Transcriptional regulation

Repression: Hypoxic Genes in Yeast

Rox1p, Tup1p, Ssn6/Cyc8p and Mot3p

Regulation of gene expression

- Almost as important as the genetic repertoire itself
  - The chimp and human gene sequences are almost identical – yet gene expression leads to very distinct results
- Five (six?)regulatory levels:
  - (DNA copy number)
  - **Transcription**
  - mRNA stability
  - Translation
  - Post-translational modifications
  - Protein stability

A yeast model for repression of gene transcription

- The transcription of the yeast ***ANB1*** gene is highly repressed in the presence of oxygen
- *ANB1* codes for the essential eIF-5A protein involved in translation initiation or mRNA export from the nucleus

- In the presence of oxygen, *ANB1* is strongly repressed, and an aerobic counterpart, *TIF51A*, which codes for and almost identical protein, is activated. Yeast needs the eIF-5A protein from one or the other gene to survive

- *ANB1* is closely linked to the yeast oxygen-activated***CYC1*** gene, which codes for the **Iso-1-cytochrome** that is required for **respiration**

Isolation of mutations affecting *ANB1* repression

-Part of the regulatory region between *ANB1* and *CYC1* was inverted

-This manipulation puts *CYC1* under the control of the *ANB1* regulatory region

-*CYC1* is highly repressed; a strain that carries a *cyc1* mutation on its genomic DNA and the plasmid with the inversion cannot grow on non-fermentable carbon sources, because no functional iso-1-cytochrome c is made

-This strain was treated with UV light or EMS (ethylmethane sulfonate, and alkylating agent) to obtain mutants in which repression by oxygen is relieved

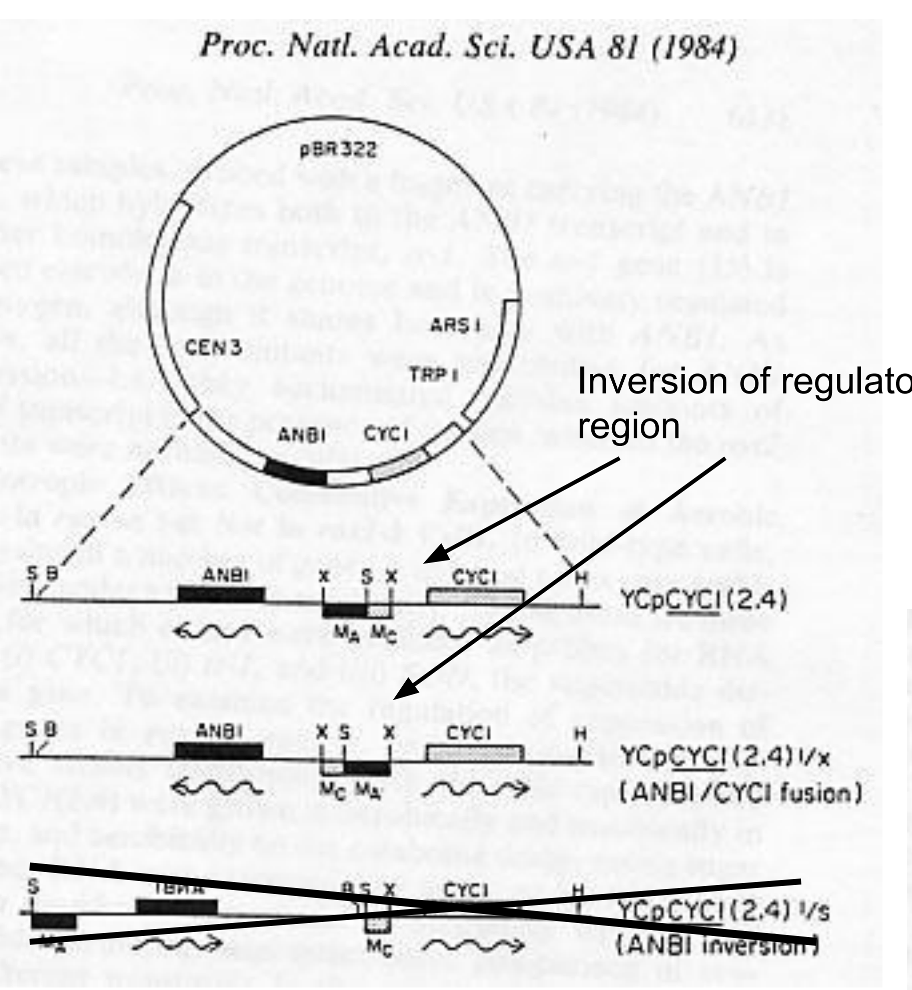


FIG. 1. Plasmids carrying the wild-type *ANB1* and *CYC1* genes, the *ANB1/CYC1* fusion, and the *ANB1* inversion. Construction of plasmids YCp*CYC1*(2.4) and YCp*CYC1*(2.4)1/X has been described (15). Plasmid YCp*CYC1*(2.4)1/S was constructed by digesting YCp*CYC1*(2.4) with *Sma* I, religating, and screening *E. coli* transformants for clones containing plasmid with the 1.6-kb *Sma* I fragment carrying the *ANB1* gene inserted in the reverse orientation. The transcriptional orientation of the two genes is indicated by the directional arrows for the different plasmid constructions. The elements M_A and M_C represent the modulator regions of the two genes, a pair of regions bounded by *Xho* I and *Sma* I sites, which contain regulatory sequences for the *ANB1* and *CYC1* genes, respectively. Restriction sites are indicated by the following abbreviations: S, *Sma* I; X, *Xho* I; B, *Bam*HI.

Characterizing mutations in *ANB1* regulation

- ***cis*-acting** mutations (mutations on the plasmid in the regulatory region) were sorted out by
  A. mating the mutant strain to the parental strain (*cyc1* Δ); *cis*-acting mutations should act dominant (□ diploid should remain respiratory competent), ***trans*-acting** loss-of-function mutants should be recessive (diploid should be unable to respire)
  B. Growing cells on non-selective media (to lose the plasmid; 5-10% loss per generation) and re-transforming the mutant with the original plasmid (mutants in *trans*-acting protein factors should still be mutant□ able to respire)

- Mutants were sorted into complementation groups

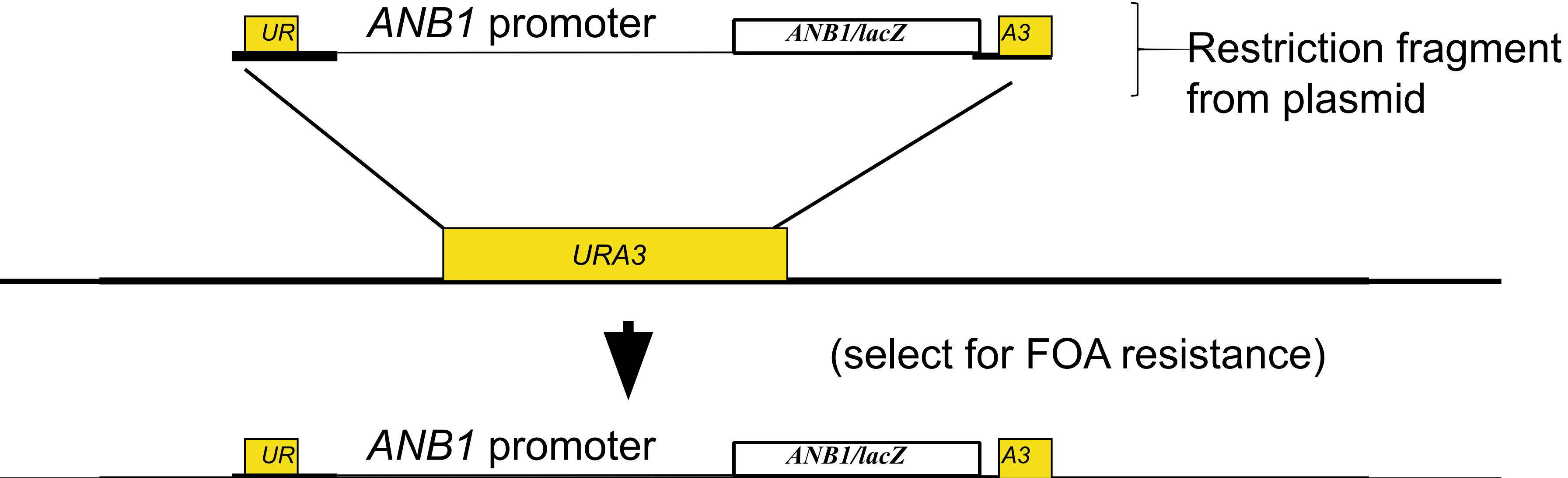
Characterization of the *rox1* mutation

- The initial *rox1* mutant displayed de-repression of the *ANB1* gene, as well as de-repression of several other oxygen repressed genes

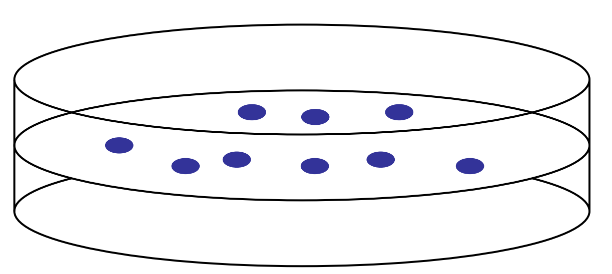
- Genetic analysis indicated the mutation was in one gene

Cloning of the *rox1* mutation

- De-repression of hypoxic genes does not have a detectable phenotype
- Creation of a reporter construct, integration into the *URA3* locus of the *rox1* mutant strain
- The resulting strain is *ura3*- and expresses the lacZ gene product (β-galactosidase) constitutively

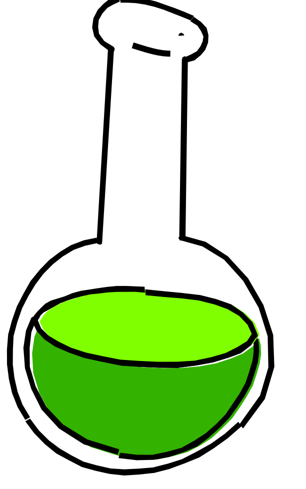


Cloning of *rox1* mutation (2)



rox1 mutant cells with integrated *ANB1-lacZ* fusion on medium containing X-gal all colonies are blue (β-galactosidase expressed)

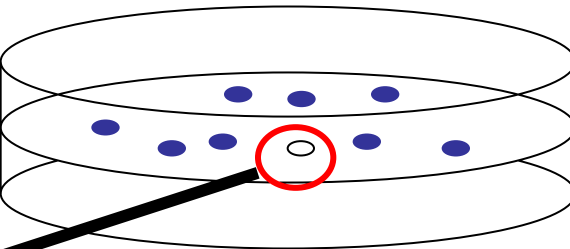
Cloning of *rox1* mutation (3)



Grow *rox1, ura3::ANB1-lacZ* mutant cells

Transform with yeast library (*URA3* marker)

Plate on SC- Ura, X-gal



Screen for white colonies

***lacZ* expression from *ANB1-lacZ* fusion repressed by *ROX1* gene from library plasmid □ no β-galcatosidase activity**

Pick colony, isolate plasmid, retransform to confirm phenotype

sequence

The Rox1 protein is the repressor of hypoxic genes

- Rox1p is a DNA – binding repressor protein with an N-terminal HMG (High Mobility Group) -DNA-binding domain and a rather undefined C-terminal “repression domain”
- The DNA – binding domain has high similarity to the DNA-binding domain of the human *Sry* gene involved in sex-determination and to proteins conferring resistance to the drug *cis*-platin used in cancer therapy
- The DNA – binding domain is roughly L-shaped and introduces 90° bends into DNA

Rox1 binding site consensus:

YYY**<u>ATTGTT</u>**CTC

Rox1p requires Ssn6/Tup1 for repression

- In a similar screen, mutations in the genes for ***ROX4*** and ***ROX5*** were isolated that caused de-repression of hypoxic genes

- Sequence analysis revealed that ***ROX4***=***TUP1*** and ***ROX5***=***SSN6/CYC8***

- Rox1p is dependent on Ssn6/Tup1 for its repression activity and recruits the Ssn6/Tup1 complex to the target promoters of the hypoxic genes

- The Tup1/Ssn6 repression complex consists of one Ssn6p subunit and three or four Tup1 subunits

Tup1/Ssn6 interacts with nucleosomes to form a repressive chromatine structure

A.

Pst I Nco I Mlu I +1 Pst I

DREs TATA

Nuc-3 Nuc-2 Nuc-1 Nuc+1 Nuc+2 Nuc+3

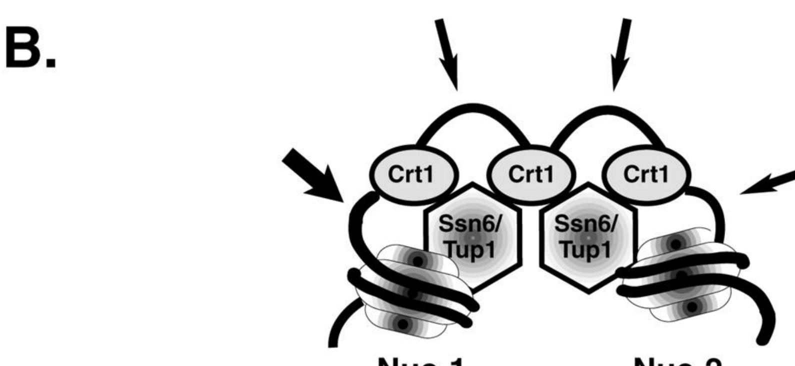
Repressed

Pst I +1 Pst I

DREs TATA

Nuc-3 Nuc-2 Nuc-1 Nuc+1 Nuc+2 Nuc+3

Derepressed



B. Li and J. C. Reese **Ssn6-Tup1 Regulates RNR3 by Positioning Nucleosomes and Affecting the Chromatin Structure at the Upstream Repression Sequence** J. Biol. Chem, September 7, 2001; 276(36): 33788 - 33797.

