



HAL
open science

FAIR sharing of molecular visualization experiences: from pictures in the cloud to collaborative virtual reality exploration in immersive 3D environments

Xavier Martinez, Marc Baaden

► To cite this version:

Xavier Martinez, Marc Baaden. FAIR sharing of molecular visualization experiences: from pictures in the cloud to collaborative virtual reality exploration in immersive 3D environments. 2020. hal-02944943v1

HAL Id: hal-02944943

<https://hal.science/hal-02944943v1>

Preprint submitted on 21 Sep 2020 (v1), last revised 24 Nov 2021 (v2)

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.

FAIR sharing of molecular visualization experiences: from pictures in the cloud to collaborative virtual reality exploration in immersive 3D environments

Authors

Xavier Martinez^{ab} and Marc Baaden^{ab*}

^aCNRS, Université de Paris, UPR 9080, Laboratoire de Biochimie Théorique, 13 rue Pierre et Marie Curie, Paris, 75005, France

^bInstitut de Biologie Physico-Chimique-Fondation Edmond de Rothschild, PSL Research University, Paris, France

Correspondence email: baaden@smplinux.de

Synopsis Visualization renders structural molecular data accessible to a broad audience. We describe an approach to share molecular visualization experiences based on FAIR principles. Our workflow is exemplified with recent Covid-19 related data.

Abstract Motivated by the current Covid-19 pandemic that has spurred a substantial flow of structural data we describe how molecular visualization experiences can be used to make these datasets accessible to a broad audience. Using a variety of technology vectors related to the cloud, 3D- and virtual reality gear, we examine how to share curated visualizations of structural biology, modeling and/or bioinformatics datasets for interactive and collaborative exploration. We discuss F.A.I.R. as overarching principle for sharing such visualizations. We provide four initial example scenes related to recent Covid-19 structural data together with a ready-to-use (and share) implementation in the UnityMol software.

Keywords: Molecular visualization; FAIR sharing; Covid-19; virtual reality; collaborative environments.

1. Introduction

The Covid-19 pandemic has spurred a wealth of new structural biology data related to the severe acute respiratory syndrome coronavirus (SARS-CoV)-2 retrovirus and its interactions with other macromolecules involved in the development and spread of the disease [1, 2]. Such structural data is transparent to experts in the field, but requires adequate visualization to become accessible to a broader audience. As put by Card et al., “visual artifacts aid thought; in fact, they are completely

entwined with cognition action” [3]. Visualizations prepared by experts may serve both educational and research purposes by highlighting key aspects of a given dataset or comparing features among several ones thereby supporting thought and reasoning. Here we aim to establish a general workflow to prepare such curated educational and research material, taking the advent of recent Covid-19 structural and modeling data as example. Our overall goal is to enable easy sharing of a given visual experience with others so that the same content can be accessed in various ways. We put focus on collaborative virtual reality (VR) as well as sharing 3D models and scenes to intuitively convey the spatial complexity of these molecular objects. Shape-related features bear particular relevance for drug design and hence for expert users, yet they can also be understood by an inexperienced person, in particular through a visual experience such as those described here. A particular aim is to treat the sharing of visual experiences with the same F.A.I.R. principles [4] as one should apply for the underlying raw chemical [5], structural and modelling data, and as is generally the case in crystallography [6]. This target is challenging by itself, as common FAIR-based sharing platforms do not provide a category that is well adapted to visual experiences.

The features of the workflow we are seeking to implement are schematically summarized in Figure 1. Starting from the original raw data that should encompass experimental structural datasets, molecular modelling results and bioinformatics findings, we aspire to prepare new shareable visual content to be set up and curated by an expert. To do so, we will use a molecular visualization software package to describe the visualization scene and all its annotations through command scripts. The experience can be customized through add-ons, for example providing dataset-specific user menus leading to a visual scene that can be explored directly in the software. This exploration does however require learning at least some basic usage of that software and how to manipulate the scene. This task can be simplified to some extent by designing appropriate simple-to-use custom user menus. As an alternative, the core content to be visualized can be exported to a variety of media such as still images, movies of various formats – possibly stereoscopic ones –, native 3D models or a fully interactive scene to be explored with advanced technology. We aim to support a broad variety of sophisticated hardware such as VR headsets, wall-size stereoscopic displays and the HoloLens, but also simple setups such as Google cardboard or even just running in a web browser. To render these media findable, accessible, interoperable and reusable, a variety of options are discussed.

