

MB&B 301b – Principles of Biochemistry II – Spring 2016

Part 2: Gene Regulation—Based on lectures by Prof. Karla Neugebauer

James Diao

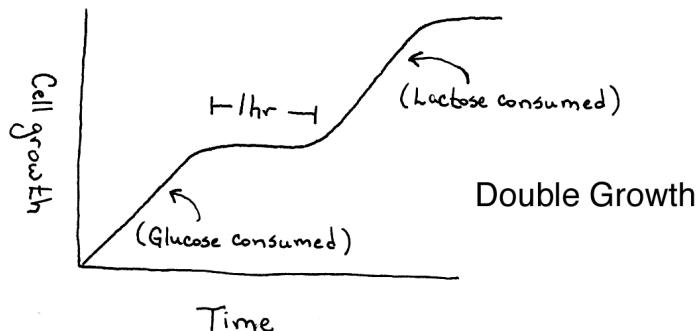
Table of Contents

Lecture 13: Prokaryotic Gene Expression I	3
Lecture 14: Prokaryotic Gene Expression II	7
Lecture 15: Regulation of Bacteriophage Lambda	13
Lecture 16: Cloning & Genomic Analysis	19
Lecture 17: Eukaryotic Genome I	29
Lecture 18: Eukaryotic Genome II	35
Lecture 19: Eukaryotic Transcription.....	41
Lecture 20: Eukaryotic RNA Processing	47
Lecture 21: Eukaryotic Translation.....	55

Lecture 13: Prokaryotic Gene Expression I

A. Monod and Lactose Biochemistry

- a. E. coli growing on glucose/lactose media have a biphasic growth curve.



- b. β -galactosidase: lactose \rightarrow glucose + galactose.

- c. Monod and Conn made artificial galactosides for (in)activation/monitoring lac induction.

- i. Original lac Efficient substrate, diminishing inducer
- ii. Types 1 & 2 Efficient substrates, no induction
- iii. Type 3 Efficient substrate, no induction, yellow color
- iv. Type 4 & 6 (IPTG) Resistant to enzyme, strong inducer
- v. Type 5 Resistant to enzyme, antagonist

Types 4-6 have a sulfur substitution that prevents cleavage.

- d. Monod 1951

- i. Grew auxotrophs on minimal media until the required AA runs out.
- ii. Added inducer- no enzyme.
- iii. Added AA- enzyme.
- iv. Conclusion: Enzyme induction requires synthesis (not pre-existing enzyme)
- v. Confirmation with pulse chase: radioactive AA found in new enzymes.

- e. Monod 1952

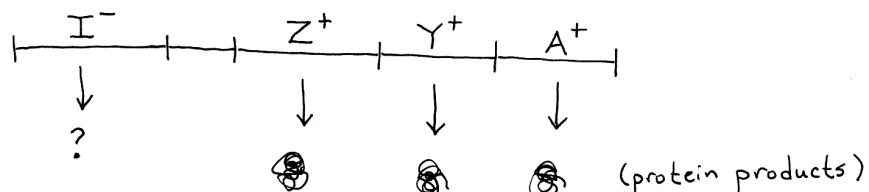
- i. Addition of IPTG (to E. coli growing on glucose medium) induced β -gal activity after 3 min.
- ii. Proved that the inducer is sufficient to produce B-gal.

- f. Cohen and Monod 1954

- i. Constitutive mutation for β -gal expression also constitutively expresses permease.
- ii. This suggests that Y and Z are induced/controlled together.

B. Inducer (I) of β -gal Activity

- a. Genetic map of E. coli places I near Z and Y



- b. In a constitutive I- mutant, all 3 are always on (no need for IPTG).

c. Models for the nature of I

1. I is a site needed for control by the inducer (**I must be DNA**).
2. I codes for an activator gp, activated by IPTG to turn on ZYA.
- I- codes for an always-active mutant.
3. I codes for a repressor gp that turns off ZYA unless IPTG stops it.
- I- codes for an inactive mutant that fails to stop ZYA.

C. Pardee, Jacob, Monod (PaJaMo) Experiment

- a. Bacterial conjugation: Hfr delivers 2nd chromosome into Hfr- to create merodiploids/partial diploids.
- b. Observe dominant/recessive nature in enzyme induction.
- c. Females were given resistance to streptomycin and T6, which killed off males to ensure that the assay only measures activity in Hfr- cells.
- d. Why do we want Y+ on the Hfr- I-/Z-/Y+?

You want there to be lac available to assay for Z activity. I would guess that conjugational recombination is not super often, so you don't want to count on both Y and Z being fixed.

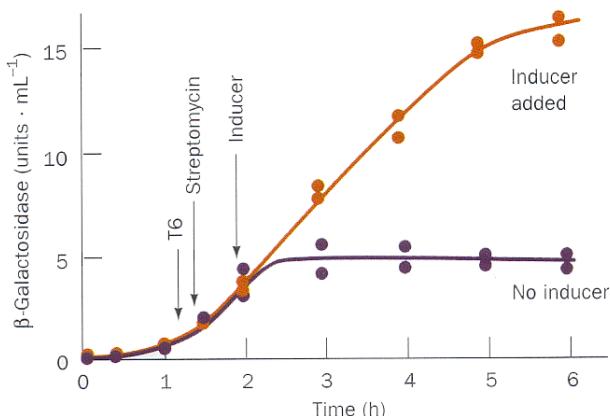
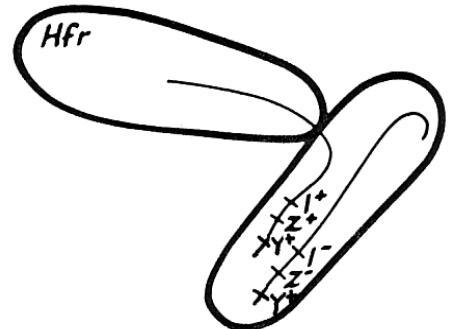


FIGURE 29-4. The demonstration of the existence of the *lac* repressor through the appearance of β -galactosidase in the transient merozygotes (partial diploids) formed by mating I^+Z^+ Hfr donors with $I^-Z^-F^-$ recipients. The F^- strain was also resistant to both **bacteriophage T6** and **streptomycin**, whereas the Hfr strain was sensitive to these agents. Both types of cells were grown and mated in the absence of inducer. After sufficient time had passed for the transfer of the *lac* genes, the Hfr cells were selectively killed by the addition of T6 phage and streptomycin. In the absence of inducer (lower curve), β -galactosidase synthesis commenced at around the time that the *lac* genes had entered the F^- cells but stopped after ~ 1 h. If inducer was added shortly after the Hfr donors had been killed (upper curve), enzyme synthesis continued unabated. This demonstrates that the cessation of β -galactosidase synthesis in uninduced cells is not due to the intrinsic loss of the ability to synthesize this enzyme but to the production of a repressor specified by the I^+ gene. [After Pardee, A.B., Jacob, F., and Monod, J., *J. Mol. Biol.* 1, 173 (1959).]

D. Additional

- a. If I is a DNA sequence, then what??? Need cis-trans assay (next page).
- b. No mixing of cytoplasm during conjugation; the gp of I is in the cytoplasm.
- c. If I is an activator:
 - i. I- is dominant over I+ because a rogue activator can't be stopped by a discriminating one.
 - ii. After conjugation, β gal+ no matter what.
- d. If I is a repressor:
 - i. I+ is dominant over I-, because a working repressor can step in for the broken one.
 - ii. After conjugation, β gal depends on IPTG.
- e. Hfr I-/Z- x F- I+/Z+ gives no β -gal because the there will always be repressor floating around from the active I+.

E. Repressor must bind IPTG

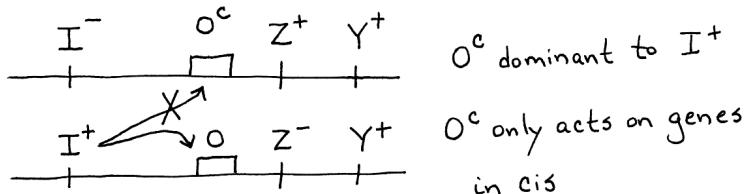
- a. I^S mutants (super-repressors) cannot be inactivated by IPTG.

b. I^S dominant over $I + IPTG$ because nothing will pop these guys off.

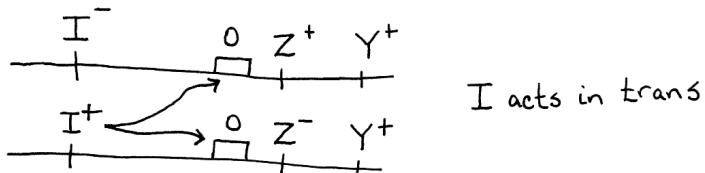
F. Repressor must act at DNA level on an operator

a. O^C is a constitutive operator that's insensitive to repressor

i. O^C acts in cis: not cytoplasmic (no gene product).



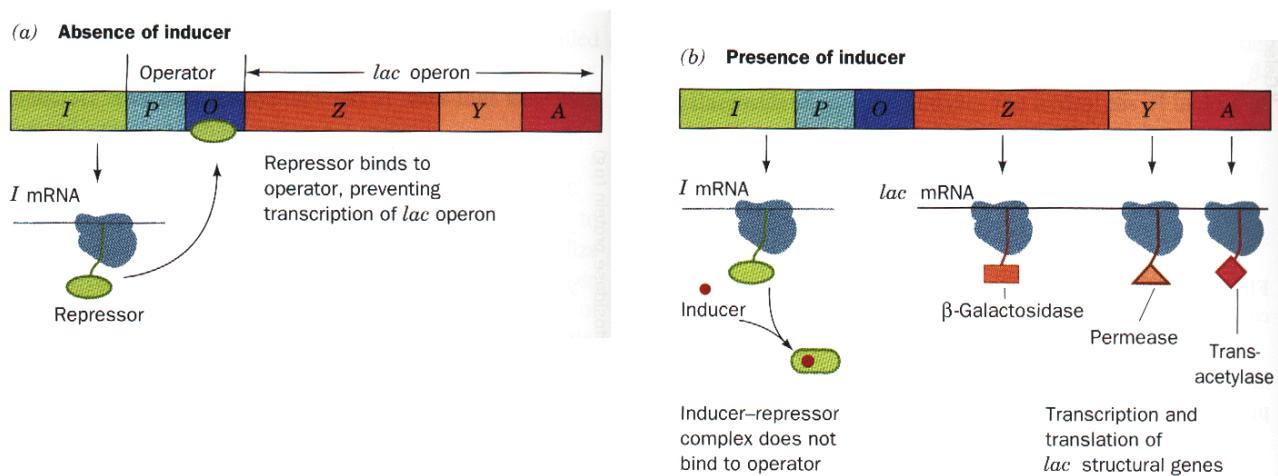
ii. I^+ acts in trans: cytoplasmic gp and can affect other chromosomes.



iii. Dominant to I^+ , because working repressor won't stop these guys.

b. Operon model of gene regulation: coordinated units of expression (operons) have an operator and gene set that are controlled by distant regulatory genes.

G. Model for Negative Regulation of Lac Operon



H. The Lac Repressor

a. Muller-Hill and Gilbert 1966 isolated lac repressor using labeled ^{14}C -IPTG

i. Tetrameric protein (1 IPTG per subunit).

ii. Affinity for DNA (K_d) is 10^9 stronger at the operator.

iii. Two domains: DNA-binding domain and IPTG-binding domain.

iv. DNA-binding rate of $10^{10} \text{ M}^{-1}\text{s}^{-1}$ much faster than diffusion-limited maximum of $10^7 \text{ M}^{-1}\text{s}^{-1}$. This suggests a sliding mechanism (1D search instead of 3D).

I. Isolating the Lac Operator (DNA segment)

a. Mix purified Lac repressor + sonicated DNA.

b. Bind the mixture with nitrocellulose (holds protein).

c. Add IPTG. The bound repressor will release DNA.

- d. Add Lac repressor to bind DNA and digest unbound/unprotected segments with DNase I.
- e. Sequence the resulting DNA fragment (convert to RNA first; we're 1960's era savages).
- f. Sequence has symmetric positions (dimeric binding) and O^c mutation sites (binding sites)

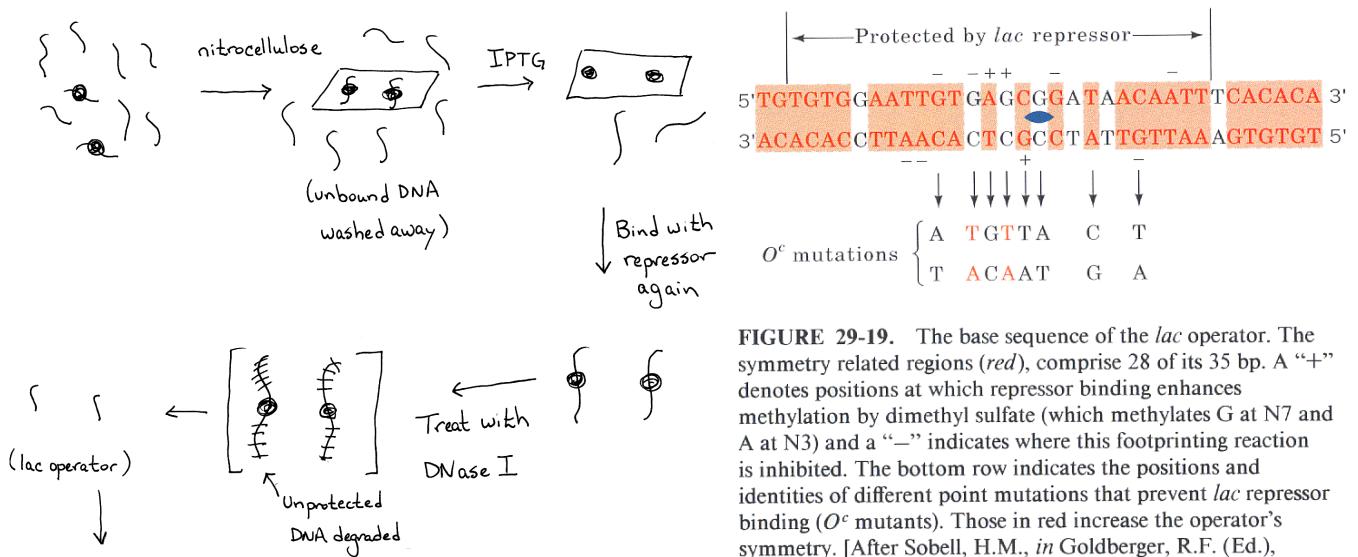
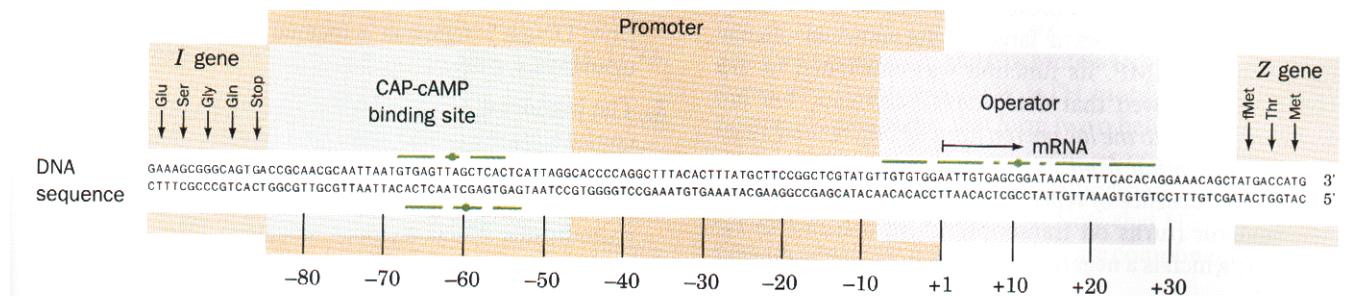


FIGURE 29-19. The base sequence of the *lac* operator. The symmetry related regions (red), comprise 28 of its 35 bp. A “+” denotes positions at which repressor binding enhances methylation by dimethyl sulfate (which methylates G at N7 and A at N3) and a “-” indicates where this footprinting reaction is inhibited. The bottom row indicates the positions and identities of different point mutations that prevent *lac* repressor binding (O^c mutants). Those in red increase the operator’s symmetry. [After Sobell, H.M., in Goldberger, R.F. (Ed.),

J. Lac Operator/Promoter Region from lacI to lacZ

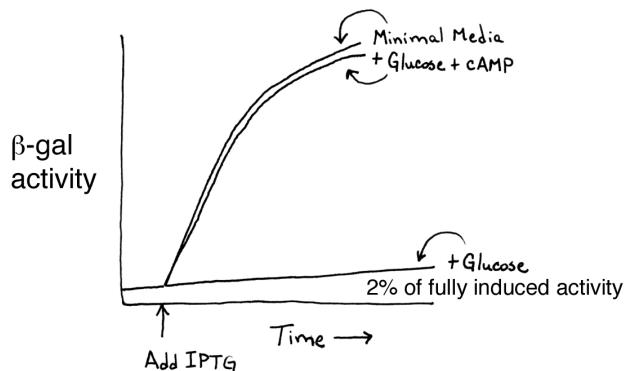
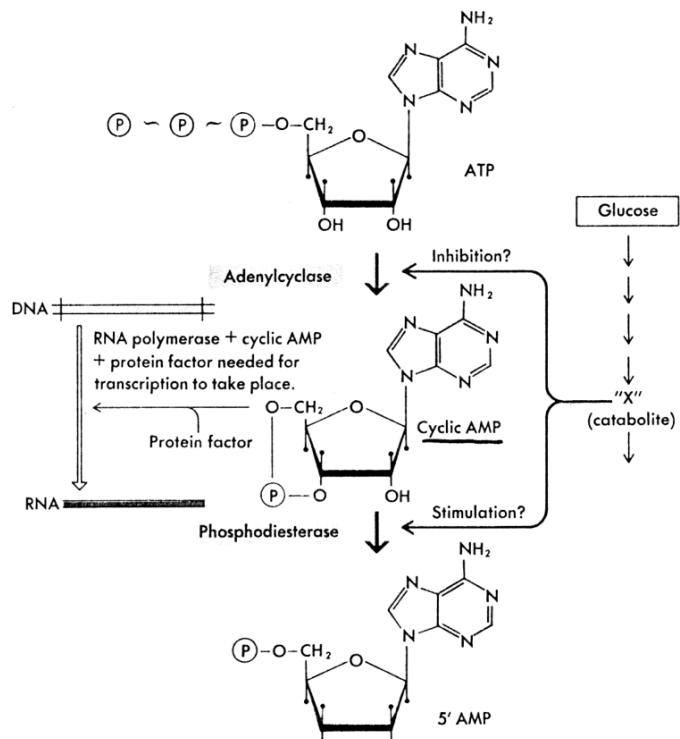
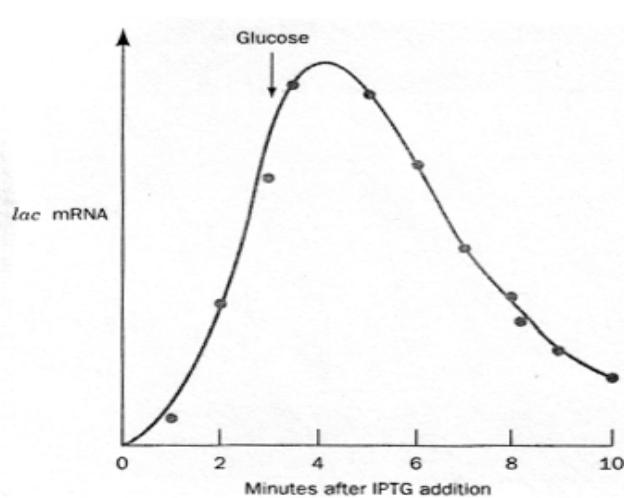


- a. Operator: positions [-7, +28]
- b. RNAP binds between [-20, +20]
- c. Lac repressor and RNAP can bind *simultaneously*.
- d. Mechanism: Lac repressor increases the specific affinity of RNAP for the promoter, so that it cannot leave and enter elongation.

Lecture 14: Prokaryotic Gene Expression II

A. Catabolite Repression and Positive Gene Regulation

- a. Lac repressor control alone does NOT fully explain the double growth curve.
- b. Glucose inhibits/decreases lac expression.



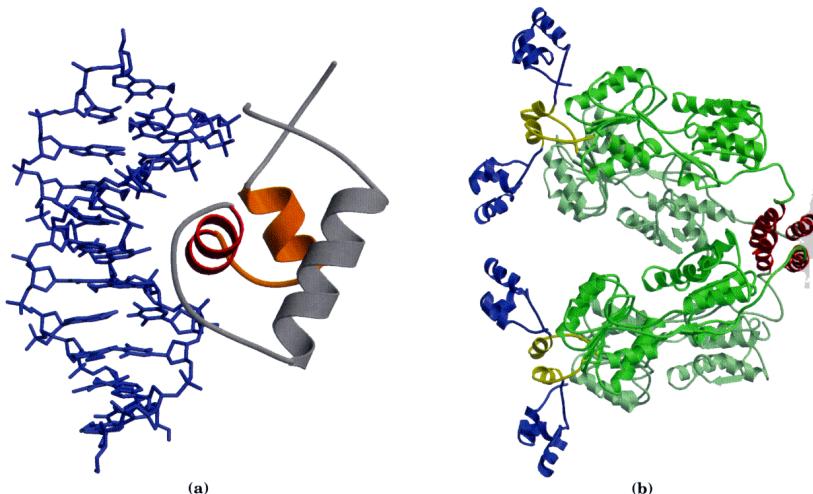
c. Effects of cAMP

- a. cAMP (5'-3' cyclic AMP) is low when glucose is present.
- b. Addition of cAMP overcomes glucose repression (it must be the ultimate effector).
- c. cAMP binds CAP (catabolite activator protein)
- d. CAP/cAMP binds promoters to activate 100+ glucose-sensitive operons, like lac, arabinose, and galactose.
- d. Thus, lac operon responds to the presence of lactose and the absence of glucose.

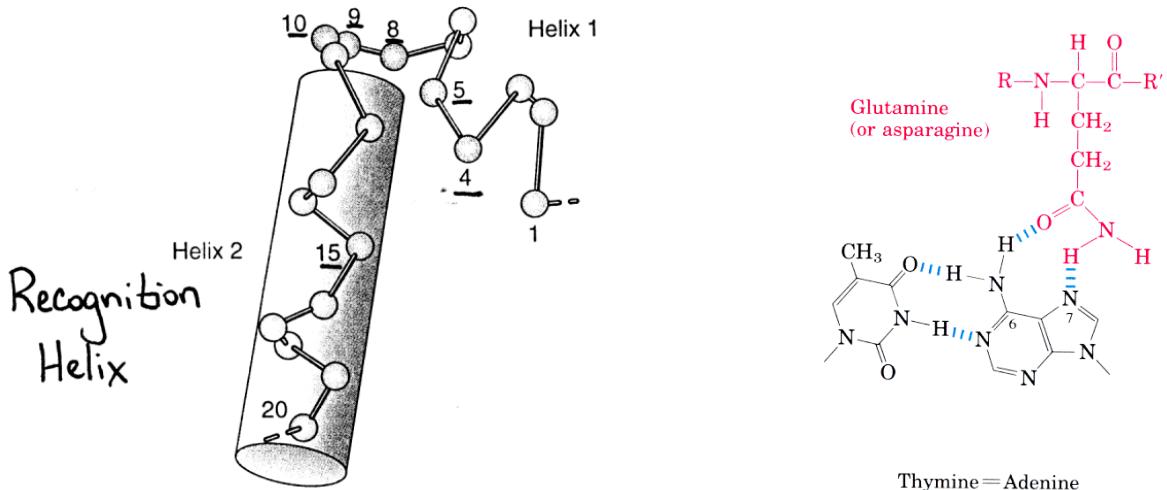
Condition	Result	CAP site	Promoter	Operator	β -galactosidase gene
1. No cyclic AMP; no inducer	No transcription because repressor blocks RNA polymerase (and CAP is not bound)				
2. Cyclic AMP present; without inducer	No transcription because repressor still blocks RNA polymerase	cAMP	CAP	Repressor	IPTG
3. No cyclic AMP; inducer present (e.g., IPTG)	Little transcription because CAP is not bound				
4. Cyclic AMP and inducer present	RNA polymerase binds and initiates lactose operon mRNA synthesis.			RNA polymerase	mRNA

B. Lewis 1996 Lac Repressor Structure with DNA and IPTG

- a. Helix-turn-helix structural motif: common for prokaryotic protein-DNA interaction.
- b. Full dimeric lac repressor. Tips are HTH motifs, connected by allosteric helices.

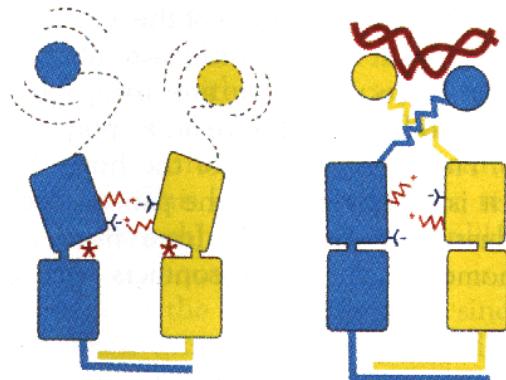


- c. The recognition helix (leftmost, red) interacts with DNA in the major groove. Recognition helices in dimeric proteins are 34 Å apart to bind adjacent major grooves. ($3.4 \text{ Å/nt} * 10 \text{ nt turns} = 34 \text{ Å/turn}$).
- d. DNA binding can be disrupted by direct mutations at the interface, or in the support structure that changes the 34 Å distance or other conformational necessities.
- e. Conserved residues
 - i. 4 hydrophobic AAs stabilize helix-helix interactions.
 - ii. 2 AAs create the bend between helices.
- f. The other AAs provide DNA specificity. Many strategies; no reliable rules for specificity.



C. Lac Repressor during IPTG-Binding

- Lac repressor binding bends DNA by $\sim 60^\circ$.
- IPTG-binding pulls apart and misaligns the 34A-apart alpha-helices at the major grooves.
- Asterisks: IPTG.
- Subdomain N-termini electrostatically cross-interact to stabilize the broken dimer. These interactions cease when IPTG dissociates.
- Where are the mutations?
 - I^S mutations should be in the IPTG/lac binding site.
 - I^- mutations should be in the lac promoter binding site.
 - O^C mutations should be at the DNA-helix interface, removing specificity.



D. CAP during cAMP-Binding

- cAMP stimulates large conformational changes in CAP.
- CAP binds DNA as a dimer in consecutive major grooves (HTH motif).
- DNA is bent by $\sim 90^\circ$ at the binding site.
- The kinks close the major groove and widen the minor groove.
- CAP recruits RNAP by stabilizing its interaction with promoter DNA.

