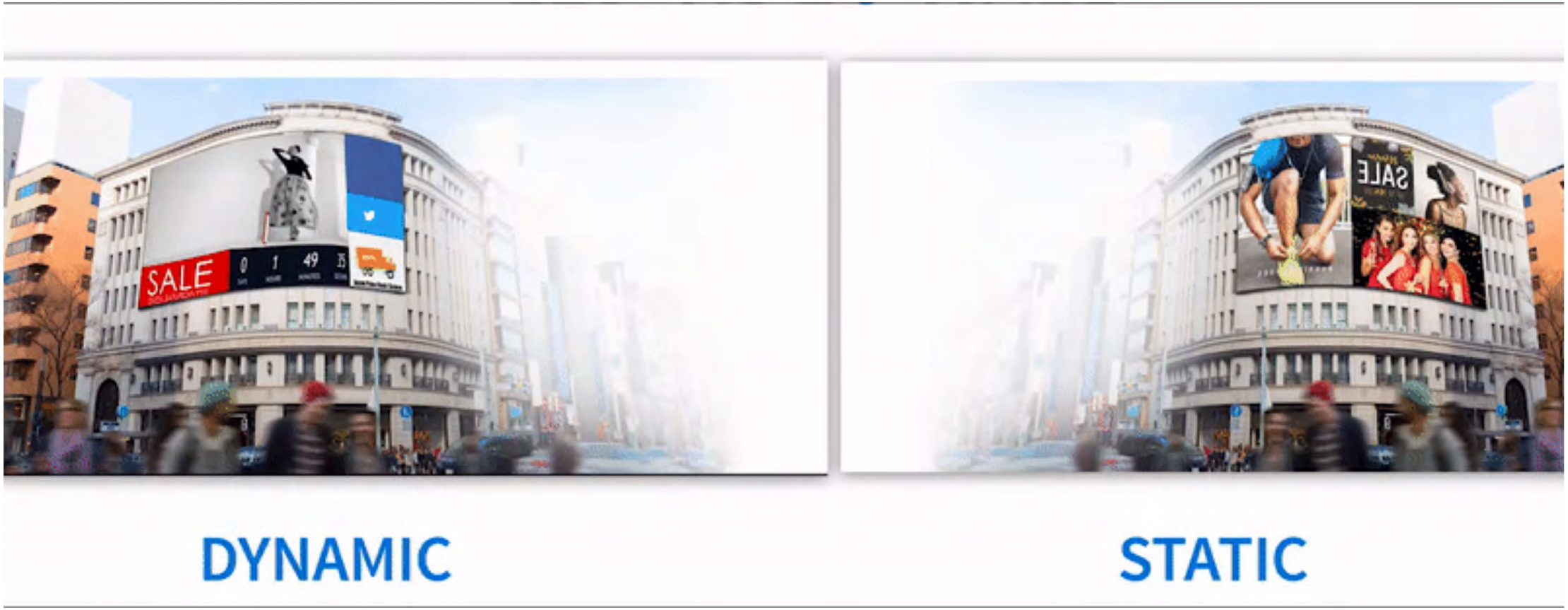
TMM 3102: Protein Structure, Function and Disease

- Structural Biology Methods: Molecular Dynamics Simulation (October 7th, 2021)

Jyh-Yeuan (Eric) Lee, Assistant Professor, BMI

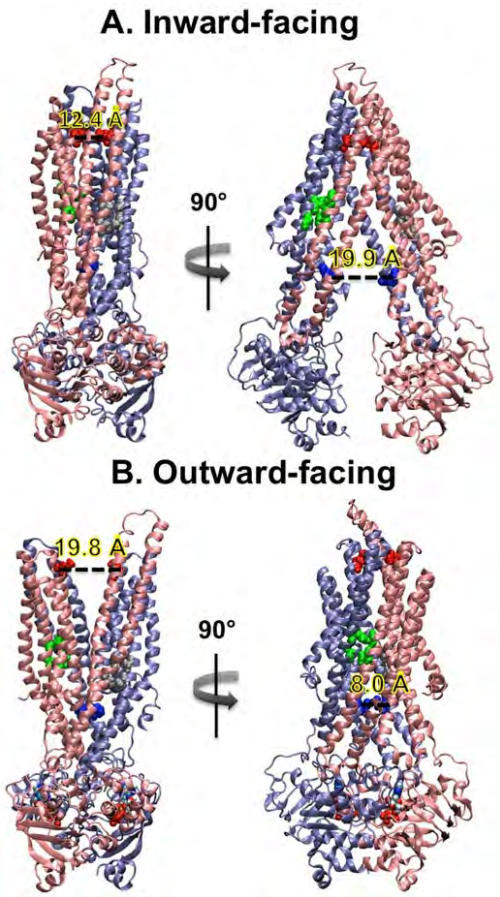
(Partially adopted from former lectures by Dr. Maria Musgaard)

Importance: Static v.s. Dynamic

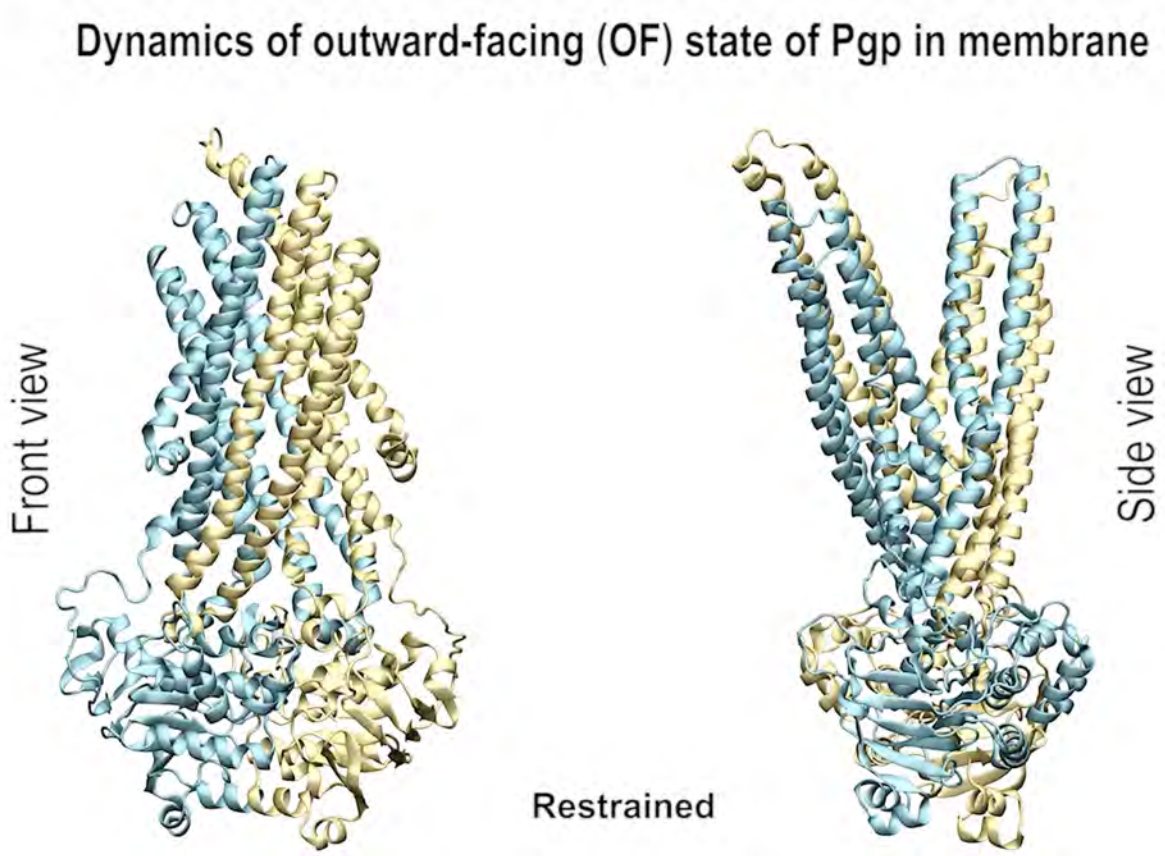


(https://digitos.io/benefits-of-dynamic-digital-signage-over-static-signage/)

