



# Present Impact of AlphaFold2 Revolution on Structural Biology, and an Illustration With the Structure Prediction of the Bacteriophage J-1 Host Adhesion Device

Adeline Goulet, Christian Cambillau

## ► To cite this version:

Adeline Goulet, Christian Cambillau. Present Impact of AlphaFold2 Revolution on Structural Biology, and an Illustration With the Structure Prediction of the Bacteriophage J-1 Host Adhesion Device. *Frontiers in Molecular Biosciences*, 2022, 9, pp.907452. 10.3389/fmolb.2022.907452 . hal-03832765

**HAL Id: hal-03832765**

**<https://hal.science/hal-03832765>**

Submitted on 27 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Present Impact of AlphaFold2 Revolution on Structural Biology, and an Illustration With the Structure Prediction of the Bacteriophage J-1 Host Adhesion Device

Adeline Goulet<sup>1\*</sup> and Christian Cambillau<sup>1,2\*</sup>

<sup>1</sup>Laboratoire D'Ingénierie des Systèmes Macromoléculaires (LISM), Institut de Microbiologie, Aix-Marseille Université—CNRS, Marseille, France, <sup>2</sup>School of Microbiology, University College Cork, Cork, Ireland

## OPEN ACCESS

### Edited by:

Andrea Mozzarelli,  
University of Parma, Italy

### Reviewed by:

Martino Bolognesi,  
University of Milan, Italy

### \*Correspondence:

Adeline Goulet  
adeline.goulet@univ-amu.fr  
Christian Cambillau  
ccambillau@gmail.com

### Specialty section:

This article was submitted to  
Protein Biochemistry for Basic and  
Applied Sciences,  
a section of the journal  
Frontiers in Molecular Biosciences

**Received:** 29 March 2022

**Accepted:** 21 April 2022

**Published:** 09 May 2022

### Citation:

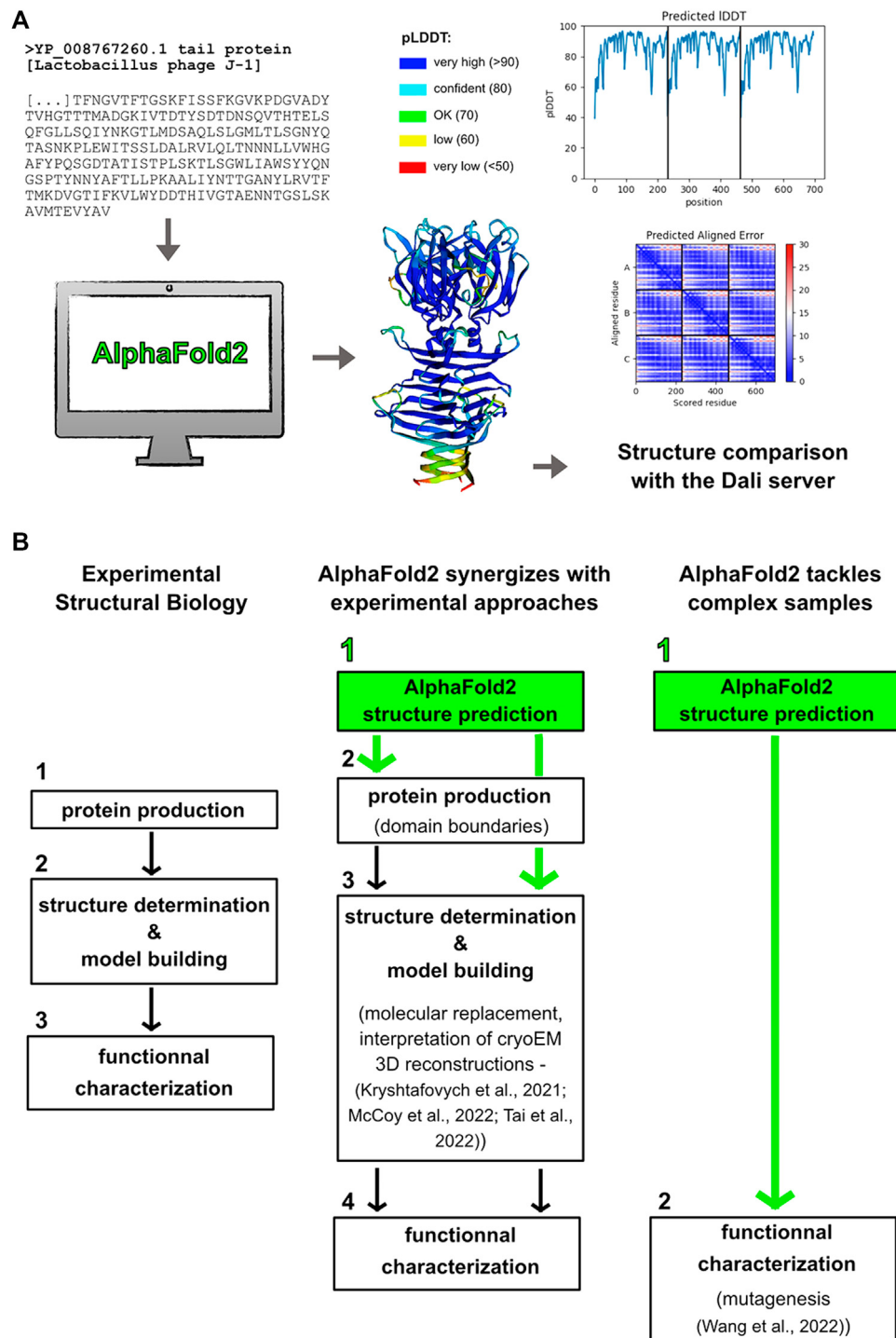
Goulet A and Cambillau C (2022)  
Present Impact of AlphaFold2  
Revolution on Structural Biology, and  
an Illustration With the Structure  
Prediction of the Bacteriophage J-1  
Host Adhesion Device.  
Front. Mol. Biosci. 9:907452.  
doi: 10.3389/fmolb.2022.907452

