# Folding kinetics and dynamics – September 17, 2015

*EG*

Home (edit)
Research: Calculations (edit)
Self (edit)

## 1 2D

### 1.1 Stability

We calculate average and median number of contacts in native sequences, according to HP-model. It is measured in the units of hydrophobic energy per $kT$. From [1] it follows that

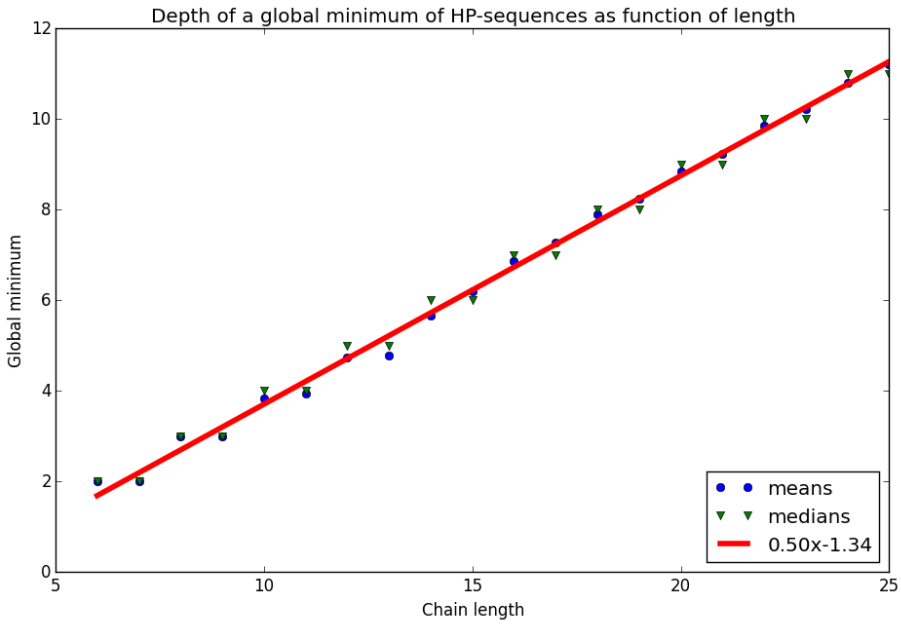


Fig. 1: Depth of the energy well is measured in the units of $e_{h\phi} = E_{h\phi}/kT$, where $e_{h\phi}$ is an energy of one hydrophobic contact per $1kT$.

free energy is composed of two terms: the enthalpy and entropy of the molecular interactions

and $\Delta G_0 = \Delta H_0 - T\Delta S_0$ and conformational entropy $\Delta G_c = -T\Delta S_c = kTNlnz$. $N$ here is a chain length and $z$ is a number of rotational freedoms.

Fitting a line on a figure 1 to the HP-model minimums yields the following dependency for the energy well depth (per $1kT$):

$$\Delta G_0 = -kNe_{h\phi} - be_{h\phi} = -0.5Ne_{h\phi} + 1.34e_{h\phi} \tag{1}$$

where $N$ is a chain length and $e_{h\phi}$ is an absolute value of the energy of one hydrophobic contact. And full free energy of folding will be (per $1kT$):

$$\Delta G = -kNe_{h\phi} - be_{h\phi} + N \ln z \tag{2}$$

## 1.2 Kinetcis

There's a theory of protein folding introduced by Ghosh and Dill in a series of papers [2, 3] and more. It is based on simple equilibrium thermodynamics and leads to the following conclusion regarding chain length distribution:

$$\ln k_f = 16.15 - 1.28\sqrt{N} \tag{3}$$

Which in turn predicts the following dependency of the rate constants: While it is likely that

