AlphaFold Predictions in Viral Research

Subjects: Virology

Contributor: Daria Gutnik, Peter Evseev, Konstantin Miroshnikov, Mikhail Shneider

AlphaFold, a modern deep-learning algorithm, enables the prediction of protein structure to a high level of accuracy. It has been applied in numerous studies in various areas of biology and medicine. Viruses are biological entities infecting eukaryotic and procaryotic organisms. They can pose a danger for humans and economically significant animals and plants, but they can also be useful for biological control, suppressing populations of pests and pathogens. AlphaFold can be used for studies of molecular mechanisms of viral infection to facilitate several activities, including drug design.

Keywords: AlphaFold ; viruses ; bacteriophages

1. Introduction

Prediction of the three-dimensional structure of proteins is a difficult task. For a long time, the main prediction methods included comparative modelling (homology modelling), threading and ab initio and machine-learning approaches [1][2]. The development of end-to-end machine-learning approaches in recent years has resulted in the emergence of new techniques that can often outperform other methods [3][4]. Moreover, recent progress associated with deep-learning methods enables speculation about a revolution in protein-structure prediction [5]. One of the most popular deep-learning techniques is Alphabet–Google DeepMind's neural network-based end-to-end solution AlphaFold2 (AlphaFold, AF2), which was presented in the CASP14 competition [6], the second iteration of the AlphaFold system entered in CASP13 [7]. AlphaFold employs a deep-learning approach and a conventional neural network. This technique is able to predict the distance and torsion distribution of proteins, using training schemes of experimentally determined PDB structures, protein primary sequences and the multiple sequence alignment (MSA) of proteins. In CASP14, AlphaFold2 structures had a median backbone accuracy of 0.96 Å $RMSD_{95}$ (Cα root-mean-square deviation at 95% residue coverage) and an all-atom accuracy of 1.5 Å $RMSD_{95}$. The corresponding values for the prediction of the best alternative method were 2.8 Å and 3.5 Å [6]. The high level of accuracy of AlphaFold2 predictions boosted the popularity of this technique. One might even talk about "AlphaFold mania", given the astonishing increase in the number of journal articles and preprints citing AlphaFold2 AI software [8].

2. Application of AF2 for Research on Eukaryotic Viruses

2.1. Application of AlphaFold for SARS-CoV-2 Research

The outbreak of severe acute respiratory syndrome caused by coronavirus 2 (SARS-CoV-2, realm *Ribozyviria*, class *Pisoniviricetes*, order *Nidovirales*, family *Coronaviridae*, genus *Betacoronavirus*) and the spread of associated infection boosted research on coronaviruses. The structure of SARS-CoV-2 spike (S) glycoprotein, the main target of antibodies, has been determined by cryo-electron microscopy and was used in the development of vaccines and inhibitors [9][10]. S glycoprotein promotes entry into the cell. Another target of drug design is main protease cutting the initial translated propeptide into functional viral proteins. The crystal structure of the SARS-CoV-2 main protease was also obtained experimentally [11].

To assist the solution of tasks related to general research and drug design, different structure prediction techniques, including AlphaFold, were used for prediction of SARS-CoV-2 proteins [12][13][14][15][16][17]. The main task was probably the investigation of the mechanism of interaction of the SARS-CoV-2 receptor-binding protein (RBP), which is the SARS-CoV-2 spike, and the angiotensin-converting enzyme 2 (ACE2) receptor. AF2 predictions enabled clarification of the structural features of monomeric and multimeric formulations of the vaccine and suggested that monomeric formulation presents more antigenic epitopes [14]. The emergence of new immune-escaping variants of SARS-CoV-2, such as Omicron BA1, made it important to study potential mutation sites that do not yet exist in nature but could increase the binding affinity of RBD and the receptor [16]. AF2 predictions were successfully used to find an explanation for the observed reduction in the neutralisation of SARS-CoV-2 variants of concern compared with other variants [15]. AF2 predictions can be combined with molecular dynamics simulations to improve modelling accuracy [18] and to predict the physical properties of proteins. Such

models can be used for studies of both qualitative and quantitative aspects of the formation of the quaternary structure of proteins [17]. AlphaFold models are useful for revealing possible ligand binding sites. Together with virtual screening and in silico validation, these approaches provide the basis for the biological testing of new drugs and for the repurposing of natural products [12].

The accuracy of predicted structures can be assessed using computational techniques [19] and via experimental methods, e.g., optical spectroscopy or measurement of solution residual dipolar couplings data (RDCs) [20][21]. A meticulous evaluation of the concordance of AF2 models of the SARS-CoV-2 homodimeric 3C-like protease (M^{pro}) with residual dipolar couplings (RDCs) measured in solution for $^{15}N–^{1}H^{N}$ and $^{13}C'–^{1}H^{N}$ atom pairs indicated the close agreement of AlphaFold predictions with experimental data [20].

