Exploring the Extreme Acid Tolerance of a Dynamic Protein Nanocage

Jesse A. Jones[1], Michael P. Andreas[1], and Tobias W. Giessen[1,2]*

[1]Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI, USA

[2]Department of Biomedical Engineering, University of Michigan Medical School, Ann Arbor, MI, USA

*correspondence: tgiessen@umich.edu

Abstract

Encapsulins are microbial protein nanocages capable of efficient self-assembly and cargo enzyme encapsulation. Due to their favorable properties, including high thermostability, protease resistance, and robust heterologous expression, encapsulins have become popular bioengineering tools for applications in medicine, catalysis, and nanotechnology. Resistance against physicochemical extremes like high temperature and low pH is a highly desirable feature for many biotechnological applications. However, no systematic search for acid stable encapsulins has been carried out, while the influence of pH on encapsulin shells has so far not been thoroughly explored. Here, we report on a newly identified encapsulin nanocage from the acid-tolerant bacterium *Acidopropionibacterium acidopropionici*. Using transmission electron microscopy, dynamic light scattering, and proteolytic assays, we demonstrate its extreme acid tolerance and resilience against proteases. We structurally characterize the novel nanocage using cryo-electron microscopy, revealing a dynamic five-fold pore that displays distinct "closed" and "open" states at neutral pH, but only a singular "closed" state under strongly acidic conditions. Further, the "open" state exhibits the largest pore in an encapsulin shell reported to date. Non-native protein encapsulation capabilities are demonstrated, and the influence of external pH on internalized cargo is explored. Our results expand the biotechnological application range of encapsulin nanocages towards potential uses under strongly acidic conditions and highlight pH-responsive encapsulin pore dynamics.

Introduction

Protein-based compartments are used by many prokaryotes to regulate and optimize their metabolism in space and time.[1] One of the most widespread families of microbial protein compartments are encapsulins. Structurally, encapsulins self-assemble from a single type of HK97 phage-like shell protein to form icosahedral nanocages.[2] Encapsulins can exhibit different triangulation numbers, T=1 (60mer, ca. 24 nm), T=3 (180mer, ca. 32 nm), and T=4 (240mer, ca. 42 nm), with pores of varying sizes located at the 5-, 3-, and 2-fold symmetry axes.[3,4] They are classified into four distinct families, with Family 1 being the first discovered and most studied,[4] and derive their name from their ability to encapsulate specific co-regulated cargo enzymes. In Family 1, encapsulation is mediated by conserved peptide sequences found at the C-terminus of all cargo proteins called targeting peptides (TPs).[5] This feature allows encapsulins to perform a variety of biological functions, such as to act as sequestration chambers for dye-decolorizing peroxidases (DyPs), involved in combating oxidative stress;[5,6] to serve as iron mineralization and storage compartments;[7-10] as well as to sequester desulfurase enzymes, likely involved in sulfur metabolism.[11]

Due to their favorable properties, encapsulins have gained much attention as bioengineering tools.[3,12,13] As such, engineered encapsulins have been used in bacteria, yeast, and human cells for various applications, including as metabolic nanoreactors,[14] cellular imaging systems,[13] and drug delivery platforms.[15,16] A recent increase in the number of studies on encapsulin systems highlights the expanding scope of the field.[2,4,10,11,17] The encapsulin shell in particular has received substantial attention in recent years.[17] Efforts aimed at increasing shell stability,[15] controlling shell assembly,[18] and modulating pore size and dynamics have recently been reported.[19,20] Encapsulin shells efficiently self-assemble under many conditions and display marked resistance against chemical or temperature denaturation, pH, and non-specific proteases. For example, the melting temperature of the T=4 shell from *Quasibacillus thermotolerans* was reported to be nearly 87°C,[10] while the encapsulin from *Brevibacterium linens* was shown to be stable across a broad range of pH values (pH 5 to 11).[21] However, a number of biotechnological and industrial processes – including lignocellulose hydrolysis for biofuel production,[22] breakdown of complex sugars for monosaccharide production,[22] and bioleaching to prevent metal contamination and enable bioremediation in the mining industry[23,24] – would benefit from modular protein cages stable at acidic conditions below pH 5. Even though evolutionary adaptations of thermostable proteins have been well characterized and include oligomerization, large hydrophobic cores, and disulfide bond formation, adaptations that lead to acid-stable proteins are poorly understood.[25,26]

Here, we carry out the first bioinformatic search for acid-stable encapsulin nanocages and subsequently characterize the structure and stability of a Family 1 encapsulin shell from *Acidopropionibacterium acidopropionici* (AaEnc) – one of the top hits identified in our *in silico* analysis. Using a combination of techniques, including cryo-electron microscopy (cryo-EM), we characterize the structure and acid stability of the AaEnc nanocage across a wide range of pH values. Our results highlight the acid stability of AaEnc and its resilience towards protease digestion under different pH conditions. Cryo-EM analysis reveals a pH-dependent dynamic 5-fold pore with defined “closed” and “open” states, with the latter

71 representing the largest pore in an encapsulin shell reported to date. Further analyses confirm the non-
72 native cargo loading capabilities of AaEnc and reveal the effects of external pH on internalized cargo.

73

74 **Results and Discussion**

75 Bioinformatic search for acid-tolerant encapsulin shells

76 Only Family 1 encapsulins have so far been used as engineering platforms as they are the most studied
77 and well-understood of the different types of encapsulins – especially with respect to non-native cargo
78 loading. Therefore, with future engineering applications in mind, we chose to focus our bioinformatic
79 search for acid-tolerant shells on Family 1.[3,4,27,28] Further arguments for focusing on Family 1 are that the
80 widespread DyP encapsulins found in Family 1 are known to optimally function under acidic conditions[29]
81 while a variety of Family 1 systems encoded by acid-tolerant and acidophilic bacteria have already been
82 previously identified.[4]

83 As acid-tolerant proteins generally possess a low calculated isoelectric point (pI) caused by a large
84 number of surface-exposed negatively charged residues,[22,25,26,30] it was hypothesized that encapsulins
85 with a low pI may exhibit increased acid tolerance. Therefore, the large set of previously identified
86 Family 1 encapsulins with molecular weights between 26 and 35 kDa – the size range of non-cargo fused
87 encapsulins – were ranked by pI (**Figure 1a, Figure 1b,** and **Supplementary Data 1**). The lowest observed
88 pIs ranged from 4.14 to 4.49 with one protein found in the pI bin centered at a pI of 4.2 and 28 proteins
89 found in the pI bin centered at a pI of 4.4. Of these encapsulins, 18 are encoded by halophiles, four by
90 acidophiles or acid-tolerant bacteria, and the rest by soil bacteria or putative pathogens (**Table S1**).

91 We chose to focus on an encapsulin encoded by one of the acid-tolerant species, namely, the DyP
92 system of the industrially relevant *A. acidopropionici* (ATCC 4875) (AaEnc; **Figure 1c, Table S1,** and
93 **Figure S1**). *A. acidopropionici* is a Gram-positive actinobacterium able to tolerate acidic conditions as
94 low as pH 4.4.[31] It is used in agricultural applications and studied for its biotechnological and industrial
95 potential due to its production of propionic acid as a primary fermentation product, with acetic acid and
96 carbon dioxide as secondary products.[31-33]

97 The pI and sequence composition of AaEnc was compared to previously characterized encapsulin shells,
98 including those from *Mycobacterium smegmatis*, *Brevibacterium linens*, *Mycolicibacterium hassiacum*,
99 *Haliangium ochraceum*,[34] *Thermotoga maritima*, and *Myxococcus xanthus* (**Figure 1**, **Figure S1, Table 1,**
100 and **Table S1**).[2,7,21,35,36] AaEnc was found to have a lower pI than all of the so far characterized Family 1
101 encapsulins, and to display the largest ratio of acidic to basic residues (**Table 1** and **Table S1**). This makes
102 AaEnc a promising test system for exploring the acid stability of encapsulin shells.

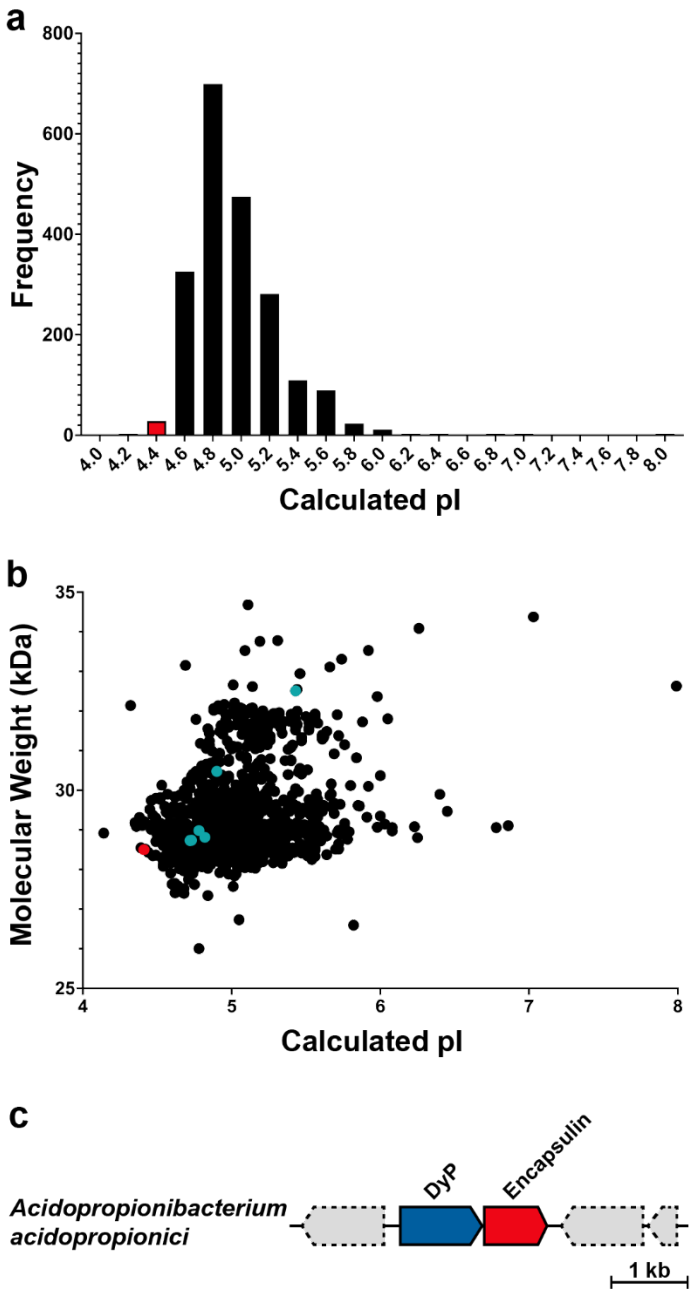


Figure 1. Bioinformatic analysis of Family 1 encapsulins. **a**) Histogram of encapsulins based on calculated isoelectric point (pI), binned at 0.2 pH units. The pI bin of interest centered at pI 4.4 and containing the *A. acidopropionici* encapsulin (AaEnc) is highlighted in red. **b**) Scatter plot of encapsulins based on calculated pI and molecular weight (MW). AaEnc (pI: 4.41, MW: 28.5 kDa) is shown in red. Previously well-characterized encapsulins from *Mycobacterium smegmatis*, *Brevibacterium linens*, *Mycolicibacterium hassiacum*, *Haliangium ochraceum*, *Thermotoga maritima*, and *Myxococcus xanthus* are shown in cyan. **c**) The *A. acidopropionici* Family 1 encapsulin operon containing a dye-decolorizing peroxidase (DyP) cargo enzyme (blue) and encapsulin shell (AaEnc, red). Functionally unrelated genes are shown in gray with dashed outlines. Scale bar: 1 kilobase (kb).

