PNAS
www.pnas.org

Supplementary Information for

Cryo-EM structure of *Mycobacterium smegmatis* DyP-loaded encapsulin

Yanting Tang, An Mu, Yuying Zhang, Shan Zhou, Weiwei Wang, Yuezheng Lai, Xiaoting Zhou, Fengjiang Liu, Xiuna Yang, Hongri Gong, Quan Wang & Zihe Rao

Corresponding author: Hongri Gong and Quan Wang
Email: gonghr@nankai.edu.cn; wangq@ibp.ac.cn.

This PDF file includes:

Other supplementary materials for this manuscript include the following:

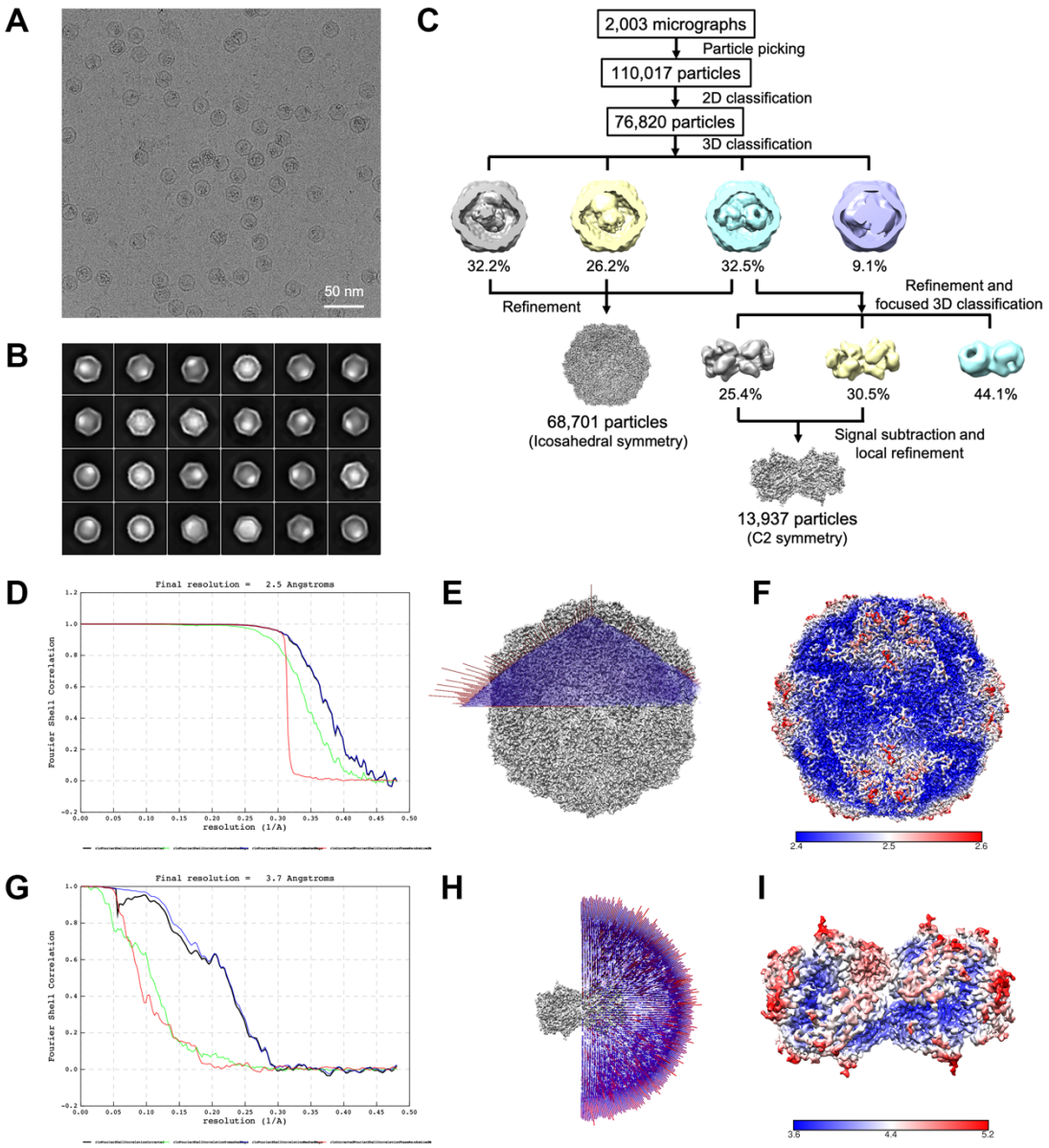

Fig. S1. Image processing procedures and cryo-EM map quality assessment. (*A*) Representative cryo-EM micrograph of *M. smegmatis* encapsulin. Scale bars: 50 nm. (*B*) 2D class averages of different views obtained from reference-free