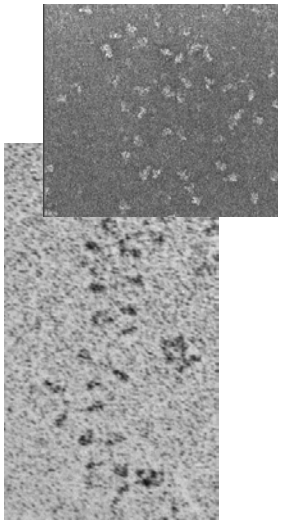
Sample preparation for structural studies

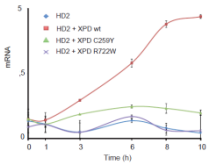
Arnaud.Poterszman@igbmc.fr

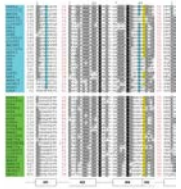
Functional characterization

Cellular imaging



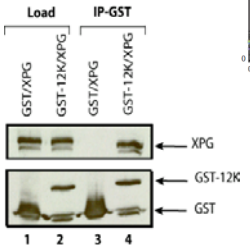
Bio computing

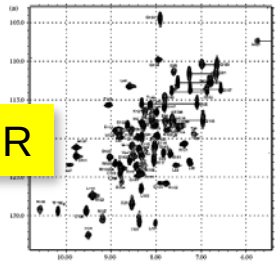




Partner identification

Cryo-EM





NMR

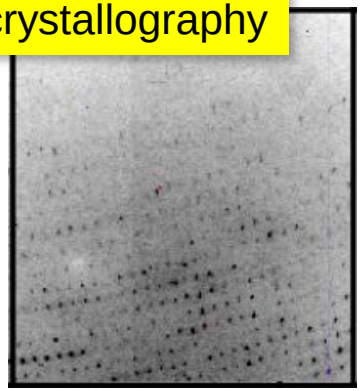
Domain/Orthologs indentification

Targets

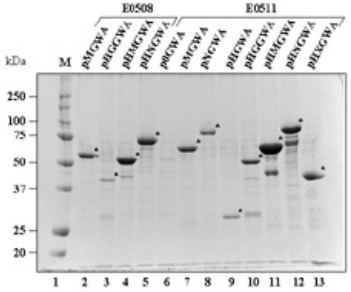
Structure/ function Analysis Drug design

Production

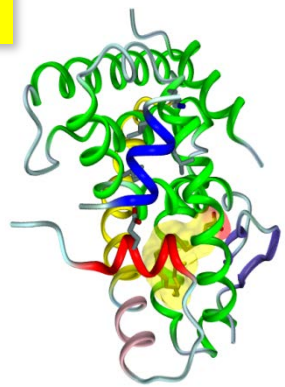
X ray crystallography



Modelling

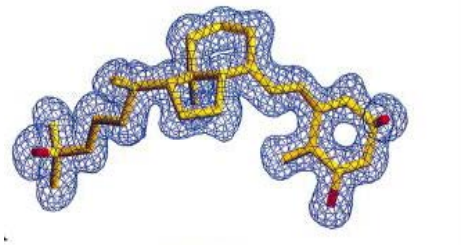


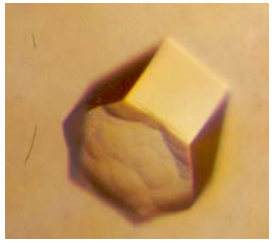
QC & Biophysical characterization



Sample requirements

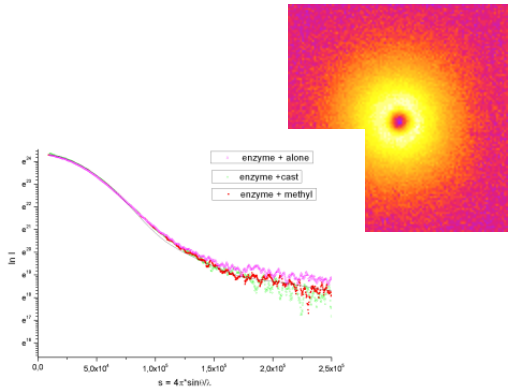
X-ray diffraction, SAXS





crystallization # 1-10 mg/ml, 100 µl

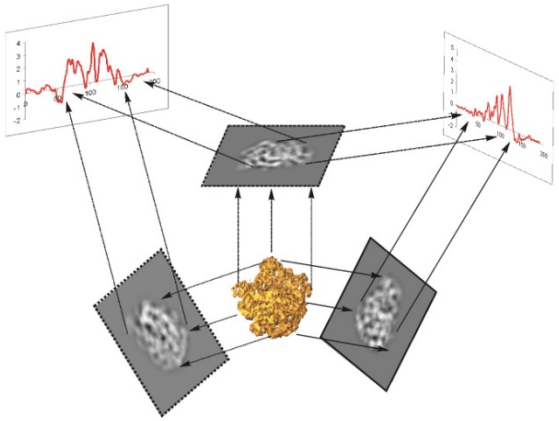
SAXS, SANS # 1-10 mg/ml, 50 µl



NMR

NMR # 100µM-2mM, 500 µl

Electron Microscopy



Un-supported cryo: # 0.5 mg/ml, 5µl/assay

Supported cryo, neg staining: # 50 µg/ml, 5µl/assay

Recombinant or endogenous?

Isolate sample from native source

Advantages – Protein solubility, authenticity

Disadvantages – Expense/effort, yield, abundancy

Popular sources: E coli, yeast, HeLa cells
Model imposed by the biological question

Bacterial expression

Advantages – Easy, great over-expression, low protease activity, no post-translational modifications

Disadvantages – Protein solubility, lack of post-translational modifications

Eukaryotic expression

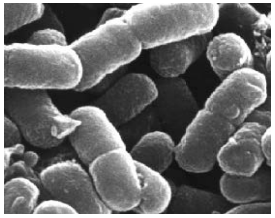
Advantages – Protein solubility, post-translational modifications

Disadvantages – Expensive, low yield, proteases, time consuming

Recombinant expression

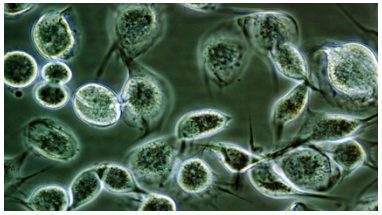
Prokaryotic

E. coli,
B. subtilis
....



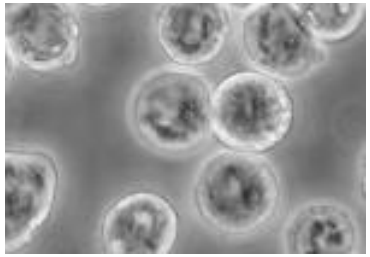
Eukaryotic

Yeast
Insect cells
Mammalian



Cell free systems:

E coli
Wheat germ,
Insect



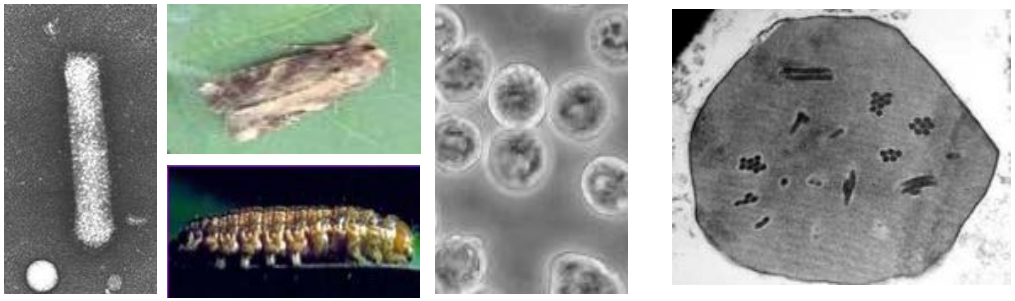
Sample preparation for structural analysis

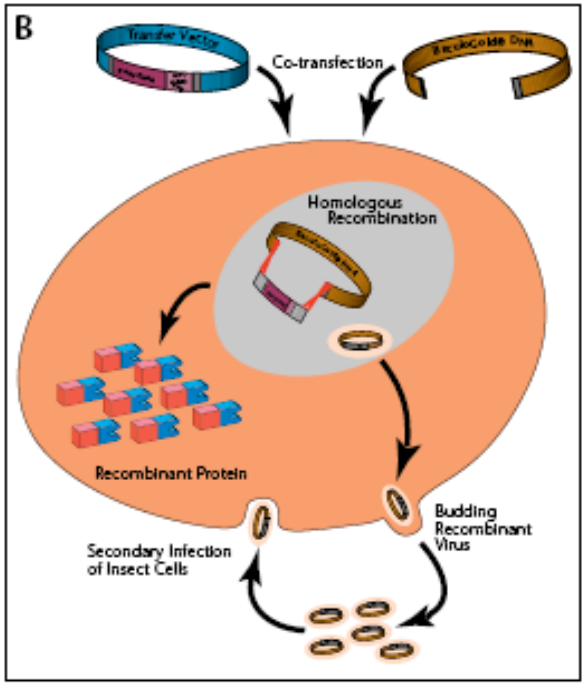
Recombinant protein expression for structural biology: Insights into the baculovirus expression system

Co-expression for reconstitution of multiprotein complexes and dissection of the protein-protein interaction network

In vivo approaches for labelling mammalian proteins to facilitate isolation of endogenous complexes and their characterization in a cellular environment

Principle: use strong late viral promoters non essential for virus replication in a cellular system



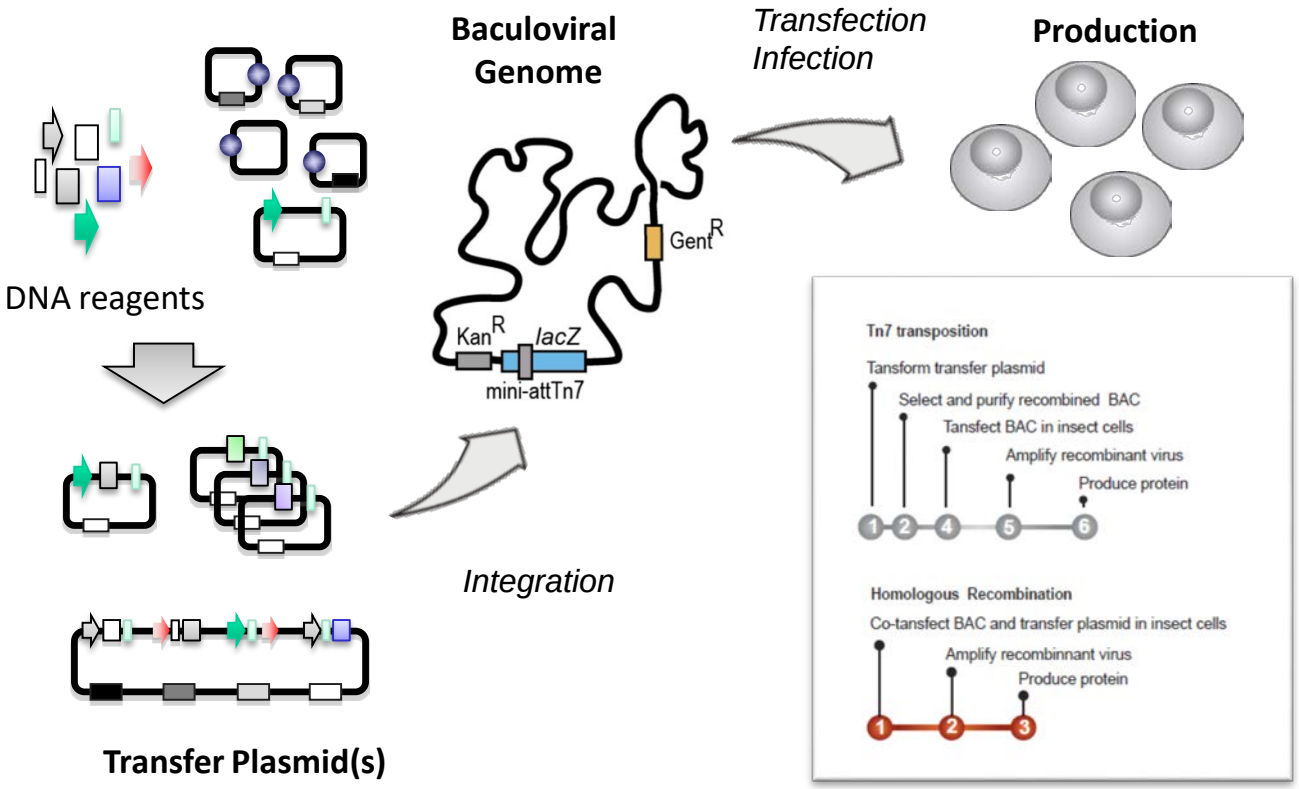


Rod-shaped, d = 40-50 nm, l = 200-400 nm
Double-stranded DNA virus (135 kb)
Virons occluded in a polyhedrin (PH) matrix*

Replace polyhedrin (PH) coding seq. with GOI
Strong promoters (PH, P10)
Protein expression 36-48 h post-infection
Cell lines grow well in suspension (27 °C, in Phosphate buffered medium, no CO_2)
Baculoviruses have a restricted host range and are safe to manipulate
High levels of heterologous expression, production of toxic proteins

* non required for replication in a cellular system

An expression flowchart for BV expression



What is needed to express a protein ?

The expression unit

- Strong promoter: PH or p10
- Kozak sequence:
- Gene of interest
- Terminator

Promoter Gene (GOI) Terminator

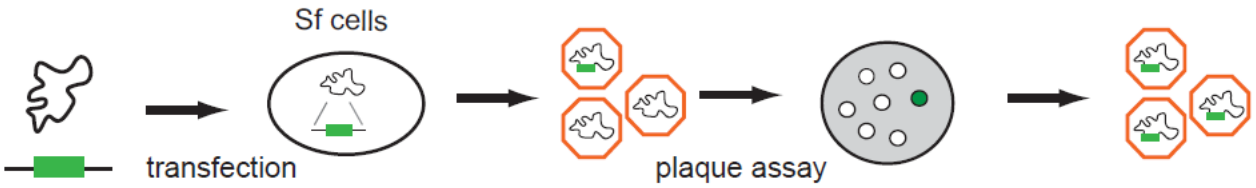
On both sides, elements that will allow the integration of the expression unit(s) into the viral genome:

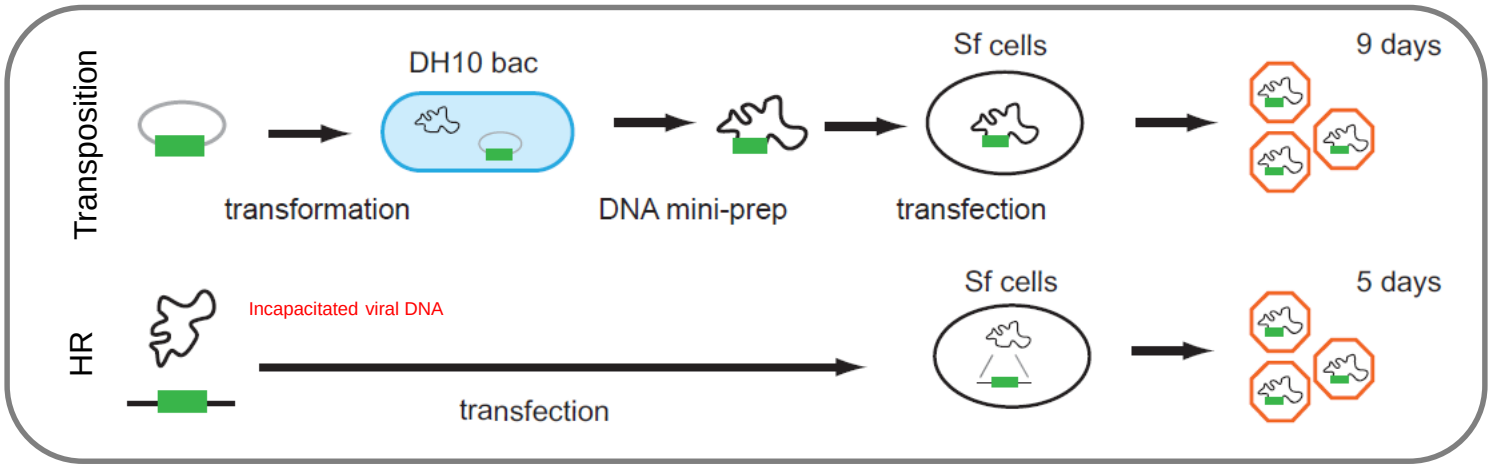
- Either segment from viral genome for homologous recombination in insect cells between the transfert vector and the viral DNA

- Or transposons (Tn7L and Tn7R) recombination sites (LoxP) when a bacmid is to be used

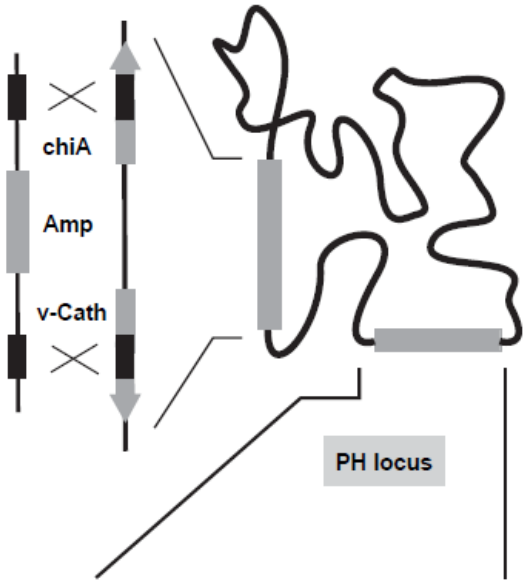
Transposition vs Homologous recombination

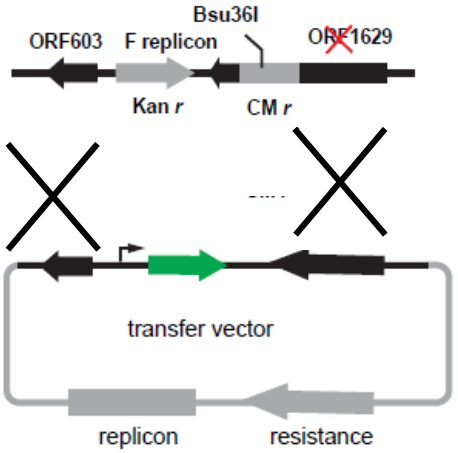
Clone into transfer vetor ▶ Recombinant baculovirus generation ▶ Small scale expression test ▶ Optimization and large scale expression

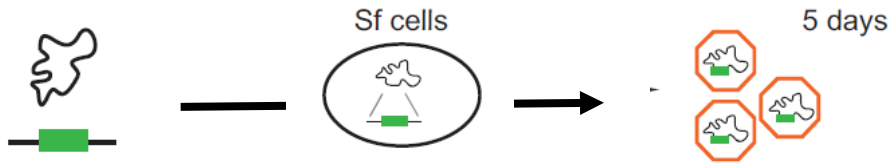




Homologous recombination in insect cells with disabled viral genome







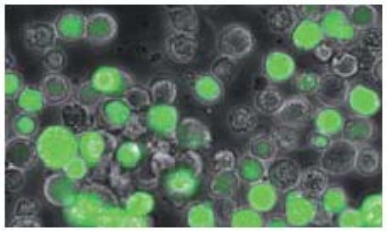
Disabled viral DNA

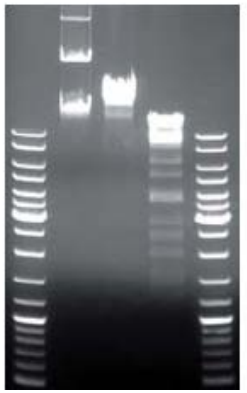
Co-transfection

Viable genome can only be formed if the truncation is bridged and repaired by recombination with a suitable transfer vector.

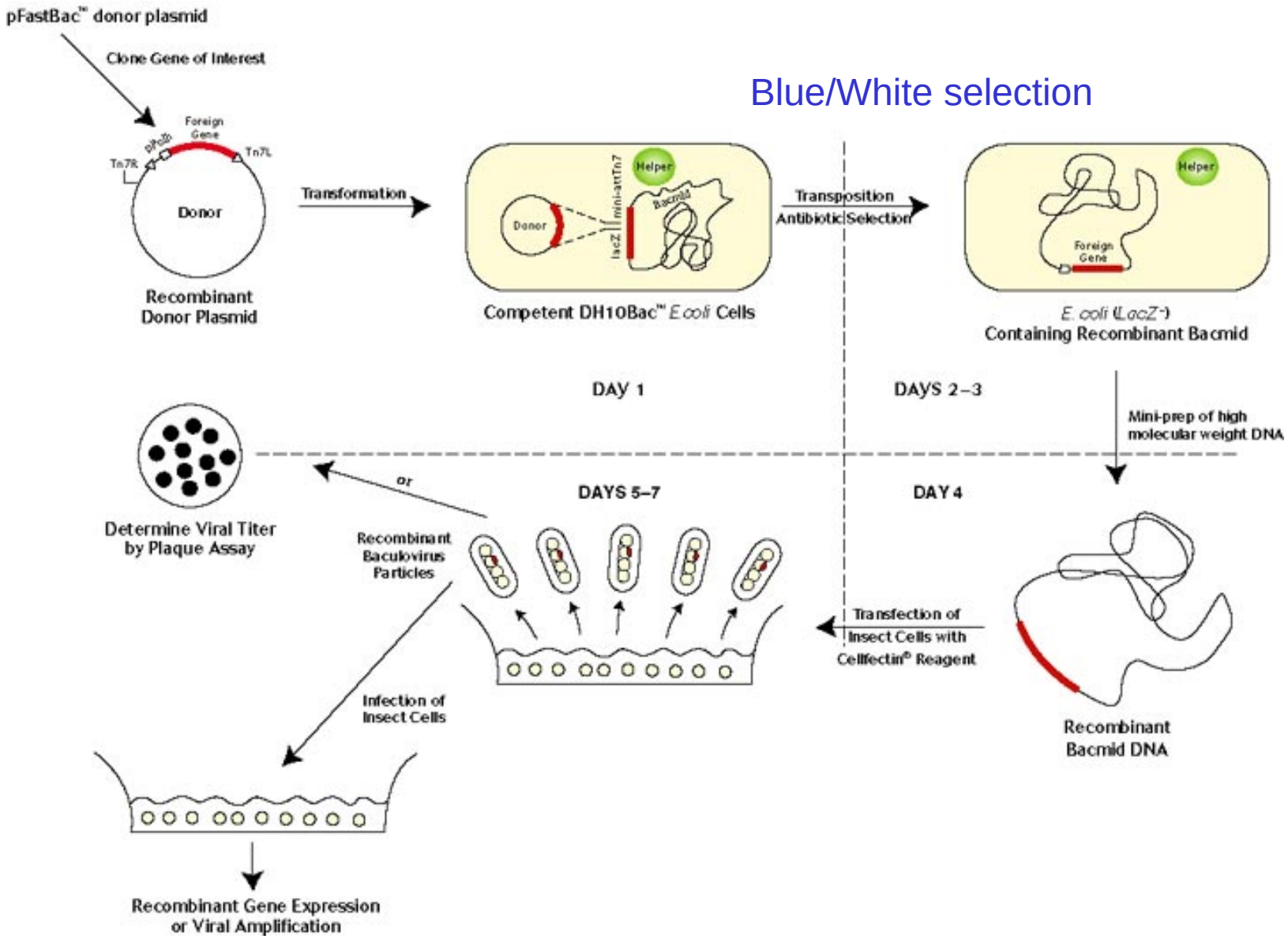
Recombination inserts the foreign gene (GFP) into the viral DNA, restores the deleted gene, allowing virus replication.

