100 % HisB (50 mM Tris, 500 ml NaCl and 500 mM imidazole, pH 8.0) was performed to elute His-tagged proteins. Fractions containing the protein of interest were identified by 15 % SDS-PAGE, pooled together and concentrated by Vivaspin centrifugal concentrator (10 kDa MWCO). Protein sample was applied to a size exclusion column (HiLoad 16/60 Superdex 200, GE Healthcare) pre-equilibrated with Buffer GF (50 mM Tris, pH 8.0, 150 mM NaCl). A 15 % SDS-PAGE was carried out to assess purity and oligomerization state and samples containing the protein were further concentrated as described before.

Isotopically depleted strepII-tagged protein production and purification

To produce isotopically depleted EncFtn proteins, competent *E. coli* BL21 (DE3) cells were transformed with the required expression plasmid. A single colony was used to inoculate 10 ml of LB media supplemented with the appropriate antibiotic, before overnight

incubation at 37 ºC. This was then used to inoculate 500 ml of 2x YT media, incubated at 37 ºC until an OD_{600} 0.6-0.8 was obtained, at which point the cell culture was centrifuged at 5000 x g at 4 ºC for 20 minutes (200 ml per expression culture required). The pellet was washed with M9 salts solution (33.9 g/L Na_2HPO_4, 15 g/L KH_2PO_4, 2.5 g/L NaCl) and suspended in 100 ml M9 minimal media. Isotopically depleted M9 minimal media was supplemented with ^{12}C (99.9%)–glucose (Cambridge Isotope Laboratories) and ^{14}N (99.99%)-ammonium sulfate. The M9 minimal media cultures were further incubated at 37 ºC for 1 hour, before protein expression was induced and stored as previously described (46).

Isotopically depleted strepII-tagged EncFtn (*ID*-EncFtn) was purified as reported in (He et al. 2019). Cells were suspended in 10x (v/w) Buffer W (100 mM Tris pH 8.0, 150 mM NaCl), sonicated and clarified by centrifugation. Cell lysate was loaded onto a Strep-Trap HP column (5 mL, GE Healthcare), equilibrated according to manufacturer's instructions, and unbound protein was washed out with Buffer W (5 column volumes). *ID*-EncFtn was eluted by Buffer E (100 mM Tris pH 8.0, 150 mM NaCl, 2.5 mM desthiobiotin).

Protein quantification

Both untagged and His-tagged EncFtn proteins were quantified using the Pierce BCA Assay kit (Thermo Fisher Scientific) following the manufacturer's instructions for the Test-tube procedure. Concentration of Enc:EncFtn proteins was determined using the Beer-Lambert equation, measuring absorbance at 280 nm with a Nanodrop Spectrophotometer (Thermo Scientific) and using an extinction coefficient calculated by ProtParam tool on the Expasy platform entering a sequence composed of 3 x Enc and 2 x Enc-Ftn sequences. This ratio was determined by examination of SDS-PAGE gels of Enc:EncFtn proteins and densitometry of bands.

Transmission electron microscopy

TEM imaging was performed as previously described (15) on purified Enc:EncFtn proteins. Proteins were diluted in Buffer GF to a final concentration of 0.1 mg/ml before being spotted on glow-discharged 300 mesh carbon-coated copper grids. Excess liquid was removed by blotting with filter paper (Whatman, UK) before the sample was washed three times with distilled water prior to being stained with 0.2 % uranyl acetate, blotted again, and air-dried. Grids were imaged using a Hitachi HT7800 transmission electron microscope and images were collected with an EMSIS Xarosa camera.

Protein crystallization and X-ray data collection

EncFtn-sH variants (E31A or E34A) were concentrated to 10 mg/ml (based on extinction coefficient calculation) and subjected to crystallization under similar conditions to the wild-type protein (WT) (Table S9). Crystallization drops were set up in 24−well Linbro plates using the hanging drop vapor diffusion method at 292 K. Glass coverslips were set up with 1–2 µl protein mixed with 1 µl well solution and sealed over 1 ml of well solution. Crystals appeared after one week to two months and were mounted using a LithoLoop (Molecular Dimensions Limited, UK), transferred briefly to a cryoprotection solution containing well solution supplemented with 1 mM $FeSO_4$ (in 0.1% (v/v) HCl) and 20% (v/v) PEG 200, and were subsequently flash cooled in liquid nitrogen.

All crystallographic datasets were collected at 10–100 eV above the experimentally determined Fe-*K*α edge on the macromolecular crystallography beamlines at Diamond Light Source (Didcot, UK) at 100 K using Pilatus 6M detectors. Diffraction data were integrated and scaled using XDS (54) and symmetry related reflections were merged with Aimless (55). Data collection statistics are shown in Table 1. The resolution cut-off used for structure determination and refinement was determined based on the $CC_{1/2}$ criterion proposed by (56).

The structures of the EncFtn-sH variants were determined by molecular replacement using decameric wild-type protein (PDB ID: 5DA5) as the search model (15). A single solution comprising three decamers in the asymmetric unit was found by molecular replacement using Phaser (57). The initial model was rebuilt using Phenix.autobuild (58) followed by cycles of refinement with Phenix.refine (59), with manual rebuilding and model inspection in Coot (60). The final model was refined with isotropic B-factors, torsional NCS restraints, and with anomalous group refinement to correctly model the iron ions in the ferroxidase center. The model was validated using MolProbity (61). Structural superimpositions were calculated using Coot (60). Crystallographic figures were generated with PyMOL (Schrödinger, LLC.). Model refinement statistics are shown in Table 1. The final models and experimental data are deposited in the PDB.

Mass spectrometry analysis

LC-MS, native MS, ion mobility MS (IM-MS) and collision induced unfolding experiments were performed on a Synapt G2 ion-mobility equipped Q-ToF instrument (Waters Corp., Manchester, UK). For LC-MS analysis, an Acquity UPLC equipped with a reverse phase C4 Aeris Widepore 50 × 2.1 mm HPLC column (Phenomenex, CA, USA) was used with a gradient of 5–95%B (where A:water + 0.1% formic acid, and B:acetonitrile + 0.1% formic acid) over 10 minutes. For LC-MS, samples were typically analyzed at 5 μM, and data analysis was performed using MassLynx v4.1 and MaxEnt deconvolution.

For native MS analysis, all protein samples were buffer exchanged into ammonium acetate (100 mM; pH 8.0) using Micro Biospin Chromatography Columns (Bio-Rad, UK) prior to analysis and the resulting protein samples were analyzed at a typical final concentration of ~5 μM (oligomer concentration). Nanoelectrospray ionization was employed using a Nanomate infusion robot (Advion Biosciences, Ithaca, NY). Instrument parameters were tuned to preserve non-covalent protein complexes and were consistent for the analysis of the EncFtn variants. After native MS optimization, parameters were: nanoelectrospray voltage 1.55 kV; sample cone 100 V; extractor cone 0 V; trap collision voltage 4 V; source temperature 60 °C; and source backing pressure 4.5 mbar. IM-MS was performed using the travelling-wave mobility cell of the Synapt G2 with nitrogen drift gas. The typical conditions used for IM-MS were: wave velocity 300 m/s; wave height 19 V; and sampling cone of 75 V. For collision induced unfolding (CIU) experiments, samples were analyzed with incremental increasing trap voltages between 10 V to 60 V, using 2 V increments. Data was processed and CIU50 values determined using CIUSuite (43) and CIUSuite 2 (44).