113 **Table 1.** Comparison of the pI and charge of AaEnc and other characterized encapsulins.

Encapsulin	pI[1]	# Acidic residues	# Basic residues	Acid/base ratio	Charge at pH 7.0[2]
A. acidopropionici	4.41	41	25	1.64	-23.47
M. smegmatis	4.72	41	31	1.32	-16.44
B. linens	4.73	40	30	1.33	-18.77
M. hassiacum	4.78	41	32	1.28	-15.50
H. ochraceum	4.82	40	32	1.25	-13.47
T. maritima	4.90	50	39	1.28	-13.05
M. xanthus	5.45	37	36	1.03	-7.50

114 [1] Theoretical isoelectric points calculated with Expasy Compute pI/MW tool (https://web.expasy.org/compute_pi/).
115 [2] Charges at pH 7.0 calculated via Geneious Prime 2020.2.4 (https://www.geneious.com).

116 Biophysical analysis of the AaEnc protein nanocage

117 To characterize the AaEnc nanocage, it was first heterologously expressed in *Escherichia coli* and then
118 purified via a combination of polyethylene glycol (PEG) precipitation, anion exchange chromatography
119 (IEC), and size exclusion chromatography (SEC), with the latter serving as an initial verification that the
120 nanocage was assembled at pH 7.5 (**Figure 2a**). Aliquots of the purified sample were then exchanged
121 into various buffers across a wide range of pH values (pH 1.5 to 10.0) while holding the salt
122 concentration constant at a physiological value of 150 mM NaCl. After incubation for 6 h, samples were
123 imaged via negative stain transmission electron microscopy (TEM) to assess the effect of pH on protein
124 aggregation and the assembly state of the AaEnc nanocage. Some assembled AaEnc shells could be
125 observed at pH values as low as 1.9 and as high as 7.5 (**Figure S2**). However, the pH range within which
126 AaEnc was close to fully assembled with only minor aggregation occurring was between pH 2.25 and 7.5.
127 Therefore, subsequent biophysical analyses were carried out at four pH values spanning this pH range,
128 namely at pH 2.25, 3.0, 5.0, and 7.5. Dynamic light scattering (DLS) analyses and negative stain TEM
129 indicated that AaEnc maintained a similar size and appearance across all four tested conditions with Z-
130 average diameters of 30.3 nm at pH 2.25, 30.2 nm at pH 3.0, 29.8 nm at pH 5.0, and 24.4 nm at pH 7.5
131 (**Figure 2b** and **Figure 2c**). The slight increase in average diameter at acidic pH values is likely due to
132 limited aggregation under these conditions, however, as can be seen in TEM micrographs, individual
133 shells at all tested pH values exhibited diameters of ca. 24 nm. Static light scattering (SLS) further
134 indicated that AaEnc is relatively stable across all tested pH values, with aggregation temperatures (T_{agg})
135 of 36.0°C at pH 2.25, 38.3°C at pH 3.0, 39.3°C at pH 5.0, and 62.4°C at pH 7.5 (**Figure 2d**). However, a
136 clear trend can be observed with lower pH values leading to decreased T_{agg} values. We next explored the
137 resistance of AaEnc against proteolytic degradation at various pH values. AaEnc proved to be relatively
138 resilient to pepsin degradation at pH 3.0 and 37°C during a 3 h incubation period, whereas the control
139 protein, bovine serum albumin (BSA), was completely degraded under the same conditions (**Figure 2e**).
140 Similarly, AaEnc was relatively resistant to degradation by trypsin and chymotrypsin at pH 7.5 and 37°C
141 over an 8 h period, with BSA being again substantially degraded under the same conditions (**Figure 2f**).
142 Overall, these results highlight the substantial acid stability of AaEnc which is significantly higher than
143 that of any other previously characterized encapsulin.

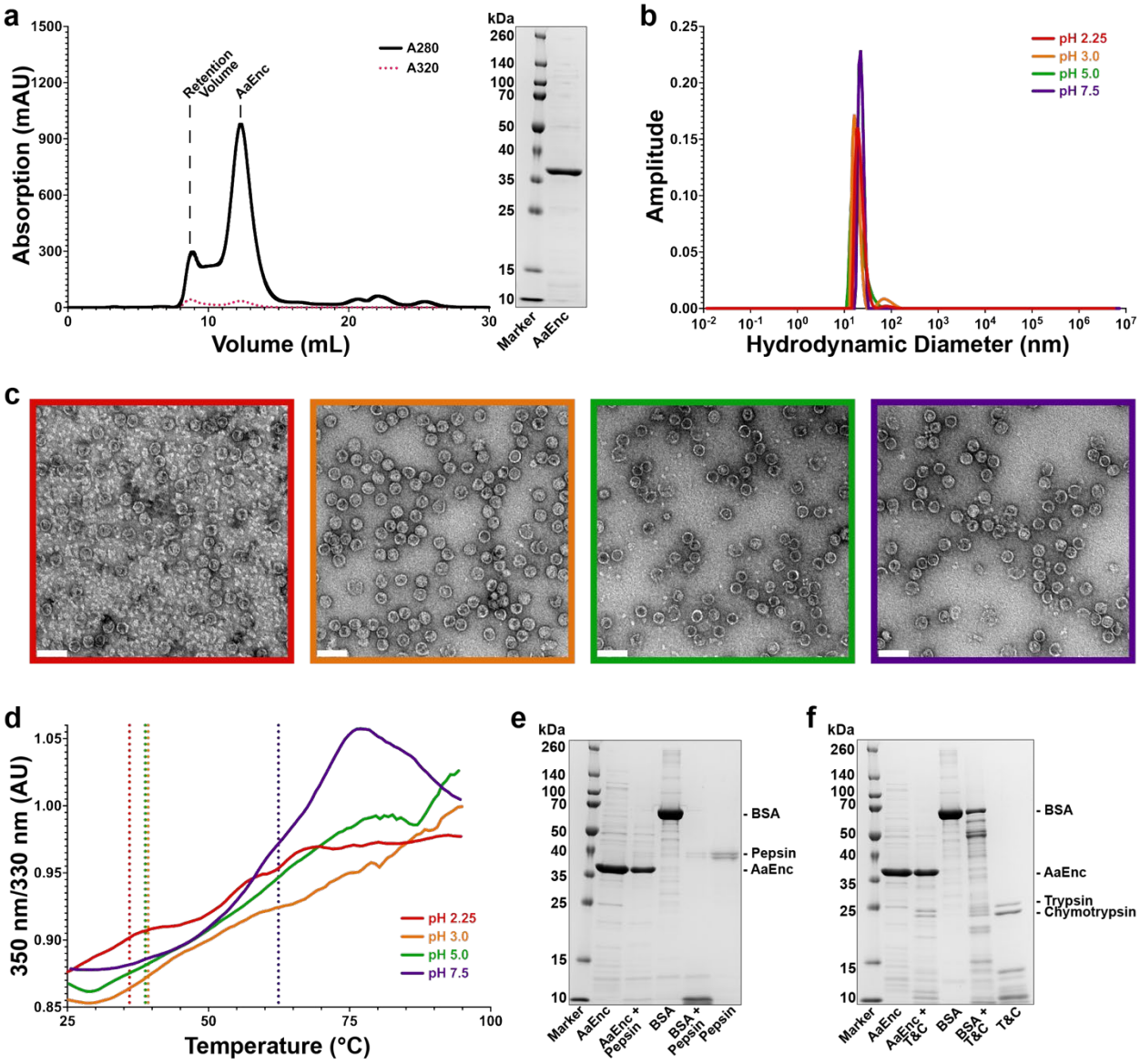


Figure 2. Biophysical analysis of the AaEnc nanocage. **A**) Size exclusion chromatography (SEC) of the AaEnc nanocage depicting elution at 12 mL suggestive of an assembled T=1 encapsulin (left), along with SDS-PAGE analysis of purified AaEnc (right). **B**) Dynamic light scattering (DLS) analysis of AaEnc at pH 7.5 (purple), pH 5.0 (green), pH 3.0 (orange), and pH 2.25 (red). **C**) Transmission electron microscopy (TEM) analysis of AaEnc after 6 h incubation at different pH values: pH 2.25 (red, far left), pH 3.0 (orange, middle left), pH 5.0 (green, middle right), and pH 7.5 (purple, far right). Scale bars: 50 nm. **D**) Representative thermal unfolding curves for AaEnc at different pH values: pH 2.25 (red), pH 3.0 (orange), pH 5.0 (green), and pH 7.5 (purple) with corresponding aggregation temperatures (T_{agg}, vertical dashed lines, respective colors). **e**) Protease stability analysis of AaEnc exposed to pepsin at pH 3.0 for 3 h with bovine serum albumin (BSA) as a control. **f**) Protease stability analysis of AaEnc exposed to trypsin and chymotrypsin (T&C) at pH 7.5 for 8 h with BSA as a control.

