

Scipy 2020, Austin TX | @virtualScipy



# Molecular infrastructure for modeling viruses with pythonic-mediated packages | pyF4all |

Horacio Andrés Vargas Guzmán, Ph.D.

Github: pyF4all | @govarguz

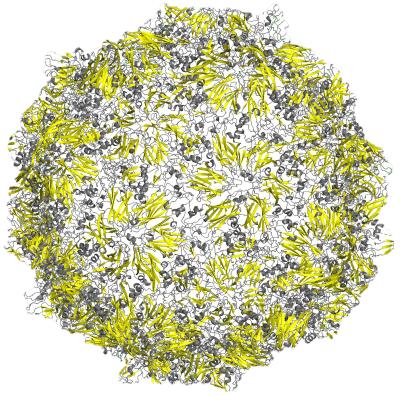
# What this talk is about...

>python | modeling | structure | dynamics | properties

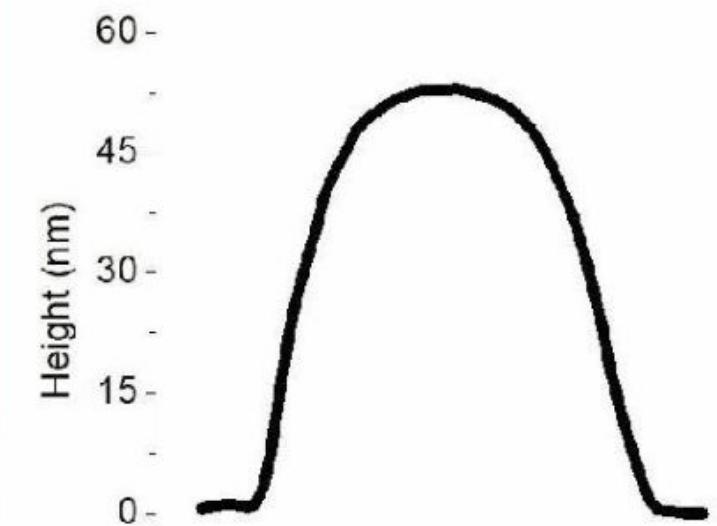
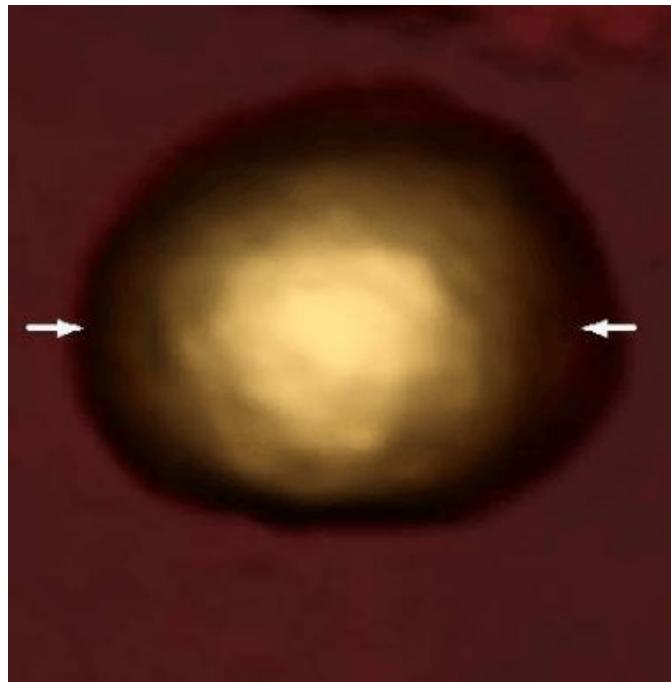
- An introduction to multiscale simulation methods for modeling viruses
- A brief discussion of tools/packages available in python to perform those simulations (and an outlook to what's coming next)
- A rationale for the use of Jupyter-notebooks for a collaborative environment towards modeling full viruses

# What this talk is about...

>structure | dynamics | properties

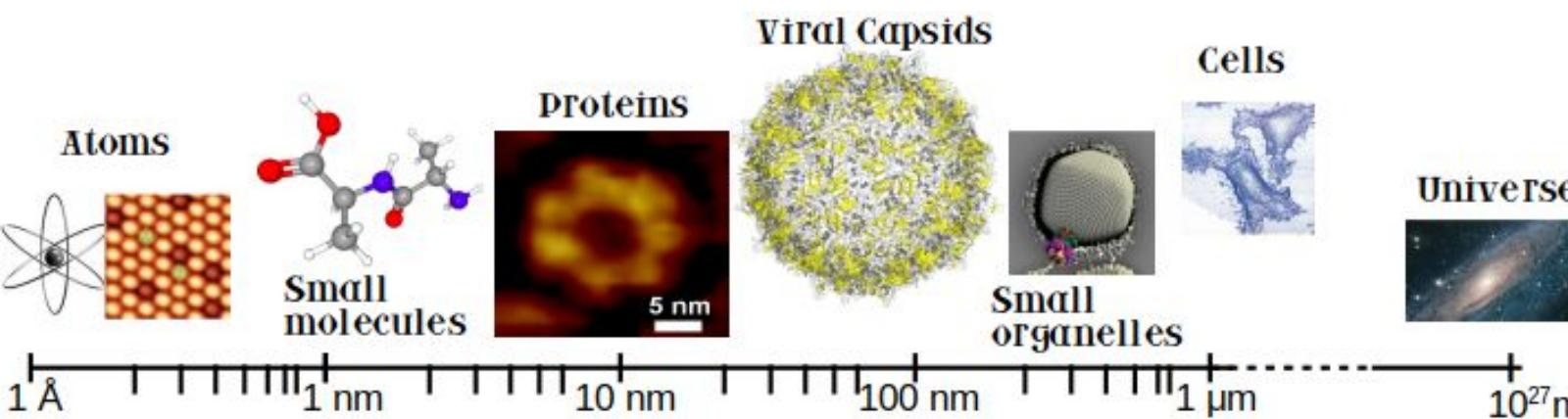


STRUCTURES

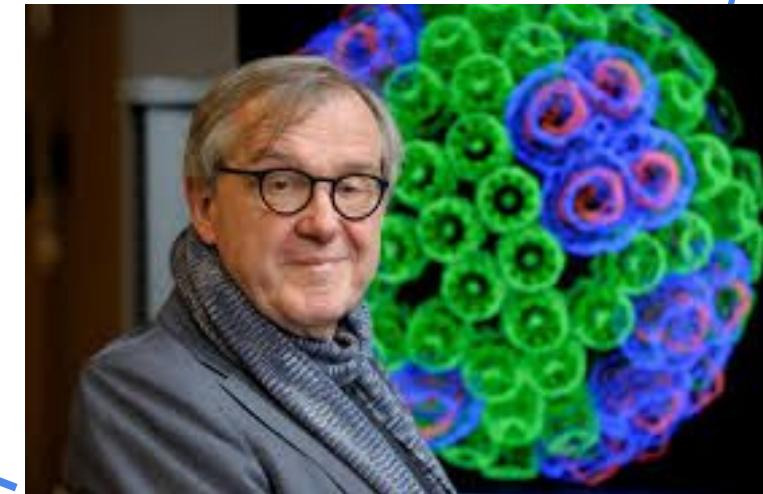


DYNAMICS

# Computational microscope | all-atom MD

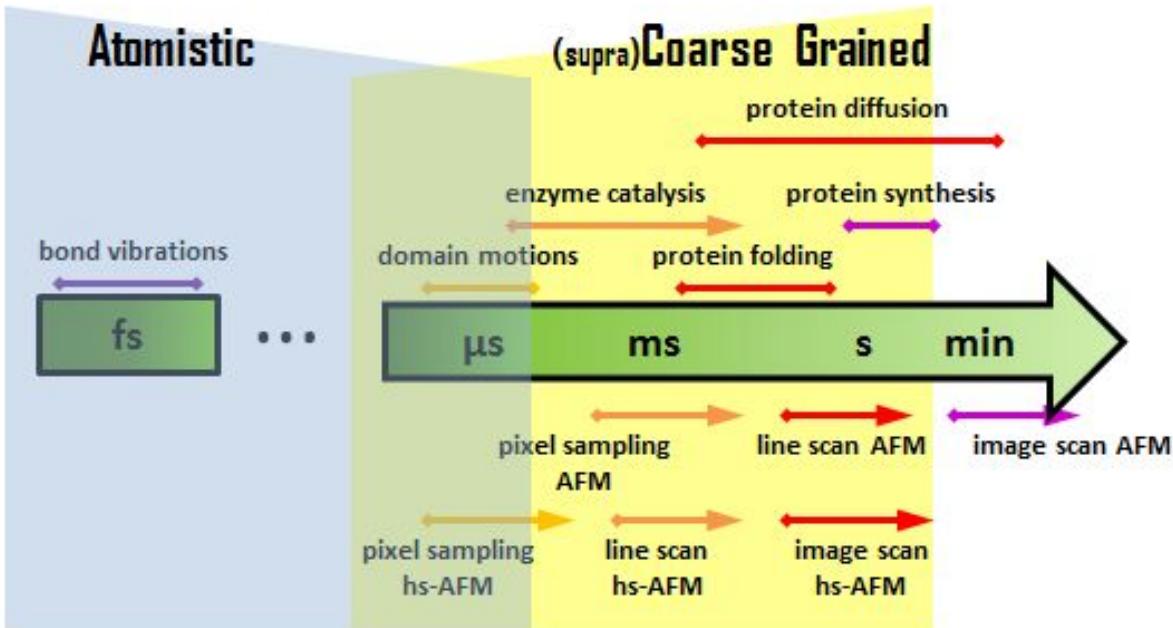


*processing speed on its  
1,572,480 cores to 94.6  
petaFLOPS (bluewaters)*



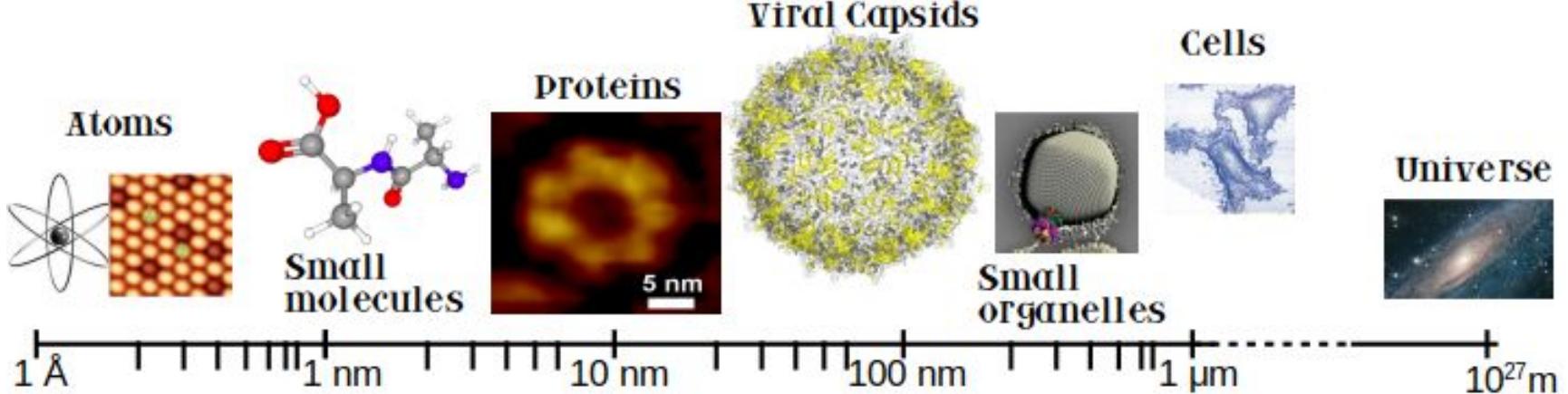
J. Perilla, ... K. Schulten, *Curr. Opinion In Struc. Biol.* (2015)

