Title

Structures from cells and viruses through endogenous cryoEM methods

Author

Zhen, James

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Structures from cells and viruses through endogenous cryoEM methods

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular Biology

by

James Zhen

2024

ABSTRACT OF THE DISSERTATION

Structures from cells and viruses through endogenous cryoEM methods

by

James Zhen

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2024

Professor Hong Zhou, Chair

As we seek to elucidate the functions and mechanisms of biological constructs, there is a growing need to learn about their structures in their native contexts as members of their molecular societies. Conventional methods of structural biology entail purification of recombinant proteins, which isolates them from their interactome and spatial localization. Thus, minimizing sample manipulation is vital for discovering biological interactions. Unique among the structural biology methods, cryogenic electron microscopy (cryoEM) is able to generate data from impure, heterogeneous samples. By combining old-fashioned endogenous methods with recent advancements in the throughput of cryoEM data collection and computation, we can perform purification *in silico*, allowing for high resolution structural determination of proteins while retaining their native binding partners.

We apply this endogenous cryoEM approach to determine the structures of two critical metabolic complexes, the icosahedral pyruvate dehydrogenase complex (PDC) and the cubic 2-

oxoglutarate dehydrogenase complex (OGDC), from mitochondrial lysates of bovine kidneys. By minimizing disruption of the sample, we identify heterogeneous arrangements of E1 and E3 peripheral subunits bound to their central E2 cores. Additionally, the substrate-shuttling lipoyl moiety of E2 is found within the E2 active site. Direct comparison of PDC and OGDC from the same organism allows us to identify critical interaction interfaces for the prevention of heterologous binding and determination of geometry despite a near-identical monomeric E2 fold.

Some systems do not have the luxury of containing multiple copies for averaging by single particle cryoEM. Instead, cryogenic electron tomography (cryoET) enables imaging of unique subjects within their fully intact environment. Here, we report structures and observations of intact cancer-causing herpesviruses, Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV), and their pleiomorphic tegument and envelope using cryoET. Within their teguments, we observe strand-like features that suggest a means of viral assembly. We also show that herpesvirus fusogen proteins assemble into triplet clusters to promote host membrane fusion.

Our results provide new insights into assembly of metabolic complexes and herpesviruses. Together, they highlight wide applicability of cryoEM and cryoET for understanding large complexes in context of their native systems.

The dissertation of James Zhen is approved.

William M. Gelbart

Joseph Ambrose Loo

Ting-Ting Wu

Jerome A. Zack

Hong Zhou, Committee Chair

University of California, Los Angeles

2024

Dedicated to my family.

Table of Contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

List of Tables

Acknowledgements

I express my foremost gratitude to my thesis advisor, Dr. Hong Zhou, for his continuous patience and support. The unwavering optimistic attitude of his guidance has helped me see the positives in both my data and my life. His absolute support of me and other trainees within his group has been inspirational, and I aspire to one day become as thoughtful and generous mentor as him.

Thank you to my collaborators, Dr. Jia Chen, Dr. Richard Longnecker, Dr. Alex Lam, Dr. Haigen Huang, and Dr. Ting-Ting Wu for making our work on gammaherpesviruses possible. Their responsiveness and vast knowledge of the subject have guided us through these projects. They have always been receptive of my ideas and opinions, allowing for great discussions.

A great thanks to Dr. Shiheng Liu and Dr. Xian Xia for allowing me to work with them on the endogenous lysates. They have been great mentors who have helped me build my foundational knowledge about electron microscope operation and cryoEM data processing. Their tutelage has enabled my experimental success.

I extend my gratitude to Dr. Fumiaki Ito and Dr. Guodong Xie, who have been of great help with their expression and cryoEM expertise with recombinant systems, making the herpesvirus gB structure possible.

I thank the members of the Zhou group, both past and present for their great collective knowledge and attitudes. In particular, a great thanks to Dr. Si Zhu and Dr. Hui Wang for teaching me cryoET. I also thank Dr. Kang Zhou, Dr. Yao He, and Dr. Yuntao Liu, from who have patiently introduced a naïve me to cryoEM. I also thank my fellow graduate student compatriots Alex Stevens and Jonathan Jih for our great discussions, both scientific and non-scientific.

I also extend my thanks to the past and present staff of the Electron Imaging Center for

Nanosystems, Judy Su, Ivo Atanasov, Wong Hoi Hui, Dr. David Strugatsky, Dr. Matthew Mecklenburg, Dr. Yanxiang Cui, and Dr. Peng Ge, for their training and maintenance of the instruments. Without them, this work would not be possible. I further extend my thanks to Dr. David Strugatsky for introducing me to cell and virus culture.

I am grateful for Dr. Yuan Yang and Dr. Juli Feigon, who have instilled a great sense of lab technique and scientific inquisitiveness into me during my short rotation with them.

I am also grateful for my summer internship manager Dr. Colin Deniston who exposed me to rigorous science outside of the academic setting and trusted me with autonomy of various projects. I also thank Dr. Cosmo Buffalo for sharing his expertise.

I thank Dr. Zhefeng Guo for introducing me to academic research and laying my foundation for structural biology.

Lastly, my greatest thanks to my parents and my brother who have supported me through this entire journey. Thank you to my friends who have accommodated me regardless of my schedule. Thank you to my partner who has supported me through and through.

Chapter 2 is adapted from Liu, S., Xia, X., Zhen, J. et al. Structures and comparison of endogenous 2-oxoglutarate and pyruvate dehydrogenase complexes from bovine kidney. *Cell Discov* **8**, 126 (2022). Springer Nature grants authors the right to reuse their article's Version of Record, in whole or in part, in their own thesis. Additionally, they may reproduce and make available their thesis, including Springer Nature content, as required by their awarding academic institution.

Chapter 3 is adapted from a manuscript submitted to *Journal of Virology*. American Society for Microbiology (ASM) Journals grants authors the right to republish discrete portions

of their article in any other publication (including print, CD-ROM, and other electronic formats), provided that proper credit is given to the original ASM publication. ASM authors also retain the right to reuse the full article in their dissertation or thesis.

Chapter 4 is adapted from a manuscript revised at *Journal of Virology* with an initial submission preprint deposited on *bioRxiv*: Zhen, J. et al. Structures of Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus virions reveal species-specific tegument and envelope features. Preprint at https://doi.org/10.1101/2024.07.09.602672 (2024).

Vita

James Zhen

Education

University of California, Los Angeles. PhD Candidate, Molecular Biology. Sept. 2019 – Sept. 2024.

University of California, Los Angeles. BS, Biochemistry. Minor in Biomedical Research. Sept. 2015 – Mar. 2019.

Academic/Professional History

University of California, Los Angeles. Hong Zhou group. Graduate Student Researcher. Sept. 2019 – Dec. 2019.

In situ atomic structures of herpesvirus/Separating native structures from cellular milieu

Novartis Institutes for BioMedical Research. Intern. June 2023 – Aug. 2023.

Cryo-EM of biotherapeutics

University of California, Los Angeles. Zhefeng Guo Group. Undergraduate Research Assistant. Jan. 2017 – Mar. 2019.

Structural studies of Aβ40 fibrils using electron paramagnetic resonance spectroscopy

Publications

1. Ito, F.*, **Zhen, J.***, Xie, G., Huang, H., Silva, J.C., Wu, T.-T., and Zhou, Z.H. Atomic structures of the Kaposi's sarcoma-associated herpesvirus gB in post-fusion conformation. Manuscript submitted.
2. **Zhen, J.**, Chen, J., Huang, H., Liao, S., Liu, S., Yuan, Y., Sun, R., Longnecker, R., Wu, T.-T., and Zhou, Z.H. Structures of Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus virions reveal species-specific tegument and envelope features. Manuscript revised at *Journal of Virology*.
3. Lam, A.K., Roshan, R., Miley, W., Labo, N., **Zhen, J.** et al. Immunization of mice with virus-like vesicles of Kaposi sarcoma-associated herpesvirus reveals a role for antibodies targeting ORF4 in activating complement-mediated neutralization. *Journal of Virology* e01600-22 (2023).
4. Liu, S.*, Xia, X.*, **Zhen, J.***, Li, Z., and Zhou, Z.H. Structures and comparison of endogenous 2-oxoglutarate and pyruvate dehydrogenase complexes from bovine kidney. *Cell Discovery* 8, 126 (2022).
5. Xiao, H., Duo, L., **Zhen, J.**, Wang, H., and Guo, Z. Static and dynamic disorder in Aβ40 fibrils. *Biochemical and Biophysical Research Communications* 610, 107-112 (2022).
6. Yang, Y., Liu, S., Egloff, S., Eichhorn, C.D., Hadjian, T., **Zhen, J.**, Kiss, T., Zhou, Z.H., and Feigon, F. Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP. *Molecular Cell* 82, 1724-1736 (2022).
7. Yoon, A., **Zhen, J.**, and Guo, Z. Segmental structural dynamics in Aβ42 globulomers. *Biochemical and Biophysical Research Communications* 545, 119–124 (2021).

(* indicates co-first authorship)

Presentations

1. Zhen, J., Chen, J., Huang, H., Liao, S., Liu, S., Yuan, Y., Wu, T.-T., Longnecker, R., Zhou, Z.H. Structures of Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus from

cryo-electron tomography. Poster presentation delivered at FASEB Science Research Conferences Virus Structure and Assembly, July 2024.

2. Zhen, J., Chen, J., Huang, H., Liu, S., Wu, T.-T., Longnecker, R., Zhou, Z.H. Structures of cancer-causing human gammaherpesviruses and comparison with alpha- and betaherpesvirus. Poster presentation delivered at SoCal Cryo-EM Symposium, October 2023.
3. Zhen, J., Chen, J., Huang, H., Liu, S., Wu, T.-T., Longnecker, R., Zhou, Z.H. Structures of cancer-causing human gammaherpesviruses and comparison with alpha- and betaherpesvirus. Poster presentation delivered at International Herpesvirus Workshop, July 2023.
4. Zhen, J., Liu, S., Xia, X., Li, Z., and Zhou, Z.H. Structure and comparison of endogenous 2-oxoglutarate and pyruvate dehydrogenase complexes from bovine kidney. Poster presentation delivered at SoCal Cryo-EM Symposium, December 2022.
5. Zhen, J., Wang, H., Xiao, K., and Guo, Z. Structural studies of Aβ40 fibrils using electron paramagnetic resonance spectroscopy. Poster presentation delivered at UCLA Research Poster Day, May 2018.

Honors/Awards

International Herpesvirus Workshop Merit-Based Award	2023
NIH T32 Interdisciplinary Training in Virology and Gene Therapy	2020 – 2021
UCLA Graduate Dean's Scholar	2019
UCLA Regents Scholar	2015 – 2019
UCLA Undergraduate Research Scholar	2017 – 2018

Teaching

University of California, Los Angeles

MIMG 101 Introductory Microbiology	Fall 2020, Fall 2021

Chapter 1

Introduction

1.1 The crude origins of structural biology

In 1960, the first atomic structure of a protein was determined for myoglobin that was isolated from sperm whale meat[1,2]. Prior to the advent of recombinant systems and tagging[3,4], proteins and other biological components were obtained from their native sources out of necessity. However, crystal growth for X-ray crystallography structure determination required further purification of these endogenous isolates, stripping them from their sea of interacting partners. At the time, obtaining the structure itself was considered a breakthrough, and the extensive network of the interactome was unknown[5].

The first demonstration of a helix by cryogenic electron microscopy (cryoEM) was performed by Richard Henderson in 1975[6]. He used purple membrane sheets extracted from halobacterium to characterize the transmembrane helices of bacteriorhodopsin. Henderson and his colleagues would then go on to solve the first cryoEM atomic structure with bacteriorhodopsin in 1990[7]. However, these structures were obtained using electron diffraction by taking advantage of the native crystalline arrangement of bacteriorhodopsin on the membrane. Atomic structures through the imaging modality by single particle cryoEM of non-crystalline samples would not be obtained until nearly two decades later for highly symmetric viruses[8–10]. Soon after, with the development of direct electron detectors, atomic resolution structures could be determined for individual proteins[11,12], breaking open the floodgates for atomic structures in their native state, free from crystal packing. Due to the profound impact and new paradigm

established by cryoEM for the field of structural biology, the 2017 Nobel Prize in Chemistry was awarded to Jacques Dubochet, Joachim Frank, and Richard Henderson for the development of cryoEM.

Since then, the popularity of cryoEM has skyrocketed due to its immense applicability for structural determination and new innovations that have increased its accessibility and breadth of potential targets. In tandem with the growing breadth of structures in the Protein Data Bank[13], our depth of knowledge on how these structures interact with one another grows alongside it. Current efforts in structural biology have shifted towards *in situ* structures to better understand how biological constructs exist, not as individuals, but as members of an interactive molecular society. In this thesis, we demonstrate a return to crude samples for the retention and discovery of biological interactions. Using single particle cryoEM of cellular isolates, we document the native core complex of our central metabolism. We also demonstrate how cryogenic electron tomography (cryoET) is used to resolve native or *in situ* structures using a viral system and how a recombinant protein approach towards a viral fusogen is limited to an undesirable conformation.

1.2 Overview of native structures by imaging crude cellular isolates by cryoEM

CryoEM has enabled structural determination of biological structures in their native states, frozen in time within a window of glass-like ice. Conventional sample preparation methods for cryoEM have their roots in X-ray crystallography and nuclear magnetic resonance, where high sample purity is required as a baseline to obtain interpretable data[14,15]. Although recombinant samples are a useful tool for understanding protein and nucleic acid structural biology, binding

partners are often omitted or limited to those with prior biochemical knowledge, thus providing an incomplete picture. In some cases, a sample of interest inherently cannot be recovered at high purity, thus stymying its structural characterization[16]. CryoEM, on the other hand, is an imaging modality that captures structural information while tolerating sample heterogeneity, and permits the presence of interactome members, both known and unknown.

CryoEM can be applied to lysate samples where the milieu of species is left intact, and the purification is instead performed computationally, *in silico*[16] (Lyumkis 2019). In the cryoEM procedure, many movies are collected of a sample specimen at high magnification. The individual frames of these movies are then aligned and summed into micrographs to correct for beam-induced motion during collection[17]. Particles can then be picked from these micrographs either manually or, more commonly, automatically. Different particle picking protocols can be used to narrow or broaden the search [18,19]. These particles are extracted into their own individual images, which can then be compared and clustered during 2D classification. During 2D classification, the particle images are computationally aligned with consideration for translation and in-plane rotation to produce class averages of repeating features that improve signal-to-noise. This classification step is critical for identifying species heterogeneity when imaging lysates. In a milieu, 2D class averages encompass both different orientations within a species and different species. Innovations in *ab initio* reconstruction and heterogeneous refinement have enabled solving 3D reconstructions from these mixtures of views and species[20].

The resolution of cryoEM reconstructions is dependent on the number of particles used in the reconstruction [21,22]. In a lysate, the particles within a micrograph are split amongst different species with some occupying more or less of the proportion dependent on expression levels within the system. As a result, each micrograph from a lysate sample contributes less towards the

completion of a 3D reconstruction than a purified recombinant sample, necessitating longer imaging sessions and for lysate samples. Additionally, extensive classification of particles entails greater computational overhead. Recent innovations in both cryoEM hardware and software have reduced the burden of imaging endogenous samples[20,21,23]. Reconstructions of unknown species can also be identified through automated processes[24–26]. These advances and those to come increase the accessibility of endogenous structural biology, which deepens our understanding of the molecular interactions that govern life.

1.3 Overview of native structures through cryoET

CryoET is an extension of cryoEM that is capable of 3D reconstruction of individual subjects of interest rather than a bulk average. Beyond the individual protein level, protein complexes and larger molecular and cellular ultrastructures contain heterogeneous elements that are not amenable to averaging by single particle cryoEM. By tilting the electron microscope stage and recording multiple perspectives of the same subject of interest, the 2D tilt series images can be backprojected into a 3D volume, a tomogram, which contains a treasure trove of structural information to be uncovered and annotated[27,28]. An intact system retains spatial context between features, making their study by cryoET a powerful for elucidating the foundational construction of cells and viruses.

To meet the demand for endogenous structural information, cryoET has undergone a recent surge in innovation to improve both throughput and interpretability[27]. Automation has been implemented for the tedious tasks of tilt series collection, alignment, and tomogram reconstruction[29–32]. Neural networks have been applied for improving tomogram quality and annotation of the wealth of data[33,34]. Newer software packages have also made cryoET more accessible[29,35–37].

1.4 Critical components of metabolism: α-keto acid dehydrogenase complexes

Metabolism is an essential requirement for all life. For higher-order organisms, the linkage of glycolysis to oxidative phosphorylation through the Krebs cycle is necessary for survival. The pyruvate dehydrogenase complex (PDC) catalyzes the oxidative decarboxylation of pyruvate to acetyl-CoA, linking glycolysis to the Krebs cycle and downstream oxidative phosphorylation. PDC is a member of the α-ketoacid dehydrogenase complex (αKDC) family, alongside 2-oxoglutarate dehydrogenase complex (OGDC) and branched-chain α-ketoacid dehydrogenase complex (BCKDC), both of which perform analogous reactions in central metabolism[38,39]. Genomic studies have identified mutations in components of these multienzyme complexes linked to severe clinical consequences, including metabolic acidosis and neurodegeneration[40–43]. The PDC also attracts interest in cancer biology for its role in modulating the Warburg effect to promote tumor anabolism[41,44–46]. Although pathological mutations in the αKDC family have been identified, it is unclear how these mutations are mapped to structural and functional changes that result in the phenotype. Much of our current understanding of αKDCs is derived from recombinantly expressed proteins. However, structures from these individual proteins are unable to fully capture how multiple components come together and function as a complex. In the absence of the native complexes, we have an insufficient understanding of how these subunits are organized into functional complexes.

PDC, OGDC, and BCKDC are each comprised of multiple copies of a substrate specific dihydrolipoamide acyltransferase (E2) core surrounded by α-ketoacid specific dehydrogenases (E1) and universal dihydrolipoamide dehydrogenases (E3)[38,39]. Each E1, E2, and E3 subunit in an α-ketoacid dehydrogenase is highly similar both in sequence and structure to its counterpart in the other classes of the family[47,48]. Despite these similarities, the subunits self-assemble into their

respective complexes. Furthermore, the PDC core adopts different geometry in lower order and higher order organisms, but OGDC and BCKDC cores retain the same geometry[48–52]. Understanding the molecular sociology of these complexes requires a holistic approach in visualizing the inter-subunit interactions at the residual level.

While there are standalone atomic models of the E2 cores and some E1 and E3 proteins, critically, there is no single complete structure containing all components for any α-ketoacid dehydrogenase to provide a complete view of the complex. CryoEM presents an exciting opportunity for structural determination of the αKDCs in native or near-native conditions to identify the interactions that govern their mechanisms of assembly and function[24].

1.5 Cancer-causing gammaherpesviruses

Herpesvirus is a family of enveloped dsDNA viruses that are highly prevalent amongst humans and account for the greatest share of chronic infections[53,54]. The herpesviruses Kaposi's sarcoma-associated herpesvirus (KSHV) and Epstein-Barr virus (EBV) are two of the seven known cancer-causing viruses[55,56]. KSHV and EBV are both members of the gammeaherpesvirus subfamily. Their infection is normally controlled by the immune system but can cause tumor outgrowth and trigger fatal lymphoproliferative disease in immunosuppressed individuals. EBV is associated with Hodgkin's lymphoma, Burkett's lymphomas, and nasopharyngeal carcinomas[56,57]. Additionally longitudinal analysis of a cohort larger than 10 million individuals has shown that EBV is a causative agent of multiple sclerosis[58]. KSHV is the etiologic agent of Kaposi's sarcoma (KS), the most common malignancy associated with infection by human immunodeficiency virus (HIV)[56,59,60]. KSHV is also causally linked to multicentric Castleman's disease and primary effusion lymphoma (PEL)[61]. Like other members of the Herpesviridae family, KSHV establishes lifelong latent infections[62,63]. Periodically, the virus may be activated

and shed, often asymptomatically. Transmission occurs through skin or mucous membrane contact with saliva or lesions from both symptomatic and asymptomatic carriers. Thus, the development of vaccines and treatments against KSHV and EBV infection is critical in preventing their spread and human health burdens.

All herpesviruses are composed of a dsDNA genome enclosed within an icosahedral tegument that is then surrounded with an amorphous tegument layer and pleiomorphic envelope decorated with envelope glycoproteins. While most cells are latently infected, lytic gammaherpesvirus replication is necessary for initial infection, virion production, and viral transmission, thus contributing to the pathogenesis of this disease[62–67].Glycoprotein-mediated membrane fusion is essential to the persistent infection and viral spread of EBV and KSHV during lytic replication. In herpesviruses, numerous glycoproteins embedded in the envelope function as the fusion machinery responsible for host entry, but most of their structures are undetermined[68]. The herpesvirus fusion machinery is more complex than that of other viruses due to several envelope glycoproteins that undergo specific conformational changes and interactions to guide the process. However, a recombinant system is unable to recapitulate the interactions between these glycoproteins necessary for infection. Tegument components within herpesvirus also guide their assembly and prepare the host environment for infection [69,70]. Unfortunately, herpesvirus tegument proteins have unable to be captured in recombinant systems.

There are currently no specific drugs or vaccines against gammaherpesvirus infection. Notably, structures of the prefusion spike glycoprotein of SARS-CoV-2 enabled the rapid and successful development of vaccines during the pandemic[71,72], and the same structure-based approach recently yielded the first respiratory syncytial virus (RSV) vaccine [73]. Similarly,

treatment of EBV and KSHV can proceed through the design of structure-based entry inhibitors and vaccines targeting EBV and KSHV envelope glycoproteins. However, there are no atomic structures of KSHV glycoproteins and Alphafold structure prediction is unreliable for viral proteins, so how these glycoproteins interact with their binding partners is unknown. Recent technological and methodological breakthroughs in both cryogenic electron microscopy (cryo-EM) and electron cryotomography (cryo-ET) now present an exciting opportunity for structural determination of KSHV tegument and glycoproteins in native or near-native conditions.

1.6 Dissertation Outline

The work presented in this thesis seeks to structurally characterize complex systems using endogenous cryoEM and cryoET approaches. Chapter 1 provides a background for the methods of cryoEM and cryoET and the motivations for endogenous structural studies of the systems at hand. Chapter 2 presents the published manuscript on the application of cryoEM for simultaneous *in silico* purification and structural characterization of endogenous PDC and OGDC from the milieu of bovine kidney mitochondrial lysates[74]. Chapter 3 describes structural characterization of KSHV gB through recombinant means and then contextualizes these findings in the virus system. Chapter 4 presents a manuscript in submission about the structural characterization of intact EBV and KSHV virions, involving contextualization of the pleiomorphic tegument and envelope through mapping back of calculated capsid vertices and portal positions[75]. Chapter 5 summarizes the work presented in Chapters 2 through 4 and future prospects stemming from them.

1.6 References

1. Kendrew, J. C. *et al.* Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature* **185**, 422–427 (1960).

2. Dickerson, R. E. Chapter 2: Myoglobin: A Whale of a Structure! *Journal of Molecular Biology* **392**, 10–23 (2009).

3. Munro, S. & Pelham, H. R. Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of Drosophila hsp 70. *EMBO J* **3**, 3087–3093 (1984).

4. Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598–599 (1975).

5. Braun, P. & Gingras, A.-C. History of protein–protein interactions: From egg-white to complex networks. *PROTEOMICS* **12**, 1478–1498 (2012).

6. Henderson, R. & Unwin, P. N. T. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257**, 28–32 (1975).

7. Henderson, R. *et al.* Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *Journal of Molecular Biology* **213**, 899–929 (1990).

8. Zhang, X., Jin, L., Fang, Q., Hui, W. H. & Zhou, Z. H. 3.3 Å Cryo-EM Structure of a Nonenveloped Virus Reveals a Priming Mechanism for Cell Entry. *Cell* **141**, 472–482 (2010).

9. Yu, X., Jin, L. & Zhou, Z. H. 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature* **453**, 415–419 (2008).

10. Settembre, E. C., Chen, J. Z., Dormitzer, P. R., Grigorieff, N. & Harrison, S. C. Atomic model of an infectious rotavirus particle. *EMBO J* **30**, 408–416 (2011).

11. Cao, E., Liao, M., Cheng, Y. & Julius, D. TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* **504**, 113–118 (2013).

12. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods* **10**, 584–590 (2013).

13. Burley, S. K. *et al.* Protein Data Bank: A Comprehensive Review of 3D Structure Holdings and Worldwide Utilization by Researchers, Educators, and Students. *Biomolecules* **12**, 1425 (2022).

14. Acton, T. B. *et al.* Preparation of Protein Samples for NMR Structure, Function, and Small Molecule Screening Studies. *Methods Enzymol* **493**, 21–60 (2011).

15. KIM, Y. *et al.* HIGH-THROUGHPUT PROTEIN PURIFICATION FOR X-RAY CRYSTALLOGRAPHY AND NMR. *Adv Protein Chem Struct Biol* **75**, 85–105 (2008).

16. Lyumkis, D. Challenges and opportunities in cryo-EM single-particle analysis. *J. Biol. Chem.* **294**, 5181–5197 (2019).

17. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **14**, 331–332 (2017).

18. Bepler, T., Kelley, K., Noble, A. J. & Berger, B. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nature Communications* **11**, 5208 (2020).

19. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochemical Journal* **478**, 4169–4185 (2021).

20. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290–296 (2017).

21. Chari, A. & Stark, H. Prospects and Limitations of High-Resolution Single-Particle Cryo-Electron Microscopy. *Annual Review of Biophysics* **52**, 391–411 (2023).

22. Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics* **28**, 171–193 (1995).

23. Fréchin, L., Holvec, S., von Loeffelholz, O., Hazemann, I. & Klaholz, B. P. High-resolution cryo-EM performance comparison of two latest-generation cryo electron microscopes on the human ribosome. *Journal of Structural Biology* **215**, 107905 (2023).

24. Ho, C.-M. *et al.* Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. *Nature Methods* 1–7 (2019) doi:10.1038/s41592-019-0637-y.

25. Chojnowski, G. *et al.* findMySequence: a neural-network-based approach for identification of unknown proteins in X-ray crystallography and cryo-EM. *IUCrJ* **9**, (2022).

26. Jamali, K. *et al.* Automated model building and protein identification in cryo-EM maps. *Nature* 1–2 (2024) doi:10.1038/s41586-024-07215-4.

27. Hong, Y., Song, Y., Zhang, Z. & Li, S. Cryo-Electron Tomography: The Resolution Revolution and a Surge of In Situ Virological Discoveries. *Annual Review of Biophysics* **52**, 339–360 (2023).

28. Hylton, R. K. & Swulius, M. T. Challenges and triumphs in cryo-electron tomography. *iScience* **24**, 102959 (2021).

29. Burt, A. *et al.* An image processing pipeline for electron cryo-tomography in RELION-5. *FEBS Open Bio* **n/a**,.

30. Eisenstein, F. *et al.* Parallel cryo electron tomography on in situ lamellae. *Nat Methods* 1–8 (2022) doi:10.1038/s41592-022-01690-1.

31. Bouvette, J. *et al.* Beam image-shift accelerated data acquisition for near-atomic resolution single-particle cryo-electron tomography. *Nature Communications* **12**, 1957 (2021).

32. Zheng, S. *et al.* AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. *Journal of Structural Biology: X* **6**, 100068 (2022).

33. Liu, Y.-T. *et al.* Isotropic reconstruction for electron tomography with deep learning. *Nat Commun* **13**, 6482 (2022).

34. Heebner, J. E. *et al.* Deep Learning-Based Segmentation of Cryo-Electron Tomograms. *JoVE (Journal of Visualized Experiments)* e64435 (2022) doi:10.3791/64435.

35. Wang, H., Liao, S., Yu, X., Zhang, J. & Zhou, Z. H. TomoNet: A streamlined cryogenic electron tomography software pipeline with automatic particle picking on flexible lattices. *Biol Imaging* **4**, e7 (2024).

36. Liu, H.-F. *et al.* nextPYP: a comprehensive and scalable platform for characterizing protein variability in situ using single-particle cryo-electron tomography. *Nat Methods* 1–11 (2023) doi:10.1038/s41592-023-02045-0.

37. Balyschew, N. *et al.* Streamlined Structure Determination by Cryo-Electron Tomography and Subtomogram Averaging using TomoBEAR. 2023.01.10.523437 Preprint at https://doi.org/10.1101/2023.01.10.523437 (2023).

38. Patel, M. S. & Roche, T. E. Molecular biology and biochemistry of pyruvate dehydrogenase complexes. *The FASEB Journal* **4**, 3224–3233 (1990).

39. Reed, L. J. A Trail of Research from Lipoic Acid to α-Keto Acid Dehydrogenase Complexes. *Journal of Biological Chemistry* **276**, 38329–38336 (2001).

40. Ævarsson, A. *et al.* Crystal structure of human branched-chain α-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease. *Structure* **8**, 277–291 (2000).

41. Roche, T. E. & Hiromasa, Y. Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer. *Cell. Mol. Life Sci.* **64**, 830 (2007).

42. Burrage, L. C., Nagamani, S. C. S., Campeau, P. M. & Lee, B. H. Branched-chain amino acid metabolism: from rare Mendelian diseases to more common disorders. *Hum Mol Genet* **23**, R1–R8 (2014).

43. Whitley, M. J. *et al.* Pyruvate dehydrogenase complex deficiency is linked to regulatory loop disorder in the αV138M variant of human pyruvate dehydrogenase. *Journal of Biological Chemistry* **293**, 13204–13213 (2018).

44. McFate, T. *et al.* Pyruvate Dehydrogenase Complex Activity Controls Metabolic and Malignant Phenotype in Cancer Cells. *Journal of Biological Chemistry* **283**, 22700–22708 (2008).

45. Tran, Q., Lee, H., Park, J., Kim, S.-H. & Park, J. Targeting Cancer Metabolism - Revisiting the Warburg Effects. *Toxicological Research* **32**, 177–193 (2016).

46. Chen, J. *et al.* Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer. *Nature Genetics* **50**, 219–228 (2018).

47. Yu, X. *et al.* Structures of the Human Pyruvate Dehydrogenase Complex Cores: A Highly Conserved Catalytic Center with Flexible N-Terminal Domains. *Structure* **16**, 104–114 (2008).

48. Forsberg, B. O., Aibara, S., Howard, R. J., Mortezaei, N. & Lindahl, E. Arrangement and symmetry of the fungal E3BP-containing core of the pyruvate dehydrogenase complex. *Nature Communications* **11**, 4667 (2020).

49. Stoops, J. K. *et al.* On the Unique Structural Organization of the Saccharomyces cerevisiae Pyruvate Dehydrogenase Complex. *J Biol Chem* **272**, 5757–5764 (1997).

50. Kyrilis, F. L. *et al.* Integrative structure of a 10-megadalton eukaryotic pyruvate dehydrogenase complex from native cell extracts. *Cell Reports* **34**, 108727 (2021).

51. Andi, B. *et al.* Structure of the dihydrolipoamide succinyltransferase catalytic domain from Escherichia coli in a novel crystal form: a tale of a common protein crystallization contaminant. *Acta Cryst F* **75**, 616–624 (2019).

52. Nagy, B. *et al.* Structure of the dihydrolipoamide succinyltransferase (E2) component of the human alpha-ketoglutarate dehydrogenase complex (hKGDHc) revealed by cryo-EM and cross-linking mass spectrometry: Implications for the overall hKGDHc structure. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1865**, 129889 (2021).

53. Virgin, H. W., Wherry, E. J. & Ahmed, R. Redefining Chronic Viral Infection. *Cell* **138**, 30–50 (2009).

54. Hoang, M. P., Rogers, B. B., Dawson, D. B. & Scheuermann, R. H. Quantitation of 8 Human Herpesviruses in Peripheral Blood of Human Immunodeficiency Virus-Infected Patients and Healthy Blood Donors by Polymerase Chain Reaction. *American Journal of Clinical Pathology* **111**, 655–659 (1999).

55. Fatahzadeh, M. & Schwartz, R. A. Oral Kaposi's sarcoma: a review and update. *Int J Dermatol* **52**, 666–672 (2013).

56. Wen, K. W., Wang, L., Menke, J. R. & Damania, B. Cancers associated with human gammaherpesviruses. *The FEBS Journal* **289**, 7631–7669 (2022).

57. Schiller, J. T. & Lowy, D. R. An Introduction to Virus Infections and Human Cancer. *Recent Results Cancer Res* **217**, 1–11 (2021).

58. Bjornevik, K. *et al.* Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science* (2022) doi:10.1126/science.abj8222.

59. Chang, Y. *et al.* Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* **266**, 1865–1869 (1994).

60. Cesarman, E., Chang, Y., Moore, P. S., Said, J. W. & Knowles, D. M. Kaposi's Sarcoma–Associated Herpesvirus-Like DNA Sequences in AIDS-Related Body-Cavity–Based Lymphomas. *New England Journal of Medicine* **332**, 1186–1191 (1995).

61. Sharp, T. V. & Boshoff, C. Kaposi's sarcoma-associated herpesvirus: from cell biology to pathogenesis. *IUBMB Life* **49**, 97–104 (2000).

62. Flore, O. *et al.* Transformation of primary human endothelial cells by Kaposi's sarcoma-associated herpesvirus. *Nature* **394**, 588–592 (1998).

63. Ballestas, M. E., Chatis, P. A. & Kaye, K. M. Efficient persistence of extrachromosomal KSHV DNA mediated by latency-associated nuclear antigen. *Science* **284**, 641–644 (1999).

64. Zhong, W., Wang, H., Herndier, B. & Ganem, D. Restricted expression of Kaposi sarcoma-associated herpesvirus (human herpesvirus 8) genes in Kaposi sarcoma. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6641–6646 (1996).

65. Blauvelt, A., Herndier, B. G. & Orenstein, J. M. Propagation of a human herpesvirus from AIDS-associated Kaposi's sarcoma. *N. Engl. J. Med.* **336**, 1837–1839 (1997).

66. Renne, R., Blackbourn, D., Whitby, D., Levy, J. & Ganem, D. Limited Transmission of Kaposi's Sarcoma-Associated Herpesvirus in Cultured Cells. *J Virol* **72**, 5182–5188 (1998).

67. Zhu, F. X., Cusano, T. & Yuan, Y. Identification of the immediate-early transcripts of Kaposi's sarcoma-associated herpesvirus. *J. Virol.* **73**, 5556–5567 (1999).

68. Connolly, S. A., Jardetzky, T. S. & Longnecker, R. The structural basis of herpesvirus entry. *Nature Reviews Microbiology* **19**, 110–121 (2021).

69. Bohannon, K. P., Jun, Y., Gross, S. P. & Smith, G. A. Differential protein partitioning within the herpesvirus tegument and envelope underlies a complex and variable virion architecture. *Proceedings of the National Academy of Sciences* **110**, E1613–E1620 (2013).

70. Guo, H., Shen, S., Wang, L. & Deng, H. Role of tegument proteins in herpesvirus assembly and egress. *Protein Cell* **1**, 987–998 (2010).

71. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).

72. Hsieh, C.-L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).

73. Joyce, M. G. *et al.* Iterative structure-based improvement of a fusion-glycoprotein vaccine against RSV. *Nat Struct Mol Biol* **23**, 811–820 (2016).

74. Liu, S., Xia, X., Zhen, J., Li, Z. & Zhou, Z. H. Structures and comparison of endogenous 2-oxoglutarate and pyruvate dehydrogenase complexes from bovine kidney. *Cell Discov* **8**, 1–15 (2022).

75. Zhen, J. *et al.* Structures of Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus virions reveal species-specific tegument and envelope features. 2024.07.09.602672 Preprint at https://doi.org/10.1101/2024.07.09.602672 (2024).

Chapter 2

Structures and comparison of endogenous 2-oxoglutarate and pyruvate dehydrogenase complexes from bovine kidney

Shiheng Liu[1,2,#], Xian Xia[1,2,#], James Zhen[1,2,3,#], Zihang Li[1,2], Z. Hong Zhou[1,2,3,*]

1 Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA

2 California NanoSystems Institute, UCLA, Los Angeles, CA 90095, USA

3 Molecular Biology Institute, UCLA, Los Angeles, CA 90095, USA

# Contributed equally

* Corresponding Author

Email: hong.zhou@ucla.edu; Phone: 310-694-7527

2.1 Abstract

The α-keto acid dehydrogenase complex family catalyzes the essential oxidative decarboxylation of α-keto acids to yield acyl-CoA and NADH. Despite performing the same overarching reaction, members of the family have different component structures and structural organization between each other and across phylogenetic species. While native structures of α-keto acid dehydrogenase complexes from bacteria and fungi became available recently, the atomic structure and organization of their mammalian counterparts in their native states remain unknown. Here, we report the cryo electron microscopy (cryoEM) structures of the endogenous cubic 2-oxoglutarate dehydrogenase complex (OGDC) and icosahedral pyruvate dehydrogenase complex (PDC) cores from bovine kidney determined at 3.5 Å and 3.8 Å resolution, respectively. The structures of multiple proteins were reconstructed from a single lysate sample, allowing direct structural comparison without the concerns of differences arising from sample preparation and structure determination. Although native and recombinant E2 core scaffold structures are similar, native structures are decorated with their peripheral E1 and E3 subunits. Asymmetric sub-particle reconstructions support heterogeneity in the arrangements of these peripheral subunits. Additionally, despite sharing a similar monomeric fold, OGDC and PDC E2 cores have distinct interdomain and intertrimer interactions, which suggests a means of modulating self-assembly to mitigate heterologous binding between mismatched E2 species. The lipoyl moiety lies near a mobile gatekeeper within the interdomain active site of OGDC E2 and PDC E2. Analysis of the two-fold related intertrimer interface identified secondary structural differences and chemical interactions between icosahedral and cubic geometries of the core. Taken together, our study provides a direct structural comparison of OGDC and PDC from the same source and offers new insights into determinants of interdomain interactions and of architecture diversity among α-keto acid dehydrogenase complexes.

2.2 Introduction

The pyruvate dehydrogenase complex (PDC; components suffixed by 'p') is an essential element of life that catalyzes the oxidative decarboxylation of pyruvate to acetyl-coenzyme A (acetyl-CoA). This key metabolic reaction links glycolysis to oxidative phosphorylation and the Krebs cycle. PDC is a member of the α-keto acid dehydrogenase complex family, alongside 2-oxoglutarate dehydrogenase complex (OGDC; 'o') and branched-chain α-keto acid dehydrogenase complex (BCKDC; 'b'), which all perform analogous reactions in central metabolism[1,2]. Genetic studies have identified mutations in components of these multienzyme complexes linked to severe clinical consequences, including metabolic acidosis and neurodegeneration[3–7]. PDC also attracts interest in cancer biology for its role in modulating the Warburg effect to promote tumor anabolism[4,8–10].

PDC, OGDC, and BCKDC are each comprised of multiple copies of a substrate-specific dihydrolipoamide acetyltransferase (E2p), dihydrolipoamide succinyltransferase (E2o), or dihydrolipoamide acyltransferase (E2b) inner catalytic (IC) core surrounded by multiple copies of the respective α-keto acid dehydrogenase (E1p, E1o, or E1b) and universal dihydrolipoamide dehydrogenase (E3) components[1,2]. Mammalian E2p and E2b is comprised of one or two lipoyl domains (LDs) followed by a peripheral subunit-binding domain (PSBD) and an IC domain that are connected by flexible linkers (Fig. 2-1a). Mammalian E2o has a similar composition but lacks the classical PSBD of E2p and E2b[11–13]. Each α-keto acid dehydrogenase complex proceeds through a three-step mechanism, starting with the decarboxylation of the α-keto acid substrate by E1 and transfer of an acetyl, acyl, or succinyl functional group to lipoate in the LD of E2. Then, the functionalized LD localizes to the IC core, where the functional group is transferred from the lipoyl moiety to CoA. Lastly, E3 reoxidizes the lipoyl moiety, which enables the process to repeat with the production of NADH. The LD translocates between these active sites through a flexible swinging arm mechanism[14].

Although each E1 and E2 subunit in an α-keto acid dehydrogenase complex is highly similar in

both sequence and overall structure to its counterpart in the other classes of the family and the same E3 protein is shared by all classes within a tissue[1,15,16], there are notable differences in the architecture of how these components assemble. E1o is a homodimer, and E1p is either a homodimer or a heterotetramer in which each polypeptide of the homodimer is divided into an alpha and a beta subunit[17,18]. In the gram-negative bacteria *E. coli* and *A. vinelandii*, PDC, OGDC, and BCKDC cores all have a 24-mer cubic architecture. However, in eukaryotes and gram-positive bacteria, PDC cores adopt a different 60-mer icosahedral architecture despite OGDC and BCKDC cores retaining a cubic geometry[16,18–24]. As an exception, actinobacteria distinctly lack assembled cores of E2p and also possess a unique OdhA protein that fuses the E2o catalytic domain to E1o[25]. Eukaryotic PDC has an additional noncatalytic E3-binding protein (E3BP) that specifically binds E3 to the core and is structurally similar to E2[16,19,26–32]. In mammals, E3BP substitutes for E2p in the core scaffold at possible stoichiometric ratios of 48:12 or 40:20 E2p:E3BP in an unknown arrangement[28,31]. Although peripheral E1 and E3 subunits are bound to the core by a discrete PSBD located in E2 or E3BP, eukaryotic E2o distinctly lacks a discrete PSBD for the binding of E1 and E3[11–13]. Instead, E3 binds an N-terminal region of E1o that shares antisera reactivity with E2p and E3BP and that is similar to them in sequence[13,33,34]. In turn, E1o binds to E2o at a region comprised of residues from the IC domain and the preceding linker[13,22]. Nonetheless, the subunits of each α-keto acid dehydrogenase complex self-assemble into their respective complexes[24].

Much of our current structural understanding of α-keto acid dehydrogenase complexes is derived from crystal structures of recombinantly expressed proteins. However, the flexible assembly of intact α-keto acid dehydrogenase complexes is unamenable to crystallography, and components were often crystallized individually. Structures from these individual proteins are unable to fully capture how multiple components come together and function as a complex. Additionally, crystallization may impose artificial order on protein side chains[35]. In the absence of the structure of native complexes, we have an insufficient understanding of how these subunits are organized into functional units. Recently, structures of native PDC from fungus and bacteria became available from cryo electron microscopy

(cryoEM)[16,23,32,36,37]. In the intact fungus PDC structures, E3BP was observed as an additional component appended to the interior of the core instead of substituting for an E2p subunit of the core like in mammals and contains a conserved fold found in other E2 proteins[16,32,36]. From bacteria cultured in minimal media to minimize the activity of PDC, cryoEM density of the lipoyl domains interacting with the PDC IC domains could be observed. Despite these exciting progresses, no such endogenous, intact α-keto acid dehydrogenase complexes have been determined for mammals.

In this study, we report the cryoEM structures of endogenous OGDC and PDC E2 IC domain cores extracted from bovine kidney tissue at 3.5 Å and 3.8 Å resolution, respectively. We observe peripheral subunits around the core and the lipoyl moiety substrate within the E2 active site for both complexes. By comparing structures of PDC and OGDC from the same source, we identify distinct interactions in the IC domain trimer and in the two-fold related intertrimer interface, which may direct self-assembly within a milieu of similar components.

2.3 Results

2.3.1 Extraction, cryoEM, and identification of native, intact OGDC and PDC

To generate samples of intact α-keto acid dehydrogenase complexes, we extracted mitochondrial lysates from bovine (*Bos taurus*) kidney tissue. We then used sucrose gradient fractionation to separate and enrich the protein species present in the mitochondrial lysates (Fig. 2-1b). Fractions were evaluated by SDS-PAGE, western blot, and mass spectrometry to confirm the presence of PDC and OGDC (Fig. 2-1b; Supplementary Fig. 2-S1a). The mass spectrometry data indicated BCKDC to be present only as a minor species compared to PDC and OGDC. Fractions containing PDC and OGDC were pooled together into a single lysate sample for EM studies. Negative stain 2D class average evaluation of the sample on a single grid showed both icosahedral classes and cubic classes (Supplementary Fig. 2-S2). Negative stain classes also show peripheral densities around the cores, suggesting that E1 or E3 are present and that intact α-

keto acid dehydrogenase complexes were recovered. Higher-order assemblies are visible in cryoEM micrographs (Supplementary Fig. 2-S3a), and smeared densities are present at the periphery of well-resolved cores in the cryoEM 2D classes (Supplementary Fig. 2-S3b). Notably, additional classes belonging to other unknown protein species were obtained (Supplementary Figs. 2-S2, 2-S3b), thus highlighting the capability of cryoEM for the study of molecular sociology. CryoEM 3D reconstruction yielded cubic and icosahedral cores surrounded by peripheral densities (Figs. 2-1b, 2-2a, 2-3a).

To structurally characterize the putative OGDC and PDC in the pooled lysate, we obtained 3D reconstructions of the cubic and icosahedral cores of the complexes by single-particle cryoEM at 3.5 Å and 3.8 Å resolution, respectively (Figs. 2-2b, 2-3b; Supplementary Fig. 2-S3; Supplementary Table 2-S1). We initially identified PDC based on its distinct icosahedral architecture. This is confirmed by atomic model building of the bovine E2p sequence into the icosahedral cryoEM density map, which showed good agreement between the sequence and the respective side chain densities. FindMySequence results also support the identity assignment of PDC[38] (E-value 4.80e-74, 3.30e-19, and 2.40e-15 for E2p, E2o, and E2b sequences, respectively).

Although PDC is distinct for its icosahedral architecture, both OGDC and BCKDC share cubic architecture, which introduces ambiguity to the identification of cubic α-keto acid dehydrogenase complexes. To confirm whether our cubic complex was OGDC, BCKDC, or a hybrid of both, we created a homology model by mutating a pre-existing human E2o model to the bovine E2o sequence[22]. After fitting and real-space refinement of the model into our cubic cryoEM density map, the side chains for the E2o model fit the density well (Fig. 2-2e). Sequence alignment of bovine E2o and E2b reveals primary structure features in E2b that cannot correspond to side chain densities in our map. Additionally, fitting of a pre-existing E2b model into our map showed discrepancies between the modeled side chains and the density[20]. FindMySequence and checkMySequence results also support the identity assignment of OGDC[38,39] (E-value 5.30e-89 and 6.30e-11 for E2o and E2b sequences, respectively). The unambiguous identification of the cubic complex as OGDC is further reflected in the low presence of BCKDC

components in the mass spectrometry data.

2.3.2 Heterogeneous arrangement of peripheral subunits around the native E2 cores

Unmasked 3D reconstructions have external densities suggestive of E1 or E3 around the core (Figs. 2-2a, 2-3a). Both OGDC and PDC have a large density over each core face. In PDC, thin densities from the trimer vertices extend toward the density above the pentameric face. OGDC has additional smaller densities along each cubic edge with thin densities of putative linker regions connecting to the E2o trimer vertices. Previous models based on hydrogen/deuterium exchange and cross-linking mass spectrometry studies propose that E1o binds at the edges and E3 binds at the faces[13,22]. The opposite orientation of the small connecting densities between OGDC and PDC suggests different predominant states of activity and a large range of motion for the flexible linker regions. However, enforced symmetry during reconstruction restricts the interpretability of these external densities for the structures and locations of E1 and E3.

