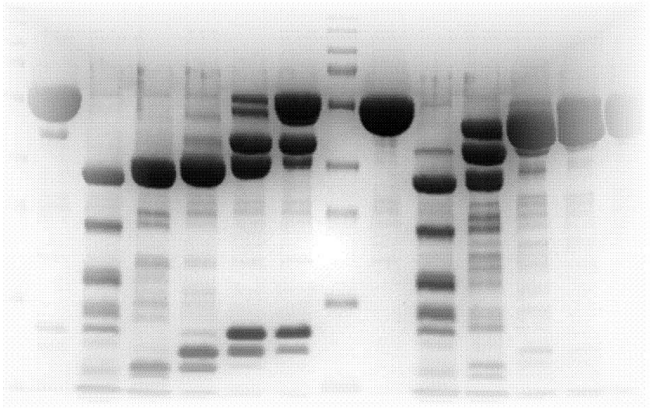
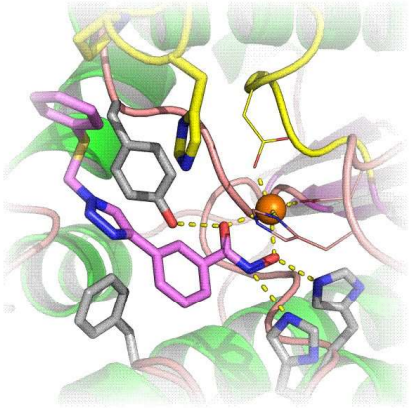
LOSCHMIDT LABORATORIES

Recombinant proteins:
from small- to large-scale expression technologies

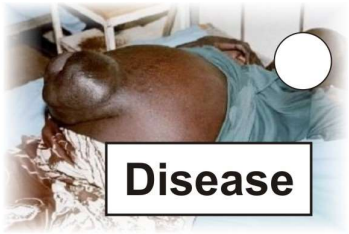
Martin Marek
Andrea Schenkmayerová



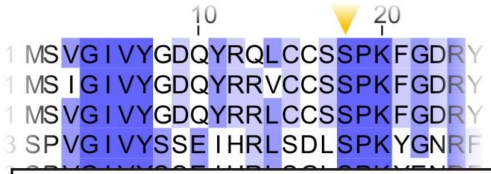


Molecular Biotechnology Practise
24 October (2018), MUNI, Brno (CZE)

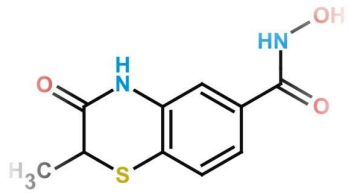
Key role of recombinant proteins in drug discovery



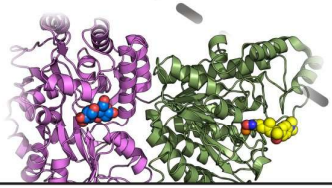
Disease



Identify target
(Genomics, biological assays)



J1038



Isolate protein involved in disease
(Molecular cloning, recombinant expression, biochemical and structural analyses)

Find a drug effective against disease protein
(virtual screening, HTS and library screening etc.)

Human clinical trials

Preclinical testing
(Animal studies, scale-up, formulation)

Protein production is a significant bottleneck in early phase drug discovery

FDA approval

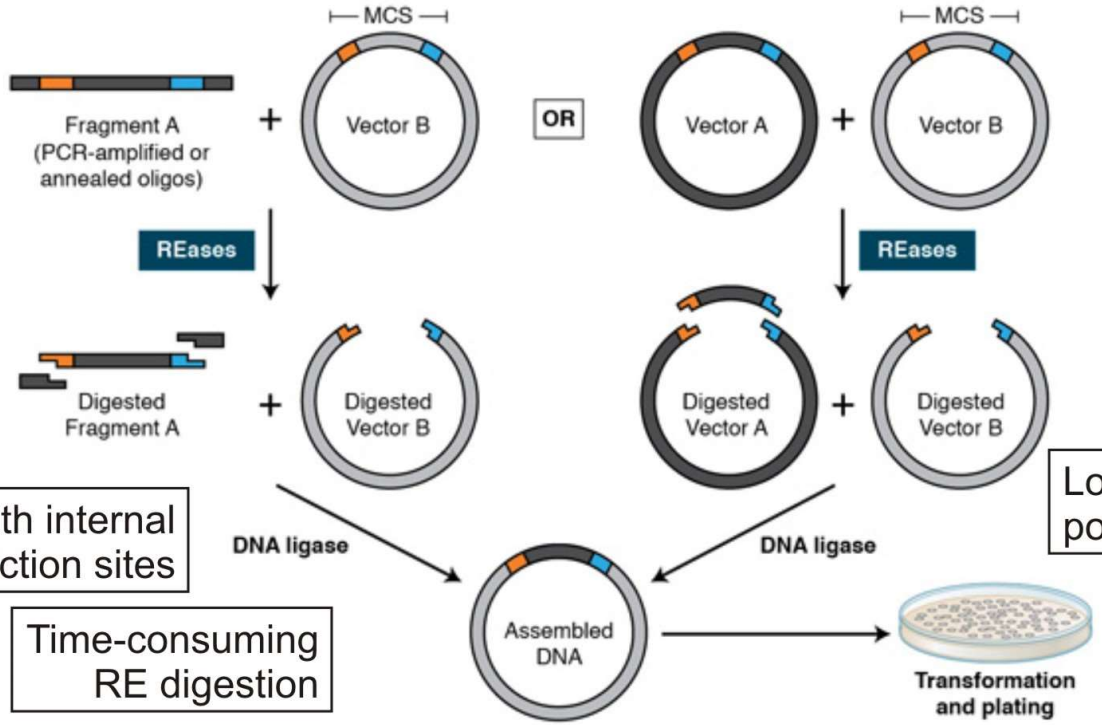
Recombinant protein production workflow

- Molecular cloning
- Protein expression
- Protein purification
- Protein characterization
- (Protein structure determination)