# A multiscale journey, time and length wise...

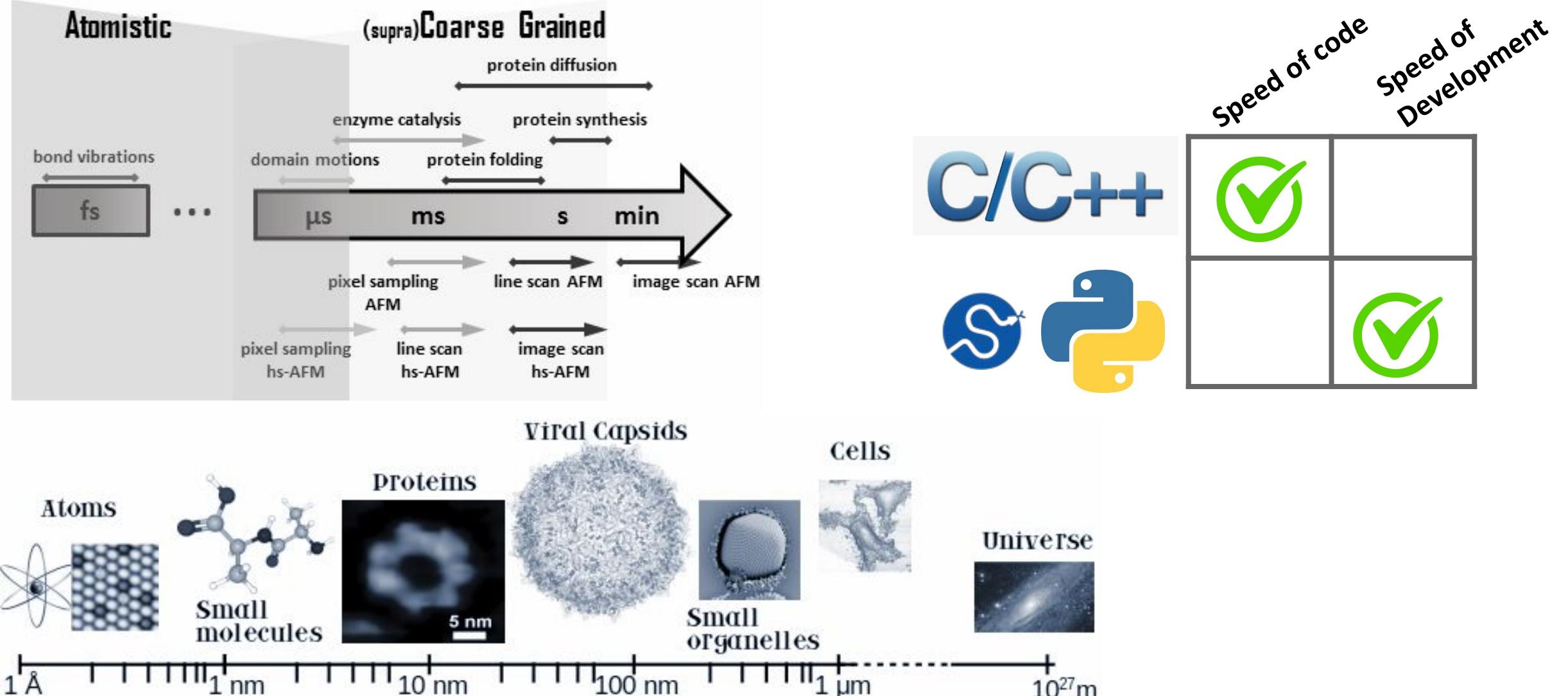


## TIME SCALES

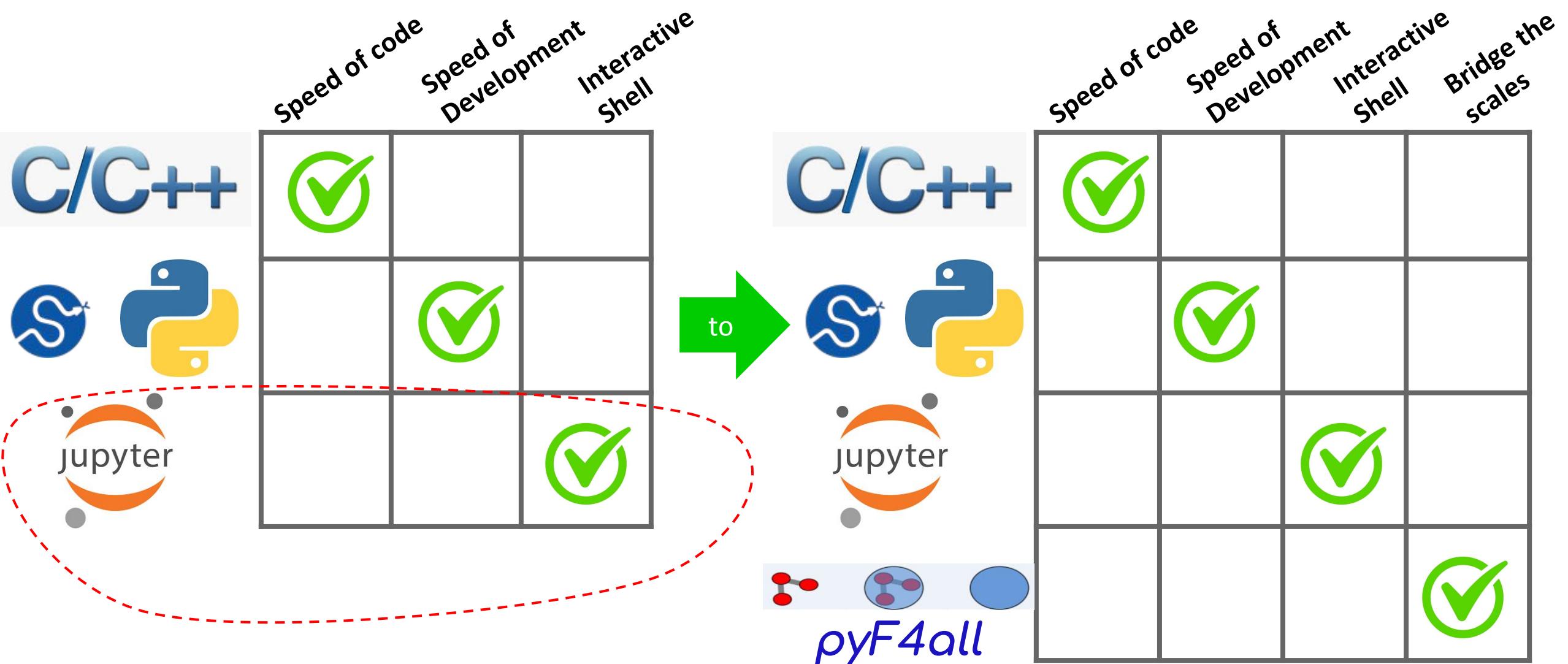
## LENGTH SCALES



# A multiscale journey, time and length wise...and coding wise

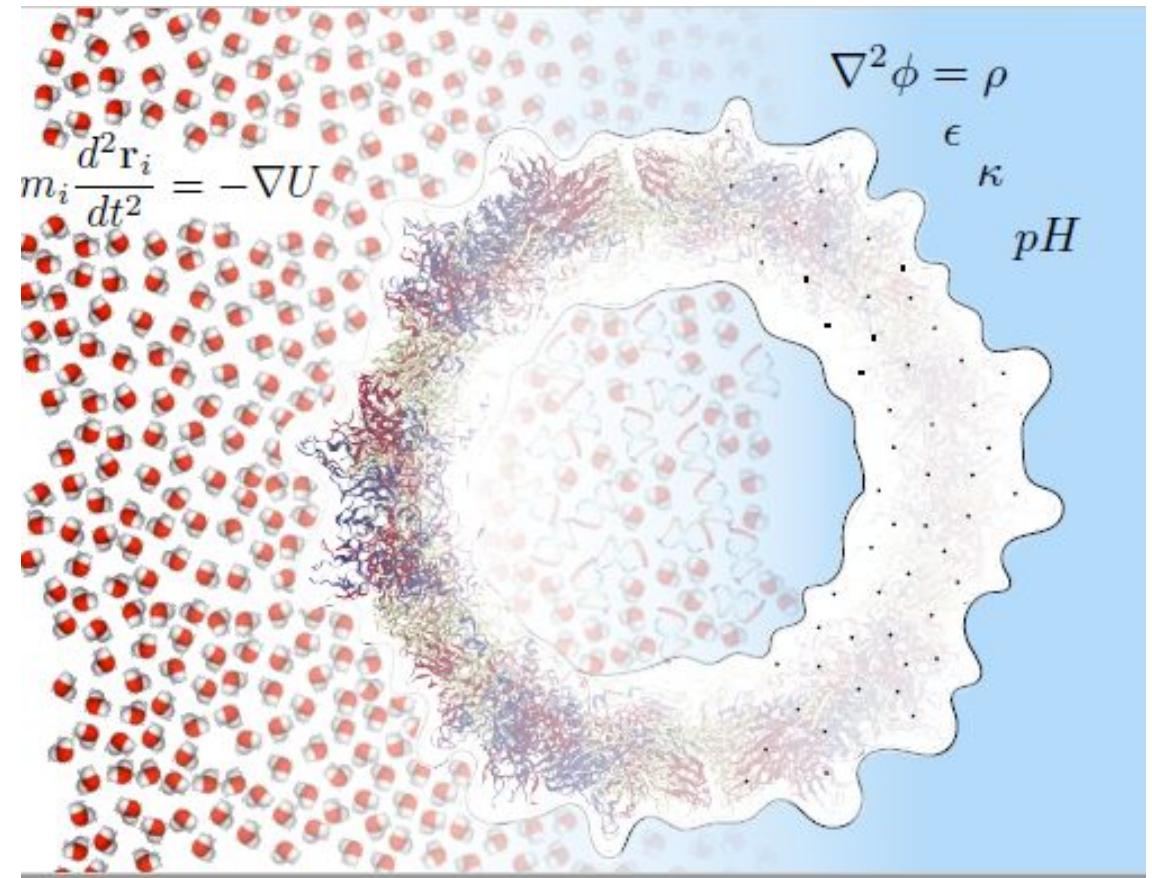
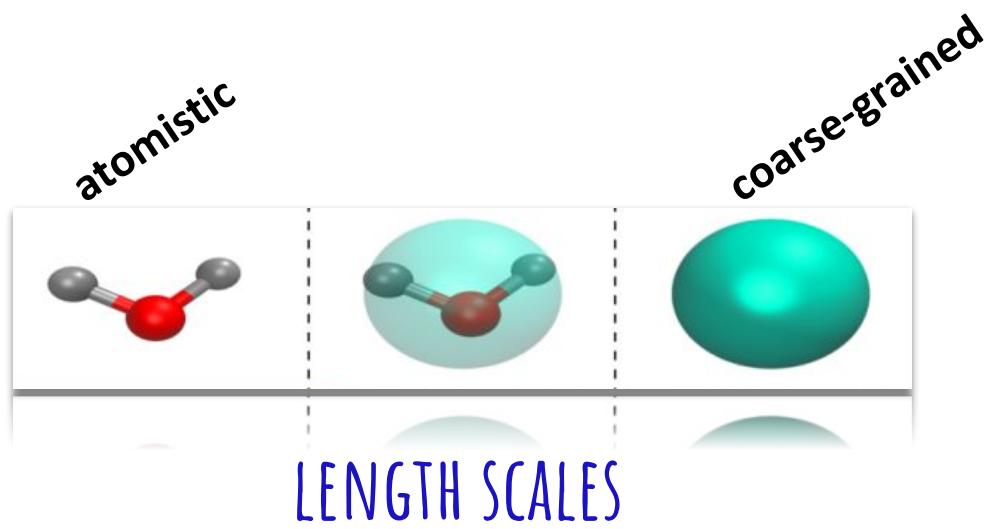


