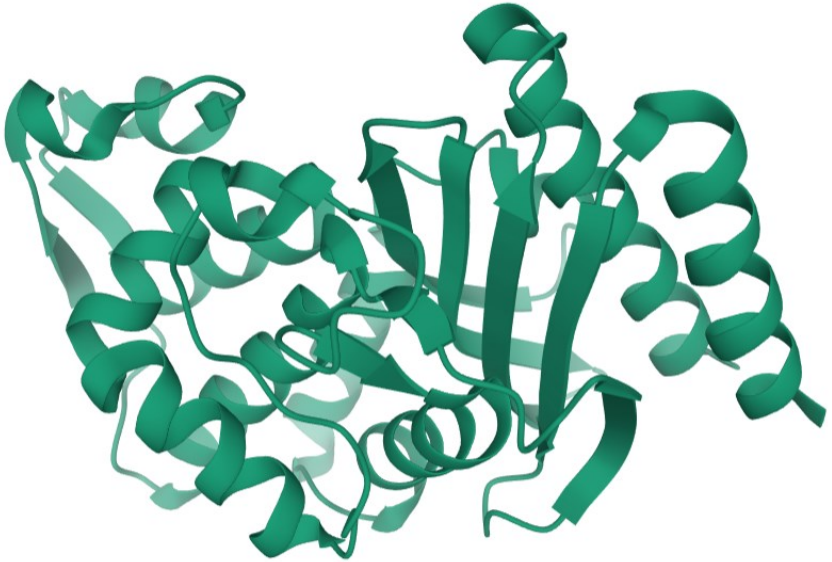
Structural biology

Practice 4: Statistical potentials and PROSA

Course 2023-2024

Practice 4: Statistical potentials and PROSA



Is this model correct?

Practice 4: Statistical potentials and PROSA

Is this model correct?

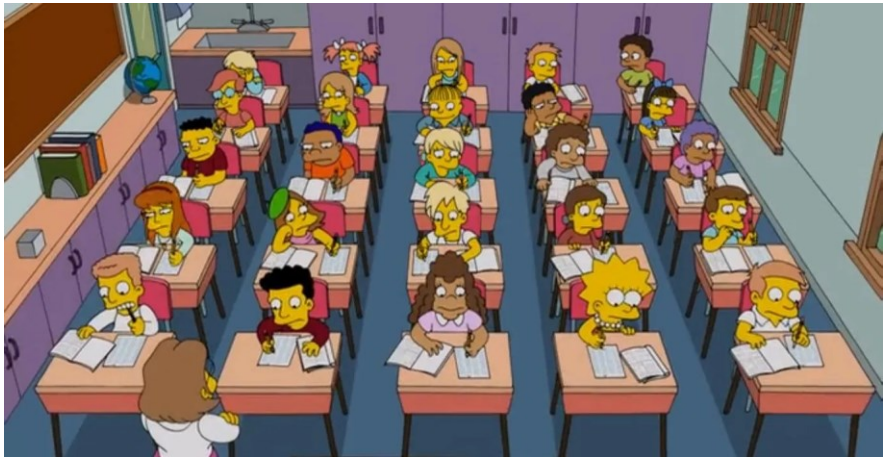
We can use statistical potentials to answer this question

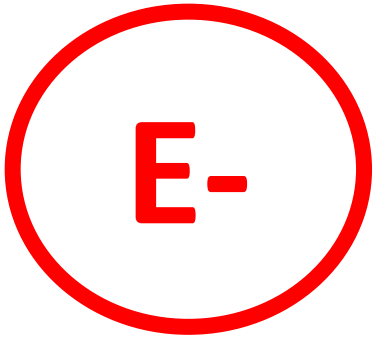
Practice 4: Statistical potentials and PROSA

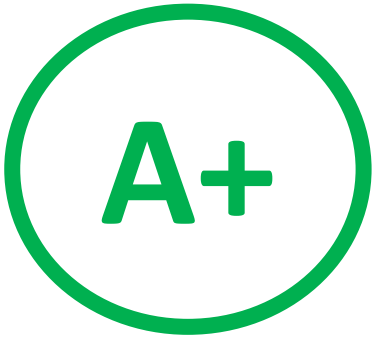
Statistical potentials are scoring functions that are derived from the analysis of experimental structures

Practice 4: Statistical potentials and PROSA

Statistical potentials are scoring functions that are derived from the analysis of experimental structures