In 2021, the release of AlphaFold2 - the DeepMind's machine-learning protein structure prediction program - revolutionized structural biology. Results of the CASP14 contest were an immense surprise as AlphaFold2 successfully predicted 3D structures of nearly all submitted protein sequences. The AlphaFold2 craze has rapidly spread the life science community since structural biologists as well as untrained biologists have now the possibility to obtain high-confidence protein structures. This revolution is opening new avenues to address challenging biological questions. Moreover, AlphaFold2 is imposing itself as an essential step of any structural biology project, and requires us to revisit our structural biology workflows. On one hand, AlphaFold2 synergizes with experimental methods including X-ray crystallography and cryo-electron microscopy. On the other hand, it is, to date, the only method enabling structural analyses of large and flexible assemblies resistant to experimental approaches. We illustrate this valuable application of AlphaFold2 with the structure prediction of the whole host adhesion device from the *Lactobacillus casei* bacteriophage J-1. With the ongoing improvement of AlphaFold2 algorithms and notebooks, there is no doubt that AlphaFold2-driven biological stories will increasingly be reported, which questions the future directions of experimental structural biology.

**Keywords:** AlphaFold2, structural biology, bacteriophage, host adhesion device, bacteriophage-host interactions, *Lactobacillus casei* bacteriophage J-1

## INTRODUCTION

A major turning point in the field of structural biology was marked when AlphaFold2 - the DeepMind's machine learning structure prediction algorithm - was made publicly available in mid-2021 (Senior et al., 2020; Jumper et al., 2021). Stunning results produced by AlphaFold2 at the CASP14 (Critical Assessment of Techniques for Protein Structure Prediction) contest indicated that deep learning-based methods are now able to predict protein structures with an accuracy comparable, in most cases, to that of experimental structures (Pereira et al., 2021). This important achievement offers great perspectives to the life science community, including experimented structural biologists as well as untrained biologists, to address challenging



**FIGURE 1 |** Making the most out of AlphaFold2 structure predictions. **(A)** Predicted structures are obtained “simply” by providing their sequences to the program. The pLDDT and PAE scores have to be carefully considered for structure-function analyses. Structure comparison using the Dali server may return structural homologs (Holm, 2020). Here are shown the results obtained for the Tai C-terminal domain (trimeric assembly) from the *L. casei* phage J-1. **(B)** AlphaFold2 is becoming an essential part of any structural biology project. AlphaFold2 can synergize with experimental approaches in that predicted structures can be used to determine domain boundaries for recombinant protein production, to solve crystal structures by molecular replacement, and to interpret cryoEM 3D reconstructions. Moreover, AlphaFold2 can be the only way to get structural models of challenging samples, which can then be used as reliable templates for functional characterization.

