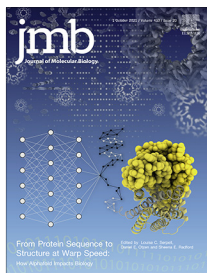




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Can We AlphaFold Our Way Out of the Next Pandemic?

Matthew K. Higgins

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

Correspondence to : matthew.higgins@bioch.ox.ac.uk

<https://doi.org/10.1016/j.jmb.2021.167093>

Edited by Louise C. Serpell

Abstract

The announcement of the outstanding performance of AlphaFold 2 in the CASP 14 protein structure prediction competition came at the end of a long year defined by the COVID-19 pandemic. With an infectious organism dominating the world stage, the developers of AlphaFold 2 were keen to play their part, accurately predicting novel structures of two proteins from SARS-CoV-2. In their blog post of December 2020, they highlighted this contribution, writing “we’ve also seen signs that protein structure prediction could be useful in future pandemic response efforts”. So, what role does structural biology play in guiding vaccine immunogen design and what might be the contribution of AlphaFold 2?

© 2021 Elsevier Ltd. All rights reserved.

Introduction

The value of structural biology in the fight against a pandemic has been brought into stark focus this past year by the way in which structural insights have guided development of vaccines to tackle COVID-19. Knowledge of the structure of the SARS-CoV coronavirus spike protein allowed design of mutants which stabilise the spike in the pre-fusion conformation.¹ As this is the form of the spike found on virus particles, vaccine immunogens which elicit antibodies that target this conformation are likely to be most effective.¹ As the SARS-CoV-2 virus, which causes COVID-19, is closely related to SARS-CoV^{2–4} this insight was transferable to the new pandemic strain.² As a result, the Pfizer/BioNTech and Moderna vaccines, amongst others, include spike-stabilising mutations in their effective designs.⁵ This is just one example of a broader field of ‘reverse vaccinology’, in which rational insight into the structures of pathogen surface proteins, and their complexes with neutralising monoclonal antibodies, guide design of improved vaccine immunogens.^{6–8} Structural biology clearly has a

part to play in tackling the pandemics of the present and of the future. But, what level of structural insight is required and how does this match the strengths of AlphaFold 2?

At the time of writing, our knowledge of AlphaFold 2 and of its performance in the Critical Assessment of Structure Prediction (CASP) 14 competition comes from press releases, blog posts and from the data released as part of the CASP competition,⁹ with peer reviewed publications, which reveal how AlphaFold 2 works, on the way. Nevertheless, the CASP competitions are stringent, rigorous and robust peer review processes in their own right¹⁰ and AlphaFold 2 was by some distance the best prediction method yet to test its arm in the competition.⁹ In particular, its performance was strongest in the challenging ‘free model’ category in CASP. Here predictions are made for proteins for which there is no related molecule of known structure available as a template for modelling.¹⁰ The CASP 14 press release states that AlphaFold 2 produced models with global distance test scores above 90 for about two thirds of the target proteins. These scores indicate a similar level of accuracy to

high-quality experimentally determined structures.¹¹

While AlphaFold 2 provides a substantial improvement in the prediction of structures of individual proteins, there are many important aspects of structural biology which it has not yet mastered. First, many proteins must undergo conformational changes, either upon binding to a partner or in response to a change in environment. Second, proteins are frequently modified through post-translational additions, such as glycosylation. Third, the majority of proteins perform their roles as part of complexes. This can be through the formation of multimeric complexes, through interaction with other proteins, such as receptors or antibodies, through interactions with other macromolecules such as DNA or through interactions with small molecules or ions. Multimerisation, conformational flexibility, glycosylation and interactions with human receptors or antibodies all play important roles in the functions of pathogen surface proteins. AlphaFold 2 has yet to master prediction of structures of protein complexes, or proteins with multiple conformational states, so how might it contribute to the design of vaccine immunogens and where do its current limitations lie?

