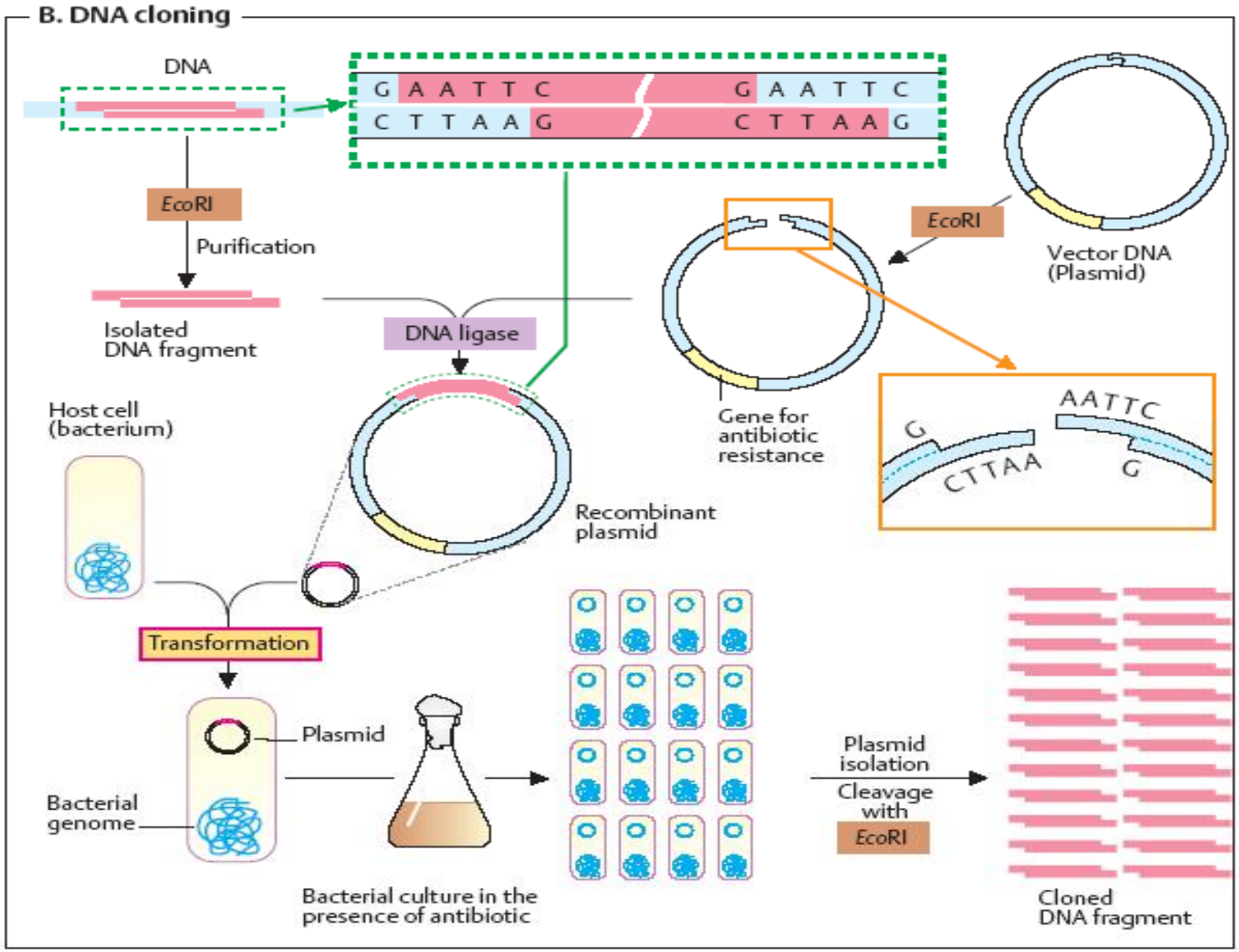
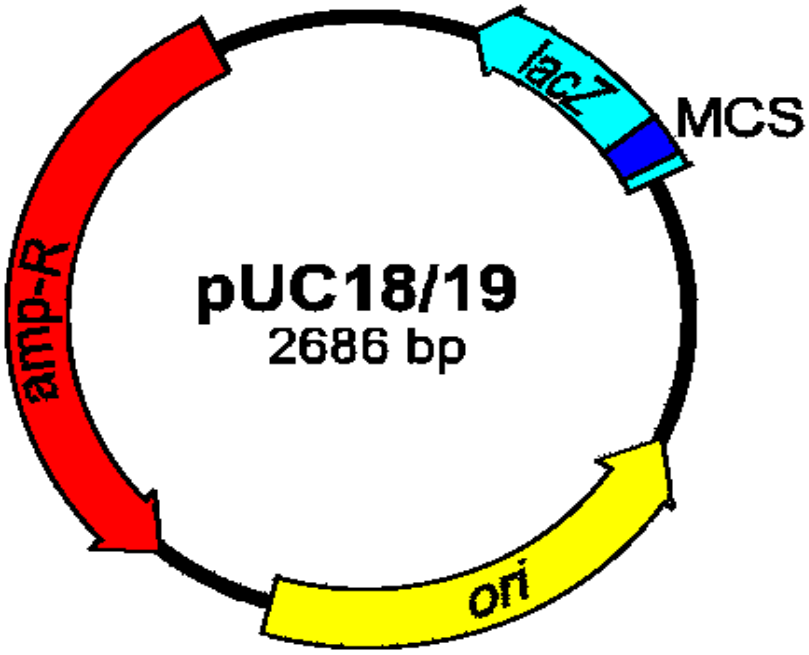
Gene Expression

By Prof Amer Jamil
Dept of Chemistry and Biochemistry
University of Agriculture Faisalabad



