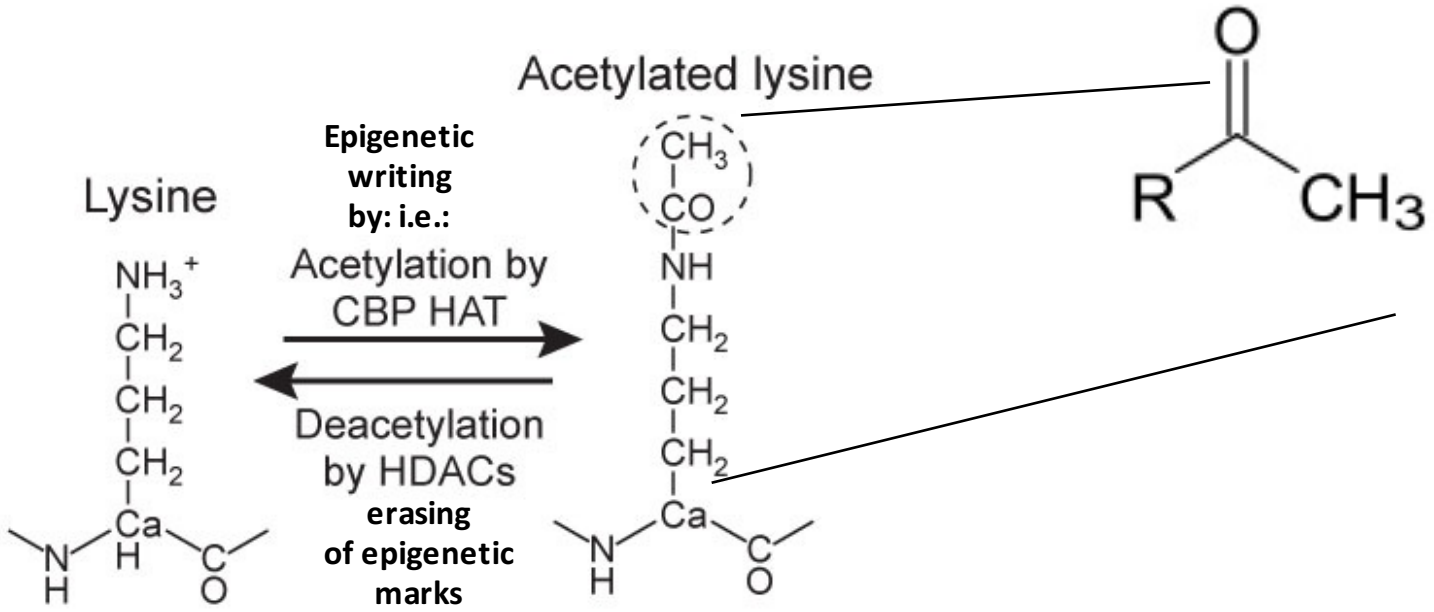
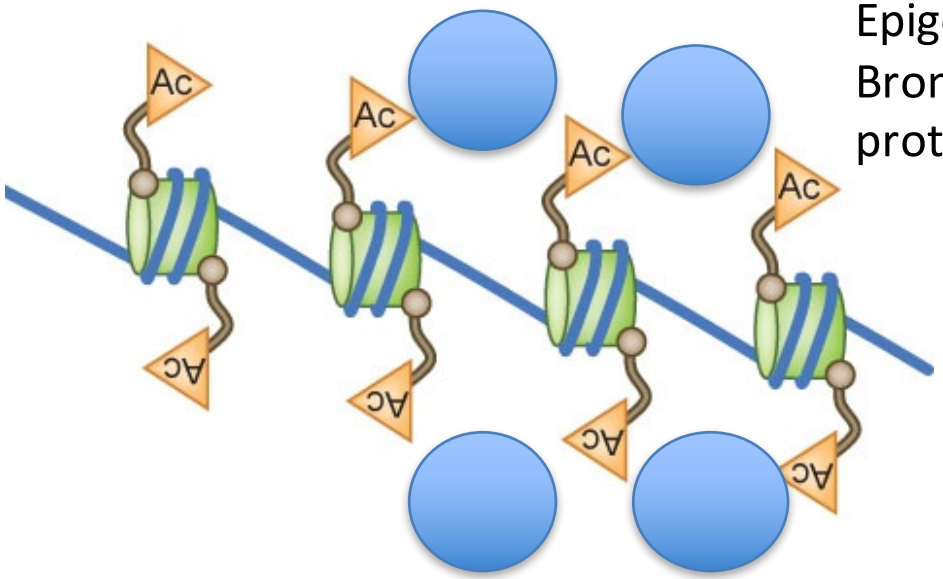
HISTONE ACETYLATION AND DEACETYLATION

Acetylation





Epigenetic reading by Bromo-domain proteins

Transcriptional Activation

Families of Histone acetyltransferases

5 conserved families of histone acetyl transferases (HATs)
also called Lysine acetyltransferases (KATs) → acetylate lysine residues

Table 1. The five major HAT families

Major HAT subfamilies	Prominent members	Key structural and biochemical properties
HAT1	yHat1	Member of the GNAT family Amino- and carboxy-terminal segments used for histone substrate binding Requires the yHat2 regulatory subunit for maximal catalytic activity
Gcn5/PCAF	yGcn5 hGCN5 hPCAF	Member of the GNAT family Uses a ternary complex catalytic mechanism Amino- and carboxy-terminal segments used for histone substrate binding
MYST	yEsa1 ySas2 ySas3 hMOZ dMof hMOF hTIP60 hHBO1	Uses a ping-pong catalytic mechanism Requires autoacetylation of a specific lysine at the active site for cognate histone acetylation
p300/CBP	hp300 hCBP	Metazoan-specific, but shows structural homology with yRtt109 Uses a ternary Theorell–Chance (hit-and-run) catalytic mechanism Contains a substrate-binding loop that participates in AcCoA and lysine binding Contains an autoacetylation loop that requires lysine autoacetylation for maximal catalytic activity
Rtt109	yR11109	Fungal-specific, but shows structural homology with p300 Contains a substrate-binding loop that participates in AcCoA and probably also lysine binding Requires autoacetylation of a lysine residue near the active site for maximal catalytic activity Requires one of two histone chaperone cofactors (Asf1 or Vps75) for maximal catalytic activity and histone substrate specificity

y, yeast; h, human; GNAT, Gcn5-related *N*-acetyltransferase.

