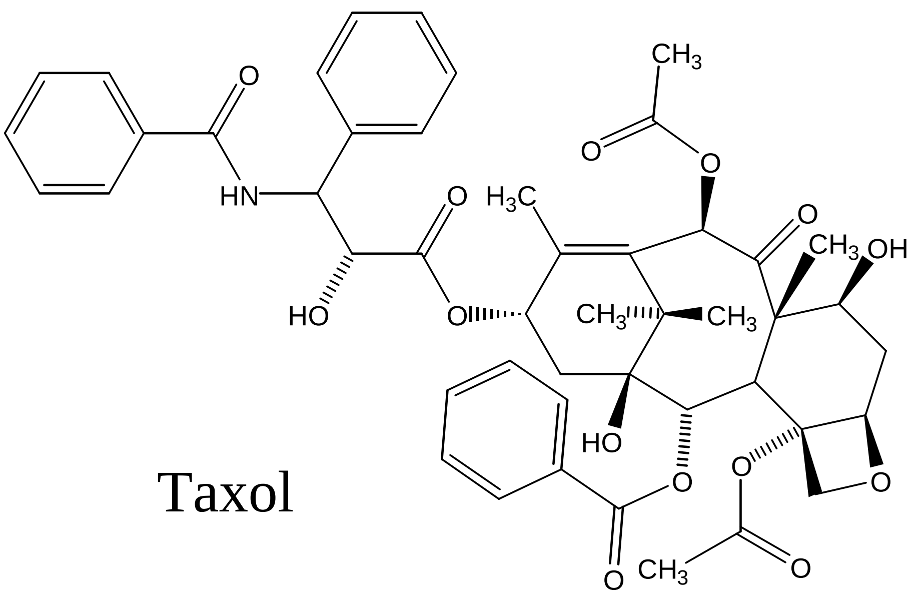
Application of Enzymes and Microorganisms for Organic Synthesis

Jay Keasling



Taxol

Taxol is extracted from the Pacific Yew

- The Pacific yew tree is an environmentally protected species and one of the slowest growing trees in the world.
- Isolation of the compound, which is contained in the bark, involves killing the tree.
- One 100-year old tree results in approximately 350 mg of taxol, just enough for one dose for a single cancer patient.

Total synthesis of taxol

Total synthesis of taxol

K. C. Nicolaou*†, Z. Yang*, J. J. Liu*, H. Ueno*, P. G. Nantermet*, R. K. Guy*, C. F. Claiborne*, J. Renaud*, E. A. Couladouros*, K. Paulvannan* & E. J. Sorensen*†

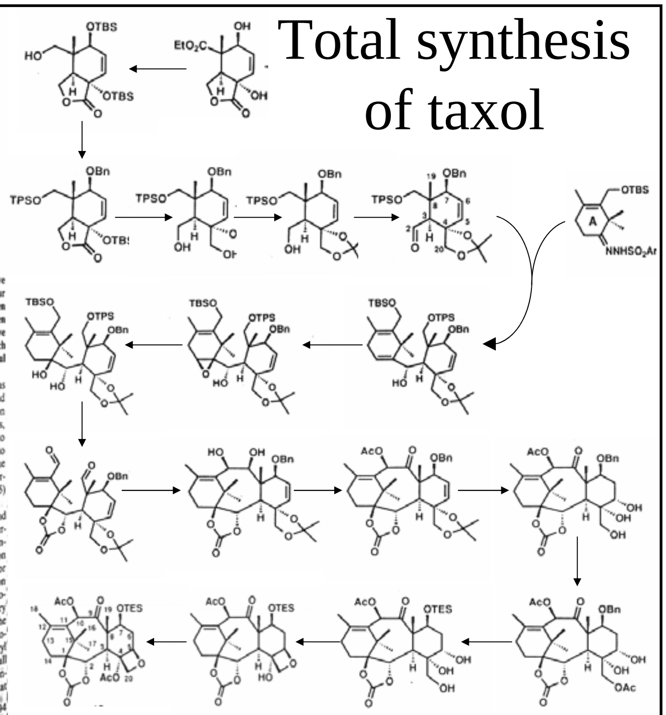
* Department of Chemistry, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA
† Department of Chemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA

TAXOL[1-4], a substance originally isolated from the Pacific yew tr (*Taxus brevifolia*) more than two decades ago, has recently bee approved for the clinical treatment of cancer patients. Hailed a having provided one of the most significant advances in canc therapy[5], this molecule exerts its anticancer activity by inhibitin mitosis through enhancement of the polymerization of tubulin ar consequent stabilization of microtubules[6]. The scarcity of tax and the ecological impact of harvesting it have prompted extensive searches for alternative sources including semisynthesis, cellular culture production and chemical synthesis[2,3]. The latter has been attempted for almost two decades, but these attempts have been thwarted by the magnitude of the synthetic challenge. Here we report the total synthesis of taxol by a convergent strategy, which opens a chemical pathway for the production of both the natural product itself and a variety of designed taxoids.

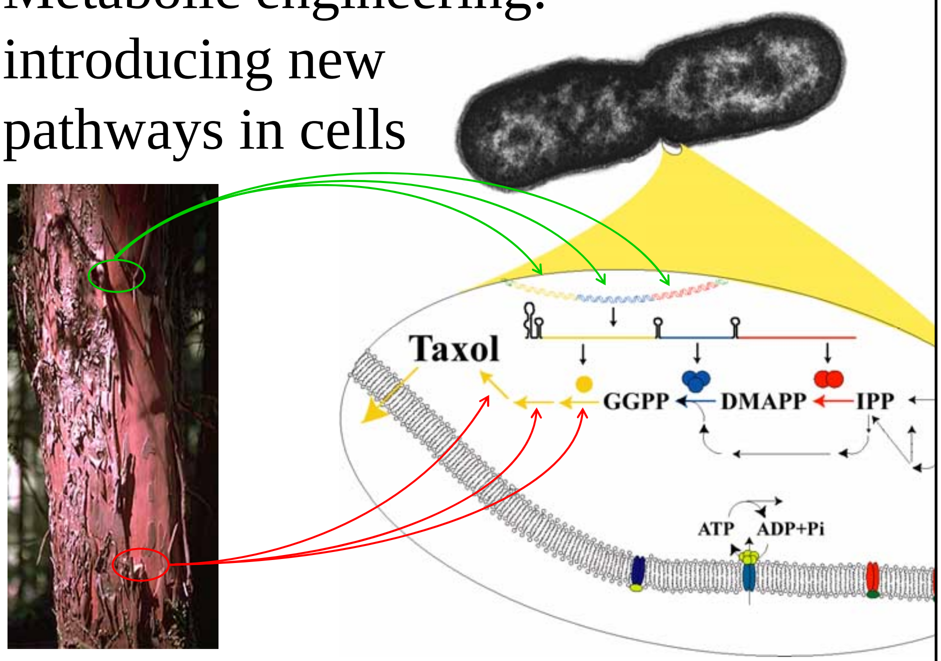
The strategy for the present synthesis of taxol (**1**, Fig. 1*a*) was based on a retrosynthetic analysis involving the bond disconnections[7] shown in Fig. 1*b*. Thus, in the synthetic direction the following key operations were proposed: (1) two fragments, representing precursors to rings A and C (see Fig. 1*a*), were to be coupled by a Shapiro reaction[8] and a McMurry coupling[9] to assemble the ABC ring skeleton; (2) instalment of the oxetane ring; (3) addition[10] of the various substituents around the peripheries of rings B and C; (4) oxygenation[10] at C-13; and (5) esterification to attach the side chain[11].

The previously reported intermediates **2** (ref. 12) (Fig. 2) and **8** (refs 7, 13) (Fig. 3) served as the starting points for the convergent synthesis of taxol reported here. Figure 2 presents the construction of the requisite C-ring aldehyde **7** from **2**. Protection of both hydroxyl groups in **2** with TBS groups (95%) (for abbreviations see figure legend) followed by selective reduction of the ester group with $LiAlH_4$ at 0 °C, furnished primary alcohol **3** (94% yield). Acid-catalysed deprotection of the secondary alcohol in **3** proceeded in a highly selective manner to give the corresponding diol (90% yield), which was then selectively protected with a TPS group at the primary position and a benzyl group at the secondary to afford compound **4** in 80% overall yield. The γ-lactone in **4** was then reductively opened with concomitant desilylation at the tertiary position using $LiAlH_4$ at

NATURE · VOL 367 · 17 FEBRUARY 1994



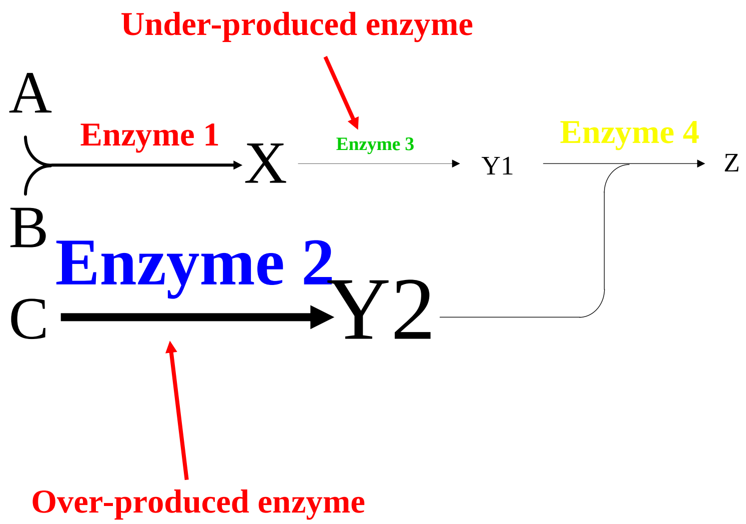
Metabolic engineering: introducing new pathways in cells



Synthesis of complex molecules

- May require several enzymes from one or more organisms
- Expression of genes must be balanced
  - Underexpression of any one gene may limit flux through the pathway and therefore product yields
  - Overexpression will lead to inefficiencies
- Precursors (from inside the cell or supplied from outside the cell) should not severely limit production of the desired product

Balancing enzymatic reactions in the cell

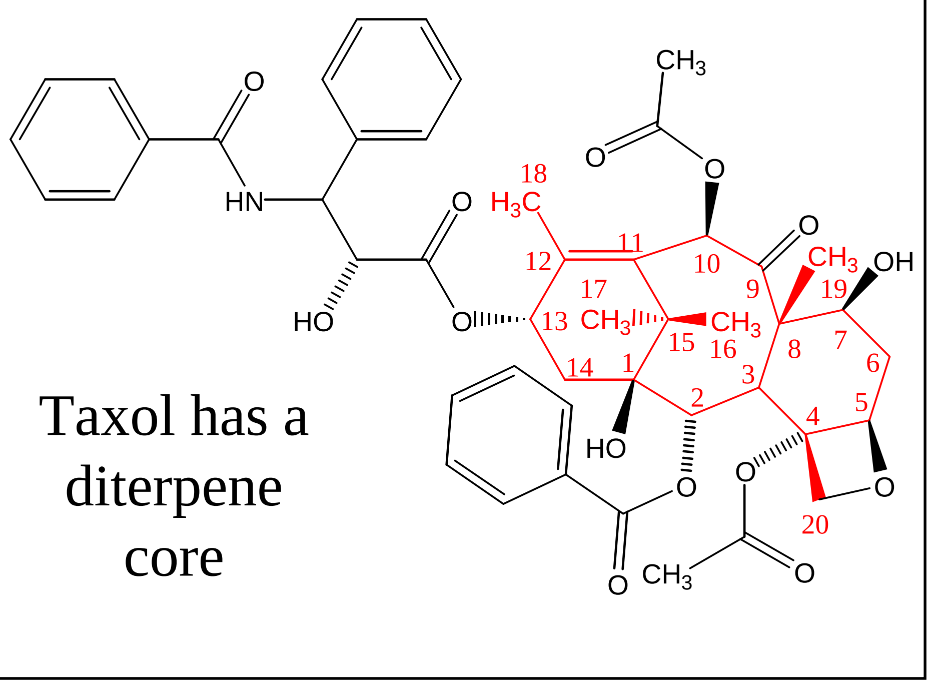


What is metabolic engineering?

- Metabolic engineering is a redirection of enzymatically-catalyzed reactions for the production of a new compound or the degradation of a compound
  - genetic modification of a single organism
  - engineering a consortium of organisms

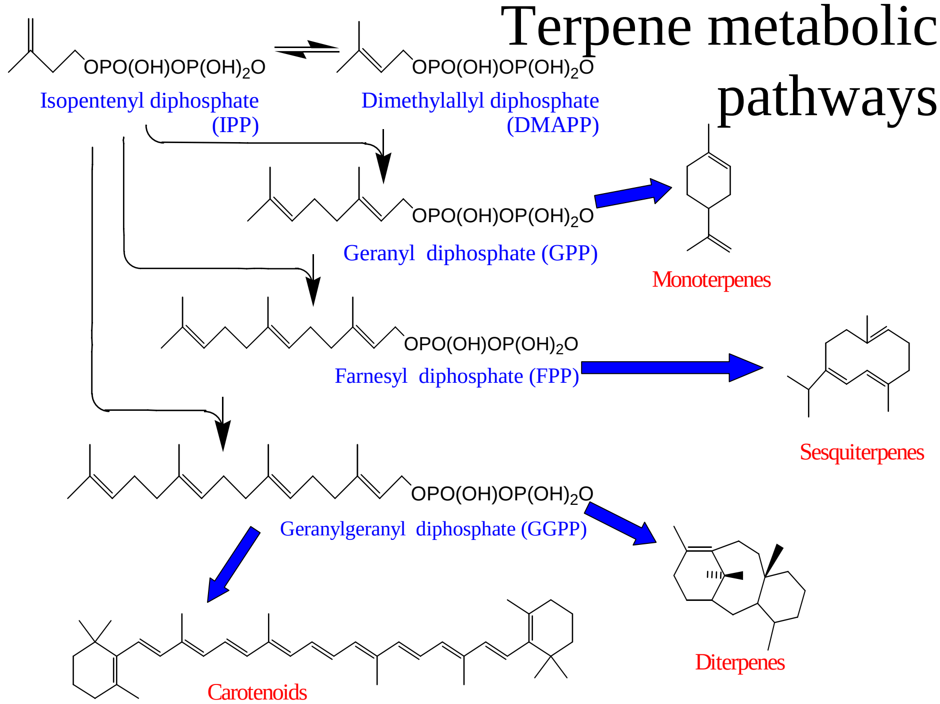
Why do metabolic engineering?

- Production of novel compounds
  - new biopolymers
  - antibiotics
- Production of existing compounds in better ways
- Bioremediation of recalcitrant compounds
  - pesticides/nerve agents
  - PCBs

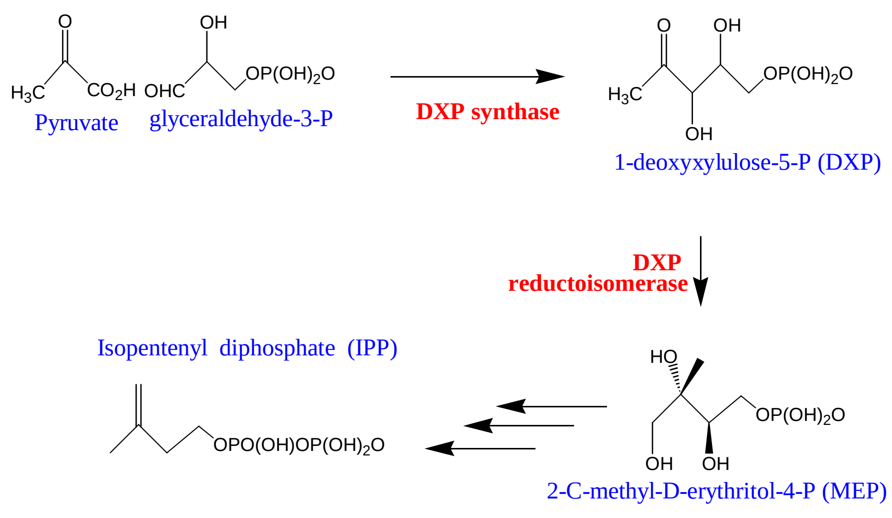


Taxol has a diterpene core

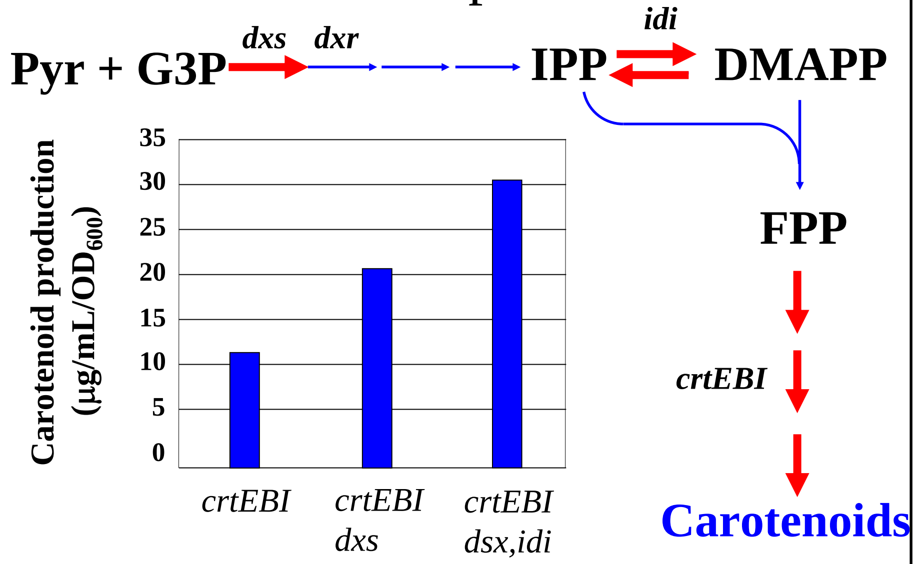
Terpene metabolic pathways



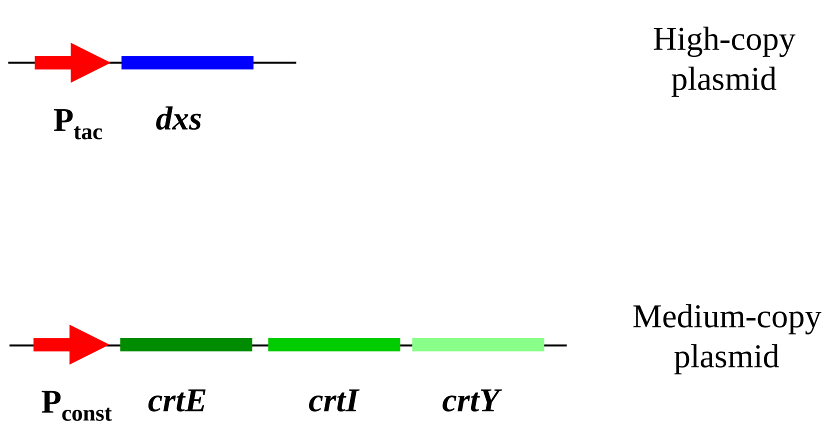
In bacteria, IPP is produced via the non-mevalonate pathway



