جامعة الملك سعود
King Saud University

STRATEGIZE PROTEIN ENGINEERING

Protein Engineering (PE)

A method of changing a protein sequence to achieve a desired result, such as:

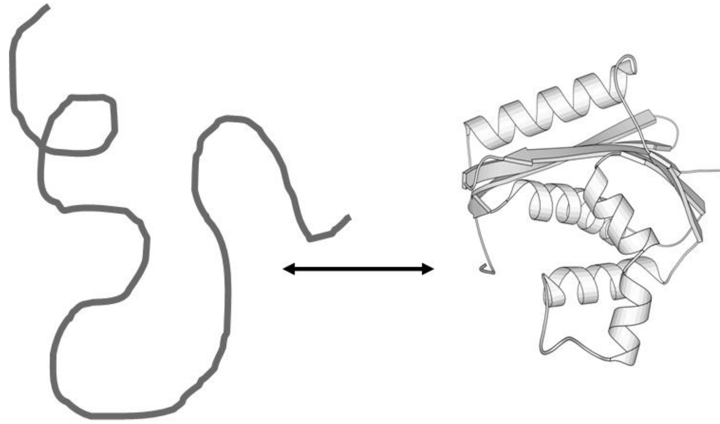
- **Change in the substrate specificity.**
- **Increased stability** to
  - → temperature,
  - → organic solvents, and/or
  - → extremes of pH.

Protein Engineering Targets

Stability

Specificity

Structural



UNFOLDED (U) NATIVE (N)

Slow down unfolding

$$N \underset{}{\overset{K_{unfold}}{\rightleftharpoons}} D \quad K_{unfold} = \frac{k_{unfold}}{k_{fold}}$$

Speed up folding

What does increase protein stability mean?

Proteins are dynamic object exists in equilibrium between the native and unfolded stat

The stabilizing interaction shifts

the folding–unfolding equilibrium toward folding

→ **the protein spends more time in the folded form**

It mean stabilizing **the folded form** or **destabilizing the unfolded form**

Stabilizing a protein does not prevent unfolding

Stabilization strategies

Stabilization to cooperative unfolding

- Restore conserved amino acids
- Destabilize unfolded ensemble
- Stabilize folded form

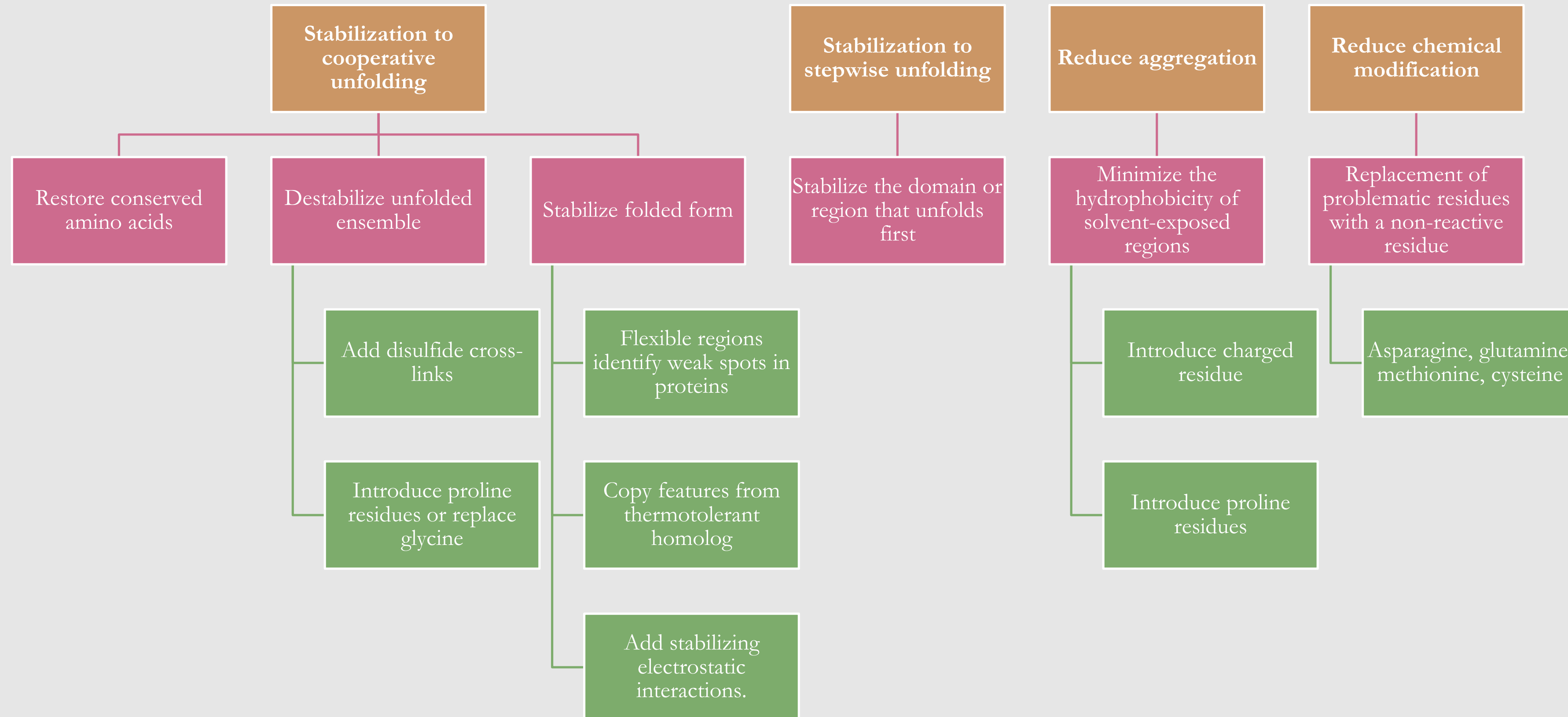
Stabilization to stepwise unfolding

- Stabilize the domain or region that unfolds first

Reduce aggregation and chemical modification

- Minimize the hydrophobicity of solvent-exposed regions
- Replacement of problematic residues with a non-reactive residue