By Prof Amer Jamil



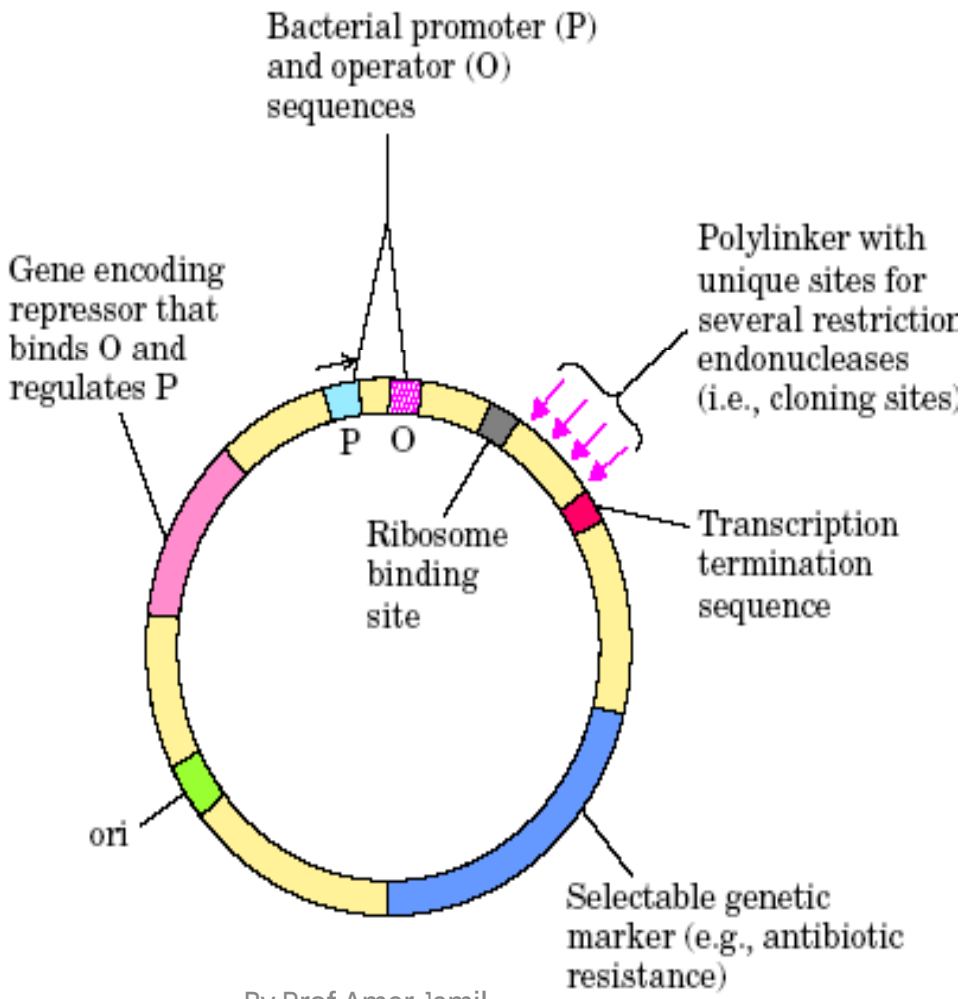
Useful Plasmid Features

- Relaxed Replication
- Selectable Markers
- Streamlined
- Polylinker or MCS
- Identification of Recombinants
- most derived from pUC or pBR322

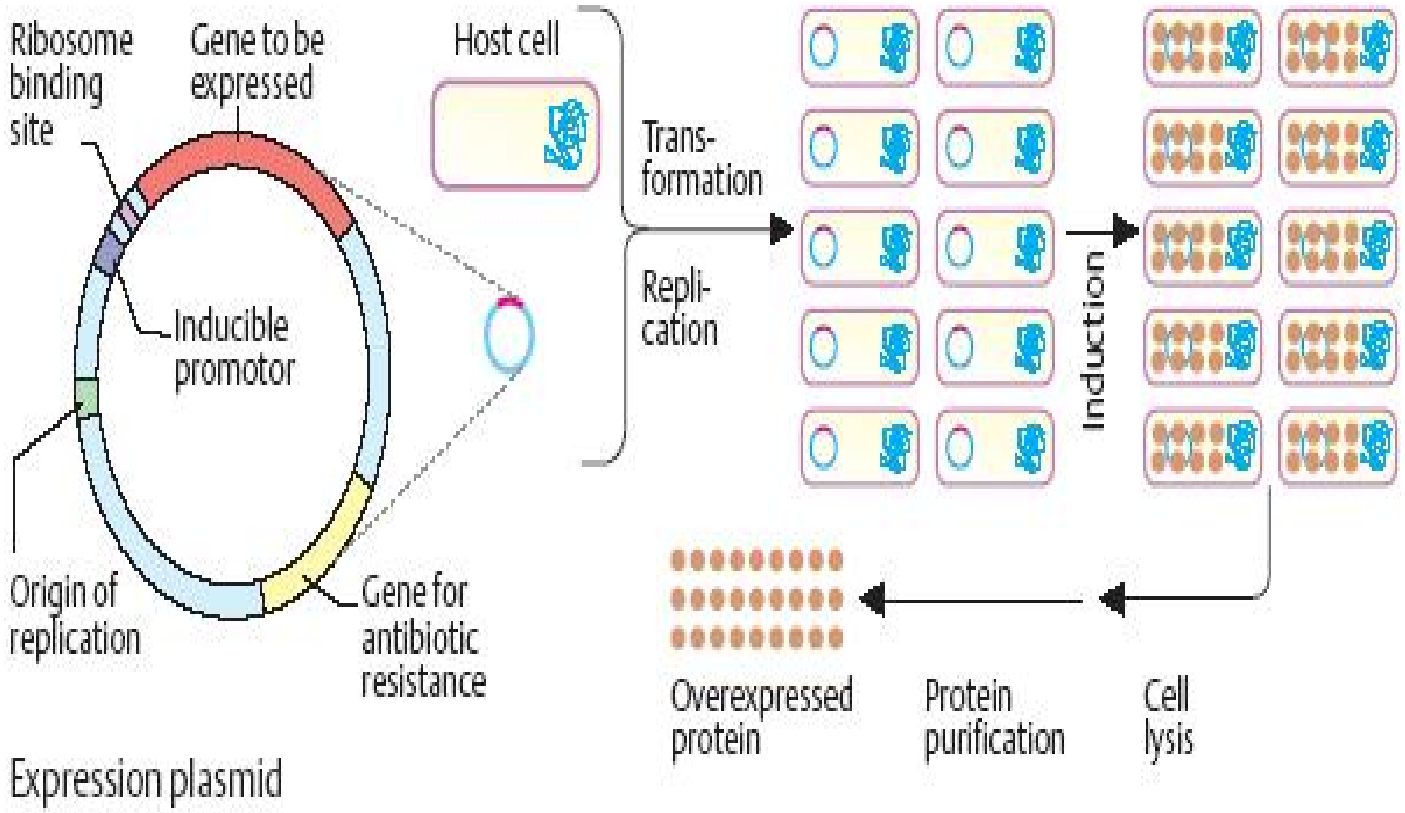
Multiple Cloning Site:

```
|SacI|  |ScII|          |XbaI||SpeI||BamH||SmaI||PstI||EcRI||EcRV||HIII||ClaI|   |SalI||XhoI|          |KpnI|
GAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACC
CTCGAGGTGGCGCCACCGCCGGCGAGATCTTGATCACCTAGGGGGCCCGACGTCCTTAAGCTATAGTTCGAATAGCTATGGCAGCTGGAGCTCCCCCCCGGGCCATGG
```

DNA sequences in a typical *E. coli* expression vector.



Overexpression of proteins



Heterologous gene expression in prokaryotes
Promoters for gene expression

Promoter	Plasmid	Host	Selection	Induction
pL promoter controlled by cI repressor	pKC30, pRK16F, pWW12	N4830 (*cI*); temp sensitive or SA1512 (*cI*)	Amp^R	Temp 30 to 42 ^oC Nalidixic acid (cleaves λ repressor)
Phage T7 promoter controlled by T7 RNA polymerase	PT7-1,2 pET1-7	HMS273 (pGP1-2) BL21 (DE3)	Amp^R Amp^R	Temp sensitive (*cI* based) IPTG-*lac* repressor

Promoter	Plasmid	Host	Selection	Induction
E. coli lac promoter controlled by *lac* repressor	pUC18, 19 pDL19	JM101	Amp^R	IPTG
Synthetic *tac* promoter controlled by *lac* repressor	ptacI ptacII	JM101	Amp^R	IPTG
Run-away replicon	pMOB45 pMOB48	DB1443 (*hsdR*) JM101	Cm^R/Tet^R Cm^R	Temp sensitive; over 100 fold expression

Increased gene dosage—multiple copies of a gene downstream of a strong promoter

Measuring overexpression

- SDS-PAGE
- Western blotting
- Biochemical Analysis

Important considerations

- Gene toxicity
  - Overproduction of any gene can be toxic to the cell
    - Transcription is repressed until the expression is desired
- Premature termination
  - Antiterminators (such as of λ phage are used) to counter the intragenic terminators
- mRNA instability
  - mRNA stability may be enhanced by
    - Trimming or addition of linker sequences in the junction between the promoter and the coding region
    - *omp*A-like leader sequence elements
    - Mutations agaisnt RNaseIII
    - Presence of strong terminators