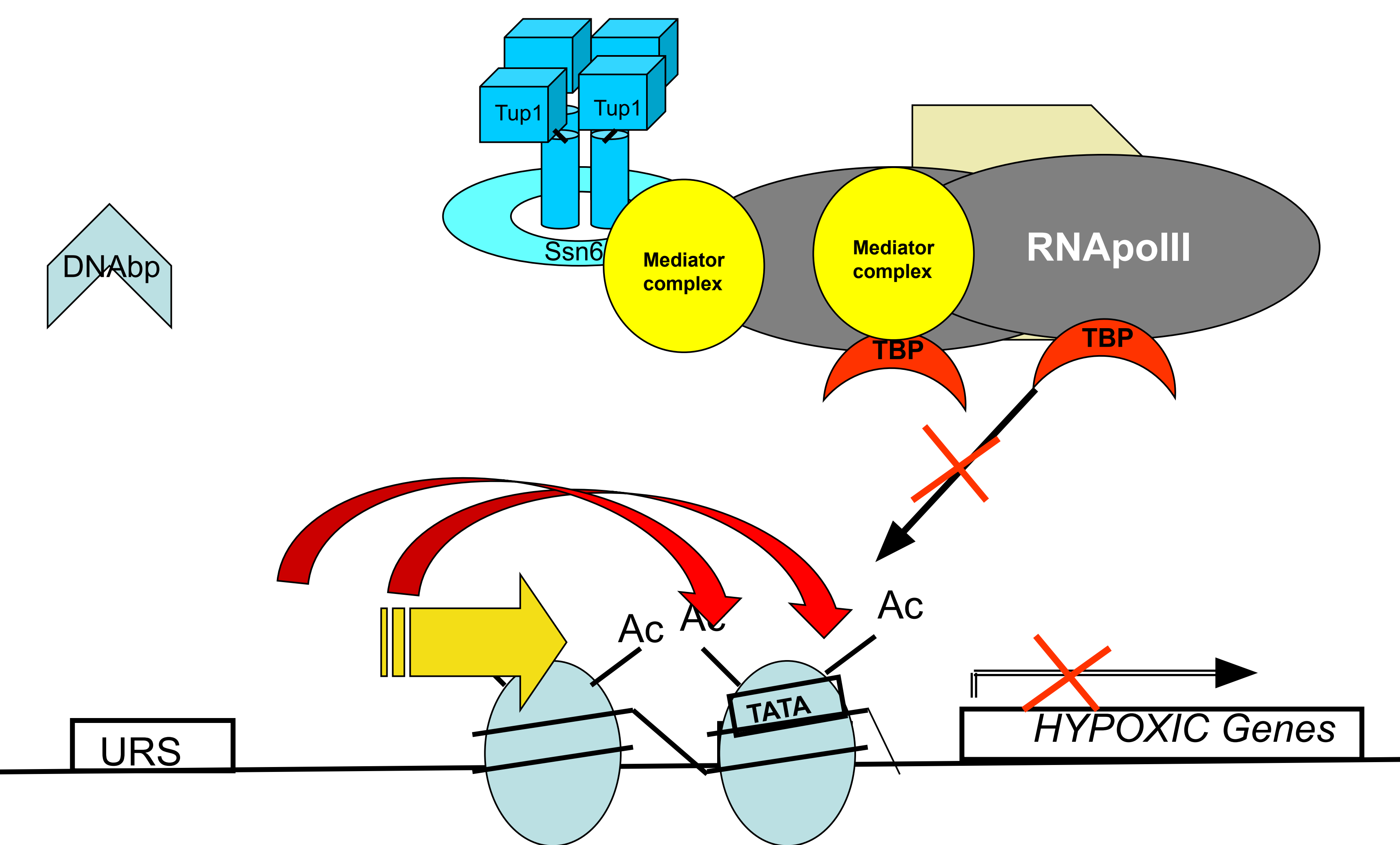
Model of protein and nucleosome interactions at the *RNR3* promoter. *A,* a schematic map of the chromatin organization over the *RNR3* promoter under the repressed and derepressed conditions. *B,* cooperative protein-DNA-nucleosome interactions at the URS. *Arrows* indicate the approximate locations of MNase hypersensitivity detected by high resolution mapping in repressed cells. The *larger arrow* indicates the position of the strongest hypersensitive site. The stoichiometry of Crt1 to the Ssn6-Tup1 complexes is not based upon experimental evidence.

Ssn6/Tup1 recruit HDACs to establish a repressive chromatin structure

- Tup1 has been demonstrated to directly interact with **H**istone-**D**e-**Ac**etylases (HDACs) Rpd3p

- Histone deacetylation causes tighter association of Histones with DNA due to the positive charge of K (Lysine) and R (Arginine) residues in the N-terminal tails of Histones H3 and H4

- Tup1 has also been demonstrated to directly interact with hypo- (under-) acetylated H3 and H4


Tup1
Tup1
Ssn6
DNAbp
Mediator complex
Mediator complex
RNApolII
TBP
TBP
Ac
Ac
Ac
TATA
URS
HYPOXIC Genes

2. Ssn6/Tup1 interacts with the RNA poymerase II mediator complex

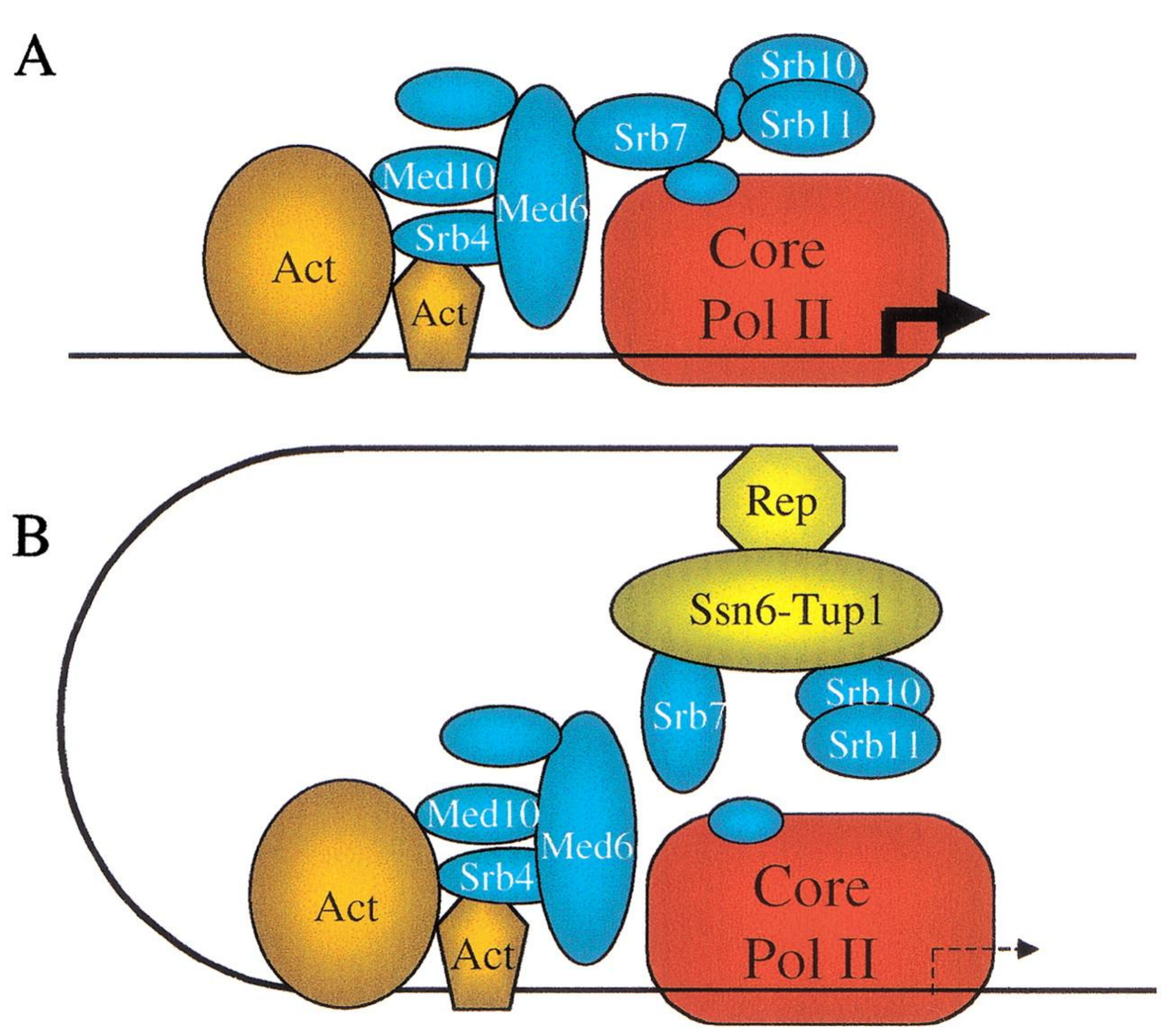
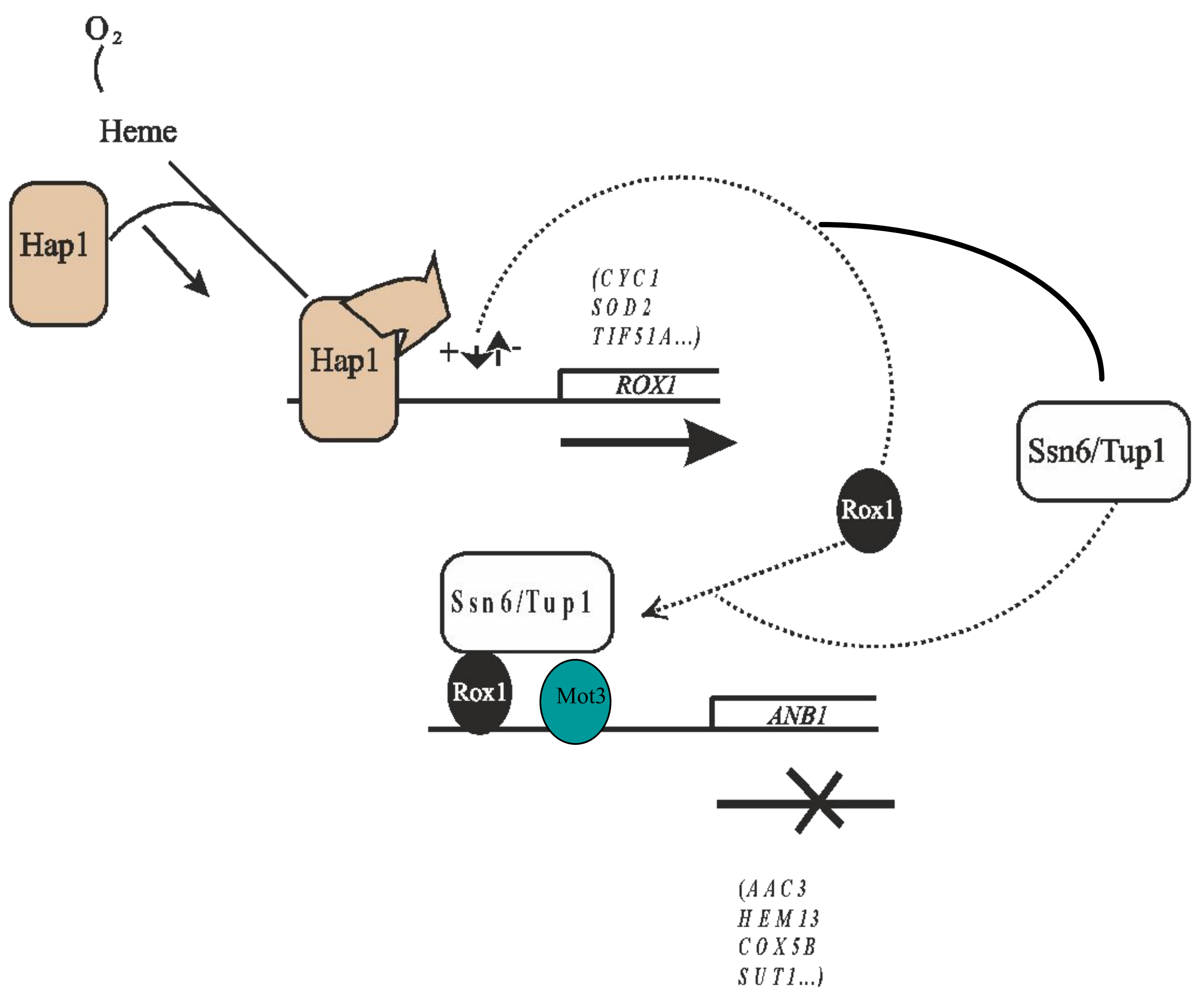


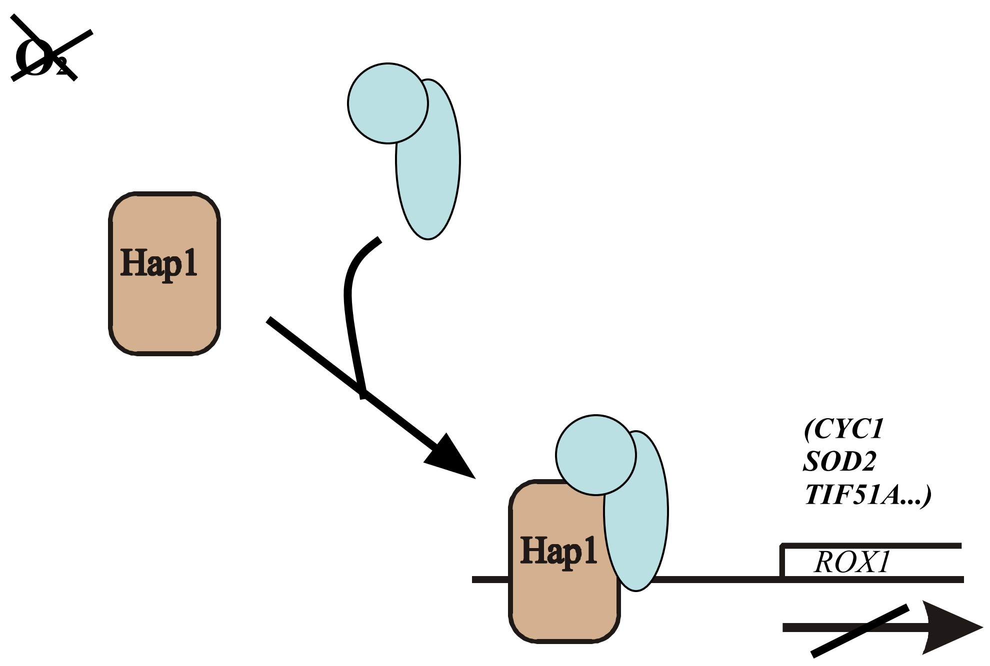
Figure 3. Interactions between Tup1 and the mediator. (*A*) The RNA polymerase II holoenzyme consists of core Pol II and a mediator, which contains multiple subunits, only a few of which are illustrated here (blue ellipses). For simplicity, the general transcription factors have been omitted. A number of activators (Act) require Med6 to activate transcription. These activators may stimulate an interaction between Med6 and Srb7, leading to activation. (*B*) After recruitment by a repressor (Rep), Tup1 (as a component of the Ssn6-Tup1 complex) may block activation by competing with Med6 for binding to Srb7. Tup1 has also been proposed to engage in an inhibitory interaction with Srb10/Srb11.

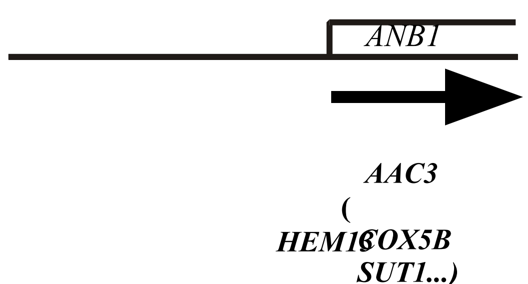
A. J. Courey and S. Jia
Transcriptional repression: the long and the short of it
Genes & Dev., November 1, 2001; 15(21): 2786 - 2796.

Oxygen regulation in yeast



Expression of Hypoxic genes





Promoter analysis

- **What determines the efficiency of repression?**

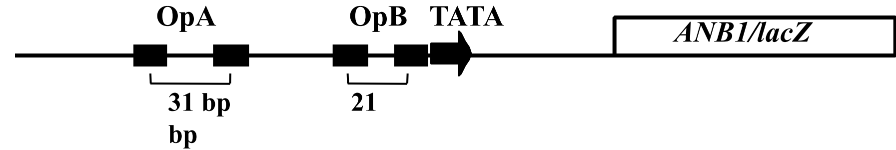
- - Sequence of repressor binding sites
- - Number of operators/ repressor binding sites
- - Position?
- - Modulating factors?