Interestingly, the high level of accuracy of AF2 predictions makes it possible to use AlphaFold predictions to determine a macromolecular structure from crystallographic diffraction experiments. It has been shown that a template-free AF2 model, generated by the AlphaFold2 group, was of sufficient quality to phase the native SARS-CoV-2 ORF8 dataset by molecular replacement, overcoming the limitations of the crystallographic phasing problem [13]. However, a comparison of RMSD (root mean square deviation of atomic positions) values of SARS-CoV-2 spike RBD, the laboratory-derived structure with both trRosetta-generated models [22] and models generated by AlphaFold v2.1.0, indicated the high level of accuracy of both methods, but the better results were obtained with trRosetta.

2.2. Application of AlphaFold to Study Eukaryotic Viruses

AlphaFold is widely used in research on other eukaryotic viruses, including monkeypox virus (MPXV) [23][24][25][26], herpes simplex virus [27][28], hepatitis E virus (HEV) [29] and other viral pathogens of humans and economically significant animals and plants [30][31][32][33][34][35]. Monkeypox virus (MPXV) represents a new serious threat to human health. MPXV has spread to 110 countries (https://www.cdc.gov/poxvirus/mpox/response/2022/world-map.html, accessed on 1 March 2023). As of 1 March 2023, there were 86,231 confirmed cases worldwide, of which 84,858 cases occurred in locations that had not previously reported MPXV cases. Monkeypox virus is classified as a member of realm *Varidnaviria*, class *Pokkesviricetes*, order *Chitovirales*, family *Poxviridae*, genus *Orthopoxvirus* and is evolutionarily close to vaccinia virus (VACV), the smallpox virus. AlphaFold-derived structures of the recombinantly expressed MPXV antigen truncations to their VACV homologues have indicated that MPXV and VACV antigens are likely to achieve similar conformations [26]. The World Health Organisation (WHO) has recommended the current anti-smallpox drugs tecovirimat, brincidofovir and cidofovir for the treatment of monkeypox [36]. Brincidofovir and cidofovir inhibit DNA polymerase (DNAP), while tecovirimat is an inhibitor for poxvirus phospholipase D (protein F13) [37], but specific antiviral treatment requires new drugs.

MPXV DNA polymerase (DNAP) is a very important antiviral drug target. The laboratory-derived structure of MPXV DNAP was deposited in the RCSB PDB database (PDB code 8HG1) in mid-November 2022, and a paper describing this structure was published in January 2023 [38]. Before that, the AF2-derived structure was obtained and used in the search and design of new inhibitors of MPXV DNAP. The molecules found were predicted to bind to the MPXV DNAP with a binding energy comparable to that of brincidofovir and cidofovir. New MPXV DNAP inhibitors are important in the context of possible drug resistance, which can arise due to mutations in proteins of the DNA replication complex (RC). Studies of the effect of mutations in MPXV RC using AF2-generated models have suggested similar mechanisms of drug resistance to cidofovir in monkeypox and vaccinia viruses [24]. It appears that the use of highly accurate AlphaFold predictions can assist the forecasting of the emergence of drug-resistant variants of concern to improve preparedness for them.

The molecular mechanism of interaction of tecovirimat with the monkeypox phospholipase D (F13) was studied using AlphaFold models and molecular dynamics simulations [25]. The results suggested a detailed mechanism of inhibition of F13 by tecovirimat and supported the efficacy of tecovirimat against monkeypox virus, emphasising the importance of the availability of precise modelling for revealing molecular mechanisms of drug action.

The development of new drugs is barely possible without an understanding of the mechanisms of viral infection. This knowledge can often require robust structural analysis, which can make use of modern deep-learning structure prediction methods. AlphaFold can facilitate the elucidation of the functionality of viral proteins.

