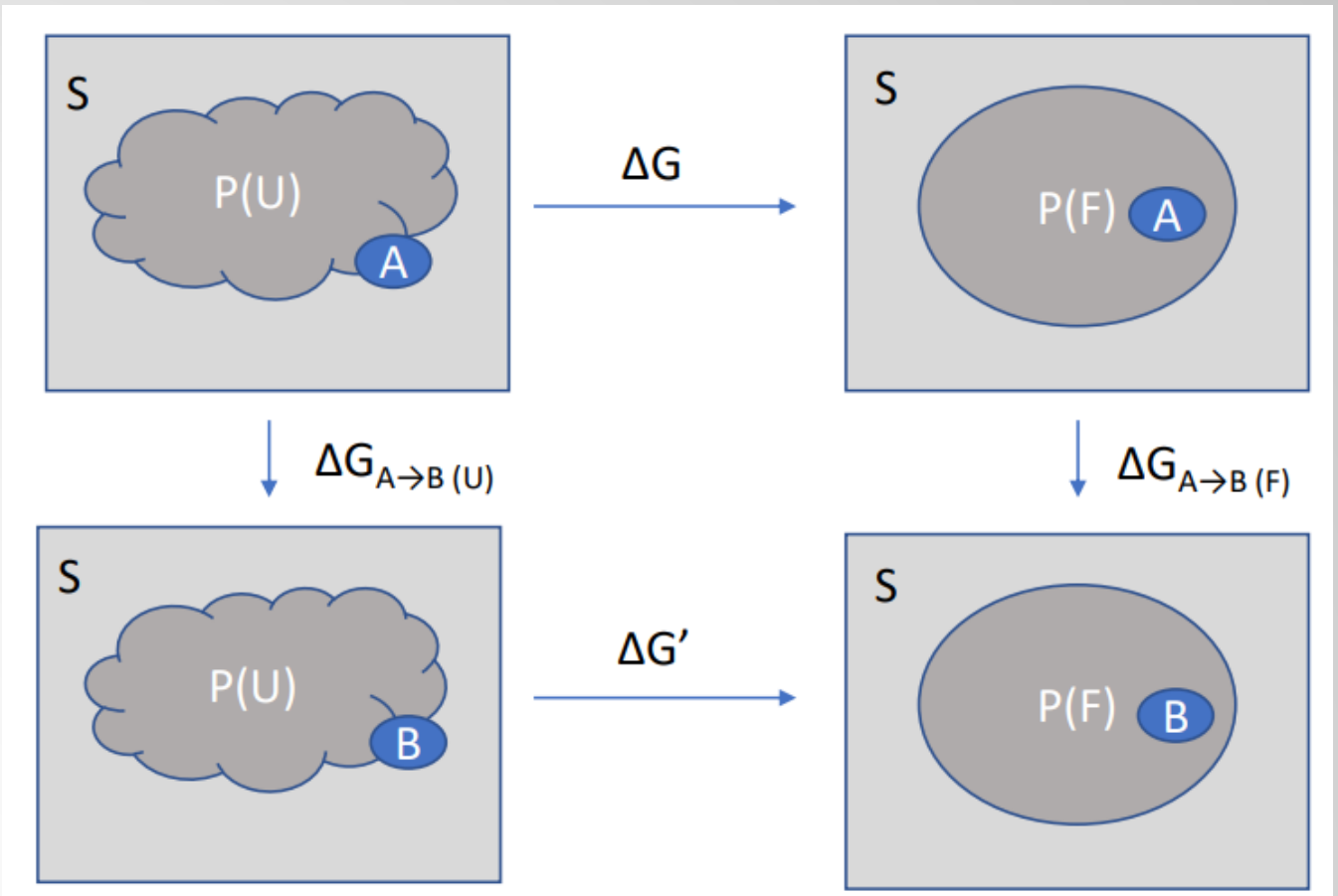
Seminar 5: Mutational analysis



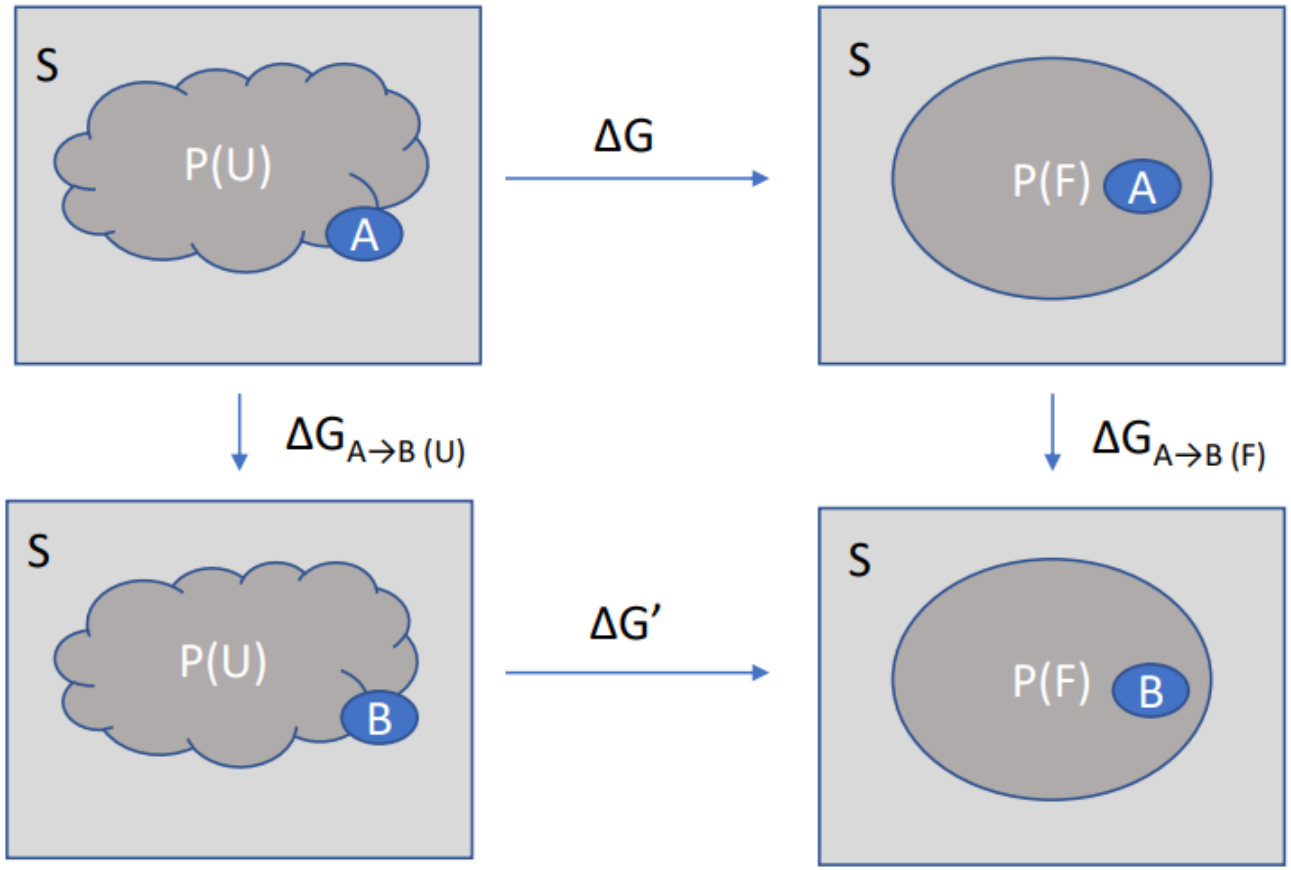
Biophysics

Course 2022-2023

Alberto Meseguer
Irene Acero

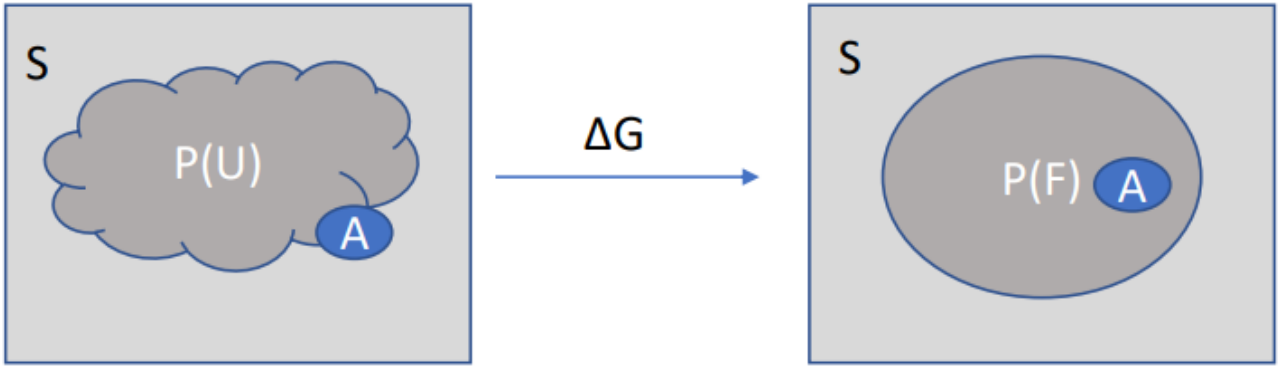
Understanding the folding process

Before understanding the effect of a mutation in the stability of a protein we are going to understand protein folding and protein stability



Understanding the folding process

Before understanding the effect of a mutation in the stability of a protein we are going to understand protein folding and protein stability



Understanding the folding process

We can think of the folding process as a chemical reaction, where the reactants are the unfolded state and the products are the folded state

Unfolded (U) ⟶ Folded (F)

As any chemical reaction, there are thermodynamic variables associated with this process:

Change in free energy (ΔG)

Change in enthalpy (ΔH)

Change in entropy (ΔS)

Understanding the folding process

We can think of the folding process as a chemical reaction, where the reactants are the unfolded state and the products are the folded state

Unfolded (U) ⟶ Folded (F)

As any chemical reaction, there are thermodynamic variables associated with this process:

Change in free energy (ΔG): will this process be spontaneous?