E. Model for Lac Repressor Interaction with the Lac Operator by DNA Looping

Binding of lac repressor tetramer bends DNA

Lac Repressor: bottom

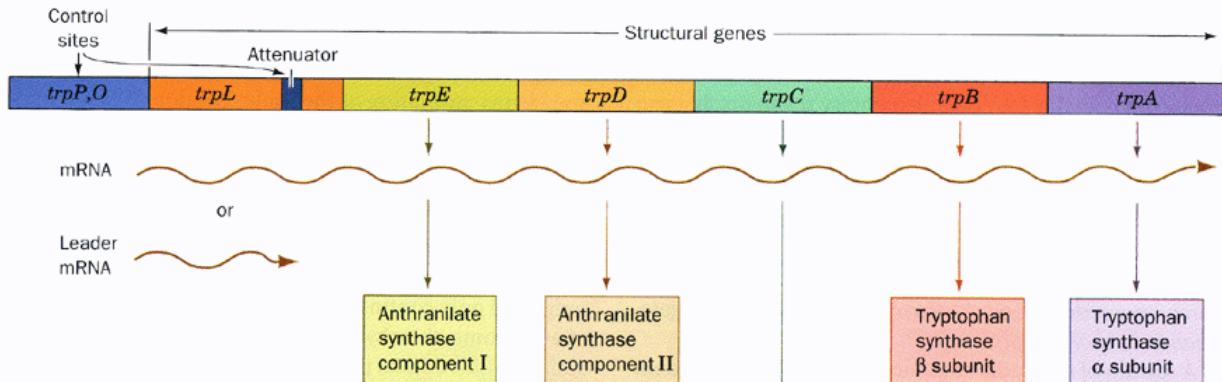
Lac Operators: 2 separate ends of DNA

Promoter: N and SW loop



F. Transcriptional Regulation of the Trp Operon

- a. Trp biosynthesis enzymes (*trpE-trpA*) are all on a single mRNA transcript
- b. Trp repressor downregulates transcription 70x.
- c. Trp repressor binding to *trp* aligns the recognition helices (HTH motif).



G. Trp Repressor Binding to DNA Operator

- a. Recognizes sequence indirectly (no direct H-bonding to bases)
 - i. Via shape of phosphate backbone (some solvent-mediated).
 - ii. H_2O mediates polar contacts with bases as an “honorary” side chain.

H. Trp Repression

- a. Charles Yanofsky: studied TrpA and TrpB Expression

trpR^+ and trp^+	most repression (1)
trpR^- and trp^+	some repression (70)
trpR^- and trp^-	no repression (600)

Conclusion: trpR^+ and trp^+ both limit expression.

*NOTE: $\Delta[\text{gene}]$ refers to a mutation in [gene]

- b. Mutations in 3 genes leads to no repression under conditions of trpR^- and trp^+

ΔtrpL (post-operator region)	no repression (550)
$\Delta\text{trpT}_{\text{TS}}$ (trp tRNA)	no repression (500)
$\Delta\text{trpRS}_{\text{TS}}$ (trp tRNA synthetase)	no repression (550)

Conclusion: The trp-mediated repression pathway involves translation and the *trpL* region.

- c. mRNA Levels

- i. $\text{trpR}^-/\text{trp}^+$ $\text{trpL} \gg \text{trpE}$ (8x)
- ii. $\text{trpR}^-/\text{trp}^-$ $\text{trpL} == \text{trpE}$

Conclusion: The trp-mediated repression pathway involves mRNA abundance.

- d. Transcriptional control by termination at the *trpL* leader peptide, which has TRP-TRP.

I. Mechanism

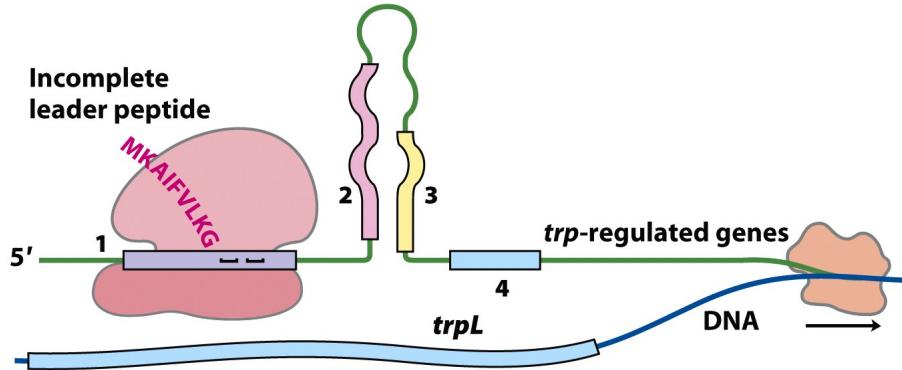
- a. 2 possible secondary structures

- i. 1-2, and 3-4 terminator
- ii. 2-3 antiterminator (1 and 4 unpaired)

- b. Low trp

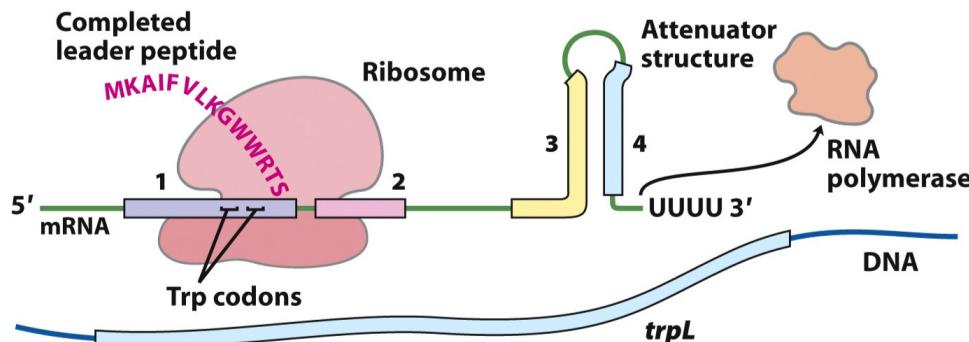
- i. Ribosome pauses at trp-trp codons in (1), because not enough are available.
- ii. (2) forms a paired structure with (3).

This stops it from forming the attenuator structure with (4).



c. High trp

- i. Enough trp is available for the ribosome to quickly translate (1) and get to (2) before (3) is transcribed. After the ribosome passes, (2) pairs with (1).
- ii. (3) is now free to pair with (4) to form a terminator-like attenuator.



J. Attenuation as a General Means of Transcriptional Control

- a. AA-operons commonly use the attenuation mechanism of containing lots of that AA.

Table 31-3 Amino Acid Sequences of Some Leader Peptides in Operons Subject to Attenuation

Operon	Amino Acid Sequence ^a
<i>trp</i>	Met-Lys-Ala-Ile-Phe-Val-Leu-Lys-Gly-TRP-TRP-Arg-Thr-Ser
<i>pheA</i>	Met-Lys-His-Ile-Pro-PHE-PHE-PHE-Ala-PHE-PHE-PHE-Thr-PHE-Pro
<i>his</i>	Met-Thr-Arg-Val-Gln-Phe-Lys-HIS-HIS-HIS-HIS-HIS-HIS-Pro-Asp
<i>leu</i>	Met-Ser-His-Ile-Val-Arg-Phe-Thr-Gly-LEU-LEU-LEU-Asn-Ala-Phe-Ile-Val-Arg-Gly-Arg-Pro-Val-Gly-Gly-Ile-Gln-His

b. Advantages:

- i. Economical
- ii. Sensitive to several AAs
- iii. Links AA synthesis to demand, not supply.
- iv. ONLY works in prokaryotes, where transcription and translation are linked.

Lecture 15: Regulation of Bacteriophage Lambda

A. Lambda Phage Cycle

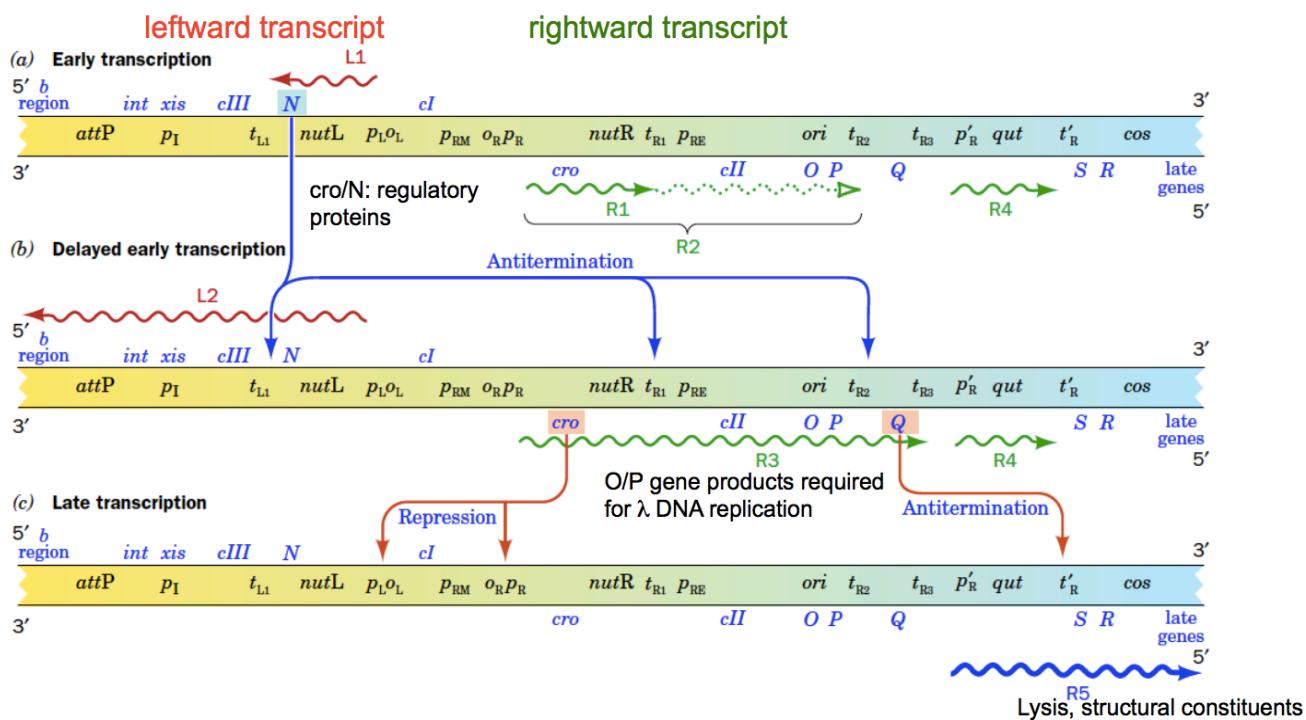
- Phage adsorbs to cell (via tail fibers on maltose transporter) and injects its DNA.
- Linear phage DNA circularizes
- Lysogenic mode: phage DNA stably integrated into host chromosome.
 - Induction: DNA damage (e.g. UV) induces excision → lytic mode.
This can (rarely) occur spontaneously (every 10^5 divisions).
 - Lytic mode: makes DNA/proteins for ~100 progeny phages; lyses host.
- Terminology: prophage (DNA) is integrated into a lysogen (host)
- Good model for gene regulation and viral latency.

B. Genome organization

- Genes highly organized into structural, regulatory, etc.
- The chromosome circularizes/linearizes at the *cos* site.

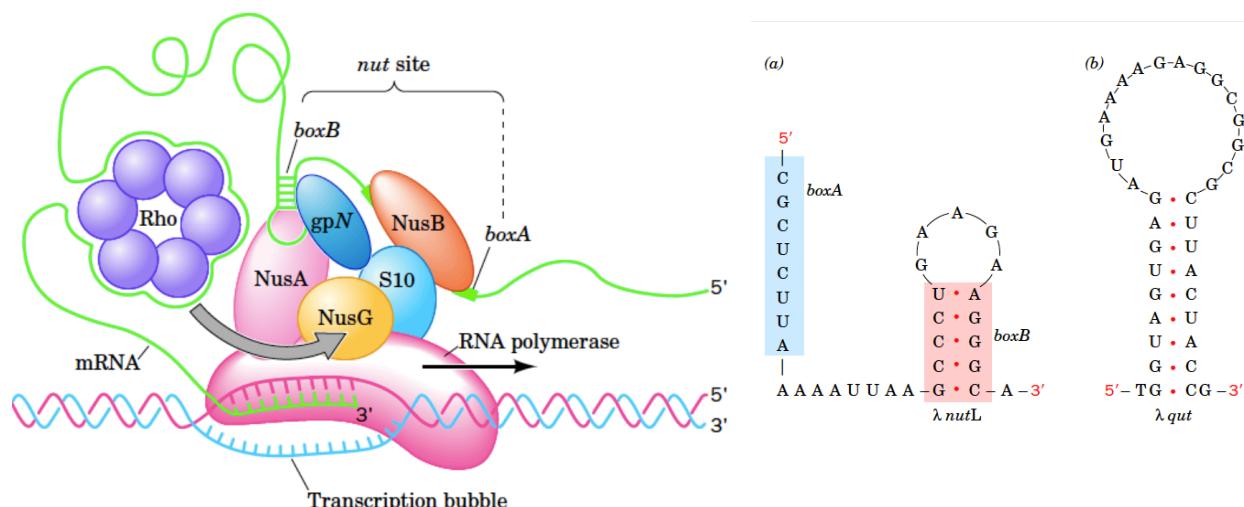
C. The lytic mode: 3-phase transcriptional program

- Early
 - Active P_L (left) makes gpN and stopped by t_L .
 - Active P_R (right) makes gpCro and stopped by t_{R1} .
 - Active P'_R (right-2) makes nothing and stopped by t'_R .
- Delayed early
 - gpN antiterminates t_L , t_{R1}/t_{R2} , turning on CII, CIII, O, P, Q.
 - gpO and gpP involved in lambda DNA replication.
- Late:
 - Cro blocks P_L P_{RM} , and P_R → blocks lysogeny
 - gpQ antiterminates t'_R → turning on S and R (lysis, structural constituents).



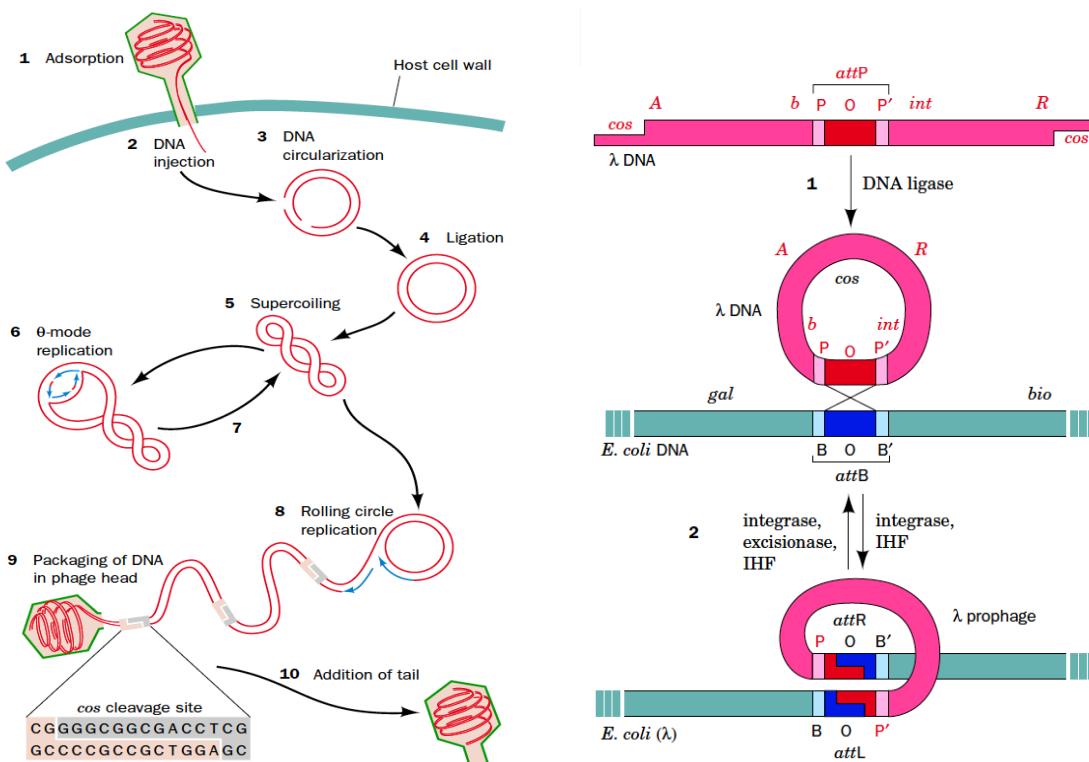
D. Anti-termination sites and mechanism

- a. nut = N utilization; qut = Q utilization
 - b. gpN + Nus proteins (host factors) form an antitermination complex on the nut site that ALSO binds transcribing RNAP.
 - c. Two possible ways that it antiterminates:
 - i. This complex inhibits RNAP from pausing at a termination site.
This stops **rho factor** from catching up and releasing the transcript.
 - ii. gpN modulates NusG-rho interaction to inhibit transcript release.



E. Cellular decision: lytic or lysogenic?

- a. Lysogens express ONLY the cl gene (lambda repressor), which stops superinfection.
 - b. Induction dependent on RecA-mediated **SOS response** (often by UV).
 - c. Gene products O and P (delayed early transcription) find ori and initiate DNA replication.

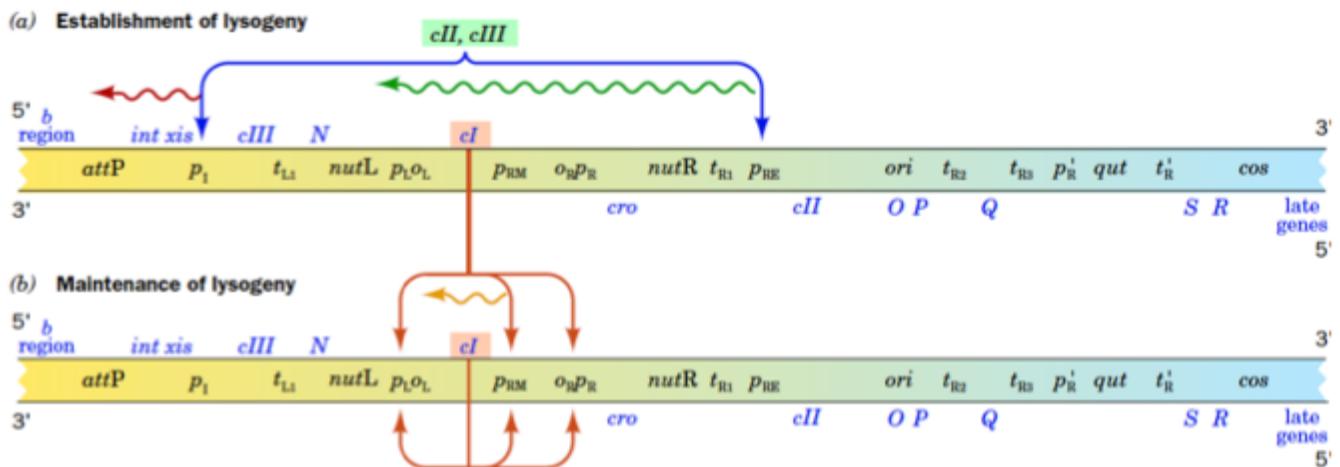


F. Lambda Assembly

- Two parallel pathways for head and tail assembly.
- Discovered by conditional lethal mutants (allow you to block assembly at specific stages)
 - ts: temperature-sensitive mutants (amber mutations)
 - Suppressor tRNA rescue nonsense mutations until you remove it.
 - Intermediates accumulate and can be inspected by electron microscopy.
- In vivo* genetic complementation: cross infection of 2 mutants to see if their mutations are on the same gene.
- In vitro* complementation: mixing incrementally more purified cell extracts from 2 different mutants that rescue each other to identify protein products.

G. Lysogenic Transcriptional Reprogramming

- cI repressor binds to operators O_R and O_L to turn off everything.
- P_{RM} starts with AUG! How does this affect the protein level?
- After CII and CIII are made in delayed early transcription. The following 2 replace late transcription.

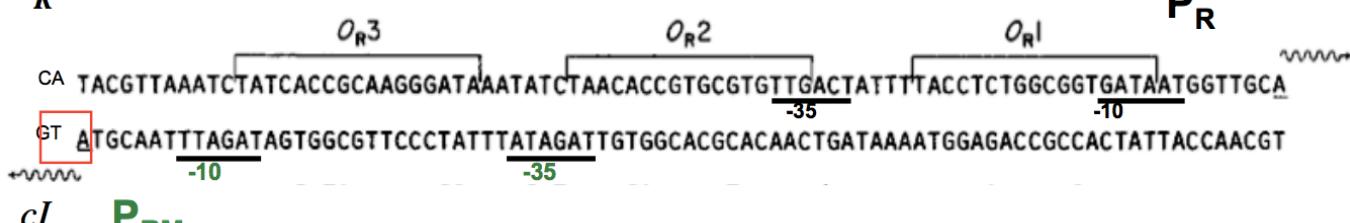


- The two cI mRNAs have different translational efficiencies
 - P_{RM} starts with AUG
 - P_{RE} works 20x better; starts with [Shine-Dalgarno]—AUG
 - P_{RE} starts it; needs to be sensitive. P_{RM} doesn't care.

H. Autogenous Regulation of cI transcription by cI at O_R

- P_{RM} lies within and overlaps with O_R .

O_R :

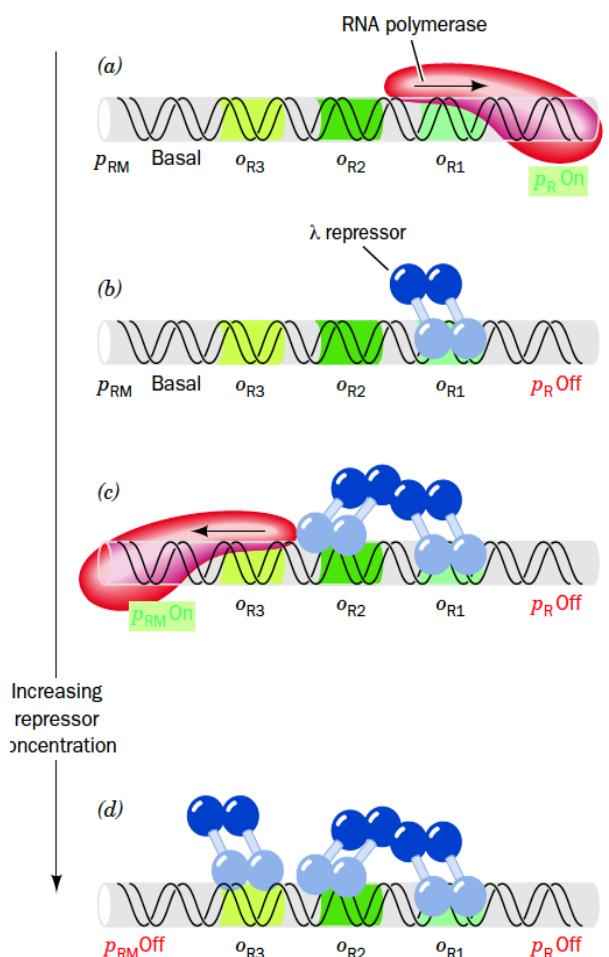
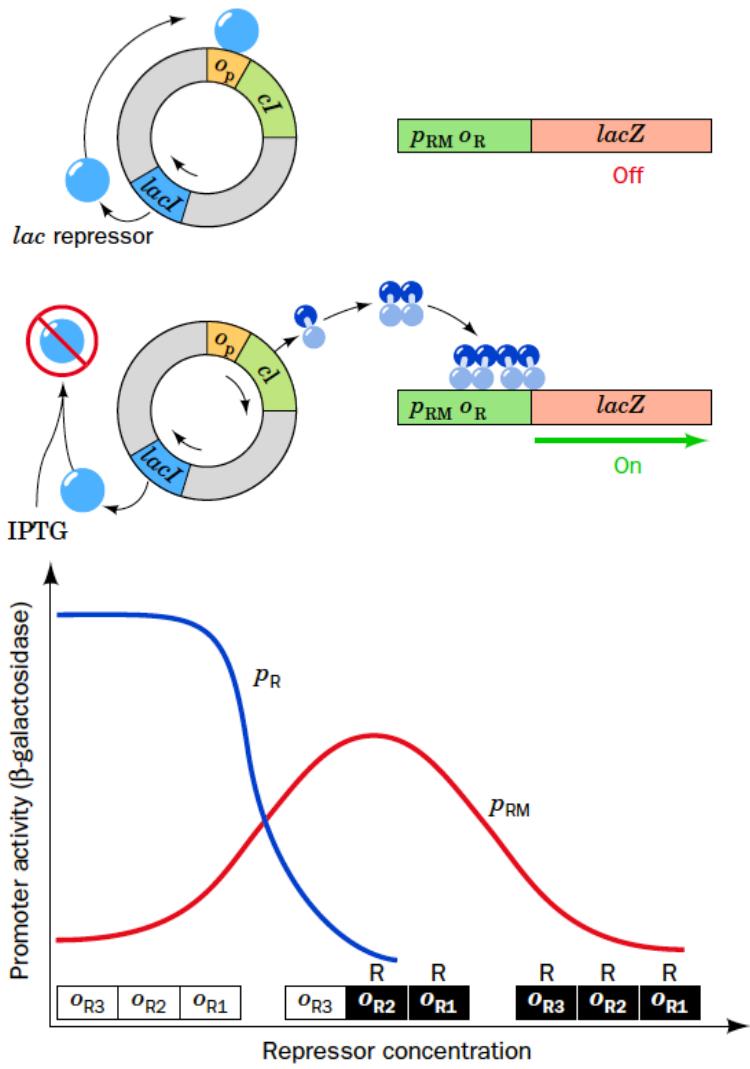


cI P_{RM}

- O_R is ~80 bp with 2 divergent promoters and 3 binding sites for lambda Cl repressor.
- Cl repressor is a dimer with a helix-turn-helix DNA-binding motif.

I. Mechanism for cI Regulation of P_{RM} and P_R

- a. Finding this is difficult because:
 - i. The regulated gp is the regulator!
 - ii. No readily assayable product.
- b. Reporter Gene Assay
 - i. Plasmid (left): lac operator-promoter (O_p) controls the cI gene. cI creates repressor, which can be blocked by IPTG.
 - ii. Prophage (right): contains promoter P_{RM} fused to $lacZ$.
Level of β -galactosidase ($gp-lacZ$) thus reflects P_{RM} activity.
 - iii. Similar experiments replaced cI with cro repressor and P_{RM} with P_R promoter.

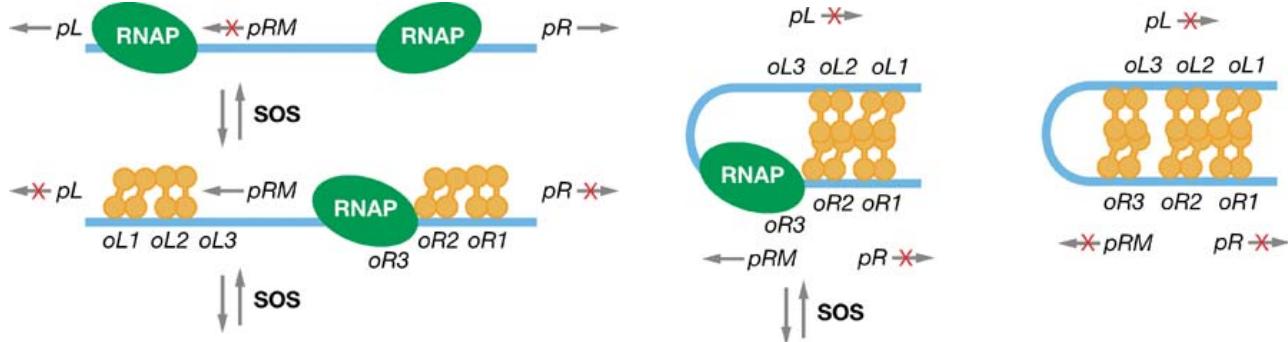


A. Promoter activity vs. [cI] repressor]

- a. No repressor: Active P_R and P_{RM} .
- b. 1st repressor on O_{R1} : Blocks (-) P_R (10x higher affinity for O_{R1})
- c. 2nd repressor on O_{R2} : Boosts (+) P_{RM} (activates RNAP by contacting sigma)
- d. 3rd repressor on O_{R3} : Blocks (-) P_{RM} (via RNAP loading); lysogen levels of cI .
- e. $cI-O_{R1}$ enhances cI binding to O_{R2} – almost always simultaneously occupied.
- f. This keeps [cI] in a narrow range for stable lysogeny.
Too high will release early, and too low will prevent induction when desired.

