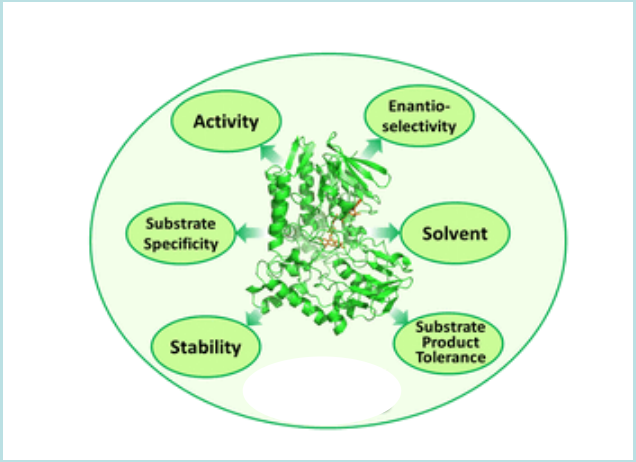
Lect#13: Enzyme Engineering and Design

1. Protein Engineering
2. Engineering Technology
3. Engineering Methods
4. Engineering proteins to be more stable enzymes
5. Engineering a more stable T4 lysozyme



6. Disulfide bonds increase T4 lysozyme stability
7. Helix stabilization (capping) of T4 lysozyme
8. Incorporation of Non-Canonical Amino Acids (NCAA) into Proteins
9. Protein Engineering by Combinatorial Methods
10. Machine-learning-guided directed evolution for protein engineering

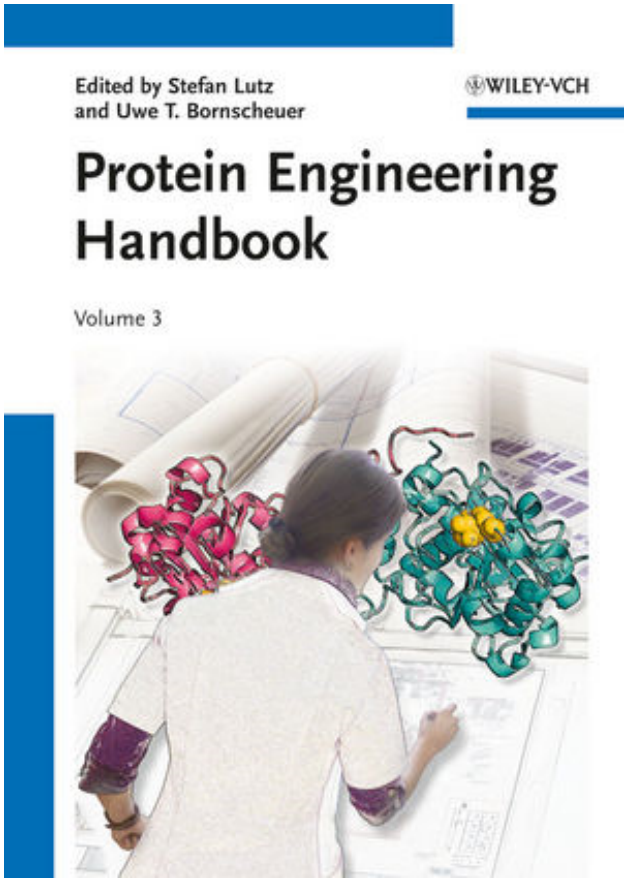
Lecture #13: **Enzyme Engineering and Design**

Enzyme ENGINEERING and DESIGN

- exciting developments in protein structure is ability to design and construct new proteins or enzymes with novel functions—"*Bionic Biochemistry*"

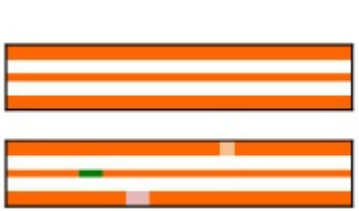
1. Protein Engineering

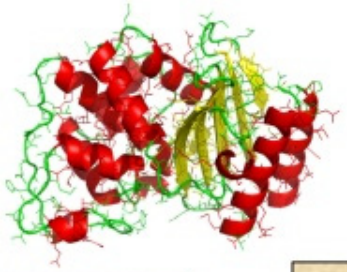
- Technology that alters protein structures to improve their properties in applications such as pharmaceuticals, environmental chemistry and biofuels
- **Challenge is to build accurate models** to predict the best substitutions to prepare candidates to insert into the parent protein
- Involves a combination of experimental and *in silico* predictions to build the best model

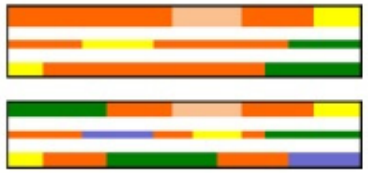


Lecture by David Baker entitled "Introduction to Protein Design" (2015)
https://www.youtube.com/watch?v=0LetJMbu7uY

Protein Engineering







Parent sequences

Scaffold/Active site identification

In silico combinatorial library

Score substitutions

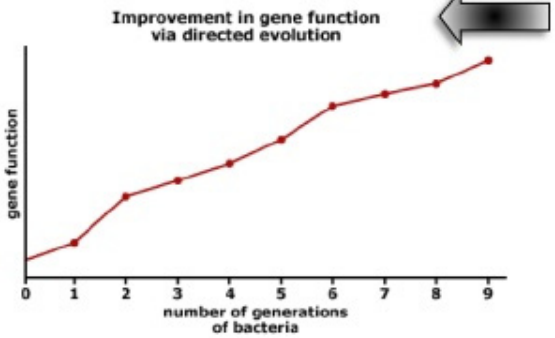
$$y = \sum_{i=1}^{N} c_i x_i + y_0$$

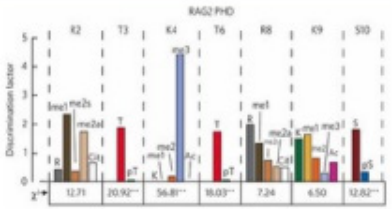
PROTEIN ENGINEERING CYCLE

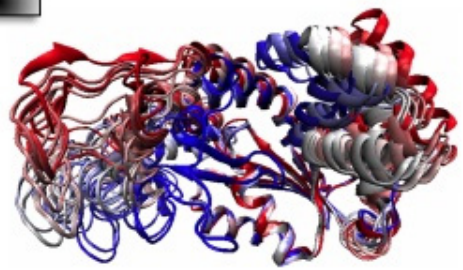
Directed evolution

Ensemble of ranked variants

Protein engineering and screening systems

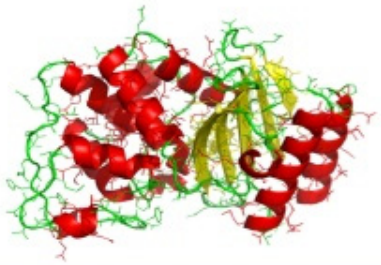






Choosing the Right Strategy

Rational Protein Design



Directed Evolution

Knowledge about structure and mechanism

Molecular modeling	Structural variation and combination	Simultaneous saturation mutagenesis	Random mutagenesis
Site-directed mutagenesis	**Shuffling** **Circular permutation**	**CASTing** **Cassette mutagenesis**	**epPCR**