High resolution native mass spectrometry was performed on a 12T SolariX 2XR FT-ICR MS (Bruker Daltonics) equipped with a nanoelectrospray source. Due to the quadrupole range of the FT-ICR MS instrument used, a truncated and strep(II)-tagged version of EncFtn-WT was used for these experiments. Protein samples were buffer exchanged into ammonium acetate (100 mM; pH 8.0) prior to direct infusion. Source conditions and ion optics were optimized to transmit native proteins ions and when required, Continual Accumulation of Selected Ions (CASI) was employed to isolate charge states of interest. Typically, 2 Megaword data was collected in QPD (2ω) mode to produce a 6 second FID, which resulted in a typical mass resolving power of ca. 300,000. The resulting data was processed using Data Analysis (Bruker Daltonics). For metal titrations, fresh iron(II) chloride tetrahydrate or zinc acetate in 0.1% (v/v) HCl was added to *ID*-EncFtn (1:1 metal:protein concentration) prior to buffer exchange into ammonium acetate (100 mM; pH 8.0).

Ferroxidase assay of EncFtn proteins

Enzymatic activity of EncFtn proteins was tested by ferroxidase assay, as previously described (14). Fe(II) aliquots were prepared under anaerobic conditions by dissolving $FeSO_4 \cdot 7H_2O$ in 0.1% (v/v) HCl. Purified protein was diluted anaerobically in Buffer H (10 mM HEPES, pH 8.0, 150 mM NaCl),

previously purged with gaseous nitrogen, to a final concentration of 10 µM (EncFtn).
Protein and ferrous iron samples were added to a quartz cuvette (Hellma) under aerobic conditions at a final concentration of 10 µM and 50 µM, respectively, corresponding to a FOC:Fe(II) ratio of 1:10.

Absorbance at 315 nm was recorded every second for 1500 s by a UV-visible spectrophotometer (PerkinElmer Lambda 35), using the TimeDrive software. The same experiment was performed in the absence of the enzyme to determine the oxidation of ferrous salts by atmospheric oxygen.
The same setup was adopted for recording the activity of EncFtn-E34A and EncFtn-E31A/E34A in a separate assay at higher concentration (20 µM EncFtn and 100 µM Fe(II)) for comparison purposes. Data presented here are the mean of three technical replicates of time zero-subtracted progress curves with standard deviations calculated from the mean.

Calculation of the EncFtn enzymatic reaction initial rate (v_0) was made by applying the Linear Regression tool in GraphPad (Prism8) on the absorbance at 315 nm measured for the first 200 s, when curves are still linear and following an order zero kinetics (Figure S8, Table S3). Slopes obtained from these curves correspond to initial reaction rate. $v_0^{variant}/\ v_0^{wild\ type}$ factors were calculated by dividing the initial rates of the variant enzymes by the initial rate of the wild-type protein (Table S5).

Ferroxidase assay of Enc:EncFtn protein complexes
Ferroxidase assay of Enc:EncFtn protein complexes was performed in the same conditions as those used for EncFtn, but with different protein and iron salt concentration. Enc:EncFtn was diluted to a final concentration of 25 µM which corresponds to 10 µM EncFtn based on a 3:2 Enc:EncFtn ratio, as previously observed by SDS-PAGE by our group (Figure 1A in (15). Fe(II) sample was diluted to 50 µM in the reaction system to maintain a final ratio of 1:10 FOC:Fe(II).

Calculation of the Enc-EncFtn enzymatic reaction initial rate (v_0) and $v_0^{variant}/\ v_0^{wild\ type}$ factors was carried out as described in the above section (Figure S8, Tables S4 and S5).

Zinc inhibition of ferroxidase activity of EncFtn proteins
Ferroxidase assay was performed as above with EncFtn (10 µM) and Fe(II) samples (50 µM) in the presence of 34 µM $ZnSO_4{\cdot}7H_2O$. The chosen concentration corresponds to the IC_{50} value previously determined for EncFtn-StrepII, under the same experimental conditions by our group (14). Oxygen-free metal samples were added to the quartz cuvette under aerobic conditions, followed by the protein sample. Data were replicated three times and means and standard deviation of time zero-subtracted progress curves were calculated. A negative control was performed by monitoring A_{315} of Zn(II) and Fe(II) salts mixed in the absence of enzyme.

Monitoring quenching of protein intrinsic fluorescence upon metal-binding
Experiments were carried out aerobically in a quartz cuvette (Hellma), using a Cary Eclipse Fluorescence spectrophotometer. Protein was diluted in Buffer GF to a final concentration of 20 µM, corresponding to 10 µM for the entry site. $ZnSO_4{\bullet}7H_2O$ stock was prepared in deionized water. The instrument excitation wavelength was set to 280 nm, corresponding to maximum absorption for tryptophan residues.

Two distinct modes of detection were used; Kinetic option on the Analysis software package was chosen to record protein fluorescence emission signal at specific wavelengths over time. Pre-scans were carried out to find the optimal emission wavelength (ranging from 322 nm to 355 nm) for each protein variant. Metal aliquots (0.15 or 0.3 molar equivalents) were added to the cuvette while pausing the data collection for fewer than five seconds. The final data represent average of 10 data points at equilibrium upon each metal addition.

The Scan option of the Analysis software package was chosen to record emission in the 290-400 nm range, allowing detection of possible shifts in emission peak maximum upon metal addition. Spectra were recorded in triplicate at equilibrium.

Data availability

Data sets supporting this paper have been deposited in appropriate public data repositories. Please see figure legends and tables for links to these.

Acknowledgements
We would like to thank staff at the Newcastle University Electron Microscopy Research Services for assistance with TEM. We would like to thank Diamond Light Source for beamtime (proposal mx9487), and the staff of beamlines I02, I03 and I24 for assistance with crystal testing and data collection. We would like to thank Prof. Dominic Campopiano and Gregor Skeldon for their constructive advice throughout this project.

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

JR, 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.

KJG, ID-MS methods development and acquisition of data, Analysis and interpretation of data

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]. Equipment for Transmission Electron Microscopy was funded through the BBSRC 17ALERT call [BB/R013942/1]. FT-ICR instrumentation was funded by BBSRC 17ALERT [BB/R013993/1].

Conflict of interest
The authors declare that no competing interests exist.

Abbreviations
Enc, encapsulin; EncFtn, encapsulated ferritin; FOC, ferroxidase center; MS, mass spectrometry; nESI, Native nanoelectrospray ionization; CIU, collision induced unfolding; FT-ICR, Fourier-transform ion cyclotron resonance; ID, isotope depletion.

