Supplementary information

Biocatalytic application and structural elucidation of robust bacterial protein nanocages

Ognjen Pećanac, [#a] Alexander Belyy, [#*b] Caterina Martin, [a] Rosalie Kleissen, [a] Marco W. Fraaije *[c] and Nikola Lončar *[a]

[a] GECCO Biotech, Zernikepark 6, 9747 AN Groningen, The Netherlands

[b] Membrane Enzymology, University of Groningen, Nijenborgh 3, 9747 AG Groningen, The Netherlands

[c] Molecular Enzymology, University of Groningen, Nijenborgh 3, 9747 AG Groningen, The Netherlands

[#] Equally contributed

[*] Corresponding authors: a.belyy@rug.nl, m.w.fraaije@rug.nl, n.loncar@gecco-biotech.com

This additional file contains:

Tables: S1-S4

Figures: S1-S4

Equation: S1

Table S1. Sequences of *E. coli* codon optimized encapsulin and cargo enzyme encoding genes.

Name of the gene	Sequence ‘3 - ‘5
CyanoPOX-cargo	ATGGGTAGCAGCCACCACCATCACCATCACTCAAGCGGGCTTGTCCCCCGTGGCTCACATGCTCAGAACCGTACTATTGATGGGA CCGACAACAACCTGTCACATAATGTTTGGGGGTCAACAAACCAACATTTAGATCGCGCAGGTCCTGCTGCATACGCAGACGGCAT GTCAATGCCAGCCGGAGGTTCGCGCCCCTCCGCACGCGCAGTCAGTAACGGGATTGCCGCGCAGACGGGTTCTATGTTGAATGA CCGTATGTTGAGTGATTGGGTATGGCAGTGGGGTCAATTCCTTGATCATGACCTGGACTTGACGGATGCAGCGAGTCCCGCGGA GAGCTTCCCAATCCCAGTTCCAATGGGTGACCCTTTCTTCGACCCCTTCAATACGGGAACCCAGACCATCGGGCTTAGCCGTTCAG CCTATGATCCTGCTACCGGATCTGTTGATGCCCGCCAGCAGATGAACCAAATCACAAGTTGGATCGATGCGTCGAATGTATACGG GTCTGATATGACTCGCGCTAACGCTCTTCGCACTATGAGCGGAGGTCGTTTAGCTACTTCAGCTGGGGATCTGCTTCCCTTCAATA CAGGGGGGCTGCCCAATGCAGGTGGGACTTCGCCAAGTCTGTTCTTGGCCGGAGACGTGCGTTCGAACGAACAGAGTGGCCTTG CTGCCGTACATACCTTATTTGTACGCGAGCATAACCGTCTGGCGGACCAAATTGCTGCAGCGAACCCCGGTATGGGAGATGAGGA CATCTACCAGCAAGCTCGCAAGATCGTGGGAGCCCAGATGCAAATCATTACGTATAATGAATTTTTGCCGGCTTTATTGGGTTCAG CCGCTCCTTCGCCTATGTCGATCGGCTATGATGACTCAATCAATCCAAACATCATGAATGAATTTGCGAACGCGTGCTACCGCGTC GGTCACACCATGCTGTCGCCGACCATTTTGCGTTTAGATAACGCTGGGAACGTAATTCCCCACGGTAATTTAGCACTGCAGGACG CGTTTTTTAACCCCAATCGTATCATTAATGAAGGGGGTATCGCCCCCATTTTAAAAGGCCTGGCCTCTCAGGCCATGCAGGAGATC GACAACAAGATCGTCGACGACGTTCGTAATTTTCTGTTTGGGCCACCAGGATCTGGCGGCCTTGATCTGGCGAGCTTGAATATTC AACGTGGTCGTGACCATGGTTTACCGGACTACAATTCGACGCGTGTAATGATGGGCTTGACCTCTGTATCTAGCTTCGCTGACATT TCTTCTGACCCTGCAGTCCAAGCCGCGTTAATGTCATTATATGGAACCGTCAATGATATTGATCTGTGGGTCGGGGCGCTTGCGG AGGACCACCTGGCTGGCTCAAGCGTAGGCGAGCTTATTGCAGCTGTCCTTGGCGAGCAGTTTACTCGTCTGCGTGATGGAGACC GTTATTGGTATGAGCGCGACGATTTTTTCGTTAACAACCCATCGTTACTGGCAGAGCTTCAGGCAACTCGTTTAAGTGACATTATT CGTCGTAATAGTGACATCACGAATATTCAGGACAATGTTTTTCTTATCCCTGAACCAGCCACCCTGGGGTTATTGATGTTCGGGGC TGCTTTTTTGCGTAAACGTCGCTCTCCACCTCCCCTTCCCGACAGTGAGCCGGACCGTGAGATCCCGGCCGATGACGGCTCCTTGG GCATCGGTTCTTTGAAGGGTACACGCTCA
