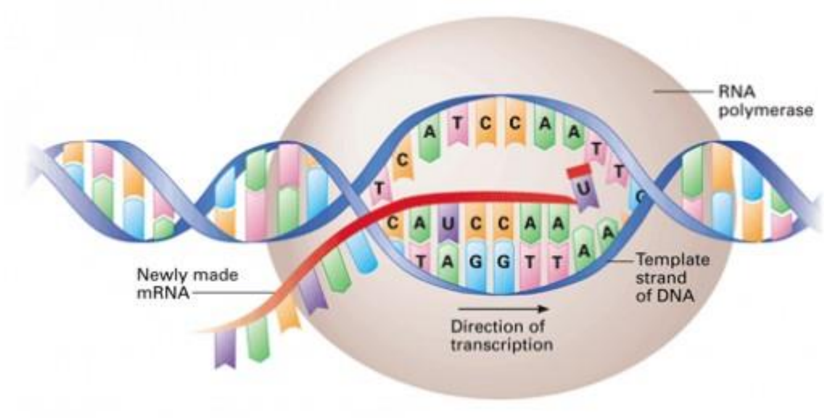
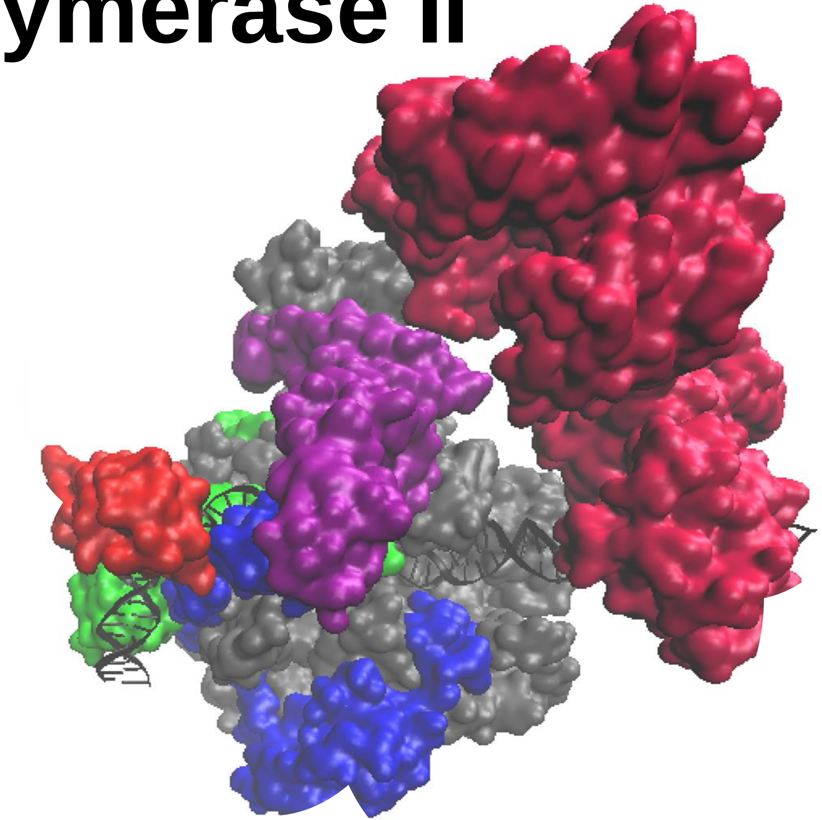
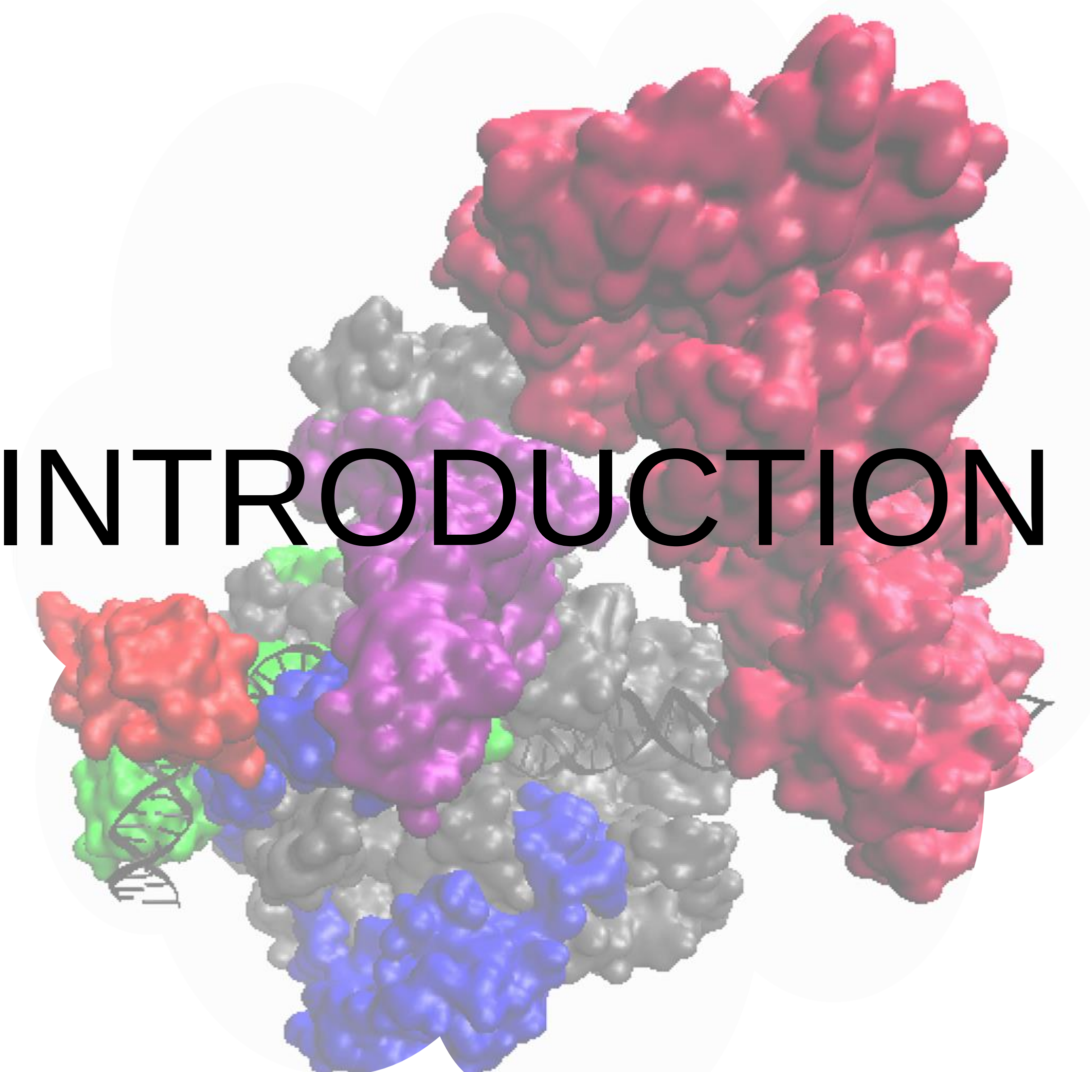
Structural basis of transcription initiation by RNA polymerase II





Josep Argerich Bergedà
Adrián González Sánchez
Miquel Àngel Schikora Tamarit



INTRODUCTION

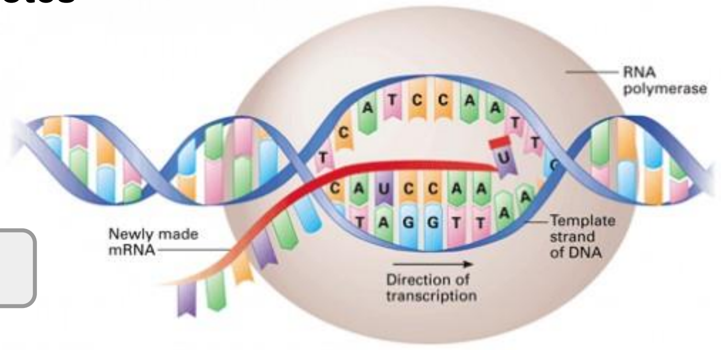
Transcription initiation is the start of gene expression in eukaryotes

➢ 3 nuclear RNA polymerases orchestrate the process

Pol-I rRNA

Pol-II mRNA

Pol-III tRNA



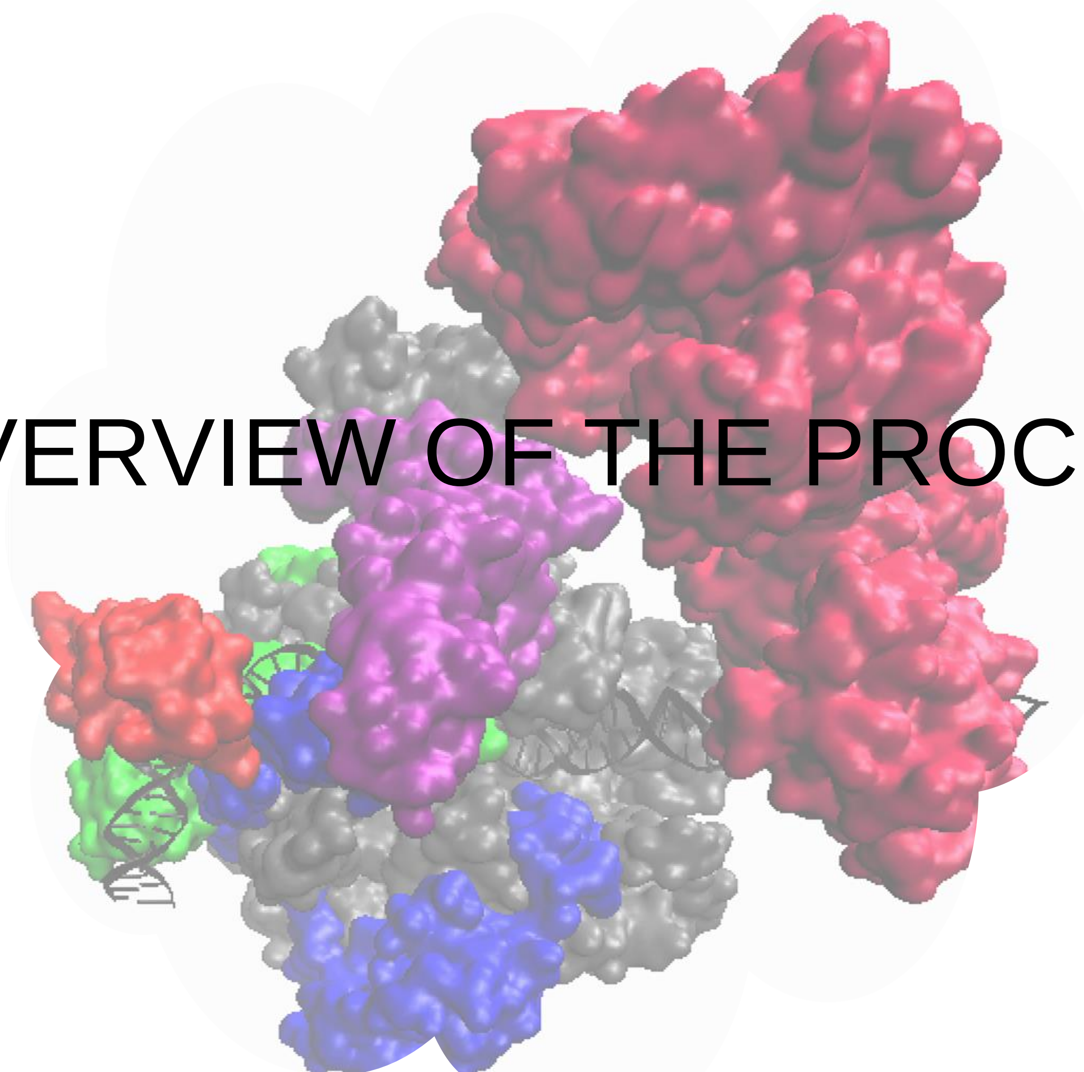
➢ Pol-II controls cell **identity**, **differentiation** and **stress response** processes

➢ 3 steps are required to produce RNA

Initiation → Elongation → Termination

➢ **Initiation** is the **starting point**, and integrates most of the transcriptional regulation

➢ The mechanism is supposed to be **highly conserved** across eukaryotes

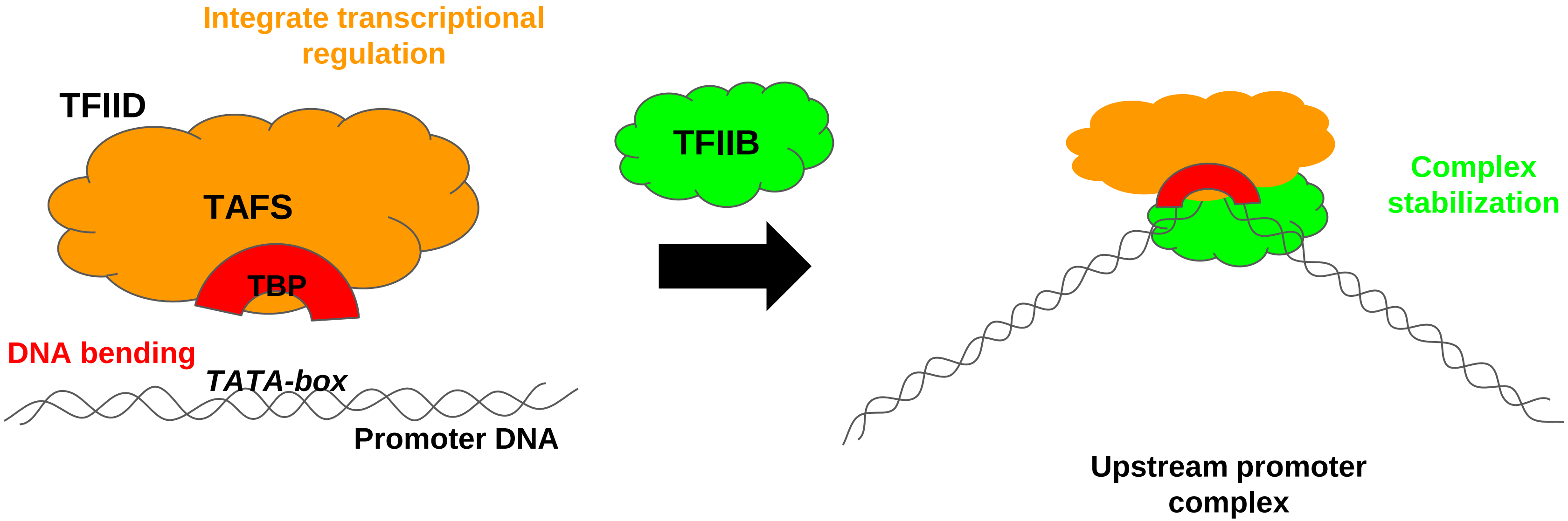


OVERVIEW OF THE PROCESS

Transcription initiation requires the assembly of the General Transcription Factors (GTF)

1. TFIID binds and bends DNA around the TATA region

2. TFIIB joins the complex



Integrate transcriptional regulation

TFIID

TAFS

TBP

DNA bending

TATA-box

Promoter DNA

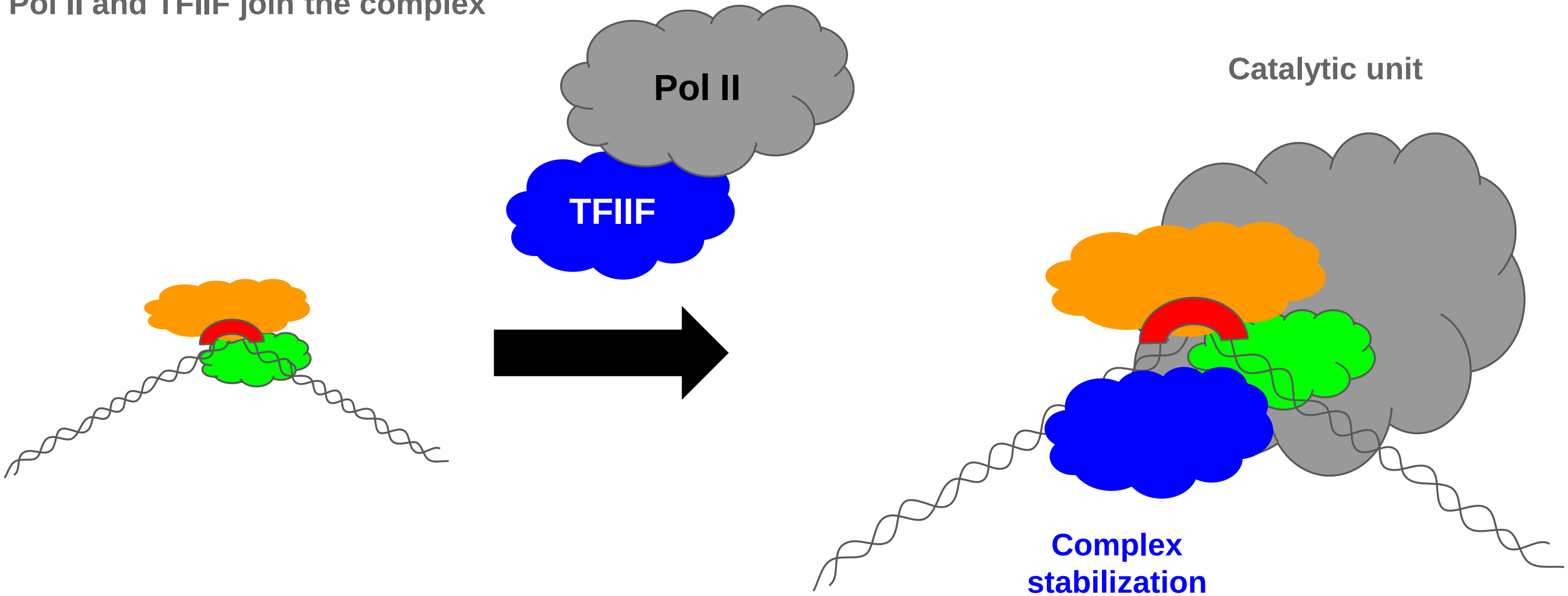
TFIIB

Complex stabilization

Upstream promoter complex

Transcription initiation requires the assembly of the General Transcription Factors (GTF)

3. Pol II and TFIIF join the complex

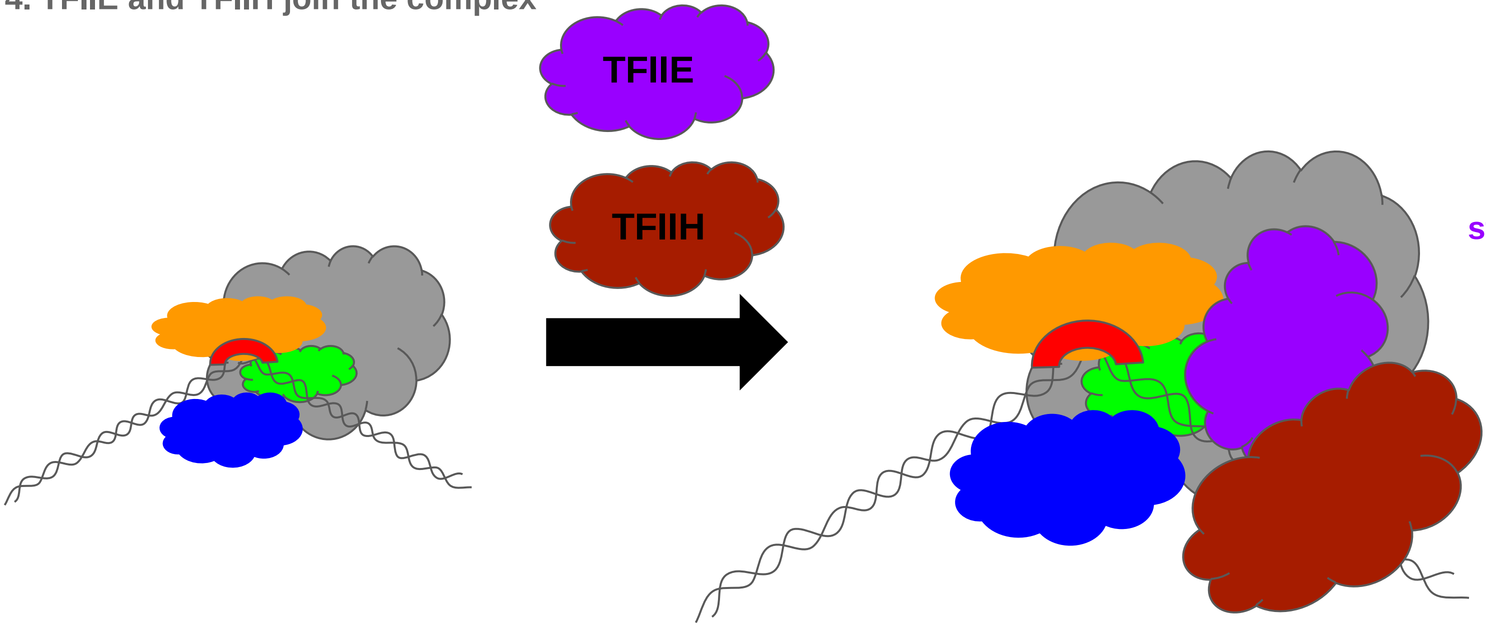


Complex stabilization

Core Preinitiation Complex (PIC)

Transcription initiation requires the assembly of the General Transcription Factors (GTF)

4. TFIIE and TFIIH join the complex



Complex stabilization

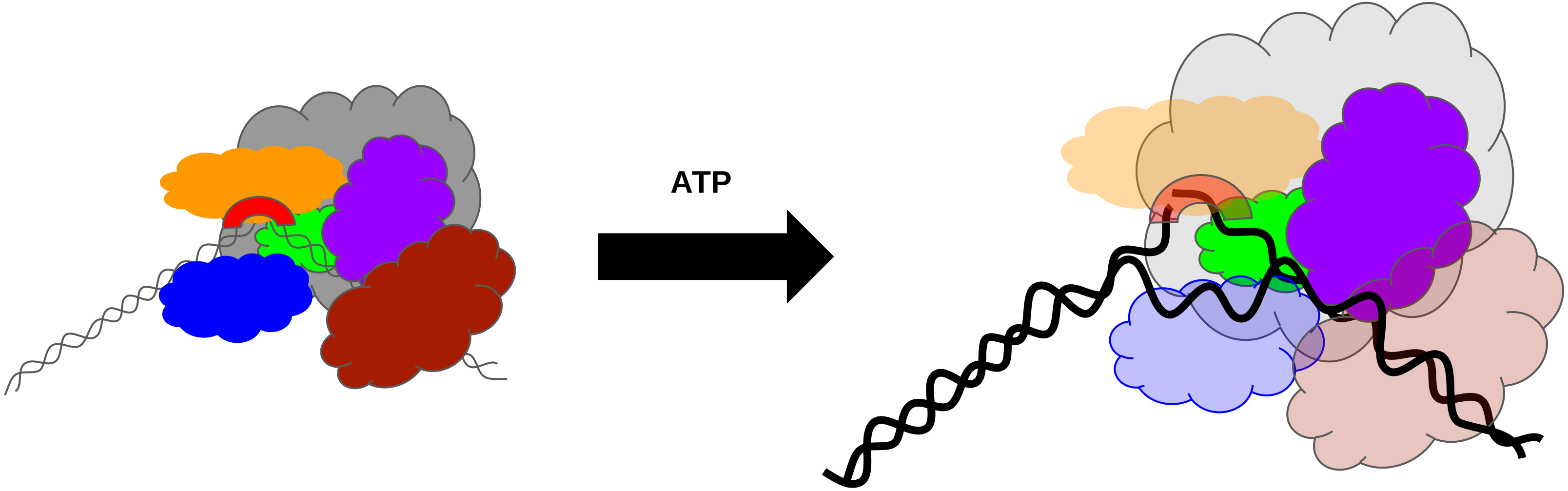
DNA unwinding

Pol-II activation

Closed PIC

Transcription initiation requires the assembly of the General Transcription Factors (GTF)

5. DNA melts around the transcription start site (TSS)

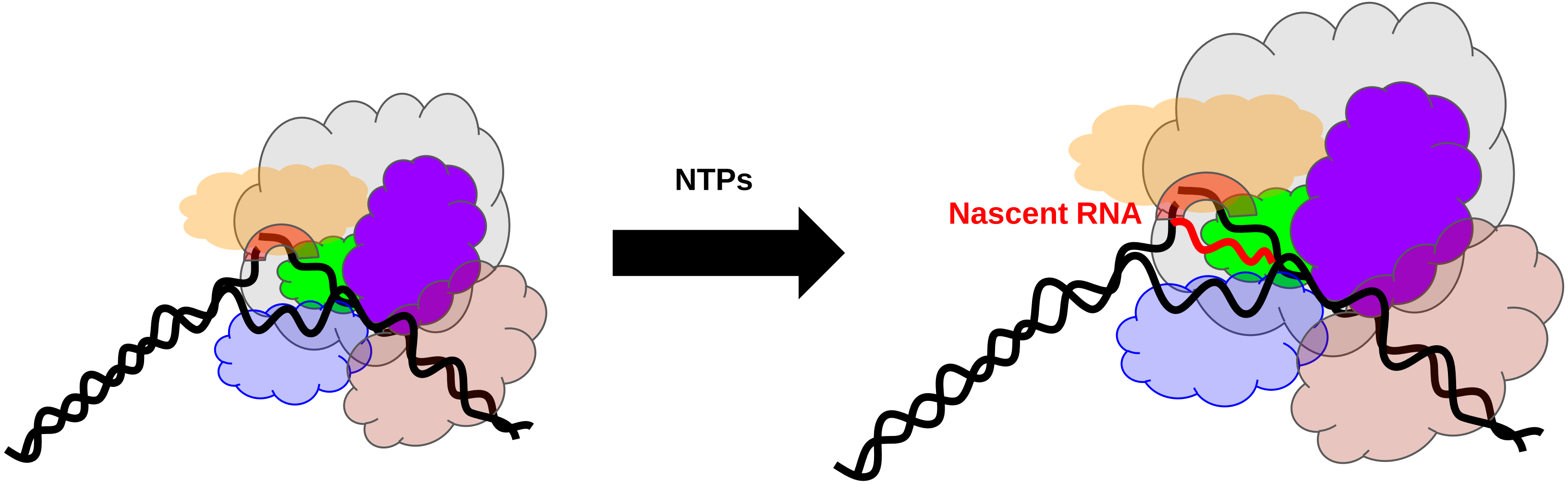


ATP

Open PIC

Transcription initiation requires the assembly of the General Transcription Factors (GTF)

6. Transcription starts

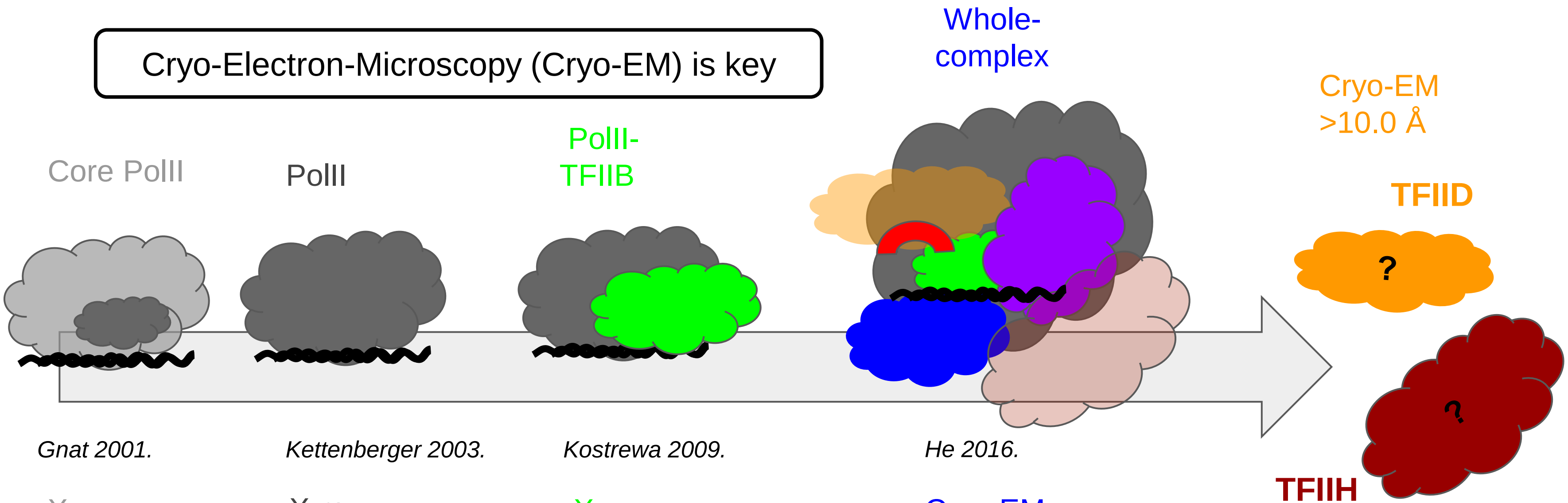


Initially transcribing
complex

The architecture of the complex is not fully understood at atomic resolution

Historical highlights on the advances in the anatomical comprehension of transcription initiation

Cryo-Electron-Microscopy (Cryo-EM) is key



Core PolII

Gnat 2001.

X-ray
3.3 Å

PolII

Kettenberger 2003.

X-ray
3.8 Å

PolII-
TFIIB

Kostrewa 2009.

X-ray
4.3 Å

Whole-
complex

He 2016.

Cryo-EM
3.9 Å

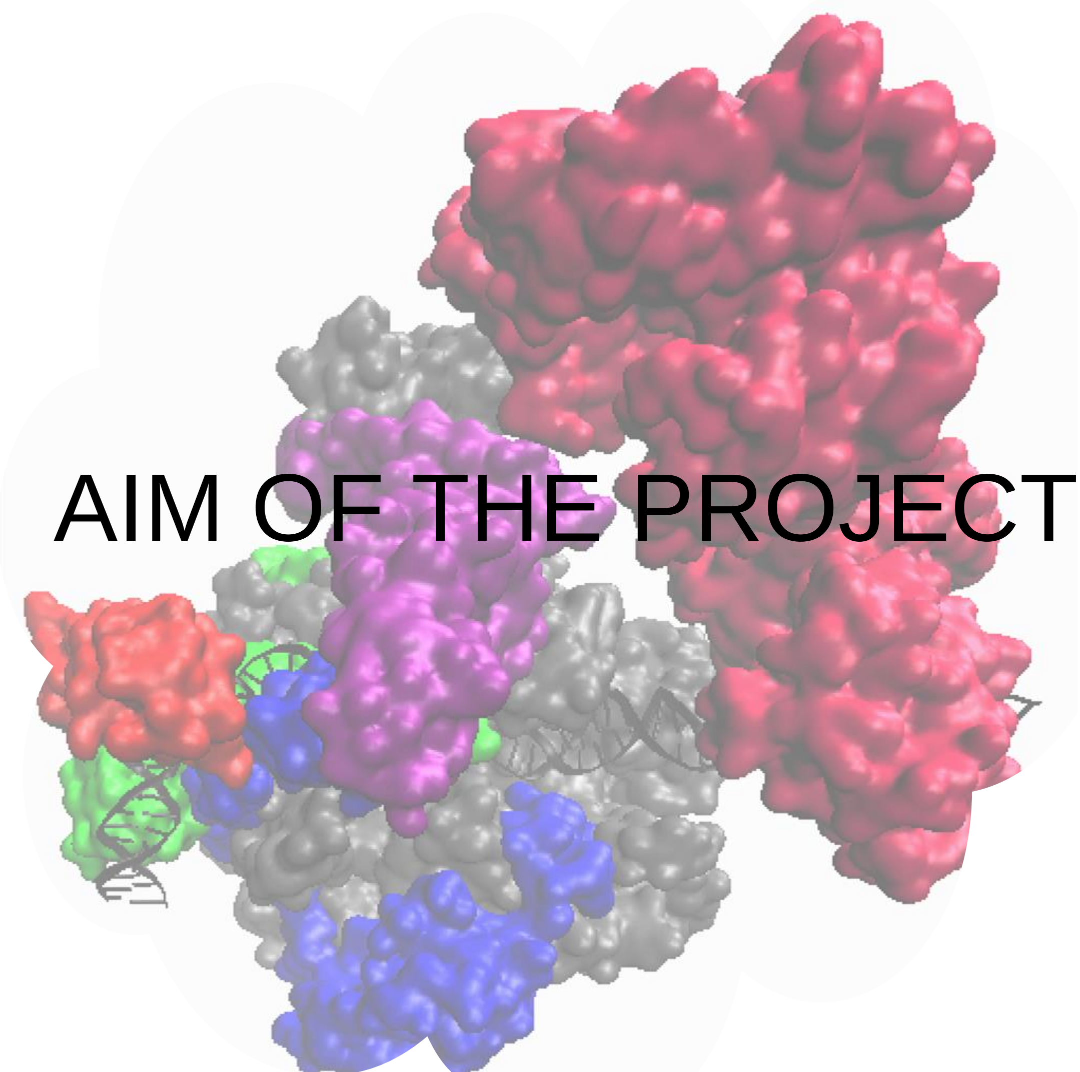
Cryo-EM
>10.0 Å

TFIID

TFIIH

Cryo-EM
7.9 - 10.0 Å

Most subunits are characterized in yeast (*S. cerevisiae*) and humans (*H. sapiens*)



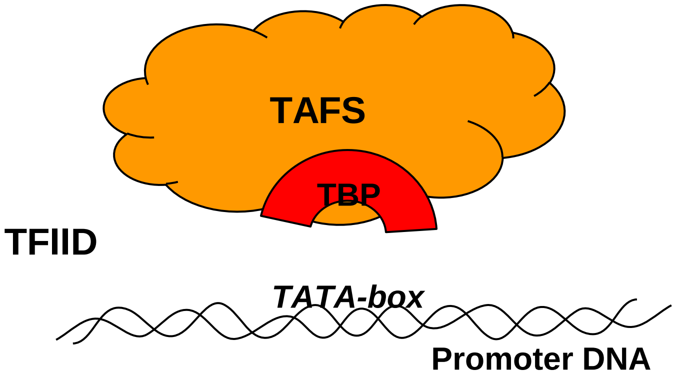
AIM OF THE PROJECT

Our project aims to review the...

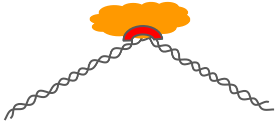
- ➢ **Architecture** of the initiation complex
- ➢ **Relevant interactions** to understand the sequential assembly of factors
- ➢ **Molecular mechanisms** beyond the structure
- ➢ **Evolutionary conservation** across eukaryotes
- ➢ The **current** structural **knowledge** about transcription initiation

Promoter recognition:

TFIID binds and bends DNA around the TATA region



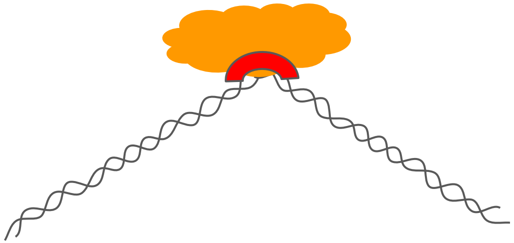
Promoter recognition: TFIID binds and bends DNA around the TATA region



TBP-associated factors

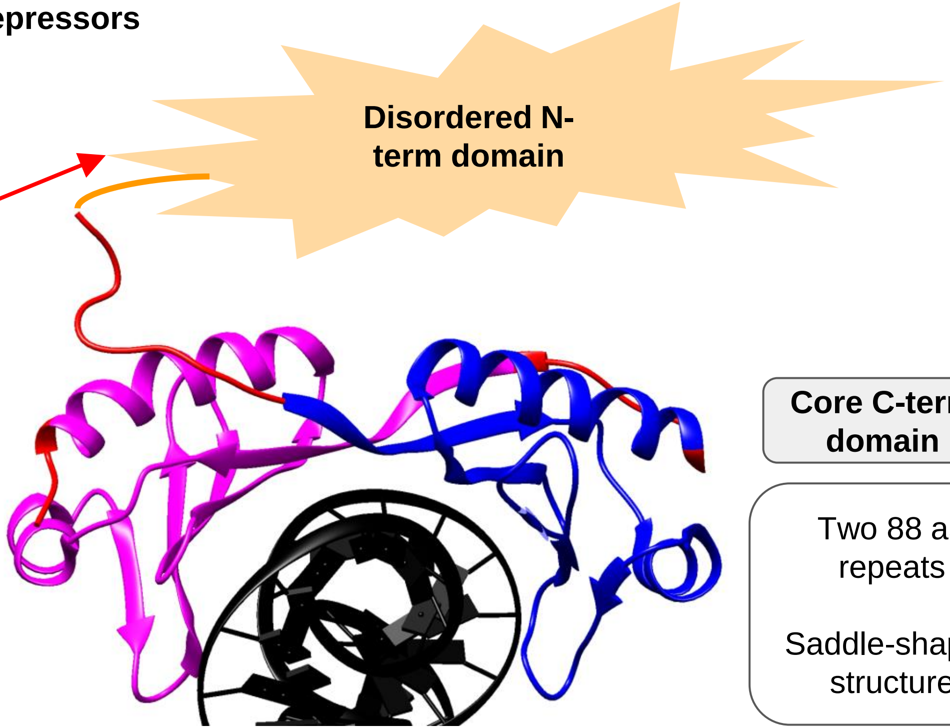
Integrate

Transcription factors
Activators
Repressors



TATA-binding protein (TBP)
(339 aa)

Cristalized part



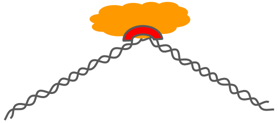
Disordered N-term domain

Core C-term domain

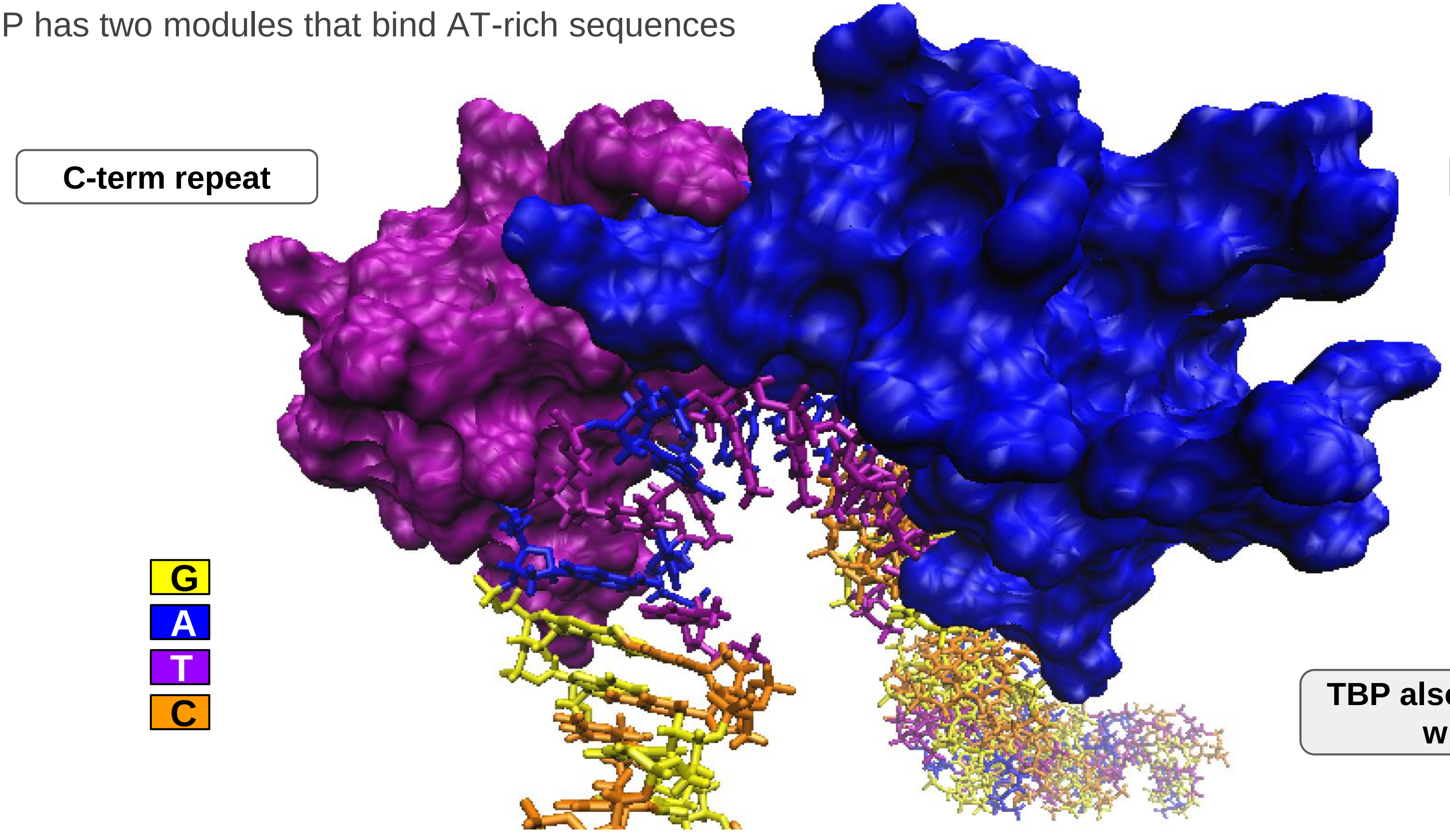
Two 88 aa repeats

Saddle-shaped structure

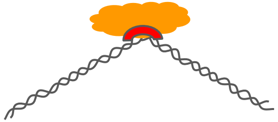
Promoter recognition: TFIID binds and bends DNA around the TATA region



TBP has two modules that bind AT-rich sequences

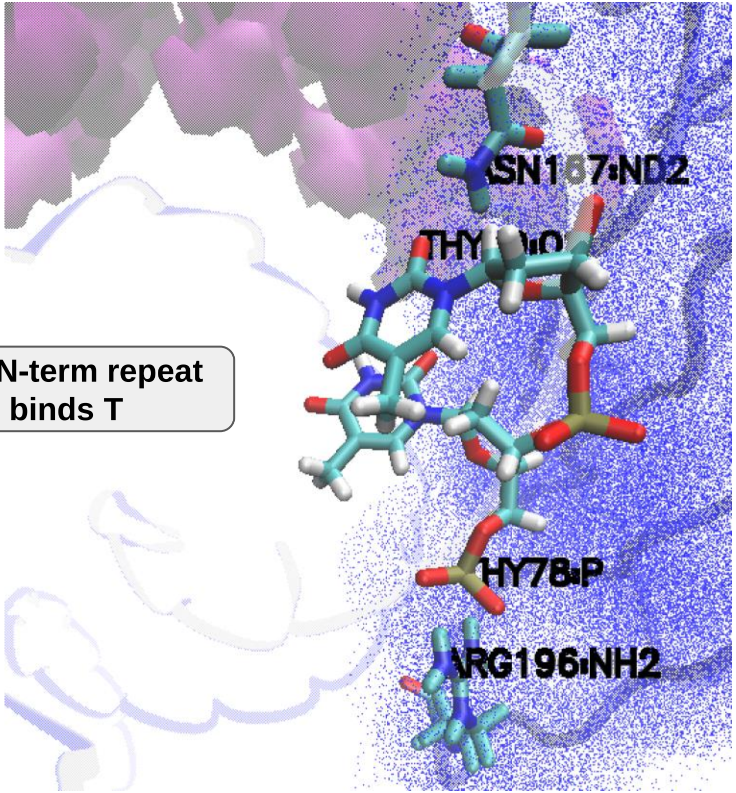


TBP also binds promoters without TATA

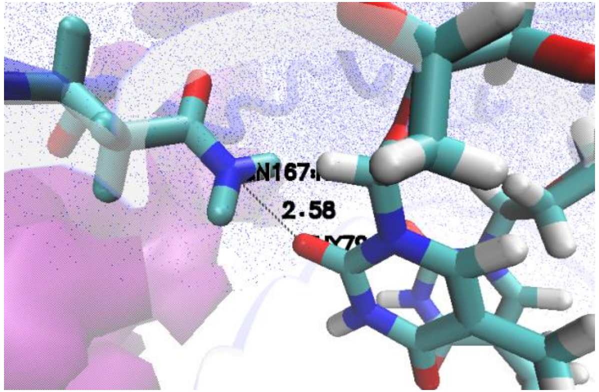


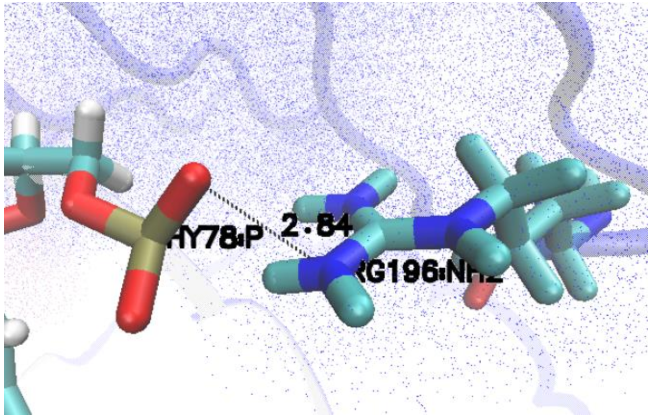
Promoter recognition: TFIID binds and bends DNA around the TATA region

H bonds and electrostatic interactions are required



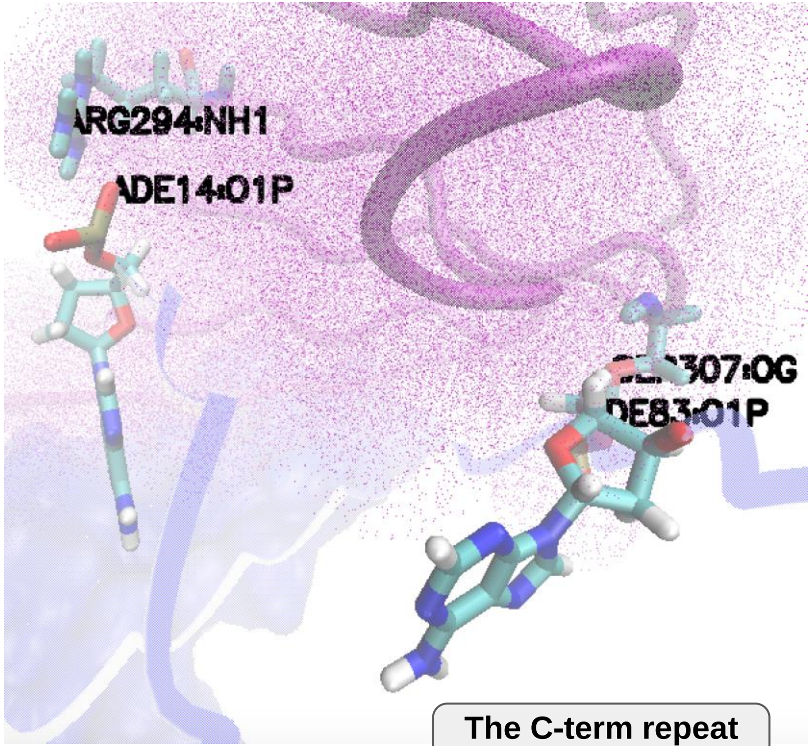
The N-term repeat binds T



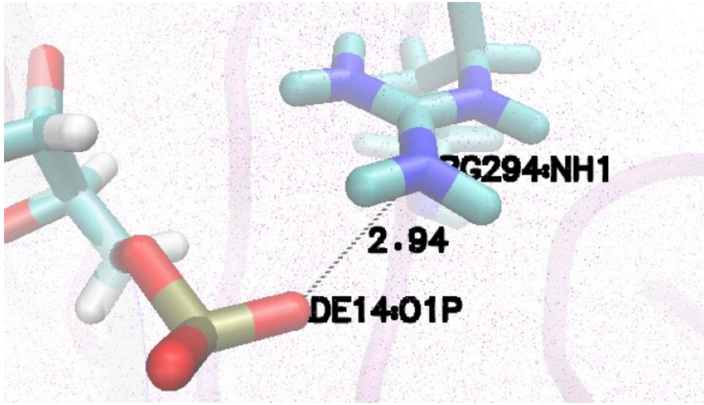


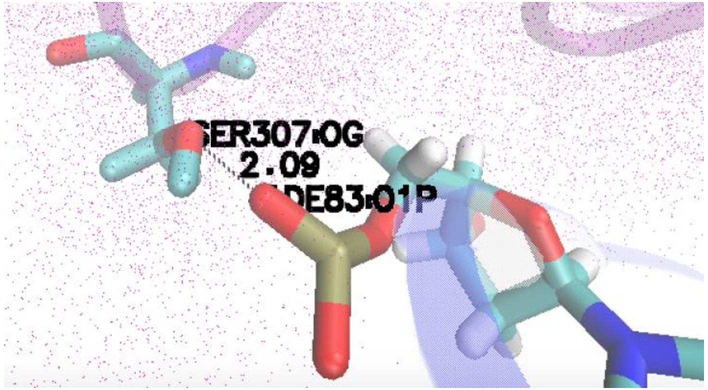
Promoter recognition: TFIID binds and bends DNA around the TATA region

H bonds and electrostatic interactions are required

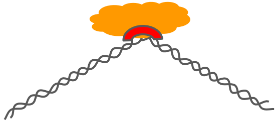


The C-term repeat binds A

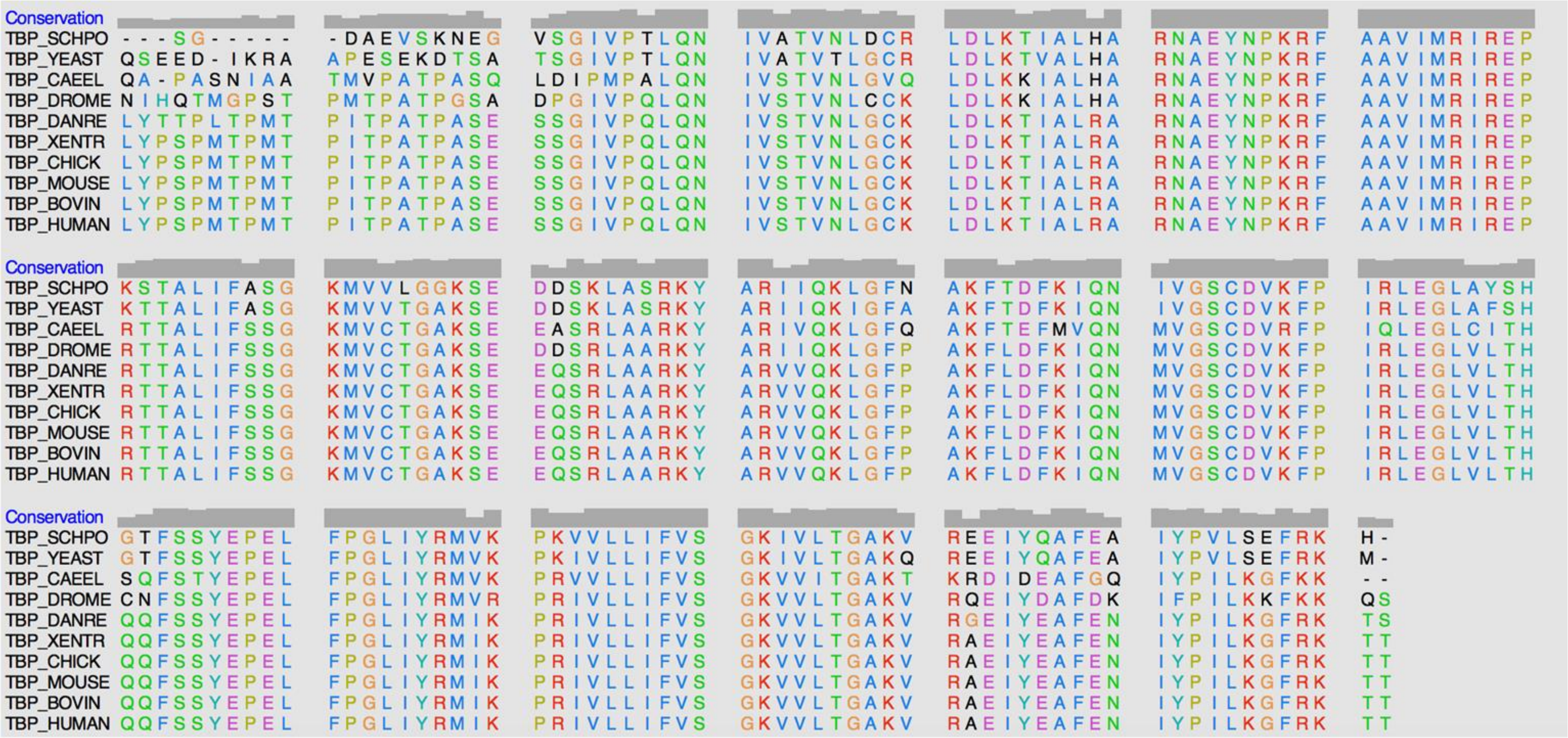




Promoter recognition: TFIID binds and bends DNA arround the TATA region



The C-terminal domain of TBP is conserved across eukaryotes

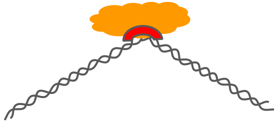


Conservation							
TBP_SCHPO	- - - S G - - - - -	- D A E V S K N E G	V S G I V P T L Q N	I V A T V N L D C R	L D L K T I A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_YEAST	Q S E E D - I K R A	A P E S E K D T S A	T S G I V P T L Q N	I V A T V T L G C R	L D L K T V A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_CAEEL	Q A - P A S N I A A	T M V P A T P A S Q	L D I P M P A L Q N	I V S T V N L G V Q	L D L K K I A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_DROME	N I H Q T M G P S T	P M T P A T P G S A	D P G I V P Q L Q N	I V S T V N L C C K	L D L K K I A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_DANRE	L Y T T P L T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_XENTR	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_CHICK	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_MOUSE	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_BOVIN	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_HUMAN	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P

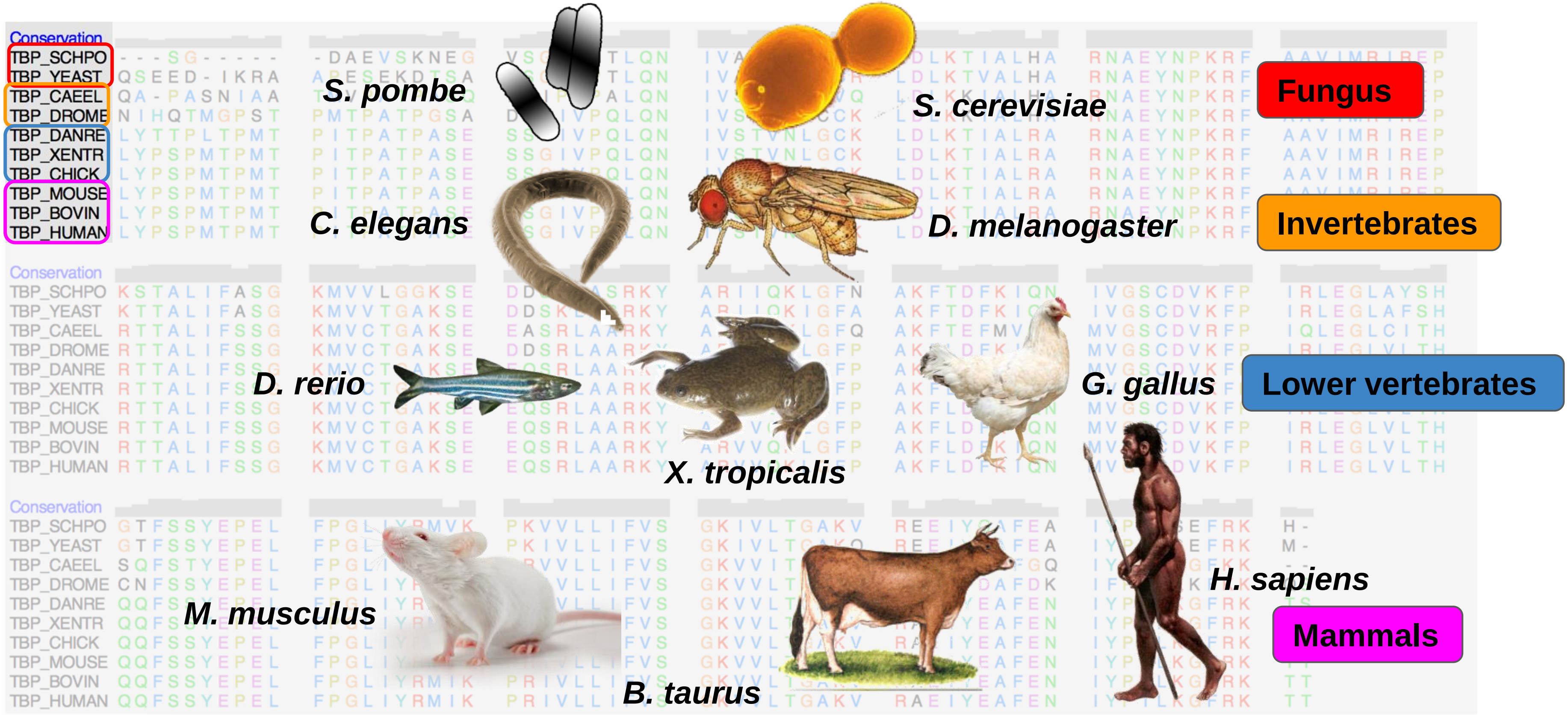
Conservation							
TBP_SCHPO	K S T A L I F A S G	K M V V L G G K S E	D D S K L A S R K Y	A R I I Q K L G F N	A K F T D F K I Q N	I V G S C D V K F P	I R L E G L A Y S H
TBP_YEAST	K T T A L I F A S G	K M V V T G A K S E	D D S K L A S R K Y	A R I I Q K I G F A	A K F T D F K I Q N	I V G S C D V K F P	I R L E G L A F S H
TBP_CAEEL	R T T A L I F S S G	K M V C T G A K S E	E A S R L A A R K Y	A R I V Q K L G F Q	A K F T E F M V Q N	M V G S C D V R F P	I Q L E G L C I T H
TBP_DROME	R T T A L I F S S G	K M V C T G A K S E	D D S R L A A R K Y	A R I I Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_DANRE	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_XENTR	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_CHICK	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_MOUSE	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_BOVIN	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_HUMAN	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H

