LOSCHMIDT LABORATORIES

Protein expression & purification technologies

Manufacturing of recombinant proteins, protein complexes and virus-based vehicles for biotechnology and basic research

Dr. Martin Marek
Loschmidt Laboratories
Faculty of Science, MUNI
Kamenice 5, bld. A13, room 332
martin.marek@recetox.muni.cz

What will we talk about

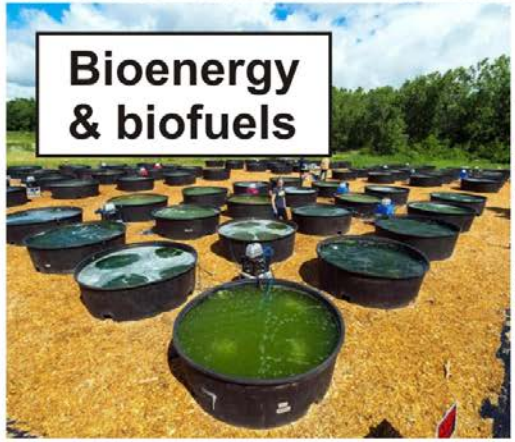
- Introduction to recombinant protein production
  - definition, goals and applications
- Protein expression in microbial cells
  - *Escherichia coli*, yeasts
- Expression and multi-expression in eukaryotic cells
  - Mammalian and insect cells, baculovirus expression
- New emerging expression tools
  - CRISPR/Cas9 technology, cell-free protein expression
- Protein purification, characterization and storage
  - Chromatography separation & QC check

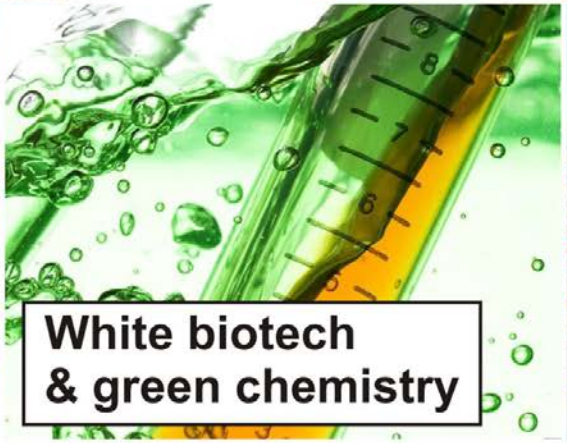
Why to purify a protein?

- To study its function
- To analyze its physical properties
- To determine its sequence
- To determine its three-dimensional structure
- For industrial or therapeutic applications

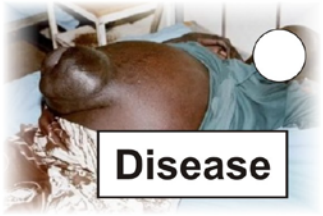
Applications of protein technologies

Why should we bother to think about protein technologies?

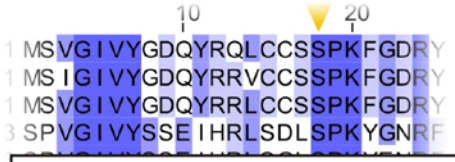




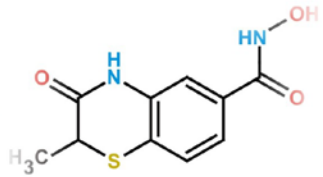
The key role of protein production in pharma industry



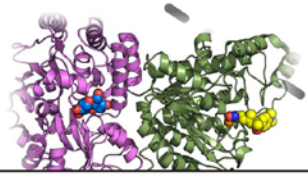
Disease



Identify target
(Genomics, biological assays)

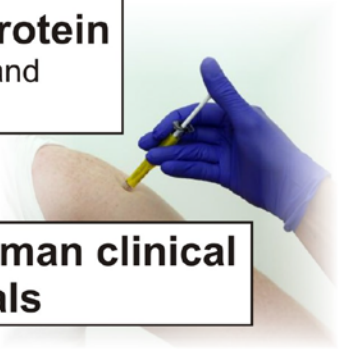


J1038



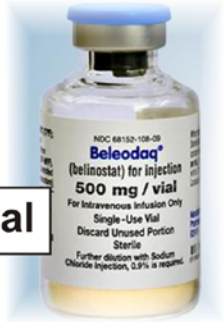
Isolate protein involved in disease
(Molecular cloning, recombinant expression, biochemical and structural analyses)

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



Human clinical trials

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



FDA approval

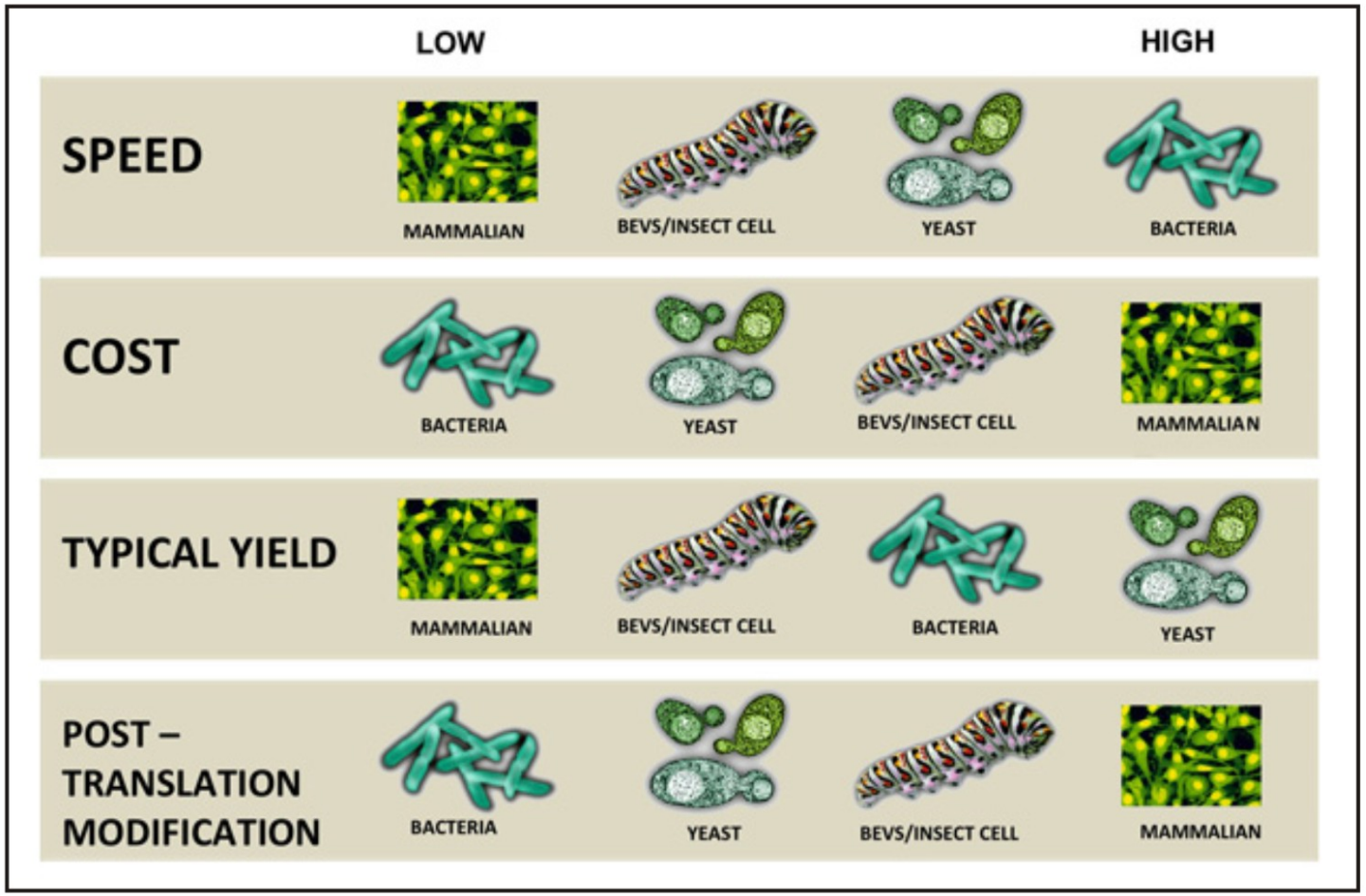
Protein production is
a significant bottleneck in early
phase drug discovery

Recombinant protein production workflow

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

Cellular protein production: selection of a host

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



Bacterial expression system

Concepts
Methods
Applications

Recombinant protein expression in *Escherichia coli*

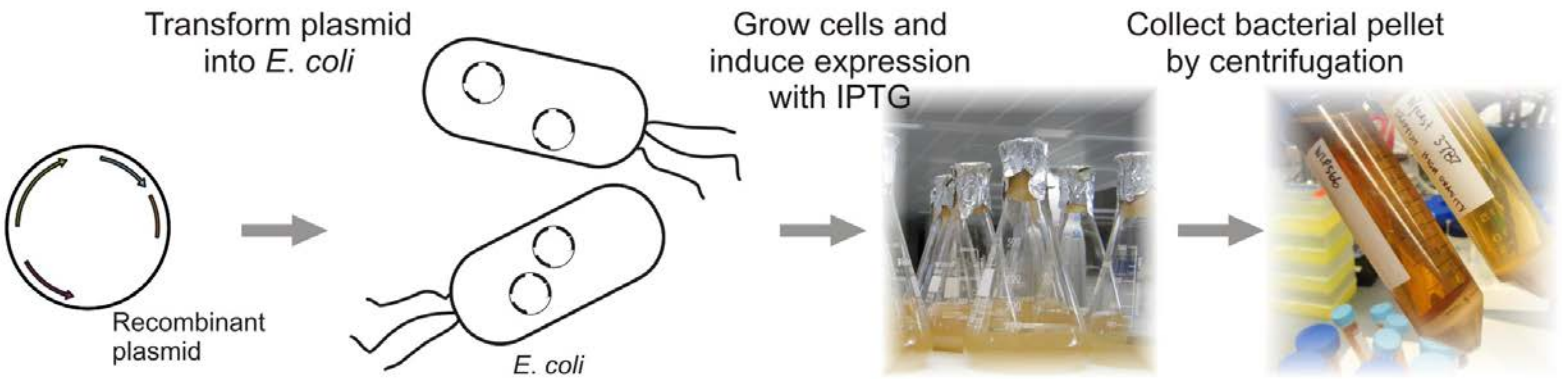
ADVANTAGES

- 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

DRAWBACKS

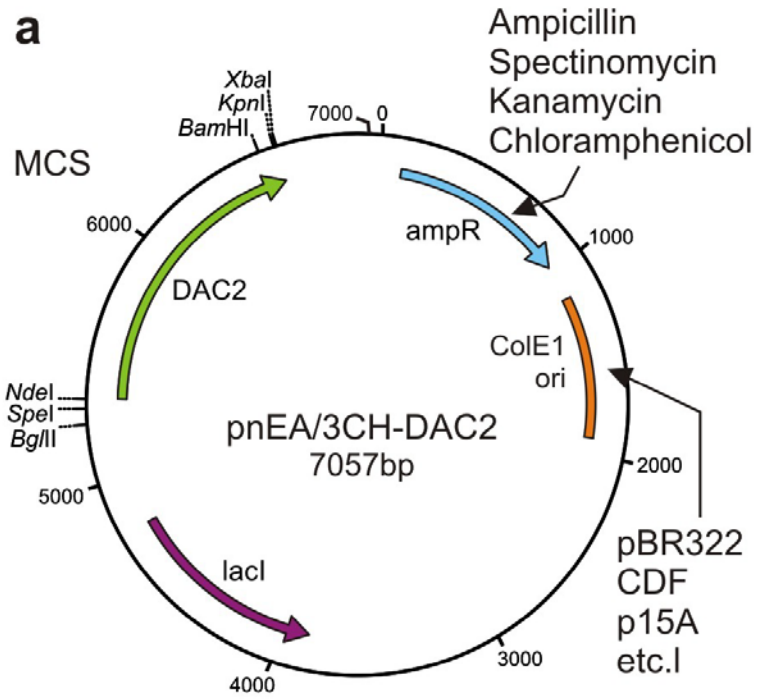
- Inability to perform post-translation modifications (PTMs)
- Limited formation of disulphide bond

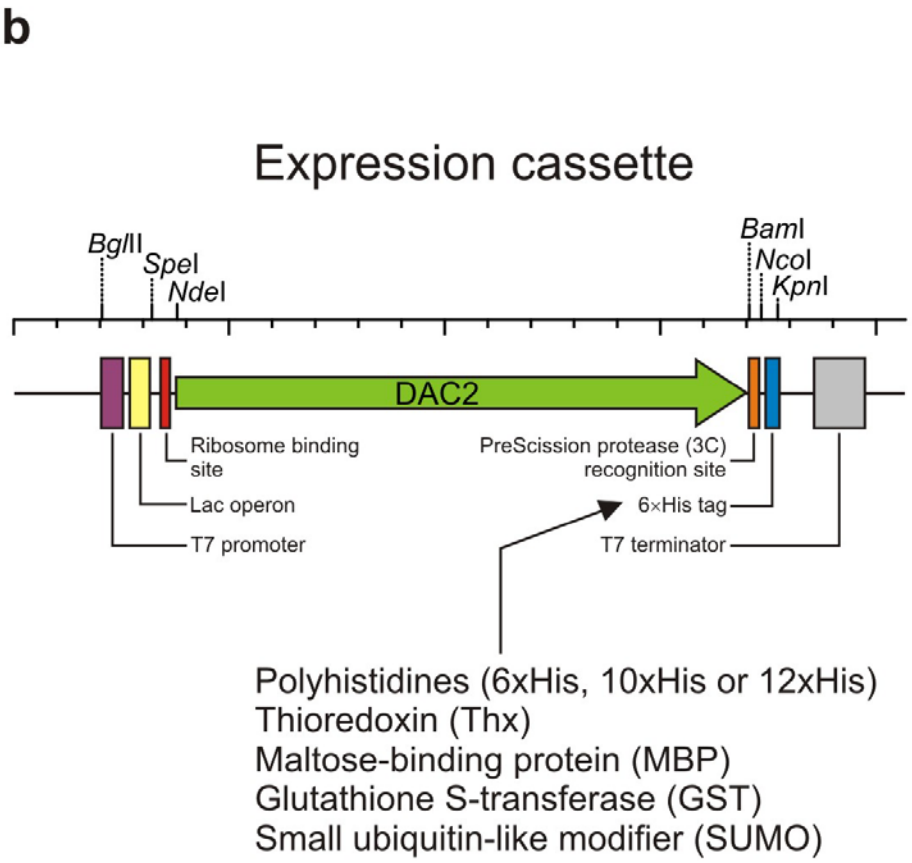
Workflow of protein expression in *E. coli*



Protein expression in *E. coli*: a plasmid backbone

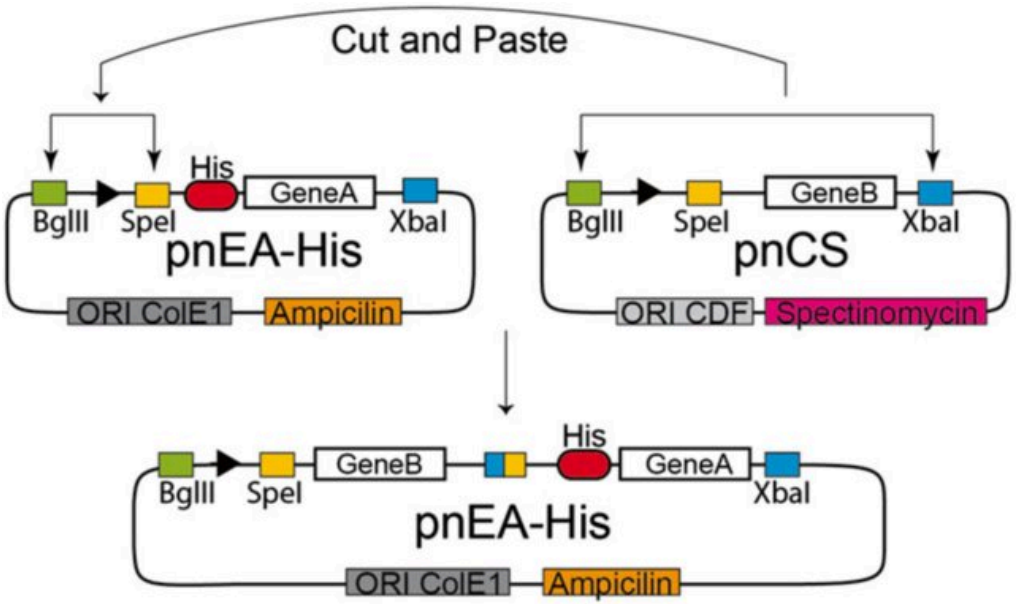
- pET
- pGEX





Co-expression of protein complexes in *E. coli*

The functional units within cells are often macromolecular complexes rather than single species. Production of these complexes as assembled homogenous samples is a prerequisite for their biophysical and structural characterization and hence an understanding of their function in molecular terms. Co-expression in *Escherichia coli* can decipher the subunit composition, assembly, and production of whole protein complexes.



- Example of the concatenation of two vectors of the pET-MCN vector series: pnEA-His (vector encoding an N-terminal poly-histidine tag in front of protein A) and pnCS (vector expressing the native protein B).
- The pnEA-His (acceptor vector) is linearized by removing with the restriction enzymes BglII and SpeI part of its promoter (T7 promoter and lacO).
- The full promoter of the pnCS (donor vector) is cut out with the restriction enzymes BglII and XbaI. After ligation of the pnCS promoter with the linearized pnEA-His vector, a new vector is obtained based on the backbone of the pnEA-His vector and whose promoter controls the genes of proteins A and B.

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

Expression in *E. coli*: tips and tricks

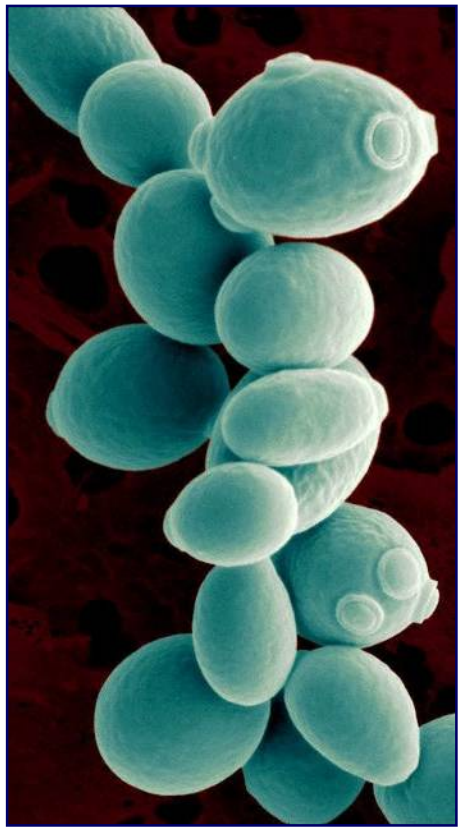
- **Growth temperature:** Typically 37°C, but lowering (25°C or 15°C) improves folding/solubility
- **Expression with a fusion tag:** polyhistidenes (6xHis, 10xHis or 12xHis)
  Thioredoxin (Thx)
  Maltose-binding protein (MBP)
  Glutathione S-transferase (GST)
  Small ubiquitin-like modifier (SUMO)
- **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

Expression in yeasts

Concepts
Methods
Applications

Expression in yeasts

- The yeast systems, such as *Saccharomyces cerevisiae* and *Pichia pastoris*, constitute another microbial systems in heterologous protein expression.



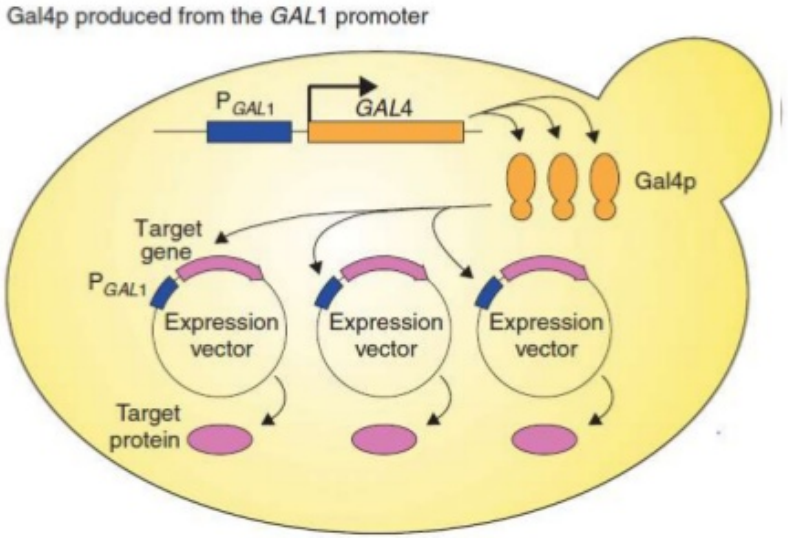
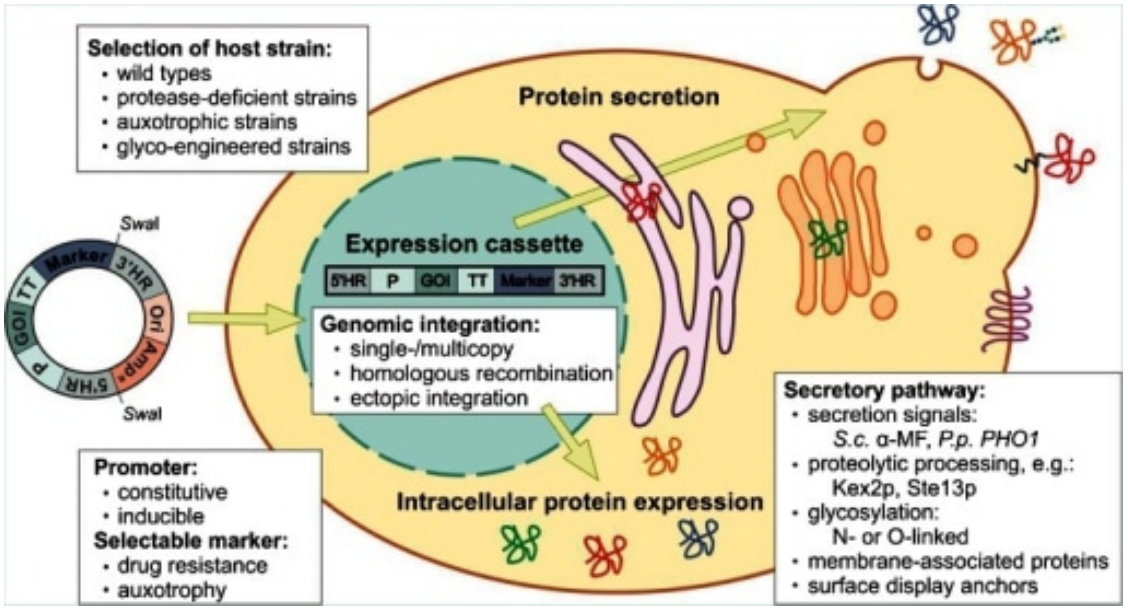
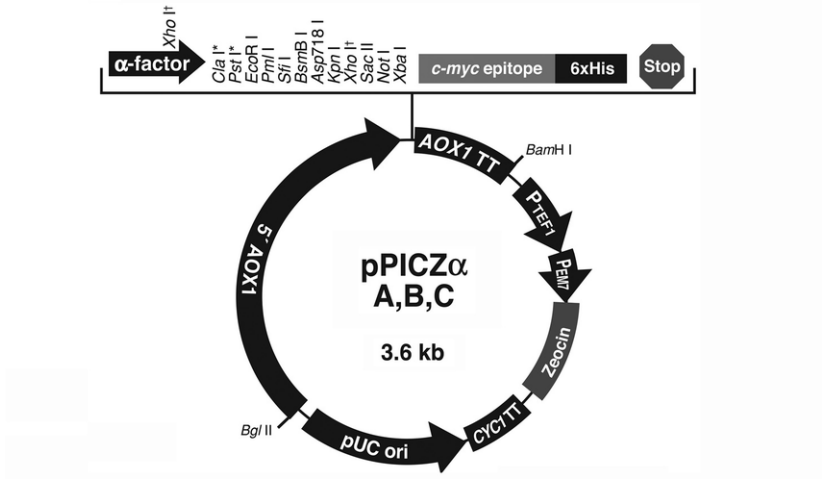