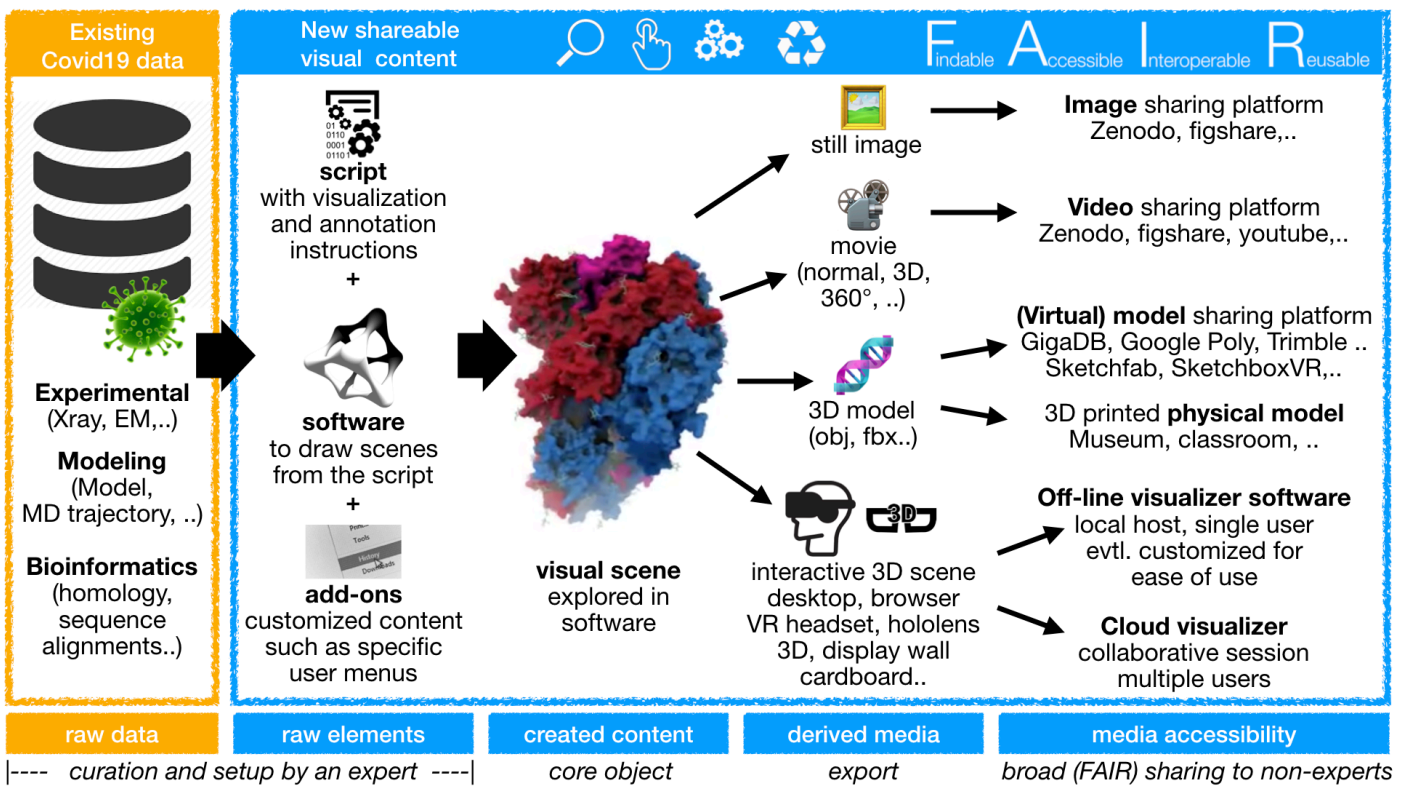
The set of (open source) software plus openly accessible scripts that we propose to provide make the visual experience as reproducible as possible [7]. Inherent limitations arise from technical hardware and software dependencies related to rapidly evolving visualization technology that are by nature short-lived with a high rate of change. We chose the UnityMol package to implement our proof of concept set of visualization experiences, as we have been developing and redesigning this software, which provides specific advantages for the tasks at hand in its latest version [8, 9]. Many alternative software options of high value could have been considered such as those recently reviewed for protein

visualization [10]. We opted for UnityMol as it provides a good compromise between feature-richness and ease of use. A wide variety of media and technology vectors is supported, while the software remains simple to use for a broad range of users due to its single-window design with only a few menus to operate.

Our ambition is to provide a few proof of concept implementations that others can build on subsequently to further improve or simply adapt them. We have implemented all functionalities that we deemed necessary to prepare the Covid-19-related materials, and present here four basic examples. Based on these example scenes, we experiment with applying F.A.I.R. principles to such visualization-centric data. One way to achieve this goal is by addressing the setup of the visualization experiences, i.e. the underlying software and the related scripts (designated as “raw elements” in Fig. 1), rather than the resulting media products. A more classical approach is to provide the derived media, which may consist in videos or images for which sharing platforms exist, although this is not enough to optimize FAIR compliance [11].

Figure 1 A scheme summarizing the suggested workflow for FAIR sharing of molecular visualization experiences, building on Covid-19-related structural data. Using existing data, new shareable visual content is created from raw elements such as a script with instructions, software to execute the script and eventual add-ons to facilitate the end-user tasks (for instance custom menus to provide only the strictly necessary shortcuts for a given experience). The visual scene that is created can be explored within the software itself or it can be exported in various ways to generate derived media such as images, movies, 3D models and even full experiences in e.g. VR contexts. These media can be rendered accessible through platforms that implement FAIR principles and ease their exploration.

Sharing of visualization experiences appears as a particularly timely topic. The case of the Covid-19 pandemic provides a stringent example for the need to render the latest research data explorable by the broad scientific community. If such data were only to be provided in its raw form, the accessibility would be quite limited. The worldwide Protein Data Bank [12] indeed provides quite a few visualizations along with the actual data files. Recently, the European PDB also proposed a visual analytics platform that is FAIR compliant [13]. Here we try to go one step further in the democratization of such data by preparing curated visual experiences as a valuable support for cross-disciplinary scientific exchange. In related fields, similar initiatives have recently emerged. Now more than ever, sharing data from molecular simulations and modeling is becoming a priority to ensure reproducibility and accelerate discovery [14]. An increasing number of initiatives aim to develop open access and reproducible molecular simulations (see for example [15–17] among other approaches).



Sharing such molecular dynamics (MD) simulation data equally needs efficient visualization and collaboration tools [18] for apprehending such data remotely to avoid unnecessary data transfers. Yet to the best of our knowledge, current initiatives focus on the actual data, not so much on the visualizations and annotation thereof. Concerning the specific case of Covid-19, many data portals and hubs were put together to provide more unified access to relevant data. A brief selection is listed in Table 1 and served as testing ground for our visualization experiments. Many of the listed portals provide a very direct access to the data we require for our visualization experiences, a few provide more general search engines with a less direct path to the actual dataset of interest. We did not include the classical routes to search data such as the PDB, GigaDB [19, 20] or Zenodo and we only included one of the many relevant initiatives by individual scientists or laboratories as example.

Table 1 A few examples of Covid-19 data portals.

Some of the data in these portals are likely to be cross-referenced or duplicated. The selection of portals is by no means exhaustive, with the number of available resources rapidly increasing.

