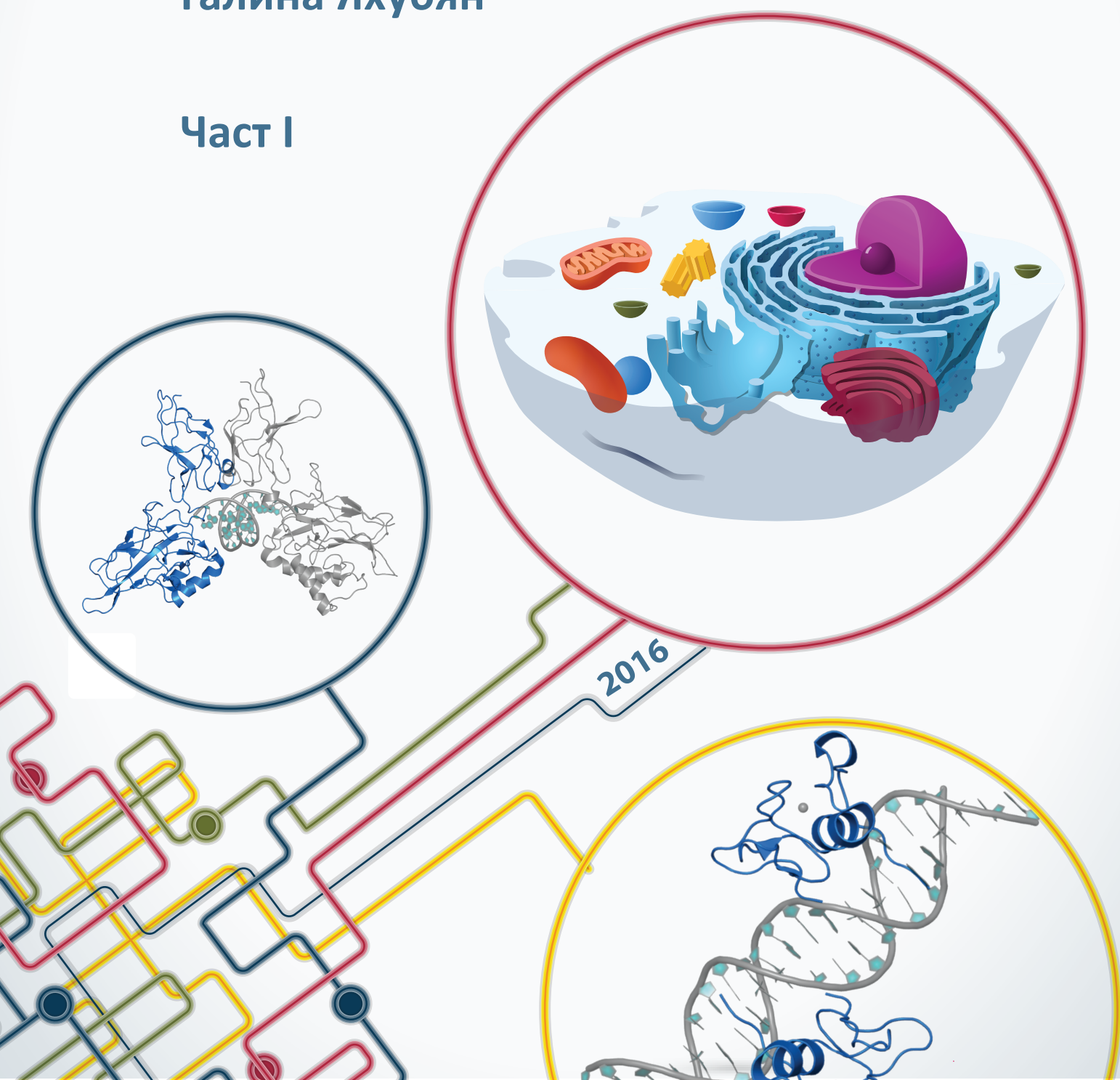


Пловдивски Университет "Паисий Хилендарски"

Регулация на генната експресия при еукариоти

Галина Яхубян

Част I



РЕГУЛАЦИЯ НА ГЕННАТА ЕКСПРЕСИЯ ПРИ ЕУКАРИОТИ

Част I: Транскрипция и Транскрипционен контрол

Галина Яхубян

Пловдивски Университет "Паисий Хилендарски"

ISBN: 978-619-202-150-4

Учебникът разглежда процесите, които са залегнали в основата на генната експресия при еукариотните организми, както и механизмите, които я контролират за да може в крайна сметка да се формират сложните многоклетъчни организми с тяхното безкрайно разнообразие от структури и функции. В Част I се описва процеса на транскрипция с неговите особености при еукариоти, представят се регулаторните елементи и транскрипционните фактори, техните взаимодействия и ролята им във фината система за транскрипционен контрол на генната експресия.

Учебникът е предназначен за студенти, които изучават биологически науки, той може да бъде използван и от студенти по медицина, както и от студенти, обучавани в биотехнологични направления.

Съдържание

I. Регулация на генната експресия при еукариоти.....	5
II. Транскрипция.....	11
1. Основи на транскрипцията.....	11
1.1. Транскрипционна единица.....	11
1.2. Етапи на транскрипцията.....	12
1.3. Базален транскрипционен апарат.....	15
1.4. Промотор.....	17
2. Базална транскрипция на РНК полимераза II	19
2.1. Фокусирана vs. нефокусирана транскрипция.....	19
2.2. <i>Cis</i> -действащи елементи на кор промотора на РНК полимераза II	18
2.3. Главни ТФ на РНК полимераза II	21
3. Базална транскрипция на РНК полимераза I.....	28
3.1. Промотор на РНК полимераза I.....	28
3.2. Главни ТФ на РНК полимераза I	29
4. Базална транскрипция на РНК полимераза III.....	31
4.1. Промотори на РНК полимераза III.....	31
4.2. Главни ТФ на РНК полимераза III.....	31
III. Транскрипционен контрол.....	33
1. Регулаторни елементи на РНК полимераза II.....	33
1.1. Къси <i>cis</i> -действащи елементи.....	33
1.1.1. Елементи на отговора	33
1.1.2. ССААТ- и CG-box.....	37
1.1.3. Други къси <i>cis</i> -действащи елементи	37
1.1.4. Механизъм на действие на късите <i>cis</i> -действащи елементи	37
1.2. Дълги регулаторни елементи.....	39
1.2.1. Енхансери.....	39
1.2.2. Локус контролиращи региони (LCR).....	41
1.2.3. Сайлънсери	42
1.2.4. Инсулатори (ICR).....	43
2. Специфични ТФ на РНК полимераза II.....	45
2.1. Регулаторен транскрипционен апарат.....	45
2.2. ДНК-свързващи мотиви и фамилии на ТФ.....	46
2.2.1. ТФ с ДНК-свързващ мотив Спирала-извивка-спирала	46
2.2.2. ТФ с ДНК-свързващ мотив „цинков пръст“.....	51
2.2.3. ТФ с основен ДНК-свързващ домен и димеризиращ мотив Левцинов цип или Спирала-примка-спирала	55
2.2.4. ТФ с други ДНК-свързващи мотиви.....	57
3. Активиране и потискане на транскрипцията на РНК полимераза II.....	58
3.1 Активиращи домени.....	58
3.2. Механизми на действие на активаторите.....	59
3.3. Потискане на транскрипцията.....	62
4. Регулация на ТФ на РНК полимераза II.....	64
4.1. Регулация на синтеза на ТФ.....	

4.2. Регулация на активността на ТФ от лиганди, които навлизат в клетката.....	64
4.3. Регулация на активността на ТФ чрез фосфорилиране, индуцирано от извънклетъчни сигнални	64
4.4. Регулация на активността на ТФ чрез други пост-транслационни модификации.....	65
4.5. Регулация на активността на ТФ чрез сигнали, които регулират процесинга на техните предшественици.....	66
5. Регулация на транскрипционната елонгация на РНК полимераза II.....	69
6. Регулация на транскрипцията на РНК полимераза I и РНК полимераза III	
Референции.....	71

Ако ДНК често се отъждествява с молекулата на живота, то генната експресия, която в крайна сметка води до изпълнение на инструкциите, записани в ДНК, със сигурност е процесът на живота. Основният клетъчен метаболизъм, клетъчното делене и невероятните процеси на клетъчна диференциация и адаптация към условията на външната и вътрешната среда, всички те зависят от контролираната експресия на определен набор от гени.

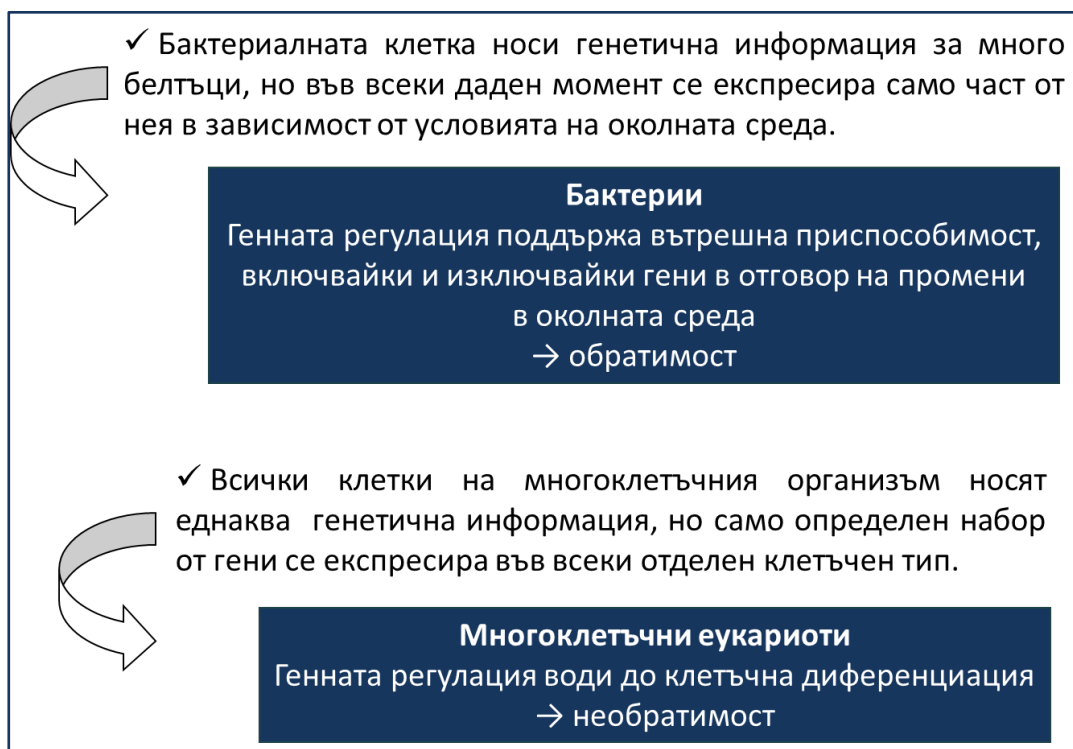
I. Регулация на генната експресия при еукариоти

Генната експресия е процес, чрез който информацията, закодирана в даден ген, се декодира във функционален генен продукт - белтък или РНК. В прокариотната клетка, този процес протича на един етап – прокариотната РНК полимераза транскрибира ДНК до мРНК, и още преди синтезата на мРНК да е завършена, рибозомите се прикрепват към растящата мРНК верига и започват да я транслират до белтък. В еукариотната клетка, транскрипцията и транслацията са разделени във времето и пространството. Синтезът на мРНК се извършва в ядрото, първичният транскрипт претърпява процесинг, който включва модификации на 5'- и 3'-краищата, сплайсинг, редактиране и др. Едва тогава зрелият транскрипт преминава в цитоплазмата, където се извършва белтъчната синтеза.

На всеки един от посочените етапи генната експресия може да бъде контролирана. Различните пътища, по които се контролират гените, могат да бъдат класифицирани в две основни категории в зависимост от нивото, на което те действат: *генна регулация на ниво транскрипция* и *генна регулация на пост-транскрипционно ниво*. В допълнение при еукариоти, има още един важно ниво за контрол на генната експресия – структурата на хроматина.



- **Основен принцип на регулация на генната експресия при еукариоти**



В сравнение с прокариотите, многоклетъчните еукариоти имат много по-голям геном, който е организиран в множество хромозоми с много по-висока височина на сложност на първичната структура (нуклеотидната последователност). Всички соматични клетки на многоклетъчния организъм носят една и съща ДНК, но тези клетки образуват тъкани с изключително различни морфология, свойства и функции. Как може да бъде обяснено това огромно разнообразие от клетъчни типове в един организъм? Отговорът на този въпрос се крие в *диференциалната генна експресия* – в различните клетъчни типове са „включени“ и „изключени“ различни гени. При това, профилът на експресирани генни продукти (РНК или белтък) може да се различава драматично, не само по това кои от тях се експресират в дадения момент и в дадения клетъчен тип, а и по нивото на експресия. Разбира се, подобно на прокариотите, еукариотите отговарят на сигналите на околната среда чрез регулиране на генната експресия. Но при многоклетъчния еукариотен организъм има и друго ниво на регулация, което е резултат на междуклетъчните взаимодействия вътре в организма и направлява индивидуалното му развитие.

Макар че генната експресия при прокариоти и еукариоти се базира на общи принципи, многоклетъчните еукариоти контролират клетъчната диференциация чрез много по-сложна и прецизна регулация на генната експресия във времето и пространството.

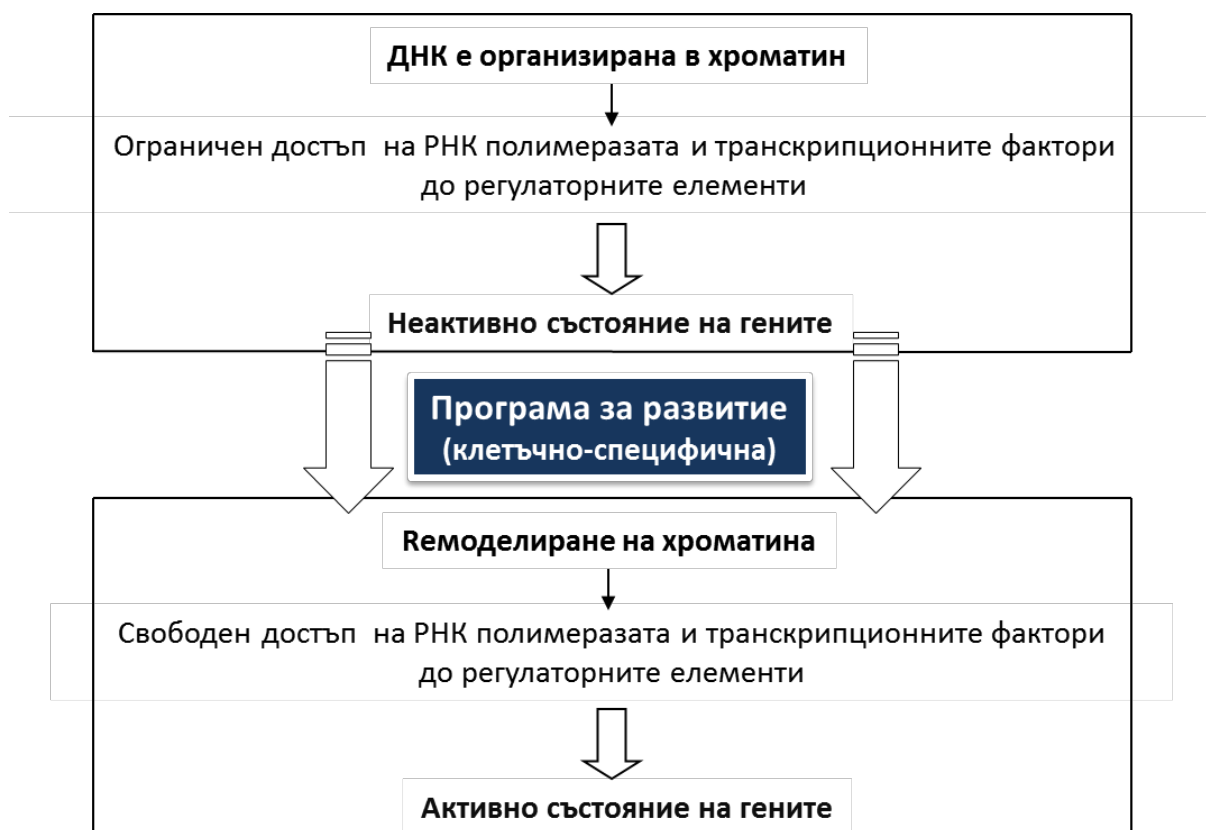
- **Модел на регулация на генната експресия при еукариоти**

Най-важната структурна разлика между прокариотната и еукариотната ДНК е свързването на еукариотната ДНК с белтъци (хистонови и нехистонови) и образуването

на хроматин в еукариотното ядро. Хроматинът обуславя напълно различно „изходно ниво“ на транскрипция в еукариотната клетка, в сравнение с прокариотната клетка.

При прокариоти, РНК полимеразата има достъп до промоторните области в ДНК веригата, поради което „изходното състояние“ на транскрипция е нерестриктивно или „включено“. Инициацията на транскрипцията се потиска, когато бъде блокирано свързването на РНК полимеразата от репресорни регулаторни белтъци.

При еукариоти, плътното пакетиране на ДНК в хроматин ограничава достъпа на РНК полимеразата до промоторните области, в следствие на което „изходното състояние“ на транскрипция е рестриктивно или „изключено“. Еукариотните РНК полимерази могат да започнат транскрипция само с помощта на белтъци, **наречени транскрипционни фактори (ТФ)**. ФНО дори когато ТФ присъстват в клетката, транскрипция не винаги може да бъде осъществена, защото много често ТФ нямат достъп до прицелните регулаторни елементи в ДНК. Затова необходимо условие за извършване на транскрипцията е преминаването на хроматина в състояние на по-хлабаво пакетиране.



Хроматиновата структура е отговорна за различни нива на сложност на генна регулация. Тя позволява едновременно регулиране на структурно или функционално свързани гени, които могат да бъдат пространствено отдалечени в еукариотната ДНК. Взаимодействието на хроматина с активиращи или репресиращи ТФ води до формиране на „отворени“ домени (хлабаво пакетиран хроматин) или „затворени“ домени (плътно пакетиран хроматин), с различни размери и стабилност. Тези вариации позволяват извършването на феномени, наблюдавани единствено в еукариотите, като генно активиране на различни етапи от развитието и епигенетична

памет при клетъчното делене. Те са отговорни също така за поддържането на диференцирано клетъчно състояние, което е от съществено значение за преживяване на многоклетъчния организъм.

Имайки предвид рестриктивното „изходно състояние“ на транскрипция, позитивната регулация чрез ТФ (активиращи ТФ или активатори) е доминиращата форма на контрол при еукариоти. Свързването на един тип ТФ може да повлияе свързването на други ТФ. Генната експресия при еукариоти е силно вариабилна в зависимост от типа на участващите активатори и сигналните молекули, които присъстват за да контролират свързването на активаторите.

• Координирано-регулираните гени не са свързани при еукариоти

При прокариоти, когато няколко гена, кодиращи отделни белтъци, се експресират в отговор на определен сигнал, то те са свързани заедно в оперон. В отговор на активирания сигнал, гените на оперона се транскрибират в една полицистронна (мултигенна) мРНК молекула, и трансляцията на всеки протеин от вътрешен AUG инициращ кодон води до координирана продукция на белтъците, кодирани от индивидуалните гени. Типичен пример за това са лактозо-индуцируемите гени, обединени в лактозния оперон (lac-оперон). Ефикасното използване на лактозата изисква не само ензима β -галактозидаза, който разрязва лактозната молекула, но също така и пермеаза, която улеснява приемането на лактозата от клетката и транс-ацетиращ ензим. Гените, кодиращи тези три белтъка, трябва да бъдат активирани в отговор на лактозата и това се постига чрез свързването им заедно в един оперон, чиято експресия е зависима от присъствието на лактозата.

При висшите еукариоти, такъв тип организираност на гени в оперони не съществува. По-скоро, всеки ген се транскрибира в индивидуална моноцистронна мРНК, кодираща единичен полипептид. Тъй като гените, чиито белтъчни продукти са необходими едновременно, не са свързани, а са представени в генома самостоятелно, тяхната експресия трябва да бъде координирана. Нещо повече, такива координирано експресирани гени могат дори да не са разположени в близост в генома, а да бъдат представени върху различни хромозоми в еукариотното ядро. Така например, образуването на функционални антитела в В клетките при бозайници изисква синтез на белтъците на тежката и леката верига на имуноглобулините. Гените, кодиращи белтъците на тежката и леката верига, се откриват върху различни хромозоми; локусът на тежката верига се намира в хромозома 14, а локусите на леката верига се откриват върху хромозоми 2 и 22. Подобно, образуването на функционална глобинова молекула изисква асоцииране на α -глобинови и β -глобинови молекули. Гените, кодиращи различните членове на α -глобиновата фамилия са локализирани върху хромозома 16 при човек, а гените кодиращи β -глобините са върху хромозома 11.

Тази съществена разлика между прокариоти и еукариоти най-вероятно отразява необходимостта от по-голяма гъвкавост в регулирането на генната експресия при еукариоти. Така организирането на гените в прокариотния геном, макар че предоставя лесен механизъм за координиране на експресията на различни белтъци, също означава най-общо, че експресията на един белтък ще води след себе си и образуване на други координирано експресирани белтъци. За разлика от прокариотната система, еукариотната система позволява α -глобинът да бъде продуциран заедно с β -глобина

във възрастния организъм, но също така да бъде свързан с друг подобен на β -глобина белтък, а именно γ -глобина, в ембриона.

- **Модел на Бритън и Дейвидсън за координирана регулация на несвързани гени**

По-голямата гъвкавост в регулирането на генната експресия при еукариоти обаче налага съществуването на начини за координиране на генната експресия, които позволяват образуването на α - и β -глобин в ретикулоцитите на възрастния индивид, на γ - и β -глобин в ретикулоцитите във ембриона, и на тежките и леките вериги на имуноглобулините в продуциращите антитела В клетки. За първи път модел на механизъм за такава координирана регулация е бил представен от Бритън и Дейвидсън през 1969 г. Те предполагат, че гените, които се регулират паралелно в отговор на определен сигнал, би трябвало да включват еднакъв регулаторен елемент, който да активира гените в отговор на сигнала. Индивидуалните гени могат да съдържат повече от един регулаторни елементи, някой от които могат да присъстват и в други гени, които на свой ред могат да притежават елементи, които не се откриват в първия ген. Специфични сигнали, причиняващи генно активиране, биха могли да действат като стимулират специфичен интегриращ ген, чиито продукт би активирал всички гени, които съдържат определен секвенционен елемент. Този механизъм би позволил наблюдаваното активиране на отделни, но припокриващи се, набори от гени в отговор на специфични стимули чрез активиране на определени интегриращи гени.

Въпреки, че този модел е създаден преди повече от 40 години, когато нашето разбиране за генната регулация на еукариоти е било значително по-ограничено отколкото понастоящем, той продължава да служи като отправна точка при формулиране на стратегията за транскрипционен контрол на генната експресия при еукариоти. *Прилагайки модерните понятия, интегриращият ген би могъл да се счита за ген, кодиращ ТФ, който се свързва с регулаторни елементи в ДНК и активира експресията на специфични прицелни гени.* Сега е ясно, че активирането на подобен фактор от даден сигнал или в специфична клетъчна тъкан може да стане или чрез *de novo* синтез на фактора, както са предположили Бритън и Дейвидсън, или чрез директно активиране на съществуващ белтък, напр. фосфорилиране в отговор на активирания сигнал. Днес, концепцията, че такива фактори действат чрез свързване с ДНК секвенции, които са общи за координирано регулираните гени, е общоприета. В оригиналния модел на Бритън и Дейвидсън се предполага, че ролята на регулаторните секвенции се изпълнява от клас повтарящи се ДНК секвенции с дължина 200–300 бд, за които е било установено, че присъстват в много копия в генома. Днес е ясно обаче, че прицелните ДНК места за свързване на ТФ са много по-къси.

- **Транскрипцията - основно ниво на регулация на генната експресия**

Процесът на транскрипция, чрез който се синтезира РНК продукт върху матрица ДНК, е от съществено значение за генната експресия. Ако този процес не протече успешно, всички стъпки, които следват след образуването на РНК като сплайсинг, транспорт към цитоплазмата и транслиране в белтък, няма как да бъдат осъществени. *Централната роля на транскрипцията в процеса на генна експресия определя транскрипцията като основно ниво на регулация на експресията на гените в различните клетъчни*

типове в отговор на специфични сигнали. Например, гените, кодиращи тежките и леките вериги на имуноглобулините се транскрибират във високи нива единствено в В клетките, които произвеждат антителата. Ключов етап в транскрипционния контрол е транскрипционната инициация, когато ДНК започва да се копира в РНК.

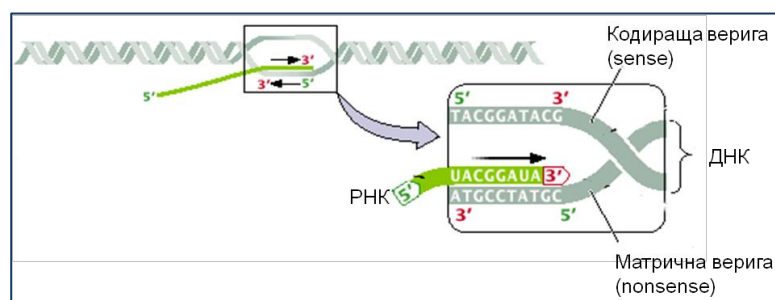
Стратегията, която еукариотната клетка използва за регулиране на генната експресия на ниво транскрипция, се основава на взаимодействието на trans-действащи фактори и cis-действащи елементи.