Figure 8.5. Galactose inducible gene expression in yeast. The expression of genes from multicopy vectors under the control of the *GAL1* promoter (P_{GAL1}) can be increased substantially if the gene encoding the transcriptional activator of *GAL1*, *GAL4*, is also placed under the control of P_{GAL1}. In this case, induction by galactose will produce more Gal4p and consequently more of the target protein

Expression in *Pichia pastoris*





- *Pichia pastoris* (*P. pastoris*) is a widely used protein expression host for the production of biopharmaceuticals and industrial enzymes
- *P. pastoris* belongs to methylotrophic yeasts, which share a common pathway to metabolize one-carbon compounds as carbon and energy sources
- Vectors: The promoters for protein expression in *P. pastoris* include inducible promoters (AOX1, FLD1, ADH1, *etc.*) and constitutive promoters (GAP, TEF1, *etc.*). For *P. pastoris*, *HIS4* (auxotrophic markers) and zeocin resistance (dominant markers) are the most popularly markers used. The heterologous proteins can be intracellular or secreted expression. *P. pastoris*has the ability to secrete high titres of proteins into culture media

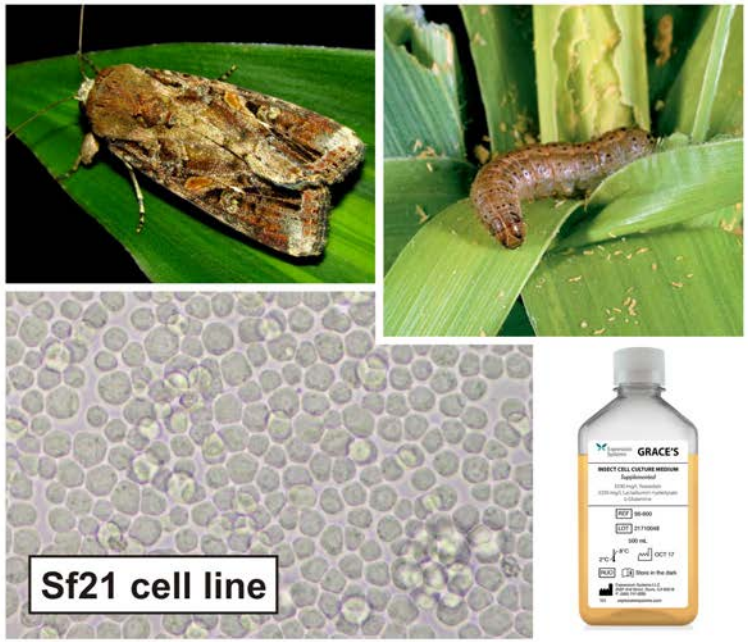
Expression in insect cells

Concepts
Methods
Applications

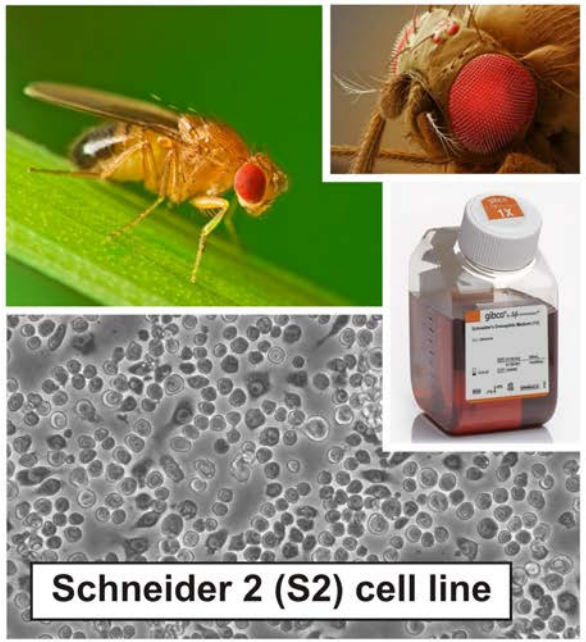
Production of proteins in insect cells

- Insect cells can efficiently express recombinant biologically active proteins and are mostly used for the development of virus-like particles and vaccines. It has been proven that insect cells are excellent platforms for the production of recombinant antibodies
- There are mainly three insect expression systems: baculovirus expression vector system (BEVS), InsectSelect (IS) system and *Drosophila* expression system (DES)

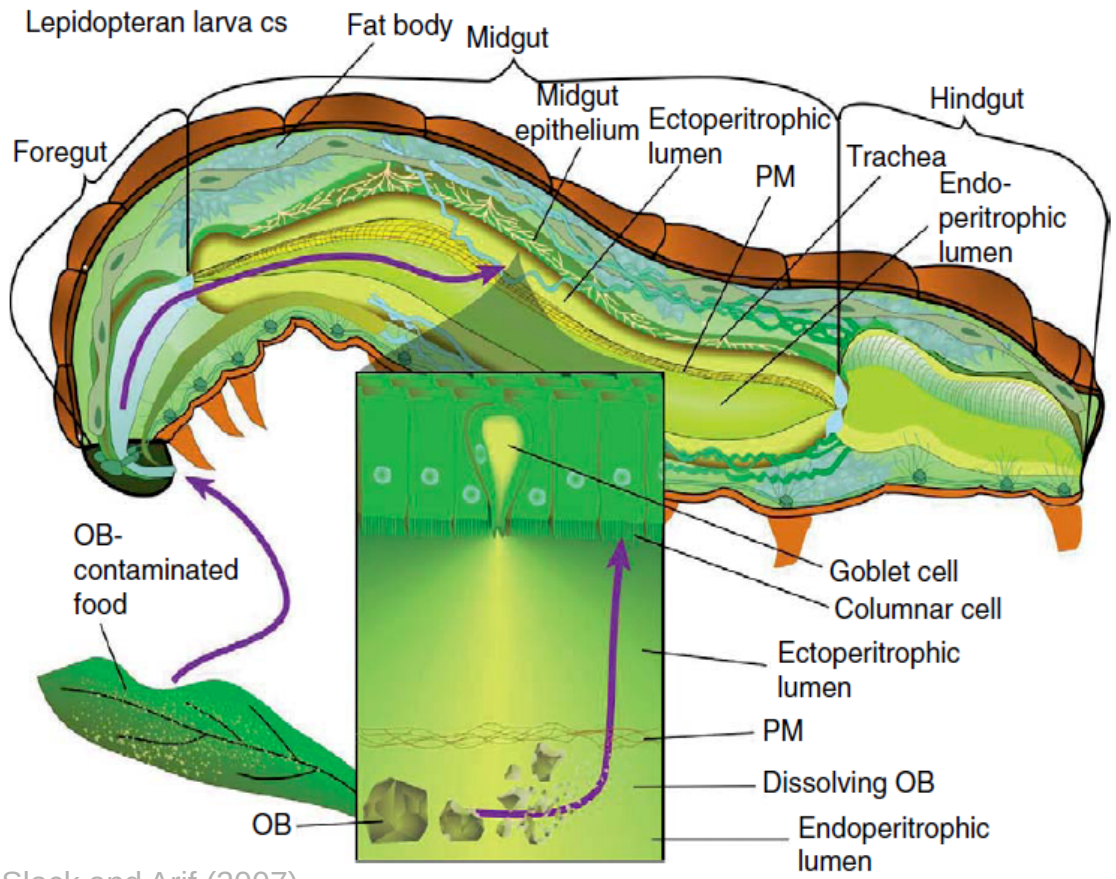
***Spodoptera frugiperda*, a fall army worm**



Drosophila melanogaster



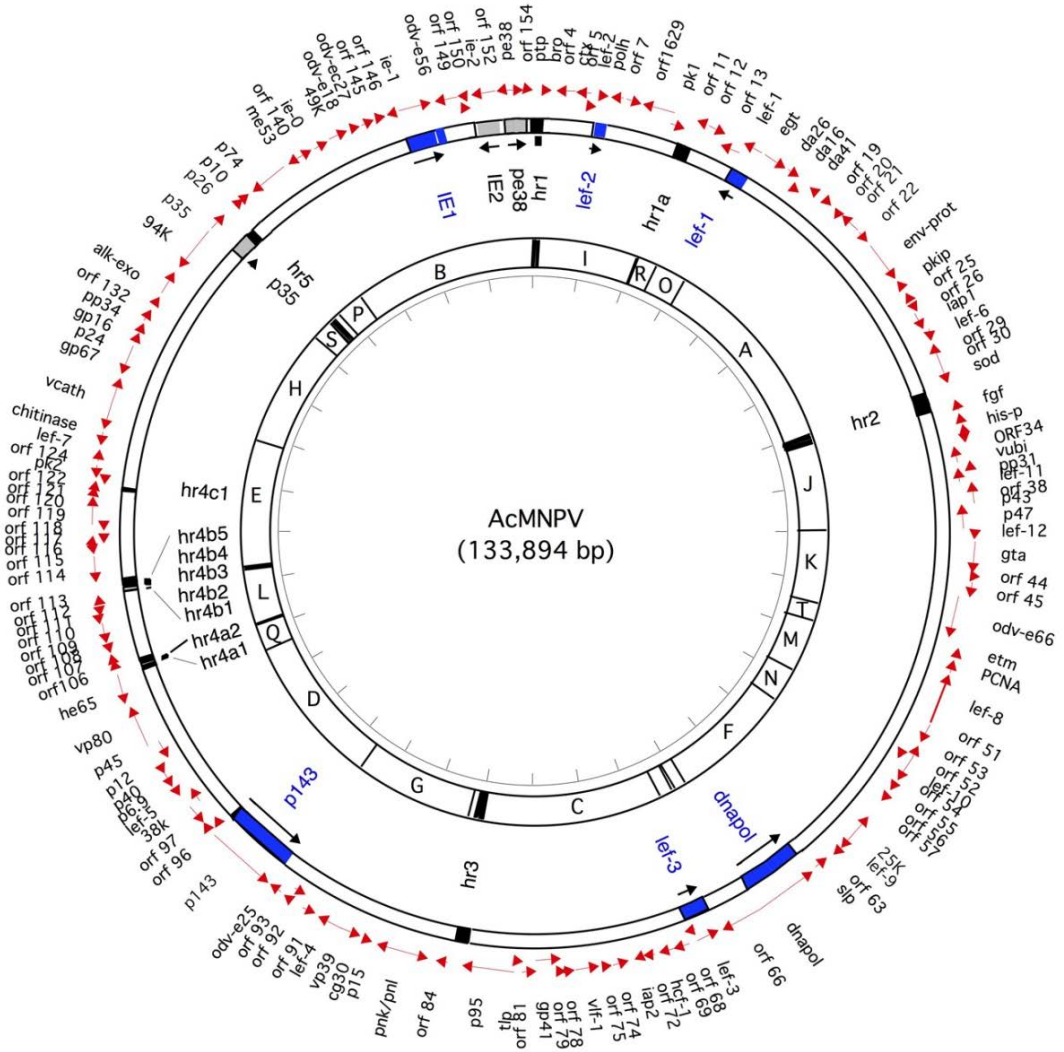
Introduction to baculovirus biology



Slack and Arif (2007)

- Insect-infecting enveloped DNA viruses
- Baculoviruses are a very diverse group of viruses with double-stranded, circular, supercoiled genomes, with sizes varying from about 80 to over 180 kb, that encode between 90 and 180 genes
- The genome is packaged in rod-shaped nucleocapsids that are 230-385 nm in length and 40–60 nm in diameter
- In the most well characterized baculoviruses, the virions are present as two types, occluded virions (ODV) and budded virions (BV).
- Although these two types of virions are similar in their nucleocapsid structure, they differ in the origin and composition of their envelopes and their roles in the virus life cycle

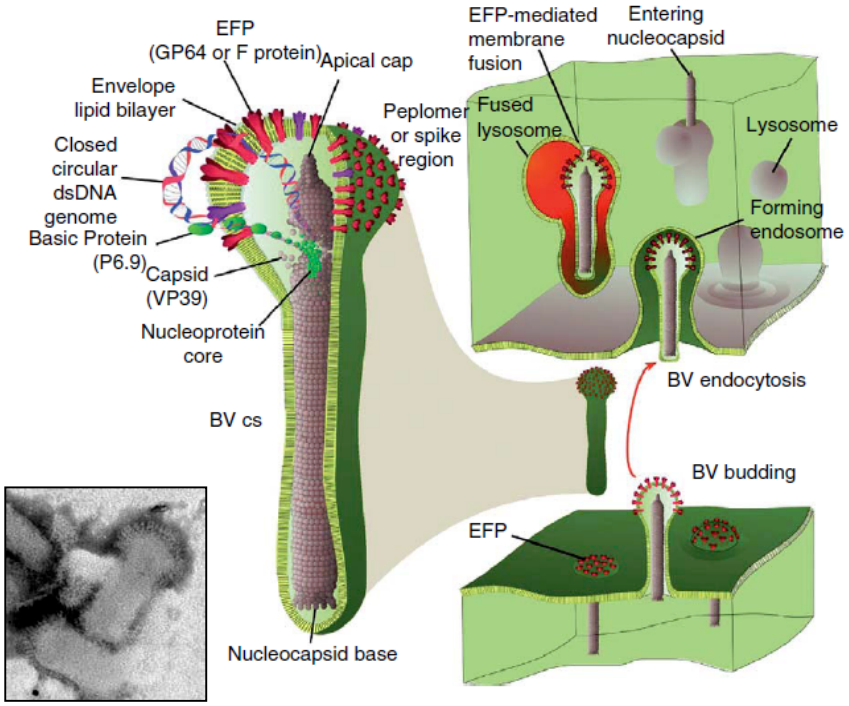
The baculovirus genome



- Circular, supercoiled dsDNA genomes, with sizes varying from about 80 to over 180 kb, that encode between 90 and 180 genes
- The genome is packaged in rod-shaped nucleocapsids that are 230-385 nm in length and 40–60 nm in diameter

Baculovirus morphogenesis: two types of virions

Budded virus (BV)

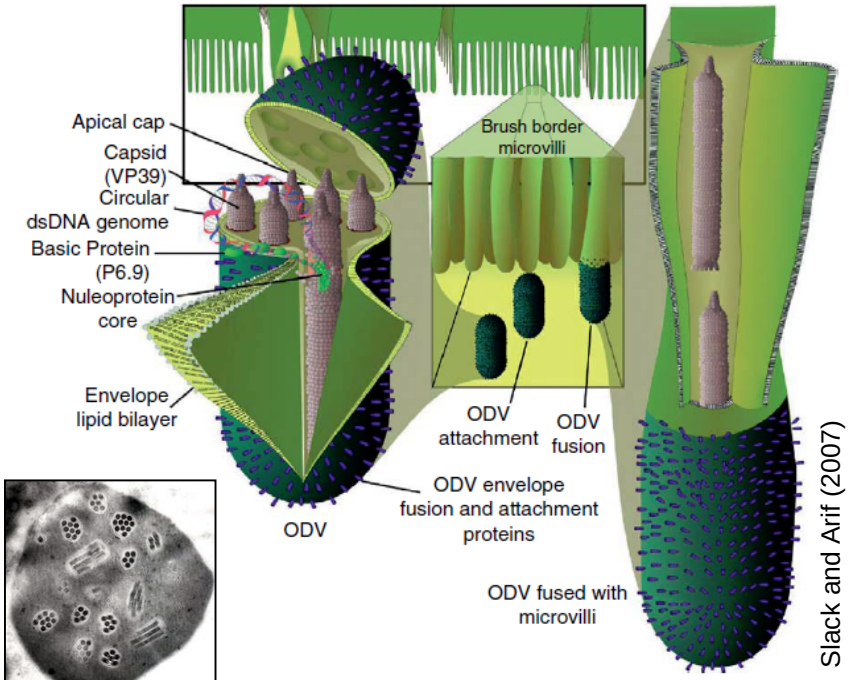


Circular dsDNA genome and associated proteins

Capsid shell (VP39, etc.)

BV envelope (GP64, F)

Occlusion-derived virus (ODV)



Circular dsDNA genome and associated proteins

Capsid shell (VP39, etc.)

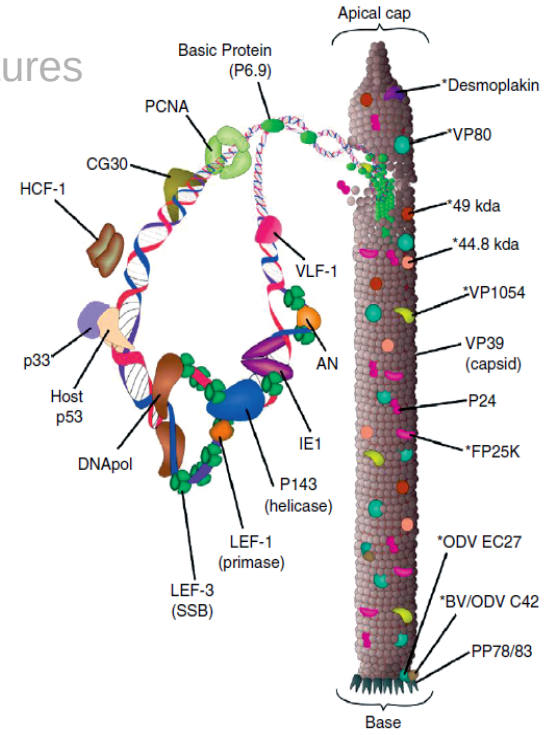
ODV envelope (P74, ODV-E66, PIFs etc)

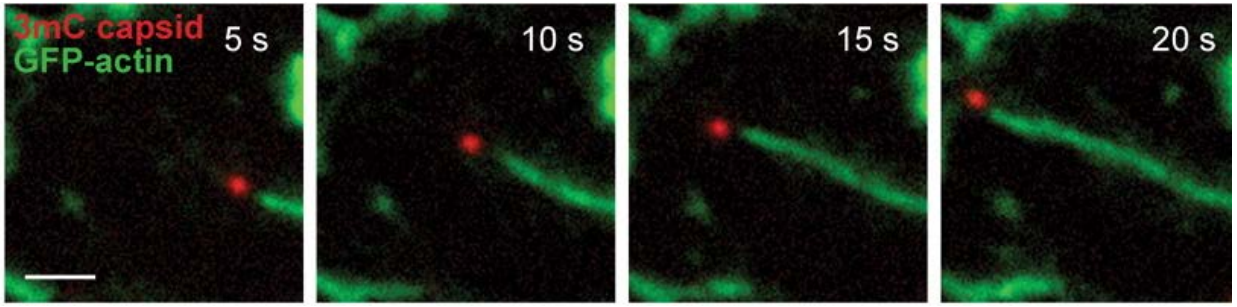
Tegument (GP41), occlusion body (Polyhedrin)

The structure of baculovirus nucleocapsid

- Rod-like capsid (VP39, etc.)
- Distinct apical and basal structures
- Circular dsDNA genome
- PP78/83: Wiskott-Aldrich syndrome protein (WASP)-like protein

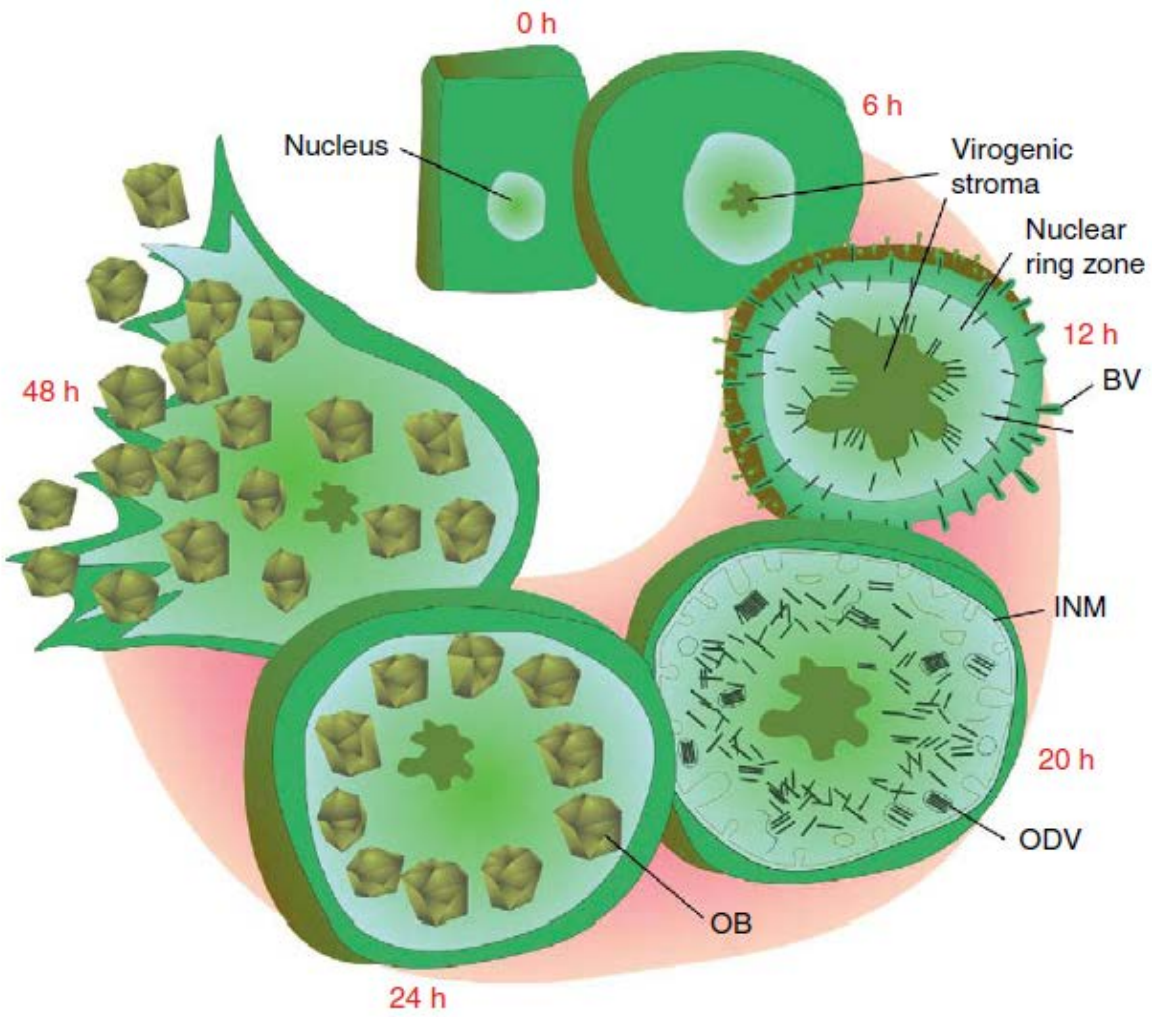
Dr. Taro Ohkawa
University of California
Berkeley, USA



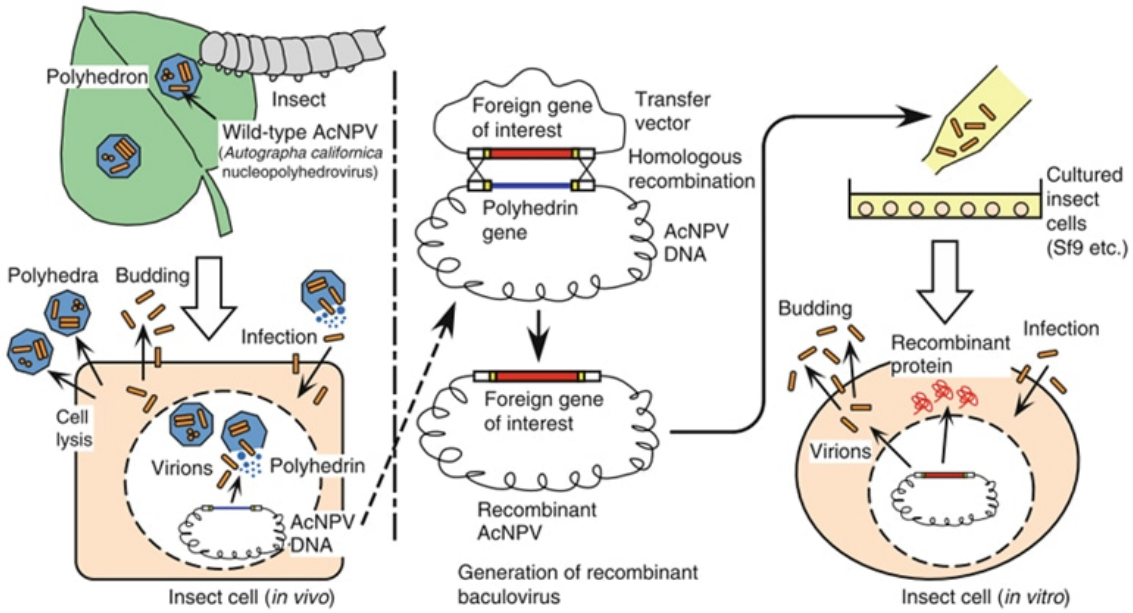


Ohkawa *et al*. (2010)
J. Cell Biol.

