Pore dynamics and asymmetric cargo loading in an encapsulin nanocompartment revealed by Cryo-EM and hydrogen/deuterium exchange mass spectrometry

— **Source link** ↗

Jennifer L. Ross, Zak McIver, Thomas Lambert, Cecilia Piergentili ...+11 more authors

Institutions: University of Edinburgh, Newcastle University

Pore dynamics and asymmetric cargo loading in an encapsulin nanocompartment

Jennifer Ross[1†], Zak McIver[2], Thomas Lambert[1], Cecilia Piergentili[2], Kelly J. Gallagher[1], Jasmine Emma Bird[2], Faye L. Cruickshank[1], Efrain Zarazúa-Arvizu[3], Louise E. Horsfall[3], Kevin J. Waldron[5], Marcus D. Wilson[4], C. Logan Mackay[1], Arnaud Baslé[5], David J. Clarke[1]*, Jon Marles-Wright[2]*

Short title: Haliangium ochraceum encapsulin structure

[1]EaStCHEM School of Chemistry, The University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, EH9 3FJ.
[2]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU.
[3]School of Biological Sciences, The University of Edinburgh, Edinburgh, EH9 3BF.
[4] Wellcome Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Kings Buildings, Mayfield Road, Edinburgh, EH9 3JR, UK
[5]Newcastle University Biosciences Institute, Medical School, Newcastle University, Newcastle upon Tyne, NE2 4HH.

†Current address: School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD.

* To whom correspondence should be addressed:
Jon Marles-Wright, +44(0)191 2084855, Jon.marles-wright1@ncl.ac.uk
David J. Clarke, +44(0)131 650 4808, Dave.clarke@ed.ac.uk

Abstract

Encapsulins are protein nanocompartments that house various cargo enzymes, including a family of decameric ferritin-like proteins. Here, we study a recombinant *Haliangium ochraceum* encapsulin:encapsulated ferritin complex using electron cryo-microscopy and hydrogen/deuterium exchange mass spectrometry to gain insight into the structural relationship between the encapsulin shell and its protein cargo. An asymmetric single particle reconstruction reveals four encapsulated ferritin decamers in a tetrahedral arrangement within the encapsulin nanocompartment. This leads to a symmetry mismatch between the protein cargo and the icosahedral encapsulin shell. The encapsulated ferritin decamers are offset from the interior face of the encapsulin shell. Using HDX-MS, we observed dynamic behavior of the major five-fold pore in the encapsulin shell, and show the pore opening via the movement of the encapsulin A-domain. These data will accelerate efforts to engineer the encapsulation of heterologous cargo proteins and to alter the permeability of the encapsulin shell via pore modifications.

Teaser

Cryo-EM and HDX analysis of an encapsulin nanocompartment shows that the pores at the five-fold icosahedral vertex of the shell are flexible.

Introduction

Cellular metabolism and reaction pathways can produce toxic by-products which damage proteins, DNA, and lipids, or can become involved in potentially harmful side-reactions. Eukaryotes use membrane-bound organelles, such as lysosomes, to prevent this damage by housing dangerous reactions in chemically privileged environments. In a similar manner, prokaryotes use large protein-based compartments to sequester such reactions and act as barrier from the cytosol (*1, 2*). Prokaryotes use a variety of compartments such as carboxysomes, which are used for carbon dioxide fixation; and ferritins, for iron oxidation and storage (*3–5*).

One compartmentalization strategy utilized by prokaryotes is the encapsulin system (*6, 7*). Encapsulin (Enc) nanocompartments are hollow icosahedral complexes which range in size from 20 nm to 42 nm (*7–9*). Encapsulin proteins are structurally related to the viral capsid protein (gp5) of the HK97 bacteriophage and self-assemble from a single monomer into one of three forms: 60 subunits (*T* = 1 capsid symmetry), 180 subunits (*T* = 3 capsid symmetry) or 240 subunits (*T* = 4 symmetry)(*6, 7, 9, 10*). Encapsulins share a common feature of housing a cargo enzyme, such as ferritin-like proteins (encapsulated ferritins, EncFtn), iron-mineralizing encapsulin-associated firmicute (IMEF), or dye-decolorizing peroxidases(*9, 11*). Cargo enzymes are directed inside the encapsulin nanocompartment by a terminal localization sequence (LS) which binds to the interior face of the encapsulin(*7, 12*). Encapsulins and their cargo proteins are found throughout the bacterial and archaeal domains in species inhabiting a range of environmental niches; consequently, the proteins are stable in diverse physical conditions(*6, 13–15*). For these reasons, the encapsulins have attracted considerable interest for biotechnological applications, through their ability to separate potentially hazardous heterologous reactions from the native cytosol(*16, 17*).

The EncFtn cargo proteins are of particular interest, as they differ from their classical ferritin relatives. Although both proteins oxidize iron using a conserved catalytically active ferroxidase center (FOC), they have remarkably different structural architectures. Classical ferritins oxidize ferrous iron, Fe(II), into a mineral ferric form, Fe(III), which is then stored within a 24-meric 12 nm nanocage (*3, 18*). In contrast to this, the EncFtn proteins have an annular structure formed from a pentamer of dimers with the FOC active sites located at a dimer interface (*8, 13*). EncFtn oxidizes iron in a similar manner to other ferritins, but due to its open structure, it must be associated with an encapsulin nanocage to act as an iron store(*8*). Together the encapsulin EncFtn (Enc:EncFtn) complex can perform both the oxidation and storage functions of classical ferritins. However, due to its increased size when compared to classical ferritins, the Enc:EncFtn complex has the potential to house significantly greater quantities of iron, and has been described as an iron megastore(*10*). For these reasons, the encapsulins have attracted considerable interest for biotechnological applications, through their ability to separate potentially hazardous heterologous reactions from the host cytosol.

Although there have been several structural studies on encapsulins, a number of key questions remain unanswered. Most notably, for the EncFtn containing encapsulin nanocompartments, the structural relationship between the encapsulin shell and the EncFtn cargo protein is unknown. Studies on Dyp loaded encapsulin nanocompartments have shown loading with either one(*19*) or two hexameric complexes(*20*). Previous models for the loading of EncFtn in encapsulin nanocompartments have suggested a symmetric arrangement of the D5 decameric encapsulated ferritin at the five-fold icosahedral vertices of the encapsulin shell, giving a theoretical maximum of twelve EncFtn decamers per encapsulin nanocage(*8, 21*).

113   Herein, we investigate the structure of the Enc:EncFtn nanocompartment from the
114   halophilic bacterium *Haliangium ochraceum,* to gain a better understanding of the
115   arrangement and stoichiometry of the complex. Using the complementary structural
116   biology techniques of cryogenic electron microscopy (cryo-EM) and hydrogen-
117   deuterium exchange (HDX) MS, we study the structural relationship between Enc and
118   EncFtn in the encapsulin nanocage.
119   We present the first cryo-EM structure of a mesophilic Enc:EncFtn
120   nanocompartment, which reveals a symmetry-breaking tetrahedral arrangement of the
121   EncFtn decamers within the encapsulin shell. Analysis of the encapsulin shell by
122   symmetry expansion of an icosahedral reconstructions and focused 3D refinement on
123   the pentameric vertex reveals a flexible pore in the shell. The dynamic nature of this
124   pore region was investigated by HDX-MS. We show that this region has a high rate of
125   H/D exchange, demonstrating its conformational flexibility. Our combination of HDX-
126   MS and cryo-EM models affords insight into the loading capacity and dynamics of the
127   Enc:EncFtn nanocompartment system.
128

Results

Recombinant *Haliangium ochraceum* encapsulin complexes form regular nanocompartments that recruit active EncFtn cargoes

In order to gain an understanding of the relationship between encapsulin nanocompartments and their EncFtn cargoes, the Enc:EncFtn nanocompartment from the halophilic mesophile *Haliangium ochraceum* was chosen as our model system. This was primarily due to the high yields and ease of purification of the recombinant nanocompartment. Constructs for the production of empty (Empty-Enc) and EncFtn loaded (Loaded-Enc) encapsulin nanocompartments were produced for recombinant protein expression in *Escherichia coli*. The protein complexes were purified by heat treatment, followed by anion exchange and size-exclusion chromatography (**Figure 1A and Figure S1Ai, Aii, Bi** and **Bii**).

The molecular masses of the protein constituents of the Empty-Enc and Loaded-Enc assemblies were determined by LC-MS (**Table S1**). MS analysis of the Empty-Enc assembly revealed a single charge state distribution corresponding to a monomer of the encapsulin protein. MS analysis of the Loaded-Enc revealed three charge state distributions present, with deconvoluted masses consistent with the encapsulin protein monomer, a monomer of EncFtn and a dimer of EncFtn (**Figure S1Aiii** and **Biii**). These results indicate that the Loaded-Enc sample contained both the encapsulin and EncFtn cargo proteins, whilst Empty-Enc has only the encapsulin protein.

The assembly of the purified encapsulin nanocompartments was confirmed through visualization by negative stain transmission electron microscopy (TEM) (**Figure 1B**). Both empty and EncFtn loaded encapsulins assembled into regular nanocompartments, with an average diameter of approximately 21 nm, consistent with other T=1 type encapsulins (**Figure S1Aiv** and **1 Biv**)(*7, 22, 23*). The solution behavior of both complexes is consistent with the TEM observations, with both complexes eluting from a size-exclusion column at the same volume, indicating similar hydrodynamic radii (**Figure S1Ai** and **Bi).** The micrographs of the EncFtn loaded encapsulin reveal a regular internal density visible within the nanocompartment, suggesting that the EncFtn cargo has been encapsulated in an organized manner.

Ferroxidase assays confirmed the ability of the EncFtn loaded encapsulin nanocompartment to convert Fe(II) to Fe(III) (**Figure 1C**). This result is consistent with our previous observations for the *Rhodospirillum rubrum* Enc:EncFtn encapsulin complex(*8*). The empty encapsulin, which lacks the EncFtn cargo is enzymatically inactive.

Taken together, these data demonstrate that functionally active EncFtn has been successfully loaded into the encapsulin nanocompartment during expression in the heterologous *E. coli* host.

The cryo-EM structure of the Loaded-Enc nanocompartment

Motivated by the apparent interior density in the Loaded-Enc sample, we performed single particle cryo-EM of the Loaded-Enc complex (**Table S2**). Consistent with previously published X-ray crystallographic and cryo-EM derived encapsulin models, an initial reconstruction was produced with imposed I1 symmetry (**Figure 2, Figure S2**, and **Figure S3**) (*6, 7, 9, 22, 24*). This resulted in a reconstruction with a global resolution of 2.5 Å as determined by the gold-standard Fourier shell correlation at 0.143 (GS-FSC) (**Figure S2D**). The reconstruction displays a $T = 1$ icosahedral arrangement of sixty encapsulin monomers, with clearly resolved secondary structure elements. Small pores are visible in the shell at, or close to, the 2-, 3-, and 5-fold

179   symmetry axes where protein monomers interact with each other (**Figure 2A**).
180   Regions of the reconstruction around the icosahedral five-fold axes displayed a lower
181   resolution than the other regions of the structure in a local resolution map (**Figure 2B**).
182   This is consistent with observations of local resolution maps for the *Quasibacillus*
183   *thermotolerans* encapsulin reconstruction(*9*). The monomer of the encapsulin
184   nanocompartment from the reconstruction displays a HK97-fold typical for encapsulins
185   (**Figure 2C**). The orientation of the E-loop (extended loop) of the HK97 phage-like fold
186   determines the topology of the nanocompartments. The $T = 1$ Family 1 encapsulins
187   have their E-loop shifted away compared to the E-loops of those from the HK97
188   bacteriophage and other encapsulins (such as, the $T = 1$ Family II Enc from
189   *Synechococcus elongatus*; the $T = 3$ Enc from *Pyrococcus furiosus* and the $T = 4$ from
190   *Quasibacillus thermotolerans*). The encapsulin monomer has an E-loop orientation
191   similar to that of the $T = 1$ Family 1 Enc from *Thermotoga maritima* (**Figure 2C**)*.*
192   With the imposition of I1 symmetry on the reconstruction, the EncFtn cargo is
193   not visible. This suggests that the organization of the EncFtn protein within the
194   encapsulin shell does not conform to icosahedral symmetry and is rotationally
195   averaged through our symmetry-imposed processing (**Figure S2C**).
196
197   **The encapsulin nanocompartment recruits four EncFtn decamers to its lumen**
198   To gain insight into the structural relationship between the encapsulin shell and its
199   EncFtn cargo protein, a reconstruction was produced with no imposed symmetry
200   averaging (**Figure 3 and Figure S3**). Separation of the dataset into five 3D classes
201   revealed a highly populated class with amorphous density in the interior **(Figure S3,**
202   **panel 5B**), and two other main classes, both containing four distinct densities,
203   arranged in a similar tetrahedral fashion within the encapsulin shell. We took one of
204   these latter classes forward for full 3D refinement to produce a final C1 reconstruction
205   with a resolution of 3.7 Å at the 0.143 FSC threshold (**Figure 3A and Figure S4A**).
206   Interestingly, the EM map revealed four distinct densities within the encapsulin
207   nanocompartment lumen, which are consistent in size and shape with four EncFtn
208   decamers (**Figure 3** and **Figure S4B**).
209   A local resolution map calculated for the asymmetric reconstruction indicates a
210   degree of flexibility at the pentameric pores compared to the trimeric pores of the
211   encapsulin shell (**Figure 3A**). The interior of the nanocompartment shows a significant
212   falloff in resolution from the inner face of the encapsulin shell to the EncFtn densities.
213   This is consistent with the tethering of the EncFtn to the encapsulin nanocompartment
214   via its localization sequence with some degree of conformational freedom of the
215   EncFtn decamer with respect to the encapsulin shell.
216   The EncFtn decamers are located approximately 3 nm away from the encapsulin
217   interior wall, which corresponds to the linker region between the main EncFtn domain
218   and the localization sequence on the EncFtn C-terminus (**Figure S4C**). The extended
219   localization sequence of the EncFtn protein acts to offset it from the inner face of the
220   encapsulin shell, an observation consistent with previous reports of the IMEF
221   encapsulin complex from *Q. thermotolerans*(*9*)*.* Due to the dynamic nature of the
222   EncFtn within the encapsulin it was not possible to trace the path of the localization
223   sequence to its binding site.
224   The four densities within the encapsulin nanocompartment are discrete, and
225   thus permit the docking of the *H. ochraceum* EncFtn crystal structure in this region
226   (PDB: 5N5F)(*13*) (**Figure 3B** and **Figure S4C**). The decameric, annular crystal
227   structure fits well into the EncFtn density; although at the observed resolution, it is not
228   possible to fix the rotational alignment around the 5-fold symmetry axis of the EncFtn.

229   Despite extensive 3D classification and attempts at multibody refinement, the
230   resolution of the observed internal density did not improve.
231   Interestingly, the four EncFtn decamers are in a tetrahedral arrangement within
232   the encapsulin nanocompartment, with the five-fold axes of the EncFtn decamers
233   aligned to the three-fold tetrahedral axes (**Figure 3B**). This results in a double
234   symmetry mismatch between the icosahedral shell and the EncFtn decamers in the
235   complex. A concurrent structural study of the *Thermotoga maritima* encapsulin
236   complex(*25*) revealed five EncFtn decamers within the encapsulin shell, with each
237   decamer found in approximation to a pentameric vertex. The overall arrangement of
238   the five EncFtn complexes within the nanocompartment is incompatible with the
239   formation of a regular platonic solid and breaks the overall icosahedral symmetry of
240   the complex.
241   The symmetry mismatches found in both of these encapsulin complexes are
242   particularly interesting in terms of the functional relationship of the cargo proteins to
243   the encapsulin shell. The pores of the encapsulin nanocompartment allow substrate
244   access to the nanocompartment interior (*26*) and given the symmetry breaking
245   arrangement of the EncFtn decamers within both the *H. ochraceum* and *T. maritima*
246   encapsulins, the former are found in non-equivalent environments in terms of iron
247   availability. Analysis of the relationship between the EncFtn decamers and the inner
248   face of the *H. ochraceum* encapsulin shell reveals a number of distinct EncFtn
249   environments (**Figure S5**). The first EncFtn environment is shared by two EncFtn
250   decamers and is in line with the five-fold pore of the encapsulin nanocompartment
251   (EncFtn 1 and 2 in **Figure S5**). The shared symmetry of the Enc nanocompartment
252   five-fold pores and of the EncFtn D5 annular structure in these positions is consistent
253   with our previously proposed hypothesis for the Enc:EncFtn relationship(*8*) and is also
254   found in the *T. maritima* encapsulin. However, the symmetry-breaking tetrahedral
255   arrangement of the EncFtn decamers in the *H. ochraceum* encapsulin creates a
256   second distinct environment shared by the remaining two EncFtn decamers, where
257   they are offset between five-fold and three-fold axes of the icosahedral encapsulin
258   shell (EncFtn 3 and 4 in **Figure S5**). With the proposed route of iron entry through the
259   5-fold pores of the encapsulin shell, the EncFtn decamers in proximity to the pores
260   would have more favorable substrate access than those found in the alternative
261   positions.
262

263   **Structural dynamics in the pentameric vertices of the encapsulin shell**
264   To further investigate the apparent conformational flexibility of the encapsulin shell
265   at the pentameric vertices in the reconstructions, we performed symmetry expansion
266   on the I1 refined particle set, followed by masked 3D-classification without alignment
267   centered on the vertex. A number of distinct conformations were revealed, and the
268   most extreme of these were subjected to 3D refinement with local searches (**Figure
269   4**, **Figure S6**, and **Figure S7**). This resulted in an 'open' pentamer conformation of 2.4
270   Å resolution and a 'closed' conformation of 2.3 Å, allowing for fitting of residue side
271   chains (**Figure S7**).
272   The 'open' conformation has a five-fold pore with an aperture diameter of
273   approximately 24 Å, while in the 'closed' conformation the aperture is reduced to 9 Å
274   diameter. To understand the structural changes taking place in the transition between
275   these conformations, an atomic model of the encapsulin protein was refined against
276   both maps (**Table S3**). The two models show a significant movement in the A-domain,
277   with a pivoting around the hinge points connecting this domain to the P-domain,
278   opening the pore like an iris (**Figure 4**). In the open conformation, the pore loop region