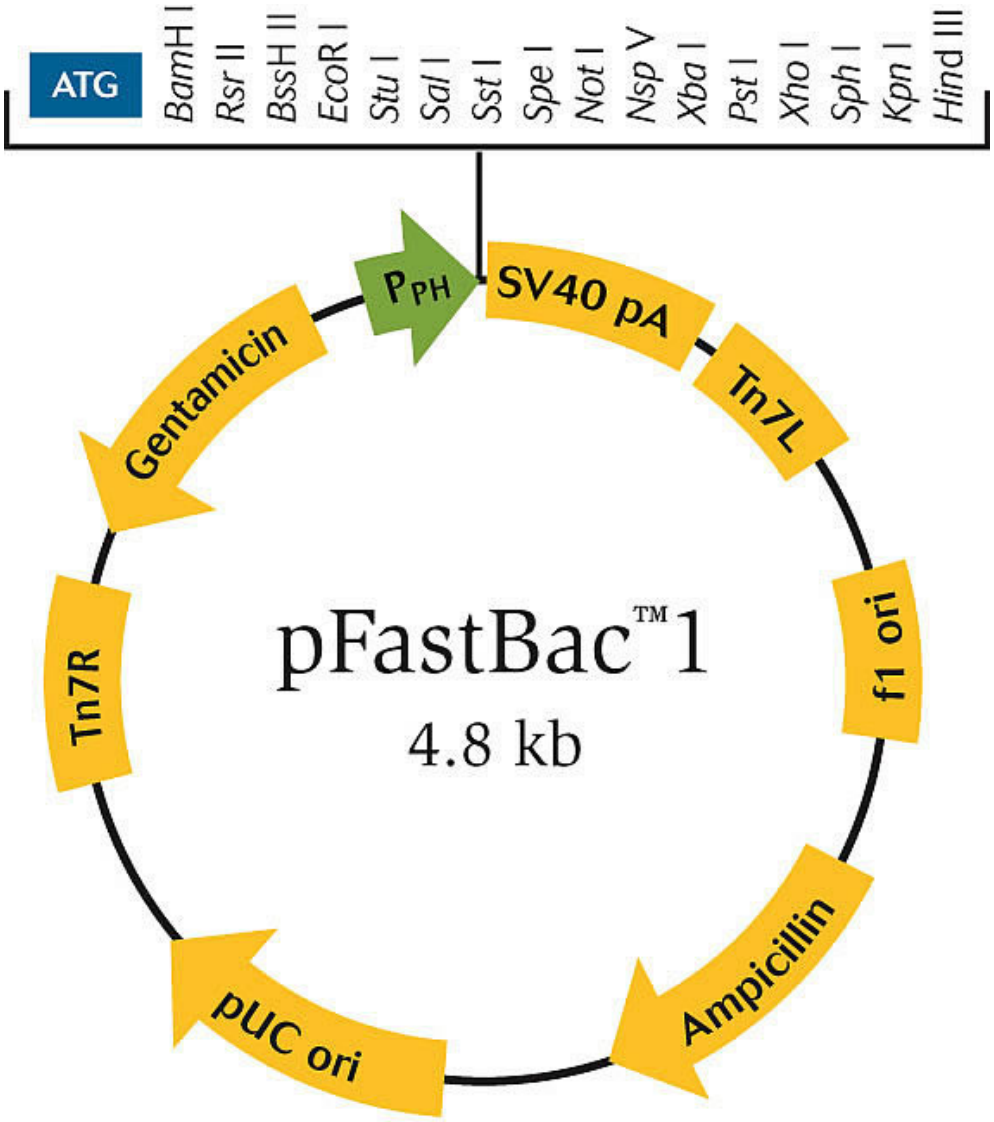
No need for plaque selection (screening) for medium size inserts





Transposition in E. Coli



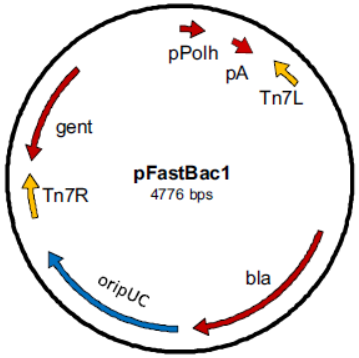

ATG
BamH I
Rsr II
BssH II
EcoR I
Stu I
Sal I
Sst I
Spe I
Not I
Nsp V
Xba I
Pst I
Xho I
Sph I
Kpn I
Hind III
P PH
SV40 pA
Tn7L
f1 ori
Ampicillin
pUC ori
Tn7R
Gentamicin
pFastBac™1
4.8 kb

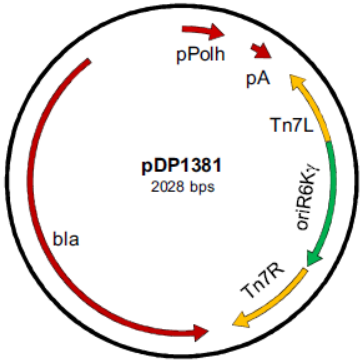
B2F vs B2B

1/ Low transposition efficiency: transposition into the E. coli genome and sub- optimal transposition system

2/ Bacmid preparation contanminated with transfer vector DNA

Bac to the Future

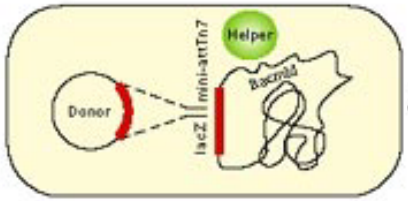




New baculovirus expression vectors: single antibiotic, conditional replication origin (oriR6Kγ)

New helper plasmid Tn7 transposase delivery vector

New E coli strain (chromosomal Tn7 site inactivated) that replaces the DH10 Bac: DE26

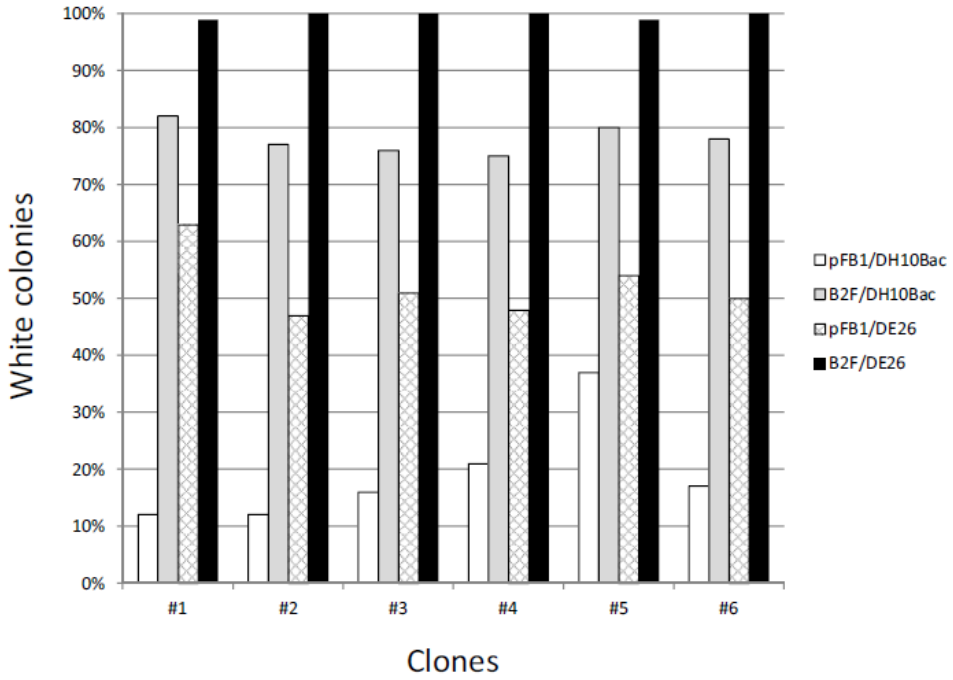


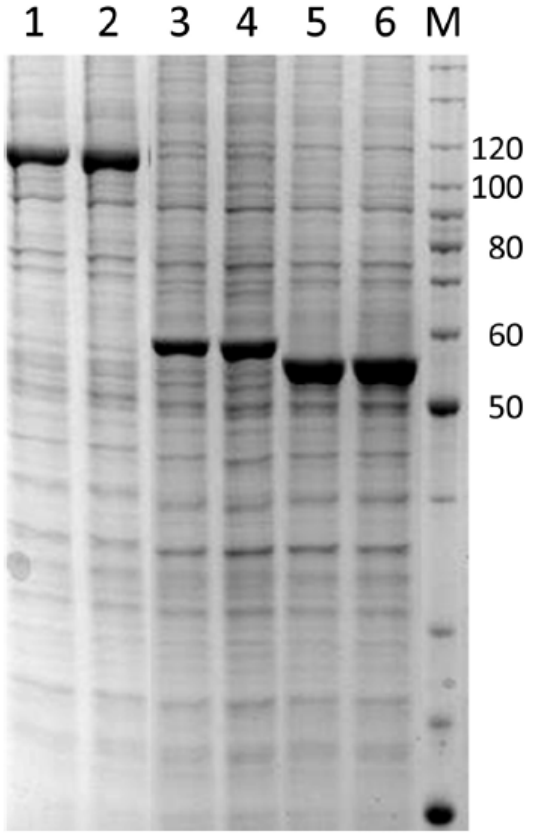
J Mehalko, D Esposito, J Biotech 2016

B2F vs B2B

B2F transfer plasmid (no transfer plasmid replication, pir-)

B2F transfer plasmid + new strain (no transfer plasmid replication and no transposition into the E. coli backbone)





J Mehalko, D Esposito, J Biotech 2016

In practice: culture conditions

For infections cells in exponential growth phase are required.

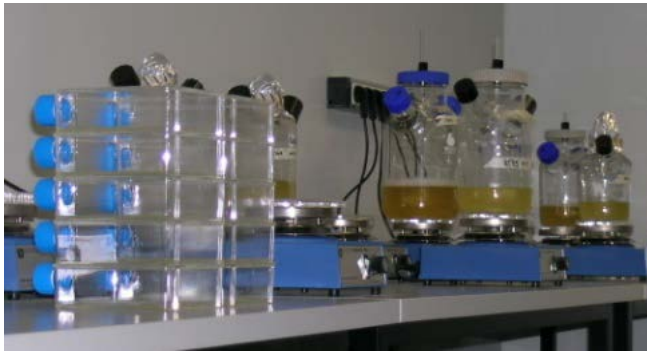
infect cells a 0.5 to 2.0 $1O^6$ cell/ml
T =27 °C, phosphate based buffer (no CO_2)
monolayers or suspension (Deep Well, Spinner, Bottles..)

Optimization of culture conditions

- harvest time post-infection: 48, 72 hrs
- multiplicities of infection: 0.1, 1, 5, 10

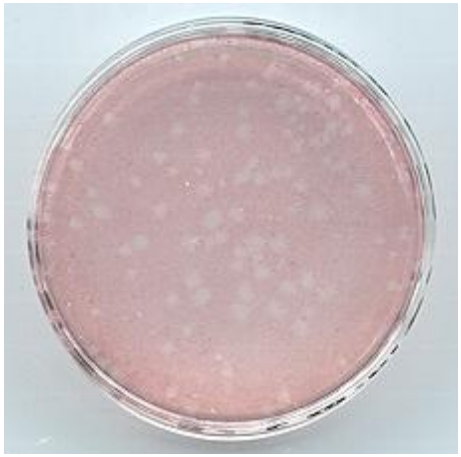
Very important for co-infections experiments

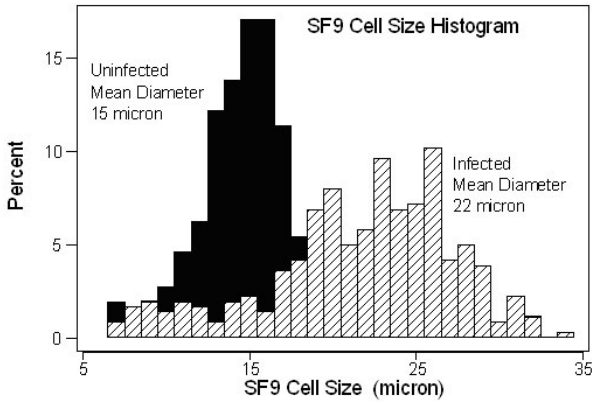
- cell line/media of choice:
  Sf9, Sf21, H5
  with or without serum



Scale: from 3 ml to several L

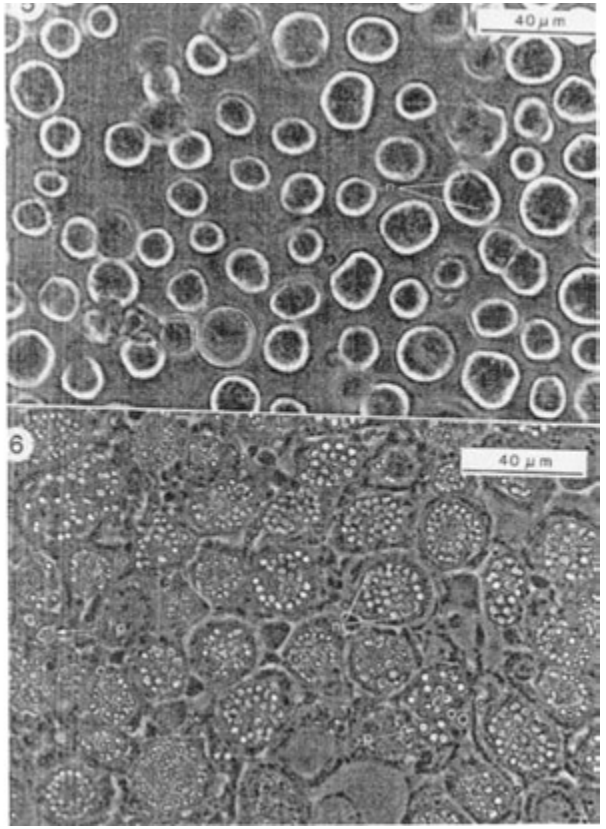
Amplification and production





Sf9 non-infected

Sf9 infected



Infected cells: stop dividing and swell

Flowchart for Baculovirus Expression

Clone the gene(s) of interest into a bacterial transfer vector

Generate the recombinant virus
Transfection/Co-transfection

Small scale expression assay

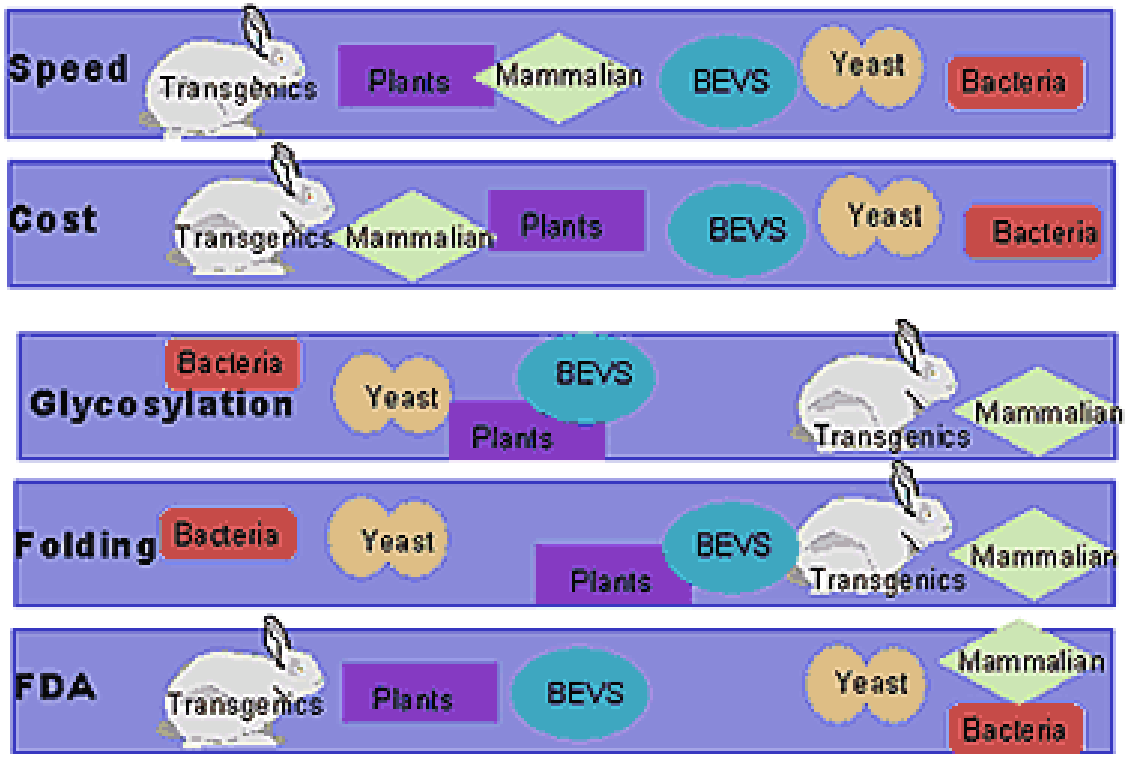
Prepare a high titer virus stock <108 pfu/ml

Optimization of the expression conditions and large scale production

YEAST

Mammalian

WORST BEST

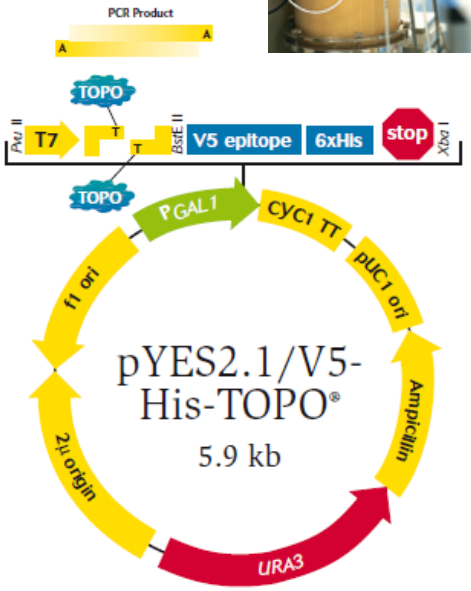


Yeast expression

Sacharomyces, *Scizosacchromyces pombe*, *Pichia pastoris* and *Hansanuela polymorpha*

Expression of recombinant proteins in *S. cerevisiae* can be done using three types of vectors: integration vectors (YIp), episomal plasmids (YEp), and centromeric plasmids (YCp).

pYES2 is a 5.9 kb YEp vector designed for inducible expression of recombinant proteins in sc. Features of the vectors allow easy cloning of your gene of interest and selection of transformants by uracil prototrophy .



- Yeast *GAL1* promoter for high level inducible protein expression in yeast by galactose and repression by glucose
- A versatile multiple cloning site for simplified cloning
- *CYC1* transcriptional terminator for efficient termination of mRNA
- *URA3* gene for selection of transformants in yeast host strains with a *ura3* genotype
- Ampicillin resistance gene for selection in *E. coli*

Mammalian expression adapted for production of extracellular and secreted proteins

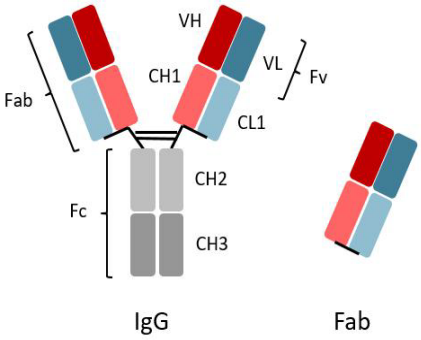
Antibodies and derivatives
Large extracellular domains

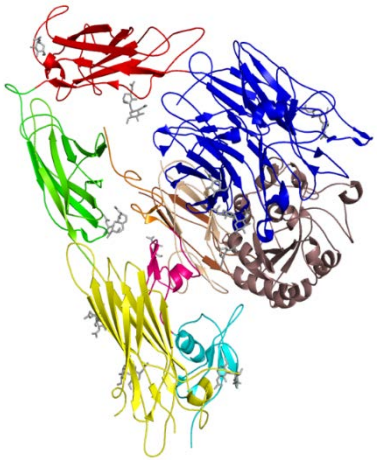
Modular multi-domain organisation
Posttranslational modifications

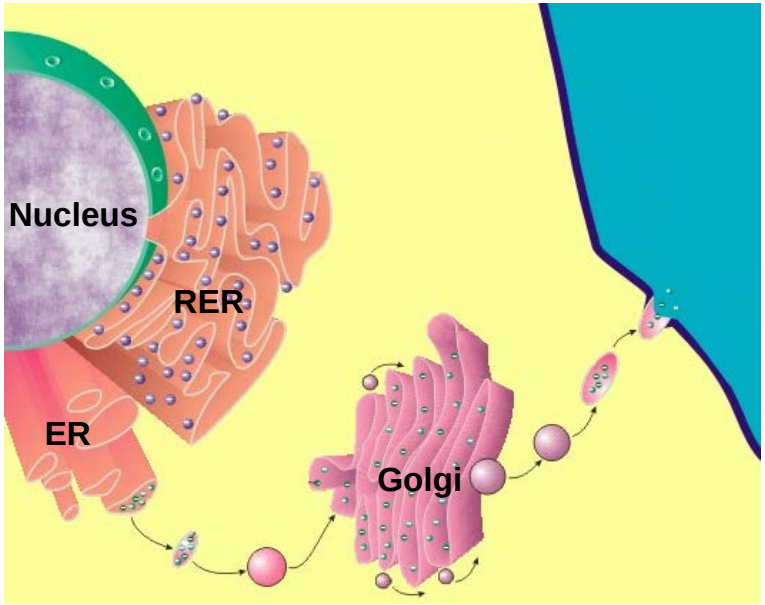
- Disulfide bridges
- Glycosylation

S-S formation (ER)
Glycosylation (ER+Golgi)
Quality control (ER)

Only correctly folded proteins are secreted





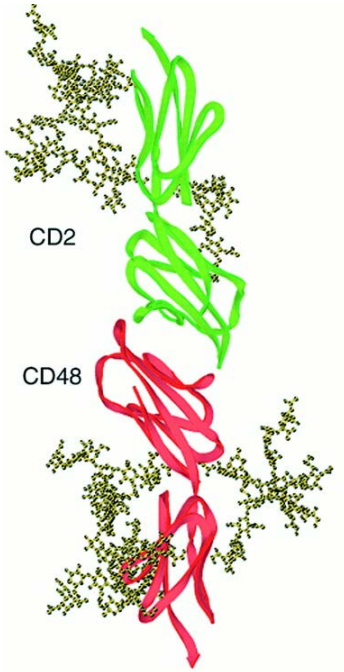


(oxford)

POST-TRANSLATIONAL MODIFICATIONS

GLYCOSYLATION

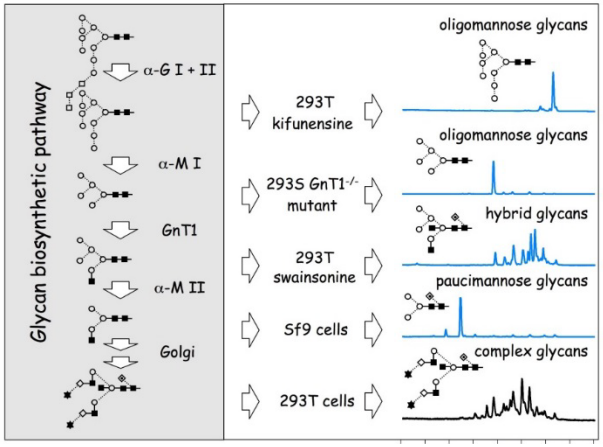
- Mammalian sugar chains have highly complex structures
- Good for functional studies
- Big problem for protein crystallization



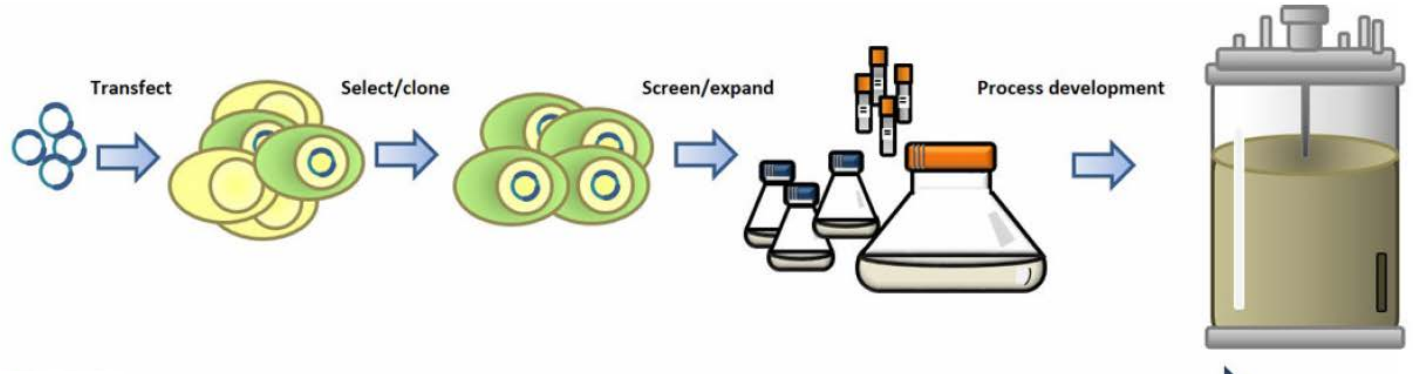
SOLUTIONS

- Mutagenesis of glycosylation sites
- Enzymatic deglycosylation
- Engineered cell lines (CHO Lec strains)
- Chemical inhibitors of glycosylation pathway
- Insect cells (simpler sugars)

(oxford)



Timelines for mammalian expression



Standard Integration/
Selection

~12 months

Transient Gene Expression

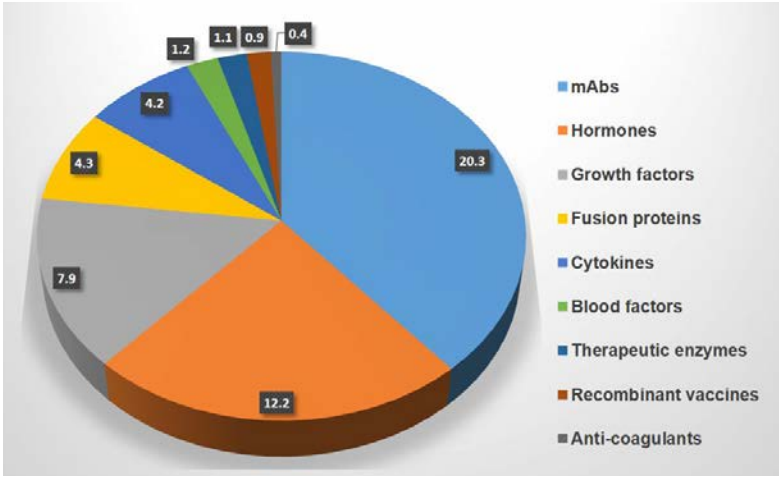
~2 weeks

Site specificintegration
(RMCE)

~2-3 months

Viral Expression
(Vaccinia, Lentiviruses,…)

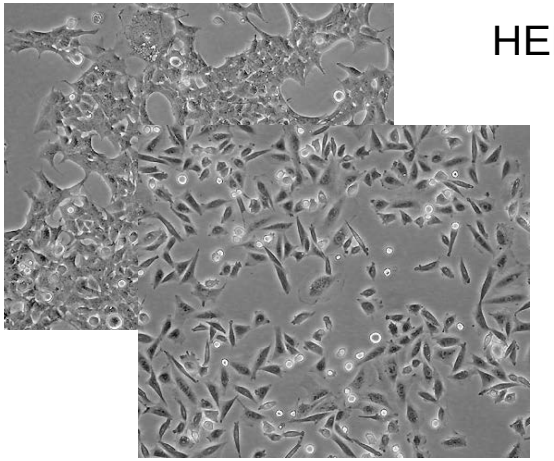
~2 months



US Top Sales Biologics

Bandaranayake and Almo, FEBS, 2014

Mammalian expression



HEK 293: Human embryonic kidney cells

CHO: Chinese Hamster Ovary cells

Most mammalian cells are adherent
Cultured in plates or flasks (suspension)
Grow in monolayer on specially treated surfaces
Medium supplemented with 5-10% Fetal Calf Serum
Laminar flow cabinet, CO_2 incubator

Large scale transfection with PEI

DNA can be introduced into a host cell by transfection with polyethylenimine (PEI), a stable cationic polymer (Boussif et al., 1995). **PEI condenses DNA into positively charged particles that bind to anionic cell surfaces**. Consequently, the DNA:PEI complex is endocytosed by the cells and the DNA released into the cytoplasm (Sonawane et al., 2003).