Time-course baculovirus infection

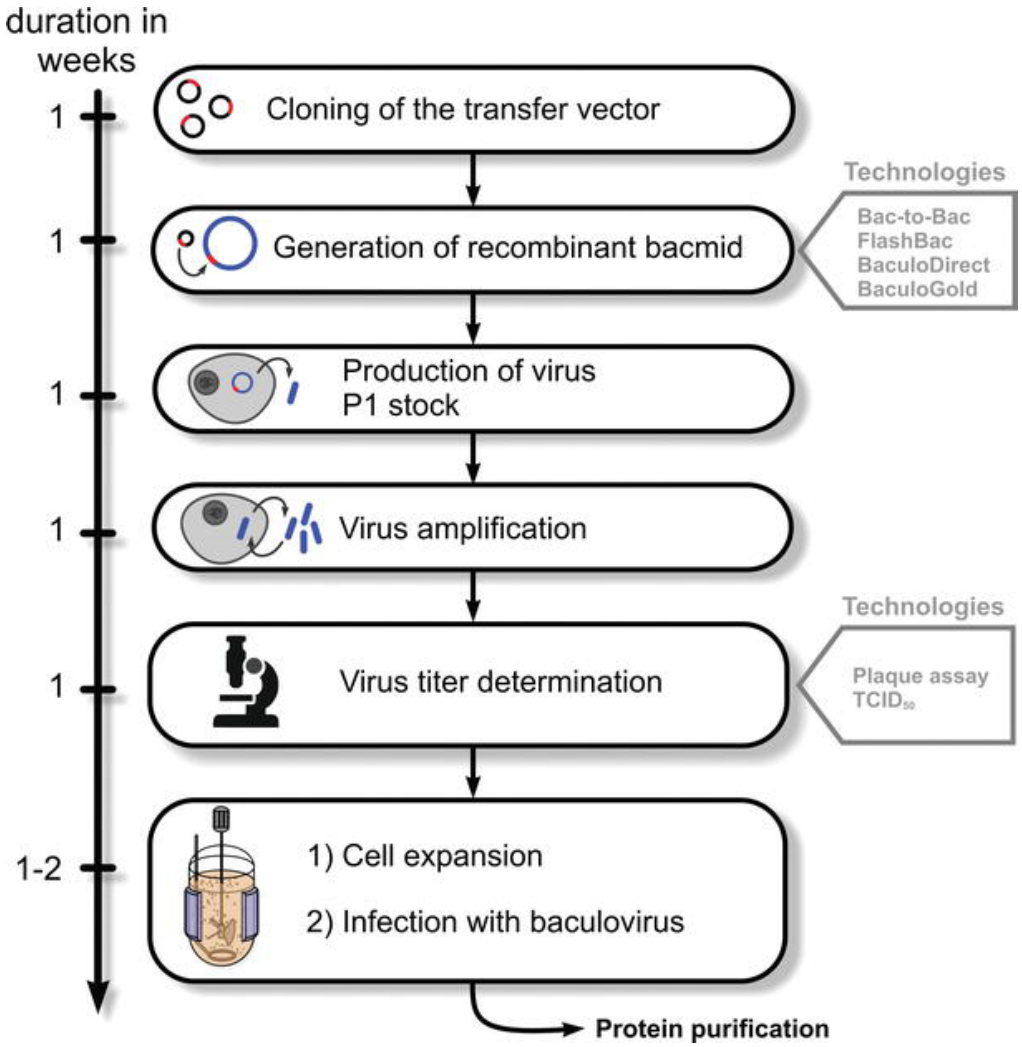


Baculovirus Expression Vector System (BEVS)



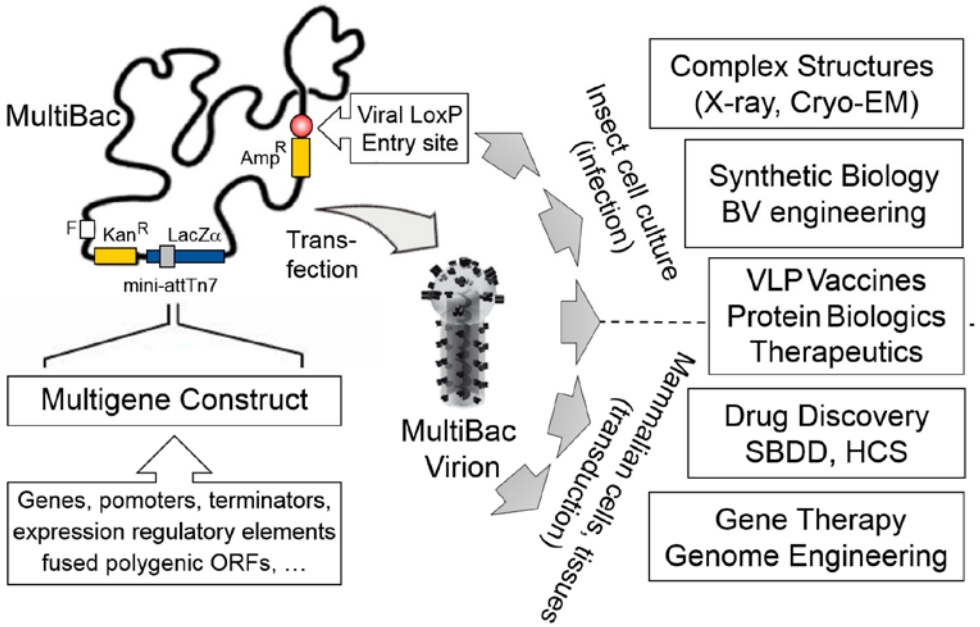
- Huge quantities of **polyhedrin** produced by baculoviruses are the fundamental features to develop the viruses as vectors for foreign protein production
- Later findings manifested that the polyhedrin protein is not necessary for baculovirus replication in infected insect cells
- So, the strategy for expressing recombinant protein by replacing viral DNA sequence encoding polyhedrin with the gene of interest is produced
- This theory became a reality when scientists performed homologous recombination of a polyhedrin region in the viral genome with a plasmid containing a foreign gene regulated by the polyhedrin promoter to produce a recombinant baculovirus, therefore triggering the birth of baculovirus-insect expression system

Workflow of protein production using BEVS



https://www.intechopen.com/books/new-insights-into-cell-culture-technology/process-optimization-for-recombinant-protein-expression-in-insect-cells

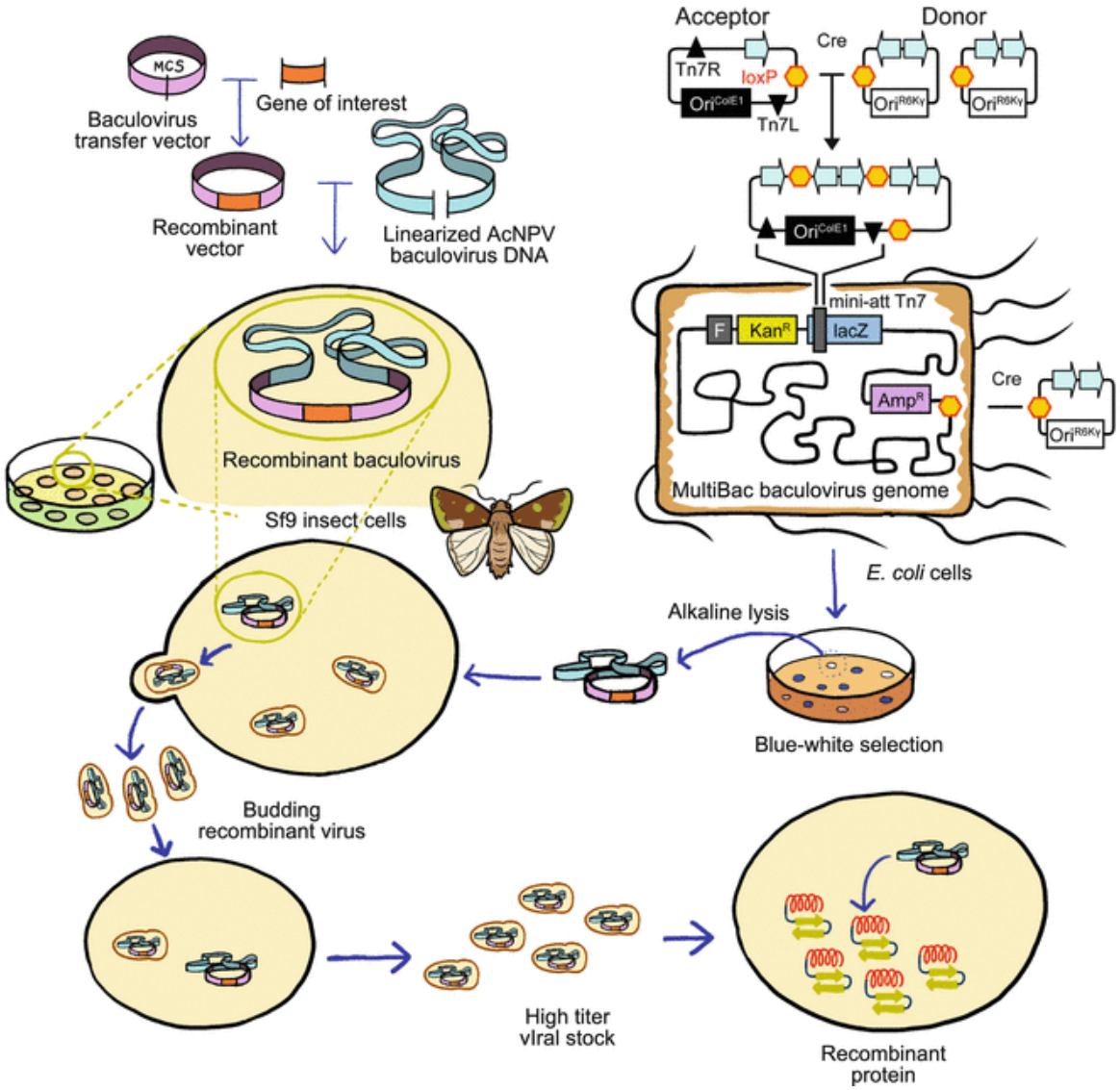
The MultiBac, a multi-expression system



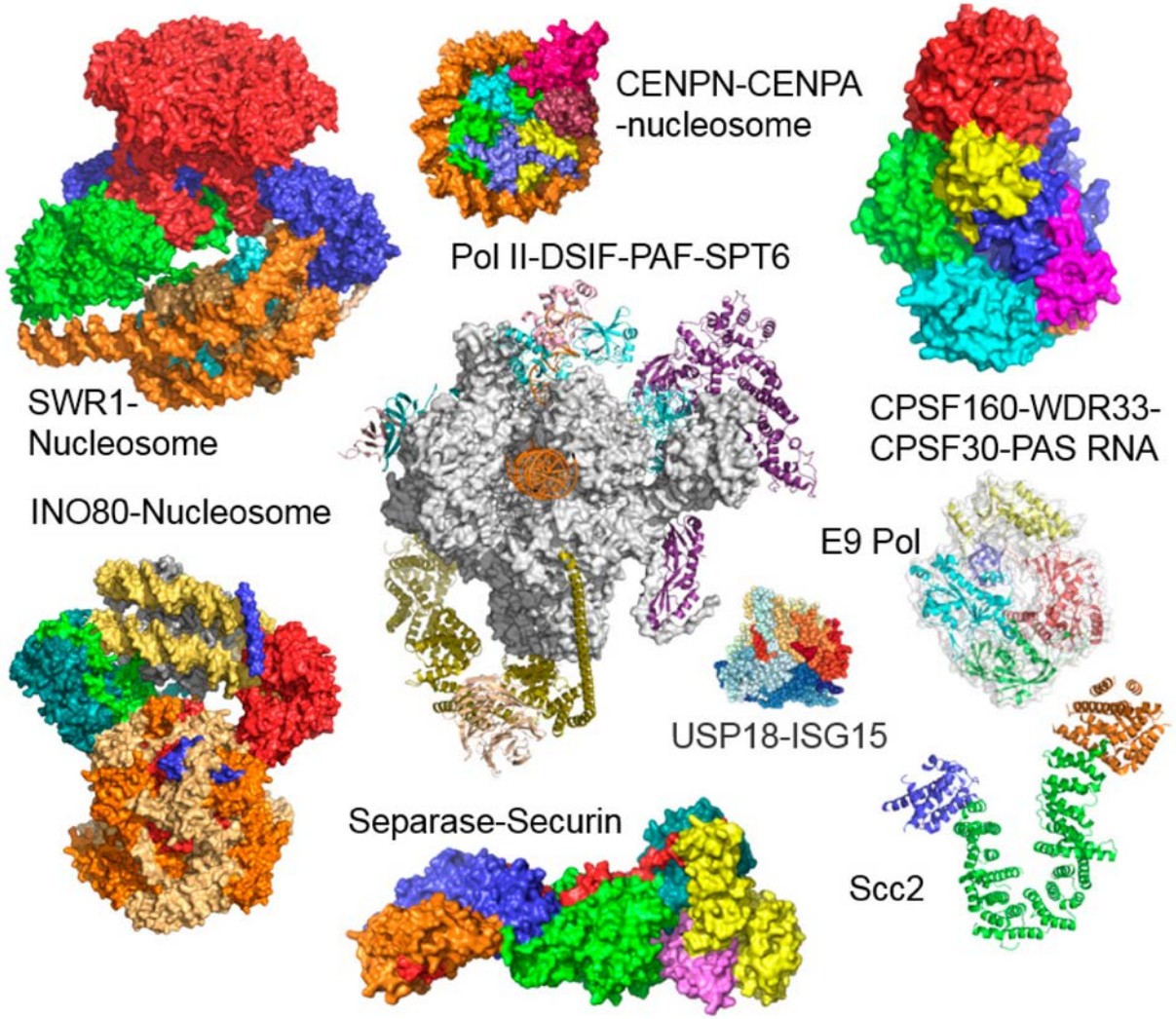
Prof. Imre Berger
School of Biochemistry & Bristol Synthetic Biology Centre, University of Bristol, UK

- MultiBac consists of a baculoviral genome optimized for multigene delivery and protein complex expression (left).
- The genome is present as a bacterial artificial chromosome (BAC) in *E. coli* cells supplying the Tn7 transposase.
- Expression cassettes are assembled into the multi-gene expression constructs and inserted into the MultiBac genome by Tn7 transposition.
- A second entry option into the viral backbone is provided distal from the Tn7 site, relying on Cre recombinase catalysed site-specific integration into a LoxP sequence (circle filled in red).
- Composite MultiBac baculoviral DNA containing all DNA elements of choice is extracted from *E. coli* cultures, followed by transfection into insect cell cultures to manufacture functional MultiBac virions.
- These are then used for a wide range of applications (right), by the infection of insect cell cultures or transduction of mammalian cells, tissues and organisms.

Application of MultiBac system to large complexes

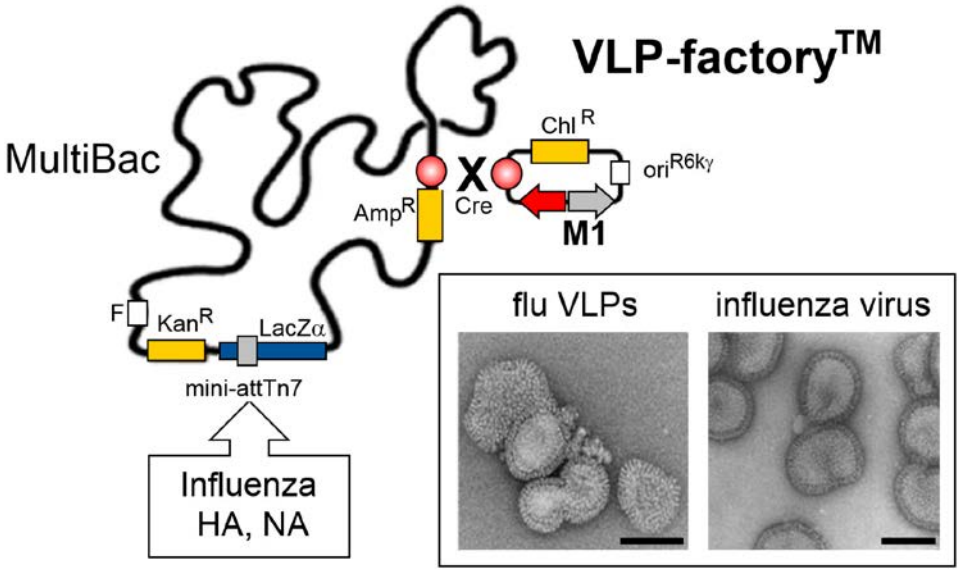


The MultiBac system: successful structural stories



- Chromatin remodeling enzymes SWR1 (14 subunits) and INO80 (11 subunits)
- The yeast polII DSIF-PAF-Spt6 cryo-EM structure
- The CENPN-CENPA-nucleosome complex
- The human CPSF-160–WDR33–CPSF-30–PAS RNA quaternary complex
- The E9 polymerase
- The USP18-ISG15 complex
- The Separase-Securin complex
- The cohesion loader Scc

The MultiBac: a factory for synthetic virus-like particles



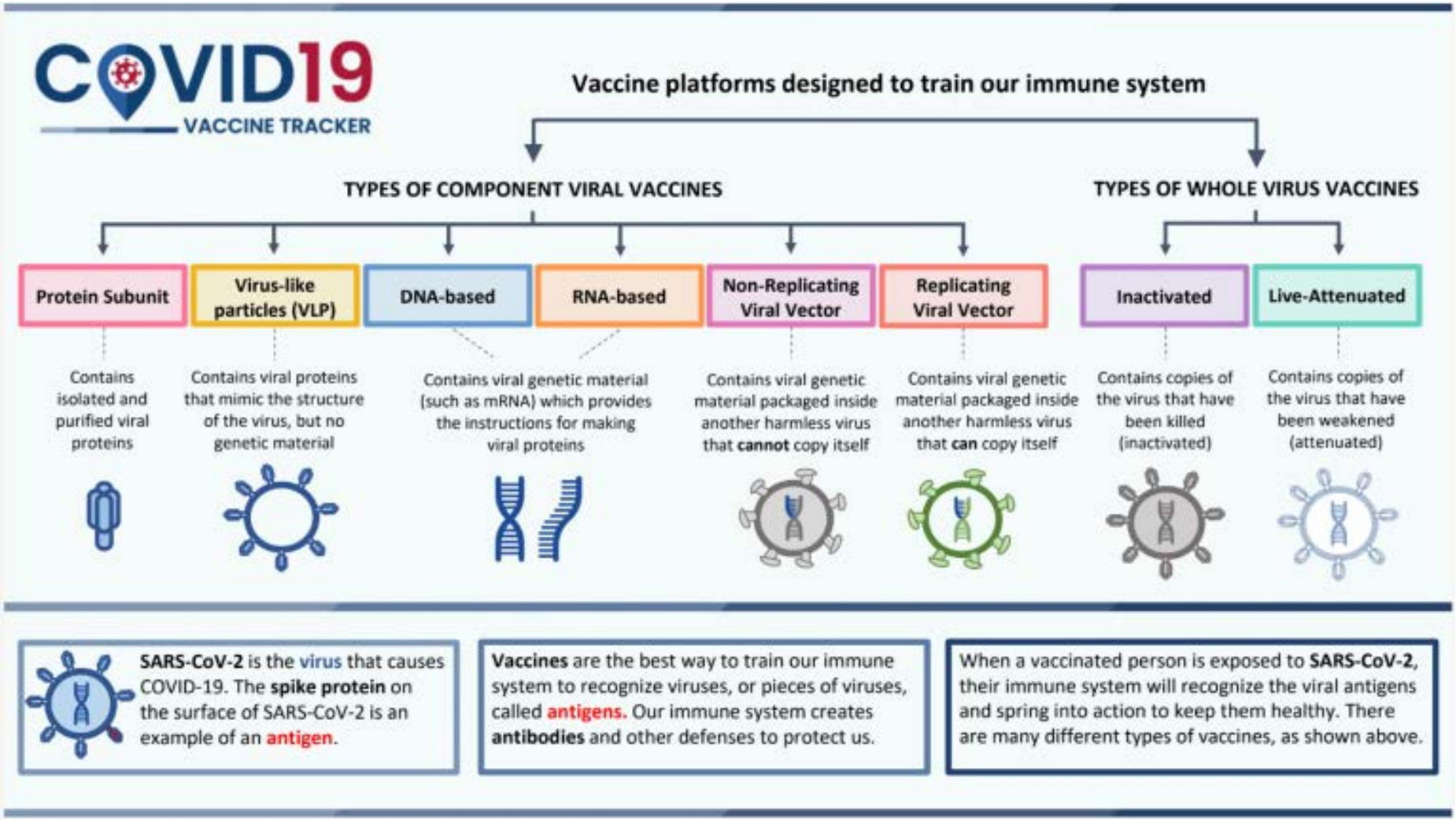
- A plasmid module comprising expression cassettes for the capsid-forming influenza H1N1 M1 protein (colored in grey) and a fluorescent protein marker, mCherry (colored in red), was introduced into the MultiBac baculoviral genome by Cre recombinase enzyme mediated plasmid fusion into the LoxP site (circle filled in red, gradient)
- Co-expression of HA, NA and M1 yields synthetic influenza virus-like particles (VLPs) resembling live influenza virus