PTDH-mFMO-cargo	ATGGGAAGTTCACATCACCATCATCACCACAGTTCCGGCTTGGTCCCTCGTGGGTCTCACATGTTGCCCAAATTAGTAATCACGCA CCGTGTTCACGAGGAAATTCTTCAGCTGCTGGCTCCTCACTGCGAATTGATCACGAACCAGACTGATTCAACCTTAACACGTGAGG AGATTTTGCGTCGCTGTCGTGATGCTCAGGCTATGATGGCCTTCATGCCGGATCGCGTTGACGCCGATTTTCTTCAAGCTTGTCCC GAATTACGCGTTATTGGATGCGCACTGAAAGGCTTTGACAATTTCGACGTAGATGCCTGCACAGCGCGTGGAGTTTGGCTGACTT TTGTACCGGACCTGTTAACTGTCCCTACAGCAGAGCTTGCGATCGGCCTGGCGGTGGGATTGGGGCGTCATTTGCGTGCGGCGG ACGCTTTCGTGCGCTCAGGTAAATTTCGCGGCTGGCAGCCGCGCTTTTACGGTACGGGATTAGACAACGCCACAGTCGGGTTTCT TGGAATGGGTGCCATTGGCTTAGCTATGGCCGATCGTTTGCAAGGGTGGGGGGCAACTTTGCAATACCATGCCCGCAAGGCATT GGACACGCAAACTGAGCAACGTTTAGGGCTTCGTCAAGTAGCGTGCTCCGAGCTTTTCGCGAGTAGTGATTTTATTCTTCTTGCTC TTCCGCTTAATGCAGACACGCTGCACCTTGTAAACGCGGAACTGCTTGCTCTTGTACGTCCAGGCGCCTTACTGGTCAACCCGTGC CGTGGCAGCGTCGTCGATGAGGCGGCTGTGCTGGCTGCTCTGGAACGCGGACAACTGGGTGGCTATGCAGCGGATGTTTTCGAG ATGGAGGATTGGGCCCGTGCGGACCGTCCTCAACAGATTGACCCGGCACTTCTTGCACACCCGAACACGTTGTTTACTCCCCACAT CGGGAGTGCGGTCCGCGCAGTTCGCTTAGAGATCGAACGCTGCGCCGCGCAAAACATTCTGCAGGCCCTGGCTGGGGAACGCCC GATTAATGCTGTGAATCGCCTTCCAAAAGCCAACCCGGCAGCCGATTCACGTTCTGCTGCTGGCATGGCGACGCGTATTGCAATCT TGGGTGCTGGGCCATCAGGGATGGCTCAATTGCGCGCATTCCAATCAGCGCAGGAGAAGGGGGCAGAAATCCCAGAACTGGTA TGTTTTGAAAAACAGGCGGATTGGGGAGGTCAGTGGAACTACACTTGGCGCACGGGTTTAGACGAGAACGGCGAGCCCGTACA CTCCTCCATGTACCGCTATCTGTGGTCAAATGGCCCAAAGGAGTGTCTTGAGTTTGCCGATTACACATTCGATGAGCACTTTGGCA AACCGATTGCCTCGTATCCGCCACGCGAGGTACTGTGGGATTATATCAAAGGGCGTGTAGAGAAAGCCGGAGTCCGTAAATATA TTCGTTTCAATACAGCGGTGCGCCATGTCGAATTCAATGAAGACTCCCAGACCTTTACTGTCACGGTGCAAGATCATACGACAGAC ACTATCTATAGCGAGGAGTTTGATTATGTGGTCTGCTGTACCGGACATTTTTCGACGCCATATGTTCCTGAATTTGAAGGGTTCGA GAAGTTCGGAGGCCGCATCTTGCATGCACACGACTTTCGTGACGCTTTGGAATTCAAAGACAAAACCGTGCTGTTAGTTGGCAGT TCCTACTCAGCTGAGGATATCGGAAGCCAGTGCTATAAGTATGGTGCCAAGAAATTAATTTCATGTTATCGTACGGCGCCTATGG GATATAAGTGGCCCGAAAACTGGGATGAGCGCCCCAACTTAGTCCGTGTCGATACTGAAAATGCATATTTCGCTGACGGCAGCTC GGAGAAAGTCGACGCCATCATCTTGTGCACGGGGTATATTCATCACTTTCCATTTCTTAACGATGATCTGCGTTTAGTAACAAATA ATCGCCTTTGGCCGCTTAACTTGTACAAAGGGGTTGTATGGGAAGACAATCCCAAATTCTTCTACATTGGAATGCAAGACCAATG GTATTCATTTAACATGTTCGATGCCCAGGCGTGGTACGCGCGCGACGTTATTATGGGCCGTTTACCCTTGCCCTCAAAAGAAGAG ATGAAGGCCGACTCTATGGCTTGGCGCGAAAAAGAGCTTACTTTGGTCACTGCTGAGGAGATGTATACATATCAAGGGGACTAT ATCCAAAATTTAATTGATATGACTGACTACCCCAGTTTTGATATTCCGGCCACGAATAAGACCTTCTTAGAGTGGAAGCATCATAA GAAAGAAAATATCATGACATTCCGTGACCATAGTTACCGCTCCCTTATGACCGGTACTATGGCACCCAAACATCATACACCCTGGA TTGATGCTTTAGACGACTCATTAGAAGCGTACTTATCGGATAAGTCGGAAATTCCGGTGGCAAAGGAAGCCCCACCCCCCTTACC CGACTCGGAGCCCGACCGTGAGATTCCGGCAGATGACGGTTCTCTGGGAATTGGGTCGTTAAAAGGGACCCGTTCT