- Translation initiation
  - AUG initiating codon; sometimes GUG, UUG and AUU
  - Shine-Dalgarno sequence positioned about 7-9 nucletides upstream of the initiator codon
  - No secondary structure within Shine-Dalgarno sequence nor any overlap with other sequences
  - Preferably AUG codon followed by GCAU or AAAA; UUAA in the fourth or fifth codon will be helpful
  - Euk cDNA does not have the sequence; therefore vectors with this sequence should be used
- Biased codon usage
  - Rare codons hinder the gene expression

- Enhancing protein stability
  - Protease deficient hosts may be used
- Secretion of cloned gene product
  - Proteins are more stable if move to periplasmic space or outside cell
  - Easy to purify if move out of cell
  - Secretion signals are used e.g., ompA, Hly (hemolysin) secretory pathway
- Fusion vectors
  - e.g. *lacZ* fusion
    - Easy to purify by affinity column
    - The fusion protein is released by a specific protease

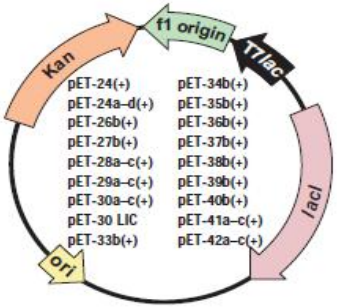
Overexpression in eukaryotes

- Why in Eukaryotic system and not in a simple prokaryotic system?
  - No acceptable environment for start codon
  - Gene toxicity
  - Synthesized peptide is biologically inactive
  - Posttranslational modifications do not exist accordingly
  - Formation of insoluble material in inclusion bodies
    - Inappropriately folded protein may not be rescuable by denaturation-renaturation

Eukaryotic systems for protein expression

- Yeast
  - *Saccharomyces cerevisiae*
  - *Pichia pastoris*
- Mammalian
  - Constitutive using strong promoters such as SV40 early promoter and enhancer
  - Transient using viral expression systems such as SV40, adenoviruses, papillomaviruses, herpesviruses etc
- Insect
  - Baculovirus-insect cell expression system

pET expression system



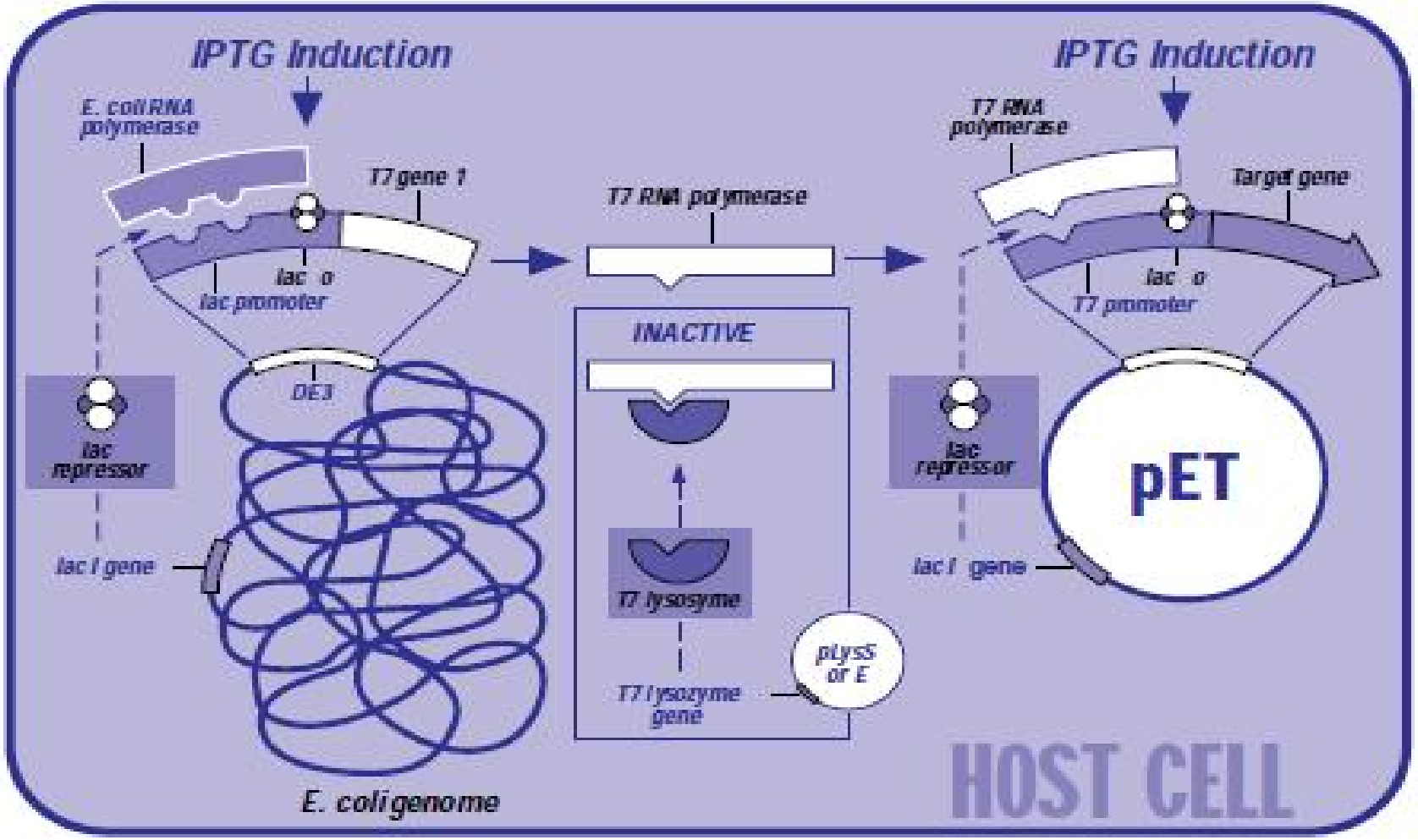


Table 1. pET System Host Strains

Strain	Derivation	Key Feature(s)	Antibiotic Resistance	Available as Competent Cells
AD494	K-12	*trxB* mutant; facilitates cytoplasmic disulfide bond formation	Kan	yes
AD494(DE3)			Kan	yes
AD494(DE3)pLysS			Kan + Cam	yes
BL21	B834	Lacks *lon* and *ompT* proteases	none	yes
BL21(DE3)			none	yes
BL21(DE3)pLysS			Cam	yes
BL21(DE3)pLysE			Cam	no
BL21*trxB*(DE3)	BL21	BL21 *trxB* mutant; facilitates cytoplasmic disulfide bond formation	Kan	yes
BL21*trxB*(DE3)pLysS			Kan + Cam	yes
BLR	BL21	BL21 *recA* mutant; stabilizes tandem repeats	Tet	yes
BLR(DE3)			Tet	yes
BLR(DE3)pLysS			Tet + Cam	yes
B834	B strain	met auxotroph; ^{35}S-met and selenomethionine labeling	none	no
B834(DE3)			none	yes
B834(DE3)pLysS			Cam	yes
HMS174	K-12	*recA* mutant Rif resistance	Rif	yes
HMS174(DE3)			Rif	yes
HMS174(DE3)pLysS			Rif + Cam	yes
HMS174(DE3)pLysE			Rif + Cam	no
NovaBlue	K-12	*recA*, *endA*, *lacI*^{q} cloning, plasmid preps	Tet	yes
NovaBlue(DE3)			Tet	yes
Origami™	K-12	*trxB*/*gor* mutant, greatly facilitates cytoplasmic disulfide bond formation	Kan + Tet	yes
Origami(DE3)			Kan + Tet	yes
Origami B(DE3)pLysS			Kan + Tet + Cam	yes
Origami B	Tuner	BL21 *lacY* deletion, *trxB*/*gor* mutant, greatly facilitates cytoplasmic disulfide bond formation; allows precise control with IPTG	Kan + Tet	yes
Origami(DE3)			Kan + Tet	yes
Origami B(DE3)pLysS			Kan + Tet + Cam	yes
			Kan + Tet + Cam	yes
Rosetta™	Tuner	Enhances expression of proteins having codons rarely used in *E. coli*, *lacY* deletion mutant	Cam	yes
Rosetta(DE3)			Cam	yes
Rosetta(DE3)pLysS			Cam	yes
Rosetta-gami™	Origami	Enhances expression of proteins having codons rarely used in *E. coli*, *trxB*/*gor* mutant	Kan + Tet + Cam	yes
Rosetta-gami(DE3)			Kan + Tet + Cam	yes
Rosetta-gami(DE3)pLysS			Kan + Tet + Cam	yes
RosettaBlue™	NovaBlue	Enhances expression of proteins having codons rarely used in *E. coli*, *recA*, *endA*, *lacI*^{q}	Tet + Cam	yes
RosettaBlue(DE3)			Tet + Cam	yes
RosettaBlue(DE3)pLysS			Tet + Cam	yes
Tuner™	BL21	BL21 *lacY* deletion mutant allows precise control with IPTG	none	yes
Tuner(DE3)			none	yes
Tuner(DE3)pLysS			Cam	yes

