BI 8 LECTURE 14

EPIGENETICS AND GENE REGULATION: NEW ELEMENTS FOR TRANSCRIPTIONAL REGULATION IN MULTICELLULAR EUKARYOTES

Ellen Rothenberg

18 February 2016

Reading: Alberts Ch.4 pp.187-207; Ch. 7, pp. 369-407

Four barriers to prokaryotic modes of gene regulation in eukaryotes

- Nuclear/ cytoplasmic distinction (space & time separation between RNA, protein synthesis)
- Genome complexity: enormous search space for TFs
- In multicellular eukaryotes, extreme biological complexity via different gene expression patterns
- Packaging of DNA in highly dense chromatin: accessibility barrier?

Histone packing of DNA in eukaryotic nuclei: not just free DNA

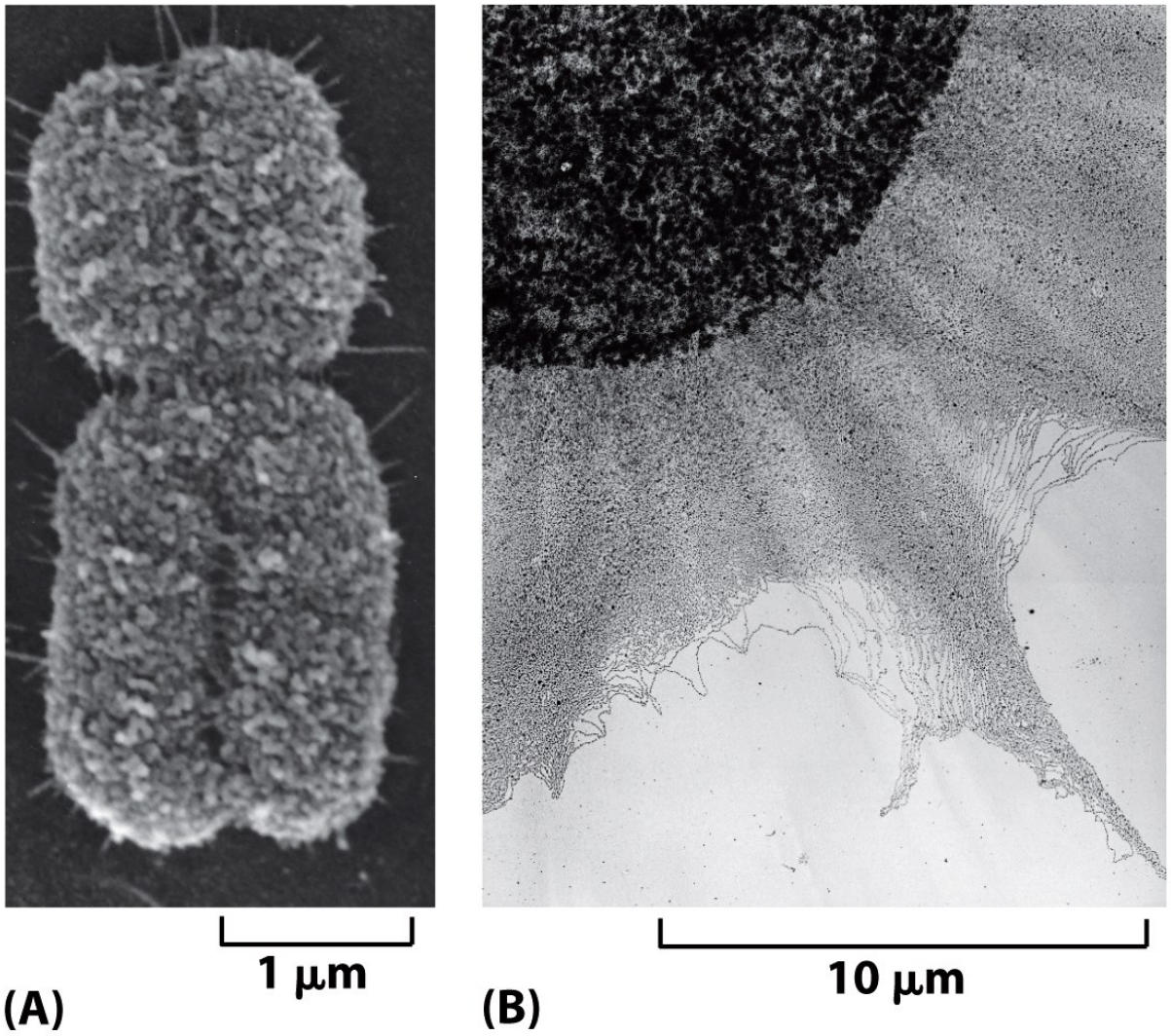


Figure 4-20 *Molecular Biology of the Cell* (© Garland Science 2008)

Large loops do not correspond to genes, but help facilitate long-range interactions ~100 Mb

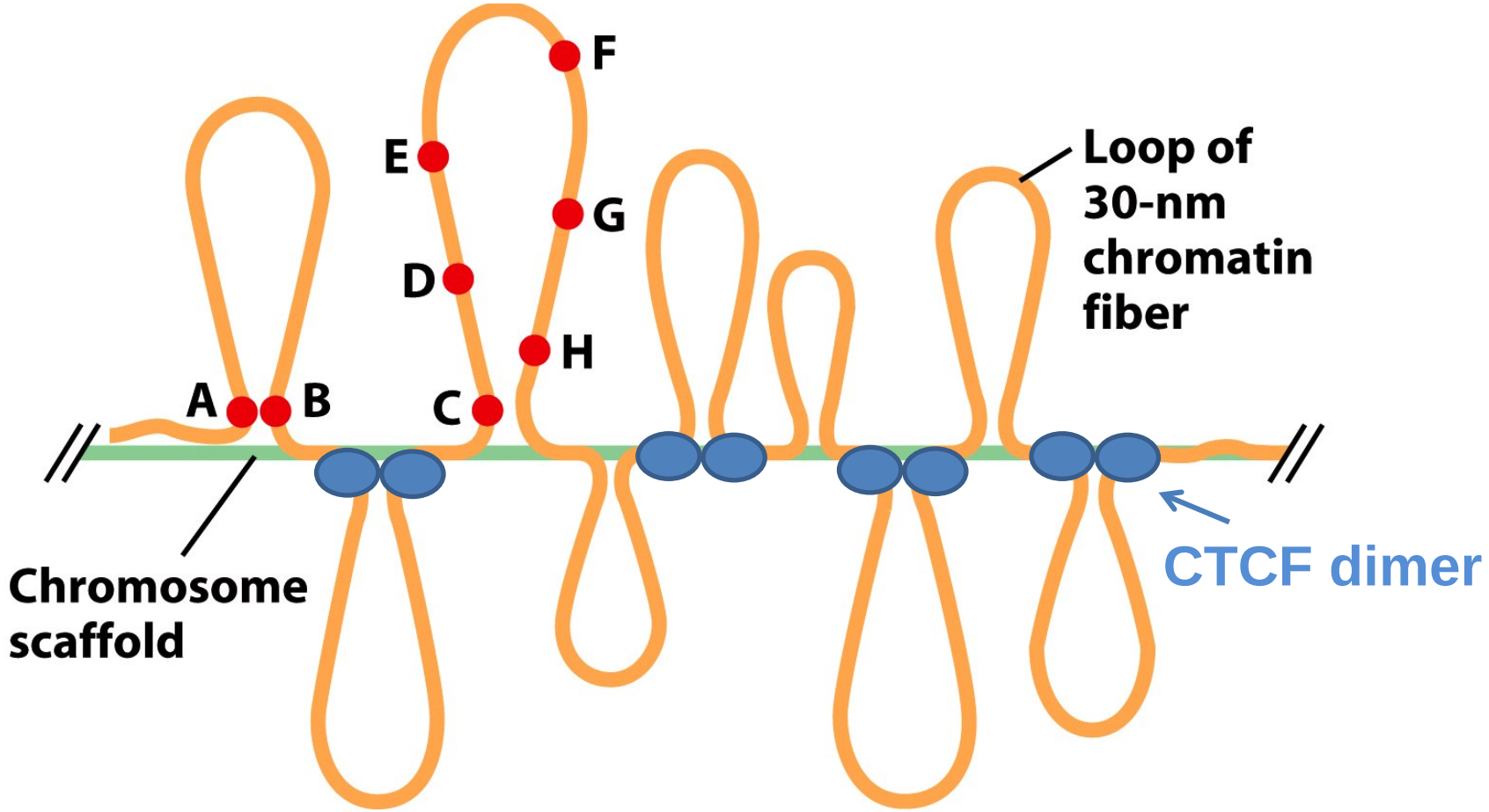
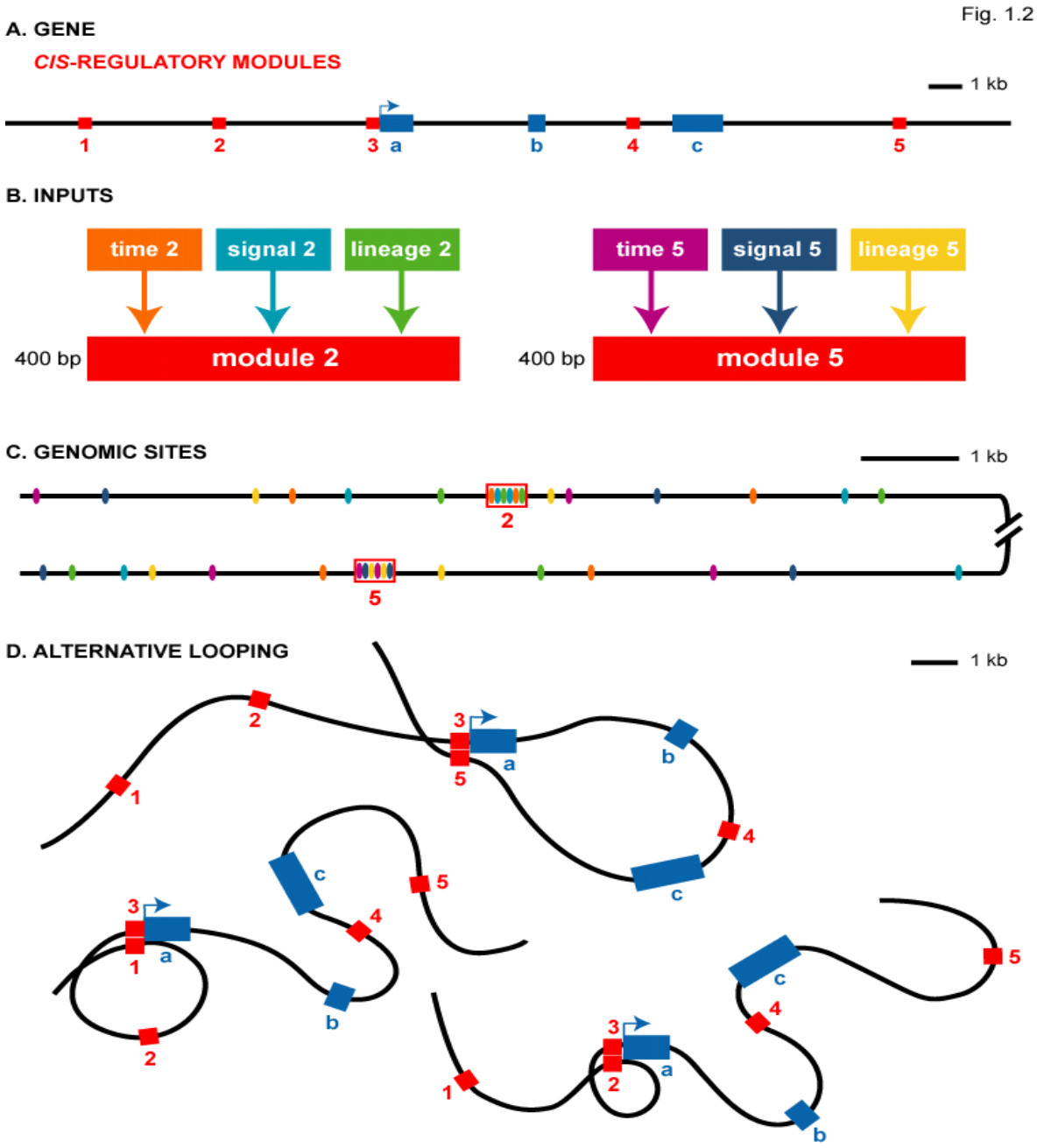