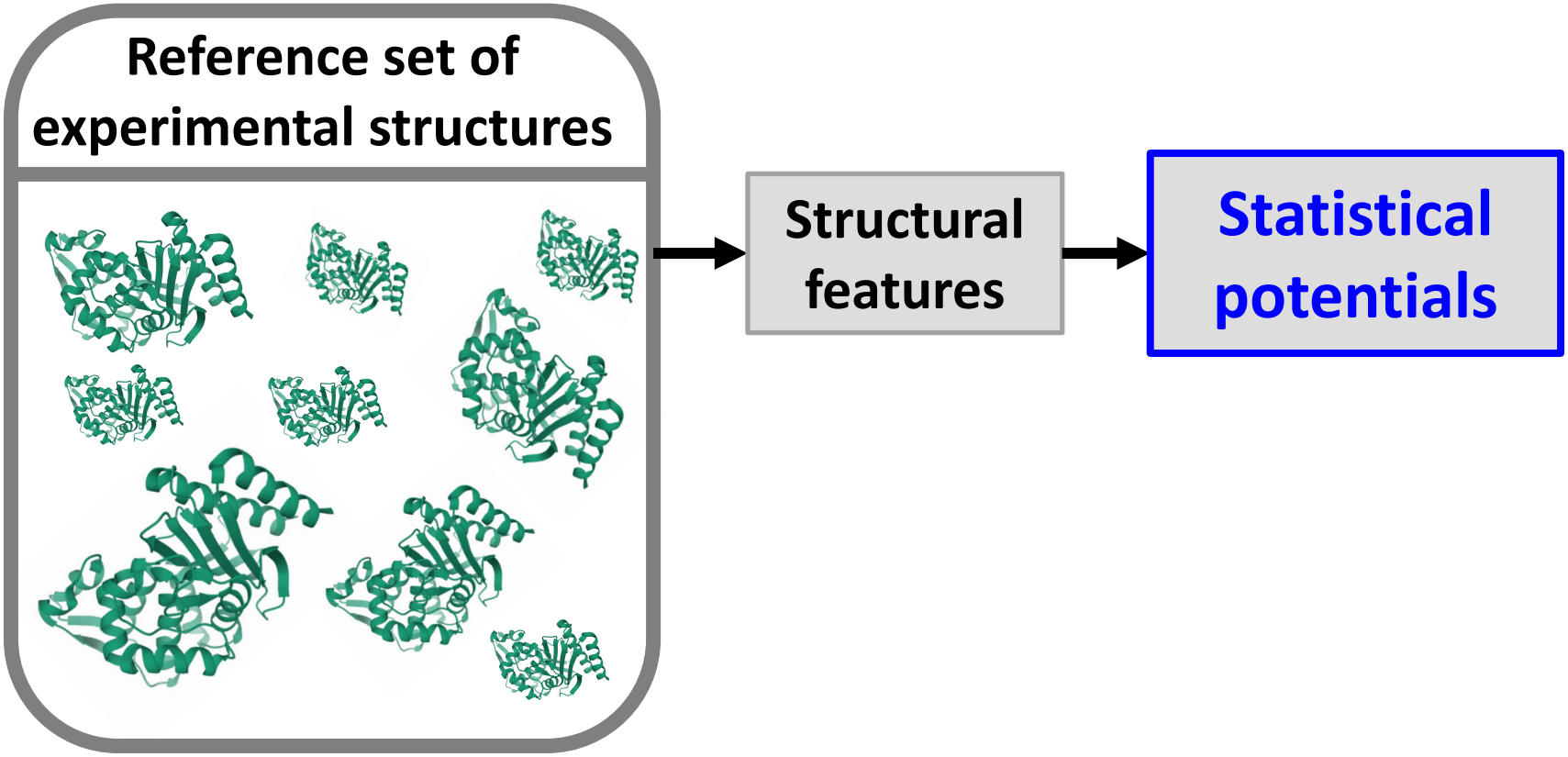




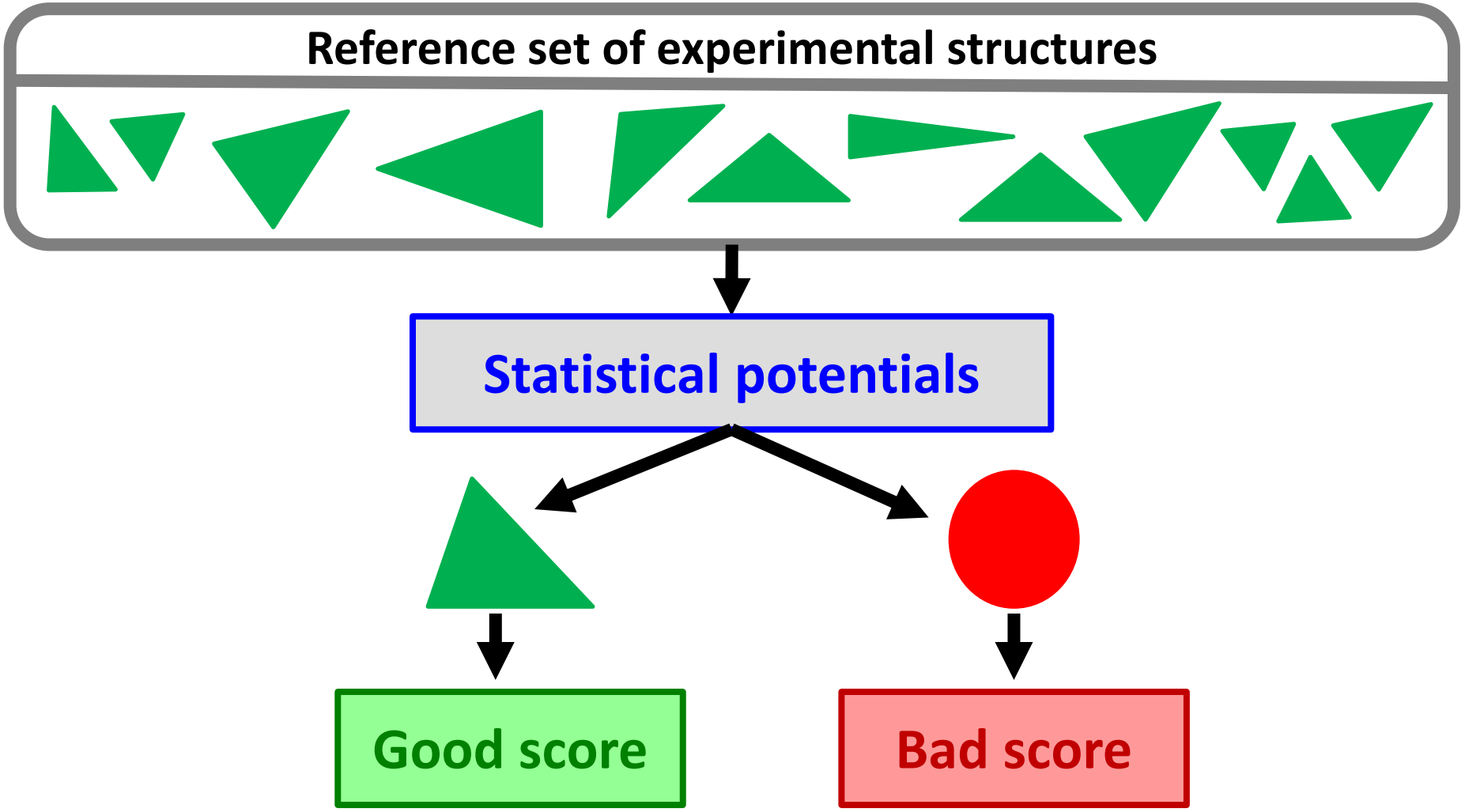


Practice 4: Statistical potentials and PROSA

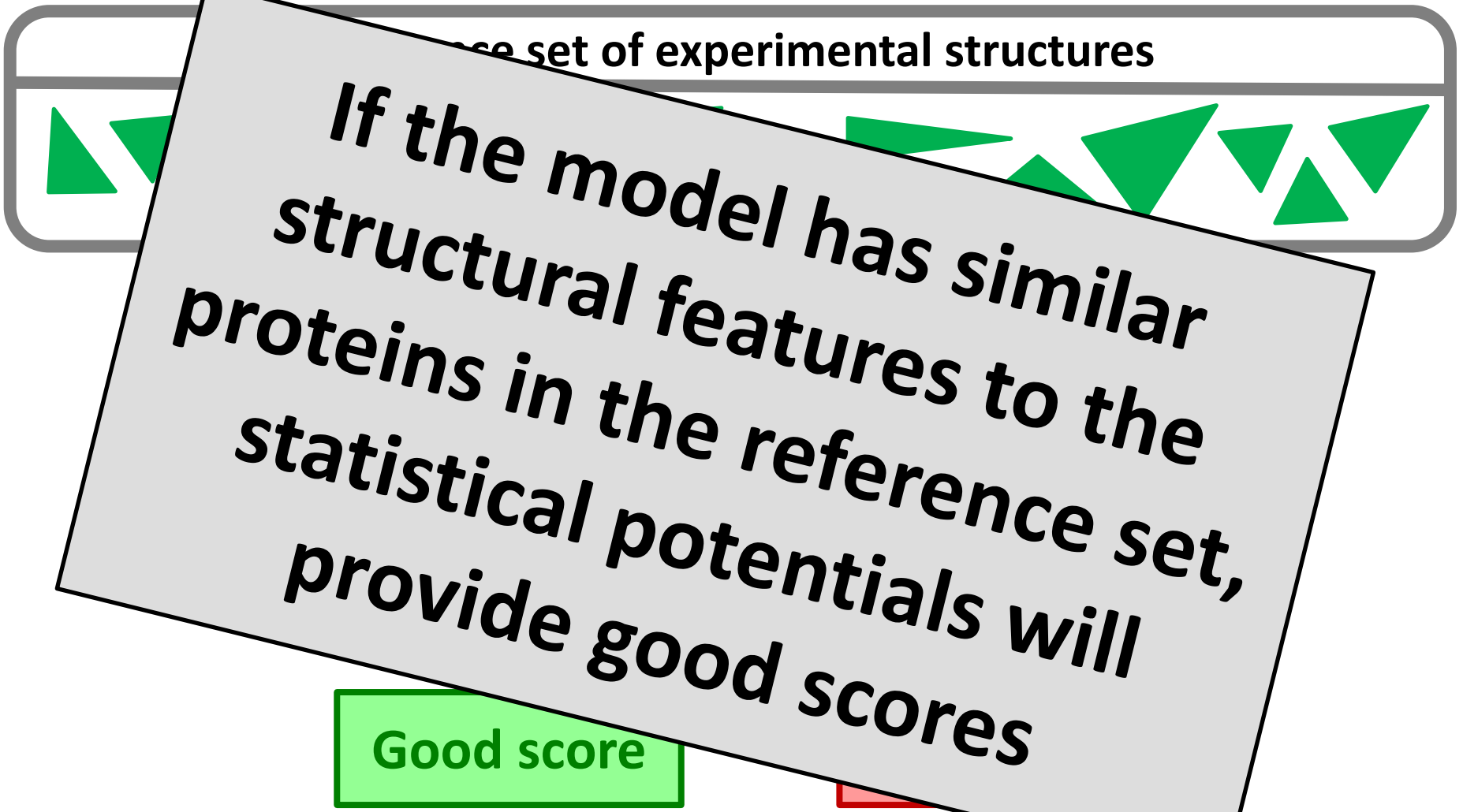
Statistical potentials are scoring functions that are derived from the analysis of experimental structures