Herpesviruses constitute an important group of pathogens that infect animals, including humans. Herpesviruses infect most vertebrates, causing a lifelong latent infection [39]. Herpesviruses belong to the realm *Duplodnaviria*, class *Herviviricetes*, order *Herpesvirales*, and comprise the families *Alloherpesviridae*, *Herpesviridae* and *Malacoherpesviridae* [40]. Human herpesviruses belong to the family *Herpesviridae*. Herpes simplex virus 1 (HSV-1) (genus *Alphaherpesviruse*), residing in sensory neurons or sympathetic neurons, has been shown to severely modify infected cells and to remodel the composition and architecture of cellular membranes [27][41][42]. One of the HSV-1 proteins, phosphatase adaptor UL21,

mediates dephosphorylation and accelerates the rate of ceramide to sphingomyelin conversion, altering cell membranes and influencing viral replication [27]. AlphaFold-Multimer modelling has revealed the details of the interaction of UL21 and viral protein UL16 and has enabled the suggestion of the functionality of domains of the latter protein using its structural features. Specific protein–protein interactions have been shown to be essential for lipid metabolism [27]. The use of AlphaFold has also shown that another HSV-1 protein, the tegument protein UL37, interacts with the cytoplasmic surface of the lipid membrane, suggesting that UL37 can be a peripheral membrane protein [28]. AlphaFold predictions have suggested the domain organisation of UL37, and assisted experimental studies and molecular dynamics simulation have clarified the structural features and molecular mechanisms of UL37 interactions.

Fundamentally similar tasks concerning research on other viral pathogens of animals, including humans, and plants can be made easier by the use of AlphaFold predictions. These tasks include mechanisms that are crucial for viral attachment, penetration, replication, release and other steps in the viral infection cycle. They can include the investigation of viral proteins and membranes [30][33][35], viral proteins and DNA [31] and studies of viral proteins, glycoproteins and their mutations [29][32][34]. It is noteworthy that AlphaFold predictions are often used as part of an integrated approach, making the planning of experiments easier and improving understanding of the results obtained.

3. Application of AlphaFold for Research on Bacteriophages

Bacteriophages (a.k.a. phages) are viruses that infect and replicate in bacterial cells alone. Bacteriophages are ubiquitous —they can be found in water, soil and various living organisms [43]. The total number of bacteriophages can be estimated at 10^{31} viral particles, which is 10–100 times the number of cells [44]. The total mass of these particles is about a trillion tons [45]. Phages are also members of plant and animal microbiomes, including humans. For example, the human gastrointestinal tract contains more than 10^{12} phage virions [46]. The ability of bacteriophages to destroy the cells of pathogenic bacteria attracted the attention of scientists as early as the beginning of the 20th century. In recent decades, interest in bacteriophage therapy has begun to grow, primarily due to the spread of antibiotic resistance. Phage therapy has important advantages [47], including sustained bactericidal activity and “autodosing”, wherein the number of phages positively correlates with the number of host bacteria. Furthermore, phages have low intrinsic toxicity, and phage therapy is characterised by minimal disruption of normal flora and the lack of cross-resistance with antibiotics.

The practical use of phages for phage therapy requires an understanding of the structural bases of interactions of the host receptor and phage receptor-binding proteins (RBPs); the latter can include tail fibre and tail spike proteins (TFP and TSP). In addition, phage RBPs, as well as endolysins and ectolysins, the proteins that cause cell lysis, can be used as antibacterial agents by themselves [48][49]. The analysis of the structural features of phage RBPs and lysins can use modern deep-learning techniques, including AlphaFold. Together with experimental studies, AlphaFold predictions can be used to elucidate the domain organisation of TFP, TSP and cell-wall degrading enzymes, to reveal the sites of phage particle binding and enzymatic domains [48][50][51][52].

As well as in the case of eukaryotic viruses mentioned above, AlphaFold predictions can contribute to building the model of the viral particle [53][54] or the virion parts, including the attachment apparatus [50][55] and phage egress machinery [56]. All the steps of phage infection are accompanied by macromolecular interactions that include proteins, so AlphaFold’s highly accurate structural predictions can assist in the elucidation of the mechanisms of the formation of the phage nucleus [57], lysogeny maintenance [58] or anti-phage defence [59][60]. AlphaFold can also be useful in the trivial but relevant task of phage genome annotation, assisting the prediction of genes’ functionality. As of January 2023, 19,499 GenBank sequences, assigned to class *Caudoviricetes*, contained 1,731,815 coding regions, 67% of which were annotated as hypothetical proteins. In some cases, BLAST search and HMM-HMM motif comparisons fail to assign a function to proteins encoded in phage genomes, but analysis of fold of AF2-derived structures can assist to clarify this function [61].

It seems that no large-scale studies have been published on the accuracy of modelling using AF2 compared with the predictions of other algorithms. However, comparing the predicted average local distance difference test (lDDT) score of the 54 AF2-derived models of the major capsid protein and ATPase subunit of phage terminase indicated an impressive level of accuracy of the predictions [61]. Interestingly, structural predictions of more conserved terminase were more accurate than those of major capsid protein, (terminase lDDT mean: 0.988, median: 0.996; major capsid protein lDDT mean: 0.907, median: 0.929). The average lDDT of the ATPase domains extracted from the ATPase subunit of phage terminase models was even higher (mean: 0.998, median: 0.999). An evaluation of models of the same major capsid proteins, carried out using a different deep-learning algorithm, RoseTTAFold, showed a lower accuracy of prediction (lDDT mean: 0.634, median: 0.649) than with the AlphaFold models.