biological questions (Goulet and Cambillau, 2021; Tunyasuvunakool et al., 2021). In particular, the AlphaFold Protein Structure Database (<https://www.alphafold.ebi.ac.uk>), developed by DeepMind and EMBL-EBI, offers open access to structure predictions for the human proteome, for proteins from key model organisms (e.g., *Mus musculus*, *Escherichia coli*, *Arabidopsis thaliana*), and for proteins from organisms related to global health (*Mycobacterium tuberculosis*, *Staphylococcus aureus*, *Trypanosoma cruzi*) (Tunyasuvunakool et al., 2021; Varadi et al., 2022). The number of structures in the AlphaFold database is expected to increase at a dizzying rate, with more than 100 million structures that should be available this year to include most representative sequences from the UniRef90 data set (Varadi et al., 2022).

However, AlphaFold2 is not substituting structural biologists. It is instead becoming a key step of any structural biology project, providing structures that must be properly analyzed. The reliability of predicted protein structures is such that they can be used as starting models for molecular replacement in X-ray crystallography, for the interpretation of cryo-electron microscopy (cryoEM) 3D reconstructions, and for the rational design of mutations and functional assays (Kryshtafovych et al., 2021; Slavin et al., 2021; Alerasool et al., 2022; Ivanov et al., 2022; McCoy et al., 2022; Paul et al., 2022). Moreover, AlphaFold2 can go a step further and predict the structure of large and flexible machineries, thereby circumventing intrinsic limitations of experimental methods (Goulet and Cambillau, 2021).

Here, we present our perspective on the present impact of AlphaFold2 on structural biology. In particular, we illustrate its power in predicting high-confidence structures of samples that cannot be studied by classical experimental approaches, using the *Lactobacillus casei* bacteriophage (phage) J-1 as model (Dieterle et al., 2017). The predicted structure of the whole J-1 host adhesion device provides important information on molecular mechanisms of phage-host interactions, and attests to the present progress of phage structural biology.

## MAKING THE MOST OUT OF ALPHAFOLD2 PREDICTIONS

AlphaFold2 generates 3D structure predictions for any given sequence (provided the length does not exceed 1,400 amino acids in freely accessible notebooks), which opens new research perspectives for many biologists. However, a solid background in structural biology and protein biochemistry is a critical prerequisite to properly interpret AlphaFold2 predictions. Basically, AlphaFold2 produces a 3D structure, using an amino acid sequence as input, by taking advantage of evolutionary information inferred from a multiple sequence alignment of homologs. This information provides inter-residue correlations that are translated into contact maps (Jumper et al., 2021). Confidence scores of predicted structures are given by pLDDT (predicted Local Distance Difference Test) per amino acid (Figure 1A). pLDDT values greater than 90% indicate an accuracy in the position of amino acid side chain comparable to that given by experimental crystal structures, while pLDDT

values lower than 50% indicate random position. These values are very informative and have to be carefully considered in evaluating the results. The pLDDT values, and hence the prediction reliability, highly depends on the number of aligned sequences. Usually, a minimum of 50–100 aligned sequences is enough to lead to a reasonable prediction. Therefore, drops in pLDDT values in a predicted protein structure correspond to drops in the number of aligned sequences. Moreover, low pLDDT values are often associated with loops and linkers as their sequences are less conserved in length and content as compared to those of the structural core of proteins. Hence, low pLDDT values usually maps well with flexible regions. The low confidence of these regions indicate that the relative orientation of the well-defined adjacent domains has to be interpreted with caution. The Predicted Aligned Error (PAE) is another metric to assess the confidence in the relative position and orientation of different domains (Varadi et al., 2022) (Figure 1A). Also, large ribbon-like structures with low pLDDT values predict intrinsically disordered regions (IDRs) and should not be interpreted as structures (Ruff and Pappu, 2021; Tunyasuvunakool et al., 2021).

## ALPHAFOLD2 IS CHANGING THE WAY WE DO STRUCTURAL BIOLOGY

The breakthrough of AlphaFold2, and certainly that of other publicly and freely available deep-learning based prediction methods like RoseTTAFold of the Rosetta suite (Baek et al., 2021), requires us to revisit our way of doing structural biology. This does not mean that experimental approaches are obsolete. Instead, AlphaFold2 expands the structural biology toolbox offering faster and more efficient ways to unravel protein structure-function relationships.