Practice 4: Statistical potentials and PROSA



Practice 4: Statistical potentials and PROSA



Practice 4: Statistical potentials and PROSA

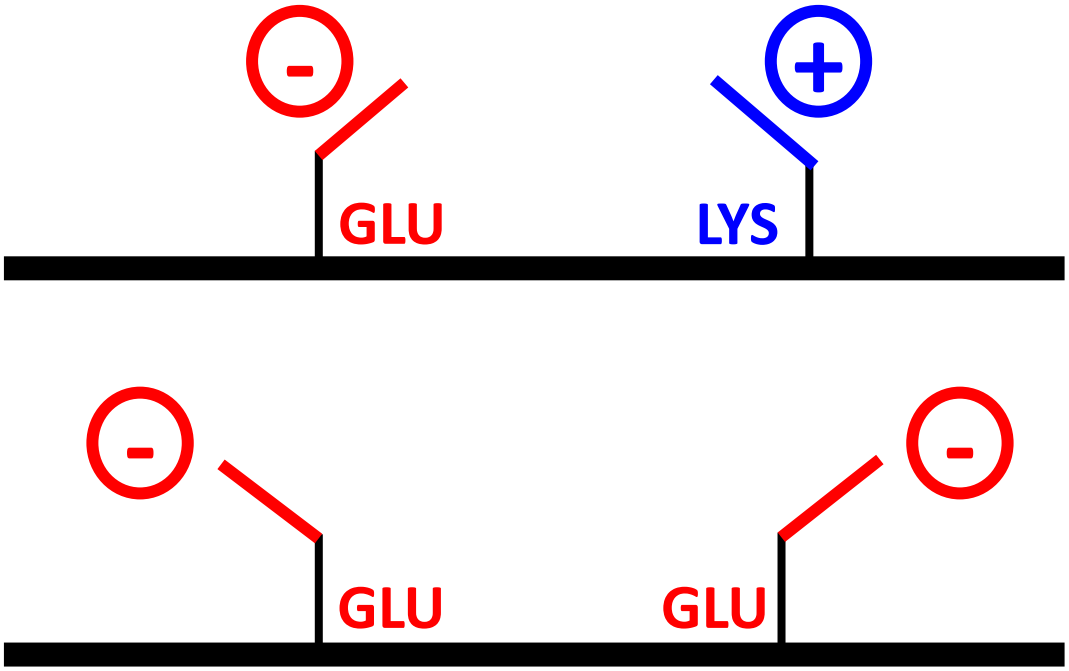
What are the structural features that statistical potentials use?