# Code-wise, open-digital infrastructure!

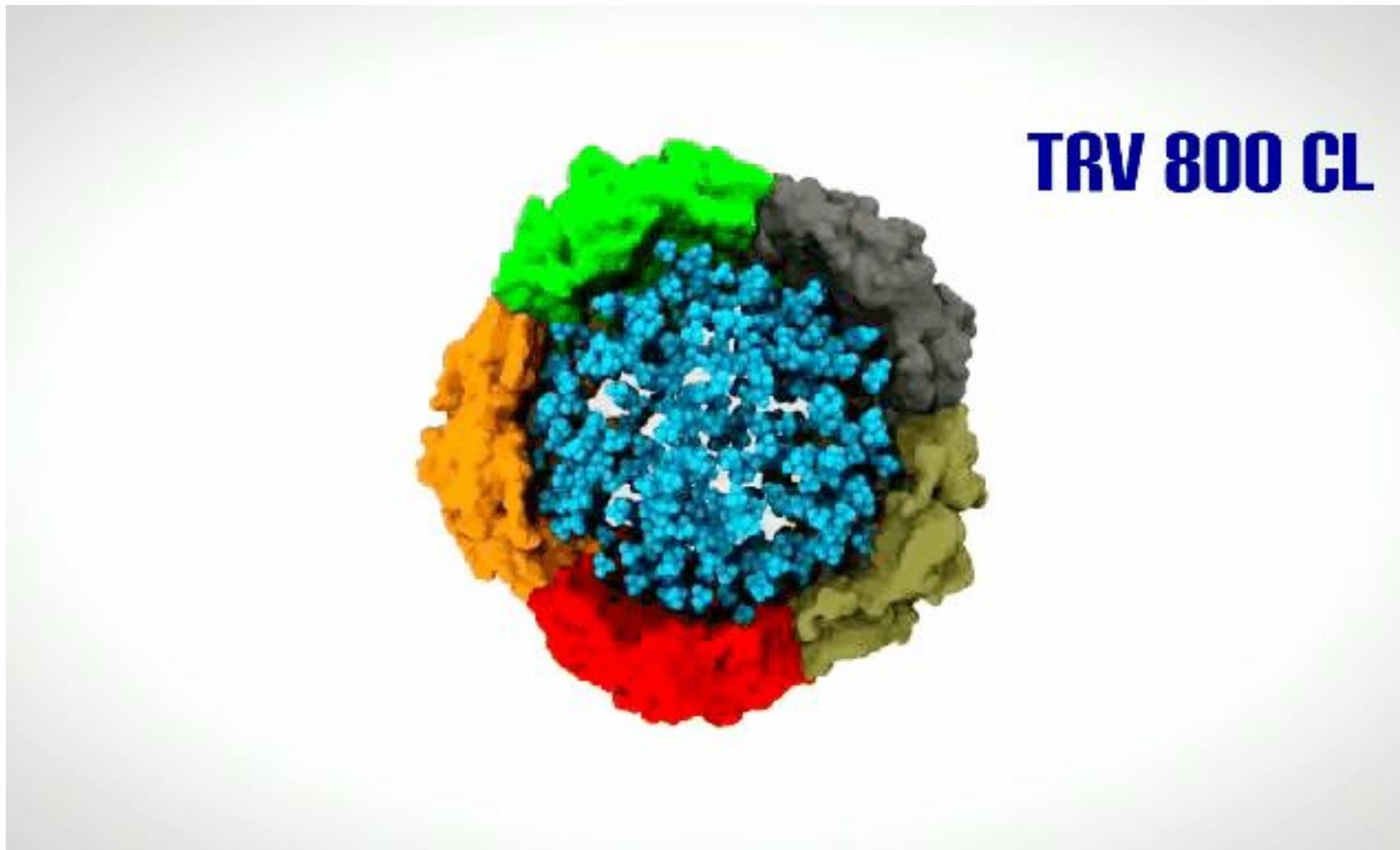


# Model premises

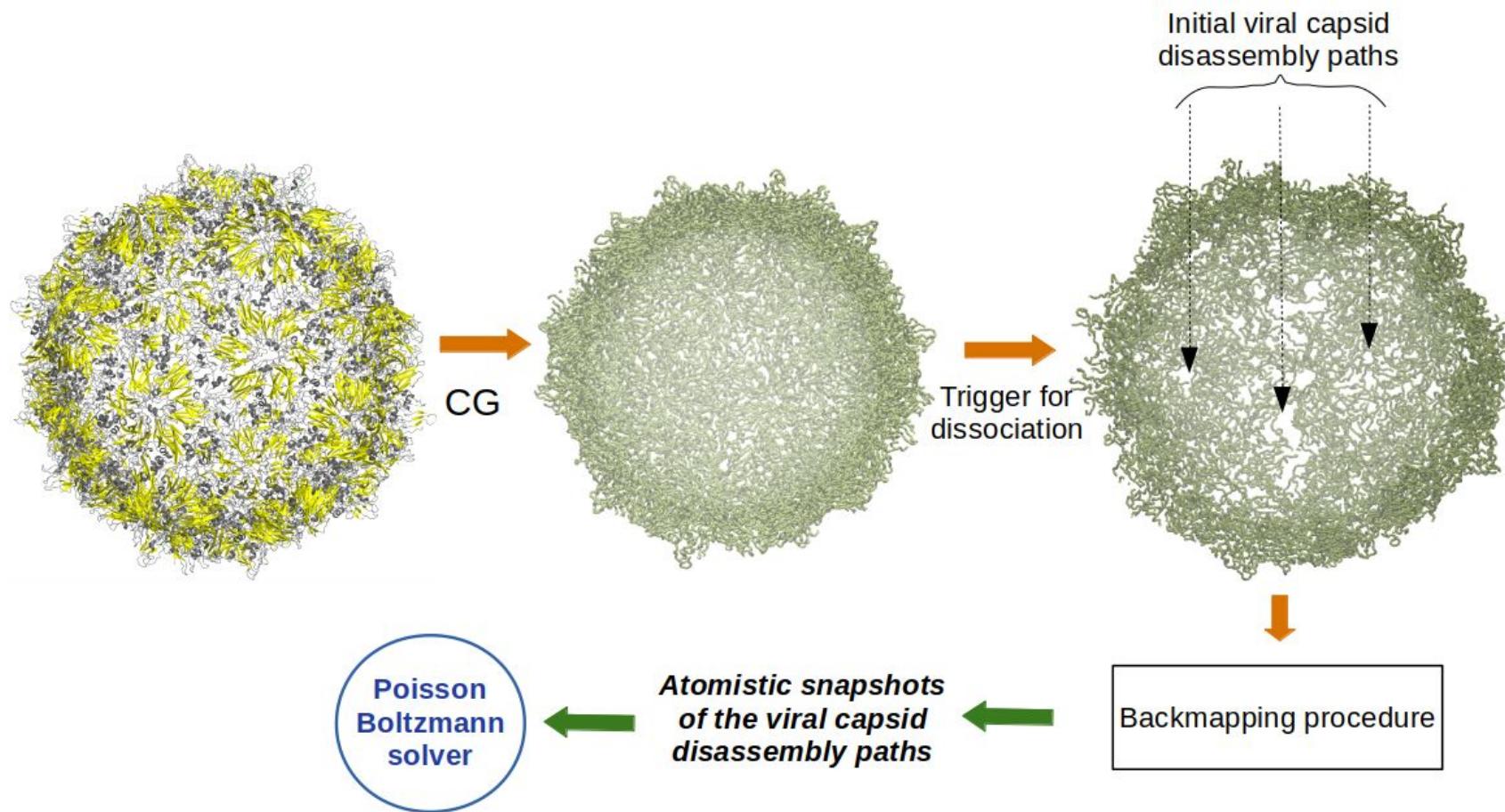
- Hypothesis: Go beyond the “proof of concept”
- Provide and interpret quantitative observables
- Go along with experimental DEVS  
on the acquisition of molecular  
resolved static and dynamic imaging