DendroEnc	ATGGACTTTCTTGATCGTAACGCGGCCCCTTTAAATGAAGGTGAGTGGCAGCGCATTGACGAGGCAGTCGTCTCGACAGCGCGTC GTACACTTGTCGCTCGCCGCATCATTGACGTTTTAGGGCCCTTAGGCTCGGGTGTTTACAGCATCCCATACTCCGTTTTCTCAGGTA AGTCACCTACGGGAATTGACATGGTCGGAGAGAACGAAGAGTTTGTGGTAGAGGCTTCGCGCCGCGCTACGATCAATCTGCCGA TCTTGTACAAGGACTTTAAGATTATGTGGCGCGACGTCGAGGCAGACCGTCACTTGGGGCTGCCCATTGACGTCAGCACAGCAGC AGTAGCAGGGAACTTTGTCGCGGTACAGGAAGATCACCTGATCTTCAACGGTAATGTAGAGCTTGGGCATGATGGCCTTTTTACG GTGAAAGACCGCCAGACTGTTGCCATTTCAGATTGGAACGAAACGGGTGCAGCGCTTGCTGACGTGGTGAAGGCGGTAGGAGC TTTGTCTCAAGCAGGTCACTACGGACCTTACGCTATGGTCGTTTCCCCGGTTTTGTTTGGGCGTATGATTCGCGTATTTGGGAATA CGGGTATGCTGGAATTAGACCAAGTTAAGGCATTGATTACTGGCGGGGTGCATTATAGCAACGTGATTGGAGGAAGTAAAGCTG TTGTCGTCGCCACCGGGTCACAGAACCTTAATCTGGCATTAGGACAAGACATGGTCACAGCATACATGGGACCGACTTCTATGAA CCACGTTTTTCGCGTCTTAGAGACCGCCGCATTATTAGTGCGCCGCCCCGACGCCATCTGCACCATCGAA
ArthroEnc	ATGAATAATCTGTACCGCGAATTAGCACCAGTCACAGACGAGGCCTGGGGTGAAATTGAACAAGAAGCTACCCGCACTTTCAAG CGTCACATTGCAGGCCGCCGCGTGGTAGACTGTACGGGACCTACTGGTGCCACTACCGCTGCAGTATCTACGGGTCATCTTCTGG ACGTAACATCACCTGGAGATGGGGTTATTGCTCACTTGCGCGATTCTCGCCCCCTTGTGCGTTTGCGTGTTCCTTTCACCATTACAC GTGATGCTATTGATGATGTAGAGCGCGGGAGCCAAGACTCCGATTGGGATCCGGTTAAGGATGCTGCCAAAAAGTTAGCGTTTG CCGAAGATCGTGCCATCTTCGAGGGGTATGCAGCTGCGAATATTCAAGGGATTCGCGCAGGTAGTAGCAATCCAGGTCTTGCACT TCCTGACGACCCCCGTGGTTATCCAGACGTTATCGCTCAGGCTCTTAGTGAATTGCGCCTTGCAGGAGTTGATGGACCCTACAGC GTCCTTCTGAGCGCCGACGCATACACTAAGGTCAGTGAGACGACGGAGCACGGGTACCCACTGCGCGAACACCTTAATCGCGTA GTTGACGGTGACATTATTTGGGCACCTGCAATCGATGGTGCGTTCGTTCTTTCAACGCGCGGTGGAGACTTCGACTTACAATTAG GTACGGATGTATCTATCGGTTATTCATCTCATGATGCCGACACGGTGCAGTTGTATCTGCAGGAGACATTTACCTTTTTGAATTAC ACAGCTGAAGCAAGTGTGGCGCTGTCTGCA

