Phylogenetic Tree of Life

Bacteria
- Spirochetes
- Proteobacteria
- Cyanobacteria
- Planctomyces
- Bacteroides
- Cytophaga
- Thermotoga
- Aquifex

Archaea
- Green filamentous bacteria
- Methanobacterium
- Methanococcus
- Thermoproteus
- Pyrodictium
- T. celr

Eucaryota
- Entamoeba
- Halophiles
- Slime molds
- Animals
- Fungi
- Plants
- Ciliates
- Flagellates
- Trichomonads
- Microsporidia
- Diplomonads
**Thermophilic organism**

**Thermophiles** are organisms that grow and thrive at temperatures (60 – 80°C) that are often too high for **mesophiles**—most thermophiles are Archaea.

Some organisms grow at even higher temperatures and pressure (> 80 °C)—these are called **hyperthermophiles**—all hyperthermophiles are Archaea.

- e.g. Pyrococcus furiosus lives at 100 °C

Other extremophiles include organisms that live at extreme pH values, pressure, high doses of ionizing radiation, desiccating conditions, salt concentration, etc.

Hyperthermophiles of Yellowstone National Park
Thermozyme

Enzymes isolated from (hyper) thermophiles are called **thermozymes**

Thermozymes have found applications in biotechnology where their unusual thermal stability makes them uniquely qualified for tasks under conditions that would easily denature mesophilic enzymes

**Polymerase chain reaction**

invented by Kary Mullis (Chemistry Nobel prize 1993)  
a biochemical reaction used to amplify nucleic acids  
requires cycling between high (~ 95 °C) and medium (~ 50 – 75 °C) temperatures  
enzymes that are active at high temperature for many hours  
Taq (Thermus aquaticus), Pfu (Pyrococcus furiosus)  
Deep Vent (half life 23 hr at 95 °C) polymerases
Inside the PCR reaction tube...

0) 55°C Before Cycle

1) 90°C DNA Denaturing

2) 55°C Primers anneal

3) 72°C Polymerization

Hint → an arrow like this implies that the arrowhead is the 3' side of the primer. This is the direction that the polymerization will take place.

Your primers must be complementary to one of the two DNA strands and the 3' sides must orient towards one another in order for the reaction to yield a two-strand product.
Why study thermophilic proteins

Extremely high thermostability can be achieved in some proteins through a combination of van der Waals, H-bond, and electrostatic interaction.

Evaluate what factors contribute to their high stability and thus learn to correlate **structural** properties with **thermodynamic** properties.

Practical concerns: When engineering proteins for **therapeutic** use, we want to maximize (thermal) stability while maintaining activity.
General observations

(Hyper)Thermophilic proteins are highly homologous to their mesophilic counterparts both in sequence (40 – 85%) and structure.

The catalytic mechanism is conserved in both classes of proteins.

Thermophilic proteins expressed in mesophilic organisms are functional—i.e. these proteins are intrinsically thermostable and do not require extrinsic factors such as glycosylation.

Genes from hyperthermophilic archaea can complement yeast mutations.

(green) E.coli CheY
(pink) T. maritima CheY

CheY: response regulator in the chemotaxis pathway.
Fundamentals of protein stability

Stabilization may require contributions from all levels of hierarchy: amino acid identity, packing, secondary and super secondary structure, domain and subunit.

Free energy of folding changes parabolically as a function of temperature.

Need a mechanism of stabilizing a protein without raising the temperature at which it becomes fully active.
Computer simulation

Rubredoxin is a small model protein containing a Fe-based redox center that participates in electron transfer.

Small size renders the protein amenable to computational studies.

Unfolding of rubredoxin from mesophilic and hyperthermophilic organisms has been simulated using MD.

Lazaridis et al, Protein Sci 6, 2589 (1997)
Lactate dehydrogenase (LDH)

Comparative study of homologous proteins from mesophilic and thermophilic organisms can help reveal the mechanism by which thermal stability of a protein is achieved.

LDH catalyzes the last step in anaerobic glycolysis by converting pyruvate to lactate.

High resolution structures are available for homologous LDH from different organisms:

- Tm: *Thermotoga maritima*
- Bs: *Bacillus stearothermophilus*
- Bl: *Bifiobacterium longum*
- Lc: *Lactobacillus casei*
- Ss: *Sus scrofa*
- Sa: *Squalus acanthia*

Auerbach et al. Structure 6, 769 (1998)
<table>
<thead>
<tr>
<th></th>
<th>Tm</th>
<th>Bs</th>
<th>Bl</th>
<th>Lc</th>
<th>Ss</th>
<th>Sa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth temperature (°C)</strong></td>
<td>50-90</td>
<td>40-65</td>
<td>37-41</td>
<td>30-40</td>
<td>37</td>
<td>20</td>
</tr>
<tr>
<td><strong>Hydrophobic/charged surface area</strong></td>
<td>0.57</td>
<td>0.66</td>
<td>0.76</td>
<td>0.73</td>
<td>0.89</td>
<td>0.86</td>
</tr>
<tr>
<td><strong>Intrasubunit ion pairs (4 Å cutoff)</strong></td>
<td>17</td>
<td>16.5</td>
<td>8</td>
<td>8</td>
<td>10.5</td>
<td>13</td>
</tr>
<tr>
<td><strong>Alpha helix (%)</strong></td>
<td>46</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>40</td>
<td>=</td>
</tr>
<tr>
<td><strong>Beta strand (%)</strong></td>
<td>23</td>
<td>22</td>
<td>19</td>
<td>20</td>
<td>19</td>
<td>=</td>
</tr>
</tbody>
</table>