### Alphafold2 Structure Prediction: The Necessary First Step

AlphaFold2 structure prediction should become the first step of any structural biology project (Figure 1B). First of all, this prediction will provide crucial information on domain boundaries to use for successful recombinant protein production, which is the pre-requisite to perform any structural and functional analyses. In the case of a multi-domain protein, drops in the curve plotting pLDDT values over the sequence will indicate sequence stretches to be avoided. Moreover, predicted structures that return structural homologs in the PDB can unveil important functional information, which often cannot be retrieved from protein sequences. This can save months, or even years, of work to determine the first, experimental structure on which all the research project will be based. For instance, catalytic residues of an enzyme can be proposed from an AlphaFold2 structure obtained in few hours (Wang et al., 2022), which then enables to focus experimental efforts on structure determination of the enzyme bound to its substrate for instance, and on its functional characterization (Figure 1B). On the other hand, this prediction can reveal features explaining an unexpected biochemical and functional behavior. For instance, the presence of a loop

blockings access to an enzyme active site may be the cause of a loss of activity. In such cases, having this kind of information early in the project will help to save time and efforts that would otherwise be spent in troubleshooting protocol conditions.

## The Interplay Between AlphaFold2 and Experimental Methods

AlphaFold2 provides accurate 3D structure predictions at the atomic level. However, it cannot predict, at the moment at least, ligand binding, folding of IDRs upon partner binding, and multi-partner complex formation. Therefore, structural biology projects can benefit from the synergy between complementary machine-learning predictions and experimental approaches. In particular, AlphaFold2 predictions can be used as starting models for crystallographic phasing by molecular replacement and for model tracing and sequence positioning in cryoEM 3D reconstructions (**Figure 1B**) (Kryshtafovych et al., 2021; McCoy et al., 2022; Tai et al., 2022). An exciting perspective associated with AlphaFold2 release was the possibility to unlock “dormant” projects for which the phase problem could not be solved. It will be interesting to assess the impact of AlphaFold2 on the number of PDB entries over the next few years.

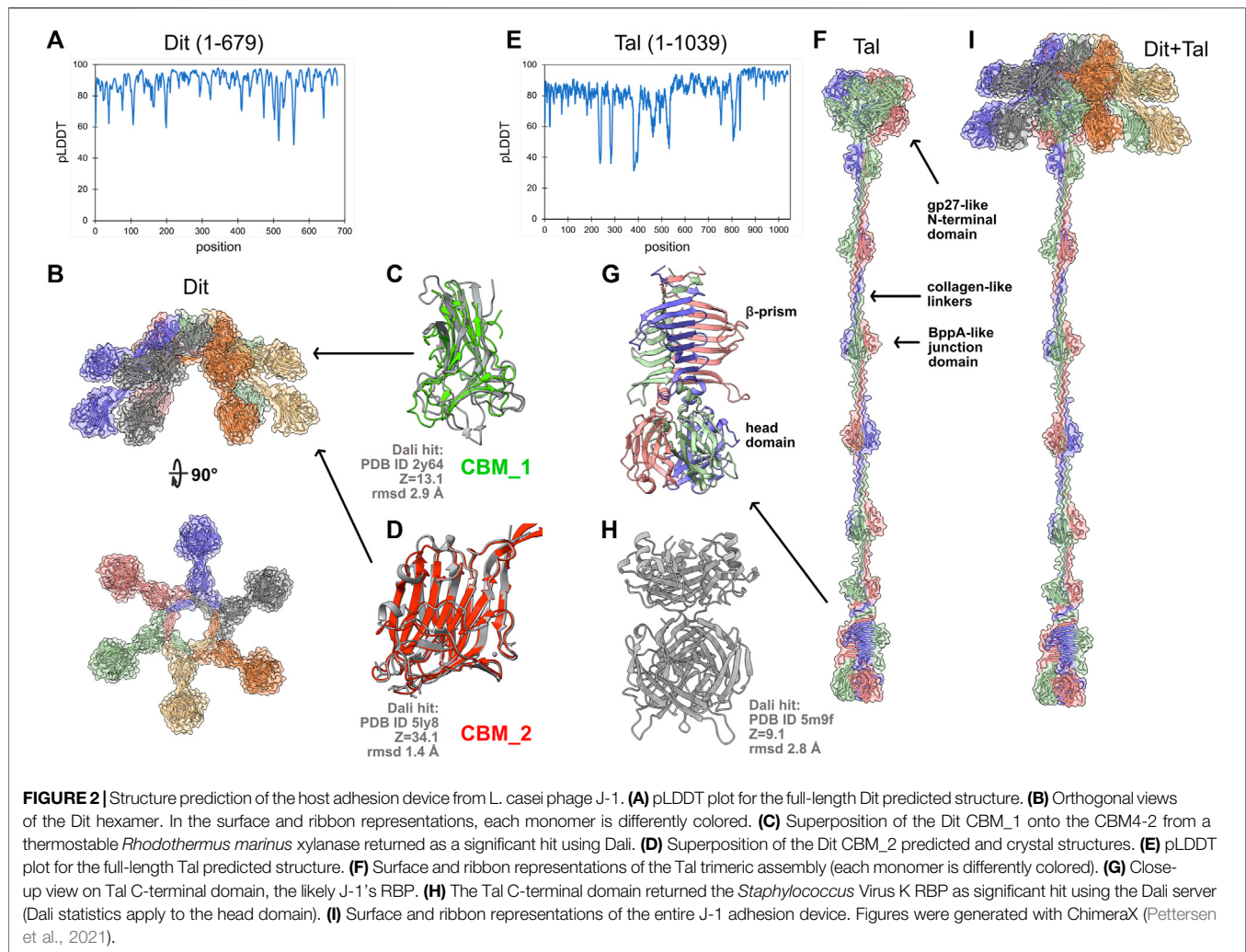
## ALPHAFOLD2 OPENS UP GREAT PERSPECTIVES TO TACKLE COMPLEX SAMPLES SUCH AS PHAGES’ HOST ADHESION DEVICES

