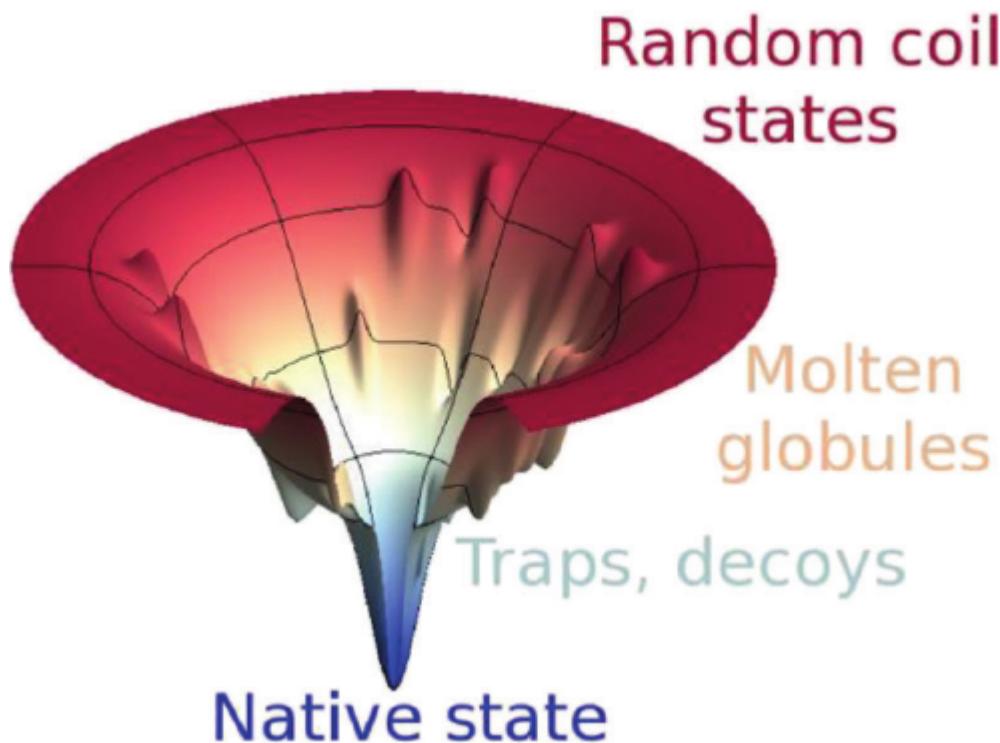


Workshop #3

Protein folding using Associative, Water-mediated, Single memory, Energy Model



ARE THERE PATHWAYS FOR PROTEIN FOLDING ?

by CYRUS LEVINTHAL

[Massachusetts Institute of Technology, Department of Biology Cambridge, Massachusetts.]

restored. This experimental result has lead to the suggestion that a native protein exists in some kind of thermodynamic configurational equilibrium, with the biologically active state being the one of lowest configurational energy. An alternative view is that the native protein is in a uniquely selected metastable state, in which the configurational energy is at a local minimum but not necessarily at an absolute minimum. In this latter model, the protein is not assumed to be in an equilibrium state, and one must postulate some sequence of events which takes place for each molecule so that the protein reaches the correct metastable state.

STUDIES ON PROTEIN FOLDING, UNFOLDING AND FLUCTUATIONS BY COMPUTER SIMULATION

I. The effect of specific amino acid sequence represented by specific inter-unit interactions

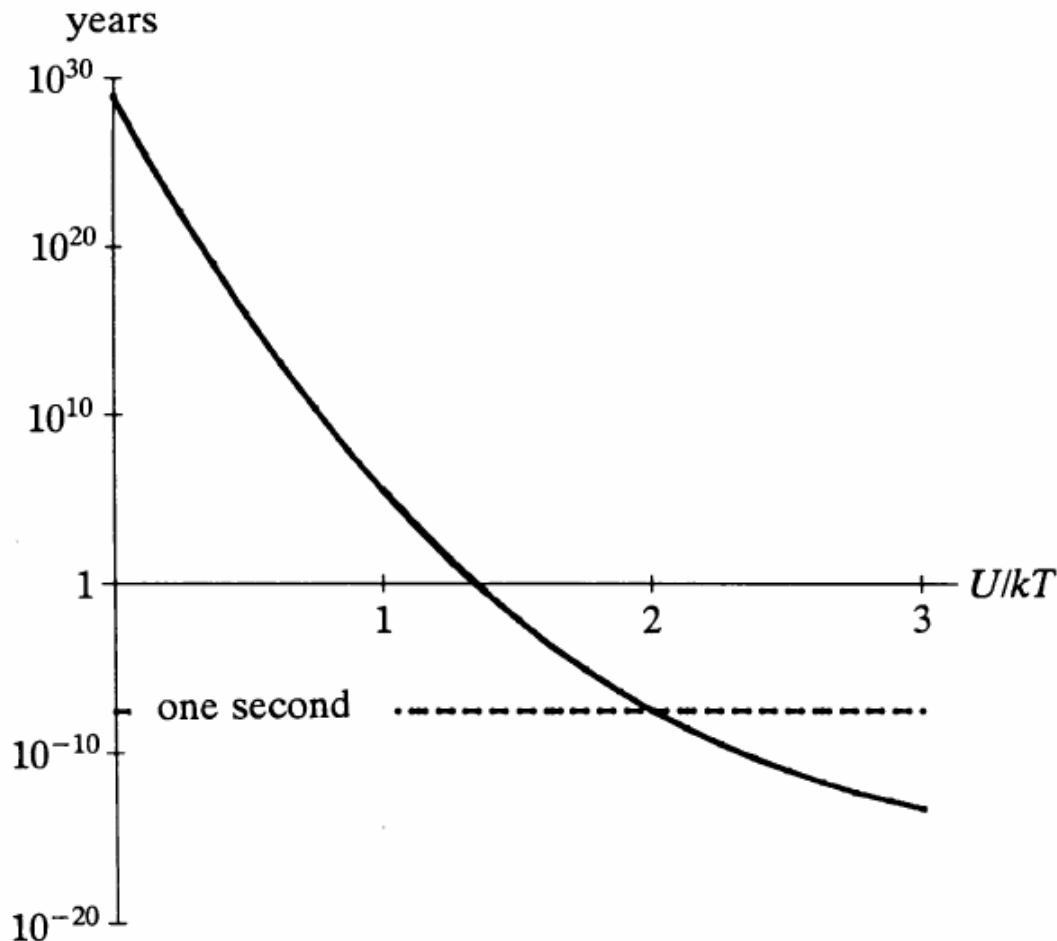
HIROSHI TAKETOMI, YUZO UEDA and NOBUHIRO Gō

Department of Physics, Faculty of Science, Kyushu University, Fukuoka, Japan

Received 9 December 1974

A lattice model of proteins is introduced. “A protein molecule” is a chain of non-intersecting units of a given length on the two-dimensional square lattice. The copolymeric character of protein molecules is incorporated into the model in the form of specificities of inter-unit interactions. This model proved most effective for studying the statistical mechanical characteristics of protein folding, unfolding and fluctuations. The specificities of inter-unit interactions are shown to be the primary factors responsible for the all-or-none type transition from native to denatured states of globular proteins. The model has been studied by the Monte Carlo method of Metropolis et al., which is now shown applied to approximately simulating a kinetic process. In the strong limit of the specificity of the inter-unit interaction the native conformation was reached in this method by starting from an extended conformation. The possible generalization and application of this method for finding the native conformation of proteins from their amino acid sequence are discussed.

A slight push in the correct direction...



$$\frac{d}{dt} [c] = -k_0[c] + k_1[i], \quad [c] + [i] = 1.$$

$$[i]_{\text{eq}}/[c]_{\text{eq}} = k_0/k_1 = K.$$

$$K = k_0/k_1 = \nu e^{-U/kT}.$$

$$\tau(S) \cong (1/Nk_0)(1 + k_0/k_1)^N$$

Levinthal's paradox

ROBERT ZWANZIG, ATTILA SZABO, AND BIMAN BAGCHI*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, Building 2, National Institutes of Health, Bethesda, MD 20892

Contributed by Robert Zwanzig, October 7, 1991

We can apply the Go postulate

$$U(\Gamma, \Gamma_0) = \sum_{\text{bonds}}^{N-1} K_b (b_i - b_{0i})^2 + \sum_{\text{angles}}^{N-2} K_\theta (\theta_i - \theta_{0i})^2 \\ + \sum_{\text{dihedrals}}^{N-3} \{K_\phi^{(1)} [1 - \cos(1 \times (\Phi_i - \Phi_{0i}))] \\ + K_\phi^{(3)} [1 - \cos(3 \times (\Phi_i - \Phi_{0i}))]\} \\ + \sum_{\text{native contacts } |i-j|>3} \left\{ \varepsilon \left[5 \left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{r_{0ij}}{r_{ij}} \right)^{10} \right] \right\} \\ + \sum_{\text{non-native contacts, } |i-j|>3} \left(\frac{C}{r_{ij}} \right)^{12} \quad (1)$$

Many flavors can be added:

Many basins

Many contact sets

Many dihedral sets

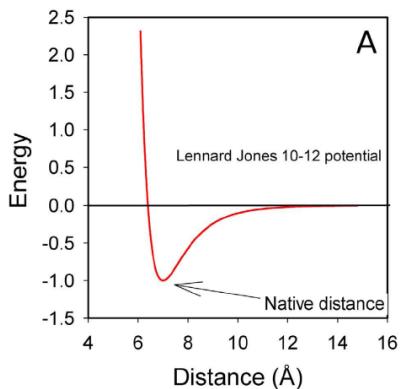
Charges (coulombic, Debye Huckel)

Single bead structures

Many bead structures

Non native interactions

And many more



An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields

Paul C. Whitford,¹ Jeffrey K. Noel,¹ Shachi Gosavi,¹ Alexander Schug,¹ Kevin Y. Sanbonmatsu,² and José N. Onuchic^{1*}

We can apply the Go postulate

$$U(\Gamma, \Gamma_0) = \sum_{\text{bonds}}^{N-1} K_b (b_i - b_{0i})^2 + \sum_{\text{angles}}^{N-2} K_\theta (\theta_i - \theta_{0i})^2 \\ + \sum_{\text{dihedrals}}^{N-3} \{K_\phi^{(1)} [1 - \cos(1 \times (\Phi_i - \Phi_{0i}))] \\ + K_\phi^{(3)} [1 - \cos(3 \times (\Phi_i - \Phi_{0i}))]\} \\ + \sum_{\text{native contacts } |i-j|>3} \left\{ \varepsilon \left[5 \left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{r_{0ij}}{r_{ij}} \right)^{10} \right] \right\} \\ + \sum_{\text{non-native contacts, } |i-j|>3} \left(\frac{C}{r_{ij}} \right)^{12} \quad (1)$$