Hypoxic regulatory sequences

Gene	Operator
HEM13	-476 TCAATTGTTTAG -465 -238 TGCTTTGTTCAA -249 -185 CCCATTGTTCTC -174
ERG11	-358 CCTATTGTGCAT -347
CPR1	- 95 TCATTTGTTCCT - 84
HMG2	-282 CGCATTGTTTTG -271 -224 CTTATTGTTCTC -235
SUT1	-243 GTTTTTGTTCCT -232 -342 AGCTTTGTTCTT -331
OLE1	-272 CCTATTGTTACG -261
COX5b	-228 TGTATTGTTCGA -217
CYC7	-333 CCTATTGTATTA -322
AAC3	-197 TTCATTGTTTGG -186 -145 TCCATTGTTCTT -134
ANB1	-316 TCCATTGTTCGT -305 -285 CCTATTGTTCTC -274 -218 TCCATTGTTCTC -207 -197 CTCATTGTTGCT -186
ROX1	-397 CCTATTGTTGCT -386 -364 CGTATTGTCTTG -353
Consensus	**YYYATTGTTCTC**

Organization of the Operators in the *ANB1* Regulatory Region

Rox1 Binding Site Rearrangements



OpA Spacing (in) OpB

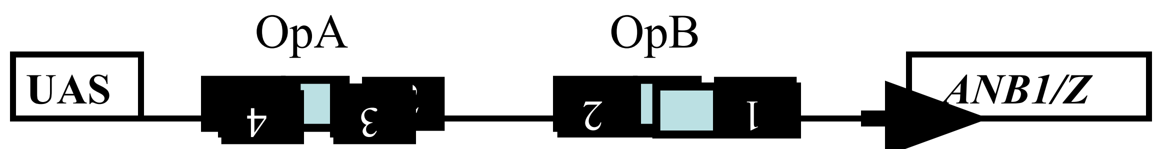
ANB1-Promoter)	$ -galactosidase Activity		Fold Repression () *rox1*/Wildtype)
	Wildtype Strain	) *rox1* Strain	
Wildtype	0.43	114	265
) OpB	1.5	114	76
+5bp OpA	2.2	87	40
-5bp OpA	2.8	83	30
+10bp OpA	2.2	89	41
-10bp OpA	16	130	8.3

Operator Substitutions

OpB+10bp) OpA	31	96	3.1

Role of position for repressor efficiency

Operator construct[1]	β-galactosidase units[2]		Fold repression
	Wild type	*rox1Δ*	
OpA in B position (+)	0.42 ± 0.17	65 ± 26	152
OpA in B position (-)	0.72 ± 0.18	44 ± 7.3	61
OpB in A position (+)	27.5 ± 11	102 ± 32	3.7
OpB in A position (-)	23.3 ± 10.5	80 ± 27	3.4
OpA	1.1 ± 0.6	84 ± 32	76
OpB	9.6 ± 3.8	81 ± 16	8.4



A neglected sequence important for high levels of repression through OpA

conserved sequence adjacent to Rox1 binding sites:

ANB1 OpA : TTC GTTGCCT

ATTGATTCGTT - - TTG GTTGCCT

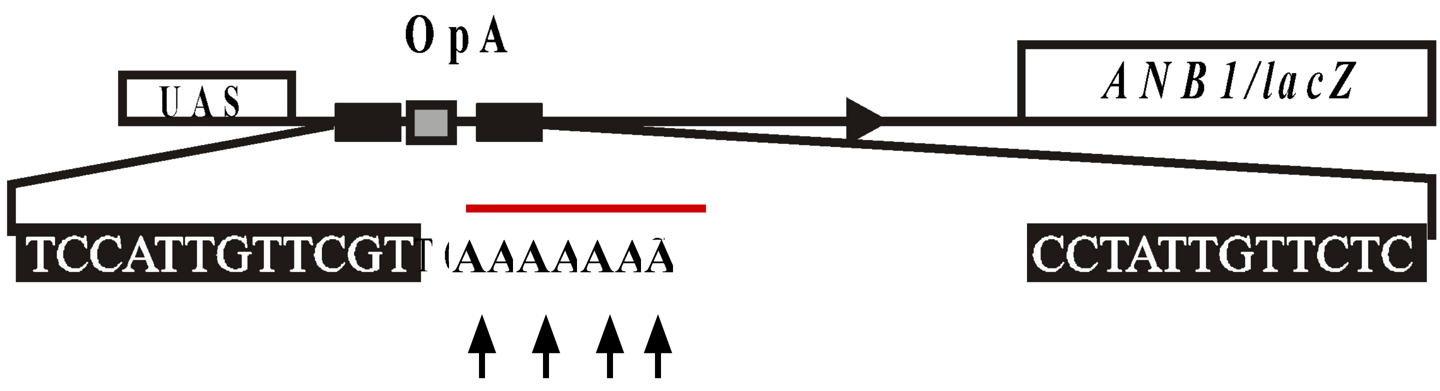
: **ATTCG**

AAC3

COX5B : ATTGTTCGATA AGGCAAC GAA

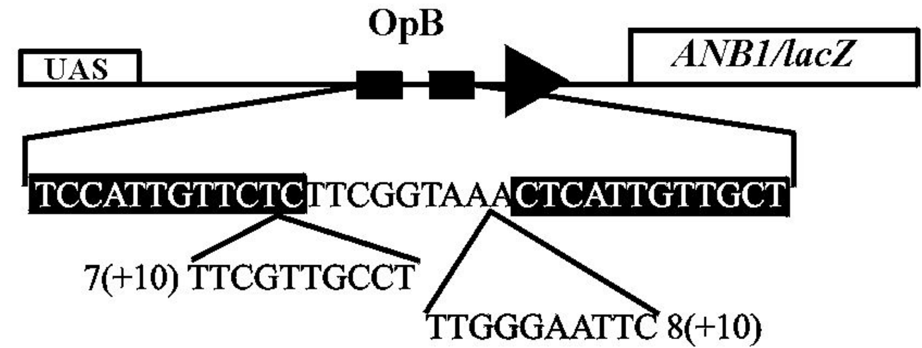
possibly also in other Rox1 regulated promoters

The sequence TGCCT is responsible for stronger repression from OpA



ANB1-Promoter[1]	β-galactosidase Activity[2]		Fold Repression (*rox1Δ*/Wildtype)
	Wildtype	*rox1Δ*	
ΔOpB OpA-wildtype	1.3 ± 0.7	97 ± 34	75
→ OpA 1(-10)	22 ± 5	116 ± 14	5
→ OpA 2(-10)	1.6 ± 0.7	88 ± 12	55
→ OpA 3	1.1 ± 0.2	93 ± 25	84
→ OpA 4	10 ± 3	101 ± 38	10
→ OpA 5	15 ± 4	102 ± 19	7
→ OpA 6	4.4 ± 3	107 ± 31	24
ΔOpA OpB-wildtype	12 ± 8	83 ± 16	7

Insertion of the conserved sequence adjacent to the OpA 5' Rox1 binding site improves repression from OpB



ANB1-Promoter[1]	β-galactosidase Activity[2]		Fold Repression (*rox1Δ*/Wildtype)
	Wildtype	*rox1Δ*	
ΔOpB OpA-wildtype	1.3 ±0.7	97 ±34	75
ΔOpA OpB-wildtype	12 ±8	83 ±16	7
OpB 7(+10)	5.6 ±2	91 ±23	16
OpB 8(+10)	31 ±2	96 ±9	3

MOT3 (**M**odulator **O**f **T**ranscription):

-Mutant derepresses *DAN1* (Delayed Anaerobic) and *ANB1*
Sertil O, Kapoor R, Cohen BD, Abramova N, Lowry CV.Synergistic repression of anaerobic genes by Mot3 and Rox1 in Saccharomyces cerevisiae. Nucleic Acids Res. 2003 Oct 15;31(20):5831-7.

- poorly characterized transcriptional regulator

- deletion with weak mutant phenotype (marginally slower growth; positive and negative effects on transcription)

- DNA binding protein with two C_2H_2 zinc fingers, localized to the nucleus

- binding site is T(A>G)CCT(G>T>A)

site in OpA: TGCCT

Does the Mot3 protein bind OpA?

Electrophoretic mobility shift assay (EMSA)

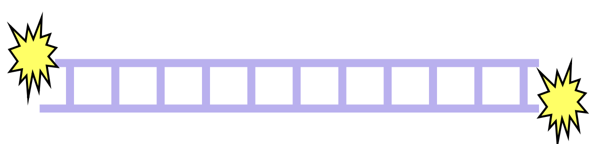
- Used in analysis of DNA binding properties of proteins
- Binding target (DNA or RNA, often a short oligomer containing protein binding sites) is labelled radioactively
- Binding of protein to DNA results in retardation of the migration of the labelled DNA band

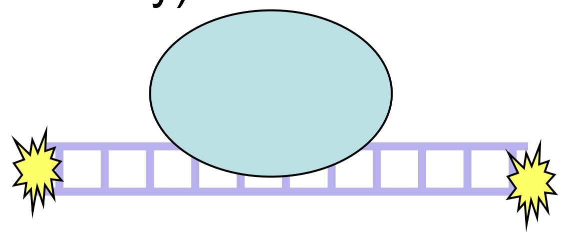
EMSA - Principle

DNA with binding site

DNA – protein complex
(High molecular weight,
bulky)

-



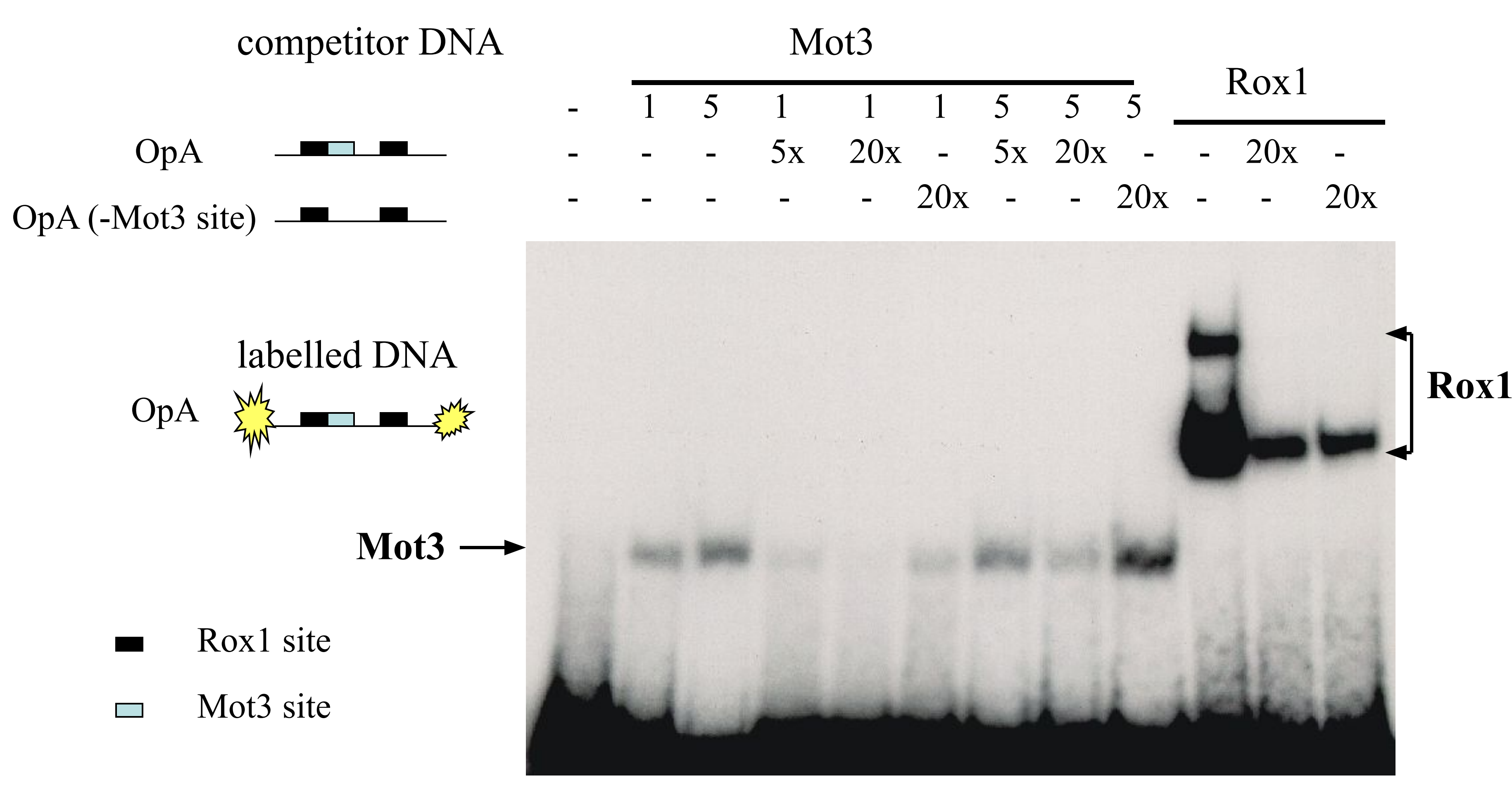


-

+

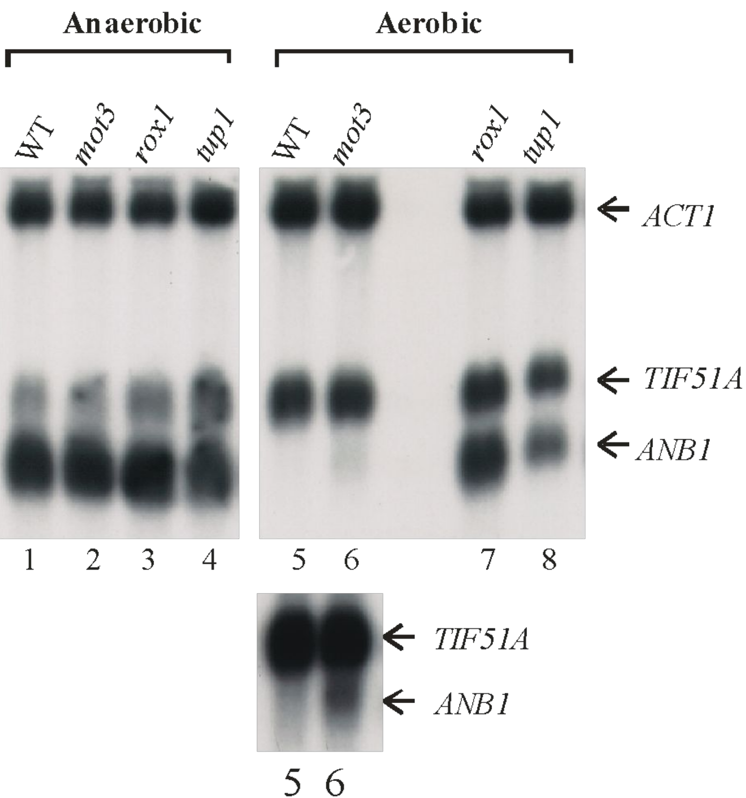
+

The Mot3 protein binds specifically to OpA in the ANB1 promoter



Does Mot3p play a role in *ANB1* repression *in vivo*?

Northern blot probing for *TIF51A/ANB1* transcripts in wild type and mutant strains



A *mot3* deletion causes mild derepression of *ANB1*

How does Mot3p exert its effect on repression?

1. Interaction with Rox1p? (cooperative binding?)

2. Interaction with the Ssn6/Tup1 general repression complex?
    - establishment complex formation?
    - aiding repression function?