Stabilization strategies



identifies amino acid residues conserved within homologous proteins, but which are missing in the target protein. Evolution conserves amino acids that contribute to protein function. This contribution may be to structure, catalysis, stability or other protein property. The consensus approach hypothesizes that conserved amino acid residues are more

Original polypeptide

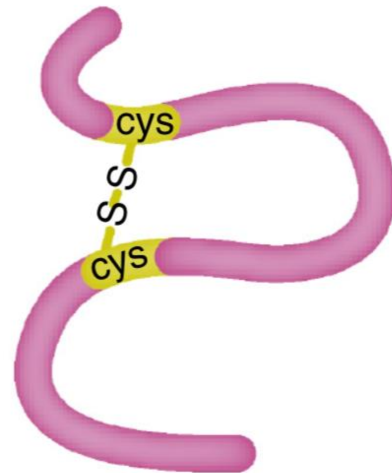
ALTER DNA

Engineered polypeptide

Cysteine

FOLDING

FOLDING

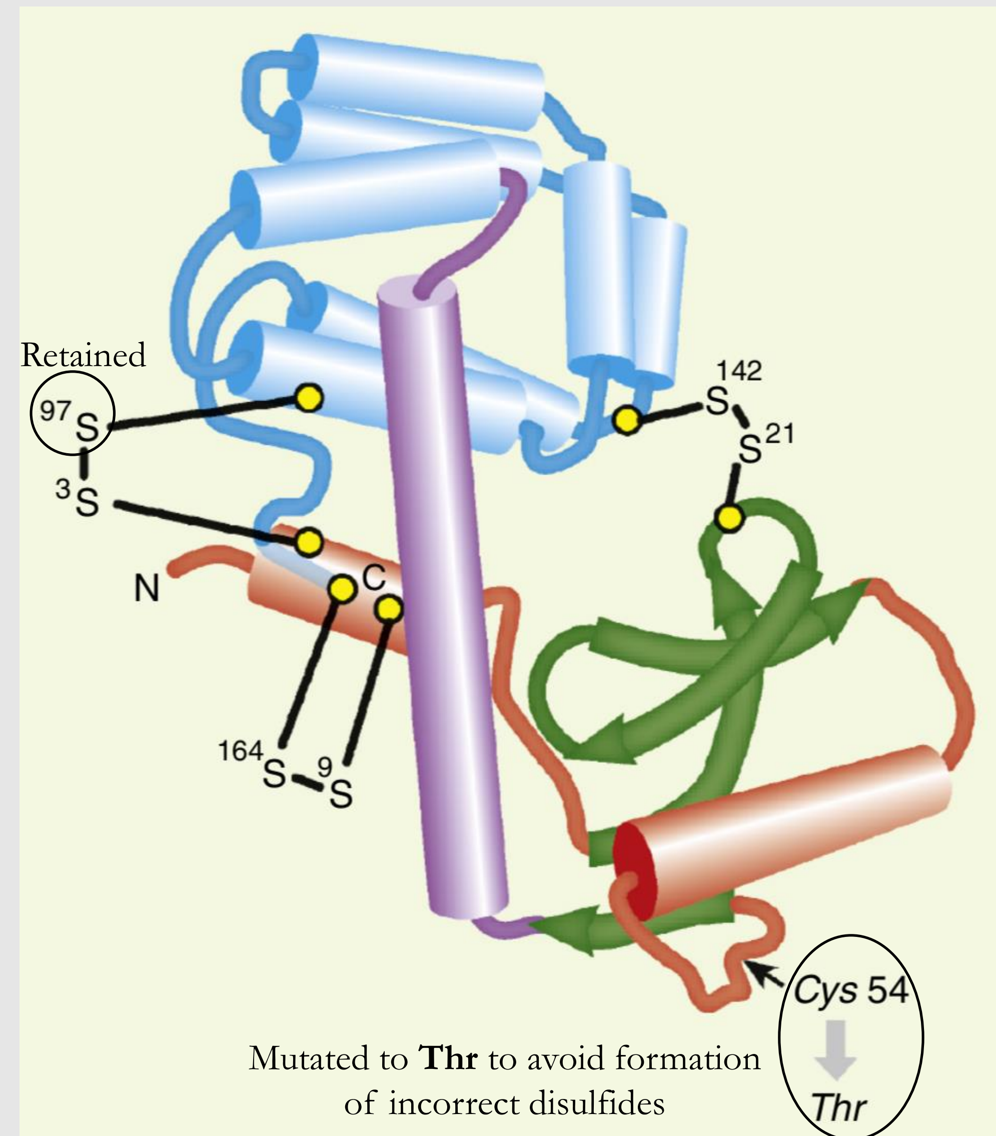


Introduce disulfide bonds

- Relatively straightforward way to increase the stability of proteins.
- A disulfide bond can be added to a protein by changing two amino acids into cysteines by site-directed mutagenesis.
- Under oxidizing conditions, the two cysteines form a disulfide bond, holding the protein together at that site.