Name	URL	Relevant data	Comments
EU COVID-19 Data Portal	https://www.covid19dataportal.org	Structures, Models	Additional information on sequences, expression data, compounds and targets
COVID-19 Molecular Structure and Therapeutics Hub	https://covid.molssi.org	Structures, Models, Simulations	Additional information on Targets and Therapeutics
BioExcel Covid-19 Research	https://bioexcel.eu/covid-19-research/	Structures, Models, Simulations	Additional information on Targets and Therapeutics
BioExcel-CV19	https://bioexcel-cv19.bsc.es/#/	Simulations	Web-access to atomistic-MD trajectories
PDBe-KB COVID-19 Data Portal	https://www.ebi.ac.uk/pdbe/covid-19	Ligand binding sites, protein-protein interaction residues	
Coronavirus Structural Task Force	https://github.com/thornlab/coronavirus_structural_task_force	Structures	
SwissModel	https://swissmodel.expasy.org/repository/species/2697049	Models	
CASP_Commons	https://predictioncenter.org	Models	

	/caspccommons/index.cgi		
COVID Moonshot	https://covid.postera.ai/covid/structures	Structures (fragment-oriented)	Links back to other databases such as Fragalysis or RCSB
OpenAIRE COVID-19 Gateway	https://beta.covid-19.openaire.eu	Various research datasets	A search engine referencing many resources
DE Shaw Research Resources	https://www.deshawresearch.com/downloads/download_trajectory_sarscov2.cgi	Simulations	

2. Materials and Method

Here we will briefly describe key functionalities implemented for the purpose of sharing molecular visualization experiences. As an overarching principle, we want to be able to generate a broad variety of media, including still images and videos based on these visual scenes describing the core of the contents, capturing them from within the running software, ideally in an executable build and for a handful of technically more involved features from within the Unity3D editor environment. For this purpose we implemented, extended and stabilized several features. A certain number of those features are still experimental only.

We added a python console and API for scripting within UnityMol, providing an easy way to formalize and share the narrative of a given experience, as well as facilitating the reproducibility of all necessary steps leading to a scene. We furthermore implemented the possibility to generate custom user menus so that complex and fine-tuned functionality can be achieved with the click of a button. A possible practical use could be to group shortcuts to relevant structural ensembles in light of Covid-19 for visual analysis, comparison or even manipulation. The latter can be achieved through interactive simulations: UnityMol implements the generic IMD protocol for steering molecular simulations through MDDriver [21], and provides a few in-app implementations such as a rigid protein-protein docking module.

The scenes can be annotated by the user through several functionalities such as text bullets or free hand drawing. A guided tour function was implemented to walk the user through a set of pre-determined key frames in a fully interactive experience.

A particular focus concerns the possibility to generate effectful high-quality graphics, for instance in order to convey the complex shapes of molecular surfaces using ambient occlusion (AO) or to draw the user's attention to a certain part of the structure of interest using photography effects such as depth-of-field (DOF). In addition to the molecular structures themselves, their properties such as electrostatic field can be depicted as well [9]. In the latest version of the software we added an

experimental feature enabling interactive raytracing using the OSPRay [22] package. This is illustrated for our first example of the spike trimer complex with the ACE2 enzyme (Figure 2A). We added a custom menu to simplify navigating through this example (Figure 2B). We explored the scene collaboratively in a 3-user virtual reality session (Figure 2C) where avatars represent each participant. To capture such sessions, still images from screenshots or movies from video capture can be created at user defined resolution. These media can be generated and exported with specific options such as a 360 degree view or a 3D stereoscopy enhancement. For these purposes, specific cameras had to be added within Unity. In general, image and video exports do however not preserve the full complexity of the initial object. For this purpose we added the export of textured 3D polygonal objects in the common obj, fbx and probably soon also glTF formats. Special effects such as AO and DOF cannot currently be exported. Such 3D models can be transformed into tangible physical objects through 3D printing [23].

Figure 2 Illustration of three types of visual experiences within UnityMol. **A)** Interactive raytracing of a simple structural view of the spike-ACE2 complex (example 1). A series of four images at distinct time intervals is shown to depict the ongoing work of the raytracer. After 25 s a denoiser pass is executed, significantly increasing the visual clarity of the scene. The raytracer continues indefinitely to improve the scene. The fluidity and interactivity are only limited by the CPU-bound performance. Indicated timing was obtained with a 2013 3,5 GHz 6-Core Intel Xeon E5. **B)** A custom menu extension for example 1, providing the possibility to toggle annotations on/off along with direct access to different representations and viewpoints. **C)** Screenshot of a collaborative multi-user session where 3 participants are examining example 1 with virtual reality headsets and controllers.

3. Results

We have started the process of building up a collection of Covid-19 example scripts with several objectives: they should be easy to understand, even for non-specialists; others should be able to re-run them under various conditions in terms of visualization hardware and modality and (in the future) it should be simple to use them as starting scene for shared multi-user sessions. A particular emphasis lies on the design of visual experiences suitable and tested for virtual reality (VR) exploration as our current research is focused on this technology. In our opinion its potential is still largely under-explored in academia and industry, be it for research, collaboration or education purposes [24]. To illustrate the sharing of such visual experiences, we have prepared four typical datasets depicted in Figure 2 and Figure 3): a simple structural view, a collection of small molecules binding to a viral protein, an MD trajectory and a bioinformatics sample with conservation as well as mutation data.

