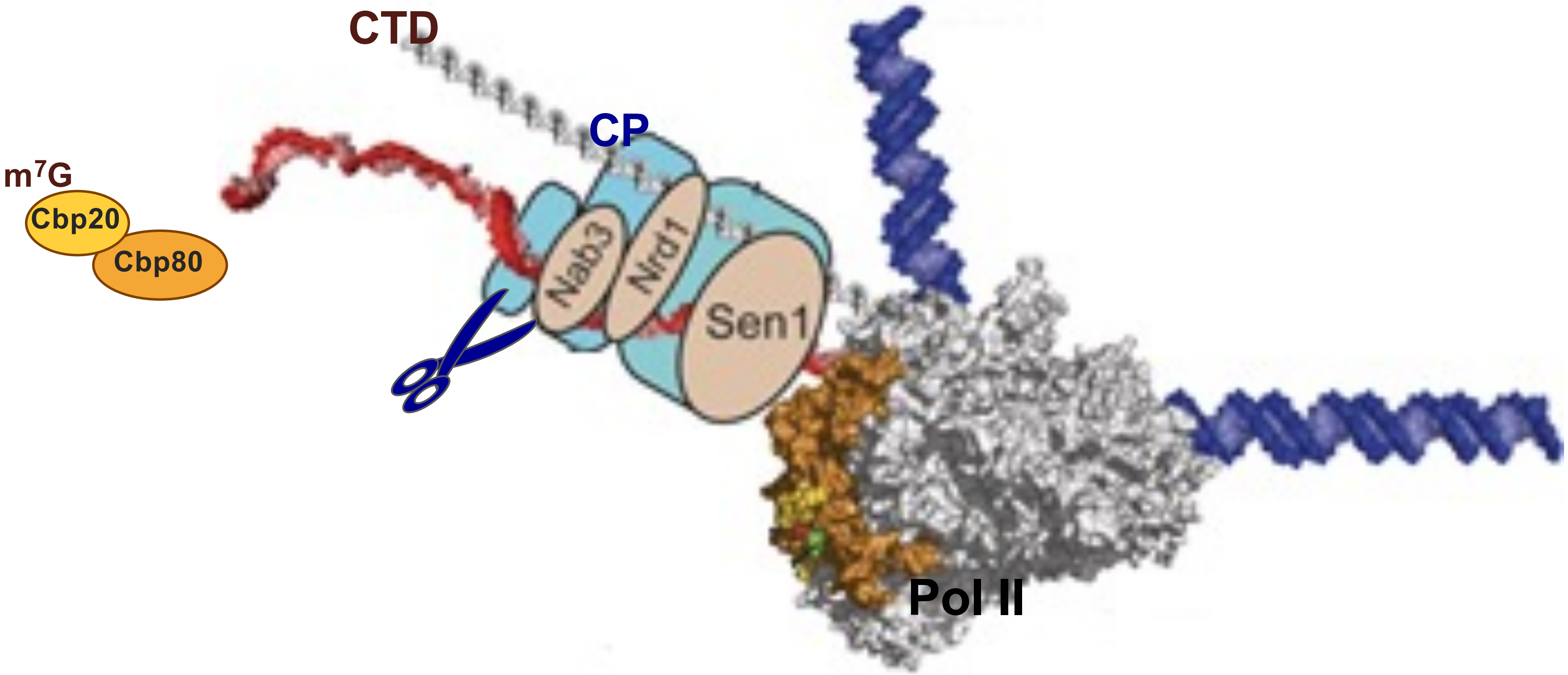
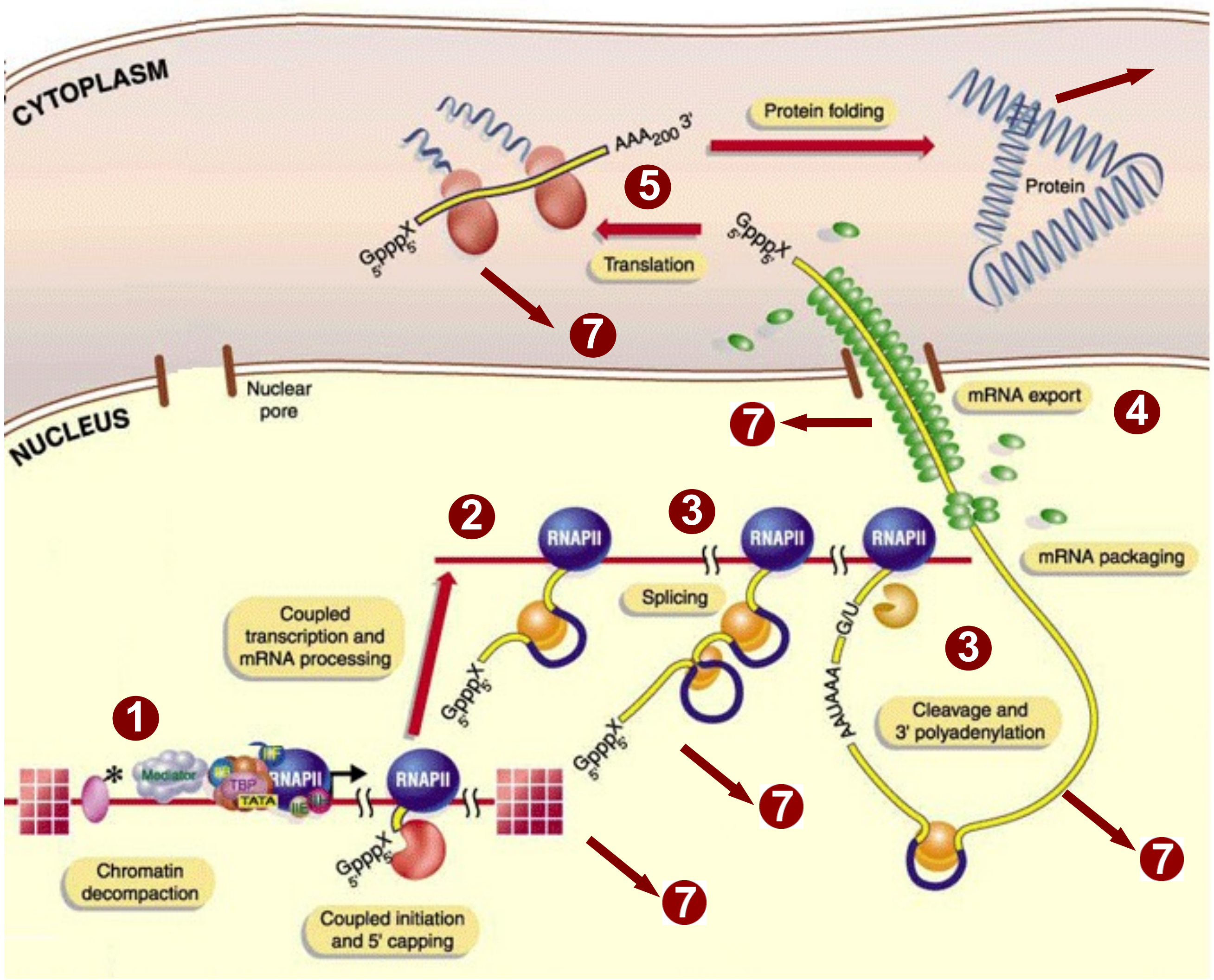
TRANSCRIPTION

How to make RNA?



REGULATION OF GENE EXPRESSION



1) chromatin

2) transcription

3) RNA processing

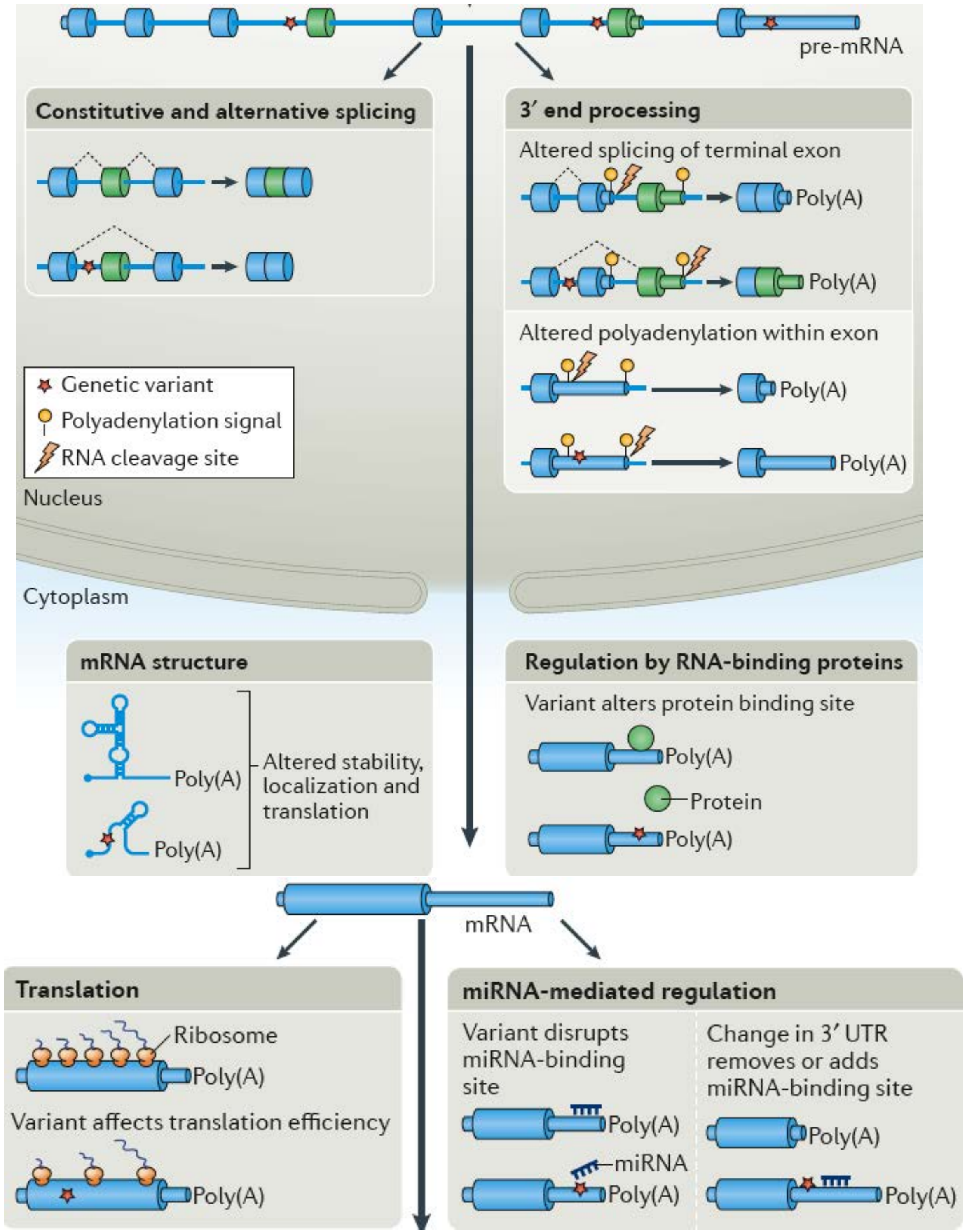
4) RNA export

5) translation (mRNA)

6) protein stability

7) RNA degradation

All steps of RNA metabolism affect gene expression

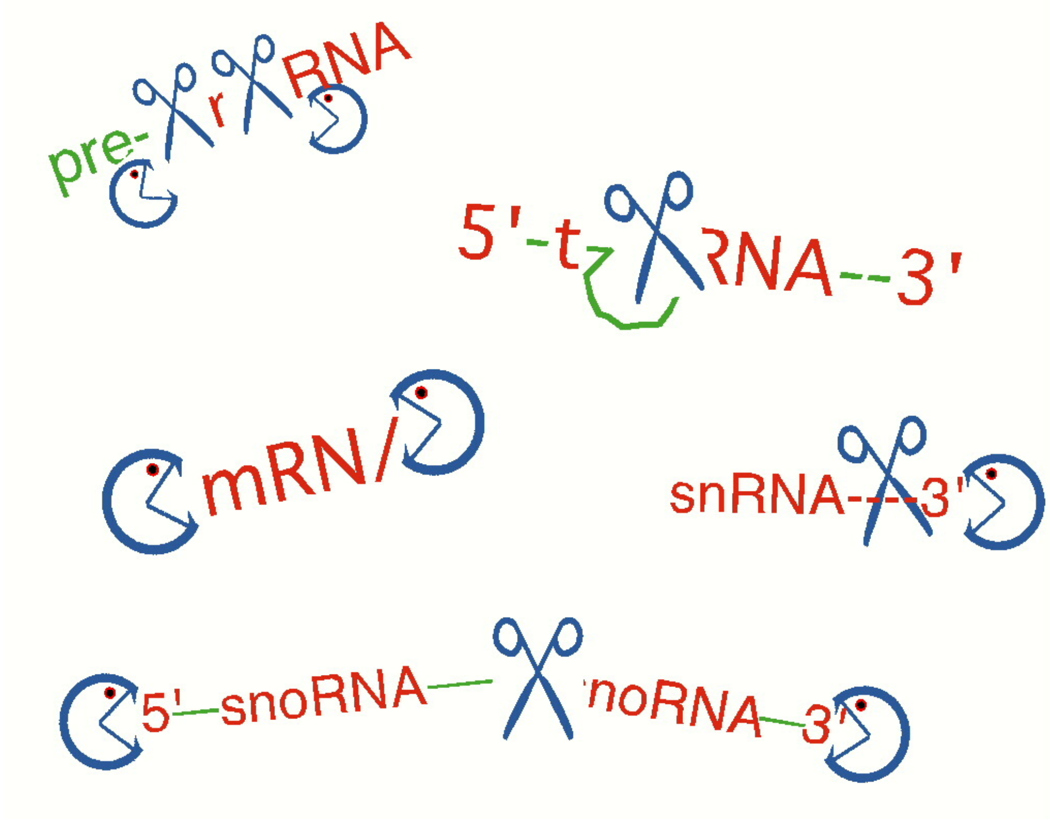


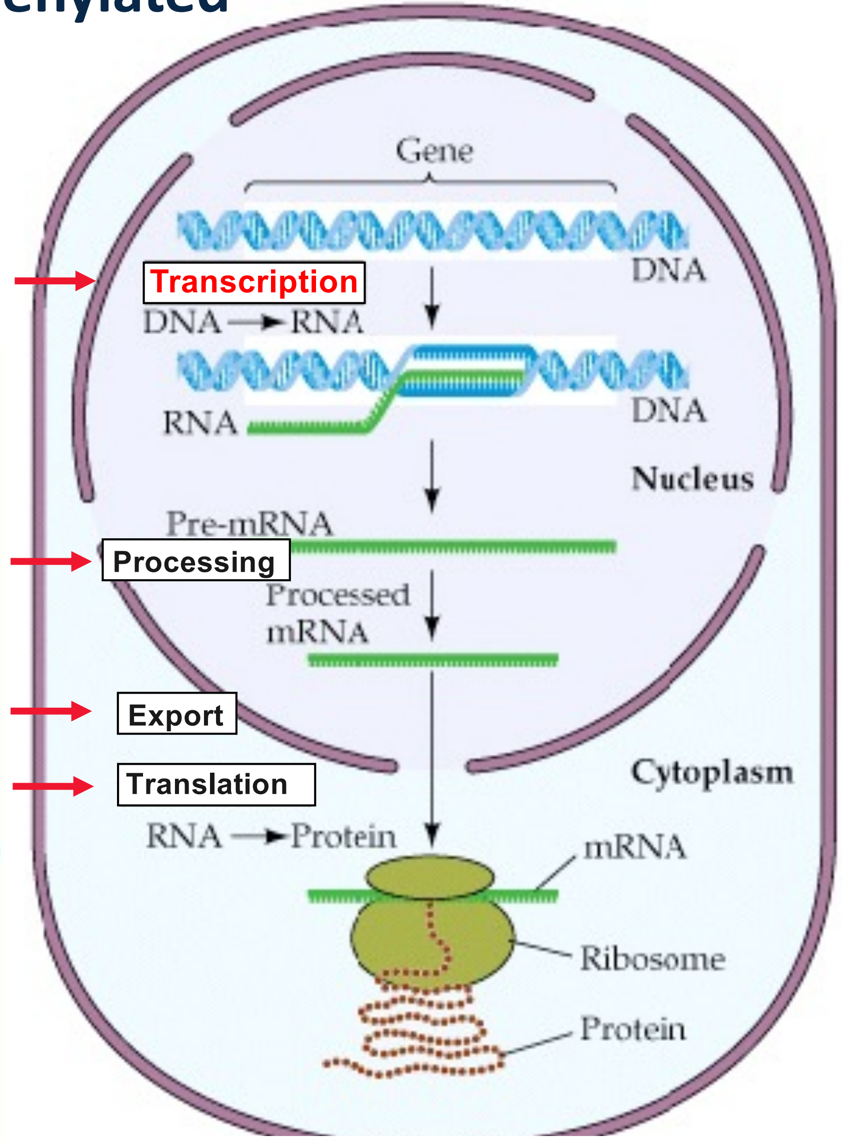
Manning et al, NatRevMolCellBiol 2016

RNA

- coding- mRNAs
- non-coding- ncRNAs
  - housekeeping
  - regulatory
  - short
  - long

- polyadenylated
- non-polyadenylated
- stable
- unstable
- capped
- uncapped



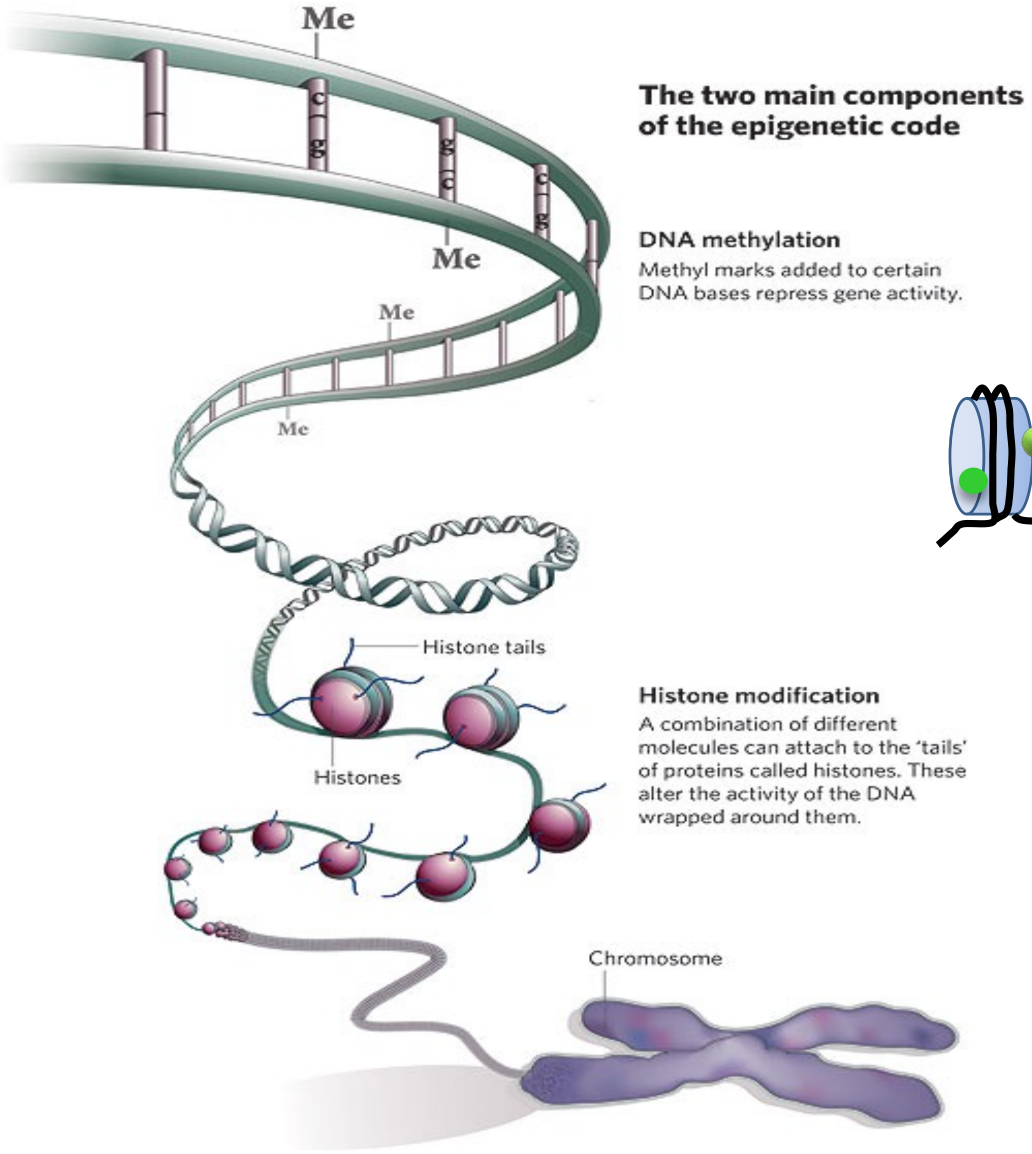


eukaryotic cell

Transcription

1. Chromatin structure and modifications, histones, nucleosomes
2. Eukaryotic polymerases
3. Promoters, activators, enhancers
4. Factors, reguators, complexes
5. Initiation, elongation, termination
6. Co-transcriptional processes

Chromatin



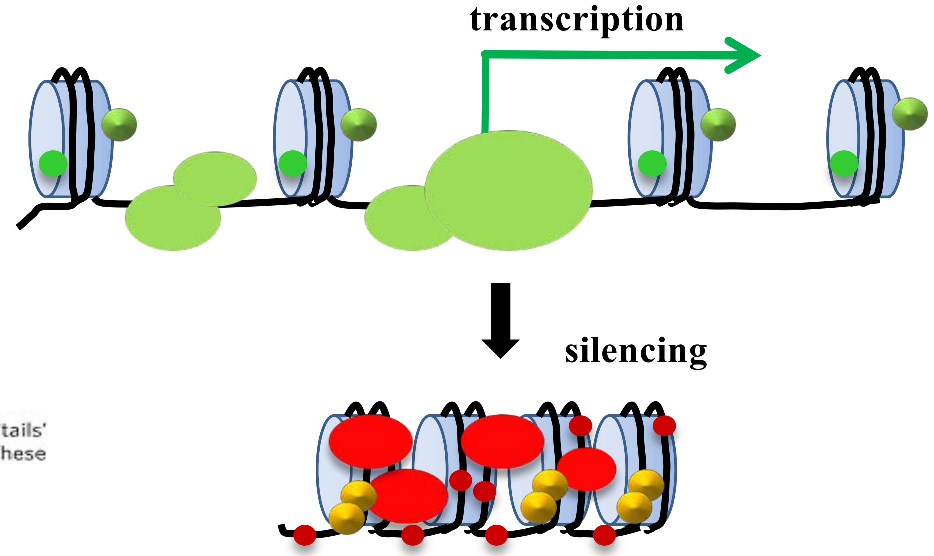
The two main components of the epigenetic code

DNA methylation

Methyl marks added to certain DNA bases repress gene activity.

Histone modification

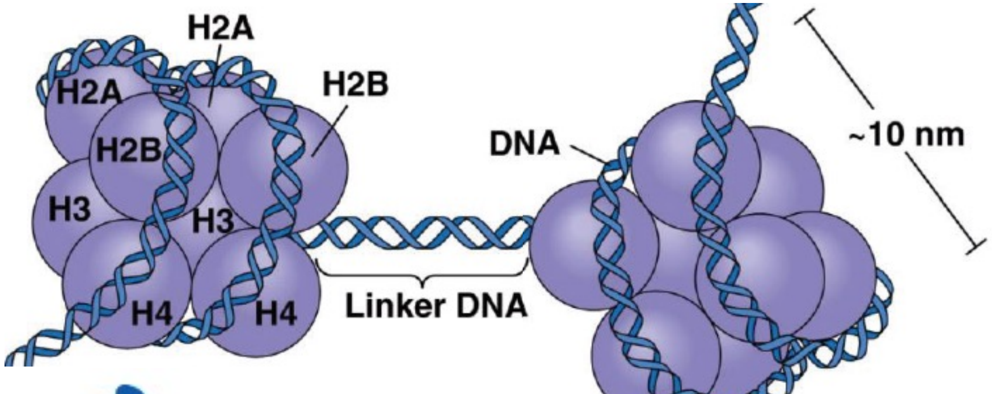
A combination of different molecules can attach to the 'tails' of proteins called histones. These alter the activity of the DNA wrapped around them.

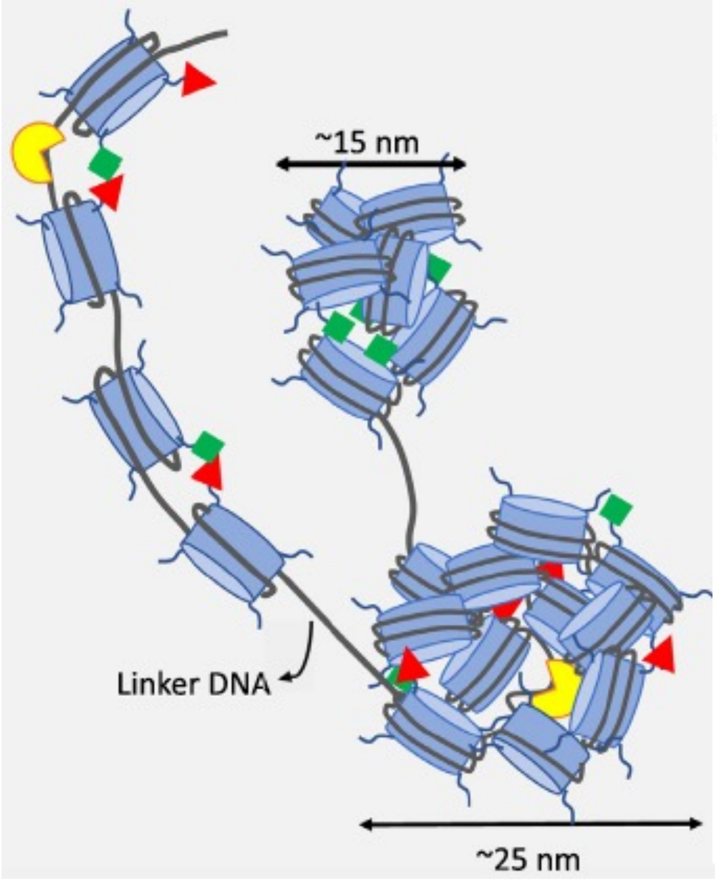


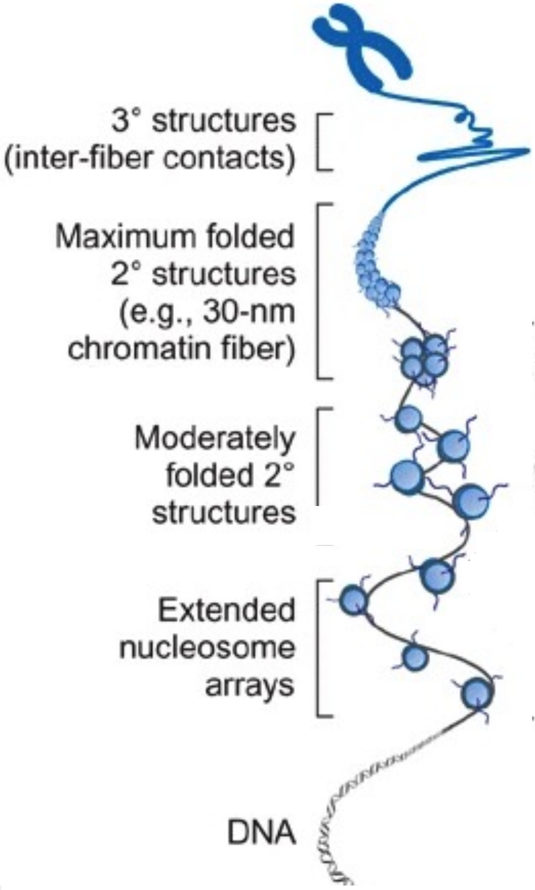
Chromatin: structure

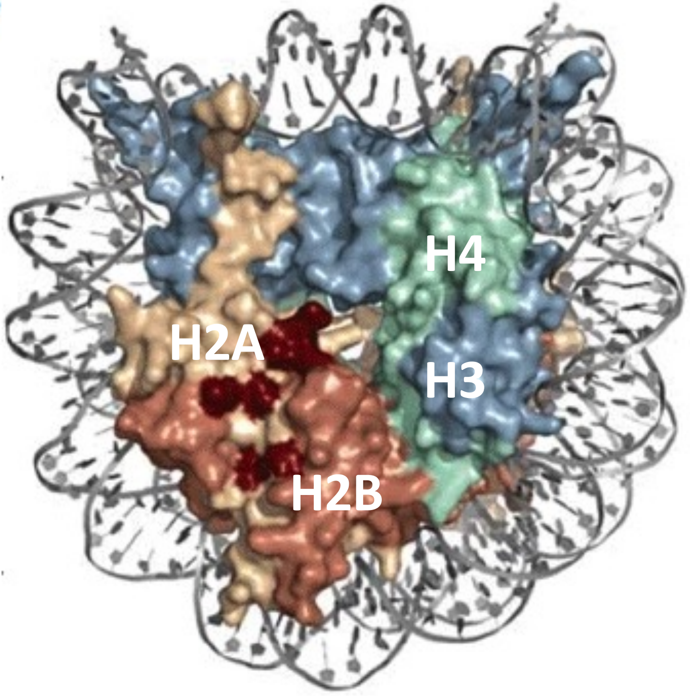
~ 147 bp DNA wrapped around histone octamer

- 8 histones:
- 2x {H2A, H2B, H3, H4}
- linker DNA (50 bp)
- linker histone H1









Caterino and Haye, 2007, Nature; Parmar and Padinhateeri, Curr Op Struct Biol, 2020

Chromatin: structure

Heterochromatin

- ***constitutive***

- always present in a cell,
- devoid of genes

(centromeric, telomeric regions)

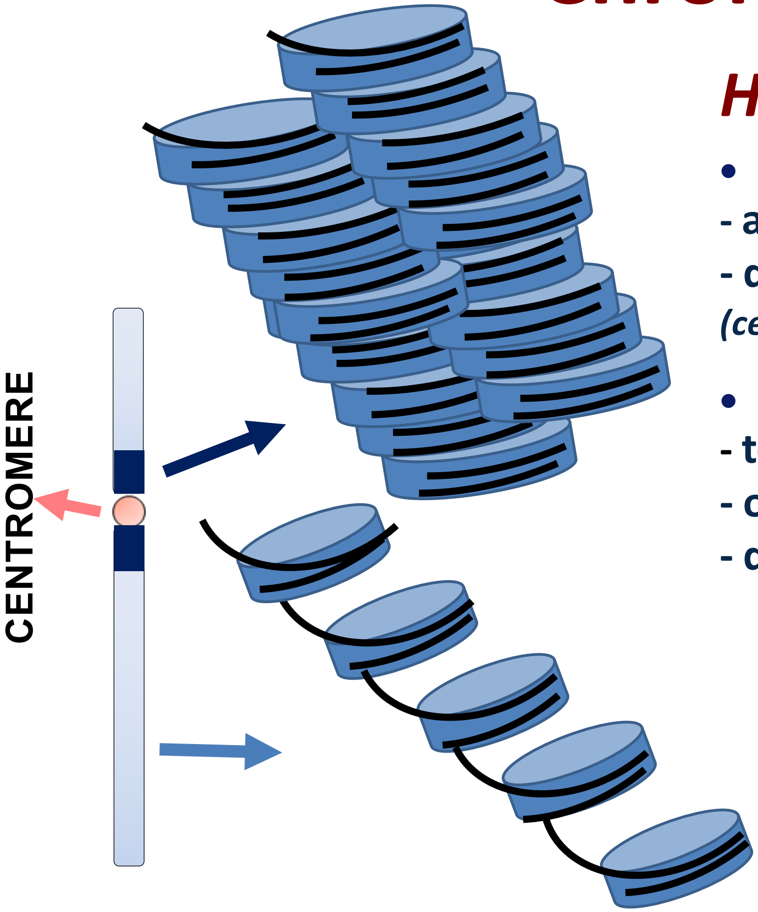
- ***facultative***

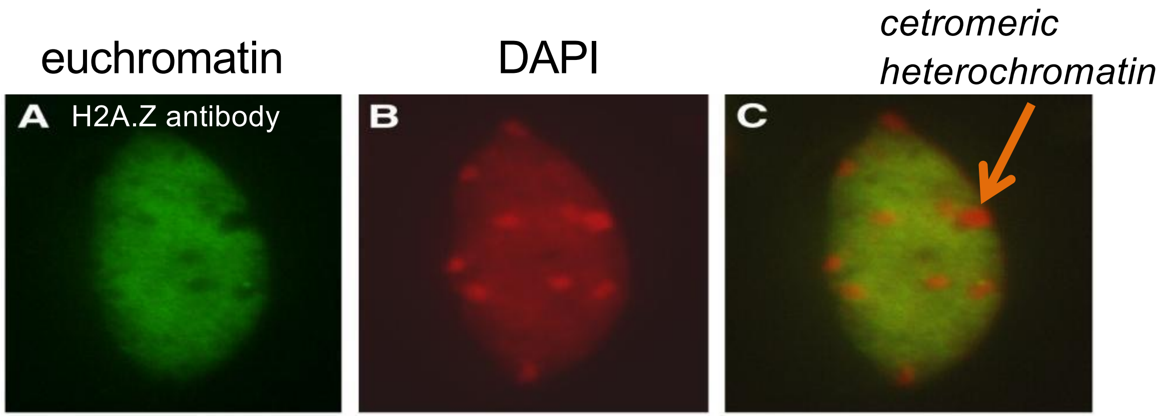
- temporary
- often tissue or cell specific
- during some cell cycle phases

(DNA during mitosis is heterochromatic)

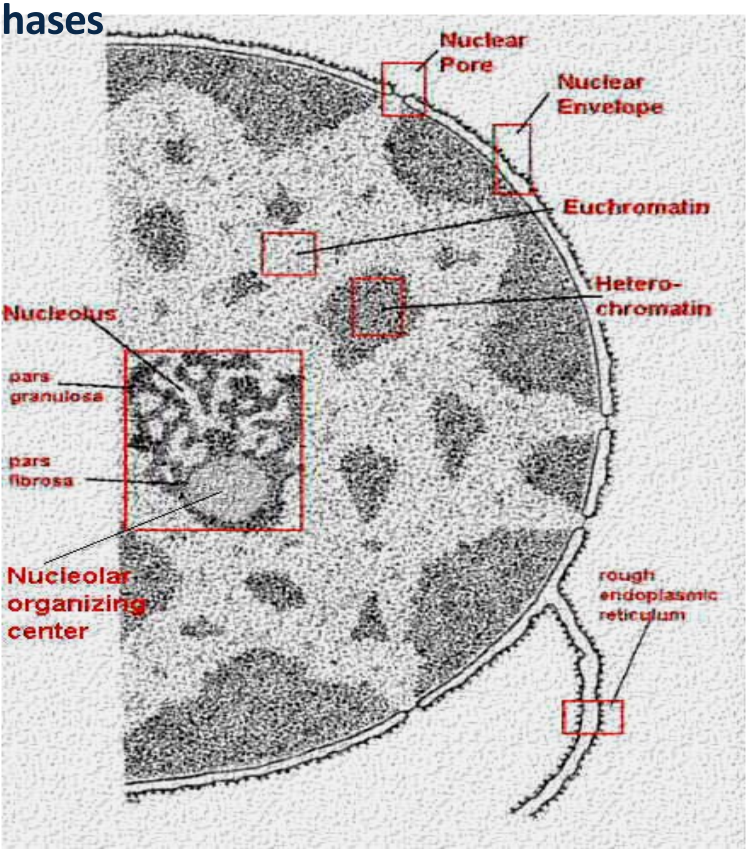
Euchromatin

- loosely packed
- contains transcriptionally active genes

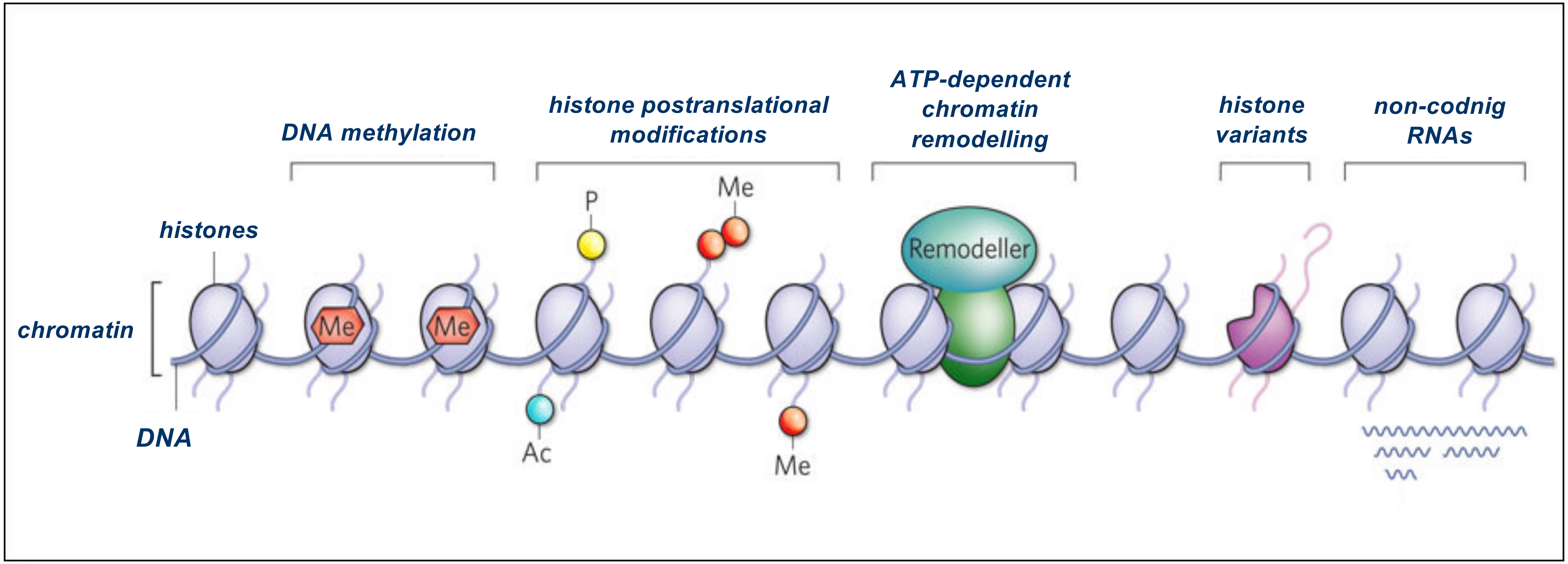




Deal, et al, 2007, Plant Cell

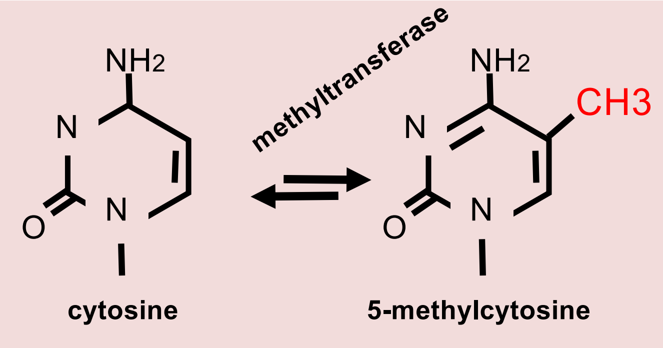


Chromatin: levels of regulation

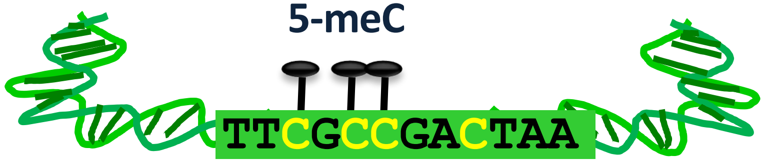


Dulac, 2010, Nature

DNA methylation



Covalent DNA modification in mammals and plants
(some protozoa, fungi and insects)



<u>Function of DNA methylation</u>: imprinting, X chromosome inactivation, embryonic development, silencing of repetitive sequences and transposons

<u>MET1</u> (*METHYLTRANSFERASE1*) – 5'-CG-3’ i 5’-CNG-3’

- ***silencing of* transposons *and DNA repeats***
- ***genomic imprinting***

<u>CMT3</u> (*CHROMOMETHYLASE3*) – 5'-CHG-3’ (*H= A, C or T*)

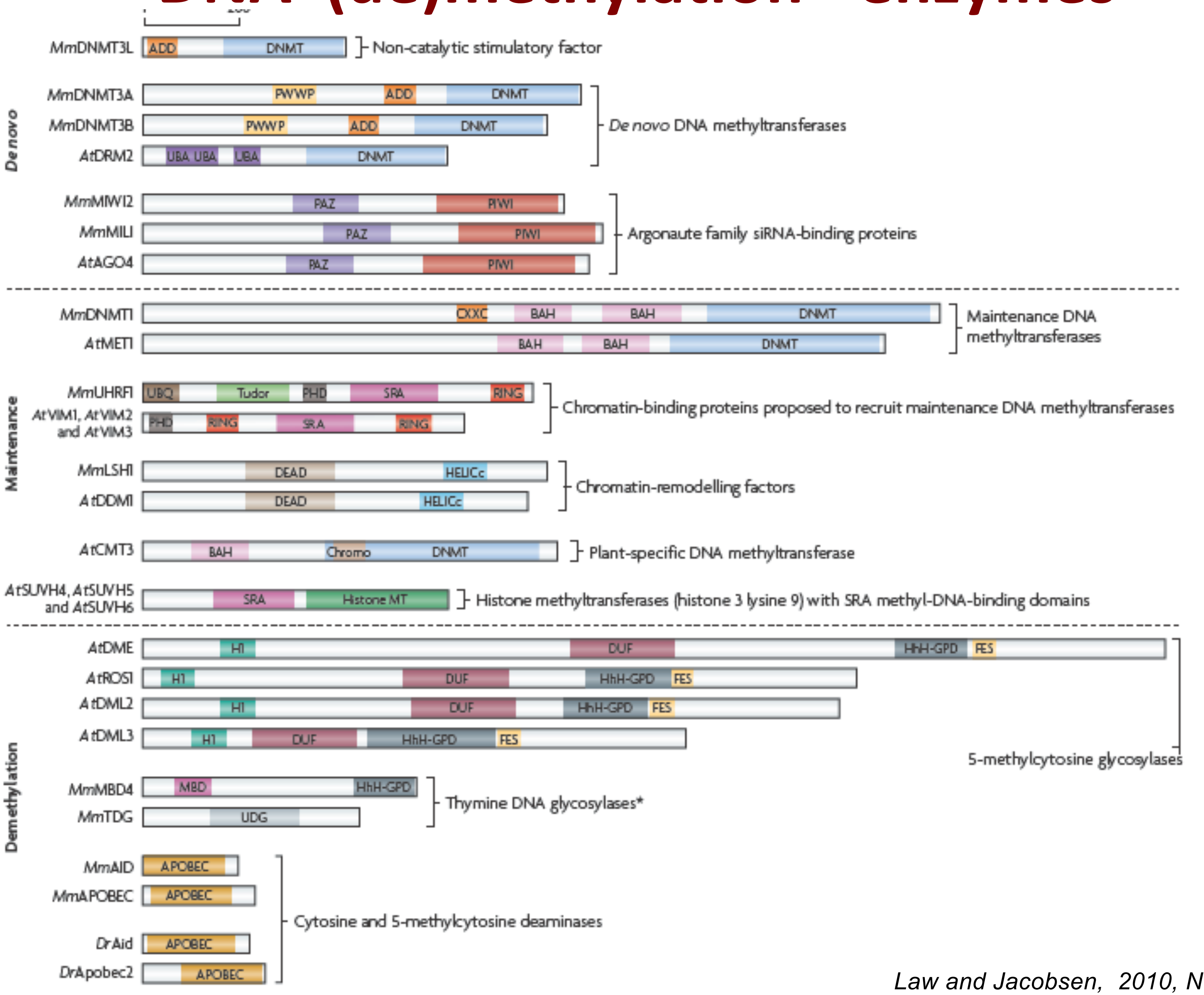
- ***plant specific***
- ***can be recruited by histone methyltransferase SUVH4 (KYP)***
- ***correlated with histone modification***

<u>DRM1/DRM2</u> (*DOMAINS REARRANGED 1/2*) – 5'-CHH-3’

- ***de novo methylation – DRM2***
- ***methylation of DNA repeats silenced by siRNA***

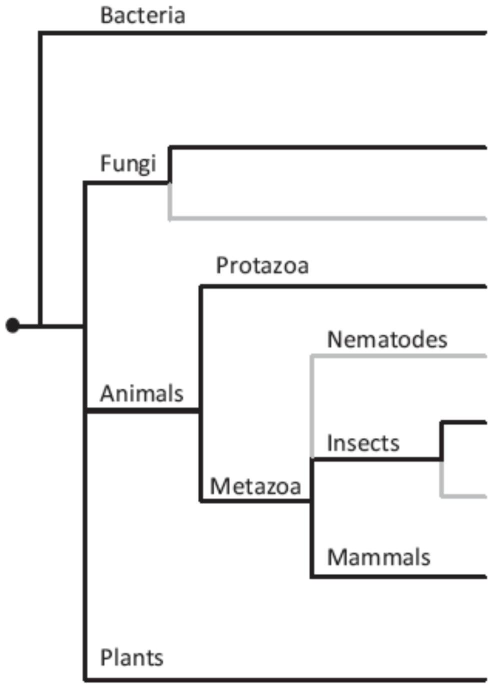
Methylation is reversible (demethylation)

DNA (de)methylation - enzymes

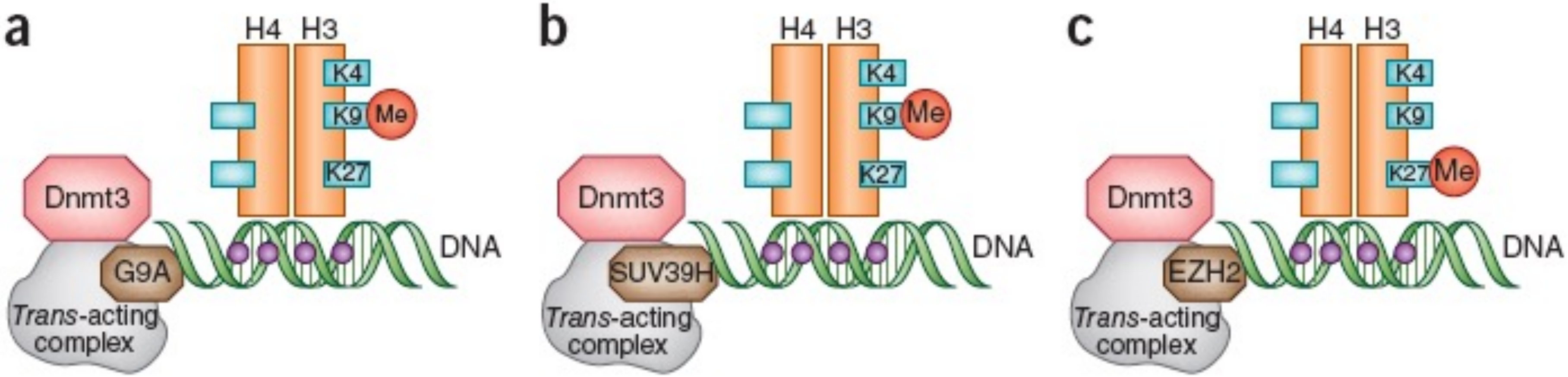


Law and Jacobsen, 2010, NatRevGenet

DNA methylation



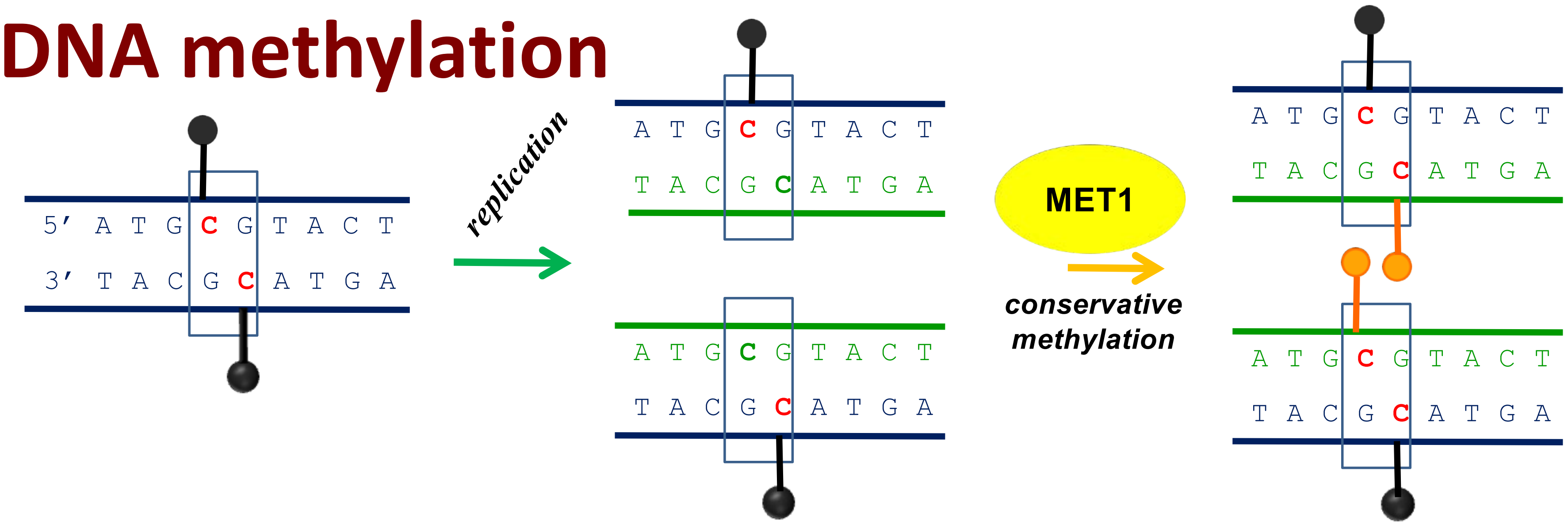
	DNA methylation			DNA demethylation	
	Described			Described	
E. coli	Yes	M-R systems		No	–
		De novo	Maintenance		Factors involved
N. crassa	Yes	Dim2	Dim2	No	–
S. pombe	No	–	–	–	–
T. brucei	Yes	–	–	No	–
C. elegance	No	–	–	–	–
A. mellifera	Yes	Dnmt3	Dnmt1.1 Uhrf1	No	–
D. melanogaster	No	–	–	No	–
M. musculus	Yes	Dnmt3a Dnmt3b Dnmt3l	Dnmt1 Uhrf1	Yes	Mbd4 Tet1 Gadd45 Tet2 Aid Tet3 Apobec Tdg
A. thaliana	Yes	Drm2	Met1 Vim1,2,3 Cmt3	Yes	Dma Ros Dml2



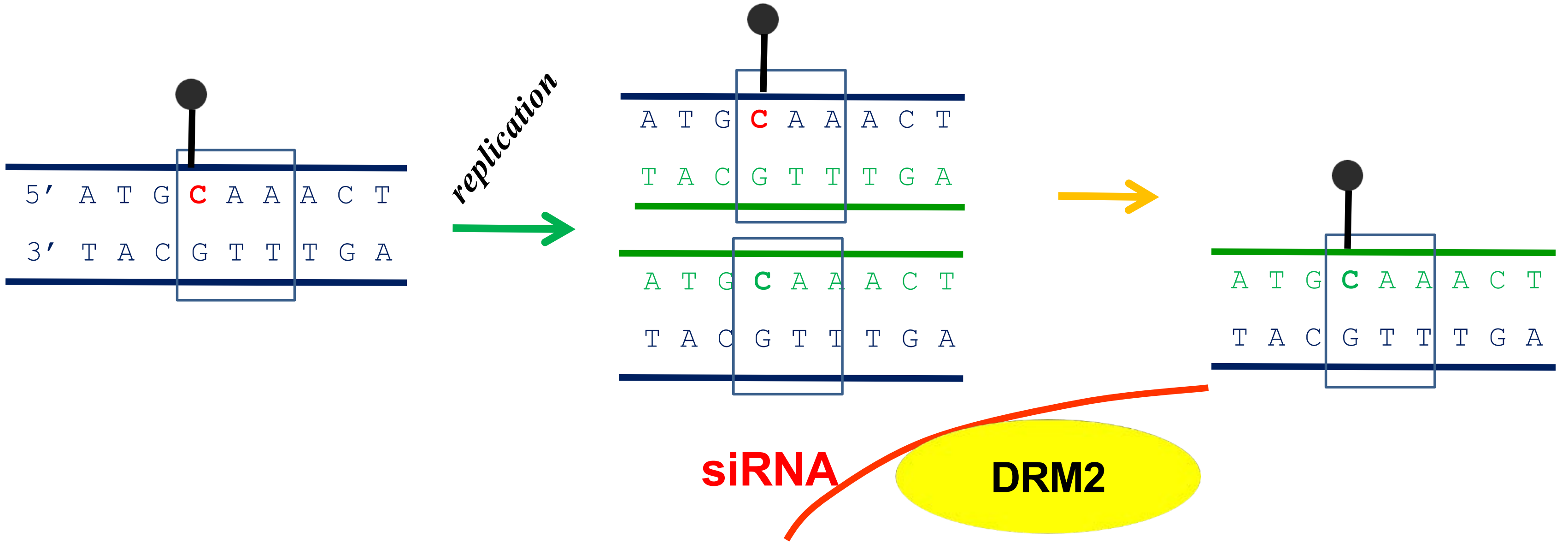
Piccolo and Fisher, 2014, TiCB; Bergman and Cedar, 2013, Nat Str Mol Biol.

