Protein Parametrizations

Ladislav Bartoš

March 29, 2021

FB810/C9926 Course

Molecular Mechanics

$$E_{\text{bond}} = \frac{1}{2}\sum_{i=1}^{N_B} K_{r,i}(r_i - r_{i,0})^2 \qquad E_{\text{electrostatic}} = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r_{ij}}$$

$$E_{\text{angle}} = \frac{1}{2}\sum_{i=1}^{N_A} K_{\theta,i}(\theta_i - \theta_{i,0})^2$$

$$E_{\text{dihedral}} = \sum_{i=1}^{N_D}\sum_{n=1}^{N_\varphi} \frac{V_{n,i}}{2}\left[1 + \cos\left(n_i\varphi_i - \varphi_{i,0}\right)\right]$$

$$E_{\text{vdW}} = \sum_{i=1}^{N}\sum_{j>i}^{N} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

- what should the **parameters** in the equations be?
- simulations should **match** the reality ← *easier said than done*

Force Field Families

AMBER

ff94
ff96
ff99
ff99SB
ff99SB*
ff99SB-ILDN
ff99SB*-ILDN
ff99SB-ILDN-φ
ff99SB-ILDN-NMR
ff03
ff03W
ff03*
ff03WS
...

CHARMM

C22
C22*
C27
C36
...

GROMOS

53a6
53a5
43a1
43a1p
...

OPLS-AA

OPLS-AA/M

Which force field to use?

Benchmark #1

- *Lindorff-Larsen et al., 2012*

- simulations of two folded proteins vs. NMR data
  → ff03, ff03*, C22, OPLS-AA ✖

- temperature dependent structural properties of proteins
  → C27, ff03 **overstabilize** helical structures ✖
  → ff99SB-ILDN **underestimates** the stability of helices ✖

- folding of α-helical & β-sheet proteins
  → ff03, C22, C27, OPLS-AA ✖

- best force fields? → **ff99SB*-ILDN** & **C22***

Benchmark #1: Force Fields

AMBER

ff94

ff96

ff99

ff99SB

ff99SB*

~~ff99SB-ILDN~~

ff99SB*-ILDN

ff99SB-ILDN-φ

ff99SB-ILDN-NMR

~~ff03~~

ff03W

~~ff03*~~

ff03WS

...

CHARMM

~~C22~~

C22*

~~C27~~

C36

...

~~OPLS-AA~~

OPLS-AA/M

GROMOS

53a6

53a5

43a1

43a1p

...

Benchmark #2

- *Beauchamp et al., 2012*

- simulations of peptides compared with a large number of NMR measurements

- **ff99SB-ILDN-φ** & **ff99SB-ILDN-NMR**
  → calculation errors ≈ experimental uncertainty

- **new force fields:**
  > Amber: ff14SB, ff19SB
  > CHARMM: C36

Benchmark #2: Force Fields

AMBER

ff94

ff96

ff99

ff99SB

ff99SB*

ff99SB-ILDN

ff99SB*-ILDN

ff99SB-ILDN-φ

ff99SB-ILDN-NMR

ff03

ff03W

ff03*

ff03WS

...

CHARMM

C22

C22*

C27

C36

...

OPLS-AA

OPLS-AA/M

GROMOS

53a6

53a5

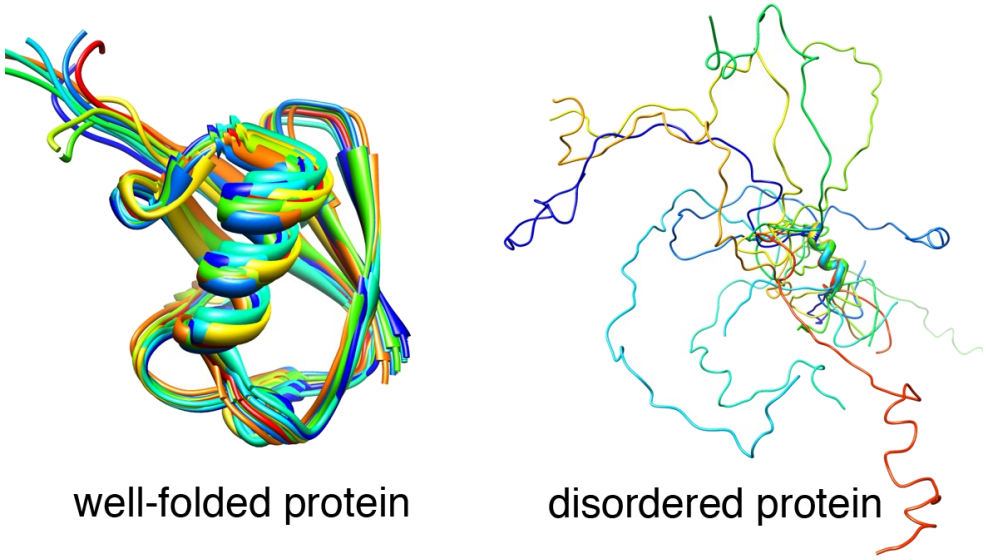
43a1

43a1p

...

Intrinsically Disordered Proteins

- intrinsically disordered = no fixed tertiary structure



well-folded protein disordered protein

- problem 1: capturing disordered regions
- problem 2: capturing **both** the folded & disordered regions
- CHARMM: C36 → **C36m** Amber: ff99SB → **ff99SB-disp**

Conclusion

- know the **limitations** of your force field

  Internal dynamics of α-helical protein?
  > FF underestimating α-helix stability 

- choose a force field **appropriate** for your system

  *A system with **no** α-helices?*
  > FF underestimating α-helix stability

- when in doubt:
  > simulate the system twice with different force fields