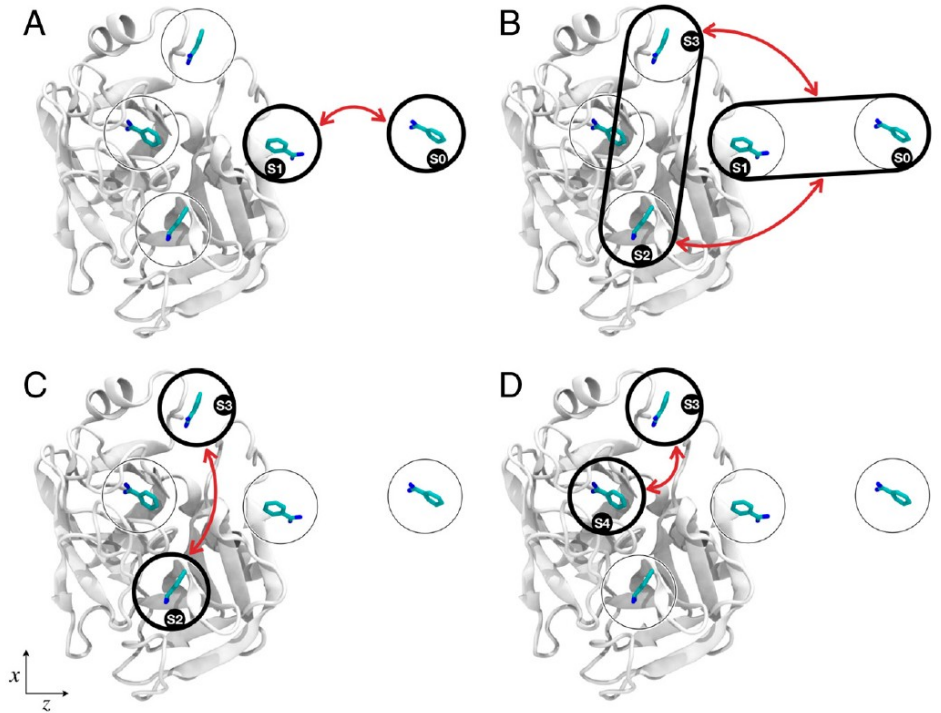
MD simulations of ligand-protein complexes

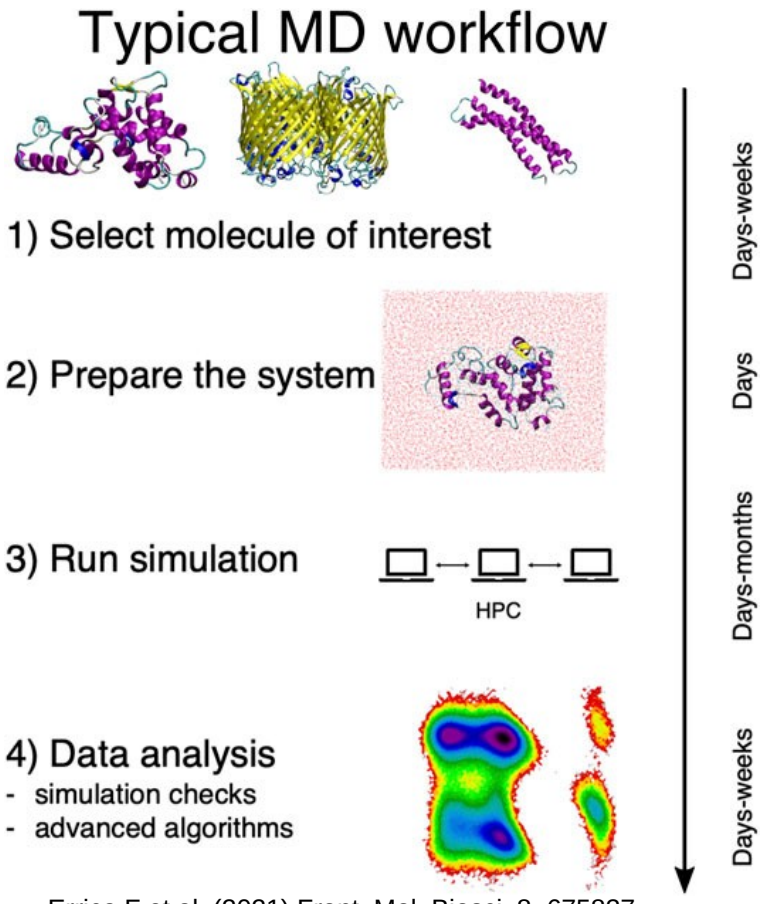


Buch I, Giorgino T, De Fabritiis G. (2011) PNAS 108, 10184–10189.

Dror RO et al. (2012) Annu. Rev. Biophys. 41, 429–452.

Why structure preparation matters?

