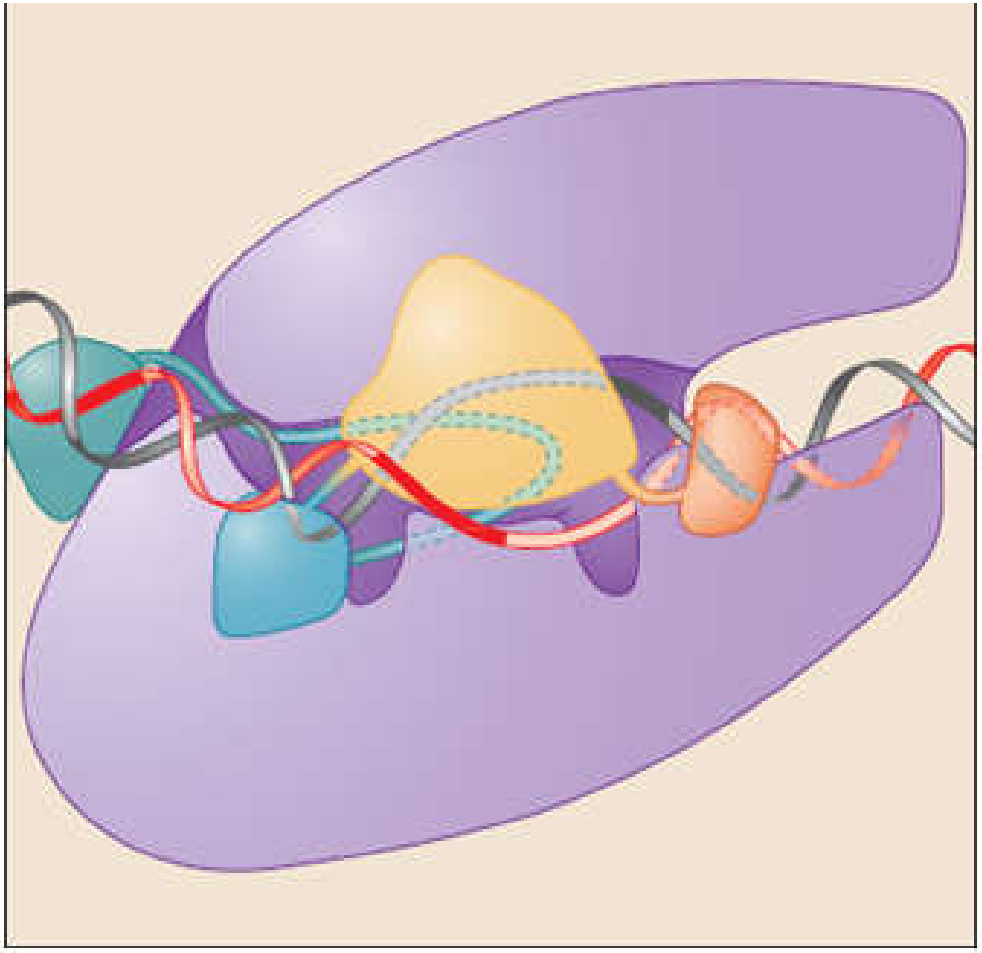
Chapter 6

Transcription in eukaryotes

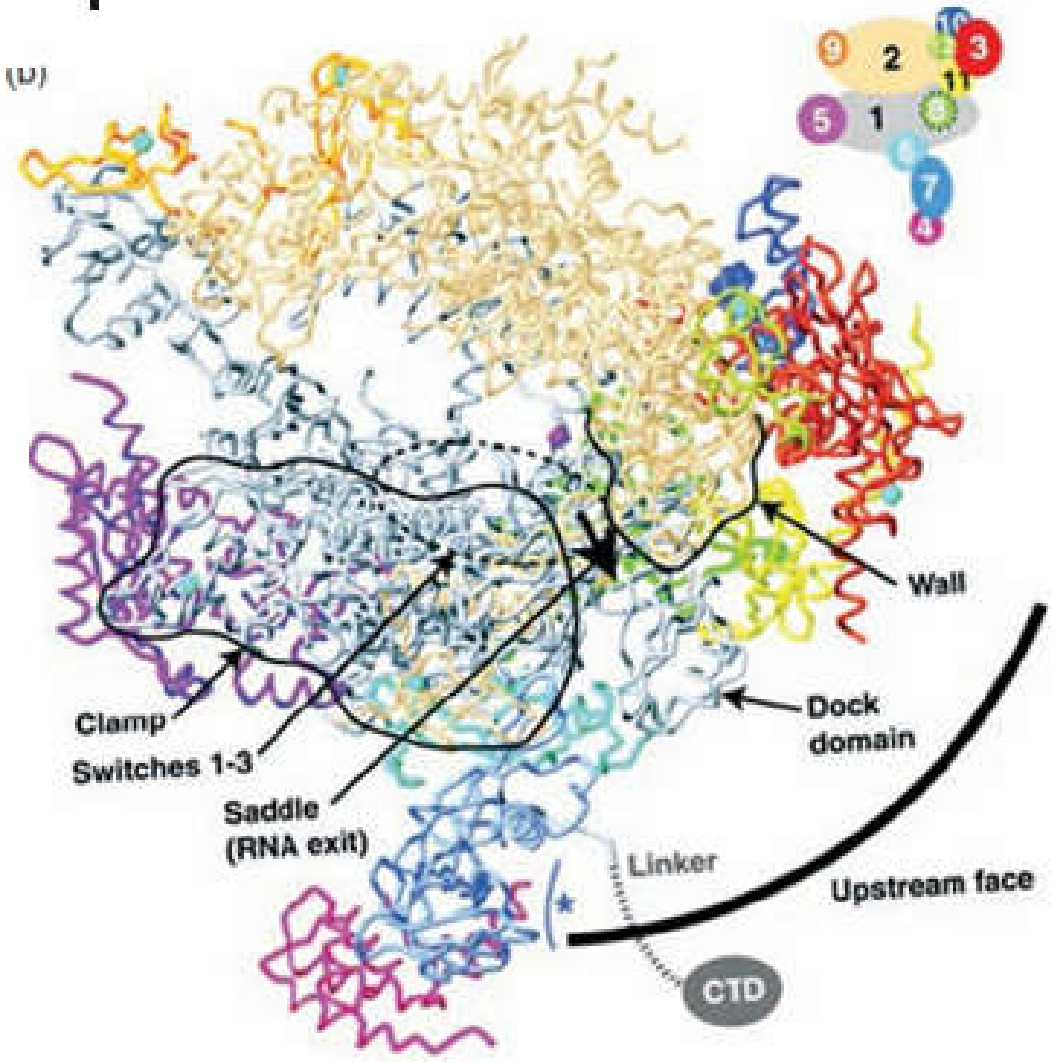


1. RNA polymerases in eukaryotes

1.1 Types of RNA polymerases in eukaryotes

	RNA Pol I	RNA Pol II	RNA Pol III
Transcription	18S, 28S, 5.8S rRNA	hnRNA, some snRNA and miRNA	tRNA, 5S rRNA and some sRNA
Location	Nucleoli (核仁)	Nucleoplasm (核质)	Nucleoplasm (核质)
α-amanitin (鹅膏蕈碱)	Not sensitive (-)	Very sensitive (+ +)	Moderately sensitive (+)

1.2 Subunits of RNA polymerase



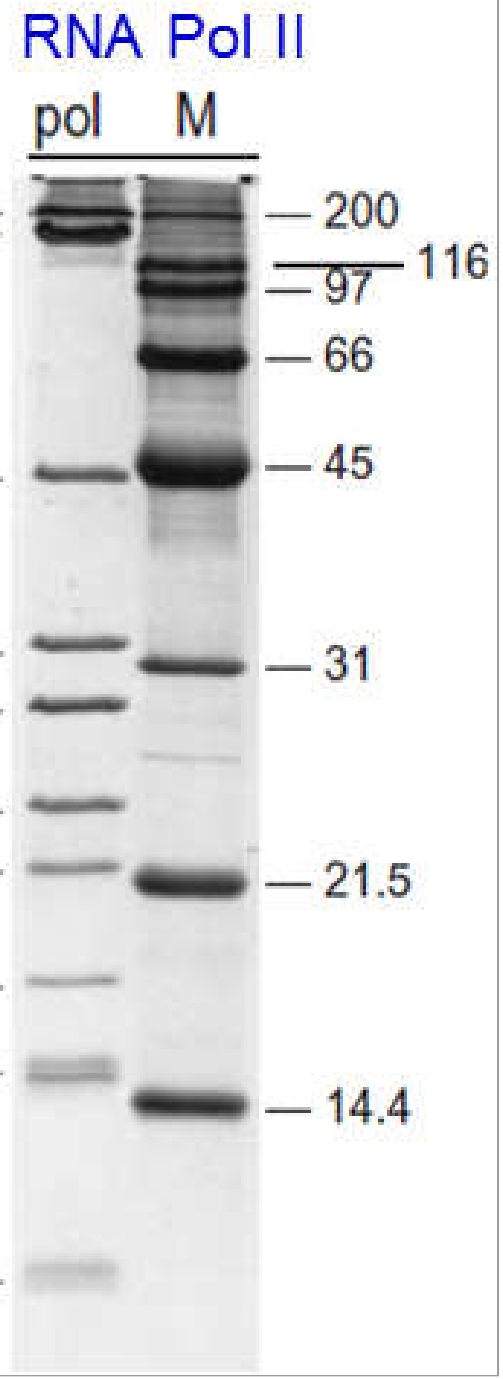
- All the polymerase structures are complex.
- Pol I 14
  Pol II 12 subunits
  Pol III 17
- Each contains two large (>100 KD) subunits, plus a variety of smaller subunits.

Core Subunits

- Related to bacterial subunit β' — Rpb1
- Related to bacterial subunit β — Rpb2
- Related to bacterial subunit α — Rpb3

Common Subunits — Rpb4, Rpb5, Rpb6, Rpb7, Rpb8, Rpb9, Rpb11, Rpb10, Rpb12

- **Rpb1 has a carboxyl-terminal domain (CTD, 羧基末端结构域) consisting of multiple repeats of a heptamer (七聚体). Tyr-Ser^P-Pro-Thr-Ser^P-Pro-Ser^P**
- **Involved in transcription and RNA processing**



2. Transcription factors

- Eukaryotic RNA polymerases are incapable of binding by themselves to their respective promoters.
- **Transcription factors** are proteins that bind to specific DNA sequences, thereby controlling (always enhancing) the transcription efficiency.

转录因子是与特定DNA序列结合(通常)促进基因转录的蛋白质。

- **Transcription factors are grouped into two classes:**

General transcription factor (GTF)

通用转录因子：吸引RNA聚合酶至相应的启动子组装成基础转录复合物，从而启动转录的蛋白质。（启动转录程度弱，仅支持本底水平的转录，但为真核细胞转录必需）

Gene-specific transcription factor or transcription activator

基因特异转录因子或转录激活因子

The GTFs combine with RNA polymerase to form an initiation complex that is competent to initiate transcription.

Class II promoters - Class II GTFs

Class I promoters - Class I GTFs

Class III promoters - Class III GTFs

3. Class II promoters and GTFs

Promoters were recognized by RNA Pol II.

包括两个部分：核心启动子和上游启动子元件。

3.1 Core promoter

3.1.1 Functions:

- Attracts general transcription factors (GTF) and RNA Pol II at a basal level

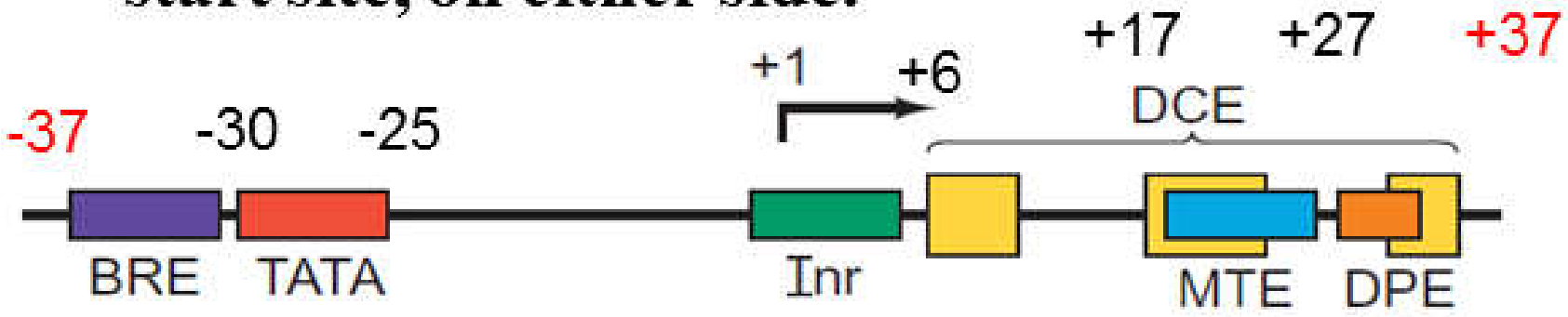
  以本底水平吸引通用转录因子和RNA Pol II

- Sets the transcription start site and directs transcription

  确定转录起始位点并指导转录

3.1.2 Structures of core promoter

- **Lying within ~37 bp of the transcription start site, on either side.**



- TATA box (Goldberg-Hogness box)
- Inr: initiator (起始子/起始元件)
- DCE: downstream core element [MTE: motif ten element (十基序元件), DPE: downstream promoter element]
- BRE: TFIIB recognition element
- At least one of these elements is missing in most promoters.

(1) TATA box

- **Structure:**
  - Consensus sequence: TATA(A/T)A(A/T)
  - Similar to the prokaryotic -10 sequenc. But position: -30 ~ -25.
  - The spacing between the TATA box and the initiation site is important.
  - The sequence around the TATA box is critical.

Function:

Specialized genes (特化基因), which encode proteins made only in certain types of cells, do have TATA boxes.

- In Class II promoters, the TATA box serves as the site where initiation complex assembly.
- Some Class II promoters require the TATA box for function, but others need it only to position the transcription start site.

TATA-less promoters:

- G and C creep in (混入)
- Housekeeping genes (持家基因) and developmentally regulated genes (发育调节基因)
- Many TATA-less promoters have Inr or DPEs that play the same role as a TATA box.

(2) Initiators (Inr, 起始子)

- **Initiators - some Class II promoters have conserved sequences around their transcription start sites that are required for optimal transcription.**
- **Mammalian initiators have the consensus sequence: PyPyΔ(T/A)PyPy. A is the transcription start point.**
- **This initiator can drive basal-level (本底水平) transcription alone, but at a low level.**

3.2 Upstream promoter elements (UPE)

UPE are also called proximal (近侧) promoter elements or upstream regulatory elements (UREs, UCEs).

3.2.1 Location

- **Upstream of core promoters** (-250 ~ -37).
- **GC box**: GGGCGG and CCGCCC, upstream of the TATA box, may not only one copy.
- **CCAAT box**

3.2.2 Function of UPE

- **Help attract GTFs and RNA polymerase**

Mutations or deletions in UPE significantly decreased promoter activity.