Targeted *de novo* site-specific DNA methylation involves also histone methylotransferases

DNA methylation



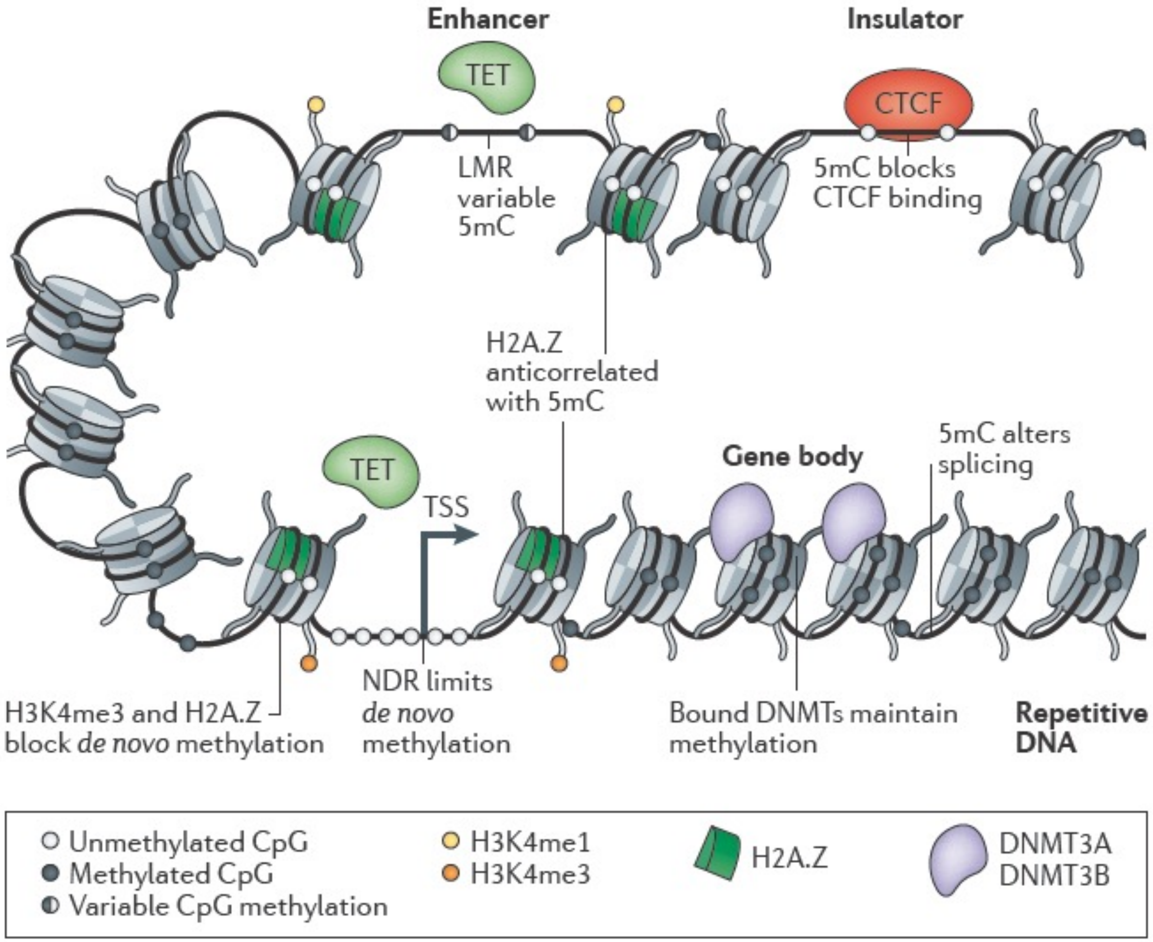
Symmetric CG methylation is maintained by replication *(typical for mammals)*



Initiation and maintenace of asymmetric methylation (CHH) depends on histone modifications and occurs via RNA-directed DNA methylation (RdDM)

Methylation of some cytosines is maintained by siRNA and RdDM

DNA methylation



CpG methylation: - TSS are unmethylated, when methylated → silencing (XCI)

- methylation of repetitive elements → silencing (transposons, LINEs, Alu)

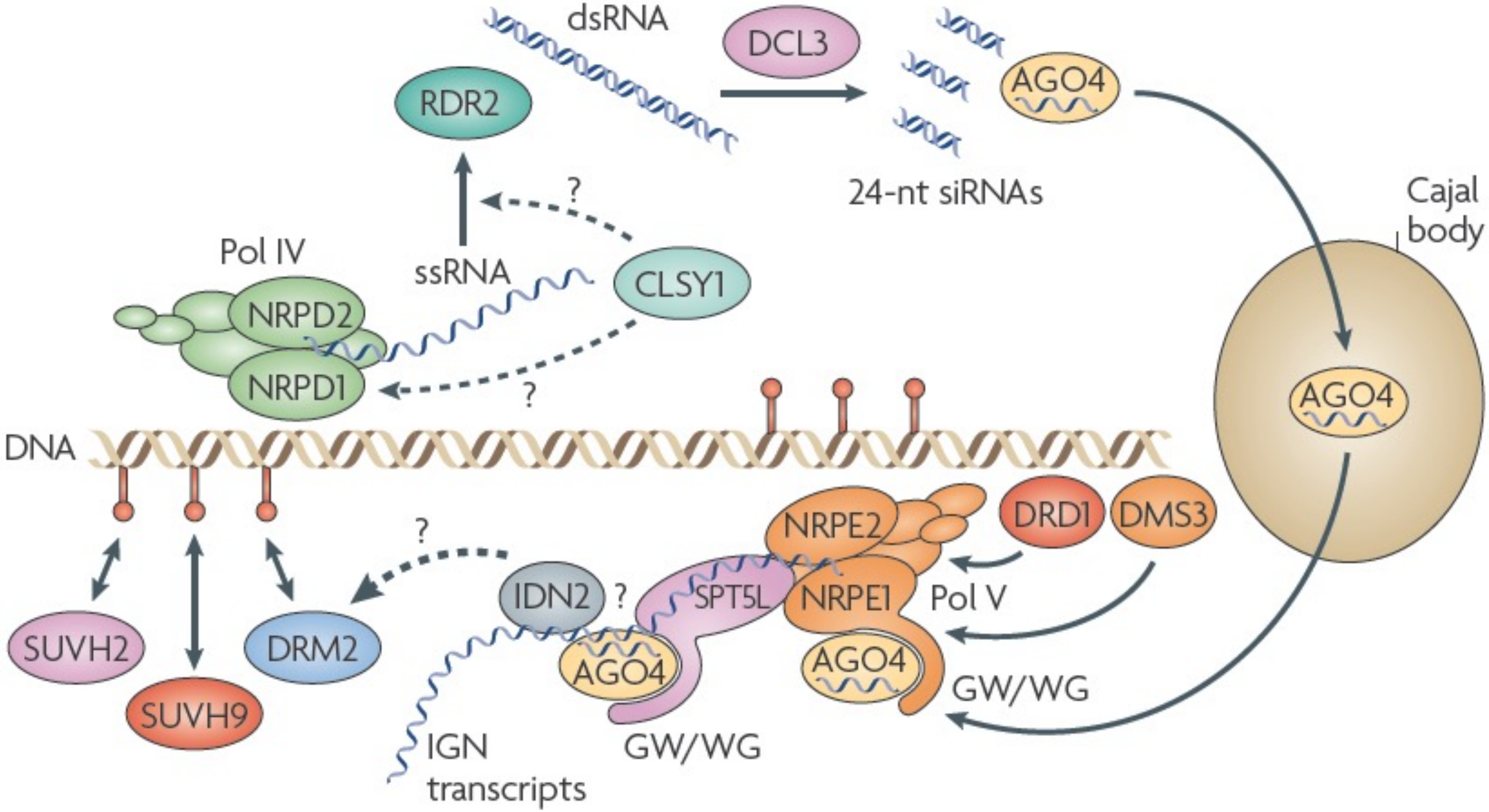
- methylation blocks transcription start not elongation

- methylation in gene bodies is not associated with repression

- other methylation sites: enhancers, insulators, splicing

Jones, 2012, Nat Rev Genet

DNA methylation: silencing (plants)



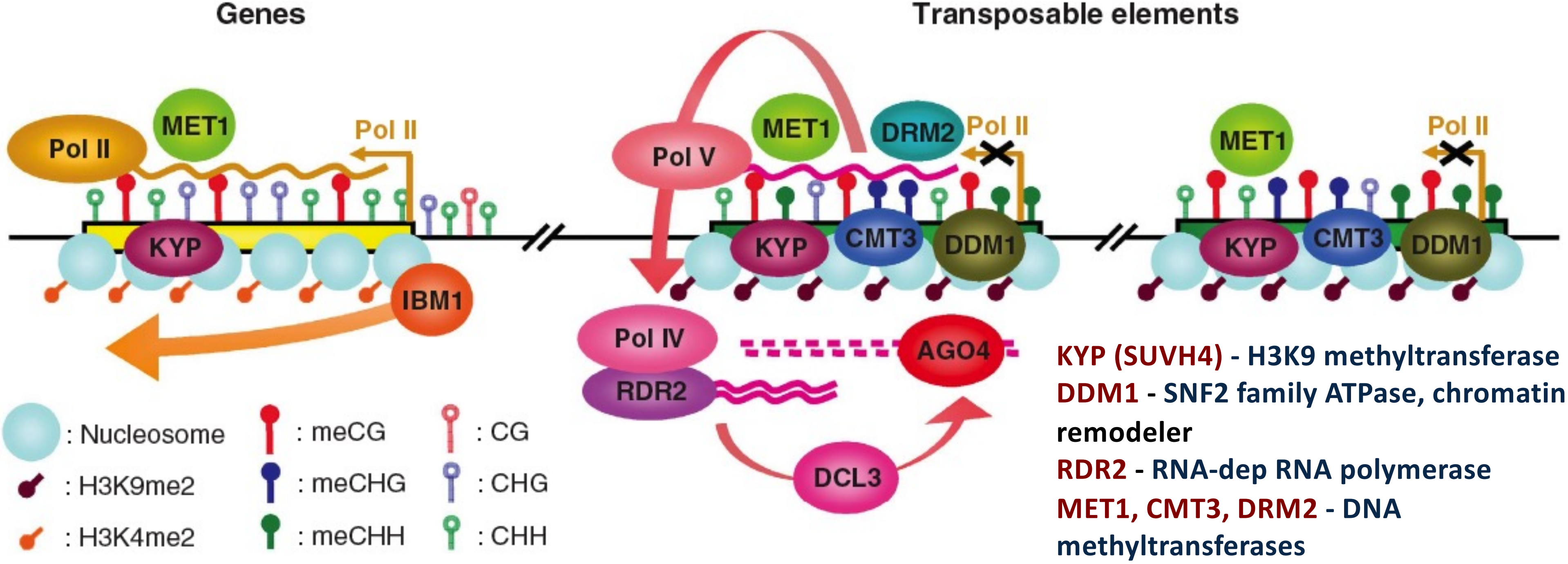
ssRNA RNA Pol IV transcripts converted to dsRNAs by RDR2 are processed to siRNAs by DCL3 and associate with AGO4. Nascent Pol V ncRNAs IGS serve as a scaffold for AGO4/siRNAs and other factors and target DRM2, SUV2, SUV9 that bind meDNA

Law and Jacobsen, 2010, NatRevGenet

DNA methylation: silencing (plants)

of repetitive sequences and transposons (TE) by RNAi

A. thaliana



KYP (SUVH4) - H3K9 methyltransferase
DDM1 - SNF2 family ATPase, chromatin remodeler
RDR2 - RNA-dep RNA polymerase
MET1, CMT3, DRM2 - DNA methyltransferases
IBM1- histone demethylase

Methylated DNA bind MBD (*methyl-CpG binding domain*) proteins which recruit histone deacetylase complex and histone methyltransferase. This leads to chromatin condensation and gene repression.

DNA methylation is affected by nucleosme positioning, methylases are targeted to nucleosomes.

Only a subset of methylated TEs are targeted by RNAi

Texeira and Colot, 2009, EMBO J.

DNA methylation

Intragenic DNA methylation prevents spurious transcription methylation

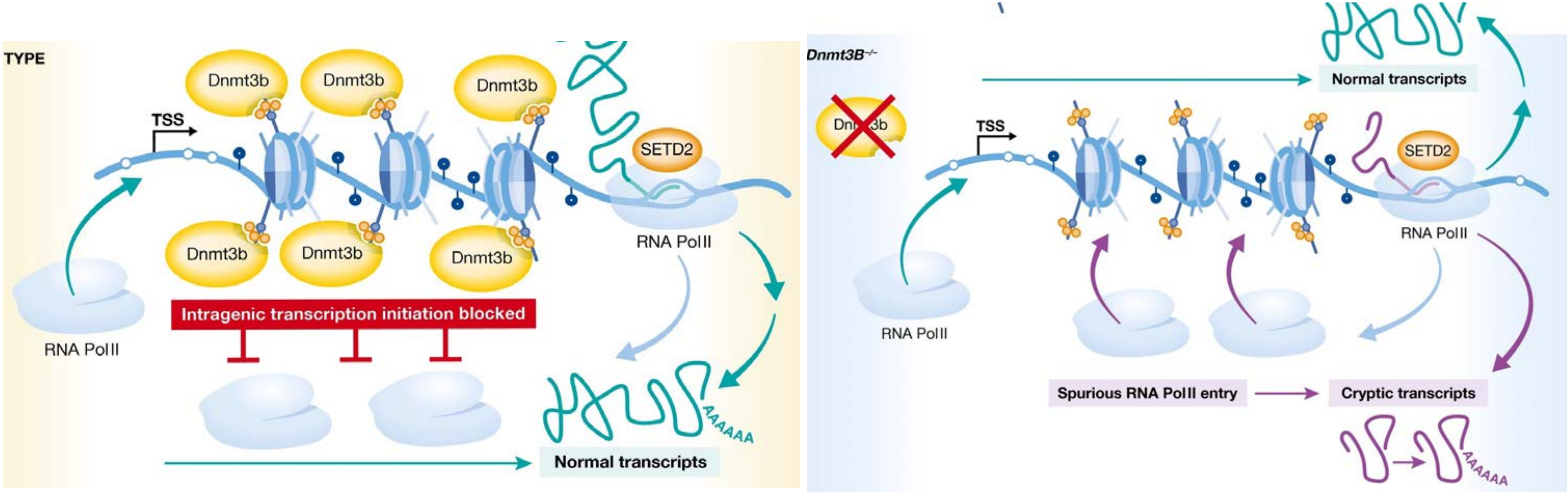
DNA methylation in mammals occurs mainly at CpG.

Methylation of the promoter suppresses gene expression.

Gene-body DNA methylation protects the gene body from spurious Pol II entry and cryptic transcription initiation (shown in mouse embryonic stem cells)

Such spurious transcripts can either be degraded by the exosome or capped, polyadenylated, and delivered to the ribosome to produce aberrant proteins.

Elongating Pol II triggers DNA methylation to ensure the fidelity of transcription initiation.



Neri et al, Nature, 2015; Teissandier and Bourchis, EMBO, 2017

Histone modifications - chromatin structure

Histones N tails outside the nucleosome are accessible to modifying enzymes

$$\text{S/T} \xrightleftharpoons[\textit{phosphatases}]{\textit{kinases}} \text{-P S/T}$$

histone code:
phosphorylation (P)

$$\text{K} \xrightleftharpoons[\textit{histone deacetylases (HDAC)}]{\textit{histone acetylotransferases (HAT)}} \text{-COCH3 K}$$

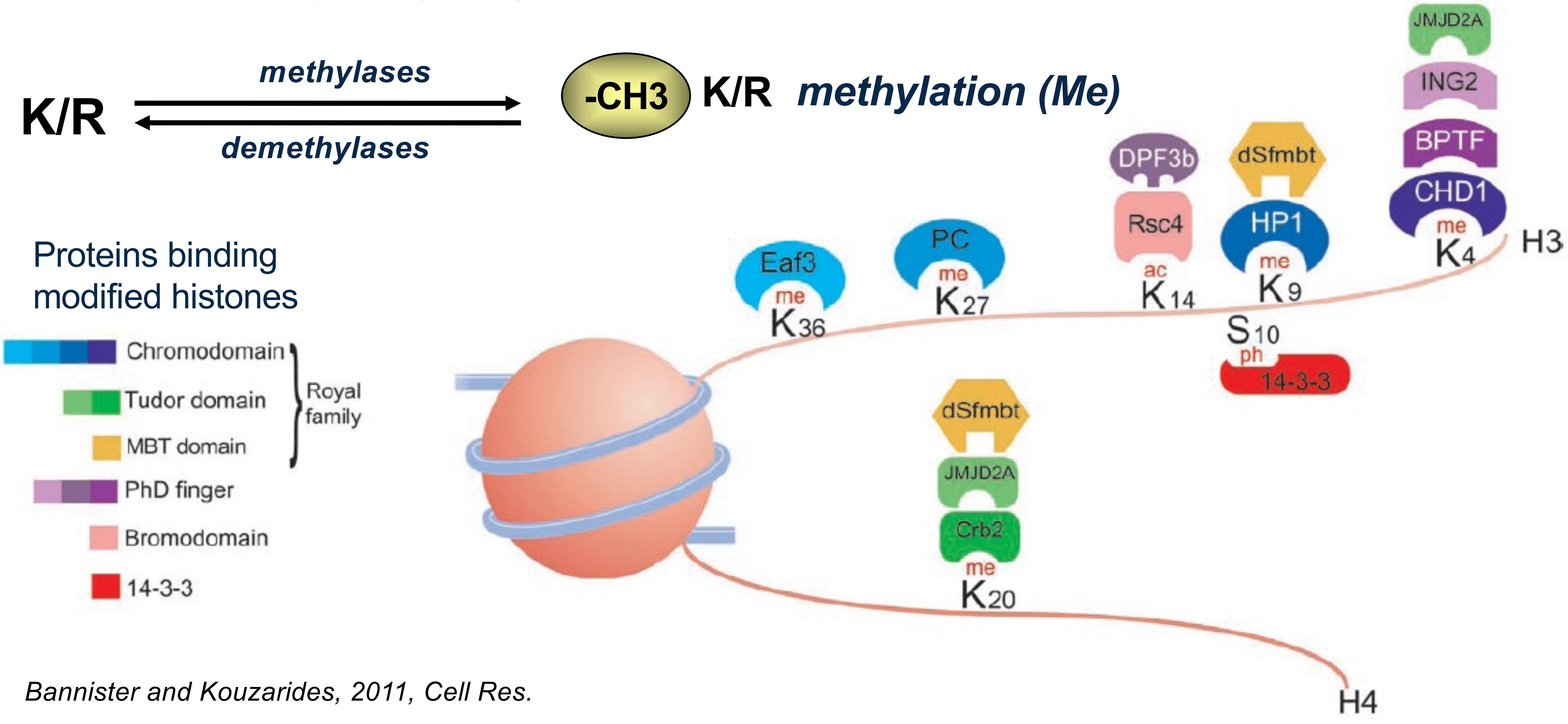
acetylation (Ac)

ubiquitination (Ub)
sumoylation (Su)

$$\text{K/R} \xrightleftharpoons[\textit{demethylases}]{\textit{methylases}} \text{-CH3 K/R}$$

methylation (Me)

Proteins binding
modified histones



Bannister and Kouzarides, 2011, Cell Res.

Histone modifications - chromatin structure

-COCH3 *Lysine-Ac*

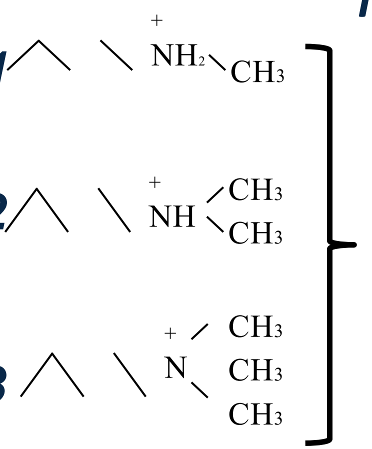
K-Ac more neutral than K
reduces DNA-histone interaction
loosens chromatin structure

-CH3 *Lysine-Me*

Kme1

Kme2

Kme3



K-Me increases hydrophobic and
cationic character of aa

-P *Serine-P*

S-P on H3S10 and H3S28 activate transcription
by inhibiting H3K9-Me and promoting K-Ac

Histone modifications affect chromatin structure or regulate binding of chromatin factors

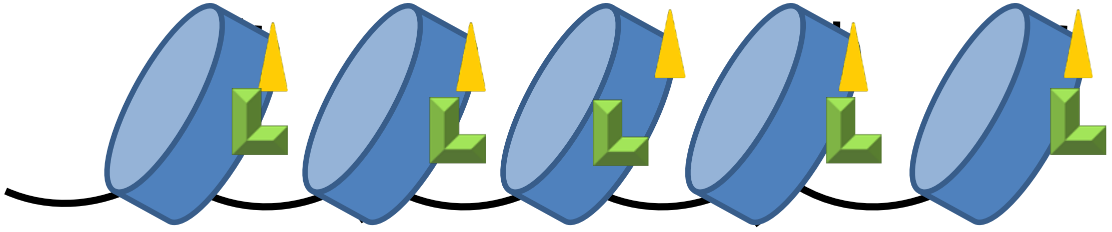
K-Ac and S-P reduce the positive charge of histones, loosen chromatin and activate transcription

K-Me (K-Ac) act mainly via protein binding, may inhibit or activate transcription:

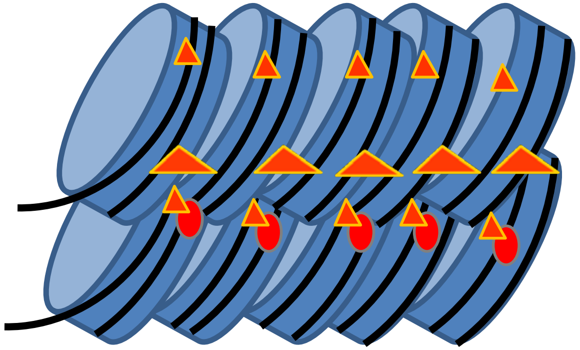
H3K4me3 – active transcription mark (recognized by PD finger proteins, can recruit DNA modifying enzymes)

H3K9me3 - repressive chromatin mark (recognized by HP1)

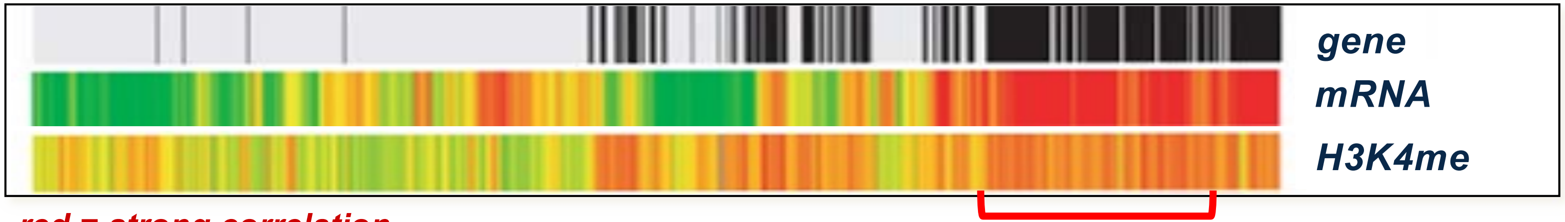
Histone modifications - chromatin structure



H3	Me	P	Ac
	K4	S10	K14

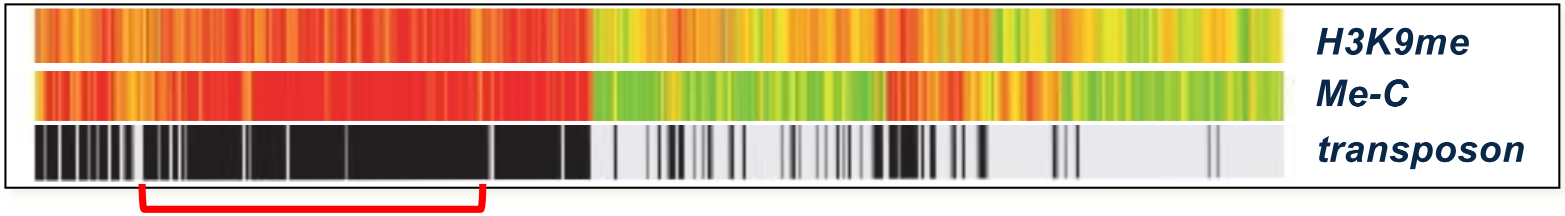


H3	Me	Me	P
	K9	K27	S28



red = strong correlation
green= weak correlation

H3K4 methylation in actively transcribed genes

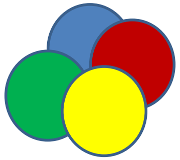
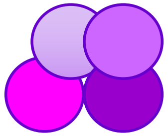
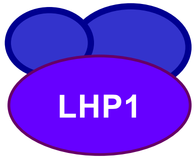


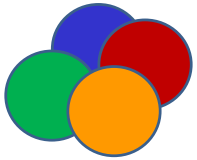
H3K9 methylation correlated with methylated DNA (meC) and trasposons

Lippman et al., 2004, Nature

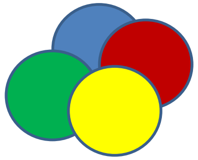
Histone modifications in coding regions and transposons differ

H3K27me3 – methylating complexes

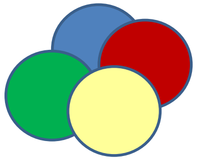
PRC2	***D. melanogaster***	***A. thaliana***
methylation establishment	*Enhancer of zeste* E(Z) *(methylase)*	*CURLY LEAF (CLF)* *MEDEA (MEA)* *SWINGER (SWN)*
Polycomb Repressive Complex 2 	*Extra sex comb* ESC	*FERTILIZATION INDEPENDENT ENDOSPERM (FIE)*
	Suppressor of zeste 12 SU(Z)12	*FERTILIZATION-INDEPENDENT SEED 2 (FIS2)* *EMBRYONIC FLOWER 2 (EMF2)* *VERNALIZATION 2 (VRN2)*
	NURF55	*MULTICOPY SUPPRESSOR OF IRA1 (MSI1,2,3,4,5)*
methylation maintenance	***PRC1*** 	***PRC1-like*** 



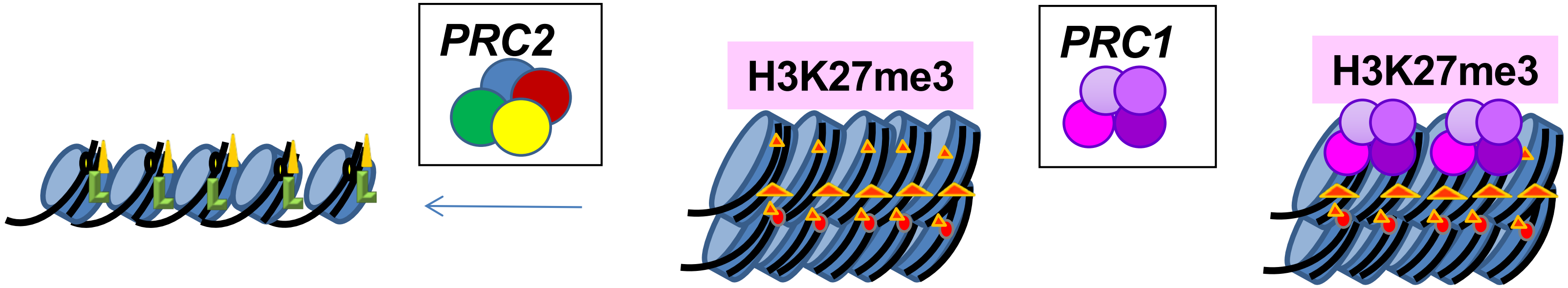
MEA + FIS2
germination



CLF/SWN + VRN2
induction of flowering

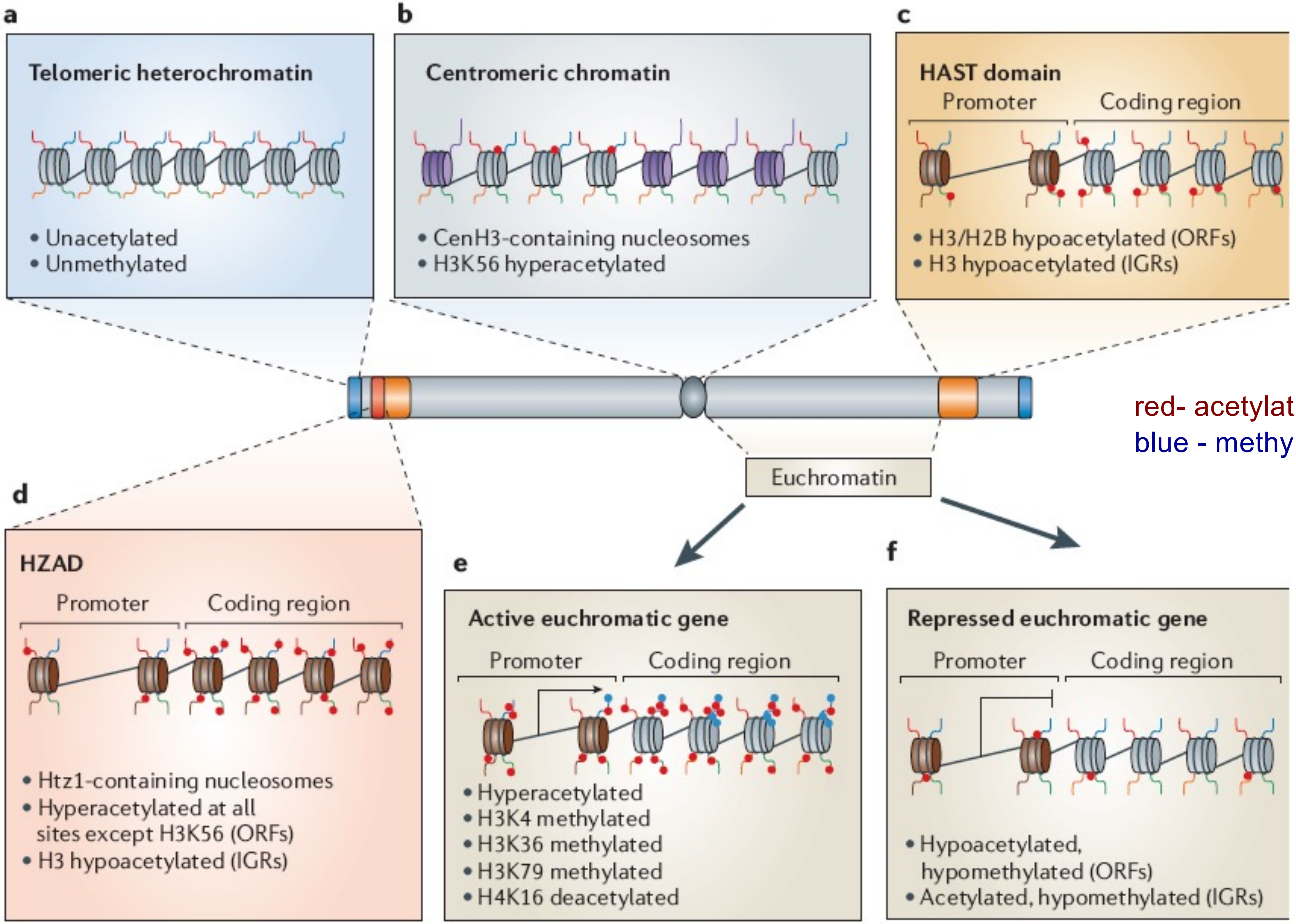


CLF/SWN + EMF2
flower development



active genes   *silenced genes*   *stably silenced genes*

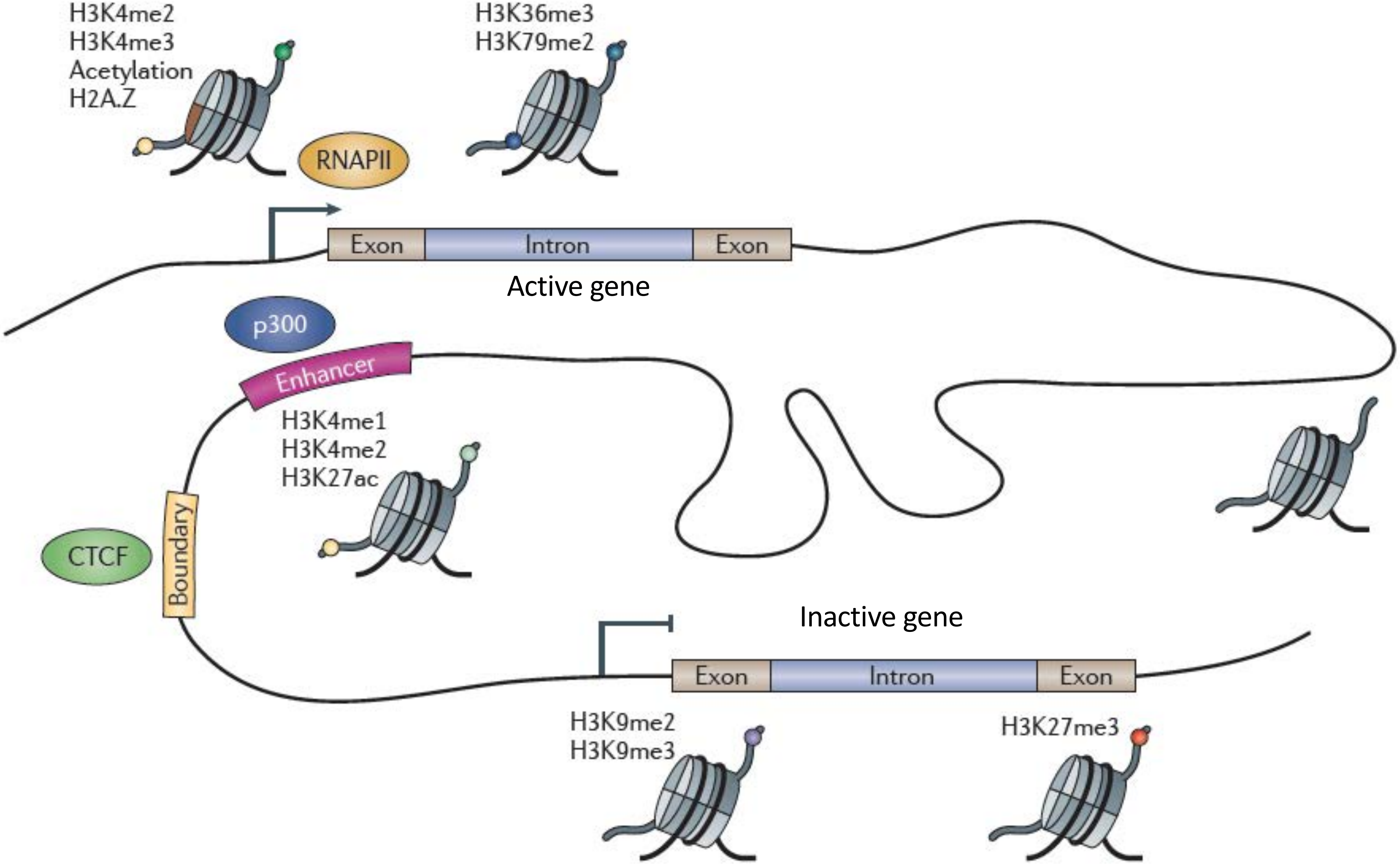
Histone code



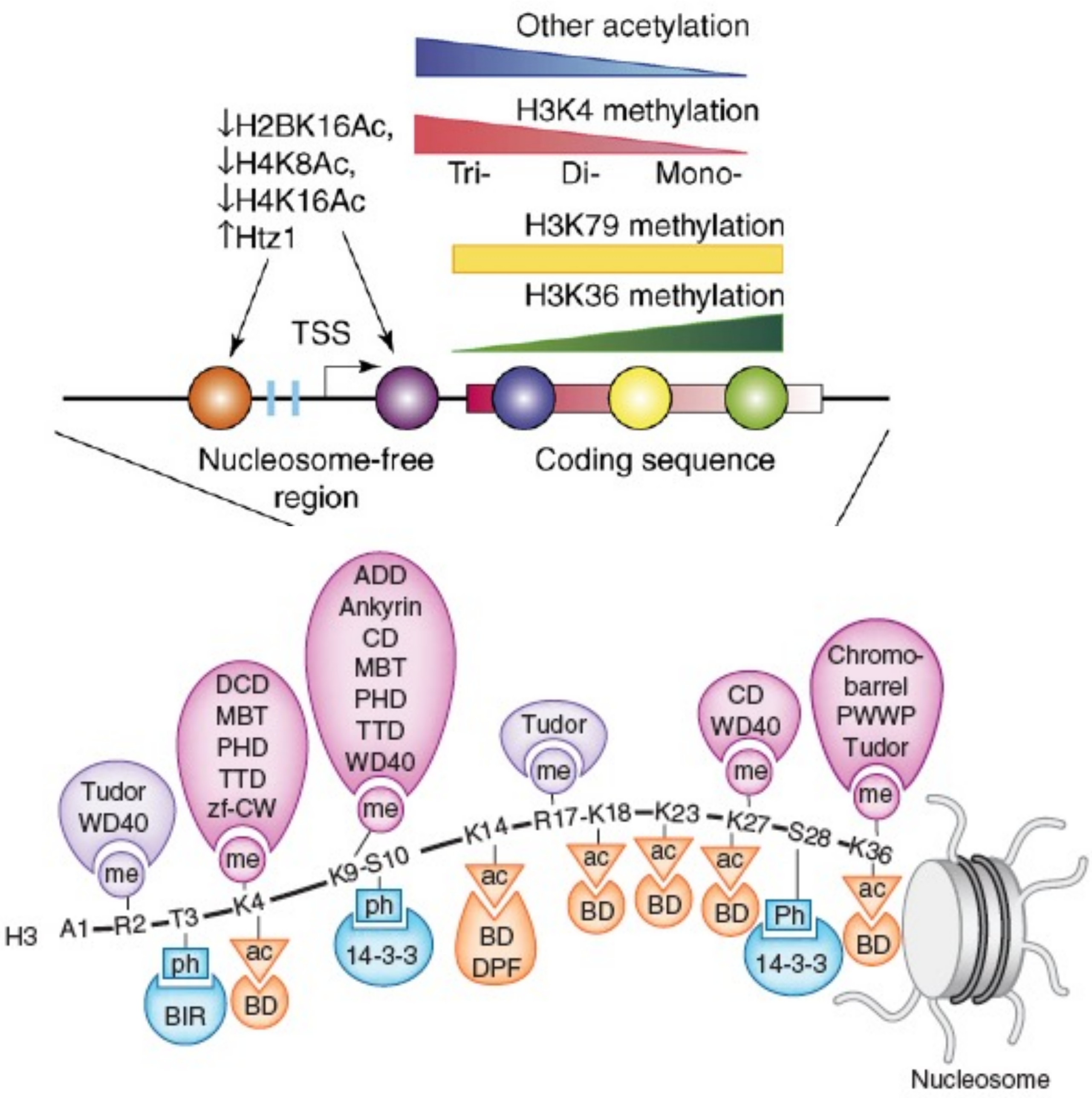
red- acetylation
blue - methylation

Every genomic element has a separate specific histone modification pattern

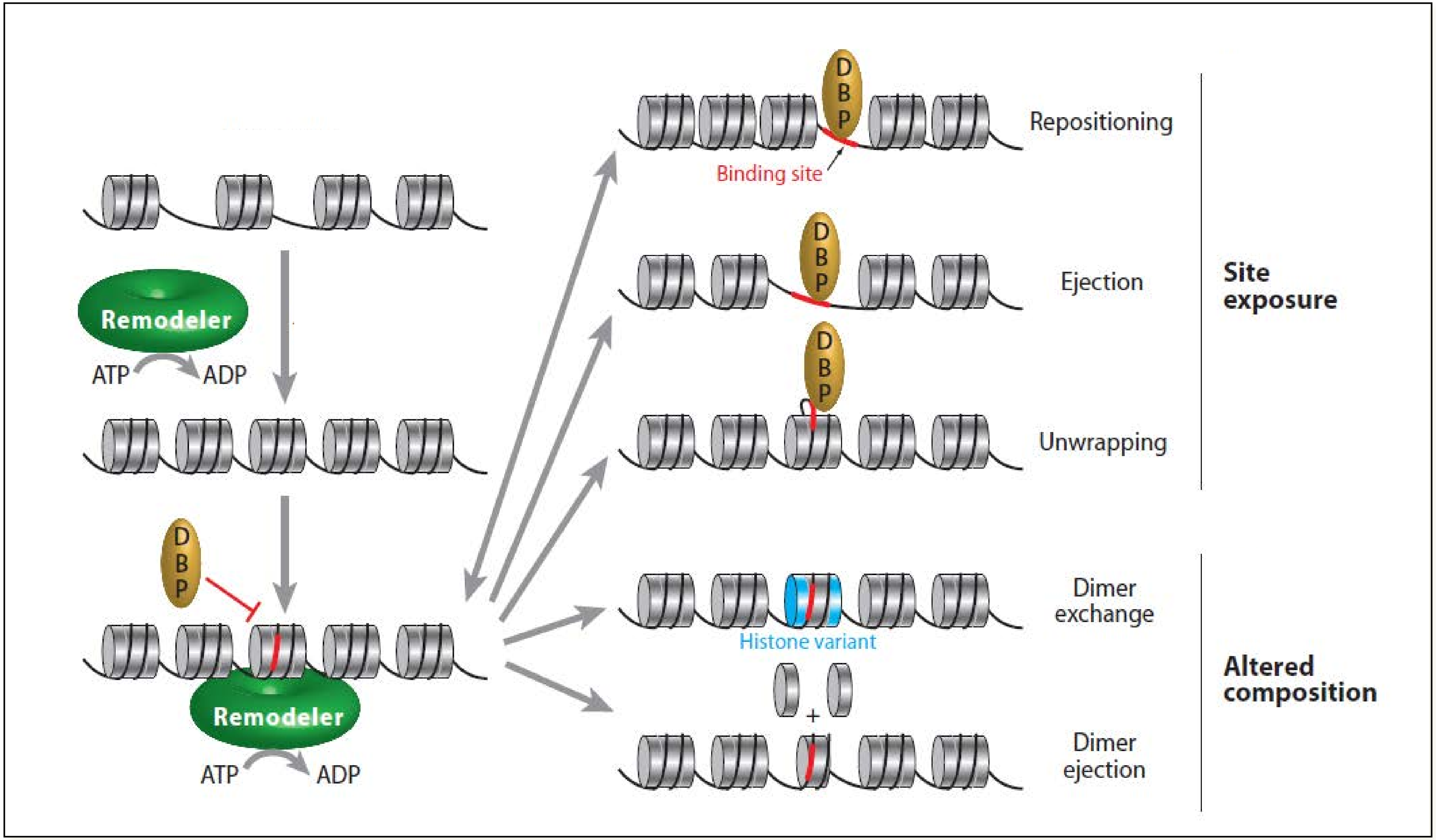
Histone code



Histone code: histone readers

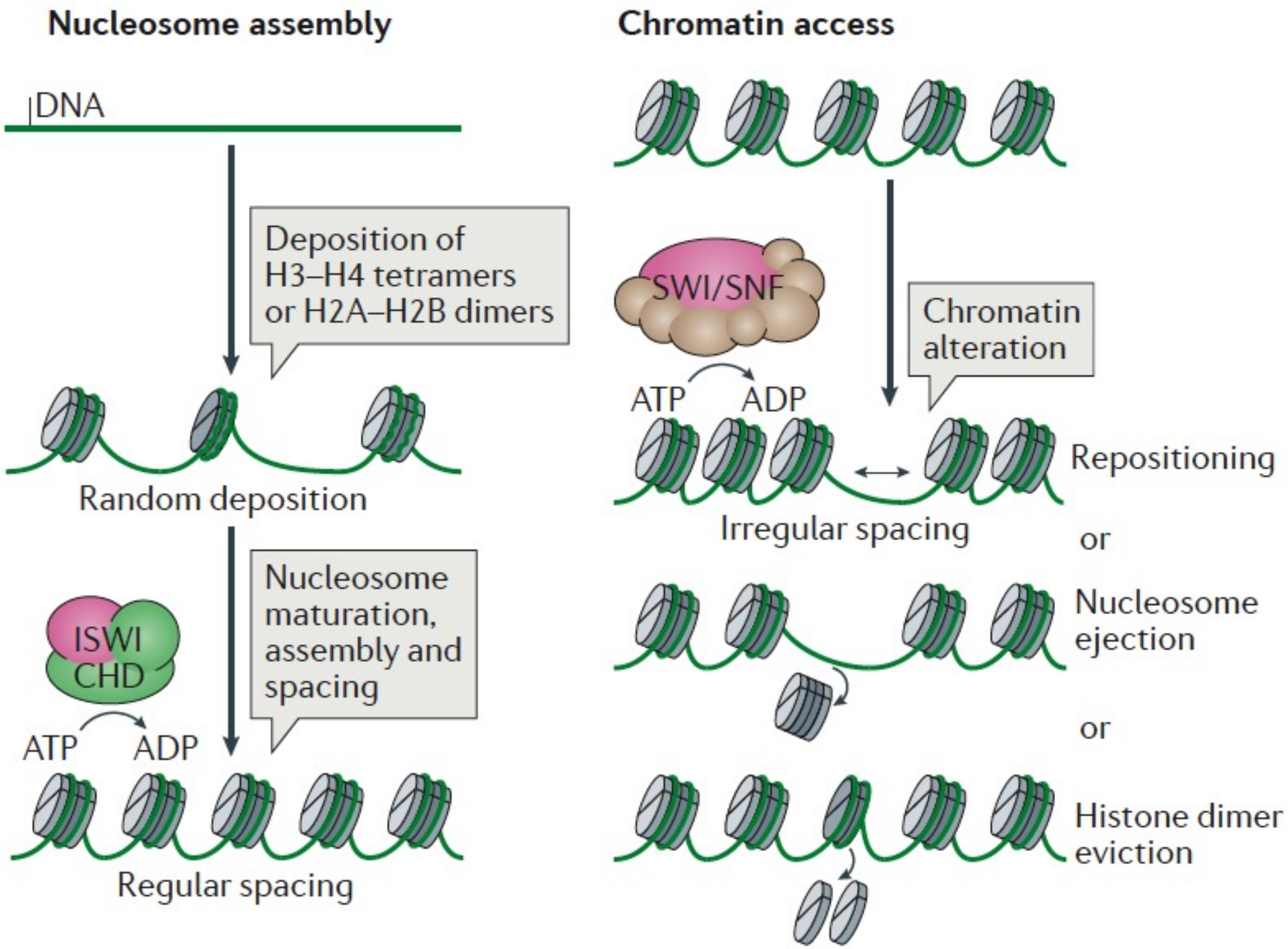


Chromatin remodeling



Clapier and Cairns, 2009, Annu Rev Biochem

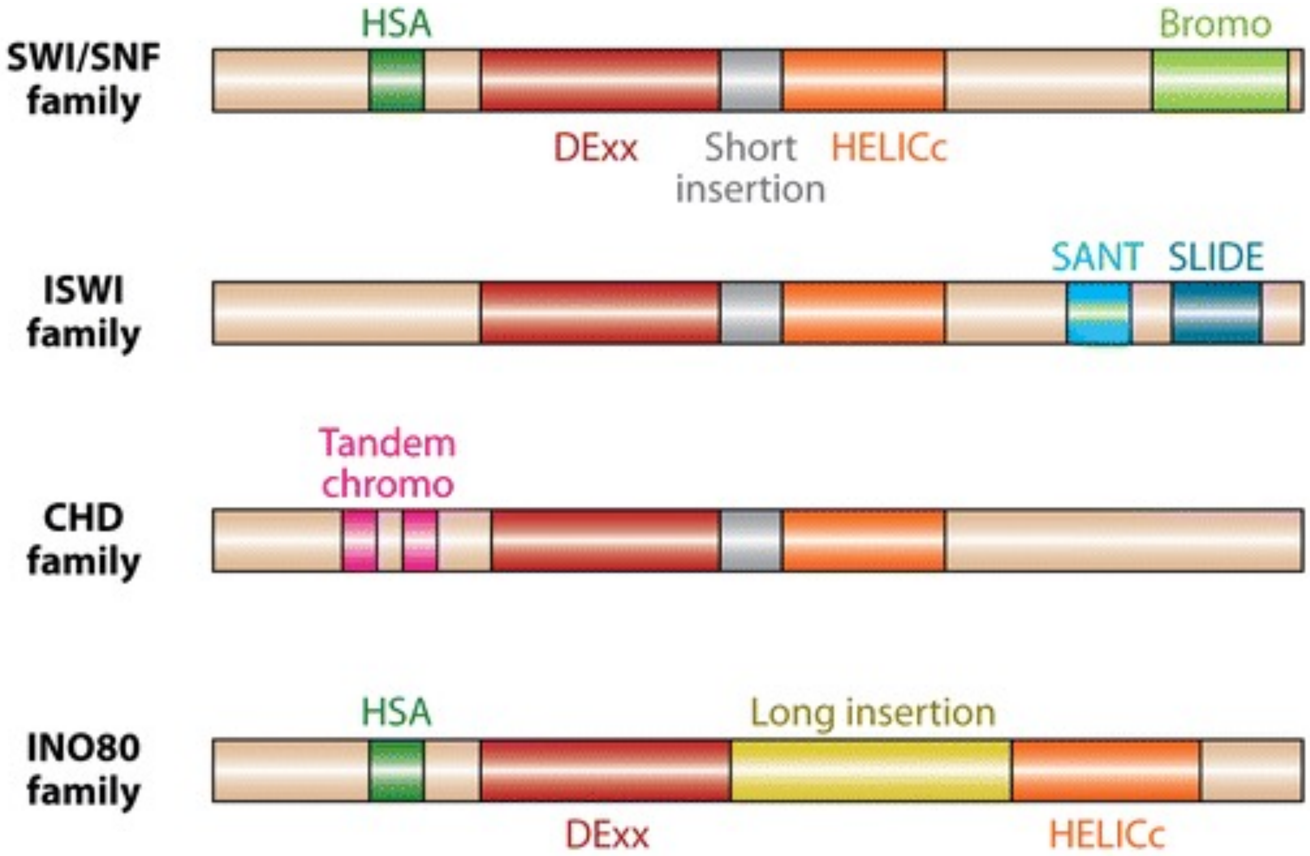
Chromatin remodeling



Clapier et al, Nat Rev Mol Cell Biol, 2017

Chromatin remodellers

ATP-dependent nucleosome-remodelling complexes



SWI2:
contain bromodomain
all remodelling types

ISWI:
nucleosome repositioning

CHD:
contain chromodomain
transcription regulation

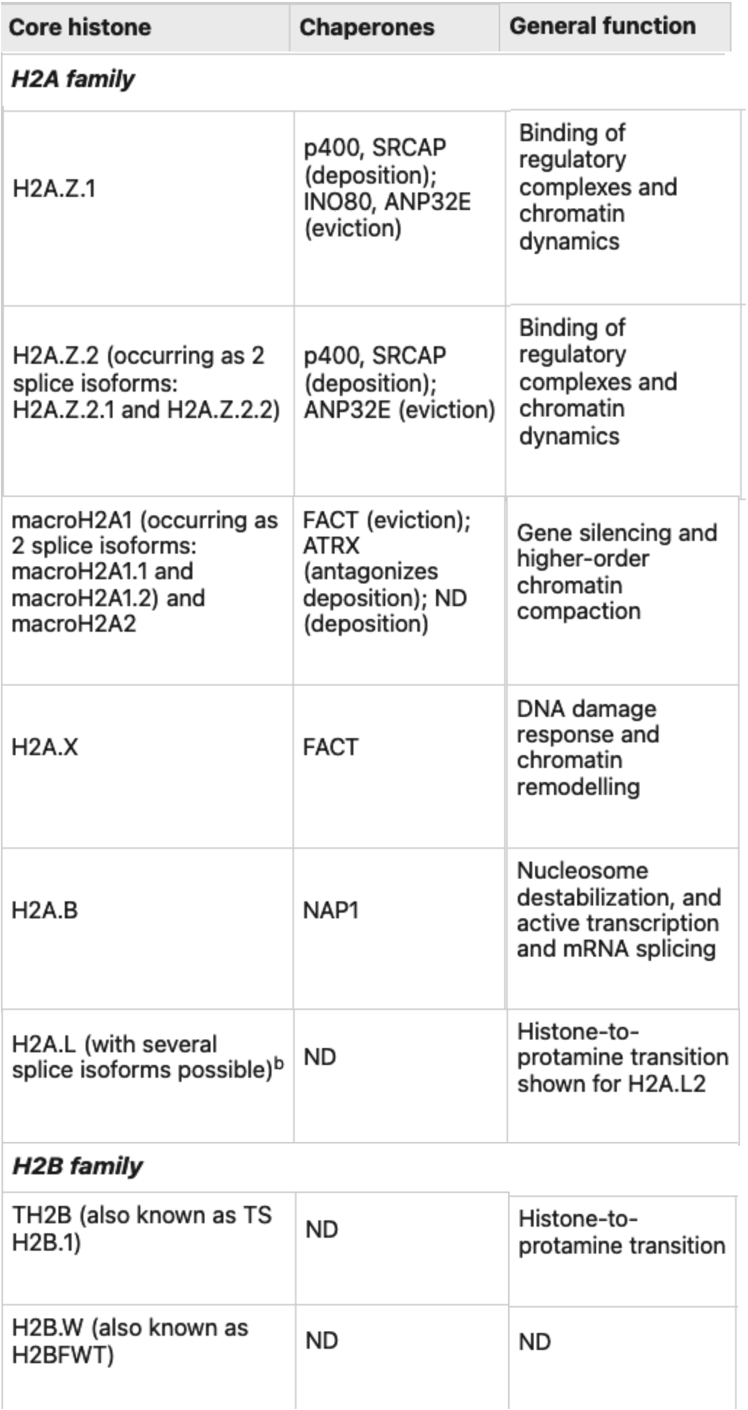
INO80/SWR:
histone exchange

Example:
Arabidopsis DDM1, ATPase from the SWI/SFN family
- **involved in transposon methylation**
- **links chromatin remodeling and DNA methylation**