Molecular cloning: classical vs. emerging technologies

- Obtain/generate a DNA template of the GOI (PCR)
- Ligate the GOI into an appropriate expression vector
- Transform the expression vector into a bacterial strain
- Analyse bacterial clones to confirm integration of the GOI
- Sequence verification

Classical cloning

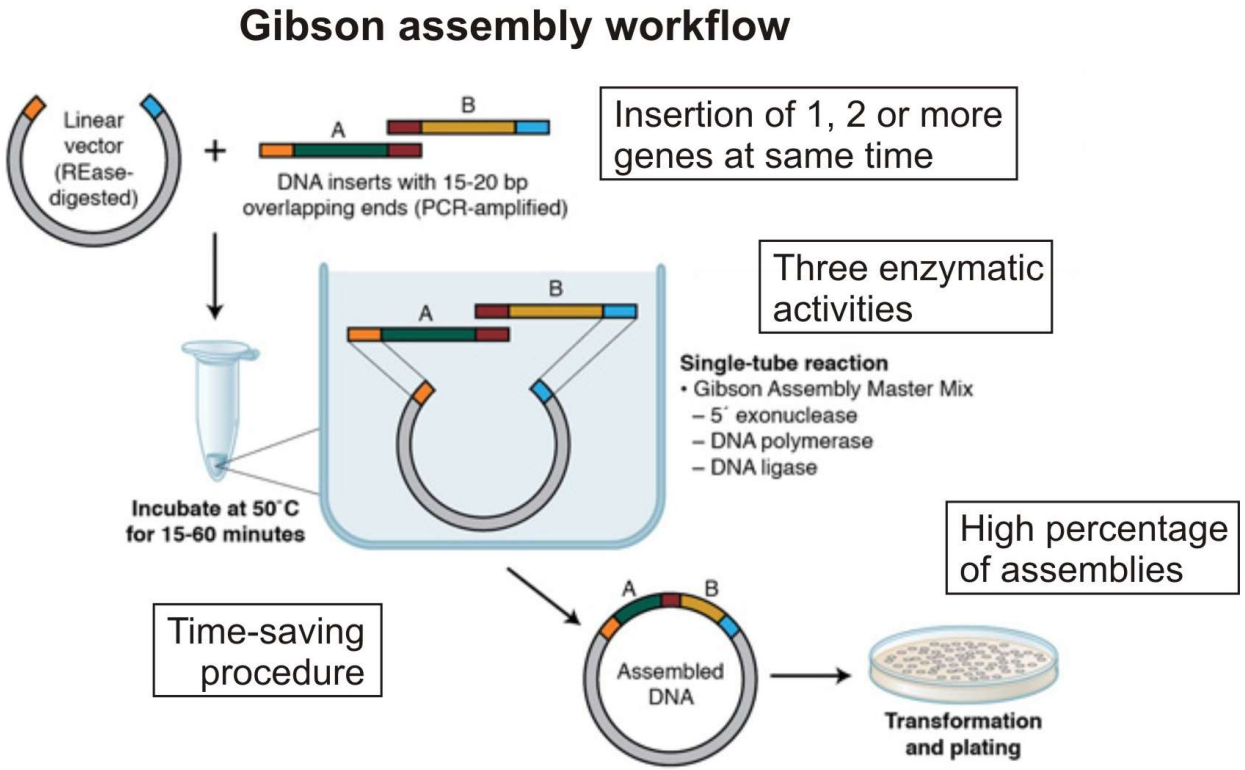


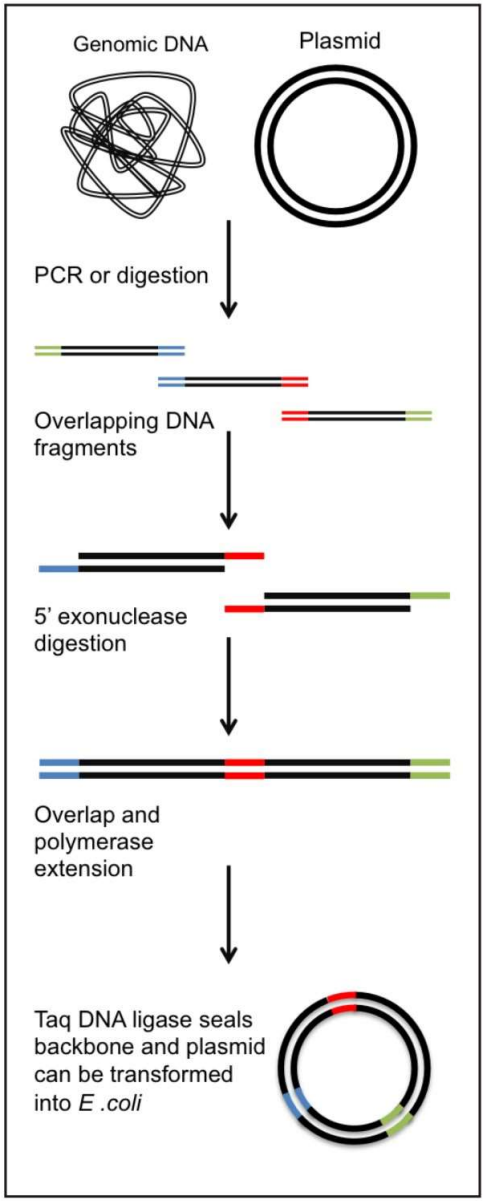
Problem with internal restriction sites

Time-consuming RE digestion

Low number of positive colonies

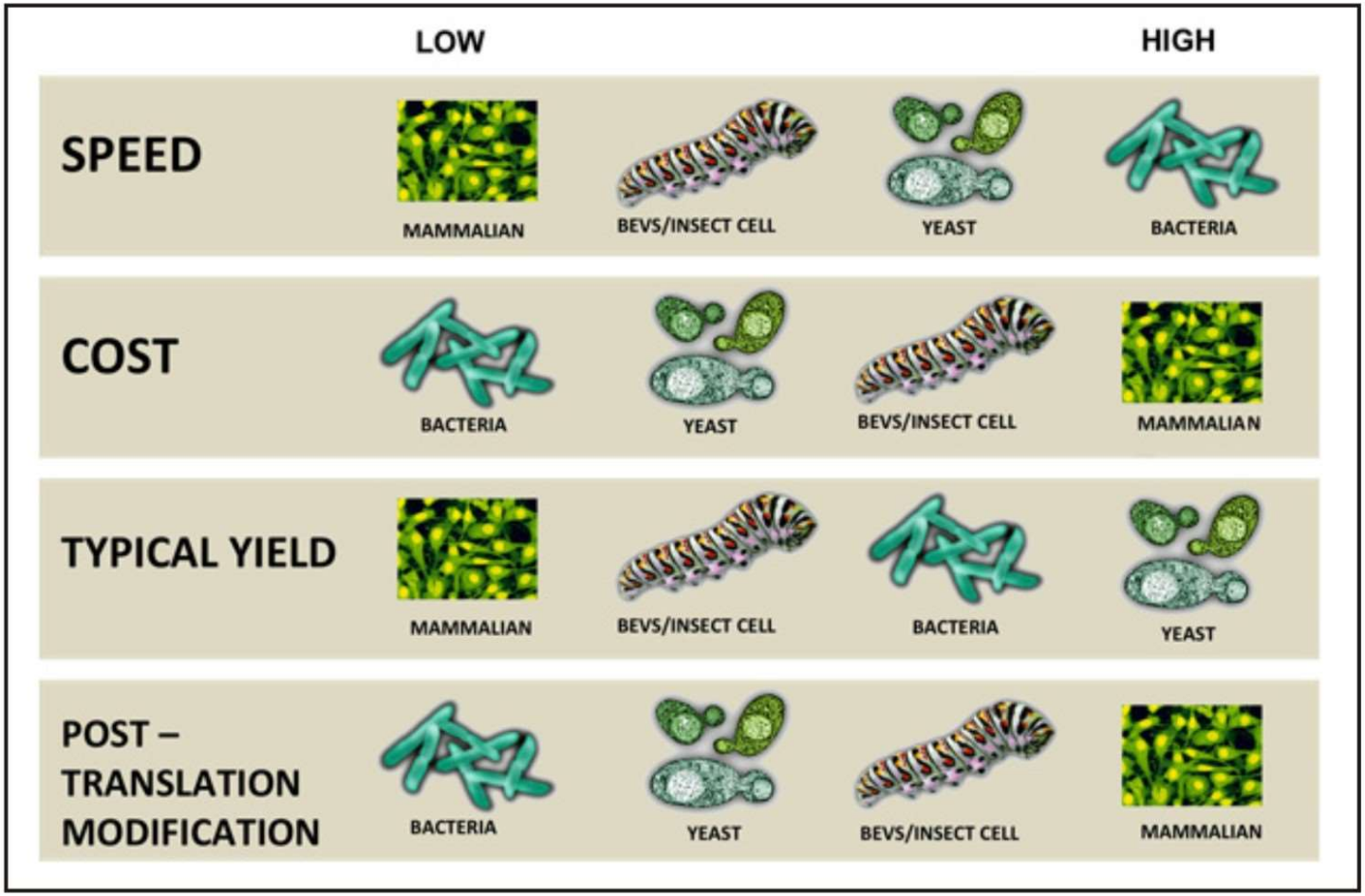
Molecular cloning: emerging techniques
