THE EMBO JOURNAL

Molecular basis of cross-species ACE2 interactions with SARS-CoV-2-like viruses of pangolin origin

Sheng Niu, Jia Wang, Bin Bai, Lili Wu, Anqi Zheng, Qian Chen, Pei Du, Pengcheng Han, Yanfang Zhang, Yunfei Jia, Chengpeng Qiao, Jianxun Qi, Wen-xia Tian, Hong-Wei Wang, Qihui Wang, and George Gao
DOI: 10.15252/embj.2021107786

Corresponding authors: George Gao (gaof@im.ac.cn) , Hong-Wei Wang (hongweiwang@tsinghua.edu.cn), Qihui Wang (wangqihui@im.ac.cn), Wen-xia Tian (wenxiatian@126.com)

Review Timeline:

Submission Date:	20th Jan 21
Editorial Decision:	2nd Mar 21
Revision Received:	29th Mar 21
Editorial Decision:	5th May 21
Revision Received:	8th May 21
Accepted:	14th May 21

Editor: Ieva Gailite

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

1st Editorial Decision 2nd Mar 2021

Thank you for submitting your manuscript for consideration by The EMBO Journal. We have now received two referee reports on your manuscript, which are included below for your information.

As you will see from the comments, both reviewers appreciate the study, while also indicating a number of concerns that would have to be addressed and clarified before they can support publication of the manuscript. In particular, reviewer #1 finds that the functional role for pangolin CoV S proteins for the virus entry into ACE2-expressing human cells would have to be shown, and both reviewers indicate that insufficient information on protein purification and structural analysis has been provided in the current version. From my side, I find these points reasonable and would like to invite you to address the concerns raised by both reviewers in a revised manuscript.

I should add that it is The EMBO Journal policy to allow only a single major round of revision and that it is therefore important to resolve the main concerns at this stage. We are aware that many laboratories cannot function at full efficiency during the current COVID-19/SARS-CoV-2 pandemic, and I would be happy to discuss the revision in more detail via email or phone/videoconferencing.

We have extended our 'scooping protection policy' beyond the usual 3 month revision timeline to cover the period required for a full revision to address the essential experimental issues. This means that competing manuscripts published during revision period will not negatively impact on our assessment of the conceptual advance presented by your study. Please contact me if you see a paper with related content published elsewhere to discuss the appropriate course of action.

Please feel free to contact me if you have any further questions regarding the revision. Thank you for the opportunity to consider your work for publication. I look forward to receiving the revised manuscript.

Referee #1:

In this report, Niu et al. have compared binding of RBDs from SARS-CoV-2 and two pangolin CoVs to a panel of ACE2 from 27 different species. Binding was measured by two independent methods - flow cytometry and SPR. They found that the pangolin CoV RBDs appeared to have a broader host range than that of SARS-CoV-2. They then determined the structures of the two pangolin CoV RBDs in complex with human ACE2 at 3.4 Å resolution by cryo-EM. The pangolin CoV RBDs bind ACE2 in a similar mode as does the SARS-CoV-2 RBD. Additional mutational data showed that Q498H mutation led to high affinity binding of the SARS-CoV-2 RBD to mouse, rat and European hedgehog ACE2s. The authors suggest that the two pangolin CoVs may infect human and require further surveillance to help prevent future pandemics.

This is a comprehensive study concerning interactions between CoV RBDs and various ACEs from different animals. Binding data derived from two different approaches are qualitatively in agreement with each other, and further supported by the structural studies. It is also interesting that a single residue seems to be responsible for a potentially broader host range for the two pangolin CoVs.