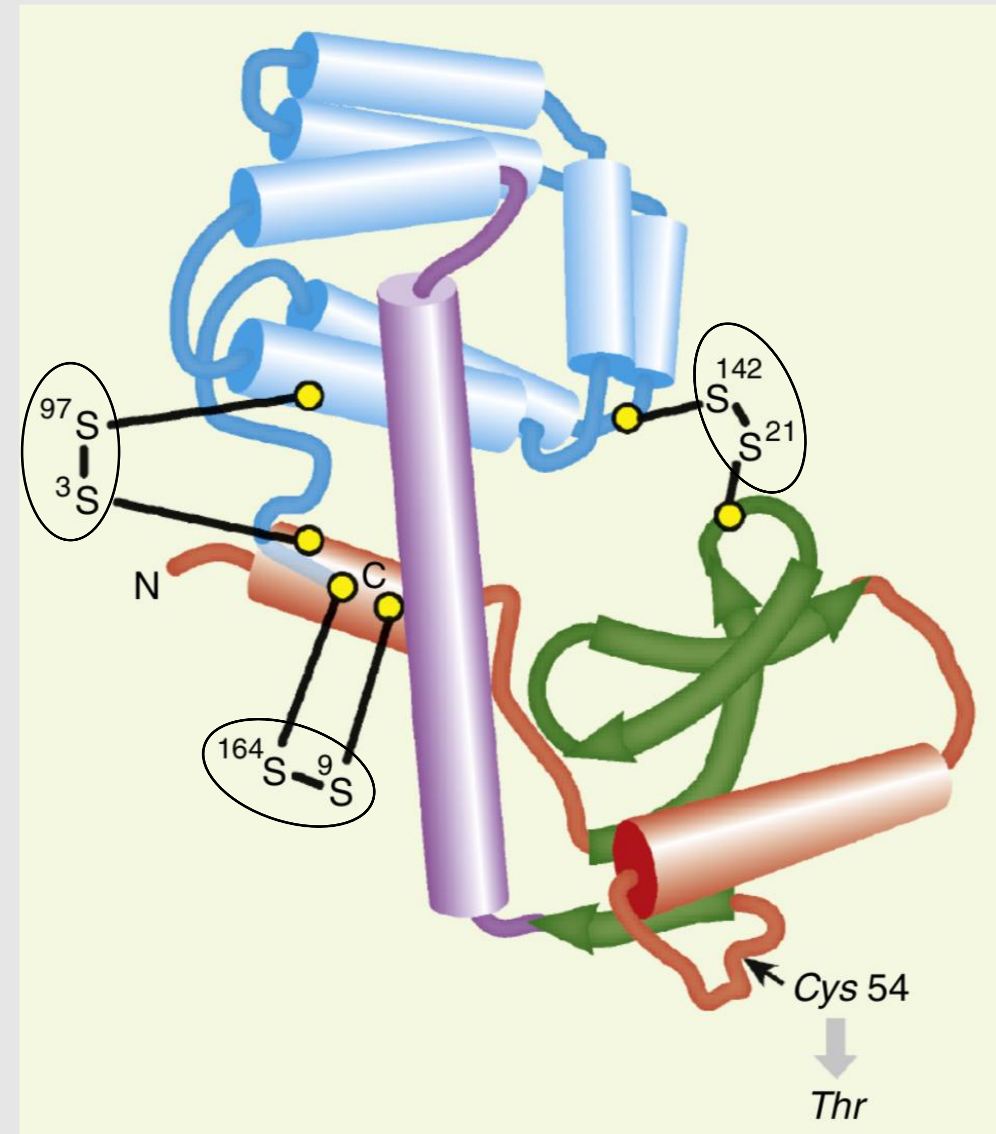
Disulfide Engineering of T4 Lysozyme

- T4 lysozyme The polypeptide chain of 164 amino acids folds into two domains
  - →**The N-terminal region** is shown in **green** and **red**
  - →**The C-terminal region** is in **blue**.
  - →These are linked by an alpha helix (**purple**).
- **Has two cysteines**, neither involved in disulfide bond formation in the wild-type protein.



Disulfide Engineering of T4 Lysozyme

- Extensive analysis of possible locations for disulfides was carried out.
- Those disulfides that might impair other stabilizing interactions in the protein were eliminated.
- This left **three possible disulfide bonds** that should theoretically promote stability.



To test these experimentally,
5 A.As (Ile3, Ile9, Thr21, Thr142, and Leu164) were converted to Cys in various combinations.