B. DNA Looping Between O_L and O_R Creates a Cooperative Network

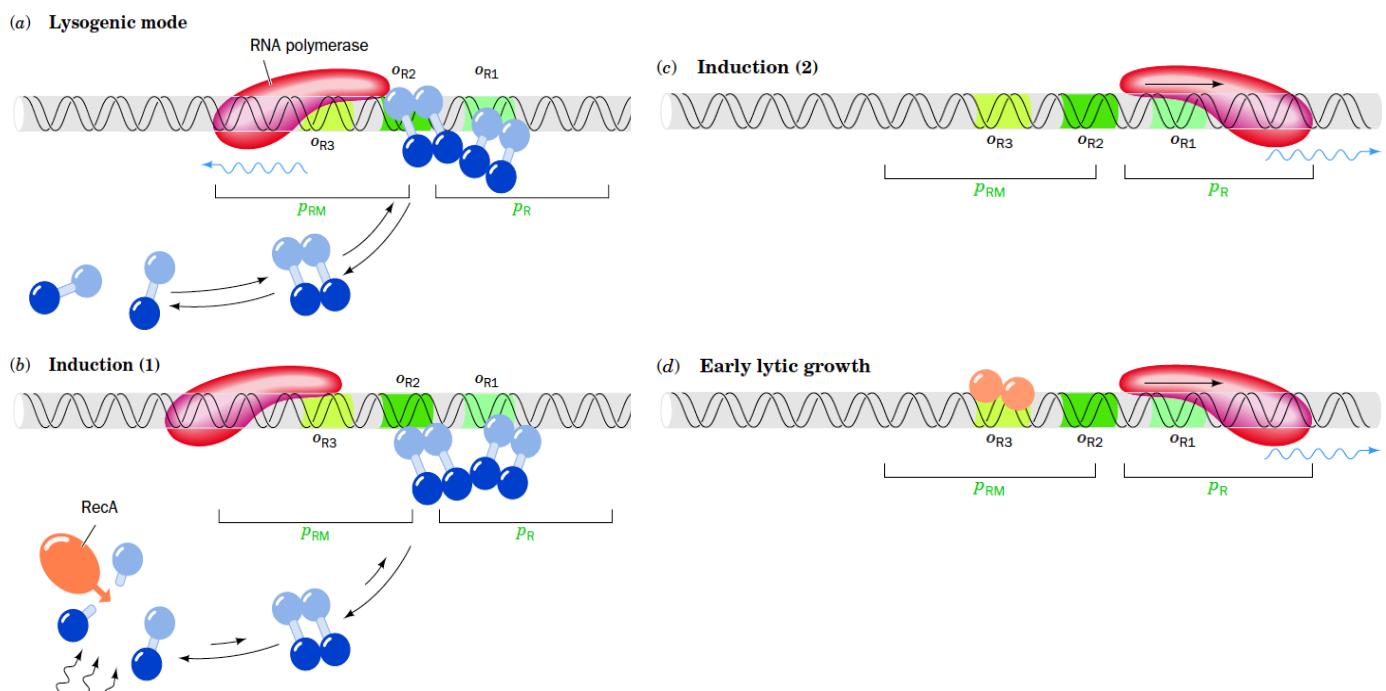
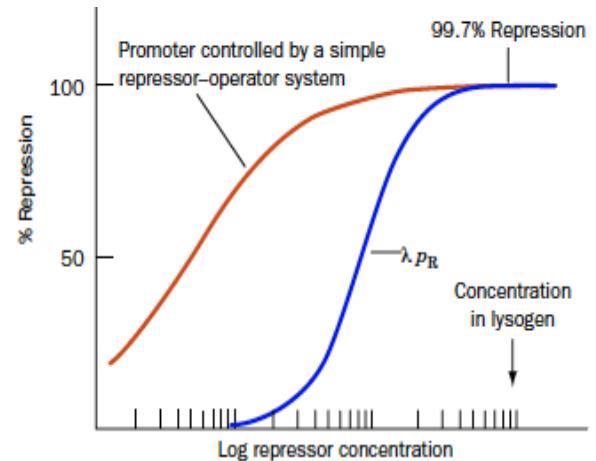
- a. cI binds to O_{L1} and O_{L2} as well, and the dimers associate in a long-range DNA loop.
- b. cI then binds cooperatively to O_{L3} and O_{R3}; allows balanced control of P_L and P_R.



- c. Based on cooperativity between O_{R1} and O_{R2}.
- d. Repression curves for lambda P_R (right) and simple system, like lac operon (left).
- e. Lambda system is more sensitive to changes in [cI repressor].

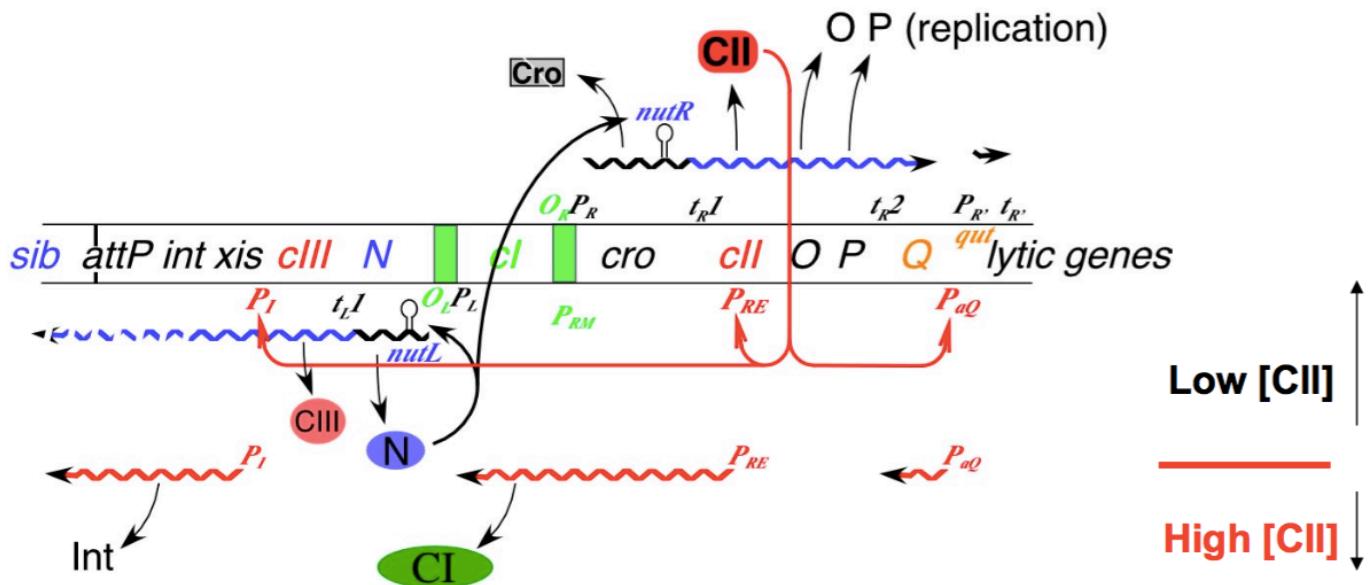
C. Induction Mechanism

- a. Lysogeny: cI binds O_R at C-terminus
cI activates RNAP at N-terminus
- b. DNA damage: RecA stimulates self-cleavage of cI into monomers.
- c. Induction: Early transcription from P_R and P_L.
- d. Cro binds O_{R3} and blocks cI synthesis → irreversible lytic cycle



D. Infection and the Lysis-Lysogeny Decision

- Low [CII] → Lysis, High C[II] → Lysogeny
- Protease FtsH normally degrades CII
 - CIII inhibits FtsH
 - High [cAMP] inhibits FtsH
- High M.O.I. → P_R and P_L are ON → N antiterminator → P_R CII and P_L CIII are active.
- Starvation conditions → high [cAMP], inhibits FtsH, stabilizes CII.
- CII is a transcriptional activator
 - Binds and activates -35 regions of P_I , P_{RE} , and P_{aQ} .
 - P_{RE} transcript efficiently translated into CI. WHY? Shine-Dalgarno.



E. Summary

- Initial lytic/lysogenic decision dependent on cII and cIII
- Normally, protease activity high, which breaks down cII → lytic cycle
- With limited nutrients, protease activity low, stabilizing cII → lysogenic cycle
- Phage is dormant until better times, when more resources are available.

Lecture 16: Cloning & Genomic Analysis

A. Molecular Cloning Propagates and Amplifies DNA

- a. Determine the sequence
- b. Make probes for experiments
- c. Express RNA or protein for structural/functional studies
- d. Make mutants for functional analysis
- e. Analyze gene expression in vivo
- f. Make transgenic animals
- g. Genome engineering
- h. Develop tests for diagnosis and treatment of disease.

B. Key Enzymes of Molecular Biology

- a. Taq DNA Pol: synthesizes new DNA strands (thermal stable)
- b. Polynucleotide kinase: phosphorylates 5'-OH of polynucleotide.
- c. Terminal transferase: adds homopolymer tails to 3'-end of a duplex
- d. Alkaline Phosphatase: removes terminal phosphates on 5' or 3' end.

C. Type II restriction enzymes

- a. Recognize and cleave dsDNA at specific 4-8 bp palindromic targets.
- b. Can be staggered ends (C2 symmetry) or blunt ends.
- c. Almost every 6-bp palindromic sequence (64 possibilities specified by 3 nt) is cut by at least 1 restriction enzyme.

D. DNA Ligase

- a. Joins blunt ends and complementary sticky ends of unrelated DNA.
- b. Powered by ATP → AMP + PP_i or NAD⁺ → NMN⁺ + AMP (species dependent).
- c. E. coli mechanism
 - i. NAD⁺ adenyl group covalently transferred onto Lys residue (phosphoamide adduct)
 - ii. The adenyl group is transferred onto the 5'-(P)-terminus of the DNA nick via 5'-PP_i linkage.
 - iii. Enzyme catalyzes the formation of the phosphodiester linkage by attack of the 3'-OH on the 5'-(P), which seals the nick and releases AMP.

E. Cloning and DNA Amplification

- a. Plasmid (extrachromosomal DNA) and source DNA are cut with the same restriction enzyme.
- b. Specificity of the annealing reaction allows specific insertions in a plasmid vector.
- c. Transformation and propagation of plasmid within host bacteria will amplify DNA.

F. Cloning cDNA in Eukaryotes

- a. Eukaryotic mRNAs are polyadenylated at 3'-ends.
- b. cDNA is made using poly-T primer + RT: this gives a weighted representation of all expressed genes in a sample.
- c. Each cDNA can be “subcloned” into a plasmid vector, and the net result is a cDNA library (of bacteria).

DNA Ligation

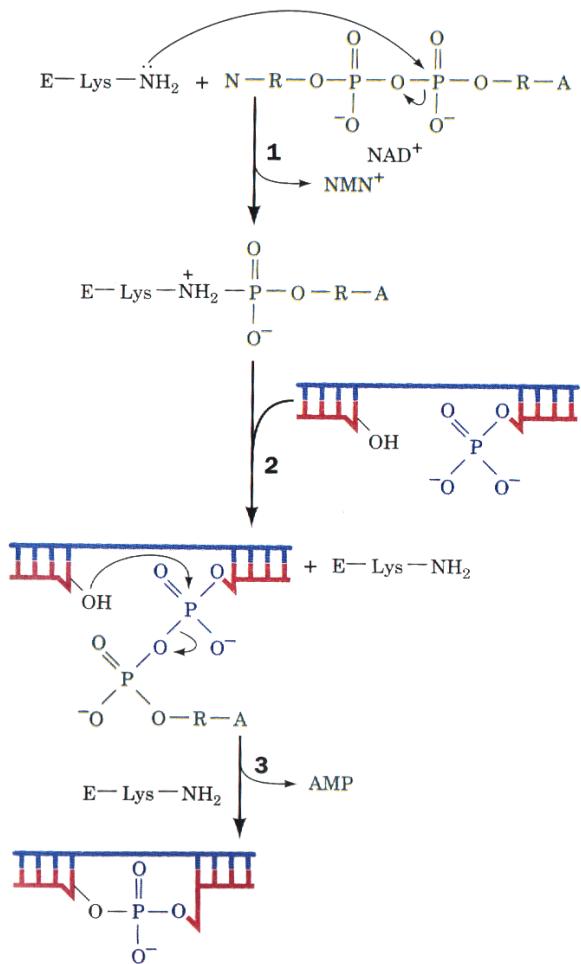
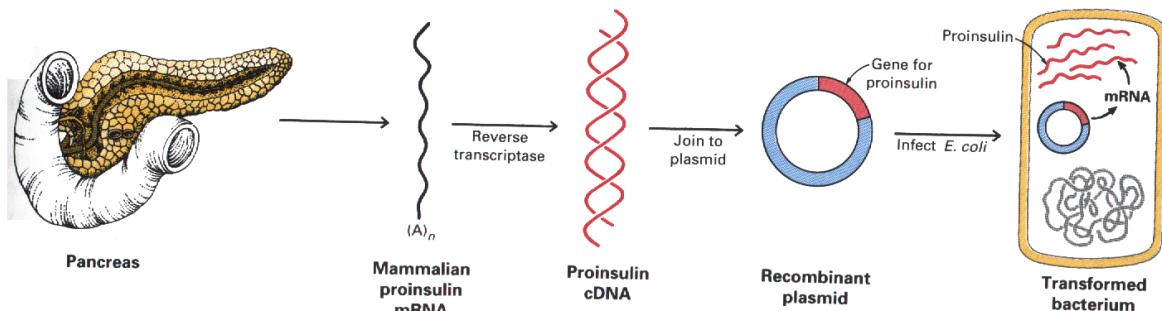
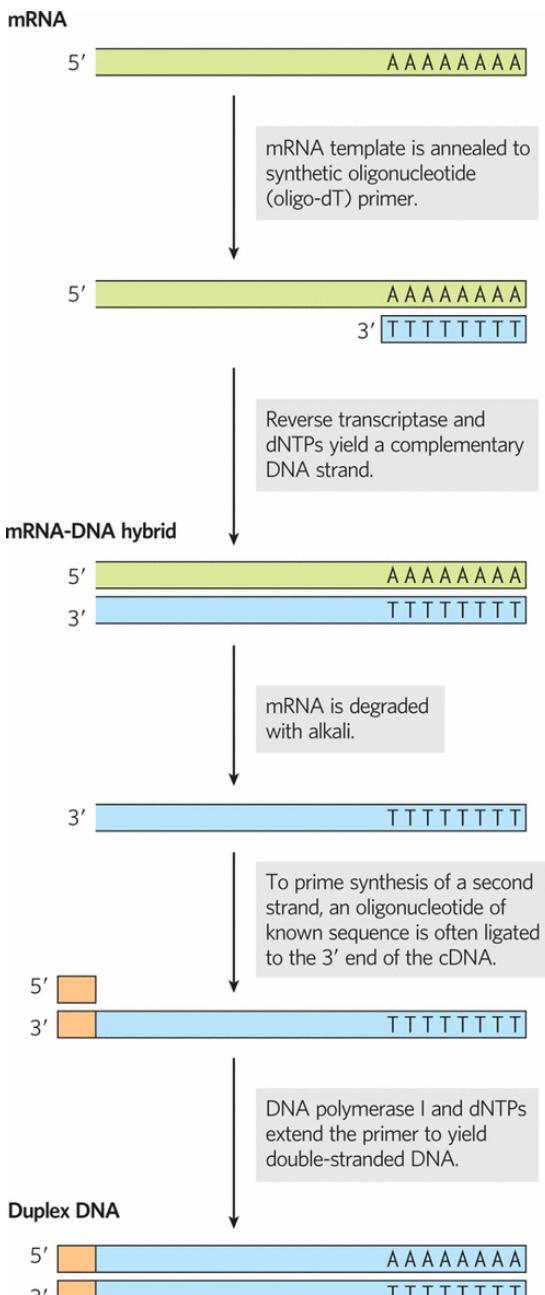


FIGURE 31-16. The reactions catalyzed by *E. coli* DNA ligase. In eukaryotic and T4 DNA ligases, NAD⁺ is replaced by ATP so that PP_i rather than NMN⁺ is eliminated in the first reaction step. The numbered steps are described in the text.

cDNA libraries



G. DNA Amplification by PCR

- Much simpler and more powerful than cloning (though less accurate).
- Developed in mid-1980s, after DNA oligo synthesis could prepare 15-100 nt sequences.
- Begin with very small amounts of genomic DNA or cDNA.
- Primer oligos anneal to opposing strands with 3' ends available.
- Thermostable DNA Pol extends primers to full oligos, and the strands are melted.
- Process repeats 25-30 times, with DNA amounts doubling each cycle.
- Product is DNA between and including the primers (other regions are lost).

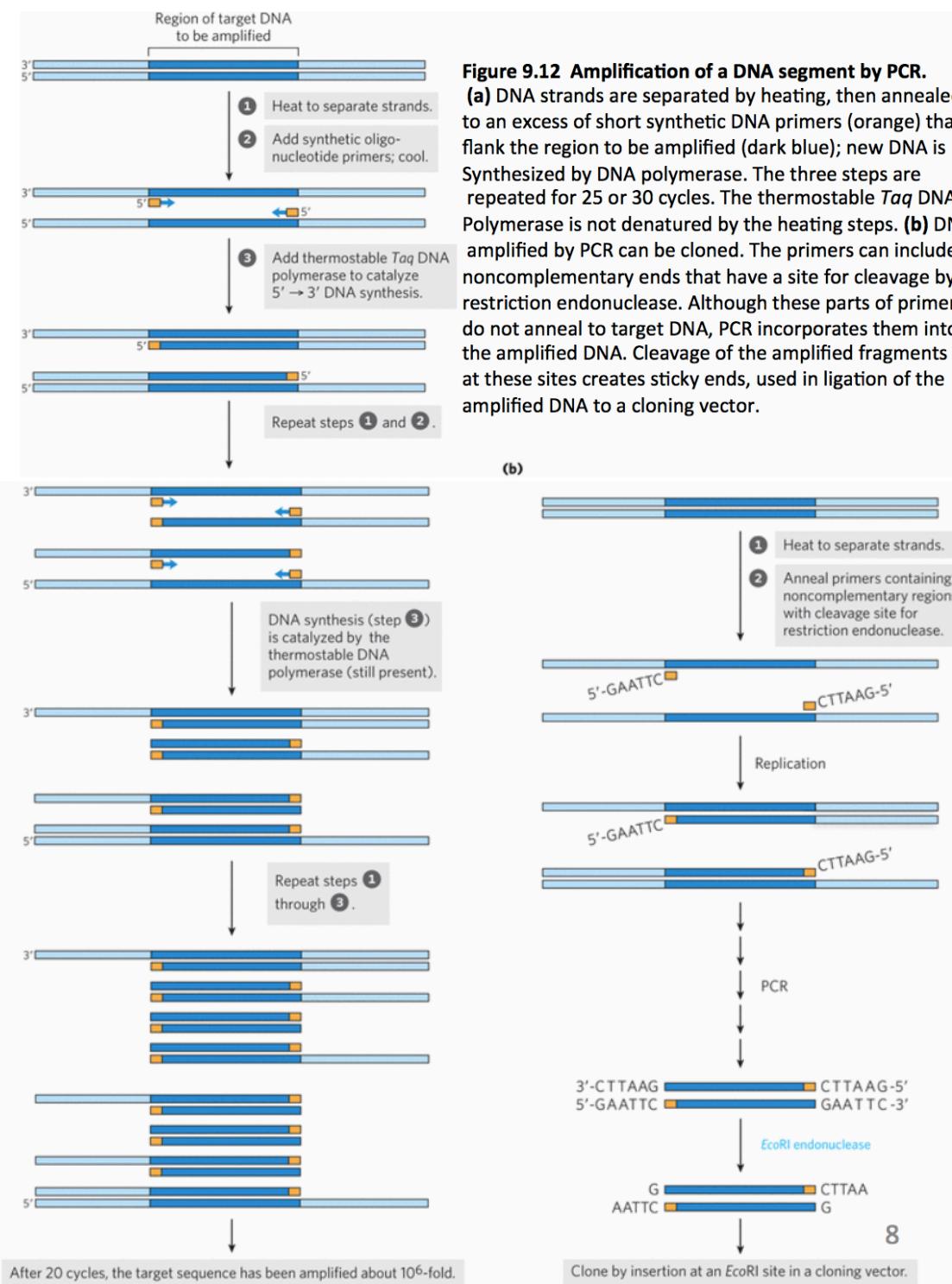


Figure 9.12 Amplification of a DNA segment by PCR.

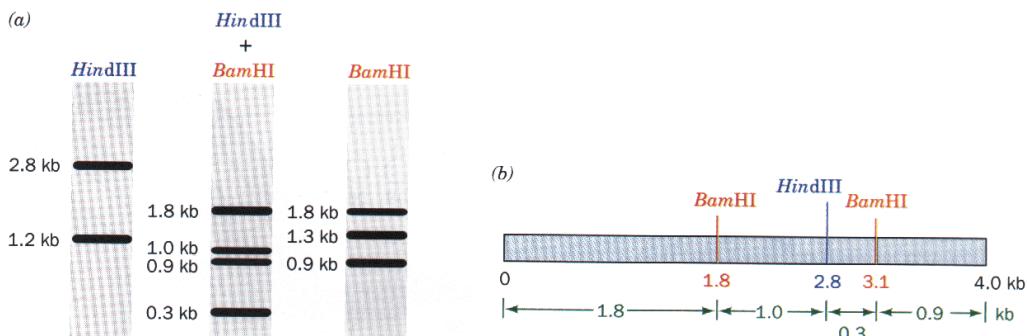
(a) DNA strands are separated by heating, then annealed to an excess of short synthetic DNA primers (orange) that flank the region to be amplified (dark blue); new DNA is synthesized by DNA polymerase. The three steps are repeated for 25 or 30 cycles. The thermostable *Taq* DNA Polymerase is not denatured by the heating steps. **(b)** DNA amplified by PCR can be cloned. The primers can include noncomplementary ends that have a site for cleavage by a restriction endonuclease. Although these parts of primers do not anneal to target DNA, PCR incorporates them into the amplified DNA. Cleavage of the amplified fragments at these sites creates sticky ends, used in ligation of the amplified DNA to a cloning vector.

H. Molecular Cloning Details

- Alkaline phosphatase (e.g. calf-intestinal phosphatase, CIP) removes 5'-phosphates to prevent self-ligation. Otherwise, you'll transform bacteria without the insert.
- Transformation by heat shock (usu Ca^{2+} competent cells) or electroporation.
- Cloning vectors may be designed to include:
 - Cloning vectors contain *ori* to allow stable replication.
 - Antibiotic resistance, β -gal, or other screening properties.
 - β -gal breaks down X-gal into blue product.
 - Polylinker site (several restriction enzyme targets in a row) for insertions.
 - Promoter sequence to drive expression of genes cloned into the plasmid.

I. Restriction Mapping

- Allows for physical map (by nt), instead of genetic one (by recombination frequency).
- Location of restriction enzyme target can be found by running digested pieces on a gel and looking at their sizes.
- Ambiguities can be resolved by multiple digestion.



J. Southern Blot Hybridization

- Multiple fragments, esp. from chromosomal DNA, can appear as a smear of DNA.
- Cut DNA, run on a gel, transfer to nitrocellulose, denature (into ssDNA) and cross-link to the membrane by UV.
- Probe with complementary labeled ssDNA (^{32}P).

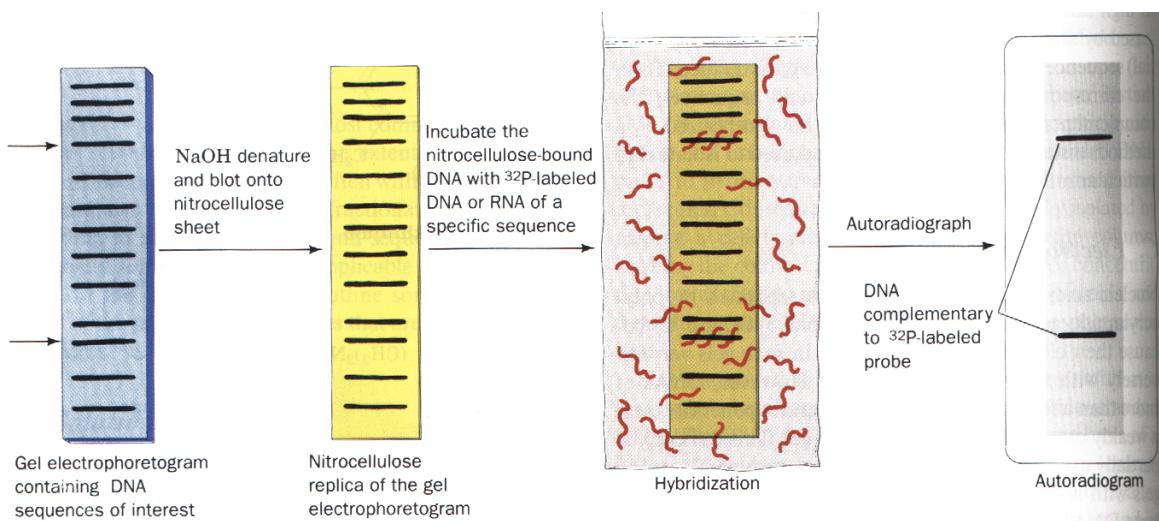


FIGURE 28-30. The detection of DNAs containing specific base sequences by the Southern transfer technique.

K. Restriction Fragment Length Polymorphism (RFLP)

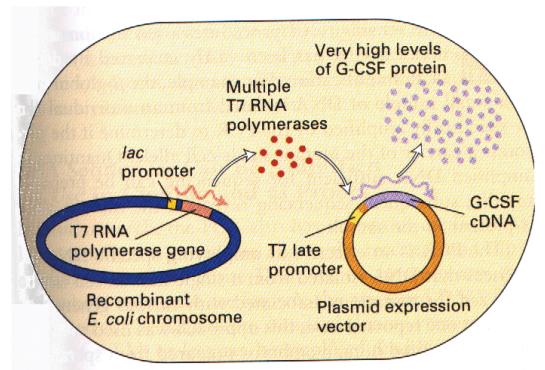
- Cut with enzyme and run fragments on a gel. Different target sites lead to different bands.
- RFLPs of highly linked genes are likely inherited together and can be used as diagnostic tools. (ideally <5% recombination frequency)
- PCR can amplify genomic DNA for RFLP

L. Random aside on Huntington's

- CAG repeat expansion results in lengthened poly-Gln sequence → misfolding → amyloid fibers.

M. Gene Overexpression

- Direct Expression: clone the gene into a vector with promoter and all. Usually by replacement.
- T7 RNAP: clone gene of interest after T7 late promoter and dual-express T7 RNAP (which may be in a different plasmid/controlled by another inducer – IPTG).