Many flavors can be added:

Many basins

Many contact sets

Many dihedral sets

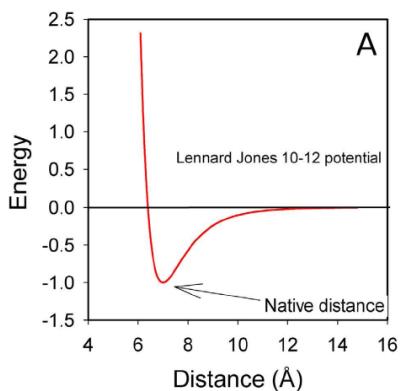
Charges (coulombic, Debye Huckel)

Single bead structures

Many bead structures

Non native interactions

And many more



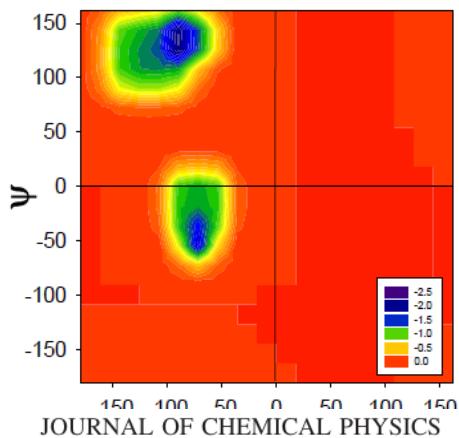
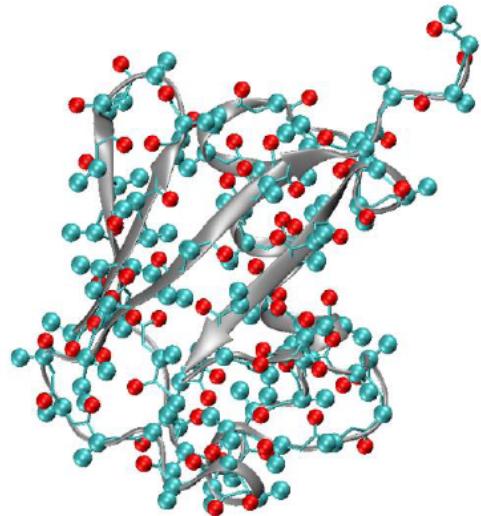
An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields

Paul C. Whitford,¹ Jeffrey K. Noel,¹ Shachi Gosavi,¹ Alexander Schug,¹ Kevin Y. Sanbonmatsu,² and José N. Onuchic^{1*}

Another form of a Go-potential with many-body pair interactions

AMH-Go

Go with non additivity



Coarse grain C α , C β , and O

$$H = H_{\text{backbone}} + H_{\text{na}}$$

$$H_{\text{backbone}} = \lambda_{\psi\phi} V_{\psi\phi} + \lambda_\chi V_\chi + \lambda_{\text{ex}} V_{\text{ex}} + \lambda_{\text{harm}} V_{\text{harm}}$$

Attractive forces native contacts

Homogeneous contact potential

Non additivity (p exponent) favors many body interactions

$$H_{\text{na}} = -\frac{1}{2} \sum_i |E_i|^p$$

$$E_i = \sum_j \epsilon_{ij}(r_{ij}) = - \sum_j \left| \frac{\epsilon}{a} \right|^{1/p} \theta(r_c - r_{ij}^N) \gamma_{ij} \times \exp \left(-\frac{(r_{ij} - r_{ij}^N)^2}{2\sigma_{ij}^2} \right)$$

$$\sigma_{ij} = |i - j|^{0.15} \text{ \AA}$$

VOLUME 114, NUMBER 10

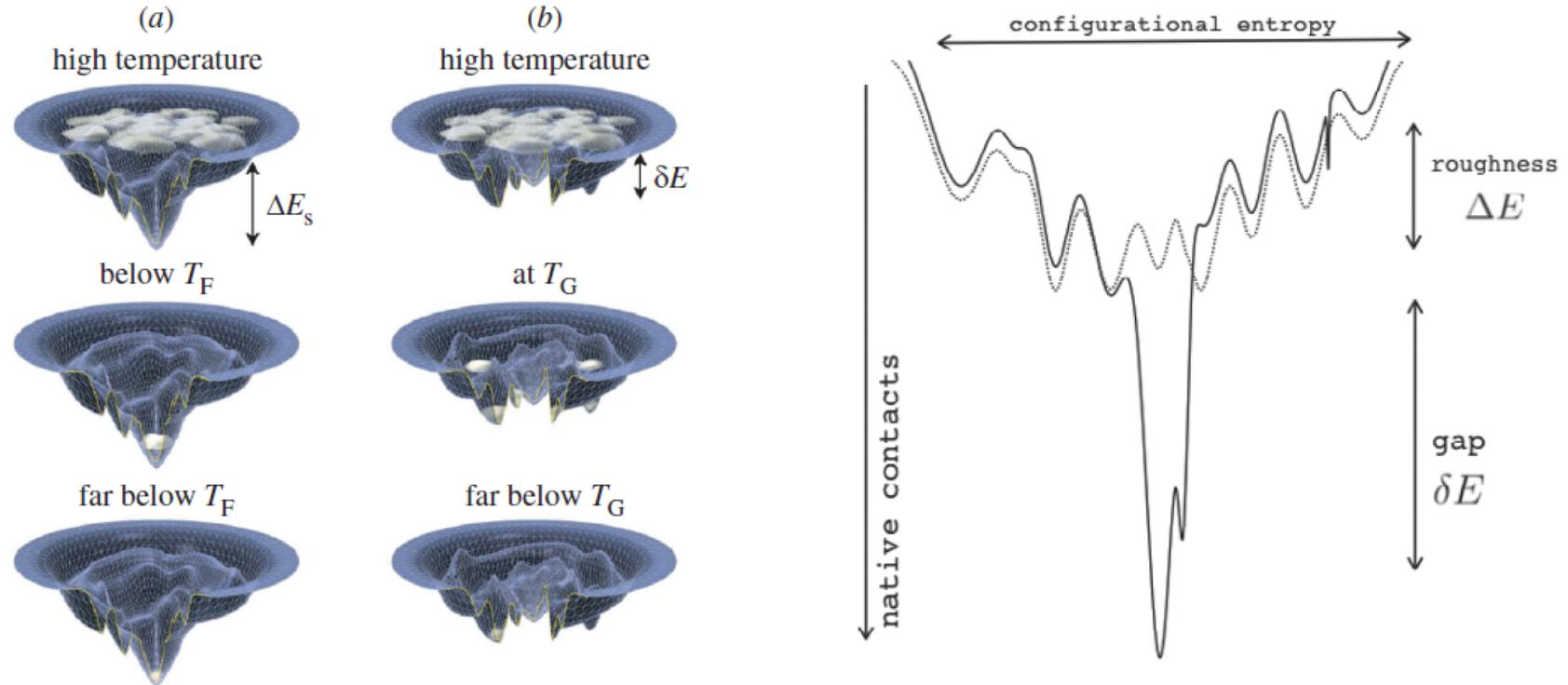
8 MARCH 2001

**Role of explicitly cooperative interactions in protein folding funnels:
A simulation study**

Michael P. Eastwood^{a)} and Peter G. Wolynes^{a),b)}

Department of Chemistry, University of Illinois, 600 S. Mathews Ave, Urbana, Illinois 61801

A few constraints following the Energy Landscape Theory

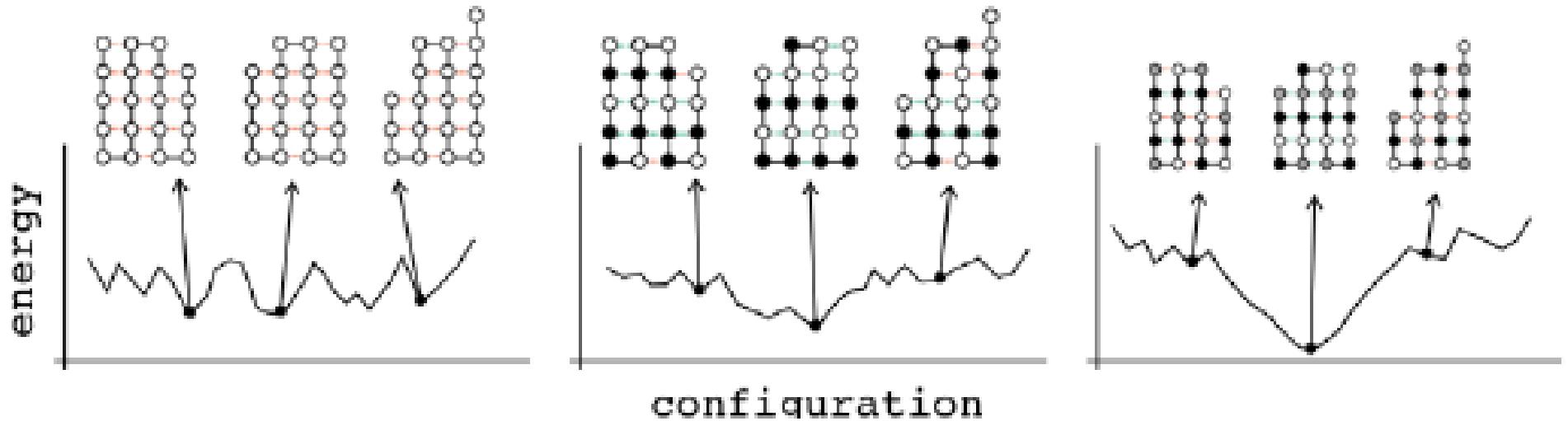


- It has been argued theoretically and experimentally that formation of native contacts in folding pathway increases the energy gap between unfolded conformations.
- Thus, as more native contacts are formed the native conformations are favored.
- Most simple structure-based models tend to funnel the landscape leaving energetically frustration to topological conflicts. These family of models have been used to understand processes such as folding, dimerization, binding, interactions.

