Interactive Animations

http://www.aw-bc.com/watson/
Click on Molecular Biology of the Gene, 6th edition
Class code: cm473170

DNA polymerase
http://media.pearsoncmg.com/bc/bc_martini_ap_slim/assets/animations/ch08_polymerization.html

Replication
http://media.pearsoncmg.com/bc/bc_martini_ap_slim/assets/animations/ch08_replication.html

Prokaryotic versus Eukaryotic Chromosomes

Prokaryotic Chromosomes

•Many prokaryotes contain a single circular chromosome.

•Prokaryotic chromosomes are condensed in the nucleoid via DNA supercoiling and the binding of various architectural proteins.

•Because prokaryotic DNA can interact with the cytoplasm, transcription and translation occur simultaneously.

•Most prokaryotes contain only one copy of each gene (i.e., they are haploid).

•Nonessential prokaryotic genes are commonly encoded on extrachromosomal plasmids.

•Prokaryotic genomes are efficient and compact, containing little repetitive DNA.

Eukaryotic Chromosomes

•Eukaryotes contain multiple linear chromosomes.

•Eukaryotic chromosomes are condensed in a membrane-bound nucleus via histones.

•In eukaryotes, transcription occurs in the nucleus, and translation occurs in the cytoplasm.

•Most eukaryotes contain two copies of each gene (i.e., they are diploid).

•Some eukaryotic genomes are organized into operons, but most are not.

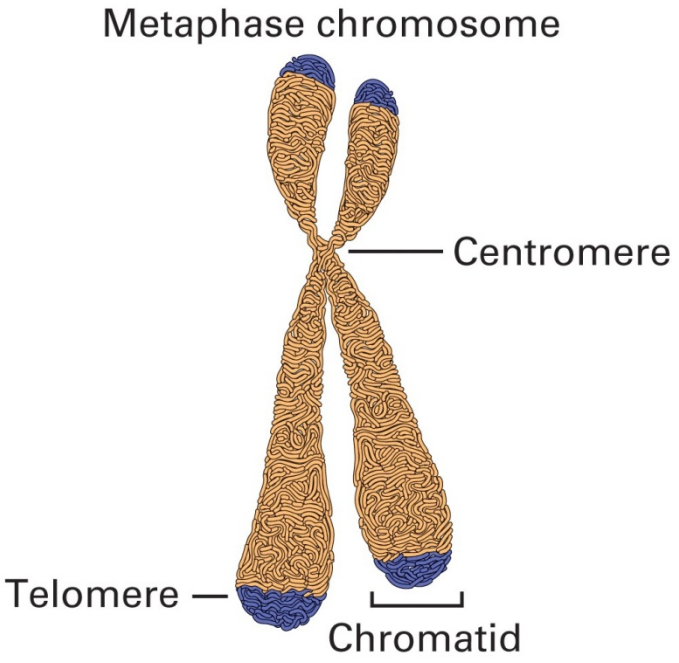
•Extrachromosomal plasmids are not commonly present in eukaryotes.

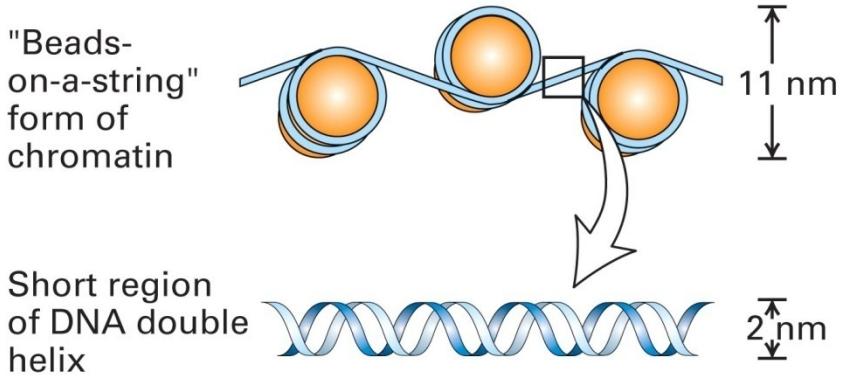
•Eukaryotes contain large amounts of noncoding and repetitive DNA.

Nucleosome assembly (eukaryotes)

http://media.pearsoncmg.com/bc/bc_martini_ap_slim/assets/animations/ch07_nucleosome.html

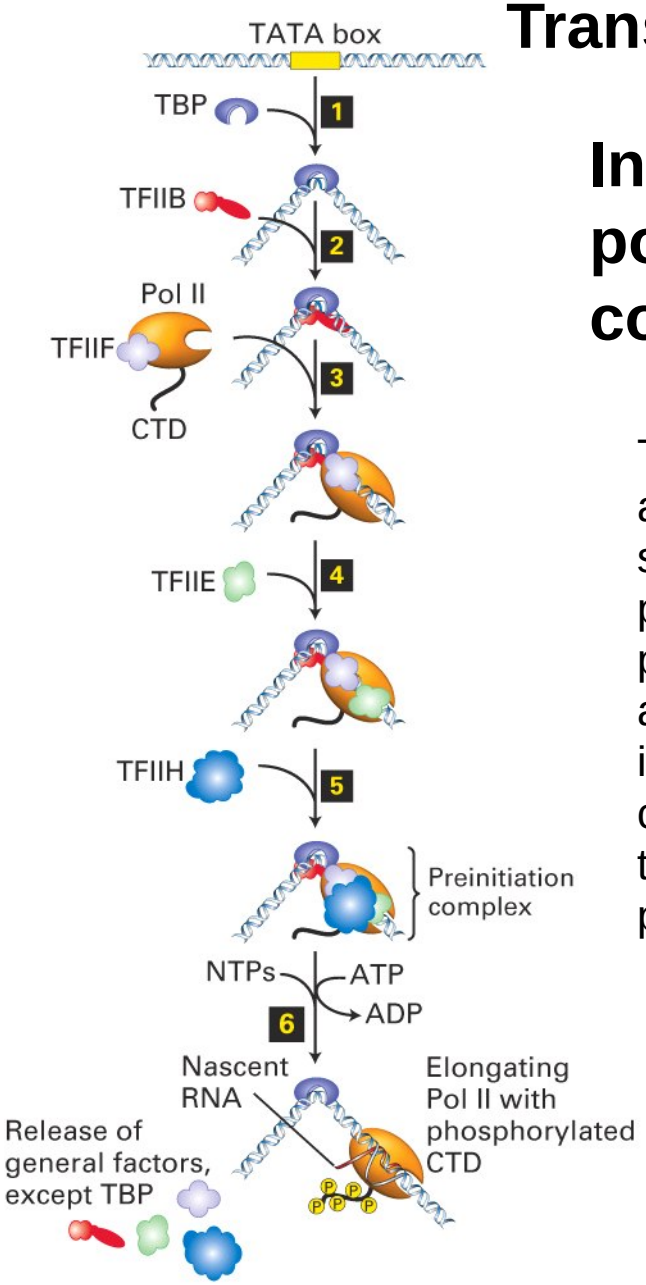
Nucleosome packaging of chromatin





The basic structural unit of chromatin is the nucleosome, which is composed of **147 bp** of DNA wrapped tightly around a disk-shape core of **histone** proteins.

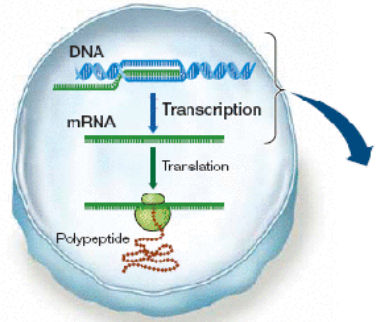
Transcription

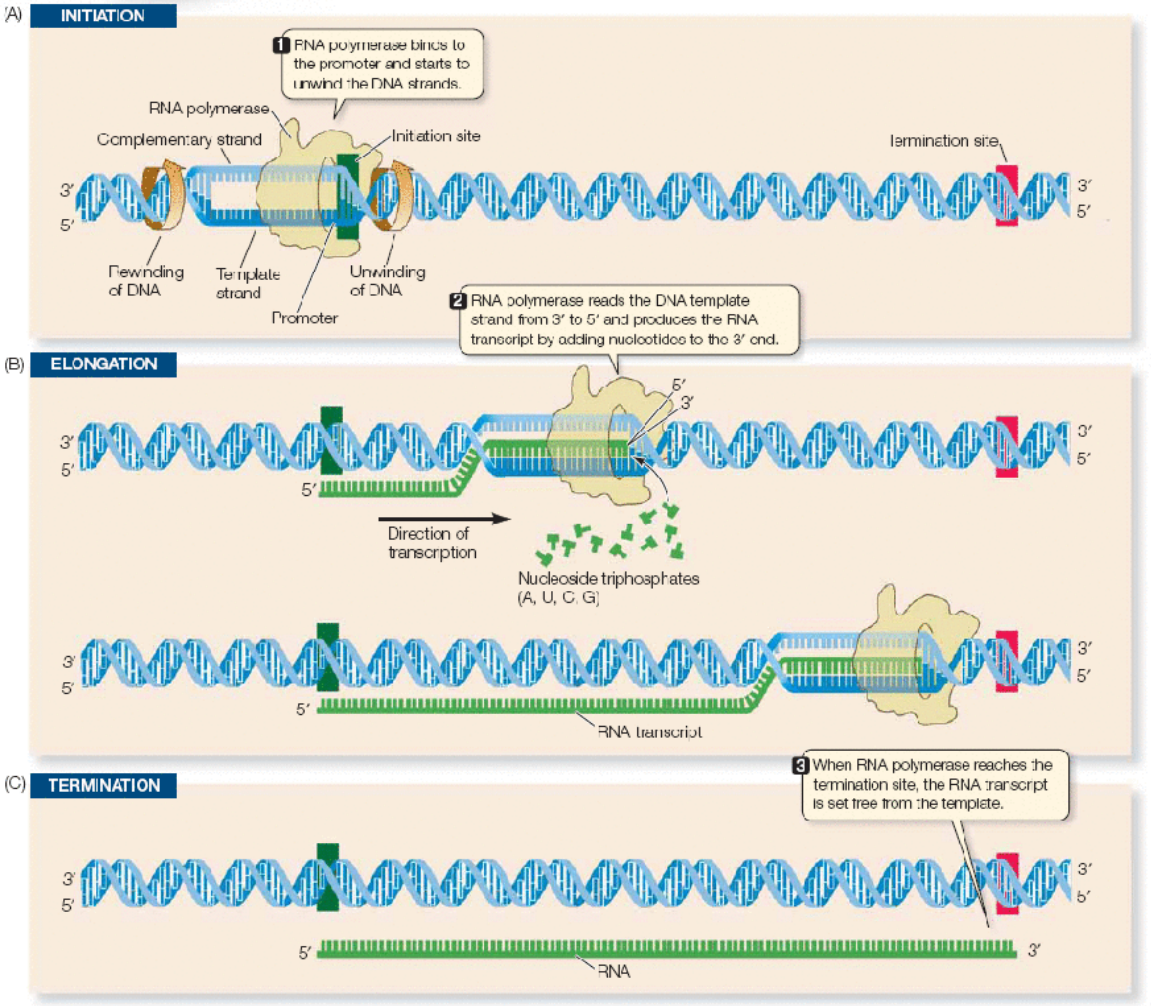


In vitro assembly of RNA polymerase II preinitiation complex

The indicated general transcription factors and purified RNA polymerase II (Pol II) bind sequentially to TATA-box DNA to form a preinitiation complex. ATP hydrolysis then provides the energy for unwinding of DNA at the start site by a TFIIH subunit. As PolII initiates transcription in the resulting open complex, the polymerase moves away from the promoter and its CTD becomes phosphorylated.

Transcription



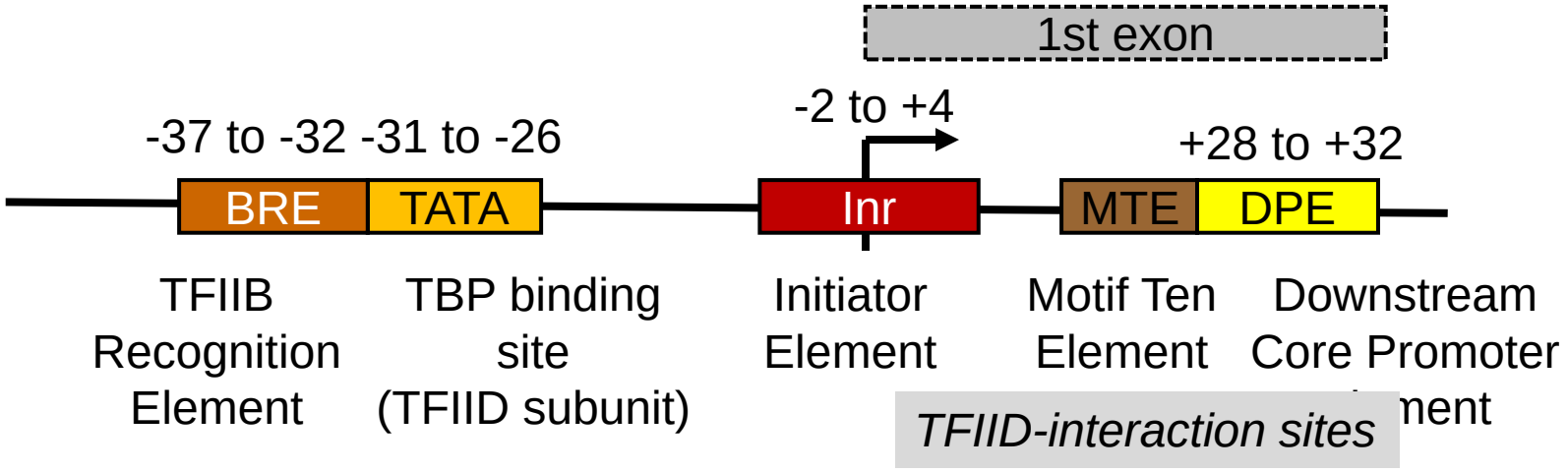


http://media.pearsoncmg.com/bc/bc_martini_ap_slim/assets/animations/ch12_transcription.html

Regulatory elements

Core Promoter

- Proximal to the transcription start site (+/- 35 bp)
- Determines timing of gene expression during development
- Interacts with pre-initiation complex

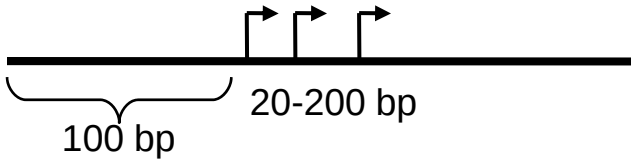


10,000 human promoters

- Inr: 50%
- DPE, BRE: 25%
- TATA: 12.5%
- None: 25%!

Maston et al., Ann. Rev. Genom. Human Genet., 2006: 7, 29

CpG island: CpG (20-50 nts)



CpG islands are characterized by: overrespresentation of the di nucleotide CpG around the transcription start site, lack of cytosine methylation, low levels of histone H1, high levels of histone acetylation and hypersensitivity to DNaseI which has been equated with nucleosome -free regions.

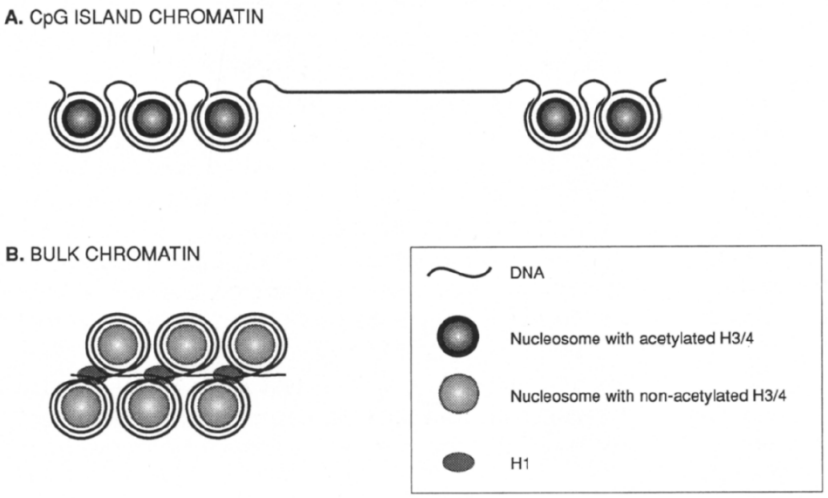


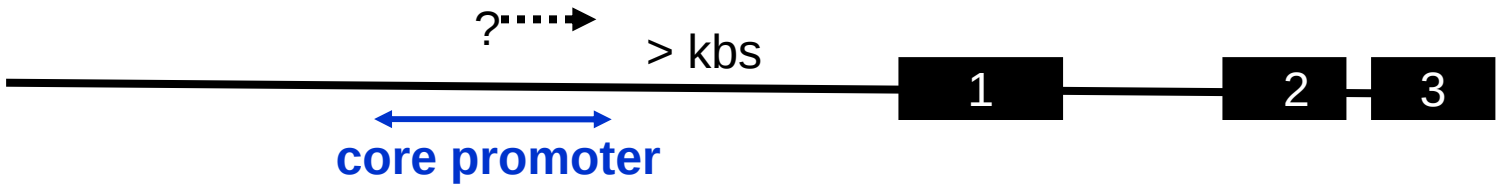
Figure 9. Diagrammatic Interpretation of the Differences between CpG Island Chromatin and Bulk Transcriptionally Inactive Chromatin

The two structures are drawn to the same scale, and both involve about 1000 bp of DNA, which is the length of DNA in an average CpG island. CpG island chromatin is considered to be in a loose beads-on-a-string configuration. The nucleosome-free region corresponds to the amount of DNA that is required to wrap around the nucleosome twice (~170 bp). DNAase I–hypersensitive sites are often about this size. Stippled circles represent either acetylated (heavy outline) or nonacetylated (light outline) nucleosomes. Small ovals represent molecules of histone H1. The compaction of the DNA in bulk chromatin is underestimated in the diagram because H1-associated nucleosomes are thought to form a helical (solenoid) structure.

Identifying network nodes: Regulatory elements

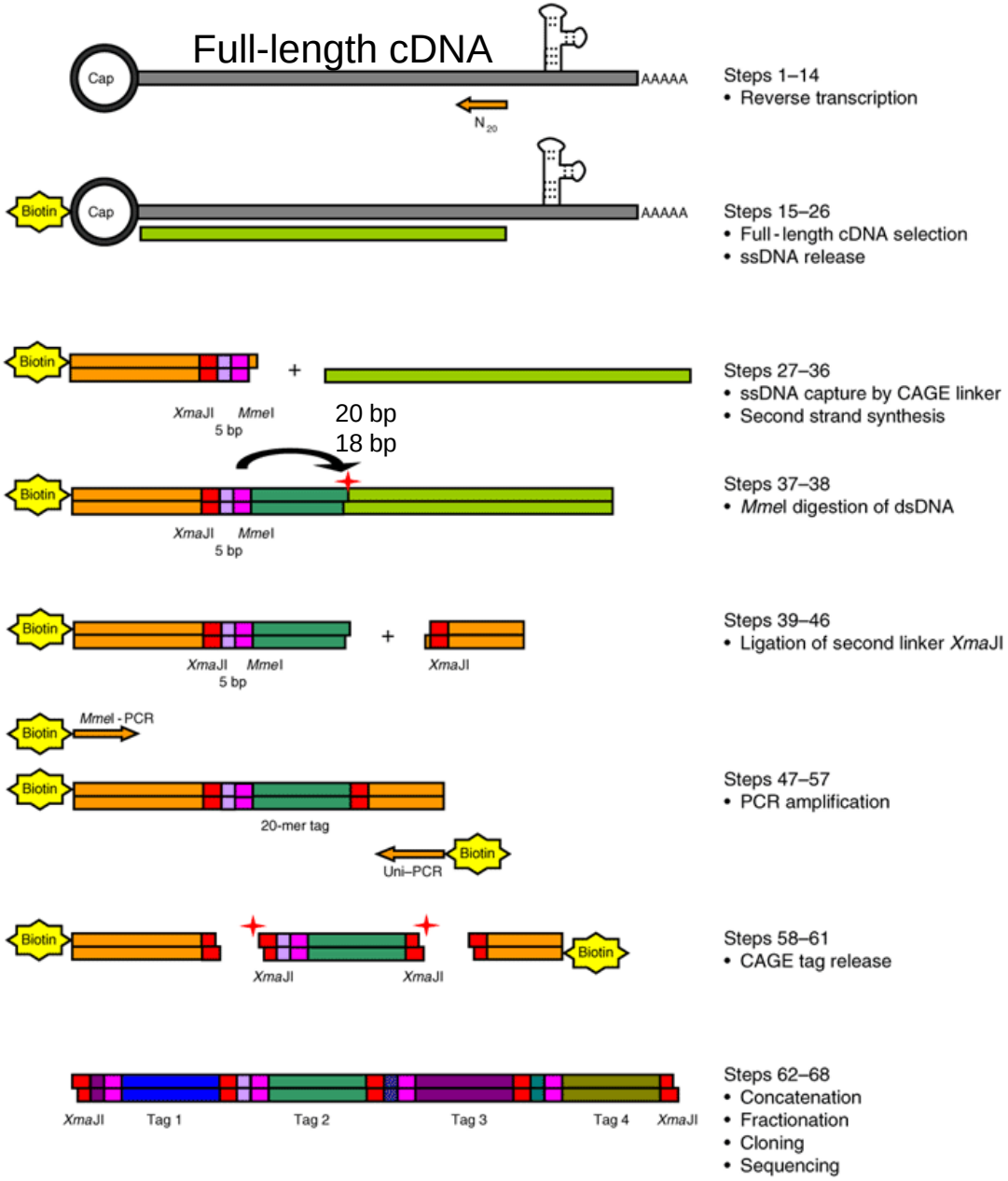
Core Promoter

Instead of focusing on core elements, try to find the transcriptional start site:



Cap Analysis of Gene Expression or CAGE analysis
Shiraki et al., PNAS, 2003: 100, 15776.