Structural characterization of the AaEnc protein nanocage

To further characterize the influence of pH on AaEnc, single particle cryo-EM analysis was carried out, initially at a physiological pH of 7.5. Results revealed the existence of two discrete structural states distinguished by an either all “closed” or all “open” conformation of the 5-fold pores within the AaEnc

158 shell (**Figure 3** and **Supplementary Video 1**). The "closed" and "open" states were determined to 2.90
159 (29,056 particles) and 3.32 Å (13,581 particles), respectively (**Figure S3**). About 68% of the used particles
160 comprise the "closed" state while about 32% exhibit an "open" state. The shell for both states consists
161 of 60 AaEnc protomers, forming a ca. 1.7 MDa, T=1 icosahedral protein cage with a diameter of 24 nm
162 and an overall negatively charged exterior surface (**Figure 3a** and **Figure 3c**). Symmetric (icosahedral, I)
163 and asymmetric (C1) refinements were carried out for both states to investigate if a given pore can be
164 "closed" or "open" independently of the other pores in the same shell or if all pore states are generally
165 correlated. Both I and C1 refinements yielded similar all "closed" and all "open" states (**Figure S3**), thus
166 confirming that under the given experimental conditions, the pore dynamics of all pentameric facets
167 within a shell appear to be strongly correlated. However, this does not necessarily exclude the possibility
168 that, under certain environmental conditions or in the presence of cargo, a single shell can contain both
169 "closed" and "open" 5-fold pores at the same time. This phenomenon – independent dynamic 5-fold
170 pores – has indeed been observed for the Familiy 1 encapsulin from *H. ochraceum* (**Figure S4**).[37] To
171 further analyze 5-fold pore dynamics, 3D variability analysis with three components was carried out in
172 cryoSPARC for both datasets, however, no states could be resolved that would indicate the
173 simultaneous presence of both "closed" and "open" pores within the shell. The 5-fold pore diameters in
174 the AaEnc nanocage are 5 Å for the "closed" and 20 Å for the "open" state (**Figure 3b, Figure 3d**, and
175 **Figure S3**). Thus, the AaEnc "open" state represents the largest pore found in an encapsulin shell to
176 date, substantially larger than the previously reported dynamic *H. ochraceum* pore which exhibited an
177 "open" state diameter of 15 Å.[37]

178 Detailed examination of the "closed" and "open" state structures indicates that the primary
179 conformational changes underlying the dynamic nature of the AaEnc pore are located at the apex of the
180 so-called axial domain (A-domain) of the encapsulin protomer (**Figure 4a**). In addition, an overall
181 backwards tilt of the "open" state protomer by 10° (**Figure 3b** and **Figure 3d**) also contributes to the
182 observed size increase of the 5-fold pore (**Figure 3c**, **Figure 3d, Supplementary Video 2**). In the "closed"
183 state, five pore residues – Asp185, His186, Gly187, Val188, and Pro189 – form a short loop between the
184 α6 and α7 helices encompassing a predicted α-turn (**Figure S5**); whereas in the "open" state, the same
185 residues form a tighter turn that loses the predicted α-turn conformation, with Val188 and Pro189
186 becoming part of and extending the α7 helix (**Figure 4c**, **Figure S5**, and **Figure S6**).[38] His186 undergoes
187 the most readily observable conformational change. It is located at the apex of the A-domain in the
188 "closed" state, yet partially buried between two adjacent protomers in the "open" state (**Figure 4b**,
189 **Figure 4c**). Furthermore, in the "open" state, intermolecular hydrogen bonding is observed between
190 His186 and Asp150 as well as Ser181 and Asn157 of adjacent protomers (**Figure 4c**). These two
191 hydrogen bonds are notably absent in the "closed" state. His186 is not strictly conserved among other
192 structurally characterized Family 1 encapsulins (**Figure S1**) and cannot be used alone as an indicator for
193 the presence of dynamic 5-fold pores in encapsulin shells, as in the *H. ochraceum* encapsulin, which also
194 displays "closed" and "open" pore states, where His186 is substituted with Asp186 (**Figure S1** and **Figure**
195 **S4**). Interestingly, in *H. ochraceum* the residue corresponding to Asp150 in AaEnc is Arg150. This could
196 indicate that analogous to the hydrogen bonding between His186 and Asp150 (AaEnc), bonding
197 between Asp186 and Arg150 (*H. ochraceum*), with swapped H-bond donors/acceptors, may be possible
198 under certain conditions. However, this was not observed in the *H. ochraceum* "open" conformation.

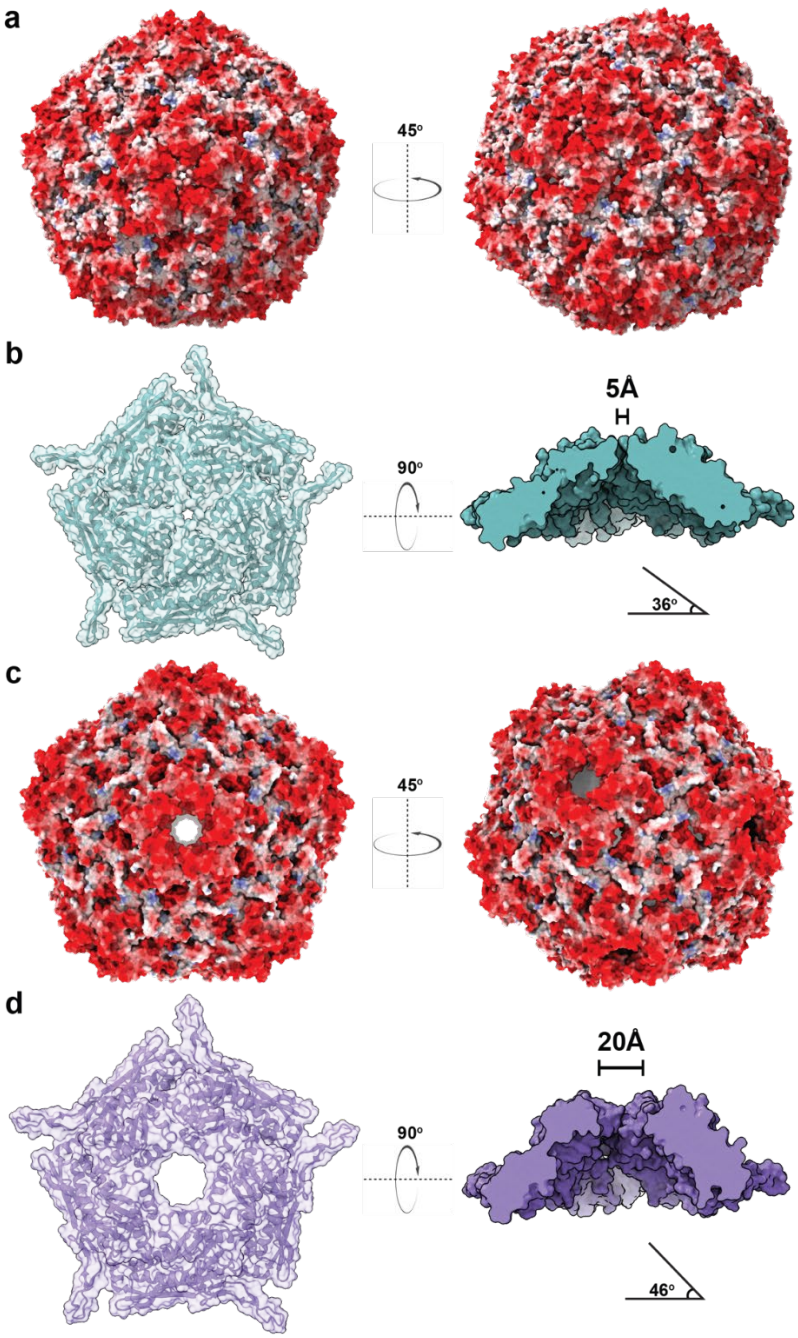


Figure 3. Structural overview of the AaEnc nanocage and 5-fold pore. **a**) Electrostatic surface representation of AaEnc in the "closed" conformation viewed down the 5-fold symmetry axis as well as at a 45° rightward turn (red, negative charge; white, neutral; blue, positive charge). **b**) Top-down ribbon and partially transparent surface representation of the "closed" AaEnc pentamer (left; cyan) and solid surface representation rotated 90° and viewed through the frontal axis plane to highlight pore size (right). The angle between protomers and the orthogonal of the 5-fold axis is highlighted. **c**) Electrostatic surface representation of AaEnc in the "open" conformation. **d**) Top-down ribbon and partially transparent surface representation of the "open" AaEnc pentamer (left; purple) and solid surface representation rotated 90° and viewed through the frontal axis plane to highlight pore size (right). The angle between protomers and the orthogonal of the 5-fold axis is highlighted.

Overall, the cryo-EM density for the loop region of the *H. ochraceum* encapsulin in the "open" conformation was not well defined, whereas for AaEnc, both "closed" and "open" states exhibit strong and well-defined densities (**Figure S6**). This could in part be due to the additional stabilization of the "open" state in AaEnc resulting from the hydrogen bonding observed between His186 and Asp150.