Amino acid contacts

Amino acid exposure

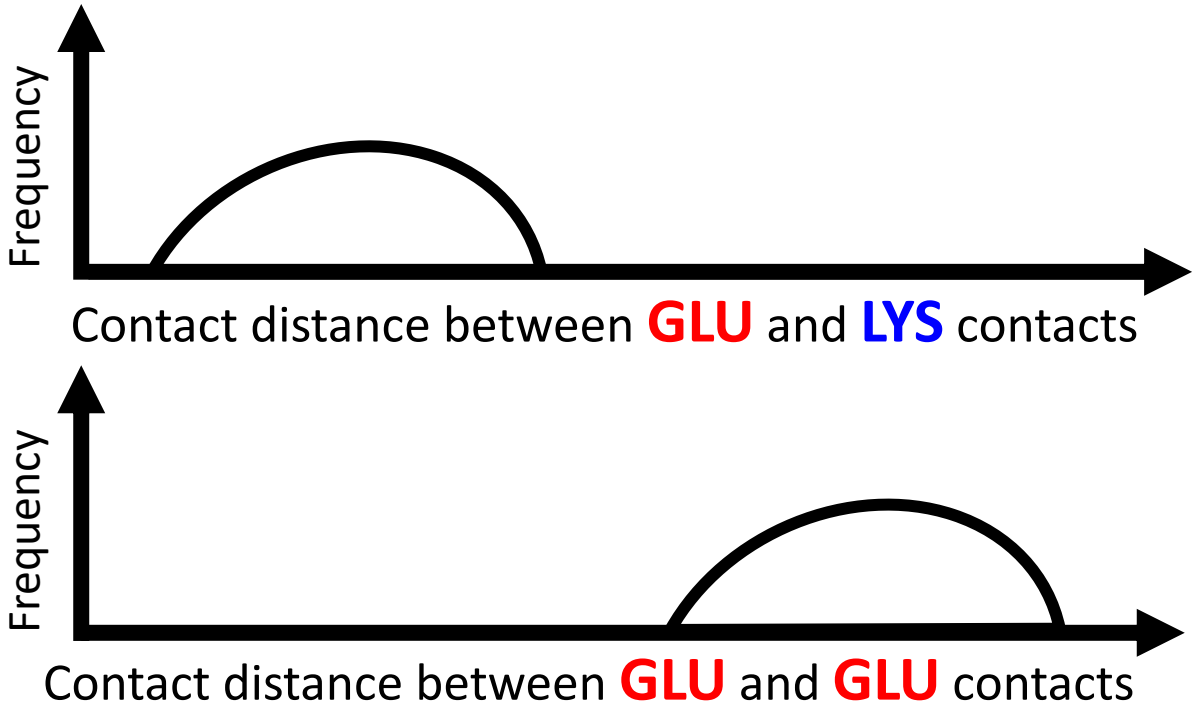
Practice 4: Statistical potentials and PROSA

Amino acid contacts



Practice 4: Statistical potentials and PROSA

Amino acid contacts



Practice 4: Statistical potentials and PROSA

Amino acid contacts

Amino acid charge is one of the factors that affect distances between pairs of amino acids

and **GLU** contacts

Practice 4: Statistical potentials and PROSA

What atom would you use to meaure distances between amino acids?

Practice 4: Statistical potentials and PROSA

What atom would you use to meaure distances between amino acids?

Beta carbon
(first carbon in the side chain)

This is a way of including information regarding side chain orientation into the potentials

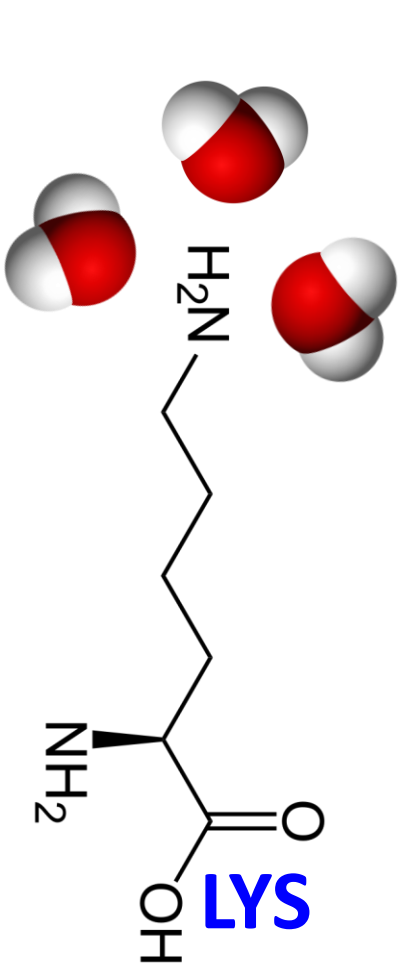
Practice 4: Statistical potentials and PROSA

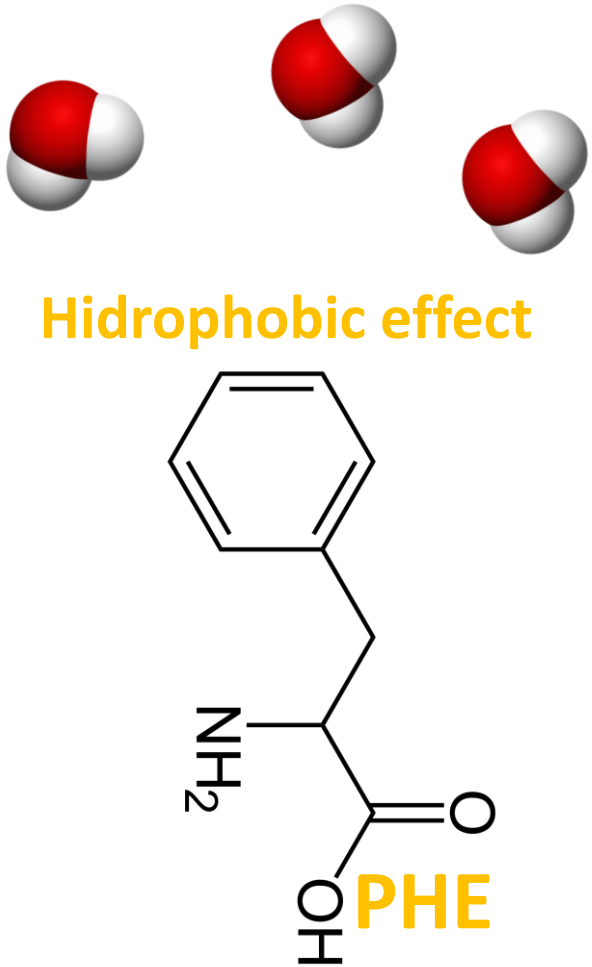
Amino acid exposure

Polar and charged are more likely to be exposed because of their tendency to interact with water molecules