Finding the TSS by CAGE



Steps 1–14
- Reverse transcription

Steps 15–26
- Full-length cDNA selection
- ssDNA release

Steps 27–36
- ssDNA capture by CAGE linker
- Second strand synthesis

Steps 37–38
- *Mme*I digestion of dsDNA

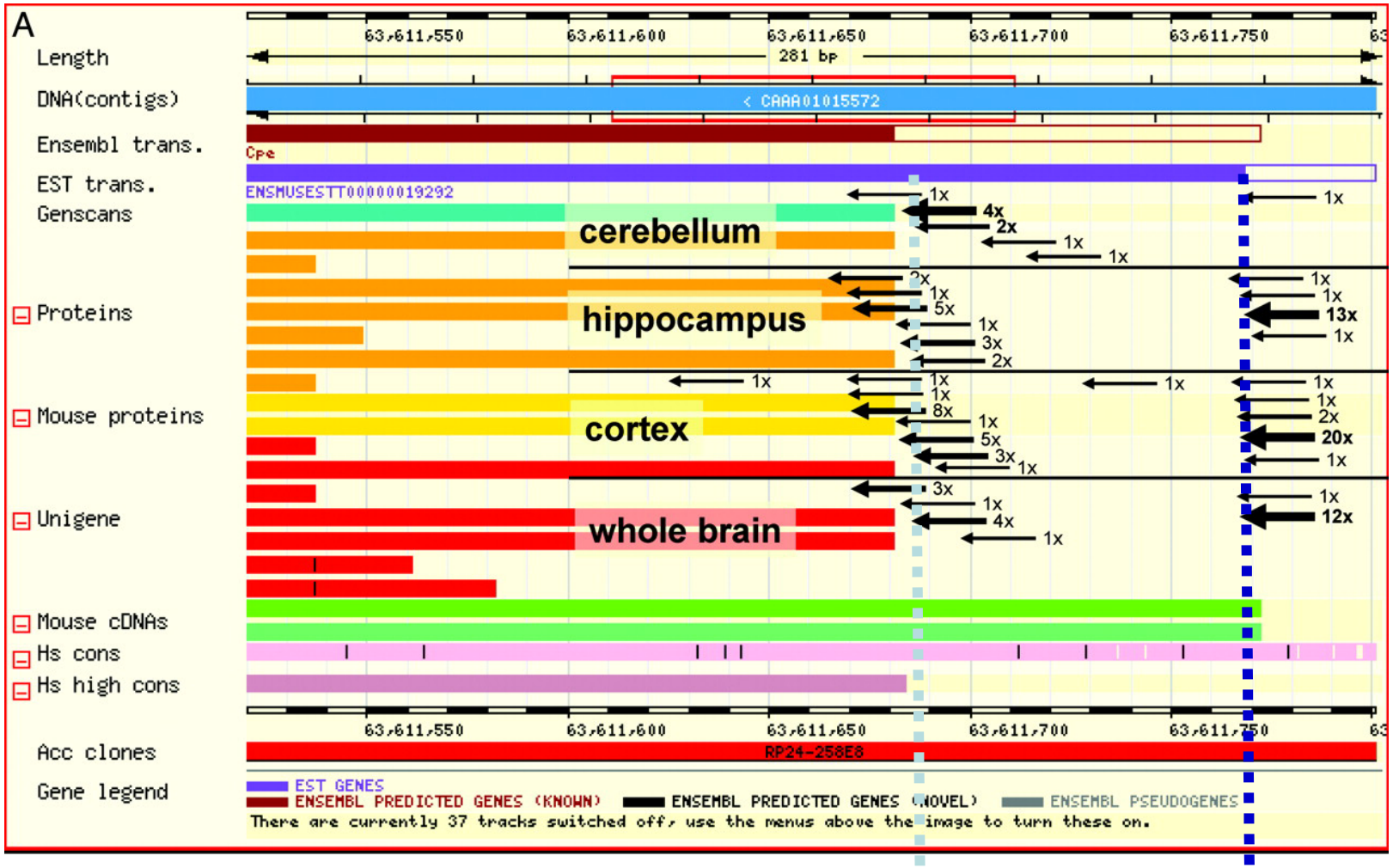
Steps 39–46
- Ligation of second linker *Xma*JI

Steps 47–57
- PCR amplification

Steps 58–61
- CAGE tag release

Steps 62–68
- Concatenation
- Fractionation
- Cloning
- Sequencing

Starting-point identification for the Carboxypeptidase E gene (119 kb)

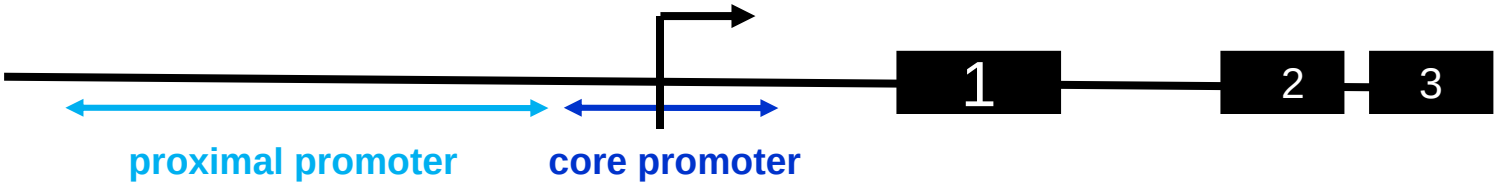


TSS 1 ~ 80 bp TSS 2

Shiraki et al., PNAS, 2003: 100, 15776.

Regulatory elements

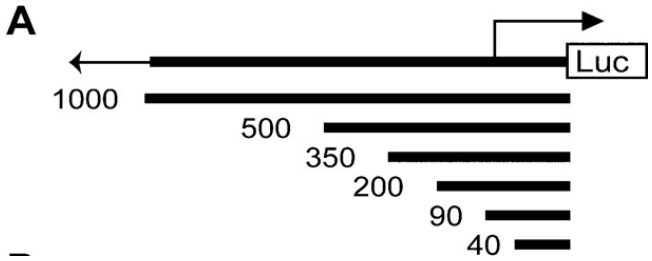
Proximal Promoter



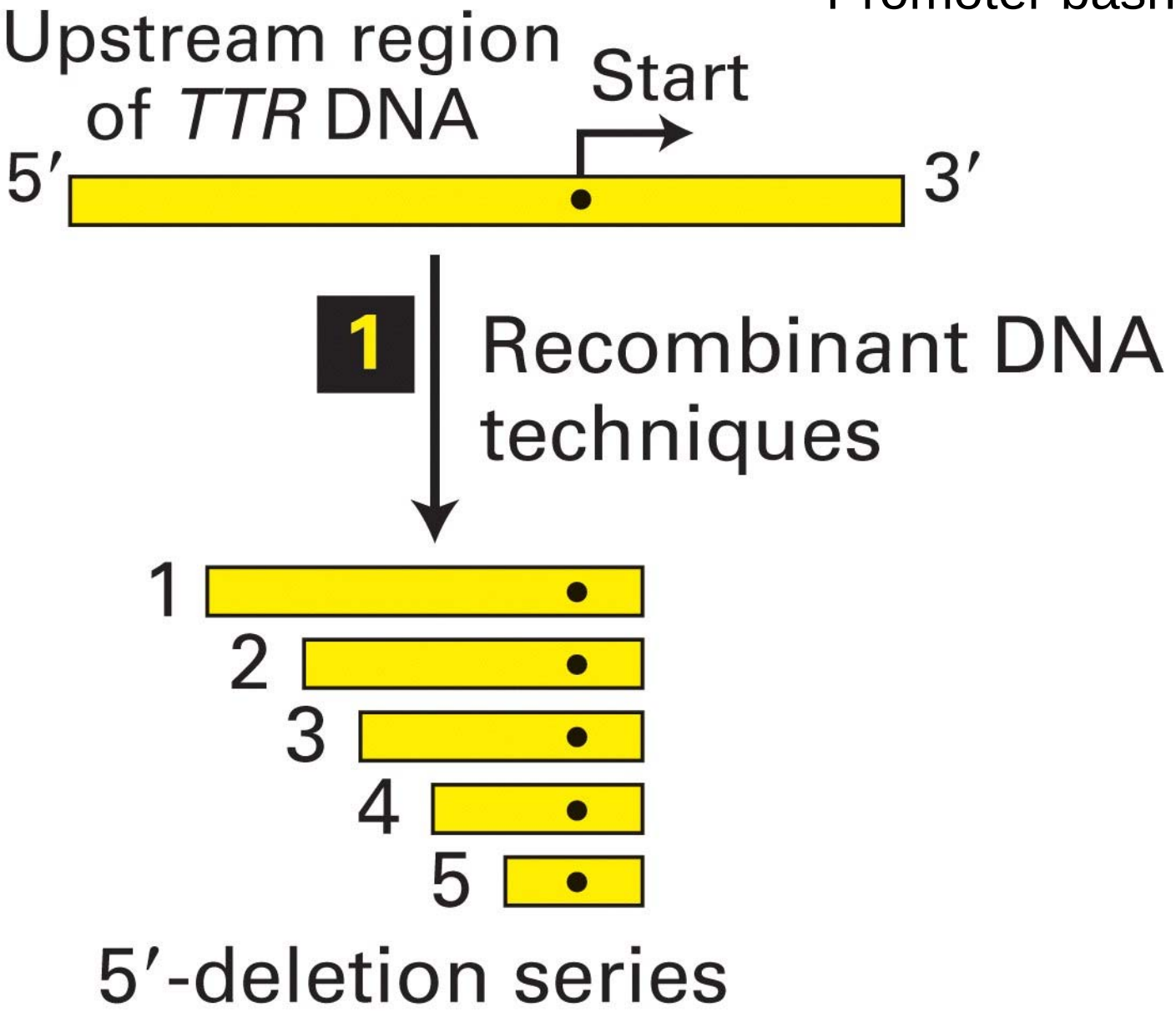
Upstream sequences (~500 bp – 1 kb)

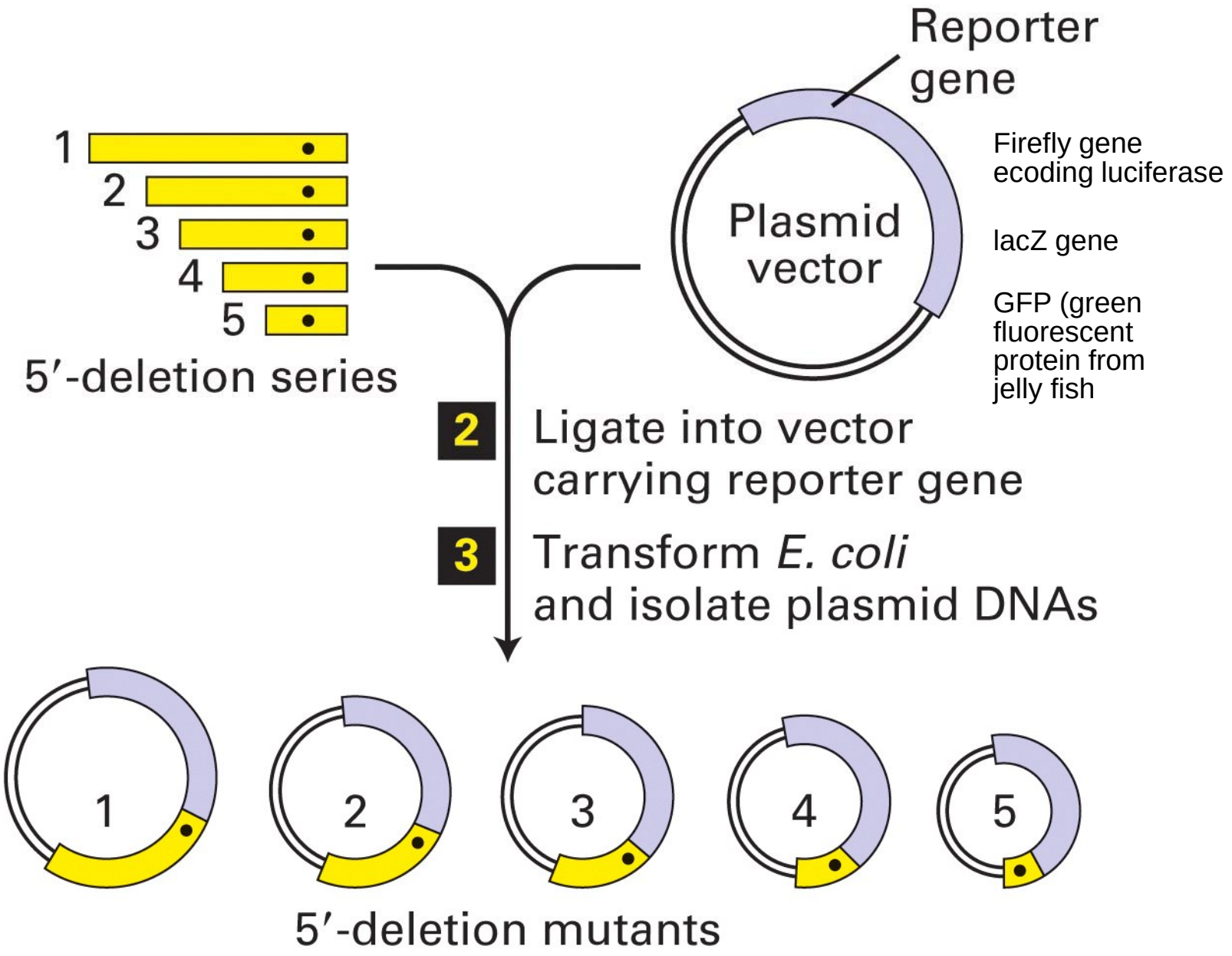
- **Augment the level of expression or choice of tissue specificity**
- **Separation from enhancers?**
  (negative elements, e.g. Cooper et al., Genome Res., 2006: 16, 1)

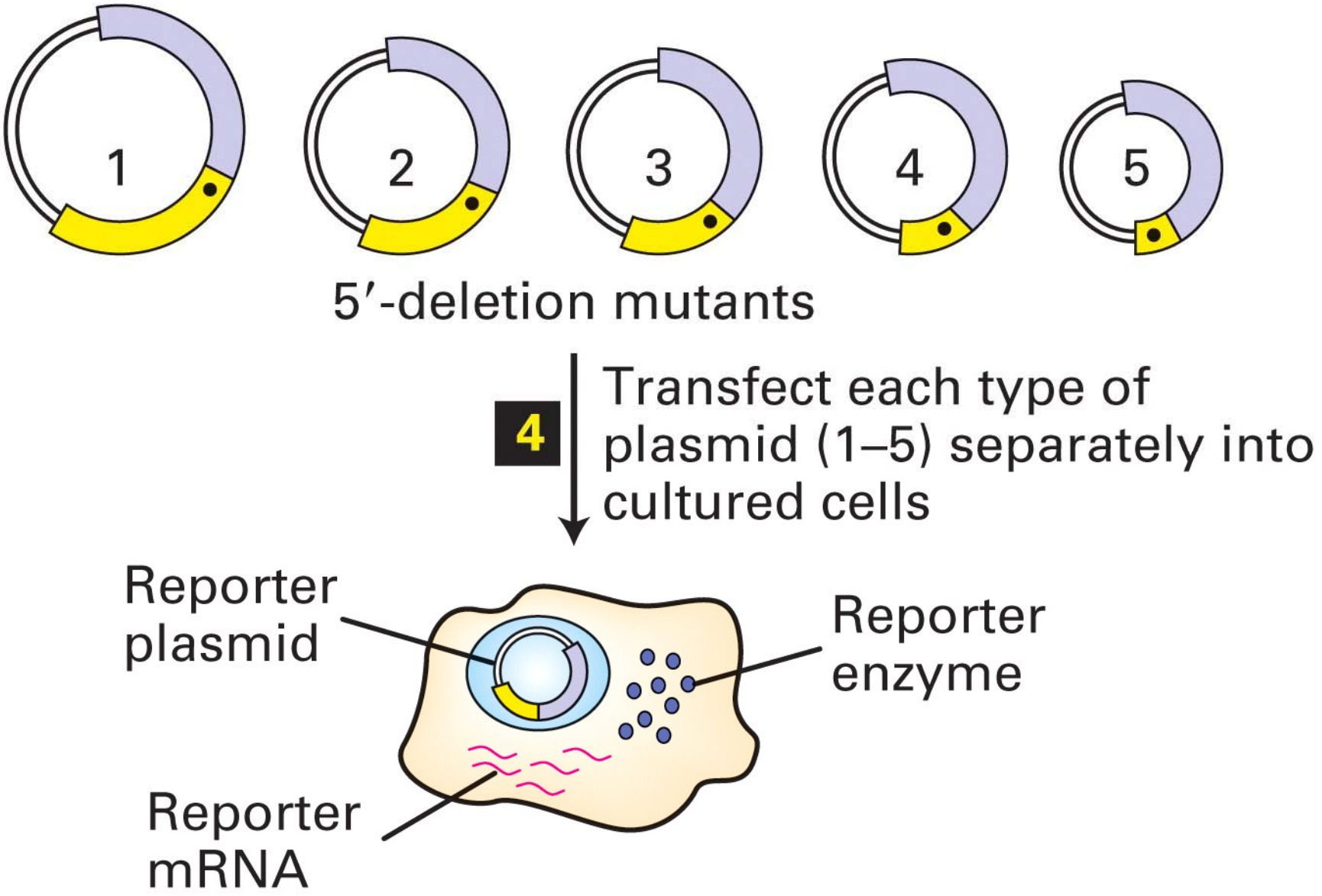
Average activity of 45 promoters across 7 cell lines



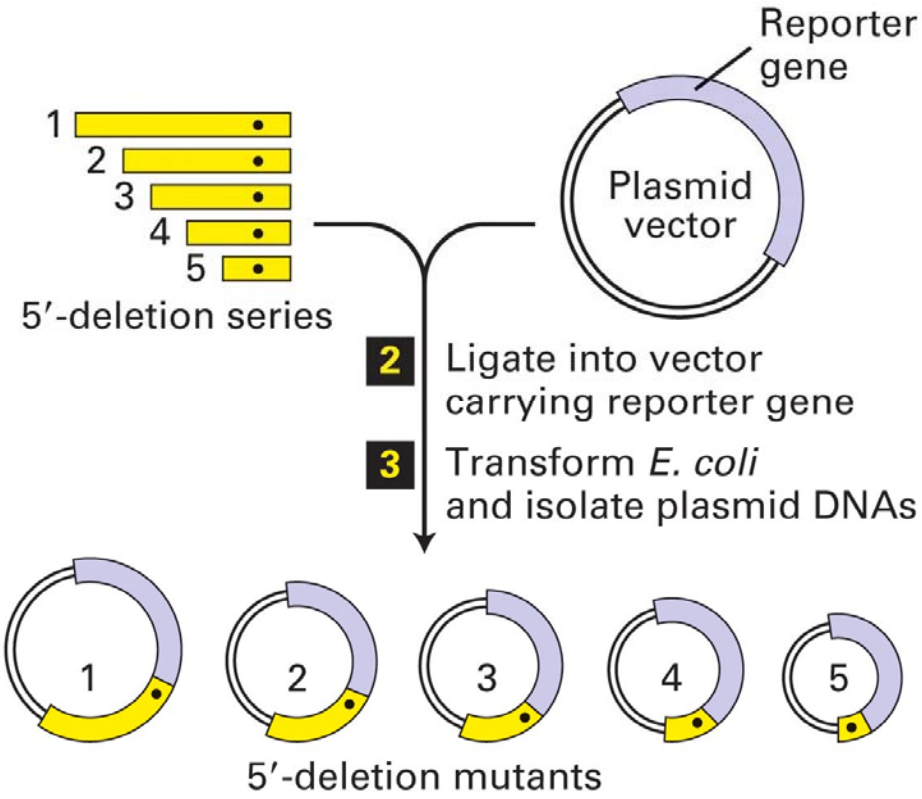
Promoter bashing




Reporter gene
Firefly gene ecoding luciferase
lacZ gene
GFP (green fluorescent protein from jelly fish
Plasmid vector
1
2
3
4
5
5′-deletion series
2 Ligate into vector carrying reporter gene
3 Transform E. coli and isolate plasmid DNAs
1
2
3
4
5
5′-deletion mutants


1
2
3
4
5
5′-deletion mutants
4
Transfect each type of plasmid (1–5) separately into cultured cells
Reporter plasmid
Reporter enzyme
Reporter mRNA