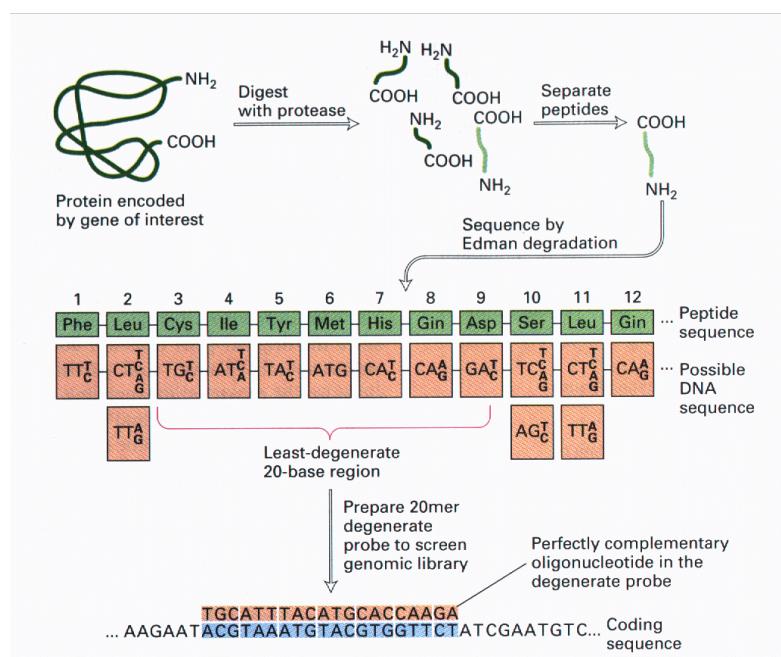


N. Fusion Proteins

- Add a gene coding for an epitope tag directly adjacent to your gene of interest.
- This may be GFP, protein A, (His)₆, B-gal, Glutathione-S-transferase (GST), Maltose-binding protein, Chitin binding domain, etc.
- IP may generate multiple bands (interacting proteins), which are identified by mass spec.

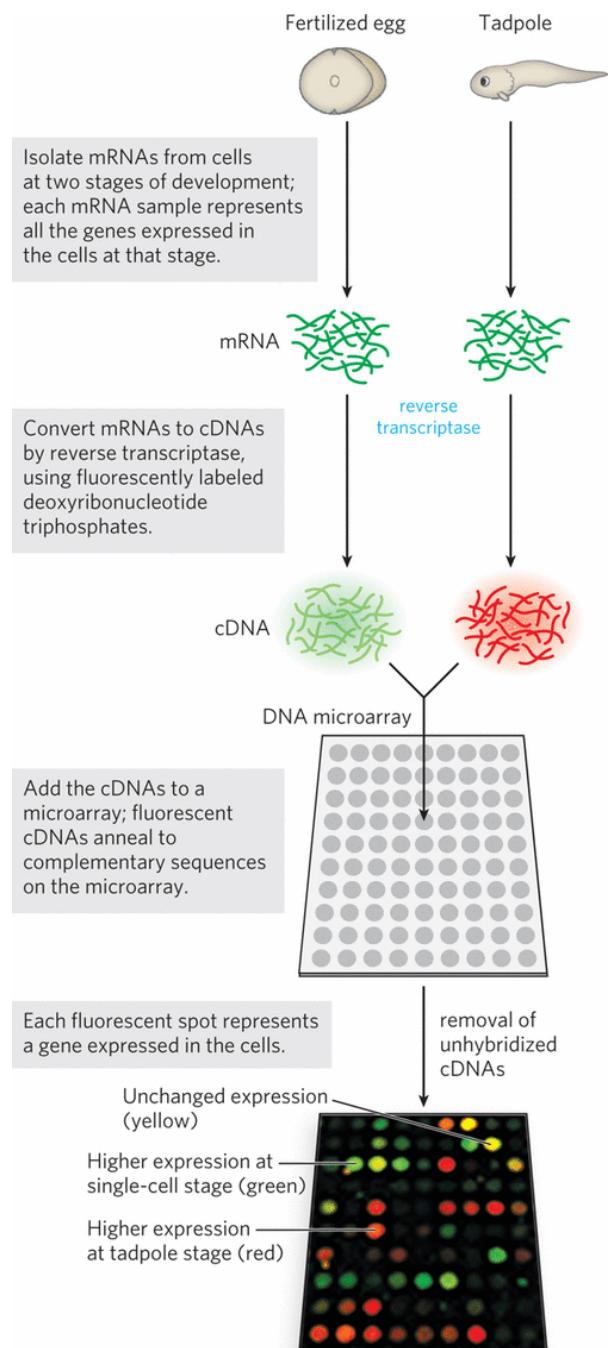
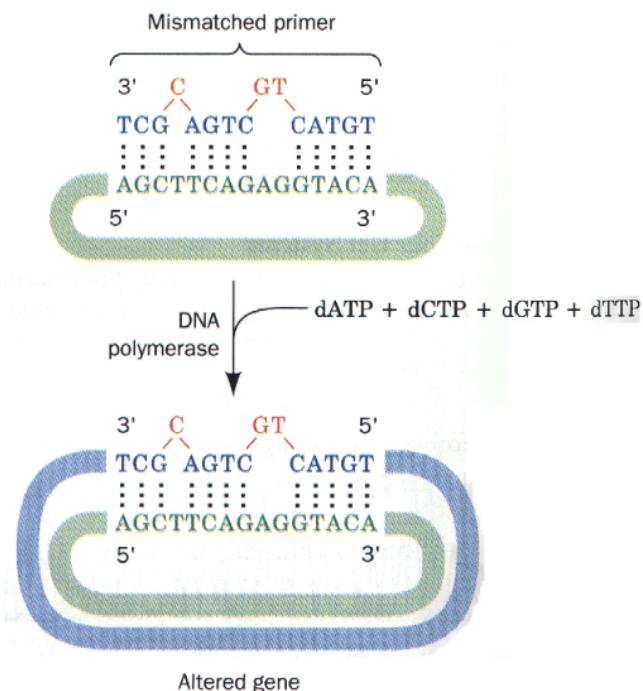
O. Oligonucleotides

- Degenerate oligos (nonspecific)
 - Collect protein, digest with protease, sequence peptides by Edman degradation.
 - Construct a plausible DNA sequence and use the least degenerate region to prepare 20-mer degenerate probes to screen the genome.



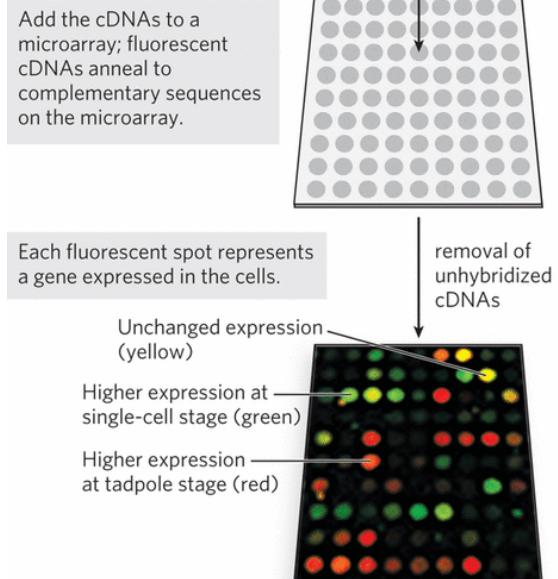
b. Site-Directed Mutagenesis

- Synthesize oligo with desired mutations.
- Hybridize to DNA encoding original gene (acts as a primer).
- Extend by DNA Pol I to generate the mutant gene.
- Insert into host to yield mutant DNA.
- More information: how do you separate/amplify it??



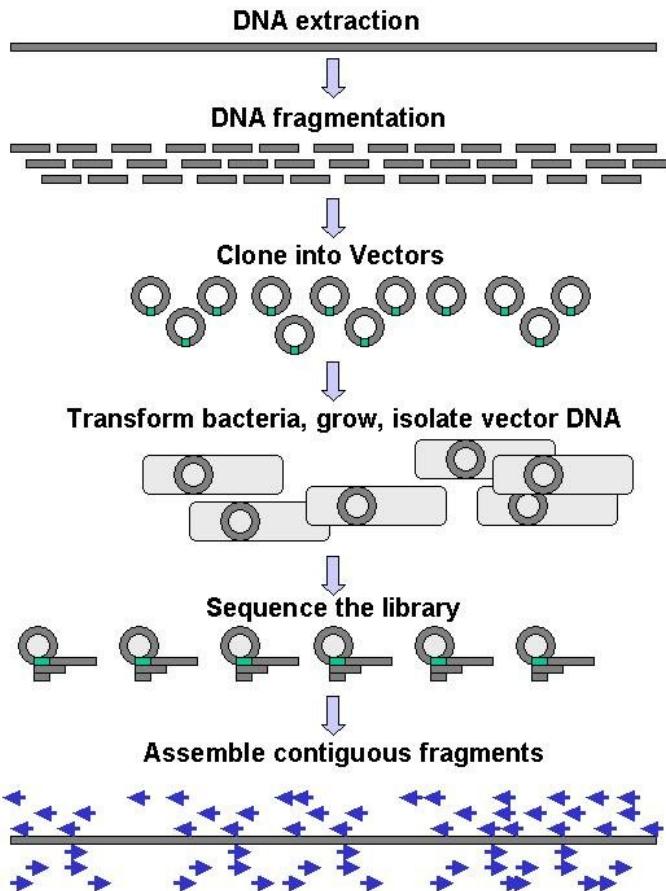
P. Microarrays and Gene Expression

- DNA chip has known sequences at specific positions on a grid.
- RNA → cDNA, then tagged with chromophore (green/red).
- Mix and anneal labeled cDNA to the chip.



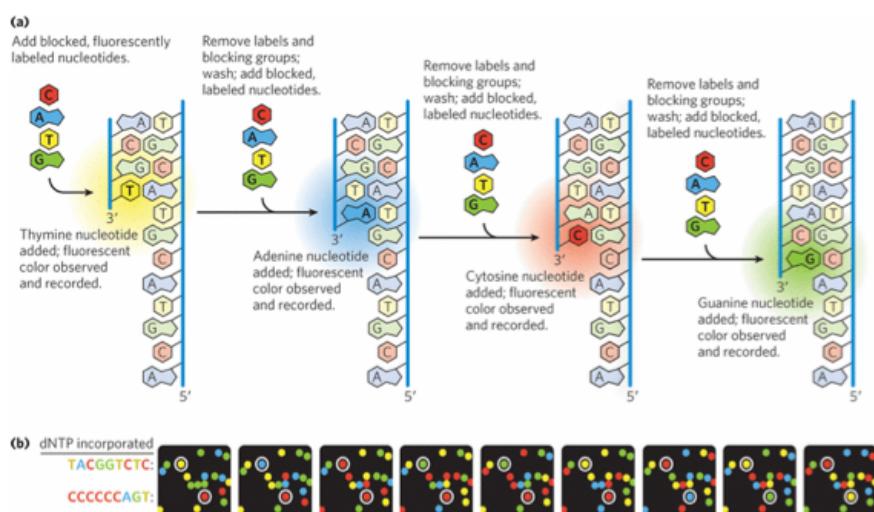
Q. Human Genome Project

- a. NIH: government and labs
- b. Celera: collection of automated DNA sequencers.
- c. Sequencing of ordered arrangement of cloned segments
- d. Shotgun sequencing and reassembly of overall sequence by computer alignment.



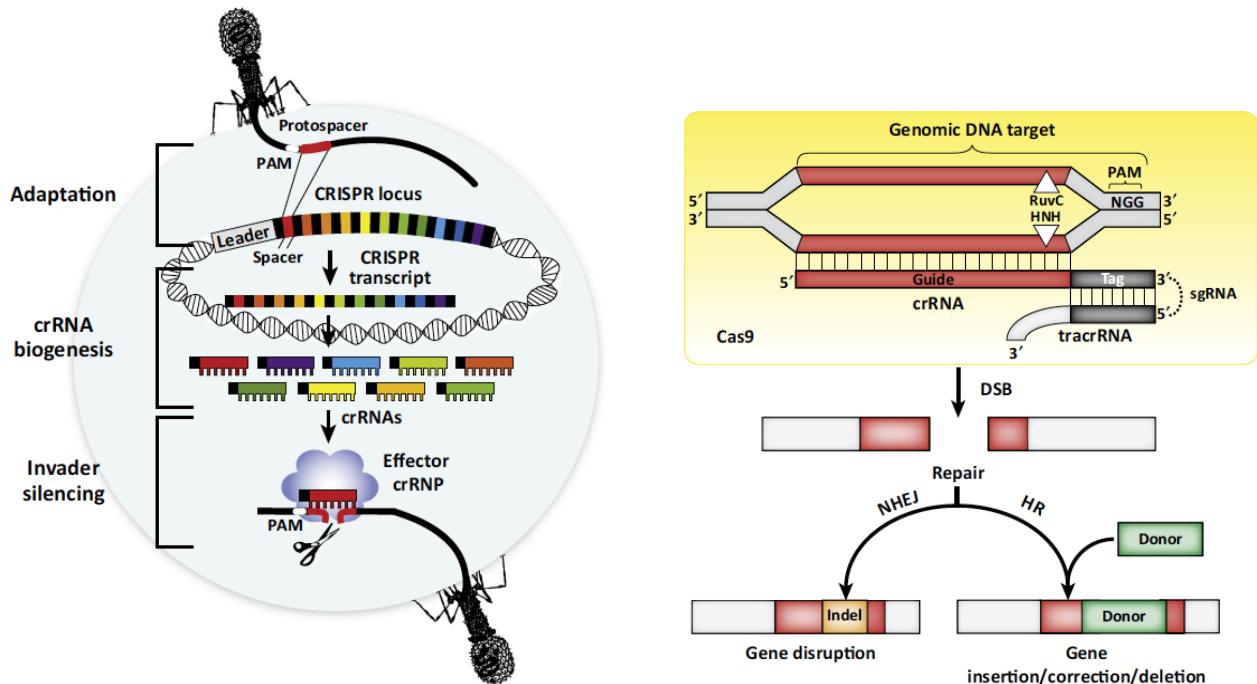
R. Next Generation Sequencing (AKA Massively Parallel Sequencing, Deep Sequencing)

- a. RNA-Seq: convert RNA to DNA, then use NGS.
- b. Whole genomes can be sequenced and aligned to a scaffold.



S. Genome Editing with CRISPR-Cas

- a. Natural Process to silence viruses and foreign genetic elements
 - i. Invader DNA is integrated into the host CRISPR locus near the leader.
 - ii. The CRISPR locus contains copies of a short direct repeat sequence (black) that separates invader-derived sequences.
 - iii. Cas genes adjacent to these loci (not shown).
 - iv. PAM (photospacer adjacent motifs) are next to photospacers (viral sequences for CRISPR integration).
 - v. CRISPR locus transcripts processed to multiple crRNAs.
 - vi. crRNAs associate with 1 of 6 Cas proteins to form crRNA-Cas effector complexes.
 - vii. crRNP recognizes foreign genetic elements and cleaves in the hybrid region.
- b. RNA-programmable genome editing by Type II Cas9 CRISPR-Cas systems
 - i. Cas9 endonuclease has two nucleases that cleave dsDNA to make blunt ends.
 - ii. crRNA contains a guide region that directs cleavage.
 - iii. Enzyme requires a short PAM adjacent to the region of homology as an auxiliary target identification signal.
 - iv. Cas9 also requires a 2nd RNA (trans-activating, or tracrRNA) that base-pairs with repeat-derived tag of the crRNA.
 - v. crRNA + tracrRNA are covalently linked into a single-guide RNA (sgRNA)
 - vi. Cas-9 induced DSB undergo repair by non-homologous end-joining (NHEJ) or homologous recombination.
 - vii. NHEJ repair results in indels (short insertions or deletions) and can be used to selectively disrupt or inactivate targets.
 - viii. When provided with homologous donor DNA, HR repair can affect sequence replacement to insert, modify, or delete sequences at the site.



Lecture 17: Eukaryotic Genome I

A. Special Eukaryotic Characteristics

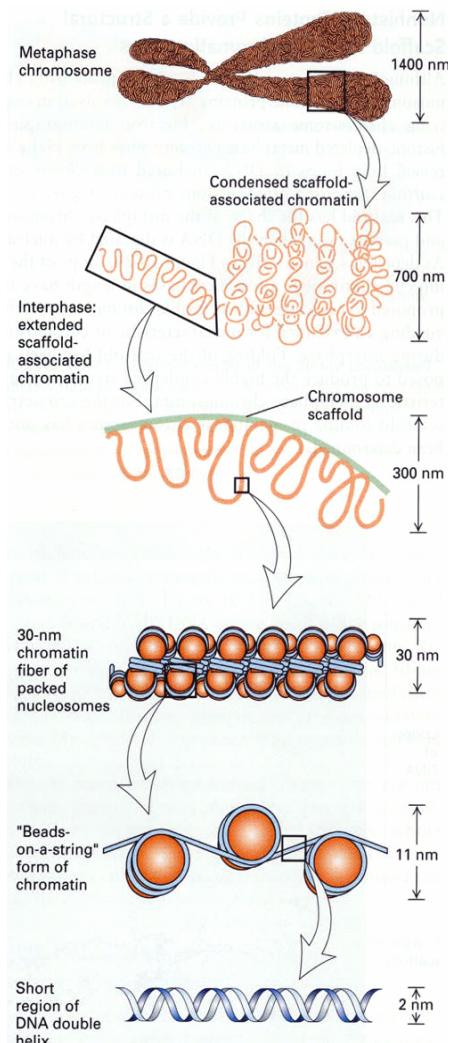
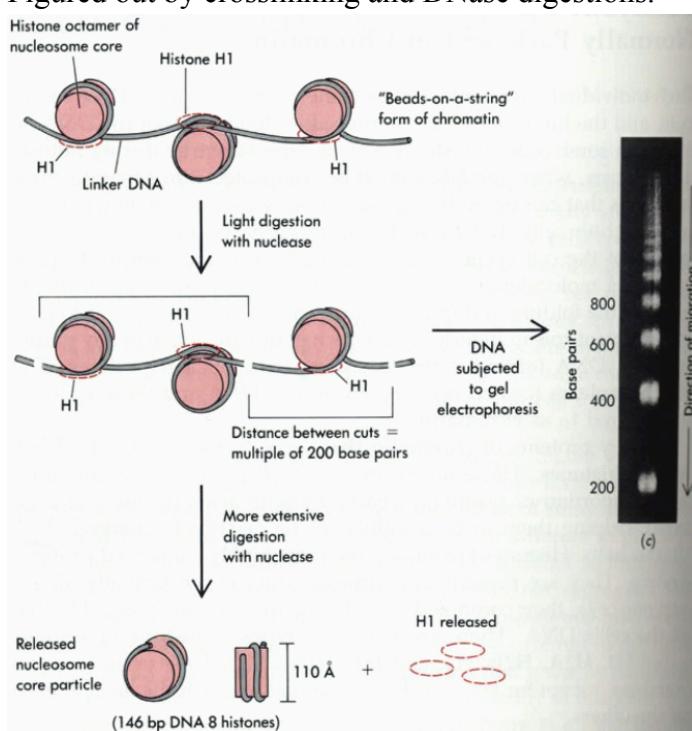
- a. Membrane-bound organelles (organelles, nucleus, mitochondria, ER, lysosomes).
- b. Fibrous systems and cytoskeleton
- c. Huge amount of DNA
 - i. C-value paradox: DNA per haploid cell is not proportional to # of genes.
- d. DNA in multiple linear chromosomes
 - i. Pulsed-field gel electrophoresis: alternatively pulses current at mirror angles.
Can separate very large DNA molecules, which take longer to change direction.
 - ii. Chromatin: G1, S, and G2.
Chromosomes: M.

B. Nucleosome and Higher Order Chromatin Structure

- a. 46 chromosomes with ~3 billion bp = 2 meters.
- b. Nucleosomes: DNA wraps around histones (100 Å).
- c. Nucleofilament: helical array of nucleosomes (300 Å).
Also: 30-nm chromatin-fiber.
- d. Metaphase chromosome: filaments arrayed on protein scaffold. TopoII is a major component of the scaffold.
- e. SMC (structural maintenance of chromosomes) proteins: cohesion and condensing.

C. Nucleosomes are DNA-protein beads on a string

- a. Protein octamer with 2x (H2A, H2B, H3, and H4).
- b. An additional H1 organizes ~200 bp.
- c. Figured out by crosslinking and DNase digestions.



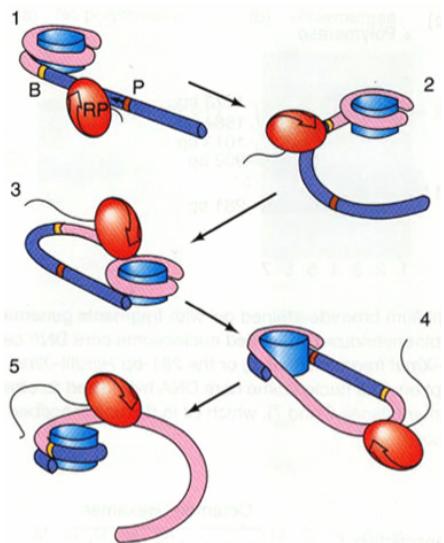
- d. Histone proteins are highly conserved:
 - i. Lys/Arg rich exterior: Basicity (+) binds DNA (-).
 - ii. Hydrophobic core for protein folding.
 - iii. DNA wraps around histones, -1 superhelical twist per nucleosome.
 - iv. 146 bp DNA wrapped around the core particle and 20-90 bp of linker DNA (between nucleosomes).

D. Nucleosome structure: solved by crystallography.

- a. H2A-H2B and H3-H4 heterodimers
- b. N-terminal histone tails are sites for modification.
Normally wrapped in minor groove; unwrapped by acetylation.

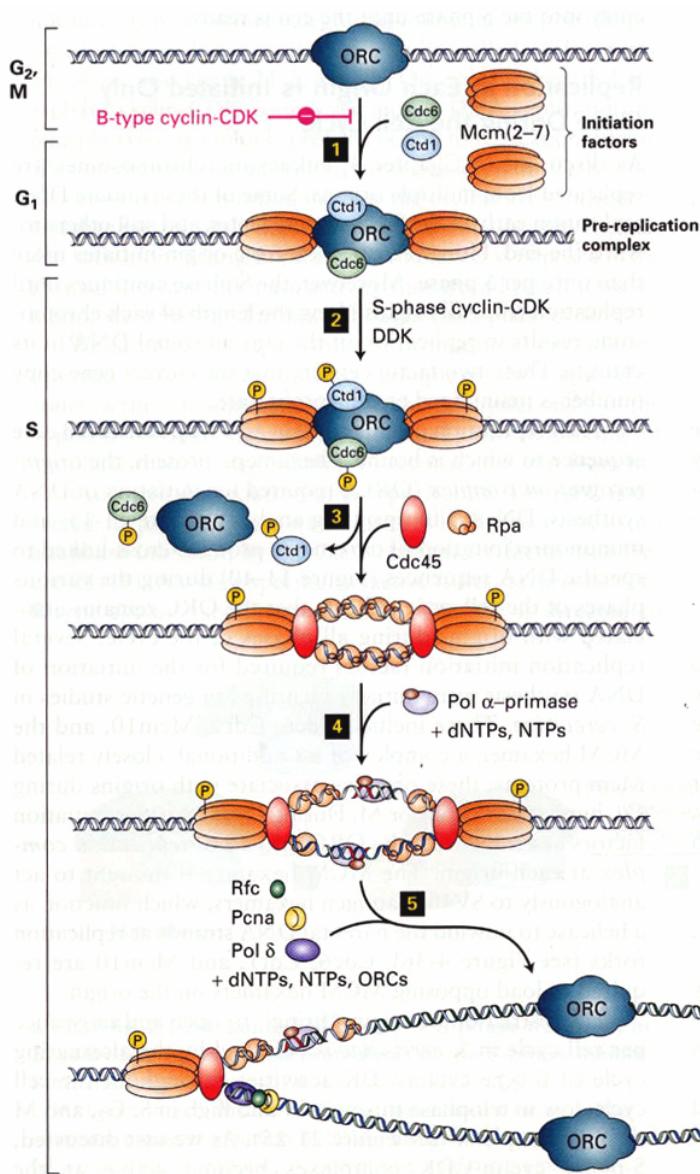
E. Nucleosomes and transcription

- a. 300 Å nucleofilament: H1 is on the inside, chromatin is inactive.
- b. Active chromatin is more sensitive to nucleases because it's less tightly wound.
- c. Hypersensitive sites near the beginning of active genes.
- d. Spooling model for nucleosome repositioning.



F. How are the giant eukaryotic DNA replicated?

- a. Multiple origins of bidirectional replication.
- b. Several different polymerases (for leading v. lagging).
- c. Controlled expression of CDKs helps coordinate replication so that each segment of DNA is replicated only once per cycle.



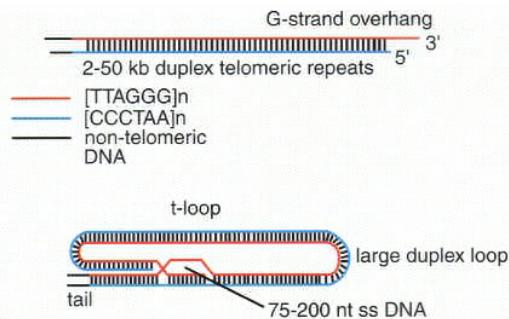
► FIGURE 21-26 Model for assembly of the pre-replication complex and its regulation by S-phase cyclin-CDK complexes in *S. cerevisiae*. Step **1**: During early G₁, unphosphorylated replication initiation factors assemble on an origin-recognition complex (ORC) bound to a replication origin to generate a pre-replication complex. Step **2**: In the S phase, S-phase cyclin-CDK complexes and DDK phosphorylate components of the pre-replication complex. Step **3**: This leads to binding of Cdc45, activation of the Mcm helicases, which unwind the parental DNA strands, and release of the phosphorylated Cdc6 and Ctd1 initiation factors. RPA binds to the resulting single-stranded DNA. Step **4**: A complex of DNA polymerase α (Pol α) and primase initiates the synthesis of daughter strands. Step **5**: DNA polymerase δ plus its accessory factors PCNA and Rfc elongate daughter strands initiated by Pol α -primase. ORC binds to the origin sequence in the daughter double-stranded DNA, but the phosphorylated initiation factors cannot assemble a pre-replication complex on it. B-type cyclin-CDK complexes maintain the initiation factors in a phosphorylated state throughout the remainder of S, G₂, and early anaphase (top). These factors cannot assemble into new pre-replication complexes until the B-type cyclins are degraded following their polyubiquitylation by the APC in late anaphase. Recent results indicate that additional proteins not shown function at step **3**. While the phosphorylation of several pre-replication complex components is represented, all the critical components whose phosphorylation by an S phase cyclin-CDK or by DDK is required for initiation have not been determined.

G. How are the ends of linear chromosomes replicated?

- The 3'-primer is removed following replication, leaving 3' overhangs.
- Telomeres are G-rich repetitive sequences at ends. These vary across species.
- Tetrahymena telomeres placed into yeast were extended with the yeast terminal sequence.
- Tetrahymena telomerase is an RNP!
 - Telomerase extends 3'-ends without a template.
 - Reverse transcribes built-in RNA template (1.5 telomeric repeats).

H. Telomeric Structures and Binding Proteins

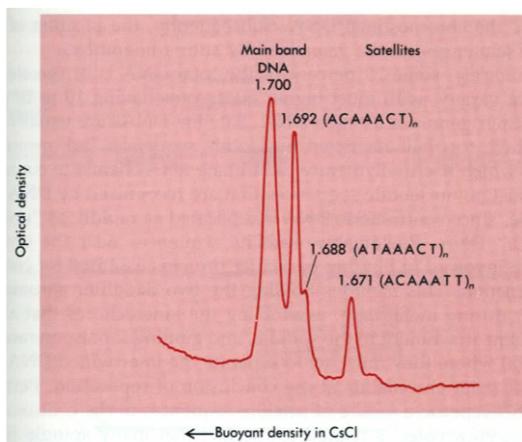
- dsDNA ends are recombinogenic and not stably inherited.
- All euk chromosomes have 3'-terminal overhangs with GT rich telomeric repeat.
- ss-polyG terminal sequences prevent end-joining by:
 - G-structures (quartets, tetraplexes) to knot the ends.
 - Recruit proteins to cap and prevent end-joining.
- T-loops: ss-G-rich stand invades telomeric duplex to cap the end.