Trans-действащите фактори са дифундиращи компоненти на транскрипционния апарат като ТФ и малки РНК молекули, които действат върху хромозома, различна от тази, в която са кодирани. За разлика от тях, участъците от ДНК, с които транскрипционния апарат контактува, се означават като *cis*-действащи елементи, тъй като те могат да повлияват транскрипцията само върху хромозомата, в която се откриват. *Cis*-действащите елементи (или регулаторни елементи) са основни компоненти на промоторите, към които се свързват компонентите на базалния транскрипционен апарат, както и на местата за свързване на активатори и репресори.

II. Транскрипция

1. Основи на транскрипцията и особености при еукариоти

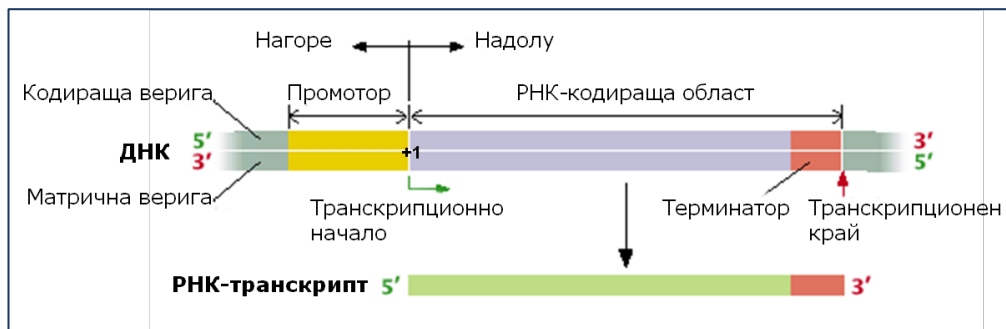
Транскрипцията е процес на пренос на генетичната информация от ДНК към РНК, който се извършва чрез синтез на РНК върху матрица ДНК. Процесът се катализира от ензима ДНК-зависима РНК полимераза (или само РНК полимераза). Ензимът се движи по едната верига на ДНК, която се нарича **матрична верига** и има 3'-5' поляритет, докато ДНК веригата с 5'-3' поляритет се нарича **кодираща верига** (или смислена) (Фигура 1). Новосинтезираната РНК е идентична с кодиращата ДНК верига с единствената разлика, че вместо тимин (Т) съдържа урацил (У). Докато прокариотната клетка използва една РНК полимераза, то еукариотната клетка използва три различни РНК полимерази – РНК полимераза I, РНК полимераза II и РНК полимераза III, а в растителната клетка се откриват още две - РНК полимераза VI и РНК полимераза V. Транскрипцията в митохондриите и пластидите се извършва от специфични за органелите РНК полимерази.



Фигура 1. Транскрипция - процес на пренос на генетичната информация от ДНК към РНК

1.1. Транскрипционна единица

Транскрипционната единица представлява участък от ДНК, в който се извършва транскрипция и се продуцира една РНК молекула като в нея обикновено се съдържа един ген (моноцистронна РНК) и по-рядко повече от един гени (полицистронна РНК). Транскрипционната единица включва три компонента: (1) промотор, (2) РНК-кодираща секвенция и (3) терминатор (Фигура 2). Нуклеотидът в кодиращата верига на ДНК, от който РНК полимеразата започва транскрипцията, и който кореспондира на първия нуклеотид в транскрибираната РНК, се нарича *Транскрипционно начало* (Transcription start site - TSS) и се бележи с +1. Много често стартовата точка е средната база на секвенцията CAT, но нейната консервативност не е достатъчна за да служи като сигнал за инициация на транскрипцията. Спрямо TSS са дефинирани посоките в транскрипционната единица – посоката, която съвпада с посоката на движение на РНК полимеразата се означава като *надолу* от TSS, а обратната на нея посока, се означава като *нагоре* от TSS. При повечето гени, промоторът се разполага непосредствено нагоре от TSS, той не се транскрибира, но има незаменима регулаторна функция, тъй като върху него се сглобява транскрипционния апарат.



Фигура 2. Транскрипционна единица

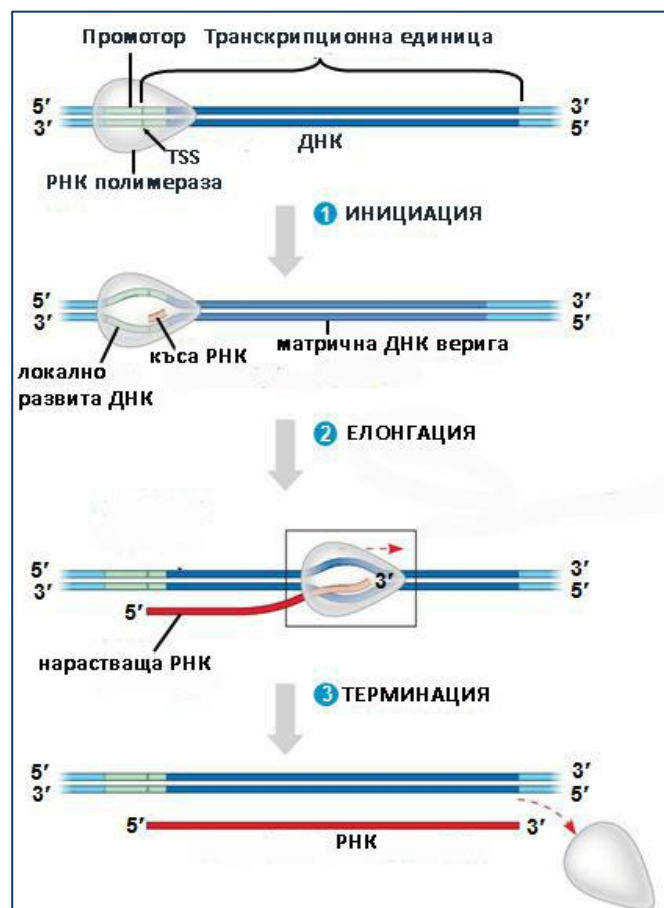
1.2. Етапи на транскрипцията

Еукариотната транскрипция, подобно на прокариотната транскрипция, включва три основни етапа (Фигура 3):

Инициация – РНК полимеразата се свързва с ДНК-матрицата в областта на промотора. Създава се *транскрипционното мехурче* чрез локално развиване на ДНК в областта на свързване на РНК полимеразата. РНК полимеразата извършва множество цикли на абортивна инициация, при които ензимът синтезира къси РНК, но не може да ги удължи. Преминването към продуктивна инициация и освобождаването на промотора се постига, когато транскриптът успее да формира стабилен елонгационен комплекс.

Елонгация - РНК полимеразата се движи върху матричната ДНК верига в посока 3'-5' и синтезира РНК в посока 5'-3' като прибавя нуклеотиди към 3'-края на нарастващия транскрипт. Синтезът на РНК се извършва в транскрипционното мехурче, което обхваща участък от ДНК с дължина 25 бд, в който само 8 нт от новосинтезиращата се РНК молекула са свързани с матричната ДНК и образуват хибриден ДНК-РНК участък. За да продължи синтеза на РНК, РНК полимеразата се движи по ДНК матрицата, откривайки нови области за комплементарно свързване. Така по веригата на ДНК се движи един локално-разплетен хибриден участък, след който веригата се затваря и РНК се освобождава.

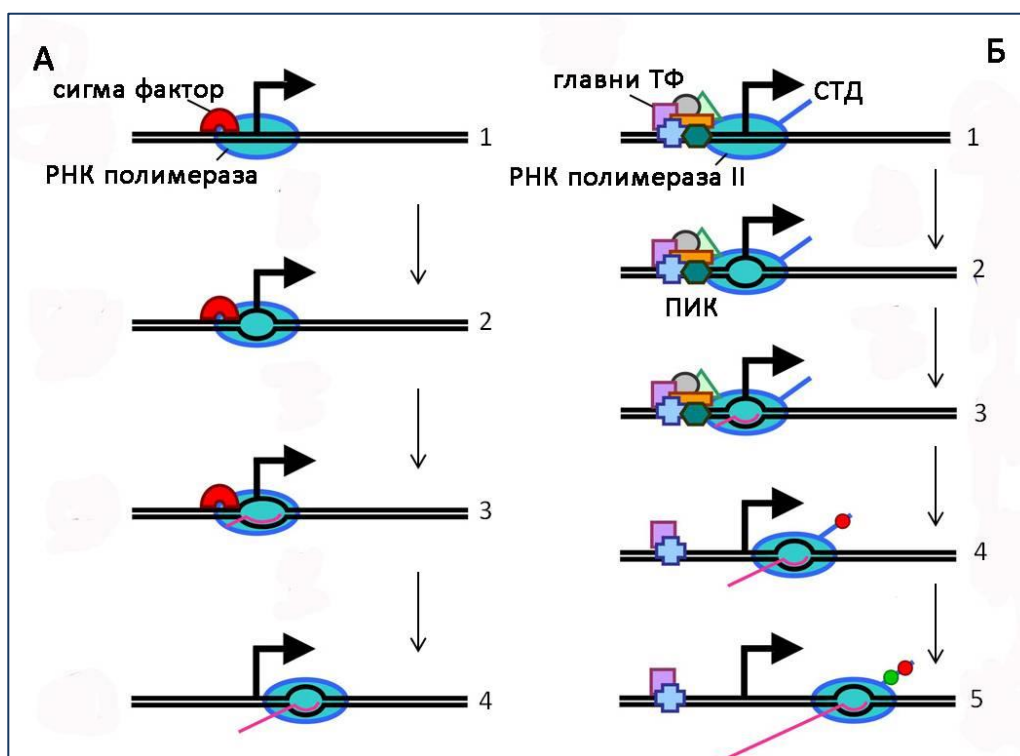
Терминация – синтезата на РНК се прекратява в определен участък на ДНК, наречен терминатор, и се освобождава новосинтезираната РНК и ензима.



Фигура 3. Етапи на транскрипцията

- **Особености при еукариоти**

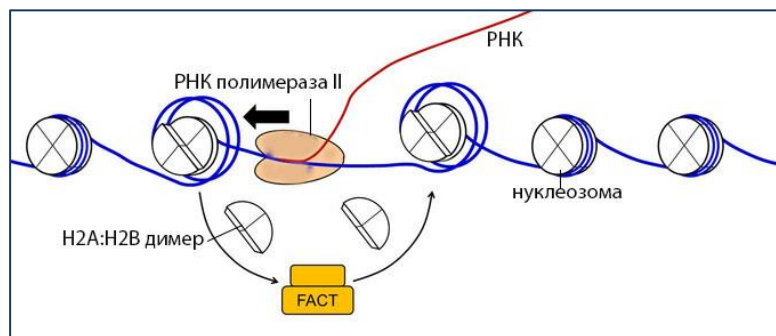
Инициация - За разлика от прокариотната РНК полимераза, която притежава вроденото свойство да разпознава промотора и може сама да се свързва към ДНК, еукариотните РНК полимерази се нуждаят от други белтъчни фактори - главни ТФ, които се свързват с промоторната област и правят възможно присъединяването на РНК полимеразата. Така напр. комплексът на РНК полимеразата II и нейните главни ТФ, който е свързан към промотора, се означава като транскрипционен преинициаторен комплекс (ПИК). Щом като двойно-верижната ДНК е била разплетена в областта на TSS, РНК полимеразата II е позиционирана върху +1 нуклеотида и е започнала катализирането на процеса на синтез на РНК, ензимът напуска промоторния регион и оставя след себе си повечето от главните ТФ (Фигура 4).



Фигура 4. Стъпки на инициацията на транскрипцията

А. Стъпки на инициацията на транскрипцията при прокариоти: (1) Образуване на преинициаторен затворен комплекс (ПИК) върху промотора от холоензимния комплекс (включващ σ фактора) на бактериалната РНК полимераза; (2) Локално развиване на ДНК в областта на транскрипционното начало и образуване на отворен комплекс; (3) Абортивна синтеза на къси РНК с дължина 2-15 нт; (4) Освобождаване на промотора от РНК полимеразата със загуба на σ фактора. **Б.** Стъпки на инициацията на транскрипцията при еукариоти: (1) Образуване на преинициаторен затворен комплекс (ПИК) върху промотора от РНК полимеразата II и главните ТФ; (2) Локално развиване на ДНК в областта на транскрипционното начало и образуване на отворен комплекс; (3) Абортивна синтеза на къси РНК с дължина 2-3 нт; (4) Освобождаване на промотора от РНК полимеразата, което е свързано с дисоцииране на повечето главни ТФ и с фосфорилиране (червено кръгче) на С-терминалния домен (СТД) на най-голямата субединица на РНК полимеразата II. При някои еукариоти ензимът спира след синтез на РНК с 20-50 нт; (5) Възстановяване на движението на РНК полимеразата II, което е свързано с допълнително фосфорилиране (зелено кръгче) на СТД на най-голямата субединица на РНК полимеразата II. (по Wade and Struhl, 2008, с модификации)

Елонгация - Въпреки, че ензимните процеси на елонгация са по същество еднакви при прокариоти и еукариоти, еукариотната ДНК матрица е много по-сложна. Дори когато еукариотните клетки не се делят, техните гени са пакетирани с хистоновите белтъци в хроматин. За да се извършва полинуклеотидния синтез, РНК полимеразата трябва да отмести хистоновите белтъци всеки път, когато се натъкне на нуклеозома. Това става с помощта на специален димерен белтък, наречен FACT (facilitates chromatin transcription, англ. – улеснява транскрипцията през хроматина). FACT частично разглобява нуклеозомата надолу от РНК полимеразата като премахва два от осемте хистона - един H2A:H2B димер (Фигура 5).

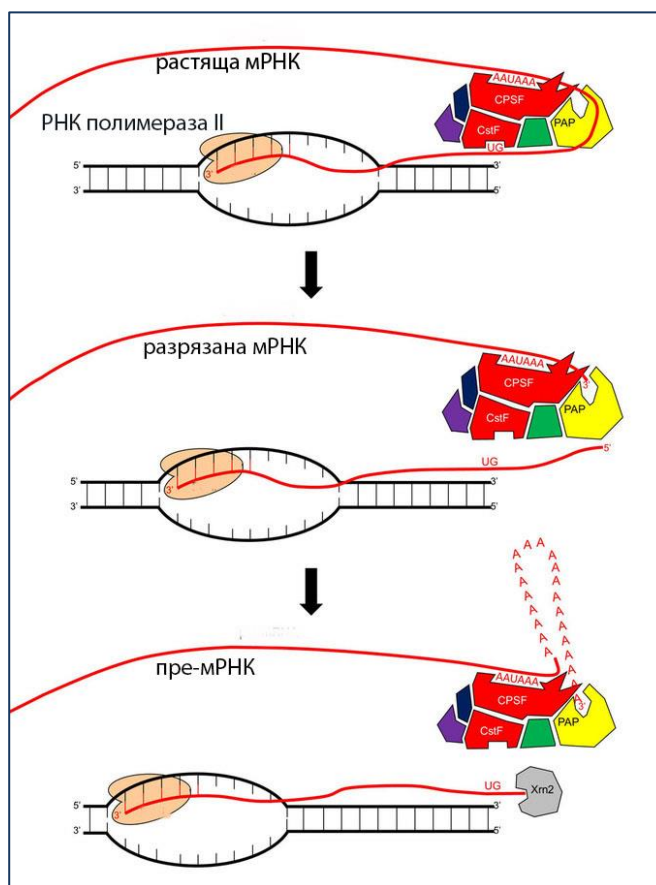


Фигура 5. Участие на FACT в елонгацията на транскрипцията при еукариоти (от [www. boundless.com](http://www.boundless.com))

Това е достатъчно да разхлаби увиването на ДНК веригата около нуклеозомното ядро, така че полимеразата да може да транскрибира разхлабения ДНК участък. След преминаването на РНК полимеразата, FACT сглобява нуклеозомата като помага на липсващите хистони да се присъединят към нея.

Терминация - Терминацията на транскрипцията е различна за трите основни еукариотни РНК полимерази. Гените на рибозомните РНК (рРНК), които се транскрибират от РНК полимераза I, съдържат специфична секвенция (11 бп при човека, 18 бп при мишка), която се разпознава от терминационния белтък TTF-1 (Transcription Termination Factor for RNA Polymerase I, англ. – транскрипционен терминаращ фактор на РНК полимераза I), който блокира транскрипцията.

При протеин-кодиращите гени, транскрибирани от РНК полимераза II, подобен специфичен сигнал не се открива. РНК полимераза II може да продължи транскрипцията след истинския край на гена с няколко до хиляди нуклеотида. Въпреки това, транскриптът се разрязва в специфичен вътрешен сайт, още преди ензимът да е приключил транскрипцията (Фигура 6).



Фигура 6. Терминация на транскрипцията при еукариоти (от [www. boundless.com](http://www.boundless.com))

Изрязаната 5'-част от транскрипта представлява първичния транскрипт (пре-мРНК), който търпи след това процесинг и се превръща в зряла мРНК, която се транслира. Остатъкът от транскрипта се разгражда от 5'-екзонуклеаза (Xrn2 при човек), докато той все още се транскрибира от РНК полимеразата II. Когато 5'-екзонуклеазата настигне полимеразата, тя подпомага освобождаването на полимеразата от ДНК матрицата.

Терминиращият сигнал на РНК полимеразата III все още не е напълно установен.

1.3. Базален транскрипционен апарат

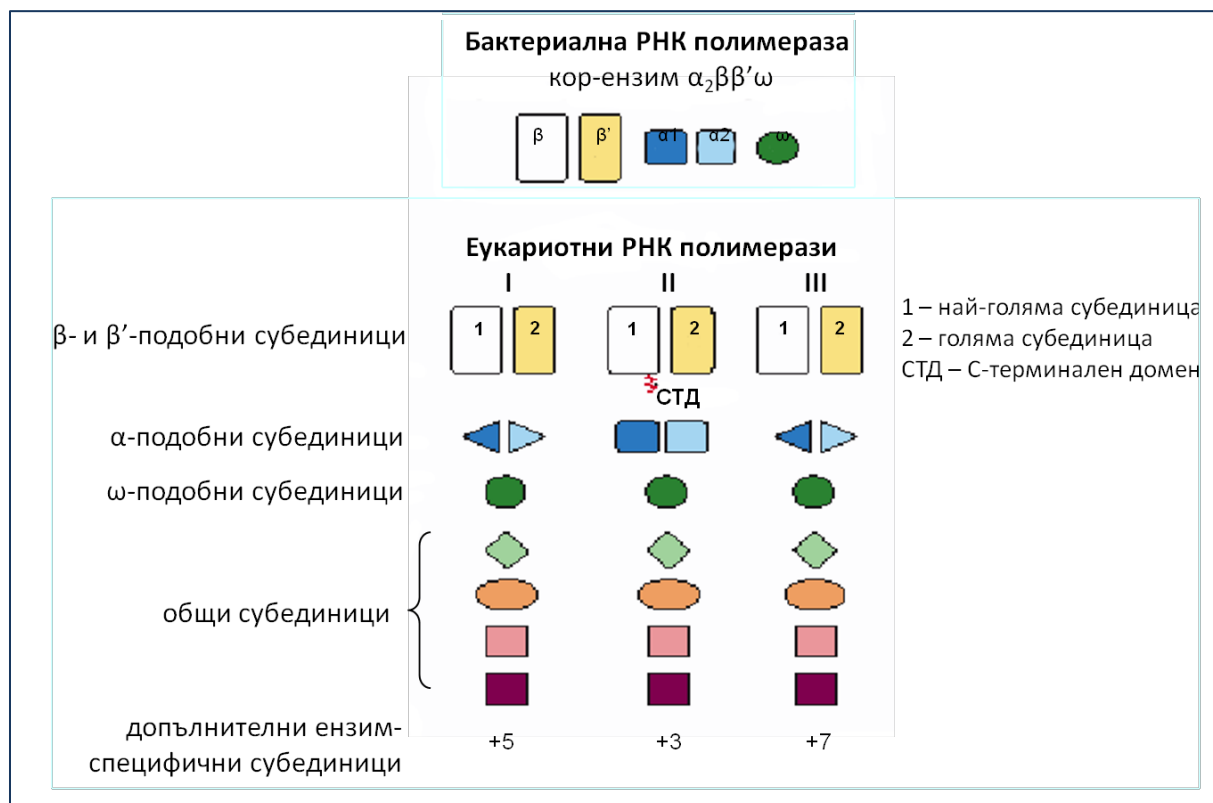
Трите основни еукариотни РНК полимерази - РНК полимеразата I, РНК полимеразата II и РНК полимеразата III осъществяват транскрипцията на различен набор от гени. РНК полимеразата II транскрибира всички гени, които кодират белтъци, както и някои гени, кодиращи малки РНК (напр. малки ядрени РНК, участващи в сплайсинга, микроРНК и др). РНК полимеразата I транскрибира гените, кодиращи 28S, 18S и 5,8S рРНК, а РНК полимеразата III транскрибира гените, кодиращи различните тРНК и 5S рРНК. Трите РНК полимерази могат да бъдат разграничени въз основа на тяхната различна чувствителност към гъбния токсин алфа-аманитин.

Ядрените РНК полимерази представляват **големи мулти-субединични ензими** (Фигура 7). Например, броят на субединиците на РНК полимеразата II при различните видове варира от 10 до 14, а размерът им е в диапазона 220 до 10 килодалтона. Клонирането на гените, кодиращи най-големите субединици на ядрените РНК полимерази, разкрива висока степен на хомоложност между тях. Експерименти, използващи химическо белязване, сочат че втората най-голяма субединица носи активното място на ензима. Поне три от по-малките некаталитични субединици са общи за трите консервативни РНК полимерази. Такива взаимоотношения са показателни за базово функционално подобие на ядрените РНК полимерази, както и за общ еволюционен произход.

В допълнение към функционалната консервативност между трите основни РНК полимерази, всека една от тях показва **строга консервативност между организмите**. Така напр. най-голямата субединица на РНК полимеразата II на бозайници има 75% хомоложност с тази на винената мушица *Drosophila*, а също така показва хомоложност с нейния еквивалент при дрожди и дори при *E. coli*. **C-терминалният домен (СТД)** на най-голямата субединица (RPB1) на РНК полимеразата II съдържа множество дегенеративни копия на хептапептида Tyr-Ser-Pro-Thr-Ser-Pro-Ser. Напр. броят на копията на секвенцията е 52 при мишка и 26 при дрожди. Секвенцията има много висока еволюционна консервативност, и както може да се очаква от това, е незаменим за точното функциониране на ензима, и от тук за клетъчната жизненост, макар, че броят на повторите може да бъде намален до определена степен без това да повлияе ензимната активност.

СТД служи като място за фосфорилиране, което е от съществено значение за функционирането на полимеразите. Дефосфорилираната форма на РНК полимеразата II е формата, която влиза в базалния транскрипционен комплекс, а фосфорилираната форма е тази, която стартира транскрипция за да произведе РНК продукт. В допълнение на това, СТД е прицелно място за транскрипционни активатори и репресори, регулиращи транскрипционната елонгация. Нещо повече, нови изследвания разкриват, че фактори, участващи в пост-транскрипционни процеси като РНК сплайсинг, се свързват с този домен, така че растящият РНК транскрипт,

продуциран от полимеразата, да бъде процесирен от фактори, като самите те са свързани с ензима.



Фигура 7. Структура на прокариотната и еукариотните РНК полимерази

Въпреки, че РНК полимеразите притежават ензимната активност, необходима за транскрипцията, те не могат да функционират самостоятелно, тъй като не могат да разпознават промотора. По-скоро, процесът на транскрипция включва множество ТФ, които взаимодействат както по-между си, така и с РНК полимеразата. Ролята на тези фактори е да разпознаят промотора и да иницират транскрипцията като организират формирането на *стабилен транскрипционен апарат*, който да включва РНК полимеразата и да е способен да осъществява повтарящи се цикли на транскрипция, в резултат на които да се продуцират много РНК копия на един и същи ген. Тъй като факторите са необходими за сглобяването на комплекса, но не и за самата транскрипция, те се дисоциират, веднага щом РНК полимеразата започне транскрипция.

Факторите, необходими за транскрипцията, варират значително от РНК полимеразата, на която служат. Бактериалната РНК полимераза разпознава промотора с помощта на единствен фактор, наречен σ фактор. При еукариоти и археи няма σ фактори, вместо тях РНК полимеразите използват за инициация множество **главни ТФ** (наричани още **базални ТФ**). Докато главните ТФ на еукариоти и археи показват родство помежду си,

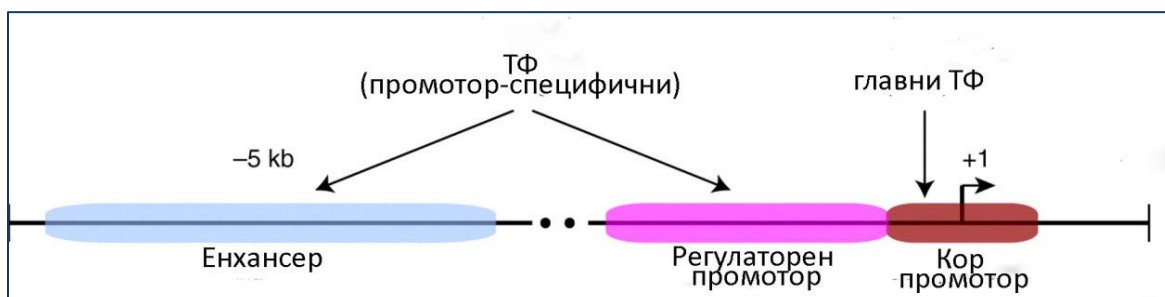
Археи – заедно с Бактерии и Еукариоти играждат трите основни домена на живите организми; представляват едноклетъчни микроорганизми с прокариотен тип на клетъчен строеж; отличават се от бактериите по много свойства – мембранен строеж, клетъчна стена, наличие на интрони в генома, нуклеотидната последователност на 16S рРНК и др.; огромно разнообразие във физиологично и екологично отношение - много от тях са способни да живеят в екстремални условия при строга анаеробиоза, в горещи и силно солени водни източници.

