LOSCHMIDT LABORATORIES

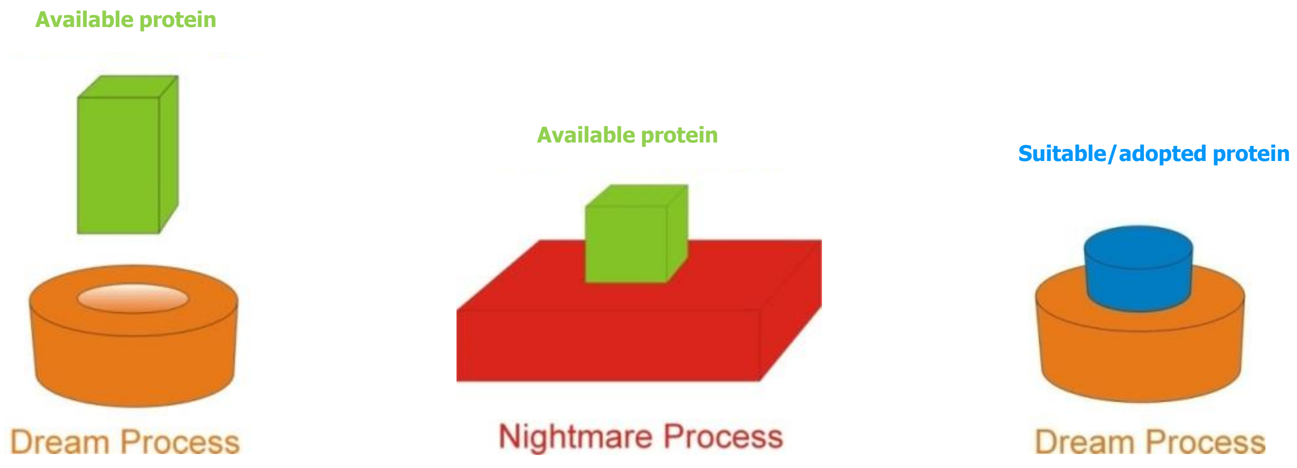
Protein Engineering

Outline

- Limitations of proteins in biotechnology processes
- Definition and aim of protein engineering
- Targeted properties of proteins
- Basic approaches in protein engineering
  - **DIRECTED EVOLUTION**
  - **RATIONAL DESIGN**
  - **SEMI-RATIONAL DESIGN**
- Examples

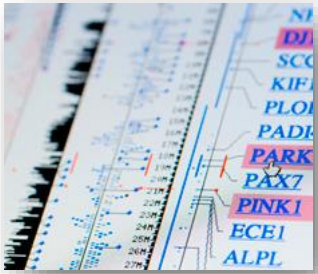
Proteins in biotechnology

- availability of optimal protein for specific process
- **traditional biotechnology** - adapt process
- **modern biotechnology** - adapt protein



Proteins in biotechnology

- **classical screening**
  - screening culture collections
  - polluted and extreme environment
- **environmental gene libraries**
  - metagenomic DNA
- **data-base mining**
  - gene databases
  - genome sequencing projects
  - numerous uncharacterised enzymes/proteins



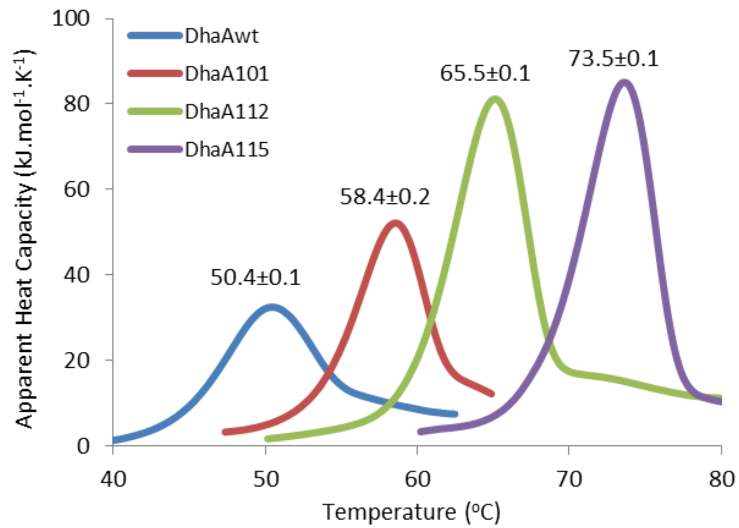
Proteins in biotechnology

- the process of **constructing novel protein** molecules by design first principles or altering existing structure
- use of genetic manipulations to alter the coding sequence of a gene and thus **modify the properties of the protein**

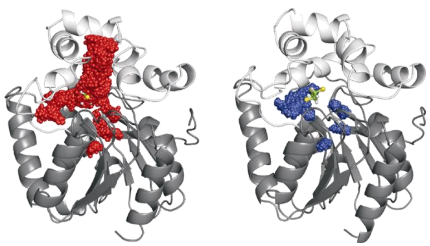
- AIMS AND APPLICATIONS
  - **technological** - optimisation of the protein to be suitable in particular technology purpose
  - **scientific** - desire to understand what elements of proteins contribute to folding, stability and function

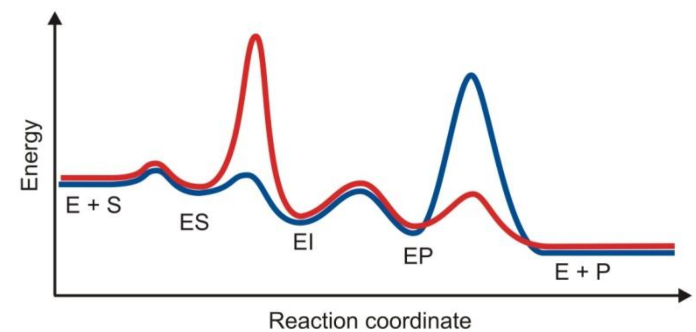
Targeted properties of proteins

- **structural** properties of proteins
  - stability (temperature, solvents)
  - tolerance to pH, salt
  - resistance to oxidative stress

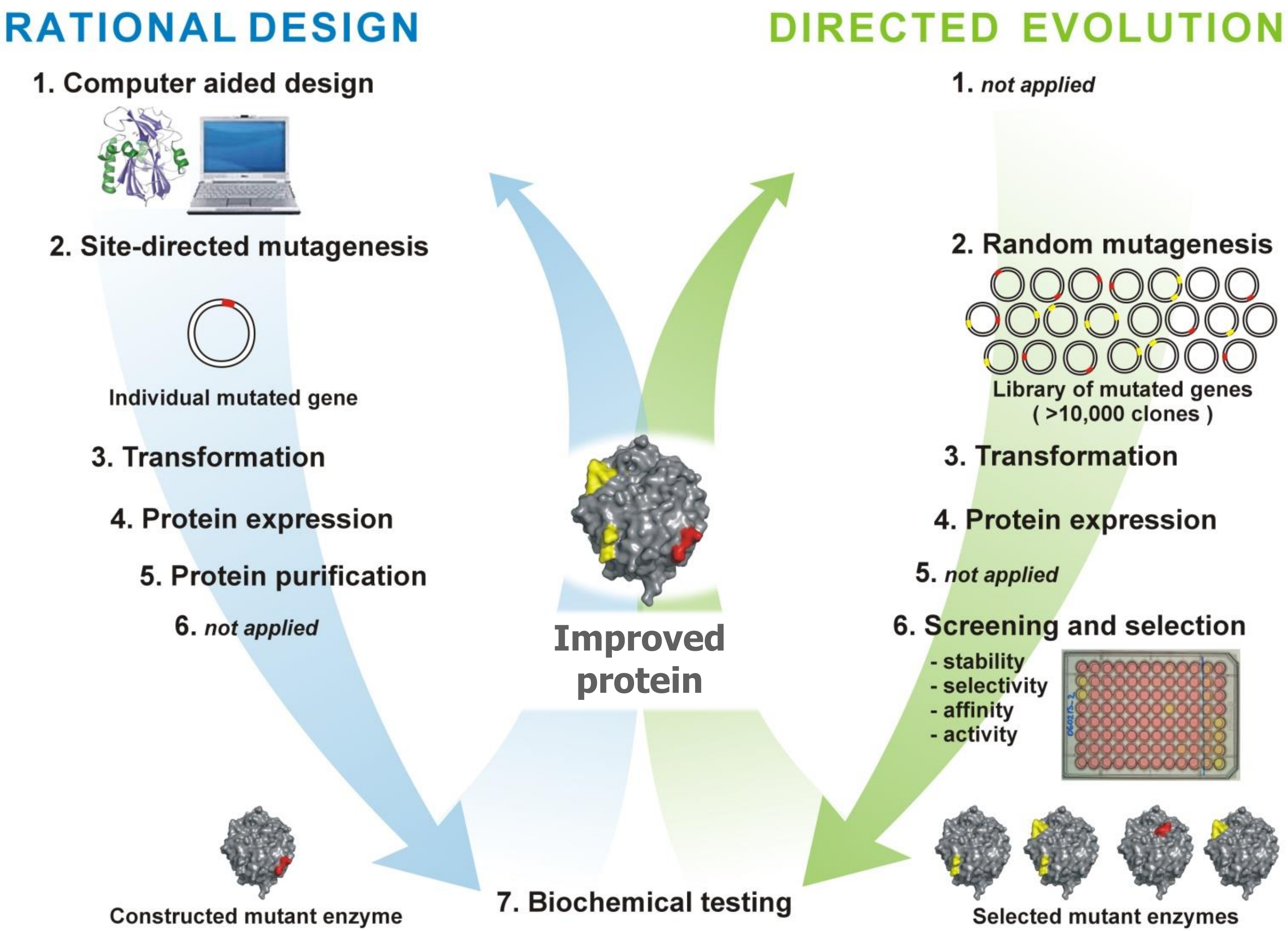


- **functional** properties of proteins
  - reaction type
  - substrate specificity and selectivity
  - kinetic properties (e.g., K_m, k_{cat}, K_i)
  - cofactor selectivity
  - protein-protein or protein-DNA interactions