References

1. Andrews, S. C. (2010) The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochim. Biophys. Acta - Gen. Subj.* **1800**, 691–705
2. Dillard, B. D., Demick, J. M., Adams, M. W. W., and Lanzilotta, W. N. (2011) A cryo-crystallographic time course for peroxide reduction by rubrerythrin from Pyrococcus furiosus. *JBIC J. Biol. Inorg. Chem.* **16**, 949–959
3. Cardenas, J. P., Quatrini, R., and Holmes, D. S. (2016) Aerobic Lineage of the Oxidative Stress Response Protein Rubrerythrin Emerged in an Ancient Microaerobic, (Hyper)Thermophilic Environment. *Front. Microbiol.* 10.3389/fmicb.2016.01822
4. Giessen, T. W., Orlando, B. J., Verdegaal, A. A., Chambers, M. G., Gardener, J., Bell, D. C., Birrane, G., Liao, M., and Silver, P. A. (2019) Large protein organelles form a new iron sequestration system with high storage capacity. *Elife*. **8**, 1–23
5. Grant, R. A., Filman, D. J., Finkel, S. E., Kolter, R., and Hogle, J. M. (1998) The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat. Struct. Biol.* **5**, 294–303
6. Pesek, J., Büchler, R., Albrecht, R., Boland, W., and Zeth, K. (2011) Structure and mechanism of iron translocation by a Dps protein from Microbacterium arborescens. *J. Biol. Chem.* **286**, 34872–82
7. Martinez, A., and Kolter, R. (1997) Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J. Bacteriol.* **179**, 5188–94
8. Gauss, G. H., Benas, P., Wiedenheft, B., Young, M., Douglas, T., and Lawrence, C. M. (2006) Structure of the DPS-like protein from Sulfolobus solfataricus reveals a bacterioferritin-like dimetal binding site within a DPS-like dodecameric assembly. *Biochemistry*. **45**, 10815–27
9. Gauss, G. H., Reott, M. A., Rocha, E. R., Young, M. J., Douglas, T., Smith, C. J., and Lawrence, C. M. (2012) Characterization of the Bacteroides fragilis bfr gene product identifies a bacterial DPS-like protein and suggests evolutionary links in the ferritin superfamily. *J. Bacteriol.* **194**, 15–27
10. Andrews, S. C. (1998) Iron storage in bacteria. *Adv. Microb. Physiol.* **40**, 281–351
11. Pfaffen, S., Abdulqadir, R., Le Brun, N. E., and Murphy, M. E. P. (2013) Mechanism of ferrous iron binding and oxidation by ferritin from a pennate diatom. *J. Biol. Chem.* **288**, 14917–25
12. Frolow, F., Kalb, A. J., and Yariv, J. (1994) Structure of a unique twofold symmetric haem-binding site. *Nat. Struct. Biol.* **1**, 453–60
13. Harrison, P. M., and Arosio, P. (1996) The ferritins: molecular properties, iron storage function and cellular regulation. *Biochim. Biophys. Acta - Bioenerg.* **1275**, 161–203
14. He, D., Piergentili, C., Ross, J., Tarrant, E., Tuck, L. R., Mackay, C. L., McIver, Z., Waldron, K. J., Clarke, D. J., and Marles-Wright, J. (2019) Conservation of the structural and functional architecture of encapsulated ferritins in bacteria and archaea. *Biochem. J.* **476**, 975–989
15. He, D., Hughes, S., Vanden-Hehir, S., Georgiev, A., Altenbach, K., Tarrant, E., Mackay, C. L., Waldron, K. J., Clarke, D. J., and Marles-Wright, J. (2016) Structural characterization of encapsulated ferritin provides insight into iron storage in bacterial nanocompartments. *Elife*. **5**, e18972
16. Ratnayake, D. B., Wai, S. N., Shi, Y., Amako, K., Nakayama, H., and Nakayama, K. (2000) Ferritin from the obligate anaerobe Porphyromonas gingivalis: purification, gene cloning and mutant studies The GenBank accession number for the sequence reported in this paper is AB016086. *Microbiology*. **146**, 1119–1127
17. Sztukowska, M., Bugno, M., Potempa, J., Travis, J., and Kurtz Jr, D. M. (2002) Role of rubrerythrin in the oxidative stress response of Porphyromonas gingivalis. *Mol. Microbiol.*

44, 479–488
18. Almiron, M., Link, A. J., Furlong, D., and Kolter, R. (1992) A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes Dev.* **6**, 2646–2654
19. Choi, S. H., Baumler, D. J., and Kaspar, C. W. (2000) Contribution of dps to Acid Stress Tolerance and Oxidative Stress Tolerance in Escherichia coli O157:H7. *Appl. Environ. Microbiol.* **66**, 3911–3916
20. Nair, S., and Finkel, S. E. (2004) Dps Protects Cells against Multiple Stresses during Stationary Phase. *J. Bacteriol.* **186**, 4192–4198
21. Chiancone, E., and Ceci, P. (2010) The multifaceted capacity of Dps proteins to combat bacterial stress conditions: Detoxification of iron and hydrogen peroxide and DNA binding. *Biochim. Biophys. Acta - Gen. Subj.* **1800**, 798–805
22. De Martino, M., Ershov, D., van den Berg, P. J., Tans, S. J., and Meyer, A. S. (2016) Single-Cell Analysis of the Dps Response to Oxidative Stress. *J. Bacteriol.* **198**, 1662–1674
23. Chasteen, N. D., and Harrison, P. M. (1999) Mineralization in Ferritin: An Efficient Means of Iron Storage. *J. Struct. Biol.* **126**, 182–194
24. Ross, J., Lambert, T., Piergentili, C., He, D., Waldron, K. J., Mackay, C. L., Marles-Wright, J., and Clarke, D. J. (2020) Mass spectrometry reveals the assembly pathway of encapsulated ferritins and highlights a dynamic ferroxidase interface. *Chem. Commun.* **56**, 3417–3420
25. Sutter, M., Boehringer, D., Gutmann, S., Günther, S., Prangishvili, D., Loessner, M. J., Stetter, K. O., Weber-Ban, E., and Ban, N. (2008) Structural basis of enzyme encapsulation into a bacterial nanocompartment. *Nat. Struct. Mol. Biol.* **15**, 939–947
26. Xu, B., and Chasteen, N. (1991) Iron oxidation chemistry in ferritin. Increasing Fe/O2 stoichiometry during core formation. *J. Biol. Chem.* **266**, 19965–70
27. Proulx-Curry, P. M., and Chasteen, N. D. (1995) Molecular aspects of iron uptake and storage in ferritin. *Coord. Chem. Rev.* **144**, 347–368
28. Yang, X., Chen-Barrett, Y., Arosio, P., and Chasteen, N. D. (1998) Reaction Paths of Iron Oxidation and Hydrolysis in Horse Spleen and Recombinant Human Ferritins †. *Biochemistry*. **37**, 9743–9750
29. Baaghil, S., Lewin, A., Moore, G. R., and Le Brun, N. E. (2003) Core Formation in Escherichia coli Bacterioferritin Requires a Functional Ferroxidase Center †. *Biochemistry*. **42**, 14047–14056
30. Ebrahimi, K. H., Hagedoorn, P.-L., Jongejan, J. a, and Hagen, W. R. (2009) Catalysis of iron core formation in Pyrococcus furiosus ferritin. *J. Biol. Inorg. Chem.* **14**, 1265–74
31. Theil, E. C., Behera, R. K., and Tosha, T. (2013) Ferritins for chemistry and for life. *Coord. Chem. Rev.* **257**, 579–586
32. Masuda, T., Goto, F., Yoshihara, T., and 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.* 10.1074/jbc.M109.059790
33. Behera, R. K., Torres, R., Tosha, T., Bradley, J. M., Goulding, C. W., and Theil, E. C. (2015) Fe2+ substrate transport through ferritin protein cage ion channels influences enzyme activity and biomineralization. *JBIC J. Biol. Inorg. Chem.* **20**, 957–969
34. Yao, H., Wang, Y., Lovell, S., Kumar, R., Ruvinsky, A. M., Battaile, K. P., Vakser, I. A., and Rivera, M. (2012) The Structure of the BfrB–Bfd Complex Reveals Protein–Protein Interactions Enabling Iron Release from Bacterioferritin. *J. Am. Chem. Soc.* **134**, 13470–13481
35. Bellapadrona, G., Stefanini, S., Zamparelli, C., Theil, E. C., and Chiancone, E. (2009) Iron Translocation into and out of Listeria innocua Dps and Size Distribution of the Protein-enclosed Nanomineral Are Modulated by the Electrostatic Gradient at the 3-fold "Ferritin-like" Pores. *J. Biol. Chem.* **284**, 19101–19109
36. Liu, X. S., Patterson, L. D., Miller, M. J., and Theil, E. C. (2007) Peptides Selected for the Protein Nanocage Pores Change the Rate of Iron Recovery from the Ferritin Mineral. *J. Biol.*