Families of Histone acetyltransferases

Best studied HATs

Coding Gene	Site of Histone Modification
HAT 1	H2AK5, H4K5, H4K12
GCN5	H3K9, H3K14, H3K56, H4K5, H4K8, H4K12, H4K16, H4K91 **H3K4**
PCAF	H3K9, H3K14
CBP	H3K14, H3K18, H3K27, H3K56, H4K5, H4K8, H4K12, H4K16
P300	H3K14, H3K18, H3K27, H3K56, H4K5, H4K8, H4K12, H4K16
TAF1	H3K14
TIP60	H2AK5, H4K5, H4K8, H4K12, H4K16
MYST3	H3K9, H3K14
MYST4	
MYST2	H3K14, H4K5, H4K8, H4K12
MYST1	H4K16
ELP3	H3K9, H3K18
GTF3C4	H3K14
NCOA1	H3K14
NCOA3	H3K14
CLOCK	H3K14
CDY1	
CDY2	
CDYL	
MGEA5	H4K8, H3K14
NAT10	

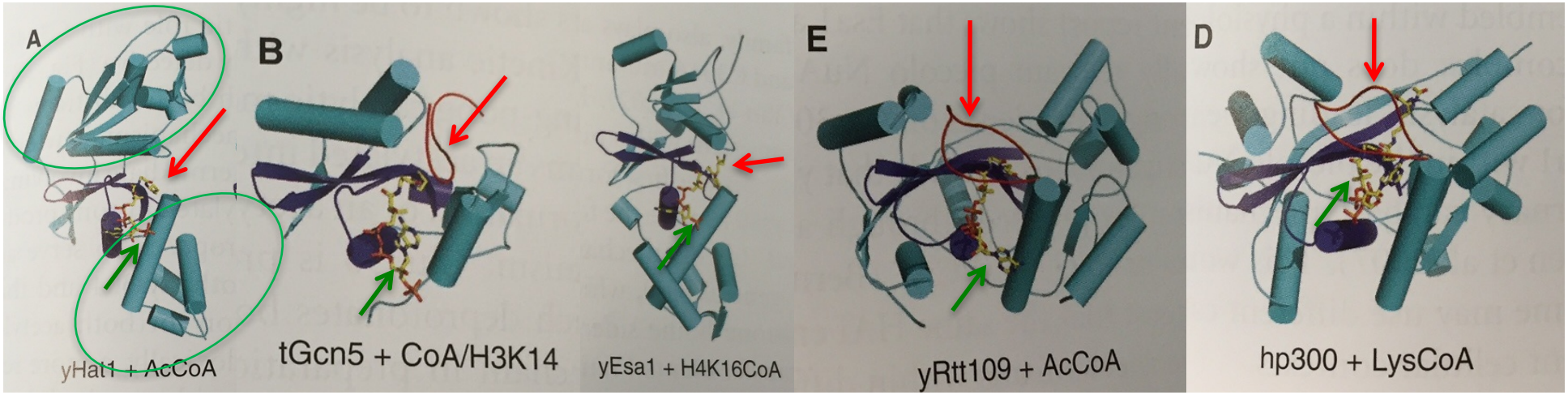
Specificity of HATs

- Many HATs have multiple lysine targets for acetylation
- A subset of lysines can be acetylated by multiple HATs
- Predominantly located in histone tails
- A subset of HAT target lysines can also be subjected to methylation → competition between epigenetic information

H3K9 → me = silent
H3K9 → ac = active

H3K27 → me = silent
H3K27 → ac = active

Structures of major HAT families



HAT1 Fam.	GCN5/PCAF Fam.	MYST Fam.	P300/CBP Fam.	Rtt109 Fam.

HATs contain a HAT domain and structurally divergent ammino-and carboxy-terminal regions

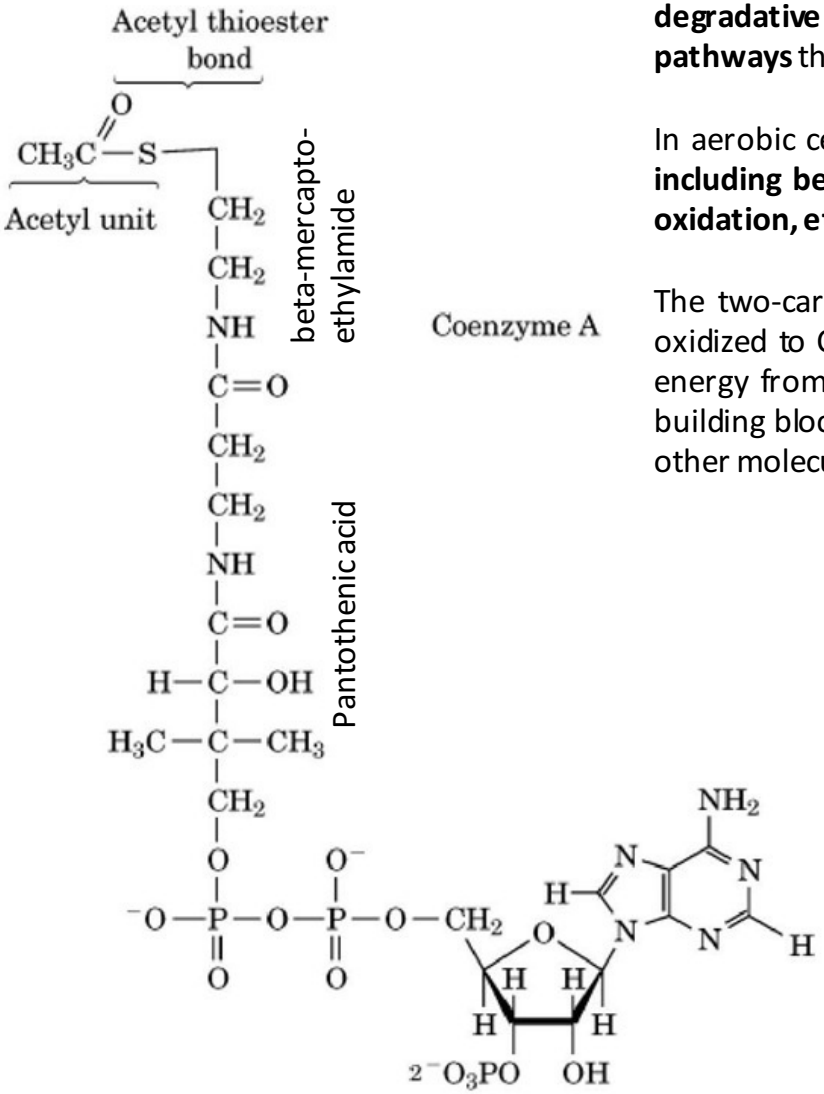
HAT domains are structurally similar and contain: 3 stranded beta-sheet and a long alpha helix
N- and C- terminal domains are divergent between HAT families

HATs contain clefts that allow the histone-substrate to access the central core domain and allow catalysis to occur (red arrow that points on red-line=histone tail)

The central core domain makes interaction with the Acetyl-CoenzymeA (AcCoA) co-factor (green arrow).