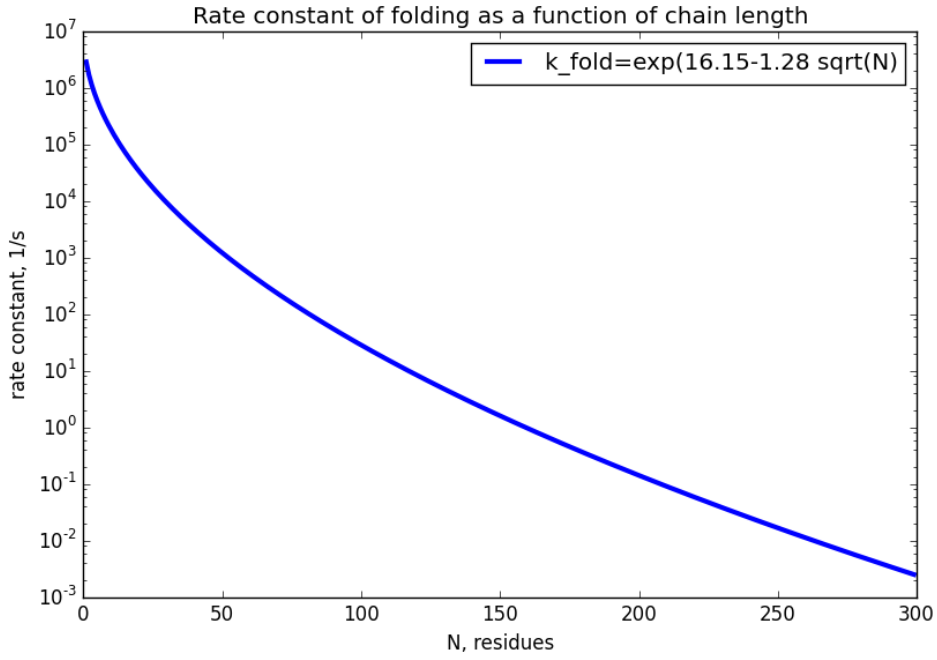


Fig. 2:

model fails at the short sequences, we can use it for a crude estimation of the folding and unfolding rates of the short HP-polymers. Though there's no real way to test if our results are true. This model suggests that the *folding* rate constants would be the following for some short chains:

- $k_f(10) = 2 \cdot 10^5$
- $k_f(20) = 3 \cdot 10^4$
- $k_f(30) = 0.9 \cdot 10^4$

Folding and unfolding rate constants relate to each other and free energy of folding.

$$\ln\left(\frac{k_f}{k_u}\right) = \ln K = -\Delta G/kT \tag{4}$$

From the section 1.1, if $\Delta G_0$ follows law from fig.1 it follows that:

$$\ln k_f - \ln k_u = \Delta G_0/kT - N \ln z \tag{5}$$

which gives

$$\ln k_u = \ln k_f - \Delta G_0/kT + N \ln z \tag{6}$$

Substituting $\ln k_f$ from equation (3) and fitted $\Delta G_0$ we get:

$$\ln k_u = 16.15 - 1.28\sqrt{N} - kNe_{h\phi} - be_{h\phi} + N \ln z \tag{7}$$

Therefore

$$k_u = \frac{1}{z^N} \exp(16.15 - 1.28\sqrt{N} - kNe_{h\phi} - be_{h\phi}) \tag{8}$$

This gives the following dependencies for $k_u$ and $k_f$: see fig.4

## 1.3 HP-model results

If we use modification of (8):

$$k_u = \exp(16.15 - 1.28\sqrt{N} - kNe_{h\phi} - be_{h\phi}) \tag{9}$$

and correspondingly $k_f = k_u \exp(\Delta G)$:

$$k_f = \frac{1}{z^N} \exp(16.15 - 1.28\sqrt{N} - kNe_{h\phi} - be_{h\phi}) \exp(n_{h\phi}E_{h\phi}) \tag{10}$$

We will get the following dependencies (it's good for $\Delta G$, but not su much for $k_f$): see fig. 5. The folding rate drops rather low. I think it affects simulation in a negative way. However, if we fix $k_u$ and then just keep $k_f$ so that $\ln(k_f/k_u) = \Delta G$, then we'll get folding rate, which grows too quickly. I used formula

$$k_u = \exp(12 - 10.1\sqrt{N} - kNe_{h\phi} - be_{h\phi}) \tag{11}$$

and correspondingly $k_f = k_u \exp(\Delta G)$ so that numbers were reasonable...

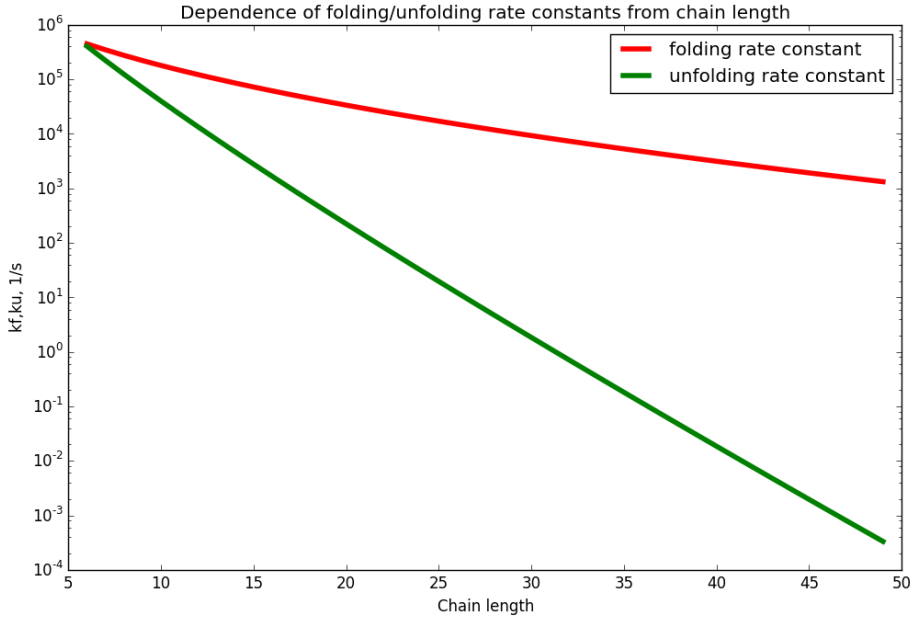


Fig. 3: Folding and unfolding rates as a function of chain length. $e_{h\phi} = 1.5, z = 1.5$ This numbers are chosen so that equilibrium constant is around $10^2 - 10^4$ when chain lengths are short (around 15-30), see figure **??**

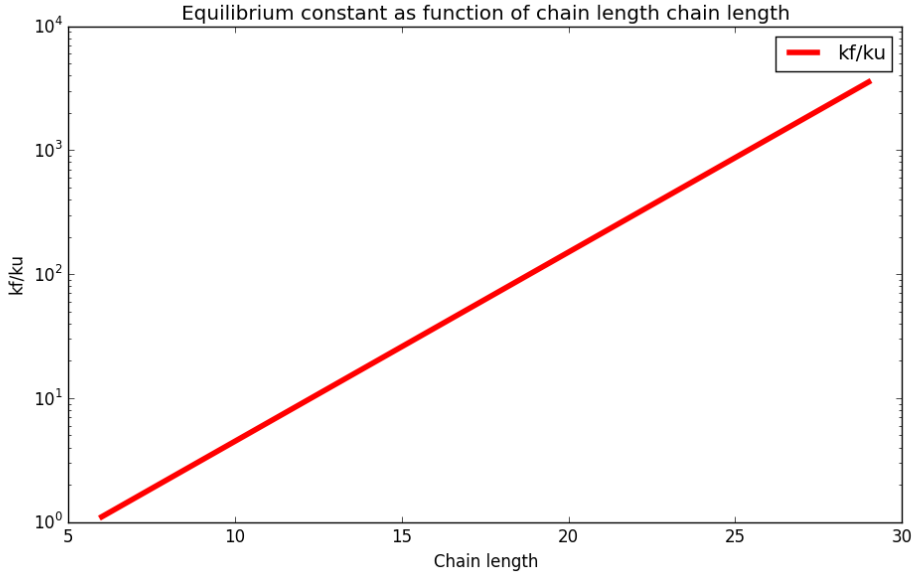


Fig. 4: Equilibrium constant as a function of chain length. $e_{h\phi} = 1.5, z = 1.5$ This numbers are chosen so that equilibrium constant is around $10^2 - 10^4$ when chain lengths are short (around 15-30)

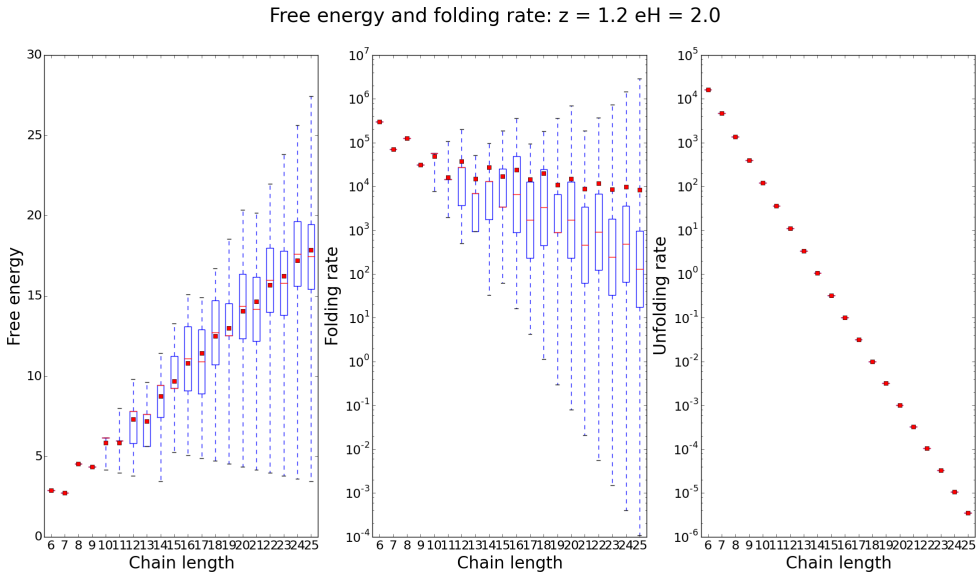


Fig. 5: Free energy, folding and unfolding rates as a function of chain length.

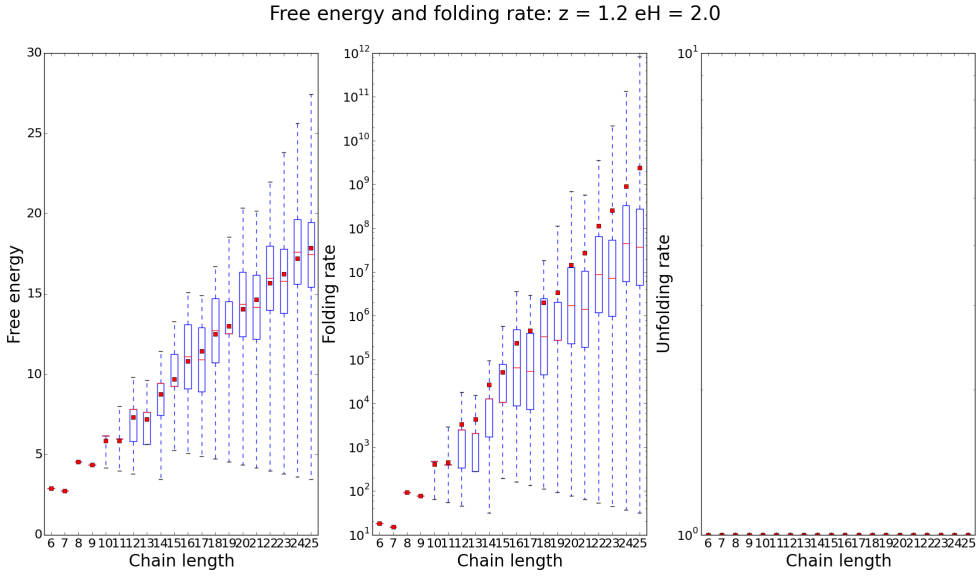


Fig. 6: Free energy, folding and unfolding rates as a function of chain length. $k_u$ is fixed: $k_f = 1.0$

## 2 3D

Let's repeat the procedure for 3-dimensional proteins. From [3] we can take a figure 1D, which has a fitted curve for $\Delta G(N)$ at 37℃. Determine parameters from it returns:

$$\Delta G(N) = 0.067N + 7.012kJ/mol \tag{12}$$

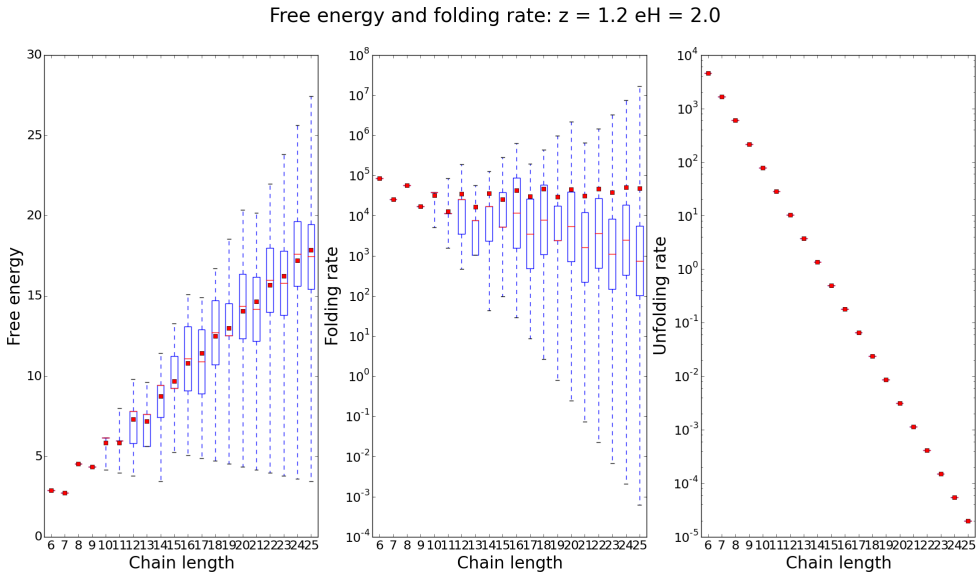


Fig. 7: Free energy, folding and unfolding rates as a function of chain length. $k_u = \exp(12 - 10.1\sqrt{N} - kNe_{h\phi} - be_{h\phi})$

Adding to it folding rate constant from (3)

$$\ln k_f = 16.15 - 1.28\sqrt{N}$$

will get

$$\ln k_f - \ln k_u = \Delta G/kT \tag{13}$$

or for $k_u$

$$\ln k_u = \ln k_f - \Delta G/kT = 16.15 - 1.28\sqrt{N} - (0.067N + 7.012)/2.5775 \tag{14}$$

so

$$k_u = \exp(13.43 - 0.026N - 1.28\sqrt{N}) \tag{15}$$

This leads to the following folding and unfolding rate constants (see fig 8)

## References

[1] Kingshuk Ghosh and Ken A Dill. Computing protein stabilities from their chain lengths. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26):10649–54, June 2009.

[2] Kingshuk Ghosh and Ken A Dill. Cellular proteomes have broad distributions of protein stability. *Biophysical Journal*, 99(12):3996–4002, 2010.

[3] Ken A Dill, Kingshuk Ghosh, and Jeremy D Schmit. Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences of the United States of America*, 108(44):17876–82, November 2011.

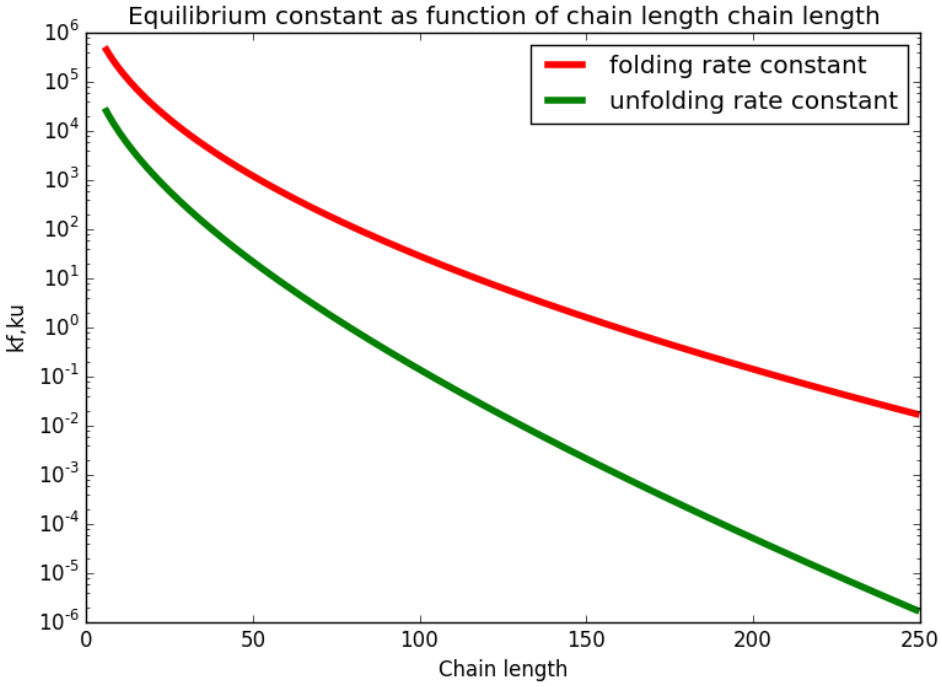


Fig. 8: Folding and unfolding rates as a function of chain length for 3D proteins. Data taken from [3]