Chem. **282**, 31821–31825
37. Trefry, A., and Harrison, P. M. (1978) Incorporation and release of inorganic phosphate in horse spleen ferritin. *Biochem. J.* **171**, 313–20
38. Treffry, A., Zhao, Z., Quail, M. A., Guest, J. R., and Harrison, P. M. (1998) How the presence of three iron binding sites affects the iron storage function of the ferritin (EcFtnA) of Escherichia coli. *FEBS Lett.* **432**, 213–218
39. Ebrahimi, K. H., Bill, E., Hagedoorn, P.-L., and Hagen, W. R. (2016) Spectroscopic evidence for the role of a site of the di-iron catalytic center of ferritins in tuning the kinetics of Fe(<scp>ii</scp>) oxidation. *Mol. Biosyst.* **12**, 3576–3588
40. Hagen, W., Hagedoorn, P.-L., and Ebrahimi, K. (2017) The working of ferritin: a crossroad of options. *Metallomics*. **9**, 595–605
41. Hempstead, P. D., Hudson, A. J., Artymiuk, P. J., Andrews, S. C., Banfield, M. J., Guest, J. R., and Harrison, P. M. (1994) Direct observation of the iron binding sites in a ferritin. *FEBS Lett.* **350**, 258–62
42. Stillman, T. J., Hempstead, P. D., Artymiuk, P. J., Andrews, S. C., Hudson, A. J., Treffry, A., Guest, J. R., and Harrison, P. M. (2001) The high-resolution X-ray crystallographic structure of the ferritin (EcFtnA) of Escherichia coli; comparison with human H ferritin (HuHF) and the structures of the Fe(3+) and Zn(2+) derivatives. *J. Mol. Biol.* **307**, 587–603
43. Eschweiler, J. D., Rabuck-Gibbons, J. N., Tian, Y., and Ruotolo, B. T. (2015) CIUSuite: A Quantitative Analysis Package for Collision Induced Unfolding Measurements of Gas-Phase Protein Ions. *Anal. Chem.* **87**, 11516–11522
44. Polasky, D. A., Dixit, S. M., Fantin, S. M., and Ruotolo, B. T. (2019) CIUSuite 2: Next-Generation Software for the Analysis of Gas-Phase Protein Unfolding Data. *Anal. Chem.* **91**, 3147–3155
45. Le Brun, N. E., Andrews, S. C., Guest, J. R., Harrison, P. M., Moore, G. R., and Thomson, A. J. (1995) Identification of the ferroxidase centre of Escherichia coli bacterioferritin. *Biochem. J.* **312**, 385–392
46. Gallagher, K. J., Palasser, M., Hughes, S., Mackay, C. L., Kilgour, D. P. A., and Clarke, D. J. (2020) Isotope Depletion Mass Spectrometry (ID-MS) for Accurate Mass Determination and Improved Top-Down Sequence Coverage of Intact Proteins. *J. Am. Soc. Mass Spectrom.* **31**, 700–710
47. Hall, Z., Hernández, H., Marsh, J. A., Teichmann, S. A., and Robinson, C. V (2013) The role of salt bridges, charge density, and subunit flexibility in determining disassembly routes of protein complexes. *Structure*. **21**, 1325–37
48. Foster, A. W., Osman, D., and Robinson, N. J. (2014) Metal Preferences and Metallation. *J. Biol. Chem.* **289**, 28095–28103
49. Tottey, S., Waldron, K. J., Firbank, S. J., Reale, B., Bessant, C., Sato, K., Cheek, T. R., Gray, J., Banfield, M. J., Dennison, C., and Robinson, N. J. (2008) Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. *Nature*. **455**, 1138–1142
50. Frausto da Silva, J., and Williams, R. (1991) The biological chemistry of the elements: the inorganic chemistry of life. *Oxford Univ. Press*. **20**, 62–63
51. Reyes-Caballero, H., Campanello, G. C., and Giedroc, D. P. (2011) Metalloregulatory proteins: Metal selectivity and allosteric switching. *Biophys. Chem.* **156**, 103–114
52. Hagen, W. R., Hagedoorn, P. L., and Honarmand Ebrahimi, K. (2017) The workings of ferritin: A crossroad of opinions. *Metallomics*. **9**, 595–605
53. Schägger, H. (2006) Tricine–SDS-PAGE. *Nat. Protoc.* **1**, 16–22
54. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 133–144
55. Evans, P. R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 282–92
56. Karplus, P. A., and Diederichs, K. (2012) Linking Crystallographic Model and Data Quality.