Change in enthalpy (ΔH) : associated with bonds made and destroyed.

Change in entropy (ΔS): associated with the number of available conformations and the hydrophobic effect.

Understanding the folding process

Now let's understand the properties of the unfolded and the folded state

Unfolded	Folded
All residues are exposed to solvent	Some residues are exposed to solvent, some are buried
All H-bonds are made with water	Many H-bonds are made within the protein
Many protein conformations are available	Few protein conformations are available

Understanding the folding process

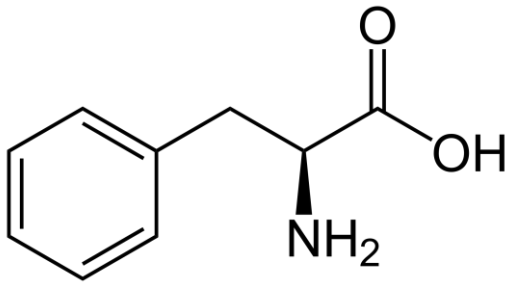
Now let's understand the properties of the

Now let's go over some examples of folding involing different amino acids

available

Folding examples

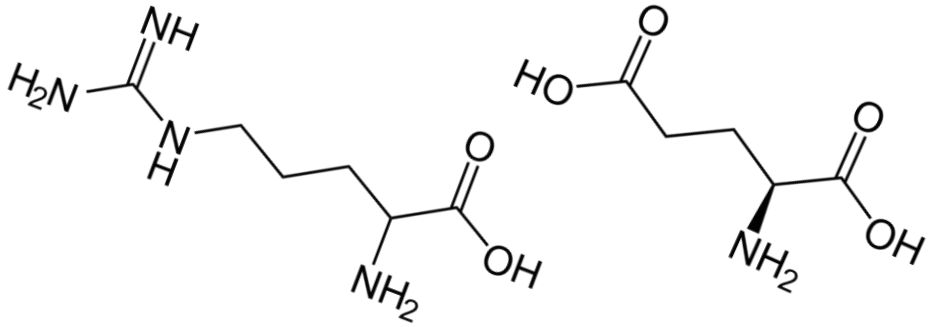
Example 1: A phenilalanine in a peptidic chain goes from unfolded to folded in a buried position



	Unfolded	Folded	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Folding examples

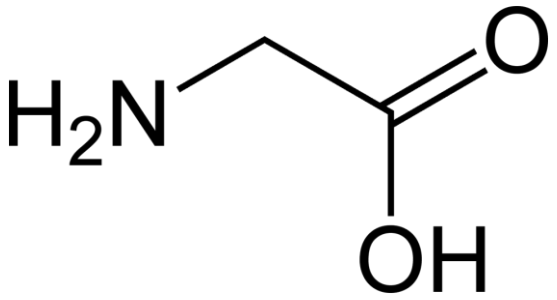
Example 2: An arginine and glutamate pair go from unfolded to a buried position



	Unfolded	Folded	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Folding examples

Example 3: 3 glycines go from unfolded to an exposed position



	Unfolded	Folded	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Glycines and conformational flexibility

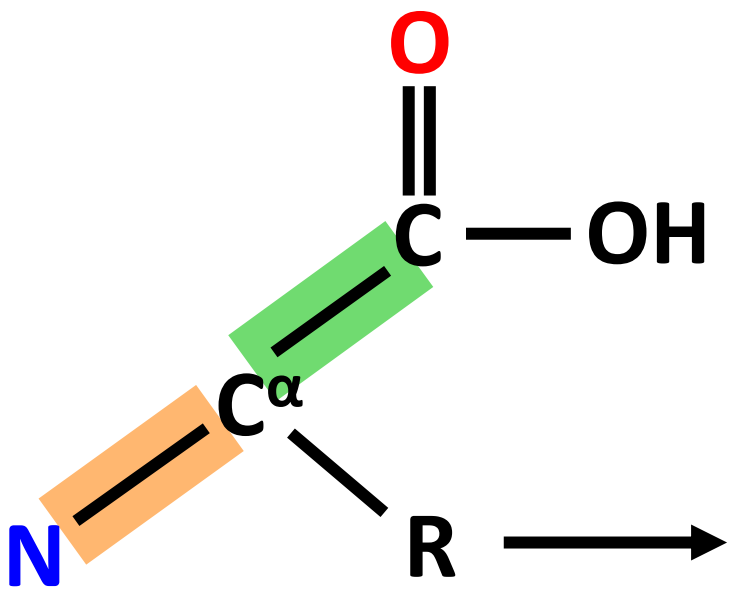
During folding the hydrophobic effect increases entropy, while the loss of conformational flexibility decreases entropy

Unfolded		**Folded**
Hydrophobic residues contact water	**ΔS >> 0** →	Hydrophobic residues hide from water
Many available conformations	**ΔS < 0** →	Few available conformations

Overall, ΔS > 0, this favors the folding to happen spontaneously

Glycines and conformational flexibility

Glycines don't have side chain (one one hydrogen). This enables glycin to adopt many rotation angles that other amino acids cannot undergo.