Disulfide Engineering of T4 Lysozyme

- Engineered proteins were tested for stability and for enzyme activity

Table 11.2 Disulfide Stabilization of T4 Lysozyme

Protein	Disulfide Bonds Present 3–97	9–164	21–142	Stability as Tm	Activity (%)
Original	–	–	–	41.9	100
1	+	–	–	46.7	96
2	–	+	–	48.3	106
3	–	–	+	52.9	0
4	+	+	–	57.6	95
5	–	+	+	58.9	0
6	+	+	+	65.5	0

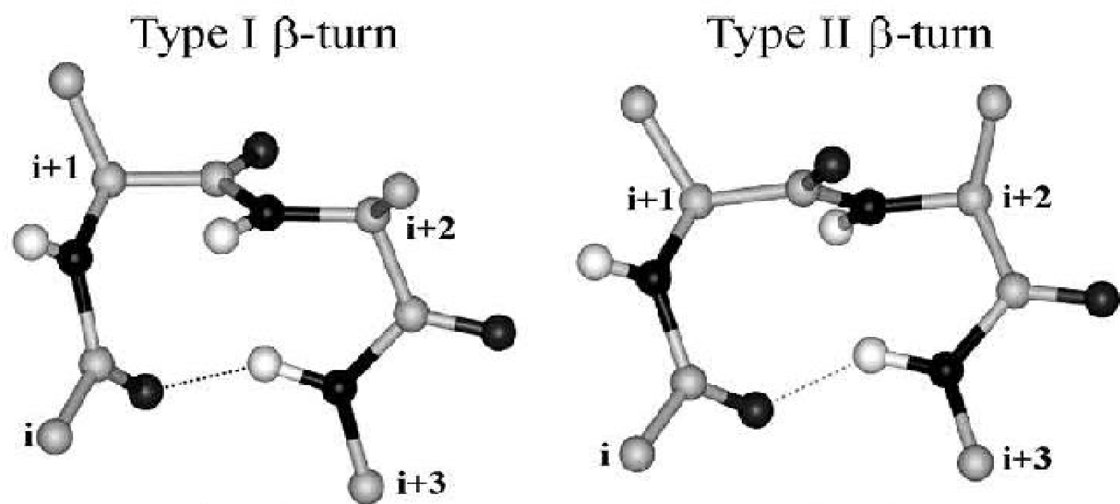
↑ increase in stability
✓**retained almost all of its enzyme activity**

↑ increase in stability
↓ **lost its enzyme activity**

Table 4 Stabilization of lysozyme from phage T4 by substitutions that reduce the flexibility of the denatured state[25]

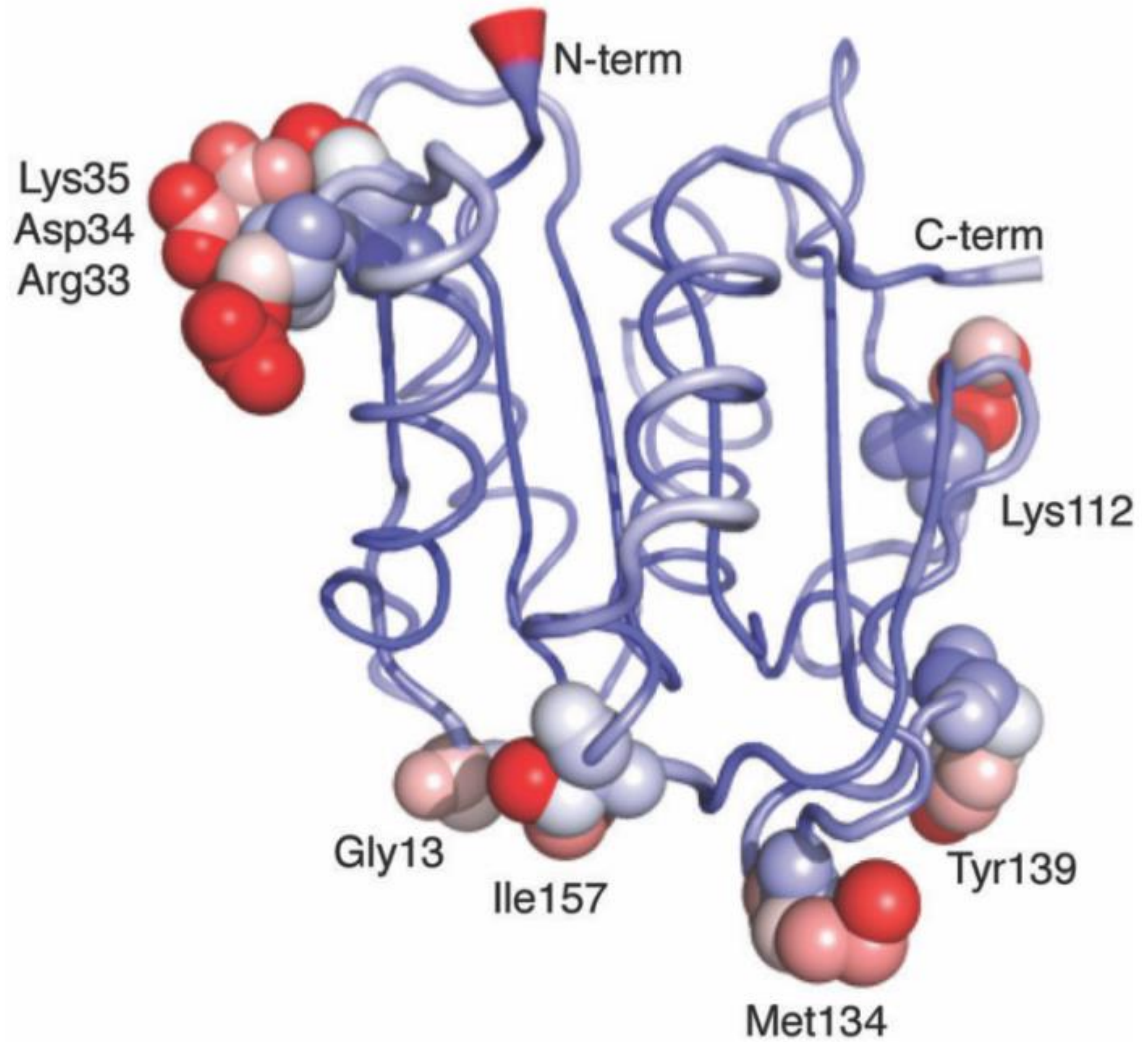
Protein	Location of substitution	T_m, °C	$\Delta\Delta G$ (kcal mol^{-1})
Wt	—	64.7	0
Ala82Pro	Near start of α-helix	66.8	0.8
Gly77Ala	Near end of α-helix[a]	65.6	0.4