Step 2.1: Prepare suspension cells for the large scale transfection

2.1.1 Grow 400 ml of suspension cells to a density of 2-3 x 10^6 cells/ml

↓

2.1.2-2.1.3 Transfer cells to sterile centrifuge bottles
Centrifuge at 200 x g, 5 min, room temperature
Aspirate the supernatant

↓

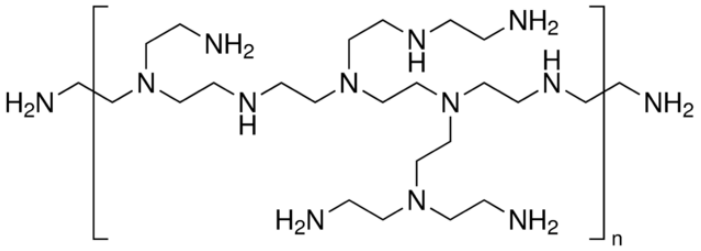
2.1.4-2.1.5 Add 25 ml of the appropriate suspension growth medium
Resuspend cells using a serological pipette

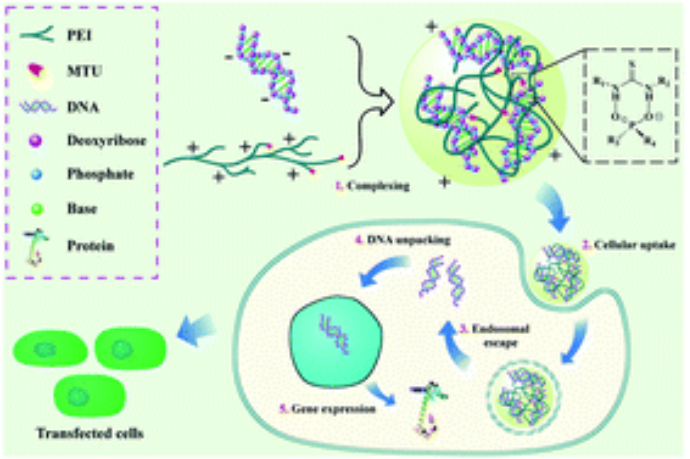
↓

2.1.6 Add cell suspension to 335 ml of suspension growth medium in a sterile square bottle

↓

2.1.7-2.1.8 Place cells in a 37°C, 8% CO_2 shaking incubator, with shaking at 130 rpm
Do not completely tighten the cap





Sample preparation for structural anaysis

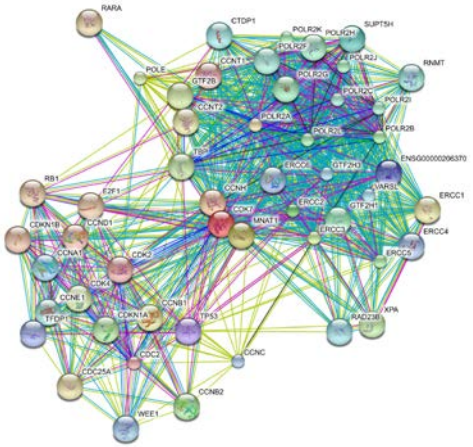
Recombinant protein expression for structural biology: Insights into the baculovirus expression system

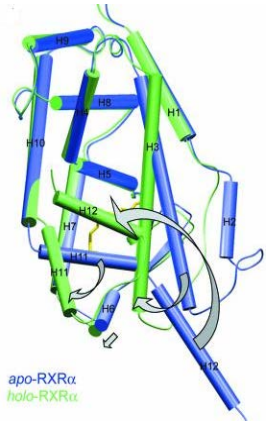
Co-expression for reconstitution of multiprotein complexes and dissection of the protein-protein interaction network

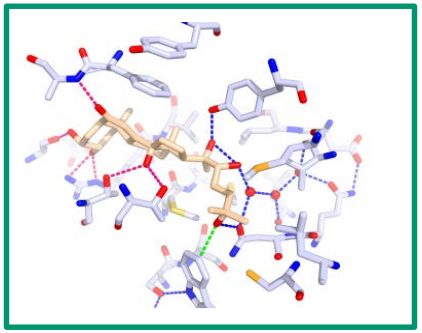
Genome engineering for labelling mammalian proteins to facilitate isolation of endogenous complexes and their characterization in a cellular environment

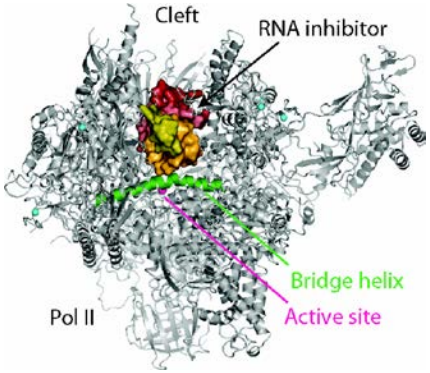
Most proteins do not function as isolated particules……

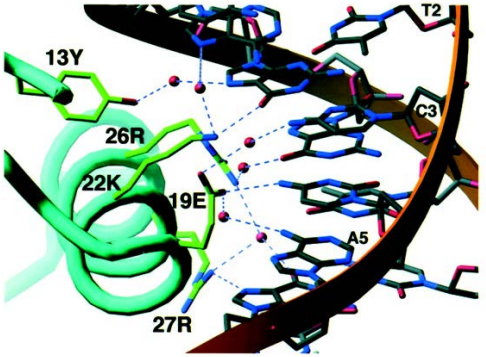
…. but interact with partners to fullfill their fonction.









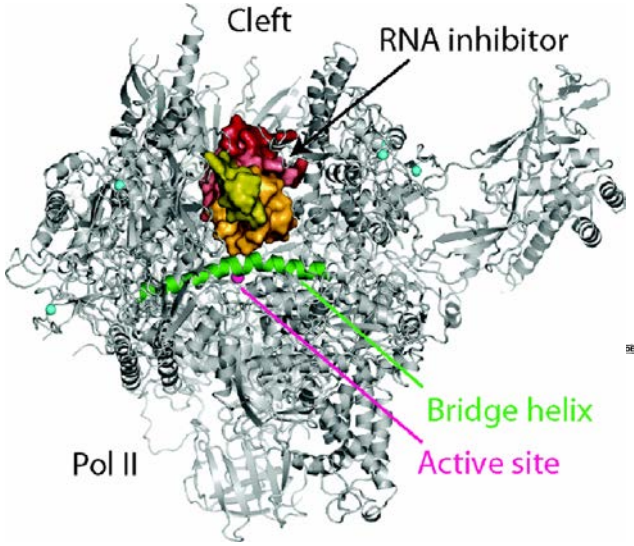


Types of complexes

Non obligate and obligate

Subunits exist independantly

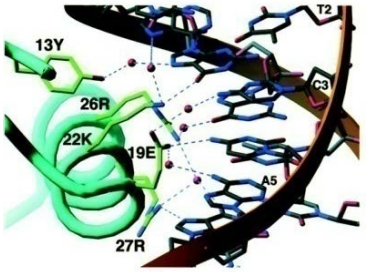
Subunits are not found as stable structures in vivo



Life-time of complexes

Transient interactions: associate and dissociate in vivo
- weak: dynamic equilibrium in solution
- strong: molecular trigger to switch on and off

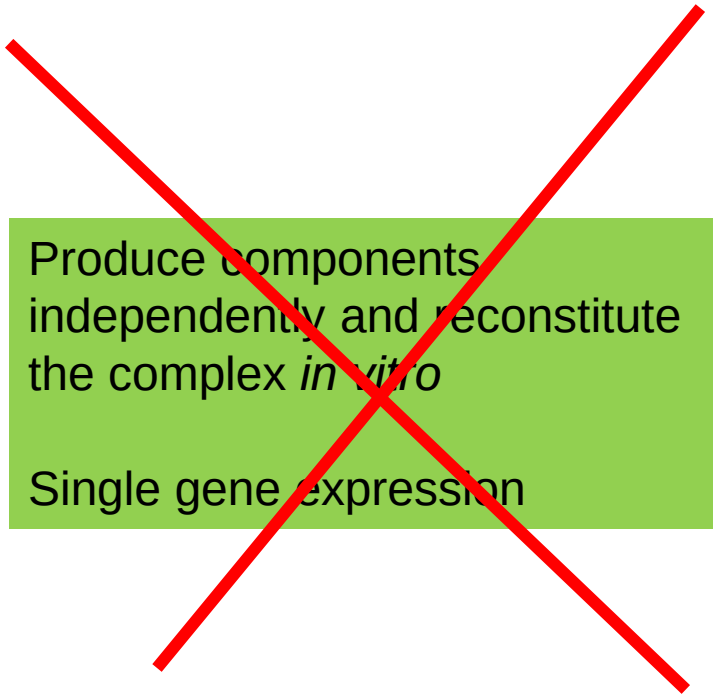
Permanant interactions: subunits only exist in complexed state. The complex can be purified



Weak	Intermediate	Strong
K_d mM-μM	K_d μM-nM	K_d nM-fM

Implications for production

Obligatory complexes



Produce components independently and reconstitute the complex *in vitro*

Single gene expression

Produce components of the complex simultanously which are assembled *in vivo*

Co-expression

Strategies for production of multi-protein complexes

Separate expression of subunits purify and mix; mix and purify

E. coli

Co-transformation with several single promoter plasmids

Transformation with multigene expression plasmid

BVES

Co-infection of insect cells by several viruses

Infection with a multigene expression virus

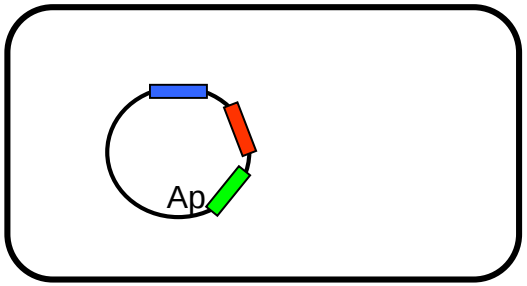
Co-expression in E. coli

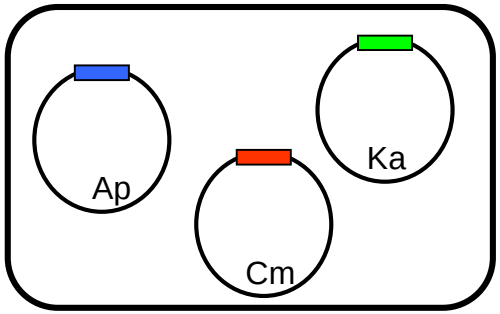
Co-transformation with several expression vectors

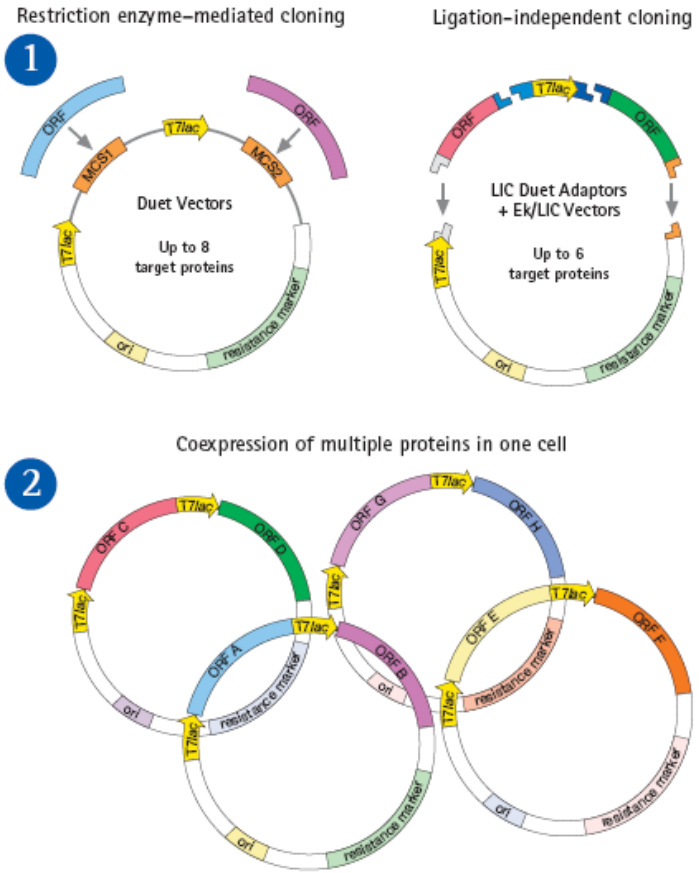
different antibiotic resistances
(different origin of replication)

One expression vector with several genes under the control of

a single promoter
several promoters (e.g. each one having a single gene





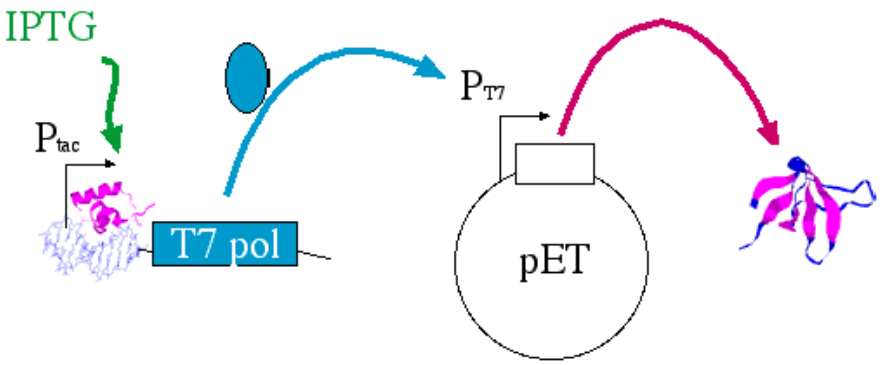


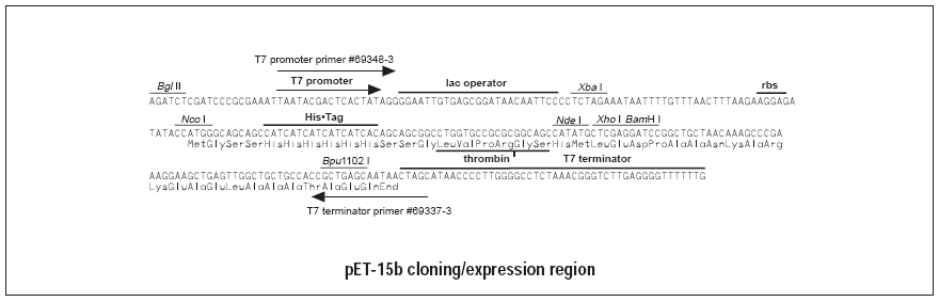
T7 based expression systems

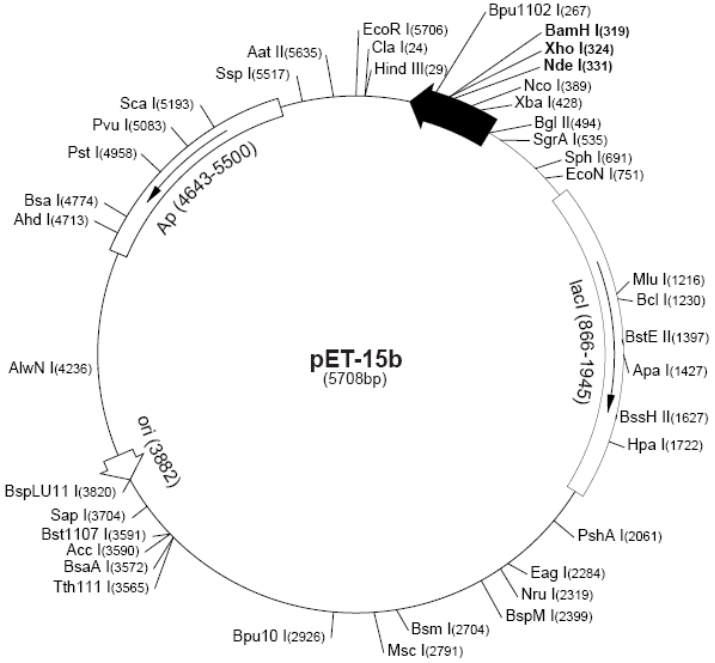
pET vectors

BL21 [DE3]
BL21 [DE3] pRARE
BL21 [DE3] pLysS
……

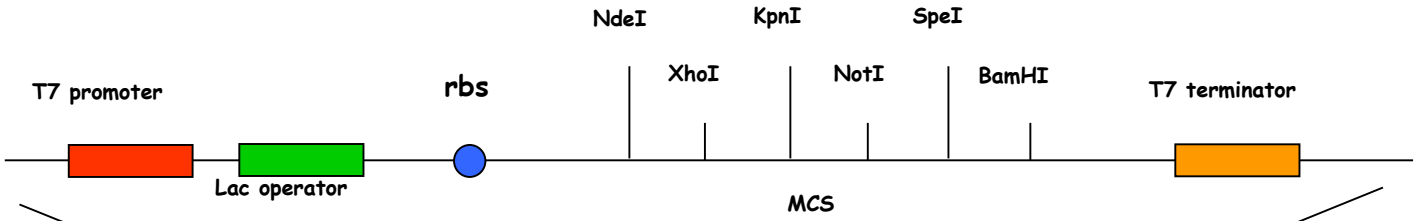
IPTG induction
Auto-inductible medium





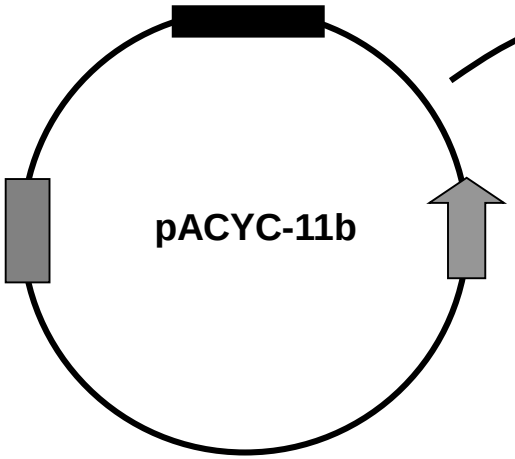


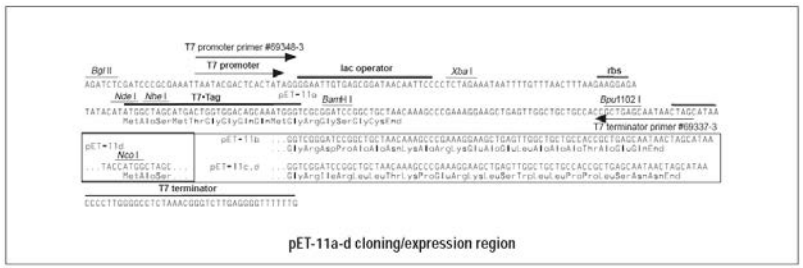
pACYC-11b

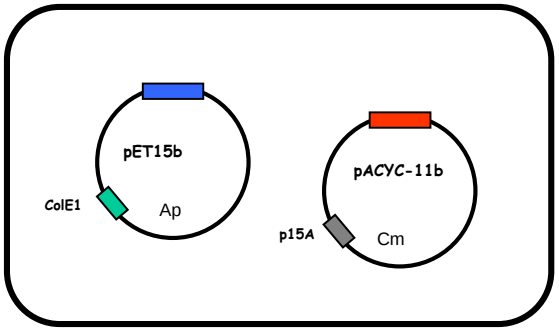


T7 expression cassette

Co-transformation and co-expression with pET (or pGEX) vectors

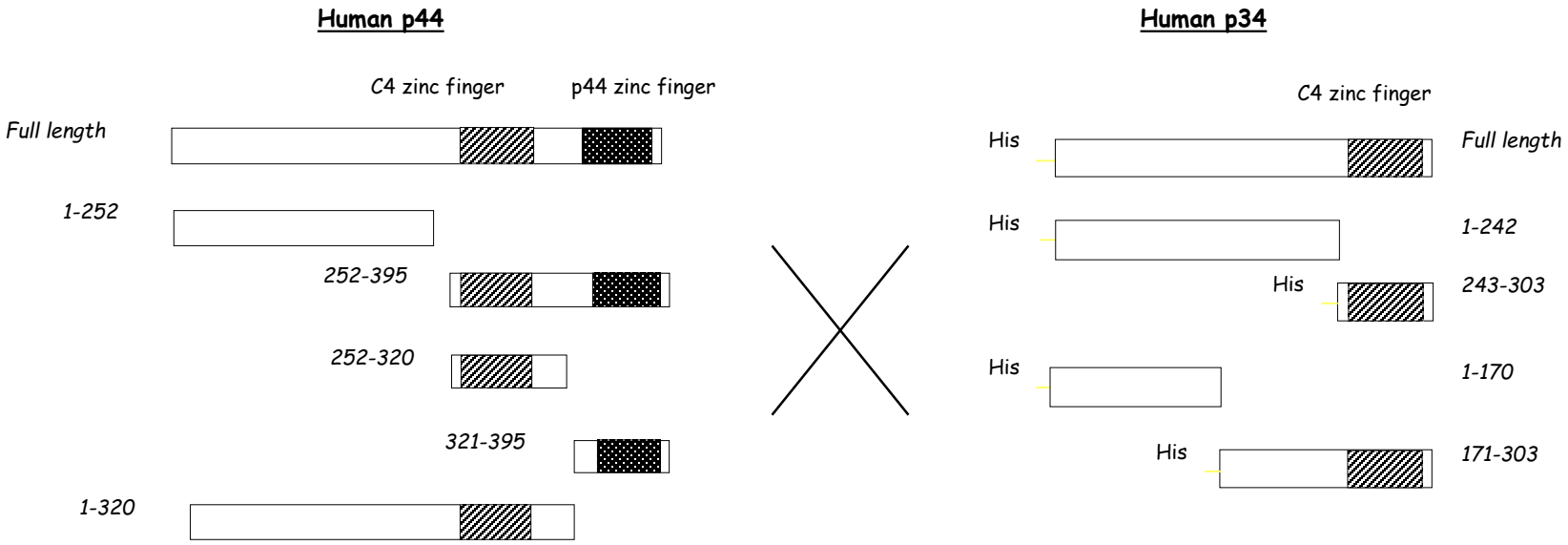




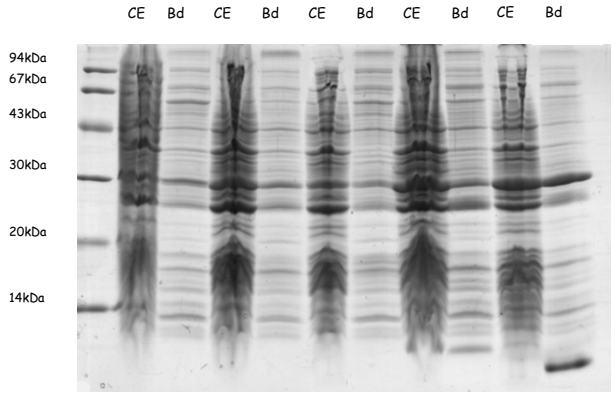


Dissection of p34-p44 Interaction by Coexpression

Aim: find a stable and soluble complex to set up crystallization trials



24 combinations were tested several times, only 1 gave a positive result



CE: crude extract

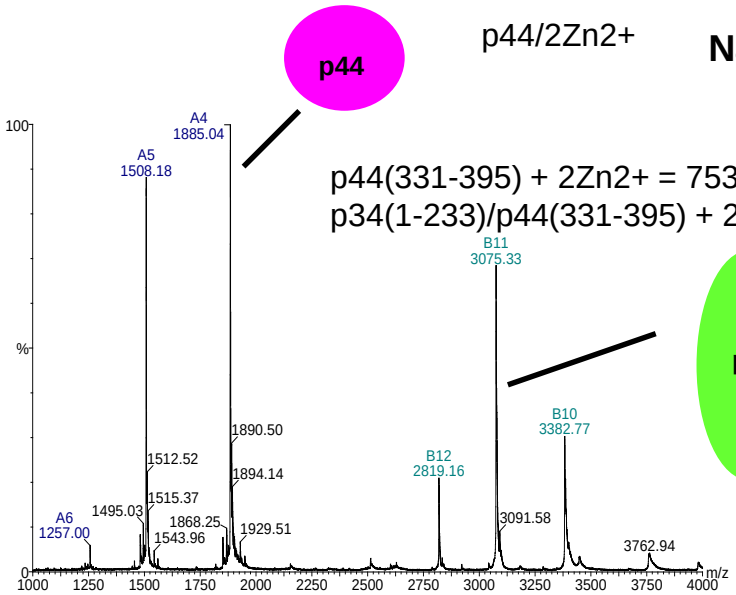
Bd: beads

His-p34(1-242)

p44(321-395)

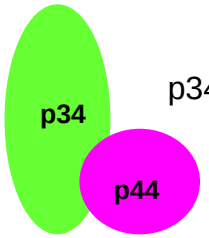
p44/2Zn2+

Native ES-MS

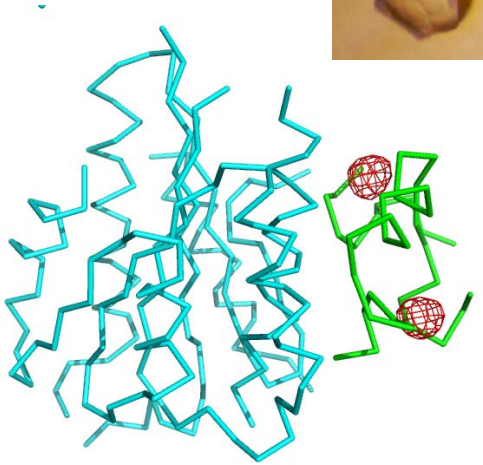


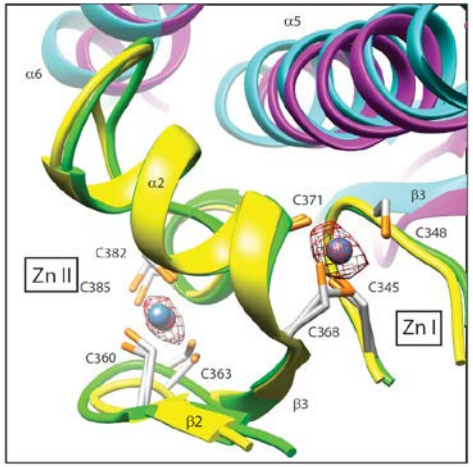
p44(331-395) + 2Zn2+ = 7536,0 ± 0,1 Da

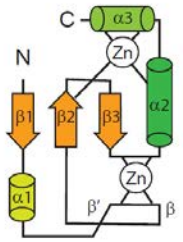
p34(1-233)/p44(331-395) + 2Zn2+ = 33817,6 ± 0,2 Da

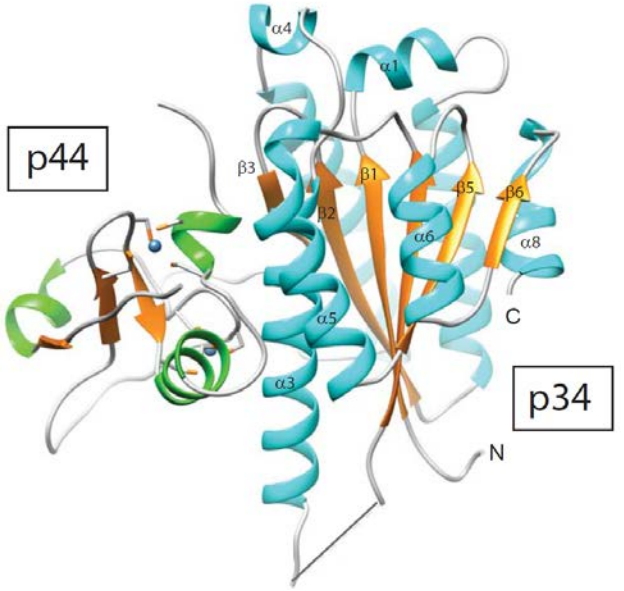


p34/p44/2 Zn2+









Radu et al. NAR 2017)

Troubleshooting: Improving expression of soluble proteins

No expression, accumulation of so-called inclusion bodies, or of incompletely synthetized proteins, degradation

Reducing the rate of protein synthesis.
Lowering the growth temperature (37 to 20). This decreases the rate of protein synthesis and usualy more soluble protein is obtained.
Use of a weaker promoter (*e.g.* **tryptophan, arabinose** instead of **T7**), of a lower copy number plasmid, lowering the inducer concentration.