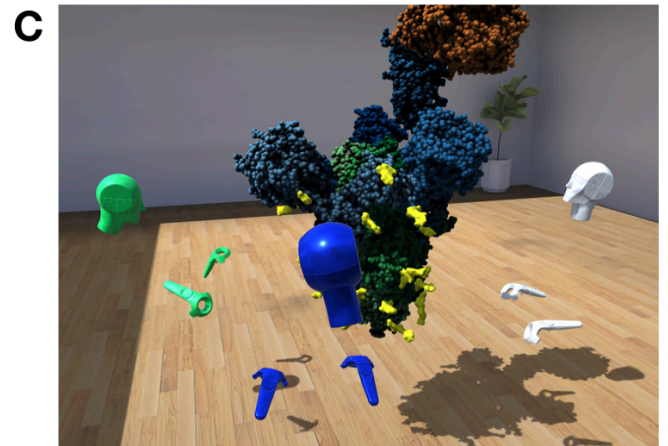
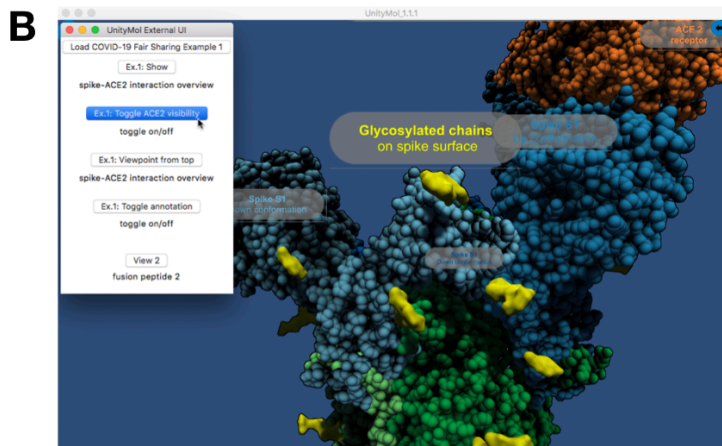
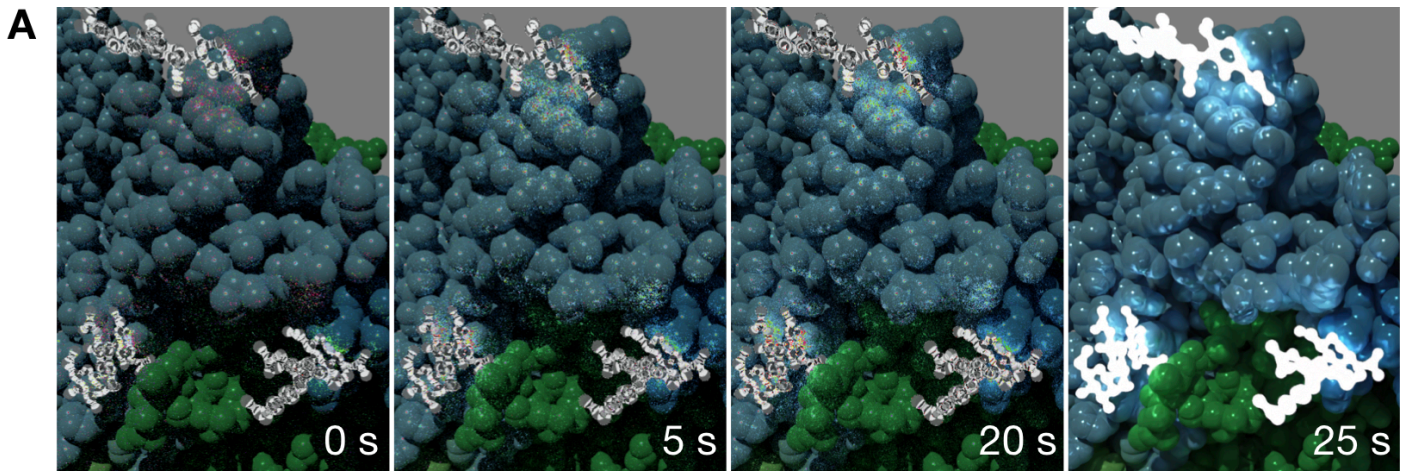


Figure 3 A glimpse into examples 2 to 4. **A)** Example 2 depicts drug molecules (shown in glowing yellow) binding to the dimeric biological assembly of Covid-19 main protease (monomers shown in blue surface and brown cartoon) in 92 protein databank structures. The image is a snapshot taken from an animation of the whole dataset. The insets show a 3D printed translucent monomer model (bottom) and how one may illuminate it (right) to highlight specific areas with a laser (red spot). The orientation of the right image is similar to the one of the brown monomer subunit shown on the left. **B)** Example 3 renders a molecular dynamics simulation of the SARS-CoV-2 spike receptor-binding domain binding to the ACE2 receptor [25]. A depth-of-field effect is used to draw attention to the binding interface. **C)** Example 4 represents results on species variability of the binding interface from a bioinformatics analysis. The interacting residues are colored in a pink color scale according to the Haddock docking score from [26] and their size is varied according to residue conservation assessed through a Shannon entropy measure.