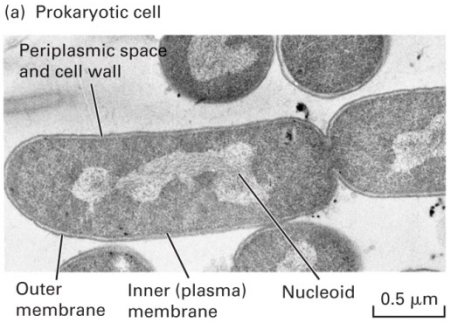
Frequently used reporter genes

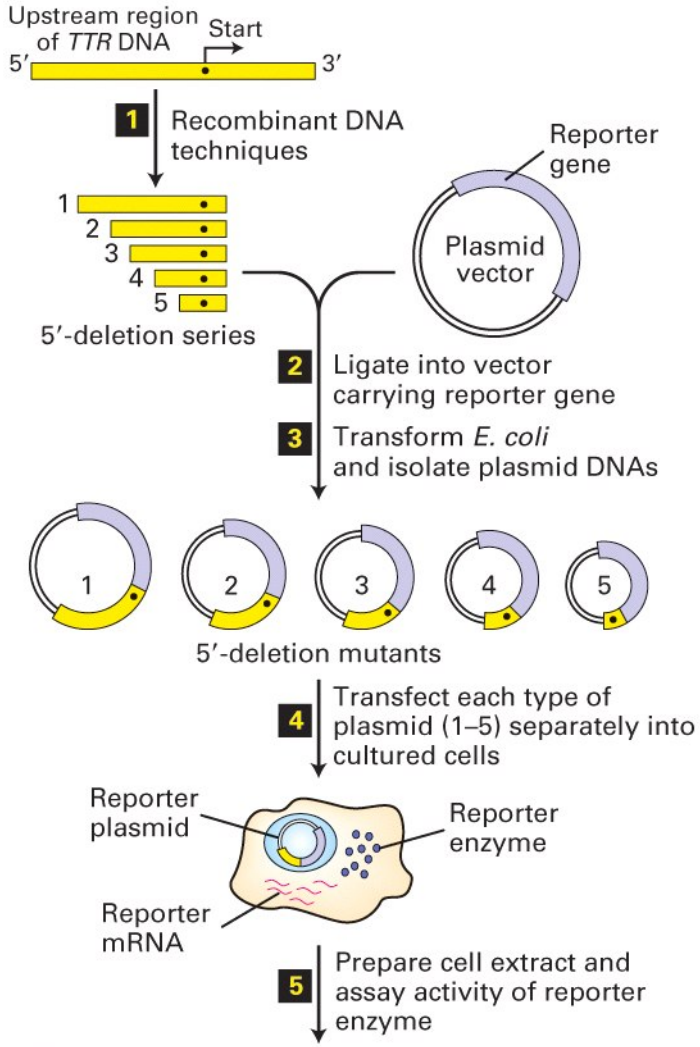


Firefly gene ecoding luciferase

GFP (green fluorescent protein from) jelly fish Aequorea victoria

lacZ gene of bacteria

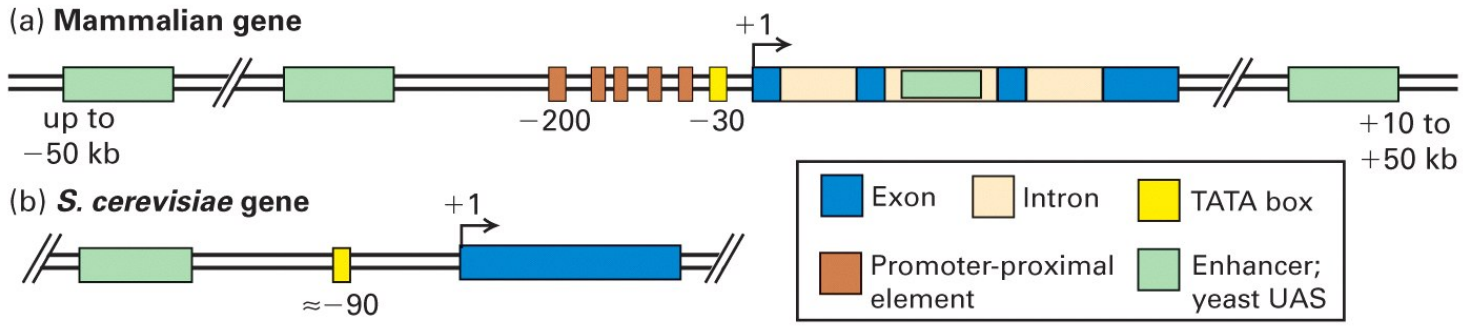




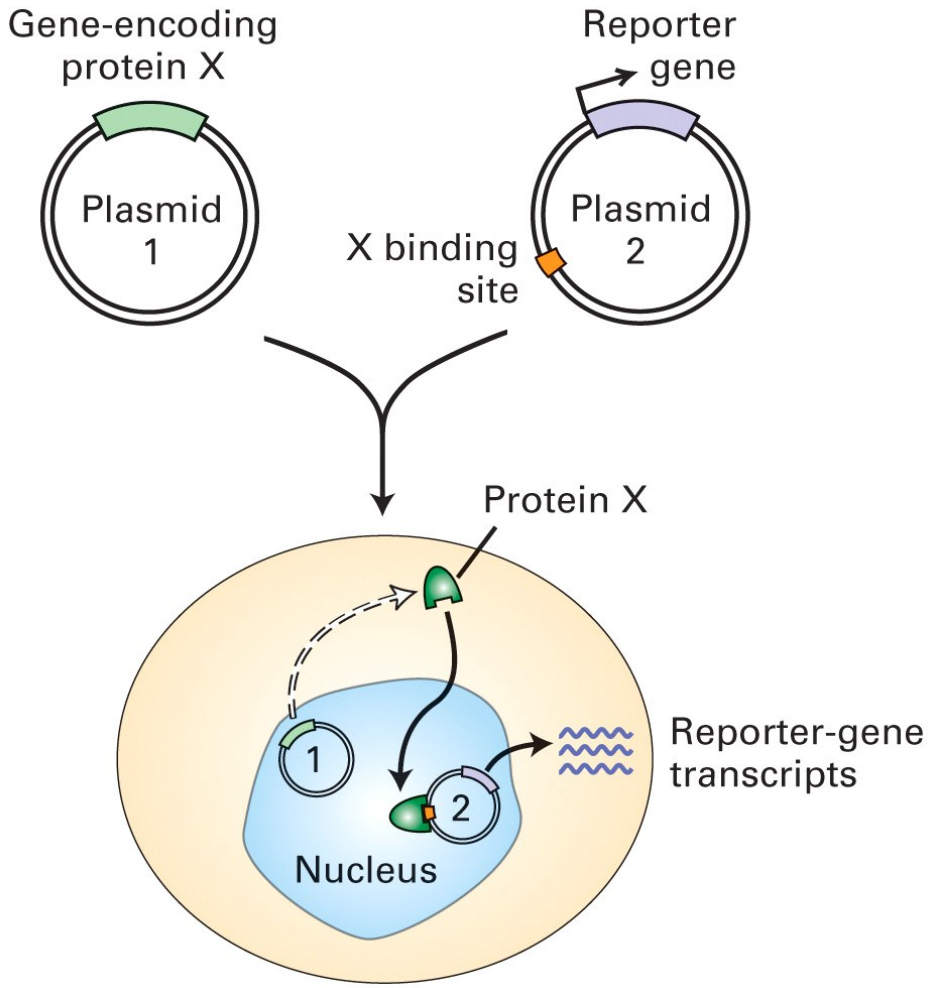
Plasmid no.	Reporter-gene expression
1	+++
2	+++
3	+
4	+
5	–

Most eukaryotic genes are regulated by multiple transcription-control elements

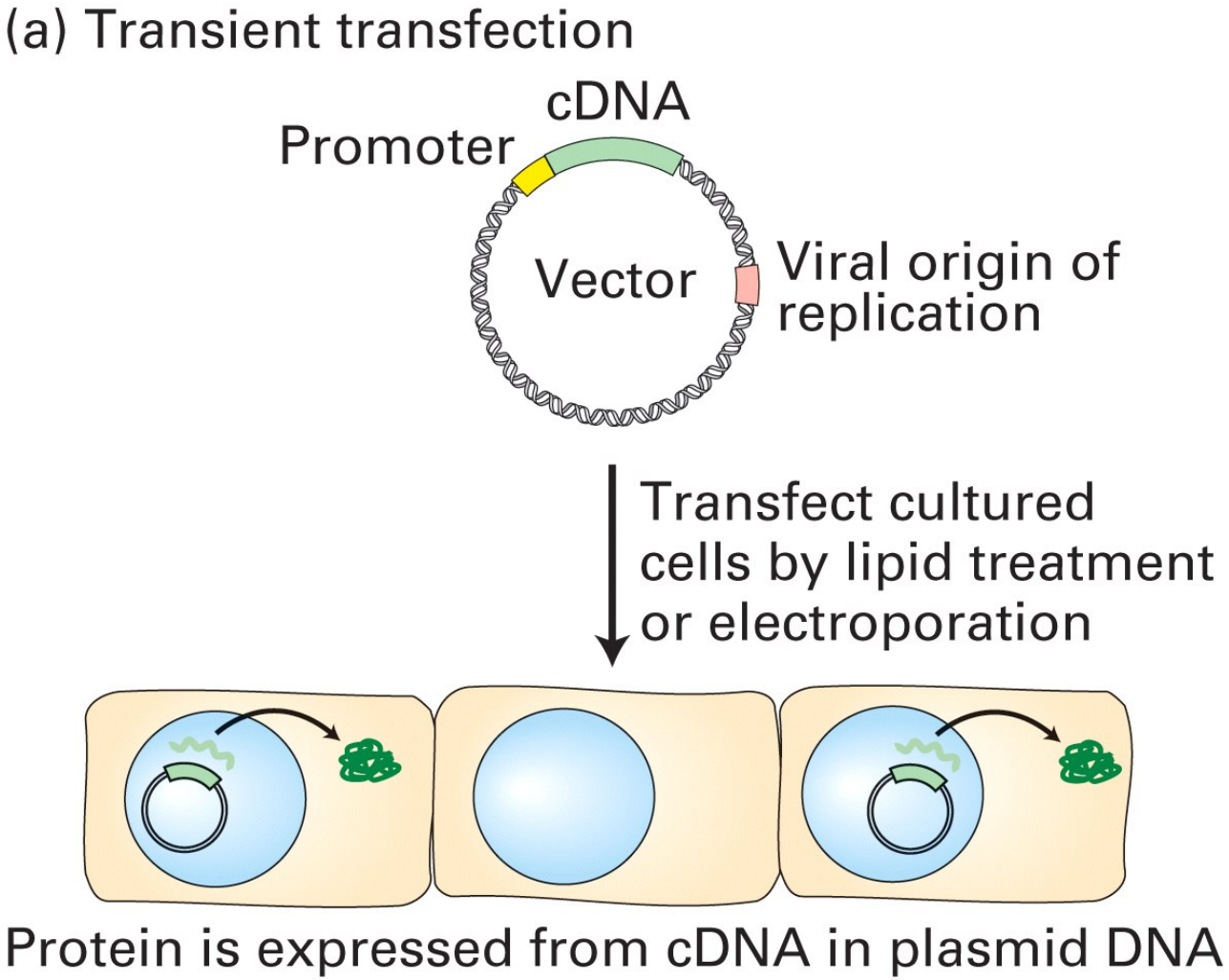
General expression pattern of control elements that regulate gene expression in multicellular eukaryotes and yeast.

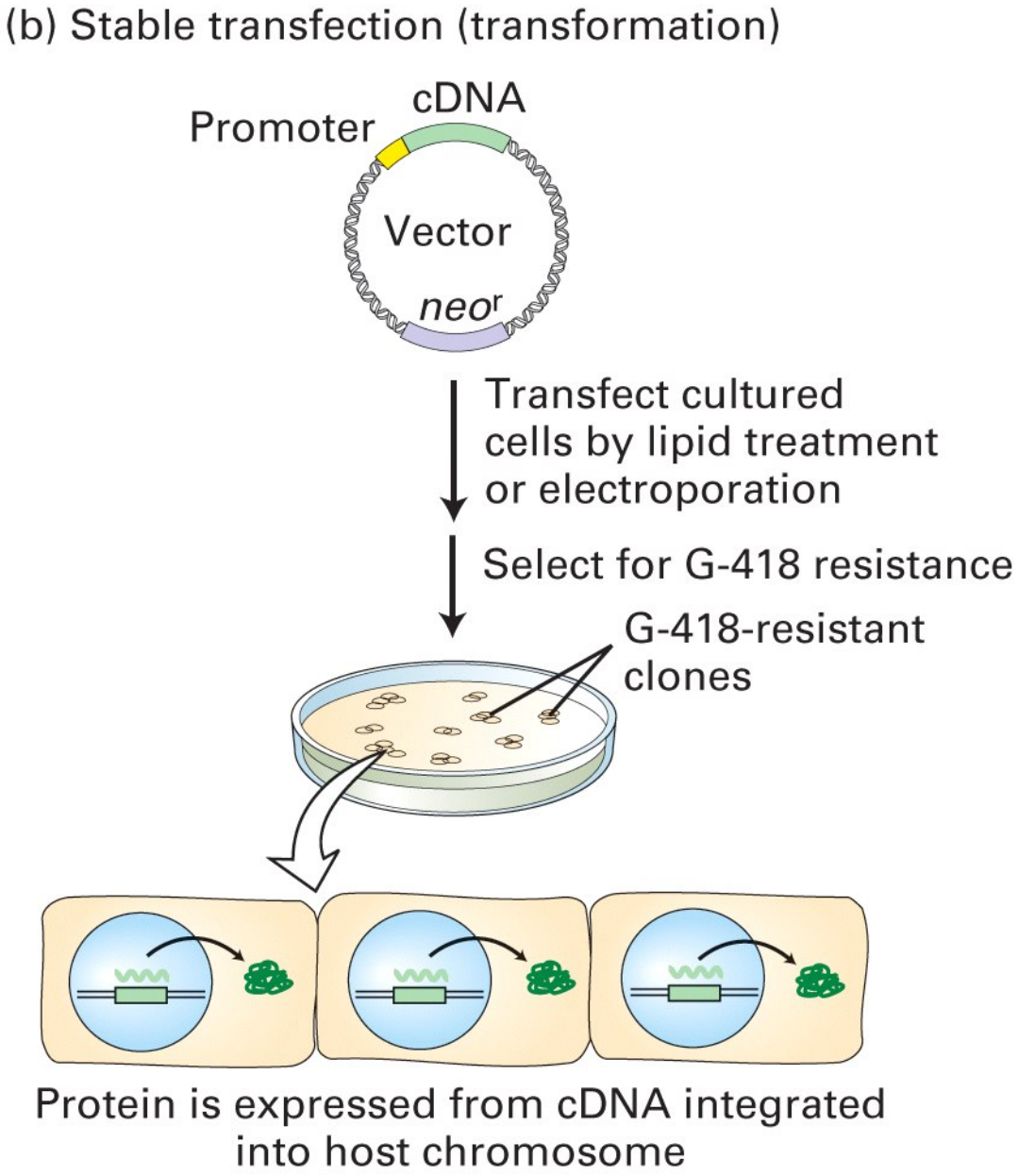


In vivo transfection assay measures transcription activity to evaluate proteins believed to be transcription factors



Transient and stable transfection with specially designed plasmid vectors permit expression of cloned genes in cultured animal cells



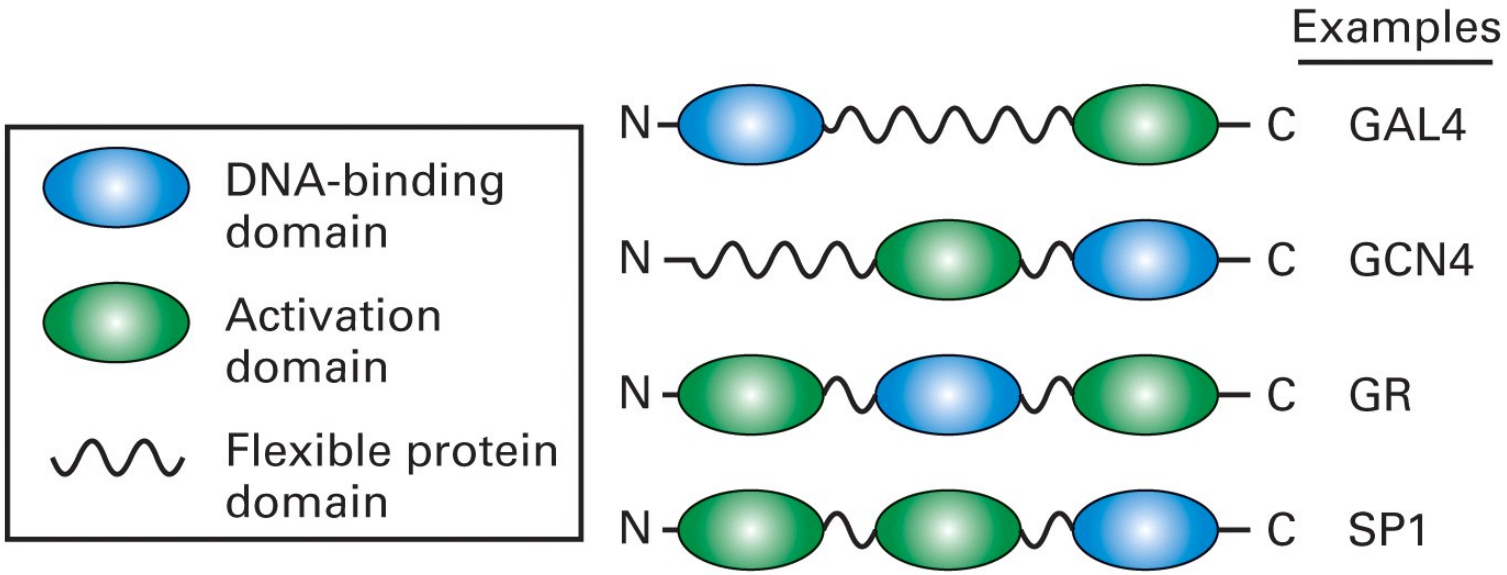


A commonly used selectable marker is the gene for **neomycin phosphotransferase** or aminoglycoside 3'-phosphotransferase gene (neo^r). The gene is of E.coli origin. The selectable drug is **G418, an aminoglycosidic antibiotic that inhibits protein synthesis**. **The neo gene allows production of the neomycin phosphotransferase which inactivates G418 by phosphorylation**

•**Puromycin resistant gene** (puro): Puromycin is an aminonucleoside antibiotic produced by *Streptomyces alboniger*. It specifically inhibits peptidyl transfer on ribosomes, therefore it inhibits the growth of various insect and animal cells. The expression of the *pac* gene (puromycin N-acetyltransferase from S. alboninger) confers puromycin resistance to transfected cells.

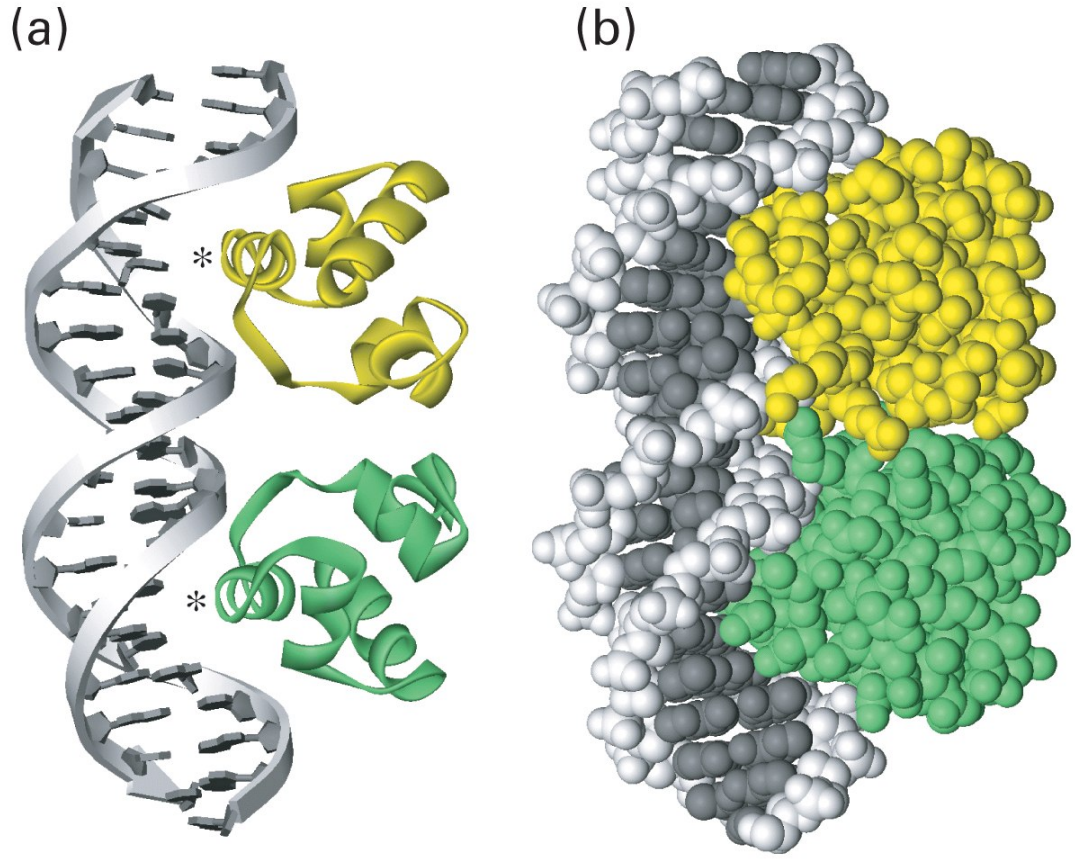
•**The hygromycin resistance gene** (hygro): Hygromycin B is an aminoglycosidic antibiotic produced by Stretomyces hygroscopicus. It is used for the selection and maintenance of prokaryotic and eukaryotic cells transfected with the hygromycin resistance gene of E. coli origin. Hygormycin B kills cells by inhibiting protein synthesis. The resistance gene encodes for a kinase that inactivates hygromycin B through phosphorylation.

