CEITEC

Central European Institute of Technology

BRNO | CZECH REPUBLIC

Protein Engineering

Pavel Mazura

CEITEC – MENDELU

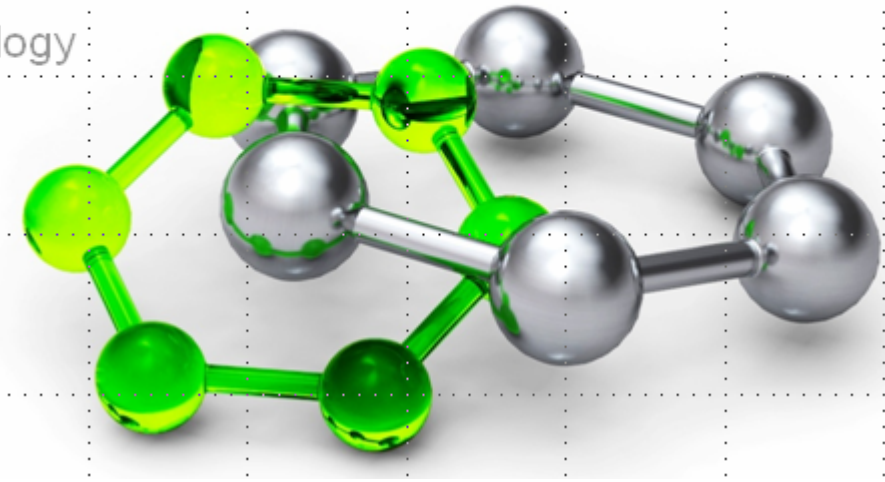
Department of molecular biology and radiobiology

Scaled Biosystems team

mazura@sci.muni.cz

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

2007–13
OP Research and Development for Innovation

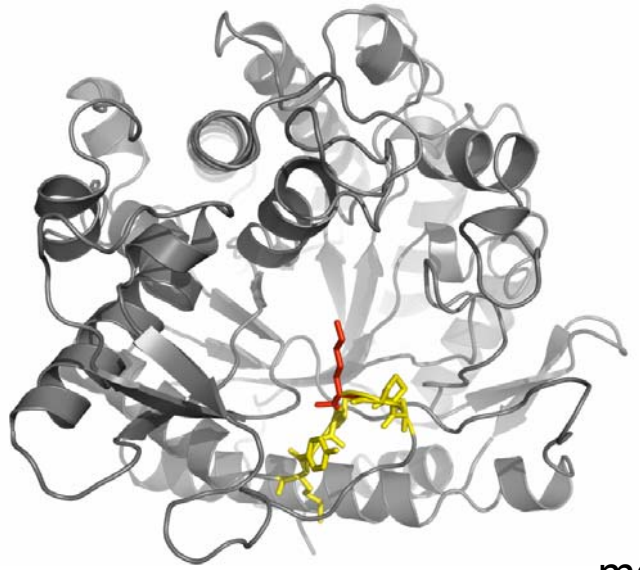


SCALED BIOSYSTEMS

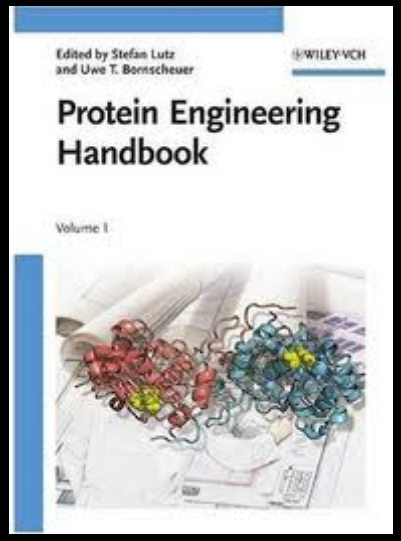
Research & Development of molecular systems for control of plant growth and morphogenesis

Screening and characterization of novel β-glucosidases
for target probing of plant hormone system

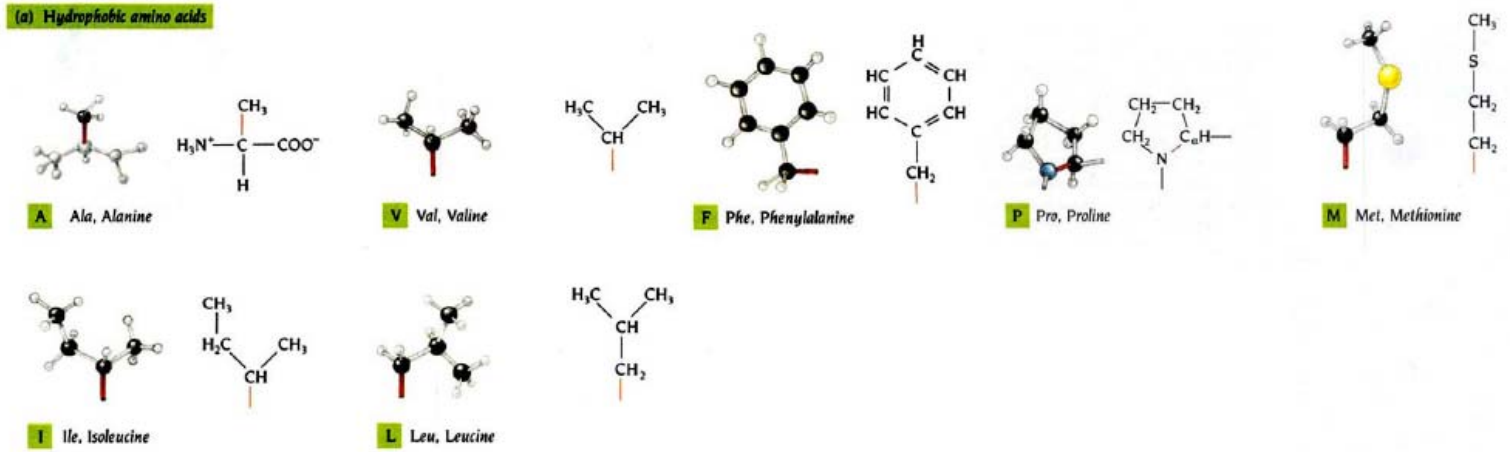
Bioinformatics, *E.coli* transformation and screening,
DNA isolation + (basic molecular biology techniques)



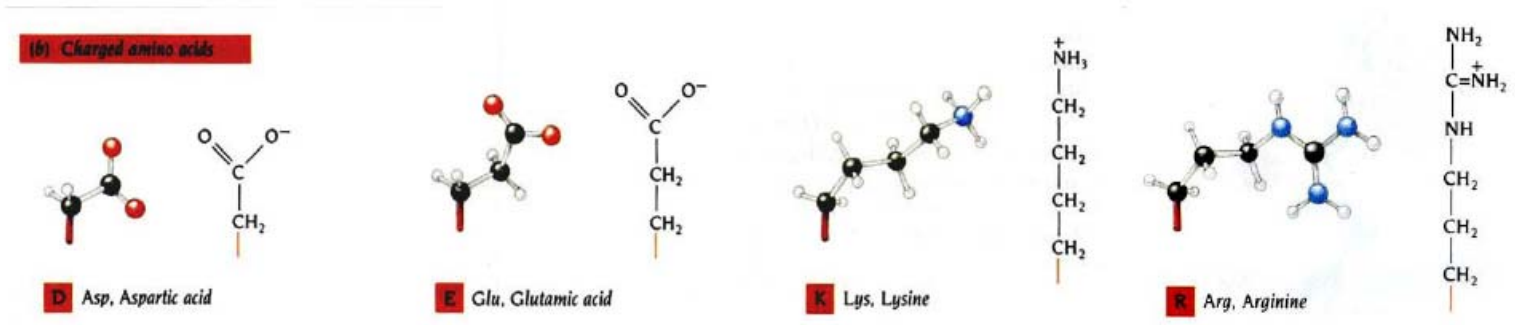
mazura@sci.muni.cz


Edited by Stefan Lutz
and Uwe T. Bornscheuer
WILEY-VCH
Protein Engineering
Handbook
Volume 1

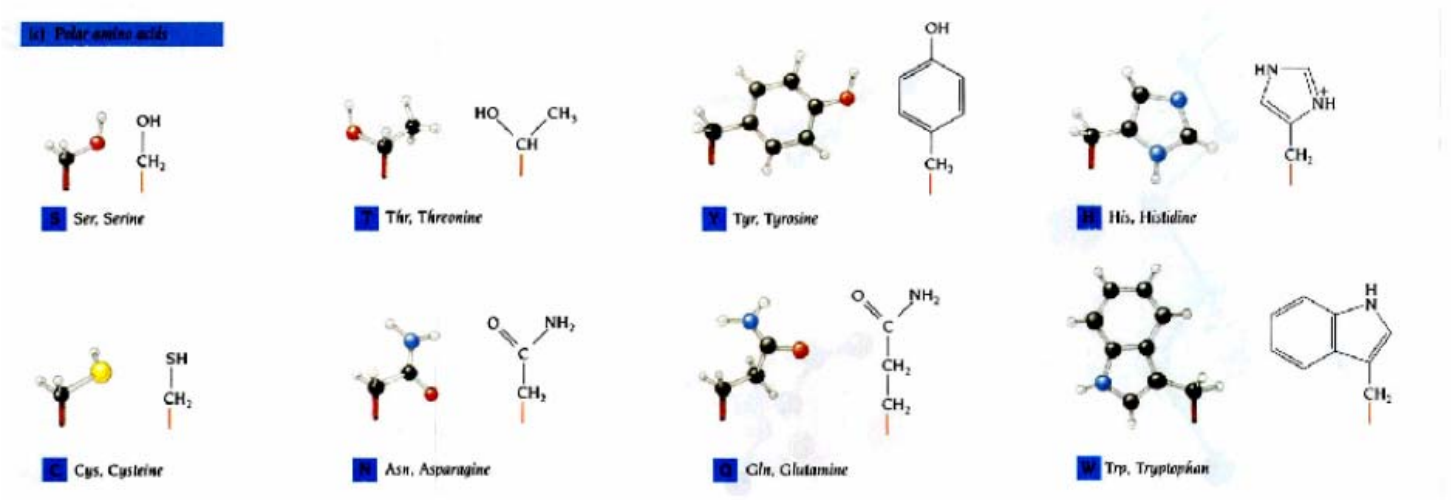
Hydrophobic amino acids



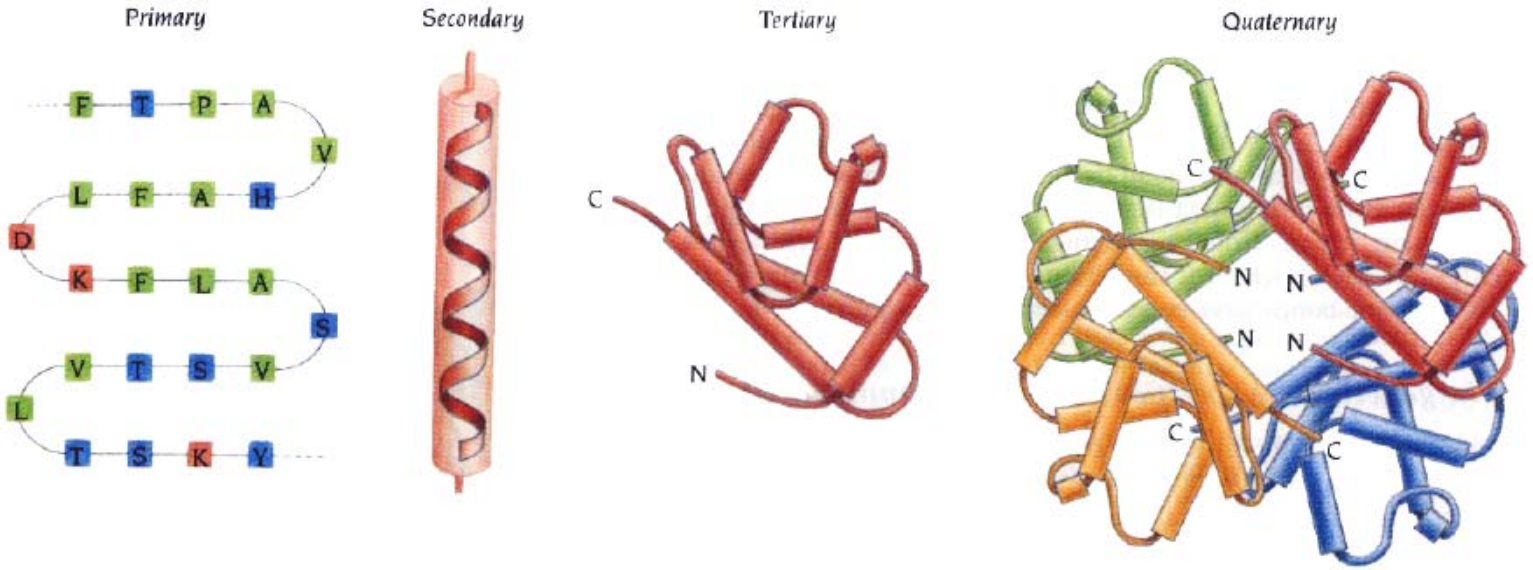
Charged amino acids



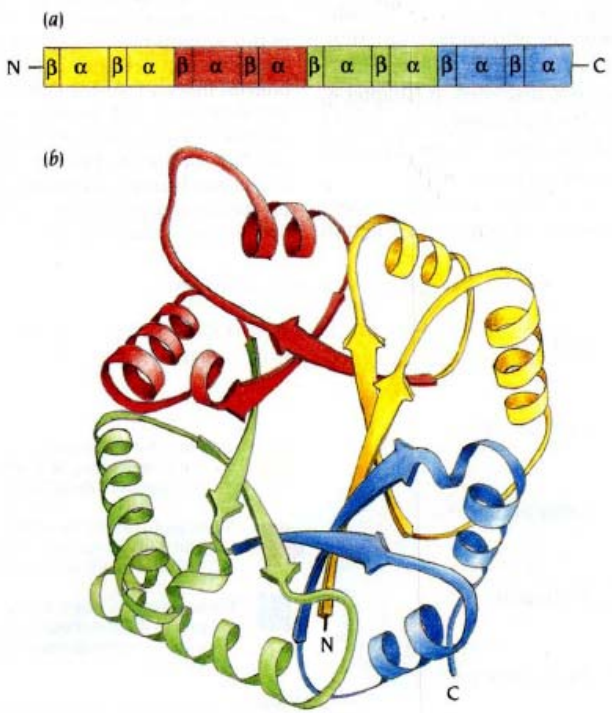
Polar amino acids



Protein structure



Domains are build from structural motifs



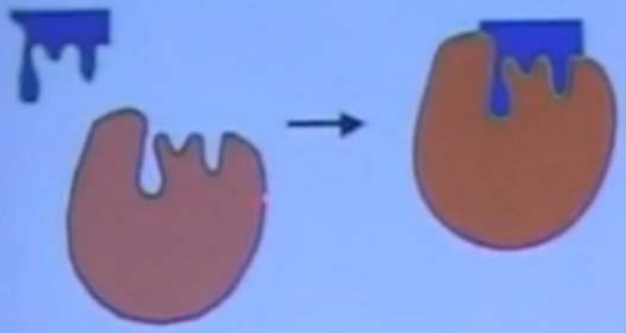
To a first approximation a polypeptide chain can be considered as a sequential arrangement of simple motifs that are formed from consecutive regions of the primary structure. For example, triosephosphate isomerase is built up from four β–α–β–α motifs that are consecutive both in the amino acid sequence (a) and in the three-dimensional structure (b).

The number of such combinations found in proteins is limited, and some combinations seem to be structurally favored.

Thus similar domain structures frequently occur in different proteins with different functions and with completely different amino acid sequences.

Proteins - the traditional view

The Lock & Key Model



one sequence =
one structure =
one function

Emil Fischer
1852-1919

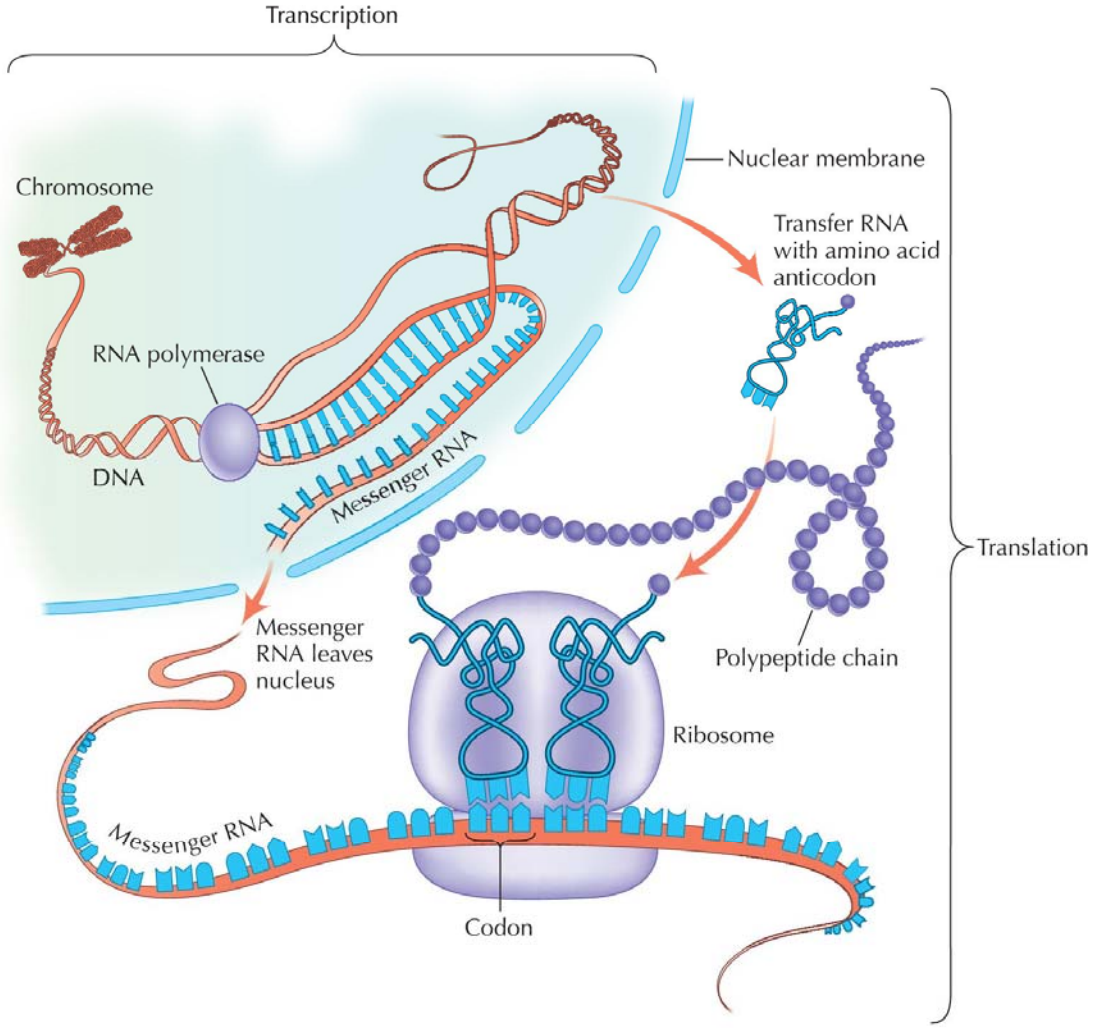
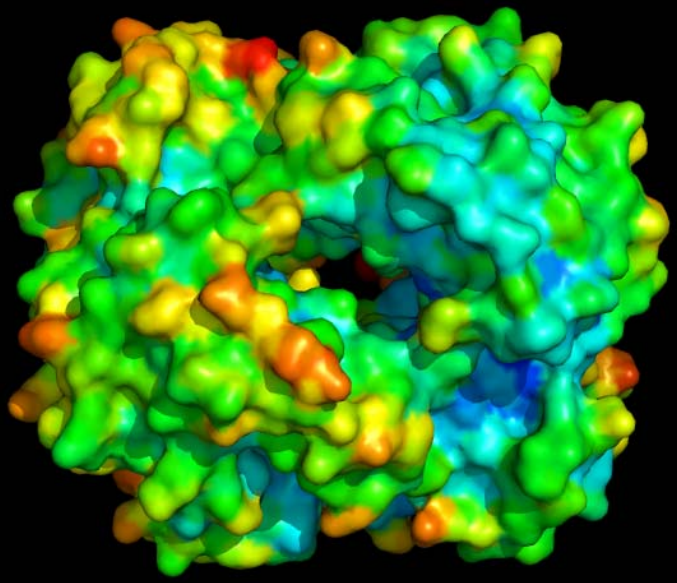
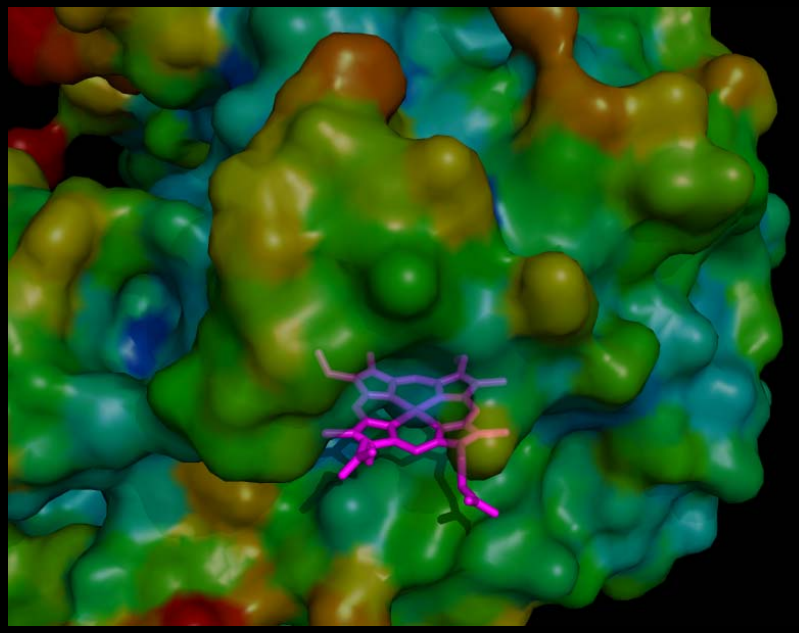


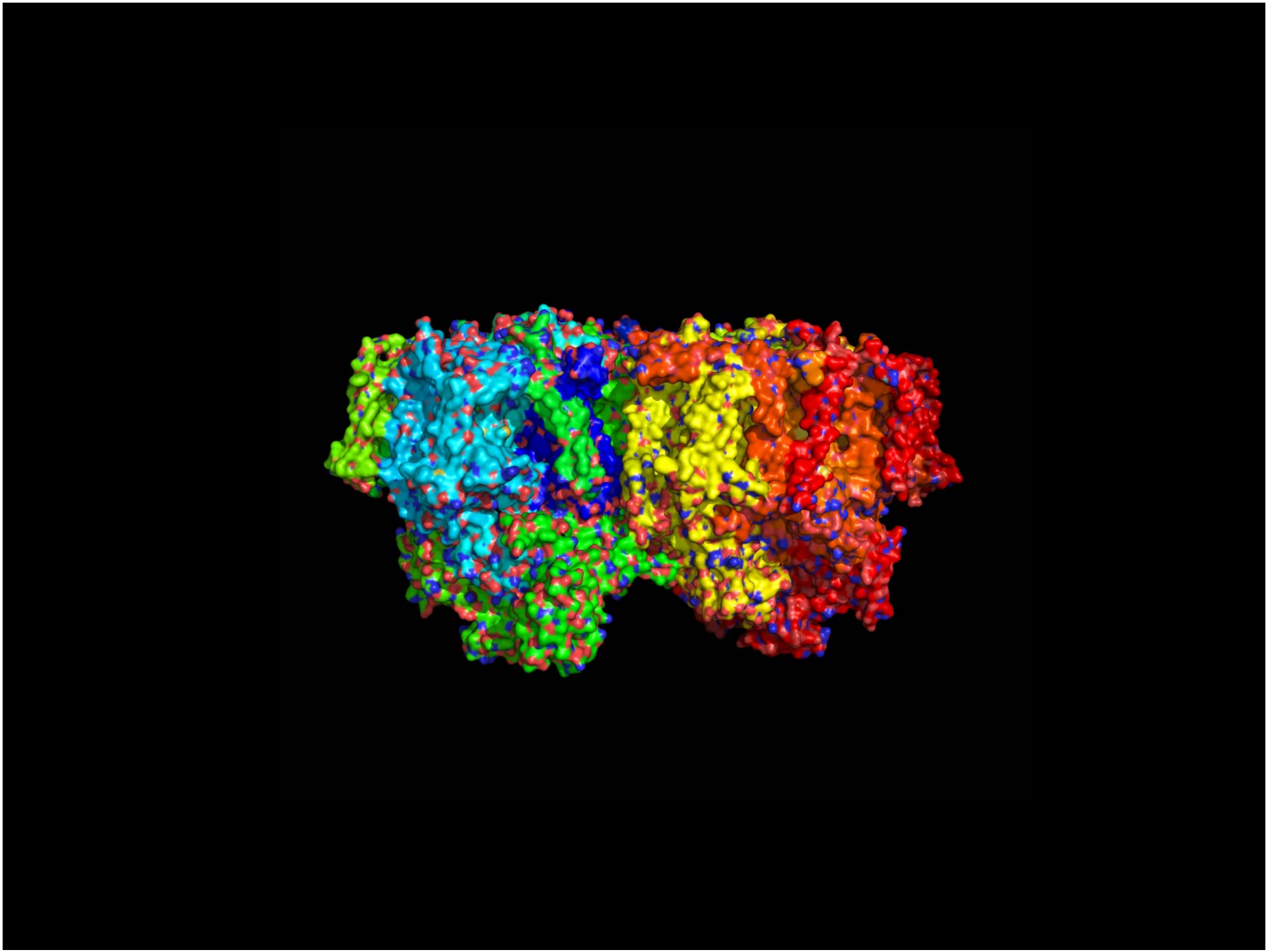
FIGURE 2.27. General mechanism of eukaryotic protein synthesis. The major steps include transcription of the DNA gene sequence into the messenger RNA template in the nucleus of the cell, translation of the DNA codons of that gene into amino acids, and their assembly into polypeptides in the cytoplasm. Important mediators of this process include transfer RNAs, splicing elements, and ribosomes.