- Molecular dynamics simulations rely on physically realistic starting structures
- Inaccuracies in protonation, missing atoms, or incorrect topologies lead to unreliable dynamics
- Structure preparation ensures chemical accuracy and compatibility with the force field

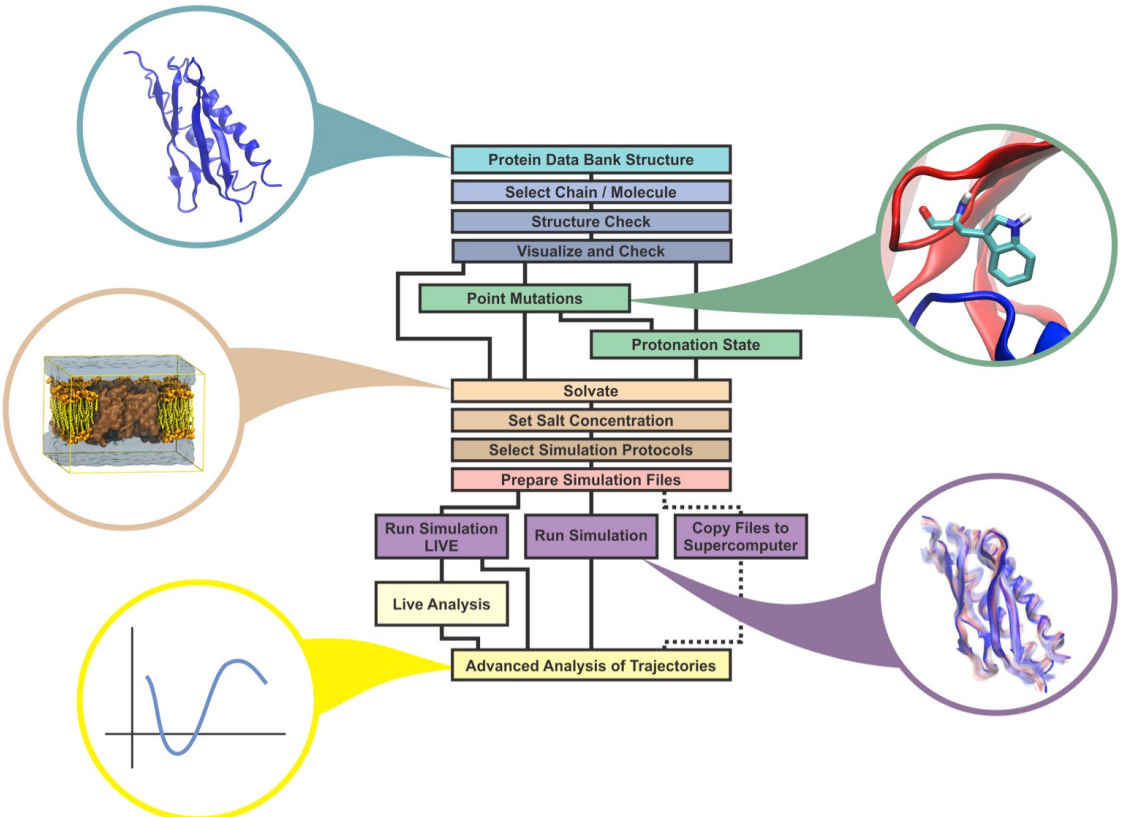


Errica F et al. (2021) Front. Mol. Biosci. 8, 675837

Karplus M et al. (2002) Nat. Struct. Mol. Biol. 9, 646–652; Ribeiro JV et al. (2016) Sci. Rep. 6, 26536

Overview of the workflow

- Key stages: inspection → protonation → hydrogen addition → minimization → minimization → solvation → parameterization
- Each step has theoretical significante, e.g., pKa affects net charge; box shape influence pressure equilibration
- Consistent system preparation improves reproducibility and interpretability

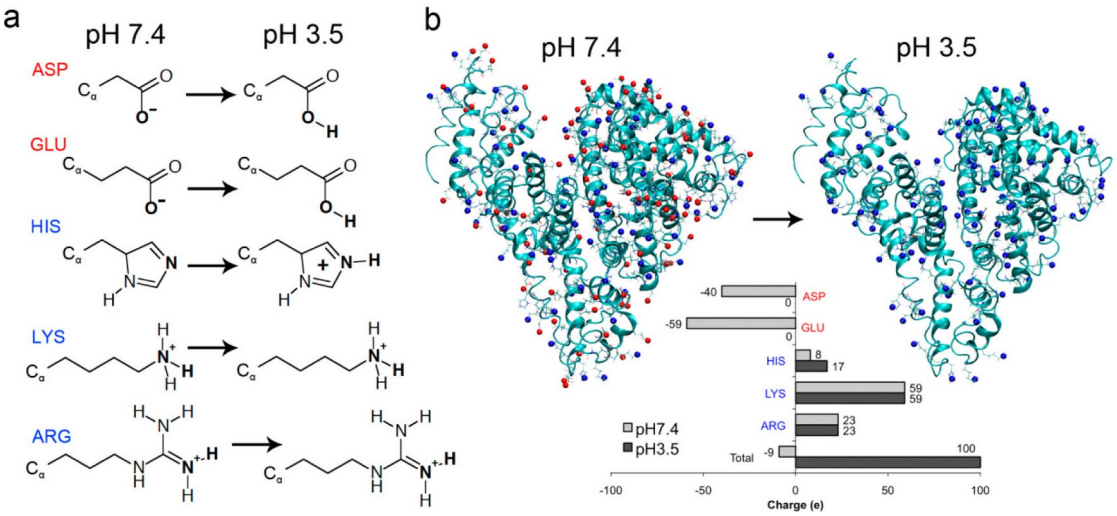


Hollingsworth SA et al. (2018) Neuron 99, 1129–1143

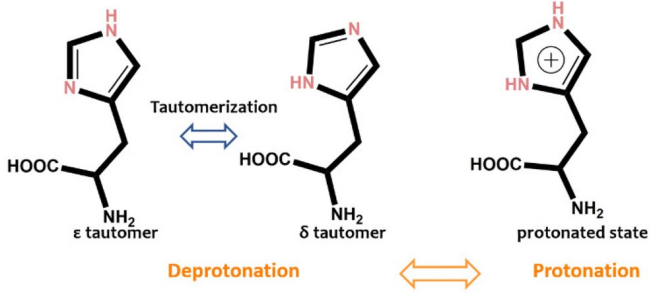
Scientific Reports volume 6, Article number: 26536 (2016)