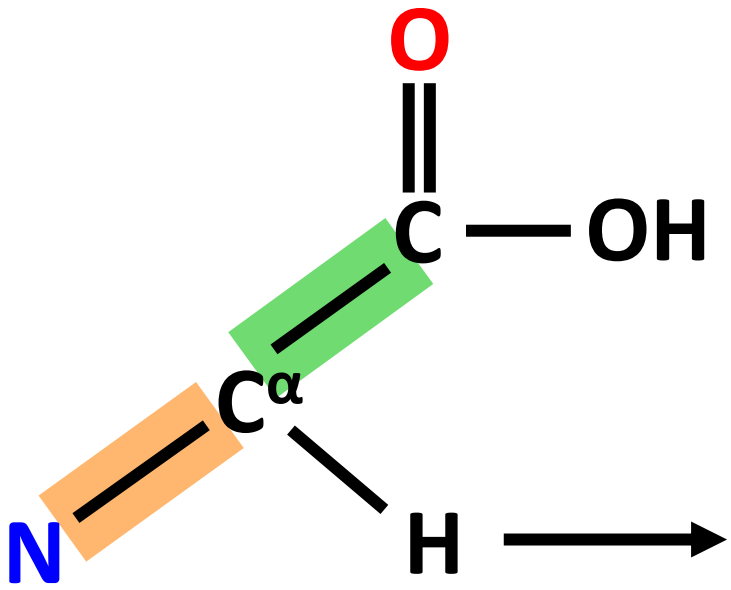


You can check this out at the following link: https://www.umass.edu/molvis/workshop/imgs/phipsian.htm

The side chains of amino acids are big and they can clash with the rest of the protein

Glycines and conformational flexibility

Glycines don't have side chain (one one hydrogen). This enables glycin to adopt many rotation angles that other amino acids cannot undergo.



A hydrogen atom is way smaller than any other side chain. Then, it will be able to adopt many conformations where the hydrogen can fit, but other side chains don't.

Glycines and conformational flexibility

Glycines don't have side chain (one one hydrogen). This enables glycin to adopt many rotation angles that other amino acids cannot undergo.

Unfolded		Folded
Hydrophobic residues contact water	$\Delta S >> 0$ →	Hydrophobic residues hide from water
Many available conformations	$\Delta S < 0$ →	Few available conformations

Glycines and conformational flexibility

Glycines don't have side chain (one one hydrogen). This enables glycin to adopt many rotation angles that other amino acids cannot undergo.

Unfolded		Folded
Hydrophobic residues contact water	**ΔS >> 0** →	Hydrophobic residues hide from water
Many available conformations **<u>Glycines</u>** provide extra flexibility and this provides more available conformations	**ΔS << 0** →	Few available conformations

Glycines and conformational flexibility

Glycines don't have side chain (one one hydrogen). This enables glycin to adopt many rotation angles that other amino acids cannot undergo.

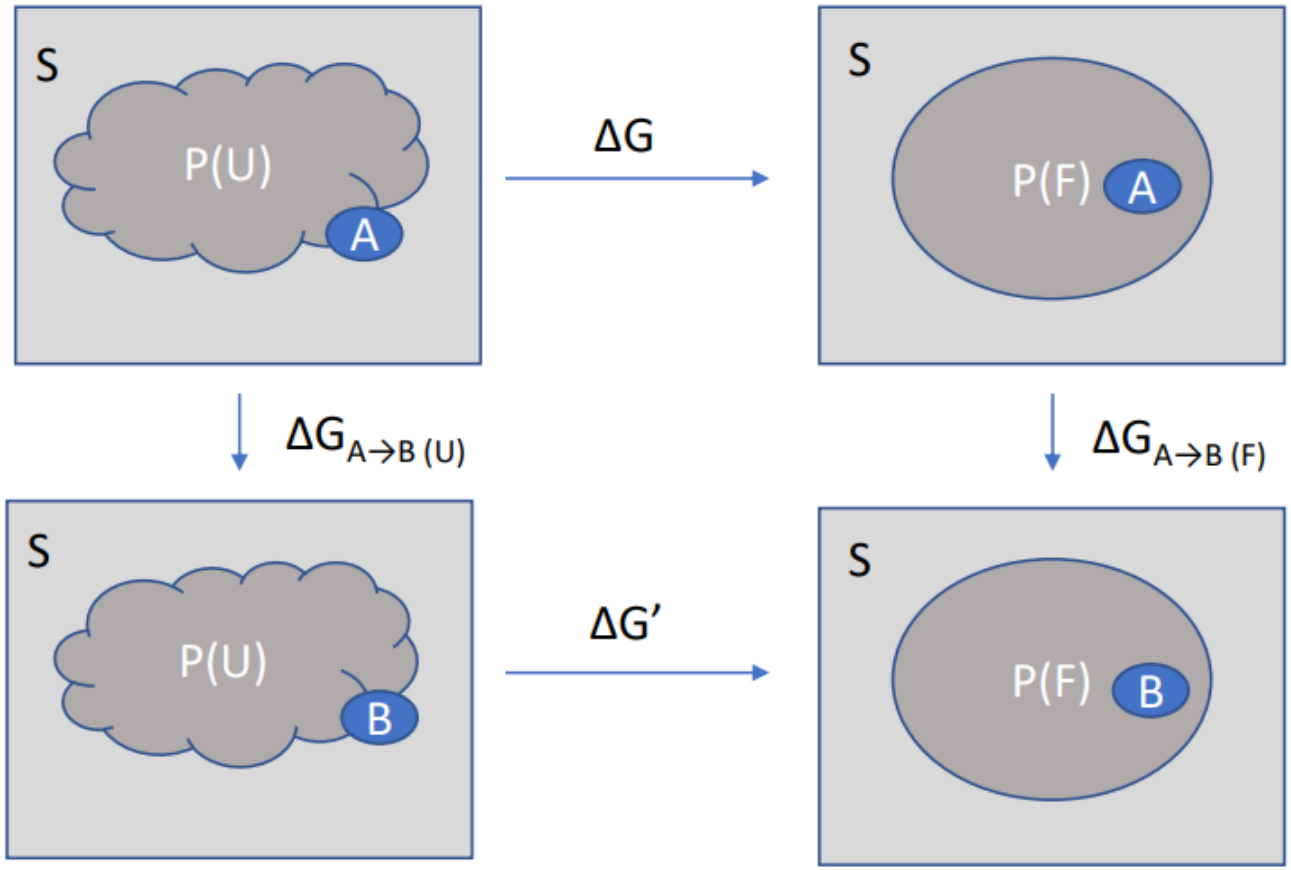
Unfolded		Folded
Hydrophobic residues contact water	**ΔS >> 0** →	Hydrophobic residues hide from water
Many available conformations	**ΔS << 0** →	Few available conformations

