ADVANCED ENZYME TECHNOLOGY

PROTEIN ENGINEERING (PART 1)

by
AIZI NOR MAZILA RAMLI
Faculty of Industrial Sciences & Technology
aizinor@ump.edu.my

Chapter Description

- Expected Outcomes
  - Describe about the general principal of protein engineering
  - Discuss about the strategies of protein engineering
  - Identify the differences between each strategies

CONTENT

- General introduction
- Strategies of protein engineering
  - *De novo* design
  - Rational design
    - Site directed mutagenesis
      - Overlap extension PCR mutagenesis
  - Directed evolution
    - Random mutagenesis
      - Error-prone PCR

Introduction

- Commonly, enzymes usage in chemical industries and other industrial applications are applied when **extremely specific catalysis are required**.

- However , in general enzymes are usually:
  - **Restricted to the amount of reactions they are able to catalyse**
  - **Reduced stability in both organic solvents and at elevated temperatures**.

- Therefore, **protein engineering** become a dynamic research area which consist of **efforts to design new enzymes with novel properties**, either through rational design or in vitro evolution.

- Protein Engineering : **Design** and **construct** proteins with **desired function**(s)

Protein X → Protein Y

↓

with improved or novel properties

Strategies of protein engineering

- *De novo* design
- Rational design
- Directed evolution

De novo design

"Knowledge-based protein design"

- **enzymes with novel catalytic activities** are **still required** to be applied in numerous industrial biocatalytic processes.
- The ultimate goal of de novo protein design is to **develop catalytic activity** such that a biocatalyst can be readily obtained for any given chemical transformation.
- A **extremely challenging approach**, offering the broadest possibility for new structure.
- Aim: to discover the best amino acid sequence to guarantee folding in a selected structure (a-helical bundle or a/β-barrel)

- **EXAMPLE:**
  - ✓ **Methionine**: prefers rigid segments, the central part of helical segments, and buried regions of natural proteins
  - ✓ **Threonine**: prefers flexible segments
  - ✓ **Lysine**: a good helix former and prefers at the C-terminal, etc.
  - ✓ **Leucine**: stabilizes helices, prefers buried regions, and is usually found in the middle positions in a helix
  - ✓ **Glutamine**: hydrophilic residue; favourable ion pairs along one side of the helix to stabilize helix formation
  - ✓ **Proline**: helix termination
  - ✓ **Arginine**: stimulate a reversal in the overall peptide chain direction

Rational Design

- Become the **earliest technique** of protein engineering
- Broadly applied to **introduce required properties into a protein of interest**.
- The **technique hinges on relating structure to function**, commonly through molecular modelling approach- Structure-function relationship of target protein.
- Advancements in rational design rely on **improvement in elucidation of structure, enhanced modelling procedures, and substantial new understanding into structure-function interactions**

- Generally, proteins can be designed by **creating calculated variations** on a solved structure of protein with its sequence.

- **Create predictions** of **protein-sequence** that will fold to specific structures.

- **Function of protein is highly related on its structure** thus rational design approach utilize this association to develop function by designing proteins that contain a target structure or fold.

- As a consequence, the **target structure or collective of structures must be identified** earlier in rational protein design approach.

- This is different to the other approach of protein engineering, for example directed evolution, where the structure information is not necessary and a various techniques are employed to identified proteins that attain a specific function.

- After that, the predicted sequences generated can be **tested experimentally**.

Hypothesis : original sequence can be revised to optimized a desired function of protein

Which suggests that:

- Protein is **NOT** at an optimum for that function
- Sequence changes **without disruption of the structure** (otherwise it would not fold)
- New sequence designed is not **excessively dissimilar** from the original sequence or else lead to protein function loss.

- There are several reasons to create specific DNA modifications
  - To **investigate changes in activity of protein** that happen as a consequence of DNA manipulation.
  - To **choose or screen for mutations** (at the DNA, RNA or protein level) that contain a target characteristic.
  - to be tuned to fit into the industrial marketplace.

- Technique for rational design:

"Site-directed mutagenesis"

Site-directed mutagenesis

- a technique to **design accurate and desired changes** in double stranded DNA of plasmid.

- The most **prevailing and broadly utilized** rational design approach, which **involves the accurate modification of specific residue(s)** in a target protein at the DNA level.

- These residue(s) are **usually recognized from the three-dimensional protein structure** (NMR or x-ray crystallography)

- **If the target protein structure is not available,**
  - the available 3D structure of a homologous protein showing high sequence identity with the protein of interest can be used to built **a structural model of the target protein (homology modelling programs)**.

- **If above options are NOT achievable**,
  - **sequence analysis programs for example NCBI BLAST can be employed to recognize the conserved residues** within a group of closely related proteins, thus suggested to be functionally important.

- Three type of site-directed mutagenesis:

Insertions	Deletions	Substitutions
• gaining one additional nucleotide	• loss of one nucleotide	• change of one nucleotide (i.e. A-> C)

- Several PCR-based mutagenesis methods have been established and one of the most extensively employed methods is the **"overlap extension PCR mutagenesis"** method.

Overlap extension PCR mutagenesis

- This strategies required two pairs of primer,
  - For each pair, one primer will comprises the mutanted codon with a mismatched sequence.

- Then, all four primers (2 pairs) will be used in the first polymerase chain reaction (PCR) step,
  - in which two PCRs will take place, and two double-stranded DNA products will be produced.

- Upon denaturation and annealing of PCR amplification, two heteroduplexes are will be formed,
  - each strand of the heteroduplex involves the desired mutagenic codon.

- During second PCR step, DNA polymerase is then utilized to complete the overlapping 3' and 5' ends of each heteroduplex and the nonmutated primer set will be used to amplify the mutagenic product.

THANK YOU