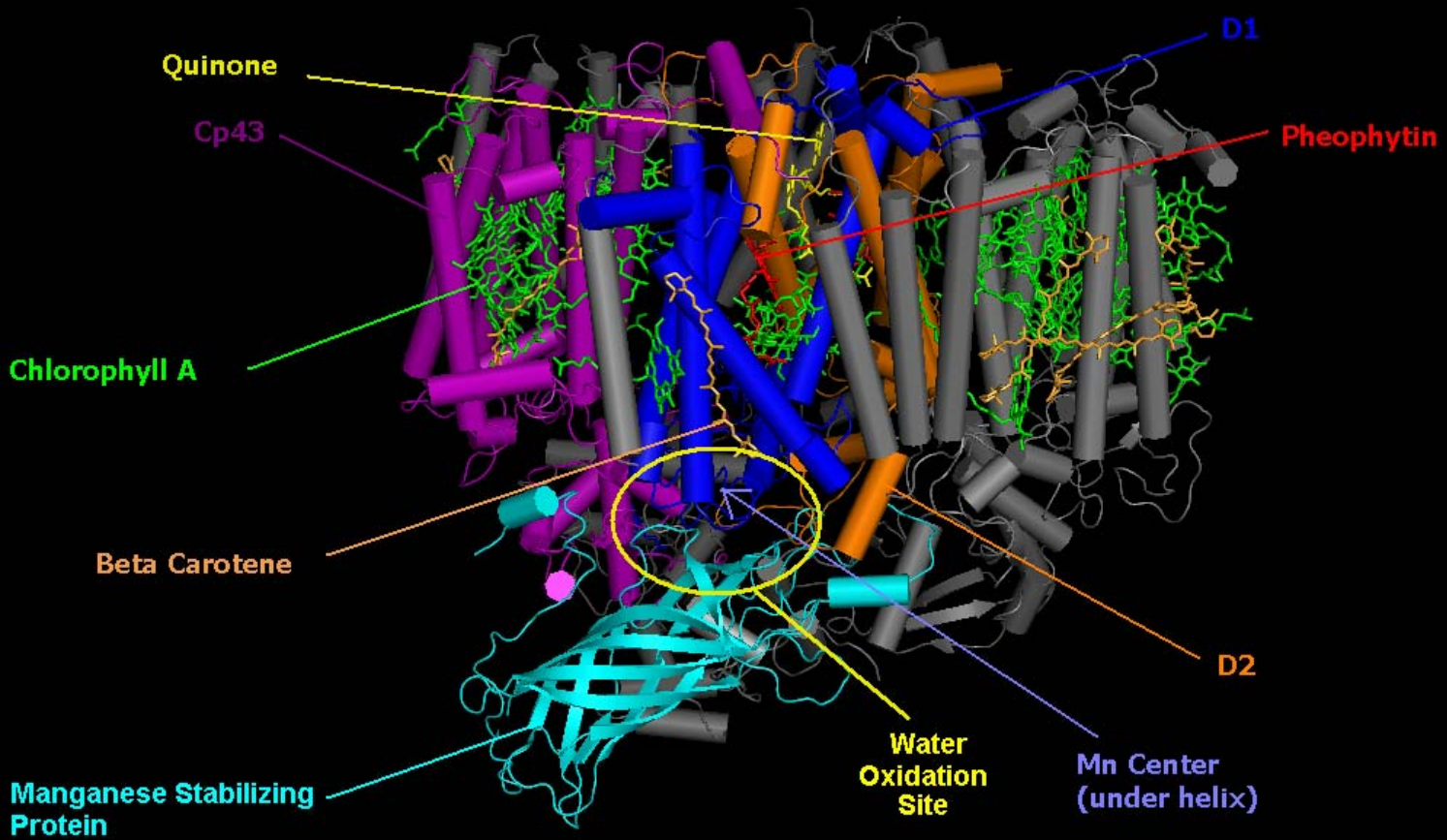
Protein Engineering

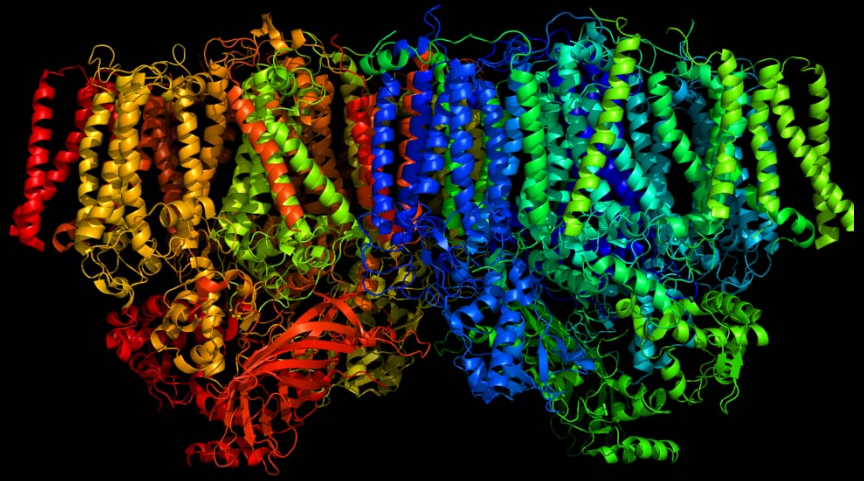
the Big picture

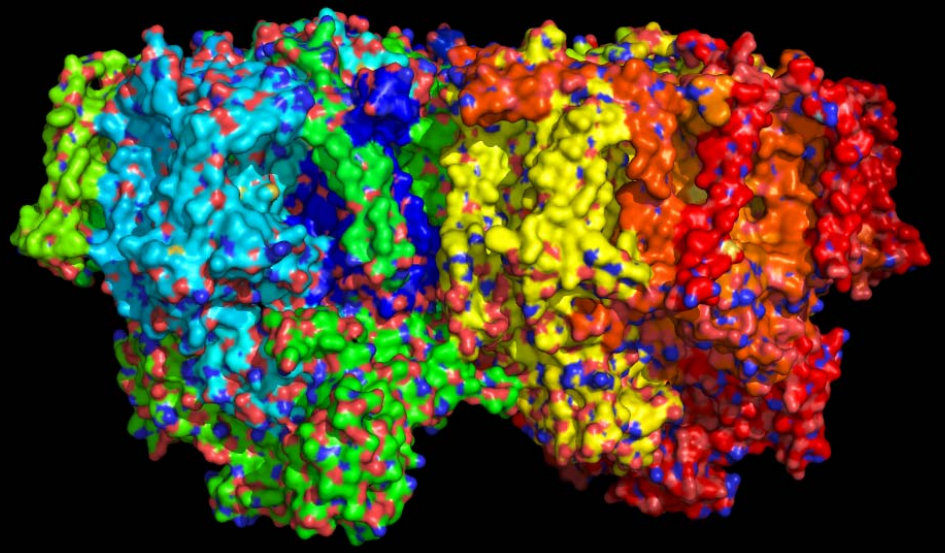


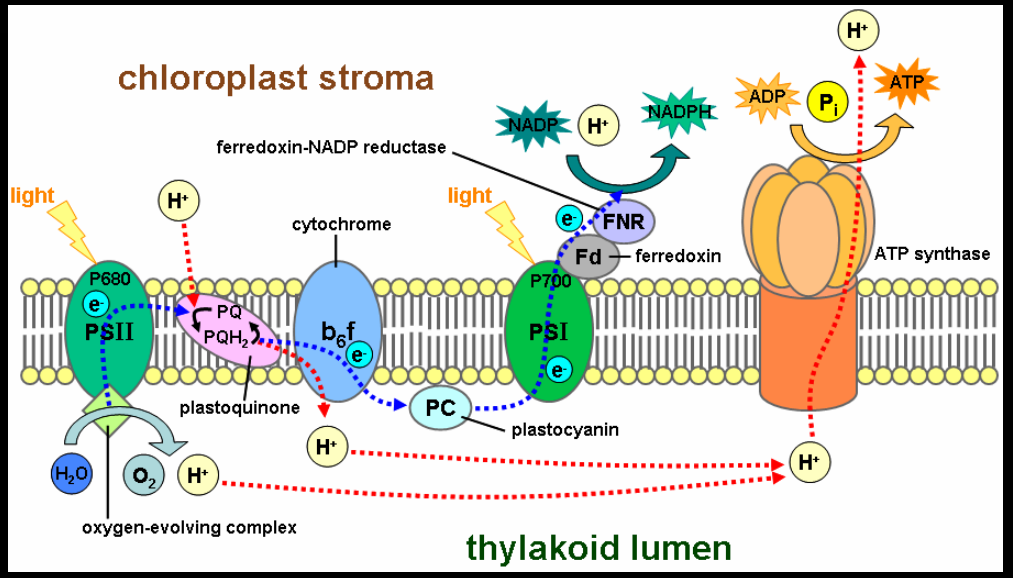


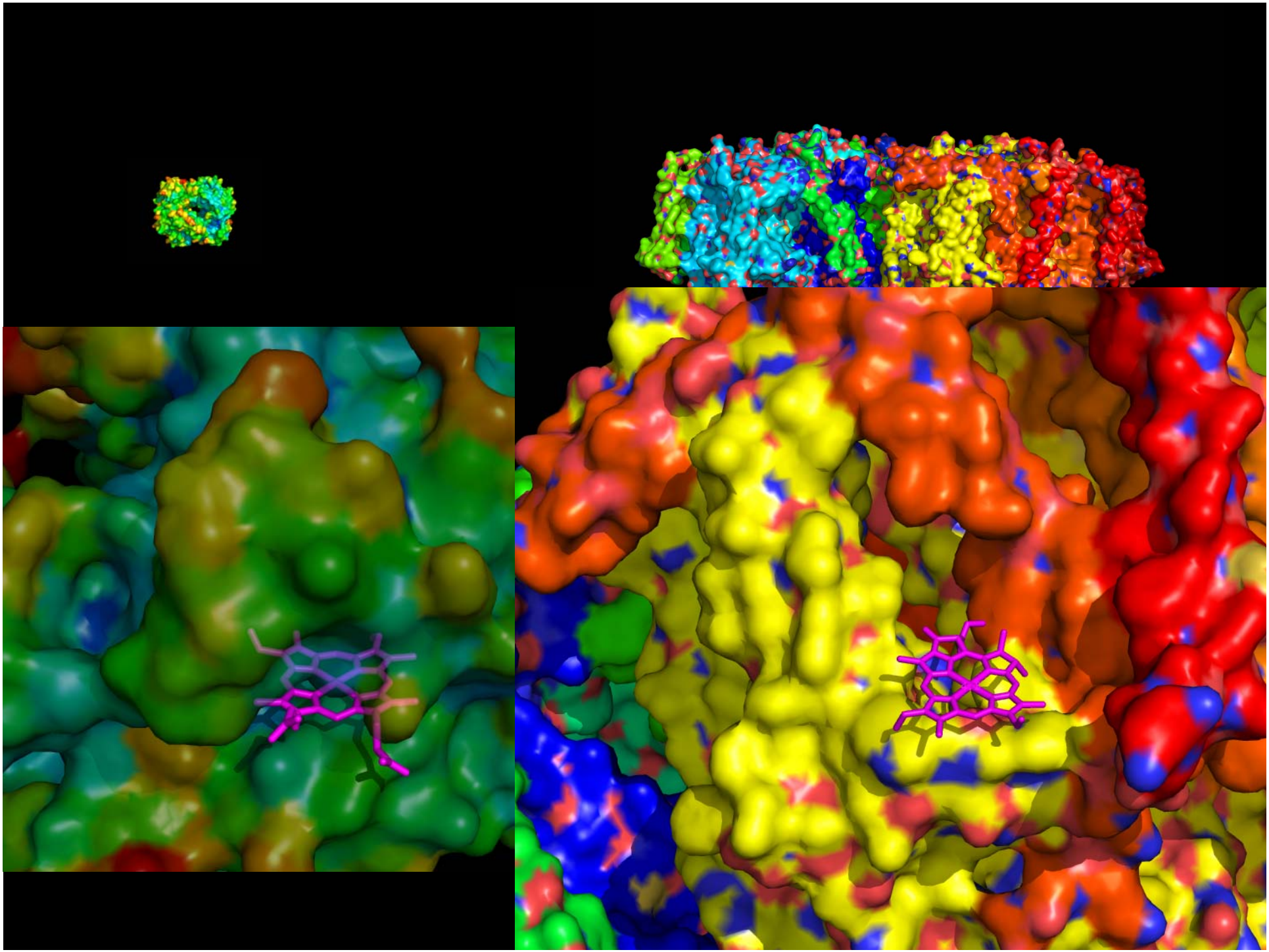



Quinone
Cp43
Chlorophyll A
Beta Carotene
Manganese Stabilizing Protein
D1
Pheophytin
D2
Water Oxidation Site
Mn Center (under helix)

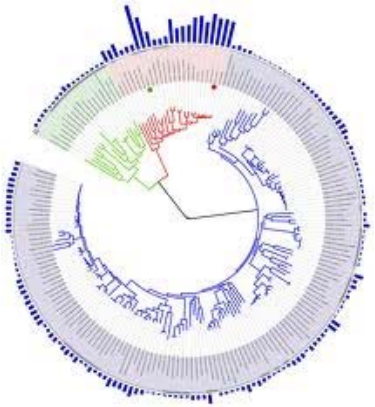


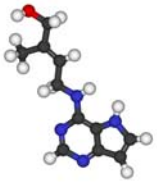



chloroplast stroma
light
H+
P680
e-
PSII
PQ
PQH2
plastoquinone
cytochrome
b6f
e-
light
P700
PSI
e-
NADP
H+
NADPH
ferredoxin-NADP reductase
e-
FNR
Fd
ferredoxin
ADP
Pi
ATP
ATP synthase
H+
PC
plastocyanin
H+
H+
H2O
O2
H+
oxygen-evolving complex
thylakoid lumen

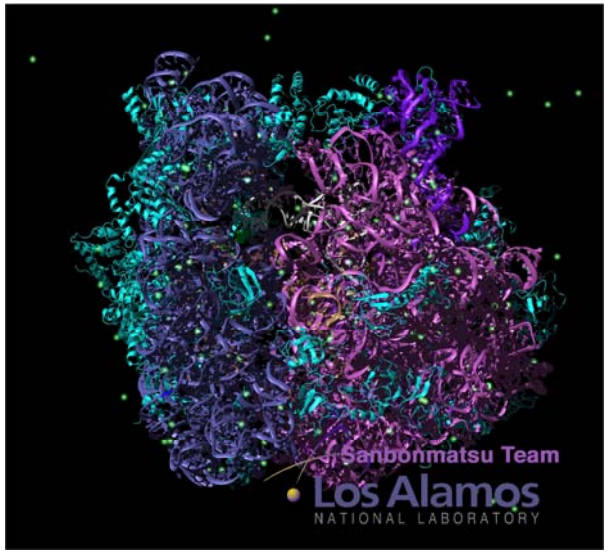


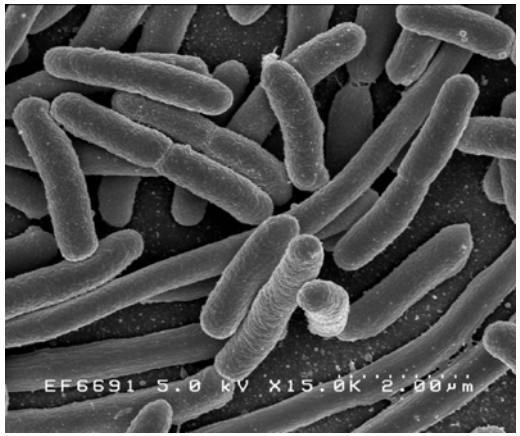
Scale is important !