то те нямат никаква родствена връзка с бактериалния σ фактор. За разлика от просто устроения базален транскрипционен апарат на прокариоти, еукариотният базален апарат може да бъде много сложно устроен. Например, РНК полимеразата II използва шест главни ТФ, изградени от повече от 20 полипептида.

1.4. Промотор

Докато при прокариоти **промоторът** се разглежда като ДНК секвенция, към която се свързва РНК полимеразата за да инициира транскрипция, то при еукариоти терминът промотор се използва за да се опишат всички секвенции, необходими за инициация на транскрипцията. За някои еукариотни гени тези секвенции могат да бъдат много на брой и с различни функции. Типичният еукариотен промотор включва **кор промотор** (наричан още базален или минимален промотор), който е мястото, върху което се сглобява базалния транскрипционен апарат (РНК полимеразата и главните ТФ), и **регулаторен промотор**, който включва разположени нагоре от транскрипционното начало *cis*-действащи елементи, към които се свързват специфични ТФ като активатори или репресори (Фигура 8). Кор промоторът обхваща областта от +40 до -50 спрямо транскрипционното начало, а регулаторният промотор се разполага нагоре от -50. Така описаната структура на промотора има най-много вариации при РНК полимеразата II, която транскрибира протеин-кодиращите гени. В транскрипционния контрол на голяма част от тези гени се включват и *cis*-действащи елементи, разположени на голямо разстояние от транскрипционното начало каквито са например енхансерите. РНК полимеразата II е обект на най-сложния транскрипционен контрол във времето и пространството. Анализът на промотора може да бъде осъществен на ниво ДНК секвенции, ТФ и биологични функции.

Кор промоторът е минималната секвенция (най-късата ДНК секвенция), която е достатъчна за еукариотната РНК полимеразата да инициира транскрипция. По принцип кор промоторът може да инициира транскрипция във всяка клетка. В действителност обаче, поставен в контекста на хроматина *in vivo*, кор промоторът, ако въобще инициира транскрипция, то тя е слаба. Кор промоторът прави възможно, с помощта на главните ТФ, позиционирането на РНК полимеразата при стартовата точка на транскрипция, определя посоката на транскрипция, както и коя от двете ДНК вериги да бъде прочетена като матрица.



Фигура 8 . Еукариотен промотор

Системният анализ на първичната структура на промотора разкрива **неговата модулна организация**. Всеки модул може да бъде разглеждан като набор от *cis*-действащи елементи, които носят свързващи места за ТФ. Модулите могат да бъдат локализирани

на по-малко или по-голямо разстояние от стартовата точка на транскрипция в посока нагоре или надолу от нея. Всеки промотор е изграден от специфичен набор от модули, които определят неговите уникални свойства. Няма универсални промоторни елементи, които да се присъстват във всеки промотор, но има такива, които се срещат преимуществено в голям брой промотори напр. Inr и TATA елементите, които се откриват поотделно в кор промоторите на много еукариотни гени.

Всяка от еукариотните РНК полимерази разпознава специфични промоторни секвенции; всъщност *специфичността на промоторните секвенции определя кой ген от коя полимераза ще бъде транскрибиран.*

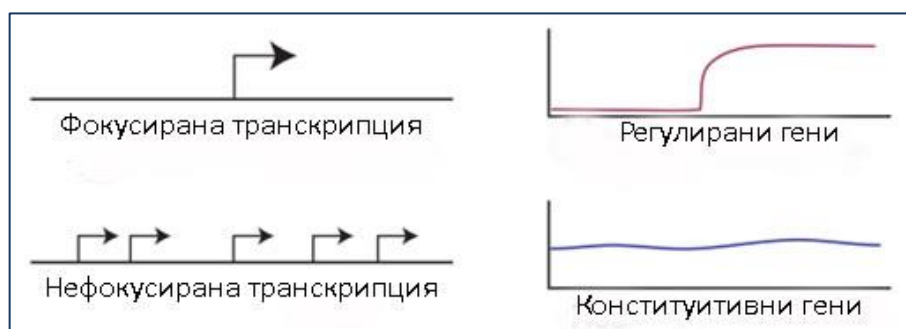
2. Базална транскрипция на РНК полимера II

РНК полимерата II е комплекс от 12 субединици, които позволяват на ензимът да изпълнява полимеразна и хеликазна активности, да поддържа развития ДНК участък в едноверижно състояние, да се движи по матричната верига, както и други функции. Следователно РНК полимеразата не се нуждае от много допълнителни белтъци за да катализира синтеза на новата РНК верига по време на транскрипционната елонгация.

2.1. Фокусирана vs. нефокусирана транскрипция

Изследванията върху моделите на транскрипционната инициация показват два различни начина на инициация на транскрипцията - фокусирана и нефокусирана инициация. При **фокусираната инициация**, транскрипцията започва от единствен нуклеотид или от няколко стартови места в тесен регион от няколко нуклеотида, докато при **нефокусираната инициация** (наричана още *дисперсна*) съществуват множество слаби стартови места в по-широк регион от около 50-100 нт (Фигура 9).

Фокусираната инициация се среща при всички организми, и изглежда е преобладаващия или единствения начин на транскрипция при низшите организми. При гръбначни обаче, около 70% от гените имат нефокусирани или дисперсни промотори, които типично се откриват в CpG островите. *Принципно, фокусираните промотори се свързват с регулираните гени, докато нефокусираните промотори се използват от конститутивните гени.* Анализите на фокусираните кор промотори са довели до откриването на елементи като TATA box, BREu, BREd, Inr, MTE, DPE, DCE and XCPE1. За разлика от фокусираните промотори, нефокусираните промотори не съдържат изброените елементи. Най-вероятно, съществуват фундаментални различия между механизмите на транскрипция, иницирана от фокусиран или нефокусиран промотор.



Фигура 9 . Фокусирана vs. нефокусирана транскрипция

2.2. Cis-действащи елементи на кор промотора на РНК полимера II

- **Инициаторен елемент Inr**

Inr елементът (initiator, англ. - инициатор) обхваща стартовата точка на транскрипция +1, и вероятно е най-често срещания елемент в кор промоторите (Фигура 10). Въпреки, че много фактори са били намерени да взаимодействат с Inr елемента, изглежда че базалната транскрипция включва свързване на TFIID към Inr. На базата на функционални анализи е установена консенсусната секвенция на Inr при човек -

YYANWYY (нуклеотиден код IUPAC) и *Drosophila* – TCAKTY, където A е +1, а при ориз и *Arabidopsis* е установен Inr мотива YR, където R е +1. По-късно, на базата на компютърно-базирани анализи консенсусната секвенция на Inr е дефинирана като YR, където R е +1, при човек, и TCAKTY (A+1) за фокусираните промотори и TCA (A+1) за дисперсните промотори на *Drosophila*.

- **Елементи TATA-box и BRE**

TATA-box е първият мотив, който е открит в еукариотните кор промотори. При метазои, консенсусната секвенция на TATA-box е TATAWAAR, където 5'Т се разполага обикновено 30-31 нт нагоре от A+1 (или G+1) в Inr елемента (Фигура 10). TATA-box се разпознава и свързва от TBP субединицата на TFIID. Както TATA-box, така и TBP, са консервативни от археи до човек. Макар, че TATA-box е един от най-добре проучените мотиви на кор промотора, той се присъства в едва 10-15% от кор промоторите на бозайници.

BRE елементът (TFIIB recognition element, англ. – елемент, който разпознава TFIIB) първоначално е идентифициран като секвенция, която се разполага нагоре от TATA-box в част от кор промоторите, съдържащи TATA-box. По-късно в други промотори, съдържащи TATA-box, е идентифициран BRE елемент, който е разположен непосредствено надолу от TATA-box, и означен като BREd (BRE downstream, англ. – разположен надолу BRE). Откриването на BREd води до преименуване на оригиналния BRE като BREu (BRE upstream, англ. – разположен нагоре BRE). Обикновено двата елемента, BREu и BREd, не се откриват едновременно в един същи кор промотор. И двата елемента обаче функционират свързано с TATA-box и могат както да повишават, така и да понижават нивата на базална транскрипция.

- **Елементи DPE и MTE**

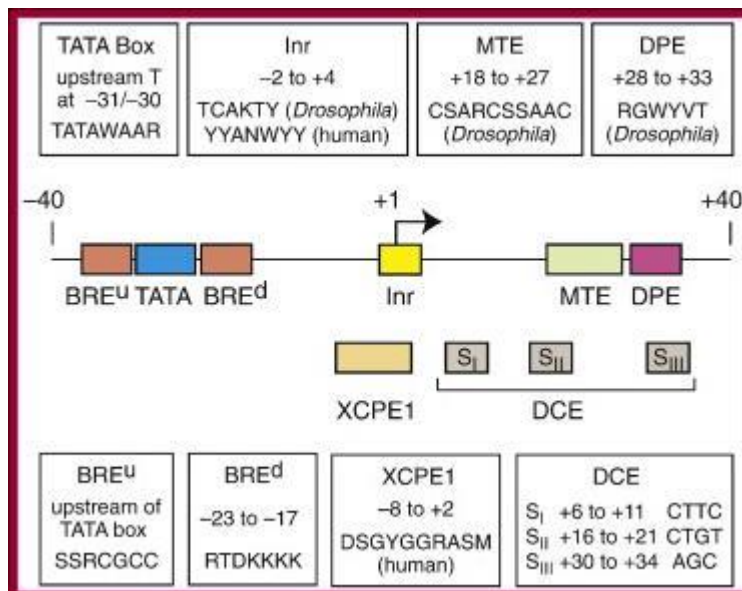
DPE елементът (downstream core promoter element, англ. – разположен надолу промоторен елемент) е открит при анализ на промотори, които не съдържат TATA-box, и е идентифициран като секвенция, която се разпознава от TFIID. DPE е локализиран точно от +28 до +33 спрямо A+1 в Inr и е консервативен за *Drosophila* и човек (Фигура 10). Разстоянието между Inr и DPE е критично за транскрипционната активност на DPE-зависимите промотори. Типичният DPE-зависим промотор съдържа по един Inr и DPE мотив. В някои случаи обаче TATA, Inr и DPE мотиви могат да бъдат намерени в един и същ промотор.

MTE елементът (motif ten element, англ. – мотив десет) е локализиран от +18 до +29 спрямо A+1 в Inr, притежава консенсусна секвенция CSARCSSAACGS и е консервативен от *Drosophila* до човек. Подобно на DPE, MTE се разпознава от TFIID. MTE функционира кооперативно с Inr, но може да действа независимо от DPE и TATA-box. Въпреки това е установено и синергично действие между MTE и DPE, както и между MTE и TATA-box. Тези изследвания са довели до създаването на супер кор промотор - СКП, който съдържа TATA, Inr, MTE, и DPE. СКП е най-силният промотор, наблюдаван *in vitro* и в клетъчни култури.

- **Елемент DCE**

За разлика от останалите елементи на кор промотора, които представляват непрекъсната ДНК секвенция, DCE съдържа три субелемента: SI, SII и SIII, чиито секвенции са съответно CTTTC, CTGTи AGC, простиращи се съответно от + 6 до + 11, от +16 до +21 и от +30 до +34 (Фигура 10). Секвенционният анализ на базата дани на

промоторните секвенции при човек показва, че промоторните елементи DCE и DPE взаимно се изключват. Crosslinking-базирани изследвания установяват, че TAF1 компонентът на TFIID контактува с всеки един от трите субелементи, което доказва, че TFIID има основна роля в установяване на промоторни контакти надолу от транскрипционното начало.



Фигура 10. Промоторни елементи на РНК полимера II.

Тези елементи се откриват обикновено във фокусираните промотори. Най-вероятно и други кор промоторни елементи предстоят да бъдат идентифицирани. Свойствата на даден кор промотор се определят от присъствието на специфични кор промоторни елементи

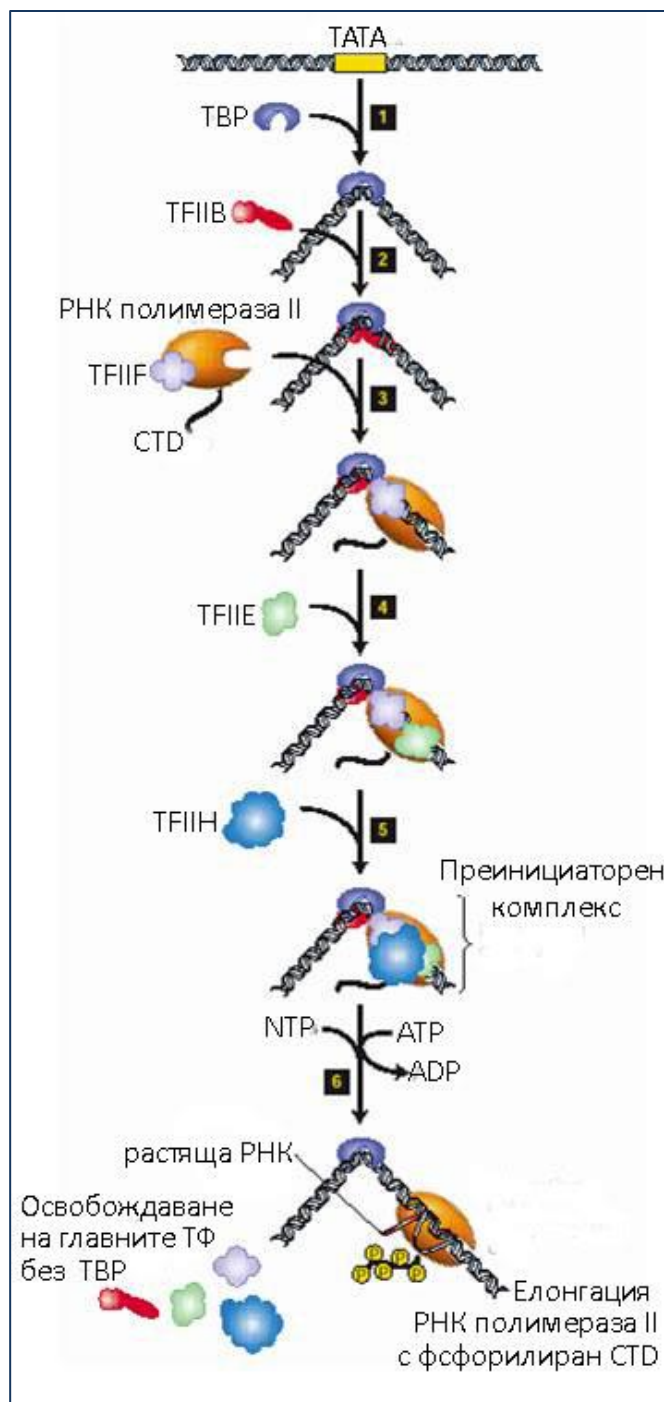
2.3. Главни ТФ на РНК полимера II

Активните изследвания на базовата организация на апарата за транскрипция на белтък-кодиращите гени започват с откритието, че пречистена РНК полимера II може да синтезира мРНК, но не може да иницира транскрипция сама, а е необходимо да се смеси с допълнителен екстракт. Пречистването на този екстракт води до идентифициране на **главните ТФ** на РНК полимера II. РНК полимера II и шест главни ТФ изграждат **преинициаторния комплекс (ПИК)**, който е необходим за инициране на транскрипцията при промоторите от клас II.

Главните ТФ се означават с TFII_X, където X идентифицира индивидуалния ТФ - **TFIIA**, **TFIIB**, **TFIID**, **TFIIE**, **TFIIF** и **TFIIN**. Подобно на субединиците на РНК полимера II, главните ТФ са консервативни при всички еукариоти. Те участват механистично в огъването на ДНК и по този начин позволяват на РНК полимера II да разпознае промотора и да иницира транскрипция.

При промоторите с TATA-box, транскрипция *in vitro* може да бъде осъществена с пречистени РНК полимераза II, TFIIB, TFIID, TFIIЕ, TFIIF и TFIIH. Тези фактори обаче не могат да опосредстват *in vitro* транскрипция от DPE-зависими. Допълнителен фактор **NC2** (negative cofactor 2, англ. – негативен фактор 2, известен още като Dr1-Drap1), който първоначално е бил идентифициран като репресор на TATA-зависимите промотори, е доказан по-късно като активатор на DPE-зависимата транскрипция. Както изглежда, TATA- vs DPE-зависимата транскрипция се контролират, поне отчасти, чрез цикъл, при който TBP активира TATA транскрипцията и потиска DPE транскрипцията, докато NC2 и Mot1 (АТФ-аза, която отделя TBP от ДНК) блокират TBP и по-този начин стимулират DPE транскрипцията, но потискат TATA транскрипцията.

Описани са два модела за сглобяване на базалния транскрипционен апарат – *модел на последователното сглобяване* и *холоензимен модел*. Според модела на последователното сглобяване инициацията на транскрипцията се извършва чрез последователно присъединяване към промотора на TFIID, TFIIB, РНК полимераза II асоциирана с TFIIF, TFIIЕ и TFIIH (Фигура 11). Според холоензимния модел, TFIID се свързва с промотора и след това към него се присъединява холоензимен комплекс на РНК полимераза II и набор от белтъци, включващ TFIIB, TFIIF, TFIIH и други белтъци, участващи в отварянето на хроматиновата структура. Независимо от особеностите, свързани със сглобяването на ПИК, и двата модела приемат, че след напускане на промотора, освобождаване на по-голяма част от главните ТФ и преминаване на РНК полимераза II към елонгация, TFIID (или поне TBP) остава свързан с кор промотора за да позволи осъществяването на повтарящи се цикли на транскрипция.



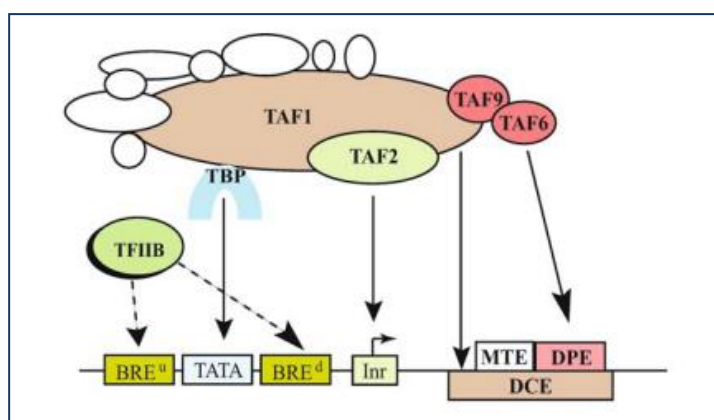
Фигура 11. Модел на последователното сглобяване на базалния транскрипционен апарат.

• TFIID

Позициониращ фактор

Първата стъпка за активиране на промоторите на РНК полимераза II е свързването на TFIID, наричан още *позициониращ фактор*. Въпреки, че първоначално TFIID е бил открит като TATA-свързващ фактор чрез *in vitro* ДНК-свързващи анализи, изглежда, че TFIID взаимодейства много активно *in vivo* при дрожди и човек и с промотори без TATA-box).

TFIID е мултисубединичен комплекс, който се състои от **TBP** (TATA-box binding protein, англ. – белтък, свързващ TATA-box) и други субединици, наречени **TAFs** (TBP-associated factors, англ. – фактори, асоциирани с TATA-box) (Фигура 12). TBP се свързва към TATA-box в малката бразда на ДНК, докато TAFs не се свързват директно с TATA-box. Съществуват множество варианти на TFIID, съдържащи различни TAFs, които разпознават промотори с различни комбинации от консервативни *cis*-действащи елементи. TAF1 и TAF2 могат да контактуват с Inr, докато TAF6 и TAF9 контактуват с DPE напр. в промотора на регулаторен фактор 1 на интерферона при човек (Shao et al., 2005), което обяснява как РНК полимераза II разпознава промотори без TATA-box. TAF1, който има няколко каталитични домена, освен че разпознава Inr, взема участие и в насочването на множество пост-транслационни модификации като фосфорилиране и ацетилиране. Голяма част от TAFs са ко-активатори, които свързват активиращи ТФ с базалния транскрипционен апарат. Някои TAFs са тъканно-специфични, други TAFs имат хомология с хистонови белтъци (H3 и H4), а една малка група асоциират с хистон-модифициращи ензими като SAGA.



Фигура 12. Разпознаване на промоторните елементи от TFIID и TFIIB (по Thomas and Chiang, 2006)

TBP – универсалният ТФ

TBP се разглежда като „универсален“ главен ТФ, тъй като той е необходим на трите еукариотни РНК полимерази (РНК полимераза I, РНК полимераза II и РНК полимераза III) за инициране на транскрипцията. Той е незаменим компонент на позициониращия фактор на всеки тип РНК полимераза. Въпреки това природата на универсално необходимата функция не е изяснена. Универсалната функция очевидно не е свързване към TATA-box, тъй като много промотори от клас II, както и почти всички промотори от класове I и III, не съдържат TATA-box. TBP има ключова роля за сглобяването на базалния транскрипционен комплекс, независимо дали се свързва директно към ДНК при промоторите с TATA-box, или се присъединява чрез белтък-белтъчни взаимодействия (напр. с различни TAFs) към промоторите без TATA-box.

TBP има уникалното свойство да се свързва към ДНК в малката бразда (болшинството ДНК-свързващи белтъци контактуват с голямата бразда). Свързването на TBP към TATA-box води до локално огъване на ДНК на $\sim 80^\circ$, а формата, която самият белтък заема,

наподобява „седло“. Свързването на изолиран TBP към ДНК *in vitro* защитава ~1 оборот на двойната спирала при TATA-box, и се разполага от -37 до -25 спрямо стартовата точка на транскрипция, докато свързването на TFIID комплекса обикновено защитава областта от -45 до -10.

ДНК-свързващият домен на TBP е разположен на С-терминалния край и е консервативен между видовете, докато вариабилният N-терминален край осъществява контакт с други белтъци. Високата степен на консервативност (~80%) между ДНК-свързващите домени на TBP при дрожди и човек прави TBP мярка за консервативността на механизма на инициация на транскрипцията.

TBP-свързани фактори

Съществуват също така механизми за разпознаване на кор промотора, които не включват каноничния TFIID комплекс, а използват фактори наречени **TRF** (TBP-related factors, англ. – фактори, свързани с TBP) - TRF1, TRF2, and TRF3, чиято функция е различна от тази на TBP.

TRF1 не съществува при човек и дрожди, но присъства в *Drosophila*, където се използва вместо TBP от РНК полимеразата III. **TRF2** (известен още като TLF, TLP, TRF и TRP) се използва от повечето еукариоти при транскрипцията на РНК полимеразата II на гени, чиито промотори не съдържат TATA-box, напр. гена, кодиращ хистон H1 при *Drosophila*. **TRF3** (известен още като TBP2 и TBPL2) присъства само в гръбначни и е най-тясно свързан с TBP – той се свързва с TATA-box и подпомага TATA-зависимата транскрипция. Установено е, че TRF3 е важен за ембрионалното развитие (Bártfai et al., 2004; Jallow et al., 2004). Особено интересна функция на TRF3 е открита при анализиране на диференциацията на миобласти до миофибрили. Установено е, че миобластите съдържат каноничен TBP-съдържащ TFIID комплекс, докато при крайната диференциация в миофибрили TFIID комплексът се замества от TRF3–TAF3 съдържащ комплекс. Тези открития разкриват, че крайно диференцираните клетки може да използват специализирани транскрипционни системи, посветени на специфични клетъчни функции.

- **TFIIA**

Свързването на TFIID към промотора се стабилизира от TFIIA. В допълнение, TFIIA може да опосредства отговора към активиращи или репресиращи молекули. Такава една роля може да бъде от особено значение, тъй като антагонизмът между позитивно и негативно действащи фактори при сглобяването на основния транскрипционен комплекс може да е от критично значение за регулиране скоростта на транскрипция.

- **TFIIB**

Веднъж щом TFIID се свързал към ДНК, към комплекса се присъединява TFIIB, който контактува с ДНК нагоре и надолу от TATA-box. Присъединяването на TFIIB към TFIID е важна стъпка при формирането на базалния транскрипционен комплекс, защото тя е предпоставка за присъединяването на РНК полимеразата II (към която вече се е свързал TFIIF). TFIIB съдържа N- и С-терминални домени, означавани като TFIIB_N и TFIIB_C, респ.

TFIIB_C и насоченост на промотора

TFIIB_C съдържа два прави повтора, които се нагъват в хомоложни субдомени, между които се образува цепнатина. В комплекса промотор/TBP/TFIIB С-крайният домен на TBP се вмъква в цепнатината между двата субдомена на TFIIB_C.

Освен че контактува с TBP, TFIIBc доменът, който е от типа HTH (helix-turn-helix, англ. – спирала-извивка-спирала), контактува с BREu нагоре от TATA-box. Взаимодействието TFIIB-BREu може да се извърши независимо от TBP и *играе важна роля за определяне насочеността на промотора*. Свързването на TFIIB с BREd надолу от TATA-box изглежда зависимо от свързването на TBP към TATA-box. Макар, че първоначално са идентифицирани в промотори с TATA-box, BREd и BREu могат да бъдат открити в много промотори, които не съдържат TATA-box, което предполага, че TFIIB най-вероятно стабилизира свързването на TBP към кор промотора като предоставя допълнителни контактни точки.