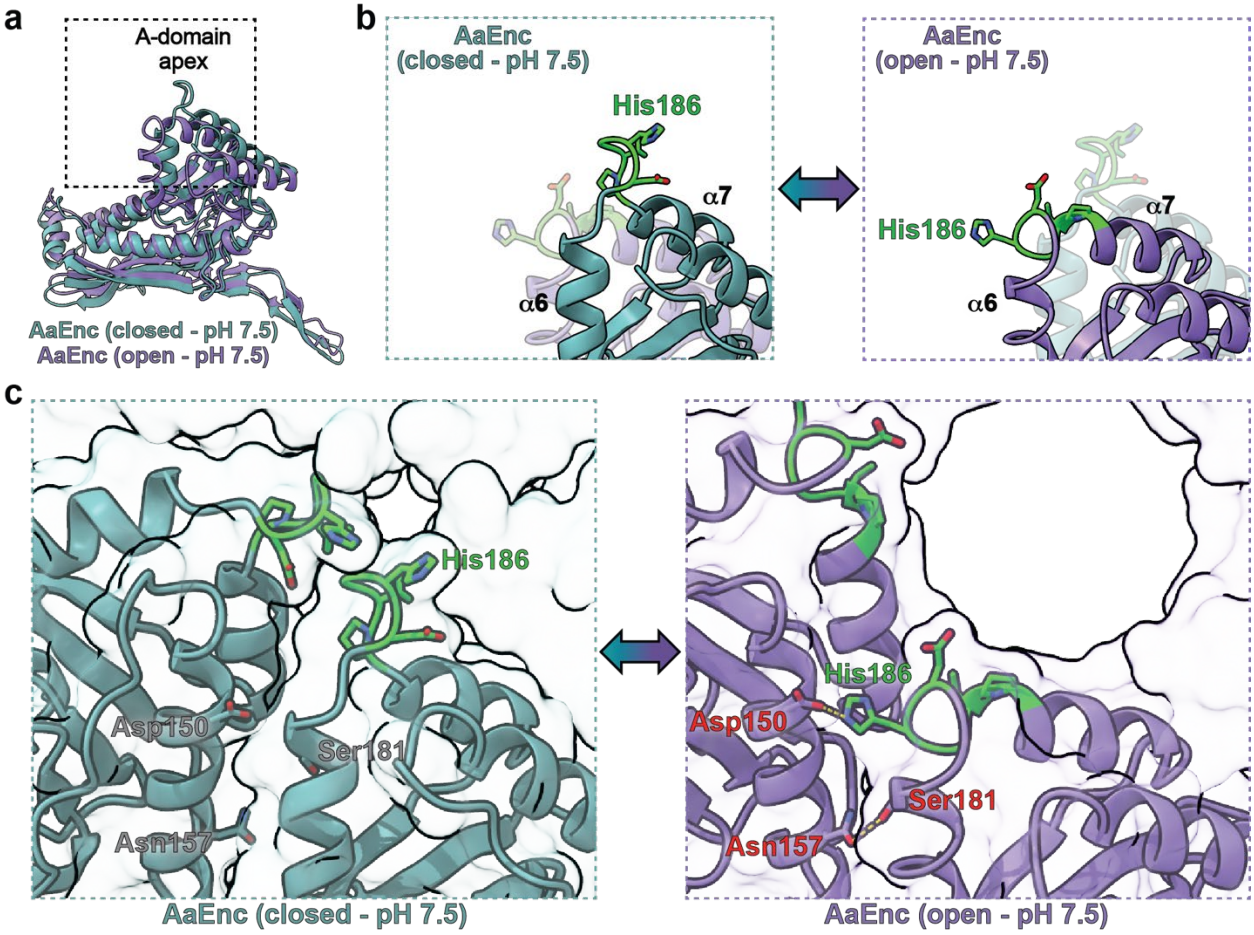


Figure 4. Detailed structural analysis of the AaEnc 5-fold pore. **a**) Aligned and overlayed ribbon representation of the "closed" (cyan) and "open" (purple) AaEnc protomers with dashed box highlighting the A-domain. **b**) Magnified ribbon representation juxtaposing the dynamic A-domain of the "closed" (cyan; left, solid; right, transparent) and "open" (purple; right, solid; left, transparent) AaEnc protomer, with the loop residues of interest–Asp185, His186 (labeled), Gly187, Val188, and Pro189–highlighted (green). **c**) Magnified solid ribbon representation of two adjacent AaEnc A-domains with transparent surface representation of the AaEnc pentamer highlighting the 5-fold pore. The "closed" state (left, cyan) exhibits a lack of hydrogen bonds between Asp150 (gray) and His186 (green) as well as Asn157 (gray) with Ser181 (gray), while the "open" state (right, purple) showcases gained hydrogen bonds between Asp150 (red) with His186 (green), as well as Asn157 (red) with Ser181 (red).

Additional cryo-EM experiments were carried out at pH 3.0 to assess the structure of AaEnc under strongly acidic conditions. At pH 3.0, only a single "closed" conformational state was observed and was determined to 2.77 Å resolution (47,164 particles) (**Figure S7** and **Table S3**). It was also found that both of the "closed" AaEnc states – at pH 7.5 and pH 3.0 – are seemingly identical, with a root-mean-square deviation (RMSD) of 0.32 between the two aligned protomers (**Figure S8**).[39]

227 *In silico* analyses were conducted using the APBS-PDB2PQR software suite to quantitatively assess the
228 hydrogen bonding and solvent exposure of the His186 and Asp150 residues at pH 7.5 and 3.0 (**Table 2**).[40]
229 The results further corroborate that His186 is more exposed – calculated as only 10% buried – in the
230 "closed" states, and more buried – calculated as 39% buried – in the "open" state due to the inter-
231 protomer hydrogen bonding described above. Furthermore, no hydrogen bonds were predicted
232 between His186 and Asp150 in either of the "closed" states, while being clearly predicted for the "open"
233 state.

234 Based on the results outlined above the "closed" state is clearly favored at low pH. It seems likely that in
235 addition to global protonation state changes throughout the AaEnc protomer, specifically the
236 protonation of Asp150 at low pH would preclude any hydrogen bonding with His186, thus making the
237 conformational change from "closed" to "open" state energetically less favorable at acidic conditions.
238 The specific molecular and biological functions of dynamic 5-fold pores in encapsulin shells is currently
239 unknown. However, as the native cargo of AaEnc is a DyP-type peroxidase, generally known to be
240 optimally active at acidic pH values, the preference of the AaEnc nanocage for the "closed" pore state at
241 low pH might have significant functional and biological implications.

242 **Table 2**. *In silico* analysis of key pore residues. Calculated buriedness, hydrogen bonds, and pKa values are shown.

State and residue	Buried	Sidechain Hydrogen Bond (Partner)[1]	Calculated pKa
Closed (pH 7.5)			
Asp150	44%	-0.57 (Gln147, intramolecular)	5.63
His186	10%	none	6.00
Open (pH 7.5)			
Asp150	12%	-0.53 (His186, intermolecular)	3.50
His186	39%	0.53 (Asp150, intermolecular)	6.63
Closed (pH 3.0)			
Asp150	43%	0.90 (Asp149, intramolecular)	7.64
His186	10%	none	5.85

243 [1] pKa shift due to respective hydrogen bond.

244 Computational analysis of continuum electrostatics and solvation of the AaEnc protomer

245 To gain deeper insights into the stability of the AaEnc nanocage under acidic conditions, an array of
246 computational analyses focused on continuum electrostatics and solvation of the AaEnc protomer was
247 carried out. In particular, we thought to investigate if the AaEnc protomer would be predicted to show
248 increased acid stability by itself outside the context of the encapsulin shell. The APBS-PDB2PQR software
249 suite was used to assess the surface electrostatics of AaEnc at different pH values ranging from pH 2.0 to
250 8.0 (**Figure 5a**).[40] The calculated protein electrostatics correlated well with the calculated pI of 4.41,
251 showing a change from an overall positive surface charge below pH 4.0 to an overall negative surface
252 charge above pH 5.0. Next, using the Protein-Sol software package, a heatmap depicting the average
253 charge per residue based on pH and ionic strength was calculated.[41] The results again correlate well with
254 the calculated pI, with a positive average charge per residue at pH 4.0 and below, and a negative
255 average charge per residue at pH 4.5 and above, regardless of ionic strength (**Figure 5b**). To investigate if
256 a discernible increase in folded state protein stability mediated by interactions between ionizable

groups might exist for the AaEnc protomer at acidic pH, further analyses using Protein-sol were carried out (**Figure 5c**). It was found that at physiological ionic strength (150 mM), per residue energies below pH 4.5 were positive, indicating decreased stability of the protomer fold below this pH threshold.

Taken together, our computational analysis of the AaEnc protomer suggests that the observed acid stability of the AaEnc nanocage is not easily attributable to a highly stable protomer building block. Instead, AaEnc acid stability is likely due to a complex combination of factors at the scale of the assembled 60mer nanocage. Important factors likely include favorable inter-protomer interactions within the context of the encapsulin shell, such as the structural dynamics and hydrogen bonding interactions discussed above.

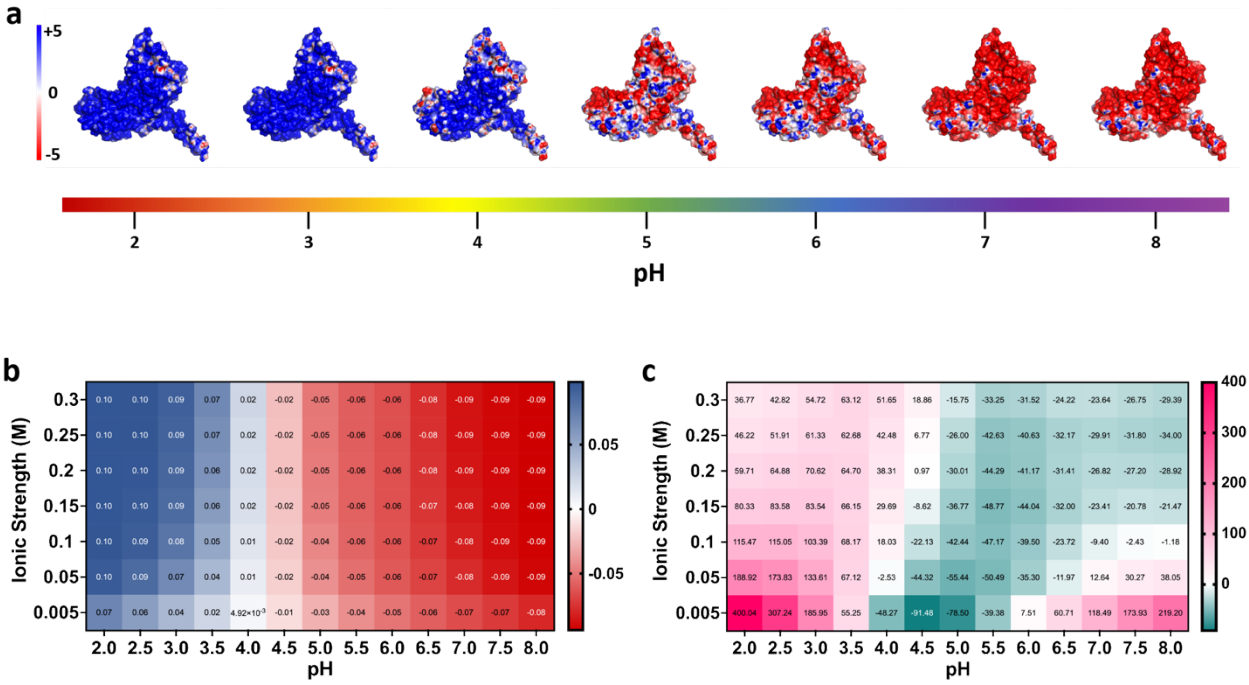
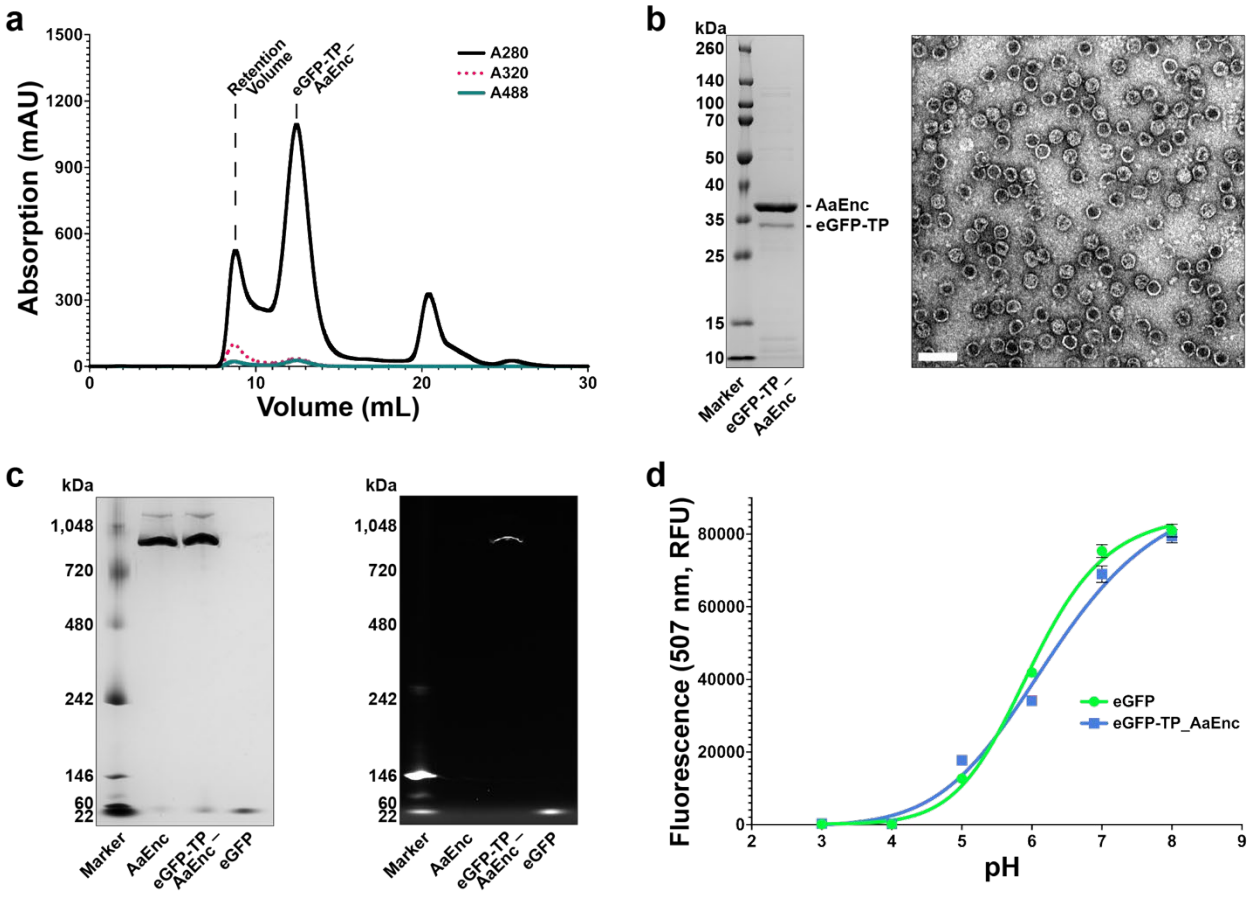