Protonation states: theory

- Ionizable residues (Asp, Glu, Lys, Arg, His) have pKA values that determine their protonation state at a given pH
- Histidine has three forms: HID (delta-protonated), HIE (epsilon-protonated), HIP (doubly protonated)
- The local microenvironment can shif pKa by >2 units (e.g., buried Glu in hydrophobic core may be protonated)



Olsson MH et al. (2011) J. Chem. Theory Comput. 7, 525–537

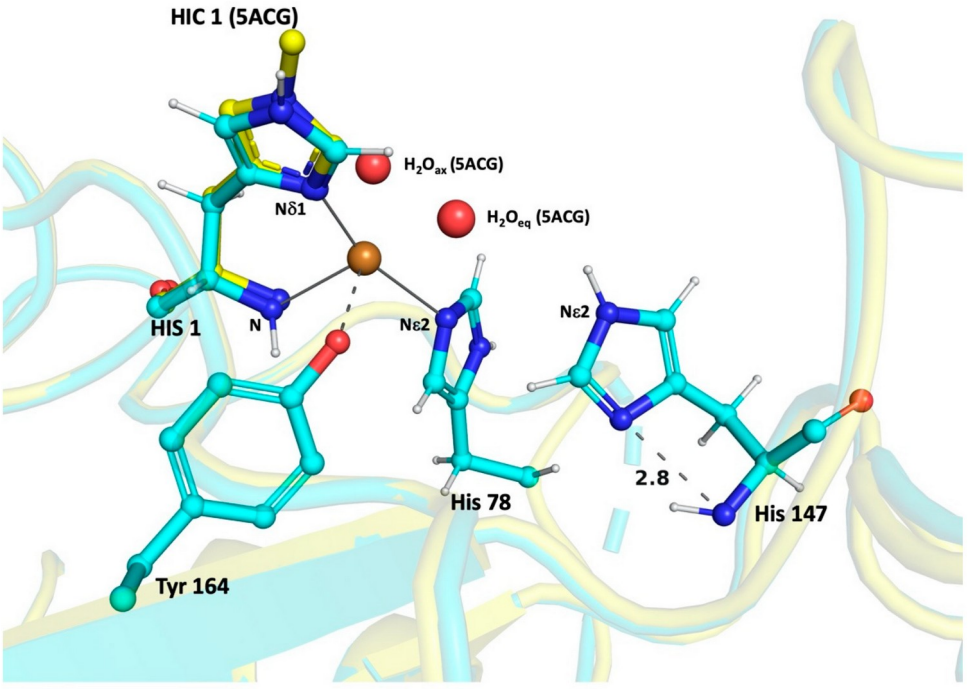


Sun Y et al. (2023) Phys. Chem. Chem. Phys. 25, 18346–18353

Olsson MH et al. (2011) J. Chem. Theory Comput. 7, 525–537

Protonation states: tools and conventions

- PROPKA and H++ use structure-based methods to estimate pKa values
- Reduce can optimize hydrogen bonding by rotating side chains (e.g., Asn, Gln, His)
- Software-specific conventions: GROMACS require explicit protonation state (via residue names), AMBER uses pdb4amber
- Manual inspection is essential for residues near the active site or metal ions

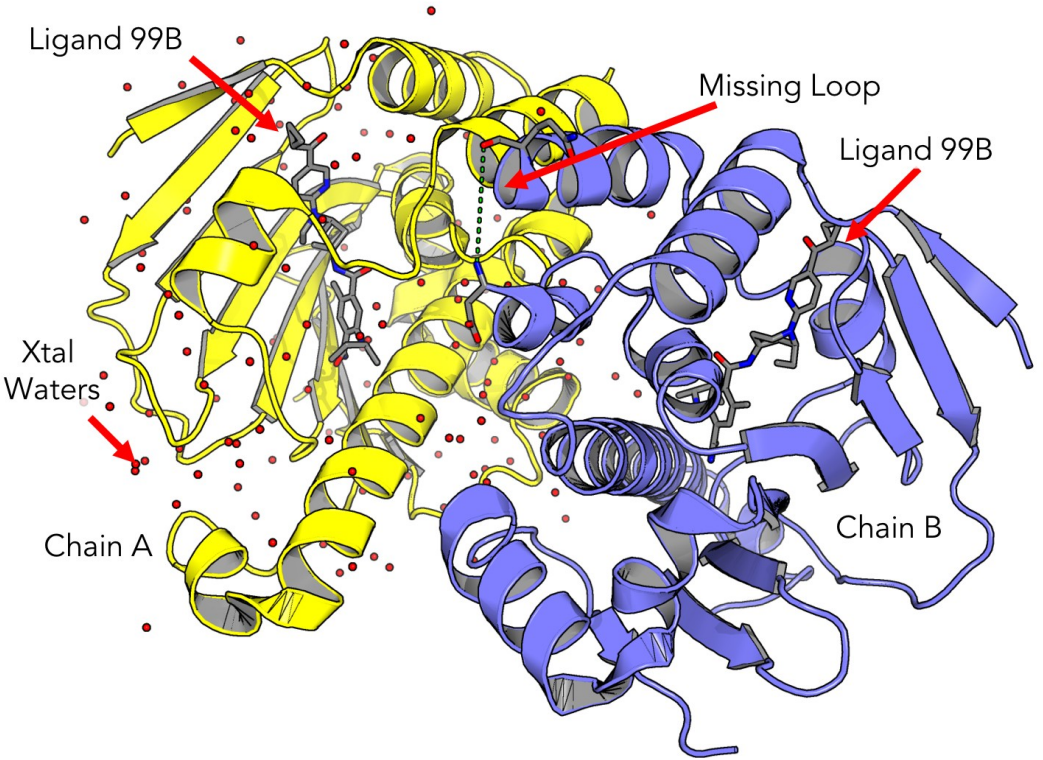


Banerjee S et al. (2022) Biomolecules 12(2), 194.