Changing the growth medium:
Addition of prostethic groups or **co-factors** which are essential for proper folding or for protein stability.
Addition of buffer to control pH fluctuation in the medium during growth.
Addition of **1% glucose** to repress induction of the lac promoter by lactose, which is present in most rich media (such as LB, 2xYT).
Addition of **polyols (*e.g.* sorbitol) and sucrose**. The increase in osmotic pressure caused by these additions leads to the accumulation of osmoprotectants in the cell, which stabilize the native protein structure.

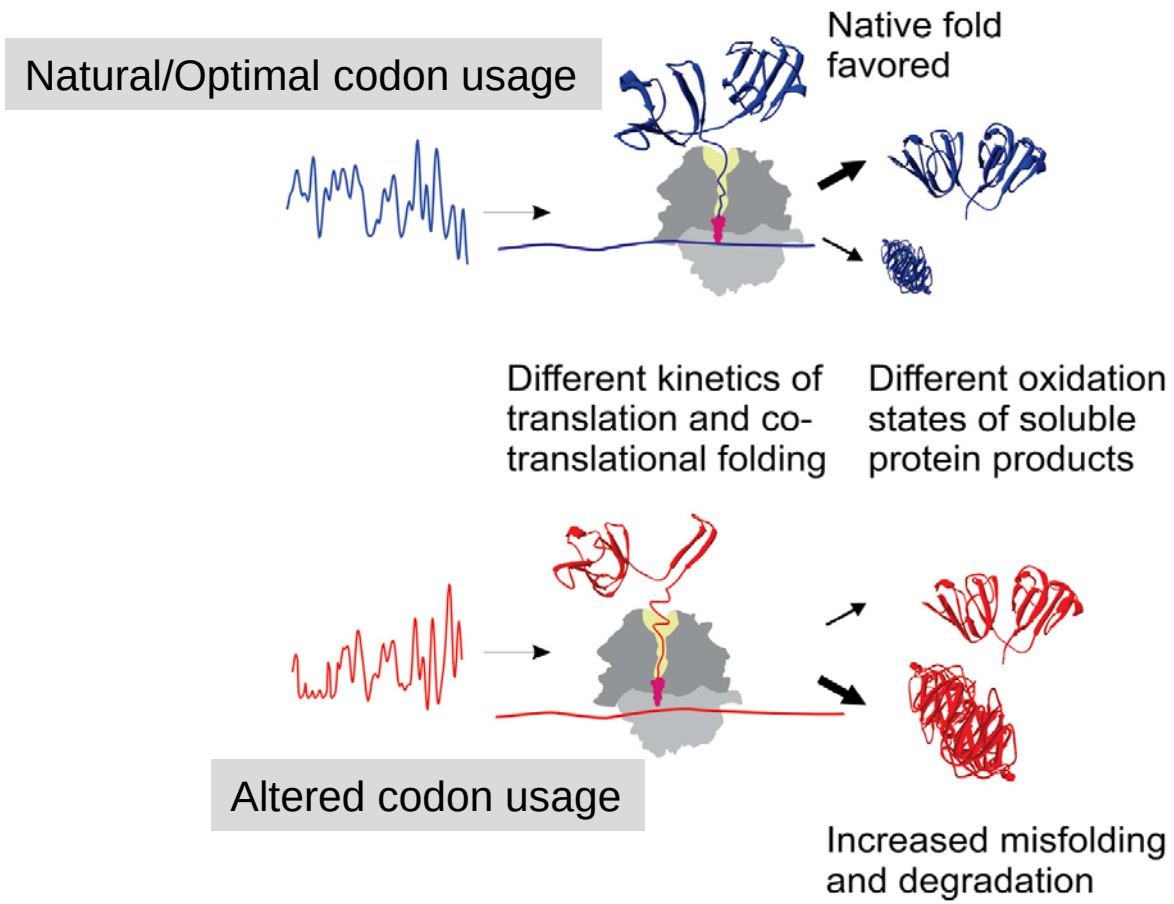
Troubleshooting: Improving expression of soluble proteins

Using specific host strains:
The solubility of disulfide bond containing protein can be increased by using a host strain with a more oxidizing cytoplasmic environment. Two strains are commercially available (Novagen):
AD494, which has a mutation in thioredoxin reductase (trxB).
Origami, a double mutant in thioredoxin reductase (trxB) and glutathione reductase (gor).

Addition of a fusion partner:
Fusion of the N-terminus of a heterologous protein to the C-terminus of a soluble fusion partner often improves the solubility of the fusion protein.

Expression of a fragment of the protein:
E. coli does not express well very large proteins (> 70 kDa). Chosing a smaller fragment of the target protein can improve expression levels and solubility.
The solubility of a poorly soluble (or insoluble) protein can also be improved by selecting only a soluble domain for expression.

Codon biasing



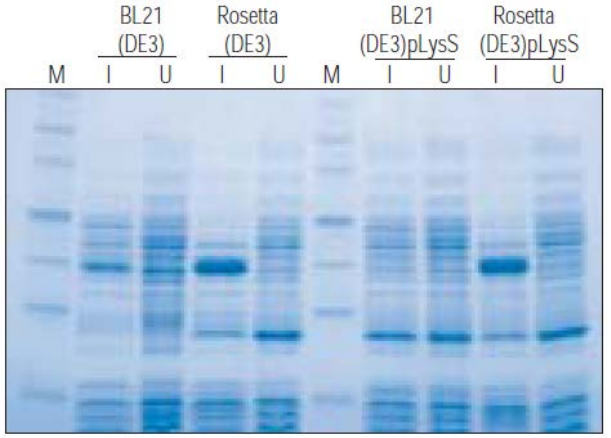
Codon biasing

In a particular organism a specific codon is preferably used to code a specific amino acid, despite of the other codons. Codon bias leads to translational stalling, premature termination, frameshifting and misincorporation

Optimization of the codon can be done by in vitro gene synthesis thus removing the rare codons.

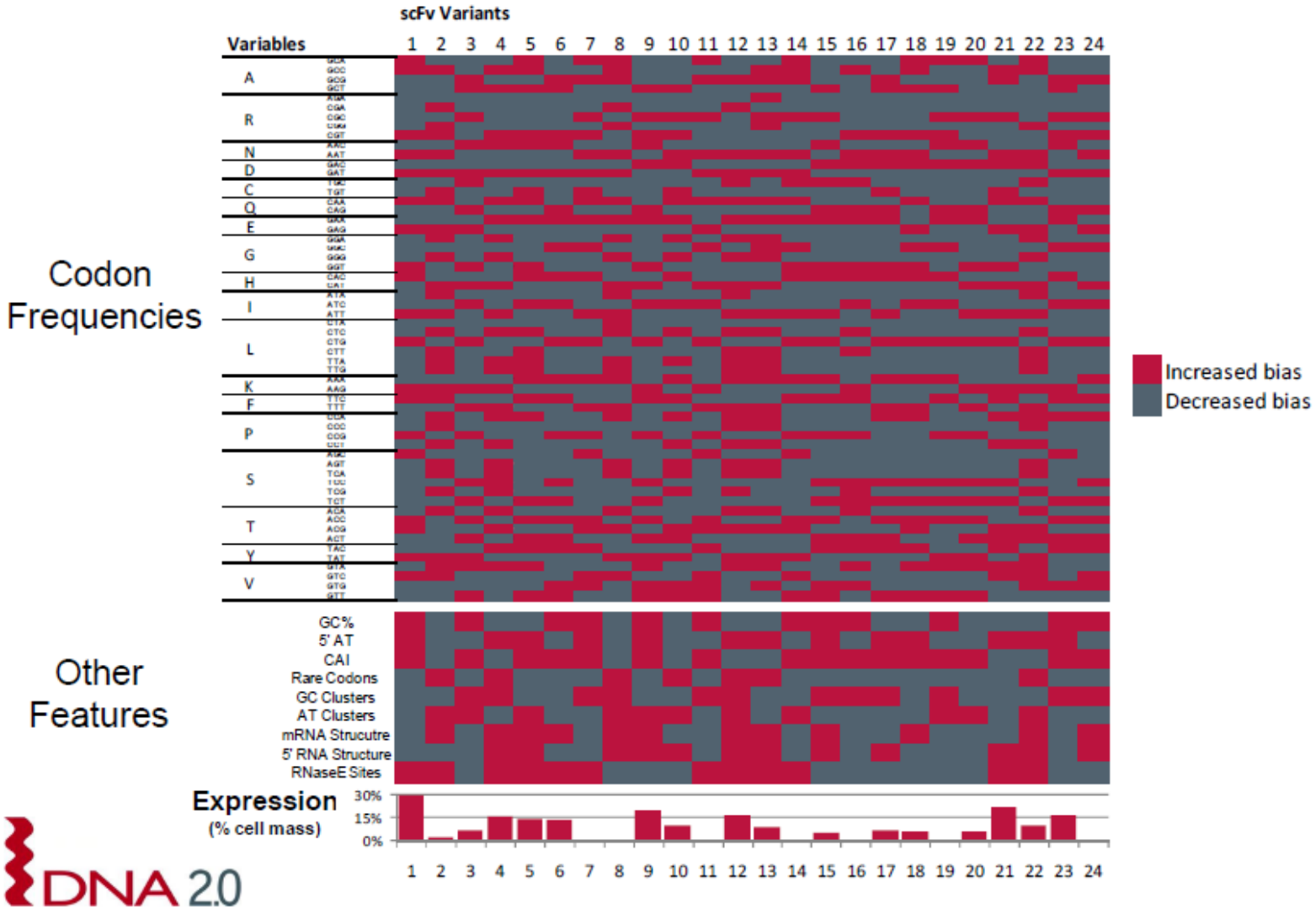
tRNAs genes coding for rare codons can be expressed to adapt the tRNA pool to the demand: chromosomic integration or compatible plasmids

Codon	Human	Drosophila	E. coli
Arginine:			
AGA	22 %	10 %	1 %
AGG	23 %	6 %	1 %
CGA	10 %	8 %	4 %
CGC	22 %	49 %	39 %
CGG	14 %	9 %	4 %
CGU	9 %	18 %	49 %



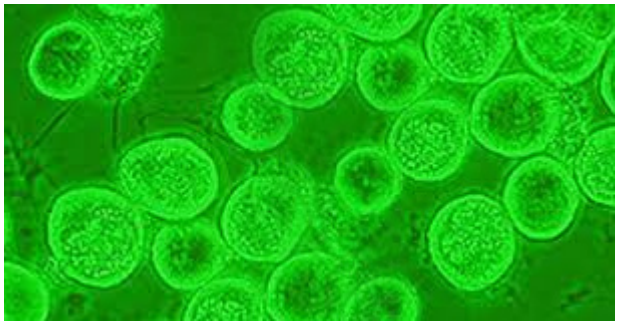
I Induced, U Un-induced

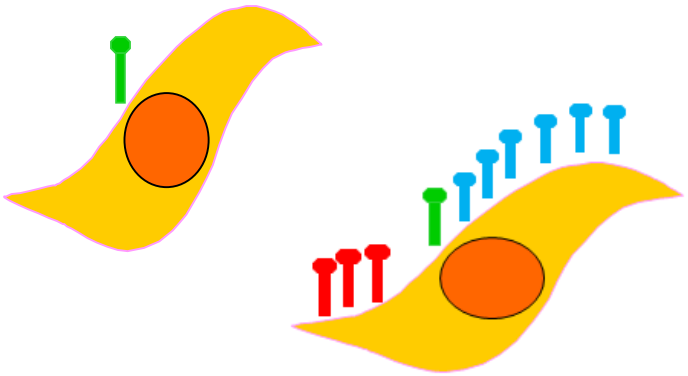
Sequence optimization

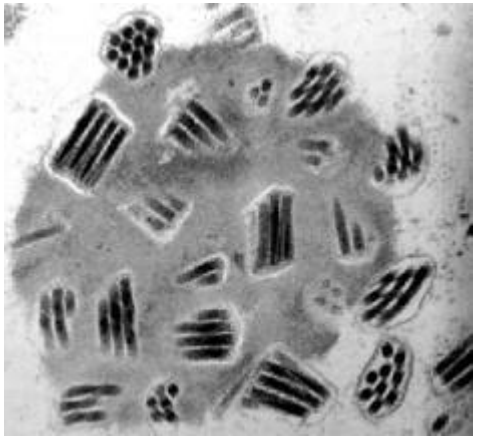


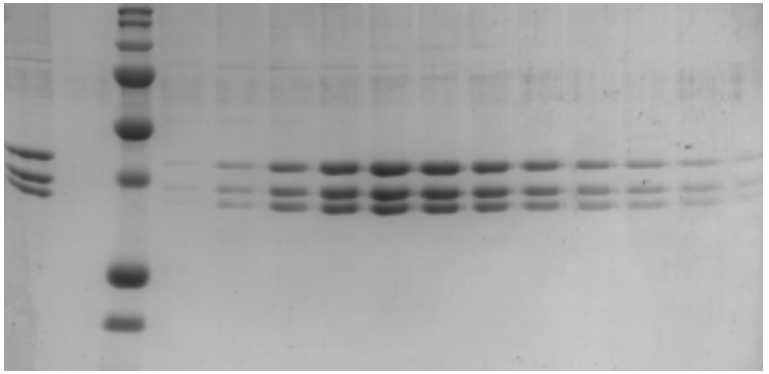
DNA 2.0

Co-expression in insect cells using the BVES

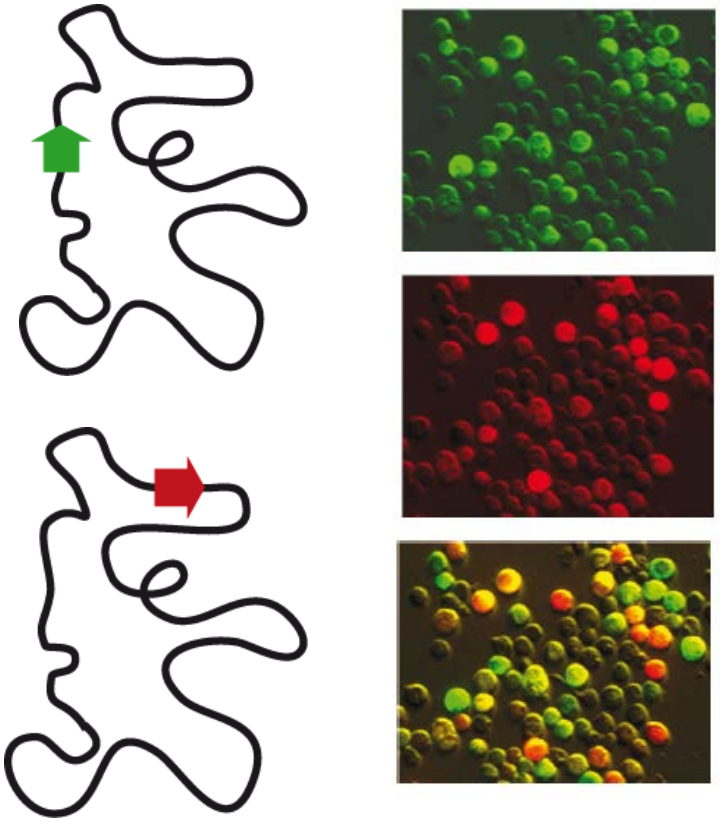




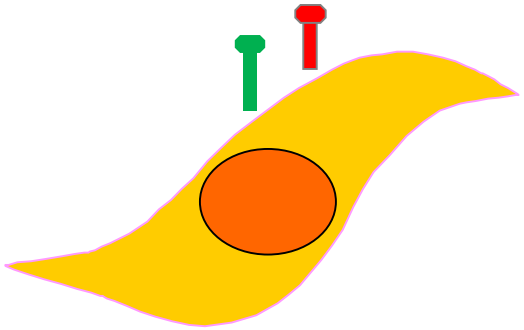




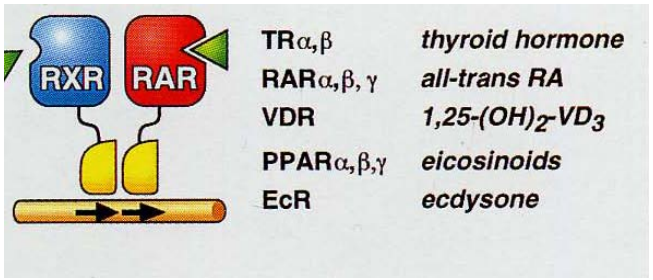
Co-infection: a simple way to co-express proteins

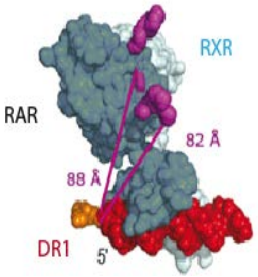


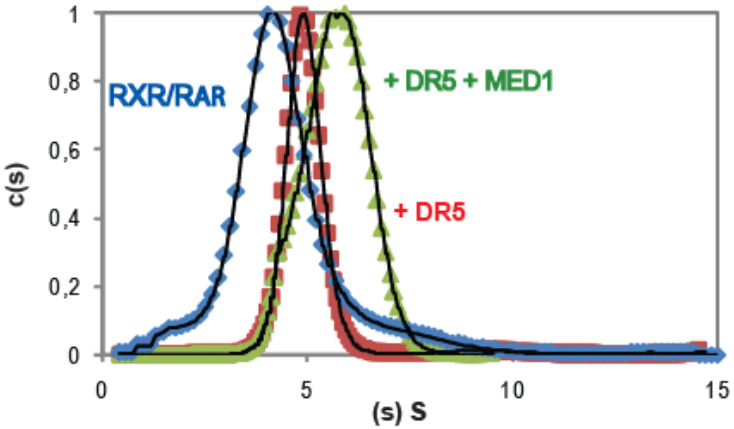
pH GFP+ p10 DsRed

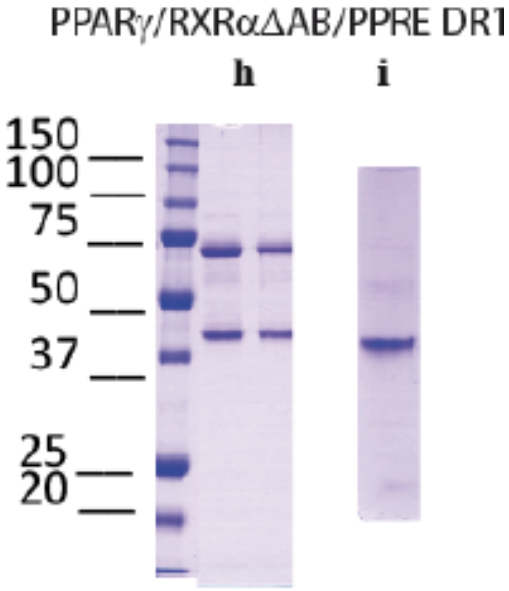


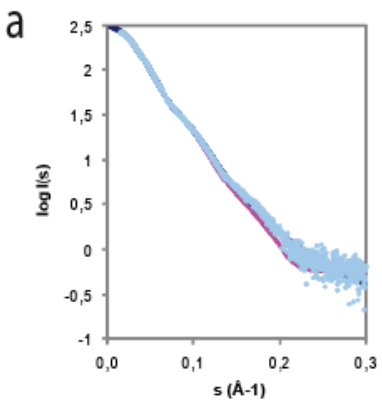
Production of nuclear hormone receptor complexes

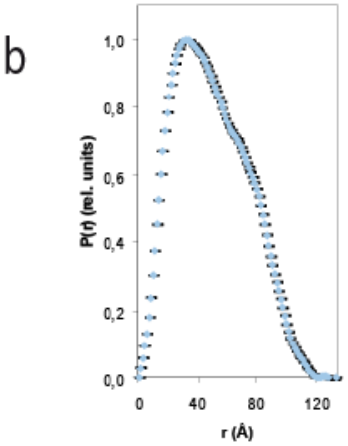










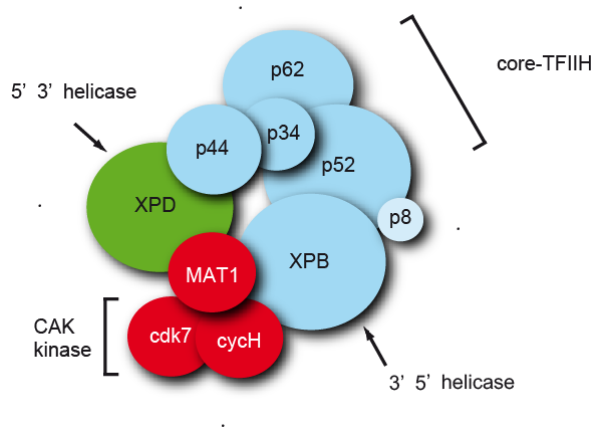


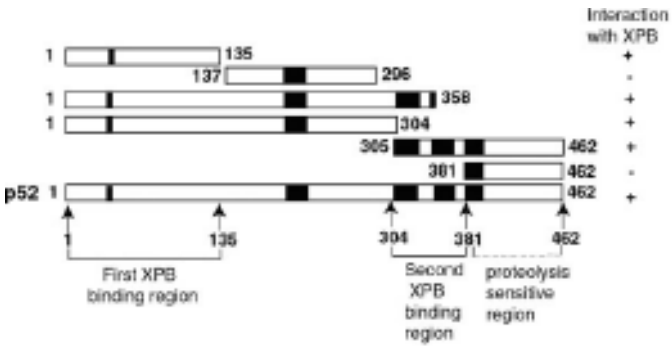
Rochel, et al. NSMB, 2011

Systematic dissection of protein-protein interactions within a complex

Generate two sets of n viruses:typically the first with a FLAG epitope and the second with an 6His tag

Test all combinations of pair-wise interactions (Flag-protein x/His-protein y)



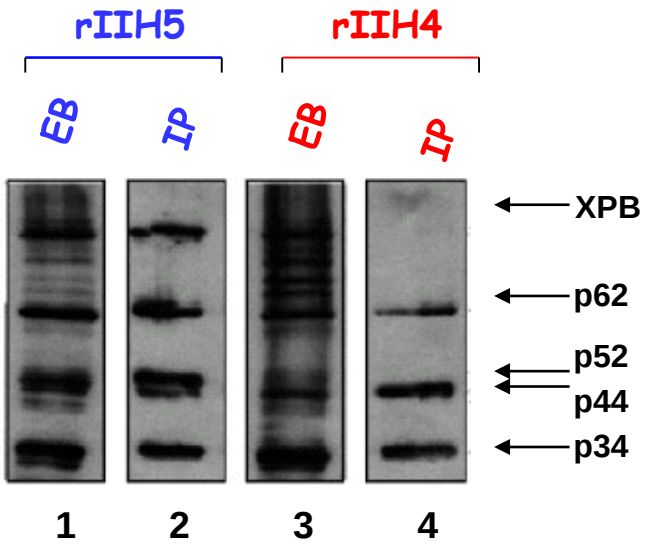


His-protein

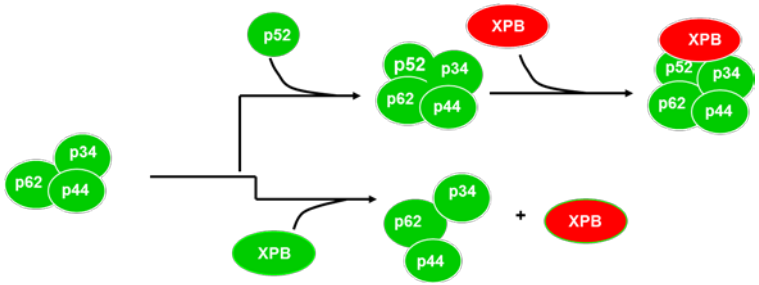
Flag-protein

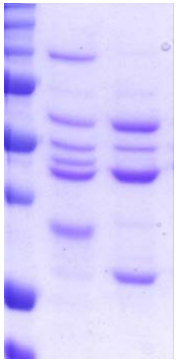
	p34	p44	p52	p62	XPB	XPD	MAT1
p34							
p44							
p52							
p62							
XPB							
XPD							

Systematic dissection of protein-protein interactions: deletion analysis



Analysis of the protein interaction network
Identification of key regulatory interactions



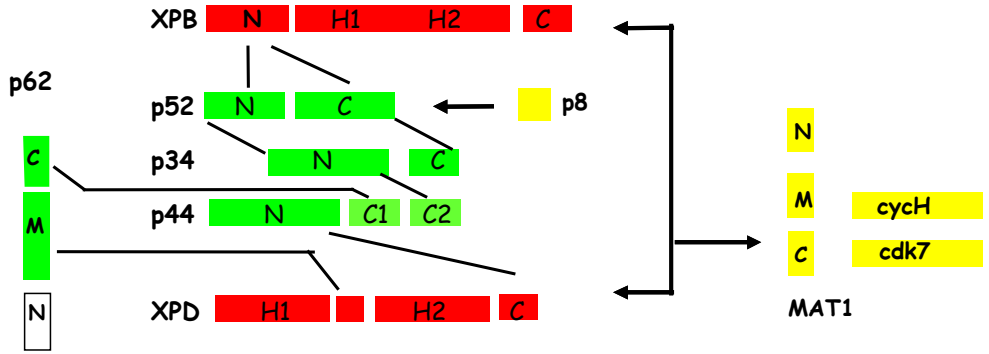


XPB

P62
*
P52,p44

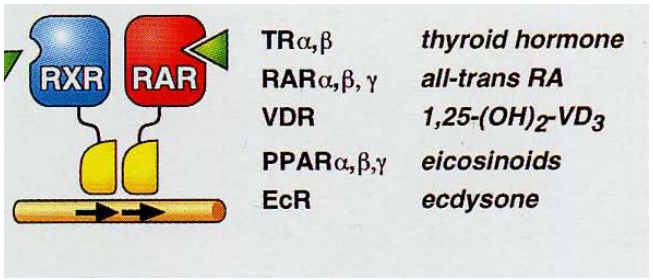
P34

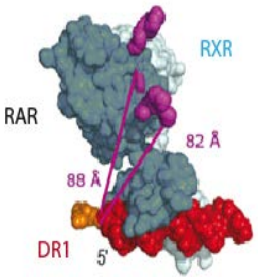
P34(1-233)