279 (residues 182 - 189) is not well defined in the density; while it is tightly locked in the
280 closed conformation, with Asp186 forming the outer boundary of the pore and Tyr188
281 and Lys192 forming the inner bounds (**Figure S8**). The tyrosine is well conserved
282 among the family 1 *T*=1 encapsulins, while the lysine is substituted for a glutamine in
283 the *R. rubrum* encapsulin (**Figure S9**). The family 2 *T*=1 encapsulin from
284 *Synechococcus elongatus* has a five-residue sequence insertion in this region, which
285 forms an extended linker between secondary structure elements, rather than a distinct
286 loop within the pore.
287   In the *H. ochraceum* encapsulin the five-fold pore has a negative charge on the
288 exterior of the encapsulin shell and positive charge on the interior in both the open and
289 closed conformations (**Figures S9 and S10**). The closed conformation is consistent
290 with observations from the crystal structure of the *T. maritima* encapsulin (*7*) and high
291 resolution cryo-EM structures of other encapsulins(*9*, *24*, *25*). However, this is the first
292 time that an 'open' pore-conformation has been observed in an encapsulin protein.
293 This observation has important implications for efforts to engineer the pores of
294 encapsulin nanocages. Where early efforts to widen the five-fold pores have
295 demonstrated an increase in mass-transport of model substrates across the
296 encapsulin shell(*26*), more recent investigations into pore modifications have shown
297 that the shell does not act as a strong barrier to the passage of the small lanthanide
298 substrates tested(*27*). Our results provide an explanation for these observations,
299 where a dynamic and flexible pore would not act as a barrier to the passage of small
300 ligands, such as divalent cations, across the shell. They would also be able to
301 accommodate a wide range of sizes of potential ligands for engineered nanocages.
302   Additionally, our focused refinements of the pentameric subunits allowed us to
303 build and sequence the $_{117}$GSLGIGSLR$_{125}$ peptide from the EncFtn protein (**Figure 5**).
304 This region of the localization sequence forms a network of hydrophobic interactions
305 with the inner face of the P-domain of a single encapsulin monomer, with further
306 stabilization by a number of water-mediated backbone contacts. The core GxLGIxxL
307 motif found in this region of the localization sequence is conserved between the *H.*
308 *ochraceum* EncFtn and other proteins in the family and is observed in the crystal
309 structure of the *T. maritima* encapsulin(*7*).
310
311 ***Dynamics of the 5-fold encapsulin pore through hydrogen/deuterium exchange***
312 ***mass spectrometry.***
313   To further investigate the dynamic nature of the five-fold pore of the encapsulin
314 shell and the docking of the EncFtn localization sequence to the interior of the
315 nanocompartment, we performed hydrogen/deuterium exchange mass spectrometry
316 (HDX-MS) on both Empty-Enc and Loaded-Enc nanocompartments. The extent of
317 backbone-amide hydrogen exchange was determined at seven time points (0
318 seconds, 10 seconds, 30 seconds, 5 minutes, 30 minutes, 4 hours, and 24 hours). By
319 calculating the rate of hydrogen exchange throughout the protein, regions that differ in
320 solvent exposure and/or dynamics can be detected.
321   HDX-MS analysis of the encapsulin nanocompartment resulted in 40 pepsin
322 peptides, which constituted a protein sequence coverage of 85%, with peptide
323 redundancy of 2.28 (**Table S4**, **Table S5** and **Figure S11**). The encapsulin
324 nanocompartment displayed variable exchange rates throughout the protein sequence
325 and regions of the protein displaying elevated H/D exchange rates were clearly
326 evident. Overlaying these local H/D exchange rates onto the cryo-EM reconstruction
327 revealed that the regions of highest exchange were located around pentameric
328 vertices. (**Figure 6**). This was most notable with the peptide spanning the region

329   between amino acids 180-196, which includes the five-fold pore loop (**Figure S12**). In
330   contrast, lower rates of HDX are observed at the 2-fold interface and the potential 3-
331   fold pore (**Figure 6**). These findings are in agreement with our cryo-EM structural
332   analyses and support the proposed conformational flexibility at the 5-fold pore.
333   Comparison of the H/D exchange rates of Empty-Enc and Loaded-Enc revealed
334   similar exchange profiles throughout the encapsulin protein sequence, suggesting that
335   cargo loading has little effect on the overall architecture and dynamics of the
336   assembled nanocompartment shell (**Table S4**, **Table S5**, and **Figure S13**). However,
337   after prolonged exchange times (4 hours), the Loaded-Enc exhibited areas with a
338   modest reduction in exchange when compared to the Empty-Enc. Notably, several
339   peptides in the N-terminal region displayed reduced exchange rates in Loaded-Enc;
340   for example, the peptide covering amino acids 21-37 displayed almost twelve percent
341   reduction. Mapping the position of this region onto our encapsulin reconstruction
342   highlights that this peptide is located on the interior face of the nanocompartment and
343   included the proposed binding site for the localization sequence of EncFtn (**Figure 5**).
344   A reduction in exchange across in this region is likely a consequence of shielding by
345   the engaged EncFtn localization sequences.

Discussion

347   Our cryo-EM reconstruction of an Enc:EncFtn nanocompartment complex reveals
348   key areas of divergence from a true icosahedral complex with important functional
349   consequences. The asymmetric reconstruction showed that the encapsulin
350   nanocompartment sequesters four decamers of EncFtn within its lumen. In our
351   recombinant system, with EncFtn produced in excess, this likely represents a
352   maximum loading capacity for the Enc:EncFtn nanocompartments. The symmetry
353   breaking tetrameric arrangement of the EncFtn decamers within the encapsulin shell
354   leads to two distinct environments for EncFtn, with two decamers aligned at the five-
355   fold symmetry axes, and the remaining two residing between three- and five-fold axes
356   (**Figure S5**). The concurrent observation of five EncFtn decamers within the *T.*
357   *maritima* encapsulin nanocompartment, aligned close to the icosahedral five-fold
358   symmetry axes of the encapsulin shell, highlights differences in cargo loading between
359   encapsulins from mesophilic and thermophilic bacteria(*28*). The enhanced cargo
360   loading seen in the *T. maritima* encapsulin may be a consequence of a more compact
361   and rigid structure adopted by thermophilic proteins as seen in comparisons of the
362   crystal structures of the *H. ochraceum* and *Pyrococcus furiosus* EncFtn proteins(*13*).
363   The interior volume of the *T=1* encapsulin lumen is around 4000 nm^3, while an
364   EncFtn decamer is only 120 nm^2; therefore, the additional cargo seen in the *T.*
365   *maritima* encapsulin should not impact the iron-storage potential significantly. The
366   order of magnitude discrepancy in the iron-loading capacity measured for the *R.*
367   *rubrum*(*8*) and *T. maritima*(*28*) encapsulins is no doubt the result of differences in
368   experimental conditions in different laboratories.
369   While the volume occupied by the EncFtn decamers represents less than 15 % of
370   the total lumen of the encapsulin nanocompartments, the consistent observation of a
371   gap between the encapsulin shell and encapsulin cargo proteins implies that the
372   loading and capacity of encapsulins is limited by broader steric effects. These would
373   include the offset of the cargo protein from the shell, and from the unengaged
374   localization sequences at the core of the nanocompartment. This has implications for
375   efforts to target heterologous proteins to the encapsulin nanocage; effectively setting
376   a limit on the volume of protein that can be accommodated within, which is much lower
377   than the total volume of the lumen of the nanocage.

378   These observations have functional implications for the oxidation and storage of
379   iron within the Enc:EncFtn nanocompartment. The EncFtn decamers are in non-
380   equivalent positions, and thus have different relationships to the pores of the
381   encapsulin shell. Therefore, if the pores limit the diffusion of substrates, the EncFtn
382   decamers would be subjected to different chemical environments. Furthermore, both
383   the engaged and unengaged localization sequences present a 'soft' steric barrier to
384   the diffusion of substrates. It is notable that the ferroxidase activity of the Enc:EncFtn
385   complex is significantly higher than the isolated EncFtn protein. While it is not possible
386   to make mechanistic conclusions from our model, the complex interactions with the
387   components of the encapsulin nanocompartment clearly enhance the iron oxidation
388   activity of the EncFtn protein.
389   Our data were collected on iron-free apo-Enc:EncFtn complexes, and thus it is not
390   possible to infer the nature of the iron mineralization pathway within the encapsulin
391   nanocage. Further careful work must be performed to titrate iron into the complex prior
392   to structural analysis to gain insight into the flow of metal ions from the exterior to the
393   interior of the encapsulin and to determine if metalation influences the conformational
394   flexibility of the EncFtn within the encapsulin nanocage, as we have demonstrated for
395   isolated EncFtn proteins(*8*, *13*, *29*). Finally, the nature of the iron mineral and its
396   localization within the encapsulin nanocage is still to be determined.
397   These cryo-EM and HDX-MS data illustrating a highly dynamic five-fold pore in the
398   encapsulin shell have major implications for efforts to engineer recombinant
399   encapsulins for improved access for both native and non-native substrates. The
400   limitations of previously published studies where the five-fold pore is modified for
401   altered substrate access can be explained by a highly dynamic pore structure that is
402   not particularly discriminatory for small molecules. Our work suggests new hypotheses
403   for engineering pore selectivity, through modifications to the hinge regions between
404   the P- and A-domains, which are responsible for the opening of the pore.
405
406

407 Materials and Methods

408

409 ***Experimental Design***

410   The objective of this study was to understand the structural relationship
411 between encapsulins and their EncFtn cargo using cryo-EM to determine the complex
412 structure and ferroxidase assays for validation of the complex activity. HDX was
413 utilised to establish differences between solvent accessibility of empty and loaded
414 encapsulins and garner insight into the impact of cargo on the interior of the encapsulin
415 nanocompartment.

416

417 ***Cloning of encapsulin expression constructs***

418   The *Haliangium ochraceum* encapsulin and encapsulated ferritin protein
419 expression constructs were based on the Hoch_3836 and Hoch_3837 genes and were
420 codon optimized for expression in *Escherichia coli* and synthesized as CIDAR MoClo
421 compatible gBlocks by Integrated DNA Technologies (IDT) (**Table S6**). The gBlocks
422 were assembled into a Level 0 CIDAR MoClo storage vector(*30*), DVA_CD, for
423 subsequent use. The coding sequences for the encapsulin and the EncFtn were
424 assembled into expression cassettes in the level 1 backbones DVK_AE and DVK_EF
425 respectively, each with T7 promoter and transcription terminator parts. The resulting
426 expression cassettes were then combined into the DVA_AF backbone to produce a
427 co-expression plasmid. All assembled plasmids were sequence verified by Sanger
428 sequencing by Eurofins Genomics. The protein sequences for each construct are
429 listed in **Table S7**.

430

431 ***Protein expression***

432   The Empty-Enc and Loaded-Enc expression plasmids were transformed into *E.*
433 *coli* BL21(DE3) cells and grown overnight at 37 °C on LB-agar plates containing
434 appropriate selection antibiotics (kanamycin for Empty-Enc and ampicillin for Loaded-
435 Enc). A single colony of cells was added to 1 L of autoinduction media(*31*)
436 supplemented (**Table S8**) with appropriate antibiotic and grown for 38 hours at 37 °C
437 with shaking at 200 rpm. Cells were harvested by centrifugation at 12,000 ´ *g*.

438

439 ***Encapsulin nanocompartment purification***

440   *E. coli* cell pellets expressing the empty-Enc and loaded-Enc constructs were
441 resuspended in 10 × v/w of lysis buffer (20 mM HEPES, pH 8; 2 mM $MgCl_2$; 1 mg/ml
442 lysozyme, and benzonase, 12.5 – 25 units/mL). Cells were lysed by sonication whilst
443 on ice; sonication was carried out in six 1-minute cycles (30 seconds sonication, 30
444 seconds rest). The lysate was clarified by centrifugation at 20,000 ´ g for 1 hour, 4 °C.
445   The supernatant from cell lysis was heated to 85 °C for 10 minutes in a water
446 bath and transferred to a 4 °C ice bath for 10 minutes. The supernatant was then
447 collected after centrifugation at 10,000 g for 1 hour.
448   Anion exchange chromatography of the clarified supernatant was performed
449 using a 1 mL HiTrap Q Sepharose FF column from Cytiva on an ÄKTA™ start. The
450 column and ÄKTA™ start system were equilibrated with QA buffer (20 mM HEPES,
451 pH 8.0) and the protein sample was loaded. Unbound proteins were removed by
452 washing with QA buffer. Bound proteins were eluted by QB buffer (20 mM HEPES, pH
453 8.0, 1 M NaCl) over a linear gradient of 0-100% QB over 15 column volumes. Flow-
454 through fractions containing the sample were subjected to SDS-PAGE to identify those

455 containing the protein of interest. These fractions were pooled and concentrated using
456 centrifugal concentrators with a 30 kDa nominal molecular weight cut off (Vivaspin).
457   Pooled and concentrated samples from the anion exchange step were loaded
458 on a gel filtration column (Sephacryl 400, Cytiva) equilibrated with SEC buffer (20 mM
459 HEPES, pH 8.0, 150 mM NaCl). Fractions eluting from the column containing the
460 desired protein, as identified by SDS-PAGE were pooled and concentrated as above.
461 Protein aliquots were flash cooled in liquid nitrogen and stored at -80 °C (**Figure S1Ai**,
462 **Aii**, **Bi and Bii**). (*8*, *13*).
463

Negative stain TEM

465   Purified encapsulin nanocompartments were initially imaged by negative stain
466 TEM. Continuous carbon/formvar coated copper grids (200 mesh) were glow-
467 discharged for 30 seconds using a Pelco glow discharge system. 5 µL Enc was
468 pipetted onto the glow-discharged grids and excess liquid was removed after 30
469 seconds with Whatman filter paper (grade 1, diameter 24.0 cm). The grids were
470 washed with distilled water three times, followed by staining with 2 % uranyl acetate
471 for 5 seconds. Grids were left to air dry and then imaged with a JEOL JEM-1400
472 transmission electron microscope. Images were collected with a Gatan CCD OneView
473 camera and analyzed using FIJI(*32*).
474

Ferroxidase activity assay

476   The enzymatic activity of Empty-Enc and Loaded-Enc were assessed by
477 ferroxidase assay, as previously described (Piergentili, 2020). Fe(II) samples were
478 prepared by dissolving $FeSO_4.7H_2O$ in HCl 0.1 % (v/v) under anaerobic conditions.
479 Protein samples were diluted anaerobically in Buffer GF (20 mM HEPES, pH 8.0, 150
480 mM NaCl) to a final encapsulin monomer concentration of 9 µM to allow comparison
481 between experiments.
482   Iron and protein aliquots were added aerobically to a quartz cuvette (Hellma)
483 resulting in a final concentration of 100 µM iron and 15 µM (Loaded-Enc), or 9 µM
484 (Empty-Enc). The cuvette was placed in a UV-visible spectrophotometer (PerkinElmer
485 Lambda 35) and the reaction sample was incubated at 21 °C for 50 s to stabilise.
486 Absorbance at 315 nm was then recorded every second for 1450 s using the Time-
487 Drive software. A control experiment was conducted by monitoring the background
488 oxidation by atmospheric oxygen of 100 µM $FeSO_4*7H_2O$ in the absence of the
489 enzyme. Each experiment was carried out in three or more technical replicates, with
490 replicate means and standard deviations calculated on the time zero-subtracted
491 progress curves.
492

Liquid Chromatography Mass Spectrometry

494   LC-MS experiments were performed on a Synapt G2 Q-ToF instrument (Waters
495 Corp., Manchester, UK) and an Acquity UPLC equipped with a reverse phase C4 Aeris
496 Widepore 50 × 2.1 mm HPLC column (Phenomenex, CA, USA). Mobile phases of; A=
497 water + 0.1% formic acid, and B=acetonitrile + 0.1% formic acid were used on a ten-
498 minute gradient from 5% B to 95% B. Samples were analysed at ~2 µM, and data
499 analysis was performed using MassLynx v4.1 and MaxEnt deconvolution.
500

501 **_Cryo-EM data collection and analysis_**

502 _Sample vitrification_

503 Holey grids (gold, 200 mesh, r 2/2 by Quantifoil) were glow-discharged for 30
504 seconds using a Pelco glow discharge system. The grids were then mounted into a
505 FEI vitrobot and 4 µL of encapsulin sample (3 mg/mL) was applied. Grids were then
506 blotted (100% humidity, 8 °C, blot force -5, wait time 10 seconds and blot time of 3
507 seconds) with Whatman filter paper (grade 1) and flash cooled in liquid ethane, cooled
508 with liquid nitrogen.

509

510 _Cryo-EM data collection_

511 Cryo-EM grid screening was performed on a FEI F20 microscope equipped with
512 a FEG electron source (200 kV) and a TVIPS F816 CMOS detector at the University
513 of Edinburgh. The dataset used for single particle reconstruction was obtained at eBIC
514 on a FEI Titan Krios microscope equipped with Gatan K3 camera (data collection
515 settings are shown in **Table S2**). Alignments, grid transfer and imaging set up was
516 performed by the eBIC local contact Dr Yun Song.

517

518 _Single particle reconstruction_

519 All processing steps were performed with the Relion 3.1 software package (_33_).
520 Super-resolution movies were binned (2 × 2 pixels for I1 reconstructions; 3 × 3 for C1
521 reconstructions) and motion corrected using MotionCor2(_34_). Defocus values of
522 imaged (summed movies) were determined by CTFFIND4 (_35_) and those with poor
523 CTF fits, or bad ice, were manually discarded. A template for autopicking was created
524 using 2D classes from manually picked particles. Autopicked particles were extracted
525 (using box sizes of 576 pixels (376 Å) and 512 pixels (501 Å) for icosahedral and C1
526 processing respectively) and subjected to three rounds of 2D classification to remove
527 bad particles. An initial 3D model was created from particles selected from 2D classes.
528 3D classes were generated both with and without icosahedral symmetry imposed ('I1'
529 and 'C1' symmetry). The best class from each was taken forward for 3D refinement
530 and then CTF refinement followed by further rounds of 3D refinement and Bayesian
531 polishing. After a final round of 3D refinement, postprocessing was performed using a
532 soft spherical mask. Local resolution estimation was performed in Relion3.1. The data
533 processing and refinement pipeline is shown in **Figure S3** with data processing and
534 refinement statistics in **Table S8**.

535 Motivated by the apparent flexibility of the 5-fold pores of the encapsulin shell,
536 the 5-fold pore pentamer was subjected to symmetry expansion and focused
537 classification. The I1 particle set was expanded using the sym_expand job in Relion3.1
538 and masked 3D classification without alignment was performed focused on the five-
539 fold symmetry axis. Of the five classes produced the two most highly populated and
540 distinct were taken forward for masked 3D refinement with local searches only, these
541 two classes represent the 'open' and 'closed' conformation of the five-fold pore. The
542 mask used in these steps was produced in Chimera using the molmap command from
543 a docked model of a pentamer of the encapsulin protein. To ensure the box-size and
544 pixel-size of the mask were correct, they were resampled onto the icosahedral map
545 using the vop resample command.