Strategies in protein engineering



Directed evolution

- directed evolution techniques emerged during mid-1990s
- **inspired by natural evolution**
- this form of "evolution" does not match what Darwin had envisioned
  - requires **outside intelligence**, not blind chance
  - does not create brand new species, macroevolution,
    but only improvements of molecules, **molecular evolution**
  - does not take millions of years, but **happens rapidly**

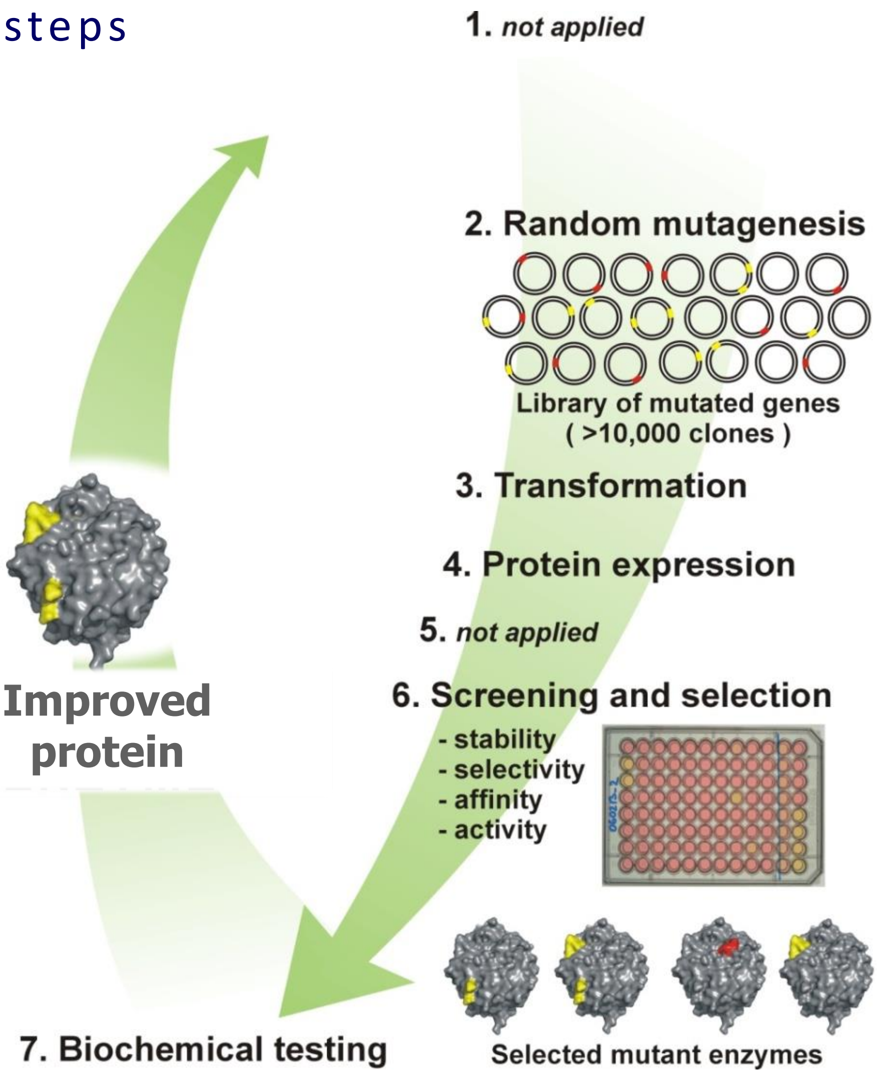
Directed evolution

- **evolution in test tube** comprises two steps
  - **random mutagenesis**

    mutant library building
  - **screening and selection**

    identification of desired biocatalyst

- **prerequisites** for directed evolution
  - gene encoding protein of interest
  - method to create mutant library
  - suitable expression system
  - screening or selection system

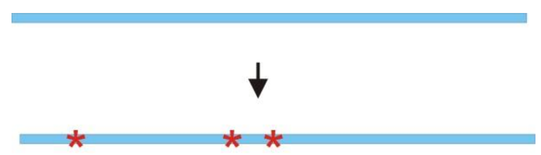


Methods to create mutant libraries

- technology to **generate large diversity**

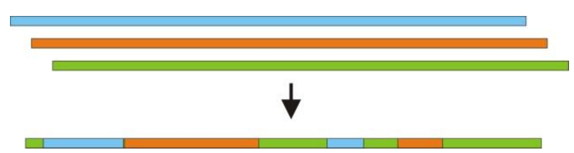
  - **NON-RECOMBINING**

    one parent gene -> variants with point mutations



  - **RECOMBINING**

    several parental homologous genes -> chimeras



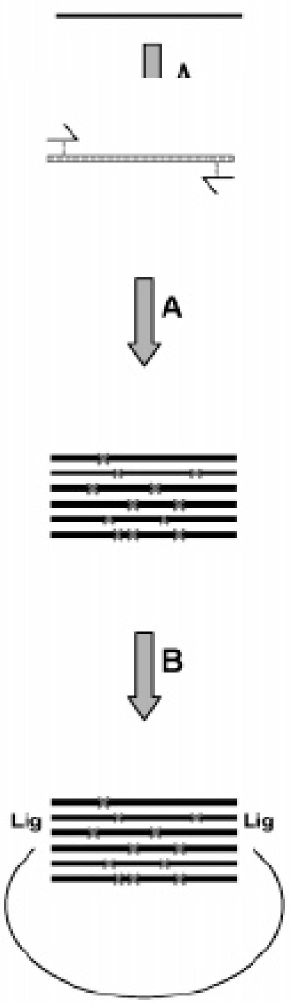
Non-recombining mutagenesis

- **UV irradiation** or **chemical mutagens** (traditional)
- **mutator strains** - lacks DNA repair mechanism

  mutations during replication (e.g., *Epicurian coli* XL1-Red)
- **error-prone polymerase chain reaction** (ep-PCR)
  - gene amplified in imperfect copying process

    (e.g., unbalanced deoxyribonucleotides concentrations,

    high Mg^{2+} concentration, Mn^{2+}, low annealing temperatures)
  - 1 to 20 mutation per 1000 base pairs
- **saturation mutagenesis**
  - randomization of single or multiple codons
- **other methods**
  - gene site saturation mutagenesis
  - cassette mutagenesis (region mutagenesis)

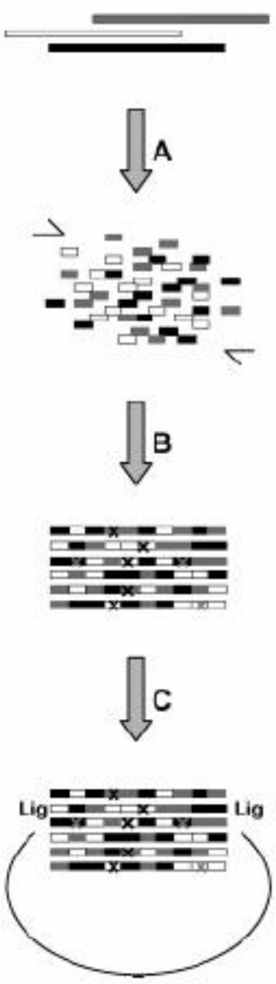


Recombining mutagenesis

- also refered to as **„sexual mutagenesis"**
- **DNA shuffling**
  - fragmentation step
  - random reassembly of segments
- **StEP** - staggered extension process
  - simpler then shuffling
  - random reannealing combined with limited primer extension
- **other methods**

  shuffling of genes with lower homology down to 70%

  (e.g., RACHITT, ITCHY, SCRATCHY)