Science (80-.). **336**, 1030–1033
57. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
58. Adams, P. D., Afonine, P. V, Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–21
59. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367
60. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501
61. Chen, V., Arendall, W., Headd, J., Keedy, D., Immormino, R., Kapral, G., Murray, L., Richardson, J., and Richardson, D. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 12–21

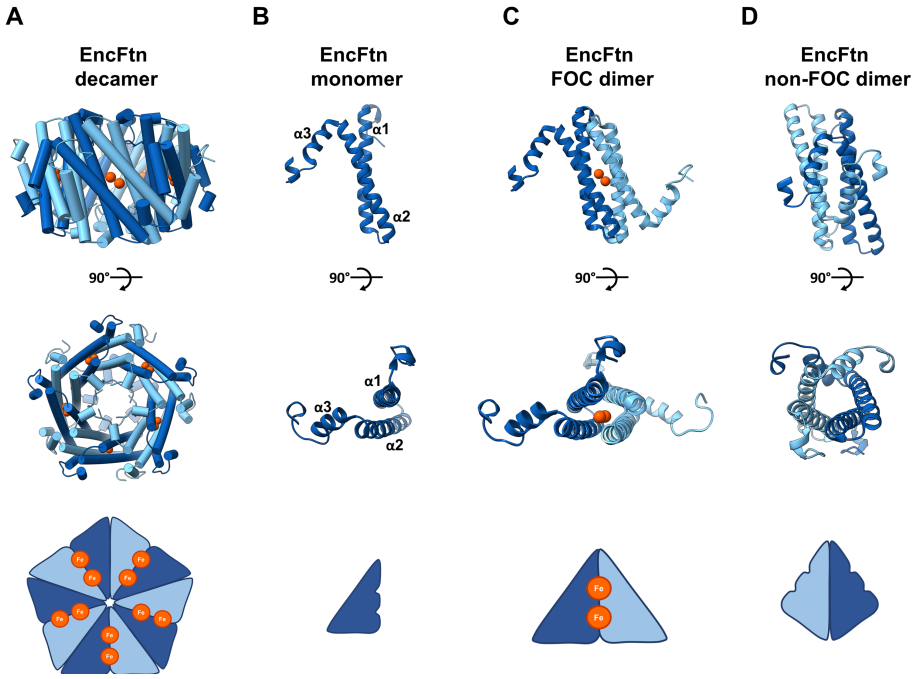

Figure 1. EncFtn subcomplex architecture

(**A**) Annular decamer structure of EncFtn shown in a side-on and top-down view, with a simplified representation of topology. The decamer structure has ten iron ions bound (orange spheres). (**B**) An EncFtn monomer with a simplified cartoon. (**C**) The FOC dimer in cartoon with two coordinated iron ions (shown as orange spheres) and a simplified FOC cartoon is shown with two iron ions as orange circles. (**D**) The non-FOC dimer with cartoon secondary structure, and a simplified representation.

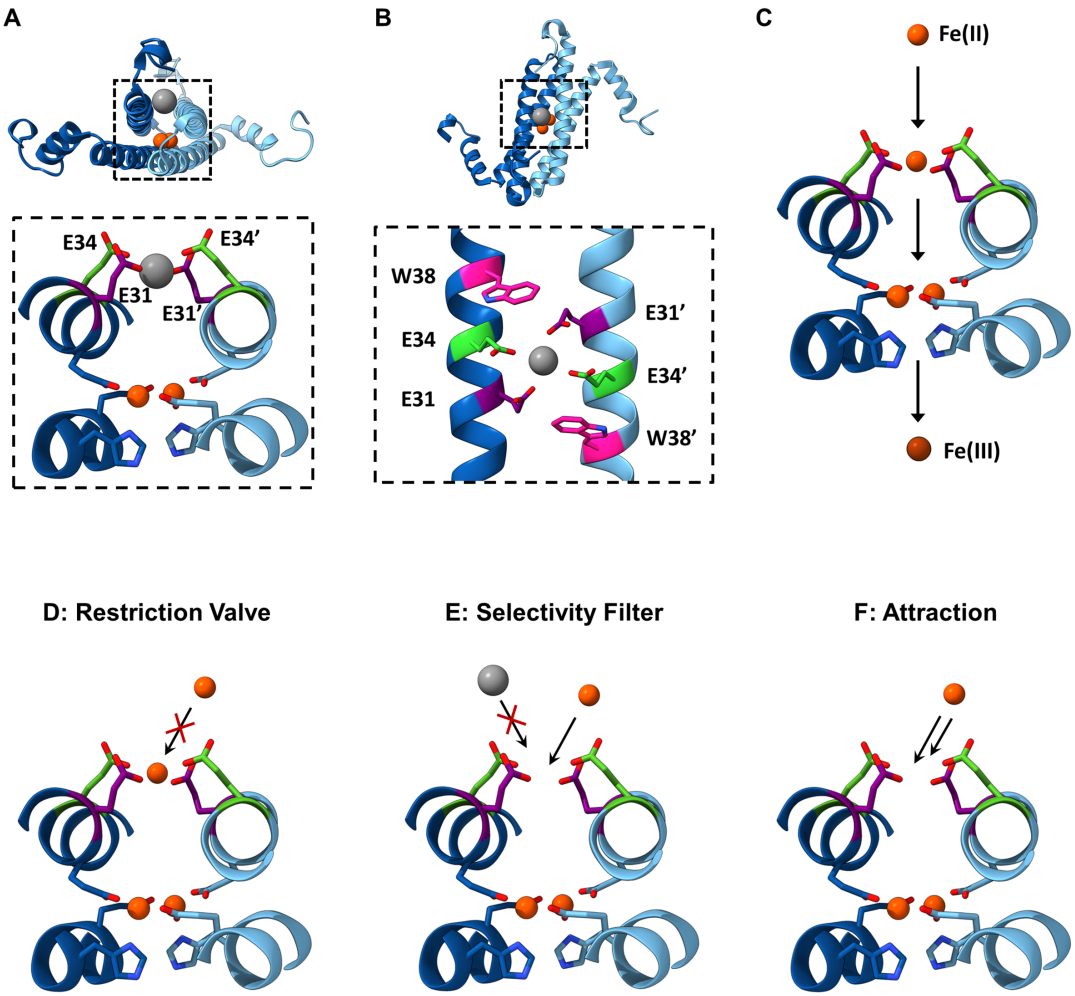

Figure 2. Hypotheses for the functional role of the entry site in EncFtn.
(**A**)Side-on view of entry site and FOC within EncFtn. The metal-binding residues of the entry site (Glue31 and Glu34 of each subunit shown in purple and green respectively) are shown coordinating a metal ion (graysphere) within the dimer interface. Iron ions coordinated in the ferroxidase center (FOC) are depicted as red spheres with FOC residues in blue. (**B**) Top-down view of the entry site showing Glu31, Glu34 and Trp38 of each subunit in purple, green and pink respectively. Glu31 and Glu34 are coordinating a divalent metal ion shown in gray. (**C**) The proposed iron pathway in EncFtn. The black arrows indicate movement of Fe(II) ions (orange spheres) in EncFtn during their oxidation process to Fe(III) (brown sphere). Fe(II) enters the entry site where the Glu31 and Glu34 side chains from both subunits are shown in purple and green. Fe(II) then proceeds to the FOC and then leaves EncFtn as Fe(III). Several hypotheses for the role played by the residues contributing to the entry site are shown in **D-F** (flow restriction valve (**D**), selectivity filter (**E**), and attraction (**F**). Ferrous and ferric ions are depicted as orange and brown circles, respectively, while non-cognate species are in gray. The entry site glutmates are shown in orange and green stick representation, and the FOC residues are shown in shades of blue.