To better evaluate the heterogeneous features of the complexes, we performed asymmetric (C_1) reconstructions of the whole OGDC and PDC followed by a 20 Å low-pass filter and color zone of 16 Å around the fitted E2 core models (Supplementary Fig. 2-S4a, b). External densities corresponding to peripheral subunits are localized at certain faces in both OGDC and PDC (Supplementary Fig. 2-S4a, b), which has been previously observed in previous C_1 reconstructions and tomography[32,37,40]. Thin densities suggestive of linkers connect certain external densities to the E2 core.

Notably, in PDC, an unidentified density at the center of the icosahedral reconstruction is also present in the C_1 reconstruction (Fig. 2-3b; Supplementary Fig. 2-S4d). This globular density radiates outward from the center of the E2p/E3BP core interior with decreasing contour level, unlike the tetrameric E3BP densities adjacent to the inner E2p core in fungus PDC[16,32,36]. The central density is absent in recombinant E2p and mixed E2p-E3BP cores and endogenous OGDC[30,31] (Supplementary Fig. 2-S4c), which suggests that the central density is unique to native PDC and does not belong to E2 nor E3BP.

Peripheral subunit densities were further evaluated by symmetry expanded C_1 sub-particle reconstructions. A 20 Å low-pass filter and color zone of 8 Å around the E2 core model were applied to the sub-particles for greater interpretability of low-resolution features (Supplementary Fig. 2-S4e, f). Although the resolution is too low for unambiguous identification or positioning of E1, E3, LD, or PBSD within the external densities, we utilized the asymmetric sub-particle reconstructions to narrow down their potential localization. Thin densities corresponding to N-terminal linkers originate from both the two-fold related intertrimer interface and the N-terminal end of H1 (Supplementary Fig. 2-S4e) in E2o, but the linkers originate from only the H1 site in E2p/E3BP (Supplementary Fig. 2-S4f). A region extending along the N-terminal end of H1 through the preceding linker (residues 224 to 240) has previously been identified as the PSBD in mammalian E2o[13,22]. In native *E. coli* E2p during the resting state, the LD binds to a cleft flanked by the N-terminal end of H1 and the C-terminal end of H4[37]. In both PDC and OGDC, extra densities are located near the H1 sites of E2/E3BP subunits not lining the inner face of interest, and fitting of *E. coli* E2p with bound LDs into the E2o and E2p/E3BP sub-particles supports binding of the LD to these sites[37]. Due to the conflicting role of the H1 site as the PSBD in E2o, the H1 site of E2o may have a dual function as both the LD binding site and the PSBD, where the binding as either role is mutually exclusive. In PDC, the extra densities at the H1 sites protrude outward and some have thin extensions (Supplementary Fig. 2-S4f, g). These densities have varying prominence and are notably absent at some IC domains (Supplementary Fig. 2-S4g), suggesting a heterogeneous binding of LDs or E1/E3.

The facial density of OGDC localizes around a single trimer vertex and is connected to the E2o IC domains by the linker originating from the H1 sites (Supplementary Fig. 2-S4e). Additional smaller edge densities protrude from the linkers originating from the two-fold related interface. In PDC, only a large facial density is observed (Supplementary Fig. 2-S4f). Similar to that of OGDC, the facial density is localized near a trimer vertex and is connected to the IC domain at the H1 site. Linkers from the two-fold related interface of E2p/E3BP like those in E2o were not observed. Trajectories of the linker between the

IC domain and PSBD from the two-fold related interface and from along the three-fold axis have been previously observed but only mutually exclusively[16,20–25,30,37,41]. In particular, the N-terminal linker of native *E. coli* E2p in the resting state folds back towards the three-fold axis of the trimer after reaching the two-fold related interface[37]. Our asymmetric sub-particle reconstructions suggest that the fold of the E2 N-terminal linker is associated with the activity or reaction step of the complex, and both orientations of the linker possibly coexist based on the binding of LD, E1 or E3 (Supplementary Fig. 2-S4e). It remains unknown whether the linker folds back to position the PSBD and LD(s) against its own IC domain or continues past the three-fold axis to a neighboring IC domain.

2.3.3 E2o and E2p utilize distinct interactions to ensure correct self-assembly into E2 trimers and prevent incorrect heterologous binding of different E2 subunits

For E2o, the backbones of residues 220-455 are fully traceable. For E2p, the backbones of residues 420-647 are fully traceable, except for residues 519 and 520, which lack densities. Both E2o and E2p share conserved secondary structure features: six α-helices (H1-6), a short C-terminal 3_{10}-helix (H7), and ten β-strands (βA-J) (Figs. 2-2c, 2-3c). The secondary structure elements described here are labeled consistently with *E. coli* E2o, bovine E2b, and human E2p[20,30,42].

The E2 IC domains of OGDC and PDC are arranged as trimers (Figs. 2-2d, 2-3d) with 51 interdomain hydrogen bonds and an average buried surface area of 4957 $Å^2$ (standard deviation (SD) = 3.47) buried surface in E2o and an average of 42.75 (SD = 2.36) interdomain hydrogen bonds and an average buried surface area of 4397 $Å^2$ (SD = 22) in E2p. Although E2o and E2p share secondary structure features and overall fold, they have notable differences in certain regions and in the interactions that stabilize their respective trimers (Supplementary Fig. 2-S5a). The N-terminus extends at a less acute angle towards the neighboring IC domain at the clockwise position in E2o than in E2p (Supplementary Fig. 2-S5b). The N-terminus wraps around a turn between βC′ and βD′ (prime symbol is used to denote the neighboring IC domain at the clockwise position from the exterior view), which is longer in E2o than

in E2p. The longer βC-βD turn in E2o enables an electrostatic interaction between the negatively charged surface of the turn and the positively charged surfaces of the N-terminus that is not present in E2p (Supplementary Fig. 2-5d). This positions βA in a β-sheet with βD′ and βC′, and the increased stability allowed for modeling of additional N-terminal residues in E2o. The extended βC-βD turn is also present in human E2o and not E2p, but it is lacking in bacteria (Supplementary Fig. 2-S5c**)**. Additionally, E2p possesses a longer interior hairpin (Supplementary Fig. 2-S5a), possibly for engaging in interactions in the larger volume of the icosahedral interior of PDC compared to the cubic interior of OGDC[43].

Within the IC domain trimer, an E2 subunit binds to its clockwise adjacent neighbor through a short β-β motif between βH and βB′. In E2p, this binding is described by three main chain-to-main chain hydrogen bonds (Fig. 2-4a). In E2o, this binding is described by two main chain-to-main chain hydrogen bonds, a main chain-to-side chain hydrogen bond, and a slipped-parallel π-stacking interaction[44] (Phe383-Phe253′). The slipped-parallel π-stacking interaction of E2o occurs near the three-fold axis of the trimer and may form a delocalized π-electron network as previously suggested in *C. thermophilum* E2p between Arg384[36]. By sequence and structure alignment, Phe383 in bovine E2o is the equivalent residue of Arg384 in *C. thermophilum* E2p. Notably, in bovine E2o, the delocalized π-electron network is extended within the three-fold axis and encompasses Phe253, Phe383, and Tyr419 (Fig. 2-4b). This delocalized π-electron network is also present in human E2o but not in bovine E2p at this structurally analogous position because the equivalent bovine E2p residues Asn575 are too distant from each other[22] (Fig. 2-4b). Like mismatched puzzle pieces, the different interfaces of E2o and E2p may prevent incorrect assembly due to co-existence of their components within the cell.

2.3.4 The lipoyl moiety is positioned near a flexible gatekeeper in the catalytic site

In the cryoEM density maps of E2p, a region of flexibility has been identified: a β-turn connecting βE and βF near the interdomain active site (Fig. 2-5a). Due to missing density at the βE-βF β-turn, two residues (Ala519 and Gly520) are unmodelled in our E2p model (Fig. 2-5b**)**. Although the corresponding density is not missing in E2o, it is still weaker relative to that of the adjacent residues. This

missing or weak density has also been observed in previous cryoEM density maps[16,22,23,25,37]. The density of the adjacent residue (Leu521 in E2p and Leu329 in E2o) is well-defined. This leucine is a conserved gatekeeper residue for the binding of the lipoyl moiety[20,30]. This gatekeeper residue is proximal to the interdomain active site, near the catalytic residues Ser566 and His620′ in E2p, and positions the dihydrolipoamide (DHLA) into the catalytic site during a previously suggested CoA-bound state[20].

Adjacent to the region of missing density, a putative density for the lipoyl moiety was identified near the catalytic residues Ser566 and His620′ and the gatekeeper residue Leu521 (Fig. 2-5b). When the contour level is decreased, the density forms a thin extension through the active site channel and accommodates a model of DHLA[45] (adopted from PDB ID: 1EAE) (Fig. 2-5c). The E2o cryoEM density map also contains a putative lipoyl moiety density (Supplementary Fig. 2-S6a). Following the trajectory of these densities, the dihydrolipoyllysine enters the active site from the exterior of the E2 core at an entrance at the end of the LD binding site cleft between the C-terminal end of H1 and a turn connecting βH and βI1, as previously depicted in *E. Coli* E2p in the resting state[37]. Densities of the lipoyl moiety are not present in cryoEM density maps of native *C. thermophilum* E2p and E2o and *N. crassa* E2p[16,23,36].

In the E2o density map at low contour level, weak cylindrical densities reminiscent of β-strands from the LD or helices from the PSBD are observed to the left and to the right of the entrance of the cleft. The left densities are located similarly to that of the LD in native *E. coli* E2p in the resting state[37] (Supplementary Fig. 2-S6b). These additional densities extend along the cleft towards the active site entrance. We attempted to fit LD and PBSD models from NMR (PDB ID: 1LAC; 1W4H) and AlphaFold prediction[46–49] (AF-P11179-F1; AF-P11181-F1), but there are mismatches between the cylindrical densities and the LD models when the dihydrolipoyllysine-containing loop is positioned adjacent to the active site channel. These observations combined with structural variation between LD domain models suggest that the β-sheets of the LD may be conformationally variable and undergo conformational and positional changes during catalysis.

To search for E3BP, we expanded the symmetry of the trimer vertex and pentameric face sub-

particles and reconstructed them asymmetrically (C_1) (Supplementary Fig. 2-S7a). Due to low resolution of the C_1 reconstructions, we are unable to utilize side chain densities to distinguish sequence identity. Sequence alignment between E2p and E3BP and fitting of predicted E3BP structures (PDB ID: 6H60; AF-P22439-F1) show two notable regions of difference in E3BP within the IC domain[31,47,48]: a three-residue shorter linker between H2 and H3 and a three-residue longer interior hairpin between βI2 and βJ (Supplementary Fig. 2-S7b). We searched the classes of the trimer vertex and pentameric C_1 sub-particles but were unable to identify any differences corresponding to the presence of E3BP (Supplementary Fig. 2-S7c). E3BP is not evenly distributed in the trimers[31], and current 3D classification methods may place too much weight on the larger features in the conserved fold of E2p and E3BP, which may reduce the quality of asymmetric reconstructions of the interior hairpin. Heterogeneity of E3BP distribution precludes identification and experimental structural determination of E3BP from native mammalian PDC.

2.3.5 Different knob-and-socket intertrimer interactions prevent heterologous self-assembly of E2 trimers

E2o and E2p have near-identical structures, yet they assemble into core scaffolds with different geometry. The cubic and icosahedral geometries are related by principles of quasi-equivalence and Euclidean geometry[24]. Each IC domain trimer vertex of the core is bound to its neighboring trimers through a palindromic two-fold related intertrimer interface[22,30] (Fig. 2-6a; Supplementary Fig. 2-S8a). The hydrophobic C-terminal 3_{10}-helix (H7) of each IC domain binds to a hydrophobic pocket formed by residues of H2*, H7*, and the N-terminal end of H4* (the asterisk is used to denote the partner IC domain across the two-fold related interface) on the opposite IC domain in a knob-and-socket interaction (Supplementary Fig. 2-S8b). This binding is further stabilized by electrostatic interactions from oppositely charged surfaces lining the exteriors of H7 and the hydrophobic pocket (Supplementary Fig. 2-S8c). In actinobacteria E2p, where H7 folds back against rather than away from the IC domain, the knob-and-socket interaction cannot occur, so the trimers are unable to assemble into a core[25].

The symmetries for bovine E2o and E2p are obtained from a different combination of interactions (Supplementary Table 2-S2). The intertrimer interface has an average of 18.7 (SD = 0.49) hydrogen bonds and an average buried surface area of 1102 $Å^2$ (SD = 0.46) in E2o and an average of 5.6 (SD = 0.63) hydrogen bonds and an average buried surface area of 825 $Å^2$ (SD = 14.18) in E2p. The exterior of the E2o socket is more charged than that of E2p (Supplementary Fig. 2-S8c), which leads to the greater number of hydrogen bonds at the E2o intertrimer interface. Additionally, the exterior surface of the socket in E2o and E2p are oppositely charged from one another. Although the residues comprising the hydrophobic sockets between the two species are similar, the shapes of the sockets are different to accommodate likewise different knobs. The E2p knob has one hydrophobic surface that binds a single hydrophobic pocket in the socket. In contrast, the E2o knob has an additional hydrophobic surface on the opposite side of the knob, and the E2o socket contains two pockets to accommodate this difference (Supplementary Fig. 2-S8b). Similar to the purpose of the interdomain interface differences to prevent heterologous assembly of E2 subunits, the oppositely charged exterior surfaces and differently shaped knob-and-socket of the intertrimer interface between E2o and E2p suggest a means of preventing heterologous assembly of E2 trimers during core formation.

To compare the relative positions of trimers at the intertrimer interface between E2o and E2p, the IC domains of the two-fold related interface were aligned by superimposition at one trimer vertex (Fig. 2-6b). The fold of H2 and the position of H7 differ between E2o and E2p. H7 of E2o (Pro448-Leu453) has a length of six residues and has two additional residues (Asp454 and Leu455) appended at the C-terminus. H7 of E2p (Pro642-Leu646) has a length of five residues and one additional residue (Leu647) at the C-terminus. H7 is positioned more closely to H4* in E2p than in E2o, which is enabled in E2p by an additional hydrogen bond between H7 and H4* and fewer hydrogen bonds between H7 and H2* (Supplementary Table 2-S2). In E2o, H2 has a length of 18 residues (Ser259-His276) and is comprised of a helix-kink-helix with the kink starting at residue 268. In E2p, H2 instead has a length of 14 residues (Gly455-Leu468) and is a single, uninterrupted helix (Fig. 2-6c). The H2 kink in E2o positions the C-

terminal end of the helix more closely against H7* for hydrogen bonds between Lys275 and His276 with Asp454* and Asp454*, respectively. The different interactions at the knob-and-socket stabilize E2 IC trimers at angles that enable the Euclidean geometry relationship between a cube and icosahedron.

2.4 Discussion

Advances in cryoEM and its integration with mass spectrometry analysis have enabled the structural characterization of intact, native complexes[38,50,51]. Endogenous methods have recently been applied to determine structures of PDC in fungi and bacteria[16,23,32,36,37]. Here, we extended these methods to the study of mammalian tissues and determined structures of native bovine OGDC and PDC at sufficient resolutions for model building of their E2 IC domains (Figs. 2-2, 2-3). Despite exhaustive classification and refinement of an enormous data set, structures beyond the E2 cores in our asymmetric sub-particle reconstructions were only resolved at low resolution (Supplementary Fig. 2-S4), indicating heterogeneity in the arrangements of the peripheral subunits attached to the core. Nonetheless, analysis of OGDC and PDC from the same source enables direct comparison between the two protein species. Although the E2 IC domains of PDC and OGDC share a nearly identical fold to conduct analogous reactions, OGDC and PDC cores differ in their interdomain (Fig. 2-4; Supplementary Fig. 2-S5) and intertrimer (Fig. 2-6; Supplementary Fig. 2-S8; Supplementary Table 2-S2) interactions. These nuanced differences suggest a means of modulating self-assembly to mitigate heterologous binding between mismatched E2 species and highlight the divergent evolution of α-keto acid dehydrogenase complexes to meet the requirements of different organisms with varied cellular contents and metabolic environments. Additionally, native structures retain biologically relevant features, such as the position of the lipoyl moiety within the E2 active site (Fig. 2-5; Supplementary Fig. 2-S6a), and present opportunities for discovery in molecular sociology (Supplementary Figs. 2-S4d, 2-S6b) that may be lost in recombinant samples.

Besides the structures reported here, native structures of α-keto acid dehydrogenase complexes

from the same endogenous source only became available early this year for the fungus *C. thermophilum*, where both PDC and OGDC were structurally characterized[23]. In contrast to our bovine E2o and E2p structures, the back-folded N-terminal region comprised of βA and H1 is absent in *C. thermophilum* E2p. Interestingly, H1 is still present in *C. thermophilum* E2o from the same size-exclusion chromatography fraction[23], and is also present in native *N. crassa* and *E. coli* E2p[16,37]. Given that the structured conformation of the N-terminal region is stabilized by electrostatic interactions between βA and H1 with the βC-βD turn (Supplementary Fig. 2-S5d), the back-folded conformation may exist transiently depending on the activity state of the complex.

The unfolded state of H1 has been suggested to be necessary for unhindered insertion of the lipoyl moiety[36]. For *E. coli* E2p in the absence of pyruvate, H1 is folded and binding of the N-terminus extends towards the three-fold axis, which was suggested to position and stabilize the E2-LD complex in a resting state[37]. Our sub-particle reconstruction of E2o shows the N-terminal region extending outwards from both the two-fold related interface and about the three-fold axis simultaneously (Supplementary Fig. 2-S4e), which indicates that multiple conformations of the PSBD-like H1 and preceding linker of E2o are present. The conformation of the linker may correspond to the transient and mutually exclusive binding of either LD to the E2o IC domain or of E1/E3 to the PSBD.

Native structures of PDC and OGDC have been captured in various activity states. While the tertiary structure of E2 remains mostly identical between structures, the conformation of the aspartate (Asp624 in PDC and Asp430 in OGDC of bovine) that is responsible for stabilizing the protonated catalytic histidine within the active site differs. Previous crystallization experiments of *A. vinelandii* E2p with reduced forms of CoA or dihydrolipoamide yielded structures with the sidechain of the aspartate-equivalent (Asn614) facing that of the catalytic histidine while other states of E2p have the aspartate-equivalent rotated away from the histidine[45]. In our E2o and E2p structures, the stabilizing aspartate faces the catalytic histidine (Fig. 2-5b, c; Supplementary Fig. 2-S6a), which suggests that substrate is bound in the active site. This asparagine is also facing towards the histidine in the native structure of *N. crassa*

E2p[16], but the asparagine is rotated away in the native structures of *C. thermophilum* E2p and E2o and *E.coli* E2p in the resting state[23,36,37]. In recombinant human E2p and E2o, the asparagine is modeled as facing towards the histidine[22,30]. However, the EM density for this asparagine side chain is missing, indicating motility of the side chain in the apo state. There could be a potential mixture of activity or assembly states of OGDC and PDC in different lysate fractions when complexes are sorted by size, and our single-particle cryoEM analysis of the entire core may have only reconstructed the most stable or populous conformation or a potential hybrid of the active and resting states.

The interdomain interactions between an E2 IC domain and its neighbor are not conserved between E2p and E2o. Bovine and human E2o contain an extensive delocalized π-electron network at the three-fold axis (Fig. 2-4b) that is absent in their E2p counterparts[22,30]. The differences in interdomain interfaces between E2p and E2o within an organism may enable correct self-assembly within a milieu of components in the mitochondrial matrix. *C. thermophilum* E2o may also have a similar delocalized π-electron network, but instead with Met350 at the analogous position of Phe383 in E2o[23]. Met-aromatic motifs can provide additional stabilization compared to purely hydrophobic interactions[52]. Rather than a delocalized π-electron network of aromatic residues, *C. thermophilum* E2p instead possesses a potential stabilizing arginine cluster that is absent in both its E2o counterpart, mammalian E2p, and mesophilic fungus *N. crassa* E2p[16,23,36]. The greater interdomain interactions could contribute towards increased thermal stability needed for thermophiles compared to mesophiles as a result of divergent evolution.

While there are pre-existing structures of α-keto acid dehydrogenase complexes, the diversity of interactions within these complexes between different organisms occludes direct comparison for determinants of assembly. Although *E.coli* E2p and bovine E2b are both cubic like bovine E2o, their interdomain interface interactions differ greatly. The knob of *E. coli* E2p has a semicylindrical hydrophobic surface that complements a narrow, elongated hydrophobic socket[37]. Bovine E2b has a double-sided knob like bovine E2o[20], but the hydrophobic surfaces are rotated about and have different sizes. Although α-keto acid dehydrogenase complexes can share the same geometry, the interdomain

interactions that hold the complex cores together differ between both family members and model organisms. Thus, structural characterization of samples from the same organism is needed to describe distinct interactions.

In mammals, E3BP substitutes for a corresponding E2 in the core scaffold as determined by biochemical analysis instead of being an additional subunit as it is in fungus[16,28,31,32]. Although there is an internal density for our native bovine PDC structure (Supplementary Fig. 2-S4d), the density is centralized and does not resemble the tetrahedral arrangement of fungal E3BP[36]. In addition to the subtle differences between E2p and E3BP in the IC domain (Supplementary Fig. 2-S7b), the linker region between the PSBD and IC domain is twice as long in E3BP compared to that in E2p (Fig. 2-1a). Due to the short length of the PSBD-IC domain linker in E2p, E1p is expected to be close to the core. Likewise, the longer linker in E3BP enables binding of E3 further away from the core, possibly to better accommodate the larger size of E3 compared to E1p. Because of the heterogeneous arrangement of E3BP and peripheral E1 and E3 subunits, PDC and other α-keto acid dehydrogenase complexes may need to be studied at the individual level. Tomography could identify peripheral subunits and avoid issues of classification for E3BP if sufficient resolution could be reached in the future.

Advancements in cryoEM have enabled atomic modeling of native α-keto acid dehydrogenase complexes and other endogenous complexes. Looking forward, faster direct electron detectors and improved cryoEM grid preparation methods will improve acquisition of particles for low population species for unambiguous identification and atomic modeling. Heterogeneous reconstruction could enable visualization of distinct protein conformations[53,54]. Machine-learning-based methods in particle picking[55], structure prediction[47], and identification can resolve unknown identities in a society of proteins[38]. Structural determination of native geometrically variable complexes from similar folds, exemplified by the cubic and icosahedral complexes presented here, should not only provide new biological insight but also inform protein engineering applications by providing an opportunity for comparison within the same environment to identify the determinants of protein architecture.

2.5 Materials and Methods

2.5.1 Preparation of bovine mitochondria

Bovine mitochondria were prepared from *Bos taurus* kidneys as described with modifications[56]. 6 kg of bovine kidneys were collected and chilled on ice immediately after slaughter. Cortical tissues were cut into small slices (0.5-1 cm) and soaked in 4 L water for 1 h and then washed with 2 L kidney buffer (20 mM potassium phosphate pH 7.6, 250 mM sucrose, 1 mM EDTA). The slices were passed through an electric meat grinder, and the ground meat was suspended in kidney buffer with 10 mM 2-mercaptoethanol and diluted to a final volume of 12 L. The suspension was homogenized using an Ultra-Turrax (IKA) homogenizer (40 s), filtered through 8 layers of cheesecloth accompanied with a sieve, and diluted to a total volume of 18 L using kidney buffer. The resulting supernatant was centrifuged at 2,000-3,000 × g for 10 min. The supernatant was decanted and further centrifuged at 5,600 × g for 25 min. The lysosome-enriched fluffy layer was carefully removed from the mitochondrial pellet. The pellets were resuspended in a total of 8 L kidney buffer and filtered again through 8 layers of cheesecloth. The suspension was homogenized using an Ultra-Turrax homogenizer (30 s), diluted to 12 L with kidney buffer, and centrifuged at 5,600 × g for 25 min. The pellets were resuspended with kidney buffer with 0.2 mM PMSF, 0.25 μg/ml aprotinin, 0.14 μg/ml pepstatin, and 1 μg/ml leupeptin in a total volume of 4 L. The suspension was treated with 0.01% digitonin for 15 min to remove outer mitochondrial membranes and diluted to 12 L using kidney buffer. The resulting mitoplasts were concentrated by centrifugation at ~25,600 × g for 20 min followed by removal of excess kidney buffer. The pellet of the final step was flash frozen in liquid nitrogen and stored at −80 °C.

2.5.2 Isolation of bovine PDC and OGDC complexes

100 g of frozen mitoplasts were thawed in 150 mL lysis buffer [20 mM potassium phosphate pH 7.6, 50 mM NaCl, 0.02 mM thiamin pyrophosphate, 20 mM $MgCl_2$, 2 mM EGTA, protease inhibitors (2 mM PMSF, 2.5 μg/mL aprotinin, 1.4 μg/mL pepstatin, and 10 μg/mL leupeptin)]. The sample was diluted to

337.5 mL with lysis buffer. 37.5 mL Triton X-100 buffer [lysis buffer components in addition with 16% (v/v) Triton X-100] was then added, and the sample was stirred for 15 min at 4 °C. The suspension was centrifuged (SLA-3000, 13,000 rpm, 20 min, 4 °C), and the supernatant was PEG precipitated in 5% (w/v) PEG 10,000 for 15 min. The precipitate was collected by centrifugation (2,500 × g, 7 min, 4 °C) and re-suspended in 35 mL lysis buffer. The suspension was homogenized using a Dounce homogenizer and centrifuged (SW28 Ti rotor, 28,000 rpm, 17 min, 4 °C). The resulting supernatant was loaded onto 50% (w/v) sucrose cushions (15 ml) and centrifuged (SW41 rotor, 40,000 rpm, 24 h, 4 °C). Pellets were each dissolved in 0.5 mL sample buffer (50 mM MOPS-KOH buffer pH 7.0, 0.02 mM thiamin pyrophosphate, 10 mM $MgCl_2$, 0.1 mM EGTA, protease inhibitors, 1mM DTT) by shaking at 230 rpm for 1 h and clarified by centrifugation (SLA-3000, 16,000 rpm, 20 min, 4 °C). The sample was loaded onto 10–40% (w/v) sucrose gradients (1 mL per gradient) and centrifuged (SW41 Ti, 26,000 rpm, 12 h, 4 °C). The gradients were fractionated and assessed by SDS-PAGE and western blot using a PDC complex antibody cocktail (Abcam) that contains four different mAbs reacting specifically with E1α, E1β, and E2/E3BP. The fractions containing PDC and OGDC complexes were pooled, dialyzed to final buffer (50 mM MOPS-KOH buffer pH 7.0, 0.02 mM thiamin pyrophosphate, 1 mM $MgCl_2$, 1 mM NAD+, 0.1 mM EGTA, protease inhibitors, 1 mM DTT), and concentrated for cryoEM analysis. Liquid chromatography-tandem mass spectrometry was performed on the sucrose gradient fractions. Gradient fractions corresponding to OGDC were subjected to gel filtration on a Superose 6 column (GE Healthcare) to remove ferritin.

2.5.3 CryoEM sample preparation and image acquisition

CryoEM grids were prepared by using an FEI Mark IV Vitrobot (Thermo Fisher Scientific). 3 μL of sample was applied onto a glow-discharged lacey carbon copper grid with a thin, continuous carbon film (Ted Pella). After waiting for 30 s, the grid was blotted (8°C, 100% humidity, 10 s blotting time, 1 blotting force) and plunge-frozen in liquid ethane. Grids were loaded into a Titan Krios (Thermo Fisher Scientific) equipped with a Gatan Imaging Filter (GIF) Quantum LS and a Gatan K2 Summit direct

electron detector. Movies were recorded with SerialEM at super-resolution mode[57]. The nominal magnification was 105,000 ×, corresponding to a calibrated pixel size of 0.68 Å at the specimen level. The total exposure time for each movie was set to 8 s and fractionated equally into 40 frames, leading to a total dosage of ~45 electrons/Å^2. Defocus was set to -1.8 to -2.6 μm.

2.5.4 Image processing

The movies were motion corrected with MotionCor2[58]. 6,959 good images were selected from a total of 9,029 by manual screening. After defocus determination by CTFFIND4[59], particles were automatically picked using Gautomatch (https://www2.mrc-lmb.cam.ac.uk/research/locally-developed-software/zhang-software/). 33,138 particles for PDC complex and 41,165 particles for OGDC complex were obtained by 2D and 3D classification in RELION-3[60]. FSC calculations were performed in RELION-3.

For PDC complex, all particles from the previous step were recentered and extracted with a box size of 576 pixels. These particles were refined with icosahedral symmetry, resulting in a map with a resolution of 3.8 Å. To get better density for the peripheral E1 and E3 subunits, a structure-guided sub-particle strategy was used. First, the STAR file from refinement was expanded with icosahedral symmetry, leading to a total of 1,988,280 particles. Second, the center of the expanded particles was shifted to one trimer or pentameric face. Sub-particles were extracted with a box size sufficient for accommodating E1 and E3. Finally, 3D classifications with a local mask focused on either the internal or external region were performed. No classes with identifiable features were obtained for the internal region. For the external region, one class with 13.4% of all particles best displays a connecting density between E2p and the putative E1/E3/LD/PSBD. The particles in this class were subjected to further classification with different parameters (skip-align classification and local search classification), but they did not yield improved density maps for E1/E3/LD/PSBD.

For OGDC complex, all particles from the previous step were recentered and extracted with a box size of 384 pixels. These particles were refined with octahedral symmetry and a mask around the E2 core,

resulting in a map with a resolution of 3.5 Å. To get better density for the peripheral E1 and E3 subunits, a structure-guided sub-particle strategy was used. First, the STAR file from refinement was expanded with octahedral symmetry, leading to a total of 987,960 particles. Second, the center of the expanded particles was shifted to one face of the cubic core, and particles were extracted in RELION with a box size of 208. Finally, 3D classification with a local mask of the external area was performed. One class with 12.7% of all particles showed the best connections between E2 and the exterior E1/E3. The particles in this class were subjected to further classification with different parameters (skip-align classification and local search classification). However, no maps with better density of exterior E1/E3/LD were obtained.

2.5.5 Atomic modeling

For model building, human E2o (PDB ID: 6H05) and human E2p (PDB ID: 6CT0) were used as templates for bovine E2o and E2p IC domains, respectively[22,30]. Each monomeric subunit model was fitted into the respective cryoEM density map of the core scaffold using ChimeraX[61], mutated and manually refined in Coot[62], and real-space refined in Phenix[63]. The monomer model was then duplicated to the appropriate stoichiometry and fit into the respective map and refined iteratively using ISOLDE and Phenix[64]. FindMySequence and checkMySequence were later used to confirm the identity of OGDC[38,39]. Electrostatic potential maps were calculated in ChimeraX. Sequence alignments were performed with Clustal Omega and EMBOSS Needle on the EMBL-EBI server[65–67]. Structure predictions were obtained from AlphaFold using ColabFold or the EBI database[47,48,68].

2.6 Data Availability

Mass spectrometry data and source gel/blot images are provided with this paper. CryoEM density maps of the cubic bovine OGDC core and icosahedral bovine PDC core have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-26649 and EMD-26650, respectively. The coordinates of the E2o IC domain and bovine E2p IC domain have been deposited in the Protein Data

Bank under accession numbers 7UOL and 7UOM, respectively.

2.7 Acknowledgements

This research was supported in part by grants from NSF (DMR-1548924) and NIH (R01GM071940). J.Z. acknowledges support from an Interdisciplinary Training in Virology and Gene Therapy training grant (NIH 5T32AI060567). We acknowledge the use of resources in the Electron Imaging Center for Nanomachines supported by UCLA and grants from NIH (S10RR23057, S10OD018111, and U24GM116792) and NSF (DBI-1338135). We thank the UCLA Proteome Research Center for assistance in mass spectrometry.

2.8 Contributions

Z.H.Z. conceived the project and supervised the research. S.L designed the experimental protocols. S.L. and X.X. prepared samples, collected cryoEM images, and determined the 3D structures. Z.L. and J.Z. built atomic models and made figures. J.Z. interpreted the results and wrote the manuscript. All authors reviewed and approved the paper.

2.9 Conflict of Interest

The authors declare no competing interests.

2.9 References

1. Patel, M. S. & Roche, T. E. Molecular biology and biochemistry of pyruvate dehydrogenase complexes. *FASEB J.* **4**, 3224–3233 (1990).
2. Reed, L. J. A Trail of Research from Lipoic Acid to α-Keto Acid Dehydrogenase Complexes. *J. Biol. Chem.* **276**, 38329–38336 (2001).
3. Ævarsson, A. *et al.* Crystal structure of human branched-chain α-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease. *Structure* **8**, 277–291 (2000).
4. Roche, T. E. & Hiromasa, Y. Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer. *Cell. Mol. Life Sci.* **64**, 830 (2007).
5. Shi, Q. *et al.* Inactivation and Reactivation of the Mitochondrial α-Ketoglutarate Dehydrogenase Complex. *J. Biol. Chem.* **286**, 17640–17648 (2011).
6. Burrage, L. C., Nagamani, S. C. S., Campeau, P. M. & Lee, B. H. Branched-chain amino acid metabolism: from rare Mendelian diseases to more common disorders. *Hum. Mol. Genet.* **23**, R1–R8 (2014).
7. Whitley, M. J. *et al.* Pyruvate dehydrogenase complex deficiency is linked to regulatory loop disorder in the αV138M variant of human pyruvate dehydrogenase. *J. Biol. Chem.* **293**, 13204–13213 (2018).
8. McFate, T. *et al.* Pyruvate Dehydrogenase Complex Activity Controls Metabolic and Malignant Phenotype in Cancer Cells. *J. Biol. Chem.* **283**, 22700–22708 (2008).
9. Tran, Q., Lee, H., Park, J., Kim, S.-H. & Park, J. Targeting Cancer Metabolism - Revisiting the Warburg Effects. *Toxicol. Res.* **32**, 177–193 (2016).
10. Chen, J. *et al.* Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer. *Nat. Genet.* **50**, 219–228 (2018).
11. Nakano, K. *et al.* Purification and molecular cloning of succinyltransferase of the rat alpha-ketoglutarate dehydrogenase complex. Absence of a sequence motif of the putative E3 and/or E1 binding site. *J. Biol. Chem.* **266**, 19013–19017 (1991).
12. Nakano, K. *et al.* Isolation, Characterization and Structural Organization of the Gene and Pseudogene for the Dihydrolipoamide Succinyltransferase Component of the Human 2-Oxoglutarate Dehydrogenase Complex. *Eur. J. Biochem.* **224**, 179–189 (1994).
13. Zhou, J. *et al.* A multipronged approach unravels unprecedented protein–protein interactions in the human 2-oxoglutarate dehydrogenase multienzyme complex. *J. Biol. Chem.* **293**, 19213–19227 (2018).
14. Perham, R. N. Swinging Arms and Swinging Domains in Multifunctional Enzymes: Catalytic Machines for Multistep Reactions. *Annu. Rev. Biochem.* **69**, 961–1004 (2000).
15. Yu, X. *et al.* Structures of the Human Pyruvate Dehydrogenase Complex Cores: A Highly Conserved Catalytic Center with Flexible N-Terminal Domains. *Structure* **16**, 104–114 (2008).
16. Forsberg, B. O., Aibara, S., Howard, R. J., Mortezaei, N. & Lindahl, E. Arrangement and symmetry of the fungal E3BP-containing core of the pyruvate dehydrogenase complex. *Nat. Commun.* **11**, 4667 (2020).

17. Frank, R. A. W., Price, A. J., Northrop, F. D., Perham, R. N. & Luisi, B. F. Crystal Structure of the E1 Component of the Escherichia coli 2-Oxoglutarate Dehydrogenase Multienzyme Complex. *J. Mol. Biol.* **368**, 639–651 (2007).

18. Reed, L. J. Multienzyme complexes. *Acc. Chem. Res.* **7**, 40–46 (1974).

19. Stoops, J. K. *et al.* On the Unique Structural Organization of the Saccharomyces cerevisiae Pyruvate Dehydrogenase Complex. *J. Biol. Chem.* **272**, 5757–5764 (1997).

20. Kato, M. *et al.* A synchronized substrate-gating mechanism revealed by cubic-core structure of the bovine branched-chain α-ketoacid dehydrogenase complex. *EMBO J.* **25**, 5983–5994 (2006).

21. Andi, B. *et al.* Structure of the dihydrolipoamide succinyltransferase catalytic domain from Escherichia coli in a novel crystal form: a tale of a common protein crystallization contaminant. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **75**, 616–624 (2019).

22. Nagy, B. *et al.* Structure of the dihydrolipoamide succinyltransferase (E2) component of the human alpha-ketoglutarate dehydrogenase complex (hKGDHc) revealed by cryo-EM and cross-linking mass spectrometry: Implications for the overall hKGDHc structure. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1865**, 129889 (2021).

23. Skalidis, I. *et al.* Cryo-EM and artificial intelligence visualize endogenous protein community members. *Structure* **30**, 575-589 (2022).

24. Izard, T. *et al.* Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. *Proc. Natl. Acad. Sci.* **96**, 1240–1245 (1999).

25. Bruch, E. M. *et al.* Actinobacteria challenge the paradigm: A unique protein architecture for a well-known, central metabolic complex. *Proc. Natl. Acad. Sci.* **118**, (2021).

26. Marcucci, O. D. & Lindsay, J. G. Component X. An immunologically distinct polypeptide associated with mammalian pyruvate dehydrogenase multi-enzyme complex. *Eur. J. Biochem.* **149**, 641–648 (1985).

27. Jilka, J. M., Rahmatullah, M., Kazemi, M. & Roche, T. E. Properties of a newly characterized protein of the bovine kidney pyruvate dehydrogenase complex. *J. Biol. Chem.* **261**, 1858–1867 (1986).

28. Hiromasa, Y., Fujisawa, T., Aso, Y. & Roche, T. E. Organization of the Cores of the Mammalian Pyruvate Dehydrogenase Complex Formed by E2 and E2 Plus the E3-binding Protein and Their Capacities to Bind the E1 and E3 Components. *J. Biol. Chem.* **279**, 6921–6933 (2004).

29. Brautigam, C. A. *et al.* Structural Insight into Interactions between Dihydrolipoamide Dehydrogenase (E3) and E3 Binding Protein of Human Pyruvate Dehydrogenase Complex. *Structure* **14**, 611–621 (2006).

30. Jiang, J. *et al.* Atomic Structure of the E2 Inner Core of Human Pyruvate Dehydrogenase Complex. *Biochemistry* **57**, 2325–2334 (2018).

31. Prajapati, S. *et al.* Structural and Functional Analyses of the Human PDH Complex Suggest a "Division-of-Labor" Mechanism by Local E1 and E3 Clusters. *Structure* **27**, 1124-1136.e4 (2019).

32. Kyrilis, F. L. *et al.* Integrative structure of a 10-megadalton eukaryotic pyruvate dehydrogenase complex from native cell extracts. *Cell Rep.* **34**, 108727 (2021).

33. Rice, J. E., Dunbar, B. & Lindsay, J. G. Sequences directing dihydrolipoamide dehydrogenase (E3) binding are located on the 2-oxoglutarate dehydrogenase (E1) component of the mammalian 2-oxoglutarate dehydrogenase multienzyme complex. *EMBO J.* **11**, 3229–3235 (1992).

34. McCartney, R. G. *et al.* Subunit Interactions in the Mammalian α-Ketoglutarate Dehydrogenase Complex: Evidence For Direct Association Of The A-Ketoglutarate Dehydrogenase And Dihydrolipoamide Dehydrogenase Components. *J. Biol. Chem.* **273**, 24158–24164 (1998).

35. Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A. & Subramaniam, S. Structure of β-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci.* **111**, 11709–11714 (2014).

36. Tüting, C. *et al.* Cryo-EM snapshots of a native lysate provide structural insights into a metabolon-embedded transacetylase reaction. *Nat. Commun.* **12**, 6933 (2021).

37. Škerlová, J., Berndtsson, J., Nolte, H., Ott, M. & Stenmark, P. Structure of the native pyruvate dehydrogenase complex reveals the mechanism of substrate insertion. *Nat. Commun.* **12**, 5277 (2021).

38. Chojnowski, G. *et al.* findMySequence: a neural-network-based approach for identification of unknown proteins in X-ray crystallography and cryo-EM. *IUCrJ* **9**, 86-97 (2022).

39. Chojnowski, G. Sequence-assignment validation in cryo-EM models with checkMySequence. *Acta Crystallogr. Sect. Struct. Biol.* **78**, 806–816 (2022).

40. Murphy, G. E. & Jensen, G. J. Electron Cryotomography of the E. coli Pyruvate and 2-Oxoglutarate Dehydrogenase Complexes. *Structure* **13**, 1765–1773 (2005).

41. Mattevi, A. *et al.* Atomic Structure of the Cubic Core of the Pyruvate Dehydrogenase Multienzyme Complex. *Science* **255**, 1544–1550 (1992).

42. Knapp, J. E. *et al.* Crystal structure of the truncated cubic core component of the Escherichia coli 2-oxoglutarate dehydrogenase multienzyme complex. *J. Mol. Biol.* **280**, 655–668 (1998).

43. Zhou, Z. H. *et al.* Direct Evidence for the Size and Conformational Variability of the Pyruvate Dehydrogenase Complex Revealed by Three-dimensional Electron Microscopy: The "Breathing" Core And Its Functional Relationship To Protein Dynamics. *J. Biol. Chem.* **276**, 21704–21713 (2001).

44. Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M. & Tanabe, K. Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc.* **124**, 104–112 (2002).

45. Mattevi, A., Obmolova, G., Kalk, K. H., Teplyakov, A. & Hol, W. G. J. Crystallographic analysis of substrate binding and catalysis in dihydrolipoyl transacetylase (E2p). *Biochemistry* **32**, 3887–3901 (1993).

46. Dardel, F., Davis, A. L., Laue, E. D. & Perham, R. N. Three-dimensional Structure of the Lipoyl domain from Bacillus stearothermophilus Pyruvate Dehydrogenase Multienzyme Complex. *J. Mol. Biol.* **229**, 1037–1048 (1993).

47. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

48. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).

49. Ferguson, N. *et al.* Ultra-fast Barrier-limited Folding in the Peripheral Subunit-binding Domain Family. *J. Mol. Biol.* **353**, 427–446 (2005).

50. Ho, C.-M. *et al.* Bottom-up structural proteomics: cryoEM of protein complexes enriched from the cellular milieu. *Nat. Methods* **17**, 79-85 (2020).

51. Ho, C.-M. *et al.* Native structure of the RhopH complex, a key determinant of malaria parasite nutrient acquisition. *Proc. Natl. Acad. Sci.* **118**, e2100514118 (2021).

52. Valley, C. C. *et al.* The Methionine-aromatic Motif Plays a Unique Role in Stabilizing Protein Structure. *J. Biol. Chem.* **287**, 34979–34991 (2012).

53. Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185 (2021).

54. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).

55. Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).

56. Pettit, F. H. & Reed, L. J. [65] Pyruvate dehydrogenase complex from bovine kidney and heart. in *Methods in Enzymology* vol. 89 376–386 (Academic Press, 1982).

57. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

58. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

59. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

60. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

61. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

62. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

63. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877 (2019).

64. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 519–530 (2018).

65. Rice, P., Longden, I. & Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000).

66. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

67. McWilliam, H. *et al.* Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **41**, W597–W600 (2013).

68. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

2.10 Figures

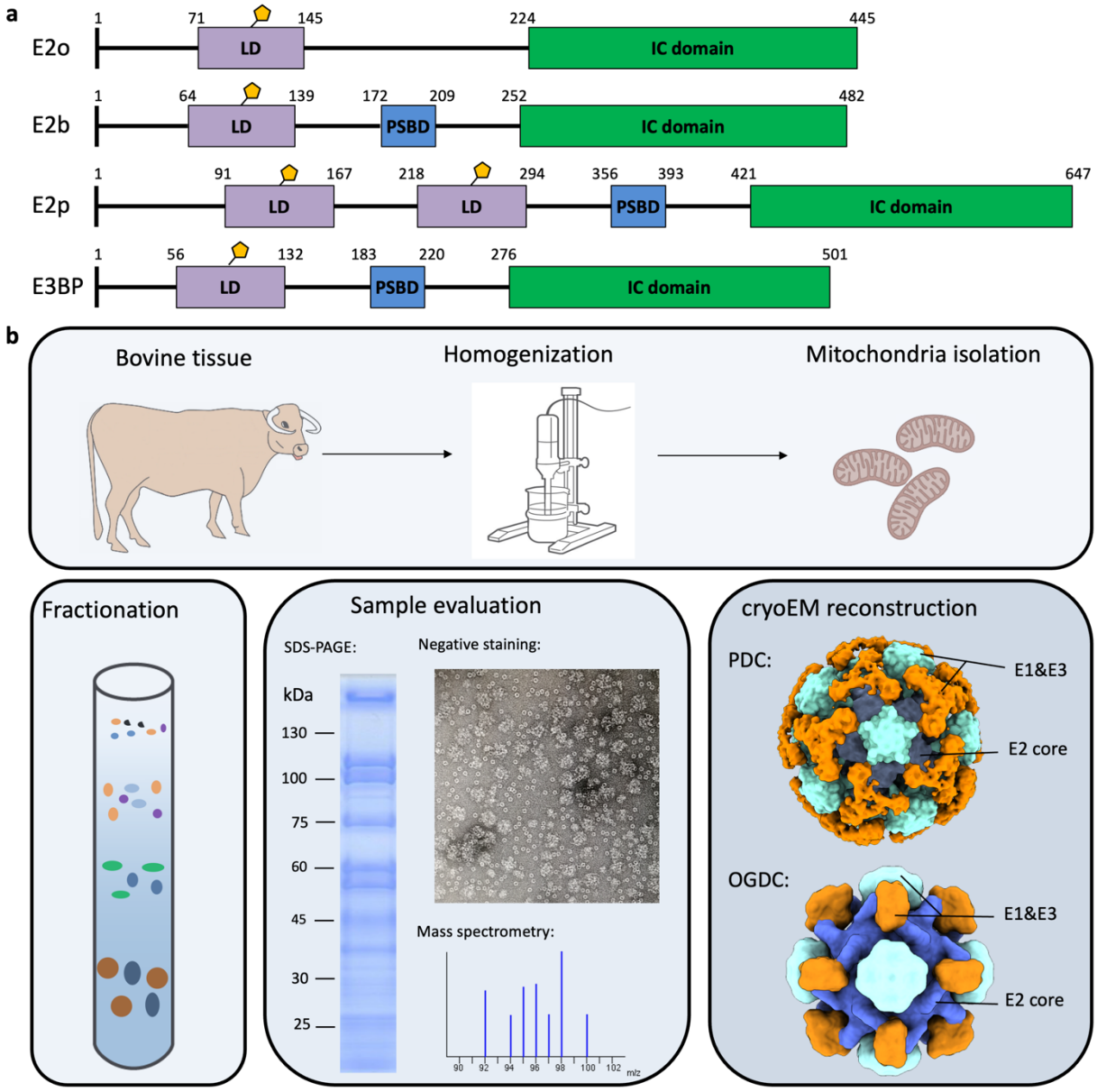


Figure 2-1. Preparation of bovine OGDC and PDC from native tissue lysate. (a) Organization of E2 and E3BP domains. **(b)** Workflow for extraction of PDC and OGDC from bovine tissue lysate. Bovine kidney was homogenized to isolate mitochondria. Mitochondrial lysates were then fractionated by sucrose gradient. Samples were evaluated for α-keto acid dehydrogenase complex by SDS-PAGE, negative stain, and mass spectrometry. CryoEM reconstruction of these samples yielded icosahedral and cubic cores surrounded by peripheral densities.