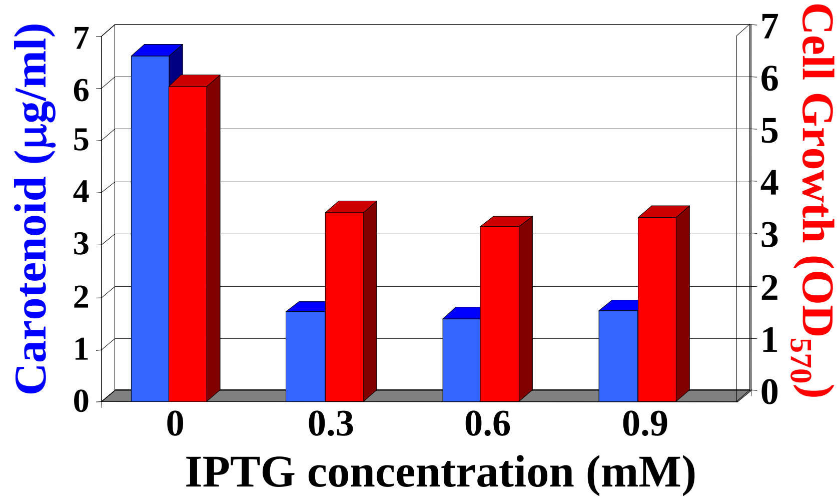
Metabolic engineering of carotenoid production

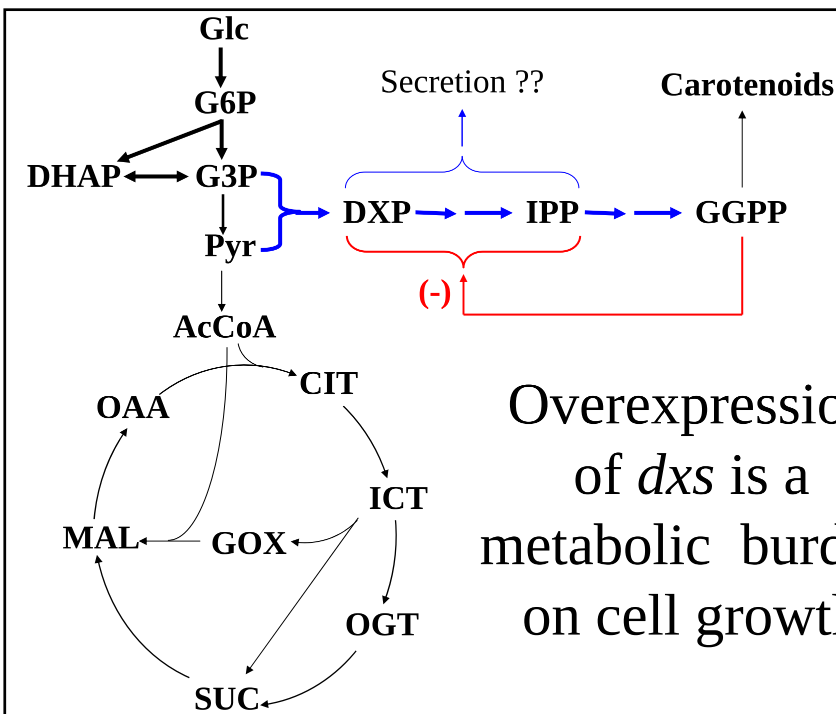


dxs and *dxr* under P_{tac} control on high-copy cloning vectors



Overexpression of *dxs* from a high-copy plasmid with a strong promoter





Overexpression of *dxs* is a metabolic burden on cell growth

Needs for Metabolic Engineering

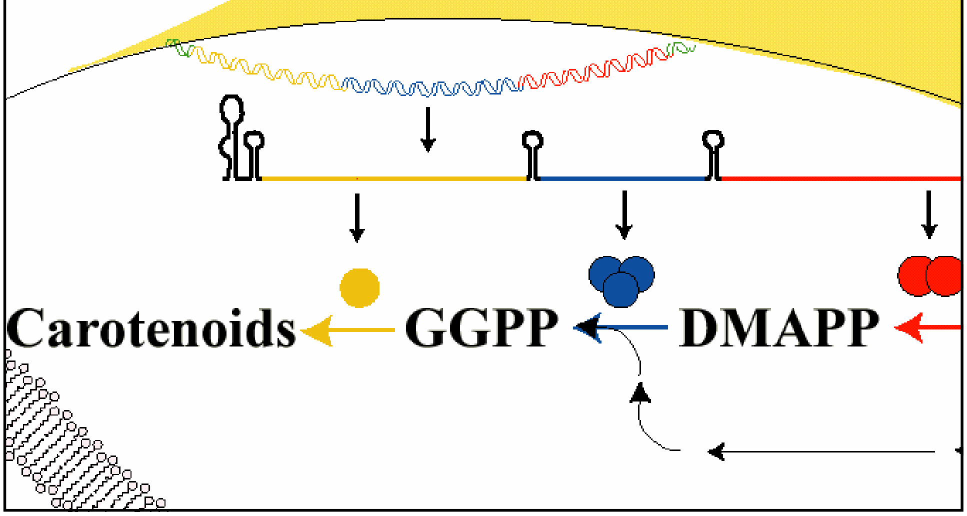
- Optimal fluxes through the heterologous metabolic pathways
- Strict control over gene expression
- Consistent control of gene expression in all cells
- Minimal burden of the heterologous genes on the host

Some key problems in gene expression control

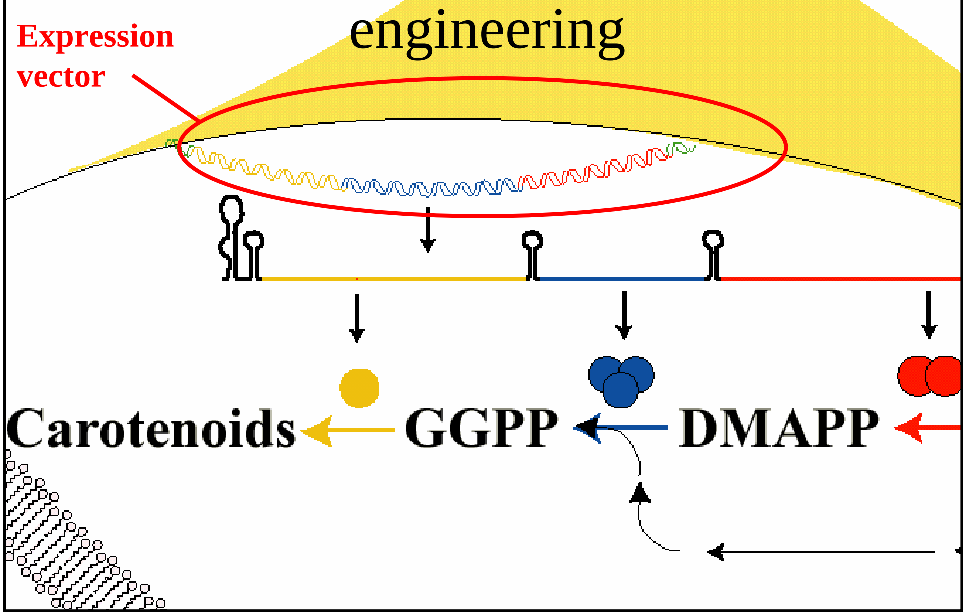
- Many expression vectors are unstable and have variable copy number in the host
- Many promoters do not allow tight and consistent control of gene expression
- There are few techniques to regulate expression of multiple heterologous genes
- Can we predict the levels of gene expression needed for flux redistribution?

Gene expression tools for metabolic engineering:

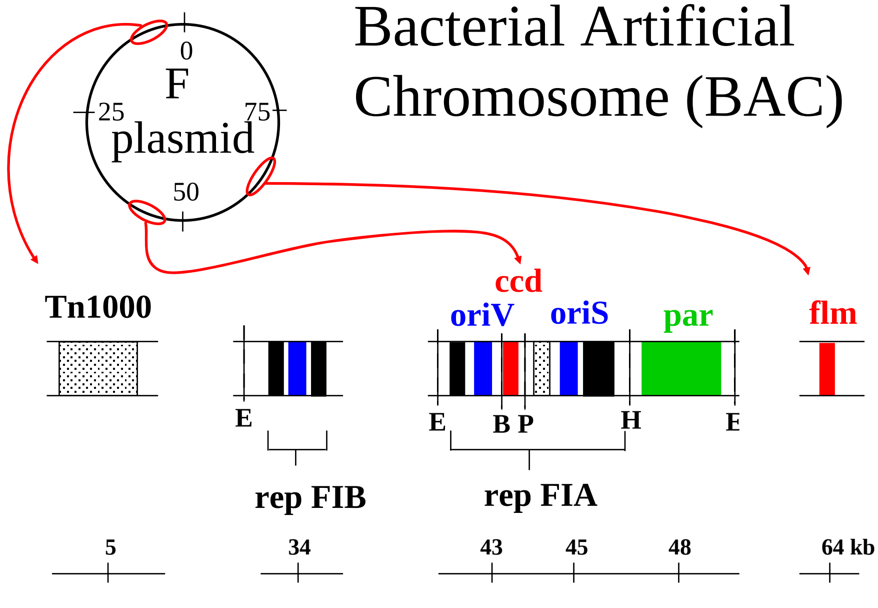
A Metabolic Engineering Toolbox



Gene expression tools for metabolic engineering

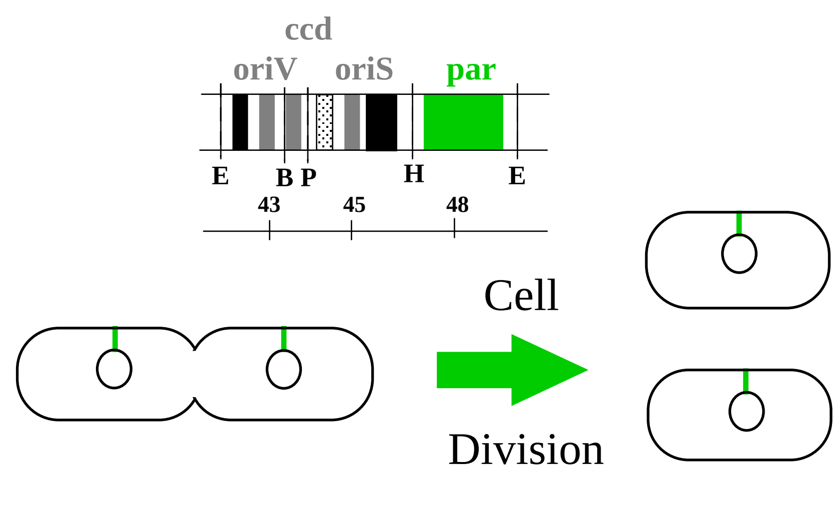


Bacterial Artificial Chromosome (BAC)

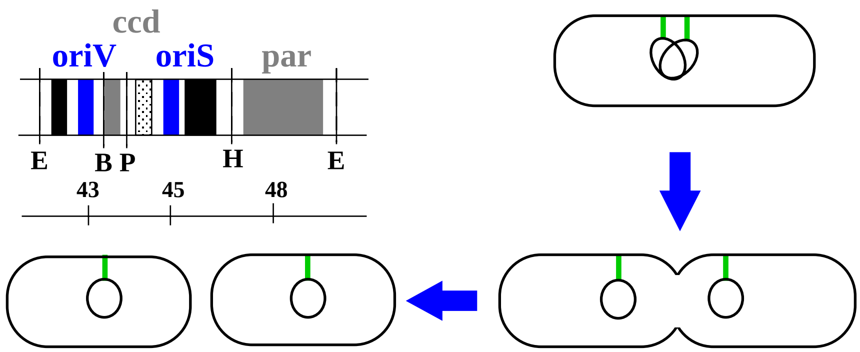


Native F plasmid of *Escherichia coli*

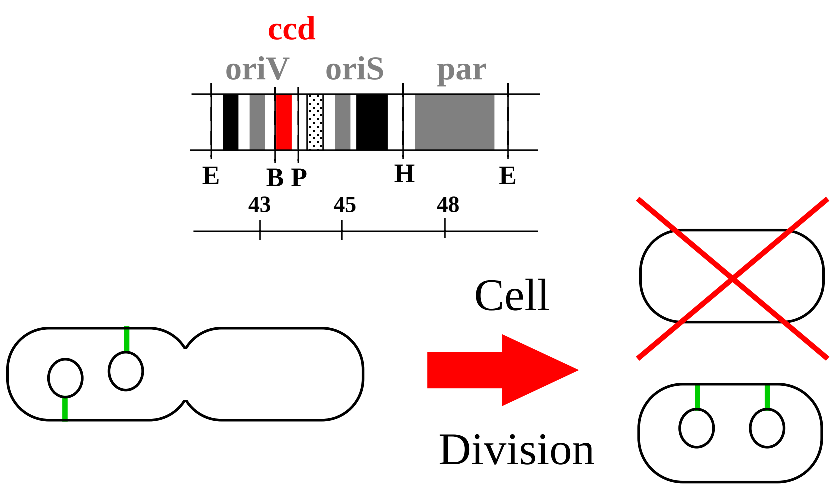
A partition element ensures that each daughter cell will receive a BAC



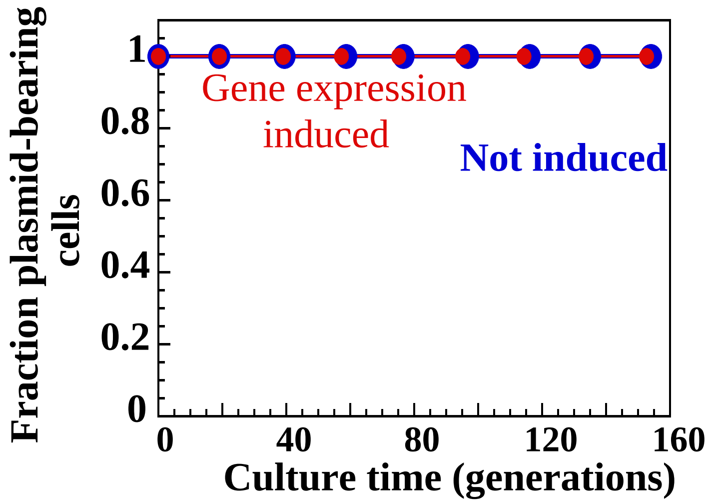
Specific replication origins time BAC replication with the cell cycle



A “Kill” element ensures that BAC-free daughter cells do not survive



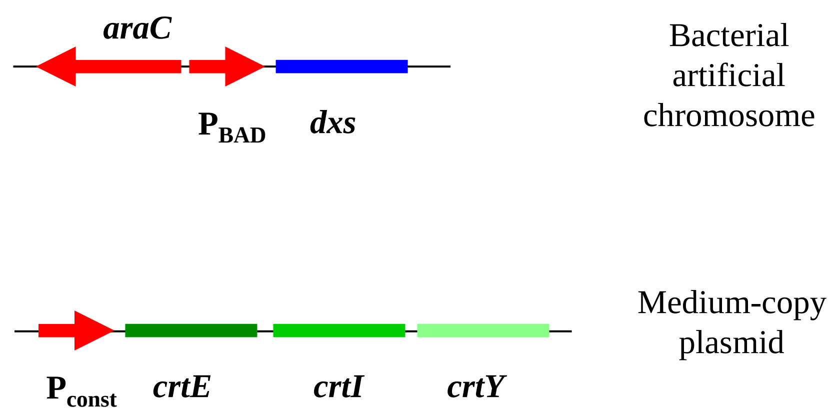
BACs are stable indefinitely in the absence of selection pressure



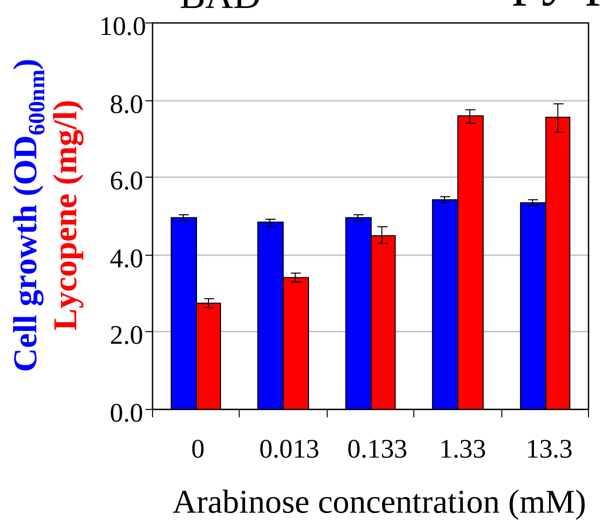
The auxiliary chromosomes have improved control of gene expression