Секвенционно-специфичното взаимодействие между TFIIB и BREu е от съществено значение за определяне посоката на транскрипция при промотори с TATA-box. Естествено срещаните се TATA-box елементи са еднопосочни – те неизменно направляват транскрипцията в една посока. Въпреки това обаче, структурната основа на тази еднопосочност не е очевидна, тъй като TATA-box е частично симетричен (т.е. неговата секвенция и тази на обратната комплементарна верига са много подобни). В резултат, TBP е способен да се свърже към TATA-box секвенциите в една от двете ориентации. От голяма важност е способността на TFIIBc да образува скоба само около C-терминалния домен, но не и върху N-терминалния домен на TBP, поради малки структурни различия в двата домена. Свързването на TFIIBc към BREu, на свой ред налага асиметрия върху TBP/TATA-box комплекса като кара C-терминалния домен на TBP да свързва с онзи TATA box, който е съседен на BREu.

TFIIBn и избор на стартова точка на транскрипция

TFIIB осъществява също контакти и с РНК полимеразата II, образувайки мост между TBP и полимеразата. TFIIBn доменът съдържа цинк-свързващ мотив („цинкова панделка“), последван от бримка („В-пръст“), за които се счита, че осъществяват важни контакти с РНК полимеразата II. Мутации в „В-пръста“ могат да променят стартовата точка на транскрипция, което разкрива важността на TFIIBn за позиционирането на ДНК в активния център на ензима и по този начин опосредства избора на стартовата точка на транскрипция.

• TFIIF

TFIIF се свързва с РНК полимеразата II преди свързването и към ДНК. Този фактор е хетеродимер, който е изграден от два типа. По-голямата субединица (RAP74) има АТФ-зависима хеликазна активност, която може да участва в локалното разплитане на ДНК при инициация. По-малката субединица (RAP30) има известна хомология с региони от бактериалния σ -фактор и дестабилизира неспецифичните взаимодействия на ензима с ДНК. TFIIF спомага за свързването на РНК полимеразата II към сглобяващия се транскрипционен апарат, и заедно с TFIIB, е необходим за избора на стартовата точка на транскрипция.

• TFIIIE

TFIIIE е хетеротетрамер, който е изграден от два типа субединици (56kD и 34kD при човек). Той се присъединява към РНК полимеразата II след свързването и с ДНК и създава място за разполагане на следващия ТФ – TFIIH. След присъединяването на TFIIH, TFIIIE взема участие в регулирането на неговите каталитични активности. Освен това, TFIIIE е необходим на РНК полимеразата II за преминаване от инициация към елонгация.

- **TFIIH**

Хеликазна активност

ДНК хеликазите катализират преминаването на ДНК от двойно-верижно в едно-верижно състояние като използват енергия, получена при хидролиза на АТФ. РНК полимераза II е единствената РНК полимераза, която не може да „отвори“ промотора без помощта на такъв ензим. Такава необходимост не е установена при бактериалната полимераза, полимеразата на археи, нито пък при РНК полимераза I и РНК полимераза III. Хеликазната активност, необходима за отварянето на клас II промоторите се осигурява от TFIIH, който е последният фактор присъединяващ се към базалния транскрипционен апарат.

TFIIH е белтъчен комплекс с молекулна маса >500kD, съдържащ 10 различни генни продукта. Целият комплекс съдържа два субкомплекса, които могат да се дисоциират един от друг – хеликазен и фосфорилиращ субкомплекс. Хеликазният субкомплекс съдържа две ДНК хеликази, XPB и XPD, като каталитичната активност на XPB е от съществено значение за образуване на „отворен“ комплекс. Макар, че TFIIH има и други активности, хеликазна активност е единствената, която е от изключителна важност при инициация на транскрипцията.

Фосфорилираща активност

Фосфорилиращият субкомплекс включва протеин киназа от семейството на циклин-зависимите кинази, а именно Cdk7. Както подсказва името, циклин-зависимите кинази са активни само свързани с белтъци, наречени циклини. Циклиновият партньор на Cdk7 е циклин H, който също е част на фосфорилиращия субкомплекс.

Инициацията на транскрипцията е свързана с фосфорилиране на СТД на РНК полимераза II, и ролята на Cdk7 е да катализира фосфорилиращата реакция. СТД не е от съществено значение за сглобяване на базалния транскрипционен апарат, но въпреки това е жизненоважна. СТД може да бъде силно фосфорилирана, особено при Ser2 и Ser5 на хептапептидните повтори, и елонгиращата полимераза съдържа поне една фосфатна група на хептапептид. Въпреки това свободните молекули на РНК полимераза II са в нефосфорилирано състояние. Това вероятно е важно за регулиране на генната експресия, защото белтъчен комплекс, наречен **Медиатор**, който има важно участие в генната регулация, се свързва преференциално с нефосфорилираната форма.

Преходът от инициация към елонгация е свързан с мащабно фосфорилиране на СТД, най-вече на Ser5, от Cdk7. Докато тази реакция не е необходима за инициация на транскрипцията, тя най-вероятно превръща СТД в платформа за фактори, необходими за процесинга на мРНК. Напр. фосфорилирането на СТД при Ser5 увеличава афинитета на СТД към кепиращите ензими, които са отговорни за прибавяне на специални структури „кеп“ (7-methyl-Gppp) към 5'-края на клас II транскриптите. Този „кеп“ от една страна предпазва транскриптите от екзонуклеазна деградация в посока към 3', а от друга страна насочва транскрипта към рибозомата.

С напредване на транскрипцията, профилът на фосфорилиране на СТД се променя. СТД фосфатази отделят фосфатните групи от Ser5, докато Ser2-специфични кинази прибавят фосфатните групи към СТД. В резултат, докато СТД на елонгиращата полимераза е фосфорилиран най-вече при Ser5 при проксималните до промотора области на

транскрипционната единица, то при дисталните области той е фосфорилиран най-вече при Ser2. Точно както фосфорилираният при Ser5 СТД е платформа за сглобяване на кепиращия 5' апарат, така и фосфорилираният при Ser2 СТД служи като платформа за сглобяване на апарата, който катализира рязането и полиаденилирането на 3'-края на транскрипта. Така координираното фосфорилиране на СТД първо при Ser5 от Cdk7 субединицата на TFIIH, а по-късно от други кинази, помага за координирано протичане на процеса на биогенеза на РНК като гарантира, че само транскриптите на РНК полимеразата II ще бъдат обект на пост-транскрипционни модификации, които са белег единствено на протеин-кодиращите транскрипти.

Поправяща активност

В допълнение към неговата роля в транскрипцията, TFIIH има отделна роля в поправката на ДНК чрез изрязване на нуклеотид NER (nucleotide excision repair, англ. – поправка чрез изрязване на нуклеотиди). При този процес, обемисти повреждания на ДНК като пиримидинови димери, образувани под действие на UV радиация, се отстраняват чрез последователно действие на ендонуклеази, които правят едноверижни прекъсвания от двете страни на повредения участък в ДНК; ДНК хеликази, които освобождават повредения едно-верижен ДНК фрагмент; ДНК полимерази, които запълват региона на изрязания участък; и ДНК лигази, които затварят останалото едноверижно прекъсване в ДНК. XPB и XPD, които са част от TFIIH, предоставят необходимата за NER хеликазна активност, доказателство за което неспособността на белтъчни екстракти, получени от клетъчни линии, мутантни по една от двете хеликази, да опосредстват NER. Активността на тези мутанти може да бъде възстановена чрез прибавяне на хеликазния субкомплекс на TFIIH.

Откриването на единичен белтъчен комплекс, необходим както за транскрипцията, така и за поправката на ДНК, предоставя възможно обяснение на феномена известен като *свързана с транскрипцията поправка на ДНК*, при който активно транскрибираните региони на генома се поправят много по-ефикасно отколкото нетранскрибираните области. Мутации на гените, кодиращи хеликазните субединици на TFIIH, XPB и XPD, са причина за генетично-детерминираното заболяване ксеродерма пигментозум (*xeroderma pigmentosum*, XP), което е свързано с изключителна чувствителност към слънчевата светлина и повишен риск от рак на кожата. Някои мутации на XPB и XPD могат да бъдат причина за развитие на синдрома на Cockayne, който се свързва с множество нарушения в развитието като умствена изостаналост и др. Докато свето-чувствителността на XP пациентите най-вероятно е резултат от дефекти в системата за поправка на ДНК, то нарушенията в развитието при синдрома на Cockayne са по-скоро следствие на неефективна транскрипция.

TFIID и TFIIA могат да останат свързани към промотора след инициацията и така да зложат основите за формиране на нов инициращ комплекс, докато TFIIIB, TFIIIE и TFIIH се освобождават. TFIIIF се движи надолу по матрицата заедно с полимеразата и, както изглежда, е важен за ефективността на елонгация.

3. Базална транскрипция на РНК полимеразата I

РНК полимеразата I осъществява транскрипция на гените, кодиращи рРНК, и отговаря за приблизително 60% от синтеза на РНК в клетката като по този начин поддържа образуването на няколко милиона рибозоми.

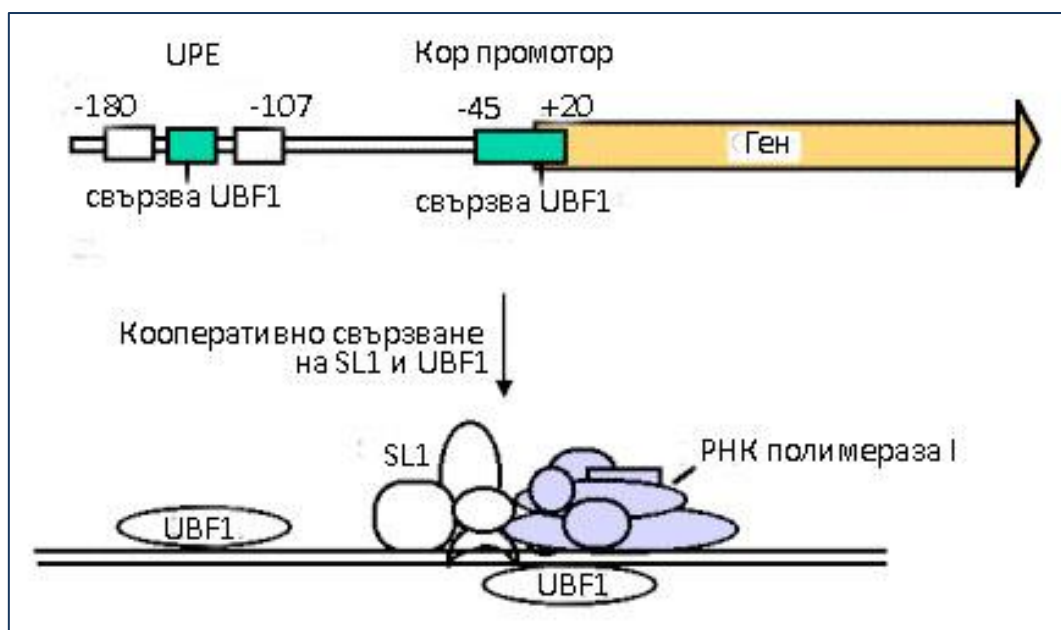
В еукариотната клетка има стотици копия на рРНК гените, които са организирани в клъстери като тандемни повтори и са основния компонент на ядрцето. При бозайници, първичният 47S рРНК транскрипт, синтезиран от РНК полимеразата I при бозайници, се процесира до зрели 18S, 5.8S и 28S рРНК, които заедно с 5S рРНК, транскрибиран от РНК полимеразата III, изграждат основния каталитичен и архитектурен компонент на 80S рибозомите. Съществува фин баланс между растежния клетъчен статус и натрупването на рРНК, който се контролира най-вече на ниво транскрипция на рДНК, която обхваща гените, които кодиращи рРНК.

В генома на дрожди са намерени около 150-200 рДНК повтори, докато при човек в диплоидната клетка има около 400 повтора. Въпреки това, само част от тези гени (~50%) се транскрибират в даден момент. Най-вероятната причина за това е алелно инактивиране на рРНК гени по време на ранното развитие. Активните и неактивните повтори се разграничават по различно състояние на хроматина, епигенетични белези, топологична организация и суб-ядрцева организация.

При бозайници, всеки рДНК повтор е с дължина около 43 kb и съдържа регулаторни елементи включващи промотори, повтарящи се енхансери и терминатори, разположени в интергенен регион дълъг около 30 kb, и един транскрибируем регион дълъг 13 kb, който съдържа областта, кодираща 47S. Въпреки, че са описани като рДНК „повтори“, нови изследвания показват, че множеството рРНК гени не са просто идентични копия на една и съща транскрипционна единица, всъщност няколко рДНК варианти съществуват и могат да бъдат диференциално експресирани и регулирани.

3.1. Промотор на РНК полимеразата I

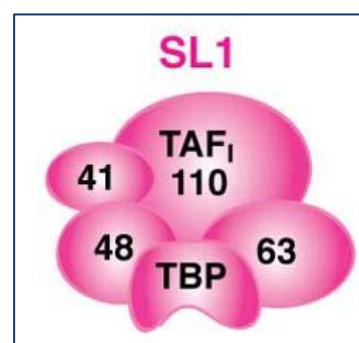
РНК полимеразата I използва само един тип промотор, който се състои от два отделни региона. **Кор промоторът** обгражда стартовата точка, простирайки се от -45 до +20, и е достатъчен за инициране на транскрипцията (Фигура 13). Обикновено той е богат на C-G (което не е типично за промоторните елементи), с изключение на къс консервативен елемент богат на A-T, разположен около стартовата точка. Ефективността на кор промотора се повишава значително от **UPE** (upstream promoter element, англ. – промоторен елемент нагоре, известен още като UCE, upstream control element, англ. – контролен елемент нагоре, който е богат на C-G и се разполага от -180 до -107. В допълнения на тези елементи са открити и дистални подобни на енхансери секвенции, които най-вероятно повишават стабилността на образуване на преинициаторен комплекс върху промотора на рРНК гените.



Фигура 13 . Промоторни елементи и главни ТФ на РНК полимераза I

3.2. Главни ТФ на РНК полимераза I

РНК полимераза I се нуждае от два ТФ за да разпознае промотора - **SL1** (Selectivity factor 1, англ. – фактор на селективност) и **UBF1** (Upstream binding factor, англ. – свързващ фактор нагоре). SL1 е комплекс на TBP и поне четири TBP-асоциирани фактори TAFI110, TAFI63, TAFI48 и TAFI41 (Фигура 14). SL1 е изключително важен за формирането на преинициаторния комплекс, тъй като той е необходим на РНК полимераза I да разпознае и да се свърже към кор промотора. Тъй като промоторът на РНК полимераза I не съдържа TATA-box, TBP не може да контактува с него. Вместо това TAFI110 и TAFI63 разпознават и свързват кор промотора. Основна отговорност на SL1 е правилното позициониране на РНК полимераза I спрямо стартовата точка на транскрипция. Освен това SL1 спомага за стабилизиране на взаимодействието между UBF и промотора на рДНК. SL1 е достатъчен сам по себе си да инициира транскрипцията на РНК полимераза I, но при много ниски базални нива.



Фигура 14 .
Мултисубединичен
комплекс на SL1

Мултисубединичният комплекс на РНК полимераза I съществува под формата на две отделни субпопулации, известни при бозайници като **Pol Iα** и **Pol Iβ**. При човека, Pol Iα която включва по-голямата част (>90%) от totalната РНК полимераза I в клетката, не е способна да инициира специфична транскрипция при промотора на рДНК дори и в присъствие на SL1, макар че тя притежава каталитична активност и може да инициира случайна транскрипция при наличие на свободни ДНК краища. Тези свойства са в съответствие с ролята на Pol Iα в транскрипционната елонгация. За разлика от нея, Pol Iβ, която представлява 10% от totalната РНК полимераза I в клетката, може да бъде инкорпорирана в преинициаторния комплекс и да инициира точна, промотор-специфична транскрипция. Това се дължи на асоциацията на Pol Iβ с фактора RRN3 (миши TIF1A). От друга страна, RRN3 се свързва и с SL1 като контактува с TAFI110, TAFI63

и TAFI41. Следователно RRN3 е мостът между Pol I β и промотор-свързващия фактор SL1, и по този начин улеснява включването на полимеразата в преинициаторния комплекс, формиращ се върху промотора на рДНК.

Активираната транскрипция, освен Pol I β и SL1, се нуждае и от UBF. Той е единичен полипептид, който се свързва като димер с C-G богатия елемент на UPE. ДНК-свързващият домен на UBF включва няколко субдомена HMG-box (high-mobility-group, англ. – силно-подвижна група). Тези субдомени показват значителна хомология с домените на ядрените нехистонови HMG белтъци HMG1 and HMG2, който са важни структурни елементи на хроматина и хромозомите. UBF е необходим за да се поддържа „отвореното състояние“ на хроматина като пречи на свързването на хистон H1 (. Ролята на UBF като архитектурен белтък се определя от неговата способност да индуцира образуването на енхансеозома, в която свързването на UBF към UPE води до огъване на ДНК на 360°. В резултат на това UPE и кор промоторния елемент се разполагат близо един до друг, с което се стимулира взаимодействието между SL1 и UBF. SL1 взаимодейства с UBF свързвайки се с неговия C-краен домен и това взаимодействие най-вероятно включва субединиците TAFI48 и TBP на SL1.

4. Базална транскрипция на РНК полимераза III

Транскрипционният апарат на РНК полимераза III отговаря за синтеза на малки по размер не-белтък кодиращи (нк) РНК, чиито транскрипционни единици често са представени в множество копия в еукариотния геном. Около 400 транскрипционни единици са прицелни за РНК полимераза III в дрожди, а при човек и мишка този брой достига 1000. Най-многобройните продукти на РНК полимераза III са различните видове тРНК, както и 5S рРНК, която обикновено е кодирана от няколко стотин идентични транскрипционни единици. В допълнение на изобилстващите нкРНК, участващи в транслацията, РНК полимераза III синтезира и малък хетерогенен набор от нкРНК, които играят роля в много РНК-зависими процеси в еукариотната клетка: процесинг на мРНК (напр. малката ядрена РНК U6, която участва в сплайсинга), тРНК (напр. РНК компонента на RNase P, която процесира 5'-края на пре-тРНК) и рРНК (напр. някои малки ядърцеви РНК), базална транслация и нейната регулация (напр. някои микроРНК и Alu РНК), ко-транслационна транслокация (напр. 7SL РНК, изграждаща скелета на сигнал-разпознаващата частица SRP), и РНК праймиране на репликацията на митохондриалната ДНК. РНК полимераза III взема участие и в синтеза на изобилстващите в геномите на бозайници транспозонни елементи SINE (Short interspersed repeated DNA elements, англ. - къси разпръснати повторени ДНК елементи), както и на вирус-кодирани РНК като аденовирусните VA-I и VA-II РНК.

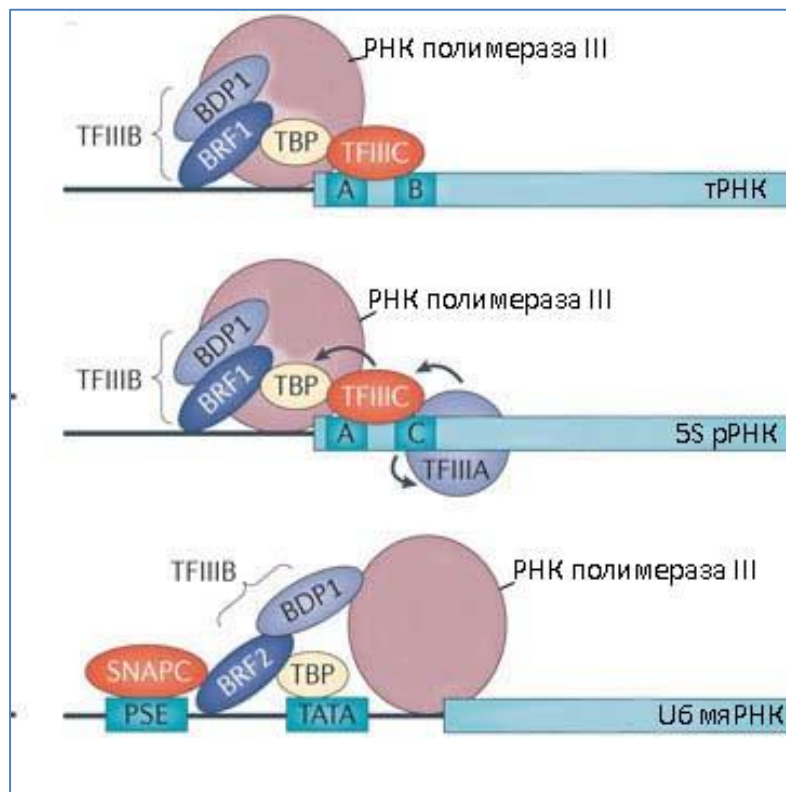
4.1. Промотори на РНК полимераза III

Промоторите на РНК полимераза III се класифицират в три основни типа, два от които се разполагат надолу от стартовата точка на транскрипция и обикновено не съдържат TATA-box а третият тип се разполагат нагоре от стартовата точка и съдържат TATA-box. **Промоторите от тип 1** се откриват в гените, кодиращи 5S рРНК, и притежават вътрешен контролен регион ICR (internal control region, англ.), който включват елементите A-box и C-box, както и разположения между тях интермедиатен елемент IE. **Промоторите от тип 2** се откриват при повечето гени на тРНК и аденовирусните VA-I и VA-II РНК включват елементите A-box и B-box като разстоянието между тях е специфично за различните гени. Типичен пример за **промотор от тип 3** е промотора на U6 мРНК (малка ядрена РНК), който при човек се състои от TATA-box и елемента PSE (proximal sequencing element, англ. – проксимален секвенционен елемент), които са необходими за базалната транскрипция *in vitro*. В допълнение, промоторът съдържа и елемента DSE (distal sequencing element, англ. – дистален секвенционен елемент), който засилва транскрипцията от базалния промотор. Промоторите от тип 3 са много подобни на промоторите на U1 and U2 при човек, които се разпознават от РНК полимераза II и включват PSE и DSE, но не и TATA-box. Освен това много промотори на РНК полимераза III са смес от елементи, съставляващи базовите 1, 2 и 3 типове промотори.

4.2. Главни ТФ на РНК полимераза III

Подобно на другите две полимерази, РНК полимераза III не може да разпознае сама своите прицелни промотори без помощта на допълнителни ТФ, които да свързват промоторните елементи и на свой ред да свързват други ТФ и полимеразата. При промоторите от тип 1, ICR се разпознава от **TFIIIA**. Образуването на комплекса TFIIIA-ДНК позволява свързването на **TFIIIC**, а след това и на **BRF1-TFIIIB**, който на свой ред прави възможно присъединяването на РНК полимераза III. При промоторите от тип 2,

TFIIIA не е необходим за образуване на преинициаторния комплекс. Вместо това, TFIIIC се свързва директно с A-box и B-box, след което последователно се присъединяват BRF1-TFIIIB и полимеразата. При промото-рите от тип 3, е необходим различен набор от ТФ. PSE се разпознава от **SNAPc** (snRNA activating protein complex, англ. – белтъчен комплекс, активиращ малки ядрени РНК), и TATA-box се разпознава от **BRF2-TFIIIB**, което позволява присъединяването на полимеразата. По този начин за всички промотори на РНК полимераза III сглобяването на базалния транскрипционен апарат води до свързването на BRF1- или BRF2-TFIIIB, което позволява



Фигура 15. Промоторни елементи и главни ТФ на РНК полимераза III

присъединяването на полимеразата (Фигура 15).

TFIIIA е първият еукариотен ТФ, който е бил клониран и неговото характеризиране е довело до концепцията за т. нар. белтъци цинкови пръсти. ТФ съдържа девет C_2H_2 цинкови пръсти. TFIIIC е голям комплекс съставен от шест типа субединици, от които само няколко са умерено консервативни от дрожди до човек. При дрожди има само един TFIIIB, който е съставен от три субединици. Едната е TBP, която е субединица на главните ТФ, специфични за всяка една от трите ядрени РНК полимерази. Втората субединица е BRF (TFIIIB-Related Factor, англ. – фактор, свързан с TFIIIB) и изпълнява функция при инициацията на РНК полимераза III, която е подобна на позициониращата функция, която изпълнява TFIIIB за РНК полимераза II. Третата субединица е BDP1 (*B double prime 1*, англ.). При човека има два типа BRF, BRF1 и BRF2, които са компоненти съответно на BRF1-TFIIIB (TBP, BRF1 и BDP1) и BRF2-TFIIIB (TBP, BRF2 и BDP1). Както беше споменато по-горе BRF1-TFIIIB се свързва към промотори тип 1 и тип 2, а BRF2-TFIIIB се свързва към промотори тип 3.

III. Транскрипционен контрол

Транскрипцията не може да се разглежда изхождайки от принципа всичко-или-нищо (базално ниво vs. активирано ниво). Всъщност, нивата на транскрипция могат да варират постепенно, от много ниско до много високо, чрез промяна например на състоянието на ДНК матрицата (от плътно пакетирана към хлабаво пакетирана ДНК), чрез броя на ТФ, които помагат на РНК полимеразата (включително и тези, които модифицират хромозомата, т. нар. хроматин-ремоделеращи фактори), както и от броя на регулаторните елементи, към които те свързват върху ДНК.