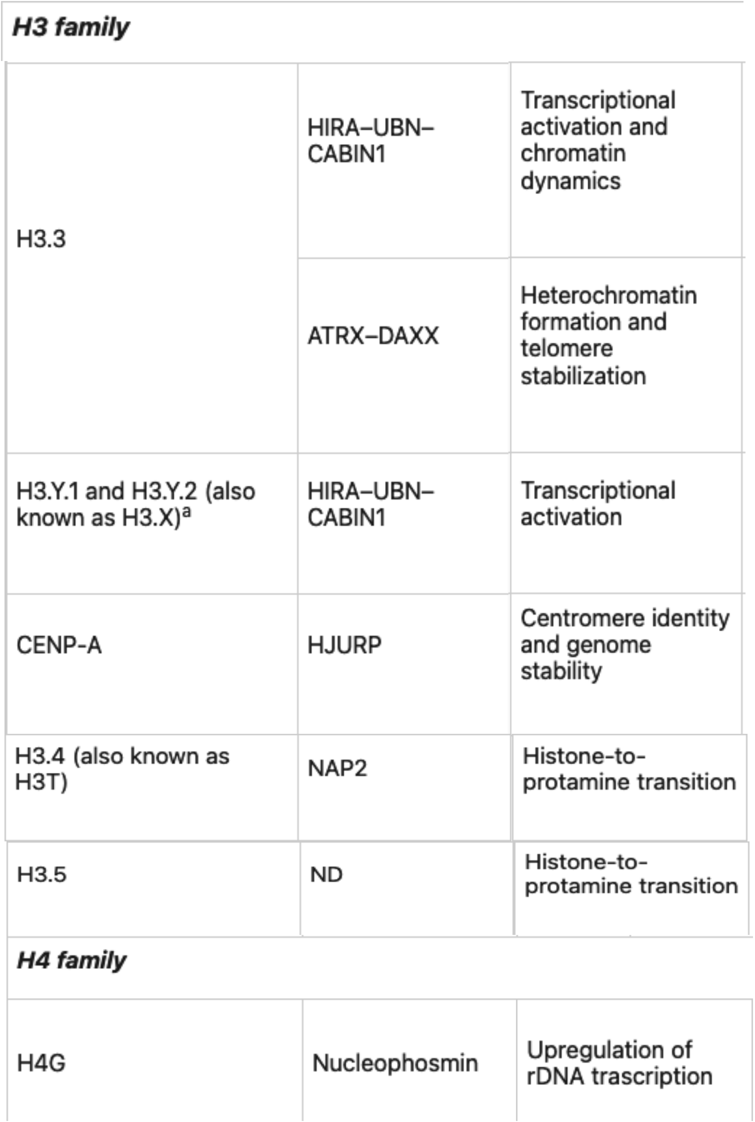
Clapier and Cairns, 2009, Annu Rev Biochem

Histone variants

- **distinct protein sequence**
- **designated chaperones**
- **different chromatin remodeling complexes**
- **specific post-translational modifications**



Core histone	Chaperones	General function
H2A family		
H2A.Z.1	p400, SRCAP (deposition); INO80, ANP32E (eviction)	Binding of regulatory complexes and chromatin dynamics
H2A.Z.2 (occurring as 2 splice isoforms: H2A.Z.2.1 and H2A.Z.2.2)	p400, SRCAP (deposition); ANP32E (eviction)	Binding of regulatory complexes and chromatin dynamics
macroH2A1 (occurring as 2 splice isoforms: macroH2A1.1 and macroH2A1.2) and macroH2A2	FACT (eviction); ATRX (antagonizes deposition); ND (deposition)	Gene silencing and higher-order chromatin compaction
H2A.X	FACT	DNA damage response and chromatin remodelling
H2A.B	NAP1	Nucleosome destabilization, and active transcription and mRNA splicing
H2A.L (with several splice isoforms possible)[b]	ND	Histone-to-protamine transition shown for H2A.L2
H2B family		
TH2B (also known as TS H2B.1)	ND	Histone-to-protamine transition
H2B.W (also known as H2BFWT)	ND	ND

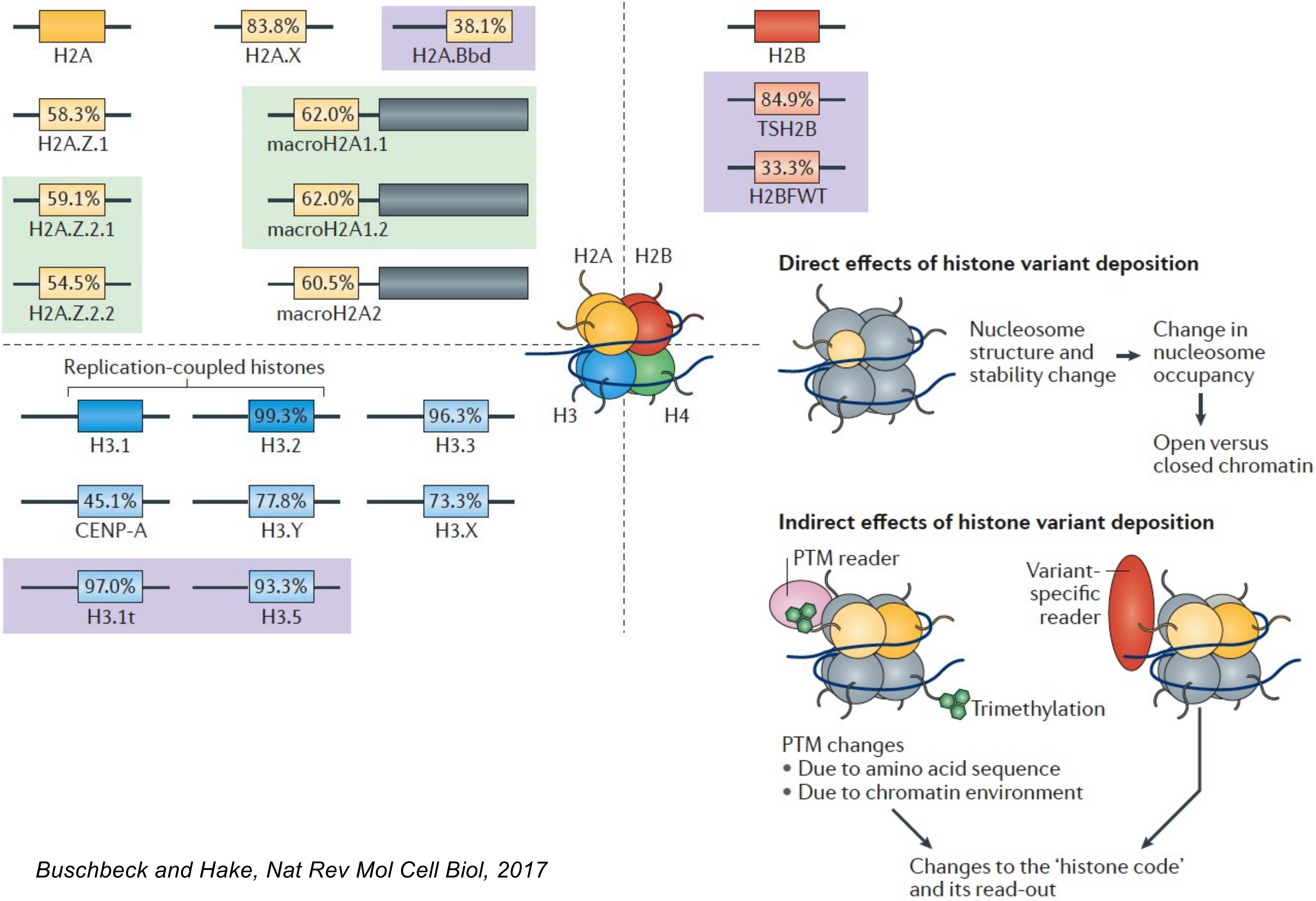


H3 family		
H3.3	HIRA–UBN–CABIN1	Transcriptional activation and chromatin dynamics
	ATRX–DAXX	Heterochromatin formation and telomere stabilization
H3.Y.1 and H3.Y.2 (also known as H3.X)[a]	HIRA–UBN–CABIN1	Transcriptional activation
CENP-A	HJURP	Centromere identity and genome stability
H3.4 (also known as H3T)	NAP2	Histone-to-protamine transition
H3.5	ND	Histone-to-protamine transition
H4 family		
H4G	Nucleophosmin	Upregulation of rDNA trascription

Martire and Banaszynski, Nat Rev Mol Cell Biol, 2020

Histone variants

human cells



Direct effects of histone variant deposition

Nucleosome structure and stability change → Change in nucleosome occupancy → Open versus closed chromatin

Indirect effects of histone variant deposition

PTM changes
- Due to amino acid sequence
- Due to chromatin environment

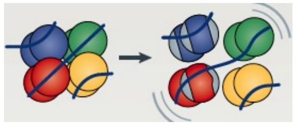
Changes to the 'histone code' and its read-out

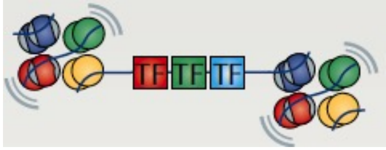
Buschbeck and Hake, Nat Rev Mol Cell Biol, 2017

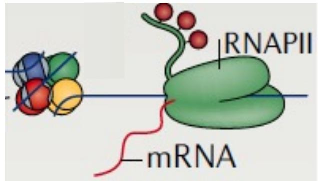
Histone variants

Histones are exchanged by INO80/SWR remodeling complexes

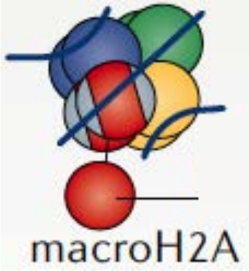
Decreased nucleosome stability → **Destabilization of DNA-histone and histone-histone interactions** → **Chromatin opening TF binding** → **Transcription**







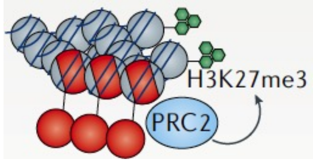
acidic patch

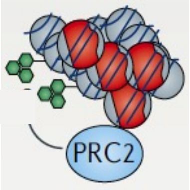


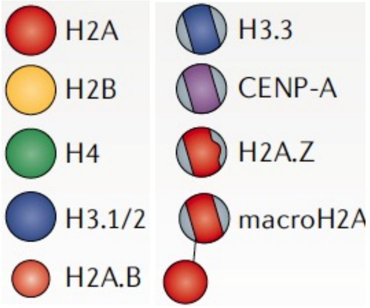
extra domain

protruding loop

Increased nucleosome stability → **Chromatin compaction Recruitment of silencing complexes** → **Silencing**

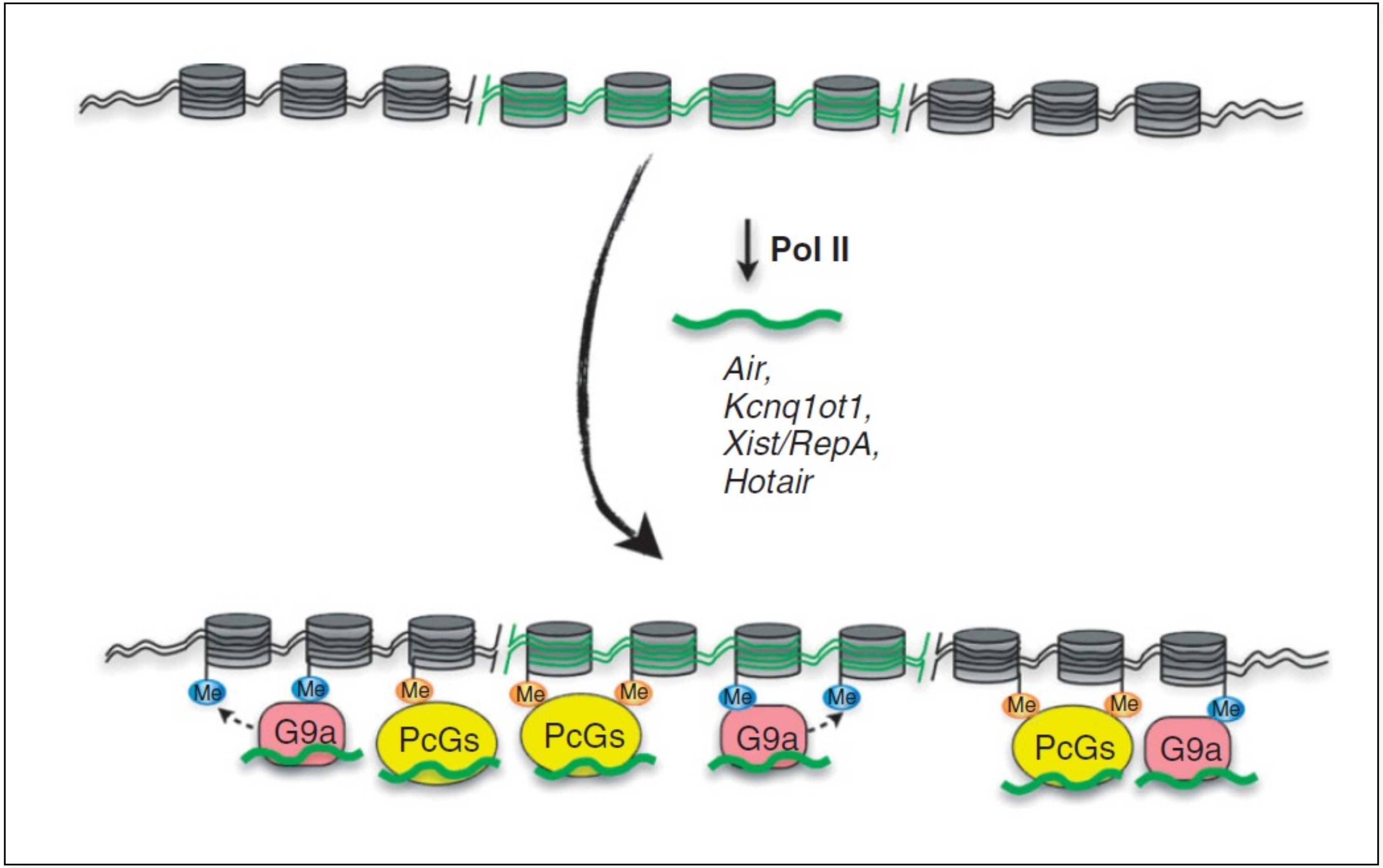






Martire and Banaszynski, Nat Rev Mol Cell Biol, 2020

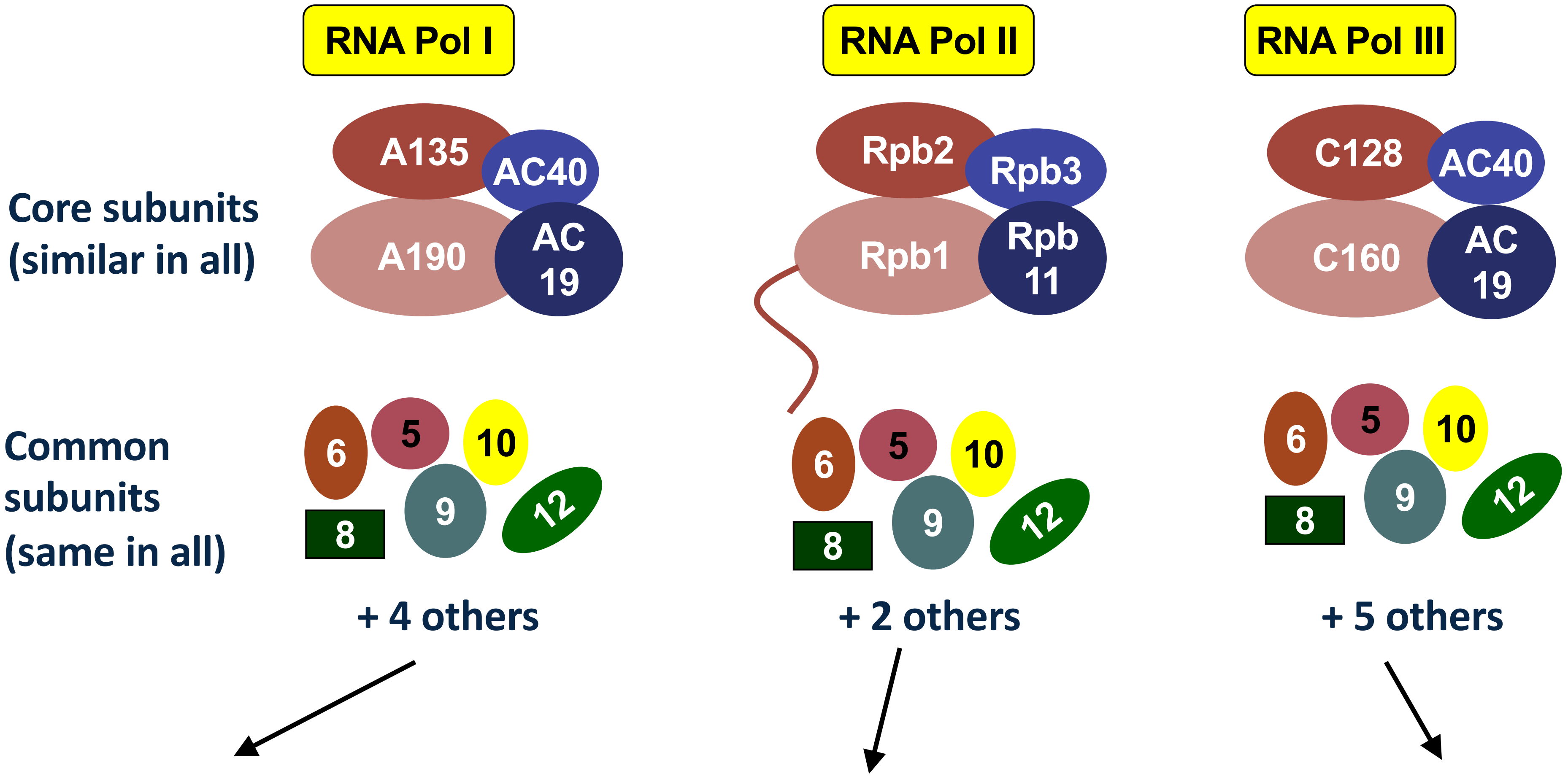
Epigenetic chromatin modification by ncRNAs



Some ncRNA (Pol II transcripts), such as siRNAs or lncRNAs, recruit silencing complexes to specific genomic loci

Chen and Carmichael, 2010, WIREs RNA

RNA Polymerases



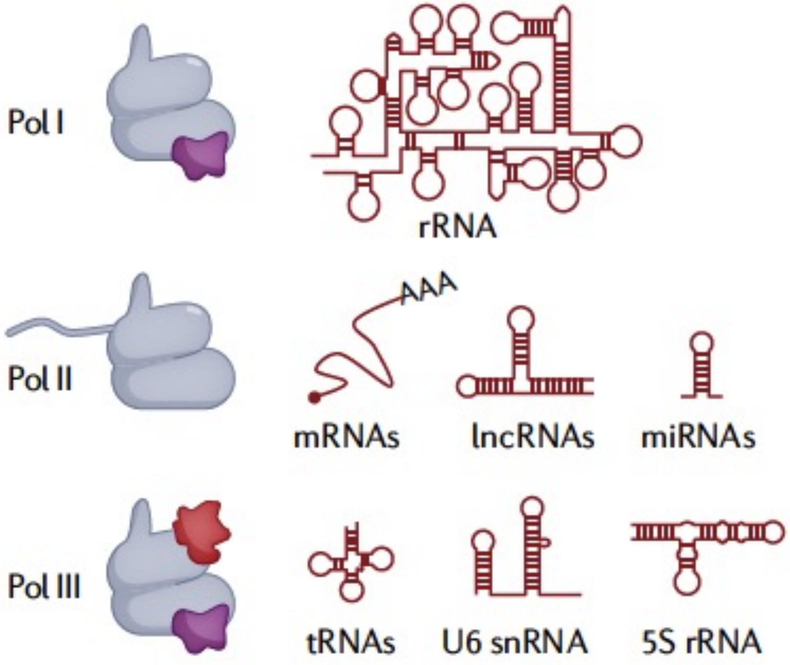
ribosomal RNA
35S precursor contains
18S, 5.8S and 25S subunits

mRNA , most snRNAs
(U1, U2, U3, U4, U5,
U11, U12, U4atac),
snoRNAs, microRNAs,
telomerase RNA

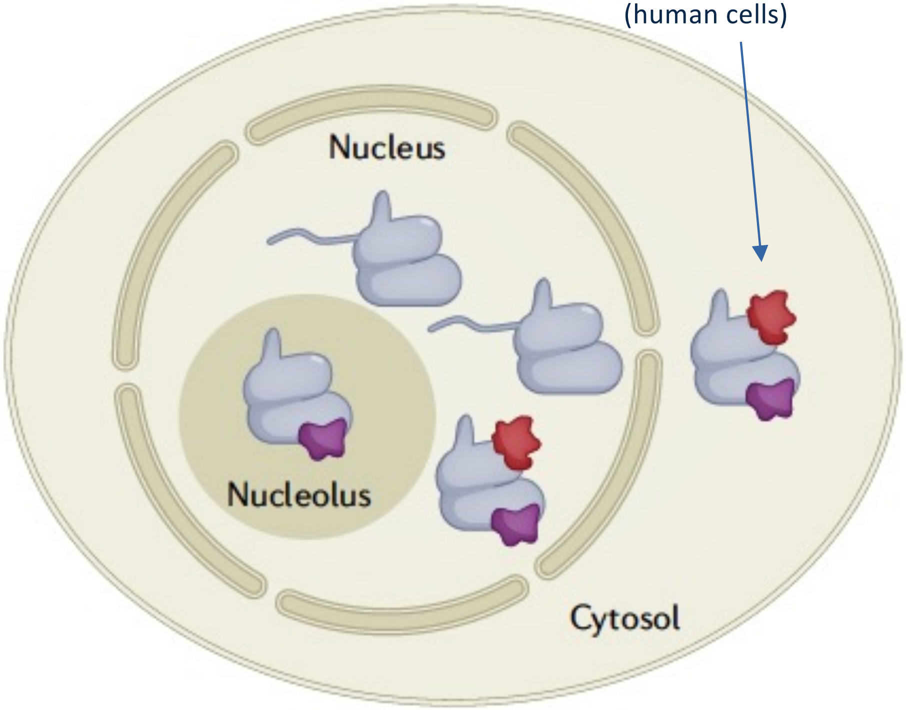
tRNA, 5S rRNA, U6
snRNA, U6atac snRNA,
7SK RNA, 7SL RNA,
RNase P RNA,
RNase MRP RNA

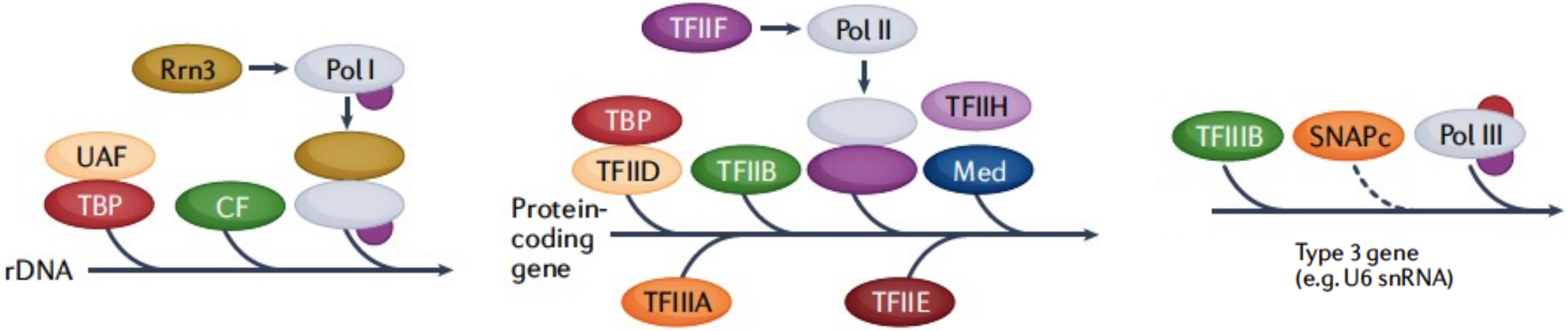
Zbigniew Dominski, lectures 2008

RNA Polymerases



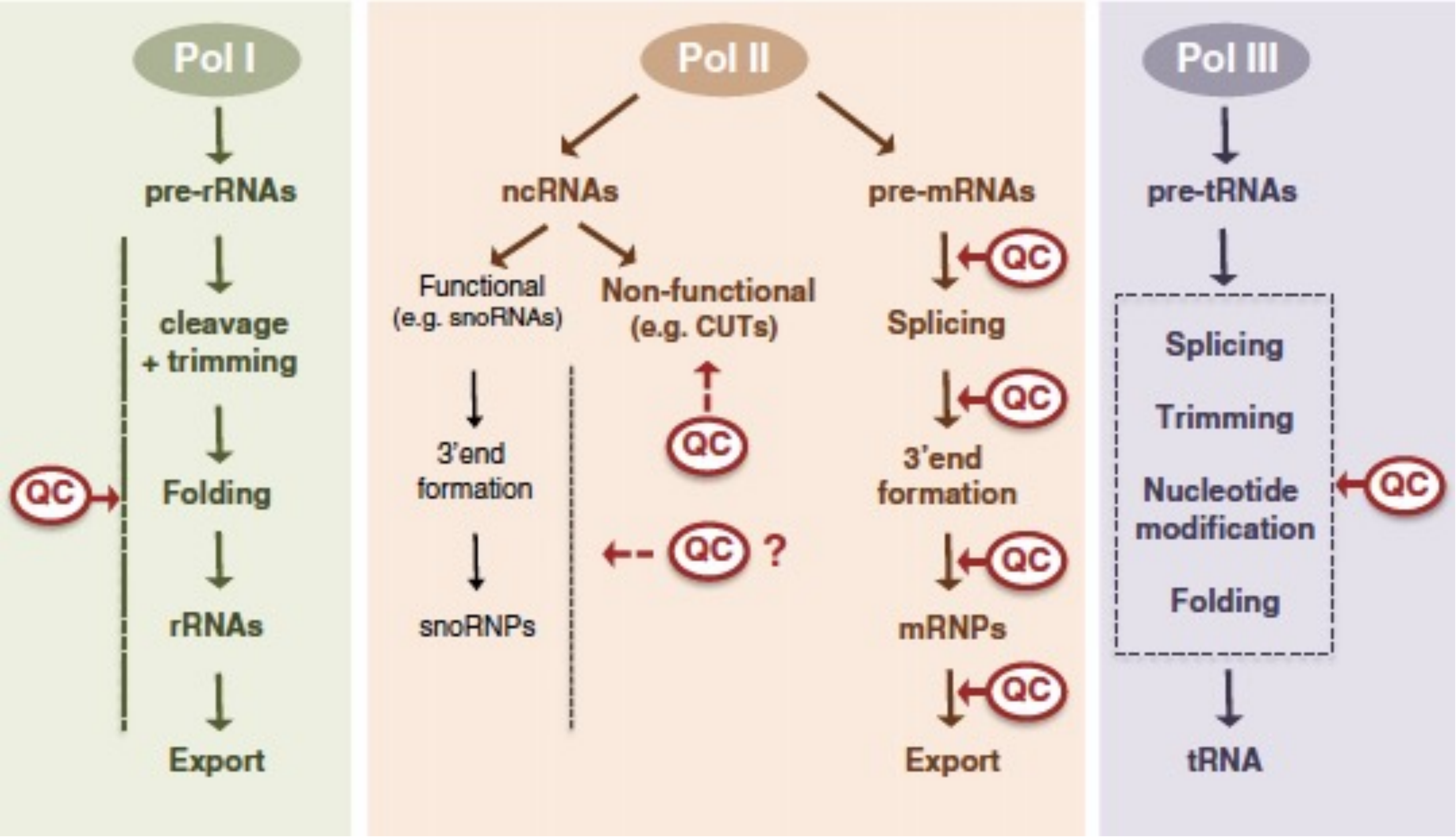
cyto Pol III functions in innate immunity
(human cells)





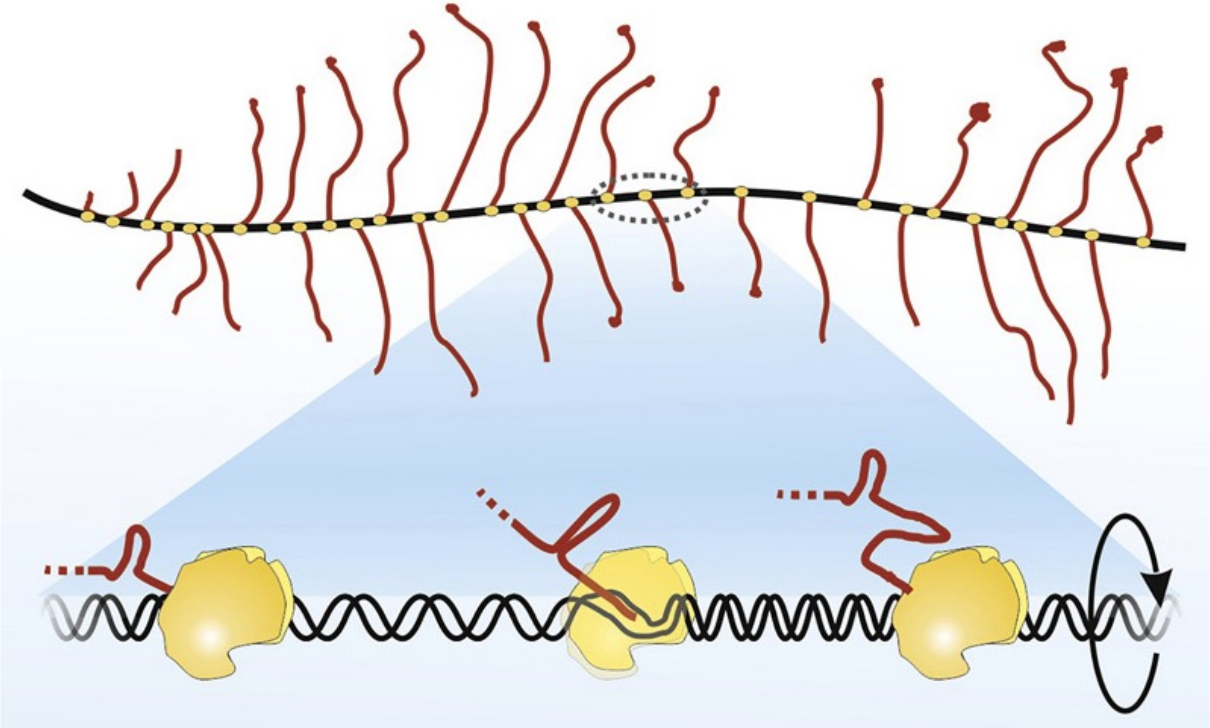
Girbig et al, 2022, Nat Rev Mol Cell Biol

RNA polymerases - transcription



Porrua and Libri, BBA, 2013

Nascent transcripts



Nascent transcript = during formation, newly formed, still bound by polymerase

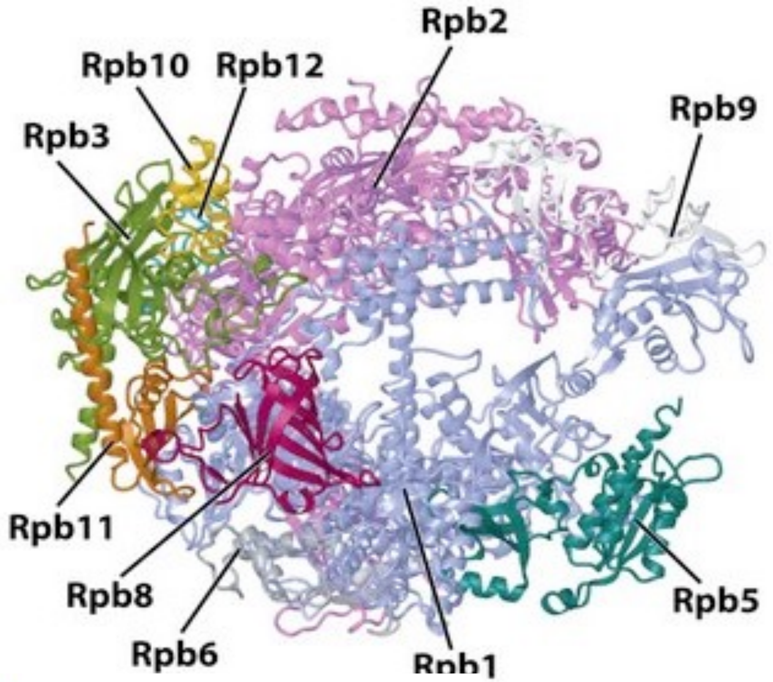
- nascent RNA couple RNA processing with transcription elongation and chromatin modification
- nascent RNA modulate binding of proteins to regulatory elements (chromatin)
- regulatory effects of nascent transcripts can be enhanced by gene looping
- high concentrations of nascent RNAs can initiate formation of nuclear bodies
- the role may be conferred by nascent transcription (activity) and not the transcript itself

Turowski et al., Mol Cell, 2020

Polymerase II (Pol II)

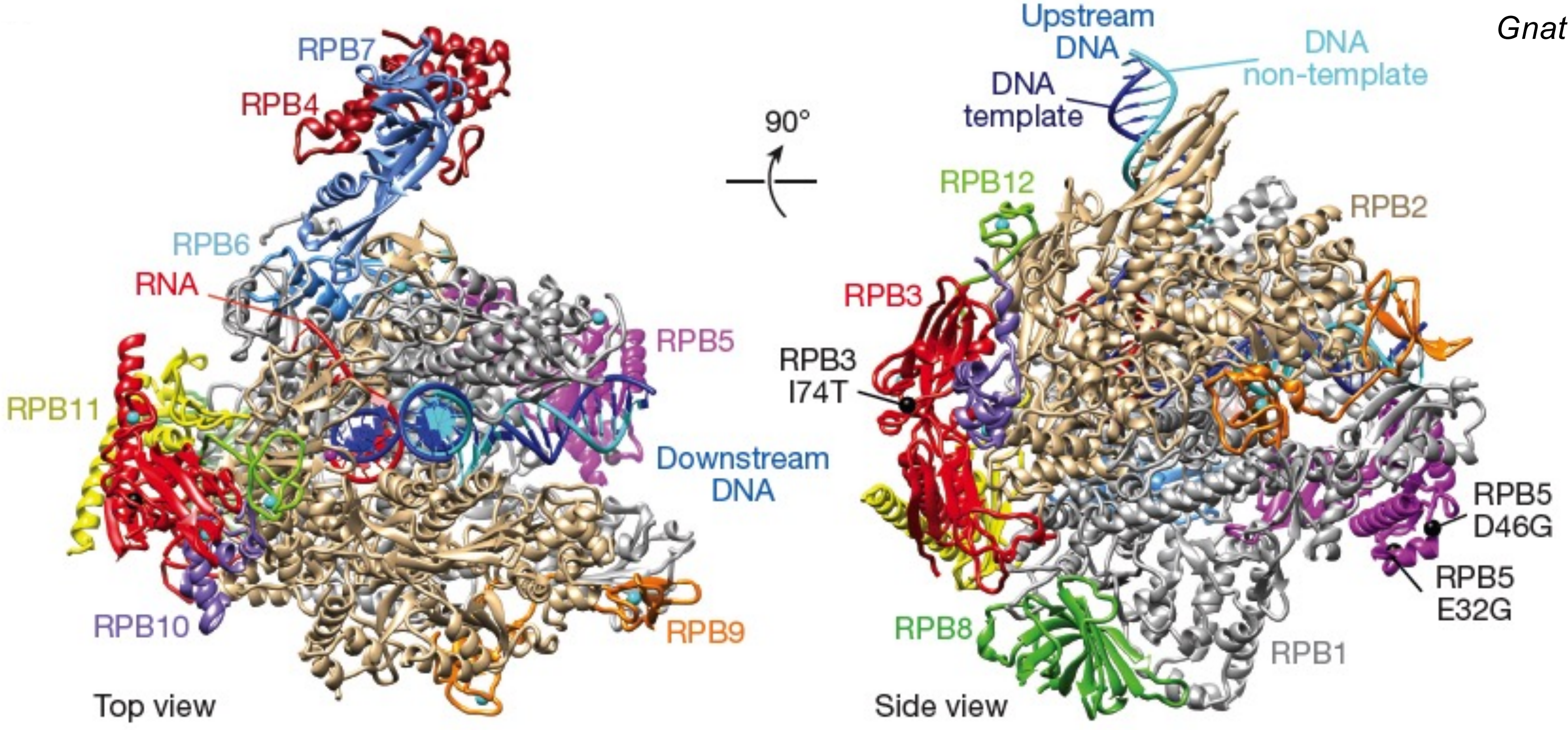
Yeast Pol II

- 12 subunits
- core by specific **Rpb1-3**, **9** and **11**
- **Rpb5-6, 8, 10** and **12** - shared by Pol I-III
- specific subcomplex **Rpb4/7** not essential
- associated factors RAP74, RAP30 (TFIIF)



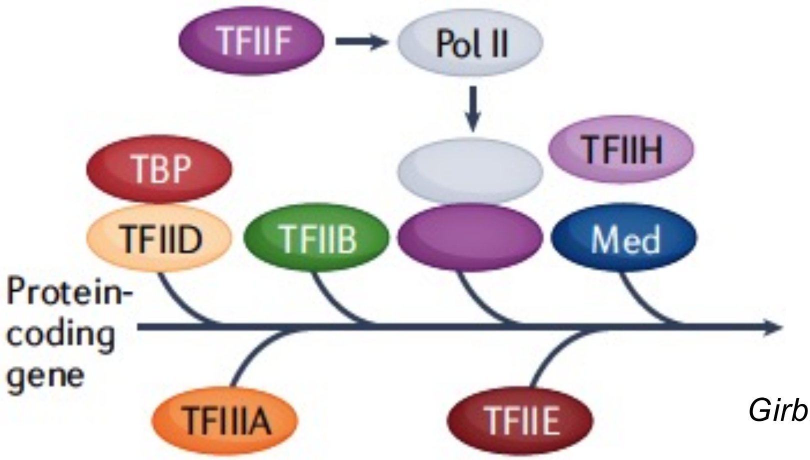
Gnatt et al, Science, 2001

Mammalian Pol II

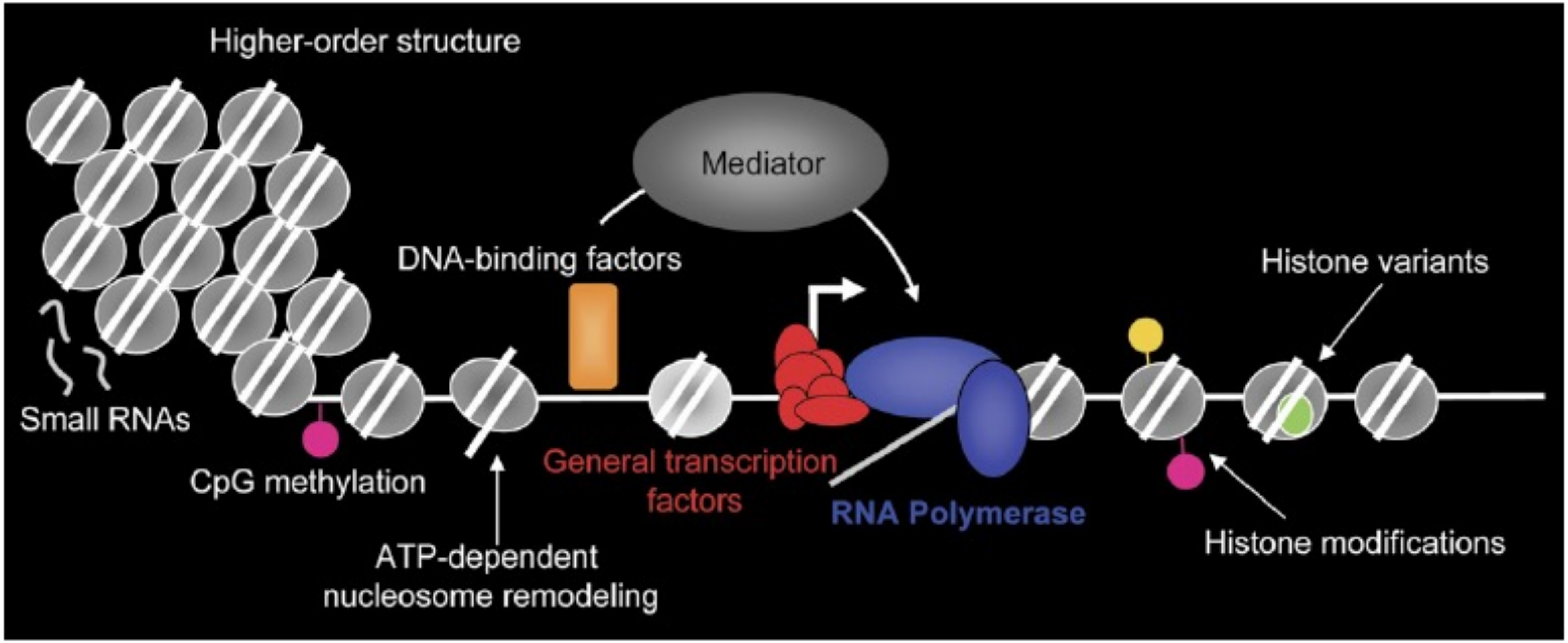


Berneckyet al, 2016, Nature

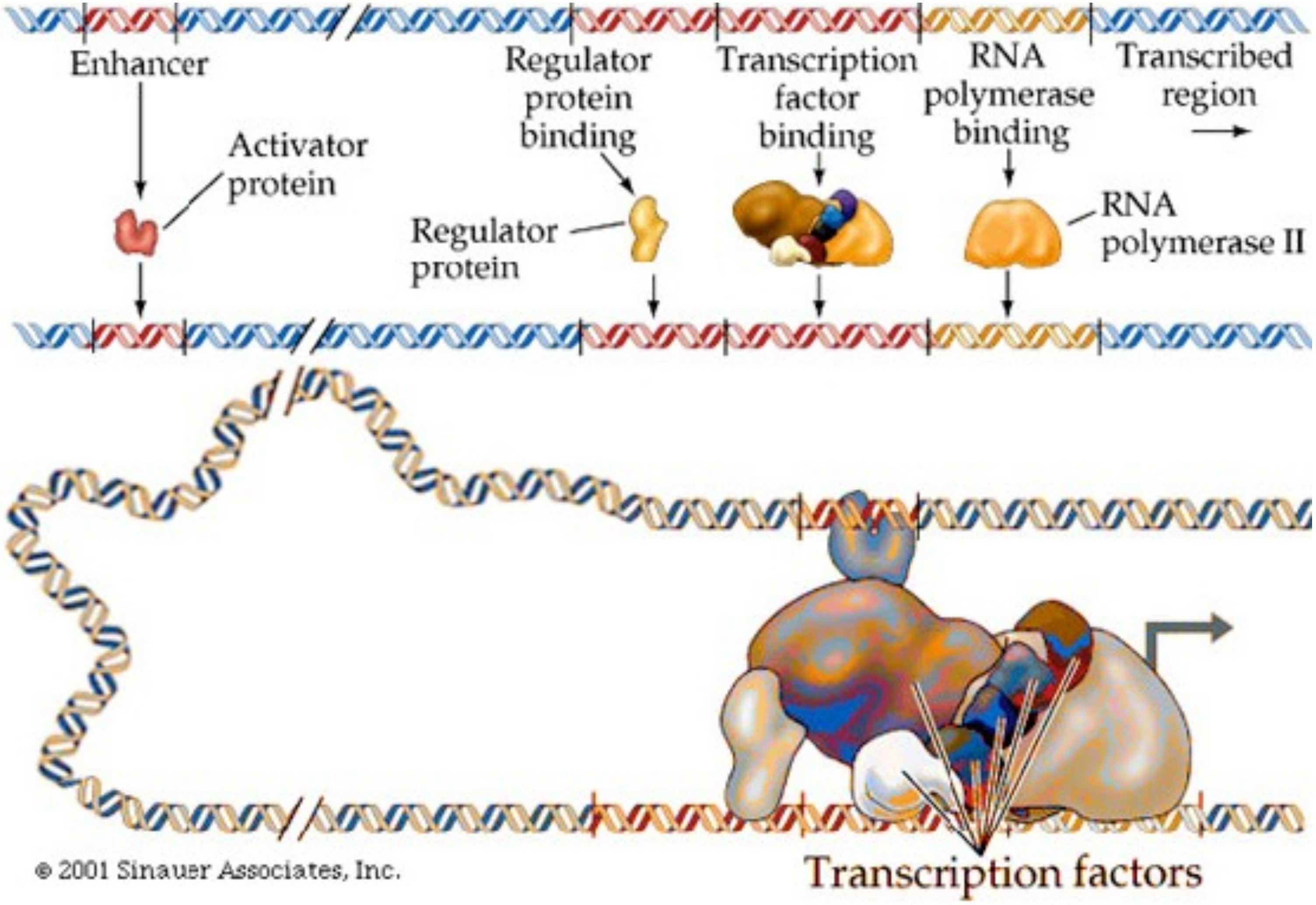
Pol II factors



Girbig et al, 2022, Nat Rev Mol Cell Biol



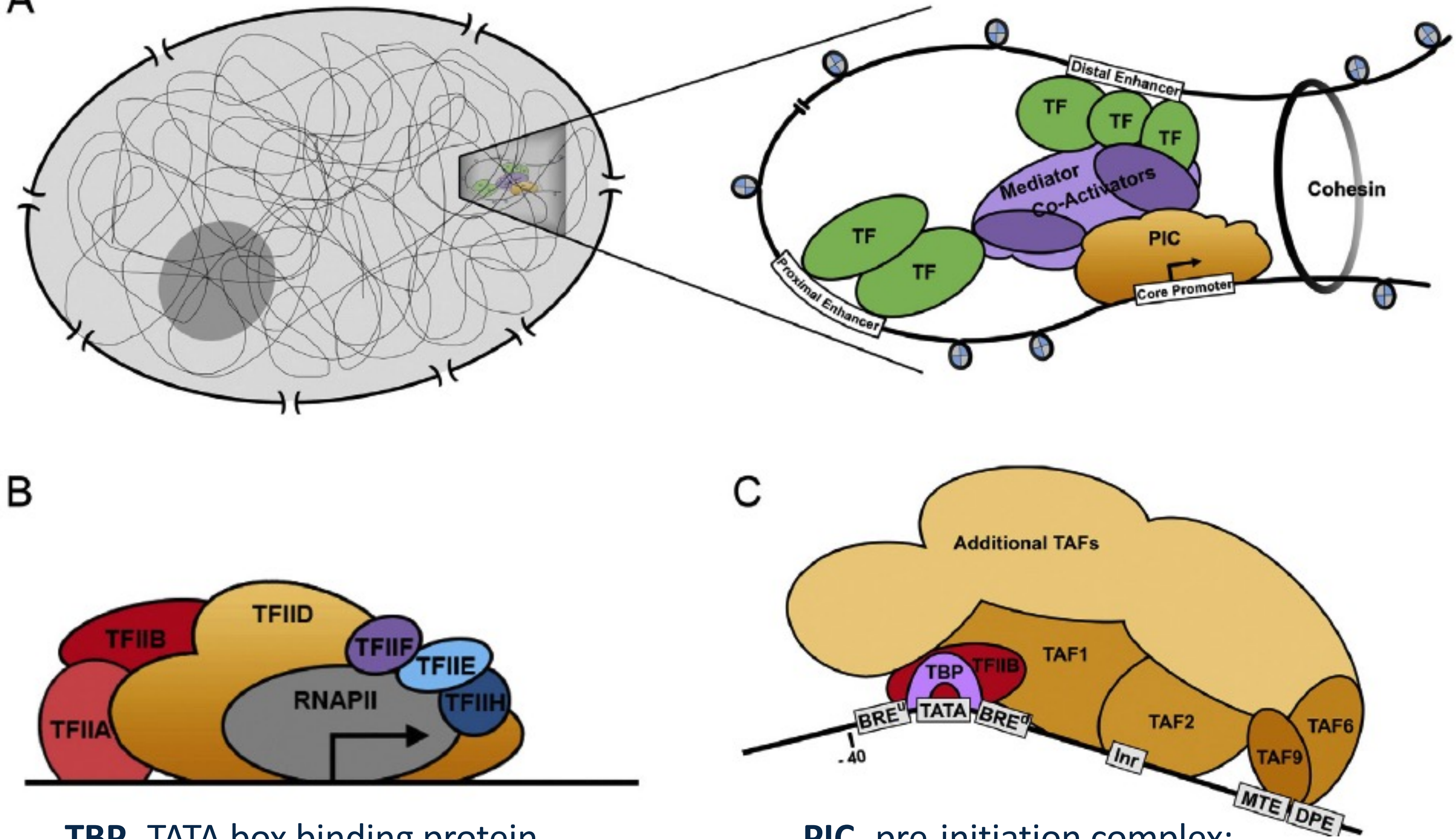
Pol II factors



© 2001 Sinauer Associates, Inc.

http://cats.med.uvm.edu/cats_teachingmod/microbiology/cats_mmg_courses_new.htm

Pol II factors



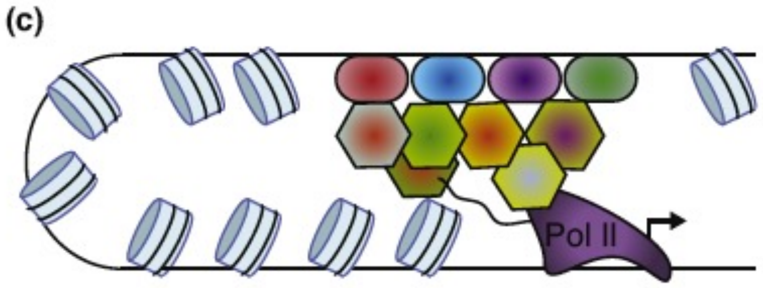
TBP- TATA box binding protein
TAF- TBP associated factors
TF- transcription factors
TSS-transcription start site

PIC- pre-initiation complex:
TFIID, TBP, TAFs (13)

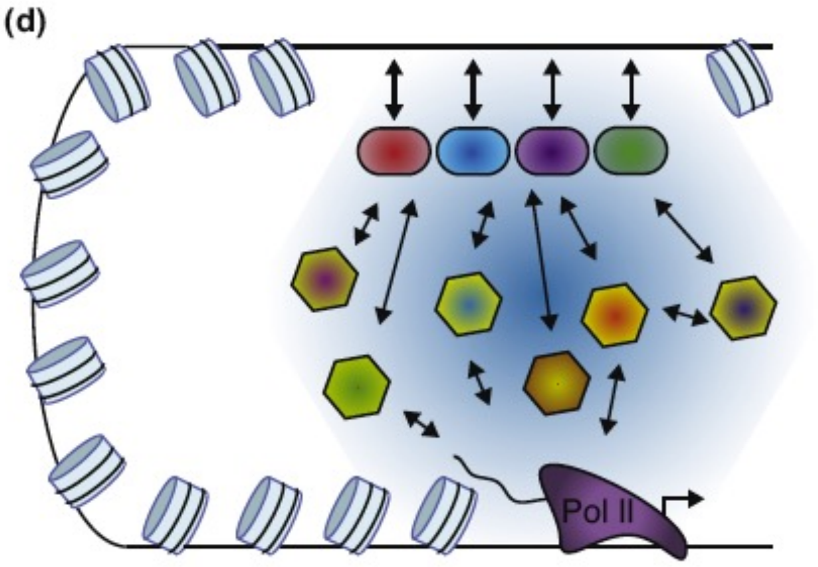
Danino et al. BBA, 2015

Pol II – network of factors

Cooperation of TF and co-factors via chromatin modification



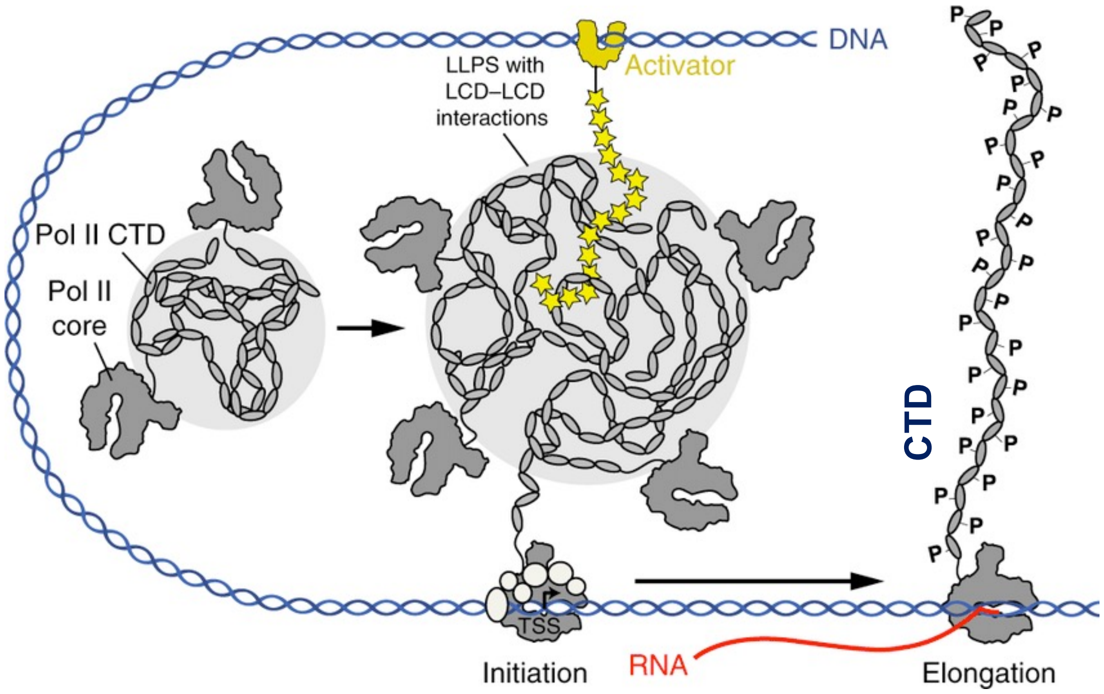
Old fixed model



New model - more flexible and dynamic (also transient) interactions of TFs and co-factors

Reiter et al, 2017, Curr Op Genet Dev

Pol II in the cell



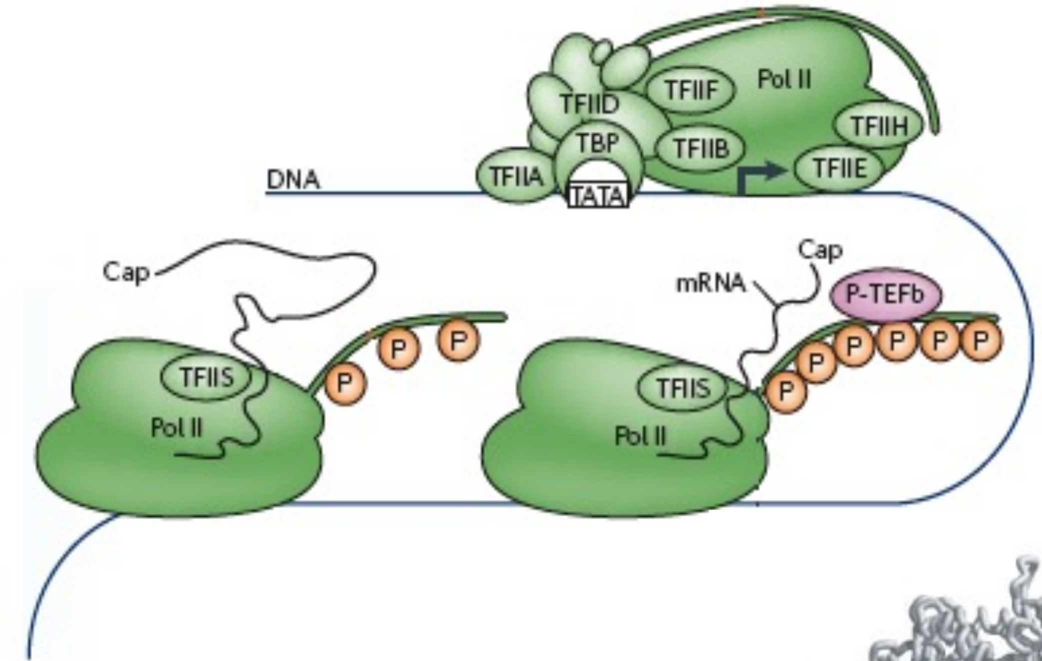
CTD-driven phase separation

Activators recruit/nucleate Pol II hubs near promoters. Initiation-coupled CTD phosphorylation removes individual Pol II enzymes for transcription elongation.