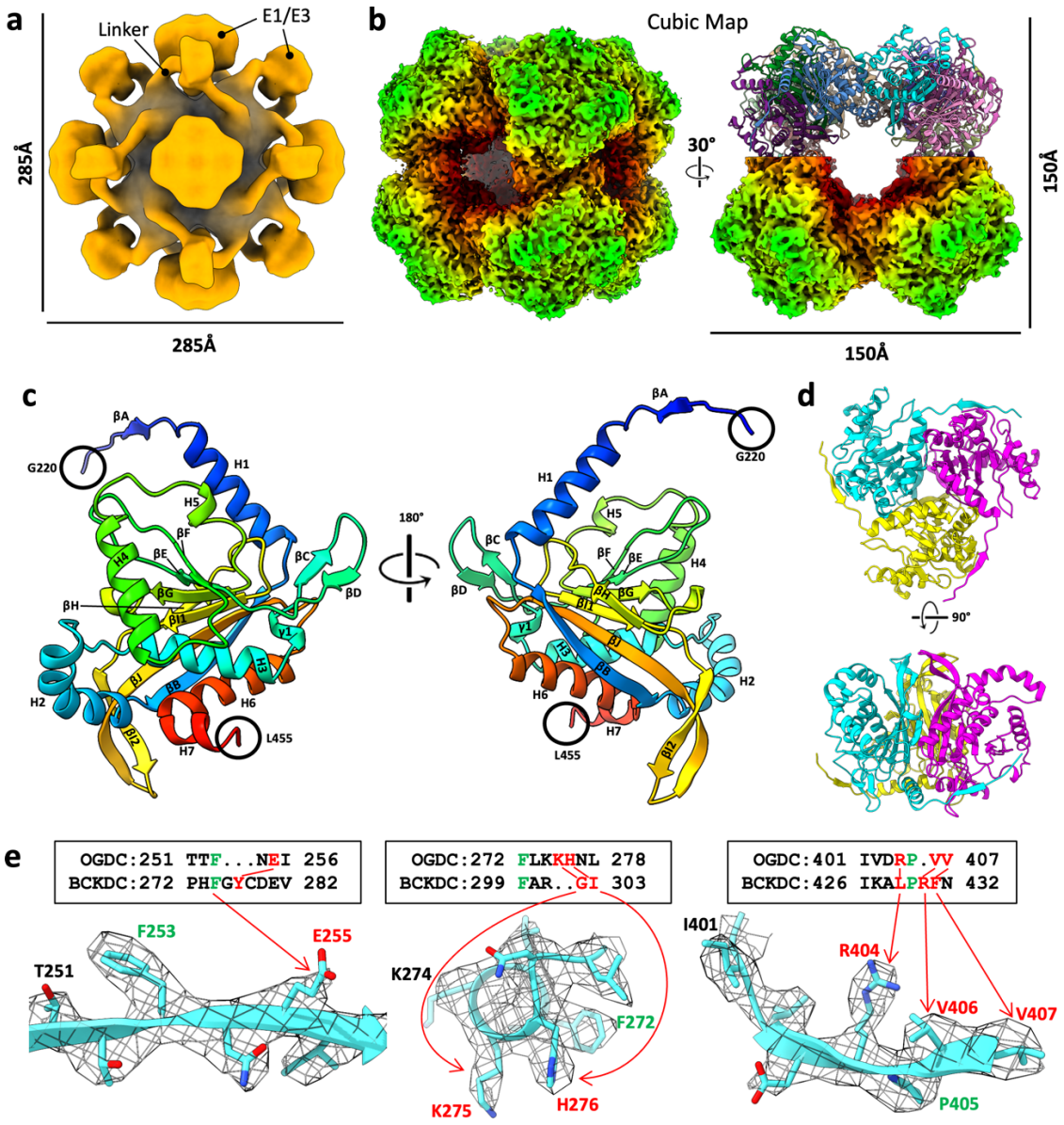


Figure 2-2. CryoEM reconstruction and identification of bovine OGDC core. (a) 3D reconstruction of the full OGDC with octahedral symmetry showing external densities (orange) around the core density (grey). Linkers extend from each E2o subunit towards a density along the edges of the core. Additional large densities are present atop the faces of the core. **(b)** Two views of the cryoEM density map of E2o core colored by radius, with half of our E2o atomic model assembly in the right. **(c)** Rainbow-colored ribbon of E2o IC domain atomic model with secondary structures labeled. Resolvable N-terminal residues (blue) and C-terminus (red) are indicated. **(d)** Structure of the E2o IC domain trimer, colored by subunits. **(e)** Sequence alignment of E2o and E2b with sidechain fit of corresponding E2o sequence into cryoEM density map. Green letters indicate conserved residues between E2o and E2b that establish landmark features in cryoEM density. Red letters and arrows indicate mismatch of residue identity in E2b with corresponding cryoEM density, thus establishing the map to be that of E2o.

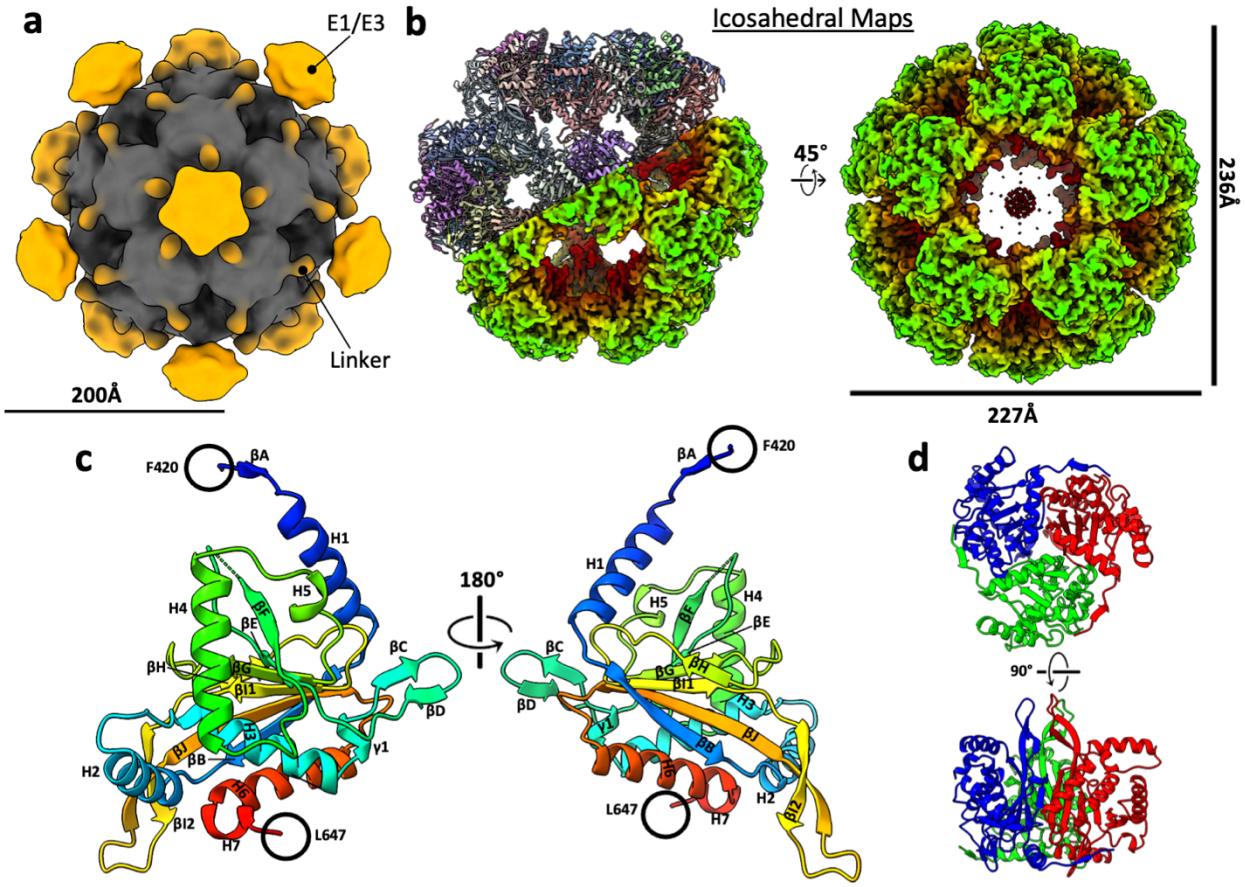


Figure 2-3. CryoEM reconstructions of bovine PDC and atomic structure of native E2p IC domain. (a) 3D reconstruction of the full PDC with icosahedral symmetry shows external densities (orange) around the core density (grey). **(b)** Two views of the cryoEM density map of E2p core colored by radius, with our E2p atomic model assembly of a hemisphere shown in the left panel. **(c)** Rainbow-colored ribbon of E2p IC domain atomic model with secondary structures labeled. Resolvable N-terminal residues (blue) and C-terminus (red) are indicated. **(d)** Two views of the atomic model of the E2p IC domain trimer, colored by subunits.

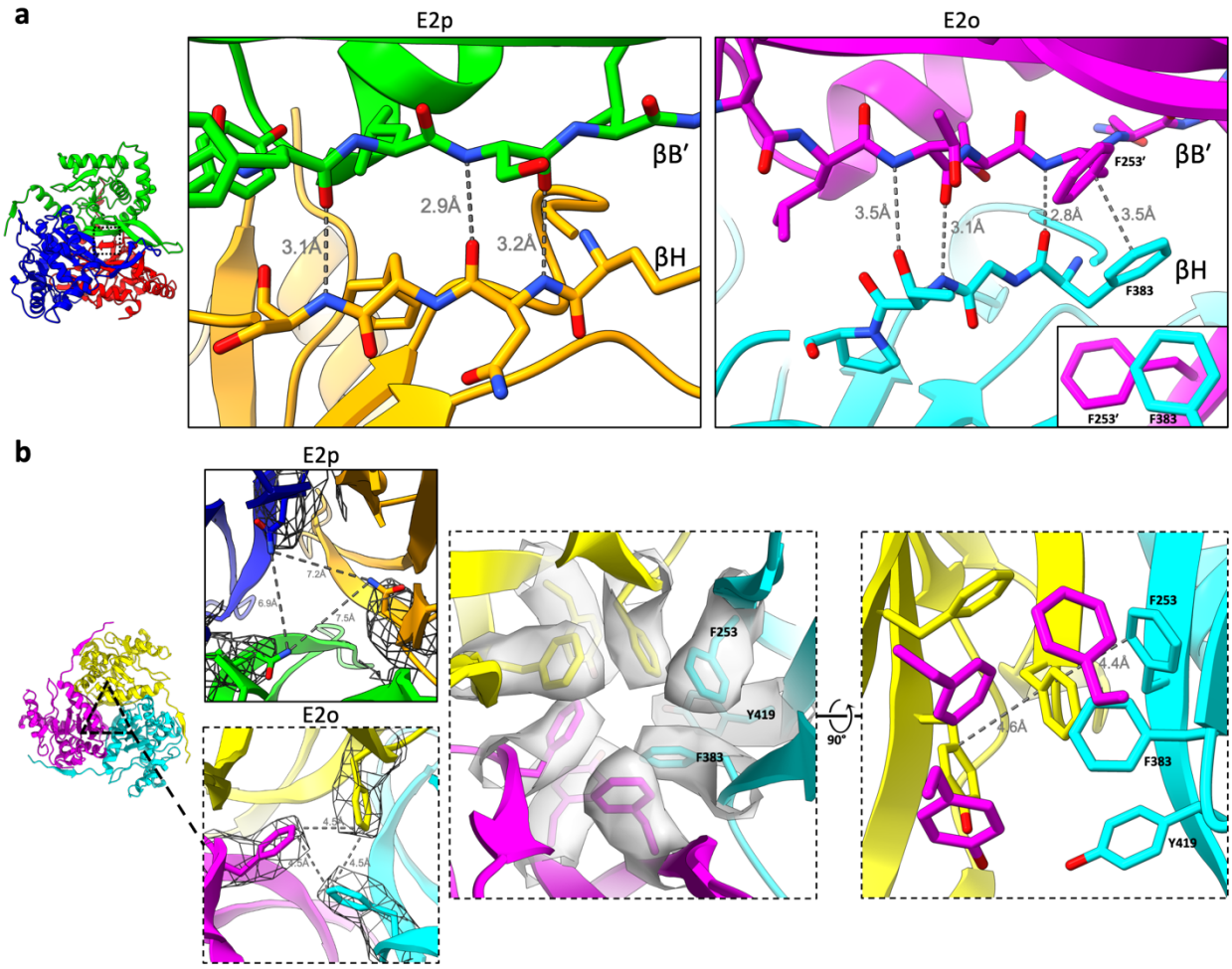


Figure 2-4. π-stacking interdomain interaction network at the three-fold interface. (a) View of the βH-βB′ interface within the interdomain trimer for E2p (left) and E2o (right). βH (orange; cyan) forms three hydrogen bonds with βB′ (green; magenta) in both E2p and E2o. The E2o interface has an additional slipped-parallel π-stacking interaction between Phe383 and Phe253′. Inset shows offset rings from view along the parallel axes. **(b)** π-stacking interactions present at the three-fold axis of E2o trimer but absent in E2p trimer. Expanded views along the three-fold axis of E2o show an extended π-electron network formed by F253, F383, and Y419.

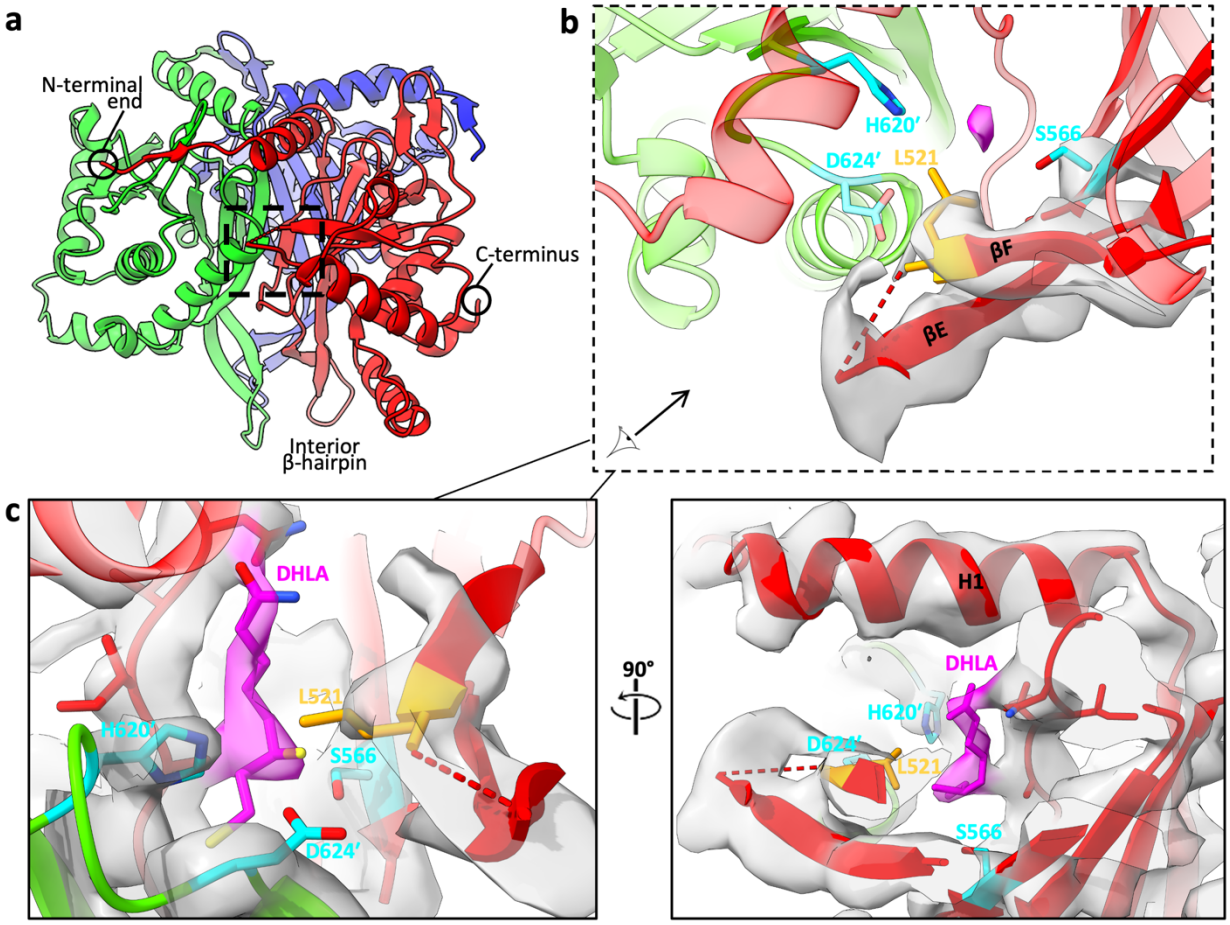


Figure 2-5. The interdomain active site within the bovine PDC E2 IC domain trimers. (a) Ribbon-model of E2p trimer (colored by subunits) with boxes highlighting the relative locations of the flexible βE-βF turn (black box). **(b)** Enlarged views of the βE-βF turn near interdomain active site relative to conserved catalytic residues (teal), Ser566 from the red subunit and His620 and Asp624 from the green subunit. Two residues (Ala519 and Gly520) are unmodeled in the turn due to absence of continuous cryoEM density (semi-transparent gray; for clarity, only those for βE-turn-βF are shown), suggesting flexibility of the residues in the turn. The conserved leucine gatekeeper (orange) is next to these unmodeled residues. A weak density (pink) bordered by catalytic and gatekeeper residues within active site is visible. **(c)** CryoEM density of the same region of (B) shown at a decreased threshold as semi-transparent pink for the lipoyl moiety or gray for the surrounding region. Atomic models for the lipoyl moiety (from PDB ID: 1EAE) and the surrounding area are shown as sticks and ribbon, respectively.

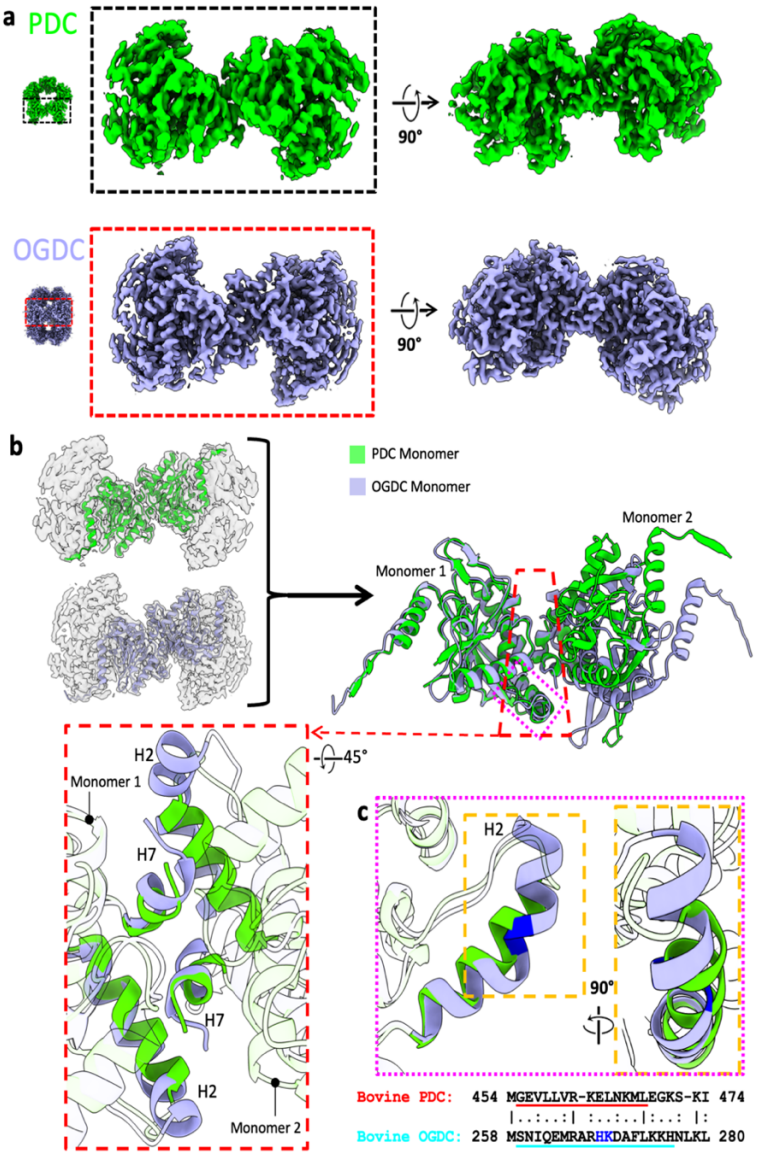
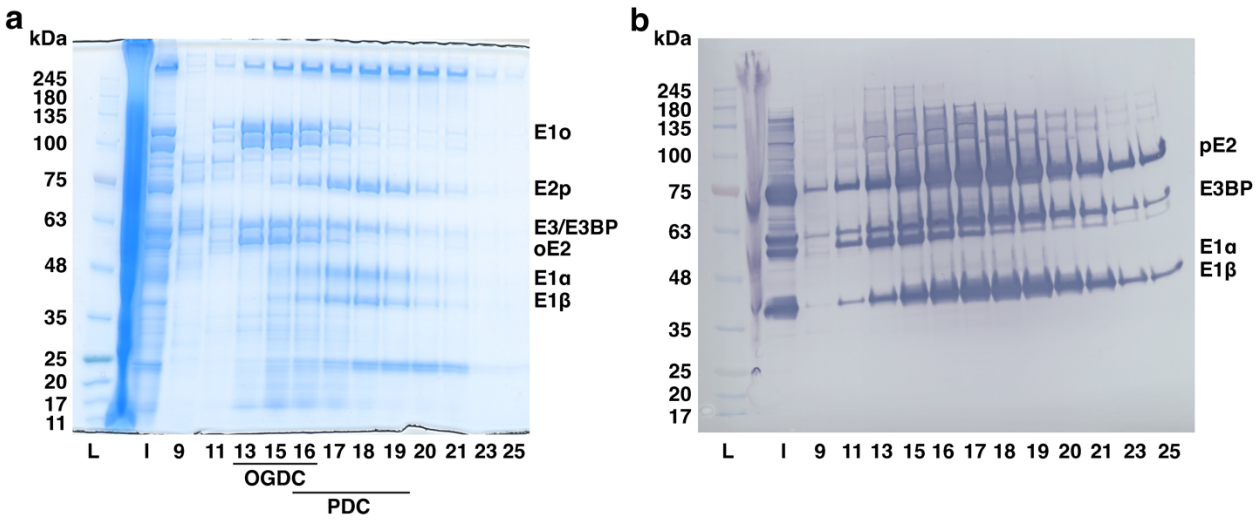
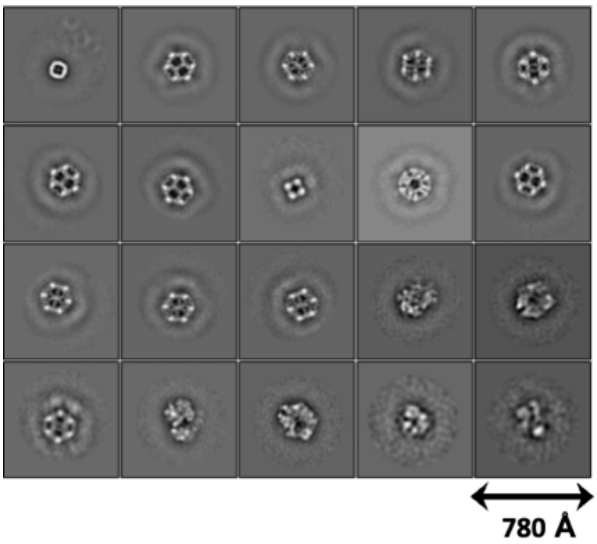


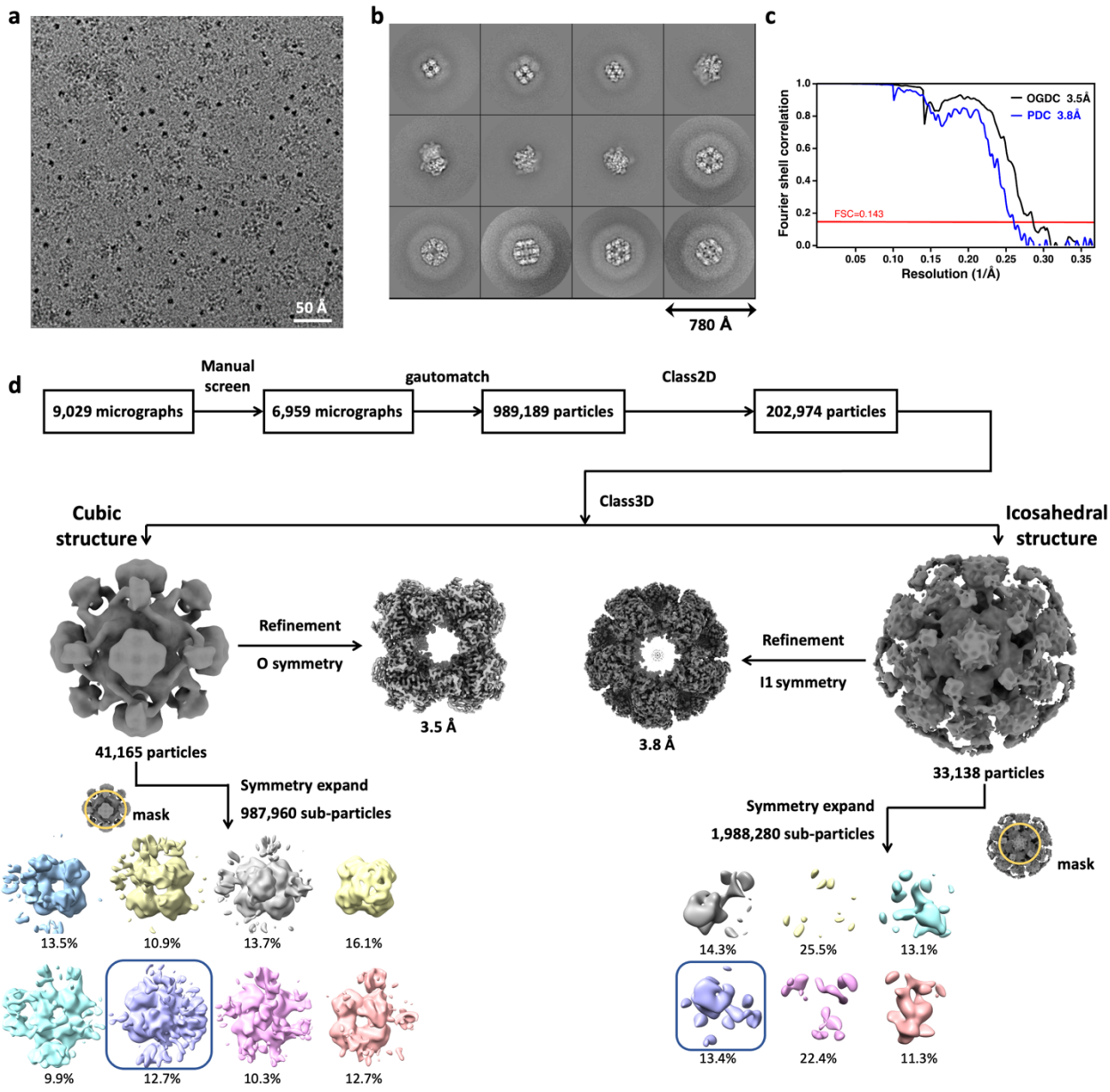
Figure 2-6. Differences in the E2 intertrimer interactions between PDC and OGDC. (a) Two orthogonal views (perpendicular and parallel to the 2-fold axis) of interacting E2 trimers from the cryoEM density maps of PDC (green) and OGDC (blue). **(b)** Comparison of the intertrimer interface in PDC (green) and OGDC (blue). Interactions involve only two subunits, one from each trimer, as apparent from the left panel where the atomic model (ribbons) of the interacting subunits are superposed in their respective densities. The two interacting subunits shown in (A) are aligned through their left subunit (monomer 1) and shown together to identify differences at the interacting interfaces with boxed regions shown as insets to highlight differences of their corresponding H2 helix. The inset in red box shows the knob-and-socket interaction of H2 and H7 (colored; rest of ribbon transparent for clarity) at two-fold interface in E2p and E2o. **(c)** Superimposition and sequence alignment of H2 in E2p and E2o. The helix-kink-helix (kink highlighted in blue) is present in E2o but absent in E2p. Inset in the orange box: Enlarged orthogonal view of the diverging H2 trajectories.



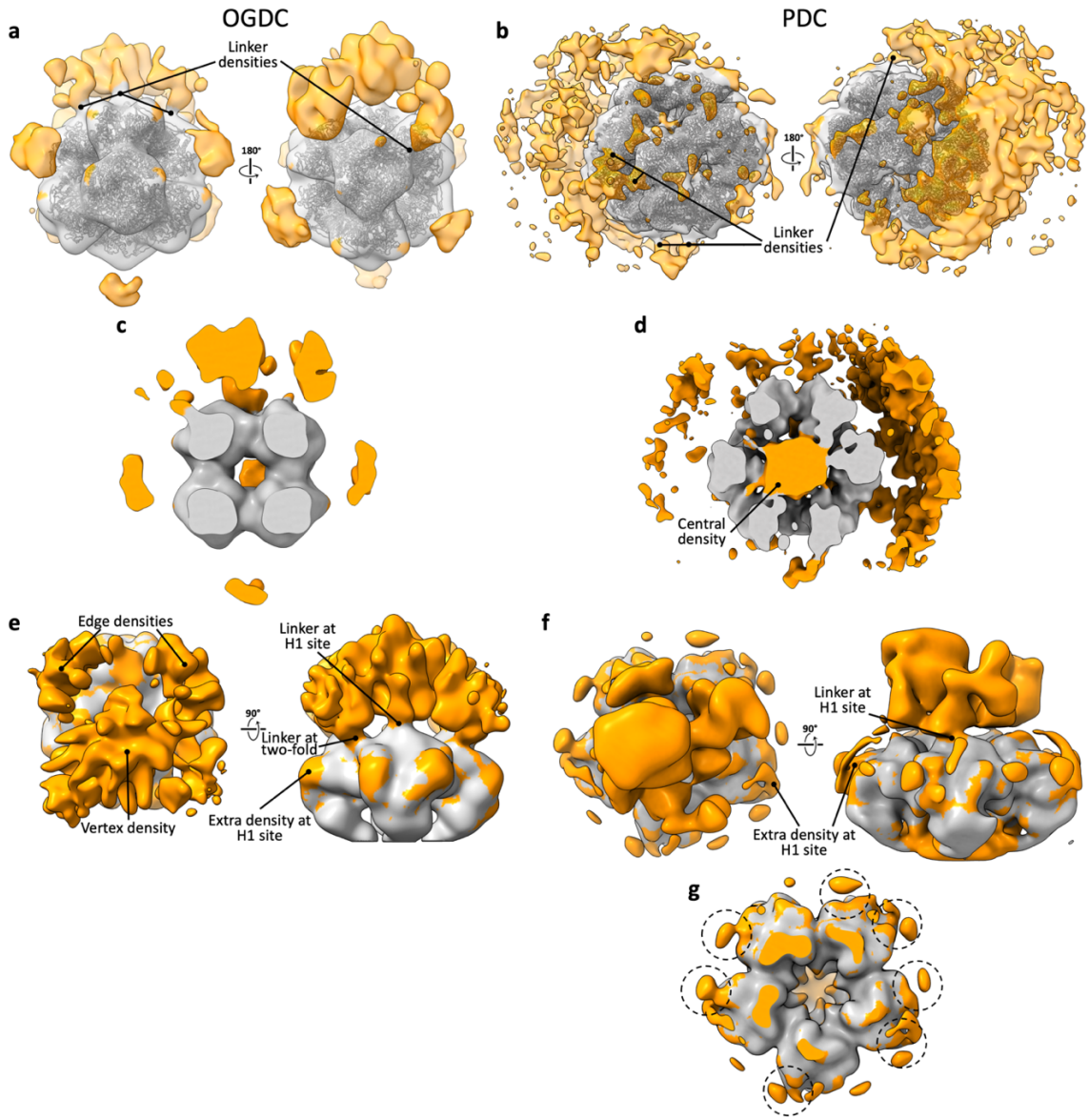
Supplementary Fig. 2-S1. Raw images of gels and blots. (a) SDS-PAGE. **(b)** Immunoblot with PDC antibody cocktail. Numbers below lanes indicate sucrose gradient fraction. Underline indicates the predominant fractions containing OGDC or PDC. Lane I shows input for sucrose gradient.



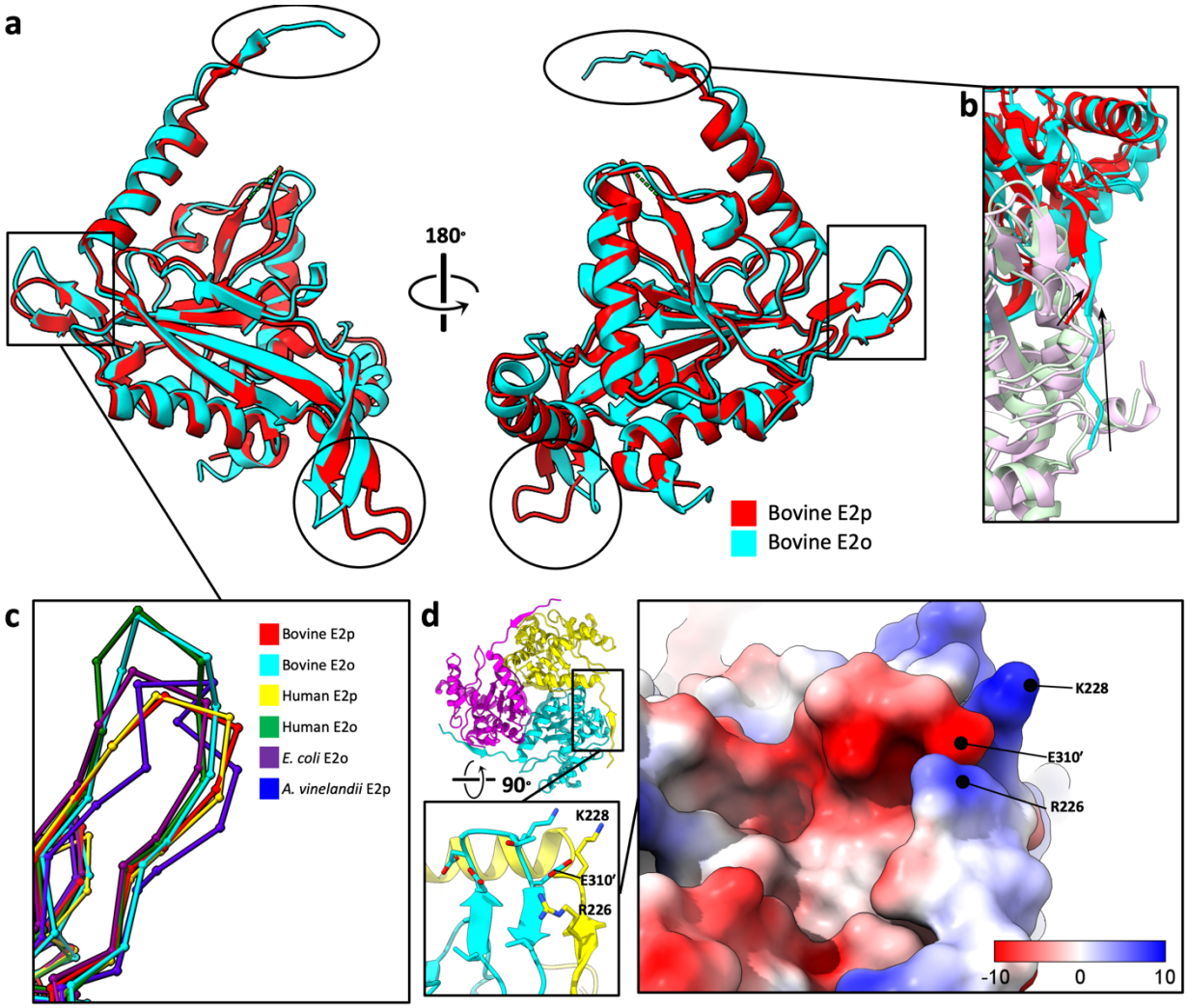
Supplementary Fig. 2-S2. Negative stain 2D class averages. Negative stain 2D class averages show cubic and icosahedral cores with peripheral density. Additional classes of unknown protein species were also observed.



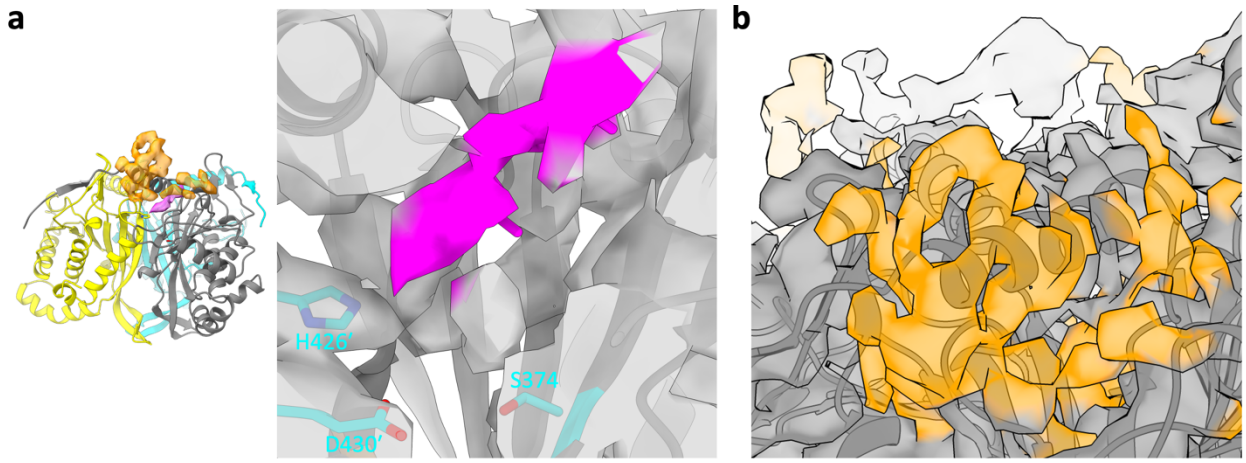
Supplementary Fig. 2-S3. CryoEM data processing workflow. (a) Representative cryoEM micrograph of lysate containing PDC and OGDC. **(b)** CryoEM 2D class averages depicting icosahedral and cubic cores with smeared peripheral density. Classes of additional unidentified protein were also obtained. **(c)** FSC of E2p and E2o core. **(d)** Data processing workflow for reconstruction of E2o and E2p cores and sub-particles.



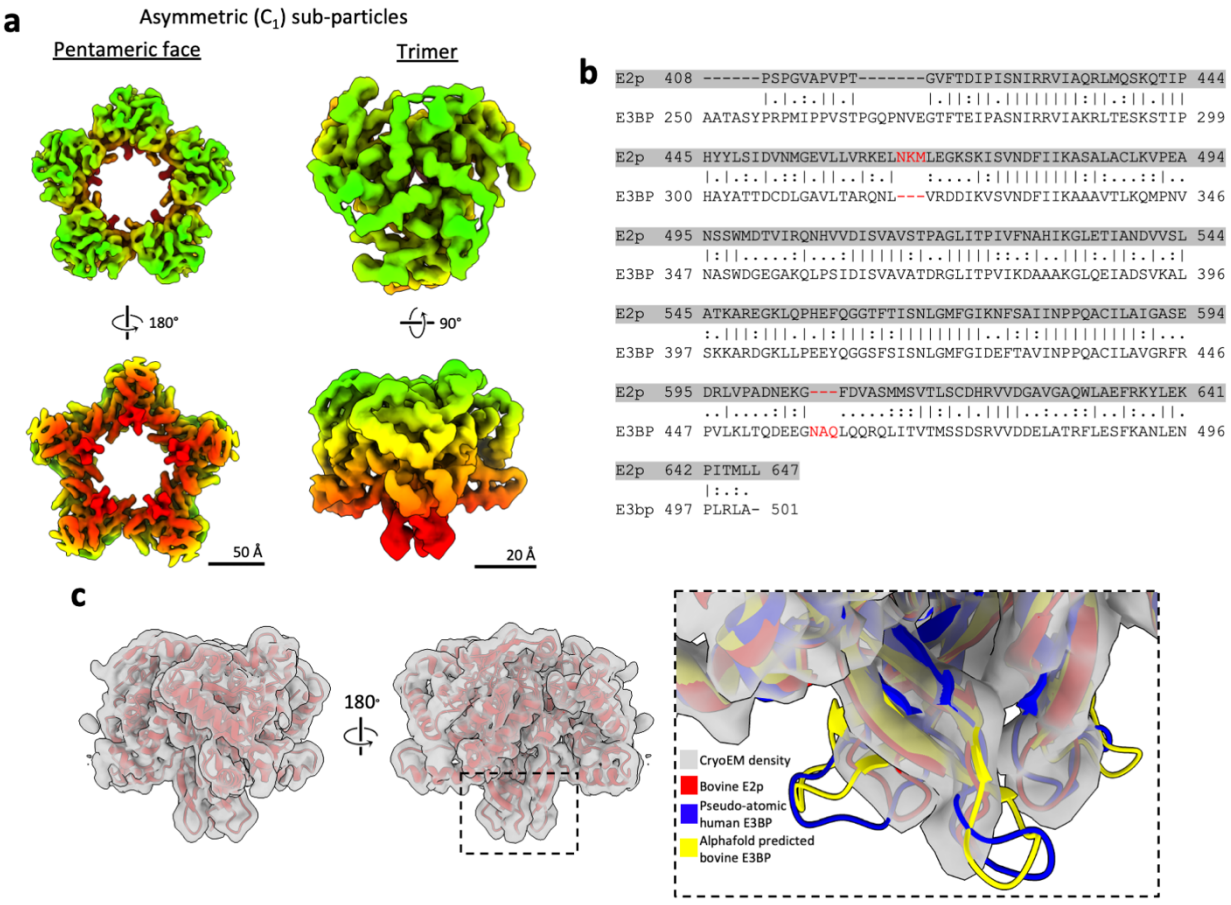
Supplementary Fig. 2-S4. Asymmetric (C_1) whole complex and sub-particle reconstructions of OGDC and PDC. (a) Two views of C_1 reconstruction of whole OGDC low-pass filtered to 20 Å and with density external to fitted E2o model (grey) in orange. **(b)** Two views of C_1 reconstruction of whole PDC low-pass filtered to 20 Å and with density external to fitted E2p model (grey) in orange. **(c)** Sliced view through center of C_1 reconstruction of whole OGDC. **(d)** Sliced view through center of C_1 reconstruction of whole PDC. Central density within E2p/E3BP core is indicated. **(e)** Top (left) and side (right) views of sub-particle reconstruction with C_1 symmetry of cubic face of OGDC. **(f)** Top (left) and side (right) views of sub-particle reconstruction with C_1 symmetry of pentameric face of PDC. **(g)** Low-pass filtered (20 Å) OGDC sub-particle map with density external to fitted E2o model in orange. **(h)** Low-pass filtered (20 Å) PDC sub-particle map with density external to fitted E2p model in orange. **(i)** Sliced view of PDC sub-particle at H1 sites. Additional densities at H1 sites are circled.



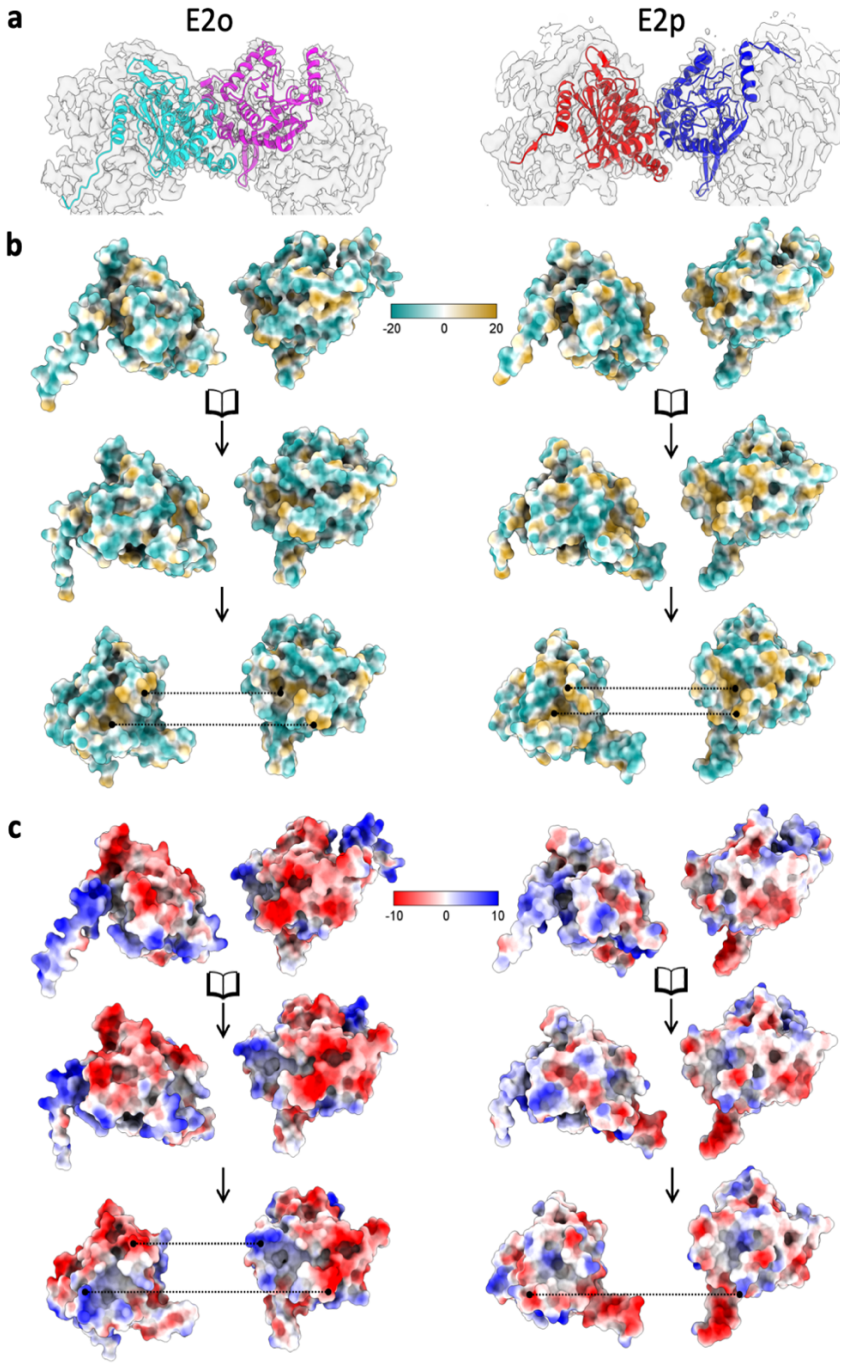
Supplementary Fig. 2-S5. Structure comparison between PDC and OGDC E2 IC domains. **(a)** Superimposition of E2p and E2o IC domains with the N-terminal linker, βC-βD turn, and interior βI2-βJ hairpin indicated by an oval, box, and circle, respectively. E2p has a longer interior hairpin than that of E2o (circle). N-terminal tail diverges at different angles between E2o and E2p (oval). E2o has a more extended βC-βD turn than that of E2p (box). **(b)** Rotated view showing N-terminal tails of E2o and E2p relative to their clockwise neighbor. Arrows indicate trajectories of N-terminal tails towards βA. **(c)** Comparison of βC-βD turn from E2o and E2p of various organisms. **(d)** Location of N-terminal linker and βC-βD turn contact in the E2o trimer. Lower inset shows residue interactions between R226 and K228 of N-terminal linker from the yellow subunit with E310′ of βC-βD turn from the cyan subunit. Right inset: electrostatic surface potential of the same area as that in the lower inset. The positively charged N-terminal linker (primarily blue) associates with the negatively charged βC-βD turn (primarily red).