Using Structure-Guided Approaches to Tackle the Challenging Vaccine Candidates Found on Pathogen Surfaces

Structure-guided vaccine design encompasses a series of approaches in which insights into the structures of pathogen surface proteins are used to guide the design of improved vaccine immunogens. These approaches are required particularly for the challenging cases, including influenza virus, human immunodeficiency virus (HIV) and respiratory syncytial virus (RSV). These pathogens have proved refractory to classical vaccine development as immunisation with unmodified viral surface proteins produce only weakly neutralising antibody responses. In such cases, the goal is to generate more focused immunogens, designed through the application of structural insight, which specifically elicit only the most neutralising antibodies.

One of the most advanced examples of the use of structural approaches to design immunogens has been seen in the development of vaccines against RSV.^{12–16} This virus causes a common respiratory illness which is frequently mild, but can be deadly for infants and younger adults. The major target for neutralising antibodies is the fusion (F) protein found on the surface of virus particles. This mediates interactions with human cells and drives membrane fusion and virus internalisation. While some antibodies against the F protein are protective,

many are not, and immunisation of human volunteers with F protein alone generates a weakly protective immune response, or can even lead to serious adverse effects.¹⁷

The challenge posed by the F protein comes from its conformational flexibility. F protein exists on virus particles in a metastable pre-fusion conformation. This readily converts into the post-fusion conformation through a dramatic conformational change which drives virus internalisation.¹⁶ While antibodies which target the post-fusion conformation can neutralise the virus and prevent invasion, the most effective antibodies target sites which are present at the tip of the pre-fusion conformation and are lost as a result of this conformational rearrangement.¹⁸ The most effective vaccine will contain immunogens which drive the generation of these pre-fusion targeting antibodies.¹⁴

The conformational dynamism of the RSV F protein is by no means unique, with the surface proteins of many viruses undergoing similarly dramatic conformational transitions.⁸ The importance of generating antibodies which target the pre-fusion state is therefore relevant to many viral diseases. Indeed, antibodies which target the post-fusion conformation can even, in some situations, be dangerous, aiding invasion of viruses into host cells and increasing the capacity of the virus to cause disease through a process known as antibody-dependent enhancement.¹⁹ Inducing the right antibodies can improve protection while inducing the wrong antibodies can be dangerous.

Using Structural Information to Design Better Vaccine Immunogens

So, how has structure-guided vaccinology helped in the design of immunogens to target RSV? First, the structure of the trimeric F protein was determined in both its pre- and post-fusion conformations, revealing the nature of its dramatic conformational change (Figure 1).^{16,20,21} Second, individual monoclonal antibodies were isolated from human volunteers, and were screened for the ability to neutralise the virus.¹⁶ Third, structures were determined of the most neutralising monoclonal antibodies bound to the F protein.²¹ This insight combined to reveal the chemical nature of the epitopes of the most desirable antibodies, guiding two different approaches to generate improved vaccine immunogens.^{12,13}

In the first approach, the goal was to produce an intact F protein trimer, locked into the pre-fusion conformation, using insight gained from comparison of the pre- and post-fusion structures.¹² The design process included introduction of cysteine residues positioned to form disulphide bonds, as well as integration of hydrophobic mutations to fill cavities. Finally, a trimeric 'foldon' tag was added to stabilise F protein monomers in a trimeric arrangement. This combination of