Conservation							
TBP_SCHPO	G T F S S Y E P E L	F P G L I Y R M V K	P K V V L L I F V S	G K I V L T G A K V	R E E I Y Q A F E A	I Y P V L S E F R K	H -
TBP_YEAST	G T F S S Y E P E L	F P G L I Y R M V K	P K I V L L I F V S	G K I V L T G A K Q	R E E I Y Q A F E A	I Y P V L S E F R K	M -
TBP_CAEEL	S Q F S T Y E P E L	F P G L I Y R M V K	P R V V L L I F V S	G K V V I T G A K T	K R D I D E A F G Q	I Y P I L K G F K K	- -
TBP_DROME	C N F S S Y E P E L	F P G L I Y R M V R	P R I V L L I F V S	G K V V L T G A K V	R Q E I Y D A F D K	I F P I L K K F K K	Q S
TBP_DANRE	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R G E I Y E A F E N	I Y P I L K G F R K	T S
TBP_XENTR	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_CHICK	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_MOUSE	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_BOVIN	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_HUMAN	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T

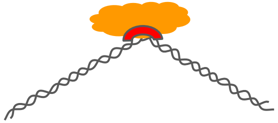
Promoter recognition: TFIID binds and bends DNA arround the TATA region



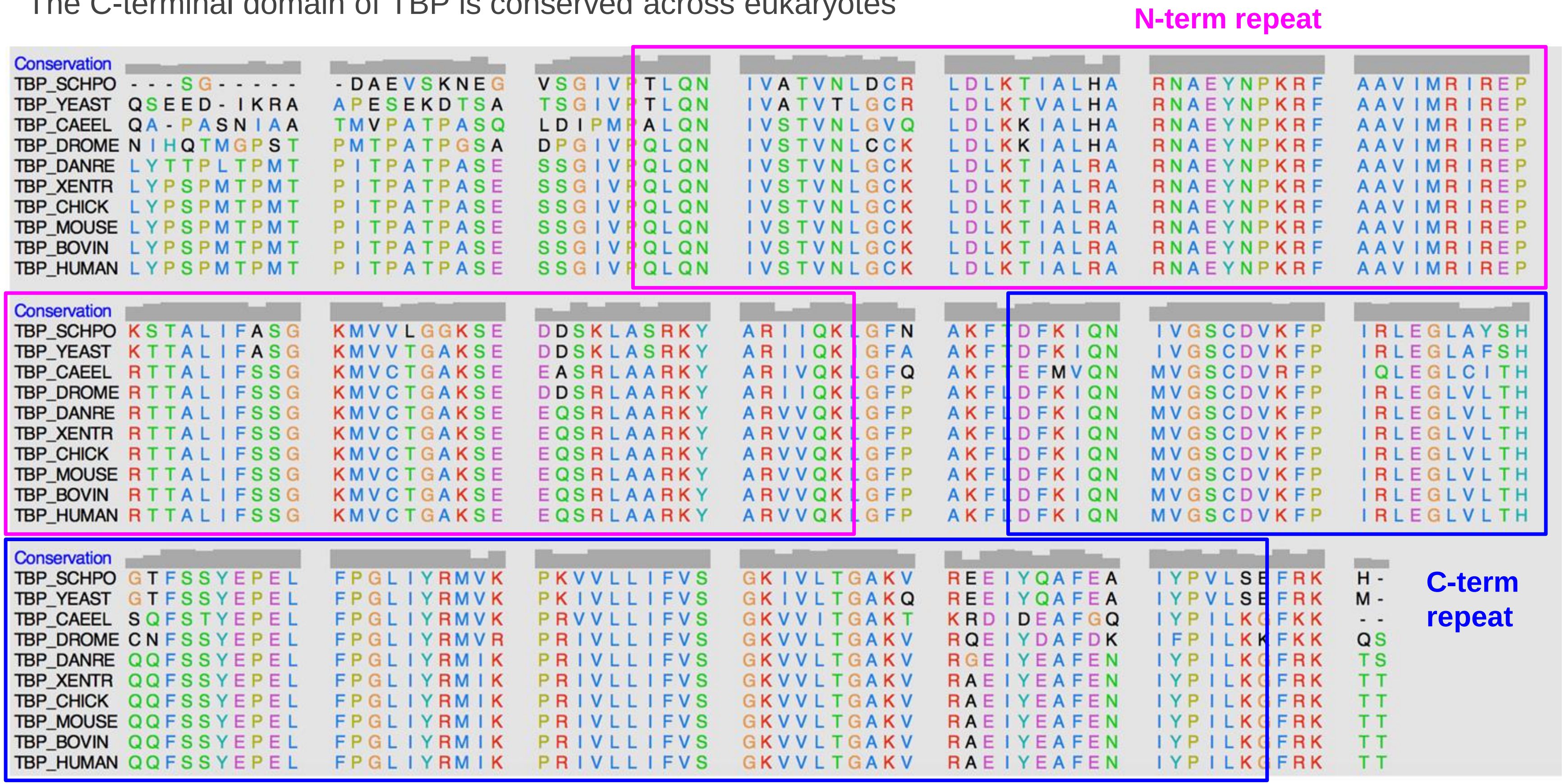
The C-terminal domain of TBP is conserved across eukaryotes



Promoter recognition: TFIID binds and bends DNA arround the TATA region



The C-terminal domain of TBP is conserved across eukaryotes

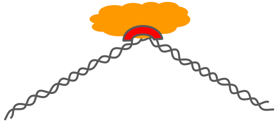


Conservation							
TBP_SCHPO	---SG-----	-DAEVSKNEG	VSGIVPTLQN	IVATVNLDCR	LDLKTIALHA	RNAEYNPKRF	AAVIMRIREP
TBP_YEAST	QSEED-IKRA	APESEKDTSA	TSGIVPTLQN	IVATVTLGCR	LDLKTVALHA	RNAEYNPKRF	AAVIMRIREP
TBP_CAEEL	QA-PASNIAA	TMVPATPASQ	LDIPMPALQN	IVSTVNLGVQ	LDLKKIALHA	RNAEYNPKRF	AAVIMRIREP
TBP_DROME	NIHQTMGPST	PMTPATPGSA	DPGIVPQLQN	IVSTVNLCCK	LDLKKIALHA	RNAEYNPKRF	AAVIMRIREP
TBP_DANRE	LYTTPLTPMT	PITPATPASE	SSGIVPQLQN	IVSTVNLGCK	LDLKTIALRA	RNAEYNPKRF	AAVIMRIREP
TBP_XENTR	LYPSPMTPMT	PITPATPASE	SSGIVPQLQN	IVSTVNLGCK	LDLKTIALRA	RNAEYNPKRF	AAVIMRIREP
TBP_CHICK	LYPSPMTPMT	PITPATPASE	SSGIVPQLQN	IVSTVNLGCK	LDLKTIALRA	RNAEYNPKRF	AAVIMRIREP
TBP_MOUSE	LYPSPMTPMT	PITPATPASE	SSGIVPQLQN	IVSTVNLGCK	LDLKTIALRA	RNAEYNPKRF	AAVIMRIREP
TBP_BOVIN	LYPSPMTPMT	PITPATPASE	SSGIVPQLQN	IVSTVNLGCK	LDLKTIALRA	RNAEYNPKRF	AAVIMRIREP
TBP_HUMAN	LYPSPMTPMT	PITPATPASE	SSGIVPQLQN	IVSTVNLGCK	LDLKTIALRA	RNAEYNPKRF	AAVIMRIREP

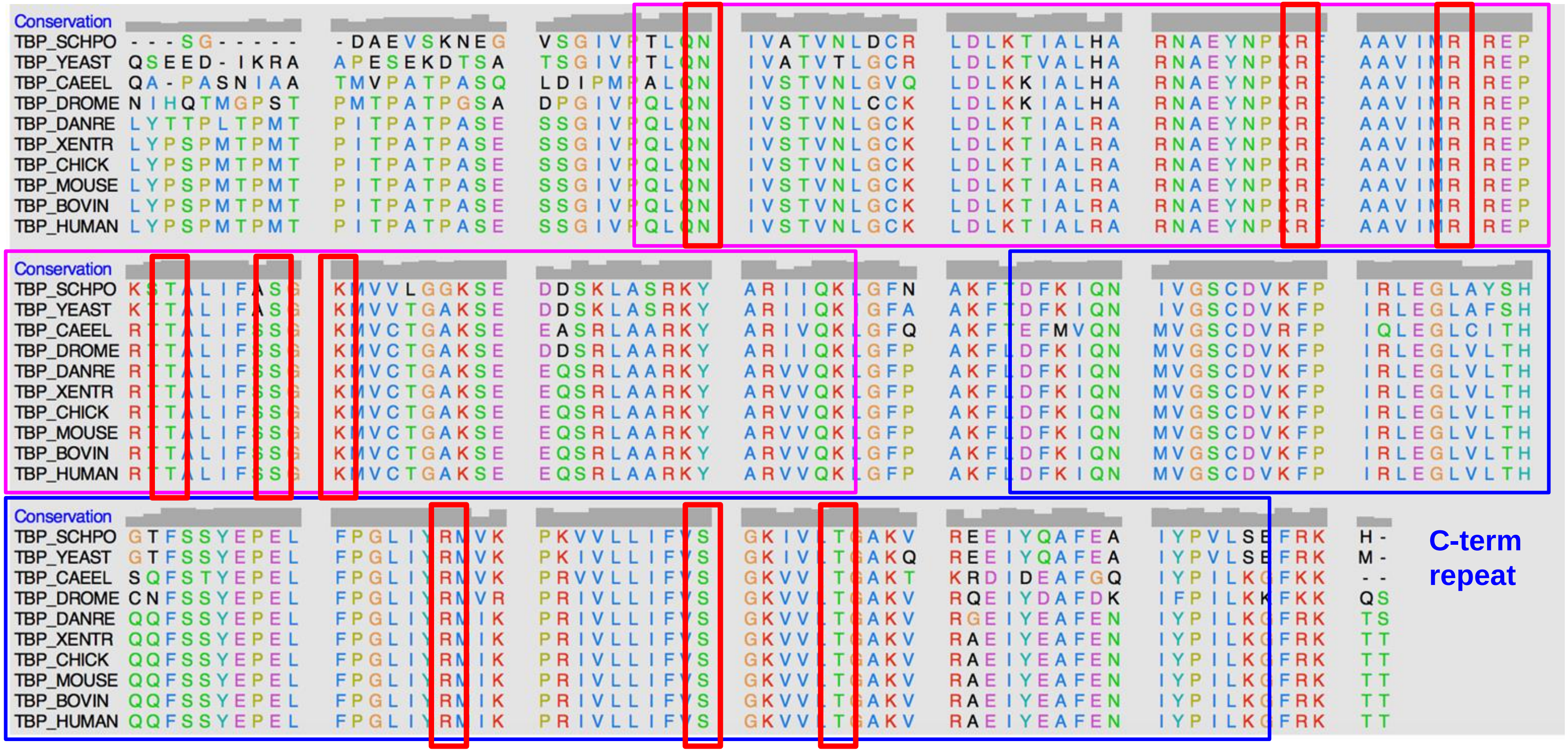
Conservation							
TBP_SCHPO	KSTALIFASG	KMVVLGGKSE	DDSKLASRKY	ARIIQKLGFN	AKFTDFKIQN	IVGSCDVKFP	IRLEGLAYSH
TBP_YEAST	KTTALIFASG	KMVVTGAKSE	DDSKLASRKY	ARIIQKLGFA	AKFTDFKIQN	IVGSCDVKFP	IRLEGLAFSH
TBP_CAEEL	RTTALIFSSG	KMVCTGAKSE	EASRLAARKY	ARIVQKLGFQ	AKFTEFMVQN	MVGSCDVRFP	IQLEGLCITH
TBP_DROME	RTTALIFSSG	KMVCTGAKSE	DDSRLAARKY	ARIIQKLGFP	AKFLDFKIQN	MVGSCDVKFP	IRLEGLVLTH
TBP_DANRE	RTTALIFSSG	KMVCTGAKSE	EQSRLAARKY	ARVVQKLGFP	AKFLDFKIQN	MVGSCDVKFP	IRLEGLVLTH
TBP_XENTR	RTTALIFSSG	KMVCTGAKSE	EQSRLAARKY	ARVVQKLGFP	AKFLDFKIQN	MVGSCDVKFP	IRLEGLVLTH
TBP_CHICK	RTTALIFSSG	KMVCTGAKSE	EQSRLAARKY	ARVVQKLGFP	AKFLDFKIQN	MVGSCDVKFP	IRLEGLVLTH
TBP_MOUSE	RTTALIFSSG	KMVCTGAKSE	EQSRLAARKY	ARVVQKLGFP	AKFLDFKIQN	MVGSCDVKFP	IRLEGLVLTH
TBP_BOVIN	RTTALIFSSG	KMVCTGAKSE	EQSRLAARKY	ARVVQKLGFP	AKFLDFKIQN	MVGSCDVKFP	IRLEGLVLTH
TBP_HUMAN	RTTALIFSSG	KMVCTGAKSE	EQSRLAARKY	ARVVQKLGFP	AKFLDFKIQN	MVGSCDVKFP	IRLEGLVLTH

Conservation							
TBP_SCHPO	GTFSSYEPEL	FPGLIYRMVK	PKVVLLIFVS	GKIVLTGAKV	REEIYQAFEA	IYPVLSEFRK	H-
TBP_YEAST	GTFSSYEPEL	FPGLIYRMVK	PKIVLLIFVS	GKIVLTGAKQ	REEIYQAFEA	IYPVLSEFRK	M-
TBP_CAEEL	SQFSTYEPEL	FPGLIYRMVK	PRVVLLIFVS	GKVVITGAKT	KRDIDEAFGQ	IYPILKGFKK	--
TBP_DROME	CNFSSYEPEL	FPGLIYRMVR	PRIVLLIFVS	GKVVLTGAKV	RQEIYDAFDK	IFPILKKFKK	QS
TBP_DANRE	QQFSSYEPEL	FPGLIYRMIK	PRIVLLIFVS	GKVVLTGAKV	RGEIYEAFEN	IYPILKGFRK	TS
TBP_XENTR	QQFSSYEPEL	FPGLIYRMIK	PRIVLLIFVS	GKVVLTGAKV	RAEIYEAFEN	IYPILKGFRK	TT
TBP_CHICK	QQFSSYEPEL	FPGLIYRMIK	PRIVLLIFVS	GKVVLTGAKV	RAEIYEAFEN	IYPILKGFRK	TT
TBP_MOUSE	QQFSSYEPEL	FPGLIYRMIK	PRIVLLIFVS	GKVVLTGAKV	RAEIYEAFEN	IYPILKGFRK	TT
TBP_BOVIN	QQFSSYEPEL	FPGLIYRMIK	PRIVLLIFVS	GKVVLTGAKV	RAEIYEAFEN	IYPILKGFRK	TT
TBP_HUMAN	QQFSSYEPEL	FPGLIYRMIK	PRIVLLIFVS	GKVVLTGAKV	RAEIYEAFEN	IYPILKGFRK	TT

Promoter recognition: TFIID binds and bends DNA arround the TATA region



The C-terminal domain of TBP is conserved across eukaryotes



N-term repeat

Conservation							
TBP_SCHPO	- - - S G - - - - -	- D A E V S K N E G	V S G I V P T L Q N	I V A T V N L D C R	L D L K T I A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_YEAST	Q S E E D - I K R A	A P E S E K D T S A	T S G I V P T L Q N	I V A T V T L G C R	L D L K T V A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_CAEEL	Q A - P A S N I A A	T M V P A T P A S Q	L D I P M P A L Q N	I V S T V N L G V Q	L D L K K I A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_DROME	N I H Q T M G P S T	P M T P A T P G S A	D P G I V P Q L Q N	I V S T V N L C C K	L D L K K I A L H A	R N A E Y N P K R F	A A V I M R I R E P
TBP_DANRE	L Y T T P L T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_XENTR	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_CHICK	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_MOUSE	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_BOVIN	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P
TBP_HUMAN	L Y P S P M T P M T	P I T P A T P A S E	S S G I V P Q L Q N	I V S T V N L G C K	L D L K T I A L R A	R N A E Y N P K R F	A A V I M R I R E P

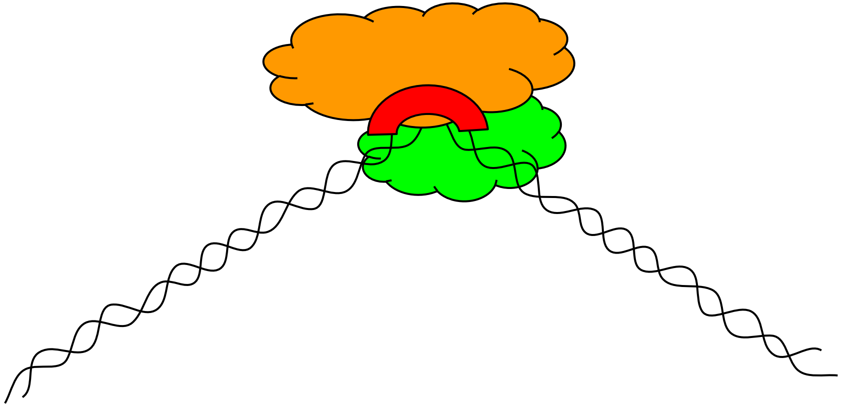
Conservation							
TBP_SCHPO	K S T A L I F A S G	K M V V L G G K S E	D D S K L A S R K Y	A R I I Q K L G F N	A K F T D F K I Q N	I V G S C D V K F P	I R L E G L A Y S H
TBP_YEAST	K T T A L I F A S G	K M V V T G A K S E	D D S K L A S R K Y	A R I I Q K I G F A	A K F T D F K I Q N	I V G S C D V K F P	I R L E G L A F S H
TBP_CAEEL	R T T A L I F S S G	K M V C T G A K S E	E A S R L A A R K Y	A R I V Q K L G F Q	A K F T E F M V Q N	M V G S C D V R F P	I Q L E G L C I T H
TBP_DROME	R T T A L I F S S G	K M V C T G A K S E	D D S R L A A R K Y	A R I I Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_DANRE	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_XENTR	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_CHICK	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_MOUSE	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_BOVIN	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H
TBP_HUMAN	R T T A L I F S S G	K M V C T G A K S E	E Q S R L A A R K Y	A R V V Q K L G F P	A K F L D F K I Q N	M V G S C D V K F P	I R L E G L V L T H

Conservation							
TBP_SCHPO	G T F S S Y E P E L	F P G L I Y R M V K	P K V V L L I F V S	G K I V L T G A K V	R E E I Y Q A F E A	I Y P V L S E F R K	H -
TBP_YEAST	G T F S S Y E P E L	F P G L I Y R M V K	P K I V L L I F V S	G K I V L T G A K Q	R E E I Y Q A F E A	I Y P V L S E F R K	M -
TBP_CAEEL	S Q F S T Y E P E L	F P G L I Y R M V K	P R V V L L I F V S	G K V V L T G A K T	K R D I D E A F G Q	I Y P I L K G F K K	- -
TBP_DROME	C N F S S Y E P E L	F P G L I Y R M V R	P R I V L L I F V S	G K V V L T G A K V	R Q E I Y D A F D K	I F P I L K K F K K	Q S
TBP_DANRE	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R G E I Y E A F E N	I Y P I L K G F R K	T S
TBP_XENTR	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_CHICK	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_MOUSE	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_BOVIN	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T
TBP_HUMAN	Q Q F S S Y E P E L	F P G L I Y R M I K	P R I V L L I F V S	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K	T T

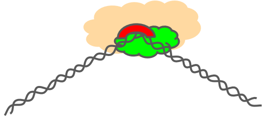
C-term repeat

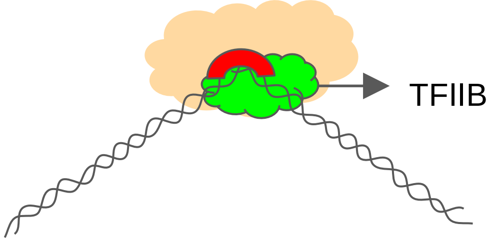
Upstream Promoter Complex formation:

TFIIB joins the complex

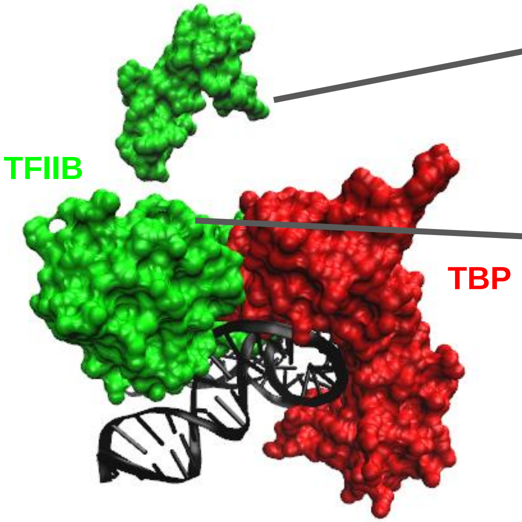


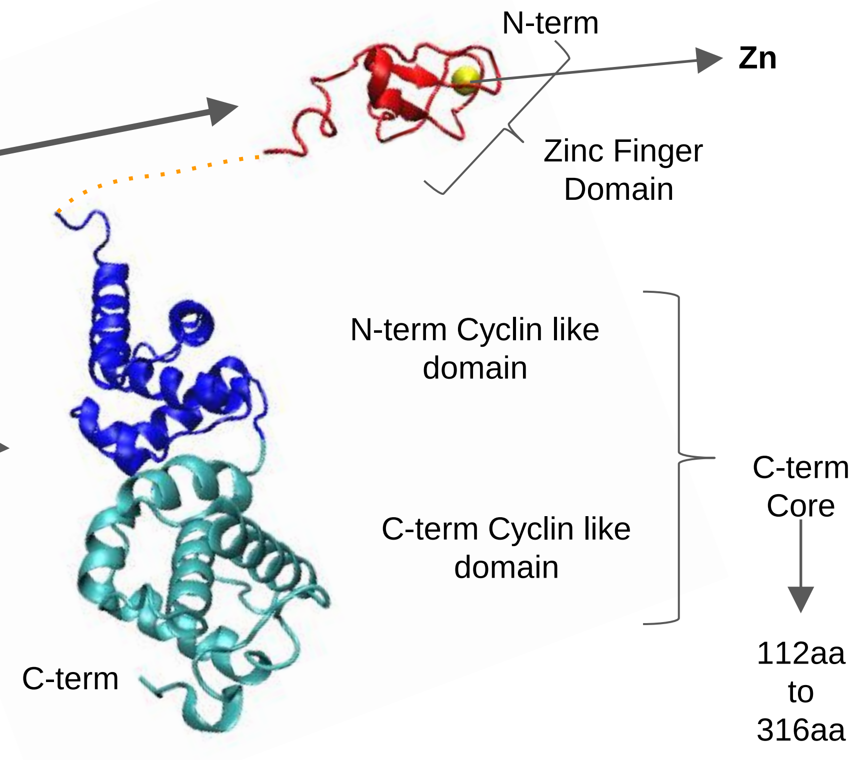
Upstream Promoter Complex formation: TFIIB joins the complex



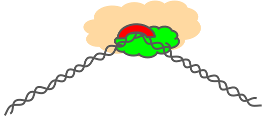


Transcription Factor II B (TFIIB) (316aa)

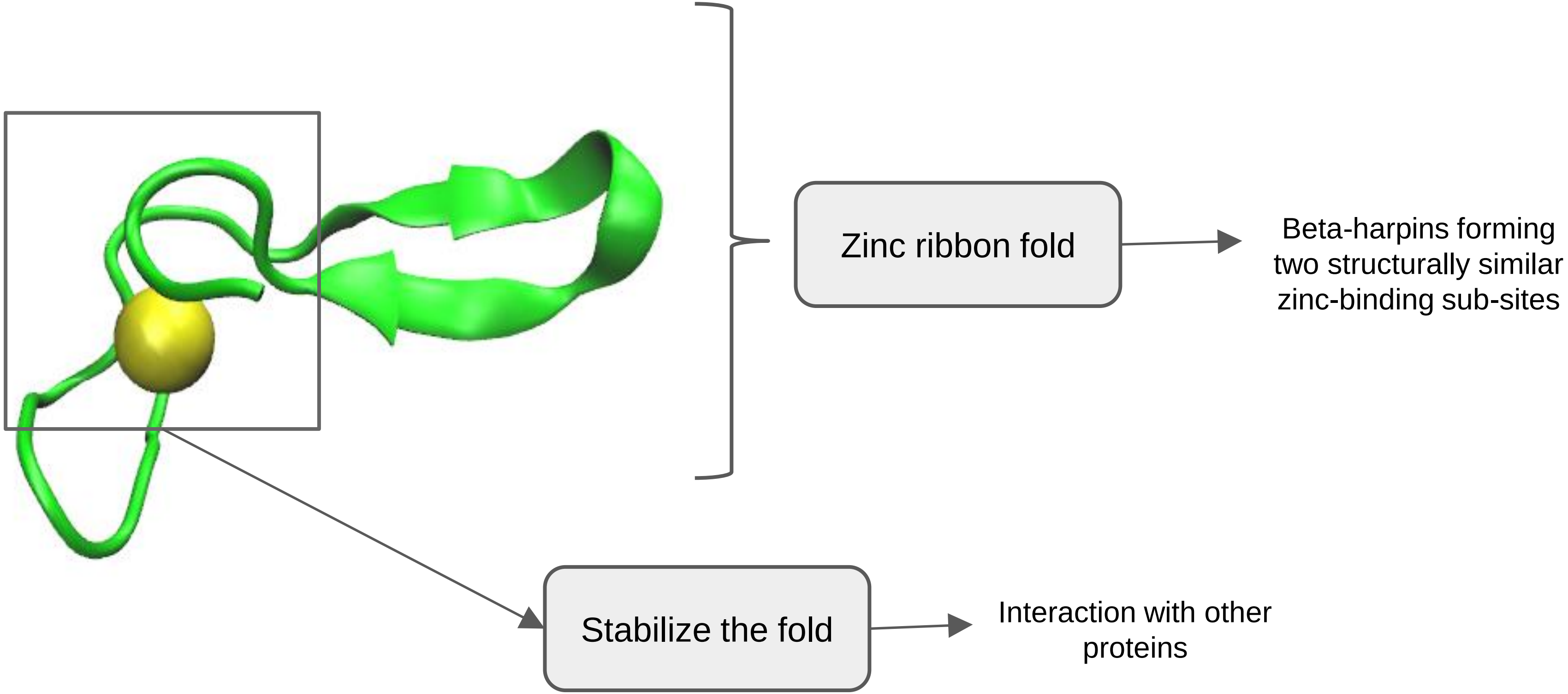




Upstream Promoter Complex formation: TFIIB joins the complex



The Zn-Finger domain is ubiquitous in transcription initiation (10 domains in the whole PIC)



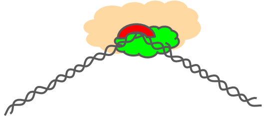
Zinc ribbon fold

Beta-harpins forming two structurally similar zinc-binding sub-sites

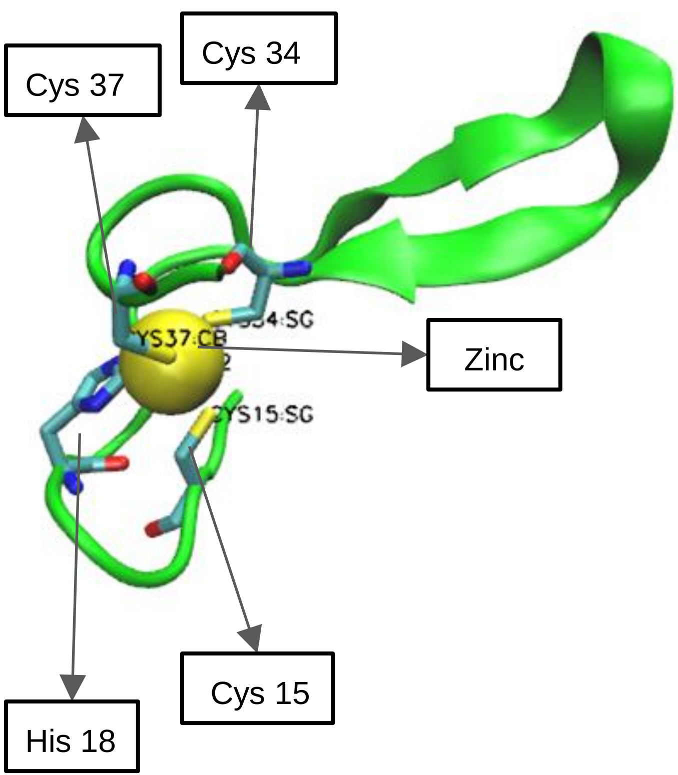
Stabilize the fold

Interaction with other proteins

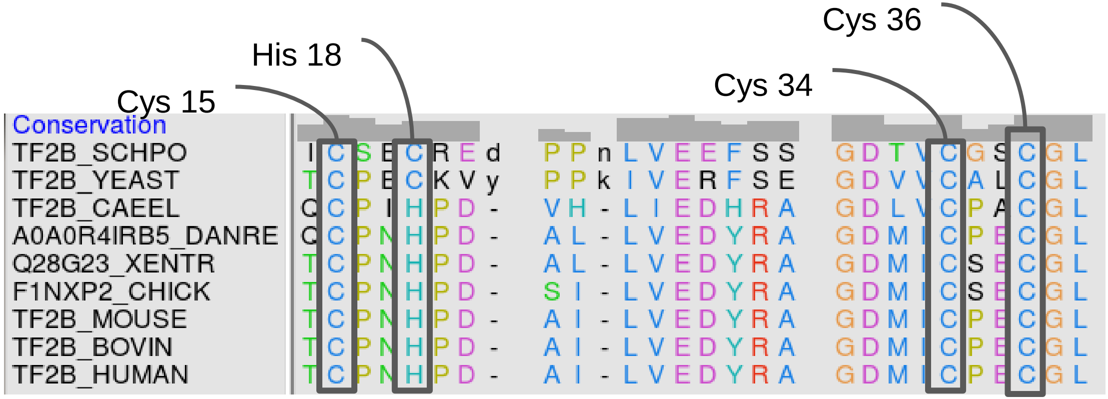
Upstream Promoter Complex formation: TFIIB joins the complex



The Zn-Finger domain in TFIIB



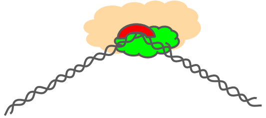
Sequence motif: Cys-Xaa_2-His/Cys-Xaa_{15}-Cys-Xaa_2-Cys



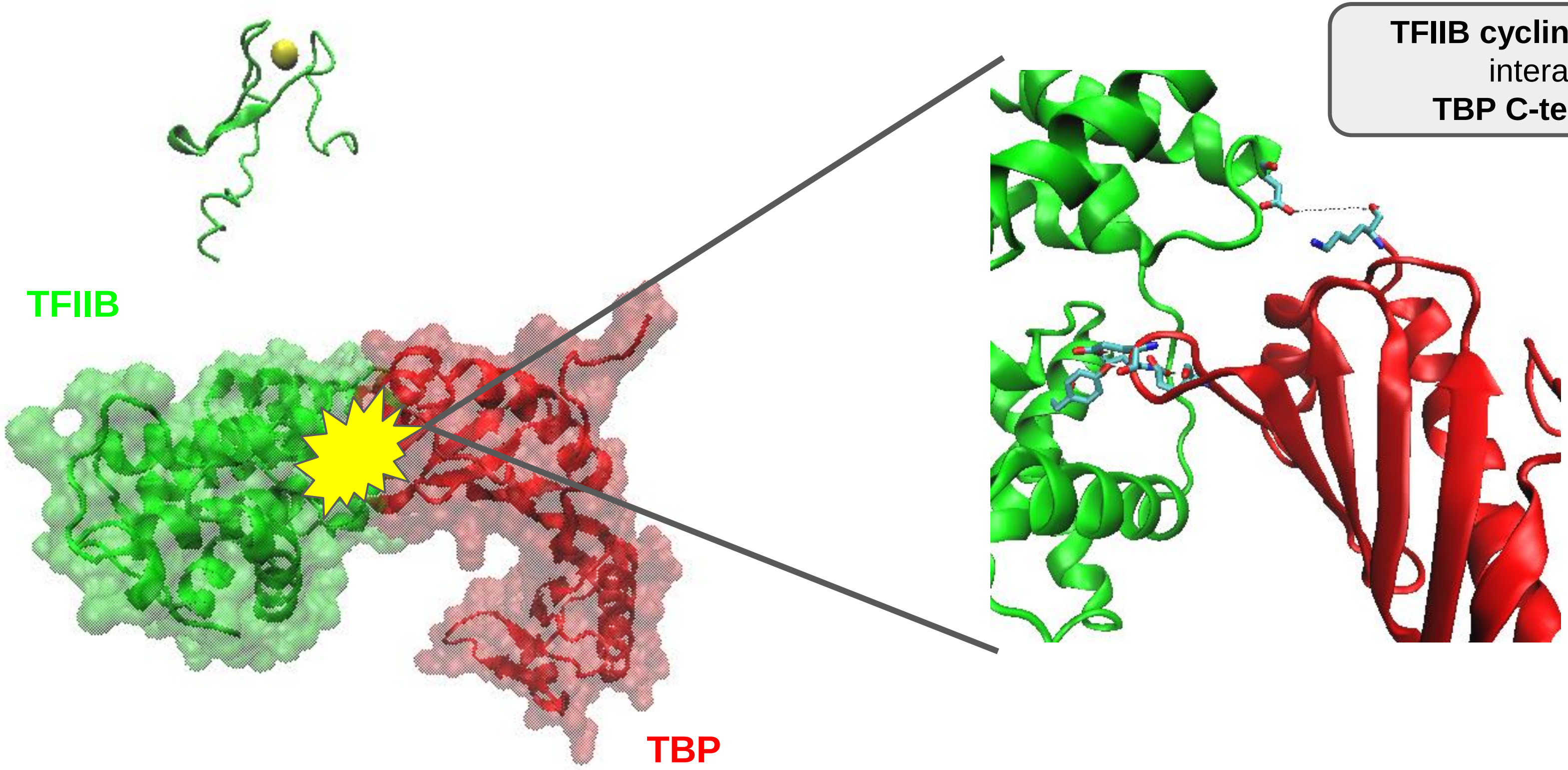
```
TF2B_SCHPO          I C S E C R E d   P P n L V E E F S S   G D T V C G S C G L
TF2B_YEAST          T C P E C K V y   P P k I V E R F S E   G D V V C A L C G L
TF2B_CAEEL          Q C P I H P D -   V H - L I E D H R A   G D L V C P A C G L
A0A0R4IRB5_DANRE    Q C P N H P D -   A L - L V E D Y R A   G D M I C P E C G L
Q28G23_XENTR        T C P N H P D -   A L - L V E D Y R A   G D M I C S E C G L
F1NXP2_CHICK        T C P N H P D -   S I - L V E D Y R A   G D M I C S E C G L
TF2B_MOUSE          T C P N H P D -   A I - L V E D Y R A   G D M I C P E C G L
TF2B_BOVIN          T C P N H P D -   A I - L V E D Y R A   G D M I C P E C G L
TF2B_HUMAN          T C P N H P D -   A I - L V E D Y R A   G D M I C P E C G L
```

The motif is conserved across eukaryotes

Upstream Promoter Complex formation: TFIIB joins the complex

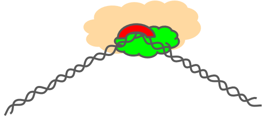


TBP recruits TFIIB to the complex

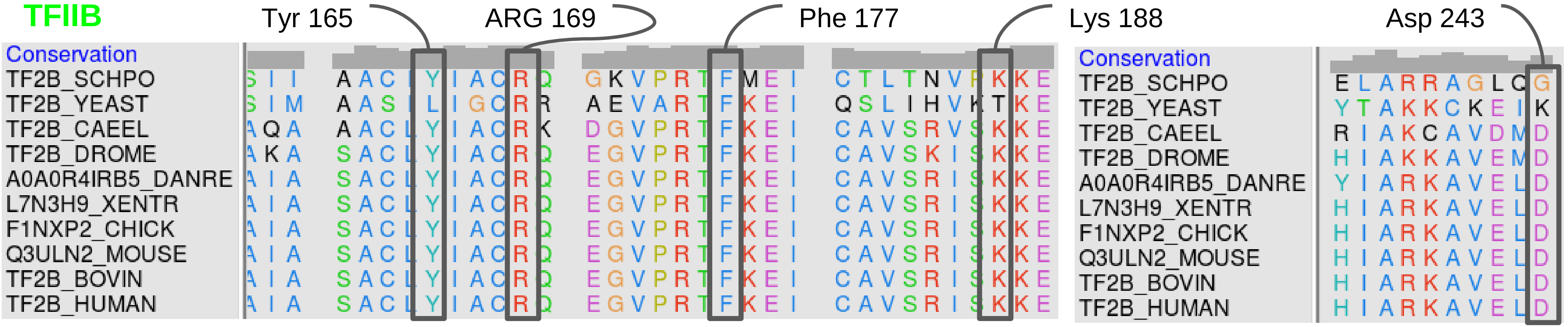


TFIIB cyclin like domains
interacts with
TBP C-term domain

Upstream Promoter Complex formation: TFIIB joins the complex

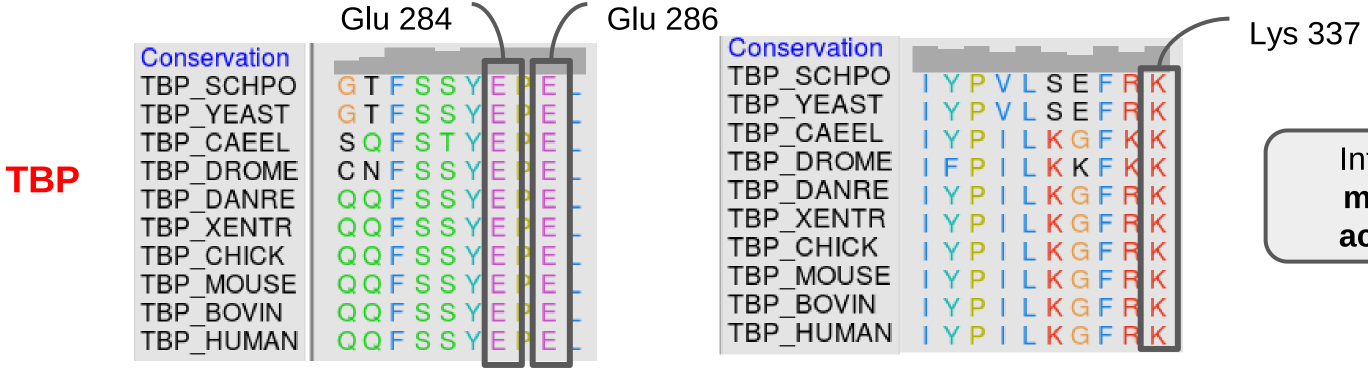


TFIIB C-term core interacts with TBP C-term domain



```
Conservation
TF2B_SCHPO          S I I   A A C I Y I A C R Q   G K V P R T F M E I   C T L T N V P K K E
TF2B_YEAST          S I M   A A S I L I G C R R   A E V A R T F K E I   Q S L I H V K T K E
TF2B_CAEEL          A Q A   A A C L Y I A C R K   D G V P R T F K E I   C A V S R V S K K E
TF2B_DROME          A K A   S A C L Y I A C R Q   E G V P R T F K E I   C A V S K I S K K E
A0A0R4IRB5_DANRE    A I A   S A C L Y I A C R Q   E G V P R T F K E I   C A V S R I S K K E
L7N3H9_XENTR        A I A   S A C L Y I A C R Q   E G V P R T F K E I   C A V S R I S K K E
F1NXP2_CHICK        A I A   S A C L Y I A C R Q   E G V P R T F K E I   C A V S R I S K K E
Q3ULN2_MOUSE        A I A   S A C L Y I A C R Q   E G V P R T F K E I   C A V S R I S K K E
TF2B_BOVIN          A I A   S A C L Y I A C R Q   E G V P R T F K E I   C A V S R I S K K E
TF2B_HUMAN          A I A   S A C L Y I A C R Q   E G V P R T F K E I   C A V S R I S K K E
```

```
Conservation
TF2B_SCHPO          E L A R R A G L Q G
TF2B_YEAST          Y T A K K C K E I K
TF2B_CAEEL          R I A K C A V D M D
TF2B_DROME          H I A K K A V E M D
A0A0R4IRB5_DANRE    Y I A R K A V E L D
L7N3H9_XENTR        H I A R K A V E L D
F1NXP2_CHICK        H I A R K A V E L D
Q3ULN2_MOUSE        H I A R K A V E L D
TF2B_BOVIN          H I A R K A V E L D
TF2B_HUMAN          H I A R K A V E L D
```

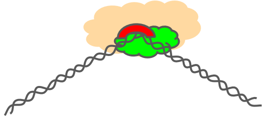


```
Conservation
TBP_SCHPO    G T F S S Y E P E
TBP_YEAST    G T F S S Y E P E
TBP_CAEEL    S Q F S T Y E P E
TBP_DROME    C N F S S Y E P E
TBP_DANRE    Q Q F S S Y E P E
TBP_XENTR    Q Q F S S Y E P E
TBP_CHICK    Q Q F S S Y E P E
TBP_MOUSE    Q Q F S S Y E P E
TBP_BOVIN    Q Q F S S Y E P E
TBP_HUMAN    Q Q F S S Y E P E
```

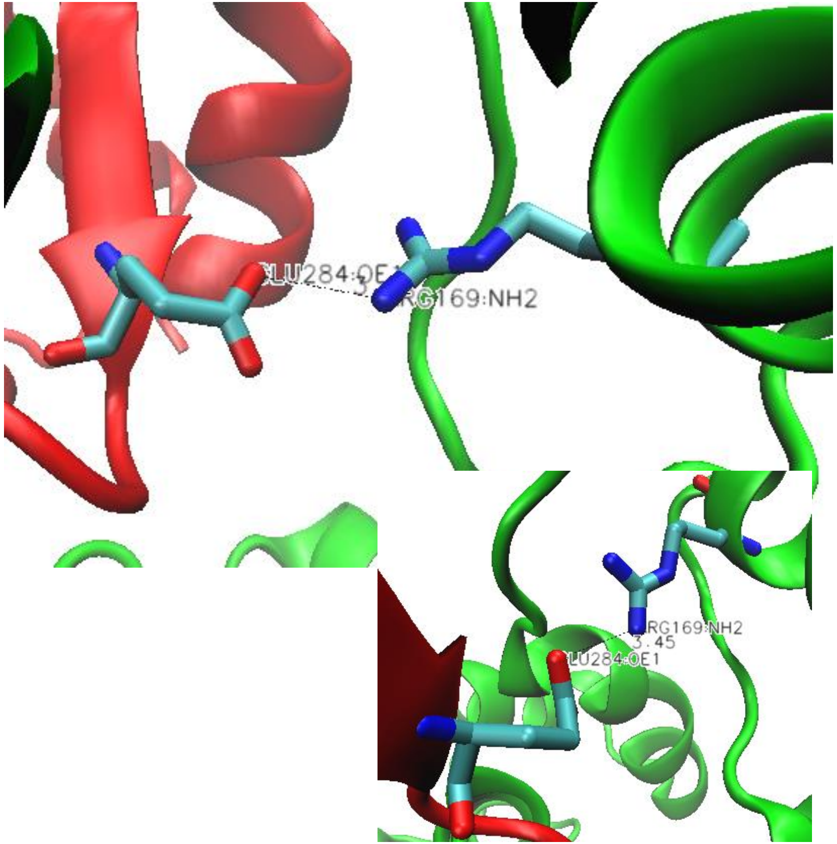
```
Conservation
TBP_SCHPO    I Y P V L S E F R K
TBP_YEAST    I Y P V L S E F R K
TBP_CAEEL    I Y P I L K G F K K
TBP_DROME    I F P I L K K F K K
TBP_DANRE    I Y P I L K G F R K
TBP_XENTR    I Y P I L K G F R K
TBP_CHICK    I Y P I L K G F R K
TBP_MOUSE    I Y P I L K G F R K
TBP_BOVIN    I Y P I L K G F R K
TBP_HUMAN    I Y P I L K G F R K
```

Interacting residues
mostly conserved
across eukaryotes

Upstream Promoter Complex formation: TFIIB joins the complex

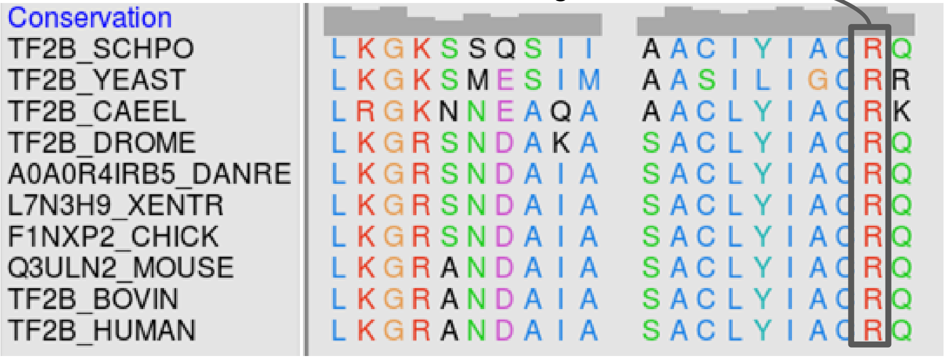


TFIIB C-term core interacts with TBP C-term domain:Electrostatic Interaction



TFIIB

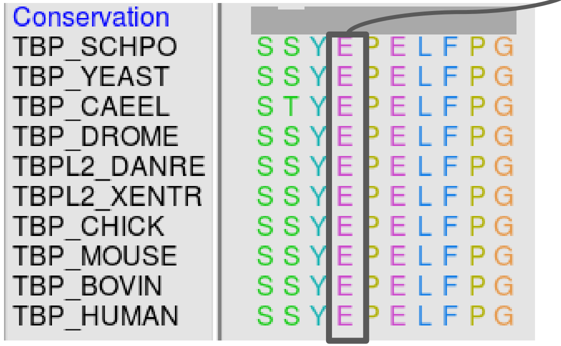
Arg 169



Conservation		
TF2B_SCHPO	L K G K S S Q S I I	A A C I Y I A C R Q
TF2B_YEAST	L K G K S M E S I M	A A S I L I G C R R
TF2B_CAEEL	L R G K N N E A Q A	A A C L Y I A C R K
TF2B_DROME	L K G R S N D A K A	S A C L Y I A C R Q
A0A0R4IRB5_DANRE	L K G R S N D A I A	S A C L Y I A C R Q
L7N3H9_XENTR	L K G R S N D A I A	S A C L Y I A C R Q
F1NXP2_CHICK	L K G R S N D A I A	S A C L Y I A C R Q
Q3ULN2_MOUSE	L K G R A N D A I A	S A C L Y I A C R Q
TF2B_BOVIN	L K G R A N D A I A	S A C L Y I A C R Q
TF2B_HUMAN	L K G R A N D A I A	S A C L Y I A C R Q

TBP

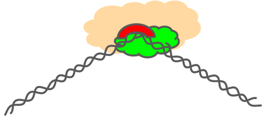
Glu 284



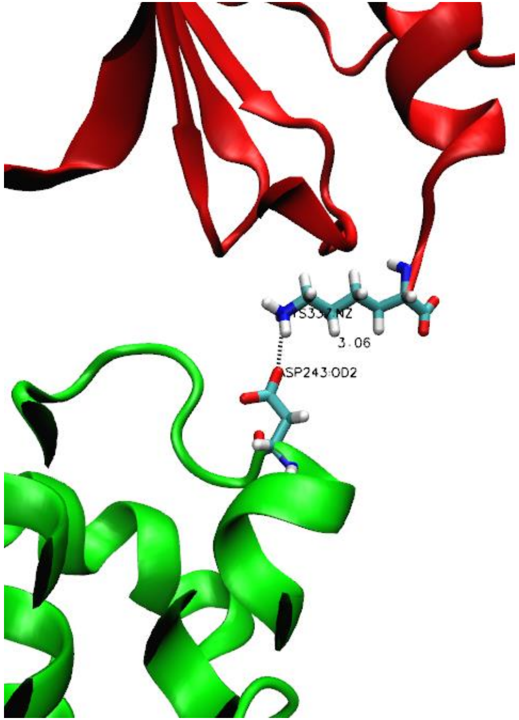
Conservation	
TBP_SCHPO	S S Y E P E L F P G
TBP_YEAST	S S Y E P E L F P G
TBP_CAEEL	S T Y E P E L F P G
TBP_DROME	S S Y E P E L F P G
TBPL2_DANRE	S S Y E P E L F P G
TBPL2_XENTR	S S Y E P E L F P G
TBP_CHICK	S S Y E P E L F P G
TBP_MOUSE	S S Y E P E L F P G
TBP_BOVIN	S S Y E P E L F P G
TBP_HUMAN	S S Y E P E L F P G

Interaction **preserved across eukaryotes**

Upstream Promoter Complex formation: TFIIB joins the complex

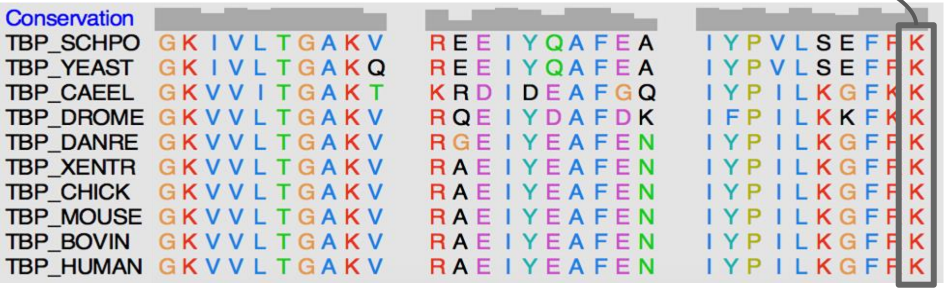


TFIIB C-term core interacts with TBP C-term domain: Salt bridge



TBP

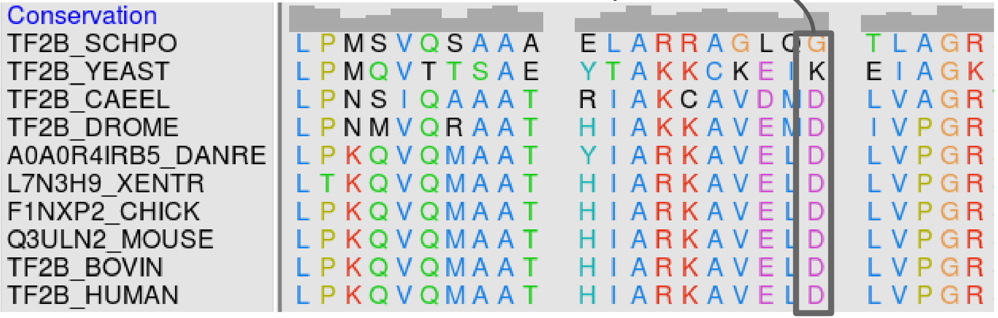
Lys 337



Conservation			
TBP_SCHPO	G K I V L T G A K V	R E E I Y Q A F E A	I Y P V L S E F R K
TBP_YEAST	G K I V L T G A K Q	R E E I Y Q A F E A	I Y P V L S E F R K
TBP_CAEEL	G K V V I T G A K T	K R D I D E A F G Q	I Y P I L K G F K K
TBP_DROME	G K V V L T G A K V	R Q E I Y D A F D K	I F P I L K K F K K
TBP_DANRE	G K V V L T G A K V	R G E I Y E A F E N	I Y P I L K G F R K
TBP_XENTR	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K
TBP_CHICK	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K
TBP_MOUSE	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K
TBP_BOVIN	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K
TBP_HUMAN	G K V V L T G A K V	R A E I Y E A F E N	I Y P I L K G F R K

TFIIB

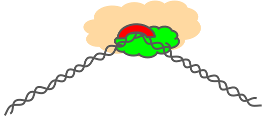
Asp 243



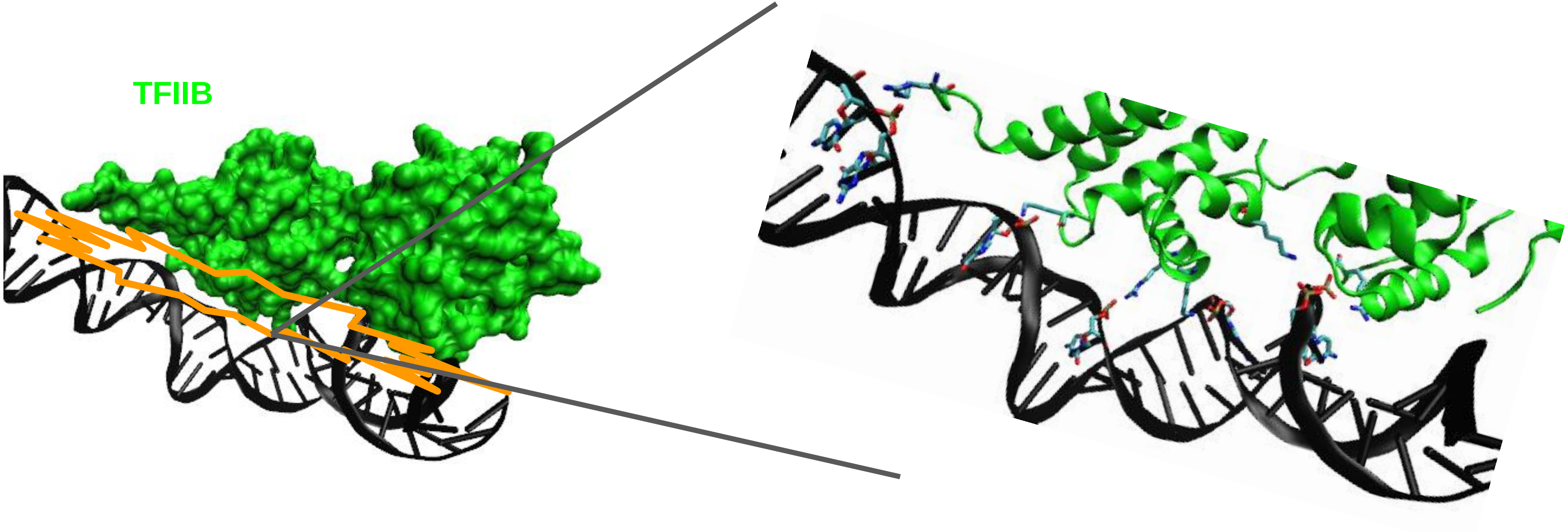
Conservation			
TF2B_SCHPO	L P M S V Q S A A A	E L A R R A G L Q G	T L A G R
TF2B_YEAST	L P M Q V T T S A E	Y T A K K C K E I K	E I A G K
TF2B_CAEEL	L P N S I Q A A A T	R I A K C A V D M D	L V A G R
TF2B_DROME	L P N M V Q R A A T	H I A K K A V E M D	I V P G R
A0A0R4IRB5_DANRE	L P K Q V Q M A A T	Y I A R K A V E L D	L V P G R
L7N3H9_XENTR	L T K Q V Q M A A T	H I A R K A V E L D	L V P G R
F1NXP2_CHICK	L P K Q V Q M A A T	H I A R K A V E L D	L V P G R
Q3ULN2_MOUSE	L P K Q V Q M A A T	H I A R K A V E L D	L V P G R
TF2B_BOVIN	L P K Q V Q M A A T	H I A R K A V E L D	L V P G R
TF2B_HUMAN	L P K Q V Q M A A T	H I A R K A V E L D	L V P G R

Interaction **preserved across metazoans**

Upstream Promoter Complex formation: TFIIB joins the complex

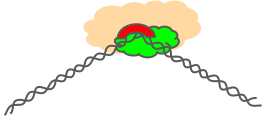


TFIIB C-term core interacts with double-strand DNA



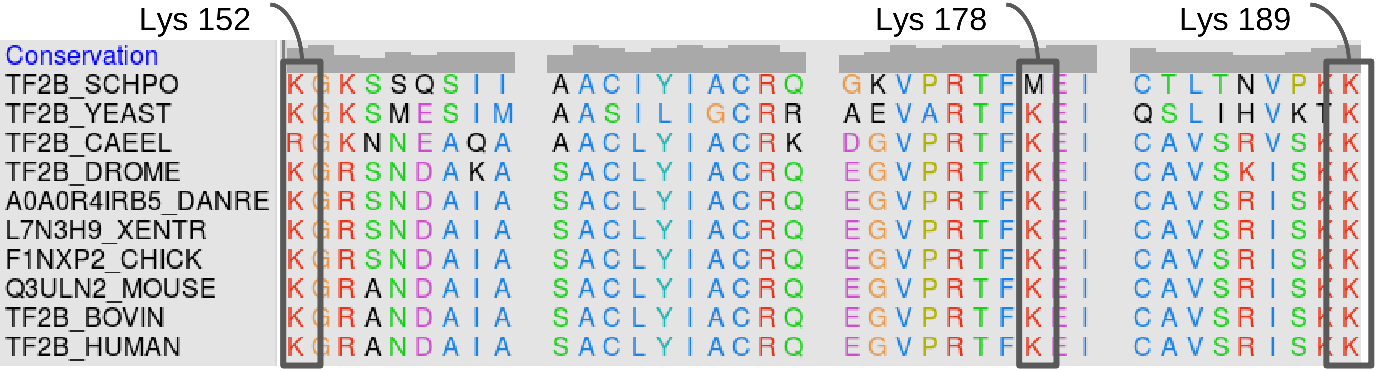
Most of the interacting
residues are Lys or Arg