- **Bind to relatively gene-specific transcription factors and stimulates transcription**

e.g. Sp1 binds to the GC boxes;

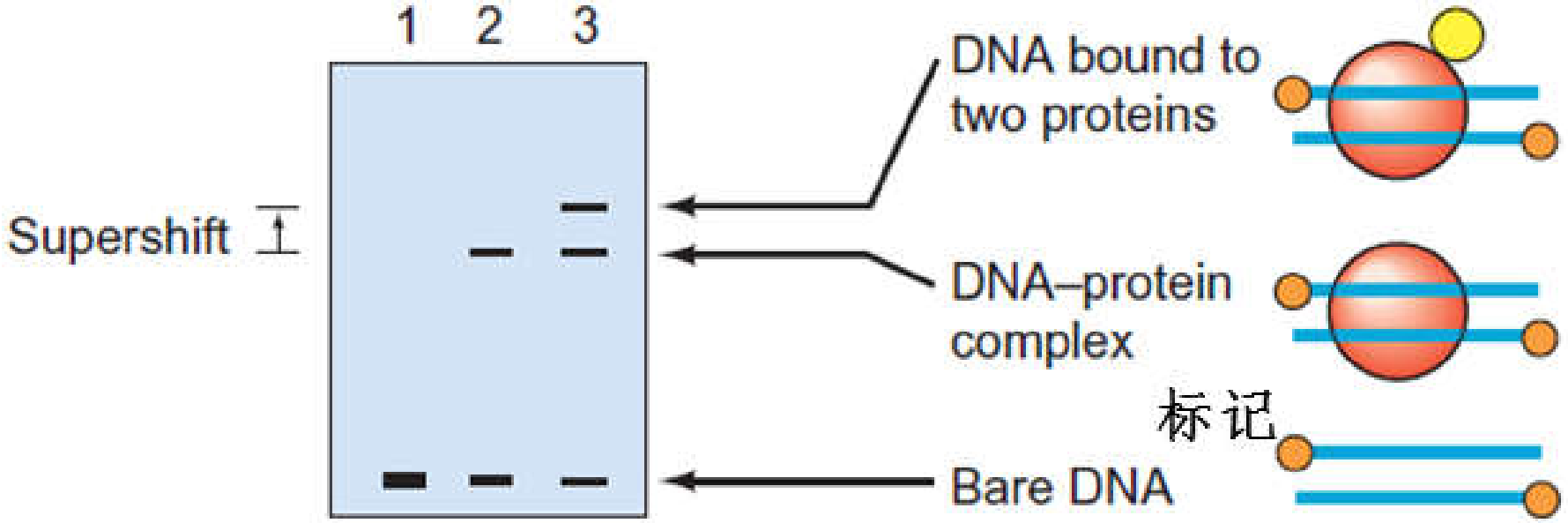
The CCAAT box must bind the CTF (CCAAT-binding transcription factor) to exert (发挥) its stimulatory influence.

3.3 The Class II initiation complex

- The Class II initiation complex contains RNA polymerase II and six GTFs named TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH.
- The Class II GTFs and RNA Pol II bind in a specific order to the growing initiation complex, at least in vitro.

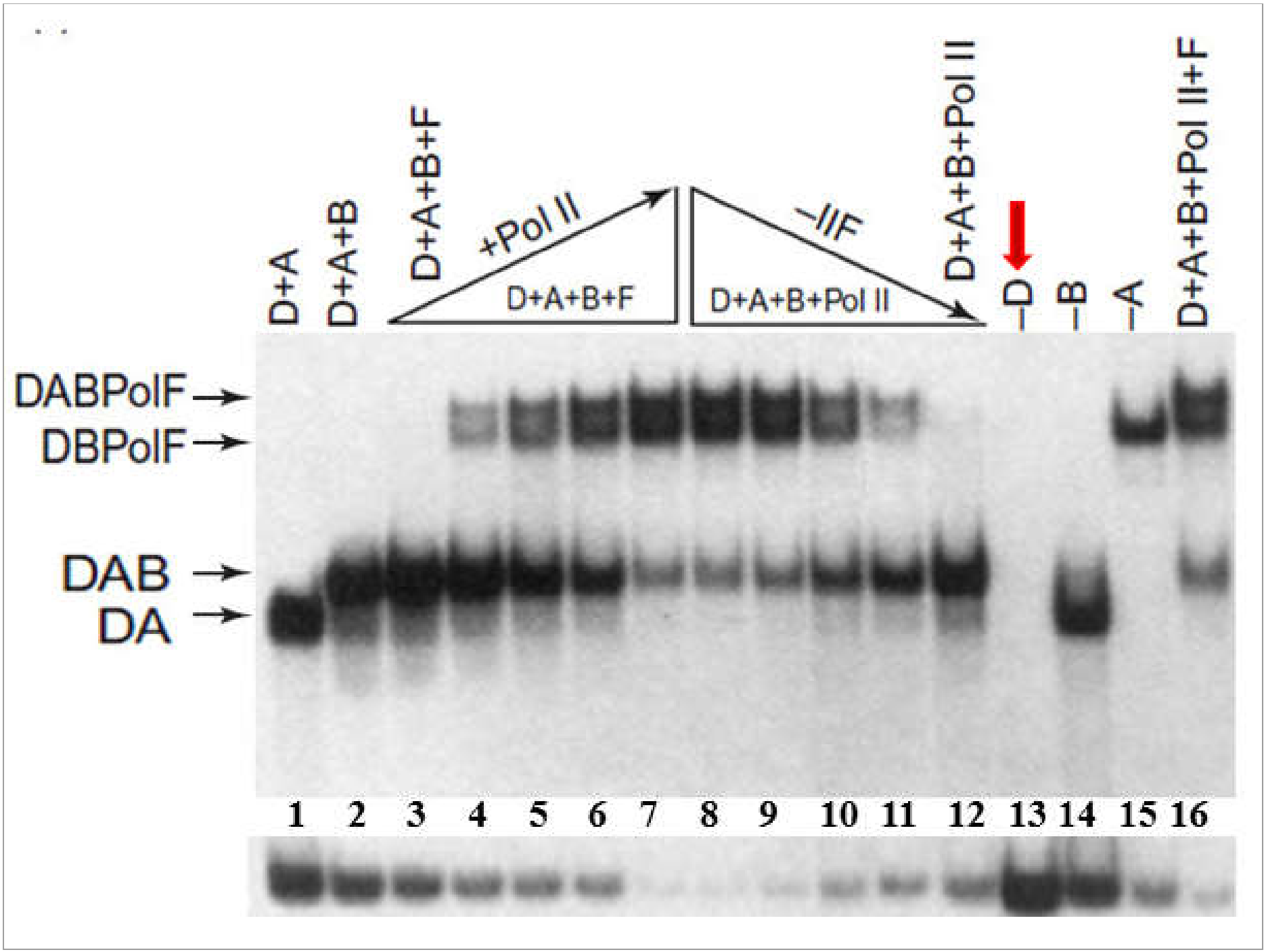
3.3.1 Binding order of Class II GTFs

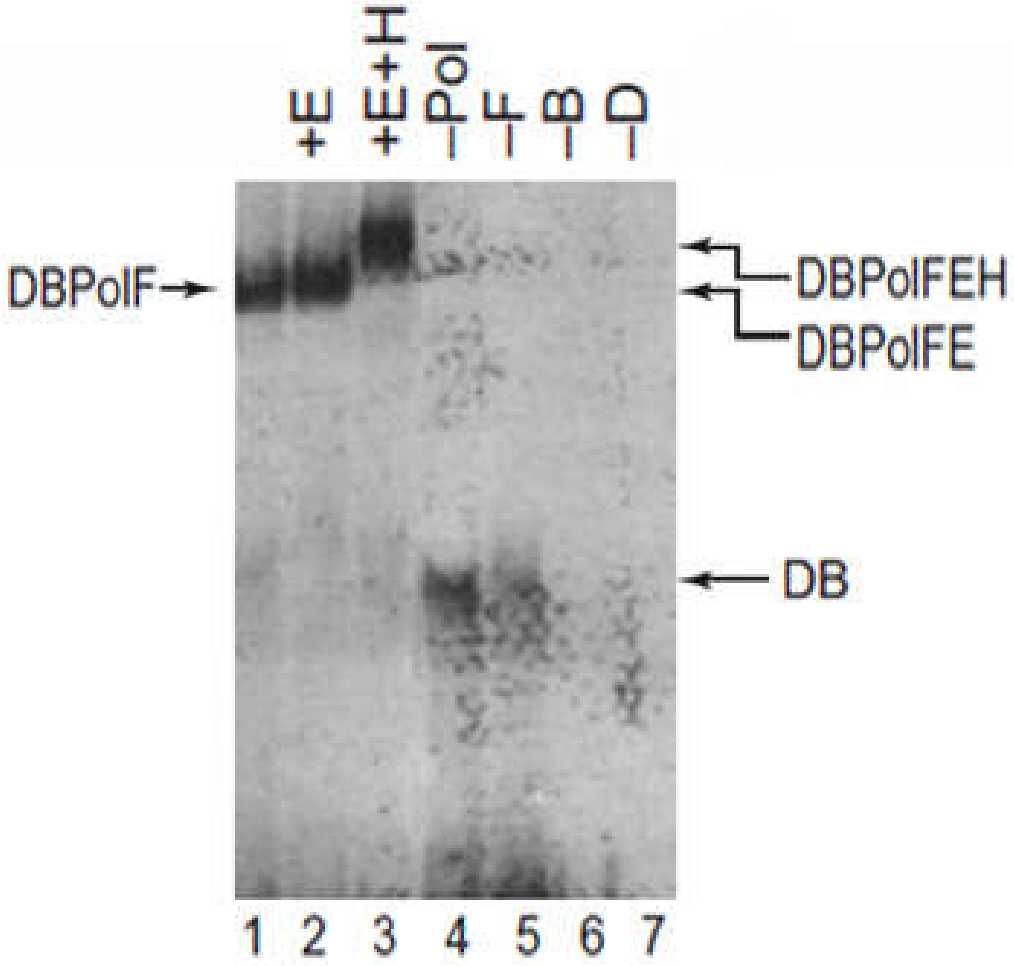
电泳迁移率变动实验 (electrophoretic mobility shift assay, EMSA)，凝胶阻滞



核酸与蛋白质结合降低了核酸分子的电泳迁移率。

DNA-蛋白质相互作用研究方法之三


D+A
D+A+B
D+A+B+F
+Pol II
D+A+B+F
-IIF
D+A+B+Pol II
D+A+B+Pol II
-D
-B
-A
D+A+B+Pol II+F
DABPolF
DBPolF
DAB
DA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

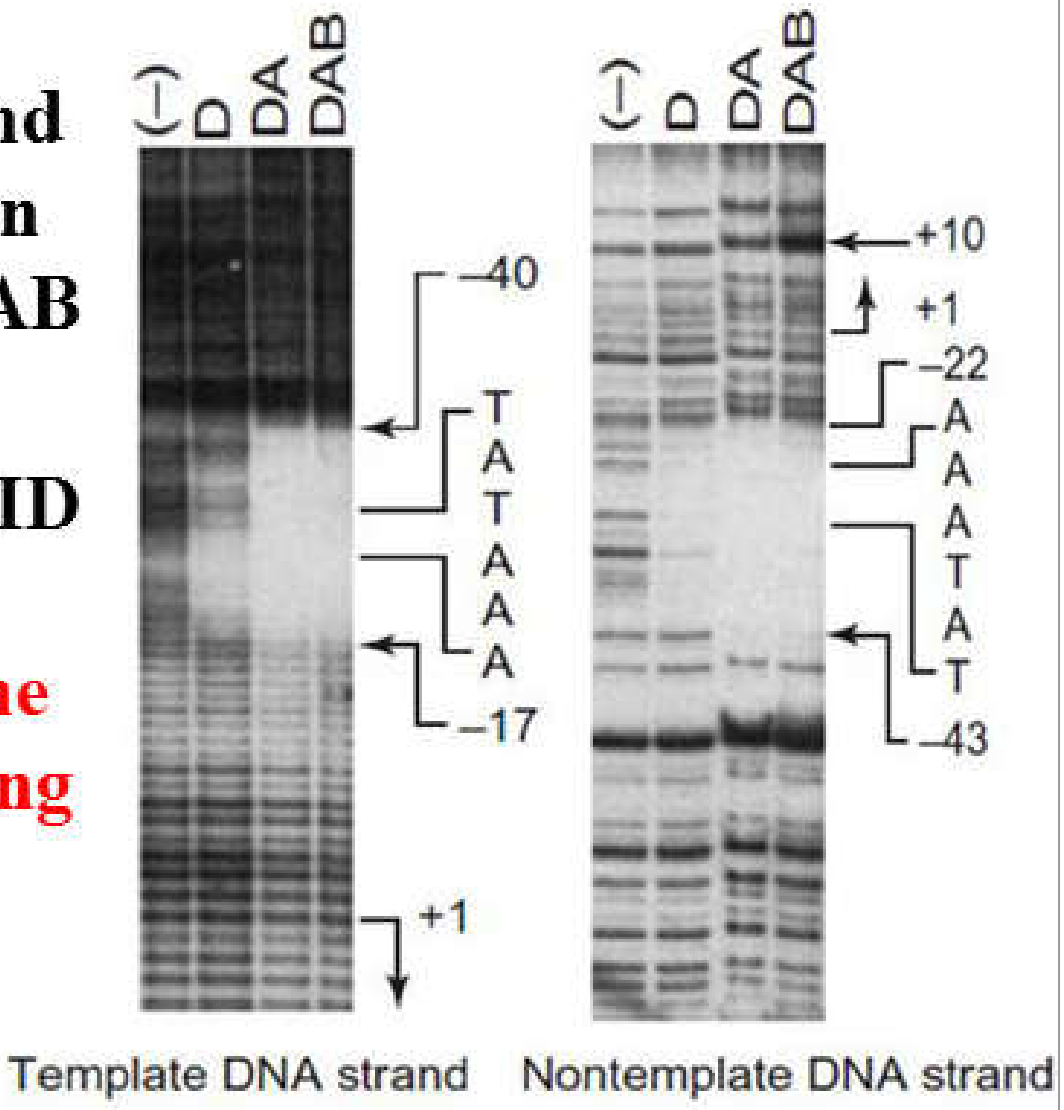


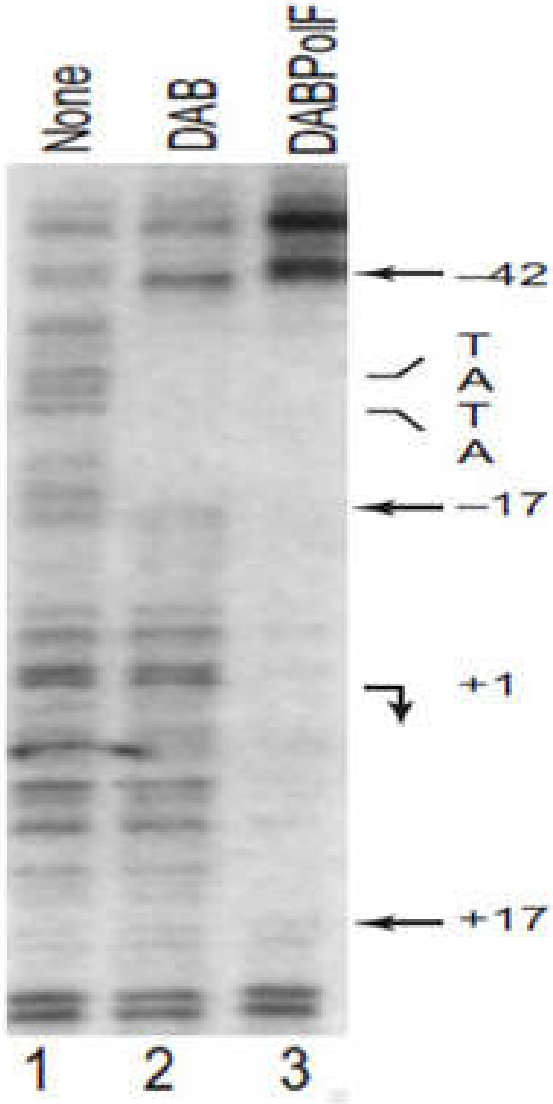
The order of addition of the GTFs and RNA polymerase to the initiation complex in vitro is as follows: TFIID, TFIIA, TFIIB, Pol II, TFIIF, TFIIE, TFIIH.