Mot3 and Rox1 do not bind DNA cooperatively *in vitro*

25ng Rox1 MBP

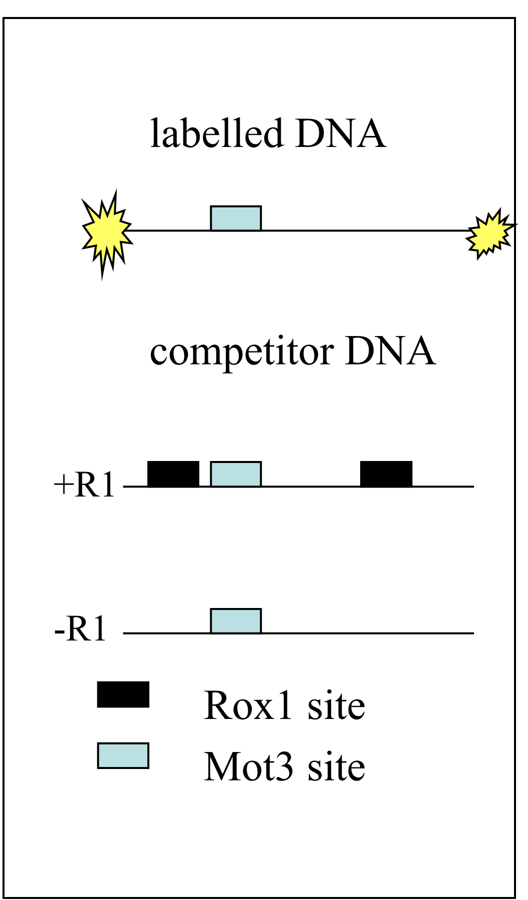
20ng Mot3

Free DNA

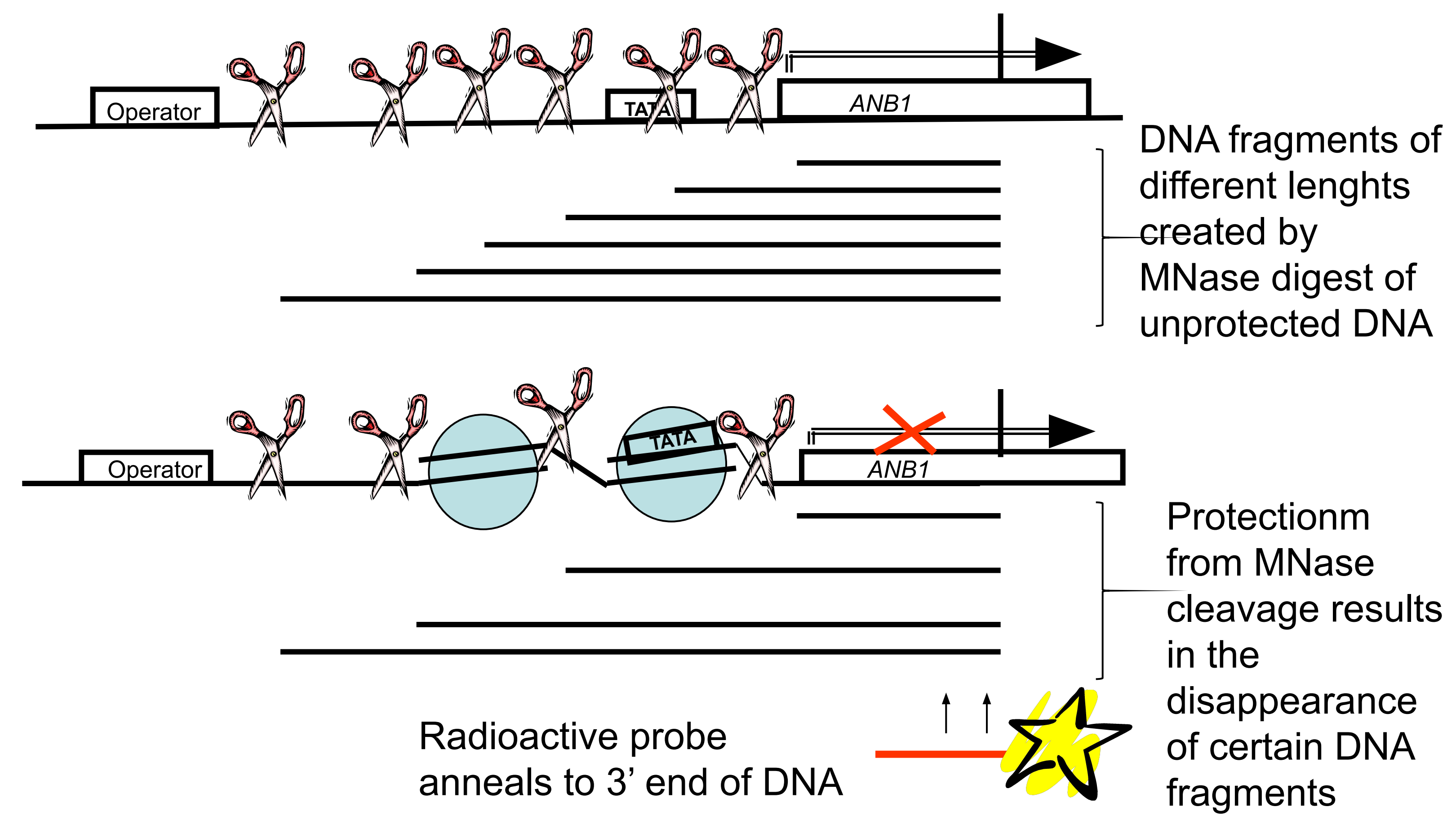
20ng GST-Mot3

+R1

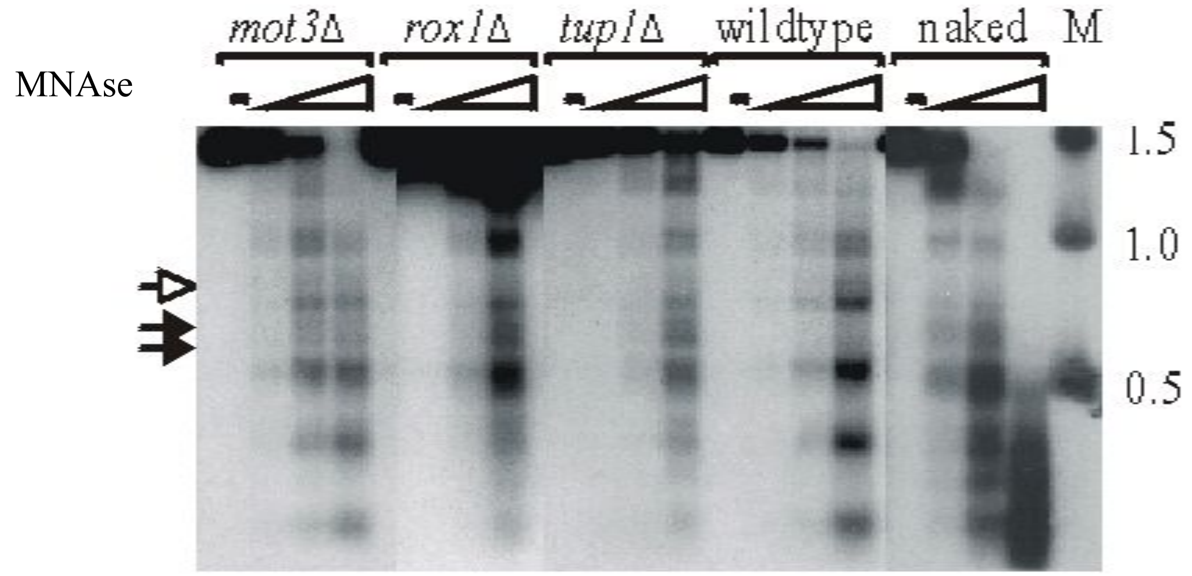
-R1

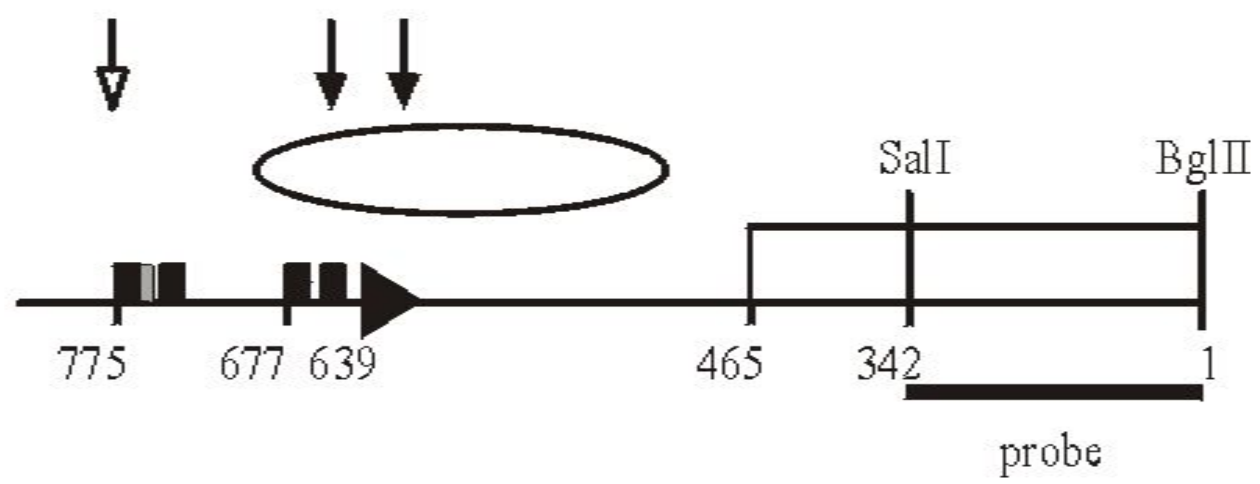


A micrococcal nuclease (MNase) digest reveals chromatin structure of regulatory regions

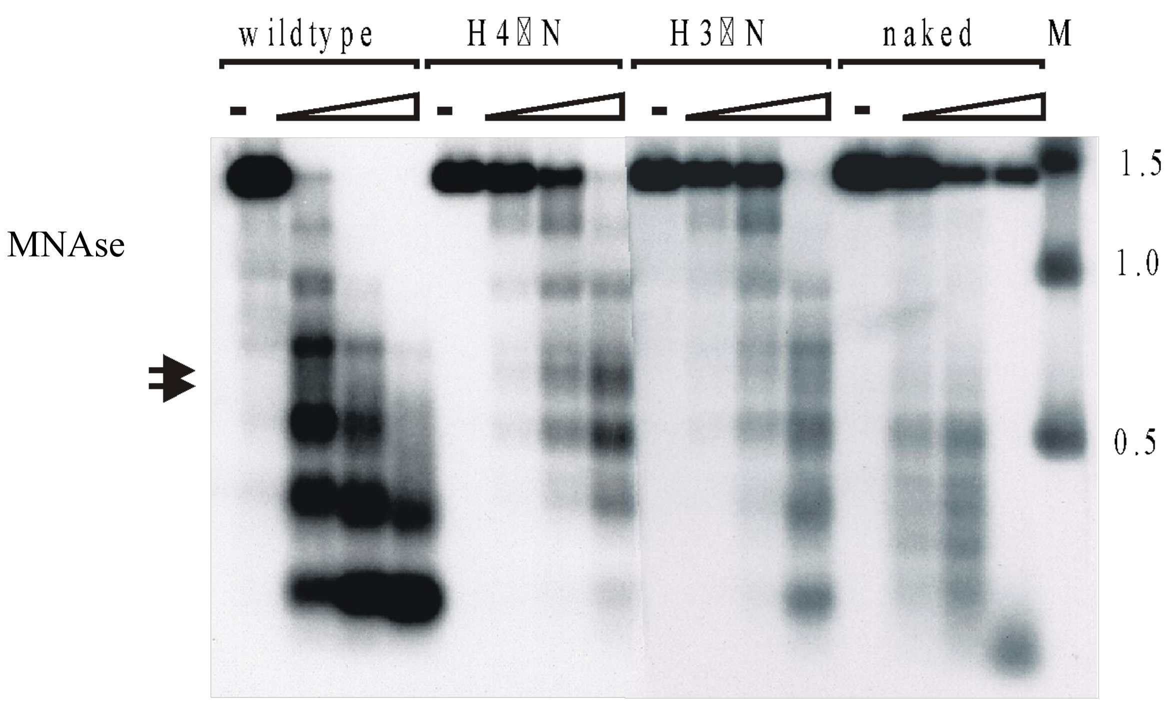


Mot3 affects the chromatin structure of the *ANB1* promoter in a similar manner as Tup1, Ssn6 and Rox1





MCNase generated digestion pattern is dependent on histone N-termini



Summary

Operator efficiency:

- operator orientation and position relative to the TATA box do only play a minor role in operator efficiency

- the sequence TGCCT between OpA binding sites is responsible for higher repression efficiency of OpA compared to OpB

- the TGCCT sequence improves repression from OpB when inserted

- the TGCCT sequence is bound specifically by the transcription factor Mot3, a zinc finger protein protein that has been reported affect the expression of various other genes

- deletion of the *MOT3* gene causes partial derepression of hypoxic genes

A Model Fungal Gene Regulatory Mechanism: The *GAL* genes of *Saccharomyces cerevisiae*

- *GAL* genes: involved in Galactose metabolism
- Early results:
  - *GAL* genes are repressed in the presence of glucose
  - “ “ “ derepressed in presence of other carbon sources
  - most *GAL* genes induced about 1000 fold upon addition of galactose to media (as long as no glucose available)

GAL mutant phenotypes:

GAL1, GAL7, GAL10, MEL1, (*GAL5*): If mutant, cells cannot utilize galactose; a specific enzymatic activity in galactose breakdown pathway missing

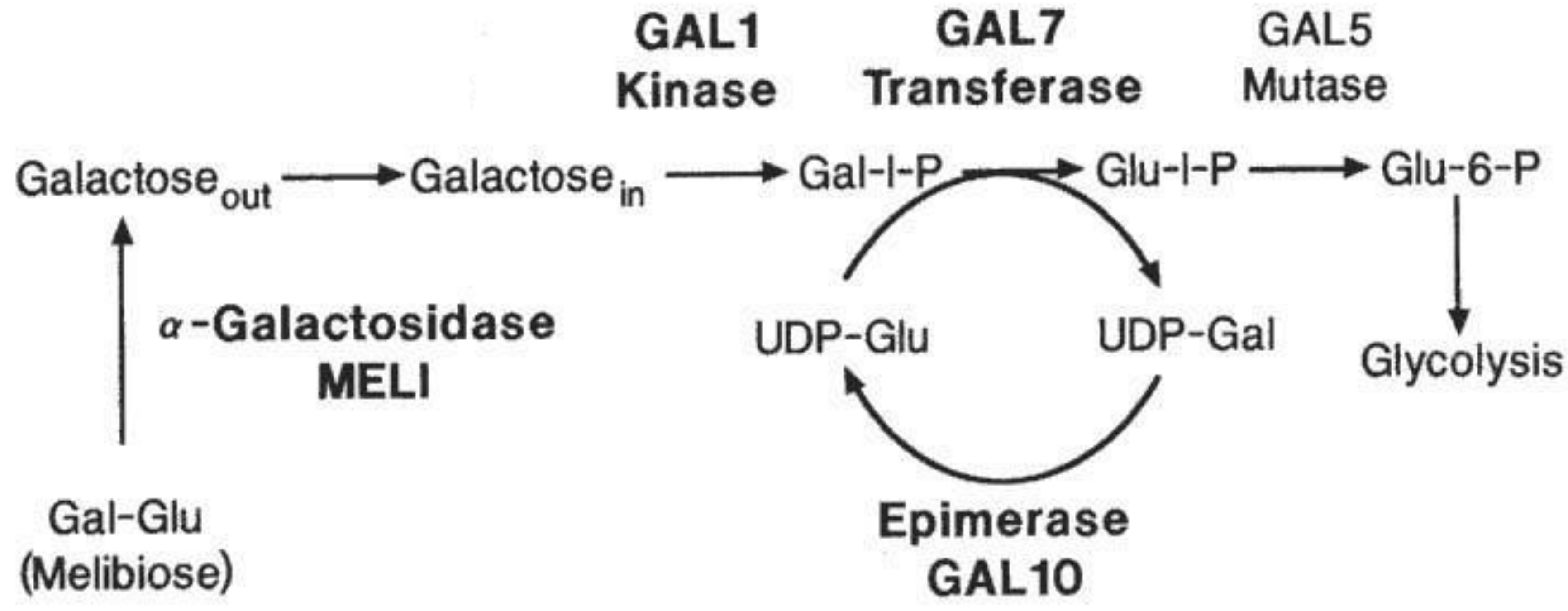
GAL2: Mutant cells cannot utilize galactose, but all enzymatic activities are present in cell extract