4. Application of AlphaFold for Evolutionary and Taxonomic Studies

Comparing structural similarity and specific structural features can clarify the evolutionary relationships between proteins. Furthermore, the emergence of new high-precision algorithms for predicting the structure of proteins, including AlphaFold, can enable the identification of evolutionary relationships between highly divergent discovered proteins, using the results of structural modelling. The evolution of proteins may be accompanied by the appearance of new domains, and comparative analysis of AF2-derived structures can help reveal patterns of protein evolution. Studies of bacteriophage tail sheath proteins, an important part of phages' contractile injection system, have enabled the identification of the common core domain, including both N-terminal and C-terminal parts. The remaining variable parts consisting of one or more moderately conserved domains have, presumably, been added during phage evolution [62].

Structural similarity is widely used to evaluate evolutionary relationships between proteins whose amino acid sequence homology level is low or cannot be determined at all [63][64]. The structural similarity between two proteins can be assessed using root-mean-square deviation (RMSD) or other metrics such as template modelling score (TM-score) and DALI Z-score; the latter two metrics have a number of advantages over RMSD [11][65]. Clustering of experimentally determined structures of major capsid proteins using the DALI Z-score has already been used to illustrate the common origin of some viral groups and to cluster prokaryotic viruses [63][66]. Integrated use of both experimental structures and AF2-derived structures can be used for elucidation of evolutionary relationships and taxonomic classification of bacteriophages and eukaryotic viruses [67][68][69]. AlphaFold modelling and subsequent clustering have been used in taxonomic studies of archaeal viruses [66]. Clustering using AlphaFold showed interesting and often biologically meaningful results [61]. Clustering using structures predicted by AlphaFold showed interesting and often biologically meaningful results. It should also be noted that the native state of viral proteins can change according to the state of the viral particle (e.g., empty, full, expanded capsids) and according to the stage of viral particle assembly [70][71][72][73]. The correlation between structural similarity and sequence identity is not absolute due to conformational plasticity, solvent effects and ligand binding [74]. Most of these limitations apply to studies that involve experimentally determined structures, but, hypothetically, they could be exacerbated by structural prediction errors. Therefore, predicting the effectiveness of using AlphaFold for the analysis of structural similarity and evolutionary history, based only on the similarity of the predicted structures, seems to be a difficult task [61].

References

1. Agnihotry, S.; Pathak, R.K.; Singh, D.B.; Tiwari, A.; Hussain, I. Chapter 11—Protein Structure Prediction. In Bioinformatics; Singh, D.B., Pathak, R.K., Eds.; Academic Press: Cambridge, USA, 2022; pp. 177–188. ISBN 978-0-323-89775-4.
2. Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. Nat. Rev. Mol. Cell. Biol. 2019, 20, 681–697.
3. Marcu, Ş.-B.; Tăbîrcă, S.; Tangney, M. An Overview of Alphafold's Breakthrough. Front. Artif. Intell. 2022, 5, 875587.
4. Dhingra, S.; Sowdhamini, R.; Cadet, F.; Offmann, B. A Glance into the Evolution of Template-Free Protein Structure Prediction Methodologies. Biochimie 2020, 175, 85–92.
5. Bouatta, N.; AlQuraishi, M. Structural Biology at the Scale of Proteomes. Nat. Struct. Mol. Biol. 2023, 30, 129–130.
6. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. Nature 2021, 596, 583–589.
7. AlQuraishi, M. AlphaFold at CASP13. Bioinformatics 2019, 35, 4862–4865.
8. Callaway, E. What's next for AlphaFold and the AI Protein-Folding Revolution. Nature 2022, 604, 234–238.
9. Walls, A.C.; Park, Y.-J.; Tortorici, M.A.; Wall, A.; McGuire, A.T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell 2020, 181, 281–292.e6.
10. Han, Y.; Král, P. Computational Design of ACE2-Based Peptide Inhibitors of SARS-CoV-2. ACS Nano 2020, 14, 5143–5147.
11. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved α-Ketoamide Inhibitors. Science 2020, 368, 409–412.
12. Abdelkader, A.; Elzemrany, A.A.; El-Nadi, M.; Elsabbagh, S.A.; Shehata, M.A.; Eldehna, W.M.; El-Hadidi, M.; Ibrahim, T.M. In-Silico Targeting of SARS-CoV-2 NSP6 for Drug and Natural Products Repurposing. Virology 2022, 573, 96–110.