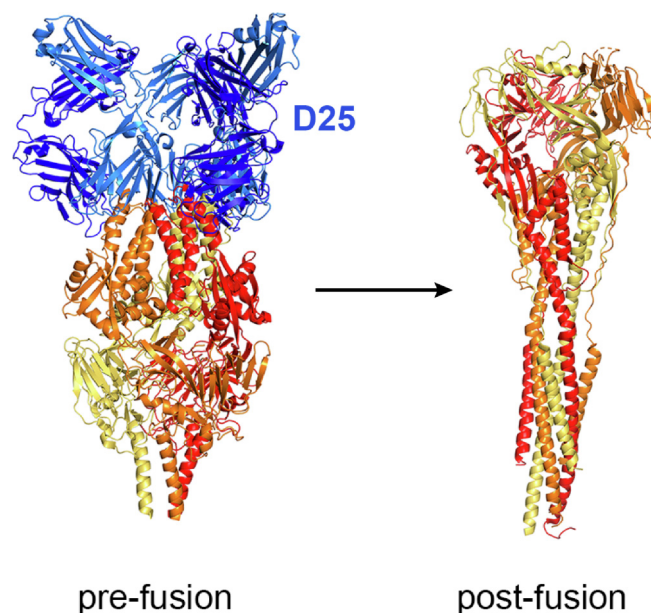


Figure 1. Respiratory syncytial virus as a paradigm of structure-guided design. The F protein of the respiratory syncytial virus undergoes a dramatic conformational change from its metastable pre-fusion conformation (PDB: 4JHW) to its post-fusion conformation (PDB: 3RRR). The F protein is a trimer with the three subunits shown in red, orange and yellow. The D25 monoclonal antibody is shown in blue. The most effective neutralising monoclonal antibodies target the pre-fusion conformation of the F protein and structure-guided vaccine design processes involve either modifications to stabilise the pre-fusion conformation or to design small *de novo* proteins which specifically present the epitopes for neutralising antibodies.

changes resulted in an F protein stabilised in the pre-fusion conformation which, when used to immunise human volunteers, generated 10-fold higher levels of neutralising antibody, in comparison to use of unmodified F protein.¹⁴ A second round of improvement involved testing over a hundred different designs and led to a further 4-fold gain in neutralising activity in an animal model.²² Structure-guided stabilisation of the pre-fusion F protein therefore provided the route to generate an immunogen which induced increased neutralising activity.

A second approach takes rational design a step further by creating small synthetic proteins which present specific antibody epitopes from the pre-fusion F protein. This was first achieved by grafting epitope surfaces onto an existing scaffold protein, followed by Rosetta-based protein design. This generates proteins which fold correctly to recapitulate these epitopes.¹³ More recently, entirely novel proteins have been designed *de novo* which present neutralising antibody epitopes.¹⁵ These designed immunogens very specifically cause the immune system to produce the desired neutralising antibodies. By combining immunogens containing different specific epitopes, a precisely tailored antibody response can be elicited. These two approaches each take different routes to generate vaccine immunogens which generate neutralising antibodies more effectively than the original viral surface protein.

Applying Structure-Guided Vaccine Design to the Coronavirus Spike

The coronavirus spike protein presents similar challenges to the RSV F protein, forming a trimeric molecule with pre- and post-fusion conformations (Figure 2).²³ Here again, the pre-fusion conformation is metastable and a large conformational change occurs as it transitions to the post-fusion conformation, driving membrane fusion. Knowledge of the structure of the pre-fusion conformations of the SARS-CoV and MERS-CoV coronavirus spikes allowed the design of mutants in which two proline residues were inserted into hinges within the spike which are predicted to become helices in the post-fusion conformation, thereby destabilising this conformational transition.¹ This 2P variant expressed around 50-fold better than the original spike and generated more effective neutralising antibody responses when used as an immunogen. As the SARS-CoV-2 spike protein is very similar in structure to that of SARS-CoV,^{2,3} the equivalent 2P mutation could be transferred across and this stabilised SARS-CoV-2 spike has been included in the successful Pfizer/BioNTech, Moderna, Novavax and Johnson and Johnson vaccines, amongst others.⁵ More recently, a second round of structure-guided mutation has been used to promote further stabilisation. A hundred structure-guided mutations were designed, including those