Acetyle CoA contains an acetyl group that is transferred by the HAT domain to the histone lysine residue located in the N-terminal histone tail

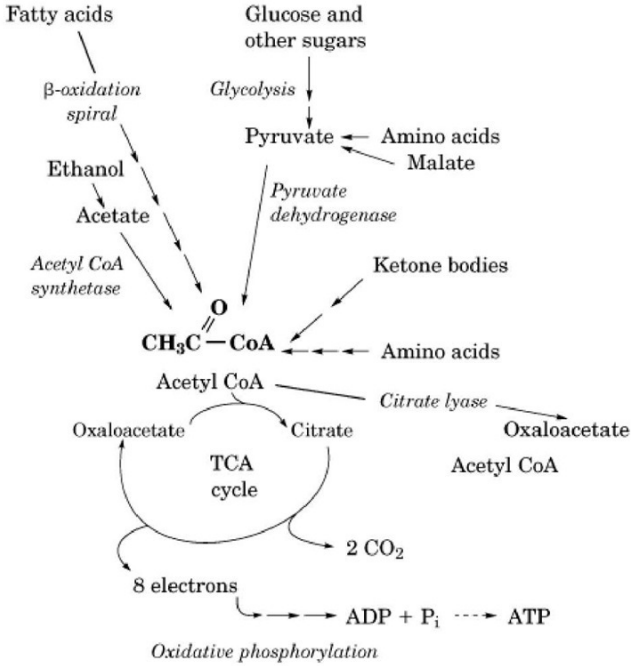
The chemistry of acetyl-transferases



Acetyl coenzyme A (acetylCoA) consists of a two-carbon activated acetyl unit attached to coenzyme A in thioester linkage. AcetylCoA is **central to energy generation from the degradative pathways of oxidative fuel metabolism and to a number of biosynthetic pathways** that utilize the activated two-carbon acetyl unit.

In aerobic cells, it is the **product of all the major catabolic pathways of fuel metabolism, including beta-oxidation of fatty acids, ketone body degradation, glycolysis and pyruvate oxidation, ethanol oxidation, and the oxidative degradation of many amino acids**.

The two-carbon acetyl unit of acetylCoA formed from these pathways can be completely oxidized to CO2 in the **tricarboxylic acid cycle (TCA cycle**), thus providing aerobic cells with energy from the complete oxidation of fuels. The acetyl unit of acetylCoA is also the basic building block of fatty acids, cholesterol, and other compounds, and it can be transferred to other molecules in acetylation reactions (eg, synthesis of N-acetylated sugars).



Catalytic mechanisms of HATs

1. GNAT family HATs

The Gcn5-related N-acetyltransferase (GNAT) family includes **Gcn5, PCAF, Hat1**, Elp3, Hpa2, Hpa3, ATF-2, and Nut1

Ordered sequential wherein both substrates (acetyl-CoA and histone) must bind to form a **ternary complex**

2. MYST family

MOZ, Ybf2 (Sas3), Sas2, and Tip60 [«MYST»], Esa1, **MOF**, MORF, and HBO1

Studies of yeast Esa1 from the MYST family of HATs have revealed a **ping-pong mechanism** involving conserved glutamate and cysteine residue

3. p300/CBP family

p300, CBP

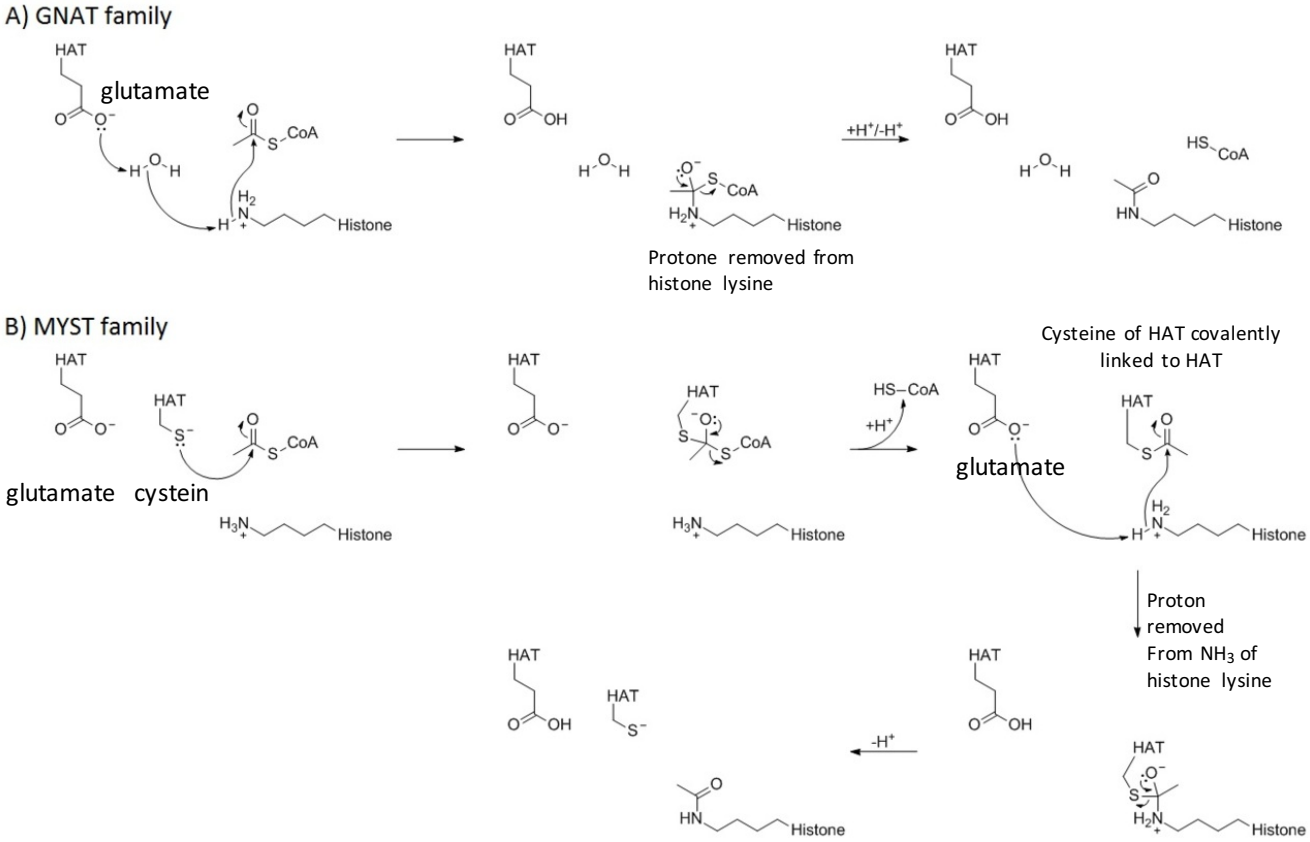
Catalysis by Theorell-Chance or **"hit-and-run"** acetyl transfer mechanism.