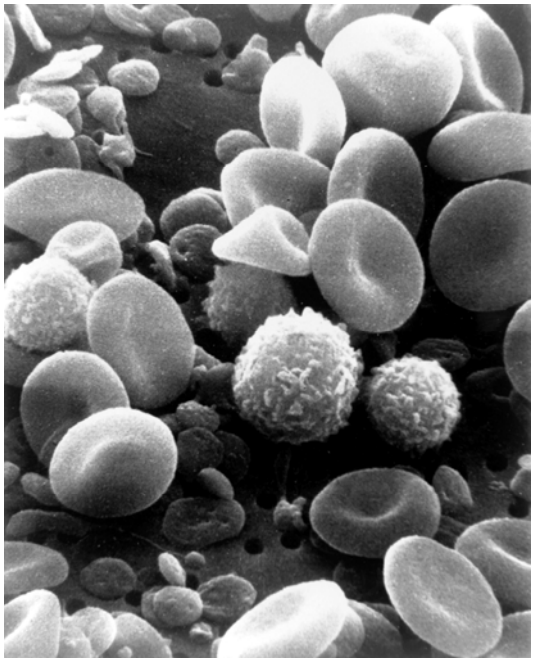


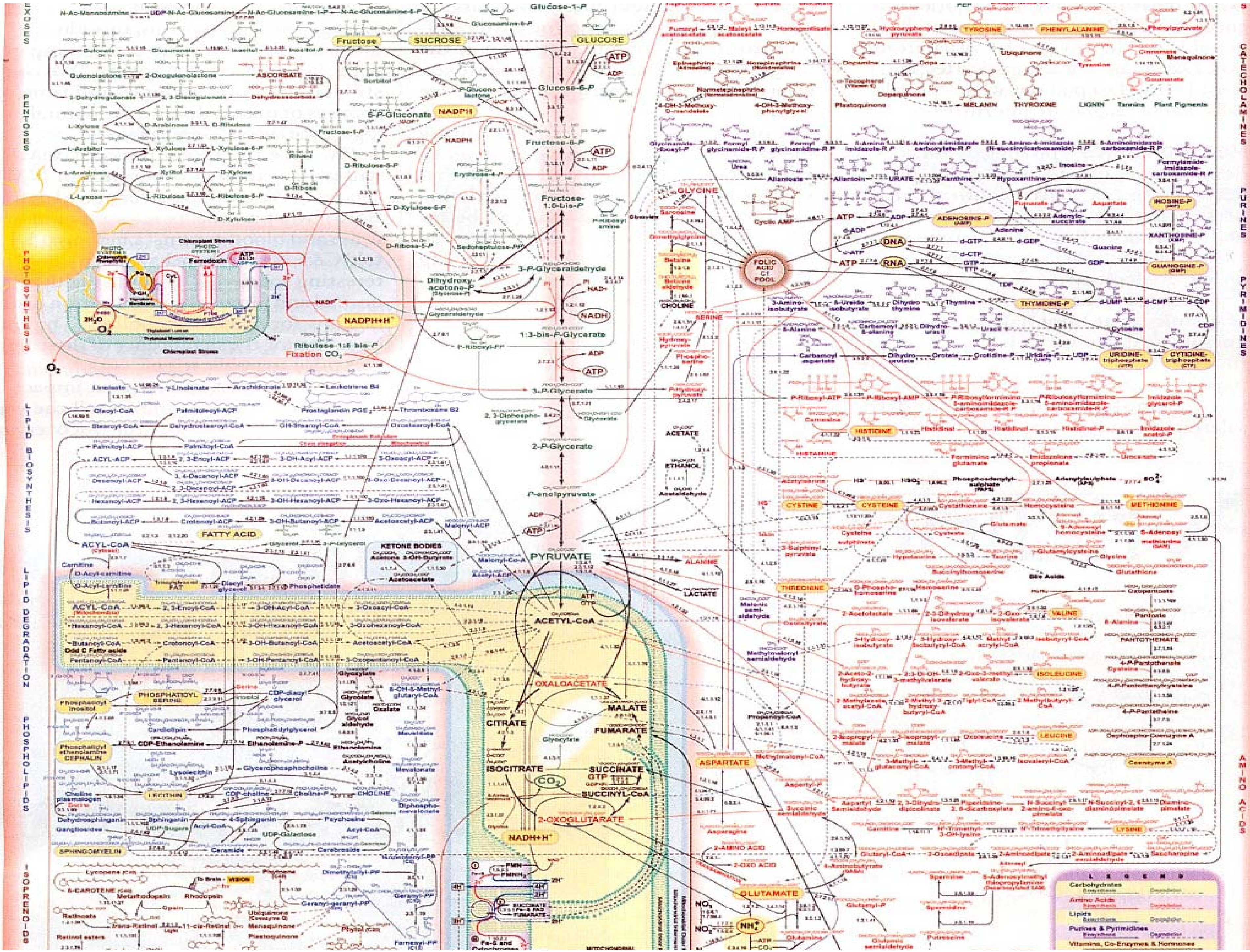


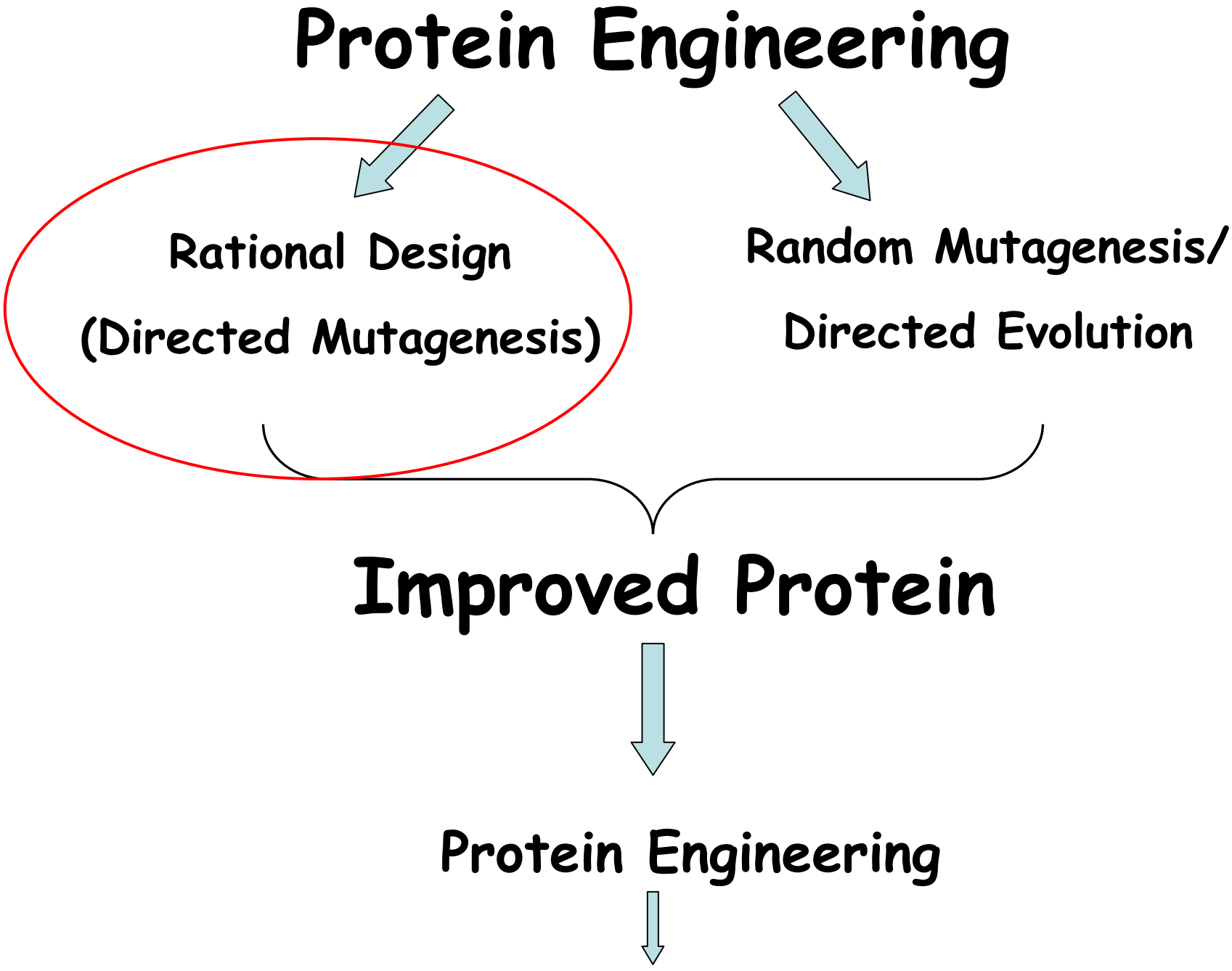
Biosystems
10E-10 → 10E7m



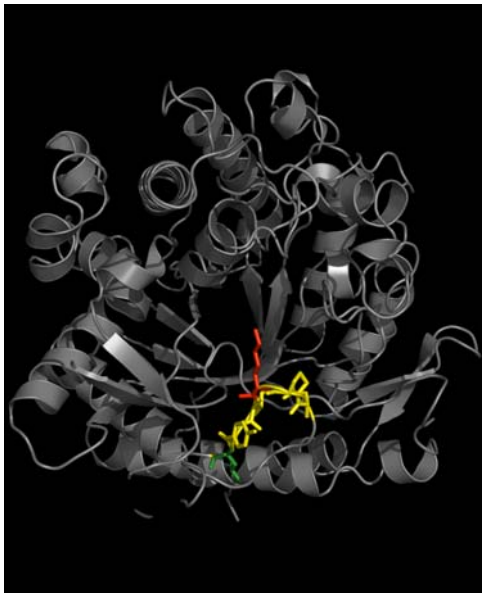





HEXOSES
PENTOSES
PHOTOSYNTHESIS
LIPID BIOSYNTHESIS
LIPID DEGRADATION
PHOSPHOLIPIDS
ISOPRENOIDS
CATECHOLAMINES
PURINES
PYRIMIDINES
AMINO ACIDS
Glucose-1-P
Fructose
SUCROSE
GLUCOSE
ATP
ADP
Glucose-6-P
6-P-Gluconate
NADPH
Fructose-6-P
Fructose-1:6-bis-P
3-P-Glyceraldehyde
Dihydroxy-acetone-P
NADPH+H+
NADH
1:3-bis-P-Glycerate
Ribulose-1:5-bis-P
Fixation CO_2
O_2
3-P-Glycerate
2-P-Glycerate
P-enolpyruvate
PYRUVATE
ACYL-CoA
FATTY ACID
KETONE BODIES
ACETYL-CoA
OXALOACETATE
CITRATE
MALATE
FUMARATE
ISOCITRATE
CO_2
SUCCINATE
GTP
SUCCINYL-CoA
2-OXOGLUTARATE
NADH+H+
LECITHIN
SPHINGOMYELIN
GLYCINE
SERINE
CHOLINE
FOLIC ACID C1 POOL
ATP
DNA
RNA
TYROSINE
PHENYLALANINE
ADENOSINE-P
INOSINE-P
GUANOSINE-P
THYMIDINE-P
URIDINE-triphosphate
CYTIDINE-triphosphate
HISTIDINE
CYSTINE
CYSTEINE
METHIONINE
ALANINE
THREONINE
VALINE
ISOLEUCINE
LEUCINE
ASPARTATE
LYSINE
GLUTAMATE
NH_4^+
ACETATE
ETHANOL
LACTATE
MELANIN
THYROXINE
LIGNIN
Coenzyme A
LEGEND
Carbohydrates
Amino Acids
Lipids
Purines & Pyrimidines
Vitamins, Co-Enzymes & Hormones


Protein Engineering
Rational Design
(Directed Mutagenesis)
Random Mutagenesis/
Directed Evolution
Improved Protein
Protein Engineering

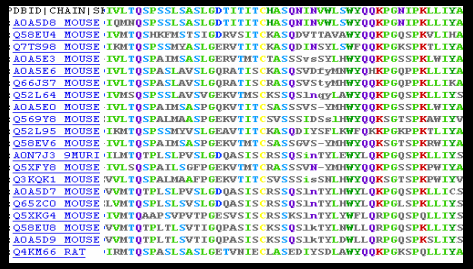
Rational Design (Directed Mutagenesis)



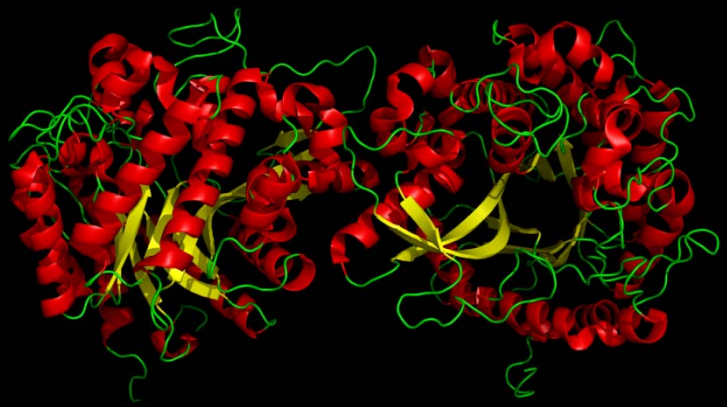
Google: AdvGentech5.ppt

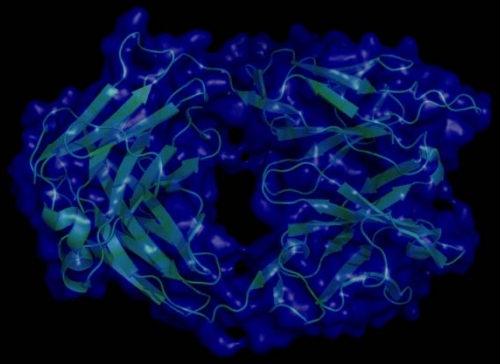
AALBORG UNIVERSITY

Bioinformatics



DNA and AA sequence information: composition, similarity, physicochemical properties, motifs, experimental design for directed or directed evolution procedure

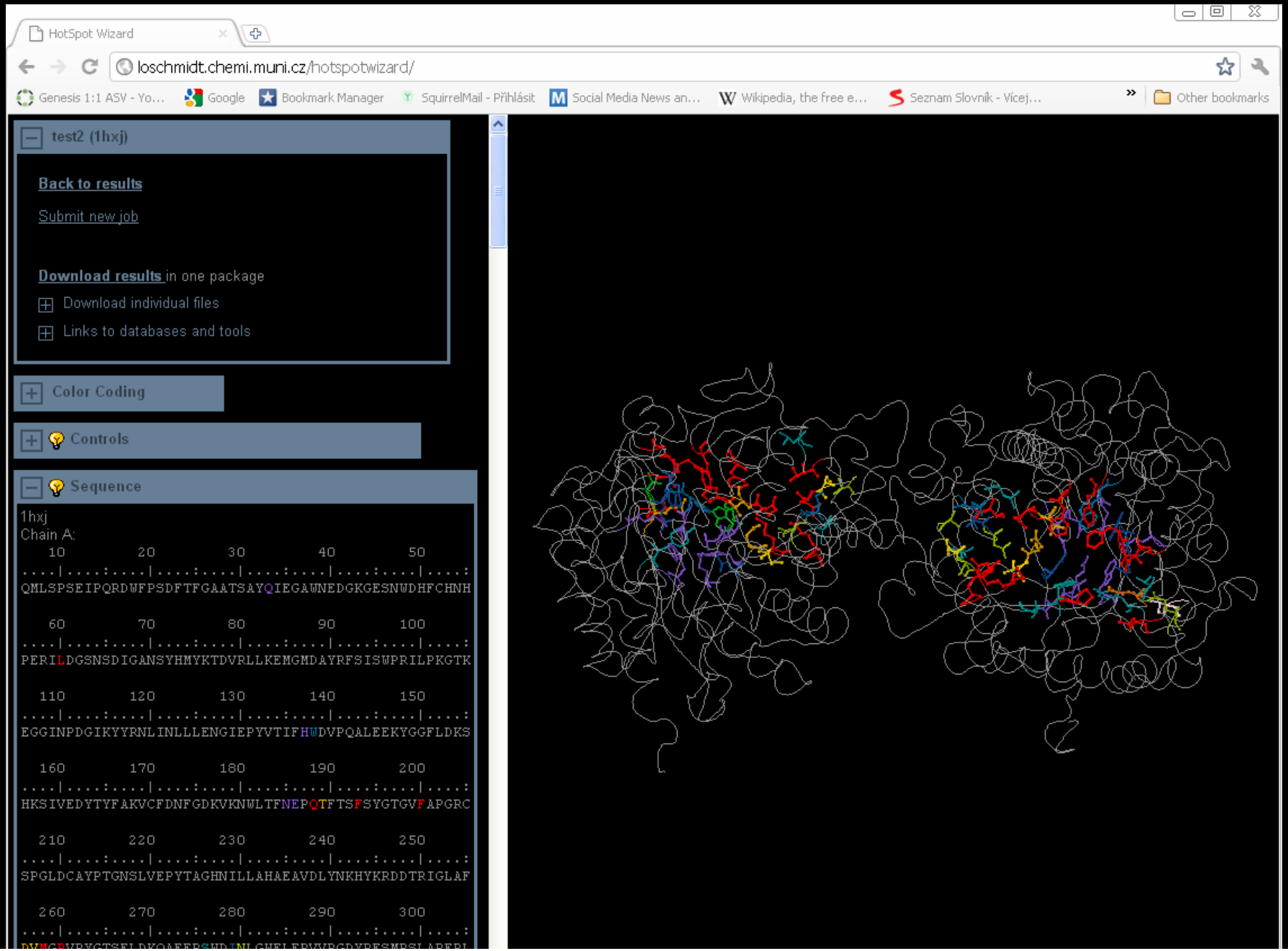




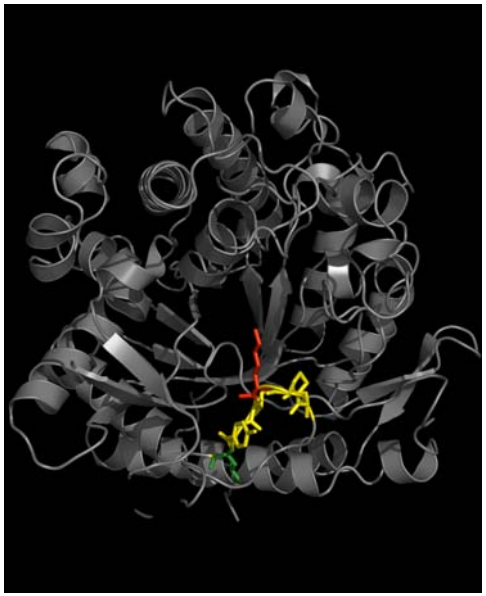
3D structure derived information: structure motifs, similarity, surfaces, canals and cavities, check and evaluation of aminoacids possible involved in misfolding

Lliterature search: current knowledge in this area, data mining, identification of possible problems, look for collaborations material etc.

HotSpot Wizard 1.7



Protein Engineering



Google: AdvGentech5.ppt

AALBORG UNIVERSITY

Protein Engineering

What can be engineered in Proteins ?

-> <u>Folding (+Structure):</u>

1. **Thermodynamic Stability**
   (Equilibrium between: Native $\Leftrightarrow$ Unfolded state)

2. **Thermal and Environmental Stability** (Temperature, pH, Solvent, Detergents, Salt …..)

Protein Engineering

What can be engineered in Proteins ?

-> Function:

1. **Binding** (Interaction of a protein with its surroundings)

How many points are required to bind a molecule with high affinity?

1. **Catalysis** (a different form of binding – binding the transition state of a chemical reaction)

Increased binding to the transition state $\Rightarrow$ increased catalytic rates !!!
Requires: Knowledge of the Catalytic Mechanism !!!

-> engineer Kcat and Km

Protein Engineering

Factors which contribute to stability:

1. Hydrophobicity (hydrophobic core)

1. Electrostatic Interactions:

   -> Salt Bridges
   -> Hydrogen Bonds
   -> Dipole Interactions

1. Disulfide Bridges

1. Metal Binding (Metal chelating site)

1. Reduction of the unfolded state entropy with X $\rightarrow$ Pro mutations

Protein Engineering

Design of Thermal and Environmental stability:

1. Stabilization of α-Helix Macrodipoles

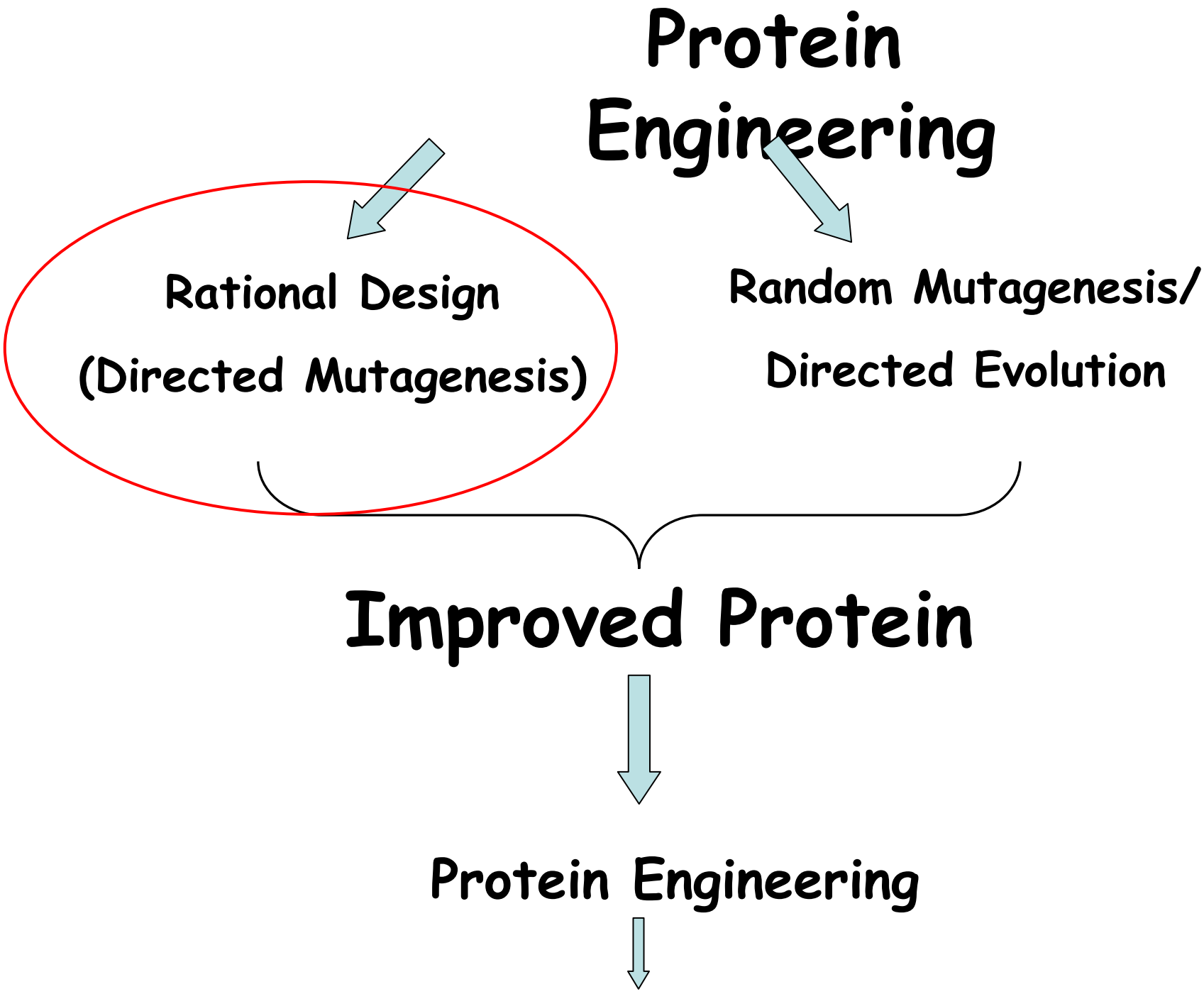
1. Engineer Structural Motifes (like Helix N-Caps)

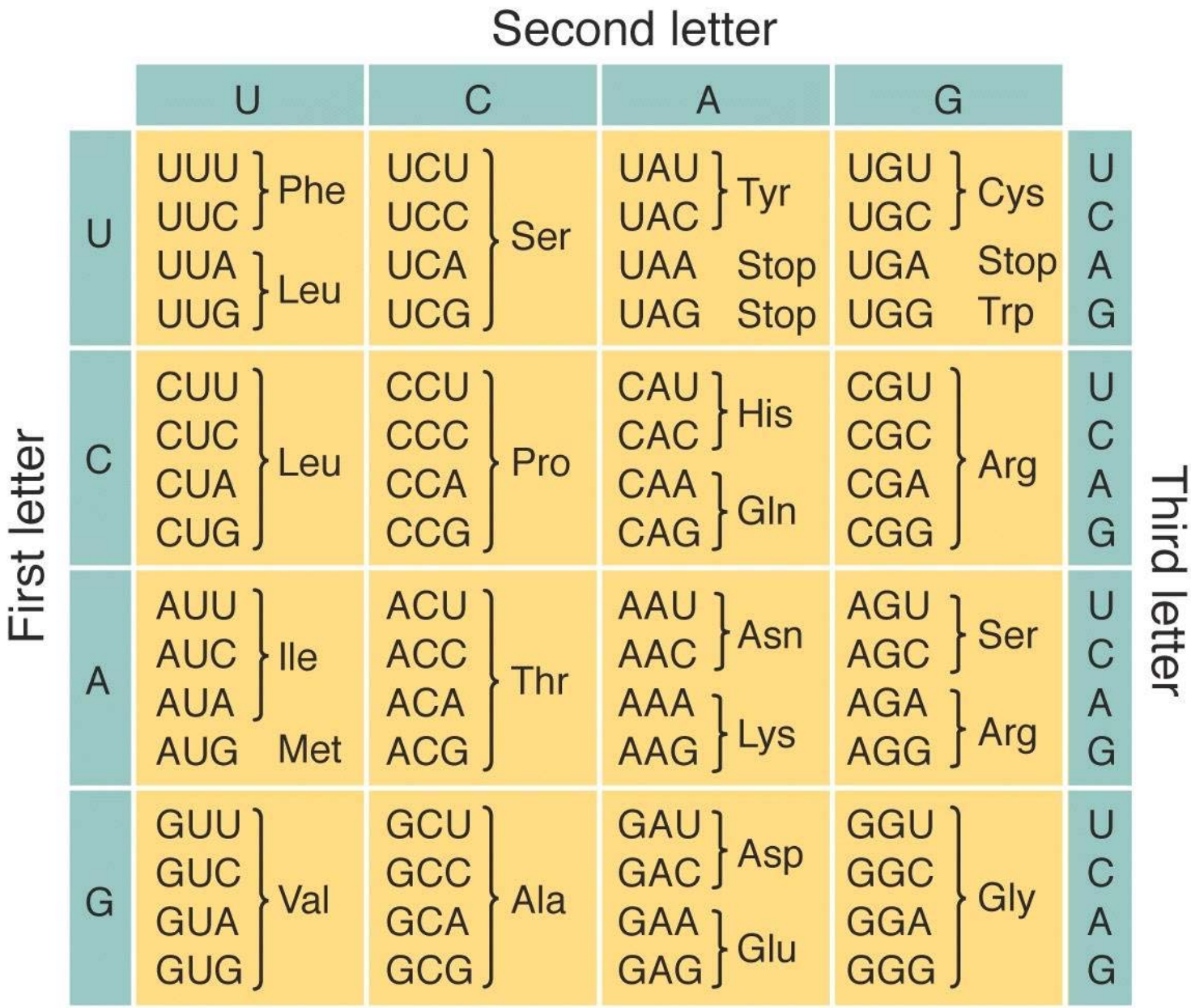
1. Introduction of salt bridges

1. Introduction of residues with higher intrinsic properties for their conformational state (e.g. Ala replacement within a α-Helix)

1. Introduction of disulfide bridges

1. Reduction of the unfolded state entropy with $X \rightarrow$ Pro mutations


Protein Engineering
Rational Design
(Directed Mutagenesis)
Random Mutagenesis/
Directed Evolution
Improved Protein
Protein Engineering


Second letter
U
C
A
G
First letter
Third letter
U
UUU Phe
UUC
UUA Leu
UUG
UCU Ser
UCC
UCA
UCG
UAU Tyr
UAC
UAA Stop
UAG Stop
UGU Cys
UGC
UGA Stop
UGG Trp
C
CUU Leu
CUC
CUA
CUG
CCU Pro
CCC
CCA
CCG
CAU His
CAC
CAA Gln
CAG
CGU Arg
CGC
CGA
CGG
A
AUU Ile
AUC
AUA
AUG Met
ACU Thr
ACC
ACA
ACG
AAU Asn
AAC
AAA Lys
AAG
AGU Ser
AGC
AGA Arg
AGG
G
GUU Val
GUC
GUA
GUG
GCU Ala
GCC
GCA
GCG
GAU Asp
GAC
GAA Glu
GAG
GGU Gly
GGC
GGA
GGG

Mutagenesis

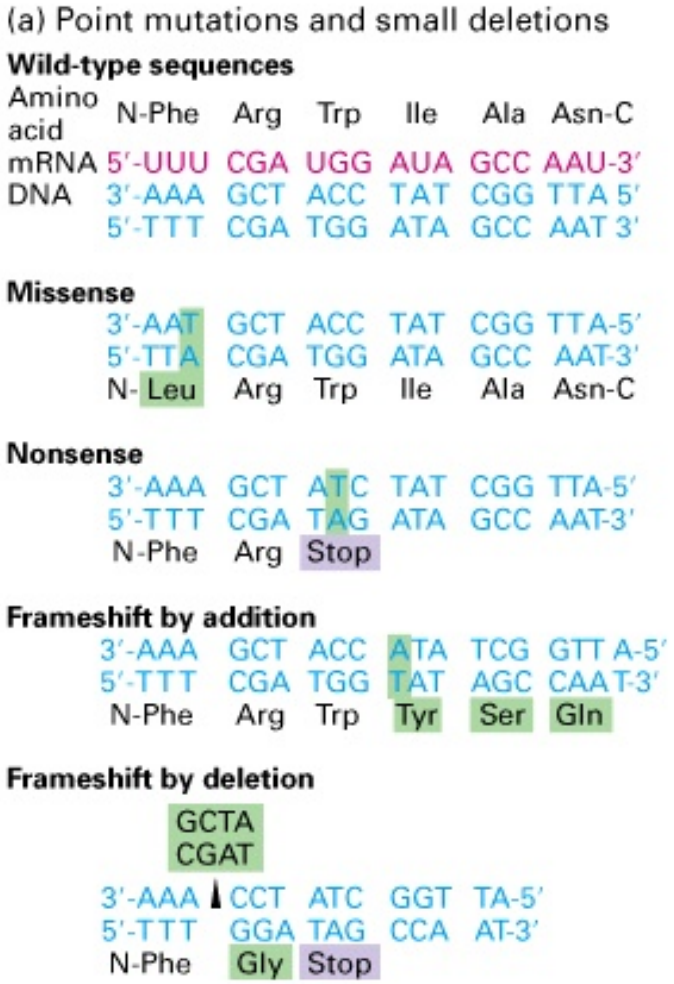
Mutagenesis -> change in DNA sequence

-> Point mutations or large modifications

Point mutations (directed mutagenesis):