Human papillomavirus-based vaccine produded by baculovirus expression system

Selected vaccines produced by BEVS

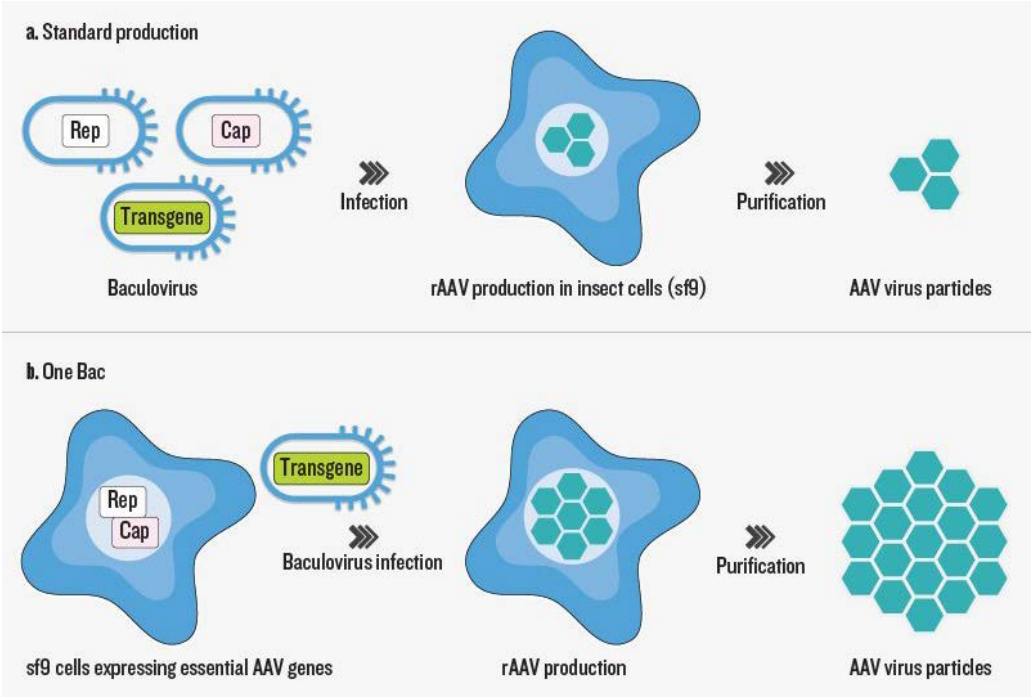
Application	Product name	Company	Stage	References
For human use				
Cervical cancer	CERVARIX®	GSK	Approved	[27]
Prostate cancer	PROVENGE®	Dendreon	Approved	[28]
Influenza	FluBlok®	Protein Sciences	Approved	[29, 30]
Influenza	A/H_5N_1 Virus-like particle	Novavax	Phase I (NCT01596725)	[31]
For veterinary use				
Procrine circovirus 2 (PCV_2)	Porcilis® PCV	Merck	Approved	[32]
PCV_2	CircoFLEX®	Boehringer Ingelheim	Approved	[33]
Swine fever	Porcilis Pesti®	Merck	Approved	[34]

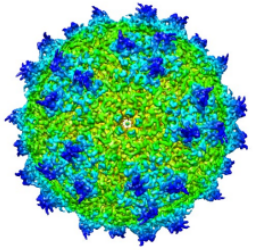
Types of COVID-19 vaccines

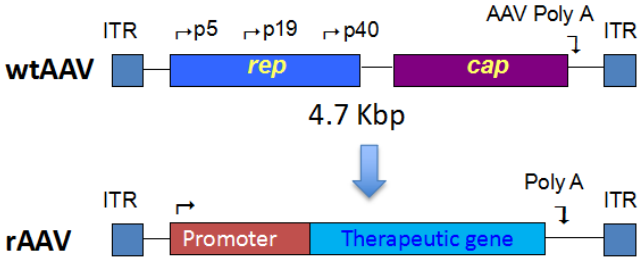


Baculovirus-mediated production of gene therapy vectors

Baculovirus-mediated production of adeno-associated viral (AAV) vectors as gene therapy vehicles



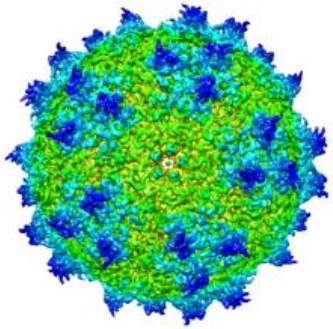


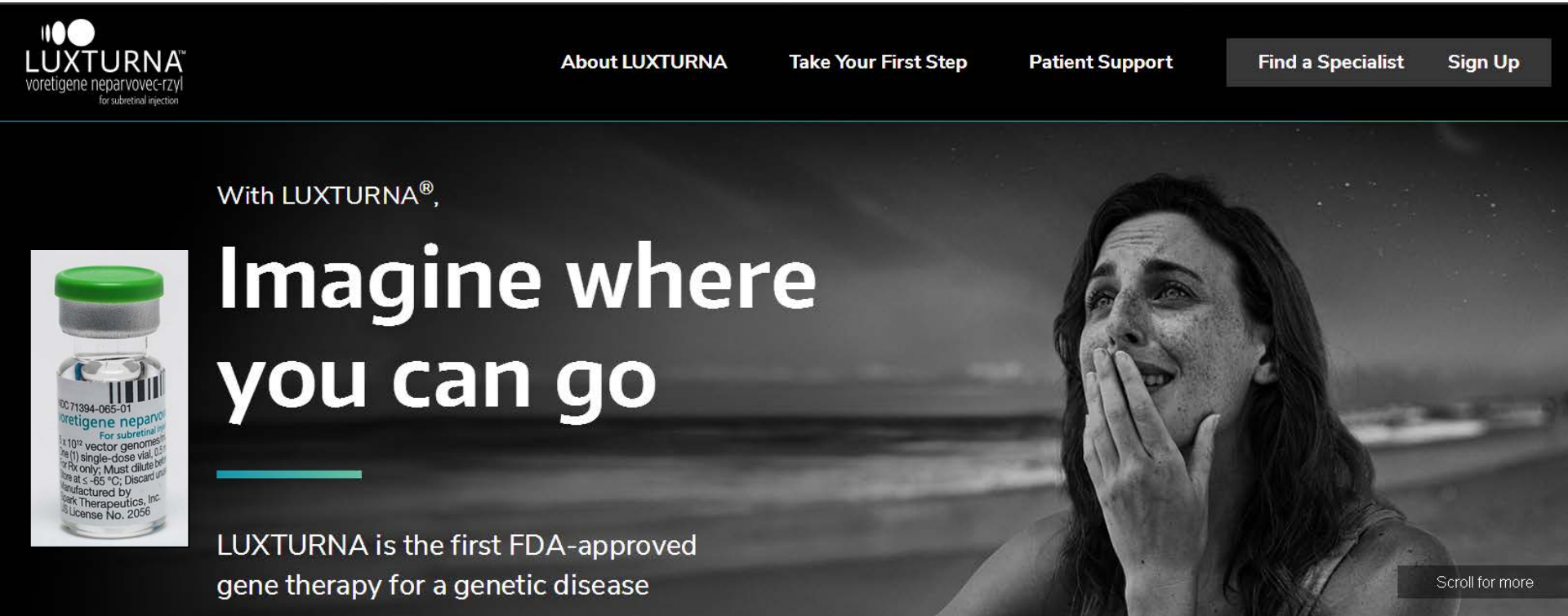


- In 2002, Prof. Robert Kotin and colleagues at the US National Heart, Lung, and Blood Institute first demonstrated its suitability for AAV manufacturing
- They infected Sf9 cell lines — derived from the fall armyworm — with three different baculoviruses: two containing essential genes for AAV particle production (rep and cap), and one containing the transgene sequence intended for delivery
- In this manufacturing process, the baculoviruses play a dual role, functioning as the ‘helper’ virus normally required for replication, as well as the vehicle for AAV genetic material
- In their initial demonstration, Kotin’s team achieved levels of productivity comparable with existing AAV manufacturing approaches — on the order of 50,000 functional viral particles per cell

Development of AAV-based gene therapy products

- In 2017, FDA-approval of the first gene therapy product targeting a disease caused by mutations in a single gene
- This product, LUXTURNA™ (voretigene neparvovec-rzyl; Spark Therapeutics, Inc., Philadelphia, PA), delivers a normal copy of the *RPE65* gene to retinal cells for the treatment of *RPE65* mutation–associated retinal dystrophy, a blinding disease
- Many additional gene therapy programs targeting both inherited retinal diseases and other ocular diseases are in development



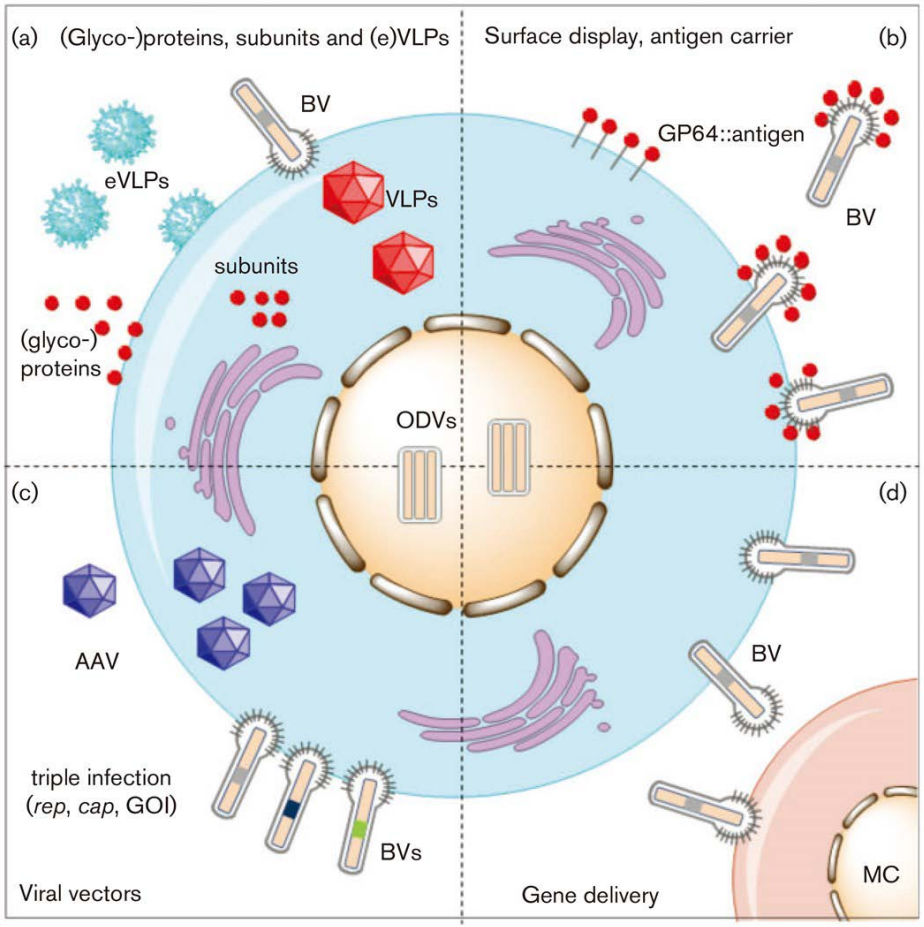


World's most efficient large-scale AAV manufacturing

VIROVEK

- Virovek has developed a patented BAC-to-AAV technology that utilizes the baculovirus expression system to produce AAV vectors in insect cells under serum-free condition
- The capability to generate over 3e+16vg of AAV vectors with a single production run, which is unmatched by any other AAV production system

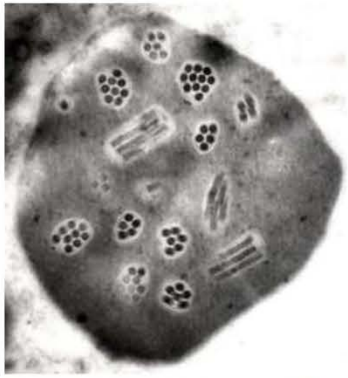
Applications of BEVS

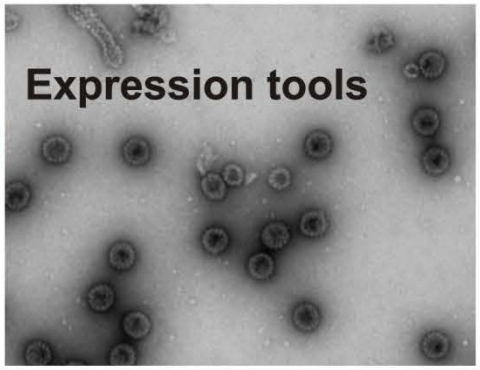


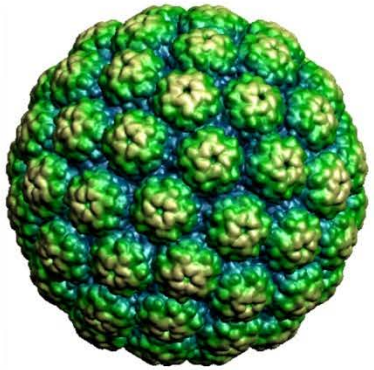
Overview of the various applications of the baculovirus expression system.

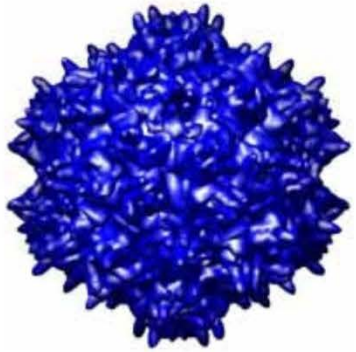
- Nowadays, baculovirus-insect expression system has been relatively mature in many labs for protein production and is becoming a leading method to produce high quality proteins among the eukaryotic expression systems
- The use of BEVS to express the protein of interest requires two steps. In the first step, the insect cells are cultured to the desired concentration. In the second step, the baculovirus is applied to infect the insect cells. The virus that infects host cells will dominate the gene expression mechanism in the host cell and trigger the production process of the target protein
- Yet, several mature expressing vectors have been developed, including *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) that is the most extensively used
- The ability to produce biologically active mammalian proteins makes the BEVS a powerful tool over yeast and bacterial expression systems

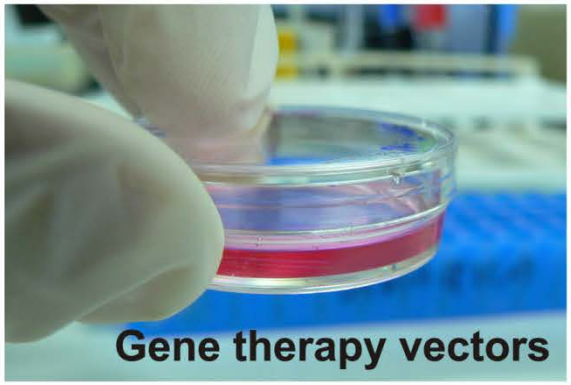
Applications of baculovirus technologies

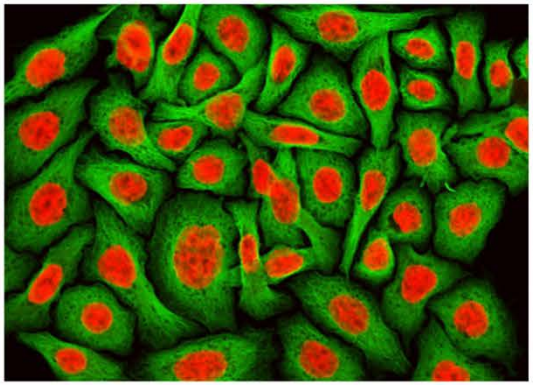


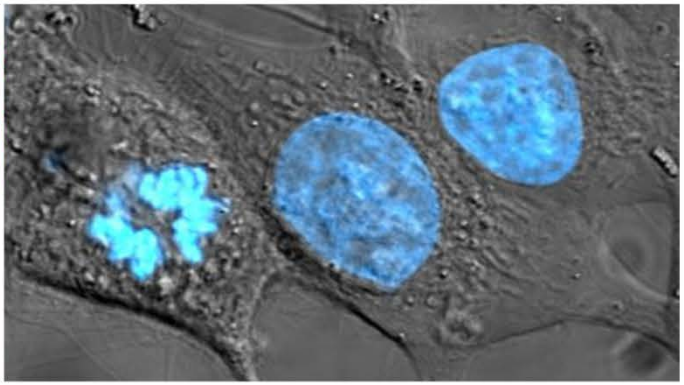












Advantages and drawbacks of baculovirus expression

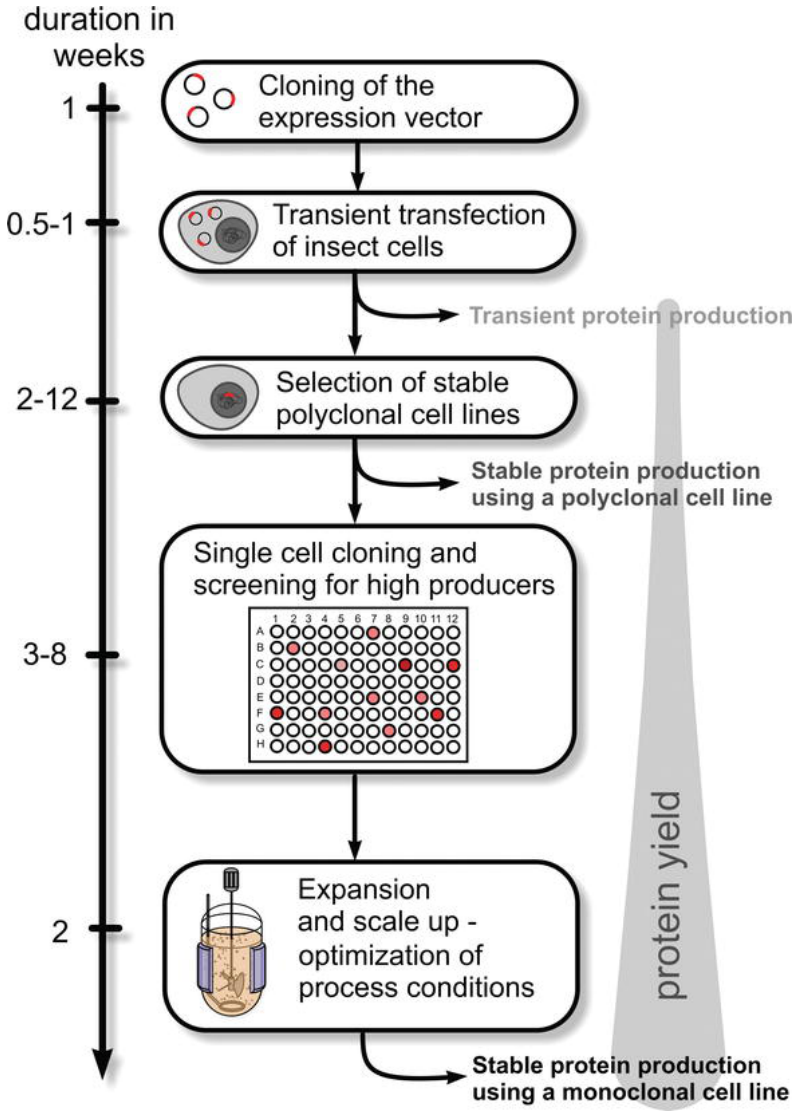
ADVANTAGES

- Recombinant protein has complete biological function, such as protein correct folding and disulfide bond
- Post-translation modification (glycosylation etc.)
- High level of expression, up to 50% of the total protein amount
- Accommodate large insert protein
- Simultaneously express multiple genes (multi-expression)
- In general, it is safer to use than mammalian virus, since it has limited host range and does not infect vertebrates

DRAWBACKS

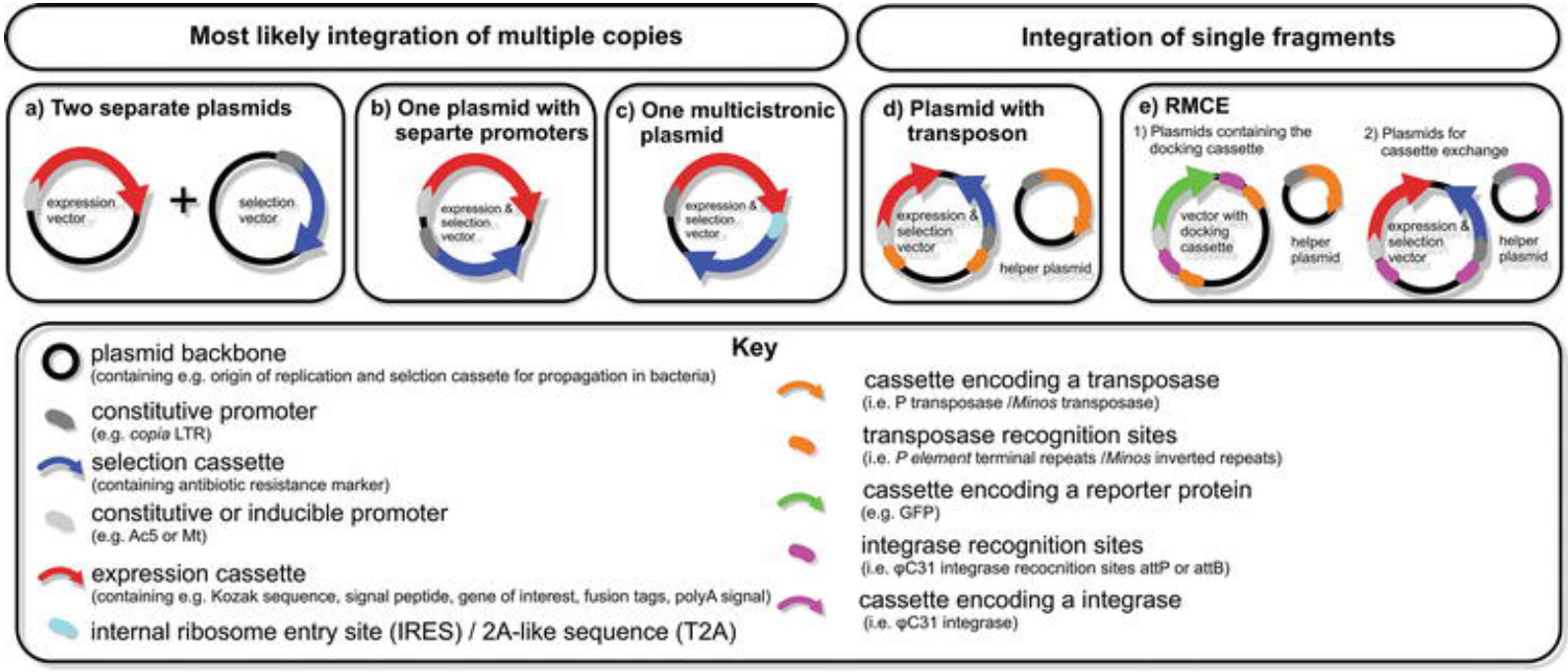
- The drawback is that exogenous protein expression is under control of the very late viral promoter, where the cells begin to die due to viral infection
- Insect expression system is normally used for production of membrane proteins, although the glycosylations may be different from those found in vertebrates

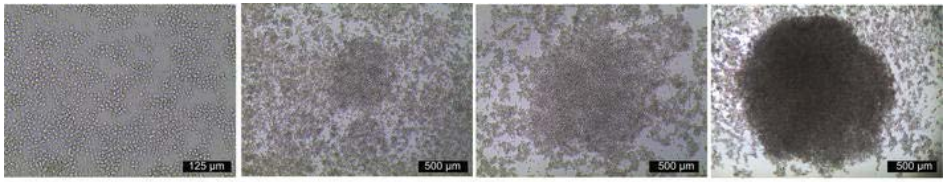
Virus-free expression in insect cells



- Overview of the general procedure to produce stably transformed *D. melanogaster* S2 cell lines for recombinant protein expression
- Protein expression can be initiated at different points, starting with transient expression immediately after transfection followed by stable expression in a polyclonal cell line and finally the selection of a highly productive monoclonal cell line

The different plasmid sets to generate stable S2 cell lines



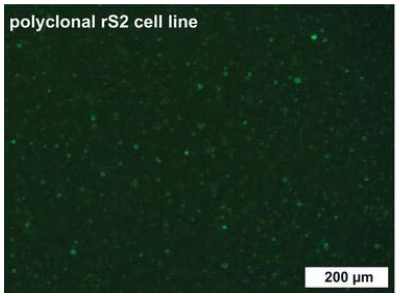


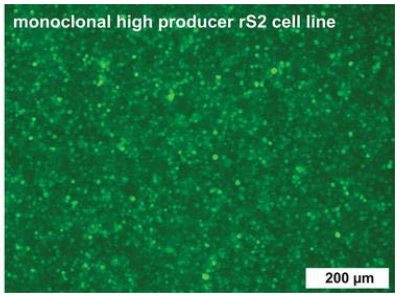
Limiting dilution cloning
Co-cultivation of untransfected feeder cells with approximatly one clone of the stable cell population in 96 well format

Start of selection
Addition of selection antibiotic to kill untransfected feeder cells and induce colony growth

Maintaining selection pressure
Feed or medium exchange to keep the cells under selective pressure and maintain sufficient nutrient supply

Pick and expansion
Most of the feeder cells are dead and macroscopically visible monoclonal colonies can be picked and used for subsequent upscaling under selection pressure