Schematic diagrams illustrating the modular structure of eukaryotic transcription factors.



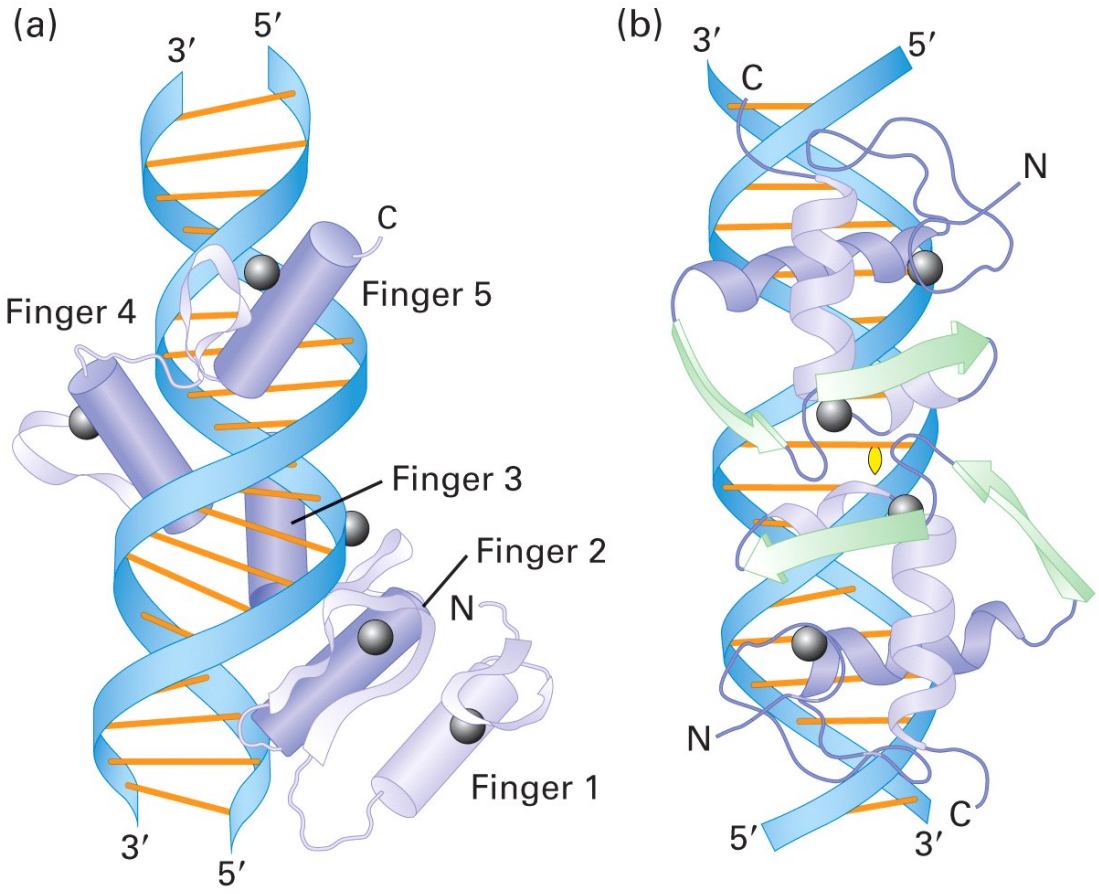
DNA-binding domains can be classified into numerous structural types

Helix-turn-Helix



Interaction of bacteriophage 434 represssor with DNA

Zinc finger Transcription factors



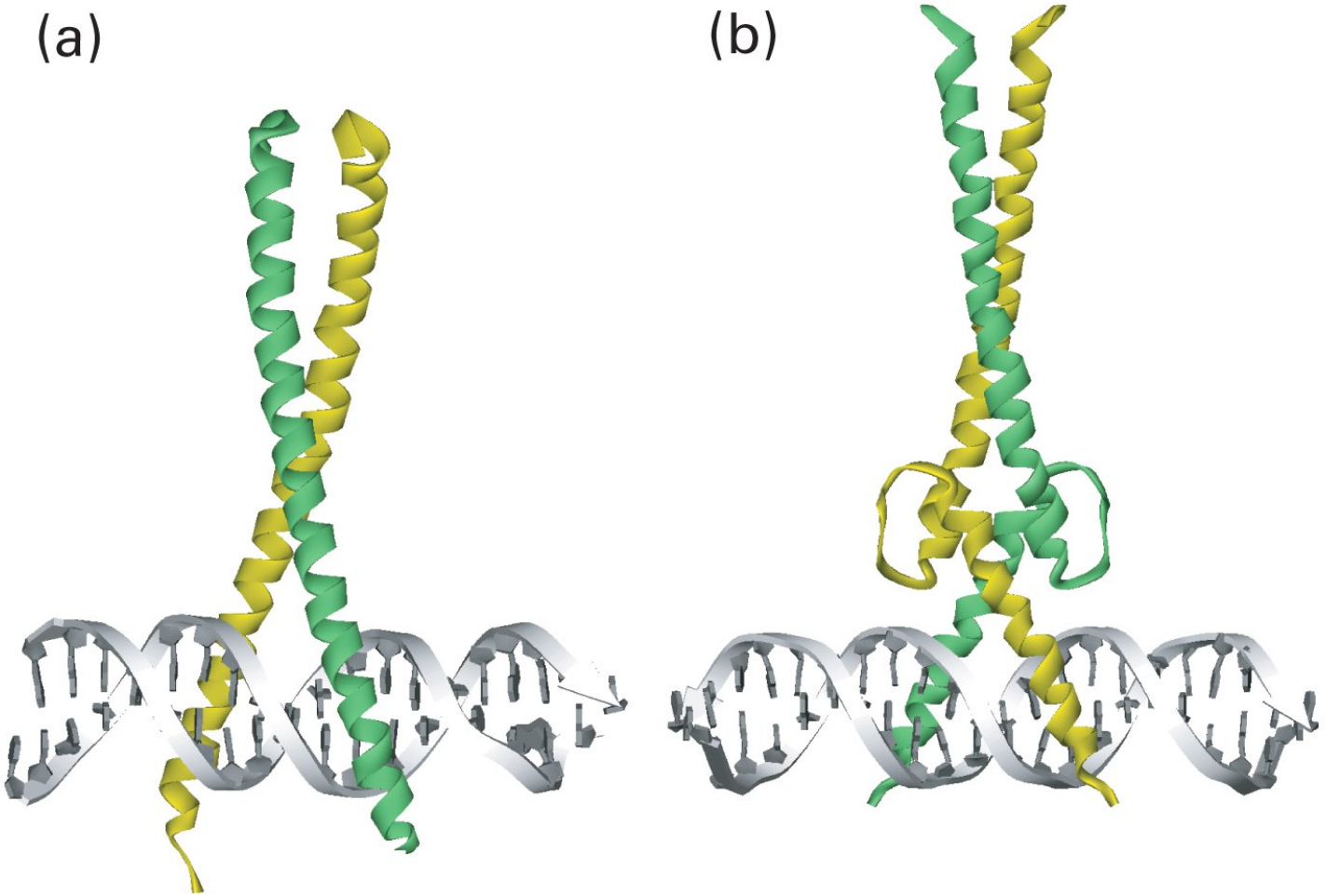
Two types of Znc finger proteins:

C2H2 zinc fingers bind as monomer.

C4 zinc fingers bind as homo-dimer (e.g.steroid receptor, such as glucocorticoid receptor)

Leucine Zipper

Basic Helix-Loop-Helix (bHLH)



Leucine Zipper: contains the hydrophobic amino acid leucine at every 7th position in the sequence

Characterizing TF binding sites

• TFs locate their target genes by binding short DNA elements, “binding sites”

• These binding sites are degenerate:

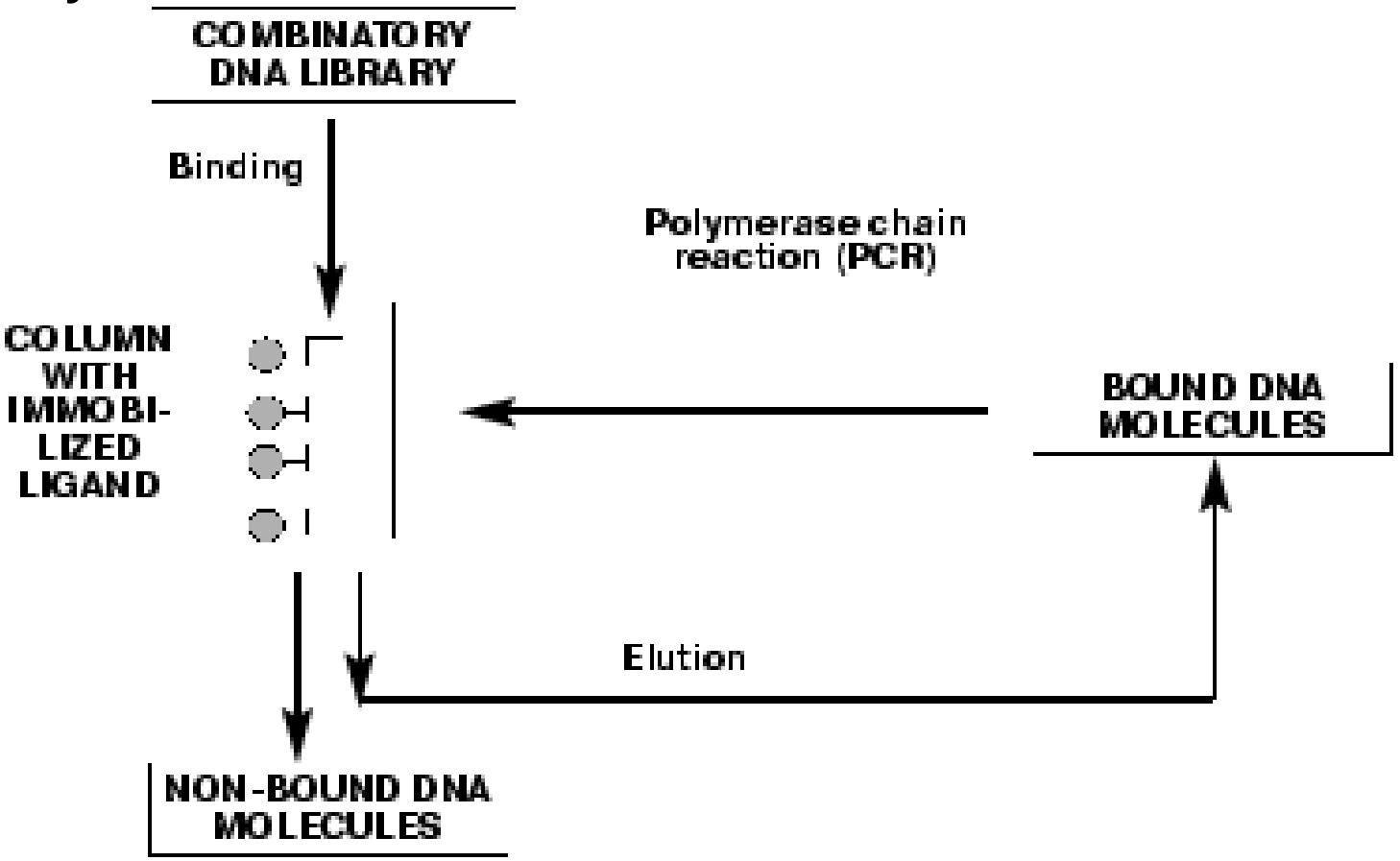
GCAGGTGAC
CCAGGTCTG
CCAGGTGTC
CCAGGTCAC
CCAGGTCAG

• How to identify these binding sites?

Characterizing TF binding sites

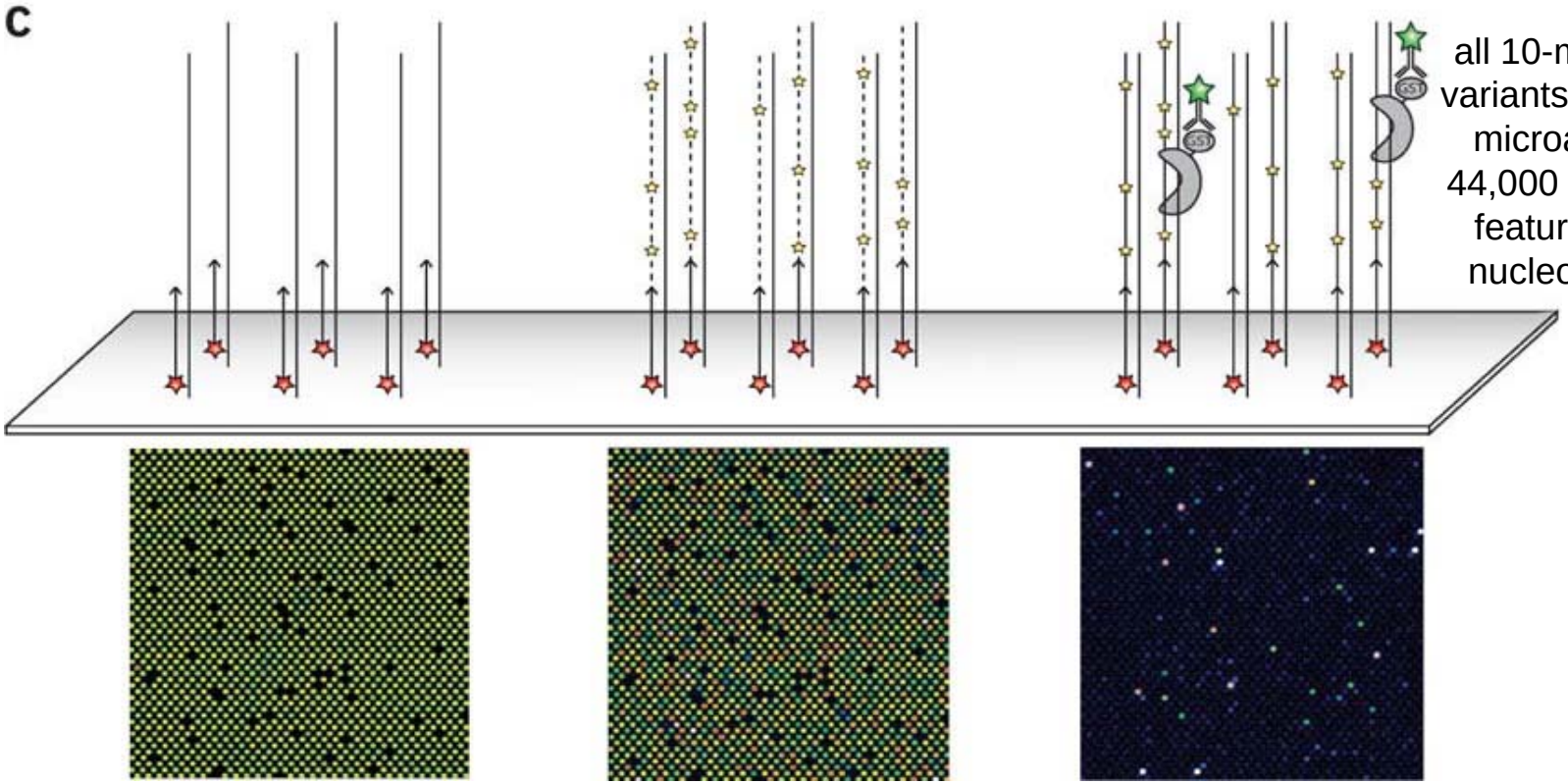
SELEX

"systematic evolution of ligands by exponential enrichment assay"



Characterizing TF binding sites

PBM
"protein-binding micro-array"



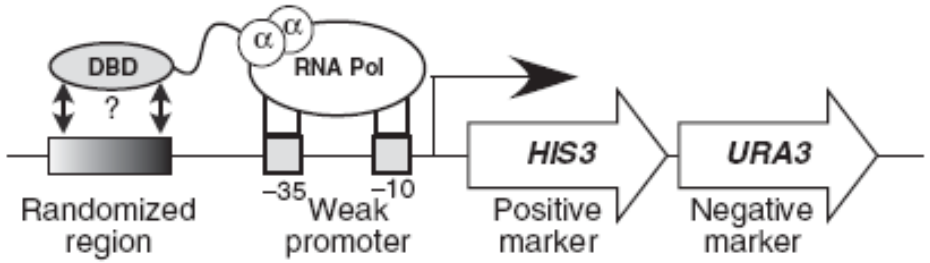
all 10-mer binding site variants, each universal microarray contains 44,000 single-stranded features that are 60 nucleotides (nt) long

Berger et al., Nat. Biotechnol., 2006: 24, 1429.

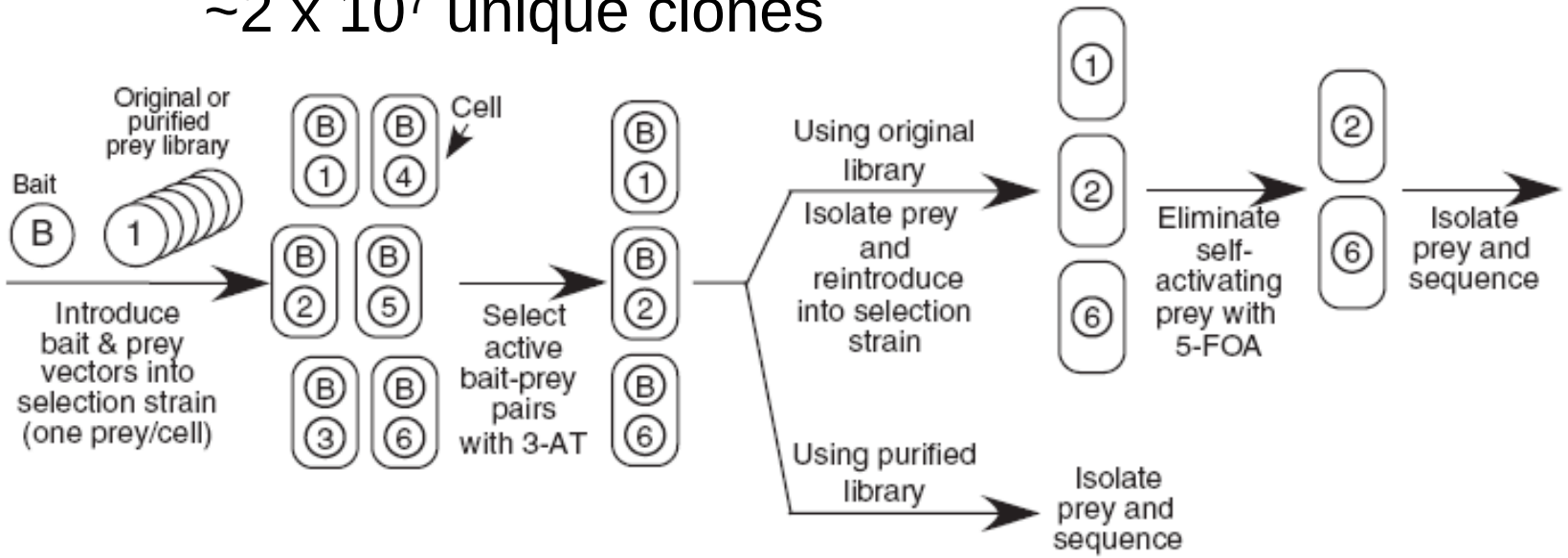
Characterizing TF network nodes: binding sites

B1H

"bacterial one-hybrid"