13. Flower, T.G.; Hurley, J.H. Crystallographic Molecular Replacement Using an in Silico-Generated Search Model of SARS-CoV-2 ORF8. Protein Sci. 2021, 30, 728–734.

14. Jansen van Vuren, P.; McAuley, A.J.; Kuiper, M.J.; Singanallur, N.B.; Bruce, M.P.; Riddell, S.; Goldie, S.; Mangalaganesh, S.; Chahal, S.; Drew, T.W.; et al. Highly Thermotolerant SARS-CoV-2 Vaccine Elicits Neutralising Antibodies against Delta and Omicron in Mice. Viruses 2022, 14, 800.

15. Singanallur, N.B.; van Vuren, P.J.; McAuley, A.J.; Bruce, M.P.; Kuiper, M.J.; Gwini, S.M.; Riddell, S.; Goldie, S.; Drew, T.W.; Blasdell, K.R.; et al. At Least Three Doses of Leading Vaccines Essential for Neutralisation of SARS-CoV-2 Omicron Variant. Front. Immunol. 2022, 13, 883612.

16. Bhowmick, S.; Jing, T.; Wang, W.; Zhang, E.Y.; Zhang, F.; Yang, Y. In Silico Protein Folding Prediction of COVID-19 Mutations and Variants. Biomolecules 2022, 12, 1665.

17. Cao, Y.; Yang, R.; Wang, W.; Jiang, S.; Yang, C.; Liu, N.; Dai, H.; Lee, I.; Meng, X.; Yuan, Z. Probing the Formation, Structure and Free Energy Relationships of M Protein Dimers of SARS-CoV-2. Comput. Struct. Biotechnol. J. 2022, 20, 573–582.

18. Heo, L.; Feig, M. High-accuracy protein structures by combining machine-learning with physics-based refinement. Proteins 2020, 88, 637–642.

19. Hiranuma, N.; Park, H.; Baek, M.; Anishchenko, I.; Dauparas, J.; Baker, D. Improved Protein Structure Refinement Guided by Deep Learning Based Accuracy Estimation. Nat. Commun. 2021, 12, 1340.

20. Robertson, A.J.; Courtney, J.M.; Shen, Y.; Ying, J.; Bax, A. Concordance of X-Ray and AlphaFold2 Models of SARS-CoV-2 Main Protease with Residual Dipolar Couplings Measured in Solution. J. Am. Chem. Soc. 2021, 143, 19306–19310.

21. Li, Z.; Hirst, J.D. Computed Optical Spectra of SARS-CoV-2 Proteins. Chem. Phys. Lett. 2020, 758, 137935.

22. Du, Z.; Su, H.; Wang, W.; Ye, L.; Wei, H.; Peng, Z.; Anishchenko, I.; Baker, D.; Yang, J. The TrRosetta Server for Fast and Accurate Protein Structure Prediction. Nat. Protoc. 2021, 16, 5634–5651.

23. Kumari, S.; Chakraborty, S.; Ahmad, M.; Kumar, V.; Tailor, P.B.; Biswal, B.K. Identification of Probable Inhibitors for the DNA Polymerase of the Monkeypox Virus through the Virtual Screening Approach. Int. J. Biol. Macromol. 2023, 229, 515–528.

24. Kannan, S.R.; Sachdev, S.; Reddy, A.S.; Kandasamy, S.L.; Byrareddy, S.N.; Lorson, C.L.; Singh, K. Mutations in the Monkeypox Virus Replication Complex: Potential Contributing Factors to the 2022 Outbreak. J. Autoimmun. 2022, 133, 102928.

25. Li, D.; Liu, Y.; Li, K.; Zhang, L. Targeting F13 from Monkeypox Virus and Variola Virus by Tecovirimat: Molecular Simulation Analysis. J. Infect. 2022, 85, e99–e101.

26. Yefet, R.; Friedel, N.; Tamir, H.; Polonsky, K.; Mor, M.; Cherry-Mimran, L.; Taleb, E.; Hagin, D.; Sprecher, E.; Israely, T.; et al. Monkeypox Infection Elicits Strong Antibody and B Cell Response against A35R and H3L Antigens. iScience 2023, 26, 105957.

27. Benedyk, T.H.; Connor, V.; Caroe, E.R.; Shamin, M.; Svergun, D.I.; Deane, J.E.; Jeffries, C.M.; Crump, C.M.; Graham, S.C. Herpes Simplex Virus 1 Protein PUL21 Alters Ceramide Metabolism by Activating the Interorganelle Transport Protein CERT. J. Biol. Chem. 2022, 298, 102589.