To date, AlphaFold2 is the only way to get structural and functional insights into complex biological samples. Indeed, some macromolecular assemblies are reluctant to experimental approaches because they cannot produce X-ray-diffracting crystals or particles amenable to 3D reconstructions by cryoEM and single particle analysis (**Figure 1B**). This is the case of often flexible phages’ host adhesion devices, which are multi-protein assemblies involved in host binding at the onset of viral infection (Veesler and Cambillau, 2011; Goulet et al., 2020). We have chosen to illustrate such AlphaFold2 powerful capabilities with structure prediction of the *L. casei* phage J-1 host-adhesion device.

Tailed phages, and in particular Siphoviridae like J-1, possess the host adhesion device, which is in large part built from conserved components, at the extremity of their flexible tail (Veesler and Cambillau, 2011; Goulet et al., 2020). The core of this machinery always consists of an hexameric Distal Tail protein (Dit), bound on one side to the last hexameric Major Tail Protein (MTP), and on the other side to a trimeric Tail associated lysin (Tal). Dit and Tal may incorporate a variable number of carbohydrate-binding domains (CBM) that trigger phage-host adhesion. On top of these, other modules are often found, such as receptor-binding proteins (RBPs), various fibers, or enzymatic modules. These machineries are often extended and highly flexible in order to explore their surrounding in search of a bacterial host (Linares et al., 2020). From a structural biology view point, due to their flexibility, these

machineries can rarely be studied as a whole (for exceptions see (Sciara et al., 2010; Veesler et al., 2012)) but mostly as isolated modules. However, we recently showed that the structure of a long and flexible Tal from the *Oenococcus oeni* phage Vinitor 162 could be predicted with AlphaFold2 (Goulet and Cambillau, 2021).