~2 x 10^7 unique clones

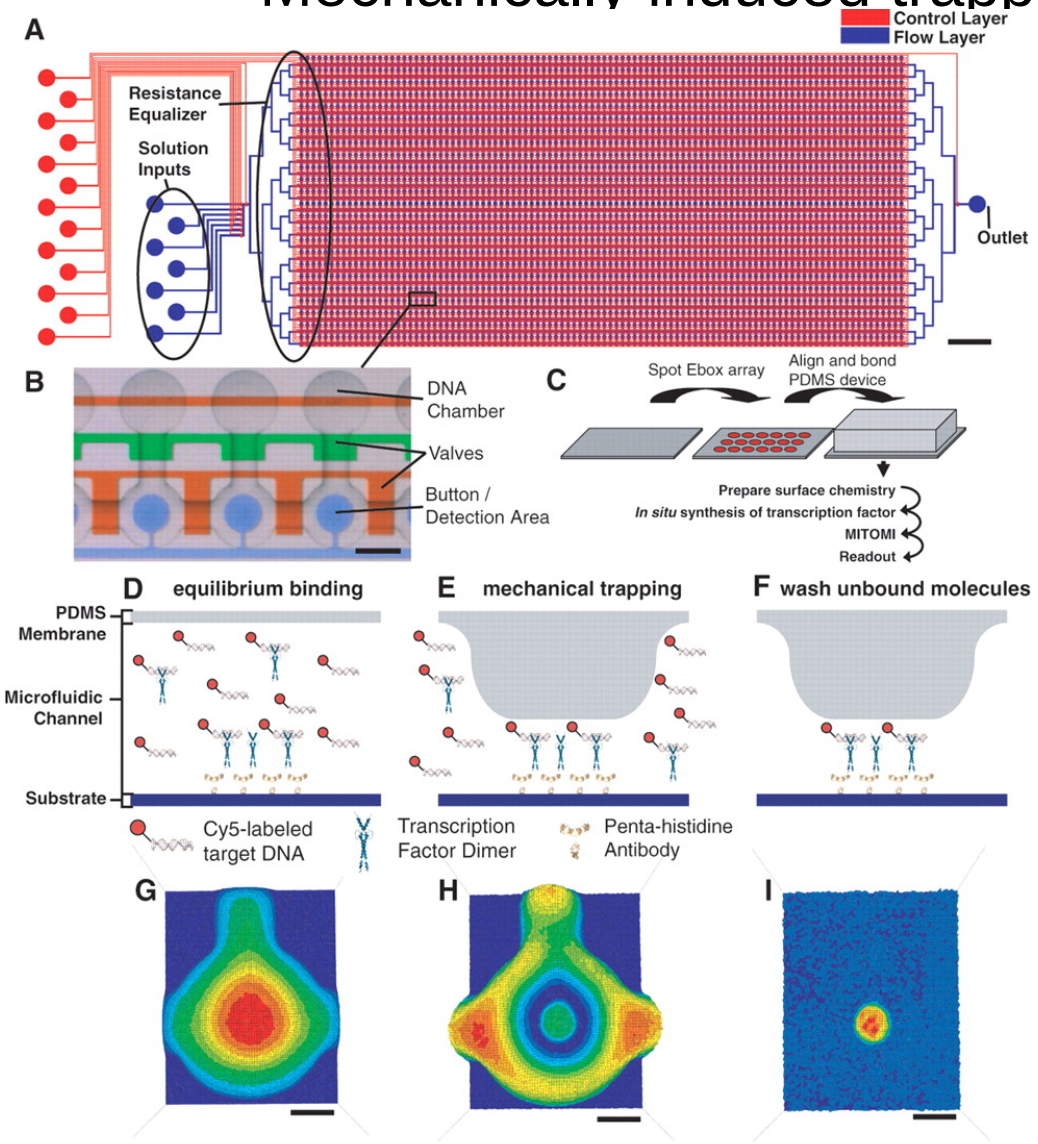


Meng et al., Nat. Biotechnol., 2005

Characterizing TF network nodes: binding sites

MITOMI

"Mechanically induced trapping of

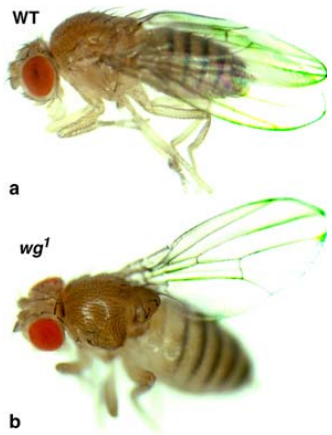


Maerkl and Quake, Science, 2007

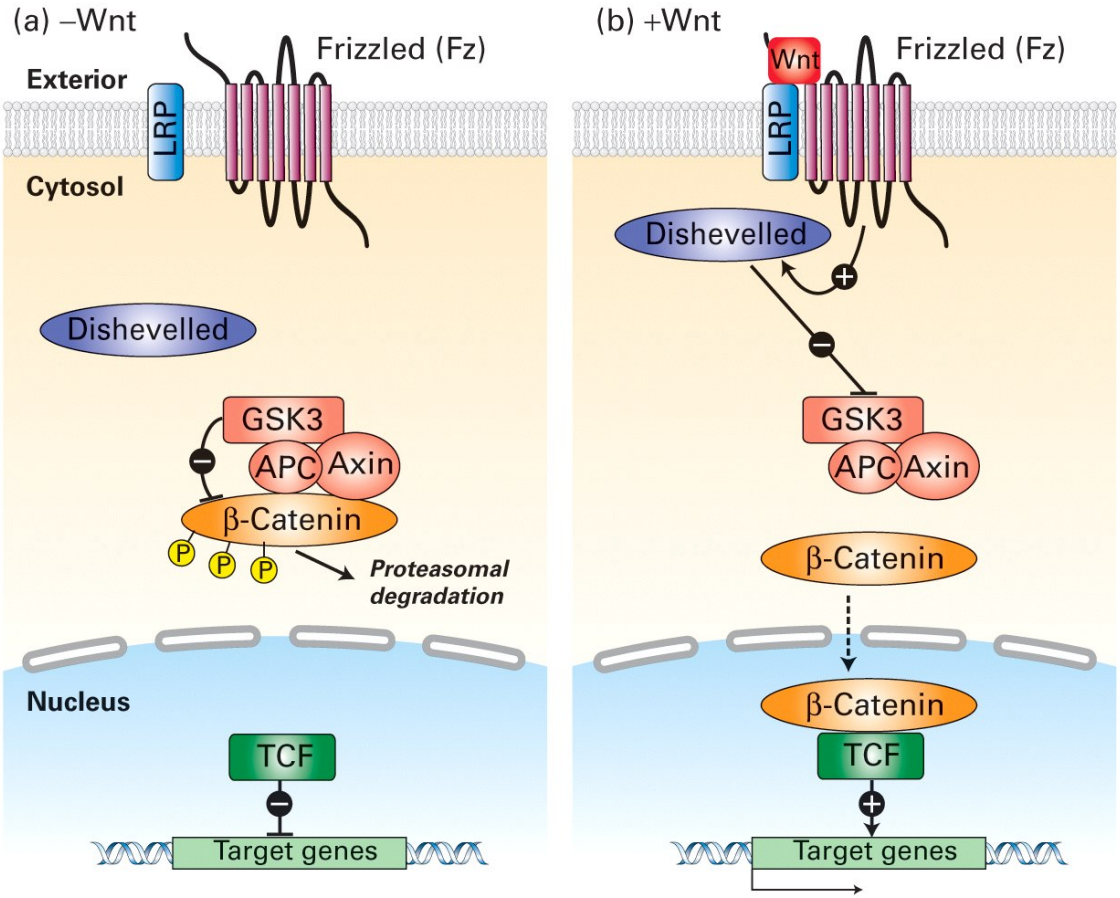
+:

- low-affinity, transient interactions can be detected through careful mechanical trapping instead of washing (eliminates the off-rate problem of array-based methods).
- absolute affinity values for protein-DNA interactions can be determined

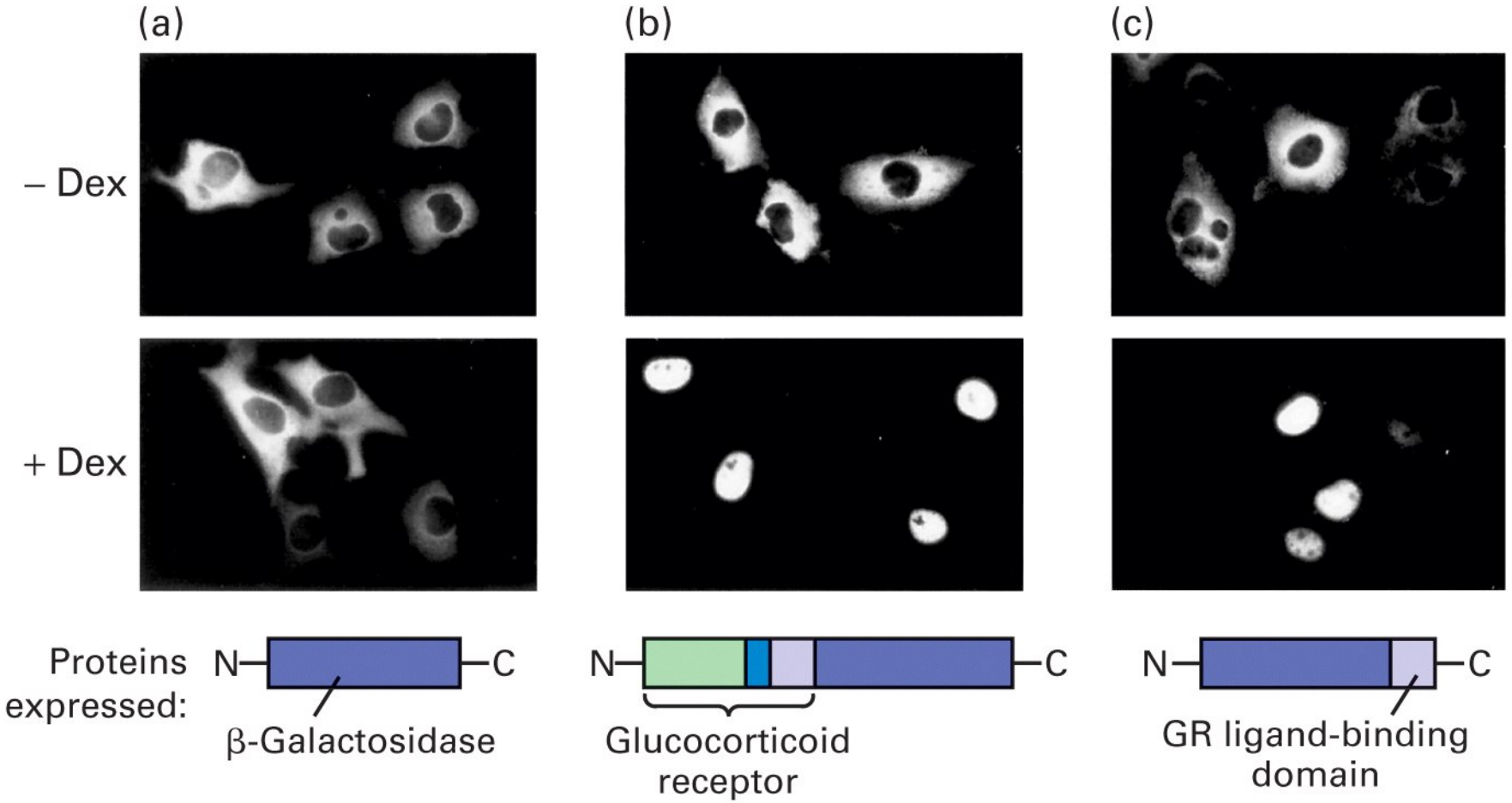
Regulation of Transcription Factor activity



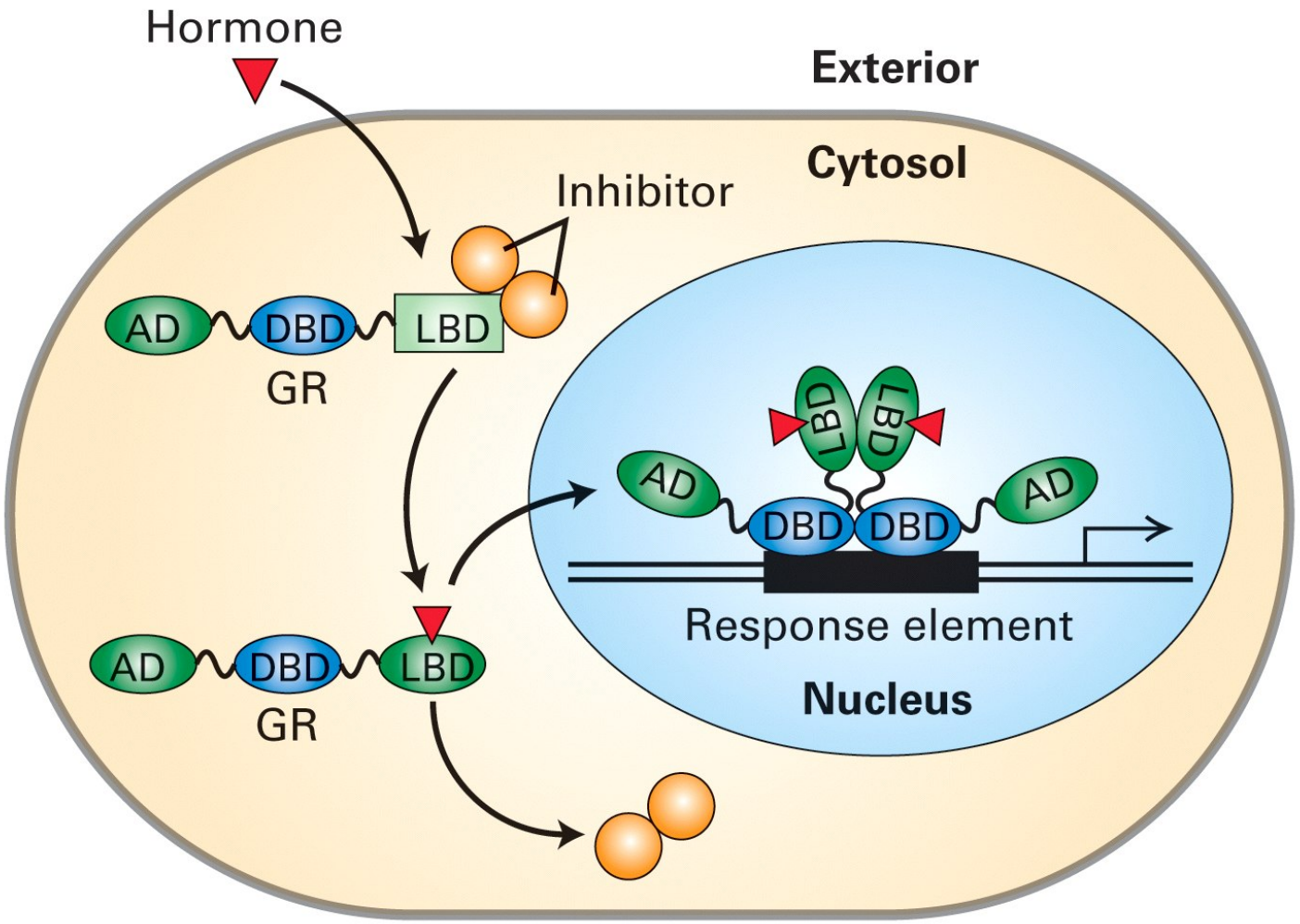
Wingless phenotype



Fusion proteins from expression vectors demonstrate that the hormone-binding domain of the glucacorticoid receptor mediates translocation to the nucleus in the presence of hormone.

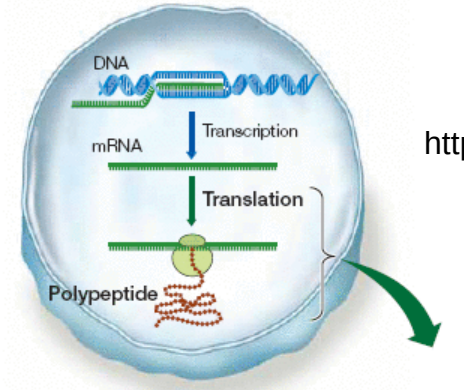


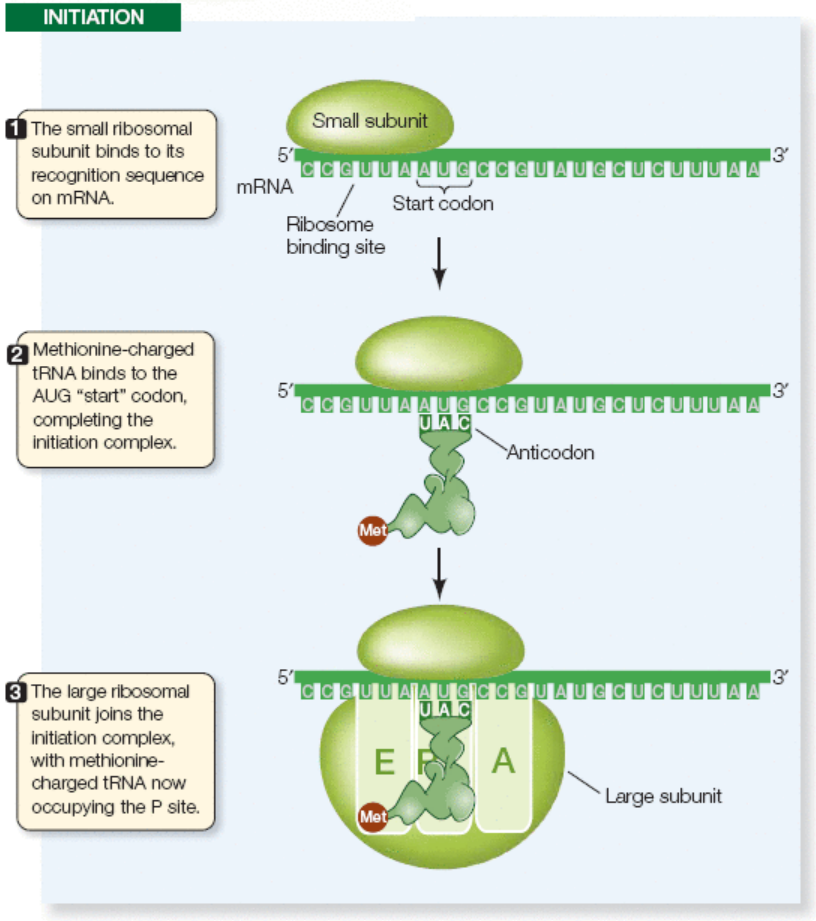
Model of hormone-dependent gene activation by a homodimeric nuclear receptor



Translation initiation

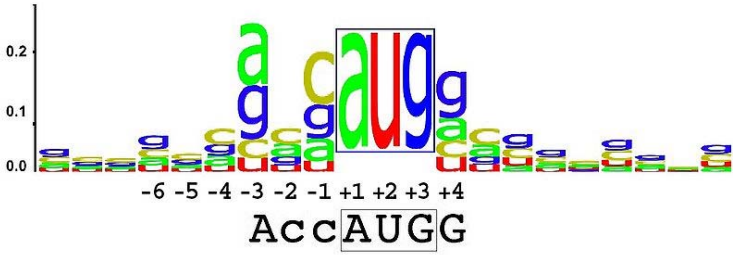
http://media.pearsoncmg.com/bc/bc_martini_ap_slim/assets/animations/ch14_translation.html





Shine-Dalgarno: $GGAGG(X)_9AUG$

Kozak in vertebrates

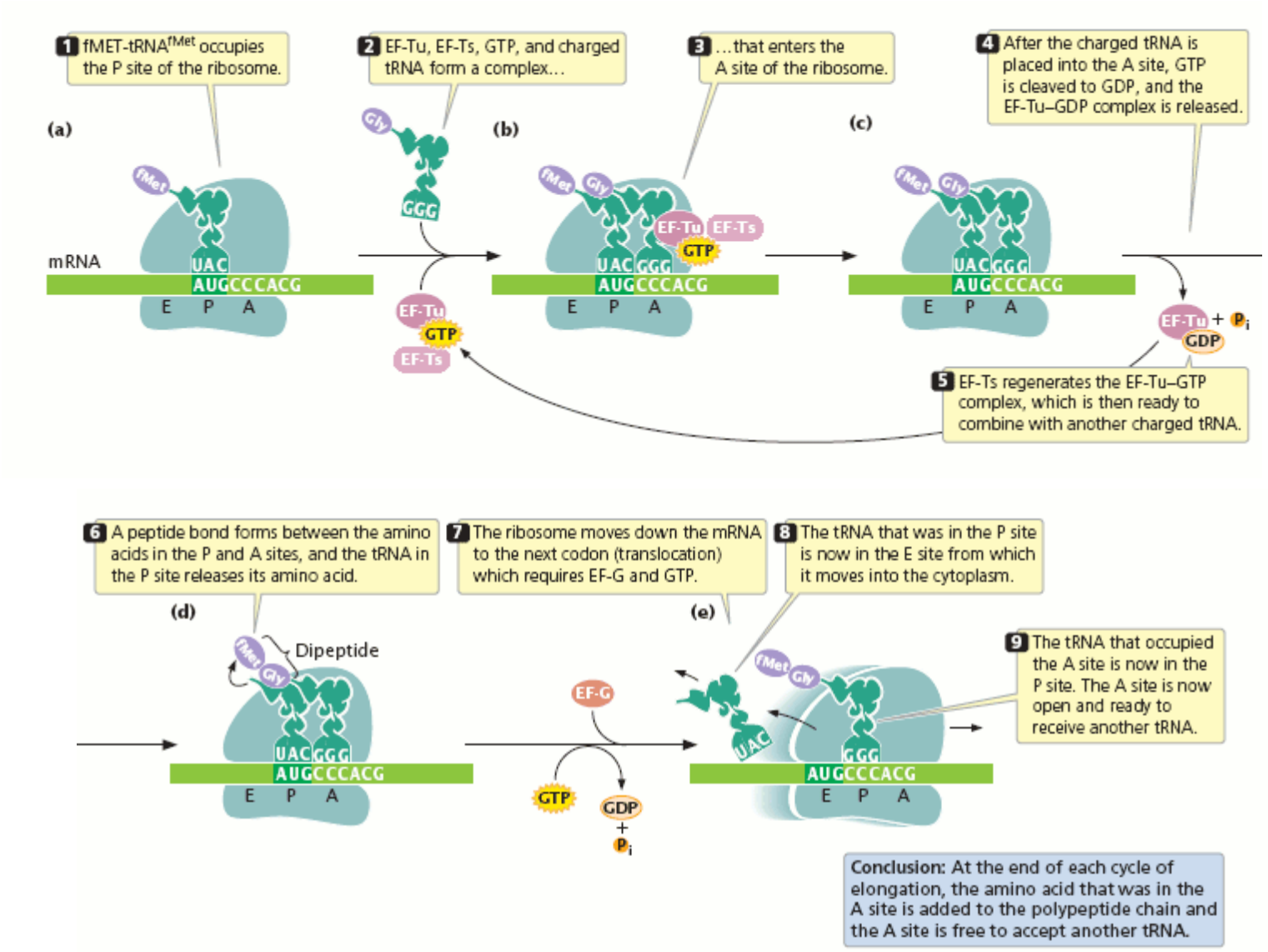


The A (amino acid) site is the location at which the aminoacyl-tRNA anticodon base pairs up with the mRNA codon, ensuring that correct amino acid is added to the growing polypeptide chain.