LLPS, droplets Liquid-liquid phase separation
Transcriptional condensates are formed by phase-separation self-assembly driven by IDR (Intrinsically Disordered Region)- containing proteins (e.g. CTD in Pol II)

Lesne et al.,2019 Genes; Boehning et al, 2018, Nat Struct Mol Biol

Pol II C-terminal domain (CTD)



Goodrich and Kugel, Nat. Rev. Mol. Biol., 2006

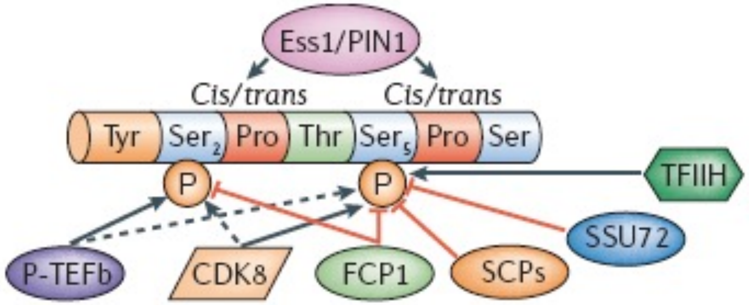
Changes of the phosho status of Pol II CTD are important for the recruitment of pre-mRNA processing machineries:

Capping Enzymes (CE)

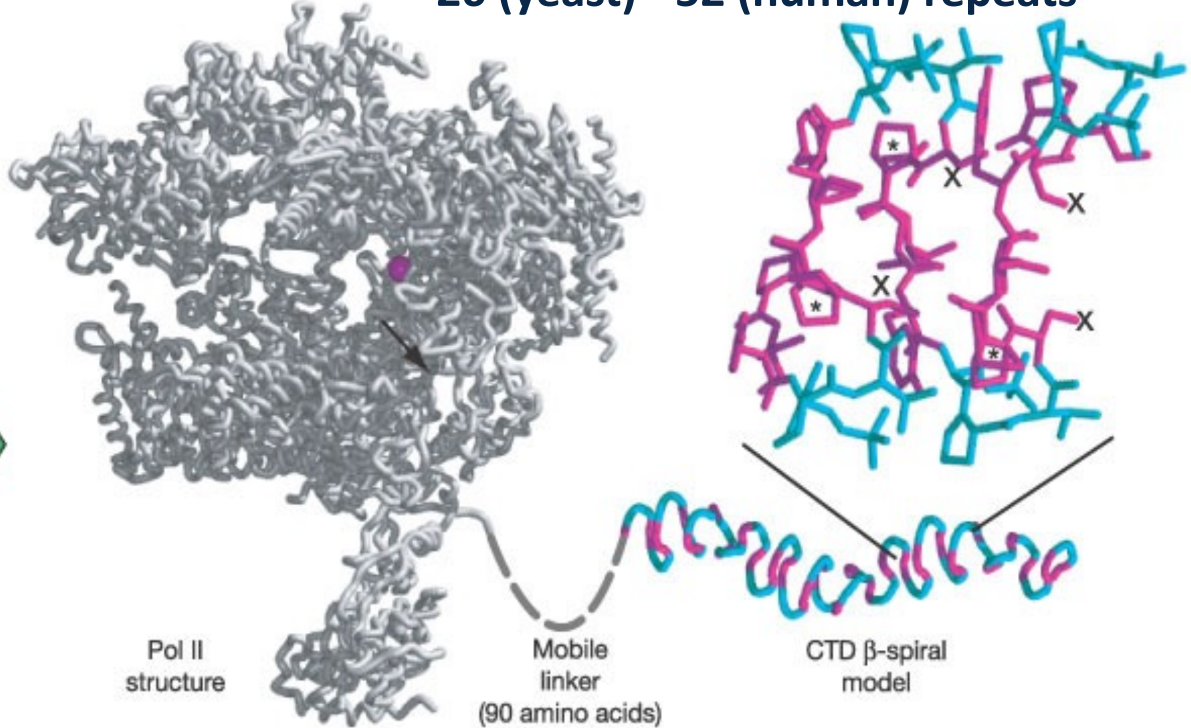
Spliceosome

3’ end formation complex (CPA)

26 (yeast) - 52 (human) repeats



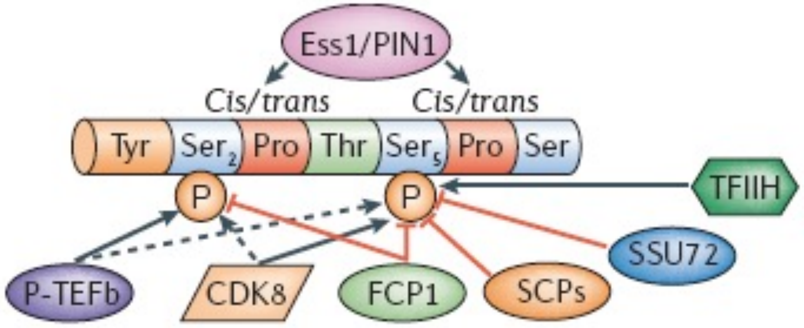
Saunders et al, 2006, Nat.Rev.Mol.Cel.Biol



Meinhart and Cramer, 2004

$Tyr_1Ser_2Pro_3Thr_4Ser_5Pro_6Ser_7$

CTD code



Ser5-P

Cyclin-dependent kinase-7 (**CDK7**) of **TFIIH** and **CDK8**

Phosphatases **SSU72**, **FCP1**

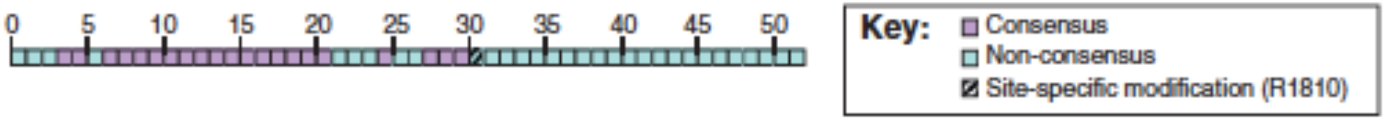
SCPs small CTD phosphatases

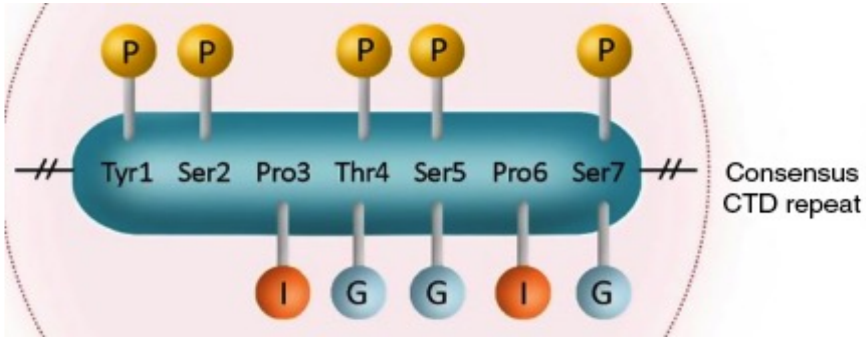
Ser2-P

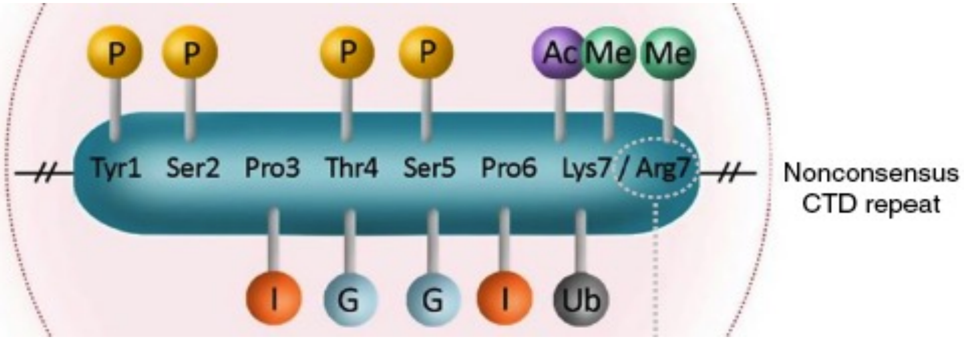
Kinases **CDK8** and **CDK9** of **P-TEFb**

Phosphatase **FCP1**

mammalian CTD repeats

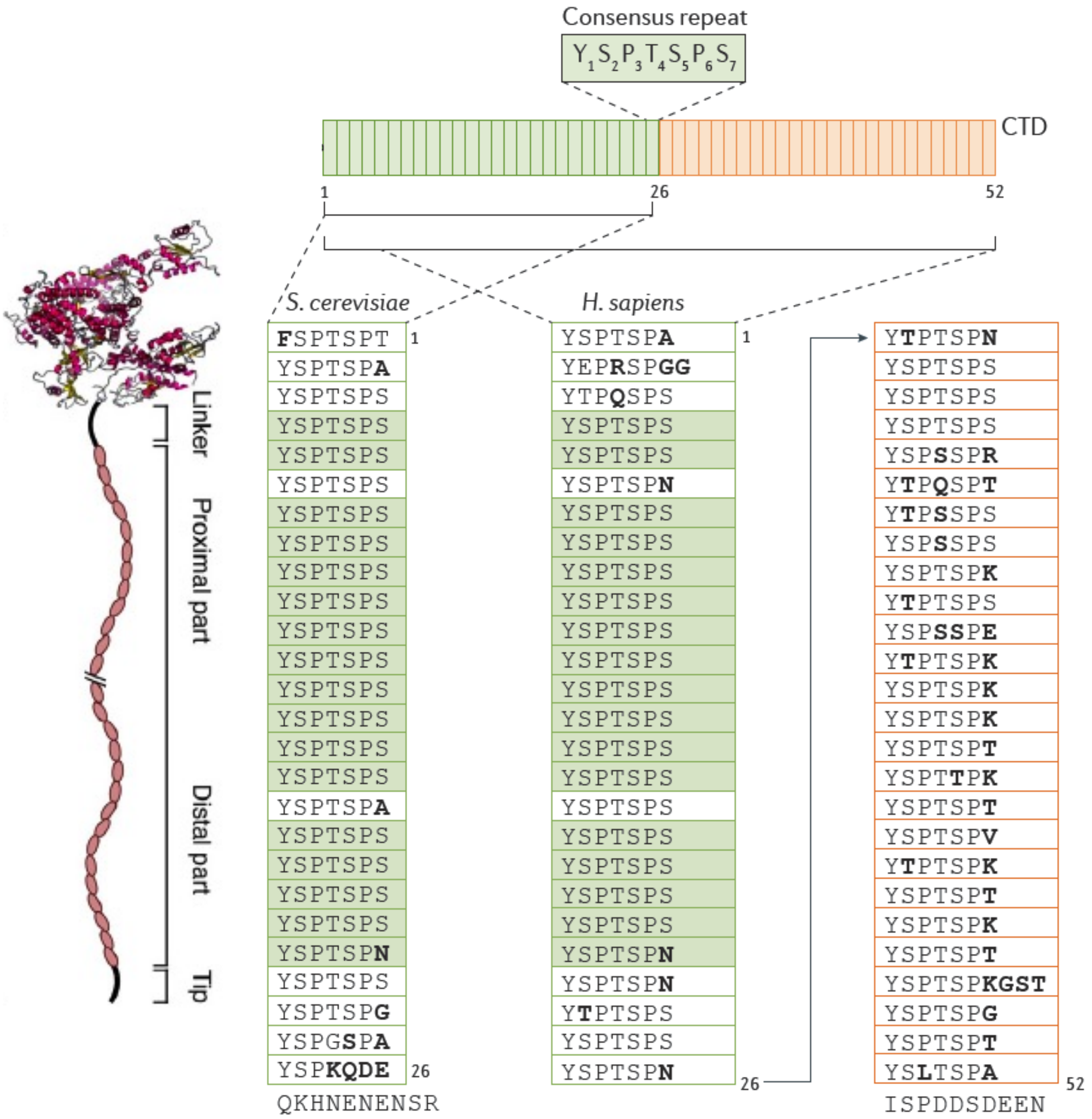






Saunders et al, 2006, Nat Rev Mol Cel Biol; Zaborowska et al, 2015, Nat Str Mol Biol

CTD code



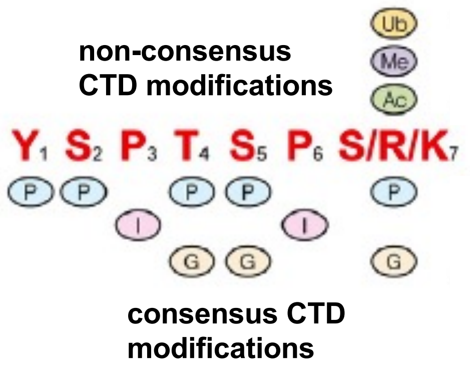
S. cerevisiae

FSPTSPT 1
YSPTSP**A**
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSP**A**
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSP**N**
YSPTSPS
YSPTSP**G**
YSPG**S**P**A**
YSP**KQDE** 26
QKHNENENSR

H. sapiens

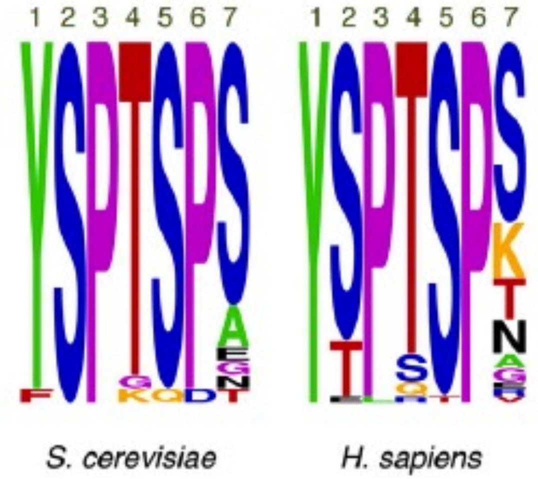
YSPTSP**A** 1
YEP**R**SP**GG**
YTP**Q**SPS
YSPTSPS
YSPTSPS
YSPTSP**N**
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSPS
YSPTSP**N**
YSPTSP**N**
Y**T**PTSPS
YSPTSPS
YSPTSP**N** 26

Y**T**PTSP**N**
YSPTSPS
YSPTSPS
YSPTSPS
YSP**S**SP**R**
Y**T**P**Q**SP**T**
Y**T**P**S**SPS
YSP**S**SPS
YSPTSP**K**
Y**T**PTSPS
YSP**SS**P**E**
Y**T**PTSP**K**
YSPTSP**K**
YSPTSP**K**
YSPTSP**T**
YSPT**T**P**K**
YSPTSP**T**
YSPTSP**V**
Y**T**PTSP**K**
YSPTSP**T**
YSPTSP**K**
YSPTSP**T**
YSPTSP**KGST**
YSPTSP**G**
YSPTSP**T**
YS**L**TSP**A** 52
ISPDDSDEEN



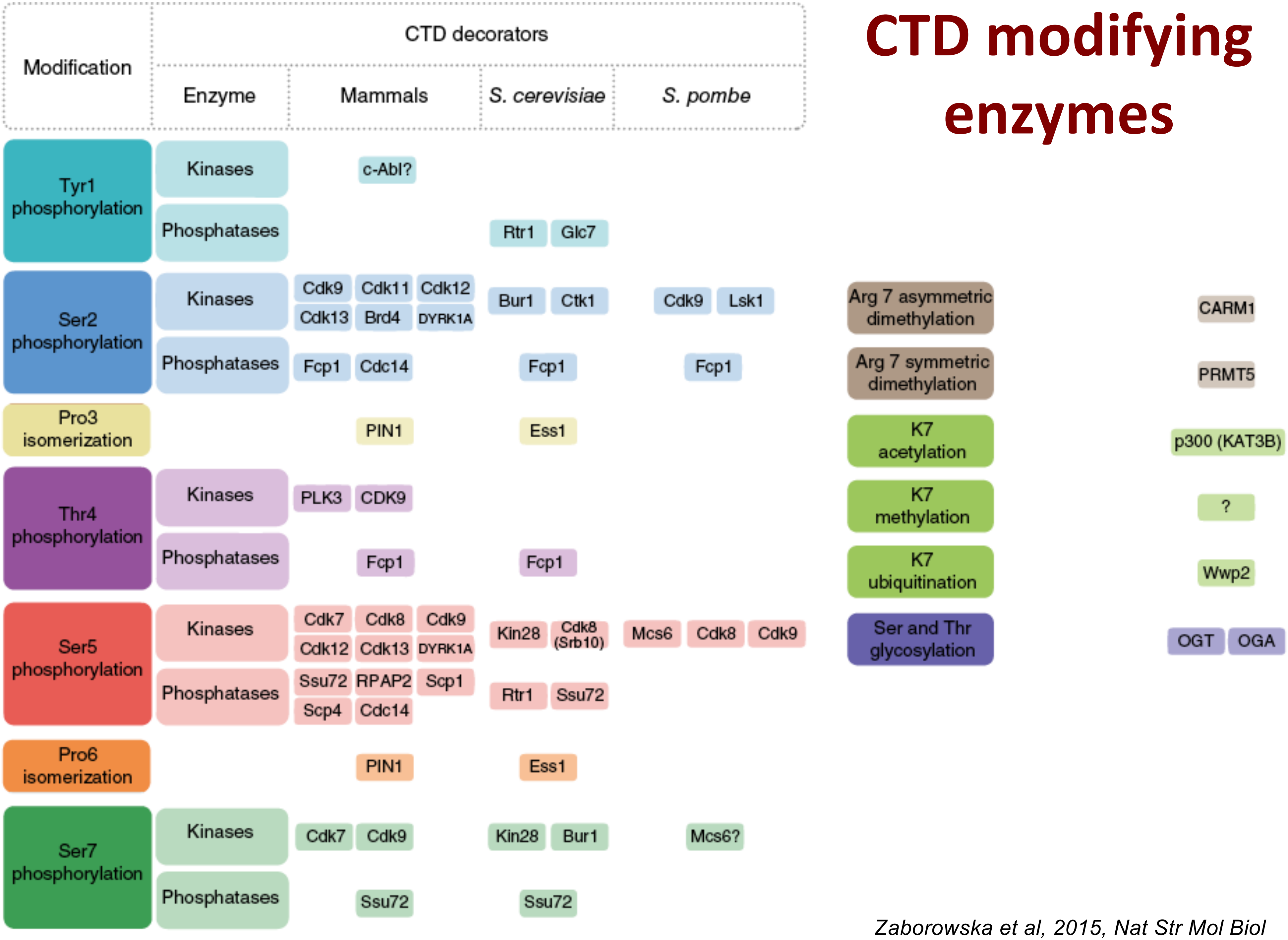
non-consensus CTD modifications

consensus CTD modifications



Srivastava and Ahn, Biotechnol Adv 2015; Harlan and Churchman, Nat Rev Mol Cell Biol, 2017

CTD modifying enzymes

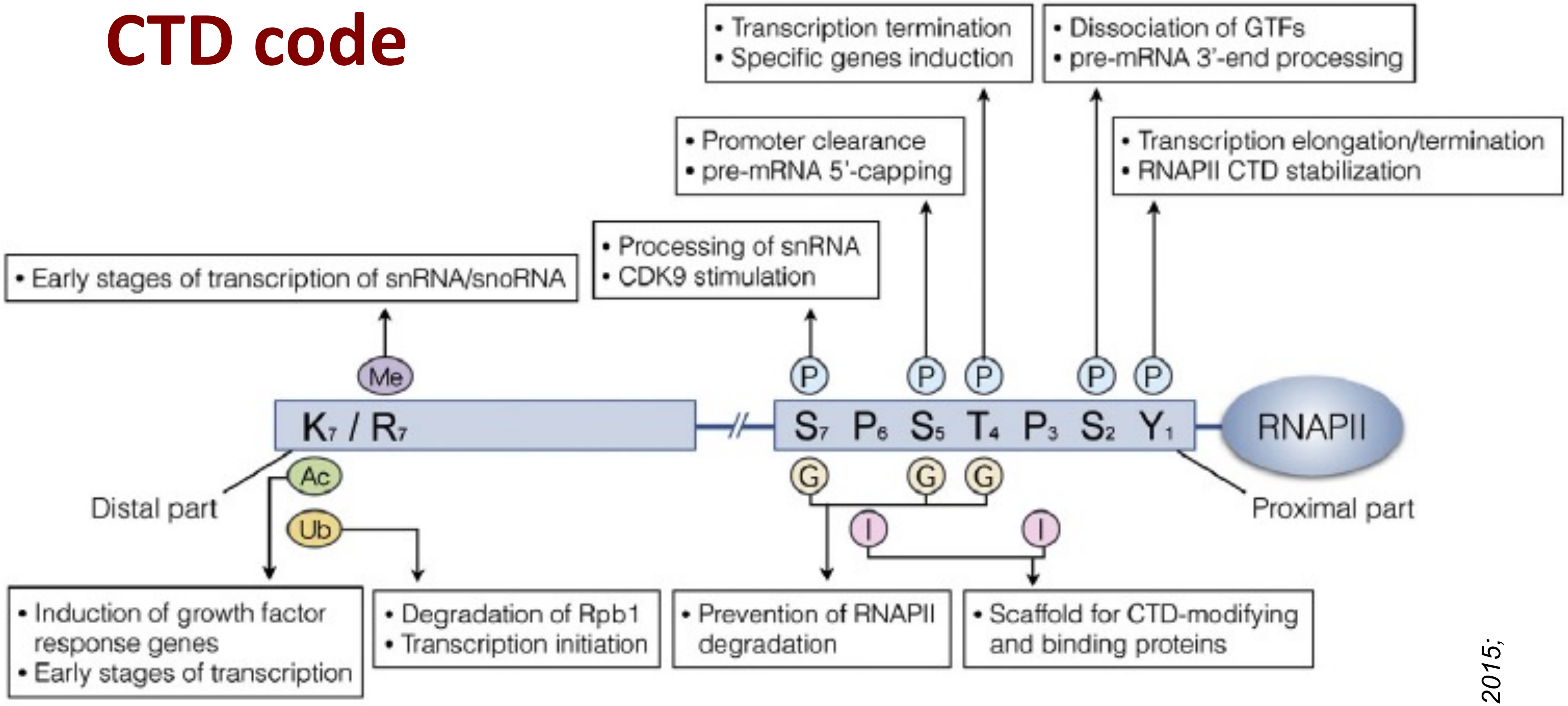


Modification	CTD decorators			
	Enzyme	Mammals	*S. cerevisiae*	*S. pombe*
Tyr1 phosphorylation	Kinases	c-Abl?		
	Phosphatases		Rtr1 Glc7	
Ser2 phosphorylation	Kinases	Cdk9 Cdk11 Cdk12 Cdk13 Brd4 DYRK1A	Bur1 Ctk1	Cdk9 Lsk1
	Phosphatases	Fcp1 Cdc14	Fcp1	Fcp1
Pro3 isomerization		PIN1	Ess1	
Thr4 phosphorylation	Kinases	PLK3 CDK9		
	Phosphatases	Fcp1	Fcp1	
Ser5 phosphorylation	Kinases	Cdk7 Cdk8 Cdk9 Cdk12 Cdk13 DYRK1A	Kin28 Cdk8 (Srb10)	Mcs6 Cdk8 Cdk9
	Phosphatases	Ssu72 RPAP2 Scp1 Scp4 Cdc14	Rtr1 Ssu72	
Pro6 isomerization		PIN1	Ess1	
Ser7 phosphorylation	Kinases	Cdk7 Cdk9	Kin28 Bur1	Mcs6?
	Phosphatases	Ssu72	Ssu72	

Arg 7 asymmetric dimethylation	CARM1
Arg 7 symmetric dimethylation	PRMT5
K7 acetylation	p300 (KAT3B)
K7 methylation	?
K7 ubiquitination	Wwp2
Ser and Thr glycosylation	OGT OGA

Zaborowska et al, 2015, Nat Str Mol Biol

CTD code



- Transcription termination
- Specific genes induction

- Dissociation of GTFs
- pre-mRNA 3'-end processing

- Promoter clearance
- pre-mRNA 5'-capping

- Transcription elongation/termination
- RNAPII CTD stabilization

- Early stages of transcription of snRNA/snoRNA

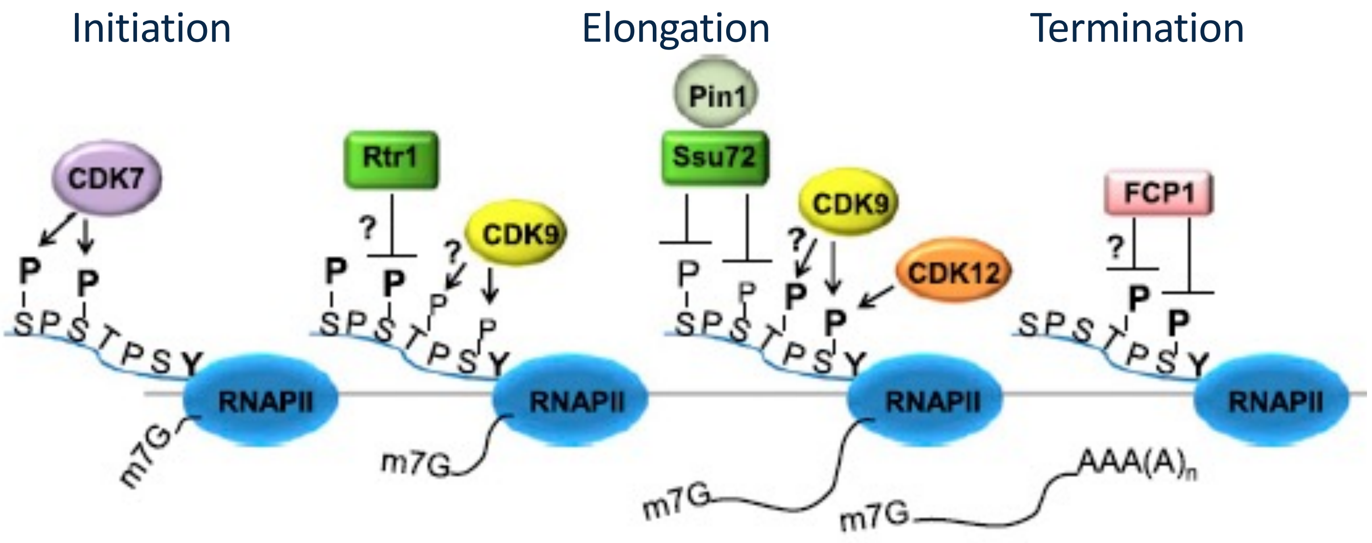
- Processing of snRNA
- CDK9 stimulation

- Induction of growth factor response genes
- Early stages of transcription

- Degradation of Rpb1
- Transcription initiation

- Prevention of RNAPII degradation

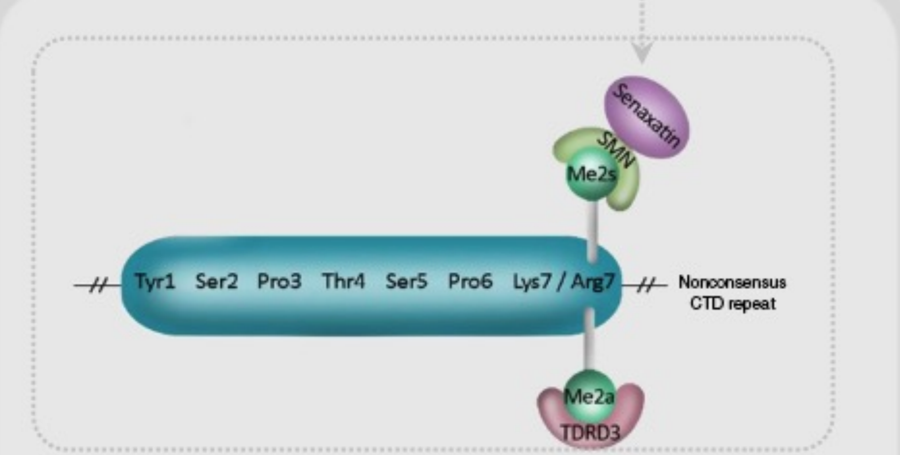
- Scaffold for CTD-modifying and binding proteins



Srivastava and Ahn, Biotechnol Adv, 2015;
Hsin and Manley, Genes Dev, 2012

CTD code

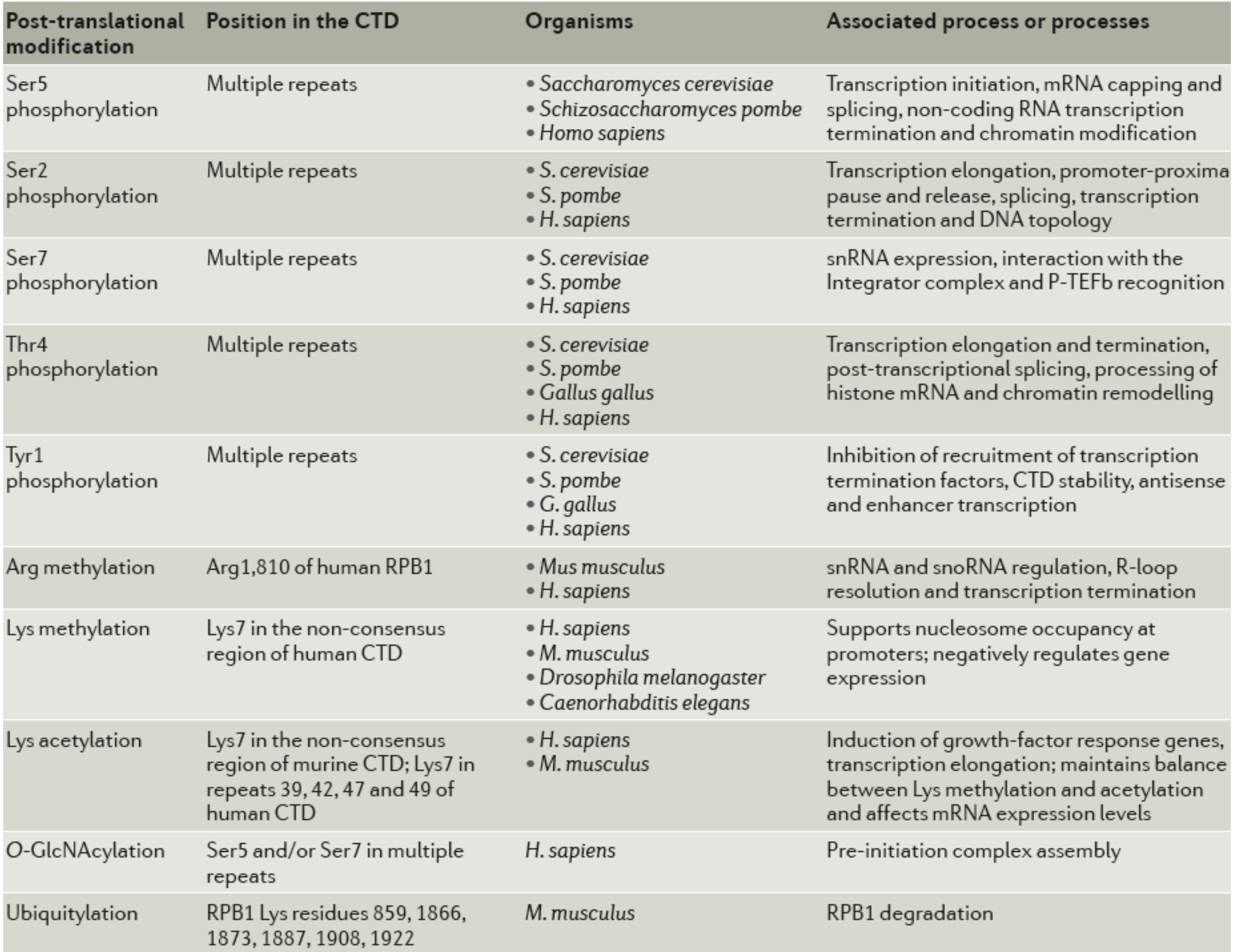
Tyr1	P	• Aids in recruitment of Spt6 (*S. cerevisiae)* • Impairs the binding of termination factors to the CTD (*S. cerevisiae)* • Important for pol II stability (chicken DT40 cells) • Associated with transcription of antisense promoters and active enhancers (human)
Ser2	P	• Aids in recruitment of splicing and polyadenylation factors • Regulates sexual differentiation (*S. pombe*) • Aids in recruitment of Set2 (H3K36me3) and Spt6 • Enhances pre-snRNA 3′-end formation (human) • Required for chromatin remodeling of genes involved in phosphate and galactose metabolism (*S. cerevisiae*)
Thr4	P	• Enhances 3′-end processing of histone transcripts (chicken DT40 cells) • Associated with efficient transcription elongation (human) • Involved in M-phase progression • Aids in termination of transcription (*S. cerevisiae*)
Ser5	P	• Marks transcription initiation and splice sites (human and yeast) • Aids in promoter escape • Aids in recruitment of the RNA 5′-end-capping machinery • Aids in recruitment of Set1 (H3K4me3) • Aids in recruitment of the Nrd1-Nab3-Sen1 termination complex (*S. cerevisiae*) • Aids in spliceosome recruitment and regulates splicing (*S. cerevisiae*)
Ser7	P	• Aids in recruitment of Integrator and RPAP2 to snRNA genes (human) • Regulates the phosphate-response pathway (*S. pombe*) • May facilitate elongation and suppress cryptic transcription (*S. cerevisiae*)
Pro	I	t • Pcf11 binds to repeats with Ser2P and prolines in *trans* c • Ssu72 and Nrd1 bind to repeats with Ser5P and prolines in *cis*
CTD	G	• May regulate initiation of transcription



Arg7	Me	• Mutation of Arg7 (R1810) causes misexpression of snRNAs and snoRNAs and defects in transcription termination • Symmetric dimethylation (Me2s) recruits SMN • Asymmetric dimethylation (Me2a) is recognized by TDRD3
Lys7	Ac	• Implicated in initiation and elongation of transcription • Required for transcription activation of c-Fos and Egr • Coevolved with metazoan lineages
Lys7	Me	• Defines the early stages of transcription
Lys7	Ub	• Regulates expression of Rpb1 in embryonic stem cells

Zaborowska et al, 2015, Nat Str Mol Biol

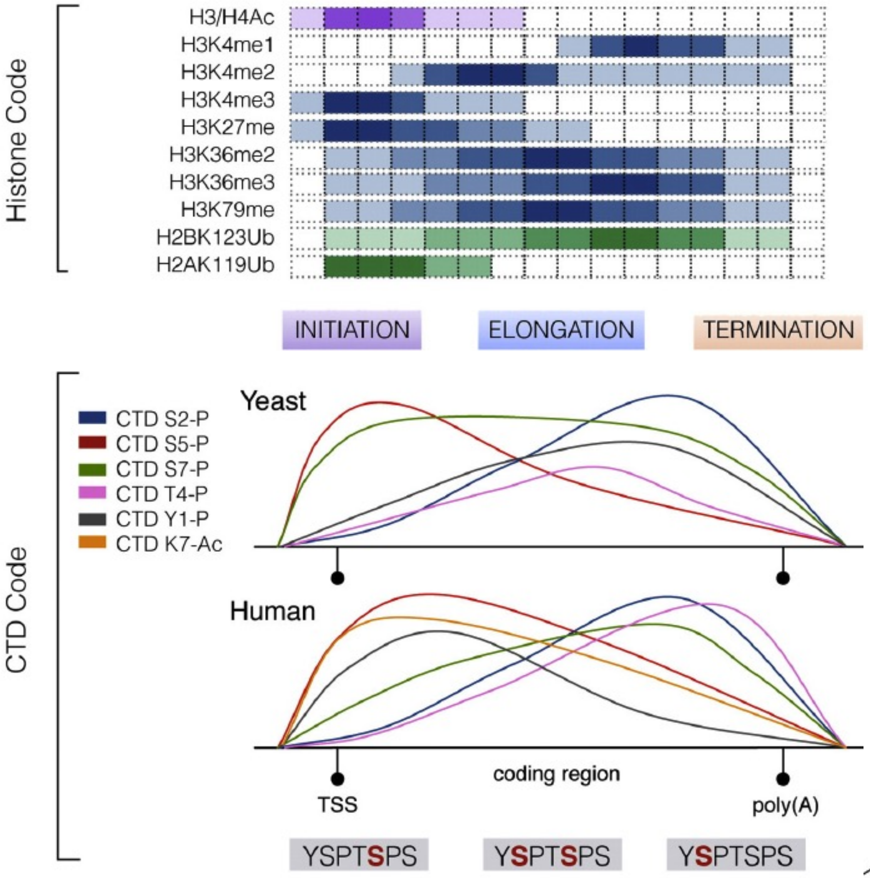
CTD code



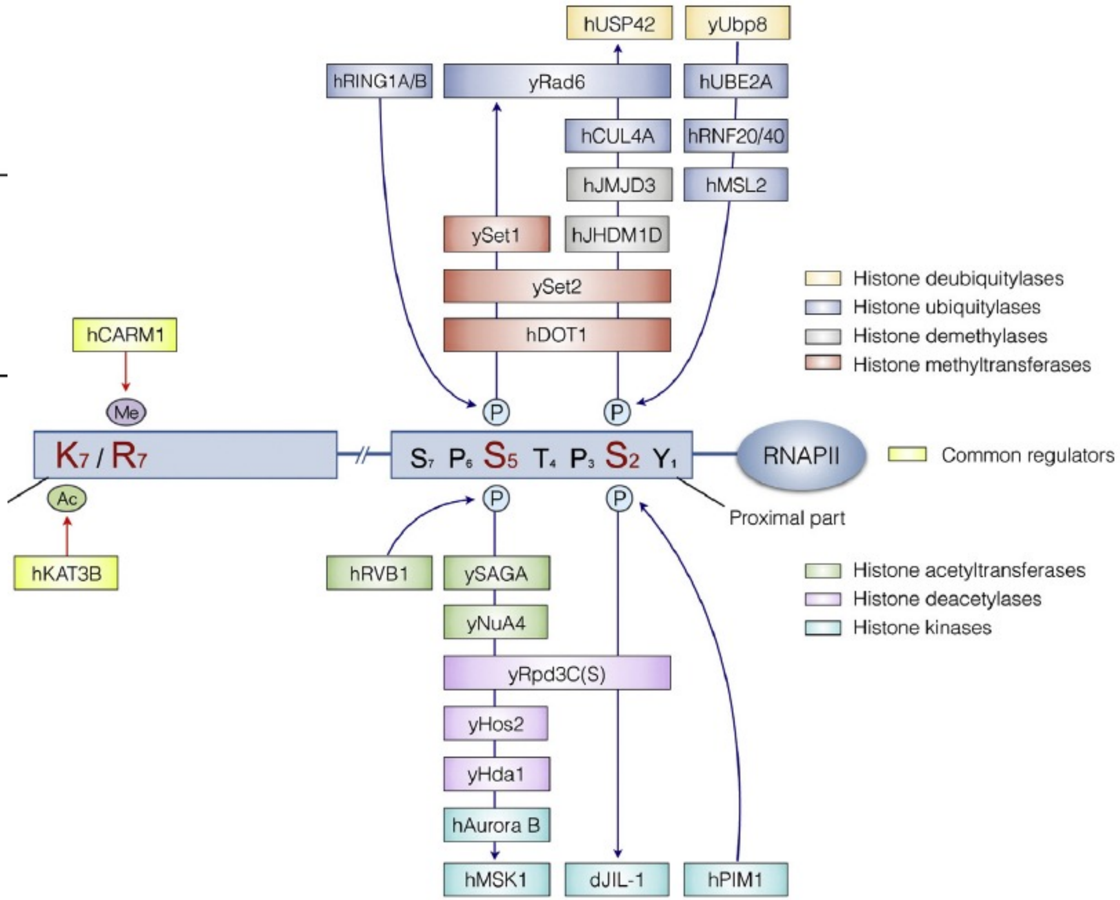
Post-translational modification	Position in the CTD	Organisms	Associated process or processes
Ser5 phosphorylation	Multiple repeats	• *Saccharomyces cerevisiae* • *Schizosaccharomyces pombe* • *Homo sapiens*	Transcription initiation, mRNA capping and splicing, non-coding RNA transcription termination and chromatin modification
Ser2 phosphorylation	Multiple repeats	• *S. cerevisiae* • *S. pombe* • *H. sapiens*	Transcription elongation, promoter-proxima pause and release, splicing, transcription termination and DNA topology
Ser7 phosphorylation	Multiple repeats	• *S. cerevisiae* • *S. pombe* • *H. sapiens*	snRNA expression, interaction with the Integrator complex and P-TEFb recognition
Thr4 phosphorylation	Multiple repeats	• *S. cerevisiae* • *S. pombe* • *Gallus gallus* • *H. sapiens*	Transcription elongation and termination, post-transcriptional splicing, processing of histone mRNA and chromatin remodelling
Tyr1 phosphorylation	Multiple repeats	• *S. cerevisiae* • *S. pombe* • *G. gallus* • *H. sapiens*	Inhibition of recruitment of transcription termination factors, CTD stability, antisense and enhancer transcription
Arg methylation	Arg1,810 of human RPB1	• *Mus musculus* • *H. sapiens*	snRNA and snoRNA regulation, R-loop resolution and transcription termination
Lys methylation	Lys7 in the non-consensus region of human CTD	• *H. sapiens* • *M. musculus* • *Drosophila melanogaster* • *Caenorhabditis elegans*	Supports nucleosome occupancy at promoters; negatively regulates gene expression
Lys acetylation	Lys7 in the non-consensus region of murine CTD; Lys7 in repeats 39, 42, 47 and 49 of human CTD	• *H. sapiens* • *M. musculus*	Induction of growth-factor response genes, transcription elongation; maintains balance between Lys methylation and acetylation and affects mRNA expression levels
O-GlcNAcylation	Ser5 and/or Ser7 in multiple repeats	*H. sapiens*	Pre-initiation complex assembly
Ubiquitylation	RPB1 Lys residues 859, 1866, 1873, 1887, 1908, 1922	*M. musculus*	RPB1 degradation

Harlan and Churchman, Nat Rev Mol Cell Biol, 2017

CTD code versus histone code

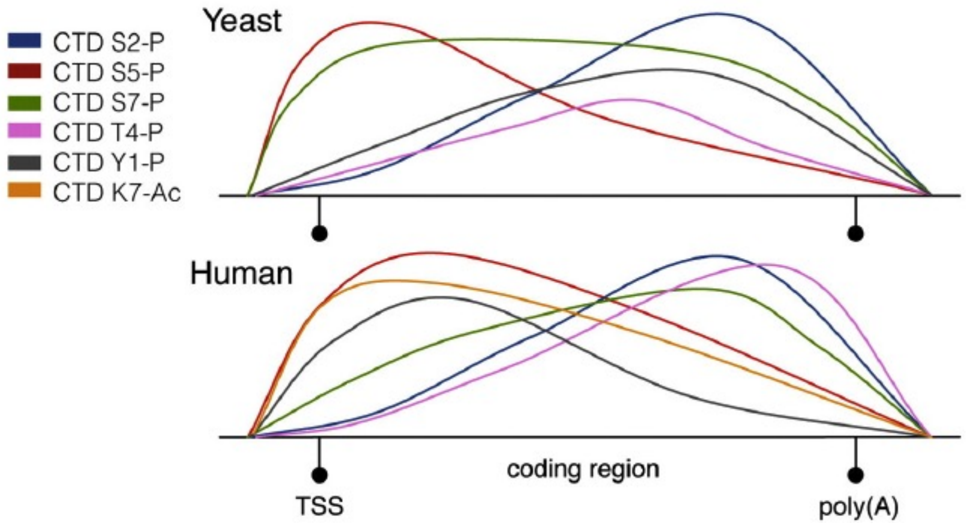


Interactions between CTD modifications and histone-modifying enzymes



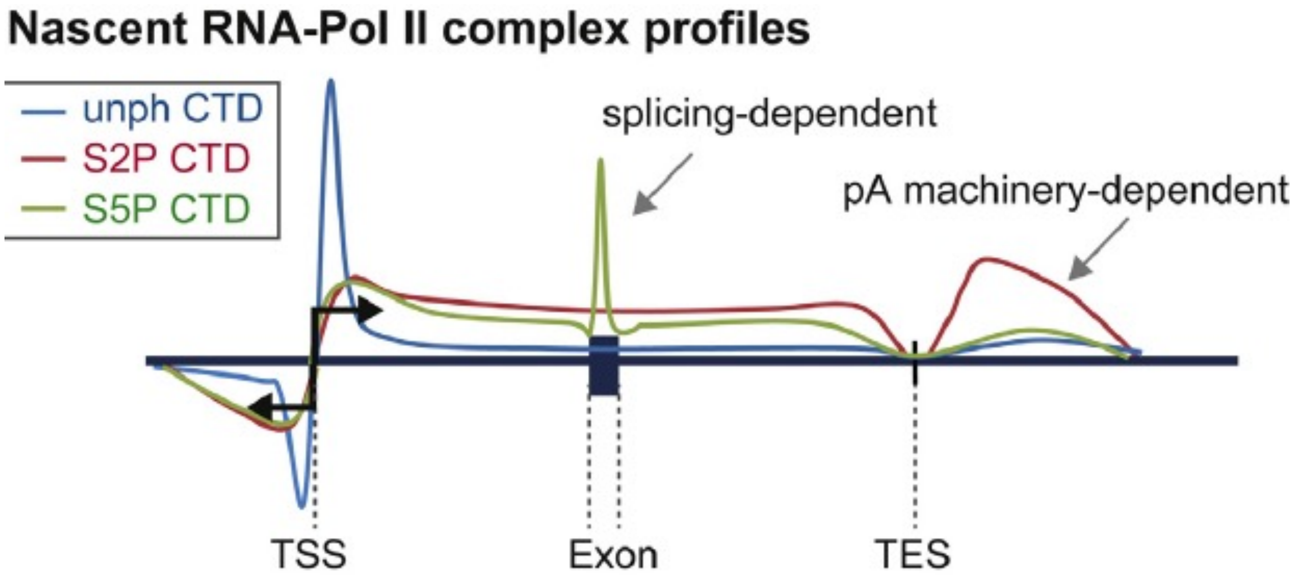
Srivastava and Ahn, Biotechnol Adv 2015

Controversy around the CTD code



Pol II ChIP
(summary)