- Substitution: change of one nucleotide (i.e. A-> C)
- Insertion: gaining one additional nucleotide
- Deletion: loss of one nucleotide

Consequences of point mutations within a coding sequence (gene) for the protein



Silent mutations:

-> change in nucleotide sequence with no consequences for protein sequence

-> Change of amino acid

-> truncation of protein

-> change of c-terminal part of protein

-> change of c-terminal part of protein

Codon Usage is different in different species

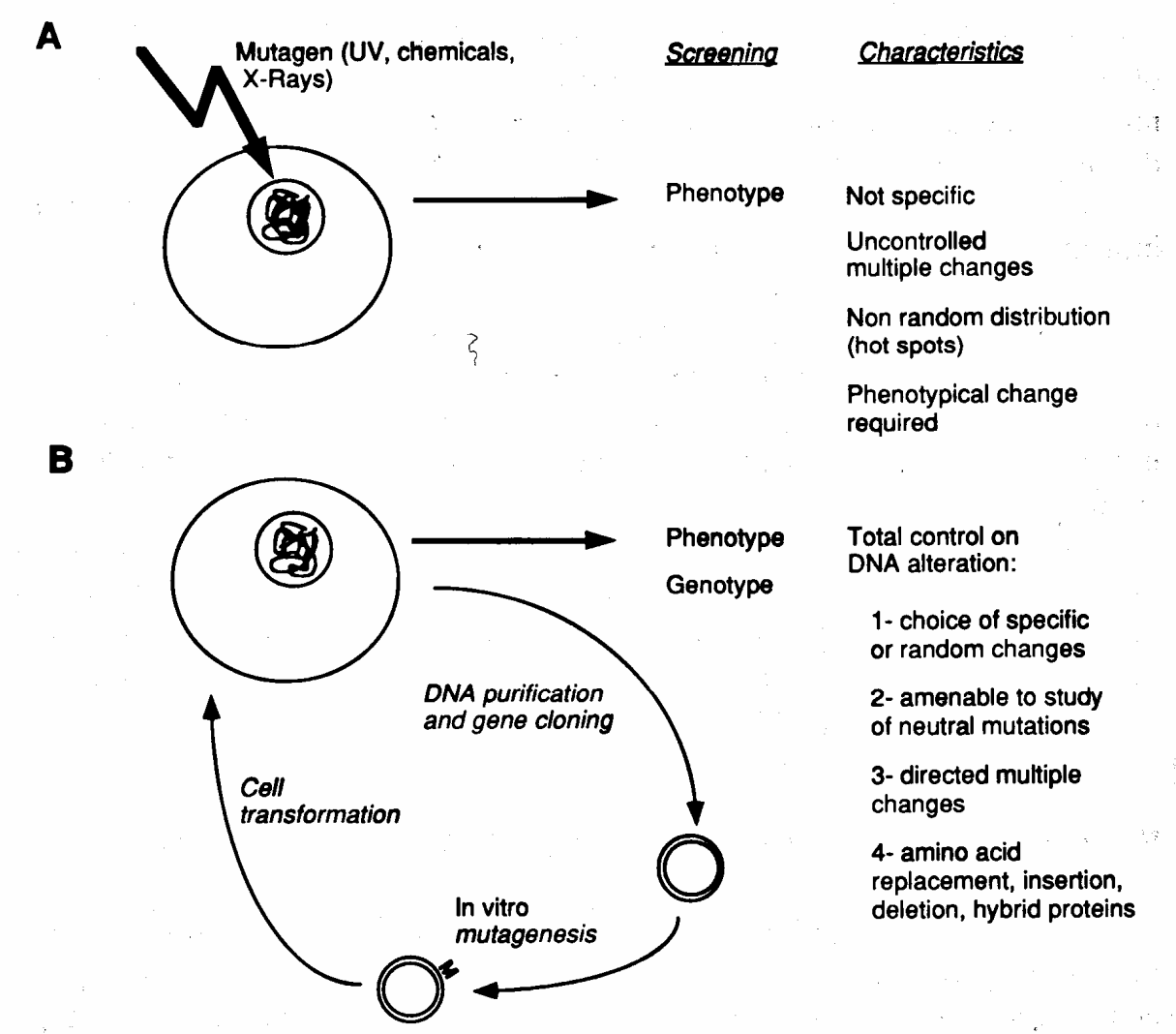
UUU F 0.57	UCU S 0.11	UAU Y 0.53	UGU C 0.42
UUC F 0.43	UCC S 0.11	UAC Y 0.47	UGC C 0.58
UUA L 0.15	UCA S 0.15	UAA * 0.64	UGA * 0.36
UUG L 0.12	UCG S 0.16	UAG * 0.00	UGG W 1.00
CUU L 0.12	CCU P 0.17	CAU H 0.55	CGU R 0.36
CUC L 0.10	CCC P 0.13	CAC H 0.45	CGC R 0.44
CUA L 0.05	CCA P 0.14	CAA Q 0.30	CGA R 0.07
CUG L 0.46	CCG P 0.55	CAG Q 0.70	CGG R 0.07
AUU I 0.58	ACU T 0.16	AAU N 0.47	AGU S 0.14
AUC I 0.35	ACC T 0.47	AAC N 0.53	AGC S 0.33
AUA I 0.07	ACA T 0.13	AAA K 0.73	AGA R 0.02
AUG M 1.00	ACG T 0.24	AAG K 0.27	AGG R 0.03
GUU V 0.25	GCU A 0.11	GAU D 0.65	GGU G 0.29
GUC V 0.18	GCC A 0.31	GAC D 0.35	GGC G 0.46
GUA V 0.17	GCA A 0.21	GAA E 0.70	GGA G 0.13
GUG V 0.40	GCG A 0.38	GAG E 0.30	GGG G 0.12

[Codon/a.a./fraction per codon per a.a.]
E. coli K12 data from the Codon Usage Database

UUU F 0.46	UCU S 0.19	UAU Y 0.44	UGU C 0.46
UUC F 0.54	UCC S 0.22	UAC Y 0.56	UGC C 0.54
UUA L 0.08	UCA S 0.15	UAA * 0.30	UGA * 0.47
UUG L 0.13	UCG S 0.05	UAG * 0.24	UGG W 1.00
CUU L 0.13	CCU P 0.29	CAU H 0.42	CGU R 0.08
CUC L 0.20	CCC P 0.32	CAC H 0.58	CGC R 0.18
CUA L 0.07	CCA P 0.28	CAA Q 0.27	CGA R 0.11
CUG L 0.40	CCG P 0.11	CAG Q 0.73	CGG R 0.20
AUU I 0.36	ACU T 0.25	AAU N 0.47	AGU S 0.15
AUC I 0.47	ACC T 0.36	AAC N 0.53	AGC S 0.24
AUA I 0.17	ACA T 0.28	AAA K 0.43	AGA R 0.21
AUG M 1.00	ACG T 0.11	AAG K 0.57	AGG R 0.21
GUU V 0.18	GCU A 0.27	GAU D 0.46	GGU G 0.16
GUC V 0.24	GCC A 0.40	GAC D 0.54	GGC G 0.34
GUA V 0.12	GCA A 0.23	GAA E 0.42	GGA G 0.25
GUG V 0.46	GCG A 0.11	GAG E 0.58	GGG G 0.25

[Codon/a.a./fraction per codon per a.a.]
Homo sapiens data from the Codon Usage Database

Mutagenesis
Comparison of cellular and invitro mutagenesis



The Nobel Prize in Chemistry 1993

"for contributions to the developments of methods within DNA-based chemistry"

"for his invention of the polymerase chain reaction (PCR) method"

Kary B. Mullis

1/2 of the prize

USA

La Jolla, CA, USA

b. 1944

"for his fundamental contributions to the establishment of oligonucleotide-based, site-directed mutagenesis and its development for protein studies"

Michael Smith

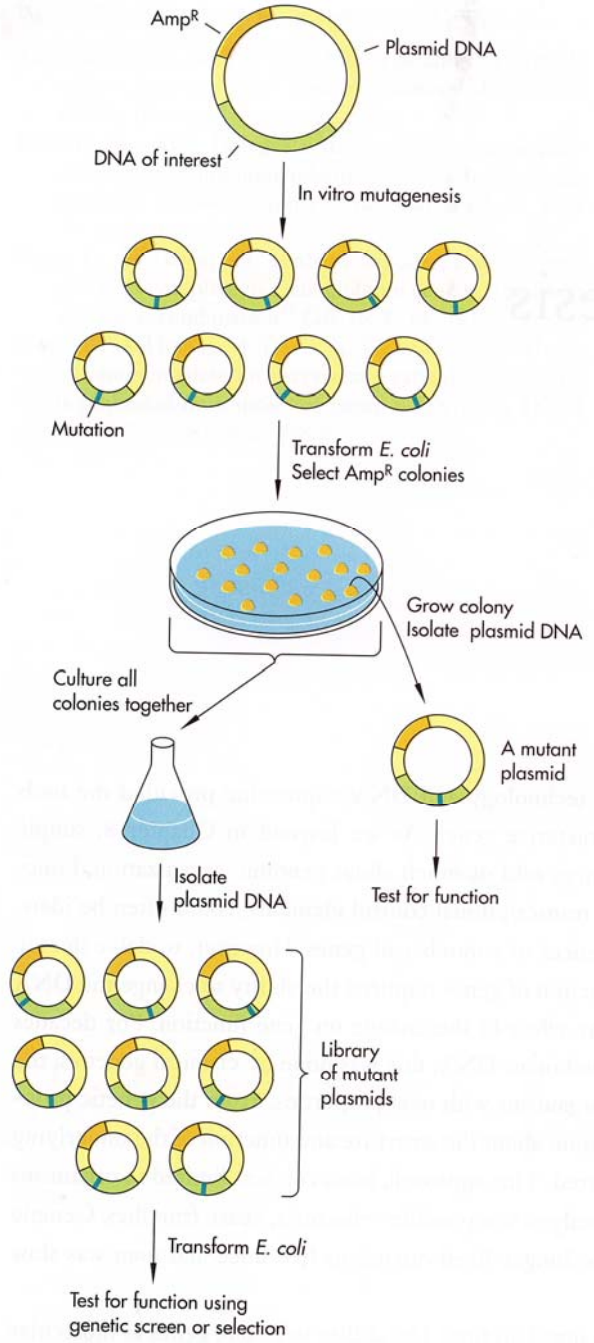
1/2 of the prize

Canada

University of British Columbia
Vancouver, Canada

b. 1932
(in Blackpool, United Kingdom)
d. 2000

Titles, data and places given above refer to the time of the award.



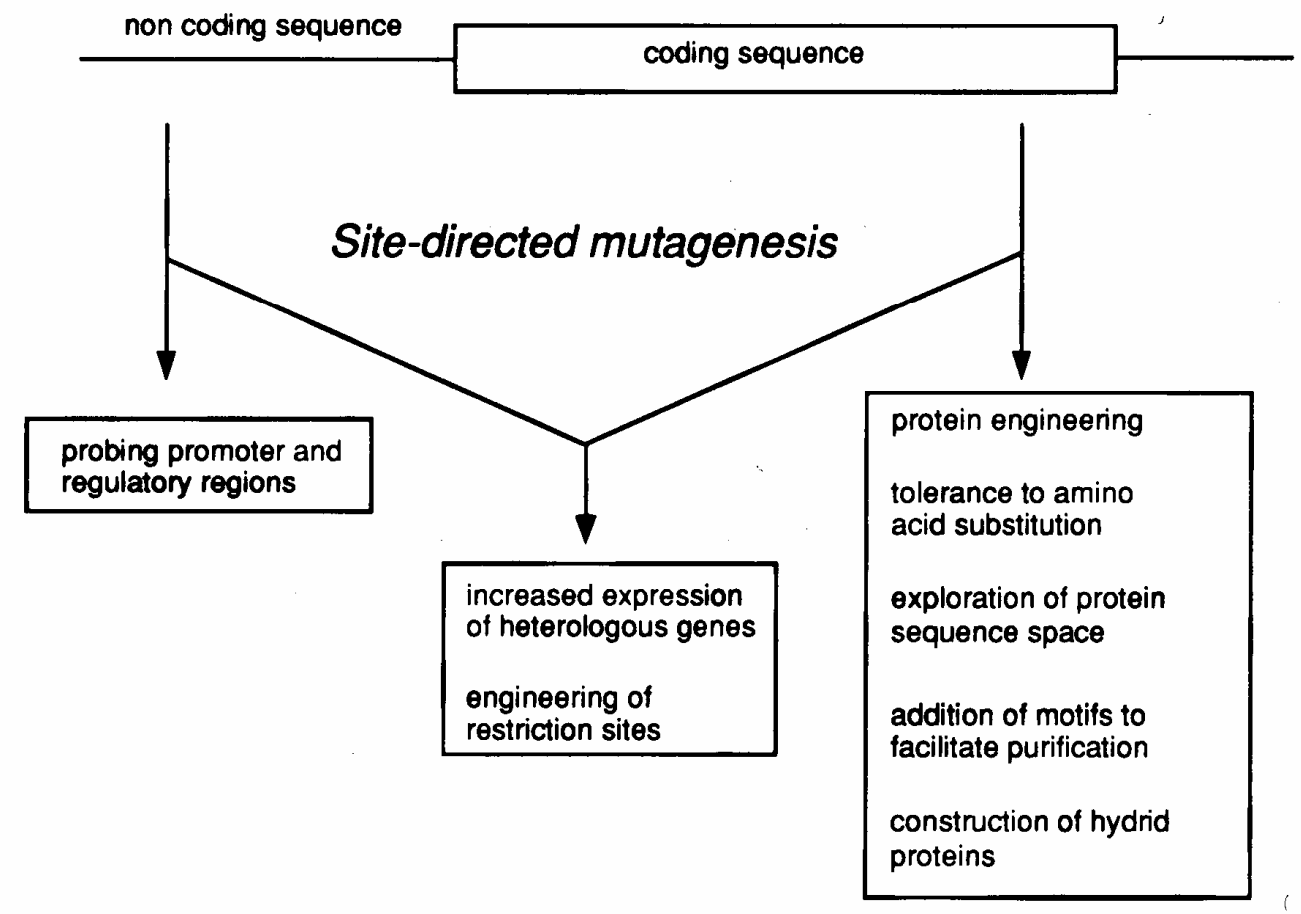
General strategy for directed mutagenesis

Requirements:

- DNA of interest (gene or promoter) must be cloned

- Expression system must be available -> for testing phenotypic change

Applications of directed mutagenesis



Approaches for directed mutagenesis

-> site-directed mutagenesis

-> point mutations in particular known area

result -> library of wild-type and mutated DNA (site-specific)

not really a library -> just 2 species

Protein Engineering

-> Mutagenesis used for modifying proteins
Replacements on protein level -> mutations on DNA level

Assumption : Natural sequence can be modified to
improve a certain function of protein

This implies:

- Protein is NOT at an optimum for that function
- Sequence changes without disruption of the structure
- (otherwise it would not fold)
- New sequence is not TOO different from the native sequence (otherwise loss in function of protein)

consequence -> introduce point mutations

Rational Protein Design

$\Rightarrow$ Site –directed mutagenesis

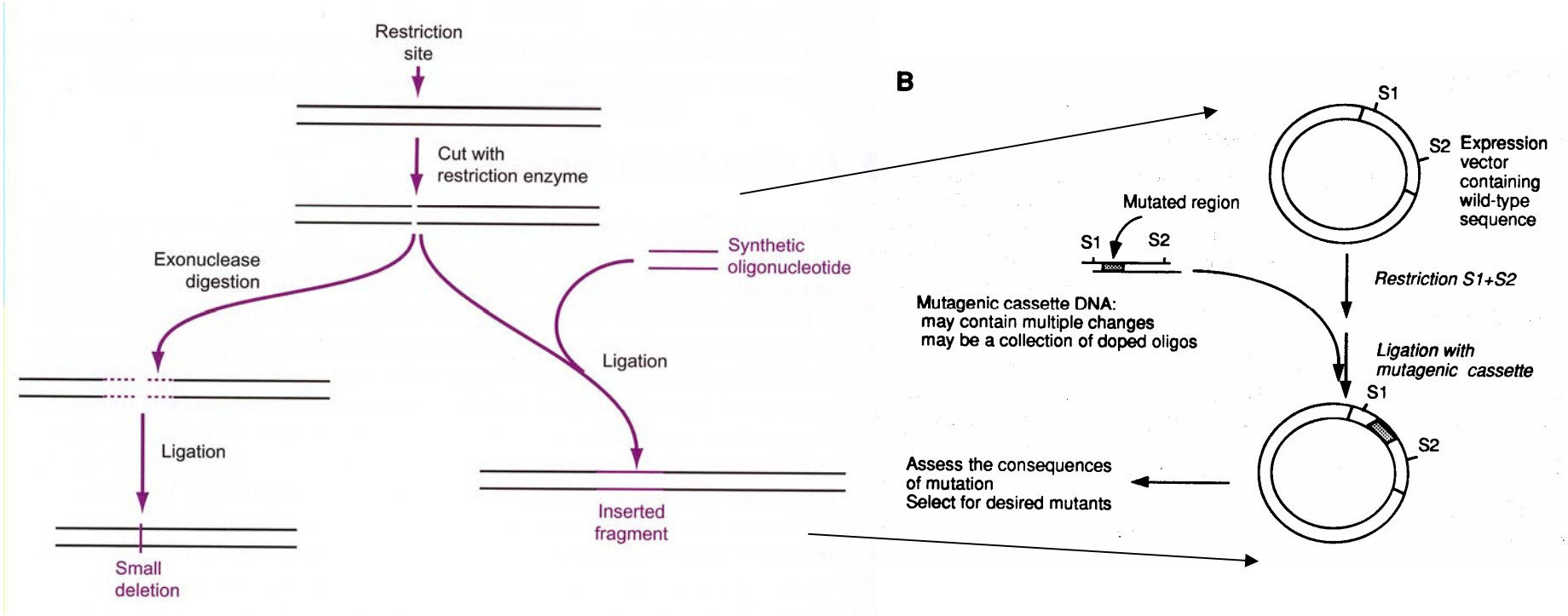
Requirements:

-> Knowledge of sequence and preferable Structure
(active site,….)

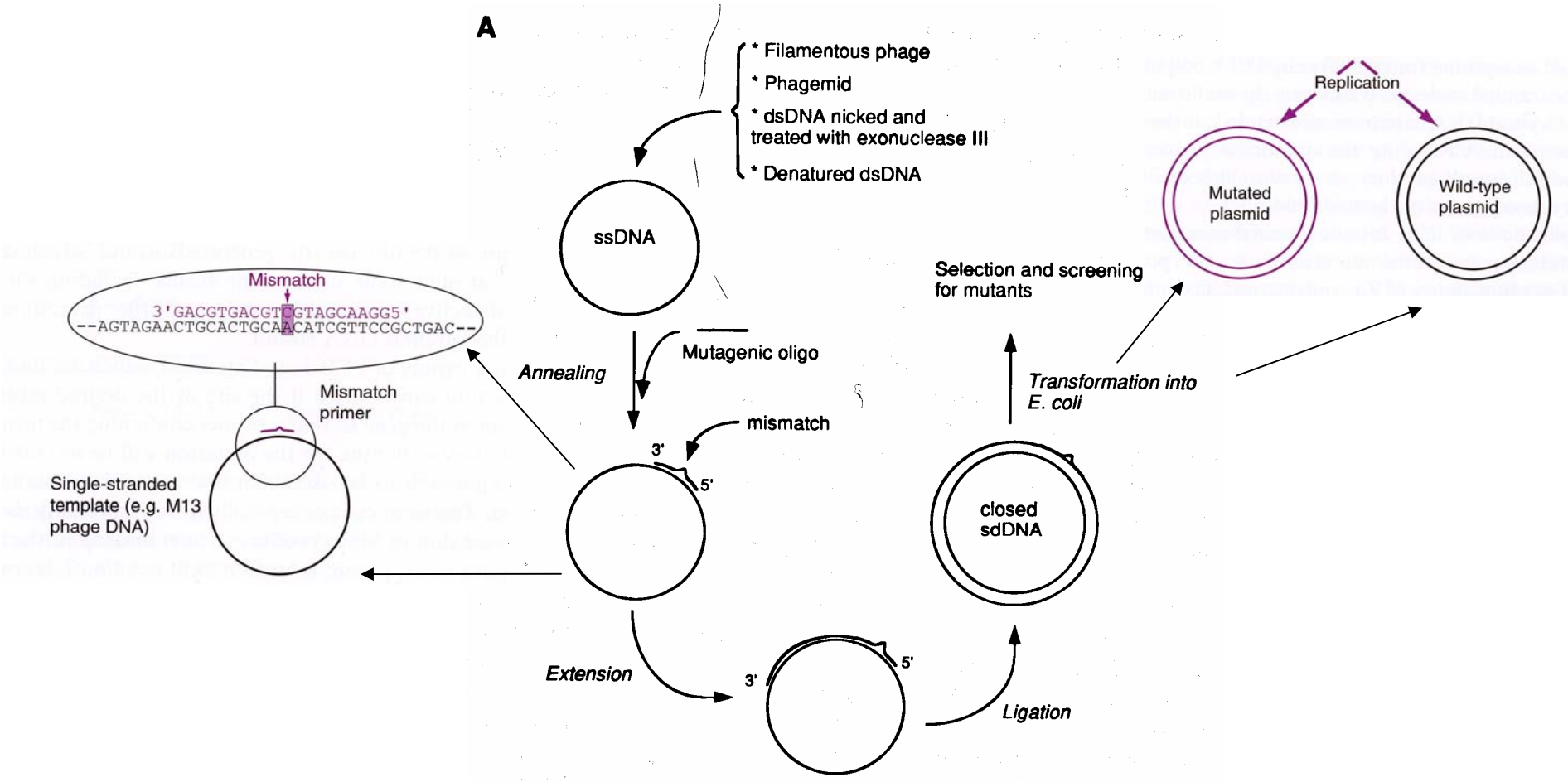
-> Understanding of mechanism
(knowledge about structure – function relationship)

-> Identification of cofactors……..