The P (polypeptide) site is the location at which the amino acid is transferred from its tRNA to the growing polypeptide chain.
(Initiator methionine tRNA is the only aminoacyl-tRNA that can bind in the P site of the ribosome)

The E (exit) site is the location at which the "empty" tRNA sits before being released back into the cytoplasm to bind another amino acid and repeat the process.

Translation elongation



Termination of Translation

Three termination codons: UAA, UAG, and UGA.

In the place of these tRNAs, release factors bind and facilitate release of the mRNA from the ribosome

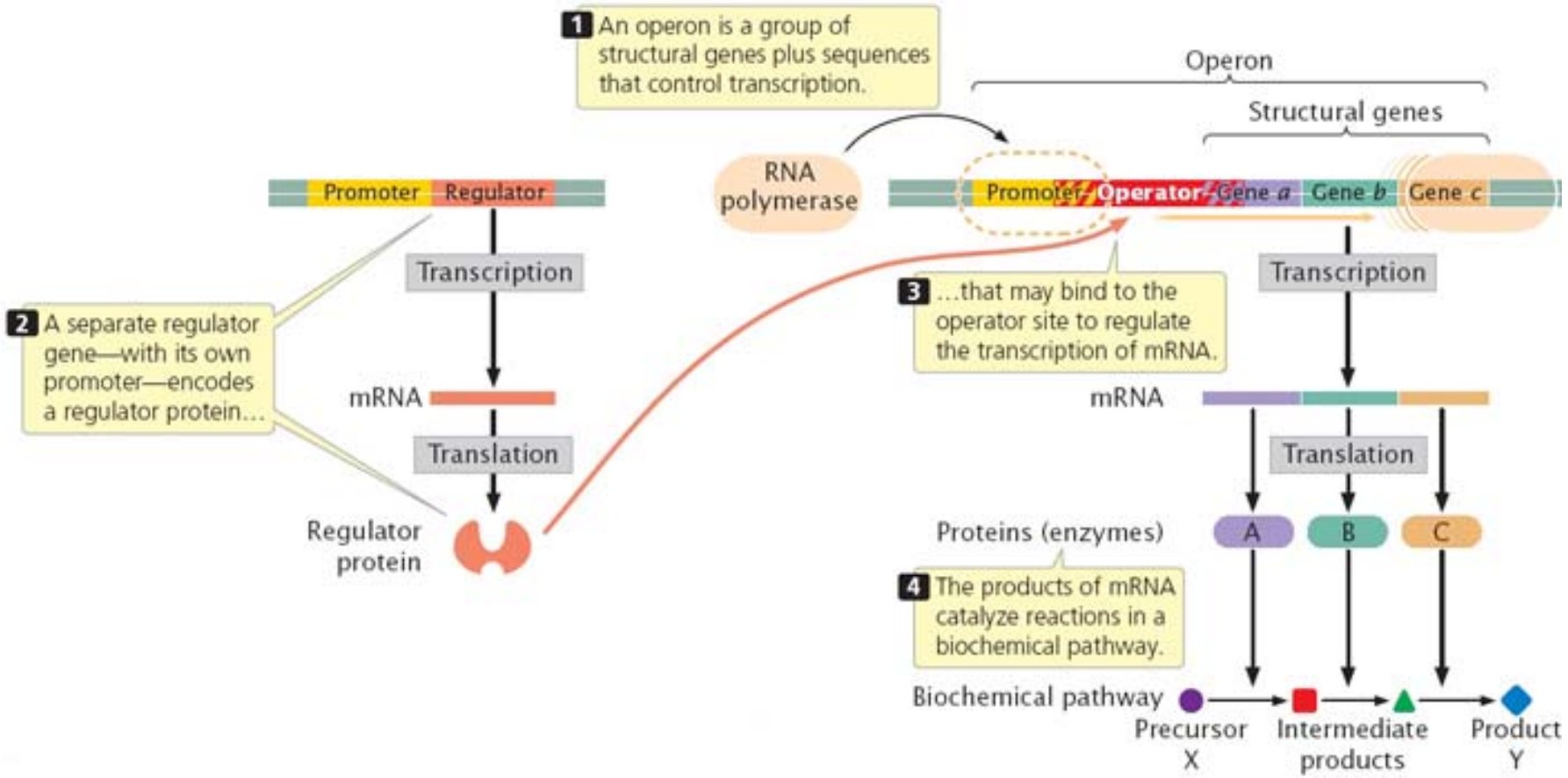
Comparing Eukaryotic and Prokaryotic Translation

The translation process is very similar in prokaryotes and eukaryotes.

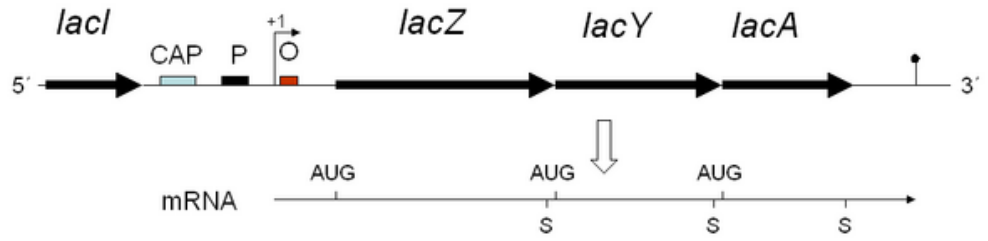
In bacteria, transcription and translation take place simultaneously, and mRNAs are relatively short-lived.

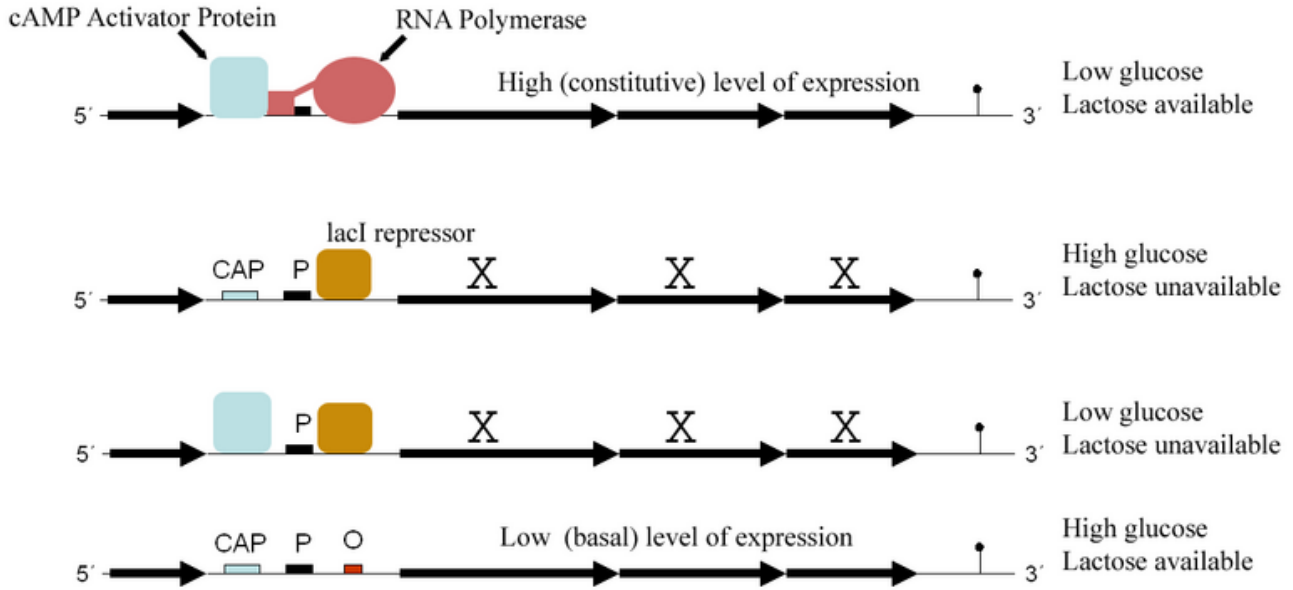
In eukaryotes, mRNAs have highly variable half-lives, are subject to modifications, and must exit the nucleus to be translated, → multiple regulatory levels of protein production

An operon is a single transcriptional unit that includes a series of structural genes, a promoter, and an operator.



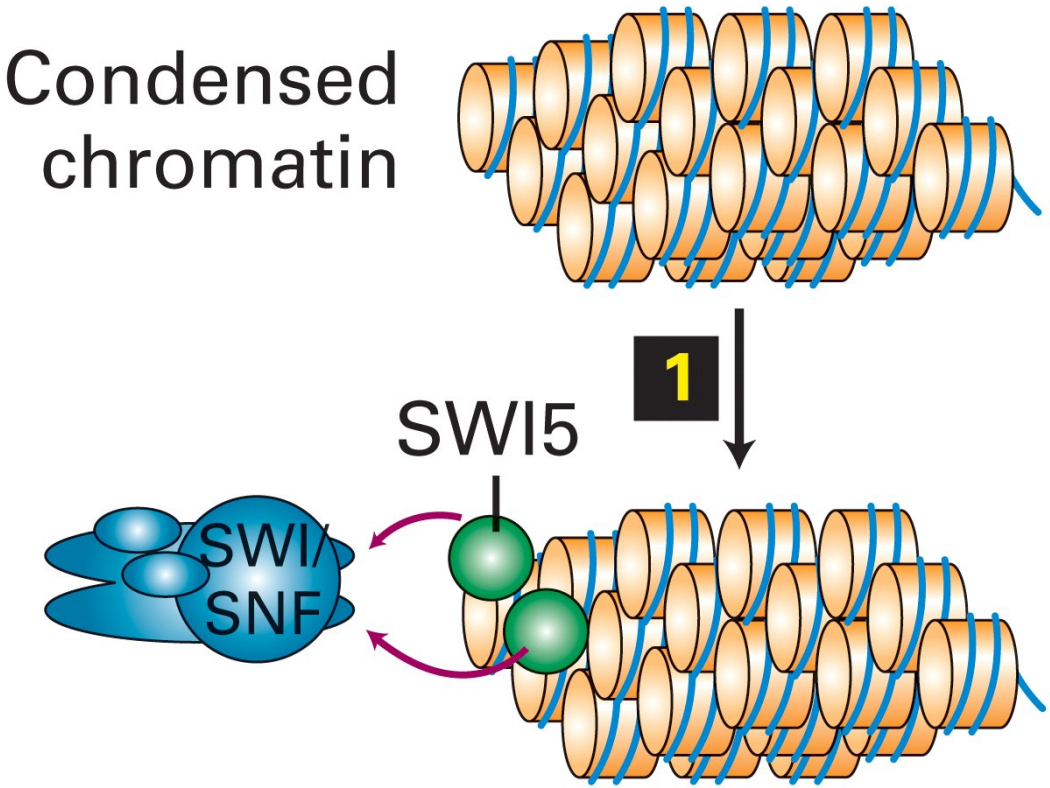
The *lac* Operon and its Control Elements





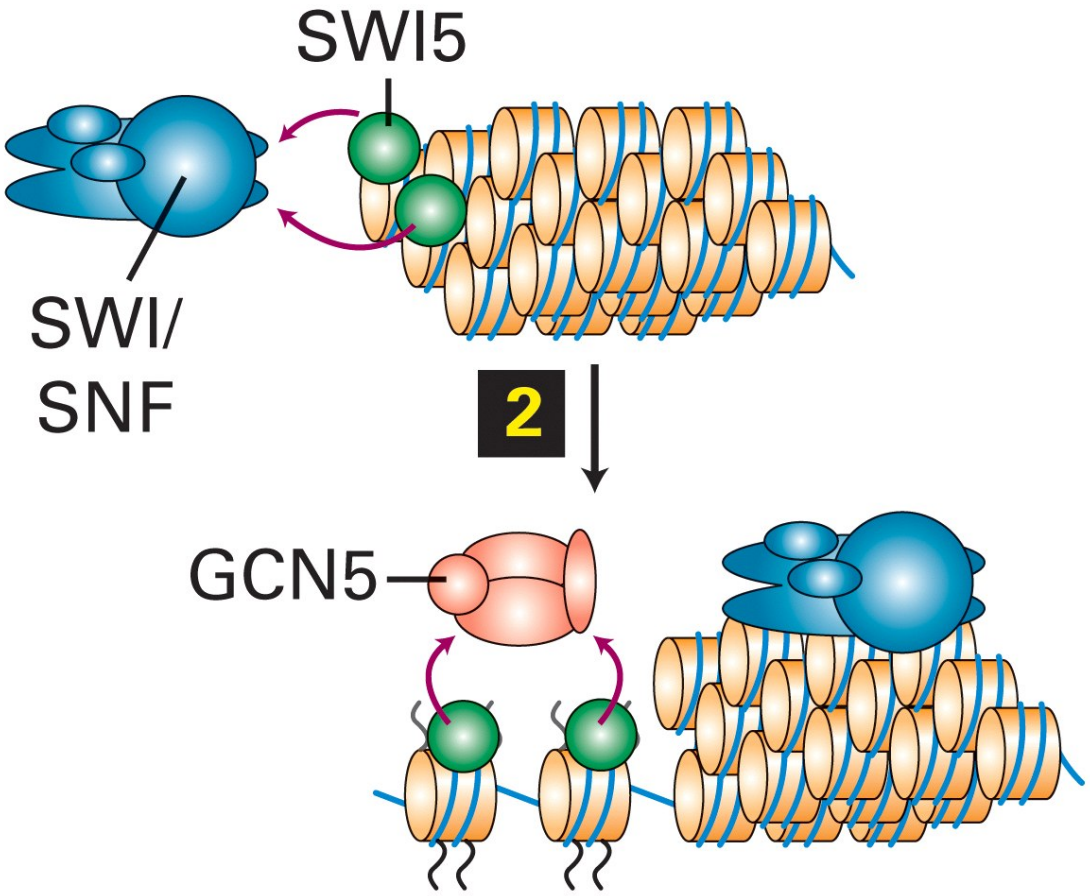
Transcription of many genes requires ordered binding of activators and action of coactivators (e.g. HO gene, yeast)

http://media.pearsoncmg.com/bc/bc_martini_ap_slim/assets/animations/ch17_HO_gene.html

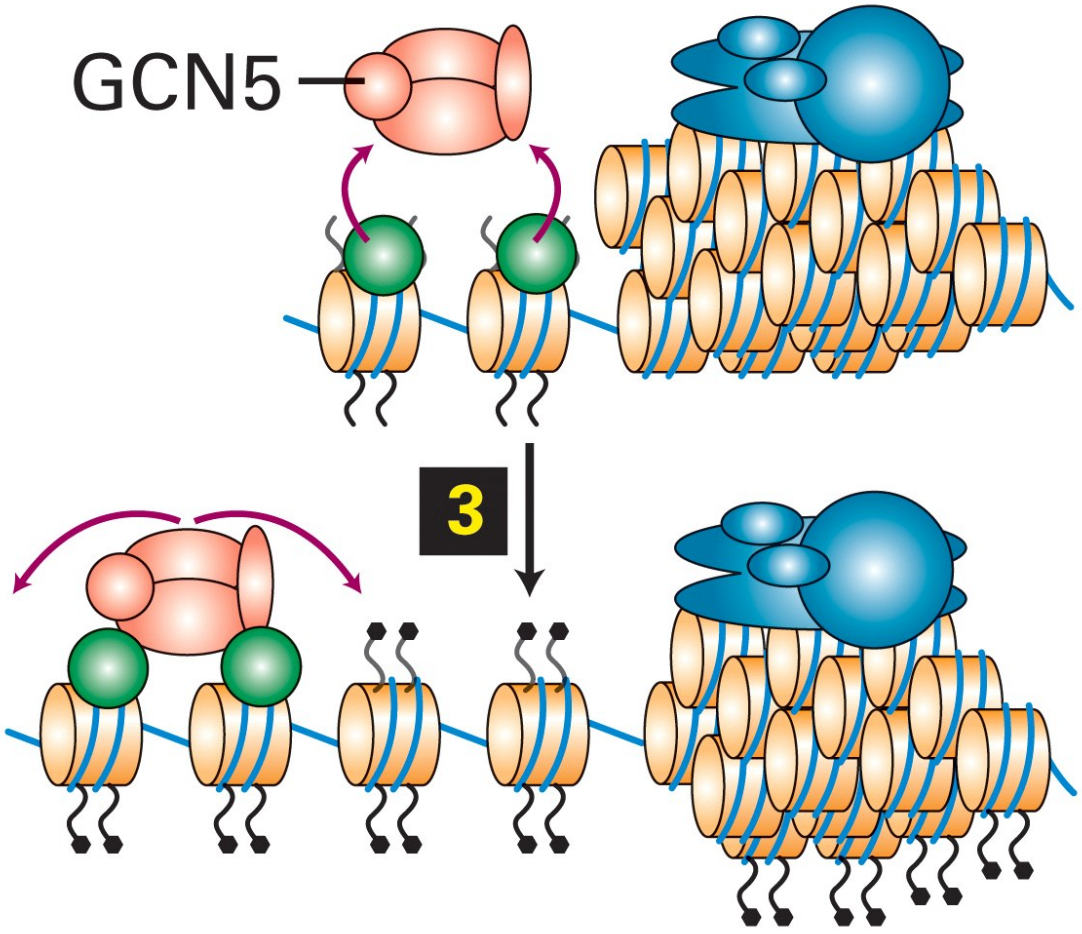


1. Initially, the HO gene is packed into condensed chromatin. Activation begins when the SWI5 activator binds to enhancer sites 1200-1400 base pairs upstream of the start site and interacts with the SWI/SNF chromatin remodeling complex

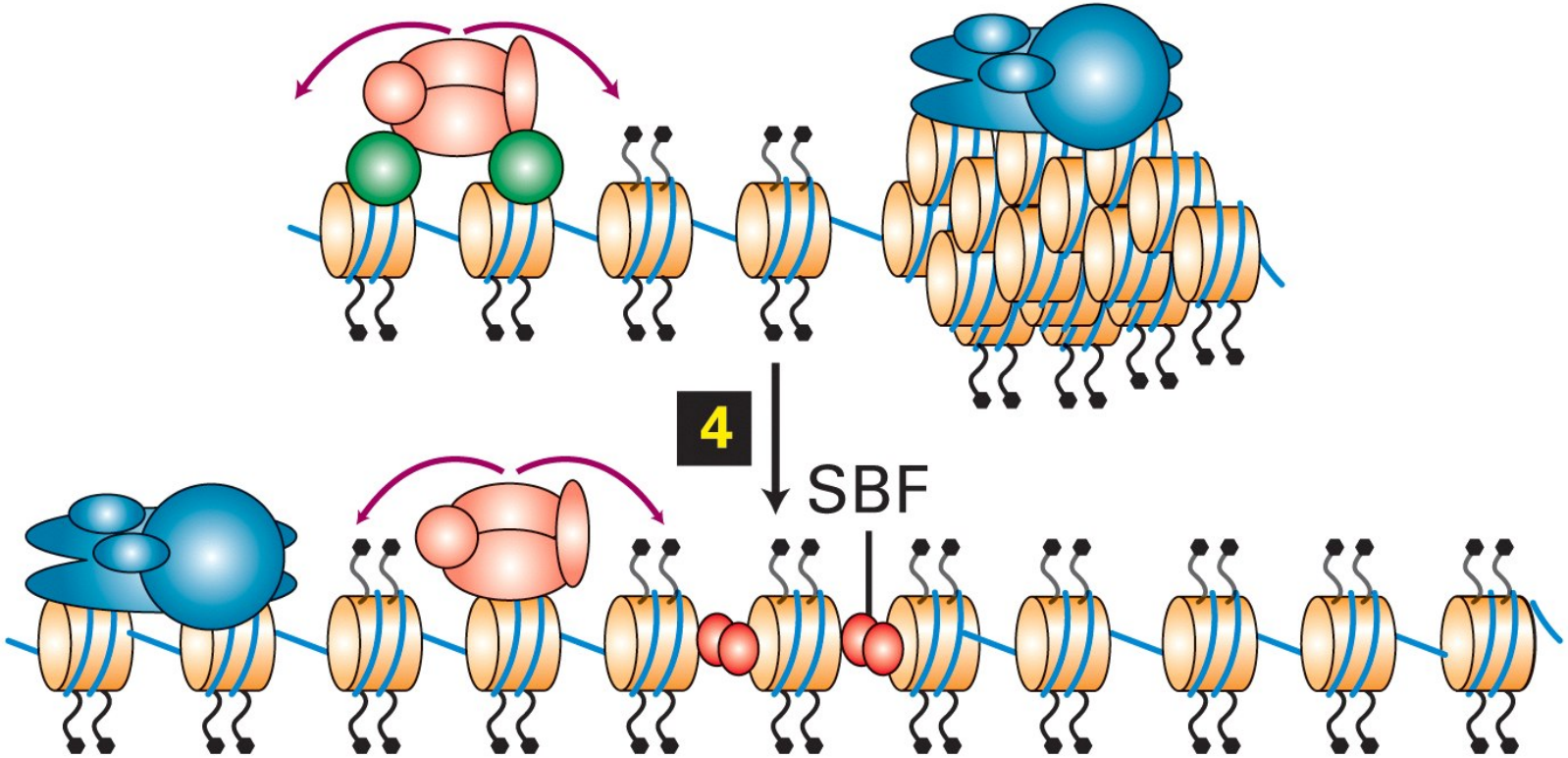
2. The SWI/SNF complex acts to decondense the chromatin, thereby exposing histone tails.