546

547 _Model building and refinement_

548 An initial homology model of the encapsulin nanocompartment monomer was
549 generated using Phyre 2.0(_36_) based on the _T. maritima_ structure. This was docked
550 into the open and closed maps using ChimeraX(_37_) and expanded to a full pentamer

551 model. The model was then fit to the map through an iterative process of automated
552 model refinement with phenix real-space refinement(*38*) and manual model building
553 in Coot(*39*), waters were added using phenix.douse and validated in in Coot. The
554 resulting models and maps were validated using Molprobity(*40*), phenix.mtriage(*41*),
555 and EM ringer(*42*) (**Table S2**). Models and maps were visualized using ChimeraX.
556
557 ***Encapsulin sequence analysis***
558 Encapsulin sequences were obtained from the Kyoto Encyclopedia of Genes
559 and Genomes (www.kegg.jp). Sequence alignments were performed using Clustal
560 Omega(*43*) and visualized using ESPript(*44*).
561
562 ***Hydrogen/Deuterium Exchange Mass Spectrometry***
563 Hydrogen/deuterium exchange mass spectrometry (HDX-MS) was performed
564 on a Synapt G2 MS system coupled to an ACQUITY UPLC M-Class UPLC with the
565 HDX manager module (Waters Corporation, Manchester, UK)(*45*). For improved
566 reliability and precision, a custom-built Leap automated platform was zutilized in all
567 sample preparation and injections. Prior to HDX-MS analysis, three buffer solutions
568 were prepared – (i) equilibration buffer (4.7 mM K_2HPO_4, 0.3 mM KH_2PO_4 in H_2O),
569 adjusted to pH 8.0 with formic acid; (ii) labelling buffer (4.7 mM K_2HPO_4, 0.3 mM
570 KH_2PO_4 in D_2O), adjusted to pH 8.0 with DCl and quench buffer (50 mM K_2HPO_4, 50
571 mM KH_2PO_4 in H_2O) adjusted to pH 2.3 with formic acid. Protein samples were diluted
572 in equilibration buffer to final stock concentration of 42 mM. The time course
573 experiments consisted of 7 timepoints: T0 (0 minute; undeuterated control), T1 (20
574 seconds), T2 (30 seconds), T3 (2 minutes), T4 (5 minutes) and T5 (30 minutes) T6 (4
575 hours) and T7 (24 hours) with each timepoint being performed in triplicate. Sample
576 preparation consisted of 5 µL protein solution, 57 µL equilibrium buffer (T0) or labelling
577 buffer (T1-7). The final concentration of deuterium during the labelling step was 91.2
578 %. Exchange was allowed to proceed at 4 °C. To arrest the exchange reaction, 50 µL
579 of quench buffer was added to this initial solution just prior to sample injection.
580 After injection, samples underwent proteolytic digestion on a 2.1 x 30 mm
581 Waters Enzymate BEH pepsin column for 3 minutes at 200 µL/min. After digestion, the
582 peptide digest was loaded on to a 2.1 x 5.0 mm Acquity BEH C18 VanGuard 1.7µm
583 C18 Trapping column to pre-concentrate the sample for 3 minutes at 200 µL/min.
584 Following trapping, the digests were separated through a 2.1 x 5.0 mm Acquity BEH
585 1.7 µm analytical column prior to MS/MS (MS^e) analysis via the Water Synapt G2 MS
586 system running MassLynx v4.1 software (Waters Corporation, Manchester, UK). The
587 separation gradient was 5-95 % acetonitrile with 0.1 % formic acid over 12 minutes at
588 40 µL/min. Both the trapping and LC separation were performed at 1 °C to minimize
589 back exchange. Post-processing was performed using Proteinlynx Global Server 3.0.3
590 and Dynamx 3.0 software to determine the average deuterium uptake for each peptide
591 at each time point. For comparative analyses, the relative fractional uptake was
592 determined by dividing the observed deuterium uptake by the number of available
593 amide exchangers on each peptide.
594

References

1. Y. Diekmann, J. B. Pereira-Leal, Evolution of intracellular compartmentalization. *Biochem. J.* **449**, 319–331 (2013).

2. C. A. Kerfeld, S. Heinhorst, G. C. Cannon, Bacterial Microcompartments (2010) (available at http://www.annualreviews.org/eprint/uv9wHHypNKsEb5yrt6z8/full/10.1146/annurev.micro.112408.134211).

3. S. C. Andrews, The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochim. Biophys. Acta*. **1800**, 691–705 (2010).

4. E. Chiancone, P. Ceci, A. Ilari, F. Ribacchi, S. Stefanini, Iron and proteins for iron storage and detoxification. *BioMetals*. **17**, 197–202 (2004).

5. T. O. Yeates, C. A. Kerfeld, S. Heinhorst, G. C. Cannon, J. M. Shively, Protein-based organelles in bacteria: carboxysomes and related microcompartments. *Nat. Rev. Microbiol.* **6**, 681–691 (2008).

6. F. Akita, K. T. Chong, H. Tanaka, E. Yamashita, N. Miyazaki, Y. Nakaishi, M. Suzuki, K. Namba, Y. Ono, T. Tsukihara, A. Nakagawa, The crystal structure of a virus-like particle from the hyperthermophilic archaeon Pyrococcus furiosus provides insight into the evolution of viruses. *J. Mol. Biol.* **368**, 1469–83 (2007).

7. M. Sutter, D. Boehringer, S. Gutmann, S. Günther, D. Prangishvili, M. J. Loessner, K. O. Stetter, E. Weber-Ban, N. Ban, Structural basis of enzyme encapsulation into a bacterial nanocompartment. *Nat. Struct. Mol. Biol.* **15**, 939–947 (2008).

8. D. He, S. Hughes, S. Vanden-Hehir, A. Georgiev, K. Altenbach, E. Tarrant, C. L. Mackay, K. J. Waldron, D. J. Clarke, J. Marles-Wright, Structural characterization of encapsulated ferritin provides insight into iron storage in bacterial nanocompartments. *Elife*. **5**, e18972 (2016).

9. T. W. Giessen, B. J. Orlando, A. A. Verdegaal, M. G. Chambers, J. Gardener, D. C. Bell, G. Birrane, M. Liao, P. A. Silver, Large protein organelles form a new iron sequestration system with high storage capacity. *Elife*. **8**, 1–23 (2019).

10. C. A. McHugh, J. Fontana, D. Nemecek, N. Cheng, A. A. Aksyuk, J. B. Heymann, D. C. Winkler, A. S. Lam, J. S. Wall, A. C. Steven, E. Hoiczyk, A virus capsid-like nanocompartment that stores iron and protects bacteria from oxidative stress. *EMBO J.* **33**, 1896–1911 (2014).

11. T. W. Giessen, P. A. Silver, Widespread distribution of encapsulin nanocompartments reveals functional diversity. *Nat. Microbiol.* **2**, 17029 (2017).

12. A. Tamura, Y. Fukutani, T. Takami, M. Fujii, Y. Nakaguchi, Y. Murakami, K. Noguchi, M. Yohda, M. Odaka, Packaging guest proteins into the encapsulin nanocompartment from Rhodococcus erythropolis N771. *Biotechnol. Bioeng.* **112**, 13–20 (2015).

13. D. He, C. Piergentili, J. Ross, E. Tarrant, L. R. Tuck, C. L. Mackay, Z. McIver, K. J. Waldron, D. J. Clarke, J. Marles-Wright, Conservation of the structural and functional architecture of encapsulated ferritins in bacteria and archaea. *Biochem. J.* **476**, 975–989 (2019).

14. R. Rahmanpour, T. D. H. Bugg, Assembly in vitro of Rhodococcus jostii RHA1 encapsulin and peroxidase DypB to form a nanocompartment. *FEBS J.* **280**, 2097–2104 (2013).

645 15. J. Snijder, O. Kononova, I. M. Barbu, C. Uetrecht, W. F. Rurup, R. J. Burnley,
646   M. S. T. Koay, J. J. L. M. Cornelissen, W. H. Roos, V. Barsegov, G. J. L.
647   Wuite, A. J. R. Heck, Assembly and Mechanical Properties of the Cargo-Free
648   and Cargo-Loaded Bacterial Nanocompartment Encapsulin.
649   *Biomacromolecules*. **17**, 2522–9 (2016).
650 16. M. C. Jenkins, S. Lutz, *ACS Synth. Biol.*, in press,
651   doi:10.1021/acssynbio.0c00636.
652 17. Y. H. Lau, T. W. Giessen, W. J. Altenburg, P. A. Silver, Prokaryotic
653   nanocompartments form synthetic organelles in a eukaryote. *Nat. Commun.* **9**,
654   1311 (2018).
655 18. S. Recalcati, E. Gammella, P. Buratti, G. Cairo, Molecular regulation of cellular
656   iron balance. *IUBMB Life*. **69**, 389–398 (2017).
657 19. M. M. Jore, M. Lundgren, E. van Duijn, J. B. Bultema, E. R. Westra, S. P.
658   Waghmare, B. Wiedenheft, Ü. Pul, R. Wurm, R. Wagner, M. R. Beijer, A.
659   Barendregt, K. Zhou, A. P. L. Snijders, M. J. Dickman, J. A. Doudna, E. J.
660   Boekema, A. J. R. Heck, J. van der Oost, S. J. J. Brouns, U. Pul, R. Wurm, R.
661   Wagner, M. R. Beijer, A. Barendregt, K. Zhou, A. P. L. Snijders, M. J.
662   Dickman, J. A. Doudna, E. J. Boekema, A. J. R. Heck, J. van der Oost, S. J. J.
663   Brouns, E. Van Duijn, J. B. Bultema, E. R. Westra, S. P. Waghmare, B.
664   Wiedenheft, Ü. Pul, R. Wurm, R. Wagner, M. R. Beijer, A. Barendregt, K.
665   Zhou, A. P. L. Snijders, M. J. Dickman, J. A. Doudna, E. J. Boekema, A. J. R.
666   Heck, J. Van Der Oost, S. J. J. Brouns, E. van Duijn, J. B. Bultema, E. R.
667   Westra, S. P. Waghmare, B. Wiedenheft, U. Pul, R. Wurm, R. Wagner, M. R.
668   Beijer, A. Barendregt, K. Zhou, A. P. L. Snijders, M. J. Dickman, J. A. Doudna,
669   E. J. Boekema, A. J. R. Heck, J. van der Oost, S. J. J. Brouns, Ü. Pul, R.
670   Wurm, R. Wagner, M. R. Beijer, A. Barendregt, K. Zhou, A. P. L. Snijders, M.
671   J. Dickman, J. A. Doudna, E. J. Boekema, A. J. R. Heck, J. van der Oost, S. J.
672   J. Brouns, U. Pul, R. Wurm, R. Wagner, M. R. Beijer, A. Barendregt, K. Zhou,
673   A. P. L. Snijders, M. J. Dickman, J. A. Doudna, E. J. Boekema, A. J. R. Heck,
674   J. van der Oost, S. J. J. Brouns, Structural basis for CRISPR RNA-guided DNA
675   recognition by Cascade. *Nat. Struct. Mol. Biol.* **18**, 529–536 (2011).
676 20. Y. Tang, A. Mu, Y. Zhang, S. Zhou, W. Wang, Y. Lai, X. Zhou, F. Liu, X. Yang,
677   H. Gong, Q. Wang, Z. Rao, Cryo-EM structure of Mycobacterium smegmatis
678   DyP-loaded encapsulin. *Proc. Natl. Acad. Sci.* **118**, e2025658118 (2021).
679 21. M. Sutter, D. Boehringer, S. Gutmann, S. Günther, D. Prangishvili, M. J.
680   Loessner, K. O. Stetter, E. Weber-Ban, N. Ban, Structural basis of enzyme
681   encapsulation into a bacterial nanocompartment. *Nat. Struct. Mol. Biol.* **15**,
682   939–947 (2008).
683 22. X. Xiong, C. Sun, F. S. Vago, T. Klose, J. Zhu, W. Jiang, Cryo-EM structure of
684   heterologous protein complex loaded thermotoga maritima encapsulin capsid.
685   *Biomolecules*. **10**, 1–13 (2020).
686 23. R. M. Putri, C. Allende-Ballestero, D. Luque, R. Klem, K.-A. Rousou, A. Liu, C.
687   H.-H. Traulsen, W. F. Rurup, M. S. T. Koay, J. R. Castón, J. J. L. M.
688   Cornelissen, Structural Characterization of Native and Modified Encapsulins as
689   Nanoplatforms for *in Vitro* Catalysis and Cellular Uptake. *ACS Nano*. **11**,
690   12796–12804 (2017).
691 24. R. J. Nichols, B. LaFrance, N. R. Phillips, D. R. Radford, L. M. Oltrogge, L. E.
692   Valentin-Alvarado, A. J. Bischoff, E. Nogales, D. F. Savage, Discovery and
693   characterization of a novel family of prokaryotic nanocompartments involved in
694   sulfur metabolism. *Elife*. **10** (2021), doi:10.7554/eLife.59288.

695   25.   B. Lafrance, C. Cassidy-Amstutz, R. J. Nichols, L. M. Oltrogge, E. Nogales, D.
696   F. Savage, *bioRxiv*, in press, doi:10.1101/2021.04.26.441214.
697   26.   E. M. Williams, S. M. Jung, J. L. Coffman, S. Lutz, Pore Engineering for
698   Enhanced Mass Transport in Encapsulin Nanocompartments. *ACS Synth. Biol.*
699   **7**, 2514–2517 (2018).
700   27.   L. Adamson, N. Tasneem, M. P. Andreas, W. Close, T. N. Szyszka, E. Jenner,
701   R. Young, L. Chen Cheah, A. Norman, F. Sainsbury, T. W. Giessen, Y. Heng
702   Lau, *bioRxiv*, in press, doi:10.1101/2021.01.27.428512.
703   28.   B. LaFrance, C. Cassidy-Amstutz, R. J. Nichols, L. M. Oltrogge, E. Nogales, D.
704   F. Savage, *bioRxiv*, in press, doi:10.1101/2021.04.26.441214.
705   29.   C. Piergentili, J. Ross, D. He, K. J. Gallagher, W. A. Stanley, L. Adam, C. L.
706   Mackay, A. Baslé, K. J. Waldron, D. J. Clarke, J. Marles-Wright, Dissecting the
707   structural and functional roles of a putative metal entry site in encapsulated
708   ferritins. *J. Biol. Chem.* (2020), doi:10.1074/jbc.RA120.014502.
709   30.   S. V. Iverson, T. L. Haddock, J. Beal, D. M. Densmore, CIDAR MoClo:
710   Improved MoClo Assembly Standard and New E. coli Part Library Enables
711   Rapid Combinatorial Design for Synthetic and Traditional Biology. *ACS Synth.*
712   *Biol.* **5**, 151019092657002 (2015).
713   31.   F. Studier, Protein production by auto-induction in high-density shaking
714   cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
715   32.   J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T.
716   Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J.
717   White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-
718   source platform for biological-image analysis. *Nat. Methods*. **9**, 676–682
719   (2012).
720   33.   T. Nakane, E. Lindahl, J. Zivanov, W. J. J. H. Hagen, S. H. W. H. Scheres, D.
721   Kimanius, B. B. O. Forsberg, T. Nakane, B. B. O. Forsberg, D. Kimanius, W. J.
722   J. H. Hagen, E. Lindahl, S. H. W. H. Scheres, New tools for automated high-
723   resolution cryo-EM structure determination in RELION-3. *Elife*. **7**, 1–38 (2018).
724   34.   S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard,
725   MotionCor2: anisotropic correction of beam-induced motion for improved cryo-
726   electron microscopy. *Nat. Methods*. **14**, 331–332 (2017).
727   35.   A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation
728   from electron micrographs. *J. Struct. Biol.* **192**, 216–21 (2015).
729   36.   L. A. Kelly, S. Mezulis, C. Yates, M. Wass, M. Sternberg, The Phyre2 web
730   portal for protein modelling, prediction, and analysis. *Nat. Protoc.* **10**, 845–858
731   (2015).
732   37.   E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I.
733   Croll, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Structure visualization for
734   researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
735   38.   P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A.
736   Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and
737   crystallography. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 531–544 (2018).
738   39.   P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of
739   Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
740   40.   V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J.
741   Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: all-atom
742   structure validation for macromolecular crystallography. *Acta Crystallogr. D.*
743   *Biol. Crystallogr.* **66**, 12–21 (2010).
744   41.   P. V. Afonine, B. P. Klaholz, N. W. Moriarty, B. K. Poon, O. V. Sobolev, T. C.

Terwilliger, P. D. Adams, A. Urzhumtsev, IUCr, New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 814–840 (2018).

42. B. A. Barad, N. Echols, R. Y.-R. Wang, Y. Cheng, F. DiMaio, P. D. Adams, J. S. Fraser, EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat. Methods*. **12**, 943–6 (2015).

43. F. Sievers, D. G. Higgins, Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **1079**, 105–16 (2014).

44. P. Gouet, X. Robert, E. Courcelle, ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3 (2003).

45. G. R. Masson, J. E. Burke, N. G. Ahn, G. S. Anand, C. Borchers, S. Brier, G. M. Bou-Assaf, J. R. Engen, S. W. Englander, J. Faber, R. Garlish, P. R. Griffin, M. L. Gross, M. Guttman, Y. Hamuro, A. J. R. Heck, D. Houde, R. E. Iacob, T. J. D. Jørgensen, I. A. Kaltashov, J. P. Klinman, L. Konermann, P. Man, L. Mayne, B. D. Pascal, D. Reichmann, M. Skehel, J. Snijder, T. S. Strutzenberg, E. S. Underbakke, C. Wagner, T. E. Wales, B. T. Walters, D. D. Weis, D. J. Wilson, P. L. Wintrode, Z. Zhang, J. Zheng, D. C. Schriemer, K. D. Rand, Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods*. **16**, 595–602 (2019).

46. M. Laguerre, M. Saux, J. P. Dubost, A. Carpy, in *Pharmaceutical Sciences* (John Wiley & Sons, Ltd, 1997), vol. 3, pp. 217–222.