Upstream Promoter Complex formation: TFIIB joins the complex



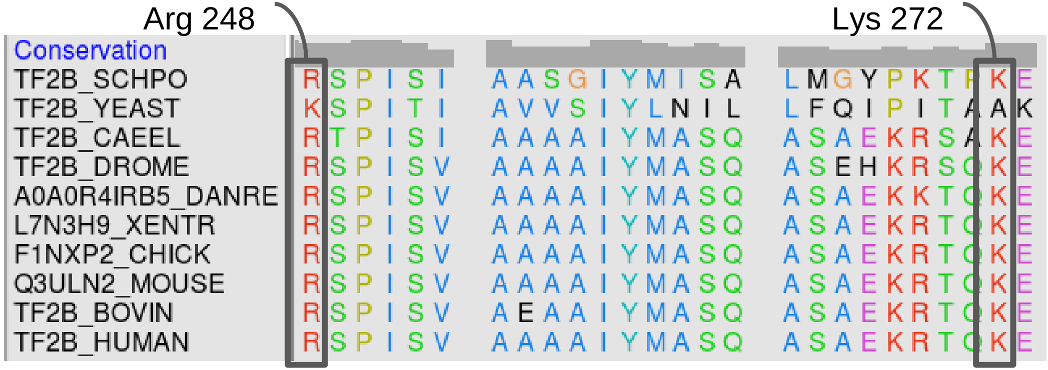
TFIIB C-term core interacts with double-strand DNA

TFIIB



Conservation				
TF2B_SCHPO	K G K S S Q S I I	A A C I Y I A C R Q	G K V P R T F M E I	C T L T N V P K K
TF2B_YEAST	K G K S M E S I M	A A S I L I G C R R	A E V A R T F K E I	Q S L I H V K T K
TF2B_CAEEL	R G K N N E A Q A	A A C L Y I A C R K	D G V P R T F K E I	C A V S R V S K K
TF2B_DROME	K G R S N D A K A	S A C L Y I A C R Q	E G V P R T F K E I	C A V S K I S K K
A0A0R4IRB5_DANRE	K G R S N D A I A	S A C L Y I A C R Q	E G V P R T F K E I	C A V S R I S K K
L7N3H9_XENTR	K G R S N D A I A	S A C L Y I A C R Q	E G V P R T F K E I	C A V S R I S K K
F1NXP2_CHICK	K G R S N D A I A	S A C L Y I A C R Q	E G V P R T F K E I	C A V S R I S K K
Q3ULN2_MOUSE	K G R A N D A I A	S A C L Y I A C R Q	E G V P R T F K E I	C A V S R I S K K
TF2B_BOVIN	K G R A N D A I A	S A C L Y I A C R Q	E G V P R T F K E I	C A V S R I S K K
TF2B_HUMAN	K G R A N D A I A	S A C L Y I A C R Q	E G V P R T F K E I	C A V S R I S K K

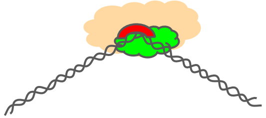
TFIIB



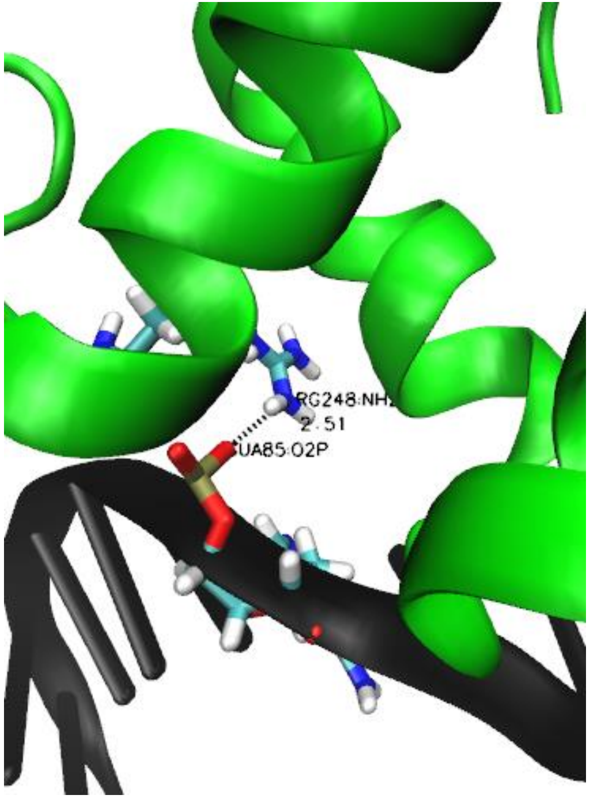
Conservation			
TF2B_SCHPO	R S P I S I	A A S G I Y M I S A	L M G Y P K T P K E
TF2B_YEAST	K S P I T I	A V V S I Y L N I L	L F Q I P I T A A K
TF2B_CAEEL	R T P I S I	A A A A I Y M A S Q	A S A E K R S A K E
TF2B_DROME	R S P I S V	A A A A I Y M A S Q	A S E H K R S Q K E
A0A0R4IRB5_DANRE	R S P I S V	A A A A I Y M A S Q	A S A E K K T Q K E
L7N3H9_XENTR	R S P I S V	A A A A I Y M A S Q	A S A E K R T Q K E
F1NXP2_CHICK	R S P I S V	A A A A I Y M A S Q	A S A E K R T Q K E
Q3ULN2_MOUSE	R S P I S V	A A A A I Y M A S Q	A S A E K R T Q K E
TF2B_BOVIN	R S P I S V	A E A A I Y M A S Q	A S A E K R T Q K E
TF2B_HUMAN	R S P I S V	A A A A I Y M A S Q	A S A E K R T Q K E

Residues **mostly conserved across eukaryotes**

Upstream Promoter Complex formation: TFIIB joins the complex



TFIIB C-term core interacts with double stranded DNA: Salt Bridge



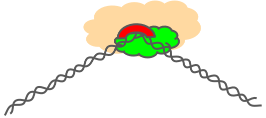
TFIIB

Arg 248

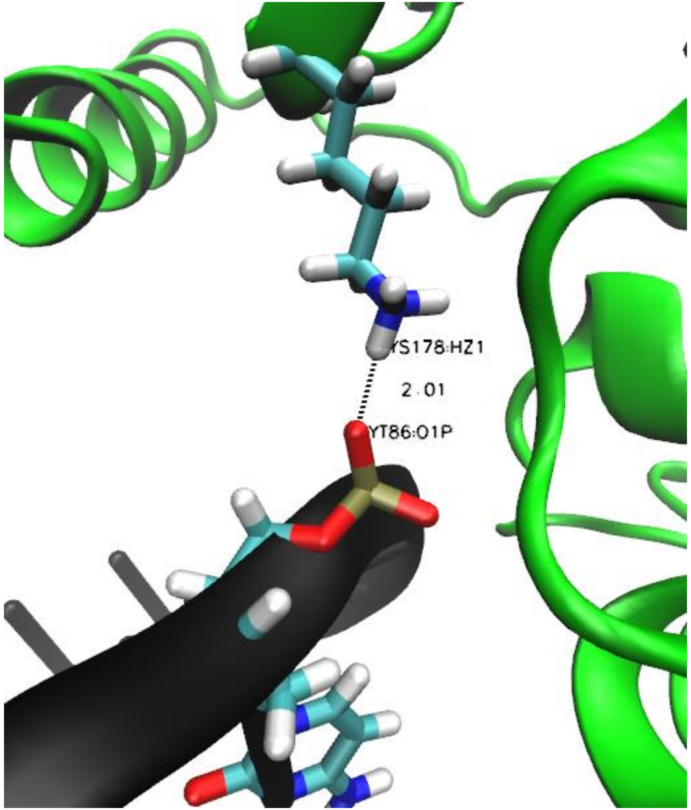
Conservation		
TF2B_SCHPO	E L A R R A G L Q G	T L A G R S P I S I
TF2B_YEAST	Y T A K K C K E I K	E I A G K S P I T I
TF2B_CAEEL	R I A K C A V D M D	L V A G R T P I S I
TF2B_DROME	H I A K K A V E M D	I V P G R S P I S V
A0A0R4IRB5_DANRE	Y I A R K A V E L D	L V P G R S P I S V
L7N3H9_XENTR	H I A R K A V E L D	L V P G R S P I S V
F1NXP2_CHICK	H I A R K A V E L D	L V P G R S P I S V
Q3ULN2_MOUSE	H I A R K A V E L D	L V P G R S P I S V
TF2B_BOVIN	H I A R K A V E L D	L V P G R S P I S V
TF2B_HUMAN	H I A R K A V E L D	L V P G R S P I S V

Arg 248 mostly preserved across eukaryotes

Upstream Promoter Complex formation: TFIIB joins the complex



TFIIB C-term core interacts with DNA: Salt bridge



TFIIB

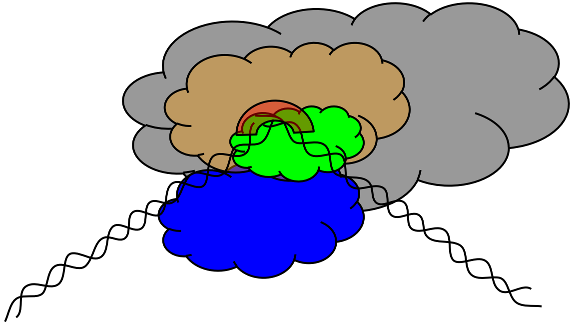
Lys 178

Conservation			
TF2B_SCHPO	S I I	A A C I Y I A C R Q	G K V P R T F M E I
TF2B_YEAST	S I M	A A S I L I G C R R	A E V A R T F K E I
TF2B_CAEEL	A Q A	A A C L Y I A C R K	D G V P R T F K E I
TF2B_DROME	A K A	S A C L Y I A C R Q	E G V P R T F K E I
A0A0R4IRB5_DANRE	A I A	S A C L Y I A C R Q	E G V P R T F K E I
L7N3H9_XENTR	A I A	S A C L Y I A C R Q	E G V P R T F K E I
F1NXP2_CHICK	A I A	S A C L Y I A C R Q	E G V P R T F K E I
Q3ULN2_MOUSE	A I A	S A C L Y I A C R Q	E G V P R T F K E I
TF2B_BOVIN	A I A	S A C L Y I A C R Q	E G V P R T F K E I
TF2B_HUMAN	A I A	S A C L Y I A C R Q	E G V P R T F K E I

Lys 178 mostly preserved across **metazoans**

Core PIC formation:

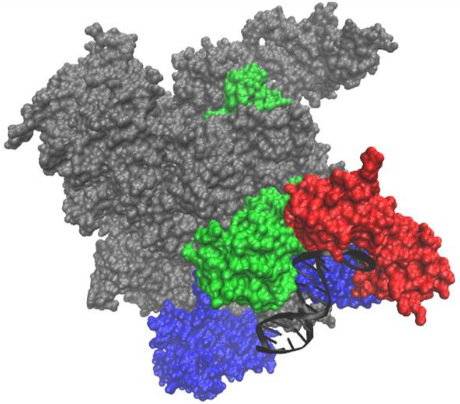
TFIIF-Pol II join the complex



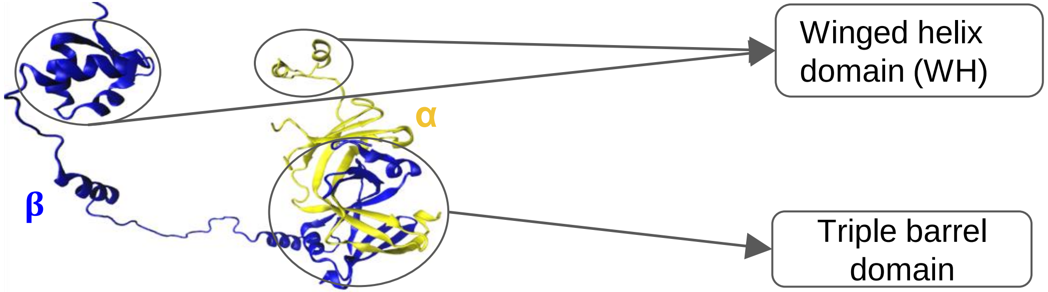
Core PIC formation: TFIIF-Pol II join the complex

Pol II contains 12 subunits and TFIIF has an α and a **β** subunit with several domains

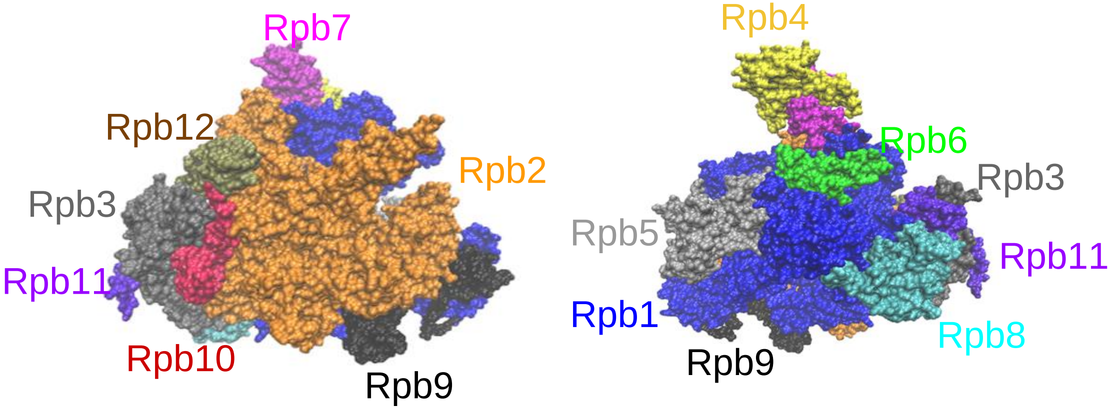
Joining of RNA pol II and TFIIF



TFIIF subunits and domains:



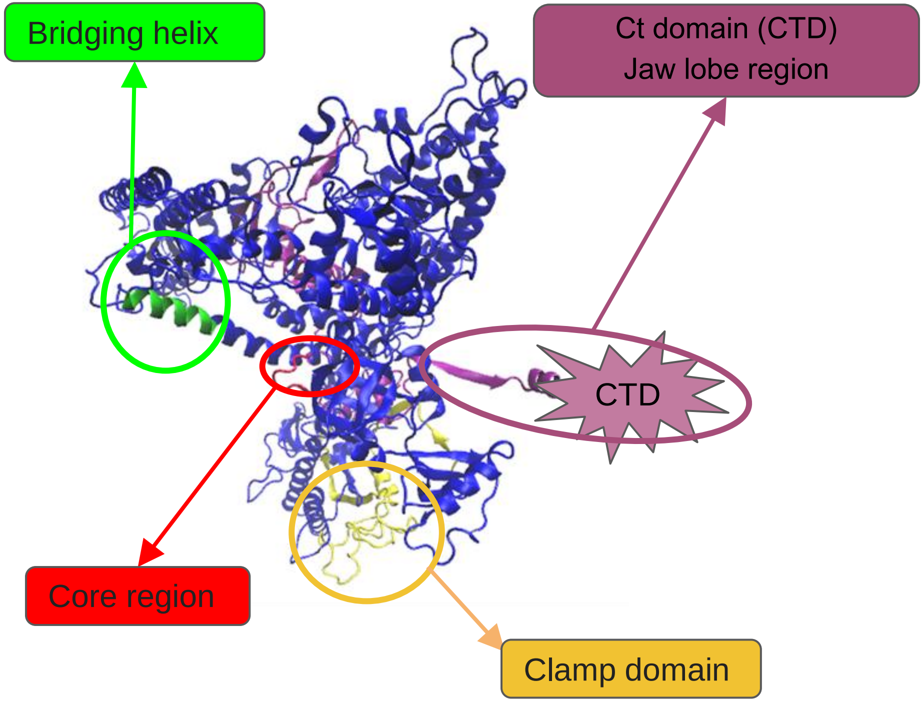
RNA pol II subunits:



Core PIC formation: TFIIF-Pol II join the complex

Rpb1 and Rpb2 subunits are the ones forming and stabilizing the core or the active center

RPB1 subunit (1970 Aa)



Bridging helix

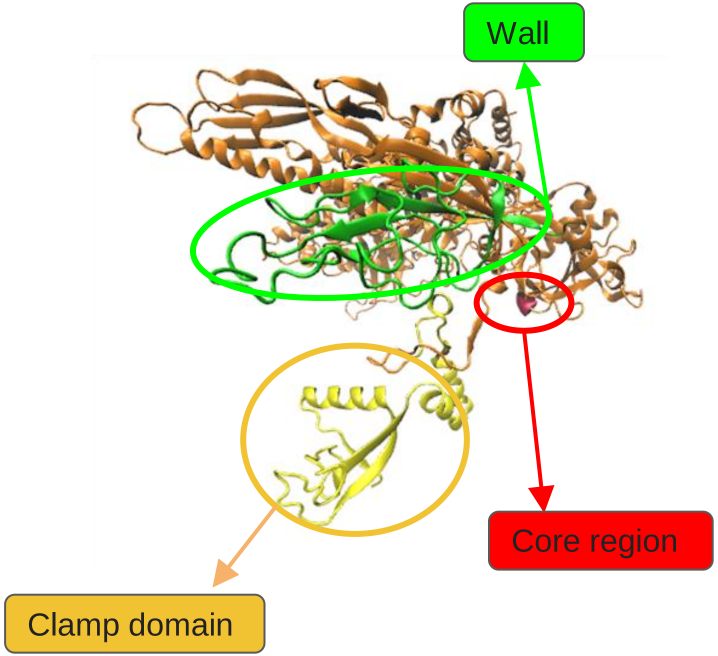
Ct domain (CTD)
Jaw lobe region

CTD

Core region

Clamp domain

RPB2 subunit (1174 Aa)



Wall

Core region

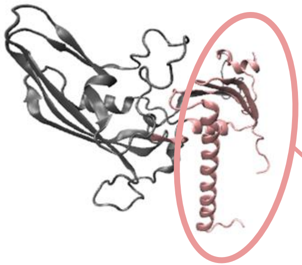
Clamp domain

Core PIC formation: TFIIF-Pol II join the complex

There are other subunits limiting the core and others acting as a wedge-lock or a blocker

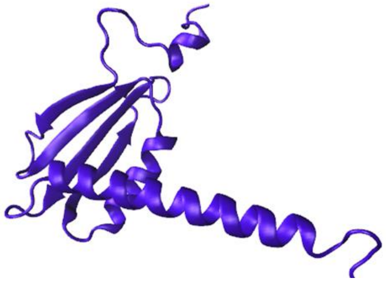
Rpb3 and Rpb11 mark the **limits of the core**

RPB3 subunit (275 Aa)



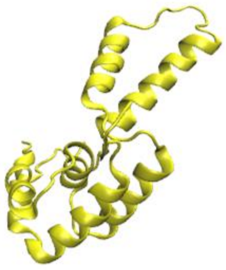
Core region

RPB11 subunit (117 Aa)

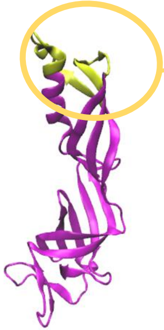


Rpb4 and Rpb7 **maintain the closing state**

RPB4 subunit (142 Aa)



RPB7 subunit (172 Aa)



Residues interacting with Rpb6 subunit

Core PIC formation: TFIIF-Pol II join the complex

There are subunits forming the Pol II structural elements and other have a role in its architecture

Forms part of the **Clamp**

RPB6 subunit (127 Aa)

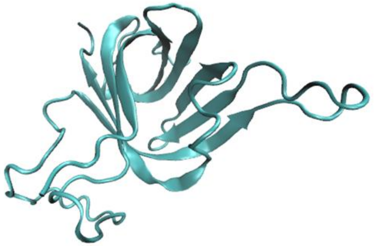
Form part of the **Jaw lobe**, changing its conformation

RPB5 subunit (210 Aa)

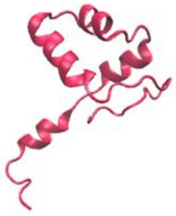
RPB9 subunit (125 Aa)

Have a role **stabilizing** the Pol II complex

RPB8 subunit (150 Aa)



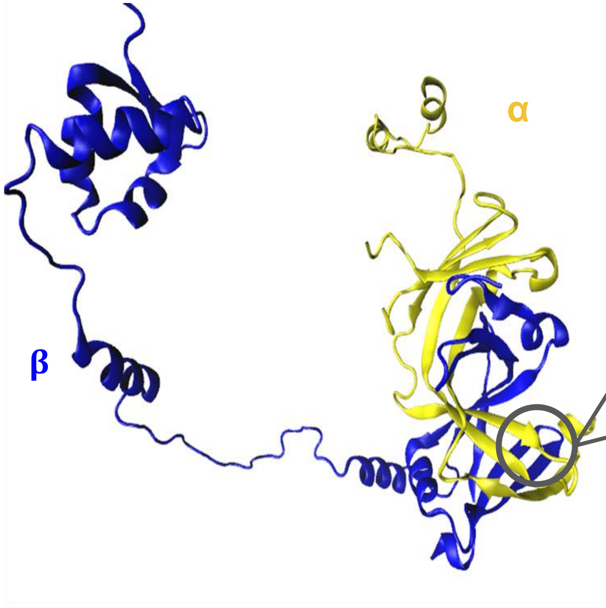
RPB10 subunit (67 Aa)

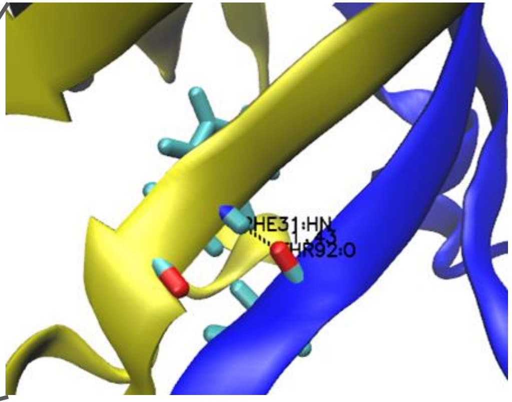


RPB12 subunit (58 Aa)

Core PIC formation: TFIIF-Pol II join the complex

TFIIF subunits interact through the triple barrel domain

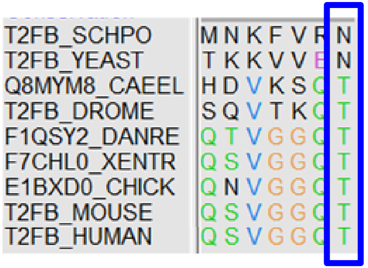




Mostly hydrogen bonds maintain the Triple barrel domain

TFIIFβ

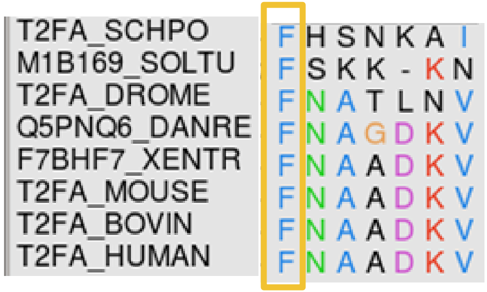
Thr92



T2FB_SCHPO	M	N	K	F	V	R	N
T2FB_YEAST	T	K	K	V	V	E	N
Q8MYM8_CAEEL	H	D	V	K	S	Q	T
T2FB_DROME	S	Q	V	T	K	Q	T
F1QSY2_DANRE	Q	T	V	G	G	Q	T
F7CHL0_XENTR	Q	S	V	G	G	Q	T
E1BXD0_CHICK	Q	N	V	G	G	Q	T
T2FB_MOUSE	Q	S	V	G	G	Q	T
T2FB_HUMAN	Q	S	V	G	G	Q	T

TFIIFα

Phe31

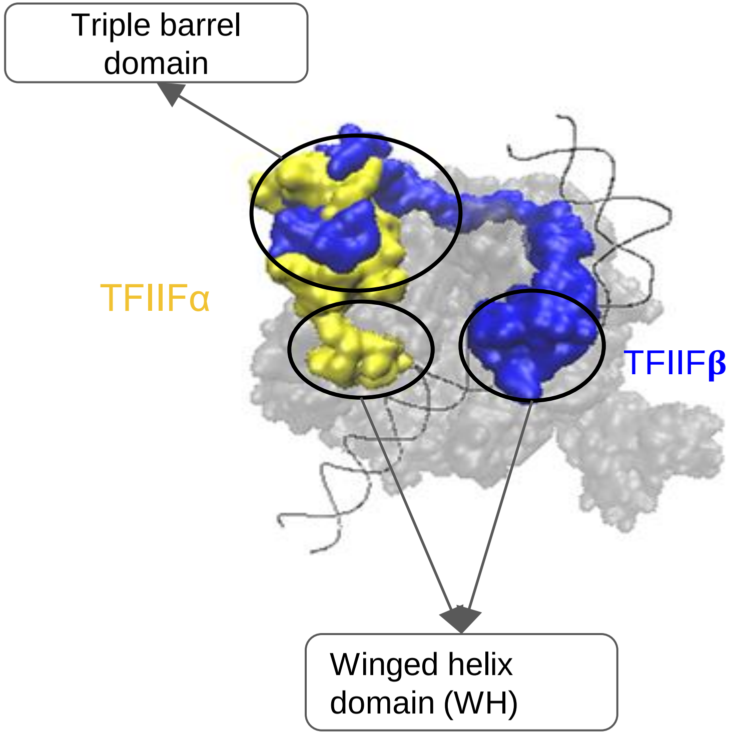


T2FA_SCHPO	F	H	S	N	K	A	I
M1B169_SOLTU	F	S	K	K	-	K	N
T2FA_DROME	F	N	A	T	L	N	V
Q5PNQ6_DANRE	F	N	A	G	D	K	V
F7BHF7_XENTR	F	N	A	A	D	K	V
T2FA_MOUSE	F	N	A	A	D	K	V
T2FA_BOVIN	F	N	A	A	D	K	V
T2FA_HUMAN	F	N	A	A	D	K	V

Conserved across metazoans

Core PIC formation: TFIIF-Pol II join the complex

TFIIF interact with different subunits of Pol II and with the DNA



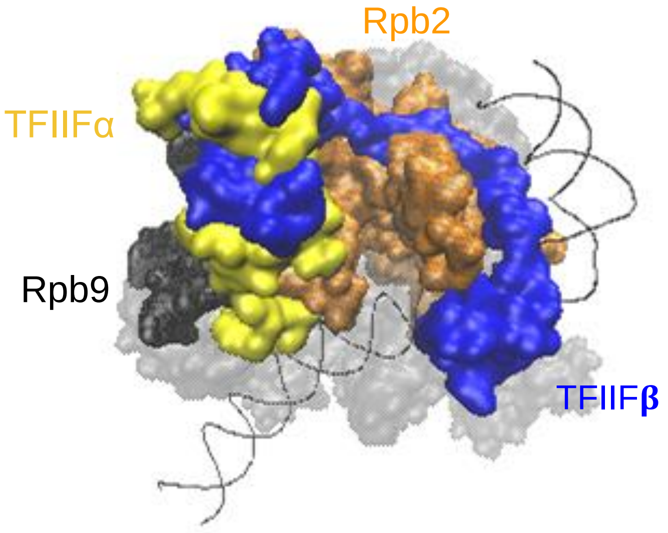
Triple barrel domain

TFIIFα

TFIIFβ

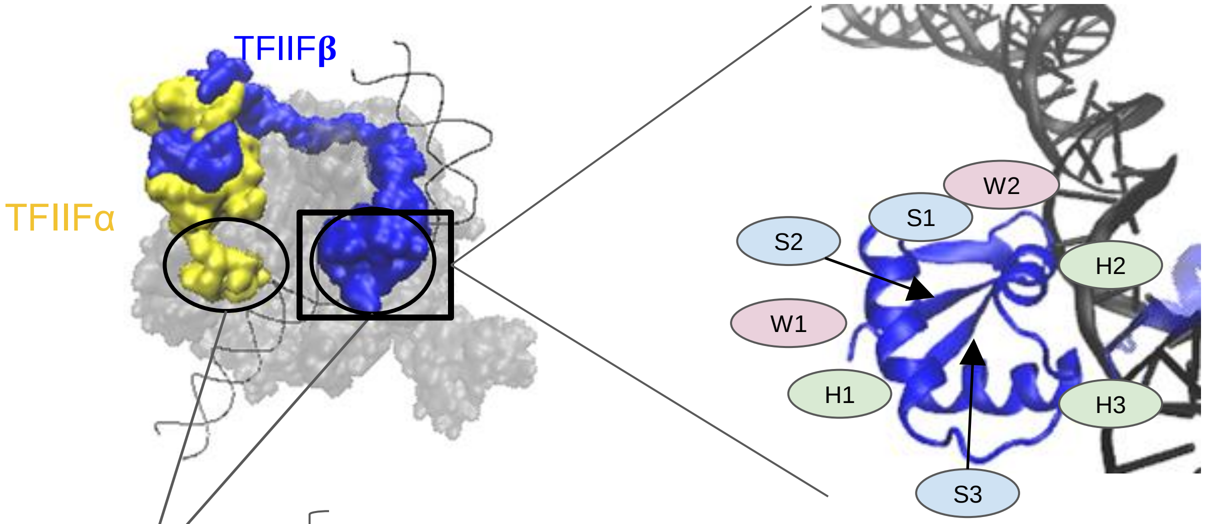
Winged helix domain (WH)

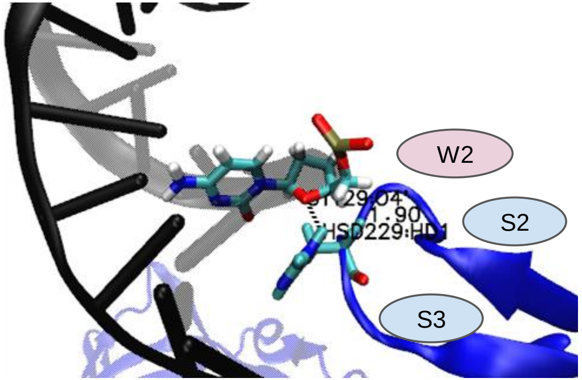
TFIIF is mainly interacting with Rpb2 subunit



Core PIC formation: TFIIF-Pol II join the complex

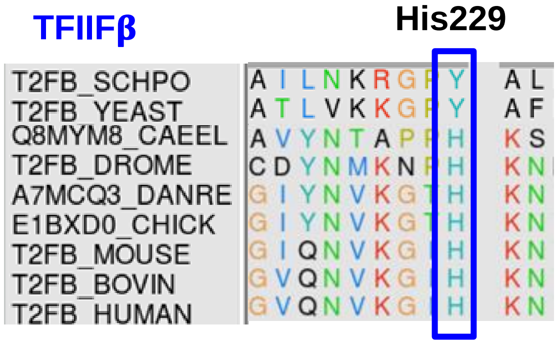
TFIIF has two winged helix domains which are important for the interaction with the DNA





Winged helix domain (WH)

- three alpha helices → Specific DNA contact
- two wings → Non-specific DNA contact
- three beta-sheets



TFIIFβ

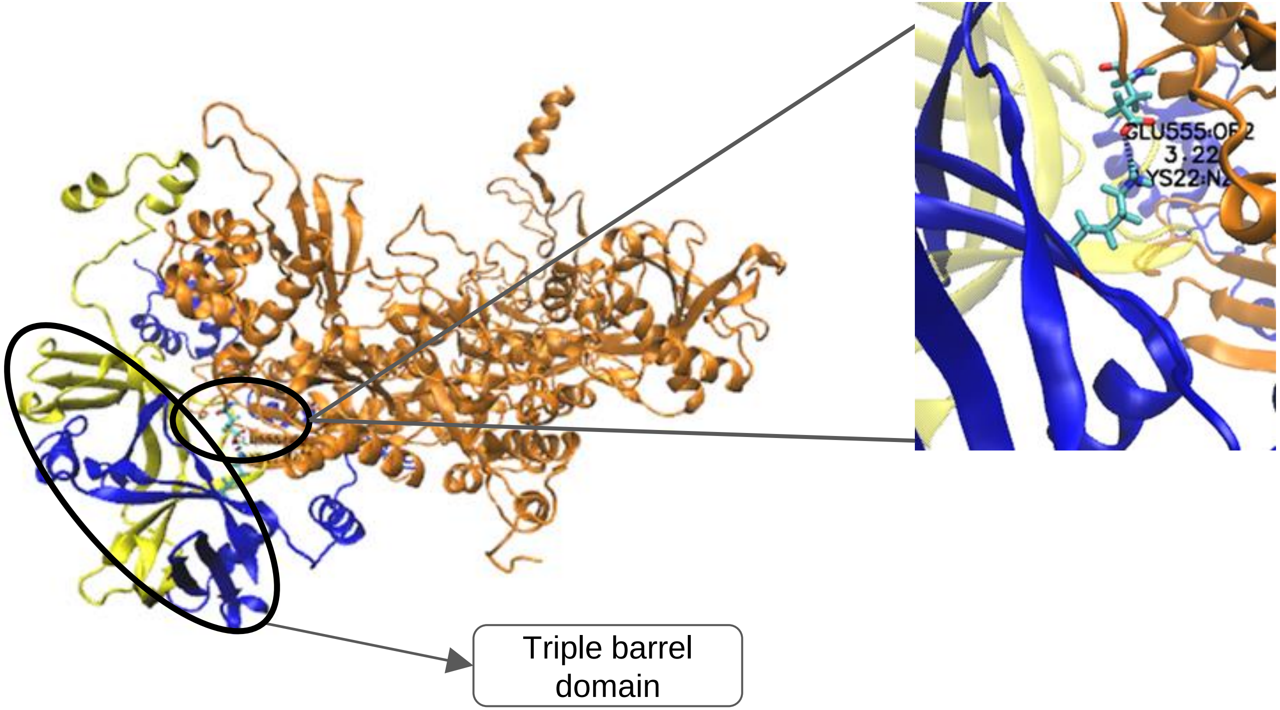
His229

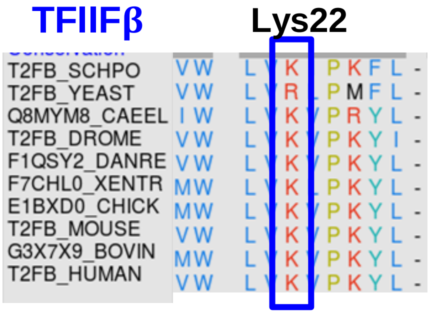
```
T2FB_SCHPO     A I L N K R G P Y   A L
T2FB_YEAST     A T L V K K G P Y   A F
Q8MYM8_CAEEL   A V Y N T A P P H   K S
T2FB_DROME     C D Y N M K N P H   K N
A7MCQ3_DANRE   G I Y N V K G T H   K N
E1BXD0_CHICK   G I Y N V K G T H   K N
T2FB_MOUSE     G I Q N V K G   H   K N
T2FB_BOVIN     G V Q N V K G   H   K N
T2FB_HUMAN     G V Q N V K G   H   K N
```

Conserved across **metazoans**

Core PIC formation: TFIIF-PolII join the complex

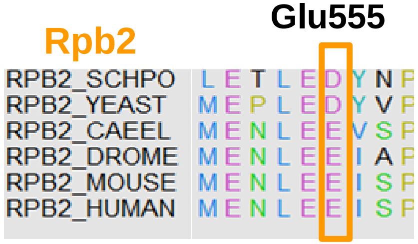
There's an important electrostatic interaction between Rpb2 and TFIIF**β** in the triple barrel domain





TFIIFβ **Lys22**

T2FB_SCHPO	V	W		L	Y	K		P	K	F	L	-
T2FB_YEAST	V	W		L	Y	R	L	P	M	F	L	-
Q8MYM8_CAEEL	I	W		L	Y	K	V	P	R	Y	L	-
T2FB_DROME	V	W		L	Y	K	V	P	K	Y	I	-
F1QSY2_DANRE	V	W		L	Y	K	V	P	K	Y	L	-
F7CHL0_XENTR	M	W		L	Y	K	L	P	K	Y	L	-
E1BXD0_CHICK	M	W		L	Y	K	V	P	K	Y	L	-
T2FB_MOUSE	V	W		L	Y	K	V	P	K	Y	L	-
G3X7X9_BOVIN	M	W		L	Y	K	V	P	K	Y	L	-
T2FB_HUMAN	V	W		L	Y	K	V	P	K	Y	L	-



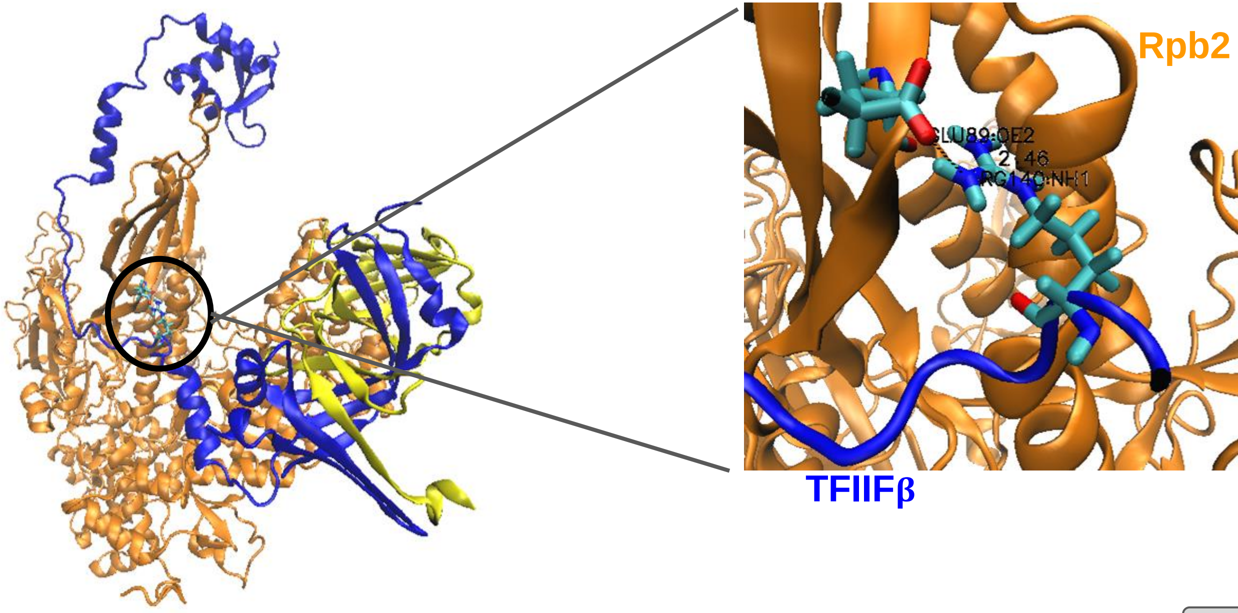
Rpb2 **Glu555**

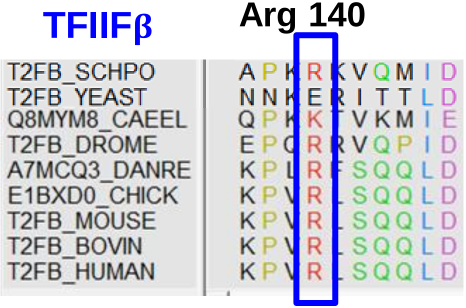
RPB2_SCHPO	L	E	T	L	E	D	Y	N	P
RPB2_YEAST	M	E	P	L	E	D	Y	V	P
RPB2_CAEEL	M	E	N	L	E	E	V	S	P
RPB2_DROME	M	E	N	L	E	E	I	A	P
RPB2_MOUSE	M	E	N	L	E	E	I	S	P
RPB2_HUMAN	M	E	N	L	E	E	I	S	P

Conserved across **eukaryotes**

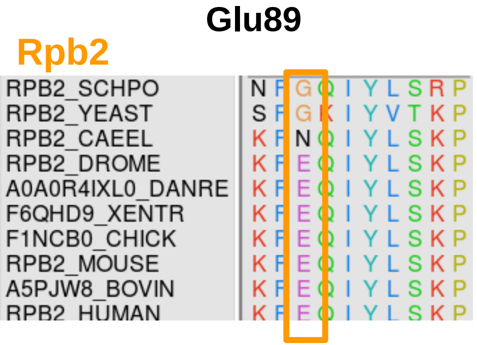
Core PIC formation: TFIIF-PolII join the complex

There is also a salt bridge interaction between Rpb2 and TFIIF**β**





T2FB_SCHPO	A	P	K	R	K	V	Q	M	I	D
T2FB_YEAST	N	N	K	E	R	I	T	T	L	D
Q8MYM8_CAEEL	Q	P	K	K	T	V	K	M	I	E
T2FB_DROME	E	P	Q	R	R	V	Q	P	I	D
A7MCQ3_DANRE	K	P	L	R	F	S	Q	Q	L	D
E1BXD0_CHICK	K	P	V	R	L	S	Q	Q	L	D
T2FB_MOUSE	K	P	V	R	L	S	Q	Q	L	D
T2FB_BOVIN	K	P	V	R	L	S	Q	Q	L	D
T2FB_HUMAN	K	P	V	R	L	S	Q	Q	L	D



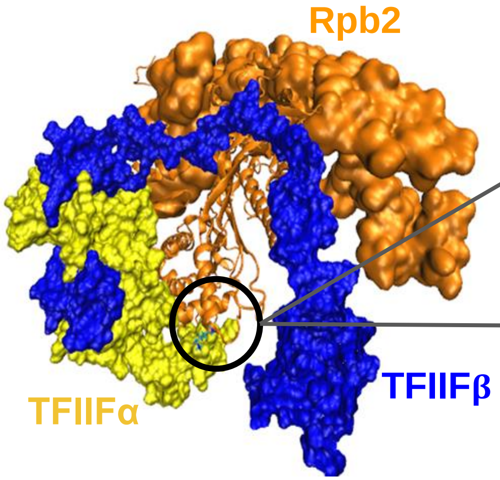
RPB2_SCHPO	N	F	G	Q	I	Y	L	S	R	P
RPB2_YEAST	S	F	G	K	I	Y	V	T	K	P
RPB2_CAEEL	K	F	N	Q	I	Y	L	S	K	P
RPB2_DROME	K	F	E	Q	I	Y	L	S	K	P
A0A0R4IXL0_DANRE	K	F	E	Q	I	Y	L	S	K	P
F6QHD9_XENTR	K	F	E	Q	I	Y	L	S	K	P
F1NCB0_CHICK	K	F	E	Q	I	Y	L	S	K	P
RPB2_MOUSE	K	F	E	Q	I	Y	L	S	K	P
A5PJW8_BOVIN	K	F	E	Q	I	Y	L	S	K	P
RPB2_HUMAN	K	F	E	Q	I	Y	L	S	K	P

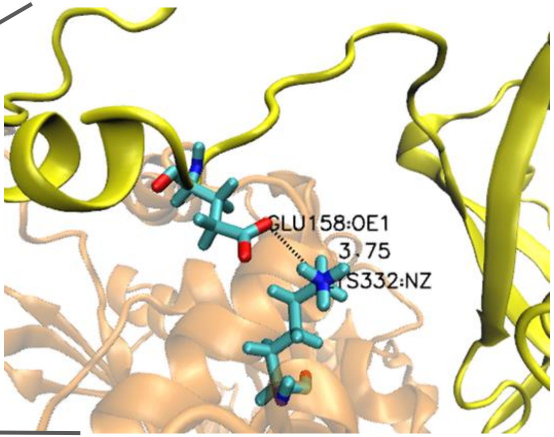
Residues conserved across **chordata**

Core PIC formation: TFIIF-PolII join the complex

An electrostatic interaction between Rpb2 and TFIIFα is required

TFIIFα is interacting lower than TFIIF**β** with Rpb2





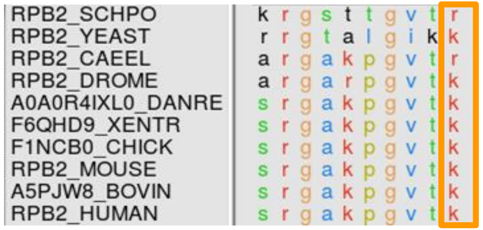
TFIIFα

Glu158

T2FA_SCHPO	D	E	A	E	A	K	M	N
B6JUZ6_YEAST	D	E	A	E	A	K	M	N
T2FA_DROME	E	E	A	E	Q	E	F	G
Q5PNQ6_DANRE	E	E	A	E	E	E	W	G
F7BHF7_XENTR	E	E	A	E	Q	E	W	E
T2FA_MOUSE	E	E	A	E	E	E	W	E
T2FA_BOVIN	E	E	A	E	E	E	W	E
T2FA_HUMAN	E	E	A	E	E	E	W	E

Rpb2

Lys332

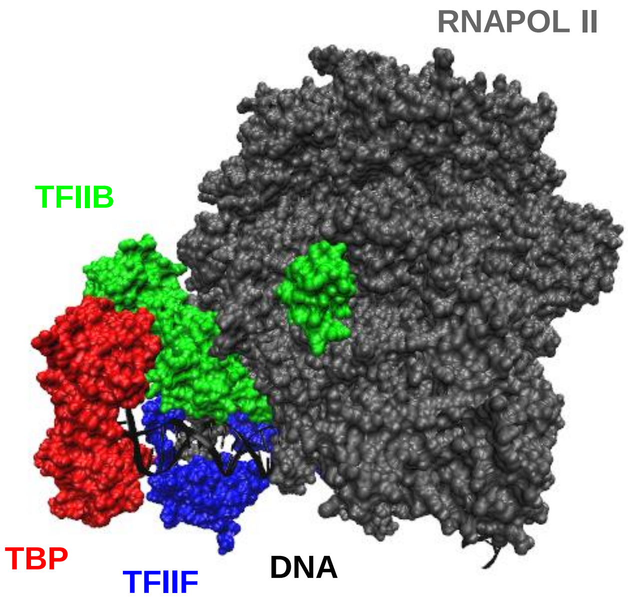


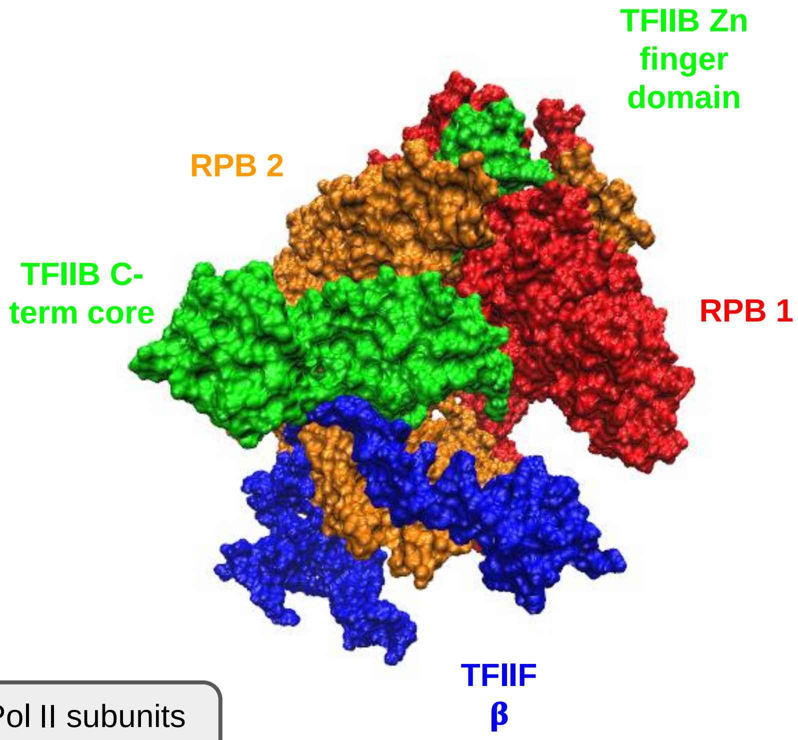
RPB2_SCHPO	k	r	g	s	t	t	g	v	t	r
RPB2_YEAST	r	r	g	t	a	l	g	i	k	k
RPB2_CAEEL	a	r	g	a	k	p	g	v	t	r
RPB2_DROME	a	r	g	a	r	p	g	v	t	k
A0A0R4IXL0_DANRE	s	r	g	a	k	p	g	v	t	k
F6QHD9_XENTR	s	r	g	a	k	p	g	v	t	k
F1NCB0_CHICK	s	r	g	a	k	p	g	v	t	k
RPB2_MOUSE	s	r	g	a	k	p	g	v	t	k
A5PJW8_BOVIN	s	r	g	a	k	p	g	v	t	k
RPB2_HUMAN	s	r	g	a	k	p	g	v	t	k

Residues conserved across **eukaryotes**

Core PIC formation: TFIIF-PolII join the complex

TFIIB recruits the TFIIF-PolII assembly to the PIC

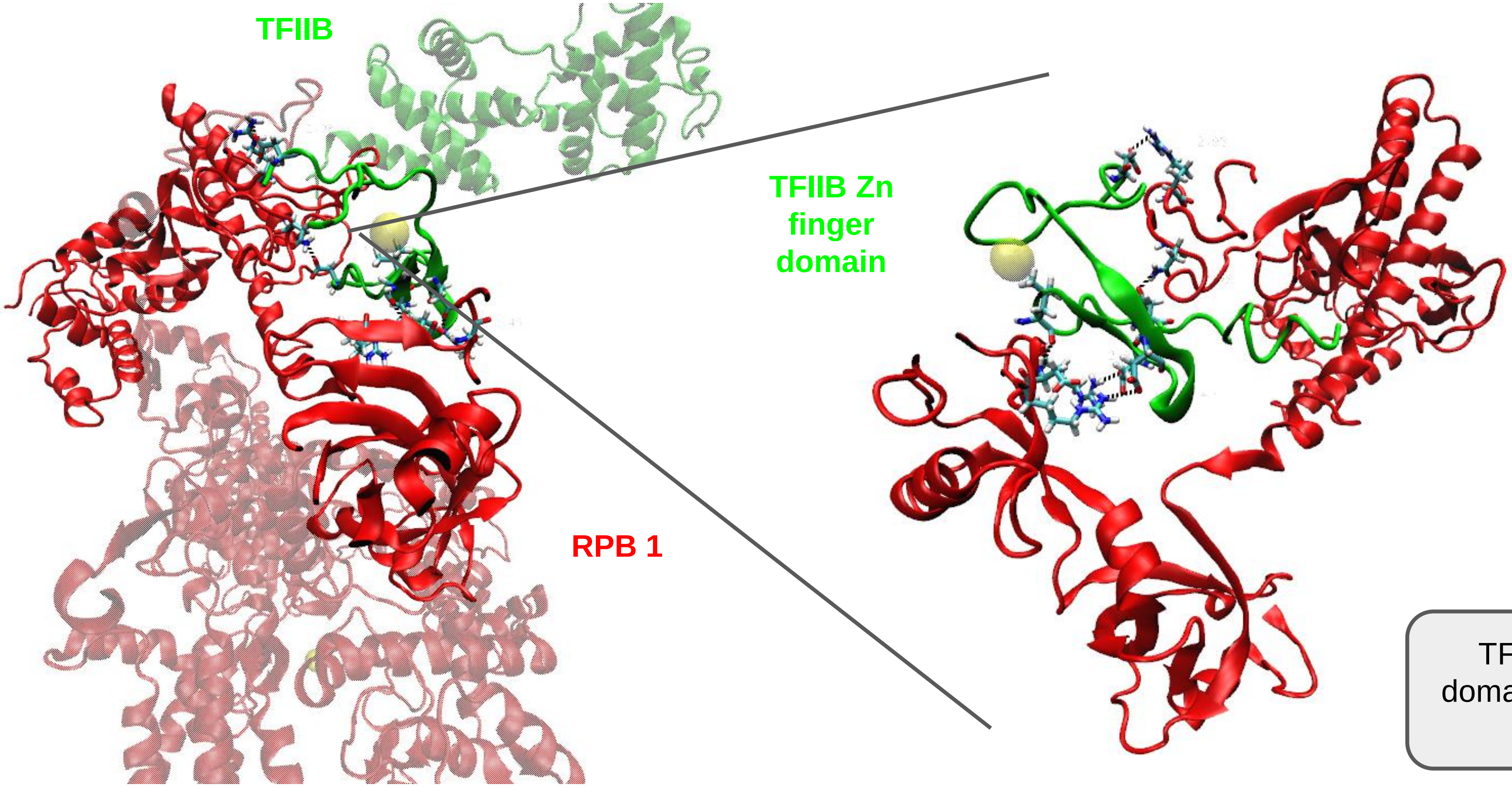




RPB 1 and **RPB 2** are the RNA Pol II subunits involved in the interaction **with TFIIF-β**

Core PIC formation: TFIIF-PolII join the complex

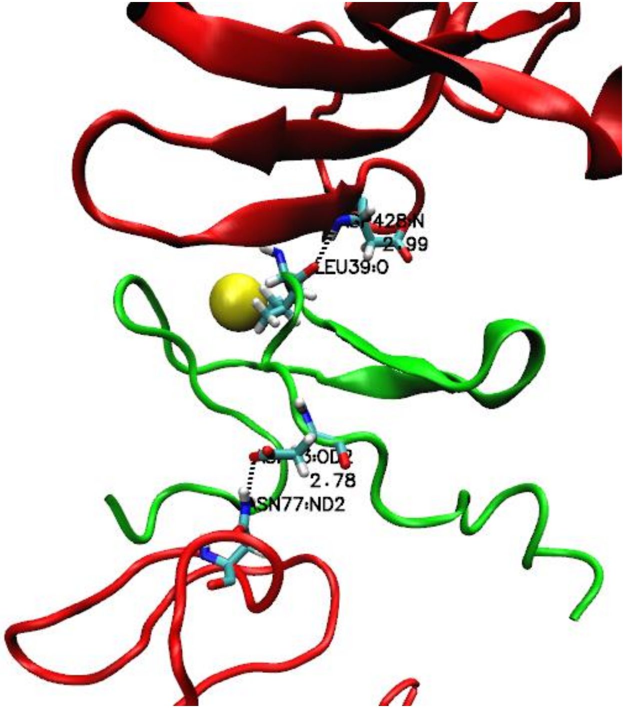
TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB interacts with Rpb1



TFIIB Zn ribbon domain interacts with RPB1

Core PIC formation: TFIIF-PolII join the complex

TFIIB recruits the TFIIF-PolII assembly to the PIC: Zn Ribbon and RPB1 (H bonds)



TFIIB

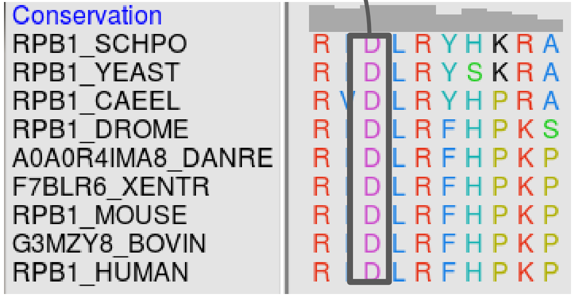
Leu 39

Asp 43

Conservation			
TF2B_SCHPO	PPnLVEEFSS	GDTVCGSCGL	VLGDRIIDTR
TF2B_YEAST	PPkIVERFSE	GDVVCALCGL	VLSDKLVDTR
TF2B_CAEEL	VH-LIEDHRA	GDLVCPACGL	VVGDRLVDVG
A0A0R4IRB5_DANRE	AL-LVEDYRA	GDMICPECGL	VVGDRVIDVG
Q28G23_XENTR	AL-LVEDYRA	GDMICSECGL	VVGDRVIDVG
F1NXP2_CHICK	SI-LVEDYRA	GDMICSECGL	VVGDRVIDVG
TF2B_MOUSE	AI-LVEDYRA	GDMICPECGL	VVGDRVIDVG
TF2B_BOVIN	AI-LVEDYRA	GDMICPECGL	VVGDRVIDVG
TF2B_HUMAN	AI-LVEDYRA	GDMICPECGL	VVGDRVIDVG

RPB1

Asp 428



Conservation	
RPB1_SCHPO	RIDLRYHKRA
RPB1_YEAST	RIDLRYSKRA
RPB1_CAEEL	RVDLRYHPRA
RPB1_DROME	RIDLRFHPKS
A0A0R4IMA8_DANRE	RIDLRFHPKP
F7BLR6_XENTR	RIDLRFHPKP
RPB1_MOUSE	RIDLRFHPKP
G3MZY8_BOVIN	RIDLRFHPKP
RPB1_HUMAN	RIDLRFHPKP

RPB1

Asn 77

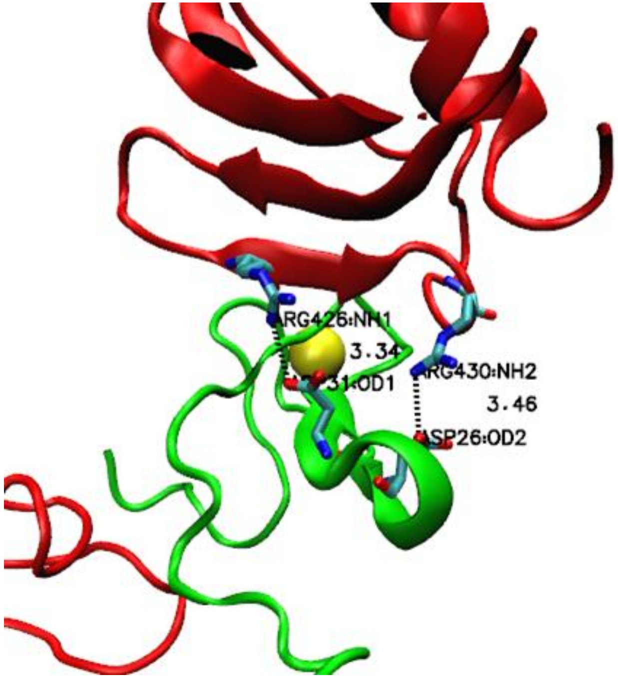
Conservation	
RPB1_SCHPO	QTCGETMADC
RPB1_YEAST	QTCQEGMNEC
RPB1_CAEEL	MTCAGNLTDC
RPB1_DROME	QTCAGNMTEC
A0A0R4IMA8_DANRE	QTCAGNMTEC
F7BLR6_XENTR	QTCAGNMTEC
RPB1_MOUSE	QTCAGNMTEC
G3MZY8_BOVIN	QTCAGNMTEC
RPB1_HUMAN	QTCAGNMTEC