[a] Stabilization likely due to methyl interactions that stabilize the helix and not due to changes in flexibility of the denatured state.



Introduce proline or replace glycine

- Introducing proline is expected to reduce the flexibility of the denatured ensemble.
- The main chain angles for proline fit well in three places:
  - near the start of an a-helix,
  - i + 1 position in a type I or II b-turn
  - i position of a type II b-turn.
- These choices of proline location consider only main chain angles; the success of any substitution also assumes minimal changes to side chain interactions.
- Replacing glycine with any other amino acid should have a similar effect.



Flexible regions identify weak spots in proteins

- Flexible regions in a folded protein identify areas where interactions with other amino acids are weak
- Loops on the surface and the N- and C-termini of the protein are often the most flexible.

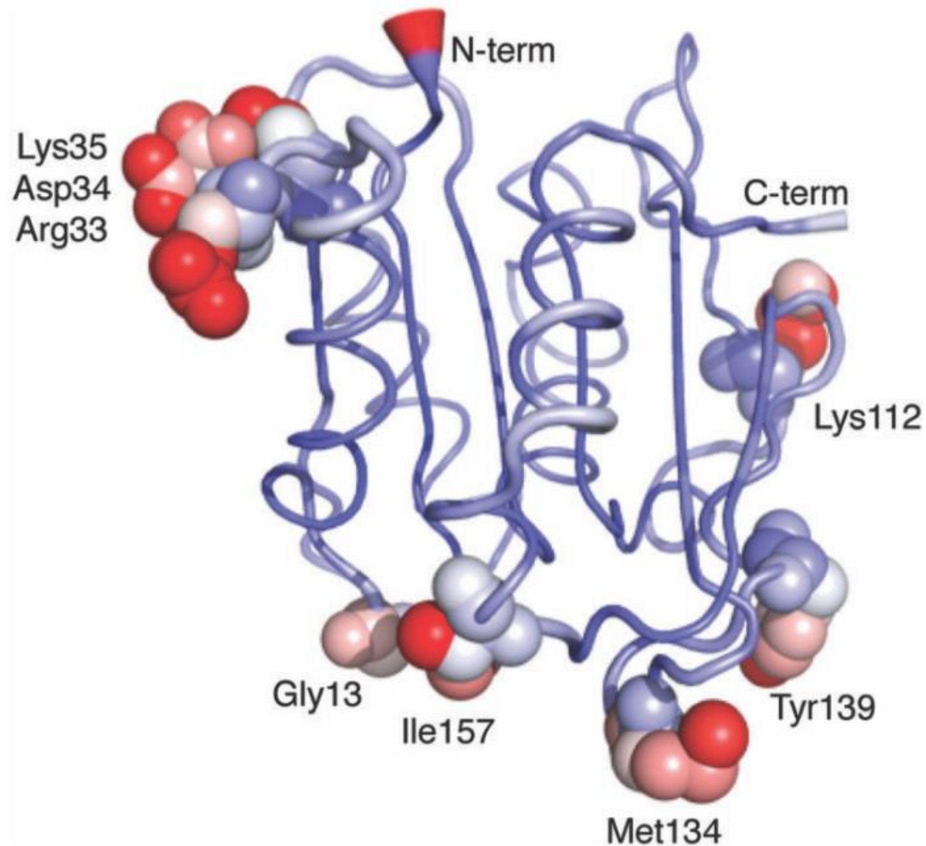


Fig. 12 X-ray structure of a lipase (pdb code 1isp) colored to show the *B*-factor from low (blue) to high (red). Reetz and coworkers[30] replaced the residues in eight regions with the highest flexibility (*B*-factor), excluding the N- and C-termini, using random mutagenesis. Stabilizing substitutions (2–4 °C increase in T_m) occurred in the six regions where the residues shown as spheres. Side chain atoms are more flexible than the main chain atoms.