GAL4: Mutant cells cannot utilize galactose, none of the enzymatic activities are present in cell extract

GAL3: In combination with mutation in *any one* mutation in *GAL1, GAL7, GAL10, MEL1* (*GAL5*), cells cannot utilize galactose, and all of the enzymatic activities are missing

GAL80: All enzymatic activities are constitutively expressed

The *GAL* structural genes



GAL1, GAL7, GAL10, induced >1000x on galactose

MEL1 induced >100 x on galactose

(*GAL5* ~ 3-4 x)

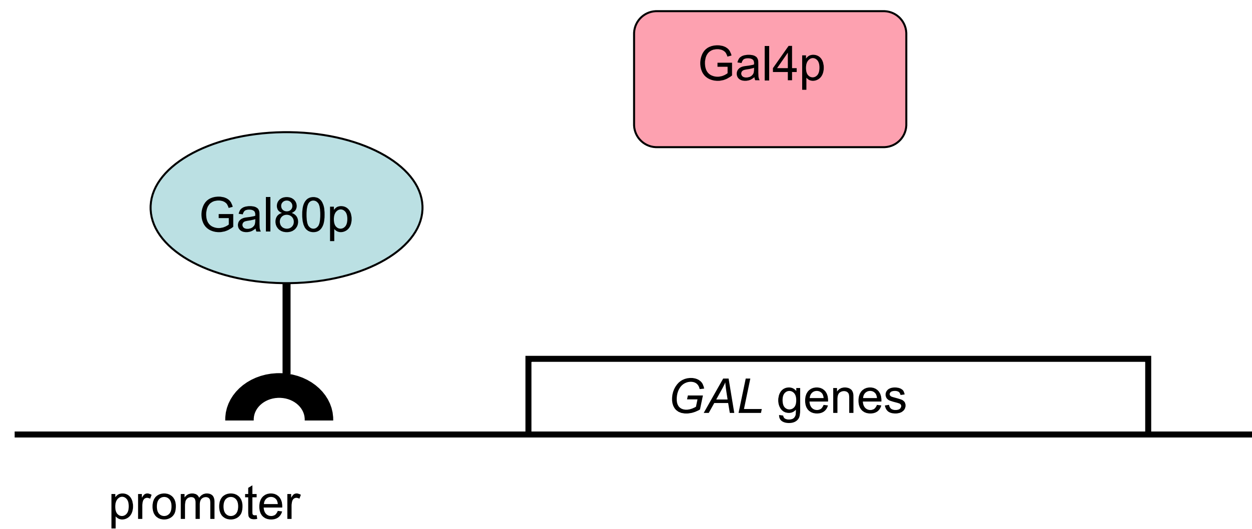
GAL4 and *GAL80* are regulatory proteins

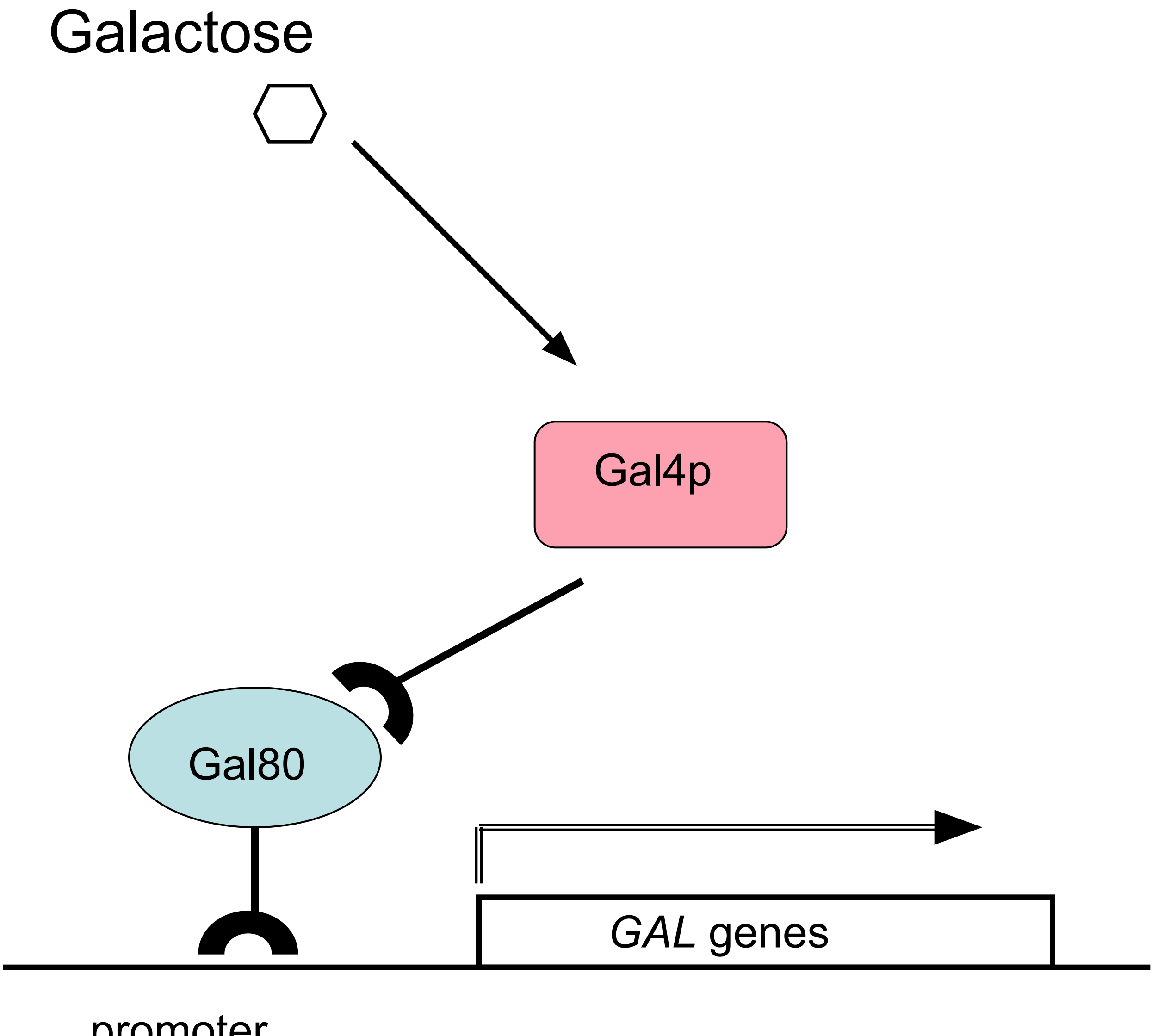
gal4- : uninducible (recessive)

gal80- : constitutive (recessive)

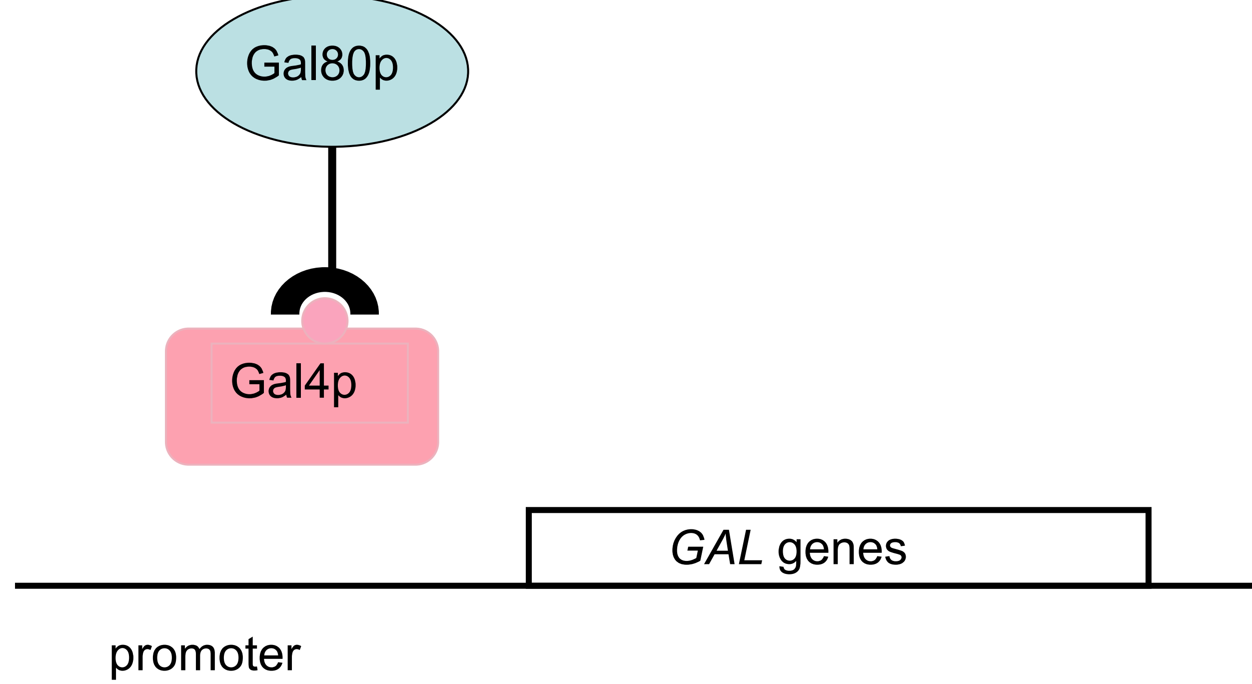
-Two (very simplified!!!) models for mode of action:

1. Gal80p is a repressor of the *GAL* genes; Gal4p inactivates Gal80p in the presence of galactose

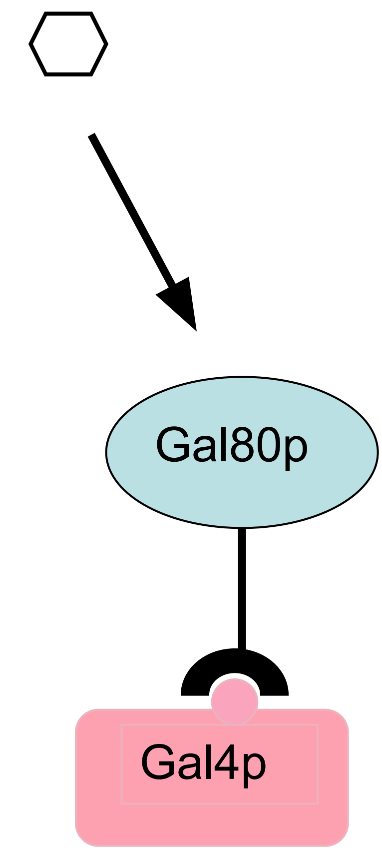



Galactose
Gal4p
Gal80
GAL genes
promoter

2. Gal4p is the activator of the *GAL* genes; Gal80p is a repressor that disables Gal4p activity in the absence of galactose:



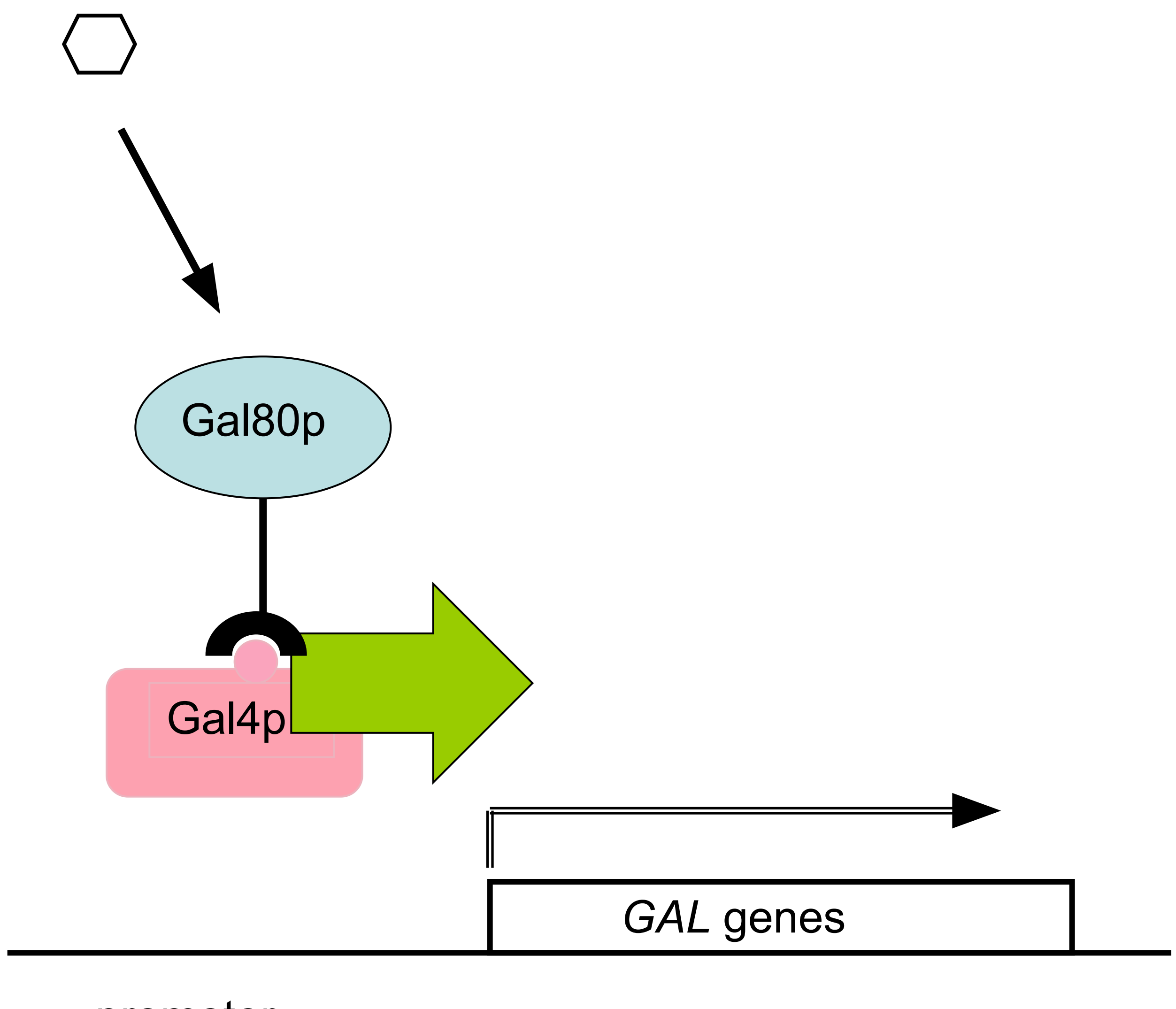
Galactose



GAL genes

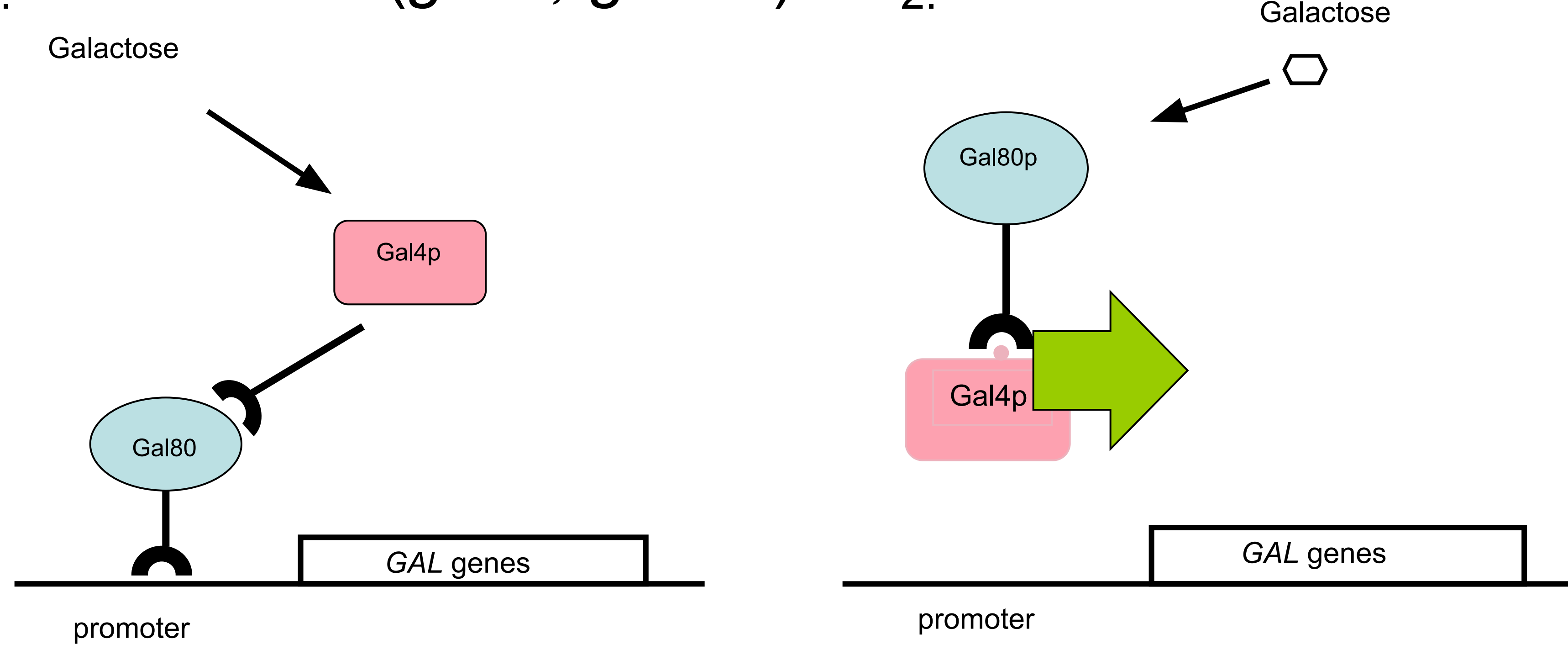
promoter

Galactose



How can we distinguish between the two models?

- Epistasis analysis of pathway
  - What would be the phenotype of the double mutant (*gal4*$^-$, *gal80*$^-$)?



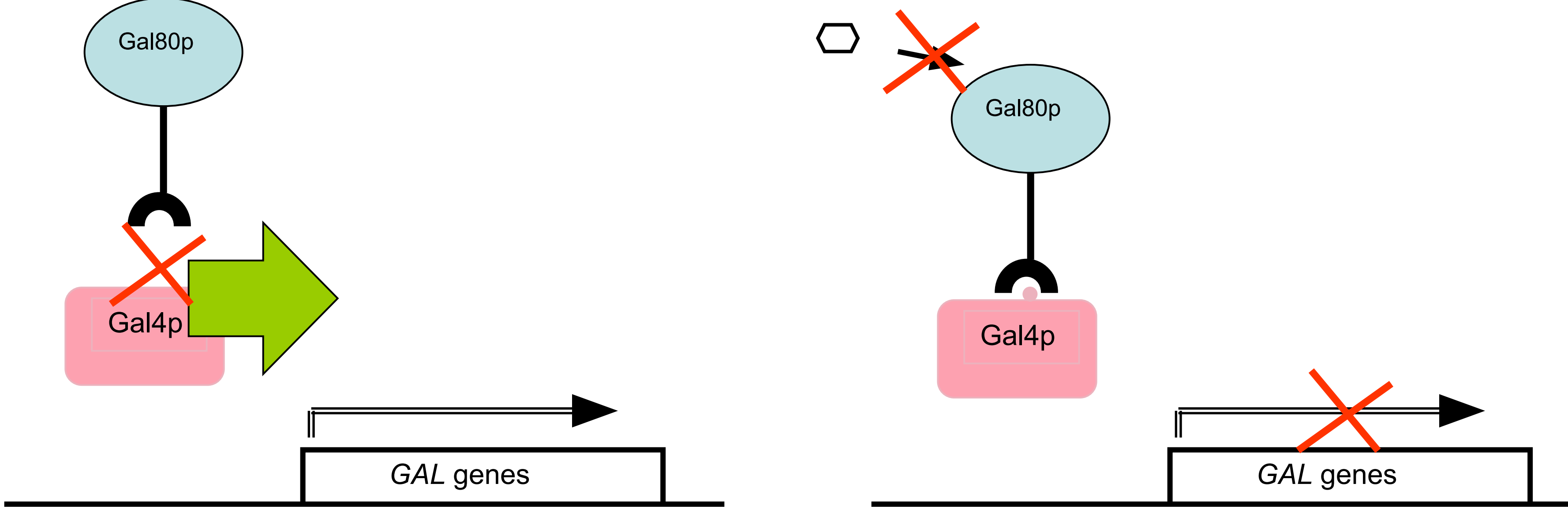
Scenario 2 is correct: the *gal4⁻/gal80⁻* mutant is uninducible

Gal4p is an activator protein, Gal80p inactivates Gal4p,

Recessiveness characteristic for loss of function mutant

GAL4ᶜ mutation: constitutive (cannot interact with *GAL80*) □ dominant mutation

GAL80ᵘ mutation: uninducible; (does not respond to galactose)□ dominant

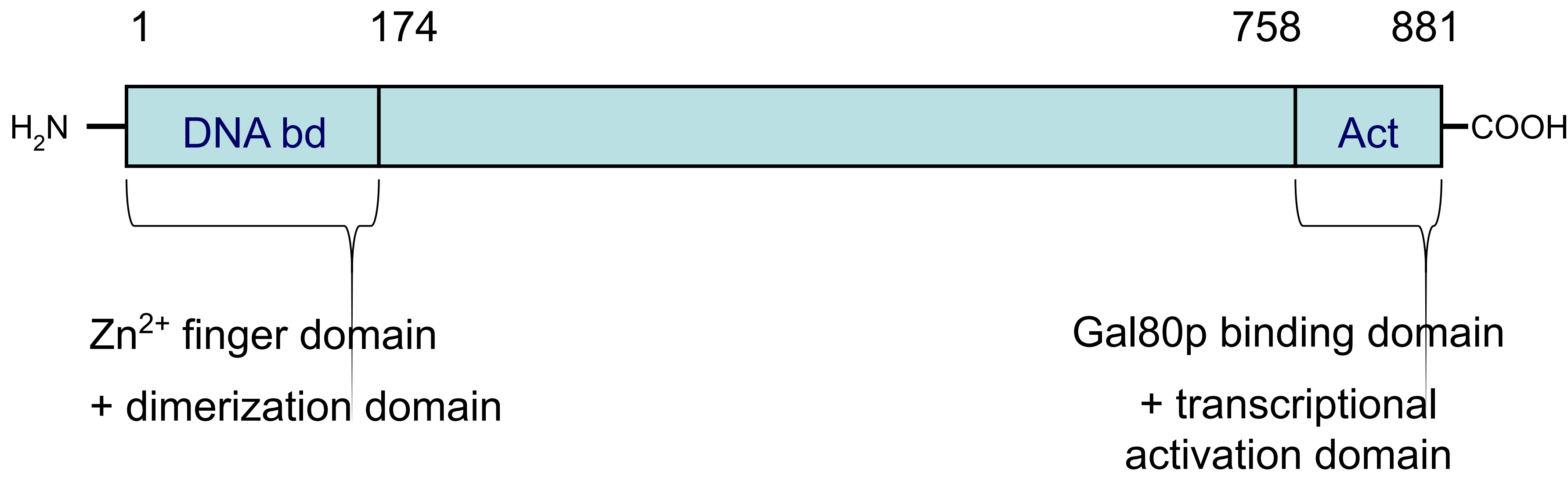


Cloning of the genes

- *gal4*$^-$ uninducible, cannot grow on plates with galactose as the sole carbon source□ transform with genomic library, plate on SCGal or YPGal - survivors should carry library plasmid with wt *GAL4*
- *gal80*$^-$ constitutive: use of inhibitor 2-deoxygalactose (kills cells that are able to metabolize galactose) □ transform cells on media with inhibitor (+ other carbon source) and select for survivors

The Gal4p Activator

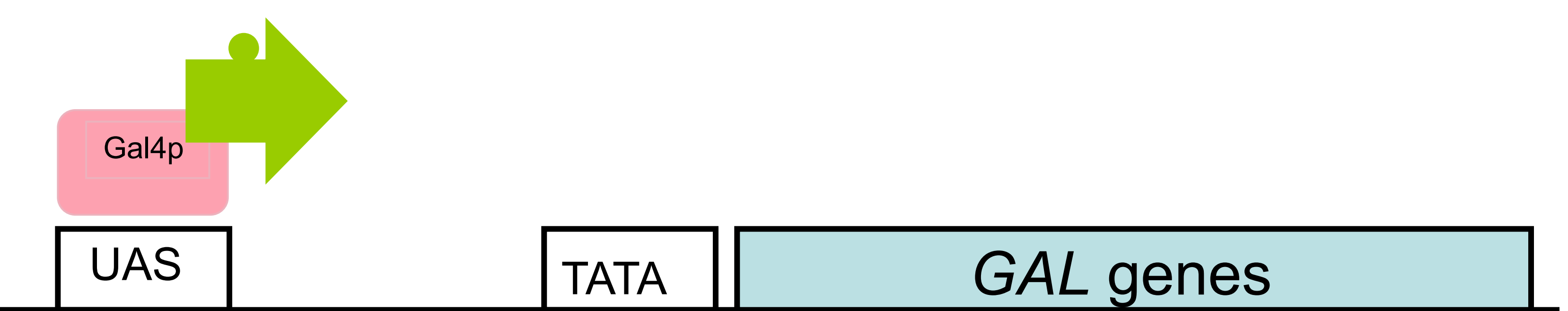
- The Gal4 protein is a DNA - binding transcriptional activator protein and binds as a dimer (Ptashne Group, Harvard late 1980s/early 1990s)



Gal4p binds UAS sequences in the regulatory region of *GAL* structural genes

UAS: upstream activation sequence

TATA – box: AT-rich sequence required for transcription machinery assembly



Deletion analysis of promoter region identified Gal4p binding sites

UAS TATA *lacZ*

"Gal4 17-mer"

Gal4p binding site:

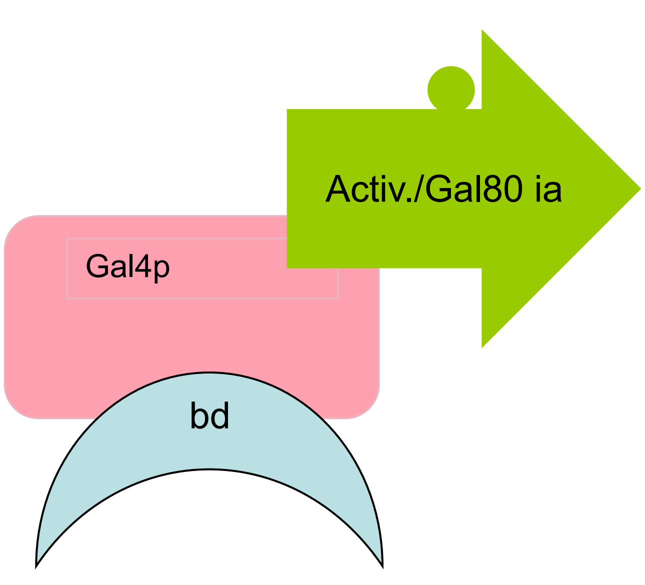
5' –CGGA$^{G}/_{C}$GACA-3'

3'TCAG$^{G}/_{C}$AGGC-5'

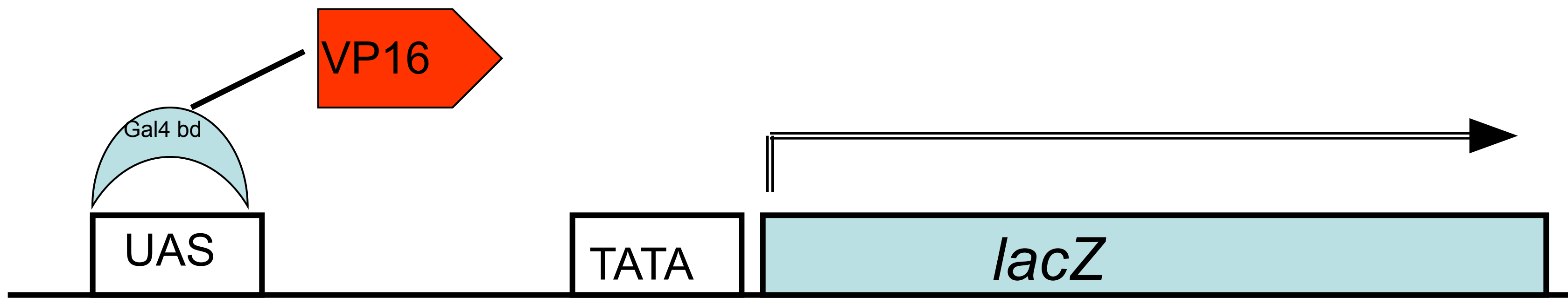
-Site is promiscuous (can function if front of many genes

-Orientation & position independent (symmetrical site, wide range of upstream region from where it can exert transactivation)