# Molecular multiscale trajectories



# A multiscale model for quantifying environment-mediated free energy landscapes



# The method: Free energies of the disassembly of viral capsids (Poisson-Boltzmann)

$$\Delta G_{\text{solv}} = \int_{\Omega_3} \rho \phi_{\text{reac}} = \sum_{k=1}^{N_q} q_k \phi_{\text{reac}}(\mathbf{r}_k)$$

$$G_{\text{coul}} = \frac{1}{2\epsilon_3} \sum_{k=1}^{N_q} \sum_{j=1}^{N_q} q_k q_j \frac{1}{4\pi |\mathbf{r}_k - \mathbf{r}_j|}$$

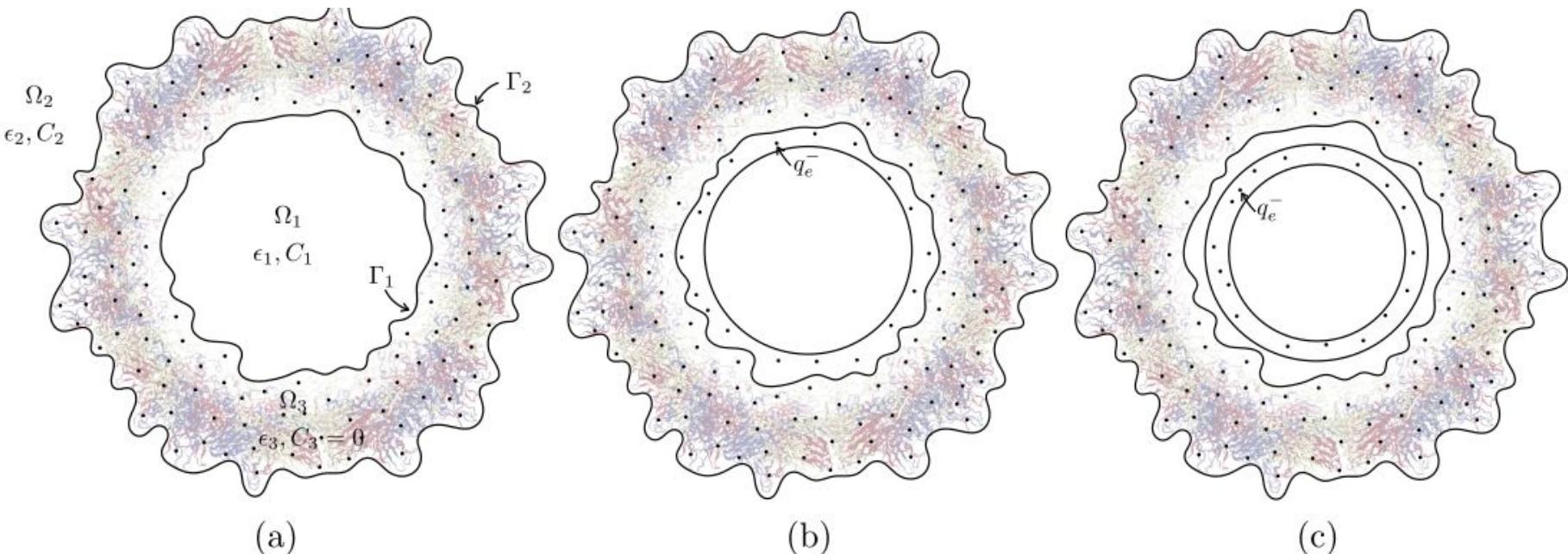
$$(\nabla^2 - \kappa_1^2) \phi_1 = 0 \text{ on } \Omega_1$$

$$(\nabla^2 - \kappa_2^2) \phi_2 = 0 \text{ on } \Omega_2$$

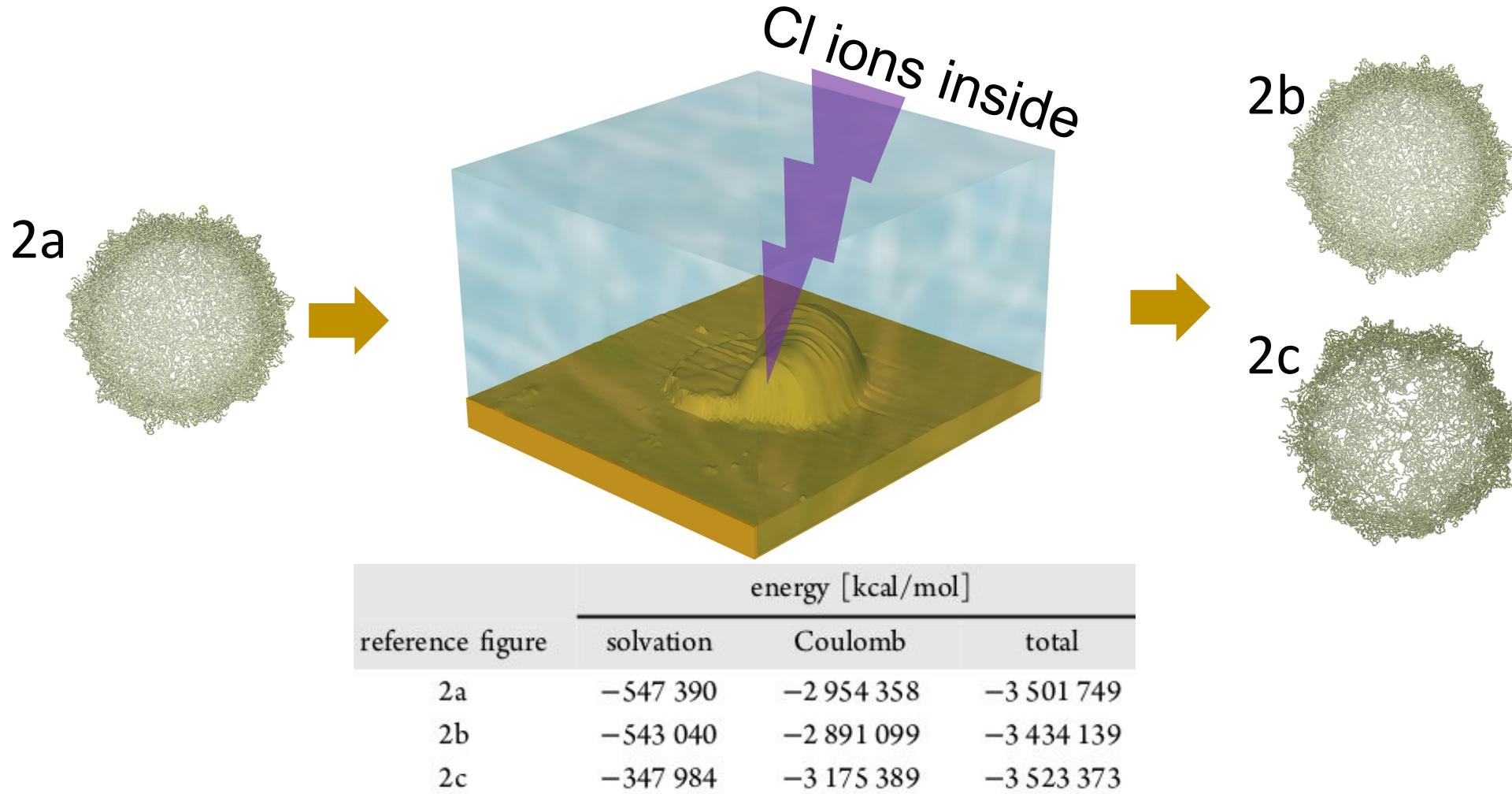
$$\nabla^2 \phi_3 = \sum_{k=1}^{N_q} q_k \delta(\mathbf{r}_k) \text{ on } \Omega_3$$

$$\phi_1 = \phi_3 \text{ and } \epsilon_1 \frac{\partial \phi_1}{\partial \mathbf{n}} = \epsilon_3 \frac{\partial \phi_3}{\partial \mathbf{n}} \text{ at } \Gamma_1$$

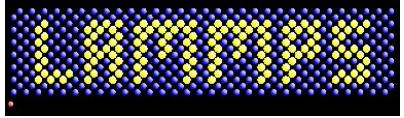
$$\phi_2 = \phi_3 \text{ and } \epsilon_2 \frac{\partial \phi_2}{\partial \mathbf{n}} = \epsilon_3 \frac{\partial \phi_3}{\partial \mathbf{n}} \text{ at } \Gamma_2$$



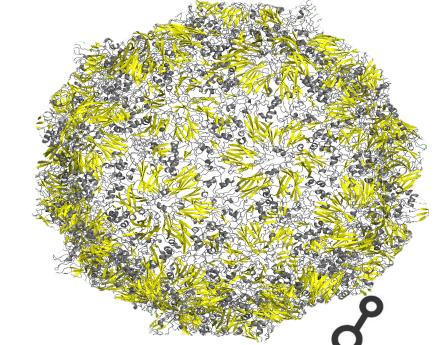
# Free energies of the disassembly of viral capsids