28. Collantes, T.M.A.; Clark, C.M.; Musarrat, F.; Jambunathan, N.; Jois, S.; Kousoulas, K.G. Predicted Structure and Functions of the Prototypic Alphaherpesvirus Herpes Simplex Virus Type-1 UL37 Tegument Protein. Viruses 2022, 14, 2189.

29. Fieulaine, S.; Tubiana, T.; Bressanelli, S. De Novo Modelling of HEV Replication Polyprotein: Five-Domain Breakdown and Involvement of Flexibility in Functional Regulation. Virology 2023, 578, 128–140.

30. Liu, H.; Peck, X.Y.; Choong, Y.K.; Ng, W.S.; Engl, W.; Raghuvamsi, P.V.; Zhao, Z.W.; Anand, G.S.; Zhou, Y.; Sivaraman, J.; et al. Identification of Putative Binding Interface of PI(3,5)P2 Lipid on Rice Black-Streaked Dwarf Virus (RBSDV) P10 Protein. Virology 2022, 570, 81–95.

31. Chen, L.; Chen, L.; Chen, H.; Zhang, H.; Dong, P.; Sun, L.; Huang, X.; Lin, P.; Wu, L.; Jing, D.; et al. Structural Insights into the CP312R Protein of the African Swine Fever Virus. Biochem. Biophys. Res. Commun. 2022, 624, 68–74.

32. Kim, S.Y.; Kwak, J.S.; Jung, W.; Kim, M.S.; Kim, K.H. Compensatory Mutations in the Matrix Protein of Viral Hemorrhagic Septicemia Virus (VHSV) Genotype IVa in Response to Artificial Mutation of Two Amino Acids (D62A E181A). Virus Res. 2023, 326, 199067.

33. Veit, M.; Gadalla, M.R.; Zhang, M. Using Alphafold2 to Predict the Structure of the Gp5/M Dimer of Porcine Respiratory and Reproductive Syndrome Virus. Int. J. Mol. Sci. 2022, 23, 13209.

34. Hötzel, I. Domain Organization of Lentiviral and Betaretroviral Surface Envelope Glycoproteins Modeled with AlphaFol d. J. Virol. 2022, 96, e01348-21.

35. Weaver, G.C.; Arya, R.; Schneider, C.L.; Hudson, A.W.; Stern, L.J. Structural Models for Roseolovirus U20 And U21: N on-Classical MHC-I Like Proteins From HHV-6A, HHV-6B, and HHV-7. Front. Immunol. 2022, 13, 864898.

36. Rizk, J.G.; Lippi, G.; Henry, B.M.; Forthal, D.N.; Rizk, Y. Prevention and Treatment of Monkeypox. Drugs 2022, 82, 957 –963.

37. Delaune, D.; Iseni, F. Drug Development against Smallpox: Present and Future. Antimicrob. Agents Chemother. 2020, 64, e01683-19.

38. Peng, Q.; Xie, Y.; Kuai, L.; Wang, H.; Qi, J.; Gao, G.F.; Shi, Y. Structure of Monkeypox Virus DNA Polymerase Holoenzy me. Science 2023, 379, 100–105.

39. Sehrawat, S.; Kumar, D.; Rouse, B.T. Herpesviruses: Harmonious Pathogens but Relevant Cofactors in Other Disease s? Front. Cell. Infect. Microbiol. 2018, 8, 177.

40. Current ICTV Taxonomy Release | ICTV. Available online: https://ictv.global/taxonomy (accessed on 9 November 2022).

41. Nahas, K.L.; Connor, V.; Scherer, K.M.; Kaminski, C.F.; Harkiolaki, M.; Crump, C.M.; Graham, S.C. Near-Native State I maging by Cryo-Soft-X-Ray Tomography Reveals Remodelling of Multiple Cellular Organelles during HSV-1 Infection. PLOS Pathog. 2022, 18, e1010629.

42. Bigalke, J.M.; Heldwein, E.E. Nuclear Exodus: Herpesviruses Lead the Way. Annu. Rev. Virol. 2016, 3, 387–409.

43. Wommack, K.E.; Colwell, R.R. Virioplankton: Viruses in Aquatic Ecosystems. Microbiol. Mol. Biol. Rev. 2000, 64, 69–11 4.

44. Simmonds, P.; Adams, M.J.; Benkő, M.; Breitbart, M.; Brister, J.R.; Carstens, E.B.; Davison, A.J.; Delwart, E.; Gorbalen ya, A.E.; Harrach, B.; et al. Consensus Statement: Virus Taxonomy in the Age of Metagenomics. Nat. Rev. Microbiol. 20 17, 15, 161–168.