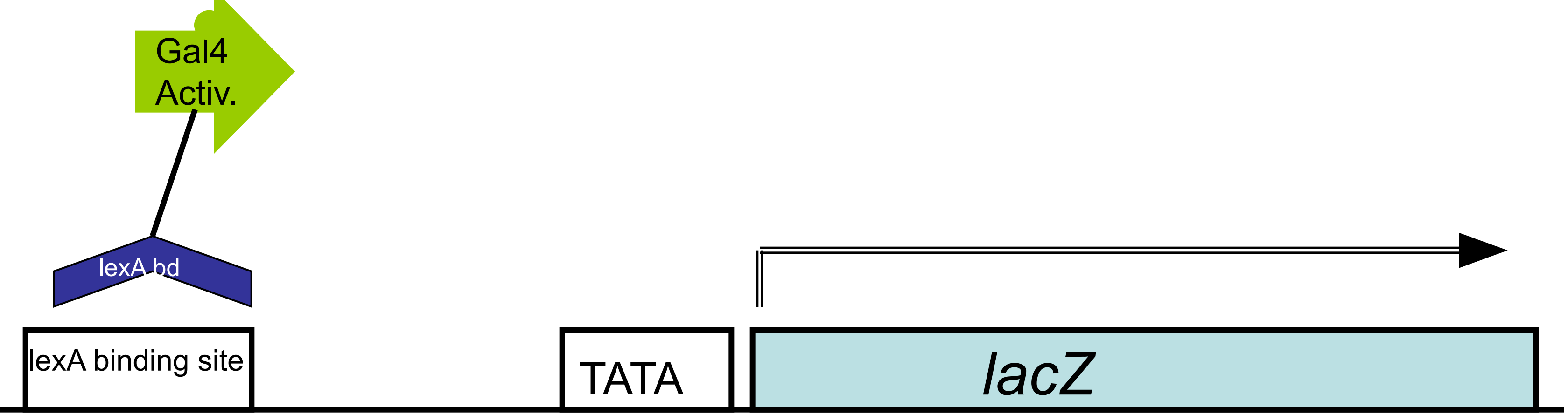
Gal4p is a modular protein



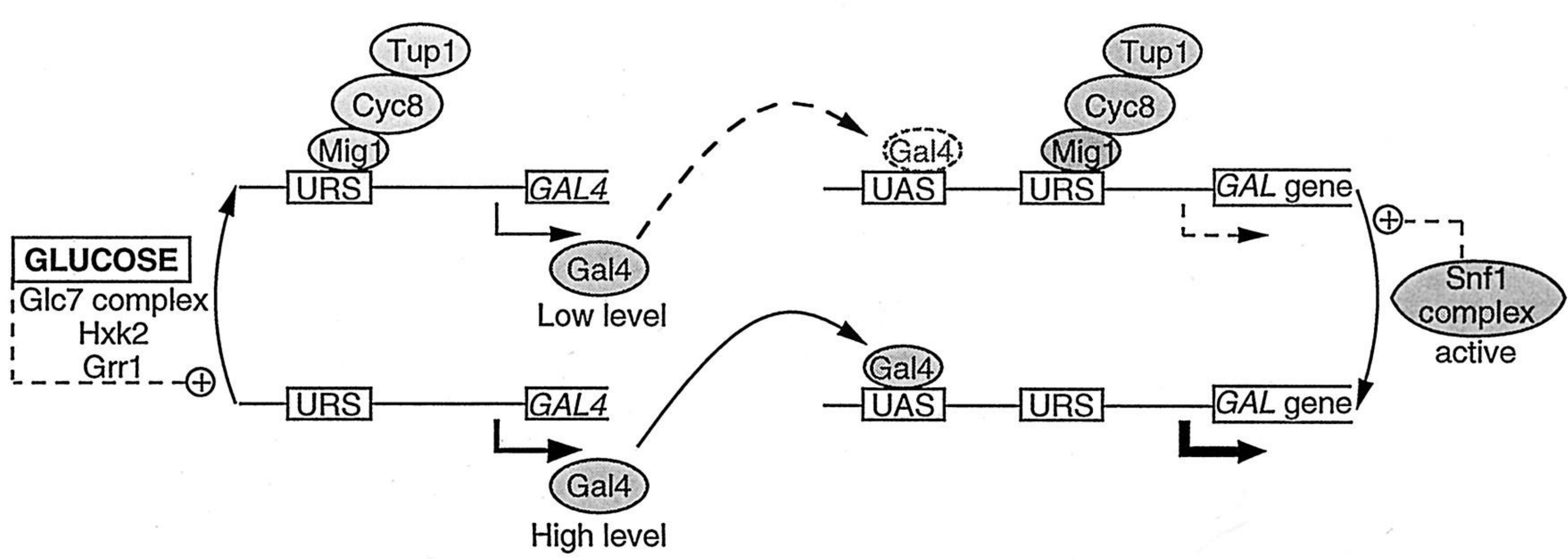
DNA binding domain (1-174) can bind DNA without the rest of the protein and can target a heterologous activation domain (VP16, viral activation domain) to promoters with a *GAL* UAS and exert transcriptional activation



Activation domain (758-881)can activate independently of the rest of the protein if fused to a heterologous DNA-binding domain (lexA bacterial DNA bd)



Expression of *GAL4* itself is regulated by glucose



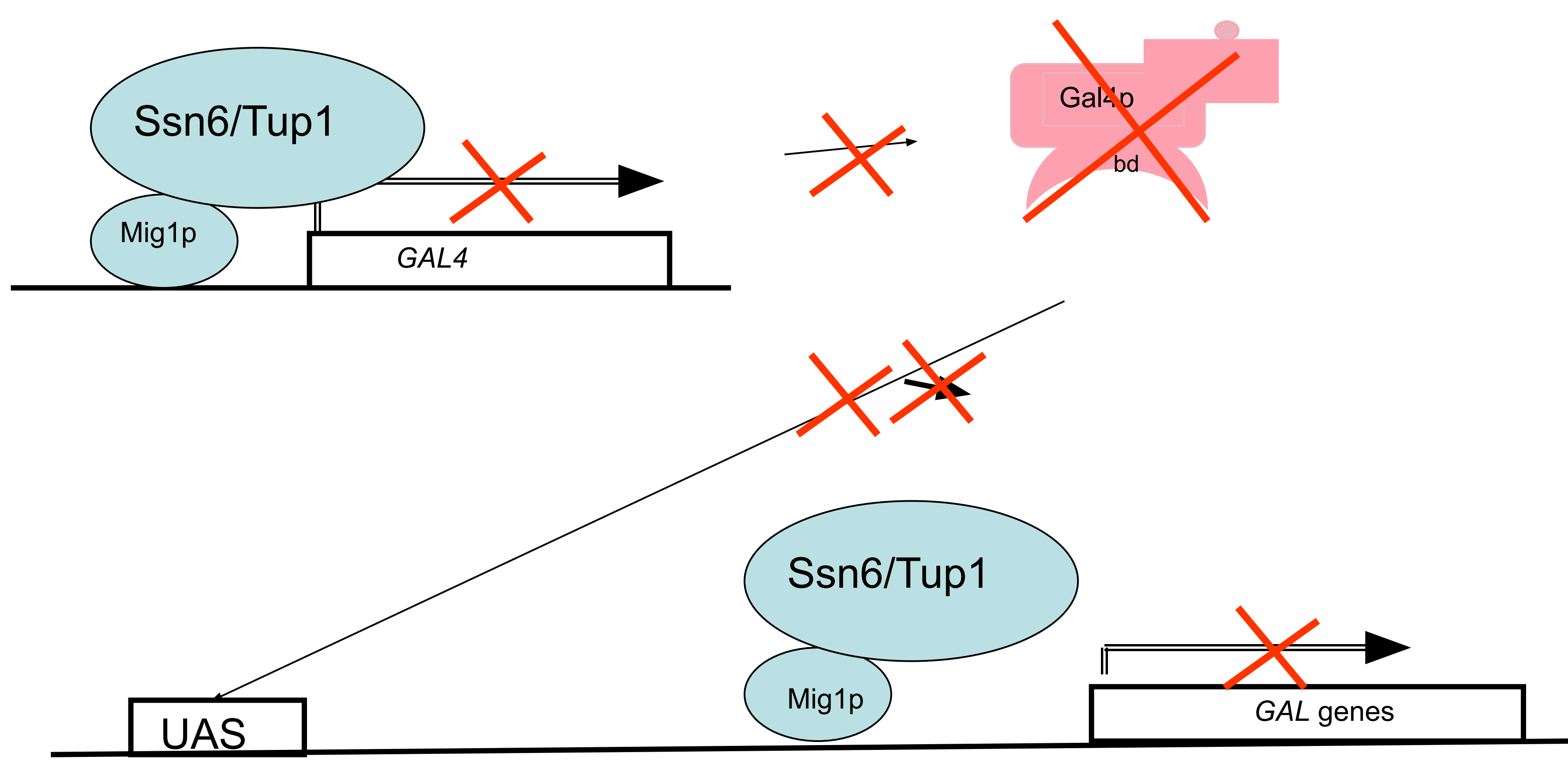
Under high glucose concentrations, the DNA – binding repressor protein Mig1p binds the regulatory region of *GAL4* and (also the other *GAL* genes) and turns off their transcription by recruiting the Tup1/Ssn6 (=Tup1/Cyc8) general repressor complex to the upstream regions of the *GAL* genes

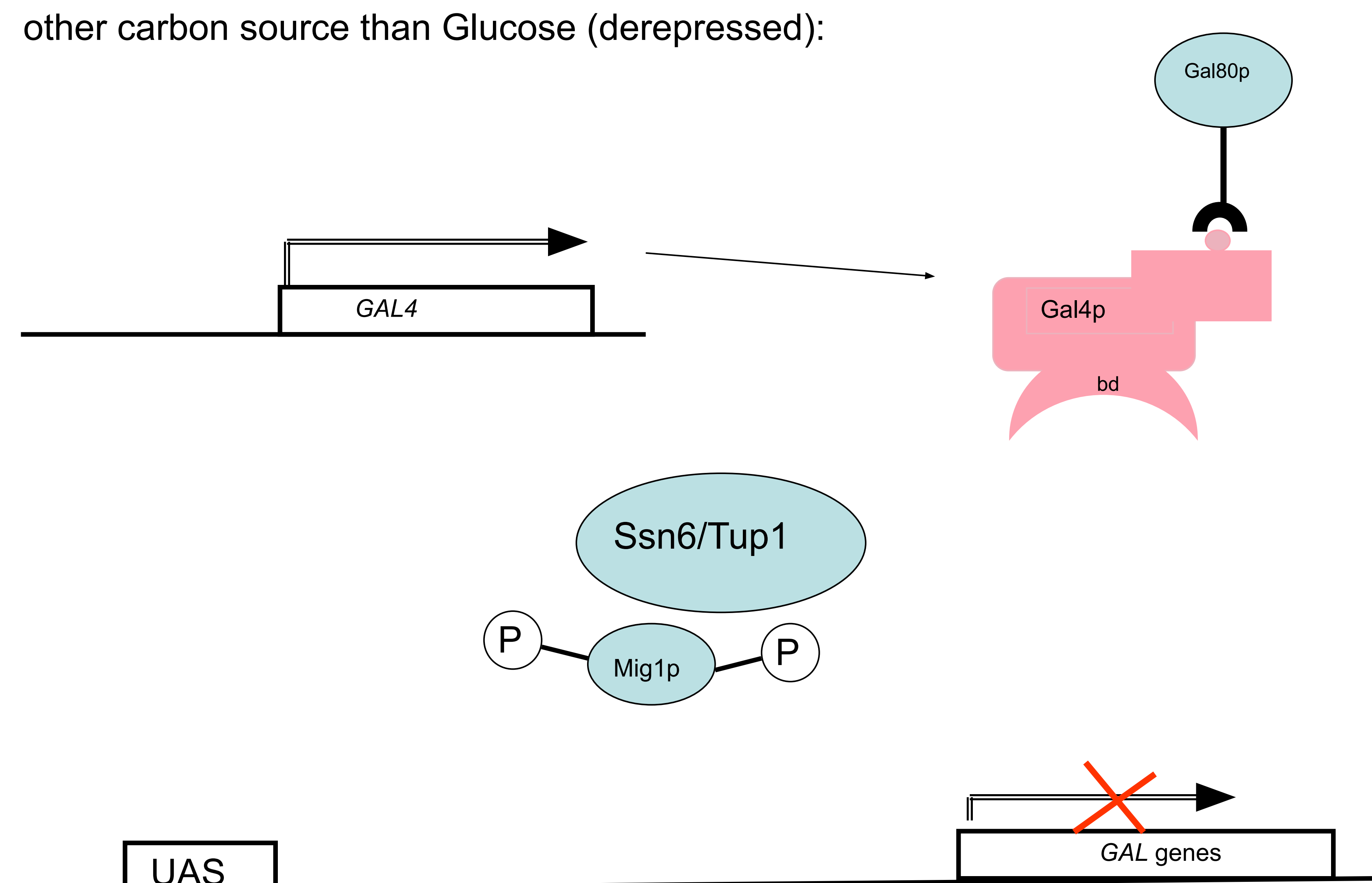
The galactose sensor: Gal3p

- Gal3p is a protein with high similarity (homology) to galactokinase
- No enzymatic activity
- In the presence of galactose, Gal3p binds the sugar and removes the Gal80p repressor from the Gal4p activator

In a nutshell….

Glucose (repressed):




other carbon source than Glucose (derepressed):
Gal80p
GAL4
Gal4p
bd
Ssn6/Tup1
P
Mig1p
P
UAS
GAL genes

Galactose:

Gal80p

GAL4

Gal4p

bd

+Galactose

Gal3

Activ./Gal80 ia

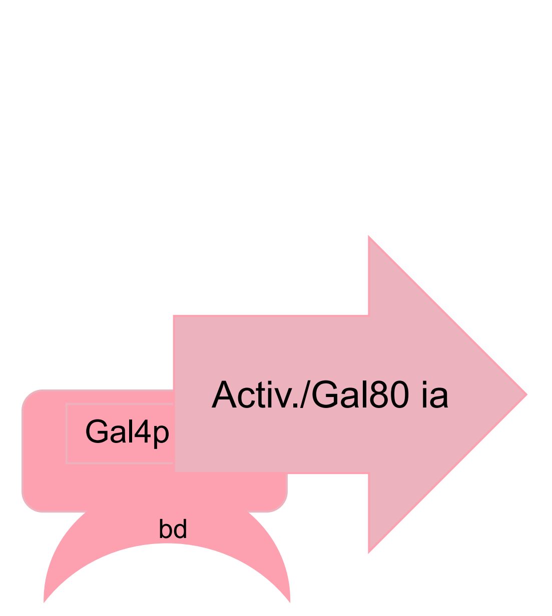
UAS

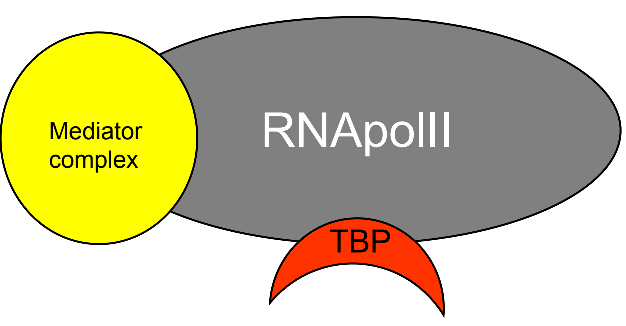
GAL genes

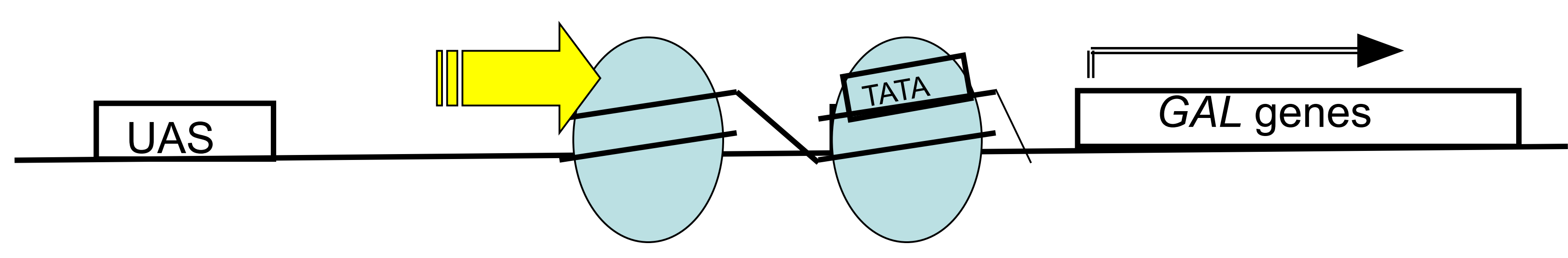
What is the mechanism of transcriptional activation by Gal4p?