Anandakrishnan R et al. (2012) Nucleic Acids Res. 40, W537–W541

Protein preparation: structural completeness

- 3D structures often have missing side chains or loops due to flexibility
- Use modeling tools like MODELLER, PyMOL, VMD, ChimeraX, or pdbfixer to rebuild missing atoms
- Disulfide bonds must be explicitly assigned; alternate locations need to be resolved (e.g., B-factors, occupancies)
- Chain continuity and terminal patches (e.g., ACE/NME) should match the selected force field

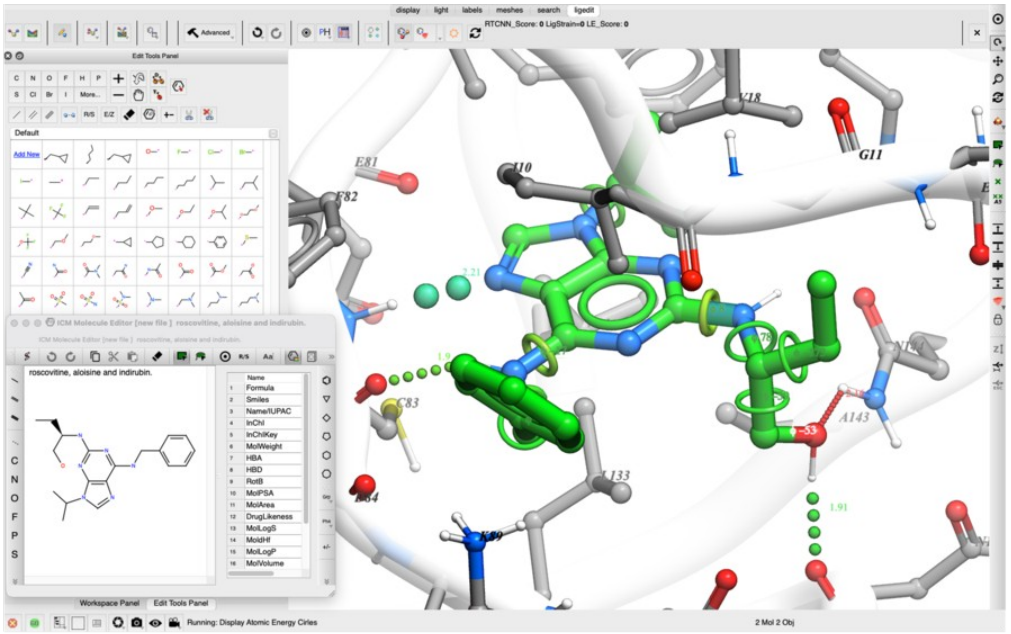


https://ctlee.github.io/BioChemCoRe-2018/system-prep/

Webb B et al. (2016) Curr. Protoc. Bioinformatics 54, 5.6.1–5.6.37

Ligand preparation: chemistry matters

- Ligand geometry and protonation must be optimized for the intended pH and tautomeric state
- Molecular mechanics uses point charges: assignment method (e.g., AM1-BCC vs RESP) greatly influences accuracy
- Aromaticity, hybridization, and stereochemistry must be checked and preserved