- При еукариоти, белтък-кодиращите гени, които е необходимо да се експресират координирано, не са тясно свързани в генома.
- Експресията на такива гени се координира от къси ДНК секвенции (**регулаторни елементи или cis-действащи елементи**), които са локализирани нагоре от стартовата точка на транскрипция и са основните компоненти на **промотора**.
- Други позитивно действащи ДНК елементи са локализирани много по-далеч от гена, който контролират, и са известни като енхансери и локус контролиращи региони.
- Всички тези секвенции действат чрез свързване на **регулаторни белтъци** (или специфични ТФ), които повлияват хроматиновата структура, или променят активността на базалния транскрипционен апарат при специфични вътрешни и външни въздействия.
- **Инициацията на транскрипцията** е основният етап, на който се осъществява транскрипционния контрол.
- За разлика от РНК полимеразата II, регулацията на транскрипцията от РНК

1. Регулаторни елементи на РНК полимеразата II

1.1. Къси cis-действащи елементи

1.1.1. Елементи на отговора (Response elements)

Елементите на отговора представляват къси консенсусни секвенции, които могат да бъдат открити в промотори или енхансери, и които отговарят на различни стимули на външната или вътрешната среда. Това е възможно, тъй като тези елементи свързват ТФ, които се синтезират или активират в отговор на индуциращ сигнал. Идентифицирани са различни елементи на отговора, които опосредстват клетъчния отговор на различни стимули (Таблица 1).

Таблица 1. Елементи на отговора, опосредстващи клетъчния отговор на различни стимули

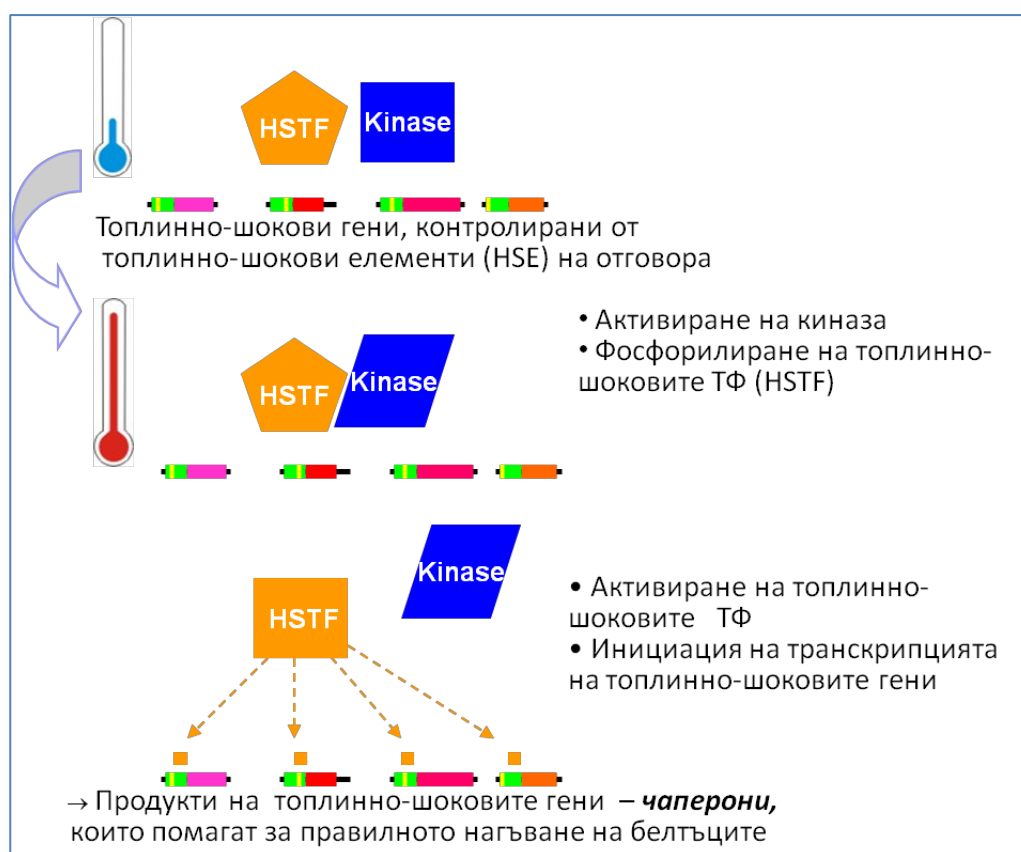
Консенсусна секвенция	Отговор на	Белтъчен фактор	Прицелен ген, съдържащ секвенцията
CTNGAATNTTCTAGA	Топлина	Топлинно-шокови ТФ	hsp70, hsp83, hsp27
T/G T/A CGTCA	Цикличен АМФ	CREB/ATF	Соматостатин, фибронектин, α -гонадотропин, c-fos, hsp70
TGAGTCAG	Форболов естер	AP1	Металотионеин IIA, α -антитрипсин, колагеназа
CC(A/T)6GG	Серумен растежен фактор	Серумния фактор на отговора (SRF)	c-fos, γ -актин
RGRACAN ₃ TGTYCY	Глюкокортикоиди	Глюкокортикоиден рецептор	Металотионеин IIA, триптофан оксигеназа, утероглобин, лизозим
RGGTCAN ₃ TGACCY	Естроген	Естрогенен рецептор	Овалбумин, коналбумин, вителогенин
RGGTCATGACCY	Тиреоиден хормон	Рецептор на тиреоидния хормон	Растежен хормон, тежка верига на миозина
TGCGCCCCGCC	Тежки метали	Mer-1	Металтионеинови гени,
AGTTTCNNTTTCNC/T	Интерферон- α	Stat-1 Stat-2	Олиго А синтетаза, гуанилат-свързващ протеин
TTNCNNNAA	Интерферон- γ	Stat-1	Гуанилат-свързващ протеин, Fc γ рецептор

R-пурин, Y-пиримидин

- **Един и същ елемент на отговора може да координира общия отговор на няколко гена към специфичен стимул**

Типичен пример за това е топлинно-шоковия елемент на отговора **HSE** (heat-shock element, англ. – топлинно-шоков елемент), който се открива в промоторите на гените, кодиращи топлинно-шоковите белтъци **HSP**. Индуцирането на топлинно-шоковите белтъци се съпровожда от свързването на белтък, известен като топлинно-шоков фактор **HSF** (heat-shock factor, англ. – топлинно-шоков фактор), който се свързва към HSE (Фигура 16). HSF е транскрипционен фактор, който всъщност не се синтезира в отговор на повишената температура (той се открива и в нестресирани клетки), а претърпява пост-транслационни модификации, в резултат на които преминава в активно състояние и свързва с HSE.

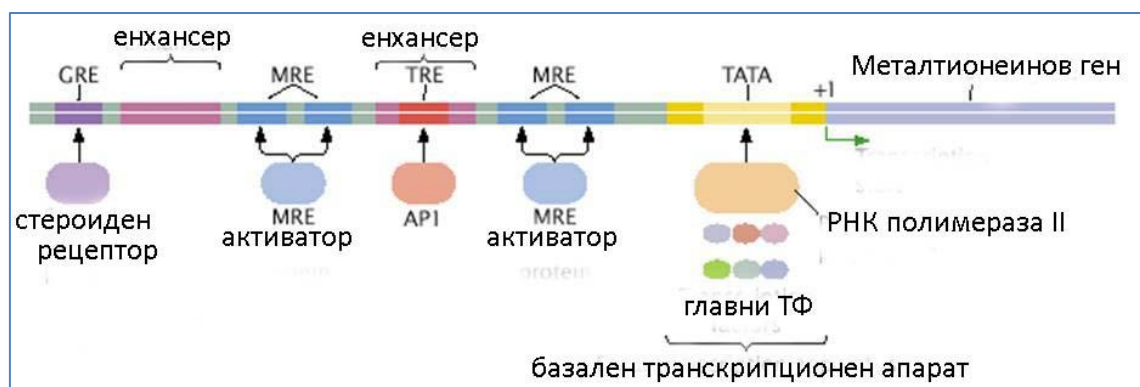
HSP фамилия (heat-shock protein, англ. – топлинно-шоков белтък) – голямо семейство белтъци, които се синтезират в отговор на повишена температура и други стресови въздействия; функционират като молекулни чаперони за да предпазват нативната структура на белтъците; високо-консервативни белтъци, присъстващи във всички живи организми.



Фигура 16. Регулация на топлинно-шовите гени от топлинно-шовия

- **Един ген може да притежава повече от един елементи на отговора, което позволява различни профили на регулация в отговор на различни стимули**

Металтионеиновият ген защитават клетката от токсични тежки метали, тъй като кодира белтък, който свързва тежките метали и спомага за отделянето им от клетката. Базалният транскрипционен апарат се сглобява около TATA-box, разположен непосредствено нагоре от транскрипционното начало на металтионеиновия ген, и иницира ниски нива на транскрипция (Фигура 17). Присъствието на тежки метали повишава значително нивата на транскрипция. Това е възможно поради присъствието на елементи на отговора, разположени нагоре от кор промотора. Напр. няколко копия на елемента на отговора на метали MRE (metal response element, англ.) се откриват нагоре в промотора, към които в отговор на тежки метали се свързва активиращ ТФ. В регулаторния промотор се откриват два енхансера; единият енхансер съдържа елемент на отговора на тиреоидните хормони TRE (th thyroid response element, англ.), който стимулира транскрипцията в присъствие на форболови естери. Трети тип елемент на отговора GRE (glucocorticoid response element, англ.) стимулира транскрипцията в отговор на глюкокортикоидните хормони.



Фигура 17. Елементи на отговора на металтионеиновия ген

- **Елементите на отговора съдържат консервативни хексамерни секвенции, които могат да бъдат организирани в различни двукомпонентни конфигурации, включващи обърнати и прави повтори**

Много елементи на отговора показват *диадна симетрия*, при която еднаква секвенция се открива в посока 5' към 3' във всяка една от веригите (*палиндром*). Така например елемента на отговора на естроген притежава секвенцията:



където двете половини на палиндрома, дълъг 10 бд, са разделени с три случайни бази образувайки *обърнат повтор*. Такава симетрия на местата, свързващите ТФ, показва, че факторите се свързват като димери, съставени от две белтъчни молекули.

Таблица 2. Връзка между елементите на отговора на различните хормони

Обърнати повтори	
Хомодимерни рецептори	
Глюкокортикоид	RGRACA N ₃ TGTYCY
Естроген	RGGTCA N ₃ TGACCY
Тиреоид	RRGTCA- - TGACCY
Прави повтори	
Хетеродимерни рецептори (с изключение на рецептора на 9-cis ретиноева киселина)	
9-cis ретиноева киселина	AGGTCAN ₁ AGGTCA
All-trans ретиноева киселина	AGGTCAN ₂ AGGTCA
	AGGTCAN ₅ AGGTCA
Витамин D3	AGGTCAN ₃ AGGTCA
Тиреоид	AGGTCAN ₄ AGGTCA
R-пурин, Y-пиримидин, N- всяка една база	

В допълнение, кор секвенцията GGTCA може да бъде организирана освен като обърнати повтори, и като два *прави повтори*. При тази организираност, разстоянието между двата повтора може да контролира кой хормон ще продуцира отговор. Така разстояние между повторите 4 бази образува алтернативен елемент на тиреоидния

хормон, 1 база детерминира отговор на 9-*cis* ретиноевата киселина, 2 или 5 бази на all-*trans* ретиноевата киселина, а разстояние от 3 бази на витамин D (Таблица 2).

Тази подобност между елементите на отговора на различни хормони е съпроводена от подобност на индивидуалните рецепторни белтъци, които образуват комплекс със съответния хормон и тогава се свързват към съответната ДНК секвенция. Всички тези рецептори са членове на голяма фамилия родствени ДНК-свързващи белтъци, които се различават по хормона, който ги индуцира, и по елемента на отговора, към които се свързват.

1.1.2. CCAAT- и GC-box

Тези елементи се откриват както в промотори с TATA-box, така и в такива без TATA-box, и се локализируют 70-150 бд нагоре от стартовата точка на транскрипция. Те се откриват в промоторните региони на голям брой гени с различни експресионни профили и свързват ТФ, които са активни във всички клетъчни типове, което подсказва, че тези елементи играят роля в стимулирането на конститутивната активност на промоторите. GC-box представлява консенсусна секвенция, богата на GC (GGGCGG), и може да бъде представен в множество копия в промоторните региони. Този елемент се открива в промоторите на редица гени като ранните гени на SV40, HIV-1, hsp70, металтионеиновите гени и др. GC-box свързва TF Sp1, който присъства във всички клетъчни типове. CCAAT-box се разпознава от различни класове ТФ като NF-Y (NF-Y, англ. – ядрен фактор Y), C/EBPs (CCAAT/enhancer binding proteins, англ. – CAAT/енхансер свързващи белтъци) и др. Самият елемент се открива в 20-30% от еукариотните промотори на конститутивно експресирани и индуцируеми гени. И двата елемента са жизнено важни за транскрипцията и тяхното елиминиране чрез мутации или делеции води до изключване на транскрипцията.

1.1.3. Други къси *cis*-действащи елементи

За разлика от елементите на отговора, които участват в отговора на специфични стимули на генната експресия, други къси секвенционни елементи или комбинация от елементи са въввлечени в контрола на тъканно-специфичните профили на генна експресия, присъщи на еукариотните гени. Така напр. октамерният мотив ATGCAAT, който се открива в промоторите на гените на имуноглобулиновите леки и тежки вериги, може да инициира специфична експресия в В клетките, когато е свързан към нерегулиран промотор. Подобно, къси ДНК секвенции, които свързват специфични за клетките на черния дроб ТФ, се откриват в албуминовия промотор при мишки, за който е известно, че участва в специфичната за клетките на черния дроб експресия на албуминовия ген.

1.1.4. Механизъм на действие на късите *cis*-действащи елементи

От изложеното по-горе става ясно, че късите *cis*-действащи елементи, разположени в близост до стартовата точка на транскрипция, играят важна роля в регулирането на генната експресия при еукариоти. Както се вижда от Таблица 1, тези секвенции опосредстват транскрипционното активиране като се свързват със специфични ТФ.

Метод на отпечатъка на ДНКаза I

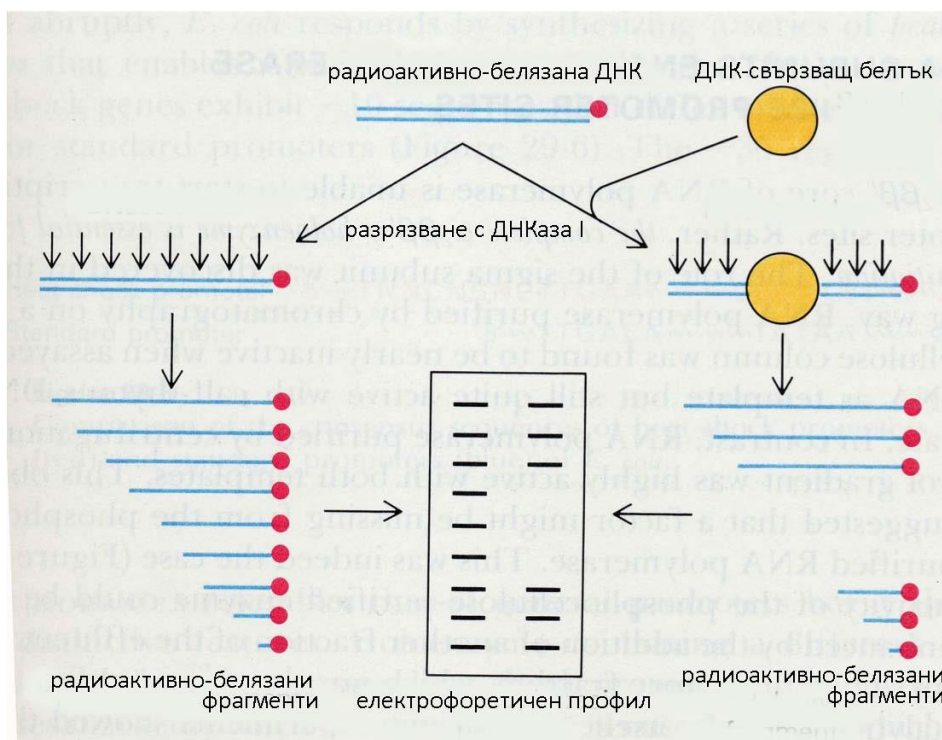
Принцип на метода – свързването на белтъчна молекула към ДНК предпазва мястото на свързване от разрязване с ДНКаза I.

Основни стъпки:

- Радиоактивно белязване на ДНК фрагмент или олигонуклеотид, съдържащ изследваната секвенция;
- Смесване на белязания фрагмент с клетъчен екстракт или ядрен екстракт, съдържащ белтък, свързващ се специфично с анализируания ДНК фрагмент ;
- Разрязване с ДНКаза I за образуване на серия от ДНК фрагменти, различаващи се по една база;
- Разделяне на ДНК фрагментите върху денатуриращ полиакриламиден гел, позволяващ разграничаване на една нуклеотидна база, и детекция на радиоактивните ивици чрез автордиография;
- Визуализация на региона на свързване на белтъка под формата на празнина или „отпечатък“ в стълбицата от ДНК фрагменти, който „отпечатък“ се дължи на белтъка, защитаващ този регион на ДНК от нарязване.

Методът позволява:

- визуализация на множество „отпечатъци“ върху дадена ДНК молекула;
- изяснява профила на белтъците, които се свързват към промотора или друг регулаторен елемент на изследвания ген;
- идентифицира хиперчувствителни на ДНКаза I участъци върху ДНК, в които хроматинът е загубил своята кондензирана структура и е в „отворено“ състояние.



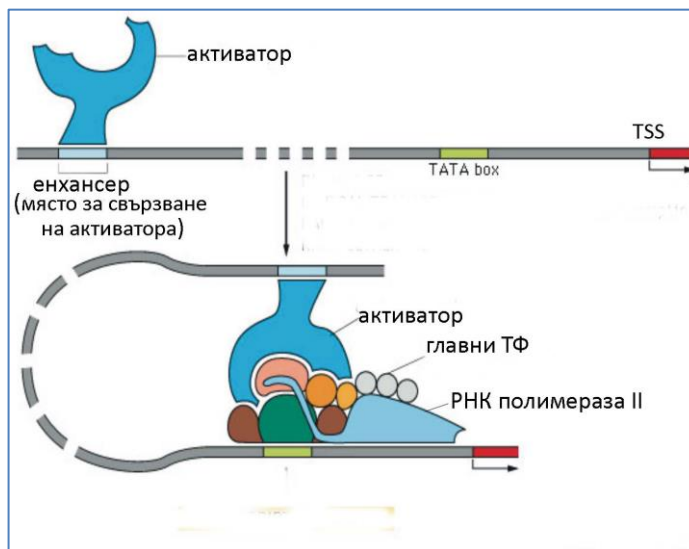
Свързването на ТФ може да повиши генната експресия по два основни начина. При единия начин, свързването на специфичния белтъчен фактор води до изместване на нуклеозоми и образуване на места хиперчувствителни на ДНКазата I, с което прави възможен достъп на други ТФ към дадения ген. При другия начин, свързаният ТФ може директно да активира транскрипцията – напр. топлинно-шоковите фактори. Тези фактори действат най-вероятно като се свързват на свой ред с белтъци, необходими за транскрипцията като ТВР или със самата РНК полимераза. Тези взаимодействия улесняват образуването на стабилен транскрипционен комплекс, който може да укрепва свързването на РНК полимеразата към ДНК или да променя нейната структура, в следствие на което да увеличава нейната активност. Описаните механизми на действия не са взаимно-изключващи се.

Така напр. глюкокортикоидният рецептор, чието свързване към специфична ДНК секвенция води до нуклеозомно изместване, съдържа също активиращ домен, способен на взаимодействия с други ТФ. Така, след свързване на рецептора, транскрипцията се стимулира чрез междофакторно взаимодействие между рецептора и друг свързан фактор. Подобно, топлинно-шоковият фактор може да индуцира ацетилиране на хистон H4, което води до отваряне на хроматиновата структура, което посочва, че той може да модулира хроматиновата структура, както и да стимулира базовия транскрипционен комплекс.

1.2. Дълги регулаторни елементи

1.2.1. Енхансери

Отличителен белег на еукариотната генна експресия е, че тя може да бъде контролирана от *cis*-действащи елементи, разположени на големи разстояние от стартовата точка на транскрипция. Тези елементи могат да бъдат локализирани нагоре, надолу или в транскрипционната единица и да функционират във всяка ориентация спрямо транскрипционното начало. Те действат чрез увеличаване на промоторната активност, макар че те самите нямат такава активност, откъдето произлиза и името им (enhance, англ. – усилвам). Някои енхансери са активни във всички клетъчни типове, докато други са тъканно-специфични и активират



Фигура 18 . Активаторът опосредства ефекта на енхансера върху транскрипцията

определени промотори в специфични клетъчни типове каквито са например енхансерите на имуноглобулиновите гени, които увеличават промоторната активност само в В клетките и водят до специфична експресия на имуноглобулините само в този клетъчен тип.

Подобно на промоторите, енхансерите имат модулна организация и съдържат множество места за ТФ, които взаимодействат помежду си за да опосредстват функционирането на енхансера (Фигура 18). В много случай елементите в енхансерите

са идентични с тези, които се откриват в регулаторния промотор. Така енхансерите на гените на имуноглобулиновите тежки вериги съдържат копия на октамерния елемент ATGCAAAT, който се открива в имуноглобулиновите промотори. Модулната организация на енхансерите е довела и до идеята за мултибелтъчен комплекс известен като *енхансеозома*, който се сглобява върху енхансера и води до транскрипционно активиране на гените, които контролира. Енхансеозомата може да включва белтъци като хистонов ацетилази и SWI/SNF комплекса, които съвместно водят до декондензация на хроматина и в последствие до свързване на активиращи ТФ.

- **Тъканно-специфична активност на енхансерите**

Скоро след откриването на енхансерите става ясно, че много гени, които се експресират в специфични тъкани, съдържат енхансери. Както беше споменато по-горе, гените, кодиращи леките и тежките вериги на молекулите на антителата, съдържат енхансер, който се разполага в пространството между регионите, кодиращи свързващия (J) и постоянните (C) региони на тези молекули. Когато този елемент се свърже към друг промотор, напр. β -глобиновия ген, той увеличава силно неговата активност, когато хибридният ген е интродуциран в В клетки. Обратно, енхансерът няма ефект върху промотора, когато хибридният ген е интродуциран в други клетъчни типове, напр. фибробласти, което показва, че енхансерът е тъканно специфичен.

Подобни тъканно-специфични енхансери са били намерени в гени, които се експресират специфично в клетките на черния дроб (фетопротеин, албумин, антитрипсин), ендокринните и екзокринните клетки на панкреаса (инсулин, еластаза, амилаза), хипофизата (пролактин, растежния хормон) и много други тъкани. Тъканно-специфичната активност на тези енхансерни елементи вероятно играе критична роля в тъканно-специфичния профил на експресия на съответния ген. Мутации в консервативните секвенции на тъканно-специфичните енхансери на гените, експресирани екзокринните клетки на панкреаса, напр. еластаза и хемотрипсин, премахват тъканно-специфичния профил на тези гени. В случая на инсулиновия ген, важноста на енхансерния елемент за тъканно-специфичната генна експресия е демонстрирана в експерименти, при които енхансерният елемент е бил свързан с гена, кодиращ големия Т антиген на еукариотния вирус SV40, чиято продукция се измерва лесно чрез използване на специфично антитяло. Полученият конструктор е бил интродуциран в оплодено яйце на мишка и експресията на големия Т антиген е била анализирана във всички тъкани на трансгенната мишка, която се развила след връщането на яйцето в яйцепровода. Експресията на големия Т антиген е била наблюдавана само панкреаса и никоя друга тъкан, и то специфично само в β клетките на панкреатичните островчета, които произвеждат инсулин. Енхансерът на инсулиновия ген, следователно, е способен да придаде специфичния профил на инсулиновата генна експресия на несвързан ген *in vivo*.

В много случаи, тъканно-специфичната експресия на даден ген се детерминира както от енхансерния елемент, така и от съседните на промотора секвенции. При специфичния за черния дроб пре-албуминов ген, генната активност се контролира както от промотора, който е активен само в клетките на черния дроб, така и разположения нагоре енхансерен елемент, който може да активира почти 10-кратно всеки промотор в клетките на черния дроб, но не и в друг клетъчен тип.

- **Механизми на действие на енхансерите**

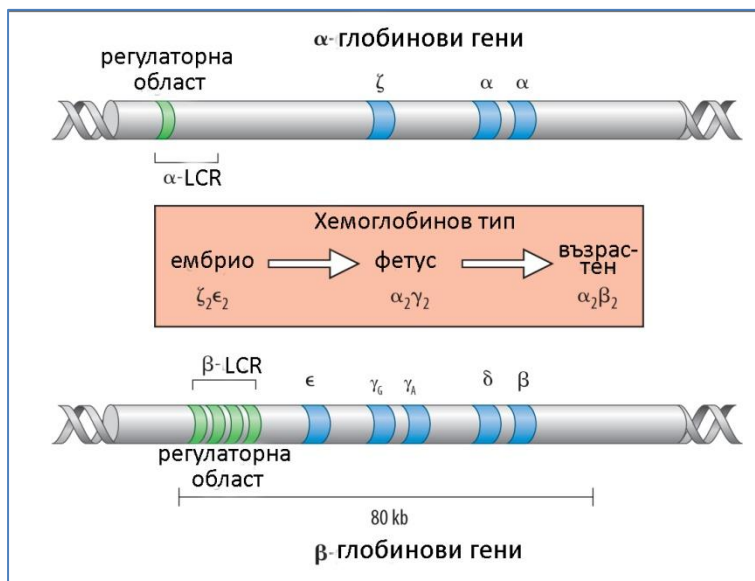
Енхансерите могат да активират генната експресия по механизми, подобни на тези, които промоторите използват, а именно чрез промяна на хроматиновата структура, което води до изместване на нуклеозоми, или чрез директно взаимодействие с белтъците на базалния транскрипционен апарат. Така енхансеозомният комплекс от регулаторни белтъци, който се сглобява върху енхансера, първо присъединява хистонови ацетилази, които на свой ред стимулират присъединяването на хроматин ремоделиращи комплекси SWI/SNF. Това води до отваряне на хроматина и позволява свързване на РНК полимеразата и асоциираните към нея белтъци, които ще направляват засилената транскрипция.

Очевидно в случаите на променена хроматинова структура, промените, причинени от свързването на регулаторен фактор към енхансера, могат да се предават на големи разстояния в двете посоки, което обяснява присъщата за действието на енхансерите независимост от разстоянието, позицията и ориентацията им спрямо транскрипционното начало. В съответствие с тази възможност, хиперчувствителни места за ДНКаза I са били картирани в множество енхансерни елементи като имуноглобулиновия енхансер. След свързването си към енхансерния елемент, регулаторният белтък осъществява контакт със свързания към кор промотора транскрипционен апарат или като се плъзва по ДНК, или чрез инициране на платформа от други белтъци, достигаща до транскрипционния апарат, или чрез огъване на ДНК между енхансерния и промоторните елементи като последният механизъм е най-вероятният.

1.2.2. Локус контролиращи региони (LCR)

Всеки ген се контролира от собствен проксимален промотор, много гени могат в допълнение да бъдат контролирани и от енхансери, а някои гени се нуждаят за своята точна експресия и от контролиращ елемент наречен **LCR** (Locus control region, англ. – локус контролиращ регион). Такива гени обикновено са организирани в домени от по няколко гена (*генен клъстер*), и експресията на всеки ген от домена се повлиява от LCR, който оказва ефект върху целия домен.

LCR обикновено регулират локуси, които показват сложни модели на генна експресия в индивидуално развитие или имат тъканна специфичност. Най-добре изученият пример за регулиран генен клъстер е този на β -глобиновите гени при бозайници. При тях, α - и β -глобиновите гени съществуват под формата на клъстери от подобни гени, които се експресират по различно време и в различни тъкани през индивидуалното развитие. *β -глобиновият клъстер* включва 5 гена – ϵ , γ^G , γ^A , δ и β (Фигура 19), които следват един след друг в съответствие с реда, по който се експресират през индивидуално-то развитие. ϵ -Генът се експресира предоминантно в примитивните еритроидни клетки на ембриона, транскриптите на двата γ -гена са доминиращи в еритропоетичните клетки на черния дроб на фетуса, където се активират δ - и β -гените, но се експресират основно перинатално. β -глобиновият клъстер е под контрола на LCR с дължина 20кб, който е разположен на 5' края на хромозомния домен и е съставен от пет хиперчувствителни на ДНКаза I места. LCR е абсолютно необходим за експресията на всеки един ген в локуса. Освен това всеки ген е под контрол и на неговите собствени контролиращи региони като промотор и енхансер.



Фигура 19. Структура на α - и β -глобиновите клъстери

Анализът на ДНК секвенции на различни LCR показва, че те са богати на мотиви, които се откриват също и в промотори и енхансери, и към които могат да се свързват ТФ. Следователно, подобно на промоторите и енхансерите, съществува принципна възможност LCR да може да функционира или чрез директно взаимодействие с транскрипционния апарат или като повлиява хроматиновата структура. На практика се оказва, че втората възможност е вярната – LCR повлияват хроматиновата структура на съседните гени

така, че те да бъдат транскрибирани..

LCR функционират като регулират хроматиновата структура на целия домен на β -глобиновия клъстер като го поддържат в „отворено състояние“ в червените кръвни клетки и техните предшественици. Хиперчувствителните на ДНКаза I места в LCR най-вероятно са свободни от нуклеозоми и се намират в суперспирализирана форма, която е обект на торзионен стрес и позволява разпространяването на променено хроматиново състояние към съседните региони, където се формират структури, характерни за активна и потенциално активна ДНК.

LCR контролира структурата на специфичен регион на ДНК, осигурявайки подходящия модел на генна експресия. Веднъж щом хроматиновият регион е отворен, промоторните и енхансерните елементи на всеки отделен ген в този регион биха могли да насочват специфичния за гена експресионен профил. При β -глобиновия клъстер, хиперчувствителните на ДНКаза I места в LCR се появяват рано в еритроидното развитие, още преди експресията на най-ранните гени в клъстера. Последователно ϵ -глобиновият ген, γ -глобиновите и β -глобиновият ген се експресират съответно в еритроидните клетки на ембриона, фетуса и възрастния индивид, като експресията им е предшествана от появата на хиперчувствителни на ДНКаза I места в промоторните/енхансерните региони на отделния ген. По този начин LCR изграждат контролен елемент, който направлява хроматиновата структура на голям регион от ДНК, и позволява последователно активиране на индивидуалните гени в клъстера чрез образуване на хиперчувствителни на ДНКаза I места и свързване на ТФ към техните индивидуални промотор-ни и енхансерни елементи.

1.2.3. Сайлънсери

От изложеното по-горе, става ясно, че енхансерите имат изцяло позитивен ефект върху транскрипцията. В тъкан, съдържаща активен енхансер-свързващ белтък, енхансерът ще активира промотора, докато в други тъкани, където този белтък липсва или е

неактивен, енхансерът няма да има ефект. За разлика от енхансерите, някои секвенции изглежда, че имат негативен ефект като инхибират гените, които ги съдържат. След първоначалното идентифициране на такъв елемент в клетъчния онкоген c-myc, подобни елементи, наречени **сайлънсери** (silence, англ. – тишина), са били намерени в гените, кодиращи колаген тип II, растежния хормон и глутатион трансфераза P. Подобно на енхансерите, сайлънсерите могат да действат от разстояние върху промотори, независимо от ориентацията си, но вместо стимулиращ, имат инхибиращ ефект върху нивото на генна експресия. Подобно на енхансерите, тези елементи действат или върху хроматиновата структура чрез свързването на фактори, които направляват плътното пакетиране на съседните ДНК участъци, или чрез свързването на белтъци, които директно взаимодействат с РНК полимеразата II и свързаните с нея ТФ.

За някои сайлънсери е установено, че предоставят места за прикрепване на ДНК към ядрения матрикс. Следователно, този тип регулаторни елементи могат да стимулират по-нататък кондензирането на хроматиновия соленоид и образуването на бримка, с която ДНК да се прикрепя към ядрения матрикс. Това може да бъде постигнато чрез свързването на специфични белтъци като поликомб РсG факторите, които направляват плътното пакетиране на съседната ДНК като индуцират убикитинизацията на хистон H2A и метилирането на лизин 27 върху хистон H3.

Освен, че действат на ниво хроматинова структура, сайлънсерите могат да действат чрез свързване на негативно-действащи ТФ, които взаимодействат с базалния транскрип-ционен комплекс. Така сайлънсерът на гена, кодиращ лизозим, както изглежда действа, поне частично, чрез свързване на рецептора на тиреоидния хормон, който при липса на хормона има инхибиращ ефект върху генната активност. *Следователно, енхансерите могат да действат на разстояние за стимулиране на транскрипцията чрез „отваряне“ на хроматиновата структура или чрез директно свързване на активиращи ТФ, а сайлънсерите могат да инхибират транскрипцията чрез „затваряне“ на хроматиновата структура или чрез директно свързване на репресиращи ТФ.*

1.2.4. Инсулатори (ICR)

Съществуването на елементи като енхансери и сайлънсери, които могат да действат на големи разстояния за да променят генната експресия, поражда въпроса как действието на тези елементи се ограничава само върху определен ген или гени. Този проблем се разрешава чрез съществуването на **инсулаторни елементи** (известни още като гранични или бариерни елементи), които както името им подсказва блокират разпространяването на активността на енхансерите и сайлънсерите, така че тя да бъде ограничена само до определен ген или гени, *дефинирайки по този начин транскрипционно независими домени.*

Инсулаторните елементи могат да изпълняват блокираща функция и/или бариерна функция. Блокиращата функция се наблюдава, когато инсулаторният елемент е разположен между енхансер/LCR и промотор, и пречи на енхансера да активира промотора (Фигура 20). Такива инсулаторни елементи са открити напр. в регион на хромозома 11 на човека, който съдържа няколко близко-родствени гени, кодиращи белтъци от β -глобинов тип. Този регион съдържа LCR, който стимулира експресията на β -глобиновите гени в еритроидните клетки. Регионът, който съдържа LCR и β -глобиновите гени, е фланкиран от два инсулаторни елемента, които не позволяват на LCR да активира

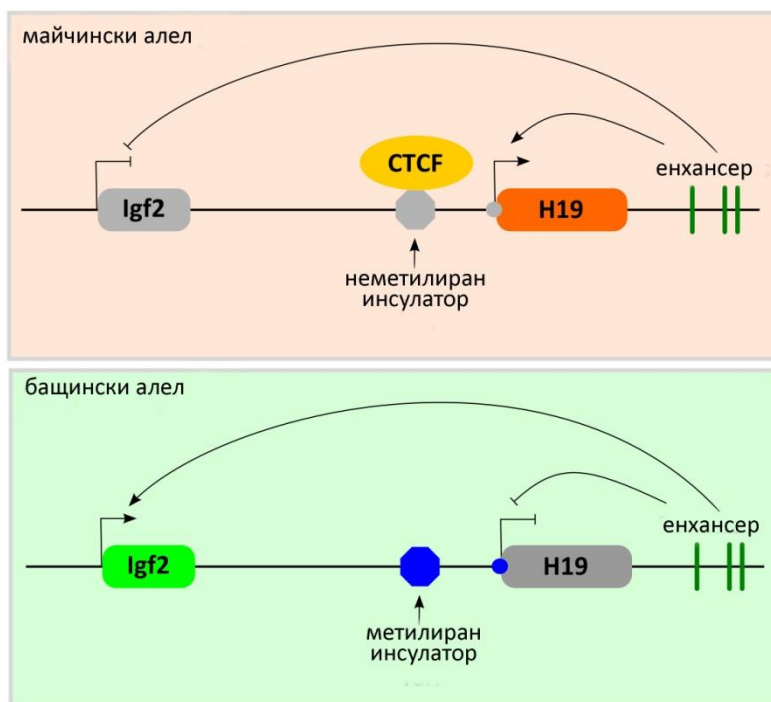
разположените в съседство гени, кодиращи фолатния рецептор и рецептора на аромати. Барьерната функция се проявява, когато инсулаторът е разположен между активен ген и хетерохроматинов регион, и се явява бариера, която защитава гена от инактивиращия ефект, който се разпространява от хетерохроматина. Така инсулаторите играят важна роля за ограничаване на разпространението на хетерохроматина от центромерните и теломерните участъци към съседни еухроматинови региони.

Активността на инсулаторните елементи може да бъде контролирана, така че те да бъдат активни в една ситуация и неактивни в друга. Така например

импринтинг-контролиращият регион не позволява на енхансера на гена H19 да активира гена Igf2 върху майчината хромозома, но няма този ефект върху бащината хромозома, където Igf2 е активен (Фигура 20).

Импринтинг-контролиращият регион действа като инсулаторен елемент блокирайки активността на енхансера да упражнява стимулиращ ефект върху съседни региони върху майчината хромозома. Елементът е активен само, когато към него се свързва регулаторния белтък CTCF. Очевидно,

инсулаторният елемент постига своя ефект чрез свързване на специфични белтъци. Ако енхансер, или сайлънсер, действа чрез индуциране на промени в хроматиновата структура, тогава инсулатор и свързаният към него белтък могат да действат за ограничаване на разпространението на промените в хроматиновата структура след мястото на свързване на белтъчния фактор.



Фигура 20 . Контрол на транскрипцията на Igf2 от енхансерен и инсулаторен елемент

2. Специфични ТФ на РНК полимераза II

Кор промоторът функционира с много ниска ефективност. **Специфични ТФ** (или само ТФ) са необходими за осигуряване на точното ниво на експресия във времето и пространството. Специфичните ТФ имат ефект само върху определени гени в отговор на специфични стимули и сигнали.

ТФ споделят **общи свойства**:

- ТФ разпознават и се свързват със специфични регулаторни елементи (напр. елементи на отговора) в промоторни области и енхансери ;
- ТФ имат **модулна структура**, в която различните домени на белтъка отговарят за различни функции. Основните домени (или модули) са ДНК-свързващ домен и активиращ (или по-рядко репресиращ) домен;
- Идентифицирани са различни ДНК-свързващи домени, които се използват за класифициране на ТФ във **фамии** със сходни ДНК-свързващи домени;
- ТФ повлияят транскрипцията **позитивно** (активатори) или **негативно** (репресори);
- **Активаторите** активират транскрипцията чрез стимулиране на сглобяването или активността на базалния транскрипционен апарат като осъществяват с него пряк контакт или косвен контакт посредством ко-активатори;
- **Репресорите** могат да потискат транскрипцията косвено като блокират стимулиращия ефект на активатор или като взаимодействат с базалния транскрипционен апарат като осъществяват с него пряк контакт или косвен контакт посредством ко-репресори;
- Обикновено ТФ **отговаря на стимул**, който сигнализира, че един или няколко гена трябва да бъдат включени;
- В случай, че ТФ модулира индуцируема, тъканно-специфична или свързана с развитието регулация на генната експресия, трябва да съществуват начини за **регулиране на синтеза или активността на самия ТФ**, така че той да е активен само в определена ситуация.

2.1. Регулаторен транскрипционен апарат

Регулаторният транскрипционен апарат на всички организми включва **активатори**, които стимулират транскрипцията, осъществявана от базалния апарат, и **репресори**, които потискат транскрипционната активност на базалния апарат. Активаторите и репресорите действат по ген-специфичен начин и така са способни да определят кои гени да бъдат транскрипционно активни във всеки отделен клетъчен тип при определени условия на вътрешната и външната за организма среда. Тъй като споменатия ген-специфичен начин зависи от способността на активаторите и репресорите да разпознават и свързват специфични ДНК секвенции в гените, които контролират, *активаторите и репресорите общо се означават като специфични ТФ*. Веднъж свързани към ДНК, тези фактори повлияват скоростта на транскрипция по

различни механизми. При бактерии и археи, и понякога при еукариоти, те осъществяват пряк контакт с базалния апарат и или пречат на процеса на инициация в случая на репресорите, или стимулират този процес в случая на активаторите.

Поради тяхната по-голяма сложност, еукариотите се нуждаят от по-фини модели на генна регулация в сравнение с другите организми. За да бъде удовлетворена тази нужда, еукариотите притежават допълнителна група ТФ, наречени **ко-активатори** и **ко-репресори**, които не се откриват в бактерии (и вероятно в археи също). За разлика от секвенционно-специфични ТФ, които се свързват с ДНК директно, тези корегулатори не контактуват с ДНК, а се свързват към секвенционно-специфични ТФ посредством белтък-белтъчни взаимодействия. Ко-активаторите и ко-репресорите повлияват скоростта на транскрипция по много начини. Подобно на активаторите и репресорите, те понякога осъществяват директен контакт с базалния апарат за да стимулират или потискат транскрипцията. Освен това, ко-регулаторите контролират транскрипцията като катализират ковалентни или нековалентни промени в структурата на хроматина. Хроматинът е изграден от ДНК и белтъци, като основния белтъчен компонент са хистоните. Последните са обект на множество пост-транслационни модификации като ацетилиране, метилиране, и фосфорилиране. Ко-активаторите и ко-репресорите често подават сигнал към другата част на транскрипционния апарат като катализират прибавянето или отстраняването на тези модификации.

Модификациите на хроматиновата структура, които определят транскрипционното състояние на самите гени, понякога са достатъчно стабилни за да преживеят митозата. Тъй като те могат да се предават от едно клетъчно поколение на друго, наподобявайки по-скоро алтернативни генни алели отколкото алтернативни форми на хроматина, тези състояния се означават като **епигенетични състояния**. Това прави епигенетичните състояния почти толкова важни колкото генотипа и околната среда при формиране на фенотипа.

Еукариотните ТФ имат модулна организация и са изградени от няколко функционални домена: **ДНК-свързващ домен**, който разпознава и се свързва със специфични регулаторни елементи на промотора (и енхансера); **активиращ домен**, който се свързва с други белтъци за да стимулира транскрипцията; **подвижен домен**, който свързва ДНК-свързващия домен и активиращия домен; **домен за ядрена локализация**, който е необходим за придвижването на ТФ от цитоплазмата в ядрото. ТФ, които образуват димери (хомо- или хетеро-димери), съдържат и **димеризиращ домен**, чрез който се свързват мономерните субединици.

2.2. ДНК-свързващи мотиви и фамилии на ТФ

ДНК-свързващите домени са изградени от структурни мотиви, които съдържат една или няколко α -спирали, които контактуват с голямата бразда на ДНК в областта на съотв. регулаторен елемент. На базата на ДНК-свързващи домени, ТФ са класифицирани във **фамилии**, в които членовете притежават сходни ДНК-свързващи домени.

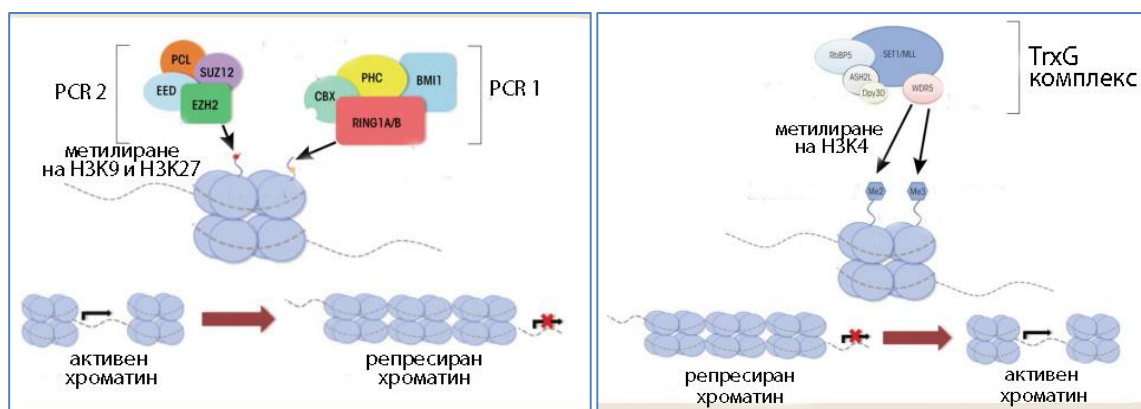
2.2.1. ТФ с ДНК-свързващ мотив Спирала-извивка-спирала

- **Хомеобокс**

Малкият размер и краткото време на размножаване на винената мушица *Drosophila melanogaster* са причина, този вид да бъде от генетична гледна точка един от най-добре характеризираните видове, и това е било свързано с идентифицирането и изучаването на голям брой мутации. Последните включват мутации, които засягат развитието на винената мушица и водят например до образуване на допълнителни крачка на местата на антените. Изучаването на този тип мутации е довело до идентифицирането на група гени, които имат незаменима роля за развитието на мушицата, особено за определяне на плана на тялото, и са известни като **хемеотични гени**.

Тези гени кодират регулаторни белтъци, които действат в точно определен етап от развитието като активират или потискат активността на други гени, необходими за образуване на определени структури. Регулаторните белтъци са способни да се свързват с ДНК, разпознавайки определена секвенция, и да индуцират активирането на транскрипцията на гените, които съдържат тези места на свързване. Така в случая на хемеотичния ген *fushi tarazu* (*Ftz*), чиито мутации водят до намален наполовина брой сегменти на тялото, е установено, че белтъкът се свързва специфично към секвенцията TCAATTAAATGA. Продуктът на друг хемеотичен ген – *engrailed*, се свързва към същата секвенция, но не води да активиране на транскрипцията, а по-скоро предотвратява активирането от *Ftz*. В следствие на това, когато в клетката се експресира само *Ftz* се активират определени гени, когато обаче се експресират и двата гена *Ftz* и *engrailed* този ефект не се наблюдава. По този начин, *взаимодействието между продуктите на хемеотичните гени, които се експресират в определени клетки, би могло да контролира експресията на специфични прицелни гени*.

Нещо повече, има доказателства, че *хемеотичните гени може би са необходими не само за формирането на специфичен клетъчен тип, но също така и за дълготрайния процес на поддържане на определен път на диференциация*. Вероятен молекулен механизъм, който прави това възможно, е установен при *Ubx* белтъка, който се свързва към своя собствен промотор и повишава своята собствена продукция. Така че след като веднъж образуването на белтъка е било индуцирано по време на формиране на клетъчния тип, то това може да продължи безкрайно и така да поддържа диференцираното състояние. Освен този механизъм, други процеси поддържат хроматиновата структура на хемеотичните гени в активно състояние, така че транскрипцията да продължи щом веднъж е била активирана. Това се постига с участието на белтъци от **trithorax групата (trxG)**. Тези фактори се свързват към активните хемеотични гени и ги поддържат в „отворена“ хроматинова структура, която обезпечава осъществяването на транскрипцията. Обратно, белтъци от **polycomb групата (PcG)** се свързват към неактивните хемеотични гени и ги поддържат в „затворена“ хроматинова структура, която не позволява тяхното активиране (Фигура 21). Както може да се очаква, мутации в гените, кодиращи белтъци от групите на *trxG* и *PcG*, водят до големи аномалии във винената мушица.



Фигура 21. *Polycomb* групата (PcG) и *trithorax* групата (TrxG) - модулатори на хроматиновата структура с антагонистично действие.

PcG фамилията участва в разпространяване на неактивното състояние на хроматина: PRC2 имат хистон-модифицираща активност и метилират H3K27 и H3K9 (репресивни белези); PRC1 разпознават горните модификации и индуцират ремоделиране на хроматина до неактивно състояние

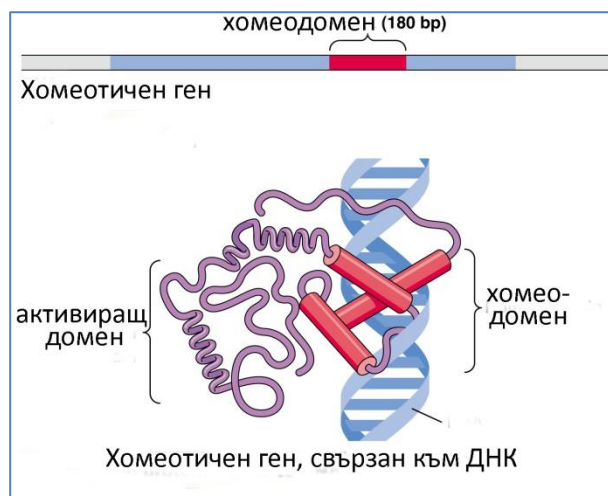
TrxG фамилията участва в разпространяване на активното състояние на хроматина: TRX има хистон-модифицираща активност и метилира H3K4 (активен белег); SWI/SNF разпознава горните модификации и индуцира ремоделиране на хроматина до активно състояние (с модификации по Veazey et al 2015)

Фактът, че продуктите на хомеотичните гени могат да регулират както собствените си гени, така и други гени, чрез специфично свързване към ДНК, е довел до широко изследване на тяхната структура за да се идентифицира региона, който опосредства това секвенционно-зависимо свързване на ДНК. Когато гените, кодиращи хомеотични фактори, са били клонирани, става ясно, че всеки един от тях съдържа къса ДНК секвенция от около 180 бд, кодираща 60 АК, която е много подобна между тях, но се фланкира от двете страни от секвенции, които се различават много силно между отделните гени. Присъствието на тази секвенция, която бива наречена **хемеобокс** или **хемеодомен**, във всички тези гени подсказва, че тя има незаменима роля за опосредстване на тяхната регулаторна функция (Фигура 22). По-късно хемеодоменът е бил използван като проба за изолиране на други все още нехарактеризирани регулаторни гени в *Drosophila*.