Figure 6-36
Molecular Cell Biology, Sixth Edition

In addition to these static loops, there are also dynamic loops…

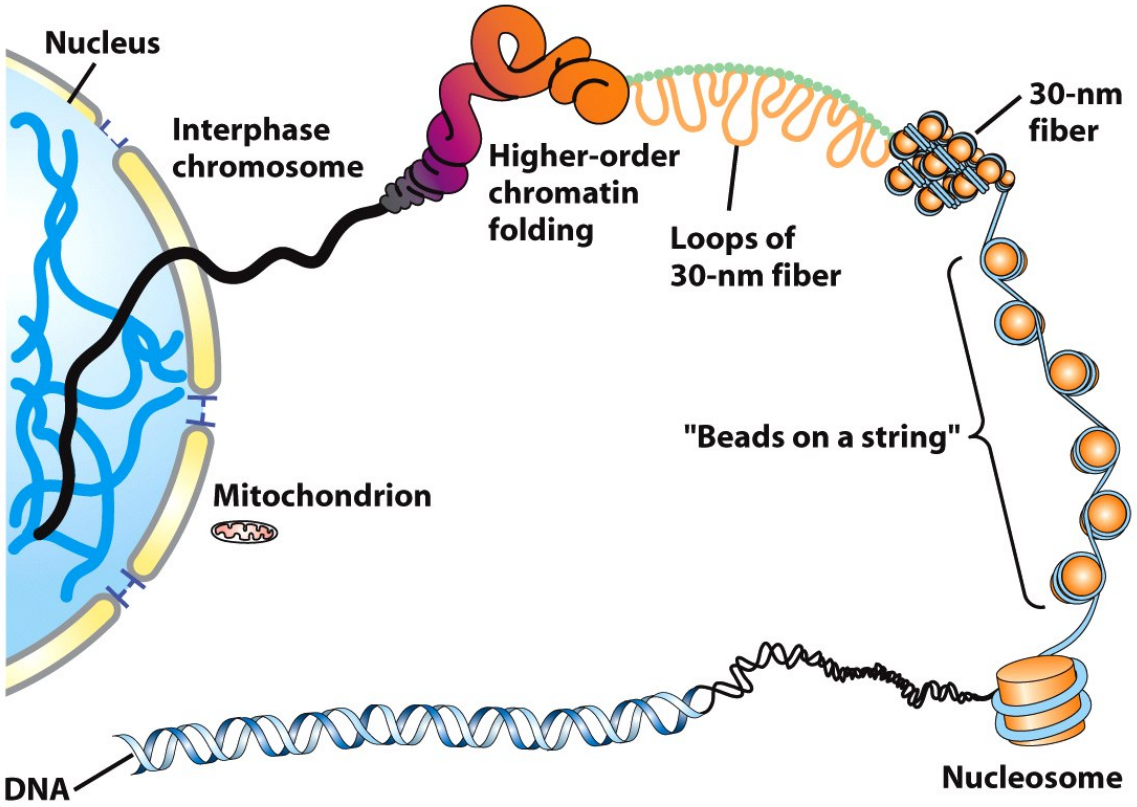
Within the larger loops, gene expression is controlled by dynamic looping of promoters to contact specific enhancer(s) that are bound by TFs in that cell at that time

(stay tuned for much more on this!)



Davidson, E.H. Genomic Regulatory Systems 2006

Multiscale packaging of DNA in eukaryotic cells



Major Types of DNA Sequence

Single-copy genes	Simple-sequence DNA
Gene families	Transposable DNA elements
Tandemly repeated genes	Spacer DNA
Introns	

Figure 6-1
Molecular Cell Biology, Sixth Edition

Nucleosomes can be decomposed into core histone octamers

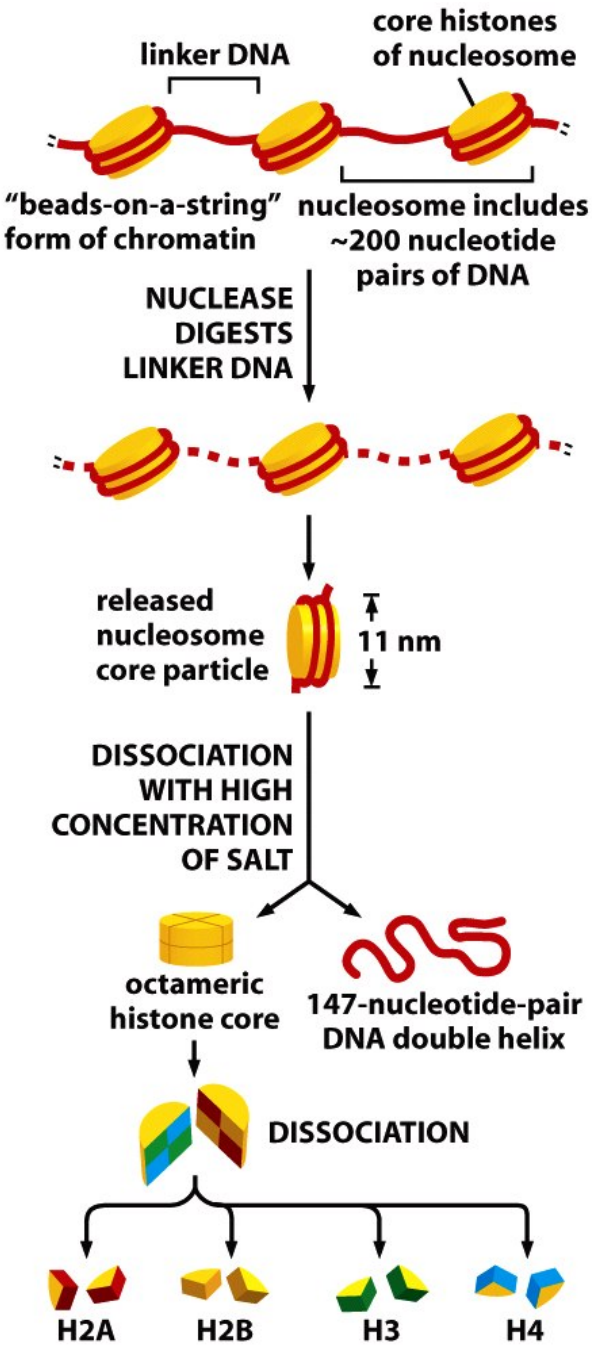


Figure 4-23 *Molecular Biology of the Cell* (© Garland Science 2008)

H3-H4 subunits and H2A-H2B subunits

Dimers of dimers assemble together with DNA wrapped around the outside… histone N-terminal tails stick out

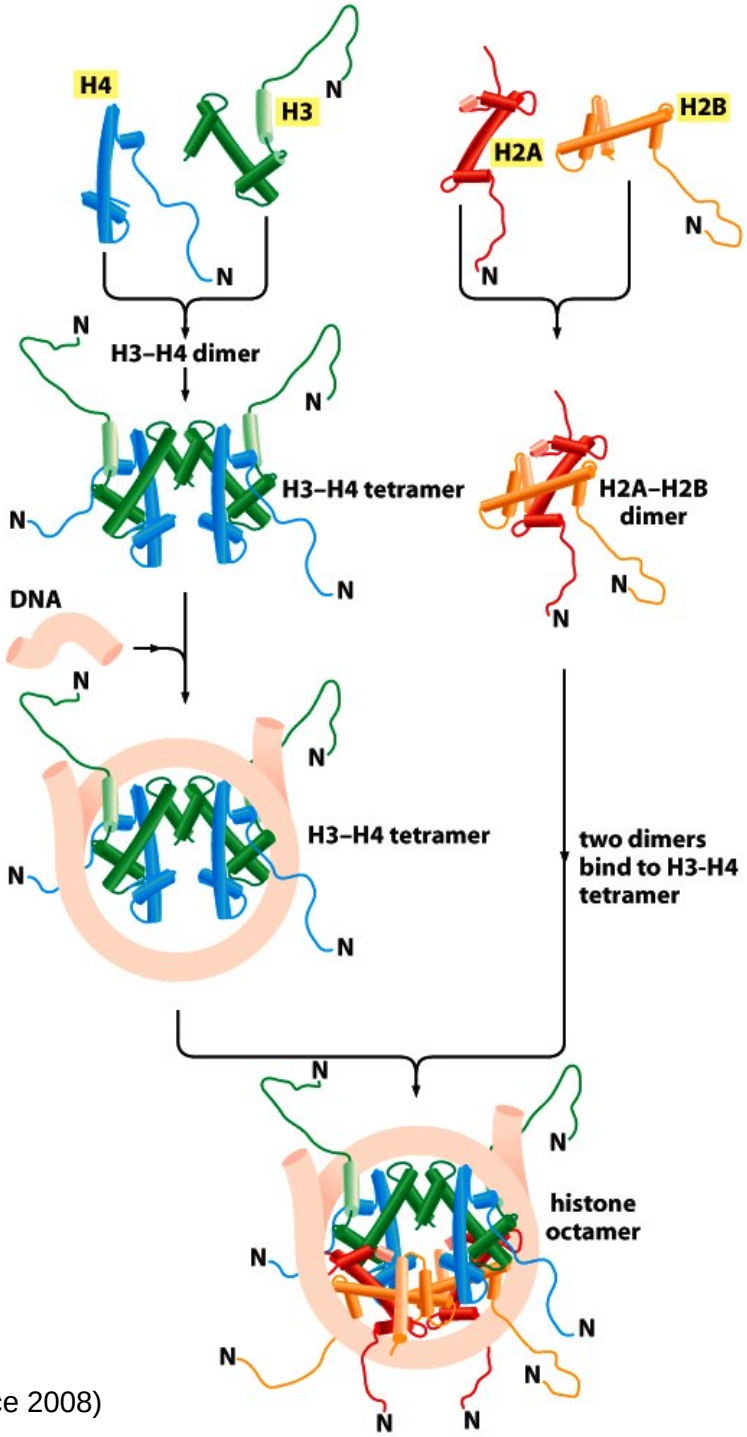
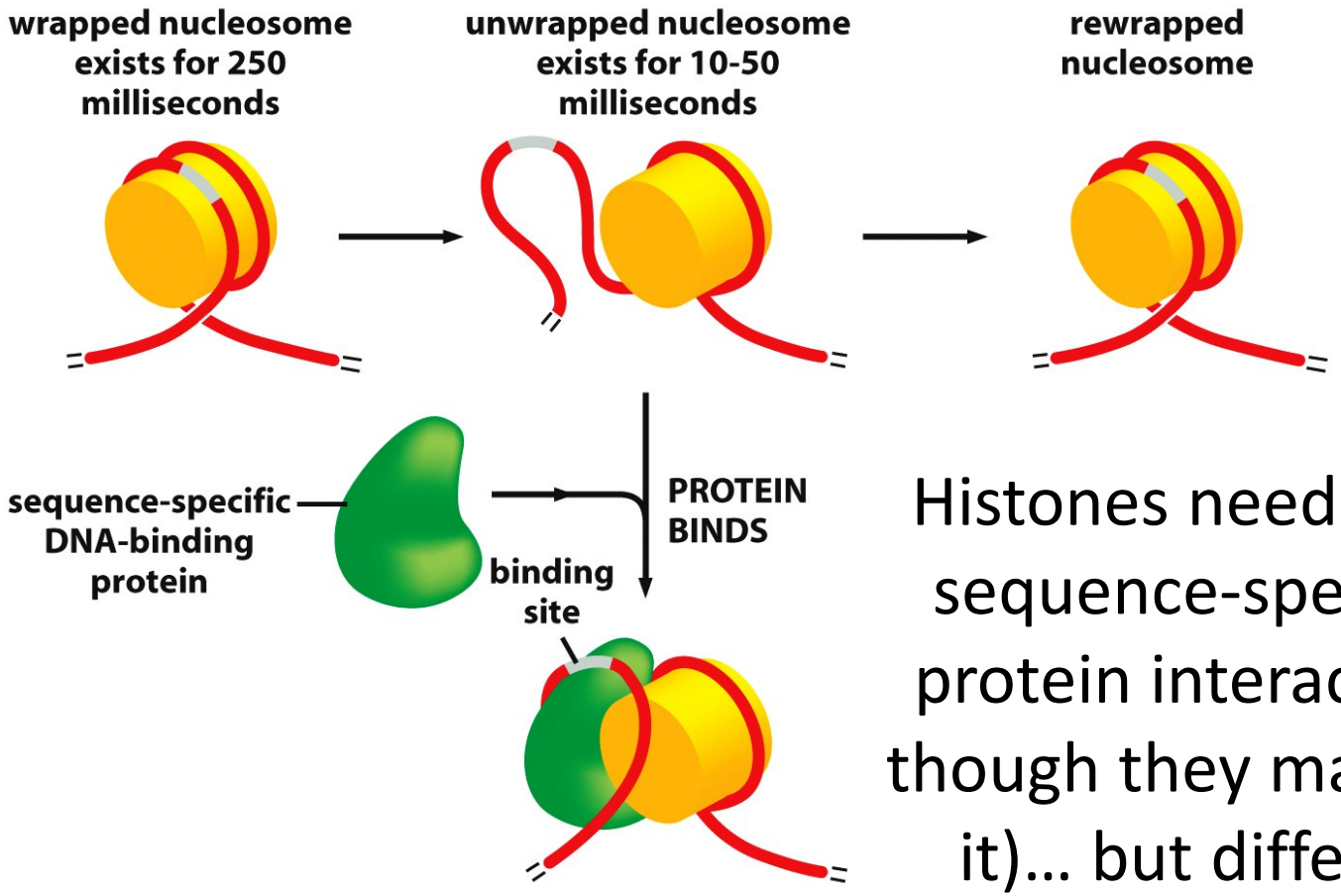