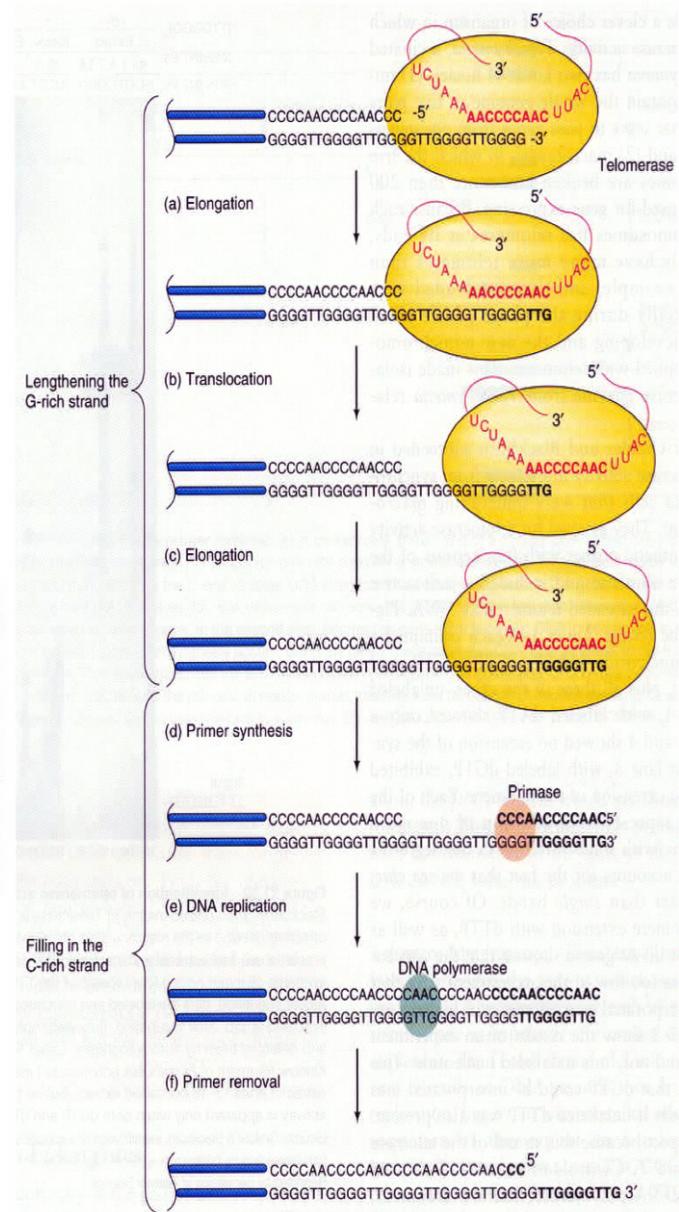
I. Telomerase activity is off in most normal cells, but appears in cancer cells.

J. Repeated Sequences at Centromeres

- Centromeres: where spindle fibers attach to metaphase chromosomes.
- Repeated sequences from centromeres appear as satellite bands (slightly fragmented total DNAs separated in CsCl gradients).



- In situ hybridization (with labeled cRNA) located these satellite sequences at centromeres.



K. rRNA/tRNA Genes Repeated

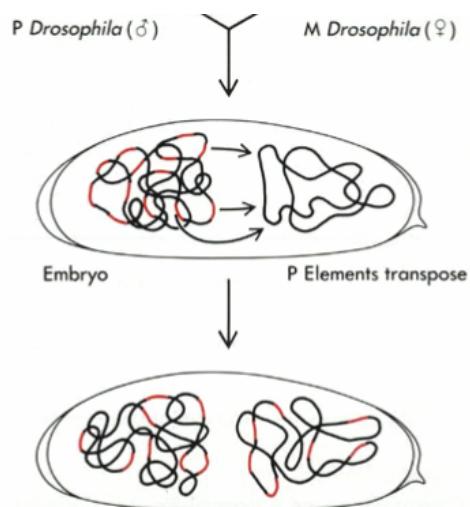
- a. rDNA repeats are clustered at several hundred copies.
- b. Protein-coding genes are single-copy:
 - i. Amplification at transcription and translation.
 - ii. Exception: histone proteins (need lots of them!).
- c. Fibroin and silkworms
 - i. 10^5 times more protein produced than mRNA transcribed.
 - ii. Isolated fibroin mRNA by synthesizing an oligo primer based on repeating AAs.
 - iii. Degenerate oligos and screen genome; hybridize to Northern blot.

L. Types and Origin of Genes in the Human Proteome

- a. Humans have only 23,000 genes, comparable to yeast and worms.
- b. Greater complexity from alternative splicing.
- c. Most human genes are not unique to humans.
- d. Many are from other sources; some from bacteria via horizontal gene transfer.

M. Transposons Inhabit Euk Genomes

- a. P-element is a cut-and-paste Drosophila transposon leading to “hybrid dysgenesis” (hybrids messed up because P elements disrupt vital genes).
- b. Expressed only in germline cells because of transposase requires tissue-specific splicing.
- c. P-element transformation:
 - i. Gene of interest inserted between ends of P element
 - ii. Inject P element into M (non-P) eggs or embryos, along with transposase.
 - iii. Insertions at multiple sites in Drosophila chromosomes.



N. Are repetitive sequence elements selfish DNA?

- a. Throughout time, wild drosophila increasingly possess P elements.

O. Human genome has lots of mobile elements.

- a. LINES (long interspersed repeat elements) include L1 retrotransposon.
- b. SINES (short interspersed repeat elements) include Alu sequences.

P. Y chromosome is unusually young.

- a. Constant rate of mutation within repetitive elements (because no selection pressure).
- b. Molecular clock: comparing sequences from 2 sources to that of a common ancestor can give age estimates.
- c. Y chromosome is largely repeated elements; has a higher mutation rate than the X chromosome; looks unusually young.

Q. Alu and L1 sequences can be mutagenic:

- a. Alu-alu recombination triggers chromosomal rearrangement.
- b. De novo insertions into genes.

Lecture 18: Eukaryotic Genome II

A. Eukaryotic mRNA post-transcriptional modifications:

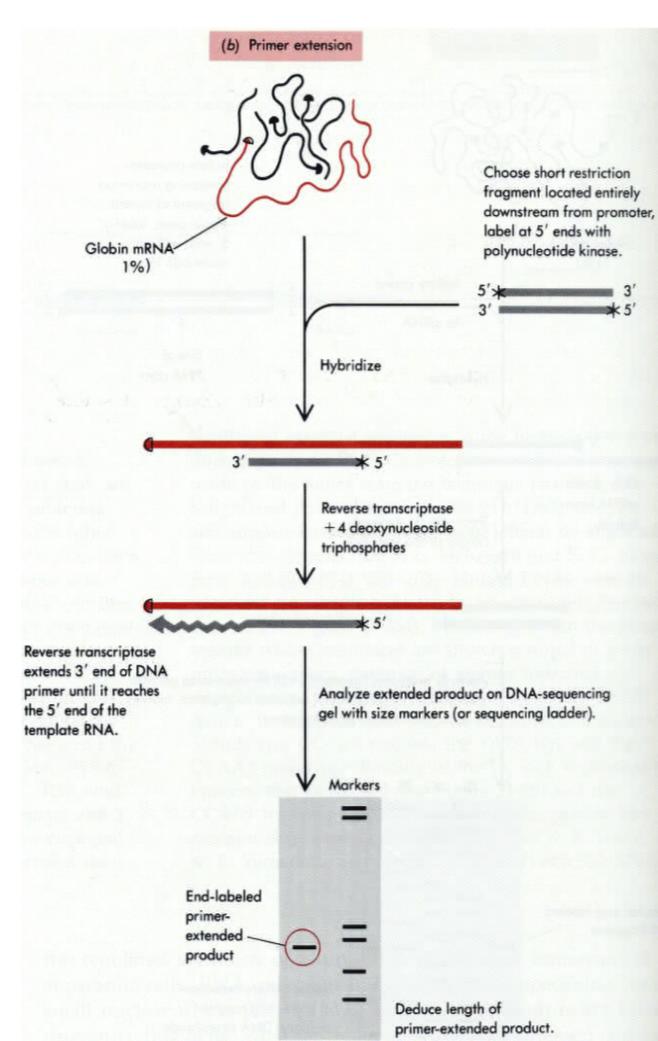
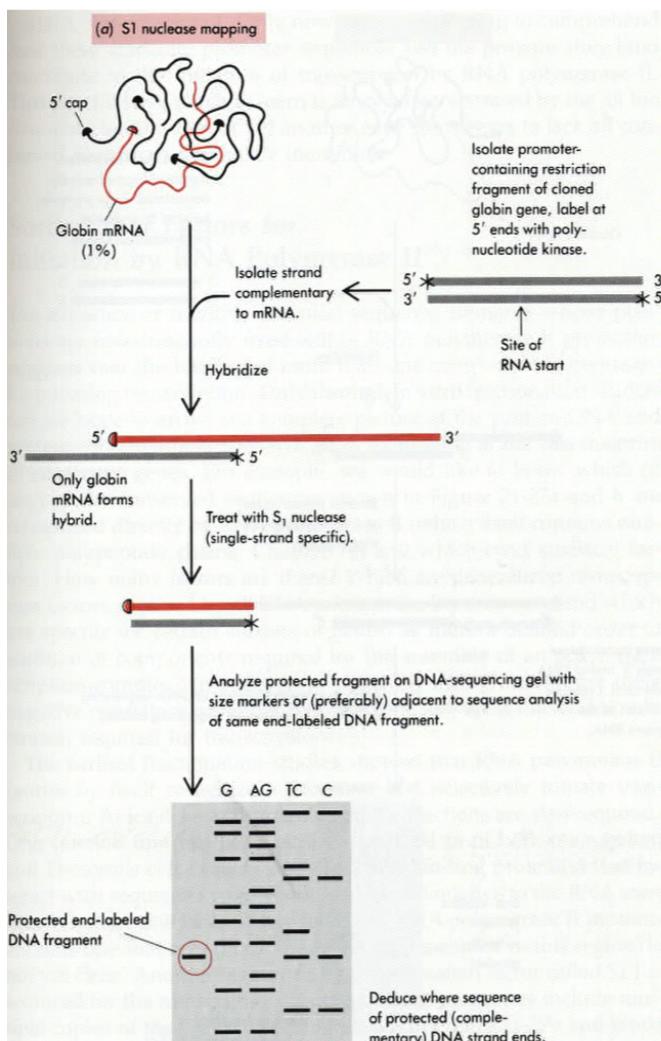
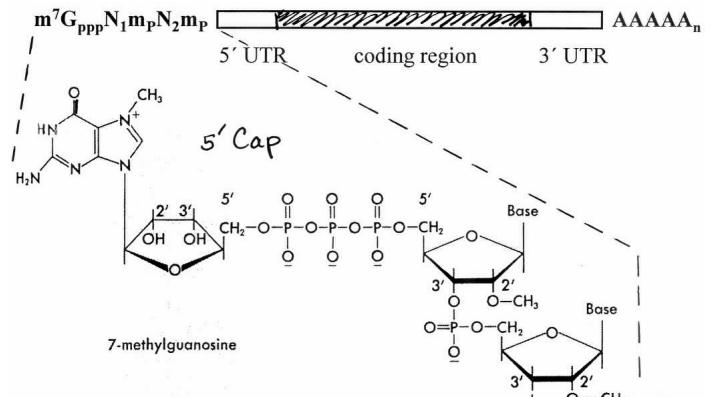
- 5' and 3' UTRs
- 5'-cap: 7meG (prevents degradation)
- 3'-polyA tails: ~200 nt added by polyA polymerase (PAP).
 - Isolate mRNA with oligo-dT.
 - Template cDNA synthesis.

B. Heterogeneous nuclear (hn)RNA/pre-mRNA

- 10x the size of mRNA.
- Have 5'-caps and polyA tails.
- Short-lived: half life of 3 min.
- 10% labeled hnRNA (exons) appears in cytoplasmic mRNA.

C. Mapping Ends and Intron-Exon Boundaries in mRNA

- S1 nuclease mapping

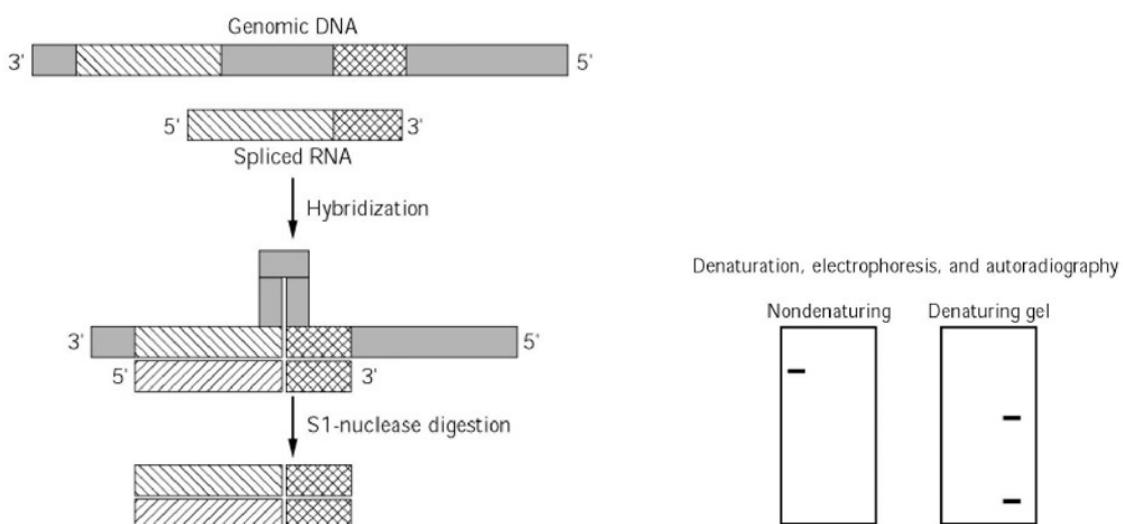


- b. S1 nuclease mapping
 - i. Isolate 5'-labeled restriction fragment that contains the mRNA 5'-start.
 - ii. Hybridize to mRNA, digest with S1.
 - iii. Can also locate mRNA 3'-end by labeling the 3'-end of the probe instead.
 - iv. Requires detailed knowledge of gene structure and sequence data.

- c. Primer extension
 - i. Isolate 5'-labeled restriction fragment downstream of start site.
 - ii. Hybridize to mRNA, extend with RT.
- d. Final gel
 - i. Length of ds segment = distance of mRNA start site from labeled 5'-end.
 - ii. Use size markers or sequencing ladder (higher resolution).
- e. RNase mapping is a variant of S1 mapping where a labeled RNA is used (instead of DNA).

D. S1 nuclease mapping of exons

- a. Hybridize mRNA to genomic DNA; the introns loop out and are digested by S1.
- b. Denature/melt the strands and run the DNA exons.

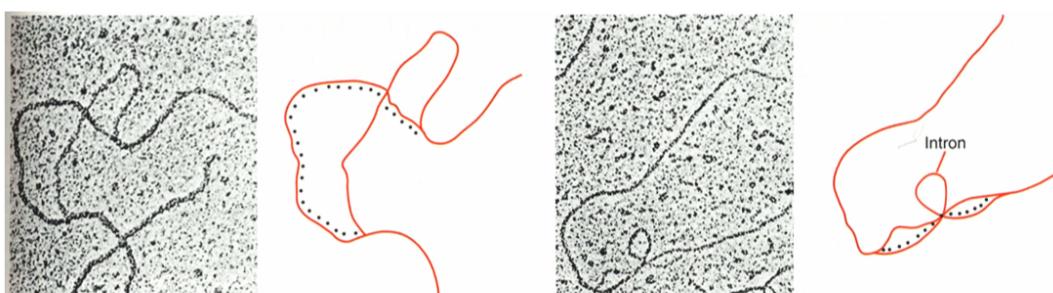


E. R-Loop Mapping Detects Introns

- a. Cloned gene DNA is incubated with mRNA in formamide (denaturant) at DNA T_m .
- b. mRNA displaces coding strand, forming an R loop.
- c. DNA-RNA is more stable because of the A-form has better base stacking.

F. Proving that introns are transcribed into RNA

- a. R loops of spliced mRNA: introns are forced to loop out.
- b. R loops of hnRNA: introns bp naturally.



G. Intron-Exon Structure of Eukaryotic Genes

- a. R loops can give an idea of intron size and location.
- b. Northern blotting to find pre-mRNA:
 - i. Isolate RNA-polyA, run on gel, transfer to blot, probe with labeled cDNA.
 - ii. Top band is primary transcript.
 - iii. Nuclear RNA bands are splicing results.

H. Exon and Intron Sizes in the Genome

- a. Most organisms have small exons (~150 nt).
- b. Humans have smaller exons and much bigger introns.

I. Things of interest

- a. Exons sometimes correspond to protein-folding domains.
- b. **Intron positions are conserved- shows that parsing the gene according to protein structure is selected for.**
- c. Ratio of non-coding to coding DNA rises with developmental complexity.

J. When did Introns originate?

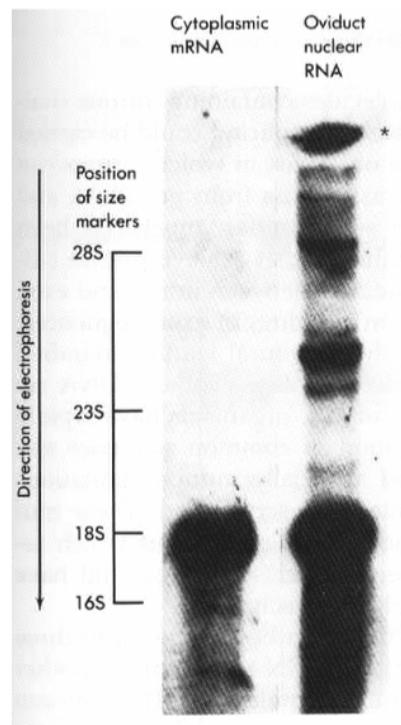
- a. Theory 1: early genomes had introns, and rapidly replicating organisms got rid of them (++generation times).
- b. Theory 2: introns have jumped into genes later.
- c. **Evidence for Early Introns**
 - i. Exon shuffling: same exon sometimes appears in different proteins because of recombination between exons (within introns).
 - ii. Exon sliding: new splice junctions can add new “exon” material.

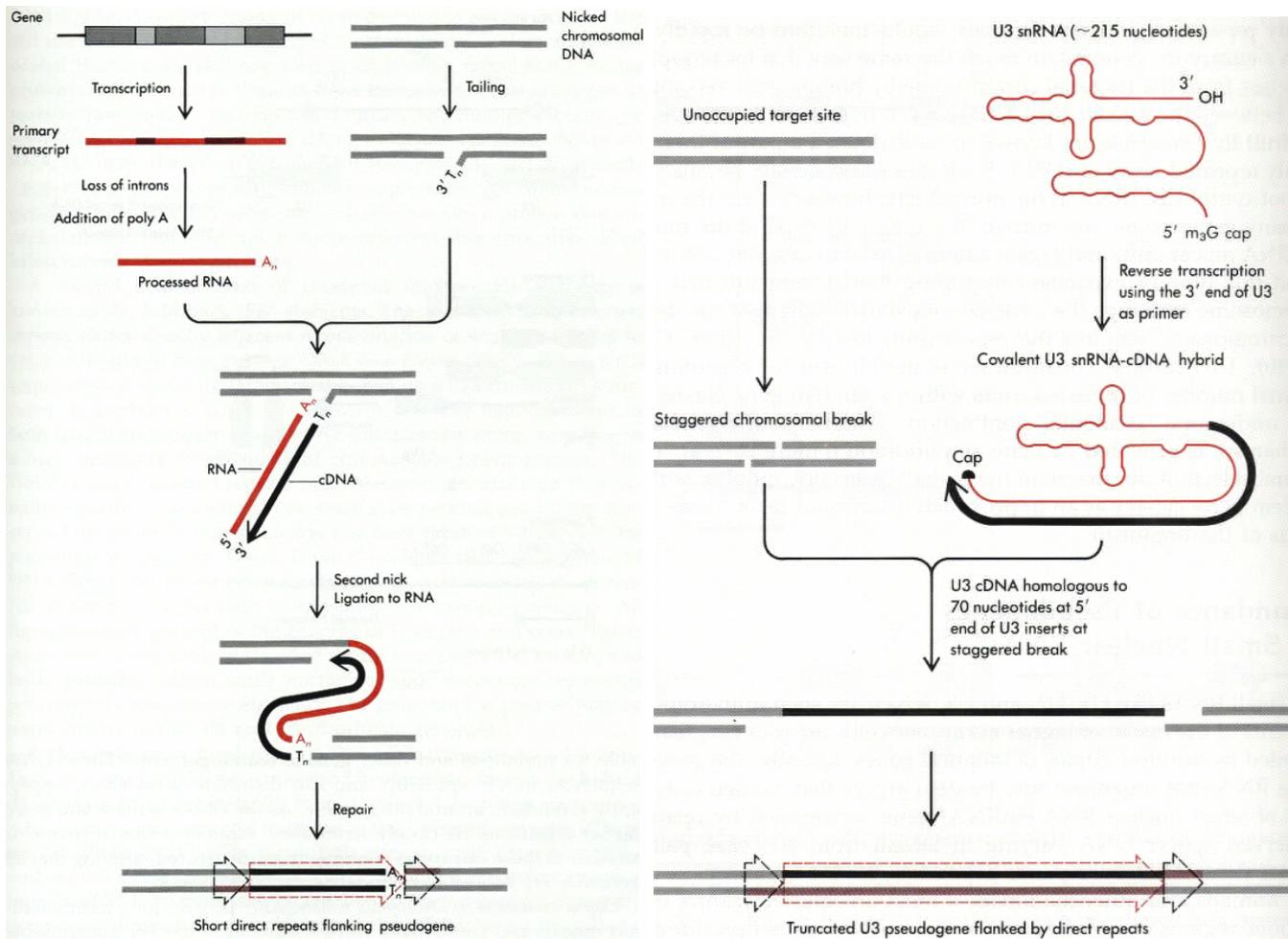
K. Creation of Pseudogenes

- a. Processed pseudogene
 - i. Nicked chromosomal DNA is tailed (with T).
 - ii. The RNA polyA pairs at the nick and reverse transcribes cDNA.
 - iii. Second nick allows incorporation (RNA ligated to DNA, gaps repaired).
 - iv. Staggered nicks create short direct repeats.
 - v. Results in spliced replica of protein-coding gene; often has polyA sequence.
- b. RNA pseudogene
 - i. The RNA (e.g. U3 snRNA) is reverse transcribed using the 3'-end as a primer.
 - ii. The cDNA inserts at a staggered double nick, and gaps repaired with direct repeats.
 - iii. Only makes a partial replica (5'-end of RNA).
- c. Processed pseudogenes are highly abundant in mammals (10X more than real genes!!!)
- d. No promoters on processed pseudogenes, but has polyA tail. Alu has promoter (RPol III). RNA pseudogenes have neither promoters nor polyA.

L. Trinucleotide repeat expansion and disease

- a. DNA Pol slips
 - i. Copies too much: new strand loop, increases repeats.
 - ii. Copies too little: old strand loop, decreases repeats.
- b. May lead to messed up splicing or protein structure.





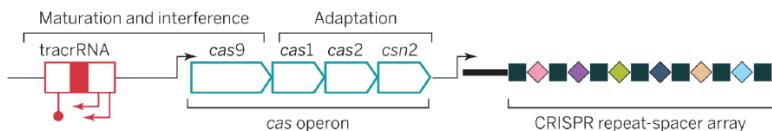
M. CRISPR: clustered regularly interspaced palindromic repeat

- CRISPR 30-40 bp “spacers” are derived from DNA of previous invaders.
- Intervening 30-40bp “repeats” are self.
- Resulting CRISPR (cr)RNAs guide Cas proteins (DNA endonucleases) to their targets.
- Many CRISPR/Cas systems - most are RNA guided and DNA-cleaving.
The crRNA hybridizes with DNA to form an R-loop.
- The PAM site (usu. NGG) is required for targeting by Cas9; absent in self DNA!

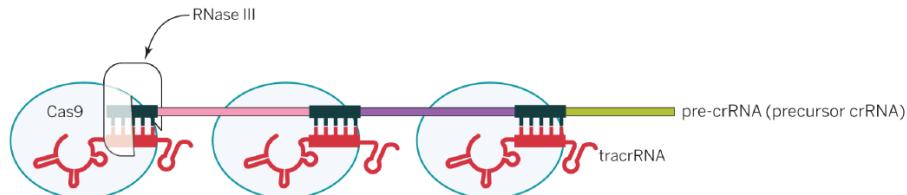
N. CRISPR for therapeutics

- Designed single guide RNA (sgRNA) find identify complementary dsDNA next to a PAM.
- sgRNA contains RNA to recruit Cas9.
- Cleavage produces a break; which is repaired in 2 ways:
 - Homology directed repair (HDR): replaces a target sequence to revert unwanted mutations, generate new alleles, insert useful domains, and introduce transgenes.
 - Non-homologous end-joining (NHEJ): leaves small insertions/deletions to inactivate the gene.

A Genomic CRISPR locus



B tracrRNA:crRNA co-maturation and Cas9 co-complex formation



C RNA-guided cleavage of target DNA

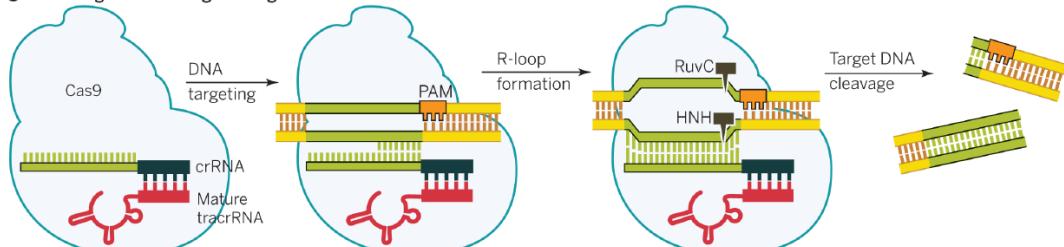
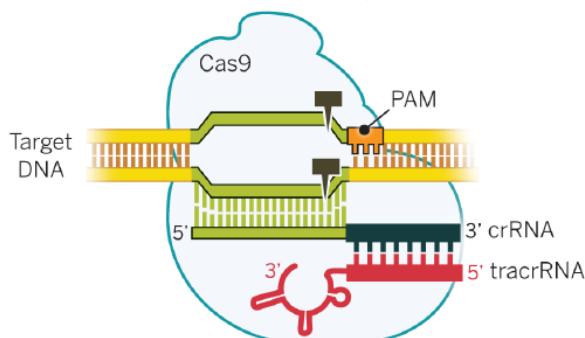
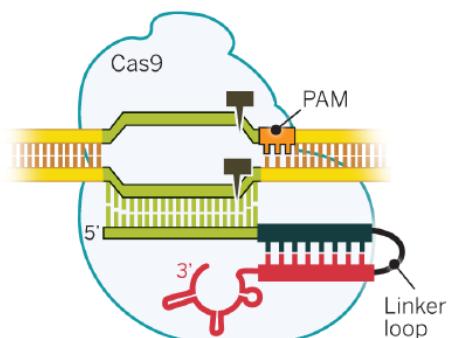


Fig. 2. Biology of the type II-A CRISPR-Cas system. The type II-A system from *S. pyogenes* is shown as an example. (A) The *cas* gene operon with *tracrRNA* and the CRISPR array. (B) The natural pathway of antiviral defense involves association of Cas9 with the antirepeat-repeat RNA (*tracrRNA:crRNA*) duplexes, RNA co-processing by ribonuclease III, further trimming, R-loop formation, and target DNA cleavage. (C) Details of the natural DNA cleavage with the duplex *tracrRNA:crRNA*.