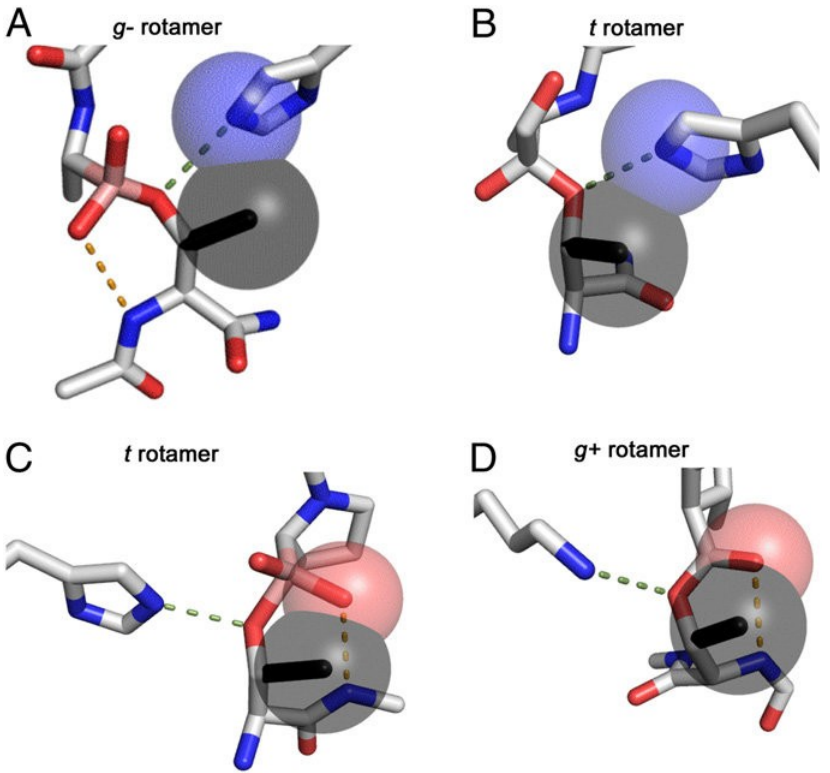
ICM-Chemist-Pro 3D Ligand Editor

https://www.molsoft.com/ligand-editor.html

Wang J et al. (2004) J. Comput. Chem. 25, 1157–1174

Energy minimization: purpose and methods

- Objective: resolve steric clashes and optimize hydrogen bonding networks before dynamics
- Steepest descent is robust for large forces; conjugate gradient is better for fine relaxation
- Minimization ensures the system is physically plausible at 0 K with negligible net forces
- Restraints alow flexible treatment of solute/solvent interactions during preparation

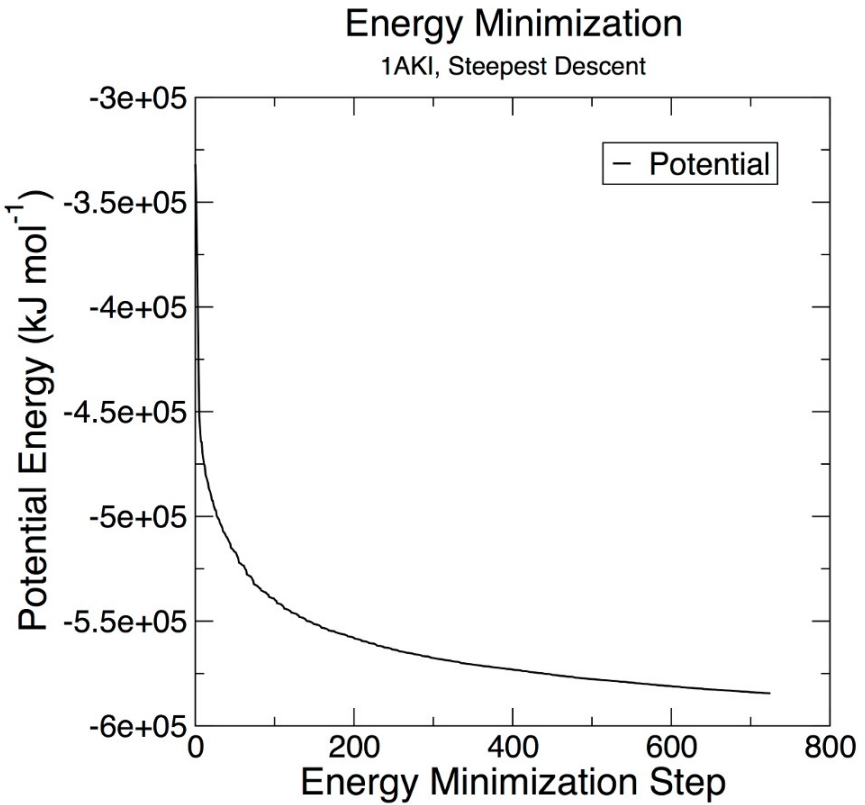


Schlick T (2010) Springer, New York

Buller A. R. et al. (2013) Proc. Natl. Acad. Sci. U.S.A. 110(8), E653–E661

Energy minimization: practical considerations

- Converge criteria: maximum force < 1,000 kJ/mol nm; total energy stabilization
- In GROMACS, use position restraint files with define = -DPOSRES
- In AMBER, use restraintmask and restraint_wt in sander
- Inspect minimization trajectory visually to confirm structural integrity

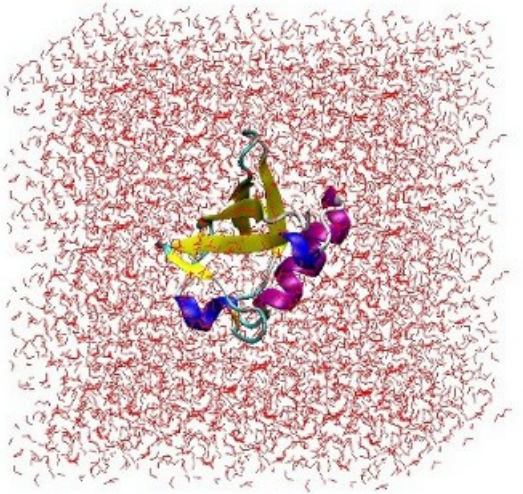


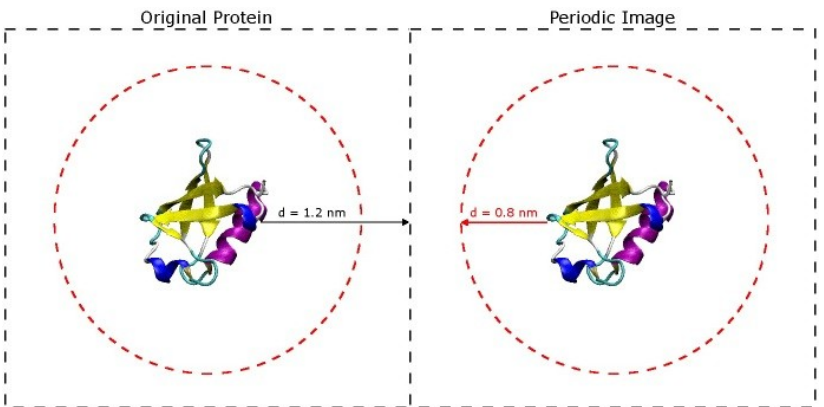
Lemkul JA (2024) J. Phys. Chem. B 128, 9418–9435

http://www.mdtutorials.com/gmx/lysozyme/05_EM.html

Solvation: why it matters

- Solvent provides dielectric screening and mimics the cellular environment
- Box size should allow ≥ 1 nm between solute and box edge (avoid artifacts)
- Shape influences atom count and equilibration time: truncated ocahedra are efficient for globular proteins
- Electrostatic calculations (PME) assume full periodicity