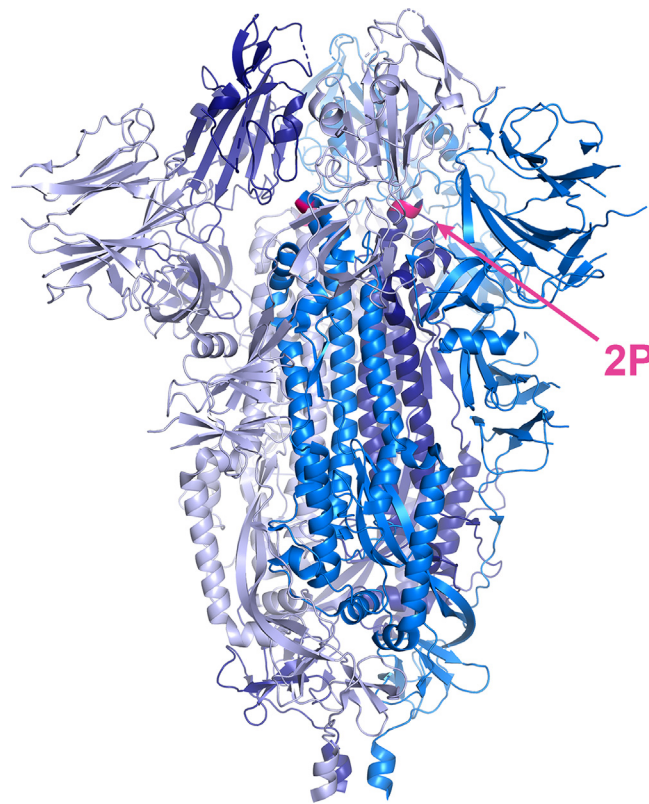


Figure 2. Structure-guided vaccinology of the coronavirus spike. The structure of the closed conformation of the SARS-CoV-2 spike (PDB: 6VSB) is shown with the three subunits of the trimer in different shades of blue. The 2P mutations (residues 986 and 987) are highlighted in pink, with an arrow pointing to one of these copies. These mutations stabilise the spike in the pre-fusion conformation and are included in a number of the effective COVID-19 vaccines.

which insert disulphides and salt bridges, or that introduce hydrophobic residues to fill internal cavities and prolines to cap helices and prevent loops from transitioning into helices.²⁴ The best of these was the HexaPro spike, which further increased stability from the 2P version and may be included in the coronavirus vaccines of the future. The use of structure-guided design appears firmly established for the generation of vaccines which more effectively target dynamic viral fusion proteins.

What Can Be Done With Single Antigens?

While tour-de-force immunogen design studies for vaccines which target RSV, SARS-CoV-2, HIV and influenza are guided by structures of dynamics proteins in multiple conformations, bound to different monoclonal antibodies, there are also cases in which the surface antigens of a pathogen are monomeric and rigid. Here, structures of individual antigens can be used to inform vaccine immunogen design, once again allowing design of immunogens which elicit only the most effective antibody responses.

An example is the PfRH5 protein from the malaria parasite, *Plasmodium falciparum* (Figure 3).^{25,26} PfRH5 is essential for this parasite to invade human erythrocytes, through its interaction with the erythrocyte receptor basigin.²⁶ As invasion is essential for parasite replication and transmission,²⁶ PfRH5 is a promising vaccine candidate.^{25,27–29} Unlike the viral fusion machinery, PfRH5 is a rigid protein which adopts the correct structure as a monomer,^{27,30} and so structural studies of monomer alone have proved instructive for immunogen design. Initial structures revealed which regions of PfRH5 form part of an ordered domain that is the target of the most neutralising antibodies, and also revealed the boundaries of a flexible loop and the flexible N-terminus, both of which can be removed to generate an immunogen which still contains neutralising epitopes.²⁷ As flexible loops often contain epitopes for non-neutralising antibodies, which may be immuno-dominant, they are often best trimmed away, and structural studies can guide this trimming process.