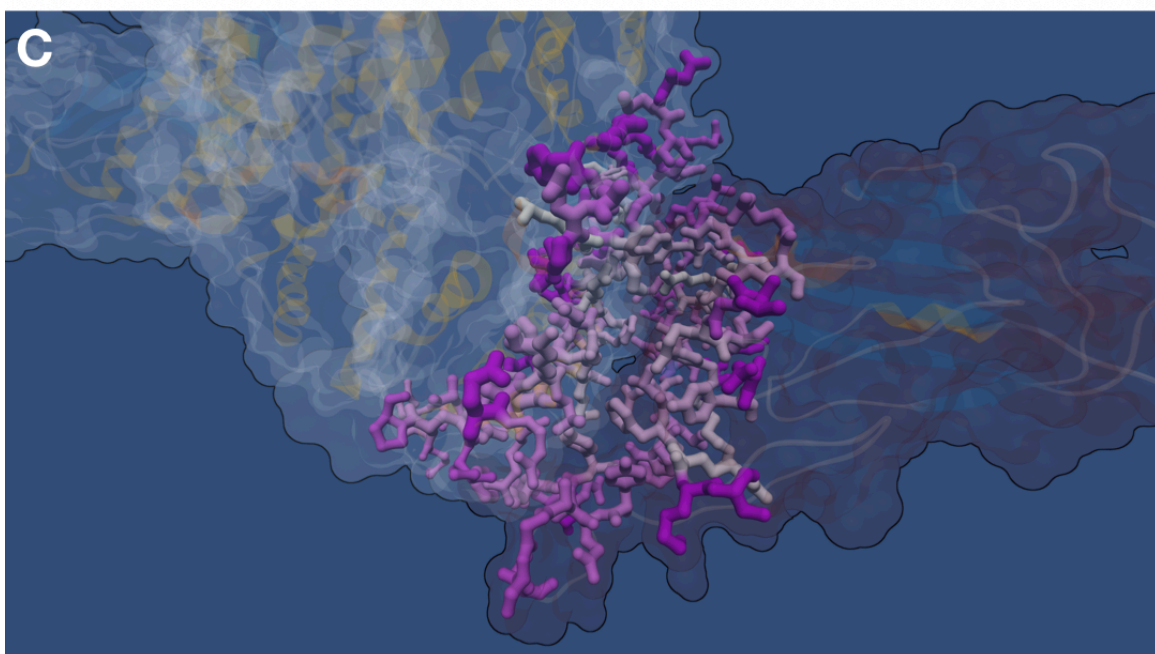
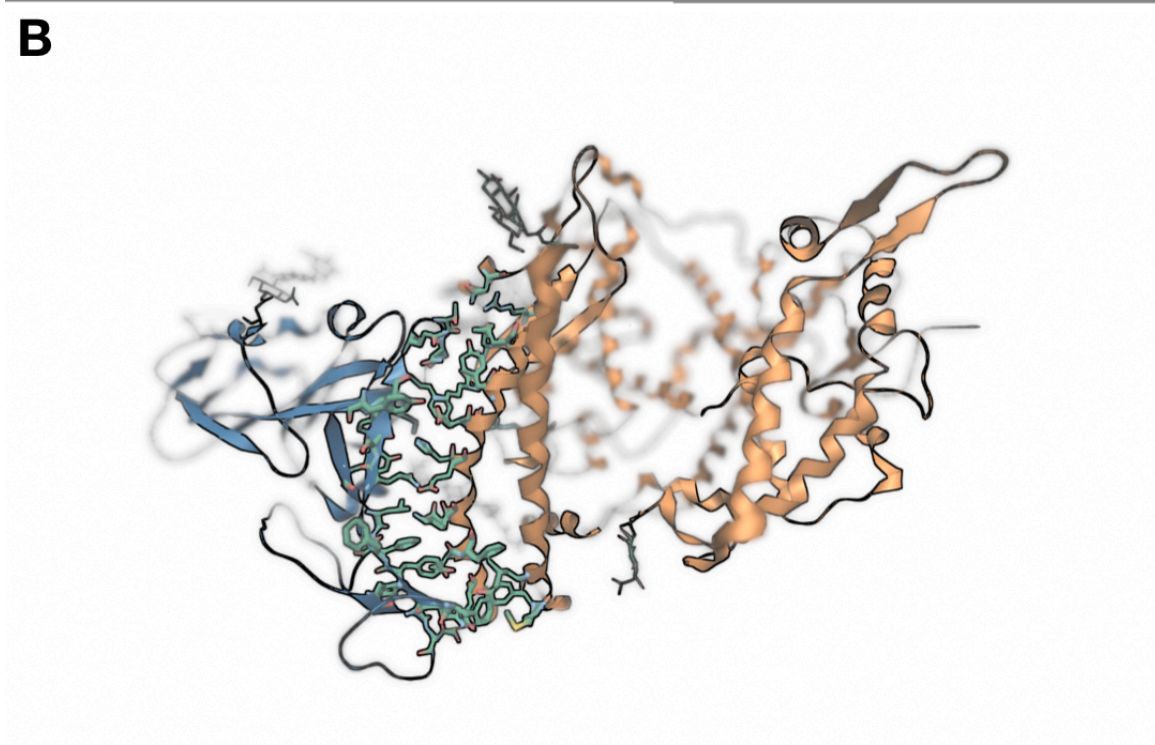
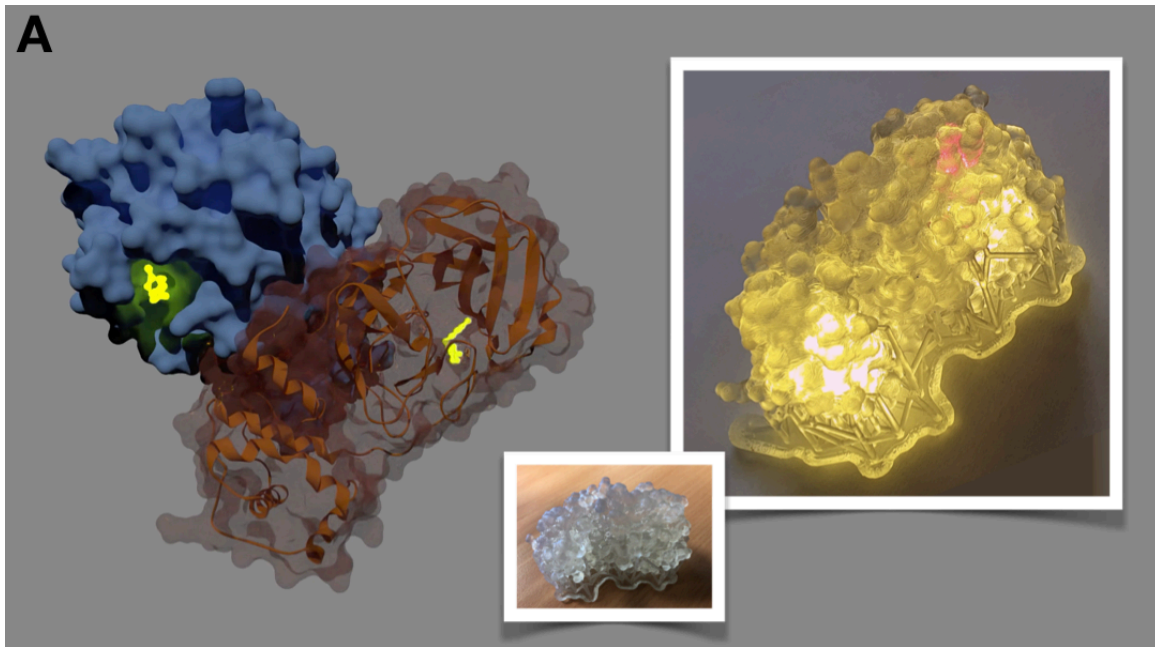
The first example depicted in Figure 2 aims at setting up a simple structural view of the SARS spike glycoprotein complex with human angiotensin-converting enzyme 2 (ACE2). The trimeric complex accessible from PDB-id 6CS2 [27] was used as starting point for many early SARS-CoV-2 studies.