Interaction preserved across **eukaryotes**

Interaction preserved across **metazoans**

Core PIC formation: TFIIF-PolII join the complex

TFIIB recruits the TFIIF-PolII assembly to the PIC: Zn Ribbon and RPB1 (Electrostatic Interactions)



TFIIB

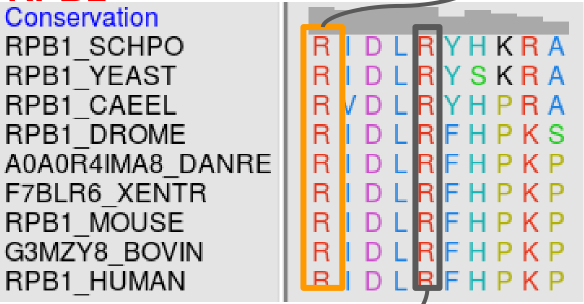
Asp 26

Asp 31

Conservation																				
TF2B_SCHPO	P	P	n	L	V	E	E	F	S	S	G	D	T	V	C	G	S	C	G	L
TF2B_YEAST	P	P	k	I	V	E	R	F	S	E	G	D	V	V	C	A	L	C	G	L
TF2B_CAEEL	V	H	-	L	I	E	D	H	R	A	G	D	L	V	C	P	A	C	G	L
A0A0R4IRB5_DANRE	A	L	-	L	V	E	D	Y	R	A	G	D	M	I	C	P	E	C	G	L
Q28G23_XENTR	A	L	-	L	V	E	D	Y	R	A	G	D	M	I	C	S	E	C	G	L
F1NXP2_CHICK	S	I	-	L	V	E	D	Y	R	A	G	D	M	I	C	S	E	C	G	L
TF2B_MOUSE	A	I	-	L	V	E	D	Y	R	A	G	D	M	I	C	P	E	C	G	L
TF2B_BOVIN	A	I	-	L	V	E	D	Y	R	A	G	D	M	I	C	P	E	C	G	L
TF2B_HUMAN	A	I	-	L	V	E	D	Y	R	A	G	D	M	I	C	P	E	C	G	L

RPB1

Arg 426



Conservation										
RPB1_SCHPO	R	I	D	L	R	Y	H	K	R	A
RPB1_YEAST	R	I	D	L	R	Y	S	K	R	A
RPB1_CAEEL	R	V	D	L	R	Y	H	P	R	A
RPB1_DROME	R	I	D	L	R	F	H	P	K	S
A0A0R4IMA8_DANRE	R	I	D	L	R	F	H	P	K	P
F7BLR6_XENTR	R	I	D	L	R	F	H	P	K	P
RPB1_MOUSE	R	I	D	L	R	F	H	P	K	P
G3MZY8_BOVIN	R	I	D	L	R	F	H	P	K	P
RPB1_HUMAN	R	I	D	L	R	F	H	P	K	P

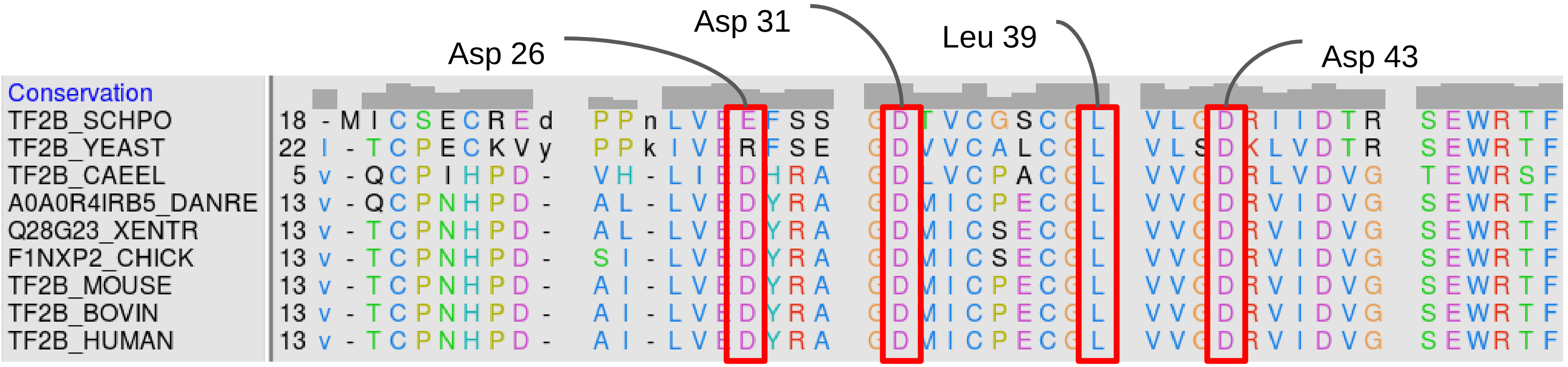
Arg 430

Both interactions are **mostly preserved across eukaryotes**

Core PIC formation: TFIIF-PolII join the complex

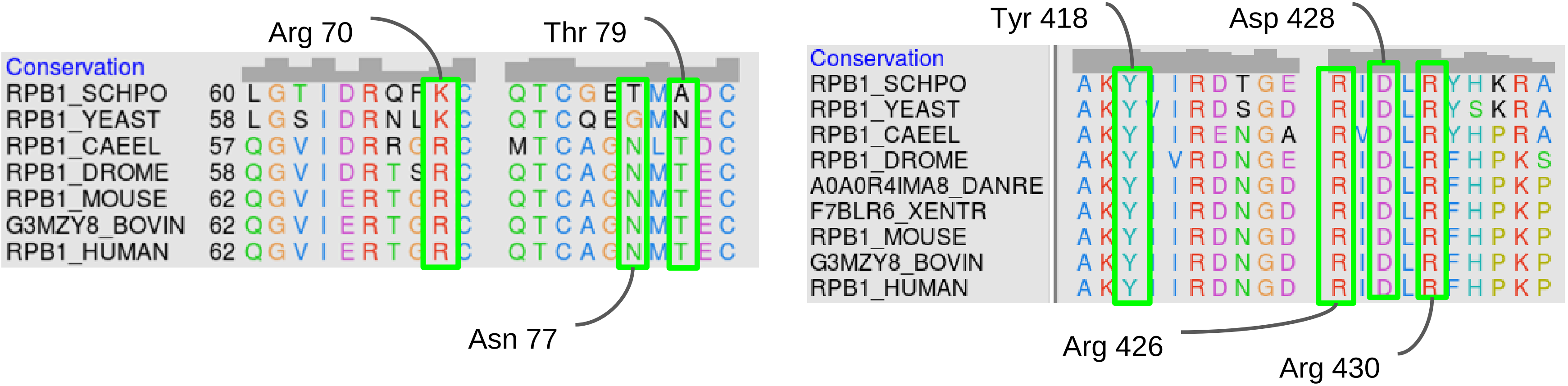
TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB interacts with RPB1

TFIIB



Conservation						
TF2B_SCHPO	18	- M I C S E C R E d	P P n L V E E F S S	G D T V C G S C G L	V L G D R I I D T R	S E W R T F
TF2B_YEAST	22	I - T C P E C K V y	P P k I V E R F S E	G D V V C A L C G L	V L S D K L V D T R	S E W R T F
TF2B_CAEEL	5	v - Q C P I H P D -	V H - L I E D H R A	G D L V C P A C G L	V V G D R L V D V G	T E W R S F
A0A0R4IRB5_DANRE	13	v - Q C P N H P D -	A L - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F
Q28G23_XENTR	13	v - T C P N H P D -	A L - L V E D Y R A	G D M I C S E C G L	V V G D R V I D V G	S E W R T F
F1NXP2_CHICK	13	v - T C P N H P D -	S I - L V E D Y R A	G D M I C S E C G L	V V G D R V I D V G	S E W R T F
TF2B_MOUSE	13	v - T C P N H P D -	A I - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F
TF2B_BOVIN	13	v - T C P N H P D -	A I - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F
TF2B_HUMAN	13	v - T C P N H P D -	A I - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F

RPB1

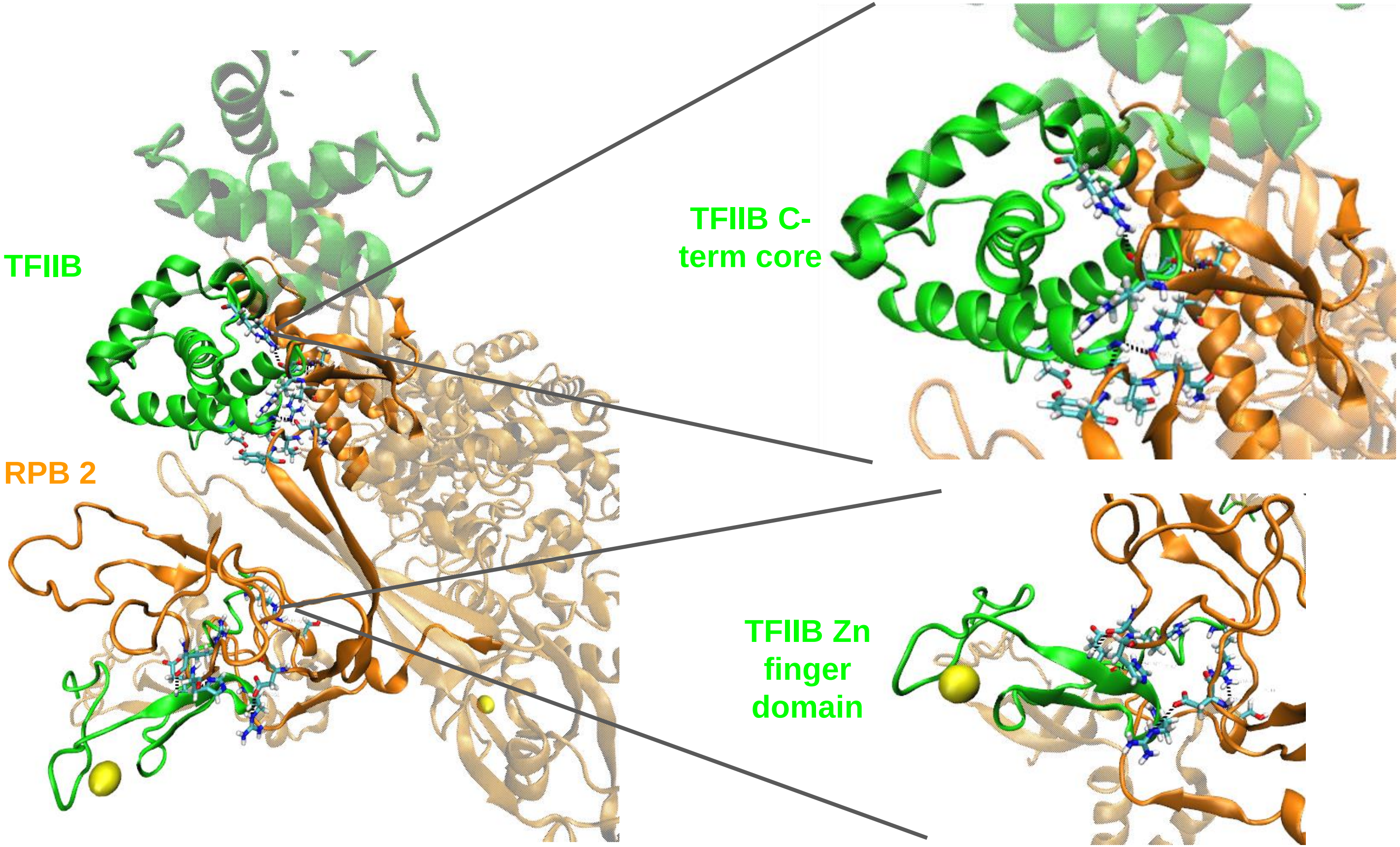


Conservation			
RPB1_SCHPO	60	L G T I D R Q F K C	Q T C G E T M A D C
RPB1_YEAST	58	L G S I D R N L K C	Q T C Q E G M N E C
RPB1_CAEEL	57	Q G V I D R R G R C	M T C A G N L T D C
RPB1_DROME	58	Q G V I D R T S R C	Q T C A G N M T E C
RPB1_MOUSE	62	Q G V I E R T G R C	Q T C A G N M T E C
G3MZY8_BOVIN	62	Q G V I E R T G R C	Q T C A G N M T E C
RPB1_HUMAN	62	Q G V I E R T G R C	Q T C A G N M T E C

Conservation		
RPB1_SCHPO	A K Y I I R D T G E	R I D L R Y H K R A
RPB1_YEAST	A K Y V I R D S G D	R I D L R Y S K R A
RPB1_CAEEL	A K Y I I R E N G A	R V D L R Y H P R A
RPB1_DROME	A K Y I V R D N G E	R I D L R F H P K S
A0A0R4IMA8_DANRE	A K Y I I R D N G D	R I D L R F H P K P
F7BLR6_XENTR	A K Y I I R D N G D	R I D L R F H P K P
RPB1_MOUSE	A K Y I I R D N G D	R I D L R F H P K P
G3MZY8_BOVIN	A K Y I I R D N G D	R I D L R F H P K P
RPB1_HUMAN	A K Y I I R D N G D	R I D L R F H P K P

Core PIC formation: TFIIF-PolII join the complex

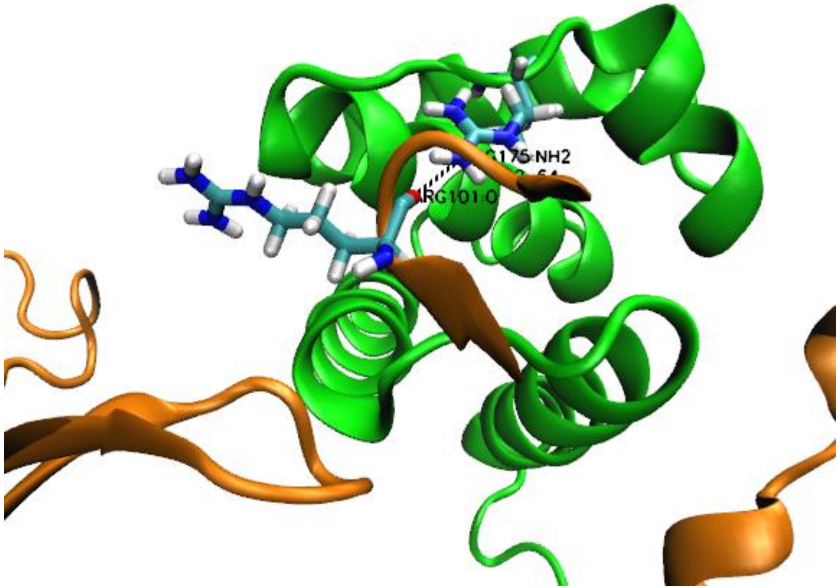
TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB interacts with Rpb2



TFIIB Zn ribbon and N-term cyclin like domains interacts with **RPB2**

Core PIC formation: TFIIF-PolII join the complex

TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB N-term cyclin like domain interacts with Rpb2 (Hydrogen bond)



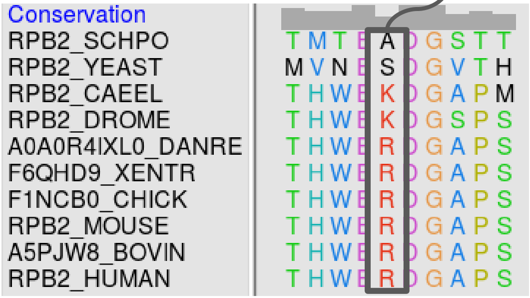
TFIIB

Arg 175

Conservation			
TF2B_SCHPO	S I I	A A C I Y I A C R Q	G K V P R T F M E I
TF2B_YEAST	S I M	A A S I L I G C R R	A E V A R T F K E I
TF2B_CAEEL	A Q A	A A C L Y I A C R K	D G V P R T F K E I
TF2B_DROME	A K A	S A C L Y I A C R Q	E G V P R T F K E I
A0A0R4IRB5_DANRE	A I A	S A C L Y I A C R Q	E G V P R T F K E I
L7N3H9_XENTR	A I A	S A C L Y I A C R Q	E G V P R T F K E I
F1NXP2_CHICK	A I A	S A C L Y I A C R Q	E G V P R T F K E I
Q3ULN2_MOUSE	A I A	S A C L Y I A C R Q	E G V P R T F K E I
TF2B_BOVIN	A I A	S A C L Y I A C R Q	E G V P R T F K E I
TF2B_HUMAN	A I A	S A C L Y I A C R Q	E G V P R T F K E I

RPB2

Arg 101

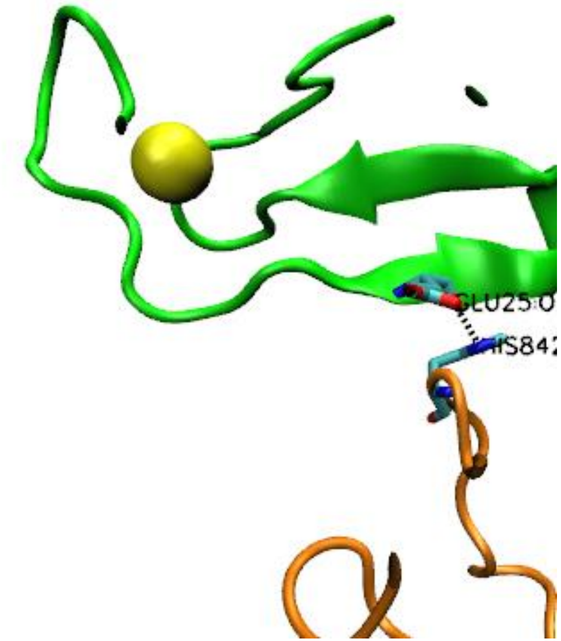


Conservation	
RPB2_SCHPO	T M T E A D G S T T
RPB2_YEAST	M V N E S D G V T H
RPB2_CAEEL	T H W E K D G A P M
RPB2_DROME	T H W E K D G S P S
A0A0R4IXL0_DANRE	T H W E R D G A P S
F6QHD9_XENTR	T H W E R D G A P S
F1NCB0_CHICK	T H W E R D G A P S
RPB2_MOUSE	T H W E R D G A P S
A5PJW8_BOVIN	T H W E R D G A P S
RPB2_HUMAN	T H W E R D G A P S

Interaction conserved across **metazoans**

Core PIC formation: TFIIF-PolII join the complex

TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB Zn-ribbon domain interacts with Rpb2 (Electrostatic Interaction)



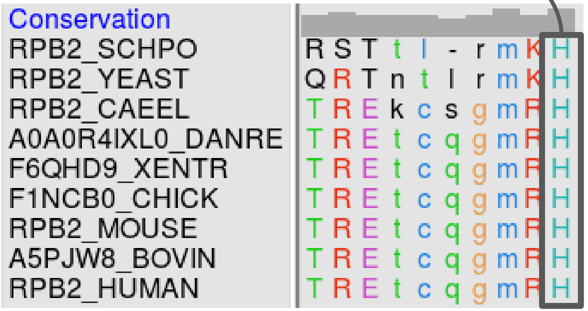
TFIIB

Glu 25

Conservation										
TF2B_SCHPO	P	P	n	L	V	E	E	F	S	S
TF2B_YEAST	P	P	k	I	V	E	R	F	S	E
TF2B_CAEEL	V	H	-	L	I	E	D	H	R	A
A0A0R4IRB5_DANRE	A	L	-	L	V	E	D	Y	R	A
Q28G23_XENTR	A	L	-	L	V	E	D	Y	R	A
F1NXP2_CHICK	S	I	-	L	V	E	D	Y	R	A
TF2B_MOUSE	A	I	-	L	V	E	D	Y	R	A
TF2B_BOVIN	A	I	-	L	V	E	D	Y	R	A
TF2B_HUMAN	A	I	-	L	V	E	D	Y	R	A

RPB2

His 842



Conservation										
RPB2_SCHPO	R	S	T	t	l	-	r	m	K	H
RPB2_YEAST	Q	R	T	n	t	l	r	m	K	H
RPB2_CAEEL	T	R	E	k	c	s	g	m	R	H
A0A0R4IXL0_DANRE	T	R	E	t	c	q	g	m	R	H
F6QHD9_XENTR	T	R	E	t	c	q	g	m	R	H
F1NCB0_CHICK	T	R	E	t	c	q	g	m	R	H
RPB2_MOUSE	T	R	E	t	c	q	g	m	R	H
A5PJW8_BOVIN	T	R	E	t	c	q	g	m	R	H
RPB2_HUMAN	T	R	E	t	c	q	g	m	R	H

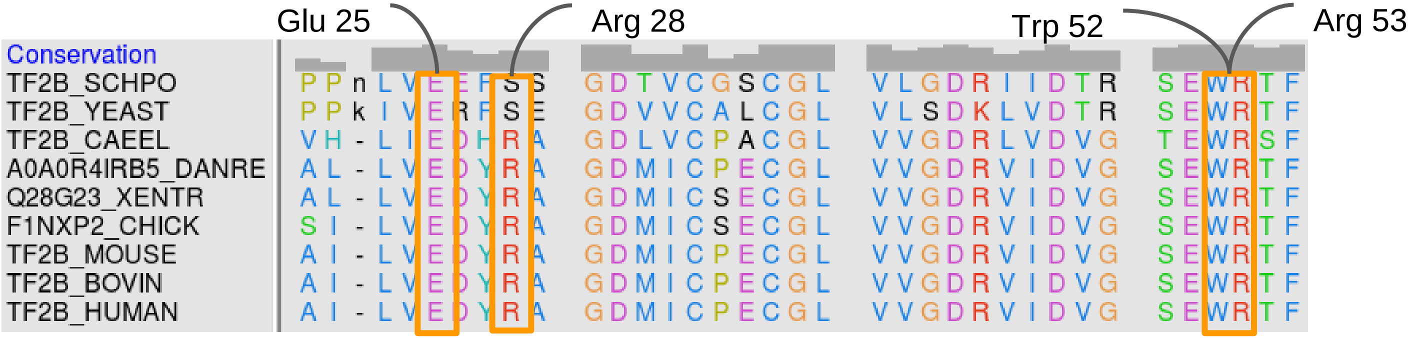
Interaction conserved across **eukaryotes**

Core PIC formation: TFIIF-PolII join the complex

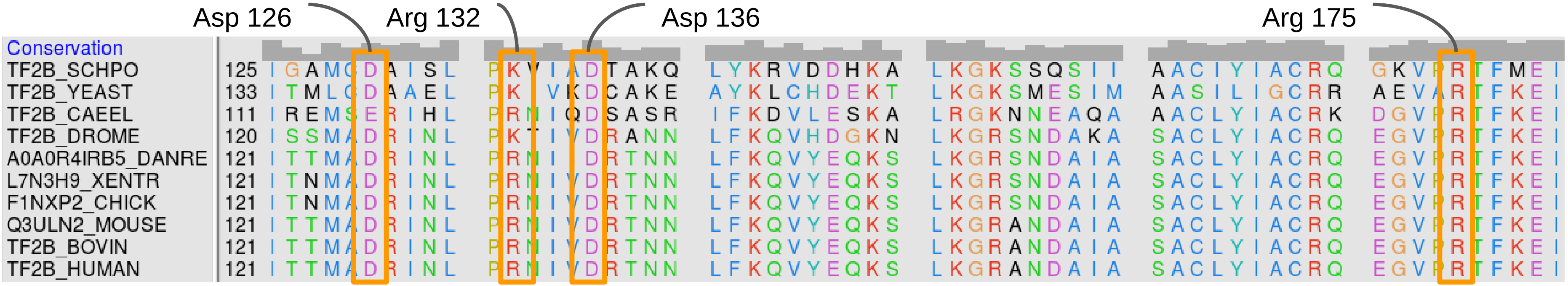
TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB interacts with Rpb2

TFIIB

Most residues conserved across eukaryotes



Conservation				
TF2B_SCHPO	P P n L V E E F S S	G D T V C G S C G L	V L G D R I I D T R	S E W R T F
TF2B_YEAST	P P k I V E R F S E	G D V V C A L C G L	V L S D K L V D T R	S E W R T F
TF2B_CAEEL	V H - L I E D H R A	G D L V C P A C G L	V V G D R L V D V G	T E W R S F
A0A0R4IRB5_DANRE	A L - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F
Q28G23_XENTR	A L - L V E D Y R A	G D M I C S E C G L	V V G D R V I D V G	S E W R T F
F1NXP2_CHICK	S I - L V E D Y R A	G D M I C S E C G L	V V G D R V I D V G	S E W R T F
TF2B_MOUSE	A I - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F
TF2B_BOVIN	A I - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F
TF2B_HUMAN	A I - L V E D Y R A	G D M I C P E C G L	V V G D R V I D V G	S E W R T F



Conservation							
TF2B_SCHPO	125	I G A M C D A I S L	P K V I A D T A K Q	L Y K R V D D H K A	L K G K S S Q S I I	A A C I Y I A C R Q	G K V P R T F M E I
TF2B_YEAST	133	I T M L C D A A E L	P K V K D C A K E	A Y K L C H D E K T	L K G K S M E S I M	A A S I L I G C R R	A E V A R T F K E I
TF2B_CAEEL	111	I R E M S E R I H L	P R N I Q D S A S R	I F K D V L E S K A	L R G K N N E A Q A	A A C L Y I A C R K	D G V P R T F K E I
TF2B_DROME	120	I S S M A D R I N L	P K T I V D R A N N	L F K Q V H D G K N	L K G R S N D A K A	S A C L Y I A C R Q	E G V P R T F K E I
A0A0R4IRB5_DANRE	121	I T T M A D R I N L	P R N I D R T N N	L F K Q V Y E Q K S	L K G R S N D A I A	S A C L Y I A C R Q	E G V P R T F K E I
L7N3H9_XENTR	121	I T N M A D R I N L	P R N I V D R T N N	L F K Q V Y E Q K S	L K G R S N D A I A	S A C L Y I A C R Q	E G V P R T F K E I
F1NXP2_CHICK	121	I T N M A D R I N L	P R N I V D R T N N	L F K Q V Y E Q K S	L K G R S N D A I A	S A C L Y I A C R Q	E G V P R T F K E I
Q3ULN2_MOUSE	121	I T T M A D R I N L	P R N I V D R T N N	L F K Q V Y E Q K S	L K G R A N D A I A	S A C L Y I A C R Q	E G V P R T F K E I
TF2B_BOVIN	121	I T T M A D R I N L	P R N I V D R T N N	L F K Q V Y E Q K S	L K G R A N D A I A	S A C L Y I A C R Q	E G V P R T F K E I
TF2B_HUMAN	121	I T T M A D R I N L	P R N I V D R T N N	L F K Q V Y E Q K S	L K G R A N D A I A	S A C L Y I A C R Q	E G V P R T F K E I

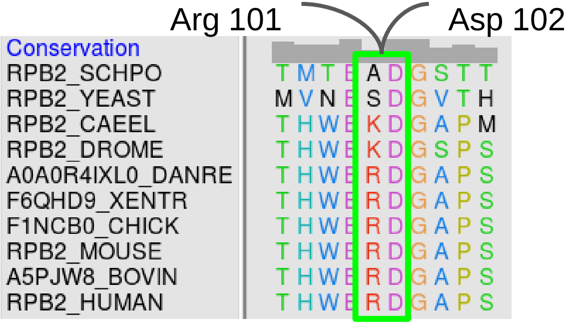
TFIIB residues relevant for Rpb2 interaction

Core PIC formation: TFIIF-PolII join the complex

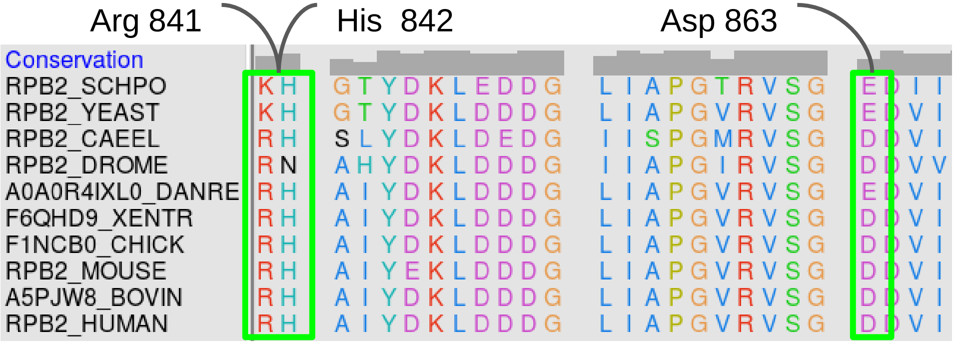
TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB interacts with Rpb2

RPB 2

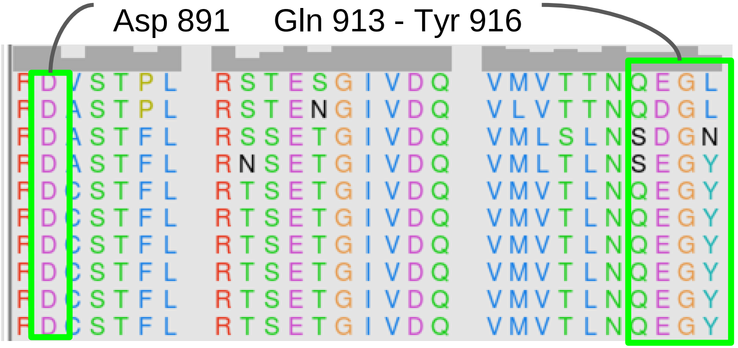
Most residues conserved across eukaryotes



Conservation			
RPB2_SCHPO	T M T E	A D	G S T T
RPB2_YEAST	M V N E	S D	G V T H
RPB2_CAEEL	T H W E	K D	G A P M
RPB2_DROME	T H W E	K D	G S P S
A0A0R4IXL0_DANRE	T H W E	R D	G A P S
F6QHD9_XENTR	T H W E	R D	G A P S
F1NCB0_CHICK	T H W E	R D	G A P S
RPB2_MOUSE	T H W E	R D	G A P S
A5PJW8_BOVIN	T H W E	R D	G A P S
RPB2_HUMAN	T H W E	R D	G A P S



Conservation				
RPB2_SCHPO	K H	G T Y D K L E D D G	L I A P G T R V S G	E D I I
RPB2_YEAST	K H	G T Y D K L D D D G	L I A P G V R V S G	E D V I
RPB2_CAEEL	R H	S L Y D K L D E D G	I I S P G M R V S G	D D V I
RPB2_DROME	R N	A H Y D K L D D D G	I I A P G I R V S G	D D V V
A0A0R4IXL0_DANRE	R H	A I Y D K L D D D G	L I A P G V R V S G	E D V I
F6QHD9_XENTR	R H	A I Y D K L D D D G	L I A P G V R V S G	D D V I
F1NCB0_CHICK	R H	A I Y D K L D D D G	L I A P G V R V S G	D D V I
RPB2_MOUSE	R H	A I Y E K L D D D G	L I A P G V R V S G	D D V I
A5PJW8_BOVIN	R H	A I Y D K L D D D G	L I A P G V R V S G	D D V I
RPB2_HUMAN	R H	A I Y D K L D D D G	L I A P G V R V S G	D D V I

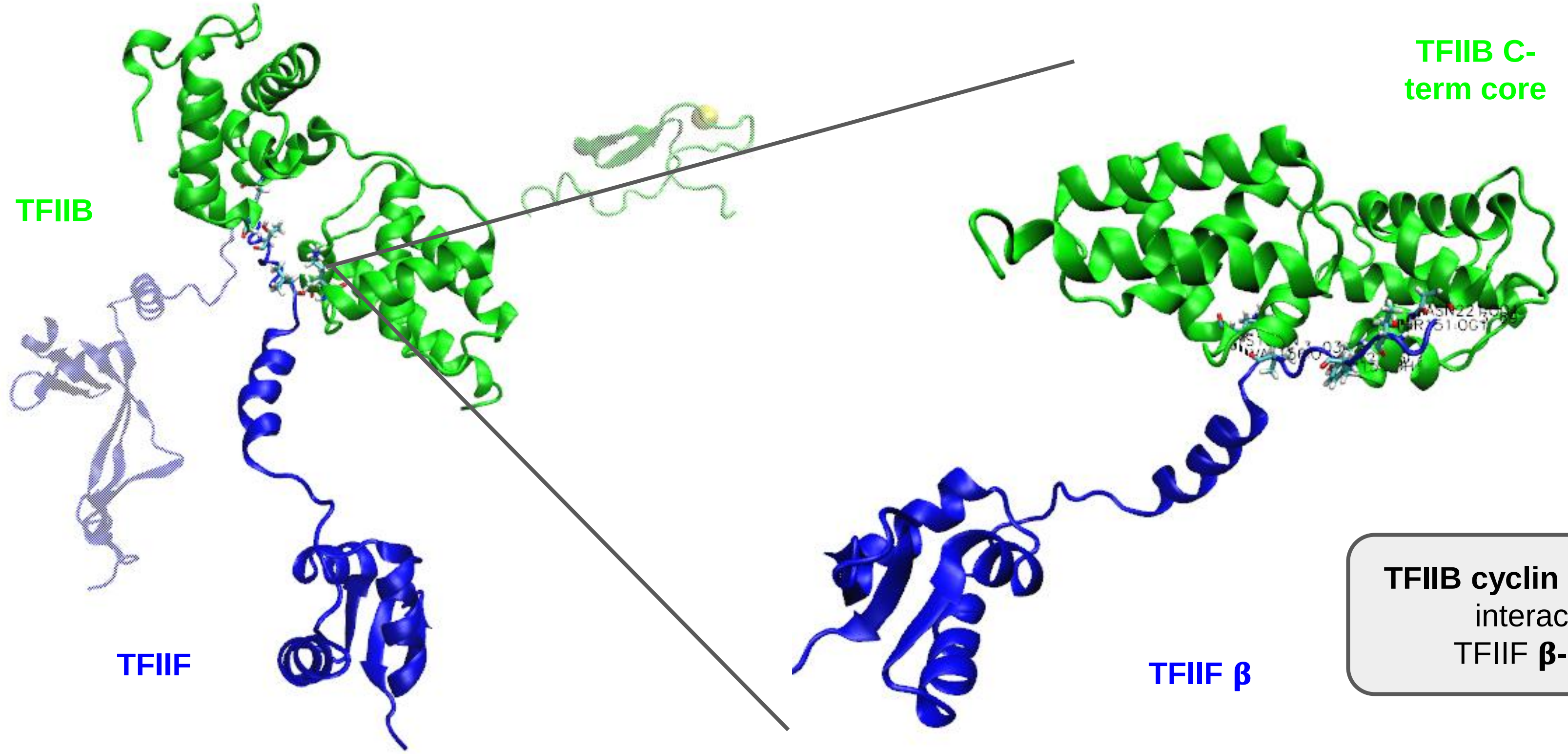


R D V S T P L	R S T E S G I V D Q	V M V T T N	Q E G L
R D A S T P L	R S T E N G I V D Q	V L V T T N	Q D G L
R D A S T F L	R S S E T G I V D Q	V M L S L N	S D G N
R D A S T F L	R N S E T G I V D Q	V M L T L N	S E G Y
R D C S T F L	R T S E T G I V D Q	V M V T L N	Q E G Y
R D C S T F L	R T S E T G I V D Q	V M V T L N	Q E G Y
R D C S T F L	R T S E T G I V D Q	V M V T L N	Q E G Y
R D C S T F L	R T S E T G I V D Q	V M V T L N	Q E G Y
R D C S T F L	R T S E T G I V D Q	V M V T L N	Q E G Y
R D C S T F L	R T S E T G I V D Q	V M V T L N	Q E G Y

Rpb2 residues relevant for TFIIB interaction

Core PIC formation: TFIIF-PolII join the complex

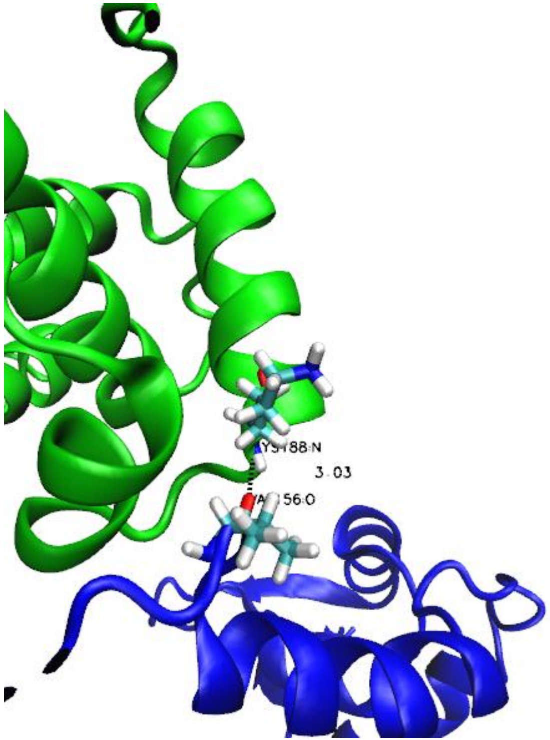
TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB interacts with TFIIF



TFIIB cyclin like domains interacts with TFIIF **β-domain**

Core PIC formation: TFIIF-PolII join the complex

TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB cyclin like domains interacts with TFIIF**β** (Hydrogen bond)



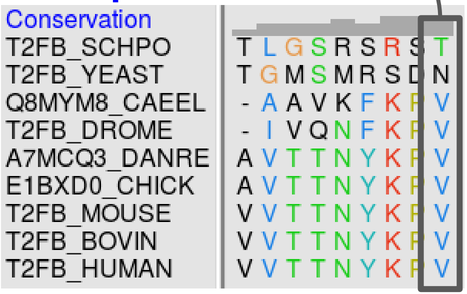
TFIIB

Lys 188

Conservation				
TF2B_SCHPO	SII	AACIYIACRQ	GKVPRTFMEI	CTLTNVPKKE
TF2B_YEAST	SIM	AASILIGCRR	AEVARTFKEI	QSLIHVKTKE
TF2B_CAEEL	AQA	AACLYIACRK	DGVPRTFKEI	CAVSRVSKKE
TF2B_DROME	AKA	SACLYIACRQ	EGVPRTFKEI	CAVSKISKKE
A0A0R4IRB5_DANRE	AIA	SACLYIACRQ	EGVPRTFKEI	CAVSRISKKE
L7N3H9_XENTR	AIA	SACLYIACRQ	EGVPRTFKEI	CAVSRISKKE
F1NXP2_CHICK	AIA	SACLYIACRQ	EGVPRTFKEI	CAVSRISKKE
Q3ULN2_MOUSE	AIA	SACLYIACRQ	EGVPRTFKEI	CAVSRISKKE
TF2B_BOVIN	AIA	SACLYIACRQ	EGVPRTFKEI	CAVSRISKKE
TF2B_HUMAN	AIA	SACLYIACRQ	EGVPRTFKEI	CAVSRISKKE

TFIIFβ

Val 156



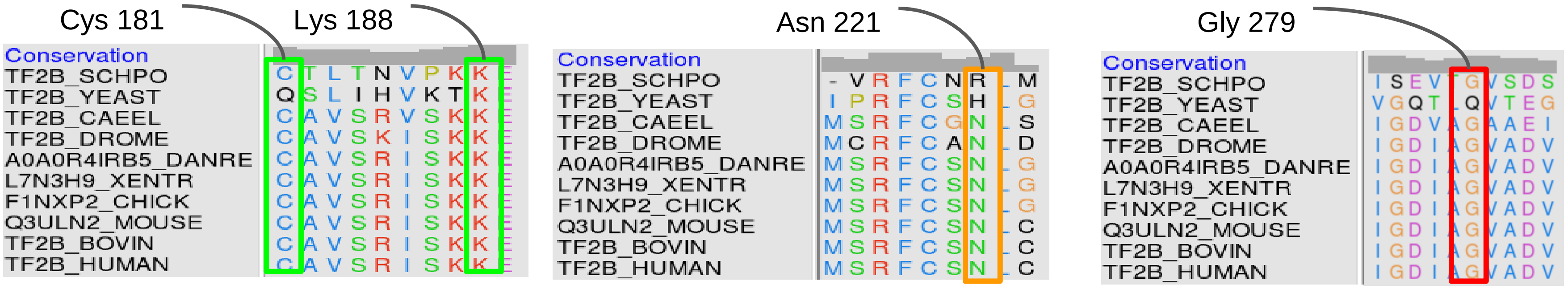
Conservation	
T2FB_SCHPO	TLGSRSRST
T2FB_YEAST	TGMSMRSDN
Q8MYM8_CAEEL	-AAVKFKPV
T2FB_DROME	-IVQNFKPV
A7MCQ3_DANRE	AVTTNYKPV
E1BXD0_CHICK	AVTTNYKPV
T2FB_MOUSE	VVTTNYKPV
T2FB_BOVIN	VVTTNYKPV
T2FB_HUMAN	VVTTNYKPV

Interaction conserved across **metazoans**

Core PIC formation: TFIIF-PolII join the complex

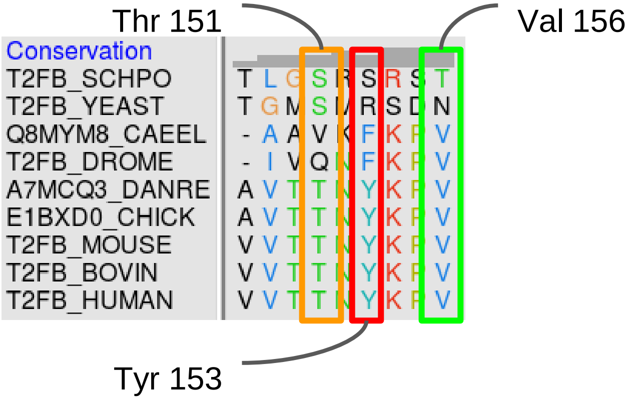
TFIIB recruits the TFIIF-PolII assembly to the PIC: TFIIB N-term cyclin domains interacts with TFIIFβ

TFIIB



Conservation			
TF2B_SCHPO	C T L T N V P K K E	- V R F C N R L M	I S E V T G V S D S
TF2B_YEAST	Q S L I H V K T K E	I P R F C S H L G	V G Q T L Q V T E G
TF2B_CAEEL	C A V S R V S K K E	M S R F C G N L S	I G D V A G A A E I
TF2B_DROME	C A V S K I S K K E	M C R F C A N L D	I G D I A G V A D V
A0A0R4IRB5_DANRE	C A V S R I S K K E	M S R F C S N L G	I G D I A G V A D V
L7N3H9_XENTR	C A V S R I S K K E	M S R F C S N L G	I G D I A G V A D V
F1NXP2_CHICK	C A V S R I S K K E	M S R F C S N L G	I G D I A G V A D V
Q3ULN2_MOUSE	C A V S R I S K K E	M S R F C S N L C	I G D I A G V A D V
TF2B_BOVIN	C A V S R I S K K E	M S R F C S N L C	I G D I A G V A D V
TF2B_HUMAN	C A V S R I S K K E	M S R F C S N L C	I G D I A G V A D V

TFIIF

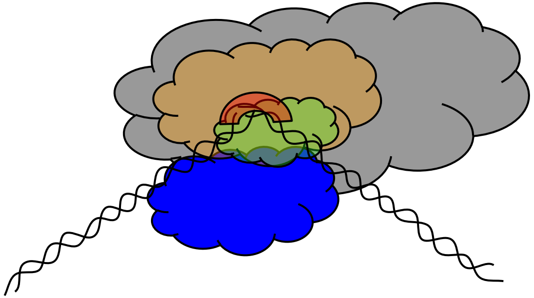


Conservation	
T2FB_SCHPO	T L G S R S R S T
T2FB_YEAST	T G M S I R S D N
Q8MYM8_CAEEL	- A A V K F K P V
T2FB_DROME	- I V Q N F K P V
A7MCQ3_DANRE	A V T T N Y K P V
E1BXD0_CHICK	A V T T N Y K P V
T2FB_MOUSE	V V T T N Y K P V
T2FB_BOVIN	V V T T N Y K P V
T2FB_HUMAN	V V T T N Y K P V

Most conserved interaction are those between TFIIB residues and TFIIF Val 156

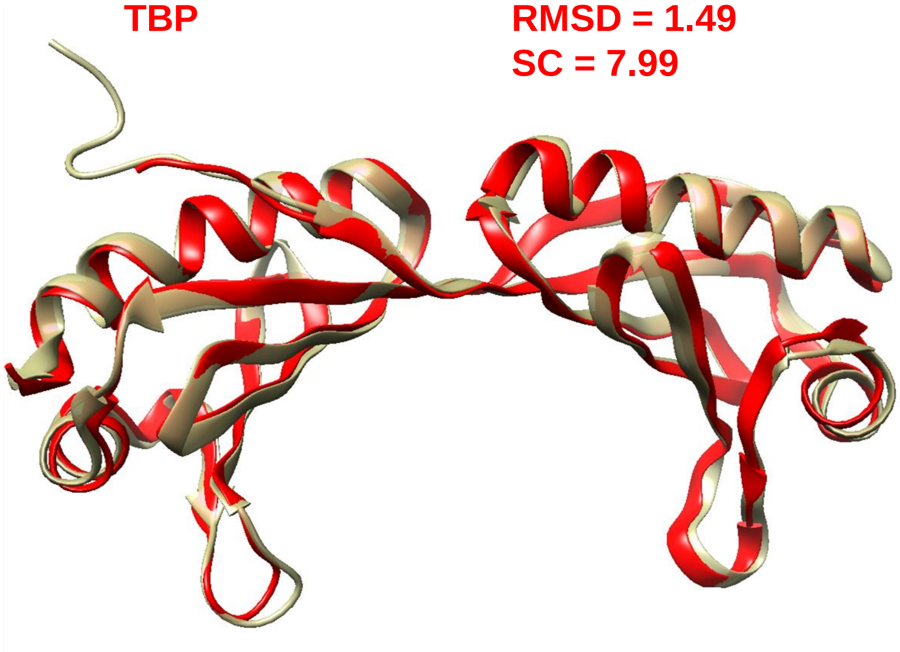
Core PIC:

Is structure conserved?



Core PIC: Is structure conserved?

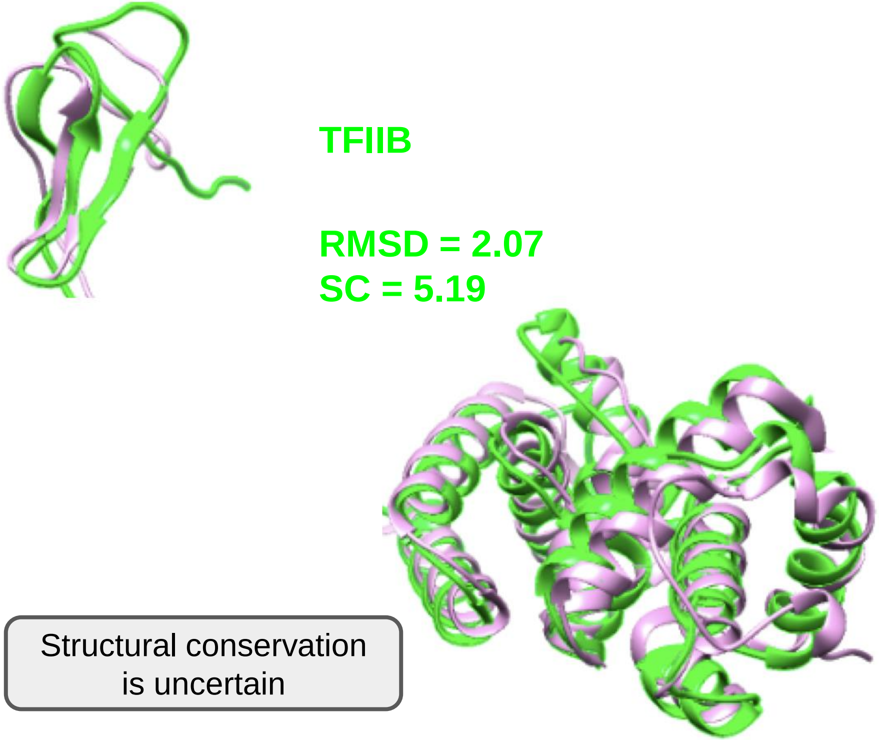
TBP and TFIIB are conserved across eukaryotes



TBP

RMSD = 1.49
SC = 7.99

Highly conserved



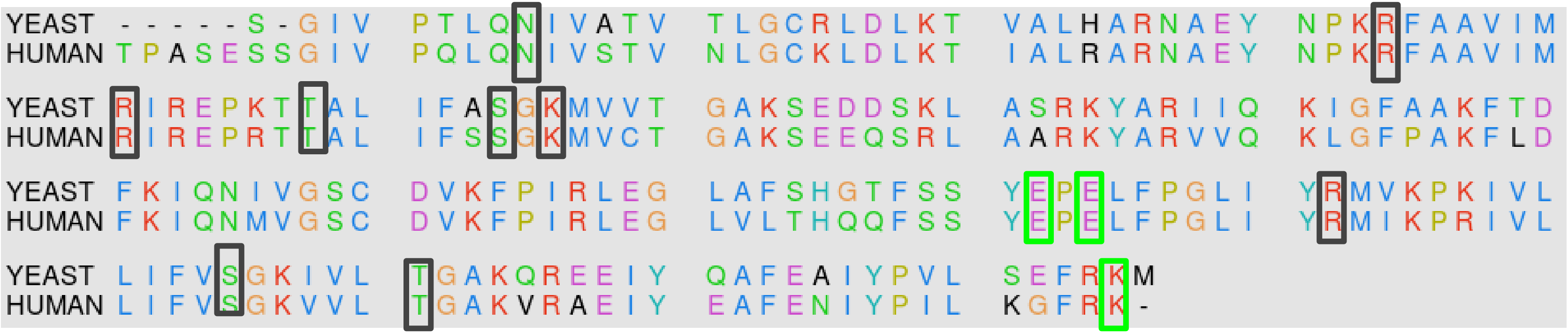
TFIIB

RMSD = 2.07
SC = 5.19

Structural conservation is uncertain

Core PIC: Is structure conserved?

All relevant residues of TBP are conserved



```
YEAST - - - - - S - G I V   P T L Q N I V A T V   T L G C R L D L K T   V A L H A R N A E Y   N P K R F A A V I M
HUMAN T P A S E S S G I V   P Q L Q N I V S T V   N L G C K L D L K T   I A L R A R N A E Y   N P K R F A A V I M

YEAST R I R E P K T T A L   I F A S G K M V V T   G A K S E D D S K L   A S R K Y A R I I Q   K I G F A A K F T D
HUMAN R I R E P R T T A L   I F S S G K M V C T   G A K S E E Q S R L   A A R K Y A R V V Q   K L G F P A K F L D

YEAST F K I Q N I V G S C   D V K F P I R L E G   L A F S H G T F S S   Y E P E L F P G L I   Y R M V K P K I V L
HUMAN F K I Q N M V G S C   D V K F P I R L E G   L V L T H Q Q F S S   Y E P E L F P G L I   Y R M I K P R I V L

YEAST L I F V S G K I V L   T G A K Q R E E I Y   Q A F E A I Y P V L   S E F R K M
HUMAN L I F V S G K V V L   T G A K V R A E I Y   E A F E N I Y P I L   K G F R K -
```

Interaction with TFIIB

Interaction with DNA

All the **important residues** involved in TBP’s interactions are **conserved at structural level**

Structure and sequence **alignments match**

Core PIC: Is structure conserved?

TFIIB-mediated Pol-II recruitment and DNA assembly might be conserved

TFIIB Zn-finger domain

```
YEAST_M  PNLNIVLT-C  PECKVYPPKI  VERFSEGDVV  CALCGLVLSD  KLVDN-----
HUMAN_M  -LDALPRVTC  PNHP--DAIL  VEDYRAGDMI  CPECGLVVGD  RVI--DVGSE
          *
YEAST_M  ----------
HUMAN_M  WRTFSNDKAT
```

TFIIB C-term core

```
YEAST_M              EVQAAFAKI   TMLCDAAELP  KIVKDCAKEA  YKLCHDEKTL
HUMAN_M              AMMNAFKEI   TTMADRINLP  RNIVDRTNNL  FKQVYEQKSL

YEAST_M  KGKSMESIMA  ASILIGCRRA  EVARTFKEIQ  SLIHVKTKEF  GKTLNIMKNI
HUMAN_M  KGRANDAIAS  ACLYIACRQE  GVPRTFKEIC  AVSRISKKEI  GRCFKLILKA

YEAST_M  LRGKS--Q--  -NLTYIPRFC  SHLGLPMQVT  TSAEYTAKKC  KEIKEIAGKS
HUMAN_M  LE-T-SVDLI  TTGDFMSRFC  SNLCLPKQVQ  MAATHIARKA  VELDLVPGRS

YEAST_M  PITIAVVSIY  LNILLFQIPI  TAAKVGQTLQ  VTEGTIKSGY  KILYEHRDKL
HUMAN_M  PISVAAAAIY  MASQASAEKR  TQKEIGDIAG  VADVTIRQSY  RLIYPRAPDL

YEAST_M  VDP-------  -------
HUMAN_M  FPTDFKFDTP  VDKLPQL
```

Interaction with TBP

Interaction with DNA

Interaction with RPB1

Interaction with RPB2

Interaction with TFIIFβ

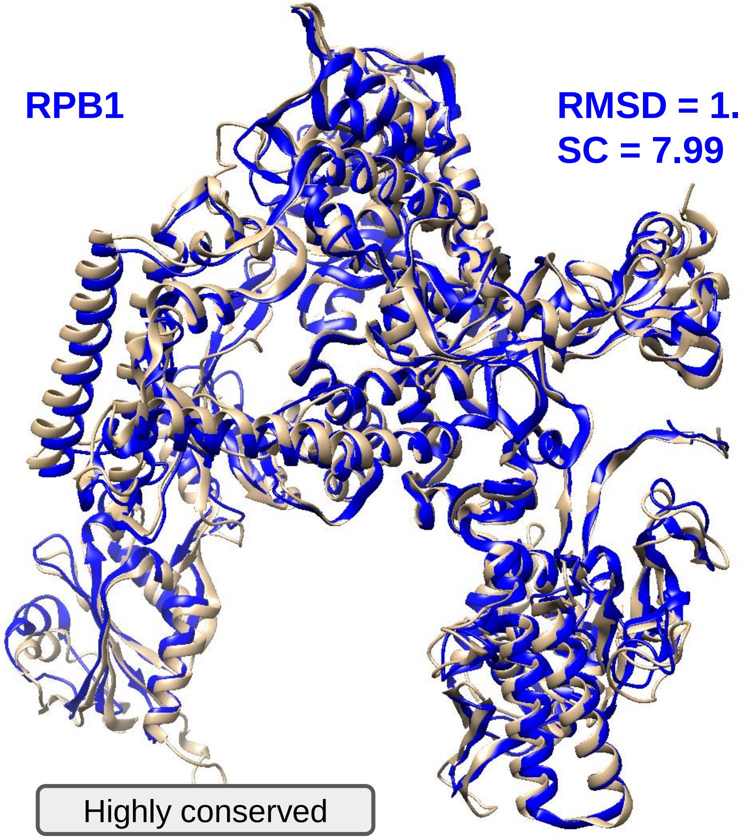
(*) Residues for which sequence and structural alignment do not match

Residues relevant for **TFIIB interactions with Rpb1**, **Rpb2** and **DNA** are the most **conserved.**

There is one **disagreement** with the **sequence alignment**

Core PIC: Is structure conserved?

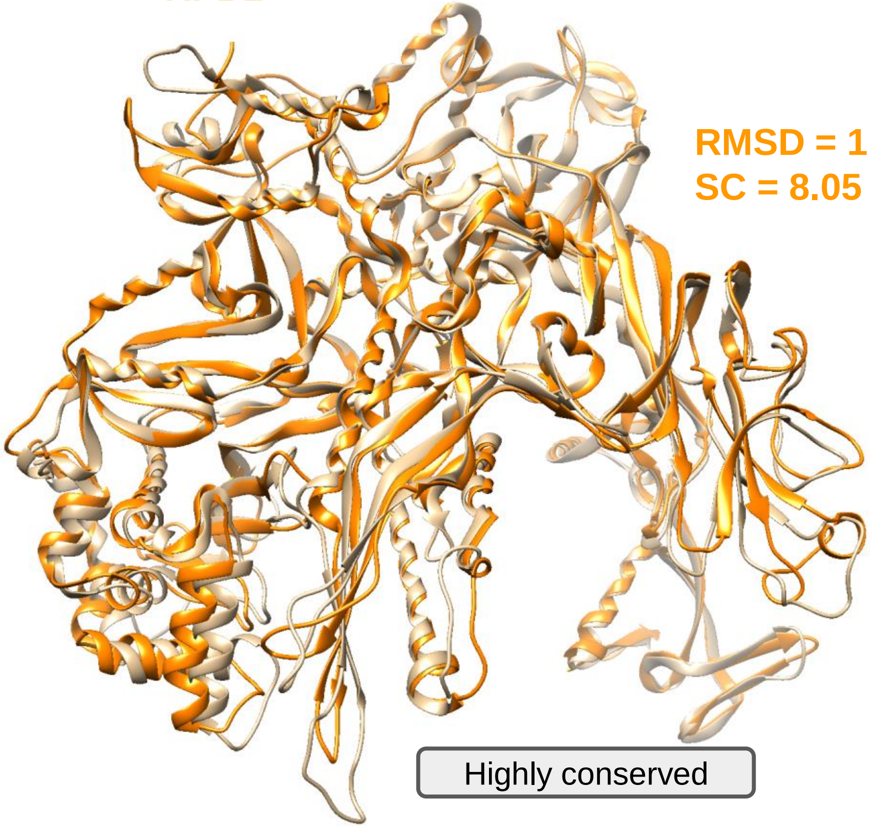
RPB1 and RPB2 are highly conserved across eukaryotes



RPB1

RMSD = 1.49
SC = 7.99

Highly conserved



RPB2

RMSD = 1.32
SC = 8.05

Highly conserved

Core PIC: Is structure conserved?