45. Hendrix, R.W.; Smith, M.C.M.; Burns, R.N.; Ford, M.E.; Hatfull, G.F. Evolutionary Relationships among Diverse Bacteri ophages and Prophages: All the World's a Phage. Proc. Natl. Acad. Sci. USA 1999, 96, 2192–2197.

46. Shkoporov, A.N.; Hill, C. Bacteriophages of the Human Gut: The "Known Unknown" of the Microbiome. Cell Host Micro be 2019, 25, 195–209.

47. Loc-Carrillo, C.; Abedon, S.T. Pros and Cons of Phage Therapy. Bacteriophage 2011, 1, 111–114.

48. Klumpp, J.; Dunne, M.; Loessner, M.J. A Perfect Fit: Bacteriophage Receptor-Binding Proteins for Diagnostic and Ther apeutic Applications. Curr. Opin. Microbiol. 2023, 71, 102240.

49. Fischetti, V.A. Bacteriophage Endolysins: A Novel Anti-Infective to Control Gram-Positive Pathogens. Int J. Med. Microb iol. 2010, 300, 357–362.

50. Goulet, A.; Cambillau, C. Structure and Topology Prediction of Phage Adhesion Devices Using AlphaFold2: The Case o f Two Oenococcus Oeni Phages. Microorganisms 2021, 9, 2151.

51. Evseev, P.; Lukianova, A.; Tarakanov, R.; Tokmakova, A.; Popova, A.; Kulikov, E.; Shneider, M.; Ignatov, A.; Miroshniko v, K. Prophage-Derived Regions in Curtobacterium Genomes: Good Things, Small Packages. Int. J. Mol. Sci. 2023, 24, 1586.

52. Eskenazi, A.; Lood, C.; Wubbolts, J.; Hites, M.; Balarjishvili, N.; Leshkasheli, L.; Askilashvili, L.; Kvachadze, L.; van Noo rt, V.; Wagemans, J.; et al. Combination of Pre-Adapted Bacteriophage Therapy and Antibiotics for Treatment of Fractur e-Related Infection Due to Pandrug-Resistant Klebsiella Pneumoniae. Nat. Commun. 2022, 13, 302.

53. Hawkins, N.C.; Kizziah, J.L.; Hatoum-Aslan, A.; Dokland, T. Structure and Host Specificity of Staphylococcus Epidermid is Bacteriophage Andhra. Sci. Adv. 2022, 8, eade0459.

54. Ouyang, R.; Costa, A.R.; Cassidy, C.K.; Otwinowska, A.; Williams, V.C.J.; Latka, A.; Stansfeld, P.J.; Drulis-Kawa, Z.; Bri ers, Y.; Pelt, D.M.; et al. High-Resolution Reconstruction of a Jumbo-Bacteriophage Infecting Capsulated Bacteria Usin g Hyperbranched Tail Fibers. Nat. Commun. 2022, 13, 7241.

55. Šiborová, M.; Füzik, T.; Procházková, M.; Nováček, J.; Benešík, M.; Nilsson, A.S.; Plevka, P. Tail Proteins of Phage SU 10 Reorganize into the Nozzle for Genome Delivery. Nat. Commun. 2022, 13, 5622.

56. Conners, R.; McLaren, M.; Łapińska, U.; Sanders, K.; Stone, M.R.L.; Blaskovich, M.A.T.; Pagliara, S.; Daum, B.; Rakon jac, J.; Gold, V.A.M. CryoEM Structure of the Outer Membrane Secretin Channel PIV from the F1 Filamentous Bacterio phage. Nat. Commun. 2021, 12, 6316.

57. Nieweglowska, E.S.; Brilot, A.F.; Méndez-Moran, M.; Kokontis, C.; Baek, M.; Li, J.; Cheng, Y.; Baker, D.; Bondy-Denom y, J.; Agard, D.A. The ΦPA3 Phage Nucleus Is Enclosed by a Self-Assembling 2D Crystalline Lattice. Nat. Commun. 20

23, 14, 927.