4. Rtt109

Not yet understood

The chemistry of acetyl-transferases

Ternary complex



Ping-pong mechanism

Nucleophilic attach on carbonyl group of acetyl-CoA to transfer acetyl group tp lysine

Members of the GNAT family have a **conserved glutamate residue** that acts as a general base for catalyzing the nucleophilic attack of the lysine amine on the acetyl-CoA thioester bond. These HATs use an ordered sequential bi-bi mechanism wherein **both substrates (acetyl-CoA and histone) must bind to form a ternary complex with the enzyme before catalysis can occur. Acetyl-CoA binds first, followed by the histone substrate.** A conserved glutamate residue (**Glu173 in yeast Gcn5) activates a water molecule for removal of a proton from the amine group on lysine,** which activates it for direct **nucleophilic attack on the carbonyl carbon** of enzyme-bound acetyl-CoA. After the reaction, the acetylated histone is released first, followed by CoA.

Studies of yeast Esa1 from the **MYST family** of HATs have revealed a **ping-pong mechanism involving conserved glutamate and cysteine residues.**
The first part of the reaction involves the formation of **a covalent intermediate in which a cysteine residue becomes acetylated following nucleophilic attack of this residue on the carbonyl carbon of acetyl-CoA.** Then, a **glutamate residue acts as a general base to facilitate transfer of the acetyl group from the cysteine to the histone substrate in a manner analogous to the mechanism used by GNATs**.

HATs have the same biochemical function
But can use slightey different chemical reactions to acetylate histones
Reason: reaction is very simple and Acetyl-CoA is very reactive

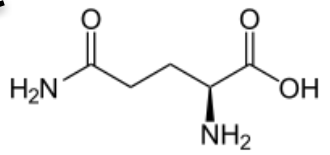
A **ternary complex** is a protein complex containing three different molecules that are bound together. In structural biology, ternary complex can also be used to describe a crystal containing a protein with two small molecules bound, for example cofactor and substrate;

The chemistry of acetyl-transferases – Gcn5/PCAF Family

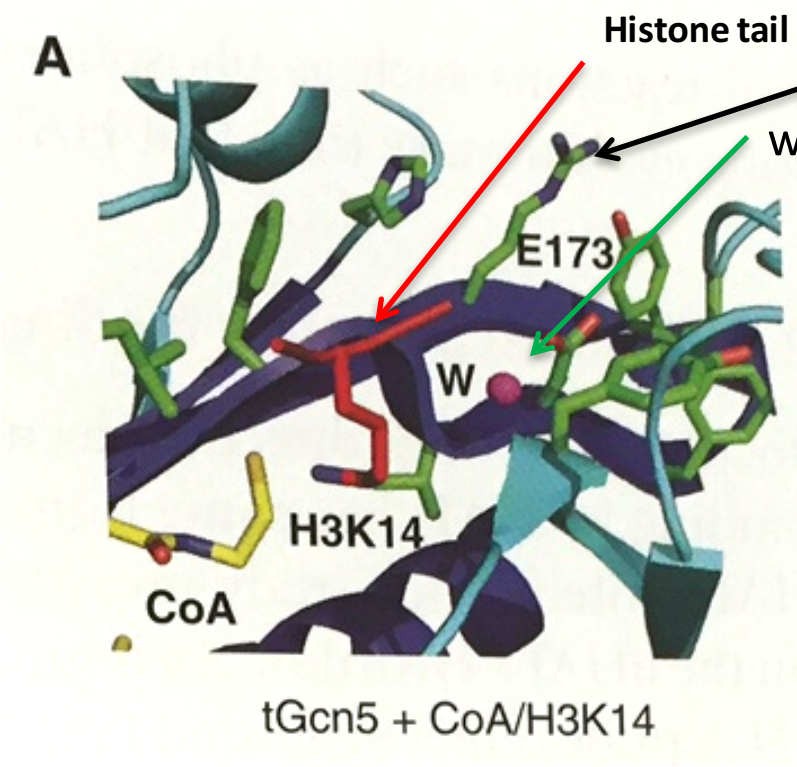
An example:

GNAT family: Gcn5/PCAF/HAT1 – histone H3K14

Glutamate 173



Ternary complex



tGcn5 + CoA/H3K14

Gcn5: glutamate at position 173 (E173) is central for HAT activity.

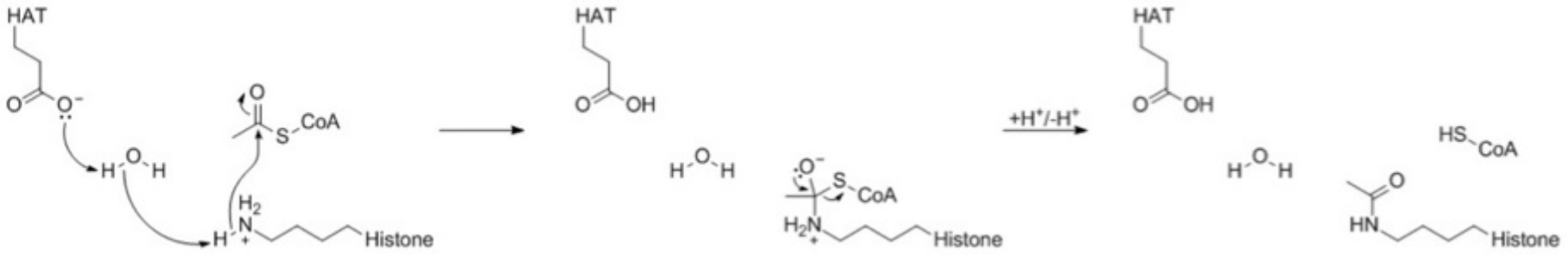
Point mutation at E173 → HAT activity is reduced 300-fold

Glutamate acts as general base for catalysis through a well ordered water molecule in the active site. Glutamate is located in a hydrophobic pocket that supports proton extraction