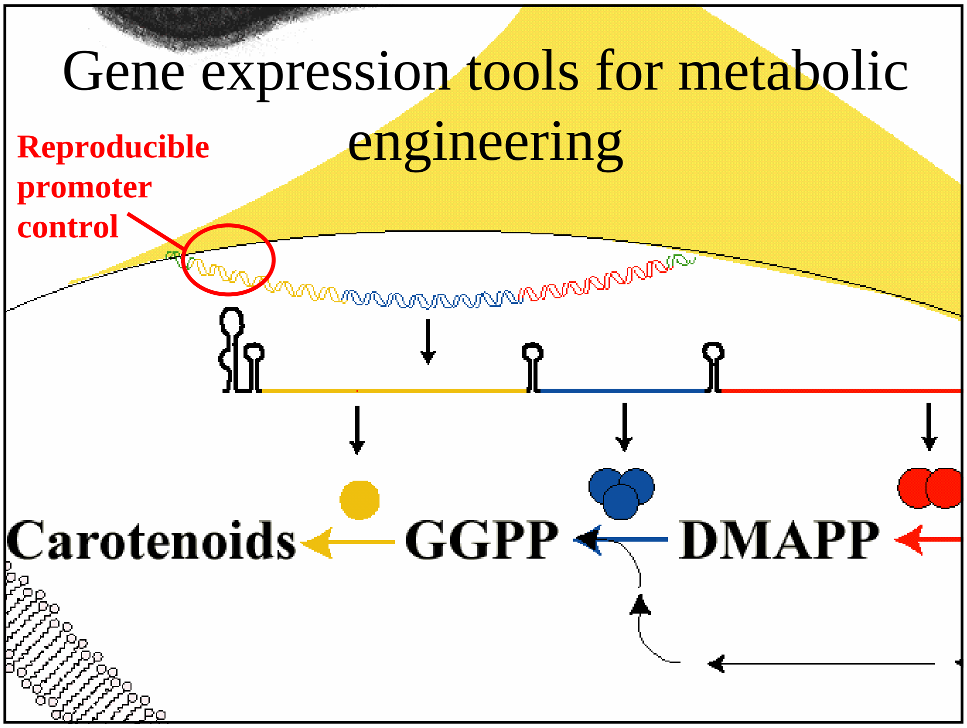
	Uninduced Expression	Induced Expression	Growth Rate of Host
BAC	15 units	4,000 units	0.69 hr^{-1}
High-Copy Plasmid	200 units	12,500 units	0.53 hr^{-1}

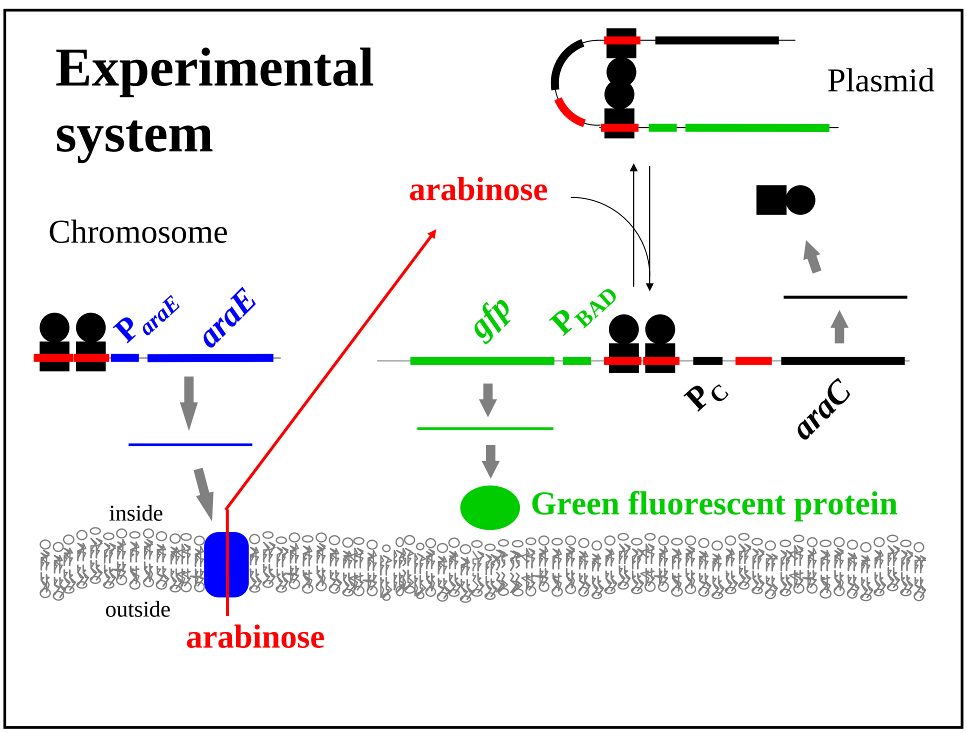
dxs and *dxr* under P_{BAD} control on bacterial artificial chromosome



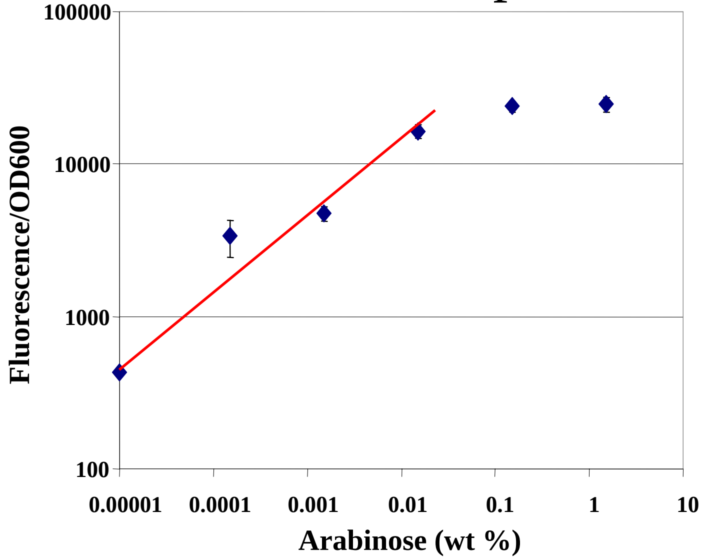
Production of lycopene: *dxs* under control of P_{BAD} on low-copy plasmid



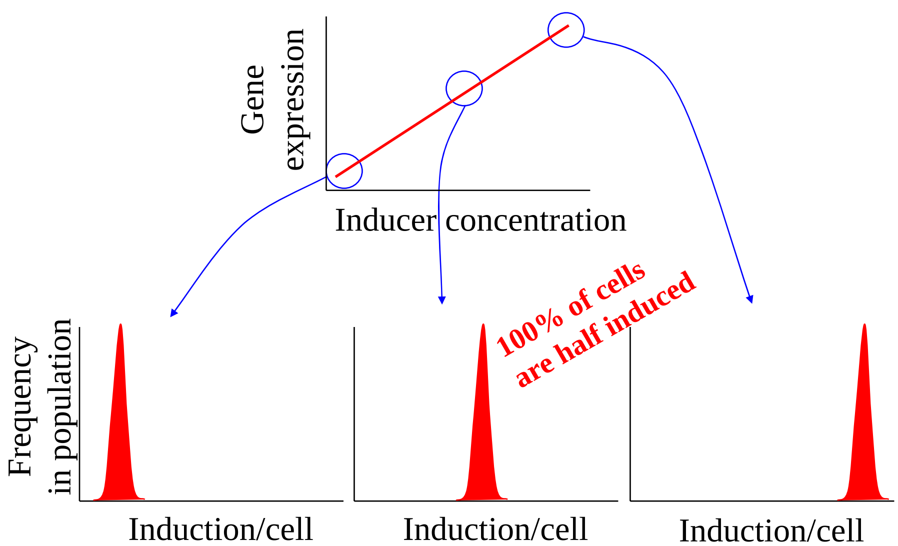

Gene expression tools for metabolic engineering
Reproducible promoter control
Carotenoids
GGPP
DMAPP


Experimental system
Plasmid
arabinose
Chromosome
P araE
araE
gfp
P BAD
P C
araC
Green fluorescent protein
inside
outside
arabinose

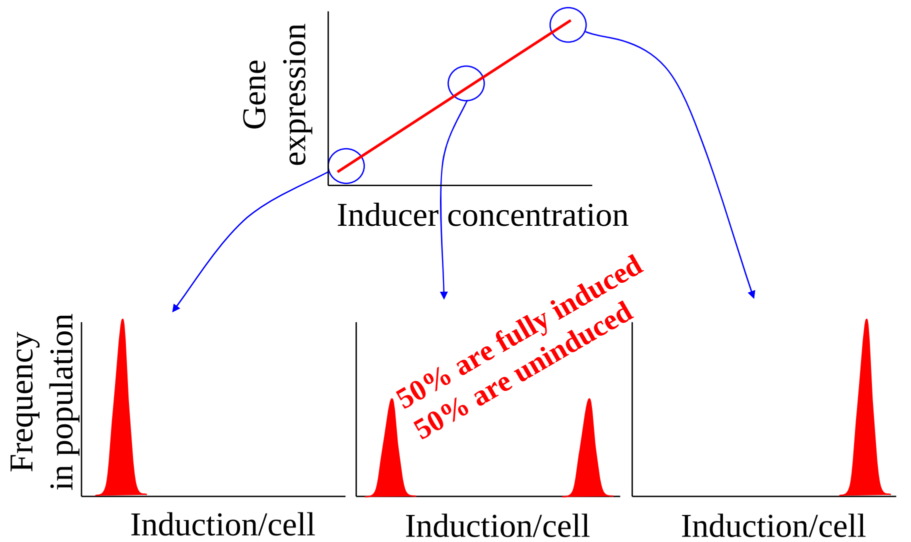
Expression of *gfp* from the arabinose-inducible promoter



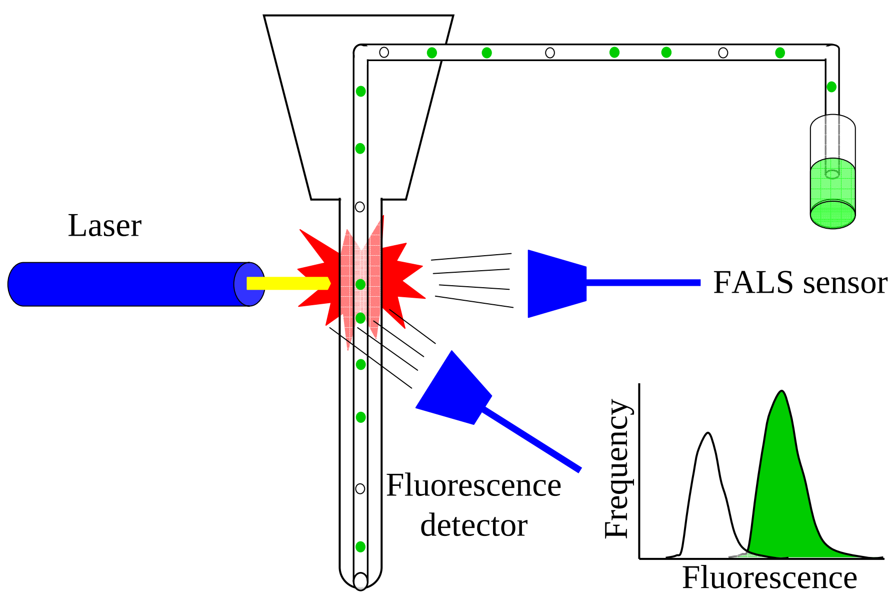
Desired gene expression in population



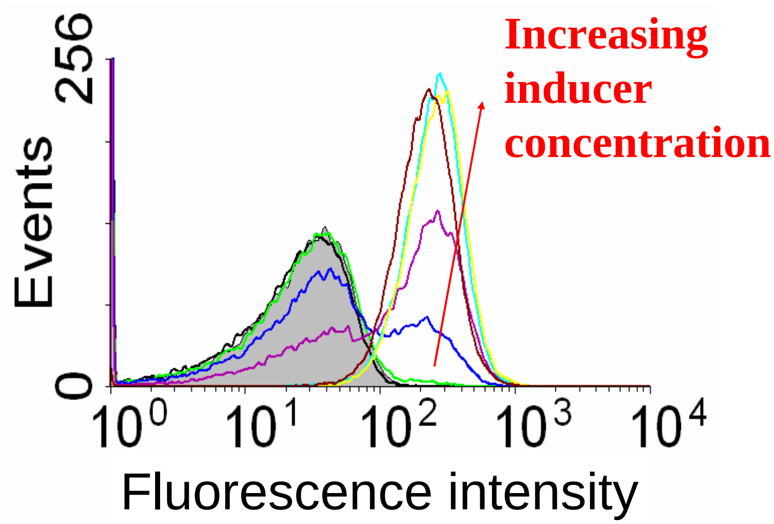
‘All-or-None’ Gene Expression



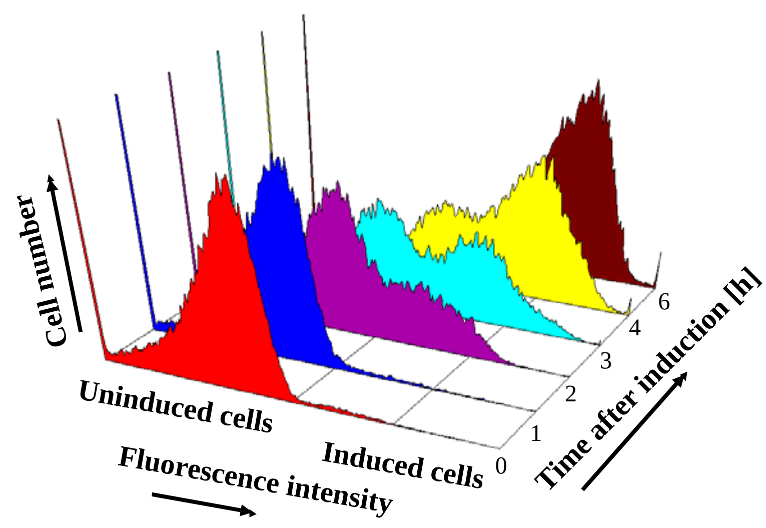
Flow cytometric analysis



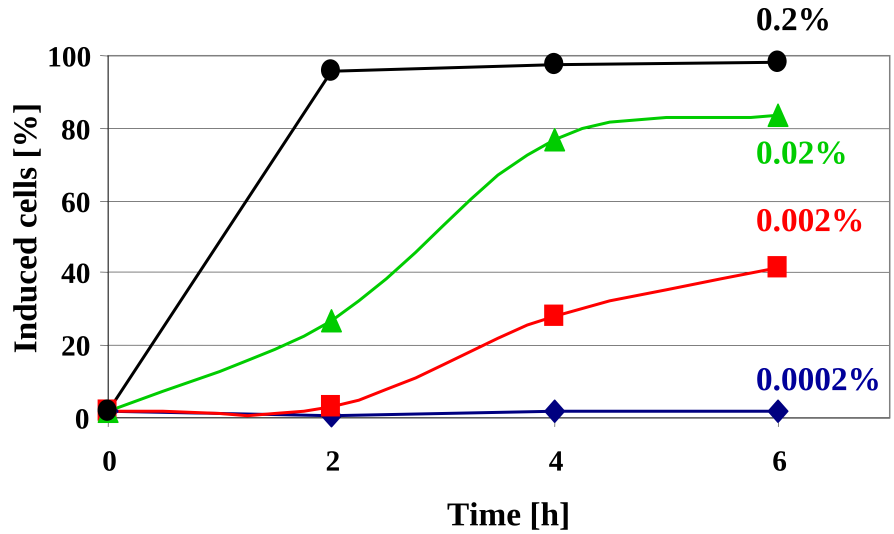
Native arabinose-inducible system gives rise to two populations

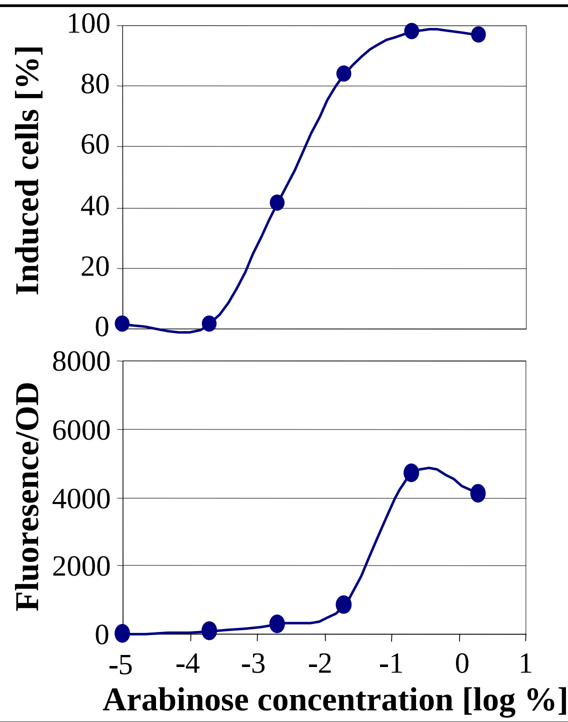


Native arabinose-inducible system gives rise to two populations



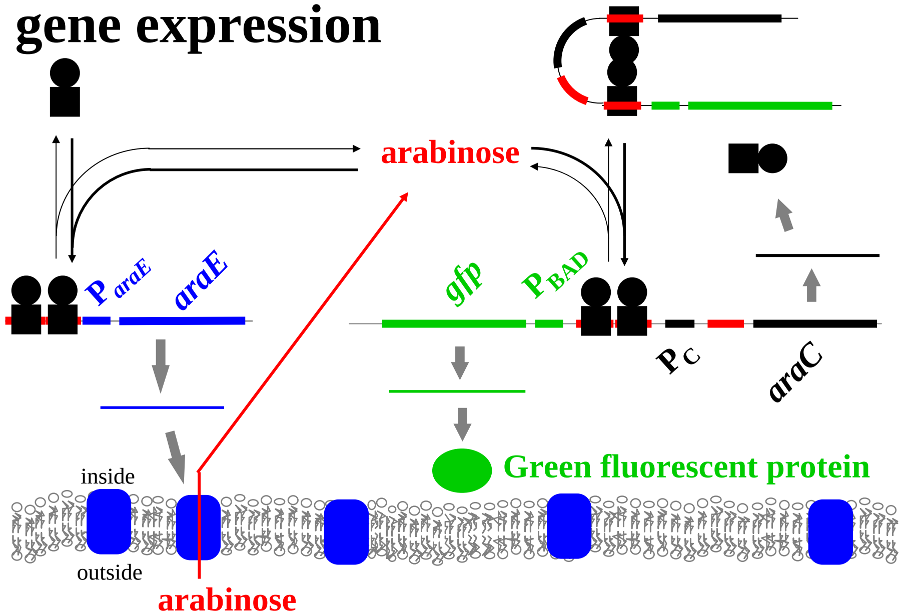
Population dynamics as a function of arabinose concentration