58. McGinnis, R.J.; Brambley, C.A.; Stamey, B.; Green, W.C.; Gragg, K.N.; Cafferty, E.R.; Terwilliger, T.C.; Hammel, M.; Hollis, T.J.; Miller, J.M.; et al. A Monomeric Mycobacteriophage Immunity Repressor Utilizes Two Domains to Recognize an Asymmetric DNA Sequence. Nat. Commun. 2022, 13, 4105.
59. Al-Shayeb, B.; Skopintsev, P.; Soczek, K.M.; Stahl, E.C.; Li, Z.; Groover, E.; Smock, D.; Eggers, A.R.; Pausch, P.; Cress, B.F.; et al. Diverse Virus-Encoded CRISPR-Cas Systems Include Streamlined Genome Editors. Cell 2022, 185, 4574–4586.e16.
60. Zhang, T.; Tamman, H.; Coppieters 't Wallant, K.; Kurata, T.; LeRoux, M.; Srikant, S.; Brodiazhenko, T.; Cepauskas, A.; Talavera, A.; Martens, C.; et al. Direct Activation of a Bacterial Innate Immune System by a Viral Capsid Protein. Nature 2022, 612, 132–140.
61. Evseev, P.; Gutnik, D.; Shneider, M.; Miroshnikov, K. Use of an Integrated Approach Involving AlphaFold Predictions for the Evolutionary Taxonomy of Duplodnaviria Viruses. Biomolecules 2023, 13, 110.
62. Evseev, P.; Shneider, M.; Miroshnikov, K. Evolution of Phage Tail Sheath Protein. Viruses 2022, 14, 1148.
63. Krupovic, M.; Koonin, E.V. Multiple Origins of Viral Capsid Proteins from Cellular Ancestors. Proc. Natl. Acad. Sci. USA 2017, 114, E2401–E2410.
64. Salemme, F.R.; Miller, M.D.; Jordan, S.R. Structural Convergence during Protein Evolution. Proc. Natl. Acad. Sci. USA 1977, 74, 2820–2824.
65. Holm, L. Using Dali for Protein Structure Comparison. Methods Mol. Biol. 2020, 2112, 29–42.
66. Liu, Y.; Demina, T.A.; Roux, S.; Aiewsakun, P.; Kazlauskas, D.; Simmonds, P.; Prangishvili, D.; Oksanen, H.M.; Krupovic, M. Diversity, Taxonomy, and Evolution of Archaeal Viruses of the Class Caudoviricetes. PLOS Biol. 2021, 19, e3001442.
67. Podgorski, J.M.; Freeman, K.; Gosselin, S.; Huet, A.; Conway, J.F.; Bird, M.; Grecco, J.; Patel, S.; Jacobs-Sera, D.; Hatfull, G.; et al. A Structural Dendrogram of the Actinobacteriophage Major Capsid Proteins Provides Important Structural Insights into the Evolution of Capsid Stability. Structure 2023, 31, 282–294.e5.
68. Hötzel, I. Deep-Time Structural Evolution of Retroviral and Filoviral Surface Envelope Proteins. J. Virol. 2022, 96, e00063-22.
69. Bisio, H.; Legendre, M.; Giry, C.; Philippe, N.; Alempic, J.-M.; Jeudy, S.; Abergel, C. Evolution of Giant Pandoravirus Revealed by CRISPR/Cas9. Nat. Commun. 2023, 14, 428.
70. Fokine, A.; Leiman, P.G.; Shneider, M.M.; Ahvazi, B.; Boeshans, K.M.; Steven, A.C.; Black, L.W.; Mesyanzhinov, V.V.; Rossmann, M.G. Structural and Functional Similarities between the Capsid Proteins of Bacteriophages T4 and HK97 Point to a Common Ancestry. Proc. Natl. Acad. Sci. USA 2005, 102, 7163–7168.
71. Fang, Q.; Tang, W.-C.; Fokine, A.; Mahalingam, M.; Shao, Q.; Rossmann, M.G.; Rao, V.B. Structures of a Large Prolate Virus Capsid in Unexpanded and Expanded States Generate Insights into the Icosahedral Virus Assembly. Proc. Natl. Acad. Sci. USA 2022, 119, e2203272119.
72. Steven, A.C.; Greenstone, H.L.; Booy, F.P.; Black, L.W.; Ross, P.D. Conformational Changes of a Viral Capsid Protein. Thermodynamic Rationale for Proteolytic Regulation of Bacteriophage T4 Capsid Expansion, Co-Operativity, and Super-Stabilization by Soc Binding. J. Mol. Biol. 1992, 228, 870–884.
73. Bowman, B.R.; Baker, M.L.; Rixon, F.J.; Chiu, W.; Quiocho, F.A. Structure of the Herpesvirus Major Capsid Protein. EMBO J. 2003, 22, 757–765.
74. Hark Gan, H.; Perlow, R.A.; Roy, S.; Ko, J.; Wu, M.; Huang, J.; Yan, S.; Nicoletta, A.; Vafai, J.; Sun, D.; et al. Analysis of Protein Sequence/Structure Similarity Relationships. Biophys. J. 2002, 83, 2781–2791.