Importance: Static v.s. Dynamic



(Pan & Aller, Sci Rep, 2015)



(Verhalen et al, Nature, 2017)

Bridging the Gap

Protein function

- Functional data
- Electrophysiology
- Substrate transport
- …
- High resolution in time

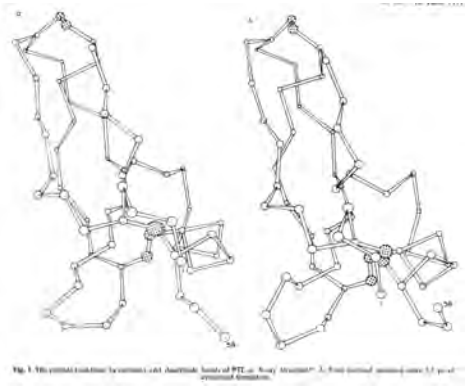
Protein structure

- X-ray
- NMR
- Cryo-EM
- ~ “snapshots”
- High resolution in space

High resolution in “space” and “time”?

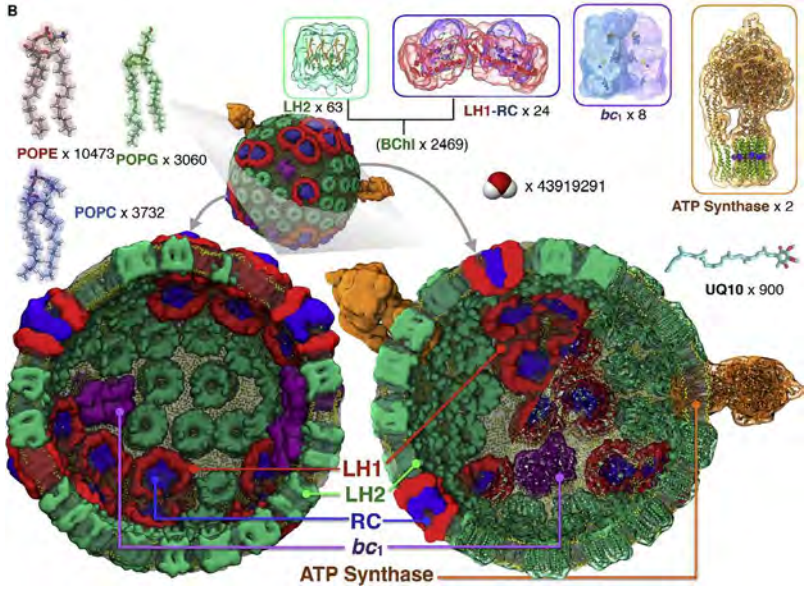
Brief History

- First MD study of proteins published in 1977
  ~60 residues, no solvent, ~9 ps



(Nature, 1977)

- 2019: full organelles, 139 million atoms, 0.5 μs



(Cell, 2019)

- Factors:
  - more structures determined
  - better algorithms
  - faster computer

Molecular Dynamics (MD): idea

- Classic mechanics (thinking of “Newton’s laws of motion”)

- Metaphor:

  If cycling at 15 km/h by Canal Rideau; keep a constant acceleration:

  - Predict how long to reach uOttawa main campus.
  - Predict where you are in 5 minutes.

- Do the same for all atoms in a protein system

Molecular Dynamics (MD): idea

Going to the next position:

- $r(t+\Delta t) = r(t) + \Delta t * v(t) + 1/2[\Delta t^2 * a(t)]$

  r(t): position at “t”
  r(t+Δt): position after Δt
  v(t): velocity
  a(t): acceleration

Molecular Dynamics (MD): idea

Acceleration:

$F = m * a$

$F = -\Delta U/\Delta r$

If we know U (potential energy), then we can calculate the force and the acceleration on each atom.

Molecular Dynamics (MD): workflow

In general, how do we do MD simulation?

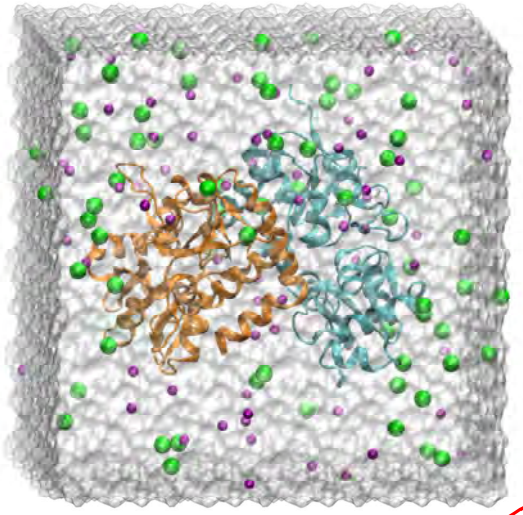
a. Find the coordinates of a known protein model from the database.

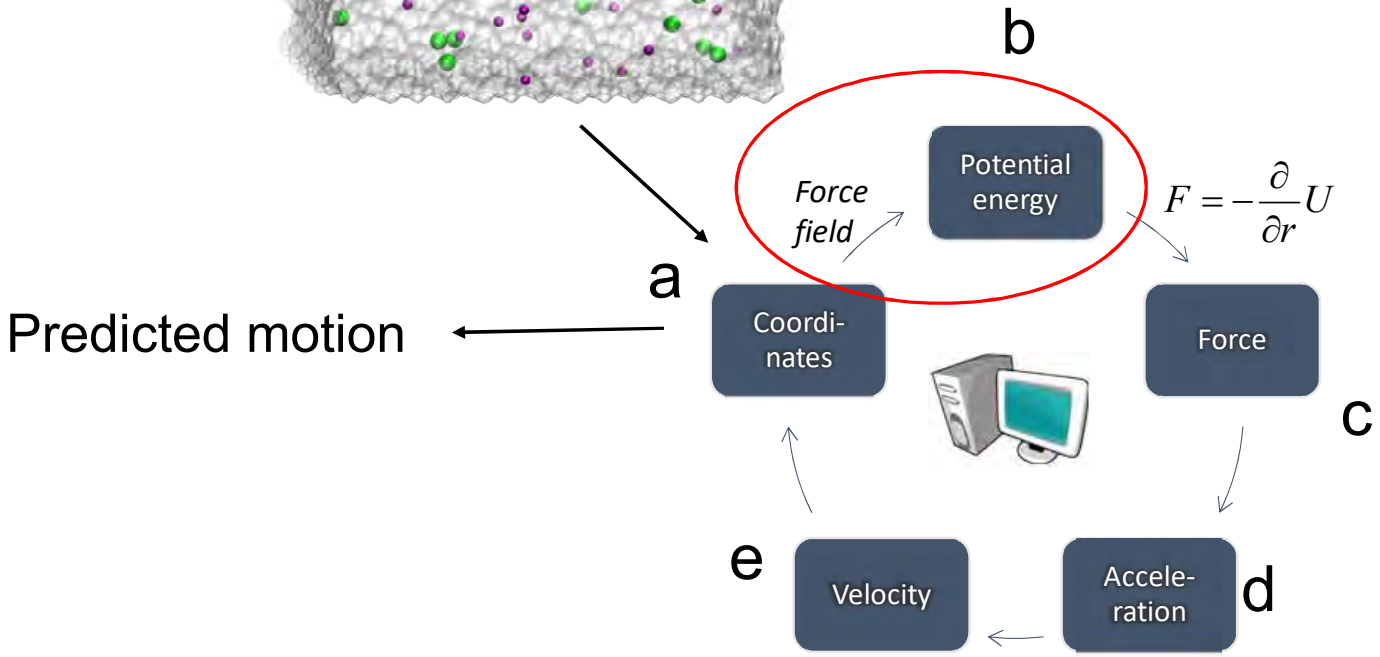
b. Choose a force field to generate energy potential for further calculation.