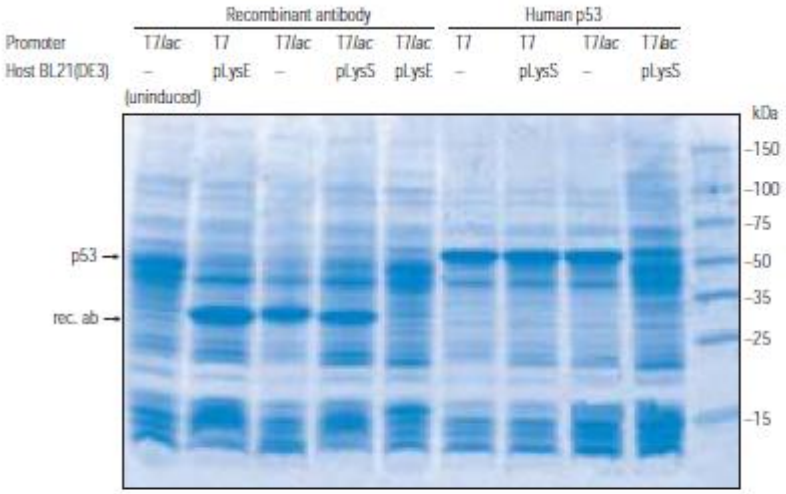
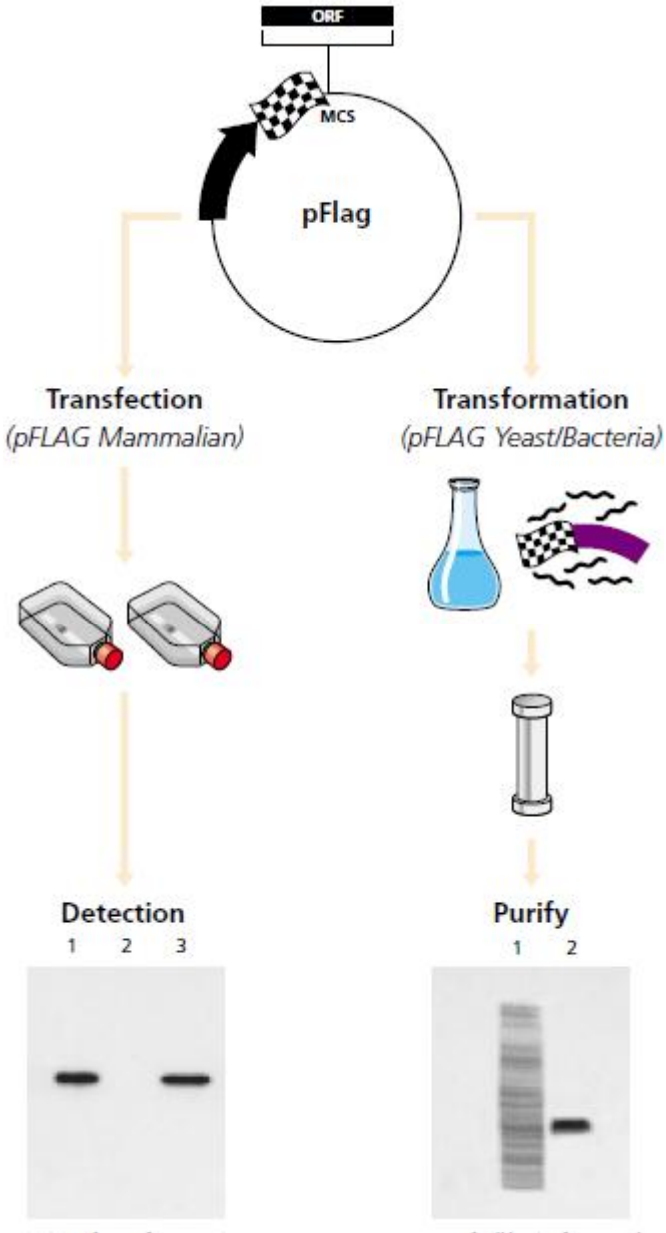


Figure 2. Effect of vector/host combination on expression levels of two proteins

The indicated cell cultures were grown at 37°C to OD_{600} of approximately 0.8 and expression induced with 1 mM IPTG for 2.5 h. Total cell protein samples were run along with Novagen's Perfect Protein™ Markers on a 4–20% SDS polyacrylamide gradient gel followed by staining with Coomassie blue. Vectors used were pET-20b(+) and pET-22b(+) for the recombinant antibody and pET-23b(+) and pET-21b(+) for p53.

Table 2. Fusion tags available for pET constructs

Tag	N/C Terminal or Internal	Size (aa)	Basis for Detection and/or Purification	Applications
T7•Tag®	N, I	11 or 260	monoclonal antibody	Western blot, immunoprecipitation, purification
S•Tag™	N, I	15	S-protein (104 aa) affinity	Western blot, quantitative assay, purification
His•Tag®	N, C, I	6, 8, or 10	metal chelation chromatography (native or denaturing)	His•Bind® resin purification
HSV•Tag®	C	11	monoclonal antibody	Western blot, immunofluorescence
pelB/ompT	N	20/22	potential periplasmic localization	protein export/folding
KSI	N	125	highly expressed hydrophobic domain	small protein/peptide production/purification, insoluble protein
Trx•Tag™	N	109	thioredoxin	soluble protein, cytoplasmic disulfide bond formation in *trxB*⁻ hosts
PKA site	N	5	protein kinase A recognition site	*in vitro* phosphorylation
CBD_{Cbs}•Tag	N	156	polyclonal antibody, cellulose binding domain	Western blot, purification, noncovalent immobilization
CBD_{cenA}•Tag	N	114	polyclonal antibody, cellulose binding domain, periplasm/media	protein export, Western blot, purification, noncovalent immobilization
CBD_{Cex}•Tag	C	107	polyclonal antibody, cellulose binding domain, periplasm/media	protein export, Western blot, purification, noncovalent immobilization
Dsb•Tag™	N	208 (DsbA) 236 (DsbC)	potential periplasmic localization	soluble protein, periplasmic disulfide bond formation, isomerization
GST•Tag™	N	220	glutathione affinity monoclonal antibody enzymatic activity	purification, Western blot, quantitative assay
Nus•Tag™	N	495	NusA	soluble protein, cytoplasmic disulfide bond formation in *trxB*⁻ hosts