Figure 5. Computational characterization of the AaEnc protomer. **a**) Surface charge visualization of the AaEnc protomer with amino acid protonation states calculated by pH with pdb2pqr and protein electrostatics calculated with APBS. **b**) Heatmap of predicted average charge per residue for the AaEnc protomer at different pH and ionic strength values. **c**) Heatmap of predicted folded state protein stability as interactions between ionizable groups in joule per residue for the AaEnc protomer at different pH and ionic strength values.

<u>*In vivo* cargo loading and pH effects on internalized cargo</u>

To investigate if the AaEnc shell has any influence on the acid stability of internalized cargo proteins, heterologous cargo loading experiments were carried out followed by pH screens. The predicted C-terminal targeting peptide (TP) of the native AaEnc DyP cargo enzyme, was genetically fused to the C-terminus of eGFP (eGFP-TP) and cloned immediately upstream of the AaEnc gene for co-expression.[4,42] eGFP was chosen as a non-native cargo due to its reliable expression, favorable solubility, simple detection, and predictable and well-reported pH sensitivity profile.[43,44] *In vivo* eGFP-TP cargo loading was confirmed via its co-purification by SEC with co-expressed AaEnc, negative stain TEM analysis, and native polyacrylamide gel electrophoresis (PAGE) (**Figure 6a**, **Figure 6b**, and **Figure 6c**). Next, the

281 fluorescence of equimolar amounts of free eGFP and AaEnc-encapsulated eGFP-TP were compared
282 across a range of pH values from pH 3.0 to 8.0 using a plate-based fluorescence assay (**Figure 6d**). Both
283 free eGFP and AaEnc-encapsulated eGFP-TP yielded very similar sigmoidal fluorescence response curves,
284 demonstrating that the interior pH of AaEnc is not appreciably different from the bulk pH. Further,
285 encapsulation within AaEnc does apparently not alter cargo pH sensitivity to a significant degree. These
286 results indicate that the AaEnc shell does not represent an effective diffusion barrier for protons. Thus,
287 buffer pH changes will result in the rapid equilibration of the external and luminal pH. Similar behavior
288 has been observed for other protein-based compartments, particularly the carbon-fixing carboxysome
289 bacterial microcompartment.[45] Stopped-flow pH colorimetry indicated a rapid equilibration of the
290 luminal carboxysome pH to that of the bulk solvent highlighting the porosity of the carboxysome shell
291 towards protons. However, considering that one of the likely primary functions of encapsulin and
292 carboxysome shells is to control the flux of specific small molecules into and out of the shell interior[45,46]
293 – possibly through the action of dynamic or gated pores – the idea that protein shells could be able to
294 control the passage of protons is not out of the realm of possibility. A molecular mechanism similar to
295 that employed by aquaporins, which allow passage of water molecules but block proton flux, could
296 certainly exist or be engineered in protein shells.[47]



297

298 **Figure 6.** Analysis of eGFP cargo-loaded AaEnc. **a**) Size exclusion chromatography (SEC) analysis of AaEnc-encapsulated eGFP
299 showing elution at 12-13 mL via protein absorbance at 280 nm and specific eGFP absorbance at 488 nm. **b**) SDS-PAGE analysis
300 of purified eGFP-loaded AaEnc (left) and negative stain TEM (right). Scale bar: 50 nm. **c**) Native PAGE gel analysis of empty

301 AaEnc, eGFP-loaded AaEnc, and free eGFP (left) along with corresponding fluorescence analysis of the same gel highlighting co-
302 elution of eGFP fluorescence with the high molecular weight encapsulin band (right). **d**) Plate-based fluorescence analysis of
303 AaEnc-encapsulated eGFP (eGFP-TP AaEnc) versus free eGFP. Data shown as means with error bars representing standard
304 deviations from three independent experiments.

305

306 **Conclusion**

307 The large number of encapsulin systems distributed across diverse bacterial and archaeal phyla –
308 including many extremophiles – represents a largely untapped source of novel biotechnological tools.[4]
309 Ongoing discoveries and research within the encapsulin field has resulted in the characterization of
310 many nano-encapsulation systems with a quickly expanding and diverse list of useful molecular
311 features.[10,11,14,18,28] However, relatively little attention has been focused on systematically exploring
312 encapsulins from extremophilic bacteria and archaea with unusual molecular characteristics and
313 stability profiles. With this study, we have taken the first step towards addressing this issue with a focus
314 on the acid stability of the AaEnc nanocage.

315 Our results highlight the difficulty of pinpointing specific protein characteristics that lead to increased
316 acid stability. Beyond the previously reported observation that acid stable proteins contain an increased
317 number of aspartate and glutamate residues, resulting in a low pI,[48] no other adaptations towards acid
318 stability are readily apparent for AaEnc. We find that at the protomer level, AaEnc does not display any
319 obvious properties beyond a low pI, that would indicate exceptionally high acid stability. This analysis is
320 necessarily purely computational as encapsulin protomers quickly self-assemble to form protein
321 nanocages and cannot be studied in isolation. It seems likely that the formation of a large 60mer protein
322 complex plays a role in the acid tolerance of AaEnc with some assembled shells still present after
323 extended incubation at pH 1.9. It can be speculated that minimizing the number of ionizable groups at
324 key subunit interfaces within the AaEnc shell would contribute towards its stability at low pH. Beyond its
325 unusual acid tolerance, AaEnc exhibits the unique feature of seemingly pH-dependent all "closed" or all
326 "open" 5-fold pore states. At physiological pH, the "closed" and "open" states exist in a ratio of 2:1,
327 whereas at low pH, the equilibrium is completely shifted towards the "closed" state. This behavior likely
328 has important functional and biological implications and will require further study. Its stability, pH-
329 responsive pores, and the fact that the "open" pore state with a diameter of 20 Å represents the largest
330 encapsulin pore reported to date, make AaEnc an interesting target for future nanocage engineering
331 applications in catalysis, nanotechnology, and medicine.

332 Specifically, acid-tolerant protein cages could offer novel opportunities in nanoreactor design and
333 engineering, particularly in the context of industrial biopolymer degradation which requires acidic
334 conditions. Enzyme encapsulation or co-localization could improve the performance of enzymes like
335 chymosins,[49] dye-decolorizing peroxidases,[50] glucoamylases,[51] and proteases,[52] already extensively used
336 in the food industry, agriculture, and biofuel production.[53-56] Further, bioleaching and bioremediation
337 approaches aimed at recovering valuable or toxic metals could benefit from acid stable nano-
338 encapsulation systems able to sequester specific metal-binding enzymes of interest.[24,57] Finally, a
339 number of biomedical applications of protein nanocages related to drug delivery or intracellular

340 targeting of acidic compartments could also benefit from robust and easily engineerable nanocages like
341 AaEnc.[58,59]

342 In sum, AaEnc is the first characterized highly acid stable encapsulin nanocage with a unique pH-
343 dependent dynamic pore and, therefore, represents a novel useful tool for the nanocage engineering
344 community.

345

346 **Methods**

347 Chemicals and biological materials

348 All chemicals were used as supplied by vendors without further purification. Imidazole, Invitrogen Novex
349 WedgeWell 14% tris-glycine Mini Protein Gels, Isopropy-ß-D-thiogalactopyranoside (IPTG), lysozyme,
350 NativePAGE™ 4 to 16% bis-tris Mini Protein Gels, NativeMark Unstained Protein Standard, Spectra™
351 Multicolor Broad Range Protein Ladder, Thermo Scientific Pierce 660 nm Protein Assay Reagent, Tris
352 base, Tris HCl, all restriction enzymes, and all cell culture media and reagents were purchased from
353 Fisher Scientific, Inc. (USA). Gibson Assembly Master Mix was purchased from NEB (USA). Amicon Ultra-
354 0.5 mL centrifugal units and Benzonase® nuclease were purchased from MilliporeSigma (USA). BL21
355 (DE3) Electrocompetent Cells used for *E. coli* expression were also purchased from MilliporeSigma (USA).
356 Bis-tris propane from Research Products International (USA) was used for the assembly buffer. Ni-NTA
357 agarose from Gold Biotechnology, Inc. (USA) was used for His-tagged protein purification.

358 Instrumentation

359 Cell lysis was conducted via sonication with a Model 120 Sonic Dismembrator from Fisher Scientific, Inc.
360 (USA). Protein was quantified on a Nanodrop Spectrophotometer from ThermoFisher Scientific, Inc.
361 (USA). Protein purification was carried out on an AKTA Pure fast liquid protein chromatography system;
362 size exclusion chromatography (SEC) was carried out with a HiPrep 16/60 Sephacryl S-500 HR and
363 Superose 6 10/300 GL columns (Cytiva, USA); anion exchange was carried out with a HiTrap Q FF column
364 (Cytiva, USA). Polyacrylamide gel electrophoresis (PAGE) and NativePAGE were performed in an XCell
365 SureLock from Invitrogen/ThermoFisher Scientific (USA). Gel images were captured using a ChemiDoc
366 Imaging System from Bio-Rad Laboratories, Inc. (USA). DLS was carried out on an Uncle from Unchained
367 Labs (USA). TEM was carried out on a Morgagni 100 keV Transmission Electron Microscope (FEI, USA).
368 Plate-based fluorescence assays were conducted on the Synergy H1 Microplate Reader from BioTek
369 Instruments (USA). EM grid glow discharging was conducted with a PELCO easiGlow™ system by Ted
370 Pella, Inc (USA). A Talos Arctica Cryo Transmission Electron Microscope by ThermoScientific, Inc. (USA)
371 equipped with a K2 Summit direct electron detector by Gatan, Inc. (USA) located at the University of
372 Michigan Life Sciences Institute was used for cryo-EM. Smaller materials are listed along with
373 corresponding methods below.