Figure 4-26 *Molecular Biology of the Cell* (© Garland Science 2008)

What is the role of histones in gene regulation?

- A neutral general obstruction? (if so, where is the specificity?)
- A local obstacle to prevent excessive transcription factor binding to “forbidden” genes?
- A specific traffic controller for transcription factor access?
- A medium on which DNA can be marked at particular sites by previous transcription factor action?

Nucleosome dynamics allow transcription factors to get to DNA …but less often where nucleosomes are densely packed



Histones need NOT block sequence-specific DNA-protein interaction (even though they may constrain it)… but differences in accessibility may explain some *preferential* binding

Figure 4-28 *Molecular Biology of the Cell* (© Garland Science 2008)

Open vs. closed chromatin: part of the answer to controlling gene accessibility?

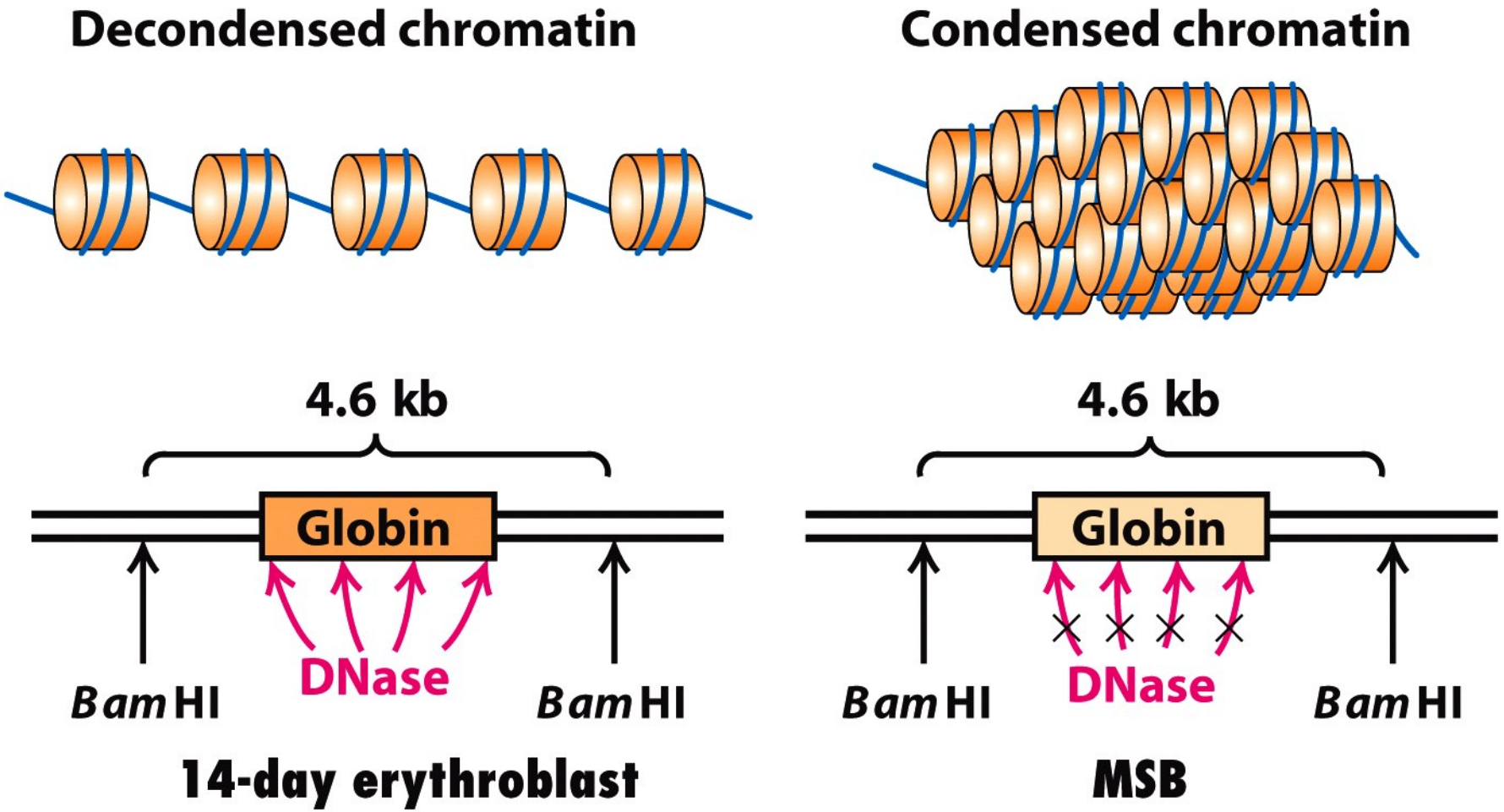
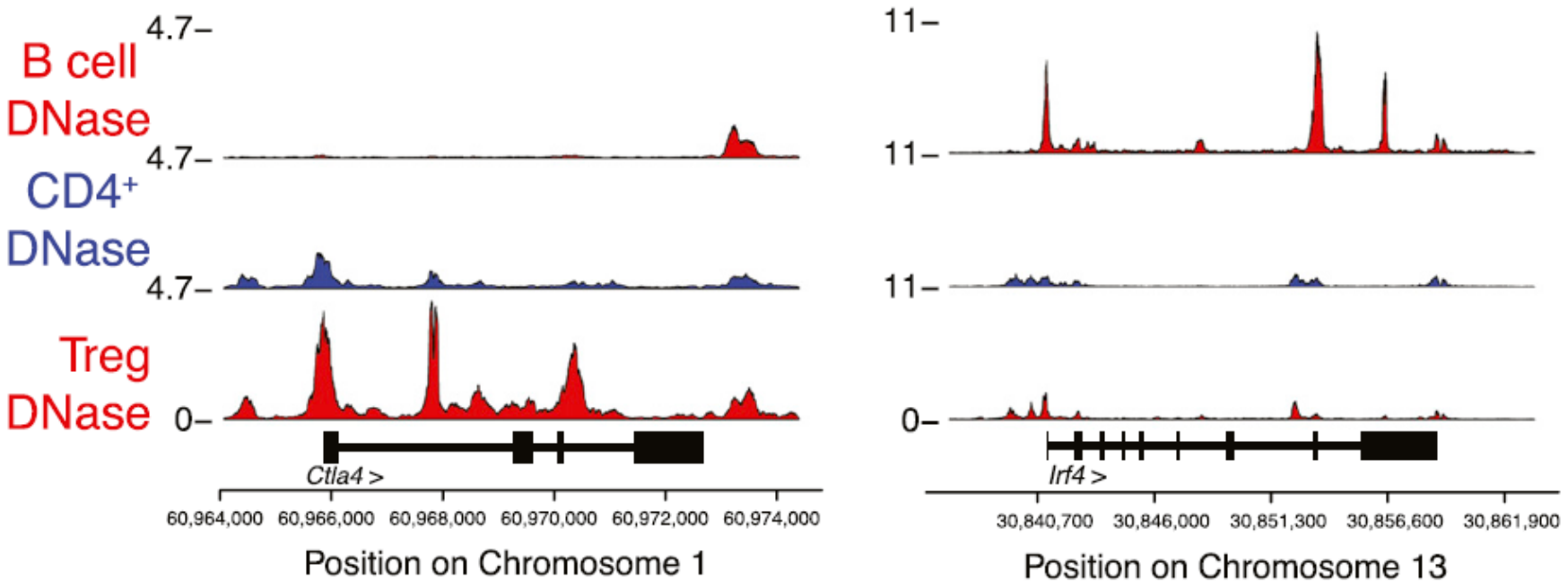


Figure 6-32a
Molecular Cell Biology, Sixth Edition

Using modern high-throughput sequencing technology, DNase hypersensitivity locates regions of DNA that may be especially active in cell-type specific gene regulation