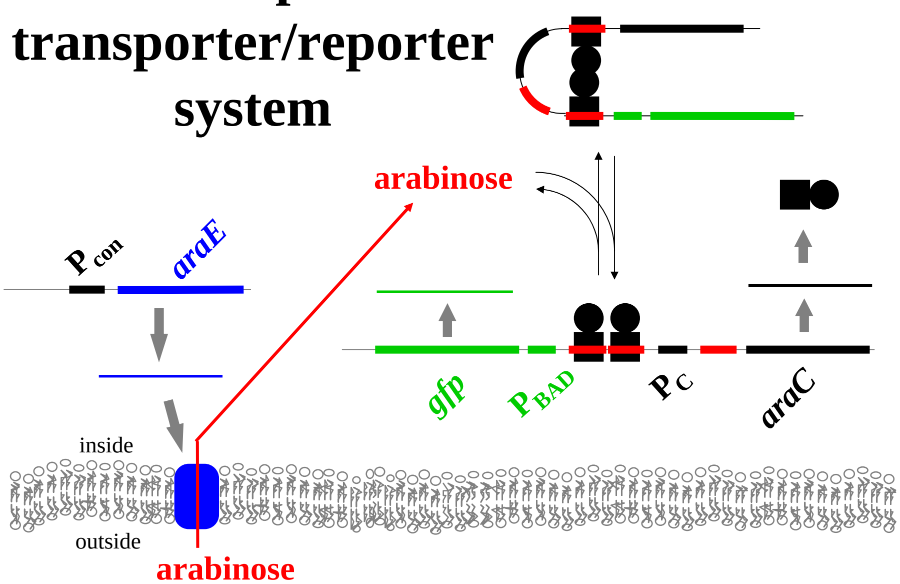


Experiment: steady-state induction levels as a function of inducer concentration

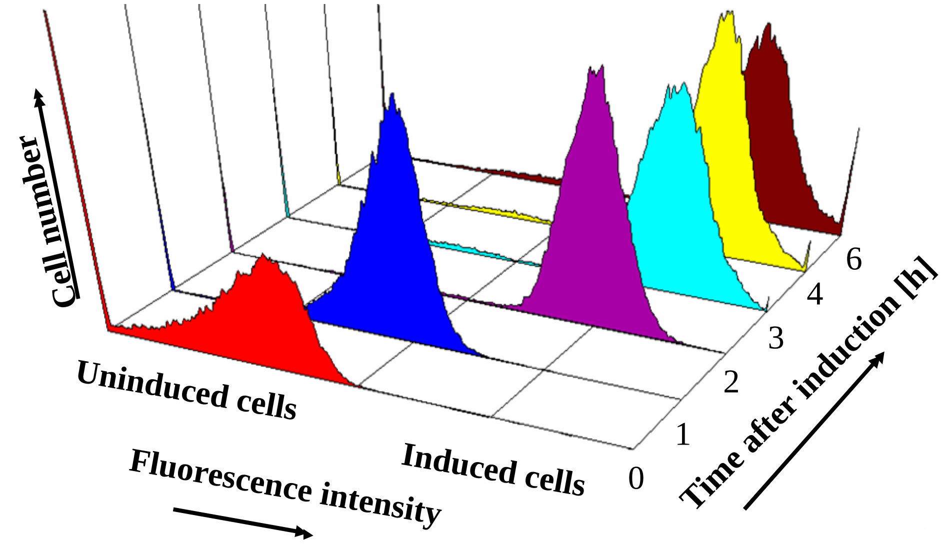
Stimulation of inducer transport gene expression



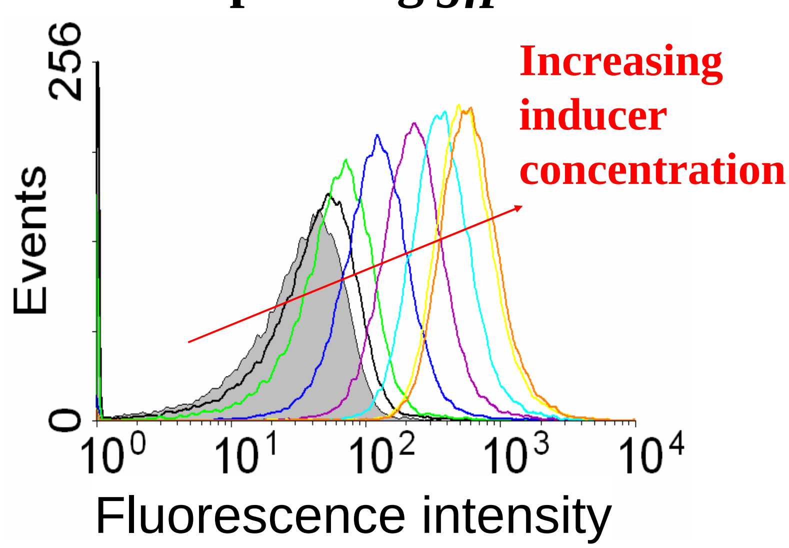
Decoupled transporter/reporter system

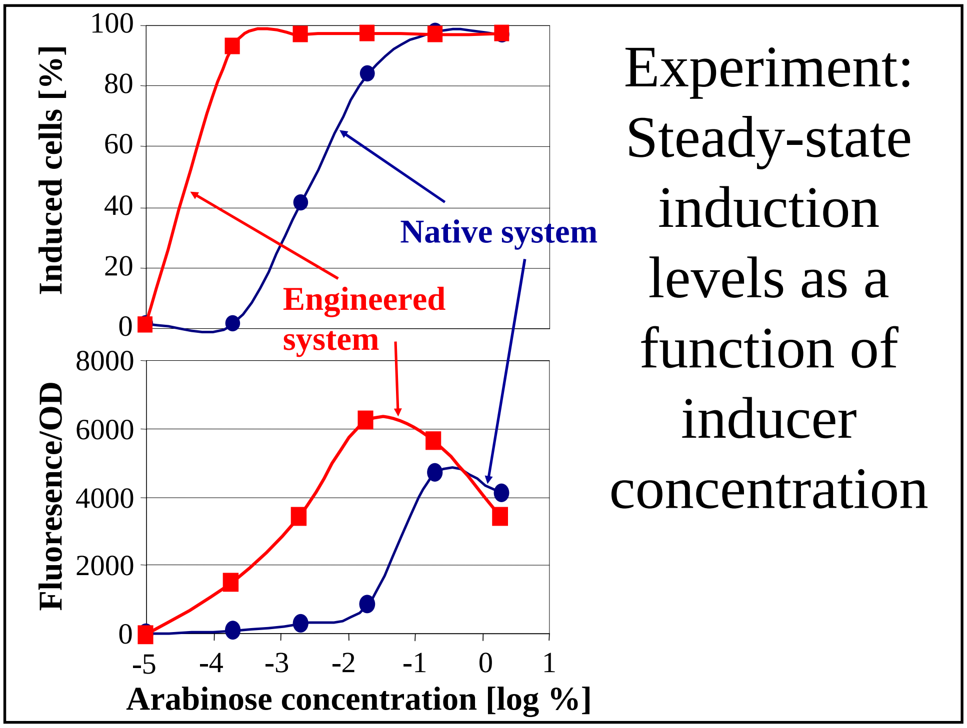


Population analysis of *E. coli* expressing *gfp*

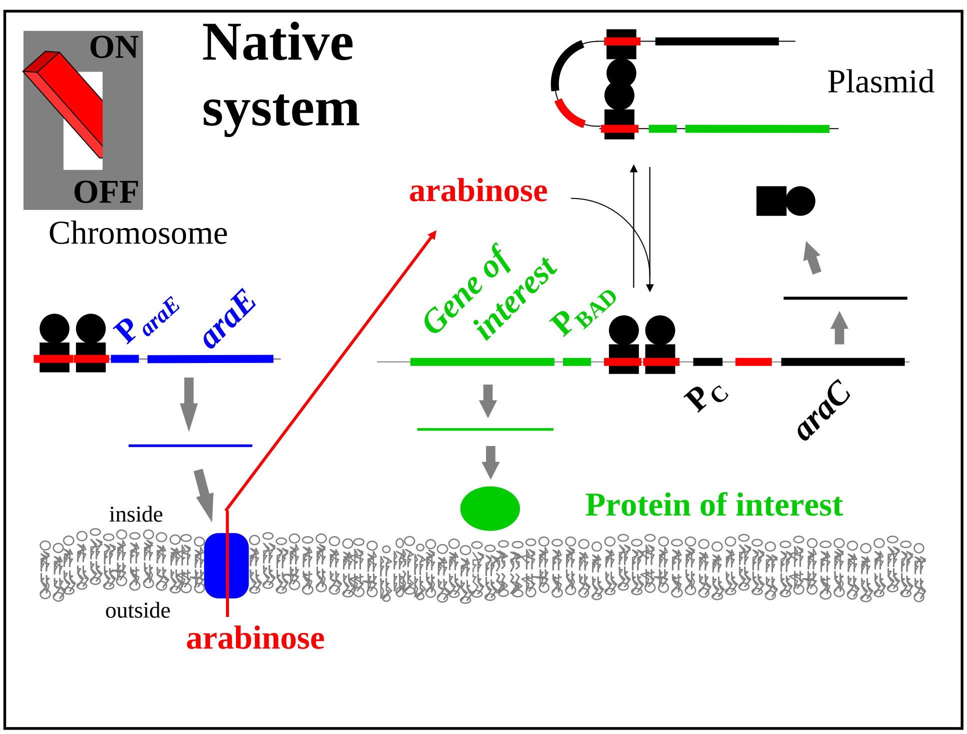


Population analysis of *E. coli* expressing *gfp*

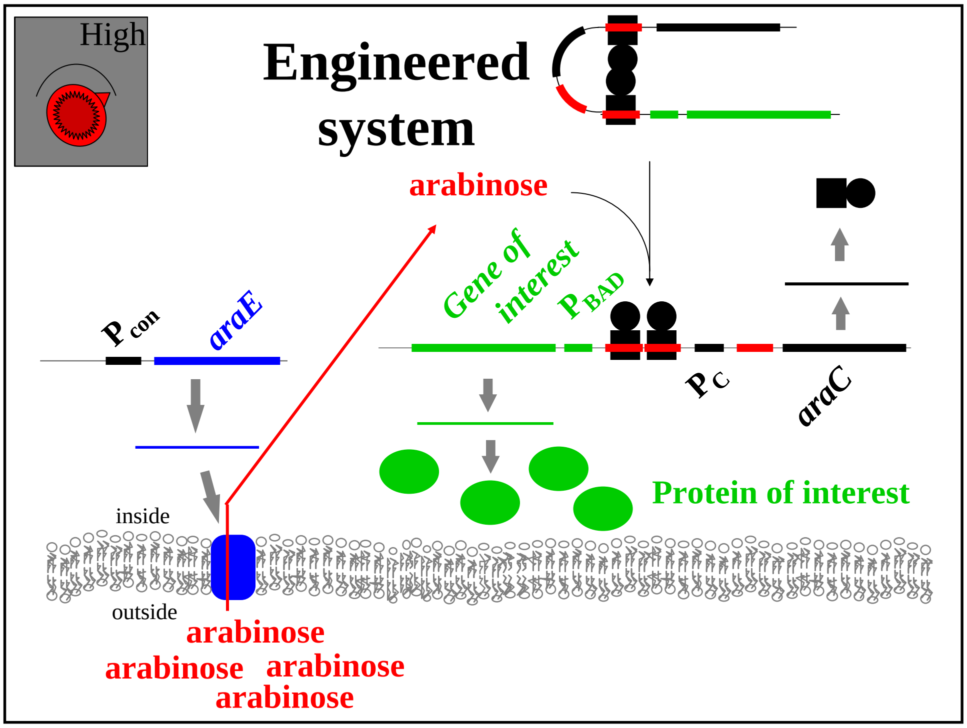


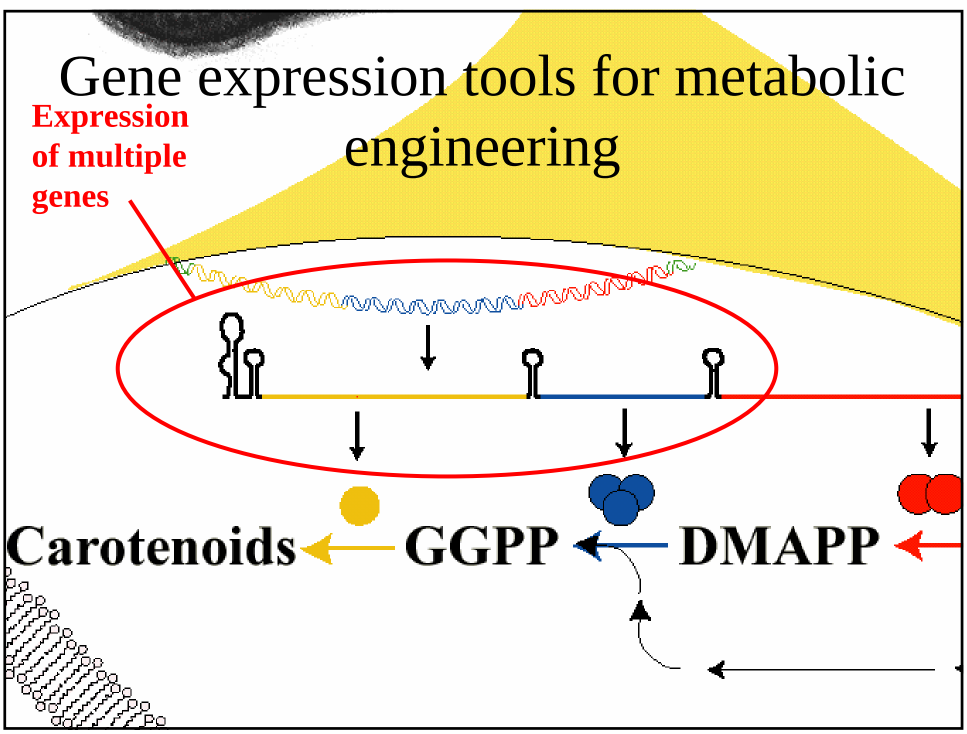


Experiment: Steady-state induction levels as a function of inducer concentration
Induced cells [%]
100
80
60
40
20
0
Fluoresence/OD
8000
6000
4000
2000
0
-5
-4
-3
-2
-1
0
1
Arabinose concentration [log %]
Native system
Engineered system

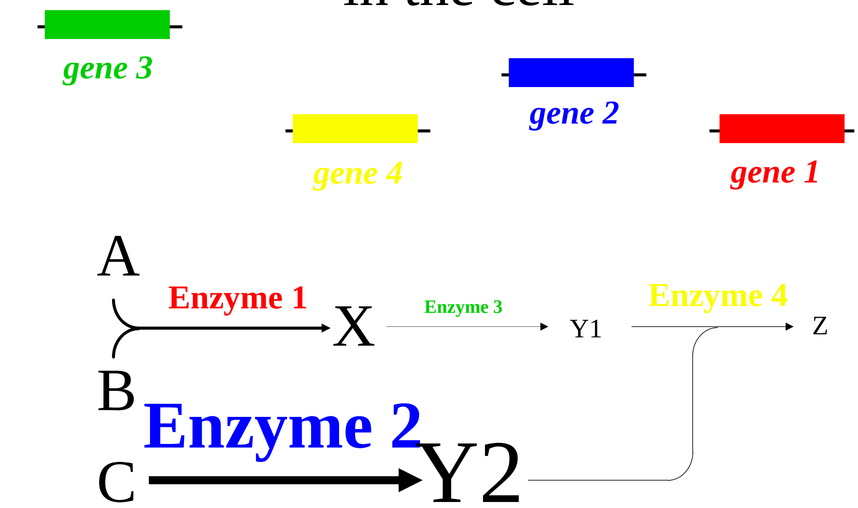


ON
OFF
Native system
Chromosome
Plasmid
arabinose
P araE
araE
Gene of interest
P BAD
P C
araC
inside
outside
arabinose
Protein of interest

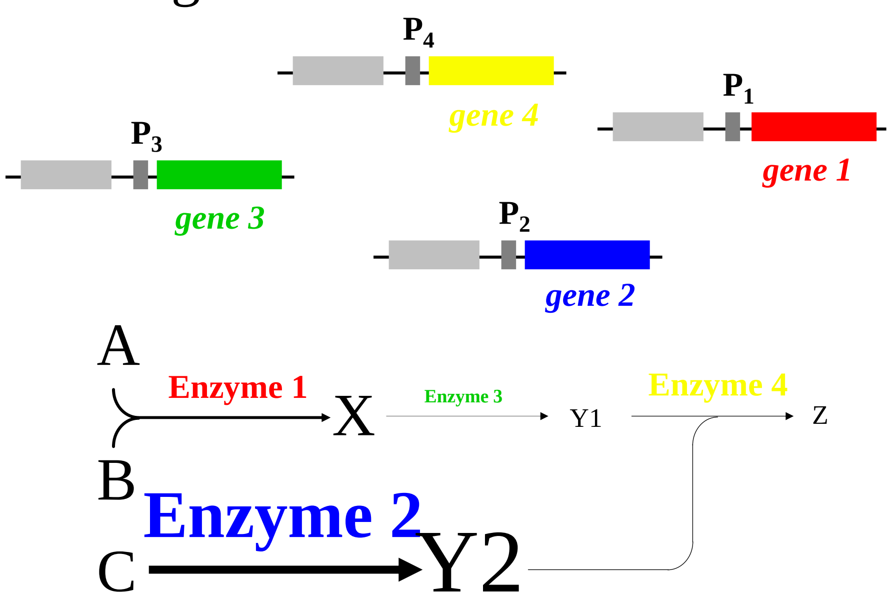

High
Engineered system
arabinose
Gene of interest
P BAD
P con
araE
P C
araC
Protein of interest
inside
outside
arabinose
arabinose arabinose
arabinose