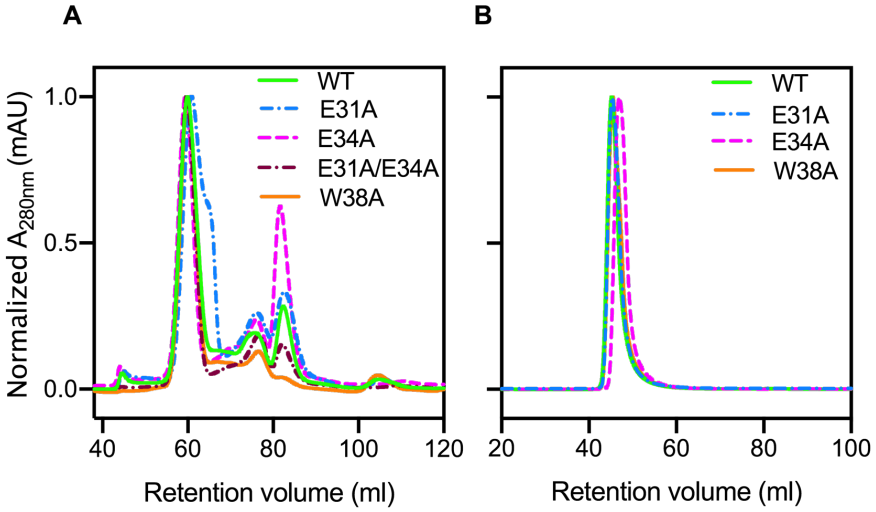

Figure 3. Purification of recombinant EncFtn and Enc:EncFtn protein complexes. Recombinant EncFtn (**A**) and Enc:EncFtn **(B)** proteins were purified by anion exchange chromatography (Hi-Trap Q-sepharose FF, GE Healthcare) and then subjected to size-exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare) previously equilibrated with 50 mM Tris-HCl, pH 8.0, 150 mM NaCl. The elution profiles in **A** reveal that all EncFtn proteins show a main peak at around 60 ml, diagnostic of oligomerization states close to 10-mer, and smaller peaks at ~ 76 ml and 82 ml (indicative of smaller assembly states, Table S2). Enc:EncFtn protein complexes elute in a single peak around 46 ml, suggesting that EncFtn is compartmentalized within the Encapsulin shell. doi. 10.6084/m9.figshare.9885557

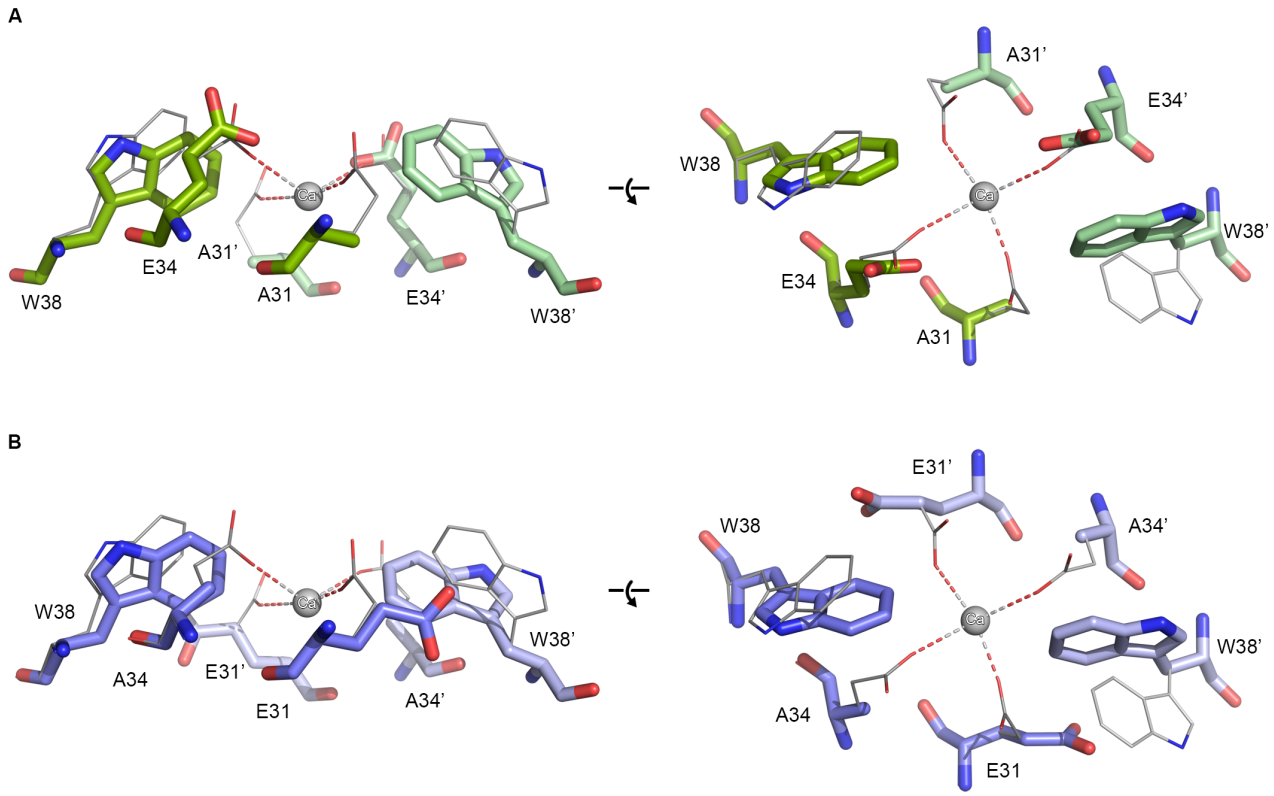

Figure 4. Comparison of side-chain positions in X-ray crystal structures of EncFtn entry site variants.
Residues in the putative metal ion entry site for the X-ray crystal structures of the E31A (**A**, green) and E34A (**B**, blue) EncFtn variants are shown with the EncFtn-sH wild-type (grey) for comparison. Residues in the metal ion entry site are shown as stick representations with a sphere indicating the position of the calcium ion found coordinated in the wild-type protein. Residues in the entry site are labelled, with residues from the second monomer indicated with prime symbols. In the E31A variant the Trp38 side chain is moved towards the putative metal ion coordinating residues, while the Glu34 side chain is in a similar conformation when compared to the wild type. The Trp38 side chain in E34A adopts a similar conformation to that seen in the E31A variant; while in the absence of the Glu34 side chain, the Glu31A side chain rotates by 120° away from the putative metal binding site.