Supplementary Fig. 2-S6. Lipoyl moiety and tentative lipoyl domain density in native E2o. **(a)** (left) Additional cryoEM density at active site (magenta) and near cleft between H1 and turn connecting βH and βI1 (orange). (right) Enlarged view of weak, thin density within active site channel of E2o corresponding to putative lipoyl moiety (magenta). Catalytic residues (teal) indicated for reference. **(b)** Extra, cylindrical density (orange) located near the previously proposed LD binding site.



Supplementary Fig. 2-S7. Search for E3BP in asymmetric sub-particle reconstructions. (a) Representative pentameric face (left) and trimer (right) sub-particle reconstructions with C_1 symmetry. **(b)** Sequence alignment between bovine E2p and E3BP. Three-residue differences are indicated in red. **(c)** Atomic model of E2p IC domain fit into trimer sub-particle reconstruction with C_1 symmetry. The βI2-βJ interior hairpin in E3BP is extended compared to that of E2p and is a key feature for distinguishing the two components. All cryoEM densities of interior hairpins from each trimer class appear identical and fit E2p closely. Enlarged view of the boxed area depicts atomic models fit into cryoEM density of interior hairpin. The extended hairpins of predicted human and bovine E3BP models (PDB ID: 6H60; AF-P22439-F1) are not accommodated by the cryoEM density.



Supplementary Fig. 2-S8. Electrostatic surface potential and hydrophobicity of E2o and E2p two-fold related intertrimer interface. (a) Two E2o (left) and E2p (right) IC domains at two-fold related intertrimer interface superposed in their respective density (transparent grey) and colored by chain. **(b**) Open-book view of hydrophobicity at intertrimer interface. A hydrophobic knob on each subunit binds to a complementary hydrophobic pocket on the partner subunit across the interface. Dotted lines indicate contacting surfaces. **(c)** Open-book view of electrostatic surface potential at intertrimer interface. Oppositely charged surfaces on the exterior mediate electrostatic interactions between the two subunits at the intertrimer interface.

Supplementary Table 2-S1. CryoEM data collection, refinement, and validation statistics

	#1 OGDC (EMDB-26649) (PDB 7UOL)	#2 PDC (EMDB-26650) (PDB 7UOM)
Data collection and processing		
Magnification	105,000 ×	105,000 ×
Voltage (kV)	300	300
Electron exposure (e^{-}/Å^2)	45	45
Defocus range (μm)	-1.8 to -2.6	-1.8 to -2.6
Pixel size (Å)	1.36	1.36
Symmetry imposed	O	I1
Initial particle images (no.)	989,189	989,189
Final particle images (no.)	41,165	33,138
Map resolution (Å)	3.5	3.8
FSC threshold	0.143	0.143
Map sharpening B factor (Å^2)	-152	-150
Refinement		
Model-to-map fit, mask CC	0.88	0.87
Model composition		
Non-hydrogen atoms	44496	104640
Protein residues	5664	13560
Ligands	0	0
Mean Isotropic *B* factors (Å^2)		
Protein	46.08	78.36
Ligand	N/A	N/A
R.m.s. deviations		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.944	0.940
Validation		
MolProbity score	1.13	1.39
Clashscore	3.05	7.12
Rotamer outliers (%)	0.5	0.09
Ramachandran plot		
Favored (%)	97.86	98.54
Allowed (%)	2.14	1.46
Outliers (%)	0.00	0.00

Supplementary Table 2-S2. Interactions at the intertrimer interface

Interactions	Cubic E2o core	Icosahedral E2p core
Hydrogen bonds (donor-acceptor)		
H2 and H7 interaction		
	Arg267-Asp447	(Lys641-Glu463)
	His268-Asp447	
	Lys275-Asp454	
	(Lys275-Asp455)	
	His276-Asp454	
	Arg449-Asn260	
	Arg449-Glu263 [×3]	
H4 and H7 interaction		
	Tyr342-Leu452	Leu534-Leu646
		Glu535-Leu647
Hydrophobic residues composing knob		
	Leu452	Ile643
	Leu453	Leu646
	Leu455	
Hydrophobic residues lining socket		
	Met264	Val460
	Phe272	Leu464
	Leu280	Phe479
	Tyr342	Leu534
	Pro448	Pro642
	Leu451	Met645

Duplicate two-fold interactions not listed. Parentheses indicate interface-variable interaction. Equivalent interactions indicated by matching colors. Arg449-Glu263 share three hydrogen bonds.

Chapter 3

Atomic Structures of the Kaposi's sarcoma-associated herpesvirus gB in post-fusion conformation

Fumiaki Ito[a,b,*], James Zhen[a,b,c,*], Guodong Xie[a,b], Haigen Huang[d], Juan C. Silva[a,b,c], Ting-Ting Wu[d], Z. Hong Zhou[a,b,c,#]

[a]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA, USA

[b]California NanoSystems Institute, UCLA, Los Angeles, CA, USA

[c]Molecular Biology Institute, UCLA, Los Angeles, CA, USA

[d]Department of Molecular and Medical Pharmacology, David Geffen School of Medicine, UCLA, Los Angeles, CA, USA

*Contributed equally

#Address correspondence to Z. Hong Zhou, hong.zhou@ucla.edu.

3.1 Abstract

Discovered in 1994 in lesions of an AIDS patient, Kaposi's sarcoma-associated herpesvirus (KSHV) is a member of the gammaherpesvirus subfamily of the *Herpesviridae* family, which contains a total of nine that infect humans. These viruses all contain a large envelope glycoprotein, glycoprotein B (gB), that is required for viral fusion with host cell membrane to initial infection. Although the atomic structures of five other human herpesviruses in their postfusion conformation and one in its prefusion conformation are known, the atomic structure of KSHV gB has not been reported. Here, we report the first atomic structure of the KSHV gB ectodomain determined by single-particle cryogenic electron microscopy (cryoEM). Despite a similar global fold between herpesvirus gB, KSHV gB possesses local differences not shared by its relatives in other herpesviruses. The glycosylation sites of gB are arranged in belts down the symmetry axis with distinct localization compared to that of other herpesviruses, which occludes certain antibody binding sites. An extended glycan chain in domain I (DI) suggests a means for host cell attachment. Local flexibility of domain IV (DIV) governed by molecular hinges at its interdomain junctions identifies a means for enabling conformational change. A mutation in the domain III (DIII) central helix disrupts incorporation of gB into KSHV virions despite adoption of a canonical fold *in vitro*. Taken together, this study reveals mechanisms of structural variability of herpesvirus fusion protein gB and informs its folding and immunogenicity.

3.2 Importance

In 1994, a cancer-causing virus was discovered in lesions of AIDS patients, which was later named Kaposi's sarcoma-associated herpesvirus (KSHV). As the latest discovered human

herpesvirus, KSHV has been classified into the gammaherpesvirus subfamily of the *Herpesviridae*. In this study, we have expressed KSHV gB and employed cryogenic electron microscopy (cryoEM) to determine its first atomic structure. Importantly, our structure resolves some glycans beyond the first sugar moiety. These glycans are arranged in a pattern unique to KSHV, which impacts the antigenicity of KSHV gB. Our structure also reveals conformational flexibility caused by molecular hinges between domains that provide clues into the mechanism behind the drastic change between prefusion and postfusion states.

3.3 Introduction

Kaposi's sarcoma-associated herpesvirus (KSHV) is the most recent human herpesvirus, discovered in the lesions of acquired immunodeficiency syndrome (AIDS) patients (1). KSHV is a member of the gammaherpesvirus subfamily, and, like its fellow gammaherpesvirus Epstein-Barr virus (EBV), is the causative agent of several devastating diseases including cancers, such as Kaposi's sarcoma, primary effusion lymphoma, and multicentric Castleman's disease (2, 3). In particular, KSHV poses a cancer burden on immunosuppressed individuals, such as AIDS patients and organ transplant recipients (4). However, there are currently no vaccines or treatments against KSHV infections to prevent its associated malignancies (5, 6).

Like all herpesviruses, KSHV is composed of a dsDNA genome within a capsid that is further surrounded by a tegument layer and enclosed inside of an envelope decorated with glycoproteins. These glycoproteins have critical roles in mediating membrane fusion during both the maturation and infection pathways (7–9). After nuclear egress of the capsid into the cytosol, the tegumented capsid acquires two sets of envelopes during the secondary envelopment process. The outer membrane layer fuses with the host membrane from the cytosolic side to release the

now single-enveloped mature virion (9). During host entry, the viral envelope attaches to and fuses with the host membrane to allow the genome-containing capsid to enter (10). Because of the critical nature of these membrane fusion events for the life cycle of KSHV, its envelope glycoproteins responsible for membrane fusion are apt targets for the development of vaccines and treatments.

Glycoprotein B (gB) is conserved across all herpesviruses and is the fusogen responsible for the physical action of membrane fusion. As a class III fusion protein (11, 12), gB undergoes a conformational change from a prefusion state to a postfusion state that draws the host and viral envelopes together to induce hemifusion. Structural determination of gB in the prefusion and postfusion conformations would provide insights into the herpesvirus fusion mechanism and epitopes. Structures of the postfusion conformation have been determined for herpes simplex virus type 1 (HSV-1), EBV, human cytomegalovirus (HCMV), and varicella-zoster virus (VZV) by crystallography (13–20). However, the structure of gB in the prefusion state has been elusive due to its metastability. Cryogenic electron tomography (cryoET) has been used to observe gB in the context of intact membranes and identify its prefusion conformation (21–25). Recently, structures of HCMV gB in the prefusion conformation have been determined at near-atomic resolution by single particle cryogenic electron microscopy (cryoEM) (26, 27).

Despite these advancements in the structural determination of gB, the structure of KSHV gB in neither post nor prefusion have been determined. In addition to a lack of KSHV gB structures, the effects of glycosylation on KSHV gB are not well understood. Glycans shield the antigen from binding by host antibodies (28, 29). Each herpesvirus gB possesses different degrees of glycosylation at varying sites, which should in turn impact the availability of sites for antibody binding. Structures of KSHV gB in tandem with the gB structures of its fellow

herpesviruses would enable a more holistic understanding of their fusion mechanism and immunogenicity.

In this study, we determined the structure of KSHV gB in the postfusion conformation by cryoEM, including glycans beyond the first sugar moiety. The glycosylation sites of KSHV gB alternate between belts of glycan-rich and glycan-poor surfaces around the ectodomain, which dictate the availability of epitope sites. KSHV gB contains local differences from the conserved folds found in gB of other herpesviruses. The domain IV (DIV) crown of KSHV gB experiences rotational flexibility governed by hinges at its interdomain junctions. A mutation in the DIII coiled coil prevents incorporation of gB into KSHV virions.

3.4 Results

3.4.1 Overall structure of KSHV gB

To understand the mechanisms of membrane fusion and cell entry mediated by KSHV, we introduced mutations in gB in an attempt to enrich for its prefusion state on otherwise native KSHV virions. Based on sequence alignment with HSV-1 and VZV gB (22, 23), we introduced the D470P or G471P into the KSHV virion and were able to isolate mature virions, though G471P had a lower yield than D470P mutants (Fig. 3-3-1A, B). Through immunoblot analysis, gB D470P is not detected in the supernatant from which secreted virions were isolated but is detected in the cellular lysates (Fig. 3-1C), indicating that the gB mutants are being expressed. On closer inspection through cryoET, KSHV gB D470P mutant virions do not have distinct densities of postfusion gB like those on KSHV WT virions (Fig. 3-1B).

To gain more insights into structural basis of these gB mutations, we expressed and purified KSHV gB containing these mutations for structural determination by cryoEM. We initially tried to use full-length constructs but encountered problems with sample preparation of this membrane protein. To alleviate this issue, we truncated the sequence and determined the structure of the soluble ectodomain of KSHV gB (amino acid residues 1-687). The ectodomain construct includes five subdomains DI-DV but excludes the C-terminal membrane proximal region (MPR), transmembrane domain (TM), and cytoplasmic domain (Fig. 3-2A). Of the homologous mutation sites at the DIII central helix (22, 23), D470P was selected for structural determination for its expression level and charged sidechain (Fig. 3-S1). The gB ectodomain was expressed in and purified from mammalian cells in its fully glycosylated form.

The structure of KSHV gB was determined at 3.3 Å resolution by single particle cryoEM (Fig. 3-S2, S3). The gB structure was reconstructed as a homotrimer with C3 symmetry aligned along the long edge. KSHV gB ectodomain exhibits an elongated rod-shaped architecture with dimensions approximately 70 Å by 70 Å by 170 Å, consistent with the postfusion conformation observed in other herpesvirus fusion proteins, such as those from HSV-1, VZV, HCMV, and EBV (13–15, 20) (Fig. 3-2B). We did not observe any 2D or 3D classes representing potential prefusion form (Fig. 3-S2), which is predicted to have a more globular dimension, highlighting the metastable nature of the prefusion conformation. Additionally, no asymmetric assembly was observed from 3D refinement without imposing symmetry, and further 3D classification did not yield any asymmetric states, indicating minimal conformational variability among the three individual protomers.

The cryoEM map resolved the densities for residues 54-677 consisting of the five core domains (Fig. 3-2C). The KSHV gB follows a canonical postfusion domain organization of the

class III fusion protein, with its core comprising three β-strand-rich modular domains (DI, DII, and DIV) and a long three-helix bundle of DIII. DI forms a trimer at the membrane-proximal end of the ectodomain with its hydrophobic fusion loops facing the membrane side. DII is juxtaposed to DI through two linkers and encircles the N-terminal end of the DIII central helical bundle. DIII bridges DII and DIV via the central helices and a DII-DIV hanging linker. DIV trimer is positioned at the membrane-distal end of the ectodomain. Finally, DV folds back and extends from DIV to DI along the DIII central helices, with its two C-terminal helices intercalating into the DI trimer interface. The C-terminal end of the DV exits between the DI-DI interface potentially allowing the subsequent MPR/TM region to form a trimeric bundle atop the DI trimer (Fig. 3-2A).

Other than these resolved regions, the first 53 residues at the N-terminus, containing signal peptide (residues 1-23) followed by a region rich in Pro, Ser, and Thr residues were invisible, indicating that this region is disordered. The residues 409-447 that link DII and DIII, including the furin cleavage site (^{437}RKRR/S^{441}) were also invisible. This region is highly variable in both sequence and length among herpesvirus gB proteins. A 14-residue patch area preceding the furin cleavage site (^{419}PTSSPPPSASPMTTS^{433}) that is also rich in Pro, Ser, and Thr could potentially be modified with O-linked glycosylation as previously shown for HSV-1 gB (30, 31). The flexibility and proteolytic cleavage in this segment are likely important for the pre- to postfusion conformational transition as the spatial orientation between DII and DIII is inverted by nearly 180 degrees during the transition (26). The five conserved disulfide bonds were readily observed including four intra-domain bonds in DI (C157-C222), DIII (C85-C484), and DIV (C68-C528 and C550-C587), and one that crosslinks between the DI-DII linker and DII (C315-C362), providing rigidity to the DI-DII junction (Fig. 3-S4).

The cryoEM map of the KSHV gB displays amino acid side chain densities for most regions (Fig. 3-2D and Fig. 3-S5). The density of the DIII central helix is shown in Fig. 3-2D. The two fusion loops in DI that are located at the edge of the ectodomain were previously missing in the cryoEM structure of closely related EBV gB (PDB ID: 7FBI) (32). Our map has well-resolved density for these regions, which allowed us to build the corresponding atomic models. Additionally, a continuous density for the DII-DIV hanging linker region of DIII, previously unresolved in EBV gB, was observed with a moderate resolution that permitted model building (Fig. 3-2D).

3.4.2 Unique glycosylation pattern of KSHV gB

Herpesvirus gB proteins are extensively glycosylated to promote the evasion of neutralizing antibody recognition and host cell attachment (33). Of the 13 hypothetical glycosylation sites based on consensus sequence in KSHV gB, we observed a total of ten N-linked glycans per protomer distributed across domains DI, DII, DIV, and DV at asparagine residues 179, 254, 275, 355, 368, 372, 385, 599, 614, and 628 (Fig. 3-3A). Each glycosylation site conforms to the canonical N-linked glycosylation motif of Asn-X-(Ser/Thr) where X represents any amino acid except proline. We could not observe glycosylation at the hypothetical sites at residues 408, 455, and 562. Asn408 and 562 are located in regions of relatively weak cryoEM density near the unmodeled DII-DIII junction and DIV, respectively. Asn455, located at the N-terminal base of the DIII central helix, is engaged in a delocalized π-electron network that stabilizes the coiled coil and sterically occludes glycosylation. No O-linked glycosylation was observed in the resolved region of the cryoEM density map. Notably, the glycosylation sites on the gB surface are predominantly clustered along a belt that extends along the symmetry axis of the gB trimer from DI to DIV (Fig. 3-3B). This surface glycosylation distribution pattern creates glycan-rich

and glycan-poor surface areas that alternate along the symmetry axis and potentially influences the antigenicity of the gB surface.

Most glycan chains displayed discernible density only for the first one or two sugar moieties, specifically N-acetylglucosamine (GlcNAc), due to the intrinsic flexibility of the sugar chains. Notably, the glycan at N179 showed an extended sugar chain density, corresponding to up to seven sugar molecules including two branching points (Fig. 3-3C). At this site, the initial two GlcNAcs form stacking interactions with the aromatic side chains of Phe158 and Tyr146, respectively. The second GlcNAc forms a hydrogen bond through the oxygen atom of its N-acetyl group with the nitrogen atom of the side chain imidazole ring of His150, thereby stabilizing the glycan chain. The observed branching pattern and sugar chain architecture indicate that this glycan chain belongs to the oligomannose type among the potential mammalian glycans including hybrid- and complex-type. The glycosylation site N179 is conserved in beta- and gammaherpesviruses and is located near the fusion loops that are proximal to the membrane, potentially indicating that this glycan may play a role in host cell attachment.

Based on sequence alignment, the glycosylation sites on KSHV gB are either unique to KSHV or only partially conserved among herpesviruses (Fig. 3-3A, S6, S7). For example, N368 is conserved only in EBV, the other gammaherpesvirus. N254 and N275 are conserved in betaherpesvirus HCMV, but not in EBV. N628 is the only glycosylation site that is conserved between alpha- and gammaherpesviruses. This unique glycosylation pattern of KSHV gB would cause steric conflicts with experimentally determined epitope sites of other human herpesviruses (15, 20, 26, 32, 34–36) (Fig. 3-4), which suggests that antibodies against KSHV gB bind different sites from those against other herpesviruses. The regions occluded from antibody binding would be larger on a fully glycosylated fusogen (Fig. 3-S7). Taken together, the KSHV

gB structure reveals by far the most extensively glycosylated among the herpesvirus gB structures reported to date and informs distinct glycan-shielded surface areas that would influence the antigenicity of the gB.

3.4.3 Structural features of KSHV gB and comparison with other herpesvirus gB

We compared our KSHV gB atomic model with prior solved atomic models of other herpesvirus gB (13–20, 26, 32, 34–39). The protomer fold is highly conserved across all herpesviruses (Fig. 3-5A). In general, the folds of each domain are nearly identical between herpesvirus species, but we observe that a loop in DI of KSHV gB at residues 252-270 deviates greatly from all currently known gB structures (Fig. 3-5A, B). In KSHV, this loop runs along the exterior of the trimer rather than being tucked in towards the interprotomer interface as in gB of all other species (Fig. 3-5B). Our cryoEM map of KSHV gB possesses density for this path and lacks density for the conformation followed by other gB (Fig. 3-5B). Sequence alignment shows that this loop has low conservation between KSHV and other herpesviruses, including its gammaherpesvirus relative EBV (Fig. 3-5C). This divergent evolution of the KSHV gB DI loop could impact the activity and immunogenicity of this fusogen.

Although the domains within gB are similar between herpesviruses, the orientation of these domains relative to each other differ. Looking down the C3 axis from DIV when the models are aligned by DI, we observe rotational differences between the domains. In KSHV and EBV, DIV lines up with DI and DII of its clockwise neighbor in an eclipsed conformation (Fig. 3-5A). In HCMV, the sole betaherpesvirus representative with atomic models of gB, DIV is slightly rotated and DII peeks out from below. Additionally, the DIV crown of KSHV gB shares a similar assembly with those of all other herpesviruses, except for HCMV, whose DIVs are bound closer together. For the alphaherpesvirus, DIV is offset from DI and DII in a staggered

conformation. By overlaying the DIV crowns, we quantify the rotation between KSHV and EBV, HCMV, and VZV as ~4°, 13°, and 36°, respectively (Fig. 3-5D). This rotation of DIV relative to DI and DII is conserved between EM and crystal structures within the same herpesvirus species, and only slight differences are present between members within a subfamily.

To observe the contributors to the rotation of DIV, we aligned the atomic models by DIV (Fig. 3-5E). While the main fold of DIV is conserved, the adjoining DII-DIV linker, DIII, and DV project from DIV with varying trajectories depending on the subfamily. In gB of alphaherpesviruses, DIII bends as it approaches DIV, which shifts the position of the short helix and initial antiparallel β-sheet composed of the β-turn following the helix and a β-strand from the DII-DIV linker (Fig. 3-5F). Although the trajectory of DV differs between species, its binding with the DIII coiled coil suggests that this difference is a result of the DIV rotation rather than a cause. For HCMV gB, the β-strand following the short helix of DIII is shorter, which shifts the β-turn such that it does not interact with the DII-DIV linker to form an antiparallel β-sheet (Fig. 3-5G). This in turn permits DIV to turn further inwards towards the C3 axis compared to KSHV. DIV transitions from being the core of the prefusion state to a distal crown in the postfusion state (26), and the postfusion rotational differences observed between herpesvirus species may have origins in differences between the prefusion states.

3.4.4 DIII-DIV interdomain dynamics

Among the five subdomains of the KSHV gB ectodomain, the membrane-distal DIV trimer region showed a lower local resolution, ranging from 3.2 to 4.4 Å, compared to the rest of the ectodomain, which showed a resolution ranging from 2.6 to 3.4 Å (Fig. 3-S3C). One possible reason for the relatively low resolution in the DIV is the presence of conformational heterogeneity. In addition, when compared with the crystal structure of EBV gB (PDB ID:

3FVC) (14), the largest structural deviations were observed around the DIV region. To understand the conformational landscape of KSHV gB present in our cryoEM data set, we performed 3D flexibility analysis, a generative neural network method that samples continuous heterogeneity within a 3D density map (40). The analysis reveals structural variability around the DIV trimer region with respect to the remaining ectodomain (Fig. 3-6A). The DIV trimer region exhibits a rotational movement around the three-fold symmetry axis, while the remaining ectodomain (DI-DIII, DV) remains largely static (Fig. 3-6A, movie S1). This rotation does not induce major structural alterations within the DIV protomer or its trimeric assembly, suggesting that the domain junctions act as molecular hinges facilitating this movement. Specifically, the flexible hinge regions are located at the DIII-DIV and DIV-DV junctions (amino acids I500-V514 and S619-T623, respectively). Atomic models derived from the first and last frames of the 3D flexibility refinement show an R.M.S.D. of 3.1 Å around the DIV and 0.98 Å across domains I, II, III, and V (Fig. 3-6B).

During the transformation from prefusion to postfusion state, herpesvirus class III fusogens undergo global conformational transitions where the DIII central helices invert relative to the DI-DII tandem domains (Fig. 3-6C). During this transition, gB hypothetically adopts an extended intermediate state in which DI-DII regions are released from the membrane-associated MPR, positioning the hydrophobic fusion loops at the tip of DI for attachment to the host membrane. The DIII-DIV interdomain flexibility potentially provides additional room for the DI fusion loops to fine-tune their positions for insertion into an optimal site on the host membrane surface. This flexibility may also facilitate the transition from the intermediate state to the postfusion form, wherein the DV refolds and folds back along the entire ectodomain, driving the membrane-anchored MPR/TM regions to merge the viral and host membranes. Due to the

flexibility needed for these processes, the DIV region remains flexible in the postfusion form even after membrane fusion has completed (Fig. 3-6C).

3.5 Discussion

Here, we determined the structure of the KSHV gB ectodomain in the postfusion conformation. Despite the DIII hinge mutation, the recombinant ectodomain conforms to the canonical postfusion global fold (Fig. 3-2). However, KSHV gB possesses a glycosylation pattern distinct from those of other herpesviruses and includes an extended glycan near the fusion loops (Fig. 3-3). The surface of the herpesvirus gB is predominantly covered by host-derived glycans, a common feature of diverse viral membrane fusion proteins (41). Because herpesvirus gB is a crucial target in vaccine development efforts, understanding the glycosylation of viral fusogens is vital for elucidating the fundamental mechanisms of viral cell entry and for informing vaccine design strategies (42, 43)(44, 45).

The KSHV gB structure in the present study displays 30 N-linked glycans in each trimer. These sites are only partially conserved across herpesviruses, with half being specific to KSHV. Each protomer possesses three glycans in DI, four in DII, two in DIV, and one in DV, with no glycan present in DIII.

Notably, the glycan at N179, located on the membrane-proximal side of the DI, showed an extended sugar chain structure as the larger part of the glycan is stabilized through the interactions with the surface residues of the gB. This is the most extensive form of glycan observed in any herpesvirus gB structure to date, as the long glycans are typically too flexible to be resolved by X-ray crystallography or cryoEM. High-mannose type glycan could be modeled most plausibly into the cryoEM density at this site, consistent with the previous study showing

that KSHV gB is rich in oligomannose (46). Besides N-linked glycans, O-linked glycans are expected in the two unresolved regions of the ectodomain based on the mass spectrometry-based glycoproteomics in herpesvirus gB homologs (30, 47): immediately after the signal peptide at the N-terminal end and in the DII-DIII linker preceding the furin cleavage site, both regions being rich in Pro, Ser, and Thr.

The extensive glycosylation on KSHV gB leaves a limited protein surface available for antibody binding. Epitopes previously identified in herpesvirus gB homologs may not be compatible with KSHV gB (Fig. 3-4). For example, the epitope region in DII observed in EBV gB would be inaccessible for antibodies in KSHV gB due to a KSHV-specific glycan at N372 (32). On the other hand, epitopes in DIV of EBV gB and DI of HCMV gB could be antigenic in KSHV gB as well (15, 32). A fully glycosylated fusogen is expected to occlude a larger surface and further limit conserved epitopes (Fig. 3-S7A). Precise mapping of antigenic surfaces on KSHV gB, however, awaits further epitope mapping study.

The DIII hinge mutation in our KSHV gB construct was added based on a previous cryoET study of an HSV-1 gB hinge mutant present on vesicles in the prefusion state (22). The prefusion state induced by this hinge mutation has been recapitulated in VZV gB expressed on vesicles (23). Notably, the HSV-1 gB hinge mutant ectodomain crystallized in the postfusion conformation, suggesting that the membrane is important for maintaining the prefusion state (22). Although we determined the structure of our KSHV gB hinge mutant ectodomain using cryoEM, which is not subject to conformational changes induced by crystal packing, we also observed the postfusion conformation, which further supports the critical function of the membrane for stabilizing the prefusion state. Indeed, even crosslinked HCMV gB with fusion inhibitor from the native virion has been found to mostly occupy the postfusion state after

purification from the membrane (26). Nonetheless, a full-length gB structures spanning from prefusion to postfusion in context of the membrane would elucidate how herpesviruses undergo membrane fusion.

Like our mutant KSHV virions, mature virions have also been observed for VZV virions containing the gB E526P or V528P DIII hinge mutation attached to the cell surface after budding (23). The recovery and observation of mature virions suggests that the virus can successfully undergo secondary envelopment and egress across the host membrane. Completion of these steps in the viral life cycle implies that fusion can still occur despite the hinge mutation. However, gB knock out HSV-1 viruses are still able to be enveloped and secreted (7), indicating that the secondary envelopment and budding steps can still occur independent of gB. Indeed, the VZV gB V528P mutant virions in particular failed to replicate due to disrupted gB biosynthesis (23), which may also be the cause of gB deficiency in our mutant virions. Our structure of KSHV gB D470P mutant still conforms to the canonical postfusion gB ectodomain (Fig. 3-5A), which suggests that the disruption in gB biosynthesis is not due to misfolding of the protein. But to truly understand the fundamental cause of this disruption, the structure of gB and its interacting partners would need to be determined in the cellular context. Moreover, due to the lack of gB on the mutant KSHV, this system cannot be used to enrich the presence of prefusion gB for the identification of prefusion gB localization and interactions. Because of the low presence of prefusion gB on WT KSHV virions and low contrast of proteins within cells, advances in structural biology methods are required to fully characterize the mechanism of herpesvirus fusion in action, from prefusion to postfusion and the intermediates in between.

3.6 Materials and Methods

3.6.1 Plasmids

ORF8 (gB) from the KSHV genome (Human herpesvirus 8 strain JSC-1 clone BAC16, complete genome GenBank: GQ994935.1) was cloned into pCAG expression vector under control of CMV enhancer and chicken β-actin promoter. For purification and isolation, a FLAG tag was placed between residues A23 and A24 after the endogenous signal sequence. A truncated version of gB, consisting of residues 1-687, was generated by removing the C-terminal membrane proximal region (MPR) and the sequence beyond it. Based on the prefusion HSV-1 gB structure study (22), a substitution of D470P in the corresponding hinge region was introduced into the truncated gB with the intention to lock the prefusion conformation. The protein expression plasmids of full-length gB with mutations D470P, G471P, or I472P of ORF8 were constructed by oligo mutagenesis through HiFi assembly approach (New England Biolabs E2621L).

3.6.2 Generation of KSHV with D470P or G471P substitution in the gB coding region

Recombinant KSHV mutant viruses were separately generated with en passant mutagenesis to substitute D470 or G471 of ORF8 with proline and add the FLAG tag at the C-terminus using the BAC16 genetic system in GS1783 *E. coli* in accordance with previously described protocols (48). BAC DNA containing the KSHV D470P genome was analyzed by restriction enzyme digestion and Sanger sequencing to ensure no gross rearrangements occurred and was transfected into 293T cells using Lipofectamine (Thermo Fisher Scientific 11668027). 293T cells stably harboring the viral genome were selected with 100 mg hygromycin. The 293T cells supplemented with wild type gB by transfection were cocultured with iSLK cells in the presence of 20 ng/mL 12-O-tetradecanoyl-phorbol-13-acetate (Fisher Scientific AAJ63916MCR) and 1 mM sodium butyrate (Fisher Scientific AC263191000) to induce virion production and subsequent infection of iSLK cells. Infected iSLK cells were gradually selected with hygromycin

from 100ug/ml to 1.2mg/ml to establish iSLK-D470P or iSLK-G417P cell line. Expression of gB protein was examined by western blotting of lysates from iSLK-D470P or iSLK-G471P cells that were treated with 1 mM sodium butyrate (NaB) and 5 μg/mL tetracycline to reactivate KSHV and induce lytic replication. To obtain KSHV virions for structural studies, the supernatant was harvested at day 5 after iSLK-D470P or iSLK-G471P cells treated with 1mM NaB and 5 μg/mL doxycycline, when >90% of cells exhibited cytopathic effects.

3.6.3 Western blotting

Western blotting was performed to detect expression of mutant gB proteins either from transfection of expression plasmids in 293T cells or from lytic replication of KSHV in iSLK-D470P or iSLK-G471P cells. For plasmid gene expression, 4 μg of plasmids were transfected into 293T cells in 6-well plates (Lipofectamine 2000) and the cell lysates in 400 ul RIPA buffer were collected 2 days post transfection. For viral gene expression, iSLK-D470P or iSLK-G471P cells were treated with NaB and tetracycline and the cell lysates were collected in RIPA buffer 2 days after reactivation. The supernatant medium was also collected to examine the release of gB protein from infected cells. Equal volume/amounts of each sample were loaded for western blots and then probed with anti-FLAG antibody.

3.6.4 Protein expression and purification

The proteins were transiently expressed in ExpiCHO™ (Chinese hamster ovary) cells (Thermo Fisher Scientific). The cell culture and DNA plasmid transfection were performed by following the manufacturer's instructions. Briefly, the ExpiCHO cells were cultured at an initial density of 0.2 million cells/mL and grown to a density of 6 million cells/mL and transfected with plasmid DNA incubated with ExpiFectamine™ CHO Transfection Kit (Gibco™ Catalog number:

A29131). Typically, 200 mL of transfected cells were harvested 72 h post-transfection. The cells were resuspended with a lysis buffer (20 mM Tris-HCl (pH 8.0) and 150 mM NaCl) supplemented with 2 mM PMSF, lysed by homogenization, and cellular debris was removed by centrifugation. The recombinant gB proteins were purified using EZview™ Red ANTI-FLAG® M2 Affinity Gel (Millipore Sigma) and competitively eluted by 250 μg/mL FLAG peptides (GenScript) in the lysis buffer. The eluted proteins were concentrated to 0.5 mg/mL and underwent buffer exchange to remove residual FLAG peptides. The freshly purified proteins were immediately used for preparing cryoEM grids.

3.6.5 Negative-stain EM

8 μl of 0.02 mg/ml purified gB sample was applied onto glow-discharged ultrathin formvar/carbon supported copper 400-mesh grids (EMS), blotted and stained with 2.0% uranyl acetate. Negative-stained grids were imaged on a Tecnai F20 transmission electron microscope (FEI) operated at 200 kV.

3.6.6 CryoEM sample preparation and data acquisition

3 uL aliquots of the purified proteins were applied to glow-discharged Quantifoil R2/1 Cu 300-mesh grids (Quantifoil, Germany). These grids were then blotted for 10 sec at 100% humidity and vitrified in a mixture of liquid ethane and propane (at a 3:7 ratio) cooled by liquid nitrogen using Vitrobot Mark IV (Thermo Fisher Scientific). Vitrified grids were screened in a Tecnai F20 (FEI) transmission electron microscope to optimize the freezing conditions. Subsequent cryoEM data was collected on a Titan Krios G3i (Thermo Fisher Scientific) equipped with a K3 direct electron detector and post-BioQuantum GIF energy filter (Gatan) operated at 300 kV in electron counting mode. Movies were collected at a nominal magnification of 105,000× in super-

resolution mode after binning by a factor of 2, resulting in an effective pixel size of 0.86 Å. A total dose of 50 $e^-/Å^2$ per movie was used with a dose rate of 15-20 e^-/pix/sec. Initial assessment of a small data set showed that KSHV gB has a preferred particle orientation, with its top view predominantly observed in the 2D class averages. To sample more diverse particle orientations, movies were collected with the stage at both zero tilt and a 30-degree tilt (49). In total, 11,911 movies were recorded without stage tilt and 2,722 movies at 30 degrees tilt by automated data acquisition with EPU version 3.5.0.

3.6.7 CryoEM data processing

The movies were imported into cryoSPARC software package (50) and subjected to patch motion correction and CTF estimation in cryoSPARC. Initially, reference-free manual particle picking in a small subset of data was performed to generate 2D templates for auto-picking and to assess the data quality. 4,041,505 and 1,501,414 particles were picked initially from the no-tilt and 30-degree tilt data set, respectively. The particles were extracted and iterative rounds of 2D classification were performed. 215,660 particles from the 2D class averages with clear secondary structure features were selected from the no-tilt data set and used for the 3D *ab initio* reconstruction to generate two initial volumes. The particles from the last round of the 2D classification of the no-tilt data set were reselected to salvage initially missed good particles, which resulted in 249,493 particles. These particles from the no-tilt data set were combined with 58,924 particles from the 2D classification of the 30-degree tilt data set and used in the following heterogeneous refinement with two copies of each of the two *ab initio* classes as starting volumes. A single class containing 21.9% of the particles showed a feature of postfusion form of class III fusion proteins. Non-uniform refinement (51) was performed with C3 symmetry to yield the final 3.3 Å resolution map. CryoEM statistics are reported in Supplementary Table 3-S1.

We did not observe any prefusion-like 2D or 3D classes in these data sets. To investigate the continuous motion around DIV region, a cryoSPARC (v.4.3.1) 3DFlex (40) training model was generated with 2 latent dimensions and a rigidity prior (lambda) of 1. 41 frames were generated to visualize the volume series.

3.6.8 Atomic model building

The crystal structure of postfusion EBV gB (PDB ID: 3FVC) (14) was used as a homology model and fitted into the cryoEM density map using ChimeraX (52). The model was then mutated and manually refined in Coot (53) followed by iterative refinement using ISOLDE (54) and Phenix real-space refinement (55). Atomic models and maps were visualized in Coot and ChimeraX. There are unknown densities near the DI-DI trimer interface and inside the cavity formed by the trimer (Fig. 3-S4B). These densities did not appear to be part of the proteins or glycans and were left unmodelled. Sequence alignments were performed with Clustal Omega and EMBOSS Needle on the EMBL-EBI server (56) and visualized in Jalview (57) or Linneao (https://github.com/beowulfey/linnaeo).

3.6.9 KSHV culture and isolation

KSHV virions were prepared as previously described with slight modification from iSLK-KSHV-BAC16 cells, received as a gift from Dr. Jae U. Jung (48, 58). In brief, iSLK-KSHV-BAC16 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Corning 10-017-CV) supplemented with 10% fetal bovine serum (FBS; 35-010-CV) and 100 U/mL of penicillin-streptomycin, 1mg/mL puromycin (Invivogen ant-pr-1), 250 µg/mL G418 (Invivogen ant-gn-1), and 1,200 mg/mL hygromycin B (Invivogen ant-hg-1). Cells were cultured to 80-90% confluency in 24 15 cm tissue culture dishes. Then, KSHV lytic replication and virion production

were induced by treatment with 1 mM sodium butyrate and 5 μg/mL tetracycline for five days, after which 720 mL of media containing KSHV virions were collected. Infectious media was clarified by centrifugation at 10,000 × g for 20 min at 4 °C twice. Virions were then pelleted using an SW28 rotor (Beckman) at 80,000 × g for 1 h at 4 °C. The supernatant was discarded and 50 μl of cold phosphate-buffered saline (PBS) was added to the pellet in each tube. The pellets were allowed to soften for 16 h on ice in a 4°C cold room before gentle resuspension by tapping. About 0.7 ml of combined pellet suspension from all ultracentrifuge tubes were loaded on 15–50% linear gradient of sucrose made in PBS and spun in SW41 (Beckman) rotor at 80,000 × g for 1 h at 4 °C. About 1 ml of the visible virus-containing band was collected into a new tube and mixed with 10 ml of sterile PBS solution in order to dilute sucrose. KSHV gB G471P did not have a visible band and was instead fractionated in 1 mL aliquots based on the band locations of other herpesviruses. The virions were concentrated at 80,000 × g for 1 h at 4 °C and 20 μl of PBS was added. After softening and gentle resuspension, the isolated virions were recovered and immediately used for cryoET sample preparation without freezing.

3.6.10 CryoET sample preparation and image acquisition

Virion samples were screened on an FEI Tecnai TF20 equipped with a 4k × 4k TVIPS F415MP CCD detector or Gatan K2 detector. To prepare EM grids for cryoET imaging, 3 μl of diluted virions were mixed with 10-nm gold fiducials and applied to Quantifoil R2/1 Cu 200 grids and plunge-frozen in liquid ethane-propane mixture using a Vitrobot mark IV (FEI Thermo Fisher Scientific) set to 100% humidity, 4 °C, and blot time of 5 s.

Imaging was performed with a Titan Krios G1 (FEI Thermo Fisher Scientific) operated at 300 kV, and movies were recorded using a Gatan K3 detector with an energy filter slit width of 20 eV. Tilt series were collected from -60° to 60° in 3° increments using a dose-symmetric

scheme (59) using SerialEM (60) at 64,000 × magnification (0.690 Å/pixel super-resolution) and a target focus of -4 μm with a total of dose of 120 e-/Å^2. Individual tilts were recorded with constant beam intensity in 10 frames.

3.6.11 Tomogram reconstruction

Each movie stack was drift-corrected using MotionCor2 (61) and defocus values were determined with CTFFIND4 (62). Tilt series alignment via gold fiducials and tomogram reconstructions were performed in IMOD with a binning factor of 4 (63). IsoNet isotropic reconstructions with a binning factor of 8 were used for observation (64).

3.7 Data Availability

The atomic model has been deposited in the PDB with accession code: 9CU4. The cryo-EM map has been deposited in the EMDB with accession codes: EMD-45927.

3.8 Acknowledgements

This research was supported in part by grants from the National Institutes of Health (R01DE025567 to Z.H.Z and T.-T.W.). J.Z. acknowledges support from an Interdisciplinary Training in Virology and Gene Therapy training grant (NIH 5T32AI060567). We acknowledge the use of resources in the Electron Imaging Center for Nanomachines supported by UCLA, Core Center of Excellence in Nano Imaging (CNI) at USC, and grants from NIH (S10RR23057, S10OD018111, U24GM116792, and T32GM145388) and NSF (DBI-1338135). We also acknowledge use of resources from the UCLA AIDS Institute, the James B. Pendleton Charitable Trust, and the McCarthy Family Foundation.

3.9 Author Contributions

J.Z. and Z.H.Z. conceived the project. Z.H.Z. and T.-T.W. supervised the research. J.Z. and F.I. designed experimental protocols. H.H. prepared constructs and cultured virions. J.Z. isolated virions and collected cryoET data. G.X., F.I., and J.S. expressed and purified samples. F.I. and G.X. prepared samples and collected cryoEM images. F.I. and J.Z. determined the 3D structure, built atomic models, interpreted the results, made figures, and wrote the manuscript. All authors reviewed and approved the manuscript.

3.10 Conflict of Interest

The authors declare that they have no conflicts of interest.

3.11 References

1. Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, Knowles DM, Moore PS. 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 266:1865–1869.

2. Cesarman E, Chang Y, Moore PS, Said JW, Knowles DM. 1995. Kaposi's Sarcoma–Associated Herpesvirus-Like DNA Sequences in AIDS-Related Body-Cavity–Based Lymphomas. New England Journal of Medicine 332:1186–1191.

3. Soulier J, Grollet L, Oksenhendler E, Cacoub P, Cazals-Hatem D, Babinet P, d'Agay M, Clauvel J, Raphael M, Degos L, Sigaux F. 1995. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease [see comments]. Blood 86:1276–1280.

4. Dittmer DP, Damania B. 2019. Kaposi sarcoma associated herpesvirus (KSHV) Associated Disease in the AIDS Patient: An Update. Cancer Treat Res 177:63–80.

5. Naimo E, Zischke J, Schulz TF. 2021. Recent Advances in Developing Treatments of Kaposi's Sarcoma Herpesvirus-Related Diseases. Viruses 13:1797.

6. Lurain K, Yarchoan R, Ramaswami R. 2022. Immunotherapy for KSHV-associated diseases. Current Opinion in Virology 55:101249.

7. Johnson DC, Wisner TW, Wright CC. 2011. Herpes Simplex Virus Glycoproteins gB and gD Function in a Redundant Fashion To Promote Secondary Envelopment ▿. J Virol 85:4910–4926.

8. Farnsworth A, Wisner TW, Webb M, Roller R, Cohen G, Eisenberg R, Johnson DC. 2007. Herpes simplex virus glycoproteins gB and gH function in fusion between the virion envelope and the outer nuclear membrane. Proceedings of the National Academy of Sciences 104:10187–10192.

9. Owen DJ, Crump CM, Graham SC. 2015. Tegument Assembly and Secondary Envelopment of Alphaherpesviruses. 9. Viruses 7:5084–5114.

10. Connolly SA, Jardetzky TS, Longnecker R. 2021. The structural basis of herpesvirus entry. 2. Nature Reviews Microbiology 19:110–121.

11. Backovic M, Leser GP, Lamb RA, Longnecker R, Jardetzky TS. 2007. Characterization of EBV gB indicates properties of both class I and class II viral fusion proteins. Virology 368:102–113.

12. Backovic M, Jardetzky TS. 2009. Class III viral membrane fusion proteins. Curr Opin Struct Biol 19:189–196.

13. Heldwein EE, Lou H, Bender FC, Cohen GH, Eisenberg RJ, Harrison SC. 2006. Crystal Structure of Glycoprotein B from Herpes Simplex Virus 1. Science 313:217–220.

14. Backovic M, Longnecker R, Jardetzky TS. 2009. Structure of a trimeric variant of the Epstein–Barr virus glycoprotein B. PNAS 106:2880–2885.

15. Chandramouli S, Ciferri C, Nikitin PA, Caló S, Gerrein R, Balabanis K, Monroe J, Hebner C, Lilja AE, Settembre EC, Carfi A. 2015. Structure of HCMV glycoprotein B in the postfusion conformation bound to a neutralizing human antibody. 1. Nature Communications 6:8176.

16. Burke HG, Heldwein EE. 2015. Crystal Structure of the Human Cytomegalovirus Glycoprotein B. PLOS Pathogens 11:e1005227.

17. Stampfer SD, Lou H, Cohen GH, Eisenberg RJ, Heldwein EE. 2010. Structural Basis of Local, pH-Dependent Conformational Changes in Glycoprotein B from Herpes Simplex Virus Type 1. Journal of Virology 84:12924–12933.