The relevant residues of RPB1 are conserved

```
YEAST_A GQQYSSAPLR TVKEVQFGLF SPEEVRAISV A-KIRF-PET MDETQTRAKI
HUMAN_A ---DSACPLR TIKRVQFGVL SPDELKRMSV TEGGIKYPET TEG-GRPKL-

YEAST_A GGL-NDPRLG SIDRNLKCQT CQEGMNECPG HFGHIDLAKP VFHVGFIAKI
HUMAN_A GGLMDPR-QG VIERTGRCQT CAGNMTECPG HFGHIELAKP VFHVGFLVKT

YEAST_A KKVCECVCMH CGKLLLDEHN ELMRQALAI- K-DSKKRFAA IWTLCKTKMV
HUMAN_A MKVLRCVCFF CSKLLVDSNN PKIKDILAKS KGQPKKRLTH VYDLCKGKNI

YEAST_A CETDVPSR-G GCGNTQPTIR KDGLKLVGSW K--K----PE LRVLSTEEIL
HUMAN_A CE-GG--GHG GCGRYQPRIR RSGLELYAEW KHVNEDSQEK KILLSPERVH

YEAST_A NIFKHISVKD FTSLGFNEVF SRPEWMILTC LPVPPPPVR- PS-ISFNESQ
HUMAN_A EIFKRISDEE CFVLGMEPRY ARPEWMIVTV LPVPPLSVRP AVVMQG---S

YEAST_A R-GEDDLTFK LADILKANIS LETLEHNGAP HHAIEEAESL LQFHVATYMD
HUMAN_A ARNQDDLTHK LADIVKINNQ LRRNEQNGAA AHVIAEDVKL LQFHVATMVD

YEAST_A NDIAGQPQAL QKSGRPVKSI RARLKGKEGR IRGNLMGKRV DFSARTVISG
HUMAN_A NELPGLPRAM QKSGRPLKSL KQRLKGKEGR VRGNLMGKRV DFSARTVITP

YEAST_A DPNLELDQVG VPKSIAKTLT YPEVVTPYNI DRLTQLVRNG PNEHPGAKYV
HUMAN_A DPNLSIDQVG VPRSIAANMT FAEIVTPFNI DRLQELVRRG NSQYPGAKYI

YEAST_A IRDSGDRIDL RYSKRAGD-- IQLQYGWKVE RHIMDNDPVL FNRQPSLHKM
HUMAN_A IRDNGDRIDL RFHP--KPSD LHLQTGYKVE RHMCDGDIVI FNRQPTLHKM
```

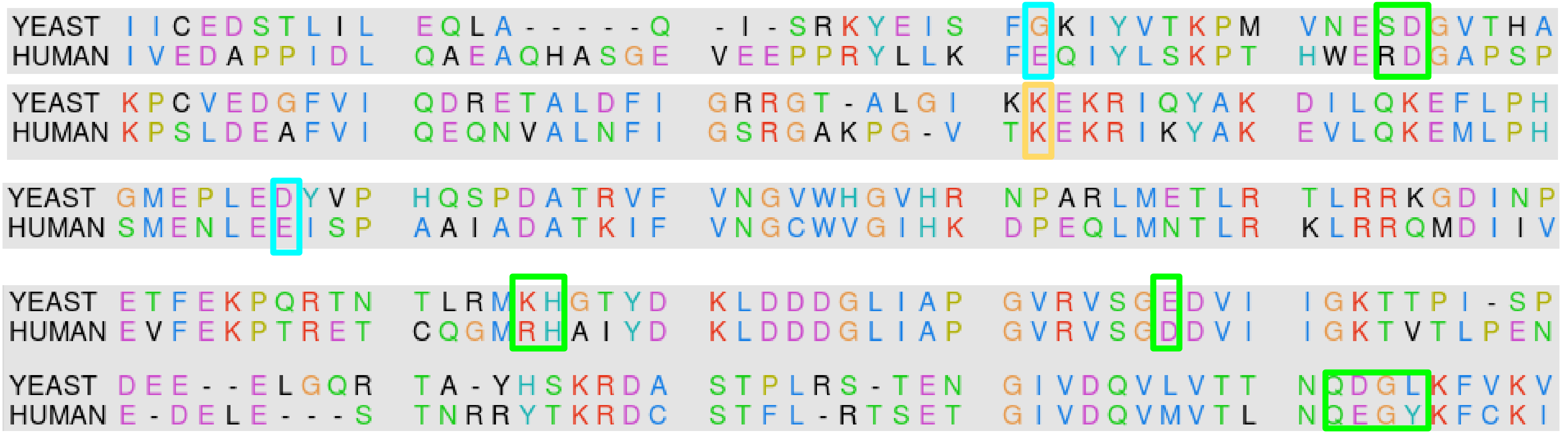
Interaction with TFIIB

Most of the **important residues** involved in Rpb1 interactions are **conserved at structural level**

Structure and sequence **alignments match**

Core PIC: Is structure conserved?

The relevant residues of RPB2 for TFIIB binding are conserved



```
YEAST  IICEDSTLIL EQLA-----Q -I-SRKYEIS FGKIYVTKPM VNESDGVTHA
HUMAN  IVEDAPPIDL QAEAQHASGE VEEPPRYLLK FEQIYLSKPT HWERDGAPSP

YEAST  KPCVEDGFVI QDRETALDFI GRRGT-ALGI KKEKRIQYAK DILQKEFLPH
HUMAN  KPSLDEAFVI QEQNVALNFI GSRGAKPG-V TKEKRIKYAK EVLQKEMLPH

YEAST  GMEPLEDYVP HQSPDATRVF VNGVWHGVHR NPARLMETLR TLRRKGDINP
HUMAN  SMENLEEISP AAIADATKIF VNGCWVGIHK DPEQLMNTLR KLRRQMDIIV

YEAST  ETFEKPQRTN TLRMKHGTYD KLDDDGLIAP GVRVSGEDVI IGKTTPI-SP
HUMAN  EVFEKPTRET CQGMRHAIYD KLDDDGLIAP GVRVSGDDVI IGKTVTLPEN

YEAST  DEE--ELGQR TA-YHSKRDA STPLRS-TEN GIVDQVLVTT NQDGLKFVKV
HUMAN  E-DELE---S TNRRYTKRDC STFL-RTSET GIVDQVMVTL NQEGYKFCKI
```

Interaction with TFIIB

Interaction with TFIIFα

Interaction with TFIIFβ

Residues relevant for **Rpb2 interactions with TFIIB** are the most **conserved.**

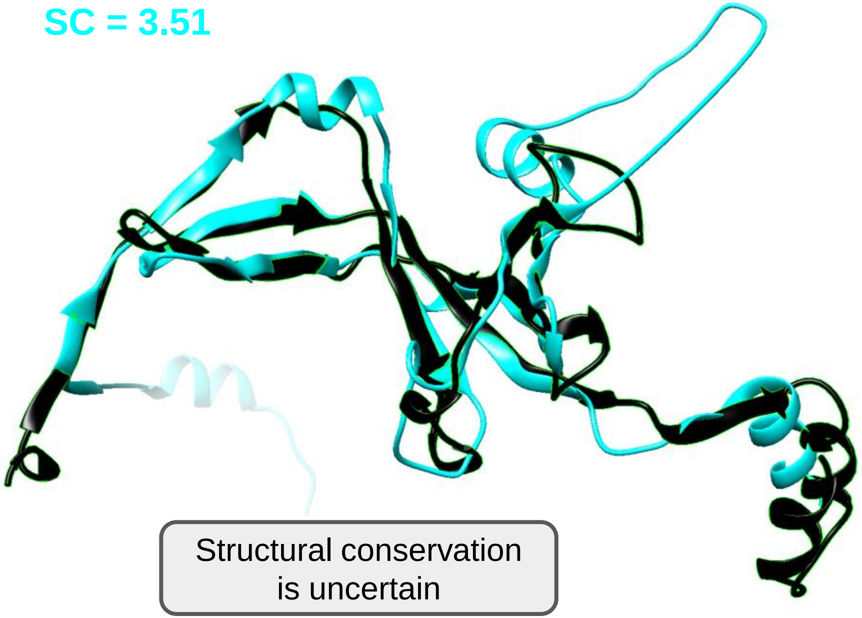
Structure and sequence **alignments match**

Core PIC: Is structure conserved?

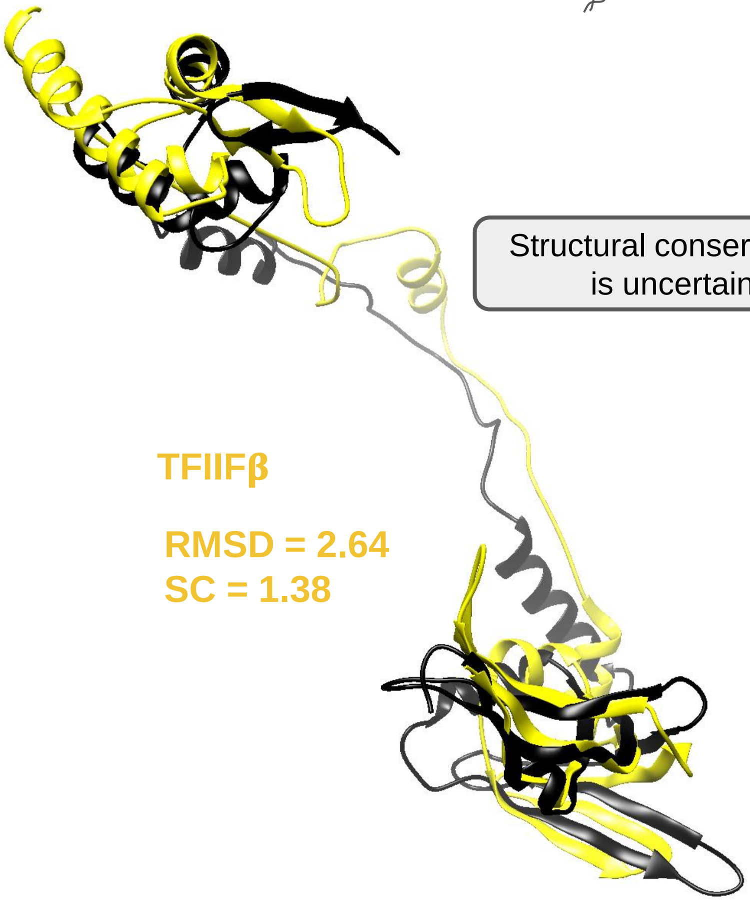
TFIIF may not be fully conserved across eukaryotes

TFIIFα

RMSD = 2.00
SC = 3.51



Structural conservation
is uncertain



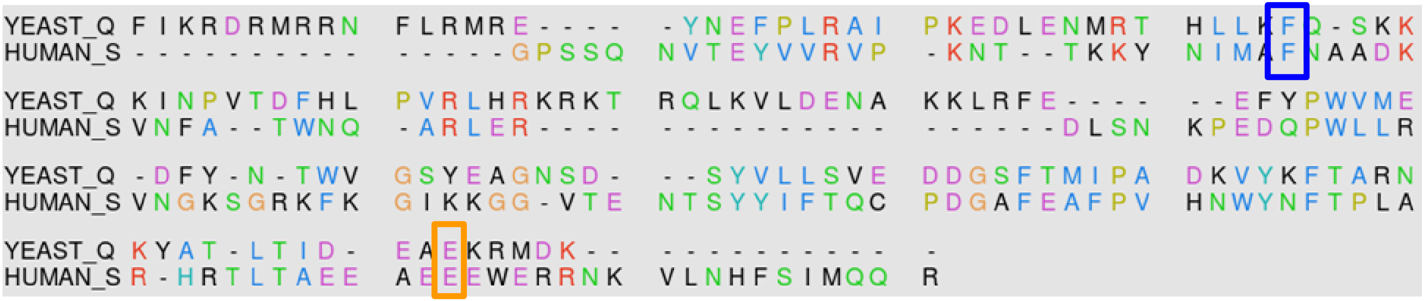
TFIIFβ

RMSD = 2.64
SC = 1.38

Structural conservation
is uncertain

Core PIC: Is structure conserved?

The relevant residues for TFIIFα assembly are conserved



```
YEAST_Q FIKRDRMRRN FLRMRE---- -YNEFPLRAI PKEDLENMRT HLLKFQ-SKK
HUMAN_S ---------- -----GPSSQ NVTEYVVRVP -KNT--TKKY NIMAFNAADK

YEAST_Q KINPVTDFHL PVRLHRKRKT RQLKVLDENA KKLRFE---- --EFYPWVME
HUMAN_S VNFA--TWNQ -ARLER---- ---------- ------DLSN KPEDQPWLLR

YEAST_Q -DFY-N-TWV GSYEAGNSD- --SYVLLSVE DDGSFTMIPA DKVYKFTARN
HUMAN_S VNGKSGRKFK GIKKGG-VTE NTSYYIFTQC PDGAFEAFPV HNWYNFTPLA

YEAST_Q KYAT-LTID- EAEKRMDK-- ---------- -
HUMAN_S R-HRTLTAEE AEEEWERRNK VLNHFSIMQQ R
```

Interaction with TFIIFβ

Interaction with Rpb2

The **important residues** involved in TFIIFα interactions are **conserved at structural level**

Structure and sequence **alignments match**

Core PIC: Is structure conserved?

The relevant residues for TFIIFβ interaction are generally not conserved

```
YEAST - - - - D L D L E R   S - - N R Q V W L V   R L P M F L A E K W   R - - - - L - - - -   G K I R I N K - - -
HUMAN A E R G E L D L T G   A K Q N T G V W L V   K V P K Y L S Q Q W   - A K A S G R G E V   G K L R I A K T Q G

YEAST - S K I T L L L N -   - - E P H E Y D L E   L T K K V V E - - N   E Y V F T E Q T I P   K - - - - K T A I V
HUMAN R T E V S F T L N E   D L A N E - H P F V   L Q S - - V G G Q T   L T V F T E - - - -   - S S S D K L S L E

YEAST G T V C H E C Q V M   - P Y H K I V E Q R   R N I - - - - - V -   - - K - - E R - I T   T L D E T V G V T M
HUMAN G I V V Q R A E C R   P A - - - - A S E N   Y M R L K R L Q I E   E S S K P V R L S Q   Q L D K - - - - V -

YEAST S H T G M S M R S -   - - D N S N F L K V   M - - - - - - - - -   - - - - - - P K K E   I L D Y L F K L F D
HUMAN V T T - - N - Y K P   V A N H - Q - - Y -   - N I E Y E R K K K   E D G K R A R A - -   D K Q H V L D M L F

YEAST E Y - D - - - Y W -   - - - S L K G L K E   R T R Q P E A H L K   E C L D K V A T L V   K K G P Y - - - - A
HUMAN S A F E K H Q Y Y N   L K D L V D I T K -   - - - Q P V V Y L K   E I L K E I G V Q N   V K - - - G I H K N

YEAST F K Y T L - R K E E   E R K A T L G E L
HUMAN T W E L K P E - - -   - - - - - - - - -
```

(*) Residues for which sequence and structural alignment do not match

Interaction with TFIIFα

Interaction with DNA

Interaction with Rpb2

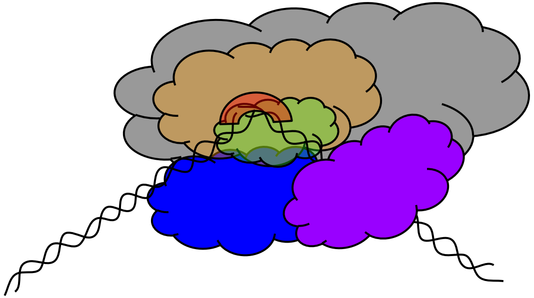
Interaction with TFIIB

Residues relevant for **TFIIFβ interactions with Rpb2** are the ones more **conserved.**

There is some **disagreement** with the **sequence alignment**

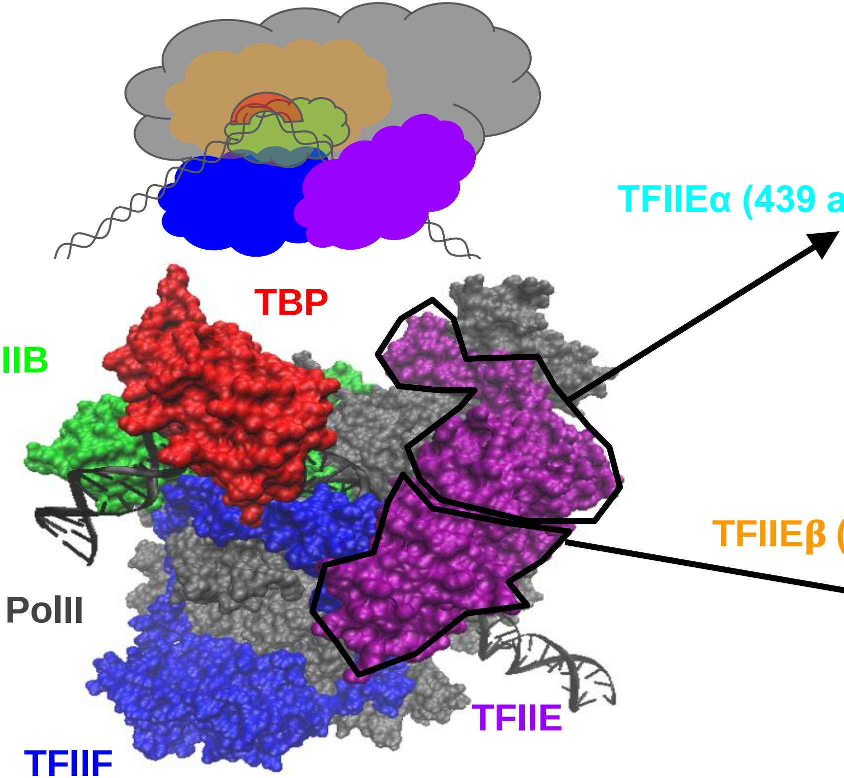
Closed PIC formation:

TFIIE joins the complex



Closed PIC formation: TFIIE joins the complex

TFIIE is an α/β dimer



TBP

TFIIB

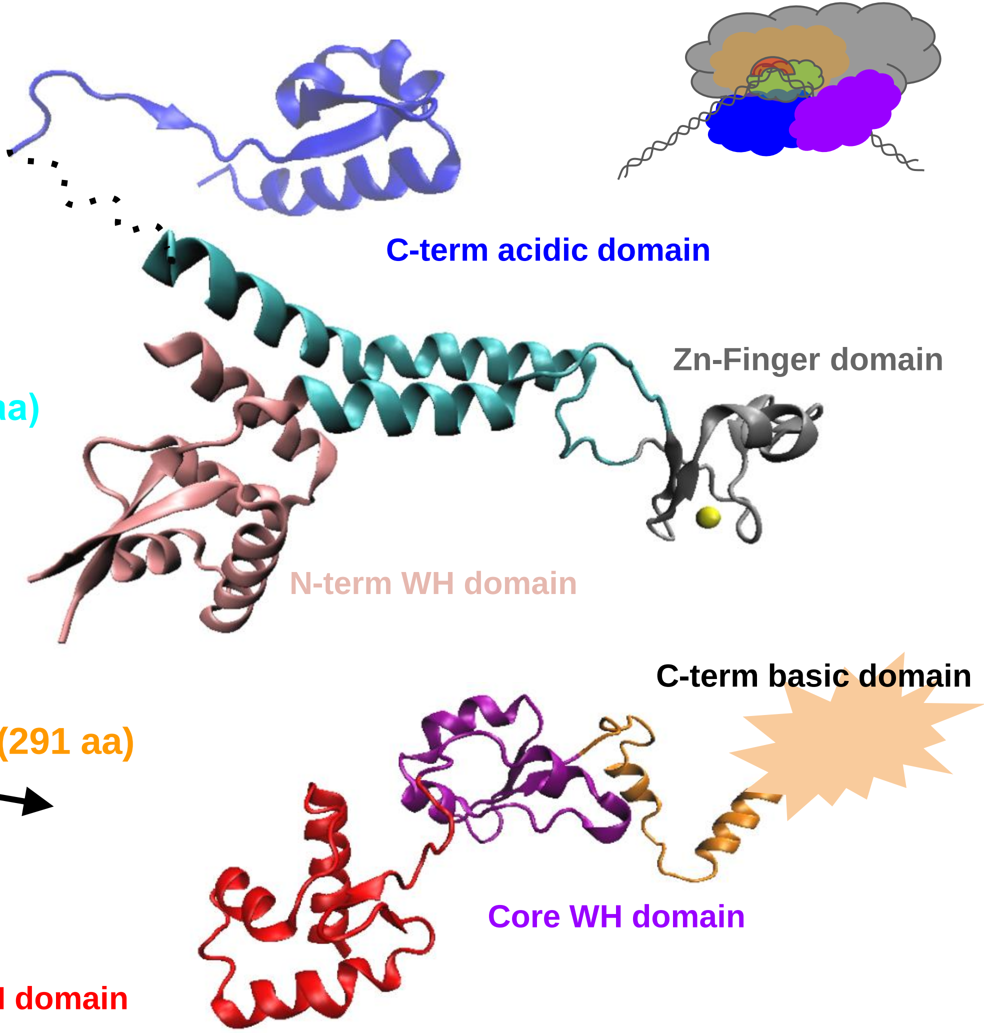
PolII

TFIIE

TFIIF

TFIIEα (439 aa)

TFIIEβ (291 aa)



C-term acidic domain

Zn-Finger domain

N-term WH domain

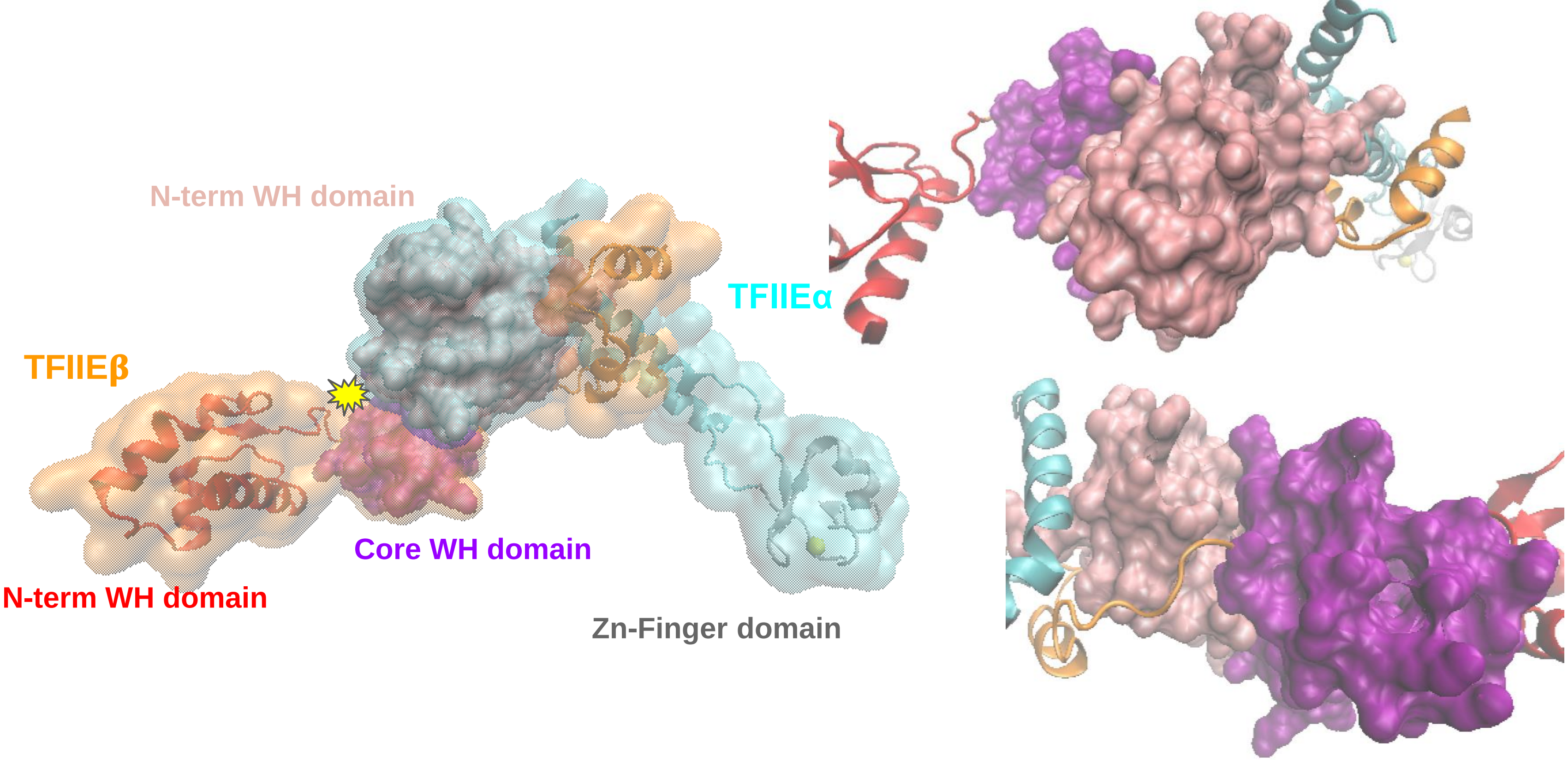
C-term basic domain

Core WH domain

N-term WH domain

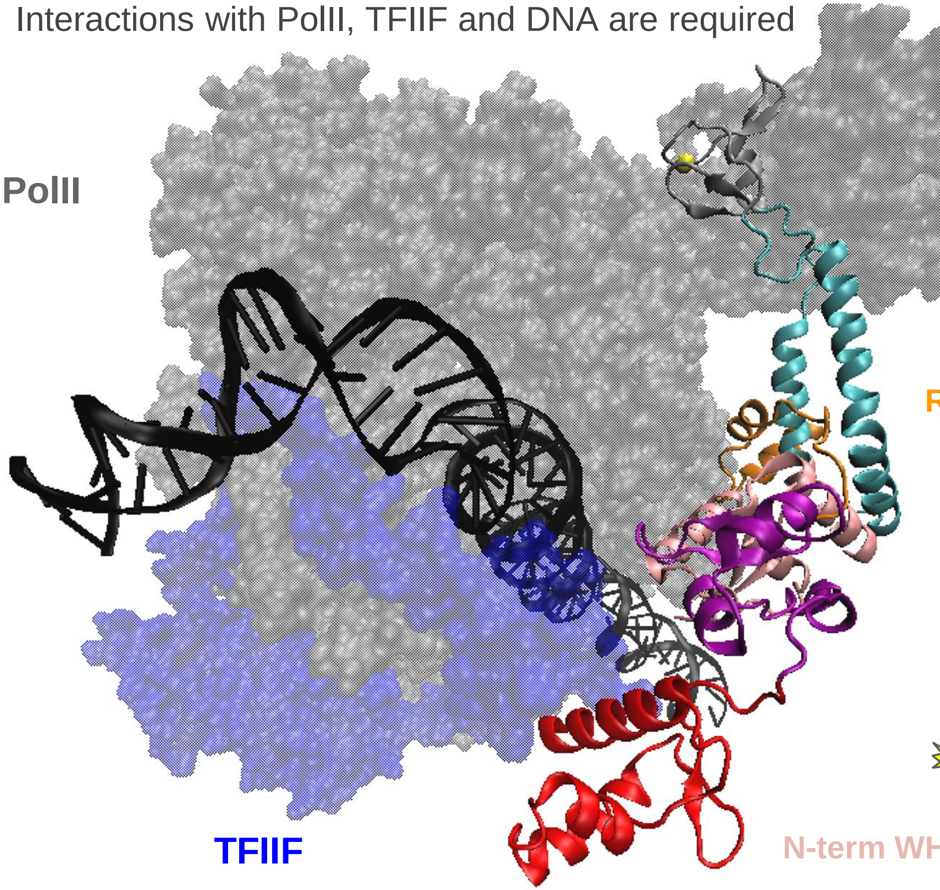
Closed PIC formation: TFIIE joins the complex

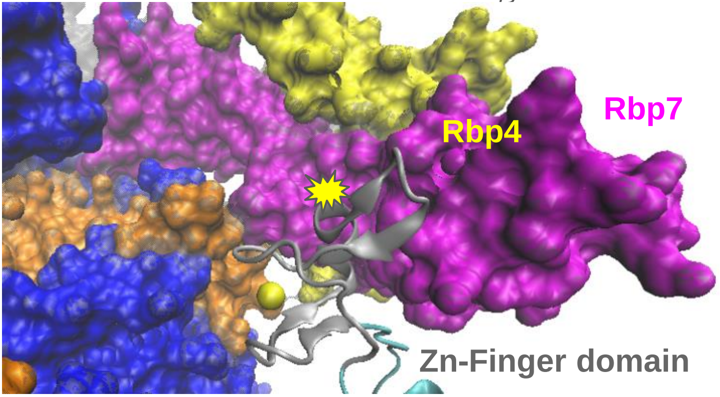
TFIIEα and TFIIE**β** assemble through two WH domains

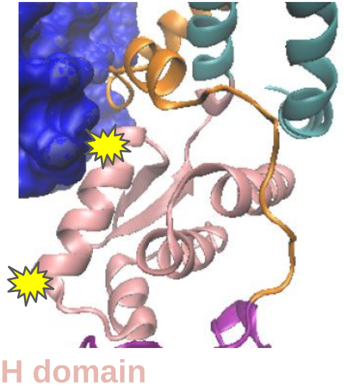


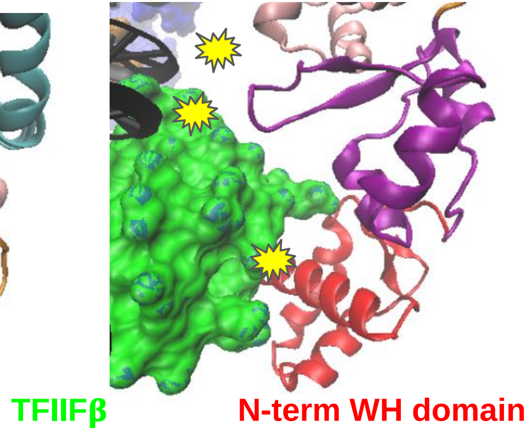
Closed PIC formation: TFIIE joins the complex

Interactions with PolII, TFIIF and DNA are required



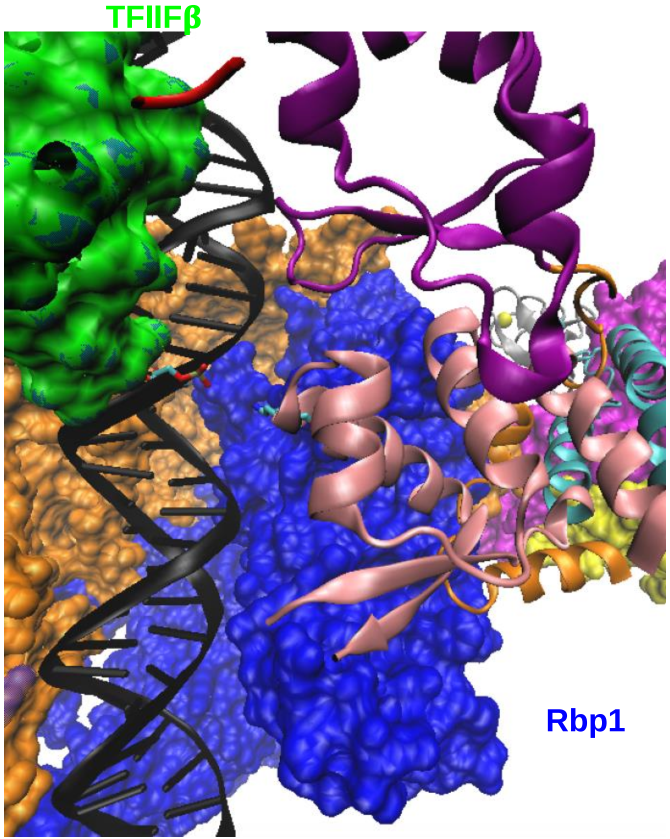


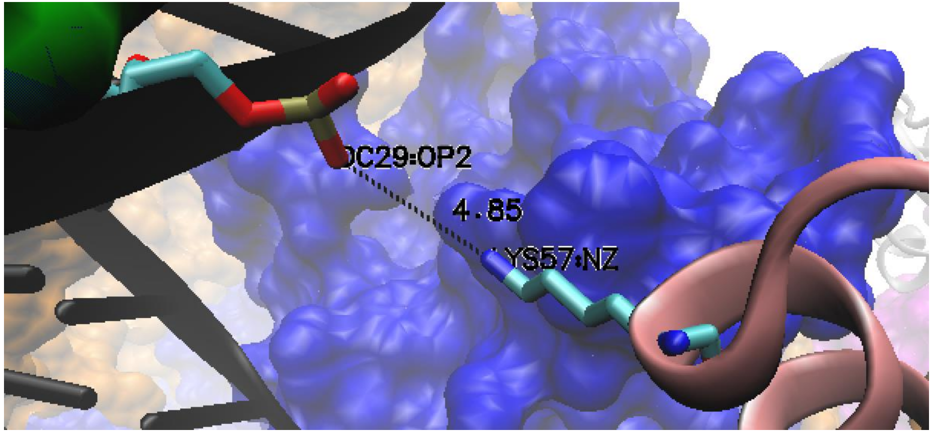


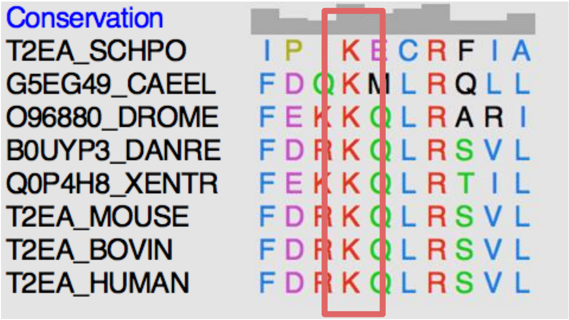


Closed PIC formation: TFIIE joins the complex

The N-term WH domain of TFIIEα is important for DNA binding







Conservation										
T2EA_SCHPO	I	P		K	E	C	R	F	I	A
G5EG49_CAEEL	F	D	Q	K	M	L	R	Q	L	L
O96880_DROME	F	E	K	K	Q	L	R	A	R	I
B0UYP3_DANRE	F	D	R	K	Q	L	R	S	V	L
Q0P4H8_XENTR	F	E	K	K	Q	L	R	T	I	L
T2EA_MOUSE	F	D	R	K	Q	L	R	S	V	L
T2EA_BOVIN	F	D	R	K	Q	L	R	S	V	L
T2EA_HUMAN	F	D	R	K	Q	L	R	S	V	L

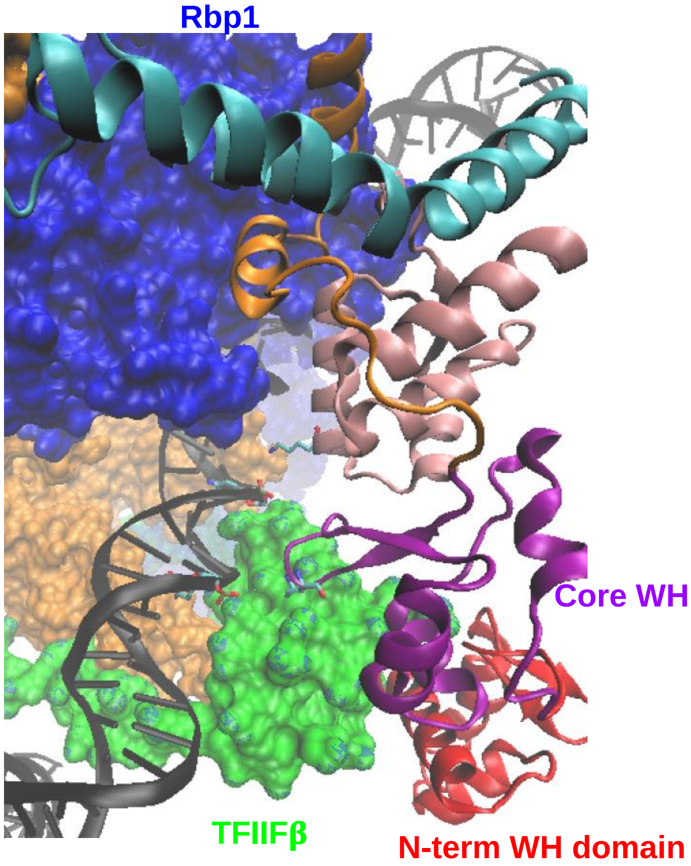
Lys 169

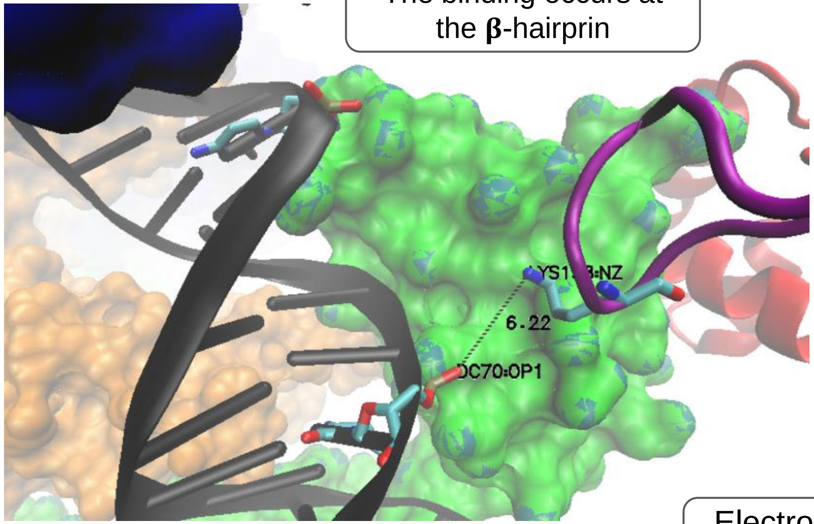
Electrostatic interaction with DNA

Conserved across eukaryotes

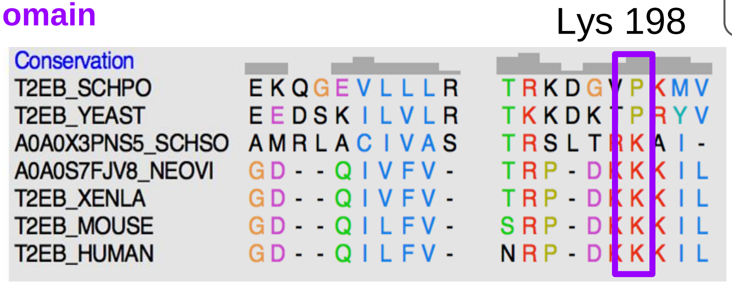
Closed PIC formation: TFIIE joins the complex

The core WH domain of TFIIEβ is important for DNA binding





The binding occurs at the β-hairprin



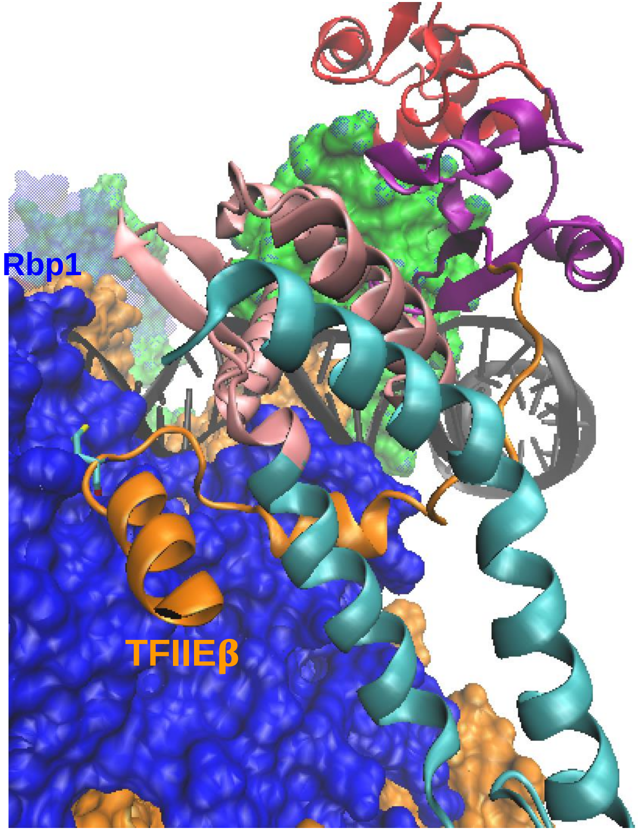
Conservation		
T2EB_SCHPO	E K Q G E V L L L R	T R K D G V P K M V
T2EB_YEAST	E E D S K I L V L R	T K K D K T P R Y V
A0A0X3PNS5_SCHSO	A M R L A C I V A S	T R S L T R K A I -
A0A0S7FJV8_NEOVI	G D - - Q I V F V -	T R P - D K K K I L
T2EB_XENLA	G D - - Q I V F V -	T R P - D K K K I L
T2EB_MOUSE	G D - - Q I L F V -	S R P - D K K K I L
T2EB_HUMAN	G D - - Q I L F V -	N R P - D K K K I L

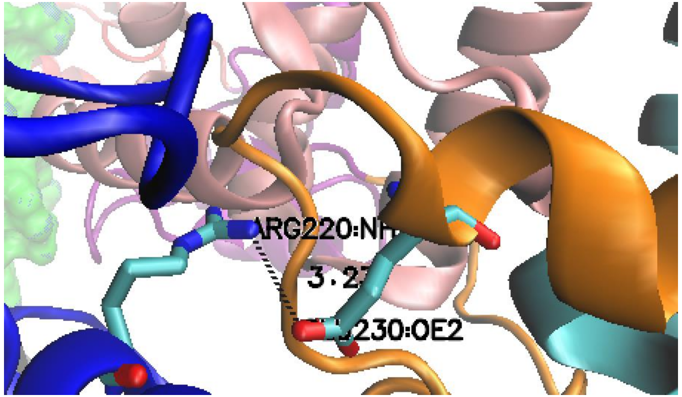
Electrostatic interaction with DNA

Conserved across metazoans

Closed PIC formation: TFIIE joins the complex

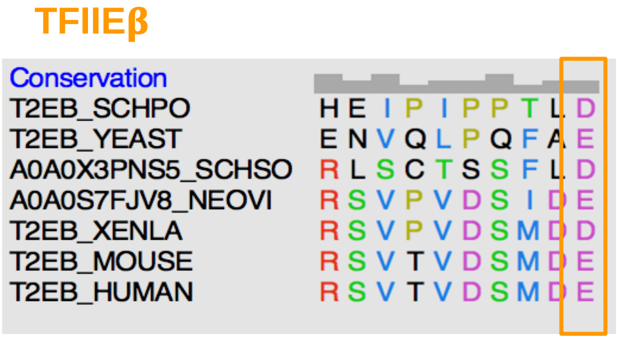
The TFIIEβ anchores to the complex through RPB1 binding





Electrostatic interaction with Rpb1

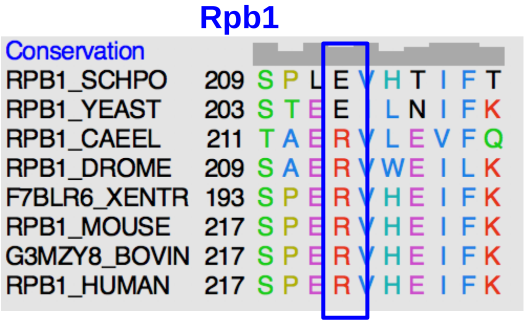
Conserved across metazoans



TFIIEβ

Conservation										
T2EB_SCHPO	H	E	I	P	I	P	P	T	L	D
T2EB_YEAST	E	N	V	Q	L	P	Q	F	A	E
A0A0X3PNS5_SCHSO	R	L	S	C	T	S	S	F	L	D
A0A0S7FJV8_NEOVI	R	S	V	P	V	D	S	I	D	E
T2EB_XENLA	R	S	V	P	V	D	S	M	D	D
T2EB_MOUSE	R	S	V	T	V	D	S	M	D	E
T2EB_HUMAN	R	S	V	T	V	D	S	M	D	E

Glu230



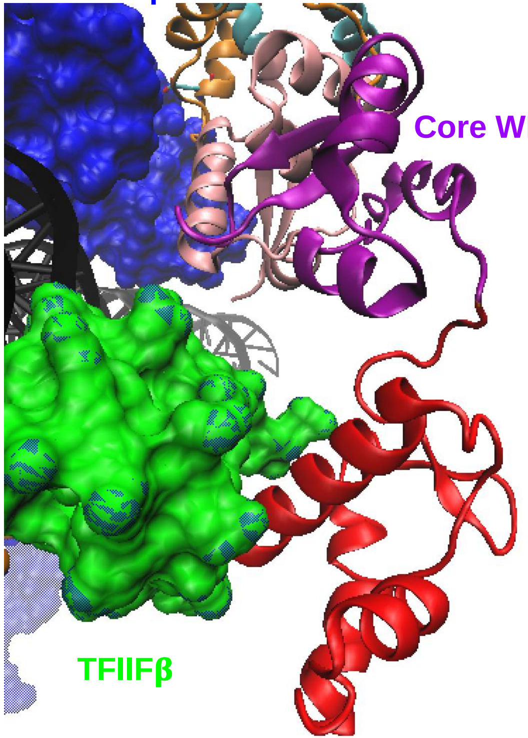
Rpb1

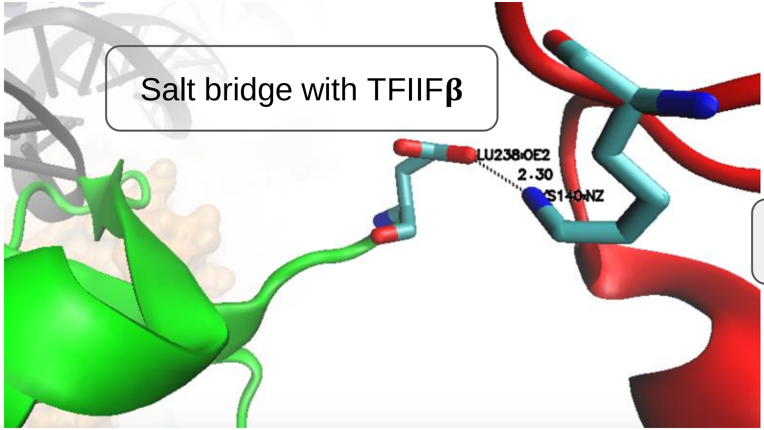
Conservation											
RPB1_SCHPO	209	S	P	L	E	V	H	T	I	F	T
RPB1_YEAST	203	S	T	E	E	I	L	N	I	F	K
RPB1_CAEEL	211	T	A	E	R	V	L	E	V	F	Q
RPB1_DROME	209	S	A	E	R	V	W	E	I	L	K
F7BLR6_XENTR	193	S	P	E	R	V	H	E	I	F	K
RPB1_MOUSE	217	S	P	E	R	V	H	E	I	F	K
G3MZY8_BOVIN	217	S	P	E	R	V	H	E	I	F	K
RPB1_HUMAN	217	S	P	E	R	V	H	E	I	F	K

Arg220

Closed PIC formation: TFIIE joins the complex

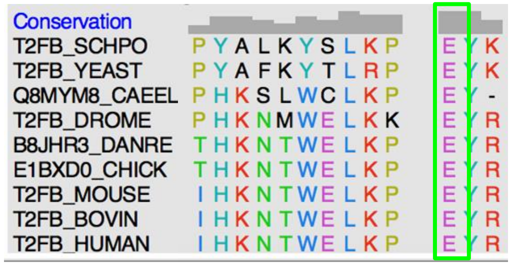
The N-term WH domain of TFIIEβ stabilizes the complex through TFIIFβ binding





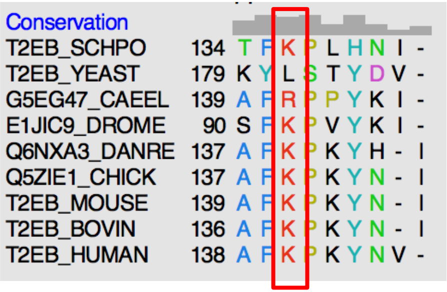
Salt bridge with TFIIFβ

Conserved across metazoans



Conservation													
T2FB_SCHPO	P	Y	A	L	K	Y	S	L	K	P	E	Y	K
T2FB_YEAST	P	Y	A	F	K	Y	T	L	R	P	E	Y	K
Q8MYM8_CAEEL	P	H	K	S	L	W	C	L	K	P	E	Y	-
T2FB_DROME	P	H	K	N	M	W	E	L	K	K	E	Y	R
B8JHR3_DANRE	T	H	K	N	T	W	E	L	K	P	E	Y	R
E1BXD0_CHICK	T	H	K	N	T	W	E	L	K	P	E	Y	R
T2FB_MOUSE	I	H	K	N	T	W	E	L	K	P	E	Y	R
T2FB_BOVIN	I	H	K	N	T	W	E	L	K	P	E	Y	R
T2FB_HUMAN	I	H	K	N	T	W	E	L	K	P	E	Y	R

Glu 238

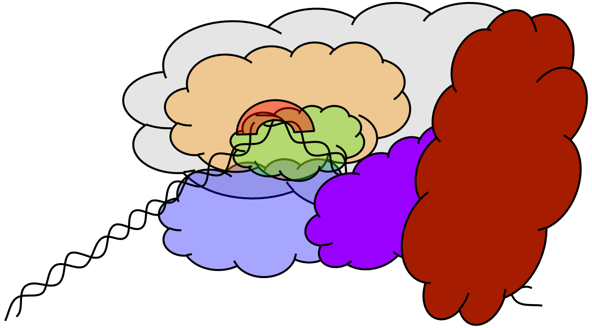


Conservation										
T2EB_SCHPO	134	T	F	K	P	L	H	N	I	-
T2EB_YEAST	179	K	Y	L	S	T	Y	D	V	-
G5EG47_CAEEL	139	A	F	R	P	P	Y	K	I	-
E1JIC9_DROME	90	S	F	K	P	V	Y	K	I	-
Q6NXA3_DANRE	137	A	F	K	P	K	Y	H	-	I
Q5ZIE1_CHICK	137	A	F	K	P	K	Y	N	-	I
T2EB_MOUSE	139	A	F	K	P	K	Y	N	-	I
T2EB_BOVIN	136	A	F	K	P	K	Y	N	-	I
T2EB_HUMAN	138	A	F	K	P	K	Y	N	V	-

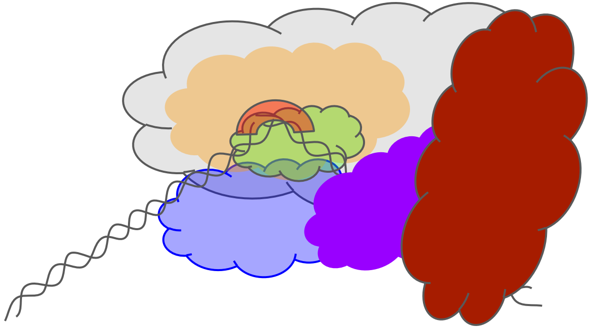
Lys 140

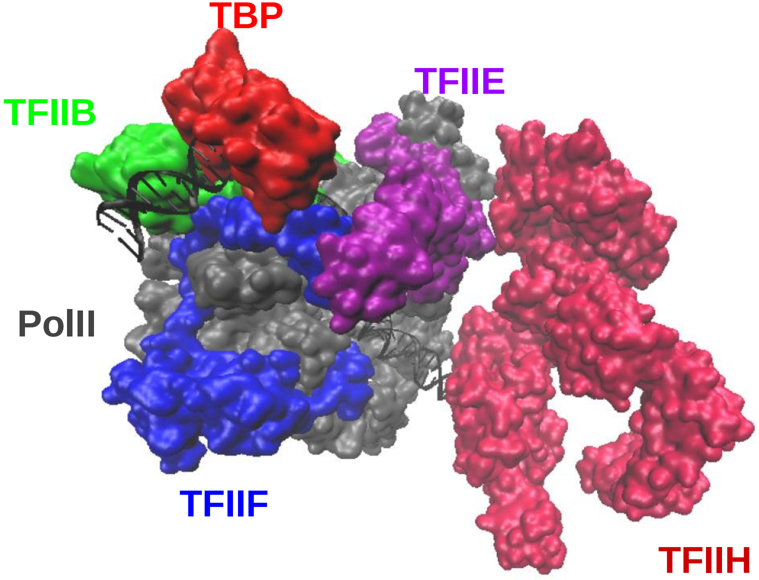
Closed PIC formation:

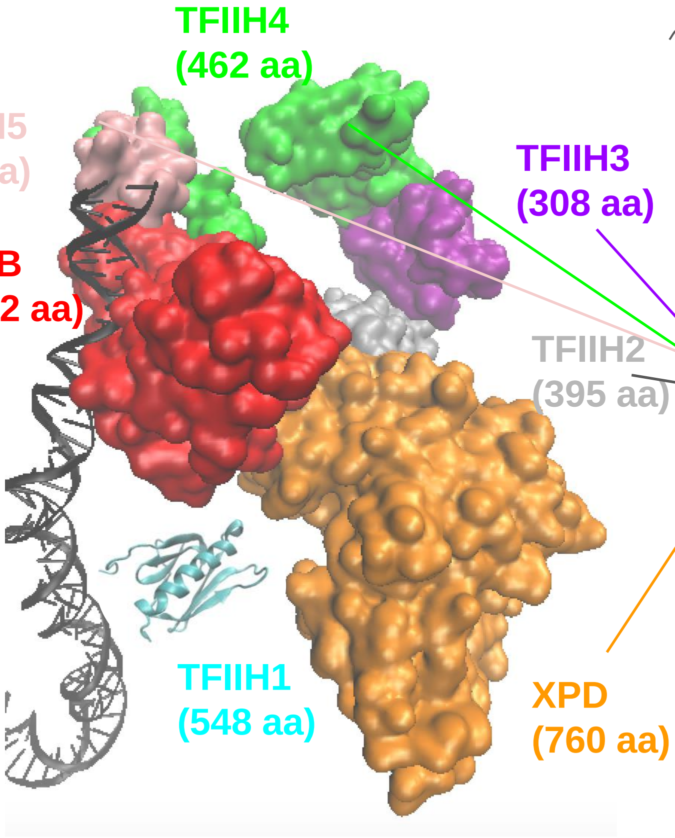
TFIIH joins the complex



Closed PIC formation: TFIIH joins the complex

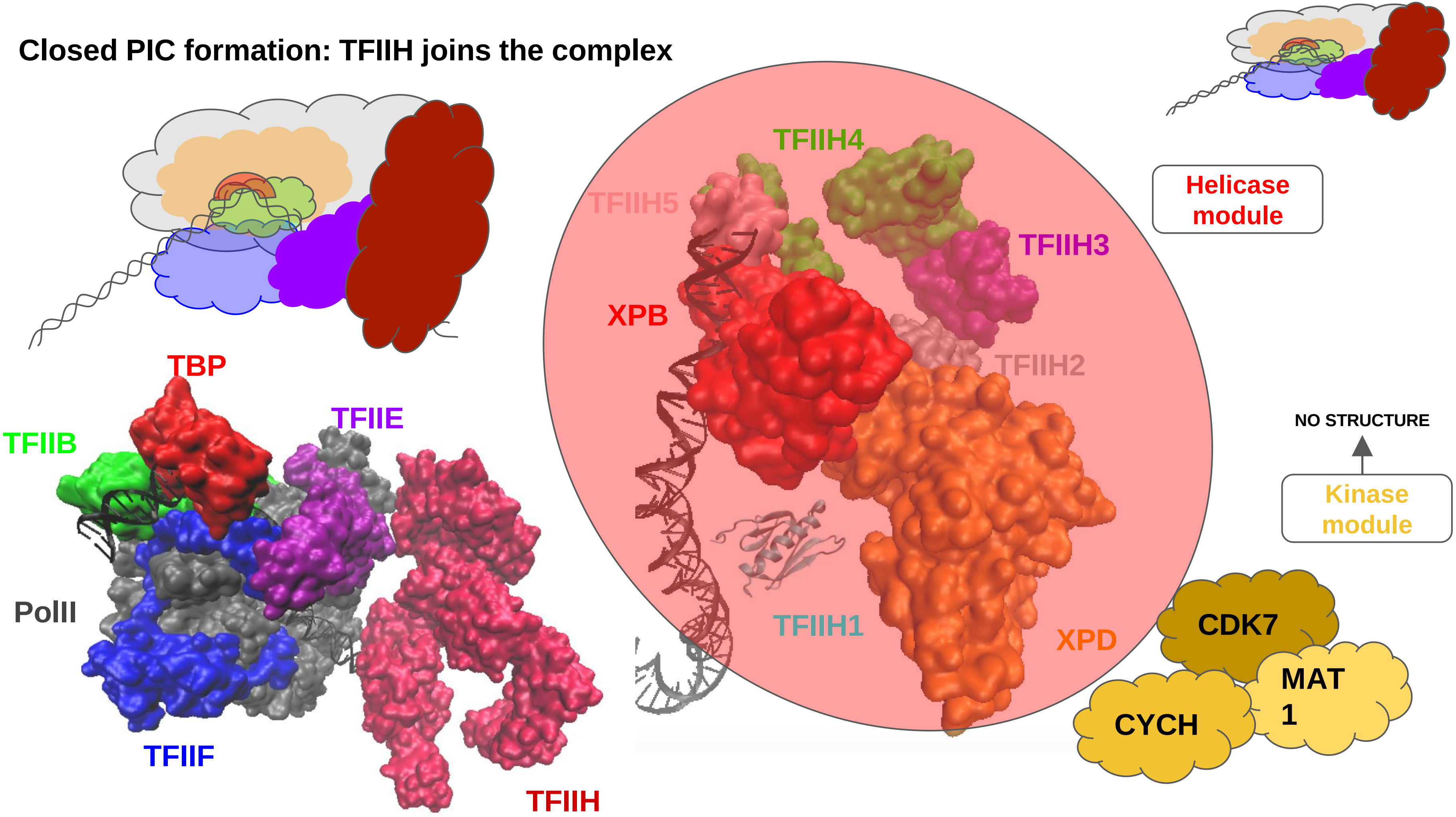






Poor structural characterization

Closed PIC formation: TFIIH joins the complex



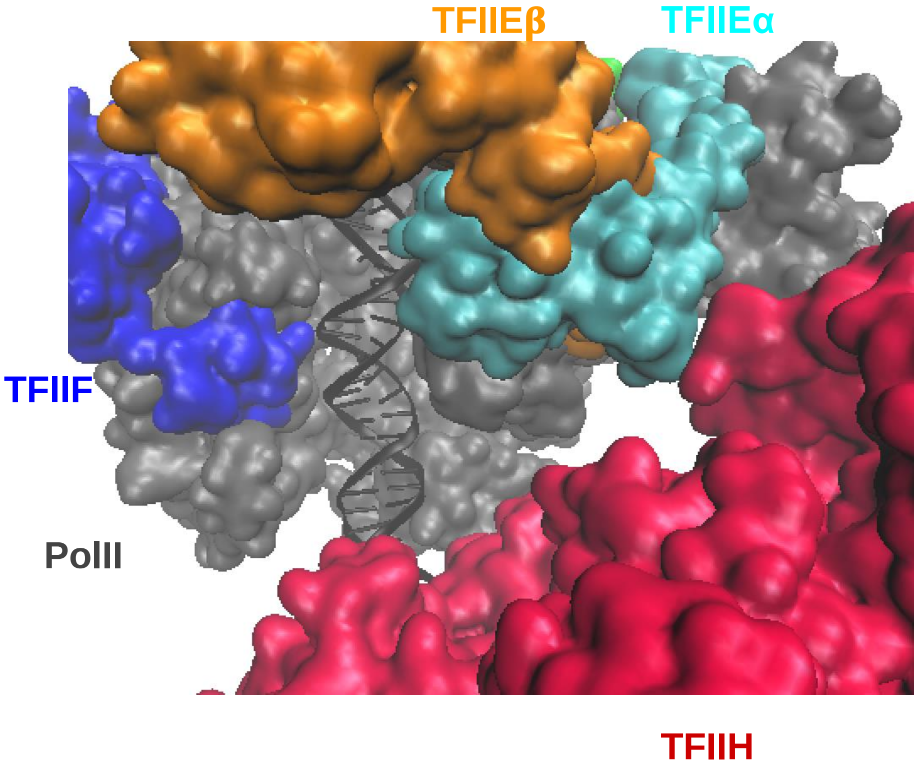
Helicase module

NO STRUCTURE

Kinase module

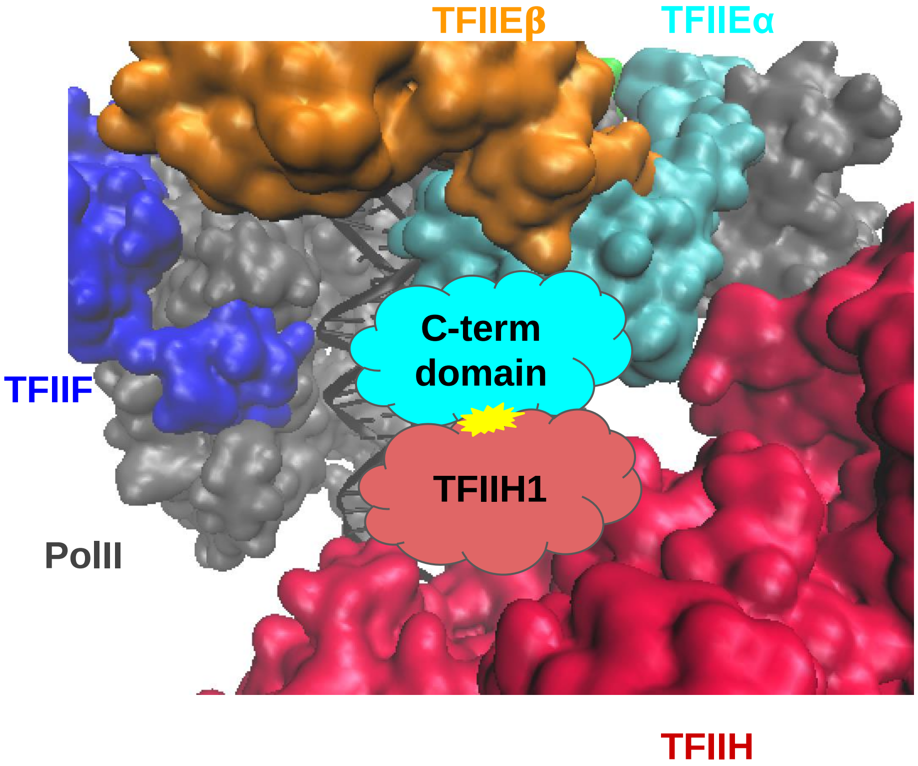
Closed PIC formation: TFIIH joins the complex

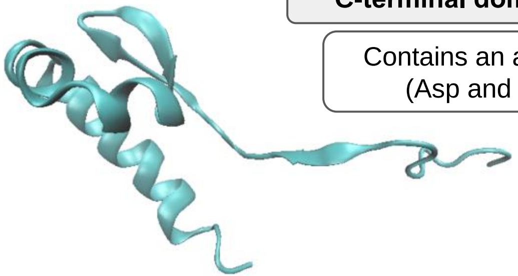
TFIIH assembles to the PIC through the C-term domain of TFIIEα



Closed PIC formation: TFIIH joins the complex

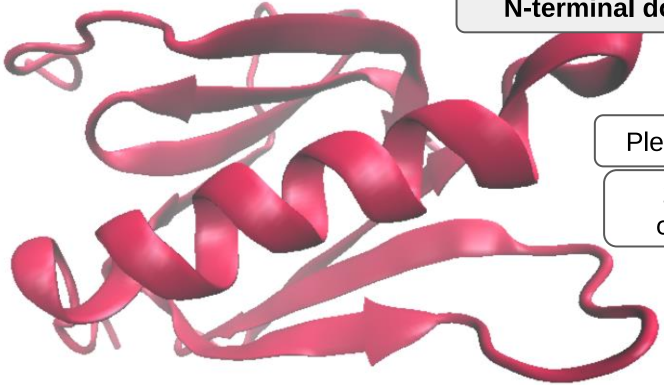
TFIIH assembles to the PIC through the C-term domain of TFIIEα





C-terminal domain of TFIIEα

Contains an acidic region (Asp and Glu rich)



N-terminal domain of TFIIH1

Pleckstrin-homology

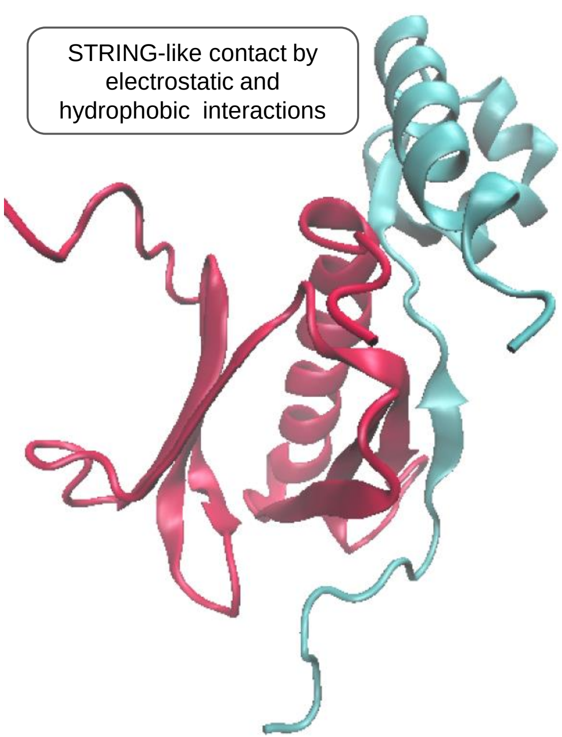
2 **β**-meanders capping α-helix

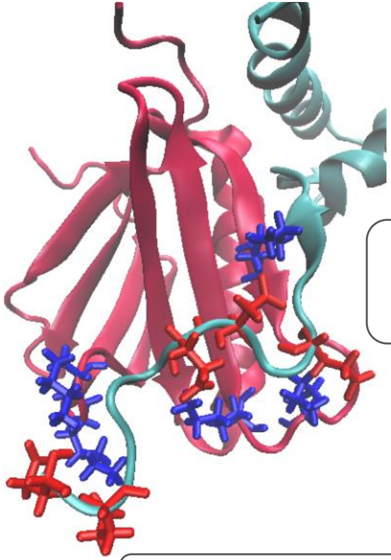
We don’t know the architecture of these domains

Closed PIC formation: TFIIH joins the complex

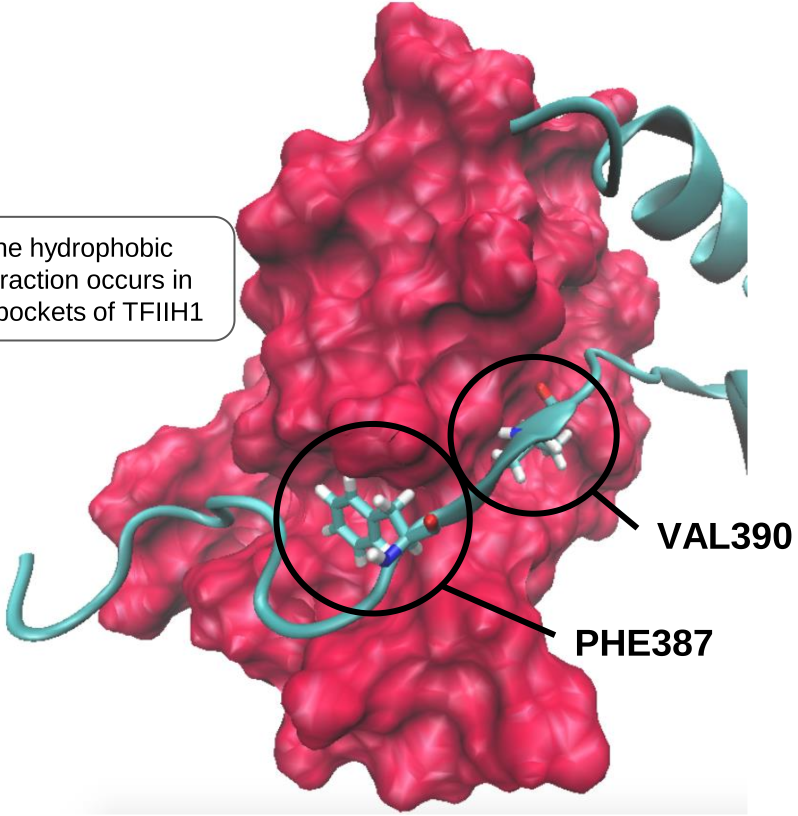
TFIIH assembles to the PIC through the C-term domain of TFIIEα

STRING-like contact by electrostatic and hydrophobic interactions





The hydrophobic interaction occurs in two pockets of TFIIH1



Asp and Glu interact with many Lys of TFIIEα

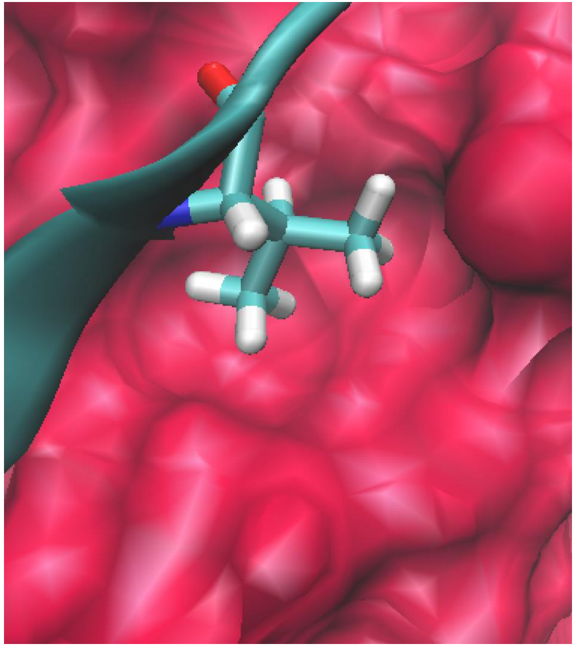
BIOCHEMICAL STUDIES

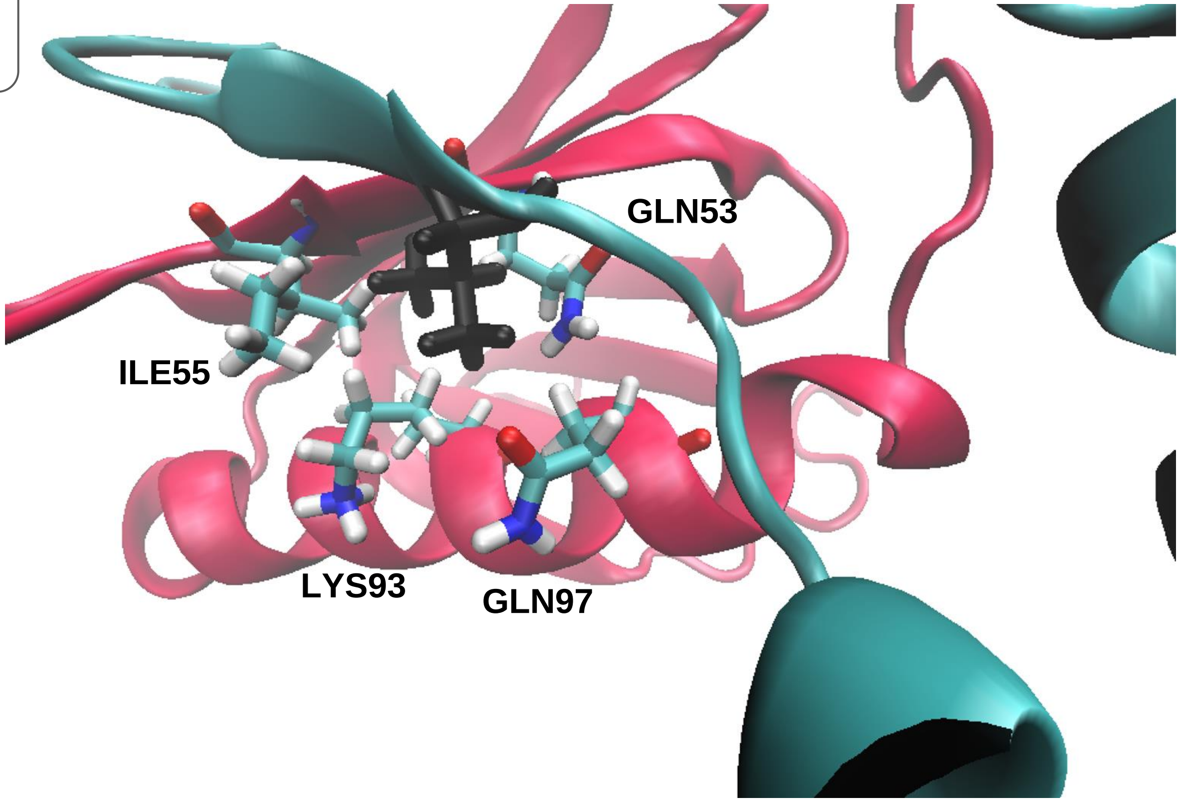
Phe387 and Val390 are essential for the interaction in humans

Closed PIC formation: TFIIH joins the complex

TFIIH assembles to the PIC through the C-term domain of TFIIEα

The interaction between Val390 of TFIIEα with TFIIH1 as an example

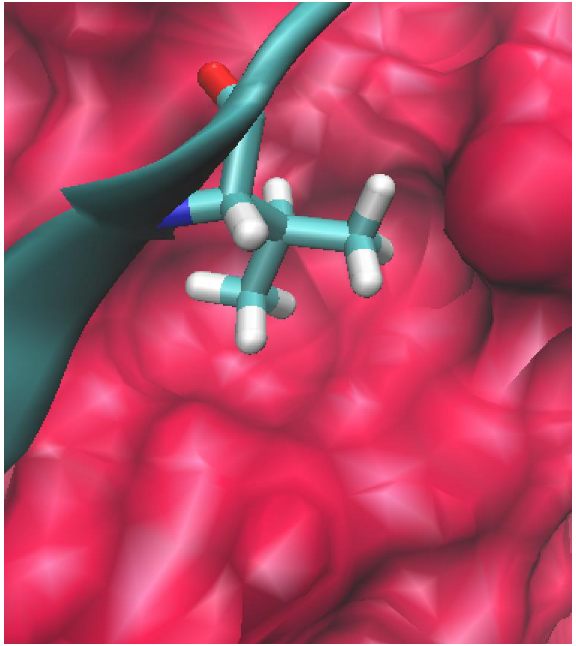


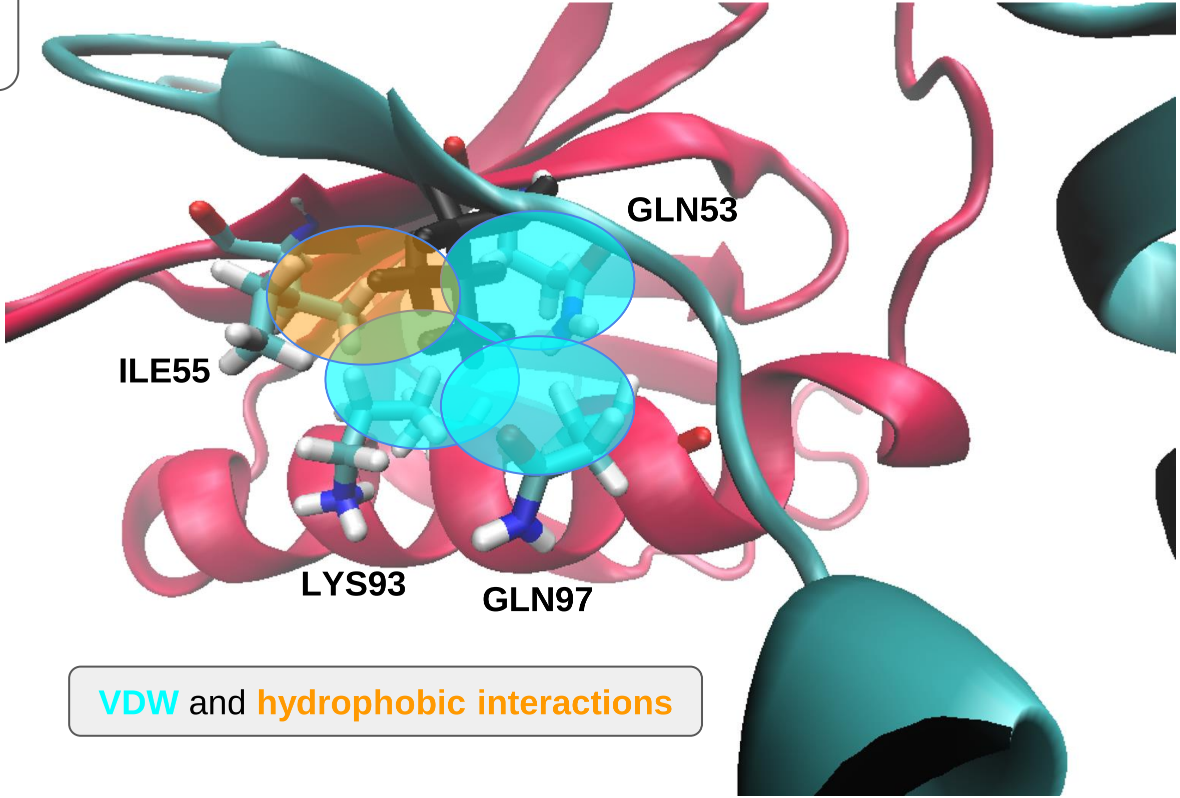


Closed PIC formation: TFIIH joins the complex

TFIIH assembles to the PIC through the C-term domain of TFIIEα

The interaction between Val390 of TFIIEα with TFIIH1 as an example

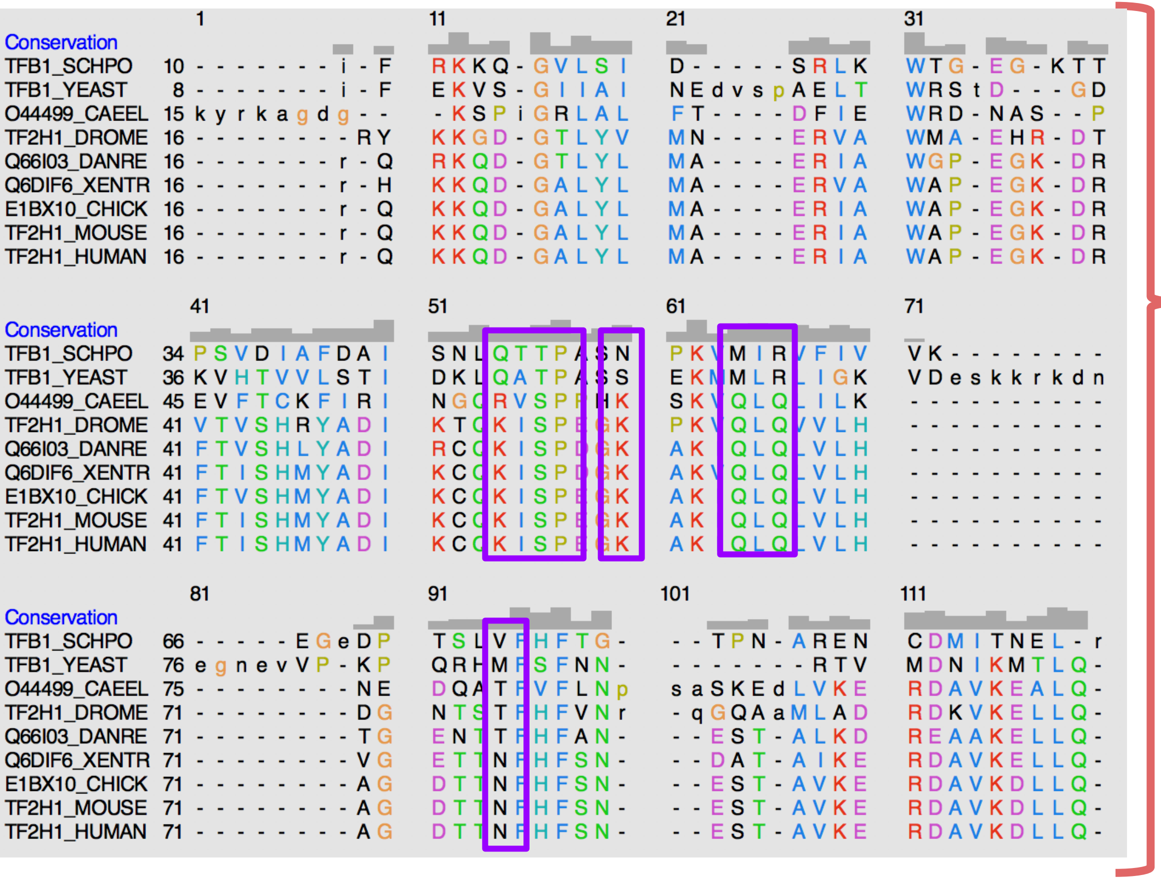




VDW and hydrophobic interactions

Closed PIC formation: TFIIH joins the complex

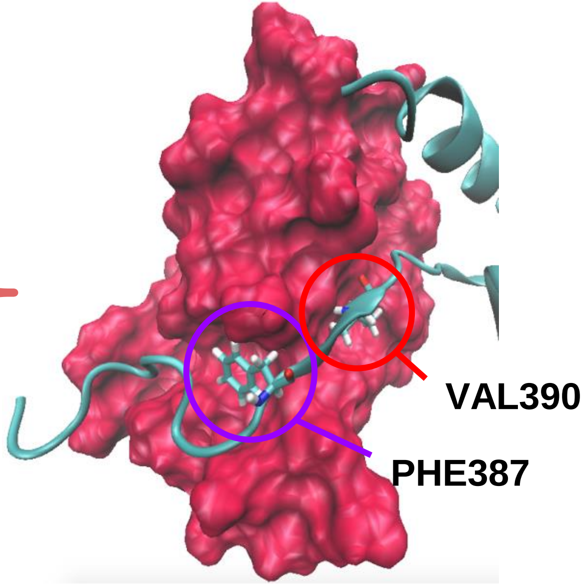
TFIIH assembles to the PIC through the C-term domain of TFIIEα



		1	11	21	31
Conservation					
TFB1_SCHPO	10	-------i-F	RKKQ-GVLSI	D-----SRLK	WTG-EG-KTT
TFB1_YEAST	8	-------i-F	EKVS-GIIAI	NEdvspAELT	WRStD---GD
O44499_CAEEL	15	kyrkagdg--	-KSPiGRLAL	FT----DFIE	WRD-NAS--P
TF2H1_DROME	16	--------RY	KKGD-GTLYV	MN----ERVA	WMA-EHR-DT
Q66I03_DANRE	16	-------r-Q	RKQD-GTLYL	MA----ERIA	WGP-EGK-DR
Q6DIF6_XENTR	16	-------r-H	KKQD-GALYL	MA----ERVA	WAP-EGK-DR
E1BX10_CHICK	16	-------r-Q	KKQD-GALYL	MA----ERIA	WAP-EGK-DR
TF2H1_MOUSE	16	-------r-Q	KKQD-GALYL	MA----ERIA	WAP-EGK-DR
TF2H1_HUMAN	16	-------r-Q	KKQD-GALYL	MA----ERIA	WAP-EGK-DR

		41	51	61	71
Conservation					
TFB1_SCHPO	34	PSVDIAFDAI	SNLQTTPASN	PKVMIRVFIV	VK--------
TFB1_YEAST	36	KVHTVVLSTI	DKLQATPASS	EKMMLRLIGK	VDeskkrkdn
O44499_CAEEL	45	EVFTCKFIRI	NGQRVSPPHK	SKVQLQLILK	----------
TF2H1_DROME	41	VTVSHRYADI	KTQKISPEGK	PKVQLQVVLH	----------
Q66I03_DANRE	41	FTVSHLYADI	RCQKISPDGK	AK[illegible]QLQLVLH	----------
Q6DIF6_XENTR	41	FTISHMYADI	KCQKISPDGK	AKVQLQLVLH	----------
E1BX10_CHICK	41	FTVSHMYADI	KCQKISPEGK	AK[illegible]QLQLVLH	----------
TF2H1_MOUSE	41	FTISHMYADI	KCQKISPEGK	AK[illegible]QLQLVLH	----------
TF2H1_HUMAN	41	FTISHMYADI	KCQKISPEGK	AK[illegible]QLQLVLH	----------

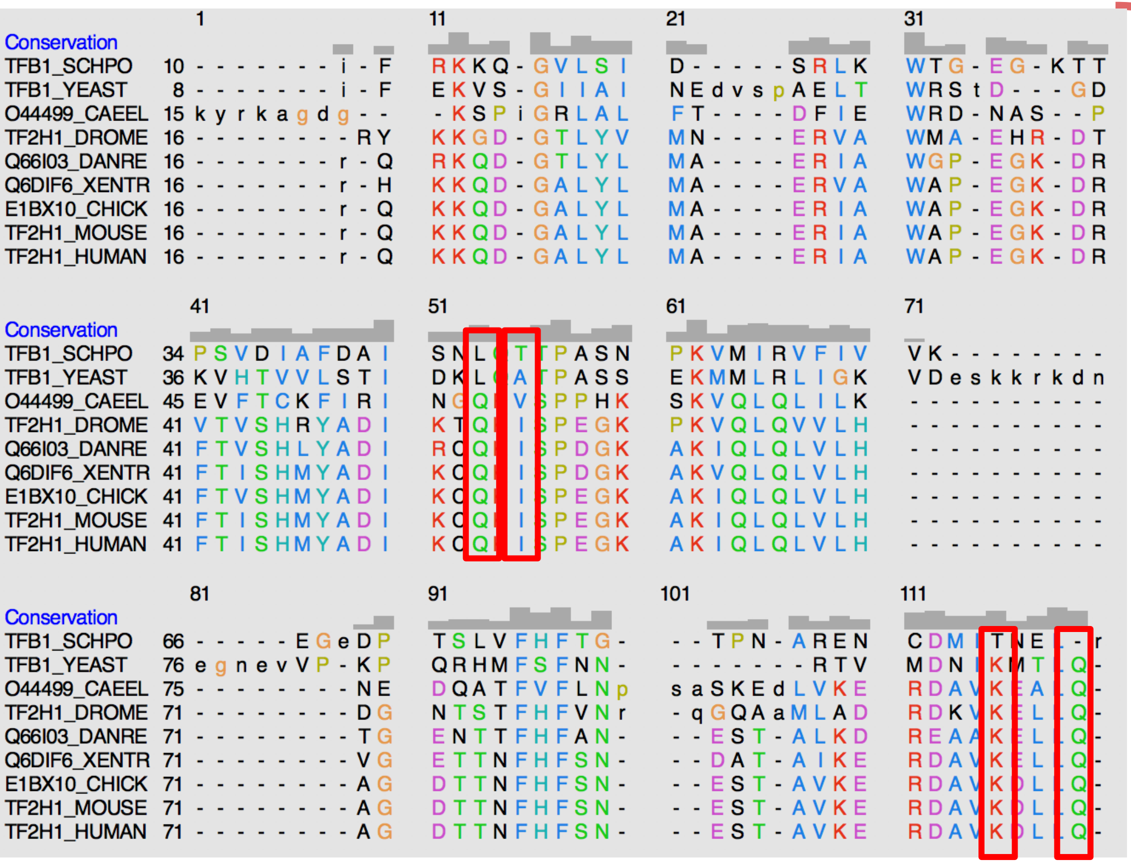
		81	91	101	111
Conservation					
TFB1_SCHPO	66	-----EGeDP	TSLVFHFTG-	--TPN-AREN	CDMITNEL-r
TFB1_YEAST	76	egnevVP-KP	QRHMFSFNN-	-------RTV	MDNIKMTLQ-
O44499_CAEEL	75	--------NE	DQATFVFLNp	saSKEdLVKE	RDAVKEALQ-
TF2H1_DROME	71	--------DG	NTSTFHFVNr	-qGQAaMLAD	RDKVKELLQ-
Q66I03_DANRE	71	--------TG	ENTTFHFAN-	--EST-ALKD	REAAKELLQ-
Q6DIF6_XENTR	71	--------VG	ETTNFHFSN-	--DAT-AIKE	RDAVKELLQ-
E1BX10_CHICK	71	--------AG	DTTNFHFSN-	--EST-AVKE	RDAVKDLLQ-
TF2H1_MOUSE	71	--------AG	DTTNFHFSN-	--EST-AVKE	RDAVKDLLQ-
TF2H1_HUMAN	71	--------AG	DTTNFHFSN-	--EST-AVKE	RDAVKDLLQ-



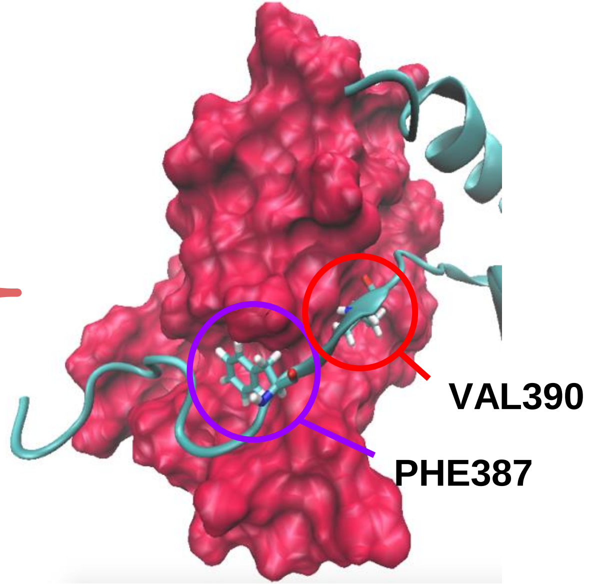
The hydrophobic pocket for PHE387 is conserved across metazoans

Closed PIC formation: TFIIH joins the complex

TFIIH assembles to the PIC through the C-term domain of TFIIEα



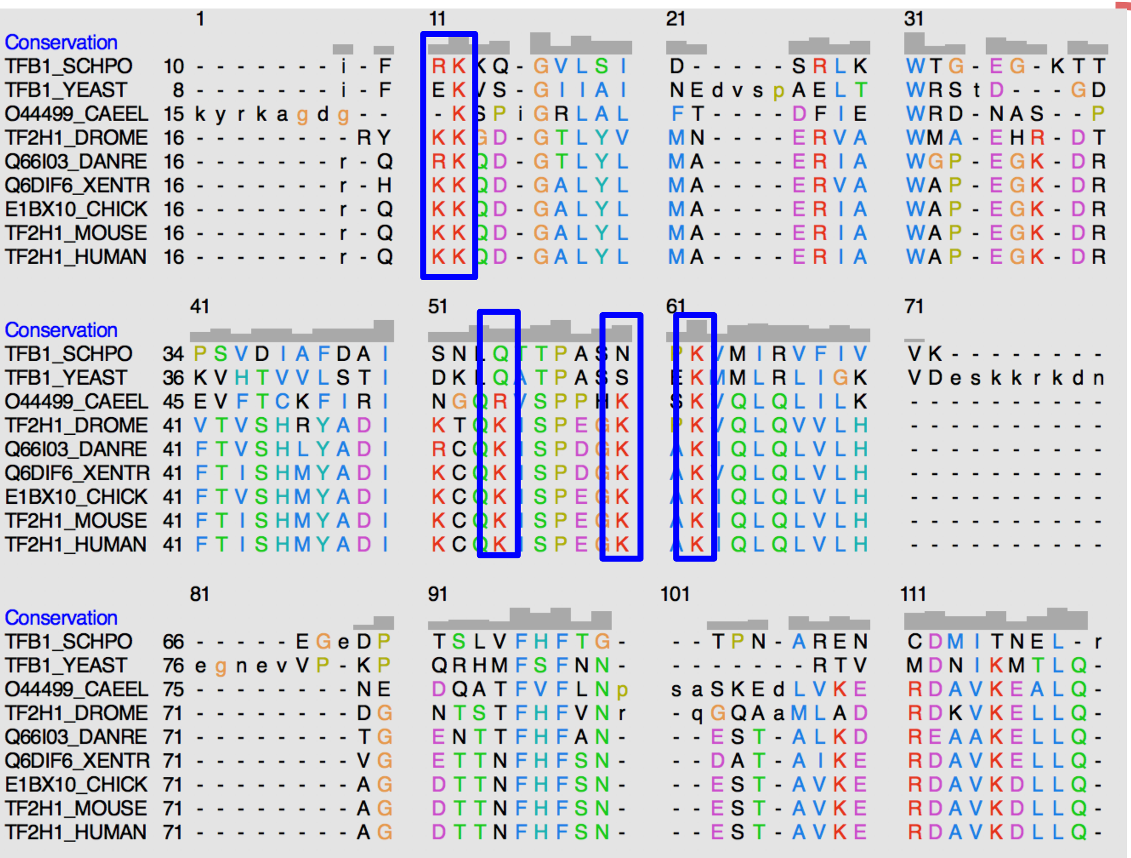
```
                    1                     11                    21                    31
Conservation
TFB1_SCHPO    10 - - - - - - - i - F   R K K Q - G V L S I   D - - - - - S R L K   W T G - E G - K T T
TFB1_YEAST     8 - - - - - - - i - F   E K V S - G I I A I   N E d v s p A E L T   W R S t D - - - G D
O44499_CAEEL  15 k y r k a g d g - -   - K S P i G R L A L   F T - - - - D F I E   W R D - N A S - - P
TF2H1_DROME   16 - - - - - - - - R Y   K K G D - G T L Y V   M N - - - - E R V A   W M A - E H R - D T
Q66I03_DANRE  16 - - - - - - - r - Q   R K Q D - G T L Y L   M A - - - - E R I A   W G P - E G K - D R
Q6DIF6_XENTR  16 - - - - - - - r - H   K K Q D - G A L Y L   M A - - - - E R V A   W A P - E G K - D R
E1BX10_CHICK  16 - - - - - - - r - Q   K K Q D - G A L Y L   M A - - - - E R I A   W A P - E G K - D R
TF2H1_MOUSE   16 - - - - - - - r - Q   K K Q D - G A L Y L   M A - - - - E R I A   W A P - E G K - D R
TF2H1_HUMAN   16 - - - - - - - r - Q   K K Q D - G A L Y L   M A - - - - E R I A   W A P - E G K - D R

                    41                    51                    61                    71
Conservation
TFB1_SCHPO    34 P S V D I A F D A I   S N L Q T T P A S N   P K V M I R V F I V   V K - - - - - - - -
TFB1_YEAST    36 K V H T V V L S T I   D K L Q A T P A S S   E K M M L R L I G K   V D e s k k r k d n
O44499_CAEEL  45 E V F T C K F I R I   N G Q L V S P P H K   S K V Q L Q L I L K   - - - - - - - - - -
TF2H1_DROME   41 V T V S H R Y A D I   K T Q L I S P E G K   P K V Q L Q V V L H   - - - - - - - - - -
Q66I03_DANRE  41 F T V S H L Y A D I   R C Q L I S P D G K   A K I Q L Q L V L H   - - - - - - - - - -
Q6DIF6_XENTR  41 F T I S H M Y A D I   K C Q L I S P D G K   A K V Q L Q L V L H   - - - - - - - - - -
E1BX10_CHICK  41 F T V S H M Y A D I   K C Q L I S P E G K   A K I Q L Q L V L H   - - - - - - - - - -
TF2H1_MOUSE   41 F T I S H M Y A D I   K C Q L I S P E G K   A K I Q L Q L V L H   - - - - - - - - - -
TF2H1_HUMAN   41 F T I S H M Y A D I   K C Q L I S P E G K   A K I Q L Q L V L H   - - - - - - - - - -

                    81                    91                    101                   111
Conservation
TFB1_SCHPO    66 - - - - - E G e D P   T S L V F H F T G -   - - T P N - A R E N   C D M I T N E L - r
TFB1_YEAST    76 e g n e v V P - K P   Q R H M F S F N N -   - - - - - - - R T V   M D N I K M T L Q -
O44499_CAEEL  75 - - - - - - - - N E   D Q A T F V F L N p   s a S K E d L V K E   R D A V K E A L Q -
TF2H1_DROME   71 - - - - - - - - D G   N T S T F H F V N r   - q G Q A a M L A D   R D K V K E L L Q -
Q66I03_DANRE  71 - - - - - - - - T G   E N T T F H F A N -   - - E S T - A L K D   R E A A K E L L Q -
Q6DIF6_XENTR  71 - - - - - - - - V G   E T T N F H F S N -   - - D A T - A I K E   R D A V K E L L Q -
E1BX10_CHICK  71 - - - - - - - - A G   D T T N F H F S N -   - - E S T - A V K E   R D A V K D L L Q -
TF2H1_MOUSE   71 - - - - - - - - A G   D T T N F H F S N -   - - E S T - A V K E   R D A V K D L L Q -
TF2H1_HUMAN   71 - - - - - - - - A G   D T T N F H F S N -   - - E S T - A V K E   R D A V K D L L Q -
```



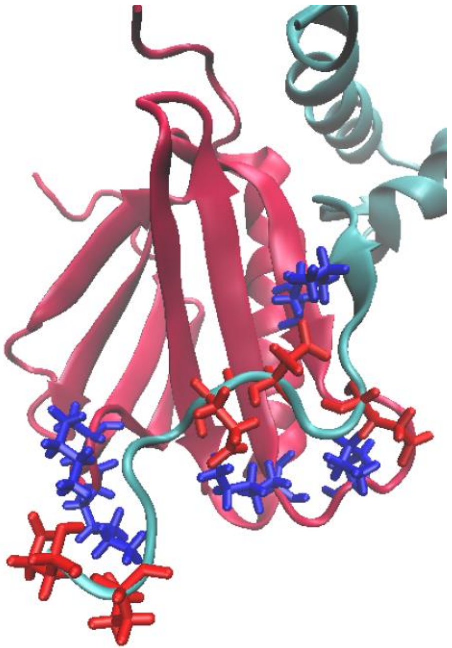
The hydrophobic pocket for VAL390 is conserved across metazoans

Closed PIC formation: TFIIH joins the complex

TFIIH assembles to the PIC through the C-term domain of TFIIEα



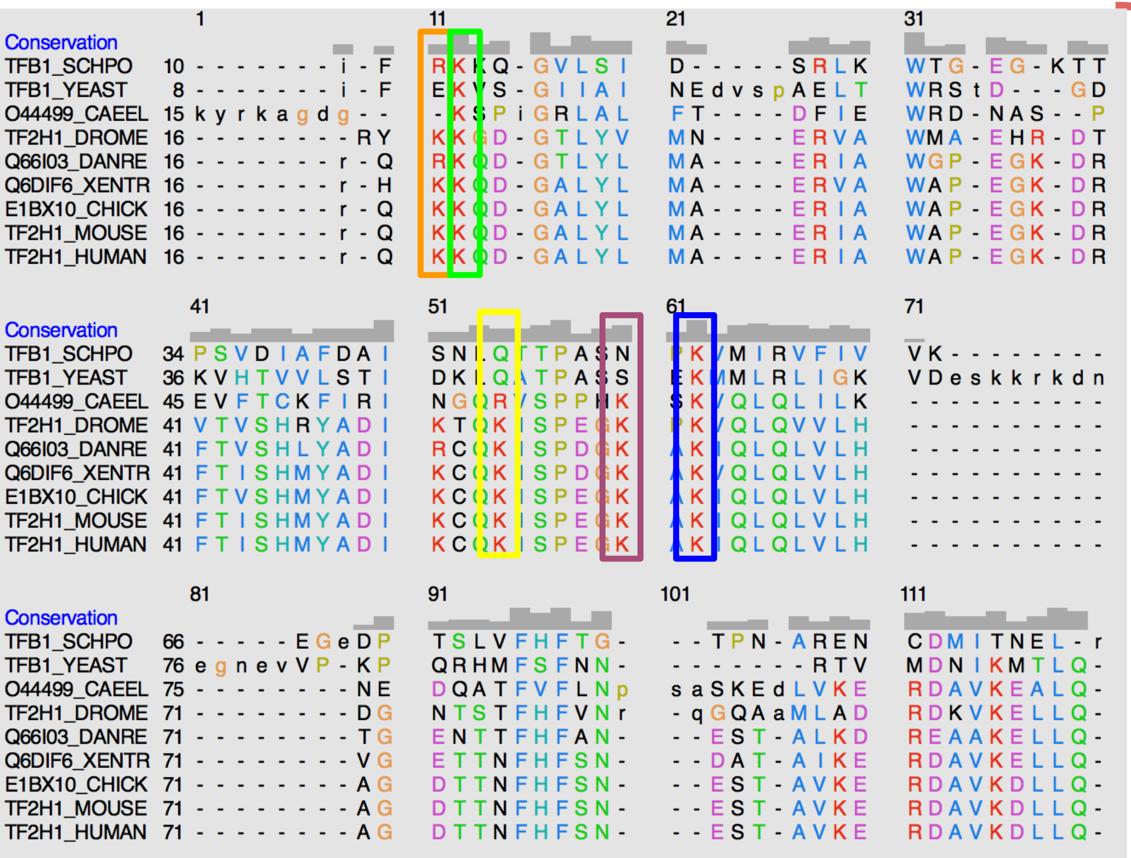
```
                   1          11         21         31
Conservation
TFB1_SCHPO    10 -------i-F RKKQ-GVLSI D-----SRLK WTG-EG-KTT
TFB1_YEAST     8 -------i-F EKVS-GIIAI NEdvspAELT WRStD---GD
O44499_CAEEL  15 kyrkagdg-- -KSPiGRLAL FT----DFIE WRD-NAS--P
TF2H1_DROME   16 --------RY KKGD-GTLYV MN----ERVA WMA-EHR-DT
Q66I03_DANRE  16 -------r-Q RKQD-GTLYL MA----ERIA WGP-EGK-DR
Q6DIF6_XENTR  16 -------r-H KKQD-GALYL MA----ERVA WAP-EGK-DR
E1BX10_CHICK  16 -------r-Q KKQD-GALYL MA----ERIA WAP-EGK-DR
TF2H1_MOUSE   16 -------r-Q KKQD-GALYL MA----ERIA WAP-EGK-DR
TF2H1_HUMAN   16 -------r-Q KKQD-GALYL MA----ERIA WAP-EGK-DR

                   41         51         61         71
Conservation
TFB1_SCHPO    34 PSVDIAFDAI SNLQTTPASN PKVMIRVFIV VK--------
TFB1_YEAST    36 KVHTVVLSTI DKLQATPASS EKMMLRLIGK VDeskkrkdn
O44499_CAEEL  45 EVFTCKFIRI NGQRVSPPHK SKVQLQLILK ----------
TF2H1_DROME   41 VTVSHRYADI KTQKISPEGK PKVQLQVVLH ----------
Q66I03_DANRE  41 FTVSHLYADI RCQKISPDGK AKIQLQLVLH ----------
Q6DIF6_XENTR  41 FTISHMYADI KCQKISPDGK AKVQLQLVLH ----------
E1BX10_CHICK  41 FTVSHMYADI KCQKISPEGK AKIQLQLVLH ----------
TF2H1_MOUSE   41 FTISHMYADI KCQKISPEGK AKIQLQLVLH ----------
TF2H1_HUMAN   41 FTISHMYADI KCQKISPEGK AKIQLQLVLH ----------

                   81         91         101        111
Conservation
TFB1_SCHPO    66 -----EGeDP TSLVFHFTG- --TPN-AREN CDMITNEL-r
TFB1_YEAST    76 egnevVP-KP QRHMFSFNN- -------RTV MDNIKMTLQ-
O44499_CAEEL  75 --------NE DQATFVFLNp saSKEdLVKE RDAVKEALQ-
TF2H1_DROME   71 --------DG NTSTFHFVNr -qGQAaMLAD RDKVKELLQ-
Q66I03_DANRE  71 --------TG ENTTFHFAN- --EST-ALKD REAAKELLQ-
Q6DIF6_XENTR  71 --------VG ETTNFHFSN- --DAT-AIKE RDAVKELLQ-
E1BX10_CHICK  71 --------AG DTTNFHFSN- --EST-AVKE RDAVKDLLQ-
TF2H1_MOUSE   71 --------AG DTTNFHFSN- --EST-AVKE RDAVKDLLQ-
TF2H1_HUMAN   71 --------AG DTTNFHFSN- --EST-AVKE RDAVKDLLQ-
```



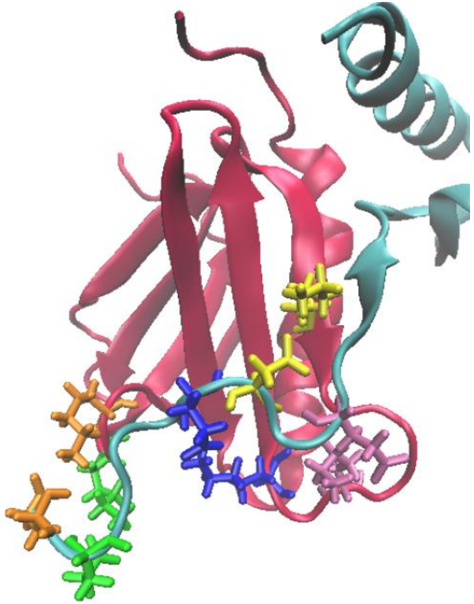
Most lysines are conserved across metazoans

Closed PIC formation: TFIIH joins the complex

TFIIH assembles to the PIC through the C-term domain of TFIIEα



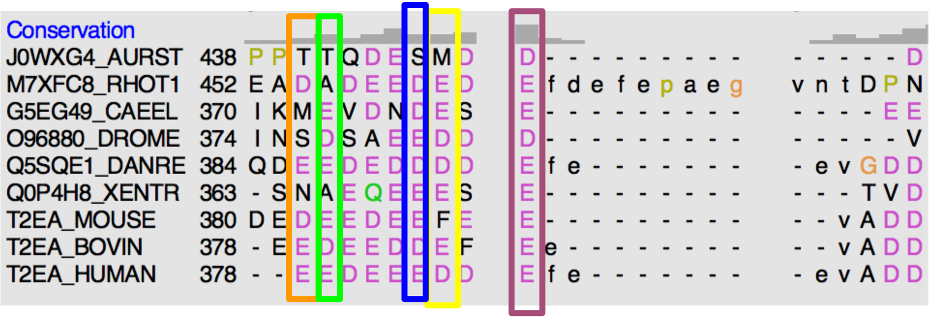
```
                    1                   11                  21                  31
Conservation
TFB1_SCHPO   10 - - - - - - - i - F   R K K Q - G V L S I   D - - - - - S R L K   W T G - E G - K T T
TFB1_YEAST    8 - - - - - - - i - F   E K V S - G I I A I   N E d v s p A E L T   W R S t D - - - G D
O44499_CAEEL 15 k y r k a g d g - -   - K S P i G R L A L   F T - - - - D F I E   W R D - N A S - - P
TF2H1_DROME  16 - - - - - - - - R Y   K K G D - G T L Y V   M N - - - - E R V A   W M A - E H R - D T
Q66I03_DANRE 16 - - - - - - - r - Q   R K Q D - G T L Y L   M A - - - - E R I A   W G P - E G K - D R
Q6DIF6_XENTR 16 - - - - - - - r - H   K K Q D - G A L Y L   M A - - - - E R V A   W A P - E G K - D R
E1BX10_CHICK 16 - - - - - - - r - Q   K K Q D - G A L Y L   M A - - - - E R I A   W A P - E G K - D R
TF2H1_MOUSE  16 - - - - - - - r - Q   K K Q D - G A L Y L   M A - - - - E R I A   W A P - E G K - D R
TF2H1_HUMAN  16 - - - - - - - r - Q   K K Q D - G A L Y L   M A - - - - E R I A   W A P - E G K - D R

                    41                  51                  61                  71
Conservation
TFB1_SCHPO   34 P S V D I A F D A I   S N L Q T T P A S N   P K V M I R V F I V   V K - - - - - - - -
TFB1_YEAST   36 K V H T V V L S T I   D K L Q A T P A S S   E K M M L R L I G K   V D e s k k r k d n
O44499_CAEEL 45 E V F T C K F I R I   N G Q R V S P P H K   S K V Q L Q L I L K   - - - - - - - - - -
TF2H1_DROME  41 V T V S H R Y A D I   K T Q K I S P E G K   P K V Q L Q V V L H   - - - - - - - - - -
Q66I03_DANRE 41 F T V S H L Y A D I   R C Q K I S P D G K   A K I Q L Q L V L H   - - - - - - - - - -
Q6DIF6_XENTR 41 F T I S H M Y A D I   K C Q K I S P D G K   A K V Q L Q L V L H   - - - - - - - - - -
E1BX10_CHICK 41 F T V S H M Y A D I   K C Q K I S P E G K   A K I Q L Q L V L H   - - - - - - - - - -
TF2H1_MOUSE  41 F T I S H M Y A D I   K C Q K I S P E G K   A K I Q L Q L V L H   - - - - - - - - - -
TF2H1_HUMAN  41 F T I S H M Y A D I   K C Q K I S P E G K   A K I Q L Q L V L H   - - - - - - - - - -

                    81                  91                  101                 111
Conservation
TFB1_SCHPO   66 - - - - - E G e D P   T S L V F H F T G -   - - T P N - A R E N   C D M I T N E L - r
TFB1_YEAST   76 e g n e v V P - K P   Q R H M F S F N N -   - - - - - - - R T V   M D N I K M T L Q -
O44499_CAEEL 75 - - - - - - - - N E   D Q A T F V F L N p   s a S K E d L V K E   R D A V K E A L Q -
TF2H1_DROME  71 - - - - - - - - D G   N T S T F H F V N r   - q G Q A a M L A D   R D K V K E L L Q -
Q66I03_DANRE 71 - - - - - - - - T G   E N T T F H F A N -   - - E S T - A L K D   R E A A K E L L Q -
Q6DIF6_XENTR 71 - - - - - - - - V G   E T T N F H F S N -   - - D A T - A I K E   R D A V K E L L Q -
E1BX10_CHICK 71 - - - - - - - - A G   D T T N F H F S N -   - - E S T - A V K E   R D A V K D L L Q -
TF2H1_MOUSE  71 - - - - - - - - A G   D T T N F H F S N -   - - E S T - A V K E   R D A V K D L L Q -
TF2H1_HUMAN  71 - - - - - - - - A G   D T T N F H F S N -   - - E S T - A V K E   R D A V K D L L Q -
```



Most lysines are conserved across metazoans

Closed PIC formation: TFIIH joins the complex

TFIIH assembles to the PIC through the C-term domain of TFIIEα

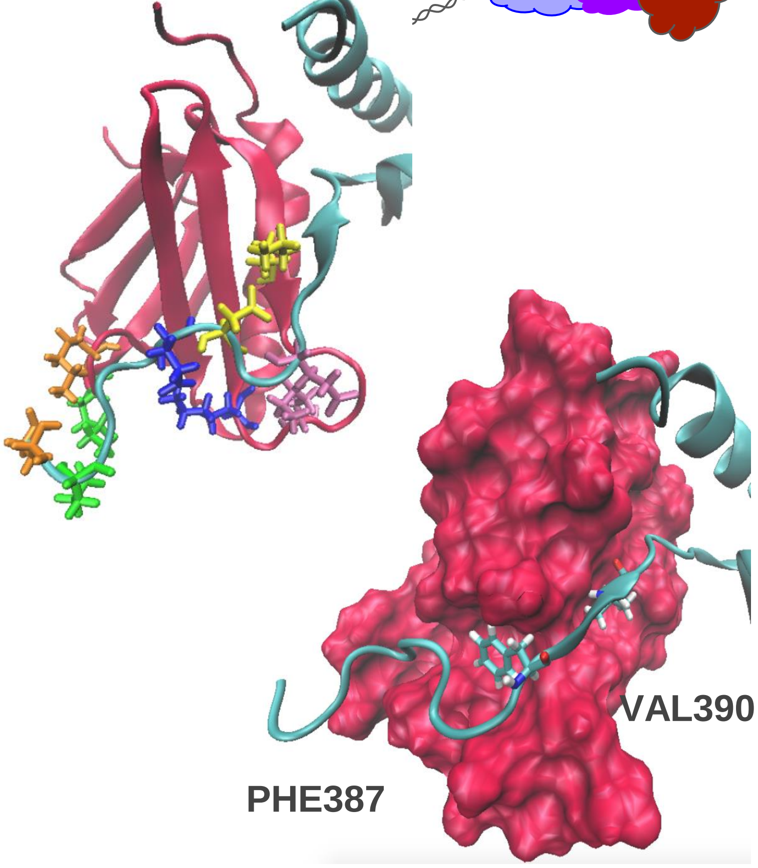


```
Conservation
J0WXG4_AURST  438 P P T T Q D E S M D   D - - - - - - - - -   - - - - - D
M7XFC8_RHOT1  452 E A D A D E E D E D   E f d e f e p a e g   v n t D P N
G5EG49_CAEEL  370 I K M E V D N D E S   E - - - - - - - - -   - - - - E E
O96880_DROME  374 I N S D S A E E D D   D - - - - - - - - -   - - - - - V
Q5SQE1_DANRE  384 Q D E E D E D D D D   E f e - - - - - - -   - e v G D D
Q0P4H8_XENTR  363 - S N A E Q E E E S   E - - - - - - - - -   - - - T V D
T2EA_MOUSE    380 D E D E E D E E F E   E - - - - - - - - -   - - v A D D
T2EA_BOVIN    378 - E E D E E D D E F   E e - - - - - - - -   - - v A D D
T2EA_HUMAN    378 - - E E D E E E D D   E f e - - - - - - -   - e v A D D
```

Glu386 is the only acidic residue “conserved”

```
Conservation
J0WXG4_AURST  438 P P T T Q D E S M D   D - - - - - - - - -   - - - - - D
M7XFC8_RHOT1  452 E A D A D E E D E D   E f d e f e p a e g   v n t D P N
G5EG49_CAEEL  370 I K M E V D N D E S   E - - - - - - - - -   - - - - E E
O96880_DROME  374 I N S D S A E E D D   D - - - - - - - - -   - - - - - V
Q5SQE1_DANRE  384 Q D E E D E D D D D   E f e - - - - - - -   - e v G D D
Q0P4H8_XENTR  363 - S N A E Q E E E S   E - - - - - - - - -   - - - T V D
T2EA_MOUSE    380 D E D E E D E E F E   E - - - - - - - - -   - - v A D D
T2EA_BOVIN    378 - E E D E E D D E F   E e - - - - - - - -   - - v A D D
T2EA_HUMAN    378 - - E E D E E E D D   E f e - - - - - - -   - e v A D D
```

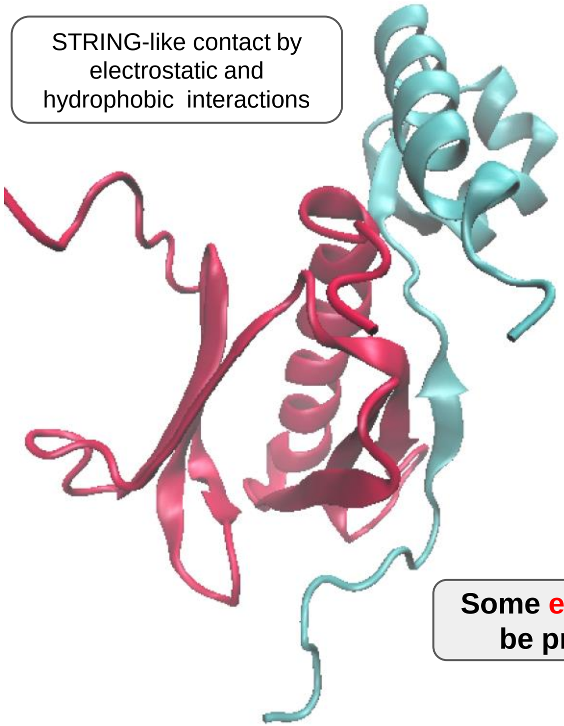
The hydrophpbic Val390 is just shared across mammals