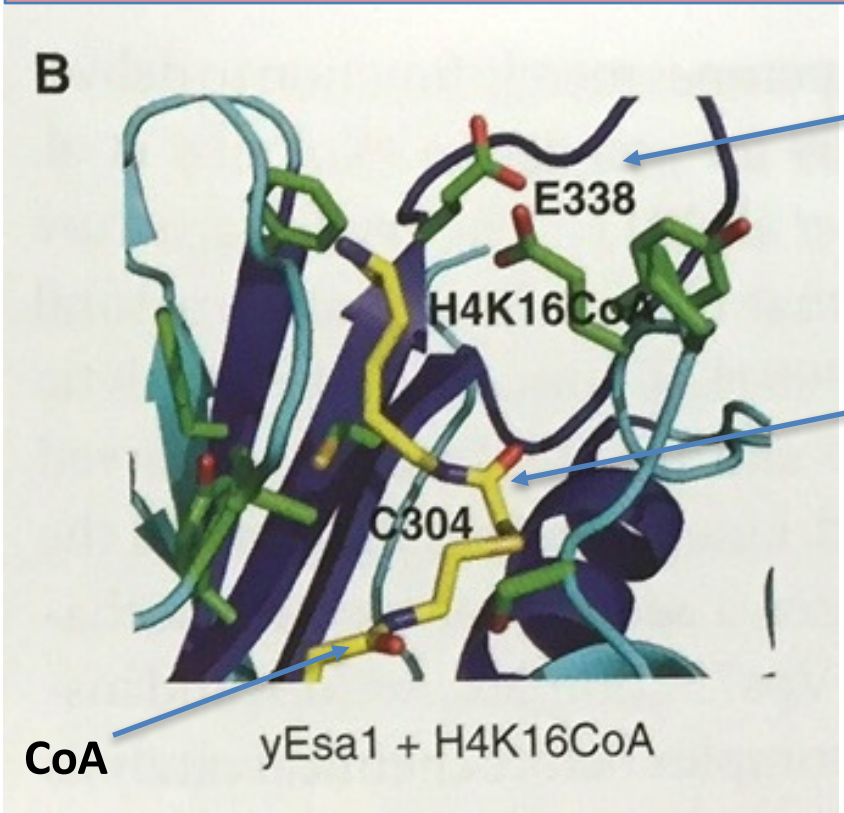
Gcn5 functions through a ternary complex mechanism in which both substrates (lysines and Acetyle-CoA) must be bound to the enzyme before catalysis can occur.

This involves a de-pronotation of the lysine substrate By the glutamate (E173) thus facilitating the direct transfer of the acetyl group from AcCoA to the lysine side chain.

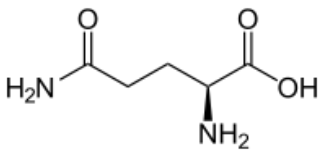
A) GNAT family



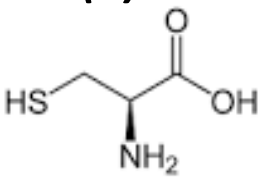
The chemistry of acetyl-transferases – MYST Family



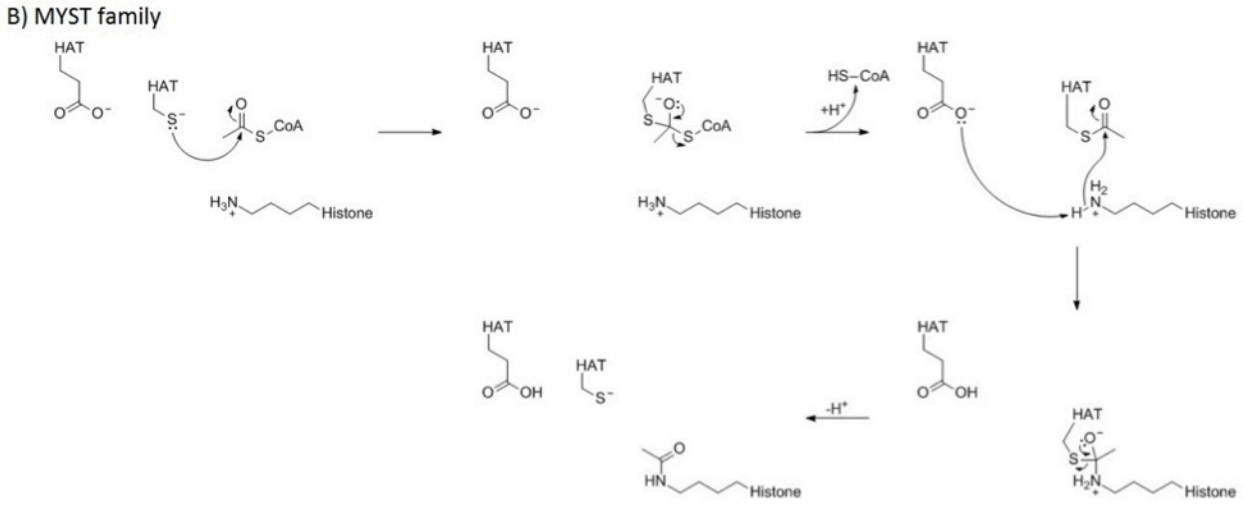
Glutamte (E)



Cystein (C)



MYST family: Important residues: Glu338 and Cys304. Mutations at these residue disrupt HAT activity. Studies of yeast **Esa1 from the MYST family of HATs** have revealed a ping-pong mechanism involving conserved glutamate and cysteine residues. **Glu338 serves as a general base to that protonates Cys304** and the histone lysine residue in preparation for both acetylation events. The first part of the reaction involves **the formation of a covalent intermediate in which a cysteine residue becomes acetylated** following nucleophilic attack of this residue on the carbonyl carbon of acetyl-CoA. Then, a glutamate residue acts as a **general base to facilitate transfer of the acetyl group from the cysteine to the histone substrat**e in a manner analogous to the mechanism used by GNATs.



The chemistry of acetyl-transferases – p300/CBP; Rtt109; HAT1 Family

p300/CBP:

- not glutamate residue for driving acetylation reaction;
- One Tyrosine (Tyr) and one Tryptophane (Trp) residues have impact on acetylation
- In human p300, Tyr1467 acts as a general acid and Trp1436 helps orient the target lysine residue of the histone substrate into the active site.

Tyr1467 mutation: 400 fold reduction in catalytic activity: Trp1436 mutation: 50 fold reduction