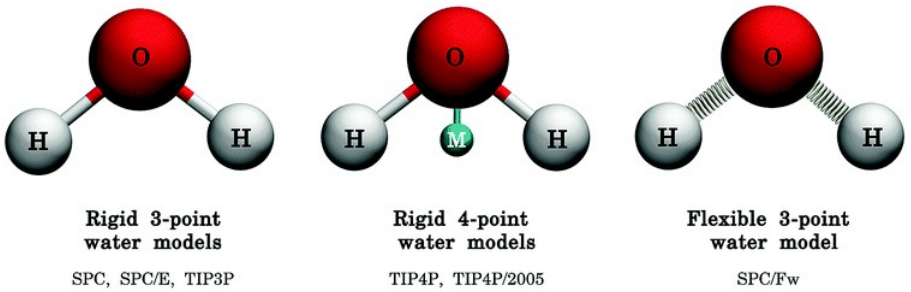


Berendsen HJC et al. (1984) J. Chem. Phys. 81, 3684–3690

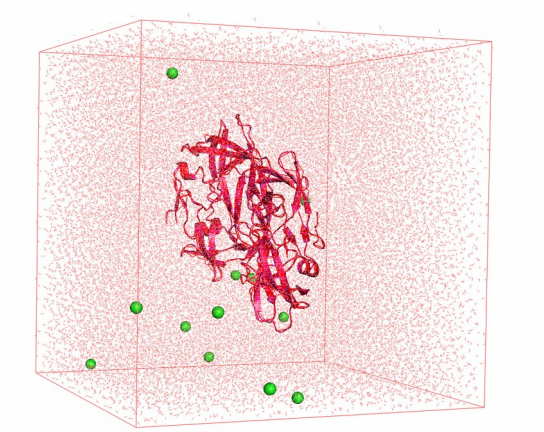
https://ringo.ams.stonybrook.edu/index.php/MD_Simulation:_Protein_in_Water

Solvation and ionization

- Commom models: TIP3P (fast, widely used), OPC (more accurate dipole), SPC/E (good for energies)
- Ions must neutralize the system
- Random or distance-based placement strategies for Na+/Cl-
- Check net charge before simulation (e.g., gmx grompp or tleap warnings)



Prasad K V et al. (2018) Phys. Chem. Chem. Phys. 20, 16005–16011



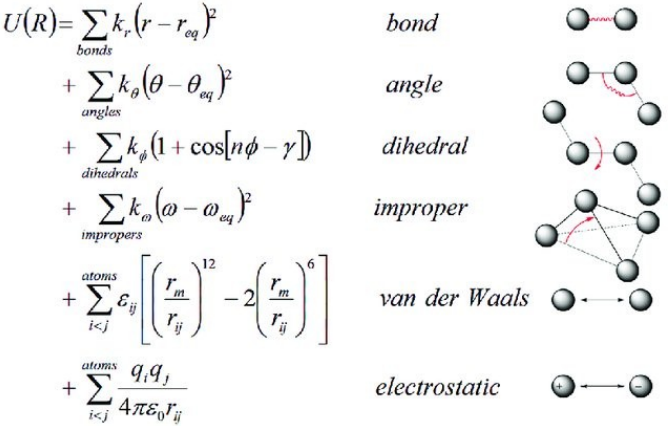
Mark P et al. (2001) J. Chem. Phys. 114, 5462–5475

https://www.compchems.com/setting-up-a-molecular-dynamics-simulation/

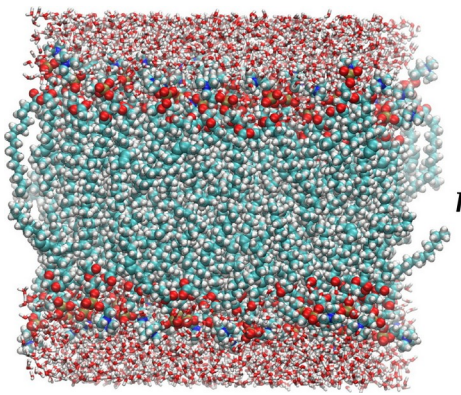
Force field theory

- Force fields describe potential energy using empirical parameters:
  - Bonded terms: bonds, angles, dihedrals
  - Nonbonded: Lennard-Jones + Coulomb
- Water models are parameterized with specific FFs (e.g., TIP3P for CHARMM36m)
- CHARMM36m and AMBER ff19SB include improved backbone dihedral sampling

$$U(R) = \sum_{bonds} k_r (r - r_{eq})^2 \quad \text{bond}$$
$$+ \sum_{angles} k_\theta (\theta - \theta_{eq})^2 \quad \text{angle}$$
$$+ \sum_{dihedrals} k_\phi (1 + \cos[n\phi - \gamma]) \quad \text{dihedral}$$
$$+ \sum_{impropers} k_\omega (\omega - \omega_{eq})^2 \quad \text{improper}$$
$$+ \sum_{i<j}^{atoms} \varepsilon_{ij} \left[\left(\frac{r_m}{r_{ij}} \right)^{12} - 2 \left(\frac{r_m}{r_{ij}} \right)^{6} \right] \quad \text{van der Waals}$$
$$+ \sum_{i<j}^{atoms} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad \text{electrostatic}$$