classification. (*C*) Overview of single-particle analysis workflow, from data collection to 3D reconstruction. The percentage of particles for each class are indicated. FSC curves of the final cryo-EM maps of the shell (*D*) and DyP (*G*). Euler angle distribution of all particles of the shell (*E*) and DyP (*H*). Local resolution of the density map of the shell (*F*) and DyP (*I*).

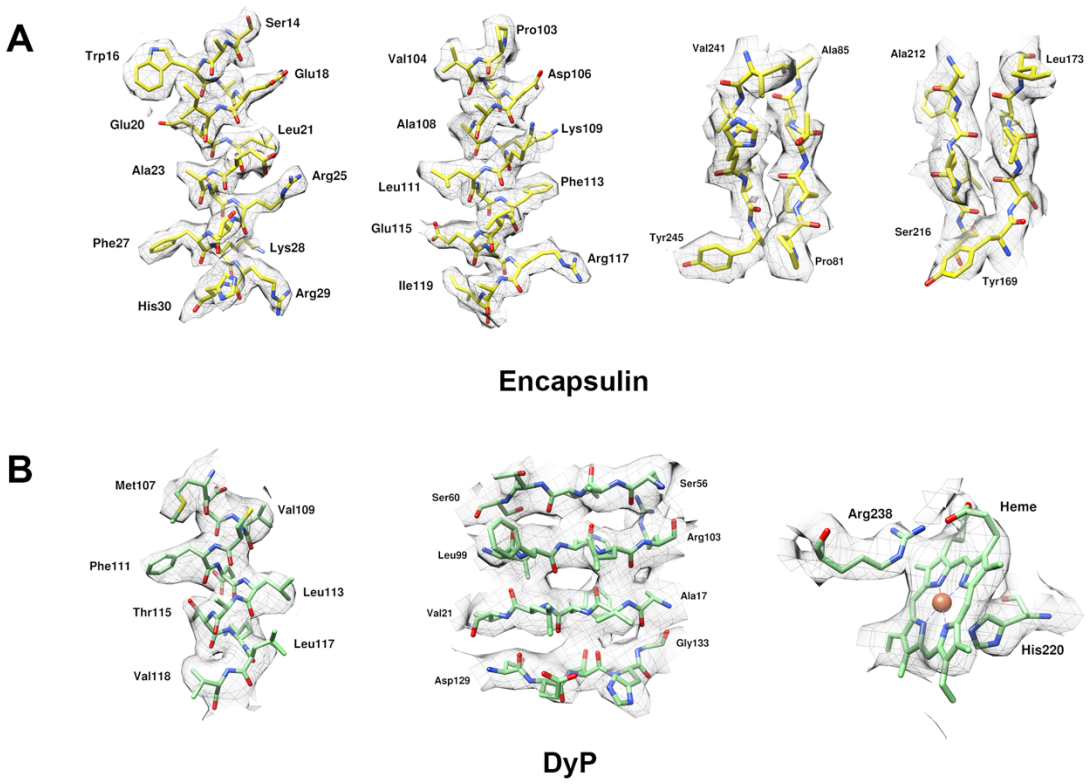

Fig. S2. Cryo-EM densities of individual domains of the encapsulin shell and DyP. (*A*) Representative cryo-EM densities of the shell monomer. (*B*) The cryo-EM densities of α-helix (left), β-sheet (middle) and the heme prosthetic group (right) of DyP, with the distal Arg238 and the proximal His220.