DABPolFEH

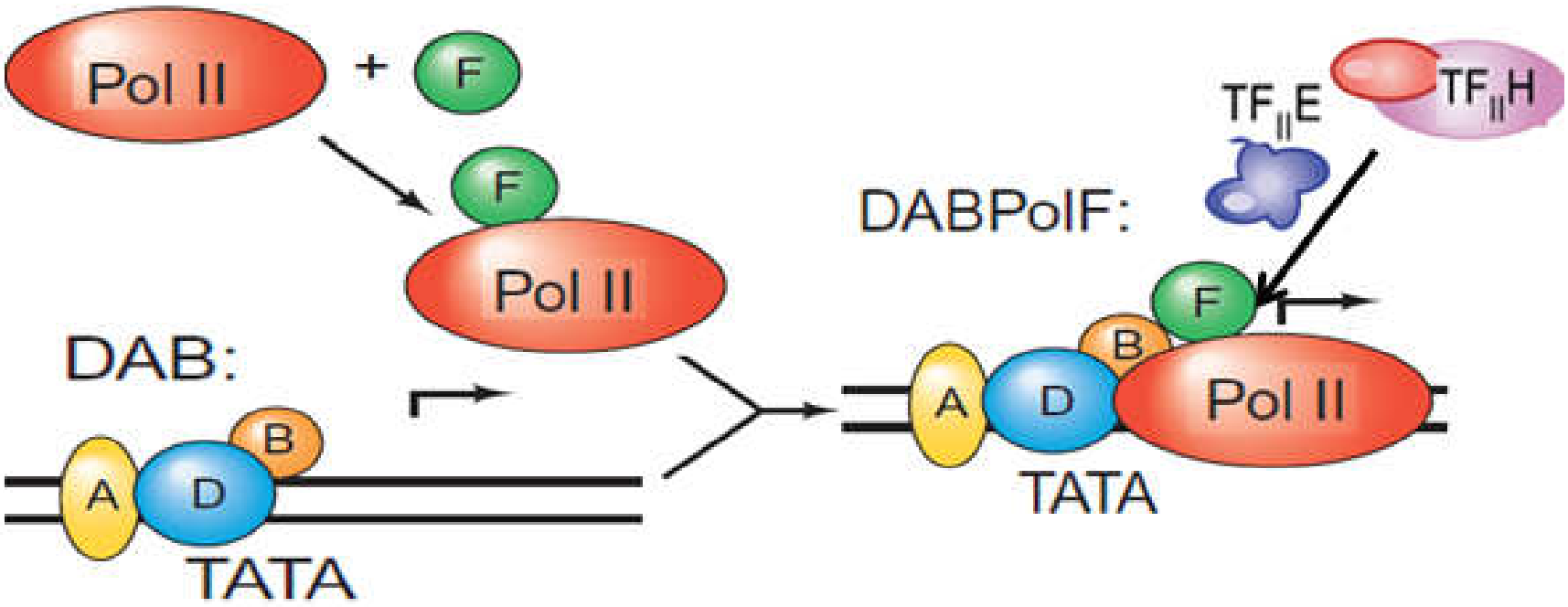
3.3.2 Binding sites of Class II GTFs

- **TFIID, A, and B bind the TATA box region (-42 ~ -17) in the DAB complex.**
- **TFIIA binds to TFIID and may enhance TFIID binding to the TATA box, stabilizing the TFIID-DNA complex.**





- **RNA Pol II and TFIIF extended the protected region.**



(1) TFIID, apparently with the help from TFIIA, binds to the TATA box of promoter.

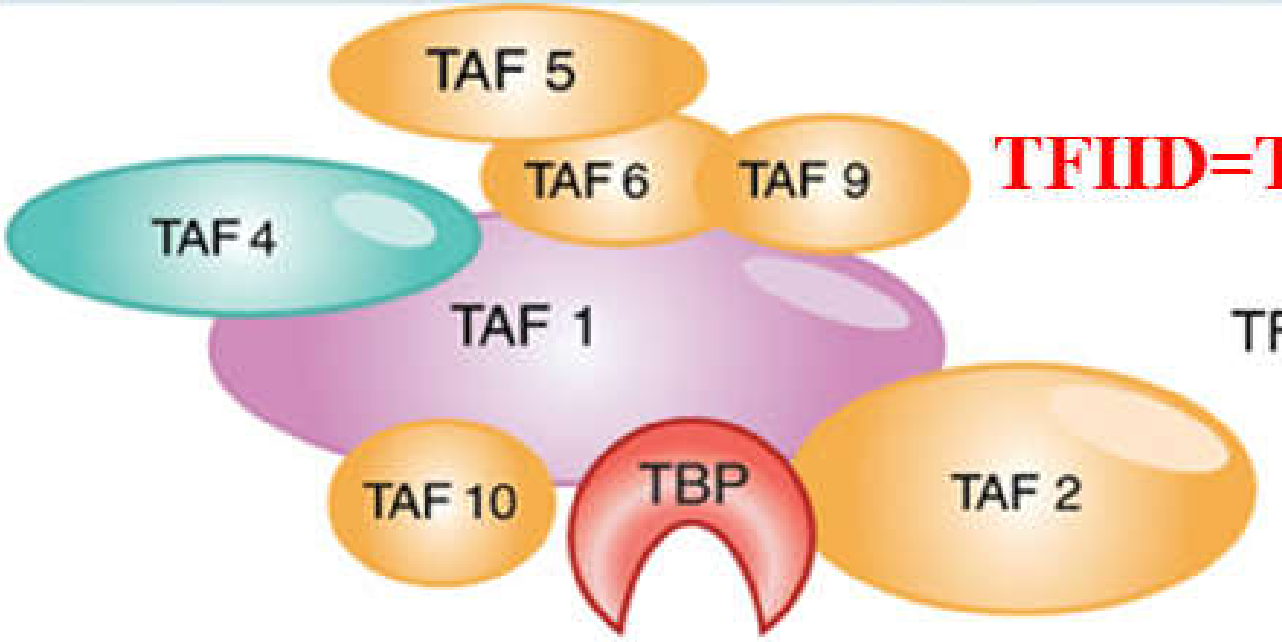
(2) TFIIB binds next.

(3) TFIIF helps RNA Pol II bind to DAB complex.

(4) Then TFIIE and TFIIH bind, forming the DABPolFEH initiation complex.

3.4 Structure and function of TFIID

TFIID is a complex protein containing a TATA-box binding protein (TBP, TATA框结合蛋白) and 13 core TBP-associated factors (TAFs, TBP相关因子 or more specifically, $TAF_{II}s$).



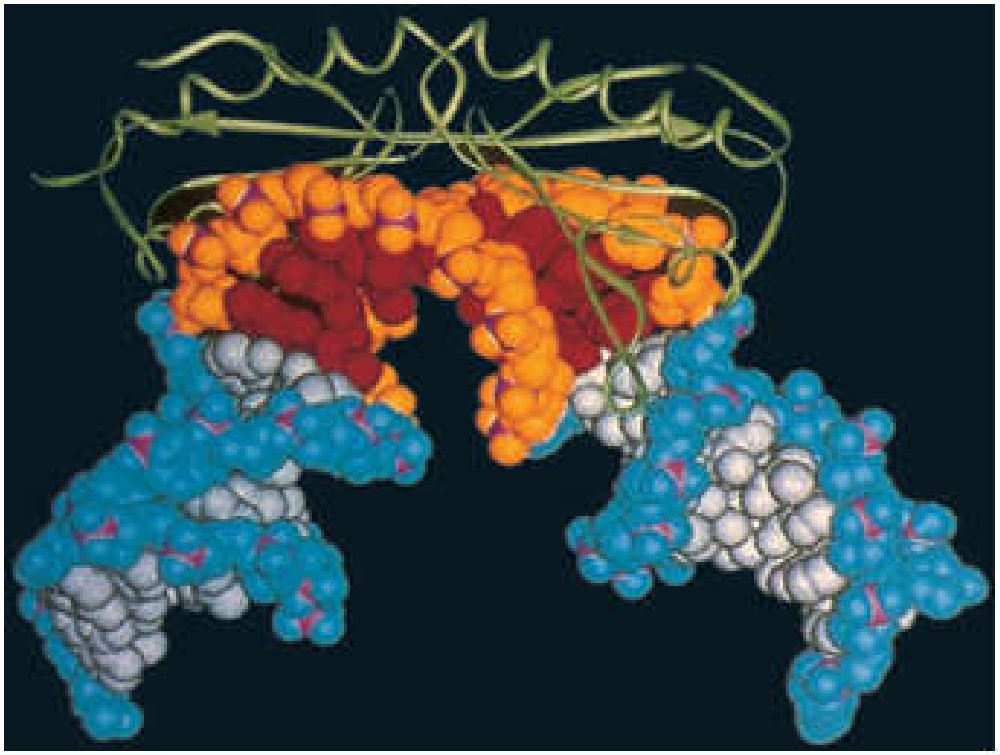
$TFIID=TBP+TAF_{II}s$

$TF_{II}D$

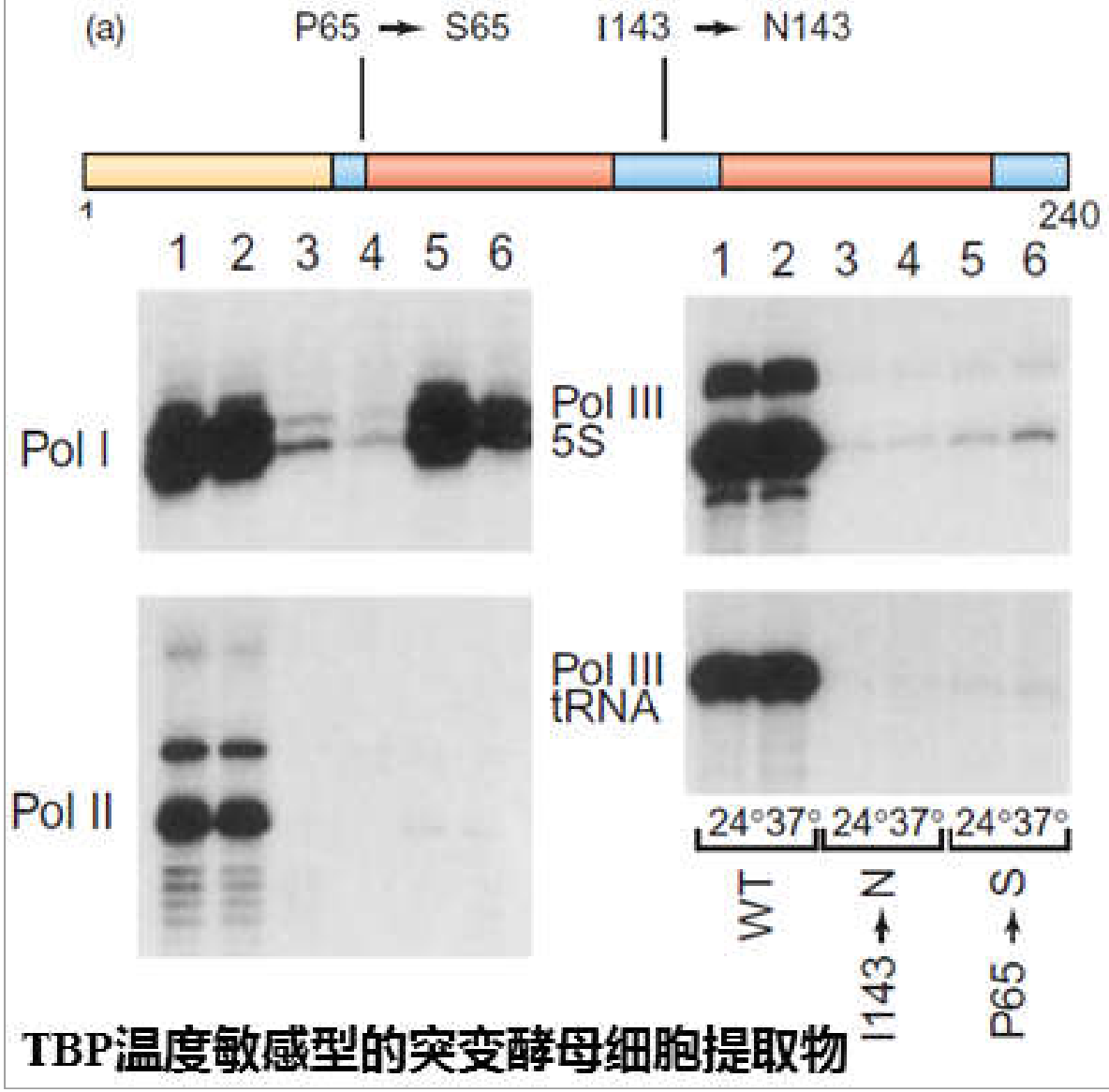
3.4.1 TATA-box binding protein (TBP)

- TBP is the first polypeptide in the TFIID complex to be characterized.
- TBP is a monomeric (单体) protein and highly evolutionarily conserved (TATA-box-binding domains that are more than 80% identical in amino acid sequence).
- TBP is an assembly factor (组装因子). It binds to the TATA box in the minor groove of DNA.

TBP forms a saddle (马鞍) around the DNA and bends it by ~80° .



The versatility (多功能性) of TBP



TBP温度敏感型的突变酵母细胞提取物

TBP is a universal transcription factor required by all three classes of genes.