374 Software

375 The following software was used throughout this work: Adobe Illustrator 2021 v25.0.0 (figures),
376 cryoSPARC v3.3.1[60] (cryo electron microscopy), Fiji/ImageJ v2.1.0/1.53c[61] (densitometric data analysis

377 and TEM images), GraphPad Prism for Mac OS v9.4.0 (chromatography, melting
378 temperature/aggregation, and fluorescence graphs), Bio-Rad Image Lab Touch Software (gel imaging),
379 Microsoft Excel for Mac v16.46 (DLS graphs), Phenix v1.19.2-4158[62] (model building), UCSF Chimera
380 v1.16[63] and ChimeraX v3[39] (cryo-EM density and model visualization), and UNICORN 7 (FPLC system
381 control and chromatography). Online software suites or tools are listed along with corresponding
382 methods below.

383 Bioinformatic search for acid-stable encapsulins

384 A curated list of Family 1 encapsulins[19] was sorted according to molecular weight, removing any entries
385 falling below 25 kDa or above 35 kDa to remove partial annotations and fusion encapsulins, respectively.
386 Results were then processed via the Expasy Compute pI/MW tool
387 (https://web.expasy.org/compute_pi/). Data was then organized according to calculated pI and binned
388 for histogram analysis or plotted for scatterplot analysis via GraphPad Prism.

389 Sequence alignments

390 Encapsulin alignments were generated with the ESPript 3 server (http://espript.ibcp.fr/) using a protein
391 sequence alignment produced with Clustal Omega, with secondary structure information based on the
392 TmEnc structure (PDB 3DKT; **Figure S1**) or the "open" and "closed" AaEnc structures (**Figure S5**).[38,64]

393 Protein production

394 For all target proteins, plasmids were constructed with target *E. coli* codon-optimized gBlock genes,
395 synthesized by IDT (USA), inserted into the pETDuet-1 vector via Gibson assembly using the NdeI and
396 PacI restriction sites (**Table S2)**. *E. coli* BL21 (DE3) was transformed with the respective plasmids via
397 electroporation per protocol and 25% glycerol bacterial stocks were made and stored at -80°C until
398 further use. Starter cultures were grown in 5 mL LB with 100 mg/mL ampicillin at 37°C overnight. For all
399 constructs, 500 mL of LB with ampicillin was inoculated with overnight starter cultures and grown at
400 37°C to an OD600 of 0.4-0.5, then induced with 0.1 mM IPTG and grown further at 30°C overnight for
401 ca. 18 h. Cells were then harvested via centrifugation at 10,000 rcf for 15 minutes at 4°C and pellets
402 were frozen and stored at -80°C until further use.

403 Protein purification

404 Frozen bacterial pellets were thawed on ice and resuspended in 5 mL/g (wet cell mass) of cold Tris
405 Buffered Saline (20 mM Tris pH 7.5, 150 mM NaCl). Lysis components were added (0.5 mg/mL lysozyme,
406 1 mM tris(2-carboxyethyl)phosphine [TCEP], one SIGMA*FAST* EDTA-free protease inhibitor cocktail tablet
407 per 100 mL, 0.5 mM $MgCl_2$, and 25 units/mL Benzonase® nuclease) and samples were lysed on ice for 10
408 min. Samples were then sonicated at 60% amplitude for 5 min total (eight seconds on, 16 seconds off)
409 until no longer viscous. After sonication, samples were centrifuged at 8,000 rcf for 15 minutes at 4°C.
410 Samples were then subjugated to 10% polyethylene glycol (PEG) 8000 precipitation (lysate brought to
411 10% PEG 8K and 500 mM NaCl and incubated for 30 minutes on ice, then centrifuged 8,000 rcf for 15
412 min). Supernatant was discarded and the pellet was resuspended in 5 mL TBS pH 7.5 and filtered using a
413 0.22 μm syringe filter (Corning, USA). The protein sample was then loaded on an AKTA Pure and purified

414 via a Sephacryl S-500 column. Sample fractions were pooled and buffer exchanged into AIEX Buffer (20
415 mM Tris pH 7.5) and loaded onto an AKTA Pure, then purified via HiTrap Q-FF by linear gradient into
416 AIEX Buffer with 1M NaCl. Sample flow-through was collected and centrifuged at 10,000 rcf for 10 min,
417 then loaded on an AKTA Pure for final purification via a Superose 6 10/300 GL column pre-equilibrated
418 with TBS pH 7.5. All proteins were stored at 4°C until use.

419 For free His-tagged eGFP purification, the sample was lysed as above in NTA Resuspension Buffer (50
420 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP, 10 mM imidazole, and 5% glycerol). Lysate was bound to Ni-
421 NTA resin pre-equilibrated with NTA Resuspension Buffer via rocking at 4°C for 45 minutes. Supernatant
422 was discarded and the bound sample was washed once with NTA Resuspension Buffer and a second
423 time with NTA Resuspension Buffer with 20 mM imidazole. Free His-tagged eGFP was then eluted three
424 times with Elution Buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP, 350 mM imidazole, 5%
425 glycerol) and stored at 4°C for future use.

426 Transmission electron microscopy

427 Samples were diluted to 0.1-0.3 mg/mL and buffer exchanged into various buffers ranging from pH 1.5
428 to pH 10.0 (**Figure 2c** and **Figure S2**) via five successive exchanges in 100 kDa MWCO Amicon Ultra-0.5
429 mL centrifugal units. Buffers used consisted of 50 mM sodium phosphate and 150 mM NaCl, pH 1.5; 50
430 mM sodium phosphate and 150 mM NaCl, pH 1.9; 50 mM sodium phosphate and 150 mM NaCl, pH 2.1;
431 50 mM sodium phosphate and 150 mM NaCl, pH 2.25; 50 mM sodium phosphate and 150 mM NaCl, pH
432 3.0; 50 mM sodium citrate and 150 mM NaCl, pH 5.0; 50 mM MES and 150 mM NaCl, pH 6.0; 50 mM bis-
433 tris propane and 150 mM NaCl, pH 9.0; and 50 mM CHES and 150 mM NaCl, pH 10.0. Samples were
434 incubated at 4°C for six hours, and then immediately stained and imaged. Additional pH 2.25 and pH 3.0
435 samples were stored at 4°C for two days and stained and imaged. Negative stain transmission electron
436 microscopy (TEM) was carried out on the various samples with 200-mesh gold grids coated with extra
437 thick (25-50 nm) formvar-carbon film (EMS, USA) made hydrophilic by glow discharging at 5 mA for 60 s.
438 Briefly, 3.5 µL of sample was added to the grid and incubated for 30 seconds, wicked with filter paper,
439 and washed once with distilled water and once with 0.75% (w/v) uranyl formate before staining with 8.5
440 µL of uranyl formate for 30 seconds. TEM images were captured using a Morgagni transmission electron
441 microscope at 100 keV at the University of Michigan Life Sciences Institute. For all TEM experiments,
442 samples were roughly 0.2 mg/mL of AaEnc monomer in appropriate buffer.

443 Dynamic and static light scattering analyses

444 All sizing and polydispersity measurements were carried out on an Uncle by Unchained Labs (USA) at 30
445 °C in triplicate. Purified AaEnc samples were adjusted to 0.4 mg/mL of monomer in the appropriate
446 corresponding buffers and centrifuged at 10,000 rcf for 10 min, then immediately analyzed via DLS
447 (**Figure 2b**). Static light scattering aggregation temperature (T_{agg}) analysis was then conducted on
448 similarly prepared samples over a 25°C to 95°C ramp at 1°C per minute (**Figure 2d**).

449 Protease assays

450 AaEnc and bovine serum albumin (BSA; ThermoScientific Pierce, USA) were individually buffer
451 exchanged into Pepsin Assay Buffer (50 mM Na_2PO_4, 150 mM NaCl, pH 3.0) and mixed in a 40:1 molar

452 ratio with commercially purchased pepsin protease (Promega, USA), then incubated at 37°C for 3 h and
453 frozen until later use. Purified AaEnc and BSA were individually buffer exchanged into TBS pH 7.5 and
454 mixed in a 40:1:1 molar ratio with commercially purchased trypsin (Promega, USA) and chymotrypsin
455 (Promega, USA), then incubated at 37°C for 8 h and frozen until later use. All samples were then rapidly
456 thawed and examined via PAGE analysis.

457 Cryo-electron microscopy

458 *Sample preparation*

459 The purified protein samples were concentrated to 3 mg/mL in 150 mM NaCl, 20 mM Tris pH 7.5 or 150
460 mM NaCl, 50 mM Na_2PO_4 pH 3.0. 3.5 µL of protein samples were applied to freshly glow discharged
461 Quantifoil R1.2/1.3 grids and plunged into liquid ethane using an FEI Vitrobot Mark IV (100% humidity,
462 22°C, blot force 20, blot time 4 seconds, drain time 0 seconds, wait time 0 seconds). The frozen grids
463 were clipped and stored in liquid nitrogen until data collection.

464 *Data collection*

465 Cryo-electron microscopy movies were collected using a ThermoFisher Scientific Talos Arctica operating
466 at 200 keV equipped with a Gatan K2 Summit direct electron detector. Movies were collected at 45,000x
467 magnification using the Leginon[65] software package with a pixel size of 0.91 Å/pixel and an exposure
468 time of 5 or 8 s, frame time of 200 ms, and total dose of 42 e^-/A^2 for the pH 3.0 sample and 41 e^-/A^2 for
469 the pH 7.5 sample. 1,357 movies were collected for the pH 3 sample and 975 movies were collected for
470 the pH 7.5 sample.