Size of the screening

Manual screening	HTS of medium-sized libraries	HTS of large libraries

Distribution of mutations in variants

Single amino acid substitutions	Hybrids and structural changes	Mutations distributed over the gene

2. Overview of Protein Engineering Technology

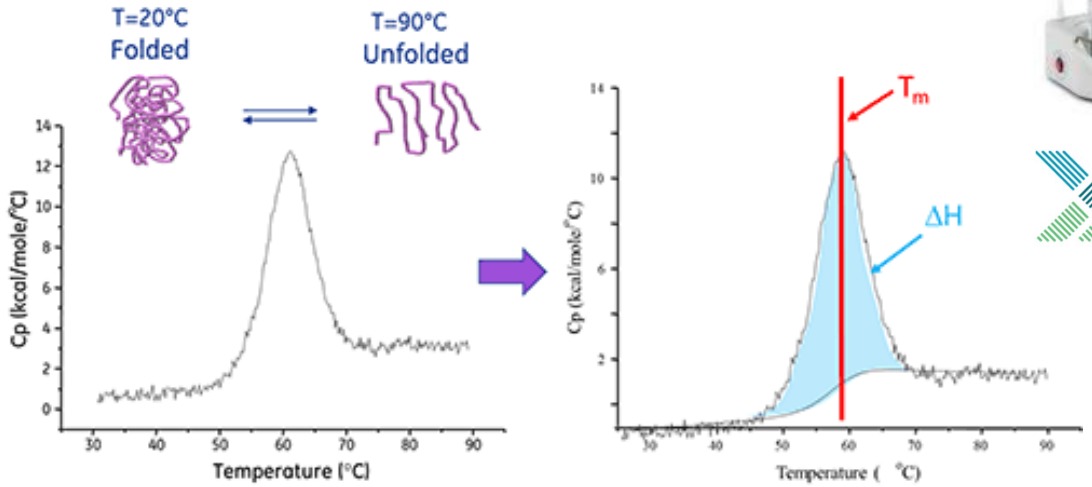
Goals:

- Increased catalytic function compared to parent enzyme
- Altered specificity, stereospecificity or affinity with interacting partners
- Increased stability

Property	**Parameters**
Thermostability	T_m, T_{50}
Catalytic activity	k_{cat}, K_M, k_{cat}/K_M
Binding specificity	K_D, K_i
Binding affinity	$\Delta G = -RT\ln(K_D^{-1})$

Goal 1: Increasing protein (enzyme) thermostability

- Thermostability quantifies the ability of a protein secondary and tertiary structures to withstand high temperatures (avoiding denaturation)
  - Typically measured by **T_m value** (calorimetry)



  - **T_{50}** is the temperature at which 50% of the proteins are inactivated in 10 min
  - Increasing thermostability is the **<u>first step</u>** in protein engineering

Goal 2: Evaluating enzyme efficiency

- k_{cat} Turnover number
- K_M Michaelis constant
- k_{cat}/K_M Specificity constant

- For an enzyme acting on two substrates S_A, S_B at rates v_A and v_B

$$\frac{v_A}{v_B} = \frac{k^A{}_{cat}/K_{M_{(A)}}[S_A]}{k^B{}_{cat}/K_{M_{(B)}}[S_B]}$$

- At $[S_A] = [S_B]$, then k_{cat}/K_M provides a measure of

substrate promiscuity efficiency

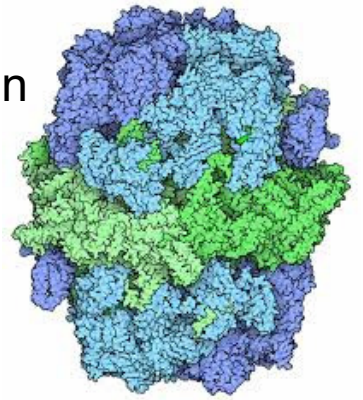
Goal 3: Protein binding affinity and specificity

- **Protein-ligand (substrate) binding:** interaction with a small molecule
  - Drug target
  - Enzyme-substrate or effector molecule

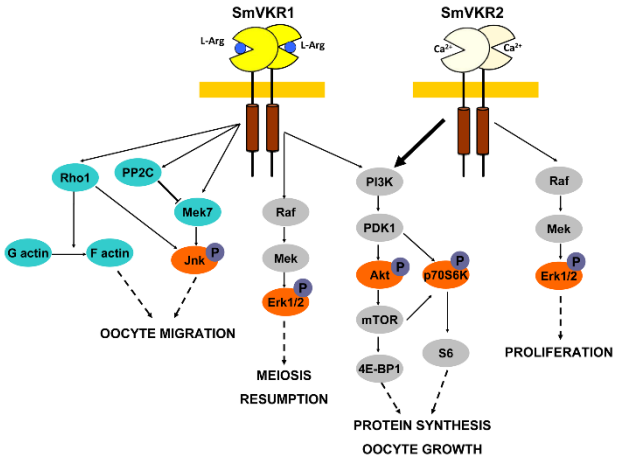
- **Protein-protein interaction:** interaction with protein partners or subunits in protein complexes

  - Permanent: in multi-unit proteins
    - Structural or functional role

  - Transient: in signalling, transport and regulation



Fatty acid synthase



3. Protein Engineering Methods

1. **Site-directed mutagenesis**

    - Point mutation, deletion or insertion

2. **Error-prone PCR (ePCR)**

    - Modifications of standard PCR methods
    - Designed to enhance the natural error rate of DNA polymerase

3. **Recombination and DNA shuffling**

    - Natural approach to making multiple mutations is recombination
    - DNA shuffling: perform functional domain or motif shuffling *in vitro*