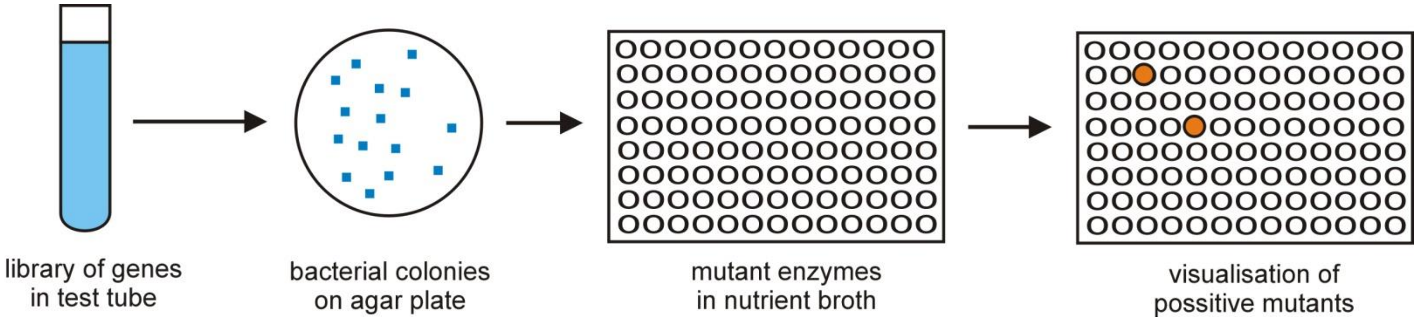


Screening and selection

- most **critical step** of direct evolution
- isolation of positive mutants hiding in library
  - **HIGH THROUGHPUT SCREENING**

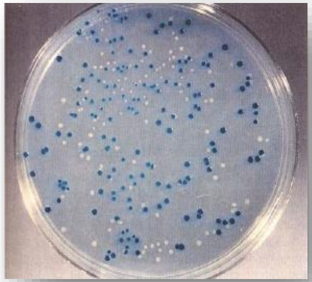
    individual assays of variants one by one
  - **DIRECT SELECTION**

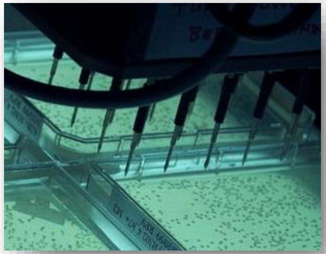
    display techniques (link between genotype and phenotype)

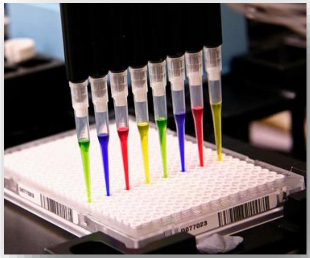


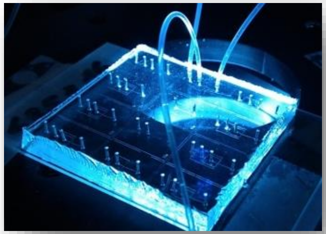
(Utra)High throughput screening

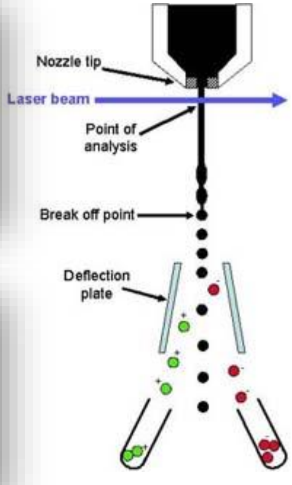
- common methods not applicable
- **agar plate (pre)screening**
- **microtiter plates screening**
  - 96-, 384- or 1536-well formate
  - robot assistance (colony picker, liquid handler)
  - 10^4 libraries
  - volume 10 – 100 uL
- **microfluidic systems**
  - water in oil emulsions (up to 10 kHz)
  - FACS sorting (10^8 events/hour)
  - 10^9 libraries
  - volume 1 – 10 pL

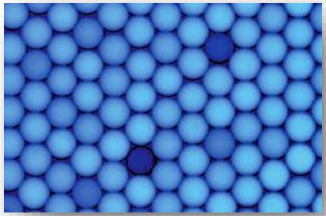






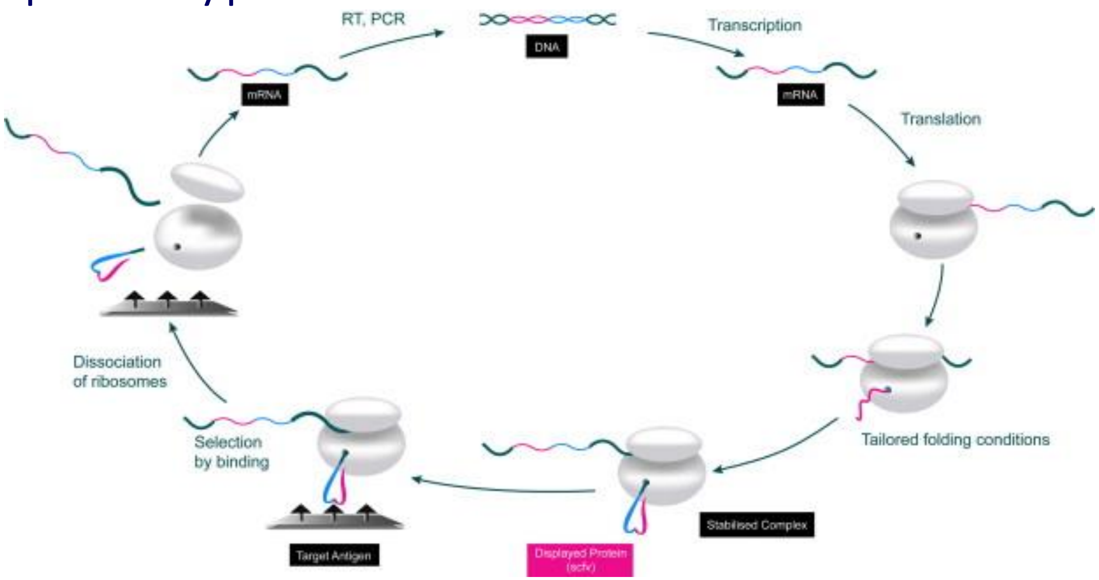


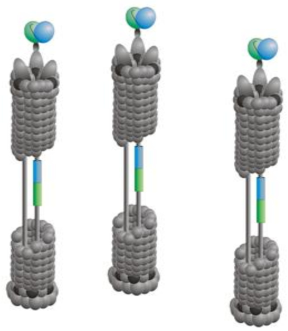




Direct selection

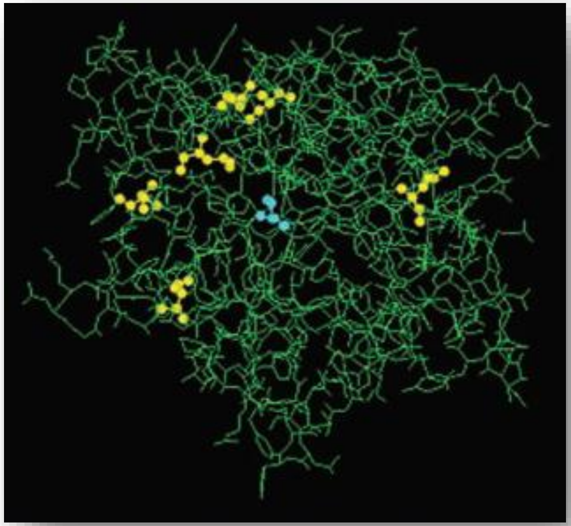
- not generally applicable (mutant libraries >10^6 variants)
- link between genotype and phenotype
- **display technologies**
  - ribosome display
  - phage display
- **life-or-death assay**
  - auxotrophic strain
  - toxicity based selection