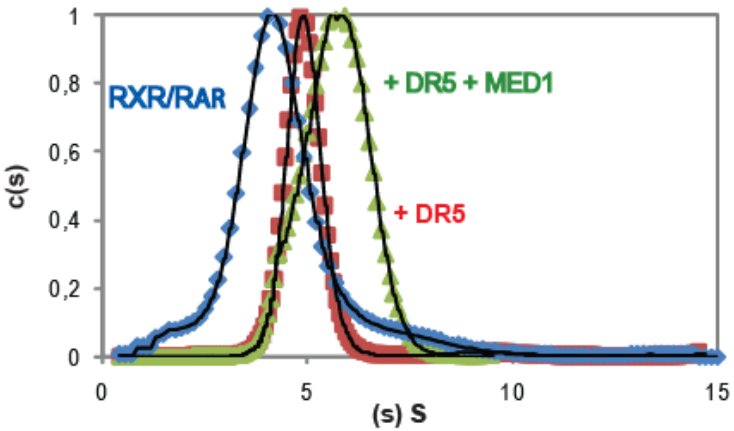


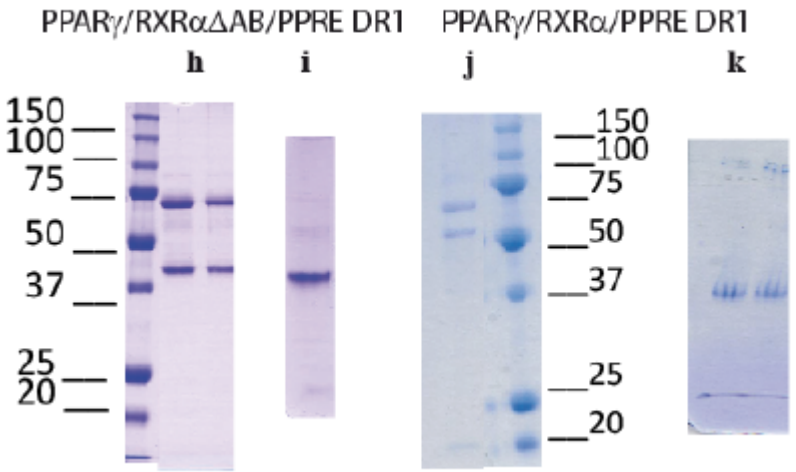
Jawahri at al. 2002, Radu et al, in prep

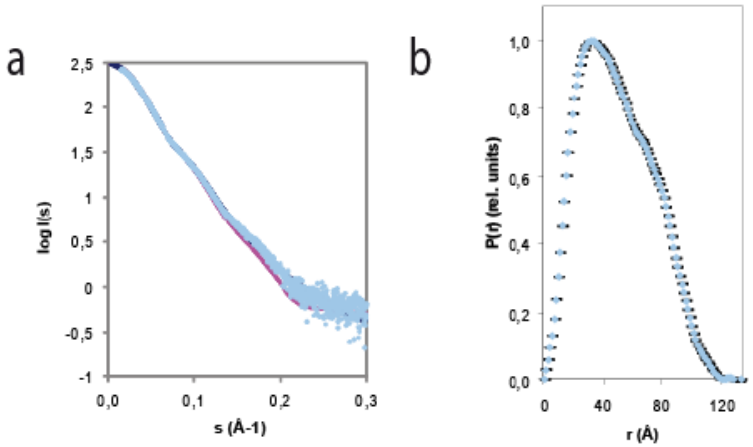
Production of nuclear hormone receptor complexes





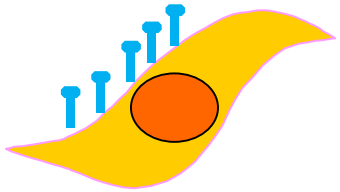


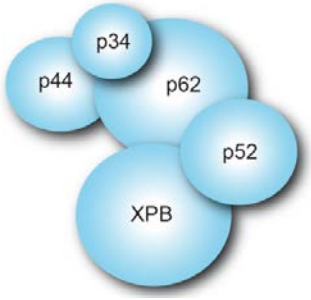




Rochel, et al. NSMB, 2011

Co-infection with multiple viruses for reconstitution of complexes



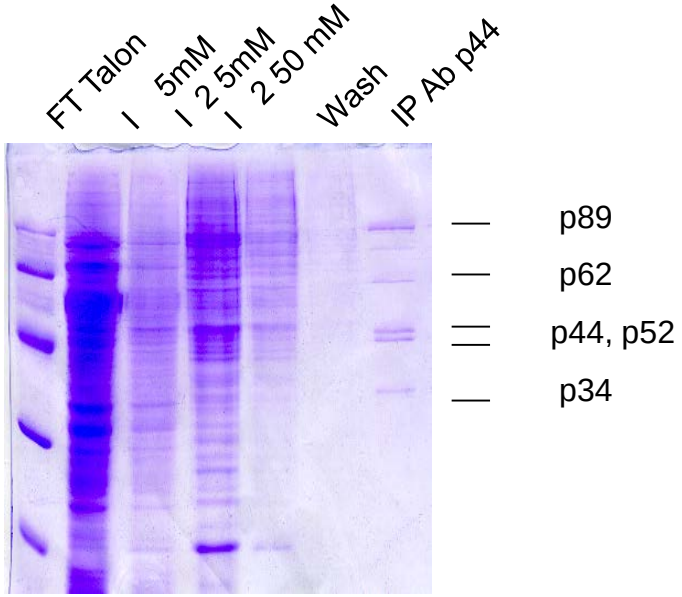


Cell extract

IMAC

Affinity purification

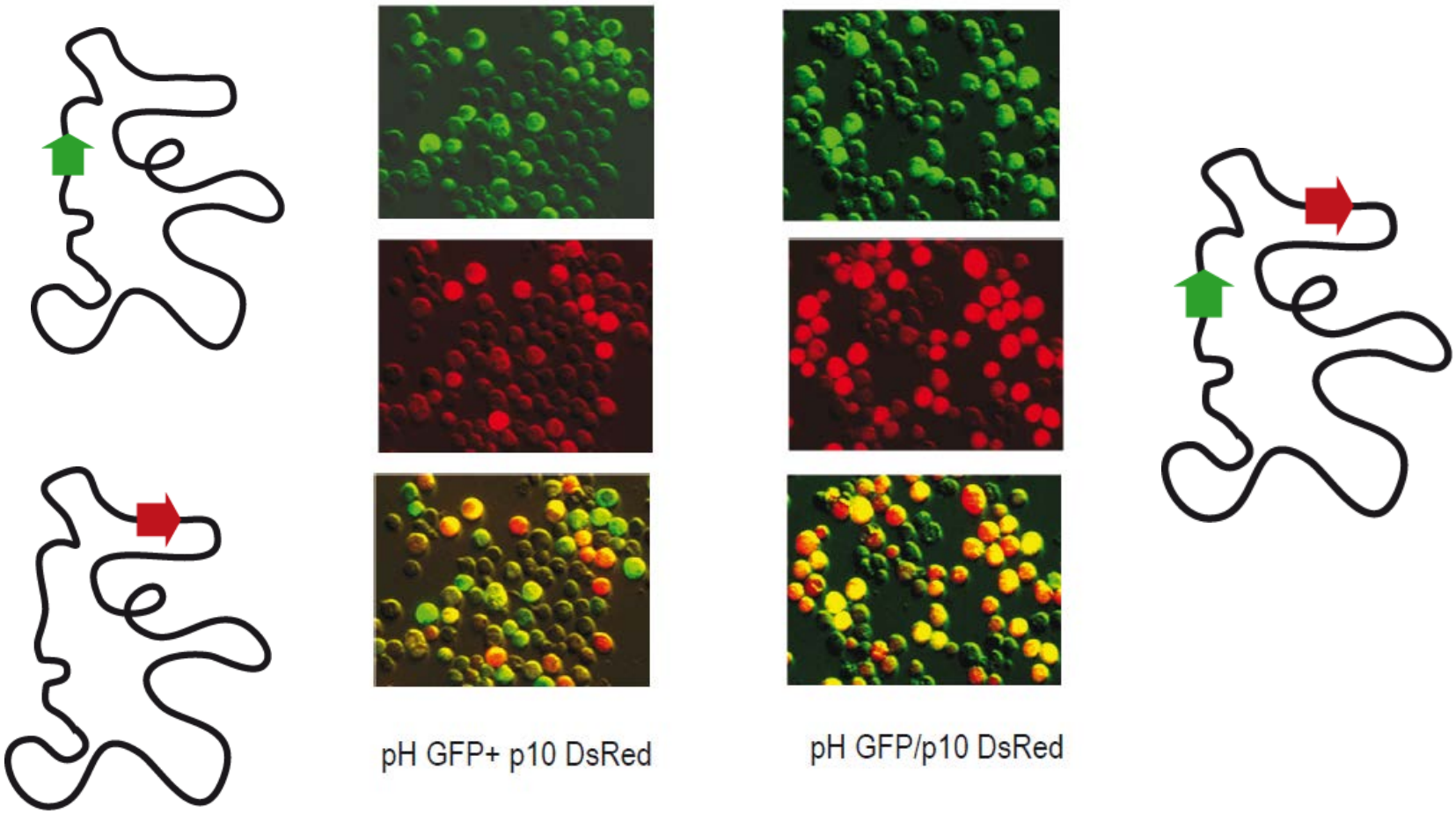
Size exclusion

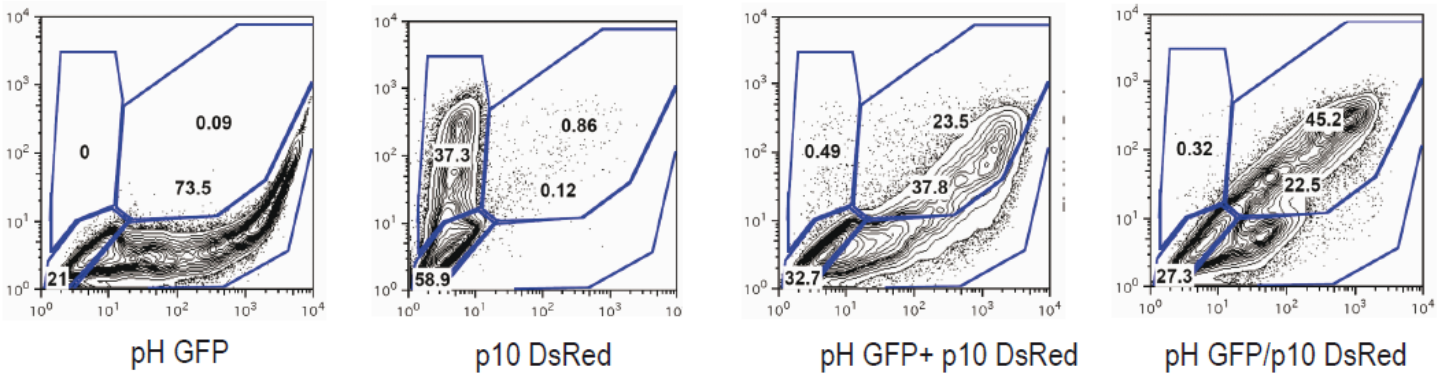


η # 50-100 μg/L

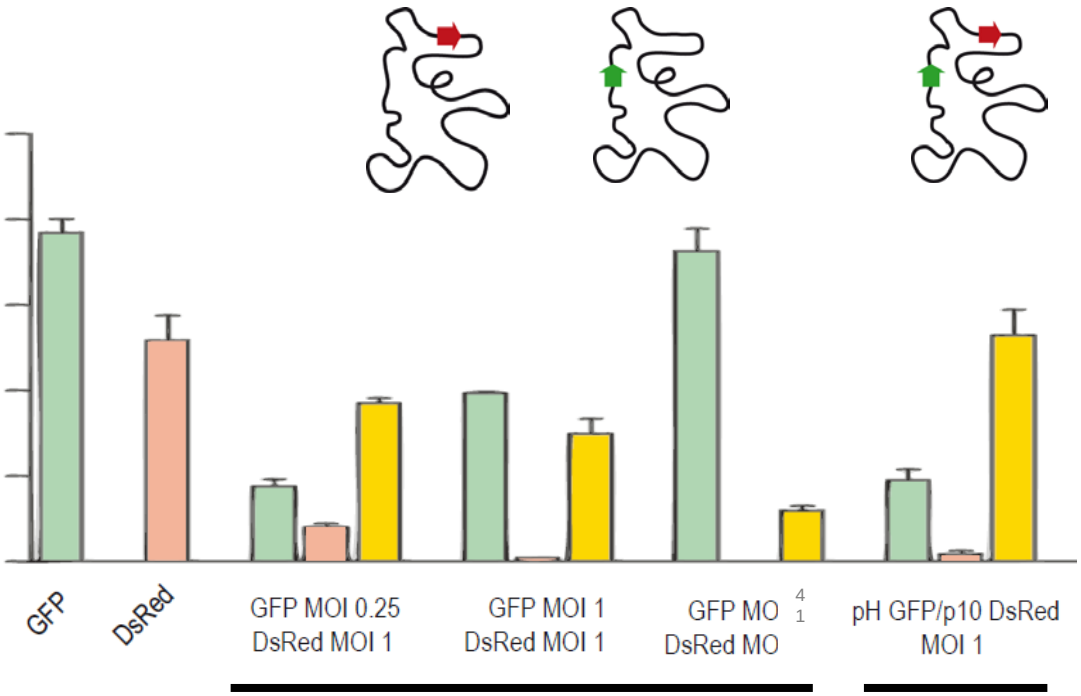
Low yields, labour intensive, poor reproducibility

Co-infection vs Multigene expression





Controls **Co-infection** **Multigene expression**

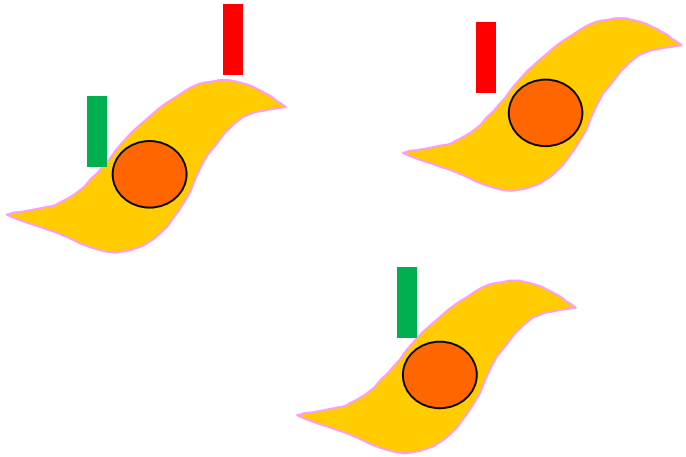


Idem for other MOIs

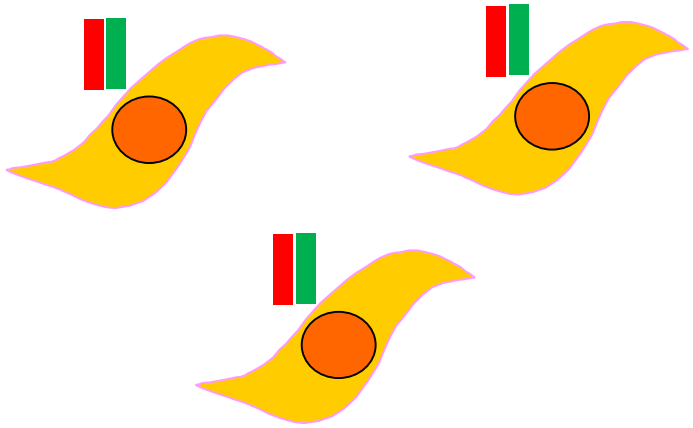
Co-infection vs Multigene expression

Two viruses encoding a single gene each

A single virus encoding the two genes



Co-infection

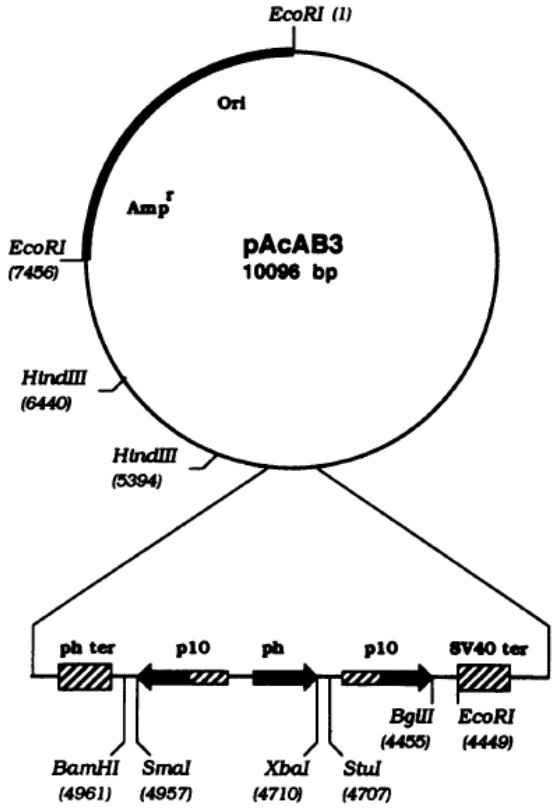


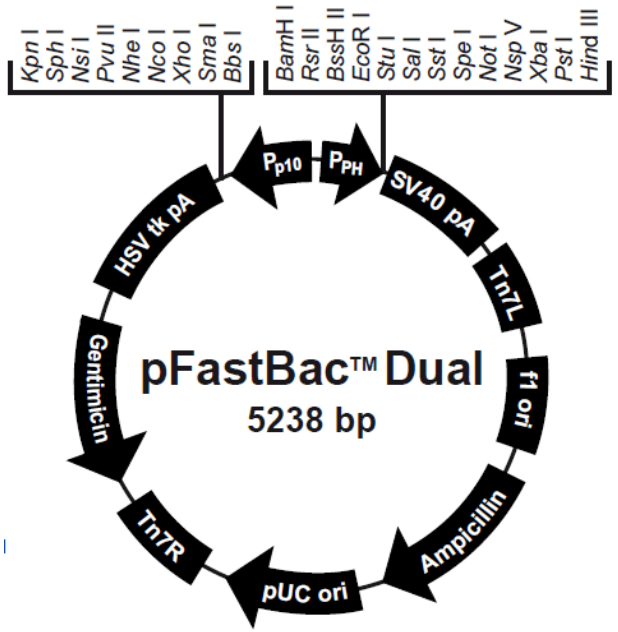
Infection by a single virus

Nucleic Acids Research, 1993, Vol. 21, No. 5 1219–1223

Development of baculovirus triple and quadruple expression vectors: co-expression of three or four bluetongue virus proteins and the synthesis of bluetongue virus-like particles in insect cells

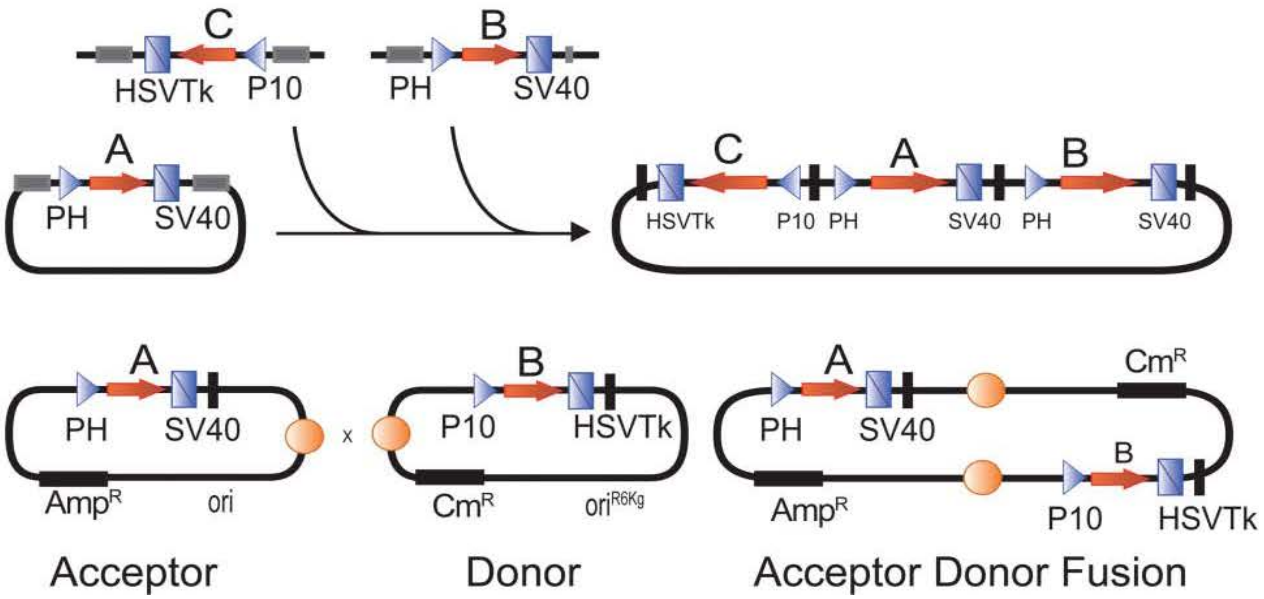
Alexander S.Belyaev[1] and Polly Roy[1,2,*]





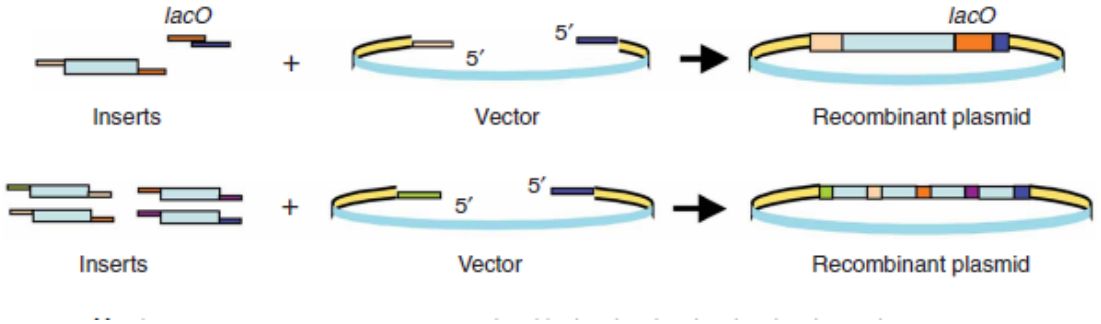
Multigene expression is conceptually trivial

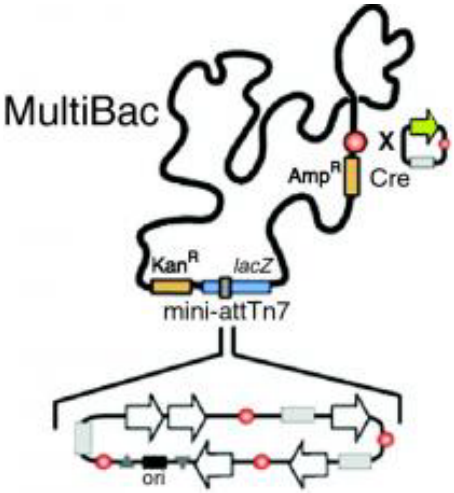
Clone genes of interest into individual expression units
Assemble the individual units into multigene transfer vectors



Adapted from Berger and Poterszman, Bioengineered. 2015

Tools to streamline cloning multigene expression transfer vectors were missing





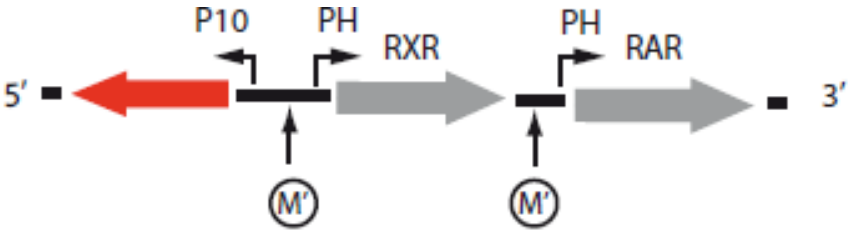
Progress in molecular and synthetic biology

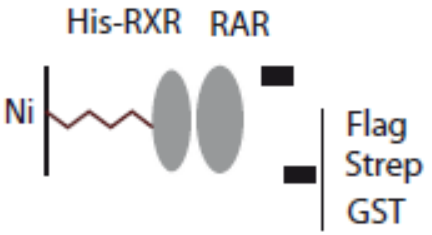
MultiBac, BigBac, MacroBac,….

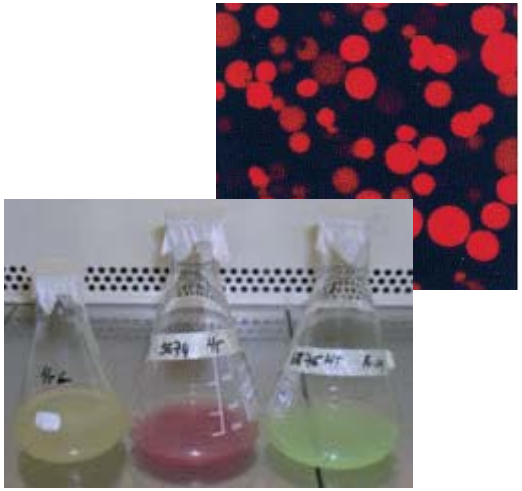
SLIC, Gibson, Infusion cloning and Cre-fusion

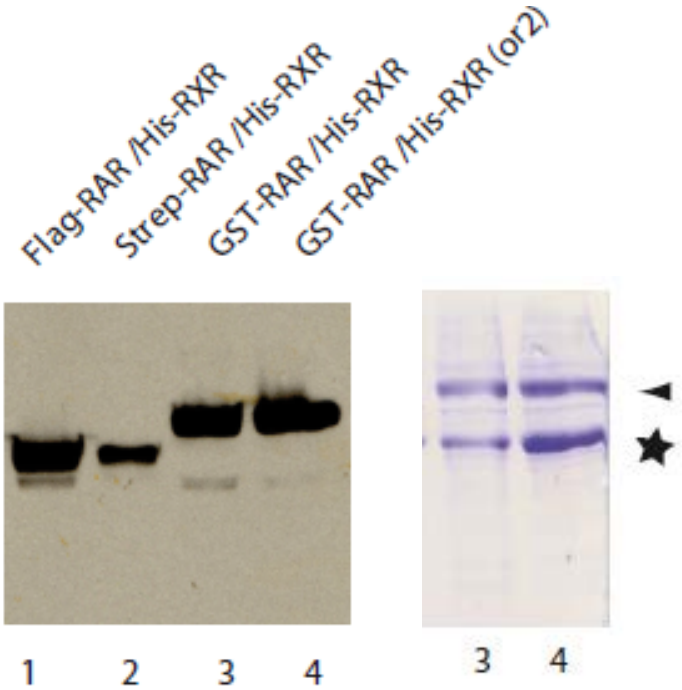
Imre BERGER, Bristol, UK

Insertion of an expression cassette into the multiplication module









Cre-LoxP recombination *in vitro*

Cre recombinase binds to the loxP sites on both the donor vector and the acceptor vector, cleaves the DNA, and covalently attaches itself to the DNA which leads to strand exchange and concatenation.