4. Engineering proteins to be more stable

- Stability of the folded structure depends on many factors

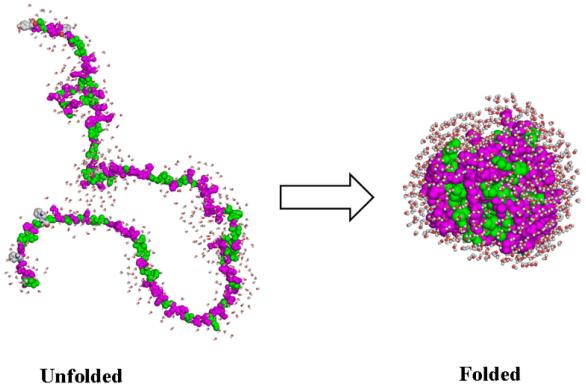
1). Unfavourable conformational entropy change upon folding

- folding is entropically unfavorable = loss of ΔS

- entropy of random coil (denatured protein) is >> than the folded native state

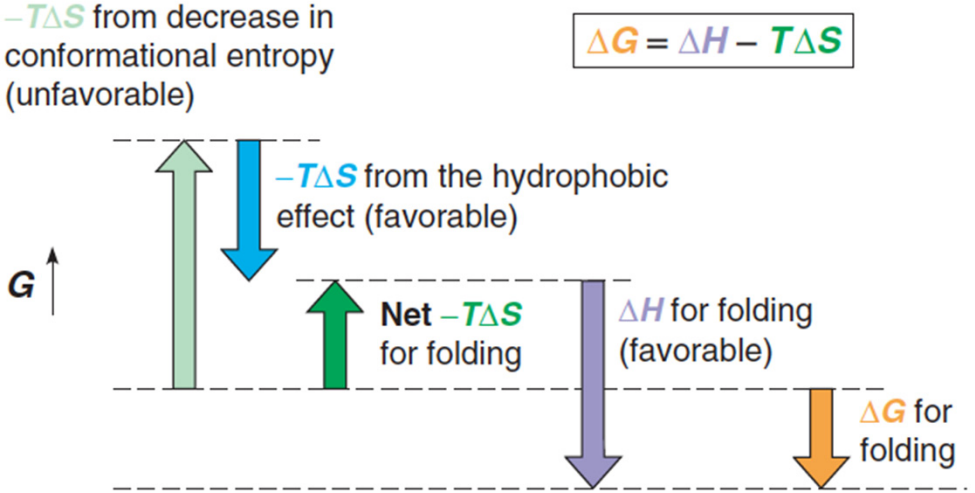
favorable → unfavorable

$$\Delta G_N = \Delta H - T\,\Delta S$$



2). Entropy loss is compensated by favorable enthalpic contacts

- hydrophobic contacts
- H-bonds
- disulfide (S-S) bonds
- van der Waals contacts
- electrostatic interactions



Matthews 4th ed, 2013

3). Stabilization of the secondary structure (α-helices, β-sheets)

- stronger forming residues

- helix capping

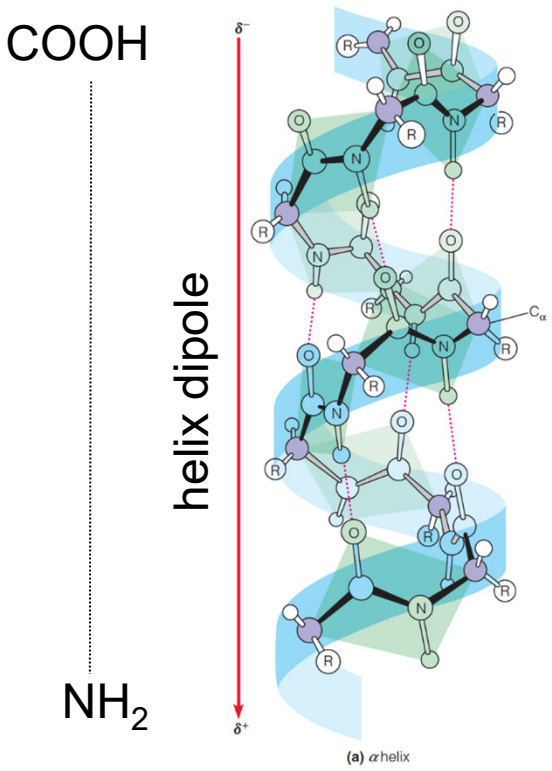


TABLE 6.8 Correspondence of amino acid residues to protein secondary structure

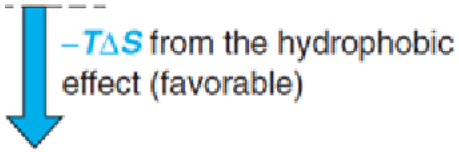
Relative probabilities of amino acid residue occurrence in different globular protein secondary structures[a]

Amino Acid	α Helix (P_α)	β Sheet (P_β)	Turn (P_t)	
Ala	1.29	0.90	0.78	Favor α helices
Cys	1.11	0.74	0.80	
Leu	1.30	1.02	0.59	
Met	1.47	0.97	0.39	
Glu	1.44	0.75	1.00	
Gln	1.27	0.80	0.97	
His	1.22	1.08	0.69	
Lys	1.23	0.77	0.96	
Val	0.91	1.49	0.47	Favor β sheets
Ile	0.97	1.45	0.51	
Phe	1.07	1.32	0.58	
Tyr	0.72	1.25	1.05	
Trp	0.99	1.14	0.75	
Thr	0.82	1.21	1.03	
Gly	0.56	0.92	1.64	Favor turns
Ser	0.82	0.95	1.33	
Asp	1.04	0.72	1.41	
Asn	0.90	0.76	1.23	
Pro	0.52	0.64	1.91	
Arg	0.96	0.99	0.88	