FLAG Peptide:

Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys

Enterokinase Cleavage Site

Protein

Detection of FLAG fusion protein with ANTI-FLAG® M2 Antibody

Lane 1. FLAG-BAP™* control protein

Lane 2. COS-7 Cell extract control

Lane 3. pFLAG™-CMV-2-BAP* transfected COS-7 cell extract

*BAP: Bacterial alkaline phospatase

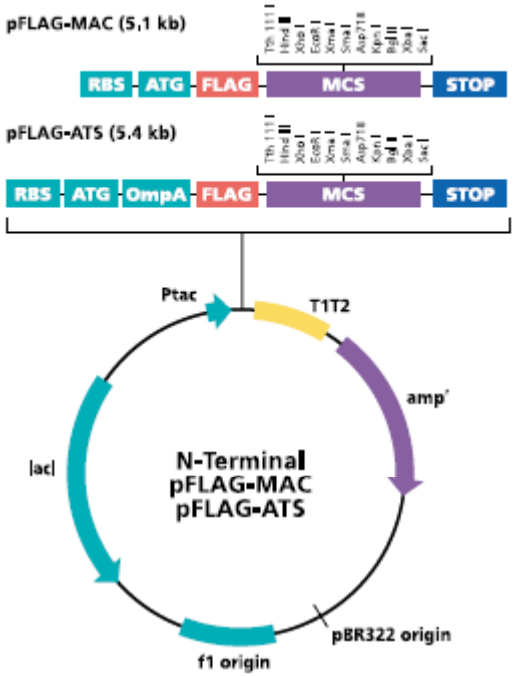
SDS-PAGE of cell lysates form *E. coli* transfected with pFLAG-ATS-BAP purified with ANTI-FLAG M2 affinity gel

Lane 1. Cell extract prior to purification

Lane 2. Affinity purified FLAG-BAP* fusion protein

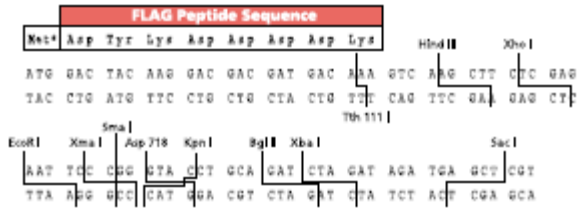
Product	Product Name	Size	Promoter	FLAG®	OmpA	Ek Site
E 5769	pFLAG-ATS™	5.4kb	tac	N	√	√
E 5644	pFLAG-MAC™	5.1	tac	N		√
E 5269	pFLAG-CTS™	5.4	tac	C	√	
E 5394	pFLAG-CTC™	5.3	tac	C		
E 8023	pFLAG-Shift12™	5.1	tac	N	√	√
E 7898	pFLAG-Shift12c™	5.1	tac	N		√
P 1118	pT7-FLAG-1™		T7/lacO	N		√
P 1243	pT7-FLAG-2™		T7/lacO	C		

N – N-terminal tag **C** – C-terminal tag **ompA** – periplasmic localization **Ek** – Enterokinase cleavage site

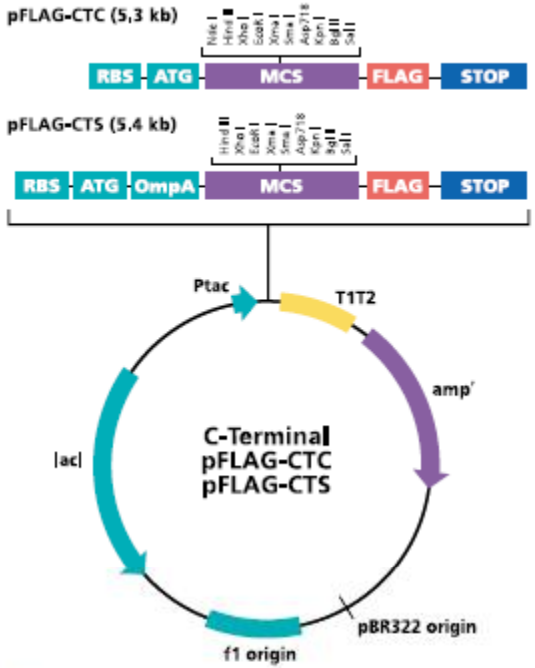


Multiple Cloning Site

(pFLAG-MAC and pFLAG-ATS*)

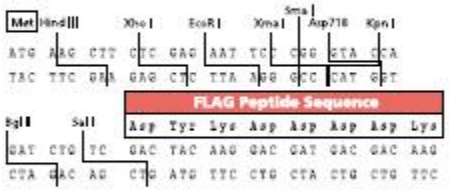


*For p-FLAG-ATS the Met-OmpA coding sequence precedes the FLAG coding sequence.



Multiple Cloning Site

(pFLAG-CTC* and pFLAG-CTS**)



*FLAG peptide coding sequence is out of phase with the ATG condon located within the NdeI site.

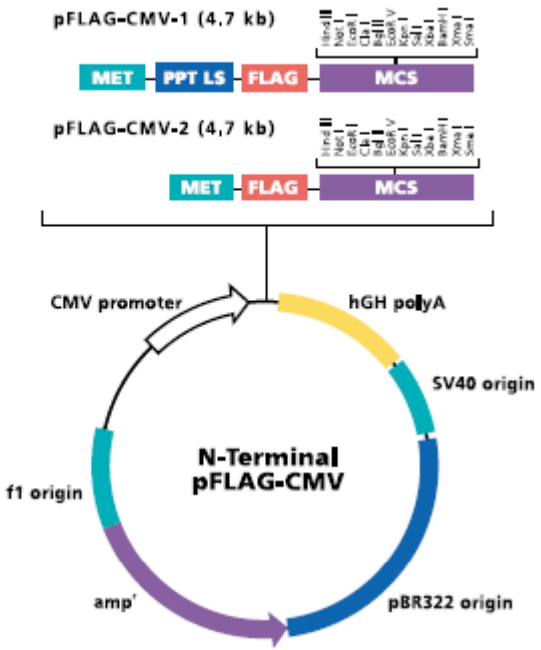
(Open reading frame must be phased with ATG and FLAG.

**For pFLAG-CTS the Met-OmpA coding sequence precedes the MCS. See product insert for pFLAG-CTS MCS.

Mammalian expression

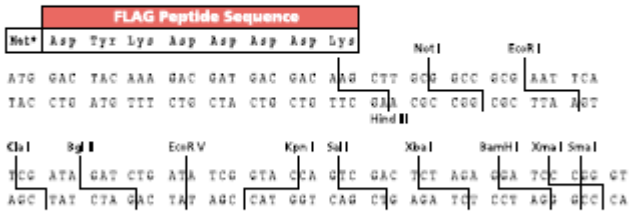
Product	Product Name	PPT	FLAG®	3xFLAG®	*c-myc*	Stable Selection
E 7273	pFLAG-CMV™-1	√	N			
E 7398	pFLAG-CMV™-2		N			
E 8770	pFLAG-CMV™-3	√	N			neomycin
E 1775	pFLAG-CMV™-4		N			neomycin
E 3762	pFLAG-CMV™-5a,b,c		C			
E 7901	pFLAG-CMV™-5.1		C			
E 2275	pFLAG-CMV™-6a,b,c		N			
E 4026	p3xFLAG-CMV™-7.1			N		
E 4151	p3xFLAG-CMV™-8	√		N		
E 4276	p3xFLAG-CMV™-9	√		N		neomycin
E 4401	p3xFLAG-CMV™-10			N		neomycin
E 4776	p3xFLAG-CMV™-13			C		neomycin
E 4901	p3xFLAG-CMV™-14	√		C		neomycin
Dual Tag Vectors						
E 5526	pFLAG-myc-CMV™-19	√	N		C	
E 5651	pFLAG-myc-CMV™-20		N		C	
E 5776	pFLAG-myc-CMV™-21	√	N		C	neomycin
E 5901	pFLAG-myc-CMV™-22		N		C	neomycin
E 6026	p3xFLAG-myc-CMV™-23	√		N	C	
E 6151	p3xFLAG-myc-CMV™-24			N	C	
E 6276	p3xFLAG-myc-CMV™-25	√		N	C	neomycin
E 6401	p3xFLAG-myc-CMV™-26			N	C	neomycin