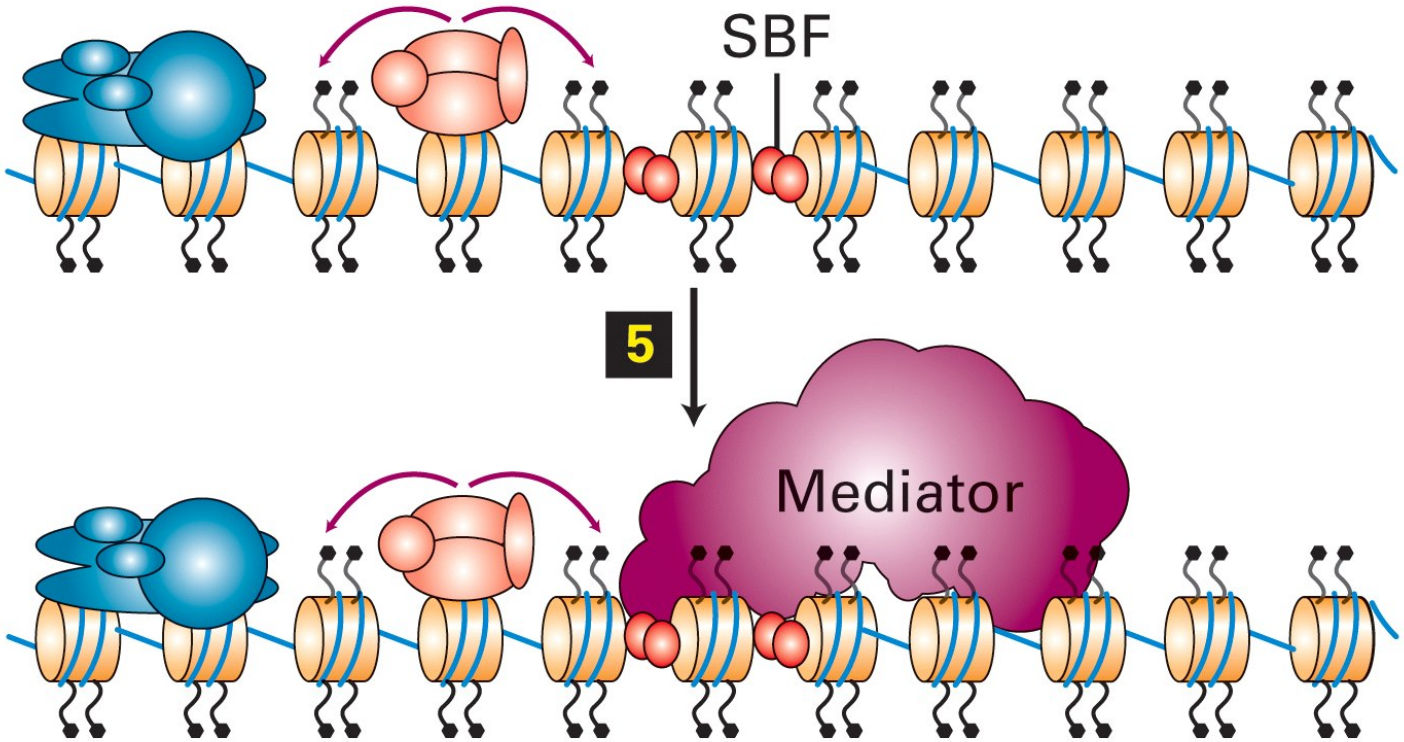
3. A GCN5-containing histone acetylase complex associates with bound SWI5 and acetylates histone tails in the HO locus as SWIF/SNF continues to decondense adjacent chromatin.



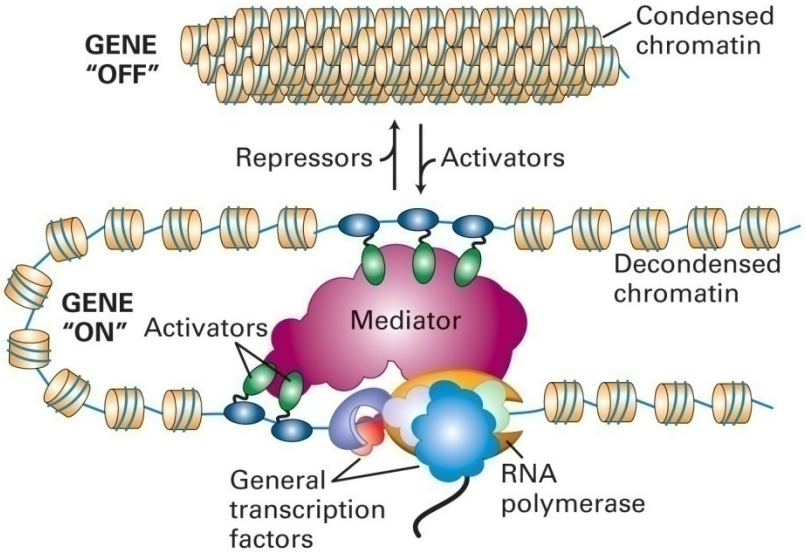
4. SWI5 is released from the DNA, but the SWI/SNF and GCN5 complexes remain associated with the HO control region. Their action allows the SBF activator to bind to several sites in the promoter proximal region.

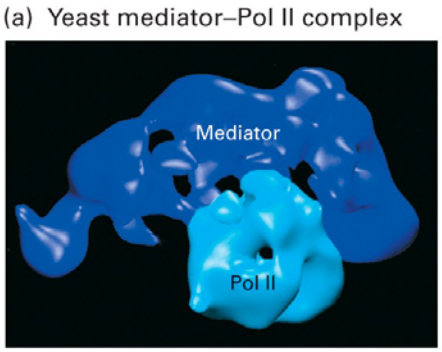


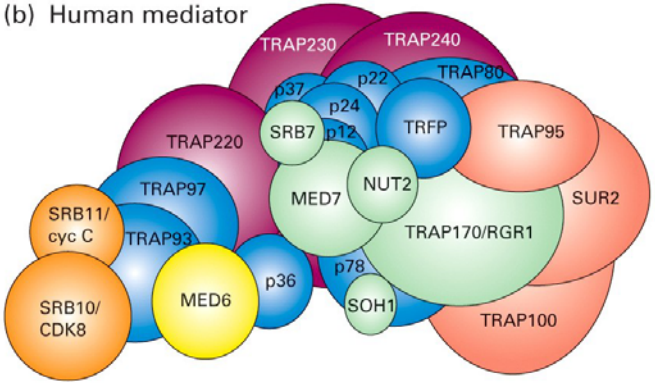
5. SBF then binds the mediator complex



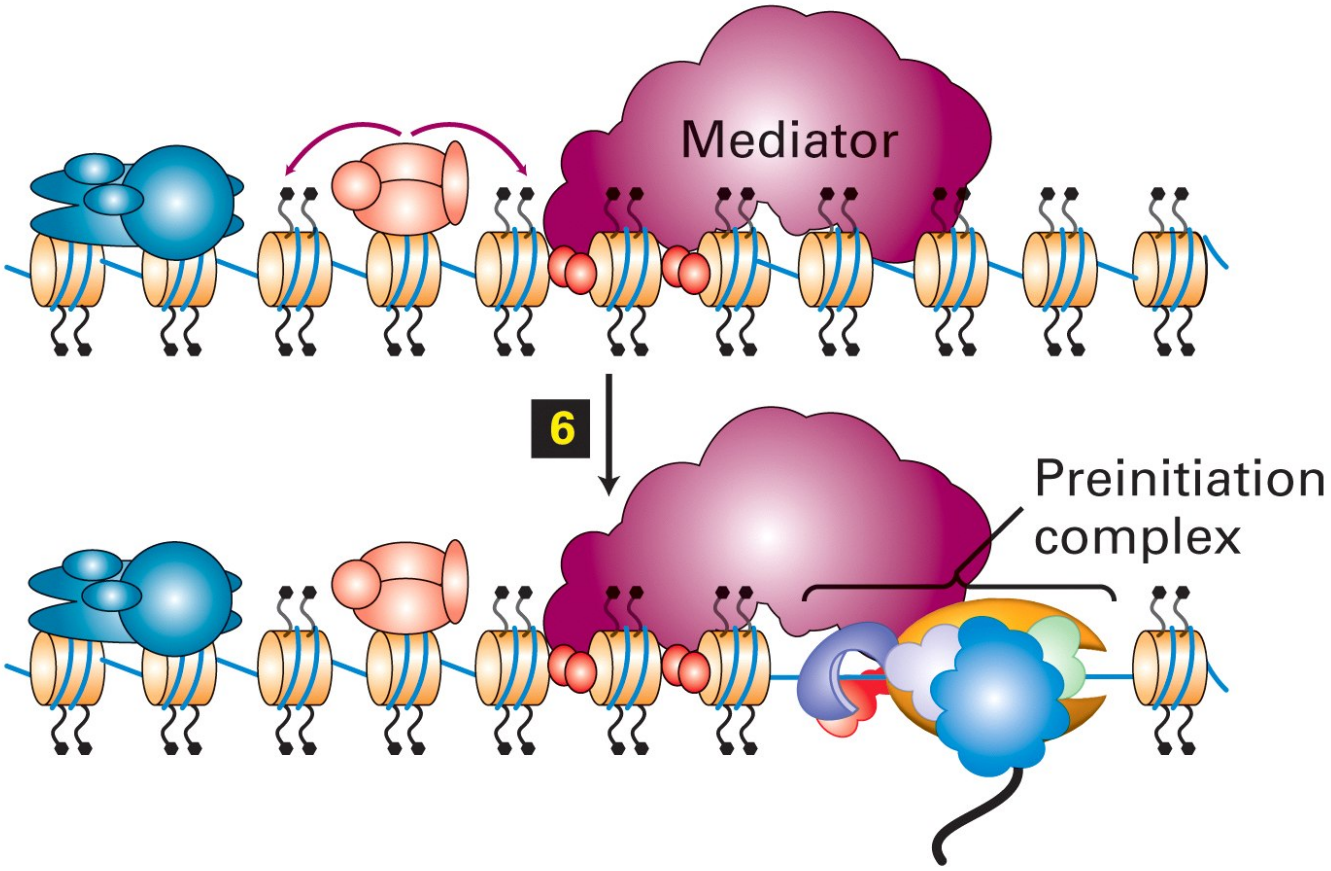
Mediator complexes are multiprotein complexes which form a molecular bridge between activation domains and Pol II





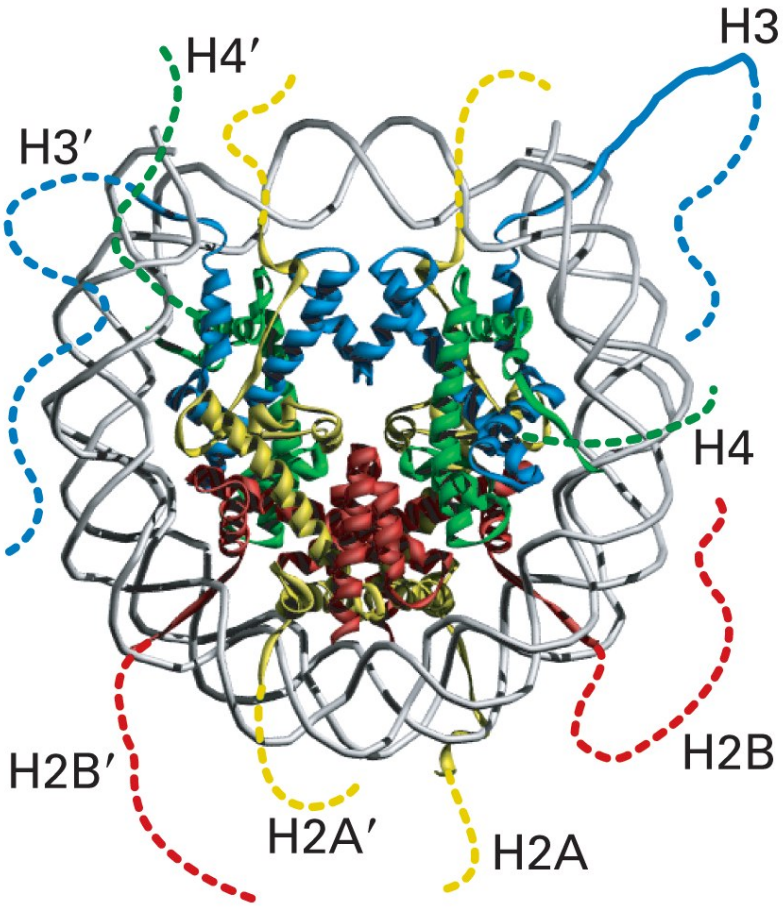


6. Subsequent binding of Pol II and general transcription factors results in the assembly of a transcription preinitiation complex.

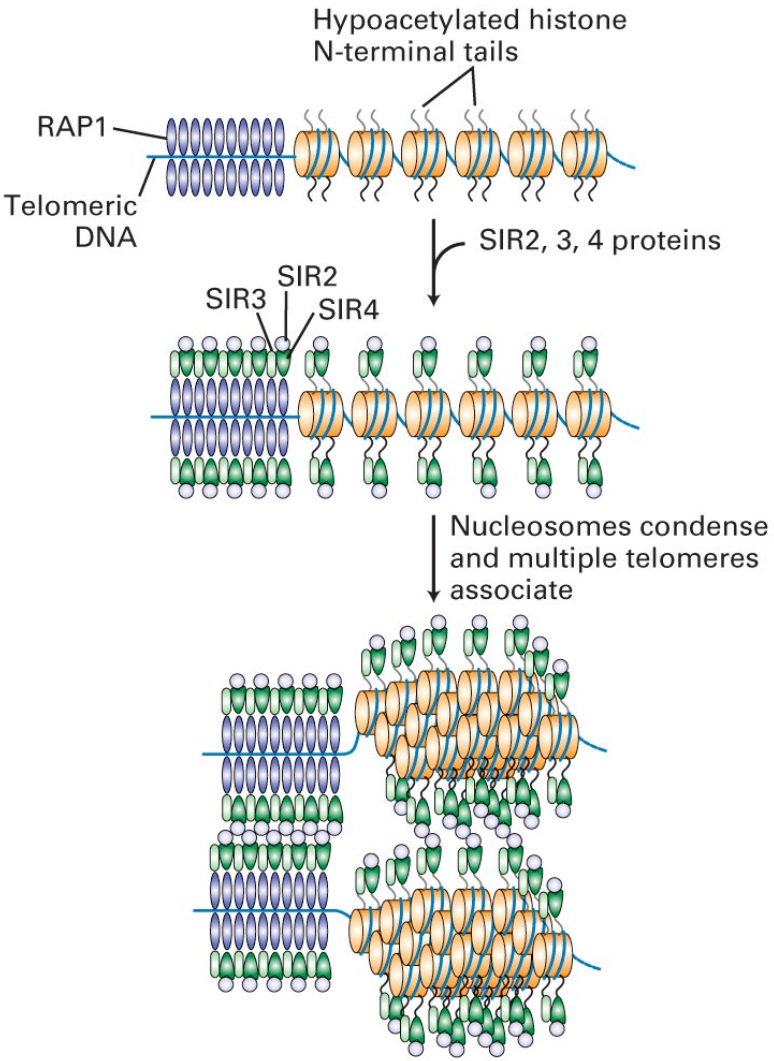


Residues within the N-terminal region of each histones, and the C-terminal region of hsitone 2A, called histone tails extend from the surface of the nucleosome and can be reversibly modified. Such modifications, espacially the acetylation of Histone H3 and H4 tails, influence the relative condensation of chromatin and thus its accessibility to proteins required for transcription initiation

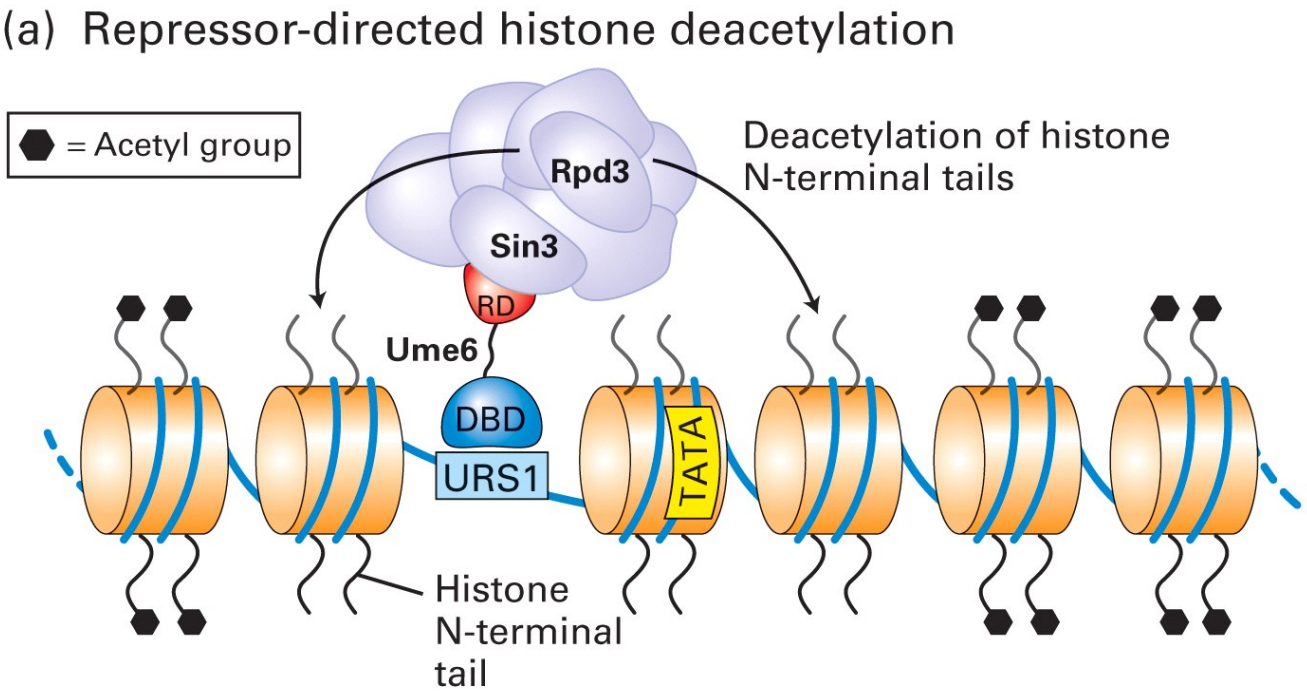
Ribbon diagram of the histones showing histone tails.