c. Calculate the force that results from the theoretical potential energy.

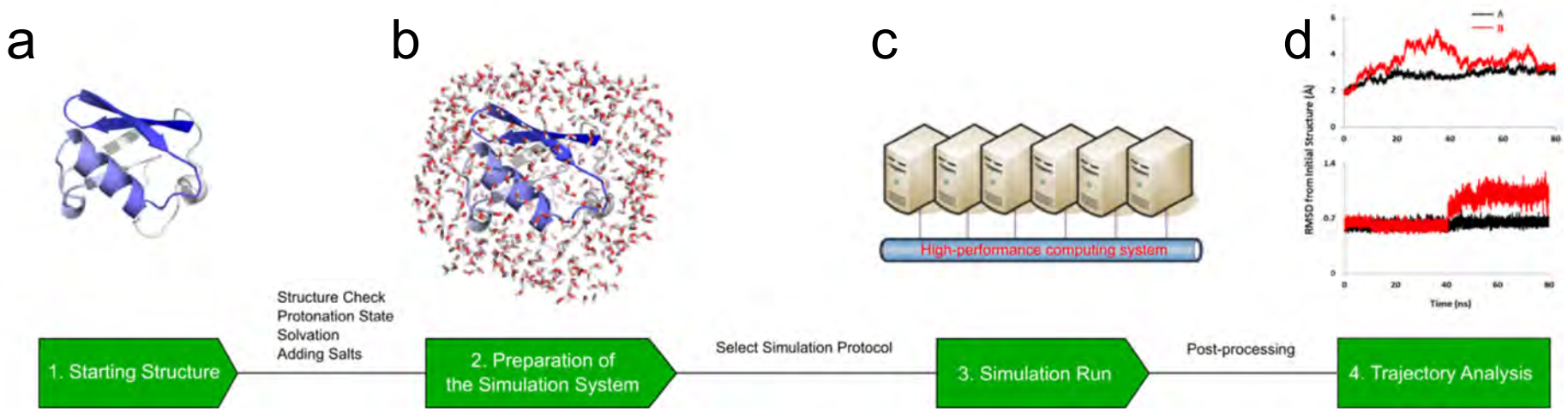
d. Find out how molecules speed up with the obtained force.

e. Calculate the speed of the molecule and where the protein move into.





Molecular Dynamics (MD): workflow



Another way to see the MD workflow:

a. Find a model template and artificially add necessary ingredients that suit the physiological condition of the target protein. This includes protonation states, salts, water, etc.

b. Prepare the simulation system by selecting the best protocol, aka force field.

c. Run the simulation using a cluster of computers.

d. Process the data and predict the where the segment of interest moves to.

Molecular Dynamics (MD): force field

What determines “force field”?

- Atoms: different in size, softness, mass, charge, …
- Bonds: different in lengths, stiffness, …
- Electrons: implicitly accounted for covalent bonds.

Molecular Dynamics (MD): force field

What is a force field used for?

- Used for large molecules or conformational studies
- Not used to break or form chemical bonds
- Empirical, so no one is most correct.
- Requires:
  - Energy equation to describe U as a function of atomic coordinates
  - Constant parameters to be used in the energy equation
  - Atom types to establish constant parameters, charges, masses, etc.

Molecular Dynamics (MD): force field

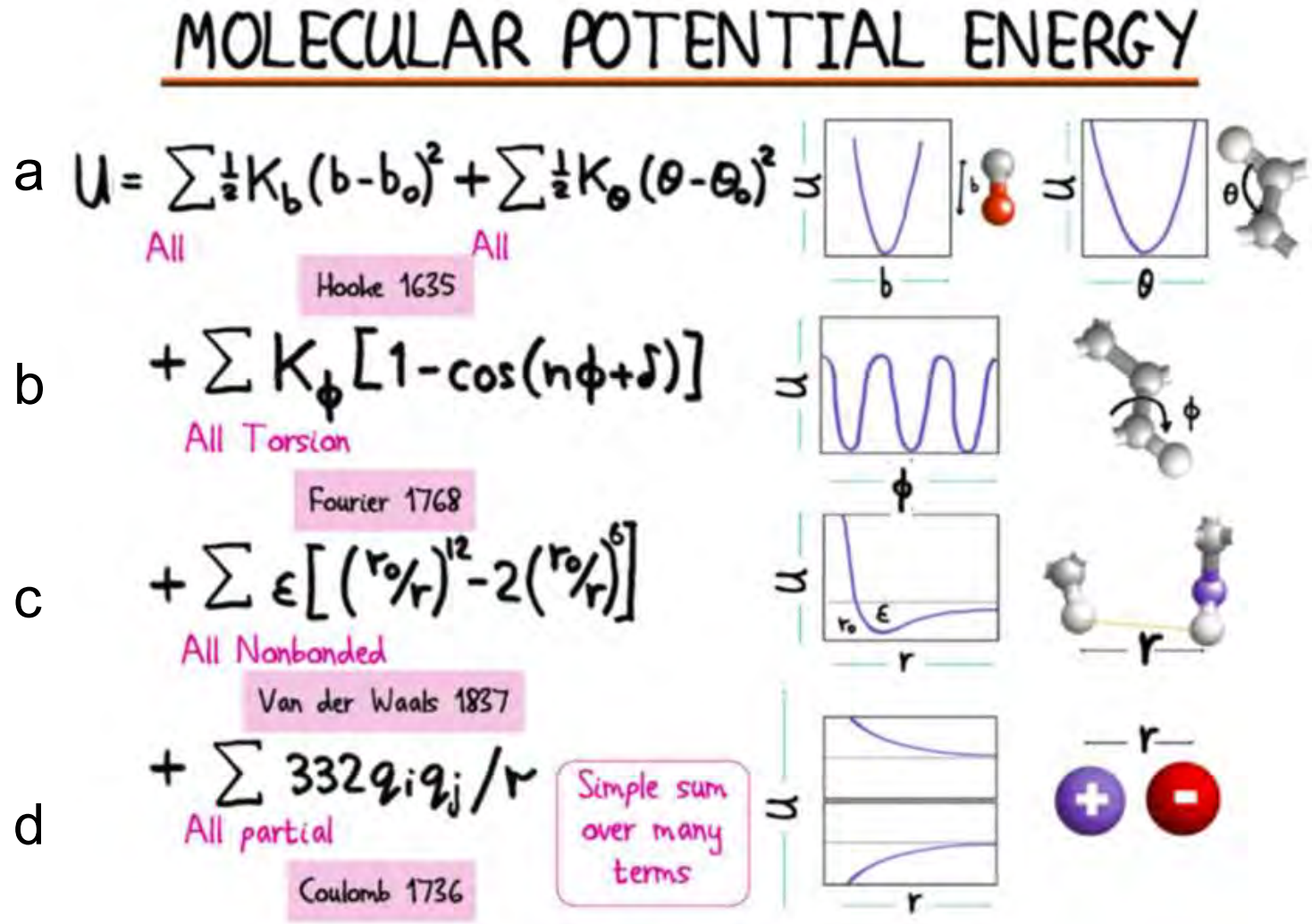
Selection of force field is like deciding what kind of potential energy to use:

a. Covalent bonds & bond angles

b. Torsion angles

c. Van der Waals interaction

d. Electrostatic force / charge-charge interaction



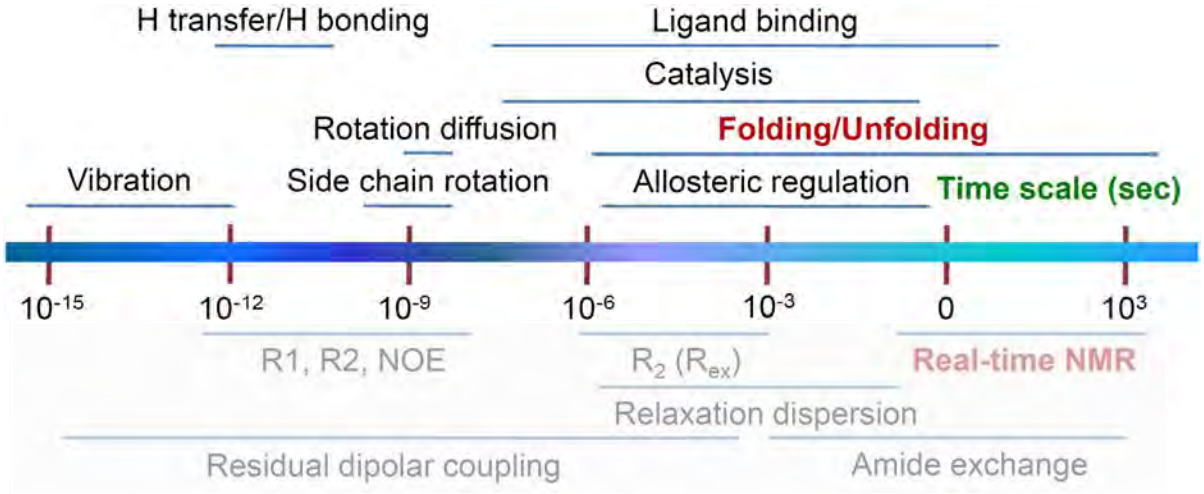
Michael Levitt, Nobel Lecture 2013

Molecular Dynamics (MD): force field

- Do’s and Don’ts
  - Never compare energies from different force fields, unless absolute energy is known
  - Never mix parameters, unless tested
  - Do simulations in the conditions similar to those used to obtain the force field
  - For new ligands, need a full set of parameters (all you can)

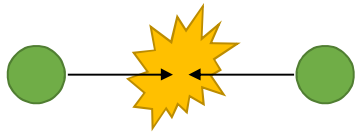
Molecular Dynamics (MD): time scale

Biological timescales

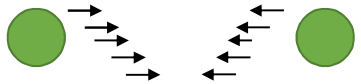


Kumar and Balbach, *Biochim. Biophys. Acta* **2015**

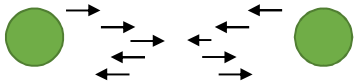
- Simulation Δt: 1-2 fs
  - Too fast:



  - Too slow:

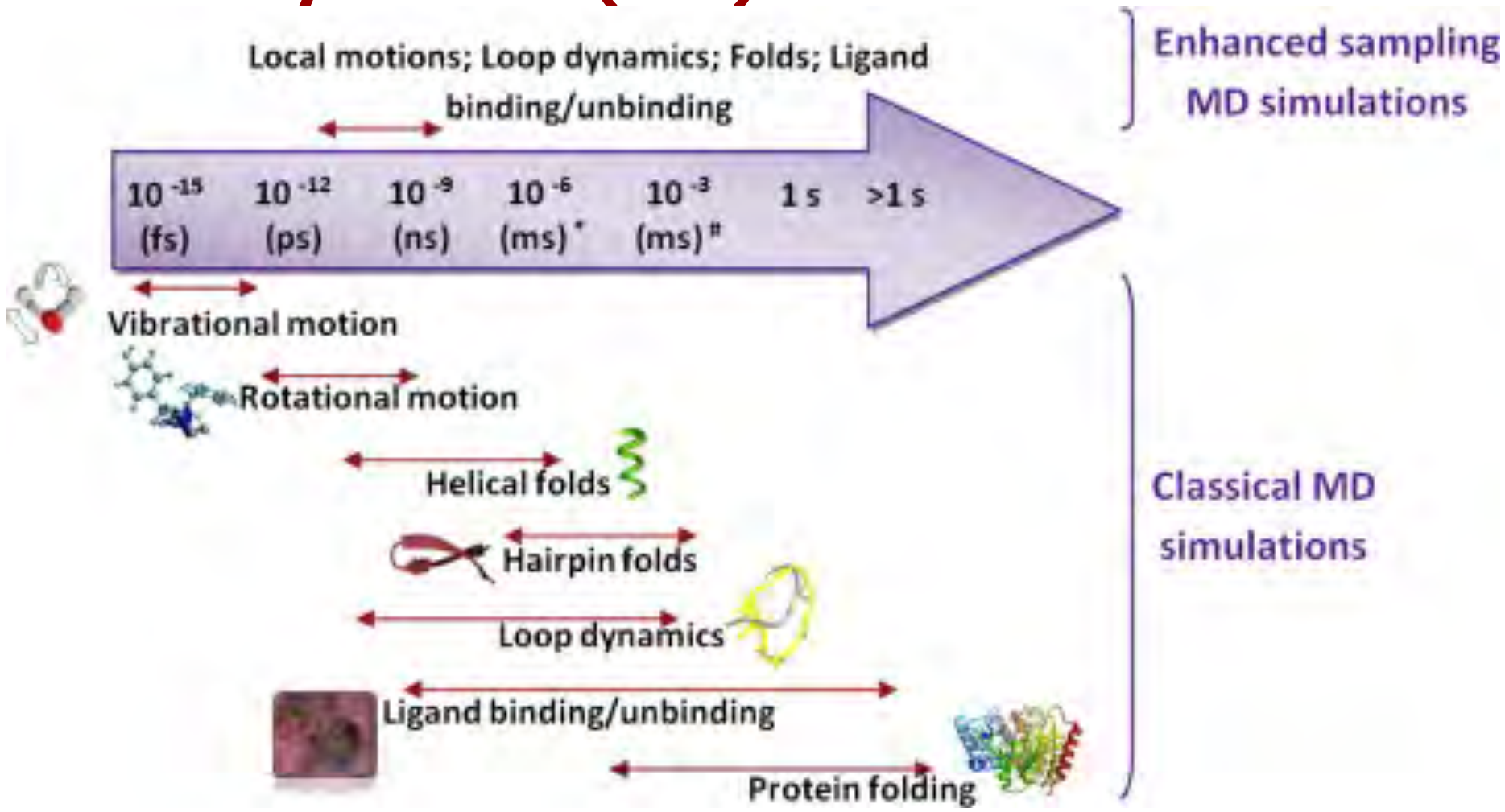


  - Good

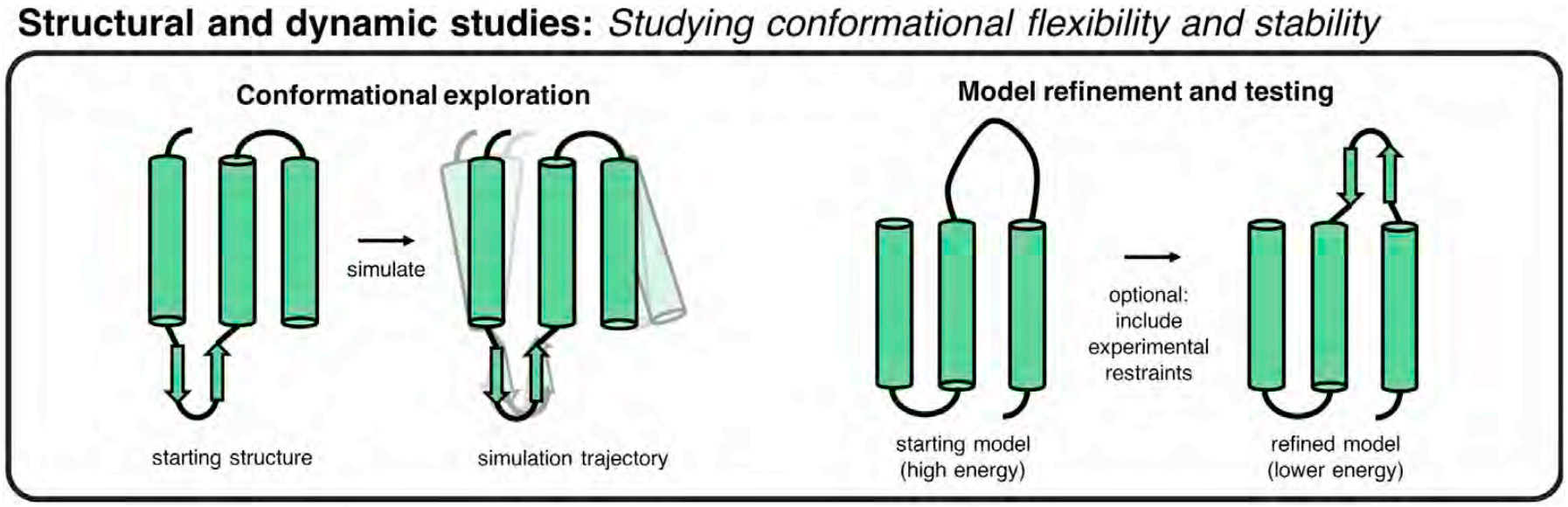


- => 0.5 to 1 million steps to reach 1 ns (!)

Molecular Dynamics (MD)



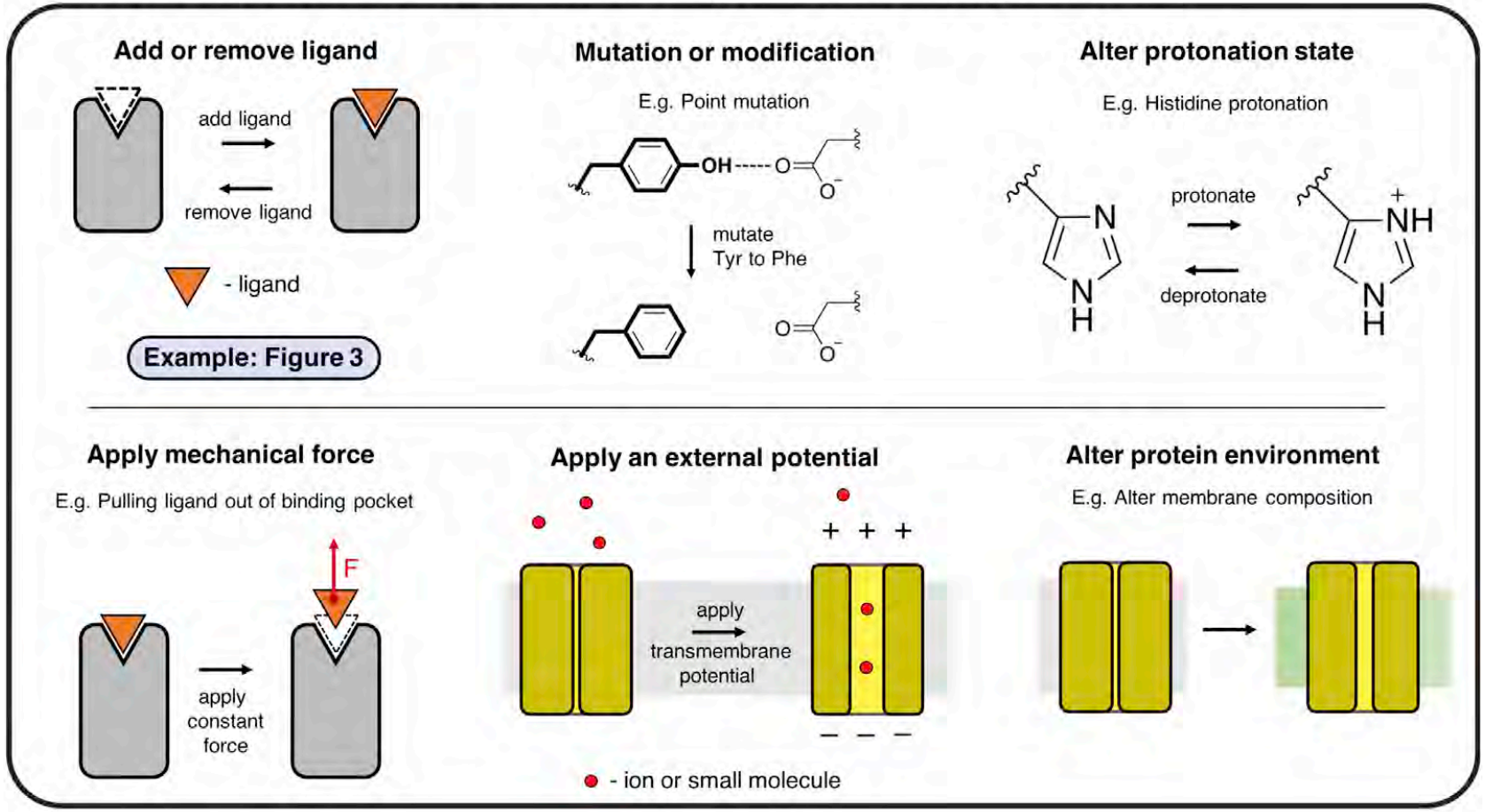
Structural Determination *in silico*



(Hollingsworth & Dror, Neuron, 2018)