Table S2. Primers used for DendroEnc N198W mutant generation.

Primer	Sequence ‘3 - ‘5
DendroEnc N198W Fw	TCGGATGGACTGGCATGCTGGAACTGGAT
DendroEnc N198W Rv	GCATGCCAGTCCATCCGAAGACGCGAATCATGCG

Table S3. Cryo-EM data collection, refinement, and validation statistics.

	ArthroEnc	**WT DendroEnc**	**N198W DendroEnc**
Microscope	Talos Arctica		
Voltage (kV)	200		
Defocus range (μm)	-0.5 to -2.5		
Camera	Gatan K3 Superresolution mode		
Pixel size (Å)	1.04		
Total electron dose (e/Å^2)	45		
Exposure time (s)	3.5		
Frames per movie	50		
Number of movies	719	4498	7874
3D Refinement			
Number of particles	15677	4549	3997
Symmetry applied	Icosahedral (I)		
Final map pixel size	1.04	1.59	1.59
Final resolution (Å)	2.9	3.4	3.6
Atomic model statistics*			
Non-hydrogen atoms	2017	8436	8464
Number of chains	1	4	4
Molprobity score	0.89	1.29	1.46
Rama distribution Z-score	-1.14 ± 0.50	-1.96 ± 0.22	-1.73 ± 0.23
Clashscore	1.51	2.13	3.43
Bond RMSD (Å)	0.003	0.002	0.003
Angle RMSD (°)	0.573	0.608	0.646
Poor rotamers (%)	0	0	0
Favored rotamers (%)	100	99.78	99.78
Ramachandran favored (%)	98.1	95.56	95.29
Ramachandran allowed (%)	1.9	4.44	4.71
Ramachandran outliers (%)	0	0	0