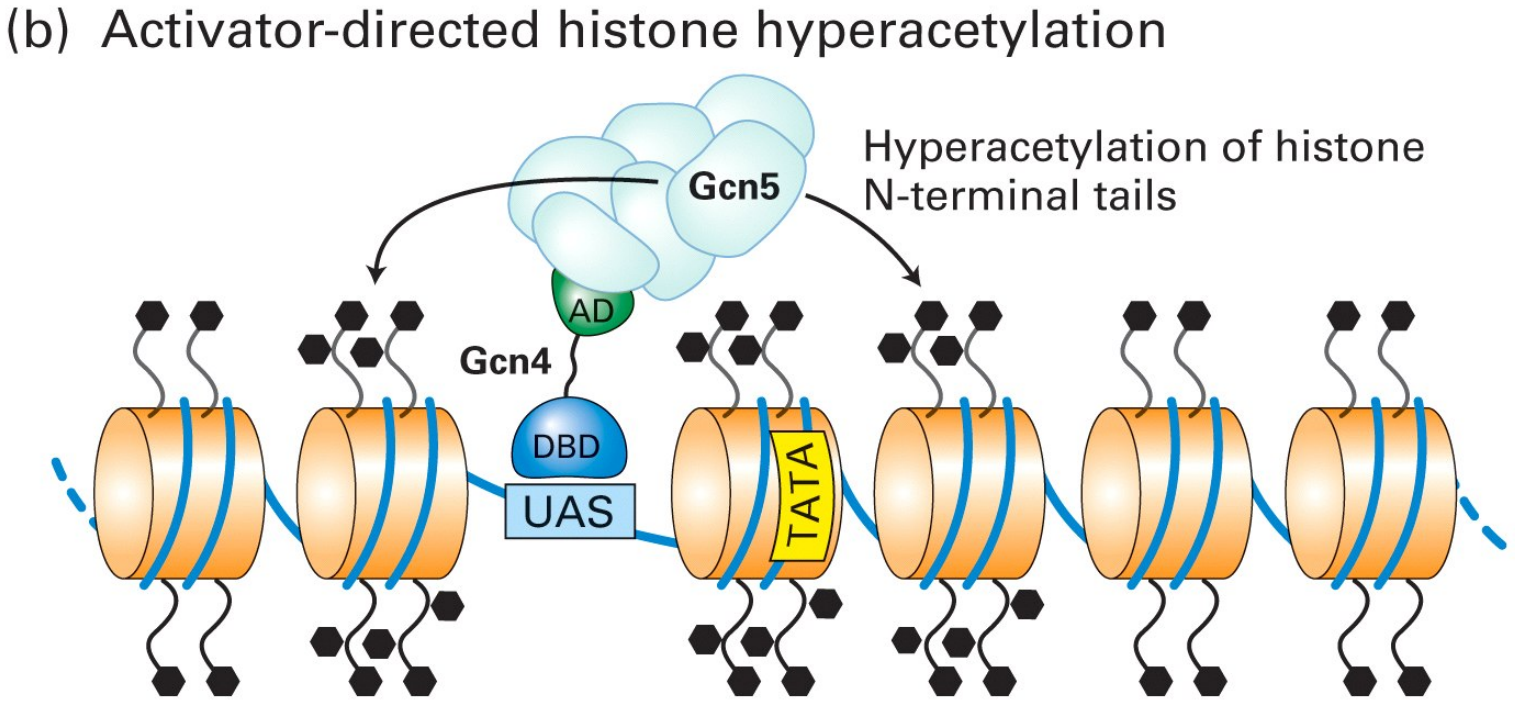
Schematic model of silencing mechanism at yeast telomeres



Multiple copies of RAP1 bind to a simple repeated sequence at each telomere region, which lacks nucleosomes. This nucleates the assembly of a multiprotein complex through protein-protein interactions between RAP1, SIR2, SIR3 and SIR4, and the hypoacetylated N-terminal tails of histones H3 and H4 of nearby nucleosomes. Sir2 deacetylates the histone tails. The heterochromatin structure at each telomere encompasses approx. 4 kb. Association of several condensed telomeres forms higher order chromatin structure.

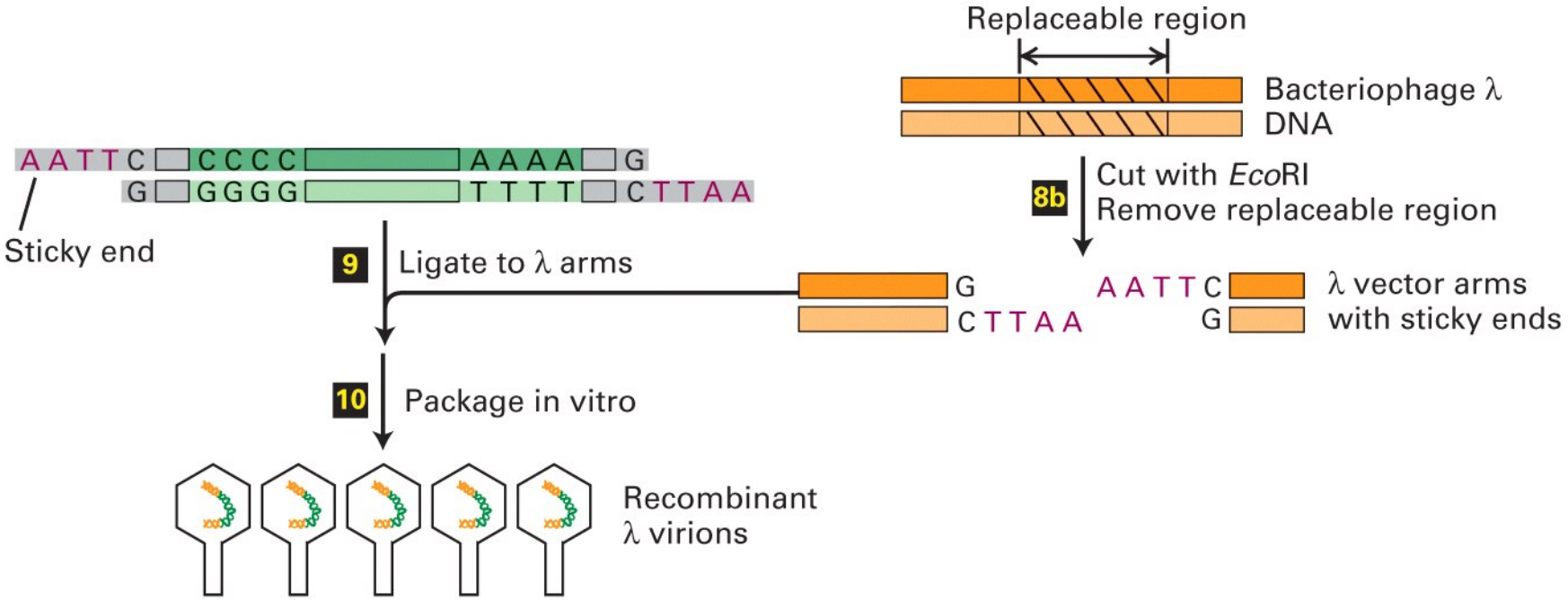


Ume6 is a transcriptional repressor which binds to a specific upstream control element (URS1). The Ume6 repression domain (RD) binds SIN3, a subunit of a multi-protein complex that includes RPD3, a histone deacetylase. Deacetylation of histone N-terminal tails on nucleosomes in the region of the Ume6 binding site inhibits binding of general transcription factors at the TATA box, thereby repressing gene expression

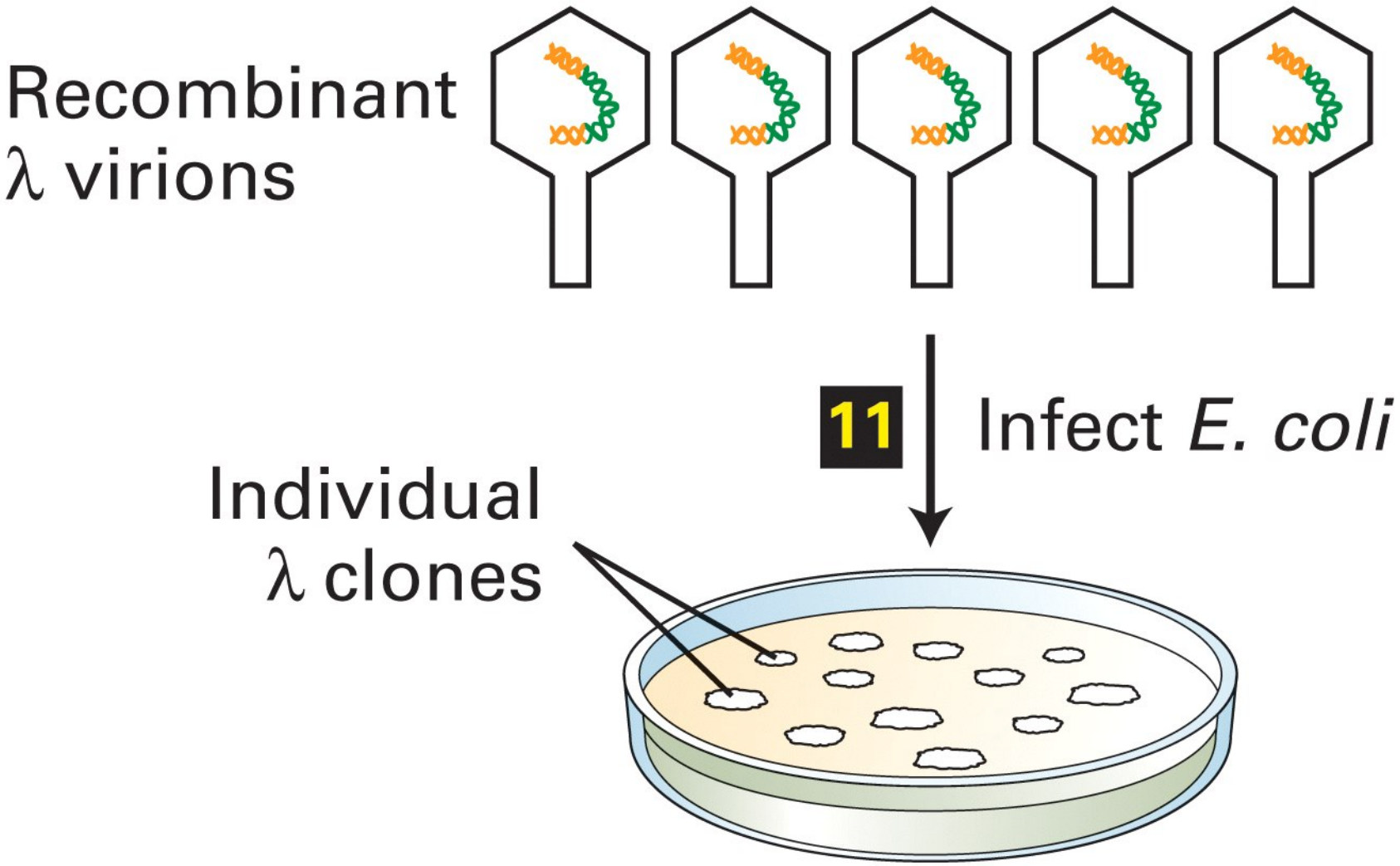


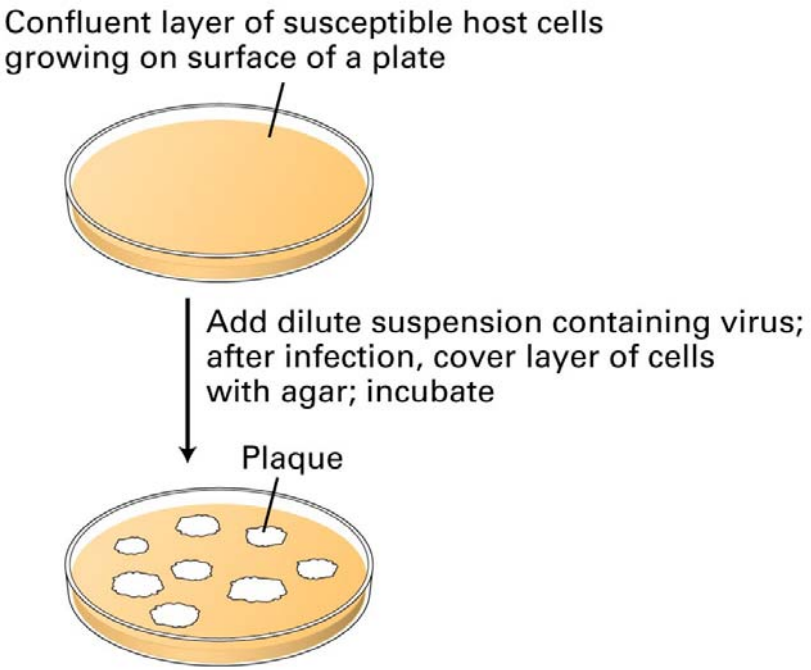
The GCN4 activation domain (AD) interacts with a multi-protein complex that includes the GCN4 catalytic subunit. Subsequent hyperacetylation of histone N-terminal tails on nucleosomes in the vincinity of the GCN4-binding site facilitates access of the general transcription factor required for initiation.

A c-DNA library can be constructed using a bacteriophage λ vector

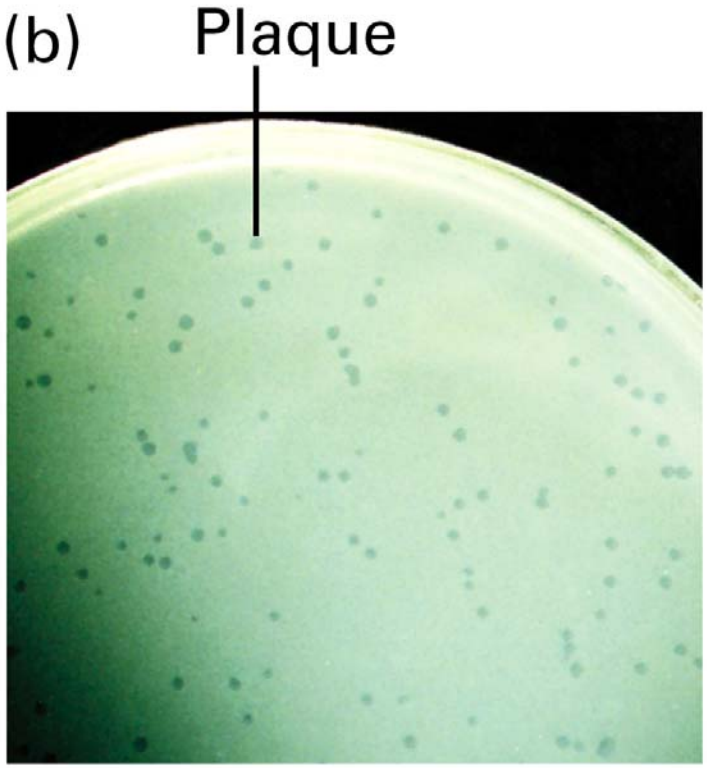


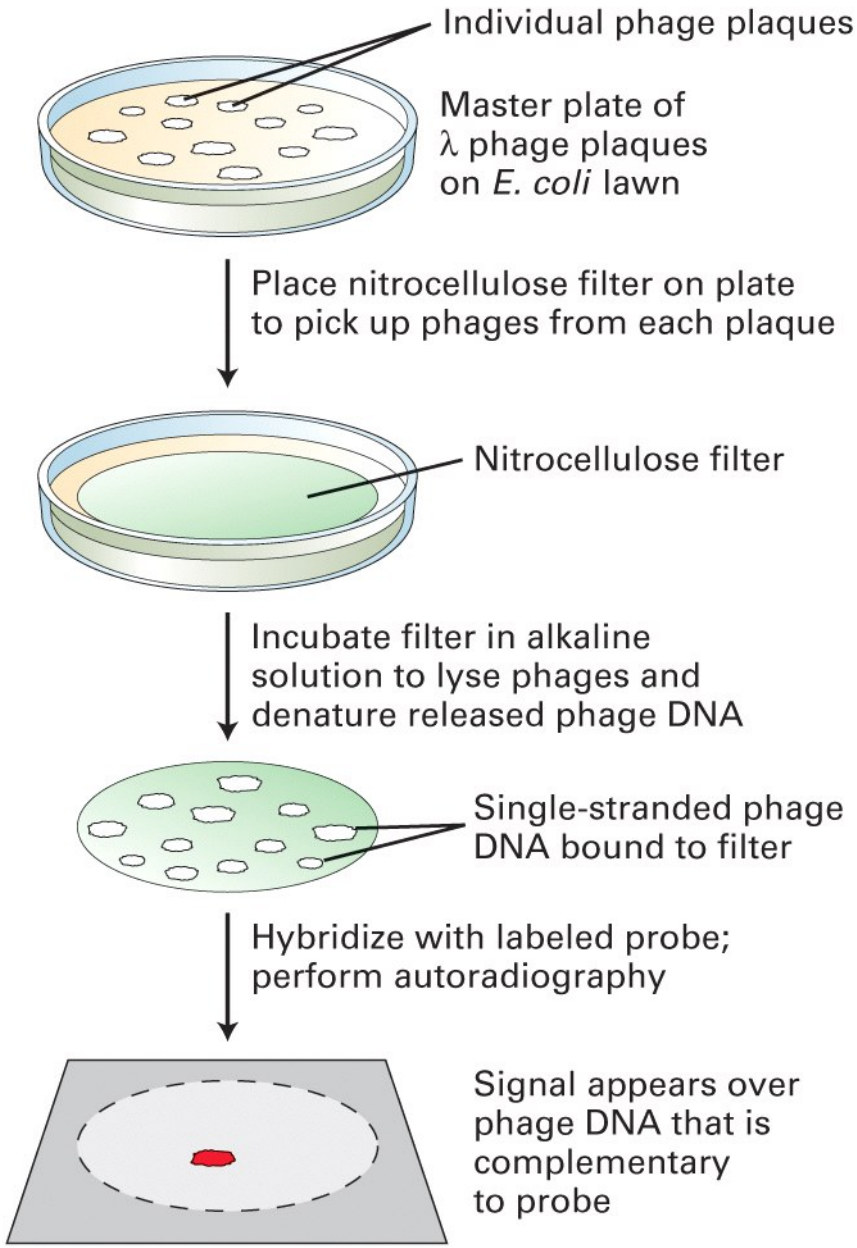
To maximaize the size of the exogenous DNA that can be inserted into the λ genome, the non essential regions of the λ genome usually are deleted. Plating the of the recombinant phage on a lawn of E.coly generates a set of cDNA clones representing all the cellular mRNAs.

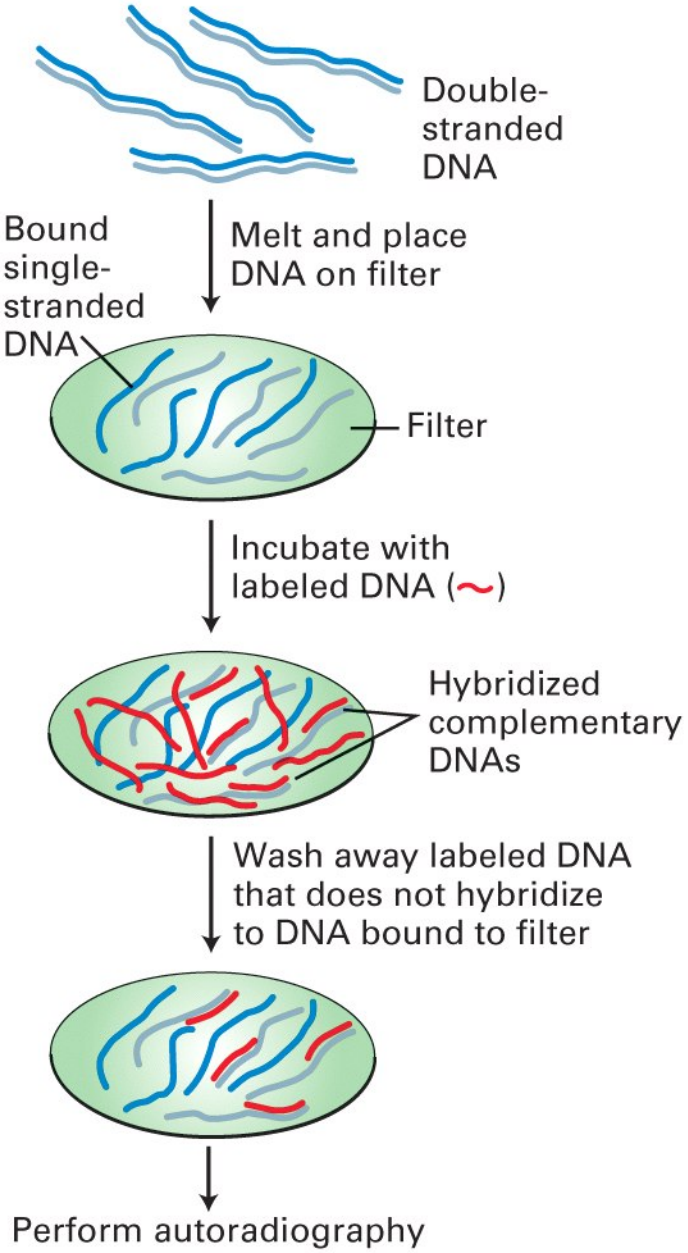

Recombinant
λ virions
11
Infect *E. coli*
Individual
λ clones



Each plaque represents cell lysis initiated by one viral particle (agar restricts movement so that virus can infect only contiguous cells)




Individual phage plaques
Master plate of
λ phage plaques
on E. coli lawn
Place nitrocellulose filter on plate
to pick up phages from each plaque
Nitrocellulose filter
Incubate filter in alkaline
solution to lyse phages and
denature released phage DNA
Single-stranded phage
DNA bound to filter
Hybridize with labeled probe;
perform autoradiography
Signal appears over
phage DNA that is
complementary
to probe


Double-stranded DNA
Bound single-stranded DNA
Melt and place DNA on filter
Filter
Incubate with labeled DNA (~)
Hybridized complementary DNAs
Wash away labeled DNA that does not hybridize to DNA bound to filter
Perform autoradiography