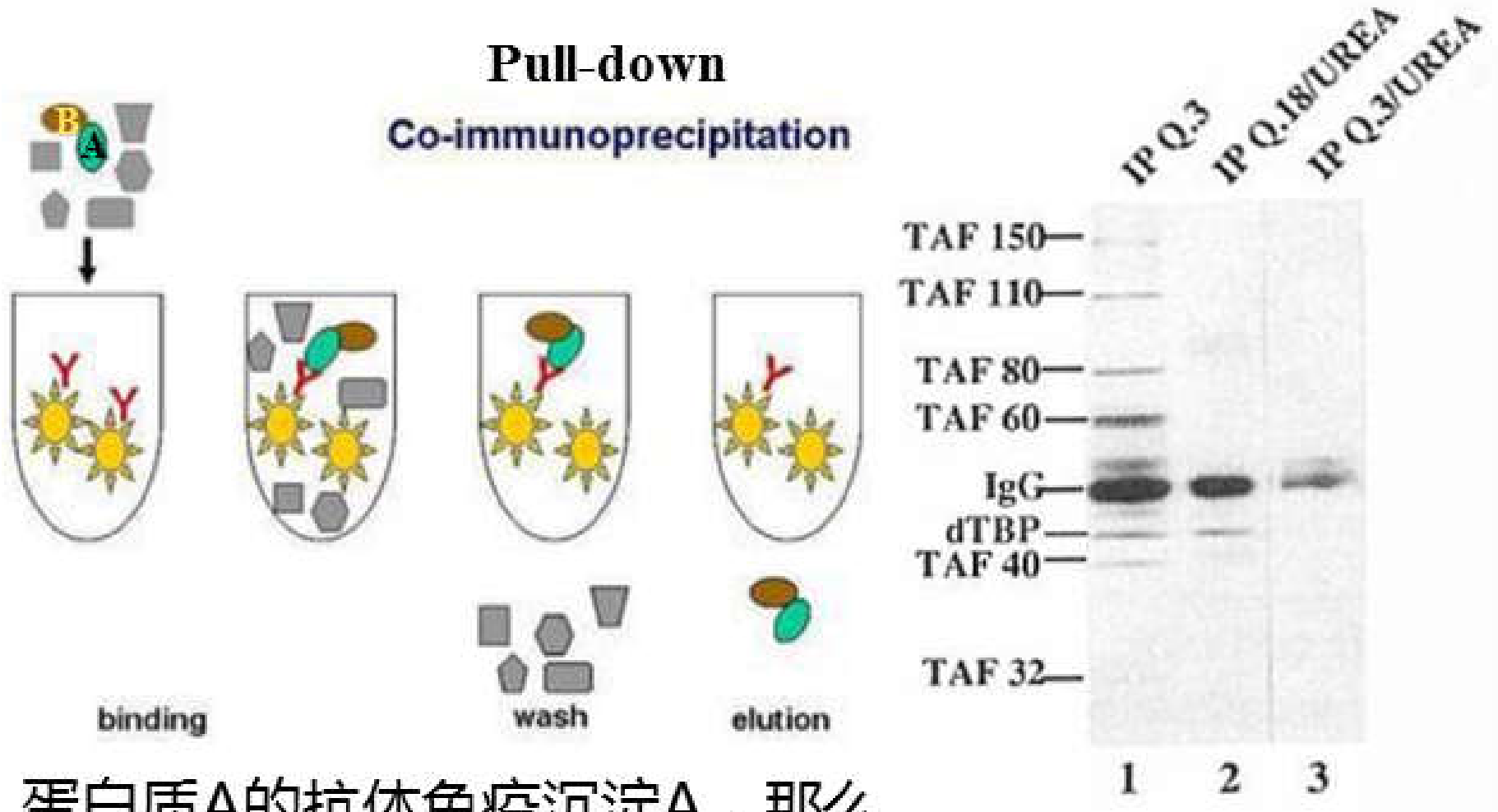
TBP是所有三类基因必需的通用转录因子。

TBP has also been found in archaea (古菌).

3.4.2 TBP-associated factors (TAFs)

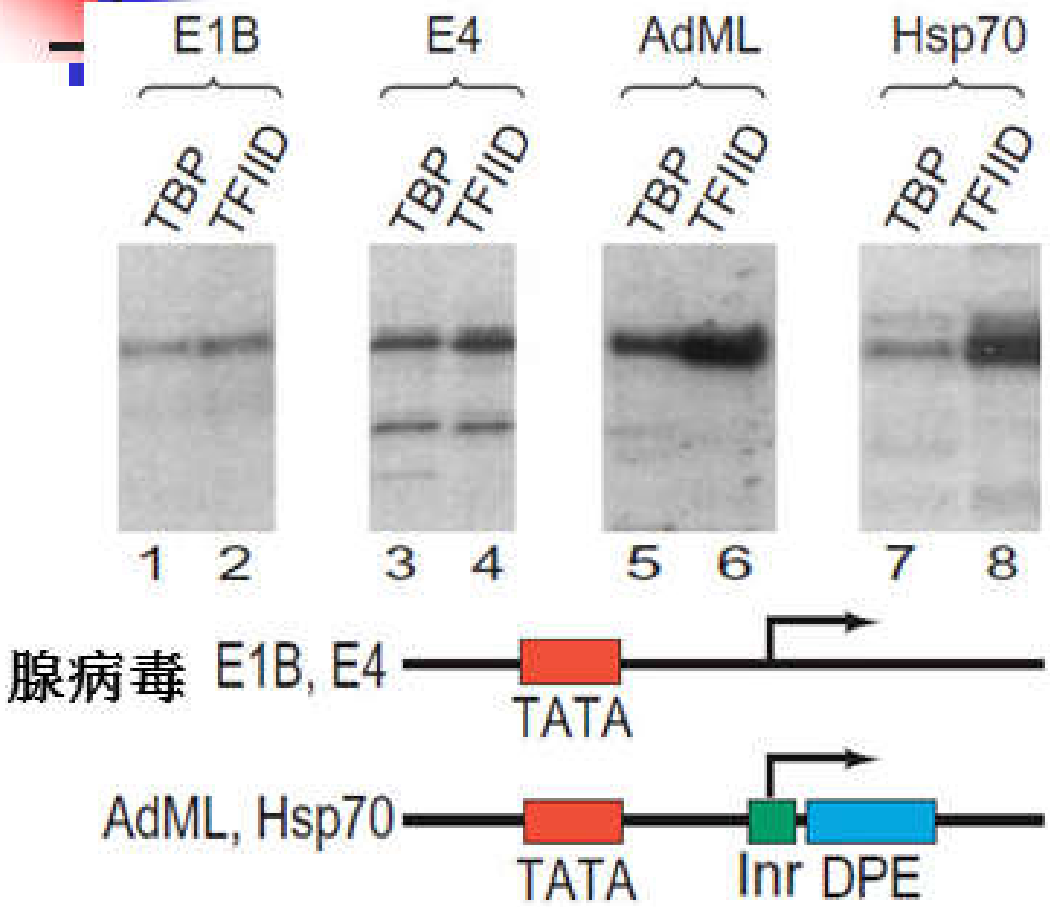
- Thirteen TAFs were identified by using an antibody specific for TBP to immunoprecipitate(免疫沉淀) TFIID.
- The core TAFs have been named according to their sizes, from largest to smallest, as TAF1 ~ TAF13.
- Two obvious functions of the TAFs are interacting with core promoter and interacting with activators.

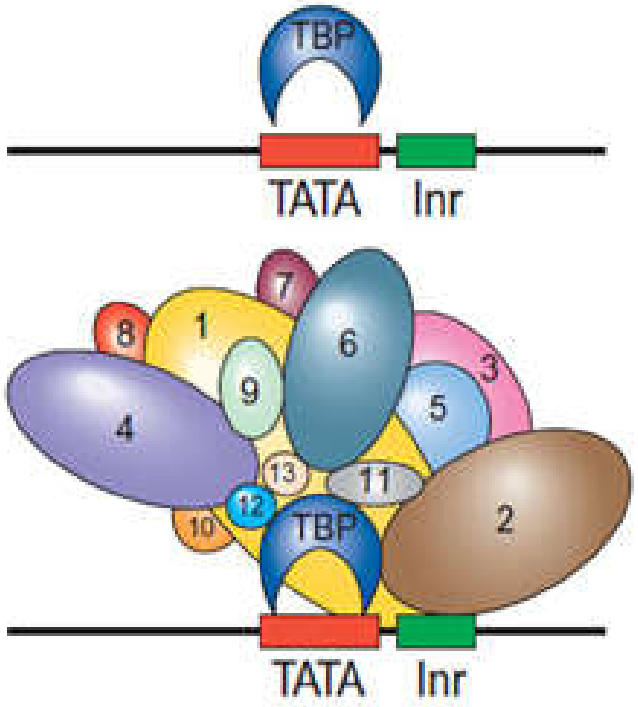
免疫共沉淀技术是研究蛋白质-蛋白质相互作用的方法之一



蛋白质A的抗体免疫沉淀A，那么与A结合的蛋白质B也能沉淀下来

(1) TATA-containing promoter

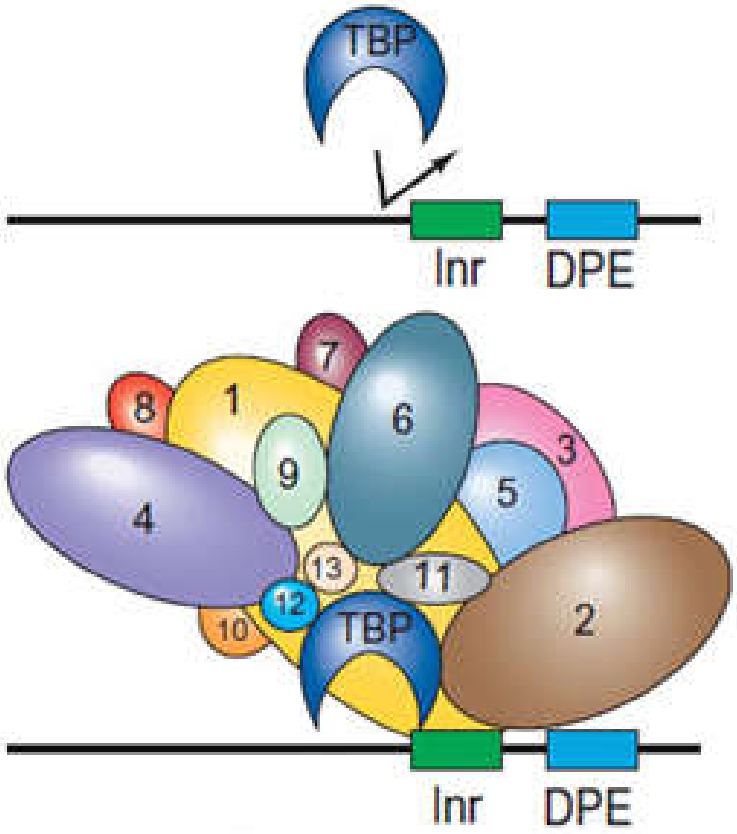




TAFs apparently help TBP facilitate (促进) transcription from promoters (especially with initiators and DPEs).

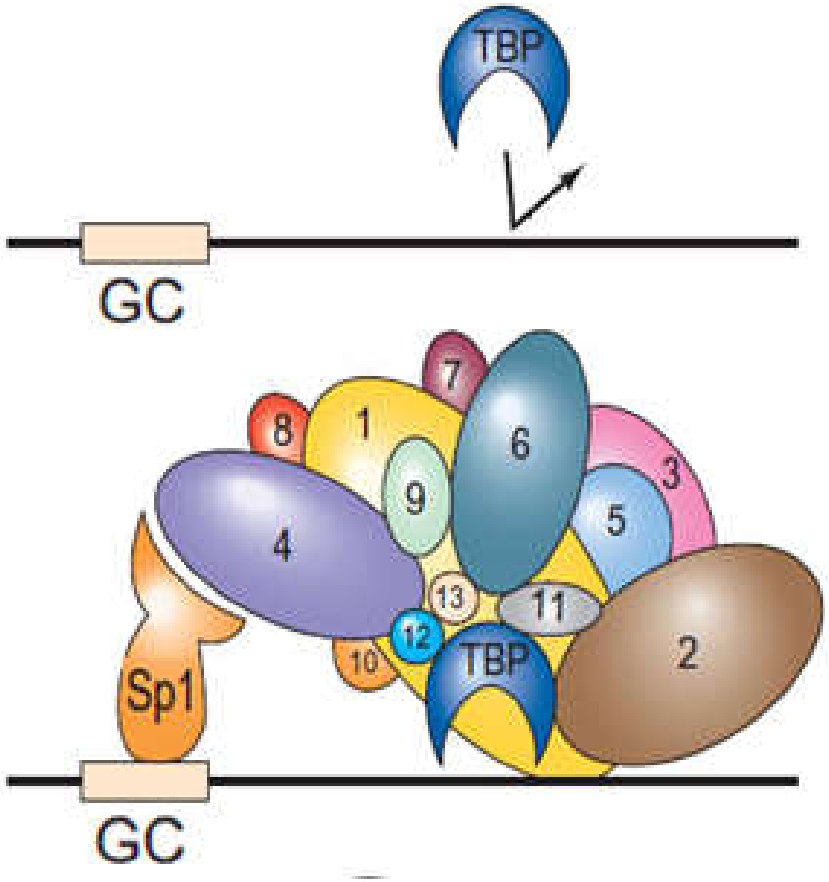
(2) TATA-less promoter with Inr and DPE

- Even though these promoters cannot bind TBP directly, most still depend on TAFs for activity.
- TAFs help TFIID bind to the Inr and DPE of the promoter and therefore can enable TBP to bind to TATA-less promoters.



(3) TATA-less promoter with GC boxes

- **TAFs help TFIID interact with Sp1 that is bound to GC boxes upstream of the transcription start site.**



3.5 Structure and function of TFIIB

- TFIIB binds to TBP at the TATA box via its C-terminal domain, and to RNA Pol II via its N-terminal domain.
- TFIIB acts as a bridge of TFIID and RNA polymerase II.
- This bridging action effects a positioning (定位) of the polymerase active center about 25~30 bp downstream of the TATA box.