Matthews 4th ed, 2013

4). Hydrophobic effect

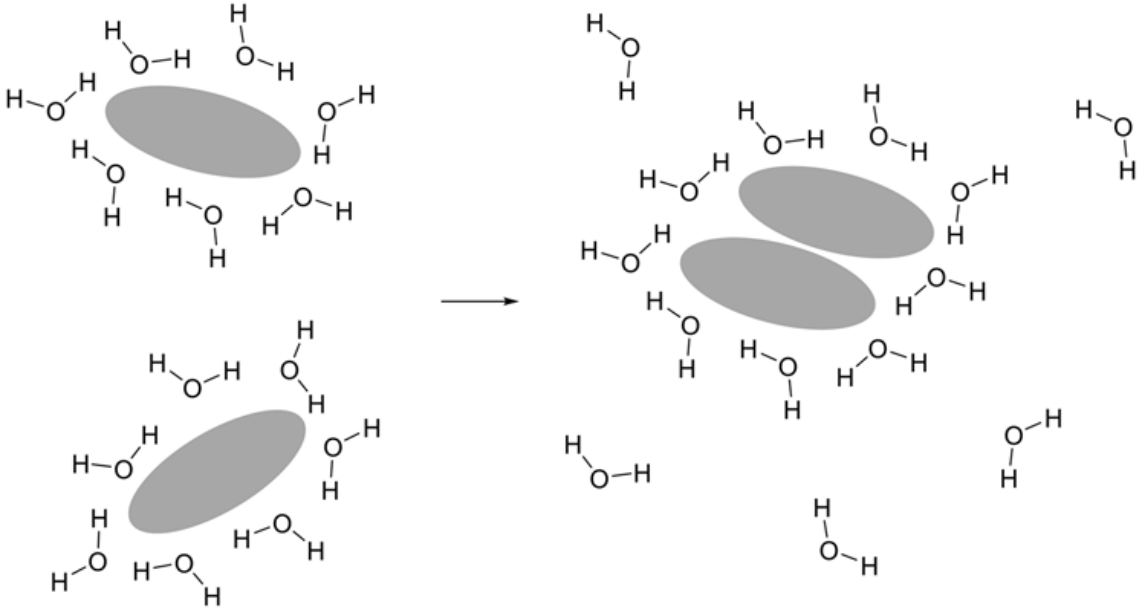
- Favourable entropy change arising from burying hydrophobic groups within the protein

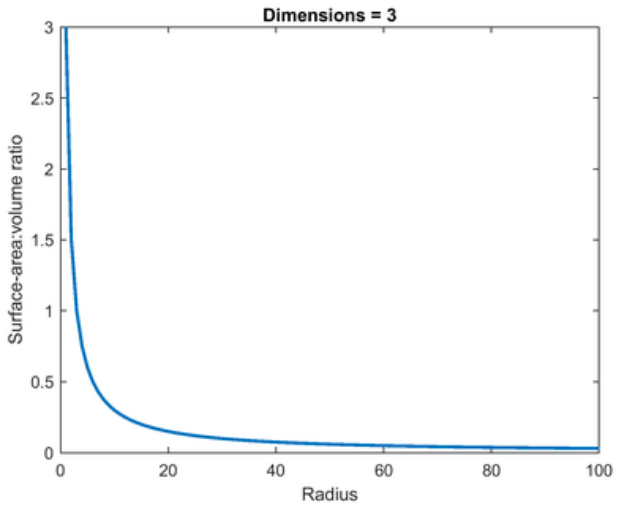


- The greater the hydrophobic buried core of a water-soluble protein, the greater will be the stability

- Methods to increase the number of hydrophobic buried residues in the protein can increase the stability