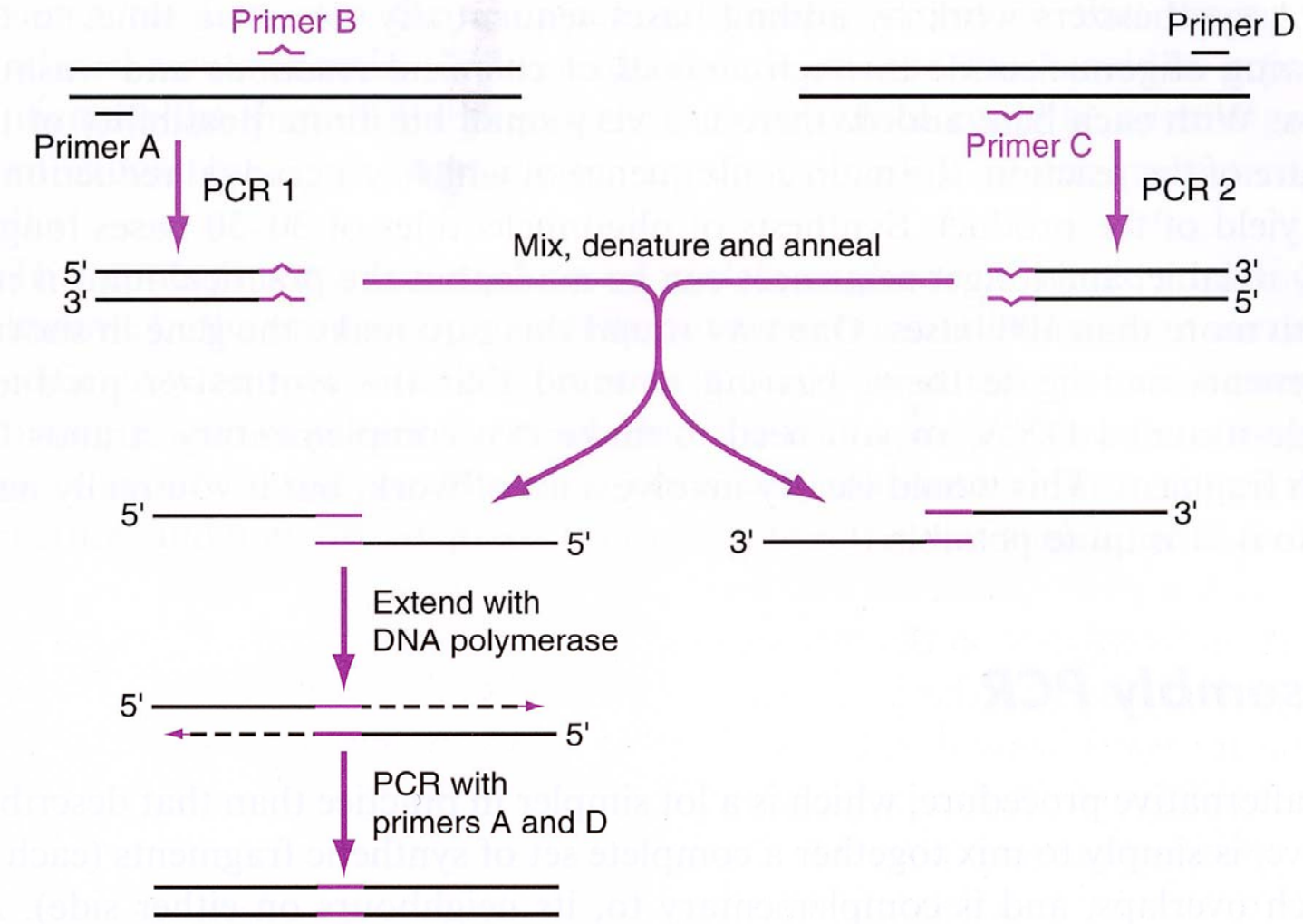
Site-directed mutagenesis methods

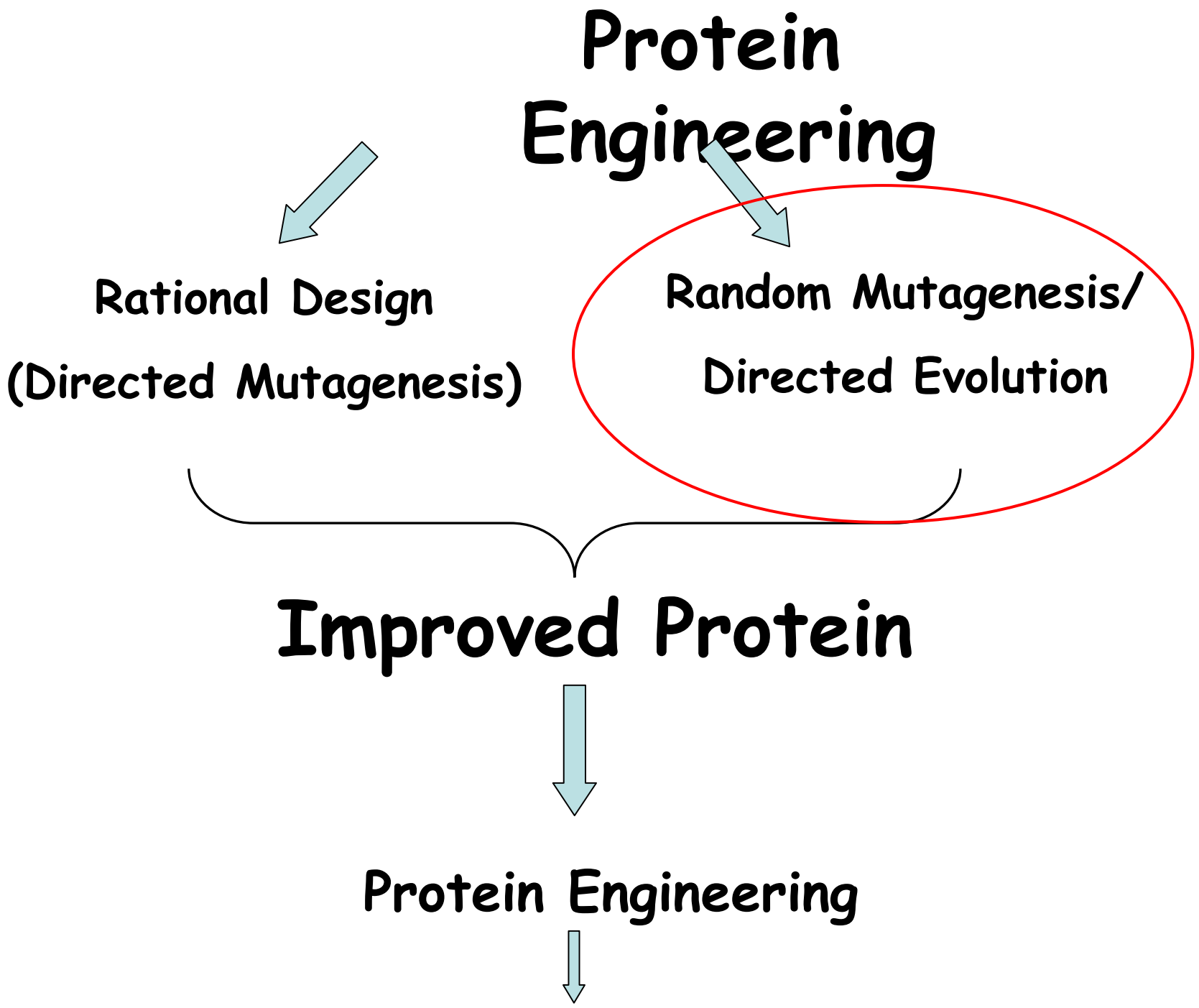


Site-directed mutagenesis methods - Oligonucleotide - directed method



Site-directed mutagenesis methods – PCR based




Protein Engineering
Rational Design
(Directed Mutagenesis)
Random Mutagenesis/
Directed Evolution
Improved Protein
Protein Engineering

The concept of laboratory-directed protein evolution is not new.

Systematic approaches to directed evolution of proteins
have been documented since the 1970s

One early example is the evolution of the EbgA protein from *Escherichia coli*, an enzyme having almost no -galactosidase activity. Through intensive selection of a LacZ deletion strain of *E. coli* for growth on lactose as a sole carbon source, the wild-type EbgA was “evolved” as a -galactosidase sufficient to replace the *lacZ* gene function (Campbell, J.1973).

Directed Evolution – Random mutagenesis

-> based on the process of natural evolution

- **NO structural information required**

- **NO understanding of the mechanism required**

General Procedure:

1. Generation of genetic diversity
   ⇒ Random mutagenesis

2. Identification of successful variants
   ⇒ Screening and seletion

Protein Engineering
Directed Evolution

Successful directed evolution has four requirements:

(i) the desired function should be physically feasible,

(ii) the function should be biologically or evolutionary feasible, i.e., a mutational pathway must exist to get from an original protein to tailored protein through ever-improving variants,

(iii) it should be possible to make libraries of mutants complex enough to contain rare beneficial mutations and

(iv) a rapid screen or selection reflecting the desired function should be available (Arnold,1998)

Approaches for directed random mutagenesis

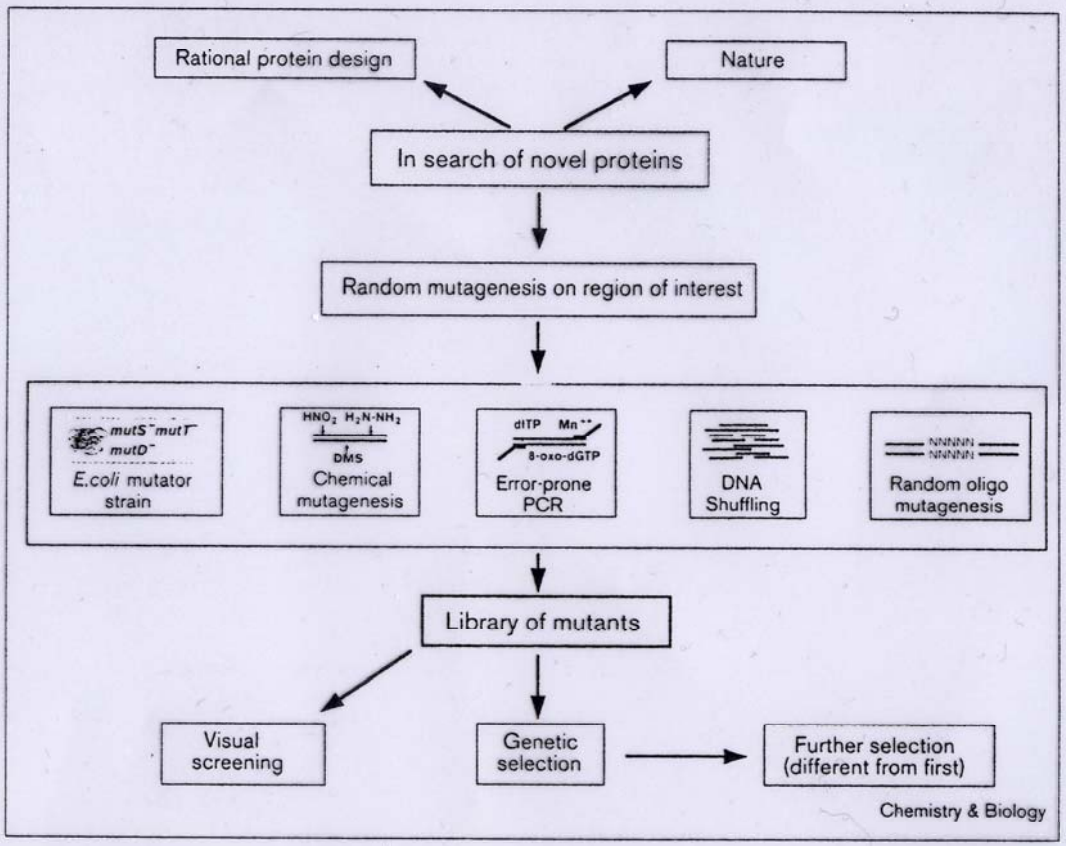
-> random mutagenesis

-> point mutations in all areas within DNA of interest

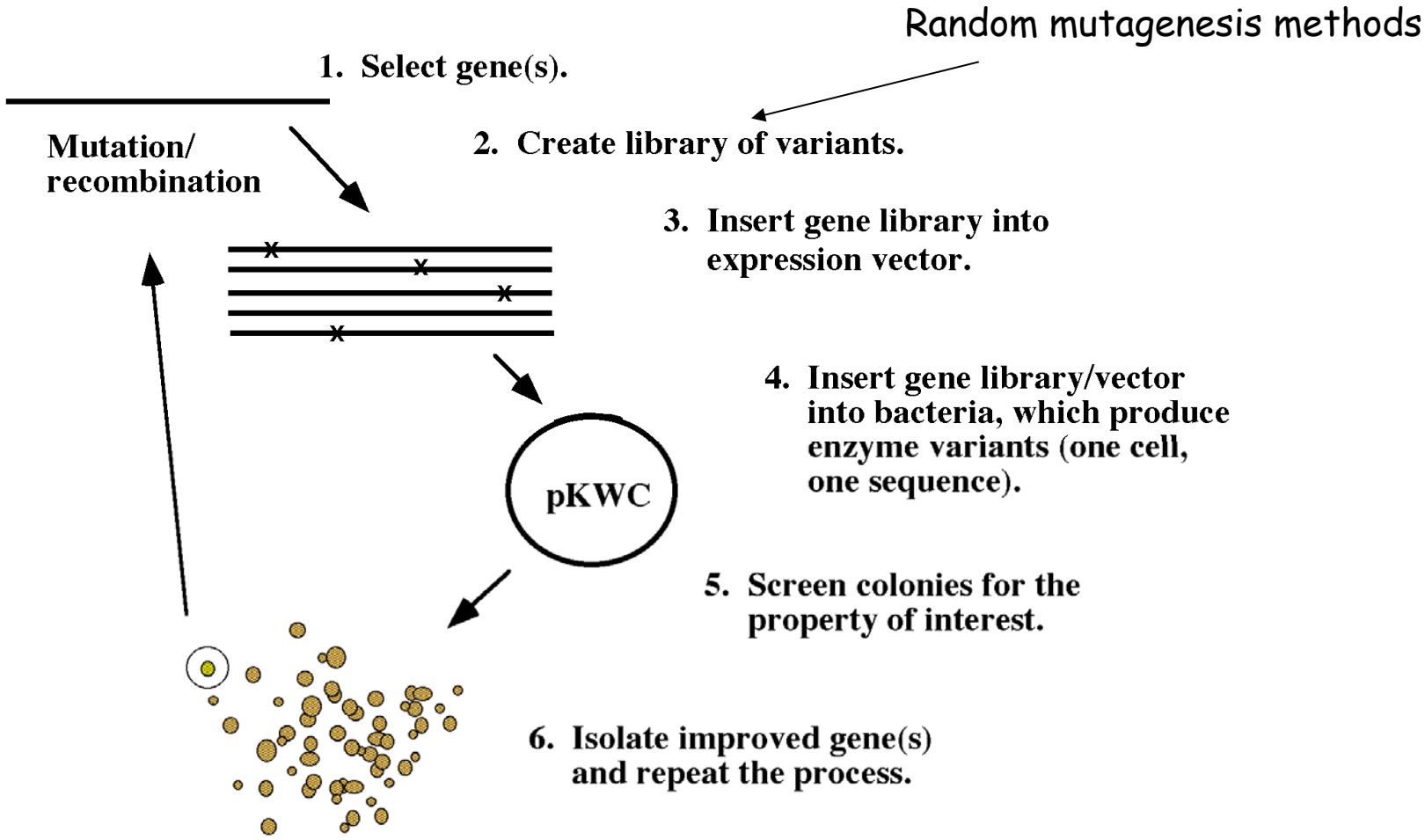
result -> library of wild-type and mutated DNA (random)
a real library -> many variants -> screening !!!

if methods efficient -> mostly mutated DNA

The creation of random libraries by applied molecular evolution. A gene target can be randomized by one of several methods, followed by ligation of the library into a vector backbone and transformation into an appropriate strain of *Escherichia coli* for selection/screening.



General Directed Evolution Procedure



Limitation of Directed Evolution

1. Evolutionary path must exist - > to be successful

1. Screening method must be available

   -> You get (exactly) what you ask for!!!

   -> need to be done in -> High throughput !!!

Typical Directed Evolution Experiment

Successful experiments involve generally
less than 6 steps (cycles)!!!

Why?

1. Sequences with improved properties are rather close to the parental sequence -> along a evolutionary path

2. Capacity of our present methods to generate novel functional sequences is rather limited -> requires huge libraries

⇒ Point Mutations !!!

Evolutionary Methods

- **Non-recombinative methods:**

  -> Oligonucleotide Directed Mutagenesis (saturation mutagenesis)

  -> Chemical Mutagenesis, Bacterial Mutator Strains

  -> Error-prone PCR

- **Recombinative methods** -> Mimic nature's recombination strategy

  Used for: Elimination of neutral and deleterious mutations

  -> DNA shuffling

  -> Invivo Recombination (Yeast)

  -> Random priming recombination, Staggered extention precess (StEP)

  -> ITCHY

Evolutionary Methods
Type of mutation – Fitness of mutants

Type of mutations:

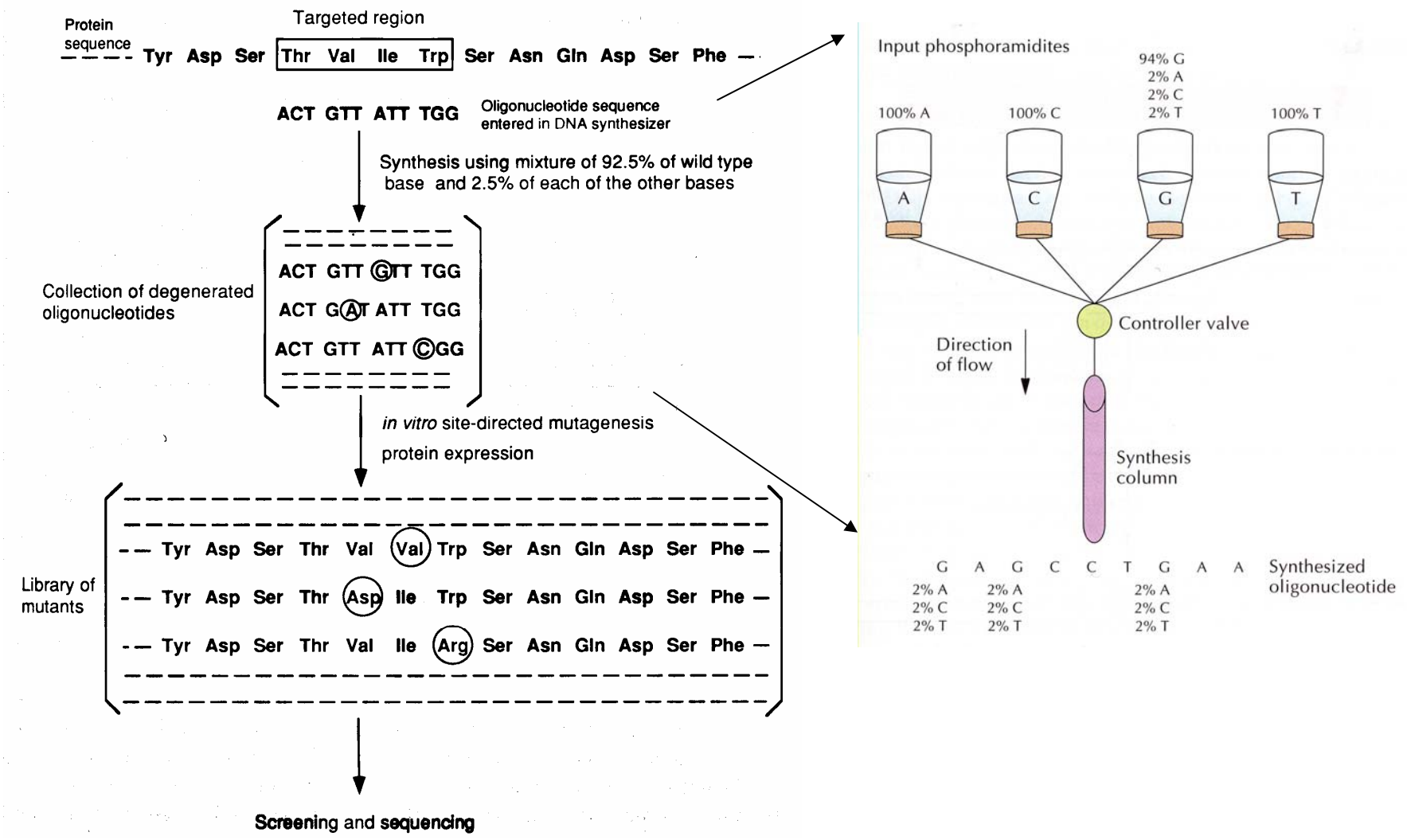
⇒ Beneficial mutations (good)
⇒ Neutral mutations
⇒ Deleterious mutations (bad)

⇒ Beneficial mutations are diluted with neutral and deleterious ones

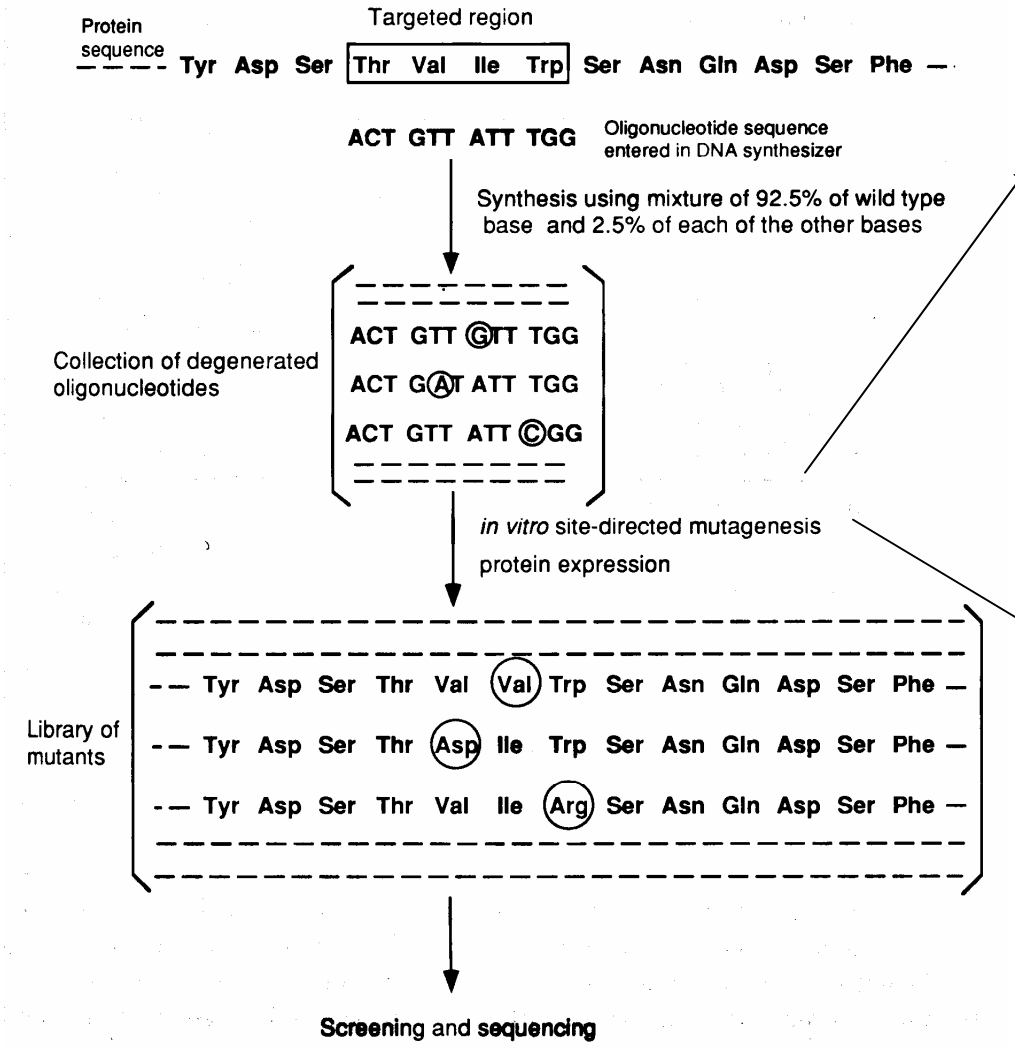
!!! Keep the number of mutations low per cycle

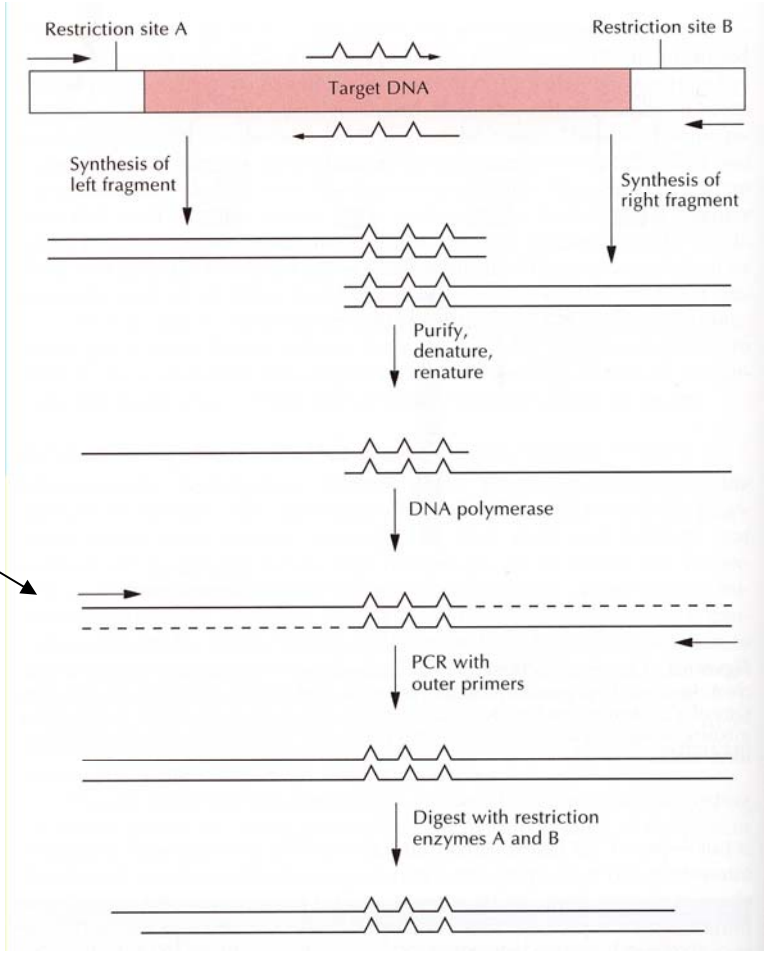
-> improve fitness of mutants!!!