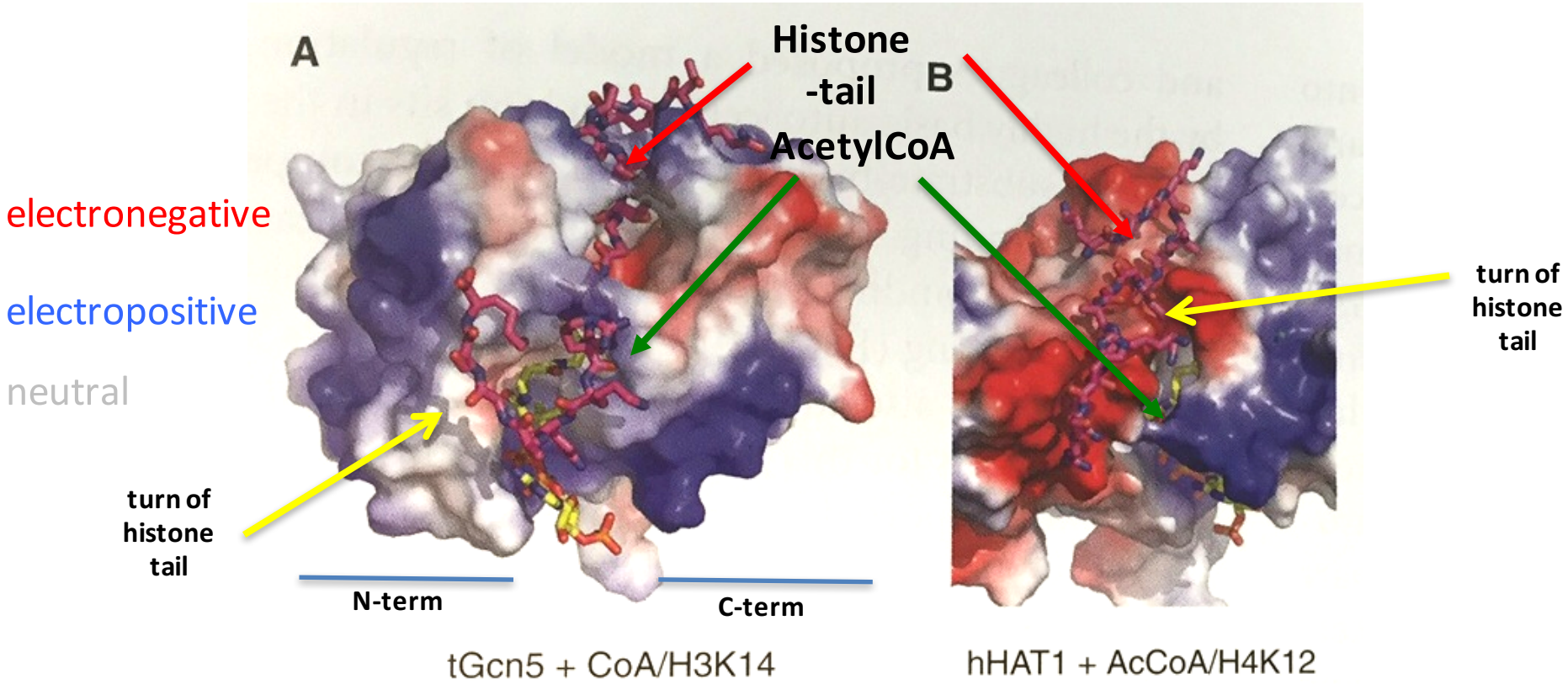
- **Hit and run catalytic mechanism (Theorell Chance)** : distinct from the classic ternary mechanism, TheorellChance is characterized that the ternary complex never accumulates and the steady-state concentrations of the ternary complex is kinetically insignificant. Following the association of Ac-CoA, the protein substrate binds transiently to the p300 surface, allowing the lysine residue to snake through the enzyme active site to receive the acetyl group, followed by rapid protein dissociation. However, answers to several important questions regarding the detailed molecular mechanism of catalysis remain unanswered

Rtt109;HAT1: less characterized

HATs have the same biochemical function
But can use slightly different chemical reactions to acetylates histones
Reason: reaction is very simple and requires "low chemical input"and Acetyle-CoA is very reactive
Long evolution time allowed to form diverse modes of acetylation

Histone substrate binding

To date information only on the binding of Gcn5 to the H3K14 region and hHAT1 to the H4K12 region



Gcn5: Histone tails are fit into a grove formed by the N- and C-terminal domains. Conserved ammino-acids form **hydrogen bonds and van der Waals interaction with H3 histone tails**. H3 tail adopts an ordered structure → basis for major specificity of Gcn5 family for histone H3 tail (Gcn5 activity towards H4 is low)

hHAT1: Histone H4 tail is fit into a grove and forms a turn structure that normally remains extended.
Two conserved hHAT1 residues (Trp199, Tyr225) interact with Gly9 and Lys8 of the turned histone tail.
Conserved amminoacids accommodate the H4 tail into the groove and Bring the K12 residue in vicinity to the active center. Other histone-tails have other peptide sequence context and to not form specific interactions with hHAT1 → specificity for H3 tails

Regulation by auto-acetylation and protein cofactors

HATs are regulated by AUTOACETYLATION and INTERACTION WITH REGUALTORY PROTEINS

1. REGULATORY PROTEINS

→Purified HATs such as GCN5 or PCAF act efficiently on free histone and histone-tail peptides but are less efficient on prepared nucleosomes

→In vivo, HATs function in *multiprotein complexes (HAT + cofactors)* to acetylate histone tails on nucleosomes. Complexes can contain 10-20 subunits that can also be shared amongst different HAT complexes

EXAMPLES: Gcn5 → SAGA complex (yeast); SLIK complex (human)
PCAF → TCTC complex (yeast); STAGA complex (human)

The role of most complex components is to support HAT specificity and activity

→HATs interact with *cofactors* in HAT complexes to increase processivity

EXAMPLE: Sas2 (MYST family) has to interact with Sas2 and Sas4 to have HAT activity

EXAMPLE: Rtt109 has no or little HAT activity; interaction with Vps75 or Asf1 (histone chaperon) increases HAT acitivity (100x) and mediates H3K9/H3K27 Aacetylation (Vps75) or H3K56 acetylation (Asf1)

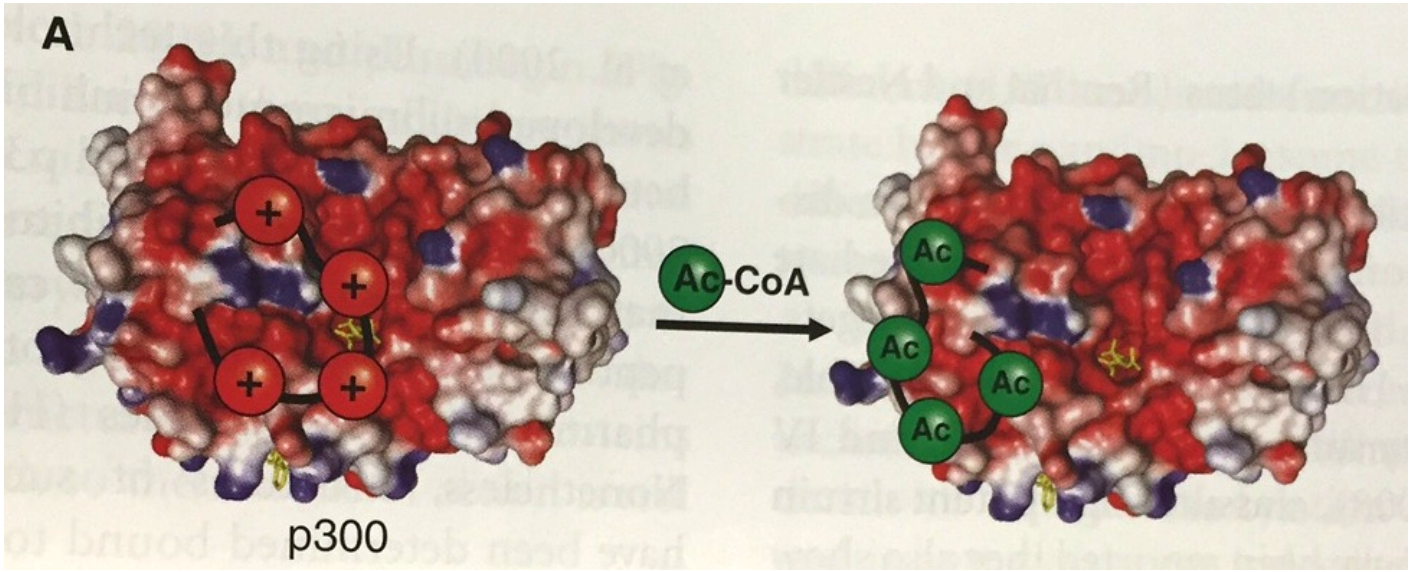
Regulation by autoacetylation and protein cofactors