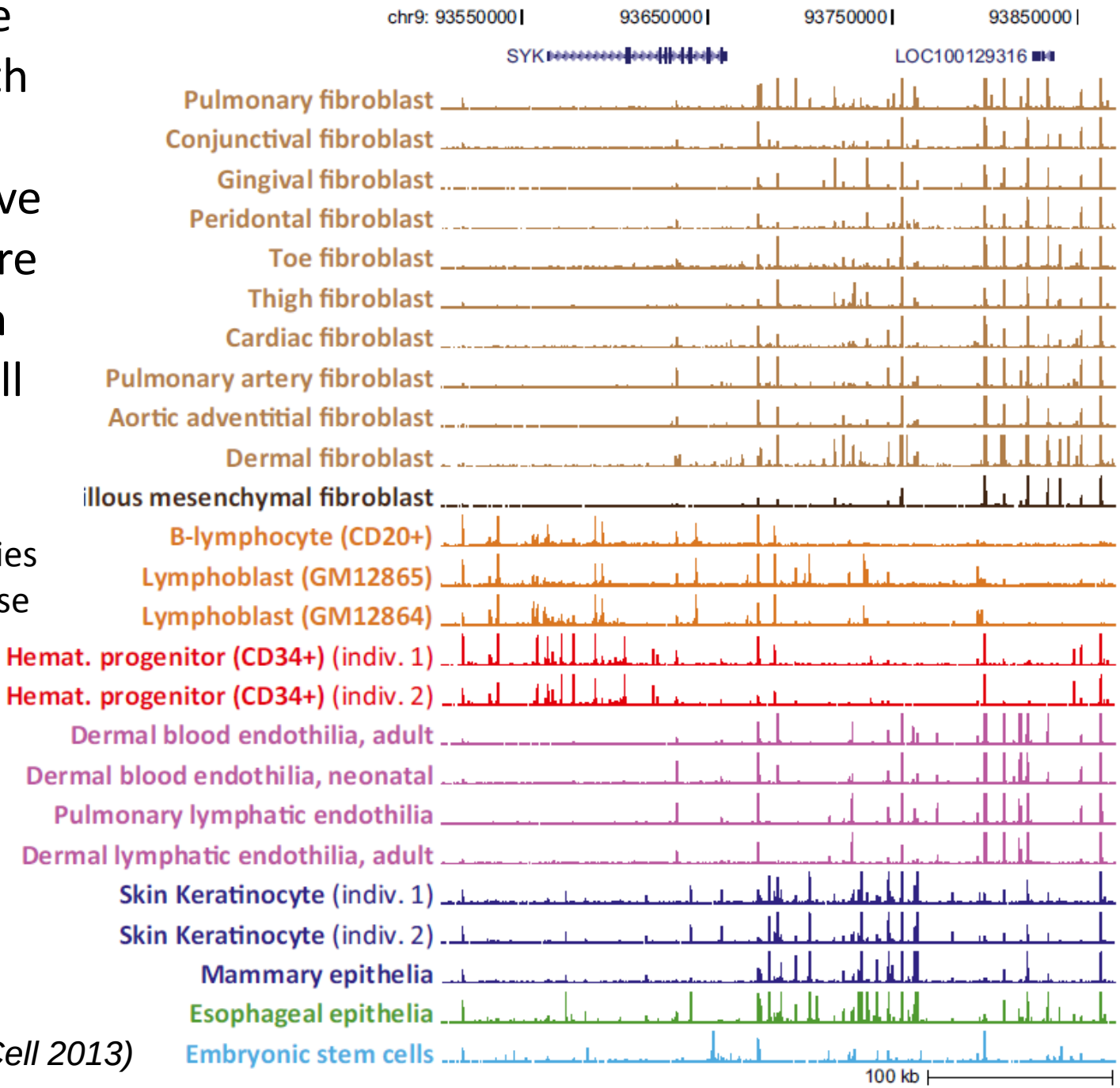
Treg specific gene

B cell specific gene

Samstein, … & Rudensky, Cell 151: 153-66, Sept 2012

Parts of the genome with DNase hypersensitive chromatin are different in different cell types…

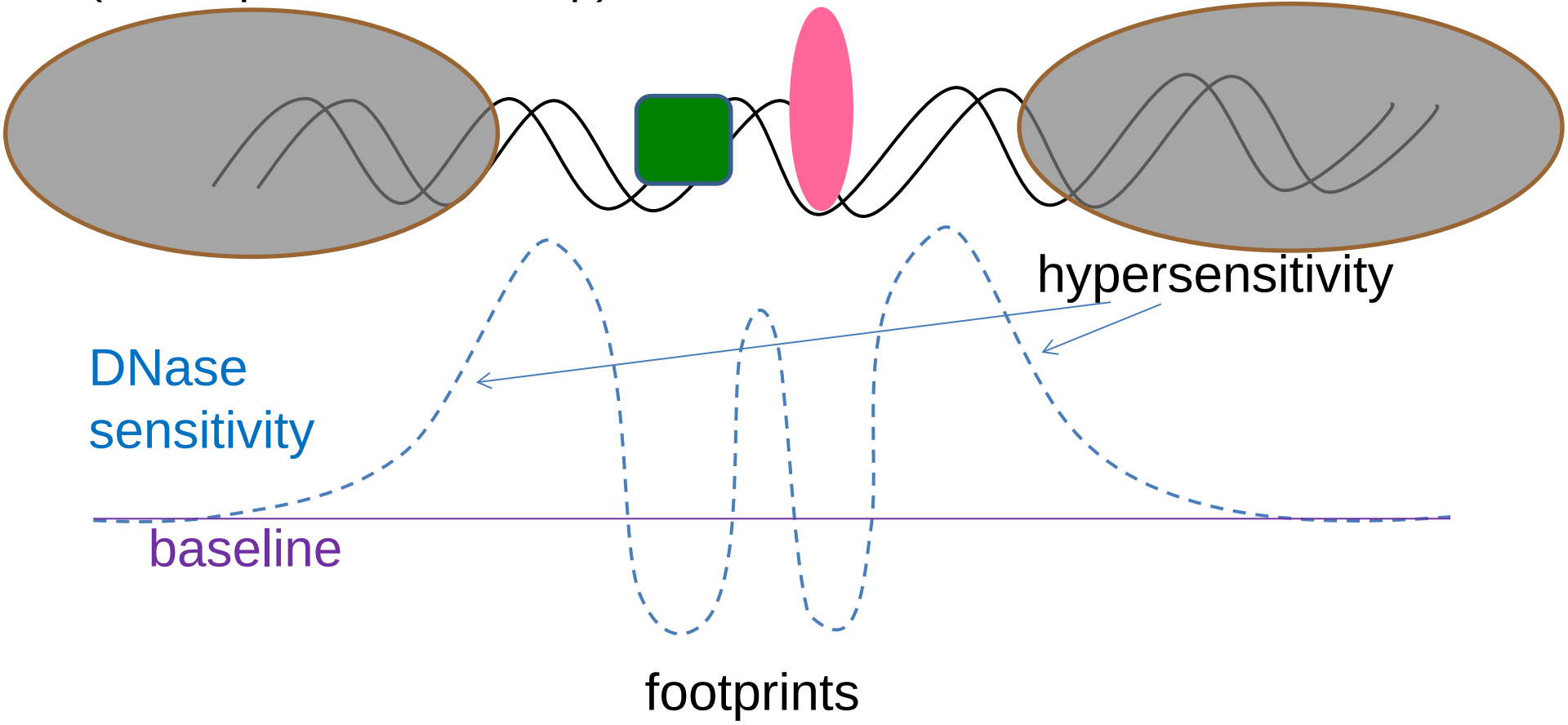
but the similarities make some sense



(Stergachis et al, Cell 2013)

One important point: these DNase hypersensitive regions are the opposite of classic DNA footprints: coarse vs. fine grain visualization

In vivo footprint is small… hypersensitive region around it is bigger (5-10 bp vs. 300-500 bp)



“Epigenetic marking”: Decoding a problematic term

- DNA in chromosomes is packaged into chromosomes that at any given moment may be different
  - in the density of histones per length of DNA (compaction/ openness)
  - in the post-translational modification of histones over particular regions (“epigenetic marks”)

- Genes that are currently being expressed are less compacted and have different marks on their histones than silent genes

- Real issue: in a differentiated cell, is all of the genome equally accessible to every transcription factor?
  - access differences based on “inertia” due to cooperativity?
  - access differences based on developmental history of the cell?
  - If not, then how are these differences caused and how reversible are they?

Distinguishing genomic regions that are/are not "in play": Two broad kinds of "epigenetic marks"

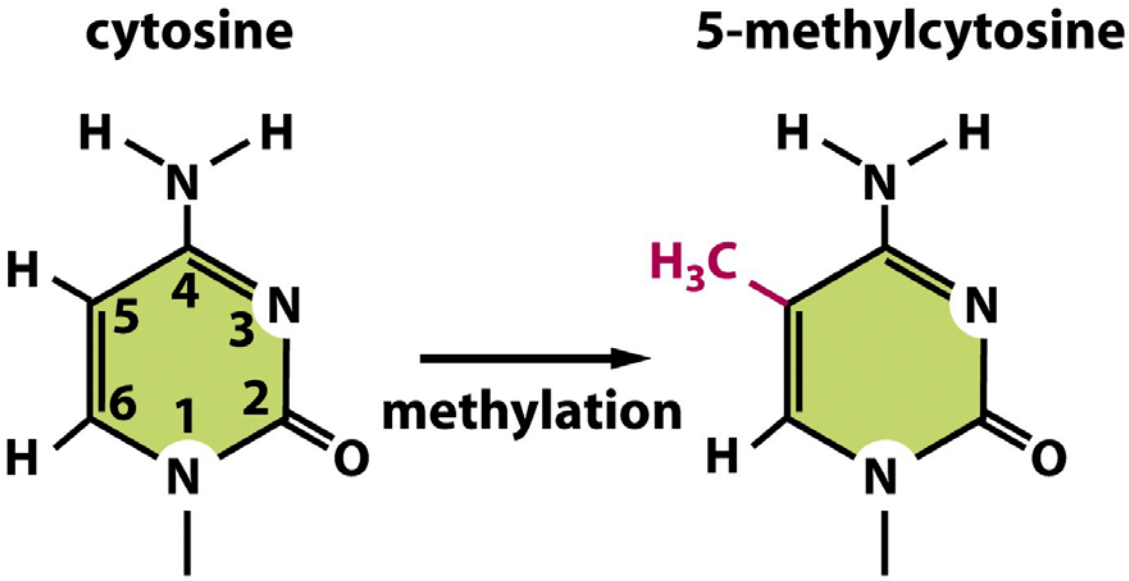
- Covalent (but reversible) methylation of DNA sequence itself on CpG dinucleotides

- Covalent (but reversible) post-translational modification of histone protein tails in nucleosomes enclosed by particular DNA sequence
  - Methylation of lysines
  - Acetylation of lysines

These have a strong correlation with ***different regulatory activity states*** of the local DNA sequences

However, ***all*** these marks can be removed as well as added

A special epigenetic modification: CpG methylation of ***DNA***



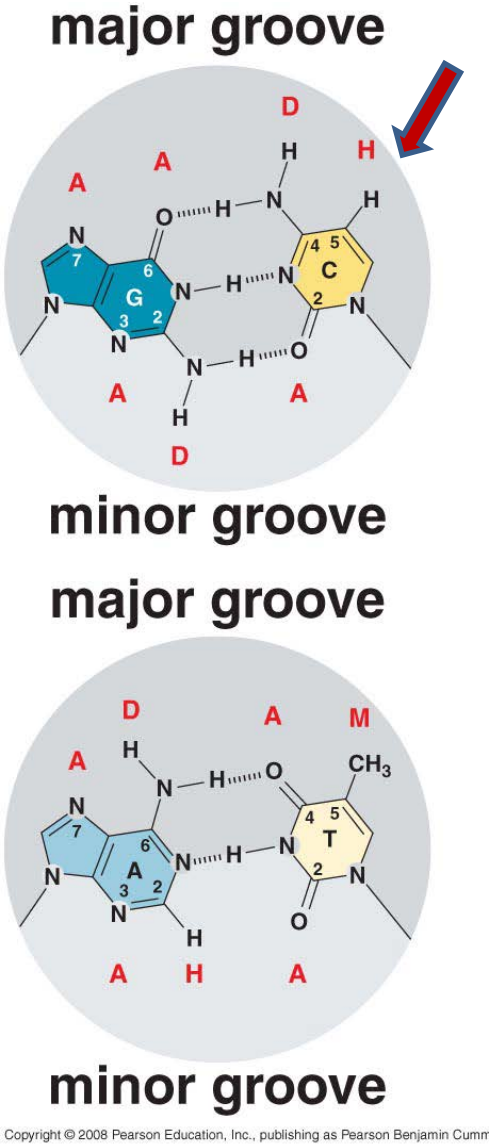


Figure 7-79 *Molecular Biology of the Cell* (© Garland Science 2008)

Doesn't block DNA base pairing, but affects recognition surface "seen" by transcription factors in major groove
(may also make duplexes a little more stable)

Effects of CpG methylation on gene regulation

- If a TF normally binds to a site with a 5'-CG-3' dinucleotide in it, methylation can block it from recognizing the site