Sphere: $V = 4/3\ \pi r^3$
$S.A. = 4\pi r^2$

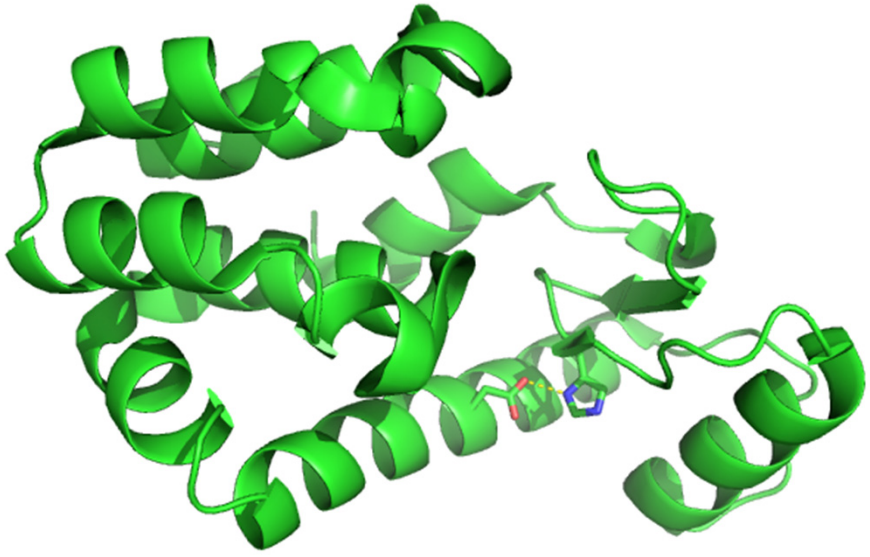




5. Engineering a more stable T4 lysozyme

Brian Matthews
HHMI
Univ. of Oregon

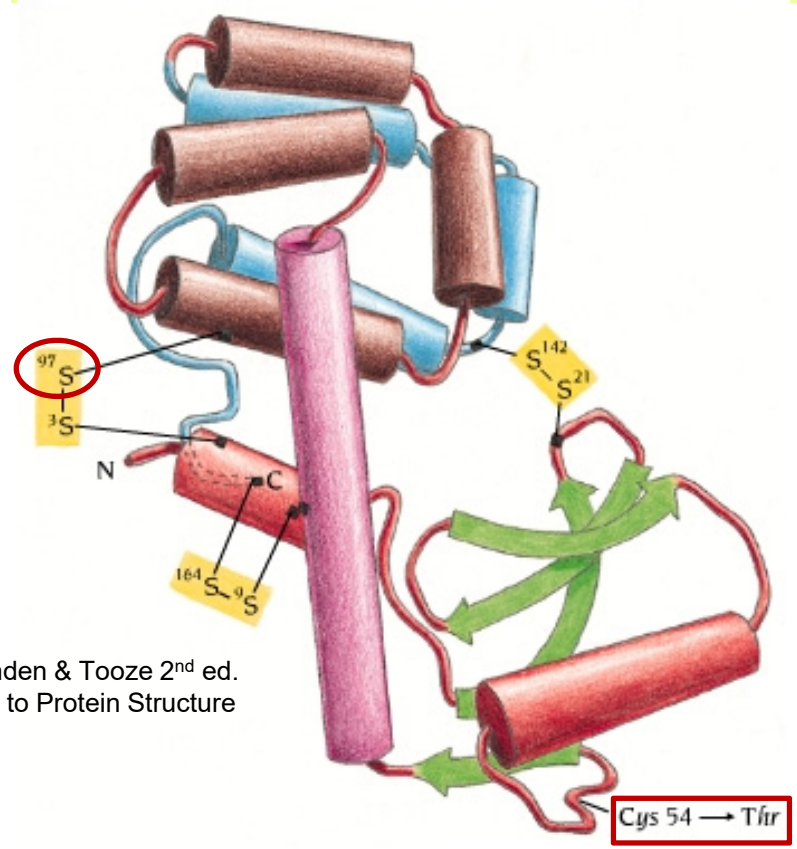
- disulfide bridges
- helix capping



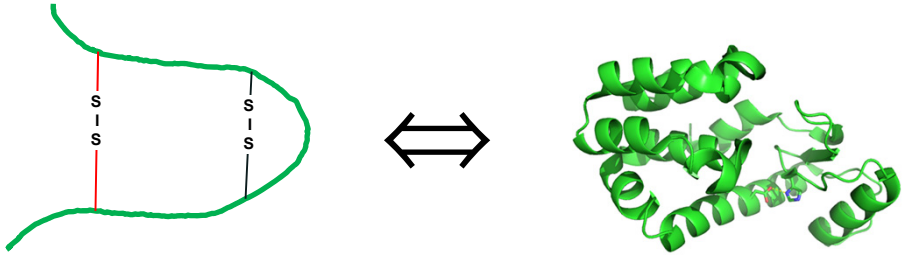
PDB:1LYD

- T4 bacteriophage infect *E. coli*
- Helps to release mature phage particles from the host cell wall by digesting the peptidoglycan
- Hydrolyzes 1,4-beta linkages between N-acetyl-D-glucosamine and N-acetylmuramic acid

6. Disulfide bridges increase T4 lysozyme stability



Branden & Tooze 2nd ed.
Intro to Protein Structure

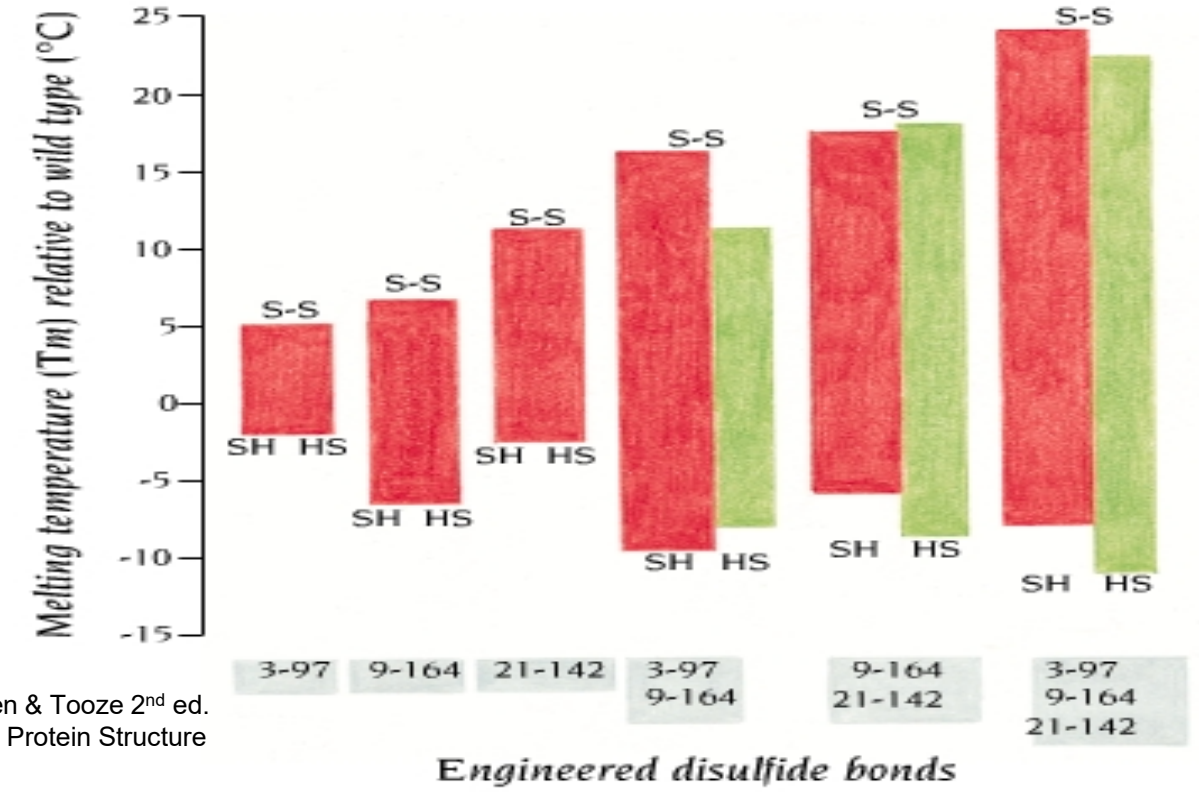


Unfolded
T4 lysozyme

folded
T4 lysozyme

Disulfide bridges

- decrease the **number of unfolded conformations** by introducing novel disulfide bridges and the **more stable** will be the newly engineered variant protein
- the **longer** the loop between the 2 Cys residues, the more **restricted** is the unfolded polypeptide
  - folded structure will be more stable
- must be introduced in **proper location** otherwise introduce **strain** into the structure, which will **decrease the stability** of the protein structure