Several concerns need to be addressed to improve the manuscript:

1. Conceptually, it is not known whether ACE2 binding alone would be sufficient for a coronavirus to infect (and replicate in) a new host. It would be really helpful to have some experimental evidence to support this hypothesis. Since the two pangolin CoV RBDs bind human ACE2 quite well, it should be straightforward to test whether the pangolin CoV S could indeed support entry into ACE2-expressing human cells using a pseudovirus assay.
2. Not data on protein production are shown. It would be useful to show the gel filtration traces of the purified protein reagents used in all the binding studies. In addition, ACE2 proteins and some RBDs were produced in insect cells and others made in mammalian cells. The two types of cells have different glycosylation, which has been reported to modulate the Spike-ACE2 interactions.
3. Details of cryo-EM are insufficient to judge whether the structural interpretations are solid or not. Only a summary table was shown. How good are the overall density and the density at the interfaces? What is the local resolution? Are all the discussions on the structural details justified at an overall resolution of 3.4Å? Is there any glycan involvement? A standard set of cryo-EM supplemental figures should be included.
4. Some arguments in Discussion do not sound convincing. For example, on p. 13, "However, no SARS-CoV-2 has yet been detected in pangolin. A recent report showed that pangolins could be infected by the pangolin CoV GD/1/2019 and showed clinical symptoms and histological changes (Xiao et al., 2020), further weakening the possibility of pangolin as intermediate hosts for SARS-CoV-2." Why would these data weaken the possibility of pangolin as an intermediate host? Adaptation of SARS-CoV-2 in humans may prevent it going back to its previous host and not sure what is the relevance here regarding pangolin CoV infecting pangolin. Another example, on the same page, "Furthermore, two pangolin CoV RBDs associate with a panel of ACE2 orthologs at a similar level to that of the SARS-CoV-2, but present a broader host range, with binding capability to ACE2 orthologs from three additional animals (mouse, rat and European hedgehog), indicating the two pangolin CoVs are likely to be at lower level in the evolutionary pathway." Why would the two pangolin CoVs be at a lower level in the evolutionary pathway? Simply because of a broader host range? If so, should bat CoVs have even a broader host range?

Referee #2:

This manuscript from Niu et al describes thorough binding and structural studies for two receptor-binding domains (RBDs) of spikes derived from pangolin-origin SARS-CoV-2-like viruses. Flow cytometry and surface plasmon resonance are used to determine binding affinities to ACE2 orthologs from various species, which revealed that the pangolin RBDs have a potentially broader host range. Cryo-EM structures of the RBDs in complex with hACE2 reveal molecular determinants of the binding interfaces, including the role of His498 to increase the affinity and binding breadth of the pangolin RBDs. Testing of a Q498H variant of the SARS-CoV-2 RBD demonstrates that it binds to hACE2 with higher affinity while also binding to ACE2s from mouse, rat, and hedgehog.

Overall, this is a thorough manuscript that is generally clear with most conclusions supported by the data. The flow cytometry and SPR binding experiments have been performed well, and the structural validation statistics are reasonable, although Table 1 is missing values for the % of poor rotamers. The major concern is the quality of the EM structures. There are no figures showing the 3D reconstruction and density at the interface, the FSC curves, and angular distribution plots. There are also no PDB validation reports provided. As such, it is very difficult to assess the quality of these structures and the conclusions drawn from them. This should be addressed in a revised manuscript.

Response to reviewers' comments

We appreciate the reviewers' supportive assessment of our work, and their comments on aspects that could be improved. Below we respond to each reviewer's points in detail, with notes as to where changes to the manuscript have been made.

Reviewer 1

Comments

1) Conceptually, it is not known whether ACE2 binding alone would be sufficient for a coronavirus to infect (and replicate in) a new host. It would be really helpful to have some experimental evidence to support this hypothesis. Since the two pangolin CoV RBDs bind human ACE2 quite well, it should be straightforward to test whether the pangolin CoV S could indeed support entry into ACE2-expressing human cells using a pseudovirus assay.

Response: Thanks for the constructive comment. We have performed the transduction assay using the three VSV-based pseudotyped viruses (SARS-CoV-2, GX/P2V/2017 and GD/1/2019). Similar amounts of the three pseudoviruses (as determined by quantitative real-time PCR) were used to transduce HeLa cells expressing hACE2. The results showed that all three pseudoviruses were unable to transduce HeLa cells, but readily transduced Hela cells expressing hACE2 (HeLa-hACE2). Moreover, The SARS-CoV-2 and GD/1/2019 pseudoviruses displayed similar transduction efficiency, and pseudotyped GX/P2V/2017 showed lower efficiency. Figure 3 has been prepared to present these new data.