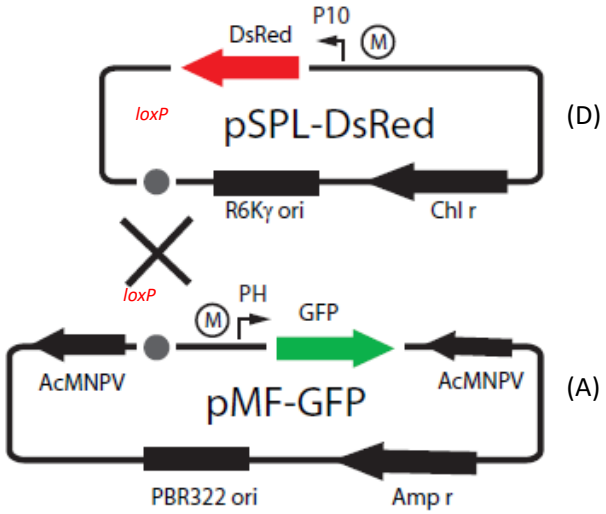
loxP sequence

8-bp overlap region

ATAACTTCGTATA GCATACATTATA CGAAGTTAT

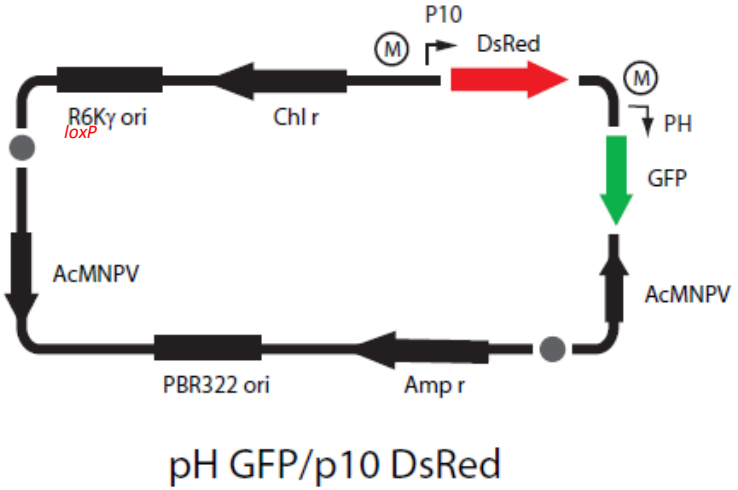
Inverted region Inverted region

- 34-bp recognition sequence
- 13-bp inverted repeats flanked by 8-bp spacer region

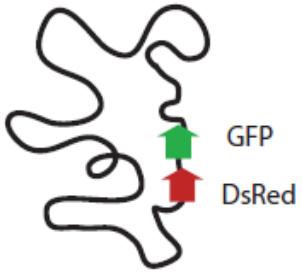


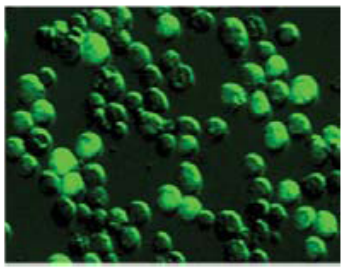
(D)

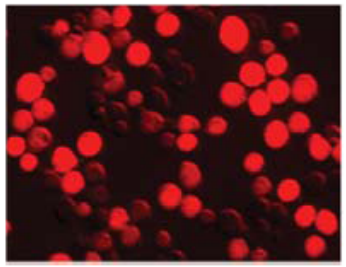
(A)

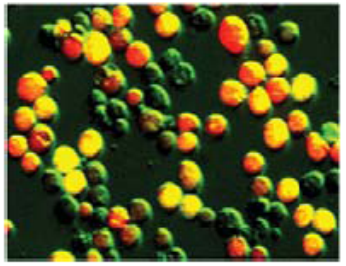


pH GFP/p10 DsRed



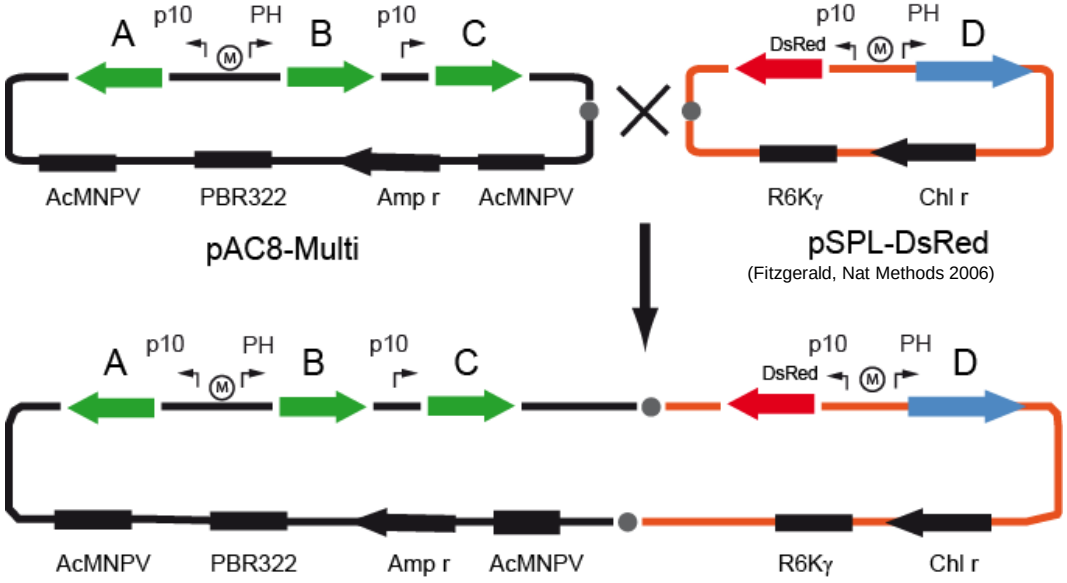






pSPL Multibac, (Fitzgerald, Nat Methods 2006)

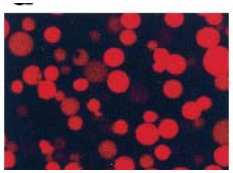
Expression of ternary and quaternary complexes with a single virus: a problem of DNA synthesis



pBac8 Multi-pSPL Fusion

A: strep-tag
D: Flag-tag

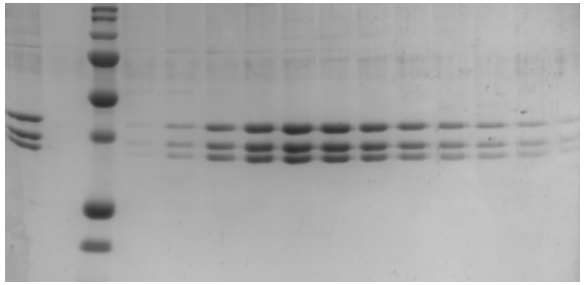
Flag-affinity



Elution with
Competitor-peptide

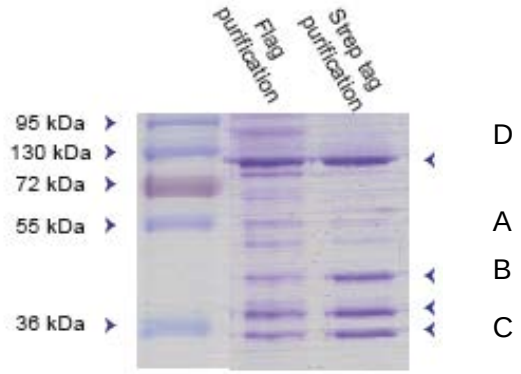
Strep-affinity

Elution with
Desthiobitin



A
B
C

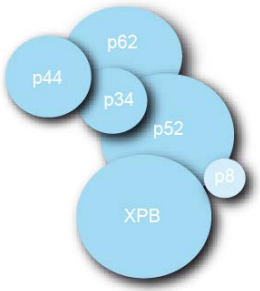
Ternary complex



Quaternary complex

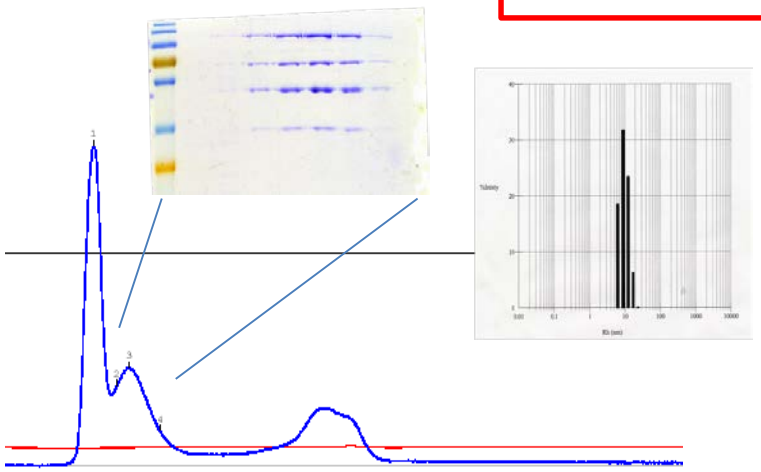
Production of core-TFIIH with a single virus

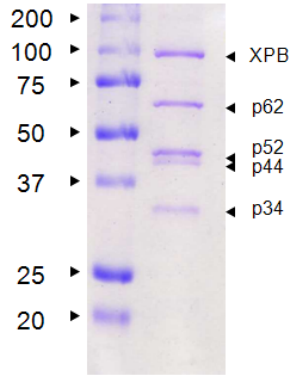
- 6 subunits: XPB, p62, p52, p44, p34, p8/TTDA (+ DsRed)
- Yield : 0.5 mg/L

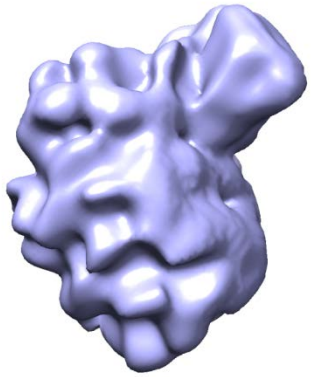


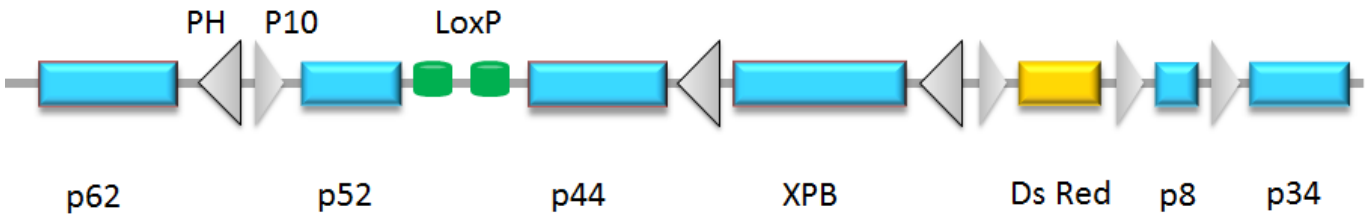
MS-cross linking, Cryo-EM and crystallization

Screening for XPB inhibitors



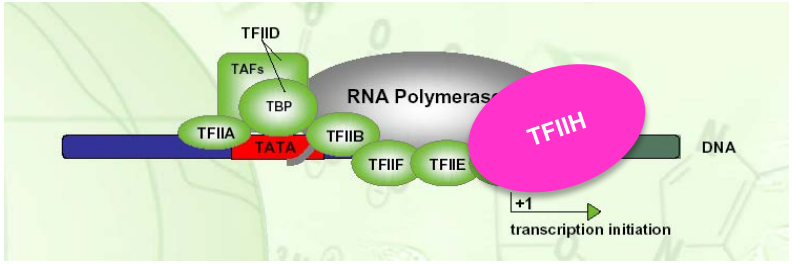


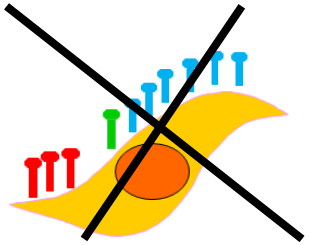


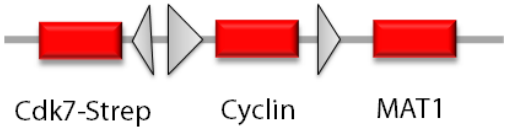


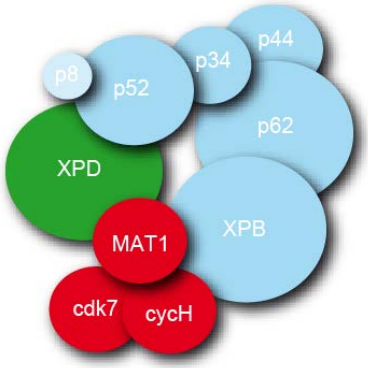
16 kbp

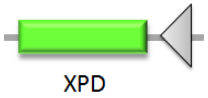
Reconstitution and in vitro assays

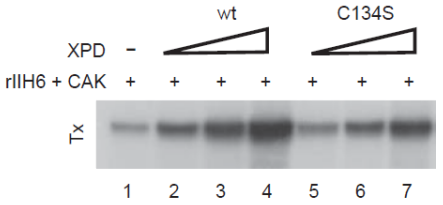


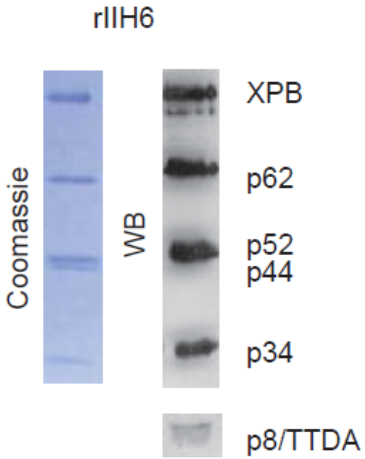












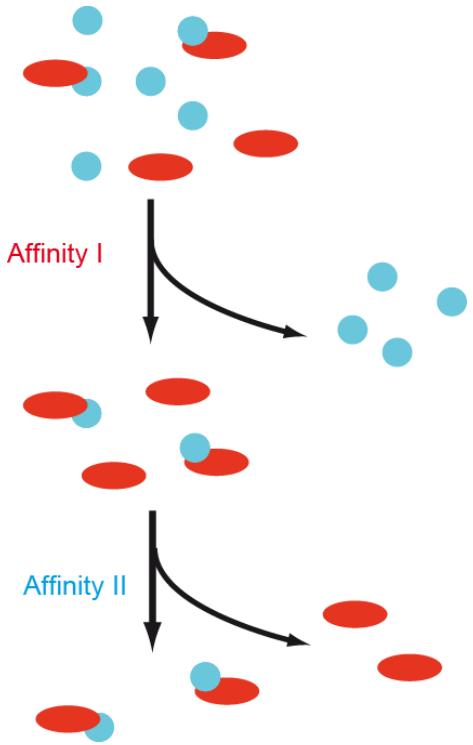
(Abdulrahman et al, PNAS 2013, Kupper et al., Plos Biology 2014, Radu NAR 2017)

Tandem Affinity Purification: Nature and position of the affinity tag

A single affinity step is usually not sufficient.

Sequential affinity steps that will select for the presence of two subunits

Position of the tag maters

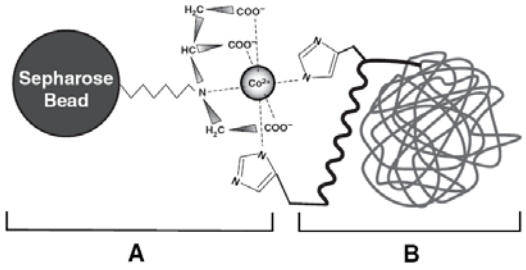


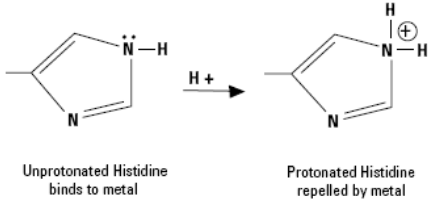
Tag	Resin	Elution	Cost/10mg
CBP	Calmodulin affinity	EGTA	181 €
TRX	Thiobond resin	ß-mercapto ethanol	n.a.
ProA	IgG Sepharose G	n.a.	275 €
FLAG	Anti flag M2 affinity gel	FLAG peptide	2343 €
GST	Glutathione sepharose 4B	Glutathione	41 €
Strep tag II	Streptactin sepharose	Desthiobiotin	67-134 €
His6	TALON Affinity	Imidazole	8-23 €

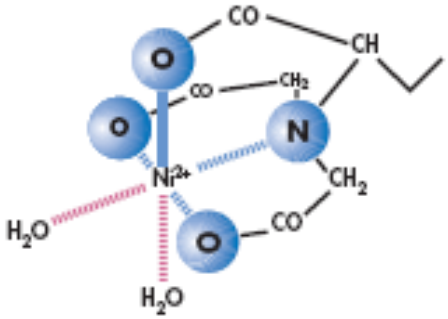
Abdulrahman, Anal Bioch, 2009
Koleschnikova et al, in prep

His-tag

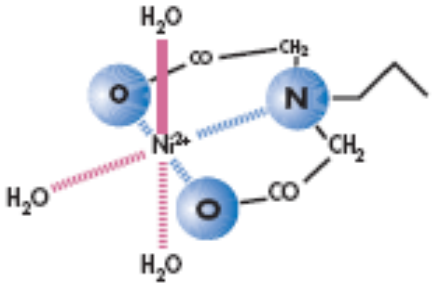
Immobilized Metal Chelate Affinity resin







Ni-NTA



Ni-IDA

Table I: Histidine Tags

Tag	Amino acids
6xHis	His – His – His – His – His – His
6xHN	His – Asn – His– Asn – His – Asn– His – Asn – His – Asn – His – Asn
HAT	Lys – Asp – His – Leu – Ile – His – Asn – Val – His – Lys – Glu – His – Ala – His – Ala – His – Asn – Lys

10xHis

Ni, Co, Fe

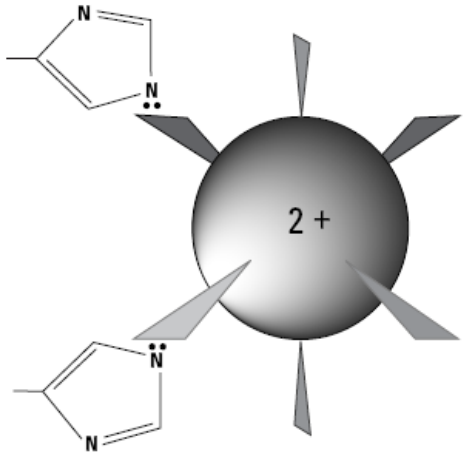
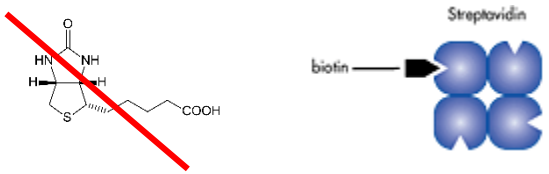
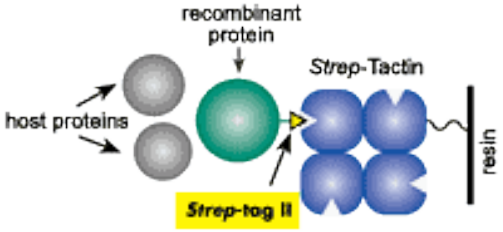


Figure 3. Binding of histidines to the TALON® Resin metal ion. Under conditions of physiological pH, histidine binds by sharing imidazole nitrogen electron density with the electron-deficient orbitals of the metal ion.

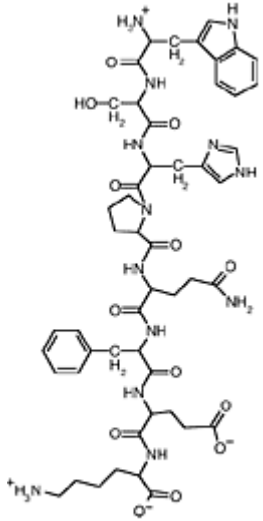
Strep tag-II

Derived from strepavidin-Biotin

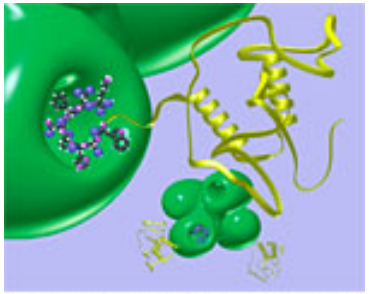




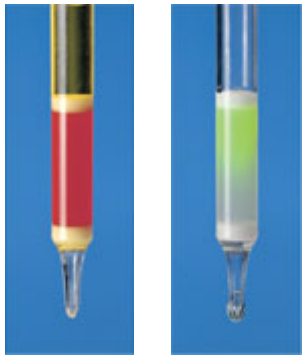
Elution with biotin analog: desthiobiotin or more recently Biotin (StrepTactin@XT)

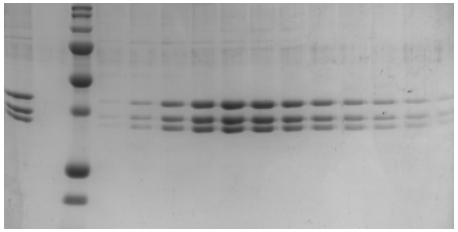


***Strep*-tag II**
NH_2 -WSHPQFEK-COOH

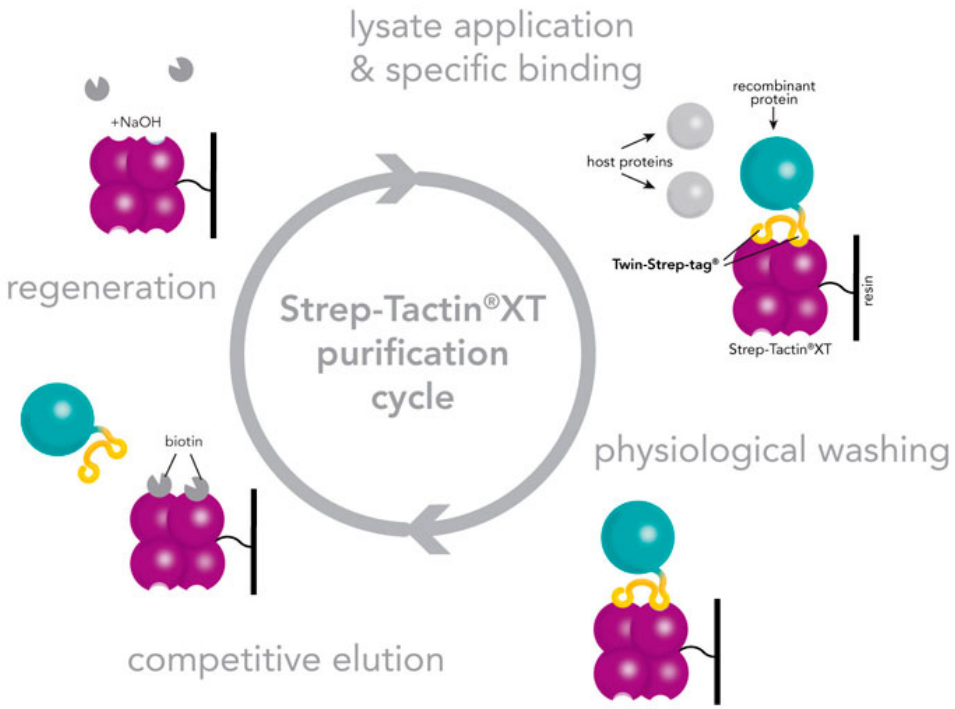


***Strep*-tag** protein is binding to a ***Strep*-Tactin** tetramer.





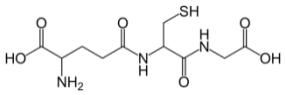
Strep-Tactin®XT purification cycle

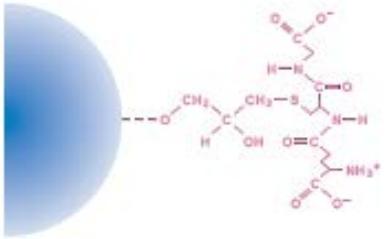


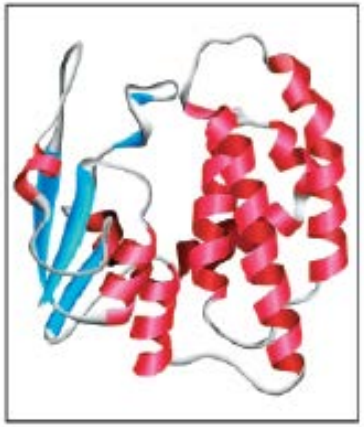
Elution with 50 mM Biotin

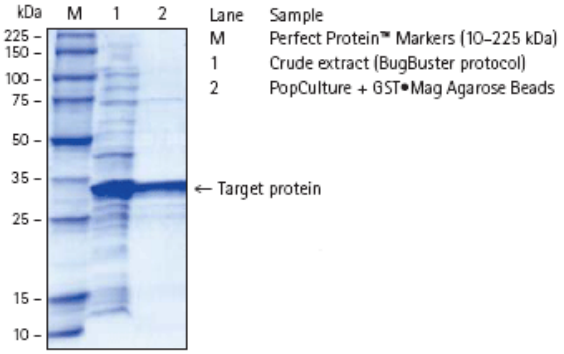
https://www.iba-lifesciences.com/tl_files/ProteinProductionAssays/3-Purification/Purification-Cycle-Strep-Tactin-XT-Twin-Strep-tag.jpg

Large tags can promote solubility and expression: GST, MBP

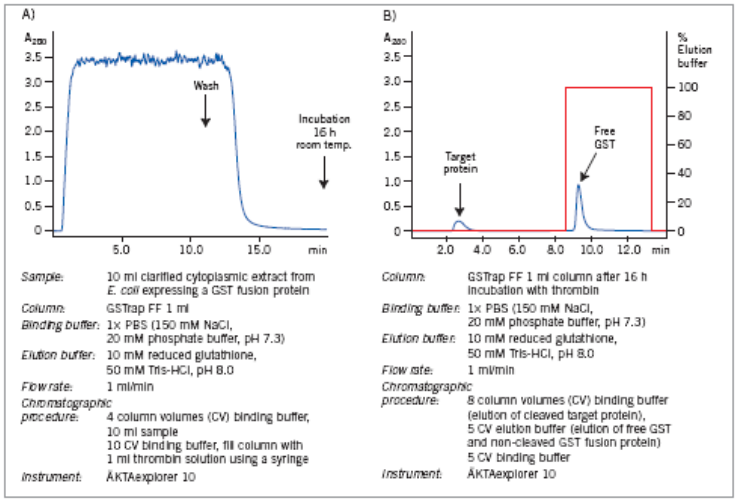








Lane	Sample
M	Perfect Protein™ Markers (10–225 kDa)
1	Crude extract (BugBuster protocol)
2	PopCulture + GST•Mag Agarose Beads