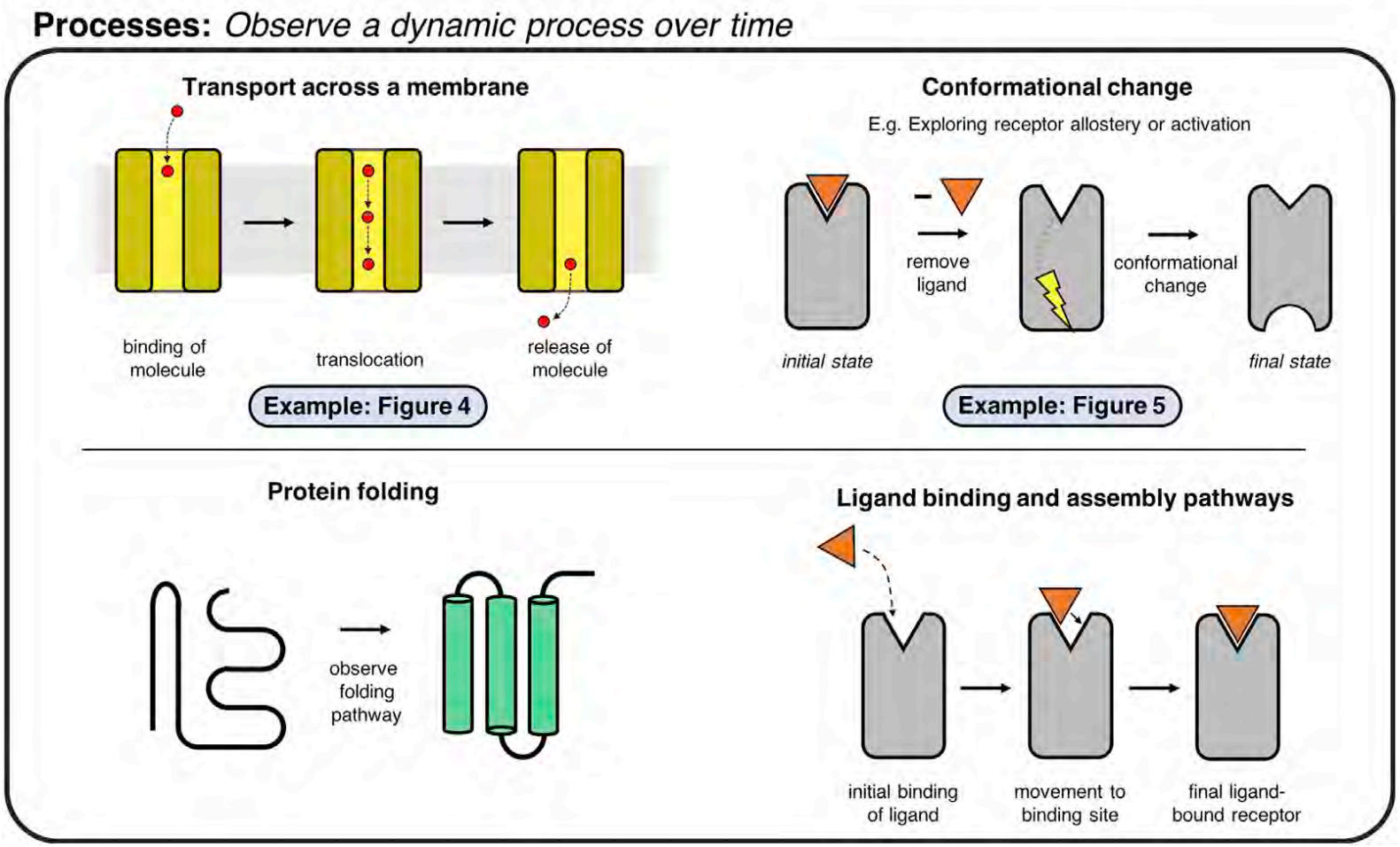
Structural Determination *in silico*

Perturbations: *Observe response following controlled change to system*



(Hollingsworth & Dror, Neuron, 2018)

Structural Determination *in silico*



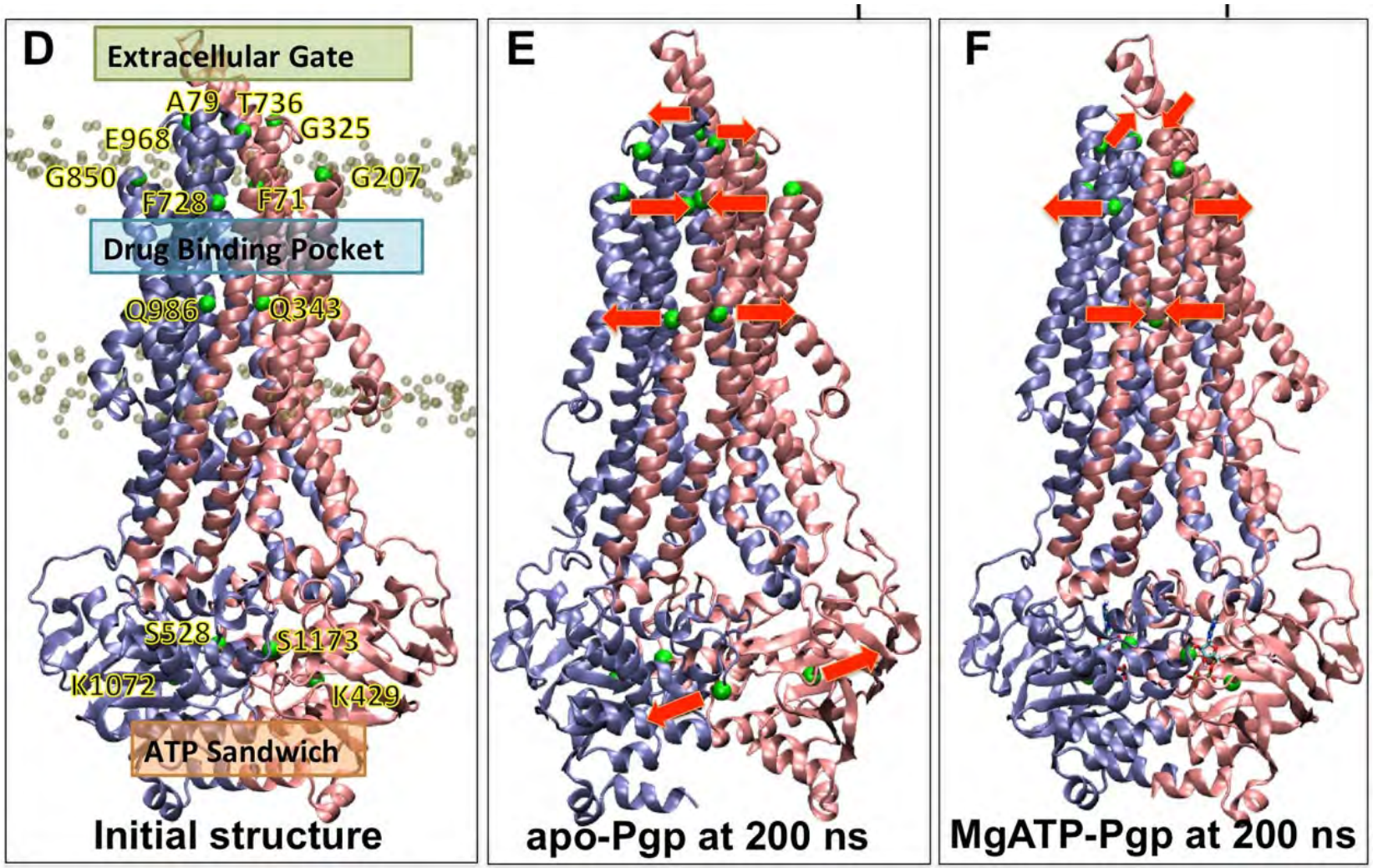
(Hollingsworth & Dror, Neuron, 2018)