Frustration in biomolecules

Diego U. Ferreiro¹, Elizabeth A. Komives^{2*} and Peter G. Wolynes³

Energy gaps and energy degeneracy



Frustration in biomolecules

Diego U. Ferreiro¹, Elizabeth A. Komives^{2*} and Peter G. Wolynes³

Energy Landscape gives a quantity to optimize

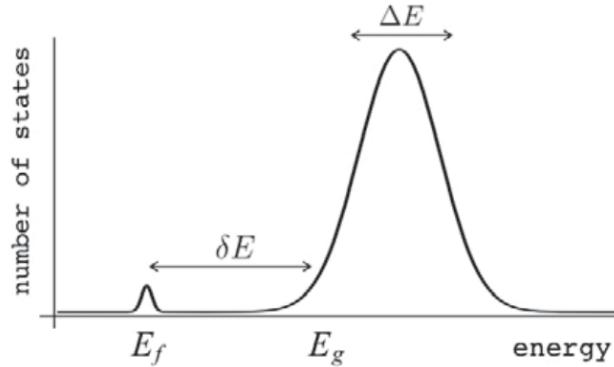
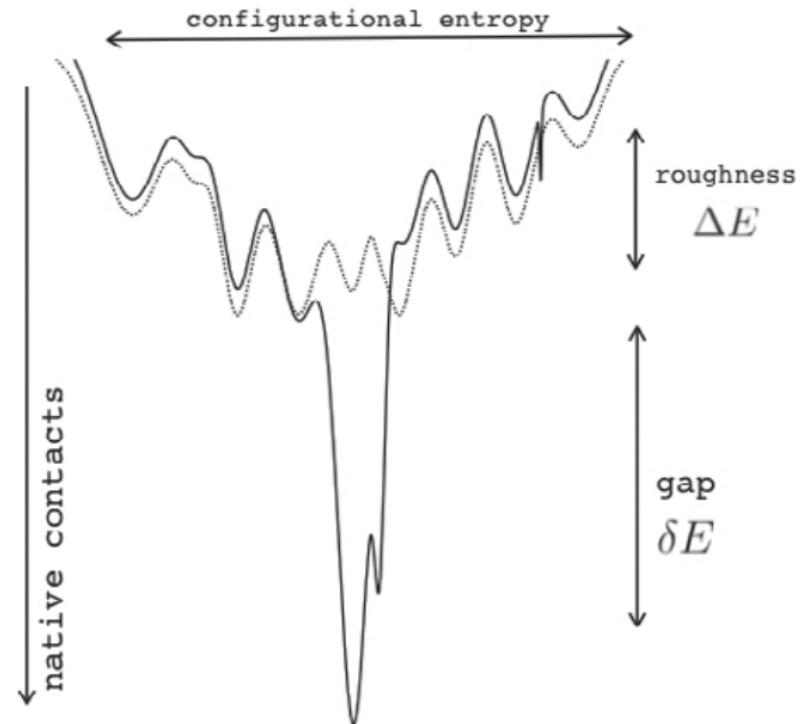


Fig. 7. The global effects of frustration can be quantified with the energy landscape theory. A frustrated heteropolymer will have the energy distribution drawn from a random energy landscape, a Gaussian with some mean \bar{E} and a standard deviation ΔE . The energy of the lowest energy states E_g can be estimated given the size of the configuration space and the variance of the energy distribution. A large protein whose energy is just E_g will have quite a few kinetic traps of nearly the same stability. Therefore, if a protein is to fold robustly, its ground-state energy must be substantially below E_g . If the energy of the completely folded state E_f is substantially below this estimate, we can predict that frustration effects will be minimal for this sequence (Bryngelson & Wolynes, 1987). The condition that E_f is below E_g is known as the *gap* condition for folding. Comparing E_f and E_g and requiring them to be well separated (δE) provides one way of quantitatively stating the Principle of Minimal Frustration.



$$\frac{T_f}{T_g} = \frac{\delta E}{\Delta E} = \frac{GAP}{ROUGHNESS}$$

Frustration in biomolecules

Diego U. Ferreiro¹, Elizabeth A. Komives^{2*} and Peter G. Wolynes³

Glassy, folding and collapse temperatures

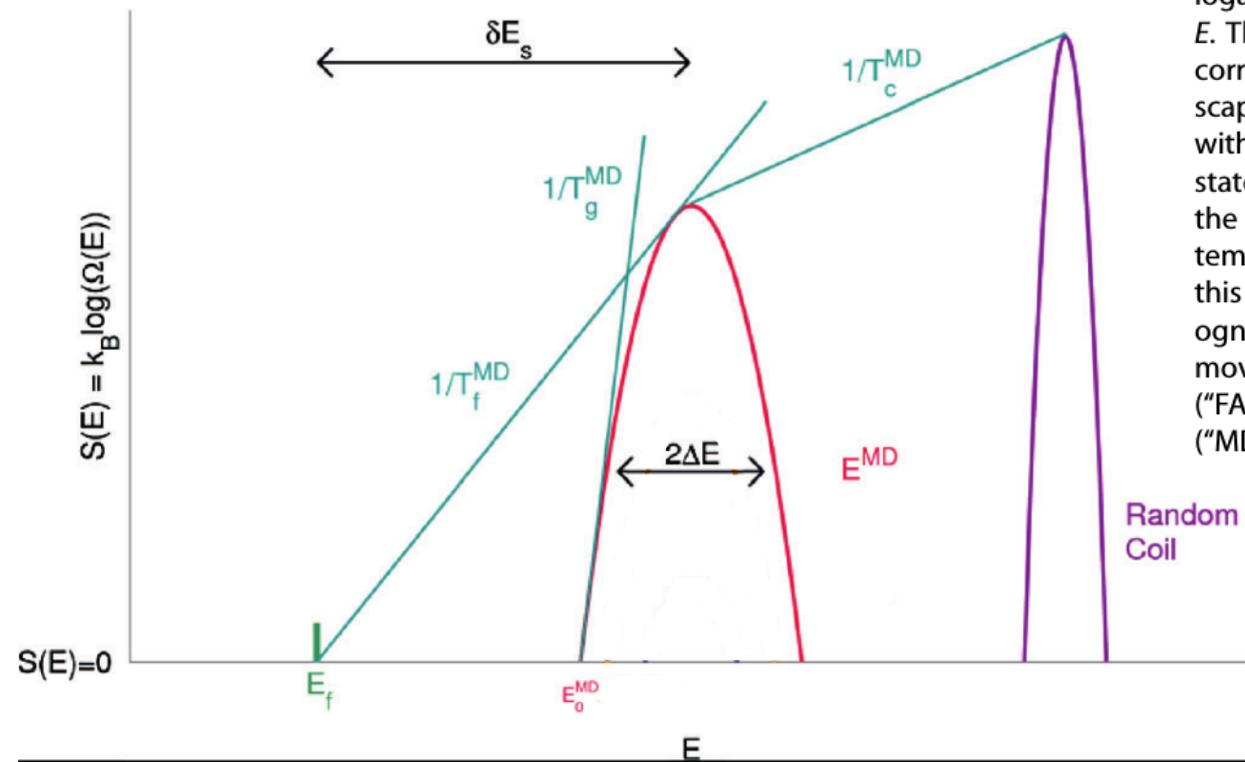


Figure 1. The theory behind this figure is described in Section 2. A logarithmic plot of the number of structures with a given energy, E . The expected ground state of a set of compact decoy structures corresponding to the molten globule can be inferred from landscape theory and is indicated by the intersection of a parabola with the abscissa. When there are more possible decoys, the trap states become more competitive and are easier to confuse with the native state at E_f . The gap is reflected also in the characteristic temperatures T_f and T_g , whose inverses are indicated as slopes on this diagram. A large T_f/T_g corresponds to a large gap and easy recognition. Recognition becomes progressively more difficult as one moves from threading-based decoys ("Thr"), to fragment assembly ("FA"), and finally to fully flexible backbone molecular dynamics ("MD").

Learning To Fold Proteins Using Energy Landscape Theory

Nicholas P. Schafer,^[a, c] Bobby L. Kim,^[b, c] Weihua Zheng,^[b, c] and Peter G. Wolynes^{*[a, b, c]}

History...

- **Associative Memory Hamiltonian (AMH)**

Short and medium distance-in-sequence contacts defined by short structural and sequence-aligned memories. Long range contacts defined by 4 letter code

- **Associative Memory with Contact (AMC)**

Long and medium range contacts defined by an optimized function using energy landscape theory. Local structural bias made by structural “memories”.

- **Associative Memory with Water (AMW)**

Papoian and Wolynes include a modification of the contact potential where distant surface contacts can interact at a larger distance based. This reflected the interaction mediated by a water molecule which improved dimer interface prediction and monomer structure prediction.

- **Associative memory, Water-mediated, Structure and Energy Model (AWSEM)**

LAMMPS implementation of the AMW potential

AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing

Aram Davtyan,[†] Nicholas P. Schafer,[‡] Weihua Zheng,[§] Cecilia Clementi,[‡] Peter G. Wolynes,^{*,‡,§} and Garegin A. Papoian^{*,†}

$$V_{\text{total}} = V_{\text{backbone}} + V_{\text{contact}} + V_{\text{burial}} + V_{\text{helical}} + V_{\text{FM}}$$

$$V_{\text{backbone}} = V_{\text{con}} + V_{\text{chain}} + V_{\chi} + V_{\text{rama}} + V_{\text{excl}}$$

$$\begin{aligned} V_{\text{excl}} = & \lambda_{\text{excl}} \sum_{ij} \Theta(r_{C_i, C_j} - r_{\text{ex}}^C) (r_{C_i, C_j} - r_{\text{ex}}^C)^2 \\ & + \lambda_{\text{excl}} \sum_{ij} \Theta(r_{O_i, O_j} - r_{\text{ex}}^O) (r_{O_i, O_j} - r_{\text{ex}}^O)^2 \end{aligned}$$

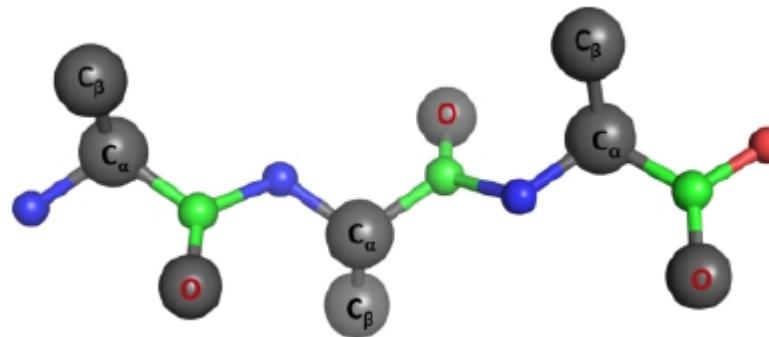
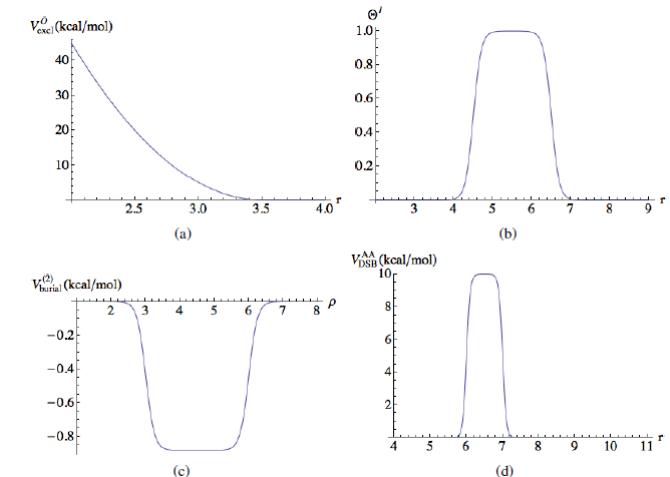
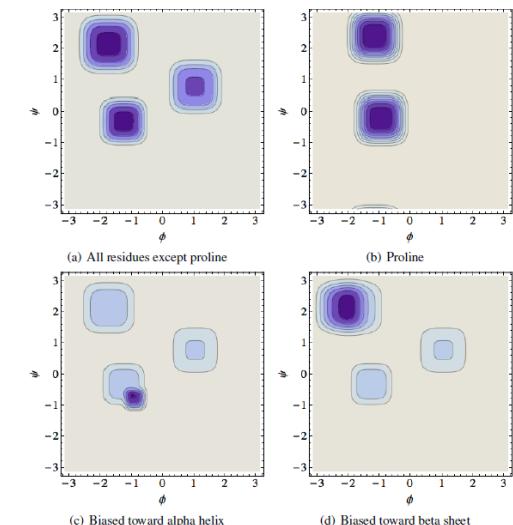


Figure S1: The connectivity of the chain is maintained by a combination of harmonic potentials. The distances constrained by V_{con} are shown as dashed lines and the distances constrained by V_{chain} are shown as double headed arrows.