Flexible regions identify weak spots in proteins

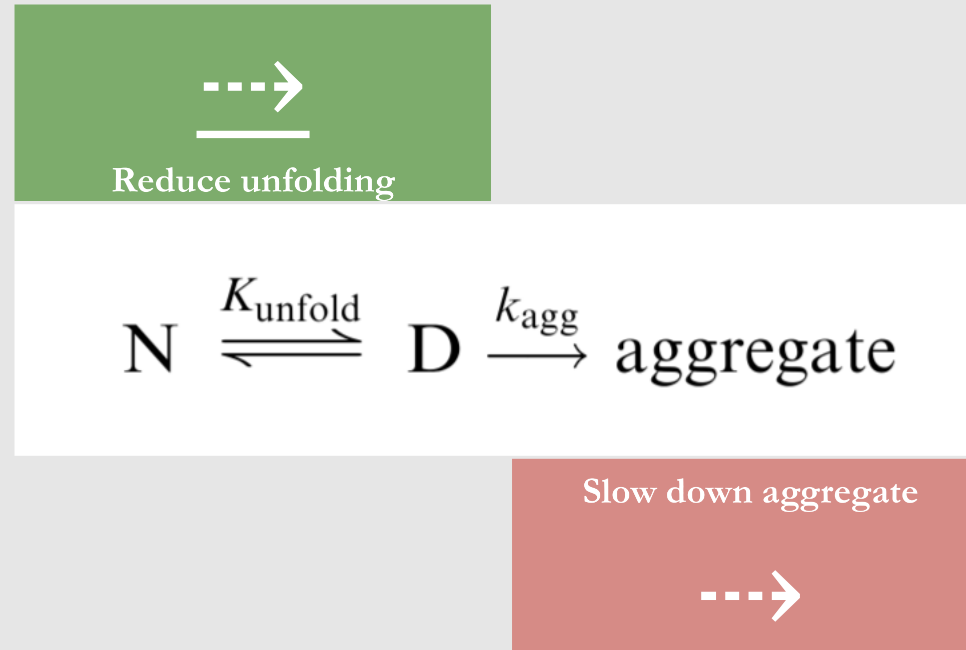
- There are two steps to this stabilization approach:
  1. **Identify the flexible regions of the protein**
     - One way to identify flexible areas is molecular dynamics simulation, which directly models protein motion
     - Another way to identify flexible regions is the high B-factors in the X-ray crystal structure

"**B-Factors** describe the spreading of electron density assigned to that atom. Maybe due to:

- movement of the atom during the X-ray analysis (temperature-dependent atomic vibrations), or
- the atom occupying several fixed positions in the structure (static disorder), which suggests motion in solution."

Reduce aggregation

Two strategies to reduce aggregation:



Reduce aggregation

Aggregation-prone regions :

- hydrophobic,
- beta-sheet forming stretches

Substitutions:

- These substitutions are charged residues (R, K, D, or E) which reduce hydrophobic interactions needed for aggregation or,
- proline, which hinders the formation of beta-sheet structures in aggregates.

Reduce chemical modifications

- **The side chains of three amino acids readily undergo spontaneous chemical modification:**
  1. asparagine deamidates to aspartate,
  2. methionine oxidizes to a sulfoxide and
  3. cysteine oxidizes to disulfide protein oligomers as well as sulfur oxides such as sulfenic acids

- **The first protein engineering of an industrial enzyme was**
  - Removal of an oxidation sensitive methionine from the **detergent protease subtilisin**.
  - The existing subtilisin could tolerate most of the harsh conditions of laundry (heat, high pH, and detergent), but it could not tolerate bleach.
  - Bleach oxidized a methionine near the active site to the methionine sulfoxide (R–S(O)–CH3), which hindered binding of the substrate proteins to inactivate the protease.
  - Replacement of this problematic **methionine with alanine** created a bleach-tolerant protease.

Reduce chemical modification application

- **A cysteine-to-serine replacement stabilizes several cytokine drugs.**
- This change does not affect biological activity, but avoids oligomerization by the formation of non-native intermolecular disulfide links between proteins.
- The specific activity of interferon-b, when expressed in *E. coli*, was about 10-fold lower than the native protein. The researchers hypothesized that oligomerization through intermolecular disulfide bonds caused the lower activity.
- Replacement of Cys17 with Ser eliminated oligomerization and restored the specific activity to that of the native protein.
- The commercial drug Betaseron for reduce the frequency of clinical exacerbations of Multiple Sclerosis contains this substitution.

Table 7 Heat-stabilizing substitutions in α-amylase from *Bacillus licheniformis*

Approach	Specifics
Prevent chemical modification	Remove oxidation (M197) and deamidation sites (N188, Q264)
Stabilize N	Bury Ca^{2+} ion (A181T), minimize electrostatic destabilization (H156Y)
Destabilize U	Introduce proline (R124P)
Random mutagenesis	M15T, H133I, N188S, A209V

Stabilizing in conclusion

- Large stabilizations require multiple stabilizing mutations
- For example, a heat-stabilized a-amylase contains at least ten amino acid substitutions

Table 1 Design strategies to stabilize proteins and web tools to implement these strategies

Strategy	Rationale	Example of implementation
Restore residues conserved in homologs	Not specified	Consensus Finder identifies conserved amino acids in homologs that are missing from target protein
Add disulfide links	Destabilize unfolded protein	SSBOND identifies suitable locations, additional analysis needed to narrow choices
Add Pro residues	Destabilize unfolded protein	Analysis of structure to identify locations suitable for proline (PROMOTIF)
Substitutions in or near flexible regions	Stabilize folded protein	Random mutagenesis in or near flexible regions (B-FITTER); or molecular dynamics modeling to identify missing interactions
Random mutagenesis & screening	Not specified	Random mutagenesis anywhere followed by screening for stabilized variants
Optimize electrostatic interactions	Stabilize folded protein	PROPKA identifies stabilizing or destabilizing electrostatic interactions
Improve hydrophobic packing	Stabilize folded protein	Modeling with Rosetta or FoldX to identify stabilizing substitutions