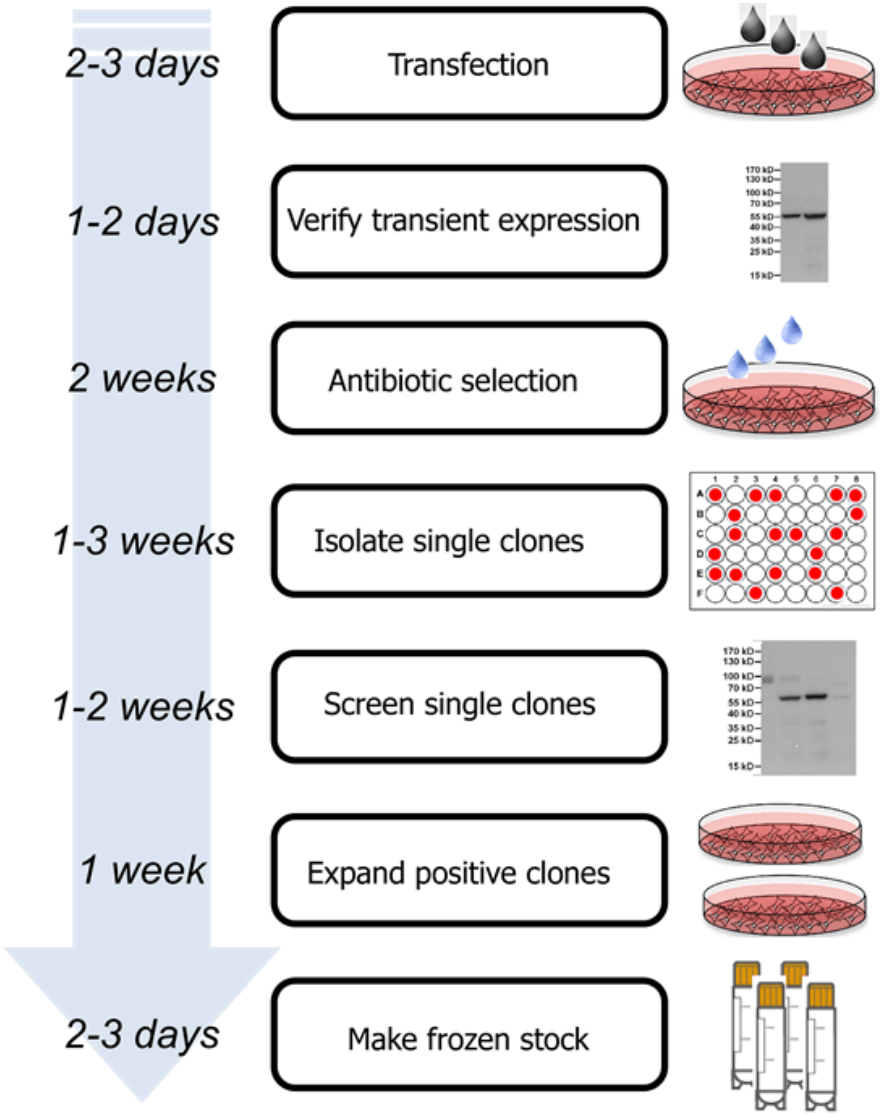
Expression in mammalian cells

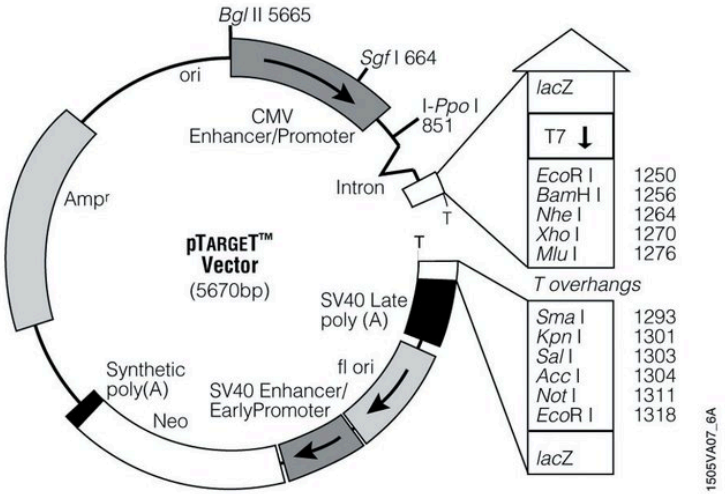
Concepts
Methods
Applications

Mammalian expression system

- Recombinant proteins expressed in mammalian cell commonly use plasmid transfection and viral vector infection. Stable cell line using plasmid transfection takes several weeks or even months, while the virus vector can quickly infect cells within a few days
- Depending on the temporal and spatial differences in protein expression, the expression system can be divided into transient, stable and induced expression systems
- Transient expression system refers that host cell cultures without selection pressure and exogenous vector gradually lost while cell division. The target protein expression duration is short. The advantage of transient expression system is simple and short experimental period
- Stable expression system means that the carrier DNA replicate and express long time stably in host cell. Due to the need of select resistance and pressure steps, stable expression is relatively time-consuming and laborious
- Induction expression system refers that the target gene begins to express when induced by foreign small molecules. The use of heterologous promoters, enhancers and amplifiable genetic markers can increase protein production
- The mammalian expression system has unique advantage in protein initiation signals, processing, secretion, glycosylation and is suitable for expressing intact macromolecules

Mammalian stable cell line generation





Expression in *Leishmania tarentolae*

Concepts
Methods
Applications

LEXSY - Eukaryotic protein expression in *L. tarentolae*

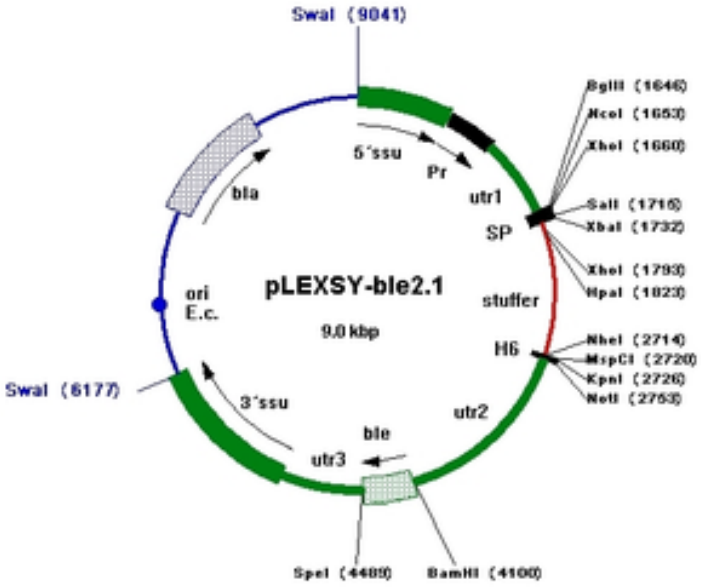
The unicellular kinetoplast protozoan *Leishmania tarentolae*, isolated from the Moorish gecko *Tarentola mauritanica*, not pathogenic to mammalians (Biosafety level 1) – was turned into the protein-producing host of eukaryotic protein expression system LEXSY:

- eukaryotic host as easy to handle as *E. coli*: no specific labware, no cell biology equipment required
- fully eukaryotic protein expression machinery with post-translational modifications, including glycosylation and disulfide bond formation
- shuttle vectors: cloning in *E. coli*, expression in LEXSY host
- constitutive or inducible, intracellular or secretory expression of target proteins
- stable expression strains for constant protein production

https://www.jenabioscience.com/lexsy-expression

Leishmania tarentolae: a little-known expression host

- *Leishmania tarentolae*, which can be easily cultured and genetically manipulated, is harmless (unless you are a gecko from Africa)
- It is closely related to other kinetoplastida, including important human pathogens such as *L. donovani* (causing kala azar), *Trypanosoma brucei* (sleeping sickness) and *T. cruzi* (Chagas disease)

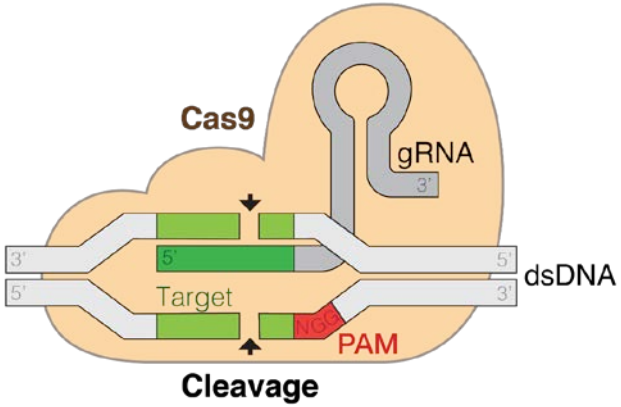


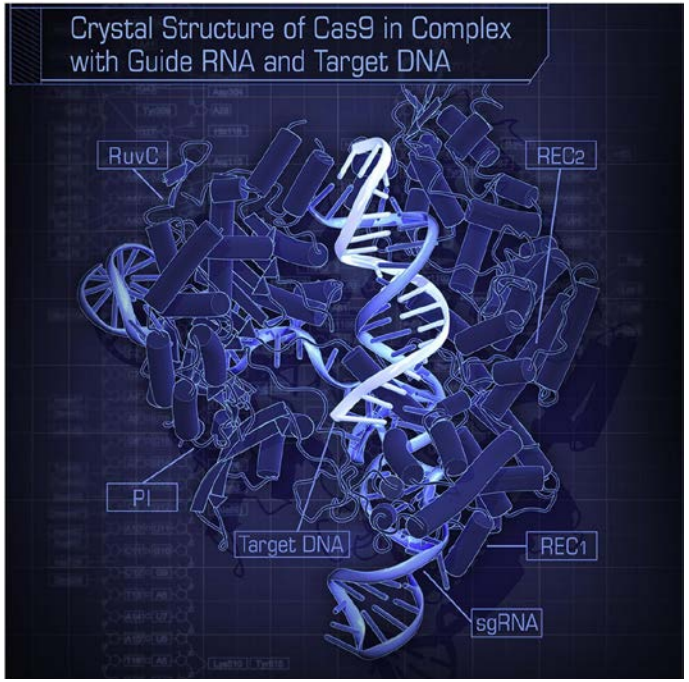
Emerging new expression tools

Concepts
Methods
Applications

The CRISPR/Cas9 system

- The **Cas9** (CRISPR associated protein 9) is a protein which plays a vital role in the immunological defense of bacteria against DNA viruses, and which is used in genetic engineering. Its main function is to cut DNA and therefore it can alter a cell's genome
- Structurally, Cas9 is an **RNA-guided DNA endonuclease enzyme** associated with CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) adaptive immunity system in *Streptococcus pyogenes*
- Cas9 performs this by unwinding foreign DNA and checking for sites complementary to the 20 bp spacer region of the guide RNA
- If the DNA substrate is complementary to the guide RNA, Cas9 cleaves the invading DNA

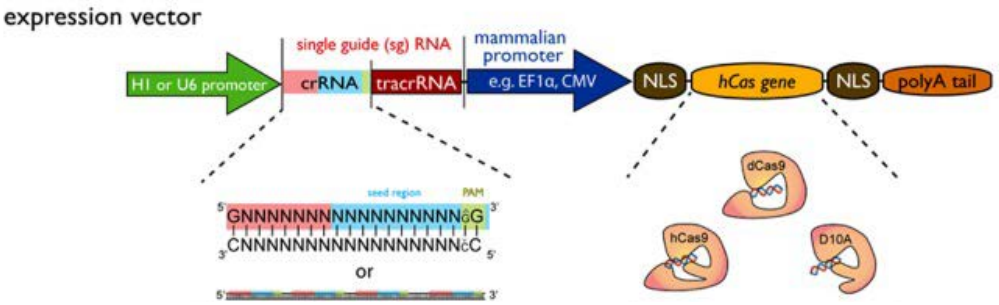


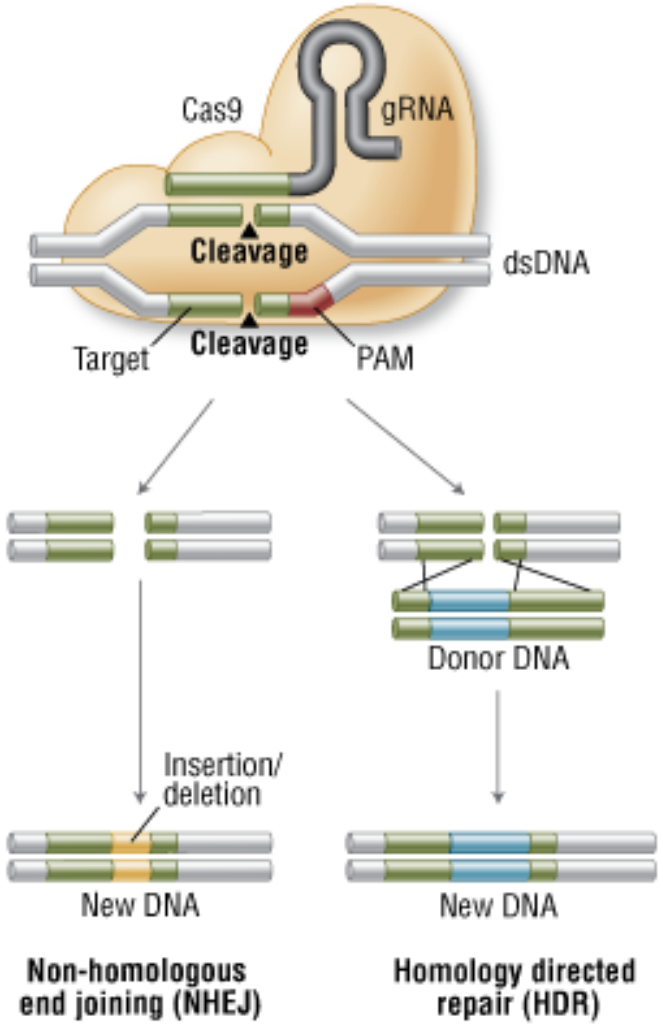


Crystal structure of *S. pyogenes* Cas9 in complex with sgRNA and its target DNA at 2.5 Å resolution, Nishimazu *et al.*, Cell 156: 935–49 (2014)

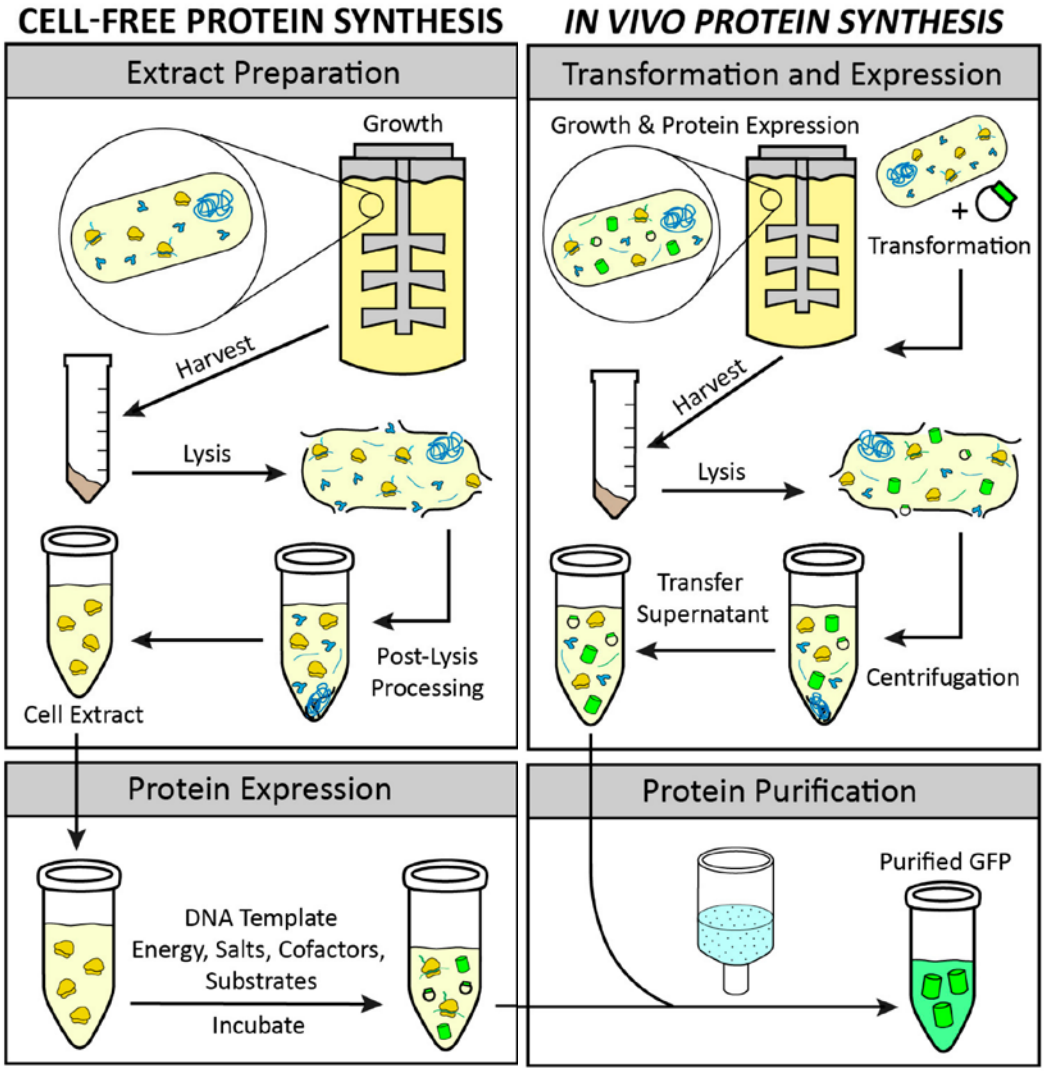
The CRISPR/Cas9 system: key elements

- Cas9 nuclease specifically cleaves double-stranded DNA activating double-strand break repair machinery
- In the absence of a homologous repair template non-homologous end joining can result in indels disrupting the target sequence
- Alternatively, precise mutations and knock-ins can be made by providing a homologous repair template and exploiting the homology directed repair pathway



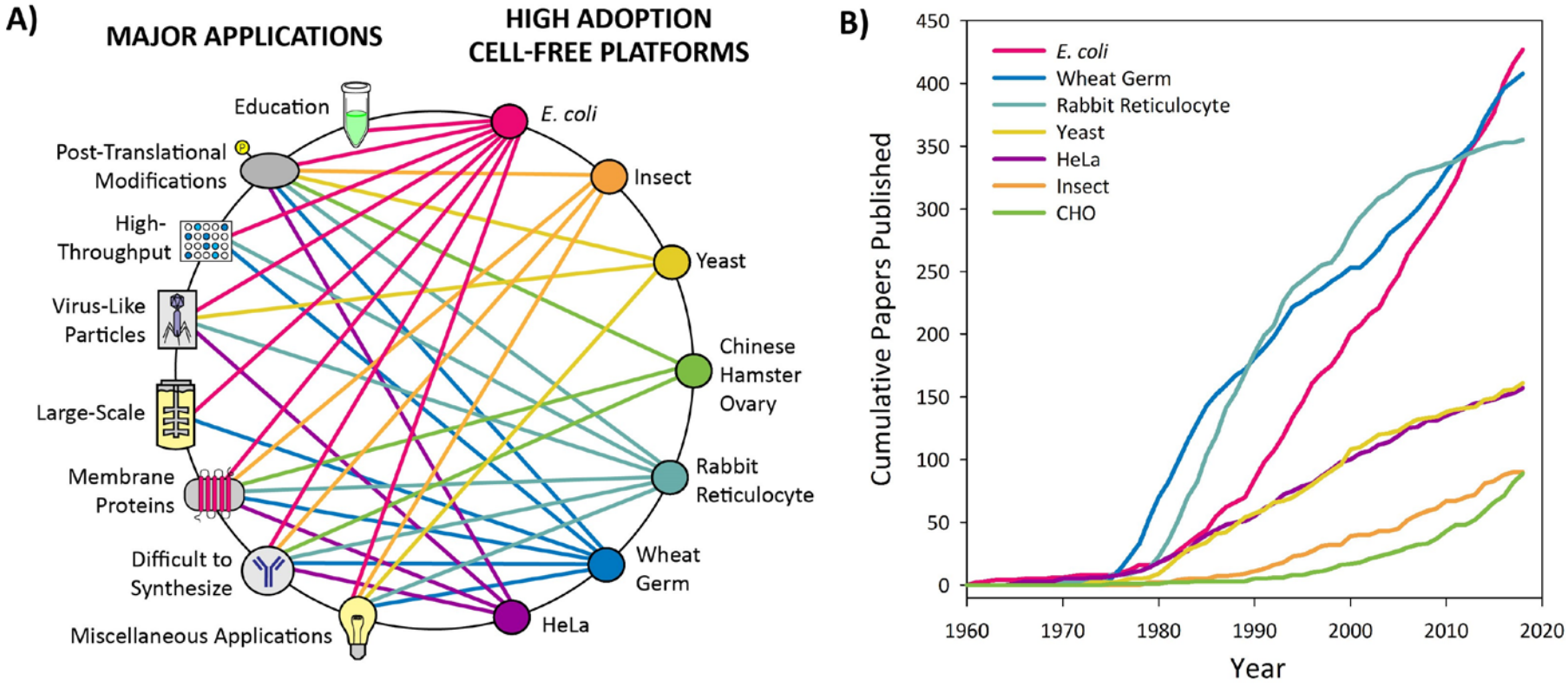


Cell-Free Protein Synthesis (CFPS)



- Cell-free protein synthesis (CFPS) is a platform technology that provides new opportunities for protein expression
- The advantages of CFPS over in vivo protein expression include its open system, the elimination of reliance on living cells, and the ability to focus all system energy on production of the protein of interest

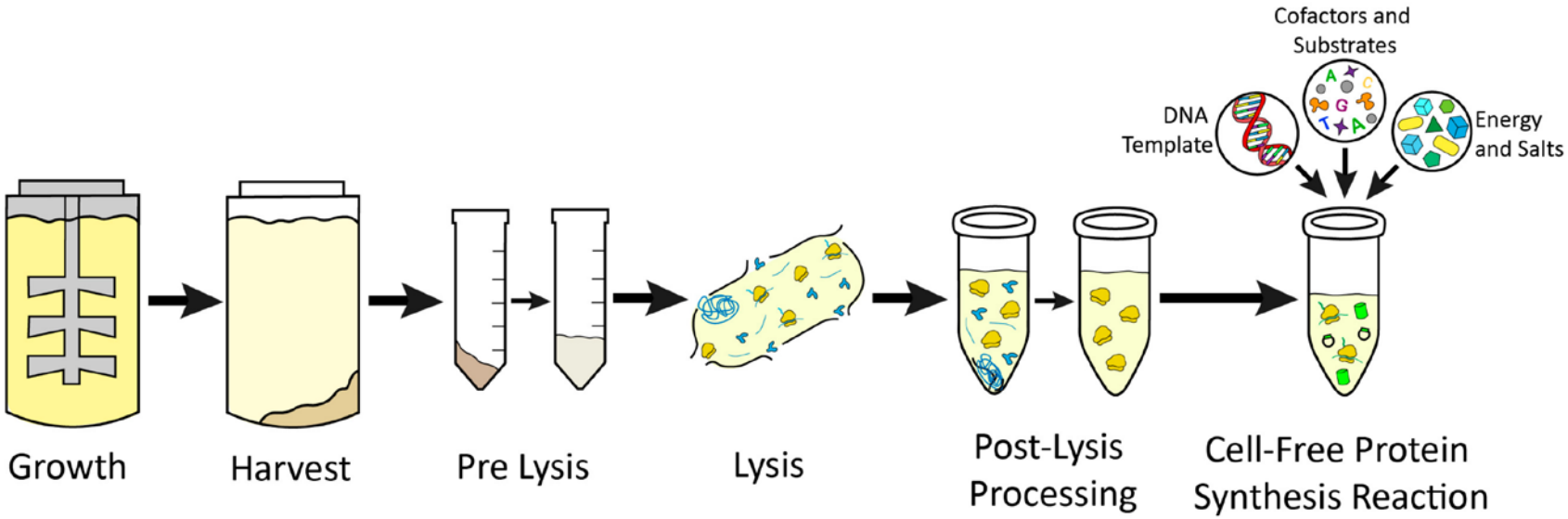
Cell-free platforms and their applications



(A) Web of the applications enabled by low adoption cell-free platforms. Connections shown are based on applications that have been published or that have been proposed in publications.