Practice 4: Statistical potentials and PROSA

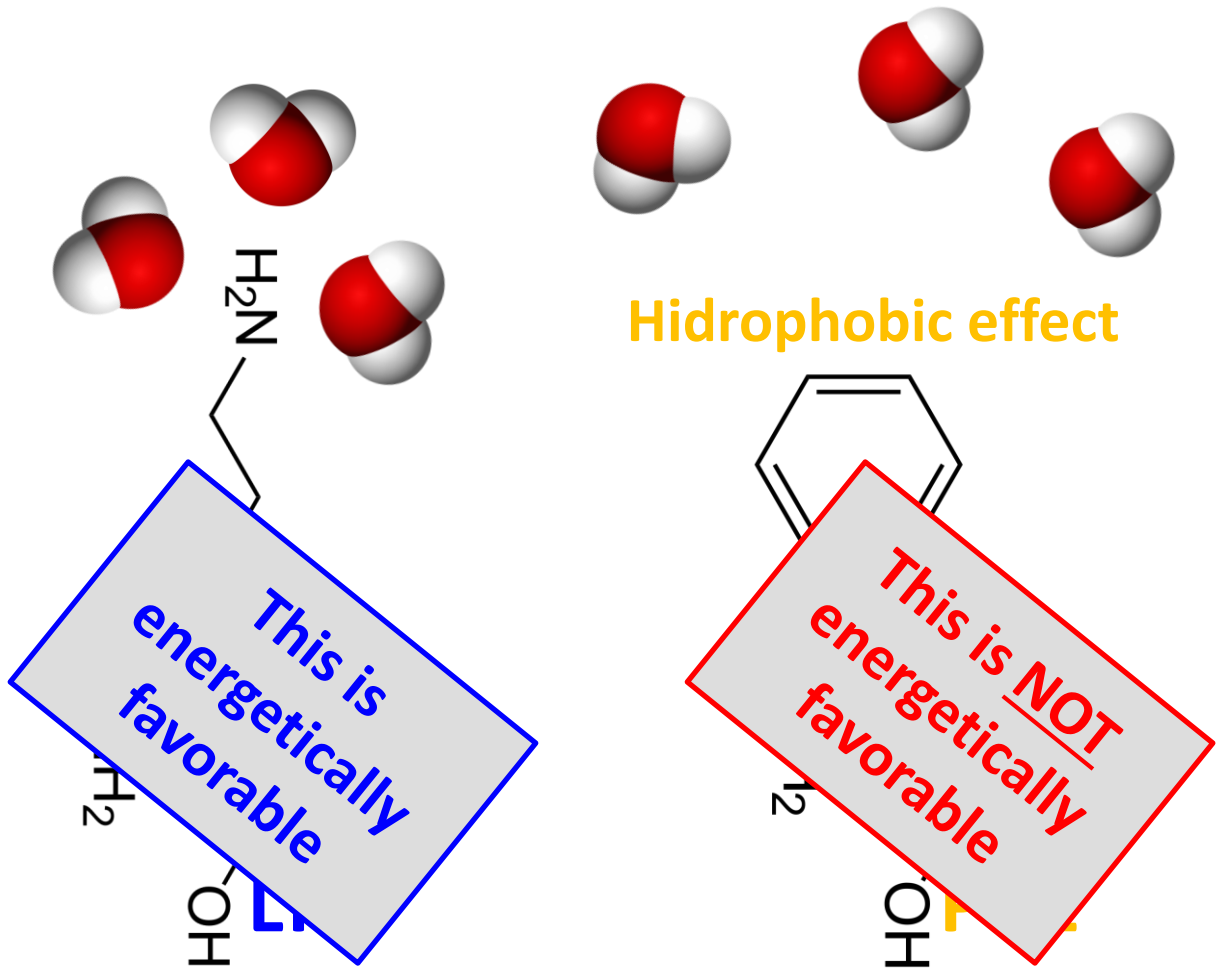
Amino acid exposure





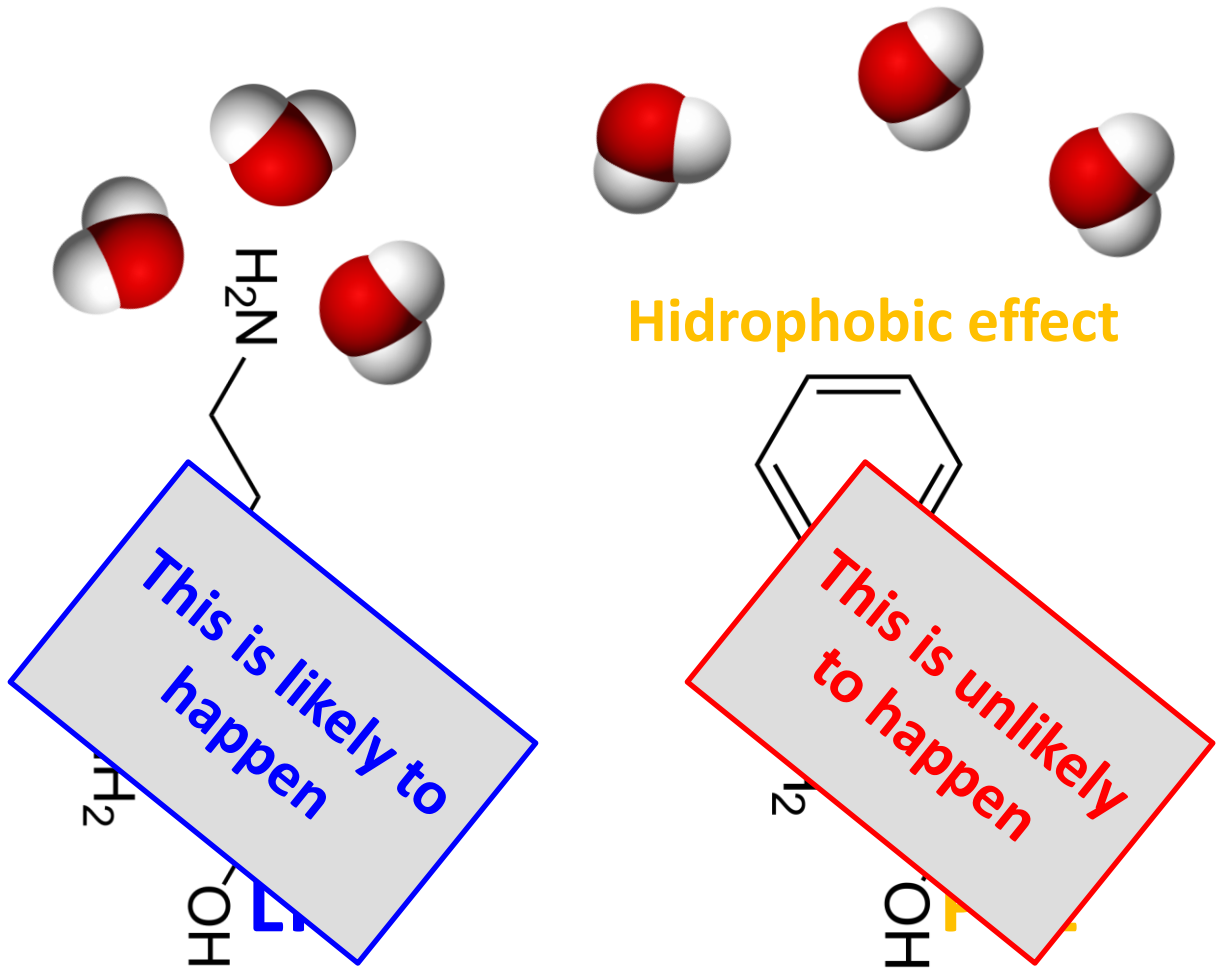
Practice 4: Statistical potentials and PROSA