Random Mutagenesis (PCR based) with degenerated primers (saturation mutagenesis)

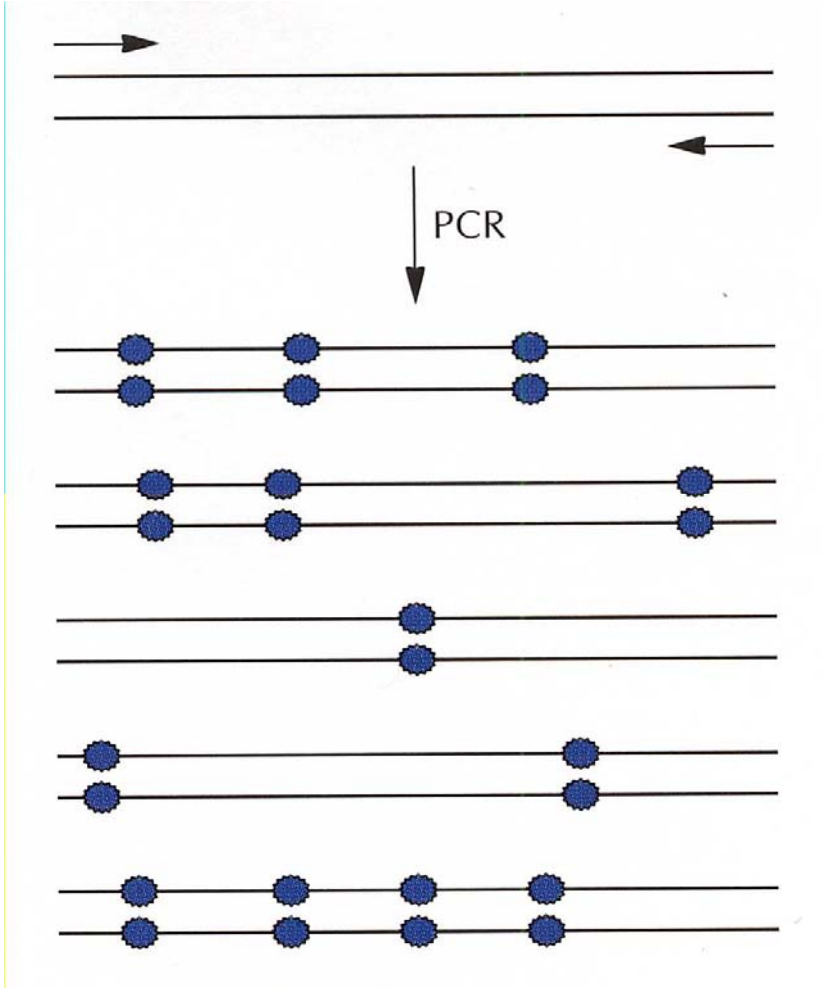


Random Mutagenesis (PCR based) with degenerated primers (saturation mutagenesis)





Random Mutagenesis (PCR based)
Error –prone PCR

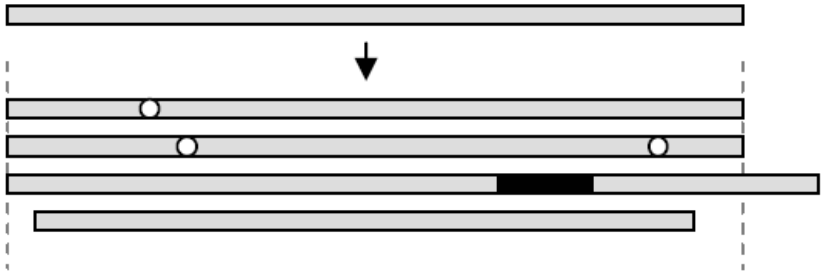


-> PCR with low fidelity !!!

Achieved by:

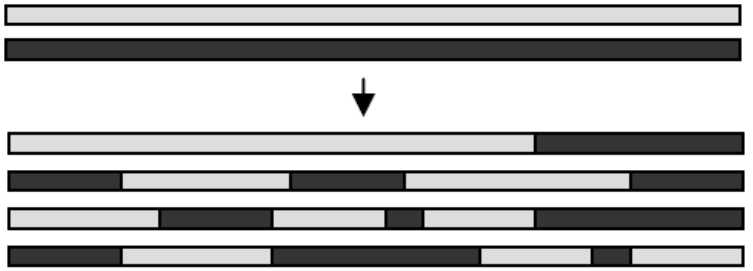
- Increased Mg2+ concentration
- Addition of Mn2+
- Not equal concentration of the four dNTPs
- Use of dITP
- Increasing amount of Taq polymerase (Polymerase with NO proof reading function)

Random mutagenesis



- Chemical mutagenesis
- Combinatorial cassette mutagenesis
- Error-prone PCR
- Mutator strains
- RID
- SeSAM
- Saturation mutagenesis
- UV irradiation

Recombination

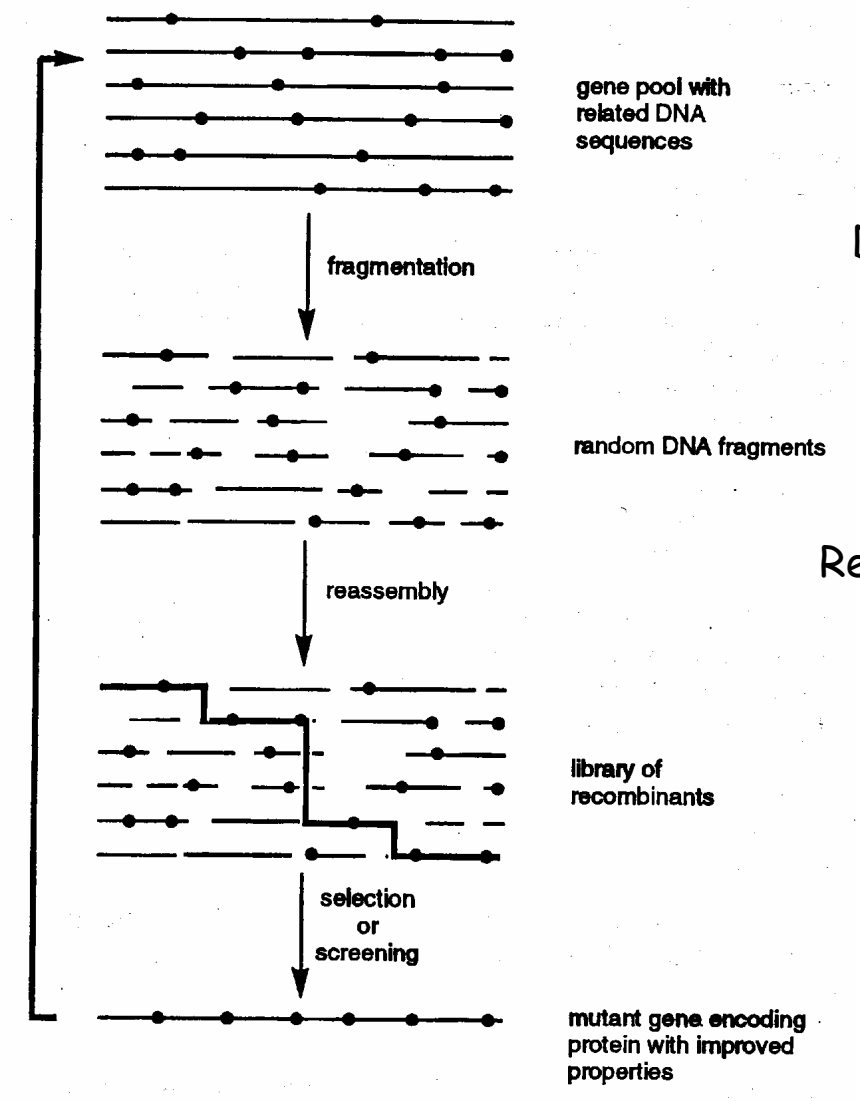


Homologous
- DNA shuffling
- DOGS
- Family shuffling
- Family shuffling with restriction enzymes
- RACHITT
- RPR
- StEP
- Genome shuffling

Non-homologous
- Exon shuffling
- DHR
- ITCHY
- THIO-ITCHY
- RM-PCR
- SCRATCHY
- SHIPREC
- SISDC
- YLBS

Fig. 1. Comparison of (a) random mutagenesis and (b) recombination strategies.

Random Mutagenesis (PCR based)
DNA Shuffling

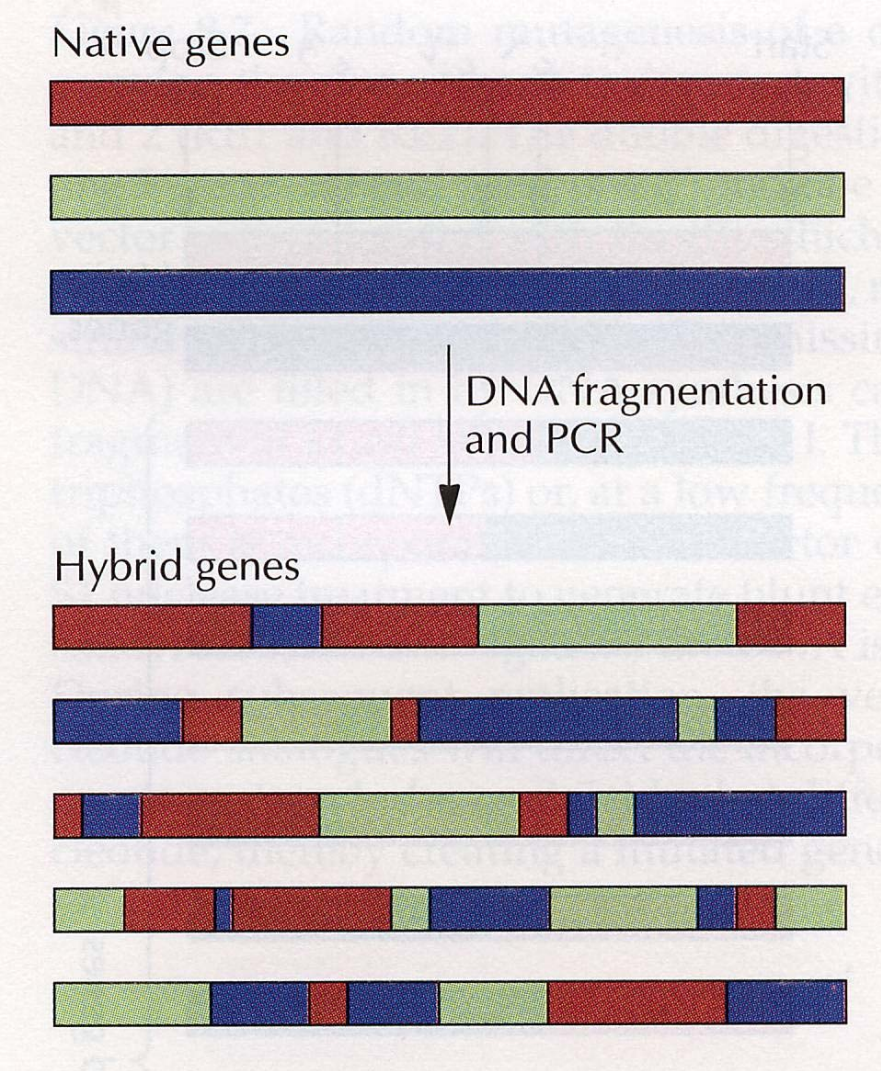


DNase I treatment (Fragmentation, 10-50 bp, Mn^{2+})

Reassembly (PCR without primers, Extension and Recombination)

PCR amplification

Random Mutagenesis (PCR based)
Family Shuffling



Genes coming from the same gene family -> highly homologous

-> Family shuffling

Random Mutagenesis (PCR based)

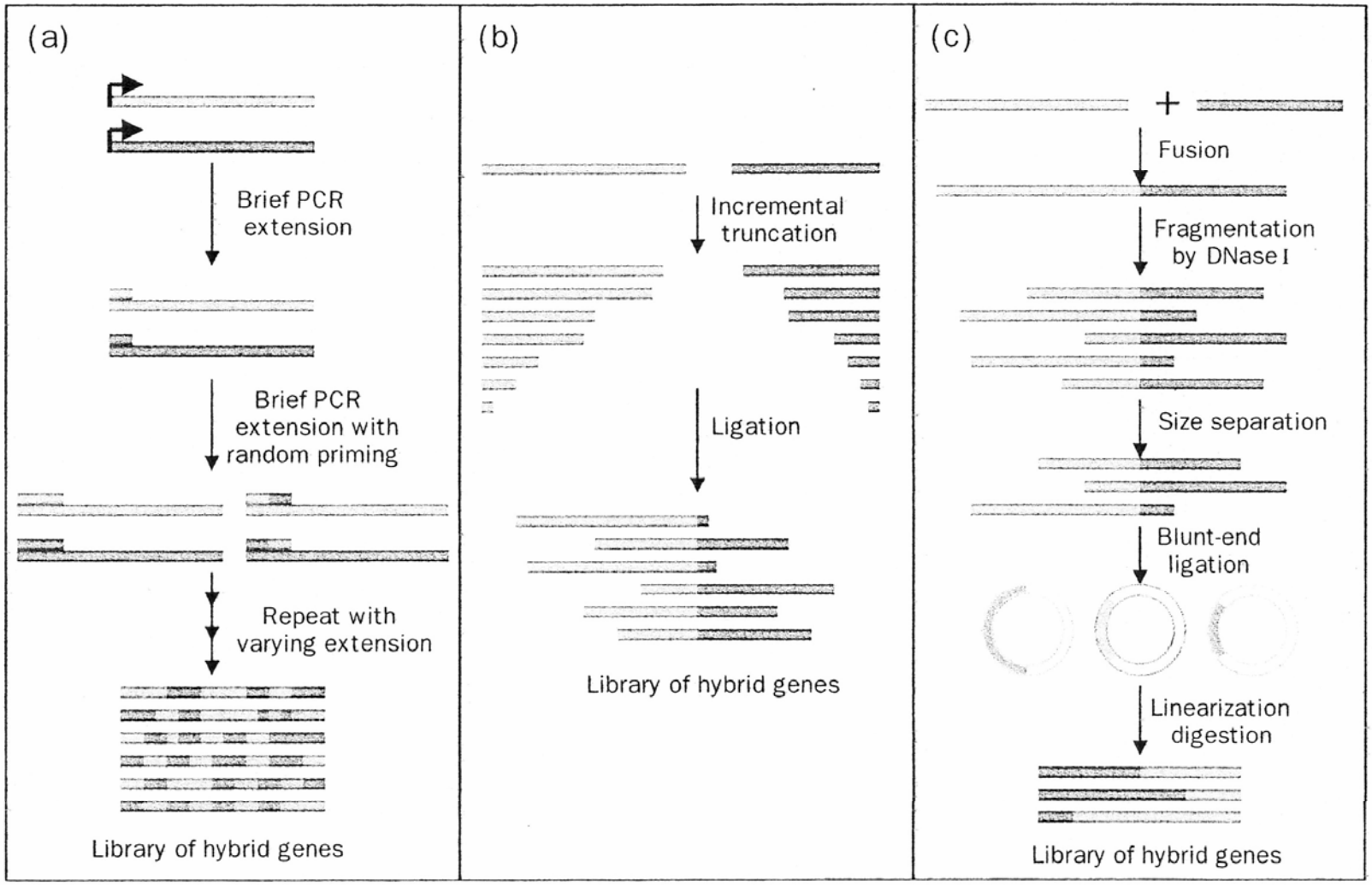


Figure 11 Procedures for the (a) StEP, (b) the ITCHY, and (c) the SHIPREC methods.

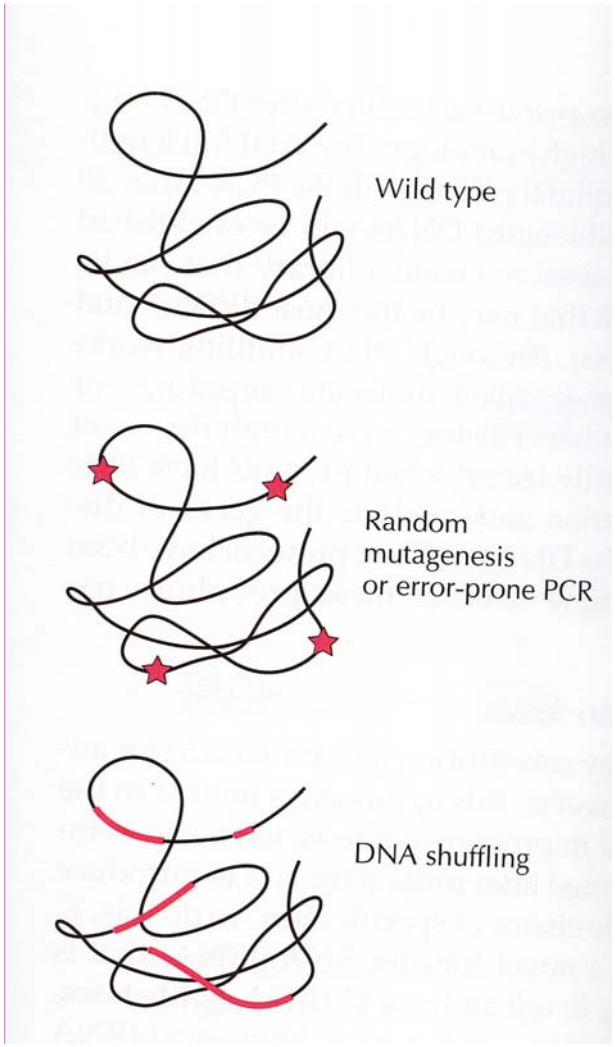
Staggered Extension Process (StEP)

Truncation for the Creation of Hybrid EnzYmes (ITCHY)

Sequence Homology-Independent Protein Recombination (SHIPREC)

Directed Evolution

Difference between non-recombinative and recombinative methods



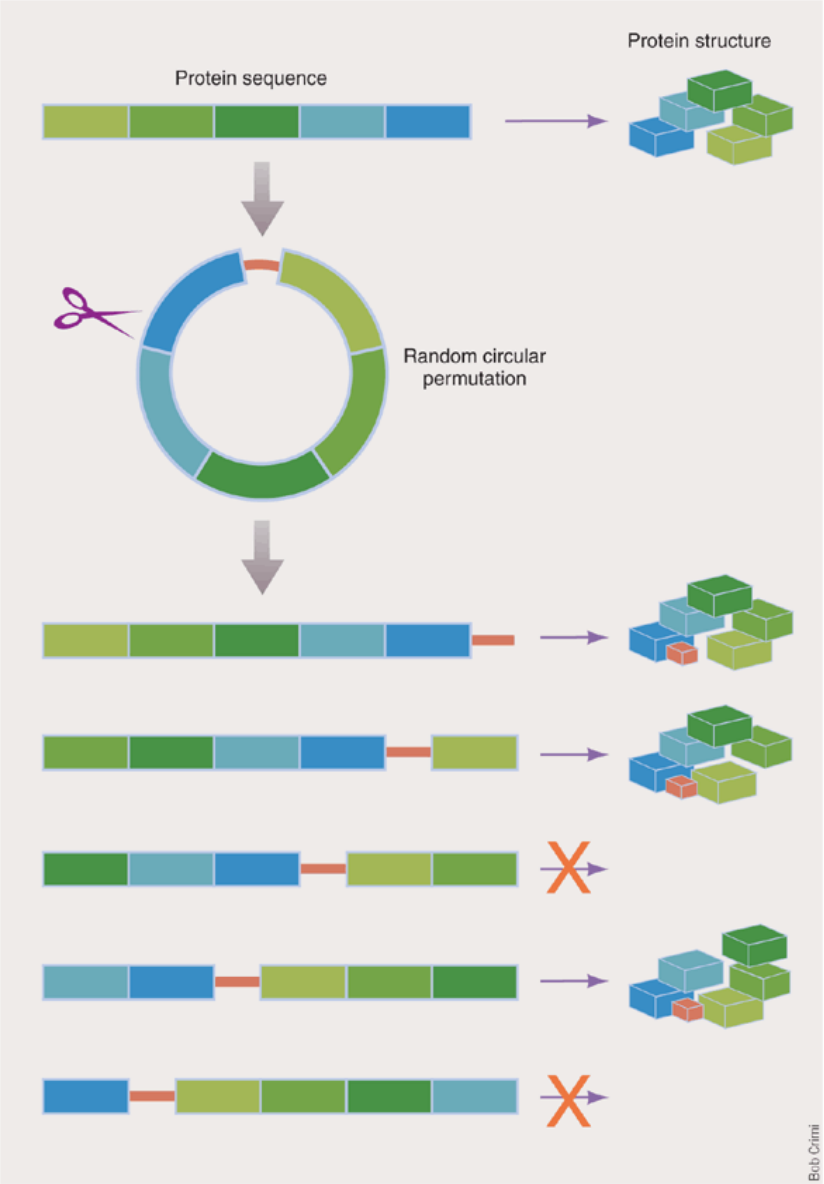
Non-recombinative methods

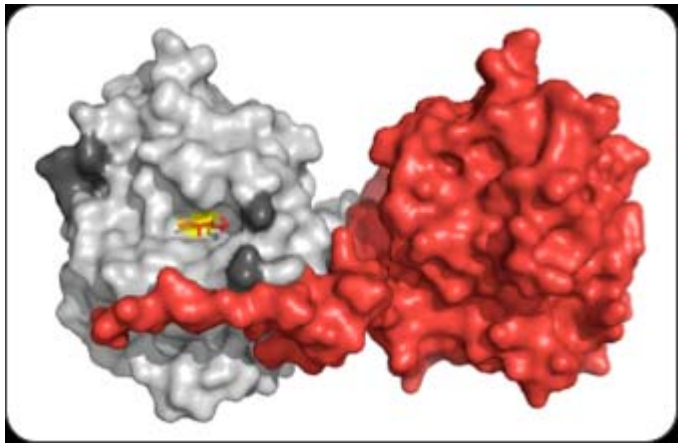
recombinative methods ->
hybrids (chimeric proteins)

...engineering proteins by circular permutation

Protein engineering must not necessarily involve the substitution of amino acids. The reorganization of a proteins' primary sequence can also change the catalytic properties. We are using a technique called circular permutation to explore the effects of termini relocation on catalysis, as well as protein stability & dynamics.

…engineering proteins by circular permutation





lipase B from Candida antarctica, an important biocatalyst in asymmetric synthesis. Upon relocating the protein termini by circular permutation, we observe up to 175-fold enhanced catalytic performance while preserving the enzyme's enantio-selectivity.