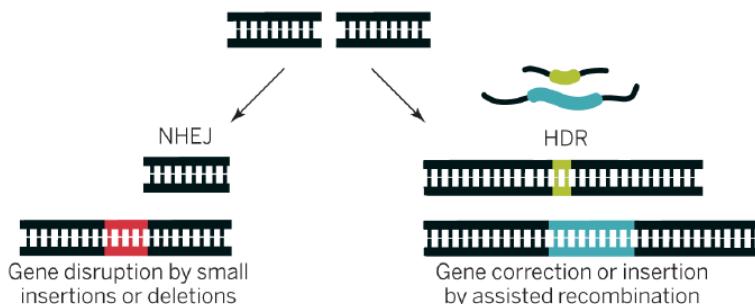
Cas9 programmed by crRNA:tracrRNA duplex



Cas9 programmed by single guide RNA



sgRNA-Cas9 double-strand DNA break



O. Too much DNA in Eukaryotes SUMMARY

- Human genome is only 1.5% coding!
- Introns make genes 1000-fold larger.
- Repeated gene families exist for tRNAs, rRNAs, snRNAs, and histones.
- Many pseudogenes resemble real genes, but do not encode gene products.
- Repeated sequence elements at centromeres and telomeres build structures critical for chromosome maintenance.
- Selfish DNA is half the human genome!
 - Includes transposons and retrotransposons (Alu and L1 sequences)
 - 100 L1s in the human genome.
 - Codes for RT and endonuclease. Nicks and initiates target-primed reverse transcription.

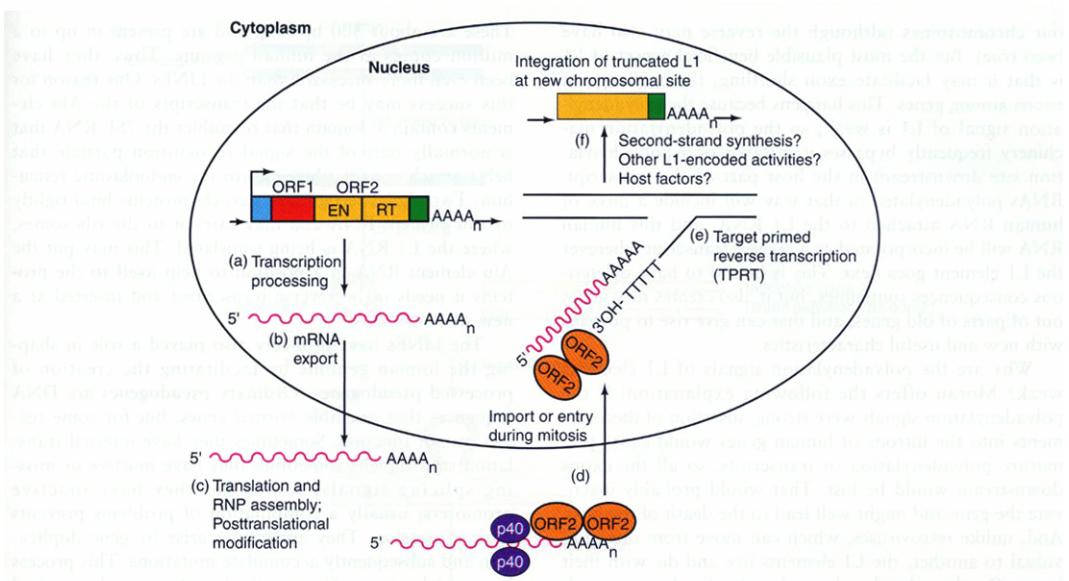


Figure 23.40 A model for L1 transposition. (a) The L1 element is transcribed and processed. (b) The transcript is exported from the nucleus. (c) The mRNA is translated to yield the ORF1 product (p40) and the ORF2 product, with endonuclease and reverse transcriptase activities. These proteins associate with the mRNA. (d) The ribonucleoprotein reenters the nucleus. (e) The endonuclease nicks the target DNA (anywhere in the genome), and the reverse transcriptase uses the new DNA 3'-end to prime synthesis of the reverse transcript. (f) In a series of unspecified steps, the second L1 strand is made and the whole element is ligated into the target DNA. (Source: Reprinted from Kazazian and Moran, "The impact of L1 retrotransposons on the human genome," *Nature Genetics* 19:19–24, 1998. Copyright 1998, with permission from Nature Publishing Group.)

Lecture 19: Eukaryotic Transcription

A. Eukaryotic v. Prokaryotic Transcription

- a. Euk RNAP cannot transcribe on their own; multiple proteins involved.
- b. Euk have 3 RNAPs.
- c. Prok have 1 RNAP, regulated by switching sigma factors (recognize different promoters).

B. Gene regulation in mammalian cells

- a. Regulation in processing and stability.
- b. 5-10k housekeeping genes.
- c. Differential expression of 100-1000s of specific genes.

C. James Darnell: Nuclear run-on experiment.

- a. Nascent transcripts in isolated nuclei are elongated with pulse of labeled NTPs.
- b. Extract RNAs and hybridize to nitrocellulose membranes with fixed cDNAs.
Wash off lone RNAs, and visualize with autoradiography.
- c. Shows how gene transcription is different across cell types.
- d. Don't forget positive (housekeeping) and negative (nonmatch) controls!

D. Bob Roeder: discovered 3 Euk RPol

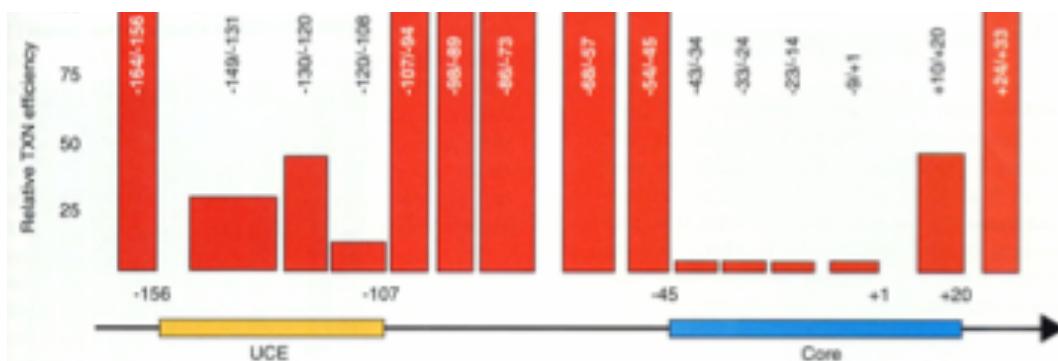
- a. Took cell extracts and separated into fractions by anion exchange chromatography.
- b. Fractions were assayed for ability to transcribe boiled DNA.
- c. RPol I: pre-rRNAs (nucleolus).
- d. RPol II: mRNA precursors and snRNA.
- e. RPol III: tRNA and 5S rRNA.

E. RPol I-III have 12-16 subunits

- a. Largest subunits homologous to the bacterial RNAP β' (polymerase activity).
In Pol II, this subunit has the CTD.
- b. 5 common subunits: others are homologous to β , 2x alpha, and omega.
- c. TBP: TATA-binding protein, for initiation.

F. RPol I

- a. Nucleolus: rRNA is made, processed, and assembled.
- b. Most abundant RPol; recognizes 1 promoter and 1 terminator.

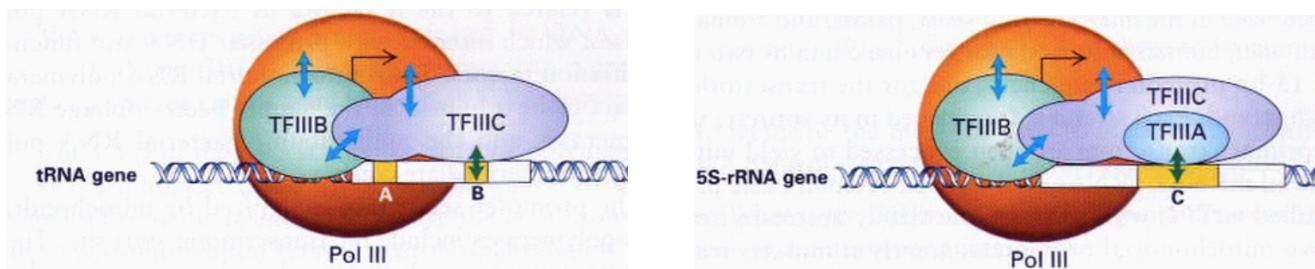


- c. Upstream control element (UCE) and core element are important for transcription.
 - i. UCE makes it optimal, but core is absolutely required.

- d. SL1 (selectivity factor 1):
 - i. Complex of 4 proteins (incl. TBP)
 - ii. Upstream binding factor (UBF): acts with SL1 to stimulate RPol I transcription.
 - iii. Strengthens RPol I-UCE binding, and opens a DNase-sensitive site in core element.
- e. Promoter elements and transcription factors are species-specific.
- f. RPol I is highly efficient; clearing the promoter is rate limiting.

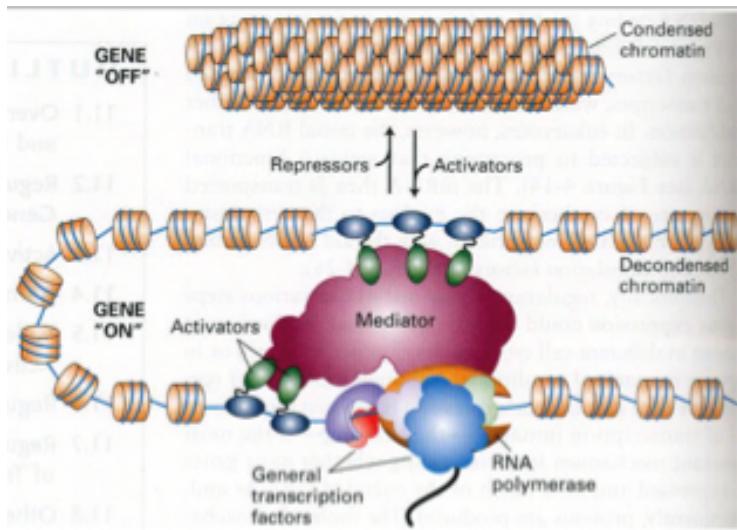
G. RPol III

- a. Transcribes small ncRNAs: pre-tRNAs, 5S rRNA, U6 snRNA, and Alu.
- b. Most promoters for Pol III are transcribed. Some upstream (U6 promoter), others internal.
- c. Multiprotein transcription factors involved:
 - i. TFIIIB contains TBP. Binds upstream of the start site. Absolutely required.
Unable to bind DNA itself (like RNAP sigma).
 - ii. A/B/C boxes: internal promoter elements.
 - iii. tRNA genes: A/B boxes recruit TFIIIC recruits TFIIIB recruits RPol III.
 - iv. rRNA genes: A/C boxes recruit TFIICIA recruits TFIIIC...
- d. Termination at Pol III occurs at polyT runs of TTTT... residues.
Transcripts must end with ...UUUU-OH-3'

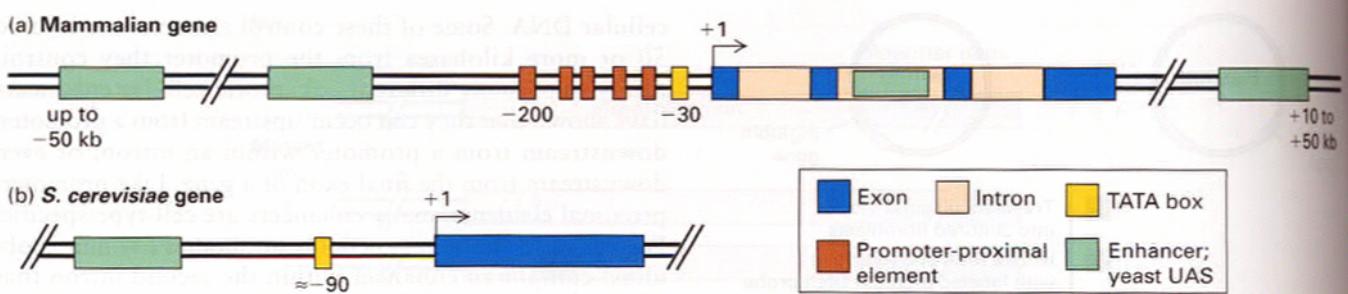


H. RPol II

- a. Transcribes mRNAs and snRNAs
- b. General transcription factors (basal TFs) bind at/near start site.
Minimal set for site-specific initiation by Pol II (e.g. TFIIA/B/D/E/F/H).
- c. Transcriptional activators: bind distant sites
 - i. Modular: DNA-binding, activation, and flexible protein domains in any order.
 - ii. Grouped by DNA-binding domain: e.g. homeobox, HLH, Zn finger, etc.
- d. Control Elements
 - i. Promoter: sequences close to the start site recognized by Pol II and TFs.
 - ii. TATA box: 25-30 bp upstream of start site.
 - iii. Weak sequence conservation at start sites.
 - iv. Promoter-proximal elements:
 1. Bind transcriptional activators (like SP1).
 2. May occur downstream of start site!
 - v. Enhancer sequences:
 1. Interact with gene-specific transcriptional activators.
 2. Can be >10 kB away, within introns, upstream or downstream.
 3. Silencers are enhancers that downregulate a gene.



TFIIB	Stabilizes TBP binding; recruits RNAP II; influences start site selection
TFIID	Recognizes TATA box; recruits
TBP	TFIIA and TFIIB; has positive and negative regulatory functions
TAFs	
TFIIE	An $\alpha_2\beta_2$ heterotetramer; recruits TFIIH and stimulates its helicase activity; enhances promoter melting
TFIIF	Facilitates promoter targeting; stimulates elongation
TFIIH	Contains an ATP-dependent helicase that functions in promoter melting and clearance



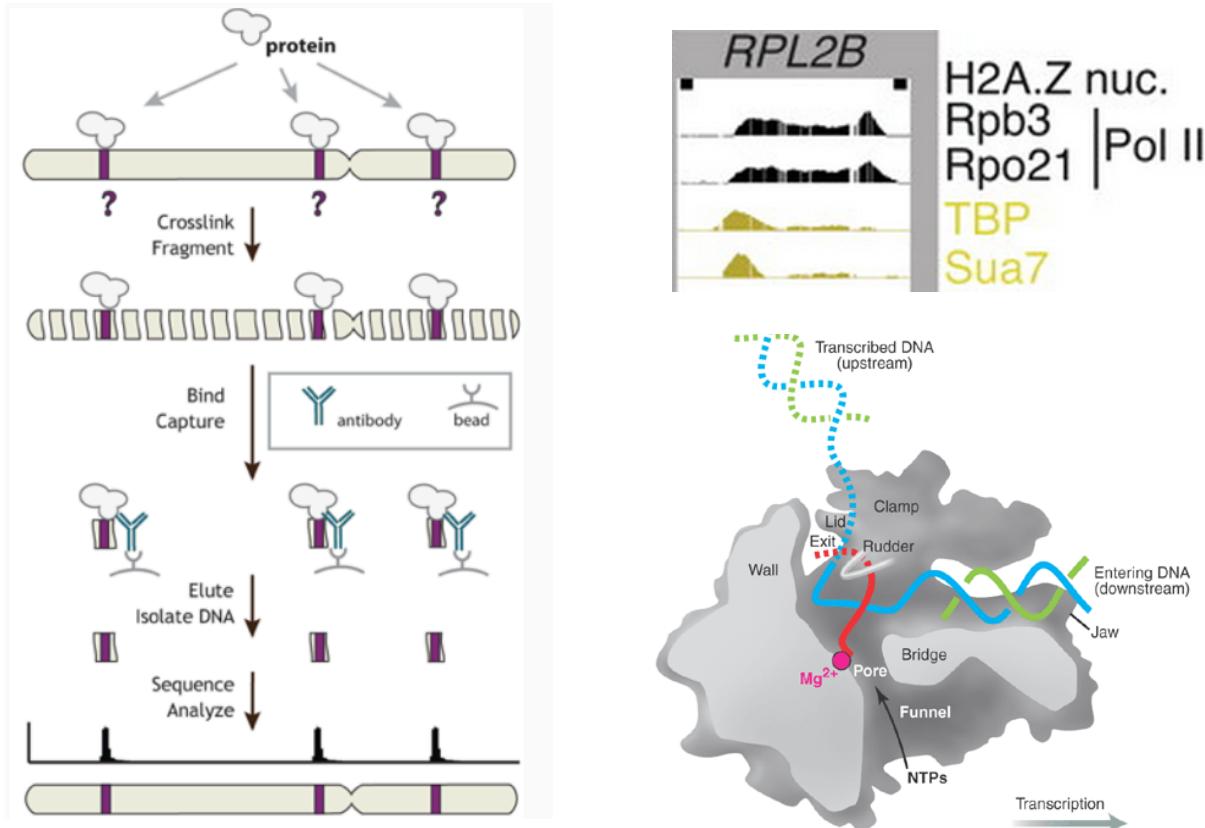
I. Assembling RPol II Complex

- TAF: TBP-associated factor
- TFIID: contains TBP and TAFs
 - Contacts the TATA box (with TBP) and the start site (with TAFs).
 - Some promoters lack a TATA box and engage TBP via other proteins instead of direct DNA binding.
- TBP: binds TATA box
 - Bends the DNA to open the minor groove and destabilize the helix.
Allows transcription from promoter on naked DNA. TFIID required *in vivo*.
 - Protein-DNA interaction via β -sheet.
 - Binds as monomer with pseudo-dimeric symmetry.
- Assembly Steps
 - TFIID binds TATA box with TBP.
 - ??? TFIIB binds downstream of the TATA box and stabilizes TBP.
Defines direction of transcription.
 - Pol II, TFIIF, TFIIIE bind.
 - TFIIH contains ATP-dependent DNA helicase; creates transcription bubble.
Also: kinase for CTD.

J. Activation of RPol II by Basal Factors

- RPol II CTD: unstructured; multiple tandem repeats of X-Ser-X-X-Ser-X-Ser.
- CTD is phosphorylated on Ser by kinases on TFIIH.
This stabilizes RPol II's interactions with TFs.

K. Chromatin Immunoprecipitation (ChIP) identifies the DNA-binding sites of proteins.



L. Structures of elongating yeast RPol II

- a. DNA bent 90°.
- b. Cramer and Kornberg solved yeast RPol II
- c. Trigger loop and Mg²⁺ coordinate dNTP
- d. Bridge promotes RNA-DNA base-pairing and translocates the active site.
- e. Rudder directs nascent RNA to exit channel.

M. Enhancers

- a. 50-200 bp in length.
- b. Multiple binding sites for distinct transcriptional activators, which interact synergistically.

N. Enhanceosome Model

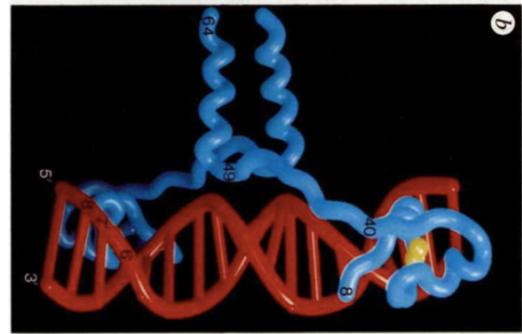
- a. Protein complex assembled at enhancers.
- b. Framework proteins stabilize bent DNA and protein-protein interactions.
- c. Modular Structure of Transcriptional Activators
 - i. DNA-binding domains: many motifs.
 - ii. Multimerization domains: where proteins join.
 - iii. Activation domains:
 1. Interact with basal transcriptional machinery.
 2. Less structured, residue redundancy.
 3. Acidic residues (Asp/Glu)
(negative noodles, acid blobs).

O. Types of DNA-binding domains

- a. Helix-loop-helix
- b. Homeodomain (helix-turn-helix binding)
- c. Zn finger
- d. Basic zipper (bZIP)

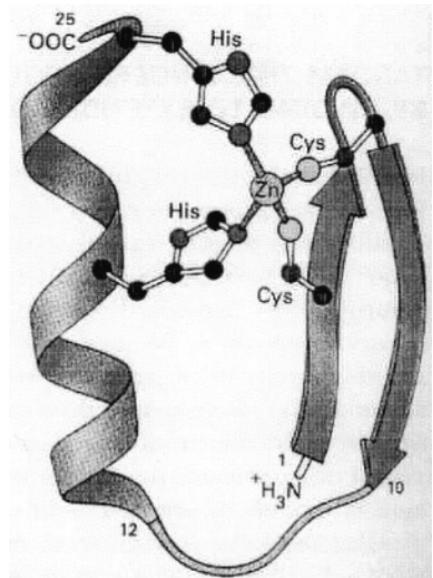
P. GAL4 is a Zn Finger Homodimer with 3 modules

- a. Zn-binding region: binds DNA.
- b. Linker region: extends along minor groove.
- c. Dimerization element: short parallel coiled-coil.



Q. Zn Finger is a protein-folding motif

- a. Zn finger motif: uses α -helical to contact and recognize DNA major groove.
- b. Zn^{2+} coordinates x2 His and x2 Cys tetrahedrally.
- c. TFIIA: 9 tandem Zn-finger repeats.
His and Cys protrude from 2 antiparallel beta sheets and an alpha helix.
- d. Recognizes 3 bp each (mostly A/G), often repeated.
Arg helps find G. NOT palindromic!



R. Mechanism of Transcriptional Activators

- a. Stimulate cooperative assembly of initiation complex.
- b. Histone acetyltransferase (HAT) complexes: recruited via protein-protein interactions with activation domains.
- c. Acetylation frees up nucleosome structure, facilitating promoter recognition by RPol II.

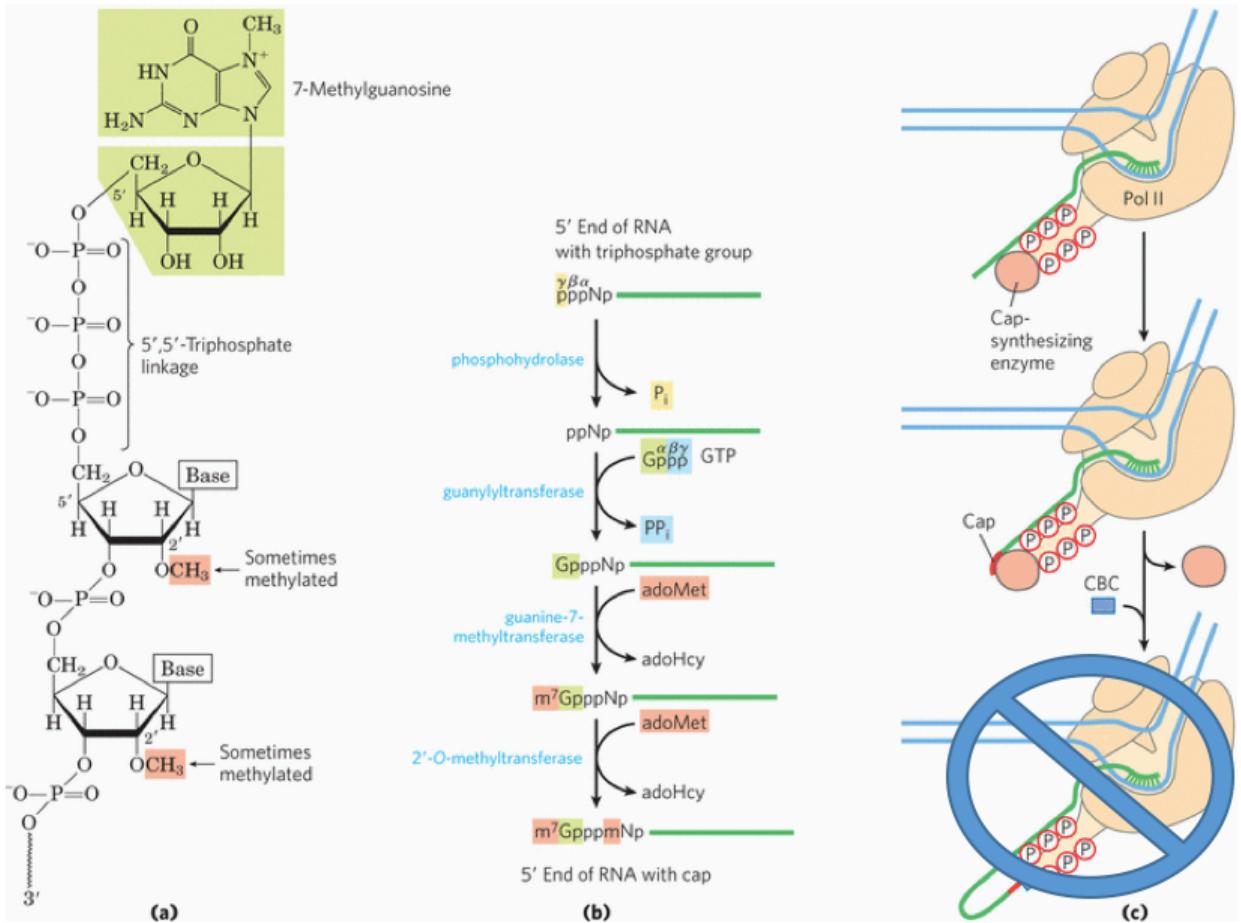
S. Histone Code Hypothesis

- a. Histone code: histone tails modifications that affect chromatin structure and transcription.
- b. Phos-CTD recruits methyltransferases.
- c. Methylation states may recruit TFIID (activation) or inhibit acetylation (suppression).

Lecture 20: Eukaryotic RNA Processing

A. 5'-capping of pre-mRNA

- a. Prevents digestion, helps translation.
- b. 7'-methylation of cap itself, then (often) 2'-methylation of nts 1/2
- c. m7G joined by 5'-5'-triphosphate linkage.
- d. Synthesis by enzymes tethered to RPol II CTD (co-transcriptional).



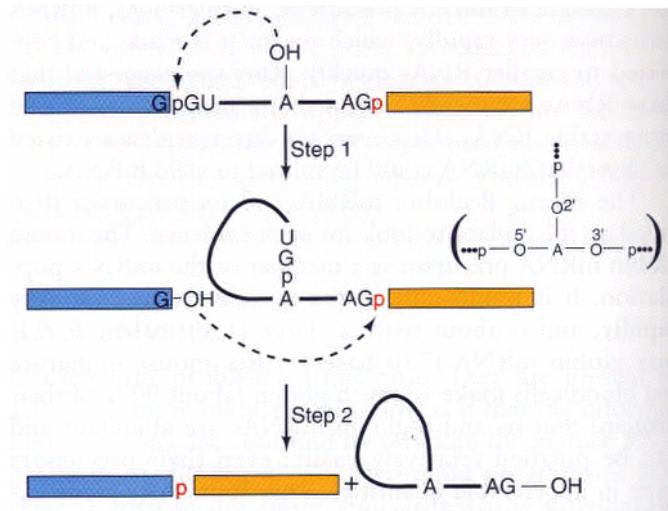
B. RNA Splicing in Euks

- a. No introns in vertebrate pre-rRNA

Final Product	RNA Polymerase Synthesizing Pre-RNA	Location of Signals Important for Splicing	Distinguishing Characteristics
rRNA	I	Intron and splice sites	RNA catalysis, guanosine cofactor
mRNA	II	Splice sites, intron branch point	Lariat formation, involvement of small nuclear RNPs
tRNA	III	Exons	ATP-activated enzyme intermediates
Organelle RNA (mitochondrial, chloroplast)	Organellar		
Group I Introns* (rRNA, tRNA, mRNA)		Intron and splice sites (structure of intron same as rRNA above)	RNA catalysis, guanosine cofactor, intron-encoded splicing proteins
Group II Introns* (mRNA)		Intron and splice sites (sequences at splice sites similar to mRNA above)	RNA catalysis, lariat formation, intron-encoded splicing proteins

C. Pre-mRNA Splicing Signals

- a. Consensus sequences at the ends of pre-mRNA introns:
 - i. Most interior nts are dispensable (to 70 nt minimum).
 - ii. Chimeric introns are spliced; splice sites are conserved (plant/animal).
 - iii. Mammals use “cryptic” splice sites (similar to the consensus sequences) when the natural splice site is disrupted. Yeast splicing just stops.
- b. Lariat intermediate
 - i. Internal intron A-OH attacks 5'-exon splice site (release of 5'-exon).
 - ii. Exon 1 G-OH attacks 3'-exon splice site (releases 3'-exon, connects to 5'-exon).