- DNA with methylated CpG's can be preferentially bundled into closed chromatin

- CpG methylation can block CTCF binding near a gene, causing the gene to end up in the wrong loop & under influence of the wrong regulatory elements

Once established, CpG methylation can be maintained over multiple cell generations by default, without additional regulatory input

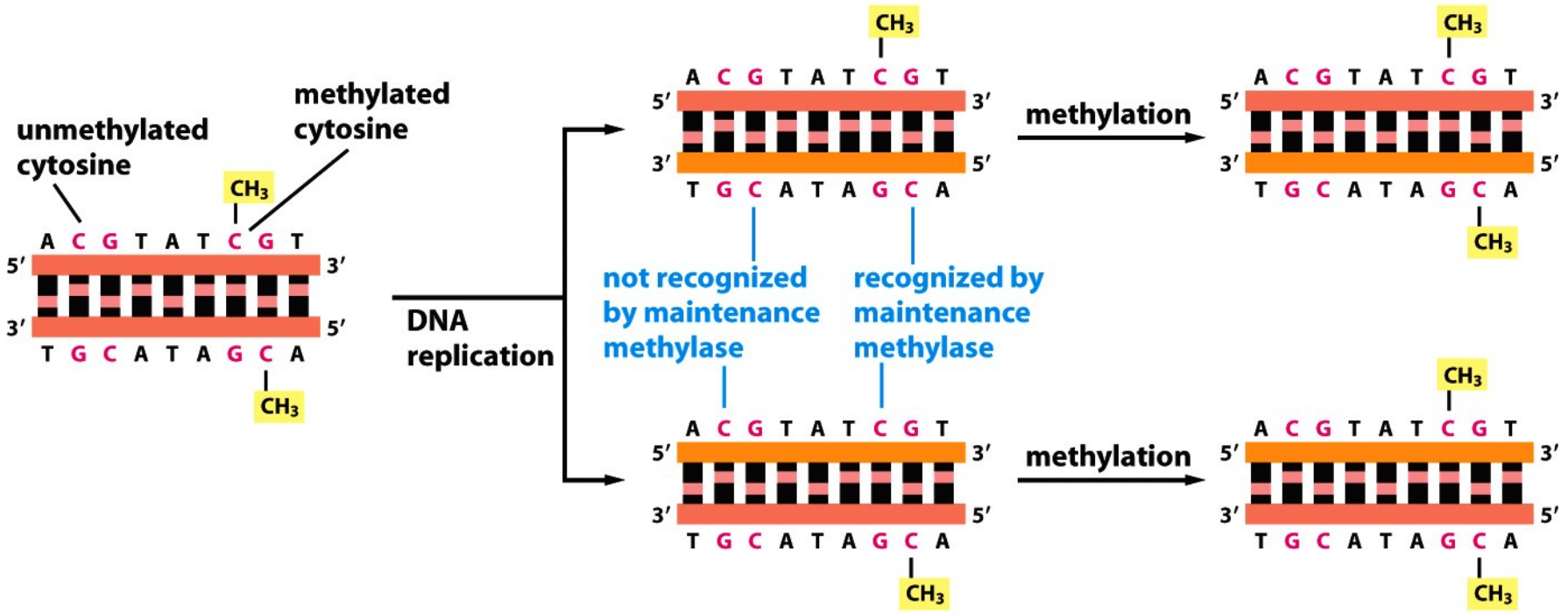
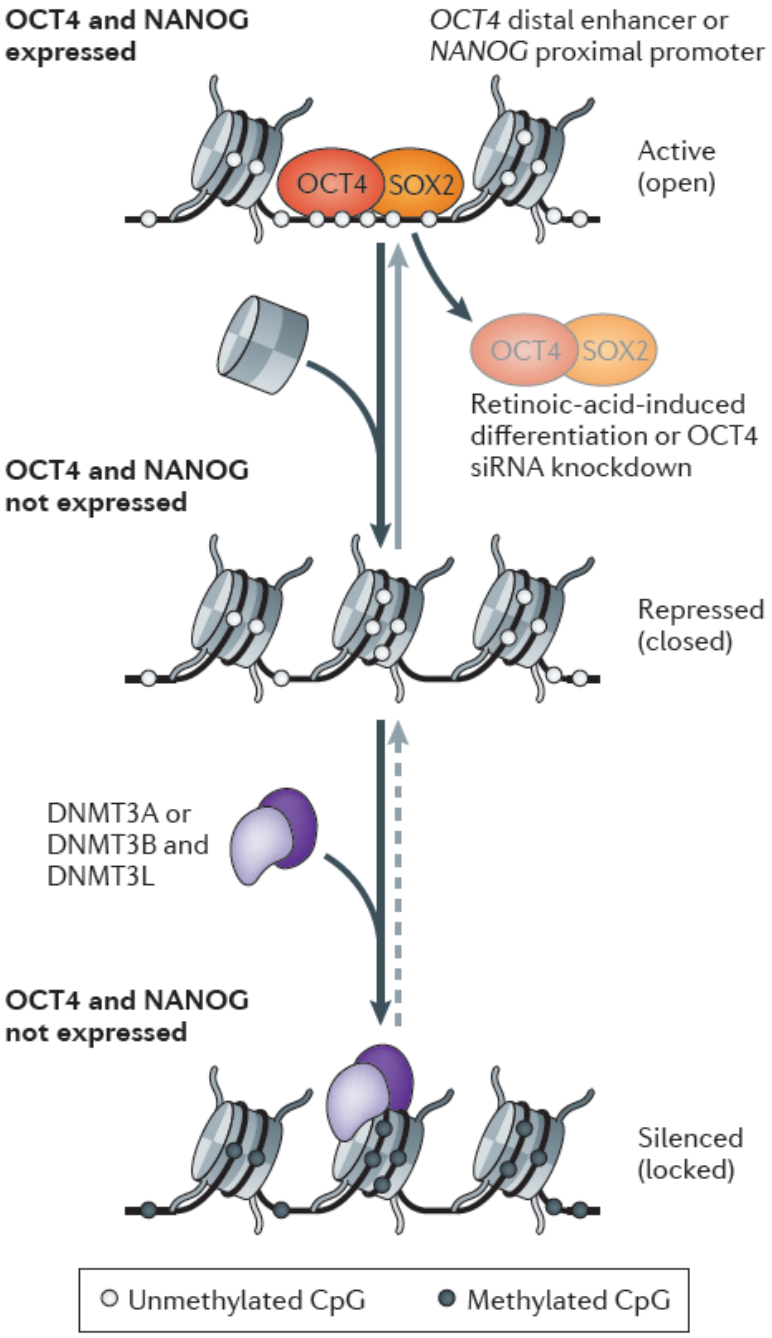


Figure 7-80 *Molecular Biology of the Cell* (© Garland Science 2008)

DNA methylation can occur as a result of transcriptional silence, not as a cause of it

Methyltransferase enzymes (Dnmt) that initiate methylation are recruited to genes that are already repressed



P. A. Jones, Nat Rev Genet, July 2012

But… Methylation can propagate and can be used to nucleate binding of DNA methyl-binding proteins… which can help to keep genes silent in future cell generations

→ MEMORY

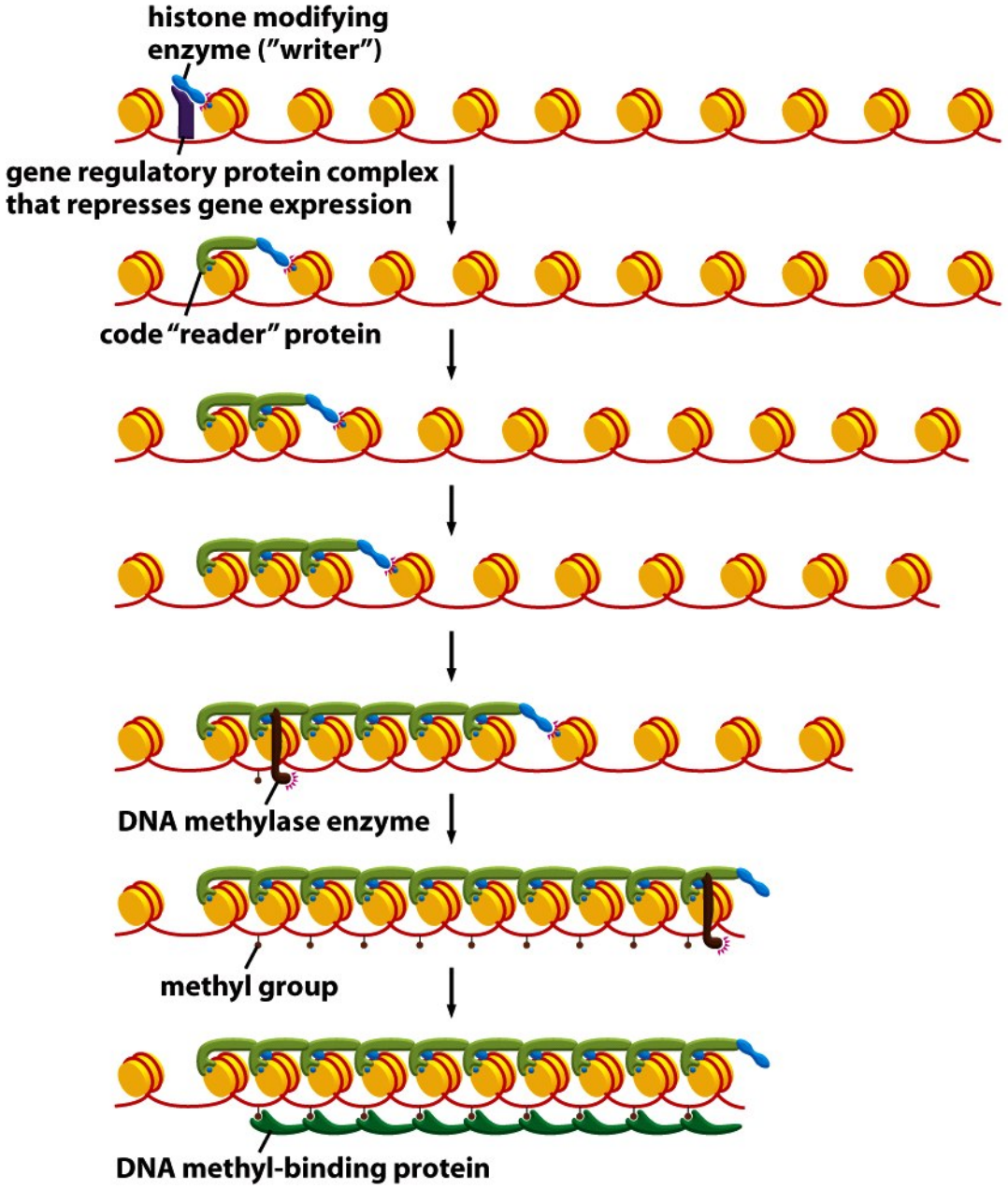
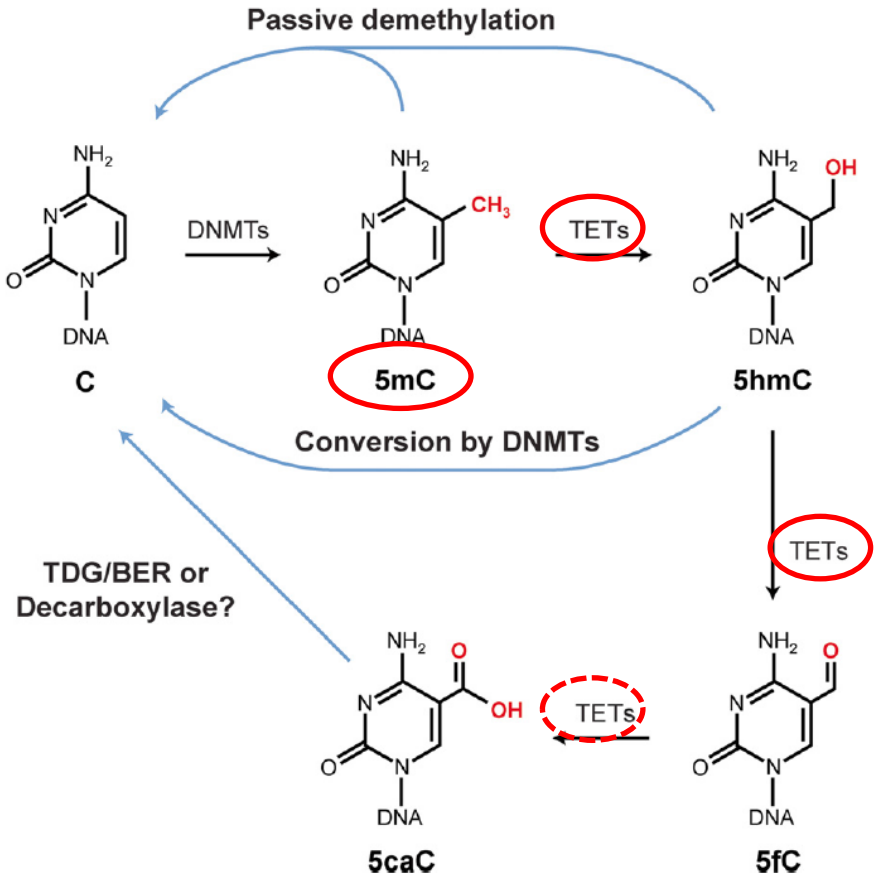