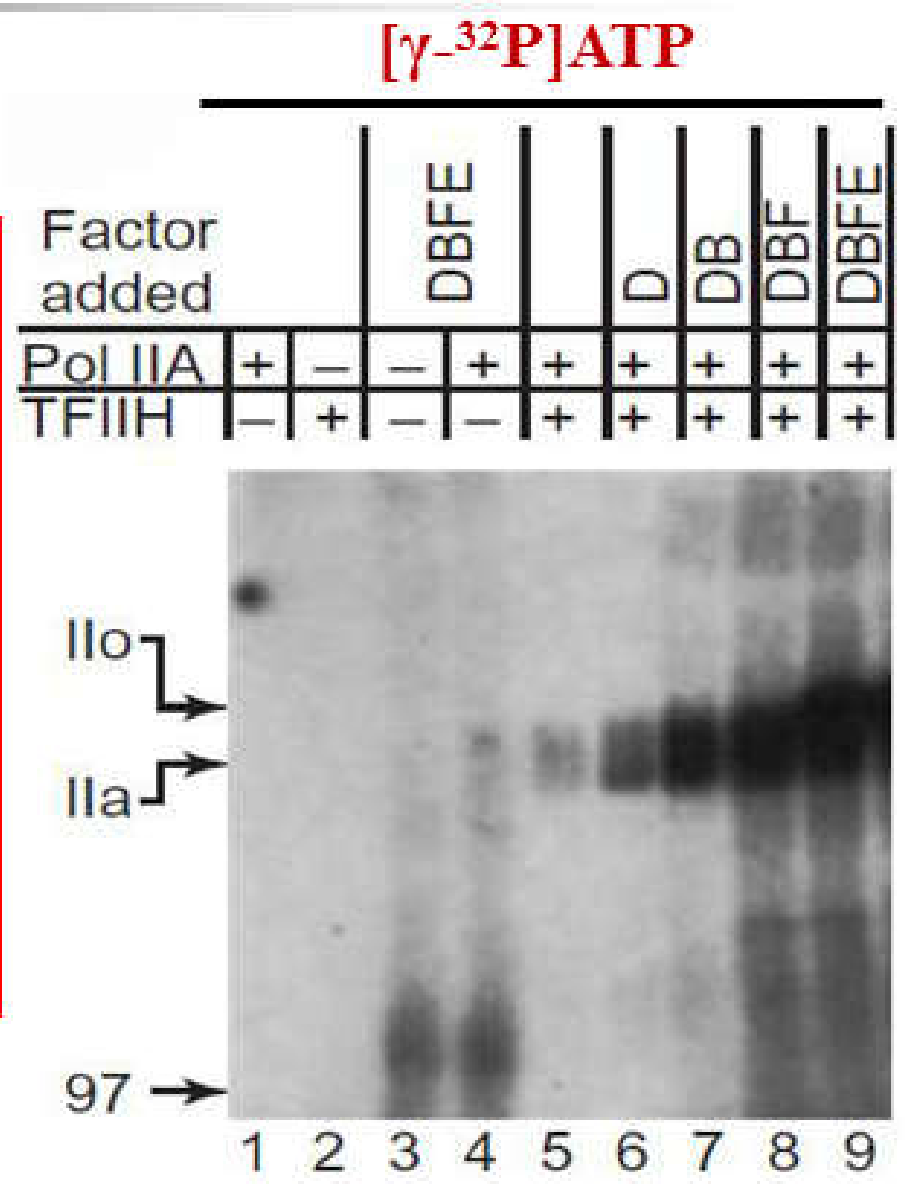
3.6 Structure and function of TFIIH

3.6.1 Structure of TFIIH

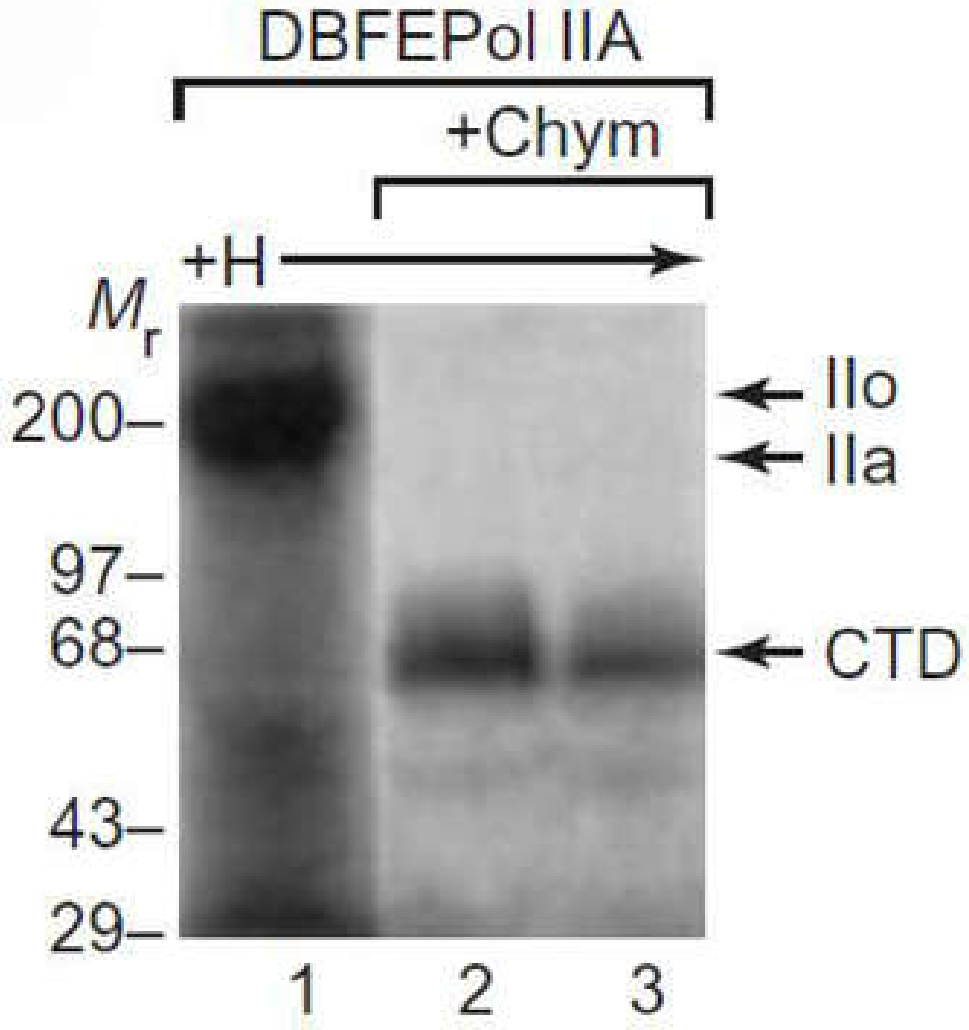
- TFIIH is a complex protein, both structurally and functionally. It contains 9 subunits.
- TFIIH can be separated into 2 complexes: a protein kinase complex composed of 4 subunits, and a 5-subunit core TFIIH complex with two separate DNA helicase and ATPase activities.

(1) Kinase activity of TFIIH

- **TFIIH by itself is capable of carrying out the phosphorylation of RNA Pol II.**
- **Other TFIIs especially TFIIE accelerates phosphorylation of RNA Pol II by TFIIH.**

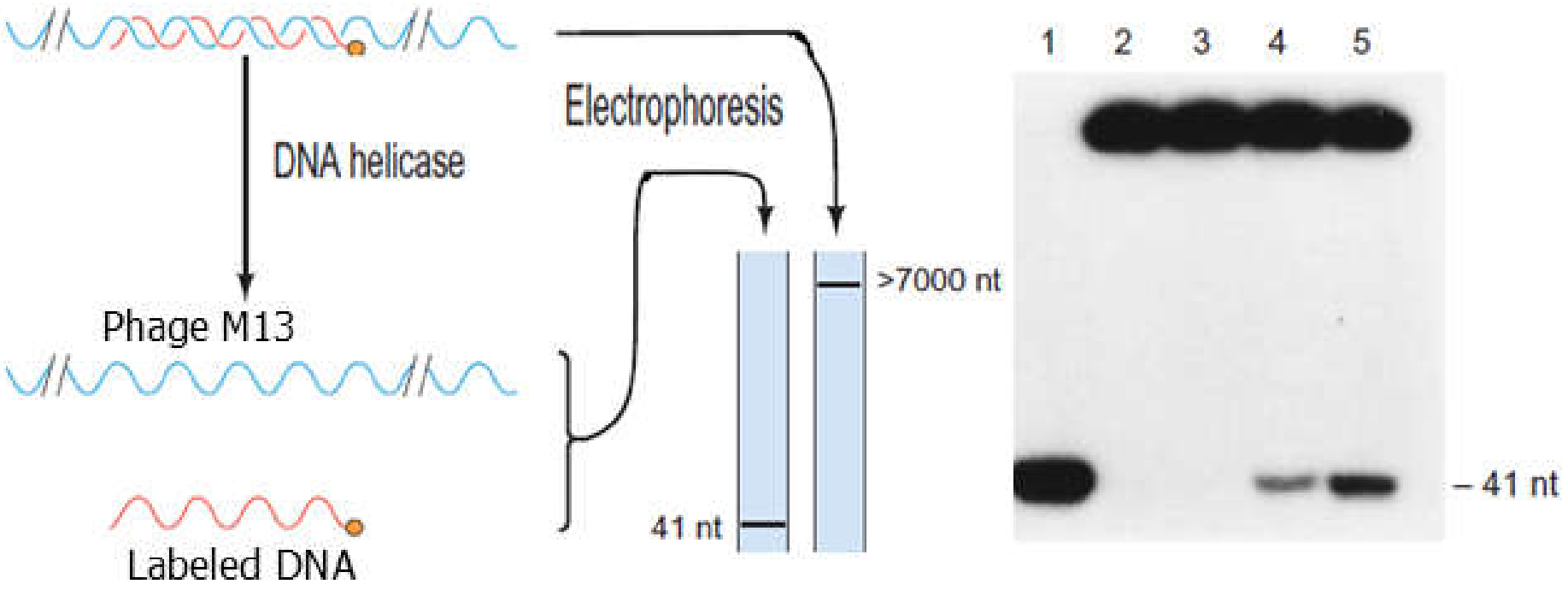


TFIIH phosphorylates the CTD of RNA Pol II



Chym：**胰凝乳蛋白酶。从磷酸化的**RNA Pol II**中切下**CTD

(2) DNA helicase activity of TFIIH



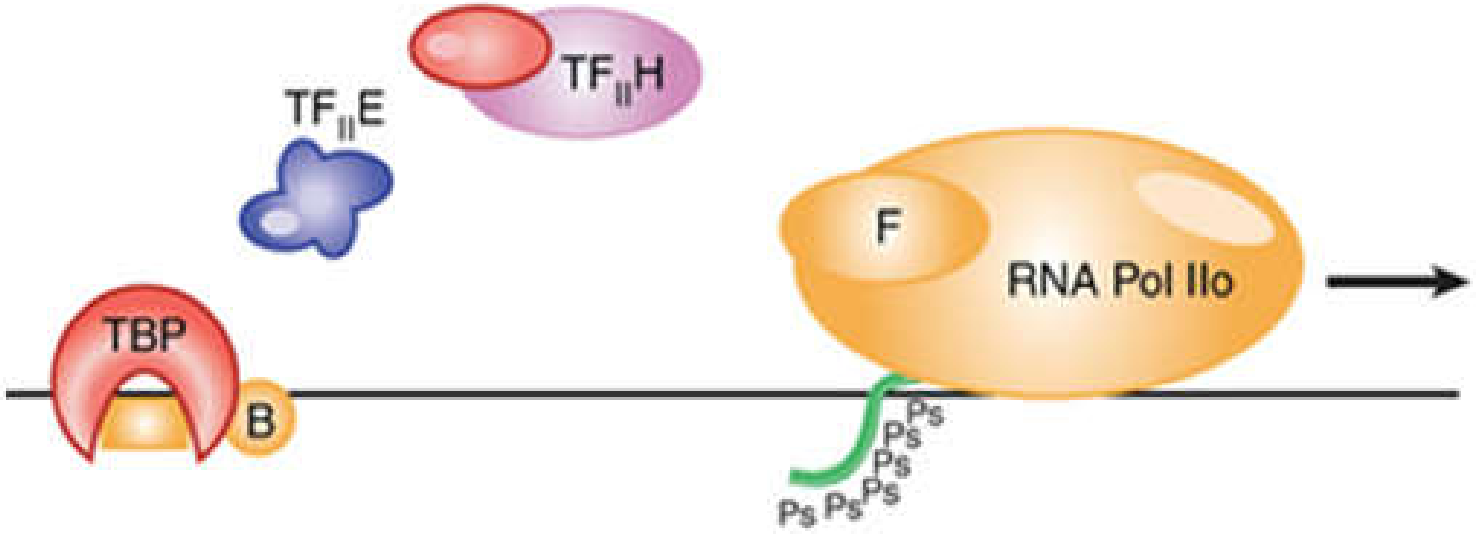
Lane 1, heat-denatured substrate; Lane 2, no protein; Lane 3, 20 ng of RAD25 (TFIIH DNA helicase) with no ATP; Lane 4, 10 ng of RAD25 plus ATP; Lane 5, 20 ng of RAD25 plus ATP

3.6.2 Function of TFIIH

TFIIH play two major roles in transcription initiation:

- Phosphorylate the CTD of RNA Pol II
  - Phosphorylation of the polymerase allows it to shift from initiation to elongation
- Unwind DNA at the transcription start site to create the "transcription bubble".
  - TFIIH DNA helicase activity causes full melting of the DNA at the promoter and thereby facilitates promoter clearance.

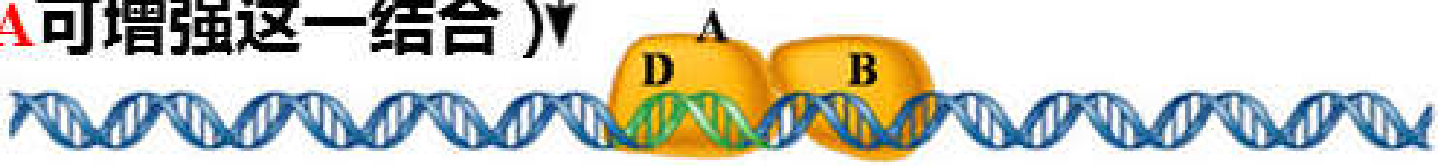
- **TFIIE and TFIIH made no difference in the elongation reaction.**



小结：GTFs参与的II类基因转录起始

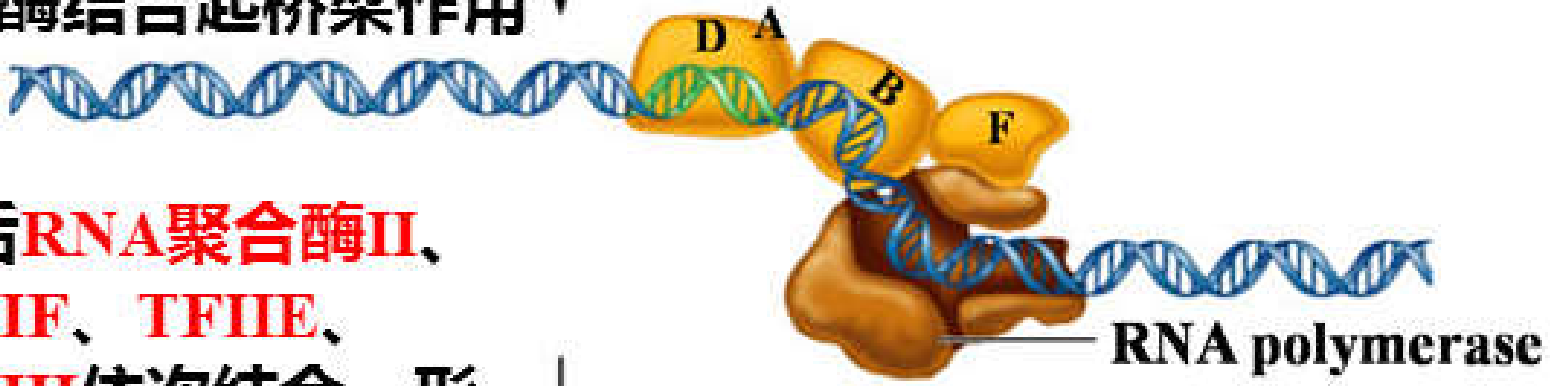
TFIID与启动子结合
（TFIIA可增强这一结合）

TFIIB binds to TFIID.