Chang C-A et al. (2016) Catalysts 6(6), 82



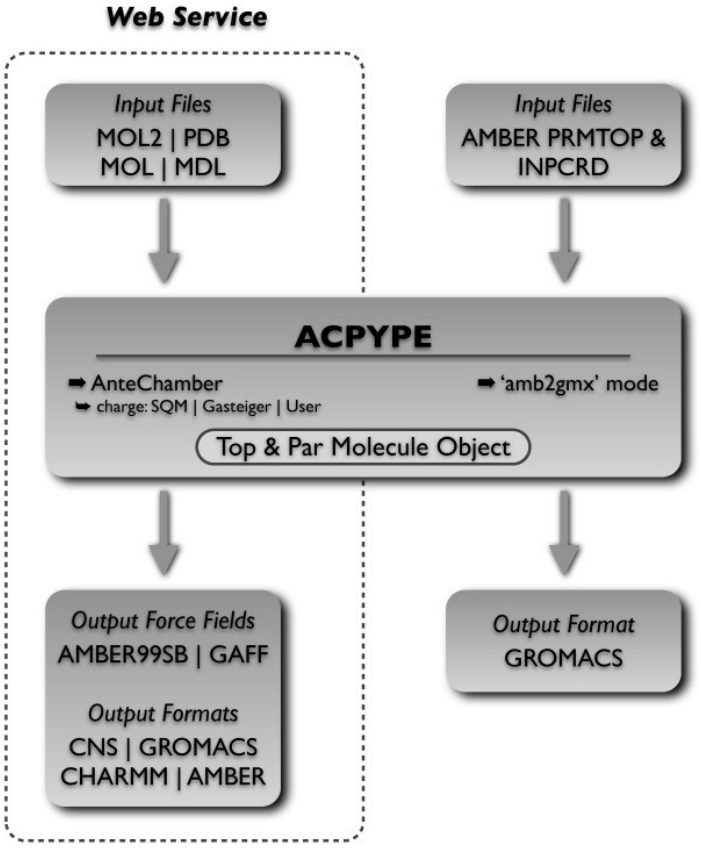
$$m\vec{a} = \vec{F} = \begin{cases} \text{CHARMM ?} \\ \text{AMBER ?} \\ \text{GROMOS ?} \\ \text{Slipids ?} \\ \ldots ? \end{cases}$$

Cornell WD et al. (1995) J. Am. Chem. Soc. 117, 5179–5197

Lyubartsev A. P. et al. (2016) Biochim. Biophys. Acta Biomembr. 1858(10)

Parameter assignment: automation and manual curation

- AMBER's LEaP combines library files and topology/coordinate generation
- GROMACS u sers can translate ligand topologies using ACPYPE (from AMBER) or CGenFF for CHARMM-based
- Manual verification: atom types, charges, torsions, and connectivity must match
- Always test ligand parameters separately (e.g., minimization in vacuum)



Sousa da Silva AW et al. (2012) BMC Res. Notes 5, 367

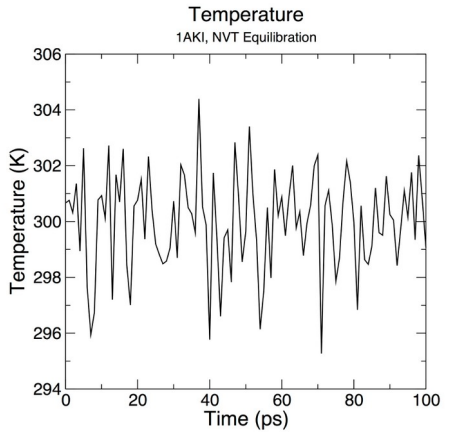
Common pirfalls and quality checks

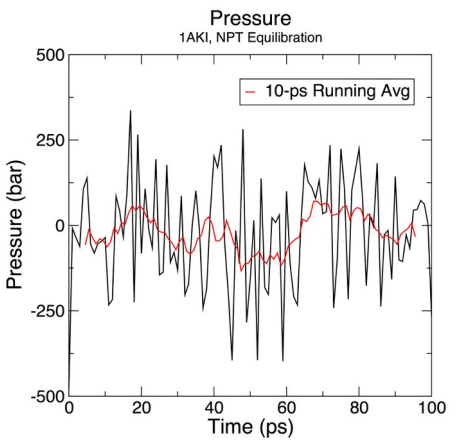
- Missing or missasigned atom types
- Wrong protonation near catalytic sites or metal cofactors
- Overlapping atoms after solvation due to insufficient box padding
- Missing or conflicting restraints between tools (e.g., default water box sizes)

Roe DR et al. (2013) J. Chem. Theory Comput. 9, 3084–3095

Equilibration: concepts

- NVT: stabilizes temperature using thermostats (e.g., velocity rescaling, Langevin)
- NPT: stabilizes pressure with barostat (Berendsen, Parrinello-Rahman)
- SHAKE/RATTLE algorithms constrain bonds involving H → allows 2 fs timestep
- Gradual release of restraints avoids destabilization of protein core