Web tool to implement stabilization strategies

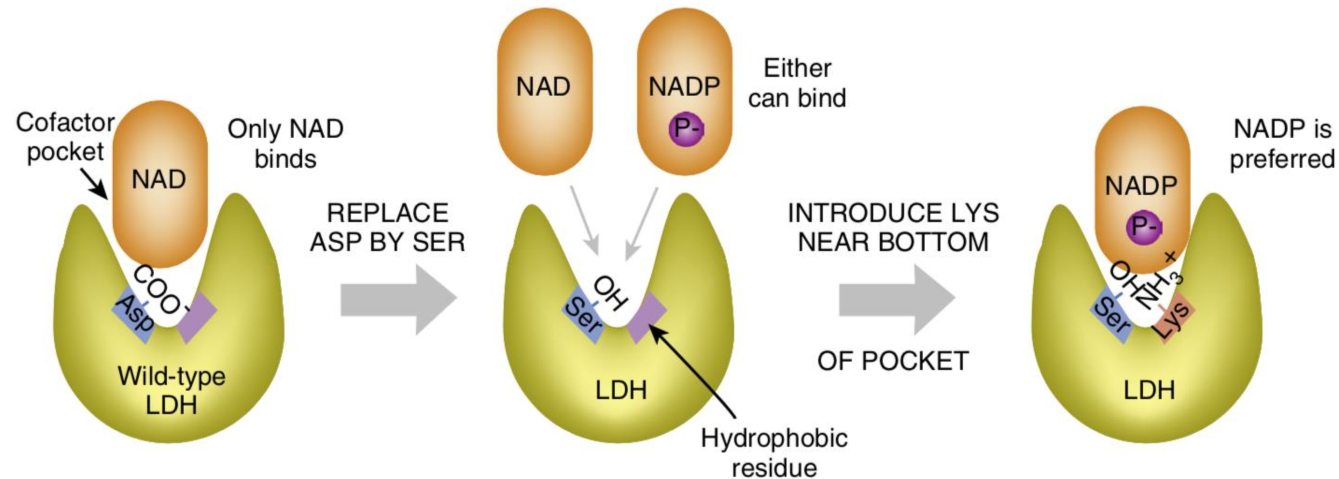
Stabilize the native protein require:

- a protein structure or,
- a homology model to start the design.

Changing binding site specificity

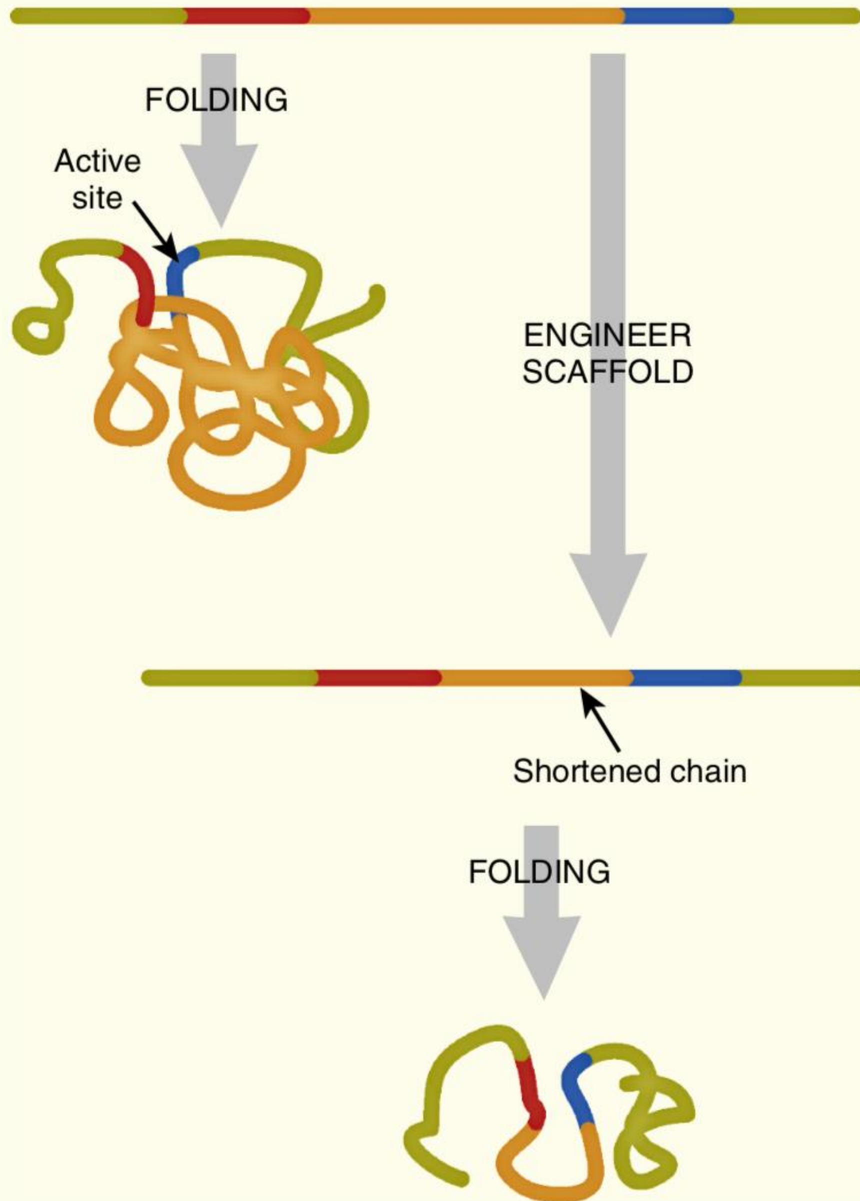
- The most straightforward alterations are those that change the binding specificity for the substrate or a cofactor **but do not disrupt the enzyme mechanism**.