TFIIB再与TFIID结合，并为RNA聚合酶结合起桥梁作用

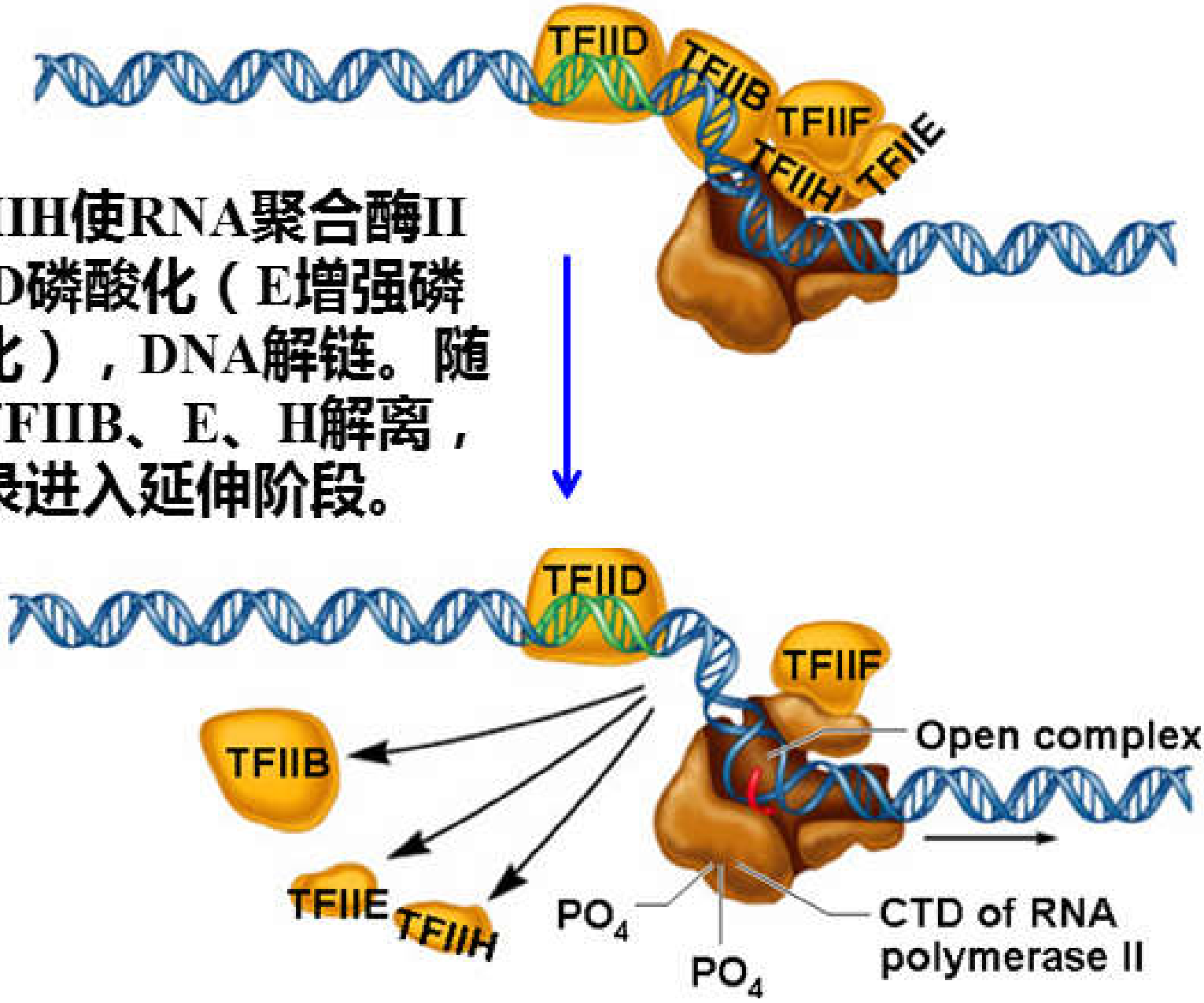
TFIIB acts as a bridge to bind to RNA Pol II/TFIIF.



随后RNA聚合酶II、TFIIF、TFIIE、TFIIH依次结合，形成转录起始复合物

TFIIE and TFIIH bind to RNA Pol II.

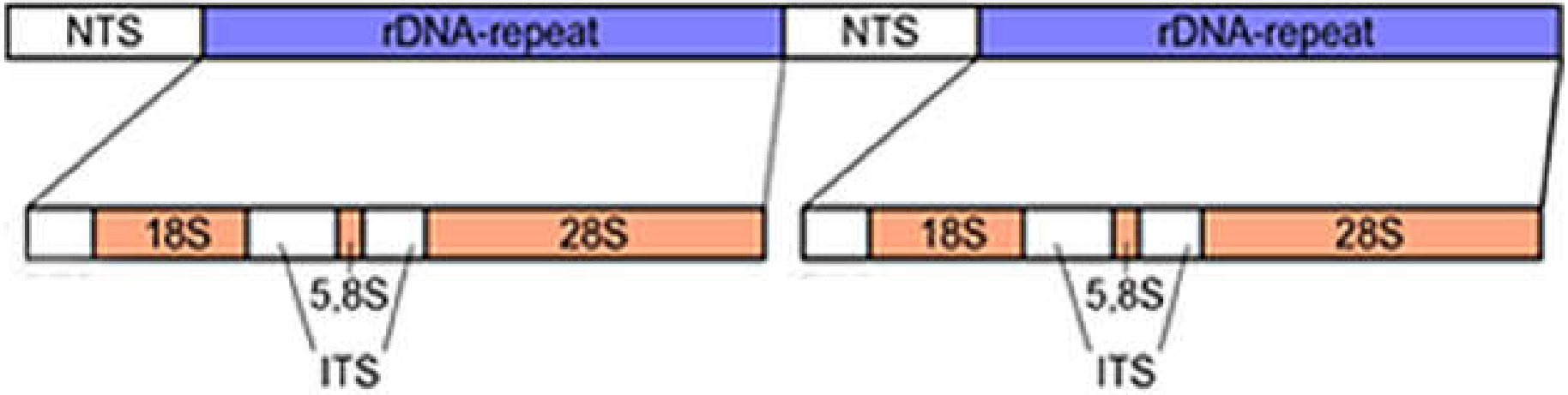
TFIIH使RNA聚合酶II CTD磷酸化（E增强磷酸化），DNA解链。随后TFIIB、E、H解离，转录进入延伸阶段。



4. Class I promoter and GTFs

4.1 rRNA gene

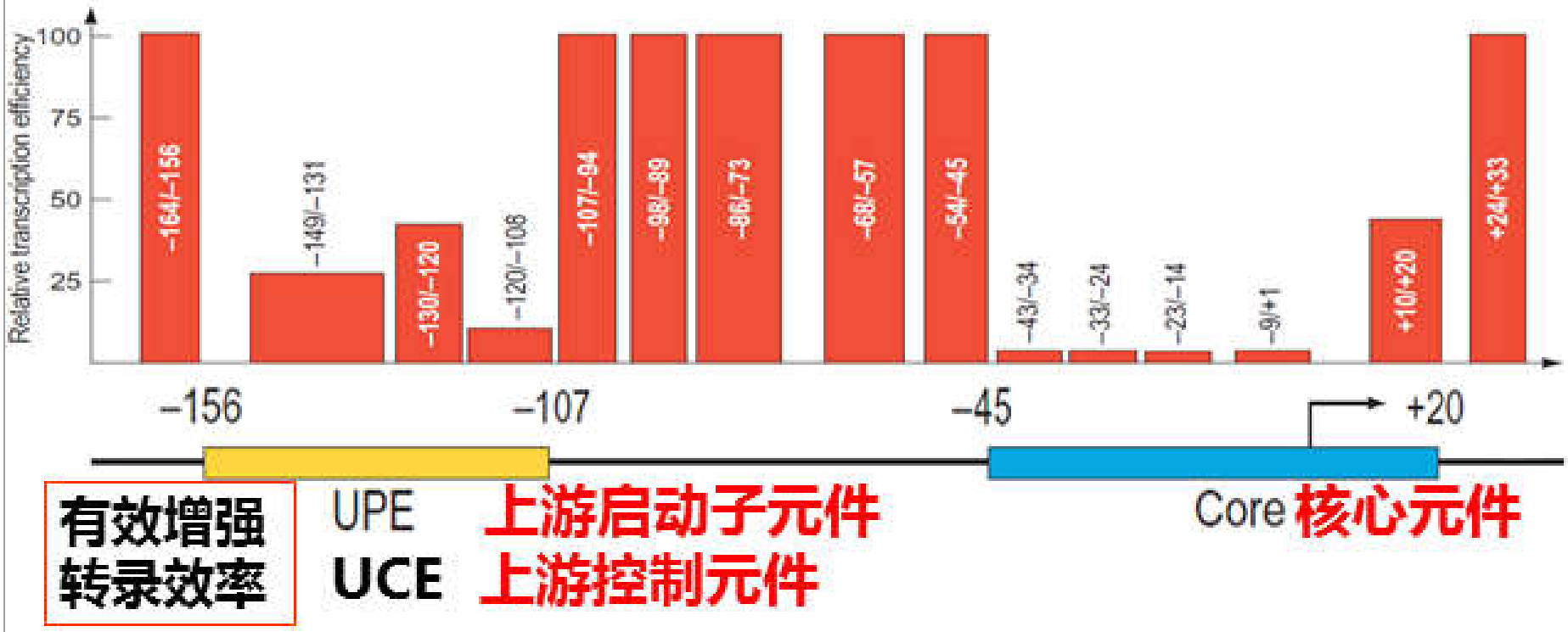
- rRNA precursor gene is present in hundreds of copies in each cell, and they all have the same promoter sequence.



- 成簇存在的rRNA基因组成了核仁组织区（nucleolar organizer region）

4.2 Class I promoter

- The promoter sequence is quite variable from one species to another.



- The spacing between the core element and UPE is important.

4.3 Class I GTFs and transcription initiation

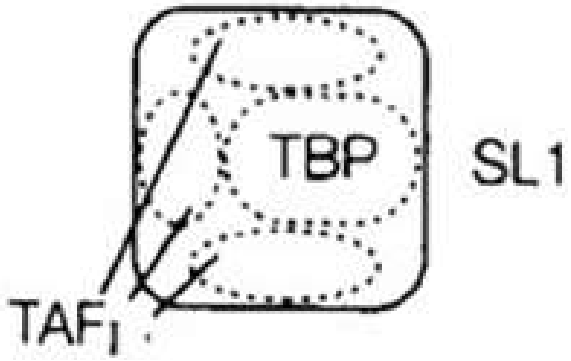
Core-binding factor (核心结合因子)
UPE-binding factor (UPE结合因子) } **GTF**

- **The core-binding factor is called selectivity factor 1 (SL1) in humans, and TIF-ⅠB in some other organisms.**
- **UPE-binding factor called upstream-binding factor (UBF) in mammals and upstream activating factor (UAF) in yeast.**

4.3.1 SL1（选择因子1）

(1) Structure of SL1

- **SL1 = TBP + TAF_I (human)**



Similar to TFIID?

TBP √

TAF ×

- **SL1 shows species specificity. TAF_I决定**

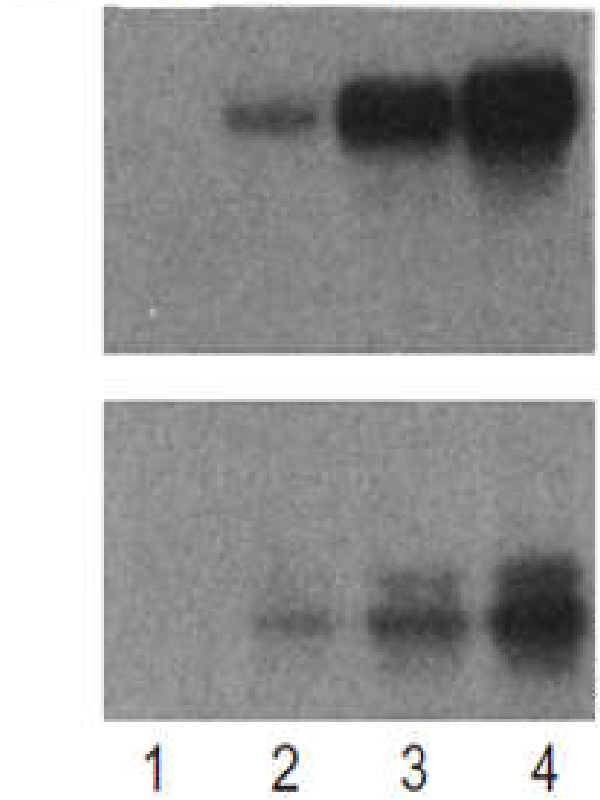
(2) Function of SL1

- SL1 is the GTF required to recruit (募集) RNA Pol I to the promoter.
- SL1 by itself cannot stimulate human RNA Pol I to start transcription efficiently. It requires the assistance of UBF.

4.3.2 UBF（上游结合因子）

(1) UBF binds to the UPE and bends the DNA dramatically.

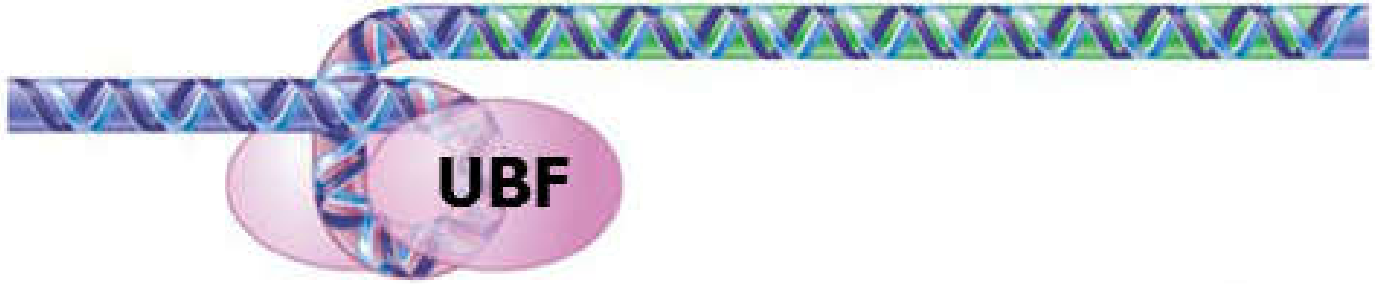
(2) UBF is an assembly factor (组装因子) that helps SL1 bind to the core promoter element.