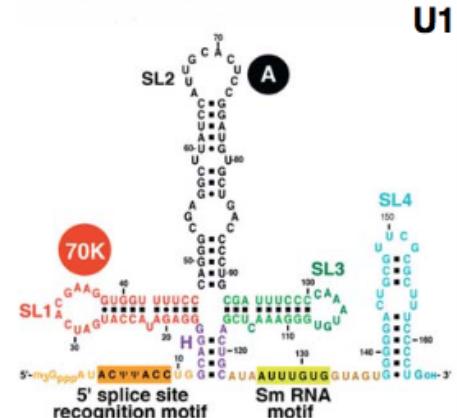


D. Splicing Mutations and Disease

- a. 15% of all genetic diseases
- b. β-Thalassaemias (β -globin mutation/deficit):
 - i. Mutation of conserved G at 5'-intron forces use of cryptic sites.
 - ii. Mutation within intron creates new 5'-splice site (uncovers 3'-cryptic splice sites).

E. snRNPs: small nuclear ribonucleoproteins

- a. snRNA complexes with 7 Sm proteins (except U6) via the Sm binding site, along with other specific proteins.
- b. Sm proteins are autoantigens in lupus, which provides antibodies against snRNPs.
- c. snRNA interacts with pre-mRNA via short absolutely conserved sequences.
- d. U6 is most highly conserved: catalytic core of the spliceosome.
- e. U4 delivers U6 to the spliceosome and leaves.
- f. snRNAs exist at 10^6 copies/cell.



F. Steitz Lab: Evidence that snRNPs Involved in Splicing

a. RNase Protection (“Bind-and-Chew”)

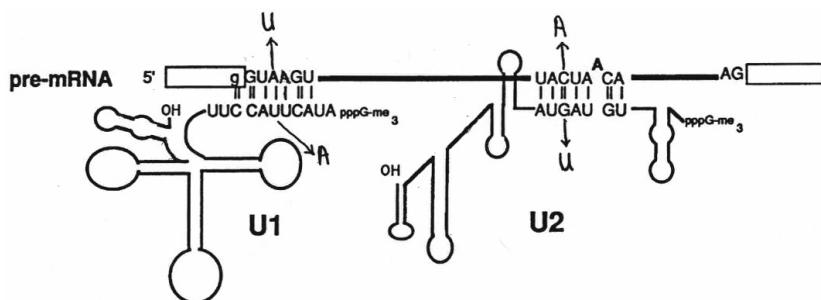
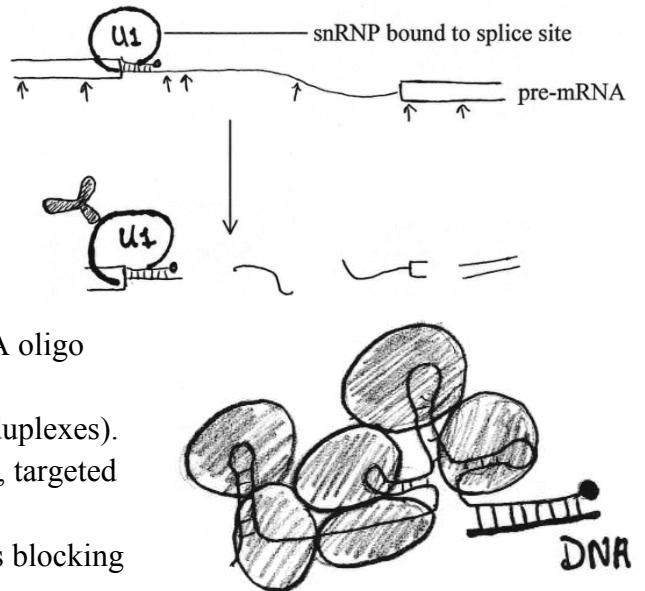
- Add ^{32}P -labeled pre-mRNA to splicing extract.
- Digest with RNase; snRNPs protect bound fragment.
- Anti-snRNP antibody selects fragment for analysis.

b. RNase H Experiments

- Incubate splicing complex with DNA oligo complementary to snRNA.
- Add RNase H (cleaves DNA-RNA duplexes).
- Add splicing substrate. If no activity, targeted snRNP is essential.
- Worries from 2° structure or proteins blocking hybridization.

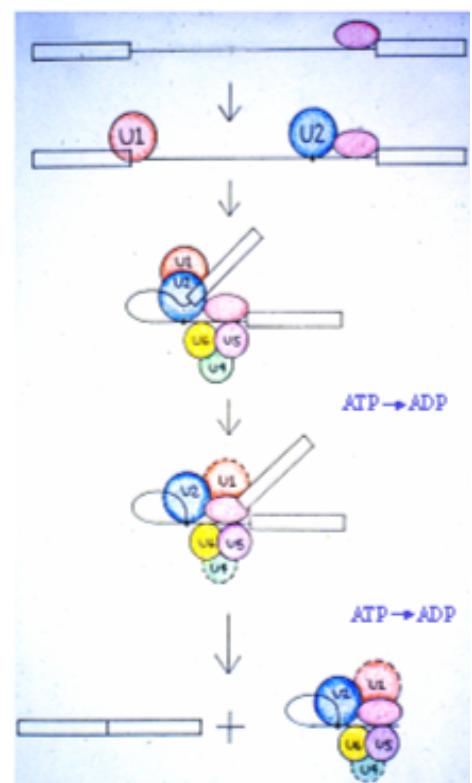
c. Genetic Suppression Experiments

- Make knockout point mutations in the 5'-splice site or branch site.
- Engineer U1 or U2 compensatory mutations. Assay splicing
- Identifies binding interactions by base-pairing.
- Identifies essential U-RNAs (1, 2, 4, 6).
- snRNP interactions with pre-mRNA explain minimum intron size (~70 nt).



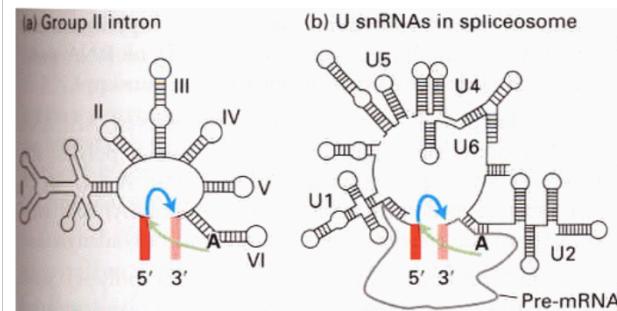
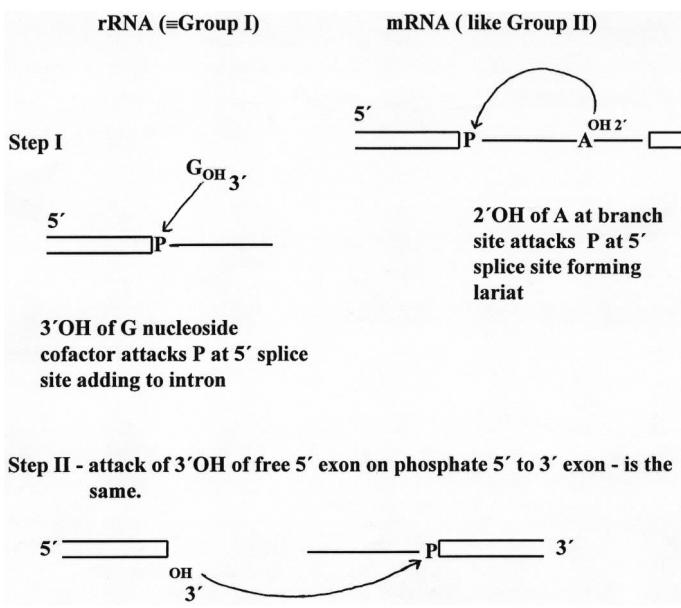
G. Spliceosome Formation

- Active spliceosome is 60S (ribosome-sized) and contains 50-100 proteins in addition to snRNP proteins.
- Splicing factor U2AF binds intron poly-C/T upstream of 3'-splice site.
- U1/U2 bind at the 5'-splice site and branch site.
- U4/U6-U5 bind as tri-snRNP.
 - U4 separates and leaves
 - U6 replaces U1 at 5'-splice site.
 - U5 aligns exons for ligation.
- After exon ligation, the snRNPs recycle.
- ATP powers conformational rearrangements, but contributes no phosphates at any step.



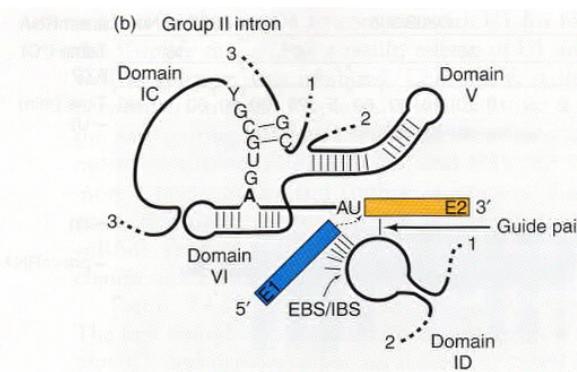
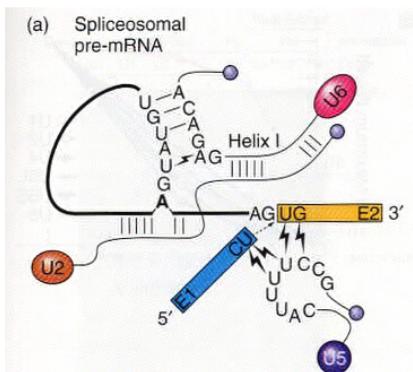
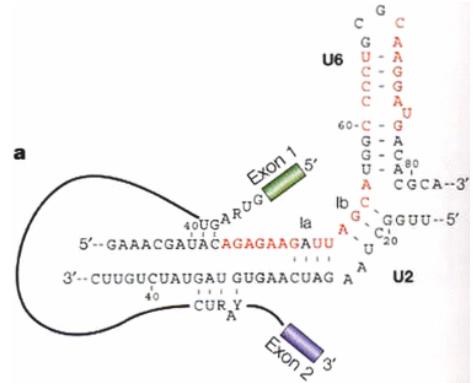
H. Spliceosome as Ribozyme?

- a. Each nt attack and displacement is a transesterification, with 2 events in splicing.
- b. Resembles Group II mechanism.



I. Dynamic Changes in RNA-RNA Interactions in the Spliceosome

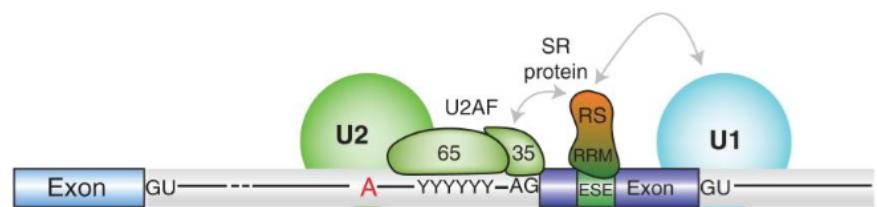
- a. Step 1
 - i. U1 snRNP identifies 5'-splice site, but replaced by U6 prior to step 1.
 - ii. U6 binds U2 to bring the 5'-splice site and branch site together for step 1.
- b. Step 2
 - i. Parallels with Group II suggest RNA catalysis.



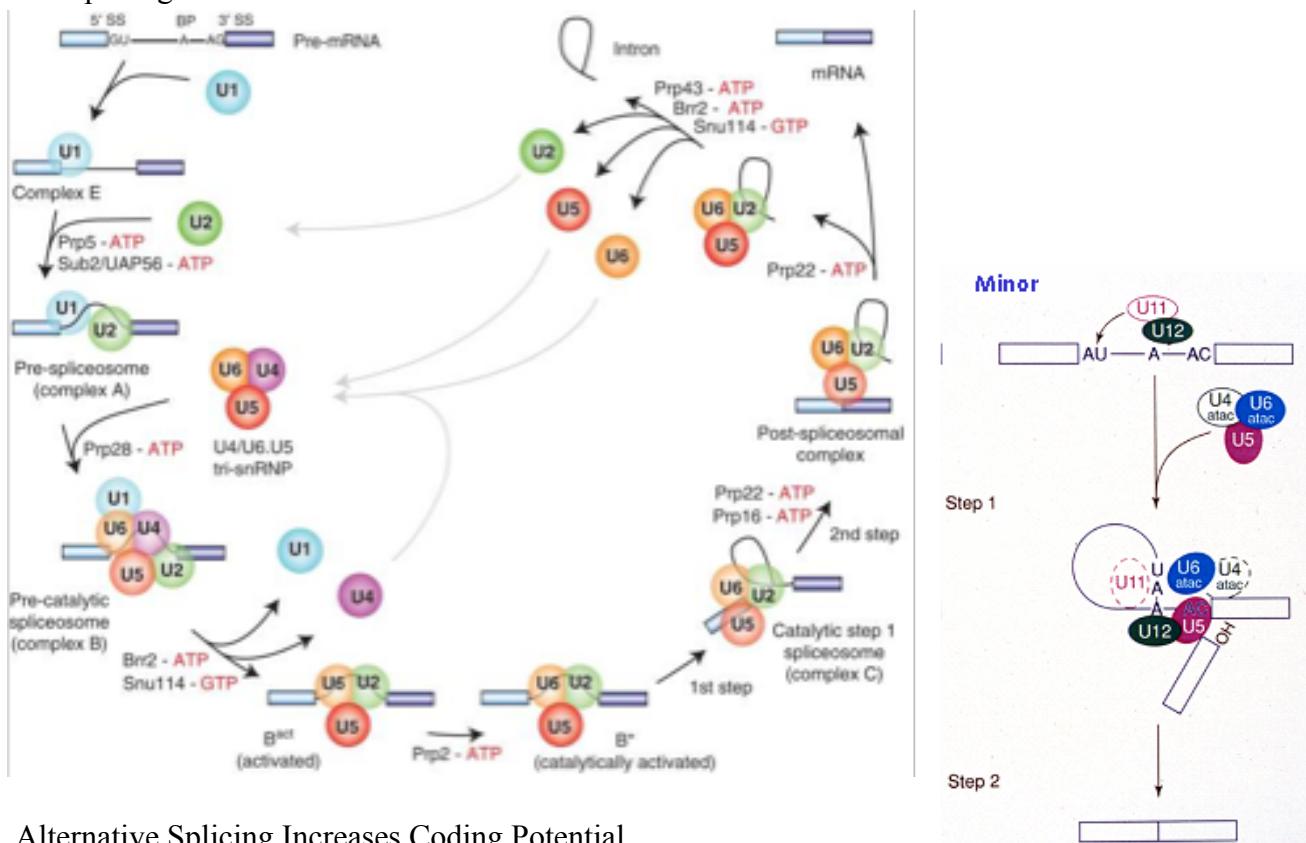
J. Role of Spliceosomal Proteins

- a. Exon and Intron Definition
 - i. Assist U1/U2 snRNP binding.
 - ii. Mammals: SR proteins (splicing factors) have Ser/Arg-rich CTDs.
 - iii. ISE: Intronic splice enhancer, ESE: Exonic splice enhancer

- iv. ~20 SR proteins bind to ISEs/ESEs and help snRNPs find the splice sites.
- v. SR proteins also assist mRNA nuclear export.
- b. Promote RNA-RNA exchanges
- i. DEAD/DEAH Box Proteins: ATPases, often helicases.



K. Full Splicing Mechanism



L. Alternative Splicing Increases Coding Potential

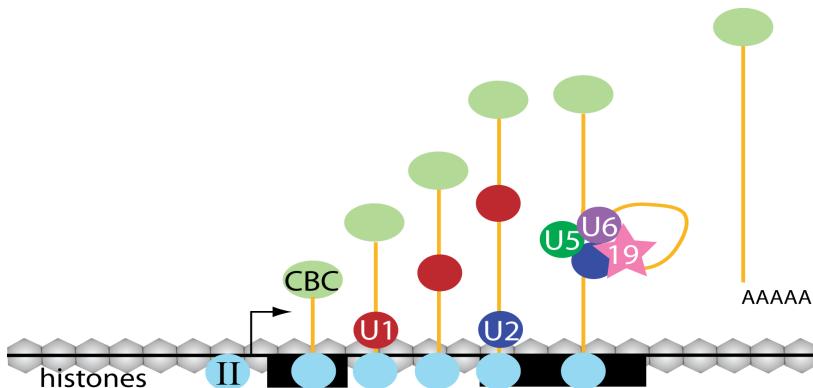
- a. 95% of human genes are alternatively spliced.
- b. Often by choosing each exon from a set of possibilities and splicing out the others.
- c. Drosophila sex determination by alternative splicing
(male transcript introduces stop codons or adds chunks).
- d. Caused by: SR proteins, hnRNPs (heterogeneous nuclear RNPs), and other proteins bind to pre-mRNA to alter spliceosome assembly.

M. Second or "Minor" Spliceosome (U12-dependent)

- a. Different consensus sequence but similar mechanism.
- b. Very old, present across plants and animals.
- c. Rare: 0.3% of human introns.
- d. snRNPs (U11, U12, U4atac, U6atac) resemble major snRNPs, but only 1% as abundant.
- e. U5 common to both spliceosomes.

N. Pre-mRNA Splicing Occurs Co-Transcriptionally

a. Processing steps associated with phases of transcription.



Transcription:	Initiation	Elongation	Termination
Processing:	Capping	Splicing	Polyadenylation

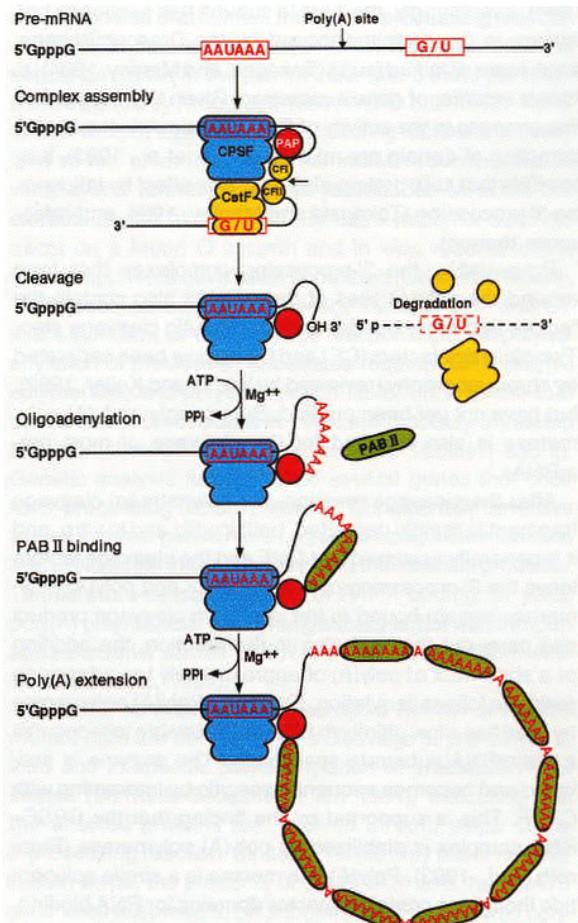
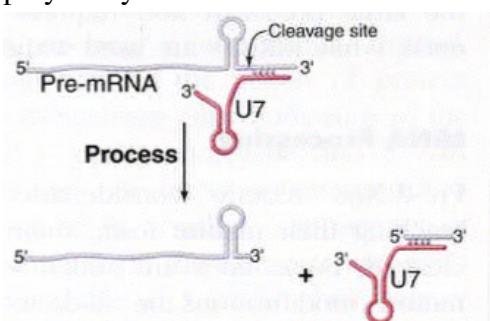
O. Vertebrate mRNA 3'-End Formation

a. Polyadenylation

- i. Only protein factors!
- ii. Pre-mRNA 3'-end includes AAUAAA signal and GU-rich region.
- iii. Complex assembly and cleavage.
- iv. First, ~10 polyA pre-tail.
- v. PAP (polyA polymerase) elongates to ~200 nts
- vi. PAB (polyA binding protein) decorates the strand 25-nt at a time.
- vii. Polyadenylation coordinated with termination/splicing via RPol II CTD.

b. Histone mRNA

- i. No polyA tails and no introns, but still processed by 3'-cleavage.
- ii. U7 snRNP binds with an element 3'- to the cleavage site and cuts it.
- iii. Involves similar factors as polyadenylation.

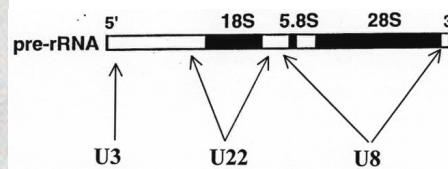


P. Summary of Nuclear Pre-mRNA Processing

Modification	catalysis	how specified	function
5' end capping	3 enzymes	Pol II CTD (trans)	stability splicing translation
Splicing	RNA (spliceosome)	RNA sequences (cis)	stability makes ORF(s) export
Polyadenylation	2 enzymes	RNA and Pol II (cis and trans)	stability alternative ORFs translation

Q. Nucleolus Inhabited by snoRNPs (lots)

- a. snoRNPs (small nucleolar RNPs) are involved in pre-rRNA processing.
- b. 3 snoRNPs involved in cleavage events.
- c. U3 snoRNP attaches near 5'-end and recruits proteins to form 5'-terminal knobs.



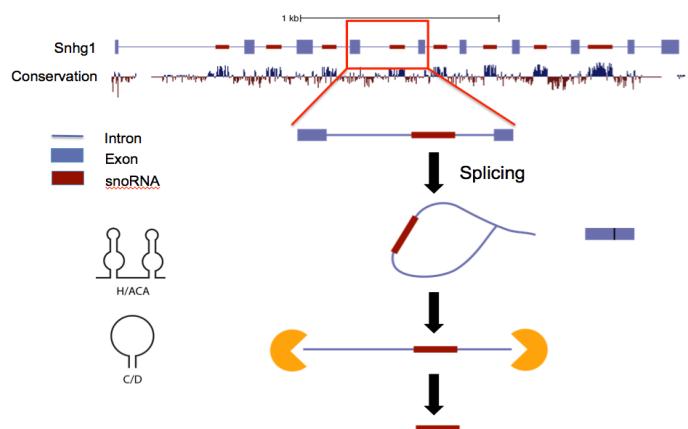
- d. U22 snoRNP excises 18S
- e. U8 snoRNP half-excises 5.8S and 28S.
- f. No intron in vertebrate pre-rRNA!

R. SnoRNPs Guide Modification of rRNA

- a. SnoRNPs base-pair to pre-rRNA and guide modifications at conserved positions.
- b. Box C/D snoRNAs direct 2'-O-methylation.
- c. Box H/ACA snoRNAs direct pseudouridylation by dyskerin.
- d. ~100 of each, so ~200 guide snoRNPs predicted.
- e. Each class uses different proteins.

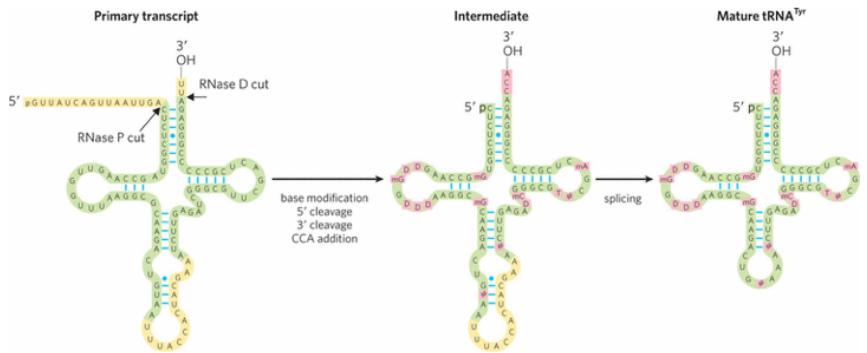
S. Most snoRNAs Encoded in Introns

- a. Introns from housekeeping protein-coding genes (riboproteins, nucleolar proteins, translation factors).
- b. SnoRNP proteins bind to snoRNA sequence during splicing.
- c. Lariat intron is debranched
- d. Exonucleases trim the ends and are stopped by snoRNP proteins.
- e. SnoRNP migrates to nucleolus.



T. Euk pre-tRNAs have introns with protein-mediated splicing.

- a. RNase P and D cut and define the acceptor stem.
- b. Base modification and CCA addition
- c. Splicing of the anticodon loop.



U. RNA Editing Alters Primary Gene Transcripts

- a. Mammalian cells:
 - i. C → U deamination (protein-catalyzed)
 - ii. A → I deamination (dsRNA-catalyzed; ADAR: adenosine deaminase RNA)
- b. Protozoa:
 - i. Site-specific insertion/deletion of Us in mitochondria.
 - ii. Involves guide-RNA-directed endonucleases (opening cut), TUTases (insertions), and 3'-U-exonucleases (deletions).

Lecture 21: Eukaryotic Translation

A. Overview of Euk Translation

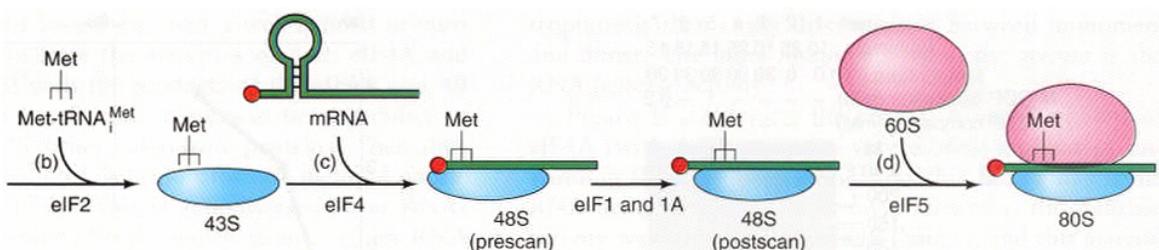
- a. Physically separated/compartmentalized from transcription.
- b. Cytoplasmic mRNA with short polyA tails are masked (not translated). Allows storage of maternal mRNAs and neural learning.
- c. Euk mRNA is monocistronic, 1 ORF.
- d. Initiation
 - i. Initiation sequences are different.
 - ii. 5'-cap and polyA tail are necessary.
 - iii. Requires ATP as well as GTP.
- e. Termination involves 1 RF for all 3 stop codons, and a stimulatory GTPase factor.
- f. Compartmental sorting of proteins (mRNA are translated where you want the protein).
- g. 80S Euk ribosomes are larger and more protein-rich.
 - i. 40S subunit (18S rRNA)
 - ii. 60S subunit (28S, 5.8S, 5S rRNAs)
- h. Proteins may undergo post-translational cleavage and modification (even splicing!)