2) …Not data on protein production are shown. It would be useful to show the gel filtration traces of the purified protein reagents used in all the binding studies...

Response: Thanks for the kind suggestion. We have included the results of gel filtration traces and SDS-PAGE of the purified protein in all experiments in the new Figures EV1 and EV2.

3) …In addition, ACE2 proteins and some RBDs were produced in insect cells and others made in mammalian cells. The two types of cells have different glycosylation, which has been reported to modulate the Spike-ACE2 interactions…

Response: In this study, SARS-CoV-2 RBD expressed in insect cells and HEK293F cells were both used in the binding assay. Our previous structure data suggest that glycosylation in the SARS-CoV-2 RBD does not involving in the binding to hACE2 (Wang et al., Cell, 2020). Similar observations were also reported by other groups (Lan et al., Nature, 2020; Shang et al., Nature, 2020). SARS-CoV-2 RBD expressed in 293F cells seems to have more glycosylation modifications than those in insect cells (Figure EV1). However, the binding affinities of the two types of SARS-CoV-2 RBDs to hACE2 are similar. Figure 2 showed the binding affinity between SARS-CoV-2 RBD expressed in insect cells to hACE2. The K_D is calculated to be 11.2 ± 0.5 nM. Figure 6 showed the binding affinity between SARS-CoV-2 RBD expressed in 293F cells to hACE2. The K_D is calculated to be 9.9 ± 2.5 nM. Thus, the different glycosylation pattens between insect cells and HEK293F cells seems to exert no effect on the interaction between SARS-CoV-2 RBD and hACE2.

4) Details of cryo-EM are insufficient to judge whether the structural interpretations are solid or not. Only a summary table was shown. How good are the overall density and the density at the interfaces? What is the local resolution? Are all the discussions on the structural details justified at an overall resolution of 3.4Å? Is there any glycan involvement? A standard set of cryo-EM supplemental figures should be included.

*Response: Tha*nk*s for the kind suggestion. For better presentation of our cryo-EM data,*

we have prepared a standard set of cryo-EM supplemental figures of the two complex structures (Figure EV3 for hACE2-GX/P2V/2017 RBD, and Figure EV4 for hACE2-GD/1/2019 RBD). In the Figures EV3 and EV4, to show the overall density and the density at the interfaces, we provided the final EM density coloured by local

resolution. In addition, there are some glycans in the two complex structures, and we showed the glycans in the new prepared Figure 4.

5) **Some arguments in Discussion do not sound convincing. For example, on p. 13, "However, no SARS-CoV-2 has yet been detected in pangolin. A recent report showed that pangolins could be infected by the pangolin CoV GD/1/2019 and showed clinical symptoms and histological changes (Xiao et al., 2020), further weakening the possibility of pangolin as intermediate hosts for SARS-CoV-2." Why would these data weaken the possibility of pangolin as an intermediate host? Adaptation of SARS-CoV-2 in humans may prevent it going back to its previous host and not sure what is the relevance here regarding pangolin CoV infecting pangolin. Another example, on the same page, "Furthermore, two pangolin CoV RBDs associate with a panel of ACE2 orthologs at a similar level to that of the SARS-CoV-2, but present a broader host range, with binding capability to ACE2 orthologs from three additional animals (mouse, rat and European hedgehog), indicating the two pangolin CoVs are likely to be at lower level in the evolutionary pathway." Why would the two pangolin CoVs be at a lower level in the evolutionary pathway? Simply because of a broader host range? If so, should bat CoVs have even a broader host range?**

Response: *We are sorry for the unsuitable arguments.*

Firstly, we rewrite the arguments about the reservoir.