Funding

Acknowledgements

Author Contributions

Conceptualization: JR, JMW, DJC
Methodology: JR, ZM, TL, CP, AMBL, JMW, DJC
Validation: JR, ZM, TL, JMW, DJC
Formal analysis: JR, ZM, TL, JMW, DJC
Investigation: JR, ZM, KG, JEB, CP, EZA
Resources: KG, FC, CLM, JEB, AMBL, EZA, MDW
Data curation: JR, TL, DJC, AMBL, JMW
Writing – original draft preparation: JR, DJC, JMW
Writing – review and editing: JR, ZM, CP, AMBL, MDW, KJW, JMW, DJC
Visualization: JR, ZM, TL, DJC, JMW
Supervision: JMW, DJC
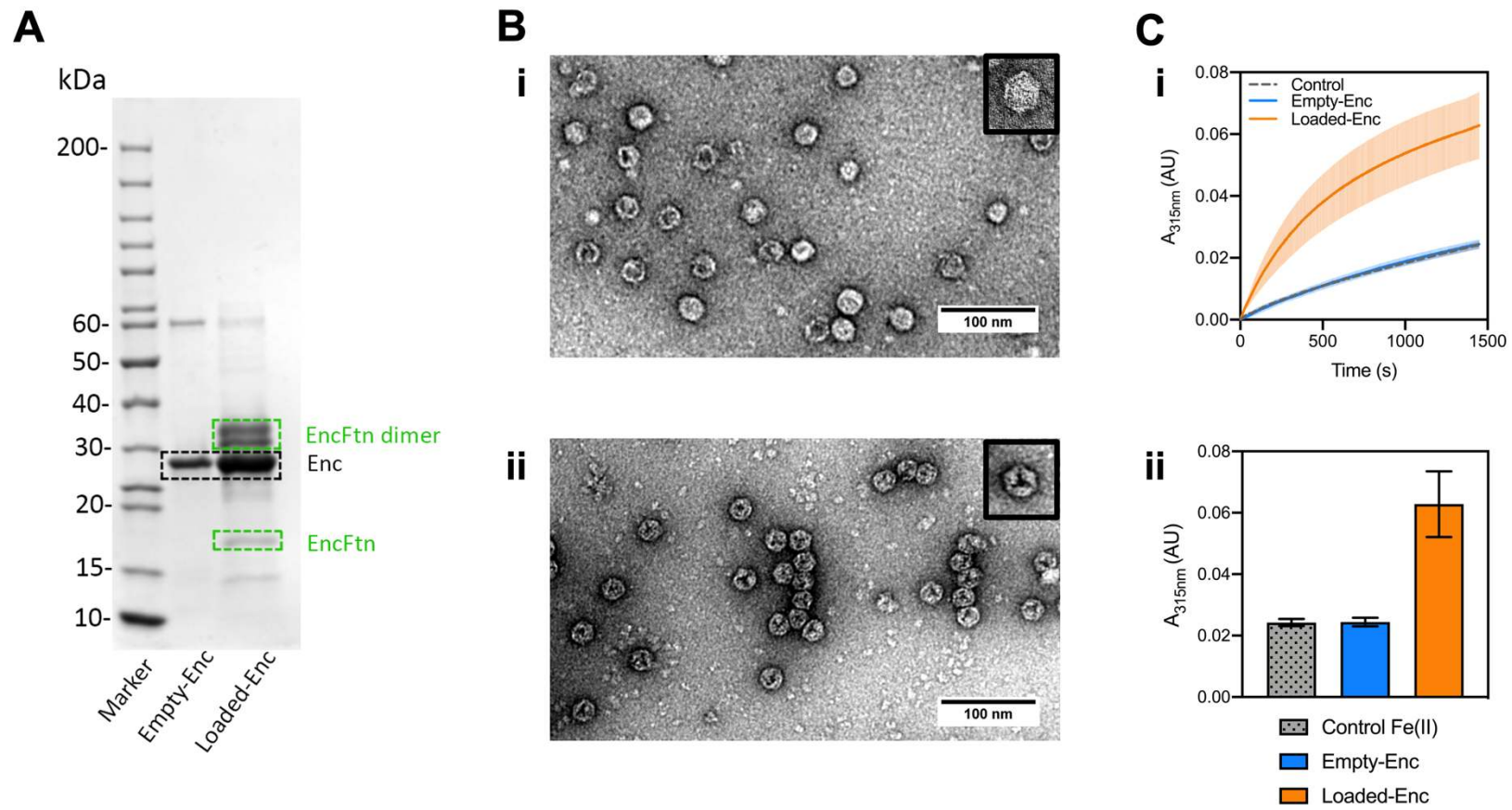Funding acquisition: MDW, LEH, KJW, DJC, JMW

Figures

Figure 1. Validation of the Assembly and Activity of Loaded-Enc and Empty-Enc.

817   **A**: SDS-PAGE of purified Empty-Enc and Loaded-Enc. Proteins resolved by 15% acrylamide SDS-PAGE and stained with
818   Coomassie blue stain. Encapsulin bands are near the 30 kDa marker and highlighted by a black dashed box. The EncFtn cargo of
819   Loaded-Enc appears as both a monomer and dimer. The purified EncFtn monomer is highlighted by a box labelled 'EncFtn'. Two
820   overlapping bands at approximately 35 kDa are highlighted by a green dashed box and labelled as 'EncFtn dimer'. **B**: Negative
821   stain transmission micrographs of Empty-Enc (**i**) and Loaded-Enc (**ii**) displaying individual particles for each complex. One
822   nanocompartment of Empty-Enc and Loaded-Enc is shown in the upper right corner of each micrograph, with a hexagonal 2D
823   geometry observed. **C: i**: Ferroxidase activity of Loaded-Enc compared to Empty-Enc. Protein samples were mixed with 100 µM
824     $FeSO_4 . _2 7H_2O$. Following an incubation period at room temperature of 50 seconds absorbance at 315 nm was measured over a
825   time-course of 1450 seconds. Control reference established using enzyme-free reaction as a measure of background iron
826   oxidation. Lines represent the mean of three technical repeats, error bars represent the standard deviation from the mean. **ii**: End
827   point ferroxidase assay comparison. Ferroxidase activity shown by the total increase in $A_{315\ nm}$ at the end point of the assay. Bars
828   represent the mean of three technical repeats, error bars represent the standard deviation from the mean.
829
830
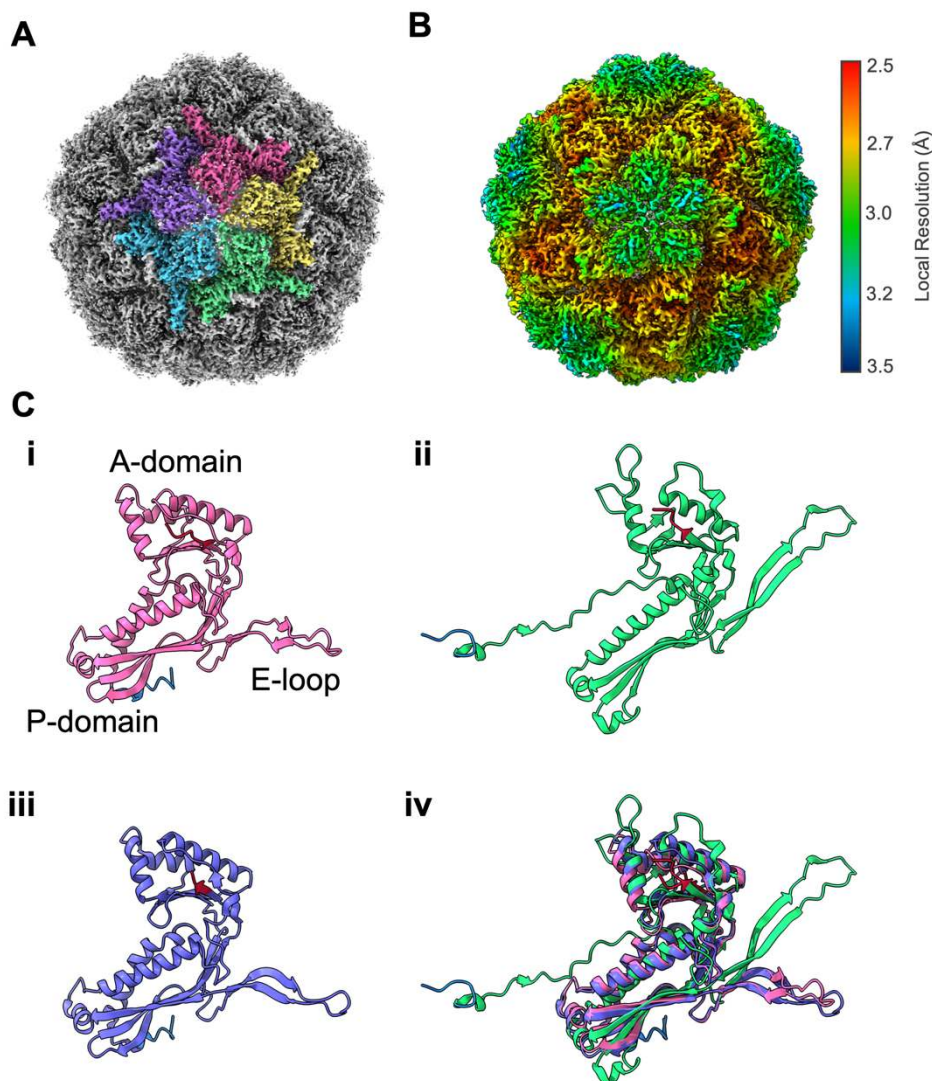
831
832
Figure 2: Architecture of the *H. ochraceum* encapsulin nanocompartment shell
Visualization of the electronic potential map of the *H. ochraceum* encapsulin from an icosahedrally averaged single particle reconstruction. **A**: The exterior of the encapsulin shell visualized at 2.4 Å resolution. Five subunits of the encapsulin nanocompartment shell have been colored to highlight the 5-fold axis. **B:** Icosahedral EM map of Loaded-Enc sharpened by local resolution estimate and colored by local resolution. The estimated resolution varies across the exterior of the encapsulin nanocompartment with the lowest resolution at the 5-fold pores. Color key of resolution mapping is shown on the right-hand side of the figure. **C**: The shared phage-like fold in the HK97 bacteriophage capsid and encapsulin proteins. Monomeric subunits of the *H. ochraceum* encapsulin protein modelled from our reconstruction are shown (pink, **C i**), with comparisons to the HK97 bacteriophage Head II *T*=7 monomer (green, **C ii**, PDB: 2FT1), and the *T. maritima T*=1 monomer (purple, **C iii**, PDB: 3DKT). The N-terminus of each monomer is highlighted in blue and the C-termius in red **C iv**: Comparison of the structures of **i** (pink), **ii** (green) and **iii** (purple) showing similar A- and P-domains. A noticeable difference between the monomers is that the E-loop of **ii** is shifted away from the others, which is typical of icosahedral encapsulin structures with a *T*-number greater than 1**.**

851
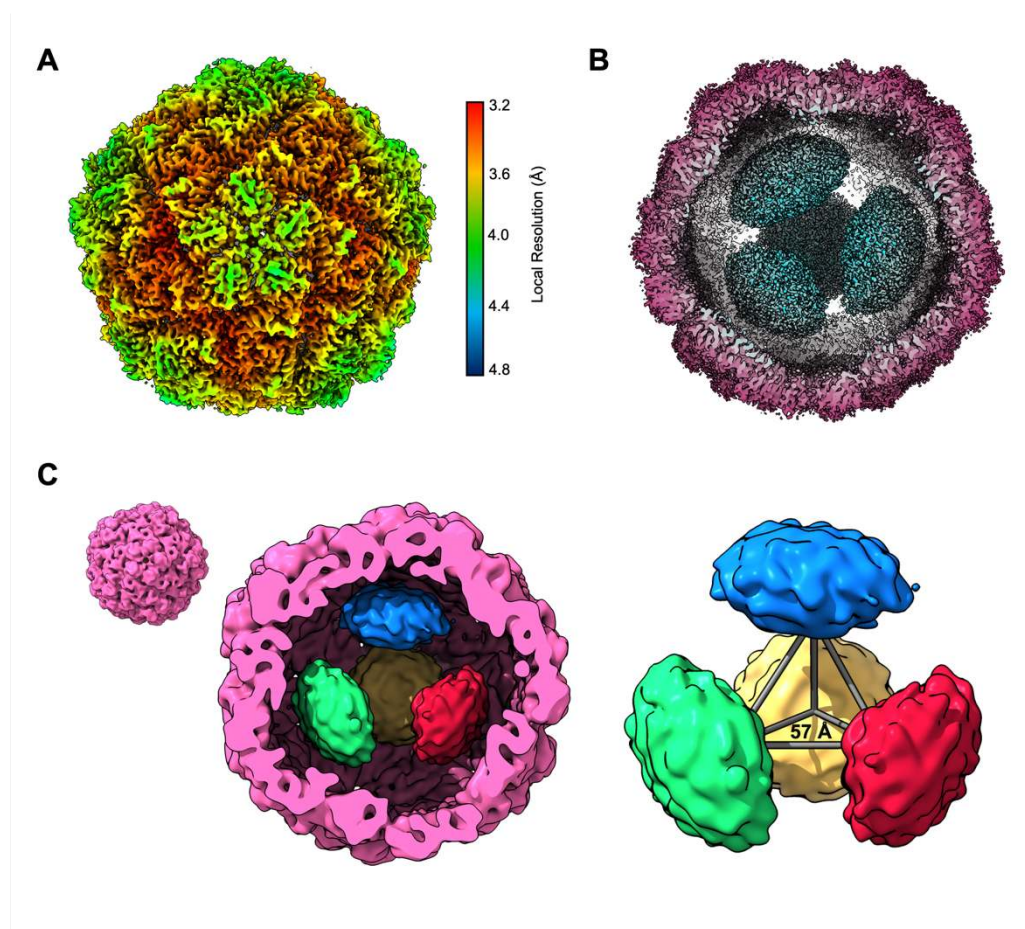
852
853
854
855 **Figure 3: Asymmetric reconstruction of the *H. ochraceum* encapsulin complex**
856 **reveals a tetrahedral arrangement of EncFtn within the encapsulin**
857 **nanocompartment.**
858 **A**: Electronic potential map of the asymmetric reconstruction of the *H. ochraceum*
859 Enc:EncFtn complex. The map is colored by local resolution with the color key for
860 the shown on the right side of **A**. **B:** Radially colored cryo-EM derived map of the
861 Loaded-Enc nanocompartment displaying the interior EncFtn (cyan). **C**: Gaussian
862 smoothed C1 map showing the four discrete EncFtn densities (red, green, yellow
863 and blue), consistent with the size of a decameric EncFtn complexes, within the
864 encapsulin nanocompartment (pink). The four EncFtn are in a tetrahedral
865 arrangement highlighted by grey lines connecting their centers of mass. The average
866 distance between each EncFtn decamer is 57 Å.
867
868

869
870

871
872
873
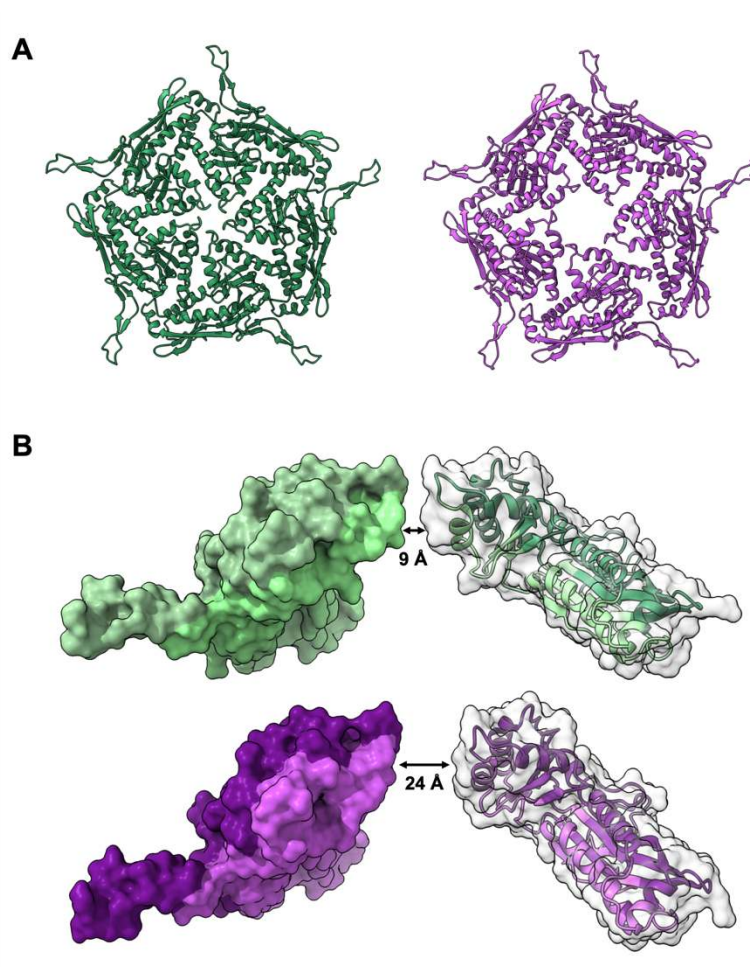
874 **Figure 4: Conformations of the dynamic 5-fold pore of the *H. ochraceum***
875 **encapsulin shell.**
876 Masked 3D refinements centered around the 5-fold pore of the icosahedral
877 reconstruction were performed after symmetry expansion of the asymmetric units.
878 **A**: Cartoon representations of the 'closed' (green) and 'open' (purple) conformations
879 of the Enc shell pentamer. The closed conformation has a 5-fold pore diameter of 9
880 Å whilst the equivalent diameter of the open conformation is 24 Å. **B**: Side-on view of
881 the 5-fold pore highlighting its increased diameter in the open conformation (purple)
882 compared to the closed conformation (green). The interior face of encapsulin for the
883 closed and open conformations is shown in light green and light purple respectively.
884
885
886
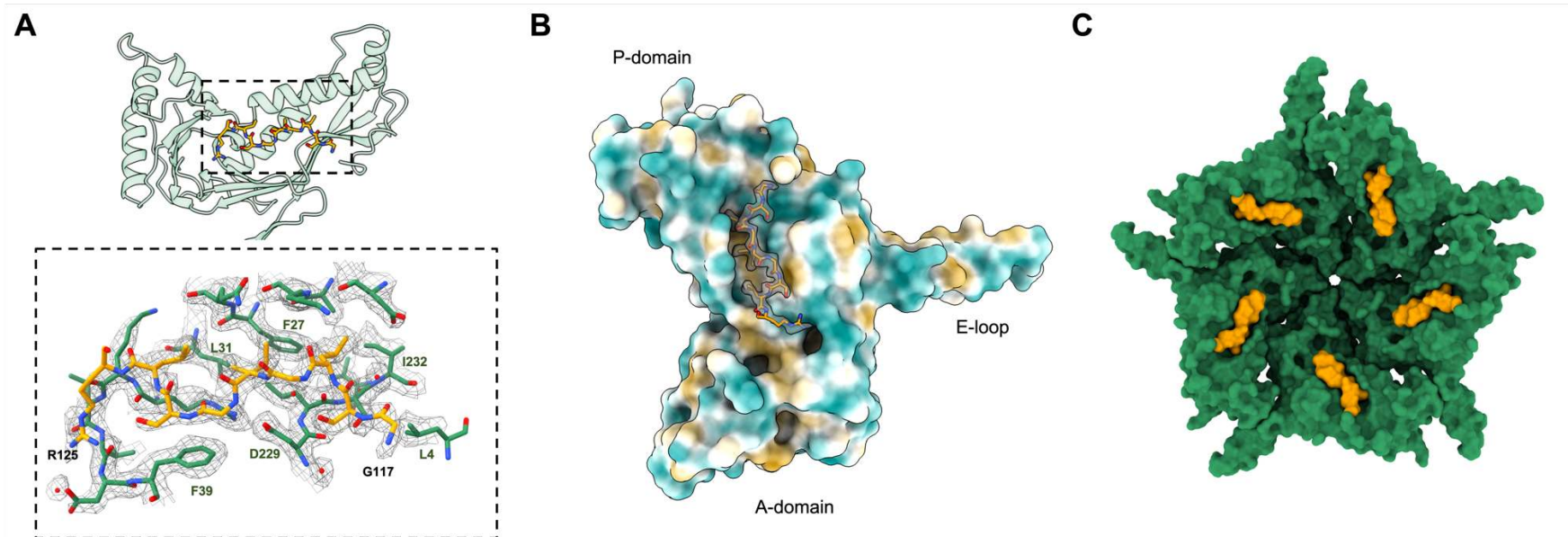
887
888
889 **Figure 5: The closed conformation of the _H. ochraceum_ five-fold pore allows docking and sequencing of the EncFtn**
890 **localization sequence to the interior face of the encapsulin monomer.**
891 **A**: Binding of localization sequence residues from the EncFtn to the interior wall of the encapsulin nanocompartment monomer
892 (green residues and transparent cartoon) and the localization sequence (yellow sticks). Lower panel: Hydrophobic residues from
893 the interior face of the encapsulin form the binding pocket for the localization sequence. The first and last residues modelled for the
894 LS have been labelled in black, and key residues from encapsulin have been labelled in green. Modelled water molecules are
895 shown as red spheres. **B:** The spatial relationship between an encapsulin monomer and the localization sequence (shown in gold
896 with its EM map density as in **A**). The Enc monomer has been colored by molecular lipophilicity potential (_46_) which ranges from
897 dark cyan (corresponding to the most hydrophilic) to white to dark gold (most lipophilic). This highlights the hydrophobic pocket on
898 the interior of the encapsulin nanocompartment where the localization sequenced binds. **C**: The 'closed' conformation pentamer
899 (green) with a localization sequence (yellow) shown on each monomer.
900
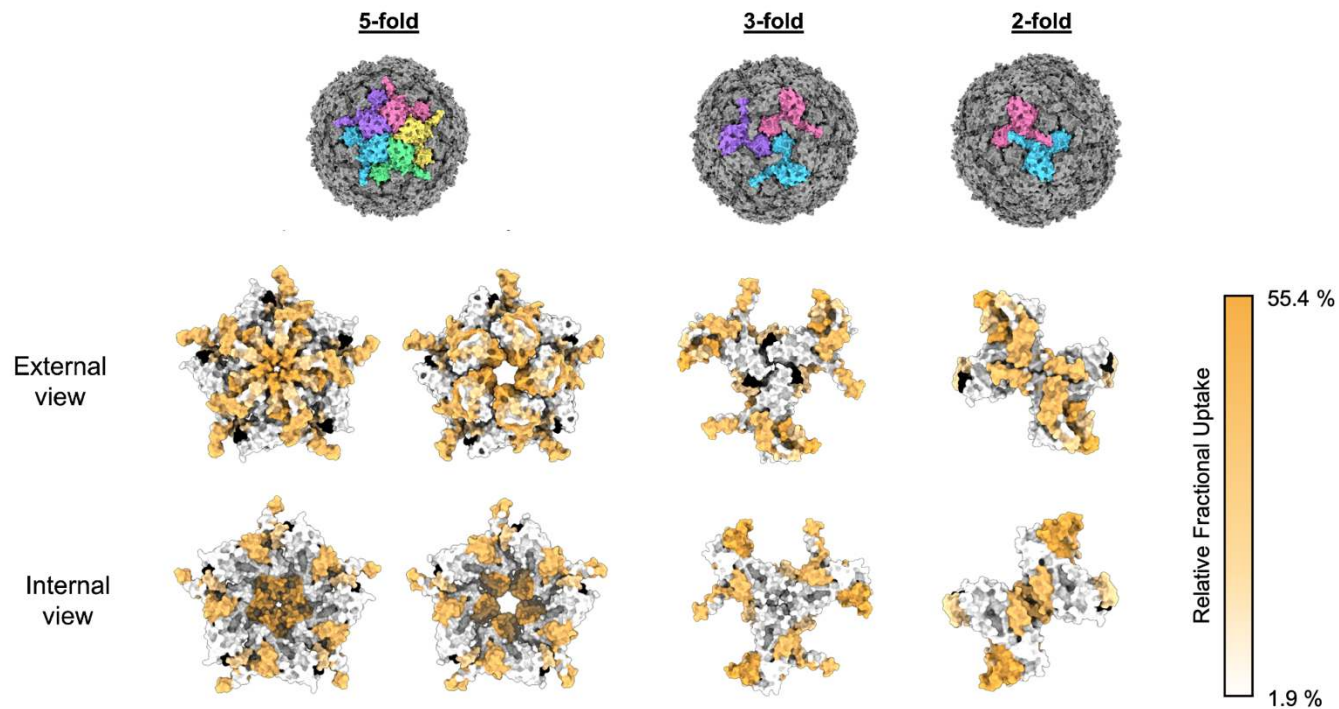
901
**Figure 6: HDX-MS fractional uptake mapped to the symmetry axes of the icosahedral reconstruction of the *H. ochraceum*
903 encapsulin complex.**
904 From top to bottom: The top row depicts molecular surface models of the *H. ochraceum* encapsulin in three orientations to show
905 the 5-fold, 3-fold and 2-fold symmetry axes. Monomers have been colored individually (pink, yellow, green, blue and purple) to
906 highlight the symmetry axes. The final two rows display the relative fractional uptake of each peptide from HDX experiments
907 displayed on the 5-fold, 3-fold and 2-fold pores of the encapsulin nanocompartment. Peptide fractional uptake is show on a white-
908 to-orange color scale with a color key show on the right-hand side of the figure. Areas colored black correspond to no peptide
909 coverage.