# pyF4all - Multiscale molecular glue (ecosystem)

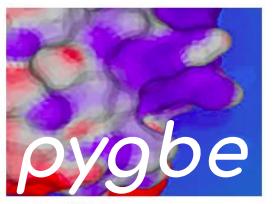
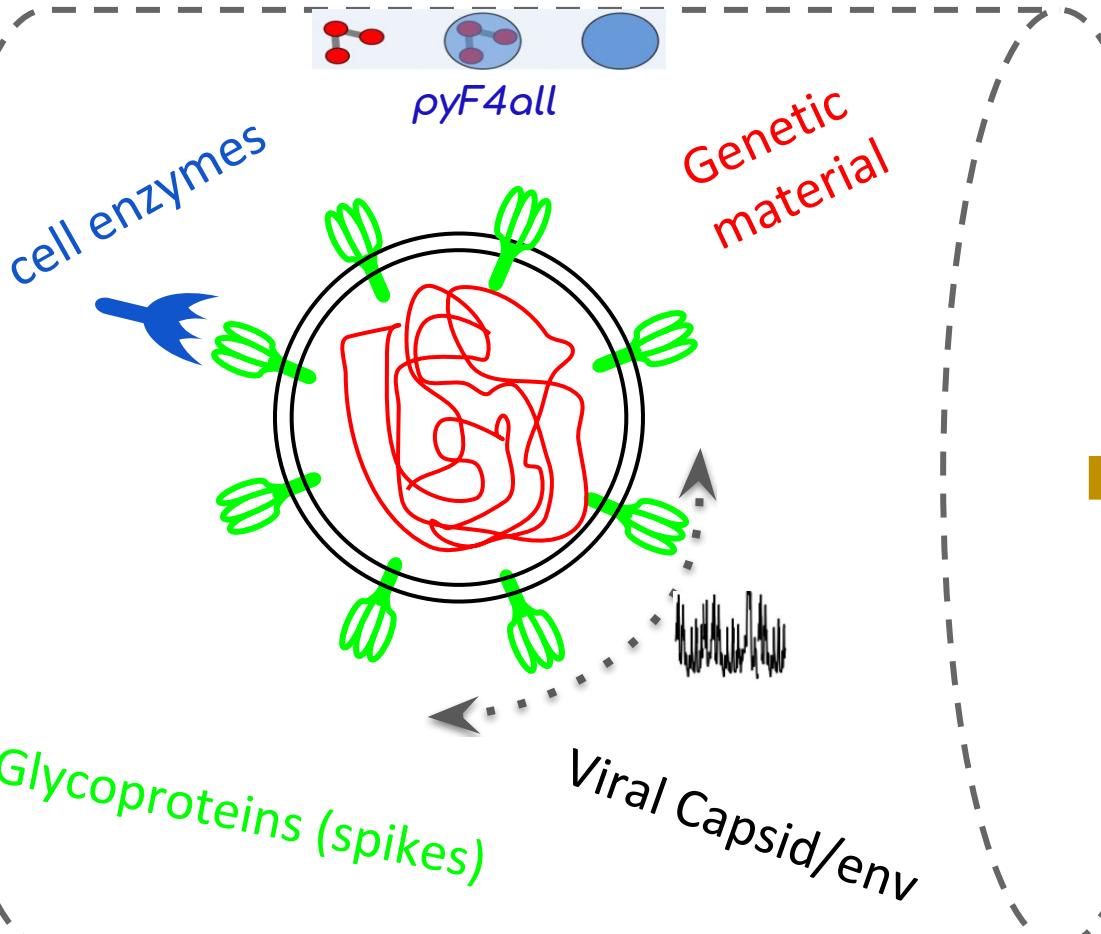


**GROMACS**  
fast, flexible & free



MolSSI  
bioexcel  
RCSB PDB  
PROTEIN DATA BANK

Among many other  
MD packages..



**NANOSHAPER**

**pdb2pqr**

Among many more...



Quantitative analysis

$$f(\Omega, \Gamma, \nabla pH, \epsilon_r, E_{eff})$$



# Take home message

- pyF4all provides free energy calculations by employing small cluster/GPU-workstation, which extends the access to biomolecular simulation and analysis (towards ubiquity)
- Multiscale approaches make biomolecular simulations feasible, once we can interact between length/time-scales rapidly and systematically
- Without the sharing initiative of sharing Biomolecular Simulation Data to help identify leads for therapies, diagnostics and vaccines. As an action to combat diseases, like COVID-19. As it is the MolSSI and BioExcel initiative

# On-going projects and status

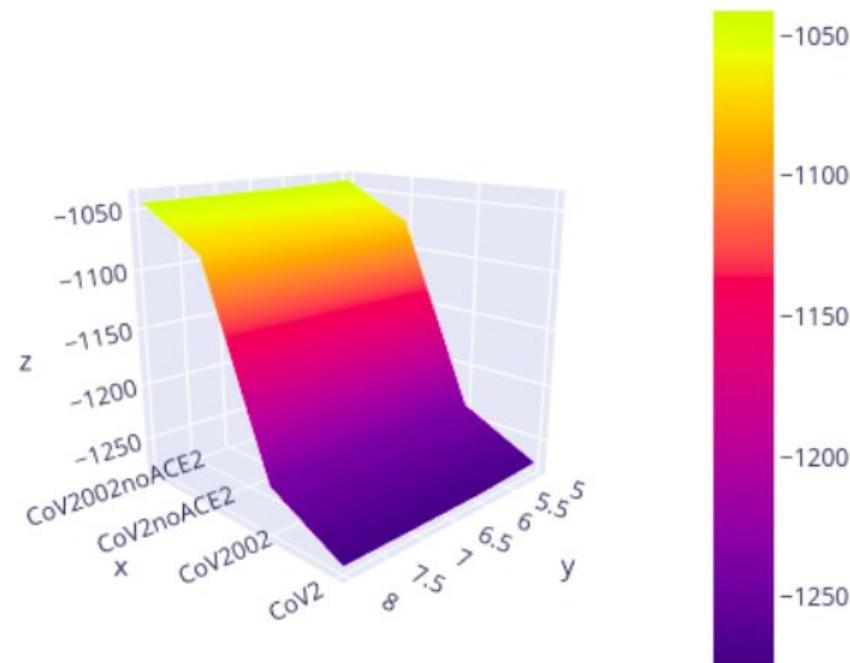
Connect to a cloud server with available GPUs (or local GPU)

```
In [ ]: import numpy,scipy,matplotlib  
!pip install pycuda,MDAnalysis,Nanoshaper,pdb2pqr  
!apt-get install swig -y  
!git clone https://github.com/barbagroup/pygbe.git
```

```
In [ ]: import pyF4all
```

Coronaviruses and comparing their thermodynamic cycles

```
In [ ]: model=pyF4all.compare.pH("SARS-COV-2","SARS-COV-2wACE2","SARS-COV-2002","SARS-COV-2002wACE2",[5.0,6.0,7.0,8.0])  
model.plot.pH(structures,pH,Free_Energies)
```



# On-going projects and status

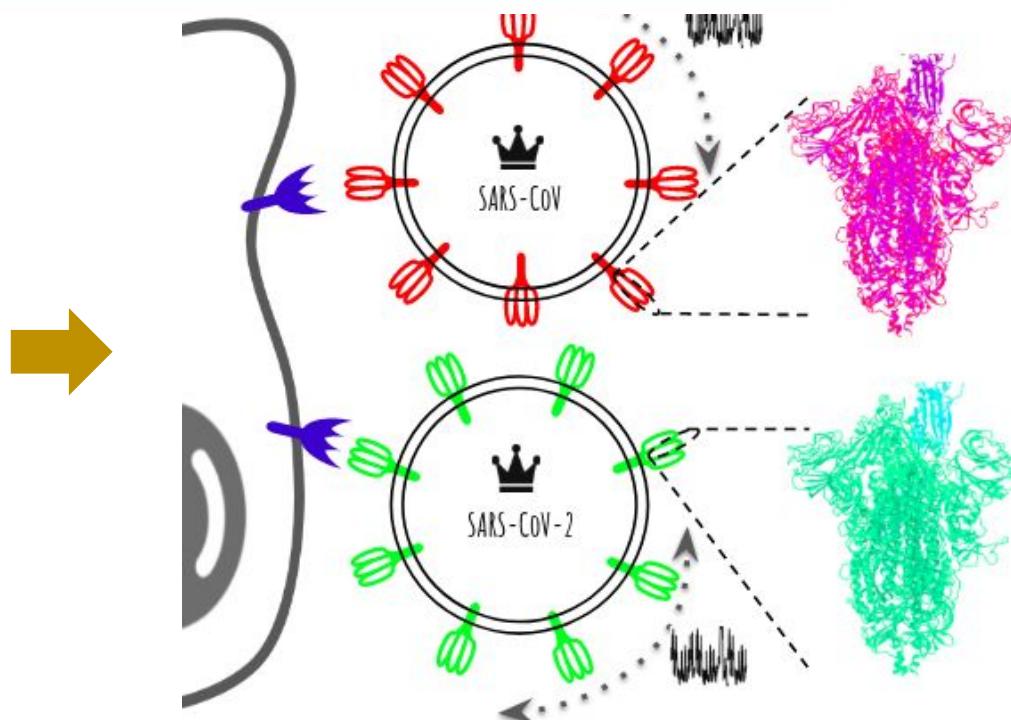
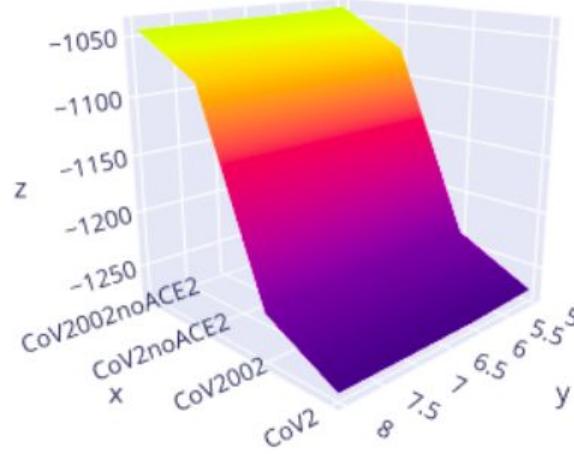
Connect to a cloud server with available GPUs (or local GPU)

```
In [ ]: import numpy,scipy,matplotlib  
!pip install pycuda,MDAnalysis,Nanoshaper,pdb2pqr  
!apt-get install swig -y  
!git clone https://github.com/barbagroup/pygbe.git
```

```
In [ ]: import pyF4all
```

Coronaviruses and comparing their thermodynamic cycles

```
In [ ]: model=pyF4all.compare.pH("SARS-COV-2","SARS-COV-2wACE2","SARS-COV-2002","SARS-COV-2002wACE2",[5.0,6.0,7.0,8.0])  
model.plot.pH(structures,pH,Free_Energies)
```