<u>Glycines</u> provide extra flexibility and this provides more available conformations

Since ΔS goes down, ΔG goes up and the folding process becomes less favorable

Mutation analysis: the concepts

The objective of a mutation analysis is to determine the effect of a mutation in the stability of the protein. We quantify this with the $\Delta\Delta G$.



Mutation analysis: the concepts

The objective of a mutation analysis is to determine the effect of a mutation in the stability of the protein. We quantify this with the ΔΔG.

We represent the change in free energy for the folding of a protein as **ΔG**.

Remember that the Δ symbol is used to indicate change. Then, the change in **ΔG** when a mutation occurs is the **ΔΔG (ΔΔG = ΔG' - ΔG)**

The **ΔΔG** tells you how is the folding of one protein affected when a mutation happens.

If ΔΔG < 0

if ΔΔG > 0

Mutation analysis: the concepts

The objective of a mutation analysis is to determine the effect of a mutation in the stability of the protein. We quantify this with the ΔΔG.

We represent the change in free energy for the folding of a protein as **ΔG**.

Remember that the Δ symbol is used to indicate change. Then, the change in **ΔG** when a mutation occurs is the **ΔΔG** ($\Delta\Delta G_{ab} = \Delta G_b - \Delta G_a$)

The **ΔΔG** tells you how is the folding of one protein affected when a mutation happens.

If ΔΔG < 0
The mutant protein is more stable, the ΔG of the mutant is lower

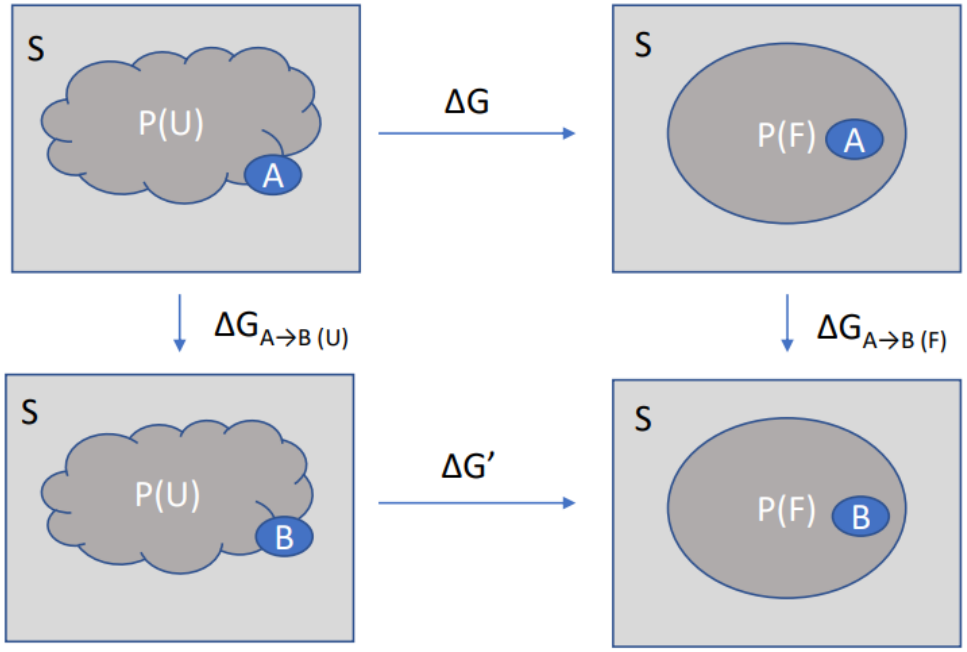
If ΔΔG > 0
The mutant protein is less stable, the ΔG of the mutant is higher