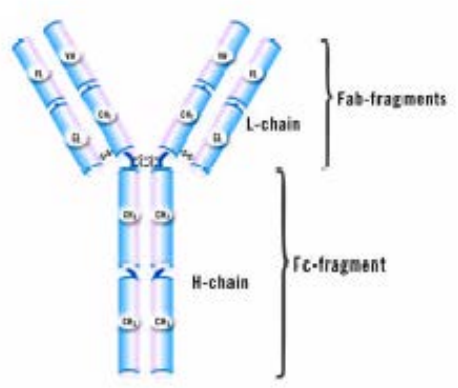


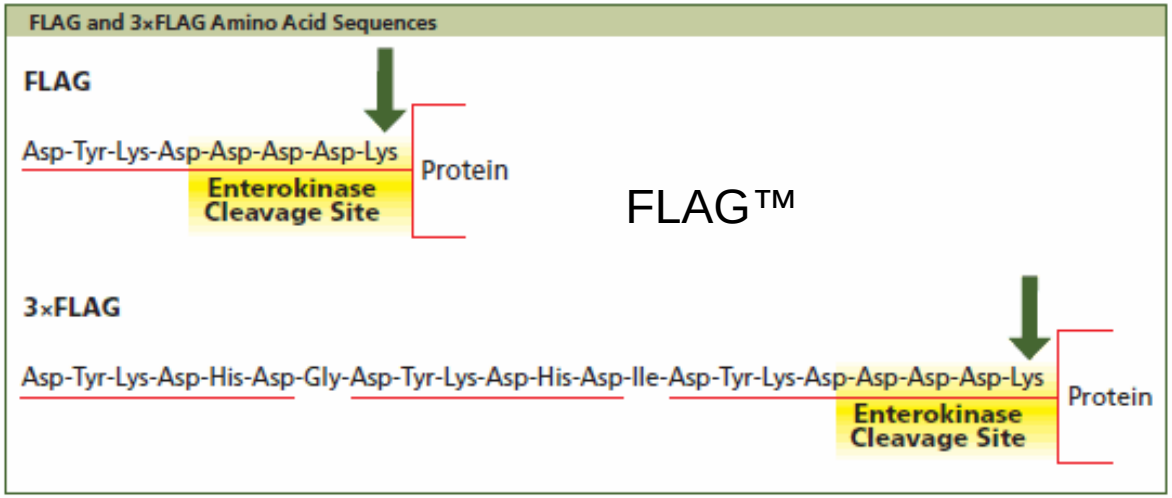
	A)	B)
Sample:	10 ml clarified cytoplasmic extract from *E. coli* expressing a GST fusion protein	GSTrap FF 1 ml column after 16 h incubation with thrombin
Column:	GSTrap FF 1 ml	
Binding buffer:	1× PBS (150 mM NaCl, 20 mM phosphate buffer, pH 7.3)	1× PBS (150 mM NaCl, 20 mM phosphate buffer, pH 7.3)
Elution buffer:	10 mM reduced glutathione, 50 mM Tris-HCl, pH 8.0	10 mM reduced glutathione, 50 mM Tris-HCl, pH 8.0
Flow rate:	1 ml/min	1 ml/min
Chromatographic procedure:	4 column volumes (CV) binding buffer, 10 ml sample 10 CV binding buffer, fill column with 1 ml thrombin solution using a syringe	8 column volumes (CV) binding buffer (elution of cleaved target protein), 5 CV elution buffer (elution of free GST and non-cleaved GST fusion protein) 5 CV binding buffer
Instrument:	ÄKTAexplorer 10	ÄKTAexplorer 10

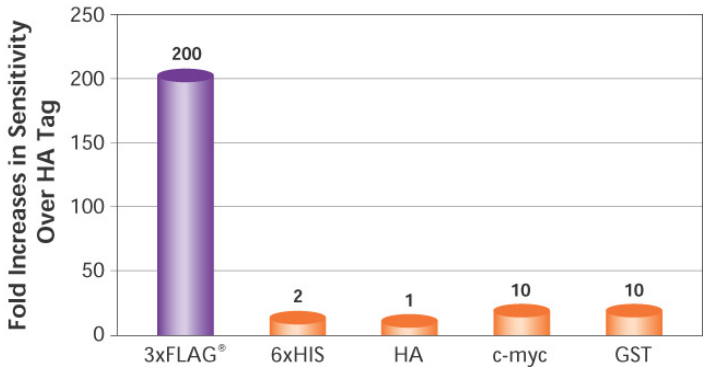
Fig 26. On-column thrombin cleavage of a GST fusion protein. A) Equilibration, sample application, and washing of a GST fusion protein on GSTrap FF 1 ml were performed using ÄKTAexplorer 10. After washing, the column was filled by syringe with 1 ml of thrombin (20 U/ml 1× PBS) and incubated for 16 h at room temperature. B) GST-free target protein was eluted using 1× PBS. GST was eluted using 10 mM reduced glutathione. The GST-free target protein fraction also contained a small amount of thrombin (not detectable by SDS-PAGE; see Fig 27, lane 6). The thrombin can be removed using a HiTrap Benzamidine FF (high sub) column.

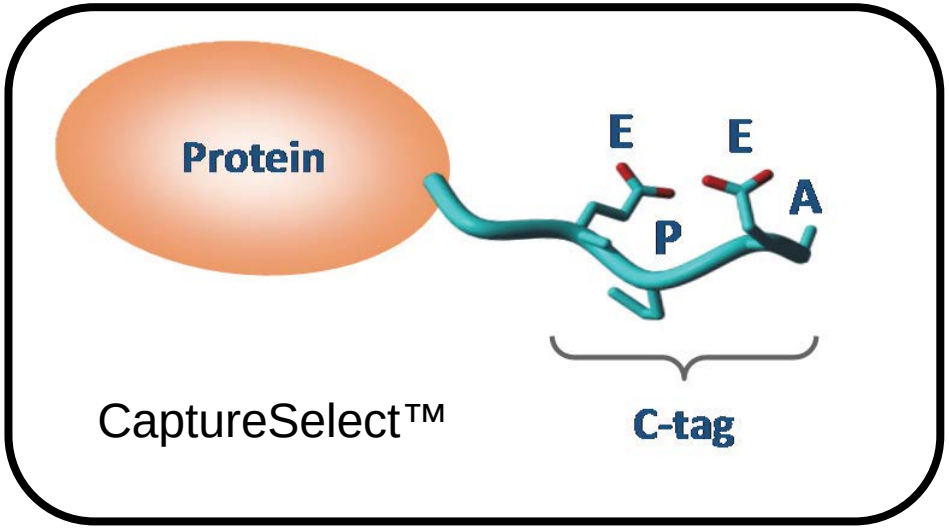
Source: See Figure 27.

FLAG and Capture Systems

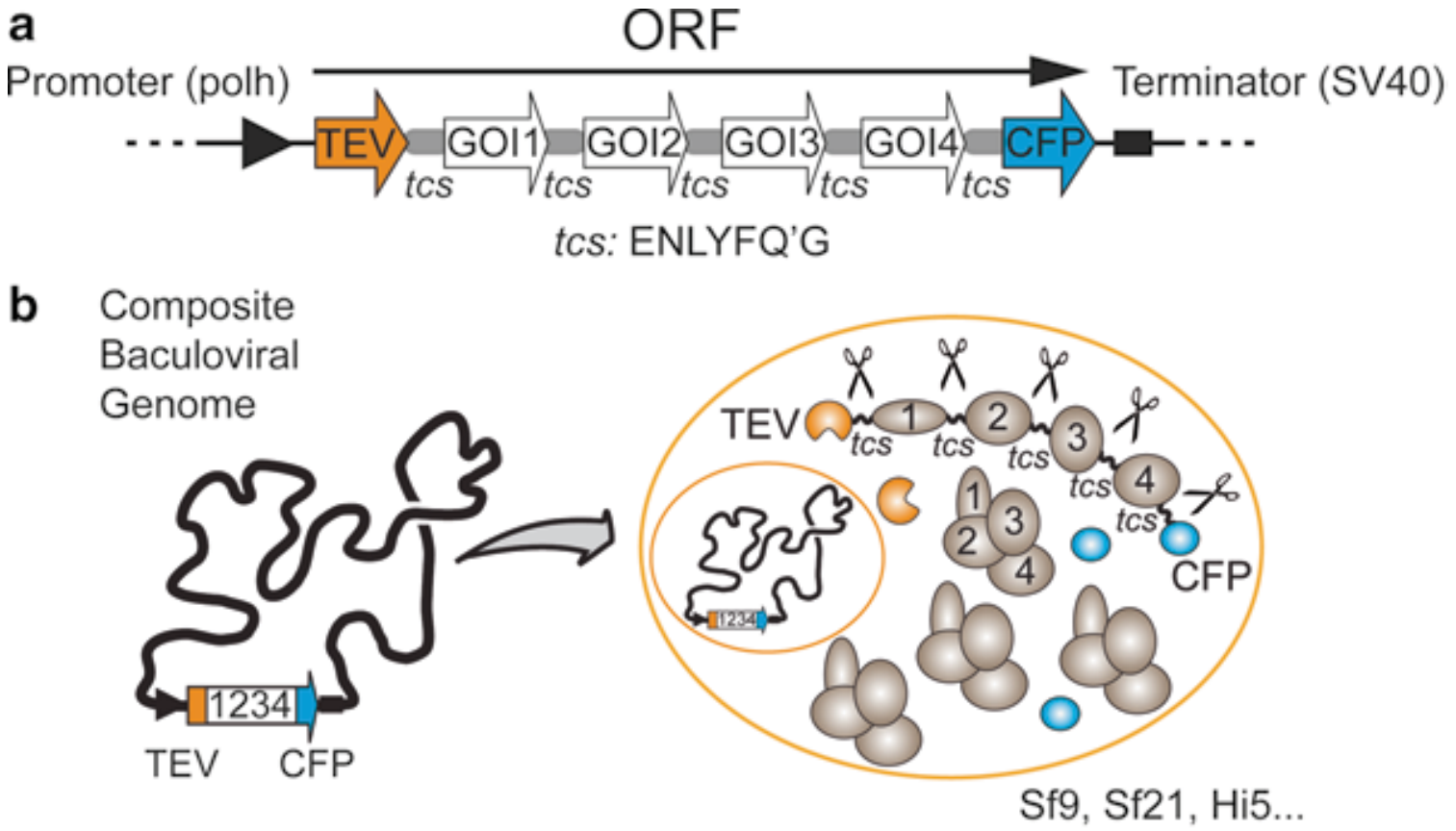








Synthetic Polyproteins: an option to control the stoichiometry



Nie, Methods Mol Biol, 2014

Production of Multiprotein complexes for biochemical & structural biology applications

Recombinant production

Bottom Up

Purification from endogenous sources

Top Down

Purification of native entities as reference

Sample preparation for biochemical an/or structural studies

Tandem Affinity Purification

Fusion of a Tandem-Affinity Purificiation (TAP) tag to the target protein which is expressed at physiological levels

Purification of the target protein in native conditions allows retrieval of associated partners

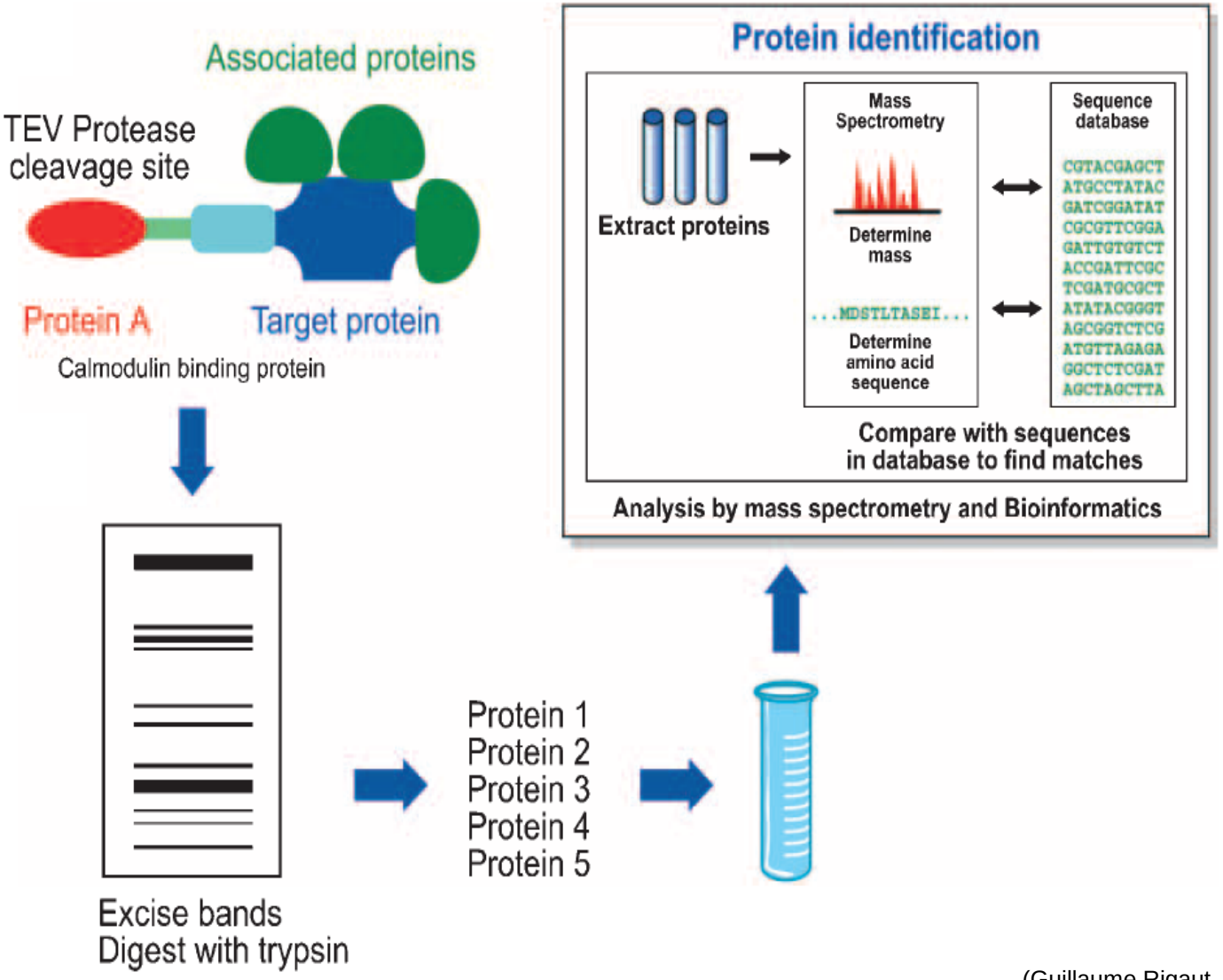
Rapid analysis of complexes without prior knowledge of the complex composition, activity, or function

Developed for large-scale studies

Ability to purify low abundant complexes from tissue/cell cultures

Sample preparation from native sources for structural/functional studies

TAP-MS



(Guillaume Rigaut, *et al.* 1999)

(Arnaud Droit, *et al.* 2005)

TAP-TAG in yeast

(Gavin, 2002)

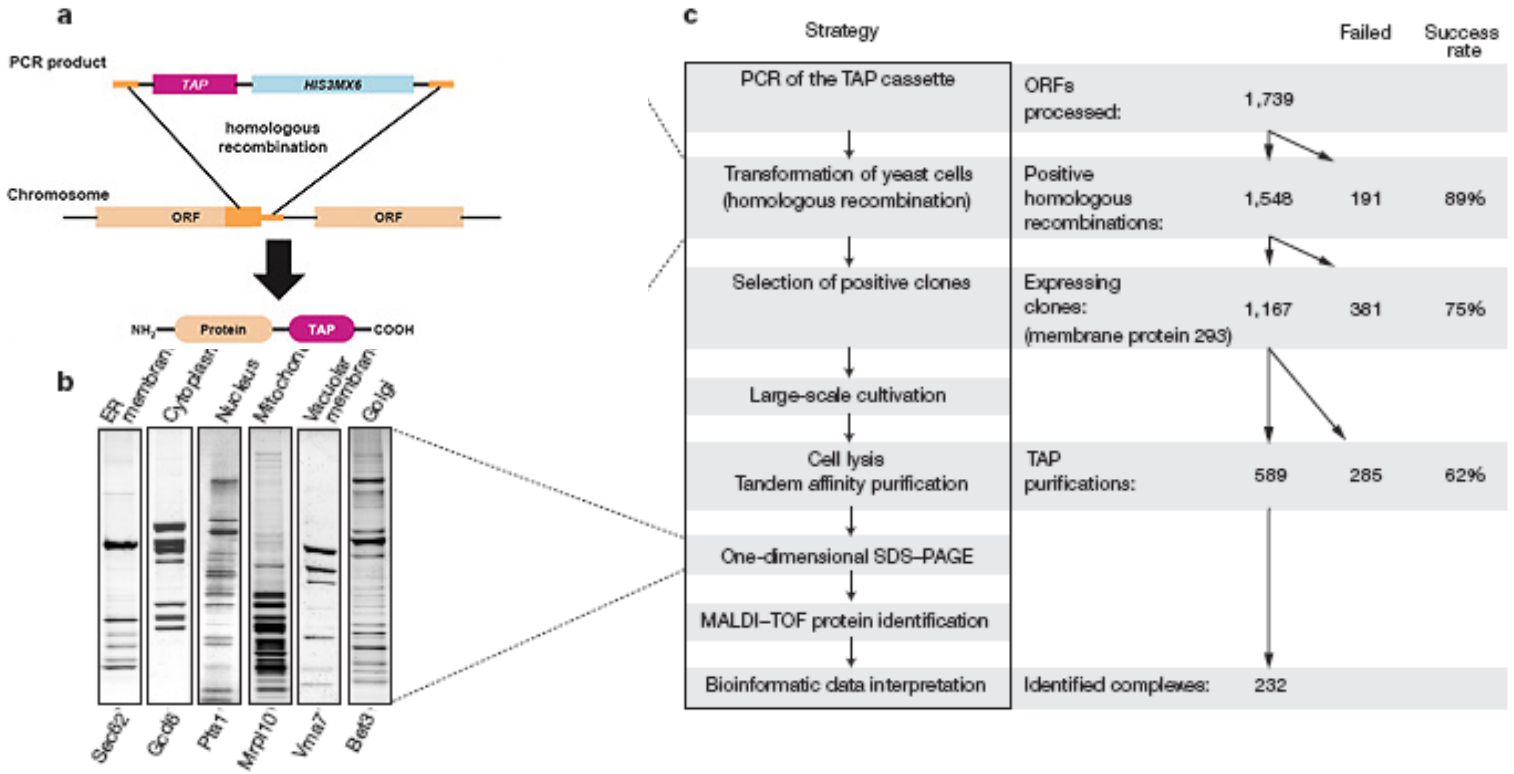
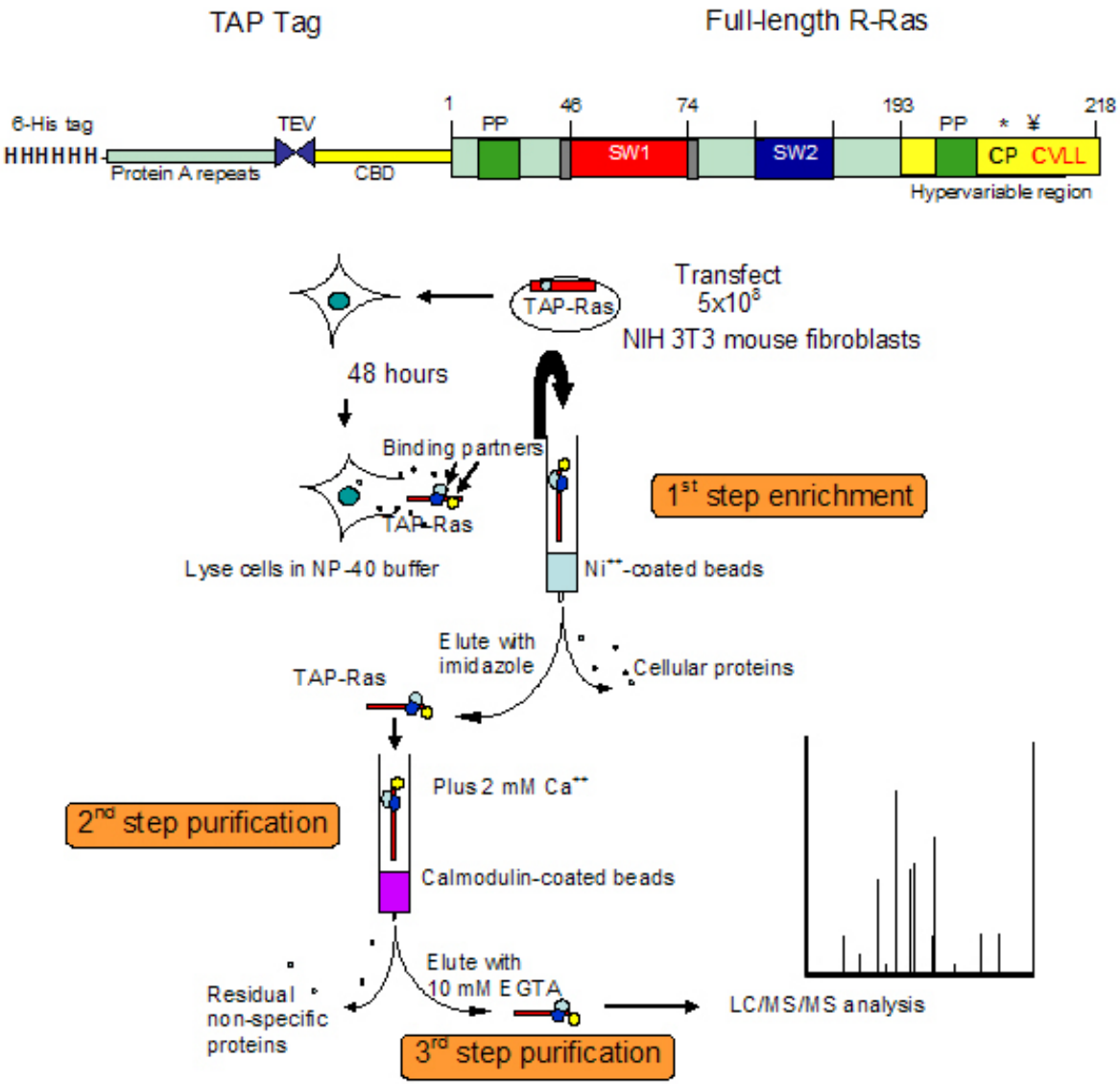


Figure 1 Synopsis of the screen. **a**, Schematic representation of the gene targeting procedure. The TAP cassette is inserted at the C terminus of a given yeast ORF by homologous recombination, generating the TAP-tagged fusion protein. **b**, Examples of TAP complexes purified from different subcellular compartments separated on denaturing protein gels and stained with Coomassie. Tagged proteins are indicated at the bottom. ER, endoplasmic reticulum. **c**, Schematic representation of the sequential steps used for the purification and identification of TAP complexes (left), and the number of experiments and success rate at each step of the procedure (right).

TAP-TAG in Mammalian cells



Engineering of Mammalian cell lines

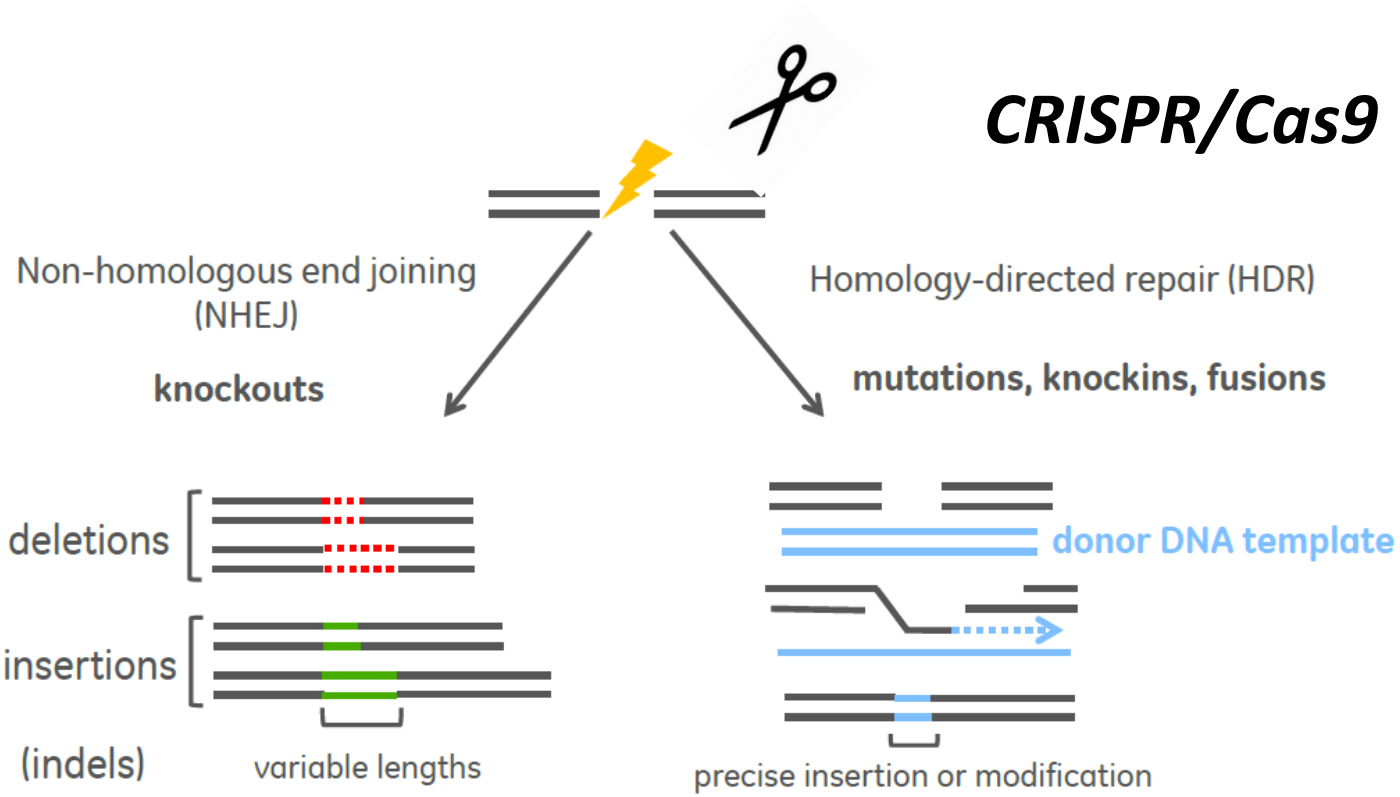
Transient expression

Random integration (antibiotic selection)

Site specific integration by RMCE

DNA Editing by repair of double-strands breaks

Genome editing by repair of double stranded breaks



Targeted insertion/deletion

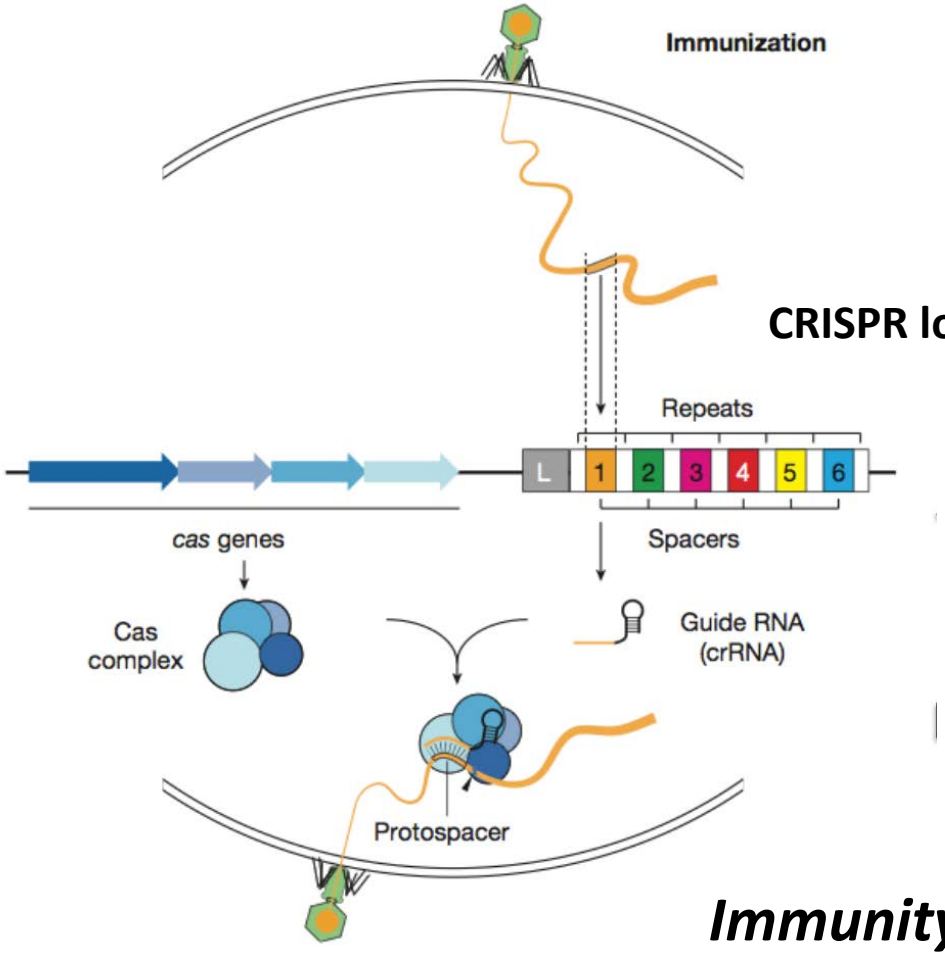
- gene KO
- enhancer deletion
- …

Precise modification

- tagging with GFP, peptide tag, …
- transgene knock-in
- correction of mutations

CRISPR/Cas systems mediate adaptive immunity in bacteria

Immunization



CRISPR locus

Mechanism of adaptive immunity in bacteria and archaea

Evolved to adapt and defend against foreign genetic material (i.e., phage, horizontal gene transfer, etc.)