Gene expression tools for metabolic engineering
Expression of multiple genes
Carotenoids
GGPP
DMAPP

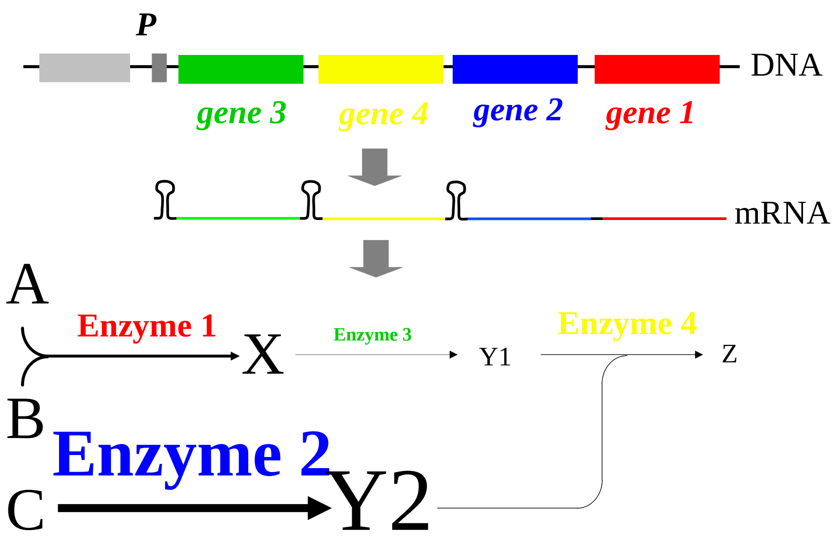
Balancing enzymatic reactions in the cell

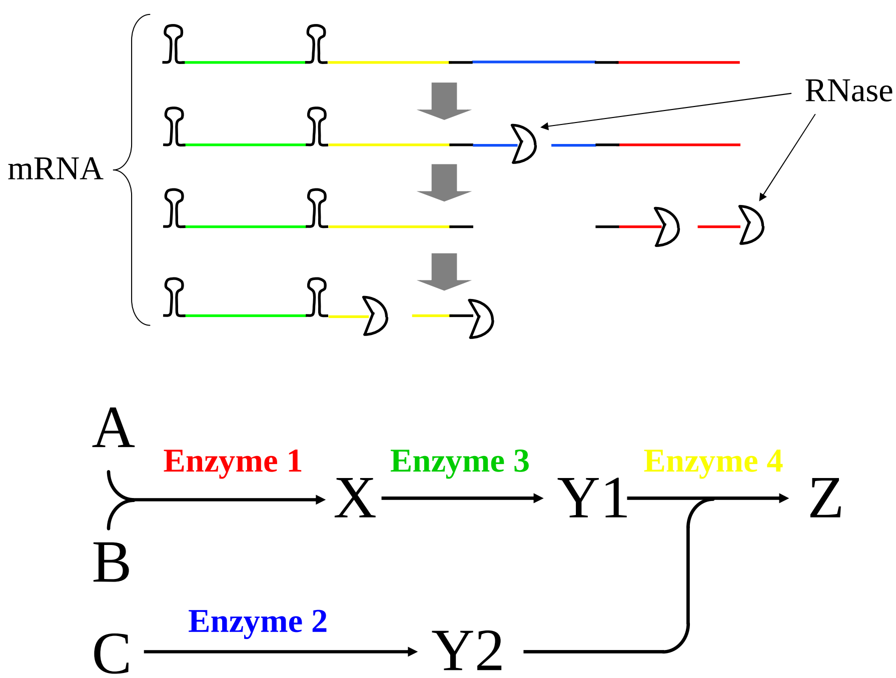


Using individual control elements

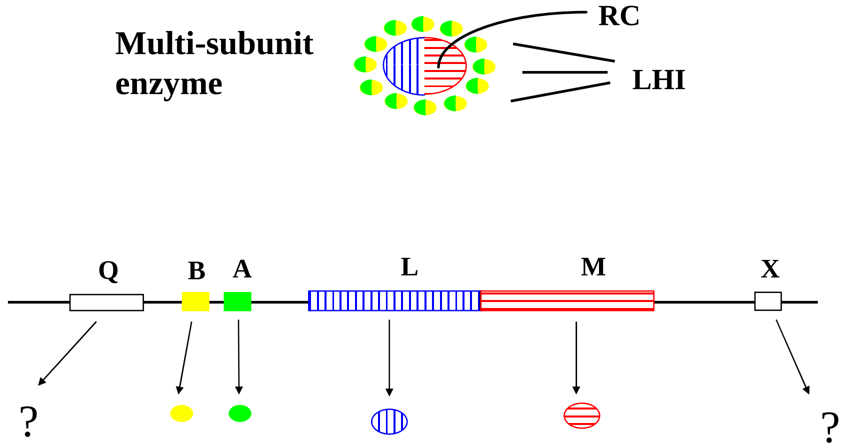


Synthetic operons

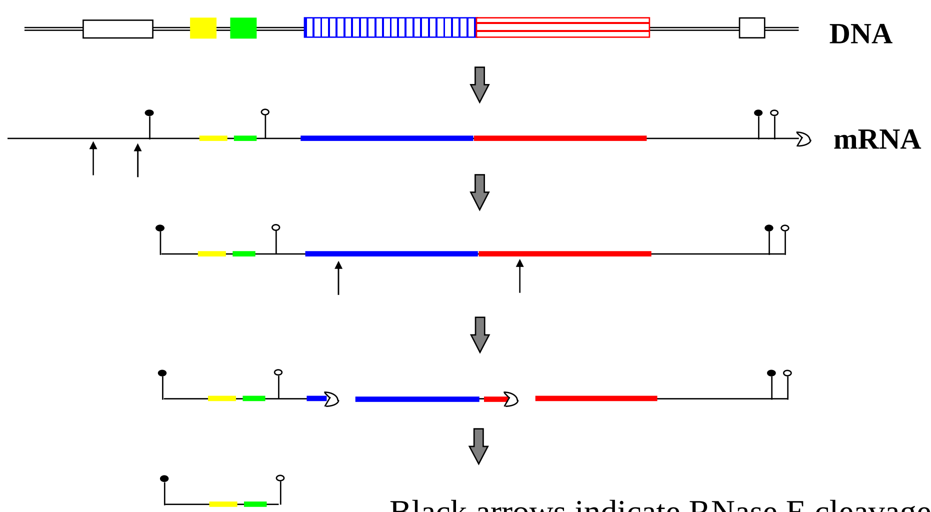




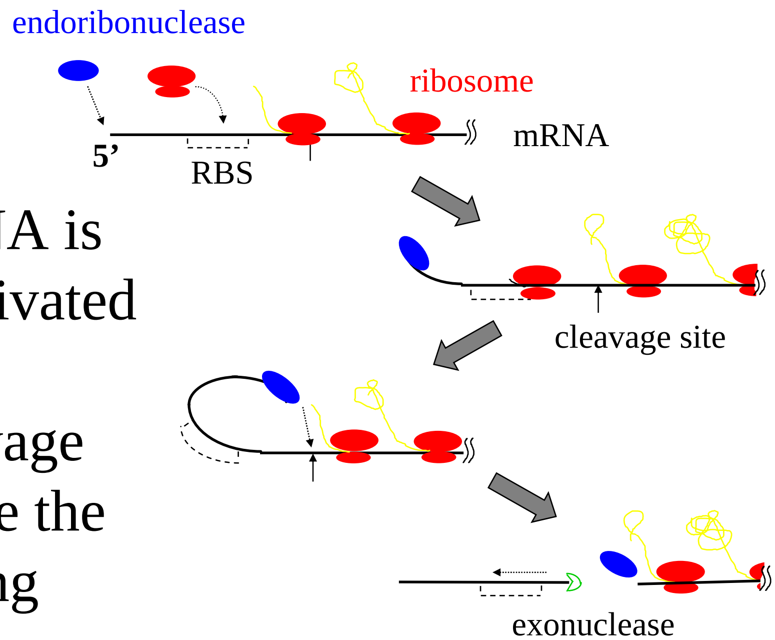
The *puf* operon in *Rhodobacter* encodes a multi-subunit enzyme



The production of enzyme subunits is controlled by mRNA stability

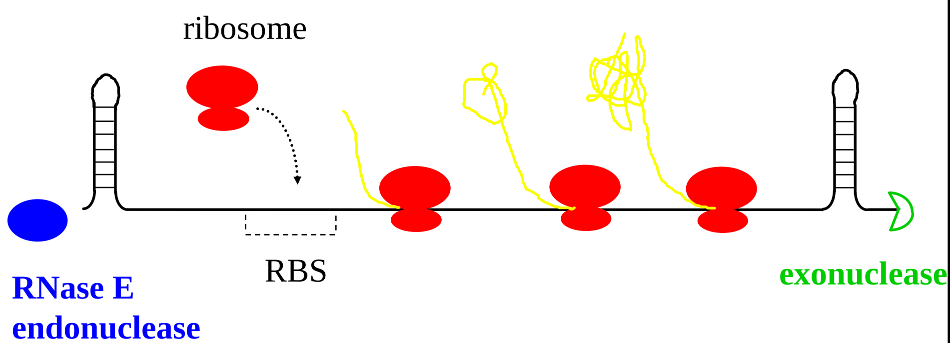


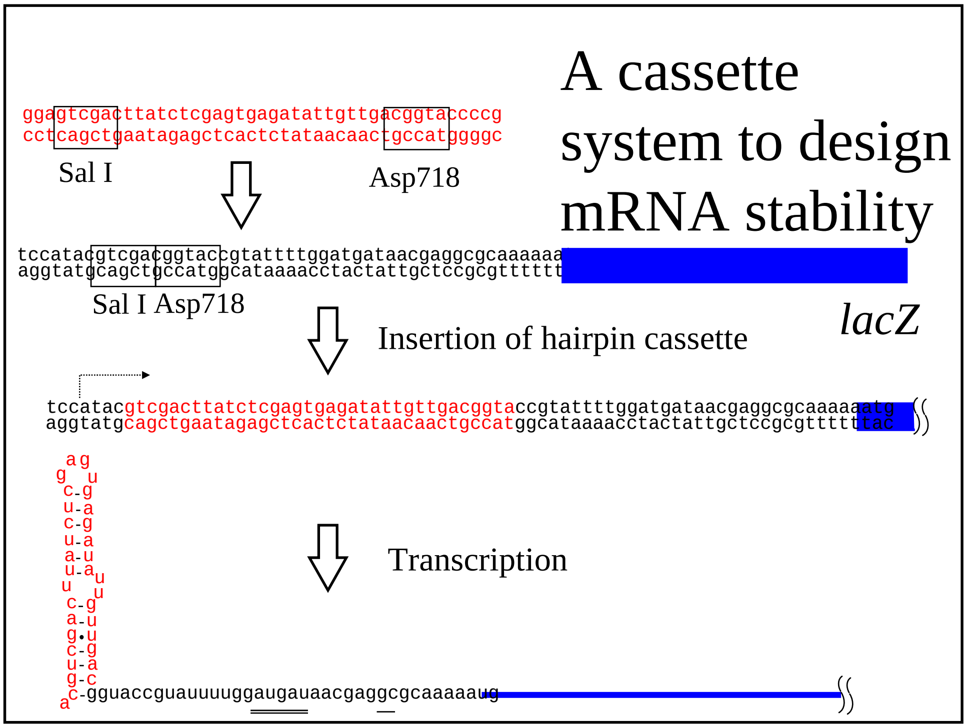
Black arrows indicate RNase E cleavages

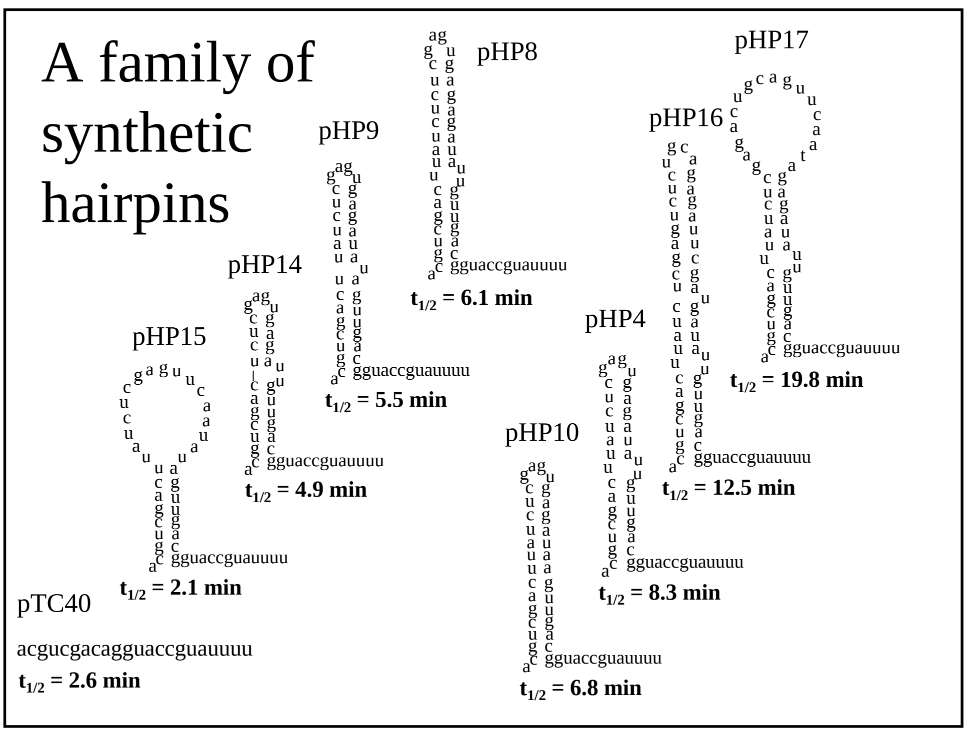


mRNA is inactivated by a cleavage inside the coding region

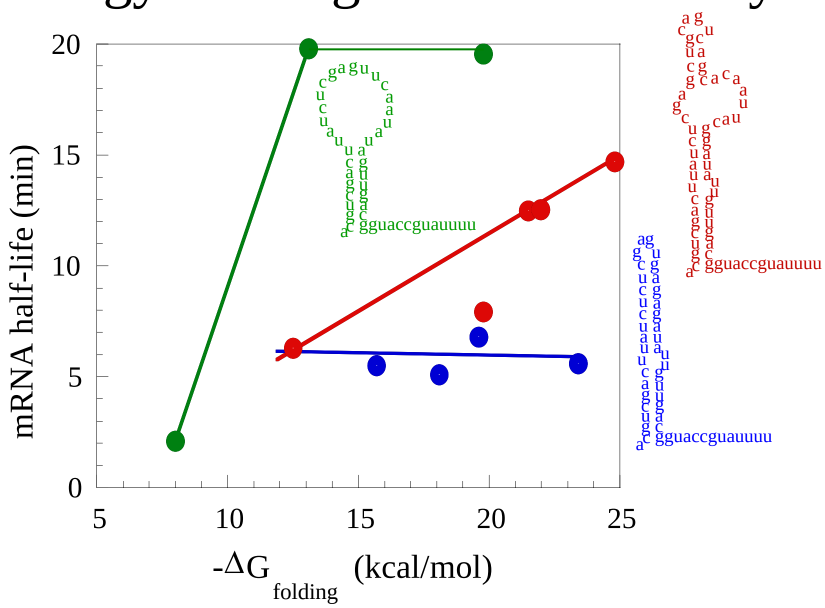
Secondary structures in the mRNA protect natural mRNAs against nucleases



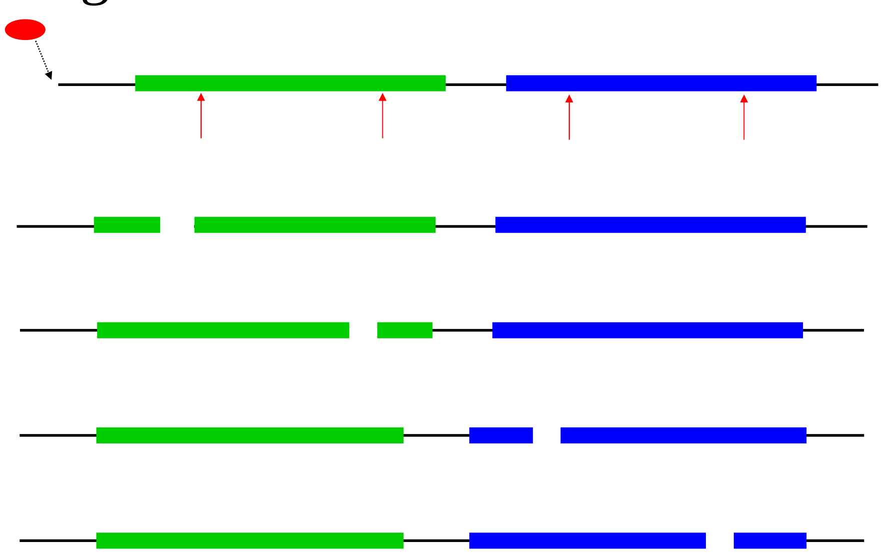

A cassette system to design mRNA stability
ggagtcgacttatctcgagtgagatattgttgacggtaccccg
cctcagctgaatagagctcactctataacaactgccatggggc
Sal I
Asp718
tccatacgtcgacggtaccgtattttggatgataacgaggcgcaaaaaa
aggtatgcagctgccatggcataaaacctactattgctccgcgtttttt
Sal I Asp718
lacZ
Insertion of hairpin cassette
tccatacgtcgacttatctcgagtgagatattgttgacggtaccgtattttggatgataacgaggcgcaaaaaatg
aggtatgcagctgaatagagctcactctataacaactgccatggcataaaacctactattgctccgcgttttttac
Transcription
c-gguaccguauuuuggaugauaacgaggcgcaaaaaug


A family of synthetic hairpins
pHP8
t1/2 = 6.1 min
pHP17
t1/2 = 19.8 min
pHP16
t1/2 = 12.5 min
pHP9
t1/2 = 5.5 min
pHP14
t1/2 = 4.9 min
pHP15
t1/2 = 2.1 min
pHP4
t1/2 = 8.3 min
pHP10
t1/2 = 6.8 min
pTC40
acgucgacagguaccguauuuu
t1/2 = 2.6 min

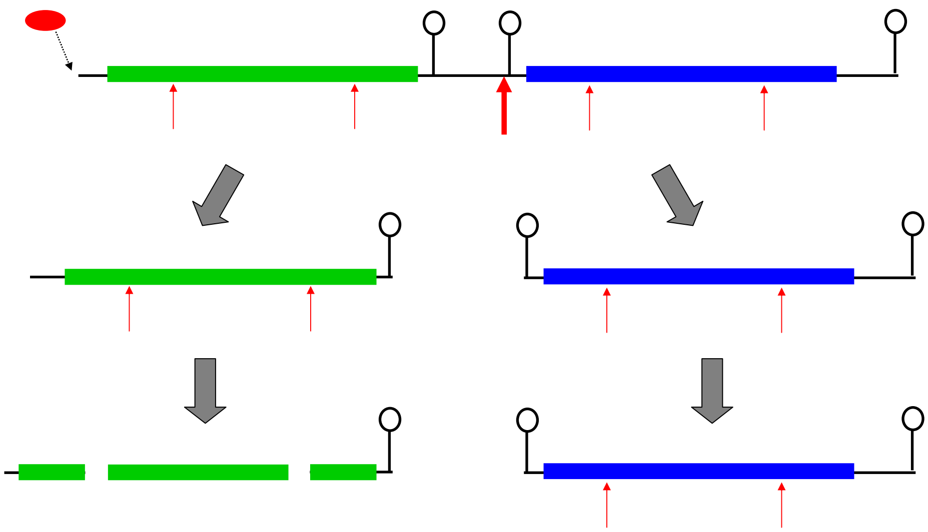
A strategy to design mRNA stability?