Gibson assembly reaction





Recombinant protein expression: selection of a host

- *Escherichia coli*
- Yeasts (*Pichia*, *Kluyveromyces*)
- Insect cells (Baculovirus expression system)
- Mammalian cells (HEK293)



Recombinant protein expression in *E. coli*

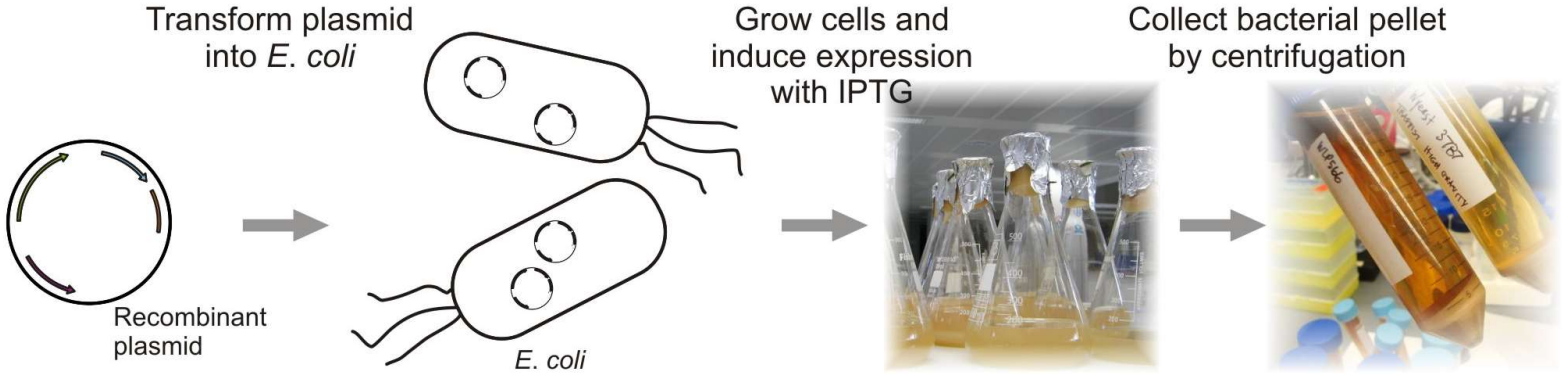
ADVANTAGES

- Inexpensive setup and running costs
- High recombinant protein production levels
- Short timeline from cloning to protein recovery (1 week)
- Limited technical knowledge required for culturing
- Scalability from small (2 mL) to very large culture (>10,000 L) volumes