The host adhesion device of J-1 comprises an hexameric ring of Dit (679 amino acids) attached to a trimer of Tal (1,039 amino acids). In 2017, we determined the crystal structure of one of the two CBMs inserted in the Dit, and showed that this domain recognizes and binds to the host cell surface (Dieterle et al., 2017). Also, we fitted this structure, together with the structures of Dit and Tal from the *Lactococcus lactis* phage p2 (Sciara et al., 2010), into a negative staining electron microscopy reconstruction of J-1 host adhesion device to produce a topological model of this multi-component assembly (Dieterle et al., 2017). It is important to note that the long and flexible extension of the Tal could not be resolved in the EM 3D reconstruction. This study is an example of the pre-AlphaFold2 way of doing structural biology to decipher the molecular mechanism of phages’ host-adhesion devices. Interestingly, the AlphaFold2-based approach, which ‘simply’ consists of submitting sequences to appropriate notebooks and comparing predictions to known structures using the Dali server (**Figure 1A**), provides, more rapidly, an insightful structure of the whole Dit-Tal assembly. We used a Github notebook (<https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb#scrollTo=XUo6foMQxwS2>) to perform the predictions, as it provides a simple and efficient service. To note, this notebook does not use PDB templates (as do ‘true’ AlphaFold2 servers), thereby providing a totally naive structure prediction. The Dit structure was predicted as a full-length monomer and assembled on the *L. lactis* phage p2 Dit hexameric ring with Coot (Emsley et al., 2010) (**Figures 2A,B; Supplementary Material**). Due to memory limitations, a monomer of Tal was first calculated (**Figure 2E**). Then, Tal segments shorter than 450 amino acids were selected for trimeric structure predictions, and assembled together with Coot to reconstitute a complete Tal trimer (**Figure 2F**). In the Dit predicted structure, two CBMs point out of the central ring by 70 Å (CBM\_1) and 40 Å (CBM\_2) (**Figures 2B–D**). Without much surprise, the Dali server reported an excellent match between the CBM\_2 predicted and crystal structures (Dieterle et al., 2017) (Z-score of 34.1, root mean square deviation (rmsd) of 1.4 Å; **Figure 2D**). CBM\_1 was found to be similar to the CBM4-2 from a thermostable *Rhodothermus marinus* xylanase (von Schantz et al., 2012) (**Figure 2C**). The Tal structure, as a trimer, has an overall length of ~650 Å (**Figures 2F,I; Supplementary Material**). The N-terminal domain (1-391) pertains to the classical T4 phage gp27 module, found in Myo- and Siphoviridae as well as in Type 6 secretion systems (Veesler and Cambillau, 2011). It is followed by six structural domains each mainly composed of three β-strands, and identified by Dali as part of *L. lactis* phage Tuc2009 BppA junction domain (Legrand et al., 2016) (**Figure 2F**). These domains are linked together by collagen-like extended structures. Following the sixth structural domain, the C-terminus is formed by a short 2-turns β-helix, followed by a 8-stranded β-prism (**Figure 2G**). Then, after



a short helical linker, a compact  $\beta$ -stranded trimeric domain forms the Tal distal end. This last domain resembles the host-binding domain of several phage RBPs (called “heads”), including that of *Staphylococcus* Virus K (**Figure 2H**), but also those of *L. lactis* phages p2 (Sciara et al., 2010) and TP901-1 (Veesler et al., 2012) or listerial phage PSA (Dunne et al., 2019). These observations make this domain a candidate as the J-1 RBP. Finally, DIT and Tal were assembled with *Coot*, using the phage p2 DIT-Tal assembly as template (**Figure 2I**), providing an overall view of J-1 host adhesion device. Based on this topology, we propose that J-1 binds to its host using, first, the unique Tal C-terminal RBP, and then, the six functional DIT's CBM\_2 (Dieterle et al., 2017), to secure binding and properly orientate the phage at the host's surface.

## DISCUSSION

Almost a decade ago, the “resolution revolution” in cryoEM, driven by major technical advances (Kühlbrandt, 2014),

dramatically improved the capabilities of structural biology. This field is experiencing another revolution, linked to the impressive improvement of machine-learning structure prediction algorithms like AlphaFold2. Accurate structures can now be predicted for virtually any protein, just from their sequence. AlphaFold2 notebooks are quite convenient and easy to use, thereby allowing structural biologists as well as non-specialists to obtain structures of their favorite samples in few minutes or few hours.

AlphaFold2 is likely to become an essential step of structure determination workflows. Indeed, it synergizes with experimental methods in assisting in the production of suitable samples and the determination/interpretation of experimental structures. Moreover, it represents, to date, the only way to characterize complex machineries at the structural level. In particular, our structure prediction of the host adhesion device from the *L. casei* phage J-1 illustrates the giant step forward that has been made in addressing technically challenging projects. Structural and functional insights can now be obtained on entire, large and flexible machineries.