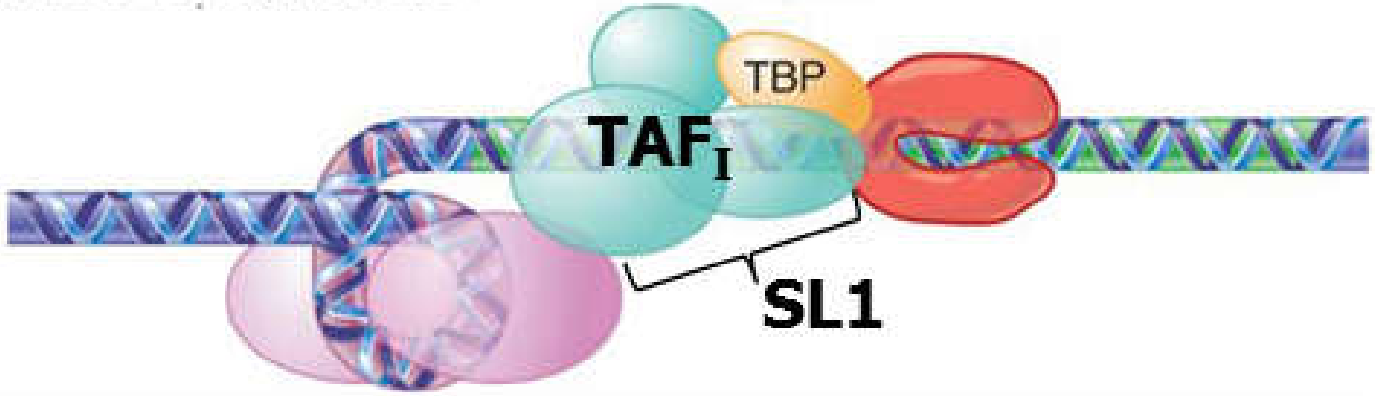
- 没有SL1转录不发生
- 没有UBF时，转录水平较低
- 没有UPE的启动子也能转录。
- UBF通过与UPE结合，增强转录效率。
- UBF可能还与核心元件或核心元件上游的DNA序列结合发挥作用。

4.3.3 Transcription initiation of Class I gene

UBF binds to upstream promoter element



RNA polymerase I holoenzyme includes core binding factor (SL1) that binds to core promoter

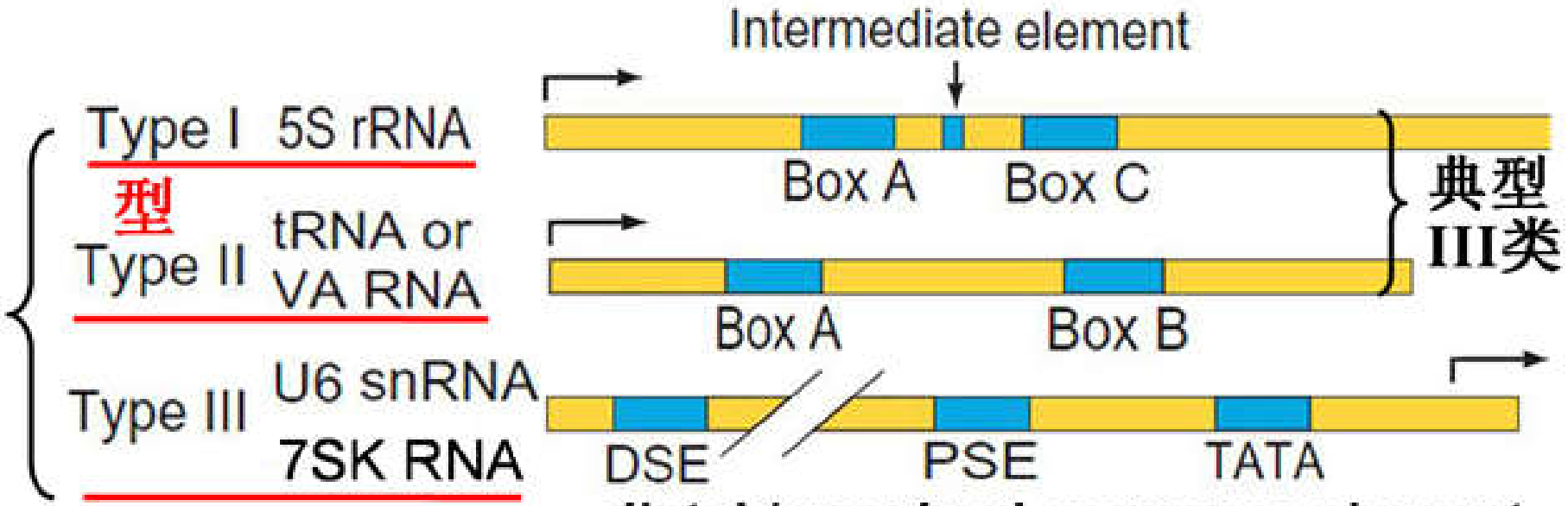


5. Class III promoters and GTFs

5.1 Class III promoters

The classical class III genes (Types I and II) have promoters that lie wholly within the genes.

—internal (内部) promoters

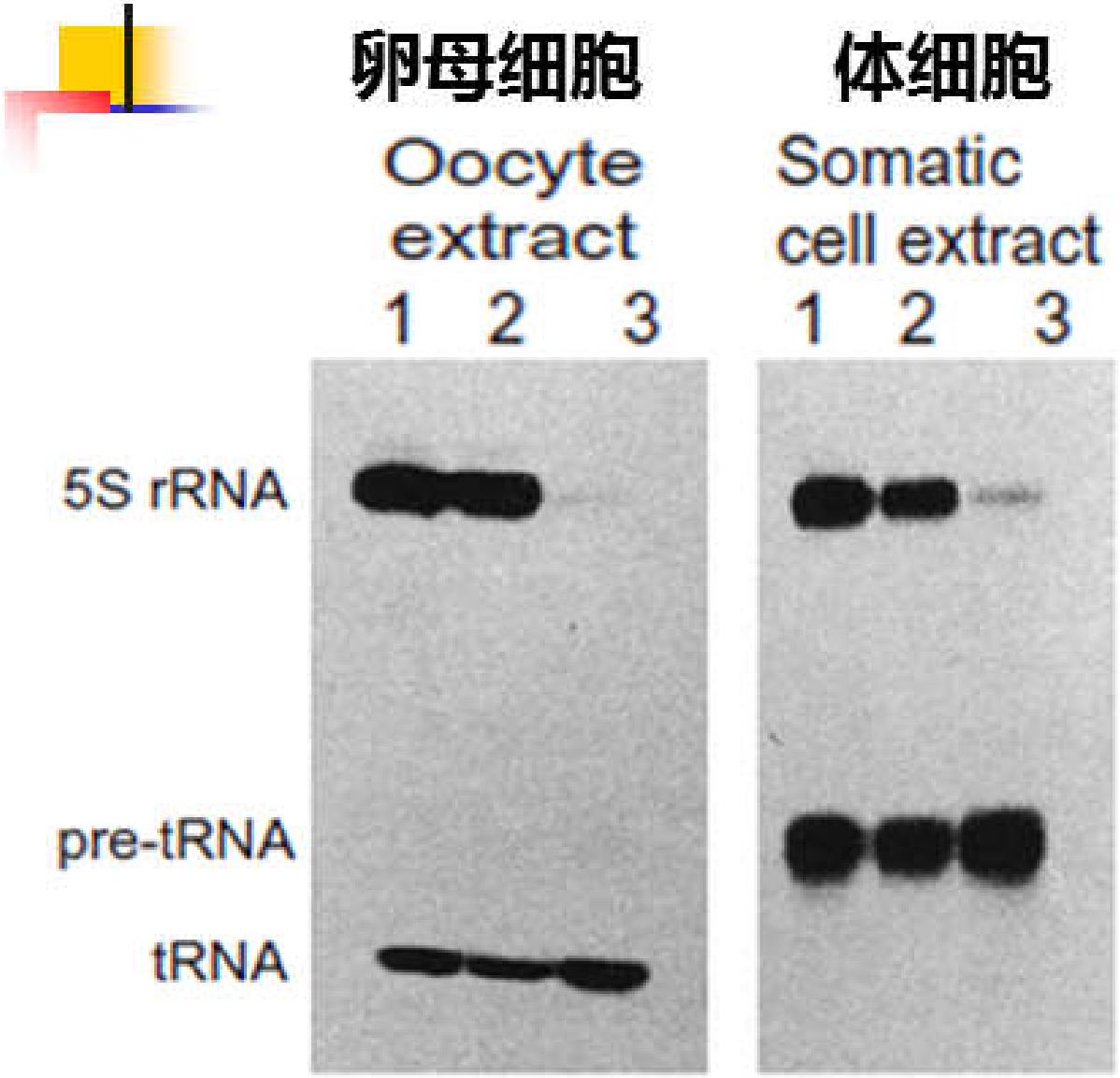


distal / proximal sequence element

远/近侧序列元件 PSE必需，DSE增强

5.2 TFIIIA

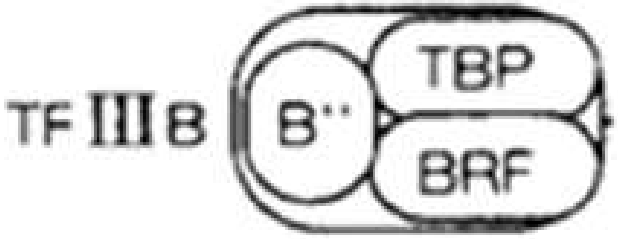
- TFIIIA is the first transcription factor discovered in eukaryotic cells.
- TFIIIA is required for transcription of the 5S rRNA genes, but not the tRNA genes.
  TFIIIA 为5S rRNA基因转录所必需，但对 tRNA则不然。
  - 体外转录系统中加入TFIIIA后5S rRNA才能合成。（抗TFIIIA的抗体可以有效阻止5S rRNA的生成，但对tRNA的合成没有影响。）



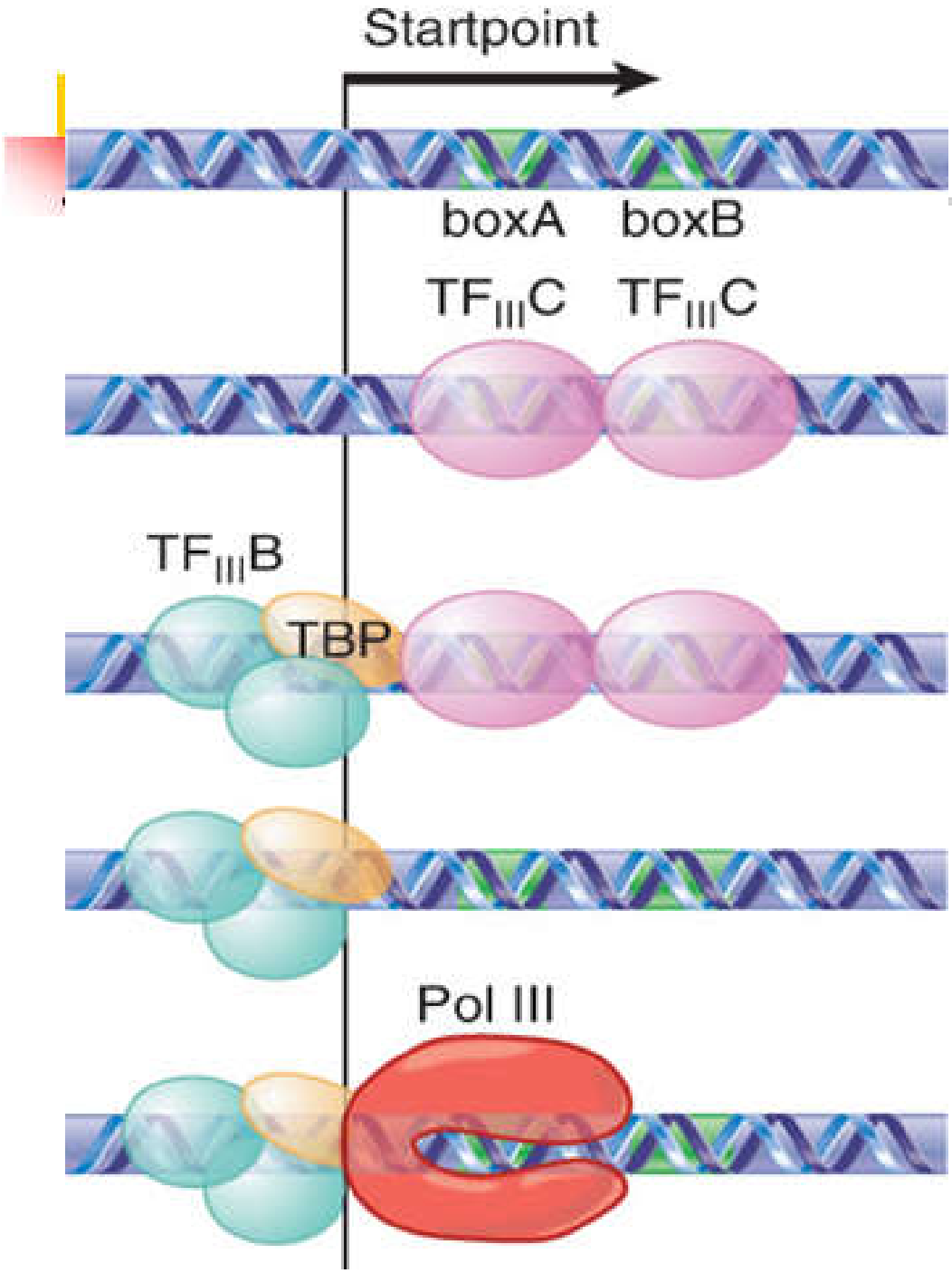
卵母细胞提取物可将tRNA前体加工为成熟tRNA

5.3 TFIIIB and TFIIIC

- **TFIIIB and C are both required for transcription of the classical Pol III genes, and these two factors depend on each other for their activities.**
  所有典型III类基因的转录都需要TFIIIB和TFIIIC.