HATs are regulated by AUTOACETYLATION and INTERACTION WITH REGUALTORY PROTEINS

2. AUTOACETYLATION

→Rtt109, p300/CBP and MYST HAT family members are controlled by auto-acetylation
HYPOACETYLATED HAT: INACTIVE
HYPERACETYLATED HAT: ACTIVE

EXAMPLE: p300 activity is controlled by acetylation of a 40 aa basic loop = autoacetylation loop



Under-acetylated "autoacetylation loop"
blocks substrate (histone tail) binding site of p300

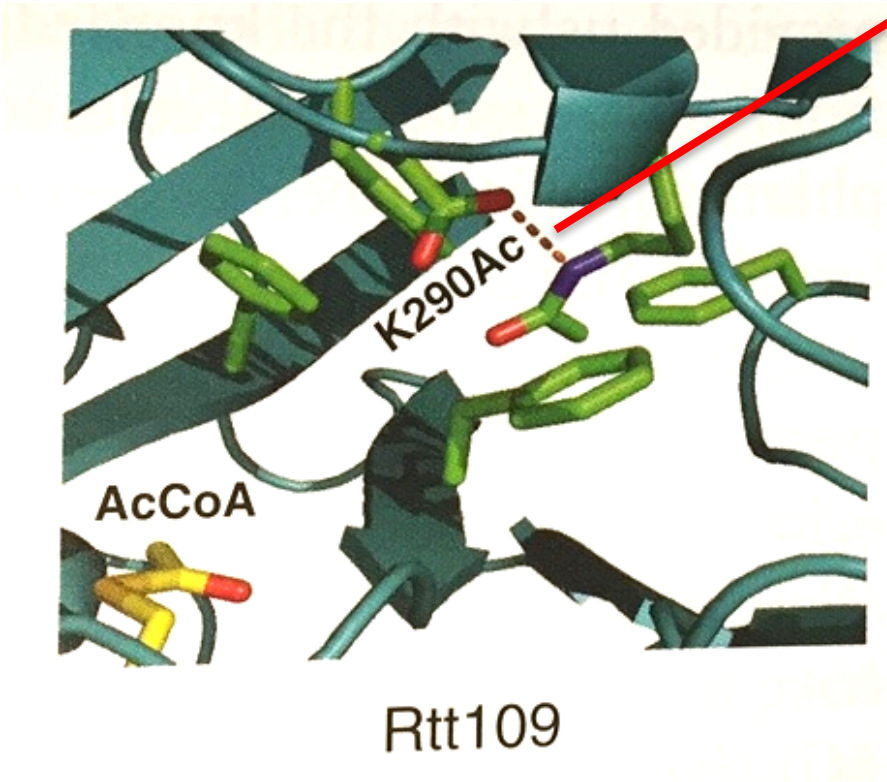
hyper-acetylated "autoacetylation loop"
enhances substrate (histone tail binding site of p300

Regulation by autoacetylation and protein cofactors

HATs are regulated by AUTOACETYLATION and INTERACTION WITH REGUALTORY PROTEINS

2. AUTOACETYLATION

→Rtt109, p300/CBP and MYST HAT family members are controlled by auto-acetylation
→HYPOACETYLATED HAT: INACTIVE
HYPERACETYLATED HAT: ACTIVE



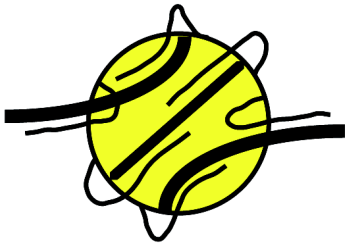
Rtt109:
Acetylation of Lys290 is required for full HAT activity.

WHY?
Acetylation of Lys290 disrupts hydrogen bonds between Lys290 and Asp288. This improves incorporation of Acetyl-CoA

Note: mutations in Asp288 increase HAT activity → presumably improved Acetyl-CoA binding

Impact of HATs on gene expression: ACTIVATORS OF GENE EXPRESSION

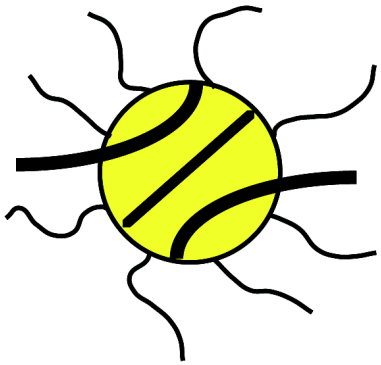
Acetylation induces a conformational change in the core histones



REPRESSED

EXAMPLE

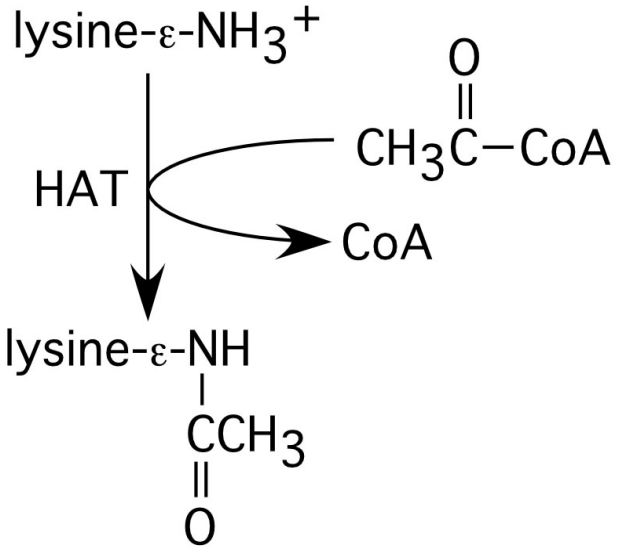
Lysine-rich tails bind tightly to DNA and repress transcription by blocking access of factors to the DNA template.



ACTIVE/COMPETENT

Acetylation of lysine residues alters chromatin structure and allows binding of transcription factors.