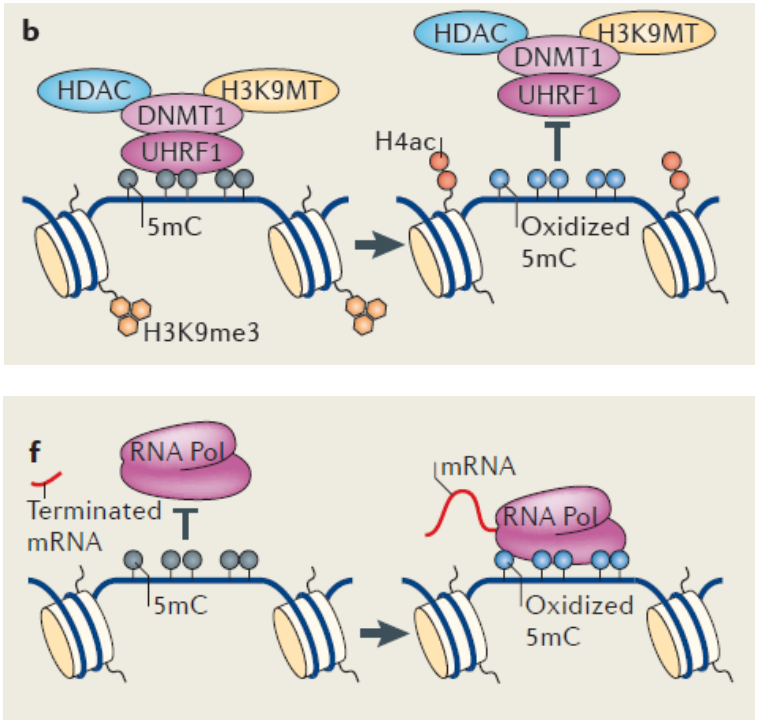
Figure 7-81 *Molecular Biology of the Cell* (© Garland Science 2008)

Still, DNA methylation (like histone modification) is reversible in a regulated way



Key transcription factors can recruit "TET" enzymes to start demethylating local DNA

Pfeifer *et al. Epigenetics & Chromatin* 2013, **6**:10



Even demethylation intermediates alter function of DNA to relieve repression

Pastor, Aravind, Rao *Nat Rev Mol Cell Biol* 2013, **14:** 341

Two broad kinds of “epigenetic marks”

- Covalent (but reversible) post-translational modification of histone protein tails in nucleosomes enclosed by particular DNA sequence
  - Methylation of lysines
  - Acetylation of lysines

- Covalent (but reversible) methylation of DNA sequence itself on CpG dinucleotides

These have a strong correlation with ***different regulatory activity states*** of the local DNA sequences

However, ***all*** these marks can be removed as well as added

Histones may be modified to alter their effects on DNA and TF interaction…

Histone N-terminal tails in many of the cell's histone octamers are post-translationally modified by phosphorylation, acetylation, or methylation

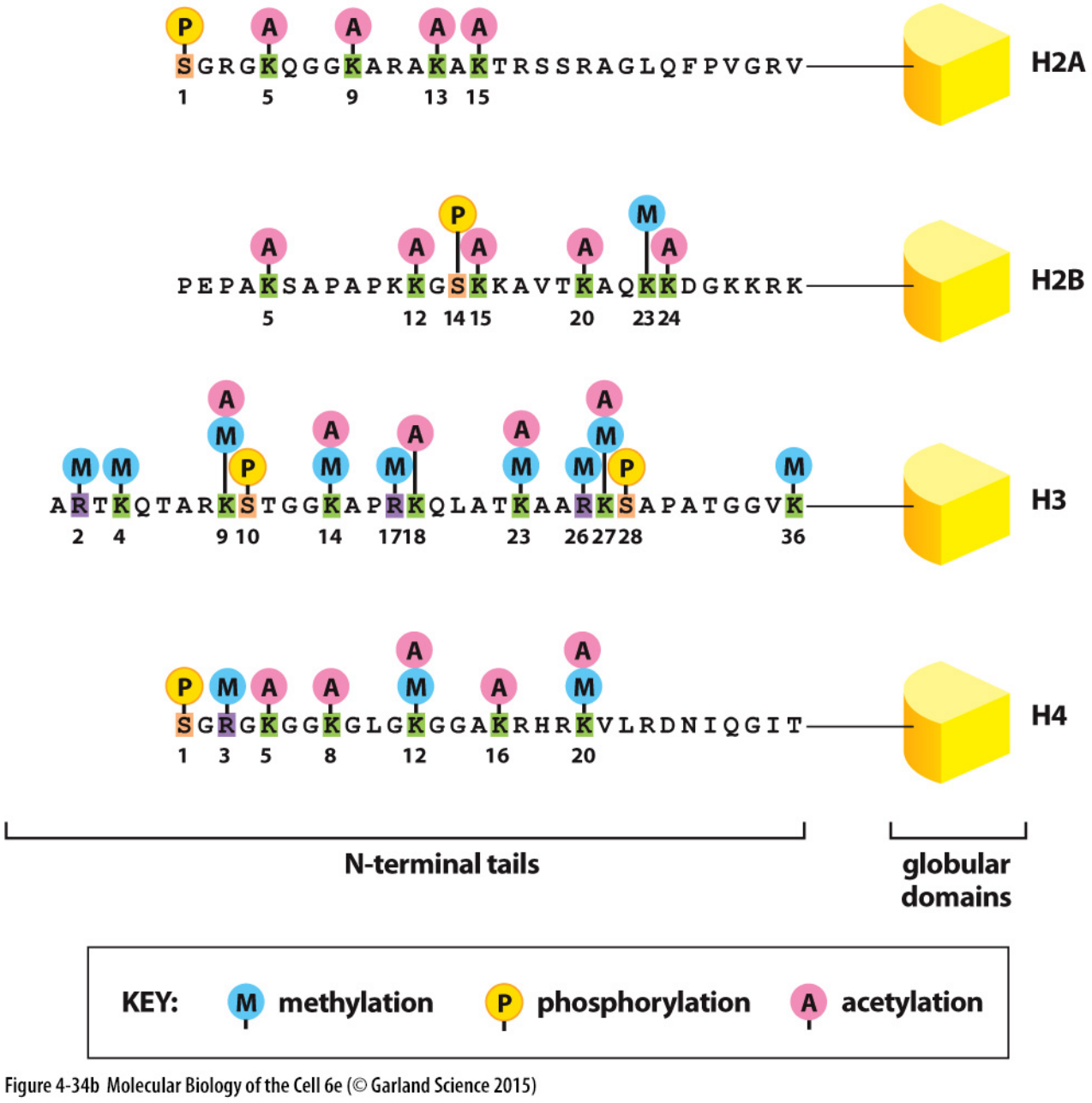


Figure 4-34b Molecular Biology of the Cell 6e (© Garland Science 2015)

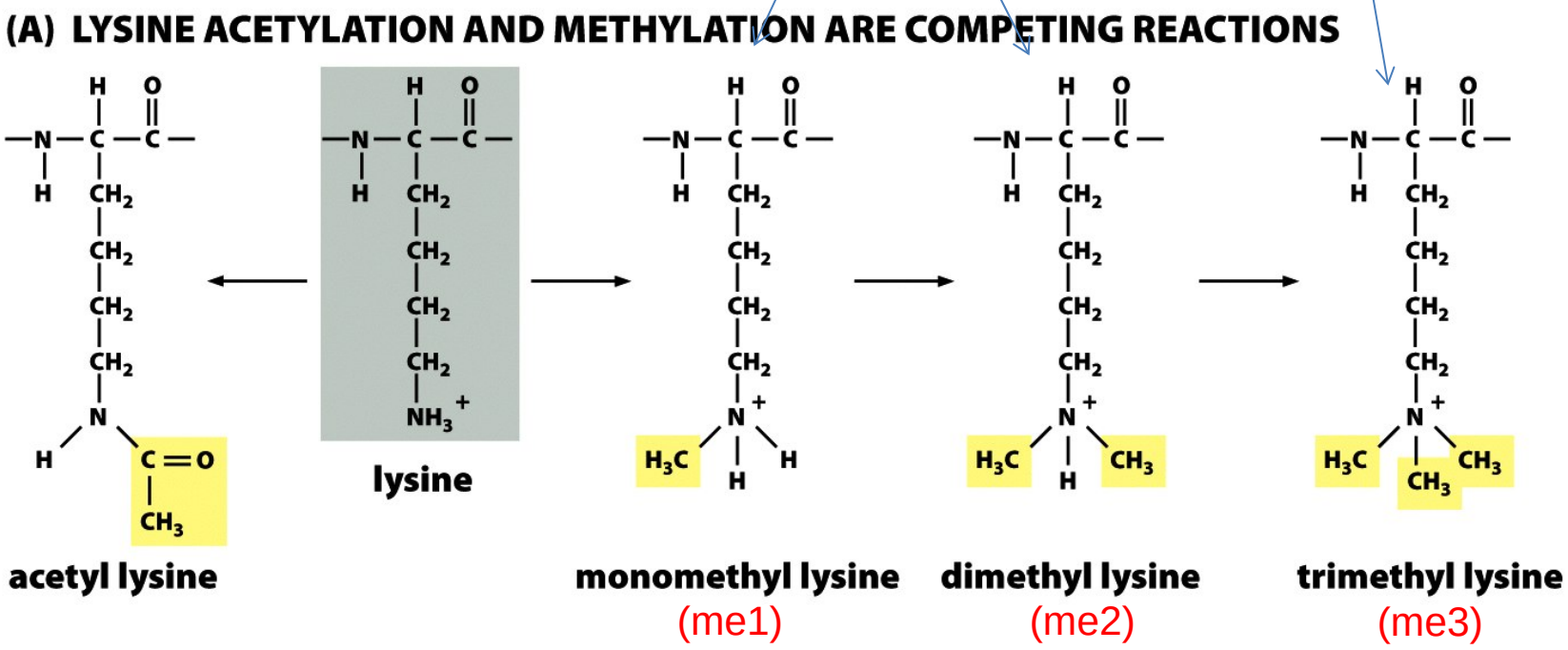
Figure 4-39b *Molecular Biology of the Cell* (© Garland Science 2008)