18. Vitu E, Sharma S, Stampfer SD, Heldwein EE. 2013. Extensive mutagenesis of the HSV-1 gB ectodomain reveals remarkable stability of its postfusion form. J Mol Biol 425:2056–2071.

19. Cooper RS, Georgieva ER, Borbat PP, Freed JH, Heldwein EE. 2018. Structural basis for membrane anchoring and fusion regulation of the herpes simplex virus fusogen gB. Nat Struct Mol Biol 25:416–424.

20. Oliver SL, Xing Y, Chen D-H, Roh SH, Pintilie GD, Bushnell DA, Sommer MH, Yang E, Carfi A, Chiu W, Arvin AM. 2020. A glycoprotein B-neutralizing antibody structure at 2.8 Å uncovers a critical domain for herpesvirus fusion initiation. 1. Nature Communications 11:4141.

21. Si Z, Zhang J, Shivakoti S, Atanasov I, Tao C-L, Hui WH, Zhou K, Yu X, Li W, Luo M, Bi G-Q, Zhou ZH. 2018. Different functional states of fusion protein gB revealed on human cytomegalovirus by cryo electron tomography with Volta phase plate. PLOS Pathogens 14:e1007452.

22. Vollmer B, Pražák V, Vasishtan D, Jefferys EE, Hernandez-Duran A, Vallbracht M, Klupp BG, Mettenleiter TC, Backovic M, Rey FA, Topf M, Grünewald K. 2020. The prefusion structure of herpes simplex virus glycoprotein B. Science Advances 6:eabc1726.

23. Zhou M, Vollmer B, Machala E, Chen M, Grünewald K, Arvin AM, Chiu W, Oliver SL. 2023. Targeted mutagenesis of the herpesvirus fusogen central helix captures transition states. 1. Nat Commun 14:7958.

24. Zeev-Ben-Mordehai T, Vasishtan D, Hernández Durán A, Vollmer B, White P, Prasad Pandurangan A, Siebert CA, Topf M, Grünewald K. 2016. Two distinct trimeric conformations of natively membrane-anchored full-length herpes simplex virus 1 glycoprotein B. Proc Natl Acad Sci U S A 113:4176–4181.

25. Fontana J, Atanasiu D, Saw WT, Gallagher JR, Cox RG, Whitbeck JC, Brown LM, Eisenberg RJ, Cohen GH. 2017. The Fusion Loops of the Initial Prefusion Conformation of Herpes Simplex Virus 1 Fusion Protein Point Toward the Membrane. mBio 8:e01268-17.

26. Liu Y, Heim KP, Che Y, Chi X, Qiu X, Han S, Dormitzer PR, Yang X. 2021. Prefusion structure of human cytomegalovirus glycoprotein B and structural basis for membrane fusion. Science Advances 7:eabf3178.

27. Sponholtz MR, Byrne PO, Lee AG, Ramamohan AR, Goldsmith JA, McCool RS, Zhou L, Johnson NV, Hsieh C-L, Connors M, Karthigeyan K, Campbell JD, Permar SR, Maynard JA, Yu D, Bottomley MJ, McLellan JS. 2024. Structure-based design of a soluble human

cytomegalovirus glycoprotein B antigen stabilized in a prefusion-like conformation. bioRxiv https://doi.org/10.1101/2024.02.10.579772.

28. Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, Salazar-Gonzalez JF, Salazar MG, Kilby JM, Saag MS, Komarova NL, Nowak MA, Hahn BH, Kwong PD, Shaw GM. 2003. Antibody neutralization and escape by HIV-1. Nature 422:307–312.

29. Fukui A, Maruzuru Y, Ohno S, Nobe M, Iwata S, Takeshima K, Koyanagi N, Kato A, Kitazume S, Yamaguchi Y, Kawaguchi Y. 2023. Dual impacts of a glycan shield on the envelope glycoprotein B of HSV-1: evasion from human antibodies in vivo and neurovirulence. mBio 14:e00992-23.

30. Wang J, Fan Q, Satoh T, Arii J, Lanier LL, Spear PG, Kawaguchi Y, Arase H. 2009. Binding of herpes simplex virus glycoprotein B (gB) to paired immunoglobulin-like type 2 receptor alpha depends on specific sialylated O-linked glycans on gB. J Virol 83:13042–13045.

31. Arii J, Wang J, Morimoto T, Suenaga T, Akashi H, Arase H, Kawaguchi Y. 2010. A single-amino-acid substitution in herpes simplex virus 1 envelope glycoprotein B at a site required for binding to the paired immunoglobulin-like type 2 receptor alpha (PILRalpha) abrogates PILRalpha-dependent viral entry and reduces pathogenesis. J Virol 84:10773–10783.

32. Zhang X, Hong J, Zhong L, Wu Q, Zhang S, Zhu Q, Chen H, Wei D, Li R, Zhang W, Zhang X, Wang G, Zhou X, Chen J, Kang Y, Zha Z, Duan X, Huang Y, Sun C, Kong X, Zhou Y, Chen Y, Ye X, Feng Q, Li S, Xiang T, Gao S, Zeng M-S, Zheng Q, Chen Y, Zeng Y-X, Xia N, Xu M. 2022. Protective anti-gB neutralizing antibodies targeting two vulnerable sites for EBV-cell membrane fusion. Proceedings of the National Academy of Sciences 119:e2202371119.

33. Luo S, Hu K, He S, Wang P, Zhang M, Huang X, Du T, Zheng C, Liu Y, Hu Q. 2015. Contribution of N-linked glycans on HSV-2 gB to cell–cell fusion and viral entry. Virology 483:72–82.

34. Kuraoka M, Aschner CB, Windsor IW, Mahant AM, Garforth SJ, Kong SL, Achkar JM, Almo SC, Kelsoe G, Herold BC. 2023. A non-neutralizing glycoprotein B monoclonal antibody protects against herpes simplex virus disease in mice. J Clin Invest 133.

35. Sun C, Yang J-W, Xie C, Fang X-Y, Bu G-L, Zhao G-X, Dai D-L, Liu Z, Zeng M-S. 2023. The structure of HSV-1 gB bound to a potent neutralizing antibody reveals a conservative antigenic domain across herpesviruses. hLife https://doi.org/10.1016/j.hlife.2023.12.004.

36. Seyfizadeh N, Kalbermatter D, Imhof T, Ries M, Müller C, Jenner L, Blumenschein E, Yendrzheyevskiy A, Grün F, Moog K, Eckert D, Engel R, Diebolder P, Chami M, Krauss J, Schaller T, Arndt M. 2024. Development of a highly effective combination monoclonal antibody therapy against Herpes simplex virus. Journal of Biomedical Science 31:56.

37. Vallbracht M, Brun D, Tassinari M, Vaney M-C, Pehau-Arnaudet G, Guardado-Calvo P, Haouz A, Klupp BG, Mettenleiter TC, Rey FA, Backovic M. 2018. Structure-Function Dissection of Pseudorabies Virus Glycoprotein B Fusion Loops. Journal of Virology 92.

38. Li X, Yang F, Hu X, Tan F, Qi J, Peng R, Wang M, Chai Y, Hao L, Deng J, Bai C, Wang J, Song H, Tan S, Lu G, Gao GF, Shi Y, Tian K. 2017. Two classes of protective antibodies against

Pseudorabies virus variant glycoprotein B: Implications for vaccine design. PLOS Pathogens 13:e1006777.

39. Oliver SL, Xing Y, Chen D-H, Roh SH, Pintilie GD, Bushnell DA, Sommer MH, Yang E, Carfi A, Chiu W, Arvin AM. 2021. The N-terminus of varicella-zoster virus glycoprotein B has a functional role in fusion. PLOS Pathogens 17:e1008961.

40. Punjani A, Fleet DJ. 2021. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. Journal of Structural Biology 213:107702.

41. Watanabe Y, Bowden TA, Wilson IA, Crispin M. 2019. Exploitation of glycosylation in enveloped virus pathobiology. Biochim Biophys Acta Gen Subj 1863:1480–1497.

42. Cao L, Diedrich JK, Kulp DW, Pauthner M, He L, Park S-KR, Sok D, Su CY, Delahunty CM, Menis S, Andrabi R, Guenaga J, Georgeson E, Kubitz M, Adachi Y, Burton DR, Schief WR, Yates JR, Paulson JC. 2017. Global site-specific N-glycosylation analysis of HIV envelope glycoprotein. Nat Commun 8:14954.

43. Behrens A-J, Harvey DJ, Milne E, Cupo A, Kumar A, Zitzmann N, Struwe WB, Moore JP, Crispin M. 2017. Molecular Architecture of the Cleavage-Dependent Mannose Patch on a Soluble HIV-1 Envelope Glycoprotein Trimer. J Virol 91:e01894-16.

44. Nelson CS, Herold BC, Permar SR. 2018. A new era in cytomegalovirus vaccinology: considerations for rational design of next-generation vaccines to prevent congenital cytomegalovirus infection. NPJ Vaccines 3:38.

45. Zhong L, Krummenacher C, Zhang W, Hong J, Feng Q, Chen Y, Zhao Q, Zeng M-S, Zeng Y-X, Xu M, Zhang X. 2022. Urgency and necessity of Epstein-Barr virus prophylactic vaccines. NPJ Vaccines 7:159.

46. Baghian A, Luftig M, Black JB, Meng YX, Pau CP, Voss T, Pellett PE, Kousoulas KG. 2000. Glycoprotein B of human herpesvirus 8 is a component of the virion in a cleaved form composed of amino- and carboxyl-terminal fragments. Virology 269:18–25.

47. Bagdonaite I, Nordén R, Joshi HJ, King SL, Vakhrushev SY, Olofsson S, Wandall HH. 2016. Global Mapping of O-Glycosylation of Varicella Zoster Virus, Human Cytomegalovirus, and Epstein-Barr Virus. J Biol Chem 291:12014–12028.

48. Brulois KF, Chang H, Lee AS-Y, Ensser A, Wong L-Y, Toth Z, Lee SH, Lee H-R, Myoung J, Ganem D, Oh T-K, Kim JF, Gao S-J, Jung JU. 2012. Construction and Manipulation of a New Kaposi's Sarcoma-Associated Herpesvirus Bacterial Artificial Chromosome Clone. Journal of Virology 86:9708–9720.

49. Tan YZ, Baldwin PR, Davis JH, Williamson JR, Potter CS, Carragher B, Lyumkis D. 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nature Methods 14:793–796.

50. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature Methods 14:290–296.

51. Punjani A, Zhang H, Fleet DJ. 2020. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat Methods 17:1214–1221.

52. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Science 27:14–25.

53. Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. ARRAY(0xab2a914). Acta Cryst D 60:2126–2132.

54. Croll TI. 2018. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. 6. Acta Cryst D 74:519–530.

55. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213–221.

56. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.

57. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189–1191.

58. Dai X, Gong D, Wu T-T, Sun R, Zhou ZH. 2014. Organization of Capsid-Associated Tegument Components in Kaposi's Sarcoma-Associated Herpesvirus. Journal of Virology 88:12694–12702.

59. Hagen WJH, Wan W, Briggs JAG. 2017. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. Journal of Structural Biology 197:191–198.

60. Mastronarde DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. Journal of Structural Biology 152:36–51.

61. Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. 4. Nature Methods 14:331–332.

62. Rohou A, Grigorieff N. 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology 192:216–221.

63. Mastronarde DN, Held SR. 2017. Automated tilt series alignment and tomographic reconstruction in IMOD. Journal of Structural Biology 197:102–113.

64. Liu Y-T, Zhang H, Wang H, Tao C-L, Bi G-Q, Zhou ZH. 2022. Isotropic reconstruction for electron tomography with deep learning. 1. Nat Commun 13:6482.

3.12 Figures

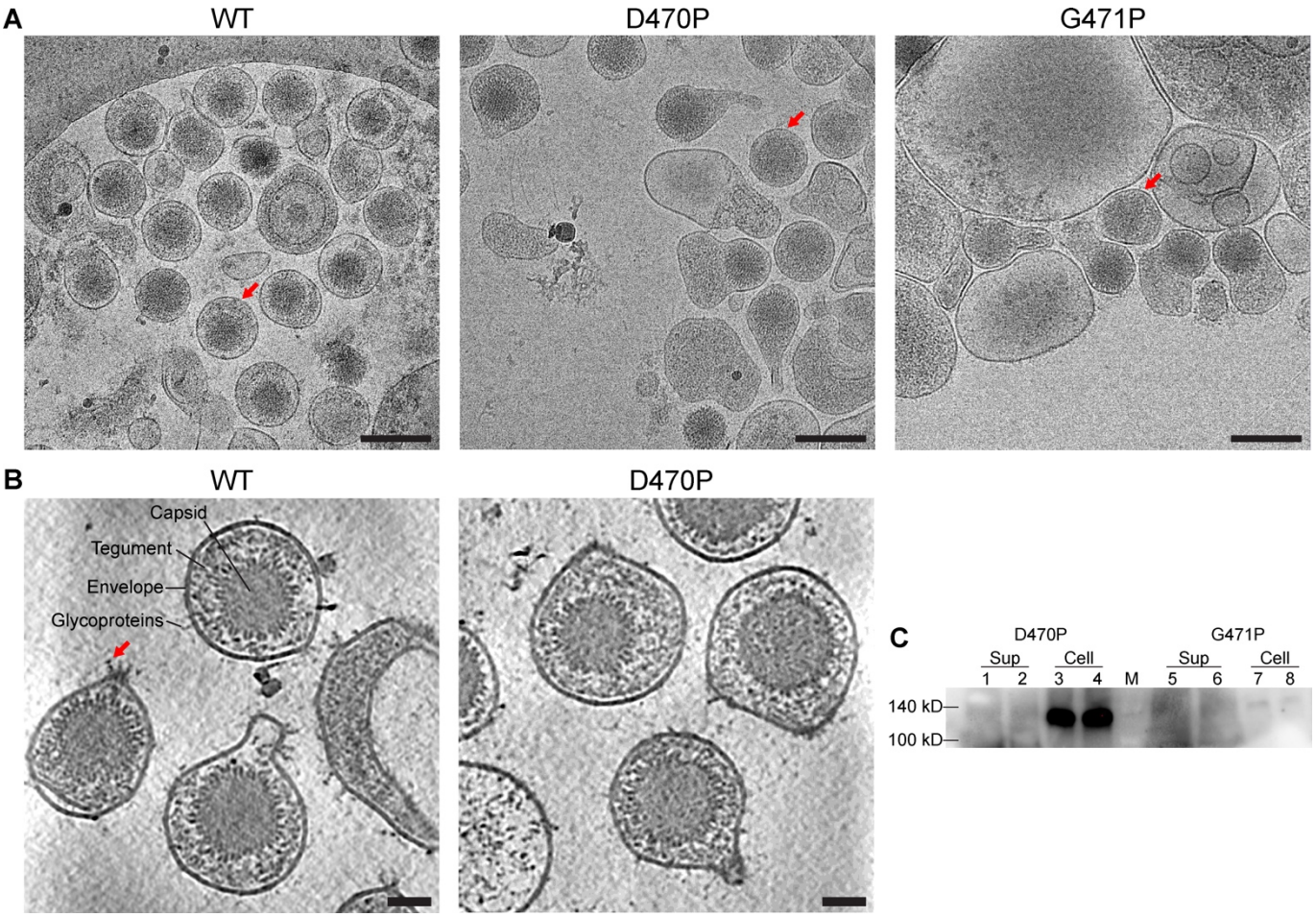


Fig. 3-1. KSHV virions. (A) CryoEM images of KSHV virions expressing WT gB (left), gB D470P (middle), and gB G471P (right). Red arrow indicates example mature virion. Scale bar, 200 nm. **(B)** Tomogram of KSHV expressing WT gB (left) and gB D470P (right). Red arrow indicates example postfusion gB. Scale bar, 50 nm. **(C)** Western blot for expression of the gB mutant in iSLK-D470P or iSLK-G471P cells. Lanes 1, 2, 5, and 6: media; Lane 3, 4, 7, and 8: cell lysates. Lanes 1-4: iSLK-D470P and Lanes 5-8: iSLK-G471P. Lanes 1, 3, 5 and 7: 1mM NaB plus 5 µg/ml tetracycline. Lanes 2, 4, 6 and 8: 1 mM NaB plus 7.5µg/ml tetracycline. M: protein ladder. The membrane was probed with an anti-FLAG antibody.

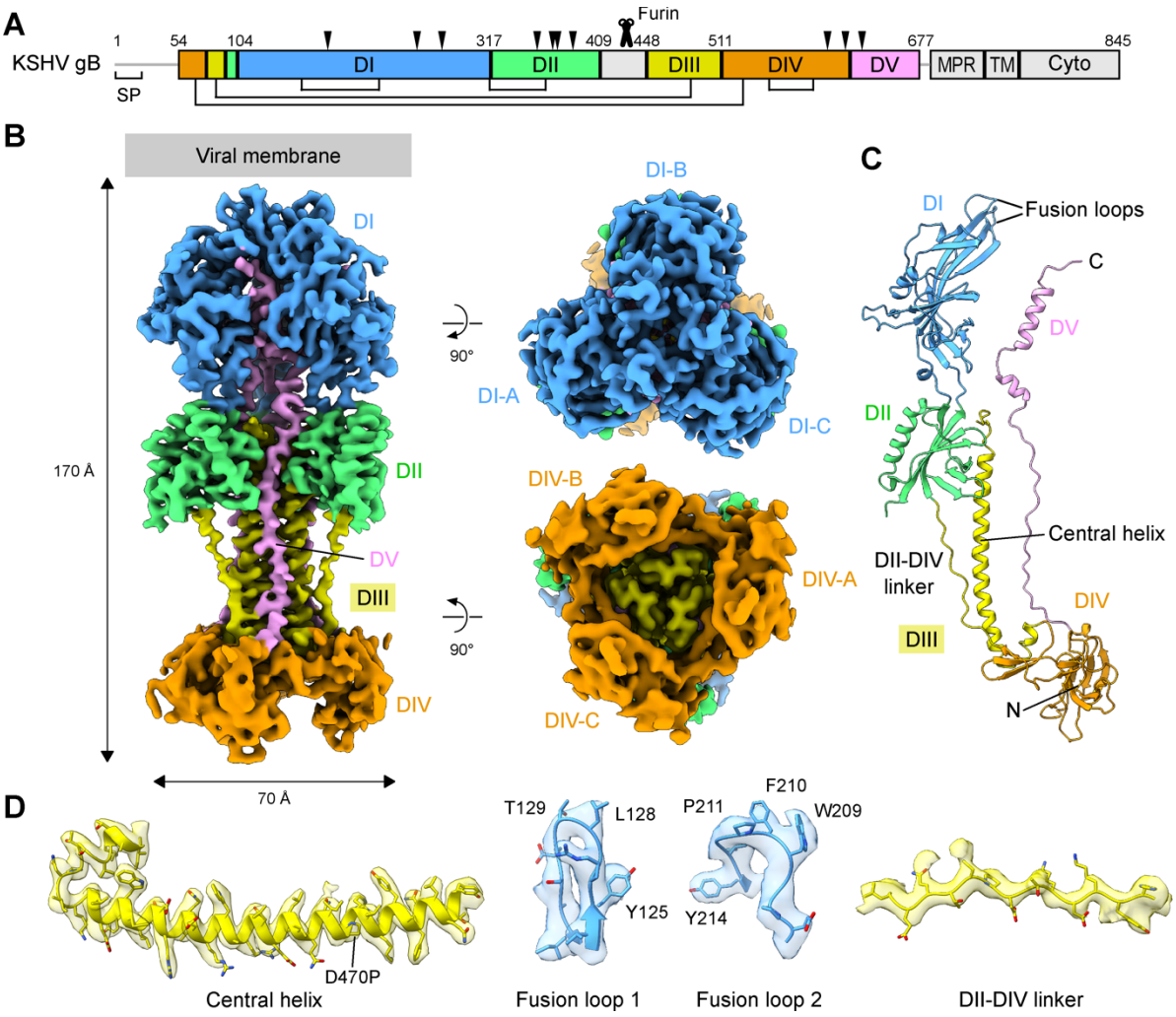


Fig. 3-2. Structure of KSHV gB ectodomain. (A) Schematic representation of the domain organization and construct design of KSHV gB. SP, signal peptide; MPR, membrane proximal region; TM, transmembrane; Cyto, cytoplasmic region. The regions either invisible in the cryoEM map or excluded from the construct are colored in light grey. N-linked glycosylation sites observed in the structure are indicated with black triangles above the schematic boxes. Five conserved disulfide bonds are depicted with black lines below the boxes. **(B)** Three orthogonal views of the cryoEM map of the KSHV gB ectodomain trimer colored by the five core domains. **(C)** Atomic model of a protomer of the gB ectodomain trimer colored by the five core domains. **(D)** Exemplary cryo-EM densities highlighting the segments of DI and DIII. The densities are shown in semi-transparent surfaces superimposed with their corresponding atomic models in ribbons and sticks.

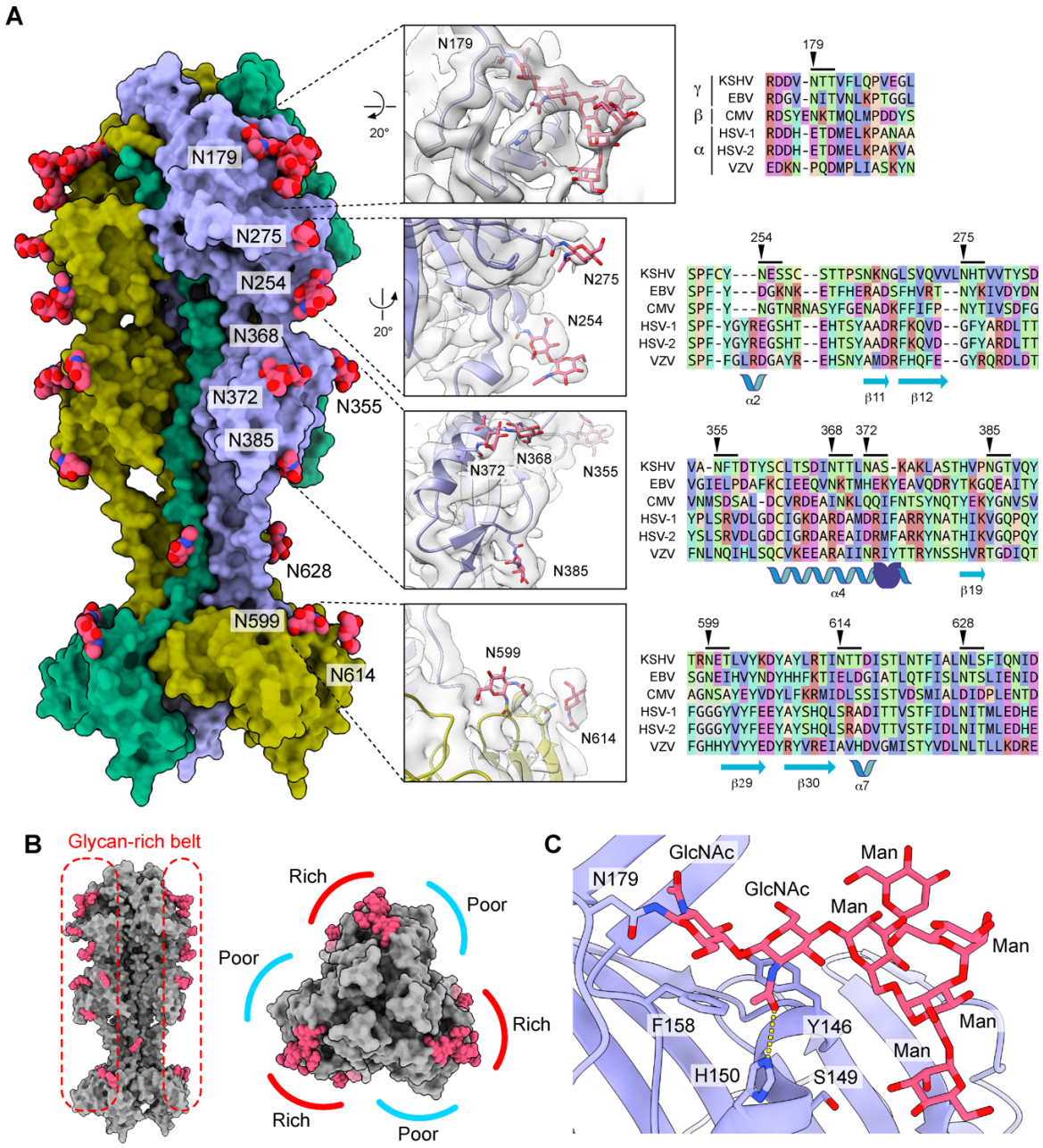


Fig. 3-3. Glycan chain of KSHV gB. (A) Overview of the surface glycosylation of the KSHV gB ectodomain trimer. The atomic models of the proteins are shown in surface representation and the glycan chains are shown in spheres. The close-up views of the glycosylation sites are shown in the insets in the center. The corresponding cryoEM densities are superimposed on the atomic models. Amino acid sequence alignments near the glycosylation sites are shown on the right to highlight the conservation of the glycosylation sites. **(B)** Glycan chain distribution pattern on the KSHV gB surface viewed from the side (left) and top (right), highlighting the glycan-rich and glycan-poor surface areas. **(C)** Close-up view of the glycan chain at N179. High-mannose glycan chain has been modeled based on the observed cryoEM density. The first two ClcNAcs form stacking interactions with the aromatic side chains of Phe158 and Tyr146, respectively.

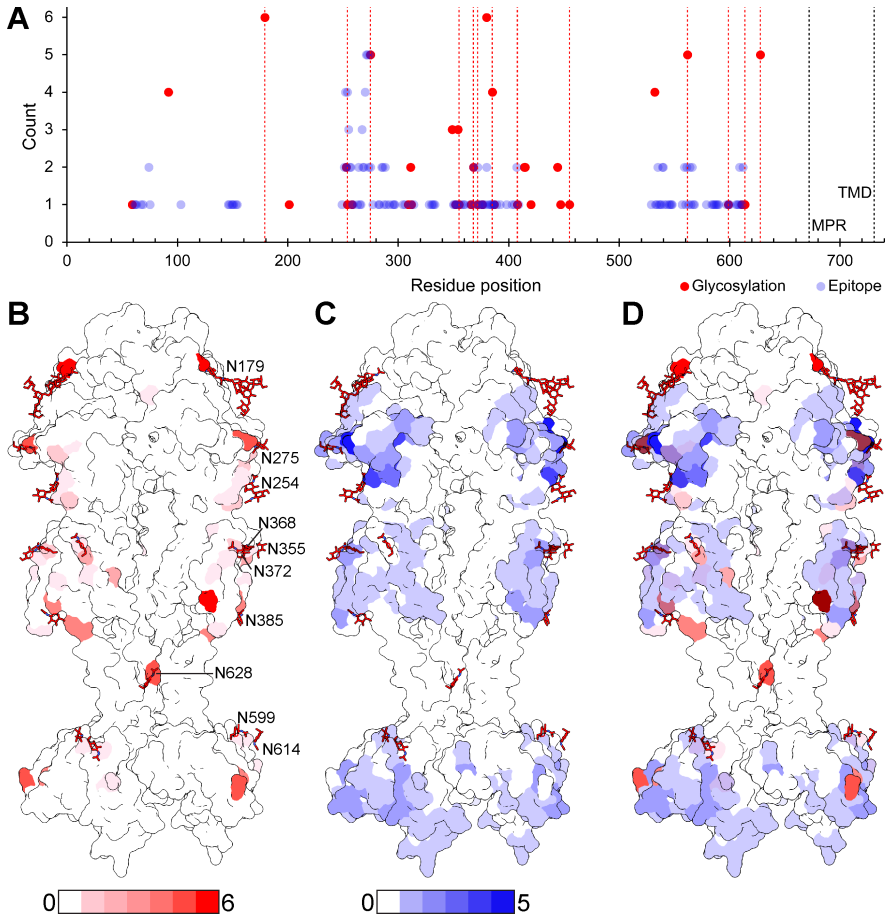


Fig. 3-4. Mapping of human herpesvirus glycosylation and epitope sites on KSHV gB. (A) Plot of count of hypothetical glycosylation and experimental epitope sites across all human herpesviruses (HHV1-8, including 6A and 6B) against equivalent KSHV gB residue position. Dashed red lines indicate positions of KSHV gB glycosylation. Red points indicate total instances of glycosylation (KSHV inclusive) and blue points indicate total instances of antibody contact (<5 Å distance) at corresponding KSHV gB residue position. **(B)** Mapping of glycosylation instances on surface render of KSHV gB atomic model. **(C)** Mapping of epitopes on surface render of KSHV gB atomic model. **(D)** Merge of (C) and (D).

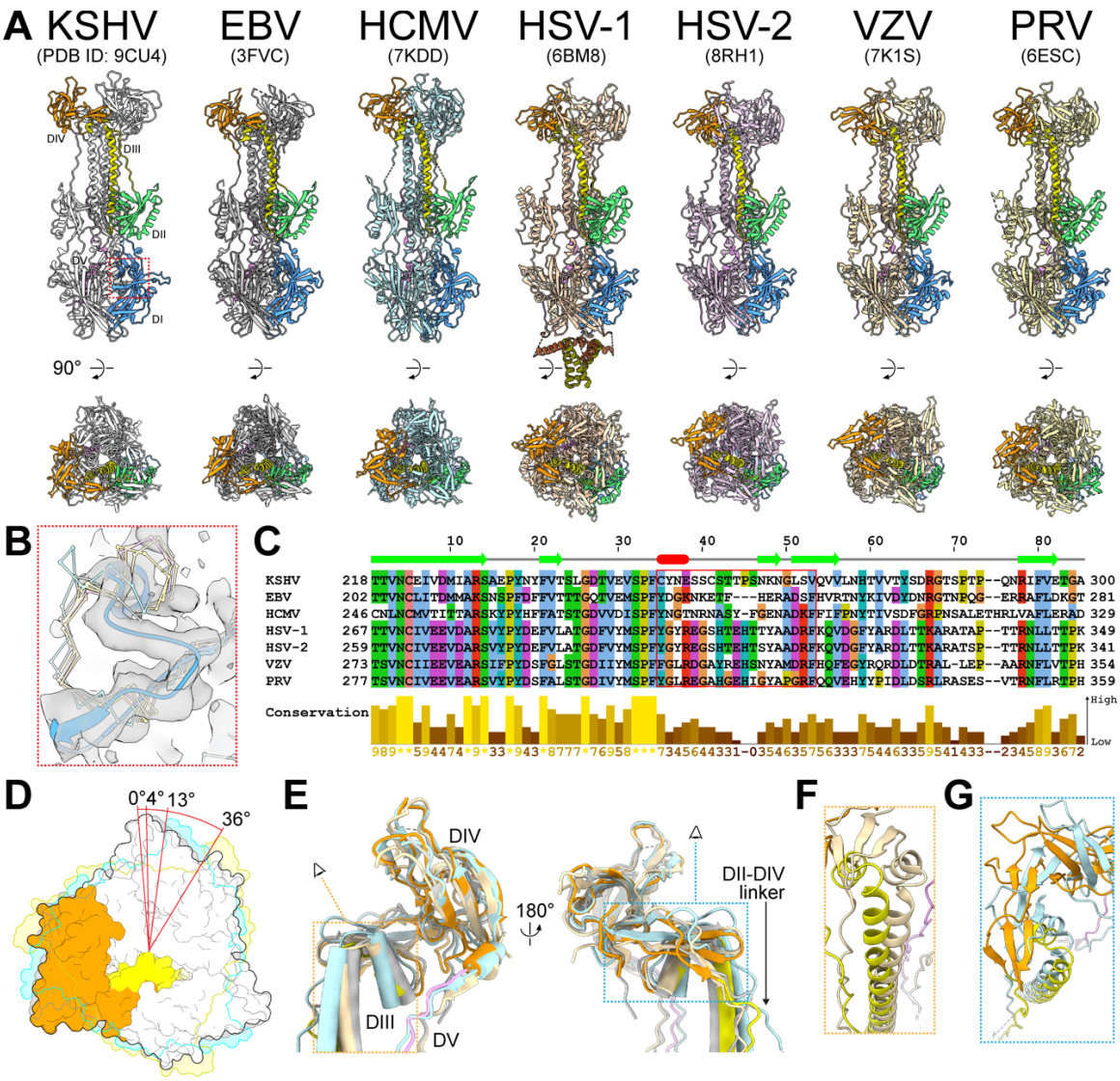


Fig. 3-5. Structural comparison of herpesvirus gB structures. (A) Array of gB trimer atomic models from different herpesvirus species. One protomer from each model is colored by domain while the other two are colored uniquely for each species. Atomic models are aligned by DI down the C3 axis. **(B)** Zoomed in comparison of atomic models of location indicated in (A) superimposed with KSHV gB cryoEM map. KSHV gB model is depicted as ribbon colored by domain, and other gB models are depicted as backbone with their representative color. **(C)** Sequence alignment of herpesvirus gB with red box indicating region where KSHV gB differs in (B). Secondary structures in KSHV gB are indicated above the corresponding residues. **(D)** Overlay of view down the C3 axis from DIV end for KSHV (white), EBV (grey), HCMV (cyan), and VZV (yellow). DIII and DIV for one KSHV gB protomer are colored by domain. Degrees of clockwise rotation of DIV in other species relative to KSHV are indicated. **(E)** Comparison of gB atomic models (KSHV colored by domain; others in representative color) aligned by DIV to identify differences in trajectory of DII-DIV linker, DIII, and DV. **(F)** Comparison of KSHV and

VZV gB DIII from perspective indicated in (E). **(G)** Comparison of KSHV and HCMV gB DIII-DIV interface helix and β-turn from perspective indicated in (E).

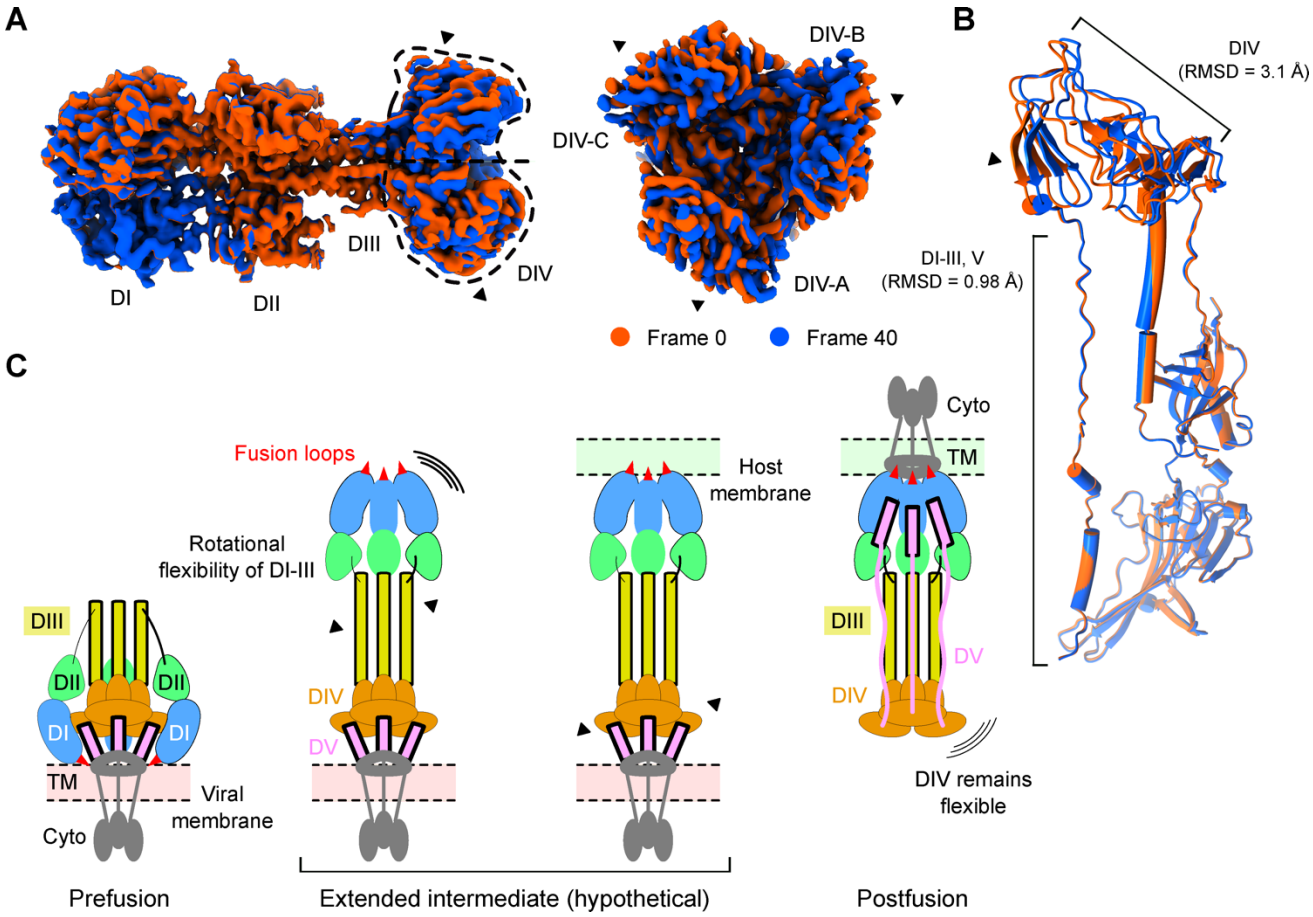


Fig. 3-6. Interdomain dynamics of DIII-DIV. (A) Representative frames (first and last) from the 3D flexibility analysis viewed from the side (left) and bottom (right). **(B)** Structural superimposition of the models derived from the first and last frames from the 3D flexibility analysis. RMSD values for the DIV and the remaining ectodomain are indicated. **(C)** Proposed role of the interdomain flexibility present in the gB during the pre- to postfusion conformation change.

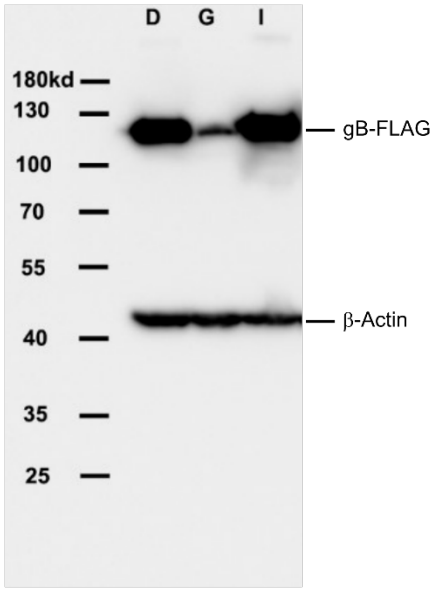


Fig. 3-S1. Expression of gB mutants in 293T cells transfected with expression plasmids. D: D470P; G: G471P; I: I472P. The expression of gB is under control of CMV enhancer + Chicken β-actin promoter from a pCAG vector. The membrane was probed with an anti-FLAG antibody.

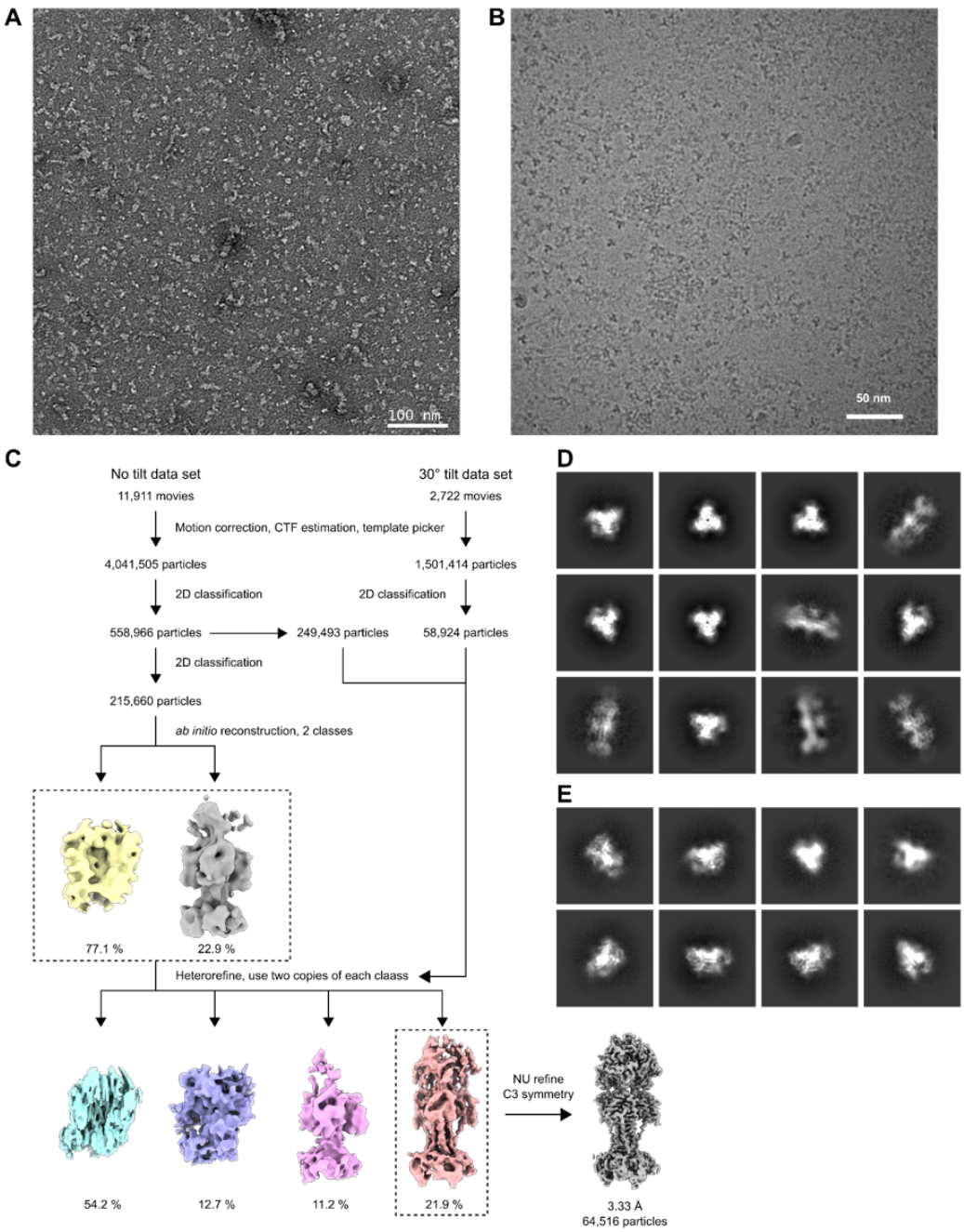


Fig. 3-S2. Workflow and intermediate results of electron microscopy imaging and 3D reconstruction of KSHV gB. (**A**) Representative negative-stain raw image. (**B**) Representative cryoEM raw image. (**C**) CryoEM image processing workflow. (**D**) Representative 2D class averages of the no-tilt data. (**E**) Representative 2D class averages of the 30-degree tilt data.

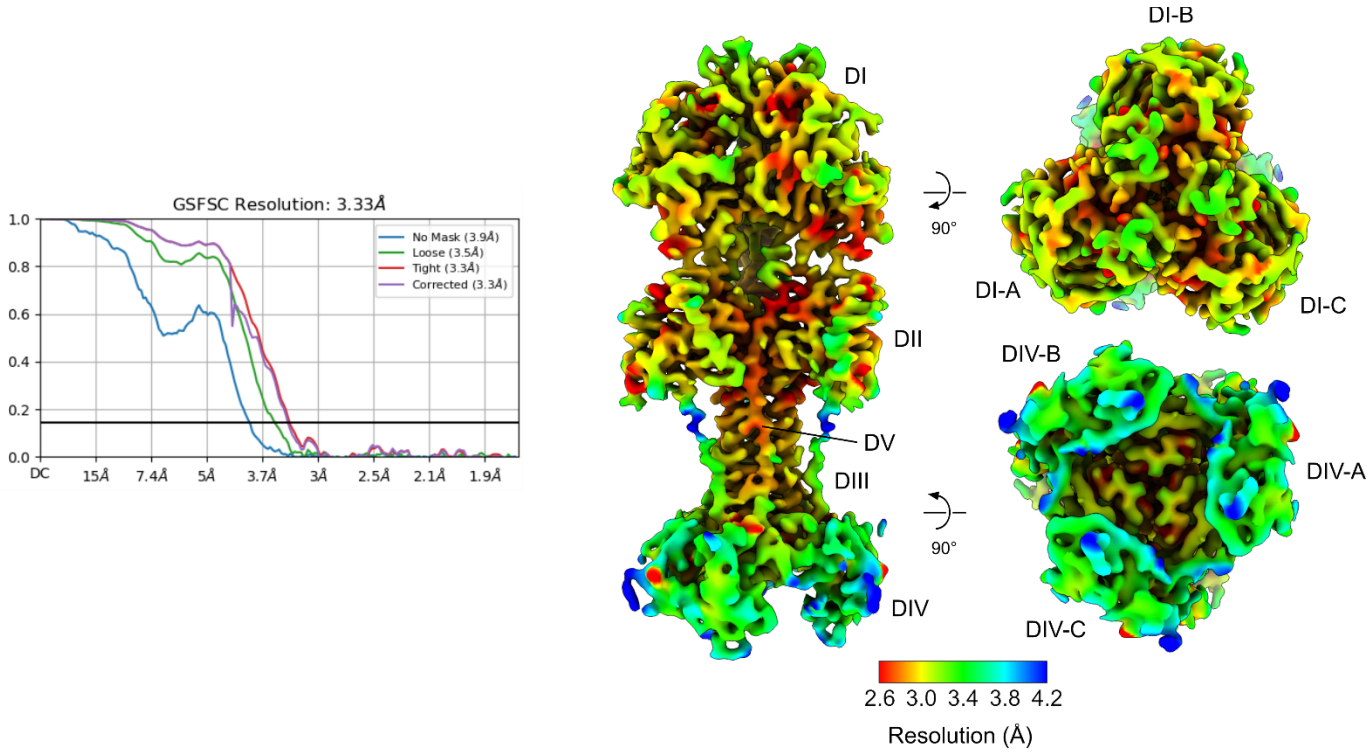


Fig. 3-S3. Evaluation of the cryoEM reconstruction of KSHV gB. Global resolution estimation (**A**), angular distribution plot of the particles (**B**), and local resolution evaluation in three orthogonal views (**C**) of KSHV gB. Resolution estimation is based on the gold standard Fourier shell correlation (FSC) coefficient of 0.143 criteria.