# On-going projects and status

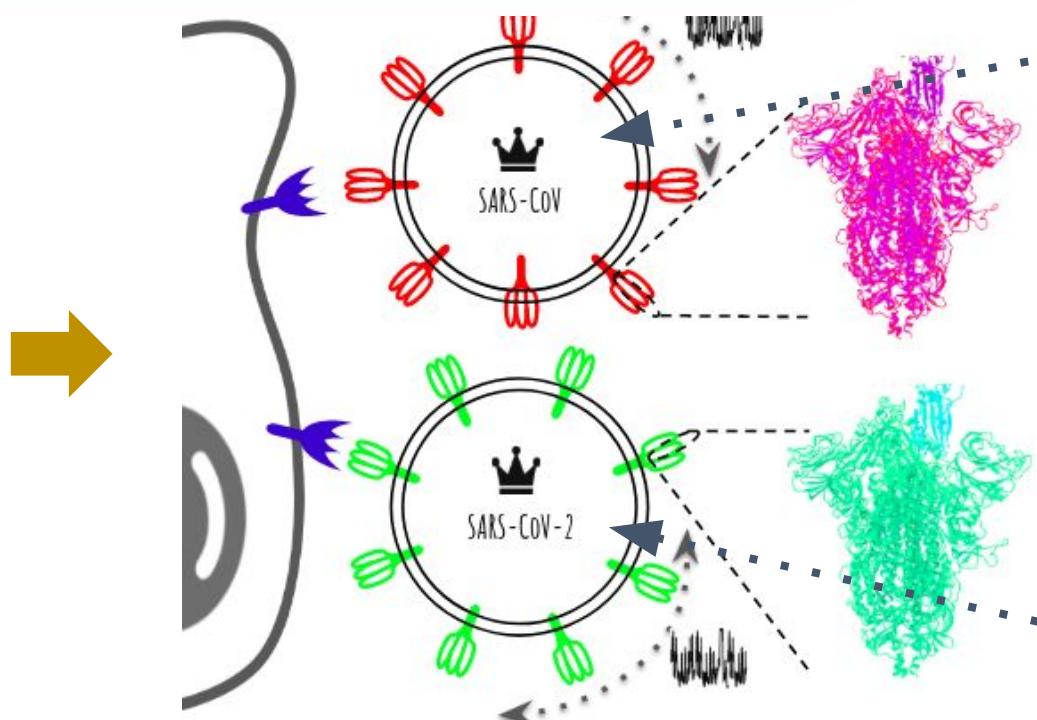
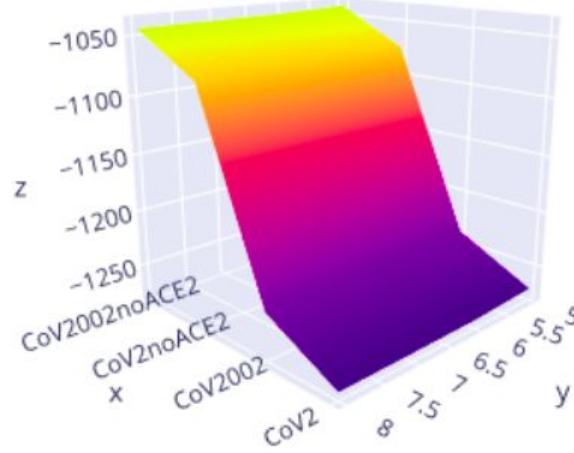
Connect to a cloud server with available GPUs (or local GPU)

```
In [ ]: import numpy,scipy,matplotlib  
!pip install pycuda,MDAnalysis,Nanoshaper,pdb2pqr  
!apt-get install swig -y  
!git clone https://github.com/barbagroup/pygbe.git
```

```
In [ ]: import pyF4all
```

Coronaviruses and comparing their thermodynamic cycles

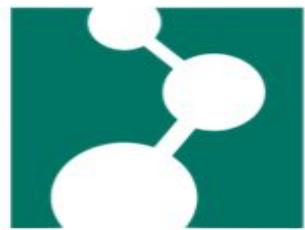
```
In [ ]: model=pyF4all.compare.pH("SARS-COV-2","SARS-COV-2wACE2","SARS-COV-2002","SARS-COV-2002wACE2",[5.0,6.0,7.0,8.0])  
model.plot.pH(structures,pH,Free_Energies)
```



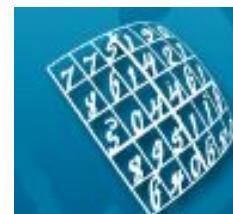
# Acknowledgments

- Dr. Christopher Cooper (UFSM)
- Prof. Kurt Kremer (MPIP)
- Prof. Ruben Perez (UAM)
- Prof. Rudolf Podgornik (CAS-CN)
- Dr. Simon Poblete (UACH)
- Dr. Matias Machado (IP-M)
- Dr. Matej Kanduc (IJS)

Max Planck Institute  
for Polymer Research  
 THEORY GROUP



 Jožef Stefan  
Institute



Slovenian  
Research  
Agency

# Acknowledgments

- Dr. Christopher Cooper (UFSM)
- Prof. Kurt Kremer (MPIP)
- Prof. Ruben Perez (UAM)
- Prof. Rudolf Podgornik (CAS-CN)
- Dr. Simon Poblete (UACH)
- Dr. Matias Machado (IP-M)
- Dr. Matej Kanduc (IJS)

Github: [pyF4all](#) | @govarguz

Twitter: @whoratz

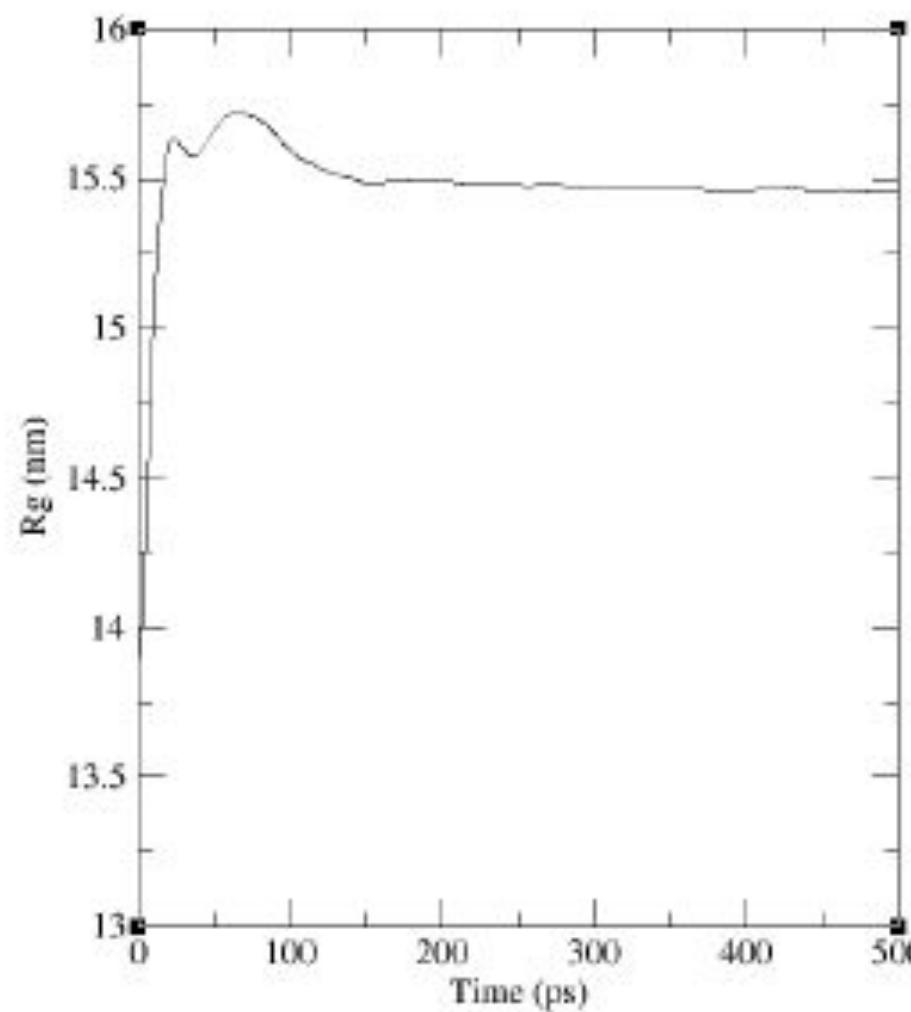
mail: [horacio.guzman@ijs.si](mailto:horacio.guzman@ijs.si)

[christopher.cooper@usm.cl](mailto:christopher.cooper@usm.cl)

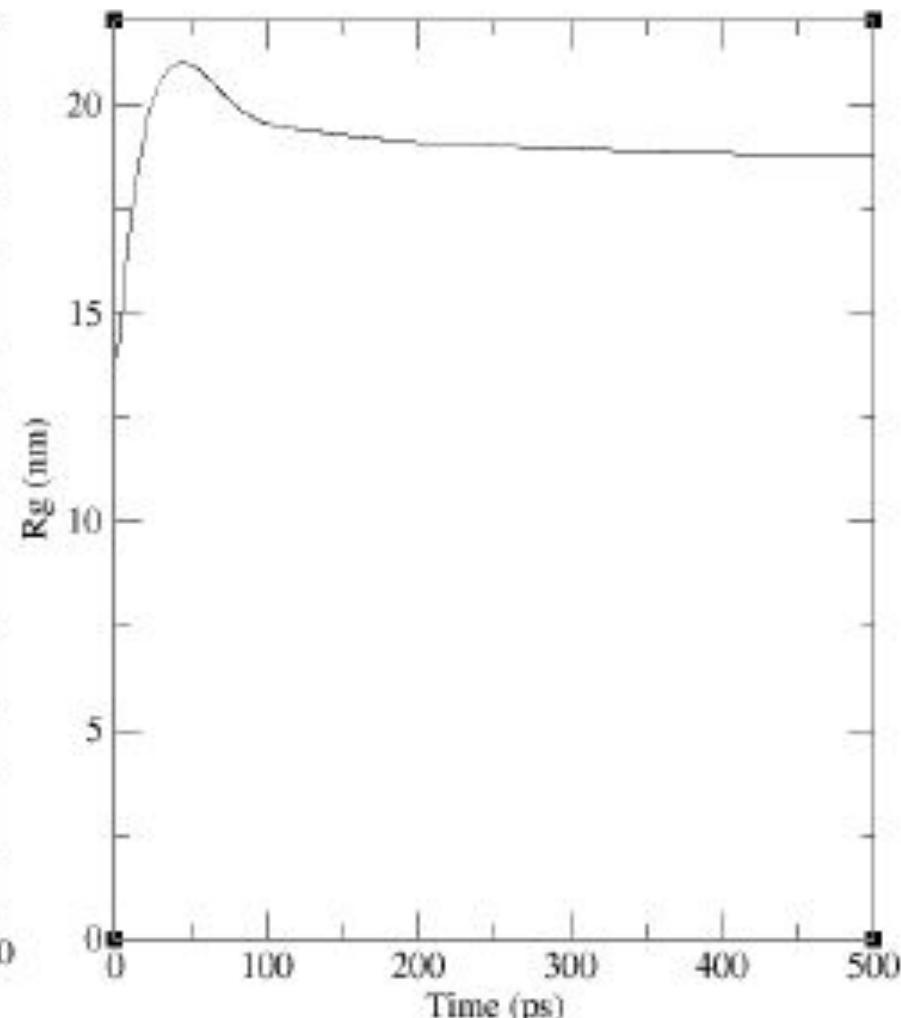
[https://github.com/pyF4all/Notebooks\\_phase1](https://github.com/pyF4all>Notebooks_phase1)