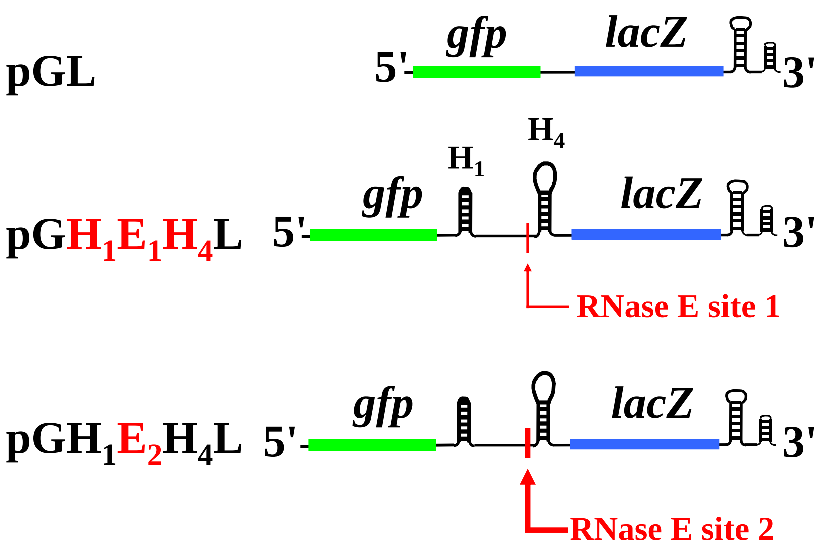
Degradation of multicistronic mRNA



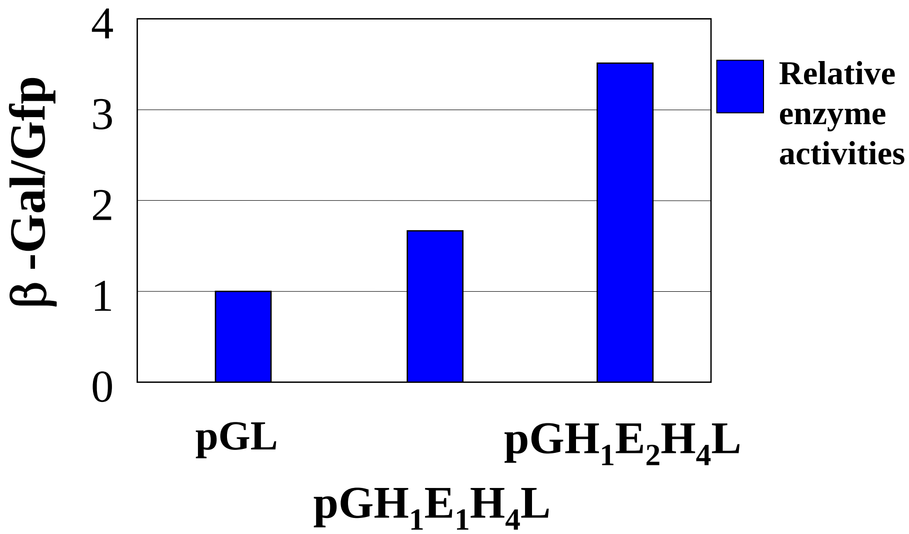
Degradation of multicistronic mRNA



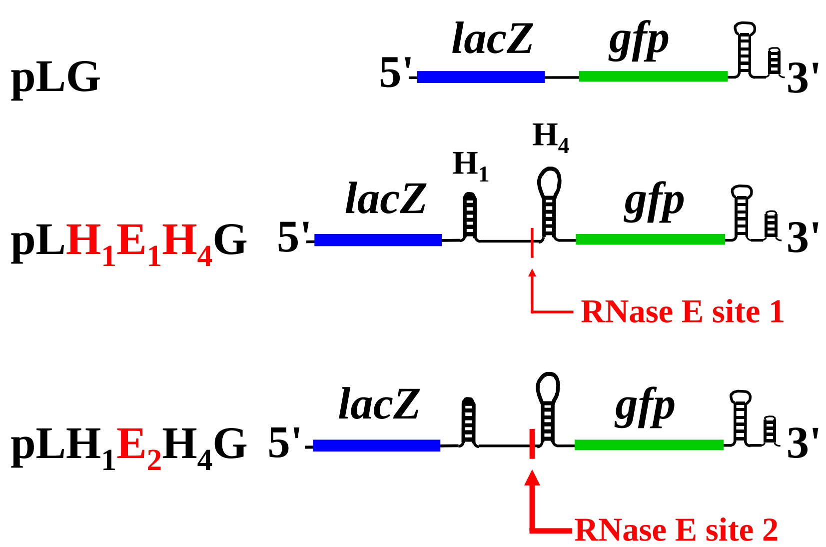
Constructs to test operons



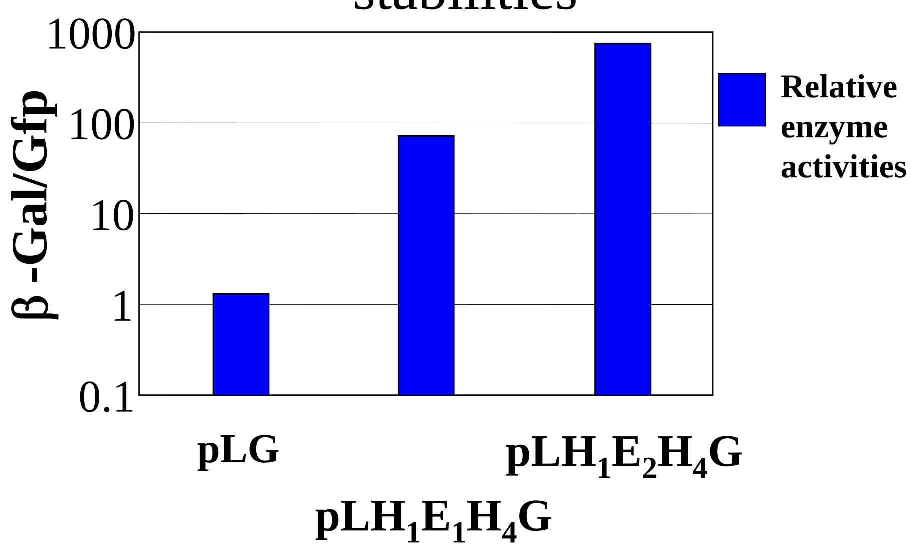
Relative protein levels/mRNA stabilities



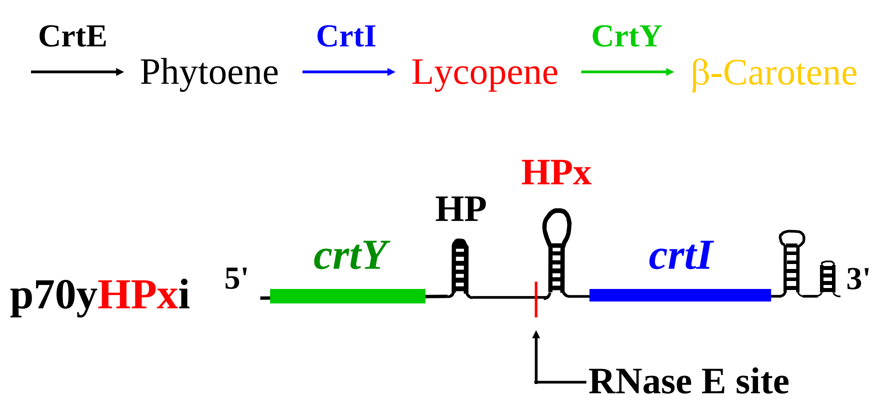
Reversing gene order



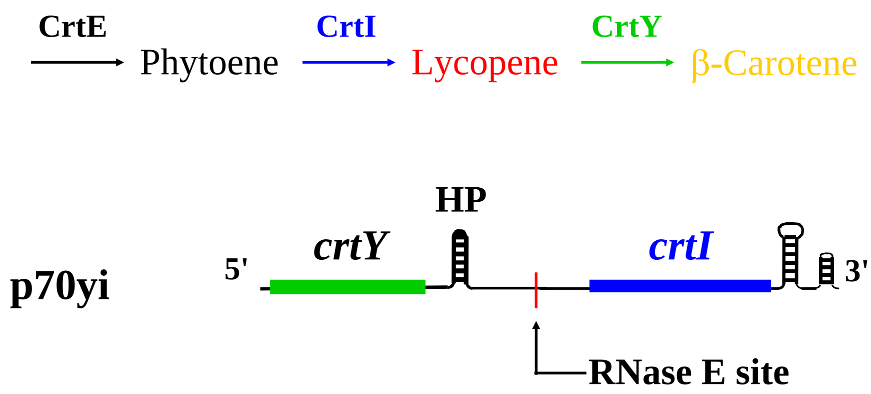
Relative protein levels/mRNA stabilities



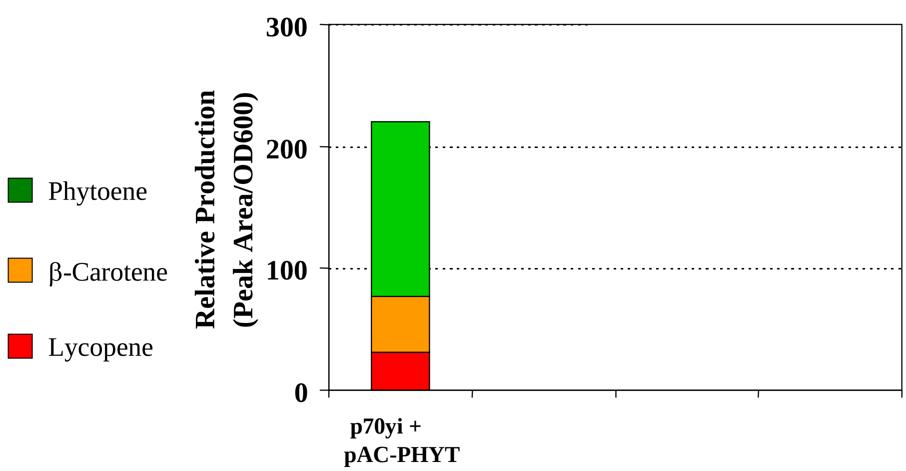
A synthetic operon for carotenoid production