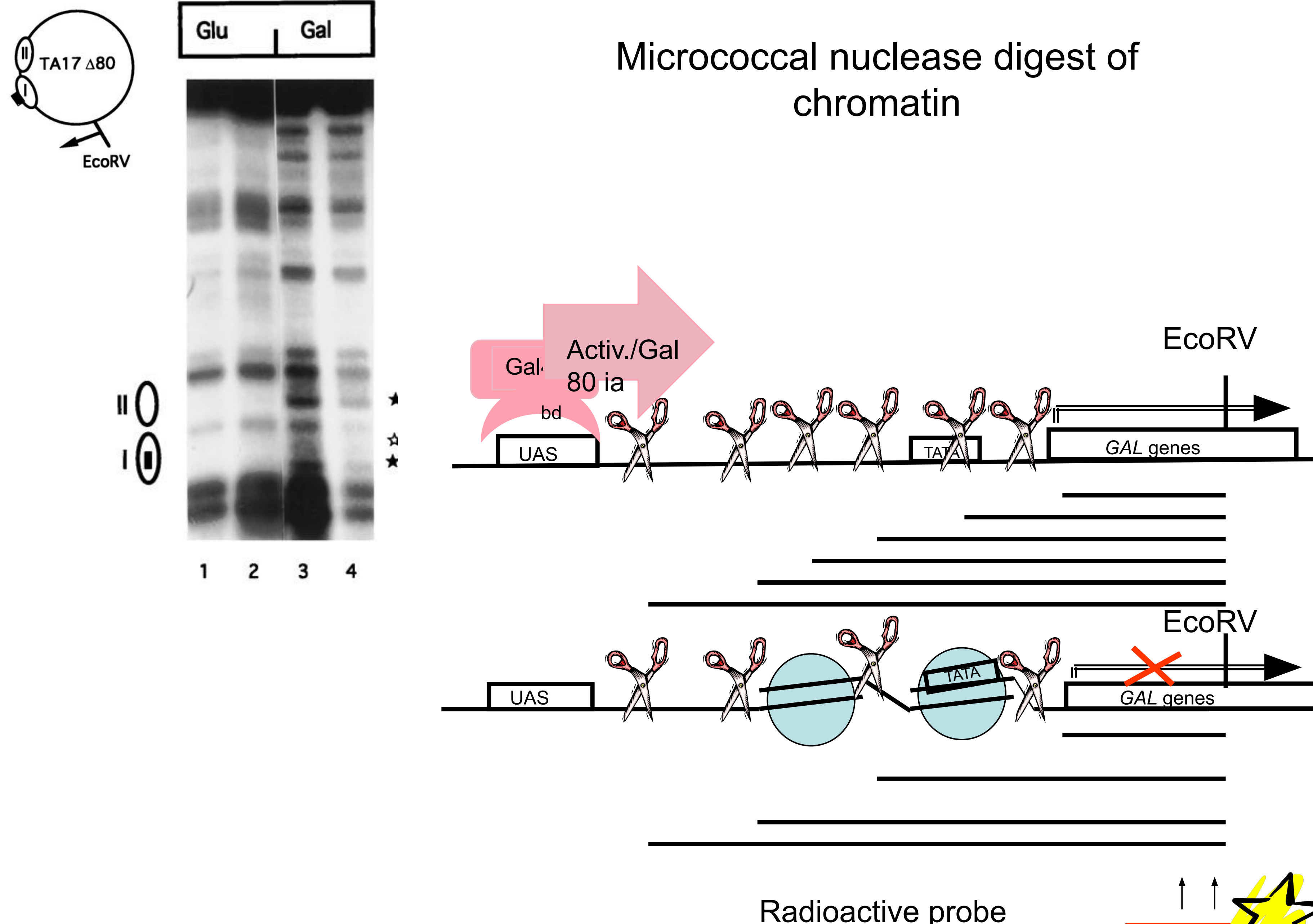
A. Gal4p activates by perturbing positioned nucleosomes that prevent access of RNA polymerase II to the promoters of the *GAL* genes

Mediator complex
RNApolII
TBP
TATA
UAS
GAL genes


Gal4p
Activ./Gal80 ia
bd


Mediator complex
RNApolII
TBP

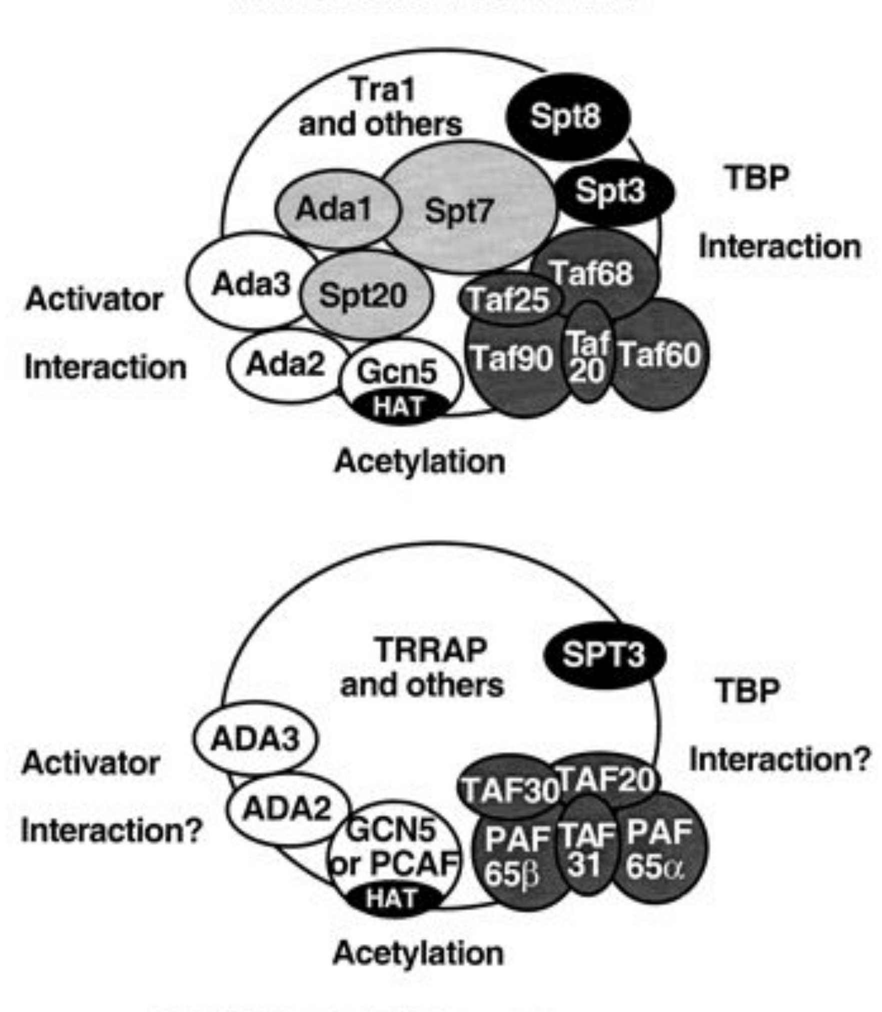

UAS
TATA
GAL genes



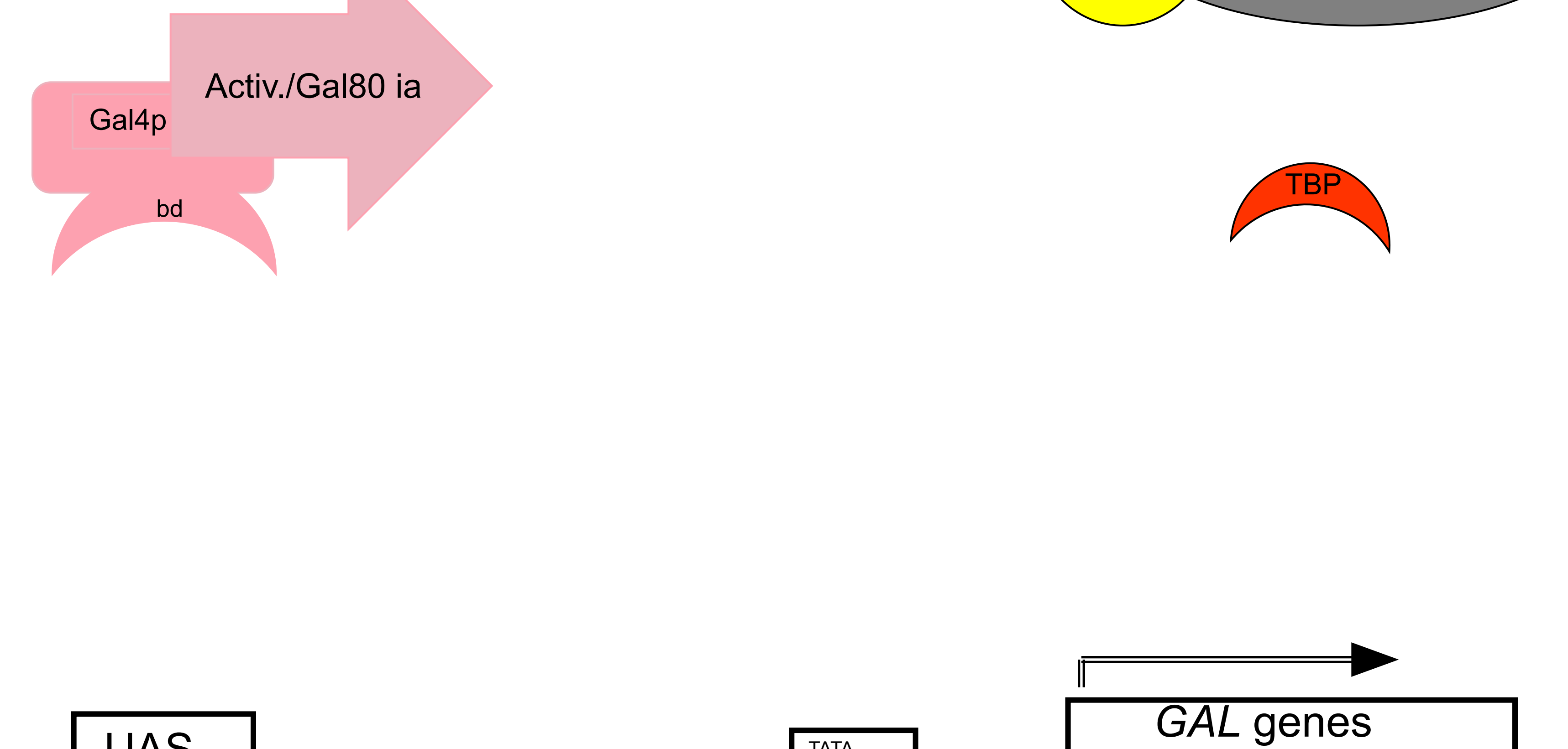
TA17 Δ80
II
I
EcoRV
Glu
Gal
1
2
3
4
Micrococcal nuclease digest of chromatin
Gal4
bd
Activ./Gal 80 ia
UAS
TATA
GAL genes
EcoRV
Radioactive probe anneals to 3' end of DNA

Nucleosome Perturbation via recruitment of Histone Acetyl-transferases (HATs)?

- Histones have positively charged N-terminal tails (K/R – rich) – interact with DNA
- Gal4 is suspected to recruit HATs (e.g Gcn5p/SAGA complex) to the promoters of the *GAL* genes and thereby locally disrupt histone-DNA interaction



B. Gal4p interacts directly with the TATA-binding protein or the polymerase II complex



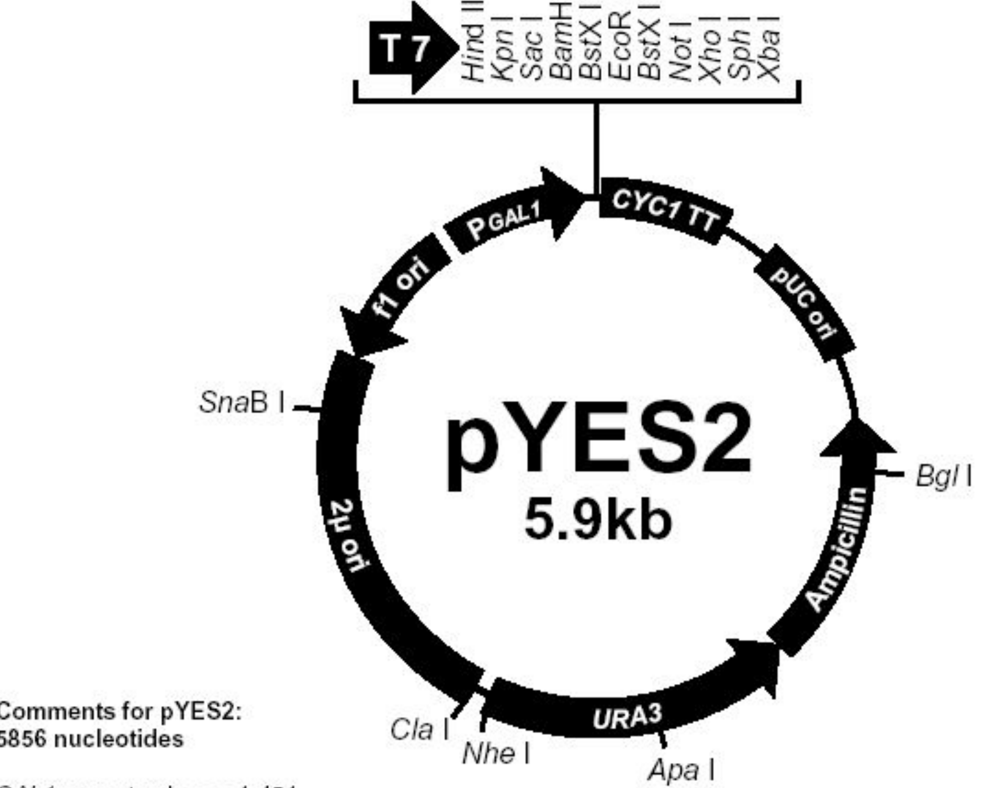
Relevance of the Gal regulation research today?

- General understanding of basic molecular principles of gene activation
- Model for the functioning of biological regulatory circuits
  - **A general mechanism for network-dosage compensation in gene circuits**. Acar M, Pando BF, Arnold FH, Elowitz MB, van Oudenaarden A. **Science. 2010** Sep 24;329(5999):1656-60

Galactose induction can be utilized to overexpress heterologous genes

- Genes of interest can be fused to the promoter and regulatory regions of galactose-regulated genes

GAL1 promoter (4 Gal4p binding sites)



GAL1 promoter: bases 1-451
T7 promoter/priming site: bases 475-494
Multiple cloning site: bases 501-600
CYC1 transcription terminator: bases 608-856
pUC origin: bases 1038-1711
Ampicillin resistance gene: bases 1856-2716 (C)
URA3 gene: bases 2734-3841 (C)
2 micron (µ) origin: bases 3845-5316
f1 origin: bases 5384-5839 (C)
(C) = complementary strand

Three expression levels:

- Repressed (2% glucose) → no expression
- Derepressed (2% Raffinose, 3% Glycerol) → intermediate expression
- Activated (2% Galactose) → high expression

Useful for:

- Overexpression for purification
- Multicopy effect studies
- Study of essential genes (genes for which deletions are lethal)

Similar: Oleate induction:

-Oleate induced genes are involved in peroxisomal proliferation and in β-oxidation

-Activator is a heterodimer of the Oaf1p/Pip2p activators which bind to oleate response elements (OREs)

-The ORE consensus is currently viewed as two inverted CGG triplets spaced by 14 (formerly 15) to 18 intervening nucleotides (N), i.e. $CGGN_3TNAN_{8}\text{-}_{12}CCG$

-Currently, the plasmid available has the promoter and terminator sequences of the oleate-induced *CTA1* (peroxisomal catalase) gene

-*CTA1* is glucose repressed similar to the *GAL* genes

-Three expression levels:

-Repressed (2% Glucose)

-Derepressed (2% Raffinose, 3% Glycerol)

-Activated (0.2% oleate, 0.02% Tween, 0.05% Glucose)

Expression from inducible promoters allows investigation of essential genes

- Essential genes are genes required for viability of the cell
- Deletions of these genes are inviable, deletion are only viable as heterozygous diploids, or deletion strains have to carry a plasmid with a wild type copy of the gene
- Shuffling in plasmids carrying mutant partial function alleles is one way of investigating the function
- Introduction of plasmids with the essential gene expressed from an inducible promoter allow more precise investigation