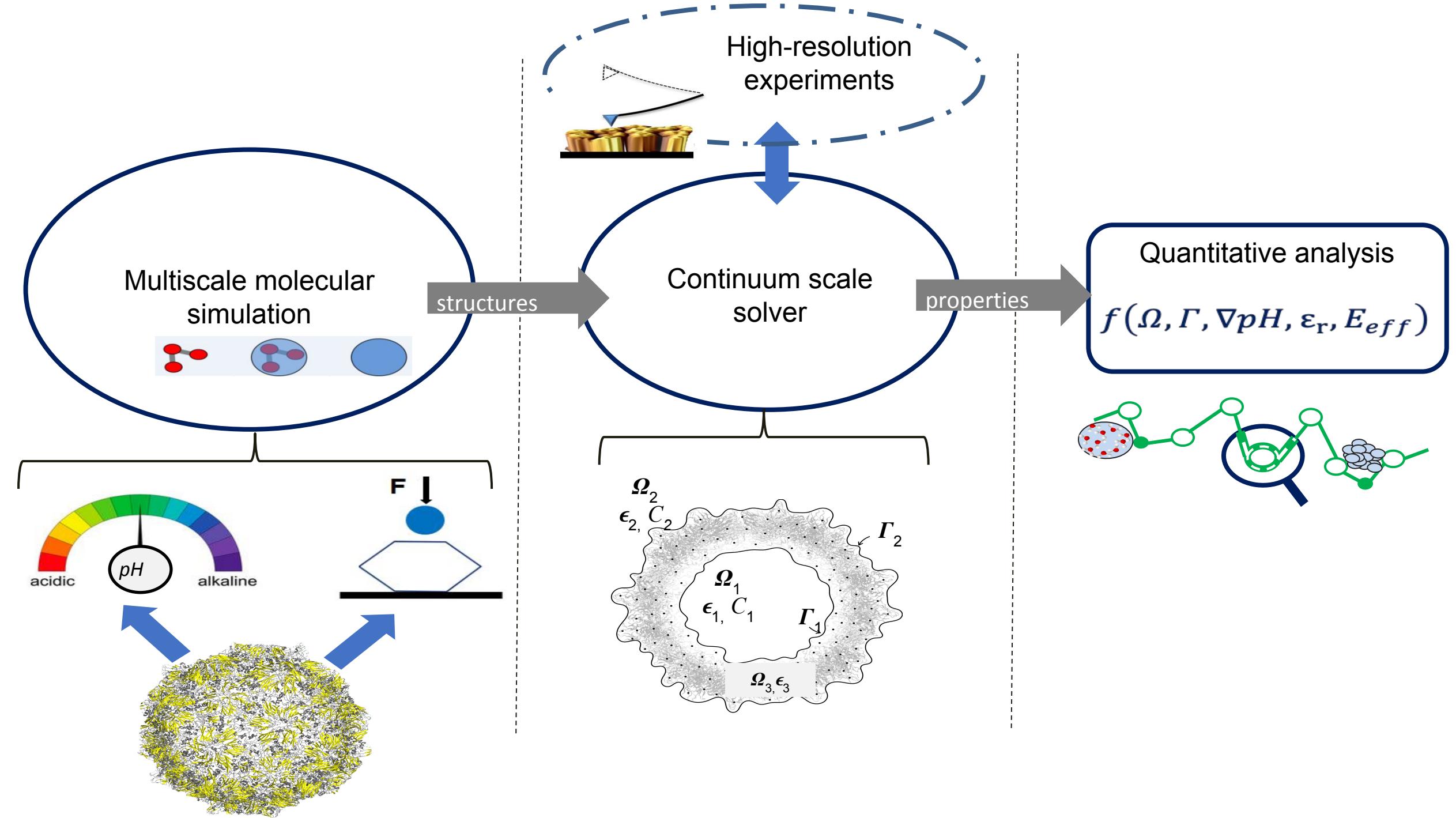
TrV with 800 Cl ions



TrV with 1600 Cl ions

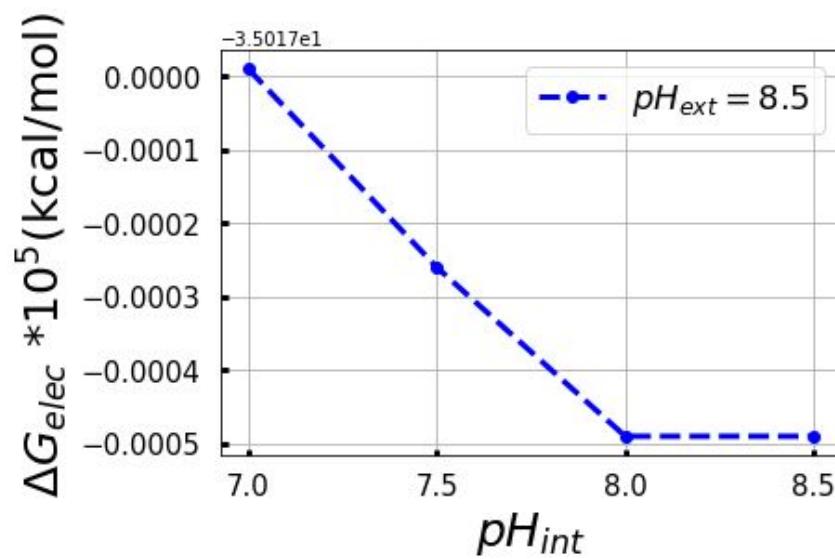
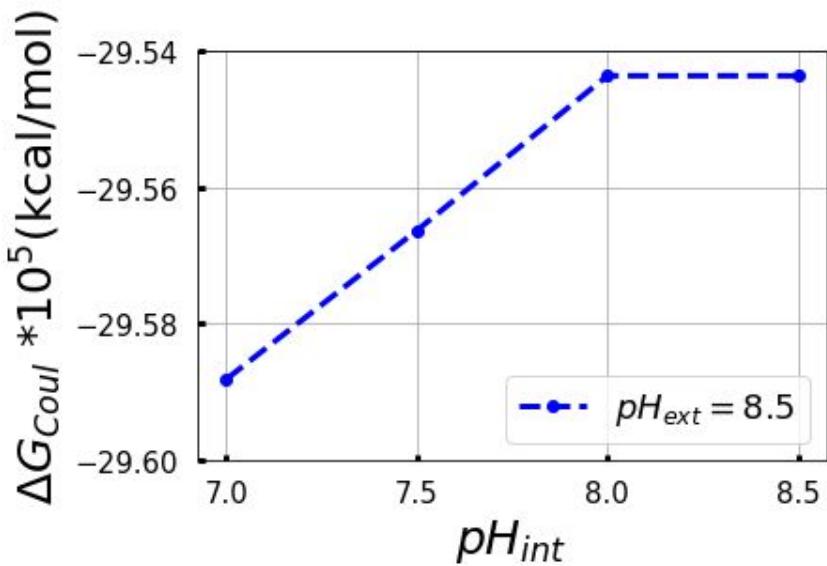
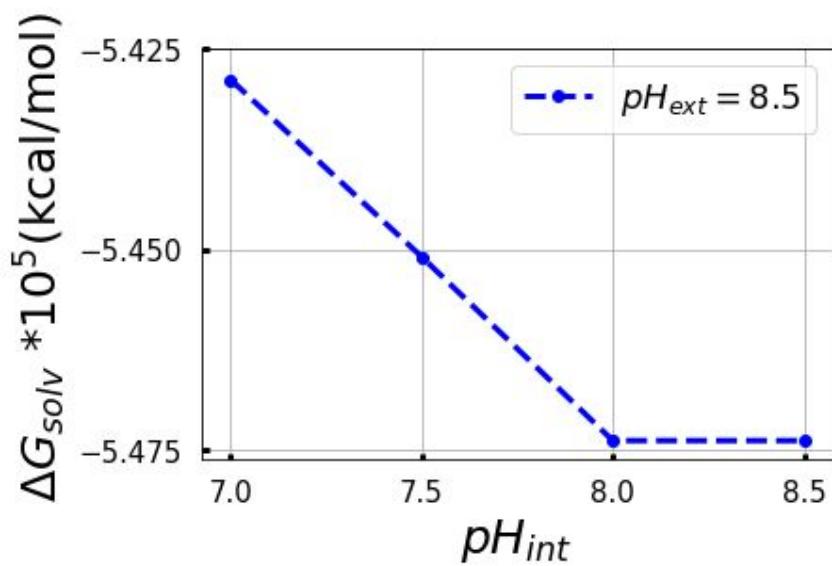