Structures can also be used to guide improvement of the physicochemical properties of a vaccine immunogen, making it more stable or more easily produced. While PfRH5 is an

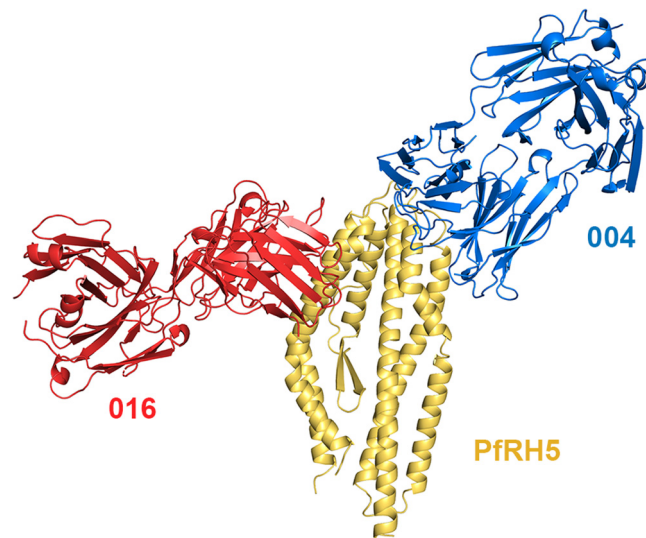


Figure 3. Structure-guided vaccinology for blood stage malaria. The PfRH5 protein is a promising antigen for a vaccine to prevent malaria caused by *Plasmodium falciparum*. Structures of PfRH5 (yellow) bound to human monoclonal antibodies such as 004 (red) and 016 (blue) (PDB: 6RCU) revealed flexible loops in PfRH5 which could be removed to ablate epitopes for non-neutralising antibodies. Structural insight also allowed the design of thermally stabilised versions of PfRH5, improving physicochemical properties.

excellent vaccine immunogen,²⁵ it is challenging to produce, with low yields from eukaryotic expression systems. It was therefore remodelled using a Rosetta-based design tool, known as the Protein Repair One Stop Shop (PROSS).^{31,32} This uses multiple sequence alignments to identify residues within a protein which are mutable. A Rosetta-based routine then predicts the effect of different residues in these locations on protein stability, before pooling predicted stabilising mutations to generate protein variants. By specifying residues in neutralising epitopes as invariant, versions of PfRH5 were designed which have unaltered immunogenic properties, but which express to higher levels and possess improved thermal stability.³¹

Can We AlphaFold Our Way to Improved Vaccines?

So, having taken a glimpse at some of the current approaches in structure-guided vaccine design, now we can ask what role might AlphaFold 2 have in guiding design of the vaccines of the future? The current strengths of AlphaFold 2 appear to be in the prediction of structures of single proteins, both in cases where there is a similar structure to act as a template, and where there is not. As seen in the case of PfRH5, structures of single antigens can be used to guide design of improved vaccine immunogens. This is best done with knowledge of the location and nature of epitopes for the most neutralising antibodies, allowing the designer to ensure that these epitopes are not changed by the design process. However, such insight could be

added to a structural model through non-structural methods, such as epitope mapping by hydrogen–deuterium exchange mass spectrometry. Rosetta-based redesign approaches have not yet been attempted using models derived from AlphaFold 2, however it appears as though many of these structures have the level of accuracy required for the successful use of these design approaches. Where such important antigens are of unknown structure, the current version of AlphaFold 2 may be able to make a valuable contribution.

For more complex antigens, as exemplified by the viral fusion proteins of RSV and COVID-19, there is still more to do before AlphaFold 2 will contribute substantially. First, the ability to predict structures of protein complexes is required. This both involves structure prediction for homo-oligomeric complexes, such as viral fusion protein trimers, and hetero-oligomeric complexes, such as those of a pathogen surface protein in complex with neutralising antibodies. It seems likely that the developers of AlphaFold will have prediction of the structures of complexes firmly in their sights, as the global analysis of sequence co-variance, which appears to lie at the heart of the AlphaFold algorithms³³ should also be highly applicable to protein complex structure prediction. Indeed, the battle between pathogen and host is one with covariance at its heart, with the pathogen evolving to avoid immune detection and the immune system adapting to keep up. AlphaFold may well develop in this direction, allowing structures of complexes to be accurately predicted.