DISADVANTAGES

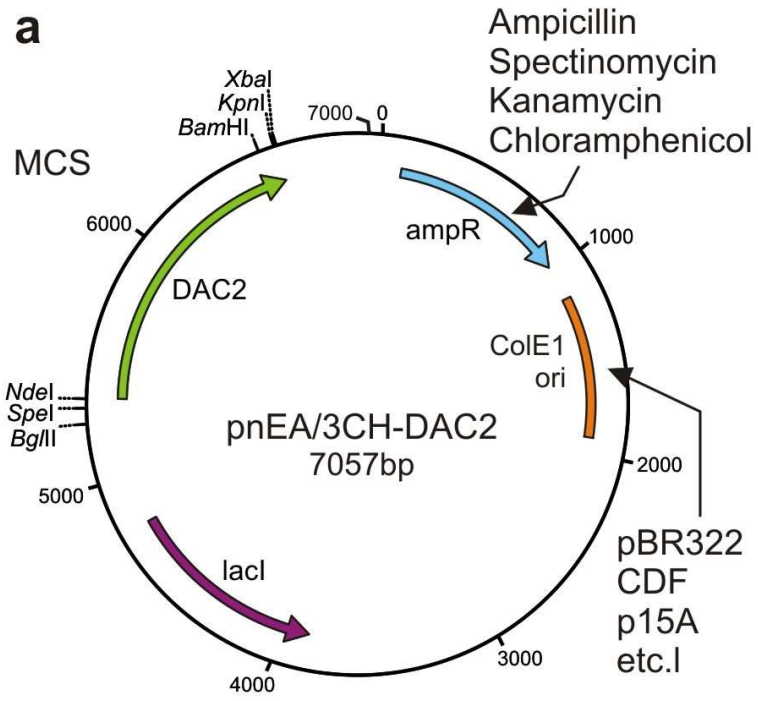
- Inability to perform post-translation modifications (PTMs)
- Limited formation of disulphide bond

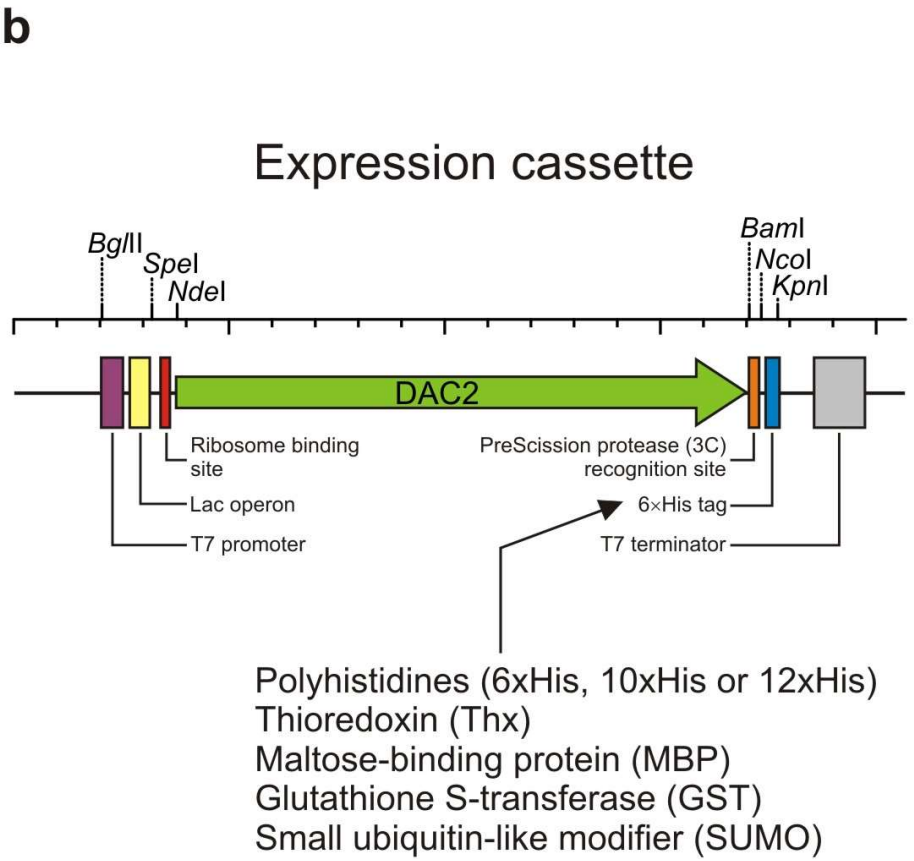
Workflow of protein expression in *E. coli*