Amino acid exposure



Practice 4: Statistical potentials and PROSA

Amino acid exposure



Practice 4: Statistical potentials and PROSA

Statistical potentials are computed using formulas coming from statistical thermodynamics

Boltzmann Law

$$P=(1/z)(e^{(-E/kT)})$$

Practice 4: Statistical potentials and PROSA

Statistical potentials are computed using formulas coming from statistical thermodynamics

Boltzmann Law

$$P = (1/z)(e^{(-E/kT)})$$

Probability

Energy

Practice 4: Statistical potentials and PROSA

Statistical potentials are computed using formulas coming from statistical thermodynamics

Boltzmann Law

$$P=(1/z)(e^{(-E/kT)})$$

High Probability	↔	Low Energy
Low Probability	↔	High Energy

Practice 4: Statistical potentials and PROSA

Statistical potentials are computed using formulas coming from statistical thermodynamics

Boltzmann Law

$$P = (1/z)(e^{(-E/kT)})$$

$$E = -KT\ln P + KT\ln Z$$

By operating with Boltzmann Law's equation we can isolate the energy

Practice 4: Statistical potentials and PROSA

St potentials are computed using formulas
statistical thermodynamics

Remember that probabilities come from our reference set of structures

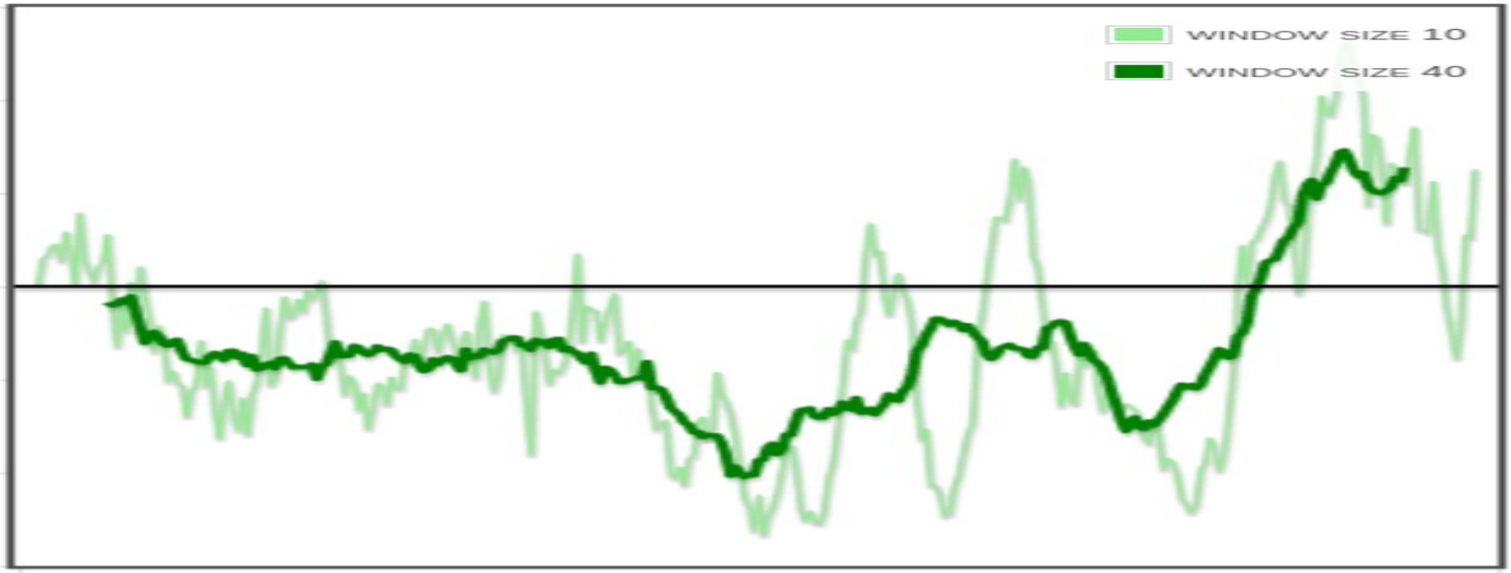
equation we

Practice 4: Statistical potentials and PROSA

We obtain one energy value for each amino acid and display it as a energy profile

$$P = (1/z)(e^{(-E/kT)})$$

$$E = -KT\ln P + KT\ln Z$$



Practice 4: Statistical potentials and PROSA

Statistical potentials are relative measurements

We use them to compare structures between them

Practice 4: Statistical potentials and PROSA

If you want to test the quality of your model, what structure would you compare with using statistical potentials?

The template you used in the modeling

Practice 4: Statistical potentials and PROSA

If you want to test the quality of your model, what structure would you compare with using statistical potentials?