**Tm LDH** *(Thermotoga maritima, most stable)*
- has the fewest and smallest cavities
- has the most intrasubunit ion pairs (salt bridges)
- has the highest percentage of Arg
- increased hydrophobic intersubunit interaction (three Phe per monomer)

No significant increase in the number of “intersubunit” ion pairs
No clear correlation with the core packing (fraction of atoms with zero ASA)
No correlation with the volume of the monomer and tetramer
No correlation with the surface to volume ratio
No correlation with the abundance of Lys/Arg near the C-terminus of helix or of Asp/Glu near the N-terminus
No correlation with the number of hydrogen bonds in the monomer and tetramer
Citrate synthase

\[\text{acetyl-CoA} + \text{oxaloacetate} + \text{H}_2\text{O} \rightarrow \text{citrate} + \text{CoA-SH}\]

The structure of citrate synthase from organisms that live at five different temperatures has been determined.

- **31 °C**
- **37 °C**
- **55 °C**
- **85 °C**
- **100 °C**

<table>
<thead>
<tr>
<th>Source organism</th>
<th>Optimum growth temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Arthrobacter DS23R</em></td>
<td>31 ^a</td>
</tr>
<tr>
<td>Pig</td>
<td>37</td>
</tr>
<tr>
<td><em>Thermoplasma acidophilum</em></td>
<td>55</td>
</tr>
<tr>
<td><em>Sulfolobus solfataricus</em></td>
<td>85</td>
</tr>
<tr>
<td><em>Pyrococcus furiosus</em></td>
<td>100</td>
</tr>
</tbody>
</table>

All CS are alpha helical homodimeric proteins with the active site located at the interface with residues from both subunits—maintaining the dimer structure is essential for enzymatic activity.

• Bacterial CS (including ArCS) have smaller surface areas than pigCS
• Psychrophilic (cold) CS has the smallest volume

• Total amount of exposed hydrophobic area decreases as the thermostability increases
• Number and size of internal cavities decreases with thermostability
• Hyperthermophilic CS has the highest percentage of buried atoms

• Conflicting results for the number of ion pairs
  – psychrophilic has the most overall
  – thermophilic has the most interfacial ion pairs
  – mesophilic has more intrasubunit interactions

• Loops are stabilized through ionic interactions in thermophilic and psychrophilic CS
• Loop length does not correlate with thermostability—for this example
No simple rules for thermostability

Thermostability cannot be explained in terms of amino acid composition as the trends are weak and there are too many exceptions.

Distribution of the residues and their interactions in the protein are important.

No universally applicable generalizations have been found.

Factors that contribute toward thermostability include:
- higher oligomeric state to modulate surface to volume ratio
- increased number of stabilizing interactions
- reduction in size of surface loops to minimize entropy
- increased hydrophobic interaction (e.g. burial of hydrophobic residues)
- remove factors that could lead to irreversible modifications—e.g. fewer Asn residues as Asn may catalyze peptide chain hydrolysis
Possible mechanisms of thermostability

- More stable protein core
  - reduction in number and volume of cavities
  - increased number of hydrophobic residues because hydrophobic effects increase with temperature
- Number of hydrogen bonds
- Tendency to form higher order oligomers
  - optimized intersubunit contacts within oligomers
- Increased secondary structure
  - stabilization of alpha helices using residues with higher helix propensity
  - redesign of the helix caps with Gly
  - but contribution from beta sheets is uncertain
  - number of helix and sheet doesn’t correlate with stability
- Decrease in conformational degrees of freedom
  - increased number of Pro
- Improved electrostatic interactions, i.e. salt bridges
  - intrahelical, interhelical, intrasubunit, intersubunit, surface, buried

Petsko, Meth Enzy 334, 469 (2001)
Evolution of thermostability

How did thermostability evolve?

what are potential physical mechanisms and how do organisms apply these mechanisms to adapt

Some organisms started inhabiting hot habitats from the very beginning

mostly archae
their enzymes are more compact and hydrophobic
“designed” their proteins to probe alternate structural repertoire

Others started out as mesophilic organisms but relocated to hot environments

mostly bacteria
these organisms had to make adjustments at the amino acid level to achieve

Thermostability may be **structural** or **sequential**

– Berezovsky and Shakhnovich PNAS 102, 12742 (2005)
– MD simulated unfolding and sequence analysis
– structural: increase in overall contacts and interactions; also more compact
– sequential: a few exceptionally strong interactions, e.g. stabilizing ionic interaction
NTN hypothesis

The codon NTN codes for hydrophobic residues: F, I, V, L, M

Genomic analysis of 14 hyperthermophiles, 7 thermophiles, and 142 mesophiles shows an increase in hydrophobic residues in thermophilic organisms by three for every 100 amino acids.

An increased frequency of NTN is observed both in archae and bacteria. Use of hydrophobic residues for thermostability deviates from the mechanistic distinctions based on phylogenetic history.