We then have a comparative look at where drug molecules bind to the main protease of Covid-19 in our second example (Figure 3A). The visualization is inspired by the animation of small molecules in 92 protein databank structures (<https://www.rbvi.ucsf.edu/chimerax/data/sars-protease-may2020/>) prepared by the ChimeraX [28] team. We created a 3D print of the protease monomer as well.

Molecular dynamics simulations provide another inestimable resource for insight into molecular mechanisms. Example 3 (Figure 3B) is based on a trajectory depicting a binding event of the receptor-binding domain (RBD) of the SARS-CoV-2 spike and the human ACE2 receptor (DESRES-ANTON-[10857295,10895671] of ref. [25]).

The fourth example illustrates a visual experience related to representing the results of bioinformatics analyses. Our example is built upon the freely available data from a recent study on cross-species transmission of SARS-CoV-2, highlighting the species variability of viral-host protein interactions [26]. We visually map this data onto the ACE2-RBD complex (Figure 3C).

The first level of sharing these four examples is to make the software and the scripts themselves available. We considered the zenodo and GigaDB platforms and chose zenodo as main platform for our experiment. Typically, no dedicated category for “visual experiences” exists in these databases, whereas the classical ones such as software, dataset, image, video, workflow or other only partially reflect this case. As a first approach, we created a Zenodo community (Figure 4A) on FAIR sharing of



molecular visualization experiences at <https://zenodo.org/communities/fair-molvisexp/>. We then linked a github repository for our visualization scripts to Zenodo and the collection. We produced a set of derived materials including pictures, movies and 3D objects. Cloud-platforms for pictures and movies are very common nowadays and we will not go into much detail on this aspect. We used both the zenodo and figshare platforms (<http://figshare.com>) for pictures and for movies, the former one for consistency and the latter because of its visually oriented interface and popularity. To regroup our productions on figshare, we created a collection (<https://doi.org/10.6084/m9.figshare.c.5101400>) shown in Figure 4B. We also experimented with tangible physical objects as a particularly accessible form for 3D models through 3D printing. Many generic 3D printing databases exist, whereas the NIH 3D print exchange [29] is a specific research initiative dedicated to bioscientific 3D prints. We deposited our models in this resource (Figure 4C). Such 3D models can also be disseminated and shared virtually, which is a less common process. We experimented with four platforms: Sketchfab (<https://sketchfab.com> ; Figure 4D), Google Poly (<https://poly.google.com> ; Figure 4E), SketchBoxVR (<http://sketchboxvr.com>) and Trimble 3D warehouse (<https://3dwarehouse.sketchup.com>). For the time being we used the dataset, software, figure and video categories to reference all information on Zenodo. We have created a versioned catalogue page of all the outcomes of these experiments at https://github.com/bam93/fair_covid_molvisexp/tree/master/overview. Once all the functionalities described here are stabilized in the UnityMol software, we will provide instructions on how to export to these platforms on the same catalogue page.

Figure 4 A visual overview of content sharing platforms hosting our visual experiences to make them FAIR-compliant. **A)** depicts a zenodo community regrouping all the content created for this work (<https://zenodo.org/communities/fair-molvisexp/>). **B)** is a figshare collection of a subset of this content (<https://doi.org/10.6084/m9.figshare.c.5101400>). **C)** illustrates the entry of our 3D printed model from Fig. 3A on the NIH 3D print exchange (<https://3dprint.nih.gov/discover/3DPX-014787>). **D)** illustrates a 3D model uploaded to Sketchfab (<https://skfb.ly/6UFOw>) with the possibility to edit AR and VR viewing settings, as the platform automatically enables access to mobile phone and desktop virtual reality rendering in addition to web-based exploration. **E)** shows the same 3D model rendered in the Google Poly web-based explorer (<https://poly.google.com/view/5zsJiglTWbm>). **C**, **D** and **E** are accessible through the respective referenced 3D model datasets on zenodo, an stl file for 3D printing and an fbx file for 3D model building.

Referencing our datasets on these platforms only represent the first, relatively easy step towards implementing the FAIR principles. The main issues may lie in the interoperability and reusability

A

zenodo

FAIR sharing of molecular visualization experiences

Recent uploads

Search FAIR sharing of molecular visualization experiences

Community

Simple structural view of human angiotensin-converting enzyme 2

Baaden, Marc; Martinez, ...

A set of 5 screenshots of Unity multi-user virtual reality sessions drive the example more effectively.

Uploaded on August 25, 2020

August 25, 2020 (91) Files

multi3.jpg (514.99 kB)

multi2.jpg (518.91 kB)

multi1.jpg (444.56 kB)

menu.png

B

figshare

Search on figshare...

U.S. Department of Health and Human Services — National Institutes of Health

NIH 3D Print Exchange

SARS-CoV-2 main protease 3D print model

Submitted by: baaden

View Edit

Download all (3.37 MB)

Simple structural view of glycoprotein complex converting enzyme 2

Figure posted on 25.08.2020, 12:18 by Marc Baaden

A set of 5 screenshots of Unity multi-user virtual reality sessions drive the example more effectively.

umol-menu.png (1.08 MB)

multi3.jpg (514.99 kB)

multi2.jpg (518.91 kB)

multi1.jpg (444.56 kB)

menu.png

C

NIH 3D Print Exchange

SARS-CoV-2 main protease 3D print model

Submitted by: baaden

View Edit

Download all (3.37 MB)

Simple structural view of glycoprotein complex converting enzyme 2

Figure posted on 25.08.2020, 12:18 by Marc Baaden

A set of 5 screenshots of Unity multi-user virtual reality sessions drive the example more effectively.

umol-menu.png (1.08 MB)

multi3.jpg (514.99 kB)

multi2.jpg (518.91 kB)

multi1.jpg (444.56 kB)

menu.png

GENERAL INFORMATION

- Model ID 3DPX-014787
- Category Proteins, Macromolecules and
- Keyword(s) unitymol fair sharing
- A 3D model for printing SARS-CoV-2
- Protein Data Bank ID 5SQN
- Model Origin Molecular data (e.g., cry

D

sketchfab.com

Currently editing Spike-ACE2 interaction overview

Help Center EDIT SAVE SETTINGS

USE UNITY + SKETCH

RESET SCALE

SAVE VIEW

This will define the default model scale for viewing in VR and AR. After saving, it will take some time to regenerate the AR versions of the model. Learn more here

Tips for world scale

- Character is 1.75m tall. A floor tile is 1m².
- A smaller world makes your model look larger.
- A large world makes your model look smaller.