Protein expression in *E. coli*: plasmid backbone

- pET
- pGEX





Wide range of *E. coli* strains

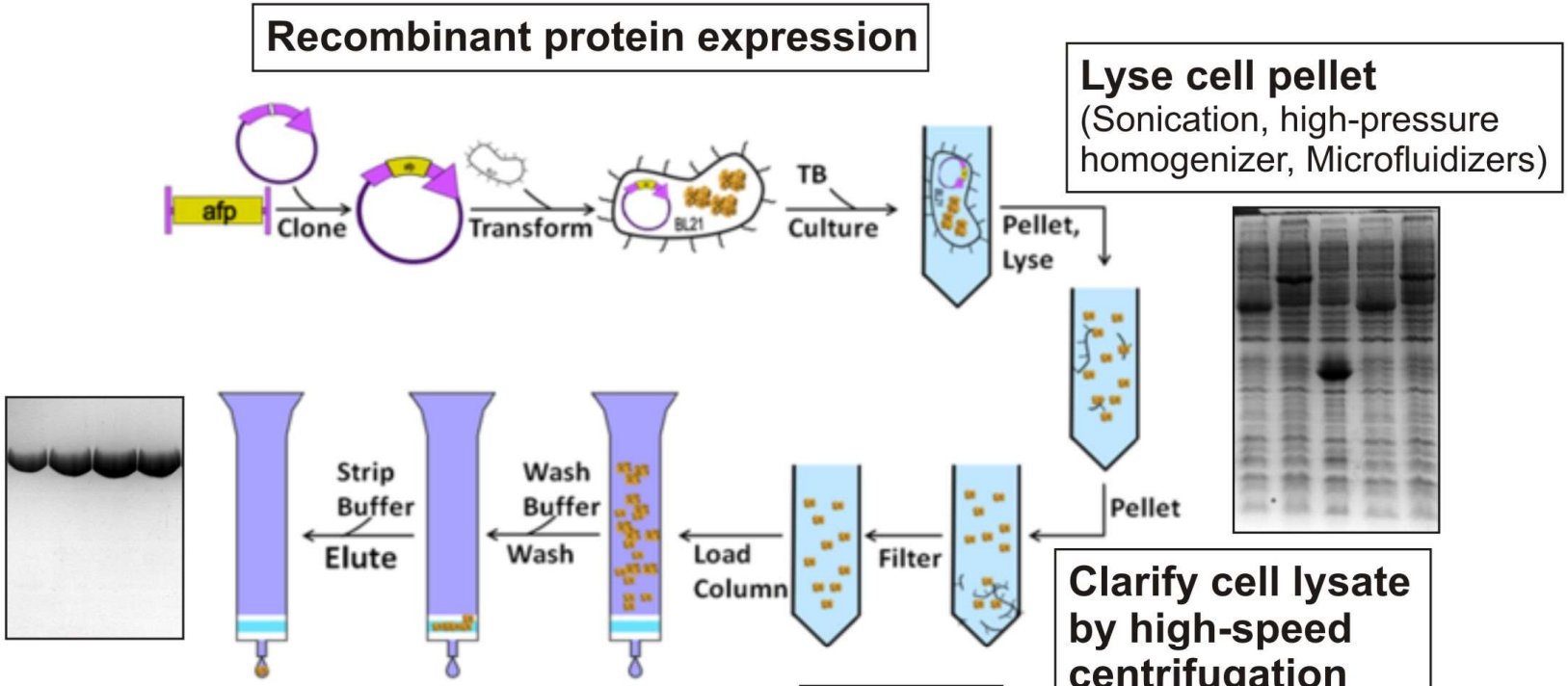
Strain	Features
BL21 (DE3)	Most common host strain, enables high-level recombinant protein expression
BL21 (DE3) pLysS	Enables high-level expression and suppression of T7 RNA Polymerase basal level expression
BL21-CodonPlus	Improved expression of genes with codons rarely used in bacteria
Rosetta	Improved expression of genes with codons rarely used in bacteria
Arctic Express	Contains a plasmid encoding chaperonins to aid in folding and allow expression at low temperature (12°C)
Shuffle T7 Express	Expresses DsbC to enable cytoplasmic disulphide bond formation

Recombinant expression in *E. coli:* tips and tricks

- **Growth temperature:** Typically 37°C, but lowering (25°C or 15°C) improves folding/solubility
- **Expression with a fusion tag:** polyhistidenes (6xHis, 10xHis or 12xHis)
  Thioredoxin (Thx)
  Maltose-binding protein (MBP)
  Glutathione S-transferase (GST)
  Small ubiquitin-like modifier (SUMO)
- **Co-expression:** Chaperonins (folding) and foldases (disulphide isomerase, DsbC)
  Co-expression with interaction partners (co-solubility effects)
- **Media:** LB Broth, 2xLB Broth (simple and cheap), Terrific Broth (TB), 2YT
- **Antibiotics:** Selection of recombinant clones and prevention of contamination
  Concentration of antibiotics in large-scale expression is decreased
- **Codon optimization:** Modifying codons in a gene sequence to match the codon usage bias of the host cell used for expression

Recombinant protein purification: workflow

The aim of a purification procedure is to obtain a highly pure and stable protein at an appropriate concentration in a buffer compatible with the intended application.



Chromatography columns in protein purification

Chromatography is the most powerful and commonly used means of purifying recombinant proteins. Each technique separates proteins based on different properties, so it is often advantageous to combine several types to maximise separation of the recombinant protein from host cell proteins.

Technique	Stage	Description
Affinity Chromatography (AC)	Capture or Intermediate	Based on a reversible interaction between the protein/affinity tag and a specific ligand
Ion Exchange Chromatography (IEX)	Capture or Intermediate	Separates proteins based on their net surface charge
Hydrophobic Interaction Chromatography (HIC)	Intermediate	Binding under high salt conditions, generally performed following an ammonium sulphate precipitation step
Size Exclusion Chromatography (SEC)	Polishing	Separates proteins based on their hydrodynamic volume (size)
Reverse Phase Chromatography (RPC)	-	High-resolution chromatography based on weak hydrophobic interactions. Harsh conditions generally only suitable for purification of peptides