Mutation analysis: the calculations

Since ΔG is a state function, it is pathway independent and this enables us to make the following calculations

$$\Delta\Delta G_{ab} = \Delta G_b - \Delta G_a$$

$$\Delta\Delta G_{ab} = \Delta G_{ab(F)} - \Delta G_{ab(U)}$$



Mutation analysis: the calculations

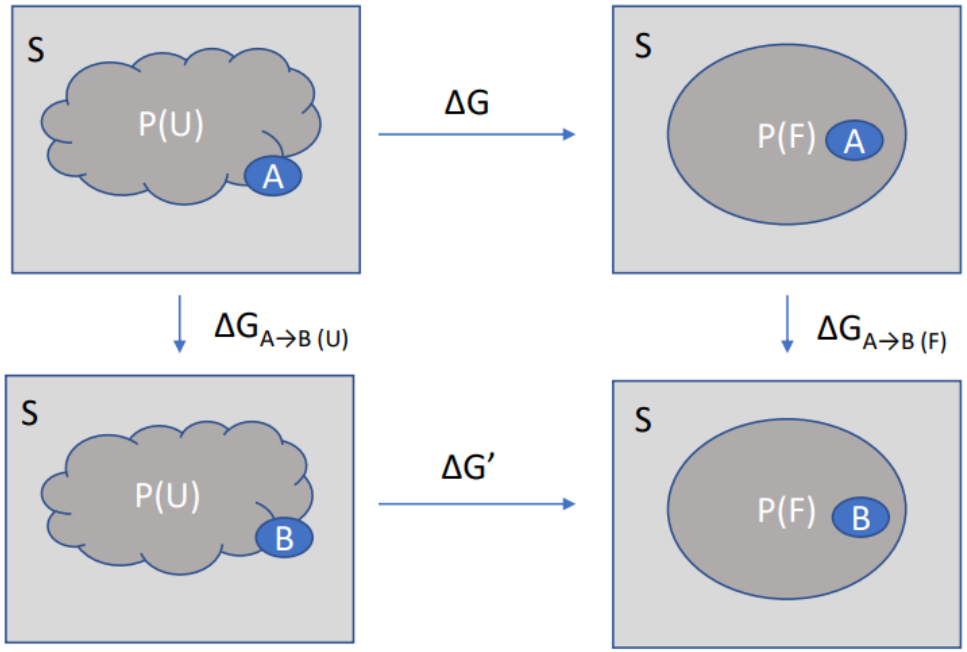
Since ΔG is a state function, it is pathway independent and this enables us to make the following calculations

$$\Delta\Delta G_{ab} = \Delta G_b - \Delta G_a$$

This is very hard to calculate

$$\Delta\Delta G_{ab} = \Delta G_{ab(F)} - \Delta G_{ab(U)}$$

This is easier to calculate



Mutation analysis: the calculations

$$\Delta\Delta G_{ab} = \Delta G_{ab(F)} - \Delta G_{ab(U)}$$

$\Delta G_{ab(F)}$ is the ΔG of going from wild type to mutant in the folded protein

$\Delta G_{ab(U)}$ is the ΔG of going from wild type to mutant in the unfolded protein

Mutation analysis: the calculations

$$\Delta\Delta G_{ab} = \Delta G_{ab(F)} - \Delta G_{ab(U)}$$

$\Delta G_{ab(F)}$ is the ΔG of going from wild type to mutant in the folded protein

$\Delta G_{ab(U)}$ is the ΔG of going from wild type to mutant in the unfolded protein

We can represent ΔG_{ab} (either for folded or unfolded proteins) as an addition of the following terms:

- Internal energy of the mutated residue ($\Delta\Delta G_{InternalAB}$)
- Internal energy of the protein ($\Delta\Delta G_{InternalP}$)
- Interaction between the mutation and the protein ($\Delta\Delta G_{InterP,AB}$)
- Solvation of the protein ($\Delta\Delta G_{SolvationP}$)
- Solvation of the mutated residue ($\Delta\Delta G_{SolvationAB}$)

Mutation analysis: the calculations

$$\Delta\Delta G_{ab} = \Delta G_{ab(F)} - \Delta G_{ab(U)}$$

If we write this equation splitting ΔG_{ab} into all the terms we defined in the previous slide, we get this expression:

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{internalAB}_{A\to B(F)} + \Delta\Delta G^{internalP}_{A\to B(F)} + \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationP}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)}$$
$$-\Delta\Delta G^{internalAB}_{A\to B(U)} - \Delta\Delta G^{internalP}_{A\to B(U)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationP}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

Mutation analysis: the calculations

$$\Delta\Delta G_{ab} = \Delta G_{ab(F)} - \Delta G_{ab(U)}$$

If we write this equation splitting ΔG_{ab} into all the terms we defined in the previous slide, we get this expression:

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{internalAB}_{A\to B(F)} + \Delta\Delta G^{internalP}_{A\to B(F)} + \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationP}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)}$$
$$- \Delta\Delta G^{internalAB}_{A\to B(U)} - \Delta\Delta G^{internalP}_{A\to B(U)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationP}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

These terms cancel out each other. We assume that there are no differences in these energies between the folded and the unfolded state.