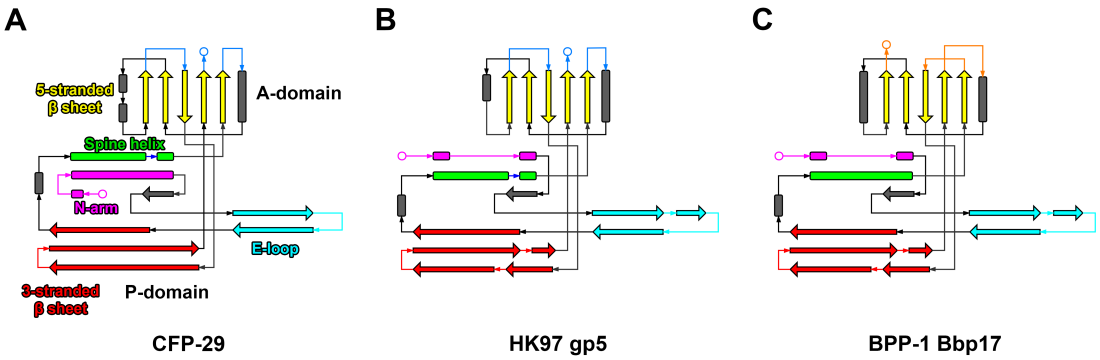

Fig. S3. Topologies of encapsulin shell CFP-29 of *Mycobacterium smegmatis* (*A*), major capsid protein gp5 of *Enterobacteria phage HK97* (*B*) and major capsid protein Bbp17 of *Bordetella phage BPP-1* (*C*). Major structural elements are colored as indicated (N-arm, magenta; E-loop, cyan; 3-stranded β sheet, red; Spine helix, green; 5-stranded β sheet, yellow).