Page 13, lines 15-22: Although not universally true, natural reservoir or the intermediate hosts tend to have coevolved with their viruses and usually do not display clinical symptom. For example, bats are natural reservoirs for a variety of emerging viruses yet rarely cause clinical disease in bats (Shi & Hu, 2008; Wang et al, 2006). Dromedary camels are thought to be the intermediate host for MERS-CoV, they could carry the virus without showing any severe disease (Peck et al, 2015). Thus, pangolins are more likely to be another victim of SARS-CoV-2-like viruses, rather than to be the natural reservoir or intermediate host.

Secondly, we removed the other unconvincing argument in Page 14 "indicating the two pangolin CoVs are likely to be at lower level in the evolutionary pathway".

Reviewer 2

Comments

1) …although Table EV1 is missing values for the % of poor rotamers…

Response: We are sorry for the mistake. We have updated the Table EV1 and added the values for the % of poor rotamers.

2) ...The major concern is the quality of the EM structures. There are no figures showing the 3D reconstruction and density at the interface, the FSC curves, and angular distribution plots. There are also no PDB validation reports provided. As such, it is very difficult to assess the quality of these structures and the conclusions drawn from them. This should be addressed in a revised manuscript…

Response: As suggested by the reviewer, Figure EV3 and Figure EV4 were added to present the EM data processing of two complex structures including 3D reconstruction, density at the interface, the FSC curves, and angular distribution plots. Furthermore, we also provided the PDB validation reports of the two complex structures in the revised version.

1st Revision - Editorial Decision 5th May 2021

Thank you for submitting a revised version of your manuscript. I apologise for the unusually protracted review process due to delayed submission of referee reports. Your study has now been seen by one of the original reviewers, who finds that their main concerns have been addressed and now recommends publication of the manuscript. Therefore, I would like to invite you to address the remaining editorial issues before I can extend the official acceptance of the manuscript.

Please let me know if you have any further questions regarding any of these points. You can use the link below to upload the revised files.

Thank you again for giving us the chance to consider your manuscript for The EMBO Journal. I look forward to receiving the final version.

Referee #1:

I think the authors have carefully addressed my previous concerns and I do not have anything further to add.

The authors performed the requested editorial changes.

2nd Revision - Editorial Decision 14th May 2021

Editor accepted the revised manuscript.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: George Fu Gao
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2021-107786

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	NA
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Our data were obtained within the range of liner proportional change and, hence, are thought to show normal distribution.
Is there an estimate of variation within each group of data?	Yes. Data are represented as mean ± SD.

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repor

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumo

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is the variance similar between the groups that are being statistically compared?	Yes.

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Company and catalgoue numbers for all antibodies included in the study are provided in the manuscript: anti-His/APC antibodies (1:500, Miltenyi Biotec, AB_2751870) and anti-mIgG antibody (ZSGB-BIO, ZF-0513).
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Sf9 cells(Invitrogen, 11496015), Hi5 cells (Invitrogen, B85502), BHK cells (ATCC, ATCC CCL-10), HEK293T cells (ATCC, ATCC CRL-3216) and HEK293F cells (ATCC) were maintained in our laboratory. Cell lines were routinely tested for mycoplasma contamination in our laboratory.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The cryo-EM density maps and corresponding atomic coordinates have been deposited in the Electron Microscopy Data Bank (EMDB), Protein Data Bank (PDB) and the China Natinal Microbiology Data Center (NMDC), respectively. The accession numbers for the cryo-EM structures reported in this paper are: GX/P2V/2017 RBD-hACE2 (EMD-30653, https://www.emdataresource.org/EMD-30653; 7DDP, https://www.rcsb.org/structure/7DDP; NMDCS0000011, https://www.nmdc.cn/resource/ncov/structure/detail/NMDCS0000011), and GD/1/2019 RBD-hACE2 (EMD-30655, https://www.emdataresource.org/EMD-30655; 7DDO, https://www.rcsb.org/structure/7DDO; NMDCS0000012, https://www.nmdc.cn/resource/ncov/structure/detail/NMDCS0000012).
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA