

Механизмы регуляции транскрипции

(Лекция 2)

Елена Васильевна Игнатьева

*к.б.н., с.н.с. лаборатории эволюционной биоинформатики и
теоретической генетики
ИЦиГ СО РАН*

ОСОБЕННОСТИ ТРАНСКРИПЦИИ У ЭУКАРИОТ:

- ✓ **1) У эукариот транскрипция осуществляется при участии РНК-полимераз трех типов, которые обозначаются как pol I, pol II, и pol III. Можно выделить три группы генов:**
 - pol I транскрибируемые: гены рРНК (18S, 5.8S и 28S)
 - pol II транскрибируемые: все белок-кодирующие гены + гены многих мяРНК + многие миРНК
 - pol III транскрибируемые : гены тРНК + гены низкомолекулярных РНК (5S рРНК, и др.)
- ✓ **2) Каждая РНК-полимераза имеет сложное, многосубъединичное строение, например, pol II включает более 10 субъединиц;**
- ✓ **3) Для точной посадки РНК-полимераз на участок, прилежащий к старту транскрипции, нужны вспомогательные белки – базальные (или общие) транскрипционные факторы;**
- 4) Для достижения необходимого уровня транскрипции каждого гена, помимо базальных (общих) транскрипционных факторов, требуются другие белки (транскрипционные факторы, коактиваторы, корепрессоры, медиаторы и др.);**
- 5) Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение**
- 6) Транскрипция генов эукариот зависит от состояния хроматина и плотности нуклеосомной упаковки.**

ОСОБЕННОСТИ ТРАНСКРИПЦИИ У ЭУКАРИОТ:

1) У эукариот транскрипция осуществляется при участии РНК-полимераз трех типов, которые обозначаются как pol I, pol II, и pol III. Можно выделить три группы генов:
- pol I транскрибируемые: гены рРНК (18S, 5.8S и 28S)
- pol II транскрибируемые: все белок-кодирующие гены + гены многих мяРНК + многие миРНК
- pol III транскрибируемые : гены тРНК + гены низкомолекулярных РНК (5S рРНК, и др.)

2) Каждая РНК-полимераза имеет сложное, многосубъединичное строение, например, pol II включает более 10 субъединиц;

3) Для точной посадки РНК-полимераз на участок, прилежащий к старту транскрипции, нужны вспомогательные белки – базальные (или общие) транскрипционные факторы;

4) Для достижения необходимого уровня транскрипции каждого гена, помимо базальных (общих) транскрипционных факторов, требуются другие белки (транскрипционные факторы, коактиваторы, корепрессоры, медиаторы и др.);

5) Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение



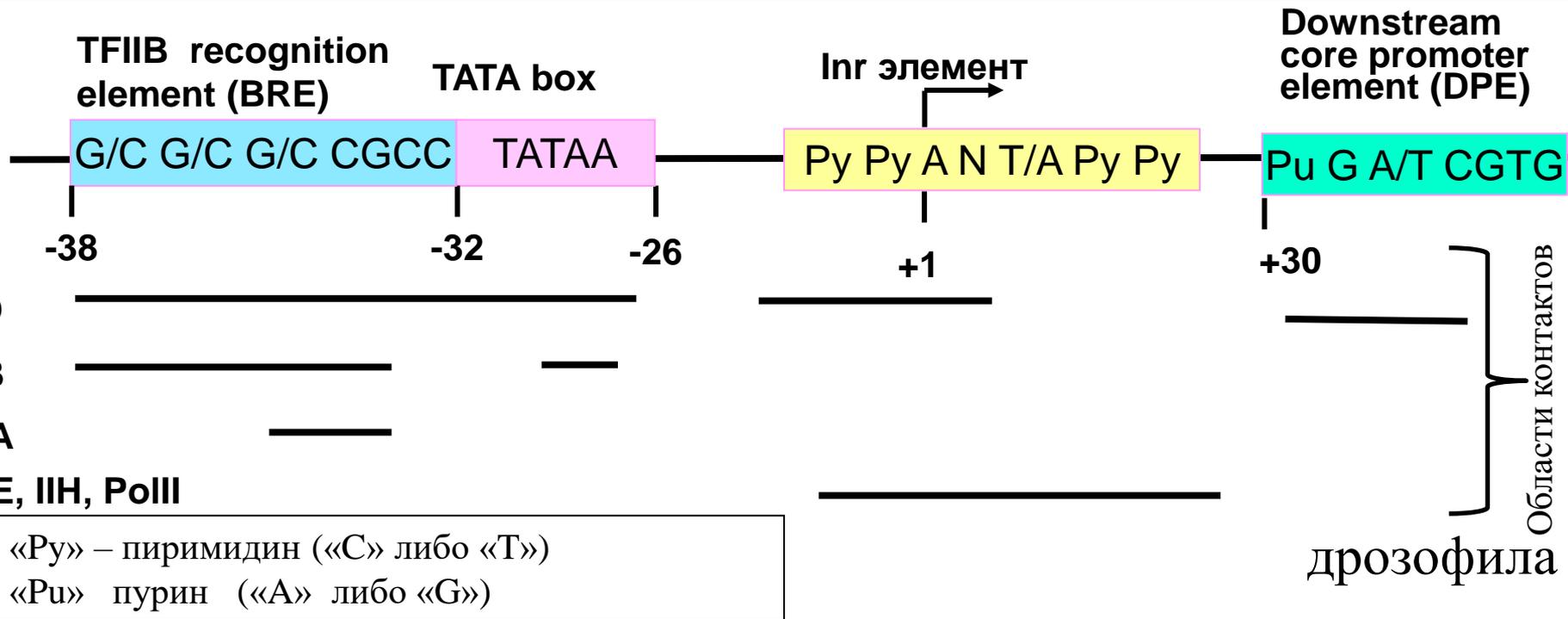
6) Транскрипция генов эукариот зависит от состояния хроматина и плотности нуклеосомной упаковки.

Коровые

5) Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение

Коровый (базальный) промотор - участок ДНК минимальной длины в окрестностях старта транскрипции гена, необходимый для формирования прединициаторного комплекса. Понятие корового промотора не содержит строгих данных о его размерах и точной локализации относительно старта транскрипции. Коровым промотором именуют либо очень короткий участок (-35 и + 35 относительно старта транскрипции) либо более протяженный район, например, от -60 до+40

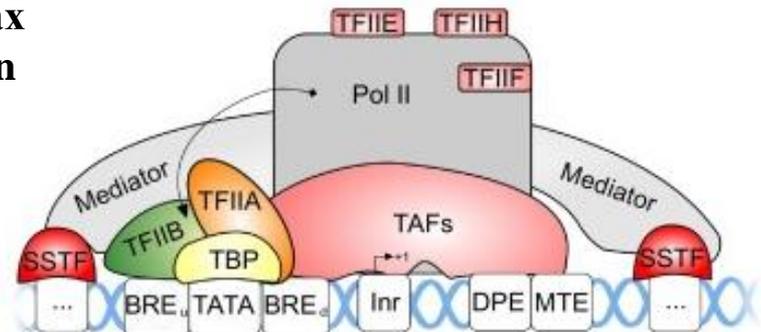
ТИПИЧНЫЕ ЭЛЕМЕНТЫ КОРОВОГО ПРОМОТОРА Pol II ТРАНСКРИБИРУЕМЫХ ГЕНОВ (2000 г.).



Transcription Regulation in Eukaryotes. Caley M., Smale S.T. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2000, 640 p. (p.8)

Модель формирования ПИК на коровых промоторах белок-кодирующих генов многоклеточных (metazoan mRNA transcription units)

MTE - motif ten element
SSTF - site-specific transcription factors
TAF - TBP-associated factors



Dergai O, Hernandez N. How to Recruit the Correct RNA Polymerase? Lessons from snRNA Genes. Trends Genet. 2019;35(6):457-469.

2018: Известные элементы корового промотора Pol II транскрибируемых генов и регуляторные белки, которые с ними связываются.

Table 1

Known core-promoter motifs and the (general) transcription factors that bind to them.

Core promoter motif	Sequence logo	Consensus sequence ^a	Position relative to ISS	Bound by	Fly	Human
TATA-box		TATAWAWR ^{49,241}	-31 to -24	TBP (REFS ^{53,242})	+	+
Inr (fly)		TCAGTY ^{56,243}	-5 to -2	TAF1/TAF2 (REF. ⁵⁷)	+	-
Inr (human)		YR ⁴⁵	-1 to +1	N.A.	-	+
		BBCABW ⁵⁸	-3 to +3			
DPE		RGWCGTG ⁵⁹ RGWYVT ⁶¹	+28 to +34 +28 to +33	TAF6 and TAF9 (REF. ⁶⁰) possibly TAF1 (REF. ⁵⁵)	+	Possibly rarely
		GCGWCKGGTTS ⁵¹	+24 to +32		+	-
MTE		CSARCSSAACGS ⁶³	+18 to +29	possibly TAF1 and TAF2 (REF. ⁵⁵)	+	-
Ohler 1		YGGTCACACTR ⁵¹	-60 to -1	M1BP (REF. ²⁴⁴)	+	-
Ohler 6		KTYRGTAIWTIT ⁵¹	-100 to -1	N.A.	+	-
Ohler 7		KNNCAKCNCTRNY ⁵¹	-60 to +20	N.A.	+	-
DRE		WATCGATW ²⁴⁵	-100 to -1	DREF (REF. ²⁴⁵)	+	+
TCT		YYCTTTY ²⁴⁶	-2 to +6	N.A.	+	+
BREu		SSRCGC ⁶⁴	-38 to -32	TFIIB (REF. ⁶⁴)	+	+
BREd		RIDKKKK ⁶⁵	-23 to -17	TFIIB (REF. ⁶⁵)	+	+
DCE (I,II,III)	N.A.	CTTC CTGT AGC ⁶⁶	+6 to +11+16 to +21+30 to +34	TAF1 (REF. ⁶⁷)	-	+
XCPE1	N.A.	DSGYGGRASM ²⁴⁷	-8 to +2	N.A.	?	+
XCPE2	N.A.	VCYCRITRCMY ²⁴⁸	-9 to +2	N.A.	?	+
Pause button		KCGRWCG ¹⁵⁵	+25 to +35	N.A.	+	?

BREd, TFIIB recognition element, downstream; BREu, TFIIB recognition element, upstream; DCE, downstream core element; DPE, downstream promoter element; DRE, DNA replication- related element; Dref, DNA replication- related element factor; Inr, initiator; M1BP, motif 1-binding protein; MTE, motif ten element; NA, not available; TAF, transcription initiation factor TFIID subunit or TBP-associated factor; TBP, TATA-box-binding protein; TCT, polypyrimidine initiator; TFIIB, general transcription factor IIB; TSS, transcription start site; XCPE, X core-promoter element.

B = C or G or T; D = A or G or T; K = G or T; M = A or C; R = A or G; S = C or G; V = A or C or G; W = A or T; Y = C or T.

ТИПИЧНЫЕ ЭЛЕМЕНТЫ КОРОВОГО ПРОМОТОРА Pol II ТРАНСКРИБИРУЕМЫХ ГЕНОВ.

Типы коровых промоторов:

TATA+ Inr+;

TATA+ Inr-;

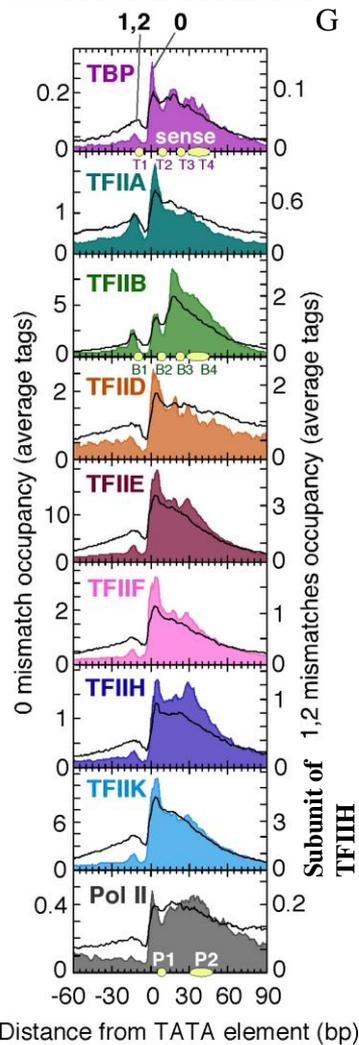
TATA- Inr+;

TATA- Inr-;

TATA бокс выявляется в 33-35% коровых промоторов человека и в 33-43% коровых промоторов дрозофилы

Расположение базальных транскрипционных факторов в районе промоторов генов дрожжей

Mismatches to TATAWAWR, где W - A or T, R - A or G



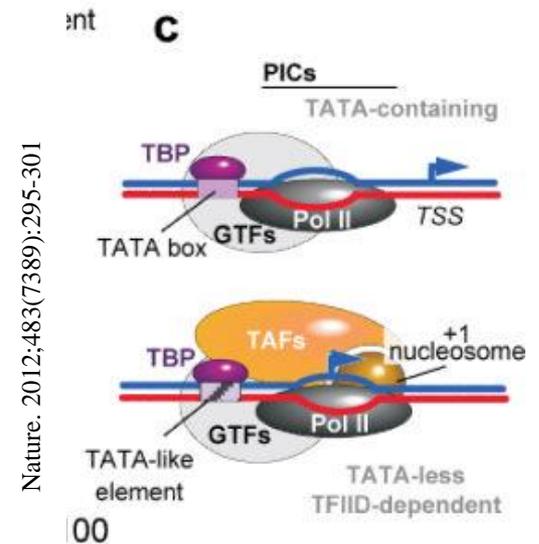
Цветные фигуры - TATA--
-containing (676 гена)

Черные линии - TATA-less
(3,229 гена)

Видны различия в
расположении компонентов
ПИК относительно позиции
мотива TATA в TATA-
содержащих и TATA-
несодержащих промоторах

Позиционирование БТФ в районе промоторов генов исследовали методом ChIP-exo (сшивка белка с ДНК, последующее расщепление свободной (незащищенной белками ДНК), далее денатурации комплексов ДНК и-белок и секвенирование)

Модель строения комплекса ПИК для TATA-содержащих и TATA-несодержащих промоторов (дрожжи)



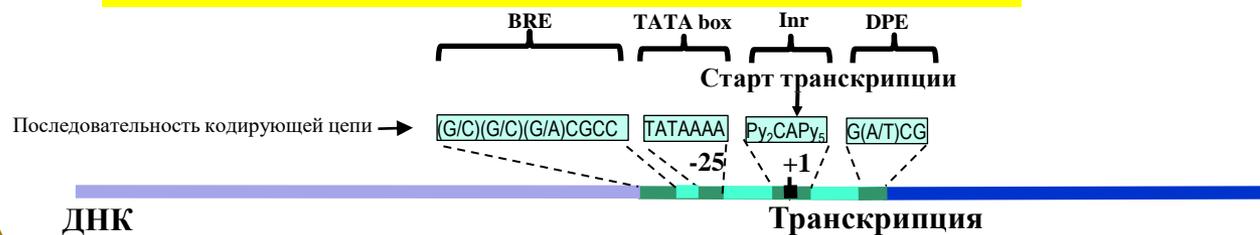
Nature. 2012;483(7389):295-301

Inr опознается комплексом ТВР-ассоциированных факторов TAF1 и TAF2. В промоторах, не содержащих TATA-бокса, Inr способен обеспечивать правильную локализацию белка ТВР в районе старта транскрипции. Показано, что TAFs, входящие в TFIID, взаимодействуют с Inr, после чего ТВР позиционируется в области -30 относительно старта транскрипции и связывается с этой областью

Dev Biol. 2010. V. 339. P. 225-229

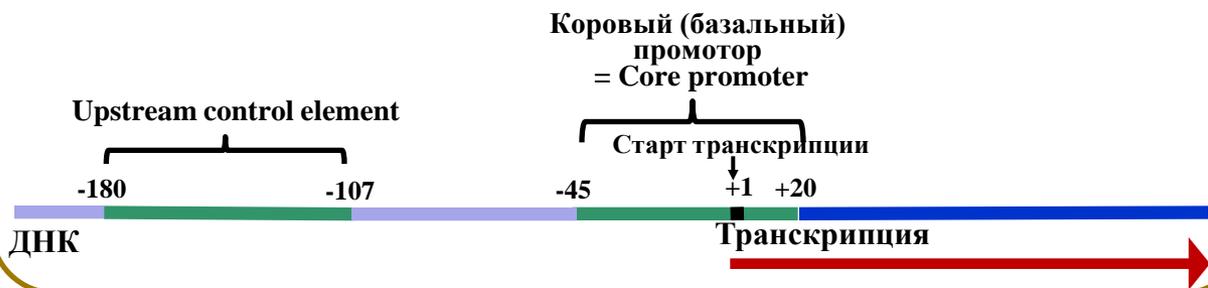
Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение

Промоторы генов, транскрибируемых РНК-полимеразой II



Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение

Промоторы генов, транскрибируемых РНК-полимеразой I



Промоторы генов, транскрибируемых РНК-полимеразой II

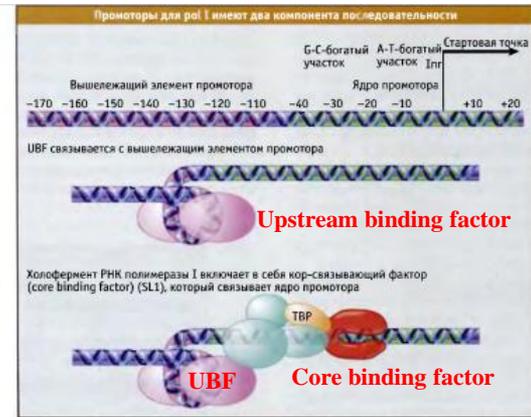
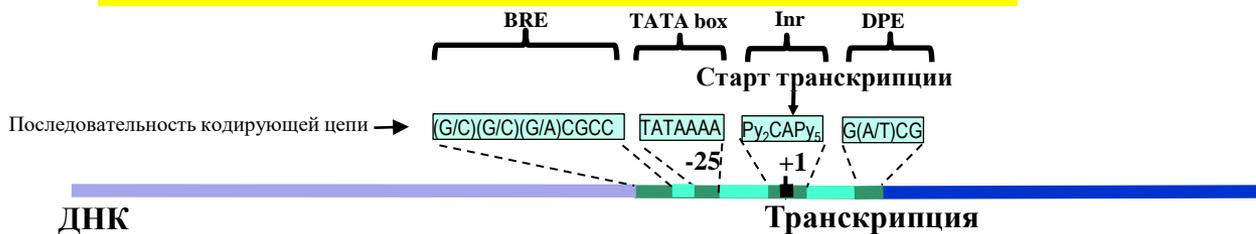


Рис. 24.3 Транскрипционная единица для РНК-полимеразы I включает в себя ядро промотора, удаленное от вышележащего промоторного элемента примерно на 70 пн. Связывание UBF с UPE увеличивает способность кор-связывающего фактора соединяться с ядром промотора. Кор-связывающий фактор (SL1) позиционирует РНК-полимеразу I на стартовой точке

Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение

Промоторы генов, транскрибируемых РНК-полимеразой I

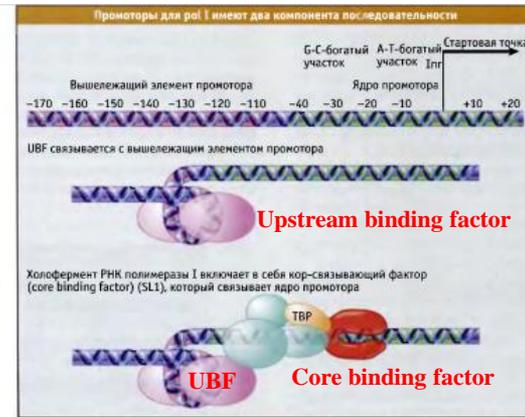
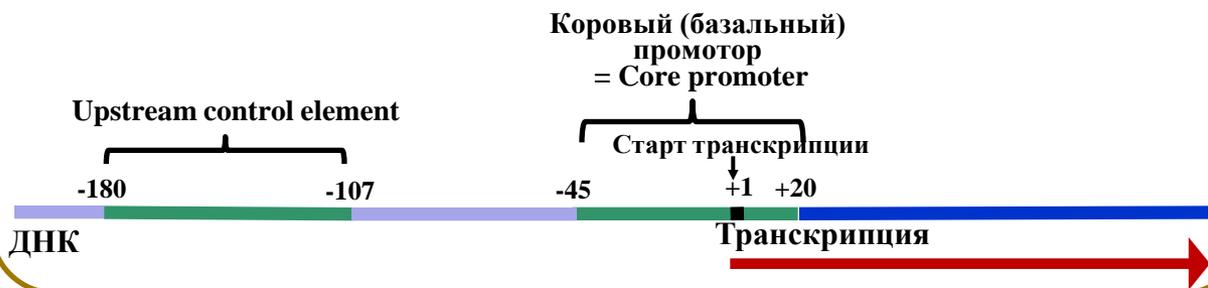
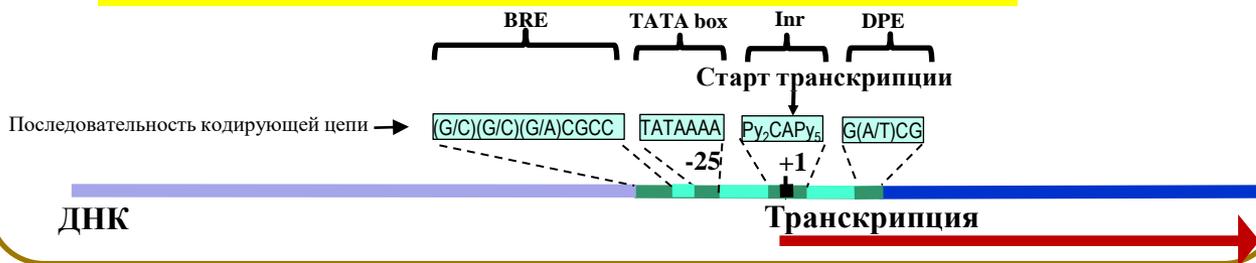
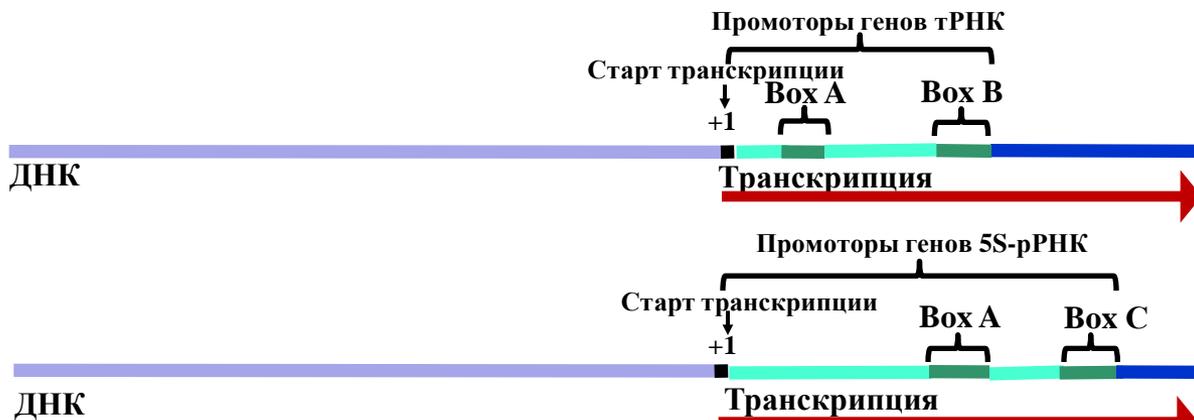


Рис. 24.3 Транскрипционная единица для РНК-полимеразы I включает в себя ядро промотора, удаленное от вышележащего промоторного элемента примерно на 70 пн. Связывание UBF с UPE увеличивает способность кор-связывающего фактора соединяться с ядром промотора. Кор-связывающий фактор (SL1) позиционирует РНК-полимеразу I на стартовой точке

Промоторы генов, транскрибируемых РНК-полимеразой II



Промоторы генов, транскрибируемых РНК-полимеразой III

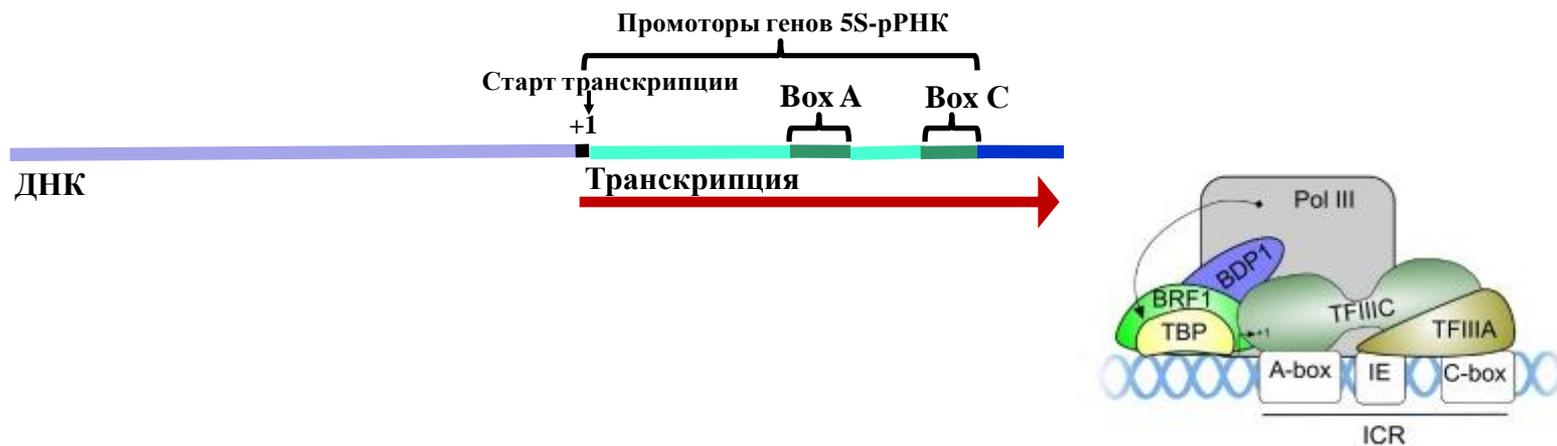
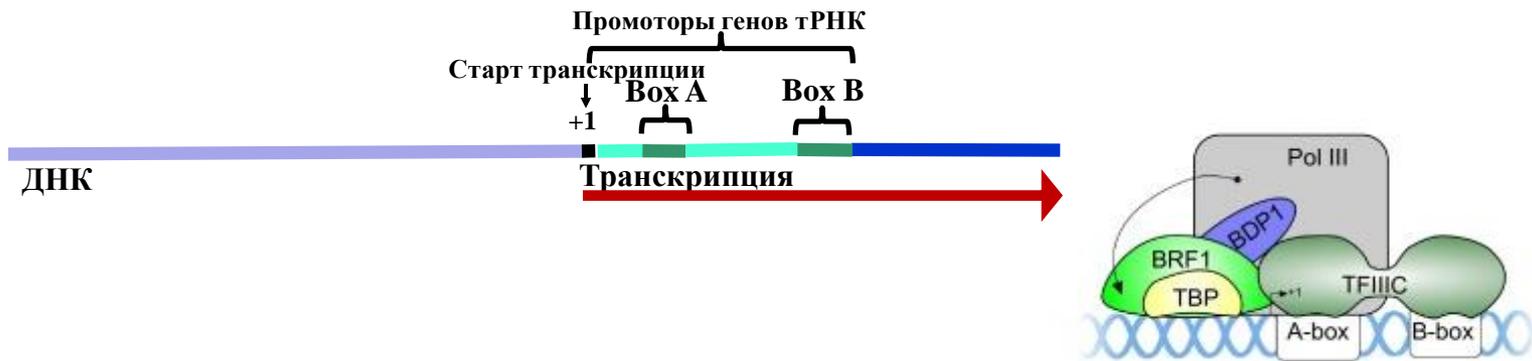


А также:

Консервативные мотивы промоторов генов мяРНК находятся выше старта транскрипции

Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение

Промоторы генов, транскрибируемых РНК-полимеразой III



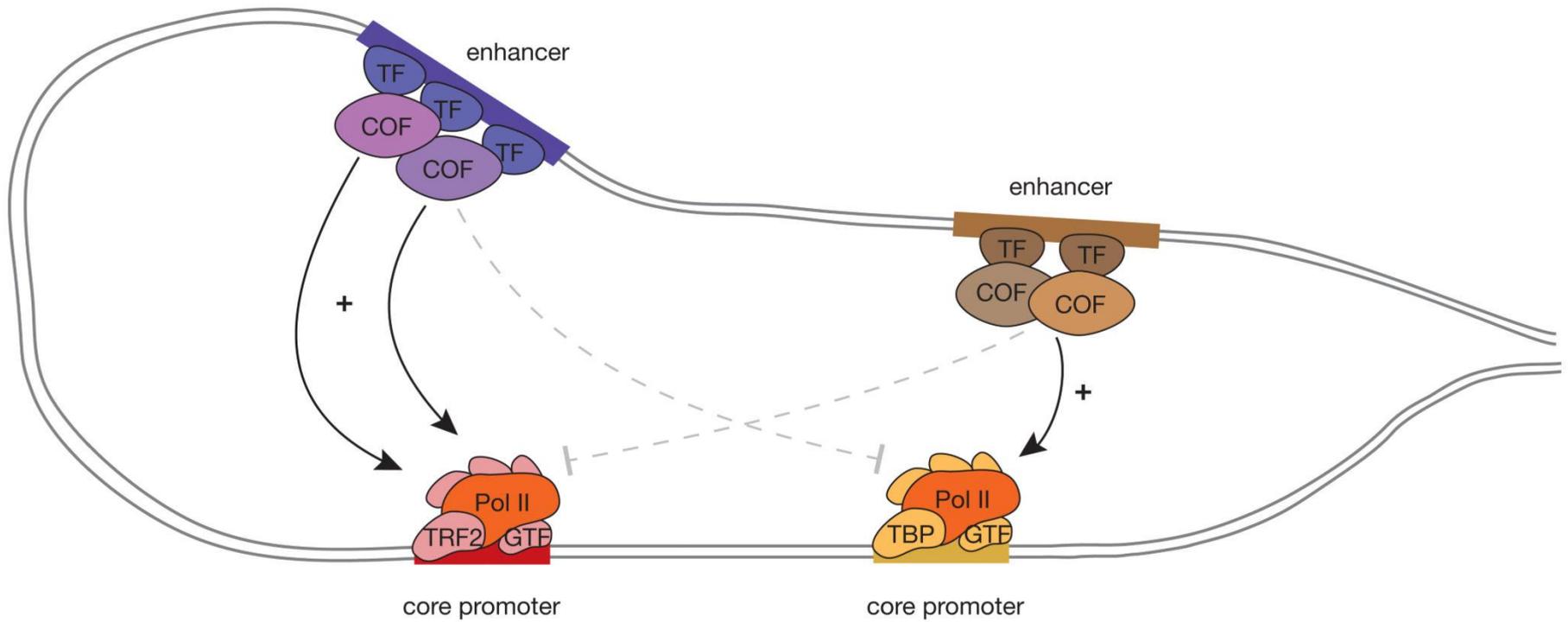
BDP1 (= TFIIB) –B double prime 1, subunit of RNA polymerase III; transcription initiation factor IIIB
BRF1 - TFIIB-related factor 1 (B - Related Factor 1); ICR - internal control region;
IE - intermediate element

Промоторный район, энхансеры, сайленсеры -
РЕГУЛЯТОРНЫЕ ЕДИНИЦЫ

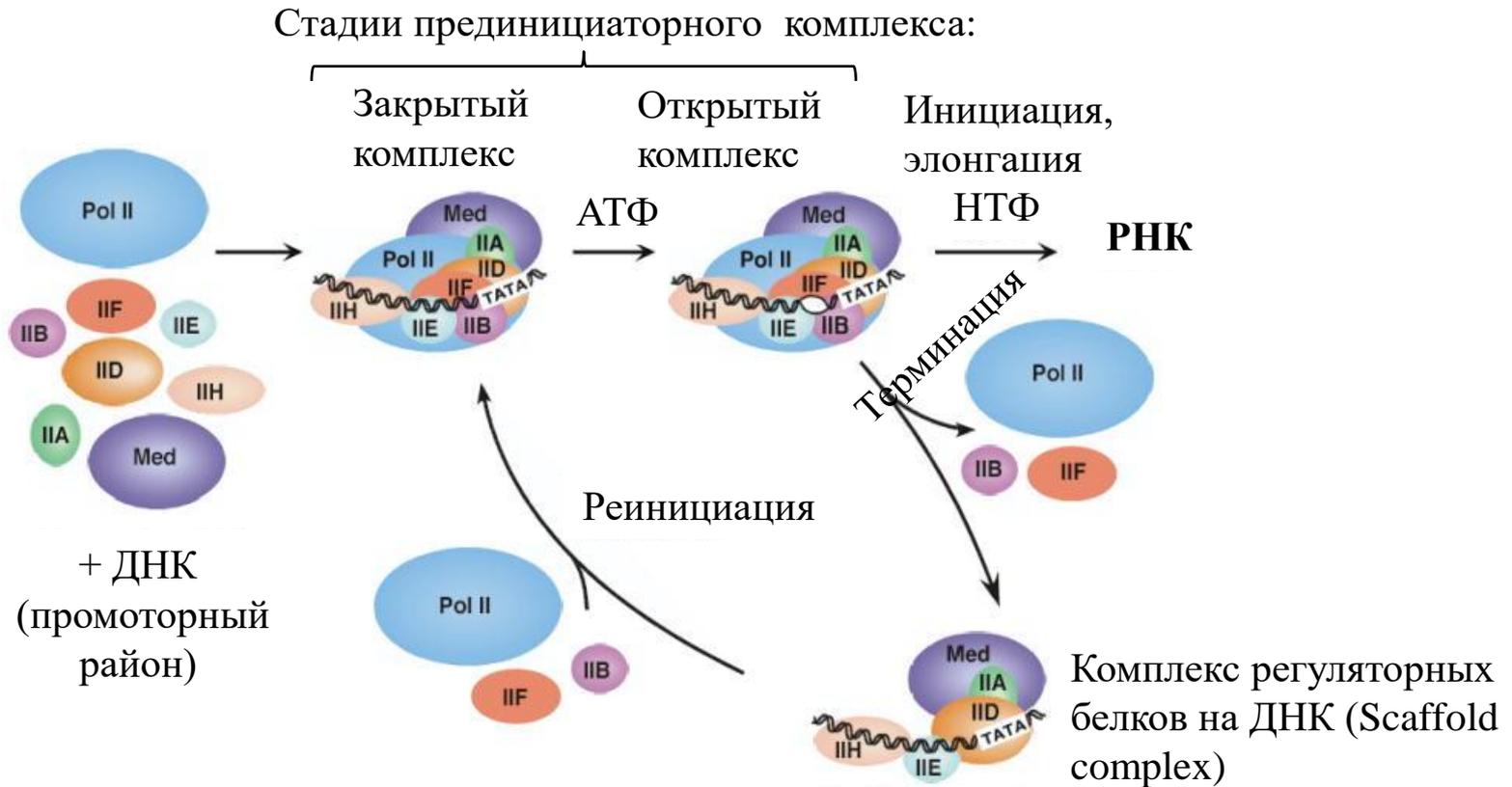
Регуляторные единицы – включают сайты
связывания транскрипционных факторов.

Транскрипционные факторы – белки, специфически
связывающиеся с ДНК и регулирующие
транскрипцию.

Коактиваторы и корепрессоры, медиаторы – белки,
не взаимодействующие с ДНК. Они также
регулируют транскрипцию, опосредуя влияние
транскрипционных факторов.



ТРАНСКРИПЦИОННЫЙ ЦИКЛ, ОСУЩЕСТВЛЯЕМЫЙ РНК-ПОЛИМЕРАЗОЙ II.



Закрытый комплекс: На этом этапе комплекс, включающий РНК-полимеразу II и базальные транскрипционные факторы, уже взаимодействует с ДНК в районе корового промотора, но еще не приобрел конформацию, необходимую для начала транскрипции.

Открытый комплекс: происходит плавление участка ДНК в окрестностях старта транскрипции (11-15 пар оснований) и матричная цепь ДНК позиционируется в районе активного центра РНК-полимеразы II.

Стадия инициации транскрипции начинается с момента образования первой фосфодиэфирной связи РНК. Во многих случаях инициации и синтезу полноразмерного РНК-продукта предшествует стадия многократной инициации и наработки abortивных продуктов - коротких цепей РНК длиной 3-10 оснований.

Стадия элонгации транскрипции с момента, как РНК-полимераза синтезировала нить РНК длиной около 30 оснований и потеряла связь с коровым промотором и транскрипционной машиной. Присоединение факторов, обеспечивающих эффективный синтез РНК, ее процессинг, а также модификацию хроматина.

Скэффолд-комплекс - часть белковых компонентов прединициаторного комплекса остается на промоторном участке гена. Он является маркером транскрибируемых генов и существенно облегчает и ускоряет формирование прединициаторного комплекса в последующих транскрипционных циклах,.

Представлено по [Hahn S. et al., 2004] с модификациями.

С-КОНЦЕВОЙ ДОМЕН РНК-ПОЛИМЕРАЗЫ II И ЕГО РОЛЬ В ТРАНСКРИПЦИОННОМ ЦИКЛЕ

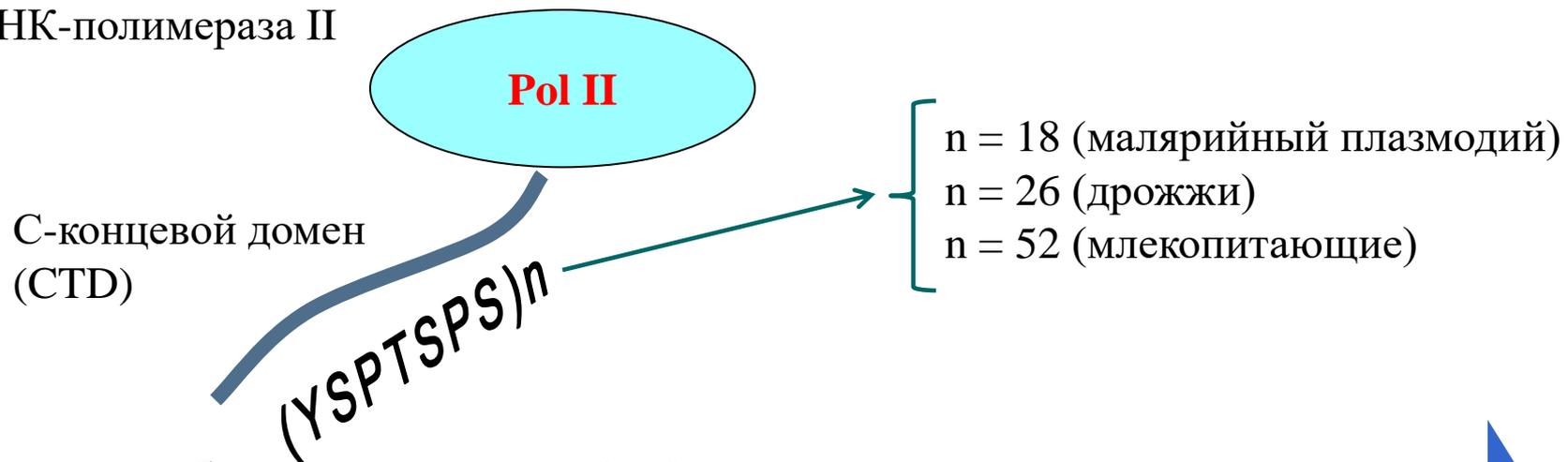
С-концевой домен = C-terminal domain (CTD) принадлежит самой большой субъединице РНК-полимеразы II (Rpb1).

Включает у каждого вида определенное число последовательно повторяющихся гептапептидов с канонической последовательностью:

Тирозин-Серин-Пролин-Треонин-Серин-Пролин-Серин
(Y1-S2-P3-T4-S5-P6-S7).

Длина CTD варьирует от 18 (у малярийного плазмодия) и 26 (у дрожжей) до 52 (у млекопитающих) повторяющихся гептапептидов и коррелирует с уровнем развития организма

РНК-полимераза II



Самые большие субъединицы ядерных РНК-полимераз I, II и III эукариот имеют общее происхождение; они представляют собой высокоомологичные белки. Однако, необычный домен найден только на С-конце субъединицы Rpb1 РНК-полимеразы II.

КОНСЕРВАТИВНОСТЬ САМЫХ БОЛЬШИХ СУБЪЕДИНИЦ РНК-ПОЛИМЕРАЗ

Rpb1 (*S. cerevisiae*)

E. coli



В субъединице Rpb1 (*S. cerevisiae*) выявлено 8 консервативных доменов, имеющих высокий уровень гомологии с определенными районами β' субъединицы РНК-полимеразы *E. coli* [Archambault J. & Friesen J.D. 1993]

Выравнивание **цинк-связывающего участка** консервативного района А самой большой субъединицы различных РНК-полимераз [Archambault J. & Friesen J.D. 1993]

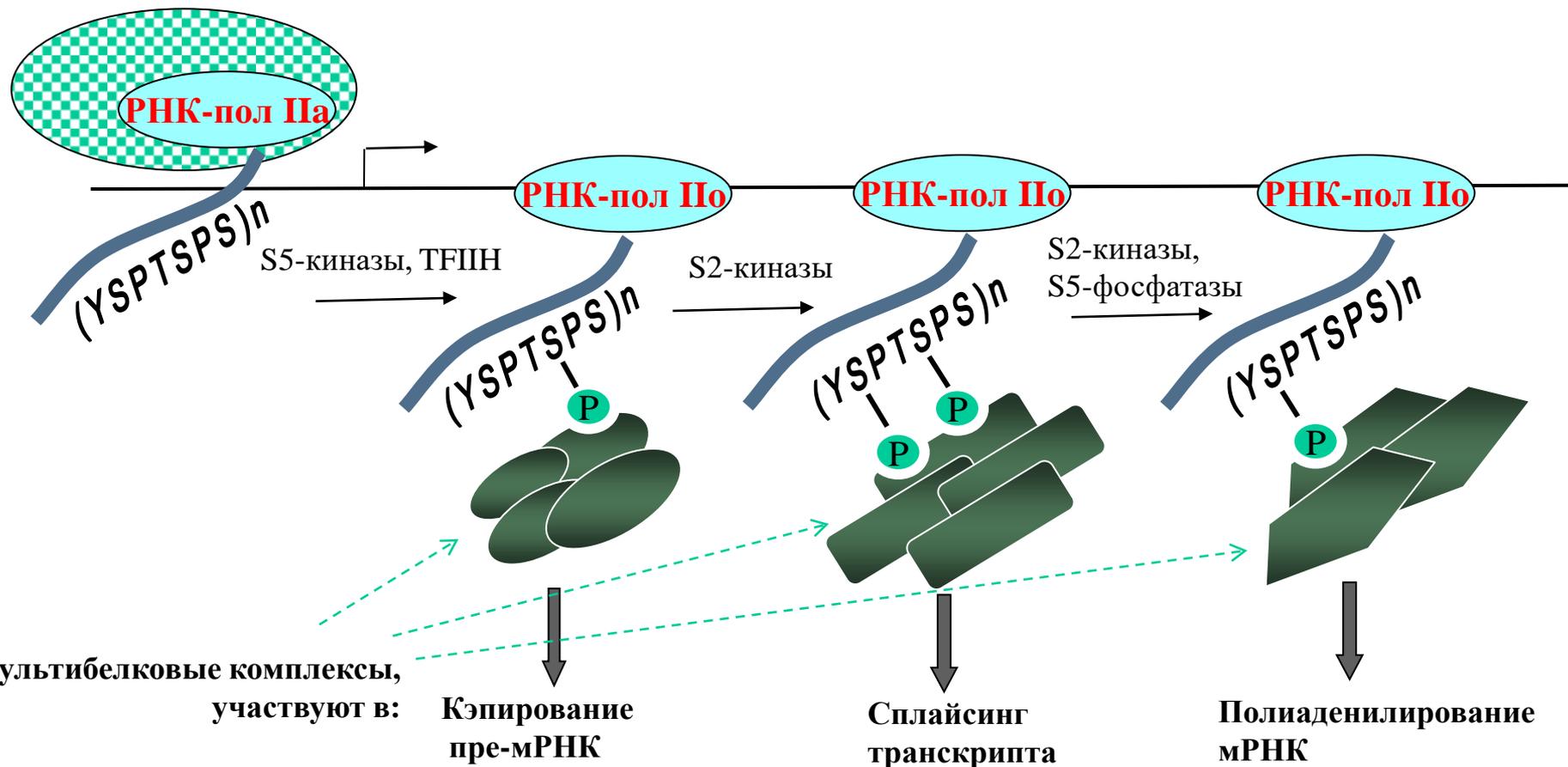
EUKARYOTIC	RNAPI	60	NL	C	ST	C	GLDEKF	C	P	GH	Q	G	HIELPVP-CYNPLFFNQLYIYLRAS	C	LF	C	HH	107	RPA1\$YEAST
		61	SV	C	AT	C	HLAERY	C	P	GH	F	G	HIVLPSP-AYHPLFFSQMYNLLRST	C	LY	C	HH	108	RPA1\$SCHPO
	RNAPII	65	LK	C	QT	C	QEGMNE	C	P	GH	F	G	HIDLAKP-VFHVGFIAKIKKVCCECV	C	MH	C	GK	112	RPB1\$YEAST
		69	GR	C	QT	C	AGNMTE	C	P	GH	F	G	HIELAKP-VFHVGFVKTMKVLRVCV	C	FF	C	SK	116	RPB1\$HUMAN
		69	GR	C	QT	C	AGNMTE	C	P	GH	F	G	HIELAKP-VFHVGFVKTMKVLRVCV	C	FF	C	SK	116	RPB1\$MOUSE
		65	SR	C	QT	C	AGNMTE	C	P	GH	F	G	HIDLAKP-VFHIGFITKTIKILRCV	C	FY	C	SK	112	RPB1\$DROME
		64	GR	C	MT	C	AGNLTG	C	P	GH	F	G	HLELAKP-VFHIGFLTTLKILRCV	C	FY	C	GR	111	RPB1\$CAEEL
		64	VK	C	ET	C	MANMAE	C	P	GH	F	G	YLELAKP-MYHVGFMKTVLSIMRCV	C	FN	C	SK	111	RPB1\$ARATH
	RNAPIII	66	TL	C	GT	C	NMNVKY	C	P	GH	F	G	HIELAKP-MYHVGFMNVLVNLRVCV	C	YH	C	GR	113	RPB1\$PLAFD
		67	SA	C	ET	C	HRKHPE	C	P	GH	F	G	YIELAEP-VFNIGVFDLVLVLRVCV	C	KT	C	GA	114	RPB1\$TRYBR
		67	SA	C	ET	C	HRKHPE	C	P	GH	F	G	YIELAEP-VFNIGVFDLVLVLRVCV	C	KT	C	GA	114	RPB2\$TRYBR
	ARCHAEBACTERIAL	65	LE	C	AT	C	HGNLAS	C	H	GH	F	G	HKLALP-VFHIGYFKATIQILQGI	C	KN	C	SA	112	RPC1\$YEAST
72		SI	C	ET	C	GLNSIE	C	V	GH	P	G	HIDLEAP-VFHLGFPTTVLRICRTI	C	KR	C	SH	119	RPC1\$TRYBR	
58		LR	C	RS	C	GAKGGE	C	P	GH	F	G	SINLARP-VIHVGFADTIHKILSSI	C	RK	C	SR	105	RPOA\$METTH	
VIRAL	60	LE	C	KT	C	GQRS GG	C	N	GH	F	G	HIELAAP-VIHVGFSKLIRRLRGT	C	RE	C	AS	107	RPOA\$HALHA	
	56	EK	C	PV	C	GNTLAG	C	P	GH	F	G	HIELIKP-VIHIGYVKHIYDFLRST	C	WR	C	GR	103	RPOA\$SULAC	
INVARIANT RESIDUES	47	AL	C	KT	C	GKTELE	C	F	GH	W	G	KVSIYKTHIVKPEFISEIIRLLNHI	C	IH	C	GL	95	RPO1\$VACCV	

Видовая принадлежность субъединиц: RPA1\$YEAST, *Saccharomyces cerevisiae*; RPA1\$SCHPO, *Schizosaccharomyces pombe*; RPB1\$YEAST, *S. cerevisiae*; RPB1\$HUMAN, *Homo sapiens*; RPB1\$MOUSE, *Mus musculus*; RPB1\$DROME, *D. melanogaster*; RPB1\$CAEEL, *C. elegans*; RPB1\$ARATH, *Arabidopsis thaliana*; RPB1\$PLAFD, *Plasmodium falciparum*; RPB1\$TRYBR and RPB2\$TRYBR, *T. brucei*; RPC1\$YEAST, *S. cerevisiae*; RPC1\$ TRYBR, *T. brucei*; RPOA\$METTH, *Methanobacterium thermoautotrophicum*; RPOA\$HALHA, *Halobacterium halobium*; RPOA\$SULAC, *Sulfolobus acidocaldarius*; RPO1\$VACCV, *vaccinia virus*.

Самые большие субъединицы ядерных РНК-полимераз I, II и III эукариот имеют общее происхождение; они представляют собой высокогомологичные белки. Однако, необычный домен найден только на С-конце субъединицы Rpb1 РНК-полимеразы II.

МОДИФИКАЦИИ С-КОНЦЕВОГО ДОМЕНА РНК-ПОЛИМЕРАЗЫ II НА РАЗЛИЧНЫХ ЭТАПАХ ТРАНСКРИПЦИИ.

Прединициаторный комплекс Инициация Ранняя элонгация → Поздняя элонгация Терминация транскрипции



В течение цикла транскрипции CTD подвергается различным структурным и ковалентным модификациям, которые могут, в свою очередь, модулировать процессы инициации, элонгации и терминации транскрипции.

ОСОБЕННОСТИ ТРАНСКРИПЦИИ У ЭУКАРИОТ:

1) У эукариот транскрипция осуществляется при участии РНК-полимераз трех типов, которые обозначаются как pol I, pol II, и pol III. Можно выделить три группы генов:

- pol I транскрибируемые: гены рРНК (18S, 5.8S и 28S)
- pol II транскрибируемые: все белок-кодирующие гены + гены многих мяРНК + многие миРНК
- pol III транскрибируемые : гены тРНК + гены низкомолекулярных РНК (5S рРНК, и др.)

2) Каждая РНК-полимераза имеет сложное, многосубъединичное строение, например, pol II включает более 10 субъединиц;

3) Для точной посадки РНК-полимераз на участок, прилежащий к старту транскрипции, нужны вспомогательные белки – базальные (или общие) транскрипционные факторы;

4) Для достижения необходимого уровня транскрипции каждого гена, помимо базальных (общих) транскрипционных факторов, требуются другие белки (транскрипционные факторы, коактиваторы, корепрессоры, медиаторы и др.);

5) Промоторы генов, транскрибируемых каждым типом полимераз, имеют характерное строение

6) Транскрипция генов эукариот зависит от состояния хроматина и плотности нуклеосомной упаковки.

ОСНОВНЫЕ КЛАССЫ БЕЛКОВ, УЧАСТВУЮЩИХ В РЕГУЛЯЦИИ ТРАНСКРИПЦИИ ГЕНОВ ЭУКАРИОТ

РНК-полимеразы

Синтез молекулы РНК,
комплиментарной и антипараллельной
матричной цепи ДНК.

РНК-полимеразы не могут функционировать сами по себе

Базальные (общие) транскрипционные факторы

Наличие ДНК-связывающего домена
НЕ ЯВЛЯЕТСЯ обязательной характе-
ристикой. ДНК-связывающий домен
имеется у белка ТВР (компонент TFIIID)

Формируют ПИК,
обеспечивают точную посадку
РНК-полимеразы на участки
ДНК, прилежащие к стартам
транскрипции

Стандартный набор для всех генов

Транскрипционные факторы

Имеют
ДНК-связывающий домен,
Специфически
взаимодействуют с ДНК

Регулируют интенсивность синтеза РНК каждого
конкретного гена в соответствии с потребностями организма
(типом ткани, стадией развития, воздействиями
окружающей среды)

Белки- медиаторы

Не имеют ДНК-связывающего домена,
Взаимодействуют с белок-белковыми
и ДНК-белковыми комплексами

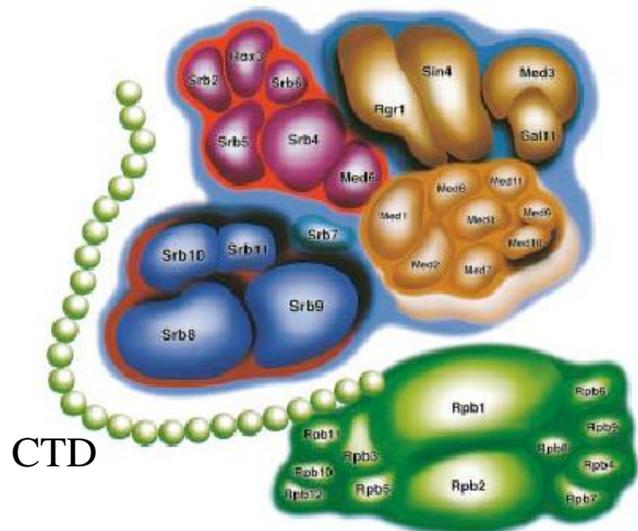
Корегуляторные белки (коактиваторы и корепрессоры)

Белки-медиаторы и **корегуляторные белки** (коактиваторы и корепрессоры) не имеют ДНК-связывающих доменов и участвуют в регуляции транскрипции без непосредственного специфического взаимодействия с ДНК

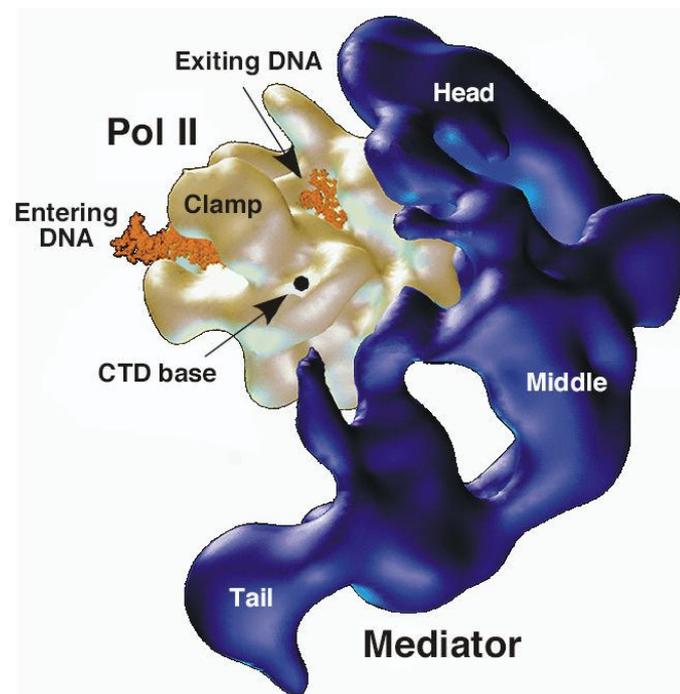
Медиаторный комплекс был впервые идентифицирован у дрожжей. Было обнаружено, что комплекс включает 24 субъединицы и необходим для поддержания как базального, так и индуцированного уровня транскрипции гена. Сходные по составу и функциям медиаторные комплексы были найдены у всех исследованных эукариотических организмов [Hahn S. et al., 2004].



СТРУКТУРА МЕДИАТОРНОГО КОМПЛЕКСА ДРОЖЖЕЙ



RNA Pol II



Медиаторный комплекс дрожжей (SRB/Med) включает 24 субъединицы:
Med1, Med2, Pgd1/Hrs1/Med3, Med4, Nut1, Med6, Med7, Med8, Cse2/Med9, Nut2/Med10, Med11, Srb8, Ssn2/Srb9, Rgr1, Gal11, Sin4, Srb4, Srb5, Rox3, Srb2, Srb7, Srb6, Srb10/Ssn3/Ume5, Srb11/Ssn8/Ume3
[Bourbon H.M. et al., 2004]

Присоединяется на средней стадии формирования ПИК. Имеет доменную структуру.

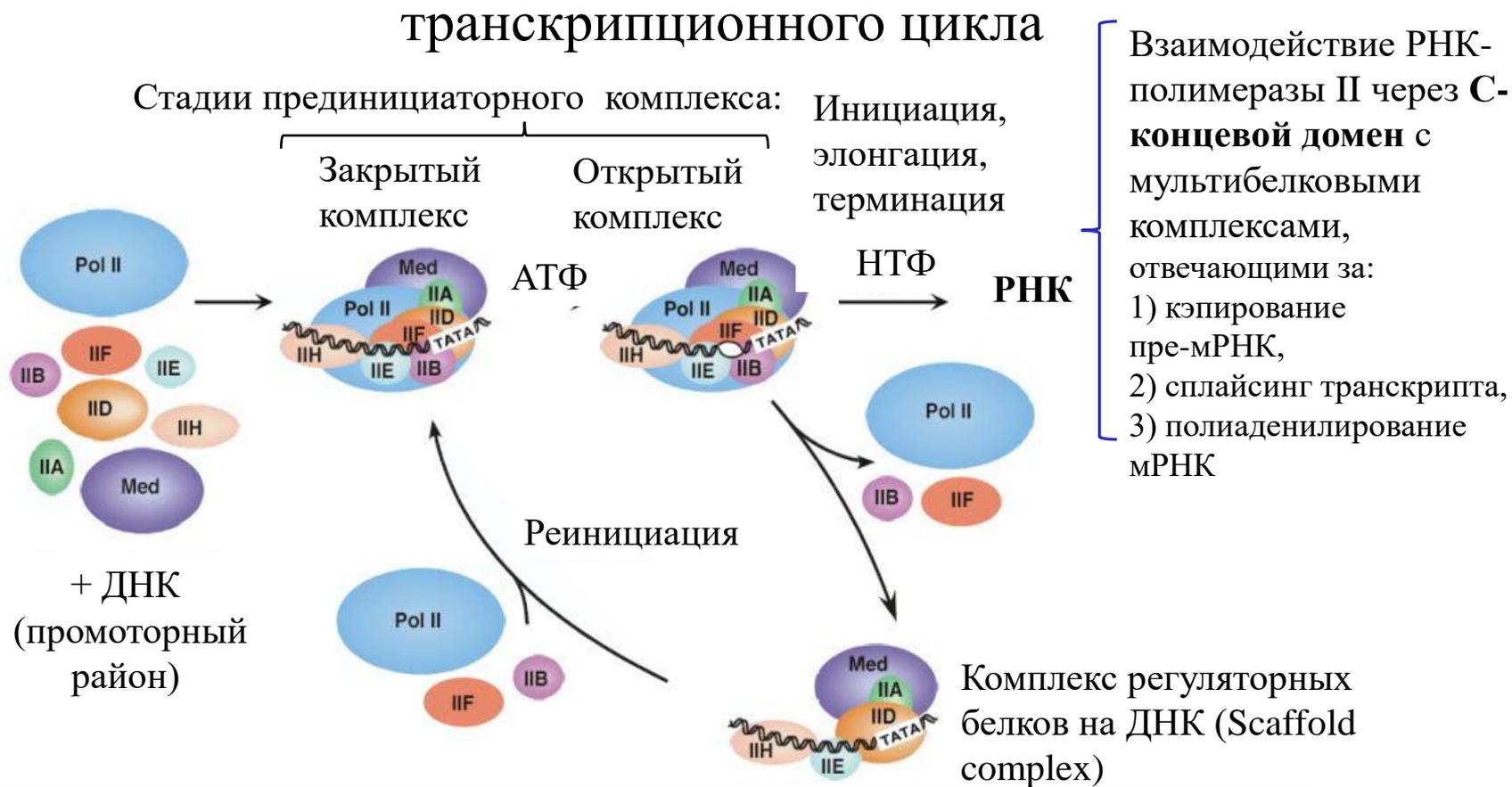
Функция: кооперативное связывание с РНК-полимеразой II и базальными транскрипционными факторами

Взаимодействие с полимеразой II осуществляется через С-концевой домен

Hampsey M, Reinberg D. RNA polymerase II as a control panel for multiple coactivator complexes. *Curr. Opin. Genet. Dev.* 1999, 9, 2, 132-139

Hahn S. Structure and mechanism of the RNA Polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* 2004 May; 11(5): 394-403.

С-концевой домен РНК-полимеразы II участвует во взаимодействии с мультибелковыми комплексами на всех стадиях транскрипционного цикла



Стадия прединициаторного комплекса: СТД <-> медиаторный комплекс

Ранняя элонгация: СТД <-> комплекс, осуществляющий кэпирование пре-мРНК

Элонгация: СТД <-> комплекс, осуществляющий сплайсинг транскрипта

Поздняя элонгация / терминация транскрипции: СТД <-> комплекс, осуществляющий полиаденилирование мРНК

РЕГУЛЯТОРНЫЕ ЭЛЕМЕНТЫ И РЕГУЛЯТОРНЫЕ ЕДИНИЦЫ (РАЙОНЫ), КОНТРОЛИРУЮЩИЕ ТРАНСКРИПЦИЮ ГЕНОВ ЭУКАРИОТ:

Регуляторные элементы:

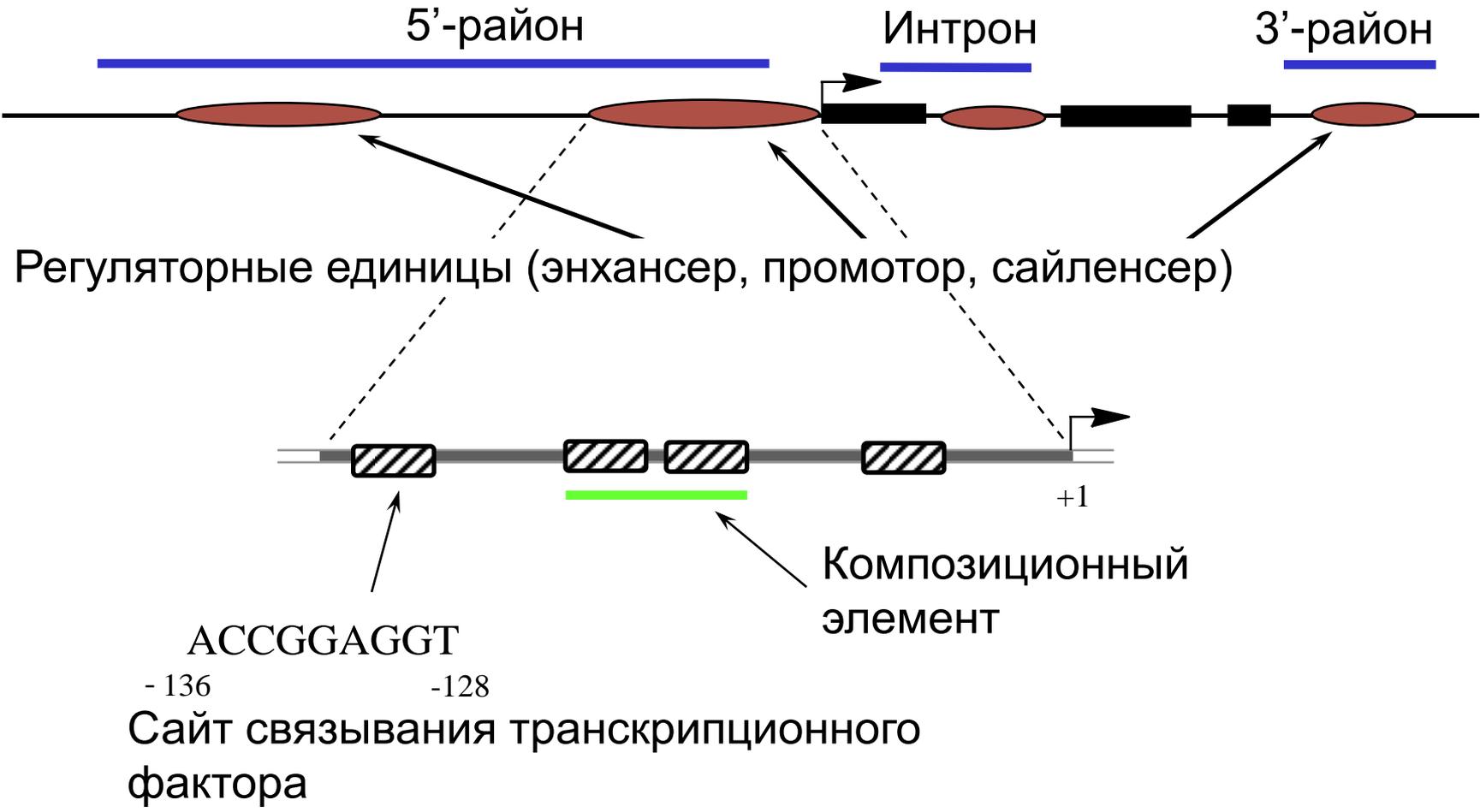
- сайты связывания транскрипционных факторов
- композиционные элементы,

Регуляторные единицы:

- промоторный район,
- энхансеры, сайленсеры



СХЕМА ОДНОГО ИЗ ВАРИАНТОВ СТРУКТУРНО-ФУНКЦИОНАЛЬНОЙ ОРГАНИЗАЦИИ РАЙОНОВ РЕГУЛЯЦИИ ТРАНСКРИПЦИИ И РЕГУЛЯТОРНЫХ ЭЛЕМЕНТОВ ГЕНОВ ЭУКАРИОТ.



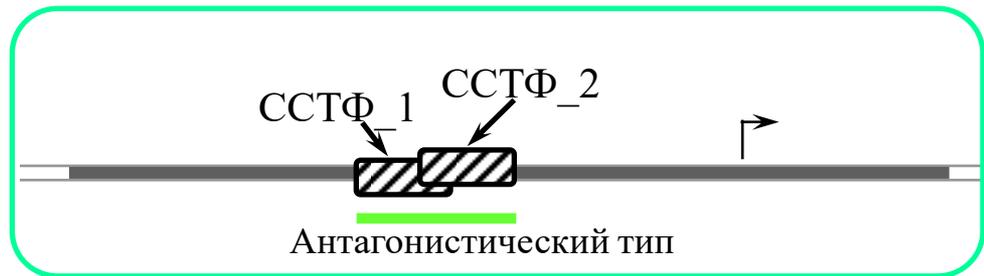
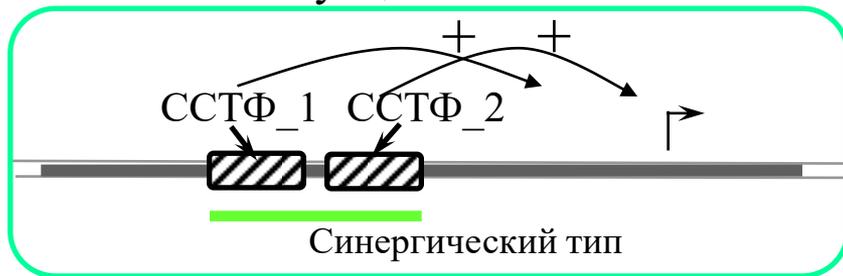
РЕГУЛЯТОРНЫЕ ЭЛЕМЕНТЫ (ОПРЕДЕЛЕНИЯ)

Сайты связывания транскрипционных факторов (ССТФ) - короткие (5-20 п.о.) последовательности ДНК, опознаваемые регуляторными белками – транскрипционными факторами. Сайты связывания транскрипционных факторов могут быть найдены в различных областях гена: 5'- и 3'-фланкирующих районах, интронах и экзонах.

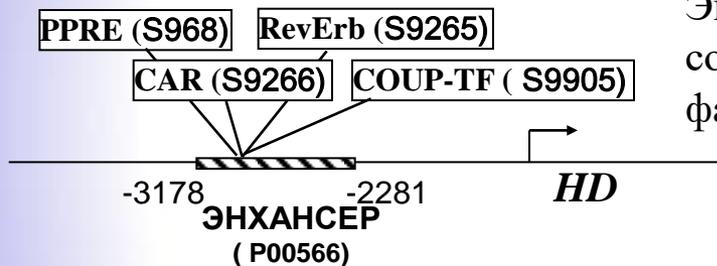
Композиционные элементы – пары сближенных или перекрывающихся сайтов связывания, которые в результате белок-белковых взаимодействий между соответствующими транскрипционными факторами приобретают новые регуляторные свойства.

Композиционные элементы **синергического типа** обеспечивают в результате белок-белковых взаимодействий неаддитивно высокий уровень активации транскрипции.

Композиционные элементы **антагонистического типа** включают перекрывающиеся либо очень близко расположенные ССТФ, взаимодействующие с факторами активаторами и репрессорами. В этой ситуации два белковых фактора не могут взаимодействовать с ДНК одновременно, поэтому при повышении концентрации одного из факторов возможна смена стимулирующего влияния фактора-активатора на ингибирующее влияние фактора-репрессора и, наоборот, в зависимости от клеточной ситуации.

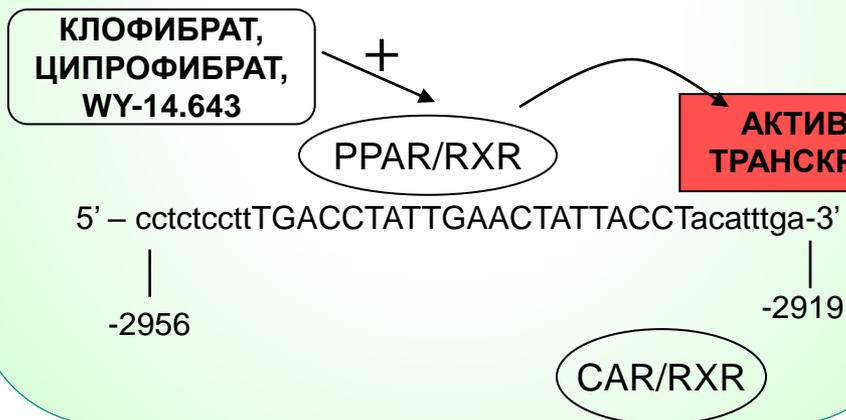


ПРИМЕР КОМПОЗИЦИОННОГО ЭЛЕМЕНТА АНТАГОНИСТИЧЕСКОГО ТИПА

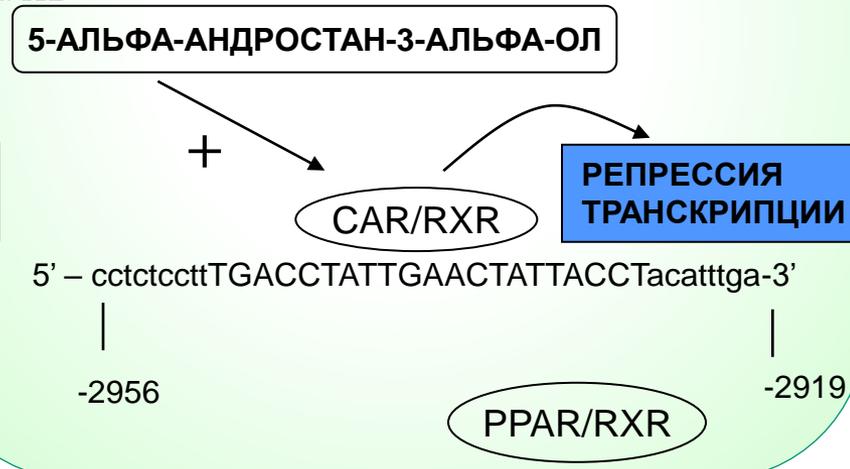


Энхансер гена гидратазы-дегидрогеназы (*HD*) крысы содержит перекрывающиеся сайты связывания факторов PPAR/RXR (=PPRE) и CAR/RXR (=CAR)

Под действием индукторов (пролифераторов пероксисом -клофибрата, ципрофибрата и WY-14.643) происходит активация гетеродимерного транскрипционного фактора PPAR/RXR, который взаимодействует с сайтом связывания в энхансере гена *HD* и усиливает его транскрипцию.



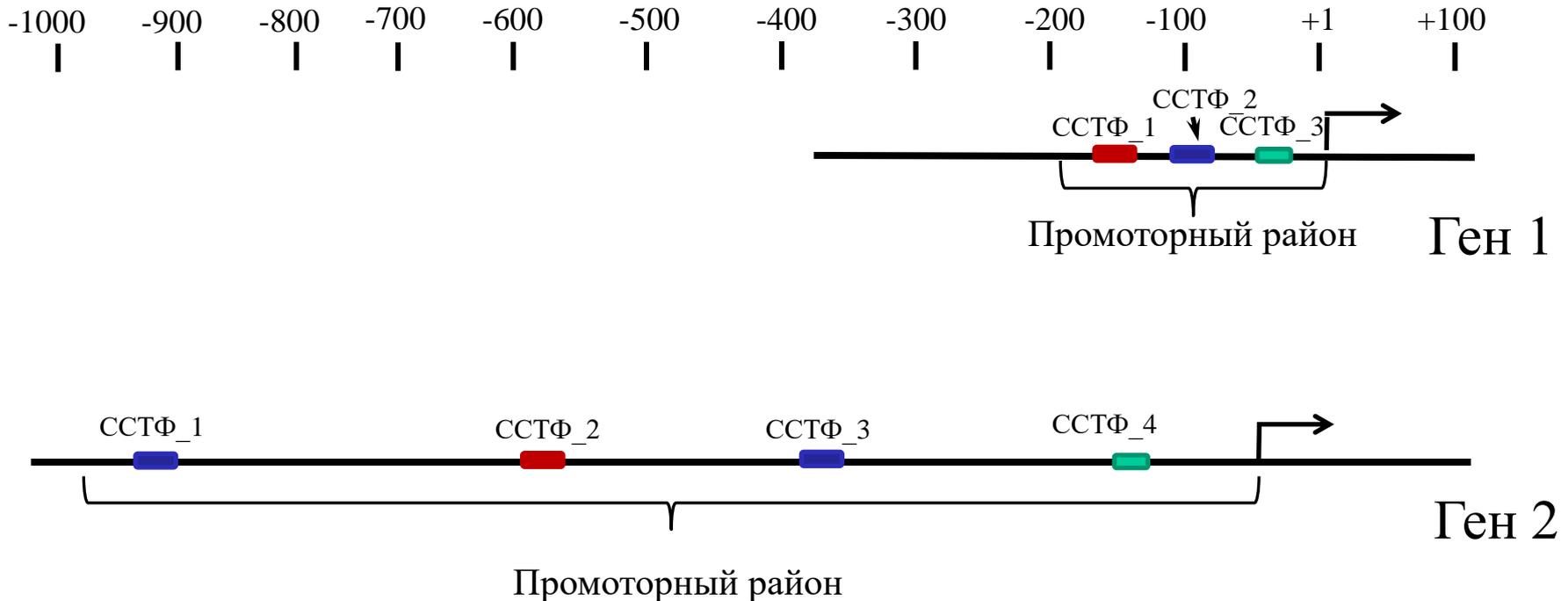
При воздействии другим веществом, 5-альфа-андростан-3-альфа-олом (5 alpha-androstan-3 alpha-ol), происходит активация транскрипционного фактора CAR (constitutive androstane receptor), который в составе гетеродимера CAR/RXR взаимодействует с соответствующим сайтом связывания в том же энхансере. Связывание белка CAR/RXR с ДНК оказывает **подавляющее** действие на транскрипцию гена *HD*



В зависимости от клеточной ситуации транскрипция гена *HD* может быть активирована, либо репрессирована

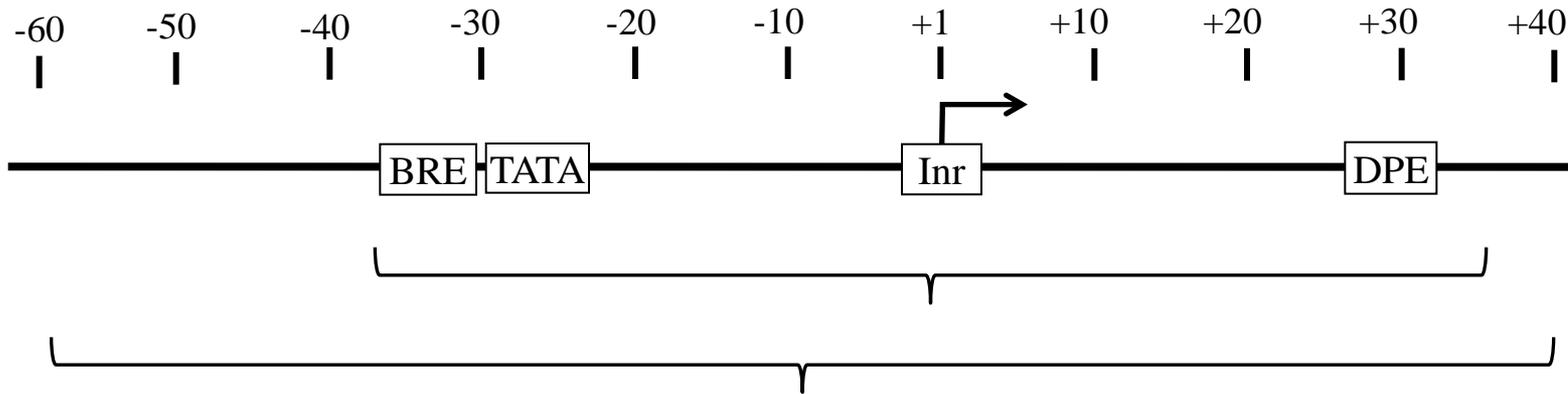
РЕГУЛЯТОРНЫЕ ЕДИНИЦЫ (ОПРЕДЕЛЕНИЯ)

Промоторный район – это 5'-фланкирующий участок гена длиной 200 - 1000 п.о., расположенный перед стартом транскрипции. Промоторный район включает коровый промотор, который обеспечивает формирование прединициаторного комплекса. Промоторный район содержит также сайты связывания транскрипционных факторов. Функция промоторного района - взаимодействие как с базальными транскрипционными факторами (компонентами ПИК), так и с транскрипционными факторами.



РЕГУЛЯТОРНЫЕ ЕДИНИЦЫ (ОПРЕДЕЛЕНИЯ)

Коровый (базальный) промотор - участок ДНК минимальной длины в окрестностях старта транскрипции гена, необходимый для формирования прединициаторного комплекса. До настоящего времени понятие корового промотора не содержит строгих данных о его размерах и точной локализации относительно старта транскрипции. В специальной литературе коровым промотором именуют либо очень короткий участок (-35 и +35 относительно старта транскрипции) [Smale S.T., and Kadonaga J.T. 2003], либо более протяженный район, например, от -60 до +40 [Zhang M.Q. et al., 1998].



Коровый (базальный) промотор

УДАЛЕННЫЕ ОТ СТАРТА ТРАНСКРИПЦИИ РЕГУЛЯТОРНЫЕ ЕДИНИЦЫ (ОПРЕДЕЛЕНИЯ)

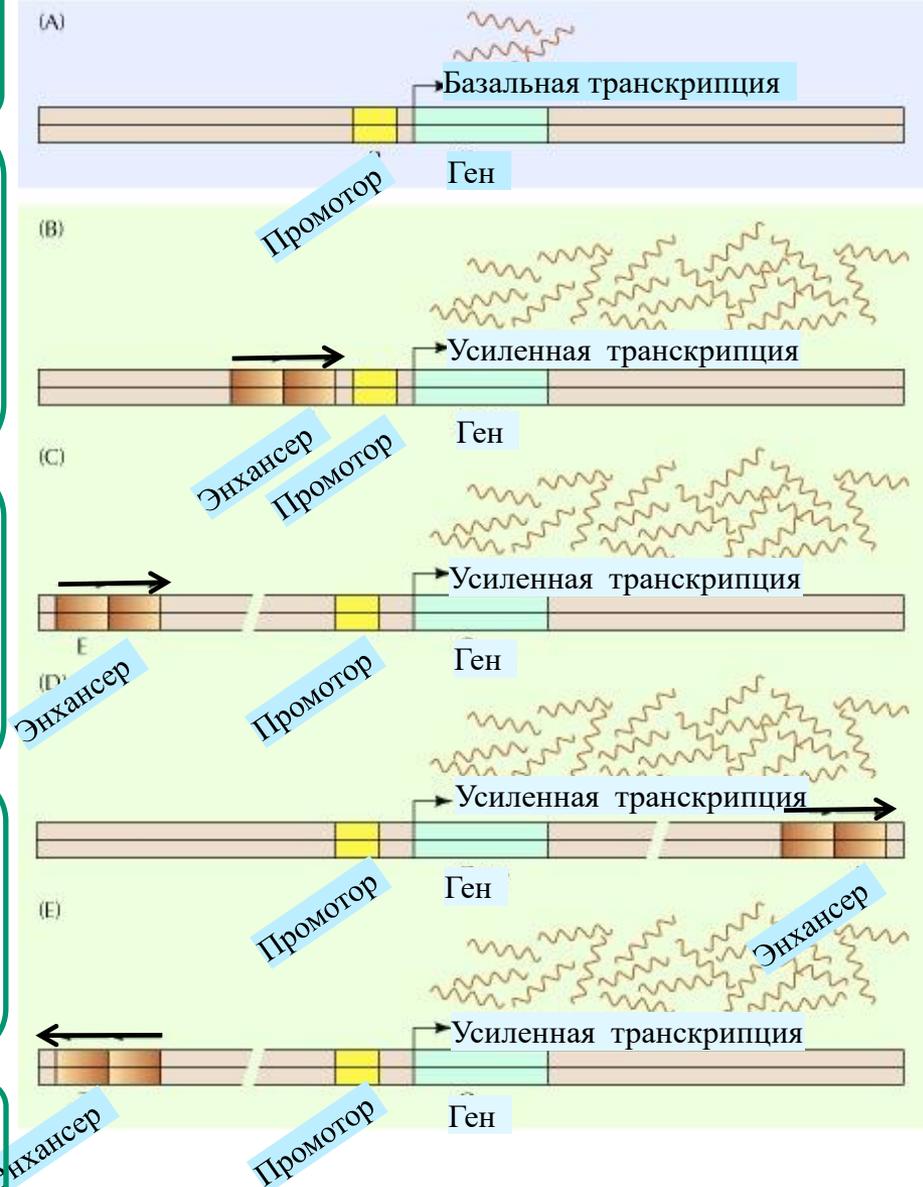
Энхансеры – регуляторные единицы, активирующие транскрипцию.

Энхансеры усиливают транскрипцию гена, управляемую определенным промотором, и оказывают свой эффект как в прямой, так и в обратной ориентации и в различной локализации (5'- либо 3'-) по отношению к промотору.

Энхансеры могут располагаться как в непосредственной близости от промоторного района (например, в области между -600 и -300) так и на значительном удалении (до нескольких тысяч п.о.).

Способность ряда энхансеров взаимодействовать со специфическими белками в дифференцированных клетках обеспечивает тканеспецифичный характер экспрессии соответствующих генов.

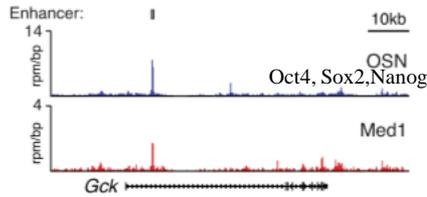
Участок энхансера может транскрибироваться с образованием энхансерной РНК (eRNA)



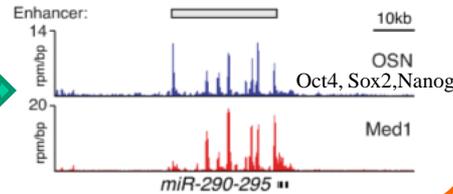
ЭНХАНСЕРЫ И СУПЕРЭНХАНСЕРЫ

Типичные энхансеры имеют протяженность в несколько сотен нуклеотидов

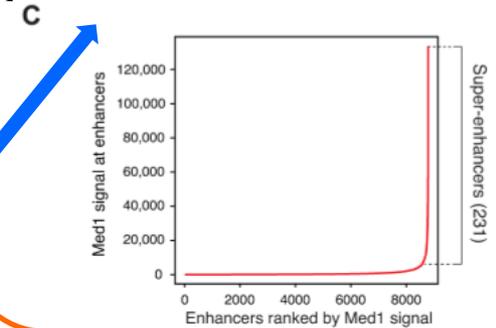
Суперэнхансеры занимают протяженные районы (до 50 кБ)



ChIP-Seq binding profiles (reads per million per base pair)

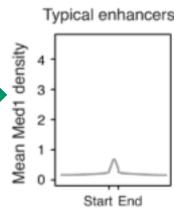


Примерно 40% всех медиаторных белков, связанных с энхансерами, были найдены в районах этих 231 суперэнхансеров

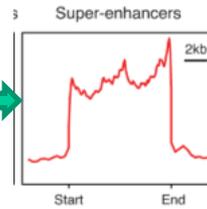


Усредненная картина распределения связывания с белком Med1

8 563 типичных энхансера (средний размер 703 п.о.)

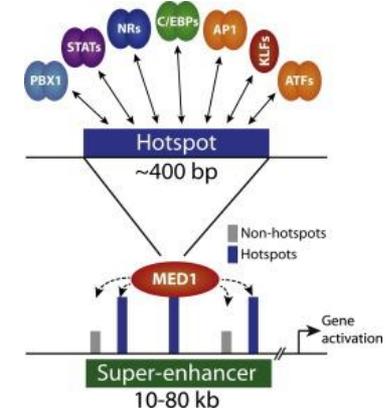
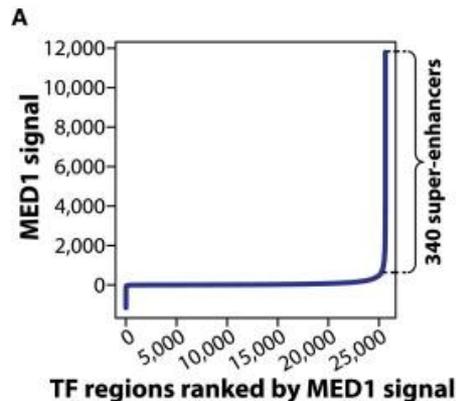


231 суперэнхансера (средний размер 8667 п.о.)



Warren A. Whyte et al., Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes Cell. 2013 Apr 11; 153(2): 307–319.

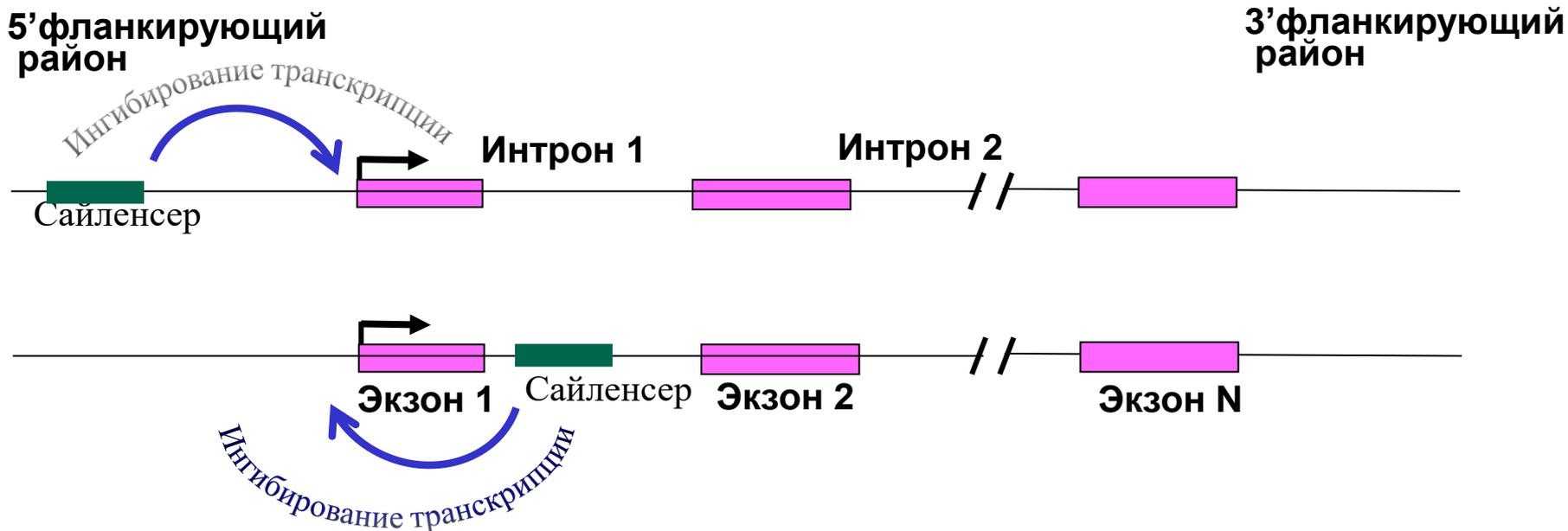
Отличительной характеристикой суперэнхансеров является размер и количество связанных с ними медиаторных комплексов, которые на порядок превышает размеры и количество, регистрируемые у типичных энхансеров



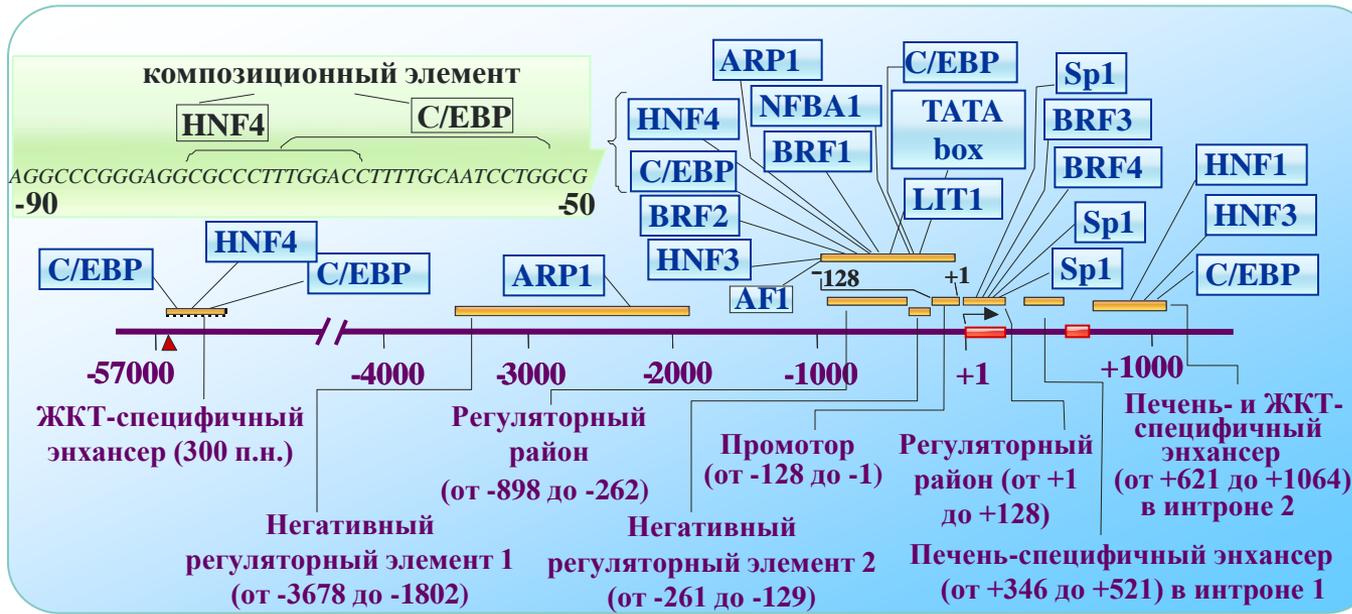
Siersbæk R. et al., Transcription Factor Cooperativity in Early Adipogenic Hotspots and Super-Enhancers. Cell Reports 2014, 7, 1443–1455

УДАЛЕННЫЕ ОТ СТАРТА ТРАНСКРИПЦИИ РЕГУЛЯТОРНЫЕ ЕДИНИЦЫ (ОПРЕДЕЛЕНИЯ)

Сайленсеры, или негативные регуляторные единицы - регуляторные последовательности ДНК, которые, также как и энхансеры, могут располагаться на значительном удалении от старта транскрипции (как в 5'-, так и в 3'-фланкирующих районах гена, а также интронах), выполняющие функцию ингибирования (вплоть до полного подавления) транскрипции гена, управляемой определенным промотором. Этот эффект реализуется за счет того, что сайленсеры (негативные регуляторные единицы) содержат сайты связывания транскрипционных факторов, взаимодействие с которыми оказывает негативное влияние на транскрипцию.

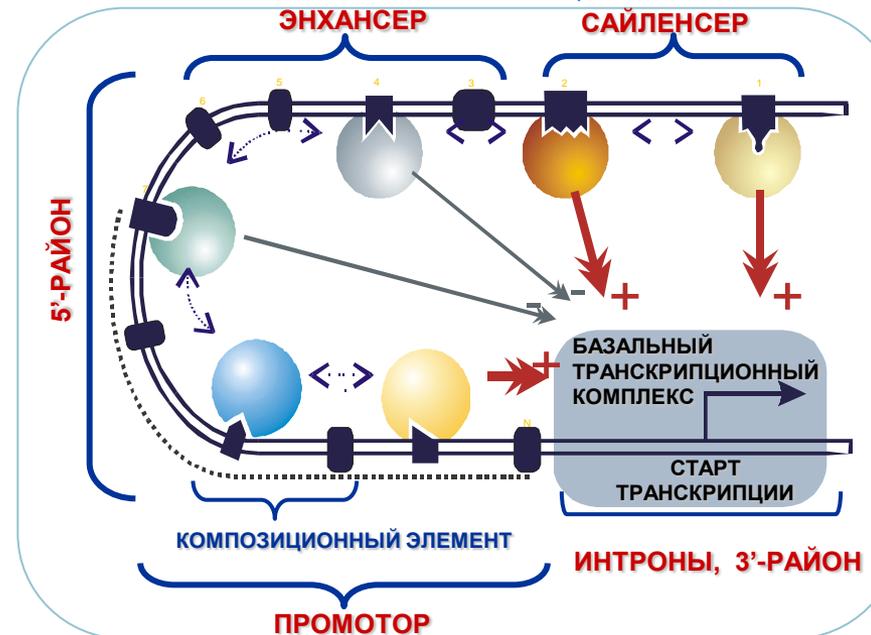


Организация регуляторных районов, контролирующих транскрипцию гена аполипопротеина В человека (База данных TRRD, ИЦиГ СО РАН)



Регуляторные районы генов эукариот содержат большое количество сайтов связывания транскрипционных факторов, что обеспечивает большое разнообразие паттернов экспрессии

Ген аполипопротеина В человека (*APOB*) содержит восемь регуляторных единиц, в числе которых, помимо промотора, имеется три тканеспецифичных энхансера. Два энхансера обнаружены в первом и втором интроне (+346/+521 и +621/+1064). Третий энхансер локализуется в 5'-фланкирующем районе на расстоянии 57 000 п.о. от старта транскрипции. Регуляторные единицы включают сайты связывания для 23 транскрипционных факторов, взаимодействие с которыми обеспечивает высокий тканеспецифический уровень экспрессии гена в печени и кишечнике



ОСНОВНЫЕ КЛАССЫ БЕЛКОВ, УЧАСТВУЮЩИХ В РЕГУЛЯЦИИ ТРАНСКРИПЦИИ ГЕНОВ ЭУКАРИОТ

РНК-полимеразы

Синтез молекулы РНК, комплиментарной и антипараллельной матричной цепи ДНК.

РНК-полимеразы не могут функционировать сами по себе

Базальные (общие) транскрипционные факторы

ДНК-связывающий домен **имеется** у белка ТВР (компонент TFIID)

Формируют ПИК, обеспечивают точную посадку РНК-полимеразы на участки ДНК, прилежащие к стартам транскрипции

Транскрипционные факторы

Имеют ДНК-связывающий домен, Специфически взаимодействуют с ДНК

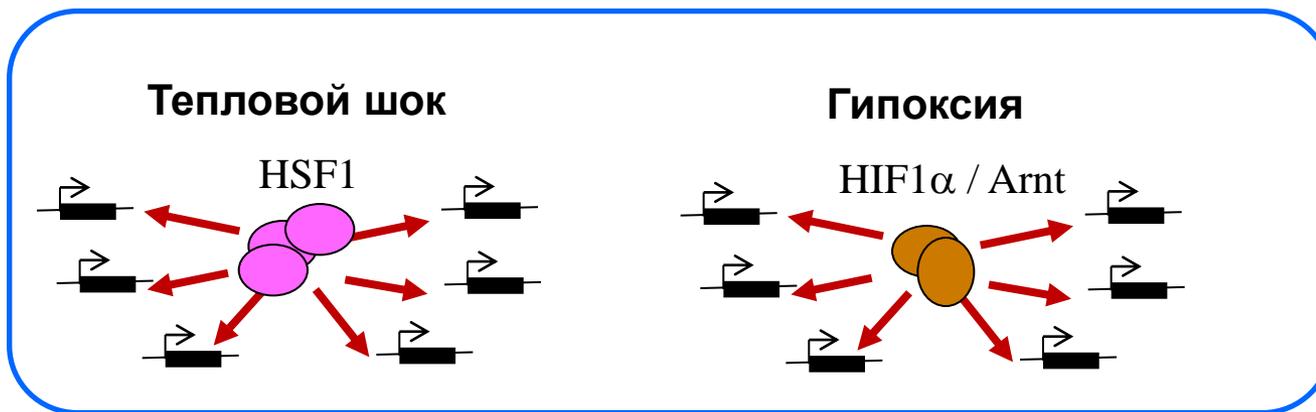
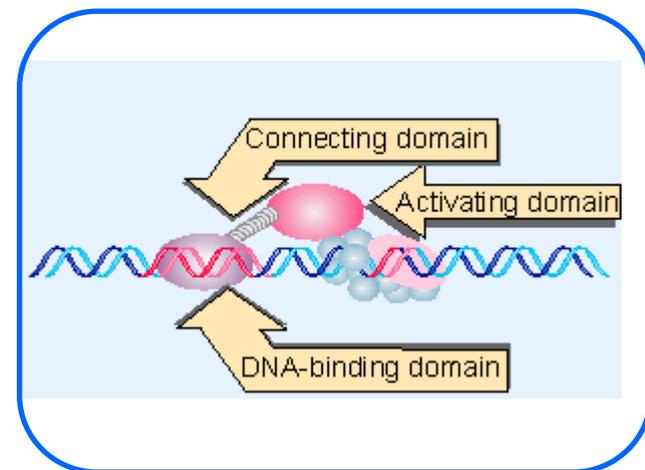
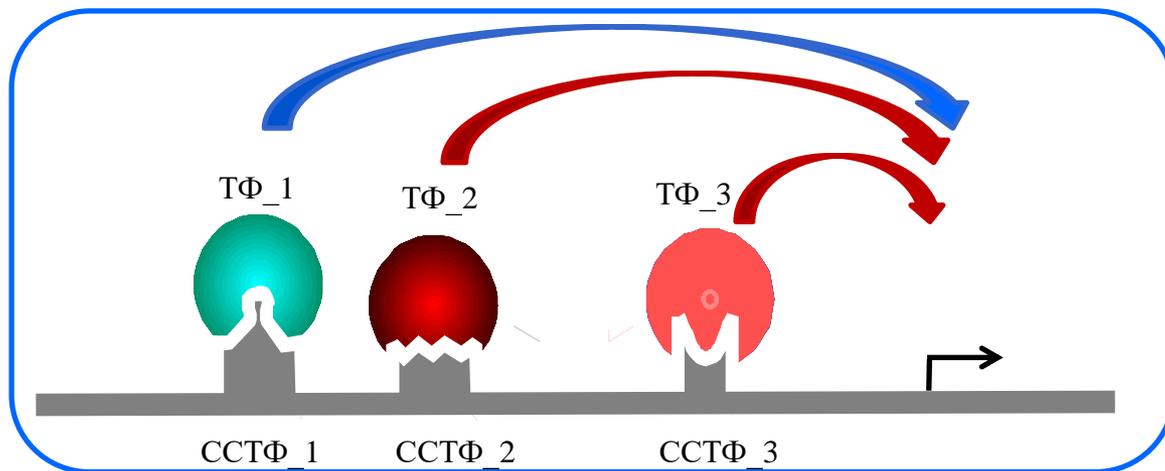
Регулируют интенсивность синтеза РНК каждого конкретного гена в соответствии с потребностями организма (типом ткани, стадией развития, воздействиями окружающей среды)

Белки-медиаторы

Не имеют ДНК-связывающего домена, Взаимодействуют с белок-белковыми и ДНК-белковыми комплексами

Корегуляторные белки (коактиваторы и корепрессоры)

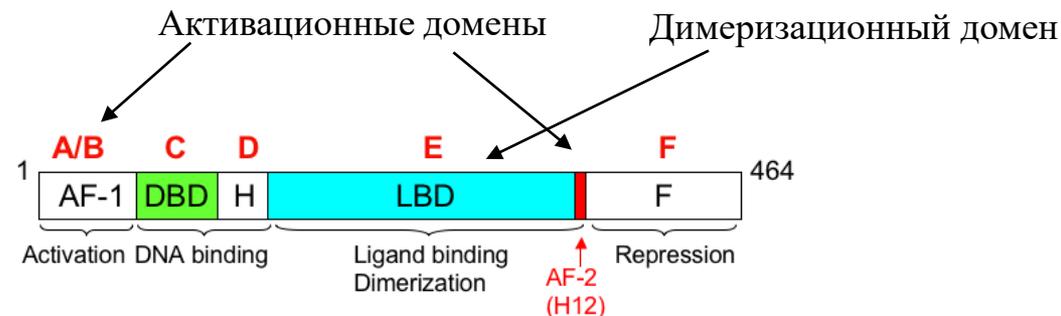
Транскрипционные факторы - белки, специфически связывающиеся с ДНК и регулирующие транскрипцию. Они взаимодействуют с короткими участками ДНК (сайтами связывания), которые расположены в различных районах генов (промоторных районах, энхансерах, сайленсерах). Транскрипционные факторы специфично регулируют экспрессию определенных групп генов



ТРАНСКРИПЦИОННЫЕ ФАКТОРЫ ВКЛЮЧАЮТ СЛЕДУЮЩИЕ ДОМЕНЫ:

- **ДНК-связывающий домен (=DBD) (обязательно имеется)**
- димеризационный (олигомеризационный) домен (многие факторы связываются с ДНК в виде димеров);
- транс-активирующий (или транс-репрессирующий) домен, через который осуществляется взаимодействие с различными белками, регуляторами транскрипции (компонентами ПИК, коактиваторами, либо корепрессорами, включая белки с хроматин-модифицирующей активностью);
- модулирующая область, которая часто является мишенью для модифицирующих ферментов, большей частью протеинкиназ;
- лиганд-связывающий домен.

Пример: доменная структура транскрипционного фактора HNF4 α человека



Современные оценки количества транскрипционных факторов в геноме человека

Компьютерная аннотация генома с целью идентификации генов, кодирующих белки, содержащие ДНК связывающие домены.

Ресурс TFClass

(<http://tfclass.bioinf.med.uni-goettingen.de/>)

Classification of Human Transcription Factors

TFClass is a classification of (so far) human transcription factors based on the characteristics of their DNA-binding domains. It comprises six levels (superclasses, classes, families, subfamilies, genera and factor species), two of which are optional (subfamilies and factor species). More detailed explanations about the classification scheme and its criteria will be given [here](#). The full classification can also be obtained [here](#) as html document and as [pdf](#) or [xls](#)-format.

When referring to this classification, please cite:
Wingender, E., Schönpf, T. and Dörfl, J.
TFClass: An expandable hierarchical classification of human transcription factors
Nucleic Acids Res. 41, D165-D170 (2013). [DOI](#)

Transcription factor classification

Superclass Class Family Subfamily
Genus Factor species

Human TF

- 1 Basic domains
- 2 Zinc-coordinating DNA-binding domains
- 3 Helix-turn-helix domains
- 4 Other all-alpha-helical DNA-binding domains
- 5 alpha-helices exposed by beta-structures
- 6 Immunoglobulin fold
- 7 beta-Hairpin exposed by an alpha/beta-scaffold
- 8 beta-Sheet binding to DNA
- 9 beta-Barrel DNA-binding domains
- 10 Yet undefined DNA-binding domains

Search:

Expand all Collapse all

Expand to:

Details

ID:

Definition:

Ресурс AnimalTFDB

(<http://www.bioguo.org/AnimalTFDB/>)

AnimalTFDB
Animal Transcription Factor Database

HOME BROWSE FAMILY BROWSE SPECIES SEARCH DOWNLOAD HELP ABOUT Quick search

Factors of Homo sapiens

This dataset collected 1544 transcription factors in 71 families, 124 chromatin remodeling factors and 302 transcription co-factors of Homo sapiens

Transcription factor family

AF-(4)	Androgen receptor(1)	AP-2(5)	ARID(15)	zinc finger(106)	C/EBP(10)
CBF(1)	CC-1(2)	COE(4)	COUP(3)	CP2(7)	CSD(8)
CR1(2)	CTF/NF(4)	CLTF(7)	DIM(7)	E2F(11)	E2f4/5/6 receptor(2)
ETS(29)	Fork head(49)	DCM(2)	GCR(1)	GTF2I(5)	HMG(50)
HMG/HMGY(2)	Homeobox(205)	HSF(8)	HTH(2)	IRF(9)	MRE(9)
MH1(8)	MYB(25)	NOTO/PhoC(1)	NF-YA(1)	NF-YB/C(2)	Nr1(1)
Nuclear orphan receptor(3)	Oestrogen receptor(1)	Orphan nuclear receptor(2)	Otrans(3)	PS(3)	PAX(9)
PC(1)	POU(21)	PPAR receptor(3)	Progesterone receptor(1)	Prmt(2)	Retroviral integrase(7)
RF(8)	RHD(10)	ROR receptor(4)	Rrm(3)	SAND(8)	SRE(6)
STAT(7)	T-box(17)	TEA(4)	TF_2ZIP(46)	TF_Oh(3)	THAP(12)
Thyroid hormone receptor(25)	TSC22(4)	Tub(5)	ZBTB(48)	z-BED(5)	z-C2HD(634)
z-C2HC(6)	z-GATA(14)	z-LITAF-like(2)	z-MIZ(7)	z-NF-X(2)	

1558 генов,
кодирующих 2904 изоформ белков,
содержащих ДНК-связывающий домен.

Из них 970 генов (62.3%) кодируют
экспериментально подтвержденные
транскрипционные факторы.

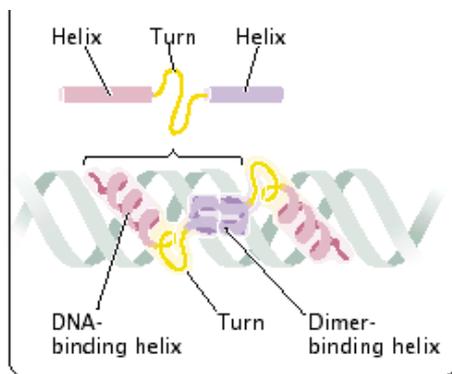
1544 гена,
кодирующих белки,
содержащие ДНК-связывающий домен.

Zhang H.M. et al., **AnimalTFDB**: a comprehensive animal transcription factor database. Nucleic Acids Res. 2012, 40(Database issue):D144-9.

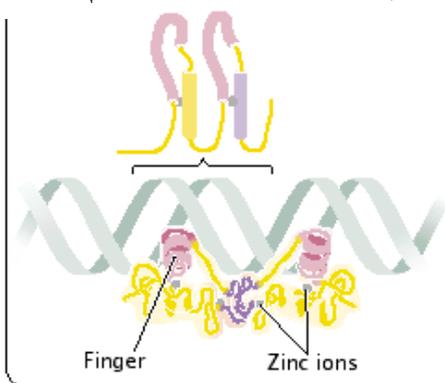
Wingender E, et al., **TFClass**: an expandable hierarchical classification of human transcription factors. Nucleic Acids Res. 2013 Jan;41(Database issue):D165-70.

НЕКОТОРЫЕ ТИПЫ ТРАНСКРИПЦИОННЫХ ФАКТОРОВ С УЧЕТОМ СТРОЕНИЯ ДНК-СВЯЗЫВАЮЩИХ ДОМЕНОВ

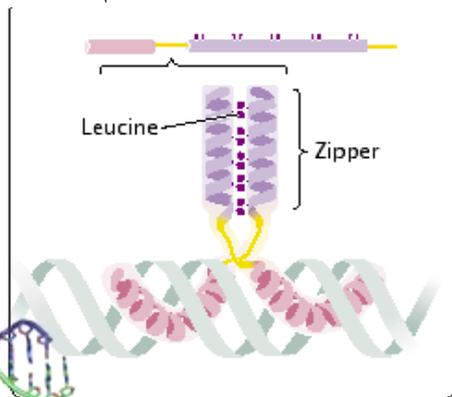
Спираль-поворот-спираль



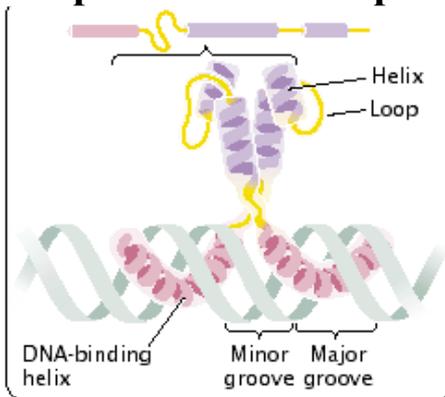
Цинковый палец



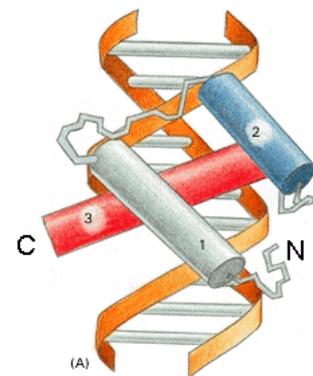
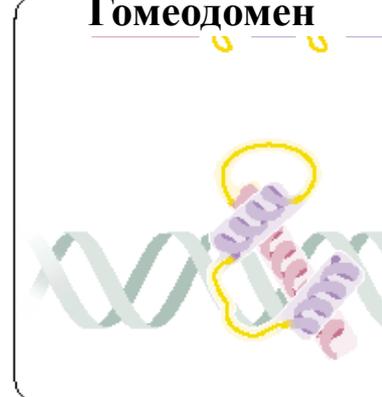
Лейциновая застежка



Спираль-петля-спираль



Гомеодомен



Специфичность взаимодействия транскрипционных факторов с распознаваемыми ими регуляторными последовательностями генов определяется преимущественно особенностями их ДНК-связывающих доменов. Классификация транскрипционных факторов, учитывающая совокупность свойств их ДНК-связывающих доменов, предложена Э. Вингендером [Вингендер Э., 1997]

САЙТЫ СВЯЗЫВАНИЯ ТРАНСКРИПЦИОННЫХ ФАКТОРОВ (ПРИМЕРЫ)

Транскрипционный фактор SF1 (Steroidogenic factor 1)

ggagtTCAAGGTAAtaagggc
 ccagctCAAGGCTAagtgaga
 gaggggggAGGTCAacactcc
 ctagctcaAGGCTAagagagg
 gtctcCCAAGGTCatccttgt
 caaagTAGAGGTCAggagga
 gctagtCAAGGTTActtcaa
 tttatTCAAGGTAatgataac
 ccggcCCAAGGTCcacttgct
 agcttTCGAGGTCatggccac
 cttggACAAGGGCGcagaggg
 gacagACAAGGTCAgaaagga
 gggtgCAAGGCCActaagca
 ggcggaCAAGGTCAgggaggt
 gggaggCAAGGCCActgggca

Последовательности сайтов связывания в
 различных генах позвоночных

Транскрипционный фактор SREBP1 (Sterol regulatory element-binding transcription factor 1)

GgcgcgtcggctccctCCACCCCTGctgagatgatgcact
 AccaggcgacgagaagACACCCACcccccaactcaccsa
 AcacccccacccccaaacTCACCCAGtgagcatcattaa
 aacggTcgccttaacaaccgcCCACTgctcgcacccgggc
 AgcttctagagtgttaTCACGCCAGtctccttccgcgact
 GtgccagttgggatctCCAAGCCACgcccaccaagagttt
 CagttccttctgggtGACACTCCACagttctcccaccgaac
 CcatccatagtccttgTCACCTGACagggggtgggtaaac
 GgcaatccccgcgcggCCACGCCACatgggctgacagctt
 TgattgcgtgaccctGTCACTCCAGagttagatagccaag
 CtctggagtgacaggGTCACGCAATcagaccacggtaatt
 TgacaggggtcacgcaATCAGACCACggtaatgacacaacg
 CgtaacccaacccttGTCACTCCAAgaccctaacttggtc
 CaagaccctaacttgGTCACTCCACagccttgcctcctct
 CacatccagccgggcTACACCCCATcactccacgggcccg

ТСААГГТСА <-Консенсусные последовательности -> ТСАСССАС

SF1

Уровень
 консерватизма



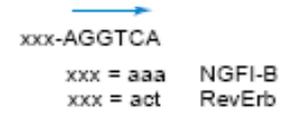
SREBP

Уровень
 консерватизма



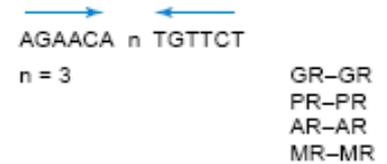
ДИМЕРНЫЕ И МОНОМЕРНЫЕ ТРАНСКРИПЦИОННЫЕ ФАКТОРЫ: ВАРИАНТЫ ОРГАНИЗАЦИИ САЙТОВ СВЯЗЫВАНИЯ НА ДНК (на примере транскрипционных факторов из класса «Cys4 zinc finger of nuclear receptor type»)

Сайт связывания
мономерного фактора

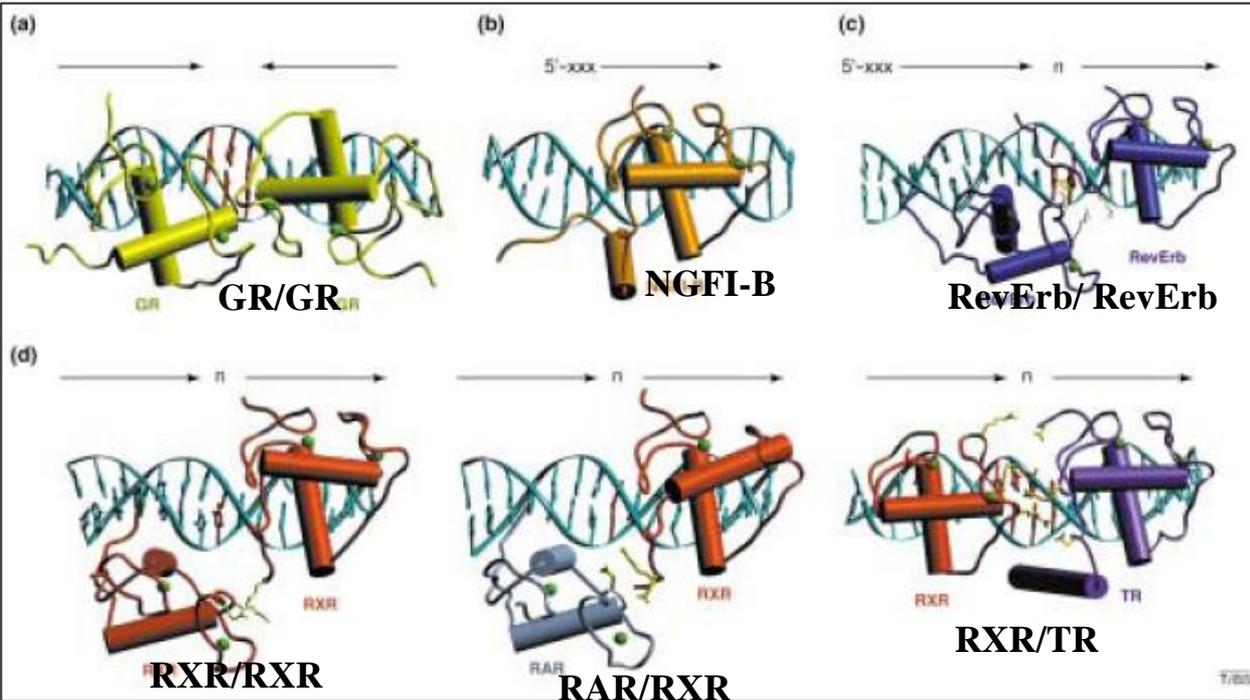
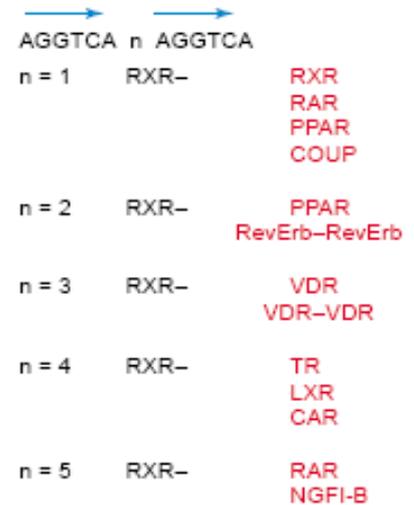


Сайты связывания димеров:
А) симметричный сайт (палиндром)

(a) Symmetric sites



Б) Прямой повтор



ФУНКЦИОНАЛЬНАЯ КЛАССИФИКАЦИЯ ТРАНСКРИПЦИОННЫХ ФАКТОРОВ:

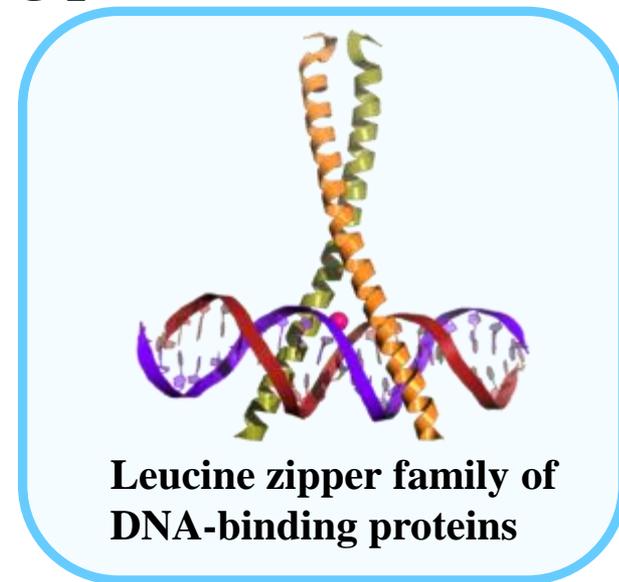
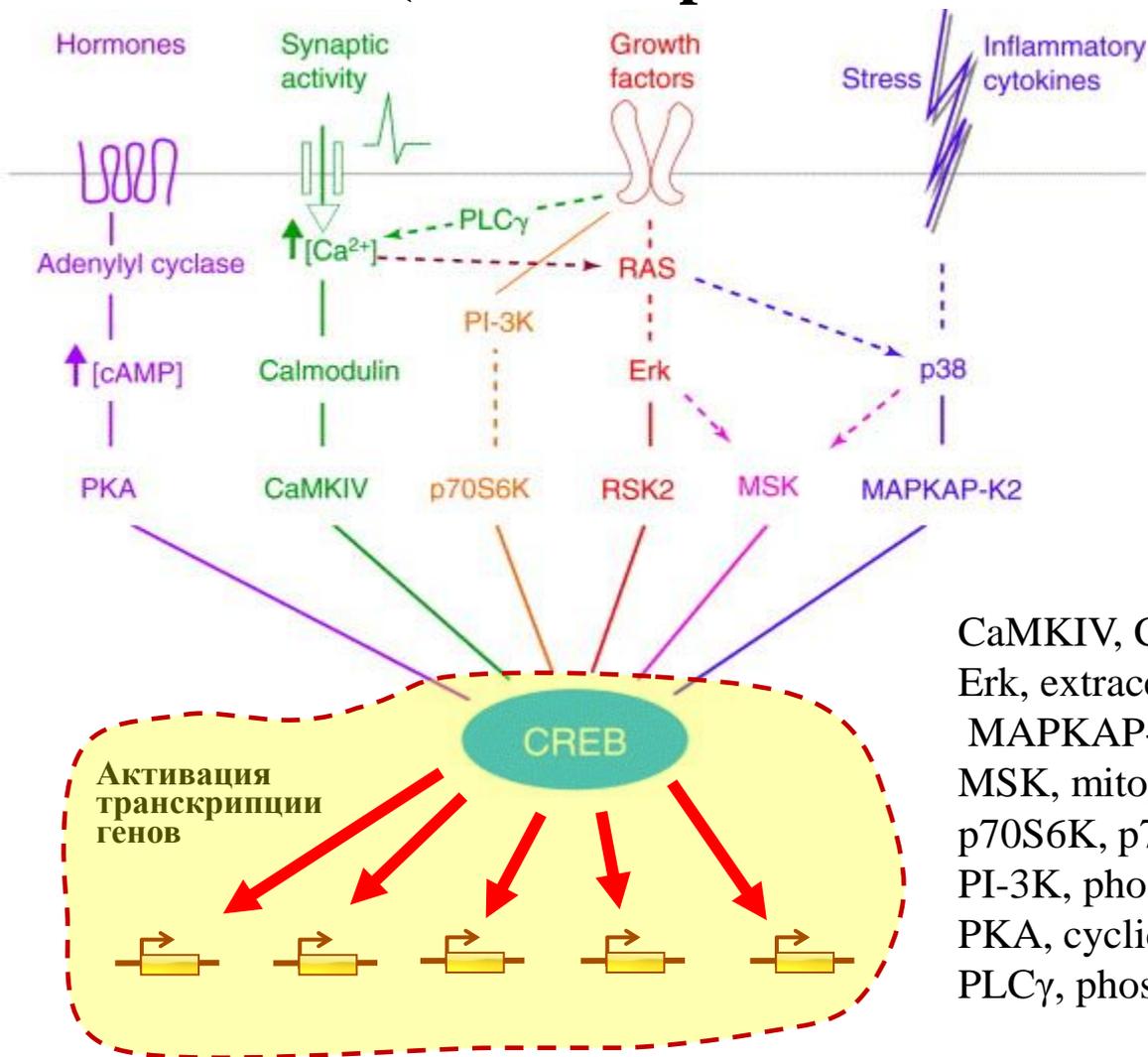


Количество транскрипционных факторов, известных на сегодняшний день, достаточно велико. Например, по современным оценкам, геном человека включает около 1 500 генов, кодирующих транскрипционные факторы. Предложена классификация транскрипционных факторов, основанная на особенностях их экспрессии и функционирования.