However, the challenges do not end with prediction of structures of complexes, as many

pathogen surface molecules have evolved to be conformationally dynamic. The metastable pre-fusion conformations of the viral surface proteins, which are the targets of many of the most neutralising monoclonal antibodies, have evolved to be short-lived to minimise their exposure to the immune system and to allow shape changes required for the viral life cycle. These transitions also often occur upon changes in environmental conditions, such as pH, on internalisation of a virus into a cell, or on contact with the membrane environment. A major challenge for protein prediction methods will be to understand how these conformational changes occur in response to environmental cues.

Finally, non-protein molecules must often be added into the mix. This review has also not touched on the use of structural biology to guide small molecule therapeutic design. Here too, structural insights can be used as part of the response to a pandemic. But here, the structures must contain not just protein, but also molecules with very different chemistry, where protein covariance is no longer a factor.

So, will AlphaFold 2 help us out in future pandemics? These impressive advances provide a major leap forwards in the ability to determine structures of single proteins. However, to make major contributions to structure-guided vaccine immunogen design, more is needed, including prediction of structures of conformationally dynamic multimers and their complexes with antibodies. It would be unwise to bet against the minds at Deepmind to make these advances in the future. However, we should also keep the x-ray and the electron beams flowing. The transformations in structural biology in the past decade have also been remarkable, as has the rapidity of the response of the structural biology community to COVID-19. These tools will continue to play their role as, for many pathogens, structural vaccinologists will not be able to rely on structure prediction alone for a while yet.

Received 3 May 2021;

Accepted 1 June 2021;

Available online 8 June 2021

Keywords:

AlphaFold;
vaccine;

structure-guided immunogenicity design

References

- Pallesen, J., Wang, N.S., Corbett, K.S., Wrapp, D., Kirchdoerfer, R.N., Turner, H.L., et al., (2017). Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E7348–E7357.
- Wrapp, D., Wang, N.S., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., et al., (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, **367**, 1260.
- Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., Velesler, D., (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein (vol 180, 281.e1, 2020). *Cell*, **183**, 1735.
- Alam, N., Higgins, M.K., (2020). A spike with which to beat COVID-19? *Nature Rev. Microbiol.*, **18**, 414.
- Fauci, A.S., (2021). The story behind COVID-19 vaccines. *Science*, **372**, 109.
- Sesterhenn, F., Bonet, J., Correia, B.E., (2018). Structure-based immunogen design - leading the way to the new age of precision vaccines. *Curr. Opin. Struct. Biol.*, **51**, 163–169.
- Graham, B.S., Gilman, M.S.A., McLellan, J.S., (2019). Structure-based vaccine antigen design. *Annu. Rev. Med.*, **70**, 91–104.
- Kwong, P.D., DeKosky, B.J., Ulmer, J.B., (2020). Antibody-guided structure-based vaccines. *Semin. Immunol.*, **50**
- Callaway, E., (2020). 'It Will Change Everything': Deepmind's Ai makes gigantic leap in solving protein structures. *Nature*, **588**, 203–204.
- Kryshtafovych, A., Schwede, T., Topf, M., Fidelis, K., Moul, J., (2019). Critical assessment of methods of protein structure prediction (CASP)-Round XIII. *Proteins*, **87**, 1011–1020.
- Zemla, A., (2003). LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Res.*, **31**, 3370–3374.
- McLellan, J.S., (2013). Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus (vol 342, pg 592, 2013). *Science*, **342**, 931.
- Correia, B.E., Bates, J.T., Loomis, R.J., Baneyx, G., Carrico, C., Jardine, J.G., et al., (2014). Proof of principle for epitope-focused vaccine design. *Nature*, **507**, 201–206.
- Crank, M.C., Ruckwardt, T.J., Chen, M., Morabito, K.M., Phung, E., Costner, P.J., et al., (2019). A proof of concept for structure-based vaccine design targeting RSV in humans. *Science*, **365**, 505.
- Sesterhenn, F., Yang, C., Bonet, J., Cramer, J.T., Wen, X. L., Wang, Y.M., et al., (2020). De novo protein design enables the precise induction of RSV-neutralizing antibodies. *Science*, **368**, 730.
- Graham, B.S., (2017). Vaccine development for respiratory syncytial virus. *Curr. Opin. Virol.*, **23**, 107–112.
- Anderson, L.J., Dormitzer, P.R., Nokes, D.J., Rappuoli, R., Roca, A., Graham, B.S., (2013). Strategic priorities for respiratory syncytial virus (RSV) vaccine development. *Vaccine*, **31**, B209–B215.
- Phung, E., Chang, L.A., Morabito, K.M., Kanekiyo, M., Chen, M., Nair, D., et al., (2019). Epitope-specific serological assays for RSV: Conformation matters. *Vaccines-Basel.*, **7**
- Takada, A., Kawaoka, Y., (2003). Antibody-dependent enhancement of viral infection: molecular mechanisms and in vivo implications. *Rev. Med. Virol.*, **13**, 387–398.
- McLellan, J.S., Yang, Y.P., Graham, B.S., Kwong, P.D., (2011). Structure of respiratory syncytial virus fusion glycoprotein in the postfusion conformation reveals preservation of neutralizing epitopes. *J. Virol.*, **85**, 7788–7796.
- McLellan, J.S., Chen, M., Leung, S., Graepel, K.W., Du, X. L., Yang, Y.P., et al., (2013). Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. *Science*, **340**, 1113–1117.