This opens up great perspectives for phage structural biology to better understand molecular mechanisms involved in phage-host interactions. Moreover, AlphaFold2 structure predictions can be used to revisit phage genome annotations and to efficiently characterize the overwhelming number of newly discovered phages, as exemplified by the human gut phageome (Shkoporov et al., 2019).

For the time being, experimental techniques like X-ray crystallography and cryoEM are the only one capable of dealing with conformational dynamics, ligand binding, and complex formation. However, future improvement of machine-learning programs, which might be combined to molecular dynamics calculations, could seriously reduce the prevalence of experimental approaches to decipher protein structure-function relationships.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

## REFERENCES

- Alerasool, N., Leng, H., Lin, Z.-Y., Gingras, A.-C., and Taipale, M. (2022). Identification and Functional Characterization of Transcriptional Activators in Human Cells. *Mol. Cell* 82, 677–695. doi:10.1016/j.molcel.2021.12.008
- Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., et al. (2021). Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* 373, 871–876. doi:10.1126/science.abj8754
- Dieterle, M.-E., Spinelli, S., Sadovskaya, I., Piuri, M., and Cambillau, C. (2017). Evolved Distal Tail Carbohydrate Binding Modules of L. Actobacillus Phage J-1: a Novel Type of Anti-receptor Widespread Among Lactic Acid Bacteria Phages. *Mol. Microbiol.* 104, 608–620. doi:10.1111/mmi.13649
- Dunne, M., Rupf, B., Tala, M., Qabrati, X., Ernst, P., Shen, Y., et al. (2019). Reprogramming Bacteriophage Host Range through Structure-Guided Design of Chimeric Receptor Binding Proteins. *Cell Rep.* 29, 1336–1350. doi:10.1016/j.celrep.2019.09.062
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and Development of Coot. *Acta Crystallogr. D. Biol. Cryst.* 66, 486–501. doi:10.1107/S0907444910007493
- Goulet, A., and Cambillau, C. (2021). Structure and Topology Prediction of Phage Adhesion Devices Using AlphaFold2: The Case of Two Oenococcus Oeni Phages. *Microorganisms* 9, 2151. doi:10.3390/microorganisms9102151
- Goulet, A., Spinelli, S., Mahony, J., and Cambillau, C. (2020). Conserved and Diverse Traits of Adhesion Devices from Siphoviridae Recognizing Proteinaceous or Saccharidic Receptors. *Viruses* 12, 512. doi:10.3390/v12050512
- Holm, L. (2020). Using Dali for Protein Structure Comparison. *Methods Mol. Biol.* 2112, 29–42. doi:10.1007/978-1-0716-0270-6\_3
- Ivanov, Y. D., Taldaev, A., Lisitsa, A. V., Ponomarenko, E. A., and Archakov, A. I. (2022). Prediction of Monomeric and Dimeric Structures of CYP102A1 Using AlphaFold2 and AlphaFold Multimer and Assessment of Point Mutation Effect on the Efficiency of Intra- and Interprotein Electron Transfer. *Molecules* 27, 1386. doi:10.3390/molecules27041386
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2
- Kryshtafovych, A., Moulton, J., Albrecht, R., Chang, G. A., Chao, K., Fraser, A., et al. (2021). Computational Models in the Service of X-ray and Cryo- Electron

## AUTHOR CONTRIBUTIONS

Conceptualisation, AG and CC; Methodology, AG and CC; Writing—original draft, AG and CC; Writing—review and editing, AG and CC; Visualization, AG and CC.

## ACKNOWLEDGMENTS

We acknowledge UCSF ChimeraX for molecular graphics that is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.907452/full#supplementary-material>