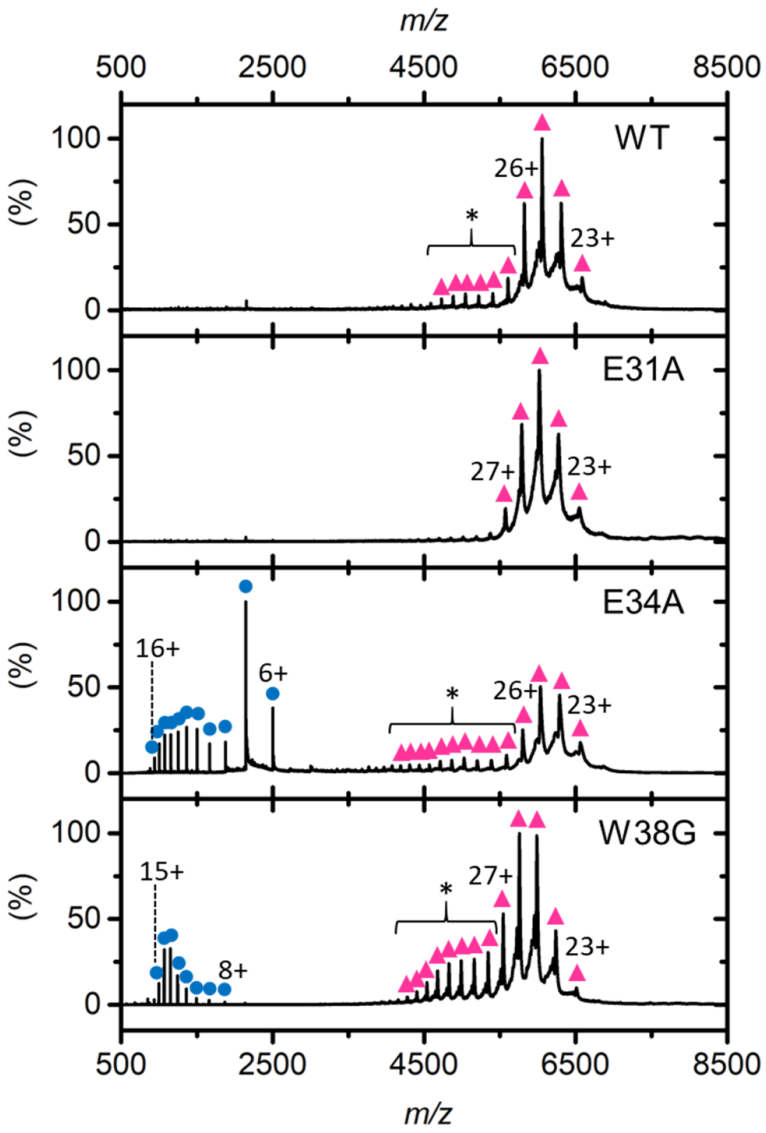

Figure 5. Native MS Analysis of EncFtn variants.
Native nESI spectra of EncFtn variants displaying decameric (pink triangles) and monomeric (blue circles) charge state distributions. The elongated decameric charge state distributions of WT and E34A are highlighted by an asterisk.

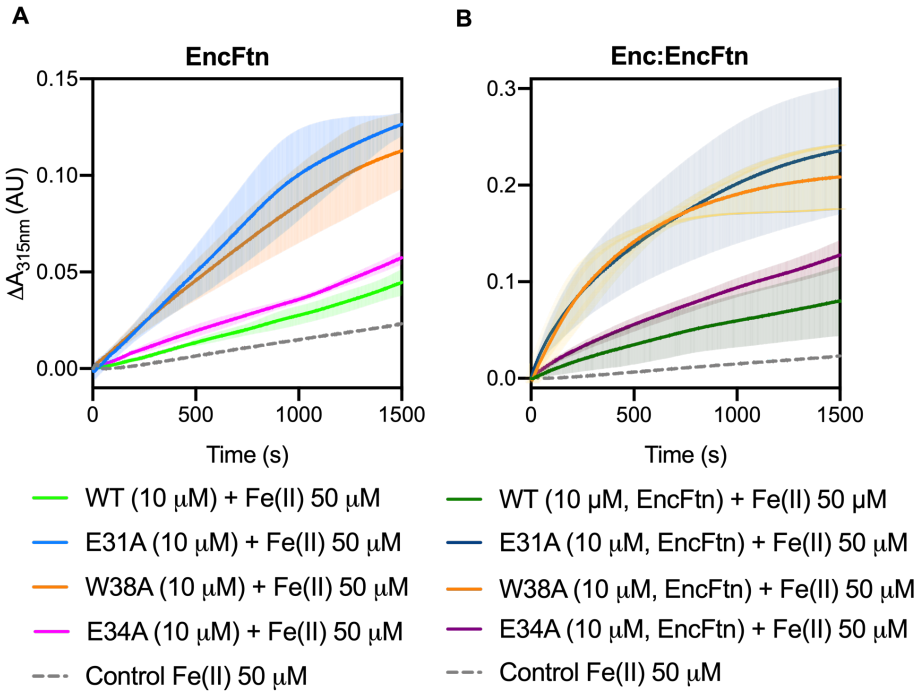

Figure 6. Ferroxidase activity of EncFtn and Enc:EncFtn complexes.
(**A**) EncFtn wild-type (green line), EncFtn-E34A (pink line), EncFtn-E31A (blue line), and EncFtn-W38A (orange line), (10 μM, monomer) were incubated with 50 μM $FeSO_4.7H_2O$ (10 times molar equivalent Fe(II) per FOC) and progress 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-WT (dark green line), Enc:EncFtn-E34A (dark pink line), Enc:EncFtn-E31A (dark blue line), and Enc:EncFtn-W38A (orange 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.
Comparison of data collected for W38A and W38G variants are shown in Figure S2, including elution profiles from gel-filtration purification steps (A/B), ferroxidase activities of Enc:EncFtn-W38A and W38G (C) and Enc:EncFtn-W38A and W38G transmission electron micrographs (D/E).
doi.10.6084/m9.figshare.9885575