Affinity chromatography: fusion tags

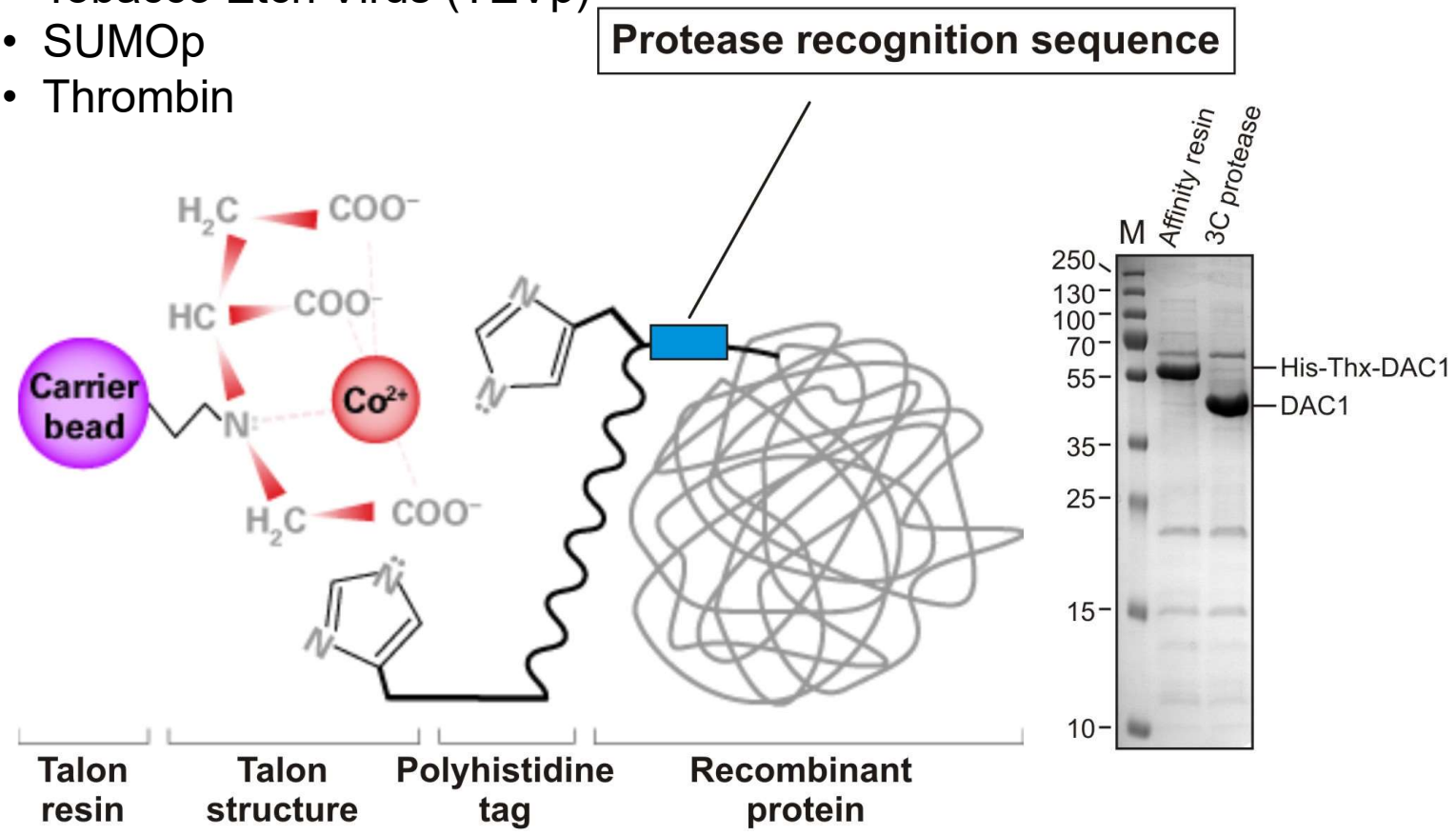
Fusion tags can improve protein expression, stability, resistance to proteolytic degradation and solubility.

Fusion tag	Function	Size (kDa)	Description
Polyhistidine (e.g. 6xHis, 10xHis)	Affinity	1-2	The most commonly used affinity tag, binds to metal ions
Strep-tag II	Affinity	1	High affinity for engineered streptavidin
Thioredoxin (Trx)	Solubility	12	Aids in refolding proteins that require a reducing environment
Small Ubiquitin-like Modifier (SUMO)	Solubility	12	Contains a native cleavage sequence enabling tag removal with SUMO protease
Glutathione S-transferase (GST)	Solubility, affinity	26	High affinity for glutathione, often needs to be removed due to large size
Maltose Binding Protein (MBP)	Solubility, affinity	41	Binds to maltose, often needs to be removed due to large size

- Fusion tag orientation (N- or C-terminus)
- Combinatorial fusion tags (Trx/GST/MBP with an affinity tag, e.g. 6xHis)

Variety of proteases for fusion tag removal

- Human Rhinovirus (HRV 3C)
- PreScission protease
- Tobacco Etch Virus (TEVp)
- SUMOp
- Thrombin

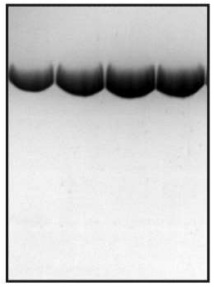


Column chromatography instrumentation

Automated chromatography systems

Chromatography steps

Affinity chromatography
Ion-exchange chromatography
Size-exclusion chromatography



Protein characterization: an aggregation problem

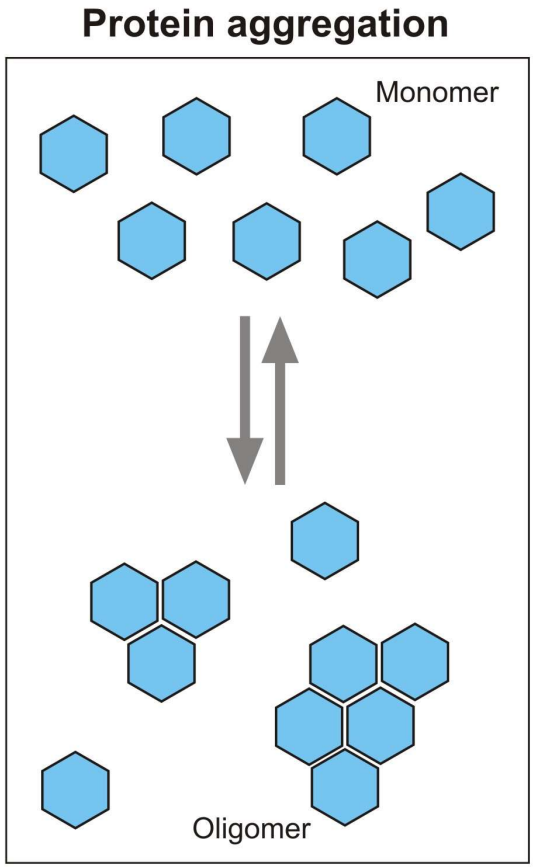
A key challenge in recombinant protein production is to maintain and store the target protein in a soluble and stable form. Protein aggregation can compromise protein function and thus it is necessary to overcome this challenge to generate functionally active protein.