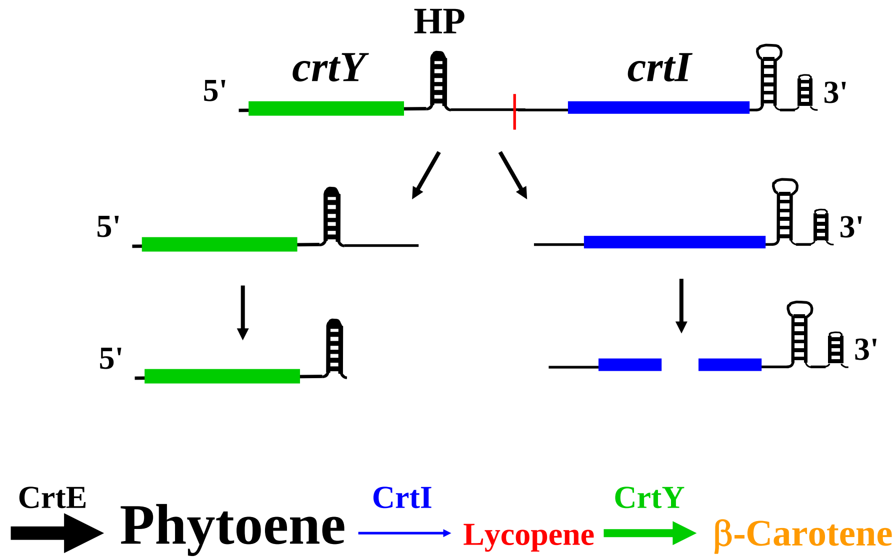


No hairpin 5’ of *crtI*

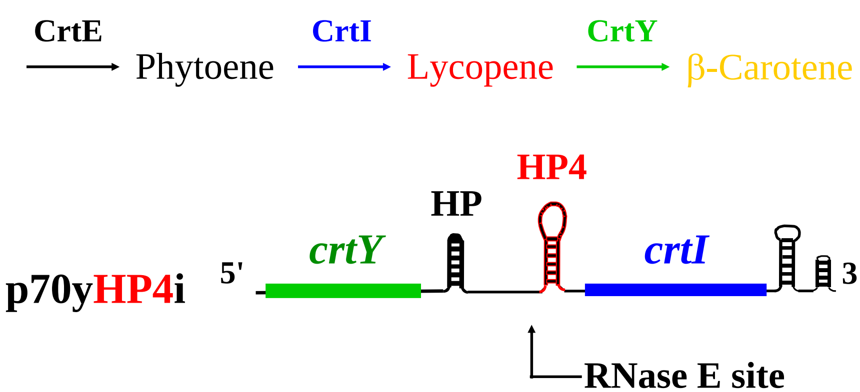


Carotenoid Production

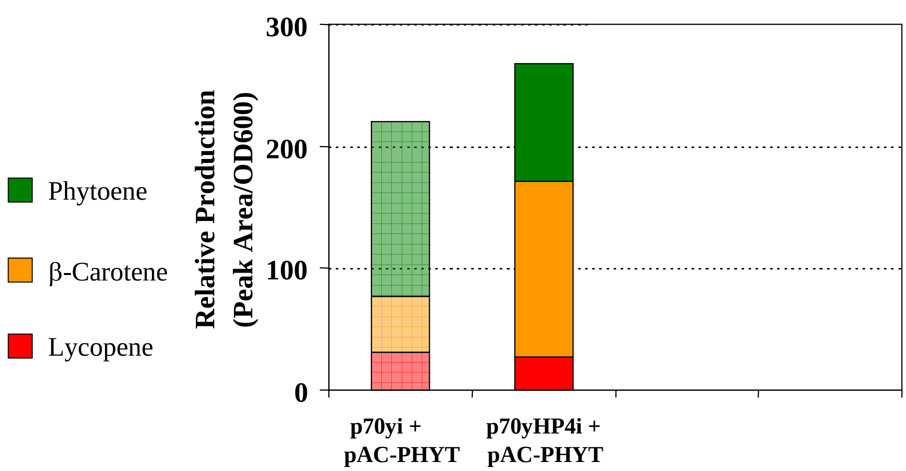


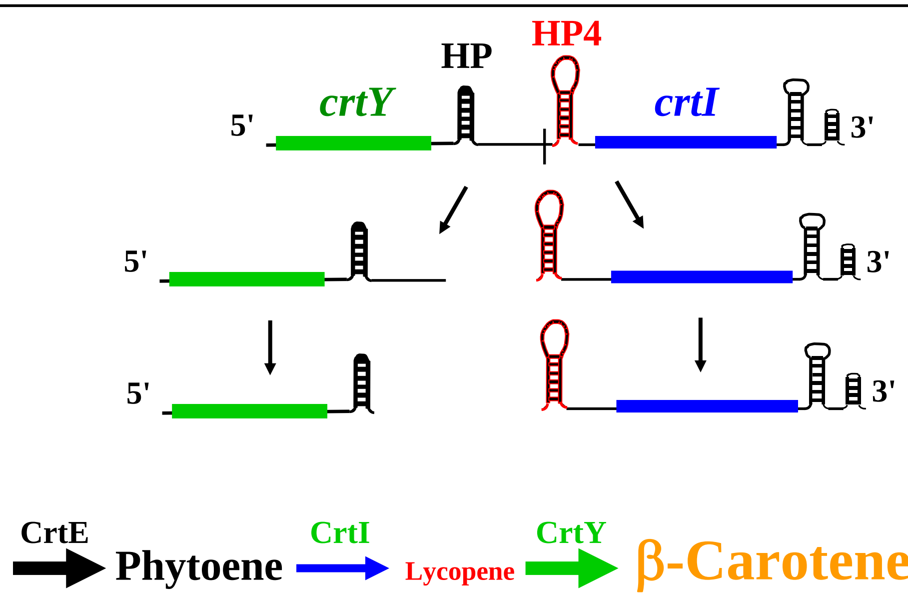


A strong hairpin 5’ of *crtI*

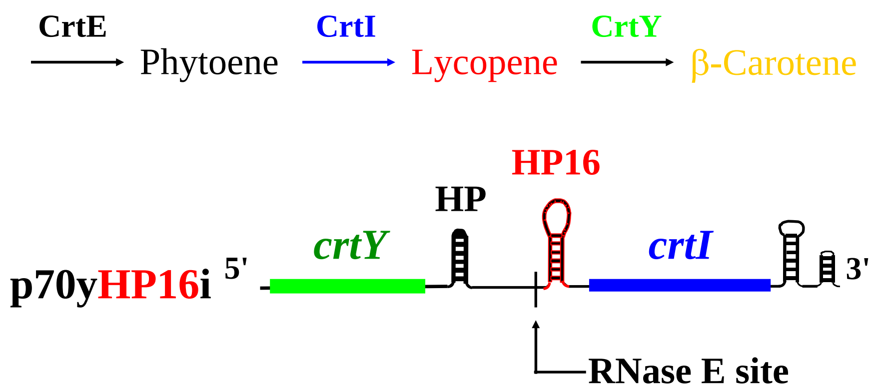


Carotenoid Production

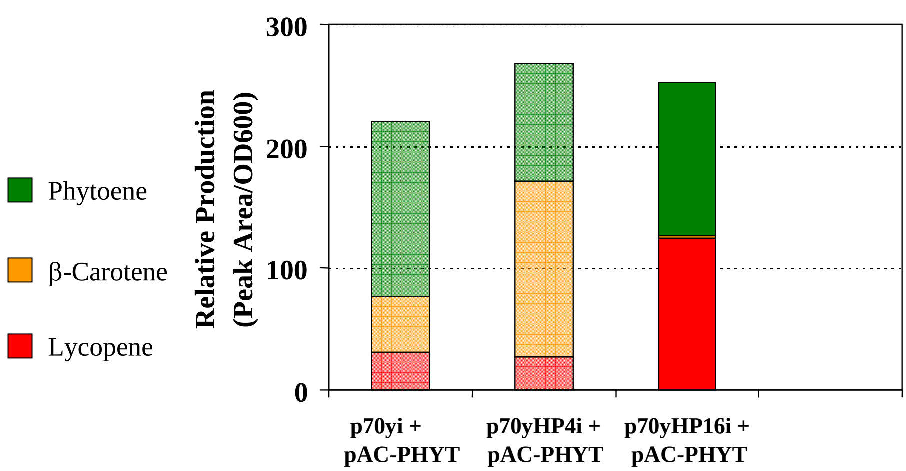


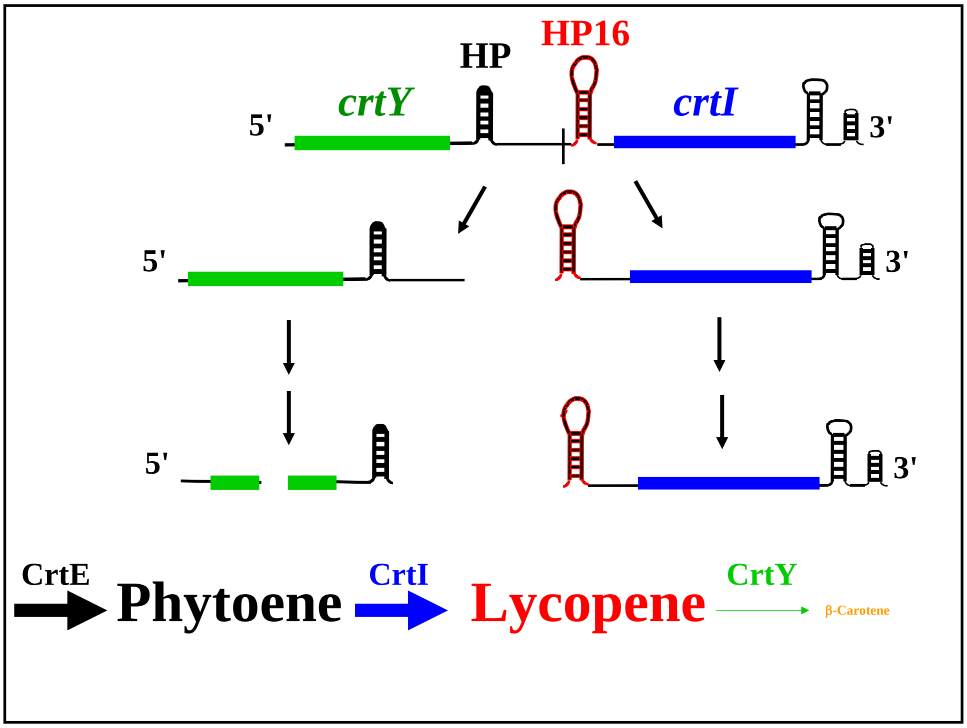


A stronger hairpin 5’ of *crtI*

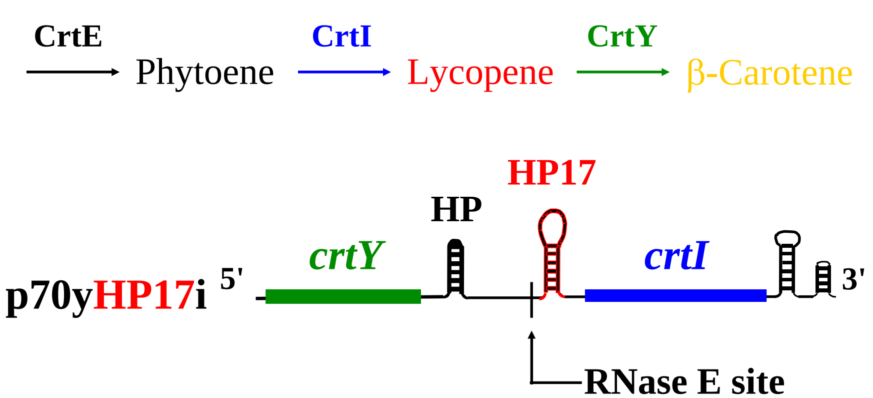


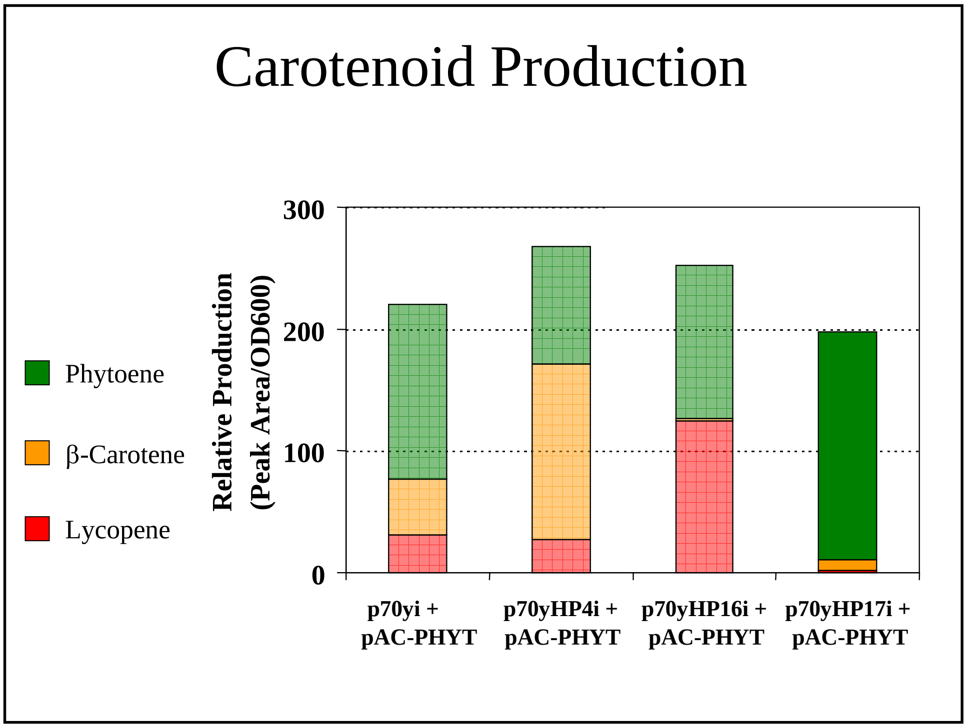
Carotenoid Production

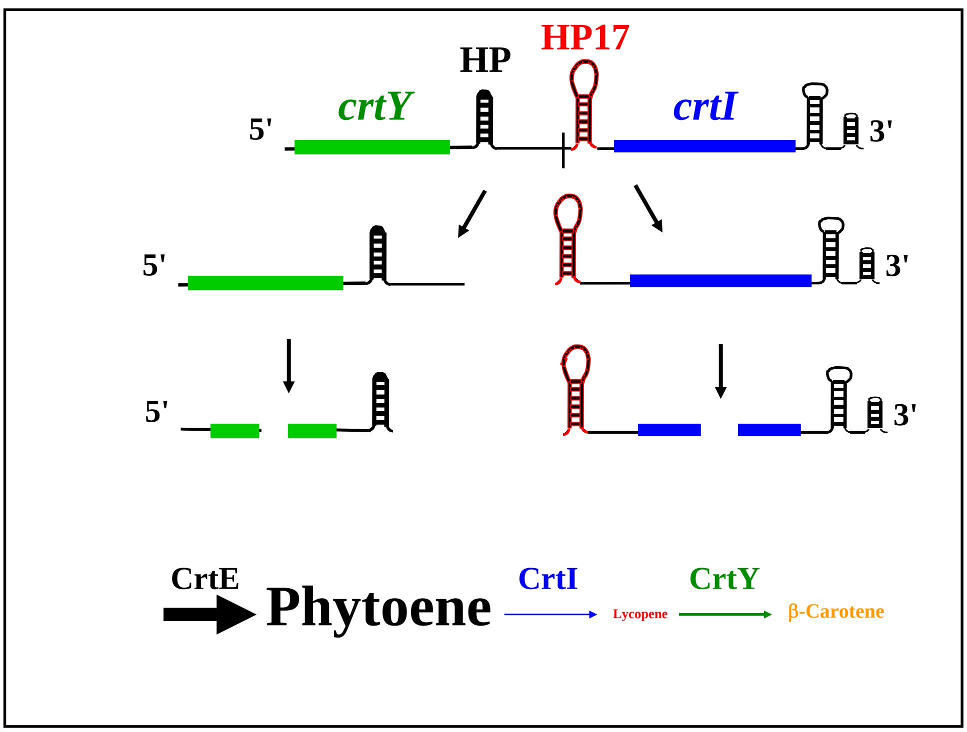


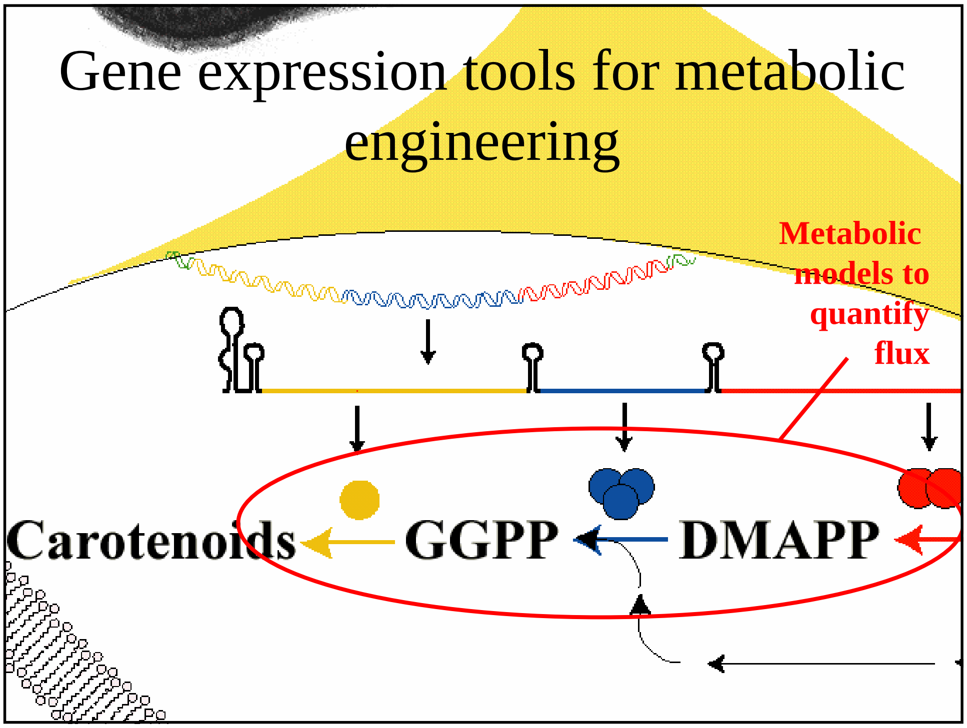


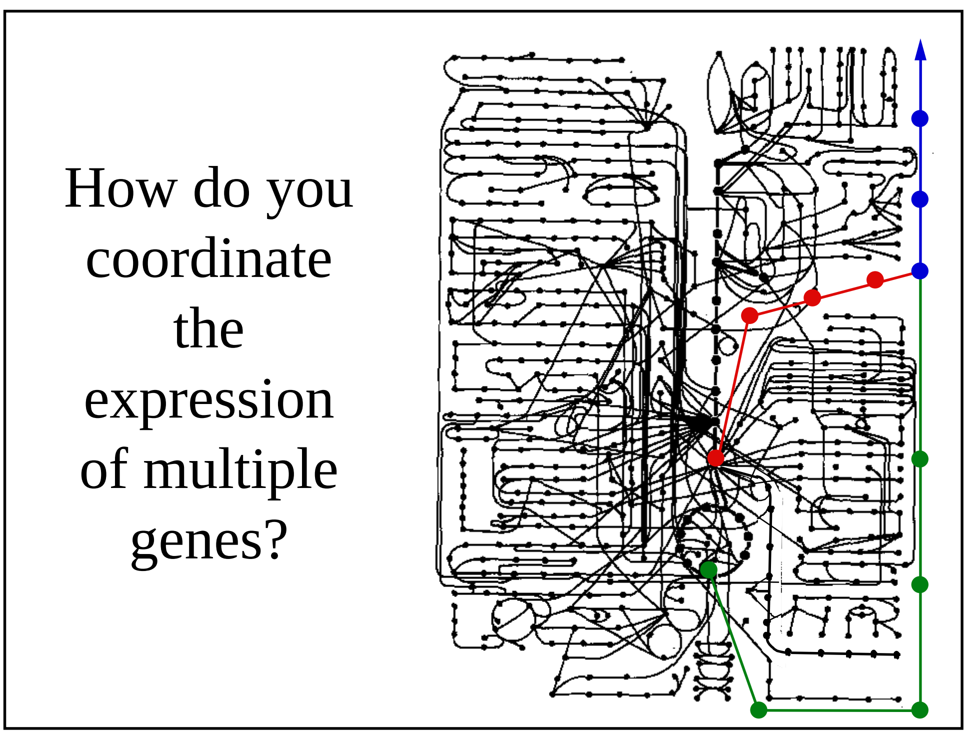
A weak hairpin 5’ of *crtI*




Carotenoid Production
Phytoene
β-Carotene
Lycopene
Relative Production
(Peak Area/OD600)
300
200
100
0
p70yi +
pAC-PHYT
p70yHP4i +
pAC-PHYT
p70yHP16i +
pAC-PHYT
p70yHP17i +
pAC-PHYT


HP17
HP
crtY
crtI
5'
3'
CrtE
Phytoene
CrtI
Lycopene
CrtY
β-Carotene


Gene expression tools for metabolic engineering
Metabolic models to quantify flux
Carotenoids
GGPP
DMAPP


How do you coordinate the expression of multiple genes?

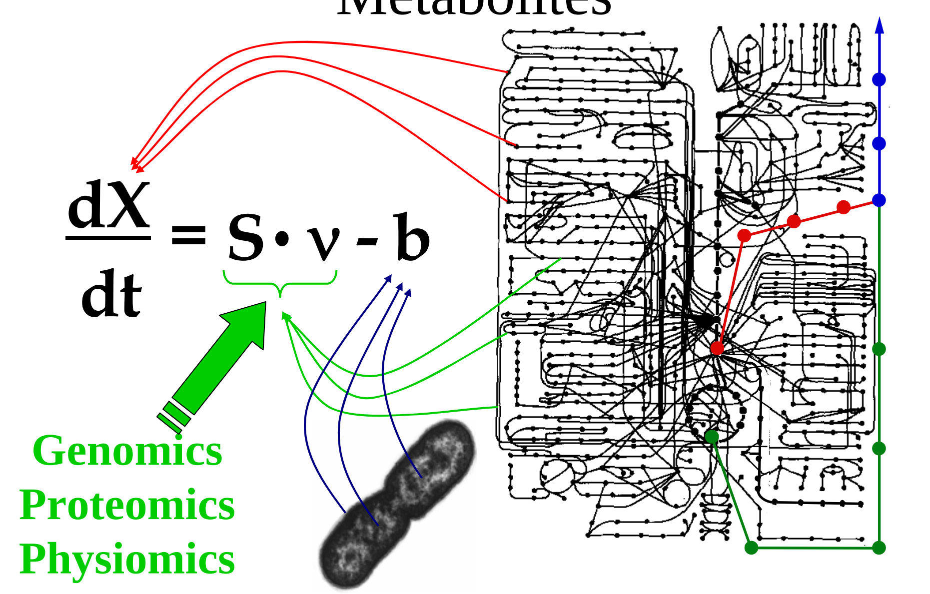
Mass Balance on Cellular Metabolites

$$\frac{d\mathbf{X}}{dt} = \mathbf{S} \cdot \boldsymbol{\nu} - \mathbf{b}$$

where

- **X** = Concentration of metabolites
- **S** = Stoichiometric matrix
  - -> known enzymatic reactions
- **b** = Uptake, secretion, and biomass synthesis
  - -> known from cell composition
- **ν** = Reaction flux vector
  - -> unknown

Mass Balance on Cellular Metabolites



Solve for fluxes using linear optimization

At steady-state: $\mathbf{S} \cdot \nu = \mathbf{b}$

number of fluxes (495) > number of metabolites (289)

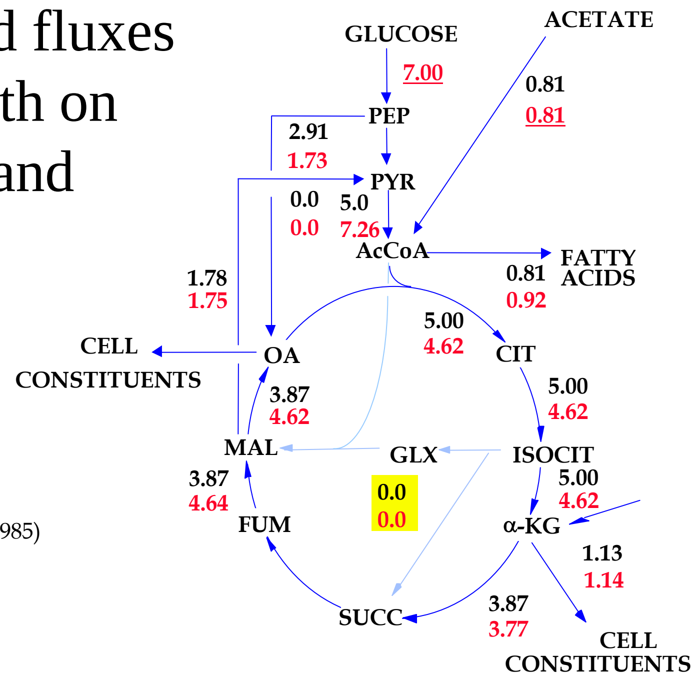
Linear Optimization:

Objective Function: $Z = \sum_i c_i \cdot \nu_i$

Constraints:

1. $\mathbf{S} \cdot \nu = \mathbf{b}$
2. Lower Bound $< \nu_i <$ Upper Bound   i = 1, 2, ...