* Statistics given for the deposited PDB model with one asymmetric unit.

Table S4. Specific activity raw data of DendroEnc CyanoPOX constructs in the presence of ABTS and guaiacol.

Enzyme	DendroEnc CyanoPOX	DendroEnc N198W CyanoPOX	DendroEnc CyanoPOX
Substrate	ABTS		guaiacol
Concentration (mM)	10		
mAbs/t (mAbs /min)	714.140	132.675	15.540
	729.740	138.225	15.990
Abs/t (Abs/min)	0.7141	0.1327	0.0155
	0.7297	0.1382	0.0160
$\frac{\Delta Abs/\Delta t}{\varepsilon * l}$ **(rate of substrate conversion (mM/min))**	0.0198	0.0037	0.0006
	0.0203	0.0038	0.0006
U/mg	0.0198	0.0037	0.0006
	0.0203	0.0038	0.0006
average U/mg	0.0201	0.0038	0.0006
standard deviation U/mg	0.00031	0.00011	0.00001

$$Specific\ activity\left(\frac{U}{mg}\right) = \frac{\left[\left(\left(\frac{\Delta Abs/\Delta t}{\varepsilon * l}\right)\right) \times\ reaction\ volume\ (mL)\right]}{protein\ concentration\ (\frac{mg}{mL})}$$

Equation S1. Equation used to calculate the specific activity of DendroEnc CyanoPOX constructs in the presence of ABTS and guaiacol. ε (molar extinction coefficient ($mM^{-1}\ cm^{-1}$). l (path length (cm)).

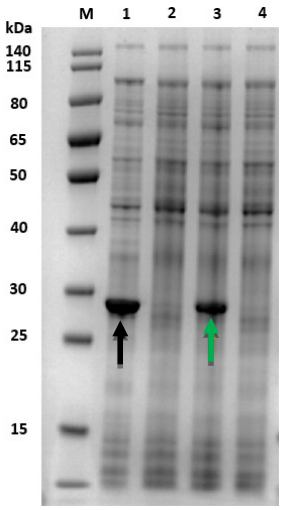

Figure S1. SDS-PAGE analysis of total *E. coli* lysates demonstrating overexpression of encapsulin proteins. Lane M: molecular weight marker. Lane 1: induced expression of ArthroEnc. Lane 2: uninduced ArthroEnc (negative control). Lane 3: induced expression of DendroEnc. Lane 4: uninduced DendroEnc (negative control). The encapsulin bands are indicated by black (ArthroEnc) and green (DendroEnc) arrows.