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

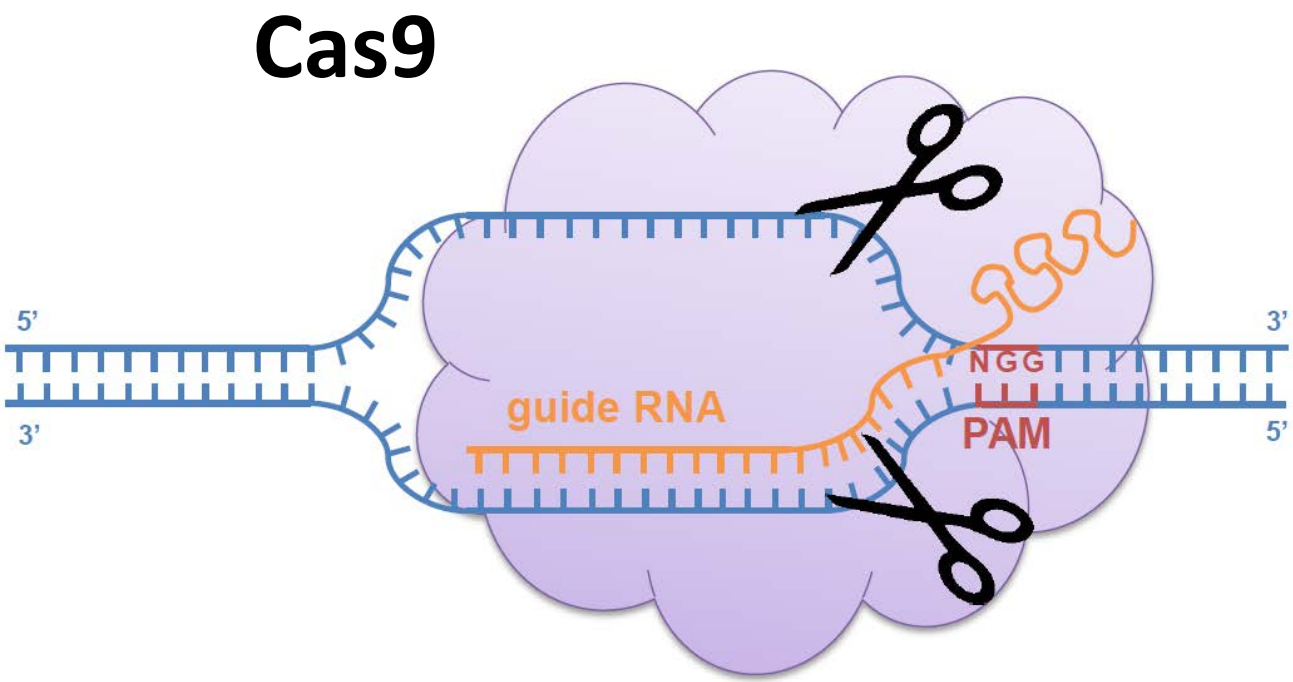
Cas: CRISPR-associated proteins

Ishino et al, 1987
Mojica et al, 2005
Barrangou et al, 2007
Jinek et al, 2012

Bhaya et al., 2011 (Ann Rev Gen)

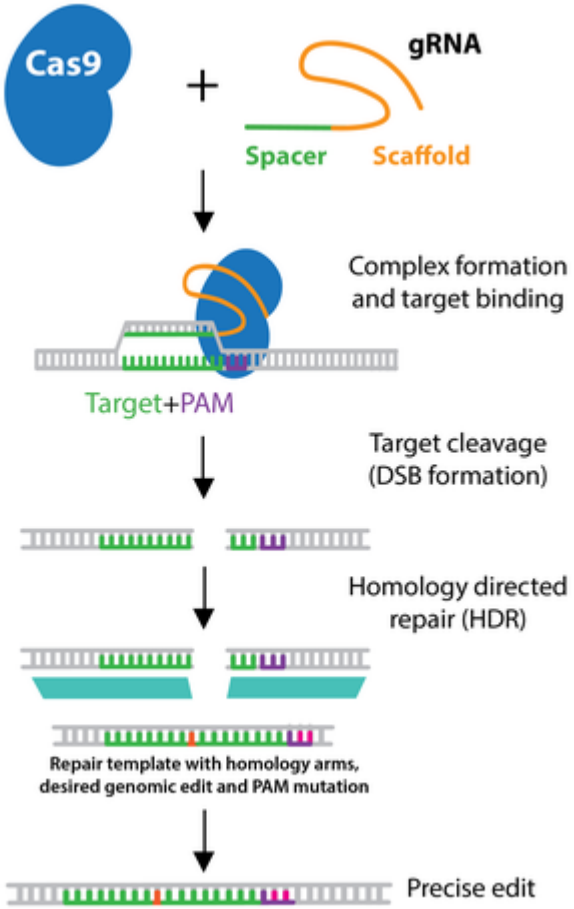
Immunity by DNA interference

CRISPR-Cas9 nucleases



- binds to PAM motif (NGG for *S. pyogenes* Cas9)
- cleaves DNA if guide RNA binds to sequence upstream of PAM

Workflow for making Precise Modifications using Homology Directed Repair (HDR)



Guide RNA (design, production, tests)

Template DNA design

Cas9 RNP and Template DNA delivery

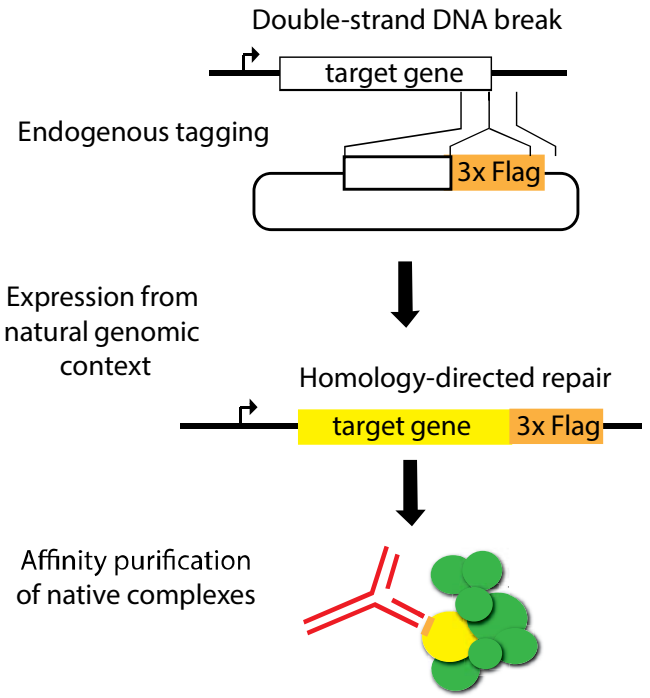
Isolation of cells with tagged gene

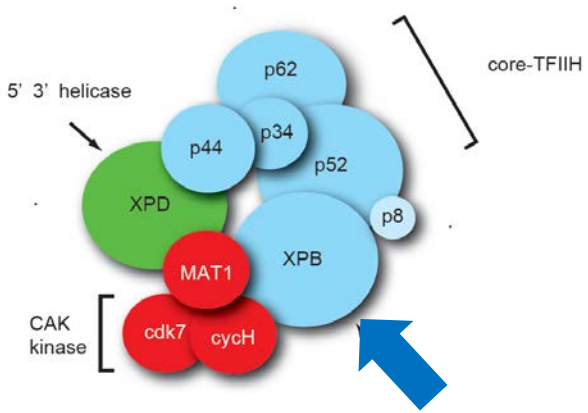
Tagging of the TFIIH XPB subunit

1/ Generate homozygous KI mammalian cell lines expressing an affinity tagged protein using the CrispR/Cas9 technology

2/ Cultivate cells to express the tagged protein and its associated partners from their natural environment

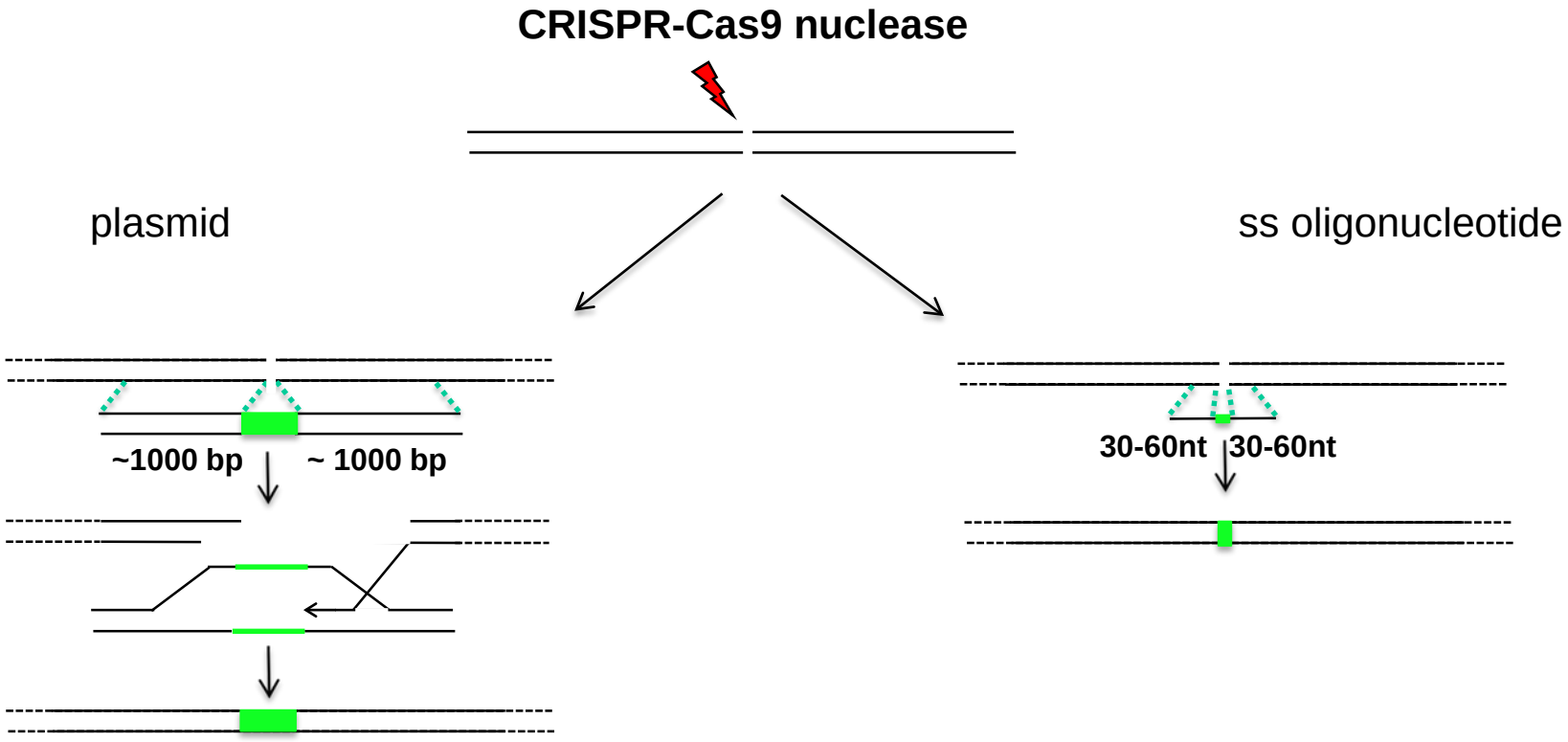
3/ Purify and characterize the corresponding endogenous complexes (composition, stability,…)





TACGENE

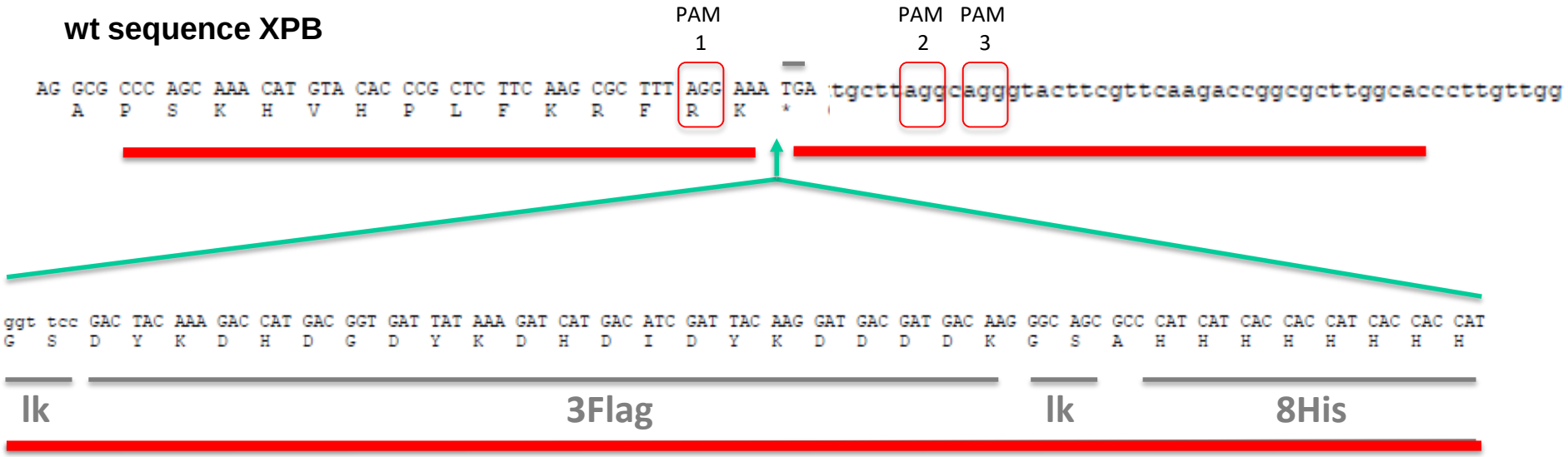
Template DNA design: Plasmid or ssODN



- **insertion** ***(GFP, transgene, …)***
- **point mutation**

- **insertion** ***(tag<120 nt, …)***
- **point mutation**

Guide RNA and Template DNA design



Peptide tag insert sequence

Guide RNA
- keep distance between cut and insert position (<20bp)
- avoid guide with unwanted off-targets
- screen for an efficient guide

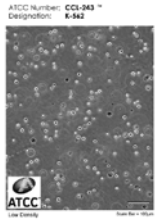
Template DNA
- 30-60 nt Homology arms
- If possible, test both orientations of ssODN donor
- New genomic sequence should not be cut by guide RNA/Cas9
- 2 PS linkages are added at each ssODN end to improve efficiency

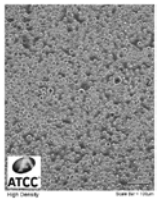
Geny, S & Concordet, JP

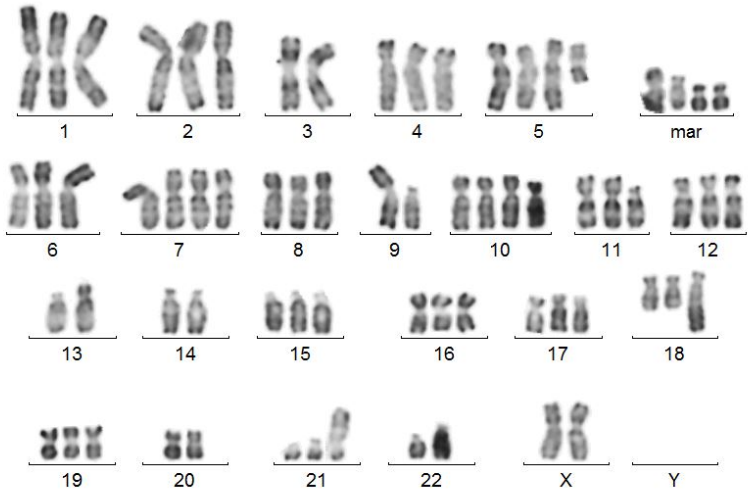
K562: an easy to manipulate cell line

Non adherent erythroleucemic cells (K562)

cultivated, expanded, and grown easily transfected with high efficiency, tolerates limite permissive to genome editing events

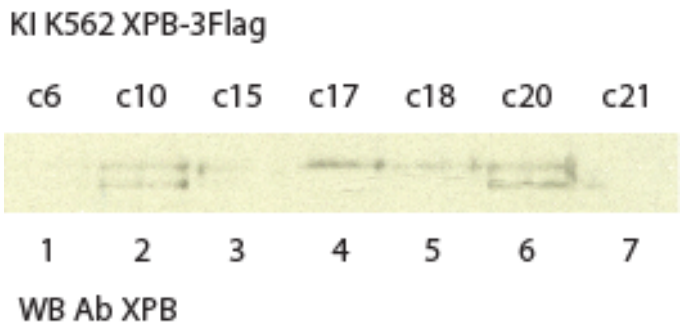


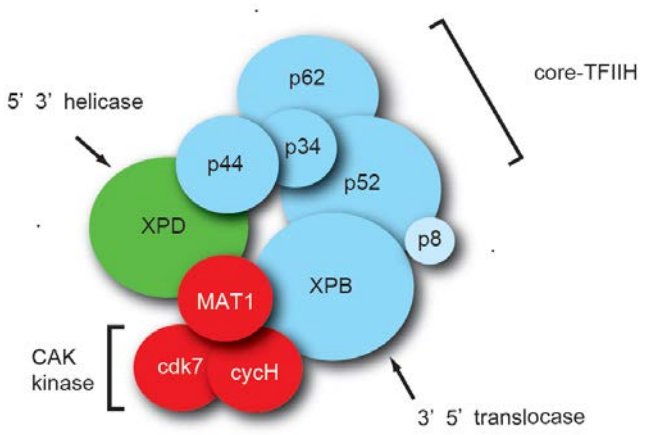


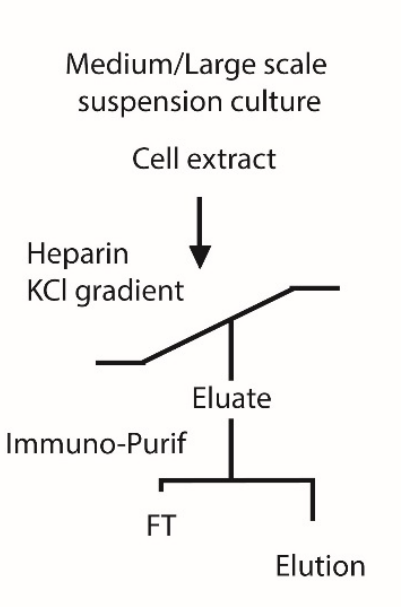


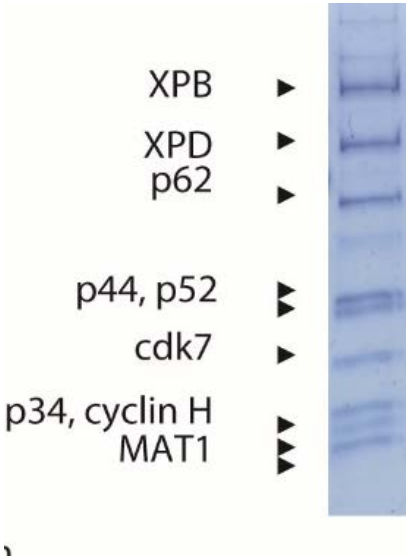
L Mauvieux and C Gervais

Tagging of the TFIIH XPB subunit









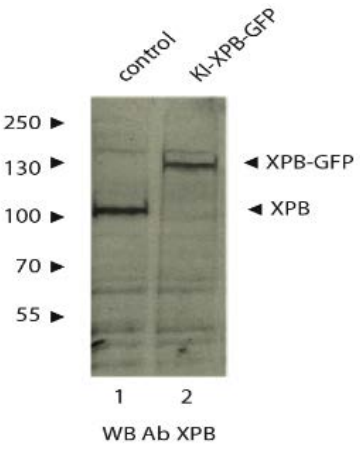
Genome editing to facilitate the purification and characterization of endogenous complexes

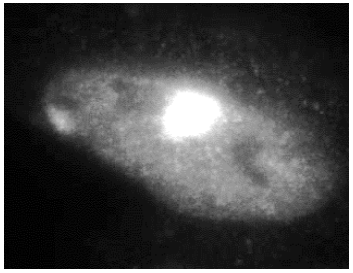
Functional characterization
(in vitro assays, CHIPs)

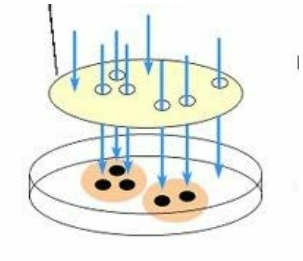
Structural analysis
(X-link-MS, cryo-EM)

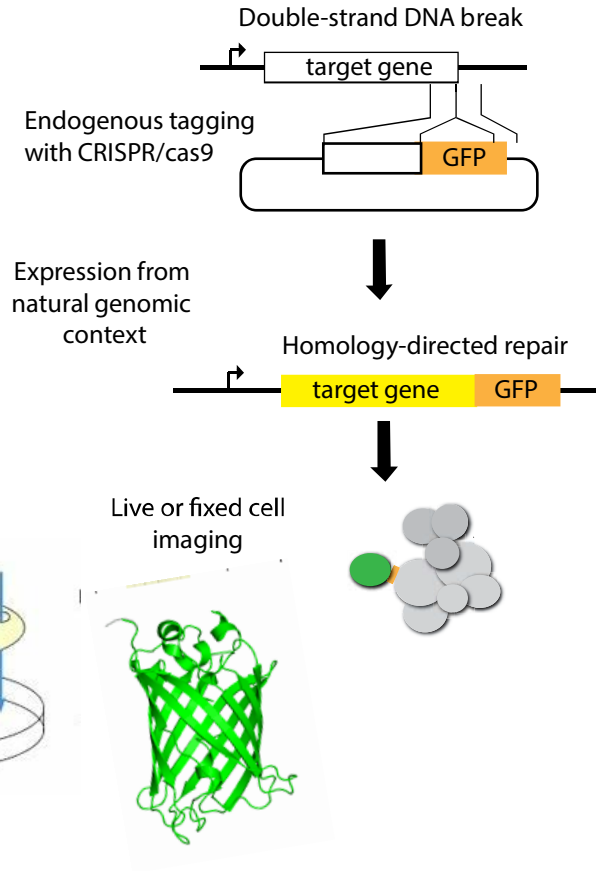
Detailled Mass Spectrometry
(PTMs, quantitative MS)

Imaging
(live cell, super-res)









Geremy et al., Nat Comm 2019

Sample preparation for structural anaysis

Recombinant protein expression for structural biology: An overview of popular expression systems

Co-expression system for reconstitution of multiprotein complexes and dissection of protein-protein interaction networks

Genome engineering for labelling mammalian proteins to facilitate isolation of endogenous complexes and their characterization in a cellular environment

Support


igbmc
institut de génétique et de
biologie moléculaire et cellulaire


ANR
AGENCE
NATIONALE
DE LA
RECHERCHE


Association
pour la Recherche
sur le Cancer


cnrs


INSTITUT
NATIONAL
DU CANCER


CONTRE
LE CANCER
LA LIGUE
101
comités
pour la vie


UNIVERSITÉ DE STRASBOURG


instruct
Integrating
Biology


Inserm


instruct
ULTRA


FRISBI
Integrated Structural Biology

igbmc
institut de génétique et de biologie moléculaire et cellulaire

cnrs UNIVERSITÉ DE STRASBOURG Inserm

Dept Genomics and Structural Biology

TFIIH (IGBMC, Strasbourg)

A Poterszman,
O Koleschnikova,
L Radu W Abdulraman*

JM Egly/F Coin
C Braun

pTefb project (ENS, Paris)

O Bensaude, L Kobbi

NRs (IGBMC, Strasbourg)

N Rochel, J Osz-Papai

MNHN (Paris)

JP Concordet,
C Giovannangeli
S Geny, K Lambriet

TACGENE

IGBMC platforms and facilities

Baculovirus expression
Molecular Biology
Structural Biology

C Birck, S Pichard
E Edelweiss, S Jacquemin
I Kolb Cheneyl; P Rossollilo
N Troeffer Charlier

FRISBI Integrated Structural Biology

instruct Integrating Biology