Stefan Lutz group

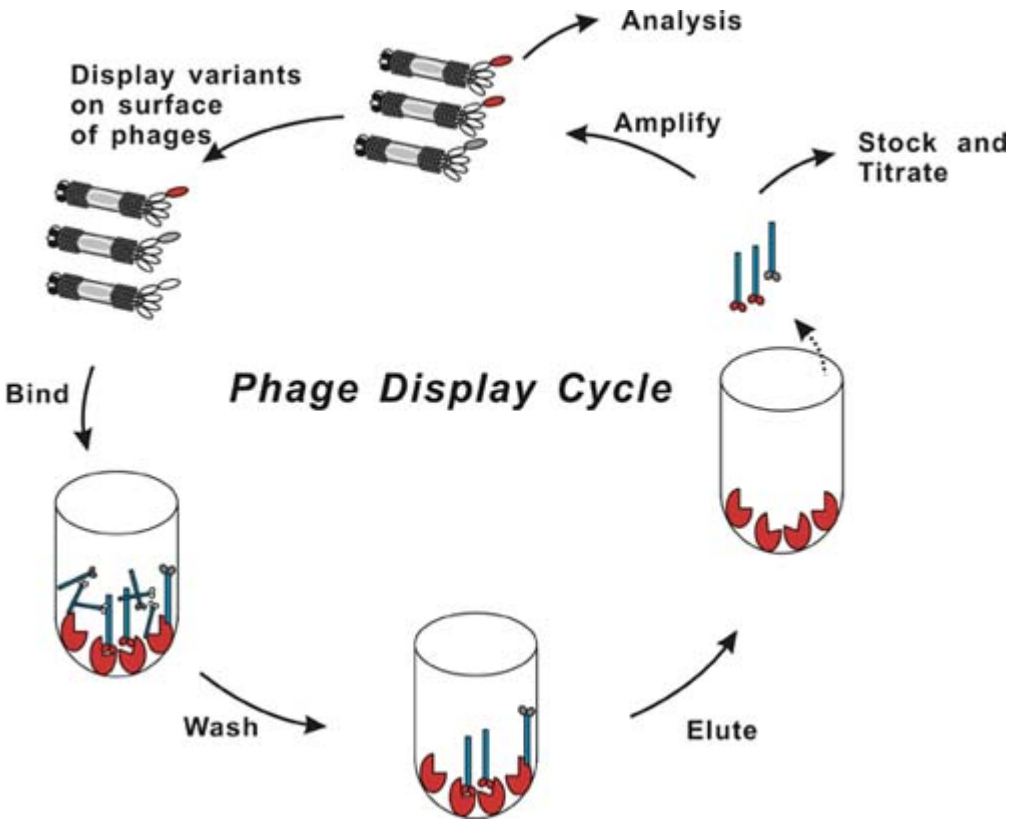
HTS (High throughput screening) System

Screening methods

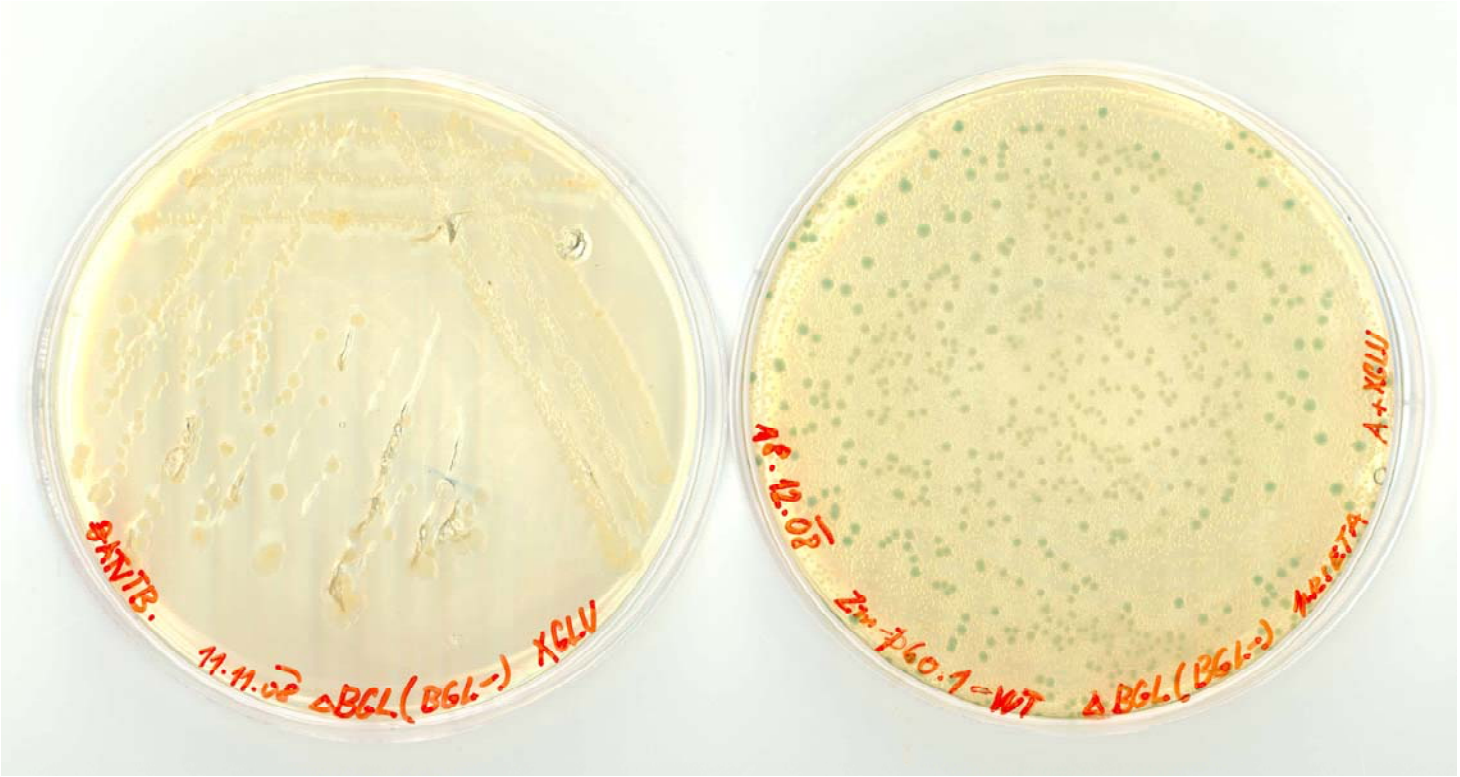
- Genetic selection
  - Growth
  - Survival
- Display technology
  - In-vitro display (cell-free translation)
  - Phage display
  - Cell surface display (Bacteria & Yeast)
- Solid or liquid-phase assay

Screening systems

- FACS (Fluorescence-activated cell sorter)
- Digital image spectroscopy
- Fluorescence detection technique


Analysis
Display variants
on surface
of phages
Amplify
Stock and
Titrate
Bind
Phage Display Cycle
Wash
Elute

Screening in *E.coli*

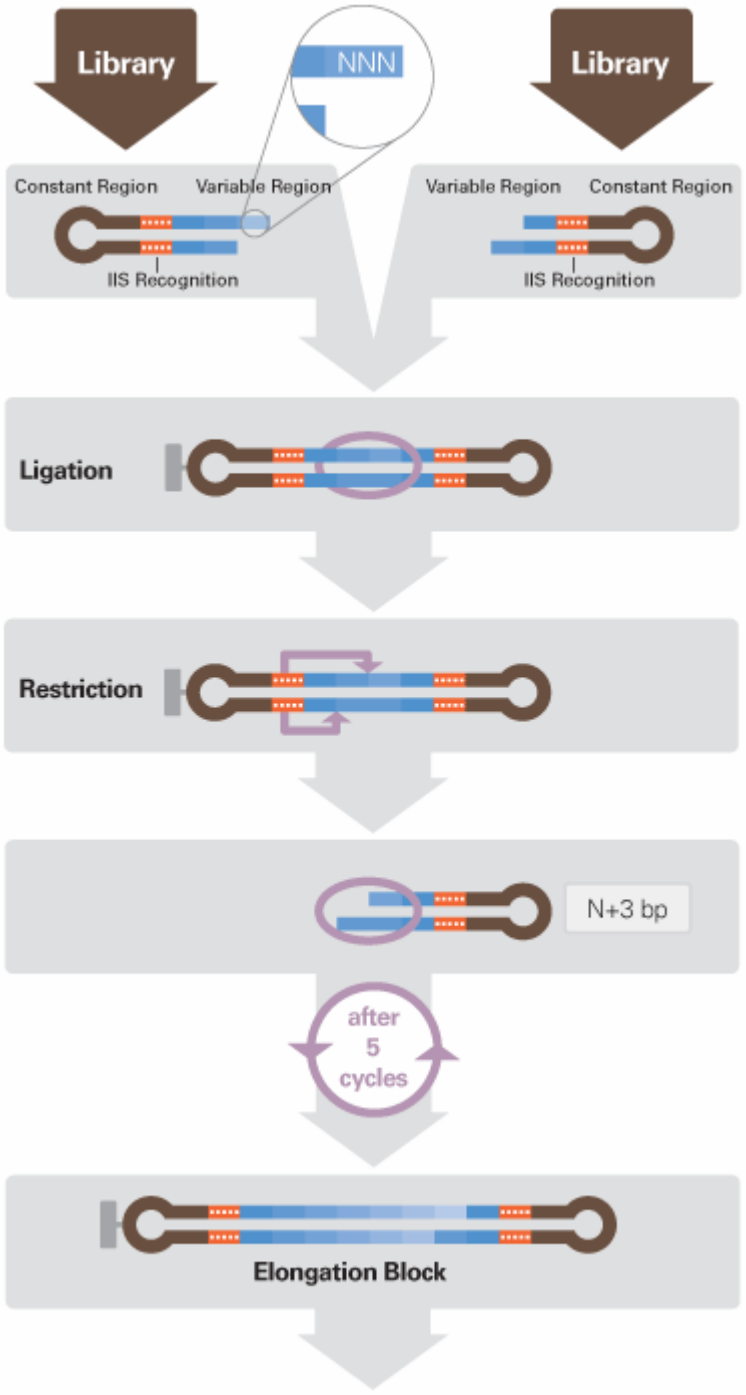


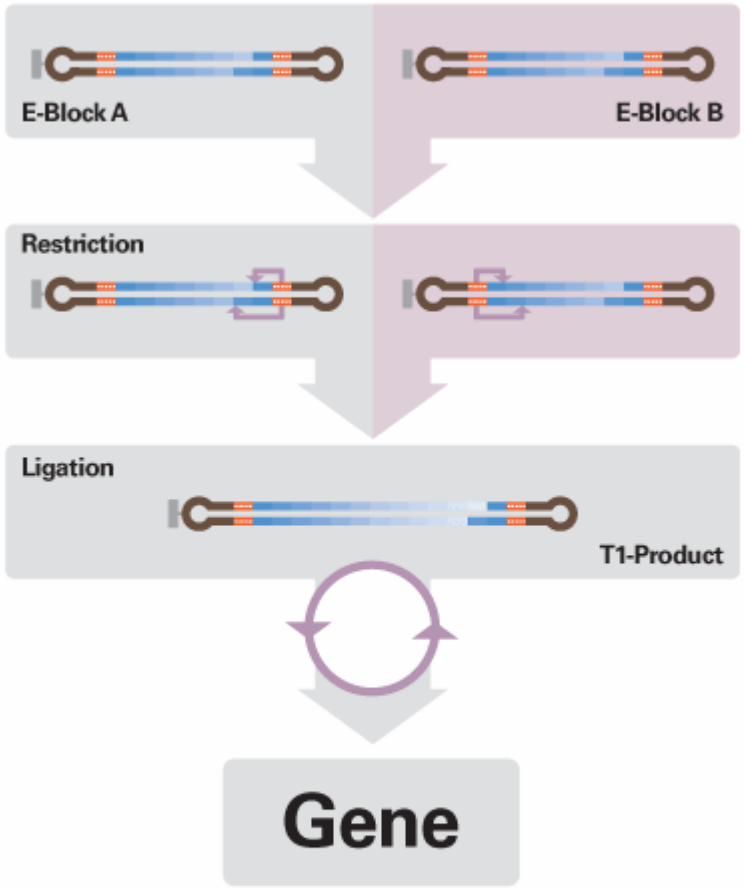
Theoretical implications for mutant library construction (size of the library)

Mutant library parameters calculated for our sample via standard equations and modified by introducing empirical coefficient.
Number of transformants *(L)*=20 and
number of possible sequence variants *(V)*=4.

	Standard	Modified
Library size containing 95% of variants	12	53
Library size with 95% chance of being complete	17.4	77

Slonomax





sloning
BioTechnology

HTS (High throughput screening) System

Screening methods

- Genetic selection
  - Growth
  - Survival
- Display technology
  - In-vitro display (cell-free translation)
  - Phage display
  - Cell surface display (Bacteria & Yeast)
- Solid or liquid-phase assay

Screening systems

- FACS (Fluorescence-activated cell sorter)
- Digital image spectroscopy
- Fluorescence detection technique

H

T

C

AGE-RELATED DISORDERS - MALCOLM A. LEISSRING

HIGH-THROUGHPUT SCREENING

Once novel drug targets have been identified and validated in vivo and appropriate assays have been developed, it becomes possible to screen large collections of compounds in search of drug-like compounds that affect the target in appropriate ways.

Dr. Leissring's lab uses robotics for high-throughput compound screening.

The laboratory of Malcolm A. Leissring, Ph.D., uses these well-characterized animal models to test novel therapeutic approaches to treating and

H

T

C

Patient Care | Health Information | MAYO CLINIC | For Medical Professionals | Research | Education

Request an Appointment | Find a Doctor | Find a Job | Log in to Patient Account | Give to Mayo Clinic

Search | Areas of Research | Faculty | Publications | Postdocs | Grants

AGE- RELATED DISORDERS MALCOLM A. LEISSRING

Hon

Facu

Cell
Stuc

Disc
Ther

High
Scre

In Vi

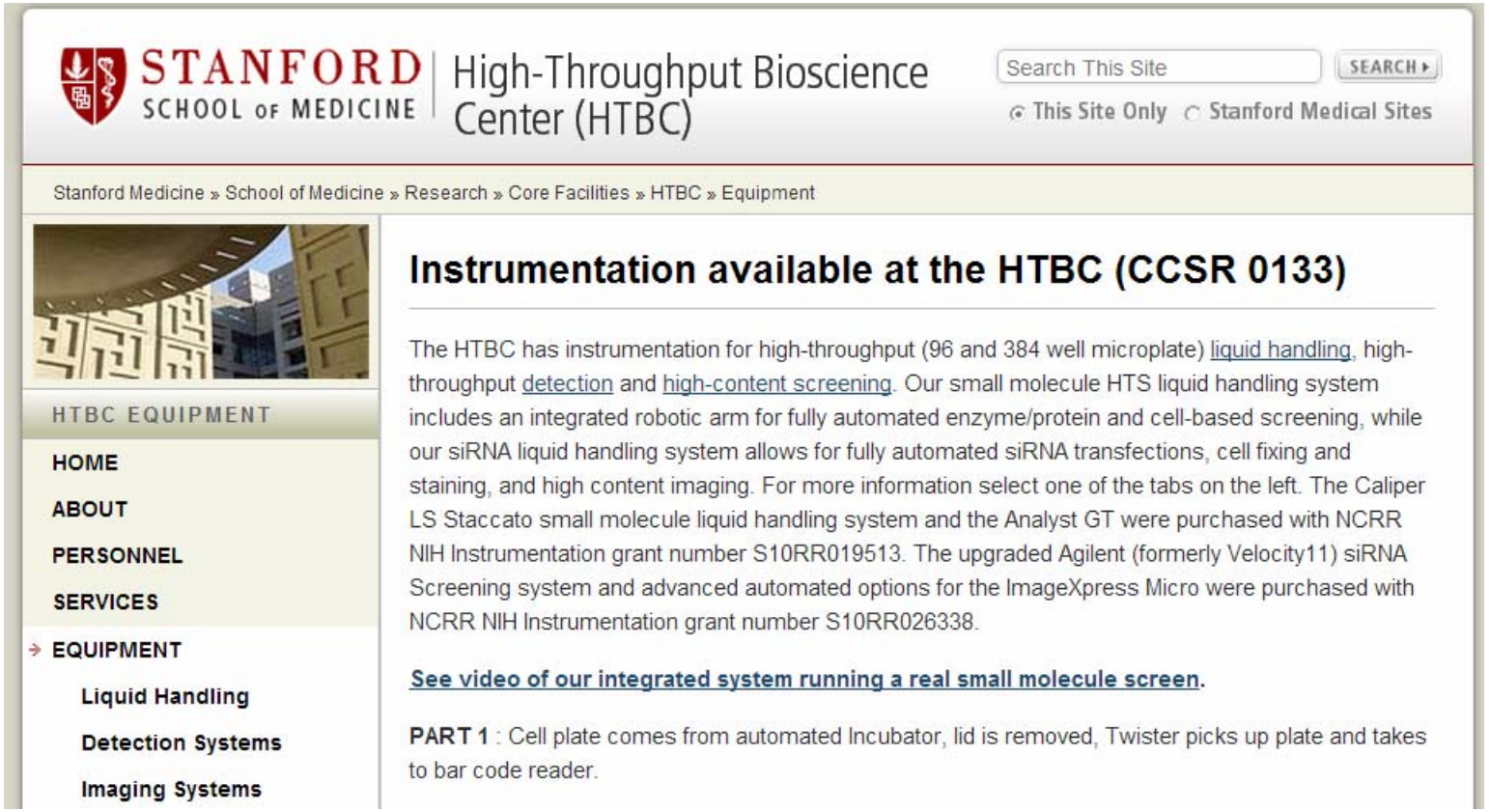
Abo

Con

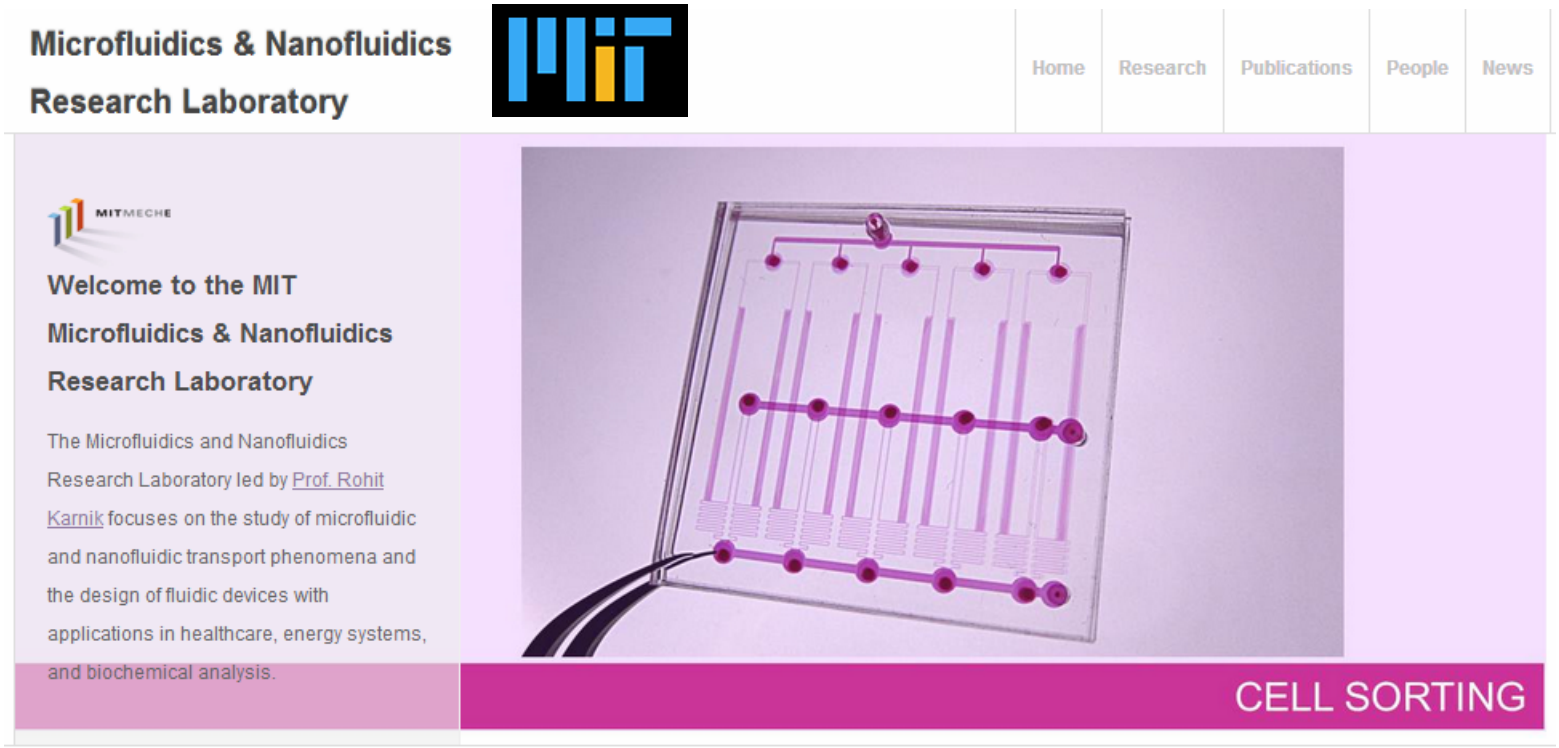
Location	Assay	Format	Library size
Harvard NeuroDiscovery Center	Fluorogenic	384-well, recombinant IDE	Approximately 37,000
Harvard NeuroDiscovery Center	FAβB and fluorescence polarization	384-well, recombinant IDE	Approximately 120,000
Southern Research Institute	FAβB and fluorescence polarization	1,536-well, recombinant IDE	Approximately 100,000
Scripps Florida	FAβB and fluorescence polarization	1,536-well, cell-based	Approximately 325,000 x 2
Genomics Institute of the Novartis Research Foundation	Other	1,536-well, cell-based	Approximately 1,200,000

and

HTS


STANFORD
SCHOOL OF MEDICINE
High-Throughput Bioscience Center (HTBC)
Search This Site
SEARCH
This Site Only
Stanford Medical Sites
Stanford Medicine » School of Medicine » Research » Core Facilities » HTBC » Equipment
HTBC EQUIPMENT
HOME
ABOUT
PERSONNEL
SERVICES
EQUIPMENT
Liquid Handling
Detection Systems
Imaging Systems
Instrumentation available at the HTBC (CCSR 0133)
The HTBC has instrumentation for high-throughput (96 and 384 well microplate) liquid handling, high-throughput detection and high-content screening. Our small molecule HTS liquid handling system includes an integrated robotic arm for fully automated enzyme/protein and cell-based screening, while our siRNA liquid handling system allows for fully automated siRNA transfections, cell fixing and staining, and high content imaging. For more information select one of the tabs on the left. The Caliper LS Staccato small molecule liquid handling system and the Analyst GT were purchased with NCRR NIH Instrumentation grant number S10RR019513. The upgraded Agilent (formerly Velocity11) siRNA Screening system and advanced automated options for the ImageXpress Micro were purchased with NCRR NIH Instrumentation grant number S10RR026338.
See video of our integrated system running a real small molecule screen.
PART 1 : Cell plate comes from automated Incubator, lid is removed, Twister picks up plate and takes to bar code reader.

Microfluidics for HTS



Microfluidics for HTS

Todd Thorsen lab

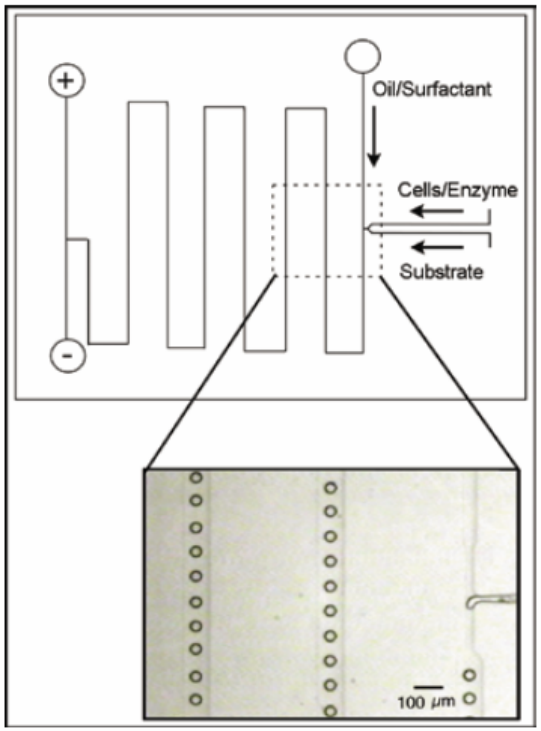


Figure 1. Droplet-generating biochemical screening chip. Monodispersed water droplets maintain uniform spacing and periodicity as they flow towards the device outlet.