Molecular Dynamics (MD)

- Advantages
  - High resolution in space and time
  - Precise simulation conditions: conformations, ± ligands, …
  - Cheap: mutations, protein-ligand, protein design, …
  - Structure-function relationship

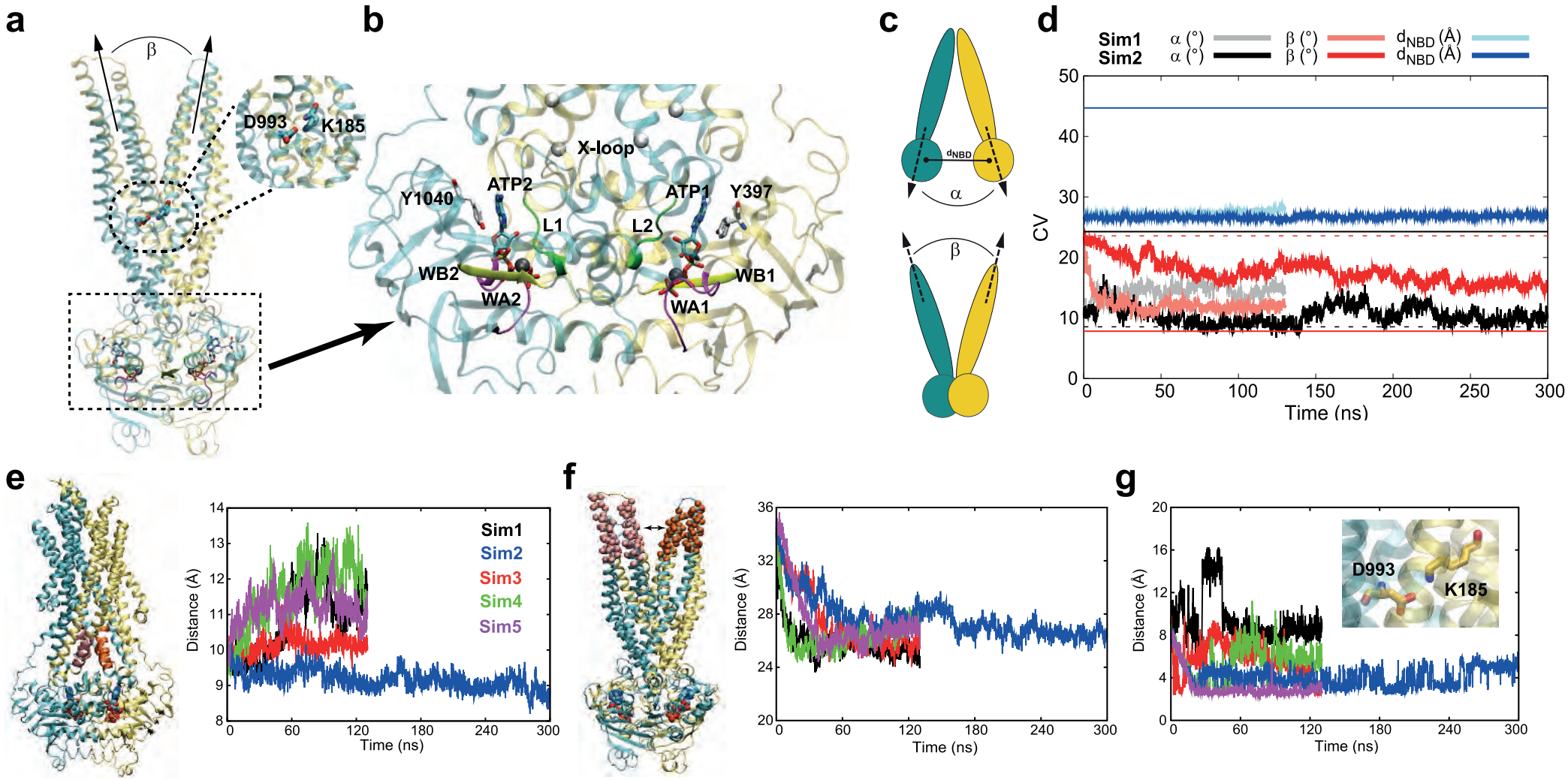
- Limitations
  - Validation: need experimental data
  - Timescale and sampling
  - Quality of starting structures
  - Force fields
  - No bond making/breaking, as it depends on protonation states

Case Study: P-glycoprotein (drug-resisance)



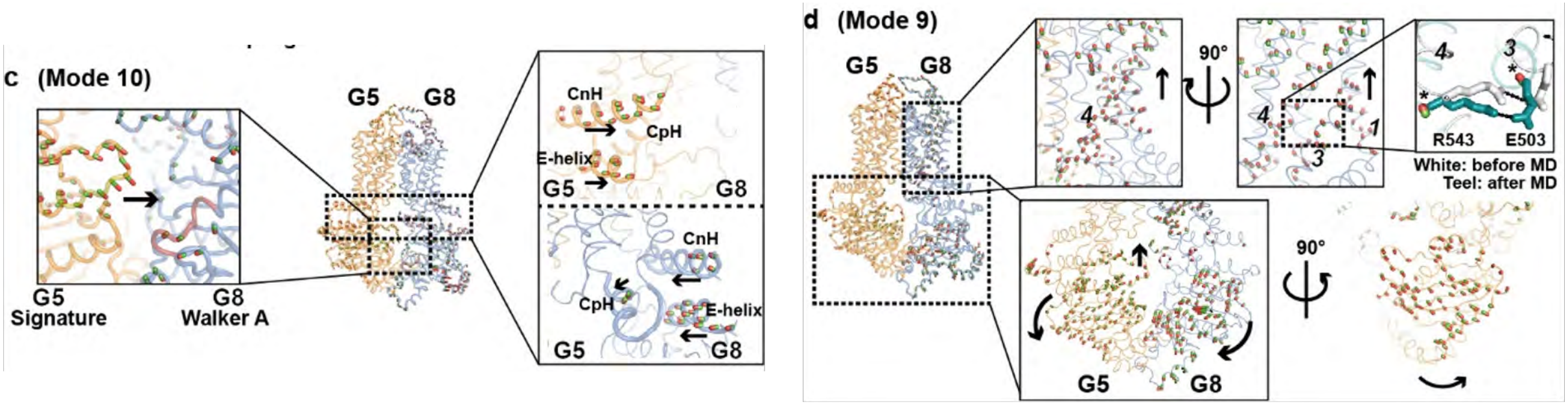
(Pan & Aller, Sci Rep, 2015)

Case Study: P-glycoprotein (drug resistance)