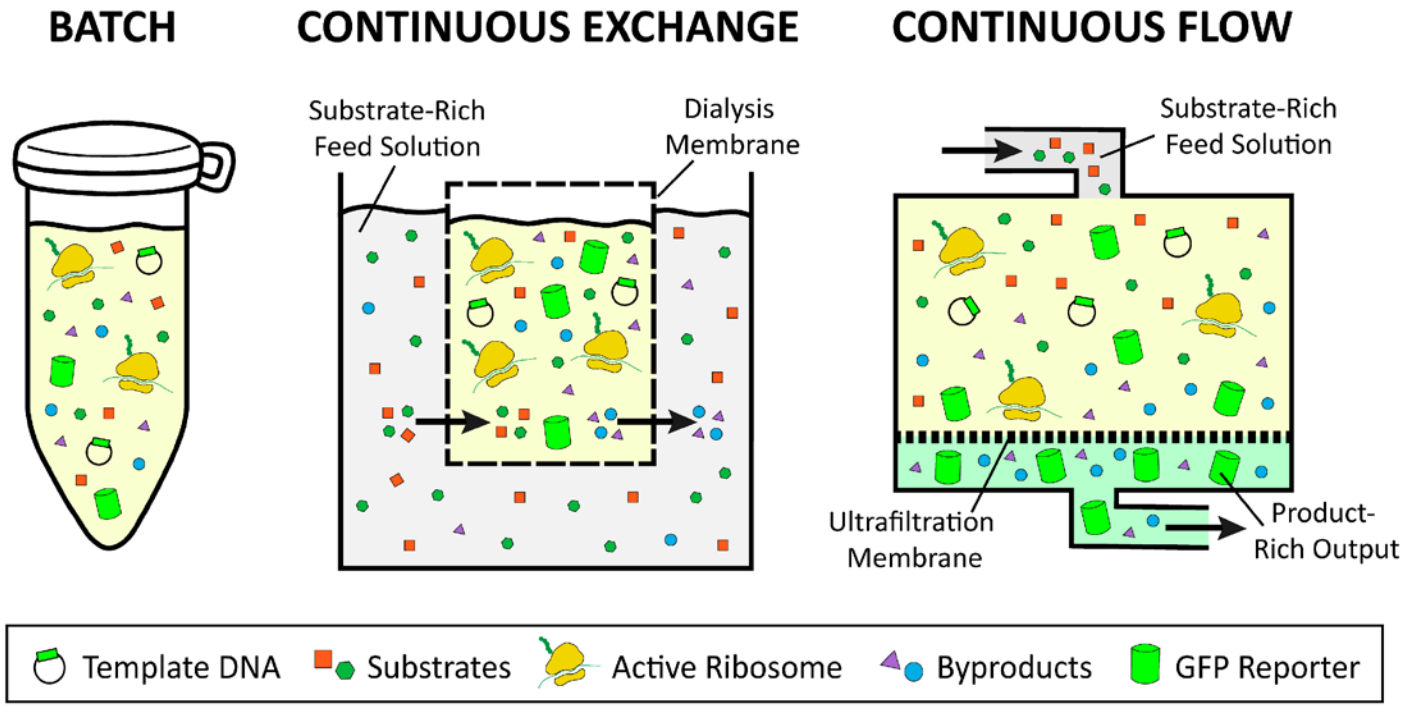
(B) Cumulative number of peer-reviewed publications over the last 60 years for cell-free platforms.

Preparation of cell-free extract and set-up of CFPS reactions



General workflow for preparation of cell-free extract and set up of CFPS reactions. A visualization from cell growth to the CFPS reaction is depicted above for a new user, highlighting the main steps involved.

Comparison of CFPS platforms

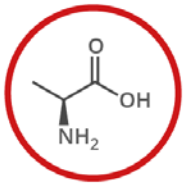


A. Batch reactions contain all the necessary reactants within a single reaction vessel.
B. Continuous exchange formats utilize a dialysis membrane that allows reactants to move into the reaction and byproducts to move out, while the protein of interest remains in the reaction compartment.
C. Continuous flow formats allow a feed solution to be continuously pumped into the reaction chamber while the protein of interest and other reaction byproducts are filtered out of the reaction.

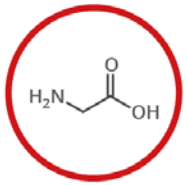
A GUIDE TO THE TWENTY COMMON AMINO ACIDS

AMINO ACIDS ARE THE BUILDING BLOCKS OF PROTEINS IN LIVING ORGANISMS. THERE ARE OVER 500 AMINO ACIDS FOUND IN NATURE - HOWEVER, THE HUMAN GENETIC CODE ONLY DIRECTLY ENCODES 20. 'ESSENTIAL' AMINO ACIDS MUST BE OBTAINED FROM THE DIET, WHILST NON-ESSENTIAL AMINO ACIDS CAN BE SYNTHESISED IN THE BODY.

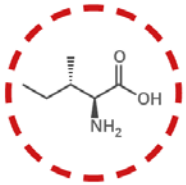
Chart Key: ALIPHATIC · AROMATIC · ACIDIC · BASIC · HYDROXYLIC · SULFUR-CONTAINING · AMIDIC · NON-ESSENTIAL · ESSENTIAL



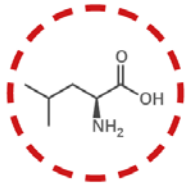
ALANINE A
Ala
GCT, GCC, GCA, GCG



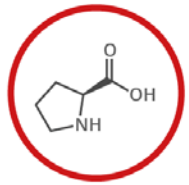
GLYCINE G
Gly
GGT, GGC, GGA, GGG



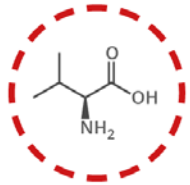
ISOLEUCINE I
Ile
ATT, ATC, ATA



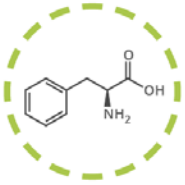
LEUCINE L
Leu
CTT, CTC, CTA, CTG, TTA, TTG



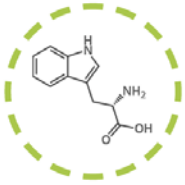
PROLINE P
Pro
CCT, CCC, CCA, CCG



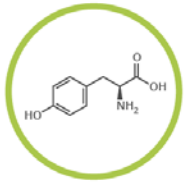
VALINE V
Val
GTT, GTC, GTA, GTG



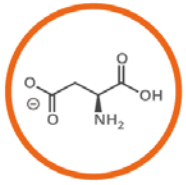
PHENYLALANINE F
Phe
TTT, TTC



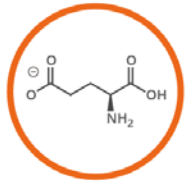
TRYPTOPHAN W
Trp
TGG



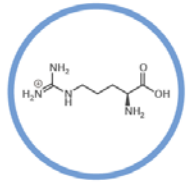
TYROSINE Y
Tyr
TAT, TAC



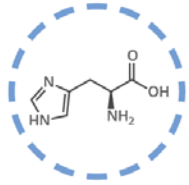
ASPARTIC ACID D
Asp
GAT, GAC



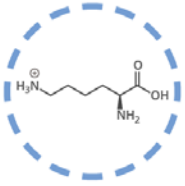
GLUTAMIC ACID E
Glu
GAA, GAG



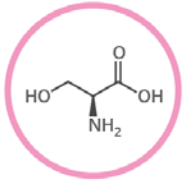
ARGININE R
Arg
CGT, CGC, CGA, CGG, AGA, AGG



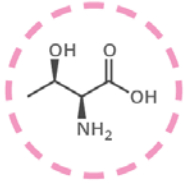
HISTIDINE H
His
CAT, CAC



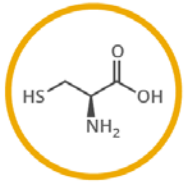
LYSINE K
Lys
AAA, AAG



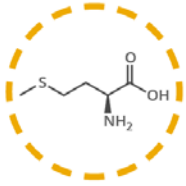
SERINE S
Ser
TCT, TCC, TCA, TCG, AGT, AGC



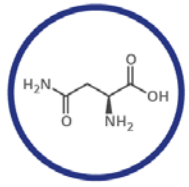
THREONINE T
Thr
ACT, ACC, ACA, ACG



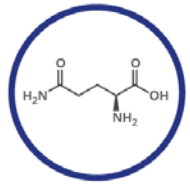
CYSTEINE C
Cys
TGT, TGC



METHIONINE M
Met
ATG



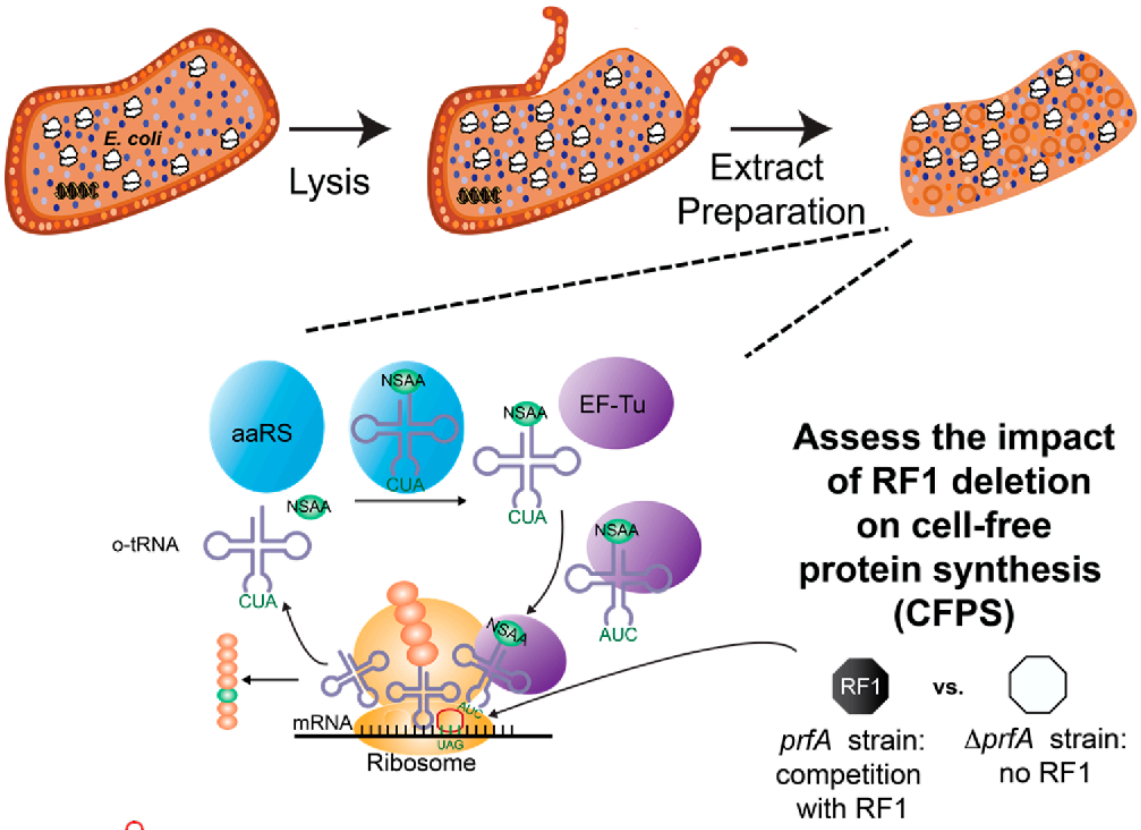
ASPARAGINE N
Asn
AAT, AAC

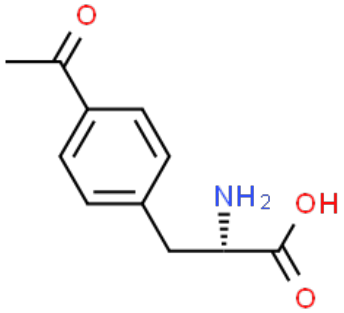


GLUTAMINE Q
Gln
CAA, CAG

Note: This chart only shows those amino acids for which the human genetic code directly codes for. Selenocysteine is often referred to as the 21st amino acid, but is encoded in a special manner. In some cases, distinguishing between asparagine/aspartic acid and glutamine/glutamic acid is difficult. In these cases, the codes asx (B) and glx (Z) are respectively used.

CFPS reaction incorporating nonstandard amino acids





p-acetyl-L-phenylalanine

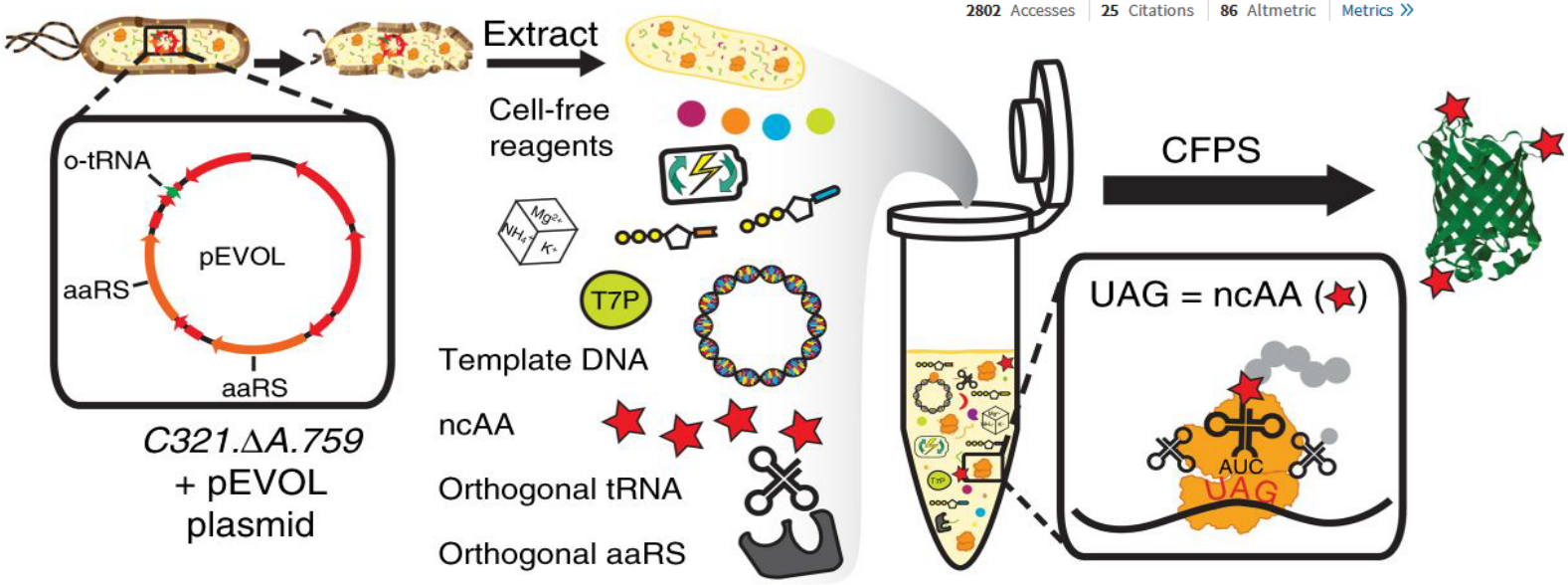
- Scheme of cell-free protein synthesis reaction incorporating nonstandard amino acid to investigate the effect of RF1
- Cell extracts containing transcription and translation machinery are prepared from rEc.E13 or rEc.E13.ΔprfA strains
- Plasmid DNA template of sfGFP containing single or multiple **amber codon** sites, **orthogonal tRNA/aaRS**, **NSAA**, T7 RNA polymerase, and other cofactors are added as necessary to activate the cell-free protein synthesis (CFPS) reaction

CFPS from extracts of a genomically recoded organism

nature COMMUNICATIONS

Rey W. Martin, Benjamin J. Des Soye, Yong-Chan Kwon, Jennifer Kay, Roderick G. Davis, Paul M. Thomas, Natalia I. Majewska, Cindy X. Chen, Ryan D. Marcum, Mary Grace Weiss, Ashleigh E. Stoddart, Miriam Amiram, Arnaz K. Ranji Charna, Jaymin R. Patel, Farren J. Isaacs, Neil L. Kelleher, Seok Hoon Hong & Michael C. Jewett

Nature Communications 9, Article number: 1203 (2018) | Download Citation
2802 Accesses | 25 Citations | 86 Altmetric | Metrics »



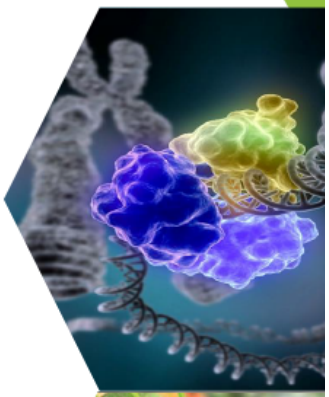
- Schematic of the production and utilization of crude extract from genomically recoded organisms with plasmid overexpression of orthogonal translation components for cell-free protein synthesis
- CFPS reactions are supplemented with the necessary substrates (e.g., amino acids, NTPs, etc.) required for in vitro transcription and translation as well as purified orthogonal translation system (OTS) components to help increase the ncAA incorporation efficiency
- aaRS, aminoacyl tRNA synthetase; ncAA, non-canonical amino acid; T7P, T7 RNA polymerase; UAG, amber codon

New course in English available!

Bi9690en

Synthetic Biology

Dr. Karel Říha
karel.riha@ceitec.muni.cz
Dr. Martin Marek
marint.marek@recetox.muni.cz



Synthetic biology is a new scientific discipline that builds on advances in molecular biology, genetic and protein engineering, systems biology and bioinformatics to re-design existing biological systems for applications in biotechnology and medicine. In this course, students will grasp concepts in synthetic biology, become familiar with basic methodological approaches and learn about its applications in science and technology.

Selected topics:

- Basic concepts of engineering in biology
- From genetic engineering to synthetic genomes
- Protein engineering and design
- Expanding building bricks of life
- From proteins to nanomachines
- Metabolic engineering, artificial organelles
- Future trends in synthetic biology and ethical issues

CEITEC

MUNI LOSCHMIDT LABORATORIES

TUESDAYS
START 21/09/2021
16:00–18:00
Room 333, Building B11
University Campus, Brno

LOSCHMIDT LABORATORIES

https://loschmidt.chemi.muni.cz/peg/

Molecular Biotechnology Practicals

Kdy: Úterý 11.10. 2022
od 10:00
od 14:00

Kde: D36-308

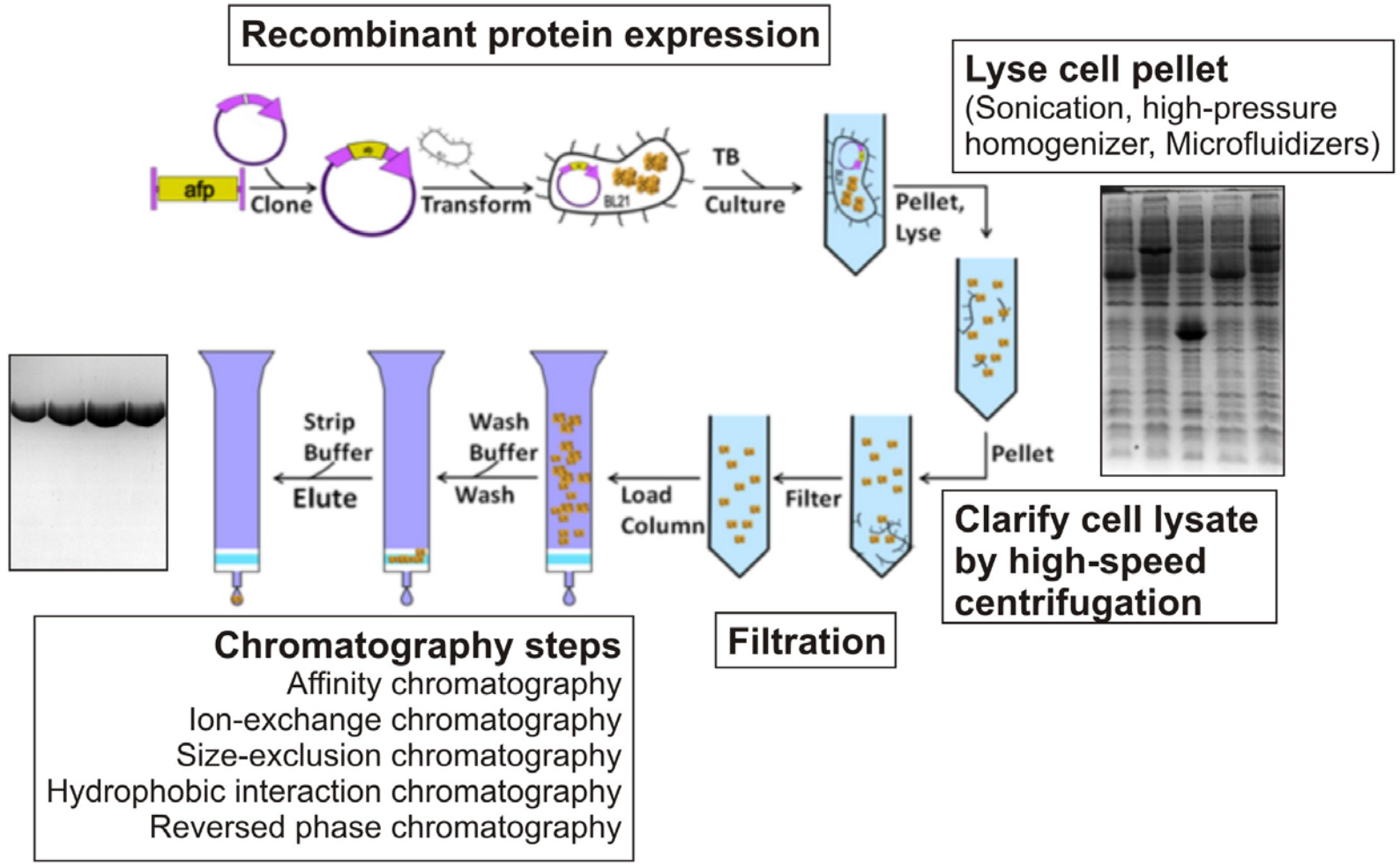
Questions

Protein purification methods

Concepts
Methods
Applications

Recombinant protein purification: step by step

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

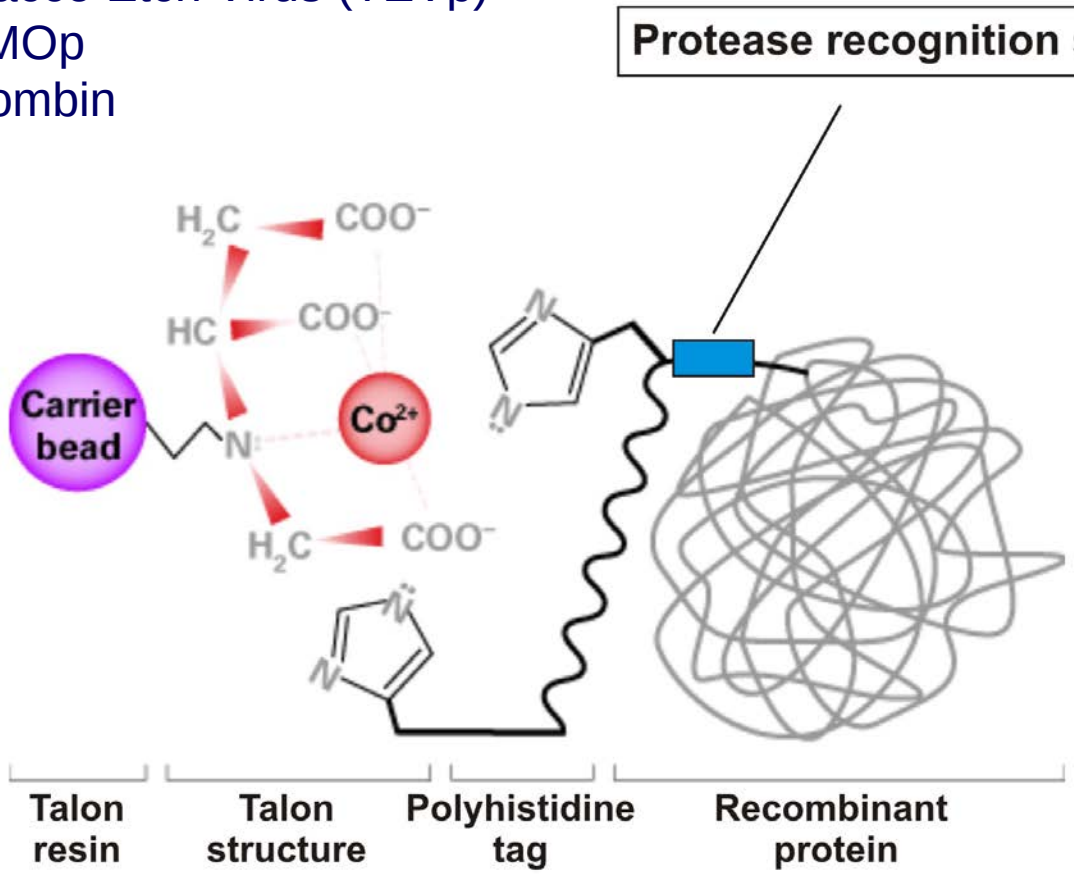
Fusion tags can improve protein expression, stability, resistance to proteolytic degradation and solubility.