Supplementary Information: Pore dynamics and asymmetric cargo loading in an encapsulin nanocompartment revealed by Cryo-EM and HDX mass spectrometry

Supplementary Figures

Supplementary Tables

Figure S1: Purification of recombinant *H. ochraceum* encapsulin protein complexes

A: i. Size exclusion chromatogram traces of Empty-Enc with normalised A_{280} values. Empty-Enc (●) eluted at a retention volume of 65 ml. Other peaks represent contaminants and encapsulin monomers. **ii.** SDS-PAGE of Empty-Enc samples at different stages of purification. Proteins were resolved by 15% acrylamide SDS-PAGE and stained with Coomassie blue stain. The encapsulin band is near the 30 kDa marker of the PageRuler Unstained Protein Ladder (highlighted with a dashed black box). **iii**: Mass spectrum of Empty-Enc displaying one charge state distribution (highlighted with purple circles), which corresponds to the mass of the encapsulin monomer (28969.7 Da). **iv:** Histograms of the size distribution of Empty-Enc from negative stain TEM. A Gaussian curve was fitted to the data by nonlinear least squares regression, showing that individual Empty-Enc nanocompartments have a mean diameter of 21.13 nm with a standard deviation of

1.80 nm. **B: i.** Size exclusion chromatogram trace of Loaded-Enc. Loaded-Enc elutes at 65 ml, the peak is stressed with an orange diamond (◆) Non-encapsulated EncFtn proteins are eluted at 80ml, exclusively in the Loaded-Enc SEC run. **ii.** SDS-PAGE of loaded-Enc at different stages of the purification. Proteins were resolved by 15% acrylamide SDS-PAGE and stained with Coomassie blue stain. Loaded-Enc encapsulin band is near 30 kDa (highlighted with a dashed black box) and the EncFtn bands are highlighted by dashed green boxes. EncFtn exists as both monomer and dimer bands, as is typical for these proteins. **iii**: Mass spectrum of proteins present in Loaded-Enc are highlighted by coloured circles; encapsulin nanocompartment protein in purple (28813.2 Da), EncFtn monomer in blue (14667.4 Da), and EncFtn dimer in green (29334.5 Da). **iv:** Histogram of the size distribution of Loaded-Enc as observed by negative stained TEM. Nanocompartments of Loaded-Enc were found to have an average diameter of 21.42 nm with a standard deviation of 1.71 nm.
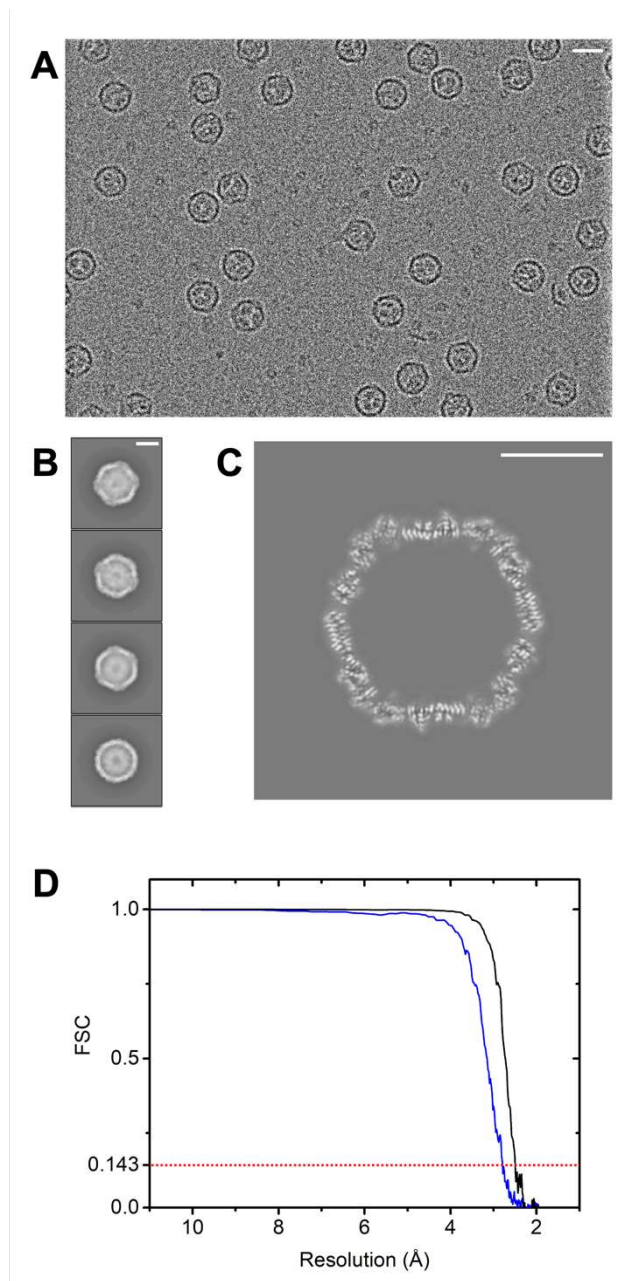
Figure S2: Representative cryo-EM micrograph and additional data from the icosahedral reconstruction of the *H. ochraceum* encapsulin complex
A: Representative cryo-EM micrograph of Loaded-Enc. **B**: 2D classes of Loaded-Enc **C**: Central slice of Loaded-Enc from icosahedral processing. A white scale bar representing 10 nm is shown in the upper right corner of **A**, **B** and **C**. **D**: Gold standard FSC curve of the masked (black) and unmasked (blue) icosahedral reconstruction half maps. The FSC 0.143 threshold is highlighted with a dashed red line.
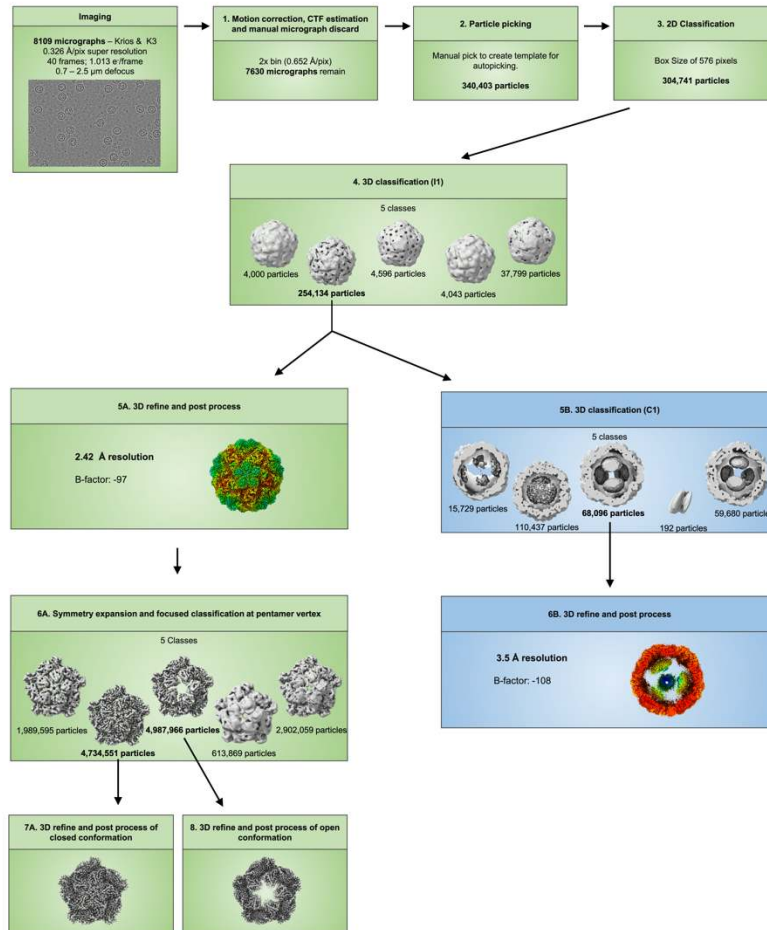
Figure S3: Cryo-EM processing workflow
The processing pipeline within Relion 3.1 used to obtain the single particle
reconstructions of Loaded-Enc. The green boxes show the processing workflow for the
initial processing steps and I1 reconstruction. The blue boxes show the diverge to C1
processing to gain insight into the EncFtn loading inside Enc.
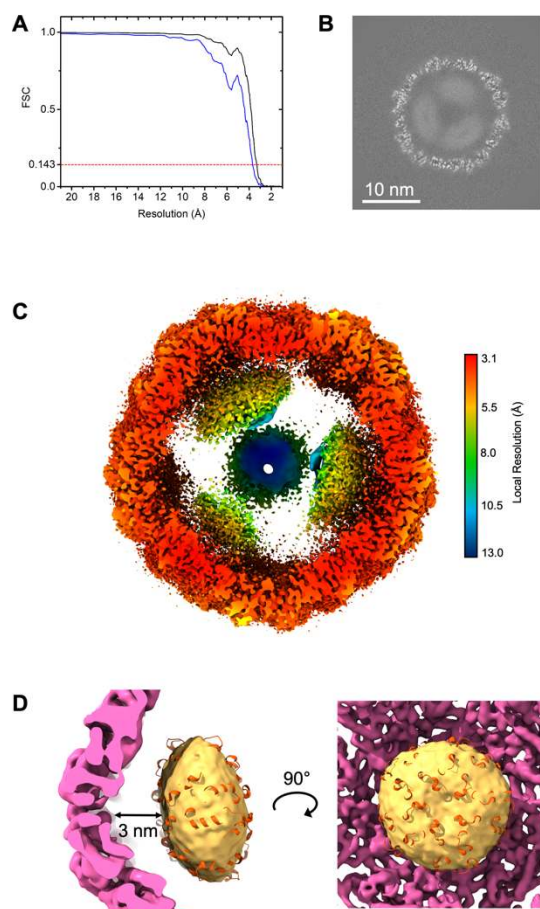
Figure S4: Supplementary cryo-EM data of Loaded-Enc C1 reconstruction
A: Gold standard FSC curve of the masked (black) and unmasked (blue) icosahedral reconstruction half maps. The FSC 0.143 threshold is highlighted with a dashed red line. **B**: Central slice of Loaded-Enc from C1 reconstruction after 3D refinement. A white scale bar representing 10 nm is shown in the lower left corner. **C:** Slice through of the locally filtered and sharped EM map of Loaded-Enc is shown colored by local resolution. The interior of the nanocompartment holds four EncFtn which are of significantly lower resolution than the shell. The color key for the local resolution is shown on the right side of the figure. **D:** The crystal structure of the EncFtn from *H. ochraceum* (orange, PDB 5N5F) docked into the interior density (yellow) of the C1 reconstruction. A 3 nm gap is observed between the Enc shell (pink) and the EncFtn (yellow and orange).
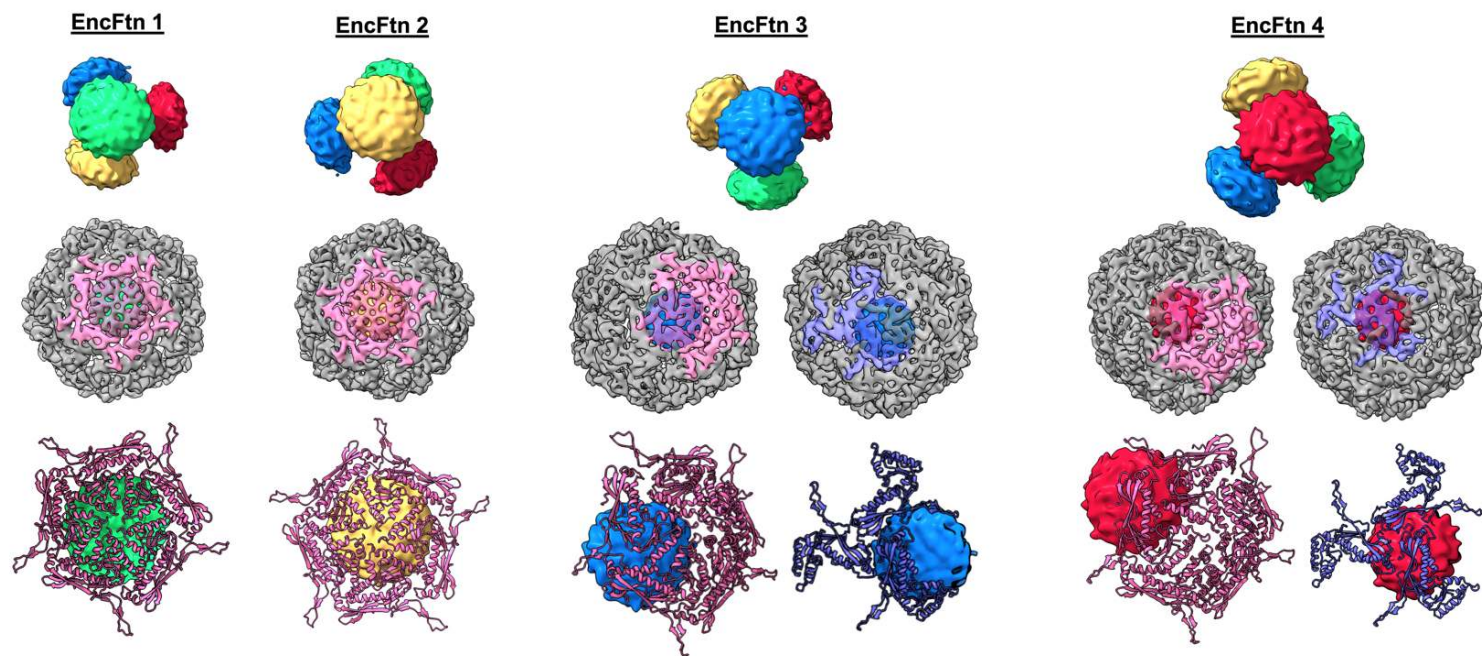
Figure S5: Visualization of the distinct EncFtn environments within the icosahedral encapsulin nanocompartment
Top panels: Each EncFtn within the encapsulin nanocompartment has been individually colored (as in **Figure 4 C**) and numbered. Four orientations of EncFtn tetrahedral are shown with a different EncFtn in the foreground. **Middle panels:** Each EncFtn of the Loaded-Enc complex as viewed from outside of the encapsulin nanocompartment. EncFtn complexes are in the same orientation as the top panels. A pentamer of the encapsulin nanocompartment has been colored pink to allow direct correlation between the EncFtn location and the 5-fold pore of encapsulin. EncFtn 3 and EncFtn 4 do not align with the 5-fold pore and so the encapsulin 3-fold pore has also been shown and colored purple. **Bottom panels:** The relationship between each EncFtn and the pores of encapsulin. The encapsulin 5-fold pore is shown by pink cartoons, the encapsulin 3-fold pore by purple cartoons and the EncFtn are colored as in the top panels. EncFtn 1 and EncFtn 2 are in broadly equivalent environments which are aligned with the 5-fold pore of the encapsulin nanocompartment.
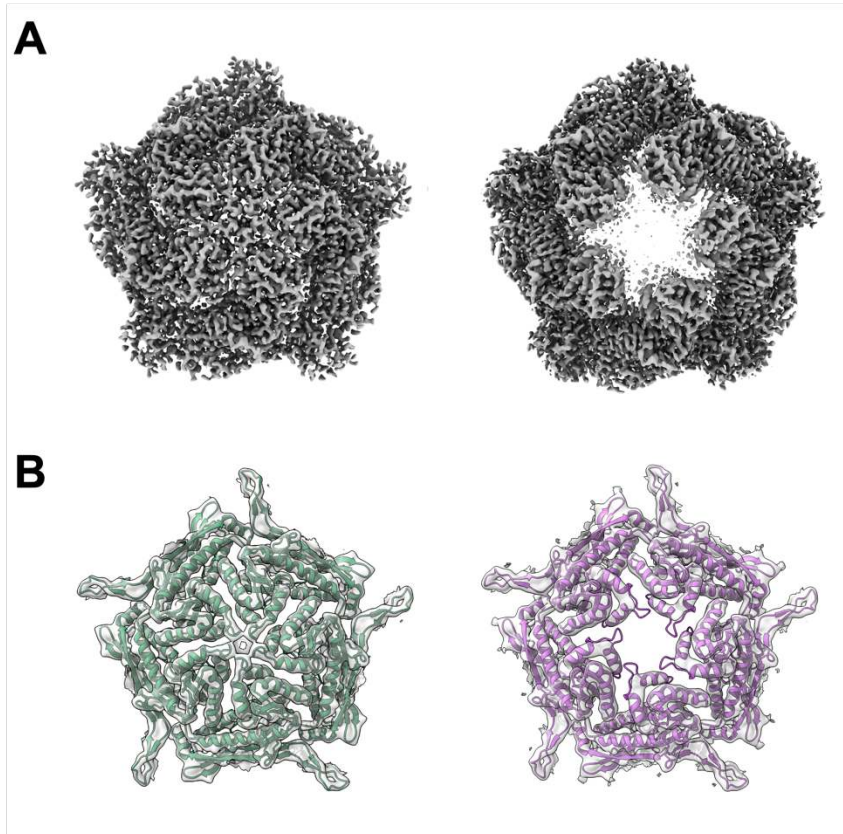
Figure S6: Electronic potential maps and models from masked 3D-classification and refinement centered on the five-fold symmetry axis of the symmetry expanded icosahedral reconstruction
A: Cryo-EM maps of the 'closed' (left) and 'open' (right) pentamer conformations from symmetry expansion of the icosahedral five-fold axis. **B**: The Gaussian smoothed cryo-EM maps with docked models of the 'closed' (green, left) and 'open' (purple, right) conformations. Smoothed EM maps allow easy visualization of the docked secondary structure.
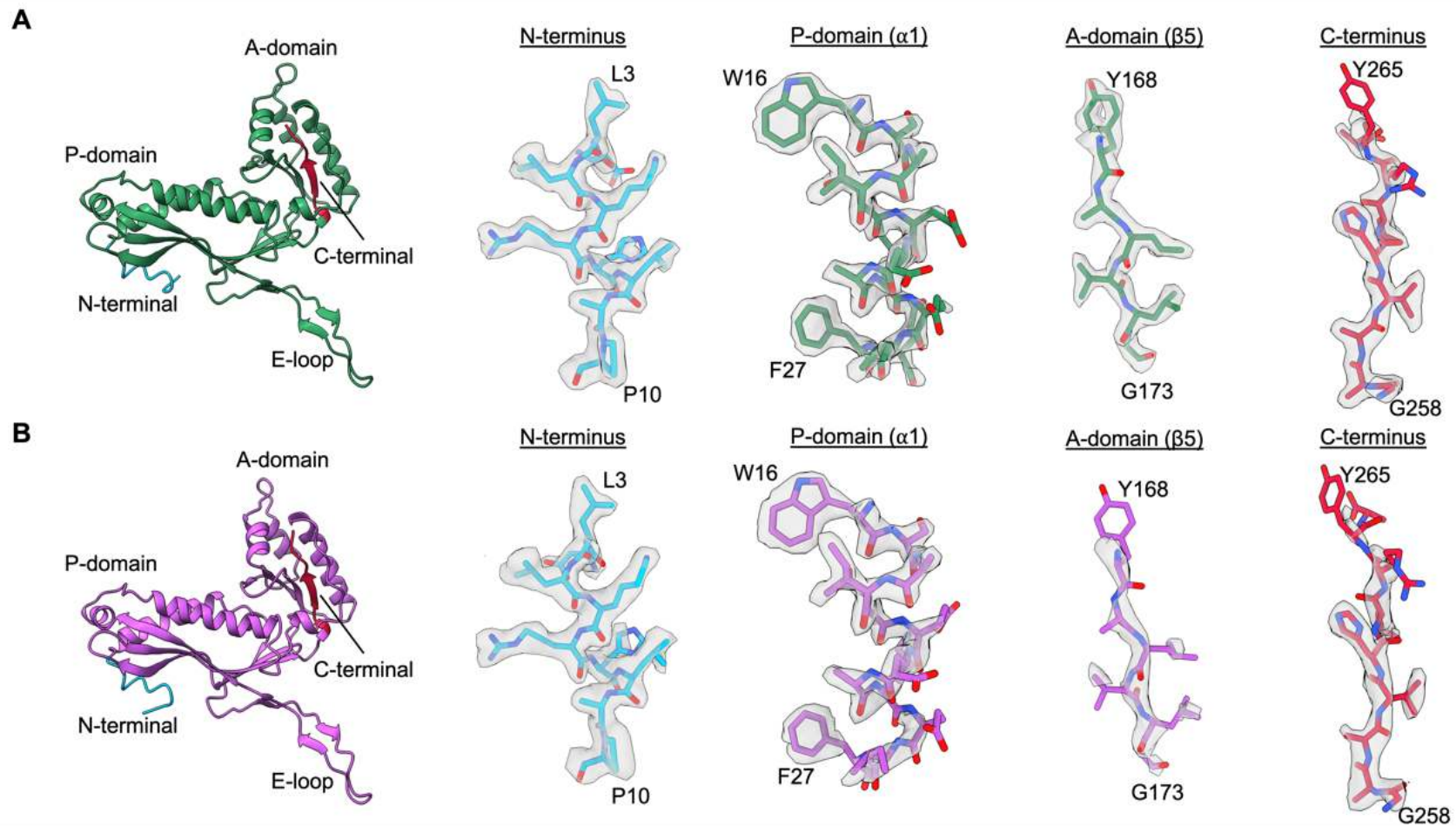
Figure S7: Model fit of the 'open' and 'closed' atomic models produced from open and closed maps
The atomic models of the monomer subunit of the 'closed' (green, **A**) and 'open' (purple, **B**) pentamer conformations. Representative maps are shown across the monomer chain illustrating the fit of the atomic model and side chains into the cryo-EM map.
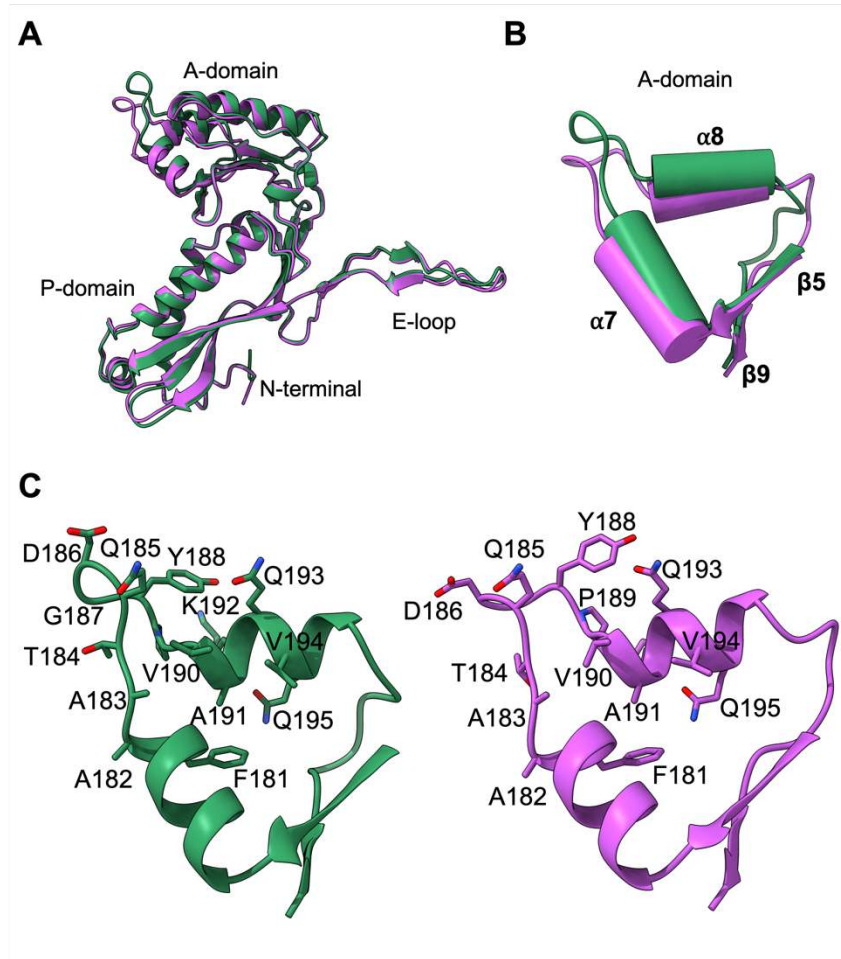
Figure S8: Comparison of the A-domain of 'closed' and 'open' pentamer structures