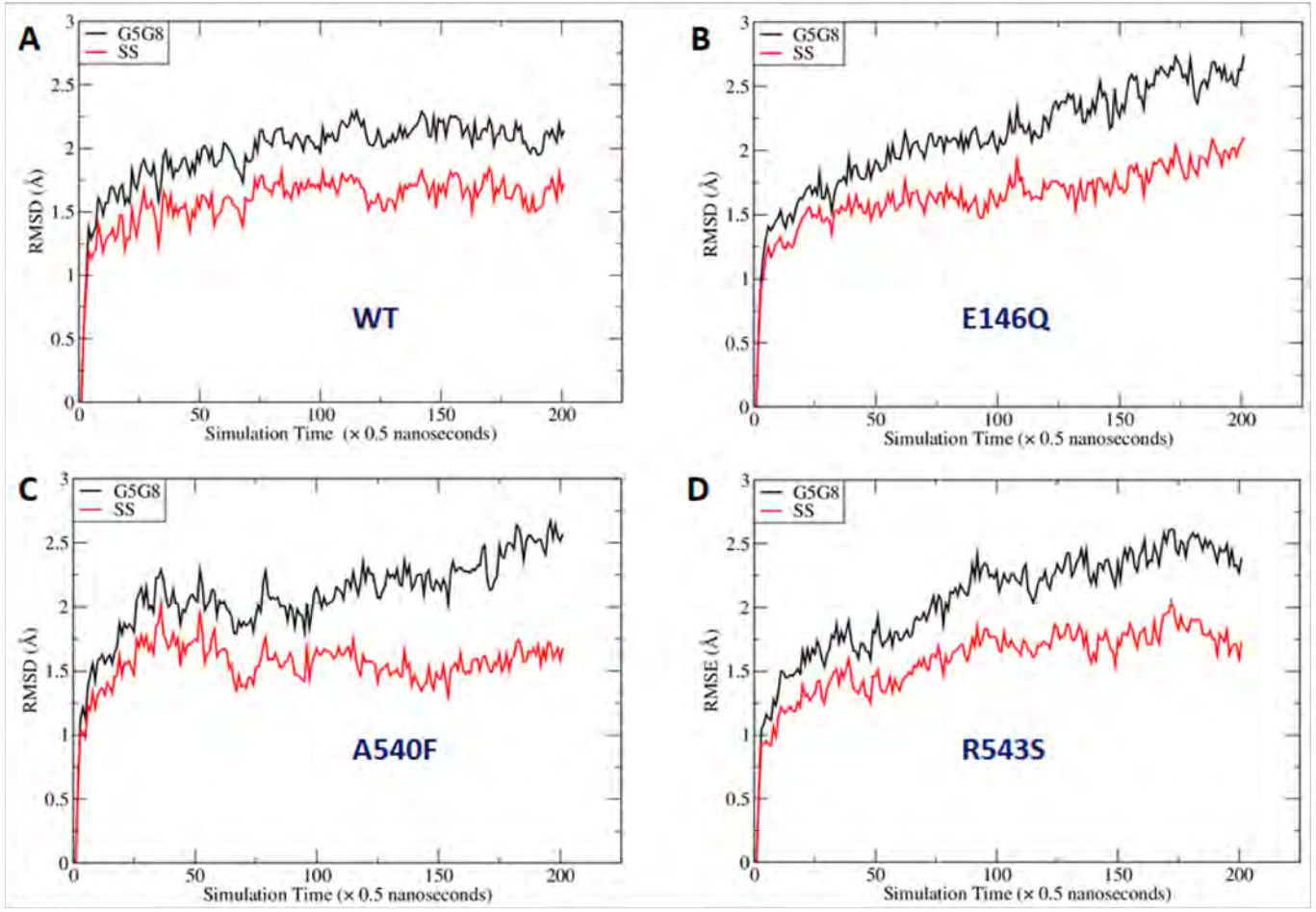
(Verhalen et al, Nature, 2017)

Case Study: ABCG5/G8 (sterol efflux)



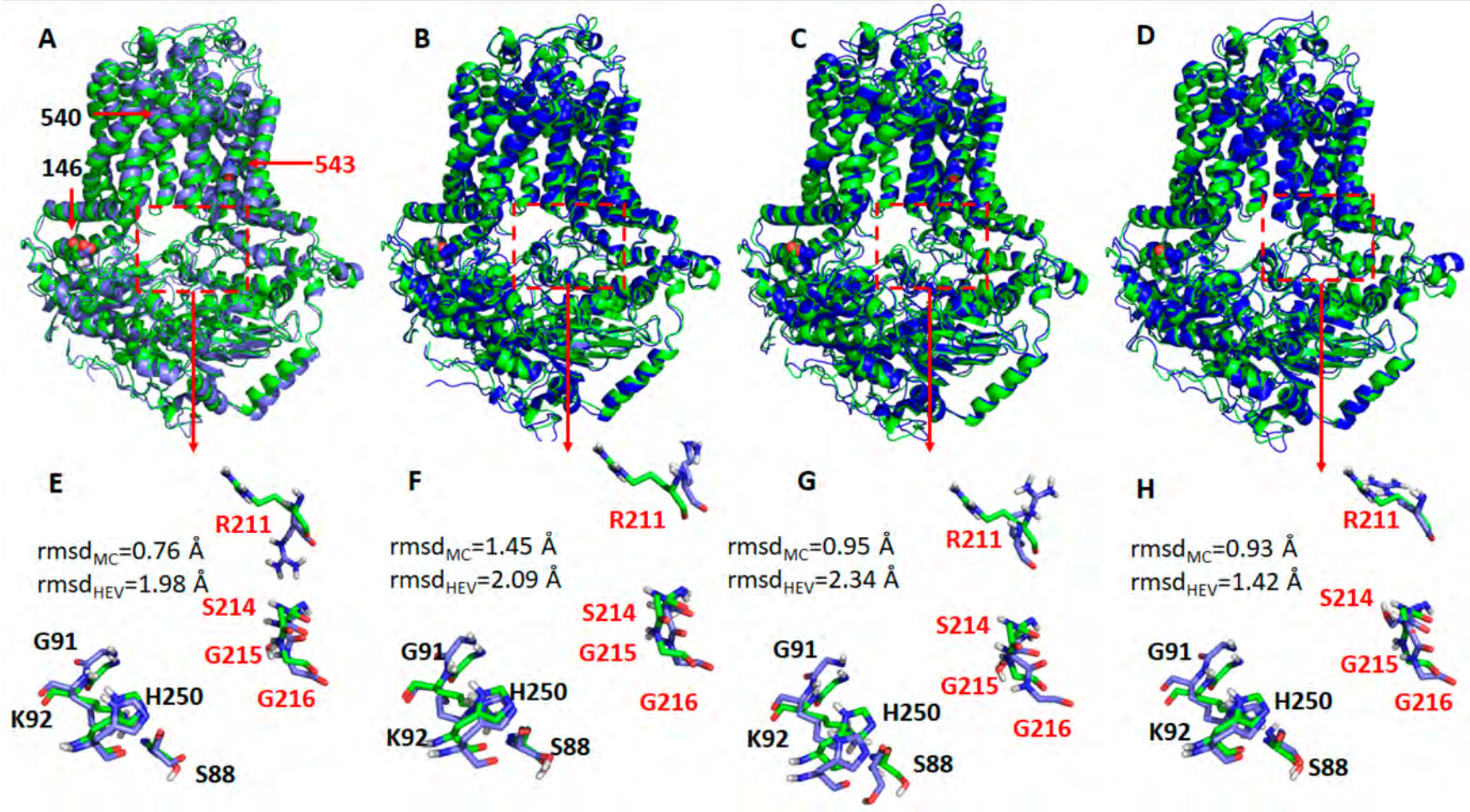
(Lee et al, Nature, 2016)

Case Study: ABCG5/G8 (sterol efflux)



(Xavier et al, IJMS, 2020)

Case Study: ABCG5/G8 (sterol efflux)



(Xavier et al, IJMS, 2020)