N – N-terminal tag **C** – C-terminal tag **PPT** – preprotrypsin leader for direct secretion ***c-myc*** – c-*myc* epitope

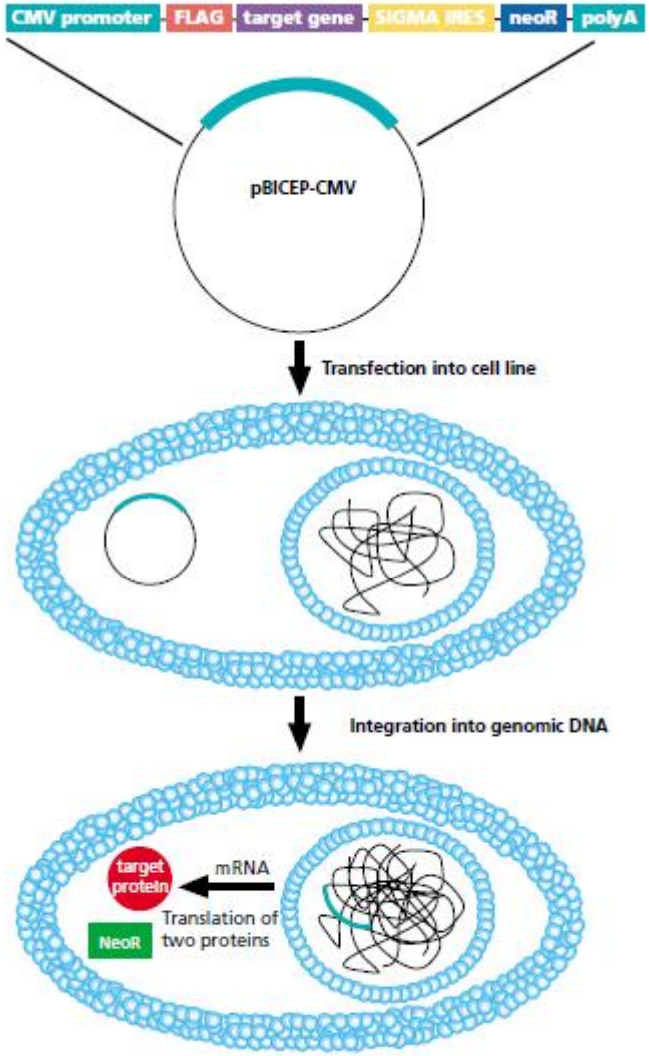


Multiple Cloning Site

(pFLAG-CMV-1* and pFLAG-CMV-2)

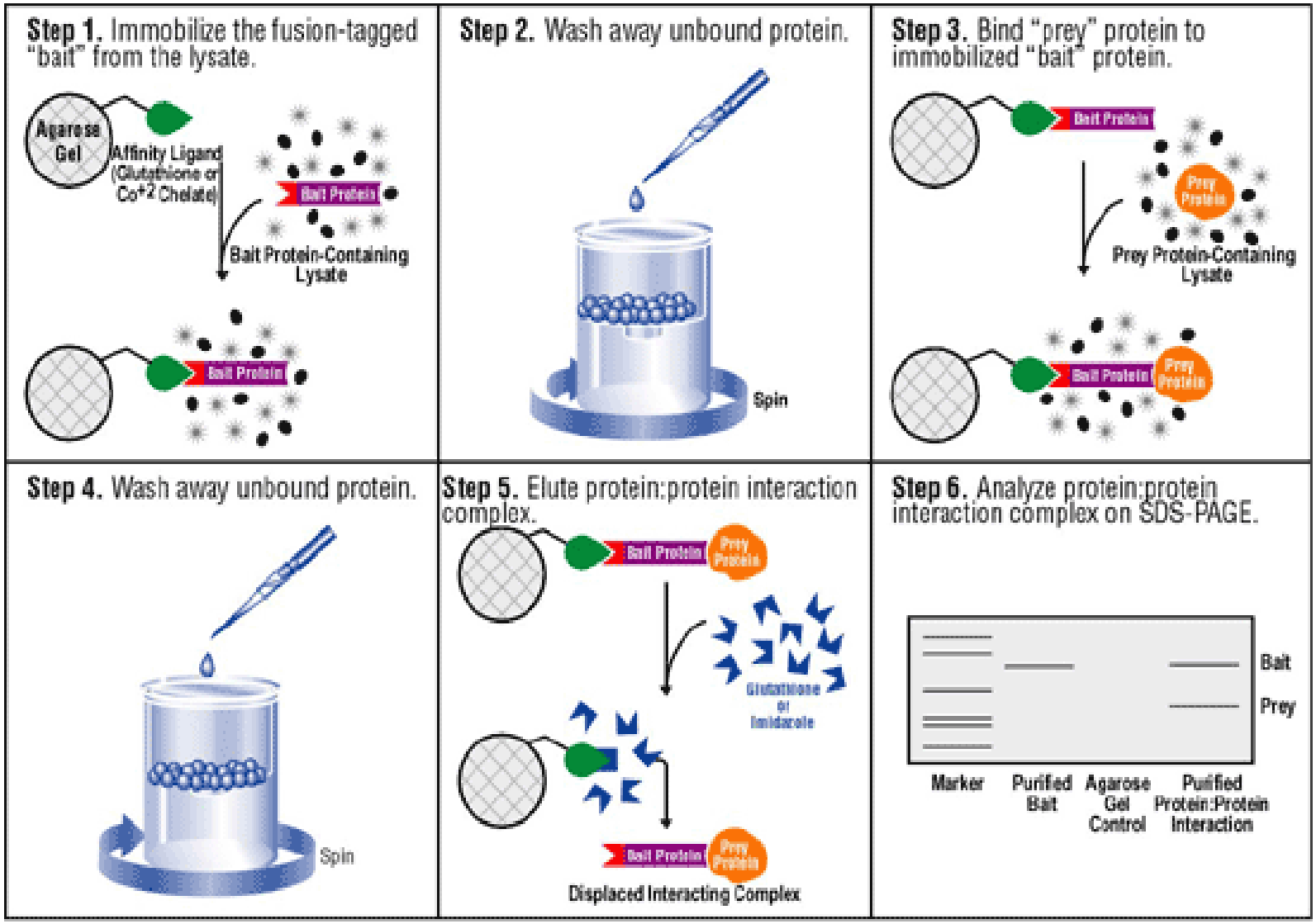


*For pFLAG-CMV-1, the Met-preprotrypsin leader sequence (PPT LS) precedes the FLAG coding sequence.