Detection of protein aggregation

- Analytical size-exclusion chromatography (SEC)
- Dynamic light scattering (DLS)
- Analytical ultracentrifugation (AUC)

Troubleshooting

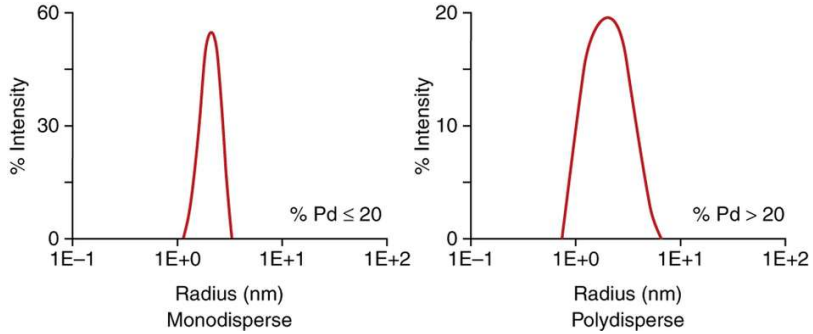
- Culture conditions (e.g. reducing temperature)
- Buffer composition (ionic strenght, pH, reducing agents)
- Presence co-factors (Acetyl-CoA, metal ions)
- Fusion tags (Trx, MBP, SUMO)
- Minimising sample handling
- Avoiding time delays between purification steps
- Performing purification steps at 4°C
- Store purified proteins in -80°C



Biophysical characterization of proteins

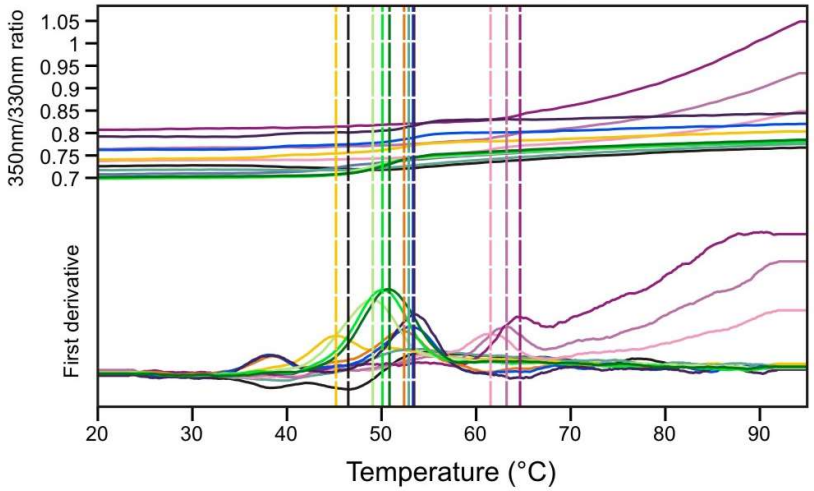
High purity and homogeneity of the protein sample are crucial for the downstream processes to be successful.

- Dynamic light scattering (DLS): To characterize the polydispersity of sample

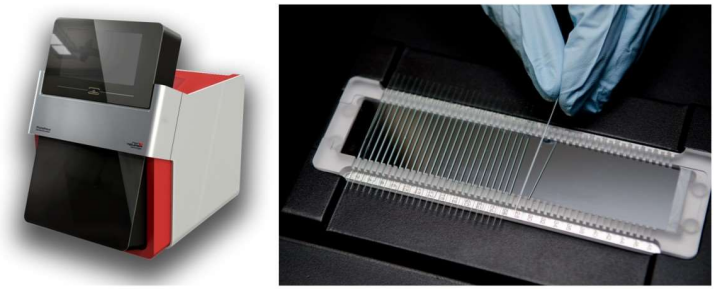


Identification of different oligomeric forms or aggregates, which are preventing crystallization

- Differential scanning fluorimetry (DSF): Analysis of protein stability



To characterize the stability of the protein in different buffers and in the presence of different ligands, which stabilize the protein for crystallization



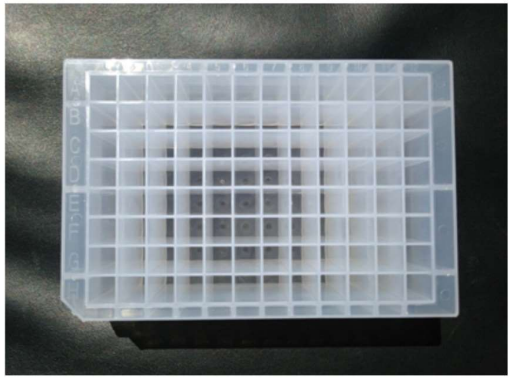
Prometheus NT.48 (nanoDSF)

Examples

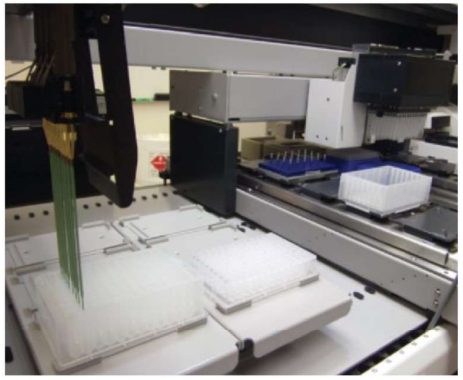
Development of expression and purification protocol for *Schistosoma mansoni* HDAC8: mini-scale tests