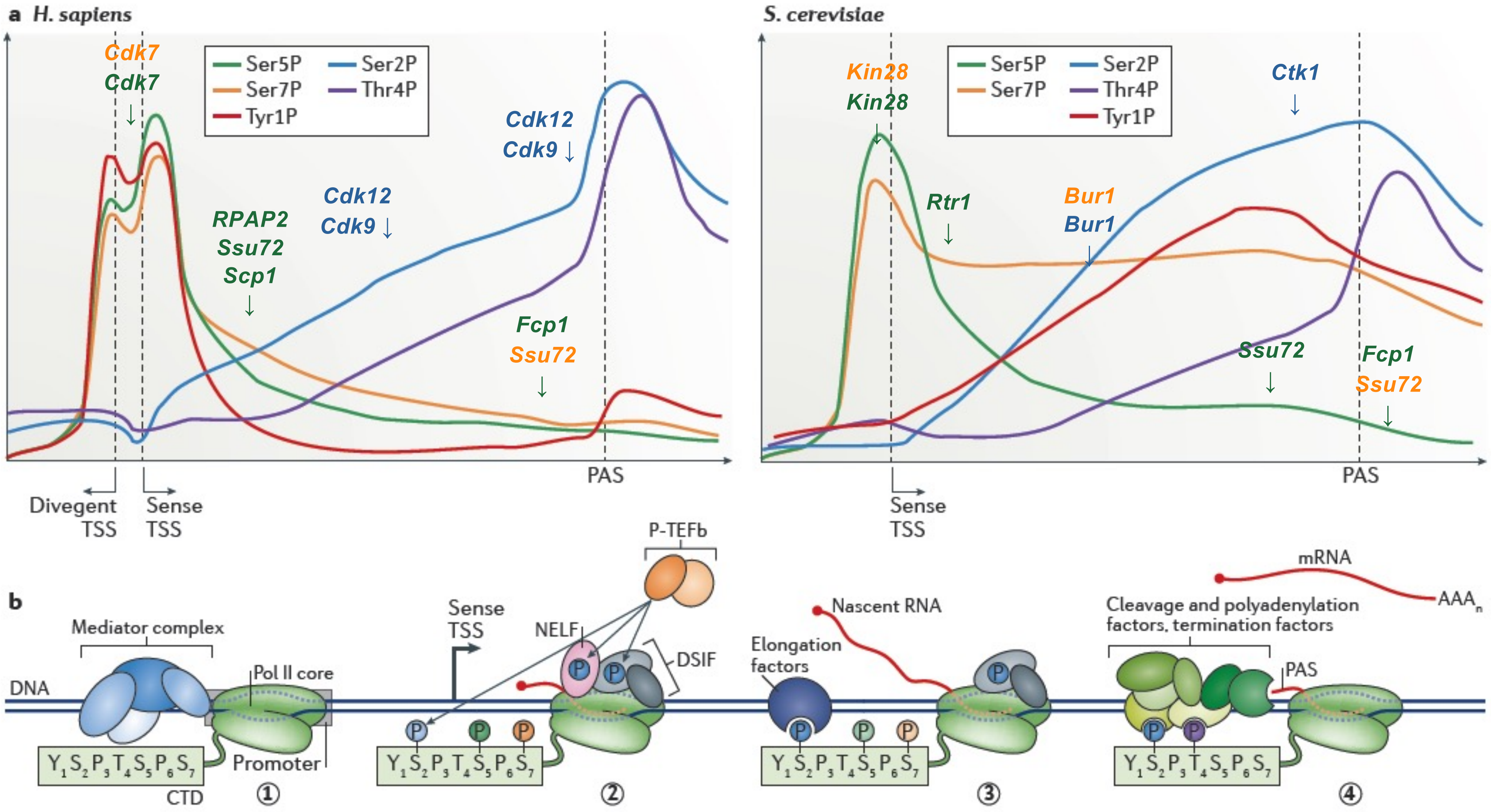
Srivastava and Ahn, Biotechnol Adv 2015



NET-Seq

Noijma et al (Proudfoot lab), Cell 2015

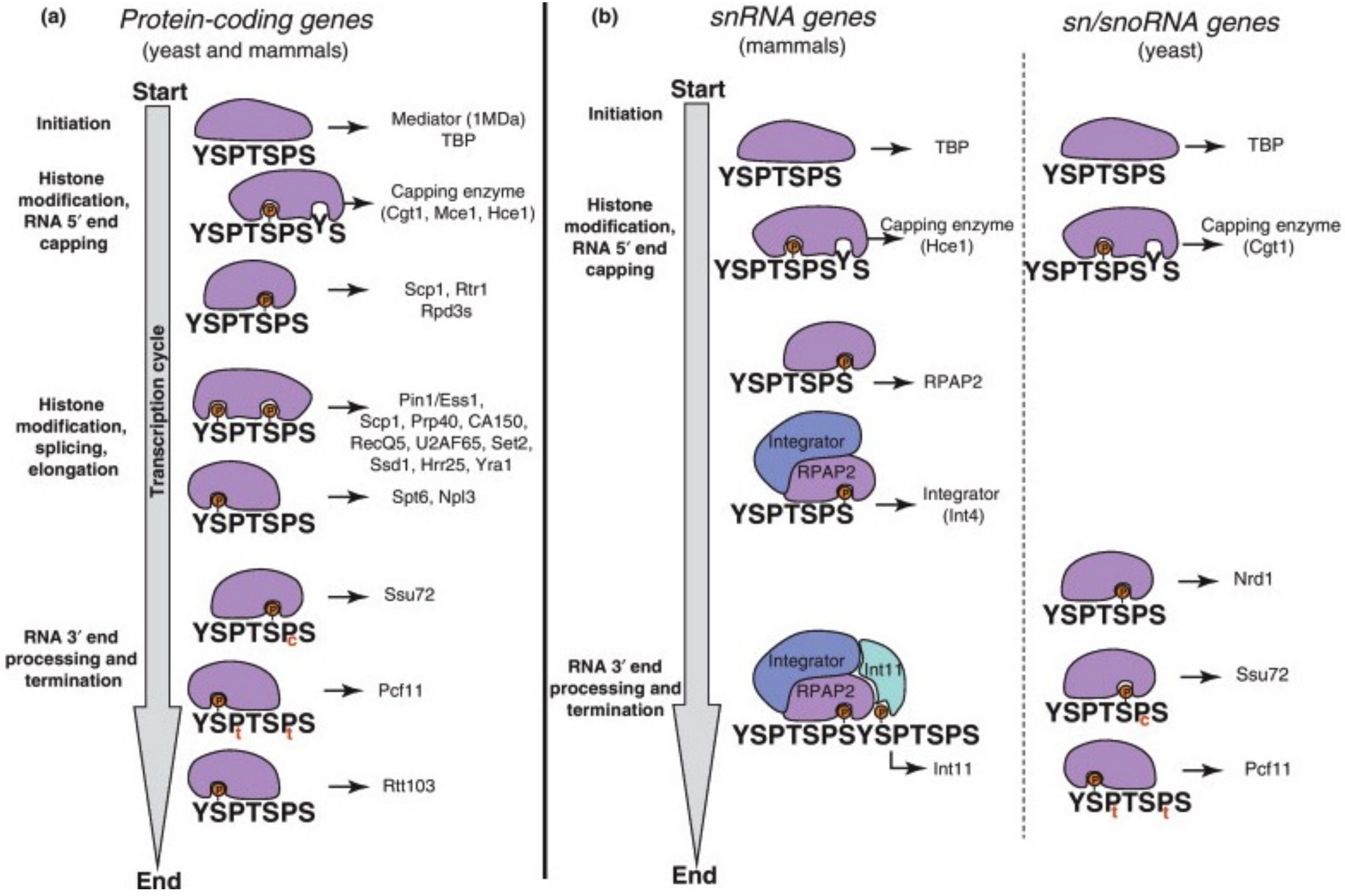
CTD - regulation of transcription



Initiation **Pausing/pausing release** **Elongation** **Termination**

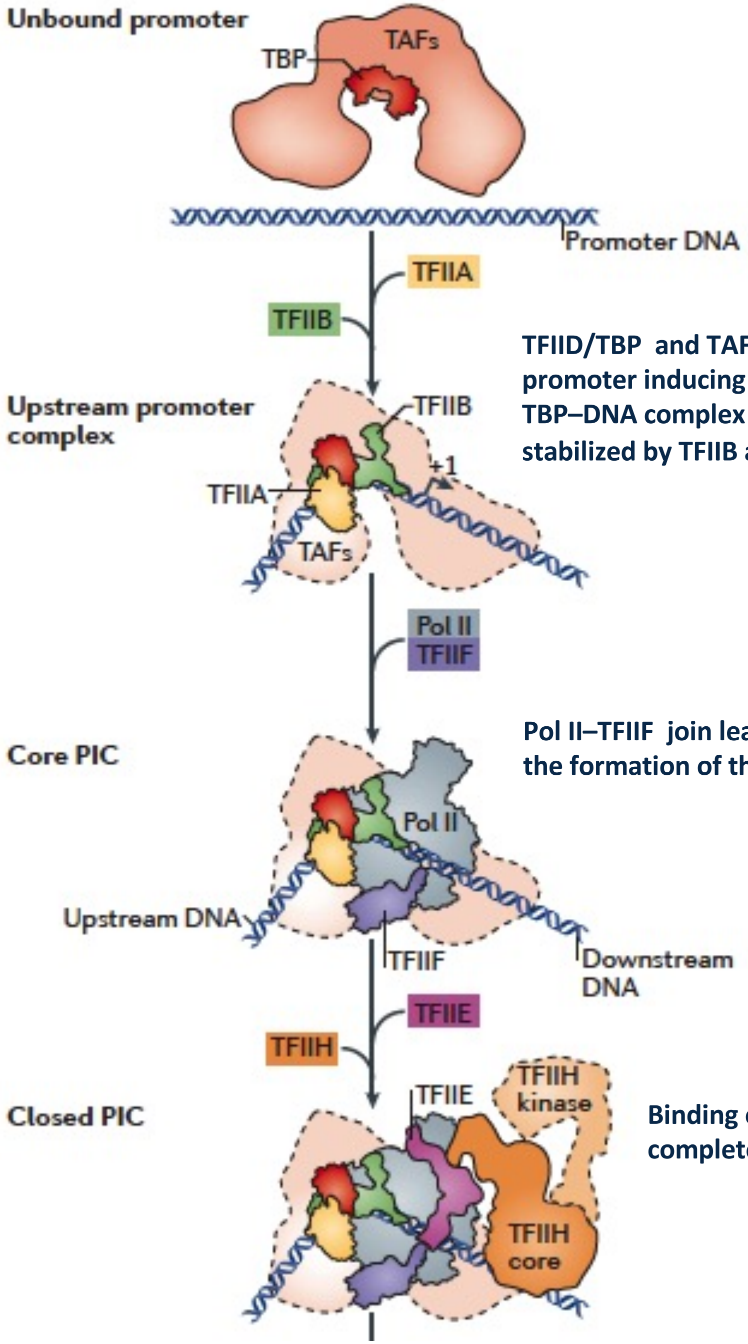
Harlan and Churchman, Nat Rev Mol Cell Biol, 2017

CTD code for different genes



Egloffs et al, 2012, TiG

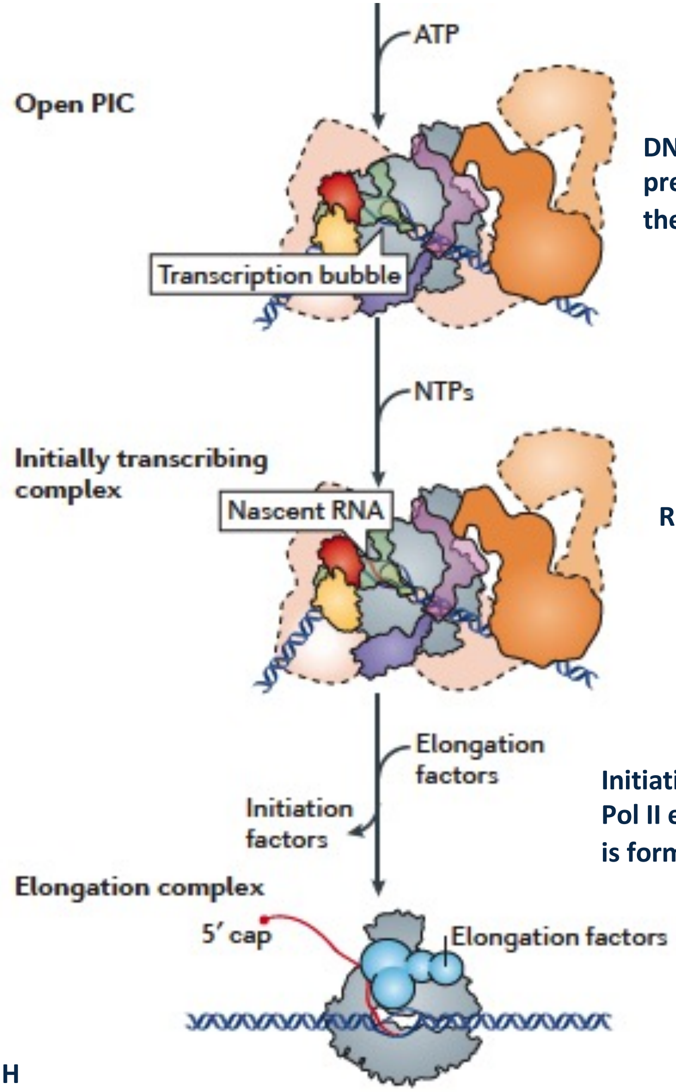
Pol II – initiation



TFIID/TBP and TAFs bind to promoter inducing a bend. TBP–DNA complex is stabilized by TFIIB and TFIIA

Pol II–TFIIF join leading to the formation of the core PIC.

Binding of TFIIE and TFIIH completes the PIC



DNA is opened in the presence of ATP, forming the transcription bubble

RNA synthesis starts

Initiation factors leave, Pol II elongation complex is formed

Most regulatory steps occur during intiation

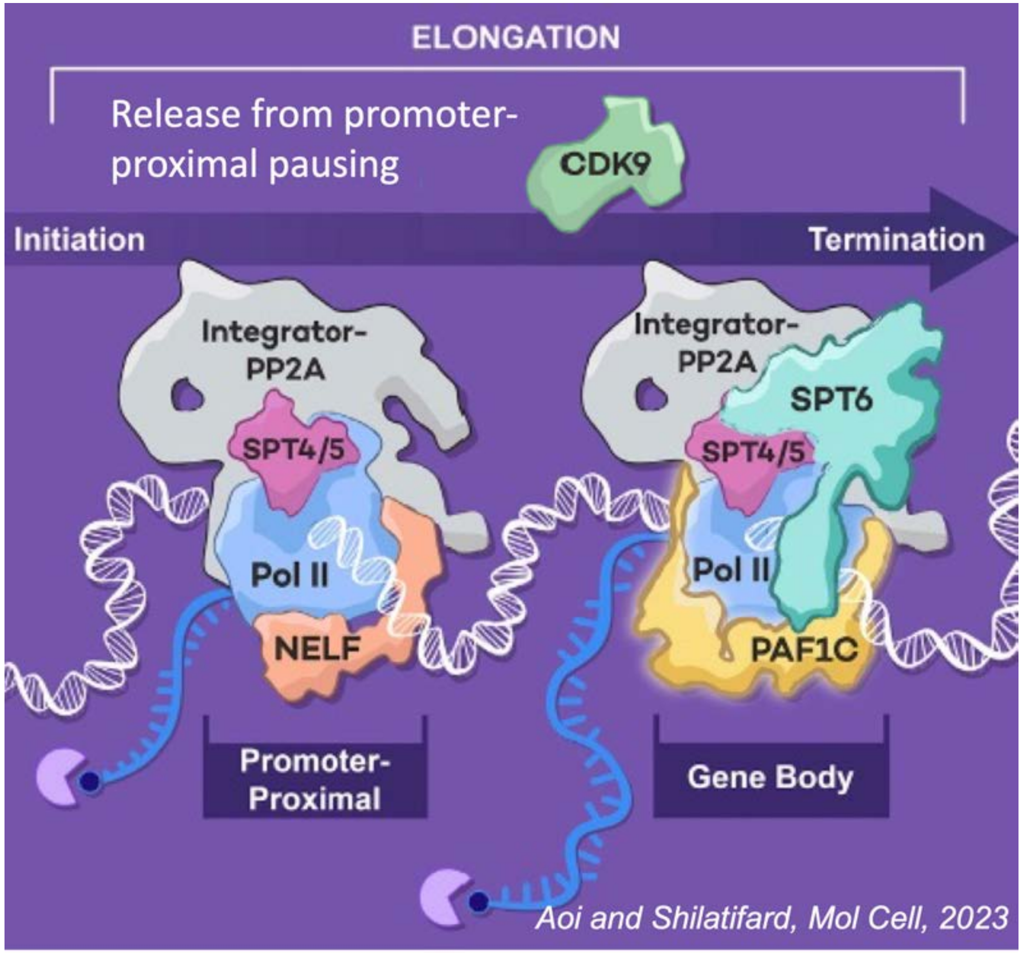
Sainsbury, Bernecky and Cramer, Nat Rev Mol Cell Biol'15

Pol II – initiation to elongation transition

P-TEFb Positive Transcription Elongation Factor b, cyclin-dependent kinase
= CDK9 (catalytic) **+ CycT** (cyclin T, regulatory)
Different P-TEFb complexes contain CycT1 or CycT2a/CycT2b and shorter or longer CDK9
There are additional CDK9 complexes: **CDK9-BDR4** and **CDK9-SEC**

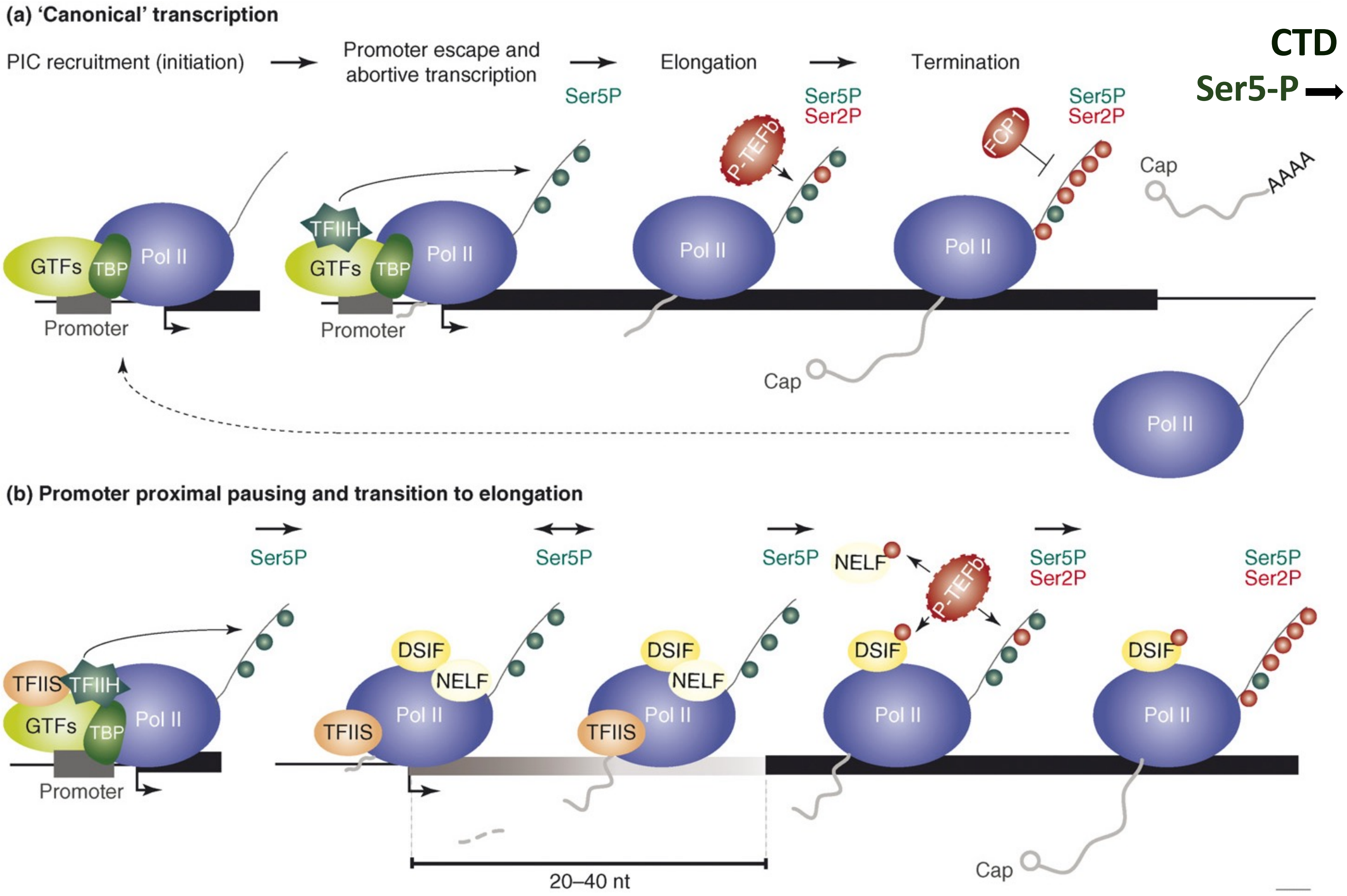
DSIF universal elongation factor = SPT4 + SPT5
PAF1C Pol II associated factor 1 complex
NELF negative elongation factor
CDK12, CDK13 elongation kinases

P-TEFb
Stimulates
SPT6 Recruits PAF1C and other elongation factors to Pol II
PAF1C Stimulates promoter-proximal pausing release and rate of elongation
DSIF Stimulates the rate and speed of Pol II
SPT5 Stimulates processive elongation. Stabilizes Pol II
NELF Recruits CBC and 3’ end processing factors (transcript stabilization). Suppresses premature termination
Integrator Multifunctional Pol II regulator



Aoi and Shilatifard, Mol Cell, 2023

Pol II – initiation to elongation transition

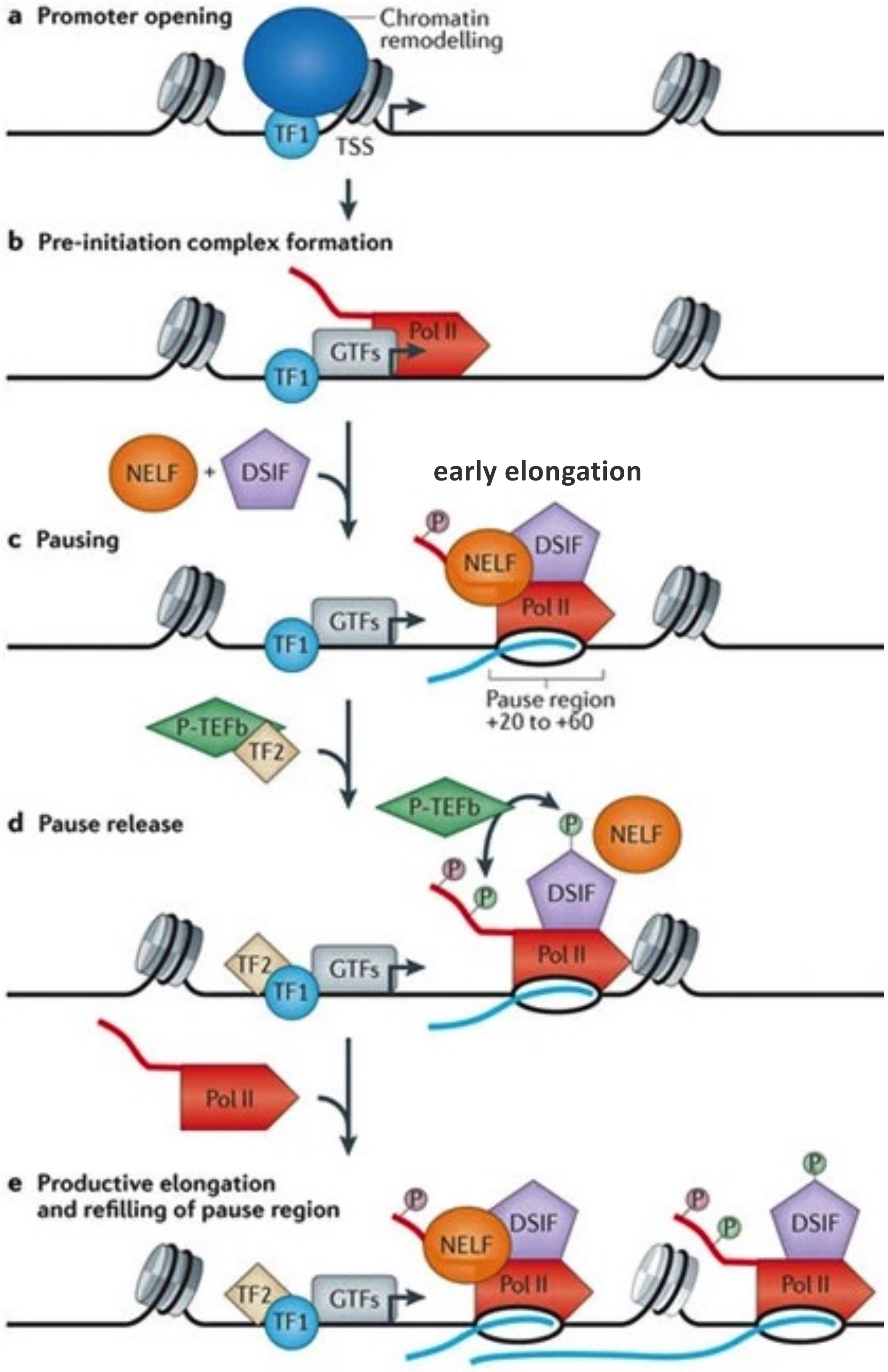


CTD
Ser5-P ➡ Ser2-P

Promoter proximal pausing involves abortive transcription. While waiting for Ser2-P, PolII transcribes short (20-40 nt) nascent RNA cleaved by elongation factor TFIIS, which allows PolII backtracking to resume transcription after arrest.

Koch et al., 2008, TiBS

Pol II – initiation to elongation transition



PIC formation

Pausing – regulatory step in metazoans

Association of NELF-DSIF (negative elongation factors) **cause Pol II pausing shortly after initiation.**

Pausing release

Recruitment of P-TEFb kinase that phosphorylates the NELF-DSIF complex triggers NELF release and transforms DSIF into a positive elongation factor.
P-TEFb also phosphorylates CTD at Ser2.
Paused Pol II is released.

Elongation

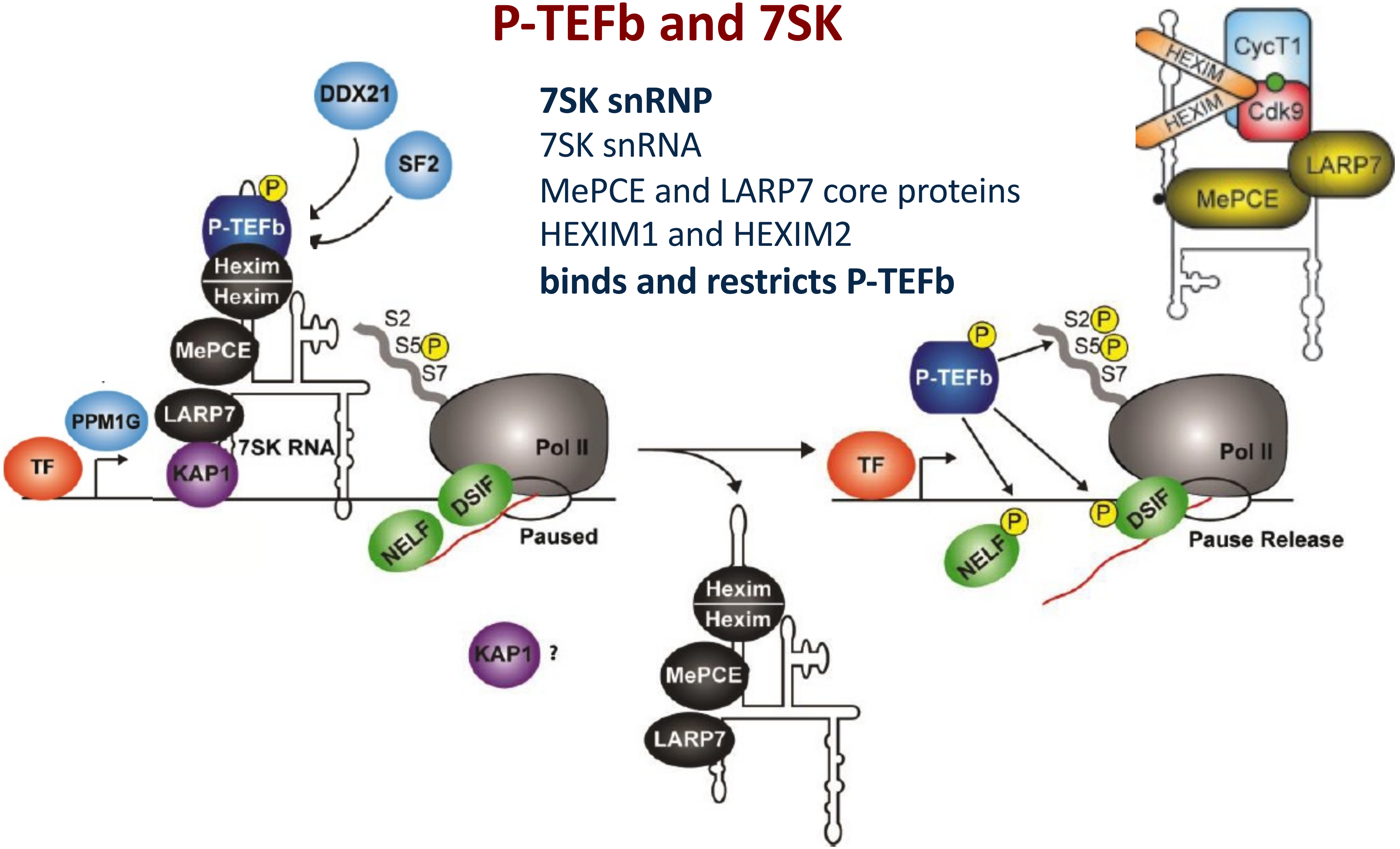
Recruitment of PAF1, Cdk12 and CycK.
Pol II enters productive elongation, followed by entry of another Pol II, leading to efficient RNA synthesis

Adelman and Lis, 2012, Nat Rev Genet

Pol II – initiation to elongation transition

P-TEFb and 7SK

7SK snRNP
7SK snRNA
MePCE and LARP7 core proteins
HEXIM1 and HEXIM2
binds and restricts P-TEFb



Quaresma et al, 2016, NAR; McNamara et al, 2016, Cell Cycle

7SK snRNP complex is present at promoters of most Pol II genes. Following stimuli P-TEFb is released and recruited to paused Pol II.

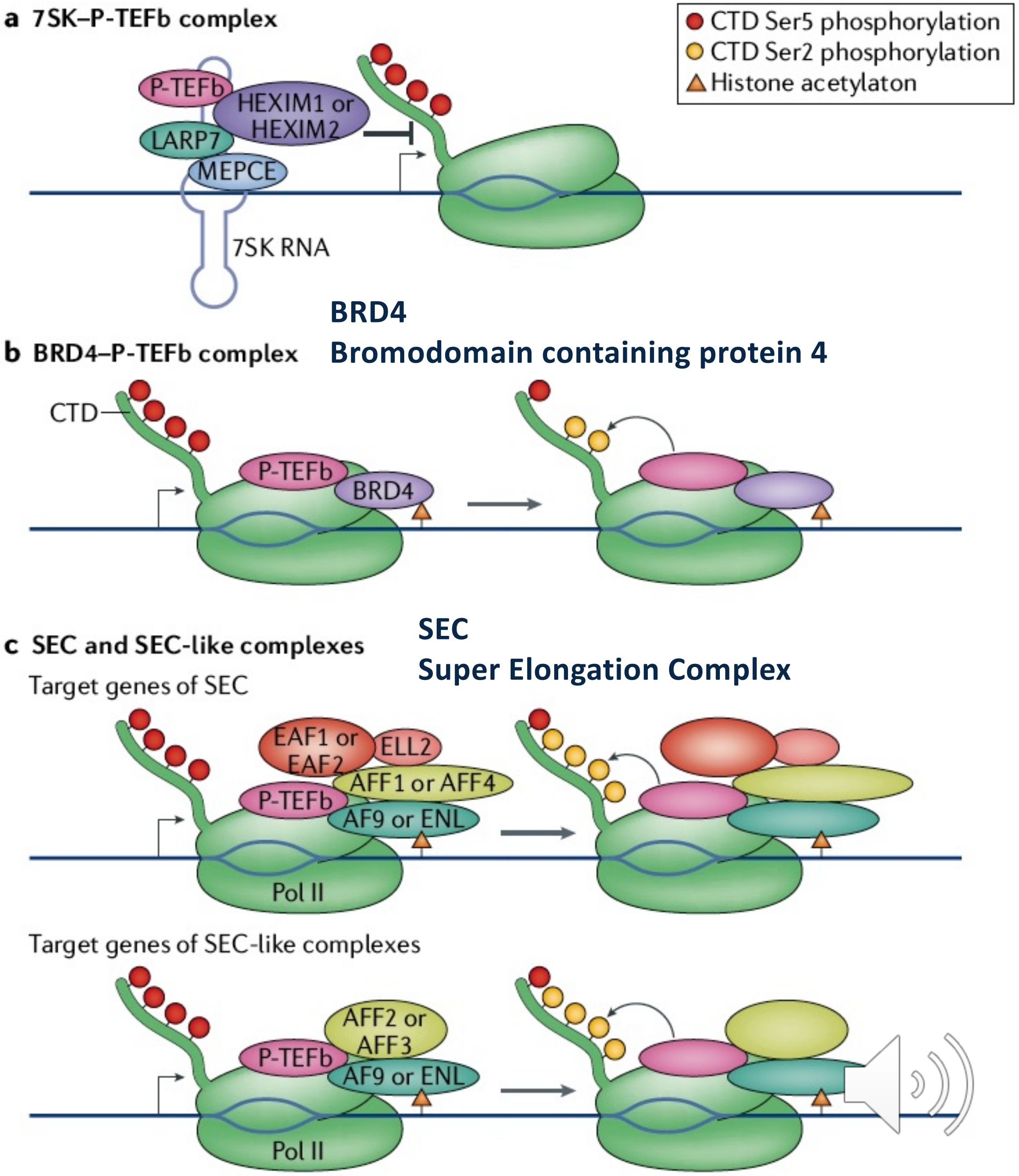
In proliferating cells, 50%–90% of P-TEFb exists within 7SK snRNP

Different P-TEFb containing complexes

7SK snRNP
represses transcription by sequestering and inhibiting up to 90% of cellular P-TEFb

SEC and BRD4 complexes activate P-TEFb in promoting Pol II release from pausing

Chen et al. Nat Rev Mol Cell Biol 2018

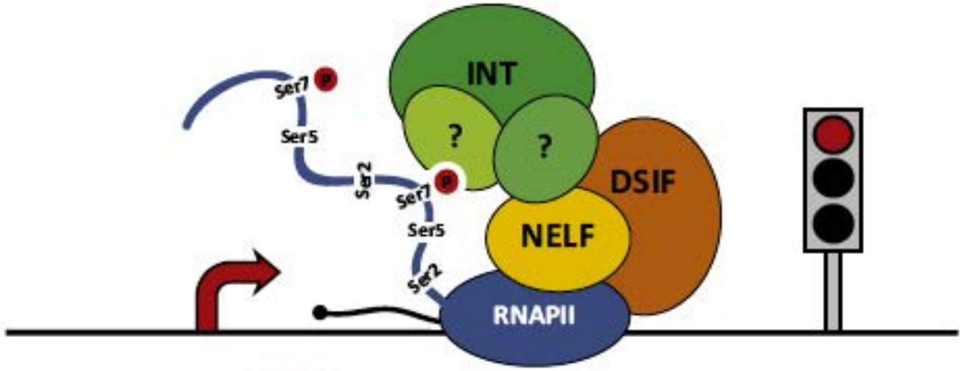


BRD4
Bromodomain containing protein 4

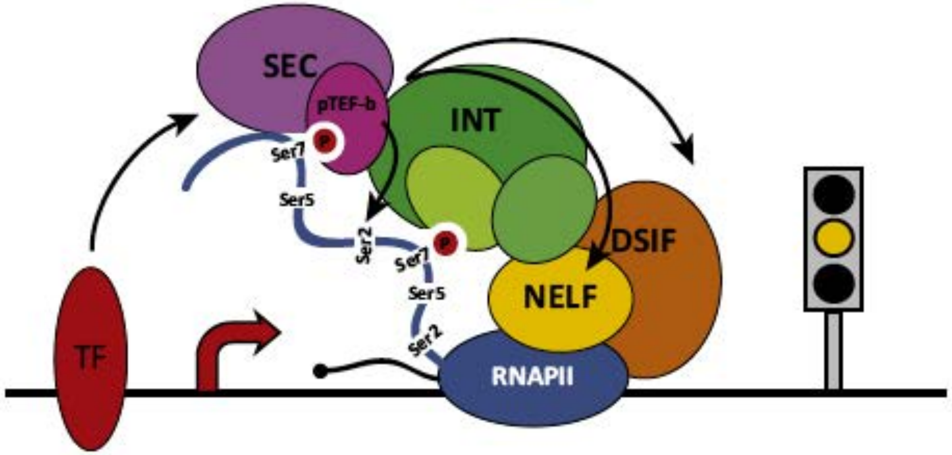
SEC
Super Elongation Complex

Integrator and promoter-proximal pause-release (INT)

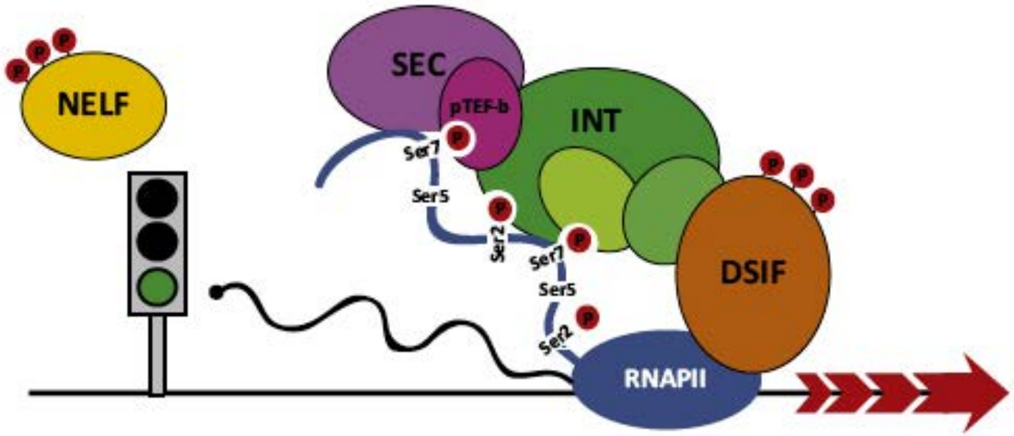
- Pol II pausing after initiation 40-60 nts downstream of TSS, with NELF and DSIF



- On activation, INT enriched at pause sites rectruits p-TEFb and SEC which phosphorylate NELF/DSIF and CTD at Ser2

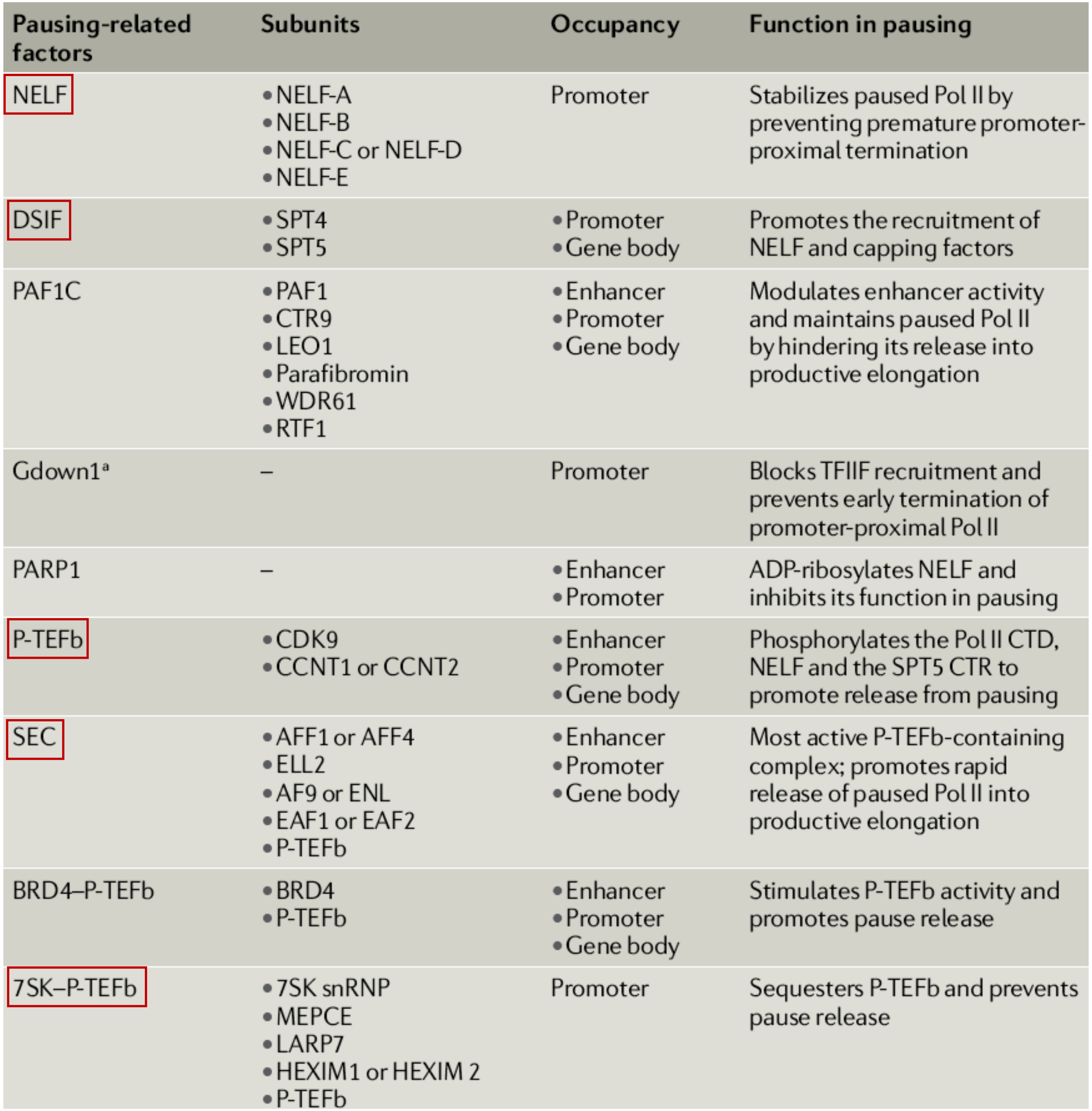


- NELF-P is displaced, DSIF-P becomes a positive elongation factor
- Pol II is elongation-competent



Baillat and Wagner, TiBS 2015

Pausing-related factors



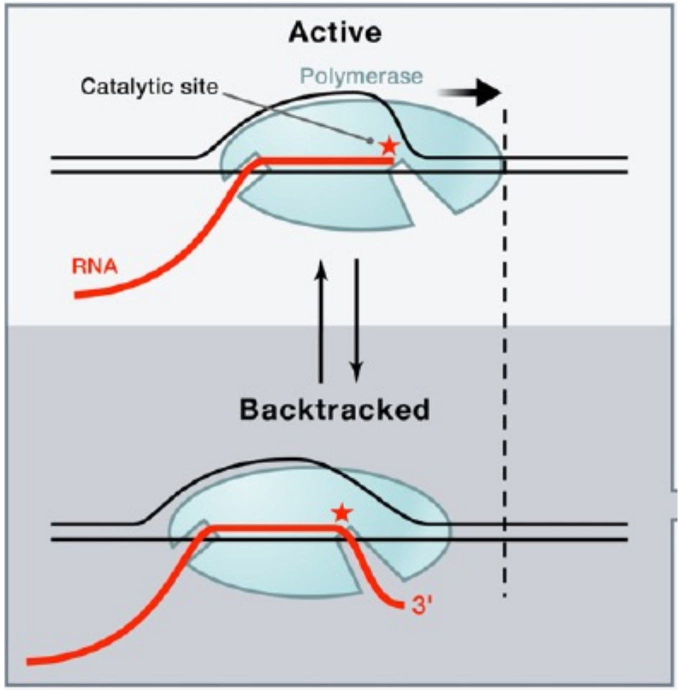
Pausing-related factors	Subunits	Occupancy	Function in pausing
NELF	• NELF-A • NELF-B • NELF-C or NELF-D • NELF-E	Promoter	Stabilizes paused Pol II by preventing premature promoter-proximal termination
DSIF	• SPT4 • SPT5	• Promoter • Gene body	Promotes the recruitment of NELF and capping factors
PAF1C	• PAF1 • CTR9 • LEO1 • Parafibromin • WDR61 • RTF1	• Enhancer • Promoter • Gene body	Modulates enhancer activity and maintains paused Pol II by hindering its release into productive elongation
Gdown1[a]	–	Promoter	Blocks TFIIF recruitment and prevents early termination of promoter-proximal Pol II
PARP1	–	• Enhancer • Promoter	ADP-ribosylates NELF and inhibits its function in pausing
P-TEFb	• CDK9 • CCNT1 or CCNT2	• Enhancer • Promoter • Gene body	Phosphorylates the Pol II CTD, NELF and the SPT5 CTR to promote release from pausing
SEC	• AFF1 or AFF4 • ELL2 • AF9 or ENL • EAF1 or EAF2 • P-TEFb	• Enhancer • Promoter • Gene body	Most active P-TEFb-containing complex; promotes rapid release of paused Pol II into productive elongation
BRD4–P-TEFb	• BRD4 • P-TEFb	• Enhancer • Promoter • Gene body	Stimulates P-TEFb activity and promotes pause release
7SK–P-TEFb	• 7SK snRNP • MEPCE • LARP7 • HEXIM1 or HEXIM 2 • P-TEFb	Promoter	Sequesters P-TEFb and prevents pause release

Additional elements

- R-loops and G4s
- Enhancers
- eRNAs
- Mediator
- Histone modifications
- Topoisomerases
- PARPs

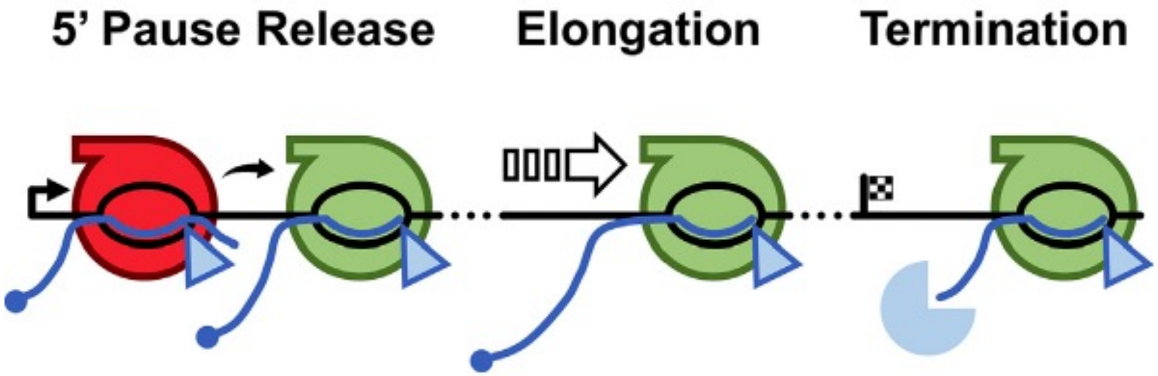
Chen et al. Nat Rev Mol Cell Biol 2018

Polymerase backtracking



Pol II arrest and backtracking occur at

- **roadblocks** (DNA binding factor or a nucleosome)
- **promoter-proximal pause sites**
- **terminators**
- **positions of base mis-incorporation**

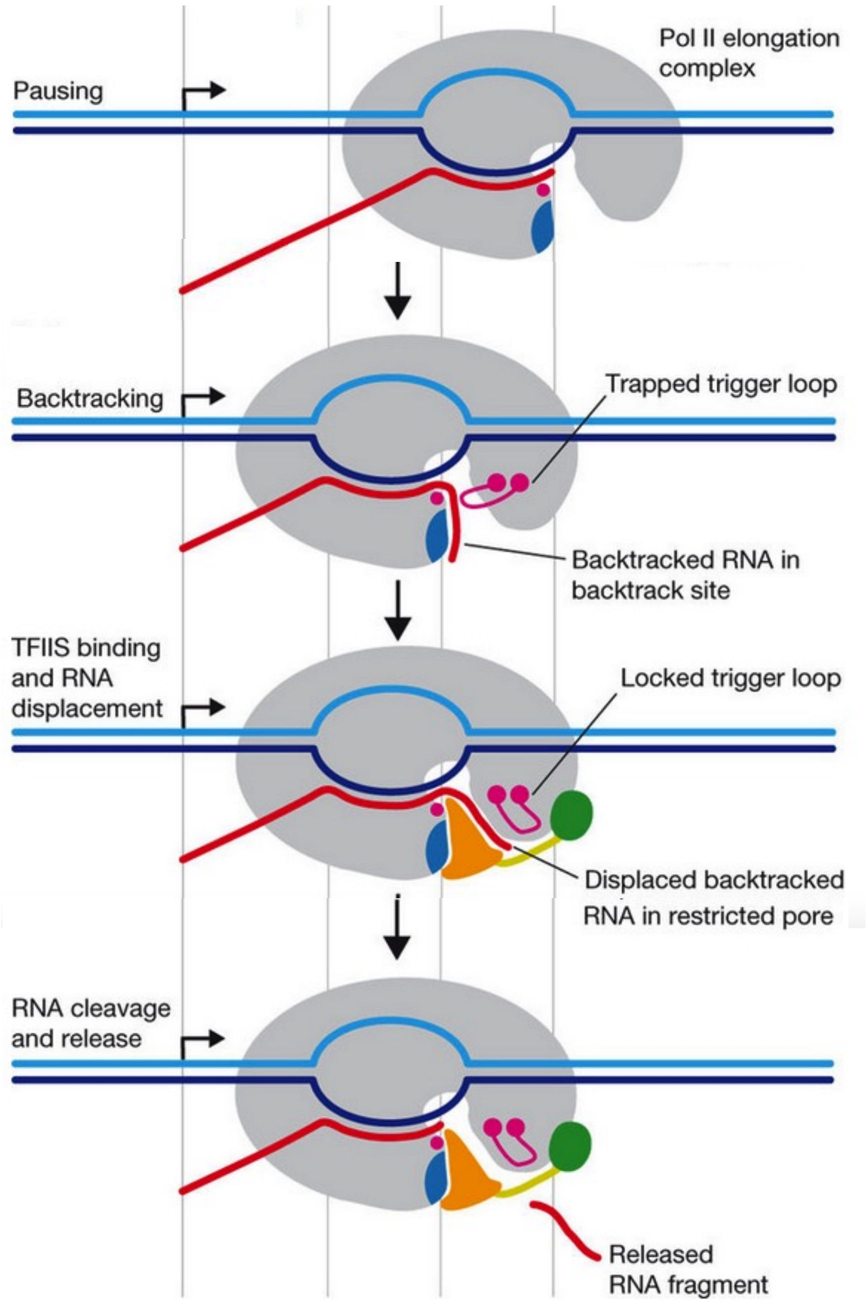


Backtracking functions

- Regulatory pauses and arrests
- Termination mechanisms
- Transcriptional fidelity
- Elongation rate control
- Co-transcriptional RNA folding and processing
- Genome stability
- Coupling transcription to translation in bacteria

Nudler, Cell, 2013; Sheridan et al, Mol Cell, 2018

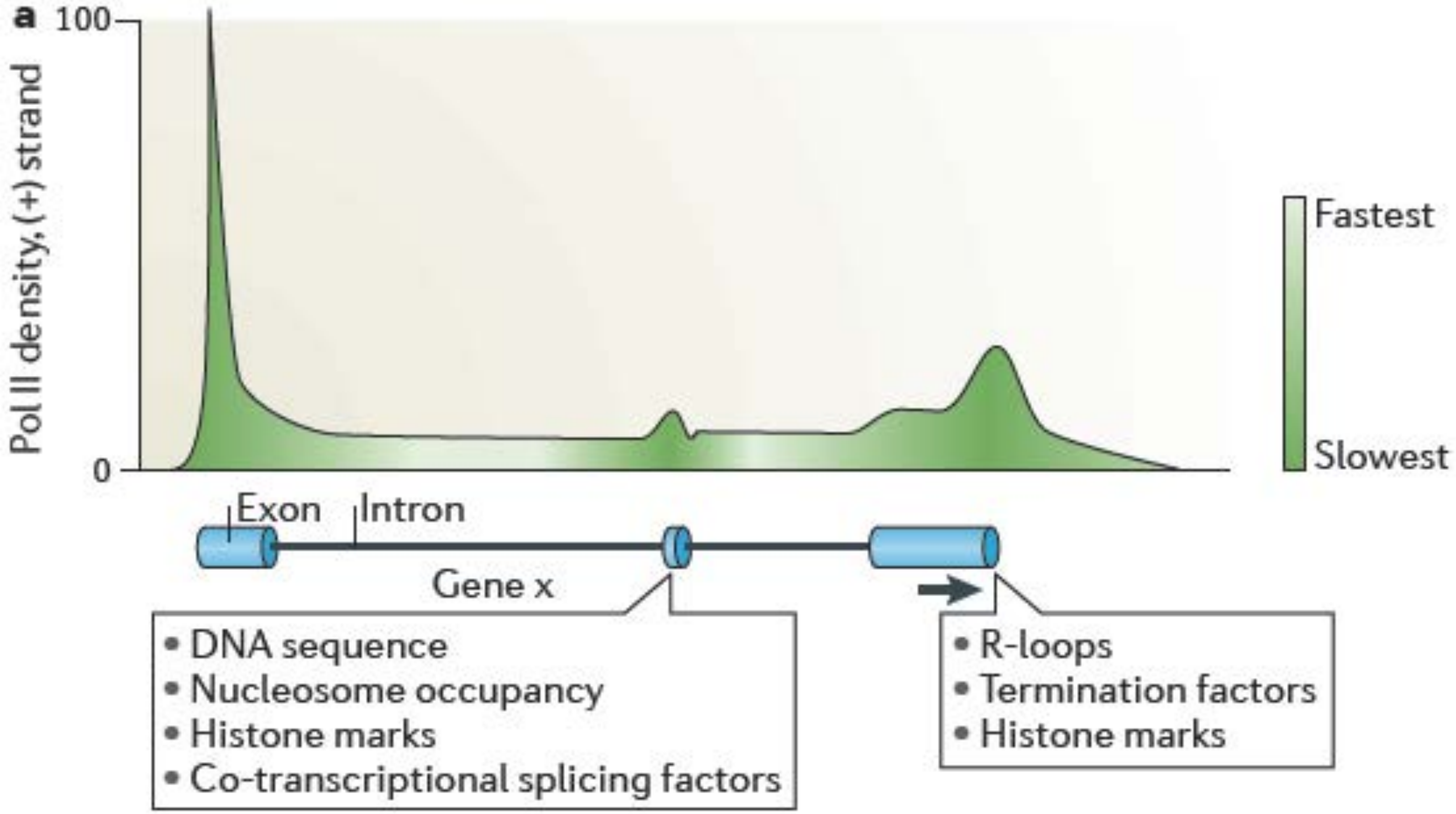
Polymerase backtracking and pause release



- **Rescue from backtracking is a major stimulus of rapid transcriptional elongation**
- **Rescue from backtracking is important for escape from promoter-proximal pause sites**
- **RNA cleavage by Pol II (TFIIS) is essential for rescue from backtracking and activation of transcription** (in particular of stress-inducible genes)

Cheung and Cramer, Nature 2011

Pol II elongation



- DNA sequence
- Nucleosome occupancy
- Histone marks
- Co-transcriptional splicing factors

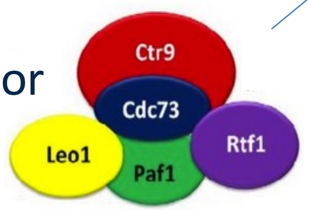
- R-loops
- Termination factors
- Histone marks

Pol II density and elongation rates are not constant but vary throughout the gene