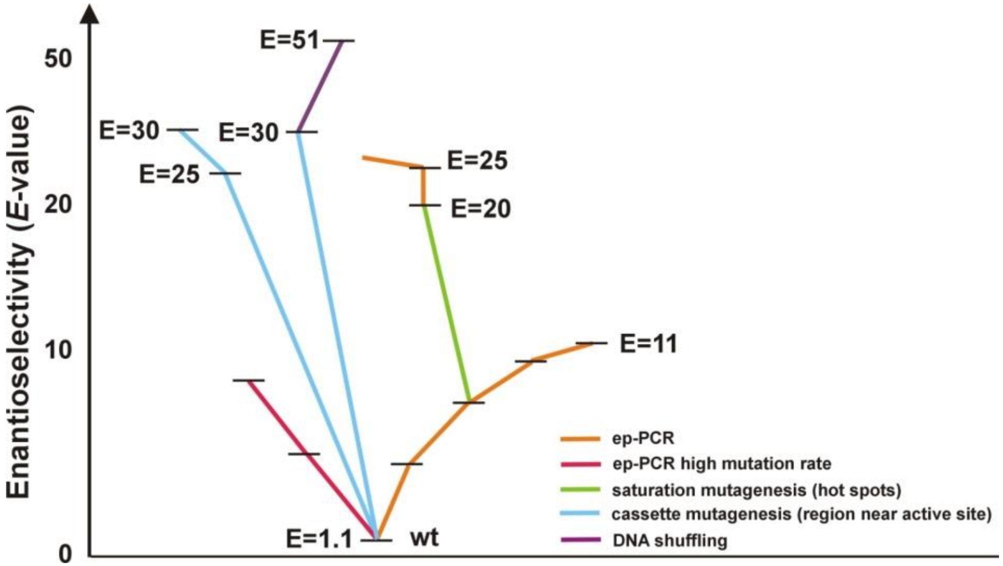




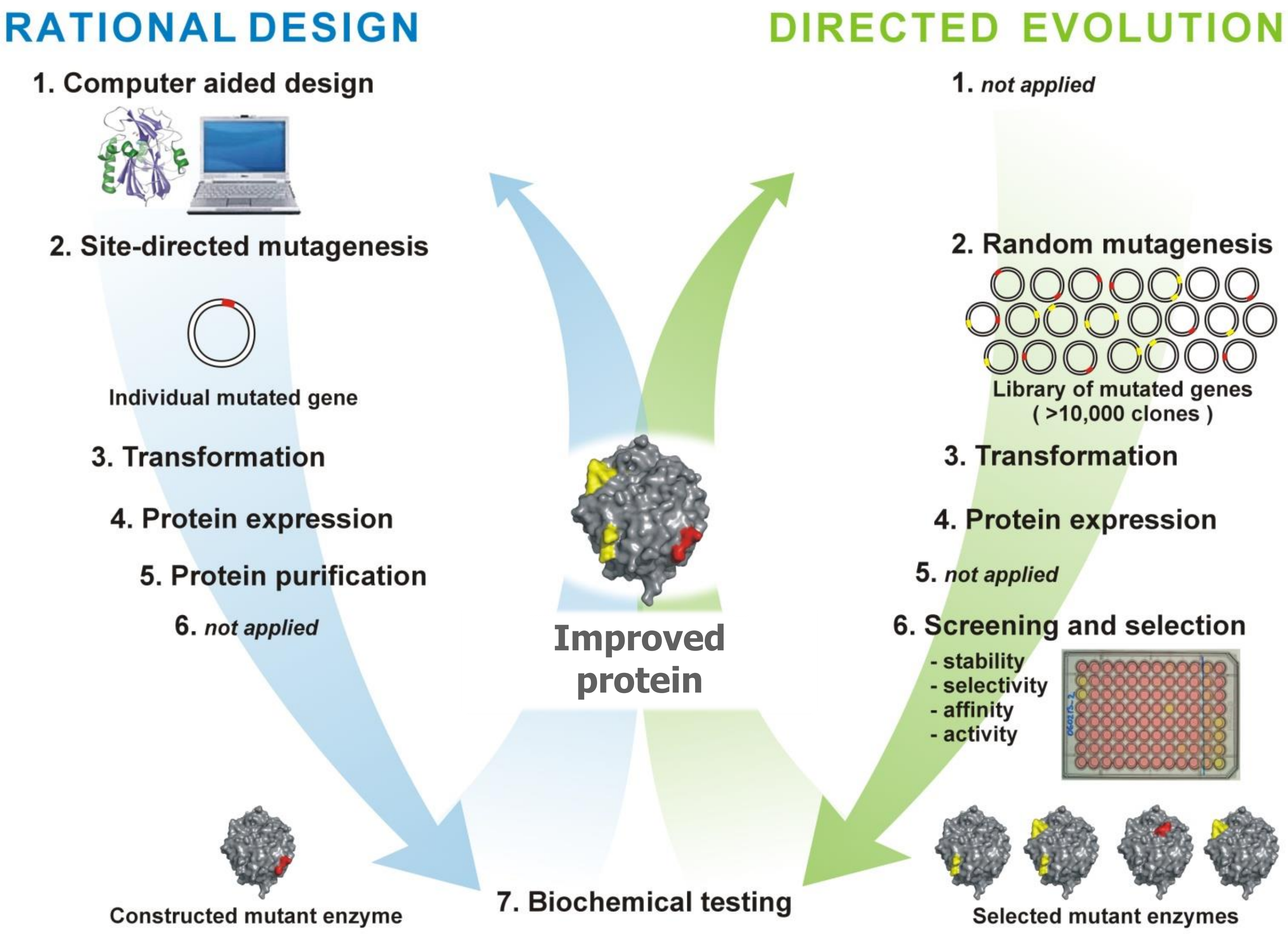
Example of Directed evolution

- directed evolution of **enantioselectivity**
  - lipase from *P. aeruginosa* (E-value improved from 1.1 into 51)
  - **spectrophotometric screening** of (*R*)- and (*S*)-nitrophenyl esters
  - **40 000 variants** screened
  - the best mutant contains six amino acid substitutions





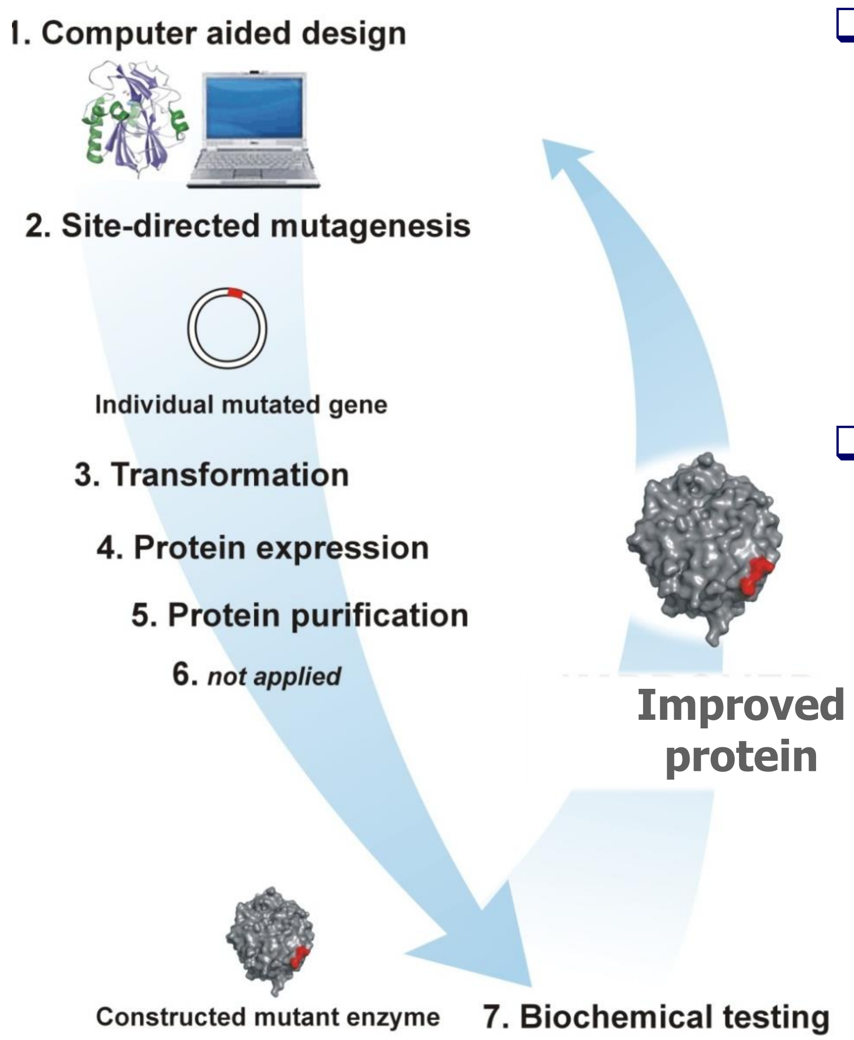
Strategies in protein engineering



Rational design

- emerged around 1980s as the original protein engineering approach
- **knowledge based** - combining theory and experiment
- protein engineering cycle:
  „structure-theory-design-mutation-purification-analysis“
- **difficulty in prediction** of mutation effects on protein property
- **de novo design**

