Structural Model of Bacteriophage T4

Victor Padilla-Sanchez¹

Abstract

Bacteriophage T4 is a virus that infects Escherichia coli, having dimensions of 90 nm in width and 200 nm in length (head and tail in extended form). It is a quite common model organism that has been studied for a century by many important virologists, and even Watson and Crick after their elucidation of DNA. Structural characterisation of the bacteriophage’s individual proteins began in the 1980s, and complexes of multiple proteins in the 1990s. However, it has not yet been possible to structurally characterise the complete phage in atomic detail (though some have begun to come closer) with multiple overall schematic models published. The increasing power of computers and the RCSB structural database have made possible the construction of a single combined model of the entire bacteriophage T4 organism with atomic resolution components as described here.

Keywords: computational biology, Escherichia coli, DNA, reconstruction

Introduction

The complete structural model for bacteriophage T4 has been constructed thanks to the determination of the structures of single proteins that constitute the virus as well as various parts of the virus. First, the capsid (head) of the virus was constructed using a 3D cryoEM reconstruction where each individual protein was fitted into the EM density: soc in orange, hoc in blue, gp23 in green, gp24 in magenta, and gp20 in red (though gp20 is hidden between the head and tail). A similar procedure was followed for the construction of the tail and tail fibers.[6][7] Although each of the component proteins has been described in detail during decades of research, this combined structural model, whilst not perfect, is the best reconstruction of the entire organism that we have today as of 2021.

Description

Further information: Wikipedia:Bacteriophage T4

This reconstruction largely used the molecular visualization software UCSF Chimera[8] in a personal computer, and where necessary for some tasks, in supercomputers like Bridges for Pittsburgh and Frontera from Texas. This combined structural model has been constructed and updated over time my work in Catholic University of America (where I started in 2007) as new structures of the components have been solved by other researchers, and therefore is the accumulated work of many years of research.

There are approximately 50 structural proteins that assemble the virus which is constructed with protein database (PDB) structures and one cryoEM reconstruction from the Electron Microscopy Data Bank (EMDB) corresponding to the brown ring between head and tail. This structure can be used for teaching at any level of education and research because is accurate at atomic resolution and therefore can be used to derive hypotheses (e.g., how antigen-display technology could be used with bacteriophage T4). The protein colours are chosen to differentiate between them and make contrast as well as showing art in science, but some colours are the same for different proteins (e.g., soc in the head and gp38 in the tail). As a reference for further publications, you can look at Bacteriophage T4 in Wikipedia where each protein is described in different research articles.

List of PDB and EMDB structures

- Head: hoc (PDB: 3shs - blue), soc (PDB: 5vff - orange), gp23 (PDB: 5vff - dark green), gp24 (PDB: 5vff - dark magenta), gp20 (PDB: 6ucn - dark red)
- Collar: brown (EM density from EMDB: 1075)
- Wac fibres: PDB: 2bsq - dark green

¹ Department of Biology, Catholic University of America: Washington, DC, US
*Author correspondence: jopadillasan@cua.edu
ORCID: [0000-0001-5759-1884]
Supplementary material: commons.wikimedia.org/location
Licensed under: CC-BY
Received 12-08-2020; accepted 05-08-2021
Figure 1 | Bacteriophage T4 Structural Model, Dr. Victor Padilla-Sanchez, PhD - Own work CC BY-SA 4.0

- Neck: brown gp33 and gp14 are represented by an EM density from EMDB: 1075
- Tail: g15 (PDB: 3jm - orange), gp18 (PDB: 3jm - orange/red)
- Tail base plate: g17 (PDB: 5jv5 - dark green/blue/dark red/purple)
- Long Tail Fibers: composites of gp37 (PDB: 2xgf) and gp34 (PDB: 4uxf) repeated to form the long tail fibre.

Note: same protein colours do not necessarily define same proteins.

Acknowledgements

1. This work used the Extreme Science and Engineering Discovery Environment (XSEDE),[5] which is supported by National Science Foundation grant number ACI-1548562. Specifically, it used the Bridges system,[10] which is supported by NSF award number ACI-1445566, at the Pittsburgh Supercomputing Center (PSC).

We would like to thank the many editors who contributed to the Wikipedia article from which this work is adapted. A full list of contributors can be found at this address: https://en.wikipedia.org/wiki/WikiJournal_of_Science/Structural_Model_of_Bacteriophage_T4

Competing interests: none declared.

References

4. Chen, Zhenguo; Sun, Lei; Zhang, Zhihong; Fokine, Andrej; Padilla-Sanchez, Victor; Hanein, Dorit; Jiang, Wen; Rossmann, Michael G et al. (2017-09-16). "Cryo-EM structure of the bacteriophage T4 isometric head at 3.3 Å resolution and its relevance to the assembly of icosahedral viruses". Proceedings of the National Academy of Sciences 114 (39): E8184–E8193. doi:10.1073/pnas.1708483114. ISSN 0027-8424. PMID 28899988. PMC PMC5625921.