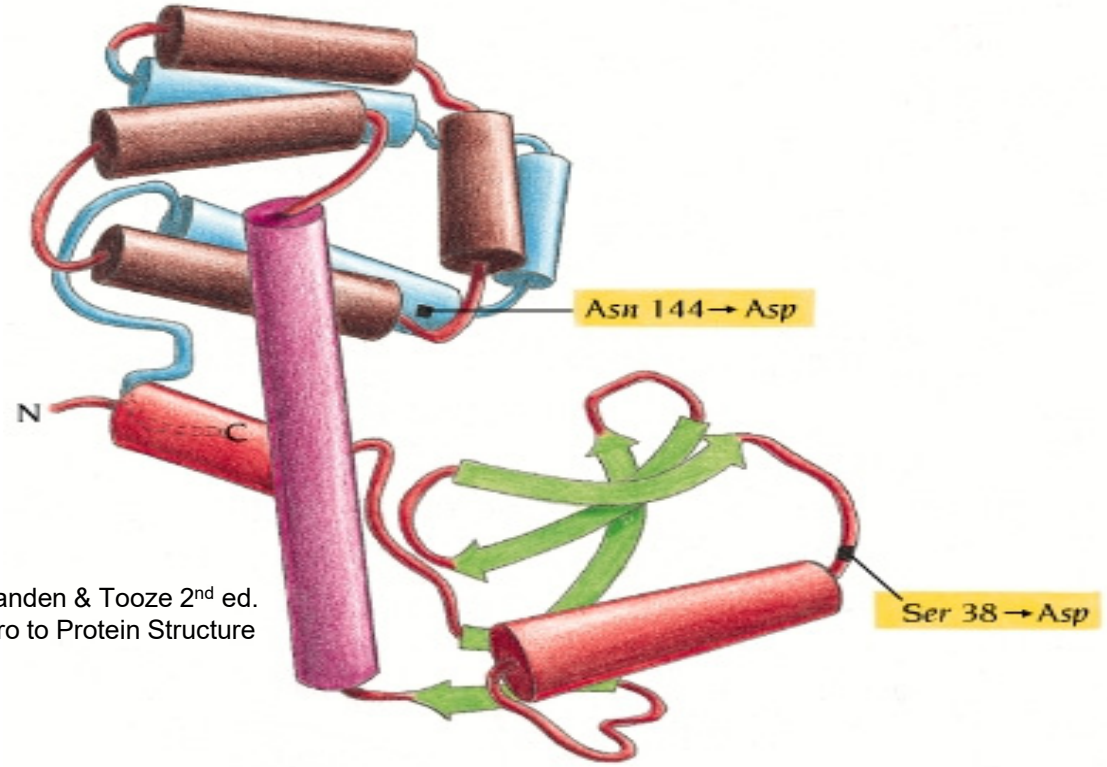


Branden & Tooze 2nd ed.
Intro to Protein Structure

- Matthews substituted Ile-3, Ile-9, Thr-21, Thr-142, Leu-164 with Cys (separate experiments)
- various combinations of single, double, and triple -S-S- bridges
- Matthews increased the stability of T4 lysozyme by 23°C for the triple variant

7. Helix stabilization (capping) of T4 lysozyme

- Two different variants with single residue substitutions at the N-terminus of two separate helices
  - Ser38 to Asp
  - Asn144 to Asp
  - Double variant (S38D/N144D)
  - Single variants showed T_M increase of 2°C
  - Double variant showed T_M increase of 2 + 2 = 4°C



Branden & Tooze 2nd ed.
Intro to Protein Structure

8. Incorporation of Non-Canonical Amino Acids (NCAA) into Proteins

ACS Biol. (2018) **13**, 854-70. Playing with the molecules of life. Peter G Schultz Scripps Res. Inst.

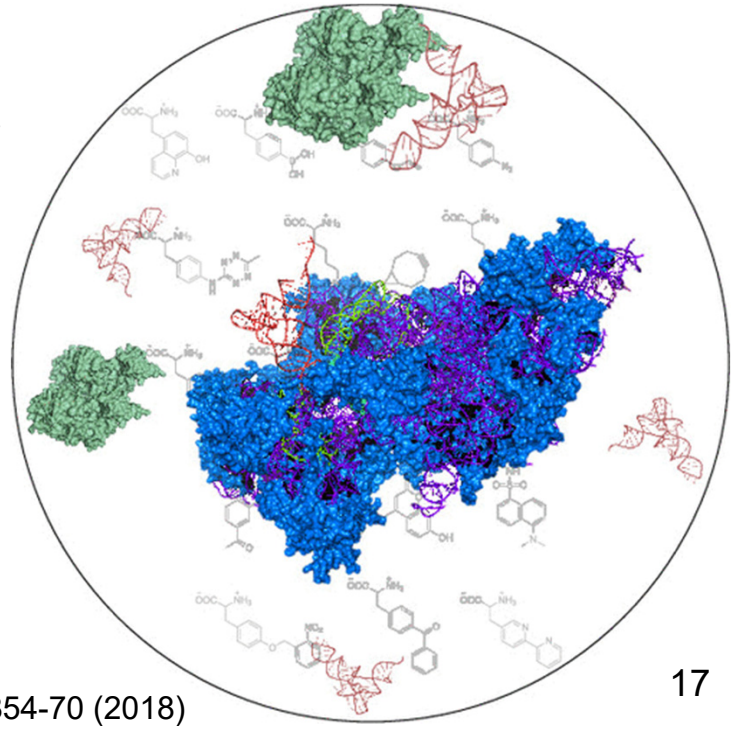
Peter G Schultz

2020 Nobel Prize??
505 pubs,
63 in *Science*
17 in *Nature*

- used to incorporate an amino acid residue that has novel chemistry not found in the naturally occurring amino acids

Applications

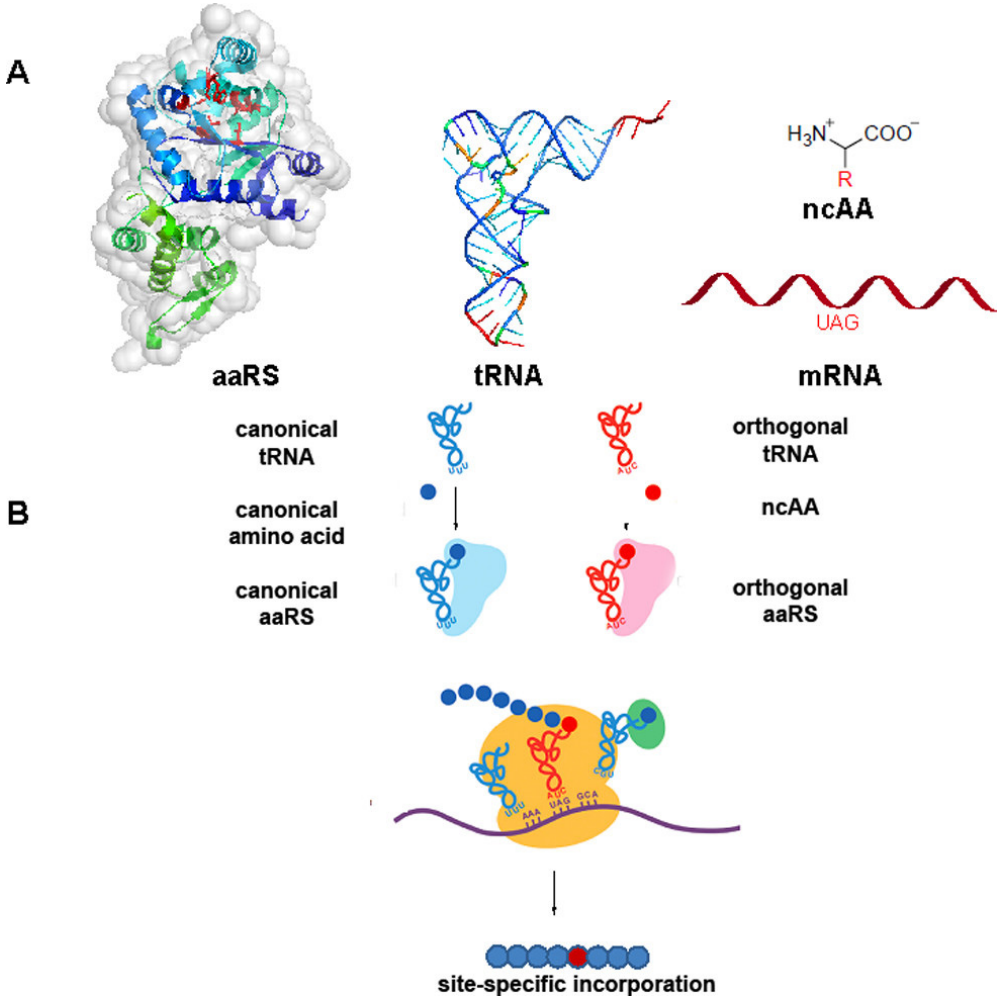
- fluorescent reporter groups → protein dynamics, conformation, interactions, etc.
- photo-reactive groups → capturing catalytic events, trapping conformations, etc.
- new chemistry for enzymes → enhance rates or provide for new reactions