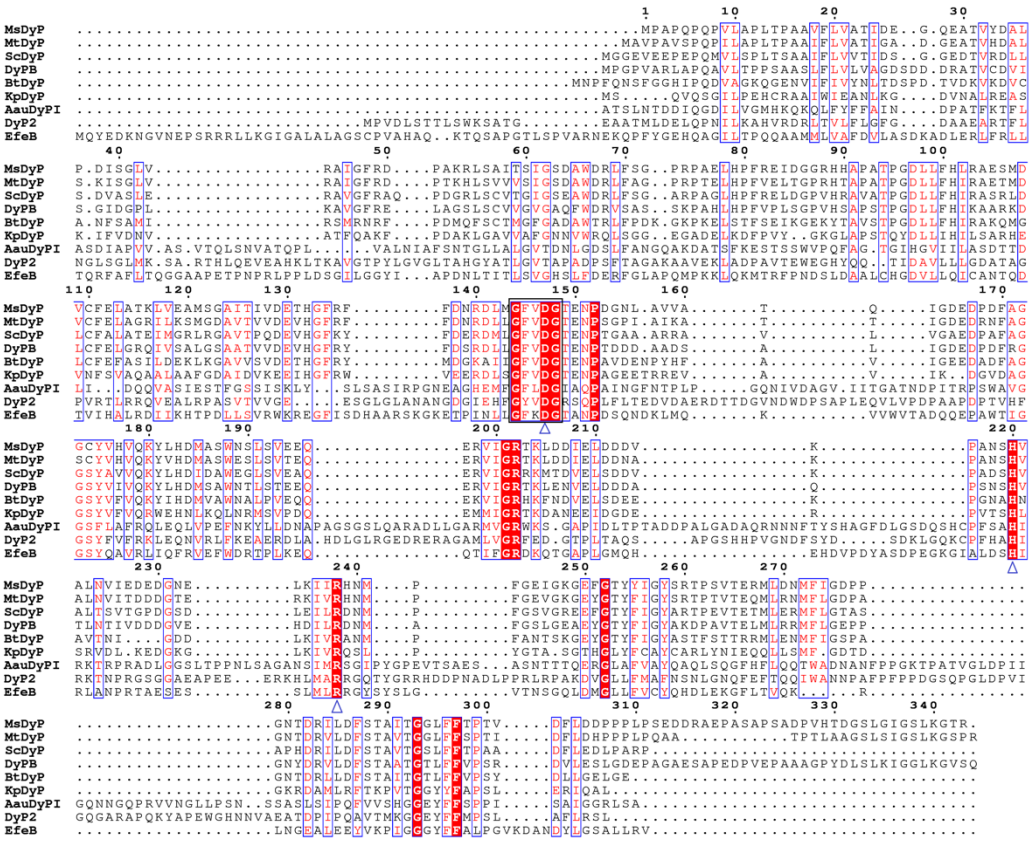

Fig. S4. Multiple sequence alignments of DyP-type peroxidase family members. The completely conserved residues are highlighted in red, and similar residues are displayed in red letters. The unique GXXDG motif is labeled with a black box. Triangles show the important residues around the heme active site. The proximal H220, the ligand for heme, forms the hydrogen-bond interaction with Asp282, and the distal Asp147-Arg238 couple plays an important role in catalytic activity. Sequence alignments were performed by Clustal Omega and ESPript. These sequences were obtained from UniProt, including MsDyP from *Mycobacterium smegmatis* (Accession No. A0R4G9), MtDyP from *Mycobacterium tuberculosis* (Accession No. A0A045KD97), DyPB from *Rhodococcus jostii* RHA1 (Accession No. Q0SE24), ScDyP from *Streptomyces coelicolor* (Accession No. Q9FBY9), BtDyP from *Bacteroides thetaiotaomicron* (Accession No. Q8A8E8), KpDyP from *Klebsiella pneumoniae* (Accession No. A0A0W8ATM9), AauDyPI from *Auricularia auricula-judae* (Accession No. I2DBY1) , DyP2 from *Amycolatopsis sp.* 75iv2 (Accession No. K7N5M8) and EfeB from *Escherichia coli* O157 (Accession No. Q8XAS4).

Table S1. The resolution of the structures of all known cargo-encapsulated encapsulins.

Organism	Method	Global resolution (Å)	Cargo protein	Cargo model built (%)	Reference
Thermotoga maritima	X-ray	3.1	Flp	7.0	(1)
Myxococcus xanthus	Cryo-EM	25	Flp	0	(2)
Brevibacterium linens	Cryo-EM	15.1	DyP	0	(3)
Brevibacterium linens	Cryo-EM	24.7	DyP	0	
Brevibacterium linens	Cryo-EM	28.3	DyP	0	
Brevibacterium linens	Cryo-EM	24.2	TFP	0	
Brevibacterium linens	Cryo-EM	13.5	TFP	0	
Myxococcus xanthus	Cryo-ET	19.6	Flp	0	(4)
Myxococcus xanthus	Cryo-ET	33	Flp	0	
Quasibacillus thermotolerans	Cryo-EM	3.85	IMEF	3.6	(5)
Synechococcus elongatus	Cryo-EM	2.2	CyD	0	(6)
Thermotoga maritima	Cryo-EM	2.86	IDM	0	(7)
Mycobacterium smegmatis	Negative stain EM	27	DyP	0	(8)
Mycobacterium smegmatis	Cryo-EM	4.1 2.5 (shell) 3.7 (cargo)	DyP	90.1	This study

Table S2. Identification of the nanocompartment components by peptide mass fingerprint spectroscopy.

Subunits			Mass Spectroscopy	
Accession /Gene Locus	Protein Description	Theoretical Molecular Weight (kDa)	Number of unique peptides	Percent sequence coverage%
WP_003897233.1 /MSMEG_5830	29 kDa antigen, CFP-29	28.7	17	70.94
WP_011730829.1 /MSMEG_5829	DyP-type peroxidase	37.2	14	57.14