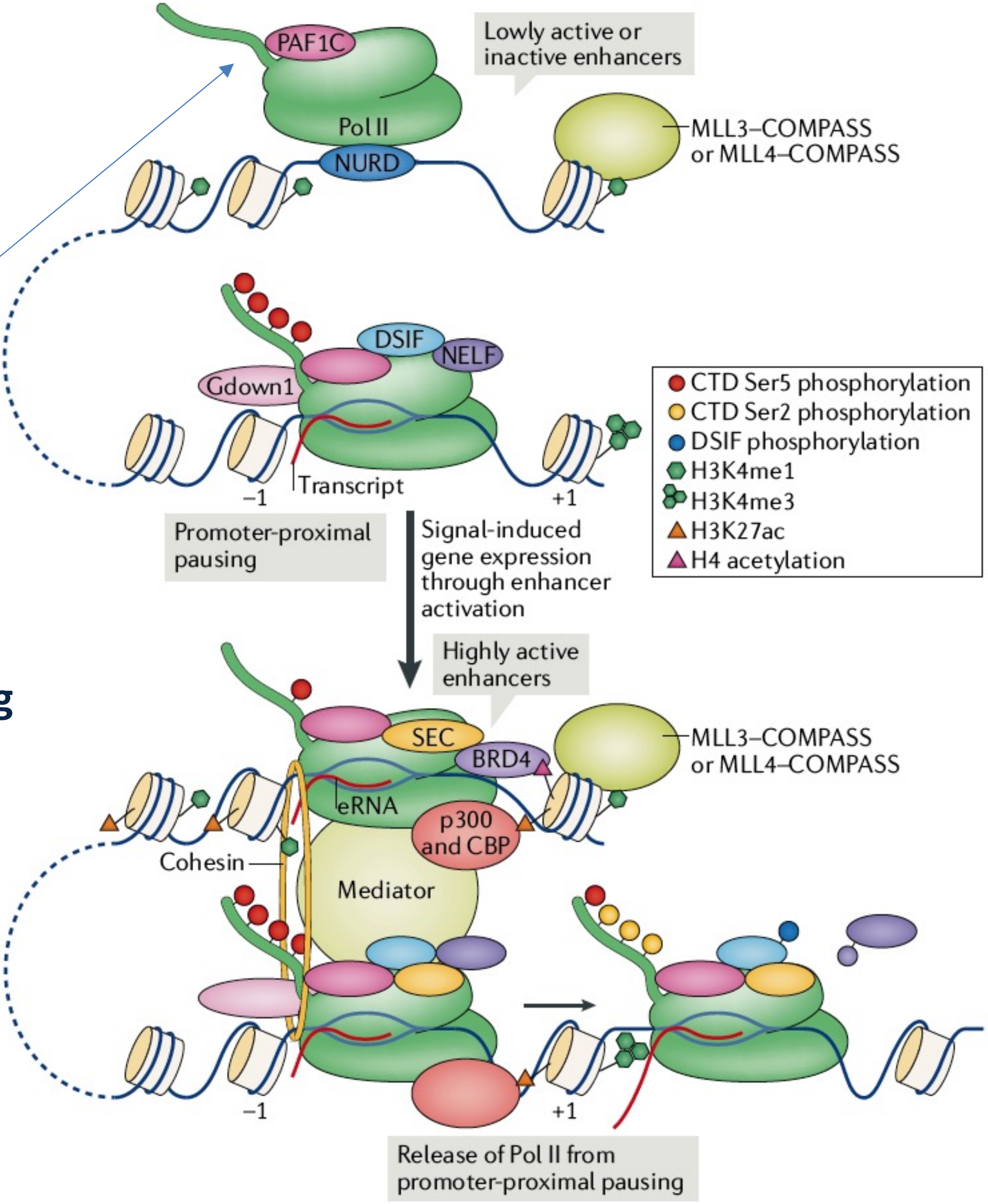
Jonkers and Lis, Nat Rev Mol Cell Biol'15

Pol II elongation is regulated by enhancers

PAF1C Pol II - associated factor pause release trx elongation



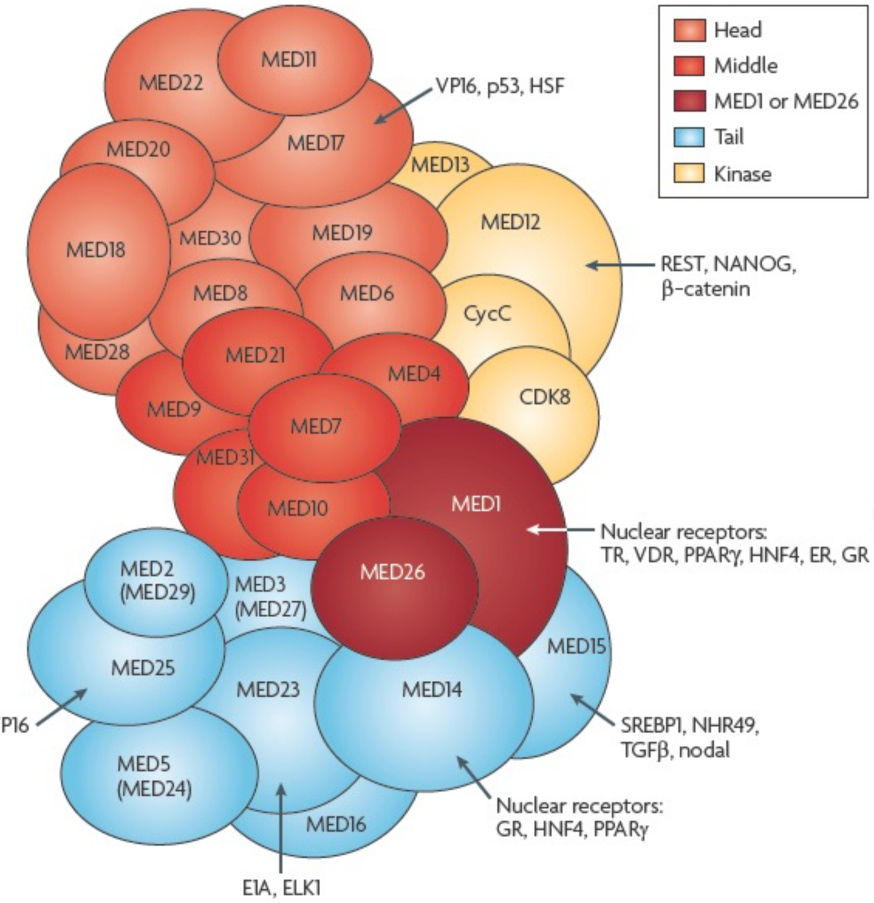
- Genes with less active enhancers have higher levels of Pol II pausing
- Activation of enhancers triggers interaction between enhancers and promoter through the Mediator complex and eRNAs
- This contributes to binding of BRD4 and SEC and pause release by P-TEFb, leading to elongation



Chen et al. Nat Rev Mol Cell Biol 2018

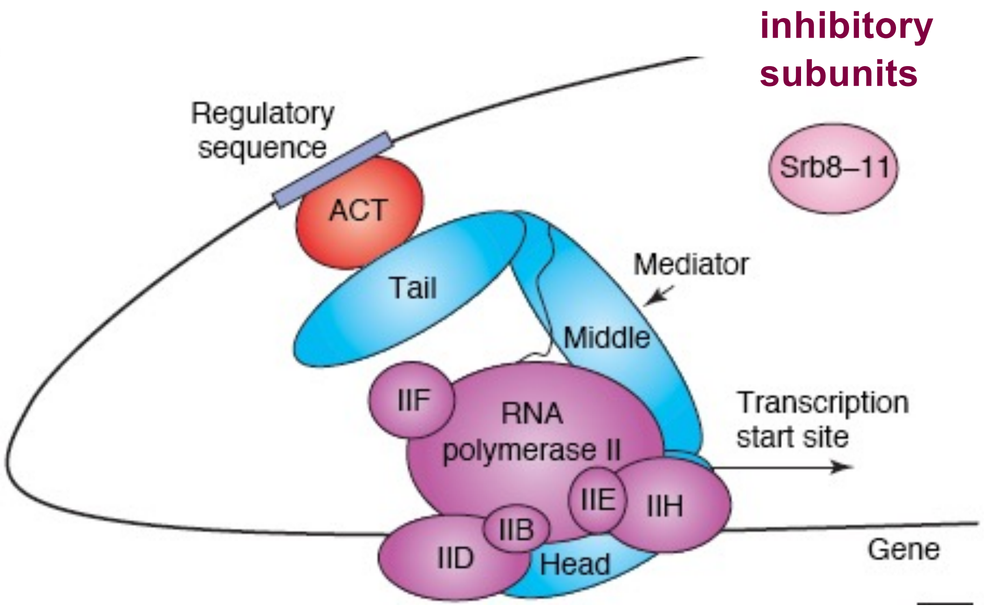
MEDIATOR - a central integrator of transcription

Metazoa



Malik and Roeder, 2010, Nat Rev Genet

Yeast, 25 subunits, 1.4 MDa



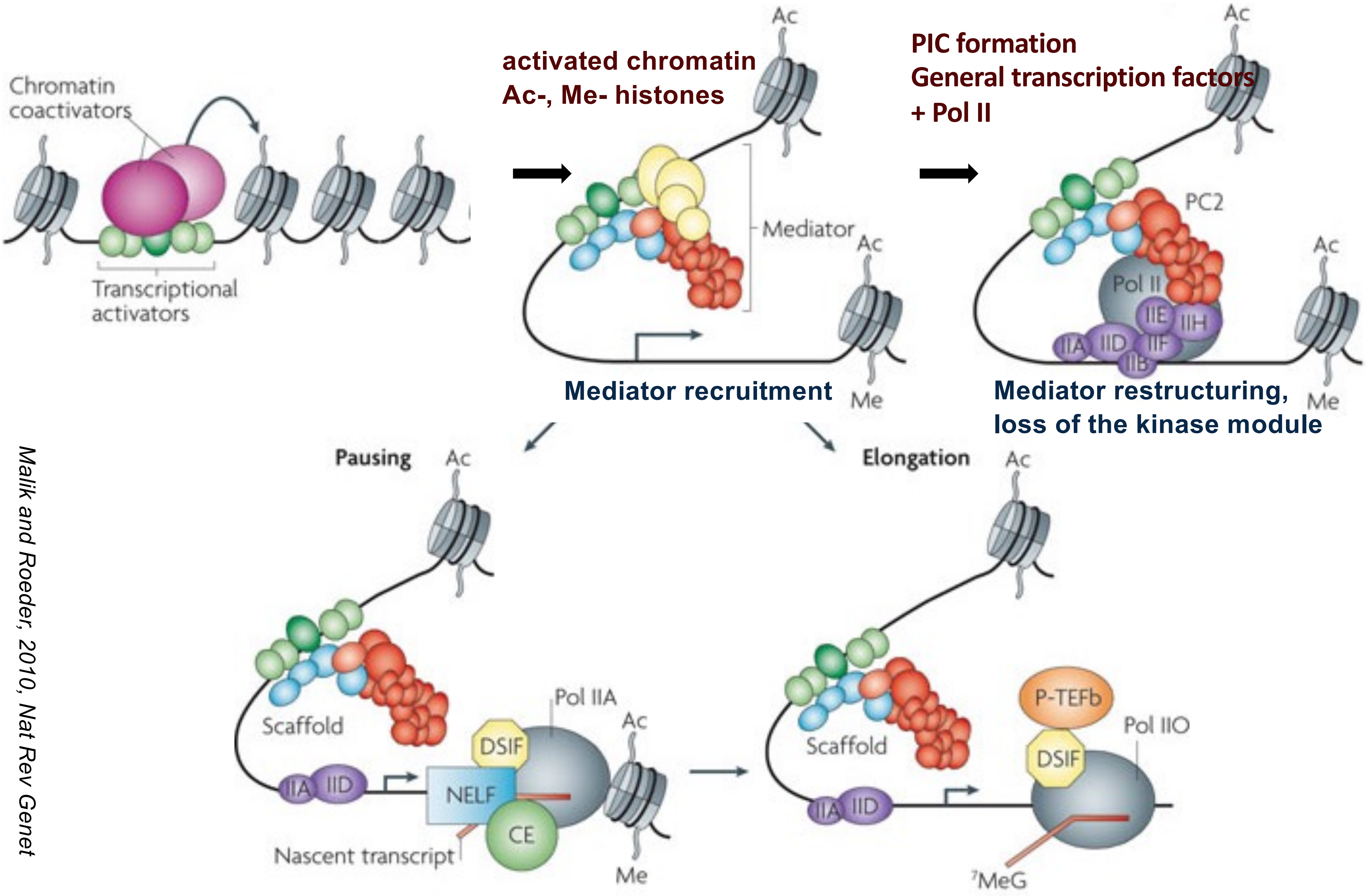
ACT - trx activator
II B D E H F – trx factors

Bjorklund and Gustafsson, 2005, TiBS

- **evolutionarily conserved, multiprotein complex**
- **transcriptional co-activator, sensor, integrator of signals**
- **also involved also in chromatin structure, formation of gene loops, gene silencing and development**

MEDIATOR - a central integrator of transcription

Mediator has a key role in chromatin architecture, PIC assembly, pausing and elongation



activated chromatin
Ac-, Me- histones

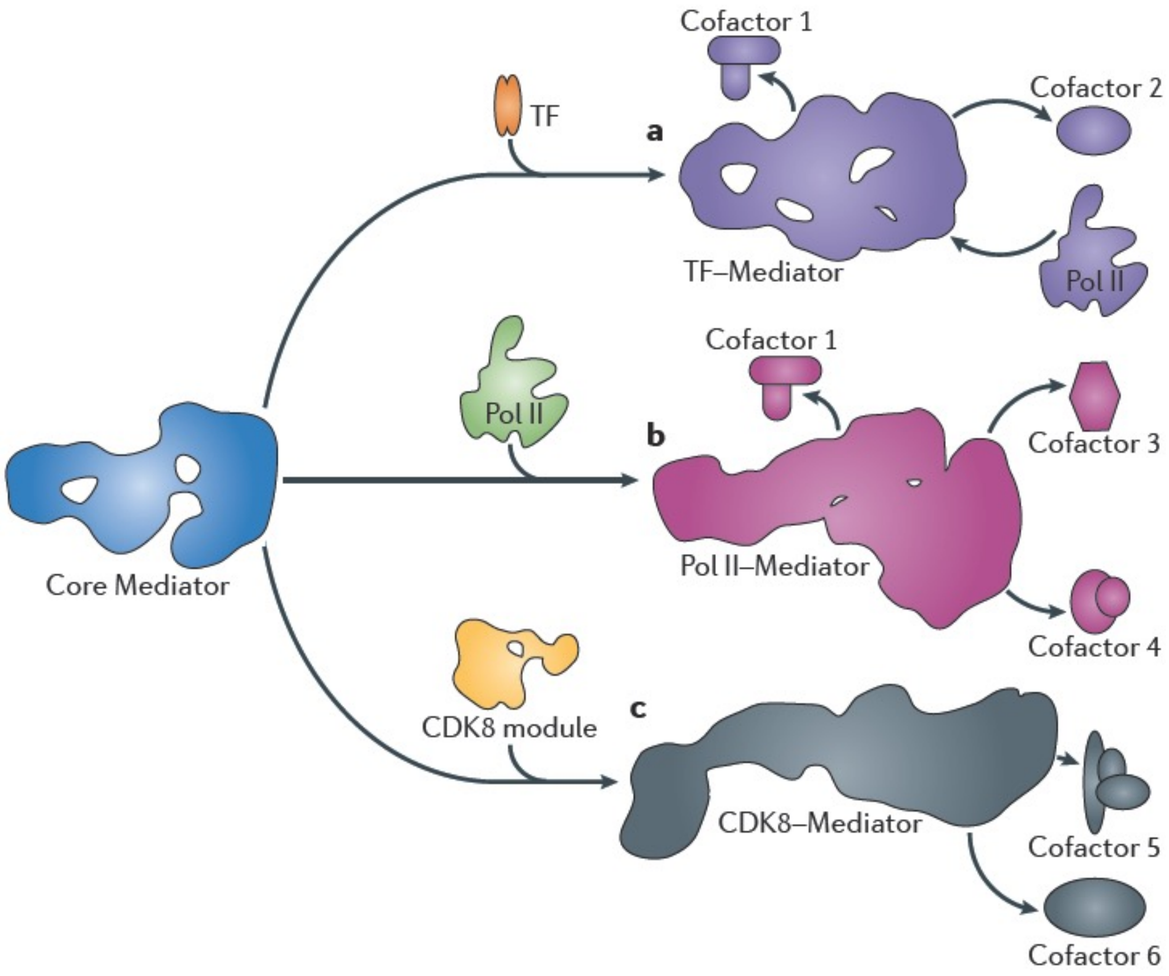
PIC formation
General transcription factors
+ Pol II

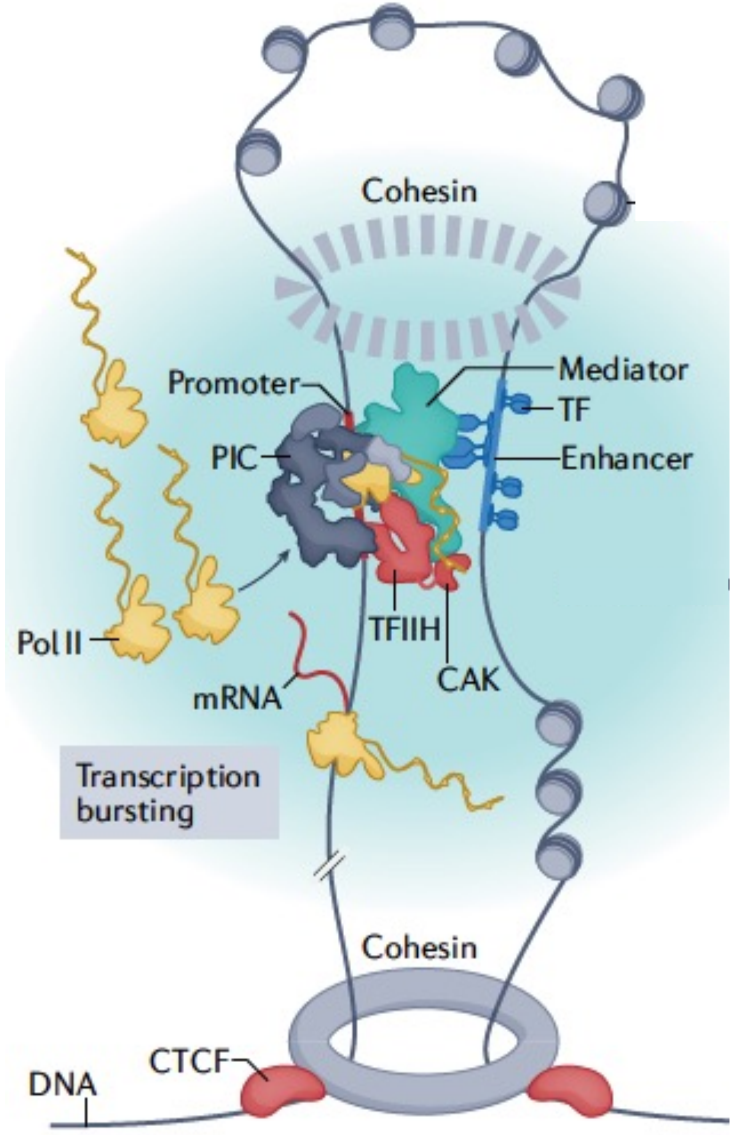
Mediator recruitment

Mediator restructuring,
loss of the kinase module

Malik and Roeder, 2010, Nat Rev Genet

MEDIATOR - a central integrator of transcription

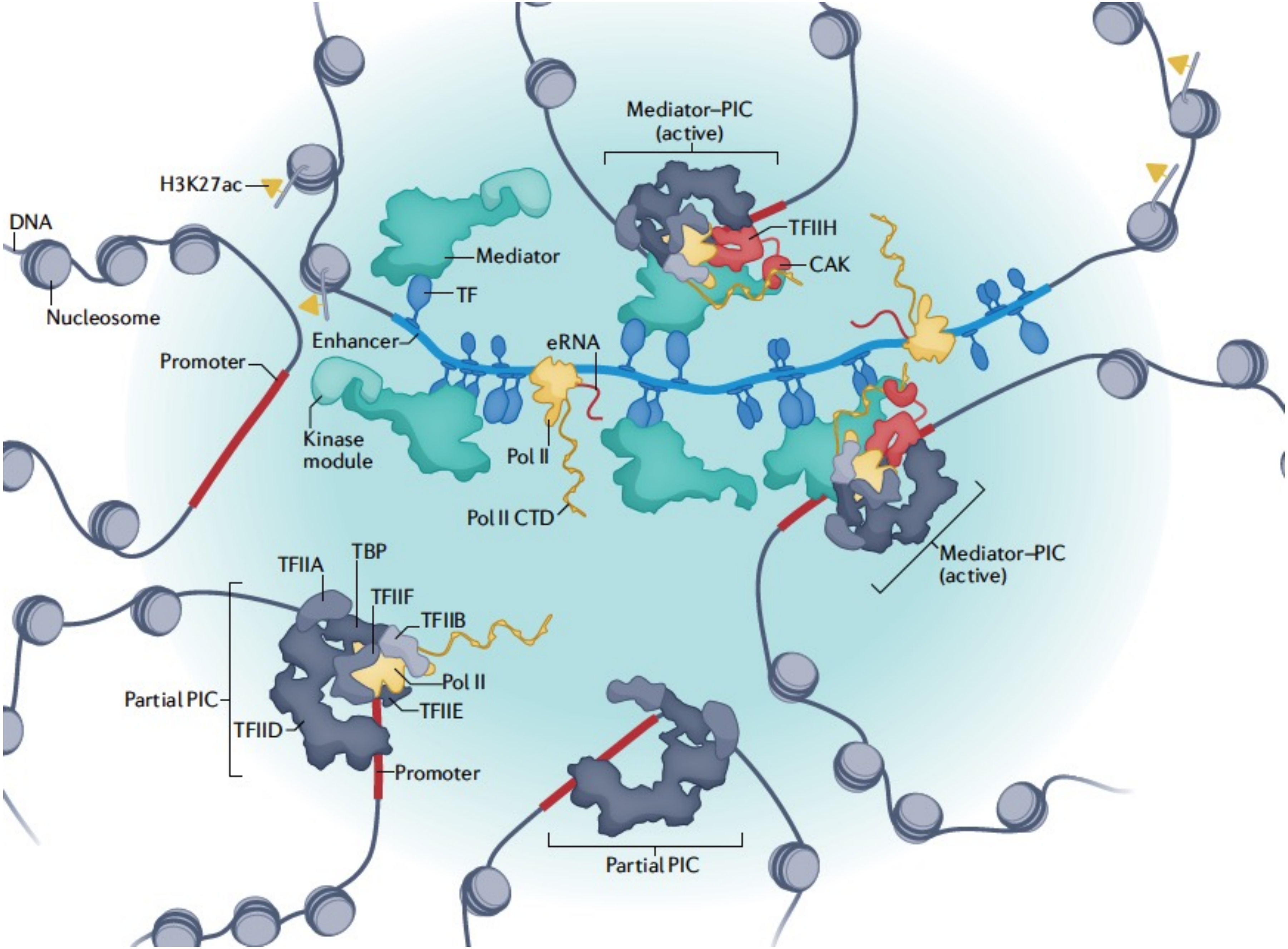




Structural changes in Mediator control its interactions with other transcription regulators

Allen and Taatjes, 2015, Nat Rev Mol Cell Biol; Chen et al. Nat Rev Mol Cell Biol 2018

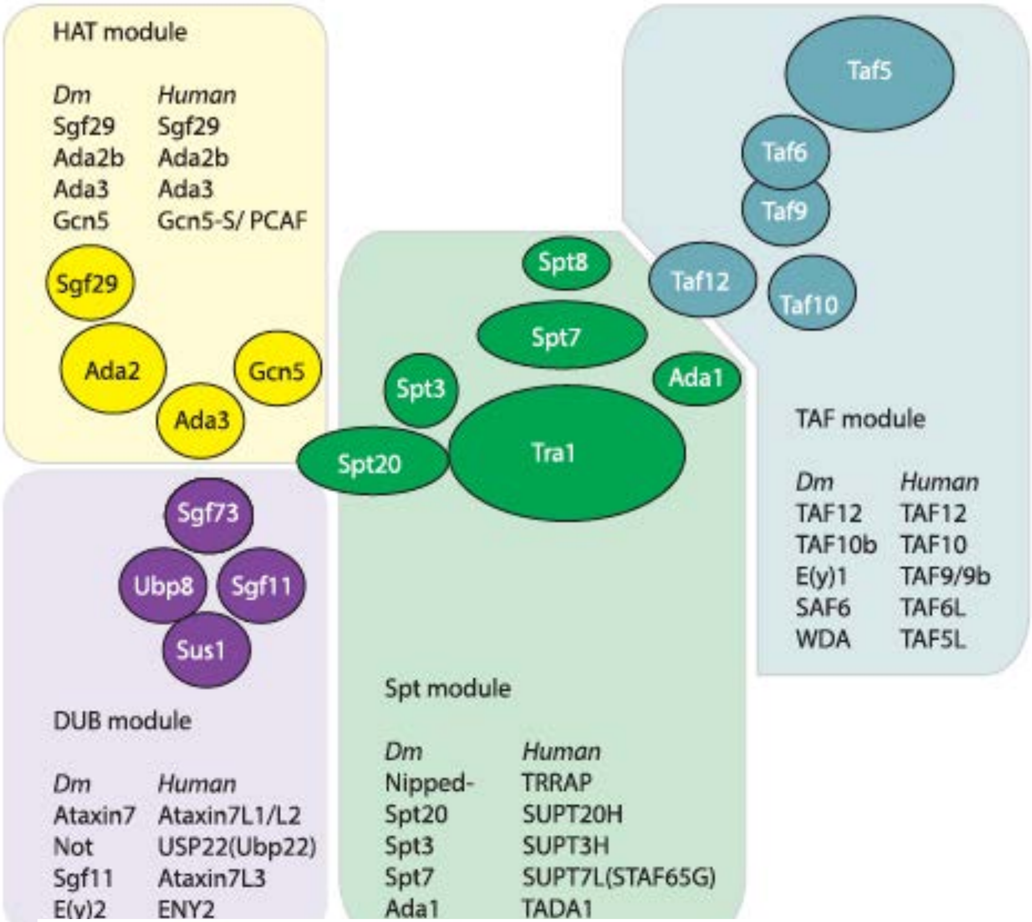
MEDIATOR - a central integrator of transcription



Chen et al. Nat Rev Mol Cell Biol 2018

SAGA

Spt–Ada–Gcn5 acetyltransferase



HAT module

Dm	Human
Sgf29	Sgf29
Ada2b	Ada2b
Ada3	Ada3
Gcn5	Gcn5-S/ PCAF

DUB module

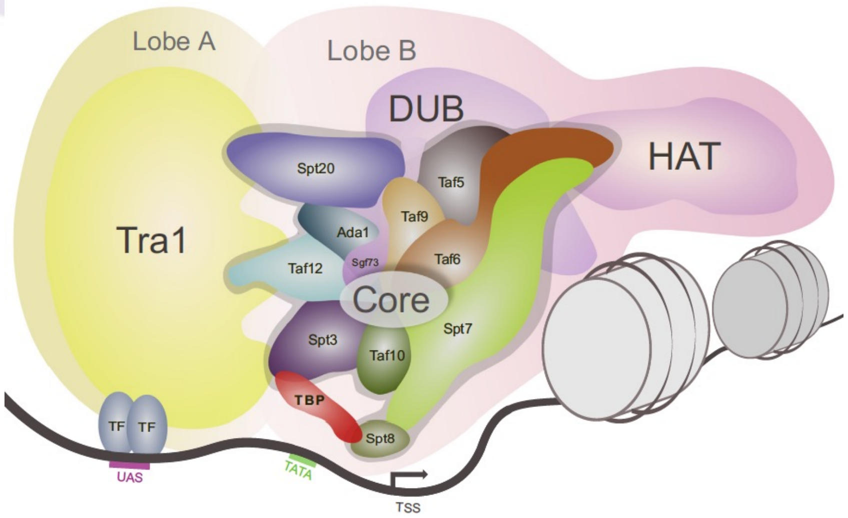
Dm	Human
Ataxin7	Ataxin7L1/L2
Not	USP22(Ubp22)
Sgf11	Ataxin7L3
E(y)2	ENY2

Spt module

Dm	Human
Nipped-	TRRAP
Spt20	SUPT20H
Spt3	SUPT3H
Spt7	SUPT7L(STAF65G)
Ada1	TADA1

TAF module

Dm	Human
TAF12	TAF12
TAF10b	TAF10
E(y)1	TAF9/9b
SAF6	TAF6L
WDA	TAF5L



- multisubunit histone modifying complex (2 MDa)
- contains four modules

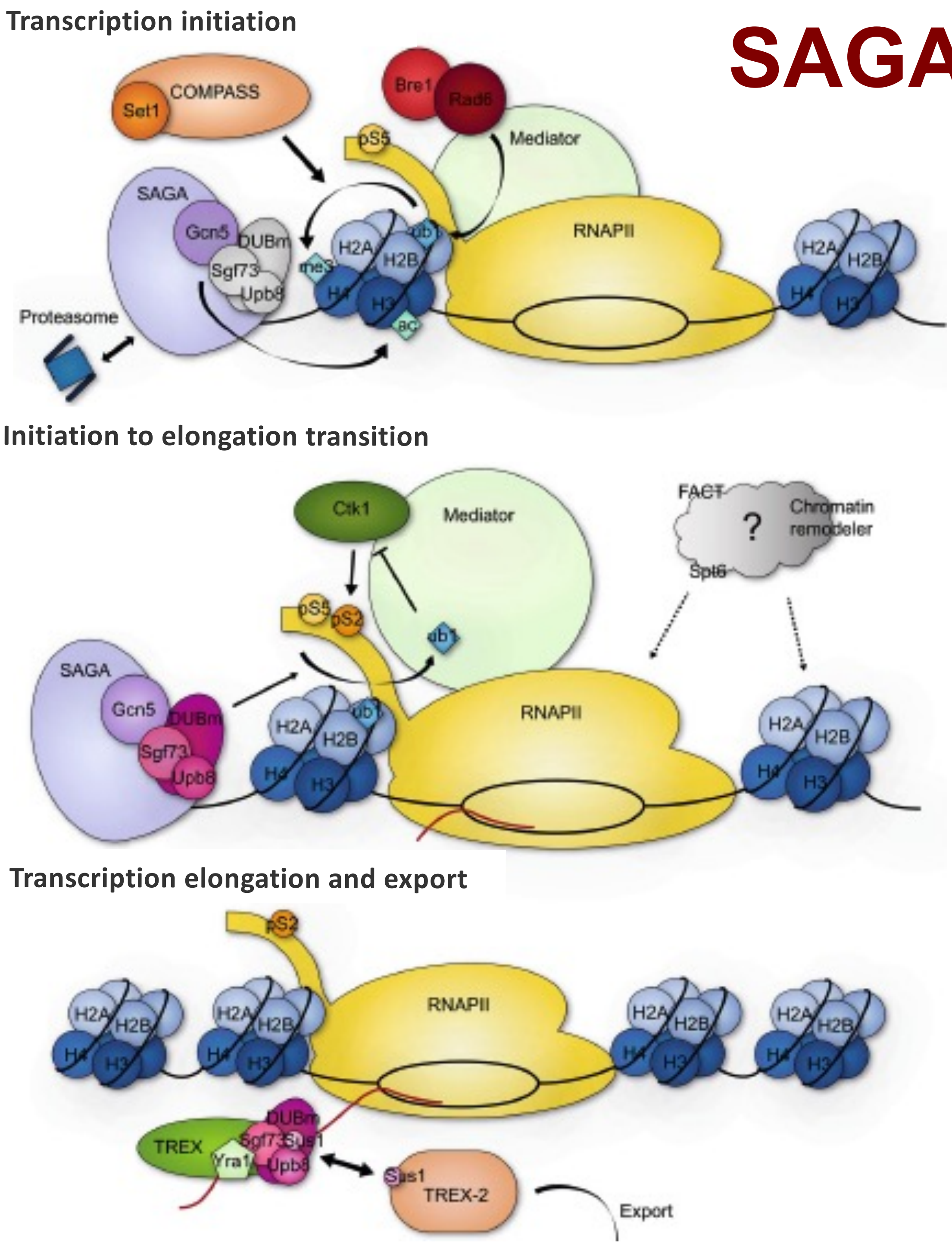
 - HAT: histone acetylation
 - DUB: histone deubiquitination
 - TAF } CORE PIC assembly
 - Spt } TFIID, TBP recruitment

- transcriptional activator
- interacts with TFs via Tra1
- involved also in

- transcript elongation
- regulation of protein stability
- telomere maintenance

Soffers and Workman, 2020, Gene Dev

SAGA in transcription


Transcription initiation

Initiation to elongation transition

Transcription elongation and export

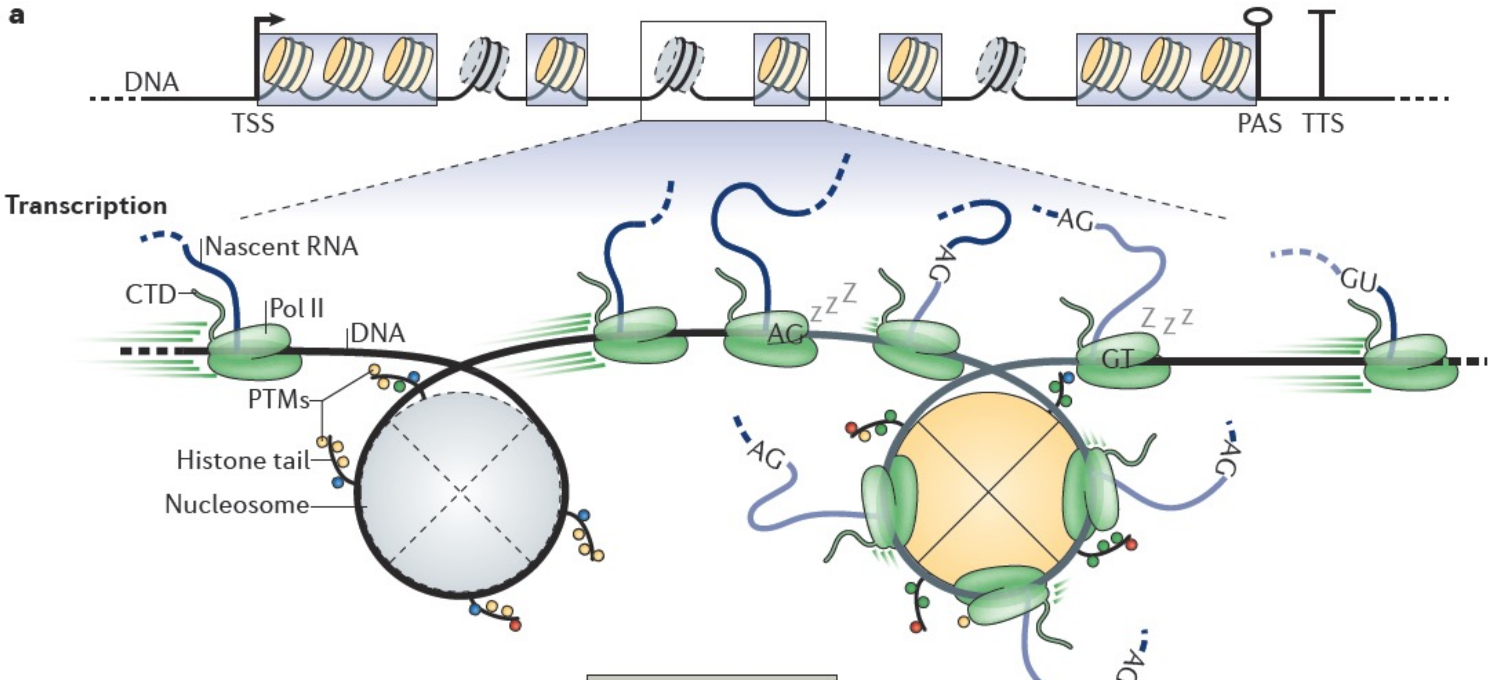
Histone acetylation by HAT activity of SAGA, followed by H3-me by Set1/Compass, promotes open chromatin, which favors initiation

Deubiquitylation of H2B-Ub, mediated by the SAGA DUB module, facilitates recruitment of Ctk1, which phosphorylates CTD at Ser2, leading to the release of paused Pol II

H2B-Ub support open chromatin and cooperate with FACT to remodel nucleosomes. H2B deubiquitylation stimulates H3-me, supporting productive transcription

Cheon et al, 2020, Exp Mol Medicine

Nucleosome positioning

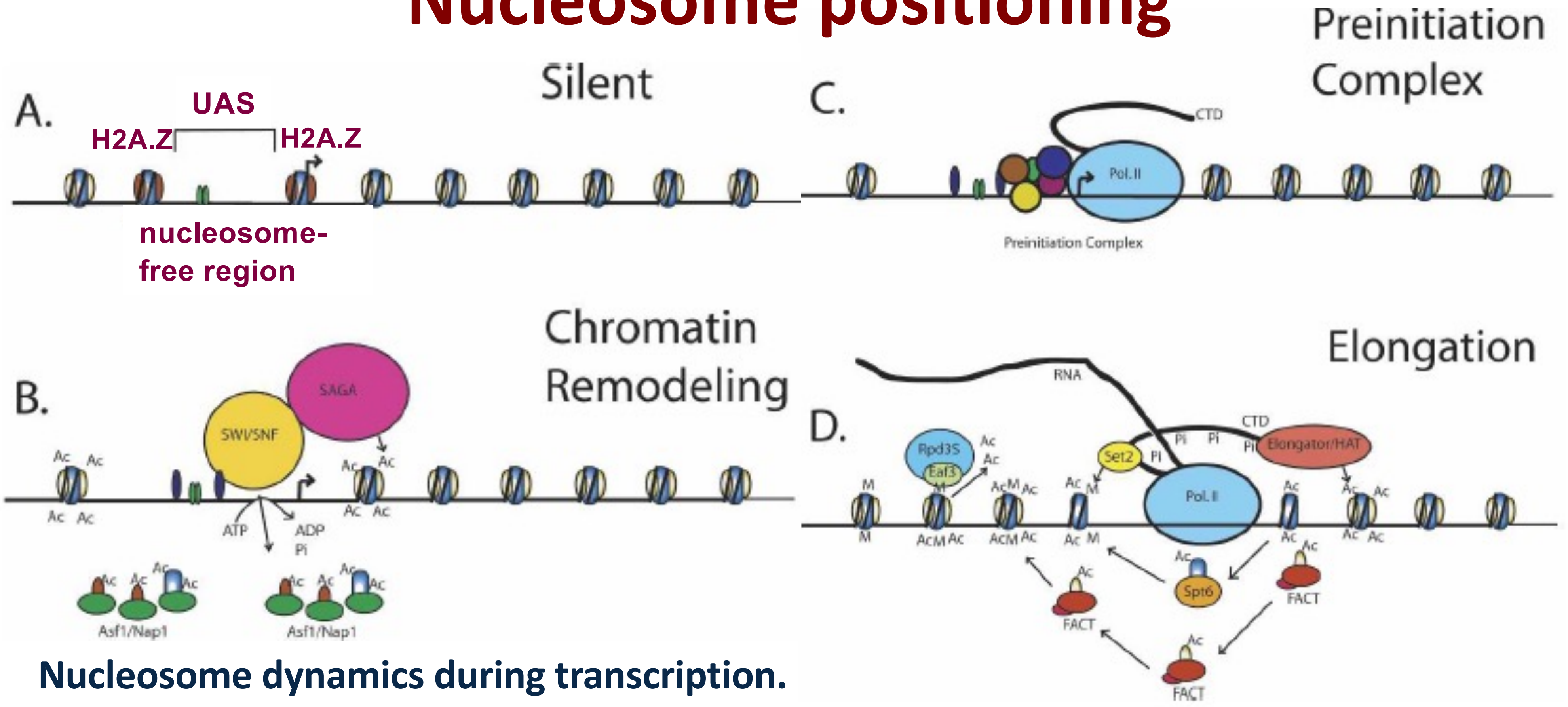


Nucleosome positioning relative to TSS (trx start site) **and TTS** (trx termination site) **and exons defines their boundaries. This provides a platform for crosstalk between chromatin, transcription and splicing**

Nucleosomes at introns are less stable (dashed lines, grey) - **Pol II is faster**
Nucleosome phasing over exons leads to slower Pol II over exons

Pol II pauses at splice sites (AG and GT)

Nucleosome positioning



Nucleosome dynamics during transcription.

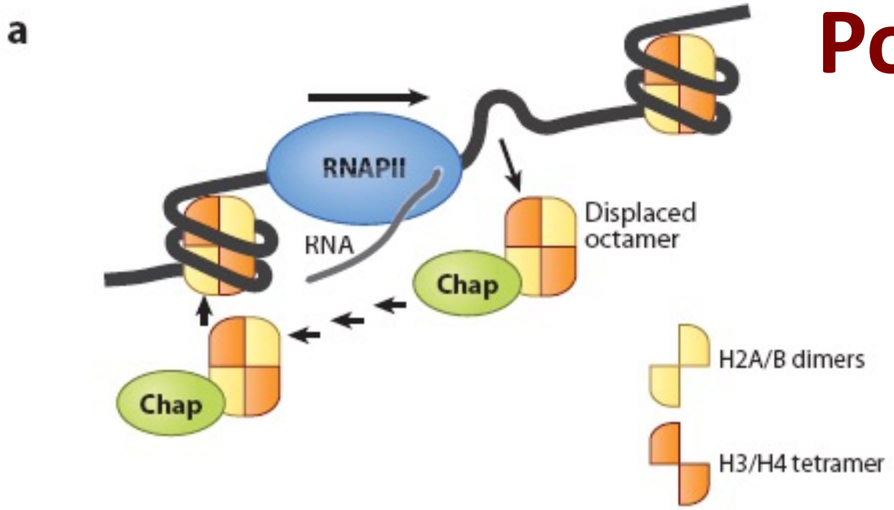
<u>Initiation:</u> DNA-binding activators at UAS recruit SAGA (acetylates nucleosomes) **and SWI/SNF** (displaces nucleosomes)**. Histones are transferred to histone chaperones. PIC/Pol II assemble at the nucleosome-free region.**

<u>Elongation:</u> Nucleosomes in front of Pol II are acetylated and displaced to Spt6/FACT chaperones, which reassemble nucleosomes behind Pol II. H3 is methylated by Set2 methyltransferase. This promotes nucleosome deacetylation by Rpd3S, which restores nucleosome stability.

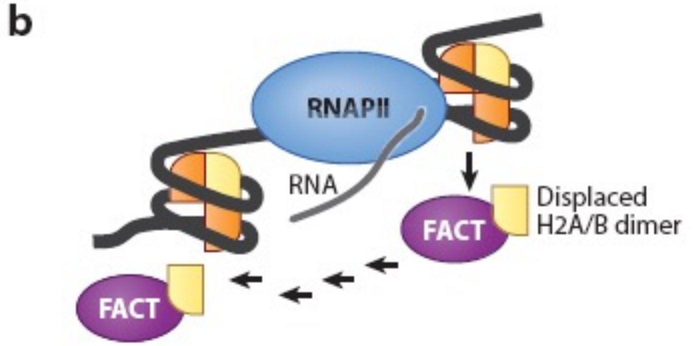
Multiple elongating polymerases displace histones and overcome nucleosomal barrier.

Workman, 2006, Gene Dev

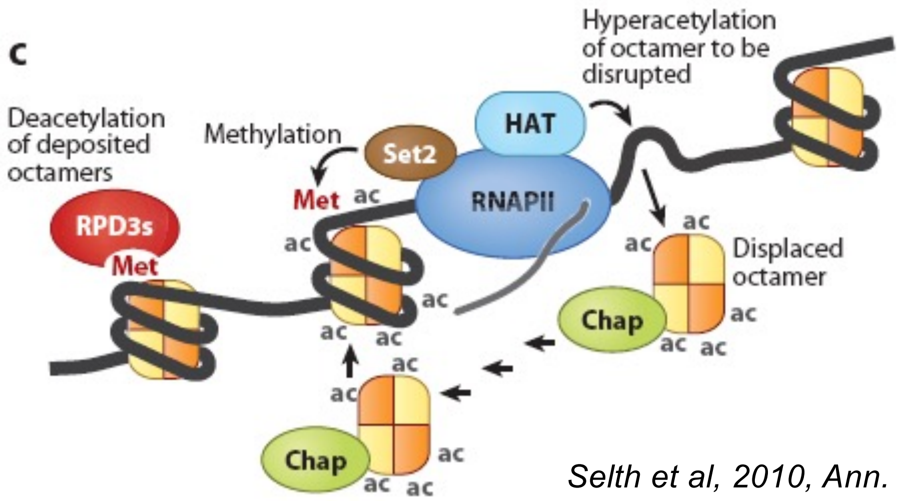
Pol II transcription through nucleosomes



Transcription through nucleosomes dislocates histone proteins to histone chaperones.



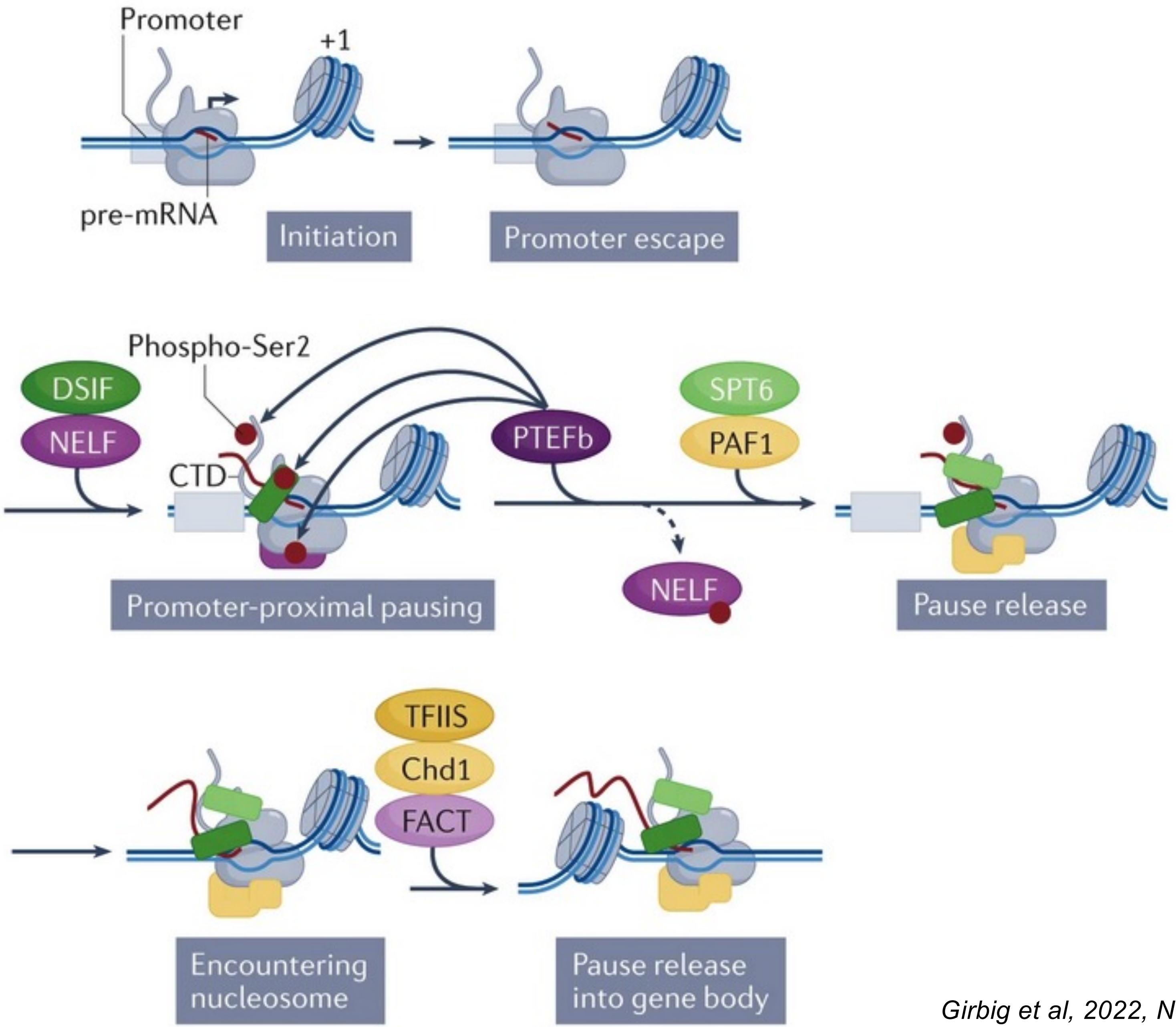
Progression of Pol II may occur without complete displacement of histone proteins. Only H2A/H2B is reloaded by FACT *(FAcilitates Chromatin Transcription)* **downstream of Pol II.**



Nucleosomes in front of Pol II are acetylated by HATs and displaced to Spt6/FACT chaperones, which reassemble nucleosomes behind Pol II.
H3 is methylated by Set2 methyltransferase.
This promotes nucleosome deacetylation by Rpd3S, which restores nucleosome stability.

Selth et al, 2010, Ann. Rev. Biochem.