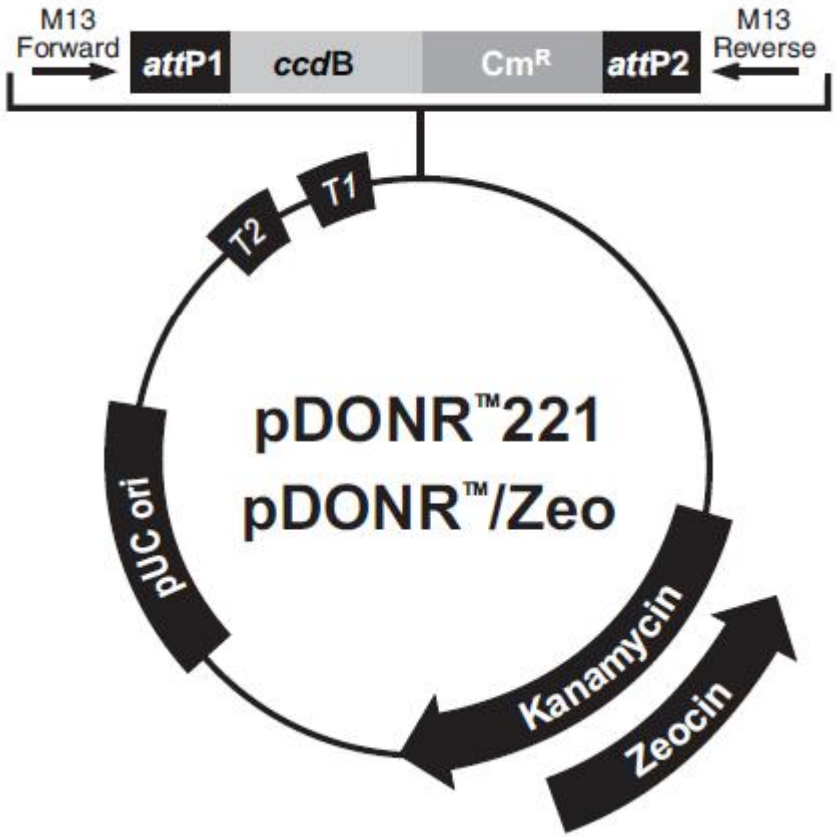
Stable Cells Retain High Levels of Expression

Procedure summary for GST Tag and His Tag Pull-Down Kits. "Pull-down" is a small-scale affinity purification technique similar to immunoprecipitation (IP), except that the antibody function of is replaced by some other affinity system. In this case, the affinity system is either a GST-tagged protein that can be captured by glutathione agarose beads or a His-tagged protein that can be captured by metal chelate (cobalt) agarose beads. The fusion-tagged protein acts as the "bait" to capture a putative binding partner (i.e., the "prey"). In a typical pull-down assay, the immobilized bait protein is incubated with a cell lysate. After the prescribed washing steps, the 'interactors" are selectively eluted for analysis in-gel or by Western blot.



= Fusion Tag (GST or polyHis)

Gateway cloning

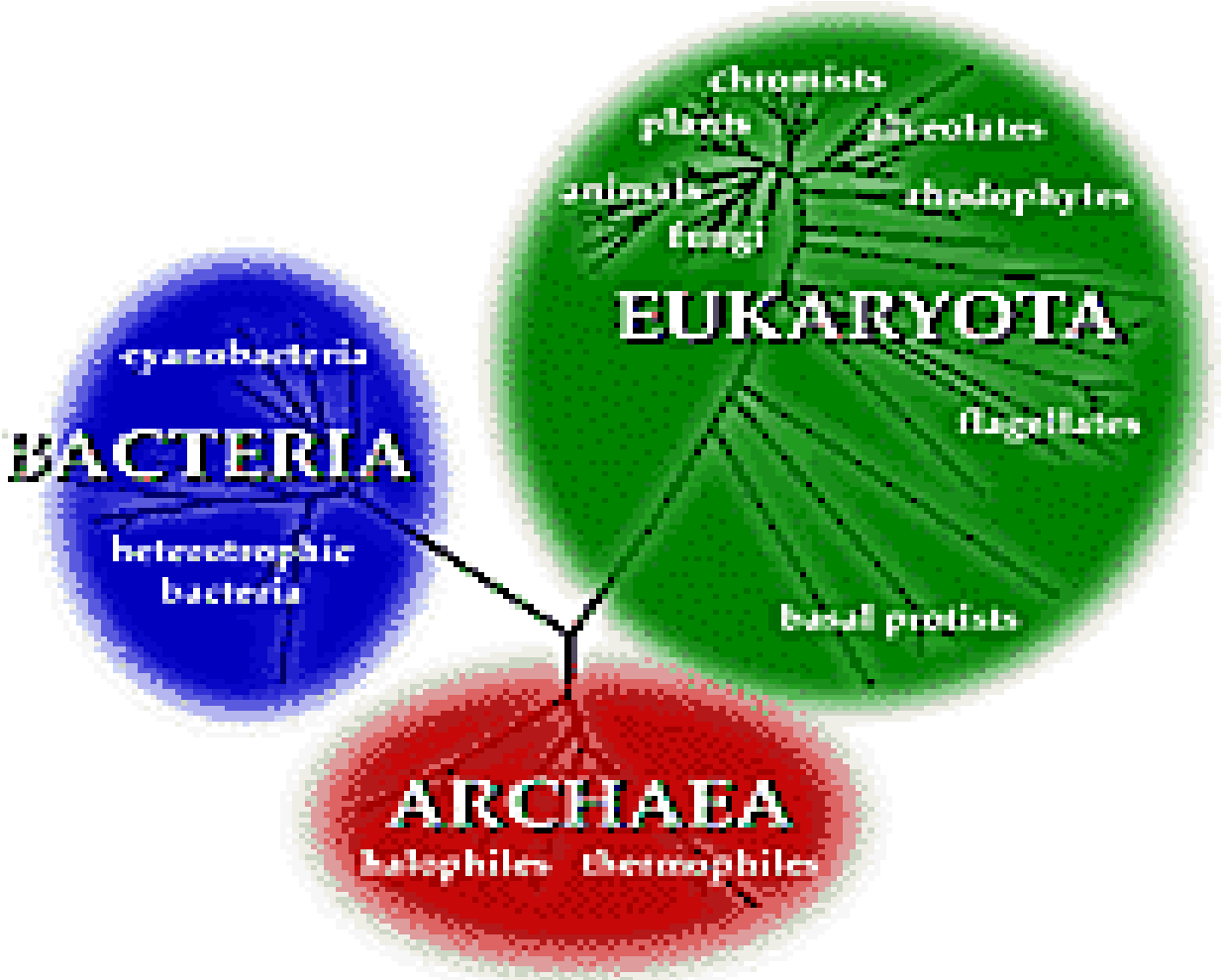


By Prof Amer Jamil

Archaea??????????