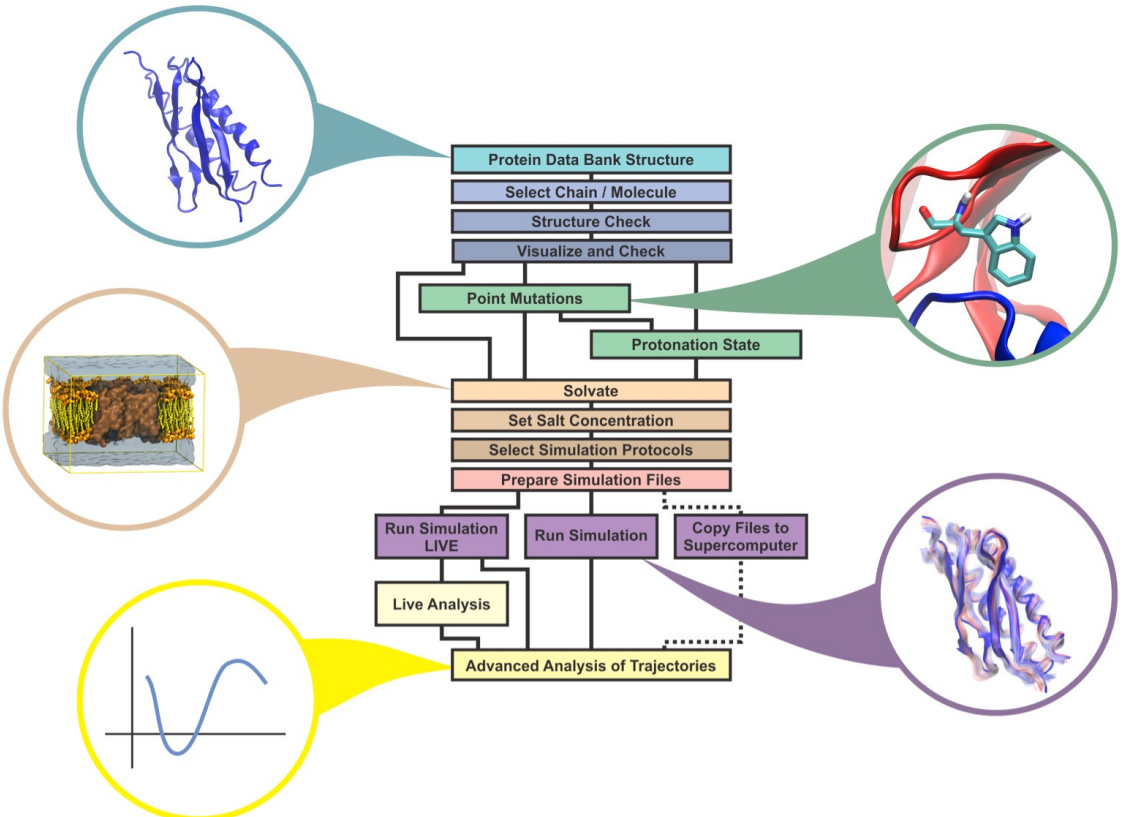


Izadi S et al. (2014) J. Phys. Chem. Lett. 5, 3863–3871

http://www.mdtutorials.com/gmx/lysozyme/06_equil.html

Best practices for robust preparation

- Check protonation, disulfides, termini, and ligands using multiple tools
- Validate topology and charges by calculating energy of minimized structure
- Save all intermediate files and document all assumptions (e.g., pH, salt)
- Visual inspection complements automated tools



Humphrey W et al. (1996) J. Mol. Graphics 14, 33–38

Scientific Reports volume 6, Article number: 26536 (2016)

Summary

- Structure preparation combines chemistry, structural biology, and physics
- Theoretical foundation is as important as automation
- Well-prepared systems yield reproducible and interpretable simulations

Leach AR. (2001) Molecular Modelling: Principles and Applications.

Rapaport DC. (2004) The Art of Molecular Dynamics Simulation.

References

- Karplus, M., McCammon, J. Molecular dynamics simulations of biomolecules. Nat. Struct. Mol. Biol. 9, 646–652 (2002). https://doi.org/10.1038/nsb0902-646
- Ribeiro, J.V. et al. QwikMD – Integrative Molecular Dynamics Toolkit for Novices and Experts. Sci. Rep. 6, 26536 (2016). https://doi.org/10.1038/srep26536
- Hollingsworth, S.A., Dror, R.O. Molecular dynamics simulation for all. Neuron 99, 1129–1143 (2018). https://doi.org/10.1016/j.neuron.2018.08.011
- Olsson, M.H.M. et al. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. J. Chem. Theory Comput. 7, 525–537 (2011). https://doi.org/10.1021/ct100578z
- Anandakrishnan, R. et al. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res. 40, W537–W541 (2012). https://doi.org/10.1093/nar/gks375
- Webb, B., Sali, A. Comparative protein structure modeling using MODELLER. Curr. Protoc. Bioinformatics 54, 5.6.1–5.6.37 (2016). https://doi.org/10.1002/cpbi.3
- Wang, J. et al. Development and testing of a general Amber force field. J. Comput. Chem. 25, 1157–1174 (2004). https://doi.org/10.1002/jcc.20035
- Schlick, T. Molecular Modeling and Simulation: An Interdisciplinary Guide. 2nd ed. Springer, New York, 2010. ISBN: 978-1441963505
- Lemkul, J.A. Introductory Tutorials for Simulating Protein Dynamics with GROMACS. J. Phys. Chem. B 128, 9418–9435 (2024). https://doi.org/10.1021/acs.jpcb.4c04901
- Berendsen, H. J. C., Grigera, J. R. T., & Straatsma, T. P. (1978). The missing term in effective pair potentials. The Journal of Physical Chemistry A, 91(24), 6269–6271. 10.1021/j100308a038