Pol II transcription cycle

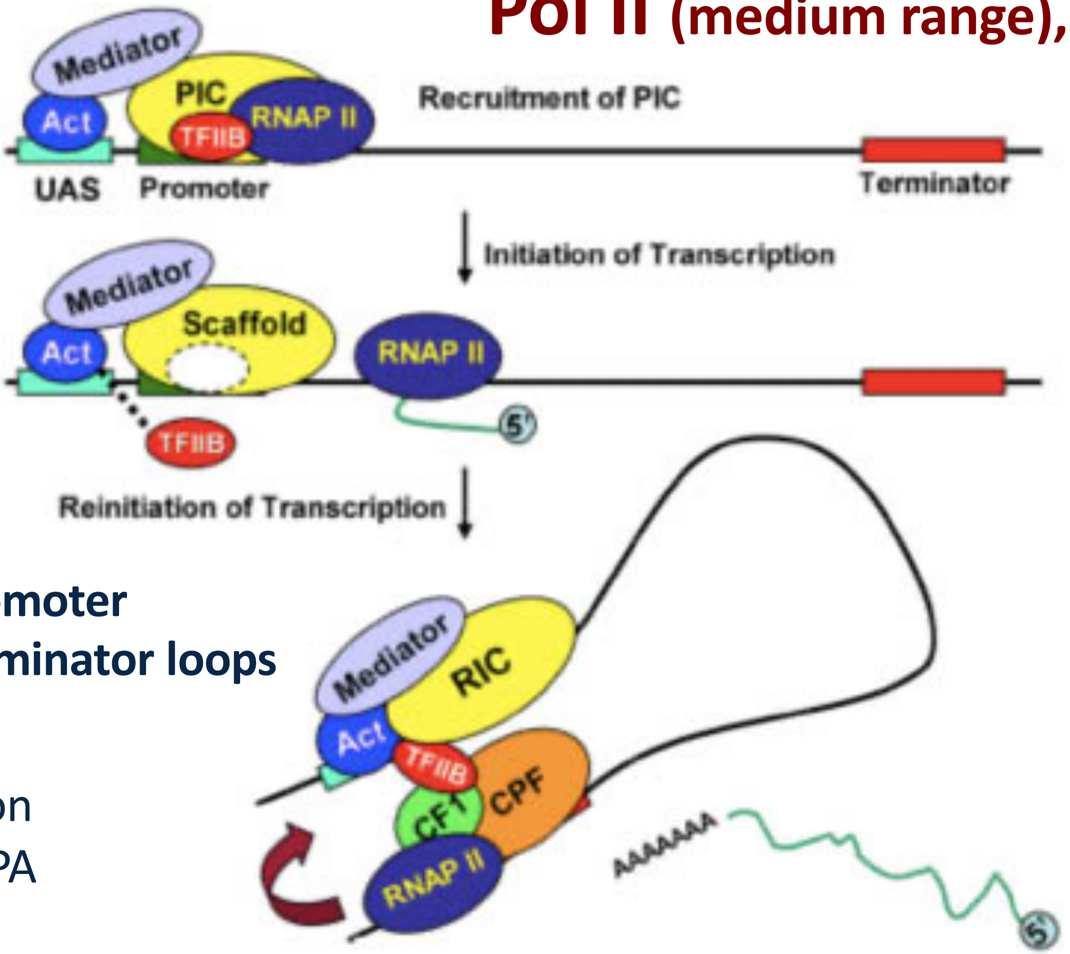


Girbig et al, 2022, Nat Rev Mol Cell Biol

Gene looping in transcription

Pol II (medium range), also Pol I

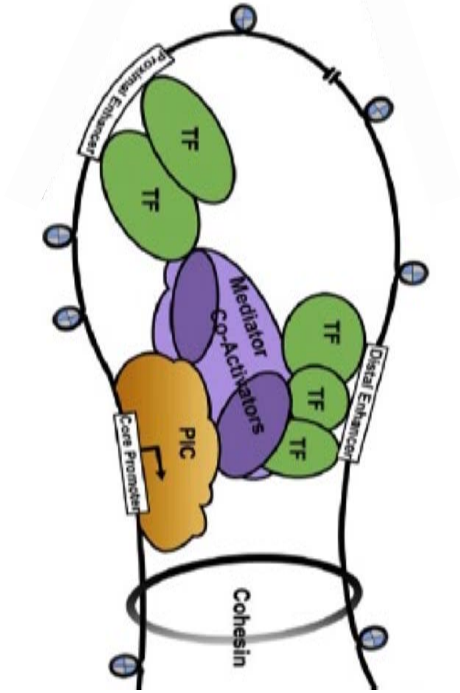
activated transcription



promoter terminator loops

Scaffold = transcription factors + CPA

Promoter-enhancer loops



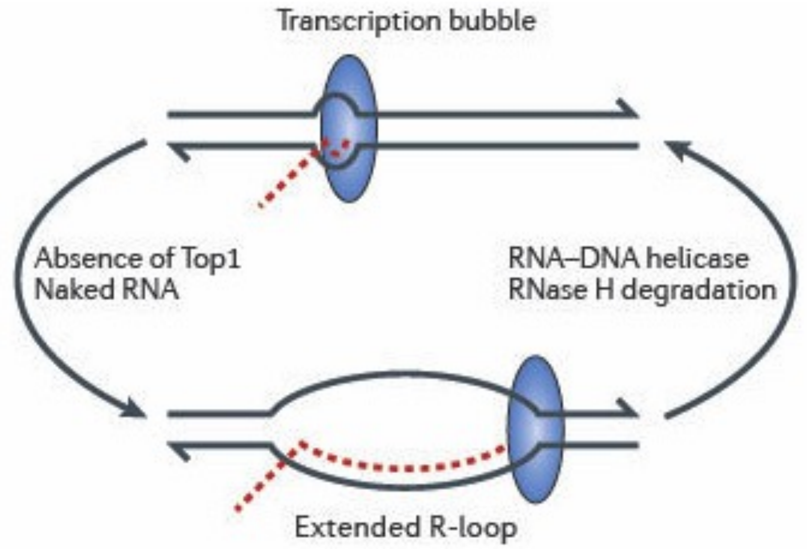
Loop formation requires interaction between factors at the promoter (THIIB) and terminator (Rna15 from CF1, Cleavage and Polyadenylation complex) /in mammals: transcription factors, nuclear receptors, insulators, chromatin remodellers, Polycombs/

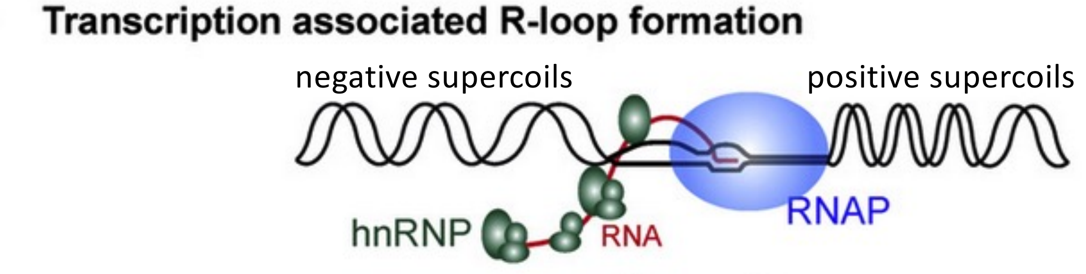
Loop function: facilitation of transcription reinitiation of Pol II, but also repression of gene expression (PcG, DNA methylation)

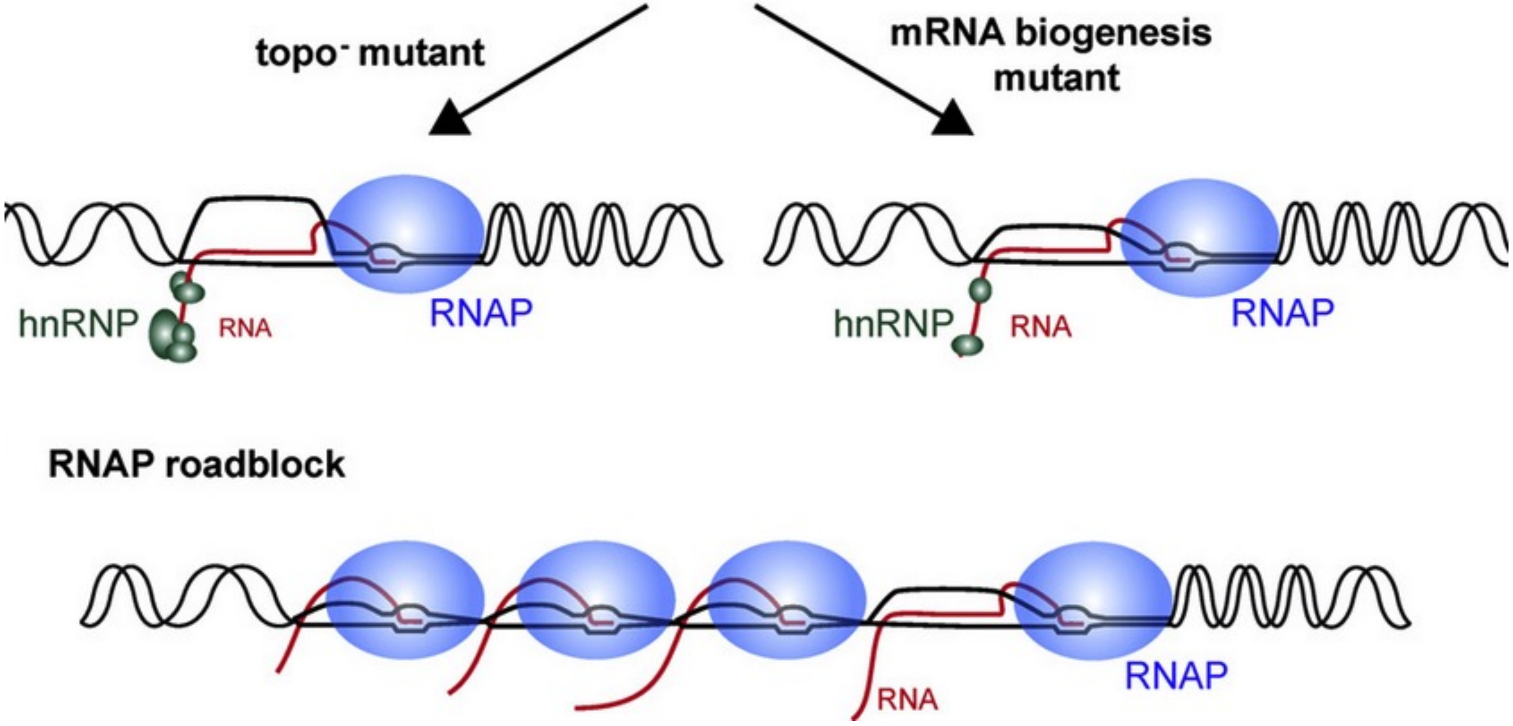
Ansari and Hampsay, GDev, 2005; El Kaderi et al., JBC, 2009

R-loops in transcription

DNA::RNA hybrids forming during transcription before RNP packaging.
Accumulate in topoisomerase or RNA biogenesis mutants (*tho*, *sen1*, mRNA export)







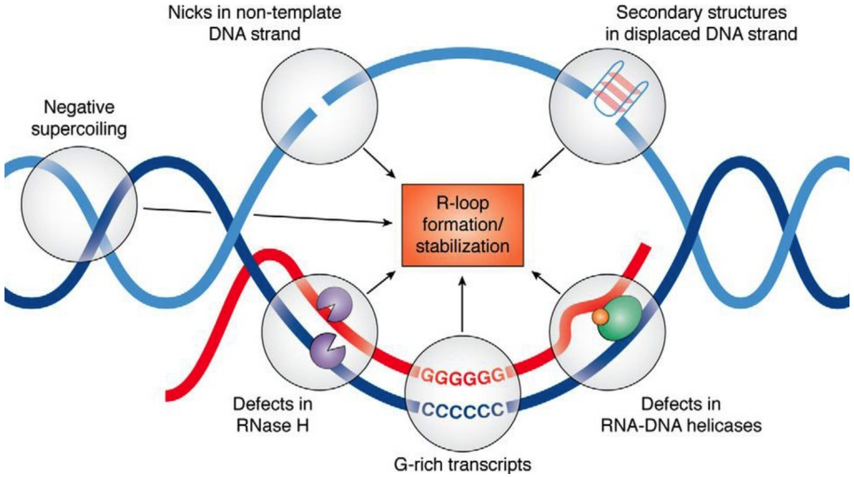
Aguilera and García-Muse, *Mol Cell*, 2012

R-loops in transcription

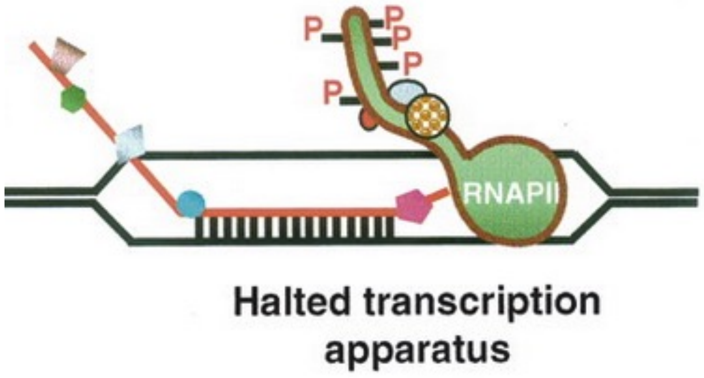
Promoting factors:

G-rich nascent RNA

Negative supercoling



- **Negative effects**
- polymerase stalling
- termination defects
- replication fork stalling
- DNA damage
- genetic instability



Preventing factors:

RNA binding proteins

Splicing factors

Cleavage and polyadenylation factors

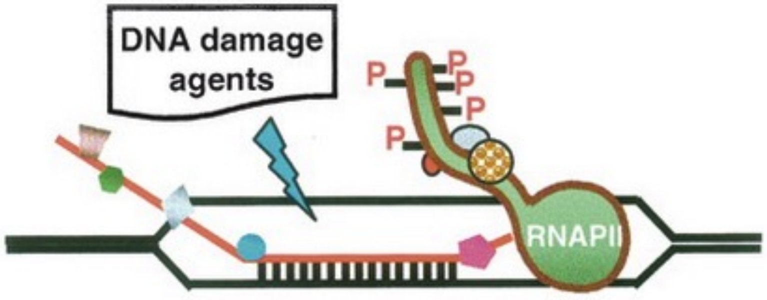
Transcription elongation factors

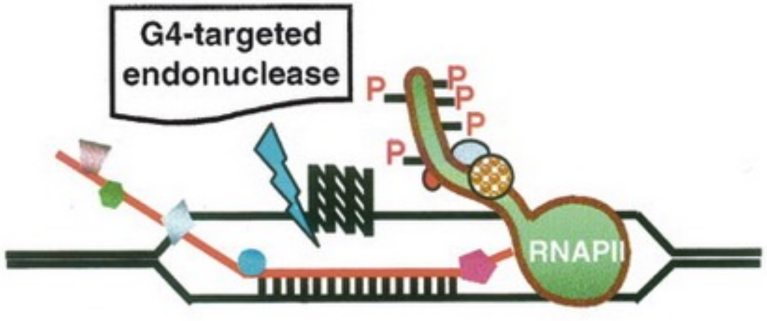
THO complex

helicase Sen1

Topoisomerases

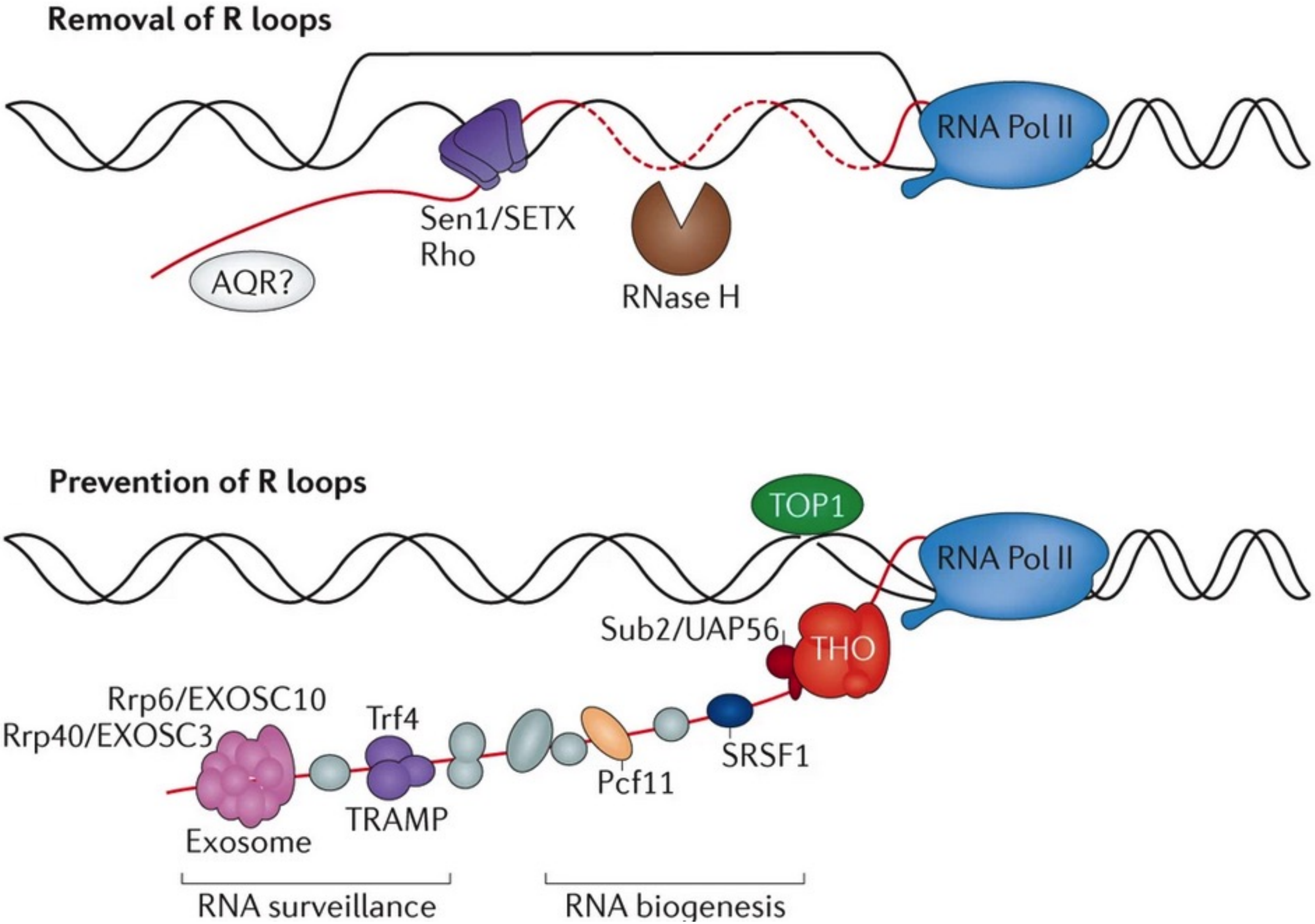
degradation by RNase H





Le and Manley, Gene Dev, 2005; Kim and Jinks-Robertson, 2012, Nat Rev Genet

R-LOOPs in transcription



Removal of R loops

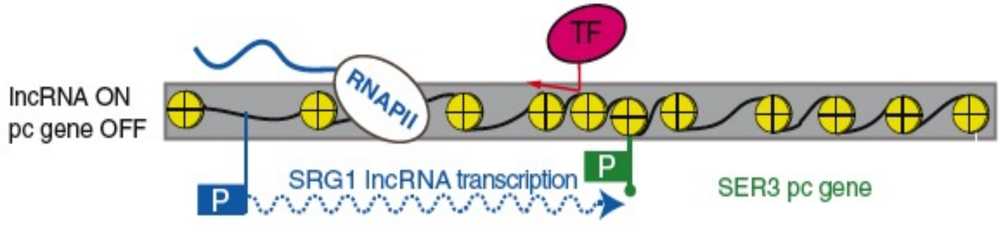
Prevention of R loops

Santos-Pereira and Aguillera 2015, Nat Re Genet

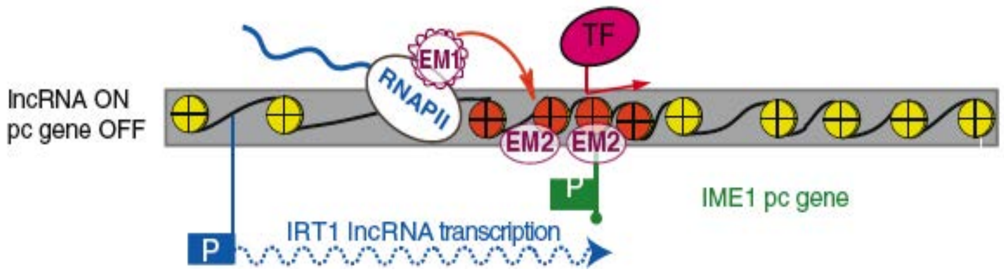
Regulation of Pol II transcription by ncRNAs

1. **RNA: siRNA-mediated TGS (transcriptional gene silencing)**
2. **lncRNA-mediated TGS via transcription itself**

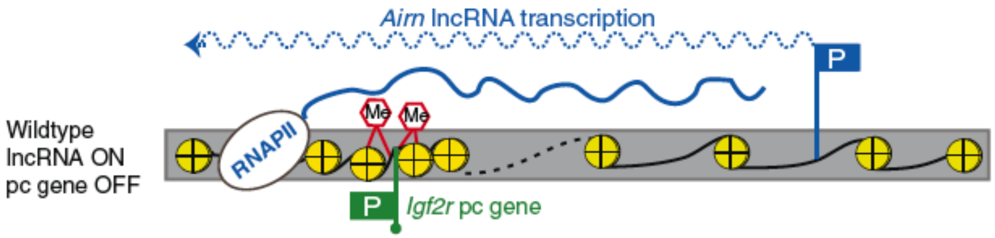
LncRNA transcription causes increased nucleosome density (yeast)



LncRNA transcription causes repressive histone modifications (yeast)



LncRNA transcription recruits DNA methylation at promoter (humans)

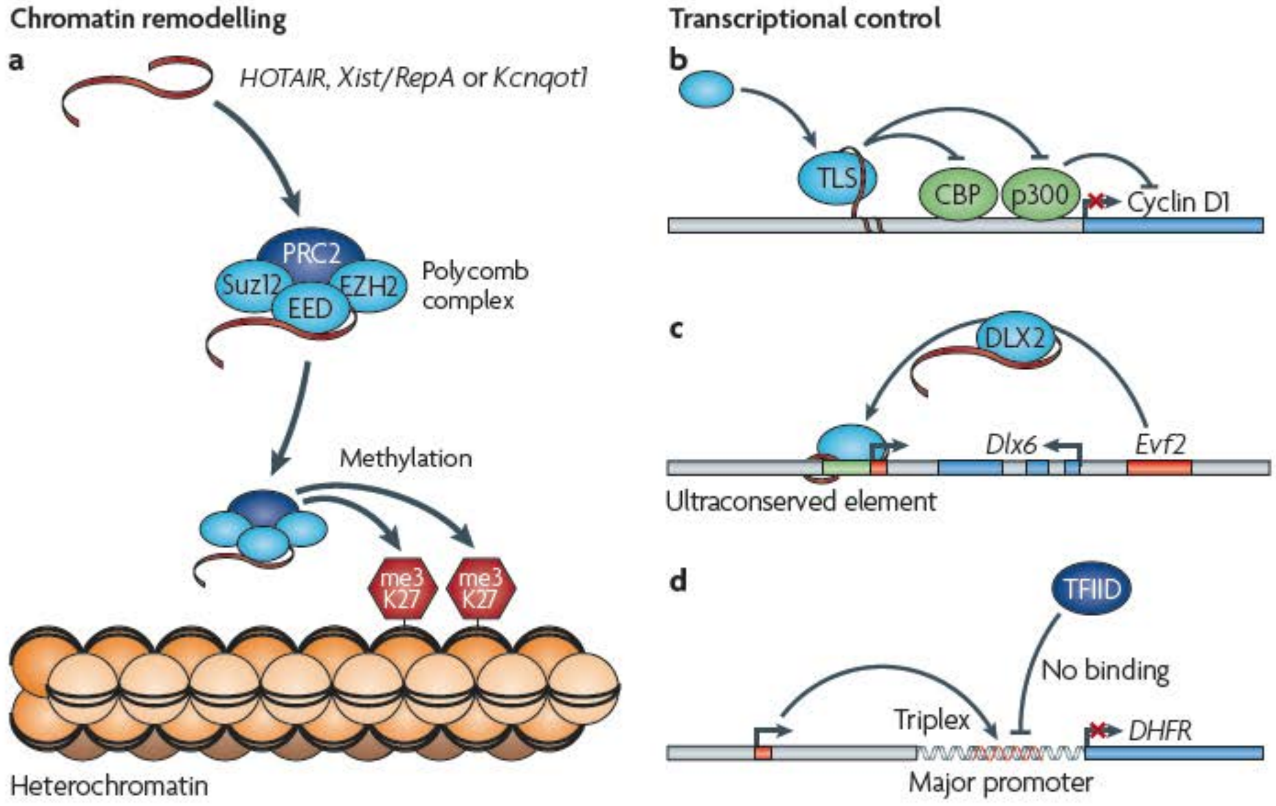


Kornienko et al, BMC Biol., 2013

Regulation of Pol II transcription by ncRNAs

lncRNA transcripts acting *in cis* or *in trans*:

TGS



recruit chromatin modifying complexes resulting in heterochromatin formation

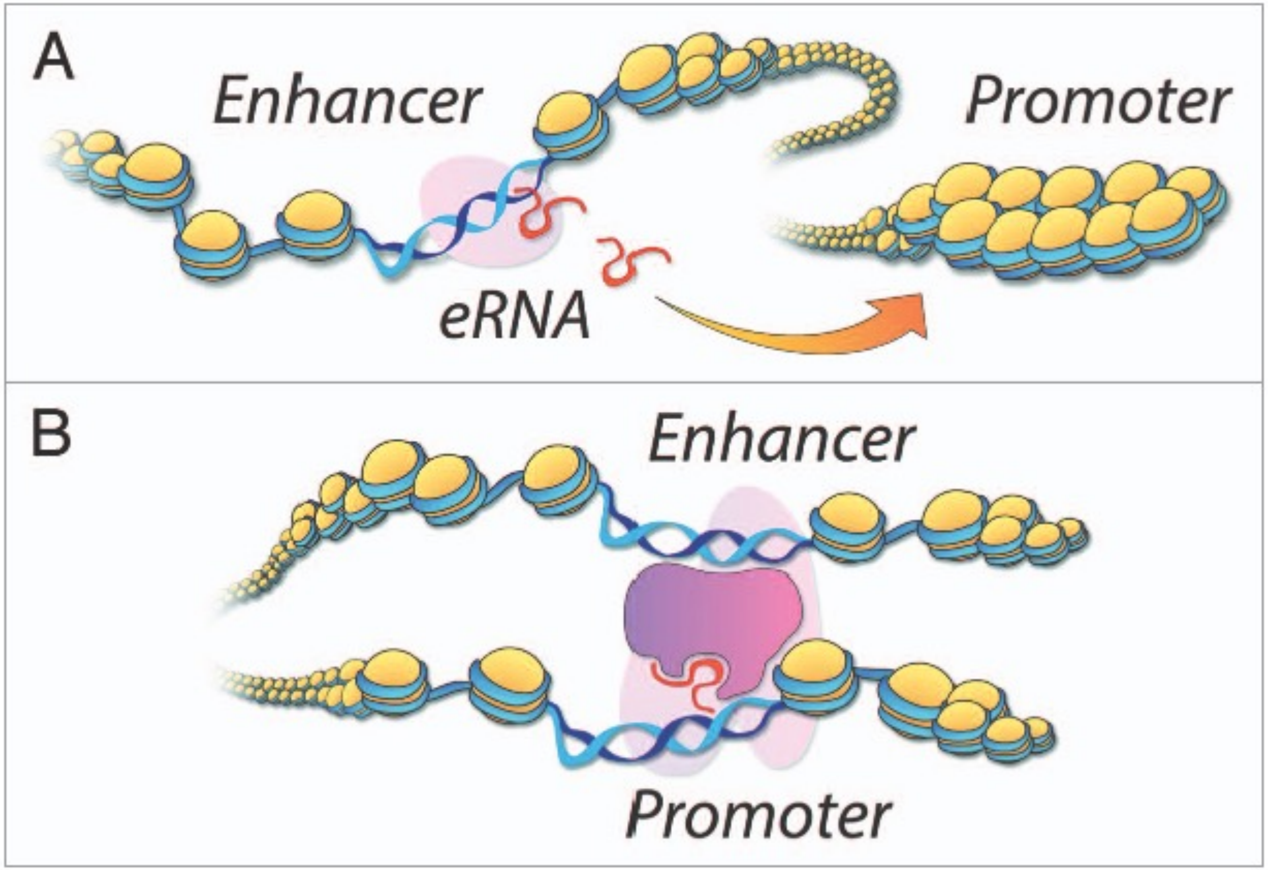
ncRNAs act as repressors or activators of transcription by binding to proteins or DNA

Mercer et al., Nat.Rev.Genet., 2007

Regulation of Pol II transcription by ncRNAs

eRNAs: enhancer ncRNAs transcribed from enhancer regions

Short but not always, some are up to 2 kb

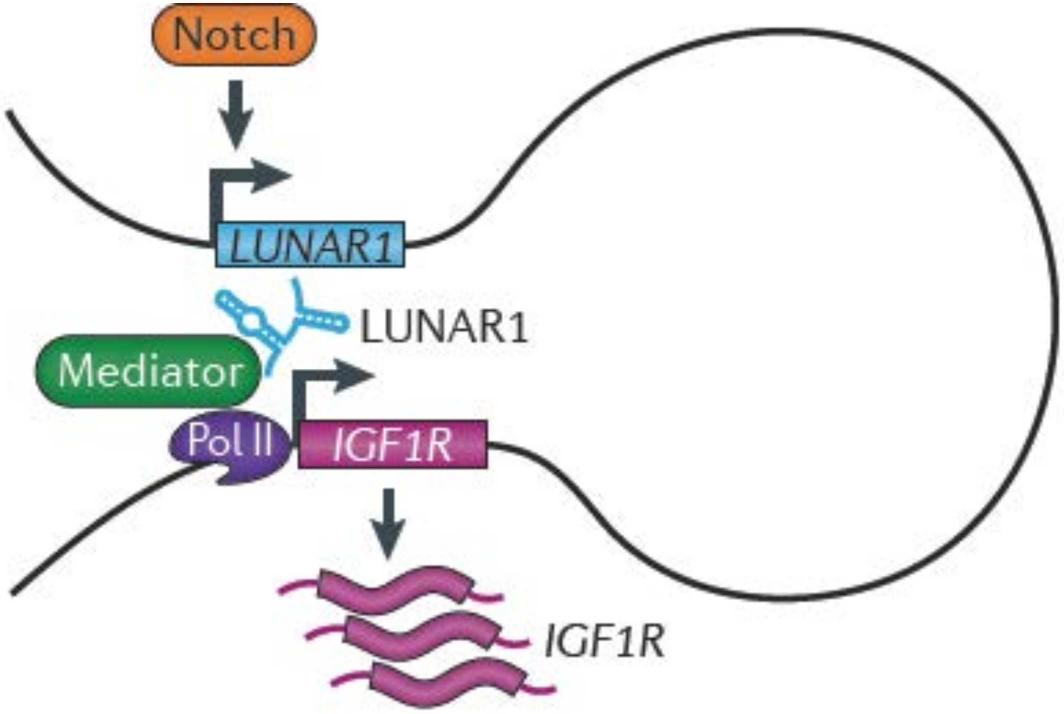


Mousavi et al., RNA Biol., 2014

- **eRNAs synthesized at enhancers are targeted to defined regulatory regions**
    - i.e. promoter (A)
- **eRNAs mediate chromatin accessibility and recruitment of factors for transcription and stabilization of enhancer-promoter contacts.**

Functions of eRNAs

Chromosome looping

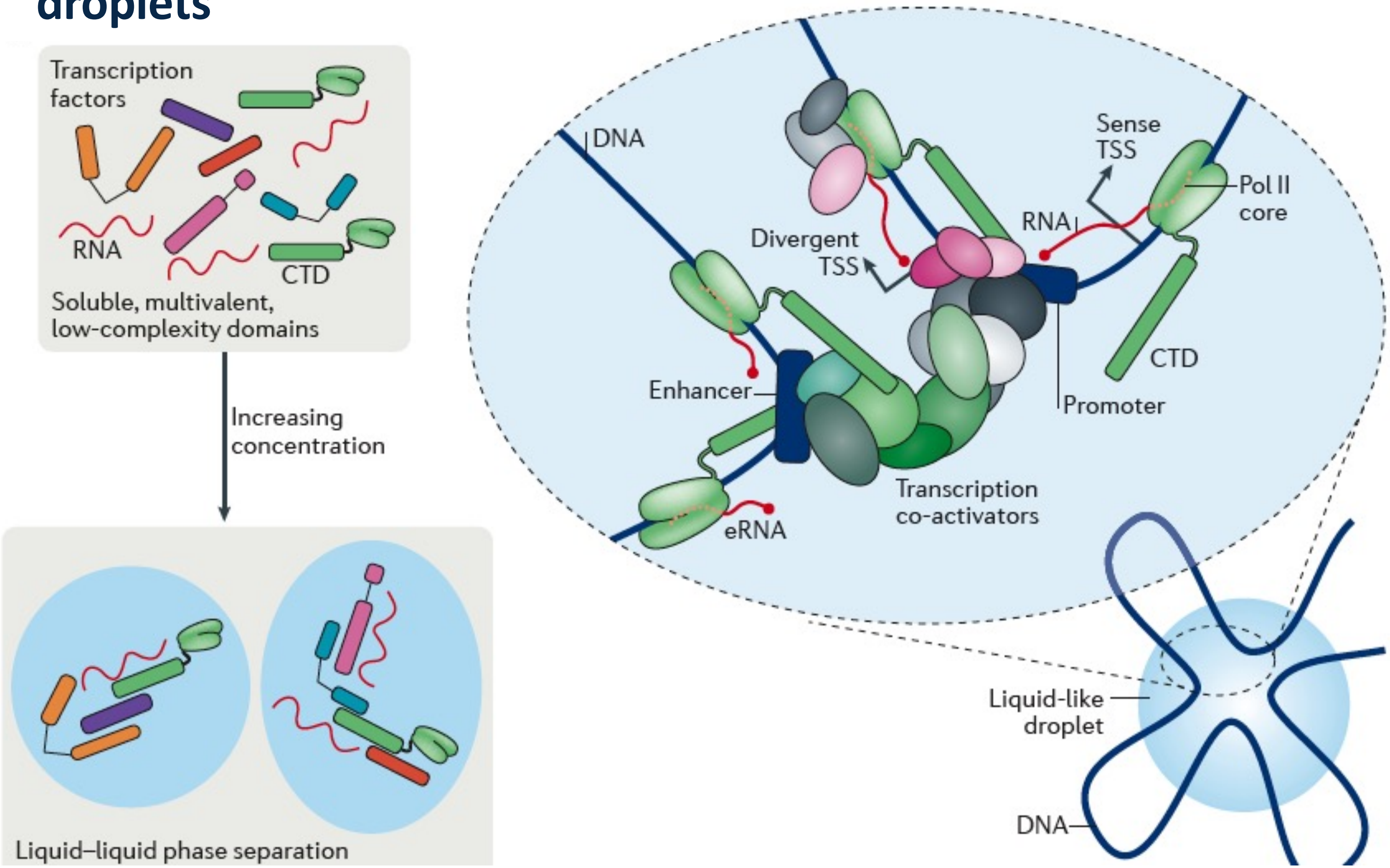


Some eRNAs (e.g. *LUNAR1* near the IGF1R locus) **mediate chromosome looping between enhancers and nearby genes via Mediator or MLL protein complexes**

Quinn and Chang, Nat Rev Genet 2015

Regulation of transcription by LLPS

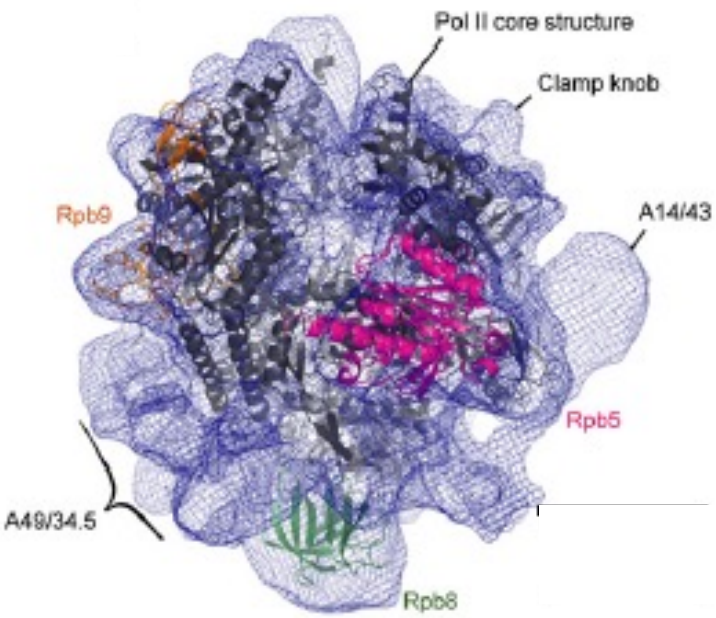
LLPS - liquid liquid phase separation
droplets



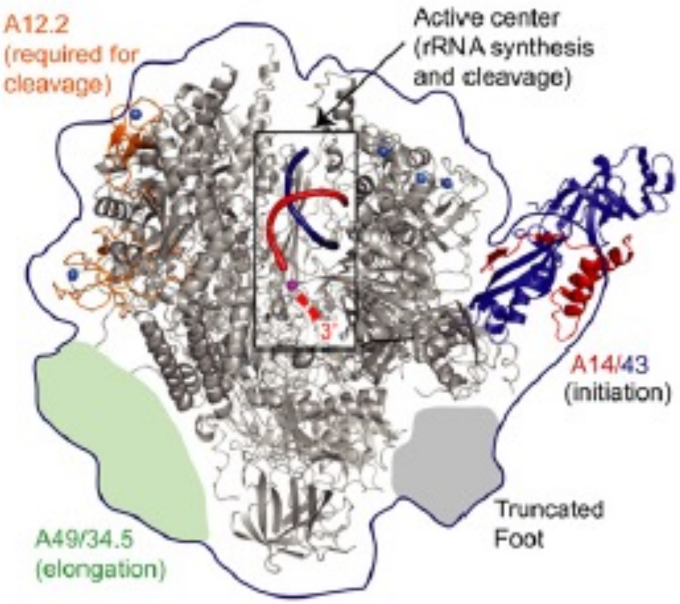
Harlan and Churchman, NatRevMolCellBiol, 2017

Unstructured, disordered or low complexity protein/RNA regions may form LLPS that can regulate different processes (e.g. transcription) by concentrating factors and enzymes (Pol II, TFs, RNA)

Yeast Pol I

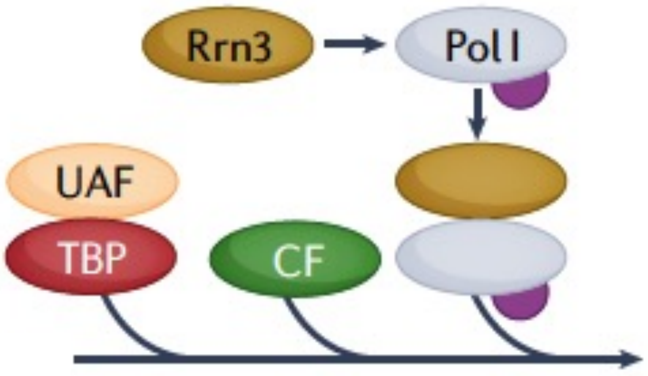


Kuhn et al, Cell, 2007 (Cramer's lab)



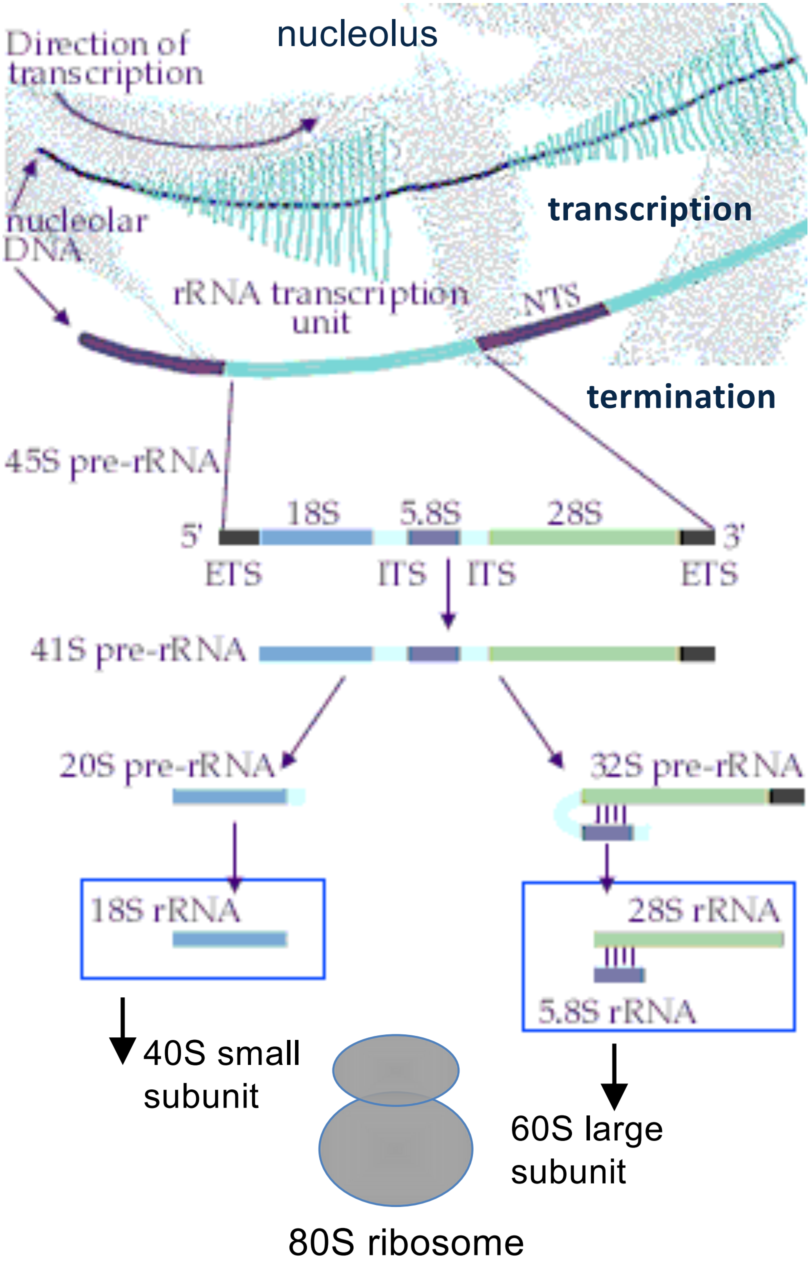
Girbig et al, 2022, Nat Rev Mol Cell Biol

- **14 subunits**
- **core, specific subunits A190, A135, AC40, AC19, A12.2**
- **Rpb5-6, 8, 10 and 12 - shared by Pol I-III**
- **specific subcomplexes A14/A43 and A49/A34.5**
- **no CTD**
- **intrinsic 3' RNA cleavage activity (A12.2/Rpa12) - possible roles in proofreading and transcription termination**



Pol I transcription

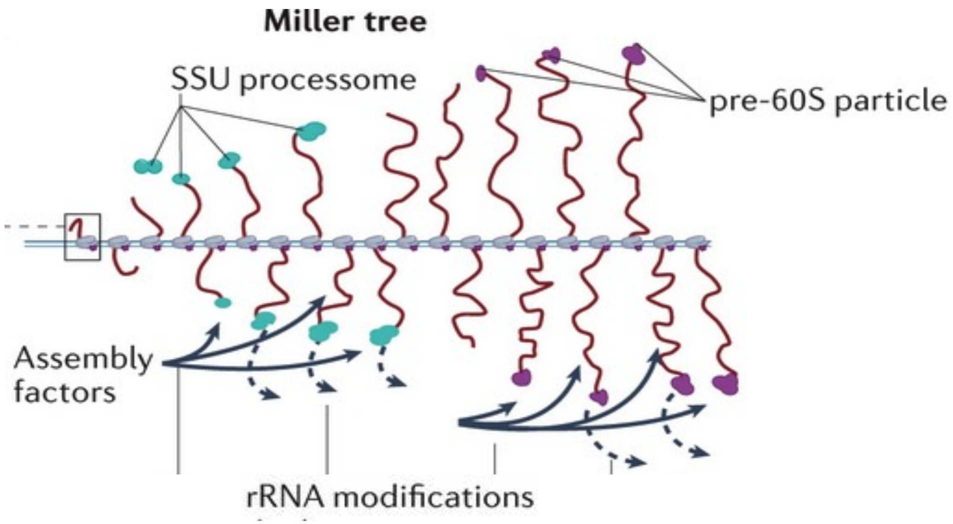
rRNA synthesis in the nucleolus

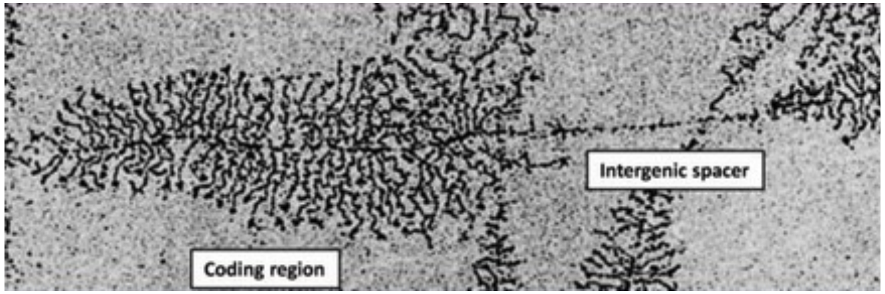


pre-rRNA processing

RNP assembly

rRNA modification
(2'-O methylation, pseudourydylation)

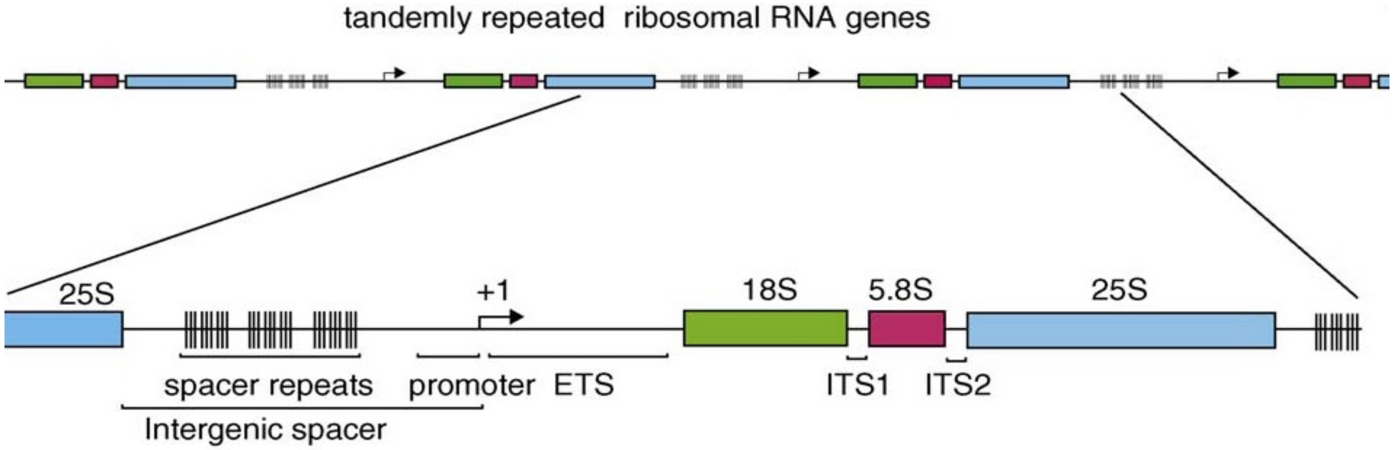




Grummt, FEBS J., 2010; Girbig et al, 2022, Nat Rev Mol Cell Biol

Pol I transcription

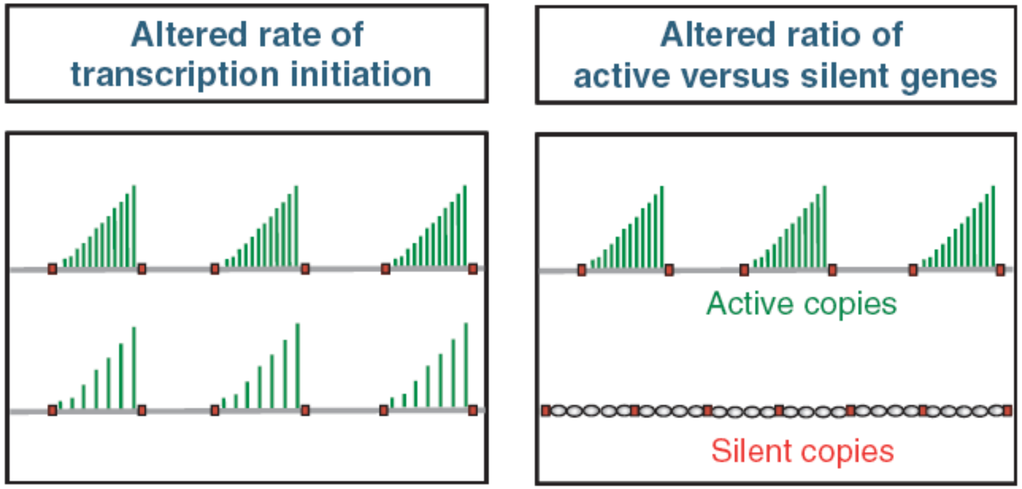
rRNA synthesis in the nucleolus



70-80% of cellular transcription is by Pol I to make rRNA

50% of Pol II transcription is for ribosomal protein genes

rDNA transcription units are arranged in tandem repeats in 150-200 copies

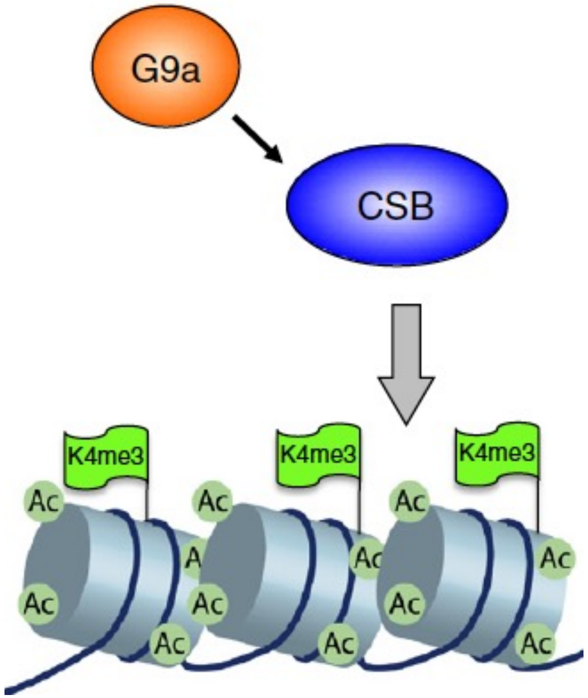


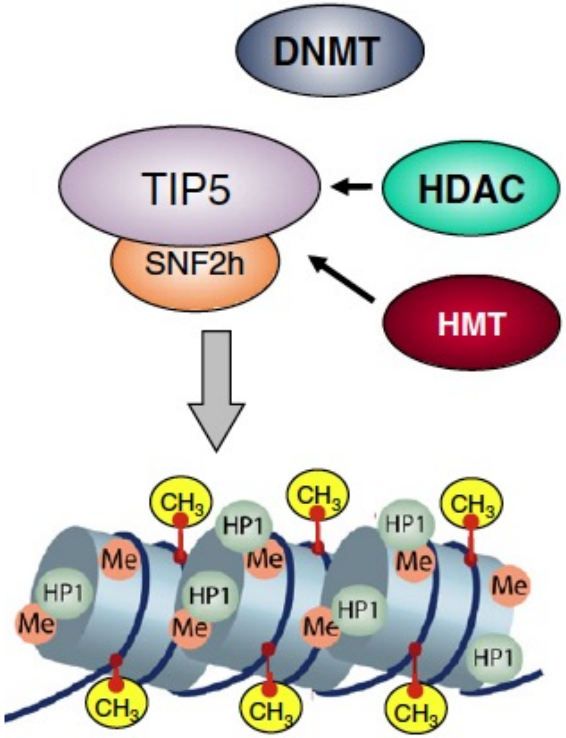
Regulation of rDNA transcription

Grummt, FEBS J., 2010;
Grummt and Längst, BBA,2013

rDNA silencing by NoRC

metazoa





Active rDNA promoter
DNA unmethylated
H3 and H4 acetylated, H3K4me3

Silent rDNA promoter (heterochromatic)
DNA methylated
H3 and H4 deacetylated, H3K9me

CSB activator
DNA-dependent ATPase

NoRC silencing complex (TIP5, SNF2)
Interacts with chromatin modifying enzymes
DNMT - DNA methyltransferase
HDAC - histone deacetylase
HMT -histone methyltransferase

Grummt and Längst, BBA,2013

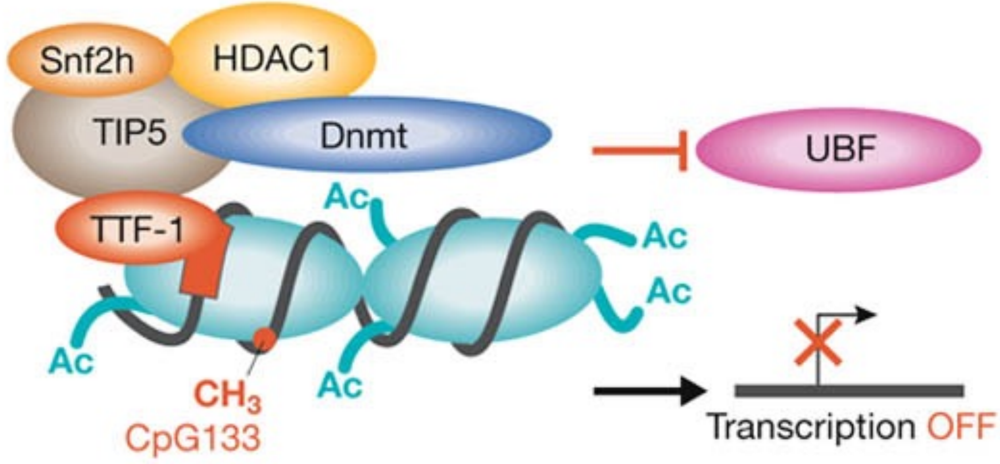
rDNA silencing by NoRC

metazoa

NoRC – mammalian nucleolar remodeling complex

Establishes and maintains heterochromatic state at promoters of silent rDNA repeats
(histone modifications and CpG methylation)

TTF-1 recruits NoRC which interacts with HDAC1, DNMT1 and HTM that modify chromatin and inhibit binding of UBF, leading to silencing.