Principal of rational design

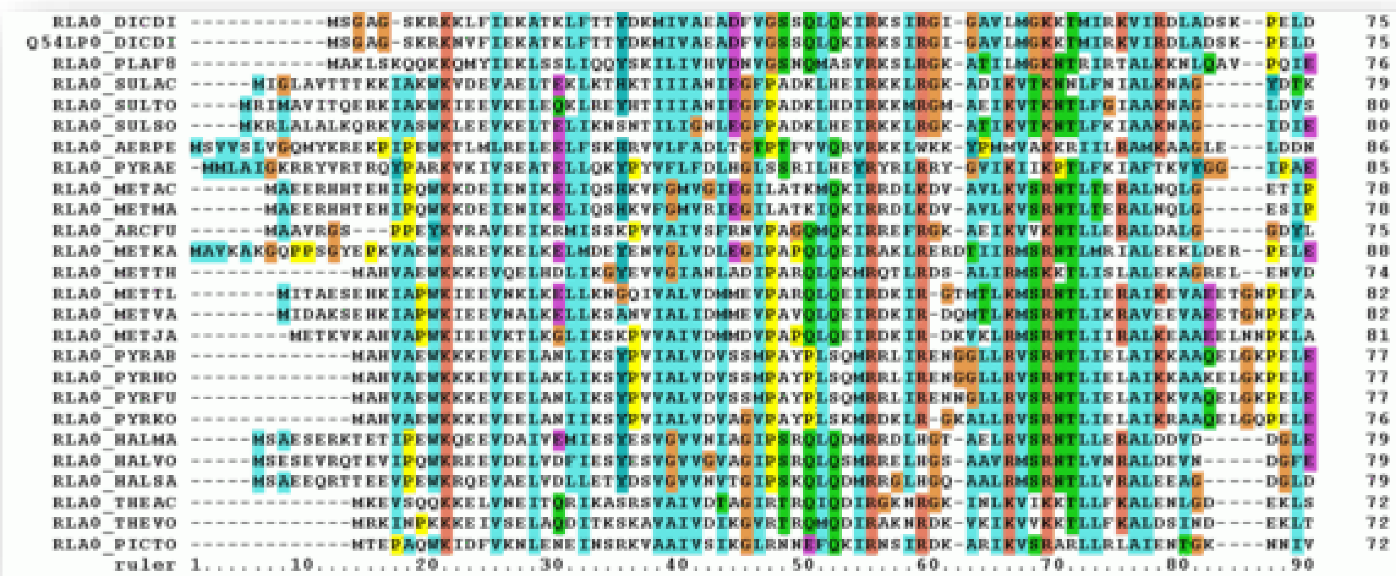


- rational design comprises:
  - **design** - understanding of protein functionality
  - **experiment** - construction and testing of mutants
- **prerequisites** for rational design:
  - gene encoding protein of interest
  - 3D structure (e.g., X-ray, NMR)
  - structure-function relationship
  - computational methods and capacity
  - (multi)side directed mutagenesis techniques
  - efficient expression system
  - biochemical tests

Design

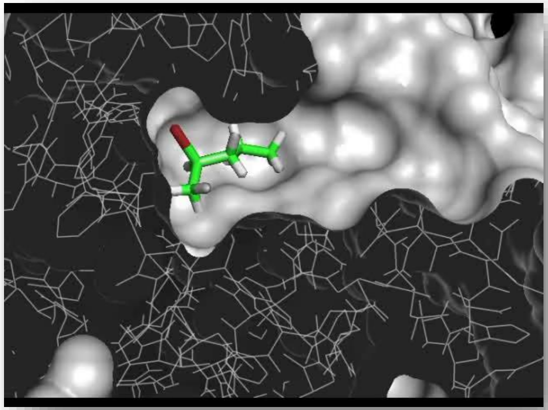
❑ HOMOLOGY APPROACH

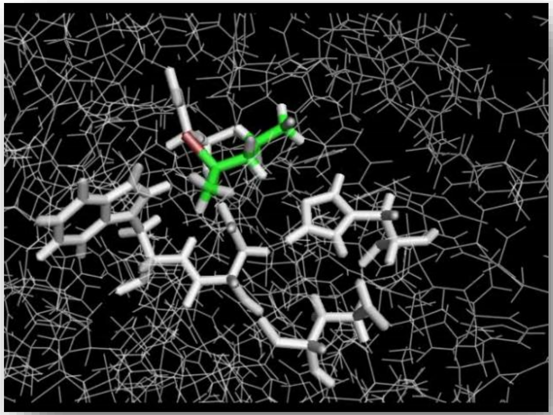
- homologous wild-type sequences are collected and compared
- identifying amino acid residues responsible for differences
- **reconstruction** - transfer differences from one enzyme to another
- **new design** - combination of possitive mutation from all parental proteins in one construct, new protein better than all parental

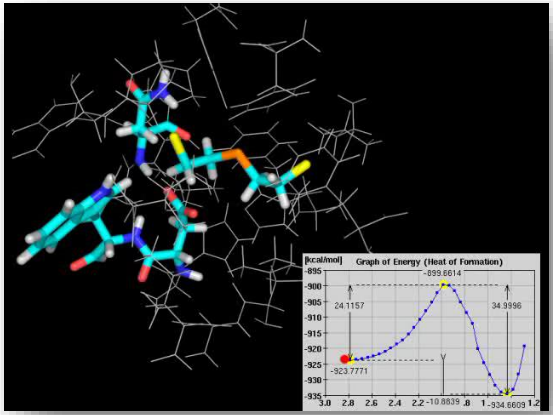


Design

- **STRUCTURE-BASED APPROACH**
  - **prediction** of enzyme function from structure alone is challenging
  - **protein structure** (X-ray crystallography, NMR, homology models)
  - **molecular modelling**
    - molecular docking
    - molecular dynamics
    - quantum mechanics/molecular mechanics (QM/MM)

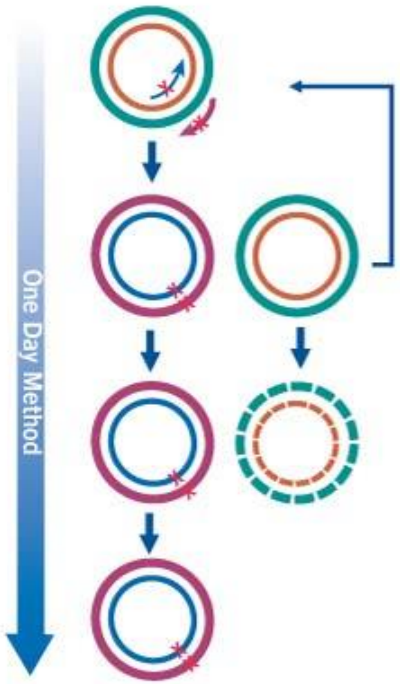






Construction

- site-directed mutagenesis
  - introducing point mutations
- multi site-directed mutagenesis
- gene synthesis
  - commercial service
  - codone optimisation



GENEART THE GENE OF YOUR CHOICE

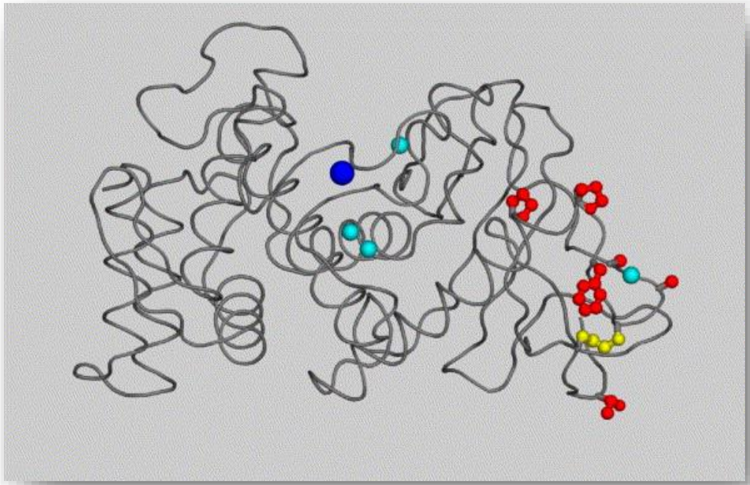
GenScript Make Research Easy

Example of rational design

- rational design of protein **stability**
  - stability to high temperature, extreme pH, proteases etc.
  - **stabilizing mutations** increase strength of weak interactions
    - **salt bridges and H-bonds**
      Eijsink et al., Biochem. J. 285: 625-628, 1992
    - **S-S bonds**
      Matsumura et al., Nature 342: 291-293, 1989
    - **addition of prolines**
      Watanabe et al., Eur. J. Biochem. 226: 277-283, 1994
    - **less glycines**
      Margarit et al., Protein Eng. 5: 543-550, 1992
    - **oligomerisation**
      Dalhus et al., J. Mol. Biol. 318: 707-721, 2002