A short “code” phrasebook: common H3 marks

modification state	“meaning”
	**heterochromatin formation, gene silencing**
	**gene expression** (promoter)
	**gene expression** (active enhancer)
	**silencing of Hox genes, X chromosome inactivation**

Poised (K4)

Activated (K4) or repressed (K9, K27)



Stepwise modification of histone lysines is mediated by different enzymes… so intermediates are often seen

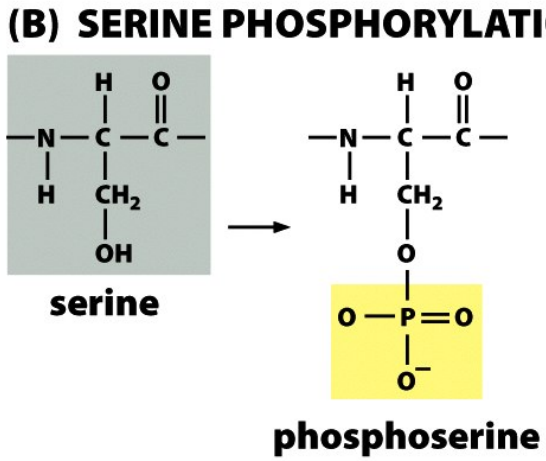


Figure 4-38 *Molecular Biology of the Cell* (© Garland Science 2008)

Some of the marks found on histones in “silent chromatin” provide docking sites for chromatin-condensing factors

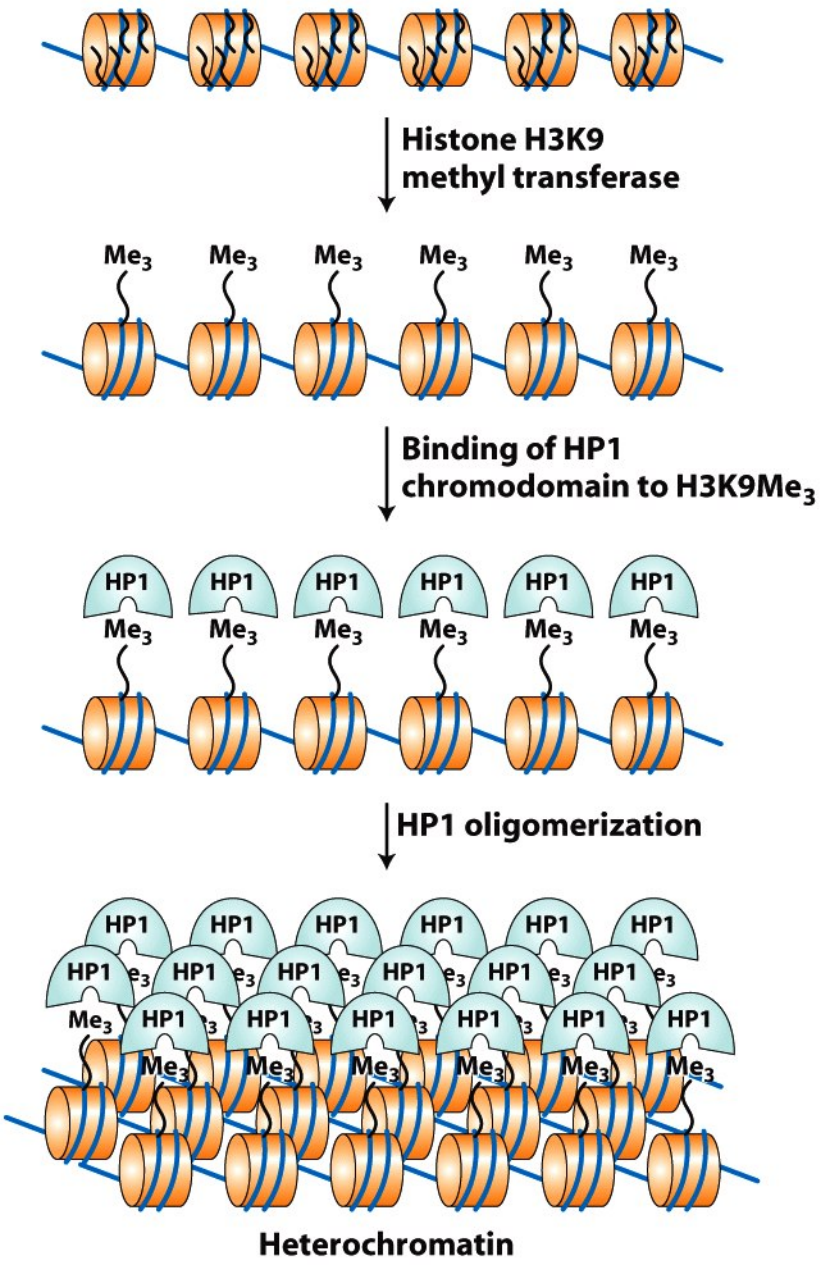


Figure 6-34a
Molecular Cell Biology, Sixth Edition

Once initiated, some modifications, especially “repressive” ones, can propagate themselves away from the initiating site through a recruitment-relay mechanism

via recruited protein complexes that act as “readers” (binding the marked histones) and “writers” (adding same marks to neighboring ones)... But not forever

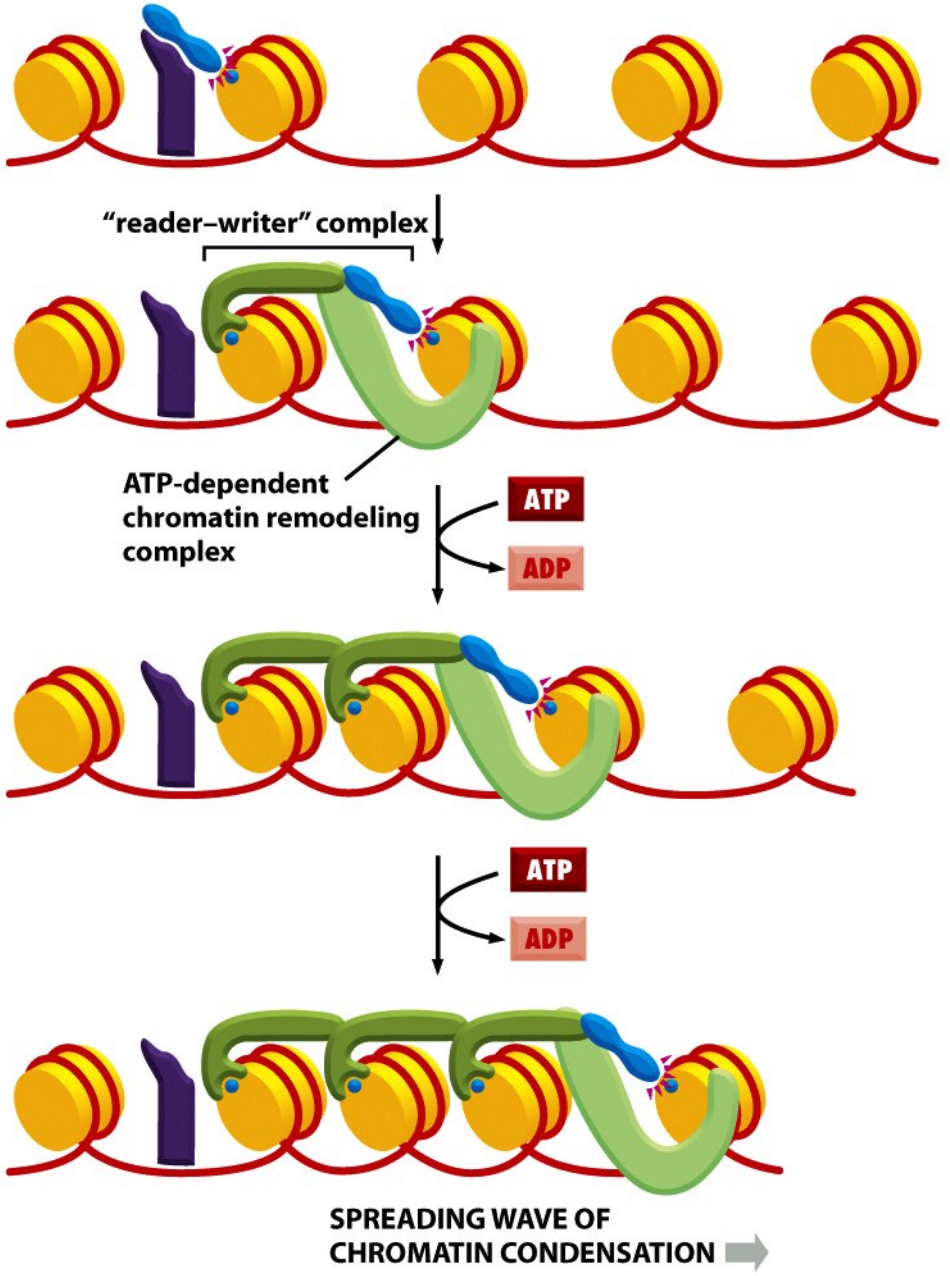


Figure 4-46a *Molecular Biology of the Cell* (© Garland Science 2008)

Ways to identify the DNA interacting with a specific protein or modified protein form, across a large number of sites

- Chromatin immune precipitation
  - Uses antibody to find an *unknown* DNA based on complex with known factor
  - Question: how do you sample all the DNAs you get?

- Depends on what you are looking for:
  - Single region may be detected by probing (but what size will it be? After restriction digest & gel electrophoresis?)
  - Array methods: see where your enriched DNA binds to array of probes to regions of known genomic sequence (ChIP-chip)
  - Direct sequencing methods (ChIP-seq)
  - Can display results vs. whole genome browser, many tracks

Figure out meaning of marks by global correlation with known promoters, RNA polymerase, and coactivators