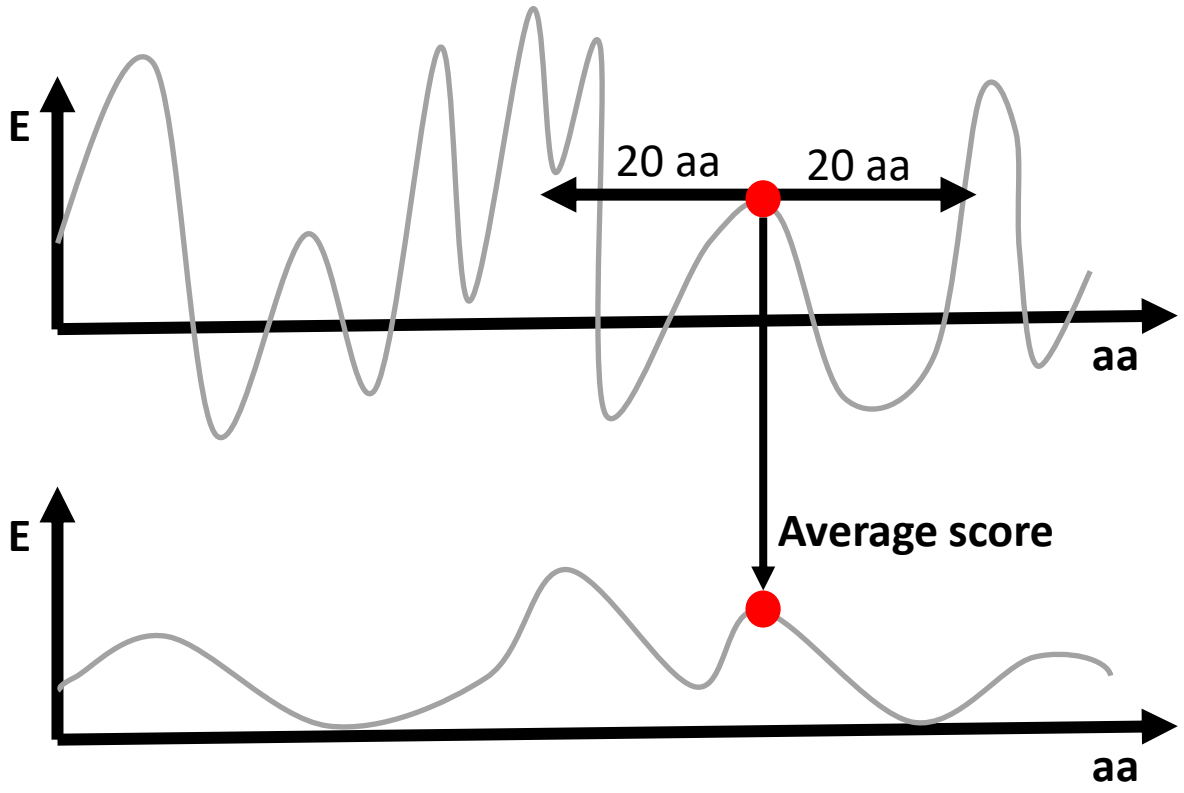
The template you used in the modeling:

- Is an experimental structure, therefore is a reference of what is right
- Is a similar protein to the one you modeled

Practice 4: Statistical potentials and PROSA

Using sliding windows

Sliding window = 40



Practice 4: Statistical potentials and PROSA

Working with Z-scores

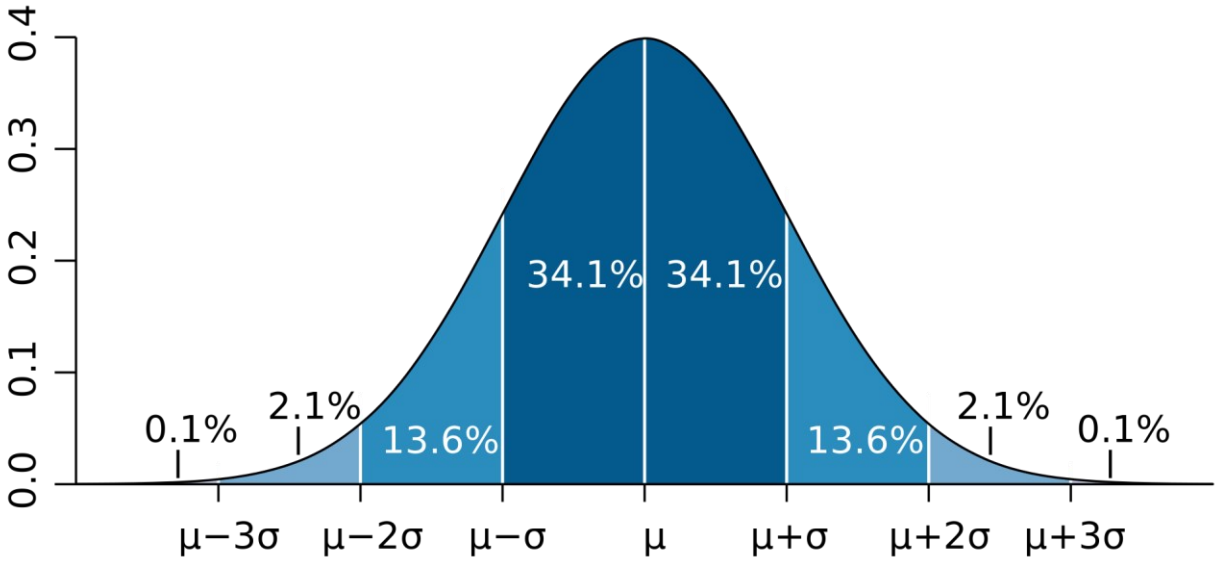
$$Z = \frac{(X - \ddot{X})}{SD}$$

How good am I in comparison with a reference distribution?

Practice 4: Statistical potentials and PROSA

Working with Z-scores

$$Z = \frac{(X - \ddot{X})}{SD}$$



Practice 4: Statistical potentials and PROSA

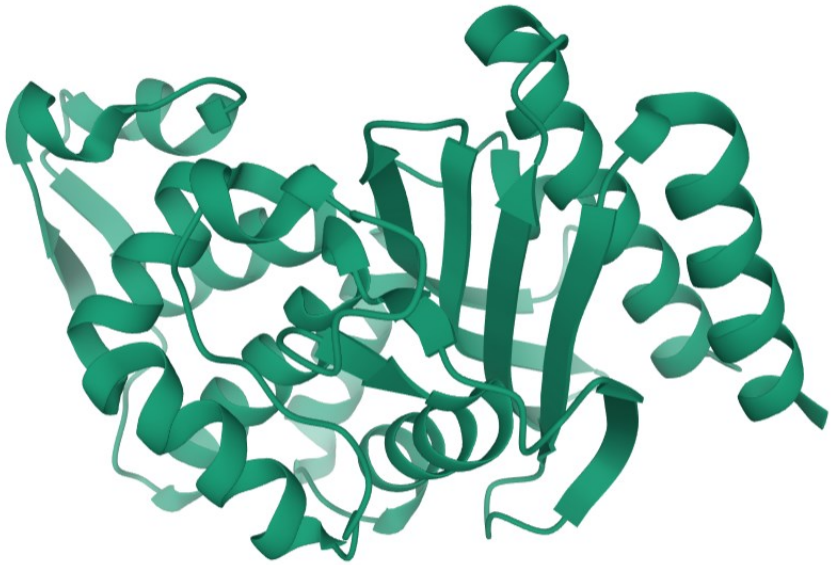
You created several models, how would you use PROSA to know what model is the best?

Practice 4: Statistical potentials and PROSA

You created several models, how would you use PROSA to know what model is the best?