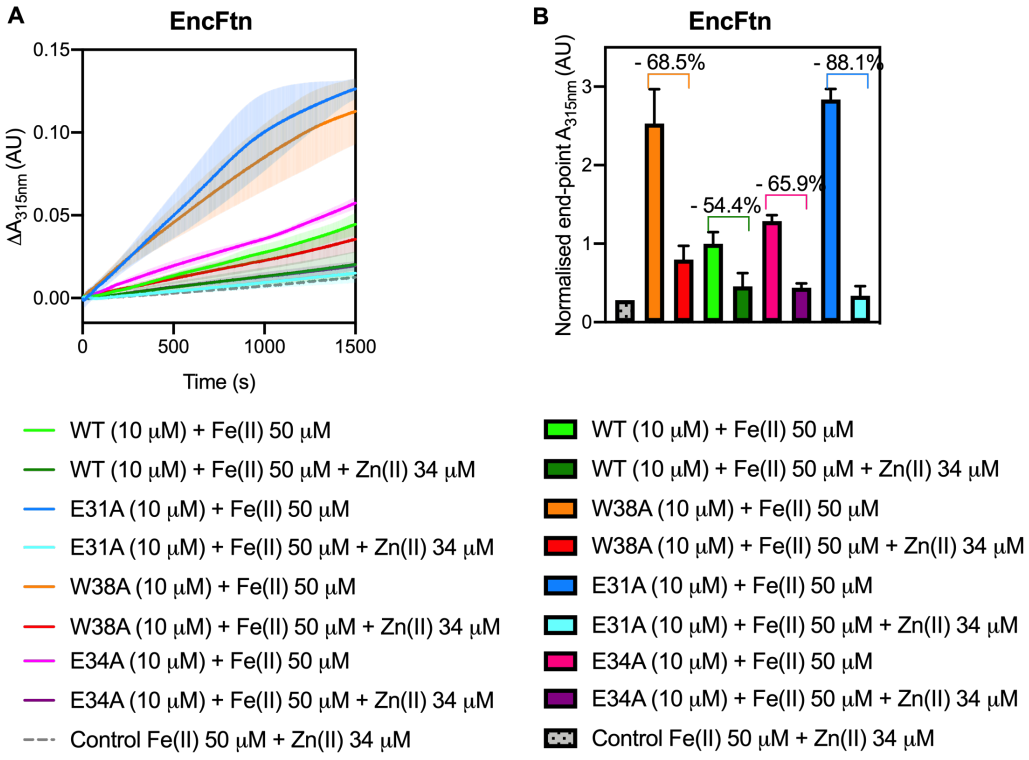

Figure 7. Ferroxidase activity of EncFtn wild-type and variants is inhibited by zinc. (**A**) Comparison between data shown in Figure 6A, collected in the absence of competing metals, are here plotted with curves from ferroxidase assays of EncFtn wild-type and variants carried out using [protein] = 10 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 EncFtn-Strep under the same experimental conditions (14). 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 7A were plotted as columns to compare inhibition by zinc. Percentage of inhibition is shown above corresponding columns. Color-coding is consistent with Figure 7A. doi.10.6084/m9.figshare.9885575

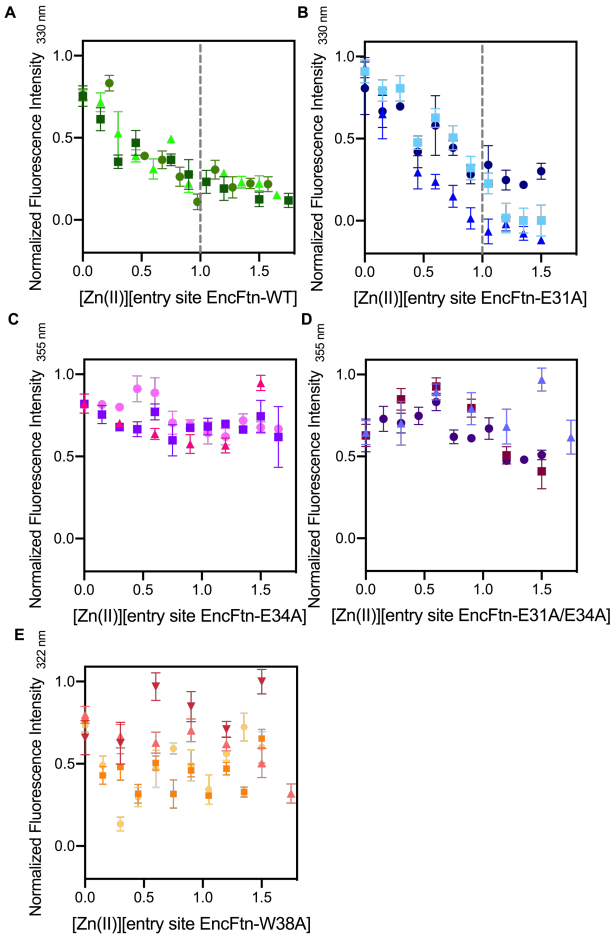

Figure 8. Intrinsic fluorescence of EncFtn wild-type and variants titrated with Zn(II). Fluorescence emission of EncFtn proteins (20 µM, corresponding to [entry-site] = 10 µM) at various wavelength (330 nm for EncFtn-WT and -E31A, 355 nm for EncFtn-E34A and E31A/E34A, and 322 nm for EncFtn-W38A) with excitation at 280 nm, and following titration with $ZnSO_4$.$7H_2O$. (**A/B**) Fluorescence is quenched upon zinc addition, with inflection around 1 molar equivalents, suggesting perturbation of environment around Trp38. (**C/D**) In the absence of entry site residues available for metal coordination no quenching was observed. (**E**) EncFtn-W38A was tested as control. Symbols represent average and standard deviation of experimental replicates, collected at equilibrium either by Scan (n >3) or Kinetic (n = 10) options on the Cary Eclipse software package. Data were normalized to allow comparison. Dotted grey lines were added at 1 molar equivalent of Zn(II) when inflection was observed. doi.10.6084/m9.figshare.11920512

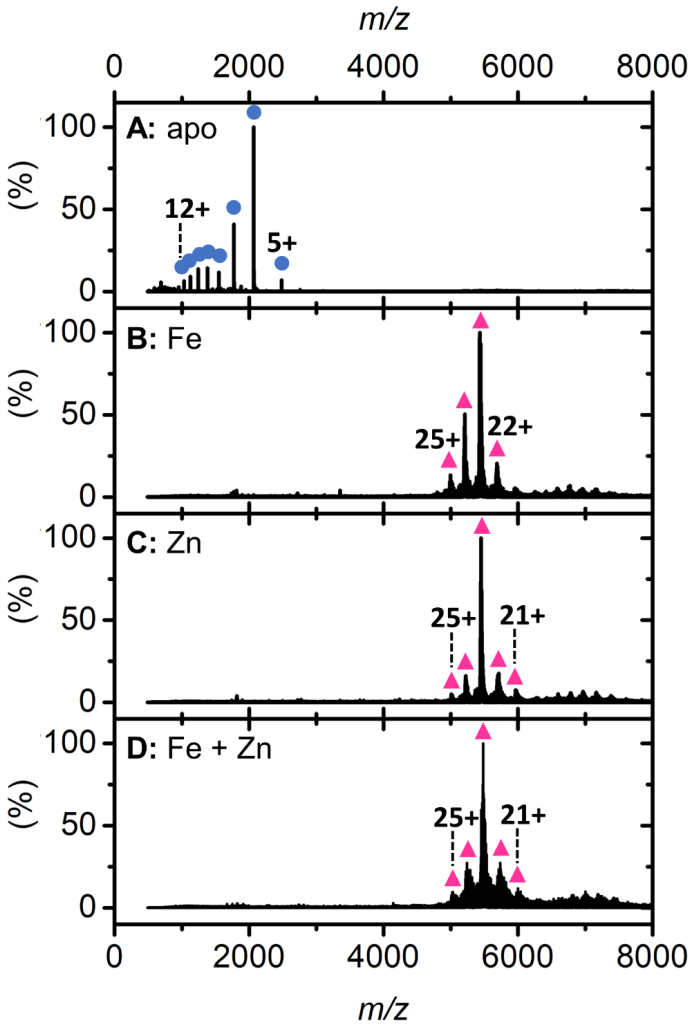

Figure 9. Native FT-ICR mass spectrum of *ID*-EncFtn
Native nESI mass spectra of: apo-*ID*-EncFtn (**A**); *ID*-EncFtn titrated with Fe(II) (**B**); *ID*-EncFtn titrated with Zn(II) (**C**); and *ID*-EncFtn titrated with Fe(II) followed by Zn(II) (**D**). Oligomerization states are stressed with colored shapes, monomer in blue circles and decamer in pink triangles.

Table 1. 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		
favored (%)	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

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