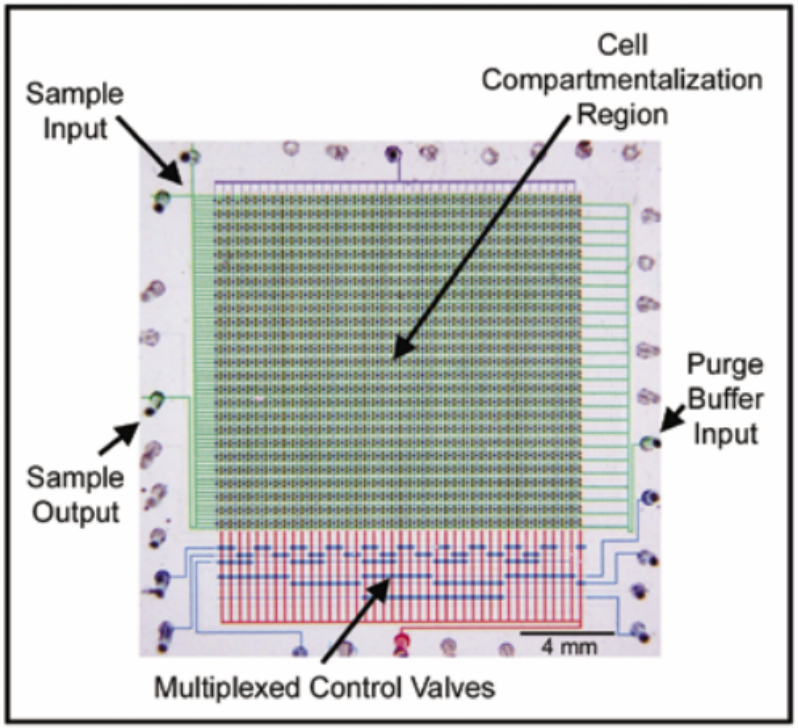
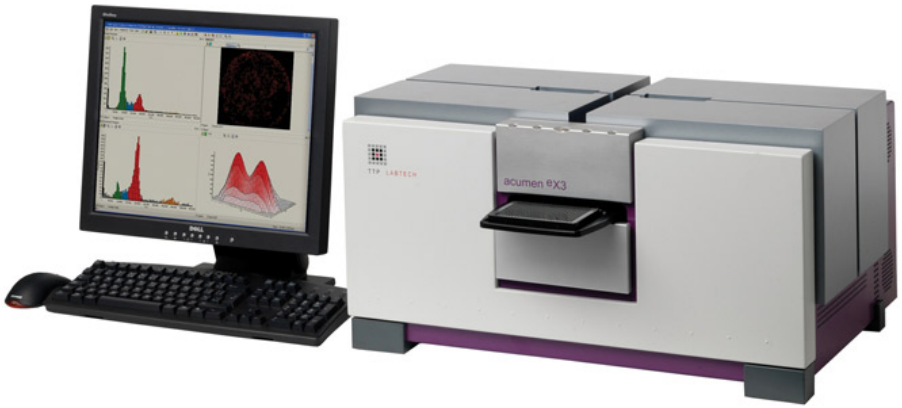


Figure 2. Silicone microfluidic array chip for high-throughput single cell assays. Chip contains 1000 independent compartments and 3574 elastomeric valves.

High Content Analysis and High Content Screening

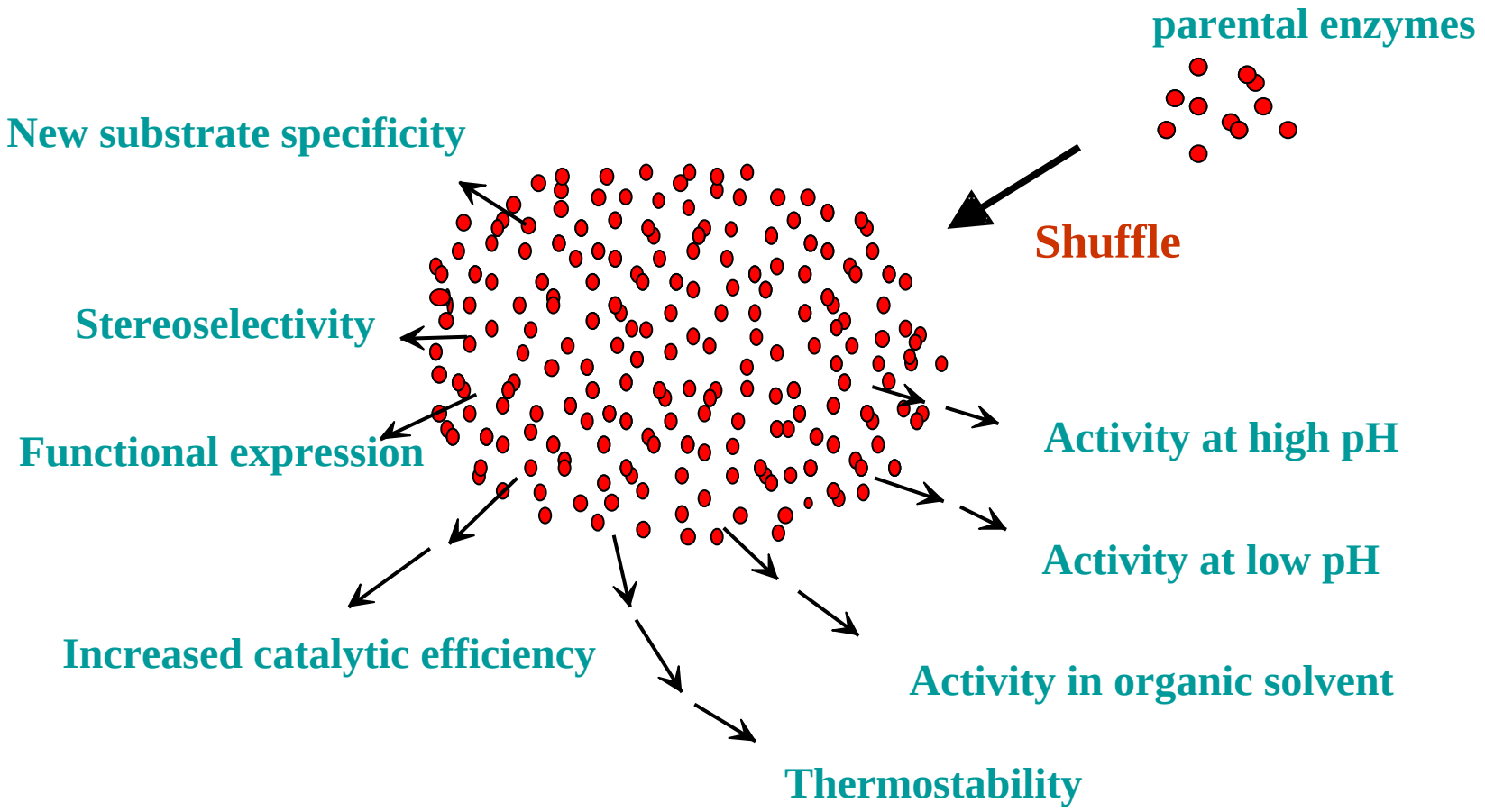
High Content Analysis (HCA) and High Content Screening (HCS) are imaging based multi-parametric approaches of cell analysis at the single-cell level. Originally developed as a complementary technology to traditional biochemical high-throughput screening (HTS) in drug discovery, today High Content Screening is established in a far broader area of the life science space as an unbiased method of imaging multiple cellular samples.



Miami Project High Content Screening Core

The core technology in the HCS uses the Cellomics VTI ArrayScan with a Thermo Fisher robot to feed the scanner.
This system can scan about thirty 96 well plates over night, acquiring 9 fields per well, with 3-4 colors. This is a lot of data.

Functional Diversity



Protein Engineering - Applications

Engineering Stability of Enzymes – T4 lysozyme

-> S-S bonds introduction

Native protein

Engineered protein

Table 8.2 Properties of T4 lysozyme and six engineered variants

Enzyme	Amino acid at position:							No. of -S-S-	% Activity	T_m (°C)
	3	9	21	54	97	142	164			
wt	Ile	Ile	Thr	Cys	Cys	Thr	Leu	0	100	41.9
pwt	Ile	Ile	Thr	Thr	Ala	Thr	Leu	0	100	41.9
A	Cys	Ile	Thr	Thr	Cys	Thr	Leu	1	96	46.7
B	Ile	Cys	Thr	Thr	Ala	Thr	Cys	1	106	48.3
C	Ile	Ile	Cys	Thr	Ala	Cys	Leu	1	0	52.9
D	Cys	Cys	Thr	Thr	Cys	Thr	Cys	2	95	57.6
E	Ile	Cys	Cys	Thr	Ala	Cys	Cys	2	0	58.9
F	Cys	Cys	Cys	Thr	Cys	Cys	Cys	3	0	65.5

Adapted from Matsumura et al., *Nature* **342**:291–293, 1989.

wt, wild-type T4 lysozyme; pwt, pseudo-wild-type enzyme; A through F, six engineered cysteine variants; -S-S-, disulfide bonds; T_m, "melting" temperature (a measure of thermostability).

Protein Engineering - Applications

Engineering Stability of Enzymes – triosephosphate isomerase from yeast

-> replace Asn (deaminated at high temperature)

Table 8.3 Stability at 100°C of the yeast enzyme triosephosphate isomerase and its engineered derivatives

Enzyme	Amino acid at position:		Half-life (min)
	14	78	
Wild type	Asn	Asn	13
Variant A	Asn	Thr	17
Variant B	Asn	Ile	16
Variant C	Thr	Ile	25
Variant D	Asp	Asn	11

Adapted from Ahern et al., *Proc. Natl. Acad. Sci. USA* **84**:675–679, 1987.

Enzyme stability is expressed as the half-life, or rate of enzyme inactivation, at 100°C. A longer half-life indicates a more stable enzyme.

Protein Engineering - Applications

Engineering Activity of Enzymes – tyrosyl-tRNA synthetase from *B. stearothermophilus*

-> replace Thr 51 (improve affinity for ATP) -> Design

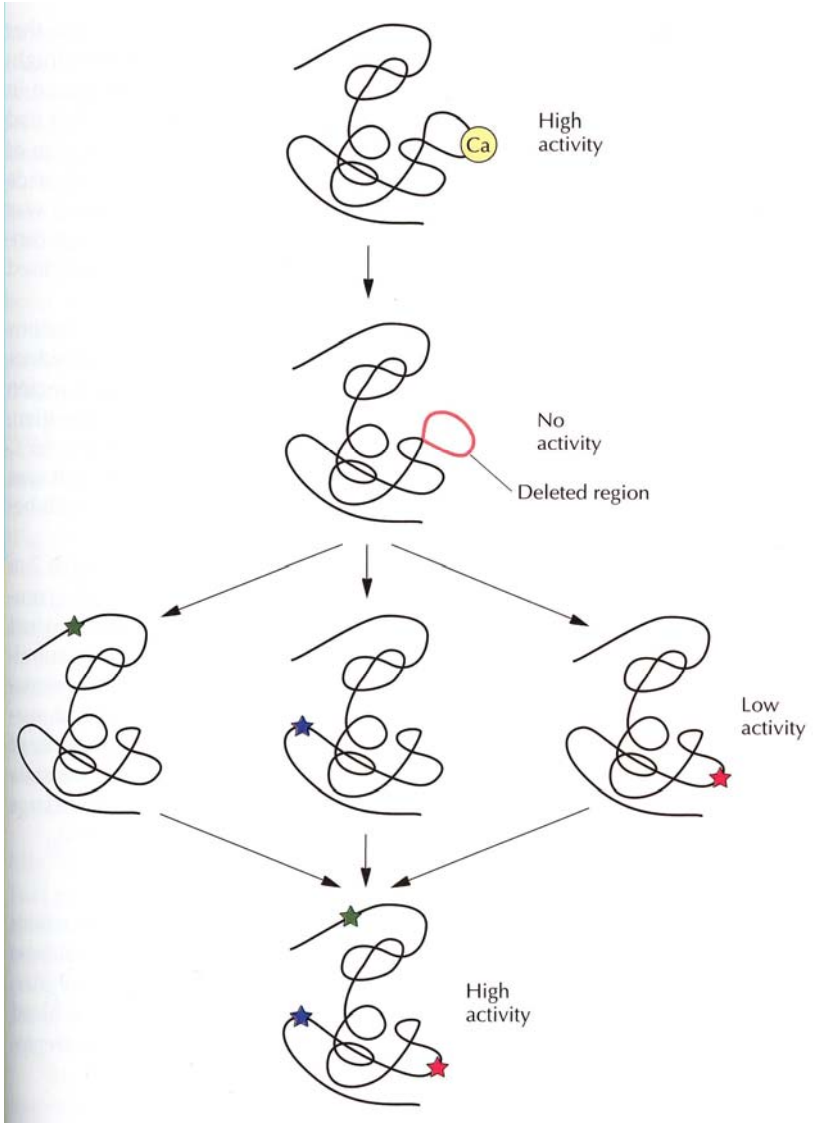
Table 8.4 Aminoacylation activity of native (Thr-51) and modified (Ala-51 and Pro-51) tyrosyl-tRNA synthetases

Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} M^{-1})
Thr-51	4.7	2.5	1,860
Ala-51	4.0	1.2	3,200
Pro-51	1.8	0.019	95,800

Adapted from Wilkinson et al., *Nature* **307**:187–188, 1984.

The units for K_m, the binding constant of the enzyme for ATP, are millimolar units (mM); the units for k_{cat}, the catalytic rate constant, are reciprocal seconds (s^{-1}); and the units for k_{cat}/K_m, the catalytic efficiency, are s^{-1} M^{-1}.

Protein Engineering - Applications



Engineering Ca-independency of subtilisin

Saturation mutagenesis -> 7 out of 10 regions were found to give increase of stability

Mutant:

10x more stable than native enzyme in absence of Ca

50% more stable than native in presence of Ca

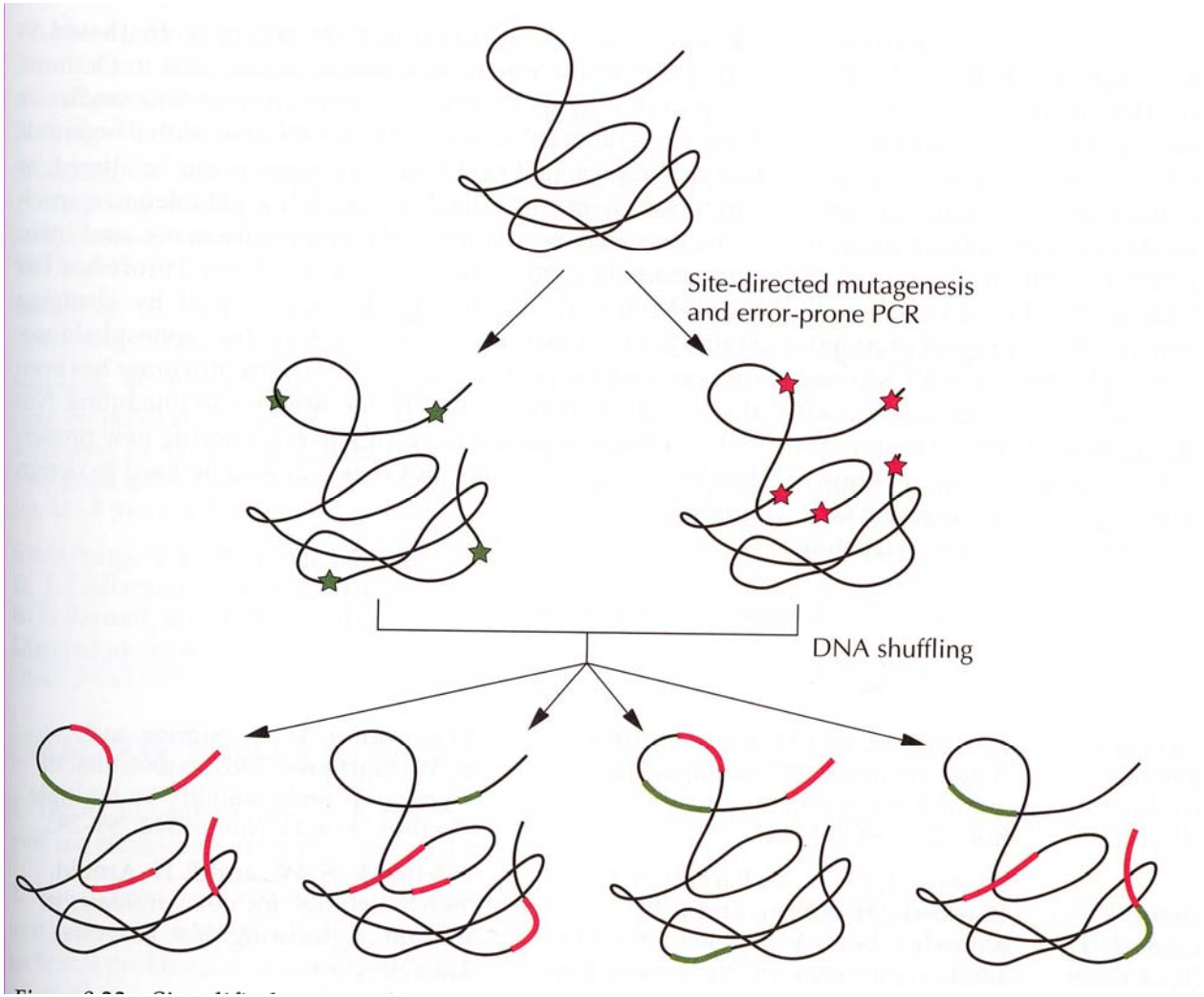
Protein Engineering - Applications

Site-directed mutagenesis -> used to alter a single property

Problem : changing one property -> disrupts another characteristics

Directed Evolution (Molecular breeding) -> alteration of multiple properties

Protein Engineering – Applications
Directed Evolution



Protein Engineering - Applications
Directed Evolution

Table 1. Examples of the diversity of properties that have been improved by evolutionary engineering

Protein	Altered function	Reference
Barley α-amylase	Thermostability: ten-fold increase of half-life at 90°C	JOYET et al. 1992
Subtilisin	Alkaline stability: doubling the autolytic half-time at pH 12	CUNNINGHAM and WELLS (1987)
Subtilisin	Tolerates loss of stabilizing divalent cations	STRAUSBERG et al. (1995)
Subtilisin E	Active in 60% DMF	YOU and ARNOLD (1996)
Streptomyces griseus protease B	Broadened substrate specificity	SIDHU and BORGFORD (1996)
Green fluorescent protein	40-fold brighter fluorescing bacterial colonies	CRAMERI et al. (1996)
Immunoglobulin constant domain	Preferential formation of heterodimers	ATWELL et al. (1997)
Immunoglobulin variable domain	Tolerates loss of structural disulfide bridge	MARTINEAU et al. (1998)

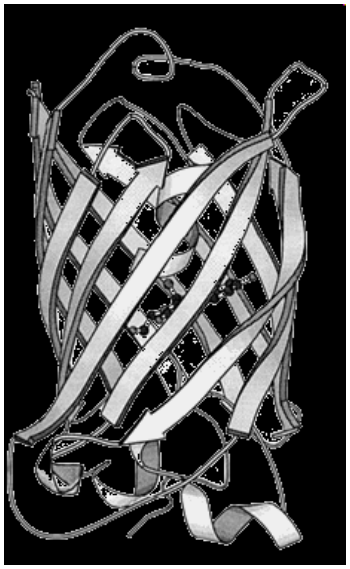
Protein Engineering - Applications Directed Evolution

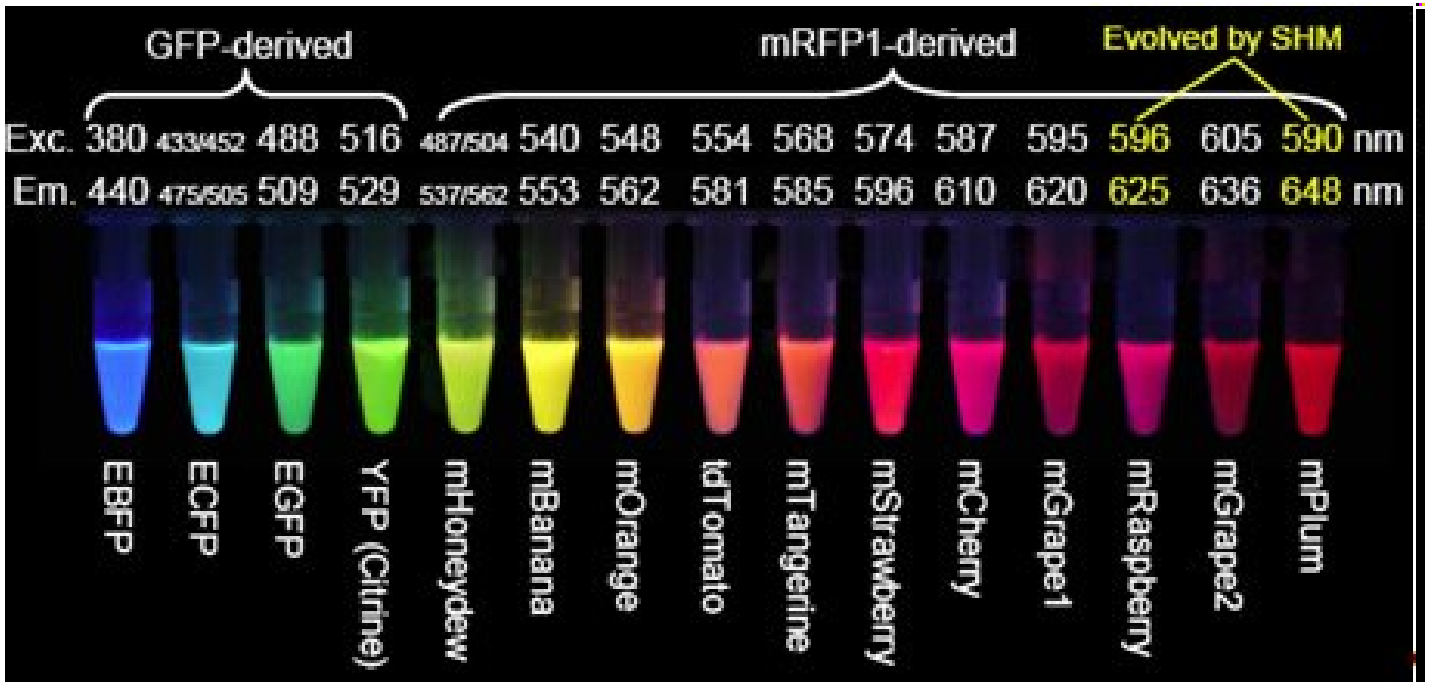
Table 2. Representative experiments using successive cycles of variation and selection

Protein	Property	Number of cycles for success	Number of nucleotide changes required	Number of amino acid changes required	Reference
β-Lactamase	Increased activity	3	4	4	STEMMER (1994b)
GFP	Improved folding and expression	3	3	3	CRAMERI et al. (1996)
Subtilisin E	Stability in aqueous DMF	2	3	3	YOU and ARNOLD (1996)
Arsenite membrane pump	Increased activity	3	3	3	CRAMERI et al. (1997)
FLP-recombinase	Thermostability	8	3–4	3–4	BUCHHOLZ et al. (1998)

In almost all cases a single nucleotide change leading to a single amino acid change was sufficient per cycle, the number of silent mutations was approximately the same. No amino acid change was reported that would have required more than one nucleotide change. Thus current protocols appear to sample sequence space in a biased fashion, in single mutation steps.

GFP




GFP-derived
mRFP1-derived
Evolved by SHM
Exc. 380 433/452 488 516 487/504 540 548 554 568 574 587 595 596 605 590 nm
Em. 440 475/505 509 529 537/562 553 562 581 585 596 610 620 625 636 648 nm
EBFP
ECFP
EGFP
YFP (Citrine)
mHoneydew
mBanana
mOrange
tdTomato
mTangerine
mStrawberry
mCherry
mGrape1
mRaspberry
mGrape2
mPlum

The Nobel Prize in Chemistry 2008

Osamu Shimomura, Martin Chalfie, Roger Y. Tsien

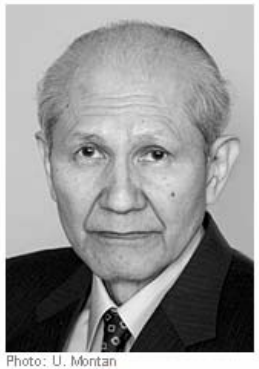


Photo: U. Montan

Osamu Shimomura

Photo: U. Montan

Martin Chalfie

Photo: U. Montan

Roger Y. Tsien

The Nobel Prize in Chemistry 2008 was awarded jointly to Osamu Shimomura, Martin Chalfie and Roger Y. Tsien *"for the discovery and development of the green fluorescent protein, GFP"*.

CEITEC

Central European Institute of Technology

BRNO | CZECH REPUBLIC

Protein Engineering

Pavel Mazura

CEITEC – MENDELU

Department of molecular biology and radiobiology

Scaled Biosystems team

mazura@sci.muni.cz

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

2007–13
OP Research and Development for Innovation