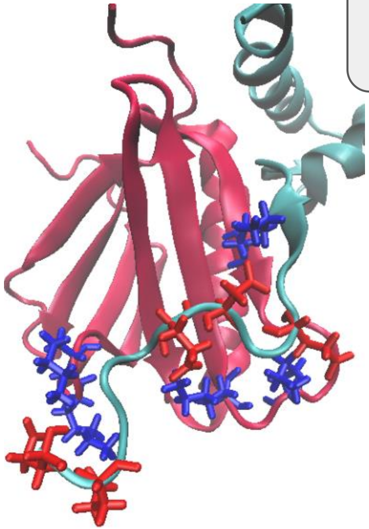


Closed PIC formation: TFIIH joins the complex

TFIIH assembles to the PIC through the C-term domain of TFIIEα

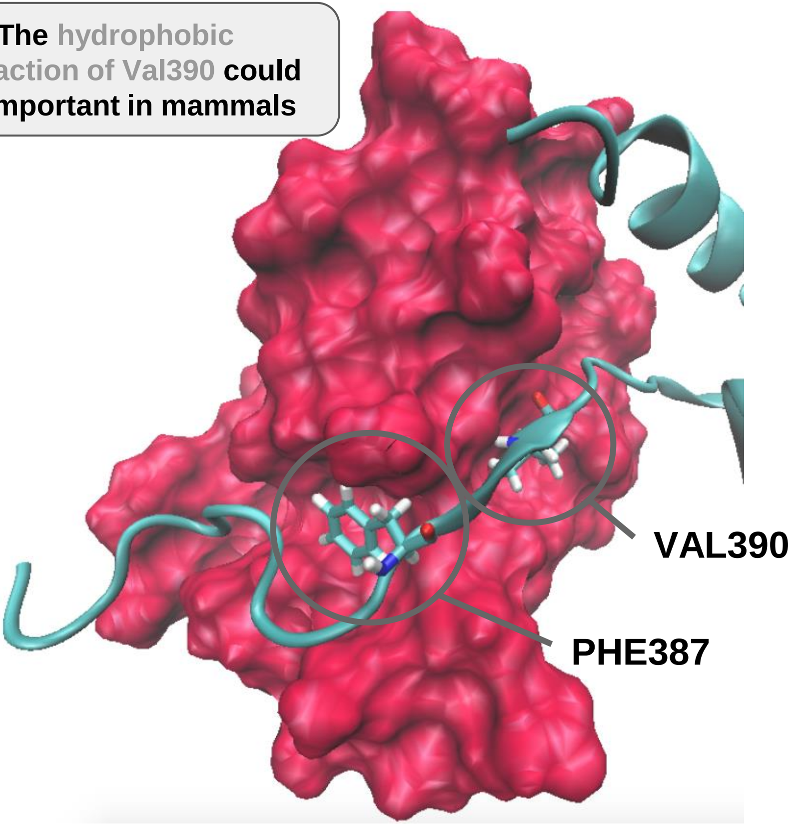
STRING-like contact by electrostatic and hydrophobic interactions





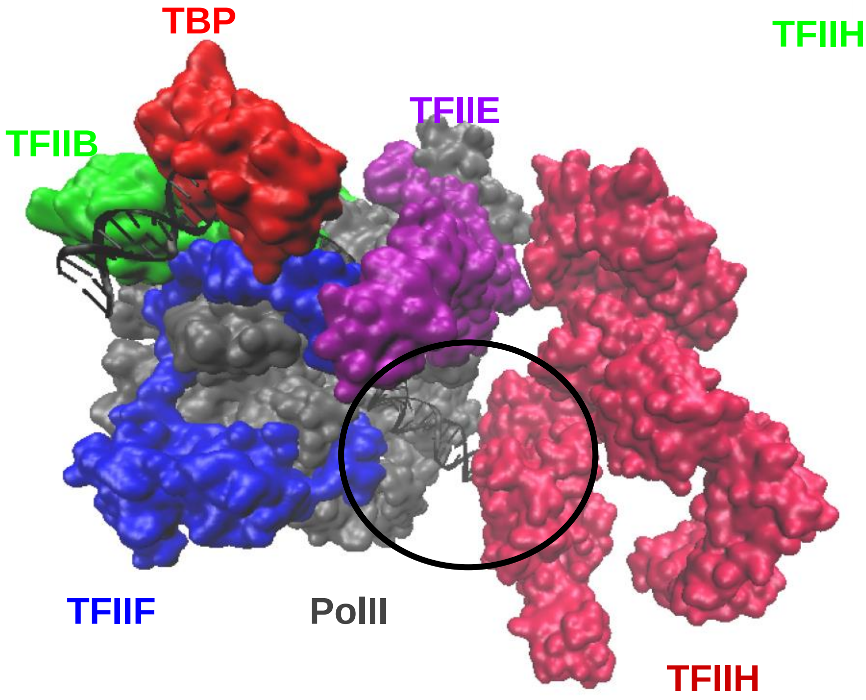
Some electrostatic interactions might be preserved across metazoans

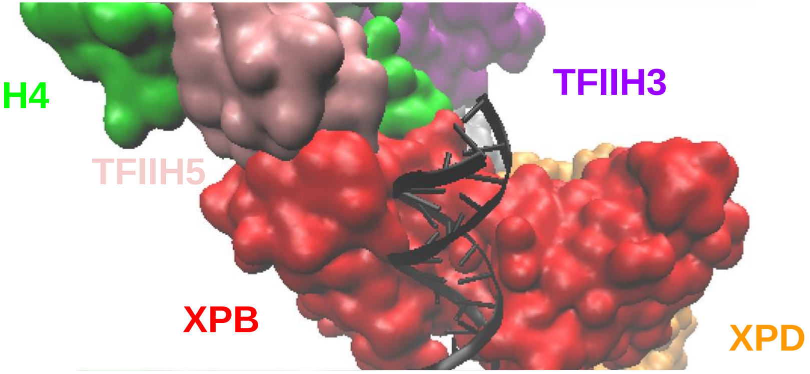
The hydrophobic interaction of Val390 could be important in mammals

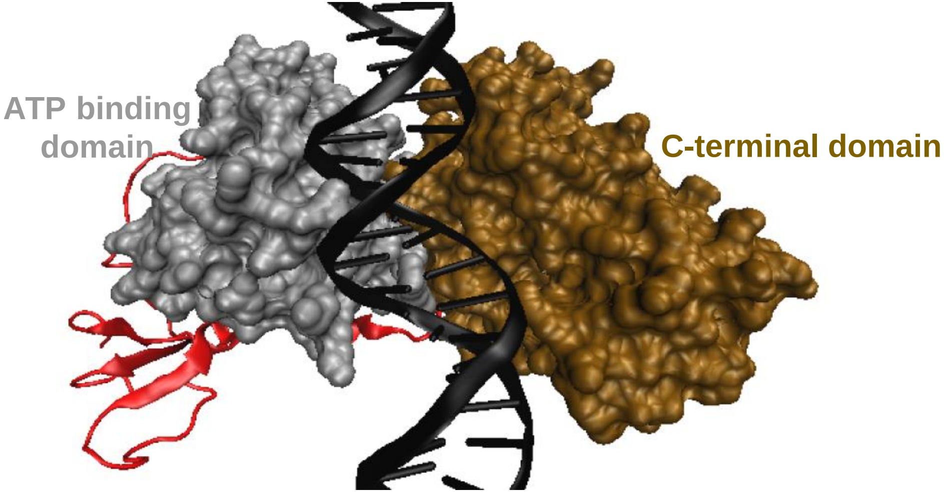


Closed PIC formation: TFIIH joins the complex

XPB is a helicase that binds upstream of the transcription start site (TSS)





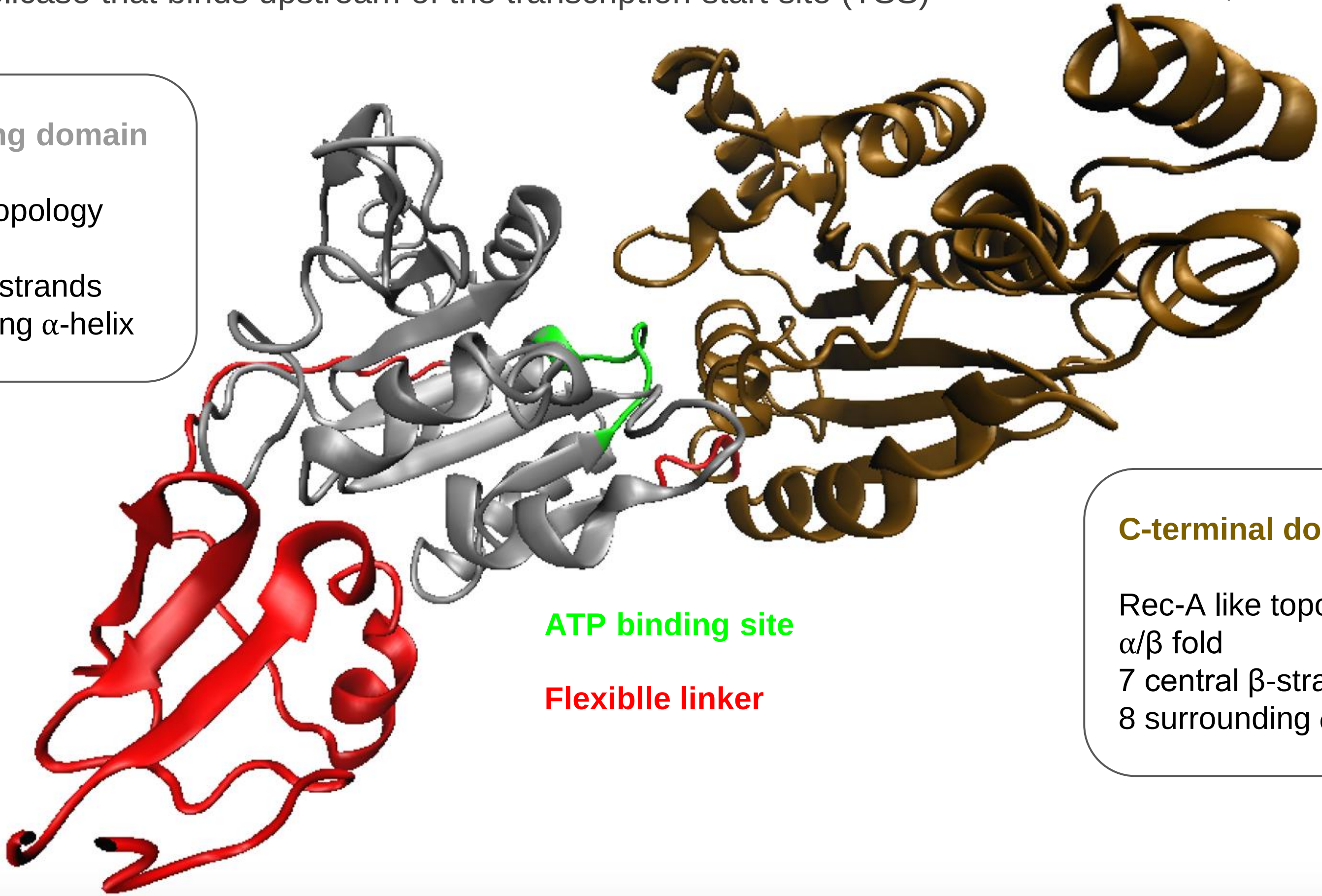


Closed PIC formation: TFIIH joins the complex

XPB is a helicase that binds upstream of the transcription start site (TSS)

ATP binding domain

RecA-like topology
α/β fold
8 central β-strands
6 surrounding α-helix



ATP binding site

Flexiblle linker

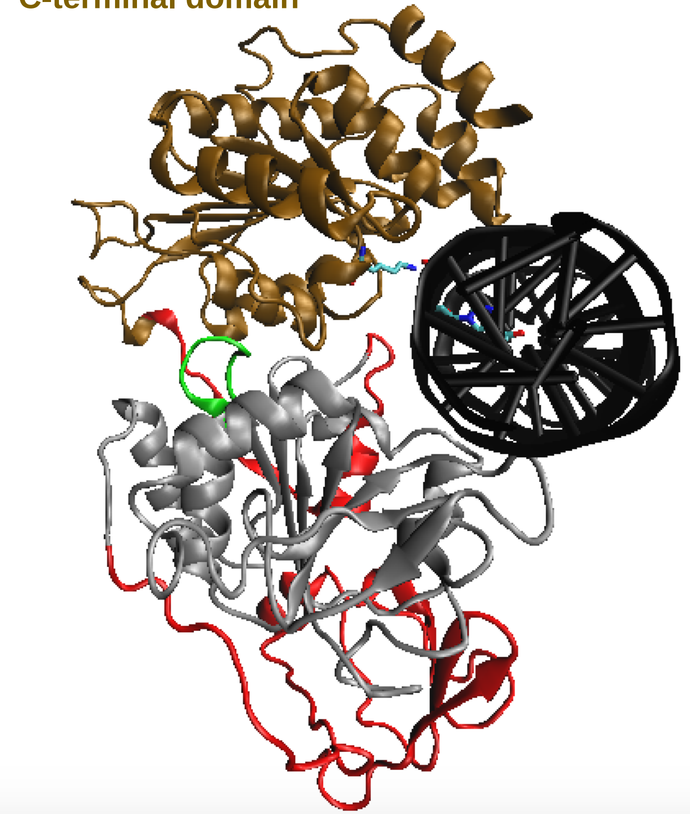
C-terminal domain

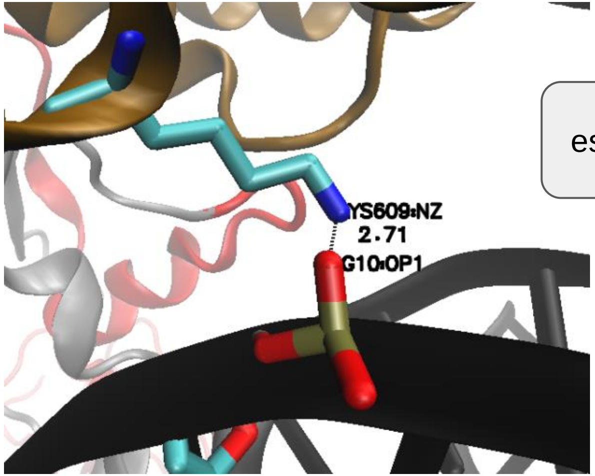
Rec-A like topolopgy
α/β fold
7 central β-strands
8 surrounding α-helix

Closed PIC formation: TFIIH joins the complex

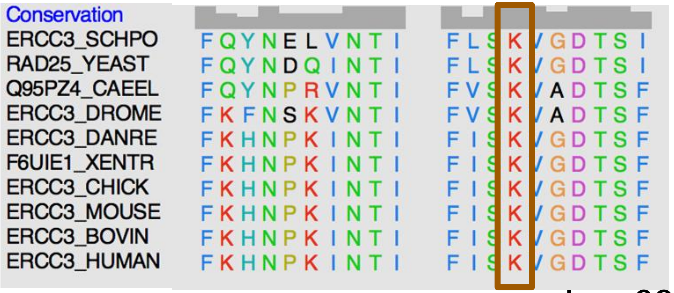
XPB is a helicase that binds upstream of the transcription start site (TSS)

C-terminal domain





The C-terminal domain establishes a salt bridge with the DNA strand



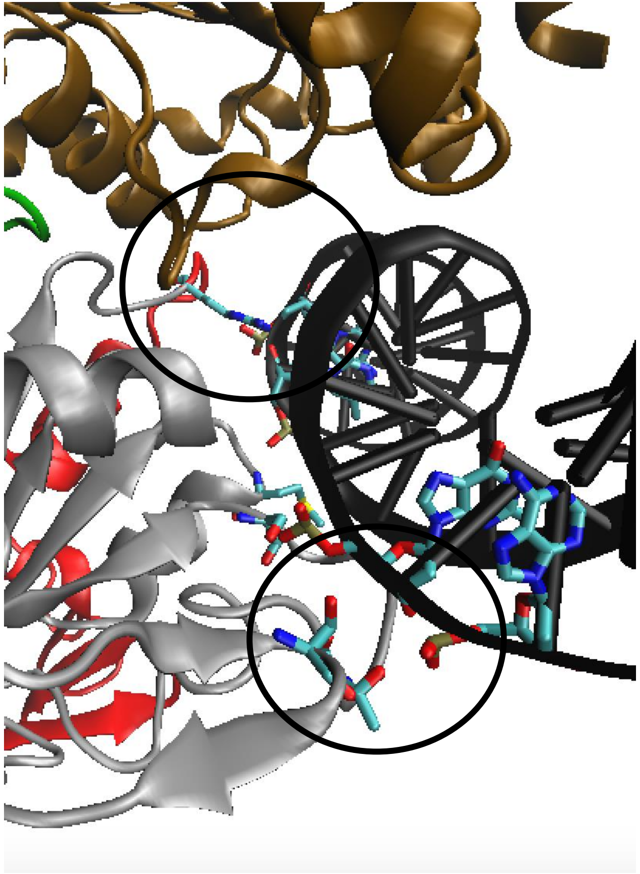
Conservation		
ERCC3_SCHPO	F Q Y N E L V N T I	F L S K V G D T S I
RAD25_YEAST	F Q Y N D Q I N T I	F L S K V G D T S I
Q95PZ4_CAEEL	F Q Y N P R V N T I	F V S K V A D T S F
ERCC3_DROME	F K F N S K V N T I	F V S K V A D T S F
ERCC3_DANRE	F K H N P K I N T I	F I S K V G D T S F
F6UIE1_XENTR	F K H N P K I N T I	F I S K V G D T S F
ERCC3_CHICK	F K H N P K I N T I	F I S K V G D T S F
ERCC3_MOUSE	F K H N P K I N T I	F I S K V G D T S F
ERCC3_BOVIN	F K H N P K I N T I	F I S K V G D T S F
ERCC3_HUMAN	F K H N P K I N T I	F I S K V G D T S F

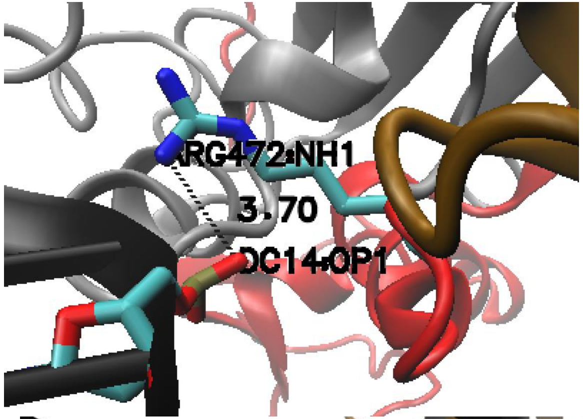
Lys 609

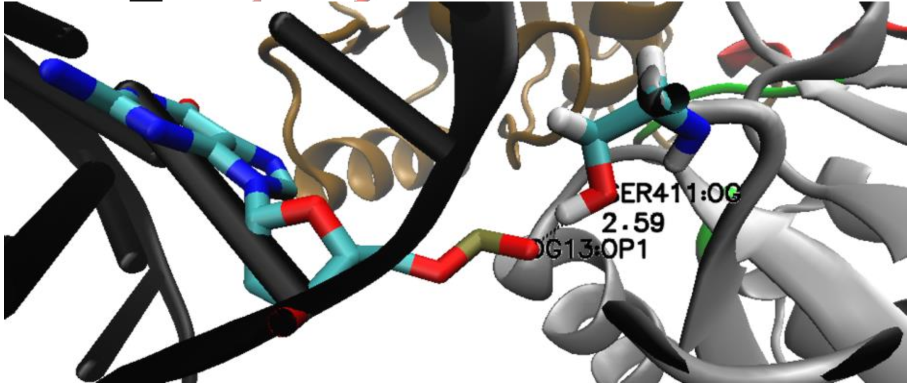
The interaction is preserved across eukaryotes

Closed PIC formation: TFIIH joins the complex

XPB is a helicase that binds upstream of the transcription start site (TSS)





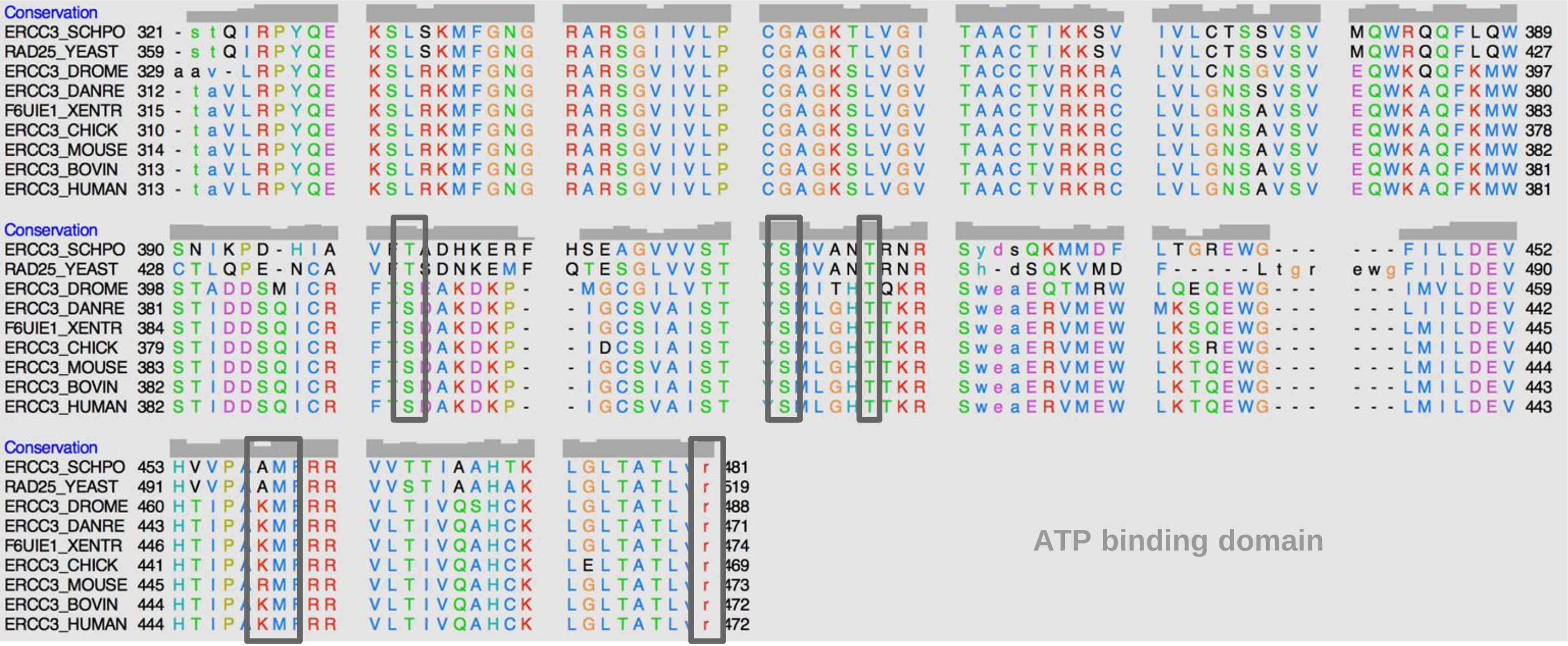


The **ATP-binding domain** establishes **electrostatic** (up) and **hidrogen bonds** (down) with DNA

Closed PIC formation: TFIIH joins the complex

XPB is a helicase that binds upstream of the transcription start site (TSS)

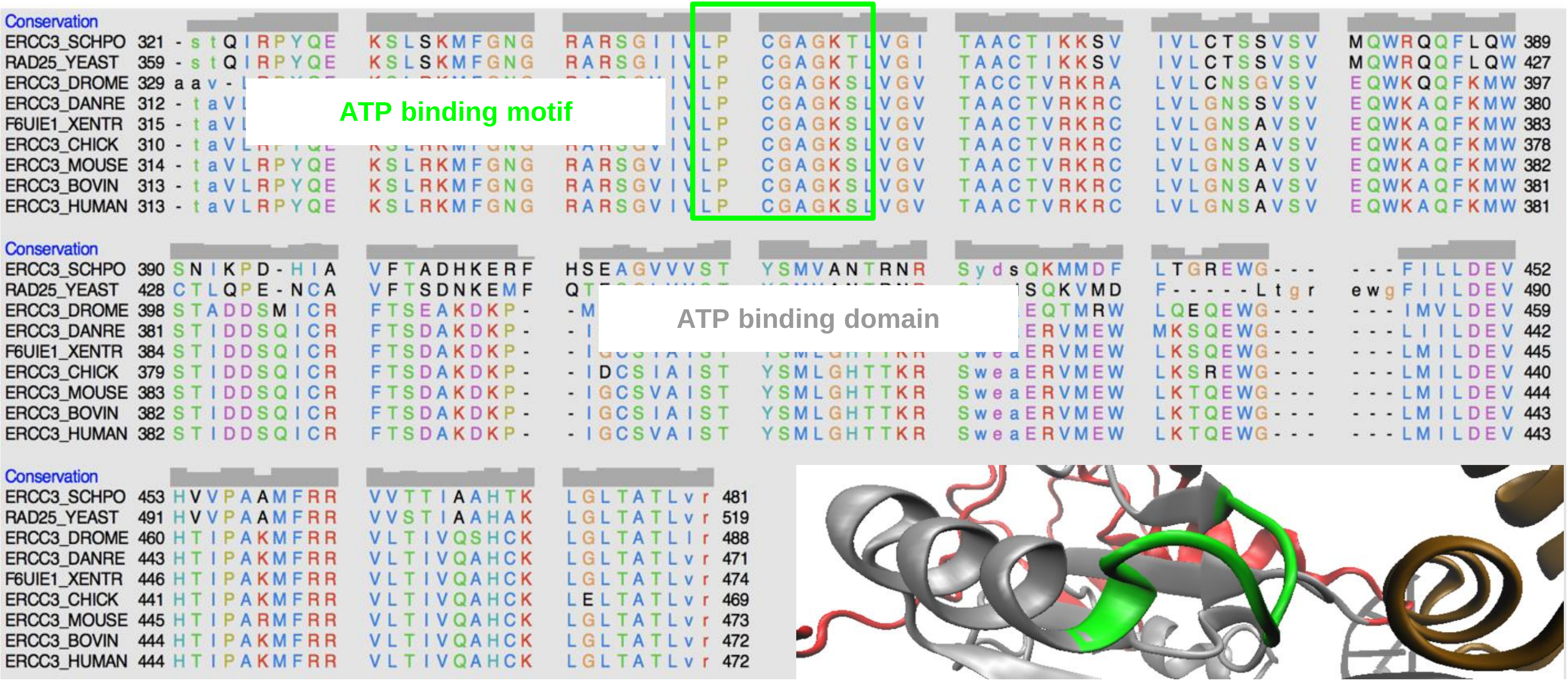
Most residues important for DNA interaction of the ATP-binding domain are conserved across eukaryotes



```
Conservation
ERCC3_SCHPO  321 -stQIRPYQE KSLSKMFGNG RARSGIIVLP CGAGKTLVGI TAACTIKKSV IVLCTSSVSV MQWRQQFLQW 389
RAD25_YEAST  359 -stQIRPYQE KSLSKMFGNG RARSGIIVLP CGAGKTLVGI TAACTIKKSV IVLCTSSVSV MQWRQQFLQW 427
ERCC3_DROME  329 aav-LRPYQE KSLRKMFGNG RARSGVIVLP CGAGKSLVGV TACCTVRKRA LVLCNSGVSV EQWKQQFKMW 397
ERCC3_DANRE  312 -taVLRPYQE KSLRKMFGNG RARSGVIVLP CGAGKSLVGV TAACTVRKRC LVLGNSSVSV EQWKAQFKMW 380
F6UIE1_XENTR 315 -taVLRPYQE KSLRKMFGNG RARSGVIVLP CGAGKSLVGV TAACTVRKRC LVLGNSAVSV EQWKAQFKMW 383
ERCC3_CHICK  310 -taVLRPYQE KSLRKMFGNG RARSGVIVLP CGAGKSLVGV TAACTVRKRC LVLGNSAVSV EQWKAQFKMW 378
ERCC3_MOUSE  314 -taVLRPYQE KSLRKMFGNG RARSGVIVLP CGAGKSLVGV TAACTVRKRC LVLGNSAVSV EQWKAQFKMW 382
ERCC3_BOVIN  313 -taVLRPYQE KSLRKMFGNG RARSGVIVLP CGAGKSLVGV TAACTVRKRC LVLGNSAVSV EQWKAQFKMW 381
ERCC3_HUMAN  313 -taVLRPYQE KSLRKMFGNG RARSGVIVLP CGAGKSLVGV TAACTVRKRC LVLGNSAVSV EQWKAQFKMW 381

Conservation
ERCC3_SCHPO  390 SNIKPD-HIA VFTADHKERF HSEAGVVVST YSMVANTRNR SydsQKMMDF LTGREWG--- ---FILLDEV 452
RAD25_YEAST  428 CTLQPE-NCA VFTSDNKEMF QTESGLVVST YSMVANTRNR Sh-dSQKVMD F-----Ltgr ewgFIILDEV 490
ERCC3_DROME  398 STADDSMICR FTSEAKDKP- -MGCGILVTT YSMITHTQKR SweaEQTMRW LQEQEWG--- ---IMVLDEV 459
ERCC3_DANRE  381 STIDDSQICR FTSDAKDKP- -IGCSVAIST YSMLGHTTKR SweaERVMEW MKSQEWG--- ---LIILDEV 442
F6UIE1_XENTR 384 STIDDSQICR FTSDAKDKP- -IGCSIAIST YSMLGHTTKR SweaERVMEW LKSQEWG--- ---LMILDEV 445
ERCC3_CHICK  379 STIDDSQICR FTSDAKDKP- -IDCSIAIST YSMLGHTTKR SweaERVMEW LKSREWG--- ---LMILDEV 440
ERCC3_MOUSE  383 STIDDSQICR FTSDAKDKP- -IGCSVAIST YSMLGHTTKR SweaERVMEW LKTQEWG--- ---LMILDEV 444
ERCC3_BOVIN  382 STIDDSQICR FTSDAKDKP- -IGCSIAIST YSMLGHTTKR SweaERVMEW LKTQEWG--- ---LMILDEV 443
ERCC3_HUMAN  382 STIDDSQICR FTSDAKDKP- -IGCSVAIST YSMLGHTTKR SweaERVMEW LKTQEWG--- ---LMILDEV 443

Conservation
ERCC3_SCHPO  453 HVVPAAMFRR VVTTIAAHTK LGLTATLvr 481
RAD25_YEAST  491 HVVPAAMFRR VVSTIAAHAK LGLTATLvr 519
ERCC3_DROME  460 HTIPAKMFRR VLTIVQSHCK LGLTATL r 488
ERCC3_DANRE  443 HTIPAKMFRR VLTIVQAHCK LGLTATLvr 471
F6UIE1_XENTR 446 HTIPAKMFRR VLTIVQAHCK LGLTATLvr 474
ERCC3_CHICK  441 HTIPAKMFRR VLTIVQAHCK LELTATLvr 469
ERCC3_MOUSE  445 HTIPARMFRR VLTIVQAHCK LGLTATLvr 473
ERCC3_BOVIN  444 HTIPAKMFRR VLTIVQAHCK LGLTATLvr 472
ERCC3_HUMAN  444 HTIPAKMFRR VLTIVQAHCK LGLTATLvr 472
```

Closed PIC formation: TFIIH joins the complex

XPB requires ATP for unwinding DNA, from yeast to humans



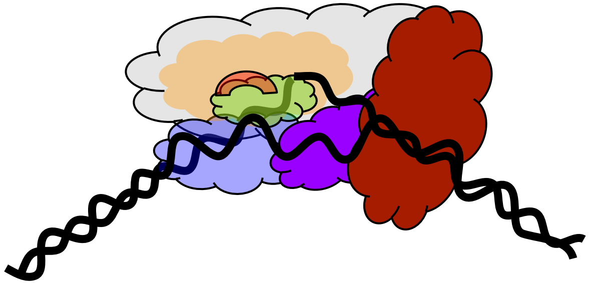
ATP binding motif

ATP binding domain

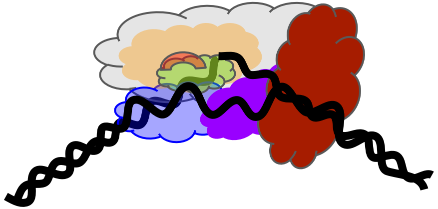
```
Conservation
ERCC3_SCHPO 321 - s t Q I R P Y Q E  K S L S K M F G N G  R A R S G I I V L P  C G A G K T L V G I  T A A C T I K K S V  I V L C T S S V S V  M Q W R Q Q F L Q W 389
RAD25_YEAST 359 - s t Q I R P Y Q E  K S L S K M F G N G  R A R S G I I V L P  C G A G K T L V G I  T A A C T I K K S V  I V L C T S S V S V  M Q W R Q Q F L Q W 427
ERCC3_DROME 329 a a v - l                                       I V L P  C G A G K S L V G V  T A C C T V R K R A  L V L C N S G V S V  E Q W K Q Q F K M W 397
ERCC3_DANRE 312 - t a V l                                       I V L P  C G A G K S L V G V  T A A C T V R K R C  L V L G N S S V S V  E Q W K A Q F K M W 380
F6UIE1_XENTR 315 - t a V l                                      I V L P  C G A G K S L V G V  T A A C T V R K R C  L V L G N S A V S V  E Q W K A Q F K M W 383
ERCC3_CHICK 310 - t a V L                                       I V L P  C G A G K S L V G V  T A A C T V R K R C  L V L G N S A V S V  E Q W K A Q F K M W 378
ERCC3_MOUSE 314 - t a V L R P Y Q E  K S L R K M F G N G  R A R S G V I V L P  C G A G K S L V G V  T A A C T V R K R C  L V L G N S A V S V  E Q W K A Q F K M W 382
ERCC3_BOVIN 313 - t a V L R P Y Q E  K S L R K M F G N G  R A R S G V I V L P  C G A G K S L V G V  T A A C T V R K R C  L V L G N S A V S V  E Q W K A Q F K M W 381
ERCC3_HUMAN 313 - t a V L R P Y Q E  K S L R K M F G N G  R A R S G V I V L P  C G A G K S L V G V  T A A C T V R K R C  L V L G N S A V S V  E Q W K A Q F K M W 381

Conservation
ERCC3_SCHPO 390 S N I K P D - H I A  V F T A D H K E R F  H S E A G V V V S T  Y S M V A N T R N R  S y d s Q K M M D F  L T G R E W G - - -  - - - F I L L D E V 452
RAD25_YEAST 428 C T L Q P E - N C A  V F T S D N K E M F  Q T                                        S Q K V M D  F - - - - - L t g r  e w g F I I L D E V 490
ERCC3_DROME 398 S T A D D S M I C R  F T S E A K D K P -  - M                                        E Q T M R W  L Q E Q E W G - - -  - - - I M V L D E V 459
ERCC3_DANRE 381 S T I D D S Q I C R  F T S D A K D K P -  - I                                        E R V M E W  M K S Q E W G - - -  - - - L I I L D E V 442
F6UIE1_XENTR 384 S T I D D S Q I C R  F T S D A K D K P -  - I                                       E R V M E W  L K S Q E W G - - -  - - - L M I L D E V 445
ERCC3_CHICK 379 S T I D D S Q I C R  F T S D A K D K P -  - I D C S I A I S T  Y S M L G H T T K R  S w e a E R V M E W  L K S R E W G - - -  - - - L M I L D E V 440
ERCC3_MOUSE 383 S T I D D S Q I C R  F T S D A K D K P -  - I G C S V A I S T  Y S M L G H T T K R  S w e a E R V M E W  L K T Q E W G - - -  - - - L M I L D E V 444
ERCC3_BOVIN 382 S T I D D S Q I C R  F T S D A K D K P -  - I G C S I A I S T  Y S M L G H T T K R  S w e a E R V M E W  L K T Q E W G - - -  - - - L M I L D E V 443
ERCC3_HUMAN 382 S T I D D S Q I C R  F T S D A K D K P -  - I G C S V A I S T  Y S M L G H T T K R  S w e a E R V M E W  L K T Q E W G - - -  - - - L M I L D E V 443

Conservation
ERCC3_SCHPO 453 H V V P A A M F R R  V V T T I A A H T K  L G L T A T L v r 481
RAD25_YEAST 491 H V V P A A M F R R  V V S T I A A H A K  L G L T A T L v r 519
ERCC3_DROME 460 H T I P A K M F R R  V L T I V Q S H C K  L G L T A T L I r 488
ERCC3_DANRE 443 H T I P A K M F R R  V L T I V Q A H C K  L G L T A T L v r 471
F6UIE1_XENTR 446 H T I P A K M F R R  V L T I V Q A H C K  L G L T A T L v r 474
ERCC3_CHICK 441 H T I P A K M F R R  V L T I V Q A H C K  L E L T A T L v r 469
ERCC3_MOUSE 445 H T I P A R M F R R  V L T I V Q A H C K  L G L T A T L v r 473
ERCC3_BOVIN 444 H T I P A K M F R R  V L T I V Q A H C K  L G L T A T L v r 472
ERCC3_HUMAN 444 H T I P A K M F R R  V L T I V Q A H C K  L G L T A T L v r 472
```

Open complex formation:

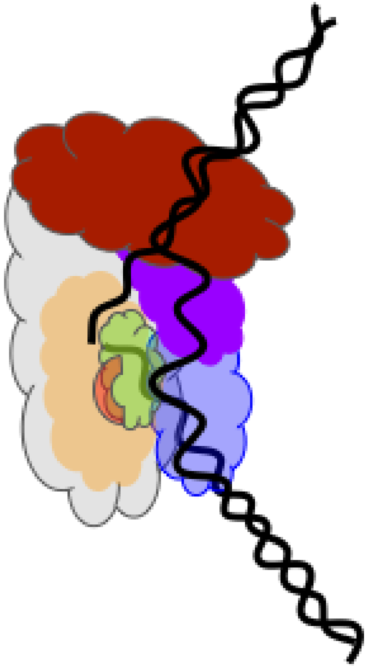
DNA melts around the transcription start site (TSS)

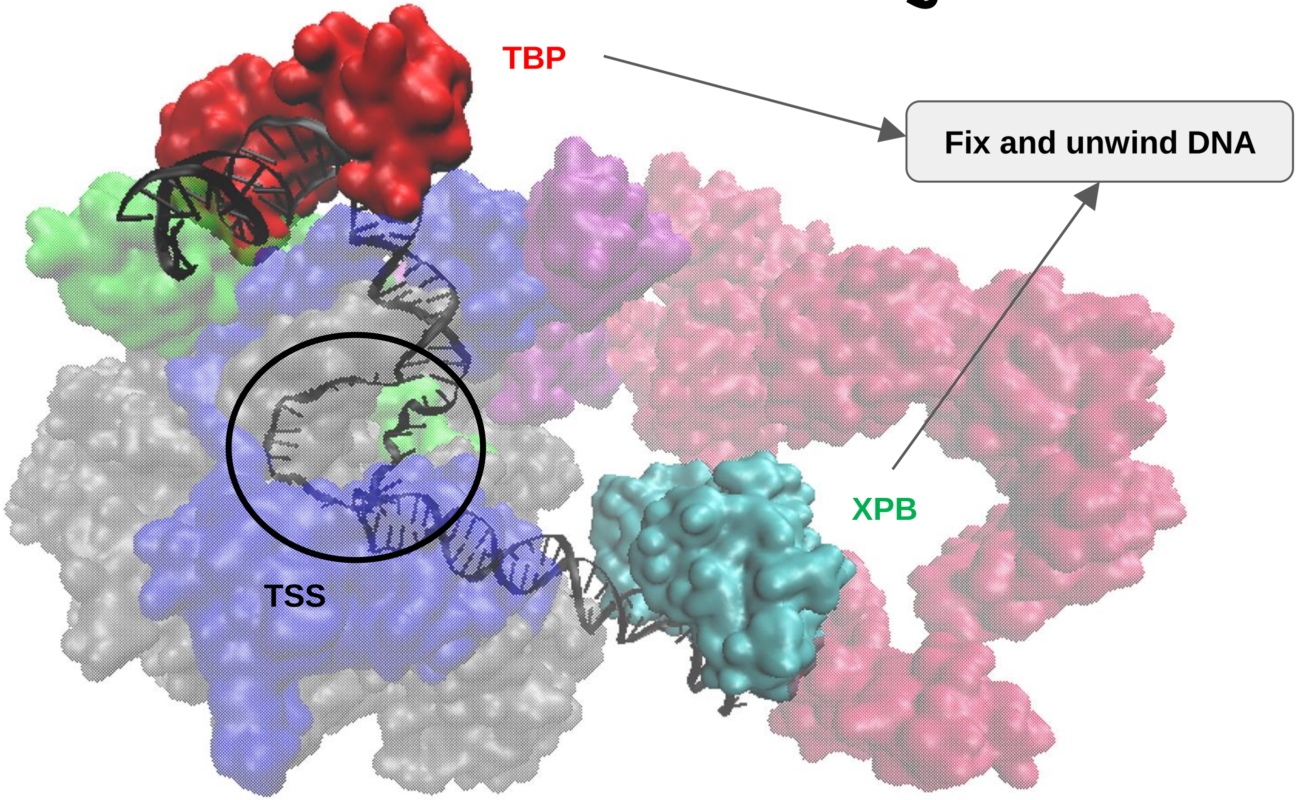


Open complex formation: DNA melts around the transcription start site (TSS)



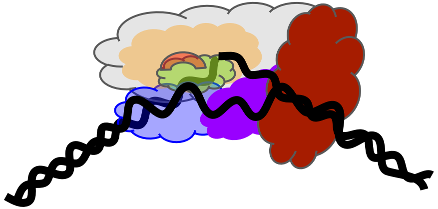
XPB is the main articulator of promoter opening



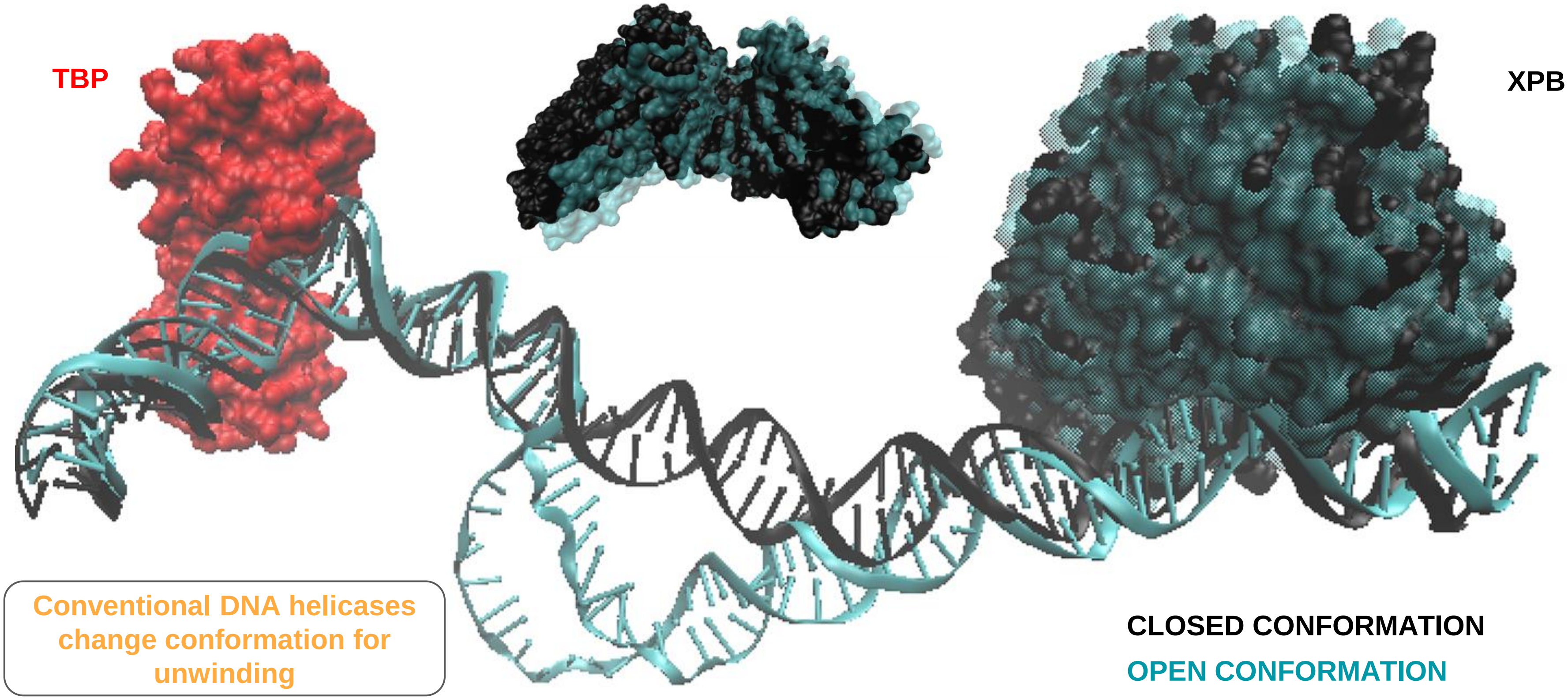


Fix and unwind DNA

Open complex formation: DNA melts around the transcription start site (TSS)

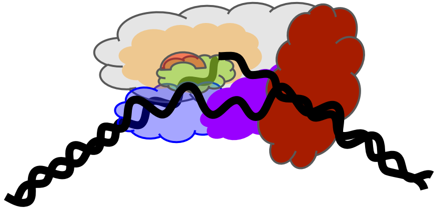


XPB does not change its conformation upon opening

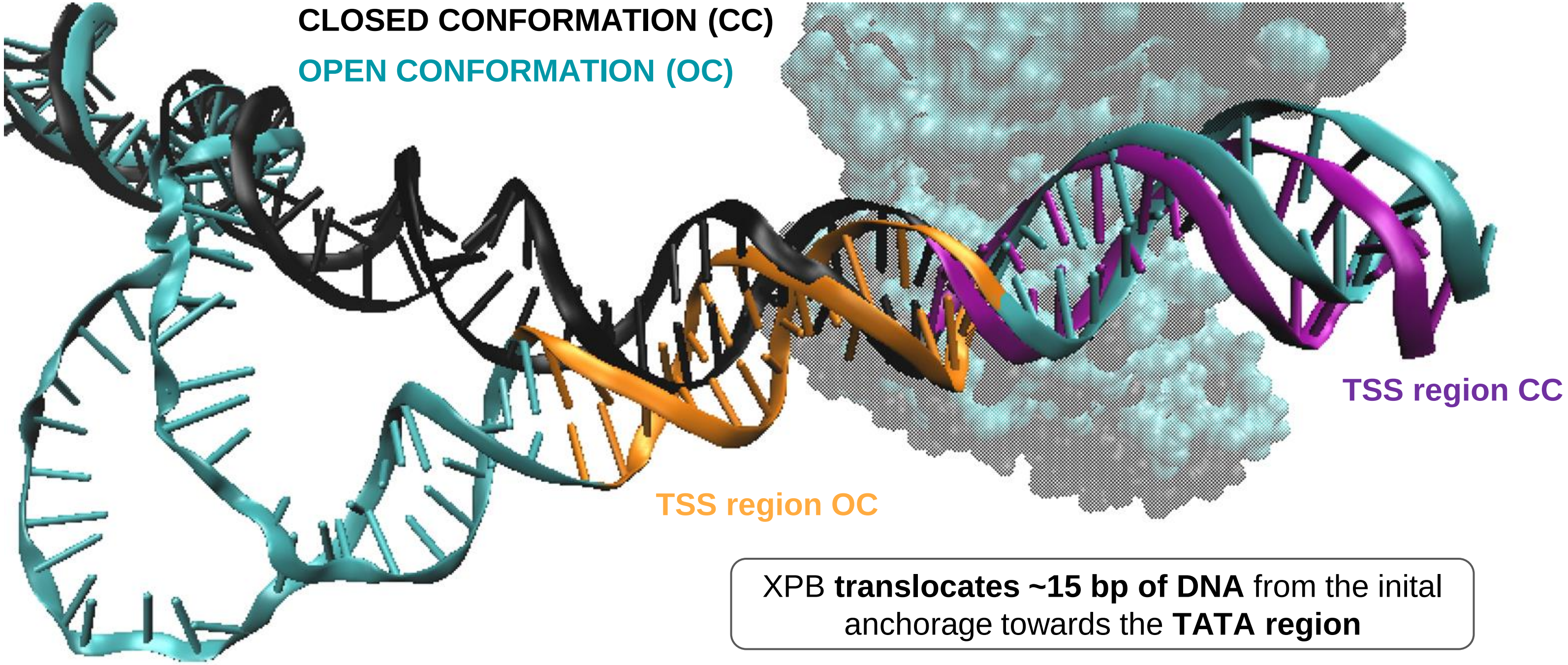


Conventional DNA helicases change conformation for unwinding

Open complex formation: DNA melts around the transcription start site (TSS)



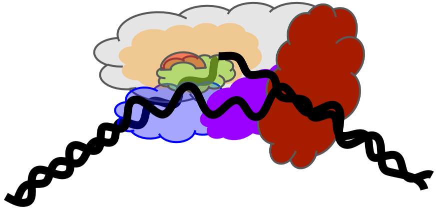
DNA is pushed towards the TATA-region complex by XPB



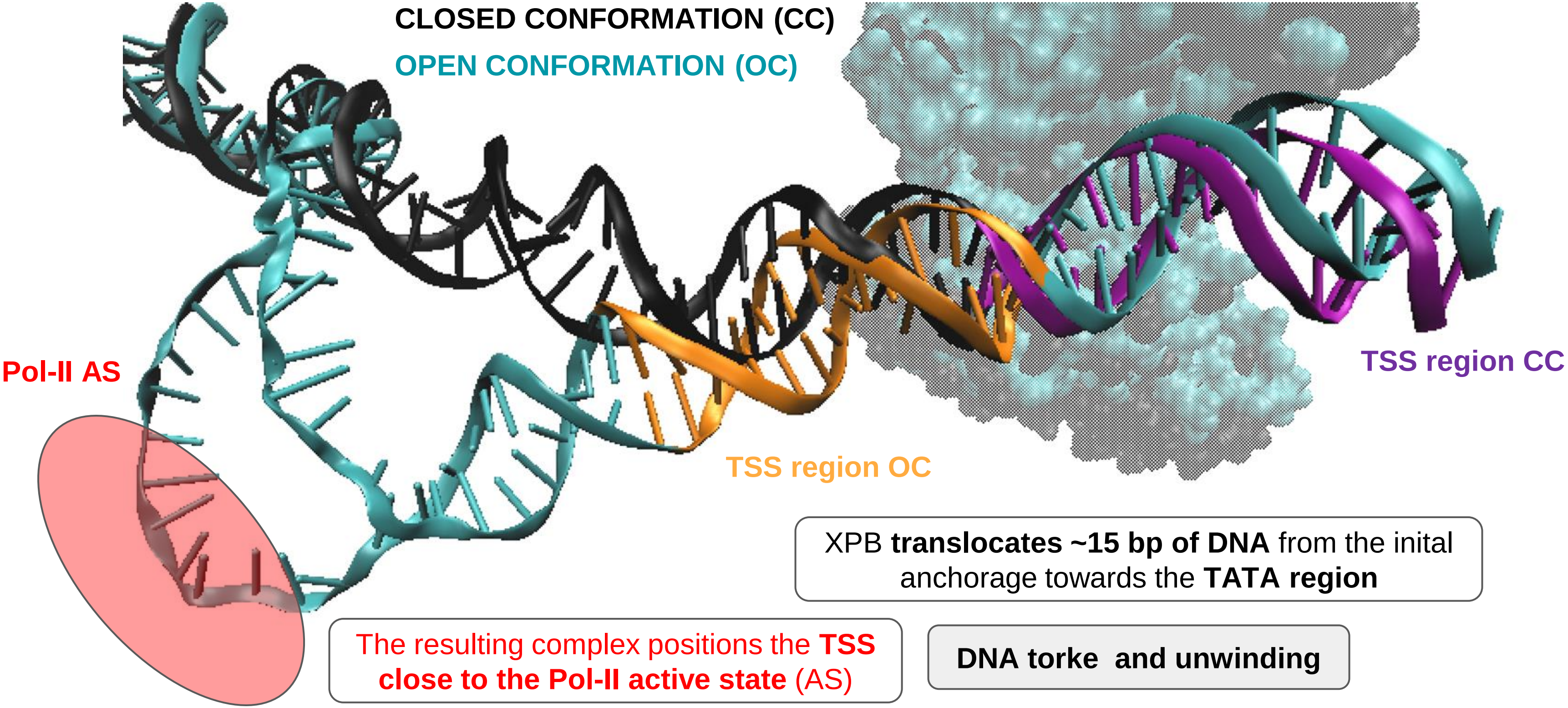
XPB **translocates ~15 bp of DNA** from the inital anchorage towards the **TATA region**

DNA torke and unwinding

Open complex formation: DNA melts around the transcription start site (TSS)



DNA is pushed towards the Pol-II complex by XPB

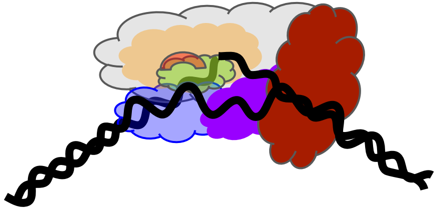


XPB **translocates ~15 bp of DNA** from the inital anchorage towards the **TATA region**

The resulting complex positions the **TSS close to the Pol-II active state** (AS)

DNA torke and unwinding

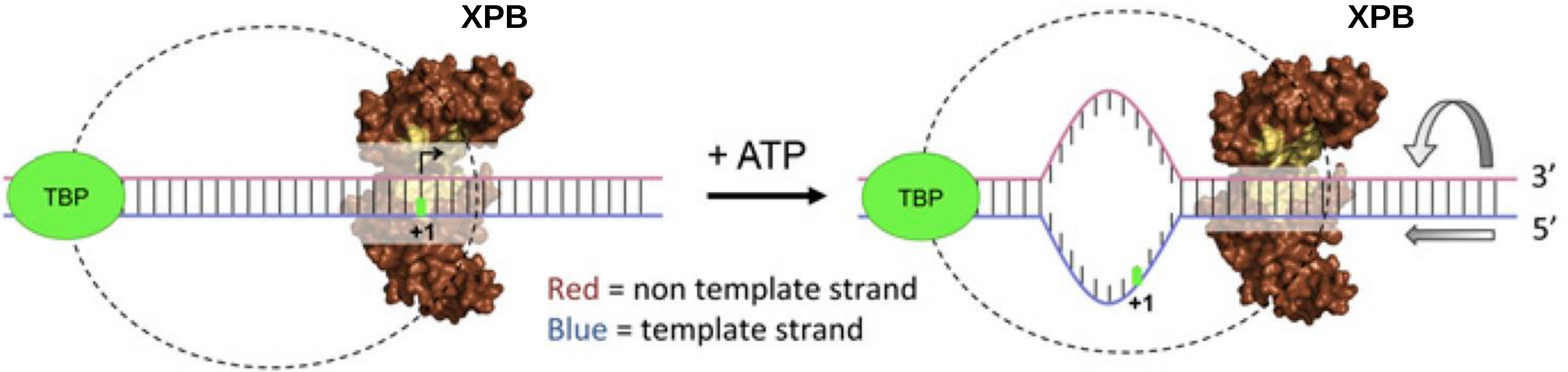
Open complex formation: DNA melts around the transcription start site (TSS)



XPB probably acts as a SWI2/SNF2 translocase

Closed complex

Open complex

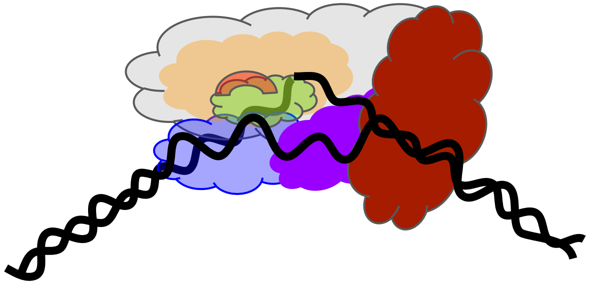


A model for XPB-mediated promoter opening

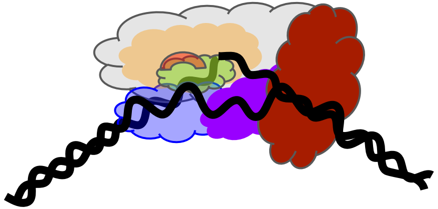
Adapted from Fishburn et al 2015

Open complex formation:

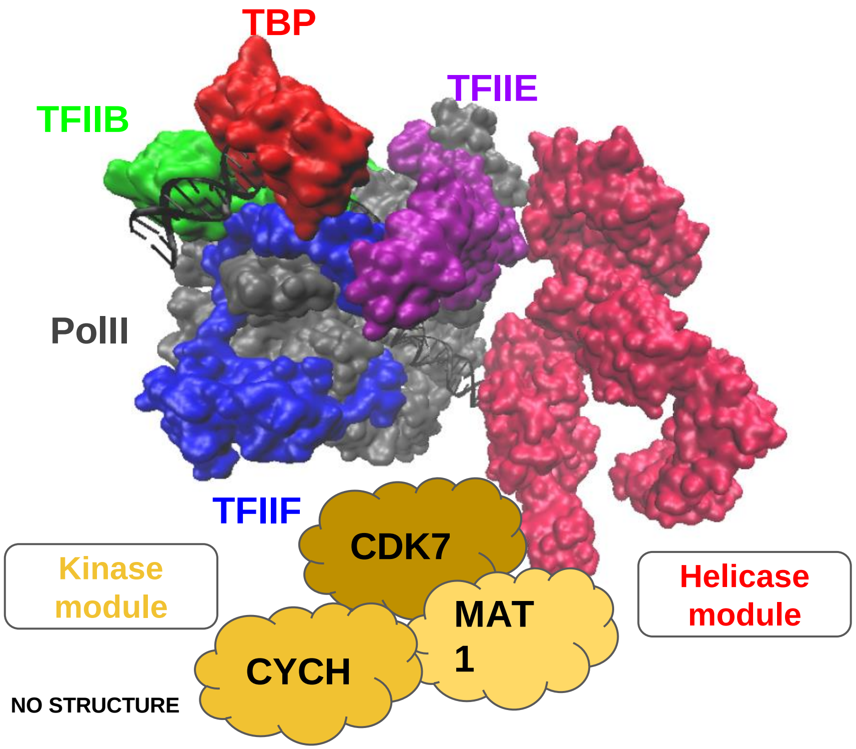
Pol-II is rearranged to allow transcription initiation

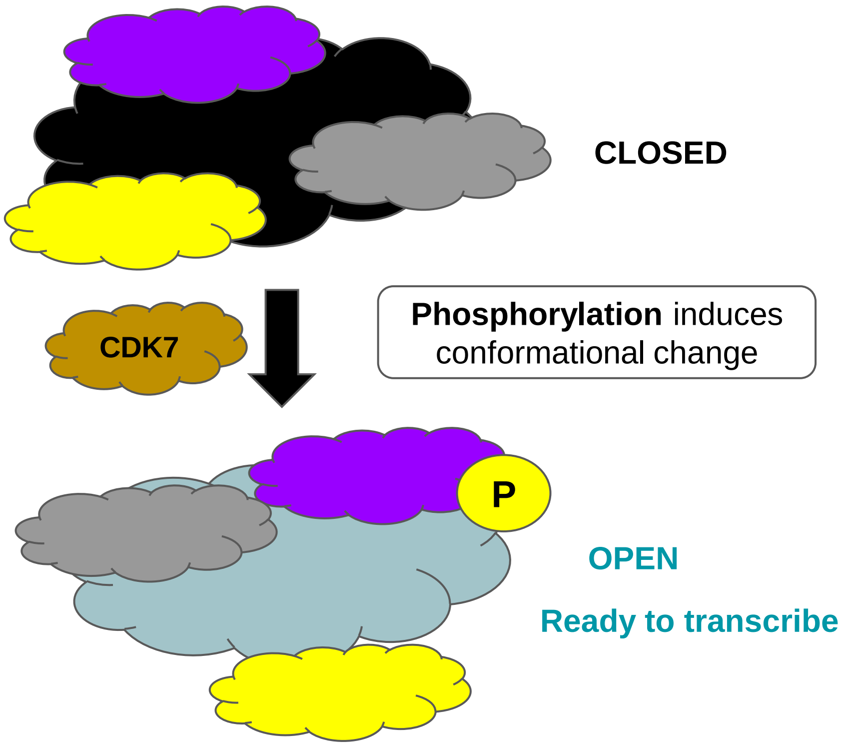


Open complex formation: Pol-II is rearranged to allow transcription initiation



In parallel, the kinase module of TFIIH induces the rearrangement of Pol-II



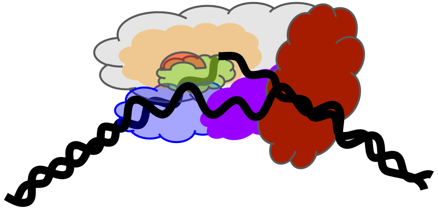


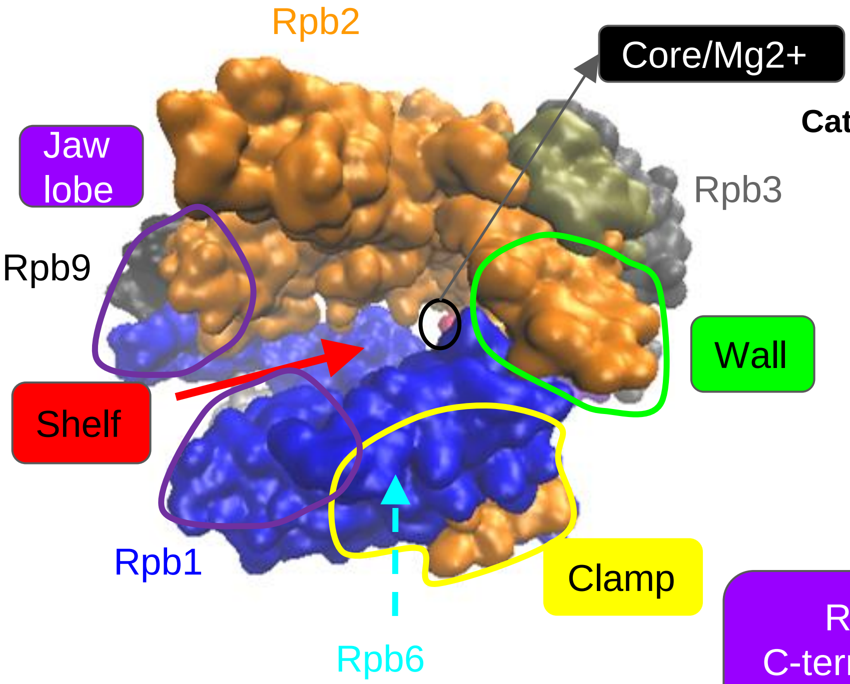
Phosphorylation induces conformational change

Ready to transcribe

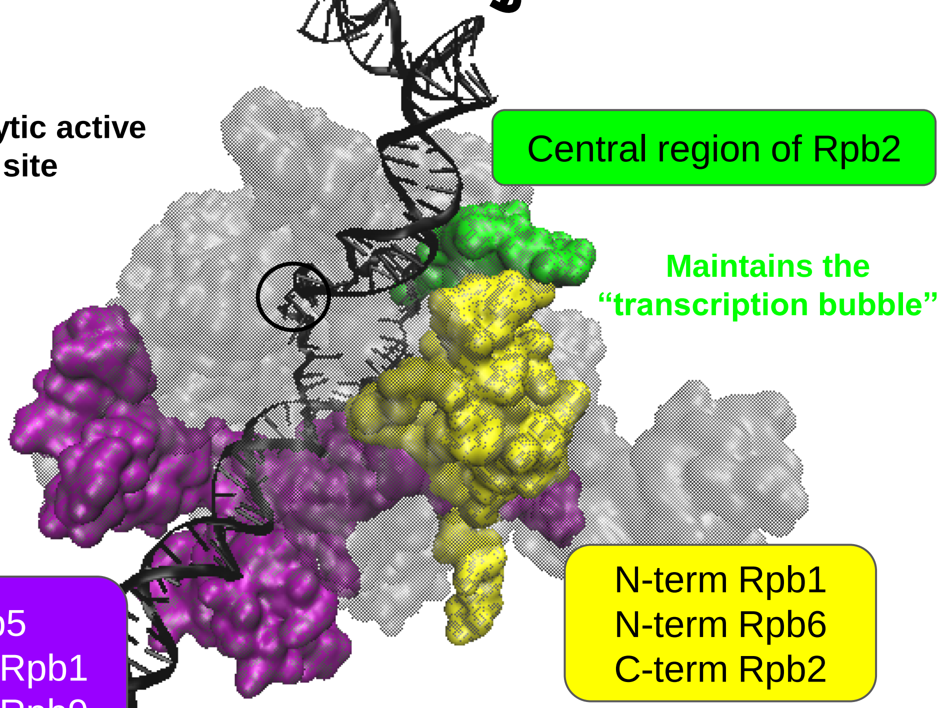
Open complex formation: Pol-II is rearranged to allow transcription initiation

Jaw lobe, clamp, core and wall the functional modules of Pol-II





Catalytic active site

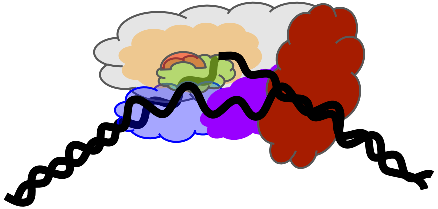


Maintains the "transcription bubble"

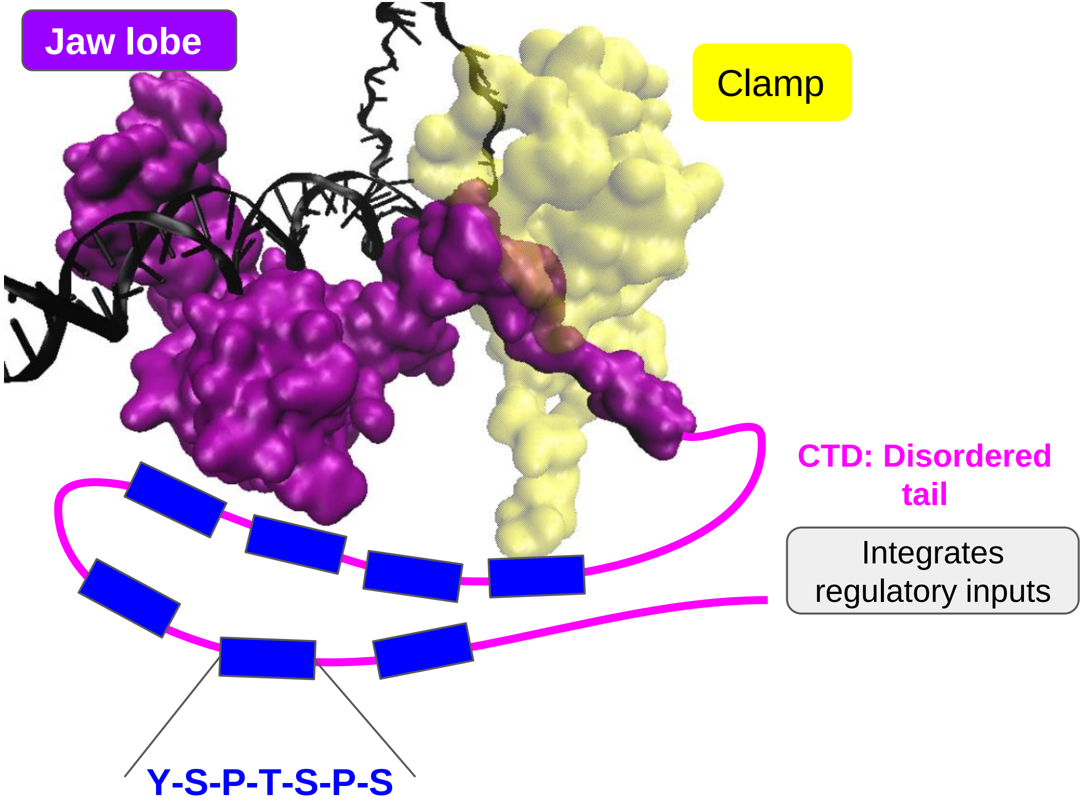
Maintains the position of DNA

Induces most of the rearrangements

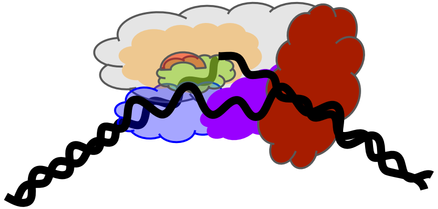
Open complex formation: Pol-II is rearranged to allow transcription initiation



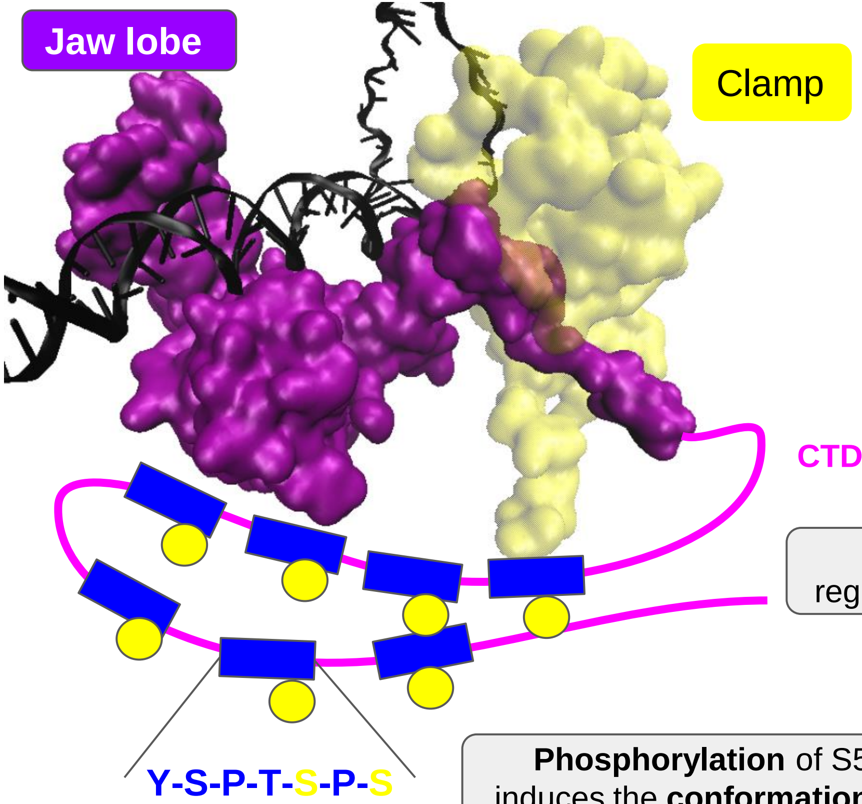
The jaw lobe is linked to the CTD of Rpb1, relevant for the conformational change



Open complex formation: Pol-II is rearranged to allow transcription initiation



The jaw lobe is linked to the CTD of Rpb1, relevant for the conformational change

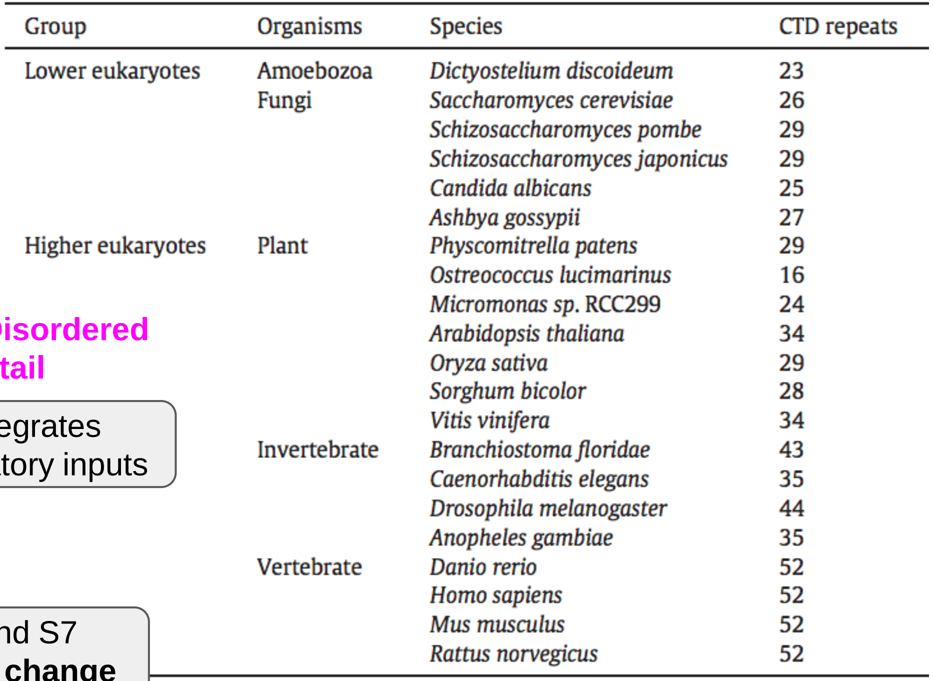


CTD: Disordered tail

Integrates regulatory inputs

Phosphorylation of S5 and S7 induces the **conformational change**

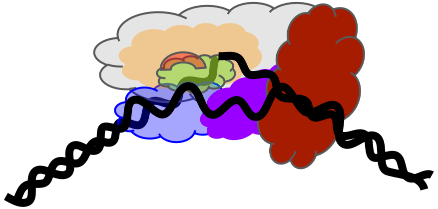
Complexity is correlated with the number of CTD repeats



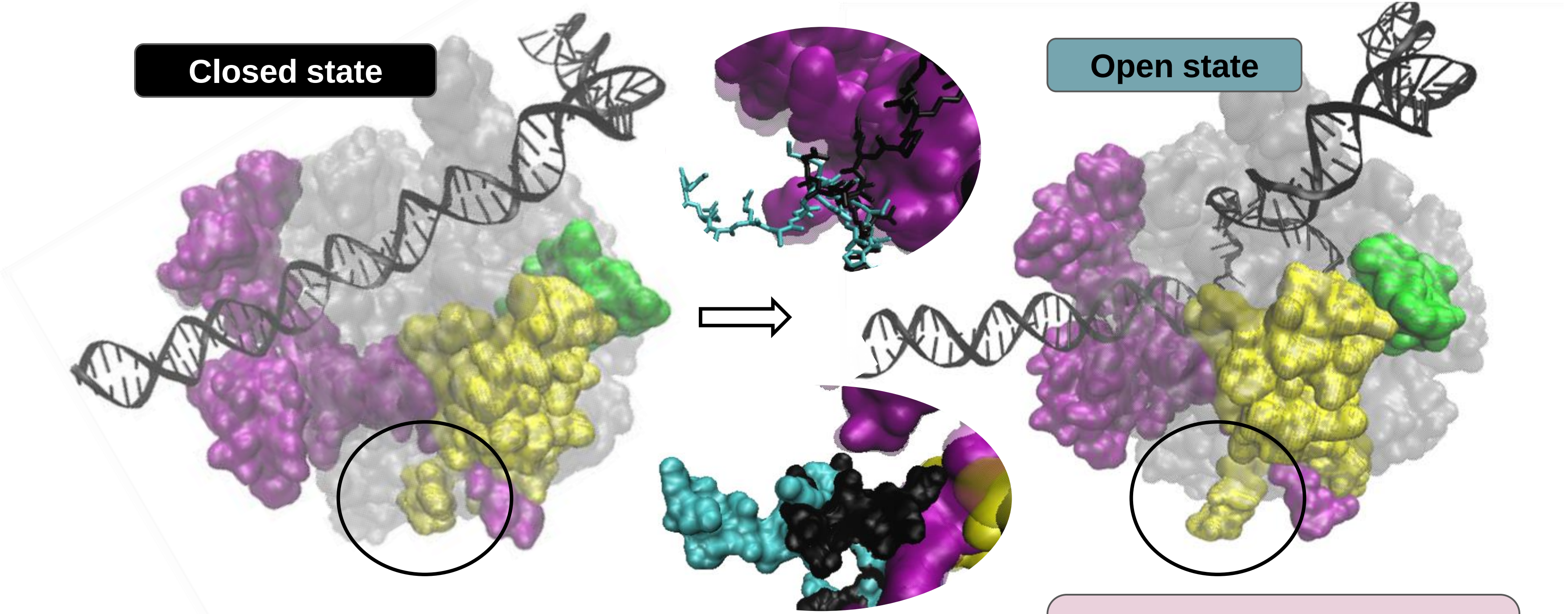
Group	Organisms	Species	CTD repeats
Lower eukaryotes	Amoebozoa	*Dictyostelium discoideum*	23
	Fungi	*Saccharomyces cerevisiae*	26
		Schizosaccharomyces pombe	29
		Schizosaccharomyces japonicus	29
		Candida albicans	25
		Ashbya gossypii	27
Higher eukaryotes	Plant	*Physcomitrella patens*	29
		Ostreococcus lucimarinus	16
		Micromonas sp. RCC299	24
		Arabidopsis thaliana	34
		Oryza sativa	29
		Sorghum bicolor	28
		Vitis vinifera	34
	Invertebrate	*Branchiostoma floridae*	43
		Caenorhabditis elegans	35
		Drosophila melanogaster	44
		Anopheles gambiae	35
	Vertebrate	*Danio rerio*	52
		Homo sapiens	52
		Mus musculus	52
		Rattus norvegicus	52

Extracted from Srivastava et al 2015

Open complex formation: Pol-II is rearranged to allow transcription initiation



A huge CLAMP movement triggers the Pol-II rearrangement



Closed state

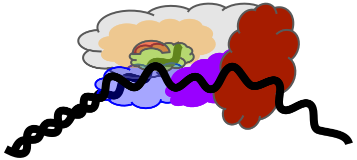
Open state

The N-term of Rpb6 changes massively

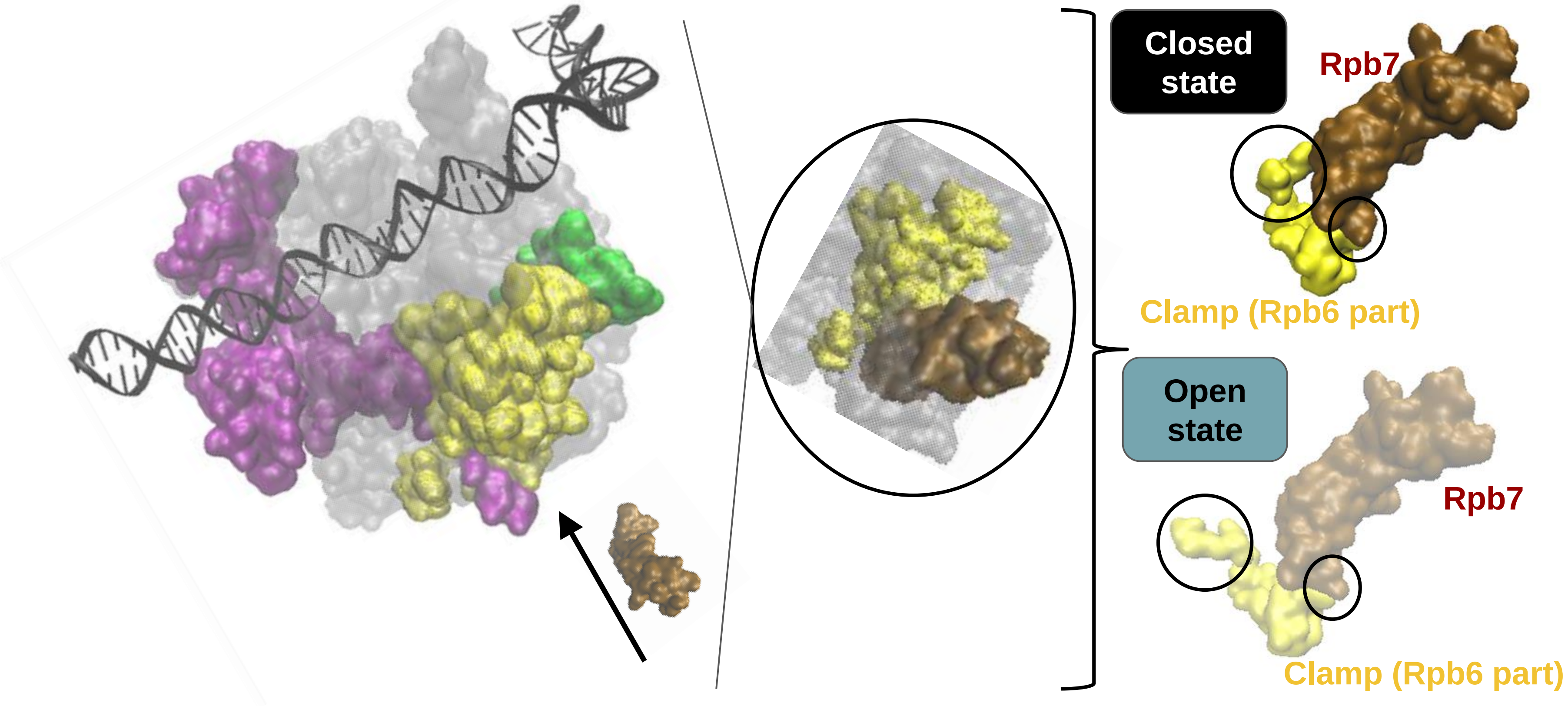
⇒

This change produces the rearrangement of Pol-II

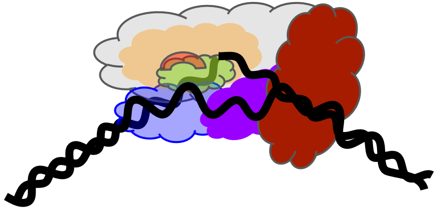
Open complex formation: Pol-II is rearranged to allow transcription initiation



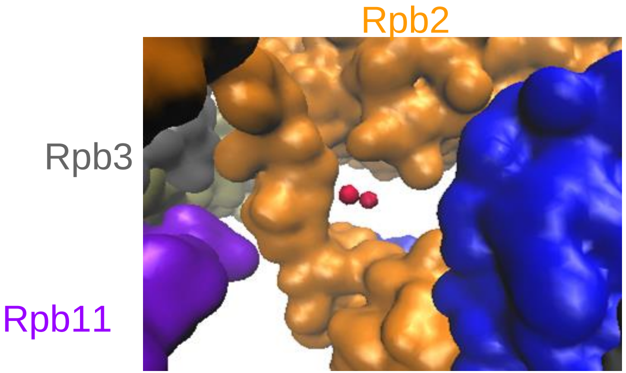
The mobility of the clamp is restricted by the presence of Rpb4/7 in the closed state

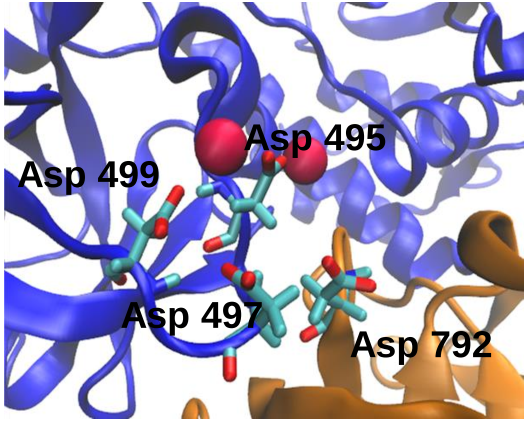


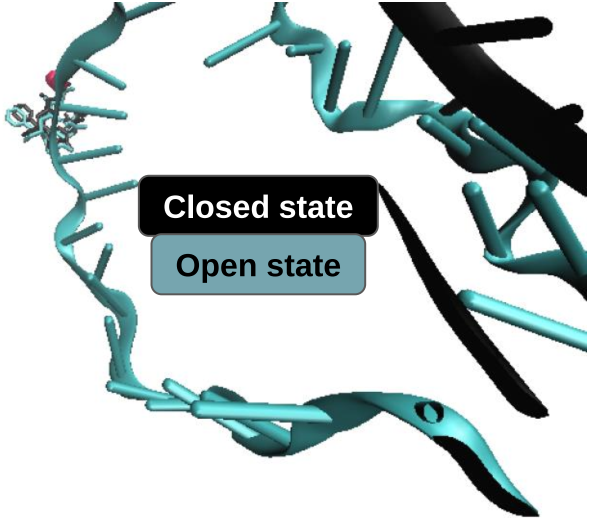
Open complex formation: Pol-II is rearranged to allow transcription initiation

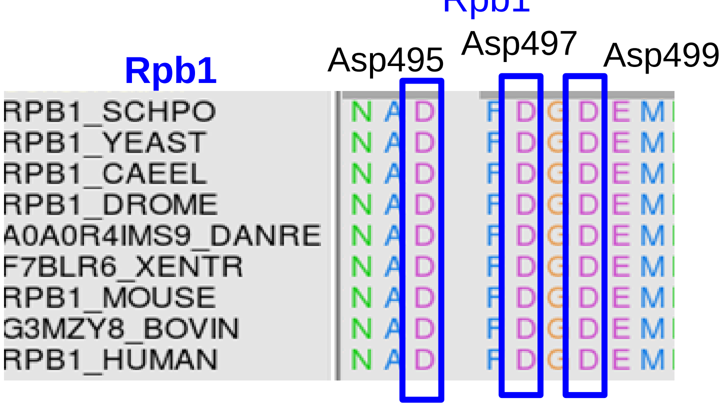


The core of Pol-II becomes close to the TSS upon promoter opening

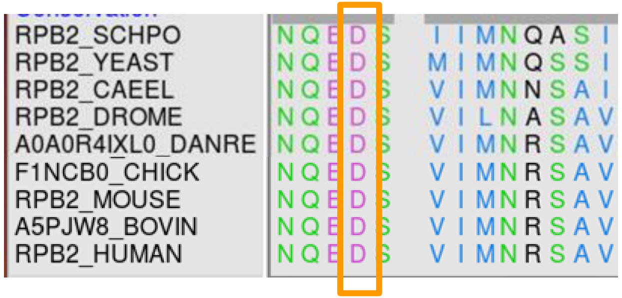








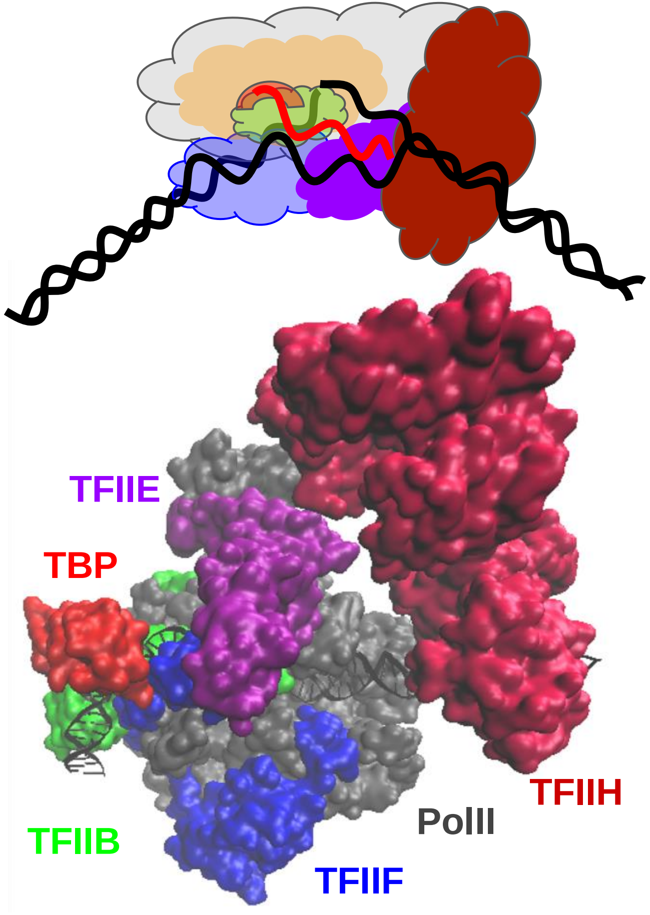
RPB1_SCHPO	N	A	D	F	D	G	D	E	M
RPB1_YEAST	N	A	D	F	D	G	D	E	M
RPB1_CAEEL	N	A	D	F	D	G	D	E	M
RPB1_DROME	N	A	D	F	D	G	D	E	M
A0A0R4IMS9_DANRE	N	A	D	F	D	G	D	E	M
F7BLR6_XENTR	N	A	D	F	D	G	D	E	M
RPB1_MOUSE	N	A	D	F	D	G	D	E	M
G3MZY8_BOVIN	N	A	D	F	D	G	D	E	M
RPB1_HUMAN	N	A	D	F	D	G	D	E	M

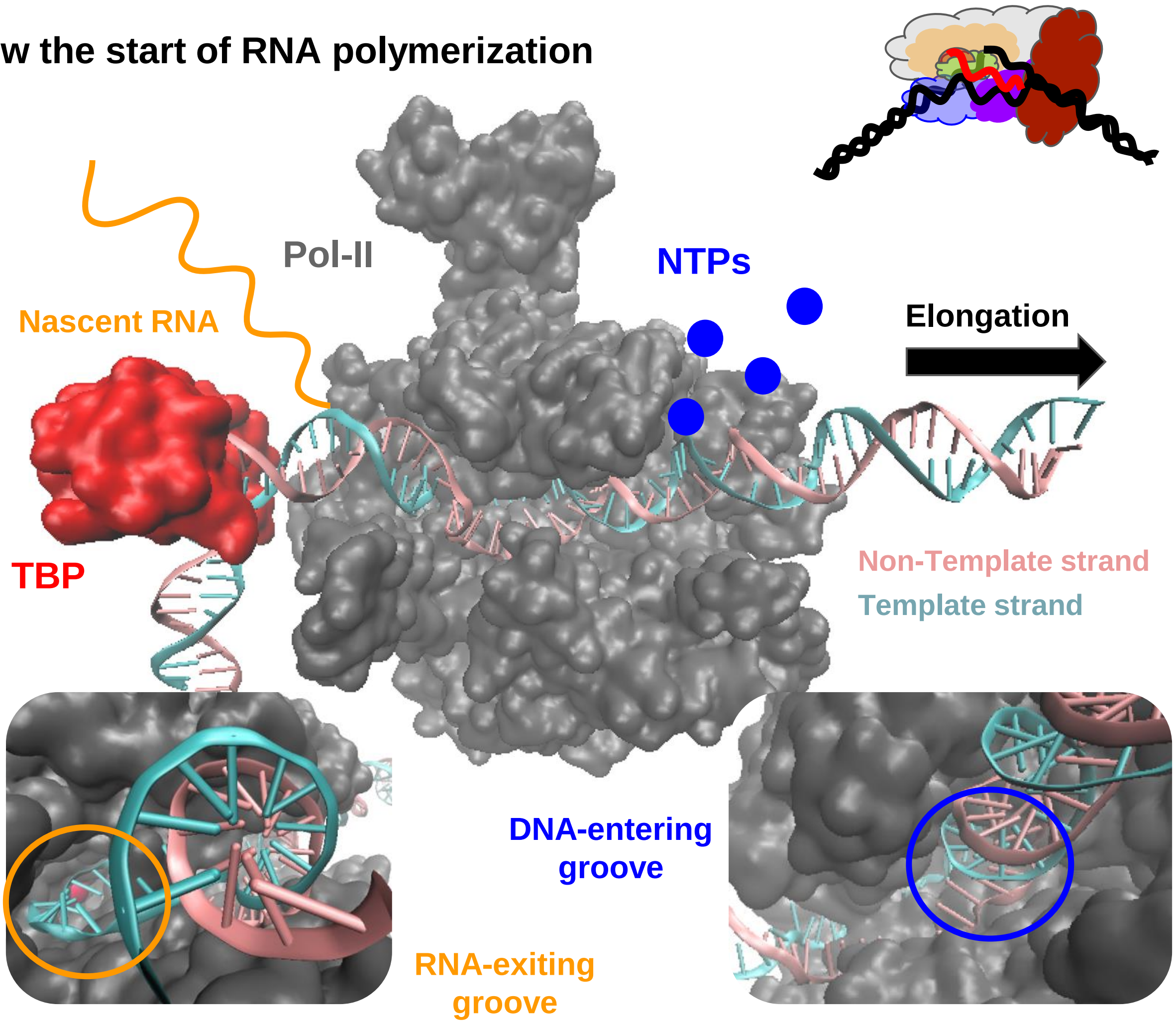


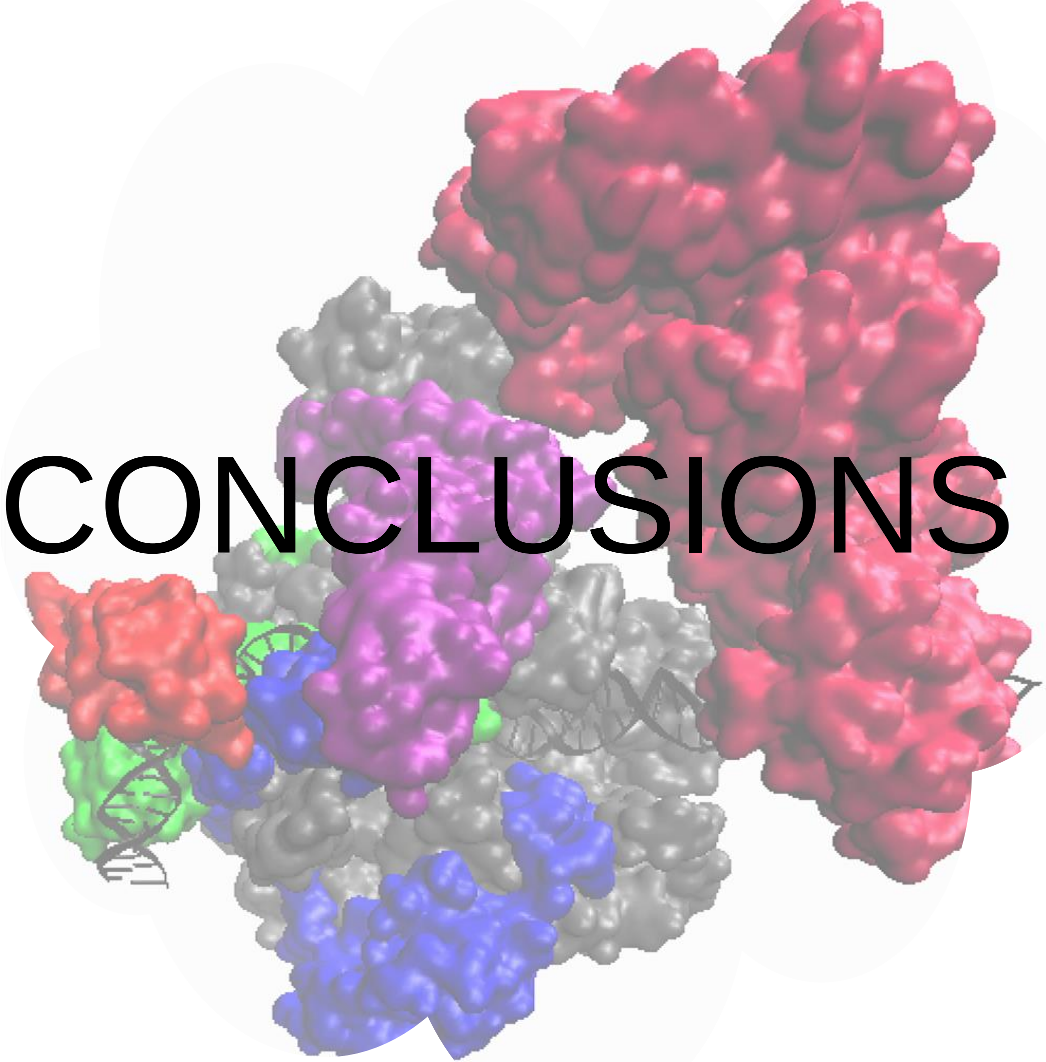
RPB2_SCHPO	N	Q	E	D	S	I	I	M	N	Q	A	S	I
RPB2_YEAST	N	Q	E	D	S	M	I	M	N	Q	S	S	I
RPB2_CAEEL	N	Q	E	D	S	V	I	M	N	N	S	A	I
RPB2_DROME	N	Q	E	D	S	V	I	L	N	A	S	A	V
A0A0R4IXL0_DANRE	N	Q	E	D	S	V	I	M	N	R	S	A	V
F1NCB0_CHICK	N	Q	E	D	S	V	I	M	N	R	S	A	V
RPB2_MOUSE	N	Q	E	D	S	V	I	M	N	R	S	A	V
A5PJW8_BOVIN	N	Q	E	D	S	V	I	M	N	R	S	A	V
RPB2_HUMAN	N	Q	E	D	S	V	I	M	N	R	S	A	V

The core is preserved across eukaryotes

All components are arranged to allow the start of RNA polymerization







CONCLUSIONS

Conclusions and future directions

- ➢ Although Pol-II is the catalytic part, the **GTFs are essential** for
    - ○ **Pol-II recruitment** to the promoter
    - ○ **Pol-II reordering** for transcription initiation
- ➢ Transcription initiation is a process **conserved across metazoans**
- ➢ **TBP, Rpb1, Rpb2 and TFIIB** are highly conserved across eukaryotes

- ➢ The **current data is limited** for understanding transcription initiation
    - ○ What happens with TATA-less promoters (90% of human promoters)?
    - ○ There is not high resolution structures for TFIIH and TFIID
    - ○ CTD is not characterized, so that we don’t know how it induces Pol-II rearrangement
    - ○ Large complexes as the PIC are still challenging to resolve

Bibliography

Noble, M. E., J. A. Endicott, N. R. Brown, and L. N. Johnson. 1997. “The Cyclin Box Fold: Protein Recognition in Cell-Cycle and Transcription Control.” *Trends in Biochemical Sciences* 22 (12): 482–87.

Nikolov, D. B., H. Chen, E. D. Halay, A. A. Usheva, K. Hisatake, D. K. Lee, R. G. Roeder, and S. K. Burley. 1995. “Crystal Structure of a TFIIB-TBP-TATA-Element Ternary Complex.” *Nature* 377 (6545): 119–28.

Ghosh, Mahua, Laura M. Elsby, Tapas K. Mal, Jane M. Gooding, Stefan G. E. Roberts, and Mitsuhiko Ikura. 2004. “Probing Zn2+-Binding Effects on the Zinc-Ribbon Domain of Human General Transcription Factor TFIIB.” *Biochemical Journal* 378 (Pt 2): 317–24.

Cramer, P., D. A. Bushnell, and R. D. Kornberg. 2001. “Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution.” *Science* 292 (5523): 1863–76.

Srivastava, Rakesh, and Seong Hoon Ahn. 2015. “Modifications of RNA Polymerase II CTD: Connections to the Histone Code and Cellular Function.” *Biotechnology Advances* 33 (6 Pt 1): 856–72.

Weipoltshammer, Klara, and Christian Schöfer. 2016. “Morphology of Nuclear Transcription.” *Histochemistry and Cell Biology* 145 (4): 343–58.

Cramer, P., D. A. Bushnell, J. Fu, A. L. Gnatt, B. Maier-Davis, N. E. Thompson, R. R. Burgess, A. M. Edwards, P. R. David, and R. D. Kornberg. 2000. “Architecture of RNA Polymerase II and Implications for the Transcription Mechanism.” *Science* 288 (5466): 640–49.

Kostrewa, D., M. E. Zeller, K-J Armache, M. Seizl, K. Leike, M. Thomm, and P. Cramer. 2009. “Crystal Structure of RNA Polymerase II in Complex with TFIIB.” doi:10.2210/pdb3k1f/pdb.

Kettenberger, Hubert, Karim-Jean Armache, and Patrick Cramer. 2003. “Architecture of the RNA Polymerase II-TFIIS Complex and Implications for mRNA Cleavage.” *Cell* 114 (3): 347–57.

Gnatt, A. L., P. Cramer, J. Fu, D. A. Bushnell, and R. D. Kornberg. 2001. “Structural Basis of Transcription: An RNA Polymerase II Elongation Complex at 3.3 A Resolution.” *Science* 292 (5523): 1876–82.

Bibliography

Fishburn, James, Eric Tomko, Eric Galburt, and Steven Hahn. 2015. “Double-Stranded DNA Translocase Activity of Transcription Factor TFIIH and the Mechanism of RNA Polymerase II Open Complex Formation.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (13): 3961–66.

Fan, Li, Andrew S. Arvai, Priscilla K. Cooper, Shigenori Iwai, Fumio Hanaoka, and John A. Tainer. 2006. “Conserved XPB Core Structure and Motifs for DNA Unwinding: Implications for Pathway Selection of Transcription or Excision Repair.” *Molecular Cell* 22 (1): 27–37.

Martinez-Rucobo, Fuensanta W., and Patrick Cramer. 2013. “Structural Basis of Transcription Elongation.” *Biochimica et Biophysica Acta* 1829 (1): 9–19.

Cheung, Alan C. M., and Patrick Cramer. 2012. “A Movie of RNA Polymerase II Transcription.” *Cell* 149 (7): 1431–37.

Fan, Li, and Kevin T. DuPrez. 2015. “XPB: An Unconventional SF2 DNA Helicase.” *Progress in Biophysics and Molecular Biology* 117 (2-3): 174–81.

Malik, Sohail, Henrik Molina, and Zhu Xue. 2017. “PIC Activation through Functional Interplay between Mediator and TFIIH.” *Journal of Molecular Biology* 429 (1): 48–63.

Dürr, Harald, Christian Körner, Marisa Müller, Volker Hickmann, and Karl-Peter Hopfner. 2005. “X-Ray Structures of the Sulfolobus Solfataricus SWI2/SNF2 ATPase Core and Its Complex with DNA.” *Cell* 121 (3). Elsevier: 363–73.

Okuda, Masahiko, Junichi Higo, Tadashi Komatsu, Tsuyoshi Konuma, Kenji Sugase, and Yoshifumi Nishimura. 2016. “Dynamics of the Extended String-Like Interaction of TFIIE with the p62 Subunit of TFIIH.” *Biophysical Journal* 111 (5): 950–62.

Hahn, Steven, and Stephen Buratowski. 2016. “Structural Biology: Snapshots of Transcription Initiation.” *Nature* 533 (7603): 331–32.

He, Yuan, Jie Fang, Dylan J. Taatjes, and Eva Nogales. 2013. “Structural Visualization of Key Steps in Human Transcription Initiation.” *Nature* 495 (7442): 481–86.

Bibliography

Grünberg, Sebastian, Linda Warfield, and Steven Hahn. 2012. “Architecture of the RNA Polymerase II Preinitiation Complex and Mechanism of ATP-Dependent Promoter Opening.” *Nature Structural & Molecular Biology* 19 (8): 788–96.

Chen, Hung-Ta, Linda Warfield, and Steven Hahn. 2007. “The Positions of TFIIF and TFIIE in the RNA Polymerase II Transcription Preinitiation Complex.” *Nature Structural & Molecular Biology* 14 (8): 696–703.

Okuda, Masahiko, Yoshinori Watanabe, Hideyasu Okamura, Fumio Hanaoka, Yoshiaki Ohkuma, and Yoshifumi Nishimura. 2000. “Structure of the Central Core Domain of TFIIEβ with a Novel Double-stranded DNA-binding Surface.” *The EMBO Journal* 19 (6). EMBO Press: 1346–56.

Miwa, Kohei, Rieko Kojima, Takayuki Obita, Yoshiaki Ohkuma, Yasushi Tamura, and Mineyuki Mizuguchi. 2016. “Crystal Structure of Human General Transcription Factor TFIIE at Atomic Resolution.” *Journal of Molecular Biology* 428 (21): 4258–66.

Okuda Masahiko, and Nishimura Yoshifumi. 2008. “[Structural biology of a general transcription factor, TFIIE and its interaction mode with TFIIH].” *Seikagaku. The Journal of Japanese Biochemical Society* 80 (6): 501–10.

Okuda, Masahiko, Aki Tanaka, Manami Satoh, Shoko Mizuta, Manabu Takazawa, Yoshiaki Ohkuma, and Yoshifumi Nishimura. 2008. “Structural Insight into the TFIIE–TFIIH Interaction: TFIIE and p53 Share the Binding Region on TFIIH.” *The EMBO Journal* 27 (7). EMBO Press: 1161–71.

Ohkuma, Yoshiaki, Hideki Sumimoto, and Others. 1991. “Structural Motifs and Potential (Omicron) Homologies in the Large Subunit of Human General Transcription Factor TFIIE.” *Nature* 354 (6352). Nature Publishing Group: 398.

Peterson, M. G., J. Inostroza, M. E. Maxon, O. Flores, A. Admon, D. Reinberg, and R. Tjian. 1991. “Structure and Functional Properties of Human General Transcription Factor IIE.” *Nature* 354 (6352): 369–73.

Blombach, Fabian, Katherine L. Smollett, Dina Grohmann, and Finn Werner. 2016. “Molecular Mechanisms of Transcription Initiation-Structure, Function, and Evolution of TFE/TFIIE-Like Factors and Open Complex Formation.” *Journal of Molecular Biology* 428 (12): 2592–2606.

Kostrewa, Dirk, Mirijam E. Zeller, Karim-Jean Armache, Martin Seizl, Kristin Leike, Michael Thomm, and Patrick Cramer. 2009. “RNA Polymerase II-TFIIB Structure and Mechanism of Transcription Initiation.” *Nature* 462 (7271): 323–30.

Bibliography

Murakami, Kenji, Kuang-Lei Tsai, Nir Kalisman, David A. Bushnell, Francisco J. Asturias, and Roger D. Kornberg. 2015. “Structure of an RNA Polymerase II Preinitiation Complex.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (44): 13543–48.

Egly, Jean-Marc, and Frédéric Coin. 2011. “A History of TFIIH: Two Decades of Molecular Biology on a Pivotal Transcription/repair Factor.” *DNA Repair* 10 (7): 714–21.

He, Yuan, Chunli Yan, Jie Fang, Carla Inouye, Robert Tjian, Ivaylo Ivanov, and Eva Nogales. 2016. “Near-Atomic Resolution Visualization of Human Transcription Promoter Opening.” *Nature* 533 (7603): 359–65.

Plaschka, C., M. Hantsche, C. Dienemann, C. Burzinski, J. Plitzko, and P. Cramer. 2016. “Transcription Initiation Complex Structures Elucidate DNA Opening.” *Nature* 533 (7603): 353–58.

Hantsche, Merle, and Patrick Cramer. 2016. “The Structural Basis of Transcription: 10 Years After the Nobel Prize in Chemistry.” *Angewandte Chemie* 55 (52): 15972–81.

Sainsbury, Sarah, Carrie Bernecky, and Patrick Cramer. 2015. “Structural Basis of Transcription Initiation by RNA Polymerase II.” *Nature Reviews. Molecular Cell Biology* 16 (3): 129–43.

Plaschka, Clemens, Kayo Nozawa, and Patrick Cramer. 2016. “Mediator Architecture and RNA Polymerase II Interaction.” *Journal of Molecular Biology* 428 (12): 2569–74.

Cramer, Patrick. 2016. “Structure Determination of Transient Transcription Complexes.” *Biochemical Society Transactions* 44 (4): 1177–82.

Yao, Jie. 2017. “Imaging Transcriptional Regulation of Eukaryotic mRNA Genes: Advances and Outlook.” *Journal of Molecular Biology* 429 (1): 14–31.

Zhang, Yinfeng, Saman M. Najmi, and David A. Schneider. 2016. “Transcription Factors That Influence RNA Polymerases I and II: To What Extent Is Mechanism of Action Conserved?” *Biochimica et Biophysica Acta* 1860 (2): 246–55.

Bibliography

Jonge, Wim J. de, Eoghan O'Duibhir, Philip Lijnzaad, Dik van Leenen, Marian Ja Groot Koerkamp, Patrick Kemmeren, and Frank Cp Holstege. 2016. "Molecular Mechanisms That Distinguish TFIID Housekeeping from Regulatable SAGA Promoters." The EMBO Journal, December. doi:10.15252/embj.201695621.

Hahn, Steven. 2004. "Structure and Mechanism of the RNA Polymerase II Transcription Machinery." Nature Structural & Molecular Biology 11 (5): 394–403.

Choder, Mordechai. 2004. "Rpb4 and Rpb7: Subunits of RNA Polymerase II and beyond." Trends in Biochemical Sciences 29 (12): 674–81.

Cramer, P., D. A. Bushnell, J. Fu, A. L. Gnatt, B. Maier-Davis, N. E. Thompson, R. R. Burgess, A. M. Edwards, P. R. David, and R. D. Kornberg. 2000. "Architecture of RNA Polymerase II and Implications for the Transcription Mechanism." Science 288 (5466): 640–49.

Cramer, Patrick. 2004. "RNA Polymerase II Structure: From Core to Functional Complexes." Current Opinion in Genetics & Development 14 (2): 218–26.

Fishburn, James, and Steven Hahn. 2012. "Architecture of the Yeast RNA Polymerase II Open Complex and Regulation of Activity by TFIIF." Molecular and Cellular Biology 32 (1): 12–25.

Luse, Donal S. 2012. "Rethinking the Role of TFIIF in Transcript Initiation by RNA Polymerase II." Transcription 3 (4): 156–59.

Čabart, Pavel, Andrea Újvári, Mahadeb Pal, and Donal S. Luse. 2011. "Transcription Factor TFIIF Is Not Required for Initiation by RNA Polymerase II, but It Is Essential to Stabilize Transcription Factor TFIIB in Early Elongation Complexes." Proceedings of the National Academy of Sciences of the United States of America 108 (38): 15786–91.

Baumann, Martina, Jens Pontiller, and Wolfgang Ernst. 2010. "Structure and Basal Transcription Complex of RNA Polymerase II Core Promoters in the Mammalian Genome: An Overview." Molecular Biotechnology 45 (3): 241–47.

Bibliography

UNIPROT: http://www.uniprot.org/

PDB: http://www.rcsb.org/pdb/home/home.do

PFAM: http://pfam.xfam.org/

INTERPROT: https://www.ebi.ac.uk/interpro/

PISA: http://www.ebi.ac.uk/pdbe/pisa/

Multiple Choice Questions

About TBP, choose which of the following statements is **FALSE**.

a) TBP’s C-term domain is formed by two repeated “shadle-shaped” structures.
b) TBP is the first factor that binds to the promoter region.
c) TBP is the last factor that binds the complex.
d) TBP mainly binds DNA through electrostatic interactions and hydrogen bonds.
e) TBP’s N-term domain is not conserved.

TFIIB interacts with…

a) TBP
b) DNA
c) RNA polymerase II
d) TFIIF
e) All are correct

Regarding to TFIIB. It is true that:

a) It has a Zinc finger domain
b) It has two cyclin like domains
c) a) and b) are correct
d) It has a winged helix domain
e) Non of them are true

Multiple Choice Questions

Regarding TFIIE. It is true that:

a) There’s a string-like interaction between its α subunit and TFIIH.
b) It has no importance in recruiting TFIIH.
c) It phosphorylates Pol-II to allow transcription initiation.
d) Transcription can occur without TFIIE.
e) It forms up the core of the RNA polymerase.

About TFIIH, choose which of the following statements is **FALSE**.

a) It cointains two modules with helicase and kinase activity.
b) It phosphorylates Pol-II to allow transcription inititation.
c) The electrostatic interaction between it’s LYS387 and ARG137 of TFIIE is key for the assembly to the pre-inititation complex.
d) It has not been charachterized yet at high resolution.
e) We do not have structural information about some of it’s subunits.

Regarding the events following promoter opening:

a) Jaw lobe, clamp, wall and core are the parts of Pol-II relevant for function.
b) The phosphorylation of the CTD of Rpb1 is on several His and Asp residues.
c) The CTD of Rpb1 is a very structured motif.
d) The CTD of Rpb1 is not important for integrating regulatory inputs.
e) Promoter opening is not important for transcription initiation.

Multiple Choice Questions

About the RNA polymerase II, choose which of the following statements is **FALSE.**

a) It is a complex formed by 7 subunits.
b) Rpb1 is the bigger subunit of the complex.
c) Rpb2 is in contact with the magnesium ion of the core.
d) Rpb1 is in contact with the magnesium ion of the core.
e) Rpb6 forms part of the clamp element.

About TFIIF, choose the **CORRECT** answer:

a) It is formed by 2 subunits (alpha and beta).
b) Has a triple barrel domain.
c) Has a winged helix domain.
d) Its interacting mainly with Rpb2 subunit of the RNA polymerase II.
e) All are correct.

Rpb7 has an important role in…

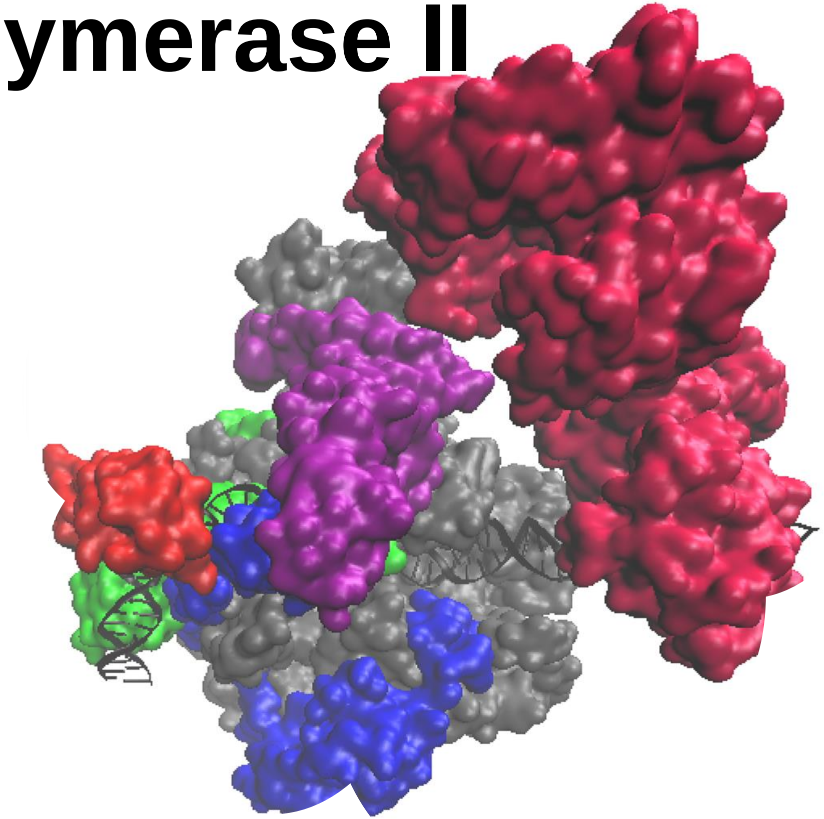
a) interacting with the core.
b) the elongation process.
c) binding the DNA in the closed state.
d) locking the clamp in the initiation process.
e) the termination process.

Multiple Choice Questions

Regarding the Core PreIniciationComplex (PIC), is formed by:

a) TBP
b) TFIIF
c) TFIIB
d) RNA polymerase II
e) All are correct

Structural basis of transcription initiation by RNA polymerase II



Josep Argerich Bergedà
Adrián González Sánchez
Miquel Àngel Schikora Tamarit