Fragment memory potential

$$V_{AM} = -\lambda_{AM} \sum_m \omega_m \sum_{ij} \gamma_{ij} \exp \left[-\frac{(r_{ij} - r_{ij}^m)^2}{2\sigma_{IJ}^2} \right]$$

$$V_{FM} = -\lambda_{FM} \sum_m \sum_{ij} \exp \left[-\frac{(r_{ij} - r_{ij}^m)^2}{2\sigma_{ij}^2} \right]$$

AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing

Aram Davtyan,[†] Nicholas P. Schafer,[‡] Weihua Zheng,[§] Cecilia Clementi,[‡] Peter G. Wolynes,^{*,‡,§} and Garegin A. Papoian^{*,†}

AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing

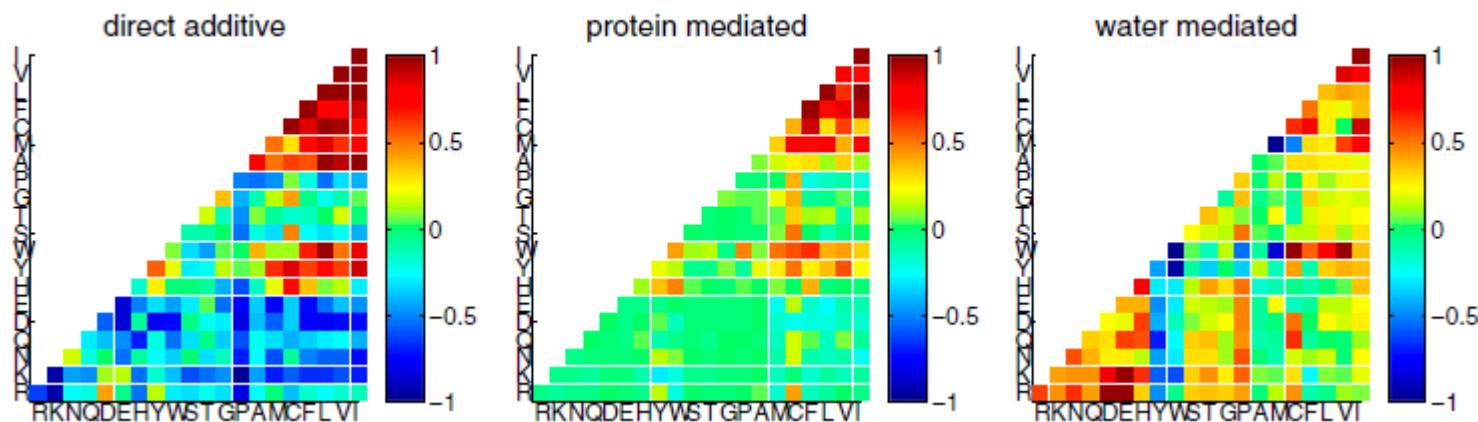
Aram Davtyan,[†] Nicholas P. Schafer,[‡] Weihua Zheng,[§] Cecilia Clementi,[‡] Peter G. Wolynes,^{*‡,§} and Garegin A. Papoian^{*†}

Mathematical optimization

$$V = \sum_i \gamma_i \phi_i$$

$$\delta E / \Delta E = A\gamma / \sqrt{\gamma B\gamma}$$

$$A_i = \langle \phi_i \rangle_{decoy} - \phi_{native}, \quad B_{i,j} = \langle \phi_i \phi_j \rangle_{decoy} - \langle \phi_i \rangle_{decoy} \langle \phi_j \rangle_{decoy}$$



$$V_{direct} = -\lambda_{direct} \sum_{j-i>9}^N \gamma_{ij}(a_i, a_j) \Theta_{ij}^I$$

$$V_{water} = -\lambda_{water} \sum_{j-i>9}^N \Theta_{ij}^{II} \left(\sigma_{ij}^{wat} \gamma_{ij}^{wat}(a_i, a_j) + \sigma_{ij}^{prot} \gamma_{ij}^{prot}(a_i, a_j) \right)$$

Eigenvalue decomposition complexity

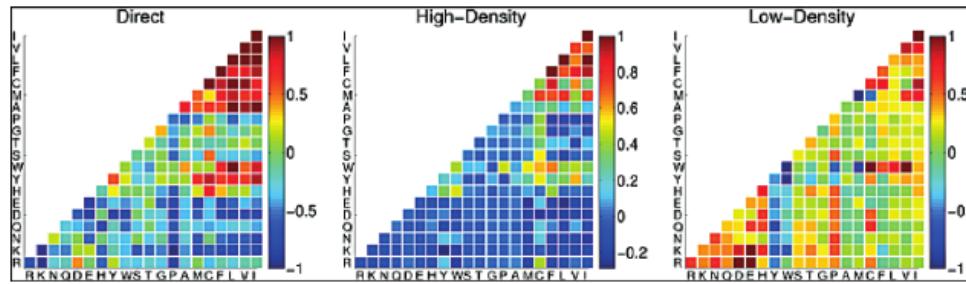
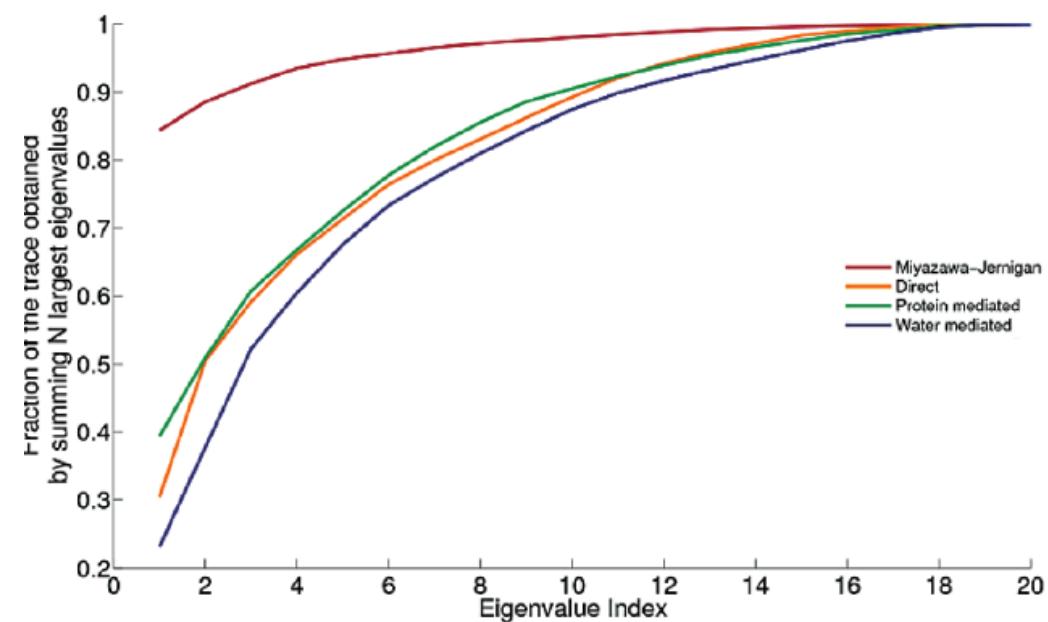


Figure 5. The three optimized interaction matrices used by the AMW/AWSEM models. The interaction weight for each pair of residue types (shown in one-letter codes) is represented in color such that red interactions are favorable and blue interactions are unfavorable. Two residues interact using the interaction weights specified in the “direct” matrix when their C_β atoms are between 4.5 and 6.5 Å apart. When the C_β atoms of two residues are between 6.5 and 9.5 Å apart, the interaction can be either “water-mediated” (low-density) or “protein-mediated” (high-density), depending on the degree of burial of the two residues.



Learning To Fold Proteins Using Energy Landscape Theory

Nicholas P. Schafer,^[a, c] Bobby L. Kim,^[b, c] Weihua Zheng,^[b, c] and Peter G. Wolynes^{*[a, b, c]}

AWSEM potential has reminiscences of physicochemical quantities

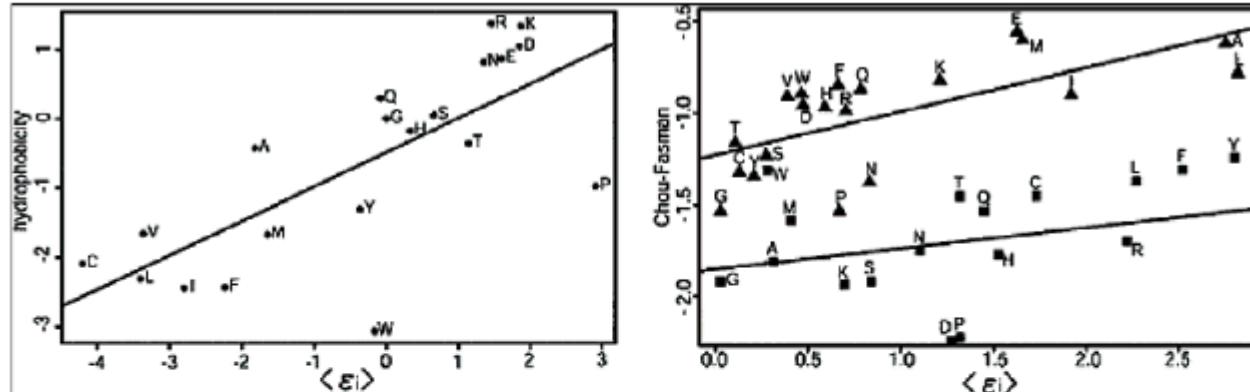
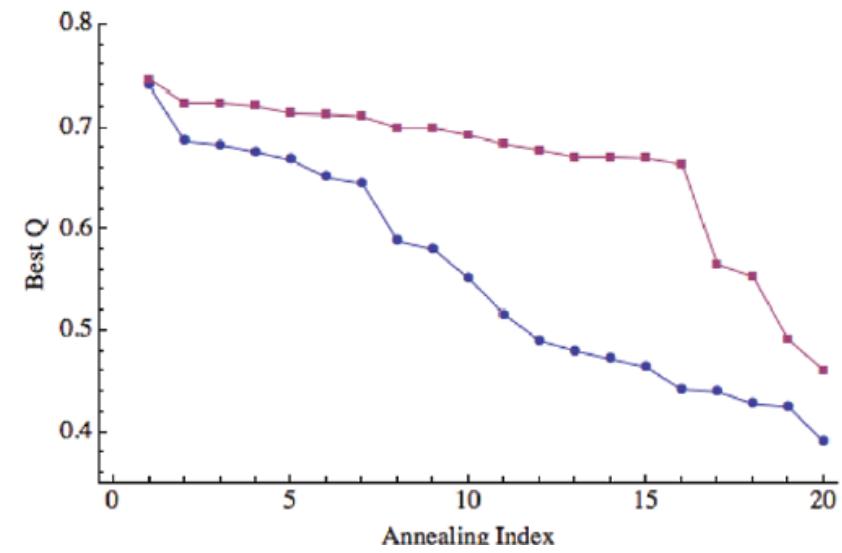
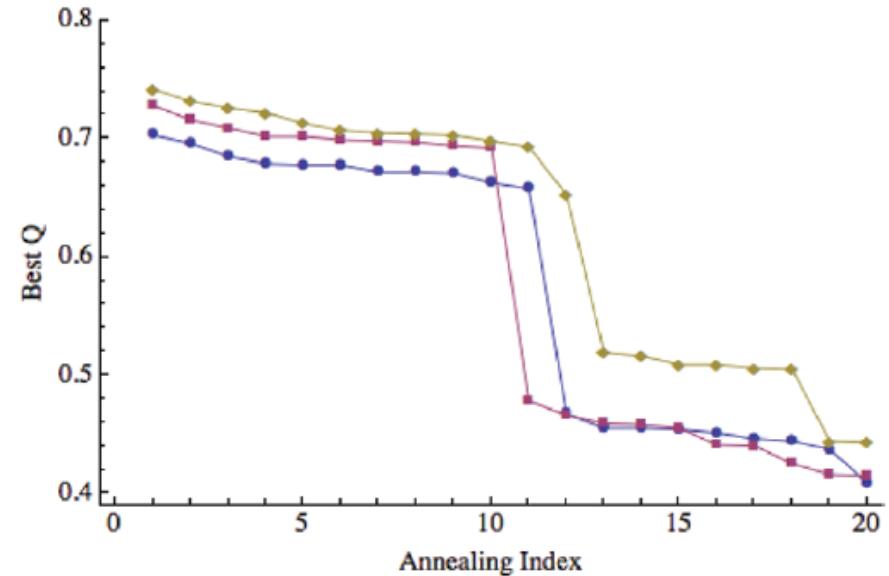
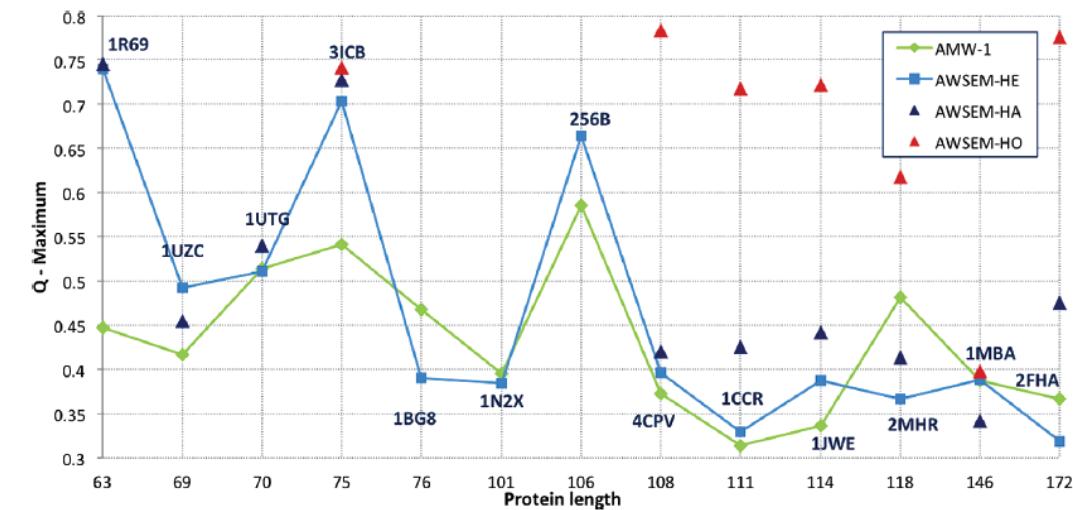


Figure 4. Two plots showing the correlation between optimized parameters in a coarse-grained Hamiltonian and experimental quantities. The burial energy is correlated with a hydrophobicity scale, and the secondary structure energies are correlated with experimentally determined secondary structure propensities. This figure was adapted from Ref. [18].

$$V_{burial} = -\frac{1}{2} \lambda_{burial} \sum_{i=1}^N \sum_{\mu=1}^3 \gamma_{burial}(a_i, \rho_i) (\tanh [\eta (\rho_i - \rho_{min}^\mu)] + \tanh [\eta (\rho_{max}^\mu - \rho_i)])$$

Learning To Fold Proteins Using Energy Landscape Theory

AWSEM structural predictions



$$Q = \frac{2}{(N-2)(N-3)} \sum_{i < j-2} \exp \left[-\frac{(r_{ij} - r_{ij}^N)^2}{2\sigma_{ij}^2} \right]$$

Different database of fragment memories yield different efficiencies

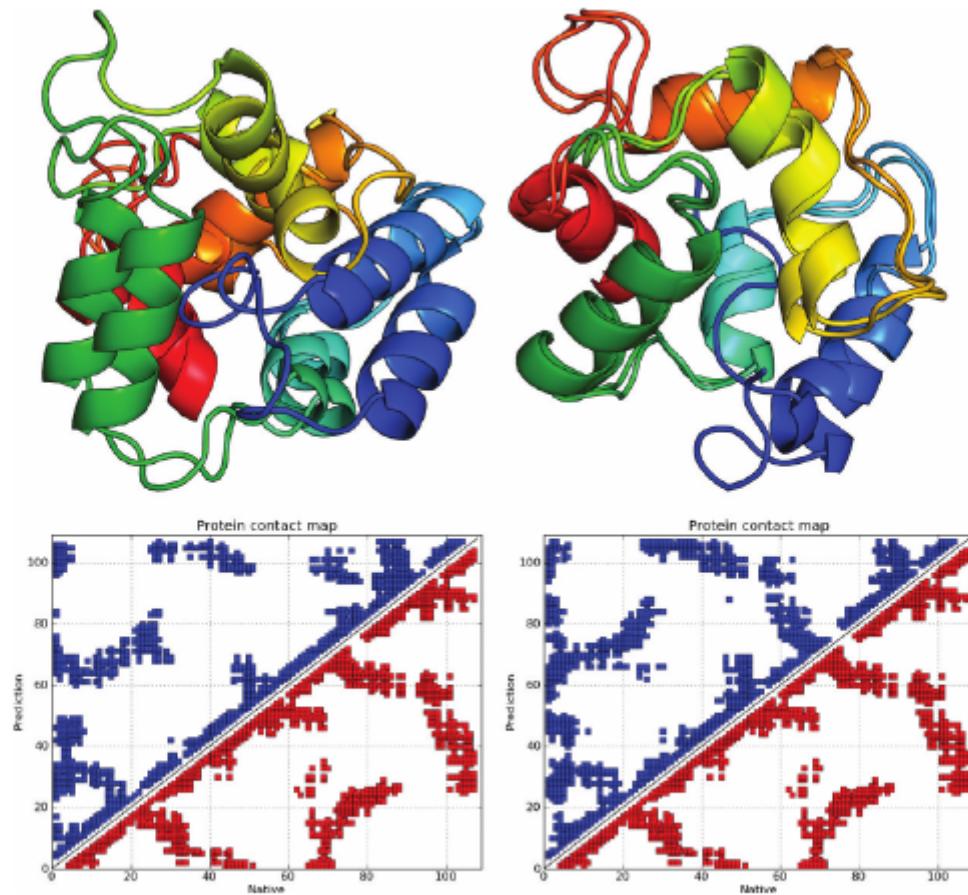


Figure 7. Structural alignments and comparative contact maps of the maximum Q score structures for 4CPV, with the "homologues excluded" prediction on the left ($Q = 0.396$, RMSD 5.8 Å) and the "homologues only" prediction on the right ($Q = 0.784$, RMSD 1.3 Å).

THE JOURNAL OF
PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing

Aram Davtyan,[†] Nicholas P. Schafer,[‡] Weihua Zheng,[§] Cecilia Clementi,[‡] Peter G. Wolynes,^{*‡§} and Garegin A. Papoian^{*‡}

AWSEM potential has recently been also trained to predict membrane protein structure

The $V_{membrane}$ Implicit Membrane Potential. We model the effects of membrane using the implicit $V_{membrane}$ potential,

$$V_{membrane} = V_{periplasmic} \delta_{zpred, periplasmic} + V_{cytoplasmic} \delta_{zpred, cytoplasmic} + V_{transmembrane} \delta_{zpred, transmembrane}$$

[S11]

$$V_{periplasmic} = - \sum_i \gamma_{peri} \Theta_{peri}(z_i) + \gamma_{cyto} \Theta_{cyto}(z_i) - \gamma_{memb} \Theta_{memb}(z_i)$$

$$V_{cytoplasmic} = - \sum_i \gamma_{peri} \Theta_{peri}(z_i) + \gamma_{cyto} \Theta_{cyto}(z_i) - \gamma_{memb} \Theta_{memb}(z_i)$$

$$V_{transmembrane} = - \sum_i -\gamma_{peri} \Theta_{peri}(z_i) - \gamma_{cyto} \Theta_{cyto}(z_i) + \gamma_{memb} \Theta_{memb}(z_i)$$

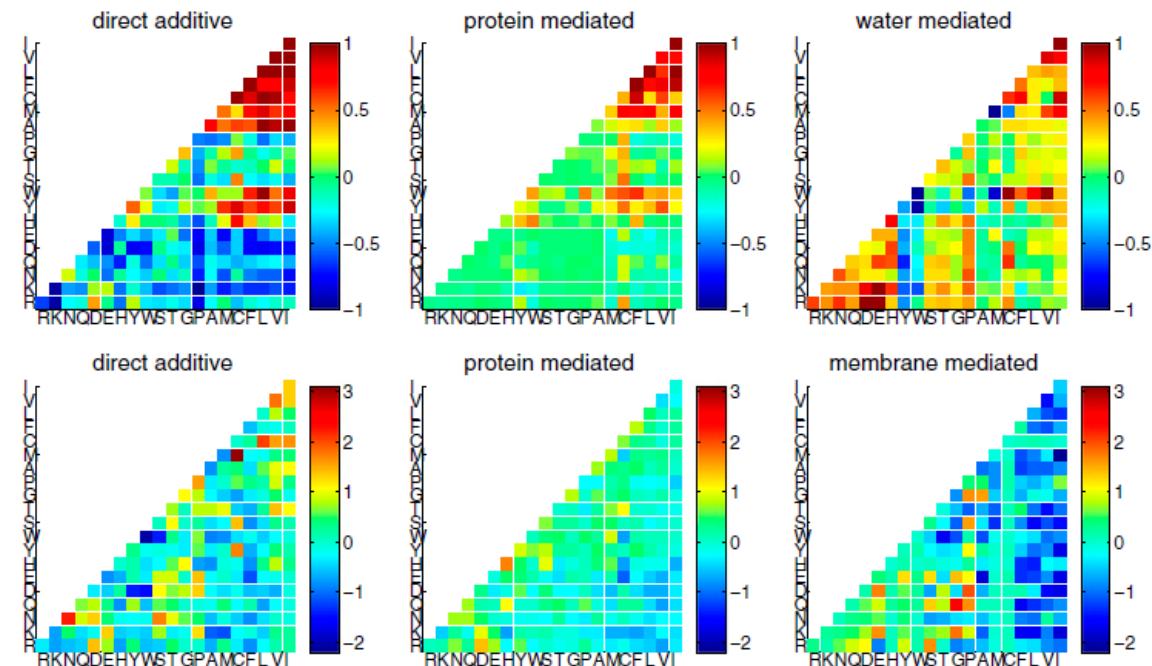
$$\Theta_{peri}(z_i) = 0.5 \left[1 + \tanh \left(k \left(z_i - \frac{\alpha}{2} \right) \right) \right]$$

$$\Theta_{cyto}(z_i) = 0.5 \left[1 + \tanh \left(k \left(-\frac{\alpha}{2} - z_i \right) \right) \right]$$

$$\Theta_{memb}(z_i) = 0.5 \left[\tanh \left(k \left(z_i + \frac{\alpha}{2} \right) \right) + \tanh \left(k \left(\frac{\alpha}{2} - z_i \right) \right) \right].$$

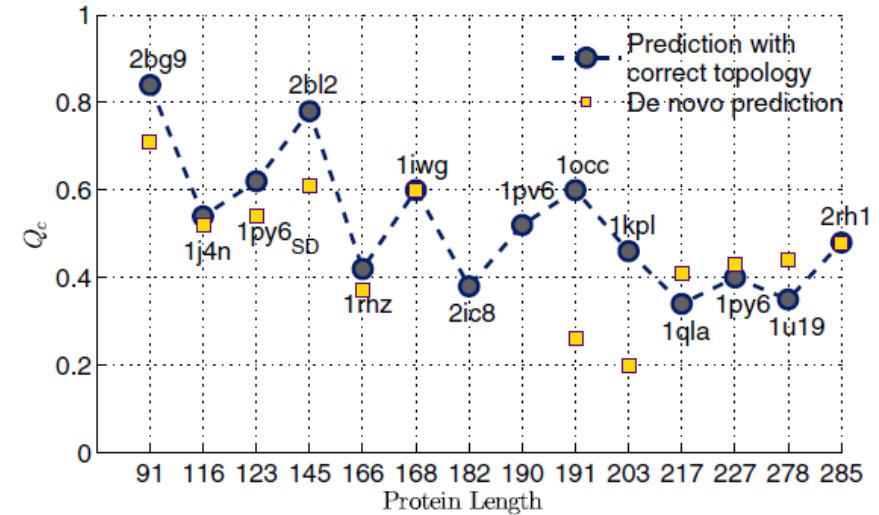
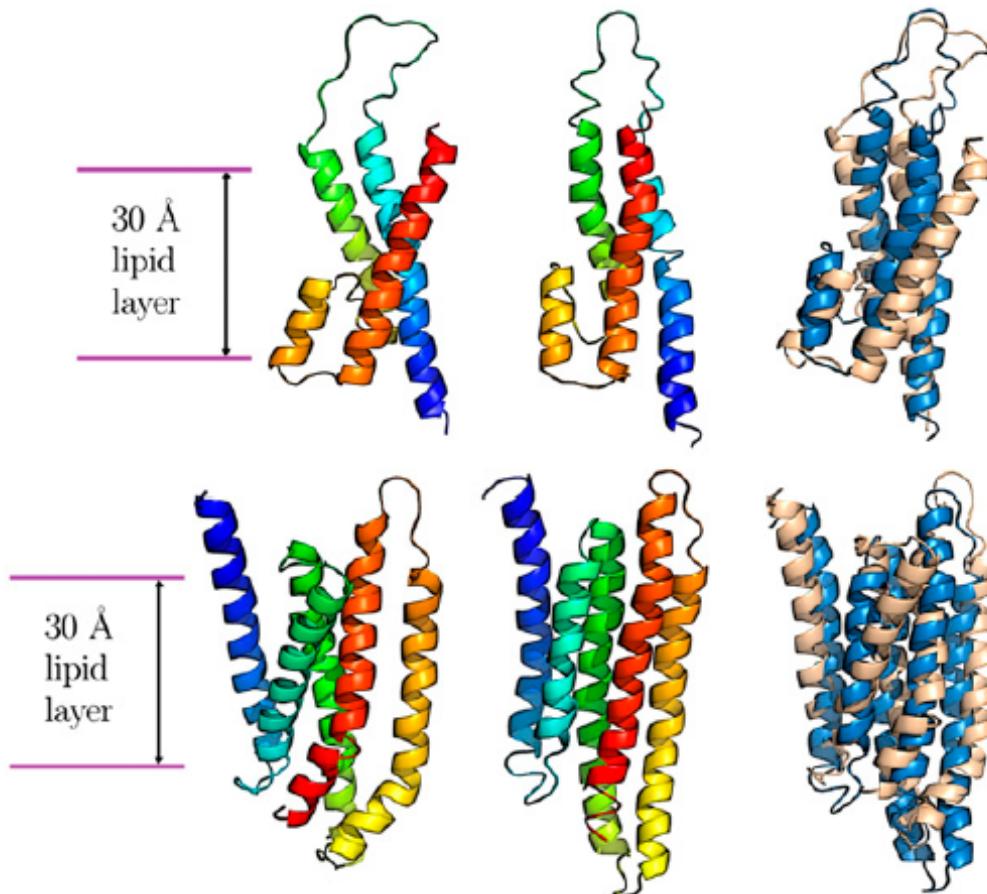
$$R_{gCyl} = \frac{1}{M} \sum_i^N m_i \left(\rho_i - \frac{1}{M} \sum_j^N m_j \rho_j \right)^2$$

$$V_{ReCyl} = K(R_{gCyl} - R_{\alpha Cyl})^2$$



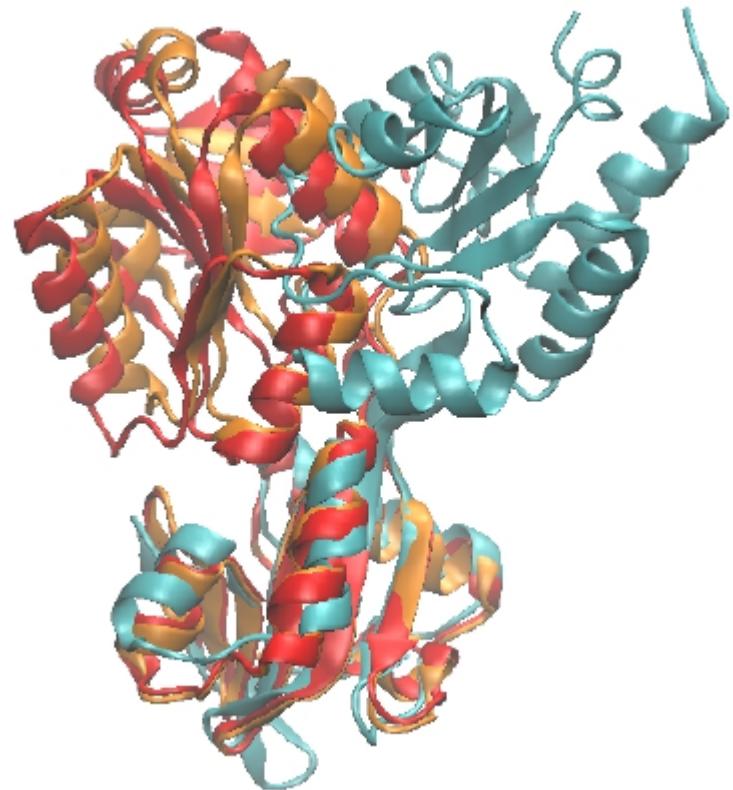
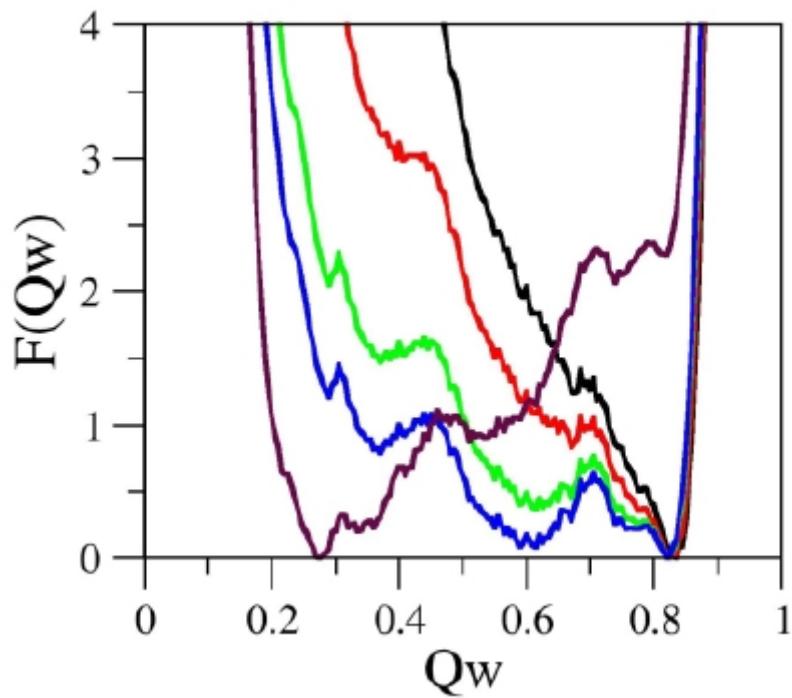
Predictive energy landscapes for folding α -helical transmembrane proteins

Membrane protein folding starts from a quasi-disorganized but good-in-topology membrane structure

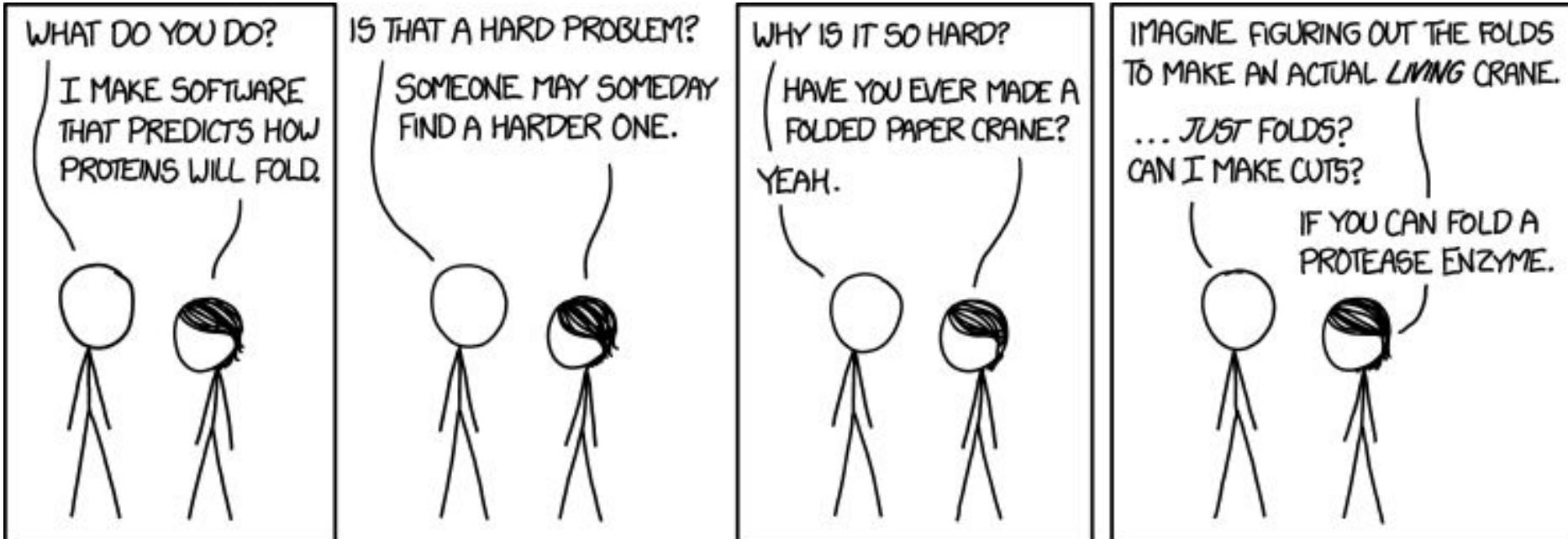


Predictive energy landscapes for folding α -helical transmembrane proteins

Structure and energy properties of the native ensemble has also been found for some proteins

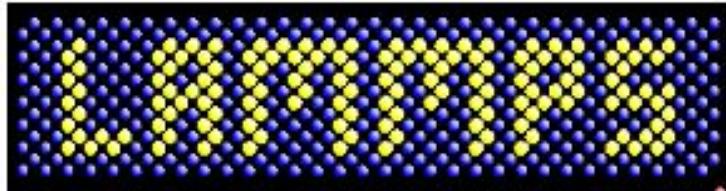


LAMMPS and AWSEM



What's LAMMPS?

Large-scale Atomic/Molecular Massively Parallel Simulator



Molecular dynamics code
written in the **C++ language**

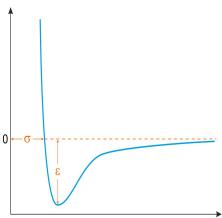
- **Open source**
- **Portable**
- **Documentation**
- **Active user Community**



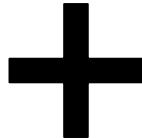
Molecular dynamics code in one slide

Positions and Velocities at time t

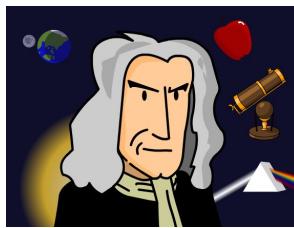
Interaction Potential



Forces

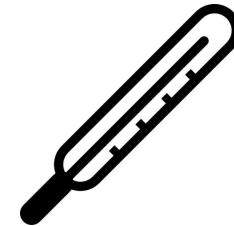


Verlet integration



$$\mathbf{F} = m \cdot \mathbf{a}$$

Nosé - Hoover



Thermostat

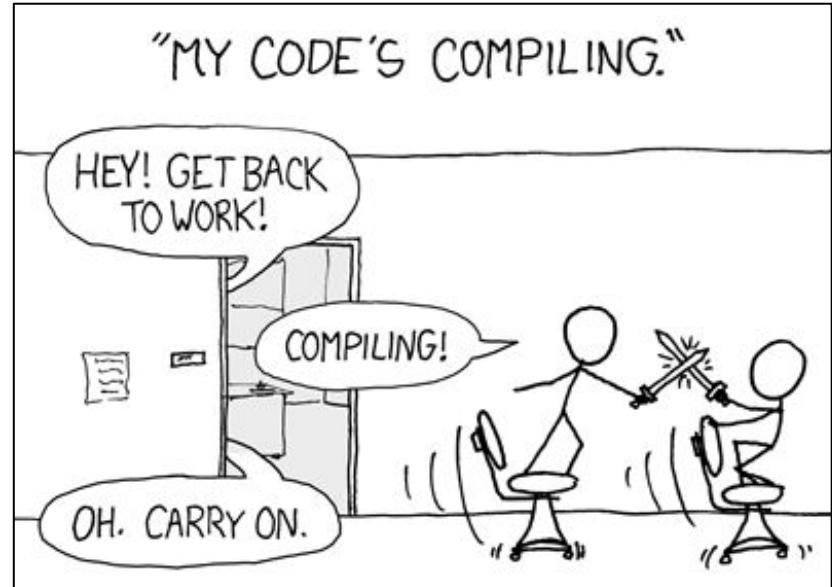
Positions and Velocities at time t + dt

Portability, Compiling, Errors, Chaos

LAMMPS is made up with a lot of C++ **text files** that contain **instructions**

C++ is a **Compiled Language**: Your **instructions** are optimized and translated into **machine language**. This takes time and is called **Compiling Time**. Usually the compilation is made with the program **make**

After this process you obtain an **Executable Program** that receives LAMMPS **input scripts** and runs your simulation (**Execution time**)



If you want to change an instruction you have to compile before executing, so compiling time is important

Backbone of a LAMMPS input script

Units and simulation characteristics



```
# LAMMPS script for 3d protein simulation
units real
timestep 5
dimension 3
boundary s s s
neighbor 5 bin
neigh_modify delay 5
special_bonds fene
atom_style peptide
bond_style harmonic
pair_style vexclude 2 3.5 3.5
read_data data.1r69
... (continues)
```

Atom style and Interaction style



Set initial configuration



LAMMPS input script is read and executed line by line

Backbone of a LAMMPS input script

Attributes of atoms



... (continues)

```
velocity all create 300.0 2349852
```

Groups of atoms



```
{ group      alpha_carbons id 1 4 ... 184 187  
group      beta_atoms    id 3 6 ... 186 189  
group      oxygens      id 2 5 ... 185 188
```

Calculations!

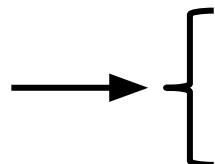


```
{ fix       1 all nvt temp 300.0 300.0 100.0  
fix       2 alpha_carbons backbone  
beta_atoms oxygens fix_backbone_coeff.data  
1r69.seq
```

... (continues)

Backbone of a LAMMPS input script

Output of simulation
quantities



```
... (continues)
thermo_style custom step temp epair emol ke
pe etotal press vol
thermo      1000
```

Write positions



```
dump      1 all atom 100 dump.lammpstrj
dump_modify 1 sort id
```

Generate **Restart** file



```
restart      1000 restart.1r69
```

RUN!



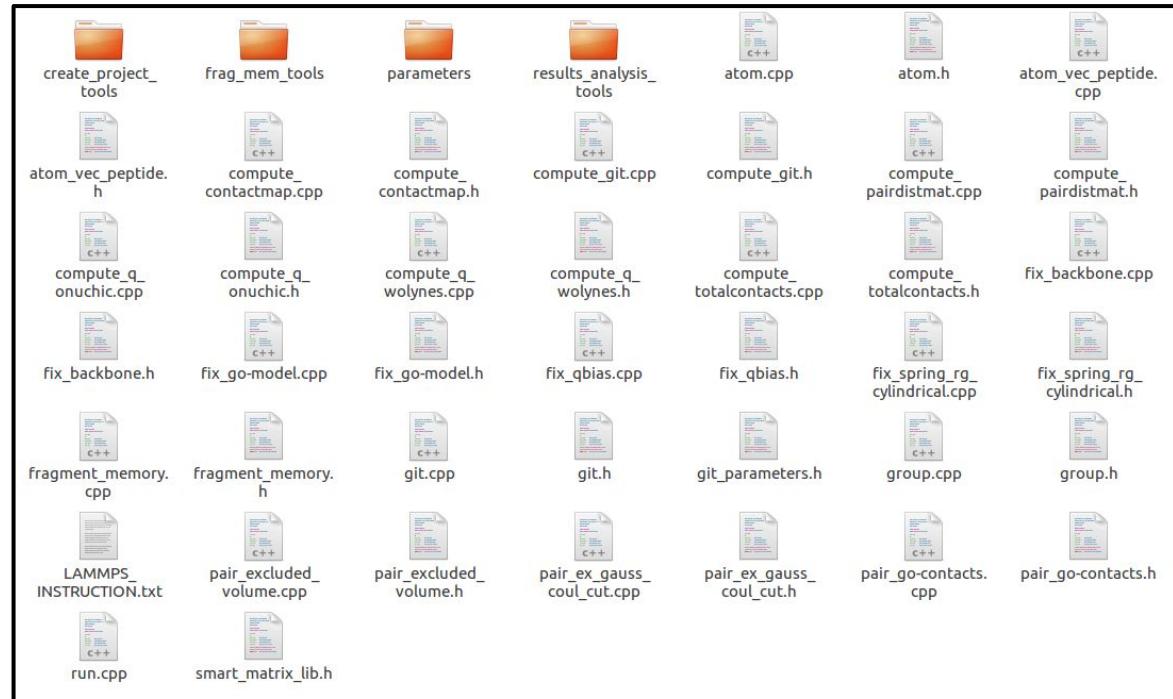
```
reset_timestep 0
run        10000
```

AWSEM and LAMMPS

The force field is written in
several C++ files

Most important file
fix_backbone.cpp

We need **more files** that
just the input script



AWSEM parameters files

File

fix_backbone_coeff.data

In this file we configure
the values of the force
field

Dash after term name

Turns off interaction

```
[Chain]  
10.0 10.0 30.0  
2.45798 2.50665 2.44973  
  
[Chi]  
20.0 -0.83  
  
[Epsilon]  
2.0  
  
[Rama]  
2.0  
  
...  
2.0 15.398 1.0 2.25 1.0 -2.16  
  
[Rama_P]  
3  
0.0 0.0 1.0 0.0 1.0 0.0  
  
...  
0.0 0.0 1.0 0.0 2.0 0.0  
  
[SSWeight]  
0 0 0 1 1 0  
0 0 0 0 0 0
```

```
[P_AP]  
0.6 0.6 0.6  
...  
4  
  
[Water]  
1.0  
...  
6.5 9.5 1  
[Helix]- ← Helix is off  
1.5  
...  
  
[Burial]  
1.0  
...  
6.0 9.0  
[Fragment_Memory]  
0.1  
1r69.mem  
1r69.gamma
```

AWSEM parameters files

Extra **parameters files**
for each interaction

- anti_HB
- anti_NHB
- anti_one
- burial_gama
- gamma
- para_HB
- para_one
- memory.gamma

Have to be copied to
simulation folder!

Where to start simulation?

- File with .data extension

Information about protein files

- Files with .seq .coord

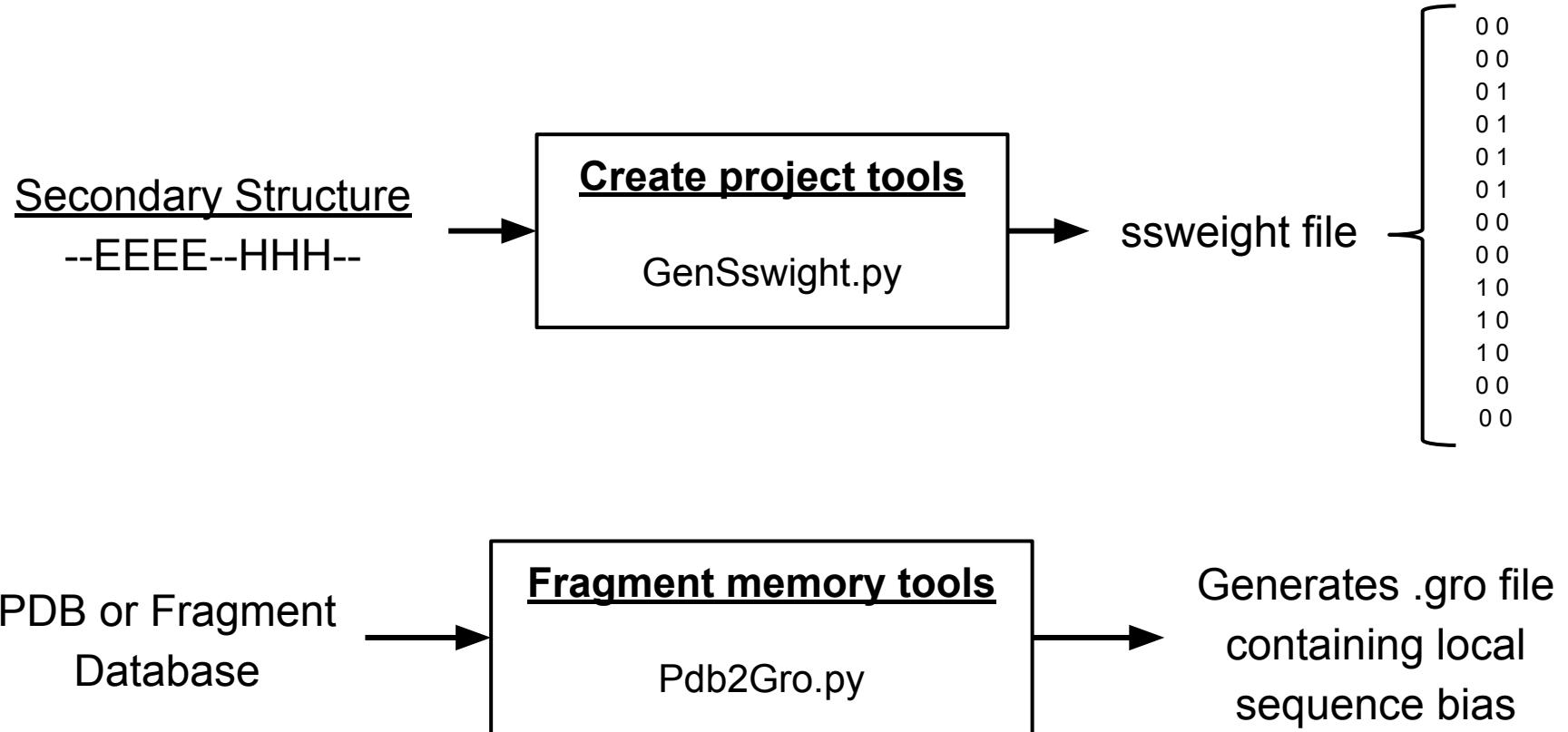
These are **generated** with
scripts from a **PDB file**

Also a lammps **input script** is
created

Create project tools

PdbCoords2Lammps.sh

Single memory run and Ssweight



Compute extra quantities during simulation

We need to generate
reference states:

- rnative.dat
- nativecoords.dat

Create project tools

GetCACADistancesFile.py
GetCACoordinatesFromPDB.py

For example, we can
compute:

- Q wolynes
- Q onuchic
- Total contacts
- Radius of Gyration

We add this calculation to the
input script

```
compute    qw alpha_carbons qwolynes rnative.dat 2 0.15
variable   qw equal c_qw
compute    rg alpha_carbons gyration
variable   rg equal c_rg
compute    qo alpha_carbons qonuchic cutoff 12.0 nativecoords.dat 1.2
variable   qo equal c_qo
compute    tc beta_atoms totalcontacts 6.5 2
variable   tc equal c_tc
variable   S_T equal step
variable   E_P equal pe
variable   T_E equal temp
```

Output our quantities to a file or screen

We can calculate different quantities

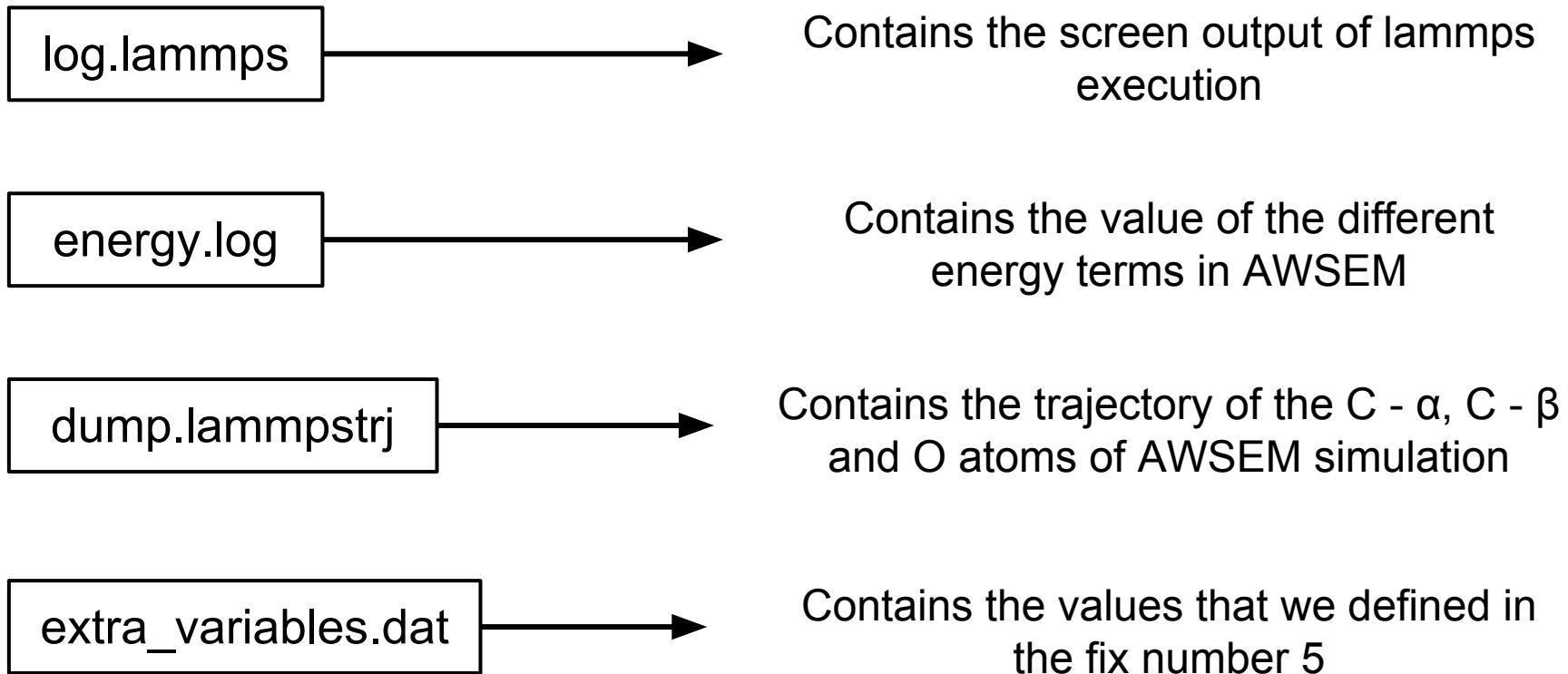
Some of them need a **compute** command

Others are calculated by the **thermo** command

```
compute      qw alpha_carbons qwolynes rnative.dat 2 0.15
variable     qw equal c_qw
compute      rg alpha_carbons gyration
variable     rg equal c_rg
compute      qo alpha_carbons qonuchic cutoff 12.0 nativecoords.dat 1.2
variable     qo equal c_qo
compute      tc beta_atoms totalcontacts 6.5 2
variable     tc equal c_tc
variable     S_T equal step
variable     E_P equal pe
variable     T_E equal temp
```

```
# Print extra quantities on screen
fix          4 all print 1000 "{$S_T} {$T_E} {$qw} {$rg} {$qo} {$tc} {$E_P}" screen yes
# Write extra quantities to a file
fix          5 all print 1000 "{$S_T} {$T_E} {$qw} {$rg} {$qo} {$tc} {$E_P}" file extra_variables.dat screen no
title "# Tstep Temp Qw Rg Qo Tc Energy"
```

What do we get after the calculations?



Data processing and visualization