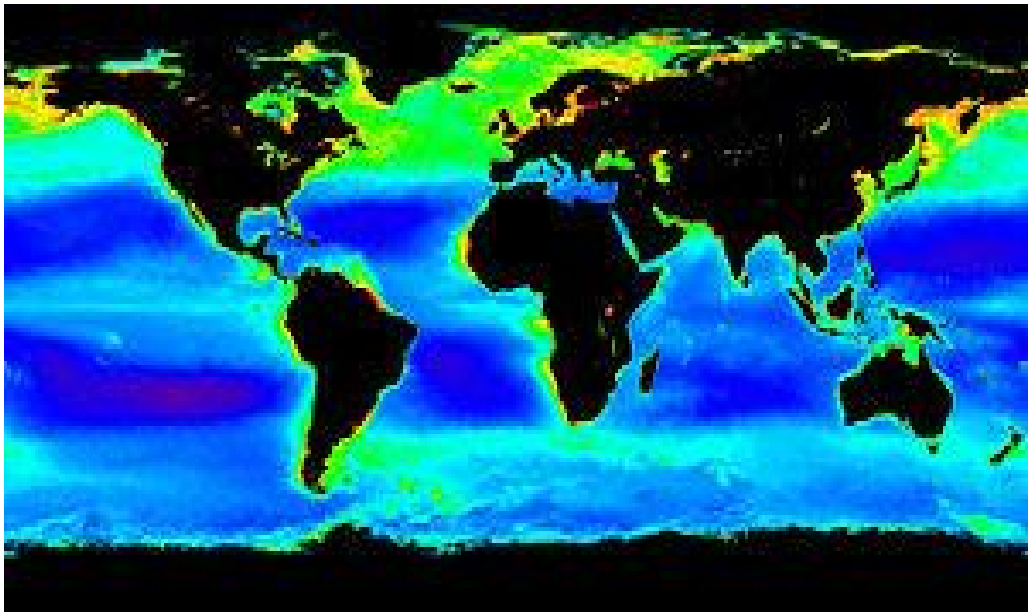
A group of single-celled microorganisms. A single individual or species from this domain is called an *archaeon* (sometimes spelled "archeon"). Archaea, like bacteria, are prokaryotes and have no cell nucleus or any other organelles within their cells.

Distinct Domains Of Life



Habitat of Archaea

- Archaea exist in a broad range of habitat
- A major part of global ecosystems
- May contribute up to 20% of the total biomass on Earth



Habitat .....

Hot springs of Yellow Stone National Park, USA

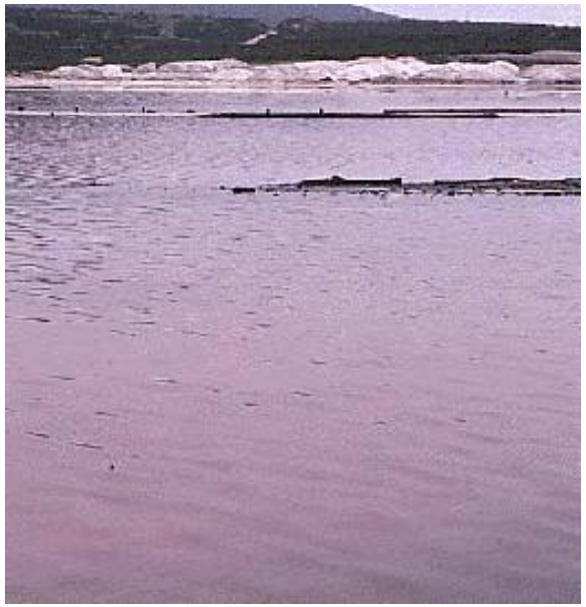
- Extremophiles:
    - high temperatures >100 as found in geysers and black smokers.
    - very cold habitats or in highly saline, acidic, or alkaline water.
- Mesophiles:
    - marshland, sewage, and soil.
- Methanogenic archaea:
    - in the digestive tracts of animals such as ruminants, termites, and humans.

(Archaea are usually harmless to other organisms and none is known to cause disease.)

Salt-Lovers……..

- Halophilic Archean sp. Near San Quentin, Baja California Norte, Maxico

- South end of San Fransisco Bay



Similarities with Bacteria

- Bacteria and archaea have several common characteristics…….

  1- Both lack internal membrane structures

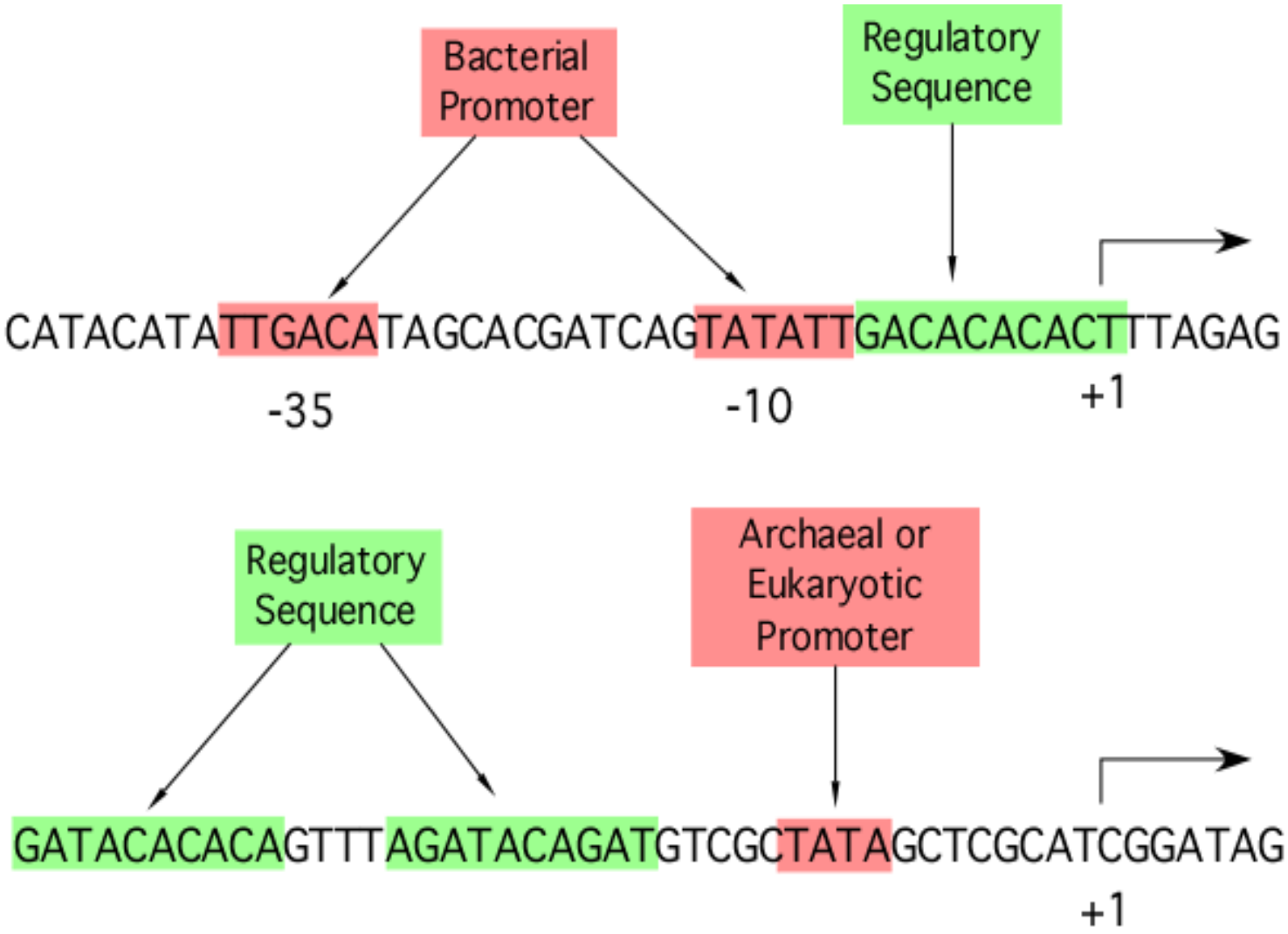
  2- Both use plasmids to pass information

  3- Both have ability to live in very extreme and hostile environments…….

  4- Ribosome of Archaea structurally similar to Bacteria

Similarities with Eukaryotes

- DNA with Histones
- DNA Replication
- Transcription
- Glycosylation
- Transcription factors to “turn on” their genes…….
- tRNA contain introns
- Sensitivity of Ribosomes to Chemical Inhibitors……
- RNA Polymerase…….
- Bacteriorhodopsin



By Prof Amer Jamil