A: Overlay of the monomers from the 'closed' (green) and 'open' (purple) pentamer conformations displaying shifted A-domains. **B**: Overlay of the A-domain of the 'closed' (green) and 'open' (purple) conformations highlighting the change in alpha helices 7 and 8. **C**: The residues present at the 5-fold pore and in alpha helices 7 and 8.
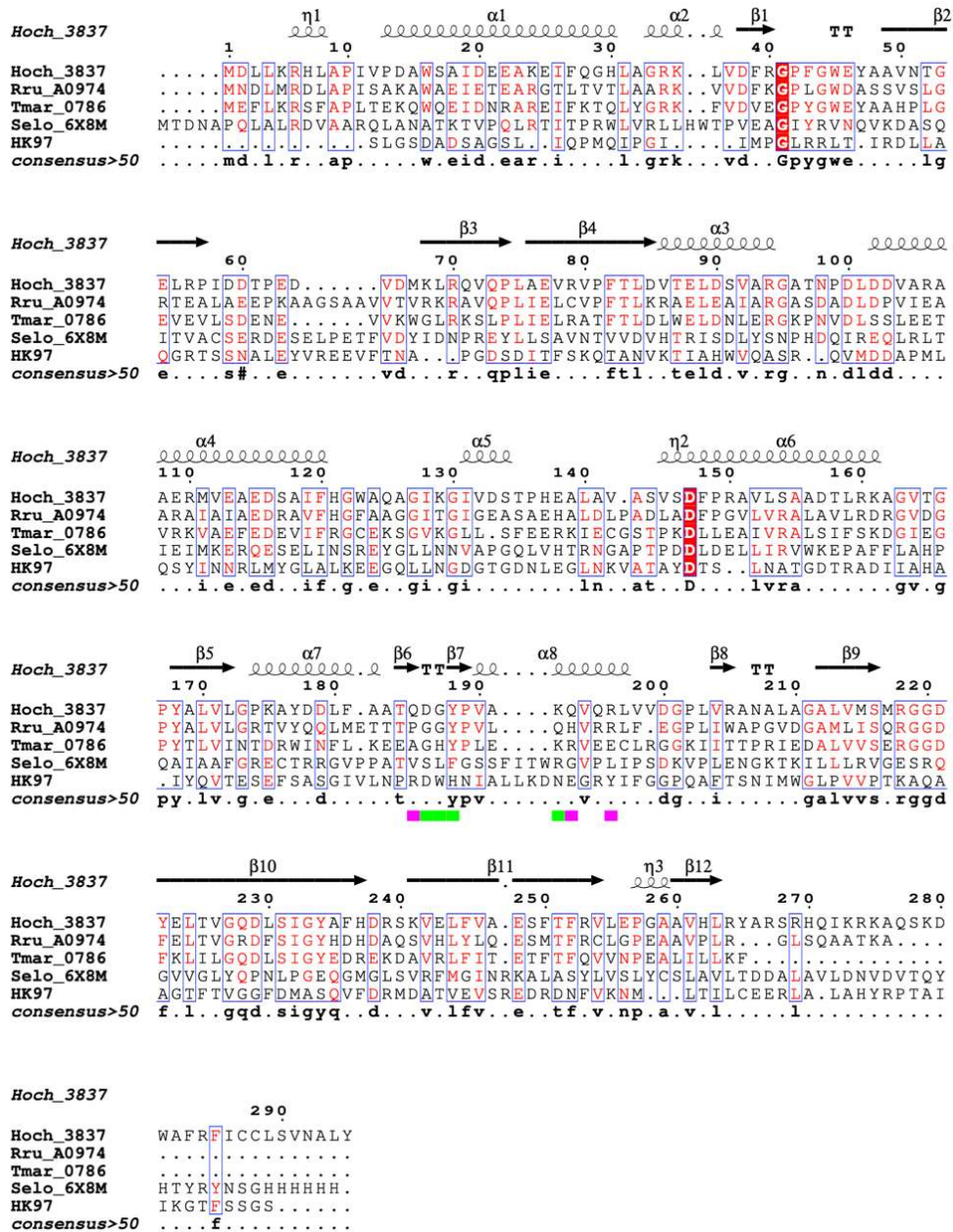
Figure S9: Sequence alignment of encapsulins and HK97

Protein sequence alignment of encapsins from the *H. ochraceum* (Hoch_3837), *R. rubrum* (Rru_A0974), *T. maritima* (Tmar_0786) and *S. elongatus* (Selo_6X8M). Encapsulins share the HK97-fold from the HK97 bacteriophage capsid which is also shown in the sequence alignment (HK97). The residues in the 5-fold pore in the 'closed' pentamer conformation are underlined in purple. The additional residues which become exposed and form the 5-fold pore in the 'open' conformation are underlined in green. Protein sequences were sourced from uniprot and the alignment was performed with Clustal Omega, and then formatted using ESPript.
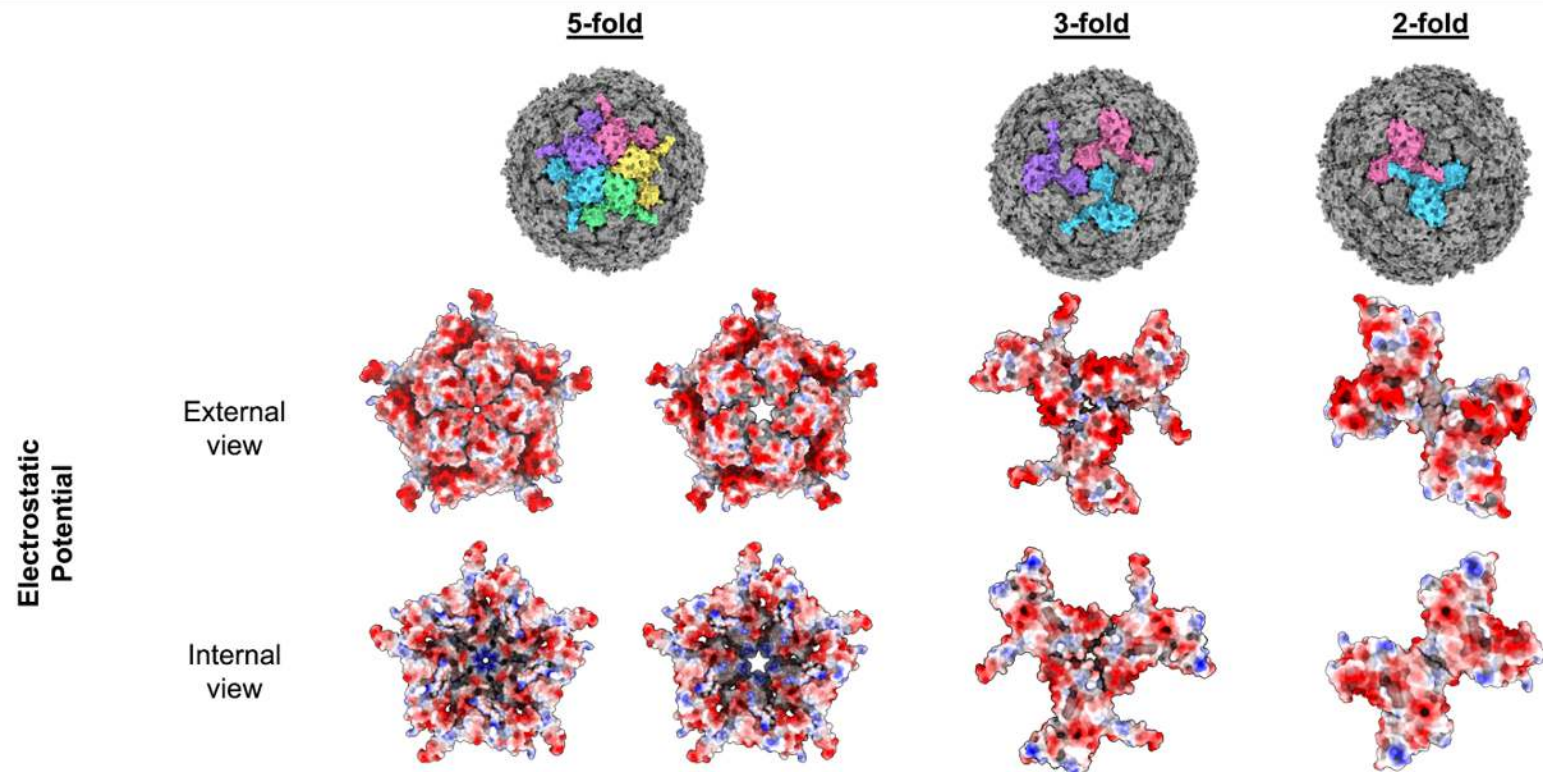
Figure S10: Electrostatic properties of the encapsulin nanocompartment pores
From top to bottom: The top row shows a T=1 encapsulin in three orientations to show the 5-fold, 3-fold and 2-fold symmetry axes. Monomers have been coloured individually (pink, yellow, green, blue and purple) to highlight the symmetry axes. The second and third row show the 5-fold (closed and open conformations), 3-fold and 2-fold Loaded-Enc symmetry axes coloured with electrostatic surfaces (positive charges shown in red and negative in blue).
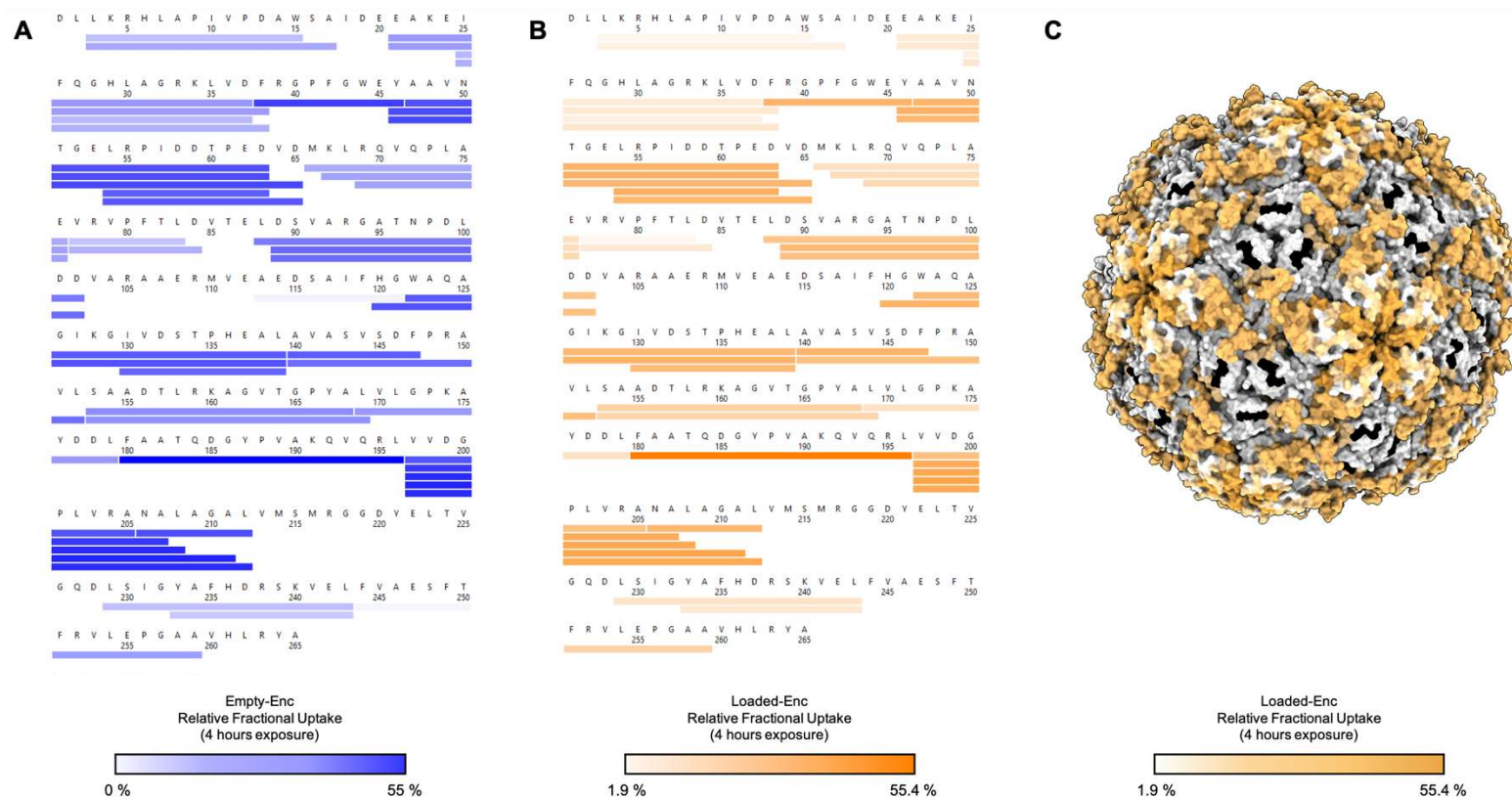
Figure S11: Sequence coverage and deuterium uptake after HDX-MS analysis of Empty-Enc and Loaded-Enc peptides