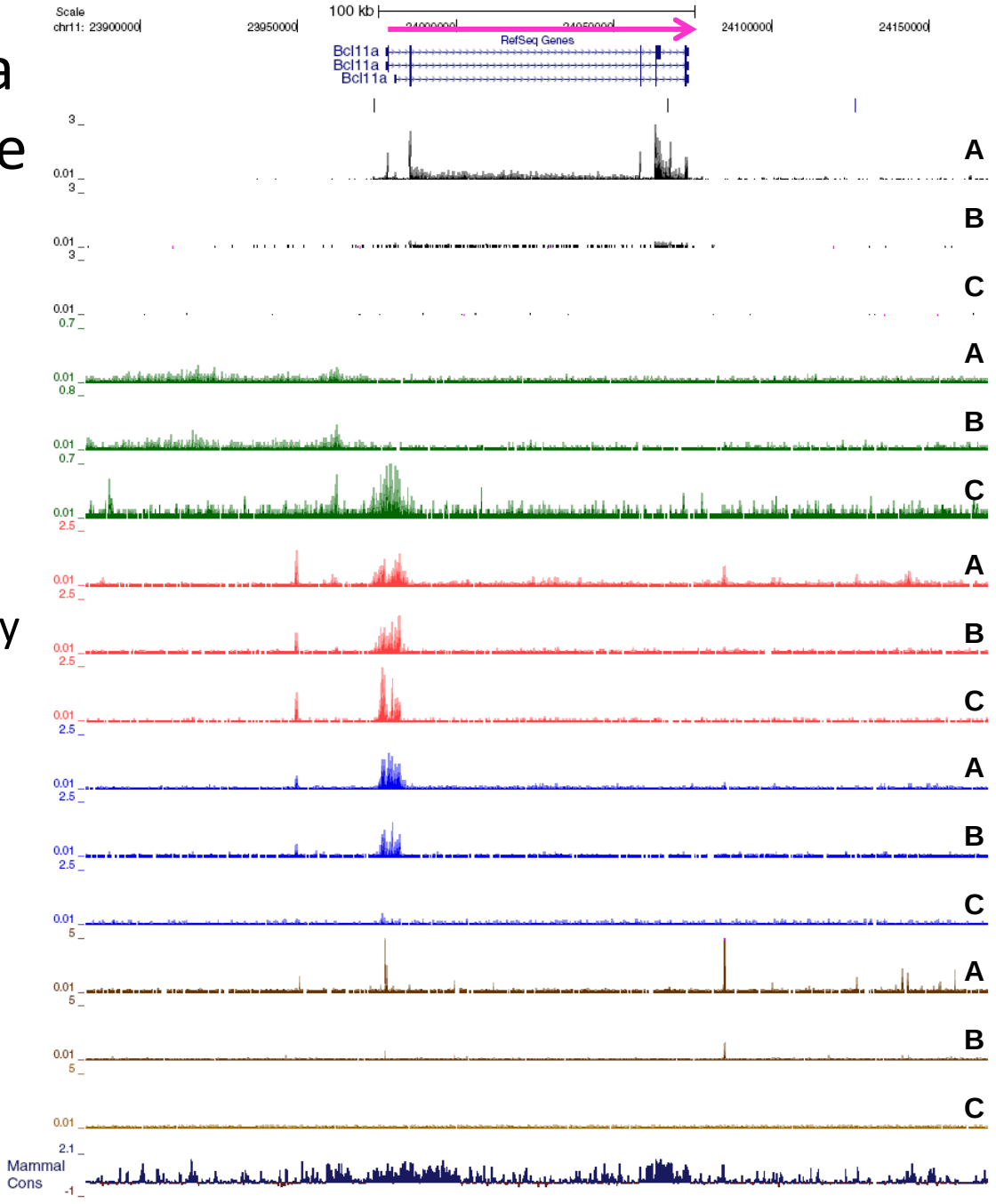
- Heintzman et al (2007): early analysis by “ChIP-chip”
- Ex post facto correlation of sequences retrieved by one antibody with sequences retrieved by another
- Correlation with known genome annotation
- Correlation with measured RNA polymerase II, binding of important coactivator proteins like p300 (to be described much more next lecture)

Ways to learn a lot about a gene in a differentiating cell type…

Transcription levels, repression marks, candidate cis-reg elements, & how they change, stages A-C

(If you have a lot of sample and $$)

Data from RNA-seq, ChIP-seq analysis of *Bcl11a* gene in immature T cells (Jingli Zhang et al., Cell 2012)



Expressed RNA: A to B down, C off

H3K27me3 "repression"

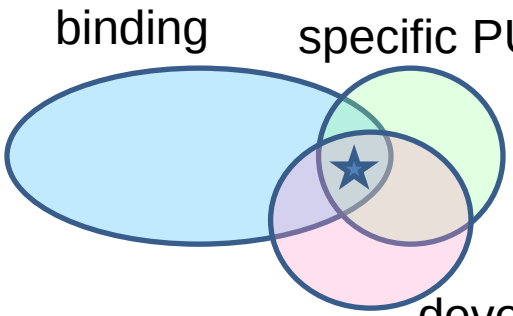
H3K4me2 "accessibility"

H3K(9,14) Ac "activation"

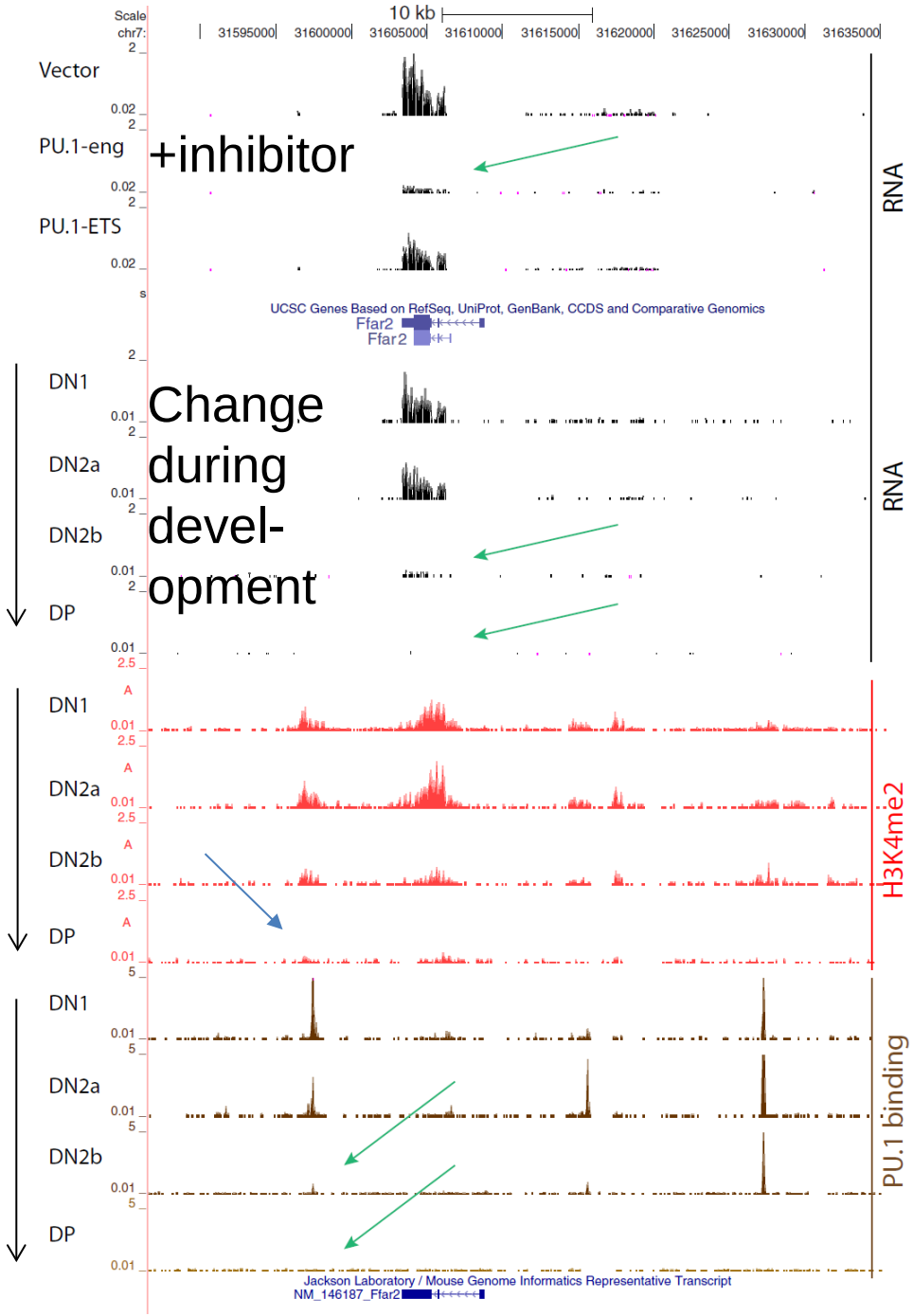
Another TF binding (PU.1): present in A, then leaving

Changes in "permissive" histone marks like H3K4me2 at sites of changing TF binding can show sites where the TF plays an important role:

Developmentally dynamic regulation of *Ffar2* by PU.1

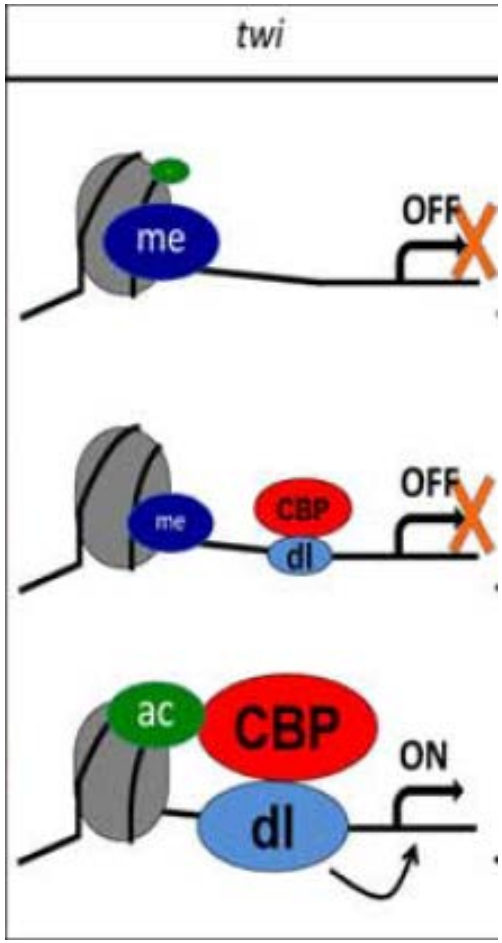


(Champhekar et al., 2015 Genes Dev)



Epigenetic marks "work" as predictors because they are ***results*** of early TF action, which recruit histone modifying enzymes… ***then*** they facilitate activity by later-arriving TFs

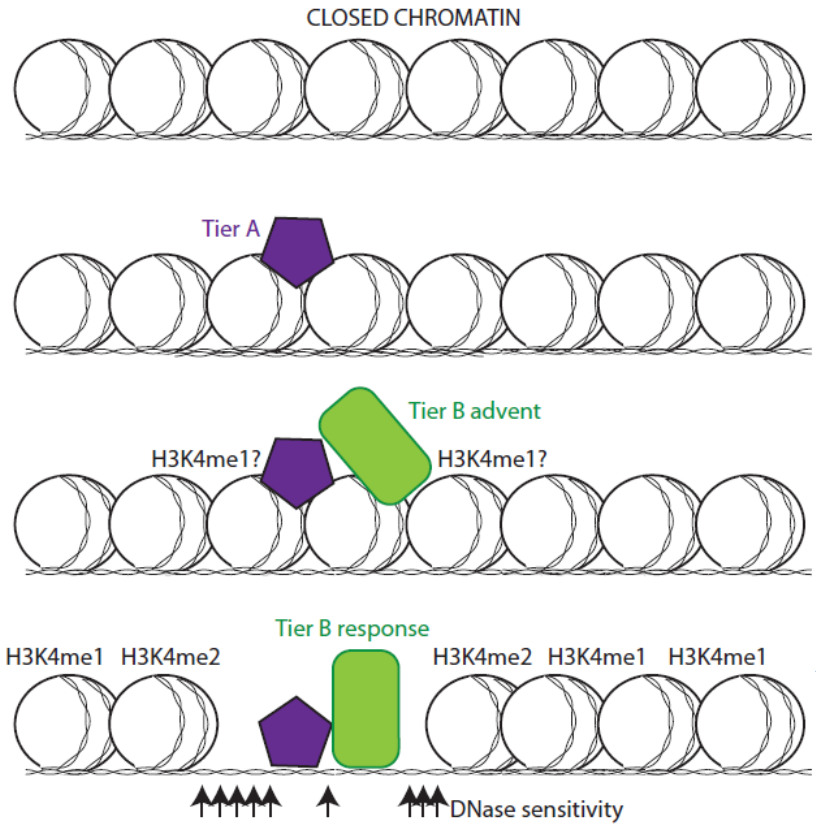
Amount of modifier recruited depends on TF binding and previous modification state



(Holmqvist…Manner-vik, PLoS Genet 2012, ***8:*** e1002769*)*

- Transcription factors can recruit MLL – histone H3K4 methyl transferases
- Transcription factors can recruit CBP or (E)P300 – histone H3K27 acetyl transferases
- Transcription factors can recruit Gcn5 (Kat2a) histone H3K9 or K14 acetyl transferases
- Transcription factors can recruit Utx or Jmjd3 histone H3K27 *demethylases:* remove repression

A stepwise interaction of transcription factors with chromatin can open & activate cis-reg elements



An open enhancer… but not yet activating a gene

Only now does the enhancer activate the target promoter