471 *Data processing*

472 pH 7.5 sample: All data processing was performed using cryoSPARC v3.3.1.[60] 975 Movies were imported
473 and motion corrected using Patch Motion Correction and CTF fits were refined using Patch CTF. 821
474 movies with CTF fits better than 8.0 Å were selected for downstream processing. Roughly 200 particles
475 were picked manually using Manual Picker and grouped into 10 classes using 2D Classification. Well
476 resolved classes were selected and used as templates for Template Picker to pick particles with a
477 specified particle diameter of 240 Å. 56,583 particles with a box size of 384 pixels were extracted and
478 subjected to 3 rounds of 2D Classification with 100 classes yielding 44,686 particles in good classes. Ab-
479 Initio Reconstruction with 6 classes and I symmetry was carried out next. The two main classes were
480 selected representing the all "closed" (29,056 particles) and all "open" (13,581 particles) states. Particles
481 from each respective state were used as inputs for separate Homogenous Refinement jobs (with I or C1
482 symmetry) with the following settings: optimize per-particle defocus, optimize per-group CTF params,
483 and Ewald Sphere correction enabled. The I refinements yielded a 2.90 Å density for the "closed" state,
484 and a 3.32 Å density for the "open" state, whereas the C1 refinements resulted in 4.84 Å and 4.44 Å
485 maps, respectively (**Figure S3**). 3D Variability Analysis with 3 components was carried out on both
486 particle sets using the C1 Refinement results as inputs, however, no components could be resolved
487 corresponding to "open" and "closed" states within the same density.

488 pH 3.0 sample: The same preprocessing procedure was used as for the pH 7.5 sample yielding 52,210
489 extracted particles with a box size of 384 pixels. 3 rounds of 2D classification with 100 classes resulted in
490 47,599 good particles. Ab-Initio Reconstruction with 6 classes and I symmetry was carried out yielding
491 one dominant class containing 47,164 particles. This was followed by I and C1 Homogenous Refinement
492 jobs using the following parameters: optimized per-particle defocus, optimize per-group CTF params,
493 and Ewald Sphere correction. The I refinement resulted in a density of 2.77 Å while the C1 refinement
494 yielded a 4.12 Å map (**Figure S7**).

495 *Model building*

496 A homology model was generated using RoseTTAFold[66] on the Robetta server and was used as a starting
497 model for all model building efforts. This starting model was manually placed into the respective cryo-
498 EM maps using Chimera v1.16,[63] and was further fit using the Fit to Volume command. The placed
499 monomeric models were then manually refined against the respective cryo-EM maps using Coot
500 v8.9.6.[67] The resulting models were further refined using Real Space Refine in Phenix v 1.19.2-4158[68]
501 with default settings and three iterations. After inspecting the refined models in Coot, symmetry
502 restraints were pulled from the maps using the Phenix.Find_NCS_from_Map command with I symmetry.
503 Complete shell models were assembled using the Phenix.Build_from_NCS command. These shell models
504 were then used as inputs for a final round of Real Space Refine with NCS restraints, 3 iterations, and all
505 other settings set to default. The models were deposited to the PDB under PDB ID 8DN9, 8DNL, and
506 8DNA; and the EMDB under EMD-27558, EMD-27573, and EMD-27560.

507 Computational electrostatics and solvation analyses

508 *In silico* hydrogen bonding and buriedness analyses were conducted using the APBS-PDB2PQR software
509 suite (https://server.poissonboltzmann.org/pdb2pqr) with PROPKA v3.2 to predict pKa values and assign
510 protonation states at the provided pH values. Analyses were conducted using the "closed" and "open"
511 state AaEnc structures at pH 7.5 as well as the "closed" AaEnc structure at pH 3.0. The AaEnc protomer
512 was further analyzed using APBS-PDB2PQR to assess the calculated surface charge of AaEnc across
513 various pH values from pH 2.0 to pH 8.0 (**Figure 5a**).[40] Calculated protomer charge and stability
514 heatmaps were generated with the Protein-Sol webtool (https://protein-
515 sol.manchester.ac.uk/heatmap) using the AaEnc protomer from the "closed" state at pH 7.5 as the input
516 (**Figure 5b** and **Figure 5c**).[41] Monomer structure alignments were carried out in ChimeraX.[39]

517 *In vivo* cargo loading, native PAGE, and fluorescence analysis

518 The AaEnc encapsulated eGFP-TP sample was co-expressed, purified, and analyzed via TEM in the same
519 manner as AaEnc as described above (**Figure 6a and Figure 6b**). The free His-tagged eGFP and the
520 AaEnc-encapsulated eGFP-TP samples were concentrated to equimolar concentrations as determined by
521 densitometric analysis via SDS-PAGE using Fiji/ImageJ.[61] Empty AaEnc and eGFP-TP-loaded AaEnc were
522 similarly concentrated to equimolar concentrations for comparative NativePAGE analysis.

523 All NativePAGE analyses were conducted in an Invitrogen XCell SureLock using NativePAGE™ 4 to 16%
524 bis-tris mini protein gels and NativeMark Unstained Protein Standard from Fisher Scientific (USA) with 1x
525 running buffer made from 10x Tris/Glycine Buffer from Bio-Rad Laboratories, Inc. (USA). 20 µg of protein

526 was loaded per well, with effort to maintain equivalent amounts across all lanes for comparative
527 analysis. NativePAGE gels were run overnight at 65 V for 16.5 hours at 4°C. The following day, gels were
528 imaged via fluorescence imaging on a ChemiDoc Imaging System by Bio-Rad Laboratories, Inc. (USA),
529 then stained with ReadyBlue™ Protein Gel Stain from Sigma-Aldrich (USA) and imaged and analyzed.

530 His-tagged eGFP and AaEnc-encapsulated eGFP-TP were buffer exchanged into varying pH buffers from
531 pH 3.0 to pH 8.0 via five successive exchanges in 100 kDa MWCO Amicon Ultra-0.5 mL centrifugal units
532 and incubated at 22°C for three hours. Endpoint eGFP fluorescence (488 nm/507 nm) was then
533 measured in a BioTek Synergy H1 microplate reader at a final volume of 100 µL in Corning® 96-well flat
534 clear bottom black polystyrene microplates (**Figure 6c**).

535 **References**

536 1 Cornejo, E., Abreu, N. & Komeili, A. Compartmentalization and organelle formation in
537 bacteria. *Curr Opin Cell Biol* **26**, 132-138 (2014).
538 https://doi.org:10.1016/j.ceb.2013.12.007
539 2 Sutter, M. *et al.* Structural basis of enzyme encapsulation into a bacterial
540 nanocompartment. *Nat Struct Mol Biol* **15**, 939-947 (2008).
541 https://doi.org:10.1038/nsmb.1473
542 3 Jones, J. A. & Giessen, T. W. Advances in encapsulin nanocompartment biology and
543 engineering. *Biotechnol Bioeng* (2020). https://doi.org:10.1002/bit.27564
544 4 Andreas, M. P. & Giessen, T. W. Large-scale computational discovery and analysis of
545 virus-derived microbial nanocompartments. *Nat Commun* **12**, 4748 (2021).
546 https://doi.org:10.1038/s41467-021-25071-y
547 5 Contreras, H. *et al.* Characterization of a Mycobacterium tuberculosis nanocompartment
548 and its potential cargo proteins. *J Biol Chem* **289**, 18279-18289 (2014).
549 https://doi.org:10.1074/jbc.M114.570119
550 6 Giessen, T. W. & Silver, P. A. Widespread distribution of encapsulin nanocompartments
551 reveals functional diversity. *Nat Microbiol* **2**, 17029 (2017).
552 https://doi.org:10.1038/nmicrobiol.2017.29
553 7 McHugh, C. A. *et al.* A virus capsid-like nanocompartment that stores iron and protects
554 bacteria from oxidative stress. *EMBO J* **33**, 1896-1911 (2014).
555 https://doi.org:10.15252/embj.201488566
556 8 He, D. *et al.* Structural characterization of encapsulated ferritin provides insight into iron
557 storage in bacterial nanocompartments. *Elife* **5** (2016).
558 https://doi.org:10.7554/eLife.18972
559 9 He, D. *et al.* Conservation of the structural and functional architecture of encapsulated
560 ferritins in bacteria and archaea. *Biochem J* **476**, 975-989 (2019).
561 https://doi.org:10.1042/BCJ20180922
562 10 Giessen, T. W. *et al.* Large protein organelles form a new iron sequestration system with
563 high storage capacity. *Elife* **8** (2019). https://doi.org:10.7554/eLife.46070
564 11 Nichols, R. J. *et al.* Discovery and characterization of a novel family of prokaryotic
565 nanocompartments involved in sulfur metabolism. *Elife* **10** (2021).
566 https://doi.org:10.7554/eLife.59288

12 Giessen, T. W. & Silver, P. A. Converting a Natural Protein Compartment into a Nanofactory for the Size-Constrained Synthesis of Antimicrobial Silver Nanoparticles. *ACS Synth Biol* **5**, 1497-1504 (2016). https://doi.org:10.1021/acssynbio.6b00117

13 Sigmund, F. *et al.* Bacterial encapsulins as orthogonal compartments for mammalian cell engineering. *Nat Commun* **9**, 1990 (2018). https://doi.org:10.1038/s41467-018-04227-3

14 Jenkins, M. C. & Lutz, S. Encapsulin Nanocontainers as Versatile Scaffolds for the Development of Artificial Metabolons. *ACS Synth Biol* **10**, 857-869 (2021). https://doi.org:10.1021/acssynbio.0c00636

15 Moon, H., Lee, J., Min, J. & Kang, S. Developing genetically engineered encapsulin protein cage nanoparticles as a targeted delivery nanoplatform. *Biomacromolecules* **15**, 3794-3801 (2014). https://doi.org:10.1021/bm501066m

16 Choi, B. *et al.* Effective Delivery of Antigen-Encapsulin Nanoparticle Fusions to Dendritic Cells Leads to Antigen-Specific Cytotoxic T Cell Activation and Tumor Rejection. *ACS Nano* **10**, 7339-7350 (2016). https://doi.org:10.1021/acsnano.5b08084

17 Altenburg, W. J., Rollins, N., Silver, P. A. & Giessen, T. W. Exploring targeting peptide-shell interactions in encapsulin nanocompartments. *Sci Rep* **11**, 4951 (2021). https://doi.org:10.1038/s41598-021-84329-z

18 Jones, J. A., Cristie-David, A. S., Andreas, M. P. & Giessen, T. W. Triggered Reversible Disassembly of an Engineered Protein Nanocage*. *Angew Chem Int Ed Engl* **60**, 25034-25041 (2021). https://doi.org:10.1002/anie.202110318

19 Adamson, L. S. R. *et al.* Pore structure controls stability and molecular flux in engineered protein cages. *Sci Adv* **8**, eabl7346 (2022). https://doi.org:10.1126/sciadv.abl7346

20 Williams, E. M., Jung, S. M., Coffman, J. L. & Lutz, S. Pore Engineering for Enhanced Mass Transport in Encapsulin Nanocompartments. *ACS Synth Biol* **7**, 2514-2517 (2018). https://doi.org:10.1021/acssynbio.8b00295

21 Putri, R. M. *et al.* Structural Characterization of Native and Modified Encapsulins as Nanoplatforms for in Vitro Catalysis and Cellular Uptake. *ACS Nano* **11**, 12796-12804 (2017). https://doi.org:10.1021/acsnano.7b07669