You can compare the models by themselves and choose the one with best scores, you don't need to compare to a template.

Practice 4: Statistical potentials and PROSA



Is this model correct?

Practice 4: Statistical potentials and PROSA

Is this model correct?

What does correct mean?

It depends on what question you want to answer with your model

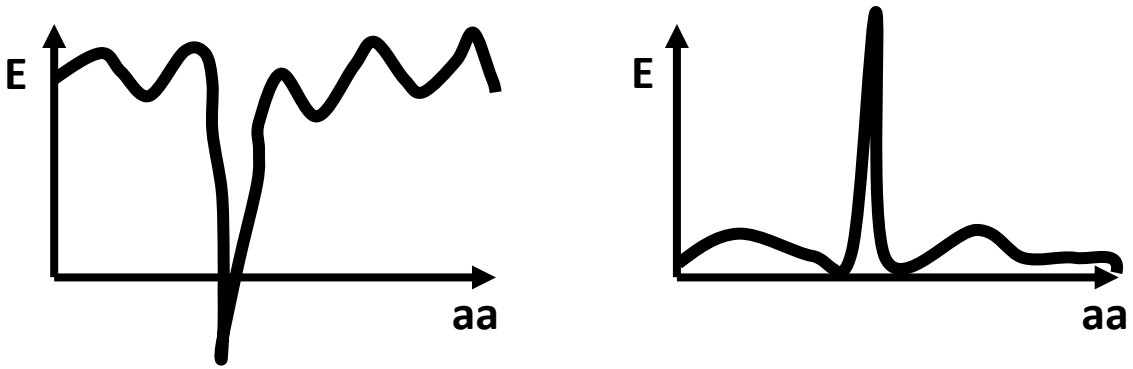
Practice 4: Statistical potentials and PROSA

It depends on what question you want to answer with your model

- **Am I interested in the model of the whole structure?**
- **Am I interested only in one region of the model?**
- **Am I interested only in one function of the model?**

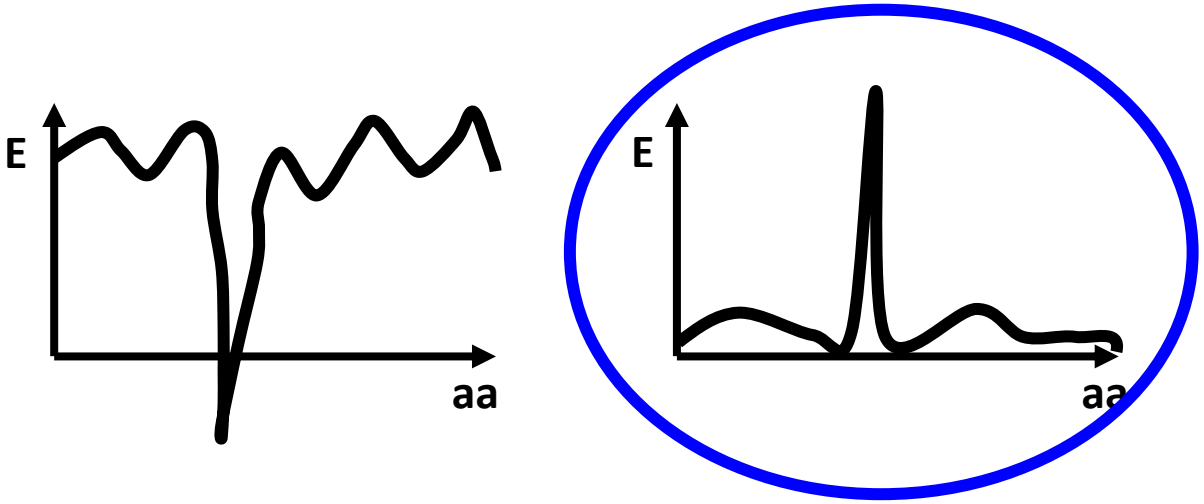
Practice 4: Statistical potentials and PROSA

What model is better?



Practice 4: Statistical potentials and PROSA

What model is better?

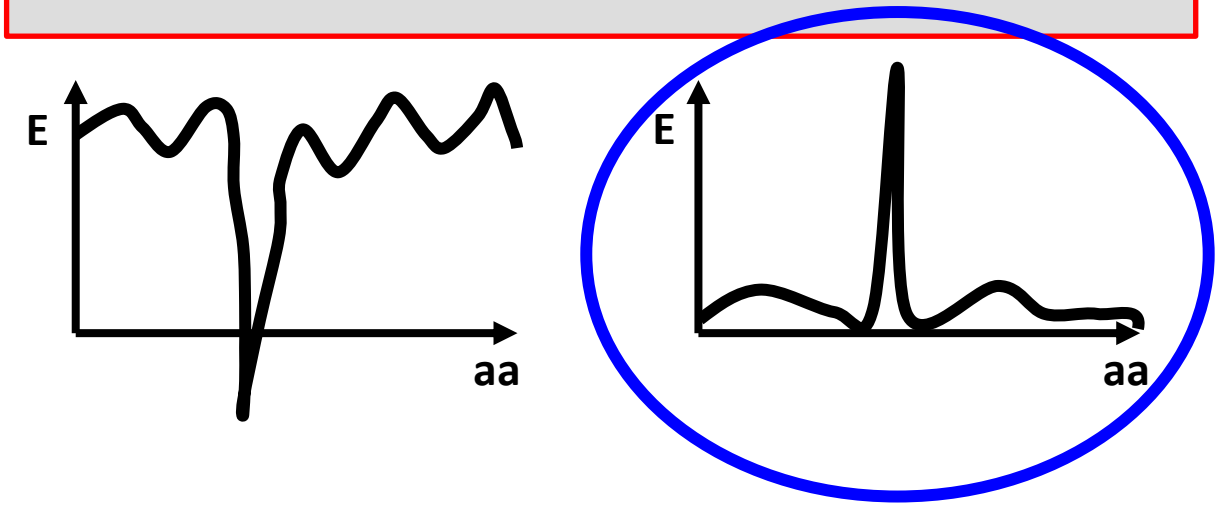


The second model only has one region with bad energies, while the rest of the model is alright. The wrong region could be corrected.

The first model only has one region right and the rest is wrong. Since most of the model is wrong it cannot be corrected. The best thing would be to make a new model.

Practice 4: Statistical potentials and PROSA

What model is better?



The second model only has one region with bad energies, while the rest of the model is alright. The wrong region could be corrected.

The first model only has one region right and the rest is wrong. Since most of the model is wrong it cannot be corrected. The best thing would be to make a new model.