	1	N-terminal arm						10	Helix 1						20							
En	Glu	Lys	Arg	Pro	Arg	Thr	Ala	Phe	Ser	Ser	Glu	Gln	Leu	Ala	Arg	Leu	Lys	Arg	Glu	Phe	Asn	Glu
Antp	Arg	Lys	Arg	Gly	Arg	Gln	Thr	Tyr	Thr	Arg	Tyr	Gln	Thr	Leu	Glu	Leu	Glu	Lys	Glu	Phe	His	Phe
Oct2	Arg	Arg	Lys	Lys	Arg	Thr	Ser	Ile	Glu	Thr	Asn	Val	Arg	Phe	Ala	Leu	Glu	Lys	Ser	Phe	Leu	Ala
En		Asn	Arg	Tyr	Leu	Thr		30	Helix 2						40							
Antp		Asn	Arg	Tyr	Leu	Thr		Arg	Arg	Arg	Arg	Glu	Leu	Ser	Ser	Glu	Leu		Gly	Leu		
Oct2		Asn	Arg	Tyr	Leu	Thr		Arg	Arg	Arg	Arg	Ile	Glu	Ile	Ala	His	Ala	Leu		Cys	Leu	
		Asn	Glu	Lys	Pro	Thr		Ser	Glu	Glu	Ile	Leu	Leu	Ile	Ala	Glu	Gln	Leu		His	Met	
En	41	Asn	Glu	Ala	Gln	Ile	Lys	Ile	Trp	Phe	Gln	Asn	Lys	Arg	Ala	Lys	Ile	Lys	Lys	Ser	Asn	
Antp		Thr	Glu	Arg	Gln	Ile	Lys	Ile	Trp	Phe	Gln	Asn	Arg	Arg	Met	Lys	Trp	Lys	Lys	Glu	Asn	
Oct2		Glu	Lys	Glu	Val	Ile	Arg	Val	Trp	Phe	Cys	Asn	Arg	Arg	Gln	Lys	Glu	Lys	Arg	Ile	Asn	

Фигура 22. Аминокиселинна секвенция на хемеодомена при хомеотичните гени. ТФ, които участват в контрола на идентичността на сегментите на тялото при винената мушица - Antennapedia (Antp) и Engrailed (En) и в индивидуалното развитие при бозайници - Oct2. (по Lewin et al. 2009)

Структурните анализи на хомеотичните гени разкриват, че те **съдържат мотив от типа спирала-извивка-спирала НТН**, който е изграден от три α -спирали и е силно консервативен между различните хомеобокс-съдържащи белтъци. Мотивът НТН включва три α -спирали, от които α -спирала 2 и α -спирала 3 разпознават ДНК-свързващия мотив ТААТ в прицелните гени, а α -спирала 1 стабилизира мотива (Фигура 23). Кристалорафски анализи на хомеодомен свързан към ДНК показват, че мотивът НТН наистина контактува с ДНК като заема голямата бразда, където може да прави специфични контакти с базите на ДНК. Ролята на мотива в разпознаването на специфична секвенция в ДНК е демонстрирана директно. Мутации, които променят лизина в позиция 9 на разпознаващата спирала на Bicoid до глутамин, който се открива на еквивалентна позиция в Antennapedia, има за резултат свързване на Bicoid към свързващи места в ДНК, специфични за Antennapedia. Следователно, не само наличието на мотив НТН е важен за свързването на ДНК, а също и точната секвенция на този мотив в различните хомеодомени.



Фигура 23. Свързване на хомеотичен ген към прицелен ген чрез НТН мотив

Очевидната значимост на хомеодомена в *Drosophila* и дрожди е стимулирала учените в търсенето на белтъци, съдържащи този мотив, в други организми. Наистина, хомеобокс-съдържащи белтъци, които се експресират в специфични клетки на ранния ембрион и играят ключова регулаторна роля, са били идентифицирани в голям брой организми, включително и бозайници. Следователно тези белтъци са жизненоважни за процесите, регулиращи развитието на различните организми.

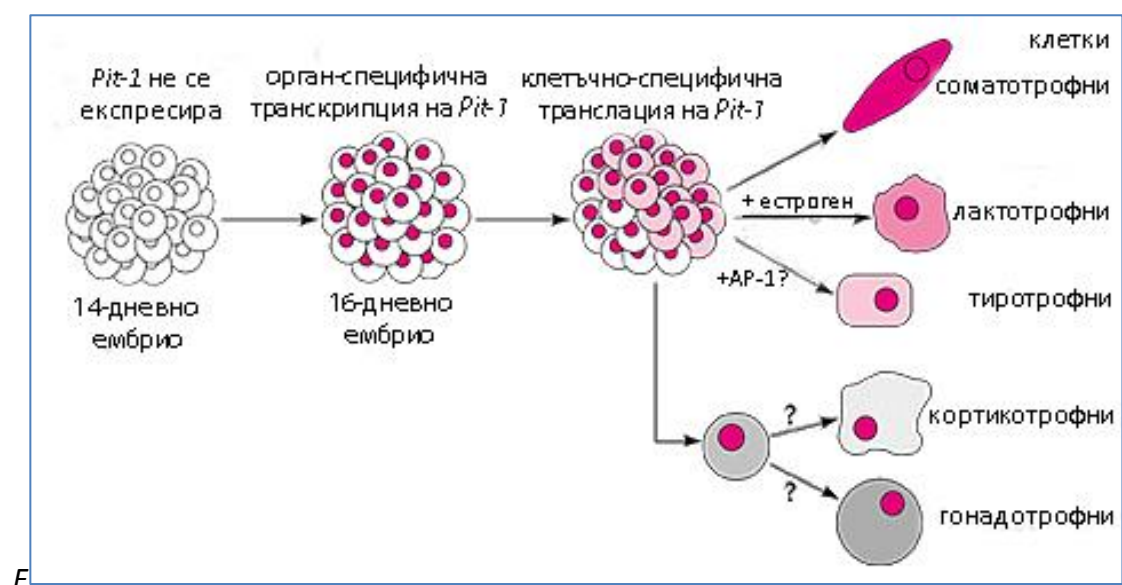
• POU домен

Подобно на хомеодоменните белтъци, друг клас регулаторни белтъци съдържат хомеодомен, който обаче е част от част от друг по-голям консервативен регион, (около 150-160 АК), известен като **POU домен**. Той е двукомпонентен домен, който включва POU-хомеодомен и POU-специфичен субдомен, като и двата са ДНК-свързващи субдомени от типа НТН. POU доменът е открит в редица ТФ.

Така регулаторните белтъци при бозайници **Oct-1** и **Oct-2** се свързват към октамерна секвенция с консенсус ATGCAAATNA в промоторите на гените, които кодират H2B и различни имуноглобулини, за да опосредстват тяхното транскрипционно активиране. Клонирането на съответните гени разкрива секвенция с дължина 150-160 АК, която по-късно е открита и в регулаторния белтък **Pit-1** в бозайници, който регулира генната експресия в хипофизата като се свързва със секвенция много сходна на октамерната секвенция. Подобна секвенция е идентифицирана и в белтъка **unc-86** при нематоди, който участва с развитието на невроните. Така доменът, който съдържа както хомеобокс-подобна секвенция, така и втора консервативна POU-специфична секвенция, бива наречен POU (Pit-Oct-Unc).

Въпреки, че има някои различия между отделните POU белтъци, изолиран хомеодомен сам по себе си е достатъчен за секвенционно-специфично свързване на ДНК, но за разлика от класическия хомеобокс, свързването има относително нисък афинитет в отсъствие на POU-специфичния домен. Следователно и двете части на POU домена са необходими за високо-афинитетно секвенционно-специфично свързване на ДНК. Предполага се, че разпознаващите спирали на POU-хомеодомена и POU-специфичния субдомен се свързват към съседни области в голямата бразда на ДНК.

POU-факторите са много интересни тъй като те илюстрират нов аспект на генна регулация – *секвенцията на свързващото място в ДНК може да повлияе ефекта на фактора върху генната експресия*. Така, когато Oct-1 POU-факторът се свърже към неговата прицелна секвенция (ATGCAAAT) в клетъчните гени, той активира слабо транскрипцията. Когато обаче факторът се свърже към друга негова прицелна секвенция (TAATGART; R = пурин) в най-ранните гени на херпес симплекс вируса (HSV), той заема различна конфигурация. Тя позволява свързването на друг фактор VP16, който се кодира от вирусния геном и е силен активатор на транскрипцията.



Фигура 24. Контрол на Pit-1 върху производството на хормони от на хипофизната жлеза.

Pit-1 свързва се с енхансери в гените, кодиращи растежния хормон (соматотропин), пролактина и др. белтъци на хипофизата и води до тяхното активиране. Самият Pit-1 се регулира в хода на индивидуалното развитие: Pit-1 мРНК се синтезира в хипофизните клетки на 16-дневен ембрион; след 2-3 дена неговата мРНК се транслира само в определени клетки на предната хипофиза: в *соматотрофните клетки*, където Pit-1 активира производството на растежния хормон; в *лактотрофните клетки*, където Pit-1 активира производството на минимални количества пролактин; производството на значителни количества от хормона се извършва само при едновременно стимулиране с естроген по време на бременност и малко след това; в *тиреотрофните клетки*, където Pit-1 стимулира производството на тиреоидни хормони. В кортикотрофните и гонадотрофните клетки на предната хипофиза Pit-1 мРНК не се транслира.

Този ефект не е ограничен до взаимодействието на клетъчен фактор с вирусен белтък. Така например, Pit-1 активира растежния хормон в соматотрофните клетки и на пролактина в лакотрофните клетки на хипофизата (Фигура 24). Промоторът на

растежният хормон, обаче, не е активен в лакотрофните клетки, въпреки свързването на Pit-1 към промотора, тъй като конформацията, която Pit-1 възприема на свързващо място на този промотор позволява на Pit-1 да присъедини клетъчния ко-репресор NCoR и допълнителни фактори, в резултат на което транскрипцията се потиска.

Следователно, за различните типове ДНК-свързващи белтъчни фактори, секвенцията на местата на свързване в ДНК може да повлиява конфигурацията на свързания фактор, и по този начин, неговата способност за свързване на други регулаторни молекули, и в зависимост от тях факторът може да активира или да потиска транскрипцията на съответните прицелни гени.

Става ясно, че POU-факторите представляват фамилия от белтъци, свързани към хомеодоменните белтъци, които най-вероятно играят важна роля в развитието. Така например, инактивирането на Pit-1 гена води до нарушения в развитието на хипофизата, в резултат на което се забавя растежа и се развива нанизъм при мишки и човек.

2.2.2. ТФ с ДНК-свързващ мотив „цинков пръст“

- „Цинков пръст“ два цистеина-два хистидина Cys₂/His₂

Транскрипционният фактор TFIIIA играе ключова роля в регулирането на транскрипцията на гените, кодиращи 5S рРНК. Когато този фактор е бил изолиран и пречистен за първи път е било установено, че белтъчната молекула има „повтаряща се“ структура и е свързана с 7-11 атома Zn. Когато генът, кодиращ TFIIIA, е бил клониран и секвениран, става ясно, че повтарящата се структура включва Tyr/Phe-X-Cys-X-Cys-X_{2,4}-Cys-X₃-Phe-X₅-Leu-X₂-His-X_{3,4}-His-X₅ и е представена от 9 копия в молекулата на TFIIIA. Повтарящата се структура съдържа два непроменливи цистенови и два непроменливи хистидинови остатъка, които свързват един атом Zn. Този мотив се нарича „**цинков пръст**“ Cys₂/His₂ на базата на предложената структура, в която бримка от 12 АК, сред които консервативни левцинови и фенилаланинови остатъци, както и няколко основни АК, настърча от повърхността на молекулата, и е закотвена в основата си чрез цистеиновите и хистидиновите остатъци, които тетраедрично координират един атом Zn. Предполаганото взаимодействие на Zn с консервативните цистеинови и хистидинови остатъци в тази структура е било по-късно потвърдено чрез адсорбционна спектроскопия на пречистен TFIIIA белтък.

Подобни цинкови пръсти Cys₂/His₂ са били открити множество ТФ, които контролират РНК полимераза II, като Sp1, който съдържа три последователни цинкови пръсти, Kruppel белтъка при *Drosophila*, който съдържа четири цинкови пръсти, както и много други, част са представени в Таблица 3.

Във всички тези случаи, цинковите пръсти изграждат ДНК-свързващия домен на ТФ.

Зависимостта на свързването на ДНК от този мотив е демонстрирана при Kruppel белтъка при *Drosophila*, който е жизненоважен за правилното гръдно-коремно развитие. В този случай, единична мутация в един от цистеиновите остатъци на пръста, при която той се заменя от серин и не може да свързва Zn, води до мутантна мушица,

Sp1 (Specificity protein 1) – ТФ, който активира или потиска транскрипцията в отговор на физиологични или патологични стимули; свързва се със GC-богати мотиви; регулира експресията на голям брой гени, участващи в процесите на растеж, програмирана клетъчна смърт, диференциация, имунен отговор, хормонален отговор; активността му се регулира чрез пост-транслационни модификации като фосфорилиране, сумоилиране, гликозилиране и др.

която не би могла да бъде различена от тази, продуцирана чрез отстраняване на целия ген, което показва незаменимата роля на цинковия пръст.

Таблица 3. ТФ с ДНК-свързващ мотив цинкови пръсти Cys₂/His₂

Организъм	Ген	Брой цинкови пръсти
<i>Drosophila</i>	<i>kruppel</i>	4
	<i>hunchback</i>	6
	<i>snail</i>	4
	<i>glass</i>	5
Дрожди	<i>ADR1</i>	2
	<i>Swi5</i>	3
<i>Xenopus</i>	<i>TFIIIA</i>	2
	<i>Xfin</i>	37
Бозойници	<i>NGF-1a (Egr1)</i>	3
	<i>MK1</i>	7
	<i>MK2</i>	9
	<i>Evi 1</i>	10
	<i>Sp 1</i>	3

ТФ, които съдържат цинковите пръсти, вземат участие в контрола на развитието на гръбначни, както и на *Drosophila*, където множество белтъци, регулиращи развитието като Krüppel, Hunchback и Snail съдържат цинкови пръсти. Взаимодействието на тези белтъци с хомеотичните белтъци, съдържащи алтернативен ДНК-свързващ мотив НТН, са от изключителна важност за развитието на *Drosophila* и най-вероятно на много други организми.

• Мултицистеинови „цинкови пръсти“ Cys₂/Cys₂

Ефектът на стероидните хормони върху генната експресия при бозайници е един от най-добре характеризираните примери на генна регулация. Стероид-регулираните гени са между първите, за които е установена регулация на ниво транскрипция чрез свързването на специфични рецепторни белтъци (рецептори) към специфична ДНК секвенция, което води до промяна в хроматиновата структура и до образуване на места хиперчувствителни на ДНК-за I. Когато гените, кодиращи ДНК-свързващите рецептори за различни стероидни хормони като глюкокортикоиди и за естроген, са били клонирани, е установено, че те изграждат фамилия от белтъци,

Стероидни хормони – липидна природа; неразтворими във вода; транспортират се в кръвта, свързани към белтъчни носители; могат да проникват в клетката през клетъчната мембрана; взаимодействат с ядрените рецептори и повлияват транскрипцията на ДНК

- **Кортикостероидни хормони** - синтезират се от кората на надбъбречната жлеза на човека; разделят се на две основни групи: (а) **минерал-кортикоидите**, които регулират баланса на соли в организма, (б) **глюкокортикоиди или кортикостероиди**, които регулират обмяната на въглехидратите (кръвната захар вкл.), мазнините и белтъците.
- **Полови хормони** – синтезират се в половите жлези; разделят се на две основни групи: (а) **мъжки полови хормони** (андрогени, вкл. тестостерон), които се произвеждат в тестисите, (б) **женски полови хормони** (естроген и прогестерон), които се синтезират в яйчниците; контролират развитието на половите белези,

кодирани от различни, но родствени гени. На свой ред, тези белтъци са свързани с други рецептори, които опосредстват отговора на клетката на такива хормони като тиреоидните хормони, ретиноевата киселина или витамин D, което поражда идеята за еволюционно свързана фамилия от гени, кодиращи ядрени рецептори, известна като фамилия на гените на рецепторите на стероид-тиреоидните хормони. При сравняване на детайлната структура на членовете на тази фамилия става ясно, че всеки от тях има мултидоменна структура която включва централен силно консервативен домен. На базата на експерименти, в които трункирани молекули на рецепторите са интродуцирани в клетки и измерена тяхната активност, се показва, че централният консервативен домен опосредства способността на рецептора да свързва ДНК, докато С-крайният регион участва в свързването на съответния хормон, а N-крайният регион е ангажиран с продуцирането на максимална индукция на транскрипцията на прицелните гени (гените, които се контролират от ядрения рецептор).

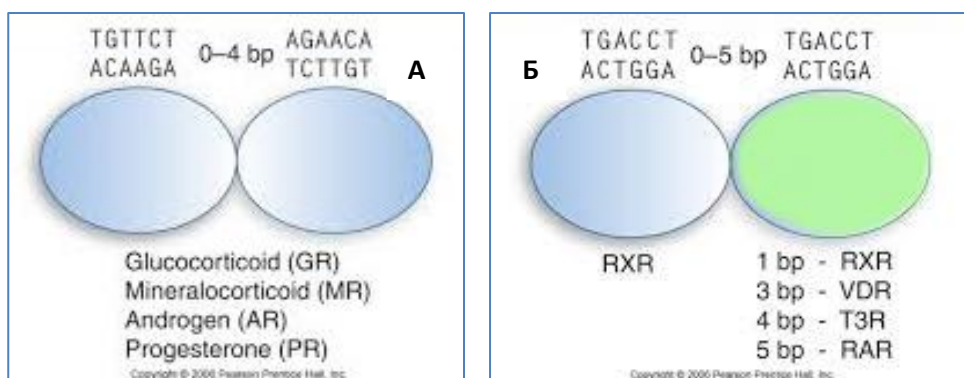
Секвенционният анализ на ДНК-свързващия домен на множество рецептори показва, че той може да бъде представен чрез консенсусната секвенция Cys-X₂-Cys-X₁₃-Cys-X₂-Cys-X₁₅₋₁₇-Cys-X₅-Cys-X₉-Cys-X₂-Cys-X₄-Cys. Подобно на цинковите пръсти Cys2/His2, свързването на ДНК от този елемент е зависимо от



Фигура 25. ДНК-свързващият домен на ядрените рецептори е изграден от два цинкови пръсти

присъствието на Zn или на други метали напр. Cd. Този елемент, означаван като „**цинков пръст**“ Cys2/Cys2, може да бъде изобразен като два обикновени цинкови пръста, в които на мястото на структурата два цистеина-два хистидина се разполагат четири цистеина за да свържат цинка и които са свързани с линкерна област от 15-17 вариабилни АК (Фигура 25). Структурният анализ също посочва, че двата цинкови пръста Cys2/Cys2 в стероидните рецептори изграждат един структурен елемент, за разлика от цинковите пръсти Cys2/His2, при които всеки пръст формира отделна структурна единица.

ДНК-свързващият домен има основна роля за определяне на типа хомодимерен рецептор (глюкокортикоиден, естрогенен и тиреоиден рецептори), който се свързва към палиндромен (обърнат) повтор, както и на типа хетеродимерен рецептор (рецептори на 9-cis ретиноевата киселина, all-trans ретиноевата киселина и витамин D, тиреоиден рецептор), който се свързва към прав повтор (Фигура 26).



Фигура 26. Свързване на ядрените рецептори към елементите на отговора на прицелния ген, който контролират.

(А) Хомодимерните рецептори се свързват с обърнати повтори (палиндроми).
(Б) Хетеродимерните рецептори се свързват с прави повтори. Разпознаването на елемента на отговора от различните рецептори зависи от разстоянието между двата повтора на елемента. (по Krebs 2011)

Съществуването на къса ДНК-свързваща област в множество различни стероидни рецепторни белтъци, които се свързват към определени, но много подобни, секвенции е позволило детайлно изучаване на елементите в тази структура, които са важни за секвенционно-специфичното свързване на ДНК. Така секвенциите, които са свързани с отговор към глюкокортикоидно и естрогенно въздействие, са много подобни една на друга. Ако богатият на цистеин регион на естрогенния рецептор се замени с този на глюкокортикоидния рецептор, полученият химерен рецептор ще има ДНК-свързващата специфичност на глюкокортикоидния рецептор, но тъй като всички останали региони ще произхождат от естрогенния рецептор, той ще продължи да свързва естроген. Така, химерният рецептор ще индуцира експресията на гените, които отговарят на глюкокортикоиди (които носят ДНК-свързващо място за глюкокортикоидния рецептор), но в този случай в отговор на естроген (към който рецепторът се свързва).

Тези изследвания са били по-нататък разширени чрез промяна на двете АК, разположени между третия и четвъртия цистеин на N-крайния цинков пръст в глюкокортикоидния рецептор с техните еквиваленти от естрогенния рецептор, която е довела до глюкокортикоиден рецептор, който индуцира гените, които отговарят на естроген. Така промяната само на две АК, в белтък, който има 777 АК, може напълно да промени ДНК-свързващата специфичност на рецептора. Специфичността на химерния рецептор към гените, които отговарят на естроген, е била още по-силно повишена и чрез промяна на една друга АК, която е локализирана в линкерната област между двата цинкови пръста, което показва, че този регион също играе роля в контрола на специфичността на свързване на ДНК. Така, *ДНК-свързващата специфичност на стероидните рецептори изглежда се обуславя от кооперирането на N-крайния цинков пръст и линкерния регион*. За разлика от ефекта на мутациите в първия пръст и линкерната област, понататъшни промени на пет АК във втория цинков пръст са били достатъчни за да може химерния рецептор да разпознава свързващите места на рецептора на тиреоидния хормон. Тъй като свързващите места на рецептора на тиреоидния хормон не се различават по секвенция, а само по разстоянието между двете половини на палиндромната ДНК секвенция на разпознаване, това показва, че *S-крайният цинков пръст в жизненоважен за опосредстване на белтък-белтъчните*

взаимодействия между двата мономера на рецептора, които се свързват към двете половини на палиндромната секвенция, и по този начин за контролиране на оптималното разстояние между двете половини при свързване на даден рецептор (който формира хомодимер).

Следователно, изучаването на множеството родствено вързани стероидни рецептори и техните взаимоотношения с ДНК секвенциите, към които се свързват, е направило възможно определянето на критичната роля на първият цинков пръст и съседния линкерен участък за контролиране на секвенцията, към която хомодимерния рецептор се свързват.

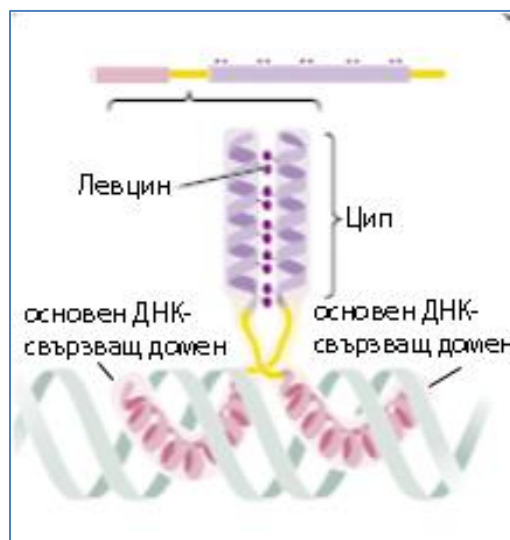
ДНК-свързващият домен играе важна роля и при детерминиране на хетеродимерните рецептори, които се свързват към прави повтори. Така когато тези повтори са разделени с една базова двойка, те могат да свързват хомодимера на ретиноид-Х-рецептора (retinoid-X-receptor, RXR) и да опосредстват отговора на 9-*cis* ретиноевата киселина, която се свързва към този рецептор. RXR хомодимерът, обаче, не може да се свързва към елемента на отговор, при който двата повтора са разделени от две, три, четири или пет базови двойки. В тези хетеродимерни комбинации, ефектът на RXR е потиснат и отговорът се определя от другия компонент на хетеродимерния рецептор. Така разстояние от две или пет базови двойки между повторите позволява свързването на хетеродимера на RXR и рецептора на ретиноевата киселина (RAR) и води до отговор на *all-trans*-ретиноевата киселина, която се свързва към RAR. Разстояние от четири базови двойки прави възможно свързването на хетеродимера на RXR и рецептора на тиреоидния хормон (TR) и води до отговор на тиреоидния хормон, а разстояние три базови двойки позволява свързването на хетеродимера на RXR и рецептора на витамин Д (VDR) и води до отговор на витамин Д.

Така, различните ДНК-свързващи домени на разнообразните ядрени рецептори продуцират множество профили на свързване на хомо- и хетеродимери към различни места на свързване (елементи на отговора), правейки възможно по този начин членовете на фамилията на ядрените рецептори да продуцират голямо разнообразие от отговори.

2.2.3. ТФ с основен ДНК-свързващ домен и димеризиращ мотив Левцинов цип или Спирала-примка-спирала

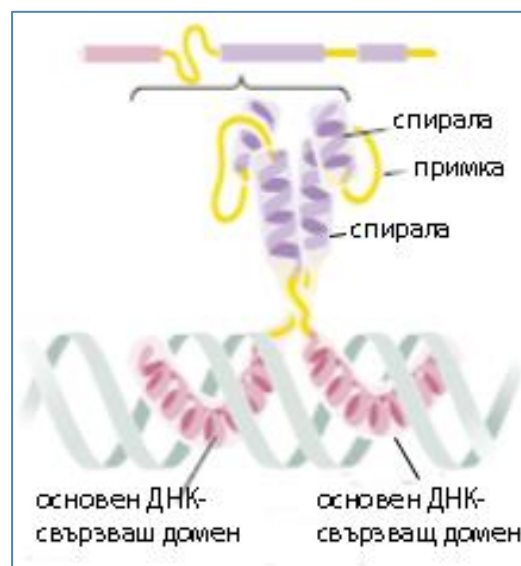
При изследване на гена, кодиращ транскрипционния фактор C/EBP, който участва в стимулиране на експресията на гени, специфични за черния дроб, се установява, че белтъчната молекула съдържа област от 35 АК, в която всяка седма АК е левцин, който бива наречен мотив **левцинов цип**. Подобна област е установена и в прото-онкогенните белтъци Мус, Fos и Jun, които първоначално са били идентифицирани на базата на тяхната способност да трансформират клетъчни култури до канцерогенен фенотип, и се е допуснало, че те действат като регулират транскрипцията на други клетъчни гени. Предположило се е, че богатият на левцин регион образува α -спирала, в която левцинов остатък се среща на всеки два оборота от една и съща страна на спиралата, и че левциновите остатъци улесняват димеризацията на две молекули на транскрипционния фактор. В съответствие с това, замяната на индивидуалните левцинови остатъци в C/EBP с валинови или изовалинови остатъци премахва способността на белтъка да образува димери, и да се свързва специфично, към ДНК елементите, които разпознава.