- Microscopy Structure Determination. *Proteins* 89, 1633–1646. doi:10.1002/prot.26223
- Kühlbrandt, W. (2014). Cryo-EM Enters a New Era. *eLife* 3, e03678. doi:10.7554/eLife.03678
- Legrand, P., Collins, B., Blangy, S., Murphy, J., Spinelli, S., Gutierrez, C., et al. (2016). The Atomic Structure of the Phage Tuc2009 Baseplate Tripod Suggests that Host Recognition Involves Two Different Carbohydrate Binding Modules. *mBio* 7, e01781. doi:10.1128/mBio.01781-15
- Linares, R., Arnaud, C.-A., Degroux, S., Schoehn, G., and Breyton, C. (2020). Structure, Function and Assembly of the Long, Flexible Tail of Siphophages. *Curr. Opin. Virology* 45, 34–42. doi:10.1016/j.coviro.2020.06.010
- McCoy, A. J., Sammito, M. D., and Read, R. J. (2022). Implications of AlphaFold2 for Crystallographic Phasing by Molecular Replacement. *Acta Cryst. Sect. D. Struct. Biol.* 78, 1–13. doi:10.1107/S2059798321012122
- Paul, B., Weeratunga, S., Tillu, V. A., Hariri, H., Henne, W. M., and Collins, B. M. (2022). Structural Predictions of the SNX-RGS Proteins Suggest They Belong to a New Class of Lipid Transfer Proteins. *Front. Cell Dev. Biol.* 10, 826688. doi:10.3389/fcell.2022.826688
- Pereira, J., Simpkin, A. J., Hartmann, M. D., Rigden, D. J., Keegan, R. M., and Lupas, A. N. (2021). High-accuracy Protein Structure Prediction in CASP14. *Proteins* 89, 1687–1699. doi:10.1002/prot.26171
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., et al. (2021). UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* 30, 70–82. doi:10.1002/pro.3943
- Ruff, K. M., and Pappu, R. V. (2021). AlphaFold and Implications for Intrinsically Disordered Proteins. *J. Mol. Biol.* 433, 167208. doi:10.1016/j.jmb.2021.167208
- Sciara, G., Bebeacua, C., Bron, P., Tremblay, D., Ortiz-Lombardia, M., Lichère, J., et al. (2010). Structure of Lactococcal Phage P2 Baseplate and its Mechanism of Activation. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6852–6857. doi:10.1073/pnas.1000232107
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., et al. (2020). Improved Protein Structure Prediction Using Potentials from Deep Learning. *Nature* 577, 706–710. doi:10.1038/s41586-019-1923-7
- Shkoporov, A. N., Clooney, A. G., Sutton, T. D. S., Ryan, F. J., Daly, K. M., Nolan, J. A., et al. (2019). The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. *Cell Host Microbe* 26, 527–541. doi:10.1016/j.chom.2019.09.009
- Slavin, M., Zamel, J., Zohar, K., Eliyahu, T., Braitbard, M., Brielle, E., et al. (2021). Targeted *In Situ* Cross-Linking Mass Spectrometry and Integrative Modeling

- Reveal the Architectures of Three Proteins from SARS-CoV-2. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2103554118. doi:10.1073/pnas.2103554118
- Tai, L., Zhu, Y., Ren, H., Huang, X., Zhang, C., and Sun, F. (2022). 8 Å Structure of the Outer Rings of the *Xenopus laevis* Nuclear Pore Complex Obtained by Cryo-EM and AI. *Protein Cell*. doi:10.1007/s13238-021-00895-y
- Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Židek, A., et al. (2021). Highly Accurate Protein Structure Prediction for the Human Proteome. *Nature* 596, 590–596. doi:10.1038/s41586-021-03828-1
- Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al. (2022). AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* 50, D439–D444. doi:10.1093/nar/gkab1061
- Veesler, D., and Cambillau, C. (2011). A Common Evolutionary Origin for Tailed-Bacteriophage Functional Modules and Bacterial Machineries. *Microbiol. Mol. Biol. Rev.* 75, 423–433. first page of table of contents. doi:10.1128/MMBR.00014-11
- Veesler, D., Spinelli, S., Mahony, J., Lichière, J., Blangy, S., Bricogne, G., et al. (2012). Structure of the Phage TP901-1 1.8 MDa Baseplate Suggests an Alternative Host Adhesion Mechanism. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8954–8958. doi:10.1073/pnas.1200966109
- von Schantz, L., Håkansson, M., Logan, D. T., Walse, B., Österlin, J., Nordberg-Karlsson, E., et al. (2012). Structural Basis for Carbohydrate-Binding Specificity-A Comparative Assessment of Two Engineered Carbohydrate-Binding Modules. *Glycobiology* 22, 948–961. doi:10.1093/glycob/cws063
- Wang, X., Li, X., Ma, Y., He, J., Liu, X., Yu, G., et al. (2022). Inhibition Mechanisms of CRISPR-Cas9 by AcrIIA17 and AcrIIA18. *Nucleic Acids Res.* 50, 512–521. doi:10.1093/nar/gkab1197

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Goulet and Cambillau. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.