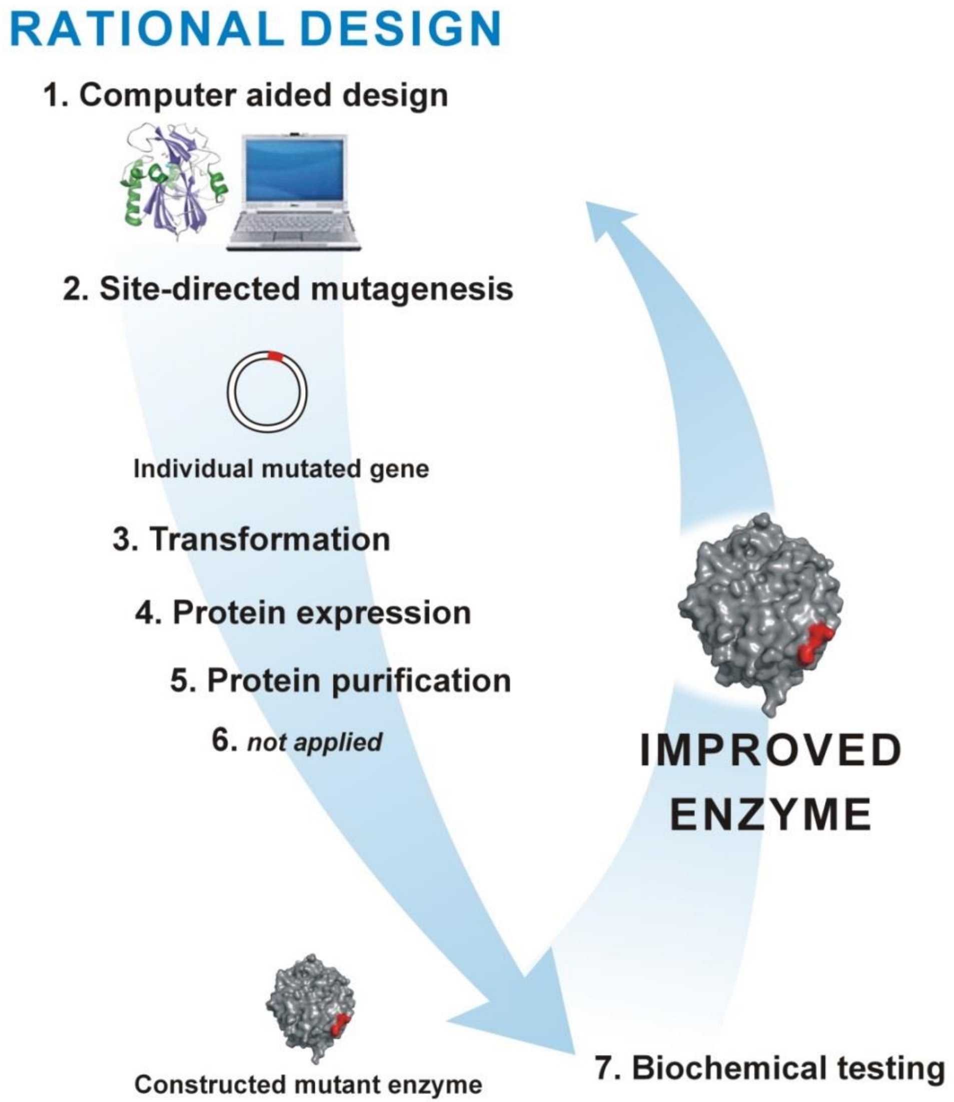
Example of rational design

- **engineering protein to resist boiling**
  - **reduced rotational freedom**
    Ser65Pro, Ala96Pro
  - **introduction of disulfide bridge**
    Gly8Cys + Asn60Cys
  - **improved internal hydrogen bond**
    Ala4Thr
  - **filling cavity**
    Tyr63Phe