B. Euk Translation Initiation

- a. No Shine-Dalgarno, but has 5'-UTR of 3 to 1000+ nts.
- b. Usually selects 1st AUG, but sequence context is important. Exceptions: viral proteins, proto-oncogenes/growth factors with IRES (internal ribosome entry sites).
- c. Kozak sequence: ACCAUGG is optimal
- d. Special Met-tRNA_i required, but no formyl group.
- e. Scanning from 5'-end
 - i. Requires 5'-cap, 40S, and eIFs to start, ATP to scan.
 - ii. Unfolds secondary structure.
 - iii. Stops when Met-tRNA_i pairs with AUG of an ORF.
 - iv. 60S attaches and eIFs fall off.

C. Euk Initiation Factors (eIFs)

- a. eIF2: binds Met-tRNA_i and GTP. Brings it to the 40S subunit to form 43S pre-primed complex which scans from cap.
- b. eIF4A is DEAD box helicase, unwinds 2° structure.
- c. GTP hydrolysis releases eIF2.
- d. Scanning can be regulated by upstream AUGs (as in GCN4) and 2° structure (as in iron).
- e. eIF2B (GEF) is the guanosine exchange factor: releases GDP from eIF2.
Analog of Ts acting on Tu-GDP.



D. Euk Translation Initiation Cycle

- Phos-eIF2 sequesters eIF2B, stopping translation.
- Need to know: eIF- 2, 2B, 4A, 4E, 4G.

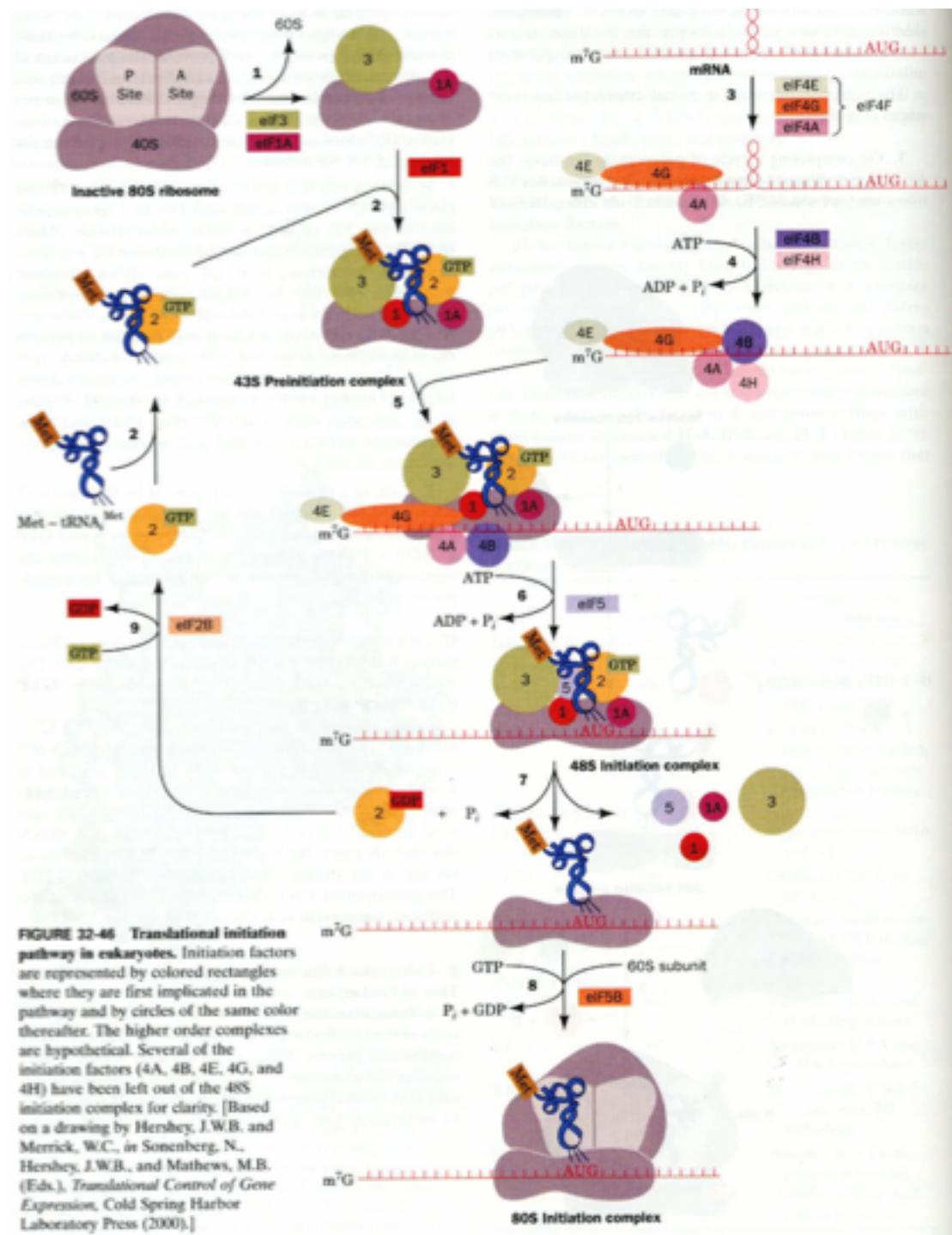
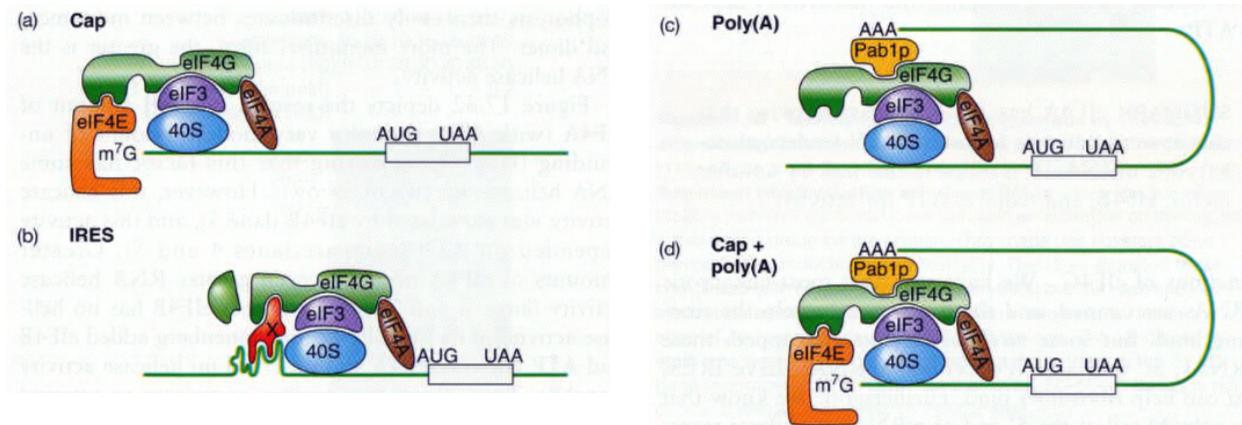


FIGURE 32-46 Translational initiation pathway in eukaryotes. Initiation factors are represented by colored rectangles where they are first implicated in the pathway and by circles of the same color thereafter. The higher order complexes are hypothetical. Several of the initiation factors (4A, 4B, 4E, 4G, and 4H) have been left out of the 48S initiation complex for clarity. [Based on a drawing by Hershey, J.W.B., and Merrick, W.C., or Sonenberg, N., Hershey, J.W.B., and Mathews, M.B. (Eds.), *Translational Control of Gene Expression*, Cold Spring Harbor Laboratory Press (2000).]

E. Interactions at mRNA Ends

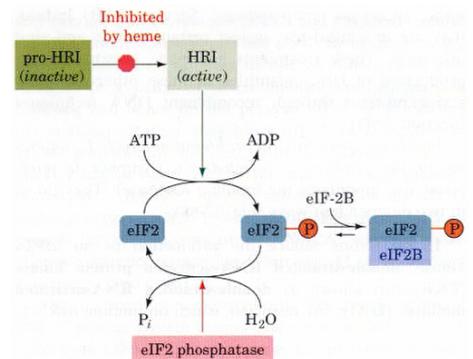
- Polysomes: multiple ribosomes translate a single mRNA.
- Circular mRNAs (linking cap and polyA) helps recycle ribosomes.
- eIF-4G as adaptor:
 - Binds eIF4E (cap), PAB (tail), and eIF3 (40S subunit).
 - Binds IRES proteins, allowing initiation at internal AUGs.



- d. Cap: recruits 40S to mRNA via eIF4E.
- e. IRES: recruits 40S to mRNA via RNA-binding protein X.
- f. PolyA: adaptor to connect 40S to polyA

F. Regulation of Initiation 1: Heme

- a. Heme + globins = hemoglobin.
- b. Reticulocytes are immature erythrocytes with a ton of globin mRNAs.
- c. Heme required for translation of globin mRNAs.
 - i. Heme inactivates an eIF2 kinase with 2 names:
HRI (heme-regulated inhibitor)
HCR (heme-controlled repressor)
 - ii. This sequesters eIF2B and stops GEF activity, stops protein synthesis.
 - iii. HRI also activated by heat shock, to prevent misfolded proteins.

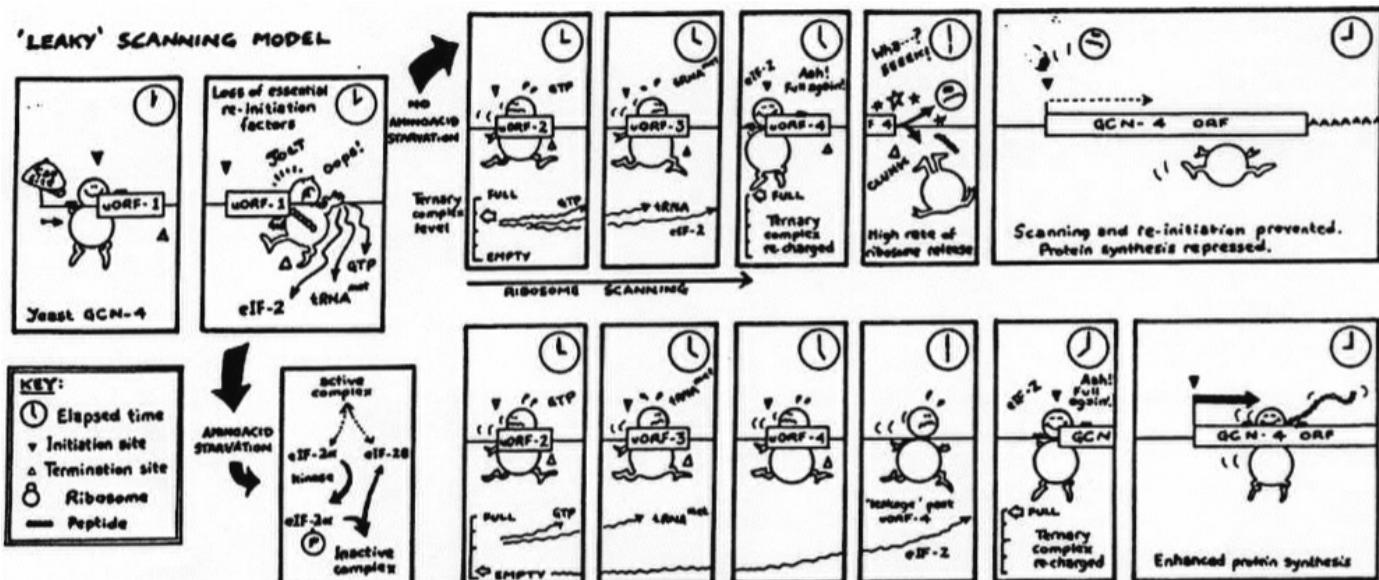


G. Regulation of Initiation 2: Interferon (glycoprotein)

- a. dsRNA activates interferon secretion, which signals antiviral state in neighbors.
- b. PKR (protein kinase dependent on dsRNA): phosphorylates α-subunit of eIF2 if dsRNA appears in the newly antiviral cell.
- c. Viral response: adenovirus encodes an imperfect dsRNA dummy that competes for binding to PKR but does not activate it.

H. Regulation of Initiation 2: GCN4 (yeast)

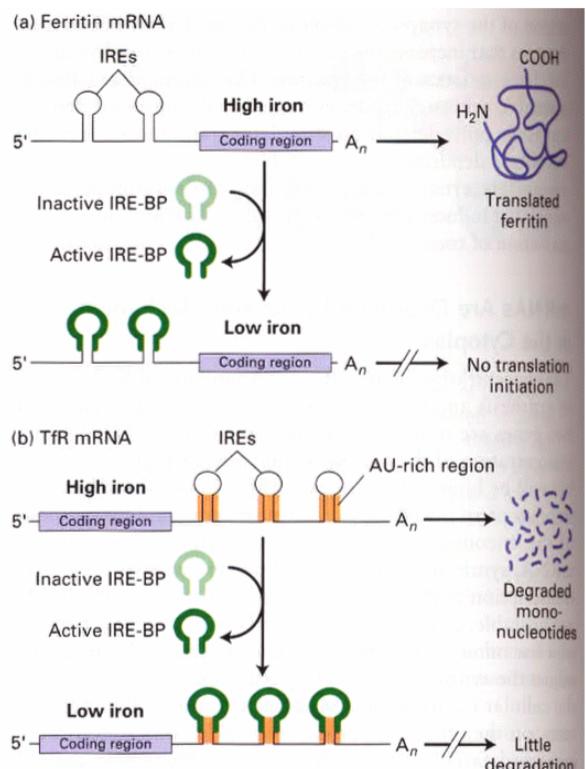
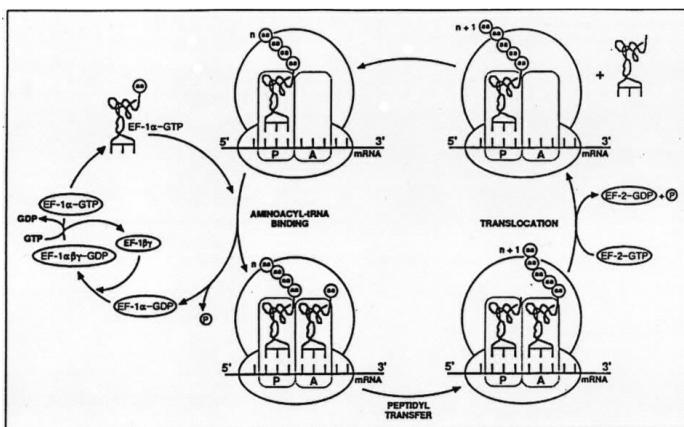
- a. GCN4 is a transcriptional activator of AA biosynthesis genes when limiting.
- b. Translational control: transcribed at a constant rate, regardless of supply.
- c. Mechanism
 - i. 5'-UTR of UCN4 mRNA contains 4 AUGs, each with a short ORF.
 - ii. Mutating/deleting these AUGs (or removing the whole 5'-UTR) leads to constitutive activation.
 - iii. Uncharged tRNAs bind and activate GCN2 kinase to phosphorylate eIF2 and reduce availability of TC (eIF2 + Met-tRNA_i + GTP)
 - iv. ORF 1 is translated, and the 40S keeps scanning. If it has TC/eIF2 available, it will restart at ORF 3 or 4 and drop. Otherwise, it will scan past and arrive at the GCN4.



I. Translational Control by Iron

- IREs (Iron Response Elements): ~10 bp conserved RNA stem loops bind IRE-BPs.
- IRE-BP is aconitase!
- Fe and IREs compete for binding.
High Fe = Free IRE, low Fe = bound IRE.
- Ferratin (iron storage protein):
BP can bind 5'-UTR IREs to stop scanning and translation.
- Transferrin (iron uptake protein):
BP can bind 3'-UTR IREs to stabilize mRNA.
- AREs (AU-rich elements): found in 3'-UTRs; trigger degradation.

J. Euk Elongation

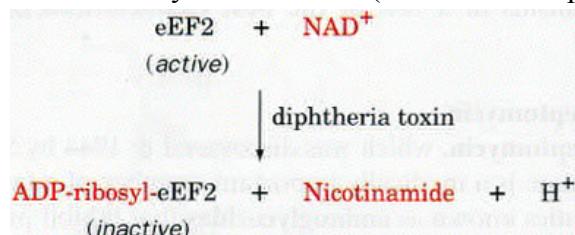


Factor	Function	Comparable Prokaryotic Factor*
EF1 α	Binds aminoacyl-tRNA to ribosomes	EF-Tu
EF1 β	Assists the recycling of EF1 α	EF-Ts
EF2	Translocation	EF-G
RF	Chain release	RF1 RF2

- a. Additional: RF3 binds GTP to stimulate termination.
 - b. Control of elongation by
 - i. Phosphorylation
 - ii. ADP-ribosylation of eEF2 by Diphtheria toxin.

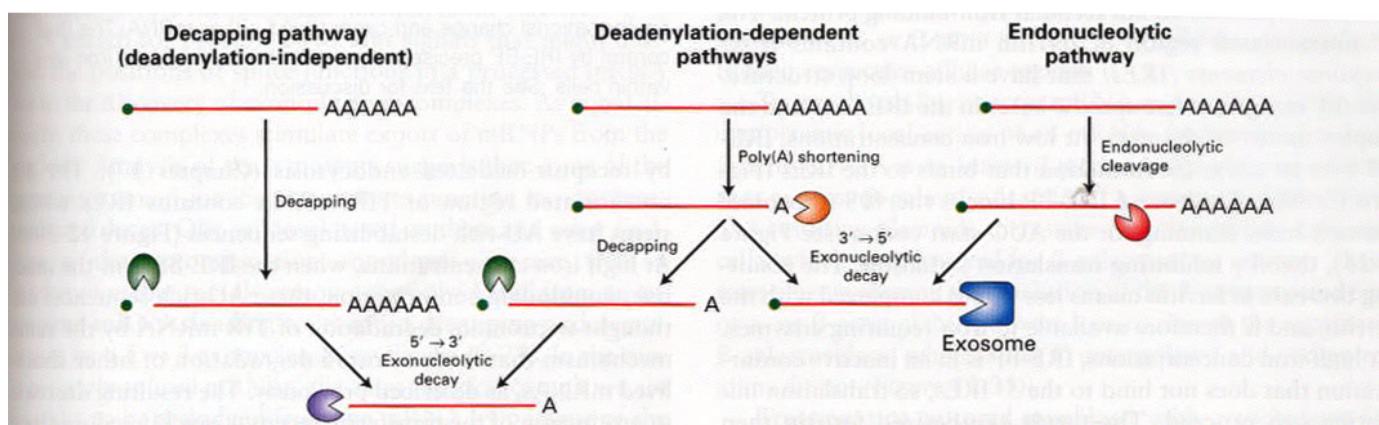
K. Inhibition of Elongation: Diphtheria Toxin

- a. ADP-ribosyl transferase: adds group to eEF2-His, inactivates factor.
 - b. Targets unique diphthamide modification in euks (not in EF-G).
 - c. Enzymatic activity: 1 molecule can poison all eEF2s in a cell.
 - d. Reversible by nicotinamide (L' Chatelier's principle).

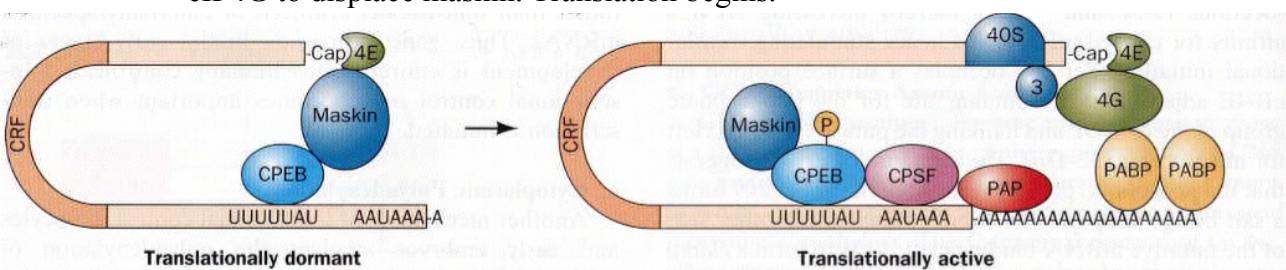


L. PolyA Tail

- a. PolyA lengthening increases translation.
 - b. PolyA trimming triggers decay. Other degradation pathways also exist.

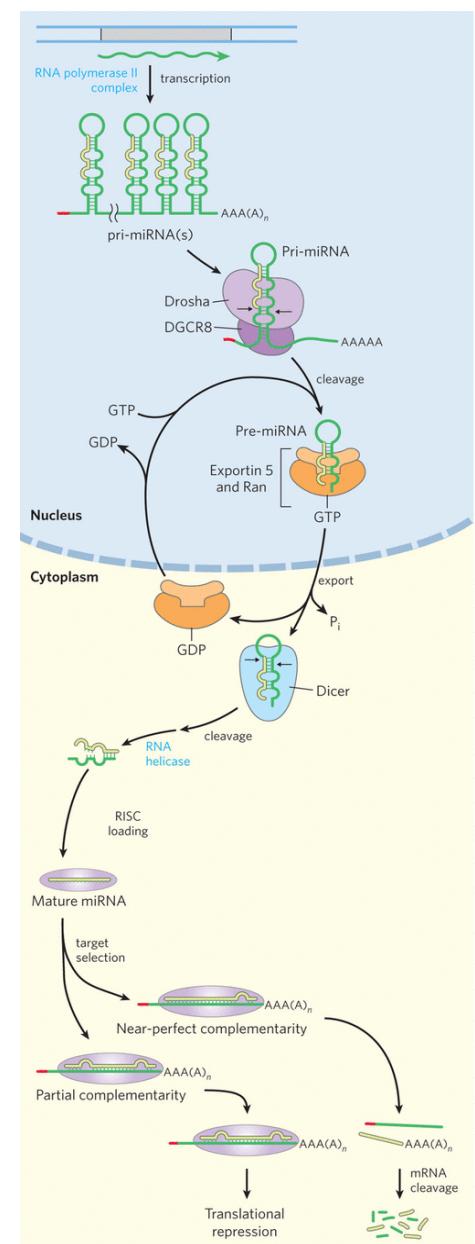
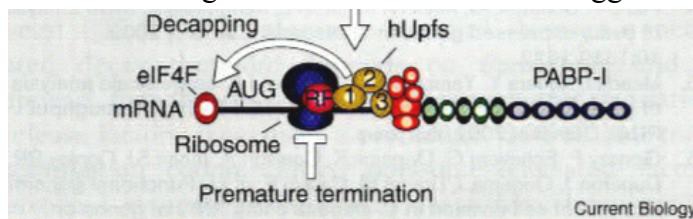


- c. Maternal mRNAs masked by proteins and short polyA.
 - i. CPEB (cytoplasmic polyadenylation element binding protein) binds Maskin and mRNA CPE. Maskin binds eIF4E-cap to stop eIF4G from initiating translation.
 - ii. Phos-CPEB recruits PAP to extend the tail. PABP binds the new tail and recruits eIF4G to displace maskin. Translation begins.



M. mRNA Stability

- a. AREs (AU-rich elements) – usually AUUUA
 - i. Target mRNA for decay.
 - ii. Present in growth factor and proto-oncogene 3'-UTRs.
 - iii. AREs can be transferred to other mRNAs to reduce half-life.
 - iv. Mechanism: factors recognize ARE and recruit exosome to for 3'-5' degradation.
 - v. Counter: other factors (e.g. IRE-BP) can competitively bind the ARE.
- b. NMD (nonsense-mediated decay)
 - i. Also called mRNA surveillance.
 - ii. Degrades mRNA to avoid producing a truncated protein.
This helps make nonsense mutation diseases recessive!
 - iii. Exon-junction complex (EJC): upstream of exon-exon junctions after splicing.
 - iv. When ribosome terminates prematurely, interaction with EJC factors leads to mRNA decapping.
 - v. Inserting an intron into the 3'-UTR triggers mRNA decapping and degradation.



N. RNA interference (RNAi)

- a. Overexpression attempts knocked out endogenous expression. dsRNA triggers degradation of complementary mRNA.
- b. Dicer: reduces dsRNA to ~22 nt siRNAs, which are loaded onto RISC.
- c. RISC: contains Argonaute (Ago) RNase H-like endonuclease that cleaves target mRNA.

O. Micro RNA (miR)

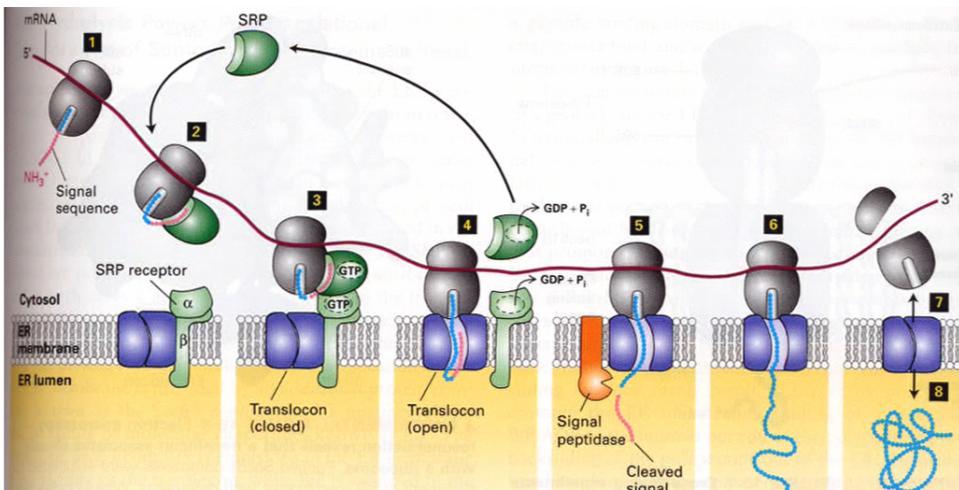
- a. Function
 - i. Perfect complementarity with target: RISC cleaves the target.
 - ii. Imperfect complementarity with target: represses translation.
- b. Biogenesis
 - i. Drosha cleaves pri-miRNA to form pre-miRNA.
 - ii. Dicer cleaves and loads miRNA into RISC.
 - iii. Low dicer associated with macular degeneration.
 - iv. Drosha and Dicer are RNase III enzymes: pseudo dimers; 2 catalytic domains per protein.

P. The Signal Hypothesis

- a. How do membrane and secreted proteins cross bilayers?
- b. Signal sequence
 - i. Targets to proper location in the cell.
 - ii. Hydrophobic core preceded by 1+ basic residues.
 - iii. Usually at N-terminus.
 - iv. Cleaved off once protein crosses membrane.

Q. Signal Recognition Particle (SRP)

- a. Transports nascent polypeptides across ER membranes.
- b. RNP complex of 7SL RNA and 6 proteins.
- c. Recognizes emerging signal sequence and binds ribosome to arrests translation.
 - i. Sequesters nasty enzymes meant for export.
 - ii. **??? Increases targeting fidelity.**
- d. SRP receptor in the membrane guides SRP to the translocon.
 - i. GTP hydrolysis triggers release of SRP and continuation of protein synthesis.
 - ii. Signal peptidase cleaves the signal.



R. Controlled Protein Degradation by the Ubiquitin Pathway

- a. **Ubiquitin**: most conserved euk protein.
- b. **E3 (ubiquitin ligase)** links many ubiquitins to internal Lys of condemned proteins.
- c. **Poly-Ub protein** is directed to proteasome (multi-protease complex).
- d. ATP hydrolysis required at many steps.
- e. **N-end rule**: E3 recognizes N-terminal residues, which determine protein half-life (can be stabilizing or destabilizing).

S. Protein Splicing

- a. Present in all kingdoms of life.
- b. Fast, autocatalytic, occurs via branched intermediate.
- c. **Inteins** (intron analog) are 100-2000 AA long, with conserved AA at N and C termini.
- d. **C-terminal exteins** contain conserved Ser.
- e. Some inteins are DNA endonucleases, **like those encoded by Group I RNA introns???**