ACS Biol **13**, 854-70 (2018)

Playing God: Expansion of the genetic code

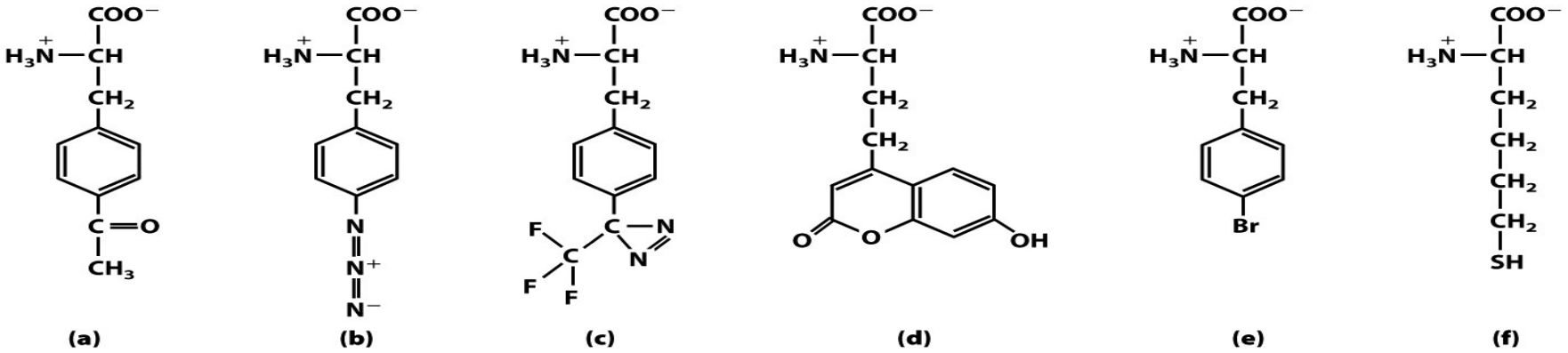
Protein



- An orthogonal tRNA/amino acyl tRNA synthetase pair and ncAA are added and encoded by a nonsense or frameshift codon in the mRNA
- The ncAA is incorporated into the protein at the nonsense site

ACS Chem. Biol. 2018, 13, 4, 854-870

Examples of Unnatural Amino Acids



Box 27-3 figure 2
Lehninger Principles of Biochemistry, Fifth Edition

Lecture by Peter G Schultz entitled “Synthesis at the Interface of Chemistry” in 2016

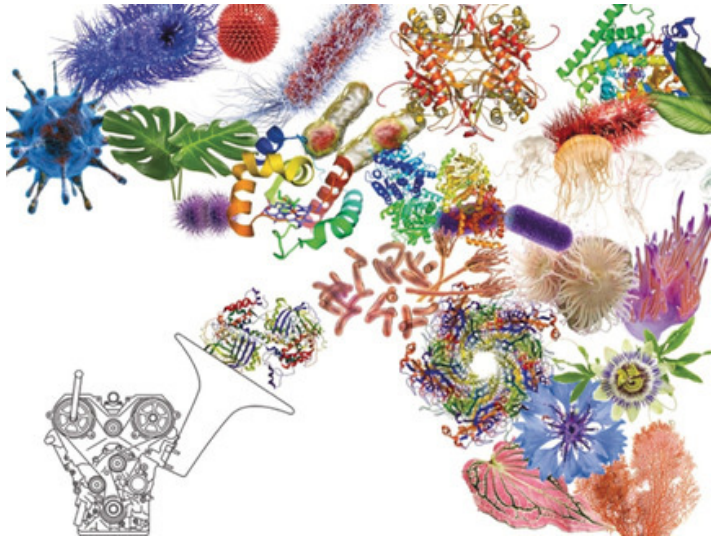
https://www.youtube.com/watch?v=X2w0FzAqm_Q

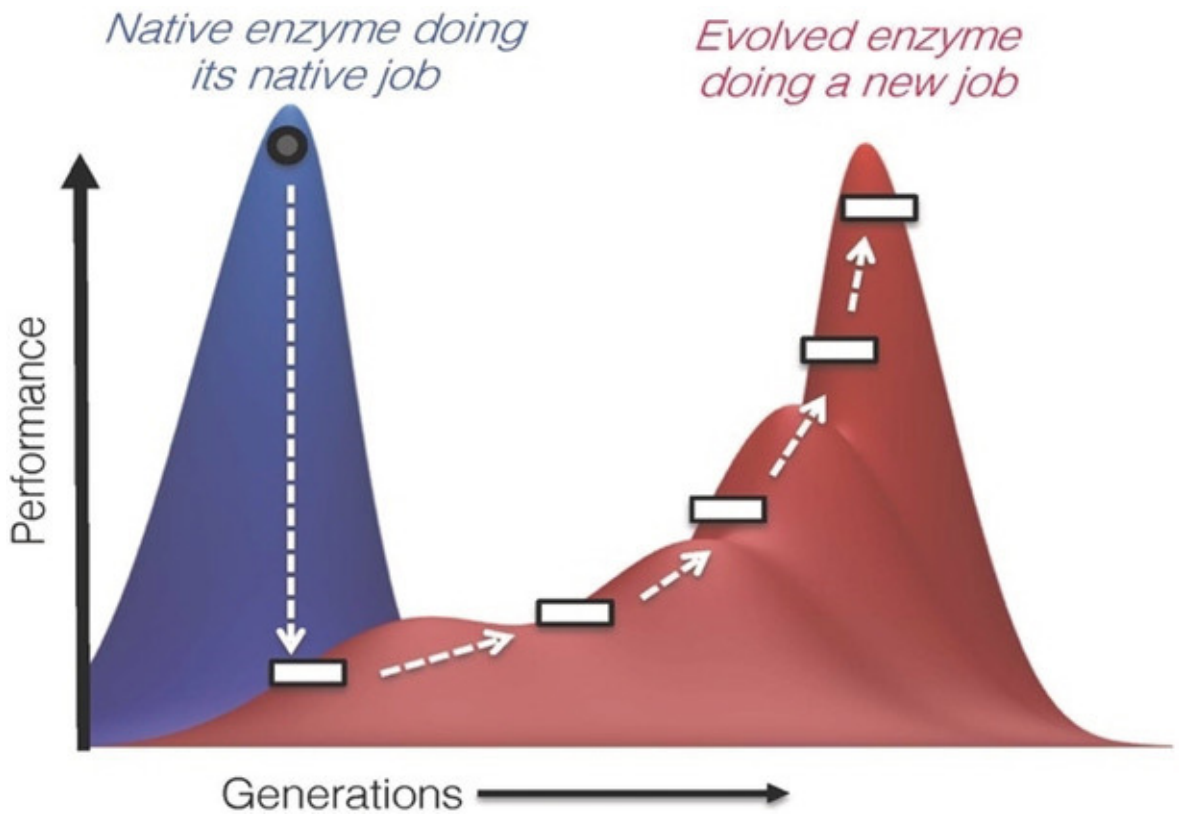
9. Protein Engineering by Combinatorial Methods