GIF credits: Prof. S. Poblete  
GIF credits: Prof. S. Poblete



# The method: Free energies of the disassembly of viral capsids (pH-mediated)

Fixed outside  
Alkaline inside

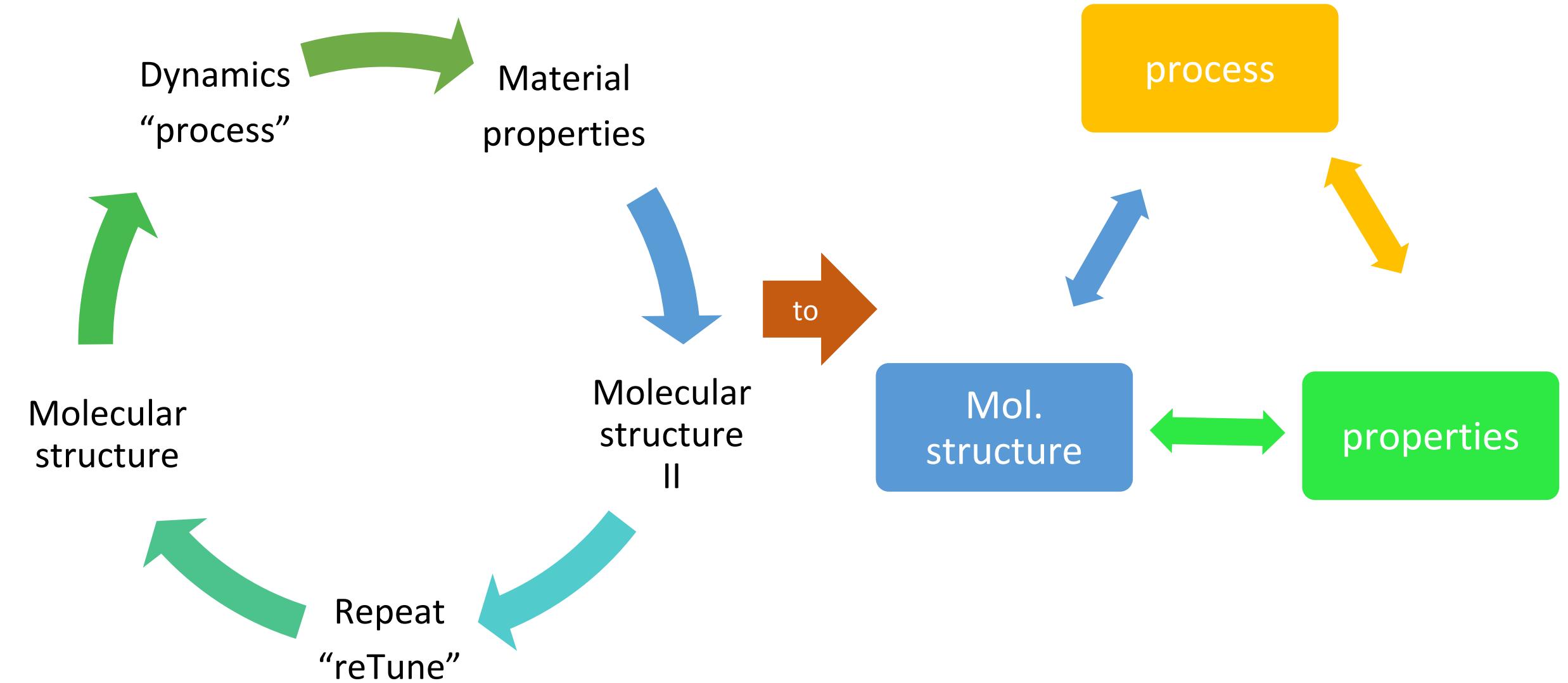


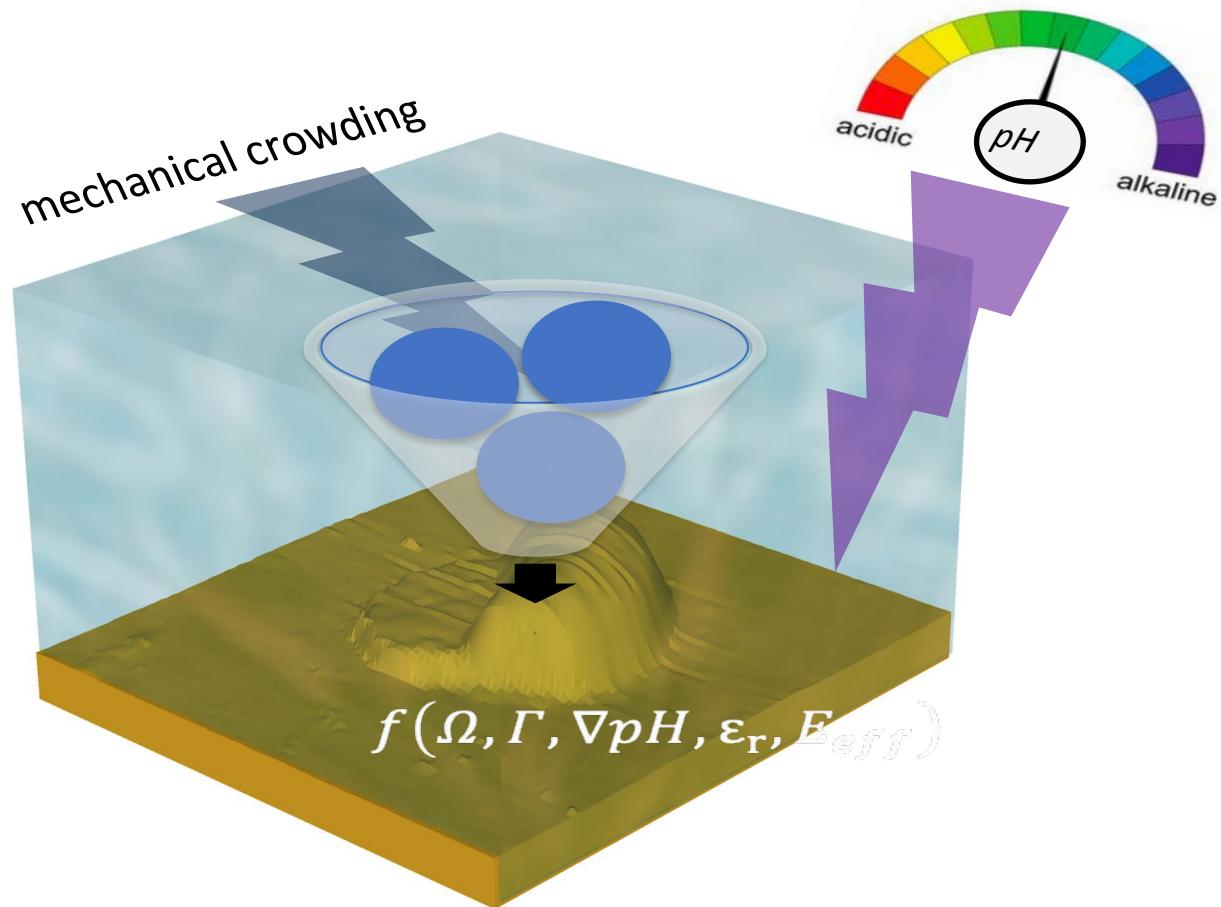


# Influence of the *pH* gradient on the free energy



# Motivation “loop”

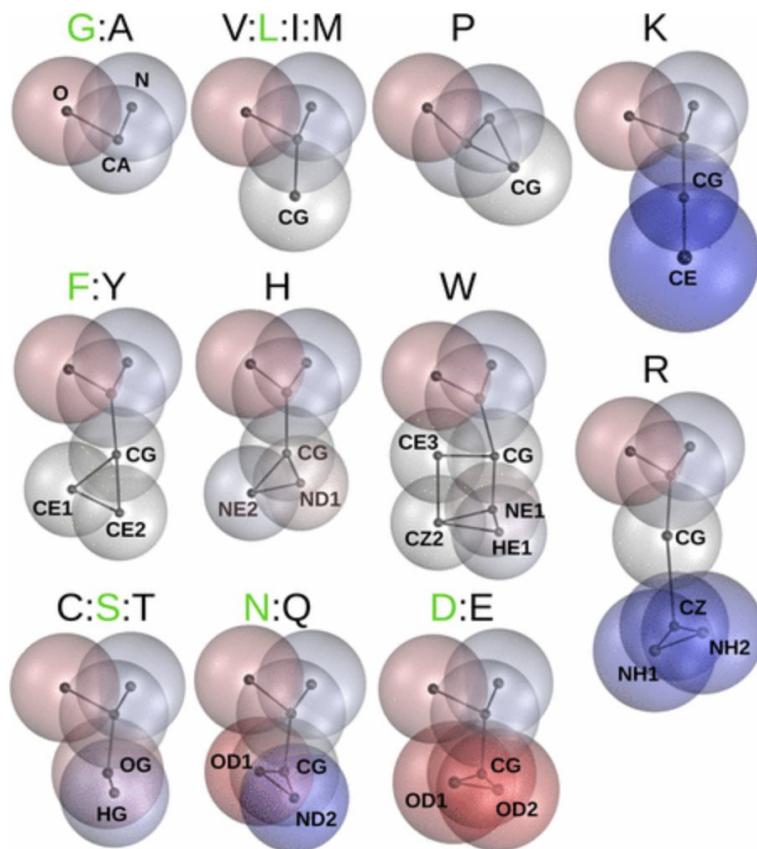
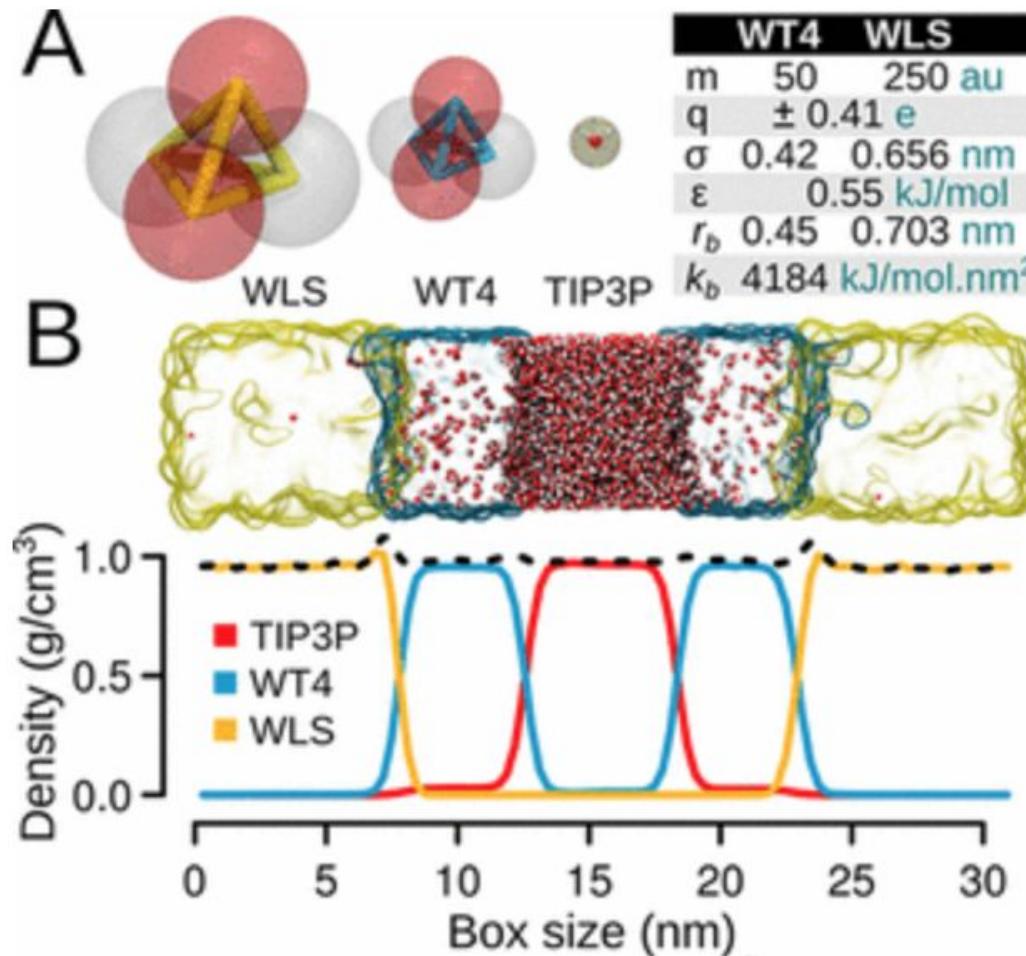




# non-concurrent top-down Multiscale

Sirah<sup>1</sup>

The CG philosophy used in here is top-down, which means that fit structural or some experimental qtty



[1] M. Machado et al., *J. Chem. Theory Comput.* **15**, 2719 (2019)

[2] L. Darre, *J. Chem. Theory Comput.*, **11**, 723 (2015)