NoRC

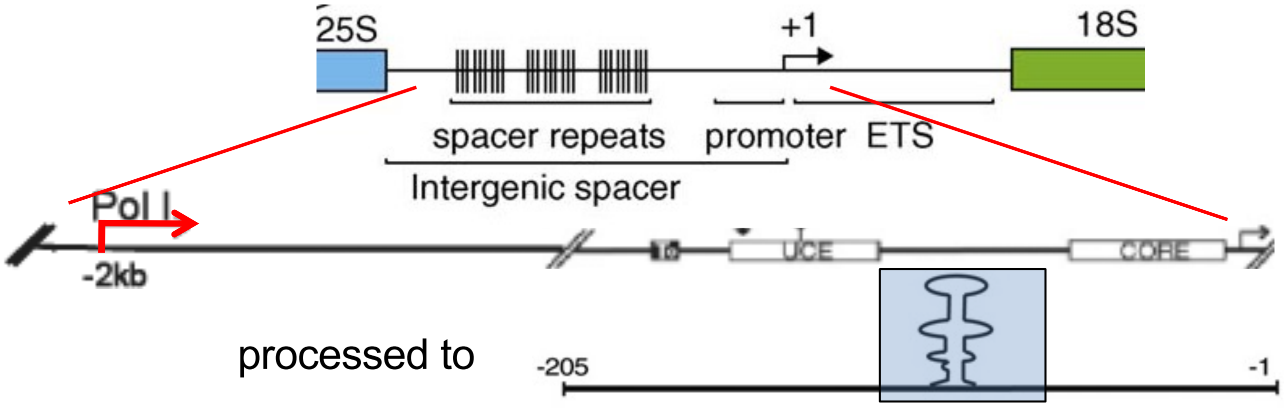
- **TIP5** TTF-I-interaction protein 5
- **SnF2** ATP-dependent chromatin remodeler

other

- **TTF-1** transcription factor I
- **UBF** upstream binding factor
- **DNMT** DNA methyltransferase
- **HDAC1** histone deacetylase

Matthews and Olsen, Embo Rep., 2006;
Tucker et al., Cur. Op. Cell. Biol., 2010

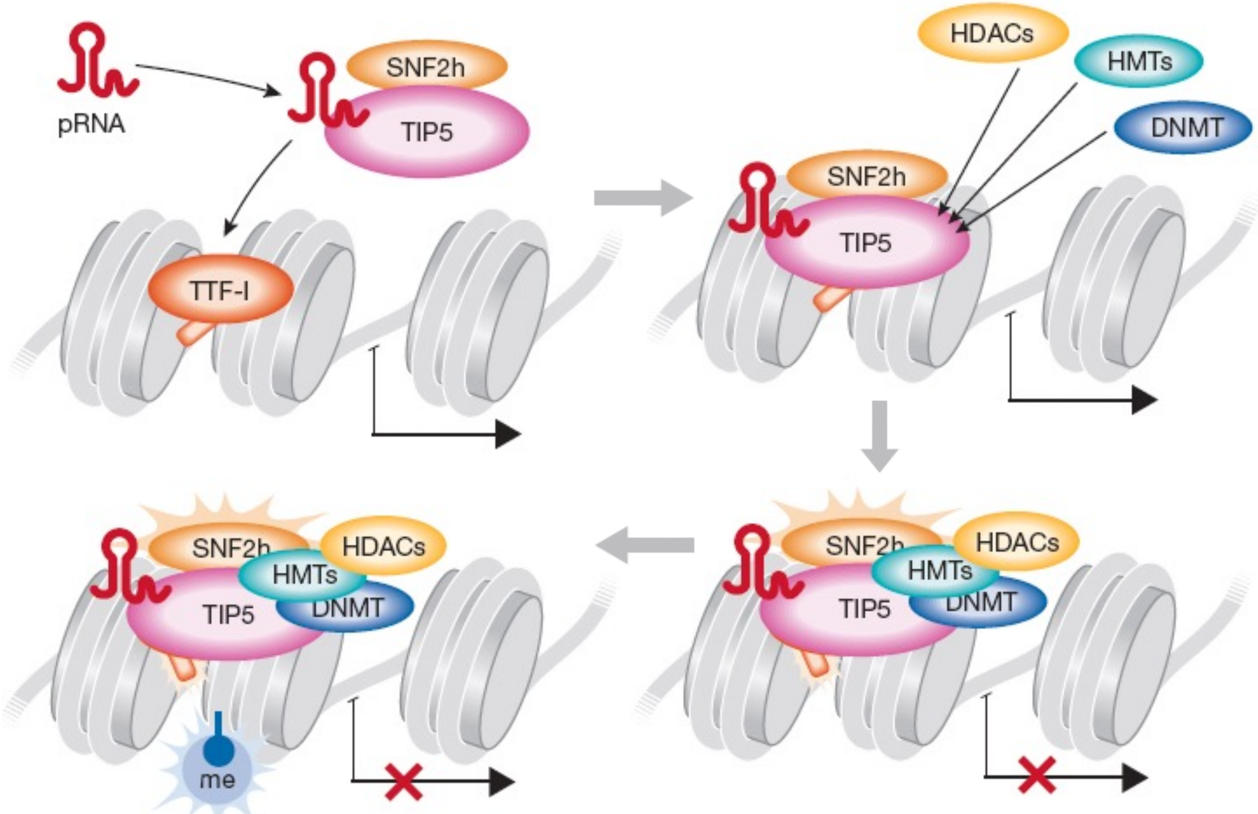
rDNA silencing by pRNA and NoRC



metazoa

Pol I intergenic transcript

pRNA

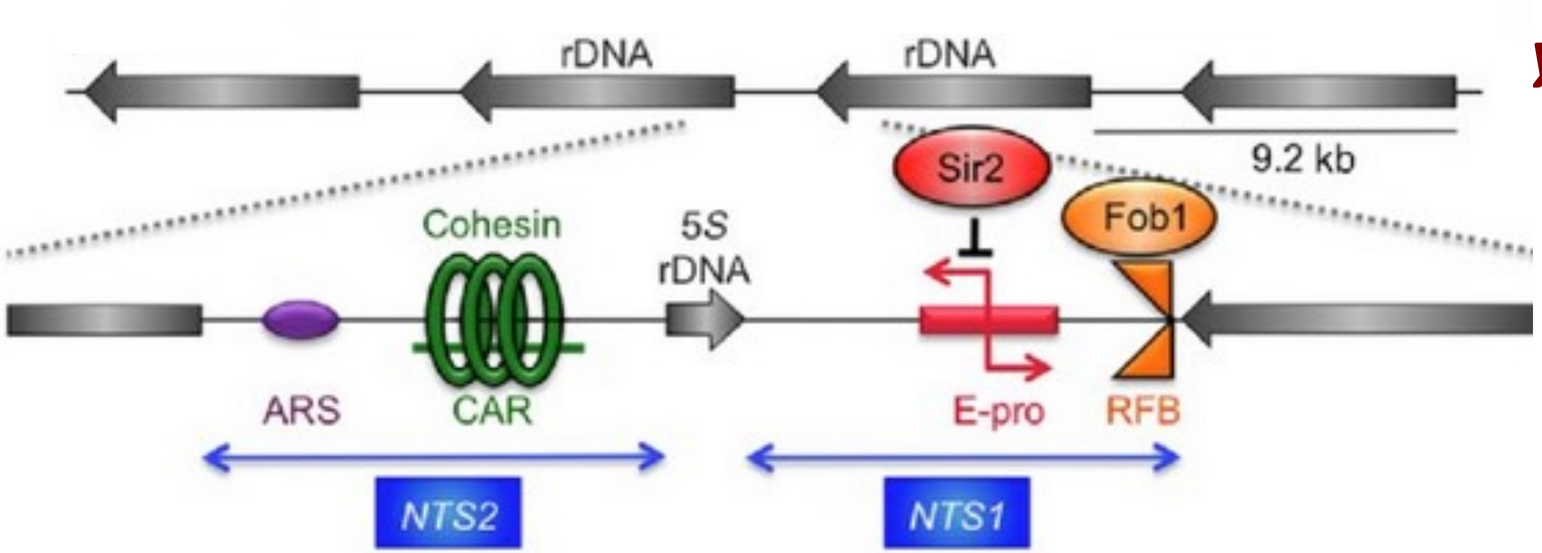


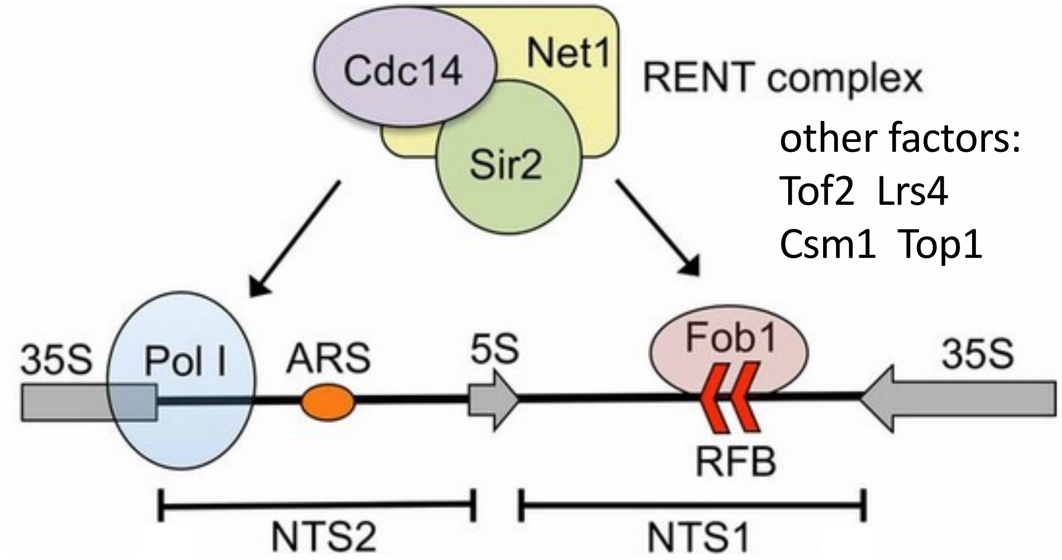
- **pRNA binds at T_0 to rDNA promoter, independently of TTF-I and other proteins, forming a triplex**
- **pRNA competes with TTF-I**
- **rDNA/pRNA triplex recruits methyltransferase DNMT3b**
- **this results in chromatin hypermethylation and rDNA silencing**

Stark and Taliansky, Embo Rep., 2008; Mayer et al., Mol. Cell, 2006; Embo Rep., 2008; Schmitz et al., Gene Dev., 2010

rDNA silencing by RENT and ncRNAs

yeast





other factors:
Tof2 Lrs4
Csm1 Top1

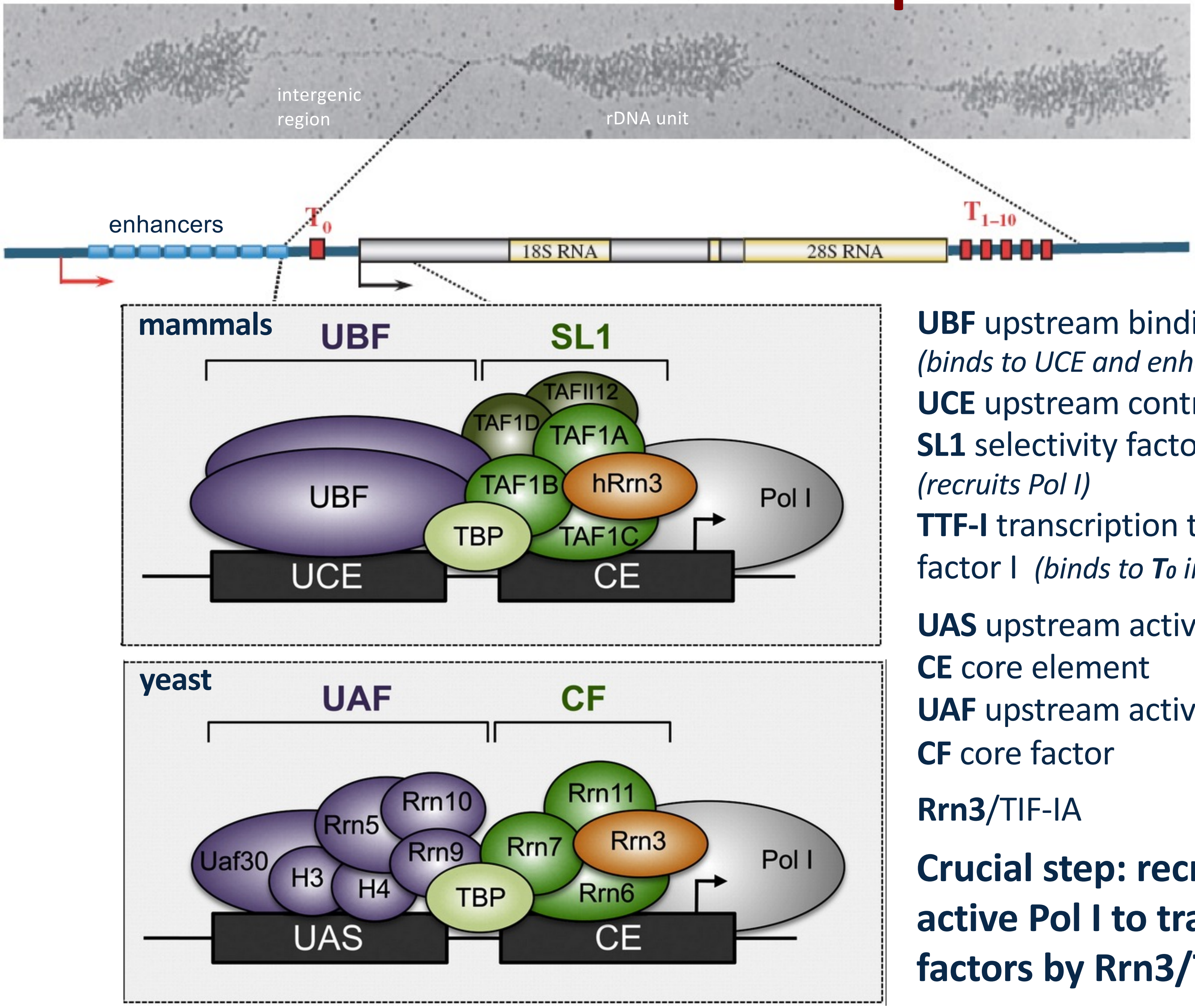
D'Alfonso et al, Sem Cell Dev Biol,2024

RENT is recruited to rDNA at the promoter and NTS1 by interacting with Pol I and Fob1

RENT silences Pol II transcription from E-pro through Sir2 histone deacetylase activity

If not silenced Pol II non-coding transcription displace cohesin and results in rDNA instability

Pol I transcription



UBF upstream binding factor *(binds to UCE and enhancers)*

UCE upstream control element

SL1 selectivity factor 1 *(recruits Pol I)*

TTF-I transcription termination factor I *(binds to* T_0 *in front of UCE)*

UAS upstream activation sequence

CE core element

UAF upstream activating factor

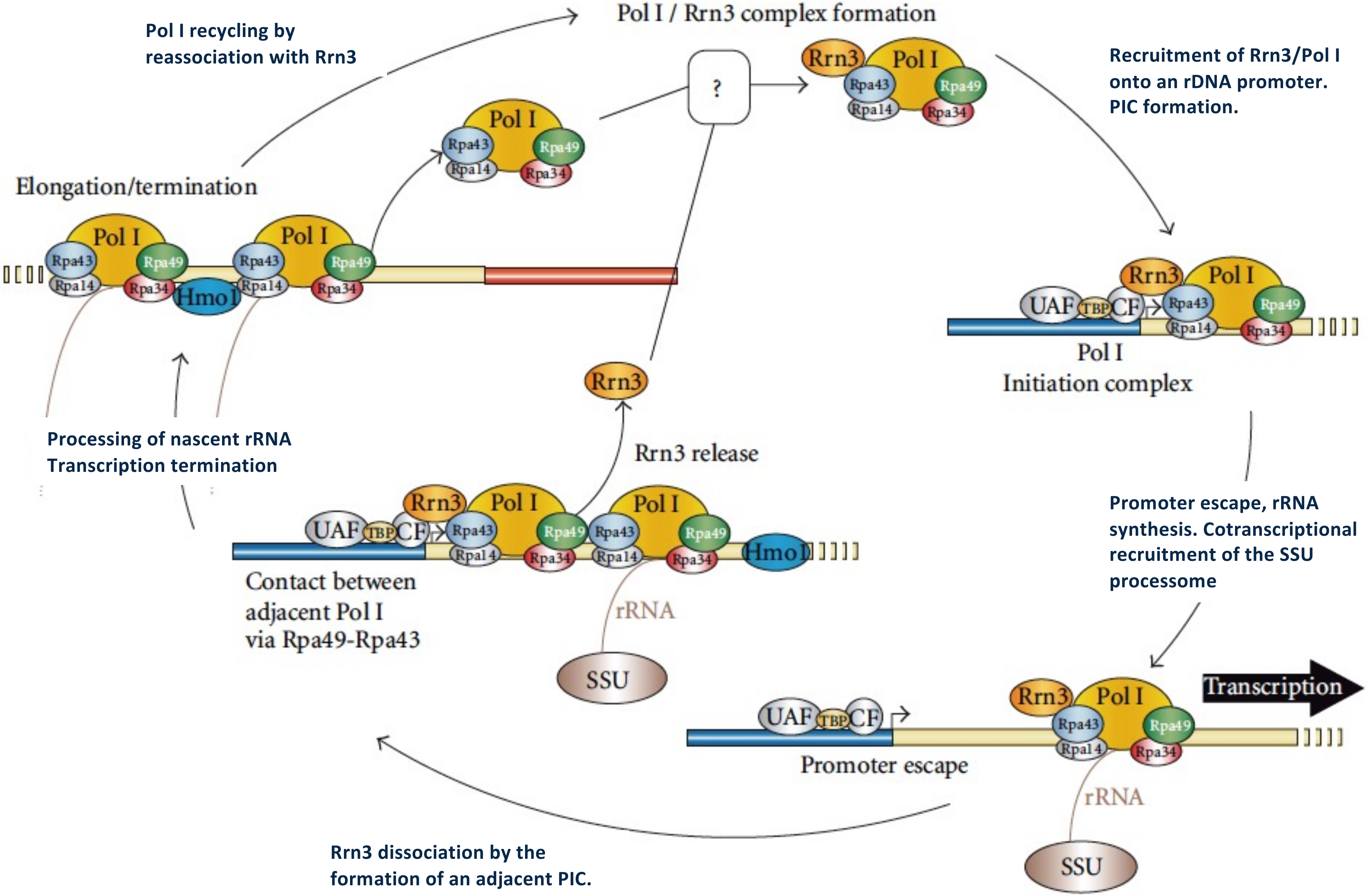
CF core factor

Rrn3/TIF-IA

Crucial step: recruitment of active Pol I to transcription factors by Rrn3/TIF-IA

Moss, Cur. Op. Gen. Dev., 2004; Grummt, FEBS J. 2010; Knutson and Hahn, BBA, 2012

Pol I transcription cycle



Pol I recycling by reassociation with Rrn3

Recruitment of Rrn3/Pol I onto an rDNA promoter. PIC formation.

Processing of nascent rRNA
Transcription termination

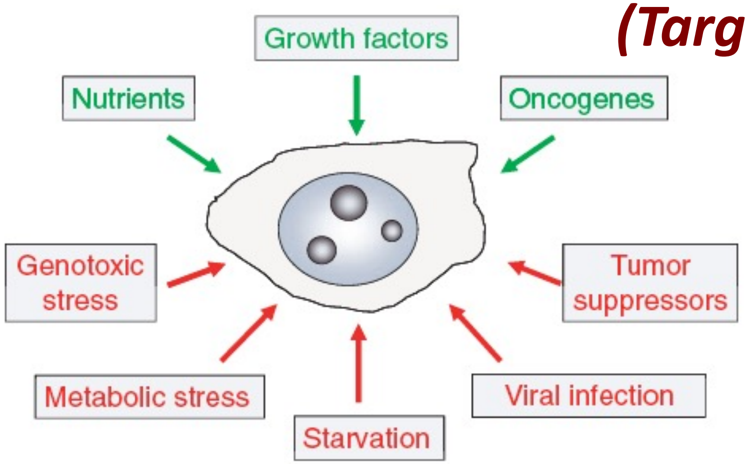
Promoter escape, rRNA synthesis. Cotranscriptional recruitment of the SSU processome

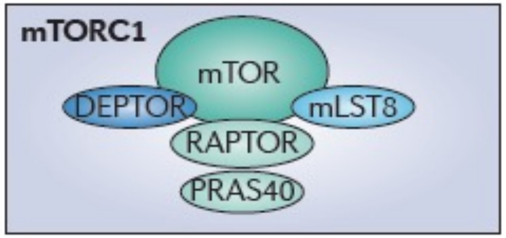
Rrn3 dissociation by the formation of an adjacent PIC.

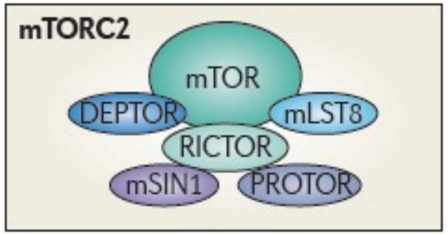
Albert et al, 2012, Genetics Res Int

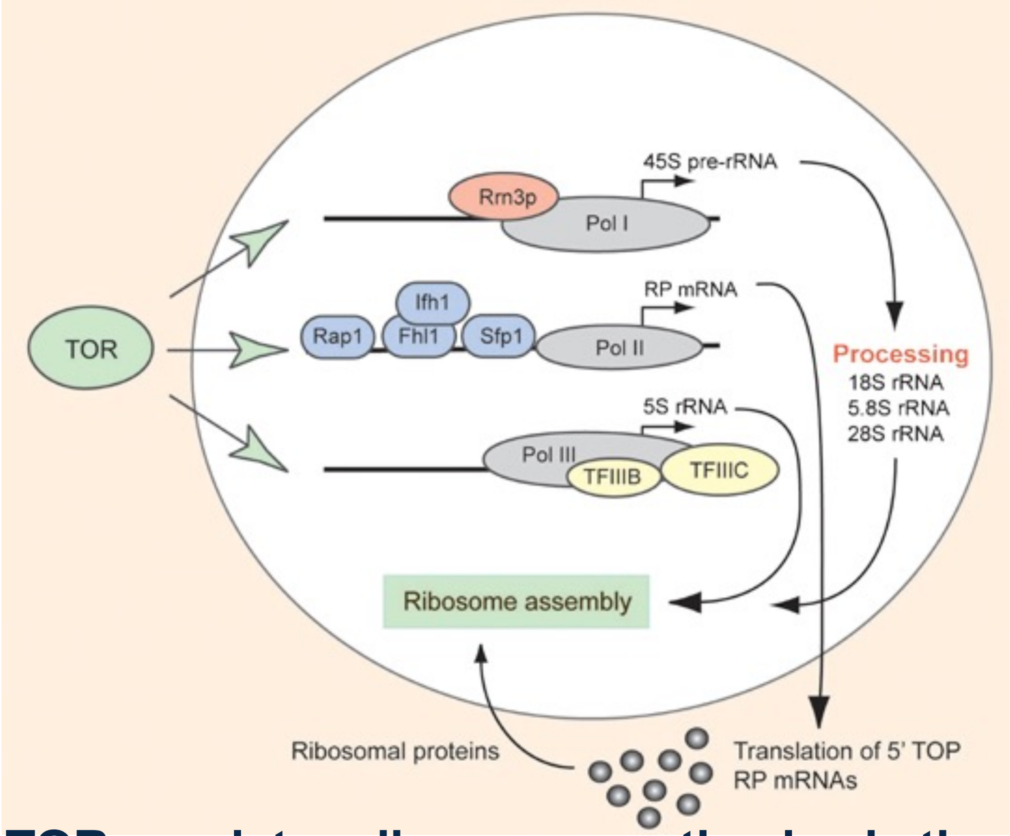
Regulation of rRNA synthesis – TORC kinase signaling

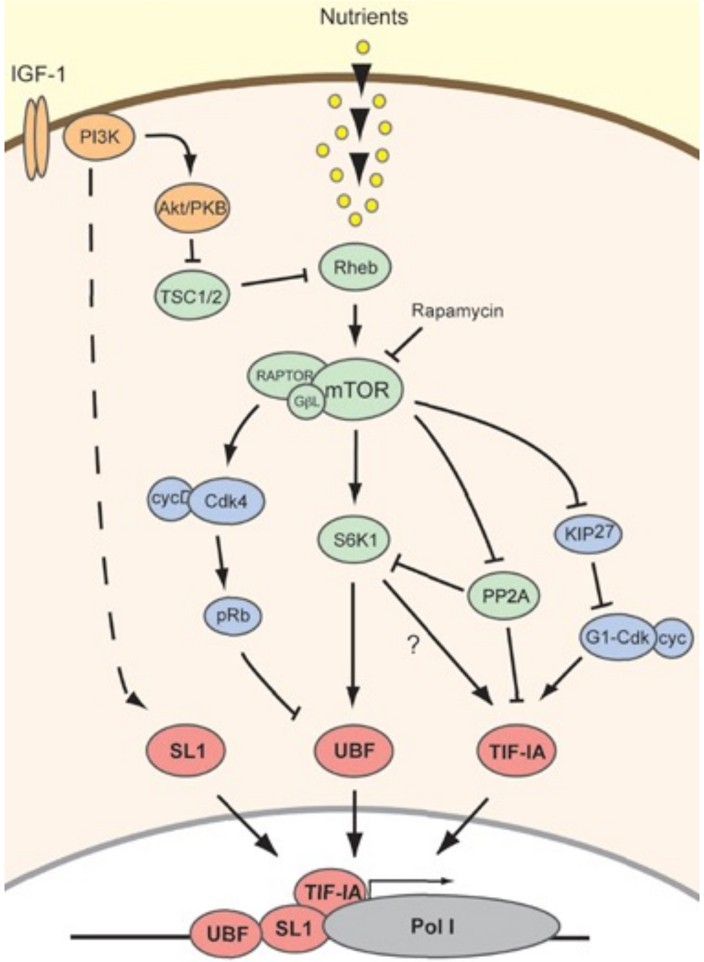
(Target of Rapamycin)







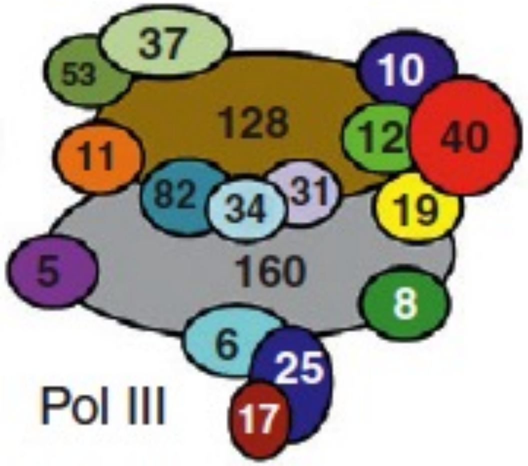




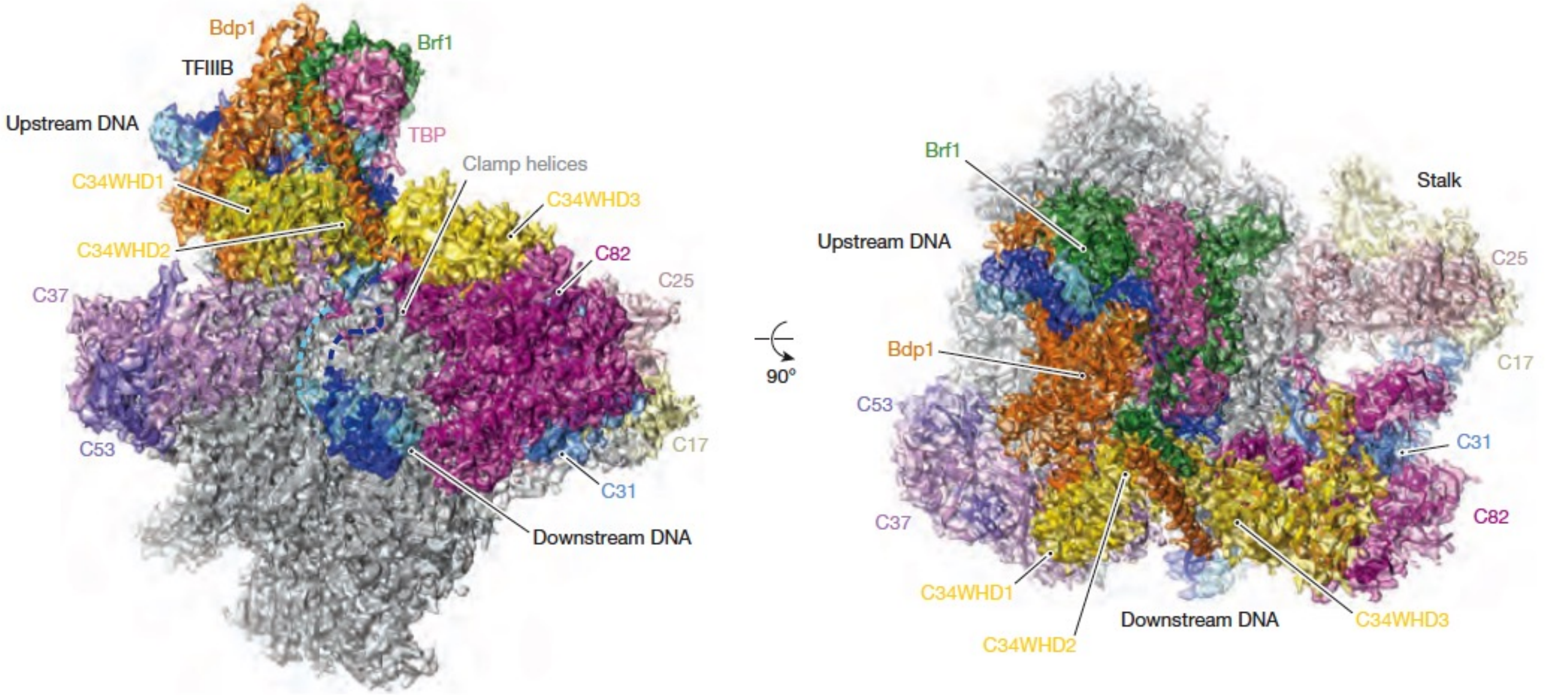
Mayer and Grummt, Oncogene, 2006

TOR regulates ribosome synthesis via three polymerases

Pol III



Pol III transcribes short genes
tRNA, 5S rRNA
U6 and U6atac snRNAs
RNase P RNA, RNase MRP RNA
7SK RNA, 7SL/SRP RNA
Y, SINEs, BC200, snaR, vault RNAs



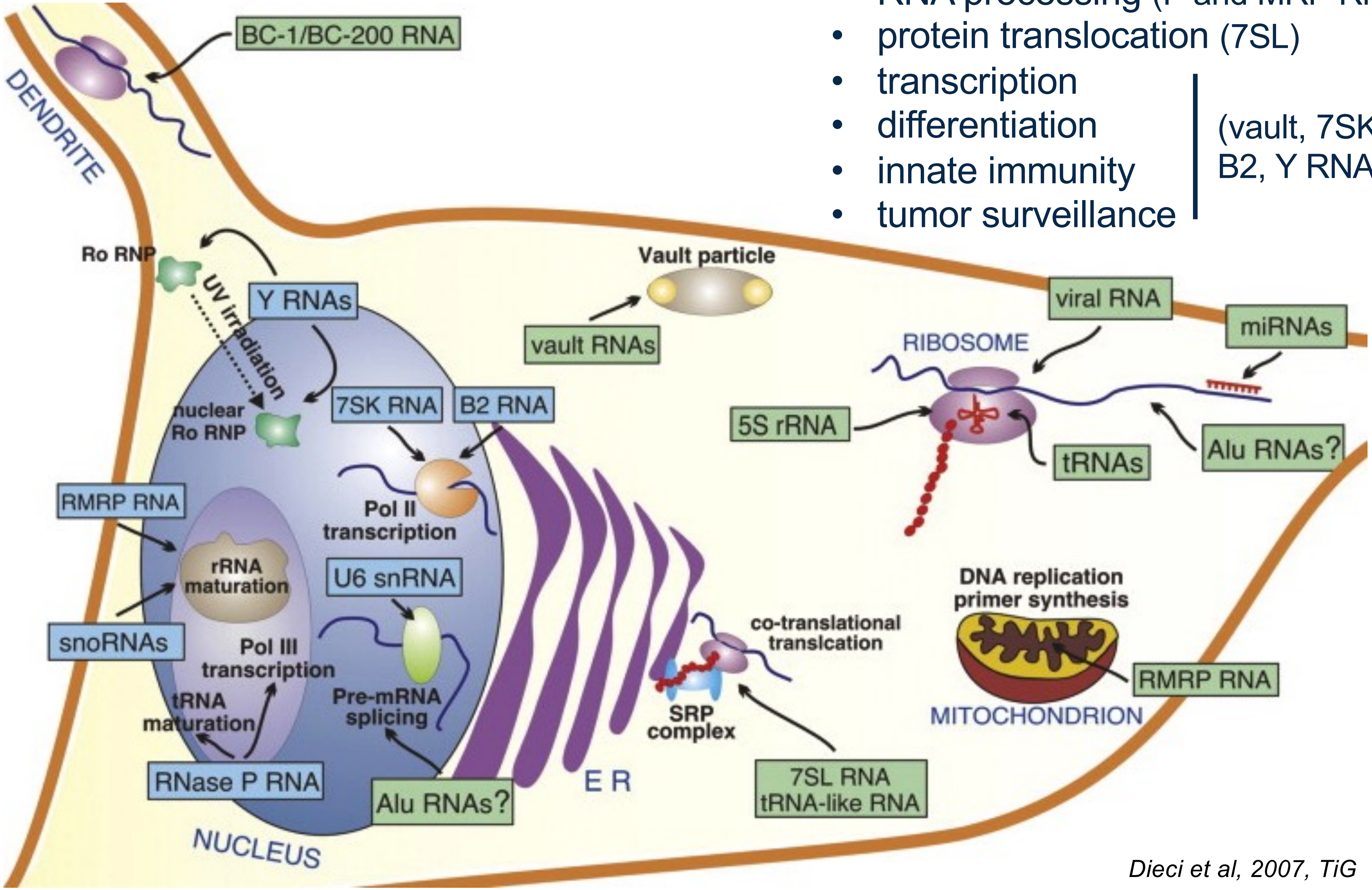
Wild and Cramer, 2012, TiBS; Abascal-Palacios et al, 2018, Nature

Pol III transcripts

Function of Pol III transcripts

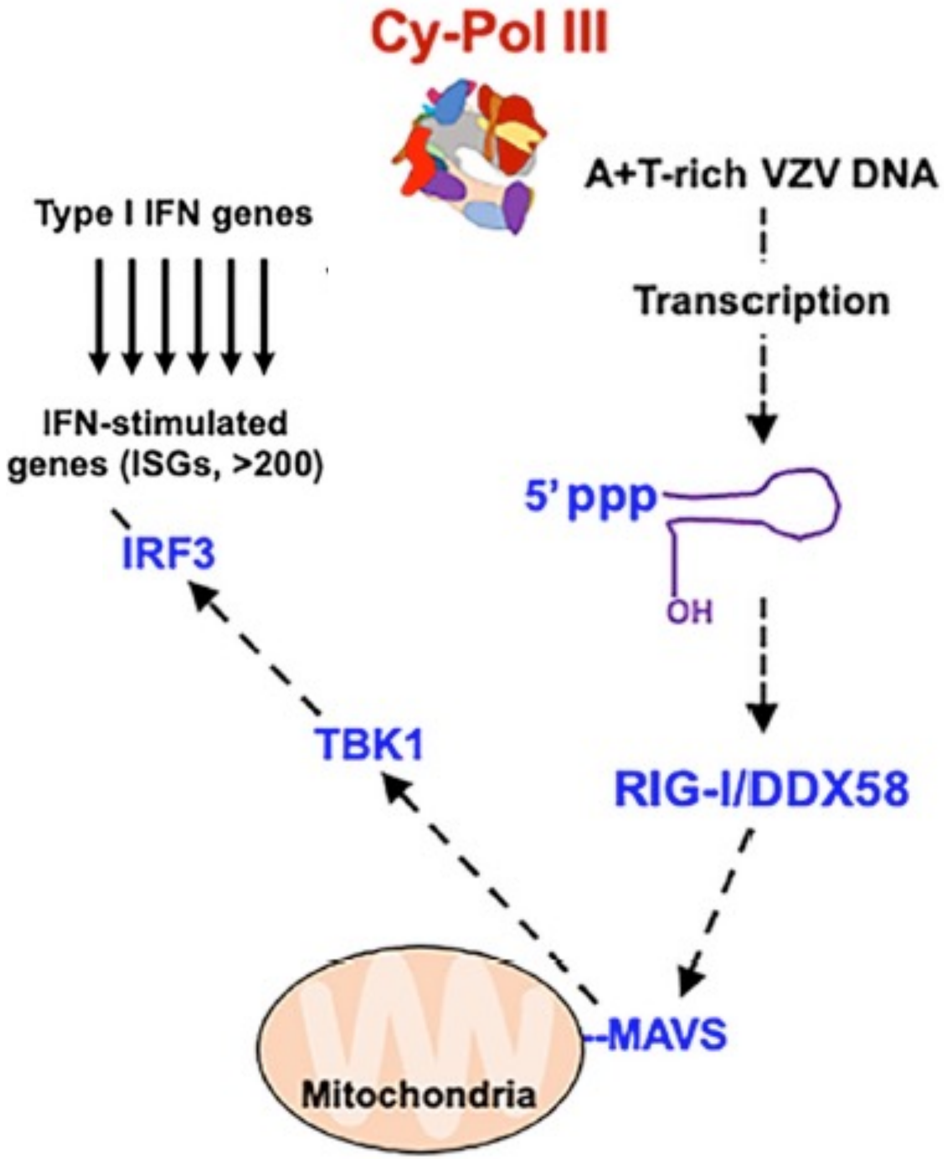
- translation (rRNA, tRNA)
- splicing (U6)
- RNA processing (P and MRP RNA)
- protein translocation (7SL)
- transcription
- differentiation
- innate immunity
- tumor surveillance

(vault, 7SK, B2, Y RNAs)



Dieci et al, 2007, TiG

Cytoplasmic Pol III



Cy-Pol III

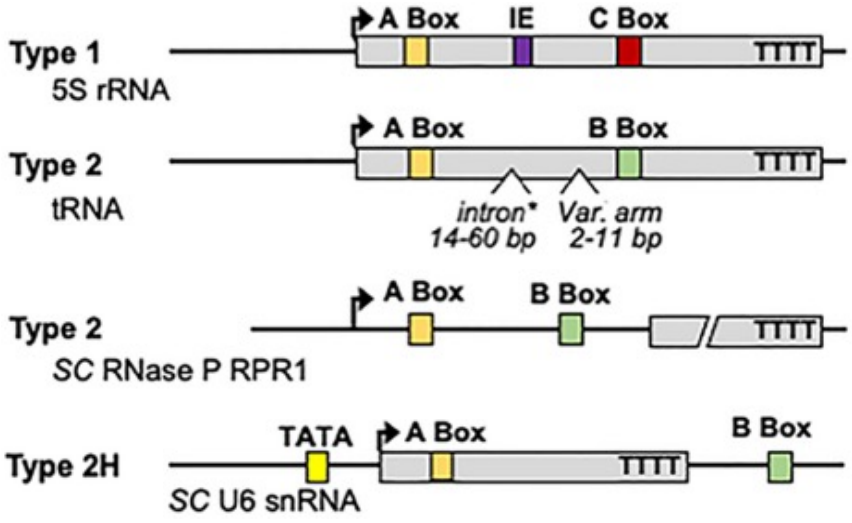
- Presumably promoter-independent
- No known TFs
- Transcribes A+T-rich DNA into a 5'ppp-RNA which is a RIG-I activating ligand
- RIG-I initiate a signaling cascade that leads to a type 1 interferon response (innane immunity)

Kessler and Maraia, NAR, 2021

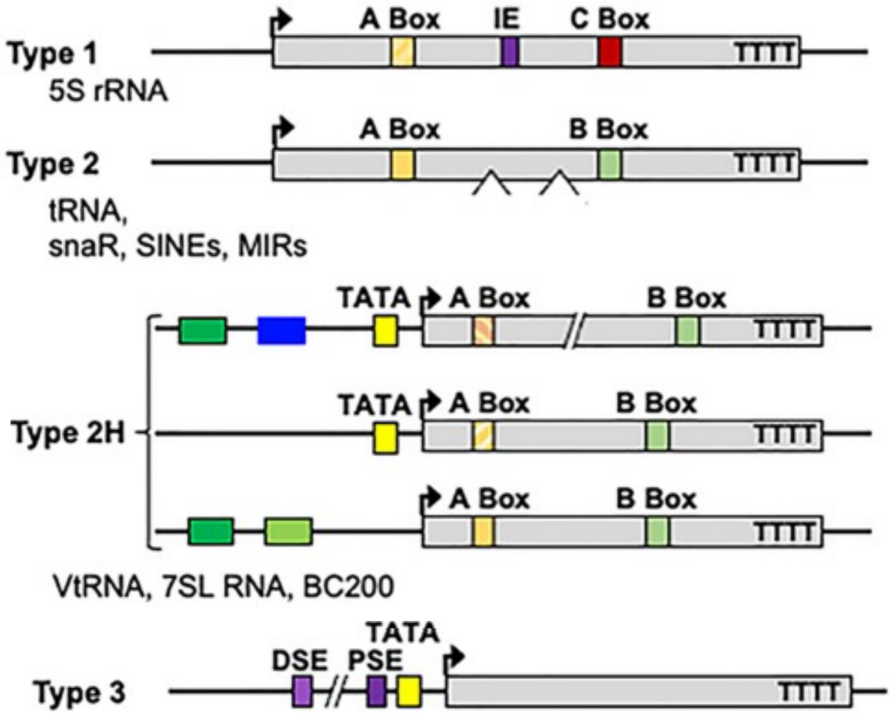
Pol III transcription

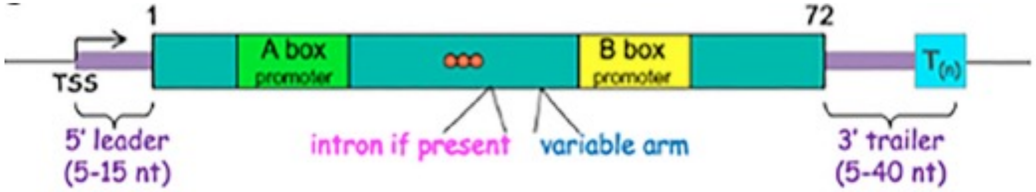
Pol III promoters of some classes are located within the gene

Yeast



Mammals





TATA - TATA box or TATA-like sequence

TBP - TATA binding protein

BRF1, BRF2 - TFIIB-related factor 1, 2

SNAPc - snRNA activating protein complex

PSE - proximal sequence element

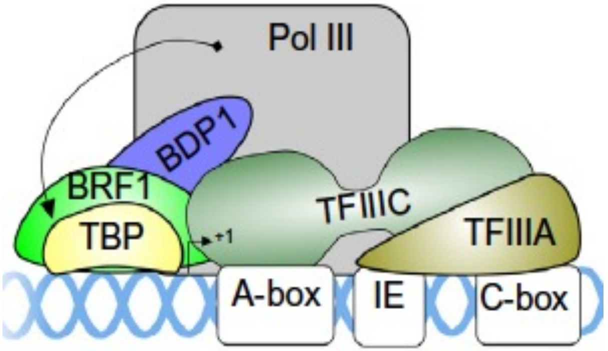
DSE - distal sequence element

TTTT - transcription termination signal

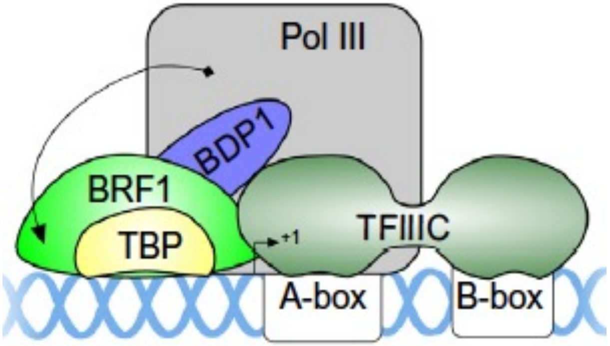
Kessler and Maraia, NAR, 2021

Pol III transcription

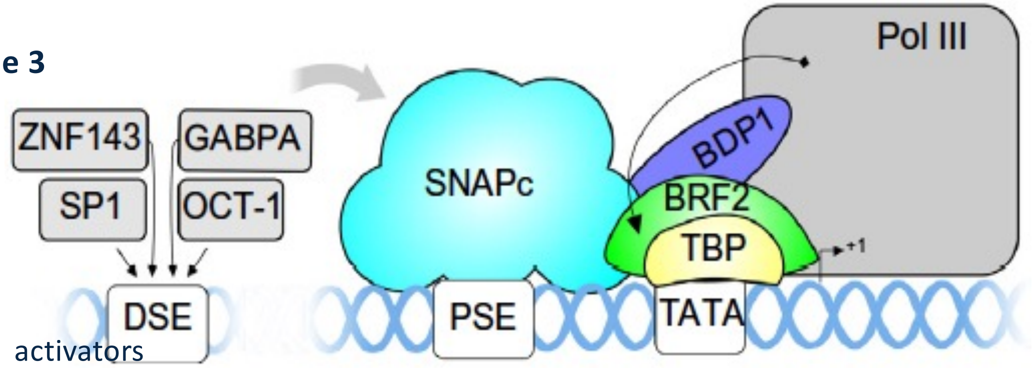
Type 1



Type 2



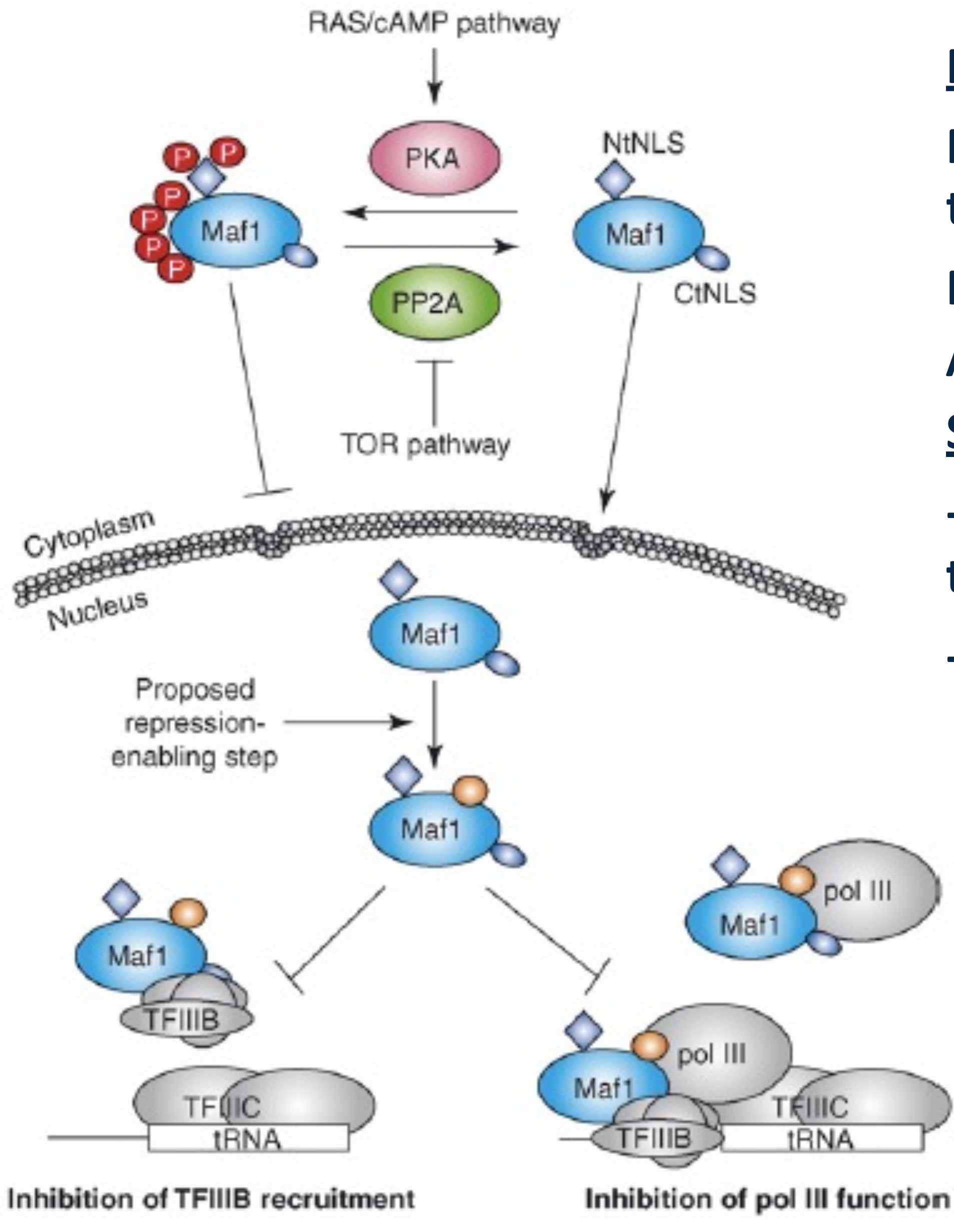
Type 3



TATA - TATA box or TATA-like sequence
TBP - TATA binding protein
BRF1, BRF2 - TFIIB-related factor 1, 2
SNAPc - snRNA activating protein complex
PSE - proximal sequence element
DSE - distal sequence element
TTTT - transcription termination signal

Dergai and Hernandez, 2019, TiG

Pol III regulation by Maf1 inhibitor



Normal growth (TORC1 active)

Maf1 is phosphorylated and remains in the cytoplasm

P-states of Maf1 are regulated by RAS /cAMP and TOR pathways

Stress (starvation, TORC1 inactive)

- Maf1 is dephosphorylated and imported to the nucleus

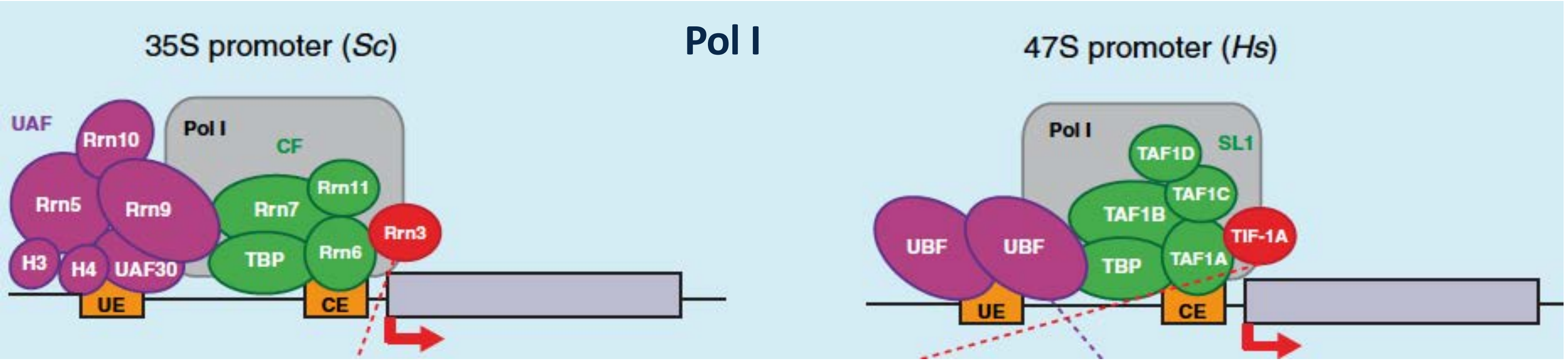
- Maf1 inhibits:

de novo assembly of TFIIIB

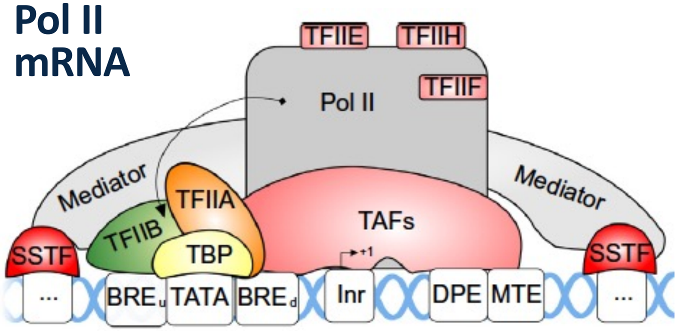
transcription by binding to Pol III

Willis and Morris, TiBS, 2007

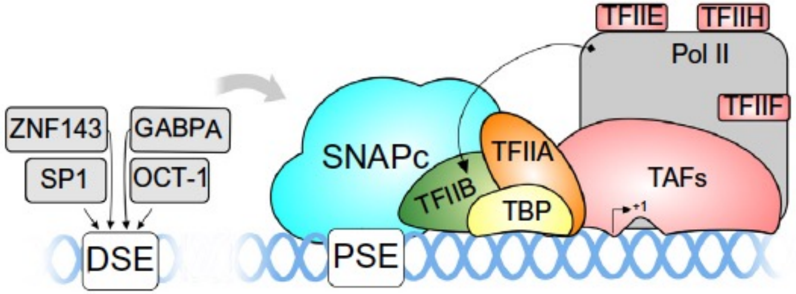
Pol I, Pol II and Pol III

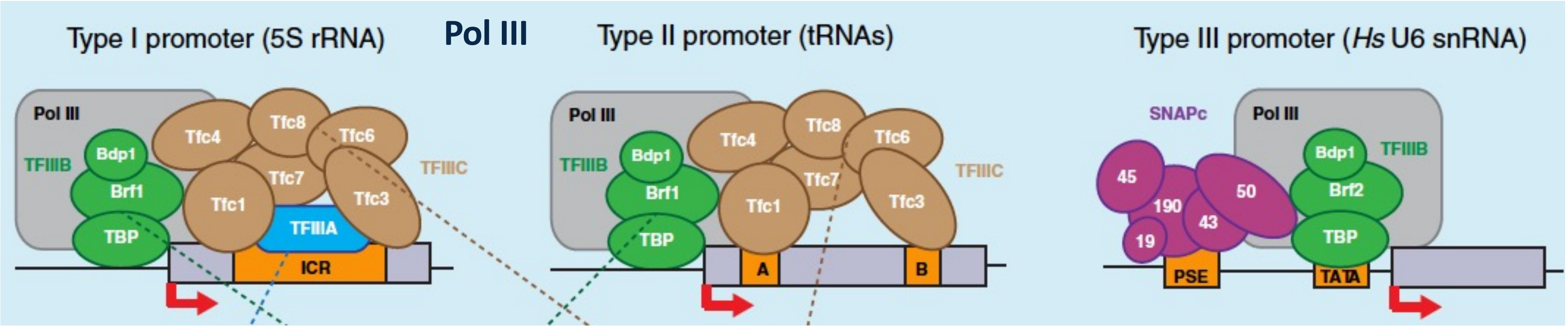


Pol II mRNA



Pol II snRNA





Vanini, BBA, 2013; Dergai and Hernandez, 2019, TiG

TAKE-HOME MESSAGE

- **Transcription of different RNAs eukaryotic is carried out by specialised RNA polymerases, I –III (*all*) and IV/V (*plants*)**
- **Transcription regulation is achieved on several levels: chromatin structure and modification, recruitment of transcription factors, silencing mechanisms (ncRNAs)**
- **Many RNA processing events occur cotranscriptionally (capping, splicing, 3' end formation, export)**
- **Transcription is regulated in response to nutrients, stress, cell cycle, development stage, etc...**