HDX coverage maps showing the relative fractional deuterium uptake for both Empty-Enc (**A**) and Loaded-Enc (**B**) after 4 hours of deuterium labelling. There are 40 observed peptides common to both states, providing 84 % protein sequence coverage with 2.35 redundancy. Color keys are shown under **A** and **B**. (**C**) The relative fractional uptake of Loaded-Enc displayed on the encapsulin nanocompartment. Areas colored black are representative of no sequence coverage
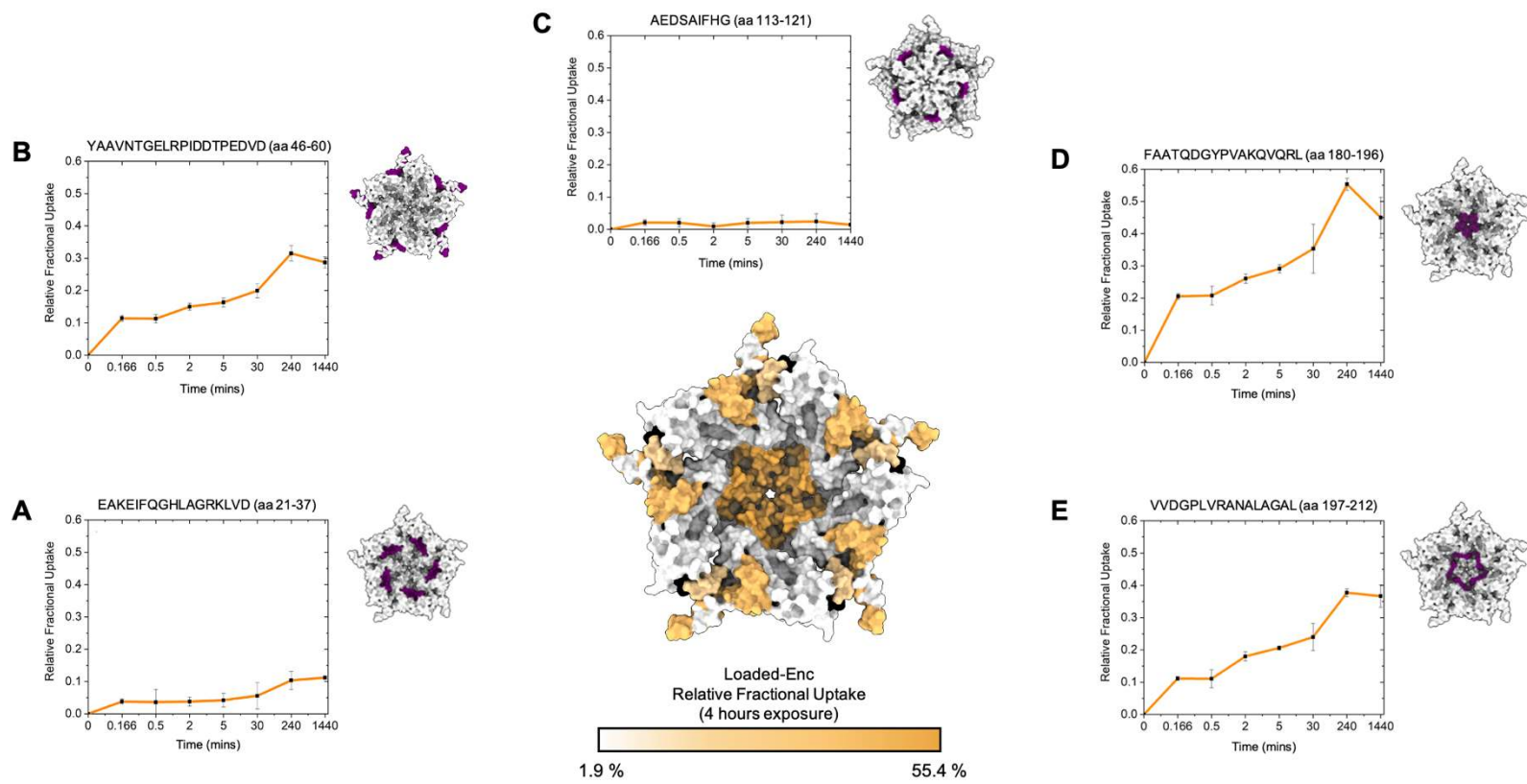
Figure S12: Deuterium fractional uptake of Loaded-Enc peptides
HDX-MS of the encapsulin 5-fold pore highlighting the amount of deuterium incorporation after 4 hours of exposure, colored according to uptake (a color key is shown at the bottom of the central figure). (**A**) to (**E**), Uptake plots for individual peptides showing the relative deuterium uptake over time. An encapsulin pentamer is shown with the corresponding peptide sequence colored purple. (**A**) residues 21-37, whose proximity is close to the proposed localization sequence for EncFtn (11.2 ± 0.2 % uptake at 4 hours). (**B**) residues 46-60, which overlays with the proposed 2-fold symmetry pore (31.7 ± 2.2 % uptake at 4 hours), (**C**) residues 113-121, a highly protected exterior peptide (0.2 ± 0 % at 4 hours), (**D**) residues 180-196 the potential 5-fold symmetry pore (55.4 ± 2.2 % at 4 hours), (**E**) interior residue 197-212 in close proximity to the potential 5-fold symmetry pore site (37.7 ± 1.2 % at 4 hours).
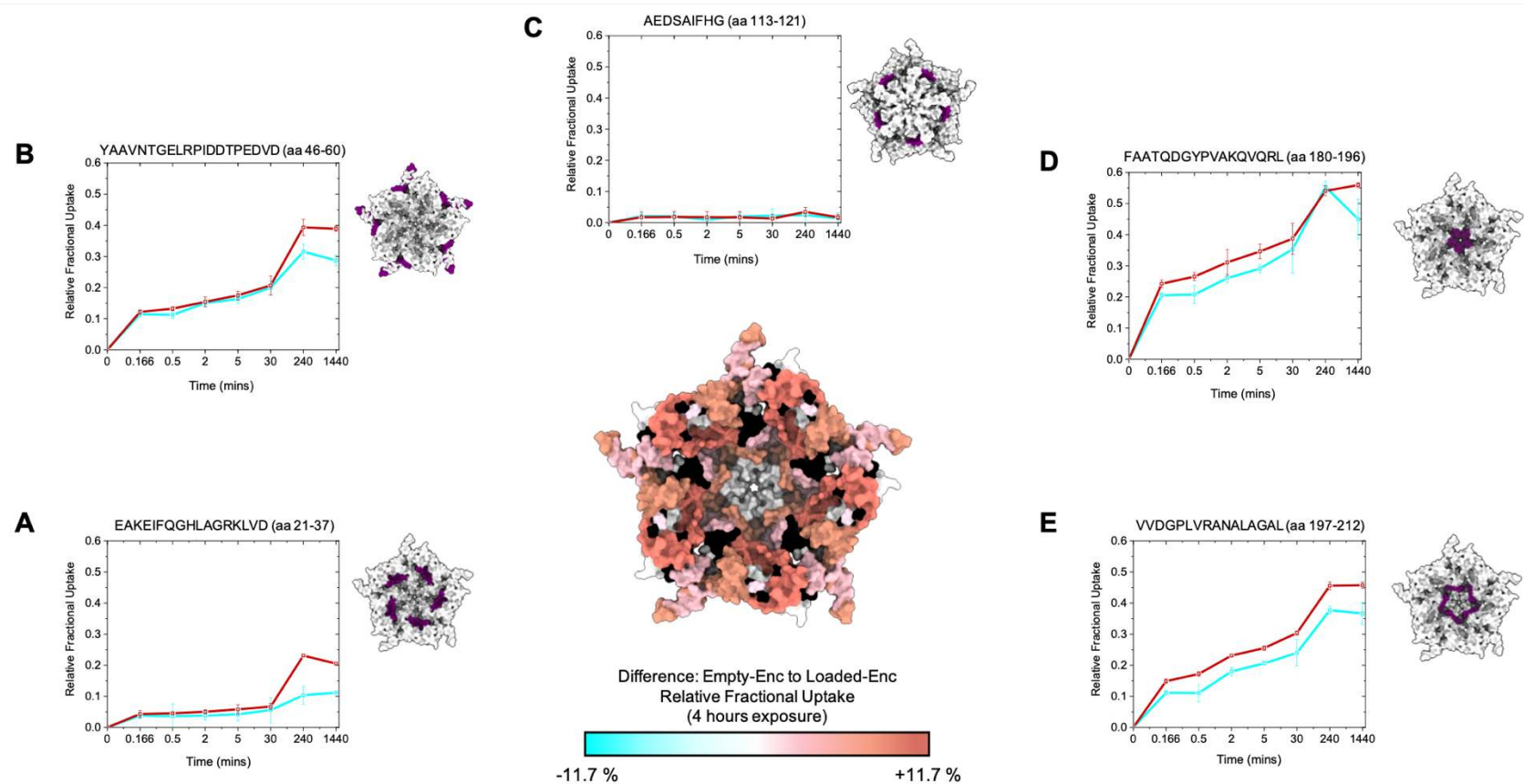
Figure S13: Differential HDX-MS analysis of Empty-Enc and Loaded-Enc.
The amount of deuterium incorporation over time is shown for five representative peptides (**A** - **E**) for both Empty-Enc (r lines) and Loaded-Enc (cyan lines). An encapsulin pentamer is shown next to each uptake chart with the corresponding peptide colored in purple. The central figure shows the difference in exchange is displayed on the structure of the 5-fold pore of encapsulin, with a color key below and areas of no sequence coverage colored black. (**A**) Location of residues 21-37 difference of +11.7 ± 1.5 %, (**B**) residues 46-65 with a difference of +7.6 ± 1.8 % (**C**) Location of residues 113-121 with a difference of 0.1 ± 0 %, (**D**) residues 180-196 with a difference of 0 ± 0.3 %, (**E**) interior residue 197-212 with a difference of 7.9 ± 1.7 %.

Table S1: Protein masses and assignments obtained by LC-MS

Sample	Protein in Sample	Observed Average Mass (Da)	Assignment	Theoretical Average Mass (Da)
Empty-Enc	Encapsulin from *H. ochraceum*	28969.67 ± 0.23	Hoch-Enc monomer	28968.84*
Loaded-Enc	Encapsulin from *H. ochraceum*	28813.24 ± 0.89	Hoch-Enc monomer	28812.66*
	EncFtn from *H. ochraceum*	14667.42 ± 0.22	Hoch-EncFtn monomer without Met	14668.20
		29334.50 ± 0.64	Hoch-EncFtn dimer without Met	29336.40

The starting methionine residue is not always retained, and this has been indicated in "Assignment". Errors generated by MassLynx v4.1.
* The difference in theoretical mass for the two *H. ochraceum* encapsulin proteins is due to an additional C-terminal arginine residue in the Empty-Enc construct which is a cloning artefact from the MoClo kit used.

Table S2: Cryo-EM data collection and processing

	Loaded-Enc Icosahedral Reconstruction	Loaded-Enc C1 Reconstruction
Data Collection		
Microscope	FEI Titan Krios (300 kV)	
Detector	Gatan K3	
Acquisition Mode	Super resolution	
Pixel Size	0.326 Å/pix (super-resolution) 0.652 Å/pix (physical)	
Total dose	40.509 e$^-$/Å^2	
Fractional dose	40 frames during a 1 s exposure	
Defocus range	0.7 – 2.5 µm (0.3 µm steps)	
Micrograph movies recorded	8109	
EMPIAR entry		
Reconstruction		
Alignment Software	MotionCor2	
Dose weighting	yes	
CTF fitting software	CTFFIND4	
Correction	full	
Particle picking method/software	Relion – templated from manual picked 2D classes	
Particles picked	340,403	
Particles in 2D classes	304,741	
Particles used in final 3D reconstruction	254,134	68,096
Alignment software	Relion 3.1	
Reconstruction software	Relion 3.1	
Box size	576 pixels (376 Å)	480 pixels (313 Å)
Voxel size (Å)	0.652	0.652
Symmetry	I1	C1
Map resolution (GS-FSC 0.143) (Å)	2.5	3.7
Sharpening B-factor (Å^2)	-97	-108
EMDB ID	EMD-12853	EMD-12873

Table S3: Model building and refinement of the 'open' and 'closed' pentamer conformations of the encapsulin nanocompartment.

Symmetry expansion of I1 dataset		
	Closed Enc pentamer	**Open Enc pentamer**
Reconstruction		
Number of asymmetric units in final model	4,734,551	4,987,966
Map resolution (GS-FSC 0.143) (Å)	2.2	2.4
Map Sharpening B-factor ($Å^2$)	-93	-117
EMDB ID	EMD-12859	EMD-12864
Coordinate refinement		
Software	Phenix	Phenix
Refinement algorithm	Real space	Real space
Resolution cutoff (Å)	2.2	2.4
$FSC_{model-vs-map}$= 0.5 (Å)	2.3	2.5
Model		
Number of amino acid residues	1375	1375
Bond length outliers	0	0
Bond angle outliers	0	0
RMS deviations		
Bonds (Å)	0.006	0.004
Angles (˚)	0.674	0.638
Validation		
Molprobity score	1.41	1.64
Clash score	3.13	6.34
Rotamer outliers (%)	2.2	0
C_β outliers (%)	0	0
CaBLAM outliers (%)	1.1	1.5
Ramachandaran (%)		
Favoured	97.79	97.05
Outlier	2.21	2.95
Model vs Data CC (mask)	0.87	0.82
EM Ringer score	5.66	4.88
PDB ID	7OE2	7OEU

Table S4: Recorded uptake of deuterium for each peptide and timepoint observed in HDX-MS of Loaded-Enc. For each peptic peptide observed in the HDX-MS analysis, the deuterium uptake (in Da) and standard deviation (SD) of triplicate data is shown for each timepoint (10 s, 30 s, 2 mins, 5 mins, 4 hours and 24 hours) for each peptide. The number of exchangeable backbone amide hydrogens is also stated (exchangers).