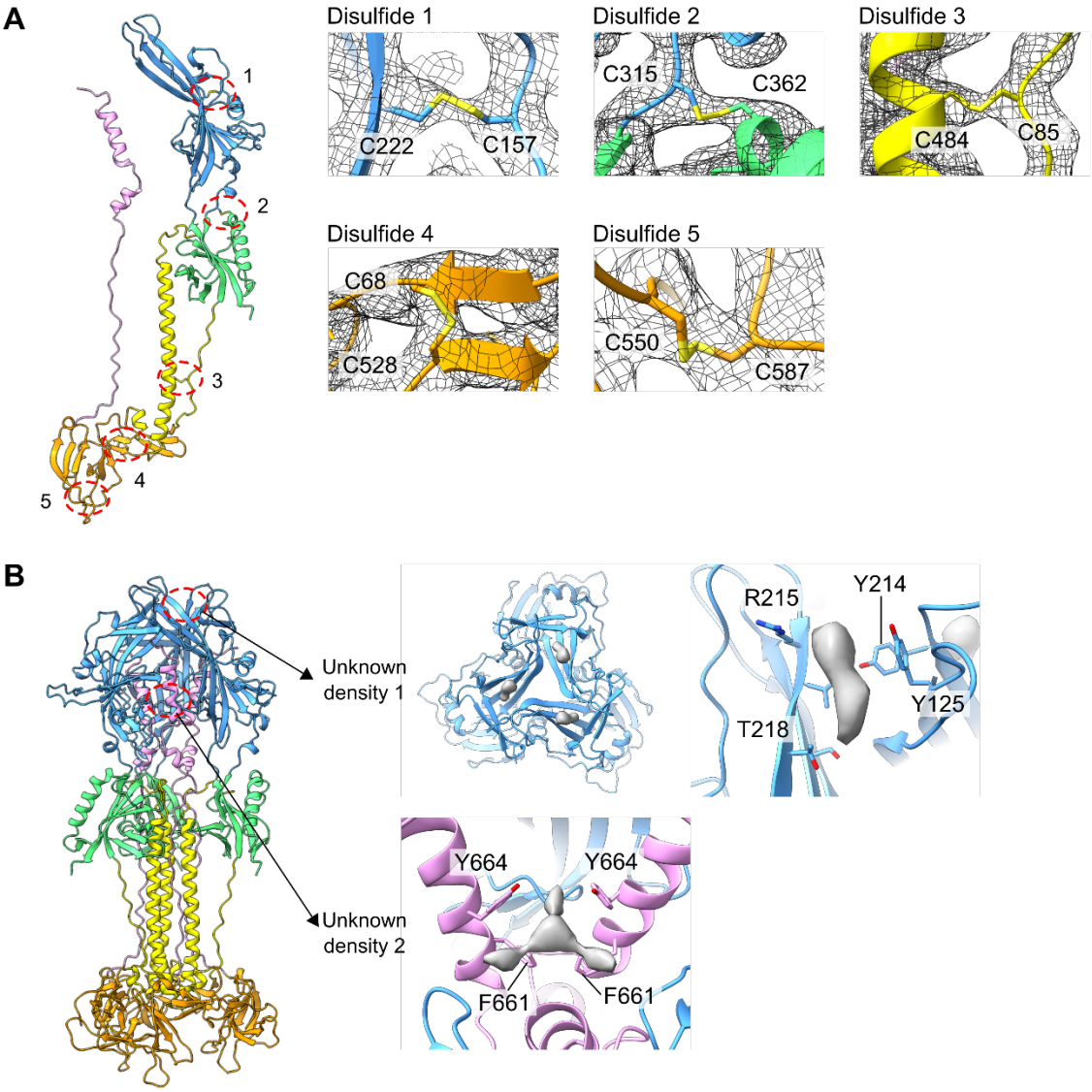


Fig. 3-S4. Intraprotomer disulfide bonds and unidentified density observed in KSHV gB (**A**) CryoEM densities of the five conserved disulfide bonds shown as mesh surface. (**B**) Two unassigned densities near the DI-DI trimer interface and inside the trimer cavity near DV.

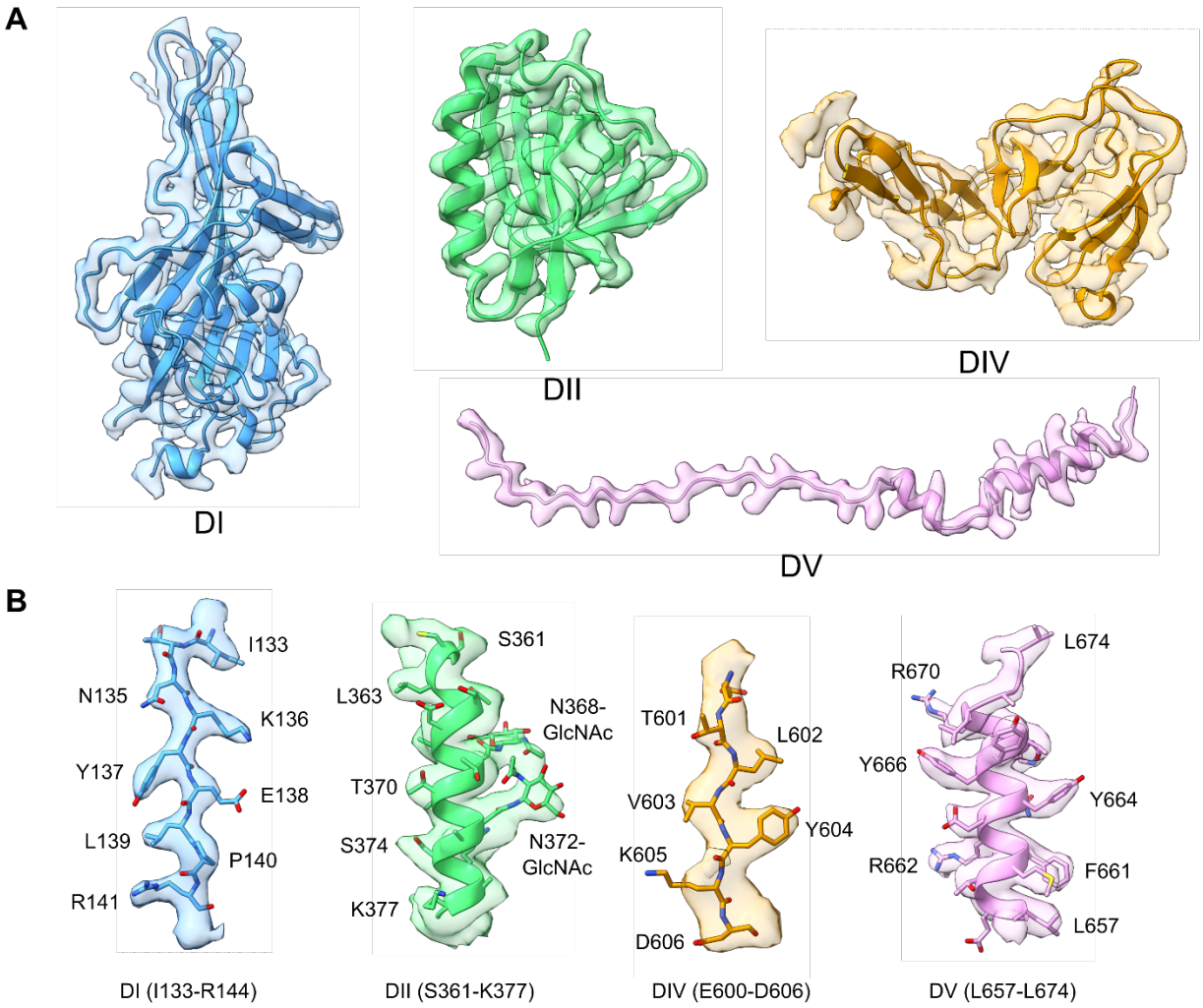


Fig. 3-S5. Regions of the cryoEM density maps superposed with the atomic model of KSHV gB. (**A**) Segmented cryoEM densities of DI, DII, DIV, and DV of KSHV gB ectodomain shown as semitransparent surfaces superposed with the corresponding atomic models in ribbons. (**B**) Segmented cryoEM densities of representative local regions of DI, DII, DIV, and DV. The atomic models of amino acid side chains are shown in sticks.

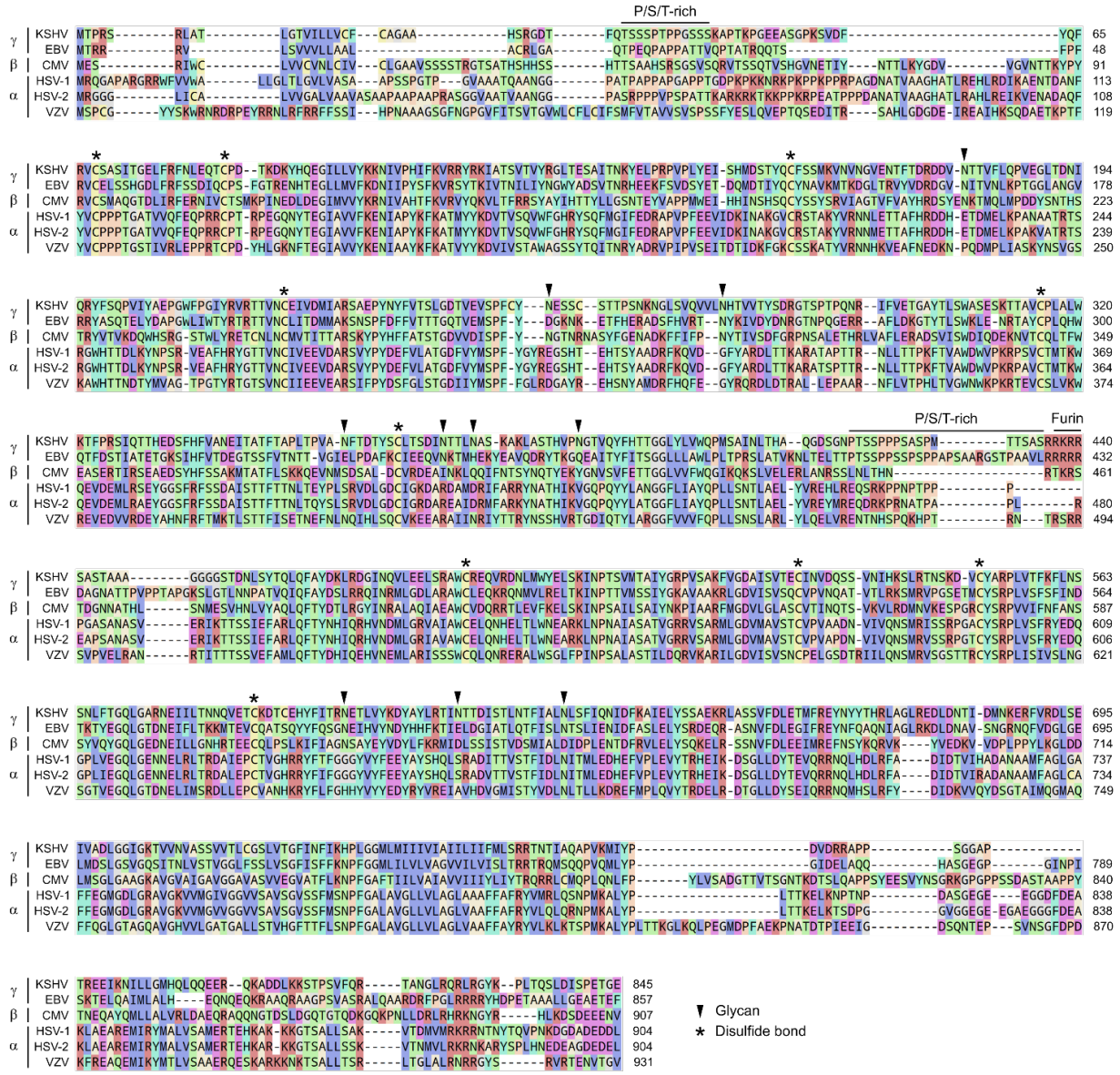


Fig. 3-S6. Sequence alignment of herpesvirus gB. The observed glycosylation sites and conserved cysteines that form disulfide bonds are highlighted with triangles and asterisks, respectively.

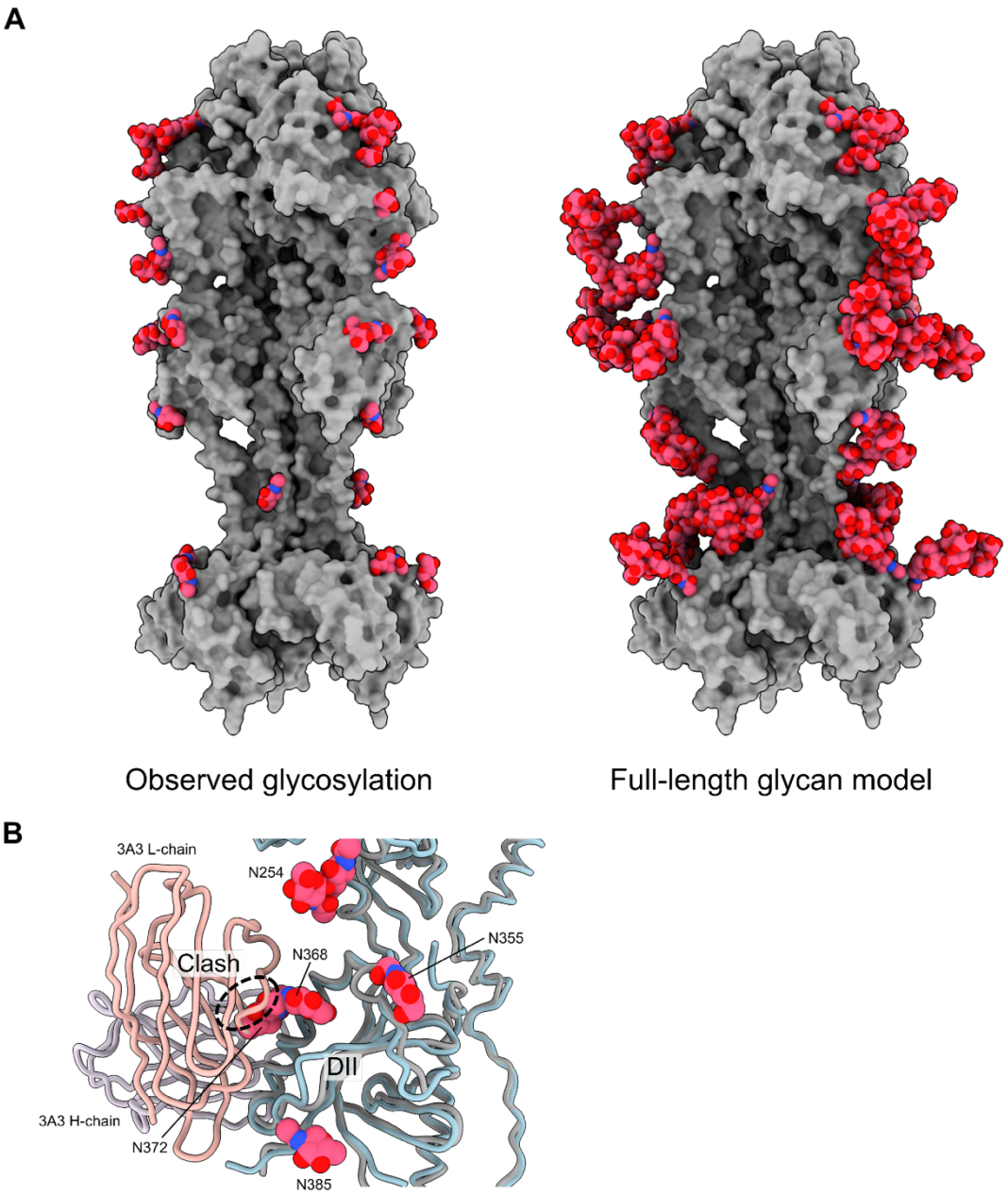


Fig. 3-S7. Insights into the glycosylation of KSHV gB. (**A**) Observed glycosylation (left) and fully-glycosylated model (right) of KSHV gB. High-mannose type glycans (Asn-N-GlcNAc2-Man8) were modeled based on the observed glycans in the cryoEM structure. (**B**) Close-up view of the DII region of KSHV gB (tubes in dark grey) overlayed with EBV gB (pale cyan) bound to neutralizing antibody 3A3 (PDB ID: 7FBI). The glycan at N372 would clash with an antibody targeting the equivalent epitope in KSHV gB.

Supplementary Table 3-S1. Cryo-EM data collection, refinement, and validation statistics

	KSHV gB postfusion (PDB: 9CU4) (EMDB: EMD-45927)
Data collection	
Magnification	105,000
Voltage (kV)	300
Electron exposure ($e^-/Å^2$)	50
Defocus range (μm)	-0.8 to -2.8
Pixel size (Å)	0.86
Symmetry imposed	C3
Initial particle images	4,041,505
Final particle images	64,516
Map resolution (Å)	3.3
FSC threshold	0.143
Map resolution range (Å)	2.6 to 4.4
Refinement	
Initial model used (PDB)	3FVC
Model resolution (Å)	3.6
FSC threshold	0.5
Map sharpening B factor ($Å^2$)	-46.0
No. non-hydrogen atoms	4885
Protein residues	585
Ligands	17
B-factors	
Protein	70.95
Ligand	78.69
R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.822
Validation	
MolProbity score	1.25
Clash score	2.08
Poor rotamers (%)	0.19
Ramachandran plot	
Favored (%)	96.04
Allowed (%)	3.96
Disallowed (%)	0.00

Chapter 4

Structures of Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus virions reveal species-specific tegument and envelope features

James Zhen[a,b,c], Jia Chen[d], Haigen Huang[e], Shiqing Liao[a,b,*], Shiheng Liu[a,b], Yan Yuan[f], Ren Sun[e], Richard Longnecker[d], Ting-Ting Wu[e], Z. Hong Zhou[a,b,c,#]

[a]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA, USA

[b]California NanoSystems Institute, UCLA, Los Angeles, CA, USA

[c]Molecular Biology Institute, UCLA, Los Angeles, CA, USA

[d]Department of Microbiology and Immunology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

[e]Department of Molecular and Medical Pharmacology, David Geffen School of Medicine, UCLA, Los Angeles, CA, USA

[f]Department of Microbiology, School of Dental Medicine, University of Pennsylvania, Philadelphia, PA, USA

#Address correspondence to Z. Hong Zhou, hong.zhou@ucla.edu.

4.1 Abstract

Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) are classified into the gammaherpesvirus subfamily of *Herpesviridae*, which stands out from its alpha- and betaherpesvirus relatives due to the tumorigenicity of its members. Although structures of human alpha- and betaherpesviruses by cryogenic electron tomography (cryoET) have been reported, reconstructions of intact human gammaherpesvirus virions remain elusive. Here, we structurally characterize extracellular virions of EBV and KSHV by deep learning-enhanced cryoET, resolving both previously known monomorphic capsid structures and previously unknown pleomorphic features beyond the capsid. Through subtomogram averaging and subsequent tomogram-guided sub-particle reconstruction, we determined the orientation of KSHV nucleocapsids from mature virions with respect to the portal to provide spatial context for the tegument within the virion. Both EBV and KSHV have an eccentric capsid position and polarized distribution of tegument. Tegument species span from the capsid to the envelope and may serve as scaffolds for tegumentation and envelopment. The envelopes of EBV and KSHV are less densely populated with glycoproteins than those of herpes simplex virus 1 and human cytomegalovirus, representative members of alpha- and betaherpesviruses, respectively. Also, we observed fusion protein gB trimers exist within triplet arrangements in addition to standalone complexes, which is relevant to understanding dynamic processes such as fusion pore formation. Taken together, this study reveals nuanced yet important differences in the tegument and envelope architectures among human herpesviruses and provides insights into their varied cell tropism and infection.

4.2 Importance

Discovered in 1964, Epstein-Barr virus (EBV) is the first identified human oncogenic virus and the founding member of the gammaherpesvirus subfamily. In 1994, another cancer-causing virus was discovered in lesions of AIDS patients and later named Kaposi's sarcoma-associated herpesvirus (KSHV), the second human gammaherpesvirus. Despite the historical importance of EBV and KSHV, technical difficulties with isolating large quantities of these viruses and the pleiomorphic nature of their envelope and tegument layers have limited structural characterization of their virions. In this study, we employed the latest technologies in cryogenic electron microscopy (cryoEM) and tomography (cryoET) supplemented with an artificial intelligence-powered data processing software package to reconstruct 3D structures of the EBV and KSHV virions. We uncovered unique properties of the envelope glycoproteins and tegument layers of both EBV and KSHV. Comparison of these features with their non-tumorigenic counterparts provides insights into their relevance during infection.

4.3 Introduction

Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) are two of the nine currently known human herpesviruses and comprise the gammaherpesvirus subfamily (1, 2). Like other herpesviruses, EBV and KSHV establish lifelong latent infections. However, unlike other herpesviruses, these gammaherpesvirus infections may manifest as cancers (1, 2). EBV is associated with Hodgkin's lymphoma, Burkett's lymphomas, and nasopharyngeal carcinomas (1, 2). Additionally, EBV has recently been identified as a causative agent of multiple sclerosis (3). KSHV is the etiological agent of Kaposi's sarcoma, a highly prevalent cancer in AIDS patients (4, 5), and is the cause of other lymphoproliferative disorders, including primary effusion lymphoma and multicentric Castleman disease (6, 7). Structural determination of EBV and KSHV would improve our understanding of their pathogeneses and inform therapies to relieve their cancer burden.

The alpha-, beta-, and gammaherpesvirus subfamilies are related primarily by their morphological similarity (8). All infectious herpesviruses are composed of a double-stranded DNA (dsDNA) genome packaged in an icosahedral capsid that is surrounded by tegument and further enclosed by an envelope decorated with glycoproteins. The capsid is assembled upon scaffolding proteins from capsomers made up of major capsid protein (MCP) arranged into hexons for the facets and pentons for all but one vertex, which are further stabilized and interlinked by small capsid protein (SCP) and triplex proteins. The unique vertex devoid of a penton is the portal vertex, through which the genome is translocated by the dodecameric portal complex into the capsid by means of a channel running through the portal turret (9). After displacement of the terminase packaging motor from atop the portal turret, the portal channel is plugged by a decameric portal cap to seal the genome inside, thus forming the nucleocapsid, also

known as C-capsid. In addition to C-capsids, herpesvirus infection also produces A- and B-capsids, which are empty capsids and scaffold-containing capsids, respectively (10, 11). These A- and B- capsids represent intermediate or abortive states of genome packaging and are also secreted alongside infectious virions containing C-capsids.

Despite their shared overall morphology, herpesviruses are diverse in their host tropisms and pathologies. Gammaherpesviruses establish latency in B cells, whereas alpha- and betaherpesviruses establish latency in neurons and mononuclear cells, respectively (12). Within the betaherpesvirus subfamily, HHV-6A, HHV-6B, and HHV-7 can also establish latency in T cells (12). To facilitate this diversity, herpesviruses possess additional proteins unique to their individual species, particularly in the tegument and on the envelope (13). The envelope proteins and tegument proteins are crucial for host entry, cell tropism, immune evasion, and conditioning of the cellular environment (14–17). The protein differences between herpesviruses should also be reflected in their structural morphologies.

Early structural studies of herpesvirus structures revealed the architecture and molecular composition of the capsid (18–24). Advances in single particle cryogenic electron microscopy (cryoEM) during the past decade have since enabled structural determination of herpesvirus capsids and capsid-associated tegument components, including those of EBV and KSHV, at near-atomic resolution (9, 25–40). However, individual intact virions are not well characterized due to the pleiomorphic nature of both their envelope and tegument, which prevents averaging.

Cryogenic electron tomography (cryoET) generates 3D reconstructions of individual intact virions and enables observation and description of structural elements on the pleiomorphic envelope and within the tegument. For example, cryoET studies of intact virions have been performed for a human alpha- and betaherpesvirus (41, 42). Tomograms of herpes simplex virus

type 1 (HSV-1) revealed a polar tegument arrangement and glycoprotein distribution around an eccentric nucleocapsid (41). CryoET has also been used to characterize the fusion protein gB in its prefusion conformation on human cytomegalovirus (HCMV) virions (42). Nonetheless, cryoET suffers from physical constraints of increasing sample thickness with increasing tilt angle, which prevents certain views of the sample from being collected. These missing views give rise to a missing wedge of information in Fourier space of the 3D reconstruction that manifests as blurring and streaking in the reconstructed map (43), thereby reducing the interpretability of the map. Recent developments in cryoET data processing, such as deep learning-based missing wedge correction (44), have increased interpretability of tomograms, including direct identification of HCMV portal complexes in enveloped A- and B-capsids (38).

Despite these advances, structural characterizations of human gammaherpesvirus virions in intact states are not available due to the difficulty of culturing them at high concentration. To date, only murine gammaherpesvirus-68 (MHV-68) virions and the naked capsids of KSHV have been interrogated (45, 46). Although EBV virions from the B95-8 system are still difficult to obtain, the development of the inducible iSLK cell line has improved KSHV production (47). This advancement, combined with recent developments in cryoET software (44, 48), enables an architectural description of human gammaherpesviruses that provide new insights into their assembly and immunogenicity.

In this study, we report structures and observations of intact EBV and KSHV determined by deep learning-enhanced cryoET. Using subtomogram averaging and tomogram-guided sub-particle reconstruction, we identify the locations of the EBV and KSHV capsid vertices and the KSHV portal to enable orientational contextualization of the tegument with respect to both the capsid and the envelope. Within the tegument, we observe thin strands along the envelope and

between the capsid and envelope, which may have structural or scaffolding roles. Gammaherpesviruses have two- to threefold fewer presenting proteins on the envelope surface compared to HSV-1 and HCMV. The fusion protein gB is found in triplet clusters, and this arrangement may further inform the mechanism of host membrane fusion.

4.4 Results

4.4.1 Architecture of EBV and KSHV virions

We isolated EBV virions by sucrose gradient centrifugation from infectious media collected from B95-8 cells. Through cryoEM imaging of the sample, we observe scant EBV virions and an abundance of vesicles (Fig. 4-4-1a). Some EBV virions deviate from their common spherical depiction, but the presence of spherical vesicles co-isolated with them suggests that such deviations are not solely attributed to centrifugation. Gammaherpesvirus infection produces both virions and virus-like vesicles (49), the latter of which obstructs data collection and quality by increasing sample thickness and obscuring virions. Nonetheless, we proceeded to collect tilt series of this suboptimal EBV sample for the purpose of high-resolution reconstruction (Fig. 4-1b). In the low-dose images of EBV virion, we observe the leaflets of the bilayer envelope and the contours of the dsDNA genome within the capsid, which indicate that moderate resolution features are present.

In addition to EBV, we isolated and imaged KSHV virions derived from the iSLK cell system (47). In contrast to the EBV sample, the KSHV sample has a much greater abundance of virions (Fig. 4-1c). Although vesicles are also present, they are less obstructive. In tilt series images of KSHV, we can also resolve individual bilayer leaflets and dsDNA contours (Fig. 4-1d). Because of the greater availability of virions in the KSHV sample relative to the EBV

sample, we could collect a more robust dataset for KSHV virions.

We used cryoET to reconstruct the 3D structures of individual EBV and KSHV virions. Initial tomograms were reconstructed through conventional weighted back-projection with a simultaneous iterative reconstructive technique (SIRT)-like filter (Fig. 4-2a,e). In these conventional tomograms, we observe the capsid, tegument, and envelope of EBV and KSHV. The trace of the dsDNA genome within the capsid can be observed in some virions (Fig. 4-S1). We are able to resolve the leaflets of the bilayer envelope, but the envelope appears distorted along the Z direction due to the missing wedge problem inherent to cryoET (Fig. 4-2a,e) (43). Features on the envelope and within the tegument layers are also difficult to discern due to poor contrast and high noise. For repeating particles of like features, subtomogram averaging is employed to overcome the missing wedge problem while simultaneously increasing the signal-to-noise ratio and resolution of the resulting map (50). However, herpesvirus tegument and envelopes are pleiomorphic, so subtomogram averaging cannot be used to remedy missing wedge artifacts and low contrast.

To overcome these problems and to improve interpretability of our herpesvirus tomograms, we applied IsoNet, a deep learning-based software for the correction of missing wedge artifacts and denoising (44) (Fig. 4-S2). Representative IsoNet-processed tomographic slices of an EBV and KSHV virion show intact viral envelopes (Fig. 4-2b,f). In spite of the remaining noise in the map, we can identify a mostly smooth surface corresponding to the viral envelope (Fig. 4-2c,g). Small protrusions connected to the smooth envelope exterior are considered to be envelope proteins. In a clipped view of EBV and KSHV segmentation (Fig. 4-2d,h), the layered architecture of the nucleocapsid, tegument, envelope, and envelope proteins are demarcated. Individual capsomeres and envelope proteins are distinguishable in the

tomographic slices and the surface rendering. Noticeably, the bulk of the tegument is offset to one side of the virion, as previously observed in HSV-1 and HCMV (38, 41). Because of pixel binning required by IsoNet, the gap between adjacent dsDNA can no longer be resolved, so the contour of the genome cannot be traced in those reconstructions (Fig. 4-S2) (25). Although the KSHV and EBV reconstructions are of insufficient resolution for atomic model building, they nevertheless enable comparison between the two species of human gammaherpesvirus.

4.4.2 Subtomogram averaging of nucleocapsids reveals distinct capsid-associated features in KSHV and EBV

To investigate the structure of the nucleocapsids, we performed subtomogram averaging of the EBV and KSHV icosahedral capsids (Fig. 4-3a,c). The EBV and KSHV capsids were reconstructed with icosahedral symmetry to ~24 Å and 17 Å resolution, respectively (Fig. 4-S3, S4). At these resolutions, the hexons, pentons, and triplexes are observed. Both of our EBV and KSHV icosahedral capsid reconstructions are similar to their single particle cryoEM maps from previous studies and can be fitted with the corresponding atomic model (31, 32, 40). Capsid vertex-specific component (CVSC) densities stemming from the triplexes adjacent to the pentons are observed in both the EBV and KSHV capsids (51–54), and their handed arrangement around the vertex validates the handedness of our subtomogram averaging reconstructions.

Viewing the KSHV capsid reconstruction at low contours, we observe additional densities spanning hexons at the 2-fold interface to hexons at the adjacent 3-fold axes between triplexes Tb and Te (Fig. 4-3c). These densities form a network that bridges the capsid facets. When models of the KSHV asymmetric unit containing MCP, SCP, and triplex proteins are fitted into the map, this density is unaccounted for and presumably belongs to some capsid-associated tegument protein that contacts loops of the adjacent MCPs at residues ^{548}PGAR^{551} (Fig. 4-3g).

This unfilled density on the KSHV capsid has previously been observed and was attributed to low-affinity secondary binding sites of CVSC protein ORF32 (55). The density is also present in other KSHV capsid cryoEM maps from intact virions (9, 40). However, it is absent in our cryoEM map of the naked KSHV capsid devoid of tegument (Fig. 4-3e,h). Additionally, this density is absent in our EBV capsid reconstruction (Fig. 4-3a) and previous cryoEM capsid reconstructions for EBV, HSV-1, and HCMV (29, 31, 32, 35, 37, 38), which possess ORF32 homologs, thus suggesting that ORF32 does not bind the capsid between Tb and Te. Combined with recent findings on herpesvirus CVSC structure and assembly (9, 38), this density between Tb and Te likely does not belong to ORF32, and the component is specific to KSHV tegument.

To analyze the 5-fold symmetric capsid vertex, the localized reconstruction method for cryoEM was adapted to cryoET (29, 56). The icosahedral capsid reconstruction was used to calculate the locations of the capsid vertices, which were then locally extracted as discrete subvolumes for subtomogram averaging. Through this method, we determined localized reconstructions of the EBV and KSHV capsid vertex with 5-fold symmetry at ~21 Å and 7.1 Å resolution, respectively (Fig. 4-3b,d; Fig. 4-S3,S4). Prior published atomic models of the EBV and KSHV asymmetric units fit their respective cryoET capsid vertex reconstruction (31, 32, 40), and distinct strands of the underlying dsDNA genome are observed. Additional densities around the capsid are visible in both the EBV and KSHV virion vertex but not in the naked KSHV capsid vertex (Fig. 4-3f), which indicates that these densities belong to the tegument.

Unlike empty capsids whose portal complexes are apparent (20, 33), the portal vertex of nucleocapsids cannot be directly observed due to obfuscation by the enclosed genome. To determine the location of the portal in the nucleocapsids, we performed classification on the capsid vertex reconstructions to separate the portal vertices from the penton vertices. Although

we were unable to distinguish the EBV portal vertex from EBV penton vertex using this approach, 3D classification successfully distinguished the KSHV portal vertex (9.8 Å resolution) from the KSHV penton vertex (7.0 Å resolution) (Fig. 4-4a,b; Fig. 4-S4). The KSHV portal vertex reconstruction contains 554 out of the 7,177 total extracted capsid vertices (7.7%), which is in line with the expected one portal vertex out of 12 capsid vertices (8.3%). In a clipped side view of the penton vertex reconstruction, we observe clear densities for the penton MCPs and underlying dsDNA genome (Fig. 4-4c). We cannot observe the penton MCPs in the KSHV portal vertex, and we instead see densities resembling the pORF19 portal cap decamer and pORF43 portal complex dodecamer (Fig. 4-4d) (9, 57). The clip turret between the portal cap and portal complex is not well defined, likely due to its plasticity (9).

Although the penton vertex cryoEM map is of higher resolution than the portal vertex map and sufficient for fitting secondary structure features (Fig. 4-4e), the density for the CVSC in the penton map is poor compared to that in the portal map (Fig. 4-4a,b,f). This difference in CVSC density suggests that the occupancy of CVSC at penton vertices is lower than at the portal vertex. Indeed, a prior cryoEM reconstruction without enforcing symmetry shows that CVSC occupancy at vertices is directional with respect to the portal (9). Further classification of the portal vertex reveals heterogeneity of the pORF19 portal cap (Fig. 4-4g). Between the two classes, the portal cap binds the CVSC with both different rotation and inverted orientations (Fig. 4-4g,h). In mature HCMV virions, variable configurations of portal cap association have been observed, including an inversion of the portal cap (38). Similar variability in portal cap configuration in KSHV and HCMV suggests a conserved mechanism of genome packaging and retention.

4.4.3 Polarity and structures of the tegument layer

Due to the pleiomorphic nature of the tegument layer, tegument features cannot be averaged, so observations must be made from individual tomograms. We sorted tegument features into three previously described categories: inner tegument, outer tegument, and membrane-proximal tegument (58). To provide further context for tegument observations, we mapped back the EBV and KSHV capsid reconstructions into the tomograms (Fig. 4-5a). From the positional information obtained from localized reconstruction of the capsid vertices, we also identify the locations of the capsid vertices in the tomograms. Since we are able to classify the portal vertices from penton vertices in KSHV, we also make use of the portal vertex to establish the orientation of the capsid within KSHV virions. Thus, capsid-relative tegument observations can be interpreted in relation to the capsid vertices for EBV and both the penton vertices and portal vertex for KSHV.

For the majority of EBV and KSHV virions, the tegument is clustered to one side of the capsid, causing the capsid to be eccentrically located within the envelope (Fig. 4-S5). This capsid eccentricity has previously been observed in HSV-1 and HCMV (38, 41) and has been recapitulated in our cryoET of HSV-1 and HCMV (Fig. 4-S5a). For KSHV, whose portal vertices we could determine, we calculated a "portal-capsid-tegument angle" (PCT angle), a previously used proxy for quantification of the angular distribution of tegument relative to the portal vertex (Fig. 4-S5b) (38). For KSHV, across 265 virions, 149 virions (56.2%) have a bulk of tegument within 40° of its portal vertex. This differs from previously published measurements for HCMV, where 38 out of 49 virions (77.6%) have a PCT angle of less than 40° (38). Although asymmetric tegumentation is a conserved feature of herpesvirus assembly, direction towards the portal vertex appears to vary, possibly due to different binding affinities of species-specific tegument components.

Distinct globular densities of ~20 nm by 15 nm by 10 nm are observed in the outer tegument near certain EBV and KSHV capsid vertices (Fig. 4-5b). More specifically for KSHV, these globular densities have an increased propensity at the portal vertex. The globular densities are offset from the center of the vertex, which suggests that these densities may belong to tegument protein species that bind to CVSC or other capsid-associated inner tegument proteins. In addition to contacts between the tegument and capsid, noticeable areas of low density are present between the tegument and capsid (Fig. 4-5b). These areas of low density indicate low tegument association or lack thereof with the capsid.

Along the virion envelope of both EBV and KSHV, a disjointed layer of membrane-proximal tegument circumnavigates the interior (Fig. 4-5c). From an orthogonal view, the densities present as long strands arranged from end to end discontinuously (Fig. 4-5d) This layer is composed of densities with a diameter of ~6 nm and varying lengths up to ~50 nm. The diameter and filamentous nature of these densities suggest that they may be actin, which is found near secondary envelopment sites and could feasibly be incorporated within the virion during secondary envelopment (59–61). Similar actin-like densities have also been observed in HSV-1 (41). Previous mass spectrometry studies of isolated herpesvirus virions have determined the presence of actin within the samples (49, 58, 62–64). A recent crosslinking-mass spectrometry of internally crosslinked intact HCMV virions shows that gamma actin crosslinks with UL94 (EBV BGLF2; KSHV ORF33) and UL99 (EBV BBLF1; KSHV ORF38) (58), conserved tegument proteins that facilitate secondary envelopment and are clustered with other membrane-proximal tegument proteins (65), which supports the presence of actin near the envelope. These filamentous components within the tegument may function as structural reinforcement of the envelope and in bridging the nucleocapsid to the host cortical actin during cellular entry.

Tegument also tends to congregate beneath regions of the envelope that contain envelope proteins, indicating interactions between the tegument and envelope proteins. This is most noticeable at protrusions with overall low proteinaceous tegument densities except for the regions containing clusters of envelope proteins, which instead have large densities below them (Fig. 4-5e). The large densities below clusters of envelope proteins are too large to be solely attributed to the endodomains of the envelope proteins, so they are likely composed of tegument proteins associated with the endodomains of envelope proteins and comprise the membrane-proximal tegument and outer tegument layers.

We observe thin densities spanning between the capsid and the envelope (Fig. 4-5f). The densities are ~4 nm width and 15-30 nm long and the capsid end primarily localizes near penton vertices. Long proteins stemming from the nucleocapsid have previously been observed in HSV-1 (pUL36) and HCMV (pUL48) (38, 66, 67). These proteins are homologous to BPLF1 and ORF64 in EBV and KSHV, respectively. Crosslinking MS experiments in HCMV have indicated tegument species that connect to both the capsid and the envelope (58), so similar spanning tegument proteins may be present in other herpesviruses, including EBV and KSHV. The major spanning protein in HCMV (UL32/pp150) does not have an equivalent in other herpesviruses, which suggests that EBV and KSHV may possess unique proteins that share a similar function as a scaffold for recruitment of the herpesvirus layers.

4.4.4 Morphology of the envelope and volume of the virion

EBV has an average volume and surface area of 2.6×10^6 nm^3 and 1.2×10^4 nm^2, respectively, and KSHV has an average volume and surface area of 2.7×10^6 nm^3 and 1.1×10^4 nm^2, respectively (Fig. 4-S5c,d). These average volumes and surfaces areas are similar to those of HSV-1 and HCMV. Both EBV and KSHV envelopes are generally round with deviations from sphericity

caused by the asymmetric association of tegument around the capsid. Additionally, some virions have protrusions of varying prominence (Fig. 4-5e,f). Some protrusions are filled with tegument while others are devoid of proteinaceous density. Although we cannot exclude the possibility that protrusions result from the purification process, previous studies of herpesvirus virions within the cytoplasm of infected cells also possess them (46, 68–70), which suggests a physiological morphology. Virions isolated by ultracentrifugation remain infectious, indicating that structural studies of herpesvirus virions prepared in this manner retain features that reflect infectivity.

Some regions of the envelope with spherical curvature do not have underlying membrane-proximal tegument (Fig. 4-5b,f). This typically occurs when the capsid is in direct contact with the envelope and occludes space for a tegument layer. Noticeably, the void of tegument also presents as a distinct gap between the envelope and the capsid as though the capsid has been pulled away from the envelope (Fig. 4-5f). The envelopes around these voids retain their curvatures, which suggests that these regions are still internally supported. Serendipitously, the lack of bulk tegument in these void regions permit clearer observation of the tegument proteins that remain, including the thin densities that span the capsid and the envelope (Fig. 4-5f). Envelope proteins appear on the opposing side of where these thin densities contact the envelope, which suggests that some species of envelope proteins function as anchors for the thin densities to bind. Due to their presence outside of the bulk tegument, we infer that these thin densities bind the capsid independently of the other tegument species and possibly function in tethering the membrane to the capsid.

In addition to intact virions, we also observe EBV and KSHV particles that deviate from the conventional depiction of herpesvirus. These include enveloped capsids with invaginations and with much larger volume than average (Fig. 4-S6). We also observe envelopes containing

two capsids. These indicate either potential fusion events between virion and vesicle or between virion and virion. We observe a potential intermediate state of virion fusion with another virion as their envelopes are connected by a thin pore (Fig. 4-S6), which suggests that virion-virion fusion may occur. Alternatively, an envelope containing two capsids may be the result of mispackaging of two capsids simultaneously during secondary envelopment (Fig. 4-S6). These deviations from the "norm" emphasize the irregularities in herpesvirus assembly.

4.4.5 Herpesvirus envelope protein organization

To characterize the envelope proteins of EBV and KSHV, we quantified them and noted their arrangements on the envelope (Fig. 4-6). Notably, both EBV and KSHV appear sparsely decorated with envelope proteins compared to HSV-1 and HCMV (Fig. 4-6a). Counting external densities of at least 5 nm protruding from the envelope surface, EBV and KSHV have two- to threefold fewer presenting envelope proteins than HSV-1 and HCMV (Fig. 4-6b). Because envelope proteins are integral for infection, a greater density of them on the envelope should increase the likelihood of encountering a host receptor and initiating membrane fusion for cell entry.

We noticed that envelope proteins tended to cluster on the herpesvirus envelope. This clustering has previously been observed in HSV-1 (71). By taking advantage of the sparsity of envelope proteins on KSHV and EBV virions visualized with the aid of deep learning-enhanced isotropic reconstructions, we better resolve the envelope protein clusters and identify their arrangements as triplets, sets of three envelope proteins (Fig. 4-6a). We performed subtomogram averaging of the envelope proteins in these triplet clusters of KSHV (Fig. 4-6c). Further classification of these particles produces only similar reconstructions, which indicates that the population of particles within these triplets is homogeneous. Subtomogram averages of envelope

proteins from triplet clusters of EBV, HSV-1, and HCMV are each similar to that of KSHV. The subtomogram average has a distinctive shape that resembles postfusion gB as determined by previous studies (42, 71–73), and the atomic model of postfusion gB fits the density well (74) (Fig. 4-6c). The pyramidal prefusion state is shorter and wider than the postfusion state and does not fit in the columnar density (72). The next largest viral envelope protein complex, gHgL, does not fit this density (75). Although the other herpesvirus glycoproteins have not been structurally determined, their molecular weights suggest that they are too small to fully account for this density. Indeed, AlphaFold structural predictions of these other glycoproteins do not fit (76). Together, these observations indicate that postfusion gB comprises the triplet clusters. Envelope glycoproteins tend to cluster on protrusions, but gB units along these protrusions tend to be in the postfusion conformation. This preference for the postfusion state can be attributed to the increased curvature at these regions inducing conformation change (77, 78). Alternatively, the prefusion state of gB may be present but less identifiable than the postfusion state due to its less distinct shape and the low contrast of the tomograms. Additionally, the gB members of triplets are angled relative to the envelope like HIV-1 gp160 spike and SARS-CoV-2 spike (79, 80). Like the other spike proteins, the angle of gB may facilitate exploration of its envelope to engage with its host receptor (80).

4.5 Discussion

CryoET reconstructions of EBV and KSHV with spatially determined capsids enable contextualized observations of their pleiomorphic tegument and envelope. Both EBV and KSHV share similar overall morphology for the dsDNA genome and capsid, but we observe differences envelope protein distribution. The tegument is not fully amorphous as commonly depicted and instead contains distinct structural features (Fig. 4-5). Tegument-mediated contacts between the

envelope and the capsid suggests an interlayer dynamic for the tegumentation and envelopment pathways during herpesvirus assembly. The envelope proteins are further organized into distinct clusters that could mediate function, and their relatively low prevalence on EBV and KSHV could explain their infectivity (Fig. 4-6).

Intracellular herpesvirus virions have been observed with deviations from sphericity, including protrusions (46, 81), which suggests that these perceived flaws may instead be a feature. Protrusions and other deviations from sphericity increase the surface area of an envelope for a given internal volume, which increases the potential interaction sites with a host cell. We notice that deviations from sphericity vary between herpesvirus species (Fig. 4-S5a), which suggests that the pliability of the envelope and its composition may differ between them. The tegument may also be involved with maintaining the shape of the envelope by anchoring it to the capsid. This is most noticeable in HCMV (42), where the envelope protrudes when underlying tegument is lacking (Fig. 4-5). The outer tegument may not be tightly associated with the capsid in order to disperse inside the cell and prepare the host environment for viral expression and assembly.

In prior cryoET studies of HSV-1 infection, envelope proteins are not observed to be recruited to the entry site, suggesting that few gB units are necessary to catalyze entry (82). In our tomograms, those few gB units appear to be clustered on the envelope into functional units to facilitate entry (Fig. 4-6a). The clustering of gB may aid their membrane fusion function, which has been observed in other fusogenic proteins. SNARE proteins require three contact points to form a pore (83). The homologous vesicular stomatitis virus G protein requires a certain number of copies clustered together to trigger hemifusion (84). The assembly of these gB clusters remains to be determined but may occur through directed localization by trans-Golgi network,

lipid rafts, tegument, or cytodomain of gB (85–87).

The different herpesvirus species have different distributions of proteins displayed on the envelope (Fig. 4-5a,b), which may explain the differing neutralizing antibody profiles between herpesviruses (88–98). Gammaherpesvirus have low immunogenic response to gB as measured by anti-gB from virion immunization or patient derived sera (92, 99, 100). The herpesvirus tomograms from this study suggests that the low prevalence of gB on the surface of EBV and KSHV may be the cause of the weak gB response by the immune system. HSV-1 and HCMV may have higher anti-gB response from sera than their gammaherpesvirus relatives due to this greater gB presence on their envelopes. Although anti-gB antibodies have neutralizing activity against EBV and KSHV like they do against HSV-1 and HCMV (101, 102), their generation via attenuated EBV or KSHV virions or virus-like particles may not be as effective compared to similar processes of neutralizing antibody generation for HSV-1 or HCMV. As such, EBV or KSHV gB overexpressed and displayed on vesicles or nanoparticles may better elicit production of antibodies with broad neutralizing activity than attenuated virions.