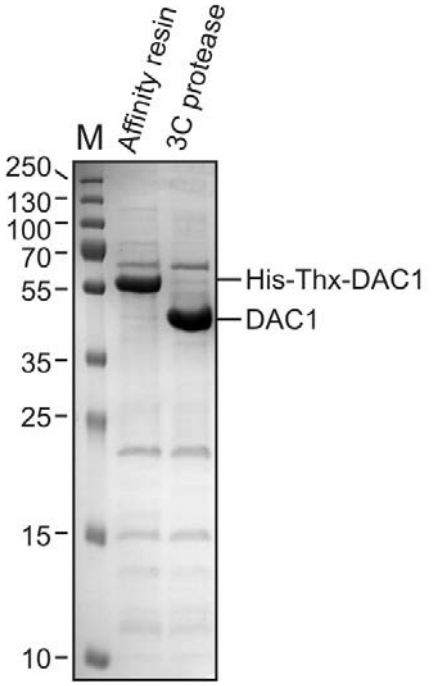
Fusion tag	Function	Size (kDa)	Description
Polyhistidine (e.g. 6xHis, 10xHis)	Affinity	1-2	The most commonly used affinity tag, binds to metal ions
Strep-tag II	Affinity	1	High affinity for engineered streptavidin
Thioredoxin (Trx)	Solubility	12	Aids in refolding proteins that require a reducing environment
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



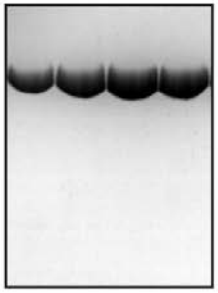


Column chromatography instrumentation

Automated chromatography systems

Chromatography steps

Affinity chromatography
Ion-exchange chromatography
Size-exclusion chromatography



Protein characterization: an aggregation problem

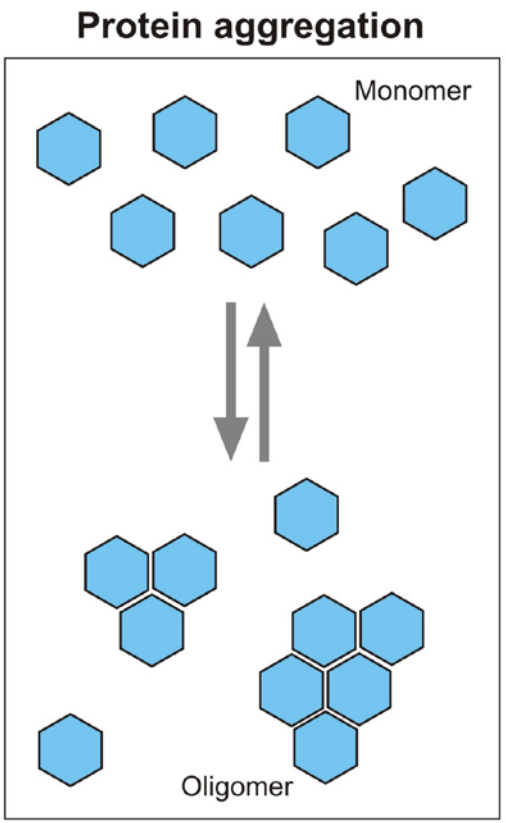
A key challenge in recombinant protein production is to maintain and store the target protein in a soluble and stable form. Protein aggregation can compromise protein function and thus it is necessary to overcome this challenge to generate functionally active protein.

Detection of protein aggregation

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

Troubleshooting

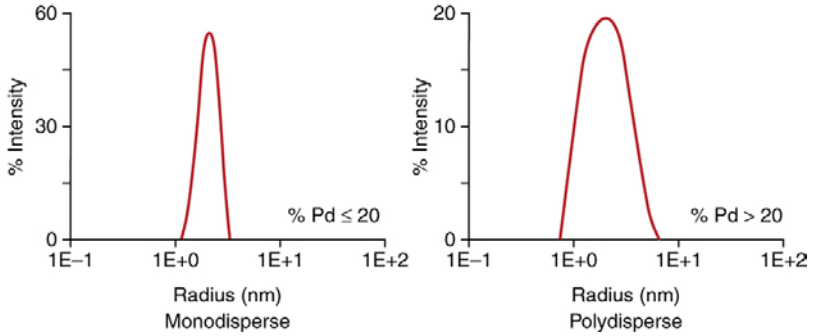
- Culture conditions (e.g. reducing temperature)
- Buffer composition (ionic strenght, pH, reducing agents)
- Presence co-factors (Acetyl-CoA, metal ions)
- Fusion tags (Trx, MBP, SUMO)
- Minimising sample handling
- Avoiding time delays between purification steps
- Performing purification steps at 4°C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



Protein quality control (QC) analyses

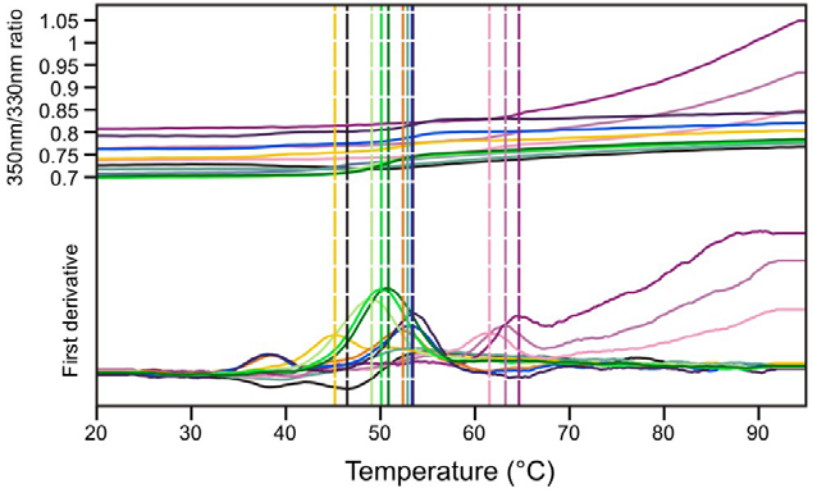
High purity and homogeneity of the protein sample are crucial for the downstream processes to be successful.

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

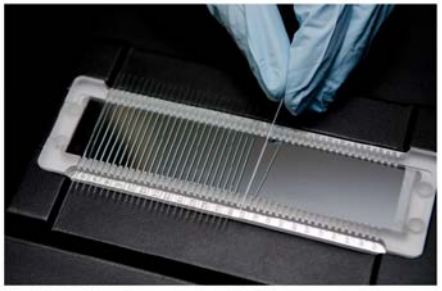


Identification of different oligomeric forms or aggregates, which are preventing crystallization

- 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)

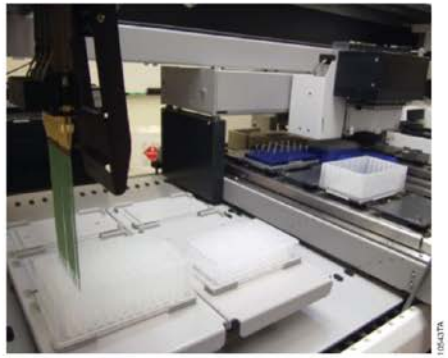
Success stories

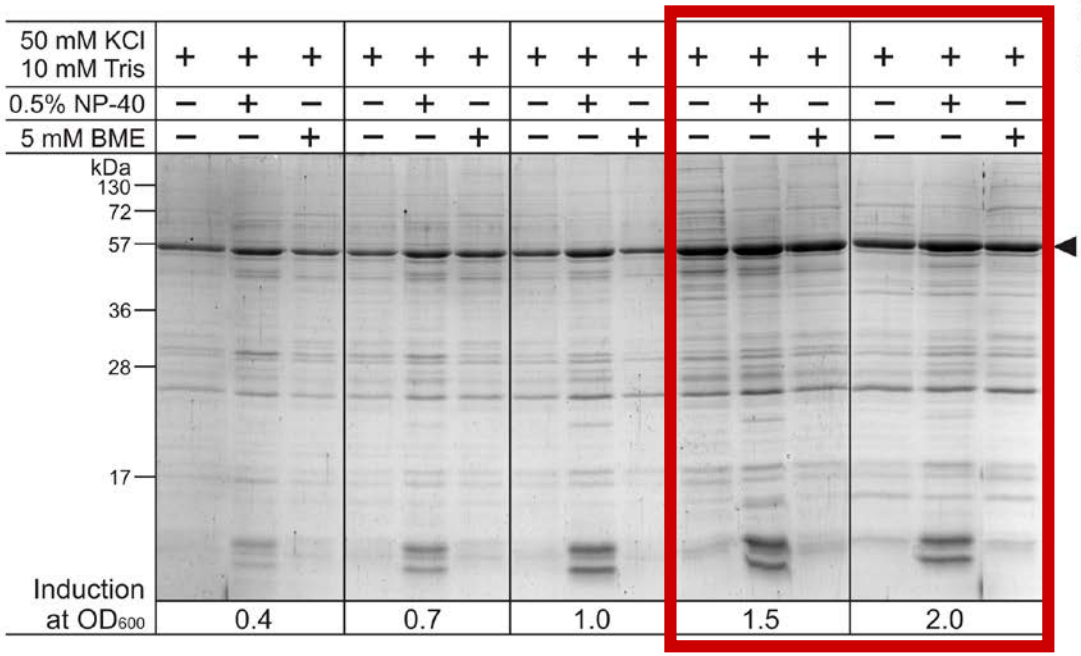
Development of expression and purification protocol for *Schistosoma mansoni* HDAC8: mini-scale tests

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

96-well deep plate for purification

Tecan robot



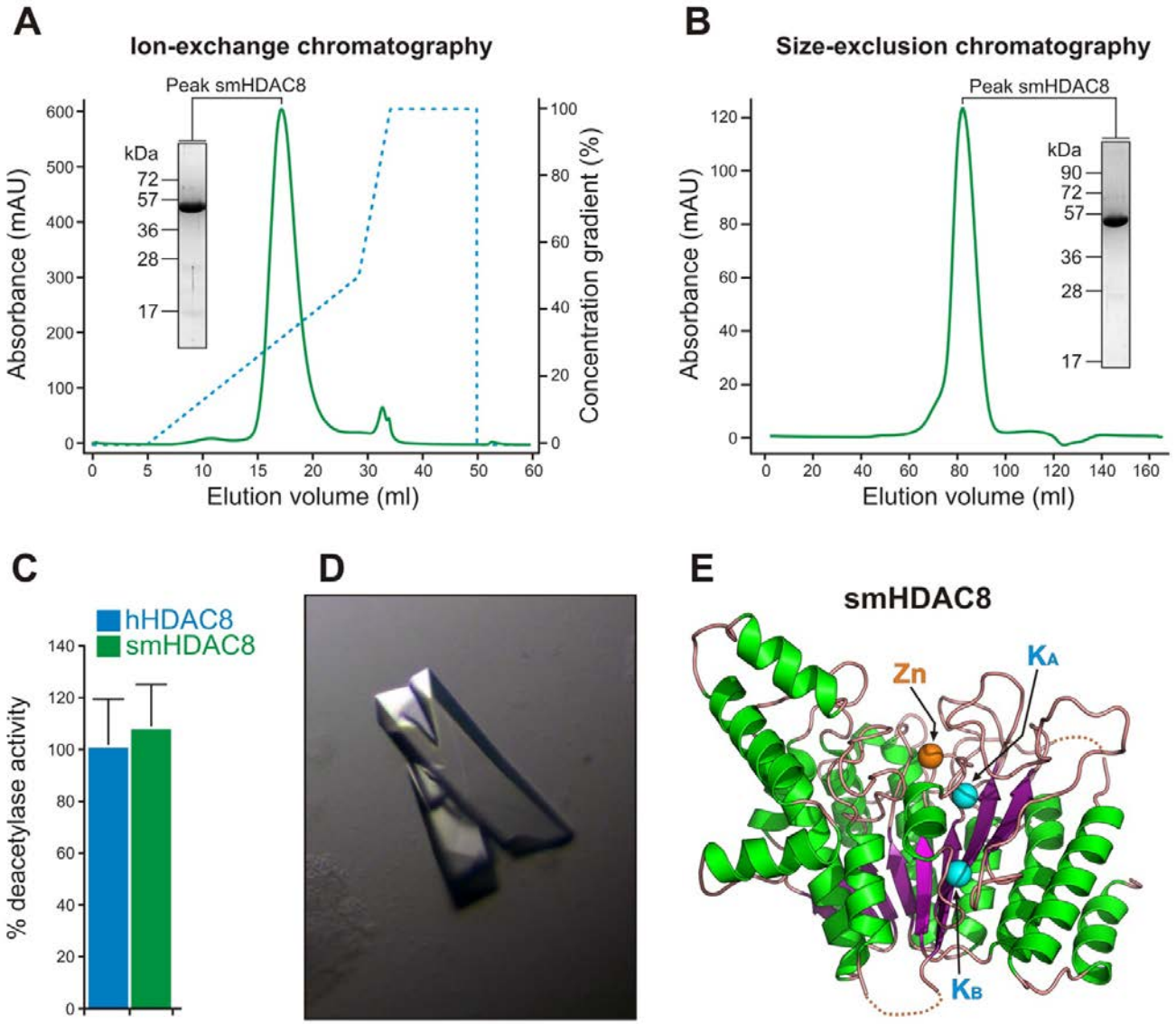


Semi-automated affinity purification

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

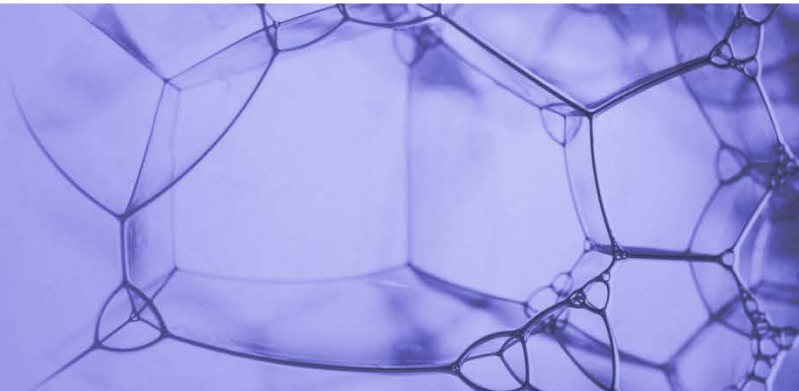
Recombinant DNA Technology Market Worth $844.6 Billion by 2025: Grand View Research, Inc.

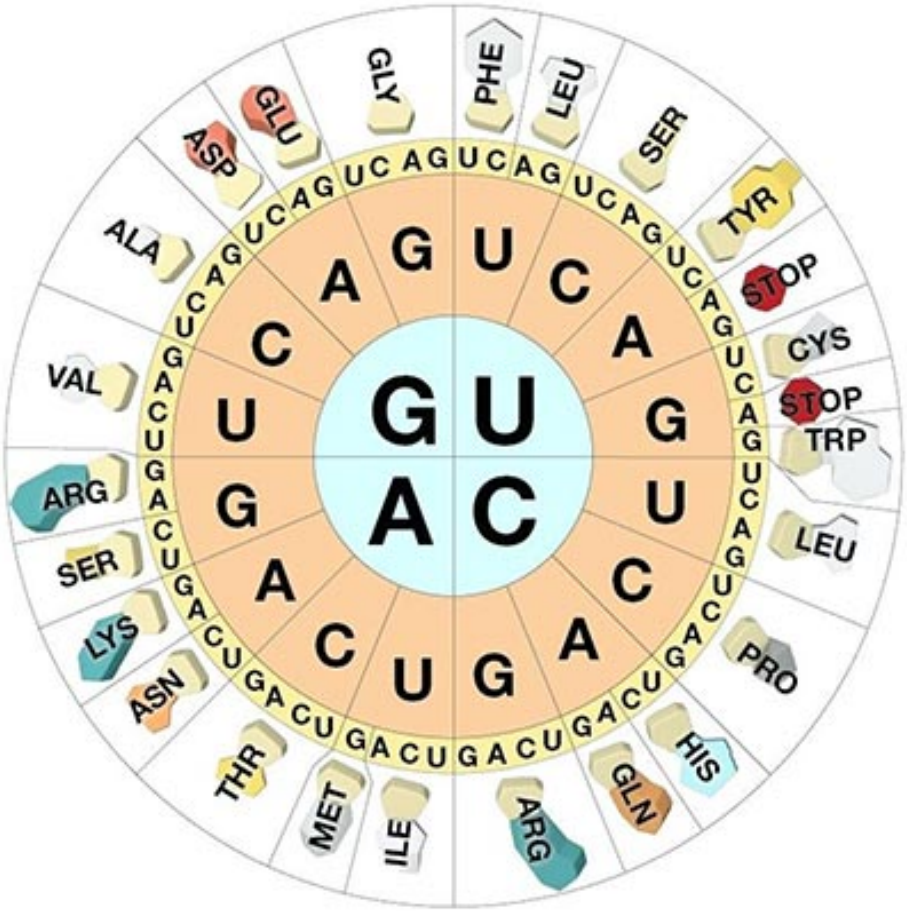
GRAND VIEW RESEARCH

Questions

Dr. Martin Marek
Loschmidt Laboratories
Faculty of Science, MUNI
Kamenice 5, bld. A13, room 332
martin.marek@recetox.muni.cz

LOSCHMIDT
LABORATORIES



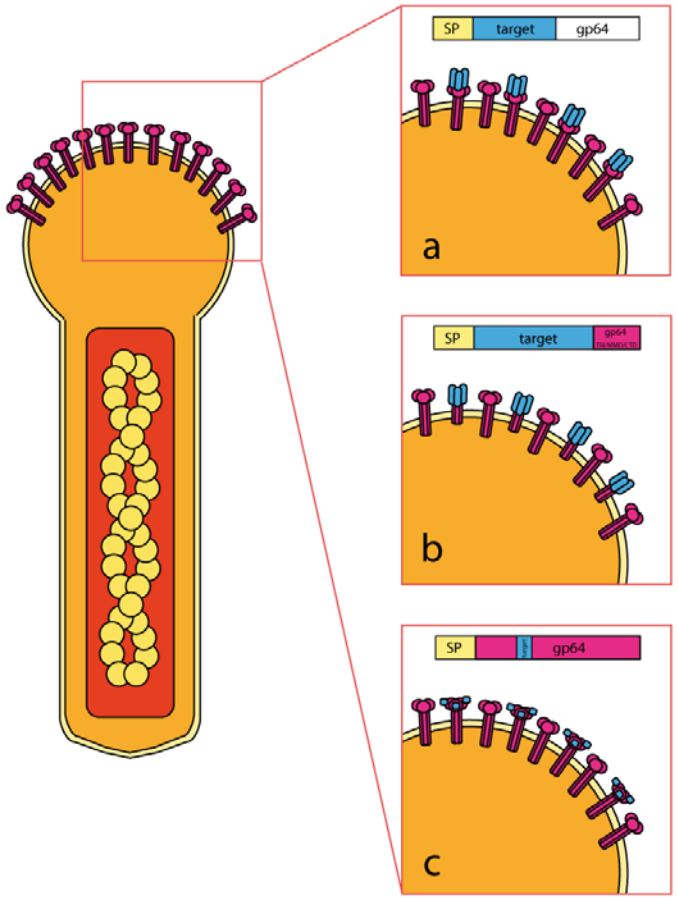


Second Letter

First Letter	T	C	A	G	Third Letter
T	TTT Phe TTC Phe TTA Leu TTG Leu	TCT Ser TCC Ser TCA Ser TCG Ser	TAT Tyr TAC Tyr TAA Stop TAG Stop	TGT Cys TGC Cys TGA Stop TGG Trp	T C A G
C	CTT Leu CTC Leu CTA Leu CTG Leu	CCT Pro CCC Pro CCA Pro CCG Pro	CAT His CAC His CAA Gln CAG Gln	CGT Arg CGC Arg CGA Arg CGG Arg	T C A G
A	ATT Ile ATC Ile ATA Ile ATG Met	ACT Thr ACC Thr ACA Thr ACG Thr	AAT Asn AAC Asn AAA Lys AAG Lys	AGT Ser AGC Ser AGA Arg AGG Arg	T C A G
G	GTT Val GTC Val GTA Val GTG Val	GCT Ala GCC Ala GCA Ala GCG Ala	GAT Asp GAC Asp GAA Glu GAG Glu	GGT Gly GGC Gly GGA Gly GGG Gly	T C A G

Baculovirus display system

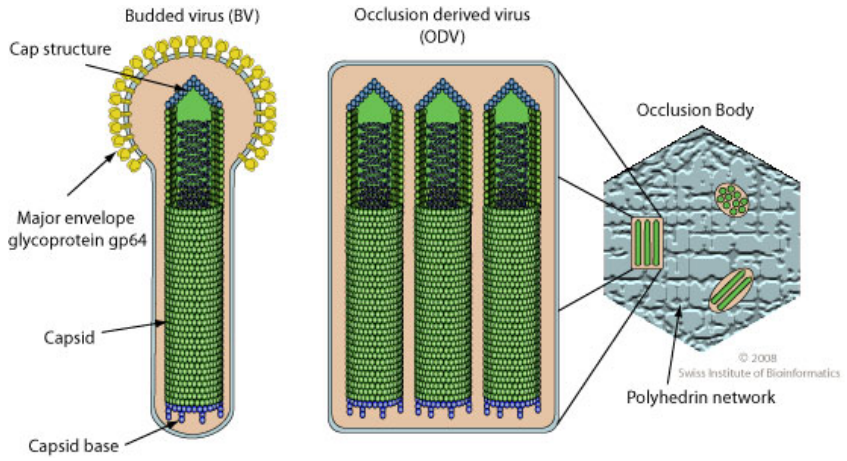
- Baculovirus is a large DNA insect-infecting virus
- Baculovirus surface glycoprotein Gp64 is expressed early and late in the infection of an insect cell
- It is a 64 kDa protein which forms trimers and locates in the BV envelope with a polarized distribution
- As Gp64 is a transmembrane protein that exposes an outer domain, it can be used to display a selected protein on the BV surface
- A chimeric Gp64 can be constructed to contain the protein of interest allowing it to be incorporated in the BV structure upon infection of insect cells