				Loaded-Enc deuterium uptake (Da)													
Sequence	Exchangers	Start	End	10s	10s SD	30s	30s SD	2 mins	2 mins SD	5 mins	5 mins SD	30 mins	30 mins SD	4 hrs	4 hrs SD	24 hrs	24 hrs SD
LKRHLAPIVPDAW	10	3	15	0.34	0.08	0.31	0.10	0.36	0.29	0.36	0.15	0.38	0.08	0.55	0.08	NaN	NaN
LKRHLAPIVPDAWSA	12	3	17	0.45	0.06	0.39	0.08	0.36	0.09	0.41	0.06	0.56	0.11	0.89	0.05	NaN	NaN
EAKEIFQGHLAGRKLVD	16	21	37	0.62	0.06	0.55	0.10	0.59	0.10	0.65	0.08	0.90	0.14	1.66	0.06	1.42	0.19
EAKEIFQGHLAGRKLVDF	17	21	38	0.62	0.03	0.56	0.06	0.62	0.06	0.79	0.05	1.08	0.22	1.90	0.03	1.52	0.16
IFQGHLAGRKLVD	12	25	37	0.34	0.08	0.32	0.08	0.31	0.11	0.39	0.09	0.43	0.09	0.95	0.13	0.69	0.13
IFQGHLAGRKLVDF	13	25	38	0.40	0.05	0.38	0.05	0.44	0.06	0.61	0.08	0.79	0.20	1.58	0.10	NaN	NaN
FRGPFGWEY	7	38	46	0.82	0.06	0.70	0.28	0.93	0.10	0.97	0.15	1.45	0.29	2.30	0.07	1.89	0.03
YAAVNTGELRPIDDTPED	15	46	63	1.83	0.13	1.72	0.44	2.32	0.18	2.54	0.10	3.11	0.57	5.13	1.70	4.05	0.34
YAAVNTGELRPIDDTPEDVD	17	46	65	1.92	0.11	1.89	0.45	2.53	0.15	2.75	0.15	3.37	0.66	5.40	0.20	4.52	0.50
AAVNTGELRPIDDTPED	14	47	63	1.81	0.05	1.66	0.45	2.41	0.13	2.54	0.14	3.04	0.47	4.63	0.18	3.68	0.27
LRPIDDTPED	7	54	63	0.97	0.08	0.97	0.29	1.20	0.15	1.29	0.05	1.70	0.29	2.48	0.11	1.94	0.08
LRPIDDTPEDVD	9	54	65	1.13	0.12	0.99	0.31	1.35	0.18	1.53	0.14	1.83	0.28	2.87	0.12	2.59	0.24
MKLRQVQPLAE	9	66	76	0.82	0.12	0.73	0.23	0.67	0.15	0.75	0.23	1.01	0.14	1.29	0.08	1.17	0.23
KLRQVQPLAE	8	67	76	0.88	0.09	0.78	0.20	0.81	0.12	0.87	0.09	1.01	0.12	1.28	0.06	1.00	0.15
RQVQPLAE	6	69	76	0.69	0.11	0.64	0.18	0.71	0.11	0.63	0.14	0.85	0.10	1.07	0.05	0.88	0.09
VRVPFTL	5	77	83	0.12	0.04	0.13	0.04	0.06	0.08	0.09	0.08	0.15	0.05	0.24	0.03	NaN	NaN
VRVPFTLD	6	77	84	0.26	0.03	0.29	0.04	0.27	0.06	0.33	0.04	0.40	0.08	0.64	0.03	0.59	0.10
LDSVARGATNPDLDD	13	88	102	1.50	0.14	1.21	0.28	1.36	0.20	1.79	0.06	2.04	0.38	3.43	0.15	NaN	NaN
DSVARGATNPDL	10	89	100	1.25	0.04	1.12	0.12	1.25	0.07	1.39	0.10	1.69	0.29	3.12	0.17	2.27	0.23
DSVARGATNPDLDD	12	89	102	1.52	0.03	1.45	0.18	1.60	0.12	1.72	0.03	2.05	0.32	3.34	0.13	2.42	0.29
AEDSAIFHG	8	113	121	0.15	0.05	0.14	0.09	0.05	0.07	0.14	0.10	0.16	0.05	0.17	0.05	-0.02	0.04
HGWAQAGIKGIVDSTPHEAL	18	120	139	2.27	0.08	2.15	0.51	2.54	0.23	2.84	0.10	3.25	0.41	6.04	0.09	5.14	0.58
WAQAGIKGIVDSTPHEAL	16	122	139	1.70	0.05	1.63	0.33	2.01	0.14	2.22	0.08	2.56	0.40	4.91	0.11	4.20	0.32
IVDSTPHEAL	8	130	139	0.68	0.07	0.72	0.12	0.86	0.08	0.97	0.04	1.07	0.19	2.39	0.12	1.85	0.21
AVASVSDFPRAVL	11	140	152	1.40	0.06	1.30	0.37	1.69	0.13	1.84	0.06	1.98	0.26	3.08	0.10	2.31	0.28
SAADTLRKAGVTGPYA	14	153	168	0.89	0.04	0.85	0.15	1.07	0.09	1.29	0.13	1.86	0.37	2.89	0.06	2.92	0.27
SAADTLRKAGVTGPYAL	15	153	169	0.85	0.07	0.82	0.14	1.08	0.09	1.25	0.08	1.75	0.44	2.85	0.08	2.92	0.28

LVLGPKAYDDL	9	169	179	0.47	0.06	0.44	0.11	0.59	0.08	0.62	0.05	0.75	0.10	1.27	0.06	1.04	0.15
FAATQDGYPVAKQVQRL	15	180	196	3.03	0.10	3.28	0.17	3.86	0.19	4.31	0.18	5.28	0.92	8.30	0.33	6.92	0.53
VVDGPLVRA	7	197	205	0.62	0.07	0.63	0.17	1.08	0.07	1.26	0.12	1.58	0.30	1.96	0.07	2.00	0.18
VVDGPLVRANA	9	197	207	1.08	0.06	1.20	0.13	1.70	0.09	2.02	0.08	2.59	0.18	3.26	0.10	2.84	0.19
VVDGPLVRANAL	10	197	208	1.09	0.04	1.02	0.28	1.83	0.18	2.01	0.19	2.75	0.21	3.84	0.10	3.53	0.31
VVDGPLVRANALAGA	13	197	211	1.58	0.06	1.45	0.44	2.50	0.27	2.86	0.09	3.76	0.32	5.20	0.15	5.17	0.29
VVDGPLVRANALAGAL	14	197	212	1.56	0.08	1.55	0.39	2.52	0.19	2.89	0.08	3.36	0.59	5.28	0.17	5.13	0.49
NALAGAL	6	206	212	0.46	0.05	0.50	0.14	0.70	0.09	0.77	0.06	1.10	0.18	1.92	0.07	2.14	0.07
LSIGYAFHDRSKVEL	14	229	243	0.28	0.13	0.31	0.13	0.38	0.15	0.34	0.16	0.93	0.23	1.85	0.12	1.59	0.19
YAFHDRSKVEL	10	233	243	0.48	0.07	0.41	0.11	0.39	0.11	0.42	0.20	0.75	0.12	1.18	0.29	0.95	0.30
FVAESFT	6	244	250	0.16	0.02	0.27	0.75	0.13	0.03	0.12	0.04	0.18	0.03	0.12	0.02	NaN	NaN
FRVLEPGAA	7	251	259	0.32	0.06	0.33	0.06	0.44	0.07	0.57	0.08	0.88	0.28	1.45	0.05	1.09	0.11

Table S5: Deuterium uptake for each peptide and timepoint observed in HDX-MS of Empty-Enc.
The deuterium uptake (in Da) is shown for each timepoint (10 s, 30 s, 2 mins, 5 mins, 4 hours and 24 hours) for each peptic peptide from Empty-Enc observed in HDX-MS. HDX-MS experiments were performed in triplicate and the standard deviation (SD) is shown for each time point. The number of exchangeable backbone amide hydrogens in each peptide is also shown (exchangers).

									Empty-Enc deuterium uptake (Da)									
Sequence	Exchangers	Start	End	10s	10s SD	30s	30s SD	2 mins	2 mins SD	5 mins	5 mins SD	30 mins	30 mins SD	4 hrs	4 hrs SD	24 hrs	24 hrs SD
LKRHLAPIVPDAW	10	3	15	0.34	0.05	0.35	0.02	0.38	0.02	0.40	0.03	0.40	0.07	1.36	0.23	1.48	0.03
LKRHLAPIVPDAWSA	12	3	17	0.40	0.05	0.42	0.04	0.47	0.05	0.55	0.10	0.56	0.15	2.25	0.33	2.57	0.06
EAKEIFQGHLAGRKLVD	16	21	37	0.71	0.08	0.75	0.05	0.85	0.02	0.98	0.03	1.10	0.18	3.53	0.24	3.31	0.06
EAKEIFQGHLAGRKLVDF	17	21	38	0.65	0.05	0.68	0.04	0.84	0.01	1.07	0.04	1.21	0.25	3.58	0.20	3.29	0.06
IFQGHLAGRKLVD	12	25	37	0.37	0.04	0.40	0.03	0.43	0.03	0.47	0.09	0.55	0.09	1.81	0.28	1.60	0.06
IFQGHLAGRKLVDF	13	25	38	0.46	0.04	0.49	0.04	0.60	0.03	0.77	0.07	0.93	0.19	2.20	0.19	2.05	0.06
FRGPFGWEY	7	38	46	0.92	0.08	0.99	0.03	1.14	0.03	1.23	0.08	1.61	0.07	2.90	0.17	3.01	0.03
YAAVNTGELRPIDDTPED	15	46	63	1.96	0.15	2.15	0.11	2.63	0.08	2.93	0.10	3.45	0.51	5.76	0.20	5.71	0.17
YAAVNTGELRPIDDTPEDVD	17	46	65	2.07	0.12	2.24	0.10	2.76	0.10	3.12	0.12	3.51	0.53	6.70	0.22	6.61	0.15
AAVNTGELRPIDDTPED	14	47	63	1.90	0.12	2.14	0.07	2.60	0.04	2.94	0.10	3.25	0.44	5.25	0.18	5.18	0.17
LRPIDDTPED	7	54	63	1.07	0.05	1.22	0.01	1.41	0.04	1.54	0.09	1.65	0.25	2.42	0.14	2.36	0.03
LRPIDDTPEDVD	9	54	65	1.14	0.07	1.23	0.06	1.44	0.02	1.69	0.09	1.79	0.40	3.25	0.07	3.24	0.06
MKLRQVQPLAE	9	66	76	0.82	0.11	0.86	0.05	0.90	0.05	0.77	0.12	0.76	0.09	1.62	0.16	1.56	0.08
KLRQVQPLAE	8	67	76	0.79	0.08	0.84	0.06	0.88	0.07	0.85	0.08	0.88	0.10	1.60	0.20	1.48	0.08
RQVQPLAE	6	69	76	0.64	0.08	0.69	0.05	0.75	0.06	0.71	0.06	0.70	0.08	1.20	0.13	1.12	0.06
VRVPFTL	5	77	83	0.10	0.03	0.11	0.03	0.11	0.04	0.11	0.06	0.09	0.07	0.63	0.12	0.70	0.03
VRVPFTLD	6	77	84	0.24	0.03	0.24	0.02	0.30	0.02	0.35	0.04	0.42	0.09	1.01	0.08	1.04	0.03
LDSVARGATNPDLDD	13	88	102	1.37	0.16	1.48	0.11	1.64	0.08	1.55	0.21	1.61	0.49	3.68	0.29	3.53	0.14
DSVARGATNPDL	10	89	100	1.27	0.07	1.33	0.03	1.52	0.04	1.54	0.06	1.61	0.33	3.24	0.08	3.15	0.08
DSVARGATNPDLDD	12	89	102	1.48	0.12	1.59	0.04	1.80	0.02	1.95	0.08	2.13	0.27	3.82	0.16	3.61	0.12
AEDSAIFHG	8	113	121	0.12	0.04	0.13	0.05	0.12	0.04	0.11	0.08	0.09	0.08	0.18	0.11	0.13	0.03
HGWAQAGIKGIVDSTPHEAL	18	120	139	2.36	0.19	2.51	0.07	2.86	0.07	3.07	0.13	3.39	0.40	6.66	0.14	6.57	0.19
WAQAGIKGIVDSTPHEAL	16	122	139	1.76	0.18	1.89	0.14	2.19	0.14	2.44	0.15	2.74	0.42	5.65	0.19	5.45	0.19
IVDSTPHEAL	8	130	139	0.72	0.08	0.80	0.08	0.93	0.09	1.02	0.04	1.08	0.12	2.65	0.17	2.59	0.07
AVASVSDFPRAVL	11	140	152	1.52	0.08	1.66	0.04	1.94	0.05	2.09	0.04	2.19	0.19	3.42	0.11	3.38	0.10

SAADTLRKAGVTGPYA	14	153	168	0.93	0.07	1.05	0.05	1.32	0.04	1.53	0.06	1.71	0.31	3.39	0.10	3.37	0.06
SAADTLRKAGVTGPYAL	15	153	169	0.89	0.07	1.00	0.05	1.26	0.03	1.51	0.05	1.72	0.28	3.48	0.11	3.43	0.06
LVLGPKAYDDL	9	169	179	0.58	0.05	0.60	0.04	0.69	0.03	0.80	0.05	0.88	0.18	1.97	0.09	1.98	0.05
FAATQDGYPVAKQVQRL	15	180	196	3.62	0.16	3.96	0.19	4.97	0.10	5.40	0.04	6.34	0.09	8.32	0.12	8.39	0.10
VVDGPLVRA	7	197	205	0.80	0.06	0.93	0.06	1.39	0.06	1.68	0.04	1.83	0.21	2.58	0.15	2.70	0.06
VVDGPLVRANA	9	197	207	1.27	0.07	1.53	0.08	2.15	0.08	2.44	0.04	2.85	0.07	3.86	0.10	3.88	0.07
VVDGPLVRANAL	10	197	208	1.28	0.09	1.51	0.09	2.16	0.10	2.62	0.07	3.22	0.06	4.59	0.11	4.55	0.12
VVDGPLVRANALAGA	13	197	211	2.03	0.09	2.37	0.08	3.15	0.09	3.59	0.20	4.06	0.52	6.15	0.20	6.24	0.13
VVDGPLVRANALAGAL	14	197	212	2.08	0.10	2.41	0.10	3.24	0.06	3.58	0.10	4.26	0.10	6.39	0.18	6.41	0.16
NALAGAL	6	206	212	0.72	0.04	0.81	0.04	0.93	0.04	1.02	0.04	1.26	0.25	2.28	0.11	2.34	0.09
LSIGYAFHDRSKVEL	14	229	243	0.53	0.05	0.64	0.06	0.86	0.07	0.87	0.10	0.89	0.14	1.90	0.23	1.55	0.08
YAFHDRSKVEL	10	233	243	0.44	0.05	0.51	0.05	0.56	0.05	0.62	0.06	0.68	0.11	1.16	0.26	0.93	0.07
FVAESFT	6	244	250	0.09	0.01	0.09	0.01	0.10	0.01	0.11	0.03	0.07	0.02	0.12	0.03	0.10	0.01
FRVLEPGAA	7	251	259	0.29	0.03	0.38	0.55	0.50	0.03	0.72	0.04	0.89	0.20	1.46	0.10	1.36	0.03

Table S6: gBlocks used in this study

Construct	Sequence
Empty-Enc	GGCGAAGACAT**AATG**GATCTGCTGAAACGTCATCTGGCACCGATTGTTCCGGATG CATGGTCAGCAATTGATGAAGAAGCCAAAGAAATTTTTCAGGGCCATCTGGCAGG TCGTAAACTGGTTGATTTTCGTGGTCCGTTTGGTTGGGAATATGCAGCAGTTAAT ACCGGTGAACTGCGTCCGATTGATGATACACCGGAAGATGTTGATATGAAACTGC GTCAGGTTCAGCCGCTGGCCGAAGTTCGTGTGCCGTTTACCCTGGATGTTACCGA ACTGGATAGCGTTGCACGTGGTGCAACCAATCCGGATCTGGATGATGTTGCCCGT GCAGCAGAACGTATGGTTGAAGCAGAAGATAGCGCAATTTTTCATGGTTGGGCAC AGGCAGGTATTAAAGGTATTGTTGATAGCACACCGCATGAAGCACTGGCAGTTGC AAGCGTTAGCGATTTTCCGCGTGCAGTTCTGAGCGCAGCAGATACACTGCGTAAA GCCGGTGTTACCGGTCCGTATGCACTGGTTCTGGGTCCGAAAGCCTATGATGACC TGTTTGCAGCAACCCAGGATGGTTATCCGGTTGCAAAACAGGTGCAGCGTCTGGT TGTTGATGGTCCGCTGGTTCGTGCAAATGCCCTGGCAGGCGCACTGGTTATGAGC ATGCGTGGTGGTGATTATGAACTGACCGTTGGTCAGGATCTGAGCATTGGTTATG CATTTCATGATCGTAGCAAAGTGGAACTGTTTGTGGCAGAAAGTTTTACCTTTCG TGTTCTGGAACCGGGTGCAGCCGTTCATCTGCGTTATGCATAA**AGGT**ATGTCTTC GTC
Loaded-Enc	Encapsulin: as above
	EncFtn: GGCGAAGACAT**AATG**AGCAGCGAACAGCTGCATGAACCGGCAGAACTGCTGAGCG AAGAAACCAAAAACATGCATCGTGCACTGGTTACCCTGATTGAAGAACTGGAAGC AGTTGATTGGTATCAGCAGCGTGCAGATGCCTGTAGCGAACCGGGTCTGCATGAT GTTCTGATTCATAACAAAAACGAAGAGGTGGAACATGCAATGATGACCCTGGAAT GGATTCGTCGTCGTAGTCCGGTTTTTGATGCACACATGCGTACCTACCTGTTTAC CGAACGTCCGATTCTGGAATTAGAAGAAGAAGATACCGGTAGCAGCAGCAGCGTT GCGGCAAGCCCGACCAGCGCACCGAGTCATGGTAGCTTAGGTATTGGTAGCCTGC GTCAAGAAGGTAAAGAAGATTAA**AGGT**ATGTCTTCCGG

BbsI restriction enzyme recognition sites are underlined, and overhang sequences are shown in bold.

1 **Table S7: Protein constructs used in this study.**

2

Protein Construct	Protein	Sequence	Amino acids	Average Molecular weight (Da)	pI	Extinction coefficient (M⁻¹ cm⁻¹)
Empty-Enc	Encapsulin shell protein from *H. ochraceum*	MDLLKRHLAPIVPDAWSAIDEEAKEIFQGHLAGR KLVDFRGPFGWEYAAVNTGELRPIDDTPEDVDMK LRQVQPLAEVRVPFTLDVTELDSVARGATNPDLD DVARAAERMVEAEDSAIFHGWAQAGIKGIVDSTP HEALAVASVSDFPRAVLSAADTLRKAGVTGPYAL VLGPKAYDDLFAATQDGYPVAKQVQRLVVDGPLV RANALAGALVMSMRGGDYELTVGQDLSIGYAFHD RSKVELFVAESFTFRVLEPGAAVHLRYA**R**	267	28968.84	4.88	26930
Loaded-Enc	Encapsulin shell protein from *H. ochraceum*	MDLLKRHLAPIVPDAWSAIDEEAKEIFQGHLAGR KLVDFRGPFGWEYAAVNTGELRPIDDTPEDVDMK LRQVQPLAEVRVPFTLDVTELDSVARGATNPDLD DVARAAERMVEAEDSAIFHGWAQAGIKGIVDSTP HEALAVASVSDFPRAVLSAADTLRKAGVTGPYAL VLGPKAYDDLFAATQDGYPVAKQVQRLVVDGPLV RANALAGALVMSMRGGDYELTVGQDLSIGYAFHD RSKVELFVAESFTFRVLEPGAAVHLRYA	266	28812.66	4.82	26930
	Encapsulated ferritin from *H. ochraceum*	MSSEQLHEPAELLSEETKNMHRALVTLIEELEAV DWYQQRADACSEPGLHDVLIHNKNEEVEHAMMTL EWIRRRSPVFDAHMRTYLFTERPILELEEEDTGS SSSVAASPTSAPSHGSLGIGSLRQEGKED	131	14799.39	4.70	13980

3 The additional Arg residue (bold, underlined) in the Hoch-Enc construct is a cloning artefact from the MoClo kit used.

Table S8: Autoinduction media components

Media Reagents	Concentration
Tryptone	1 % (w/v)
Yeast Extract	0.5 % (w/v)
Glycerol	0.5 % (v/v)
Glucose	0.05 % (w/v)
α-D-lactose	0.2 % (w/v)
$(NH_4)_2SO_4$	25 mM
KH_2PO_4	50 mM
$Na_2PO_4.7H_2O$	50 mM
Magnesium Sulfate	2 mM