- Changing the specificity for a cofactor or substrate may be useful, **either to make the product of the enzyme reaction less costly or to change it chemically.**



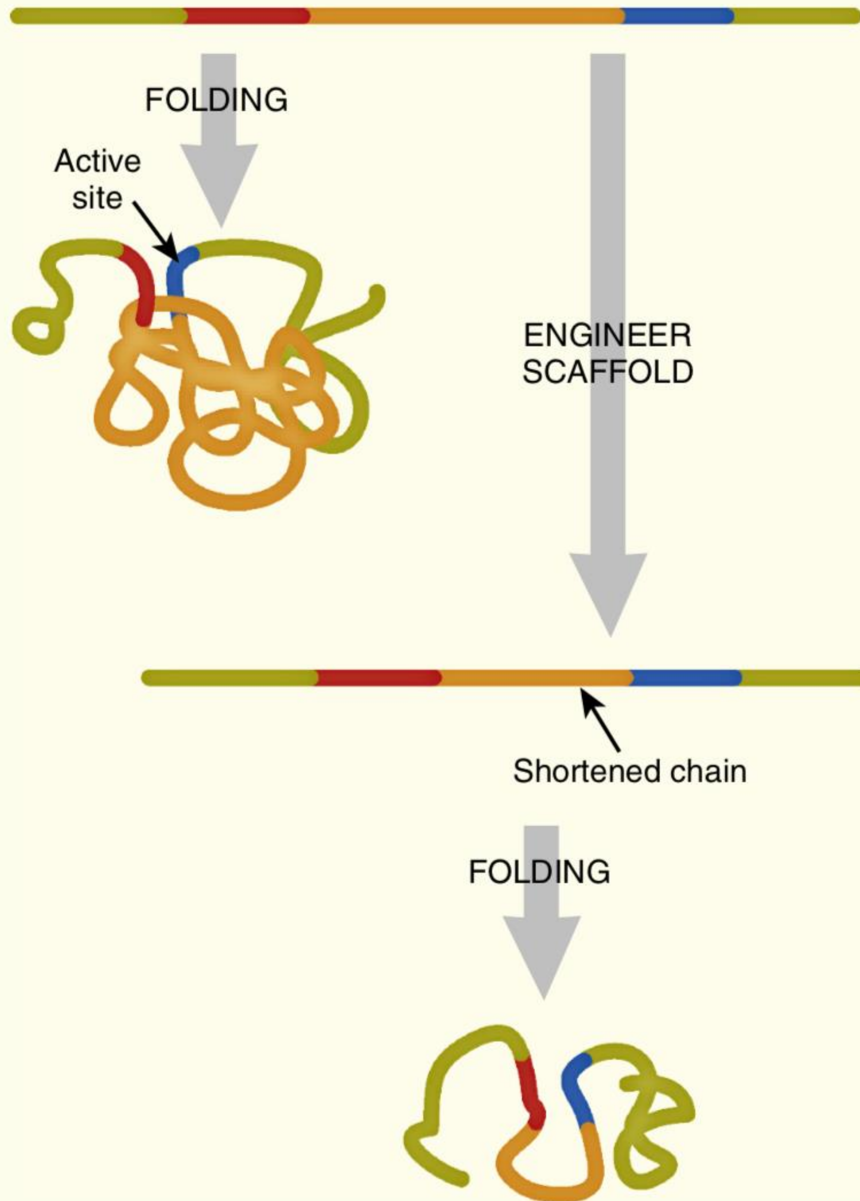
Changing Cofactor Preference of Lactate Dehydrogenase

- Lactate dehydrogenase (LDH) preferentially binds NAD because the binding pocket has an aspartic acid.
- The negatively charged carboxyl repels the negatively charged phosphate of NADP Changing the aspartic acid to serine allows either NAD or NADP to bind to LDH.
- Adding a positively charged lysine makes the pocket more attractive to the NADP



Structural scaffolds

- Relatively **few of the amino acid residues** in a protein are actually involved in the active site.
- Most of the protein provides the 3D platform or scaffold needed to correctly position the active site residues.
- **Quite often the scaffold is much larger than really necessary.**



Structural scaffolds

- The β-galactosidase of *Escherichia coli* (LacZ protein) has approximately 1000 amino acids, whereas most simple hydrolytic enzymes have only 200 to 300.
- **Presumably it should be possible to redesign a functional β-galactosidase that is only 25% to 30% the size of LacZ protein**.
- From an industrial viewpoint, such a smaller protein would obviously be more efficient.

References

- Clark, D. P., & Pazdernik, N. J. (2016). *Biotechnology*. Amsterdam: Academic/Cell.
- Kazlauskas, R. (2018). Engineering more stable proteins. *Chemical Society Reviews*, *47*(24), 9026–9045. doi: 10.1039/c8cs00014j
- Robic, S. (2010). Mathematics, Thermodynamics, and Modeling to Address Ten Common Misconceptions about Protein Structure, Folding, and Stability. *CBE—Life Sciences Education*, *9*(3), 189–195. doi: 10.1187/cbe.10-03-0018

THANK YOU