Predicted fluxes for growth on glucose and acetate



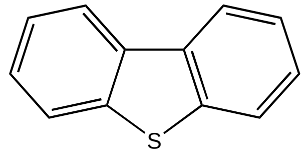
Experimental Data
Walsh & Koshland (1985)

Model Predictions

Some environmental examples of metabolic engineering

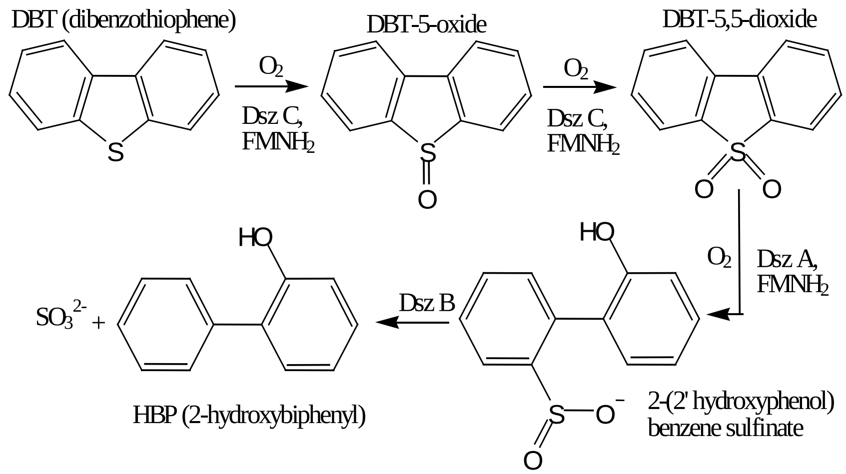
Application of metabolic engineering to biodesulfurization of fossil fuels

- Dibenzothiophene (DBT) is typical of the organic sulfur compounds found in fossil fuels.
- DBT is recalcitrant to hydrodesulfurization.
- Used extensively in biodesulfurization studies

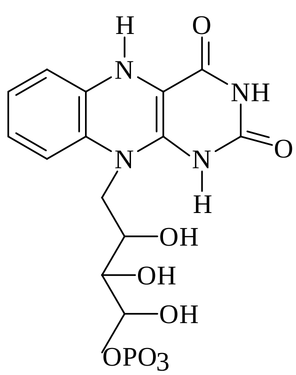


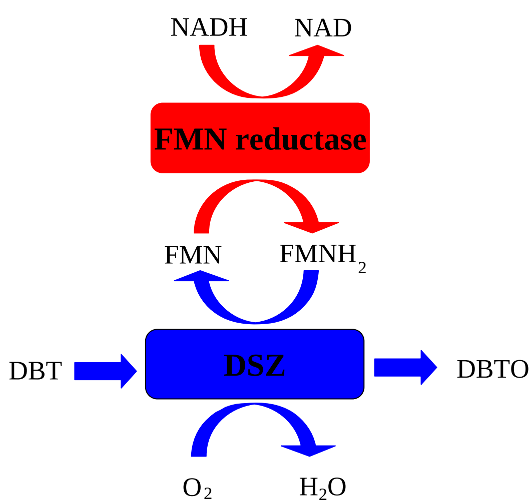
DBT (dibenzothiophene)

Rhodococcus erythropolis IGTS8 desulfurization pathway



Role of reducing equivalents



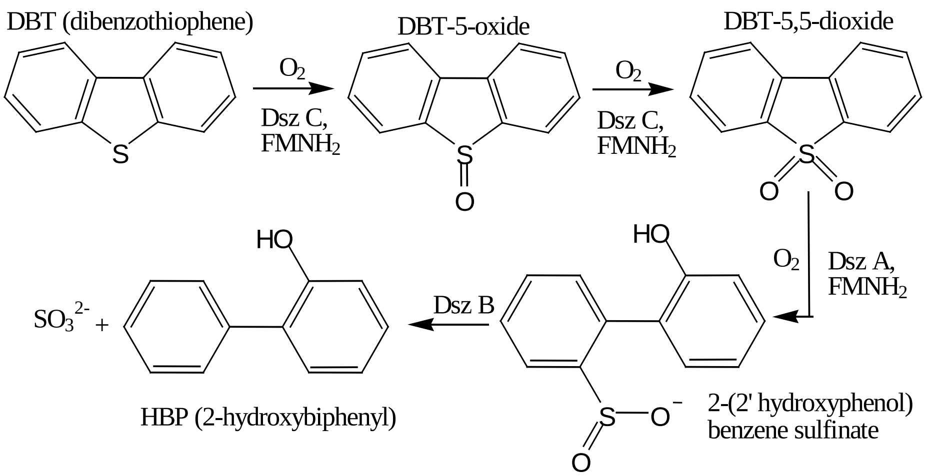


- DszA & DszC require reduced flavin ($FMNH_2$).

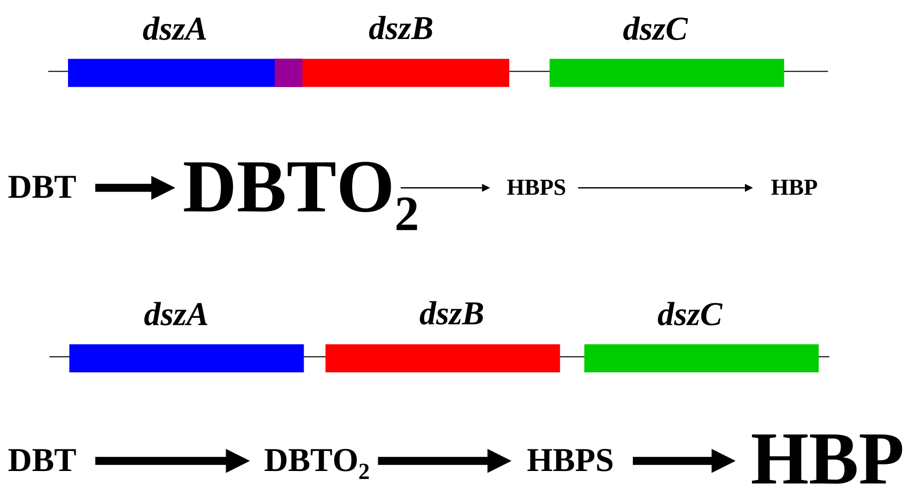
Comparison of degradation rates

Organism	Rate (mg/hr/g dcw)	Source
***E. coli* pRED/pDSZ**	**51**	**This work**
Rhodococcus erythropolis IGTS8	3	Gallardo (1997)
Rhodococcus erythropolis H-2	5	Oshiro (1996)

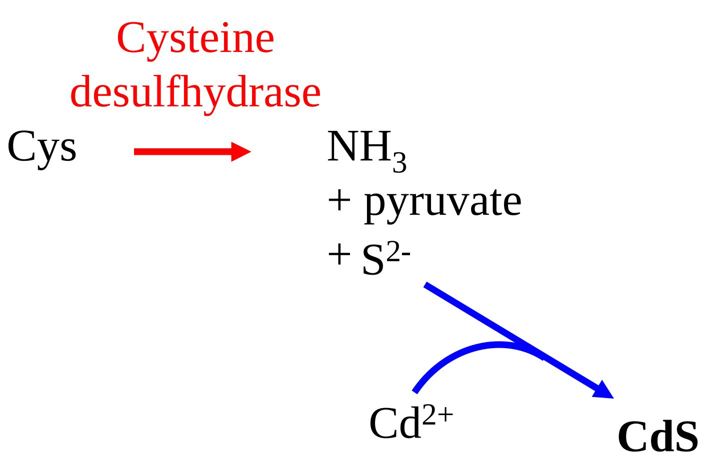
Relieving bottlenecks in the desulfurization pathway



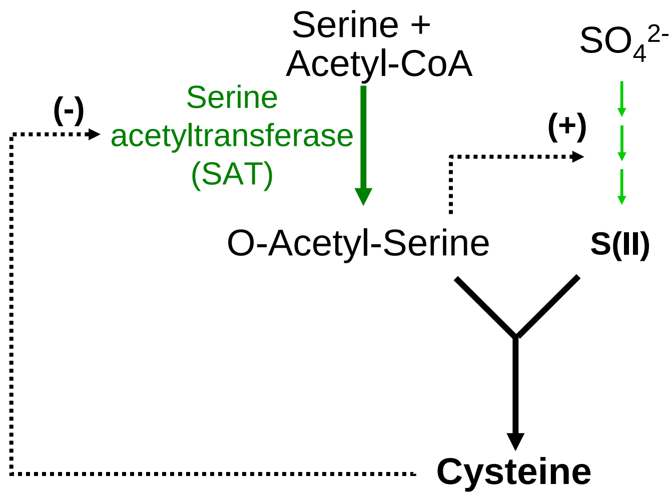
Overlapping reading frames in *dszAB* may limit flux through the pathway



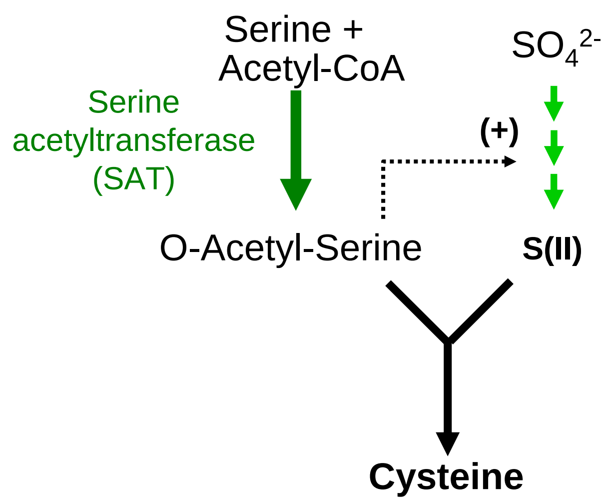
Aerobic precipitation of cadmium



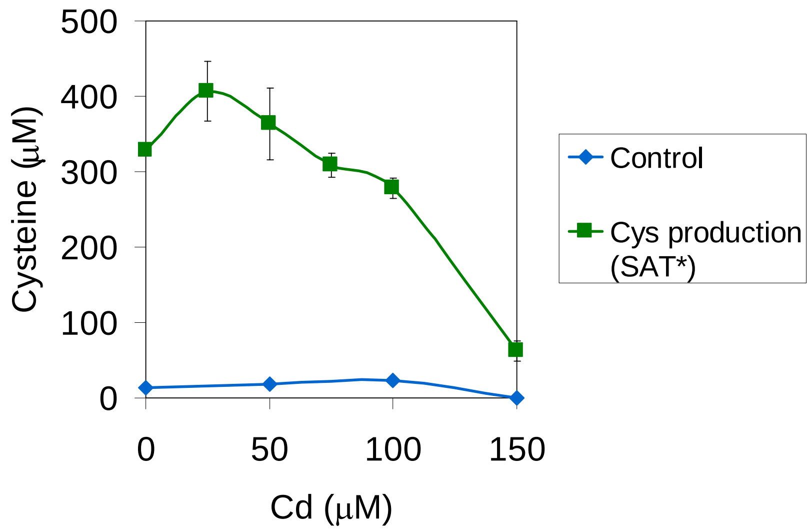
Native cysteine biosynthesis pathway



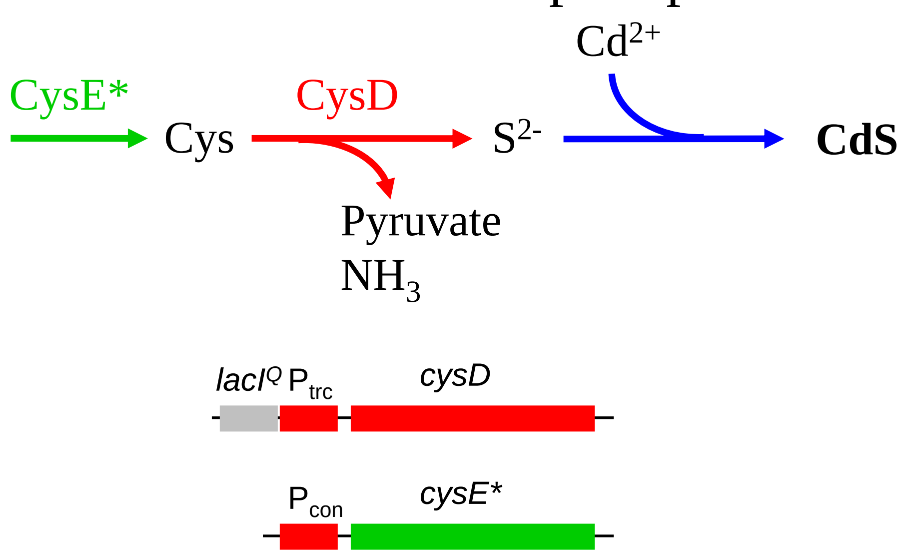
Deregulated cysteine biosynthesis pathway



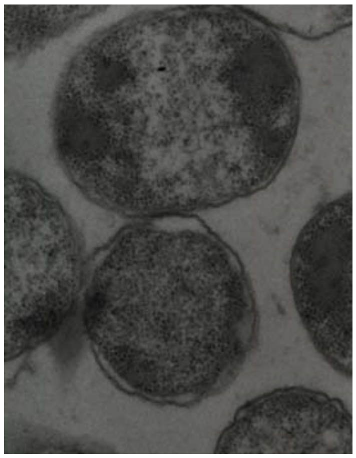
Cysteine production is inhibited by cadmium



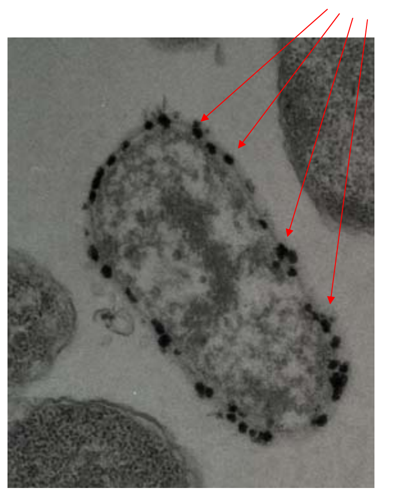
Expression of *cysE** and *cysD* for aerobic cadmium precipitation



Precipitation of CdS by the engineered strain

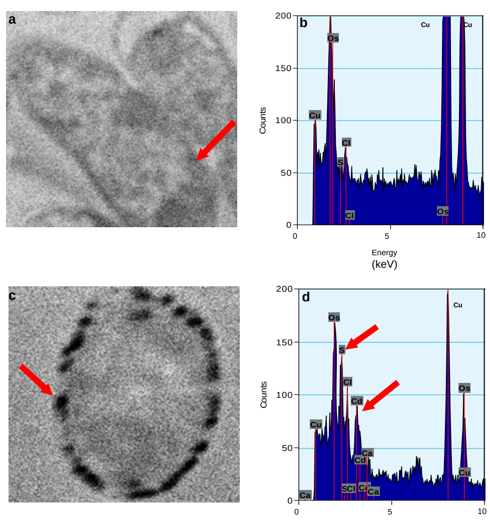


Control



Engineered strain

Cadmium precipitates on the cell wall as cadmium sulfide



Acknowledgements

Graduate Students

Jaya Pramanik
Doug Bolesch
Robert Pape
Natalya Eliashberg
Stephen Van Dien
Salomeh Keyhani
Trent A. Carrier
Kristala Jones
R. Brent Nielsen
Piper Trelstad
Clifford Wang
Ilana Aldor
David Reichmuth
Jessica Hittle
Andrew Walker
Nichole Goeden
Neil Renninger
Christina Smolke
Sundiep Tehara
Stacie Cowan
Cynthia Gong
Michel Maharbiz
David Lubertozzi
Katherine McMahon
Doug Pitera
Kai Wang
Sydnor Withers
Lance Kizer
Brian Frushour
Brian Pfleger

Post-docs

Susan T. Sharfstein
Wubin Pan
Sang-Weon Bang
Eric Gilbert
Artem Khlebnikov
Yet-Pole I
Seon-Won Kim
Vincent Martin
Guangyi Wang

Funding

- National Science Foundation
- Office of Naval Research
- DARPA
- Department of Energy
- National Institutes of Health
- Maxygen
- Merck