Mutation analysis: the calculations

$$\Delta\Delta G_{ab} = \Delta G_{ab(F)} - \Delta G_{ab(U)}$$

If we write this equation splitting ΔG_{ab} into all the terms we defined in the previous slide, we get this expression:

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{internalAB}_{A\to B(F)} + \Delta\Delta G^{internalP}_{A\to B(F)} + \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationP}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)}$$
$$- \Delta\Delta G^{internalAB}_{A\to B(U)} - \Delta\Delta G^{internalP}_{A\to B(U)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationP}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

These terms cancel out each other. We assume that there are no differences in these energies between the folded and the unfolded state. This gives way to the following expression:

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

Mutation analysis: the calculations

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

We can decompose the interaction terms into the different interactions that can be made by one amino acid and the rest of the protein

Mutation analysis: the calculations

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

We can decompose the interaction terms into the different interactions that can be made by one amino acid and the rest of the protein:

- Polar interactions
- Electrostatic interactions
- Van der Waals interactions

Mutation analysis: the calculations

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

We can decompose the interaction terms into the different interactions that can be made by one amino acid and the rest of the protein:

- Polar interactions
- Electrostatic interactions
- Van der Waals interactions

Is this familiar to you???

Mutation analysis: the calculations

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

We can decompose the interaction terms into the different interactions that can be made by one amino acid and the rest of the protein:

- Polar interactions
- Electrostatic interactions
- Van der Waals interactions

	Unfolded	Folded	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Mutation analysis: the calculations

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

To apply this formula you have to use a slightly version of these tables:

Folded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

And use one for the folded protein and another for the unfolded protein.

Mutation analysis: exercises

Analyze qualitatively the energetic consequences of the following sequence variants (consider 2 cases: a: the original residues are exposed to solvent, b: the original residues are buried)

Q-S to E-S

Mutation analysis: exercises

Analyze qualitatively the energetic consequences of the following sequence variants (consider 2 cases: a: the original residues are exposed to solvent, b: the original residues are buried)

Q-S to E-S

Q is the amino acid that gets mutated

S is doesn't get mutated. Try to get the interaction between the mutation and this amino acid.

Mutation analysis: exercises

Q-S to E-S, exposed

Folded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Unfolded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Mutation analysis: exercises

Q-S to E-S, exposed

Folded protein	Overall ΔG
Polar int.	0
Electrostatic int.	0
Van der Waals int.	0
Solvation	-
Total	-

Unfolded protein	Overall ΔG
Polar int.	0
Electrostatic int.	0
Van der Waals int.	0
Solvation	-
Total	-

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

Mutation analysis: exercises

Q-S to E-S, exposed

Folded protein	Overall ΔG
Polar int.	0
Electrostatic int.	0
Van der Waals int.	0
Solvation	-
Total	-

Unfolded protein	Overall ΔG
Polar int.	0
Electrostatic int.	0
Van der Waals int.	0
Solvation	-
Total	-

$$\Delta\Delta G_{A\to B} = \Delta\Delta G^{interP,AB}_{A\to B(F)} + \Delta\Delta G^{solvationAB}_{A\to B(F)} - \Delta\Delta G^{interP,AB}_{A\to B(U)} - \Delta\Delta G^{solvationAB}_{A\to B(U)}$$

Mutation analysis: exercises

Q-S to E-S, buried

Folded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Unfolded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Mutation analysis: Try new mutations

W-W to G-G, buried

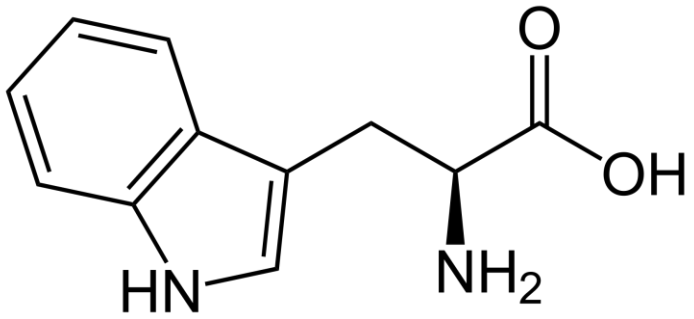
Folded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Unfolded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