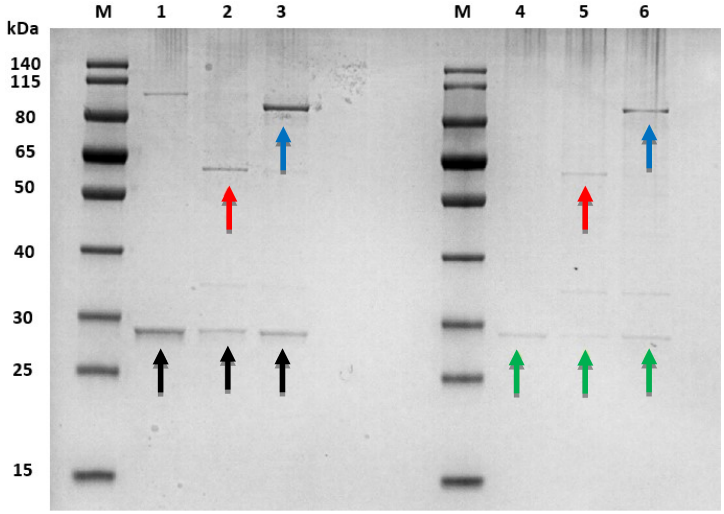

Figure S2. SDS-PAGE gel of purified DendroEnc and ArthroEnc constructs after SEC. Lane M: molecular weight marker. Lane 1: DendroEnc (black arrow). Lane 2: DendroEnc with encapsulated CyanoPOX. CyanoPOX is shown with the red arrow. Lane 3: DendroEnc with encapsulated PTDH-mFMO. PTDH-mFMO is shown by the blue arrow. Lane 4: ArthroEnc (green arrow). Lane 5: ArthroEnc with encapsulated CyanoPOX. CyanoPOX is shown with the red arrow. Lane 6: ArthroEnc with encapsulated PTDH-mFMO. PTDH-mFMO is shown by the blue arrow.

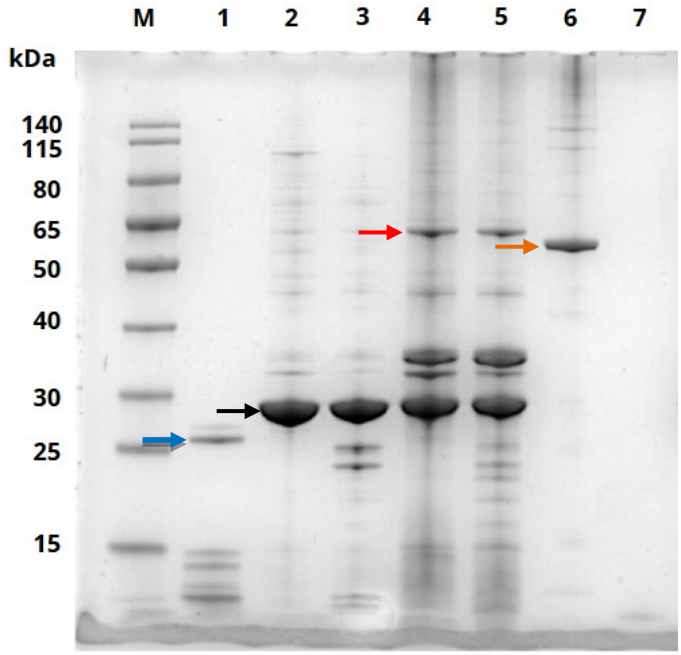

Figure S3. SDS-PAGE gel showing proteolytic stability of empty DendroEnc, CyanoPOX-loaded DendroEnc, and free CyanoPOX after incubation with chymotrypsin. Lane M: molecular weight marker. Lane 1: chymotrypsin, 20x amount compared with the digestions shown in lane 3, 5 and 7. (blue arrow). Lane 2: DendroEnc (black arrow: encapsulin protein). Lane 3: DendroEnc incubated with chymotrypsin. Lane 4: DendroEnc with encapsulated CyanoPOX (red arrow: CyanoPOX with encapsulin targeting sequence). Lane 5: DendroEnc with encapsulated CyanoPOX incubated with chymotrypsin. Lane 6: CyanoPOX (red arrow: CyanoPOX). Lane 7: CyanoPOX incubated with chymotrypsin.

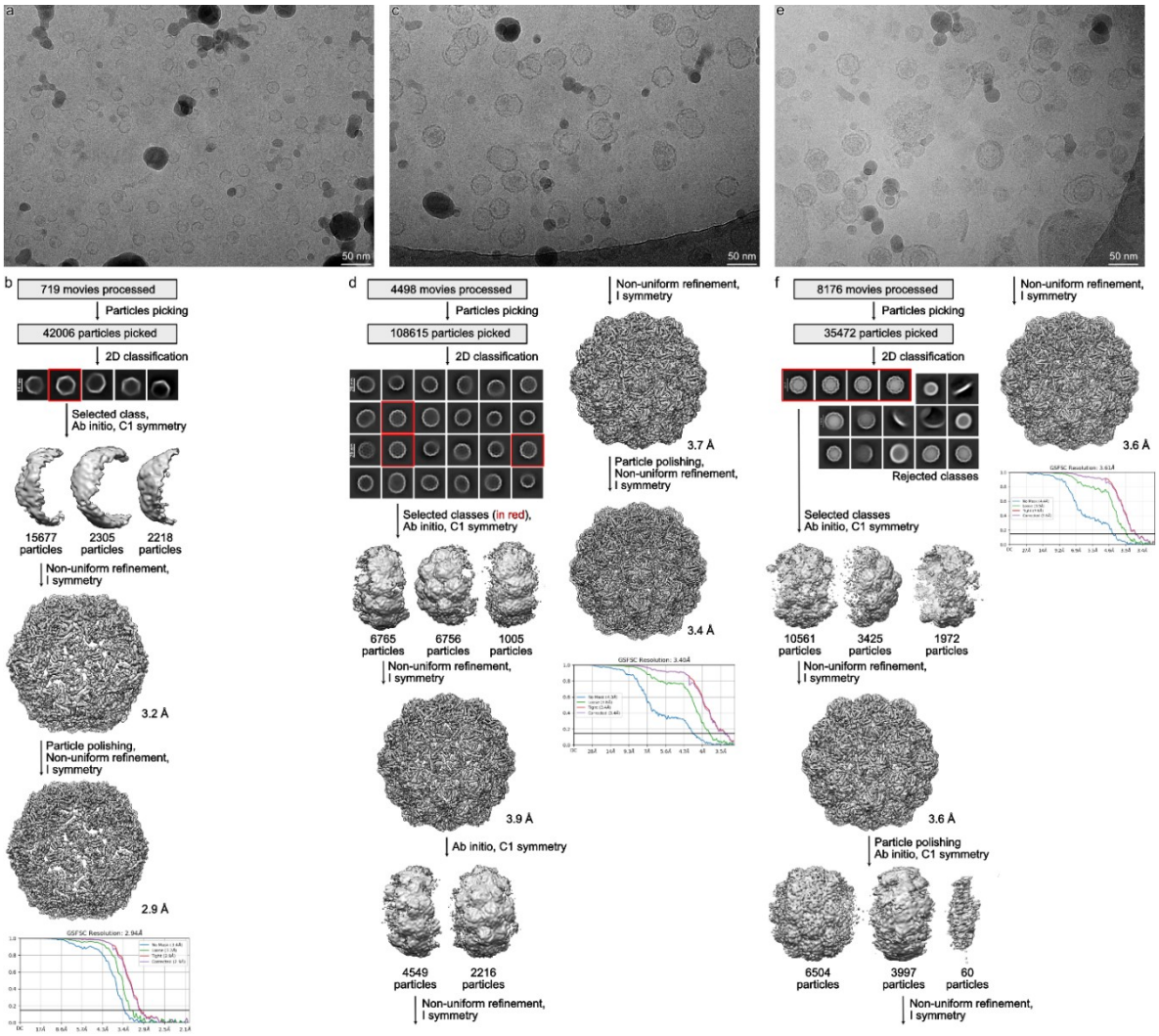

Figure S4. Cryo-EM data processing. (a), (c) and (e) Exemplary cryo-EM micrographs of ArthroEnc, WT DendroEnc and N198W DendroEnc, respectively. (b), (d) and (f) Processing overview and fourier shell correlation (FSC) curves of final reconstructions for ArthroEnc, WT DendroEnc and N198W DendroEnc, respectively.