Display floor for teleportation

Floor level: 0m

Tips for floor level

- Position the floor right below large models.
- Position the floor further down for

E

polypeople.com

Search for things

Spike-ACE2 interaction overview

Like Share Edit Download Create GIF

aspects for the (meta)data associated with the visual experiences. For example, metadata is needed to describe software version, script purpose, underlying datasets, dependencies and the variants of output that can be produced. There are definitely many ways to go about implementing the FAIR principles. In this early attempt at visual experiences, we first and foremost want to raise awareness about this yet overlooked category, but do not ambition to provide a full-fledged solution. Concerning the fact that our data should be Findable, the choice to target Covid-19 makes this goal relatively easy to achieve as there is an abundance of hubs for increasing the findability of such data as described briefly above (Table 1). The Accessible attribute should be taken care of by our choice of well-established FAIR-compliant databases for most of the produced media. Interoperability is clearly an area for further improvement. In particular we consider bridges between different visualization environments as an important way forward. Initiatives fostering universal protocols that can be interchanged such as MolQL for the query language to define selections [30] need to be extended to describe all ingredients needed for a visual experience, in particular the molecular scene representation. As a modest step in this direction we have experimented with a Pymol session interpreter for UnityMol (<https://github.com/LBT-CNRS/PymolToUnityMol>), so that Pymol users can re-use their previously prepared scenes. Reproducibility is intrinsically a difficult aim for visual experiences. However, by providing scripts with full instructions for the setup of the experience and by making the corresponding visualization software available as versioned open source project along with ready-to use executable builds for many operating systems and platforms, reproducibility and accessibility are maximized. Furthermore, the software is conceived within a game engine, which by design is easily extensible, with a shallow learning curve.

4. Outlook

From this early experiment, we plan to expand the range of example visual experiences, be it based on the context of the rapidly growing Covid-19 data or more generally the overwhelming structural, modeling and dynamic data of macromolecular assemblies. We are trying to optimize the generated objects so that they can run on devices with limited hardware specifications. An example would be the HoloLens, with good ergonomics but limited graphical power [31] for which we want to be able to export simplified FBX/gltf polygonal models to run smoothly. Most of our efforts will be dedicated to refine the multi-user collaborative features to enable visual sessions for joint exploration.

More generally, our contribution calls for considering whether visual experiences should be assigned a specific category to be added to FAIR sharing platforms in order to take into account the specificities of such interactive, computer-supported graphical representations of research data.

Acknowledgements This work was supported by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011 and grant "CACSIce", ANR-11-EQPX-0008). X.M. and M.B. thank Sesame Ile-de-France for co-funding the display wall used for data analysis. X.M. and M.B. thank UCB Biopharma for support. We thank Nawel Khenak for help with 3D printing and Nicolas Férey and Antoine Taly for experimentations with laser and LED-lighting of translucent 3D-printed molecules.

References

1. Baker EN (2020) Visualizing an unseen enemy; mobilizing structural biology to counter COVID-19. *Acta Crystallogr Sect Struct Biol* 76:311–312 .
<https://doi.org/10.1107/S2059798320004830>
2. Dömling A, Gao L (2020) Chemistry and Biology of SARS-CoV-2. *Chem* S2451929420301959 . <https://doi.org/10.1016/j.chempr.2020.04.023>
3. Card SK, Mackinlay J, Shneiderman B (1999) Readings in information visualization: using vision to think. 1999. San Franc Morgan Kauffman
4. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3:160018 .
<https://doi.org/10.1038/sdata.2016.18>
5. Blanke, Gerd, Doerner, Thomas, Lynch, Nick (2020) Chemical Data in Life Sciences R&D and the FAIR Principles. <https://doi.org/10.5281/ZENODO.3970745>
6. Helliwell JR (2019) FACT and FAIR with Big Data allows objectivity in science: The view of crystallography. *Struct Dyn* 6:054306 . <https://doi.org/10.1063/1.5124439>
7. Cohen-Boulakia S, Belhajjame K, Collin O, Chopard J, Froidevaux C, Gaignard A, Hinsén K, Larmande P, Bras YL, Lemoine F, Mareuil F, Ménager H, Pradal C, Blanchet C (2017) Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. *Future Gener Comput Syst* 75:284–298 . <https://doi.org/10.1016/j.future.2017.01.012>
8. Lv Z, Tek A, Da Silva F, Empereur-mot C, Chavent M, Baaden M (2013) Game On, Science - How Video Game Technology May Help Biologists Tackle Visualization Challenges. *PLoS ONE*