22 Sharma, A., Kawarabayasi, Y. & Satyanarayana, T. Acidophilic bacteria and archaea: acid stable biocatalysts and their potential applications. *Extremophiles* **16**, 1-19 (2012). https://doi.org:10.1007/s00792-011-0402-3

23 Rohwerder, T., Gehrke, T., Kinzler, K. & Sand, W. Bioleaching review part A: progress in bioleaching: fundamentals and mechanisms of bacterial metal sulfide oxidation. *Appl Microbiol Biotechnol* **63**, 239-248 (2003). https://doi.org:10.1007/s00253-003-1448-7

24 Pande, V., Pandey, S. C., Sati, D., Bhatt, P. & Samant, M. Microbial Interventions in Bioremediation of Heavy Metal Contaminants in Agroecosystem. *Front Microbiol* **13**, 824084 (2022). https://doi.org:10.3389/fmicb.2022.824084

25 Reed, C. J., Lewis, H., Trejo, E., Winston, V. & Evilia, C. Protein adaptations in archaeal extremophiles. *Archaea* **2013**, 373275 (2013). https://doi.org:10.1155/2013/373275

26 Dumorné, K., Córdova, D. C., Astorga-Eló, M. & Renganathan, P. Extremozymes: A Potential Source for Industrial Applications. *J Microbiol Biotechnol* **27**, 649-659 (2017). https://doi.org:10.4014/jmb.1611.11006

27 Giessen, T. W. Encapsulins. *Annu Rev Biochem* (2022). https://doi.org:10.1146/annurev-biochem-040320-102858

611 28 Rodríguez, J. M., Allende-Ballestero, C., Cornelissen, J. J. L. M. & Castón, J. R.
612 Nanotechnological Applications Based on Bacterial Encapsulins. *Nanomaterials (Basel)*
613 **11** (2021). https://doi.org:10.3390/nano11061467
614 29 Shrestha, R., Huang, G., Meekins, D. A., Geisbrecht, B. V. & Li, P. Mechanistic Insights
615 into Dye-Decolorizing Peroxidase Revealed by Solvent Isotope and Viscosity Effects.
616 *ACS Catal* **7**, 6352-6364 (2017). https://doi.org:10.1021/acscatal.7b01861
617 30 Baker-Austin, C. & Dopson, M. Life in acid: pH homeostasis in acidophiles. *Trends*
618 *Microbiol* **15**, 165-171 (2007). https://doi.org:10.1016/j.tim.2007.02.005
619 31 Guan, N. *et al.* Understanding of how Propionibacterium acidipropionici respond to
620 propionic acid stress at the level of proteomics. *Sci Rep* **4**, 6951 (2014).
621 https://doi.org:10.1038/srep06951
622 32 Parizzi, L. P. *et al.* The genome sequence of Propionibacterium acidipropionici provides
623 insights into its biotechnological and industrial potential. *BMC Genomics* **13**, 562 (2012).
624 https://doi.org:10.1186/1471-2164-13-562
625 33 Hsu, S. T. & Yang, S. T. Propionic acid fermentation of lactose by Propionibacterium
626 acidipropionici: effects of pH. *Biotechnol Bioeng* **38**, 571-578 (1991).
627 https://doi.org:10.1002/bit.260380603
628 34 Ross, J. *et al.* Pore dynamics and asymmetric cargo loading in an encapsulin
629 nanocompartment. *Sci Adv* **8**, eabj4461 (2022). https://doi.org:10.1126/sciadv.abj4461
630 35 Tang, Y. *et al.* Cryo-EM structure of Mycobacterium smegmatis DyP-loaded encapsulin.
631 *Proc Natl Acad Sci U S A* **118** (2021). https://doi.org:10.1073/pnas.2025658118
632 36 Loncar, N., Rozeboom, H. J., Franken, L. E., Stuart, M. C. A. & Fraaije, M. W. Structure
633 of a robust bacterial protein cage and its application as a versatile biocatalytic platform
634 through enzyme encapsulation. *Biochem Biophys Res Commun* **529**, 548-553 (2020).
635 https://doi.org:10.1016/j.bbrc.2020.06.059
636 37 Ross, J. *et al.* Pore dynamics and asymmetric cargo loading in an encapsulin
637 nanocompartment. *Sci Adv* **8**, eabj4461 (2022). https://doi.org:10.1126/sciadv.abj4461
638 38 Robert, X. & Gouet, P. Deciphering key features in protein structures with the new
639 ENDscript server. *Nucleic Acids Res* **42**, W320-324 (2014).
640 https://doi.org:10.1093/nar/gku316
641 39 Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and
642 analysis. *Protein Sci* **27**, 14-25 (2018). https://doi.org:10.1002/pro.3235
643 40 Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite.
644 *Protein Sci* **27**, 112-128 (2018). https://doi.org:10.1002/pro.3280
645 41 Hebditch, M. & Warwicker, J. Web-based display of protein surface and pH-dependent
646 properties for assessing the developability of biotherapeutics. *Sci Rep* **9**, 1969 (2019).
647 https://doi.org:10.1038/s41598-018-36950-8
648 42 Kamiyama, D. *et al.* Versatile protein tagging in cells with split fluorescent protein. *Nat*
649 *Commun* **7**, 11046 (2016). https://doi.org:10.1038/ncomms11046
650 43 Shinoda, H., Shannon, M. & Nagai, T. Fluorescent Proteins for Investigating Biological
651 Events in Acidic Environments. *Int J Mol Sci* **19** (2018).
652 https://doi.org:10.3390/ijms19061548
653 44 Shinoda, H. *et al.* Acid-Tolerant Monomeric GFP from Olindias formosa. *Cell Chem Biol*
654 **25**, 330-338.e337 (2018). https://doi.org:10.1016/j.chembiol.2017.12.005

45 Menon, B. B., Heinhorst, S., Shively, J. M. & Cannon, G. C. The carboxysome shell is permeable to protons. *J Bacteriol* **192**, 5881-5886 (2010). https://doi.org:10.1128/JB.00903-10

46 Penrod, J. T. & Roth, J. R. Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. *J Bacteriol* **188**, 2865-2874 (2006). https://doi.org:10.1128/JB.188.8.2865-2874.2006

47 Burykin, A. & Warshel, A. What really prevents proton transport through aquaporin? Charge self-energy versus proton wire proposals. *Biophys J* **85**, 3696-3706 (2003). https://doi.org:10.1016/S0006-3495(03)74786-9

48 Tokmakov, A. A., Kurotani, A. & Sato, K. I. Protein pI and Intracellular Localization. *Front Mol Biosci* **8**, 775736 (2021). https://doi.org:10.3389/fmolb.2021.775736

49 Lambré, C. *et al.* Safety evaluation of the food enzyme rennet paste from the abomasum of suckling goats, lambs and calves. *EFSA J* **19**, e07006 (2021). https://doi.org:10.2903/j.efsa.2021.7006

50 Salvachúa, D., Prieto, A., Martínez, Á. & Martínez, M. J. Characterization of a novel dye-decolorizing peroxidase (DyP)-type enzyme from Irpex lacteus and its application in enzymatic hydrolysis of wheat straw. *Appl Environ Microbiol* **79**, 4316-4324 (2013). https://doi.org:10.1128/AEM.00699-13

51 Parashar, D. & Satyanarayana, T. Engineering a chimeric acid-stable α-amylase-glucoamylase (Amy-Glu) for one step starch saccharification. *Int J Biol Macromol* **99**, 274-281 (2017). https://doi.org:10.1016/j.ijbiomac.2017.02.083

52 Razzaq, A. *et al.* Microbial Proteases Applications. *Front Bioeng Biotechnol* **7**, 110 (2019). https://doi.org:10.3389/fbioe.2019.00110

53 Mohanty, A. K., Mukhopadhyay, U. K., Grover, S. & Batish, V. K. Bovine chymosin: production by rDNA technology and application in cheese manufacture. *Biotechnol Adv* **17**, 205-217 (1999). https://doi.org:10.1016/s0734-9750(99)00010-5

54 Xu, Q. S., Yan, Y. S. & Feng, J. X. Efficient hydrolysis of raw starch and ethanol fermentation: a novel raw starch-digesting glucoamylase from. *Biotechnol Biofuels* **9**, 216 (2016). https://doi.org:10.1186/s13068-016-0636-5

55 Johnston, D. B. & McAloon, A. J. Protease increases fermentation rate and ethanol yield in dry-grind ethanol production. *Bioresour Technol* **154**, 18-25 (2014). https://doi.org:10.1016/j.biortech.2013.11.043

56 Li, L. *et al.* Revealing two important tryptophan residues with completely different roles in a dye-decolorizing peroxidase from Irpex lacteus F17. *Biotechnol Biofuels* **14**, 128 (2021). https://doi.org:10.1186/s13068-021-01978-y

57 Wiltschi, B. *et al.* Enzymes revolutionize the bioproduction of value-added compounds: From enzyme discovery to special applications. *Biotechnol Adv* **40**, 107520 (2020). https://doi.org:10.1016/j.biotechadv.2020.107520

58 Theillet, F. X. *et al.* Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem Rev* **114**, 6661-6714 (2014). https://doi.org:10.1021/cr400695p

59 Li, S. S. *et al.* Monitoring the Changes of pH in Lysosomes during Autophagy and Apoptosis by Plasmon Enhanced Raman Imaging. *Anal Chem* **91**, 8398-8405 (2019). https://doi.org:10.1021/acs.analchem.9b01250

60 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017). https://doi.org:10.1038/nmeth.4169

61 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012). https://doi.org:10.1038/nmeth.2019

62 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010). https://doi.org:10.1107/S0907444909052925

63 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004). https://doi.org:10.1002/jcc.20084

64 Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-W641 (2019). https://doi.org:10.1093/nar/gkz268

65 Suloway, C. *et al.* Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151**, 41-60 (2005). https://doi.org:10.1016/j.jsb.2005.03.010

66 Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876 (2021). https://doi.org:10.1126/science.abj8754

67 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **66**, 486-501 (2010). https://doi.org:10.1107/S0907444910007493

68 Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213-221 (2010). https://doi.org:10.1107/S0907444909052925

Acknowledgements

We gratefully acknowledge funding from the NIH (R35GM133325). Research reported in this publication was supported by the University of Michigan Cryo-EM Facility (U-M Cryo-EM). U-M Cryo-EM is grateful for support from the U-M Life Sciences Institute and the U-M Biosciences Initiative. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Author Contributions

J.A.J. and T.W.G. designed the project. J.A.J. conducted the laboratory experiments and negative stain transmission electron microscopy, while M.P.A. collected and analyzed cryo-EM data. J.A.J. and M.P.A. built the AaEnc structural models. J.A.J. wrote the manuscript. T.W.G. processed cryo-EM data, edited the manuscript and oversaw the project in its entirety.

Competing Interests

The authors declare no competing interests.

Additional Information

Correspondence and requests for materials should be addressed to T.W.G.