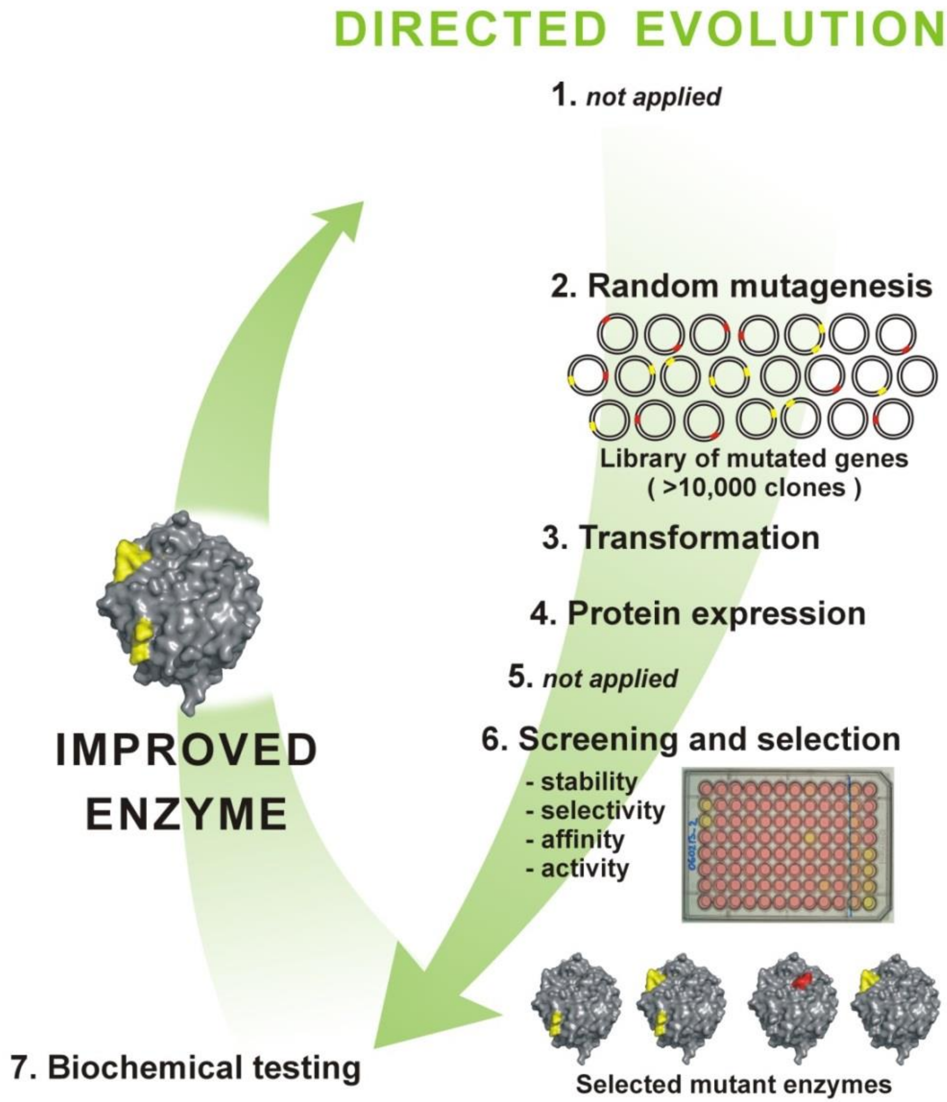


Half-lifes (min.)	80°C	100°C
wild type	17.5	>0.5
8-fold mutant	stable	170

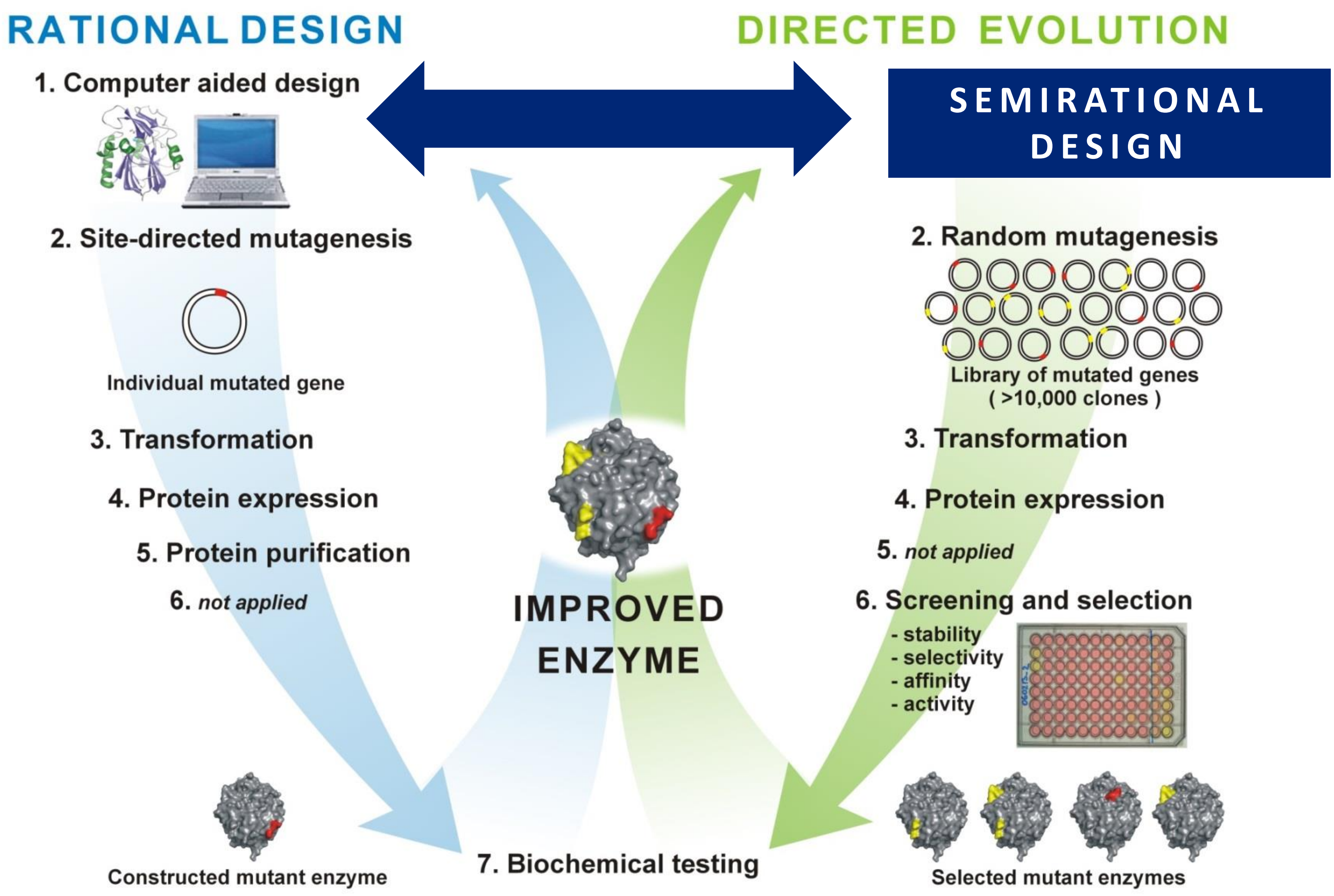
Strategies in protein engineering



Strategies in protein engineering

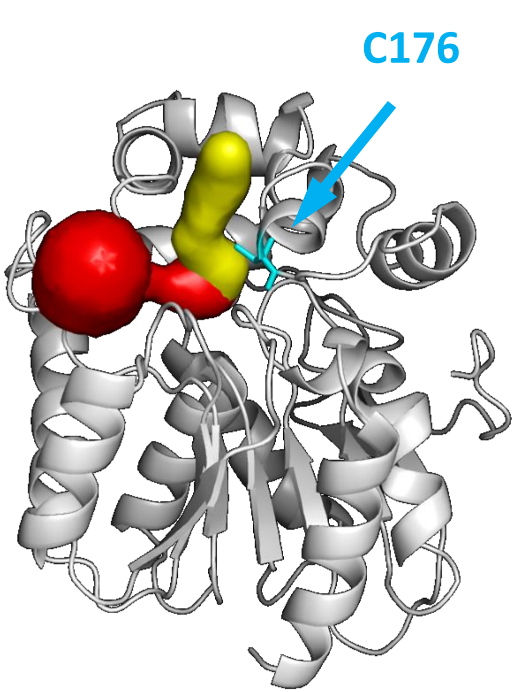


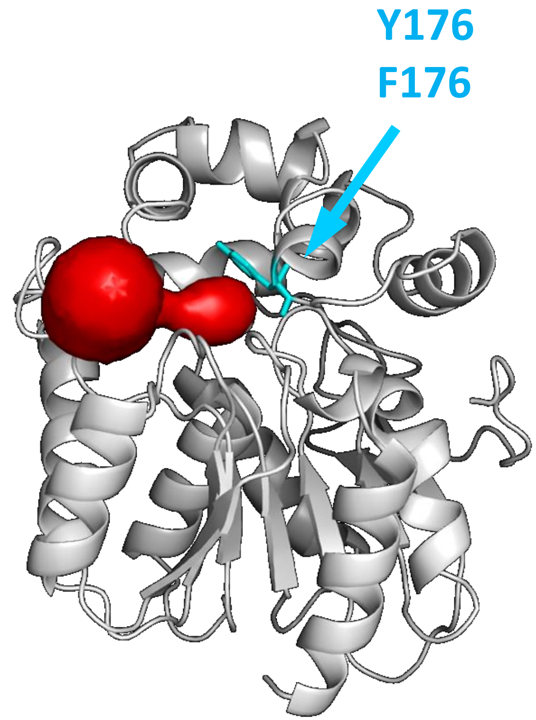
Strategies in protein engineering



Example of rational design

- conversion of 1,2,3-trichloropropane
  by DhaA from *Rhodococcus erythropolis* Y2
- **DIRECTED EVOLUTION -** importance of access pathways



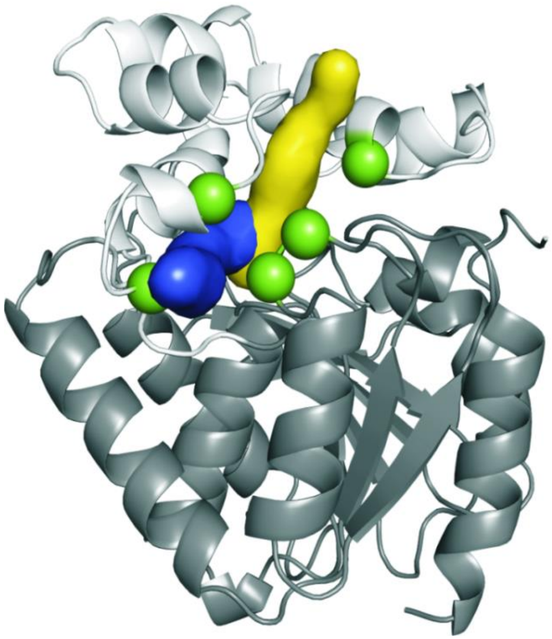


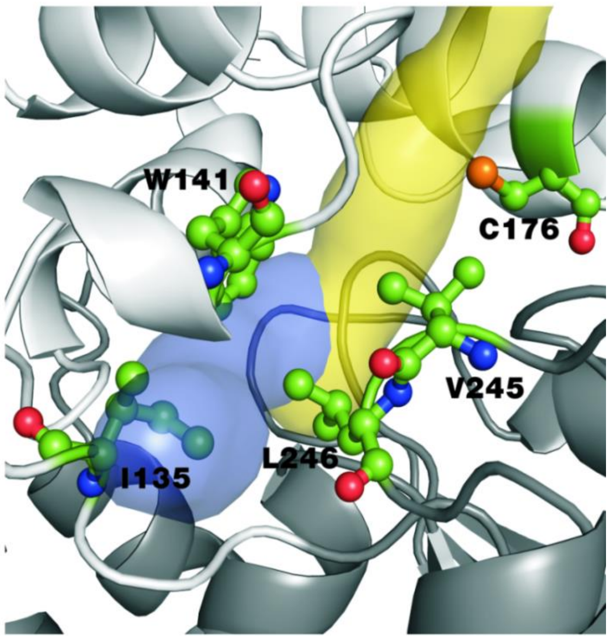
Bosma, T., et al. 2002: AEM 68: 3582-87

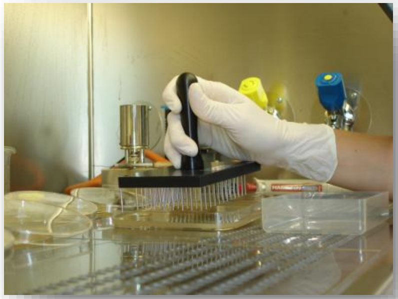
Gray, K.A., et al. 2003: Adv. Appl. Microbiol. 52: 1-27

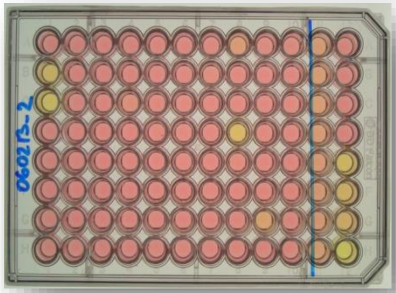
Example of rational design

- conversion of 1,2,3-trichloropropane by DhaA from *Rhodococcus erythropolis* Y2
- **DIRECTED EVOLUTION -** importance of access pathways
- **SEMI-RATIONAL DESIGN** - hot spots in access tunels
- library of **5,300 clones** screened