ПРИМЕРЫ ТРАНСКРИПЦИОННЫХ ФАКТОРОВ И КОНСЕНСУСНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ИХ САЙТОВ СВЯЗЫВАНИЯ НА ДНК

Название транскрипционного фактора (семейства факторов)	Мультимерное состояние фактора	Тип ДНК-связывающего домена	Функциональный класс	Консенсусная последовательность сайта связывания на ДНК
Sp1 (Specificity protein 1)	Мономер	Домен, содержащий координированные атомы цинка / цинковые пальцы Cys2His2 типа	Конститутивный, экспрессируется в широком круге тканей	GGGCGG
GATA1	Мономер	Домен, содержащий координированные атомы цинка / цинковые пальцы Cys4 типа	Тканеспецифичный: эритроидные клетки	GATA
SREBP	Гомодимер	Основной домен / спираль-петля-спираль, «лейциновая застежка» (bHLH-ZIP)	Активность регулируется в зависимости от уровня холестерина в клетке	TCACCCCAT
PPAR	Гетеродимер (PPAR/RXR)	Домен, содержащий координированные атомы цинка / цинковые пальцы Cys4 типа	Активность регулируется в зависимости от внутриклеточного содержания жирных кислот и их производных (эйкозаноиды, простагландины и лейкотриены)	AGGTCANAGGTCA

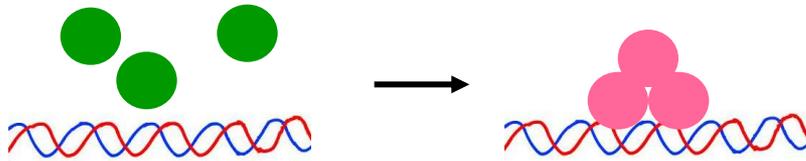
ПУТИ СИГНАЛЬНОЙ ТРАНДУКЦИИ, ПРИВОДЯЩИЕ К ФОСФОРИЛИРОВАНИЮ ФАКТОРОВ СЕМЕЙСТВА CREB (cAMP response element-binding protein)



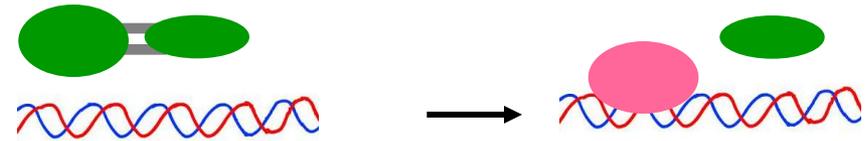
CaMKIV, Ca²⁺-calmodulin-dependent kinase IV;
 Erk, extracellular regulated kinase;
 MAPKAP-K2, MAP-kinase-activated protein kinase 2;
 MSK, mitogen- and stress-activated kinase;
 p70S6K, p70 S6 kinase;
 PI-3K, phosphoinositide 3-kinase;
 PKA, cyclic-AMP-dependent protein kinase;
 PLCγ, phospholipase C; RSK2, ribosomal S6 kinase 2.

ПУТИ РЕГУЛЯЦИИ АКТИВНОСТИ ТРАНСКРИПЦИОННЫХ ФАКТОРОВ

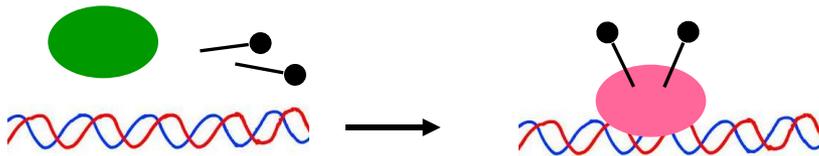
Мультимеризация (HSF1)



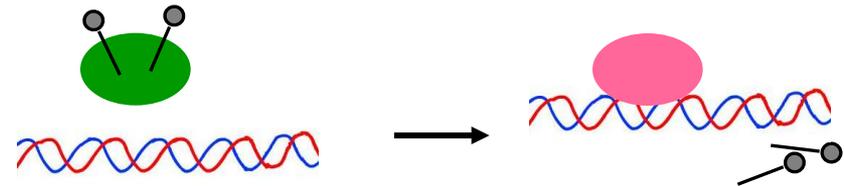
Протеолитическое расщепление неактивного предшественника (SREBP1)



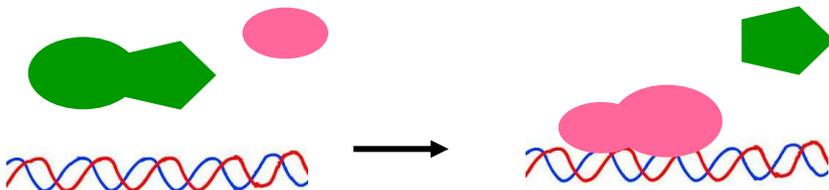
Фосфорилирование (CREB1)



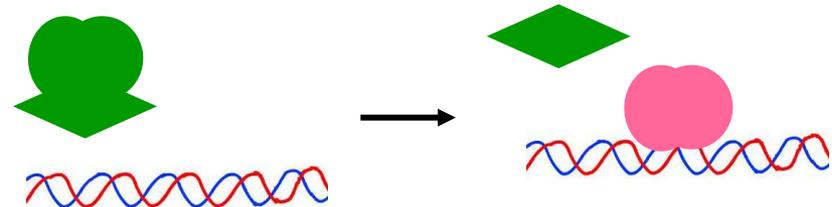
Дефосфорилирование (HNF4alpha)



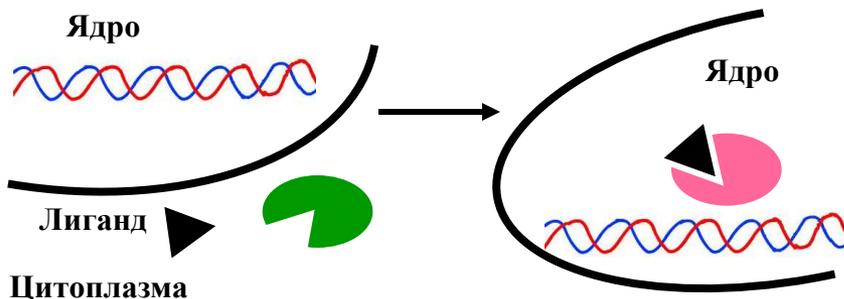
Вытеснение ингибирующего партнера активным (MyoD1)



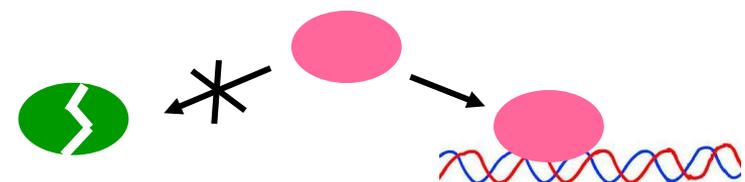
Распад комплекса с ингибиторным белком (E2F-1/DP-1)



Связывание с лигандом (ER, GR)



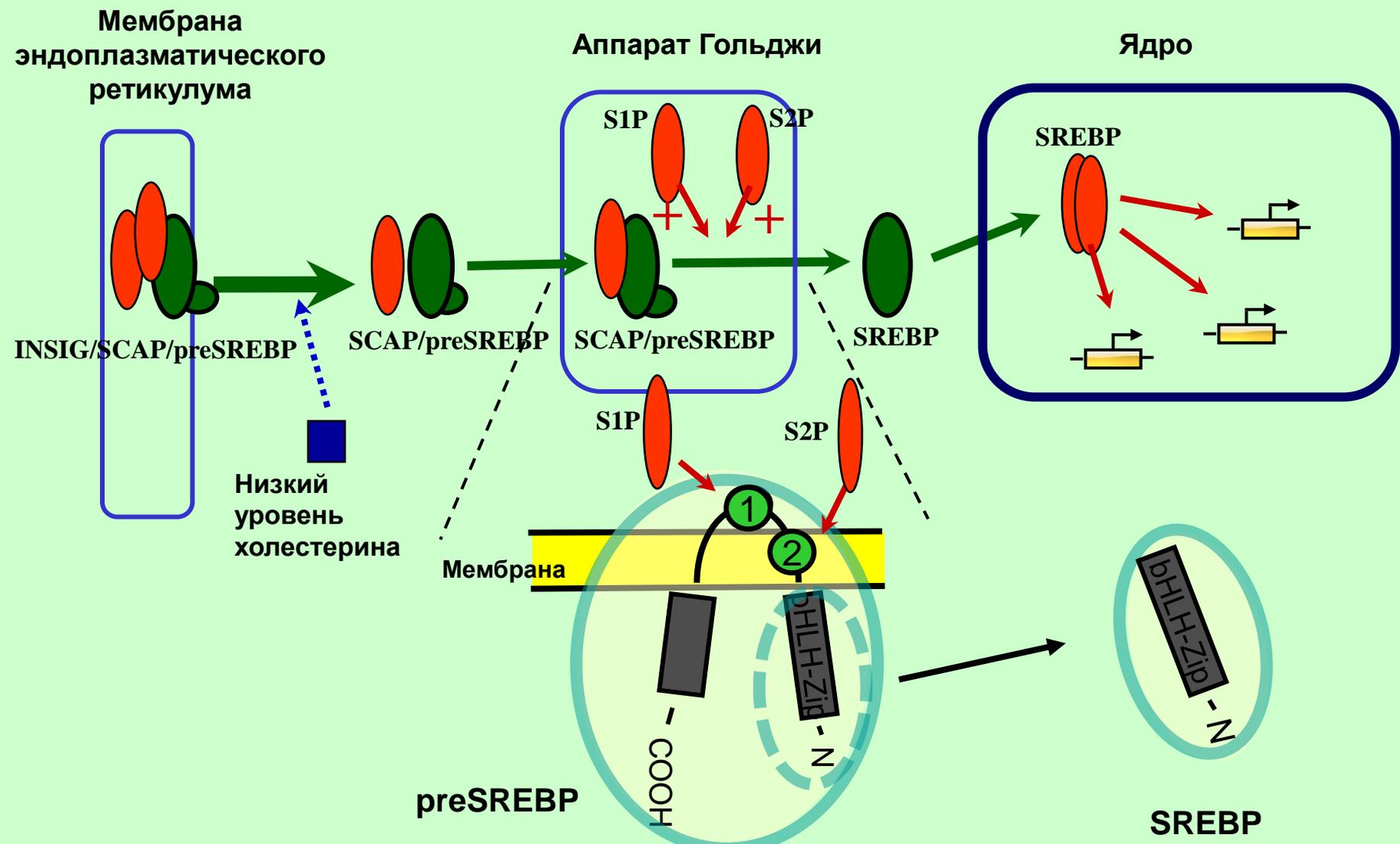
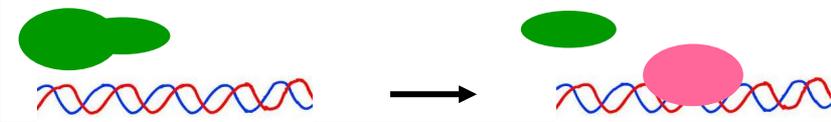
Ингибирование деградации (HIF1alpha)



Зеленым цветом обозначено неактивное состояние белка, красным - активное

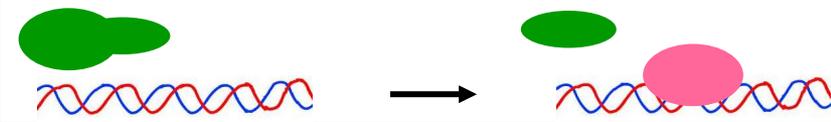
ПРИМЕР: холестерин регулирует активность транскрипционных факторов семейства SREBP (Sterol regulatory element-binding proteins)

Расщепление неактивного предшественника

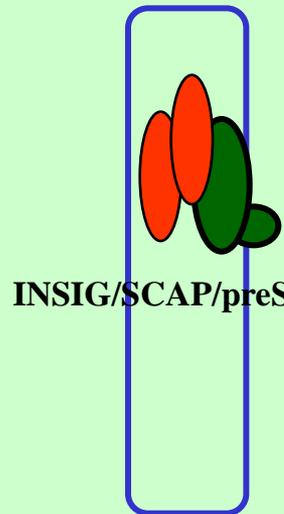


ПРИМЕР: холестерин регулирует активность транскрипционных факторов семейства SREBP

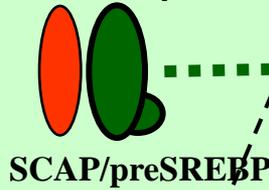
Расщепление неактивного предшественника



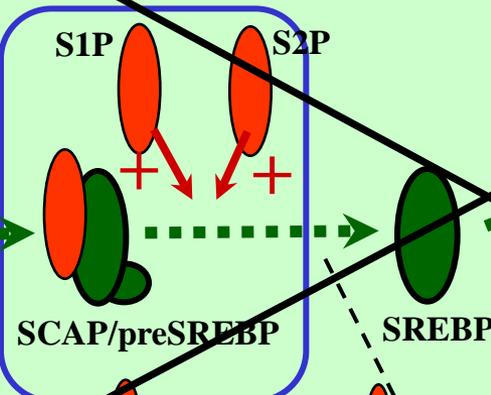
Мембрана
эндоплазматического
ретикулула



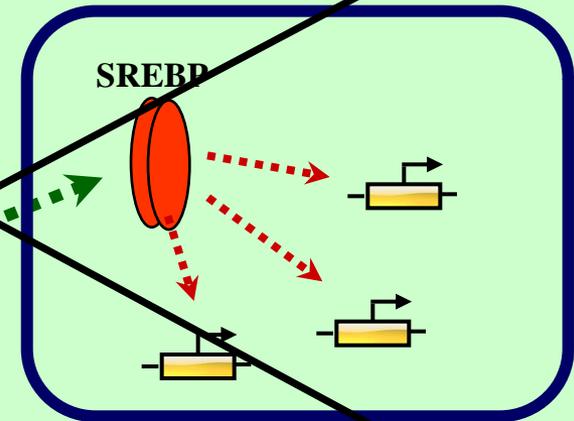
Высокий
уровень
холестерина



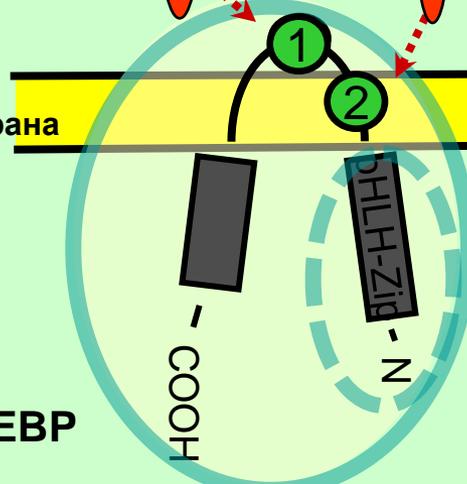
Аппарат Гольджи



Ядро



Мембрана



preSREBP

COOH



SREBP

N

Конец 2-ой лекции