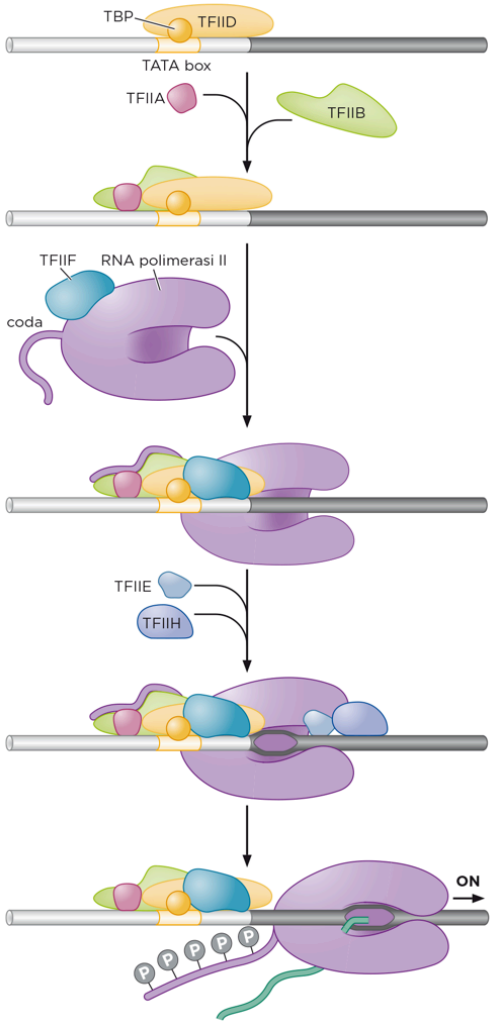
Note: acetylation neutralizes the positive charge of lysine

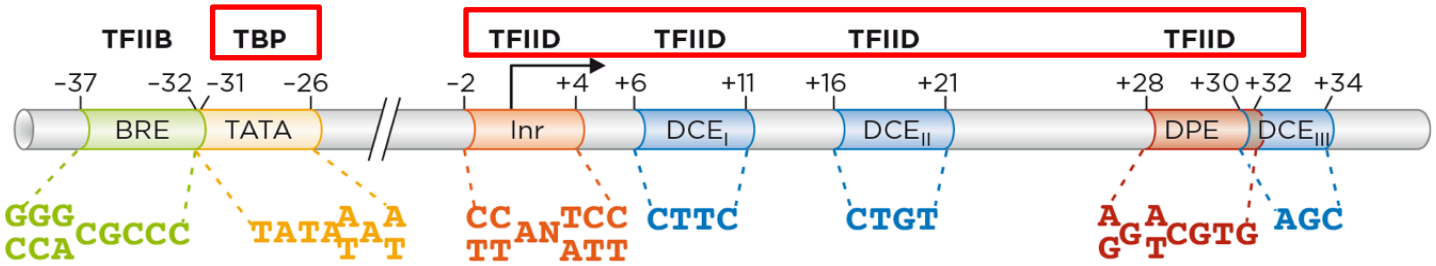


HAT: Histone Acetyltransferase

Transcription by RNA Polymerase II (RNAPII)
The RNAP II core promoter

CENTRAL PROMOTER ELEMNETS + GENERAL TRASNCRIPTION FACTORS
ca 60 nt





BRE	**TATA-box**	**Inr**	**DCE**	**DPE**
TFIIB recognition element	Bound by the TATA-box binding protein (TBP)	Initiator (l'iniziatore)	downstream core element	downstream promoter element

GTFs:

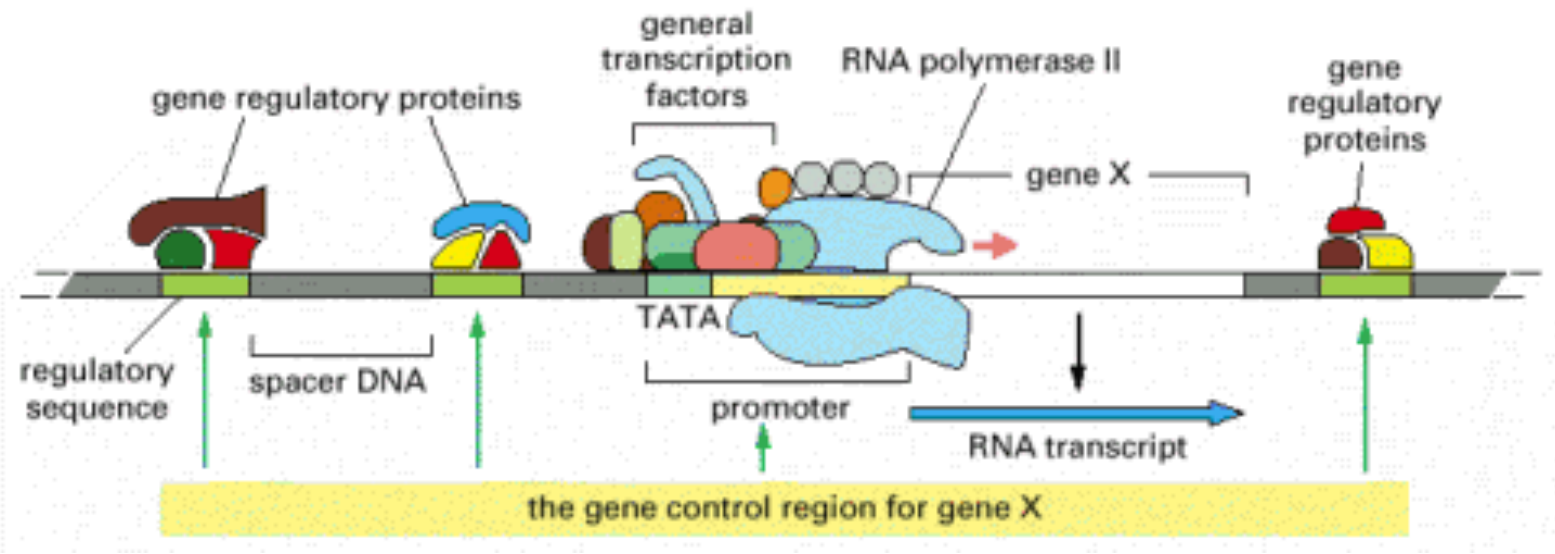
→Help RNA Pol II to bind promoter

→Promoter melting

(→→→ same function like sigma factor in bacterial gene activation)

→ Pass on to elongation phase of transcription

A complex interplay of regualtory sequences and transcription factors control the basal transcription complex



The gene control region of a typical eucaryotic gene.
The ***promoter*** is the DNA sequence where the general transcription factors and the polymerase assemble. The ***regulatory sequences*** serve as **binding sites** for **gene regulatory proteins**, whose presence on the DNA affects the rate of transcription initiation.

These sequences can be located **adjacent** to the promoter, far **upstream** of it, or even **within introns** or **downstream** of the gene. These sites can be bound by cell type specific factors that define promoter specificty or enhance specificity **DNA looping** is thought to allow gene regulatory proteins bound at any of these positions to interact with the proteins that assemble at the promoter.

Whereas the **general transcription factors** that assemble at the promoter are **similar for all** polymerase II transcribed **genes**, the **gene regulatory proteins** and the **locations** of their binding sites relative to the promoter are **different for each gene**.