22. Joyce, M.G., Zhang, B.S., Ou, L., Chen, M., Chuang, G.Y., Druz, A., et al., (2016). Iterative structure-based improvement of a fusion-glycoprotein vaccine against RSV. *Nature Struct. Mol. Biol.*, **23**, 811–820.
23. Walls, A.C., Tortorici, M.A., Snijder, J., Xiong, X.L., Bosch, B.J., Rey, F.A., et al., (2017). Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 11157–11162.
24. Hsieh, C.L., Goldsmith, J.A., Schaub, J.M., DiVenere, A. M., Kuo, H.C., Javanmardi, K., et al., (2020). Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science*, **369**, 1501.
25. Ragotte, R.J., Higgins, M.K., Draper, S.J., (2020). The RH5-CyRPA-Ripr complex as a malaria vaccine target. *Trends Parasitol.*, **36**, 545–559.
26. Crosnier, C., Bustamante, L.Y., Bartholdson, S.J., Bei, A. K., Theron, M., Uchikawa, M., et al., (2011). Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. *Nature*, **480**, 534–U158.
27. Alanine, D.G.W., Quinkert, D., Kumarasingha, R., Mehmood, S., Donnellan, F.R., Minkah, N.K., et al., (2019). Human antibodies that slow erythrocyte invasion potentiate malaria-neutralizing antibodies. *Cell*, **178**, 216.
28. Draper, S.J., Sack, B.K., King, C.R., Nielsen, C.M., Rayner, J.C., Higgins, M.K., et al., (2018). Malaria vaccines: Recent advances and new horizons. *Cell Host Microbe*, **24**, 43–56.
29. Payne, R.O., Silk, S.E., Elias, S.C., Miura, K., Diouf, A., Galaway, F., et al., (2017). Human vaccination against RH5 induces neutralizing antimalarial antibodies that inhibit RH5 invasion complex interactions. *JCI Insight*, **2**.
30. Wright, K.E., Hjerrild, K.A., Bartlett, J., Douglas, A.D., Jin, J., Brown, R.E., et al., (2014). Structure of malaria invasion protein RH5 with erythrocyte basigin and blocking antibodies. *Nature*, **515**, 427.
31. Campeotto, I., Goldenzweig, A., Davey, J., Barfod, L., Marshall, J.M., Silk, S.E., et al., (2017). One-step design of a stable variant of the malaria invasion protein RH5 for use as a vaccine immunogen. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 998–1002.
32. Goldenzweig, A., Goldsmith, M., Hill, S.E., Gertman, O., Laurino, P., Ashani, Y., et al., (2016). Automated structure- and sequence-based design of proteins for high bacterial expression and stability. *Mol. Cell*, **63**, 337–346.
33. 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.