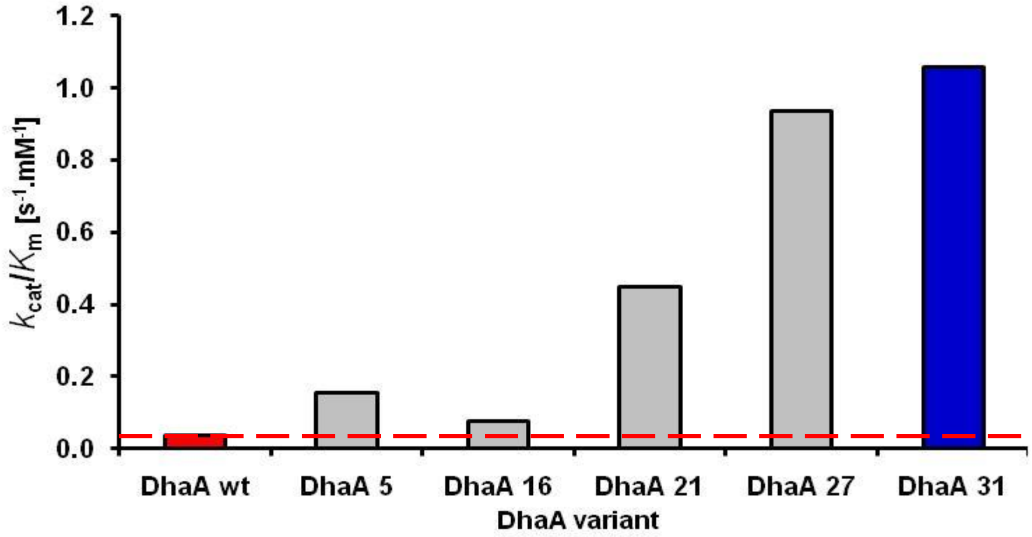


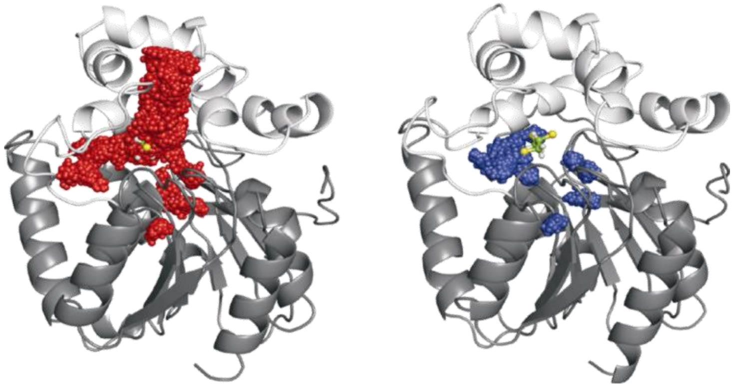


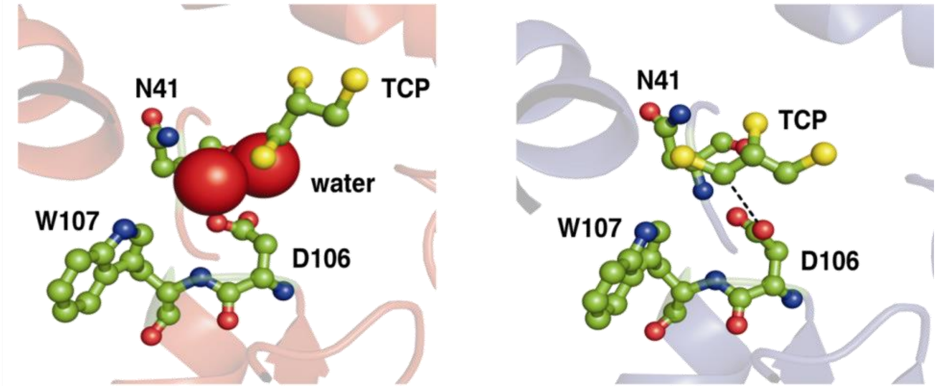


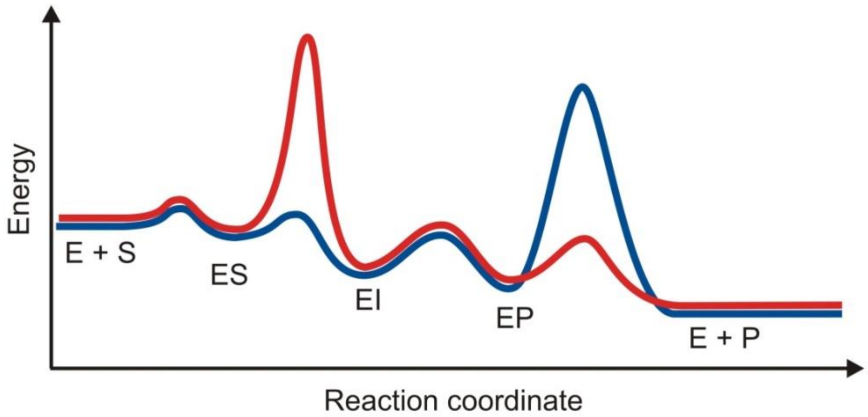
Pavlova, M., Klvana, M., Prokop, Z., et al. 2009: Nature Chem. Biol. 5: 727-733

Example of rational design



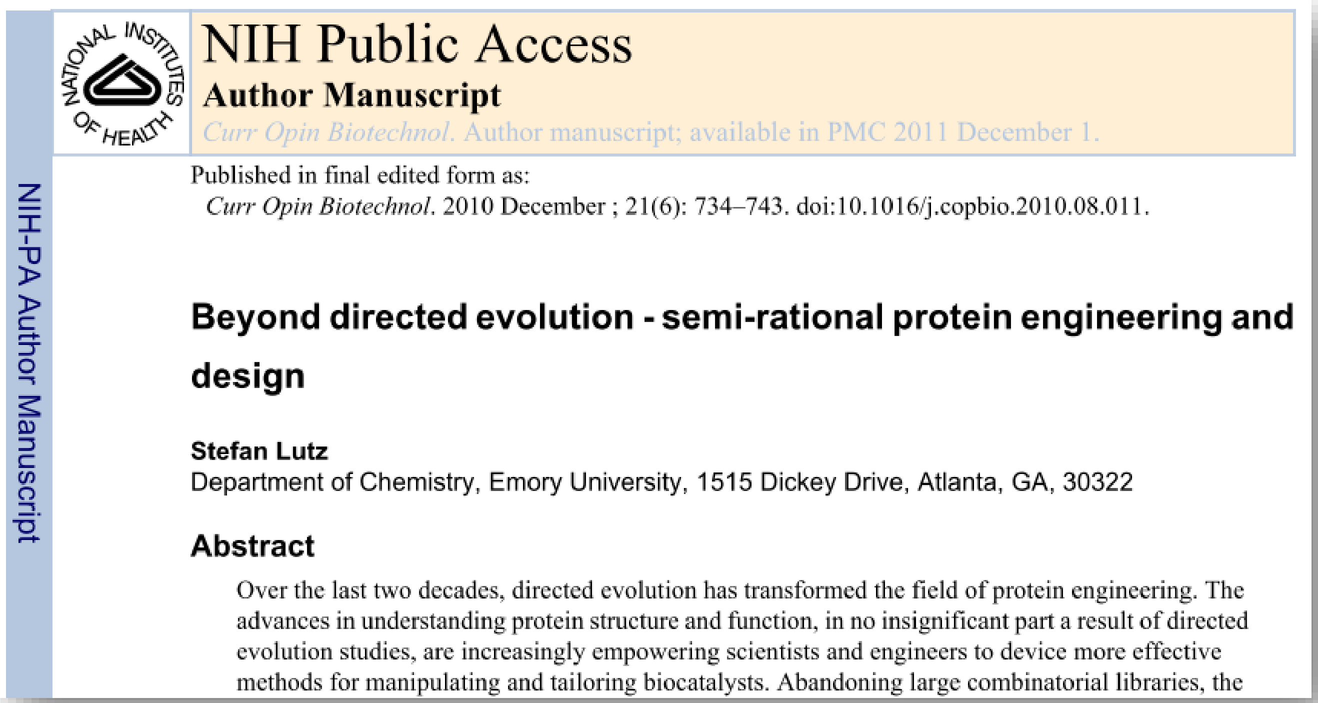






Reading

- Lutz, S. 2010: **Beyond directed evolution - semi-rational protein engineering and design**. *Curr Opin Biotechnol*. 21(6): 734–743
- *Computational enzyme redesign* and *Computational de novo enzyme design* (**page 5-7**)



NIH-PA Author Manuscript

NIH Public Access

Author Manuscript

Curr Opin Biotechnol. Author manuscript; available in PMC 2011 December 1.

Published in final edited form as:
Curr Opin Biotechnol. 2010 December ; 21(6): 734–743. doi:10.1016/j.copbio.2010.08.011.

Beyond directed evolution - semi-rational protein engineering and design

Stefan Lutz
Department of Chemistry, Emory University, 1515 Dickey Drive, Atlanta, GA, 30322

Abstract

Over the last two decades, directed evolution has transformed the field of protein engineering. The advances in understanding protein structure and function, in no insignificant part a result of directed evolution studies, are increasingly empowering scientists and engineers to device more effective methods for manipulating and tailoring biocatalysts. Abandoning large combinatorial libraries, the