24-well deep plate for *E. coli* cultures

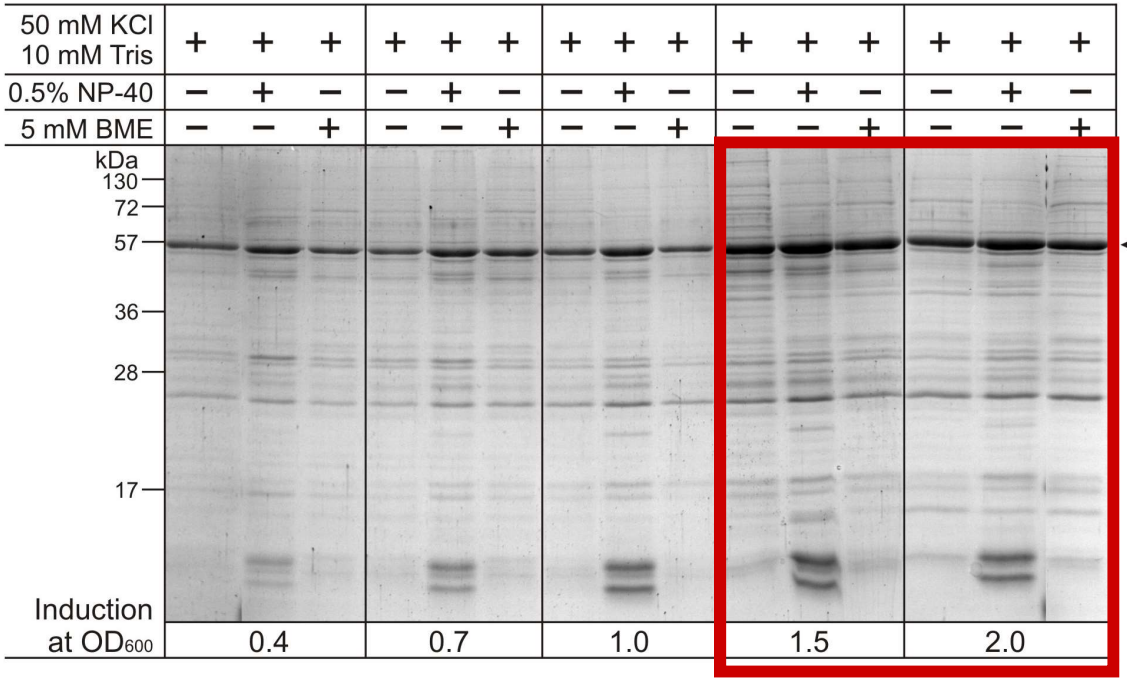
96-well deep plate for purification



Tecan robot



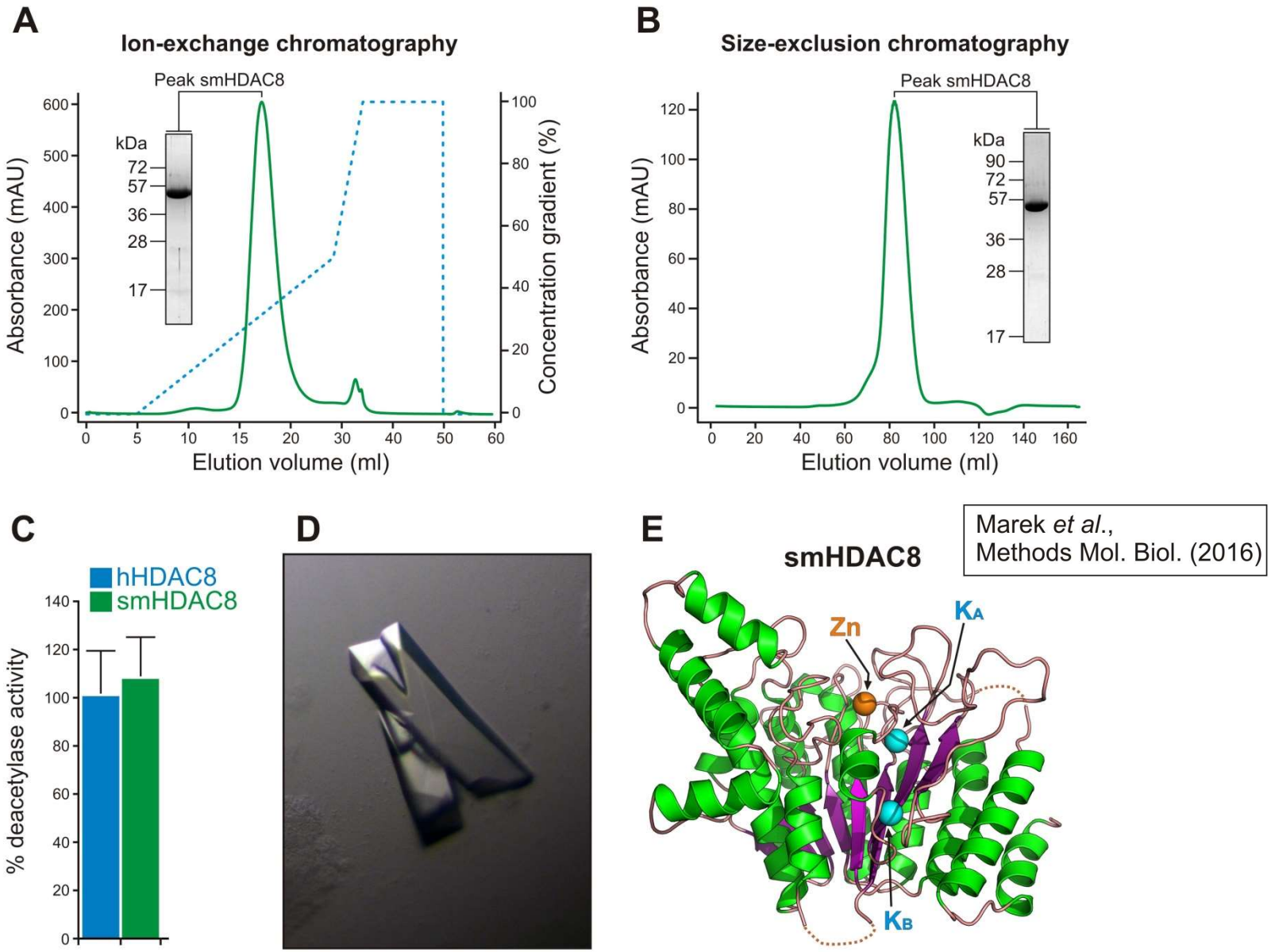
Semi-automated affinity purification



Induction at high O.D. results in higher yield

Harvest cells 1 h post induction

Large-scale production and crystallization of smHDAC8



Conclusions

- The project design is ultimately determined by the end-use of the recombinant protein
- The overall success of a project lies in an effective project design