The most recent 3D structural characterization of an intact gammaherpesvirus virion is a cryoET study of MHV-68 (46), a model system for EBV and KSHV due to its relative ease of culture (103). In MHV-68, a distinct boundary just above the capsomere towers demarcates the inner tegument layer and distinguishes it from the outer tegument layer (46). In our tomograms of EBV and KSHV, this boundary manifests instead as a lighter region of spaces between the capsomeres and a darker region of the bulk tegument (Fig. 4-5), blurring the distinction between inner and outer tegument. Occasionally, the tegument bulk infiltrates the space between capsomeres, suggesting an intermingling between the capsid-associated inner tegument and the outer tegument, which is supported by crosslinking mass spectrometry of HCMV (58).

Additionally, a previous comparative study between HCMV and its animal model, murine cytomegalovirus (MCMV), shows that the inner tegument binding pattern is different between the cytomegaloviruses that infect different species (104). Differences in inner tegument organization would affect that of the adjoining outer tegument. Analogously, MHV-68 and the human gammaherpesviruses may also have species-dependent differences in tegument organization. The differences in visualization could be attributed to a combination of intrinsic structural differences between herpesvirus species and improved tomogram-level resolution due to hardware and software advancements since the previous study.

Although cryoET has enabled observation of structures and features of EBV and KSHV, identification of individual proteins is inhibited by the limited resolution of the tomograms. The resolution is primarily limited by poor signal-to-noise ratio from the need to work around radiation damage of the sample (105). This problem is further exacerbated by the inherent thickness of the sample to accommodate the large virions (106). The volta phase plate has been used to obtain higher contrast (42), but they contradictorily reduce the high-resolution signal (107). Deep learning enhancement currently down samples the tomograms and prevents recovery of information for small features. Future developments in cryoET technology are required to increase attainable resolution of thick samples at the tomogram level in order to truly understand herpesvirus architecture and underlying molecular sociology.

4.6 Materials and Methods

4.6.1 EBV culture

7.5×10^5/mL of B95-8 cells (ATCC CRL-1612) infected with EBfaV-GFP (recombinant EBV reporter virus expressing GFP) (108) were grown in 100 mL RPMI 1640 (Corning 10-040-CV)

with 10% FBS, penicillin-streptomycin, and 30 ng/mL 12-*O*-tetradecanoylphorbol-13-acetate (TPA; Sigma) at 37°C and 5% CO_2 for 4 days. The supernatant was collected by centrifugation at 1,500 RPM for 10 minutes.

4.6.2 KSHV culture

KSHV virions were prepared as previously described with slight modification from iSLK-KSHV-BAC16 cells, received as a gift from Dr. Jae U. Jung (47, 55). In brief, iSLK-KSHV-BAC16 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Corning 10-017-CV) supplemented with 10% fetal bovine serum (FBS; 35-010-CV) and 100 U/mL of penicillin-streptomycin, 1mg/mL puromycin (Invivogen ant-pr-1), 250 μg/mL G418 (Invivogen ant-gn-1), and 1,200 mg/mL hygromycin B (Invivogen ant-hg-1). Cells were cultured to 80-90% confluency in 24 15 cm tissue culture dishes. Then, KSHV lytic replication and virion production were induced by treatment with 1 mM sodium butyrate and 5 μg/mL tetracycline for five days, after which 720 mL of media containing KSHV virions were collected.

4.6.3 HCMV culture

HCMV virions were prepared as previously described with slight modification (37). Human fibroblast MRC-5 cells (ATCC CCL-171) were seeded on 18 T175 flasks in Eagle's Minimum Essential Medium (EMEM; ATCC 30-2003) supplemented with 10% FBS (R&D Systems S11150H) and 100 U/mL of penicillin-streptomycin (Gibco 15140122) and cultured until they reached ~90-100% confluency. The cells were then infected with HCMV (AD169 strain, ATCC VR-538). The growth media were replaced 3 days after infection, and the media containing secreted HCMV virions were collected at 7 days post-infection.

4.6.4 HSV-1 culture

HSV-1 virions were prepared as previously described with slight modification (26).
Vero cells (ATCC CCL-81) were seeded on 12 T175 flasks in DMEM (ATCC 30-2002) supplemented with 10% FBS and 100 U/mL of penicillin-streptomycin. At 100% confluence, the cells were infected with HSV-1 (KOS strain, ATCC VR-1493). The media containing secreted HSV-1 virions were collected 3 days post-infection.

4.6.5 Herpesvirus virion isolation

Herpesvirus virions were isolated as previously described with slight modification (109). Infectious media was clarified by centrifugation at 10,000 × g for 20 min at 4 °C twice. Virions were then pelleted using an SW28 rotor (Beckman) at 80,000 × g for 1 h at 4 °C. The supernatant was discarded and 50 µl of cold phosphate-buffered saline (PBS) was added to the pellet in each tube. The pellets were allowed to soften for 16 h on ice in a 4°C cold room before gentle resuspension by tapping. About 0.7 ml of combined pellet suspension from all ultracentrifuge tubes were loaded on 15–50% linear gradient of sucrose made in PBS and spun in SW41 (Beckman) rotor at 80,000 × g for 1 h at 4 °C. About 1 ml of the visible virus-containing band was collected into a new tube and mixed with 10 ml of sterile PBS solution in order to dilute sucrose. EBV did not have a visible band and was instead fractionated in 1 mL aliquots based on the band locations of other herpesviruses. The virions were concentrated at 80,000 × g for 1 h at 4 °C and 20 to 40 µl of PBS was added depending on the pellet size. After softening and gentle resuspension, the isolated virions were recovered and immediately used for cryoET sample preparation without freezing.

4.6.6 CryoET sample preparation and image acquisition

Virion samples were screened on an FEI Tecnai TF20 equipped with a 4k × 4k TVIPS F415MP

CCD detector. To prepare EM grids for cryoET imaging, 3 µl of diluted virions were mixed with 5- or 10-nm gold fiducials and applied to Quantifoil R2/1 Cu 200 grids and plunge-frozen in liquid ethane-propane mixture using a Vitrobot mark IV (FEI Thermo Fisher Scientific) set to 100% humidity, 4 °C, and blot time of 5 s (110).

Imaging was performed with a Titan Krios G1 (FEI Thermo Fisher Scientific) operated at 300 kV, and movies were recorded using a Gatan K3 detector with an energy filter slit width of 20 eV. Tilt series were collected from -60° to 60° in 3° increments using a dose-symmetric scheme (111) using SerialEM (112) at 42,000 × magnification (1.044 Å/pixel super-resolution) and a target defocus of -5 µm or at 64,000 × magnification (0.690 Å/pixel super-resolution) and a target focus of -4 µm with a total of dose of 120 e-/Å^2. Individual tilts were recorded with constant beam intensity in 10 frames. An HCMV dataset was collected with a Volta phase plate conditioned to induce an initial phase shift of π/3 and a defocus target of -0.6 µm. Data collection statistics are listed in Table 4-S1.

4.6.7 Tomogram reconstruction and subtomogram averaging

Each movie stack was drift-corrected using MotionCor2 (113) and defocus values were determined with CTFFIND4 (114). Tilt series alignment via gold fiducials and tomogram reconstructions were performed in IMOD with a binning factor of 4 (115). IsoNet isotropic reconstructions with a binning factor of 8 were used for particle picking and observations of non-repeating features (44). Capsid reconstructions of EBV and KSHV were performed in RELION-4 with a binning factor of 5 and 6, respectively (48). Tomogram-guided sub-particle reconstruction of the EBV and KSHV vertices at binning factor 2 was done as previously described for single particle cryoEM with slight modification for tomography data (29, 56). The formula for defocus change was instead used to shift the Cartesian position z. RELION-4

accounts for the depth-of-focus problem by using a defocus gradient for tomogram CTF correction (48). Classification of the EBV and KSHV vertices was performed with a cylindrical local mask focused over the 5-fold symmetry axis. For KSHV, one class with 7.7% of the particles contained portal complex features. Individual portal particles were mapped back in their respective tomograms to verify that each parent capsid contained at most one portal. Redundant particles were removed based on _rlnMaxValueProbDistribution score. No distinct portal class was identified for the EBV particles. Subtomogram averaging of clustered envelope proteins was performed in PEET with the tilt orientation of particles coarsely determined using stalkInit (116). Visualization, segmentation, and measurements were performed in IMOD, ChimeraX, and ArtiaX (115, 117, 118).

4.6.8 Infection assay

Infectious media was filtered using a 0.45 μM membrane (Millipore SE1M003M00), aliquoted into 1 mL portions, and stored at −80 °C for infection experiments. For testing infectious viability after ultracentrifugation, aliquots were centrifuged at 80,000 × g for 30 minutes at 4°C, and the pellets were resuspended in 100 μL of 10% FBS DMEM prior to infection. HEK293 cells (7.5×104 cells per well) were seeded in 100 μL of 10% FBS DMEM medium in a 48-well plate and infected on the second day with the indicated amounts of EBV, KSHV, and HSV viruses in 100 μL of 10% FBS DMEM. Forty-eight hours post-infection, the cells were either visualized and captured using an EVOS fluorescence microscope for GFP reporter signal. GFP was under control of CMV, EF1α, and Egr-1 promoter in EBV, KSHV, and HSV-1, respectively (47, 108, 119).

4.6.9 KSHV capsid isolation and cryoEM

KSHV capsid isolation, data collection, and data processing was performed as previously described (24).

4.7 Data Availability

CryoEM maps have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-46986, EMD-46987, EMD-46988, EMD-46989, EMD-46990, EMD-46991. The data that support this study are available from the corresponding author upon reasonable request.

4.8 Acknowledgements

We thank Ana L. Alvarez-Cabrera and Zhu Si for assistance with HSV-1 sample preparation and imaging, Yun-Tao Liu for assistance with IsoNet, and Jonathan Jih and Alexander Stevens for insightful discussion. This research was supported in part by grants from the National Institutes of Health (R01DE025567 to Z.H.Z and T.-T.W., R21AI175798 to R.L., R01DE027901 to Y.Y.), grant IRG-15-173-21 (J.C.) from the American Cancer Society, a supplement from P30CA060553 (R.L.), and the Third Coast Center for AIDS Research pilot award (J.C.). J.Z. acknowledges support from an Interdisciplinary Training in Virology and Gene Therapy training grant (NIH 5T32AI060567). We acknowledge the use of resources in the Electron Imaging Center for Nanomachines supported by UCLA and grants from NIH (S10RR23057, S10OD018111, and U24GM116792) and NSF (DBI-1338135). We also acknowledge use of resources from the UCLA AIDS Institute, the James B. Pendleton Charitable Trust, and the McCarthy Family Foundation.

4.9 Contributions

Z.H.Z., R.L., T.-T.W., R.S., and Y.Y. initiated and supervised the research. J.Z. prepared samples, recorded cryoEM and cryoET data, processed and interpreted the data, and prepared the manuscript. J.C. cultured EBV samples and performed infection assays. H.H. cultured KSHV samples. S. Liu prepared HSV-1 samples and obtained the HSV-1 tomograms. S. Liao processed KSHV data. All authors reviewed and edited the manuscript.

4.10 Conflict of Interest

The authors declare that they have no conflicts of interest.

4.11 References

1. Schiller JT, Lowy DR. 2021. An Introduction to Virus Infections and Human Cancer. Recent Results Cancer Res 217:1–11.

2. Wen KW, Wang L, Menke JR, Damania B. 2022. Cancers associated with human gammaherpesviruses. The FEBS Journal 289:7631–7669.

3. Bjornevik K, Cortese M, Healy BC, Kuhle J, Mina MJ, Leng Y, Elledge SJ, Niebuhr DW, Scher AI, Munger KL, Ascherio A. 2022. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. Science https://doi.org/10.1126/science.abj8222.

4. Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, Knowles DM, Moore PS. 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 266:1865–1869.

5. Cesarman E, Damania B, Krown SE, Martin J, Bower M, Whitby D. 2019. Kaposi sarcoma. 1. Nature Reviews Disease Primers 5:1–21.

6. Cesarman E, Chang Y, Moore PS, Said JW, Knowles DM. 1995. Kaposi's Sarcoma–Associated Herpesvirus-Like DNA Sequences in AIDS-Related Body-Cavity–Based Lymphomas. New England Journal of Medicine 332:1186–1191.

7. Soulier J, Grollet L, Oksenhendler E, Cacoub P, Cazals-Hatem D, Babinet P, d'Agay M, Clauvel J, Raphael M, Degos L, Sigaux F. 1995. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease [see comments]. Blood 86:1276–1280.

8. Davison AJ. 2007. Overview of classification, p. . *In* Arvin, A, Campadelli-Fiume, G, Mocarski, E, Moore, PS, Roizman, B, Whitley, R, Yamanishi, K (eds.), Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge University Press, Cambridge.

9. Gong D, Dai X, Jih J, Liu Y-T, Bi G-Q, Sun R, Zhou ZH. 2019. DNA-Packing Portal and Capsid-Associated Tegument Complexes in the Tumor Herpesvirus KSHV. Cell 178:1329-1343.e12.

10. Tandon R, Mocarski ES, Conway JF. 2015. The A, B, Cs of Herpesvirus Capsids. Viruses 7:899–914.

11. Deng B, O'Connor CM, Kedes DH, Zhou ZH. 2008. Cryo-electron tomography of Kaposi's sarcoma-associated herpesvirus capsids reveals dynamic scaffolding structures essential to capsid assembly and maturation. Journal of Structural Biology 161:419–427.

12. Cohen JI. 2020. Herpesvirus latency. J Clin Invest 130:3361–3369.

13. Mocarski Jr. ES. 2007. Comparative analysis of herpesvirus-common proteins, p. . *In* Arvin, A, Campadelli-Fiume, G, Mocarski, E, Moore, PS, Roizman, B, Whitley, R, Yamanishi, K (eds.), Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge University Press, Cambridge.

14. Connolly SA, Jardetzky TS, Longnecker R. 2021. The structural basis of herpesvirus entry. 2. Nature Reviews Microbiology 19:110–121.

15. Lee H-R, Lee S, Chaudhary PM, Gill P, Jung JU. 2010. Immune evasion by Kaposi's sarcoma-associated herpesvirus. Future Microbiol 5:1349–1365.

16. Sathish N, Wang X, Yuan Y. 2012. Tegument Proteins of Kaposi's Sarcoma-Associated Herpesvirus and Related Gamma-Herpesviruses. Frontiers in Microbiology 3.

17. Murata T. 2023. Tegument proteins of Epstein-Barr virus: Diverse functions, complex networks, and oncogenesis. Tumour Virus Research 15:200260.

18. Schrag JD, Prasad BVV, Rixon FJ, Chiu W. 1989. Three-dimensional structure of the HSV1 nucleocapsid. Cell 56:651–660.

19. Booy FP, Newcomb WW, Trus BL, Brown JC, Baker TS, Steven AC. 1991. Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. Cell 64:1007–1015.

20. Newcomb WW, Trus BL, Booy FP, Steven AC, Wall JS, Brown JC. 1993. Structure of the Herpes Simplex Virus Capsid Molecular Composition of the Pentons and the Triplexes. Journal of Molecular Biology 232:499–511.

21. Zhou ZH, Dougherty M, Jakana J, He J, Rixon FJ, Chiu W. 2000. Seeing the Herpesvirus Capsid at 8.5 Å. Science 288:877–880.

22. Steven AC, Roberts CR, Hay J, Bisher ME, Pun T, Trus BL. 1986. Hexavalent capsomers of herpes simplex virus type 2: symmetry, shape, dimensions, and oligomeric status. J Virol 57:578–584.

23. Trus BL, Heymann JB, Nealon K, Cheng N, Newcomb WW, Brown JC, Kedes DH, Steven AC. 2001. Capsid Structure of Kaposi's Sarcoma-Associated Herpesvirus, a Gammaherpesvirus, Compared to Those of an Alphaherpesvirus, Herpes Simplex Virus Type 1, and a Betaherpesvirus, Cytomegalovirus. Journal of Virology 75:2879–2890.

24. Lo P, Yu X, Atanasov I, Chandran B, Zhou ZH. 2003. Three-Dimensional Localization of pORF65 in Kaposi's Sarcoma-Associated Herpesvirus Capsid. Journal of Virology 77:4291–4297.

25. Yu X, Jih J, Jiang J, Zhou ZH. 2017. Atomic structure of the human cytomegalovirus capsid with its securing tegument layer of pp150. Science 356.

26. Dai X, Zhou ZH. 2018. Structure of the herpes simplex virus 1 capsid with associated tegument protein complexes. Science 360.

27. Yuan S, Wang J, Zhu D, Wang N, Gao Q, Chen W, Tang H, Wang J, Zhang X, Liu H, Rao Z, Wang X. 2018. Cryo-EM structure of a herpesvirus capsid at 3.1 Å. Science 360:eaao7283.

28. Wang J, Yuan S, Zhu D, Tang H, Wang N, Chen W, Gao Q, Li Y, Wang J, Liu H, Zhang X, Rao Z, Wang X. 2018. Structure of the herpes simplex virus type 2 C-capsid with capsid-vertex-specific component. 1. Nature Communications 9:3668.

29. Liu Y-T, Jih J, Dai X, Bi G-Q, Zhou ZH. 2019. Cryo-EM structures of herpes simplex virus type 1 portal vertex and packaged genome. Nature 570:257–261.

30. Zhang Y, Liu W, Li Z, Kumar V, Alvarez-Cabrera AL, Leibovitch EC, Cui Y, Mei Y, Bi G-Q, Jacobson S, Zhou ZH. 2019. Atomic structure of the human herpesvirus 6B capsid and capsid-associated tegument complexes. 1. Nature Communications 10:5346.

31. Li Z, Zhang X, Dong L, Pang J, Xu M, Zhong Q, Zeng M-S, Yu X. 2020. CryoEM structure of the tegumented capsid of Epstein-Barr virus. Cell Research 1–12.

32. Liu W, Cui Y, Wang C, Li Z, Gong D, Dai X, Bi G-Q, Sun R, Zhou ZH. 2020. Structures of capsid and capsid-associated tegument complex inside the Epstein–Barr virus. Nature Microbiology 1–14.

33. Sun J, Liu C, Peng R, Zhang F-K, Tong Z, Liu S, Shi Y, Zhao Z, Zeng W-B, Gao GF, Shen H-J, Yang X, Luo M, Qi J, Wang P. 2020. Cryo-EM structure of the varicella-zoster virus A-capsid. 1. Nat Commun 11:4795.

34. Wang W, Zheng Q, Pan D, Yu H, Fu W, Liu J, He M, Zhu R, Cai Y, Huang Y, Zha Z, Chen Z, Ye X, Han J, Que Y, Wu T, Zhang J, Li S, Zhu H, Zhou ZH, Cheng T, Xia N. 2020. Near-atomic cryo-electron microscopy structures of varicella-zoster virus capsids. Nat Microbiol 5:1542–1552.

35. Li Z, Pang J, Dong L, Yu X. 2021. Structural basis for genome packaging, retention, and ejection in human cytomegalovirus. Nat Commun 12:4538.

36. Li Z, Pang J, Gao R, Wang Q, Zhang M, Yu X. 2023. Cryo-electron microscopy structures of capsids and in situ portals of DNA-devoid capsids of human cytomegalovirus. 1. Nat Commun 14:2025.

37. Liu Y-T, Strugatsky D, Liu W, Zhou ZH. 2021. Structure of human cytomegalovirus virion reveals host tRNA binding to capsid-associated tegument protein pp150. Nat Commun 12:5513.

38. Jih J, Liu Y-T, Liu W, Zhou ZH. 2024. The incredible bulk: Human cytomegalovirus tegument architectures uncovered by AI-empowered cryo-EM. Science Advances 10:eadj1640.

39. Cao L, Wang N, Lv Z, Chen W, Chen Z, Song L, Sha X, Wang G, Hu Y, Lian X, Cui G, Fan J, Quan Y, Liu H, Hou H, Wang X. 2024. Insights into varicella-zoster virus assembly from the B- and C-capsid at near-atomic resolution structures. hLife 2:64–74.

40. Dai X, Gong D, Lim H, Jih J, Wu T-T, Sun R, Zhou ZH. 2018. Structure and mutagenesis reveal essential capsid protein interactions for KSHV replication. Nature 553:521–525.

41. Grünewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, Steven AC. 2003. Three-Dimensional Structure of Herpes Simplex Virus from Cryo-Electron Tomography. Science 302:1396–1398.

42. Si Z, Zhang J, Shivakoti S, Atanasov I, Tao C-L, Hui WH, Zhou K, Yu X, Li W, Luo M, Bi G-Q, Zhou ZH. 2018. Different functional states of fusion protein gB revealed on human cytomegalovirus by cryo electron tomography with Volta phase plate. PLOS Pathogens 14:e1007452.

43. Baumeister W. 2005. From proteomic inventory to architecture. FEBS Letters 579:933–937.

44. Liu Y-T, Zhang H, Wang H, Tao C-L, Bi G-Q, Zhou ZH. 2022. Isotropic reconstruction for electron tomography with deep learning. 1. Nat Commun 13:6482.

45. Deng B, O'Connor CM, Kedes DH, Zhou ZH. 2007. Direct Visualization of the Putative Portal in the Kaposi's Sarcoma-Associated Herpesvirus Capsid by Cryoelectron Tomography. Journal of Virology 81:3640–3644.

46. Dai W, Jia Q, Bortz E, Shah S, Liu J, Atanasov I, Li X, Taylor KA, Sun R, Zhou ZH. 2008. Unique structures in a tumor herpesvirus revealed by cryo-electron tomography and microscopy. Journal of Structural Biology 161:428–438.

47. Brulois KF, Chang H, Lee AS-Y, Ensser A, Wong L-Y, Toth Z, Lee SH, Lee H-R, Myoung J, Ganem D, Oh T-K, Kim JF, Gao S-J, Jung JU. 2012. Construction and Manipulation of a New Kaposi's Sarcoma-Associated Herpesvirus Bacterial Artificial Chromosome Clone. Journal of Virology 86:9708–9720.

48. Zivanov J, Otón J, Ke Z, von Kügelgen A, Pyle E, Qu K, Morado D, Castaño-Díez D, Zanetti G, Bharat TA, Briggs JA, Scheres SH. 2022. A Bayesian approach to single-particle electron cryo-tomography in RELION-4.0. eLife 11:e83724.

49. Gong D, Dai X, Xiao Y, Du Y, Chapa TJ, Johnson JR, Li X, Krogan NJ, Deng H, Wu T-T, Sun R. 2017. Virus-Like Vesicles of Kaposi's Sarcoma-Associated Herpesvirus Activate Lytic Replication by Triggering Differentiation Signaling. Journal of Virology 91.

50. Wan W, Briggs JAG. 2016. Chapter Thirteen - Cryo-Electron Tomography and Subtomogram Averaging, p. 329–367. *In* Crowther, RA (ed.), Methods in Enzymology. Academic Press.

51. Trus BL, Newcomb WW, Cheng N, Cardone G, Marekov L, Homa FL, Brown JC, Steven AC. 2007. Allosteric Signalling and a Nuclear Exit Strategy: Binding of UL25/UL17 Heterodimers to DNA-filled HSV-1 Capsids. Mol Cell 26:479–489.

52. Conway JF, Cockrell SK, Copeland AM, Newcomb WW, Brown JC, Homa FL. 2010. Labeling and localization of the herpes simplex virus capsid protein UL25 and its interaction with the two triplexes closest to the penton. J Mol Biol 397:575–586.

53. Toropova K, Huffman JB, Homa FL, Conway JF. 2011. The herpes simplex virus 1 UL17 protein is the second constituent of the capsid vertex-specific component required for DNA packaging and retention. J Virol 85:7513–7522.

54. Cockrell SK, Huffman JB, Toropova K, Conway JF, Homa FL. 2011. Residues of the UL25 Protein of Herpes Simplex Virus That Are Required for Its Stable Interaction with Capsids. Journal of Virology 85:4875–4887.

55. Dai X, Gong D, Wu T-T, Sun R, Zhou ZH. 2014. Organization of Capsid-Associated Tegument Components in Kaposi's Sarcoma-Associated Herpesvirus. Journal of Virology 88:12694–12702.

56. Ilca SL, Kotecha A, Sun X, Poranen MM, Stuart DI, Huiskonen JT. 2015. Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. 1. Nature Communications 6:8843.

57. Naniima P, Naimo E, Koch S, Curth U, Alkharsah KR, Ströh LJ, Binz A, Beneke J-M, Vollmer B, Böning H, Borst EM, Desai P, Bohne J, Messerle M, Bauerfeind R, Legrand P, Sodeik B, Schulz TF, Krey T. 2021. Assembly of infectious Kaposi's sarcoma-associated herpesvirus progeny requires formation of a pORF19 pentamer. PLOS Biology 19:e3001423.

58. Bogdanow B, Gruska I, Mühlberg L, Protze J, Hohensee S, Vetter B, Bosse JB, Lehmann M, Sadeghi M, Wiebusch L, Liu F. 2023. Spatially resolved protein map of intact human cytomegalovirus virions. 9. Nat Microbiol 8:1732–1747.

59. Ibiricu I, Huiskonen JT, Döhner K, Bradke F, Sodeik B, Grünewald K. 2011. Cryo Electron Tomography of Herpes Simplex Virus during Axonal Transport and Secondary Envelopment in Primary Neurons. PLOS Pathogens 7:e1002406.

60. Liu F, Zhou ZH. 2007. Comparative virion structures of human herpesviruses, p. . *In* Arvin, A, Campadelli-Fiume, G, Mocarski, E, Moore, PS, Roizman, B, Whitley, R, Yamanishi, K (eds.), Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge University Press, Cambridge.

61. Baldick CJ, Shenk T. 1996. Proteins associated with purified human cytomegalovirus particles. J Virol 70:6097–6105.

62. Michael K, Klupp BG, Mettenleiter TC, Karger A. 2006. Composition of Pseudorabies Virus Particles Lacking Tegument Protein US3, UL47, or UL49 or Envelope Glycoprotein E. J Virol 80:1332–1339.

63. Varnum SM, Streblow DN, Monroe ME, Smith P, Auberry KJ, Paša-Tolić L, Wang D, Camp DG, Rodland K, Wiley S, Britt W, Shenk T, Smith RD, Nelson JA. 2004. Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome. Journal of Virology 78:10960–10966.

64. Loret S, Guay G, Lippé R. 2008. Comprehensive Characterization of Extracellular Herpes Simplex Virus Type 1 Virions. Journal of Virology 82:8605–8618.

65. Phillips SL, Bresnahan WA. 2012. The Human Cytomegalovirus (HCMV) Tegument Protein UL94 Is Essential for Secondary Envelopment of HCMV Virions. Journal of Virology 86:2523–2532.

66. Newcomb WW, Brown JC. 2010. Structure and Capsid Association of the Herpesvirus Large Tegument Protein UL36. Journal of Virology 84:9408–9414.

67. Fan WH, Roberts APE, McElwee M, Bhella D, Rixon FJ, Lauder R. 2014. The Large Tegument Protein pUL36 Is Essential for Formation of the Capsid Vertex-Specific Component at the Capsid-Tegument Interface of Herpes Simplex Virus 1. J Virol 89:1502–1511.

68. Liu Y-T, Shivakoti S, Jia F, Tao C-L, Zhang B, Xu F, Lau P, Bi G-Q, Zhou ZH. 2020. Biphasic exocytosis of herpesvirus from hippocampal neurons and mechanistic implication to membrane fusion. 1. Cell Discov 6:1–12.

69. Flomm FJ, Soh TK, Schneider C, Wedemann L, Britt HM, Thalassinos K, Pfitzner S, Reimer R, Grünewald K, Bosse JB. 2022. Intermittent bulk release of human cytomegalovirus. PLOS Pathogens 18:e1010575.

70. Nanbo A, Noda T, Ohba Y. 2018. Epstein–Barr Virus Acquires Its Final Envelope on Intracellular Compartments With Golgi Markers. Frontiers in Microbiology 9.

71. Maurer UE, Zeev-Ben-Mordehai T, Pandurangan AP, Cairns TM, Hannah BP, Whitbeck JC, Eisenberg RJ, Cohen GH, Topf M, Huiskonen JT, Grünewald K. 2013. The Structure of Herpesvirus Fusion Glycoprotein B-Bilayer Complex Reveals the Protein-Membrane and Lateral Protein-Protein Interaction. Structure 21:1396–1405.

72. Liu Y, Heim KP, Che Y, Chi X, Qiu X, Han S, Dormitzer PR, Yang X. 2021. Prefusion structure of human cytomegalovirus glycoprotein B and structural basis for membrane fusion. Science Advances 7:eabf3178.

73. Fontana J, Atanasiu D, Saw WT, Gallagher JR, Cox RG, Whitbeck JC, Brown LM, Eisenberg RJ, Cohen GH. 2017. The Fusion Loops of the Initial Prefusion Conformation of Herpes Simplex Virus 1 Fusion Protein Point Toward the Membrane. mBio 8:e01268-17.

74. Backovic M, Longnecker R, Jardetzky TS. 2009. Structure of a trimeric variant of the Epstein–Barr virus glycoprotein B. PNAS 106:2880–2885.

75. Su C, Wu L, Chai Y, Qi J, Tan S, Gao GF, Song H, Yan J. 2020. Molecular basis of EphA2 recognition by gHgL from gammaherpesviruses. 1. Nature Communications 11:5964.

76. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 1–11.

77. Zeev-Ben-Mordehai T, Vasishtan D, Hernández Durán A, Vollmer B, White P, Prasad Pandurangan A, Siebert CA, Topf M, Grünewald K. 2016. Two distinct trimeric conformations of natively membrane-anchored full-length herpes simplex virus 1 glycoprotein B. Proc Natl Acad Sci U S A 113:4176–4181.

78. Vollmer B, Pražák V, Vasishtan D, Jefferys EE, Hernandez-Duran A, Vallbracht M, Klupp BG, Mettenleiter TC, Backovic M, Rey FA, Topf M, Grünewald K. 2020. The prefusion structure of herpes simplex virus glycoprotein B. Science Advances 6:eabc1726.

79. Yang S, Hiotis G, Wang Y, Chen J, Wang J, Kim M, Reinherz EL, Walz T. 2022. Dynamic HIV-1 spike motion creates vulnerability for its membrane-bound tripod to antibody attack. bioRxiv https://doi.org/10.1101/2022.05.09.491169.

80. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, Zhang J, Weng T, Zhang Z, Wu Z, Cheng L, Shi D, Lu X, Lei J, Crispin M, Shi Y, Li L, Li S. 2020. Molecular Architecture of the SARS-CoV-2 Virus. Cell 183:730-738.e13.

81. Peng L, Ryazantsev S, Sun R, Zhou ZH. 2010. Three-Dimensional Visualization of Gammaherpesvirus Life Cycle in Host Cells by Electron Tomography. Structure 18:47–58.

82. Maurer UE, Sodeik B, Grünewald K. 2008. Native 3D intermediates of membrane fusion in herpes simplex virus 1 entry. Proceedings of the National Academy of Sciences 105:10559–10564.

83. Shi L, Shen Q-T, Kiel A, Wang J, Wang H-W, Melia TJ, Rothman JE, Pincet F. 2012. SNARE Proteins: One to Fuse and Three to Keep the Nascent Fusion Pore Open. Science 335:1355–1359.

84. Kim IS, Jenni S, Stanifer ML, Roth E, Whelan SPJ, van Oijen AM, Harrison SC. 2017. Mechanism of membrane fusion induced by vesicular stomatitis virus G protein. PNAS 114:E28–E36.

85. Kawabata A, Tang H, Huang H, Yamanishi K, Mori Y. 2009. Human herpesvirus 6 envelope components enriched in lipid rafts: evidence for virion-associated lipid rafts. Virol J 6:127.

86. Bender FC, Whitbeck JC, Ponce de Leon M, Lou H, Eisenberg RJ, Cohen GH. 2003. Specific Association of Glycoprotein B with Lipid Rafts during Herpes Simplex Virus Entry. Journal of Virology 77:9542–9552.

87. Lee GE, Church GA, Wilson DW. 2003. A Subpopulation of Tegument Protein vhs Localizes to Detergent-Insoluble Lipid Rafts in Herpes Simplex Virus-Infected Cells. Journal of Virology 77:2038–2045.

88. Goswami R, Shair KHY, Gershburg E. 2017. Molecular diversity of IgG responses to Epstein–Barr virus proteins in asymptomatic Epstein–Barr virus carriers. Journal of General Virology 98:2343–2350.

89. Coghill AE, Pfeiffer RM, Proietti C, Hsu W-L, Chien Y-C, Lekieffre L, Krause L, Teng A, Pablo J, Yu KJ PhD, Lou P-J MD, Wang C-P, Liu Z, Chen C-J, Middeldorp J, Mulvenna J, Bethony J, Hildesheim A, Doolan DL. 2018. Identification of a Novel, EBV-Based Antibody Risk Stratification Signature for Early Detection of Nasopharyngeal Carcinoma in Taiwan. Clinical Cancer Research 24:1305–1314.

90. Coghill AE, Fang J, Liu Z, Chen C-J, Jarrett RF, Hjalgrim H, Proietti C, Yu KJ, Hsu W-L, Lou P-J, Wang C-P, Zhao Y, Doolan DL, Hildesheim A. 2022. Identifying Epstein–Barr virus peptide sequences associated with differential IgG antibody response. International Journal of Infectious Diseases 114:65–71.

91. Labo N, Miley W, Marshall V, Gillette W, Esposito D, Bess M, Turano A, Uldrick T, Polizzotto MN, Wyvill KM, Bagni R, Yarchoan R, Whitby D. 2014. Heterogeneity and Breadth of Host Antibody Response to KSHV Infection Demonstrated by Systematic Analysis of the KSHV Proteome. PLOS Pathogens 10:e1004046.

92. Lam AK, Roshan R, Miley W, Labo N, Zhen J, Kurland AP, Cheng C, Huang H, Teng P-L, Harelson C, Gong D, Tam YK, Radu CG, Epeldegui M, Johnson JR, Zhou ZH, Whitby D, Wu T-T. 2023. Immunization of Mice with Virus-Like Vesicles of Kaposi Sarcoma-Associated Herpesvirus Reveals a Role for Antibodies Targeting ORF4 in Activating Complement-Mediated Neutralization. Journal of Virology 0:e01600-22.

93. Semmes EC, Miller IG, Wimberly CE, Phan CT, Jenks JA, Harnois MJ, Berendam SJ, Webster H, Hurst JH, Kurtzberg J, Fouda GG, Walsh KM, Permar SR. 2022. Maternal Fc-mediated non-neutralizing antibody responses correlate with protection against congenital human cytomegalovirus infection. J Clin Invest 132.

94. Greijer AE, van de Crommert JMG, Stevens SJC, Middeldorp JM. 1999. Molecular Fine-Specificity Analysis of Antibody Responses to Human Cytomegalovirus and Design of Novel Synthetic-Peptide-Based Serodiagnostic Assays. Journal of Clinical Microbiology 37:179–188.

95. Schoppel K, Kropff B, Schmidt C, Vornhagen R, Mach M. 1997. The Humoral Immune Response against Human Cytomegalovirus Is Characterized by a Delayed Synthesis of Glycoprotein-Specific Antibodies. The Journal of Infectious Diseases 175:533–544.

96. Shibamura M, Yoshikawa T, Yamada S, Inagaki T, Nguyen PHA, Fujii H, Harada S, Fukushi S, Oka A, Mizuguchi M, Saijo M. 2020. Association of human cytomegalovirus (HCMV) neutralizing antibodies with antibodies to the HCMV glycoprotein complexes. Virol J 17:120.

97. Mann DR, Hilty MD. 1982. Antibody Response to Herpes Simplex Virus Type 1 Polypeptides and Glycoproteins in Primary and Recurrent Infection. 3. Pediatr Res 16:176–180.

98. Kalantari-Dehaghi M, Chun S, Chentoufi AA, Pablo J, Liang L, Dasgupta G, Molina DM, Jasinskas A, Nakajima-Sasaki R, Felgner J, Hermanson G, BenMohamed L, Felgner PL, Davies DH. 2012. Discovery of Potential Diagnostic and Vaccine Antigens in Herpes Simplex Virus 1 and 2 by Proteome-Wide Antibody Profiling. Journal of Virology 86:4328–4339.

99. Kirchmeier M, Fluckiger A-C, Soare C, Bozic J, Ontsouka B, Ahmed T, Diress A, Pereira L, Schödel F, Plotkin S, Dalba C, Klatzmann D, Anderson DE. 2014. Enveloped Virus-Like Particle Expression of Human Cytomegalovirus Glycoprotein B Antigen Induces Antibodies with Potent and Broad Neutralizing Activity. Clinical and Vaccine Immunology 21:174–180.

100. Cairns TM, Huang Z-Y, Gallagher JR, Lin Y, Lou H, Whitbeck JC, Wald A, Cohen GH, Eisenberg RJ. 2015. Patient-Specific Neutralizing Antibody Responses to Herpes Simplex Virus Are Attributed to Epitopes on gD, gB, or Both and Can Be Type Specific. J Virol 89:9213–9231.

101. Zhang X, Hong J, Zhong L, Wu Q, Zhang S, Zhu Q, Chen H, Wei D, Li R, Zhang W, Zhang X, Wang G, Zhou X, Chen J, Kang Y, Zha Z, Duan X, Huang Y, Sun C, Kong X, Zhou Y, Chen Y, Ye X, Feng Q, Li S, Xiang T, Gao S, Zeng M-S, Zheng Q, Chen Y, Zeng Y-X, Xia N, Xu M. 2022. Protective anti-gB neutralizing antibodies targeting two vulnerable sites for EBV-cell membrane fusion. Proceedings of the National Academy of Sciences 119:e2202371119.

102. Zhong L, Zhang W, Krummenacher C, Chen Y, Zheng Q, Zhao Q, Zeng M-S, Xia N, Zeng Y-X, Xu M, Zhang X. 2023. Targeting herpesvirus entry complex and fusogen glycoproteins with prophylactic and therapeutic agents. Trends in Microbiology 31:788–804.

103. Pedro Simas J, Efstathiou S. 1998. Murine gammaherpesvirus 68: a model for the study of gammaherpesvirus pathogenesis. Trends in Microbiology 6:276–282.

104. Liu W, Dai X, Jih J, Chan K, Trang P, Yu X, Balogun R, Mei Y, Liu F, Zhou ZH. 2019. Atomic structures and deletion mutant reveal different capsid-binding patterns and functional

significance of tegument protein pp150 in murine and human cytomegaloviruses with implications for therapeutic development. PLOS Pathogens 15:e1007615.

105. Glaeser RM, Hall RJ. 2011. Reaching the Information Limit in Cryo-EM of Biological Macromolecules: Experimental Aspects. Biophys J 100:2331–2337.

106. Dickerson JL, Lu P-H, Hristov D, Dunin-Borkowski RE, Russo CJ. 2022. Imaging biological macromolecules in thick specimens: The role of inelastic scattering in cryoEM. Ultramicroscopy 237:113510.

107. Buijsse B, Trompenaars P, Altin V, Danev R, Glaeser RM. 2020. Spectral DQE of the Volta phase plate. Ultramicroscopy 218:113079.

108. Speck P, Longnecker R. 1999. Epstein-Barr virus (EBV) infection visualized by EGFP expression demonstrates dependence on known mediators of EBV entry. Arch Virol 144:1123–1137.

109. Dai X, Zhou ZH. 2014. Purification of Herpesvirus Virions and Capsids. Bio-protocol 4:e1193–e1193.

110. Tivol WF, Briegel A, Jensen GJ. 2008. An Improved Cryogen for Plunge Freezing. Microsc Microanal 14:375–379.

111. Hagen WJH, Wan W, Briggs JAG. 2017. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. Journal of Structural Biology 197:191–198.

112. Mastronarde DN. 2005. Automated electron microscope tomography using robust prediction of specimen movements. Journal of Structural Biology 152:36–51.

113. Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. 4. Nature Methods 14:331–332.

114. Rohou A, Grigorieff N. 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology 192:216–221.

115. Mastronarde DN, Held SR. 2017. Automated tilt series alignment and tomographic reconstruction in IMOD. Journal of Structural Biology 197:102–113.

116. Heumann JM, Hoenger A, Mastronarde DN. 2011. Clustering and Variance Maps For Cryo-electron Tomography Using Wedge-Masked Differences. J Struct Biol 175:288–299.

117. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Science 27:14–25.

118. Ermel UH, Arghittu SM, Frangakis AS. 2022. ArtiaX: An electron tomography toolbox for the interactive handling of sub-tomograms in UCSF ChimeraX. Protein Science 31:e4472.

119. Tanaka M, Kodaira H, Nishiyama Y, Sata T, Kawaguchi Y. 2004. Construction of recombinant herpes simplex virus type I expressing green fluorescent protein without loss of any viral genes. Microbes and Infection 6:485–493.

4.12 Figures

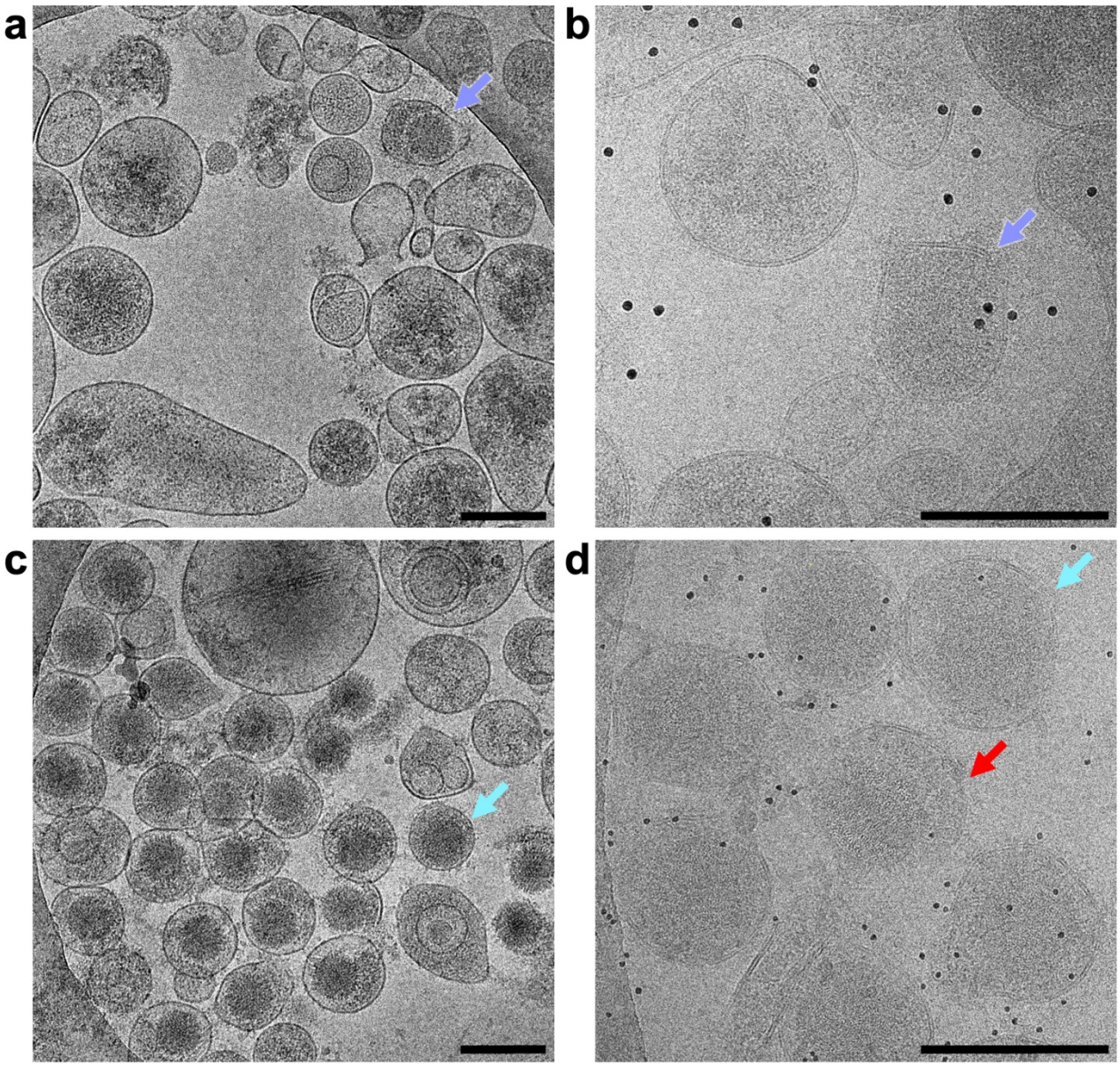


Figure 4-1. Isolated EBV and KSHV virion samples. (a) CryoEM screening image of isolated EBV virion sample. Intact EBV virion is indicated by arrow (orchid). Scale bar, 200 nm. **(b)** Low-dose image of EBV virion sample with 10 nm gold fiducials at a tilt angle of zero degrees for tomography. Intact EBV virion is indicated by arrow (orchid). Scale bar, 200 nm. **(c)** CryoEM screening image of isolated KSHV virion sample. Example of intact KSHV virions is indicated by arrow (sky). Scale bar, 200 nm. **(d)** Low-dose image of KSHV virion sample with 5 nm gold fiducials at a tilt angle of zero degrees for tomography. Example of intact KSHV virions is indicated by arrow (sky). Red arrow indicates broken virion. Scale bar, 200nm.