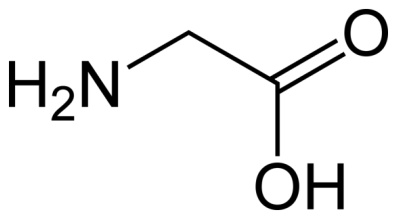
Mutation analysis: Try new mutations

W-W to G-G, buried

Tryptophan is the amino acid with the biggest side chain



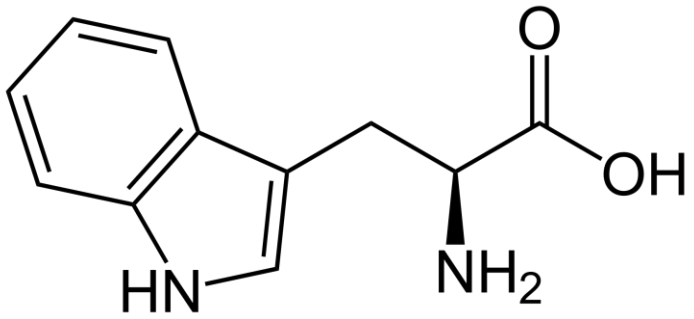
Glycine is the amino acid with the smallest side chain



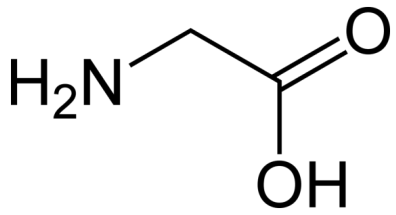
Mutation analysis: Try new mutations

W-W to G-G, buried

Tryptophan is the amino acid with the biggest side chain



Glycine is the amino acid with the smallest side chain



G-G has way more possible conformations in the unfolded state than W-W. This makes ΔS get lower in the G-G pair, thus leading to a higher ΔG.

If the mutant has higher ΔG, then the mutation decreases protein stability and ΔΔG > 0

Glycines and conformational flexibility

Glycines don't have side chain (one one hydrogen). This enables glycin to adopt many rotation angles that other amino acids cannot undergo.

Unfolded		Folded
Hydrophobic residues contact water	**ΔS >> 0** →	Hydrophobic residues hide from water
Many available conformations	**ΔS << 0** →	Few available conformations

<u>Glycines</u> provide extra flexibility and this provides more available conformations

Since ΔS goes down, ΔG goes up and the folding process becomes less favorable

Mutation analysis: Try new mutations

F-K to E-K, exposed

Folded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

Unfolded protein	Wild type	Mutant	Overall ΔG
Polar int.			
Electrostatic int.			
Van der Waals int.			
Solvation			
Total			

A reflection on protein stability

Is protein stability good or bad for biological systems?

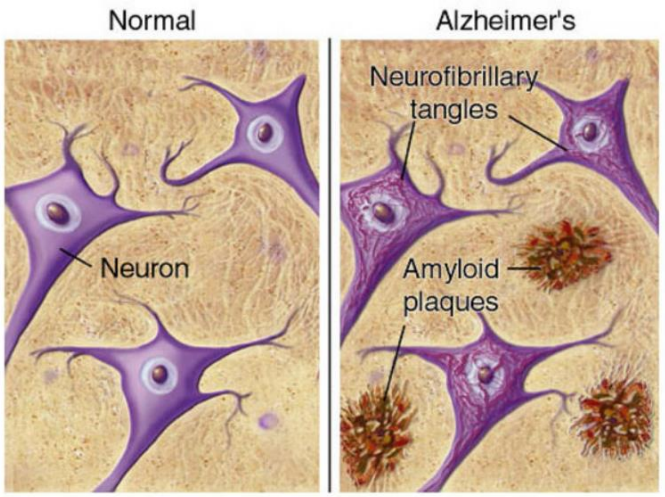
Is maximum stability for proteins the best for biological systems?

A reflection on protein stability

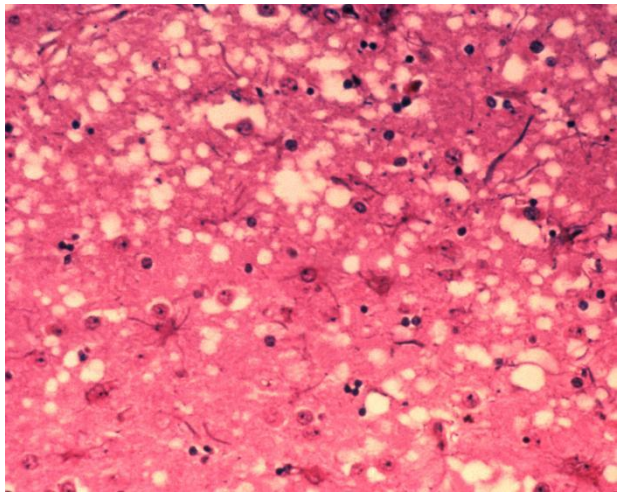
Superstable proteins are a problem for the cell, because it cannot degrade them, leading to the accumulation of such proteins in the cell

This mechanism is at the basis of diseases such as:

Alzheimer's disease
Beta amyloid accumulates creating extracelular plaques



Bovine spongiform encephalopathy (AKA mad cow disease)



Have a terrifying Halloween

If you want a really terrifying story for Halloween, investigate the origin of the mad cow disease...