8:e57990 . <https://doi.org/10.1371/journal.pone.0057990>

9. Laureanti J, Brandi J, Offor E, Engel D, Rallo R, Ginovska B, Martinez X, Baaden M, Baker NA (2020) Visualizing biomolecular electrostatics in virtual reality with UnityMol-APBS. *Protein Sci Publ Protein Soc* 29:237–246 . <https://doi.org/10.1002/pro.3773>
10. Martinez X, Chavent M, Baaden M (2020) Visualizing protein structures — tools and trends. *Biochem Soc Trans* 48:499–506 . <https://doi.org/10.1042/BST20190621>
11. Weber T, Kranzlmuller D (2018) How FAIR Can you Get? Image Retrieval as a Use Case to Calculate FAIR Metrics. In: 2018 IEEE 14th International Conference on e-Science (e-Science). IEEE, Amsterdam, pp 114–124
12. Berman H, Henrick K, Nakamura H (2003) Announcing the worldwide Protein Data Bank. *Nat Struct Mol Biol* 10:980–980 . <https://doi.org/10.1038/nsb1203-980>
13. PDBe-KB consortium, Varadi M, Berrisford J, Deshpande M, Nair SS, Gutmanas A, Armstrong D, Pravda L, Al-Lazikani B, Anyango S, Barton GJ, Berka K, Blundell T, Borkakoti N, Dana J, Das S, Dey S, Micco PD, Fraternali F, Gibson T, Helmer-Citterich M, Hoksza D, Huang L-C, Jain R, Jubb H, Kannas C, Kannan N, Koca J, Krivak R, Kumar M, Levy ED, Madeira F, Madhusudhan MS, Martell HJ, MacGowan S, McGreig JE, Mir S, Mukhopadhyay A, Parca L, Paysan-Lafosse T, Radosky L, Ribeiro A, Serrano L, Sillitoe I, Singh G, Skoda P, Svobodova R, Tyzack J, Valencia A, Fernandez EV, Vranken W, Wass M, Thornton J, Sternberg M, Orengo C, Velankar S (2020) PDBe-KB: a community-driven resource for structural and functional annotations. *Nucleic Acids Res* 48:D344–D353 . <https://doi.org/10.1093/nar/gkz853>
14. Amaro RE, Mulholland AJ (2020) A Community Letter Regarding Sharing Biomolecular Simulation Data for COVID-19. *J Chem Inf Model* *acs.jcim.0c00319* . <https://doi.org/10.1021/acs.jcim.0c00319>
15. Abraham M, Apostolov R, Barnoud J, Bauer P, Blau C, Bonvin AMJJ, Chavent M, Chodera J, Čondić-Jurkić K, Delemotte L, Grubmüller H, Howard RJ, Jordan EJ, Lindahl E, Ollila OHS, Selent J, Smith DGA, Stansfeld PJ, Tiemann JKS, Trellet M, Woods C, Zhmurov A (2019) Sharing Data from Molecular Simulations. *J Chem Inf Model* 59:4093–4099 . <https://doi.org/10.1021/acs.jcim.9b00665>
16. Elofsson A, Hess B, Lindahl E, Onufriev A, van der Spoel D, Wallqvist A (2019) Ten simple rules on how to create open access and reproducible molecular simulations of biological systems. *PLOS Comput Biol* 15:e1006649 . <https://doi.org/10.1371/journal.pcbi.1006649>
17. Abriata LA, Lepore R, Dal Peraro M (2020) About the need to make computational models of biological macromolecules available and discoverable. *Bioinformatics* *btaa086* . <https://doi.org/10.1093/bioinformatics/btaa086>
18. Tiemann JKS, Guixà-González R, Hildebrand PW, Rose AS (2017) MDsrv: viewing and sharing molecular dynamics simulations on the web. *Nat Methods* 14:1123–1124 . <https://doi.org/10.1038/nmeth.4497>

19. Sneddon TP, Li P, Edmunds SC (2012) GigaDB: announcing the GigaScience database. *GigaScience* 1:11, 2047-217X-1-11 . <https://doi.org/10.1186/2047-217X-1-11>
20. Xiao SZ, Armit C, Edmunds S, Goodman L, Li P, Tuli MA, Hunter CI (2019) Increased interactivity and improvements to the *GigaScience* database, GigaDB. *Database* 2019: . <https://doi.org/10.1093/database/baz016>
21. Delalande O, Férey N, Grasseau G, Baaden M (2009) Complex molecular assemblies at hand via interactive simulations. *J Comput Chem* 30:2375–2387 . <https://doi.org/10.1002/jcc.21235>
22. Wald I, Johnson G, Amstutz J, Brownlee C, Knoll A, Jeffers J, Gunther J, Navratil P (2017) OSPRay - A CPU Ray Tracing Framework for Scientific Visualization. *IEEE Trans Vis Comput Graph* 23:931–940 . <https://doi.org/10.1109/TVCG.2016.2599041>
23. Gillet A, Sanner M, Stoffler D, Olson A (2005) Tangible Interfaces for Structural Molecular Biology. *Structure* 13:483–491 . <https://doi.org/10.1016/j.str.2005.01.009>
24. Morrison J (2016) Will chemists tilt their heads for virtual reality? *Chem Eng News* 94:22–23
25. D. E. Shaw Research (2020) Molecular Dynamics Simulations Related to SARS-CoV-2
26. Rodrigues JP, Barrera-Vilarmau S, Teixeira JM, Seckel E, Kastiris P, Levitt M (2020) Insights on cross-species transmission of SARS-CoV-2 from structural modeling. *Bioinformatics*
27. Kirchdoerfer RN, Wang N, Pallesen J, Wrapp D, Turner HL, Cottrell CA, Corbett KS, Graham BS, McLellan JS, Ward AB (2018) Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci Rep* 8:15701 . <https://doi.org/10.1038/s41598-018-34171-7>
28. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci Publ Protein Soc* 27:14–25 . <https://doi.org/10.1002/pro.3235>
29. Coakley MF, Hurt DE, Weber N, Mtingwa M, Fincher EC, Alekseyev V, Chen DT, Yun A, Gizaw M, Swan J, Yoo TS, Huyen Y (2014) The NIH 3D Print Exchange: A Public Resource for Bioscientific and Biomedical 3D Prints. *3D Print Addit Manuf* 1:137–140 . <https://doi.org/10.1089/3dp.2014.1503>
30. Rose AS, Sehnal D, Bliven S, Burley SK, Velankar S (2018) Molql: Towards a Common General Purpose Molecular Query Language. *Biophys J* 114:342a . <https://doi.org/10.1016/j.bpj.2017.11.1910>
31. Müller C, Krone M, Huber M, Biener V, Herr D, Koch S, Reina G, Weiskopf D, Ertl T (2018) Interactive Molecular Graphics for Augmented Reality Using HoloLens. *J Integr Bioinforma* 15: . <https://doi.org/10.1515/jib-2018-0005>

