Module 1 overview

*lecture*

1. Introduction to the module
2. Rational protein design
3. Fluorescence and sensors

*lab*

1. Start-up protein eng.
2. Site-directed mutagenesis
3. DNA amplification

**PRESIDENT’S DAY**

4. Review & gene analysis
5. Protein expression
6. Purification and protein analysis
7. Binding & affinity measurements
8. High throughput engineering

4. Prepare expression system
5. Gene analysis & induction
6. Characterize expression
7. Assay protein behavior
8. Data analysis
Lecture 8: High throughput engineering

I. General requirements for HT engineering
   A. High throughput vs. rational design
   B. Generating libraries

II. Selection techniques
   A. Phage display and related techniques
   B. Selection for properties other than affinity
Rational protein design:
Knowledge-based, deterministic engineering of proteins with novel characteristics

- design/modeling (often computer-aided)
- generate required DNA constructs
- express proteins
- purify proteins
- assess proteins for desired characteristics

“Irrational” high throughput protein engineering:
Selection for desired properties from libraries of random variants

- design/modeling (often computer-aided)
- generate library of DNA constructs
- express proteins
- screen proteins in high throughput assay
- assess “hits” for desired characteristics
Methods for generating mutant protein libraries:
  • site-directed mutagenesis with degenerate primers
  • error-prone PCR
  • gene shuffling

Degenerate primers

- not all combinations of AA's possible at each position
- number of combinations expands exponentially
- degenerate primers synthesized by split-pool method
- standard primer design criteria must be considered
PCR polymerase and conditions may be chosen to promote mutations

<table>
<thead>
<tr>
<th>Polymerase</th>
<th>Template doublings ($d^b$)</th>
<th>lacI$^{-}$ plaques$^b$ (% ± SD)</th>
<th>Mutation load$^c$ (per kilobase) (±SD)</th>
<th>Error rate$^d$ (per base) ($\times 10^{-6}$ ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pfu-Pol (exo$^+$)</td>
<td>12.3</td>
<td>0.61 ± 0.09</td>
<td>0.017 ± 0.002</td>
<td>1.4 ± 0.2</td>
</tr>
<tr>
<td>Pfu-Pol (exo$^-$)</td>
<td>11.8</td>
<td>20 ± 1.7</td>
<td>0.58 ± 0.05</td>
<td>49 ± 4</td>
</tr>
<tr>
<td>Taq-Pol</td>
<td>11.6</td>
<td>3.9 ± 0.16</td>
<td>0.12 ± 0.006</td>
<td>10 ± 0.5</td>
</tr>
</tbody>
</table>

error rate = mutation load ÷ template doublings

---

**normal PCR**

- Taq
- dCTP, dTTP
- dGTP, dATP
- Mg$^{2+}$

**error-prone PCR**

- Taq
- dCTP, dTTP $\uparrow$
- dGTP, dATP $\downarrow$
- Mg$^{2+}$ $\uparrow$
- Mn$^{2+}$ $\uparrow$

*some mutations are more likely than others*

---

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Pfu-Pol (exo$^-$) D473G$^a$</th>
<th>Taq-Pol (Mg$^{2+}$/ unbalanced dNTPs)$^b$</th>
<th>Taq-Pol (Mn$^{2+}$/ unbalanced dNTPs)$^c$</th>
<th>Taq-Pol (unnatural mutagenic bases)$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A$\rightarrow$T/T$\rightarrow$A</td>
<td>28</td>
<td>40.9</td>
<td>11.4</td>
<td>0.2</td>
</tr>
<tr>
<td>A$\rightarrow$C/T$\rightarrow$G</td>
<td>7.4</td>
<td>7.3</td>
<td>3.3</td>
<td>8.4</td>
</tr>
<tr>
<td>A$\rightarrow$G/T$\rightarrow$C</td>
<td>19.2</td>
<td>27.6</td>
<td>60.9</td>
<td>78.3</td>
</tr>
<tr>
<td>G$\rightarrow$A/C$\rightarrow$T</td>
<td>22</td>
<td>13.6</td>
<td>18.1</td>
<td>13.2</td>
</tr>
<tr>
<td>G$\rightarrow$C/C$\rightarrow$G</td>
<td>7.3</td>
<td>1.4</td>
<td>4.3</td>
<td>0.7</td>
</tr>
<tr>
<td>G$\rightarrow$T/C$\rightarrow$A</td>
<td>10.3</td>
<td>4.5</td>
<td>1.8</td>
<td>0.0</td>
</tr>
<tr>
<td>Insertion</td>
<td>2.9</td>
<td>0.3</td>
<td>Not given</td>
<td>$\sim$0</td>
</tr>
<tr>
<td>Deletion</td>
<td>2.9</td>
<td>4.2</td>
<td>Not given</td>
<td>$\sim$0</td>
</tr>
</tbody>
</table>
Gene shuffling techniques mimic diversity due to meiotic recombination:
  • fragments of homologous genes combined using “sexual PCR”
  • diversity may arise from error prone PCR or multiple genes

How are libraries of mutant proteins screened?

All major methods include a strategy to keep DNA sequence info associated with the proteins that are being screened.

Phage display is a versatile high throughput method to do this:

protein “displayed” on the coat of a bacteriophage, by fusing to a natural phage coat protein

M13 phage (contains DNA)
Application: phage-displayed peptides that bind to GaAs

selected sequences

phages patterned on target substrate

Yeast display: similar to phage display, but with proteins fused to a *Saccharomyces* cell wall protein (DNA in yeast)

What would you expect advantages to be, compared with phage display?

In this example, a population of scFvs was screened for binding to an antigen

*left*: selection criterion for FACS assay

*right*: comparison of wt (blue) and selected (red) scFv binding

Ribosome display: mRNA and synthesized proteins held together non-covalently on a ribosome

What are advantages of this technique over phage/yeast display methods?

- screening not in the presence of large particles
- incorporation of unnatural amino acids

What about properties other than affinity?

A simple example: screen for dsRed variants with different excitation and emission wavelengths—how could this be done?

Directed evolution of enzymatic activity: screen is a fluorescence assay

Which type of screening method to use?

<table>
<thead>
<tr>
<th>screen method</th>
<th>throughput</th>
<th>other notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SELEX</td>
<td>$10^{15}$</td>
<td>selection of DNA/RNA</td>
</tr>
<tr>
<td>ribosome display</td>
<td>$10^{15}$</td>
<td>in vitro protein synthesis</td>
</tr>
<tr>
<td>phage display</td>
<td>$10^{11}$</td>
<td>best for small proteins/peptides</td>
</tr>
<tr>
<td>yeast display</td>
<td>$10^8$</td>
<td>compatible w/eukar. proteins</td>
</tr>
<tr>
<td>plate assays</td>
<td>$&lt;10^5$</td>
<td>versatile but more complex</td>
</tr>
</tbody>
</table>

number of variants in a protein library

- $x$ residues = $20^x$ possible variants
- 12 residues = $4 \times 10^{15}$ variants

lesson: impossible to cover sequence space except with short sequences (or few positions) and only the most high throughput techniques