- also known as "**Directed Protein Evolution**"
- combinatorial methods invented to analyze libraries of related proteins
- sorting libraries can select a **small number of active proteins** from millions of inactive variants
- in nature random DNA mutations ⇒ changes in protein structure are **rare**
- combinatorial methods **<u>accelerate evolution!</u>**
- approach is to use **random mutagenesis**
  - oligonucleotide-directed mutagenesis
  - error-prone PCR
  - mutagenic strains
  - DNA shuffling
- mutagenesis must be followed by

a selection process for the **desired function**

Frances Arnold, CalTech
Nobel Prize in Chemistry
2018



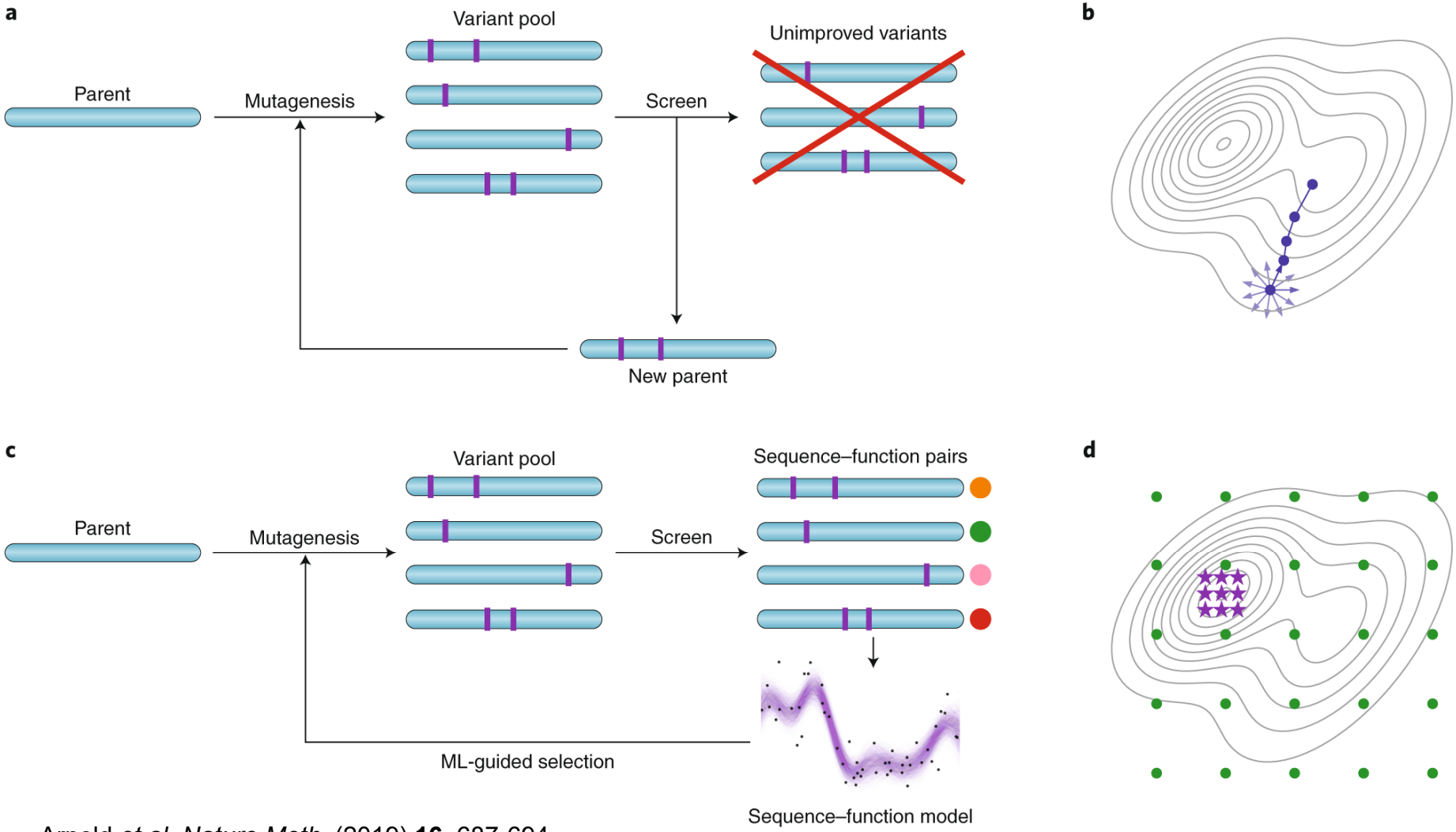


- An enzyme whose function is optimized for its native job performs **poorly** in a new role
- Directed evolution via rounds of mutation and screening can discover changes in sequence that improve performance climbing to a new "fitness" peak

Francis Arnold Nobel Lecture 2018: "Innovation by evolution: Bringing new chemistry to life"

https://www.youtube.com/watch?v=6hOZ5e0g9Uo&t=1s

10. Machine-learning-guided directed evolution for protein engineering



Arnold *et al. Nature Meth.* (2019) **16**, 687-694

(1) Directed evolution uses iterative cycles of diversity generation and screening to find improved variants

- information from unimproved variants is discarded

(2) Directed evolution is a series of local searches on the function landscape

(3) Machine learning (ML) methods use the data collected in each round of directed evolution to choose which mutations to test in the next round

- careful choice of mutations to test decreases the screening burden and improves outcomes

(4) ML guided directed evolution often rationally chooses the initial points to maximize the information learned from the function landscape

- allows future iterations to quickly converge to improved sequences