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

	Encapsulin shell	Encapsulated DyP-type peroxidase	Cargo-loaded encapsulin
PDB entry	7BOJ	7BOK	N/A
EMDB entry	EMD-30130	EMD-30131	EMD-30132
Data collection			
Microscope	Titan Krios	Titan Krios	Titan Krios
Voltage (kV)	300	300	300
Magnification	130,000x	130,000x	130,000x
Detector	Gatan GIF-K2	Gatan GIF-K2	Gatan GIF-K2
Data collection software	SerialEM	SerialEM	SerialEM
Electron exposure (e^{-}/Å^2)	60	60	60
Defocus range (μm)	1.5-2.5	1.5-2.5	1.5-2.5
Pixel size (Å)	1.04	1.04	1.04
Data processing			
Number of micrographs (no.)	2,003	2,003	2,003
Final particle images (no.)	68,701	13,937	25,020
Symmetry imposed	I1	C2	C1
Map resolution (Å)			
FSC threshold 0.143	2.5	3.7	4.1
Refinement			
Initial model used (PDB code)	3DKT	4GU7	N/A
Model resolution (Å)			
FSC threshold 0.143	2.4	3.6	N/A
Map sharpening B factor (Å^2)	-91.54	-70.94	N/A
Map correlation coefficient			
CC (mask)	0.86	0.82	N/A
CC (box)	0.42	0.75	N/A
CC (peaks)	0.24	0.66	N/A
CC (volume)	0.75	0.81	N/A
Mean CC for ligands	N/A	0.87	N/A
Model composition			
Non-hydrogen atoms	2,025	14,562	N/A
Protein residues	265	1,854	N/A
Ligands	0	6 (HEM)	N/A
B factors (Å^2)			
Protein	18.20	79.95	N/A
Ligand	N/A	74.07	N/A
R.M.S. deviations			
Bond lengths (Å)	0.007	0.006	N/A
Bond angles (°)	0.904	1.108	N/A
Validation			
MolProbity score	1.09	1.77	N/A
Clashscore	2.49	8.38	N/A
Poor rotamers (%)	0	0	N/A
Ramachandran plot			
Favored (%)	97.72	95.39	N/A
Allowed (%)	2.28	4.61	N/A
Disallowed (%)	0	0	N/A
Cβ outliers (%)	0	0	N/A
Peptide plane (%)			
Cis proline/general	0.0/0.0	0.0/0.0	N/A
Twisted proline/general	0.0/0.0	0.0/0.0	N/A
CaBLAM outliers (%)	0.77	1.69	N/A

Movie S1. The cryo-EM density maps of the DyP and encapsulin shell. Structural analysis was performed using the multi-body refinement in RELION. The relative motion between DyP and encapsulin shell is described based on the first component of the PCA analysis.

SI References

1. M. Sutter *et al.*, Structural basis of enzyme encapsulation into a bacterial nanocomp artment. *Nat. Struct. Mol. Biol.* **15**, 939-947 (2008).
2. C. A. McHugh *et al.*, A virus capsid-like nanocompartment that stores iron and prot ects bacteria from oxidative stress. *EMBO J.* **33**, 1896-1911 (2014).
3. R. M. Putri *et al.*, Structural characterization of native and modified encapsulins as nanoplatforms for in vitro catalysis and cellular uptake. *ACS Nano* **11**, 12796-12804 (2017).
4. F. Sigmund *et al.*, Bacterial encapsulins as orthogonal compartments for mammalian cell engineering. *Nat. Commun.* **9**, 1990 (2018).
5. T. W. Giessen *et al.*, Large protein organelles form a new iron sequestration system with high storage capacity. *eLife* **8**, e46070 (2019).
6. R. J. Nichols *et al.*, Discovery and characterization of a novel family of prokaryotic nanocompartments involved in sulfur metabolism. *bioRxiv* [preprint] (2020). 10.1101/2 020.05.24.113720.
7. X. Xiong *et al.*, Cryo-EM structure of heterologous protein complex loaded *Thermoto ga maritima* encapsulin capsid. *Biomolecules* **10**, 1342 (2020).
8. A. M. Kirykowicz, J. D. Woodward, Shotgun EM of mycobacterial protein complexes during stationary phase stress. *Curr. Res. Struct. Biol.* **2**, 204-212 (2020).