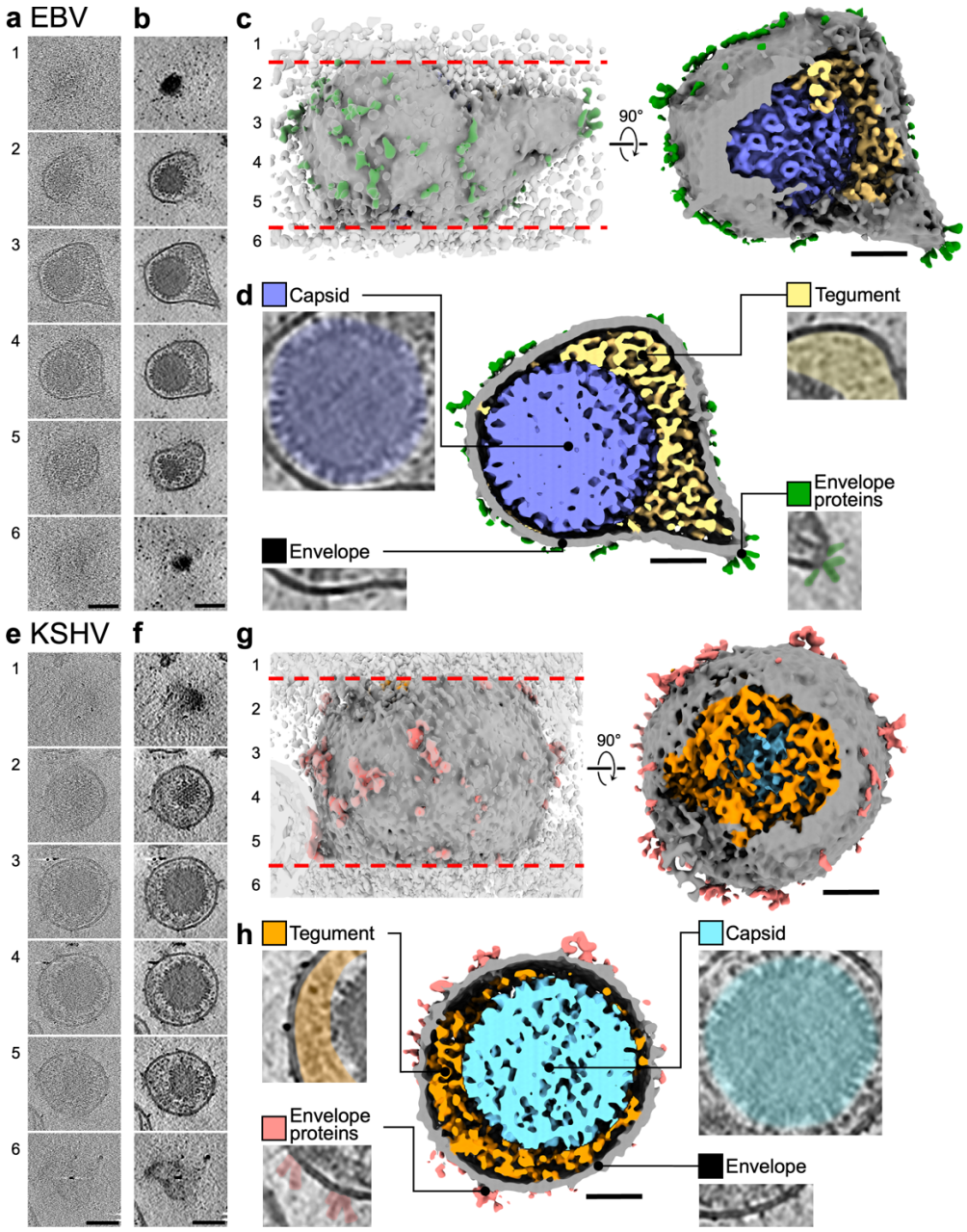


Figure 4-2. CryoET of EBV and KSHV. (a,e) Series of tomogram slices from weighted-back projection reconstruction with SIRT-like filter along z-axis of virion. Numbers correspond to location in surface rendering depiction (c,g). Scale bar, 80 nm. **(b,f)** Series of tomogram slices from isotropic reconstruction along z-axis of virion. EM density at top and bottom represent envelope, which indicates intact virion. Scale bar, 80 nm. **(c,g)** Surface renderings of virion reconstructions were cropped at the top and bottom poles (red lines) prior to segmentation to improve interpretability. Numbers indicate locations of corresponding tomogram slices (b,f). Scale bar, 40nm. **(d,h)** Sectional view of 3D surface rendering of virion colored by different components. Tomogram slices of corresponding virion component densities are indicated with their respective color. Scale bar, 40 nm.

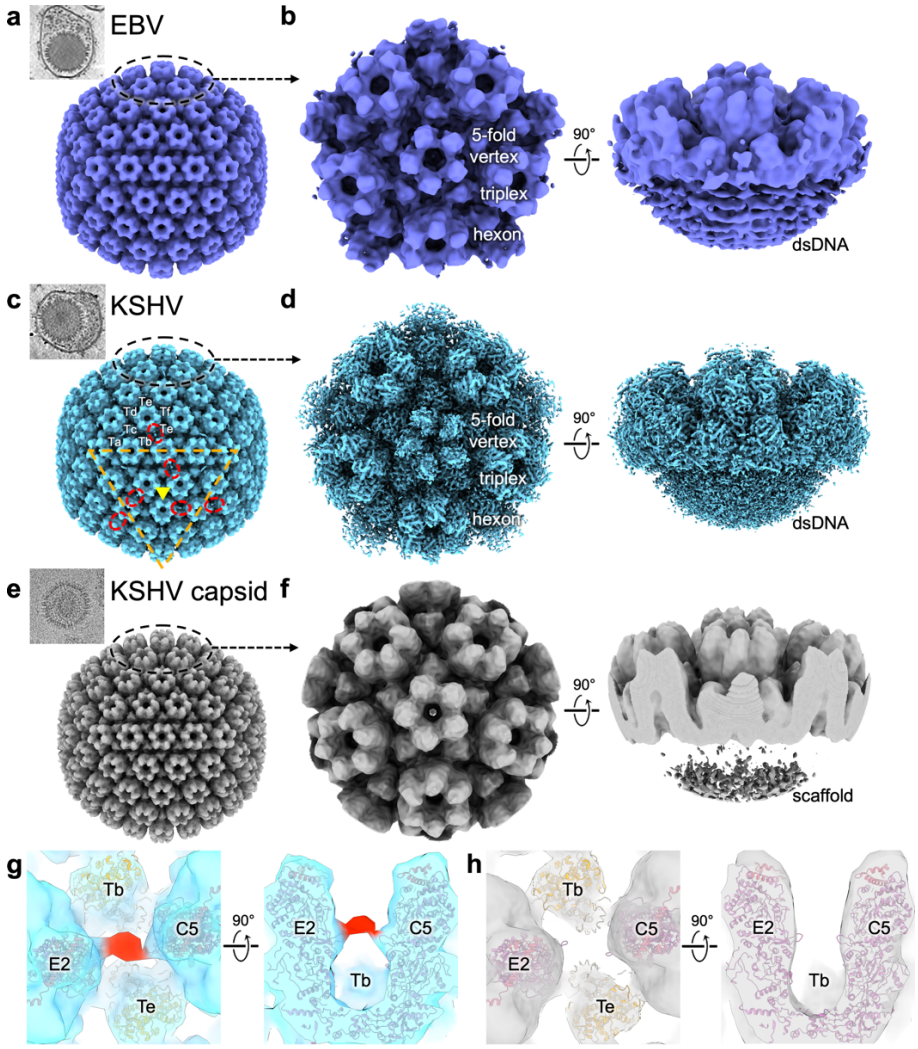


Figure 4-3. Subtomogram average reconstructions of EBV and KSHV capsid. (a,c) Subtomogram average of virion capsid with icosahedral symmetry imposed. Top-left shows tomogram slice of representative intact virion used for capsid reconstruction. In KSHV, additional densities of unknown identity are indicated (red circle). Two-fold and three-fold symmetry axes are represented by orange lines and yellow triangle, respectively. Triplex nomenclature is annotated. **(b,d)** Orthogonal views of tomogram-guided sub-particle reconstruction of capsid vertex with C_5 symmetry imposed. Hexon, triplex, 5-fold vertex, and dsDNA are labeled. **(e)** CryoEM map of naked KSHV capsid devoid of tegument with icosahedral symmetry imposed lowpass filtered to the same resolution as KSHV subtomogram average. Top-left shows cryoEM image of representative naked KSHV capsid devoid of tegument used for reconstruction. **(f)** Orthogonal views of capsid vertex density cropped from naked KSHV capsid reconstruction. **(g)** Overhead and side view of unfilled KSHV density from (c) indicated in red. Atomic models of neighboring triplexes (Tb, Te) and MCP/SCP (E2, C5) (PDB ID: 6B43) are fitted around the unfilled density. **(h)** Overhead and side view of region shown in (g) for naked KSHV capsid.

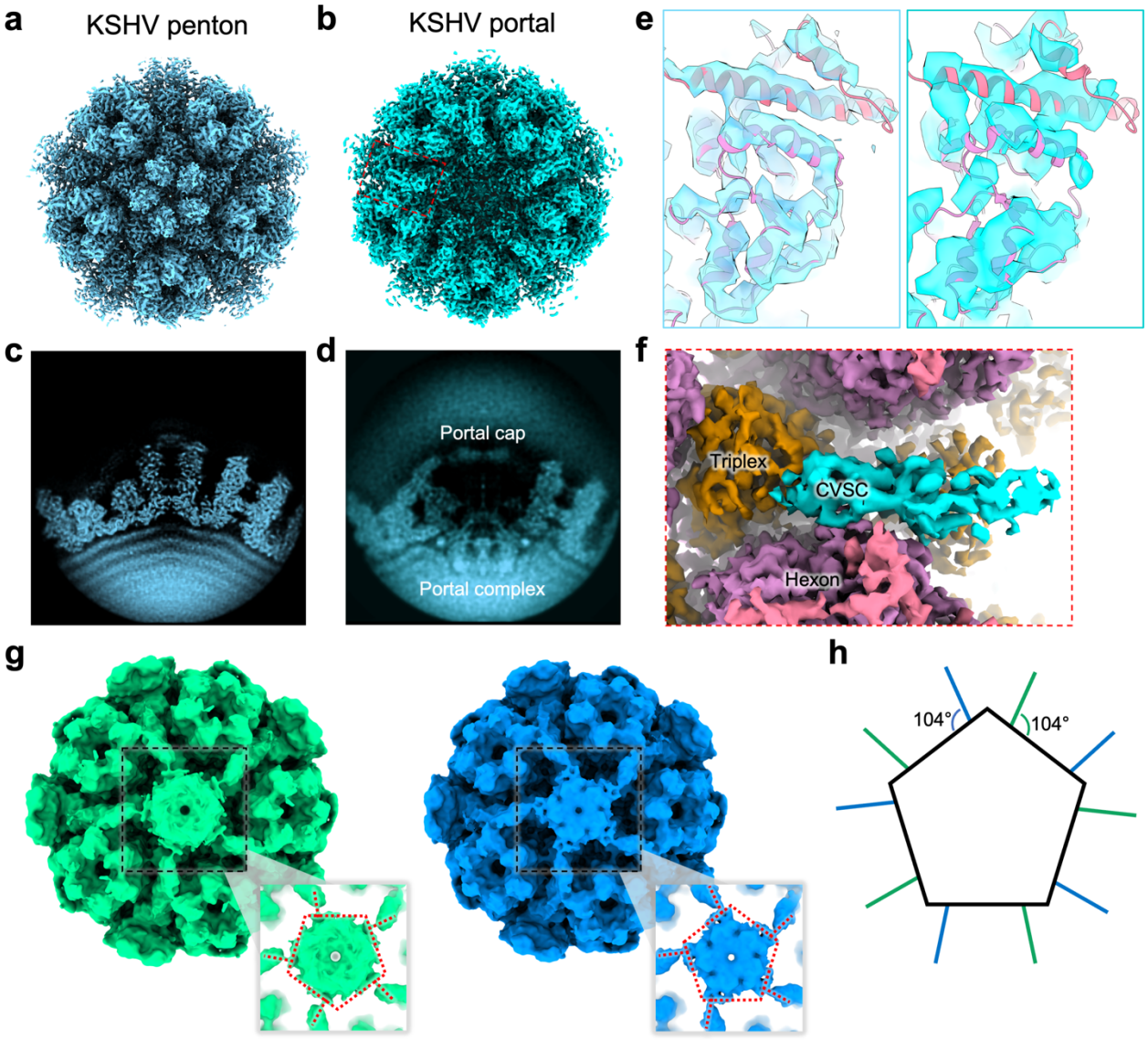


Figure 4-4. Tomogram-guided sub-particle reconstruction of KSHV penton and portal vertices. (a) Tomogram-guided sub-particle reconstruction of KSHV penton vertex with C_5 symmetry (light blue). **(b)** Tomogram-guided sub-particle reconstruction of KSHV portal vertex with C_5 symmetry (turquoise). **(c)** Side projection of KSHV penton vertex. **(d)** Side projection of KSHV portal vertex. Portal cap and portal complex are indicated. **(e)** Secondary structure fit of KSHV MCP and SCP atomic model (PDB ID: 6B43) into cryoEM map of KSHV penton vertex (left) and KSHV portal vertex (right). **(f)** Segmentation of CVSC density (teal) in KSHV portal vertex. Triplex, MCP, and SCP are colored orange, purple, and pink, respectively. **(g)** Two classes of KSHV portal vertex (green, blue). Inset shows region of portal cap density with portal cap and binding sites of CVSC indicated by red dotted lines. **(h)** Overlay of portal cap and CVSC binding site schematic from (g) adjusted for rotation of portal cap. CVSC binding sites are colored by color of respective class (green, blue). The angles between the portal cap and the connecting density to CVSC are indicated.

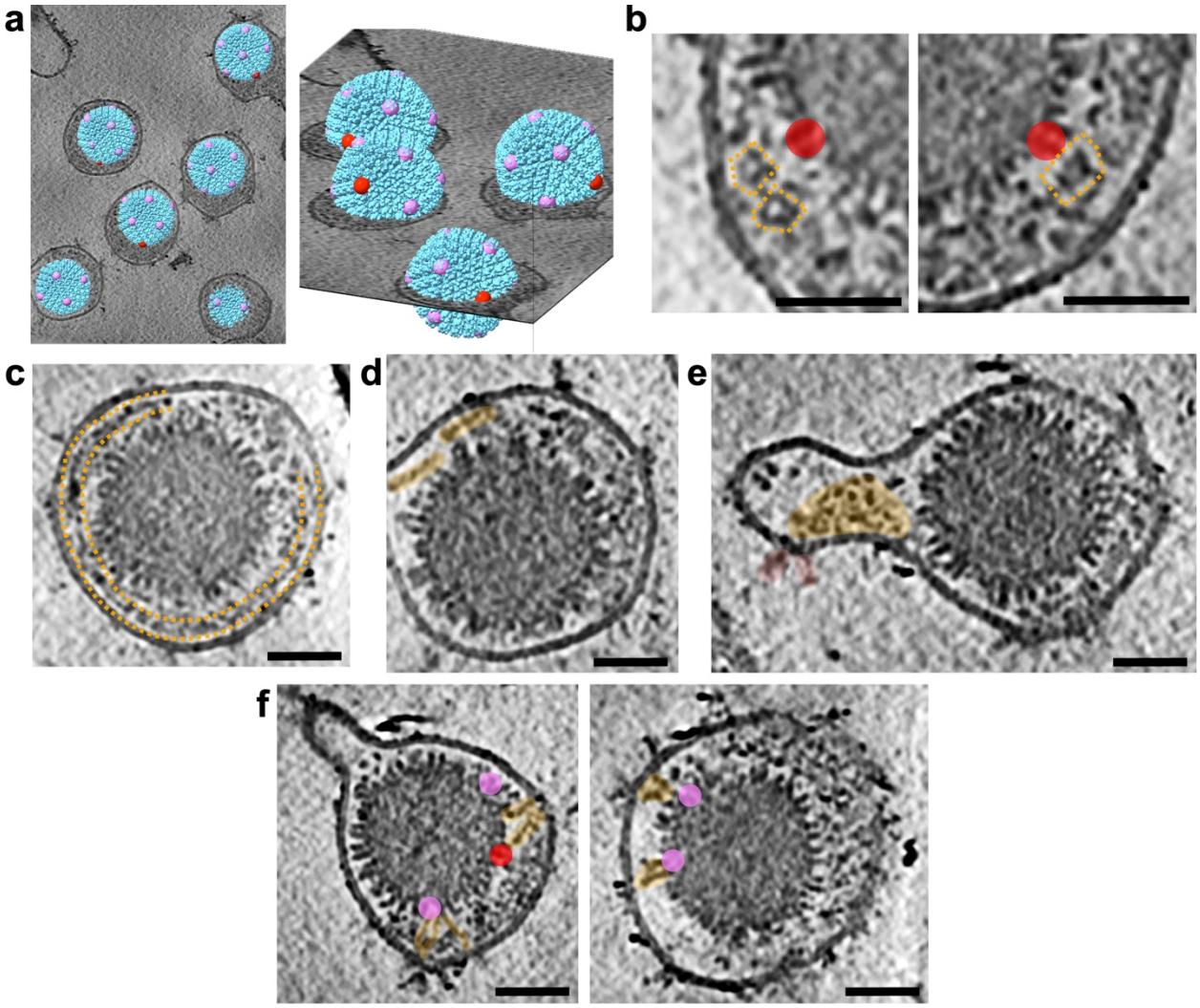


Figure 4-5. KSHV tegument in context of the portal and envelope. (a) The location and orientation of KSHV capsid 3D reconstruction (sky) mapped back into a tomogram. Portal vertex (red) and penton vertices (pink) for each virus are indicated. **(b)** Globular densities in tegument (orange) located near the portal (red) for two representative virions. **(c)** Circumnavigation of envelope by discontinuous density resembling a dotted line (between two orange lines). **(d)** Long, thin densities in outer tegument bordering envelope. **(e)** Outer tegument (orange) associated with envelope protein (red) endodomains at a protrusion with few tegument densities elsewhere. **(f)** Long, thin densities spanning from capsid to envelope. Portal vertex (red) and adjacent penton vertices (pink) are indicated. Spanning densities are present in both presence (left) and absence (right) of tegument bulk. Scale bars, 40 nm.

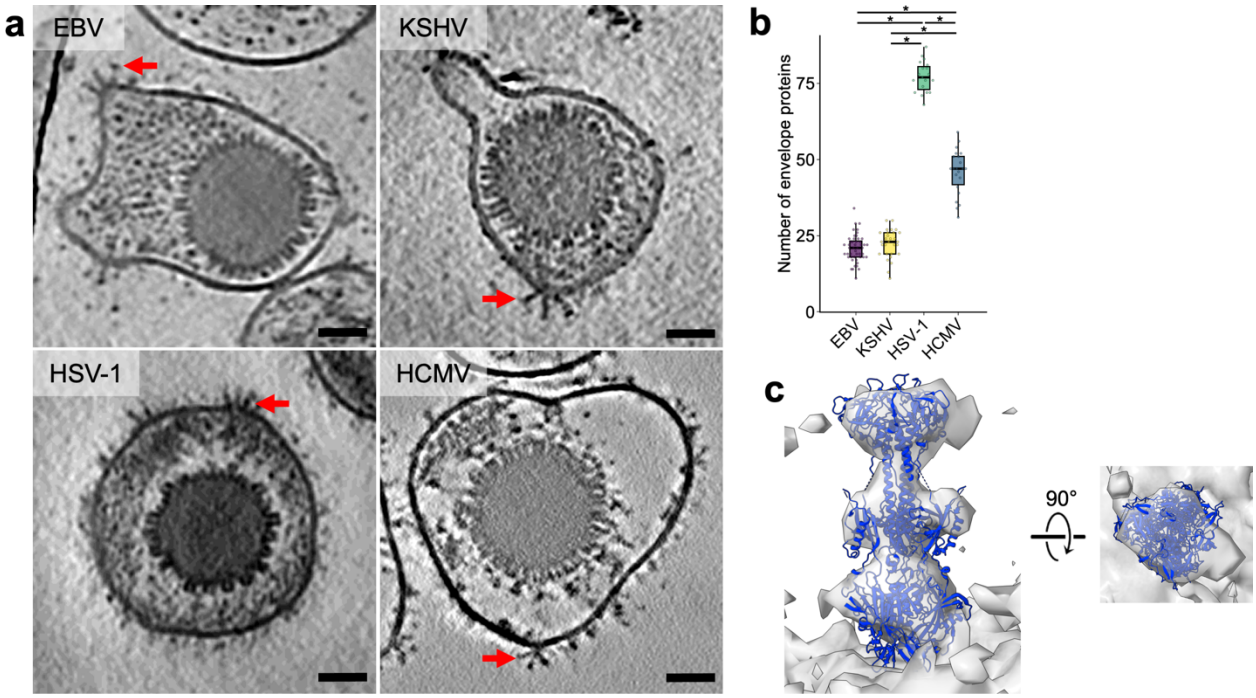
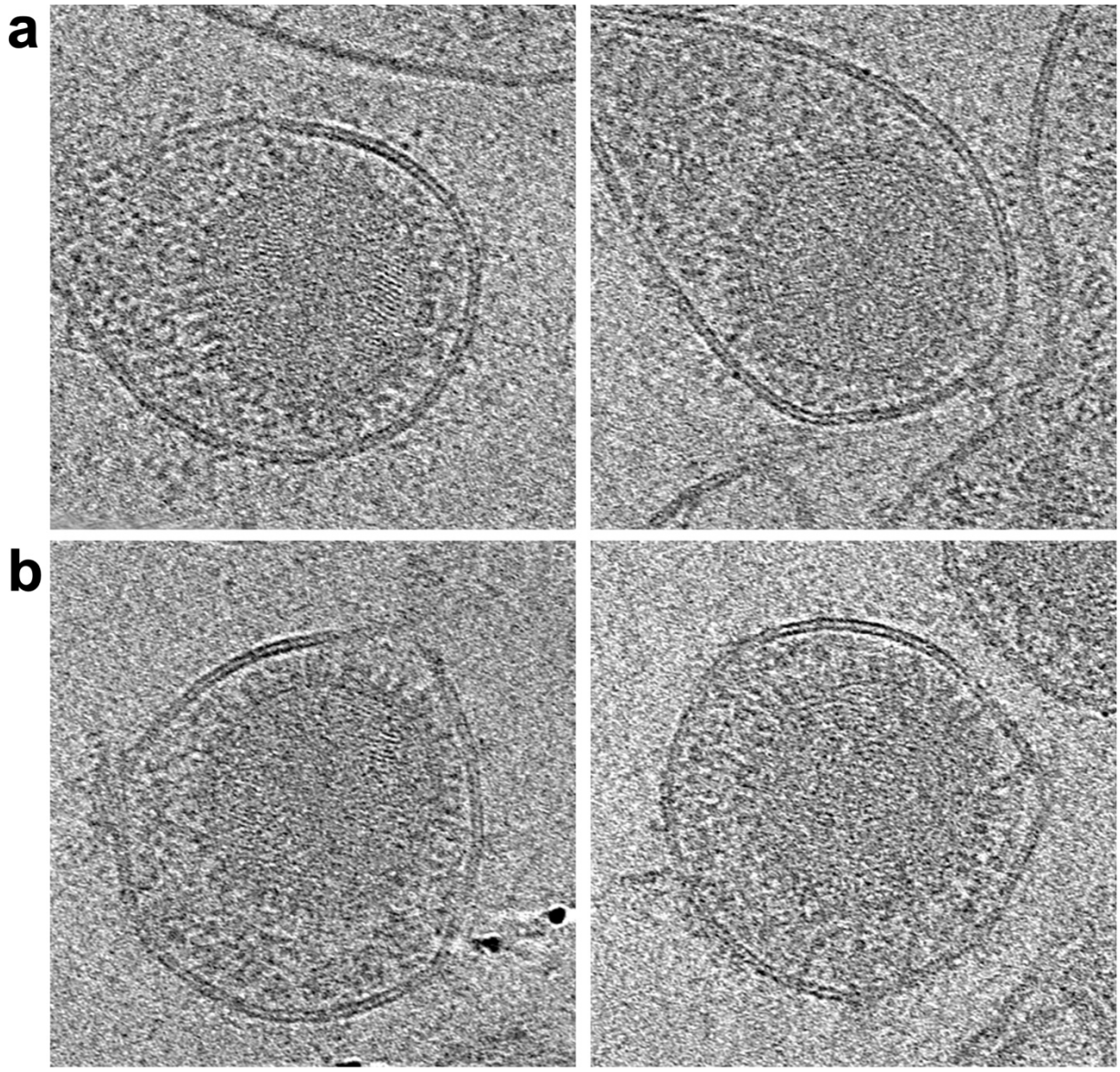
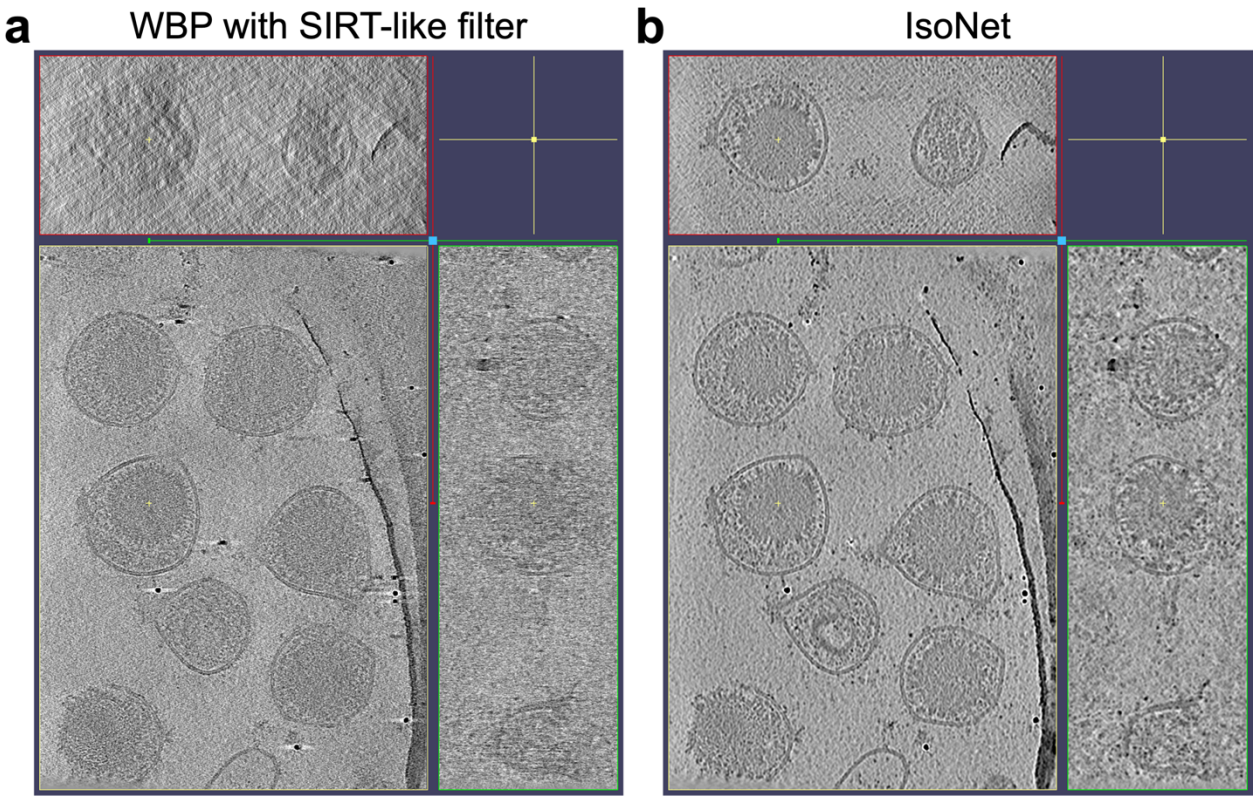


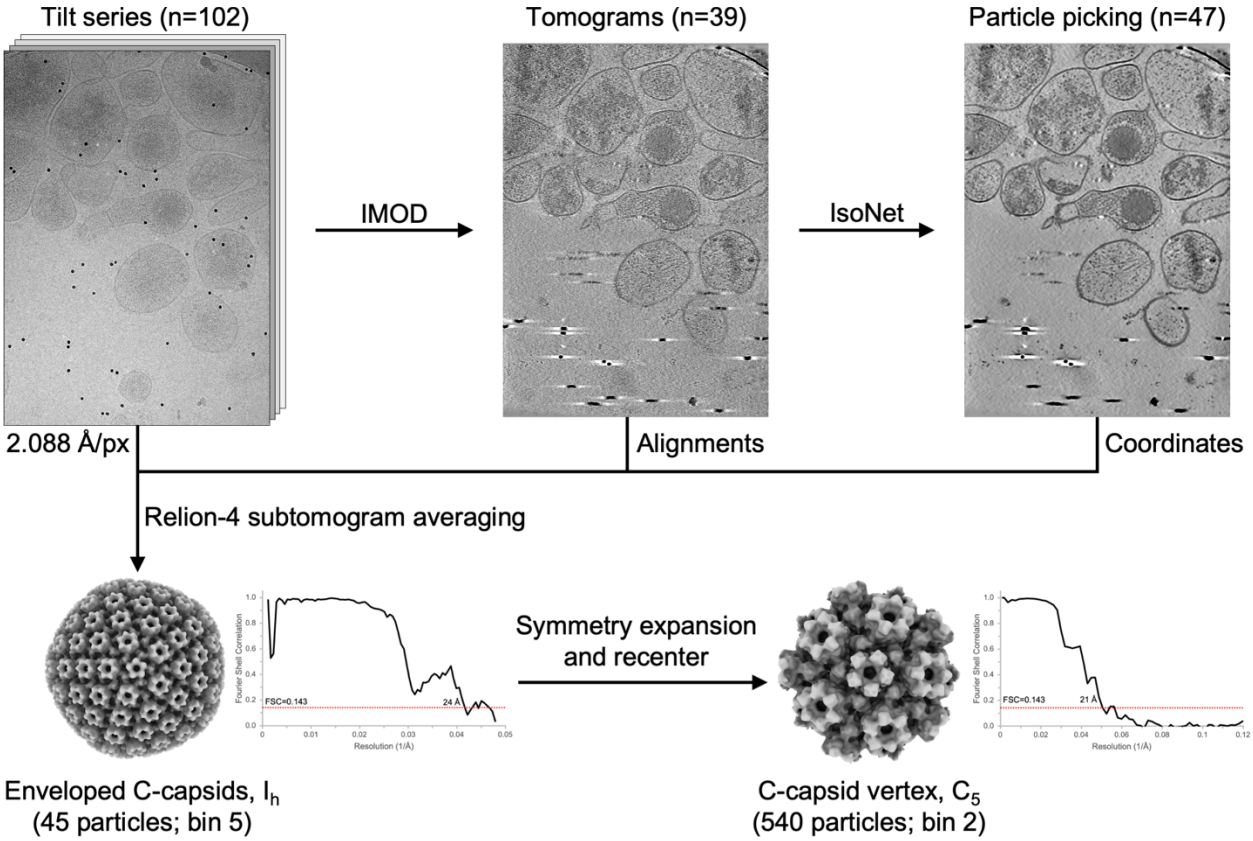
Figure 4-6. Envelope protein organization. (a) Comparison of tomogram slices between EBV, KSHV, HSV-1, and HCMV. Clusters of envelope proteins are indicated by red arrow. Scale bar, 40 nm. **(b)** Count of envelope proteins per virion for indicated herpesvirus species. Statistical analysis used ordinary one-way analysis of variance (ANOVA) with Tukey's test for multiple comparisons (*, $P < 0.05$). **(c)** Subtomogram average reconstruction of large envelope proteins from clusters in KSHV. Postfusion gB atomic model (blue; PDB ID: 3FVC) fits the cryoET density.



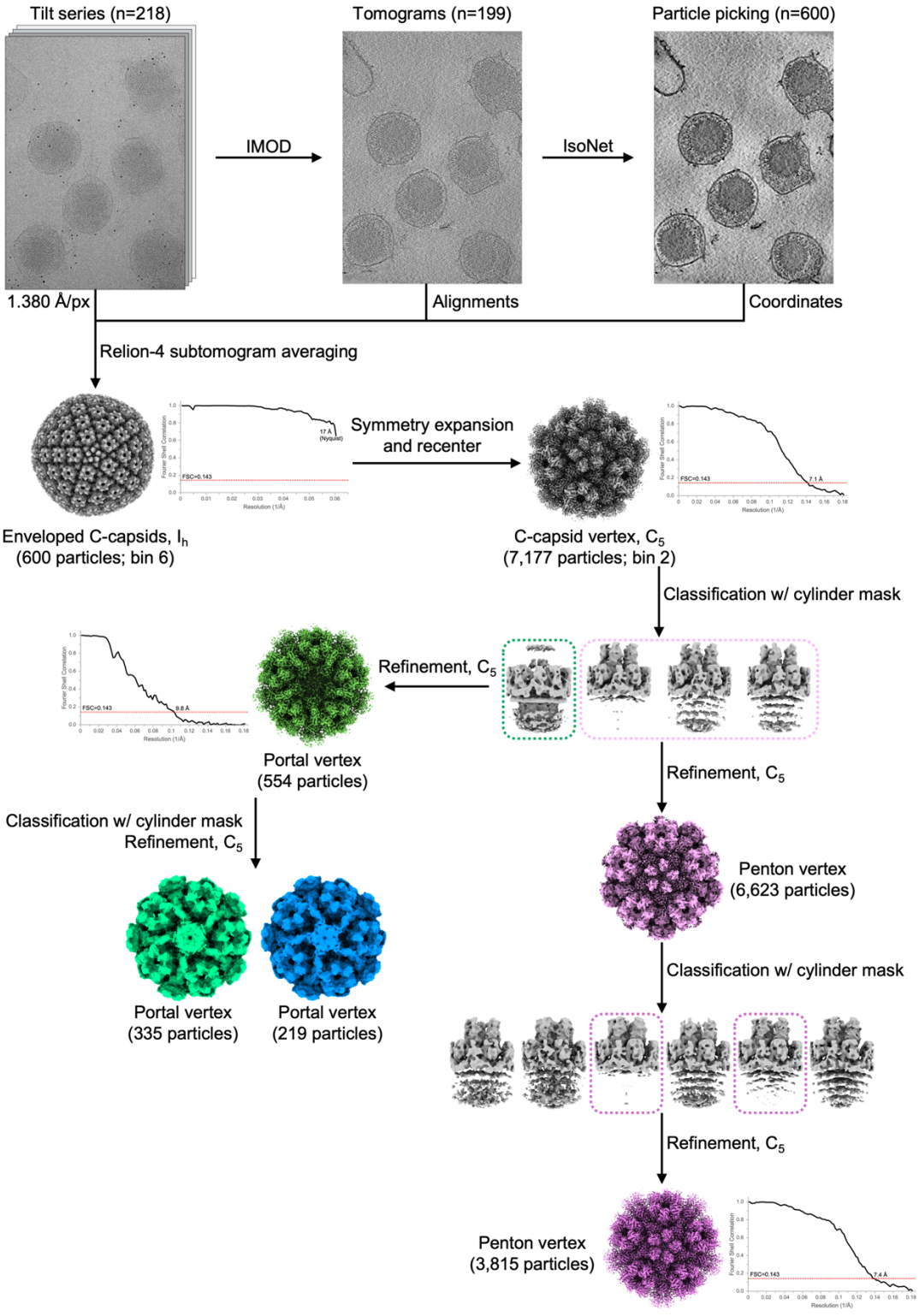
Supplementary Figure 4-S1. Tomograms capture genome strand contours. (a) Weighted back-projection with a simultaneous iterative reconstructive technique (SIRT)-like filter tomogram slice of EBV virion with genome oriented latitudinally (left) and circumferentially (right). **(b)** SIRT-like filter tomogram slice of KSHV virion with genome oriented latitudinally (left) and circumferentially (right).



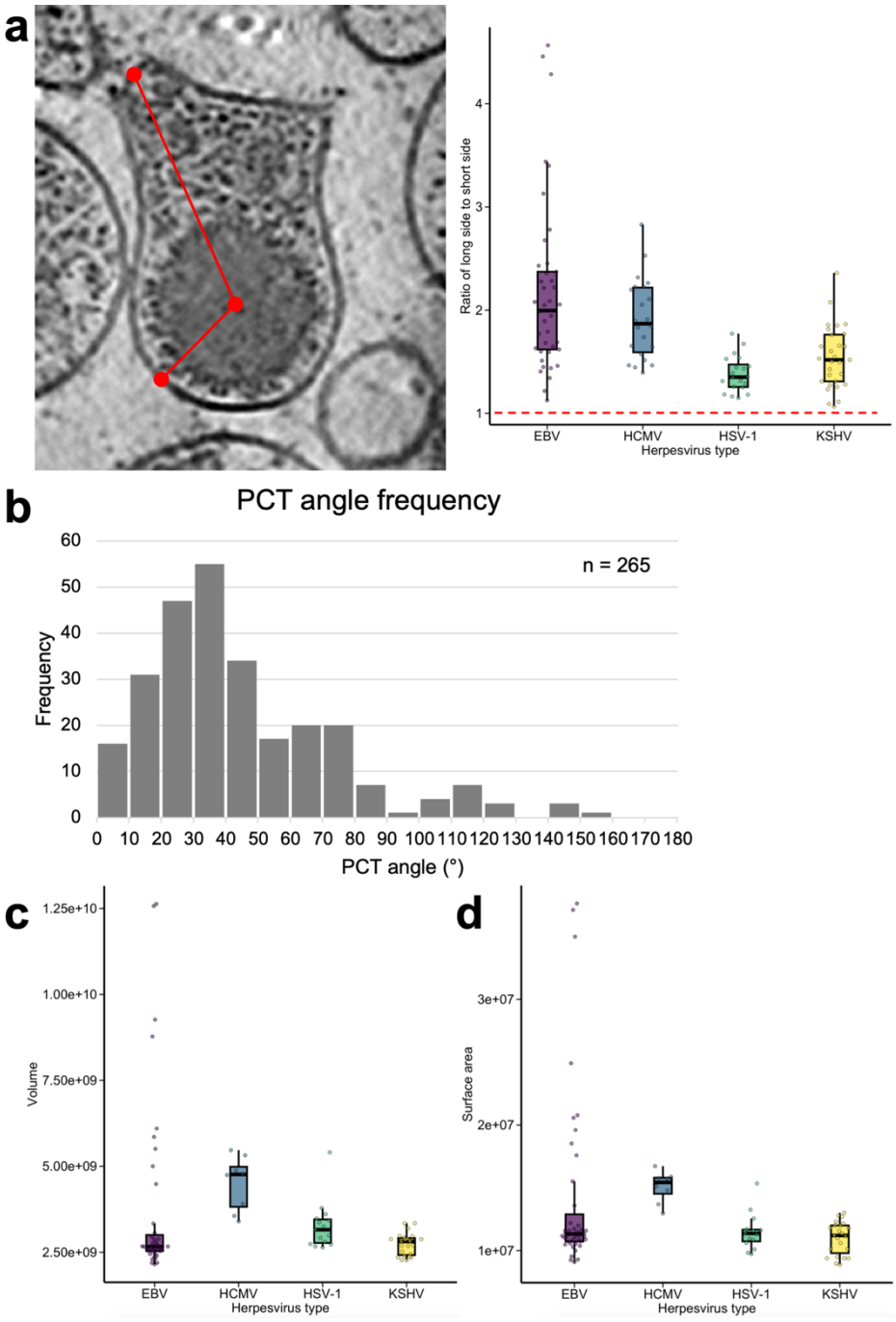
Supplementary Figure 4-S2. IsoNet processing of KSHV tomogram. (a) Weighted back-projection with a simultaneous iterative reconstructive technique (SIRT)-like filter tomogram displayed in XY, XZ, and YZ planes. **(b)** IsoNet processed tomogram displayed in XY, XZ, and YZ planes.



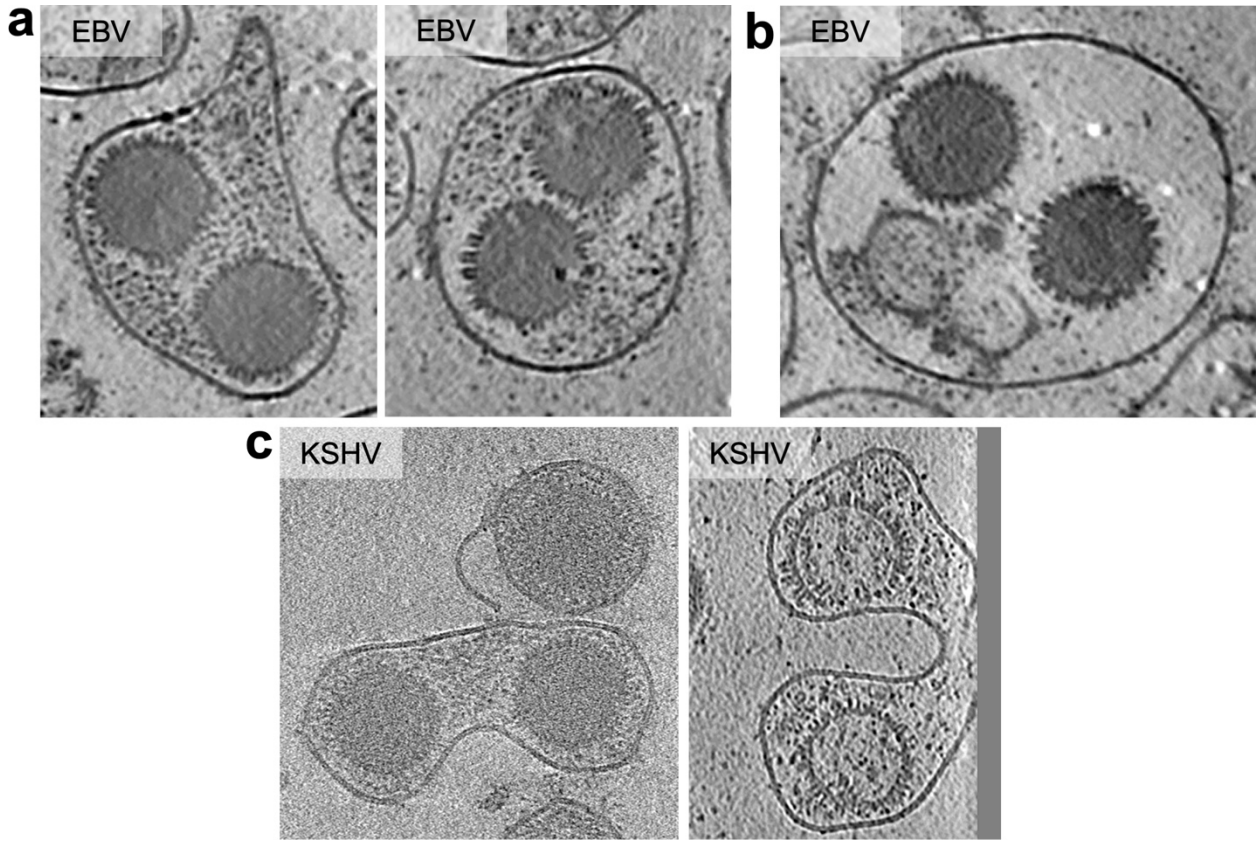
Supplementary Figure 4-S3. EBV processing workflow. Flowchart shows processing pipeline of tomogram-guided sub-particle reconstruction and reconstruction to resolve capsid vertex from icosahedral capsid.



Supplementary Figure 4-S4. KSHV processing workflow. Flowchart shows processing pipeline of tomogram-guided sub-particle reconstruction and reconstruction to resolve capsid vertex from icosahedral capsid. Subsequent focused classification separates distinct portal vertex and penton vertex reconstructions.



Supplementary Figure 4-S5. Herpesvirus subfamily morphology comparison. (a) Eccentricity of virion measured by ratio of long side to short side (left). Plot of ratios of long side to short side for virions of EBV, KSHV, HSV-1, and HCMV (right). Red dashed line indicates ratio for sphere for reference. **(b)** Histogram showing portal-capsid-tegument (PCT) angle frequency of KSHV virions. **(c)** Plot of volume ($Å^3$) encompassed by envelope of intact virions. (d) Plot of surface area ($Å^2$) of intact virion envelopes.



Supplementary Figure 4-S6. Non-standard morphologies observed in EBV and KSHV. (a) EBV particles containing two C-capsids within the same intact envelope and surrounded by tegument-like densities (2/65 virions). **(b)** EBV particle containing two C-capids within an envelope with low proteinaceous density (1/65 virions). **(c)** KSHV particles containing two C-capsids (left) or capsids devoid of genome (right) within the same intact envelope that is pinched inward between the capsids (2/600 virions).

Supplementary Table S1. CryoET data collection and processing statistics

Data collection	EBV	KSHV
Magnification	42,000 ×	64,000 ×
Voltage (kV)	300	300
Total electron exposure (e^{-}/Å^2)	120	120
Defocus target (μm)	-5	-4
Pixel size (Å)	2.088	1.380
Tilt series range (°)	±60	±60
Tilt increment (°)	3	3
Tilt scheme	Dose-symmetric	Dose-symmetric
Accepted tomograms (no.)	39	199

Data and processing	EBV capsid (EMD-46986)	EBV mixed vertex (EMD-46988)	KSHV capsid (EMD-46989)	KSHV mixed vertex (EMD-46990)	KSHV portal vertex (EMD-46987)	KSHV penton vertex (EMD-46991)
Symmetry	I_h	C_5	I_h	C_5	C_5	C_5
Final particles (no.)	45	540	600	7,177	554	3,815
Map resolution (Å)	24	21	17	7.1	9.8	7.4
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143

Chapter 5

Conclusion

To highlight the strength of cryoEM for structural characterization of systems in their native context, we applied endogenous cryoEM methods to two systems: (1) *in silico* purification and structural determination of αKDCs from bovine kidney mitochondrial lysate and (2) cryoET of intact cancer-causing EBV and KSHV virions, including a KSHV virion expressing a mutant gB. The approaches used here can be extrapolated to other systems where retention of molecular sociology can deepen our understanding.

By applying lysates directly to an EM grid for data collection, we determined the structure of two critical metabolic complexes, PDC and OGDC. By solving their native structures, we are able to observe the physiologically relevant clustered distribution of peripheral E1 and E3 subunits around the E2 core, contrary to the common symmetric depiction. From a functional perspective, the clustering is advantage for minimizing the distance that substrates need to be shuttled between subunits. By determining the two structures from the same organism, we are able to also directly compare them to identify the modes for correct self-assembly into their respective homogeneous icosahedrons and cubes.

In juxtaposition with the endogenous αKGDC structures, we used cryoEM to determine the structure of the recombinant KSHV gB ectodomain. Although the structure is in the postfusion conformation despite a prefusion-stabilizing mutation, it is nonetheless the first reported structure of KSHV gB. Our structure reveals a glycosylation pattern for KSHV gB that is distinct from those of other herpesviruses and resolves an extended glycan chain when resolving just the first sugar moiety is the norm. We also identify flexibility of the DIV crown

that may inform the mechanism of conformational change from prefusion state to postfusion state. Importantly, when the prefusion-stabilizing mutation is applied to gB within the context of the virion, we observe a diminished presence of postfusion gB. This discrepancy between the recombinant and endogenous data highlights the need for structural determination within the native context.

Through the application of cryoET, in addition to resolving symmetric structures to sub-nanometer resolution, we are able to study the pleiomorphic features of EBV and KSHV. In doing so, we observe interlayer interactions between the capsid, tegument, and envelope glycoproteins. We are also able to see congregations of the gB fusogen in triplet arrangements, which we could not observe when studying gB in isolation. Only by keeping the herpesvirus systems intact are we able to reveal the molecular interactions that underly their assembly and infection.

To gain even further context for a system, the next logical step is to take a step back to the cellular level: metabolic complexes within the mitochondria, viruses within the infected cell. However, due to the thickness of these samples, focused ion beam (FIB) milling is required to ablate samples into thin lamellae suitable for EM imaging. As methods innovate, FIB milling paired with cryoET becomes more accessible. Although the data collected from this combination is more robust in terms of context than single particle cryoEM *in silico* purification, cryoEM remains more accessible and is the higher throughput option for native structural determination for the time being. Most species in the proteome are too small for contemporary cryoEM methods. Due to electron dose limitations governing the contrast and visibility of particles, single particle cryoEM will benefit first from breakthroughs in imaging techniques. However, to observe non-repeating phenomena, the resolution of cryoET requires improvements at the

tomogram level. Additional interpretative burden will accompany the additional detail, which will necessitate improvements in automated annotation of cryoET data. Tandem improvements in both cryoEM hardware and software will be required to be able to see and fully appreciate our underlying molecular societies at large.