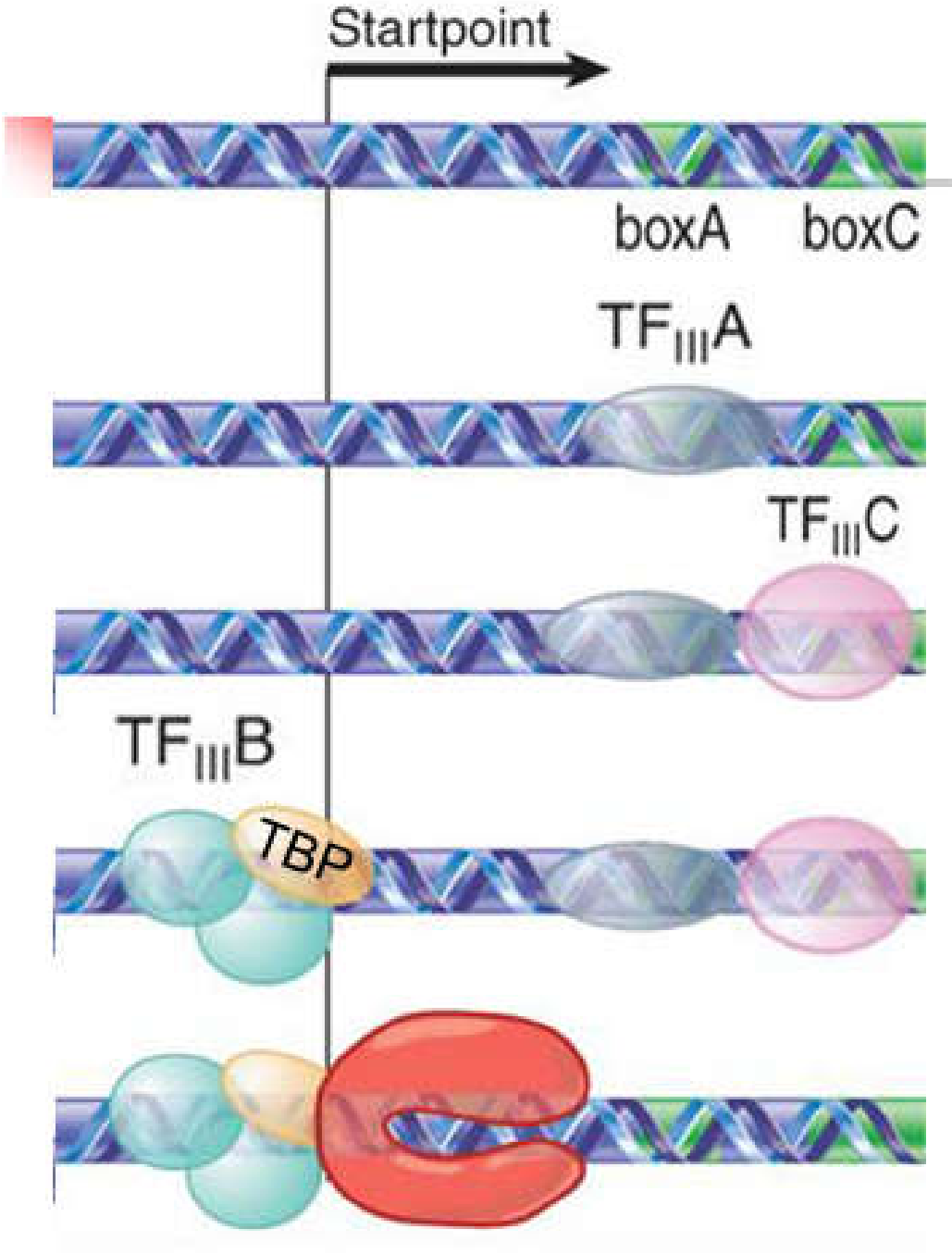


- **TFIIIB = TBP + TAF_{III}**



5.4 transcription initiation of tRNA gene

TFIIIC binds to both Box A and Box B. This enables TFIIIB to bind to the upstream region. Then, RNA Pol III can bind at start point.



5.5 Transcription initiation of 5s rRNA gene

TFIIIA must bind at Box A to enable TFIIIC to bind at Box C. Once TFIIIC has bound, TFIIIB binds to the upstream region and RNA Pol III joins the complex.

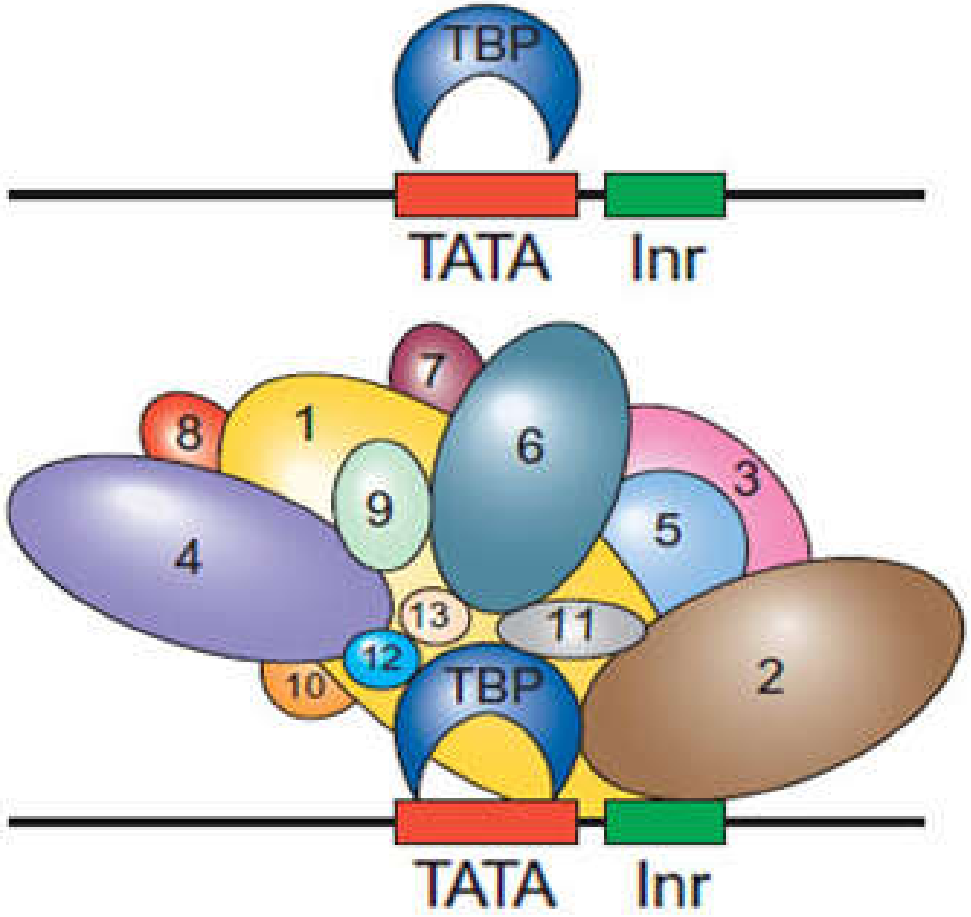
6. Unifying principles (普遍规律) of transcription initiation in eukaryotes

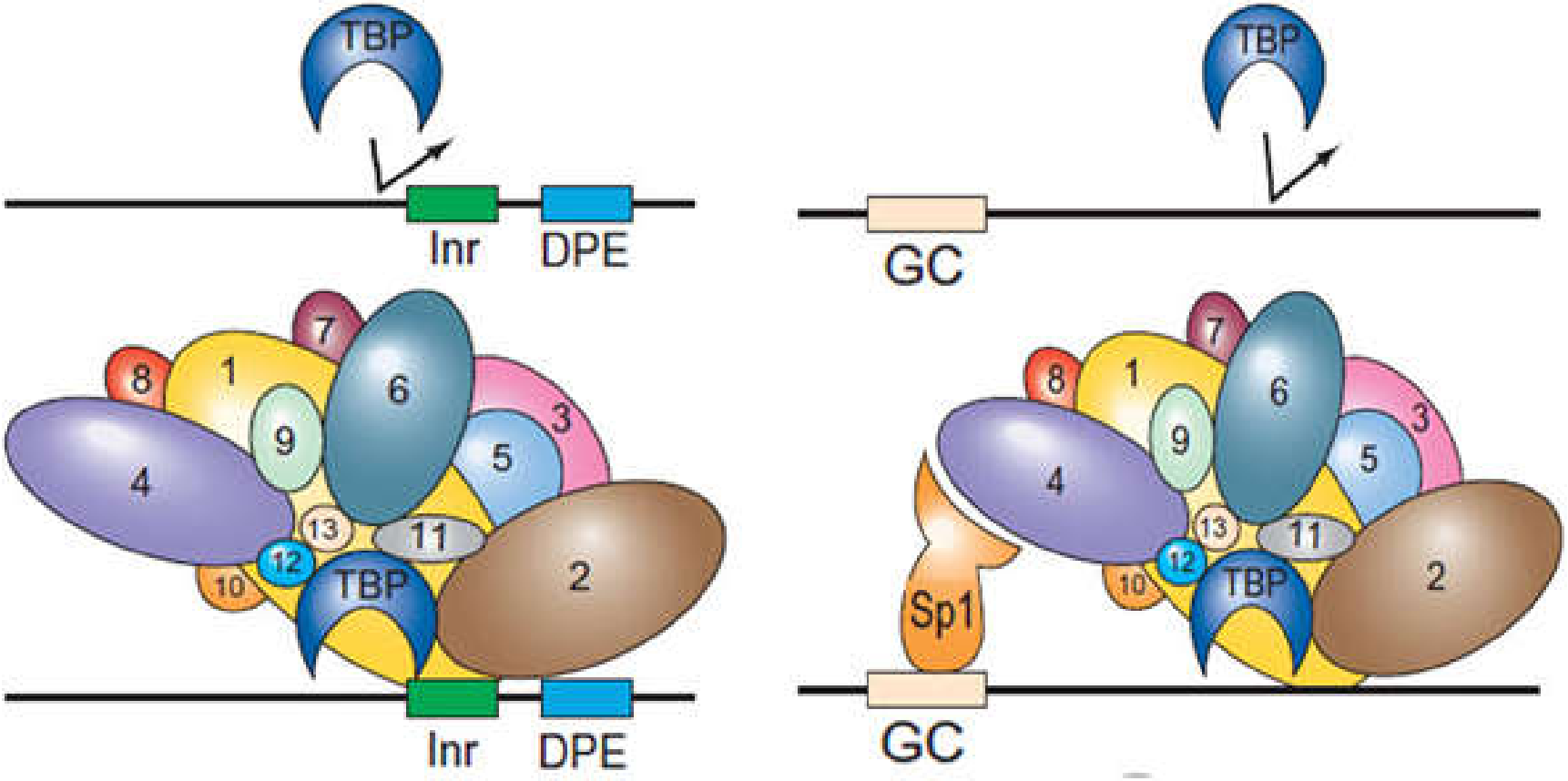
（1）转录起始复合物的形成始于组装因子，它能识别并结合启动子内特定的结合位点，然后由该蛋白募集其他组分。

（2）在大多数真核启动子中，TBP在形成起始复合物中起组织者的作用(在Pol之前结合到启动子)。

（3）TBP的专一性由与之相伴的TAF决定，TBP在与不同类型的启动子结合时，伴有不同类型的TAF.

- 含有TATA框的II类和III类启动子：组装因子通常为TBP。
- 无TATA框的II类启动子：(1) TFIID中的TAF可以结合到起始位点（有Inr）；(2) Sp1先结合到GC框。
- I类启动子：组装因子是UBF，它首先结合到UPE上，然后吸引含有TBP的SL1结合到核心元件。
- 典型的III类启动子：TFIIIC或TFIIIA（5S rRNA基因）作为组装因子结合到内部启动子上，并吸引含TBP的TFIIIB结合到起始位点上游。

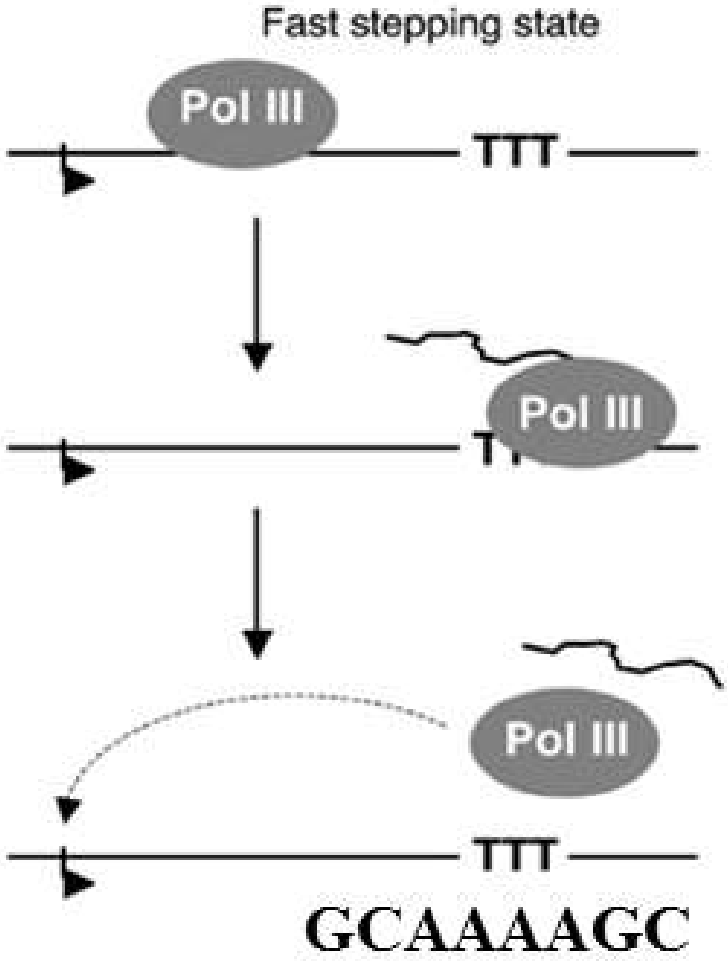

TBP
TATA
Inr
7
8
1
6
3
9
5
4
13
11
12
10
TBP
2
TATA
Inr


TBP
Inr
DPE
TBP
GC
1
2
3
4
5
6
7
8
9
10
11
12
13
TBP
Sp1
Inr
DPE
GC

7. Termination of transcription

7.1 RNA Pol III termination

- Termination of transcription by RNA Pol III appears only to require polymerase recognition of a simple nucleotide sequence.
- This consists of clusters of dA residues whose termination efficiency is affected by surrounding sequence.



7.2 RNA Pol II termination

(1) 依赖于poly(A)

绝大部分mRNA的转录终止与其3’端加尾(特定位点剪切和聚腺苷酸化)相偶联。

↑ Ser2-P ↓ Ser5-P

(2) 依赖Sen1复合物的转录终止

Sen1与Rho具有相似的结构与功能，所以Sen1复合物可通过解开RNA-DNA杂交双链终止转录。 ↓ Ser2-P ↑ Ser5-P

两种转录终止模式的选择受RNA Pol Ⅱ的CTD磷酸化的调控。

8. Differences between prokaryotic and eukaryotic transcription

	原核生物	真核生物
RNA聚合酶	1种	3种
启动子结构	小	大，多种
mRNA	多顺反子mRNA 半衰期短	单顺反子mRNA
转录因子	很少	多
终止方式	不依赖或依赖Rho	……
转录与翻译	同时	不同时 在不同部位
RNA加工修饰	连续基因 基本不需	断裂基因 需要

9. Inhibitors of nucleic acid biosynthesis

9.1 嘌呤和嘧啶类似物抑制剂

抑制或干扰核酸的合成

- 抑制核苷酸合成酶
  - 如:氮尿嘧啶二核苷酸能够抑制核苷酸磷酸化酶
- 掺入核酸分子中---功能改变
  - 如:5-氟尿嘧啶能够掺入RNA中

9.2 与DNA模板作用的抑制剂

- 烷基化剂
  - 单功能基团---DNA烷基化干扰DNA复制
  - 双功能基团---DNA交联抑制模板
- 放线菌素：放线菌素D插入dsDNA的dG-dC之间，抑制DNA复制及RNA转录
- 嵌入剂：如:吖啶, EB等

9.3 与RNA聚合酶作用的抑制剂

利福霉素	抑制原核转录起始
利迪链菌素	抑制原核转录延伸
α - 鹅膏蕈碱	抑制真核

Summary

1. Structures and functions of GTFs in three transcriptional initiation complex (especially Class II GTFs)

2. Differences between prokaryotic and eukaryotic transcription

还有哪些方法可以用来研究DNA-蛋白质的相互作用以及蛋白质-蛋白质的相互作用？原理是什么？