The MultiBac: a factory for synthetic virus-like particles

- A plasmid module comprising expression cassettes for the capsid-forming influenza H1N1 M1 protein (colored in grey) and a fluorescent protein marker, mCherry (colored in red), was introduced into the MultiBac baculoviral genome by Cre recombinase enzyme mediated plasmid fusion into the LoxP site (circle filled in red, gradient)
- Co-expression of HA, NA and M1 yields synthetic influenza virus-like particles (VLPs) resembling live influenza virus

Baculovirus biology

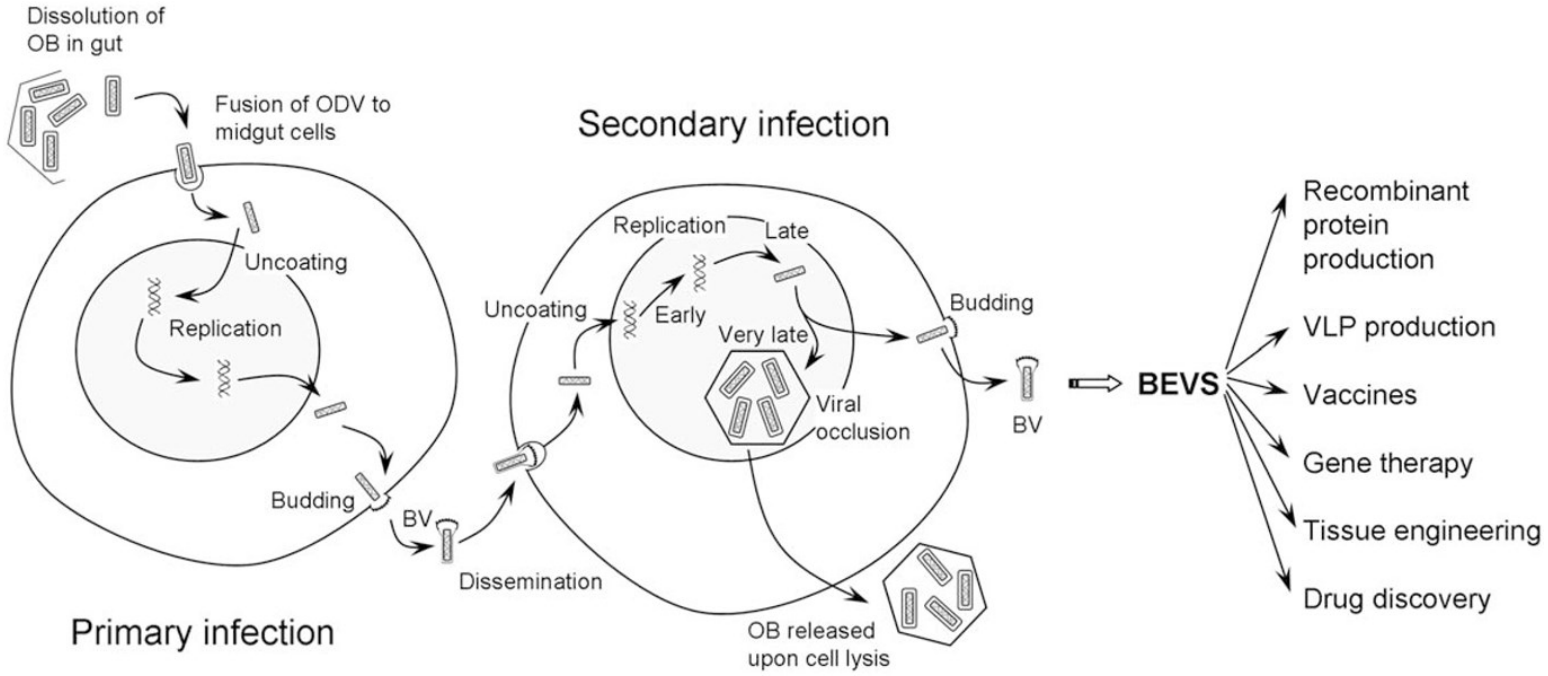


- Extracellular enveloped baculovirus virions can be found in two forms: OV (occluded virus) and BV (budded virus). The nucleocapsid is about 21 nm x 260 nm.
- Circular dsDNA genome , 80-180 kb in length, encoding for 100 to 180 proteins.

BACULOVIRUS INFECTION CYCLE

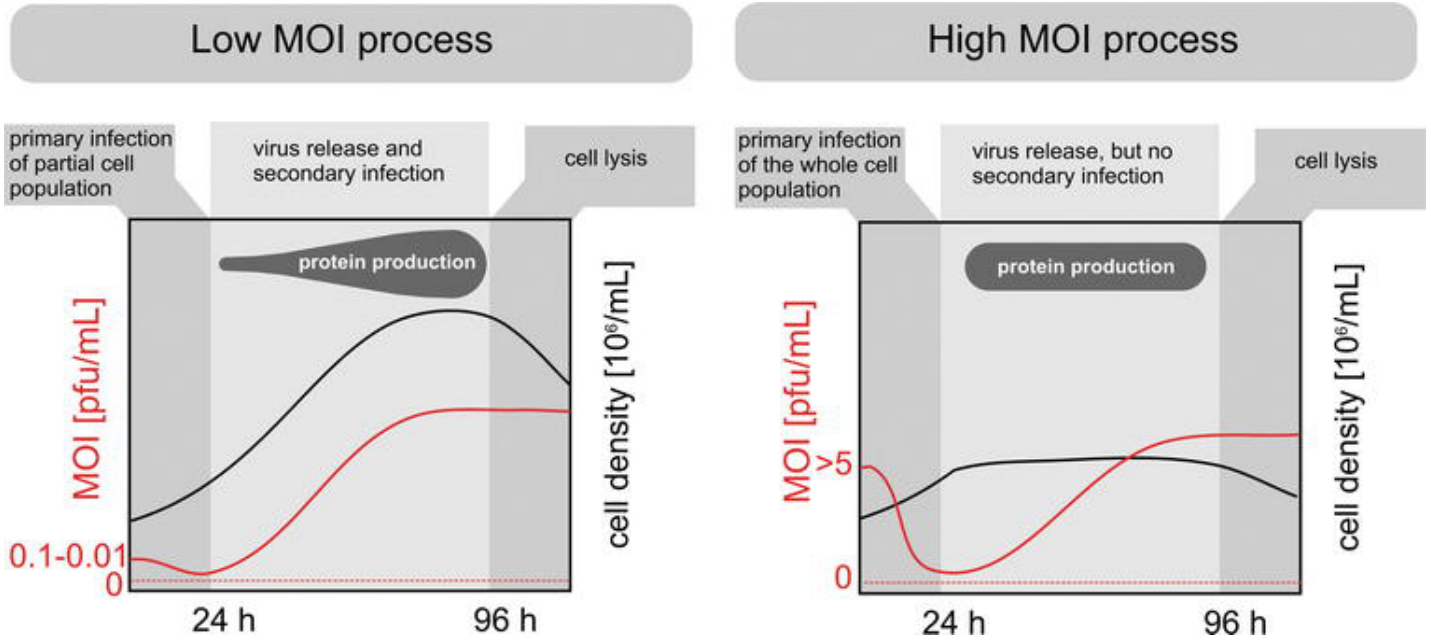
- Attachment of the viral glycoproteins to host receptors mediates endocytosis of the virus into the host cell.
- Fusion with the plasma membrane.
- The DNA genome is released into the host nucleus.
- Immediate early phase: host RNA polymerase transcribes viral genes involved in the regulation of the replication cascade, prevention of host responses and viral DNA synthesis.
- Late phase: The virally encoded RNA polymerase expresses late genes.
- Replication of the genome by rolling circle in nuclear viral factories.
- Nucleocapsids are formed which can either bud out through the cellular membrane and disseminate the infection or be occluded for horizontal transmission.
- Occlusion phase: the virus becomes occluded in the protein polyhedrin and the polyhedral envelope (calyx) is produced.
- Lysis of the cell releases the occluded virus.

The replication cycle of baculovirus



- Baculovirus is a member of the insect-borne virus family and was originally used as a biopesticide with a large double-stranded circular DNA genome. The virus is enveloped, rod-shaped particles, ranging from 30 to 60 nm in diameter and from 250 to 300 nm in length.
- The natural hosts for these viruses are insects, and Autographa californica (AcMNPV) is the most studied member of the family, which was developed into a recombinant baculovirus vector and is still under application today.
- AcMNPV and other baculoviruses are able to generate ‘‘polyhedra’’ or ‘‘occlusion bodies’’ in the infected cells of insects. In the late phase of infection, dozens of polyhedra are formed in the host cell nuclei, which accommodate progeny virions encased by a protective paracrystalline array comprised of virus-encoded protein-polyhedrin.

A comparison of BEVS processes with low and high MOI



How can baculoviruses express very late genes at such high levels?

There is no evidence that the baculovirus RNA polymerase has an intrinsic ability for high level gene expression. In fact, it may not bind to the very late promoter region with high levels of tenacity as reflected in the fact that the footprint of the polymerase on genomic DNA has never been reported despite an extensive effort by at least one laboratory.

It is likely that high levels of gene expression are influenced by several features of baculovirus biology. These include: i) the amplification of genes by DNA replication; ii) the shutoff of most late transcription, possibly by DNA binding proteins that coat the DNA and thereby make RNA polymerase available for very late transcription; iii) the efficiency of the late polymerase and VLF-1 in recognizing and initiating from very late promoter elements; iv) the efficiency of LEF-4 in capping the mRNA a possible role for LEF-2 and PK-1. As mentioned above, the 5' untranslated region of p10 mRNA appears to be capable of facilitating cap-independent translation, which may reduce the reliance of these transcripts on LEF-4 activity. Other factors that might enhance translation of very late expressed mRNAs have not been identified.

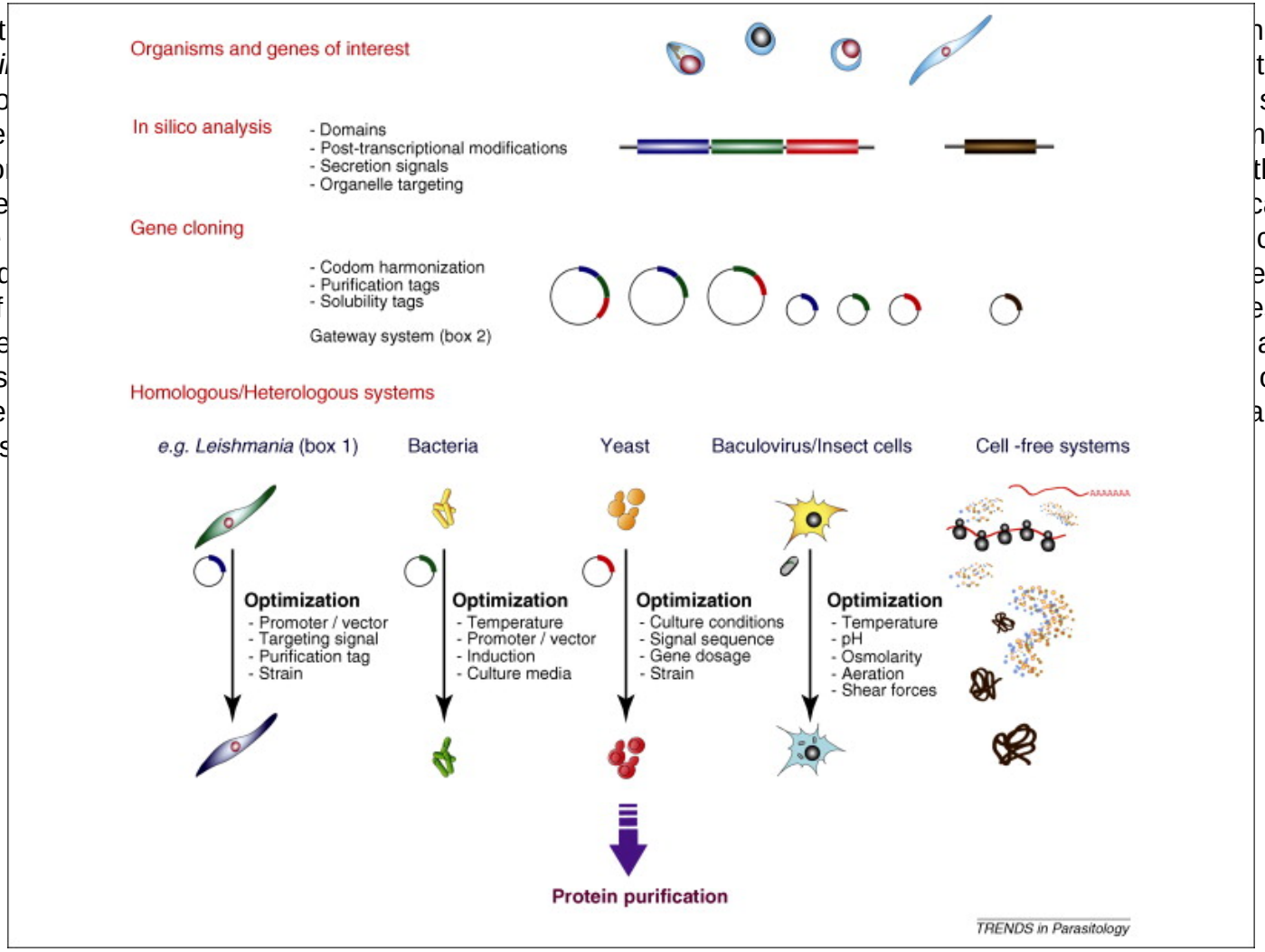
Baculovirus-mediated production of AAV vectors

Gene transfer and gene therapy are powerful approaches for many biological research applications and promising avenues for the treatment of many genetic or cancer diseases. The most efficient gene transfer tools are currently derived from viruses. Among them, the recombinant **adeno-associated viruses (AAVs)** are vectors of choice for many fundamental and therapeutic applications. The increasing number of clinical trials involving AAVs demonstrates the need to implement production and purification processes to meet the quantitative and qualitative demands of regulatory agencies for the use of these vectors in clinical trials. In this context, the rise of production levels on an industrial scale appeared essential. The introduction, in 2002, of an AAV process using a baculovirus expression vector system (BEVS) has circumvented this technological lock. The advantage of BEVS in expanding the AAV production in insect cells has been to switch the process to bioreactor systems, which are the ideal equipment for scaling up. We describe here a method for producing AAV vectors using the BEVS which can be easily used by research laboratories wishing to overcome the difficulties associated with the scaling up of production levels. The method provides sufficient quantities of AAV vectors to initiate preclinical projects in large animal models or for research projects where a single batch of vectors will consolidate the repeatability and reproducibility of in vitro and especially in vivo experimental approaches.

Methods Mol. Biol. 2019; 1937: 91-99. doi: 10.1007/978-1-4939-9065-8_5

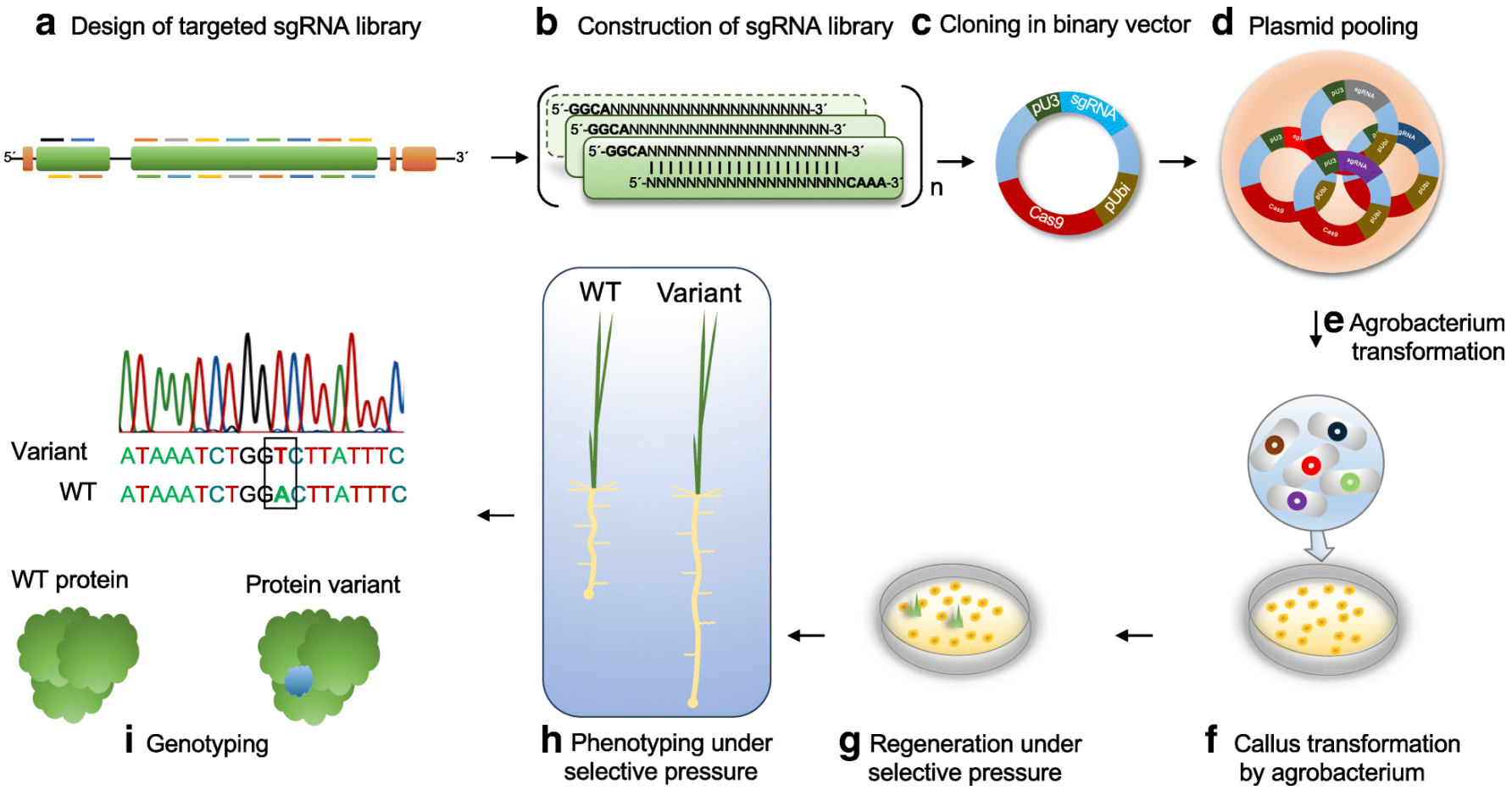
LEXSY - Eukaryotic protein expression in *L. tarentolae*

Most frequent … ne of interest, an extensive *i*… targeting and post-translatio… selection of the cloning ve… noter, into a suitable vecto… that can be grown in large… cale needed precludes the … clude removing prec… especting the boundaries of … esidues at the N- and/or C-te… antigenicity or the bacterial s… organism or cell-free syste… a suite of clones, strains…



https://www.cell.com/trends/parasitology/fulltext/S1471-4922(10)00025-5

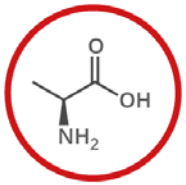
CRISPR/Cas9-directed evolution (CDE) in plants



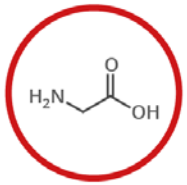
A GUIDE TO THE TWENTY COMMON AMINO ACIDS

AMINO ACIDS ARE THE BUILDING BLOCKS OF PROTEINS IN LIVING ORGANISMS. THERE ARE OVER 500 AMINO ACIDS FOUND IN NATURE - HOWEVER, THE HUMAN GENETIC CODE ONLY DIRECTLY ENCODES 20. 'ESSENTIAL' AMINO ACIDS MUST BE OBTAINED FROM THE DIET, WHILST NON-ESSENTIAL AMINO ACIDS CAN BE SYNTHESISED IN THE BODY.

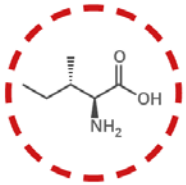
Chart Key: ALIPHATIC · AROMATIC · ACIDIC · BASIC · HYDROXYLIC · SULFUR-CONTAINING · AMIDIC · NON-ESSENTIAL · ESSENTIAL



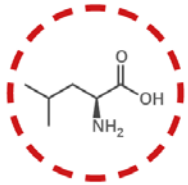
ALANINE A
Ala
GCT, GCC, GCA, GCG



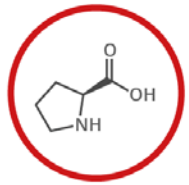
GLYCINE G
Gly
GGT, GGC, GGA, GGG



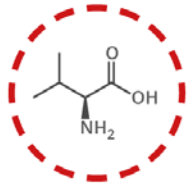
ISOLEUCINE I
Ile
ATT, ATC, ATA



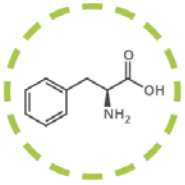
LEUCINE L
Leu
CTT, CTC, CTA, CTG, TTA, TTG



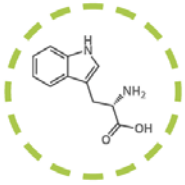
PROLINE P
Pro
CCT, CCC, CCA, CCG



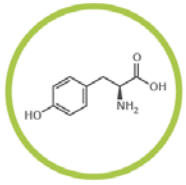
VALINE V
Val
GTT, GTC, GTA, GTG



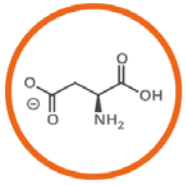
PHENYLALANINE F
Phe
TTT, TTC



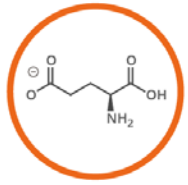
TRYPTOPHAN W
Trp
TGG



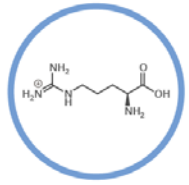
TYROSINE Y
Tyr
TAT, TAC



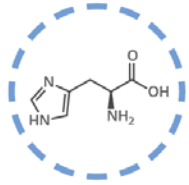
ASPARTIC ACID D
Asp
GAT, GAC



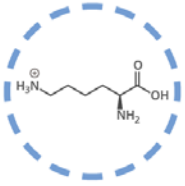
GLUTAMIC ACID E
Glu
GAA, GAG



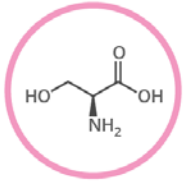
ARGININE R
Arg
CGT, CGC, CGA, CGG, AGA, AGG



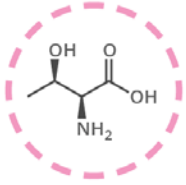
HISTIDINE H
His
CAT, CAC



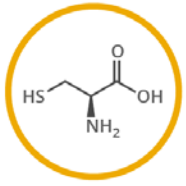
LYSINE K
Lys
AAA, AAG



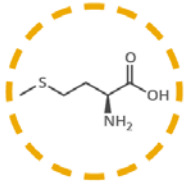
SERINE S
Ser
TCT, TCC, TCA, TCG, AGT, AGC



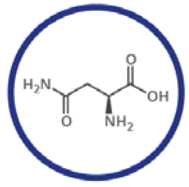
THREONINE T
Thr
ACT, ACC, ACA, ACG



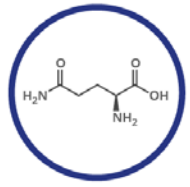
CYSTEINE C
Cys
TGT, TGC



METHIONINE M
Met
ATG



ASPARAGINE N
Asn
AAT, AAC



GLUTAMINE Q
Gln
CAA, CAG

Note: This chart only shows those amino acids for which the human genetic code directly codes for. Selenocysteine is often referred to as the 21st amino acid, but is encoded in a special manner. In some cases, distinguishing between asparagine/aspartic acid and glutamine/glutamic acid is difficult. In these cases, the codes asx (B) and glx (Z) are respectively used.