За разлика от ДНК-свързващите домени от типа НТН, левциновият цип не се свързва директно с ДНК. По-скоро той участва в димеризирането на транскрипционния фактор и осигурява правилната белтъчна структура за свързване на ДНК от съседни на левциновия цип региони в белтъка, които са богати на основни АК, наречени **основни ДНК-свързващи домени**, които могат да взаимодействат директно с киселата ДНК (Фигура 27). В съответствия с това, мутации в богатия на основни АК регион премахват способността на белтъка да свързва ДНК без обаче да повлиява димеризиращата способност. Подобно разполагане един до друг на основни ДНК-свързващи домени и левцинов цип са били също открити във Fos и Jun онкогенните белтъци и дрождевия транскрипционен фактор GCN4, в които един регион от 60 АК носи всичката информация необходима за димеризация и секвенционно-специфичното свързване на ДНК. По такъв начин левциновият цип играе роля, подобна на тази на втория цинков пръст при стероидните рецептори, който не се свързва директно с ДНК, а по-скоро модулира активността на съседния ДНК-свързващия регион.



Фигура 27. Транскрипционен фактор с основен ДНК-свързващ домен и димеризиращ мотив Левцинов цип.

Въпреки, че първоначално основният ДНК-свързващ домен е бил идентифициран в белтъци, съдържащи левцинов цип, по-късно той е бил открит на базата на хомология и в множество други ТФ, които не съдържат левцинов цип. Тези фактори включват транскрипционния фактор MyoD, който играе ключова роля при активиране на гени, специфични за скелетните мускули, както и E12 и E47 факторите, които имат важна роля за развитието на имуноглобулин-продуциращите В лимфоцити и за развитието на нервната система. В тези случаи, основният ДНК-свързващ домен е разположен в съседство на регион, който може да образува мотив **спирала-примка-спирала** (от англ. helix-loop-helix, **HLH**). Този мотив, обаче, е различен от НТН мотива на хомеобокса, тъй като може да образува две амфипатични спирали, на които всички натоварени АК се намират на едната страна на спиралата (Фигура 28). Мотивът играе роля подобна на левциновия цип, тъй като позволява димеризация на молекулите на ТФ и така улеснява свързването на съседния основен ДНК-свързващ мотив. В съответствие с това, делеции или мутации на основния домен на



Фигура 28. Транскрипционен фактор с основен ДНК-свързващ домен и димеризиращ мотив HLH.

MyoD не влияят на димеризацията, но пречат на ДНК свързването. Делеции или мутации на мотива HLH обаче, премахват както димеризиращата, така и ДНК-свързващата способност на този ТФ.

Установено е, че замяната на три АК в основния регион на E12 с техните еквиваленти на MyoD позволява основният регион на E12 да активира мускулно-специфични гени. Кристалната структура на MyoD белтъка, свързан към ДНК, подсказва, че тези АК играят критична роля, свързана със специфичната структурна конфигурация на основния домен, в която той може да взаимодейства с други активиращи транскрипционни фактори. В съответствие с това, замяната на тези три АК позволява на мутантния E12 белтък да свързва друг транскрипционен фактор, който е мускулно-специфичен - MEF2A. *Следователно, подобно на POU домена, основният домен функционира като ДНК-свързващ домен и като място за белтък-белтъчни взаимодействия, които са критични за транскрипционното активиране.*

Мотивите левцинов цип и HLH опосредстват димеризацията на транскрипционните фактори, които ги съдържат, при което образувалата се димерна молекула е способна да свързва ДНК чрез разположения в съседство на мотивите основен ДНК-свързващ домен. Тази способност предоставя допълнителен аспект в регулацията, осъществявана от тези фактори. *Така в допълнение на образуването на димери между две идентични молекули, е възможно образуването на димери между два различни фактора, които могат да имат различни свойства от гледна точка на секвенционно-специфичното свързване и генното активиране при сравнение с хомодимерите на единия или другия от двата фактора.* Пример за това са онкобелтъците Fos и Jun. Jun може да се свързва като хомодимер към мястото на разпознаване на AP1 (TGAGTCAG), който опосредства индуцирането на транскрипцията в отговор на форболови естери. Fos не може да се свързва към ДНК сам, но може да свързва с Jun. Хетеродимерът се свързва към мястото на разпознаване на AP1 с афинитет 30-пъти по-голям в сравнение с този на хомодимера на Jun и е значително по-ефективен за индуциране на транскрипцията на гените, съдържащи регулаторния елемент. Образуването на хомодимера, както и на хетеродимера, са зависими от мотива левцинов цип, който е открит и в двата белтъка.

2.2.4. ТФ с други ДНК-свързващи мотиви

Макар, че болшинството ДНК-свързващи мотиви, които са били идентифицирани в известни ТФ, попадат във фамилиите, дискутирани по-горе, други мотиви остават извън тази класификация. С характеризирането на нови ТФ, се появяват нови ДНК-свързващи мотиви и нови взаимодействия между ДНК-свързващите мотиви при различните ТФ, което води до дефиниране на нови фамилии ТФ. Така напр. ТФ на рРНК – UBF, съдържа ДНК-свързващ домен, който е открит и в други фактори като белтъците на силно-подвижната група (HMG, high mobility group), поради което доменът е наречен HMG box, а ДНК-свързващият домен на p53 белтъка, е свързан с този на NFκB фамилията.

3. Активиране и потискане на транскрипцията на РНК полимераза II

ТФ могат да повлияят инициацията на транскрипцията или позитивно като я стимулират (*активатори*) или негативно като я потискат (*репресори*). Активаторите и репресорите оказват съответния ефект като взаимодействат с базалния транскрипционен апарат на РНК полимераза II, с други ТФ или хроматин-регулирущи комплекси.

3.1 Активиращи домени

Активиращият домен е специфичен за всеки активатор и участва в активиране на транскрипцията след свързването му към ДНК. Тези домени са били идентифицирани чрез така нар. *“домен разменящи” експерименти* (domain swap, англ. – размяна на домени), при които различни области на даден ТФ се свързват с ДНК-свързващия домен на друг ТФ и се изследва способността на химерния белтък да активира транскрипцията. Така напр. домен разменящи експерименти са идентифицирали два домена в глюкокортикоидния рецептор, един на N-края на молекулата и друг в близост до C-края, които могат независимо един от друг да опосредстват генното активиране, когато са свързани с ДНК-свързващия домен на друг ТФ. C-крайният активиращ домен е разположен в близост на хормон-свързващия домен на рецептора и може да посредничи за активиране на транскрипцията само в присъствие на хормон. Следователно, този домен има важна роля в стероид-зависимото активиране на транскрипцията при наличие на хормон.

Интересен пример за модулната организация на ТФ е взаимодействието на клетъчния ТФ Oct-1 с *trans*-активиращия белтък VP16 на херпес симплекс вируса. Така въпреки, че VP16 съдържа много силен активиращ регион, който може да индуцира транскрипцията, когато е изкуствено свързан с ДНК-свързващия домен на дрождевия ТФ GAL4, макар че самият VP16 не съдържа ДНК-свързващия домен и не може да се свързва към ДНК. Транскрипционното активиране от VP16 след вирусна инфекция е зависимо от неговата способност да формира белтък-белтъчен комплекс с клетъчния конститутивно експресиран ТФ Oct-1. Този комплекс след това се свързва към октамер-свързания мотив TAATGARAT (R-пурин) на най-ранните вирусни гени чрез ДНК-свързващия домен на Oct-1 и тогава активира транскрипцията чрез активиращия домен на VP16. Следователно в този случай, домомените за свързване на ДНК и транскрипционно активиране са локализирани върху различни белтъци в ДНК-свързващия комплекс. Друг подобен пример е взаимодействието на Oct-1 с молекулата на клетъчния ко-активатор OCA-B, който участва в регулацията на имуноглобулиновите гени.

Херпес симплекс вирус - вирус на обикновения херпес (herpes simplex virus HSV); носители на вируса са 90% от жителите на земята, а 30% от тях имат проявления на симптомите поне веднъж годишно; вирусът е представен от два типа- HSV1, причиняващ лабиален херпес (херпес на устните или кожен херпес) и HSV2, причиняващ генитален херпес (засяга най-често лигавиците на половите органи).

- **Кисели домени**

Сравняването на различни активиращи домени показва липса на силна хомология между аминокиселинните секвенции, но разкрива високо съдържание на кисели аминокиселини, което води до силен отрицателен товар. Така напр. активиращият домен на глюкокортикоидния рецептор съдържа 82 АК, от които 17 са с киселинен

характер. Предполага се, че в случая с VP16, отрицателният товар на неговия активиращ домен позволява установяване на електростатични взаимодействия с TFIID31 компонента на TFIID, с който той взаимодейства за активира транскрипцията. Оказва се, че киселите активиращи домени съдържат консервативни хидрофобни АК, които имат важна роля за активиране на транскрипцията. Така, когато активиращият домен на VP16 взаимодейства с TFIID31, същият претърпява конформационни промени, в следствие на което три хидрофобни АК се разполагат в съседство и контактуват директно с TFIID31. Следователно, както киселите, така и хидрофобните АК остатъци, са необходими за действието на активиращите домени.

- **Домени богати на глутамин**

Анализът на конститутивния ТФ Sp1, който се свързва с свързващото място на много генни промотори, показва, че двата му много силни активиращи домена съдържат около 25% глутаминови остатъци и много малко негативно натоварени остатъци. Глутамин-богатите мотиви са изключително важни за транскрипционното активиране, тъй като тяхната делеция от активиращия домен води до премахване на активирането. Нещо повече, транскрипционната активност може да бъде възстановена чрез замяна на глутамин-богатите региони на Sp1 с глутамин-богатия регион на ТФ Antennapedia на *Drosophila*, който няма хомология със секвенцията на Sp1. Следователно, подобно на киселите активиращи домени, активиращата способност на богатите на глутамин домени не се дефинира от първичната структура, а от природата им да бъдат богата на специфични АК остатъци. Подобни мотиви са открити в октамер-свързващите ТФ Oct-1 and Oct-2, както и в много други ТФ при различни видове.

- **Домени богати на пролин**

Изследванията на конститутивния ТФ CTF/NF1, разкриват, че неговият С-краен регион е изключително богат на пролинови остатъци, които представляват около една четвърт от АК в него. Регионът е способен да активира транскрипцията, когато е свързан към ДНК-свързващ домен на друг ТФ. Подобно на богатите на глутамин региони, непрекъснатата последователност от пролинови остатъци е способна да опосредства активирането, което показва, че функционирането на този тип домени зависи най-вече от изобилието на пролина. Подобни пролин-богати региони са открити и в други ТФ като Jun, AP2 и Oct-2. Всъщност Oct-2 е пример за ТФ, който съдържа два типа активиращи домени, единият от които е богат на пролинови, а другият на глутаминови остатъци.

3.2. Механизми на действие на активаторите

- **Взаимодействие на активаторите с TFIID комплекса**

Свързването на TFIID комплекса (съдържащ TBP и TAFs) към ДНК е първата стъпка както при модела на последователното сглобяване на базалния транскрипционен апарат, така и при холоензимния модел на сглобяване. Активаторите могат да стимулират свързването на TFIID към ДНК, както и да установят контакт с TFIID след неговото разполагане върху ДНК. Свързването на активатора към TFIID обикновено води до промени в неговата конфигурация, в следствие на което да се повиши способността му да свързва други компоненти на базалния транскрипционен апарат, включително и на РНК полимера II.

Въпреки, че съществуват доказателства за свързване на активатори към TBP, в повечето случаи активирането на транскрипцията изисква взаимодействие между активатор и един или повече TAFs.

Установено е, че различни типове активатори могат да контактуват с различни TAFs – глутамин-богати активатори като Sp1 могат да се свързват с TAF_{II}110, кисели активатори като VP16 свързват TAF_{II}31 и пролин-богати активатори свързват TAF_{II}55. Взаимо-действията между активаторите и TAFs са още по-сложни поради съществуването на тъканно-специфични TAFs. Например В-лимфоцитите експресират форма на TAF_{II}130 известна като TAF_{II}105, която не се експресира в други клетъчни типове.

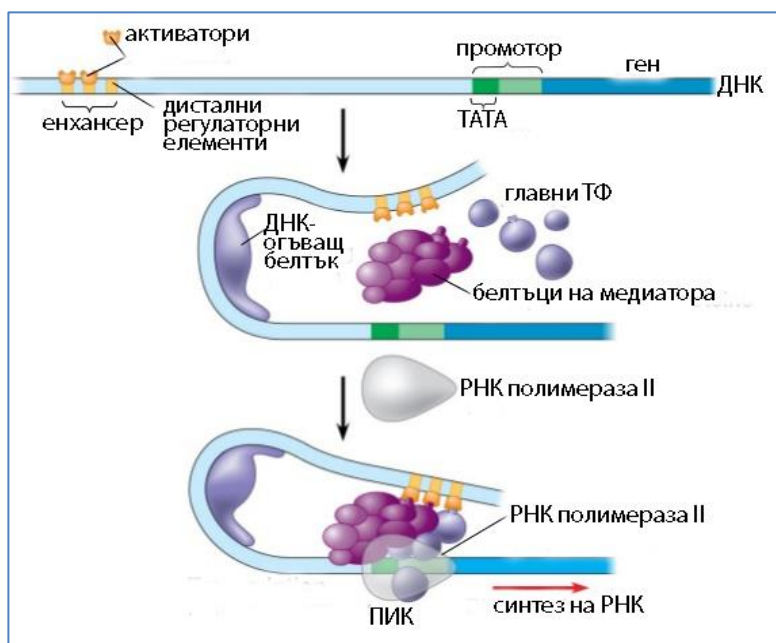
Очевидно, TAFs функционират като посредник между активаторите и базалния транскрипционен апарат, осъществявайки белтък-белтъчни взаимодействия. Те често биват отнасят към групата на ко-активаторите, тъй като те не контактуват директно с ДНК, а опосредстват ефекта на ДНК-свързаните активатори.

- **Взаимодействие на активаторите с медиатора и SAGA**

Свързването на TFIID или TFIIB към ДНК прави възможно присъединяването на самата РНК полимераза II, носеща TFIIF. Установено е, че СТД на най-голямата субединица (RPB1) на ензима, който участва в етапите на инициацията и елонгацията на транскрипцията, взаимодейства с транскрипционни активатори. Изследвания на мутации в този регион, който е съставен е повтарящ се хексапептиден елемент, показват, че делециите потискат транскрипционното активиране, докато увеличаването на броя на повтарящия се елемент води до активиране на транскрипцията.

Експерименти, проведени в клетки на дрожди и бозайници, показват, че активаторите не осъществяват директен контакт с РНК полимераза II. Идентифициран е комплекс от около 20 полипептида, наречен **медиатор** (Фигура 29), който е свързан, и дори частично обвива СТД на RPB1. Става ясно, че активаторите взаимодействат с медиаторния комплекс, който на свой ред стимулира активността на полимеразата.

Оказва се, че различни класове активатори взаимодействат с различни компоненти на медиаторния комплекс. Има данни, че ТФ, които репресират транскрипцията, също контактуват с медиатора. Това е довело до идеята, че медиаторът интегрира сигналите, идващи от транскрипционните активатори и репресори, свързани към специфични ДНК секвенции



Фигура 29. Взаимодействие на медиатора с активаторите и базалния транскрипционен апарат

на дадени гени, и след това взаимодействат с РНК полимеразата II за да продуцират необходимите нива на транскрипция.

В допълнение на медиатора, друг мултипротеинен комплекс, известен като **SAGA**, също играе ключова роля в транскрипционното активиране. Комплексът съдържа компоненти като Gcn5, които могат да ацетилира хистоните, и да променят хистоновата структура.

- **Взаимодействие на активаторите с ко-активаторите**

От изложеното по-горе става ясно, че транскрипционните активатори оказват активиращ ефект върху базалния транскрипционен апарат без да контактуват директно с неговите компоненти. Сигналите от активатора към базалния транскрипционен апарат се предават с помощта на посредник, чиято роля се изпълнява от **ко-активатора**, който осъществява белтък-белтъчно взаимодействие със ДНК-свързания активатор и някой от компонентите на базалния транскрипционен апарат.

Един от най-добре проучените ко-активатори е **CBP** (CREB-binding protein, от англ. - белтък, свързващ CREB). Както подсказва неговото име, CBP е дефиниран първоначално като белтък, който свързва ТФ CREB (cAMP response element-binding protein, от англ. - белтък, свързващ елемента на отговора на цАМФ), който активира специфични гени в отговор на цикличния АМФ. CBP може да свърже CREB само, когато CREB е фосфорилиран при сериновия остатък на позиция 113. По-късно става ясно, че CBP играе ключова роля в опосредстването на транскрипционното активиране от множество ДНК-свързани ТФ като фамилията на ядрените рецептори, NFκB, p53 и MyoD. Широкото използване на CBP от ТФ, активирани от различни сигнални пътища, може да води до конкуренция между сигналните пътища при ограничени количества на CBP. Това на свой ред води до феномен, при който едновременно стимулиране на два различни пътя, напр. като тези стимулирани от глюкокортикоидния хормон и форболовия естер, води до взаимно инактивиране на двата пътя и липса на транскрипционно активиране.

Въпреки, че CBP взаимодейства с членове от фамилията на ядрените рецептори, той не е единствения ко-активатор, участващ в тяхното активиране. В различни случаи, при хормонално въздействие специфични ко-активатори се свързват към съответните рецептори, което позволява на рецептора да осъществи хормонално-зависимо активиране на транскрипцията.

Най-вероятно ко-активаторите могат да стимулират транскрипцията поне чрез два различни механизма. Така напр. CBP може да се свързва с различни компоненти на базалния транскрипционен апарат, а също така, тъй като притежава хистон ацетил-трансферазна активност, да отваря хроматиновата структура.

- **Синергизъм на активаторите**

Активаторите могат да активират транскрипцията като взаимодействат (пряко или косвено) с удивителен набор от прицелни молекули, включващ РНК полимеразата II/медиаторен комплекс, ко-активатори, хроматин-ремоделиращи комплекси, и компоненти на TFIID, както и други компоненти на базалния транскрипционен апарат като на TFIIA, TFIIЕ и TFIIH. Най-вероятно, всеки един от тези компоненти може да бъде прицелен както за един и същ активатор, така и за различни активатори. Голямото разнообразие от активатори и множеството потенциални прицелни молекули са важна предпоставка за продуциране на силно **синергично активиране на транскрипцията**.

В много случаи свързването на два ТФ към различни места в промотора може да продуцира много по-силно транскрипционно активиране, в сравнение с ефекта, продуциран при свързване само на един от двата ТФ. Подобно, усилено транскрипционно активиране се наблюдава, когато множество копия на един същ ТФ се свържат към прицелната ДНК. Тези синергични ефекти, най-вероятно, играят ключова роля в продуцирането на силна транскрипционна активност в отговор на специфични стимули и/или на специфични етапи от индивидуалното развитие.

3.3. Потискане на транскрипцията

Макар, че болшинството ТФ, описани до сега, повлияват позитивно транскрипцията, са установени и ТФ, които имат инхибиращ ефект върху транскрипцията, който те могат да упражнят по няколко възможни механизма.

- **Репресорите могат да действат косвено чрез инхибиране на позитивния ефект на активаторите**

Репресорът може да пречи на свързването на активиращата молекула към ДНК, като това може да бъде постигнато чрез *стимулиране от репресора на образуването на плътна пакетирана хроматинова структура*, която да ограничава достъпа до ДНК. Типичен пример за това са репресорите от Polycomb фамилията на *Drosophila*, които се свързват към хомеотичните гени и ги организират в „затворена“ хроматинова структура.

Алтернативно, *репресорът може да се свързва специфично към свързващото място на активатора* и така да пречи на неговото свързване. Такъв ефект се наблюдава в промотора на β -интерферона, където свързването на няколко позитивно действащи фактори е необходимо за неговото активиране. Друг фактор действа негативно чрез свързване към този регион на ДНК като по този начин пречи на свързването на активаторите. При вирусна инфекция, репресорът се инактивира, което позволява на свързването на активаторите и активиране на транскрипцията.

Репресорът може да попречи на активатора да стимулира транскрипцията и чрез механизъм известен като *погасяване* (от англ. quenching). Това става, когато *репресорът се свързва към участък на ДНК, съседен на мястото на свързване на активатора*. Такъв е случая с промотора на онкогена c-myc, при който негативният ТФ myc-PRF се свързва в съседство на позитивния ТФ myc-CF1 и не му позволява да активира експресията на c-myc.

Потискане на транскрипцията може да бъде постигнато и чрез *образуване на комплекс между репресора и активатора* в разтвор, което пречи на свързването на активатора с ДНК. Този механизъм е установен при инхибиращия фактор Id, който димеризира с необходимия за мускулната детерминация фактор MyoD, и пречи на неговото свързване с ДНК. Подобен ефект може да бъде постигнат и когато репресорът се свърже с активатора след неговото свързване към ДНК и блокира активиращата му способност. Така напр. онкогенният белтък MDM4 се свързва с анти-онкогенния белтък p53 и маскира неговия активиращ домен.

Освен, че може да блокира свързването на активатора или да потисне активиращата му способност, *репресорът може също така да предизвиква деградация на активиращата молекула*. Така онкогенният белтък MDM2, вместо да „погасява“

активирането от p53, индуцира разграждането на p53 като катализира неговото убикитиране и така го насочва към протеазни ензими.

- **Репресорите могат да действат директно чрез инхибиране на сглобяването или на активността на базалния транскрипционен апарат**

Механизмът, при който репресорът директно потиска инициацията на транскрипцията, е много добре изучен при **рецептора на тиреоидния хормон (c-erbA)**, който е член на фамилията на ядрените рецептори. В отсъствие на тиреоиден хормон, тиреоидният ядрен рецептор се свързва към своя специфичен елемент на отговора (TRE) върху ДНК, присъединява ядрения ко-репресор NCo-R (от англ. Nuclear receptor co-repressor) и инхибира транскрипцията на прицелните гени. Установено е, че NCo-R се свързва към Sin3/RPD3 белтъчния комплекс, който деацетитира хистоните и индуцира по-плътно пакетирание на хроматиновата структура. В присъствие на хормона обаче, рецепторът претърпява конформационни промени, в следствие на което освобождава ко-репресора и свързва ко-активатори като CBP, и активира транскрипцията.

Този пример разкрива изключително важната роля на ко-активаторите и ко-репресорите за регулиране на генната експресия. Така тиреоидният хормон може да модулира генната експресия като контролира алтернативното свързване към рецептора на ко-репресора и ко-активатора. Подобно, различните конфигурации, които ТФ може да възприема при свързване към различни места на свързване върху ДНК, могат да променят неговия ефект върху транскрипцията като променят способността му да свързва ко-активатори или ко-репресори.

Рецепторът на тиреоидния хормон е пример за това, че разграничаването между активатори и репресори не винаги е ясно. В някои случаи, един и същ фактор може да стимулира или инхибира транскрипцията в зависимост от обстоятелствата, и този ефект обикновено се регулира в зависимост от това дали факторът присъединява ко-активатор или ко-репресор.

Рецепторът на тиреоидния хормон е представен от две алтернативни форми - α -1 и α -2, които са резултат на алтернативен сплайсинг и се различават по това, че в α -2 част от хормон-свързващия домен е представена от алтернативна АК секвенция и не може да свързва тиреоидния хормон. Формата α -2, когато присъства в клетката, пречи на свързването на α -1 и блокира индукцията на гените в отговор на тиреоидния хормон. Тези две алтернативни форми, които се образуват в различни количества в различните клетъчни типове, опосредстват противоположни ефекти на зависимата от тиреоидния хормон генна регулация.

Докато рецепторът на тиреоидния хормон може да упражнява негативен ефект върху транскрипцията след свързването си към ДНК и с помощта на ко-репресор, други репресори могат да свързват директно към базалния транскрипционен апарат и да инхибират транскрипционната активност. Пример за това е ТФ Dr1, който се свързва към TBP и пречи на присъединяването на TFIIIB и по този начин потиска сглобяването на базалния транскрипционен апарат върху промотори, съдържащи ТАТА-бокс.

4. Регулация на ТФ на РНК полимераза II

ТФ играят незаменима роля в процеса на транскрипция – свързвайки се към специфични секвенции в ДНК, те водят до активиране или потискане на транскрипцията на други гени (прицелни гени). Транскрипцията е ключов процес в контрола на генната експресия, чрез който се формират различни клетъчни типове и тъкани и който прави възможен отговора на клетката на специфични стимули. За да постигне това, е необходимо самите ТФ да бъдат диференциално активни в различните клетъчни типове, а също така трябва тяхната активност да бъде модулирана от специфични сигнали. ТФ могат да бъдат регулирани чрез контрол върху синтеза им или чрез контрол върху активността им.

4.1. Регулация на синтеза на ТФ

Този механизъм на регулация на ТФ е широко използван в ситуации, където е необходимо да се поддържа специфичен профил на генна експресия през продължителен период и това може да бъде лесно постигнато чрез поддържане синтеза на специфичен ТФ на определено ниво през този период. По този механизъм се регулират ТФ, които контролират тъканно-специфични гени или гени, свързани с развитието на организма. Така, гените, кодиращи този тип ТФ, се транскрибират в специфичен клетъчен тип или на специфичен етап от индивидуалното развитие. Обикновено, контролът върху транскрипцията на гените на тези ТФ се опосредства от други ТФ, които се синтезират само в определени тъкани и клетъчни типове.

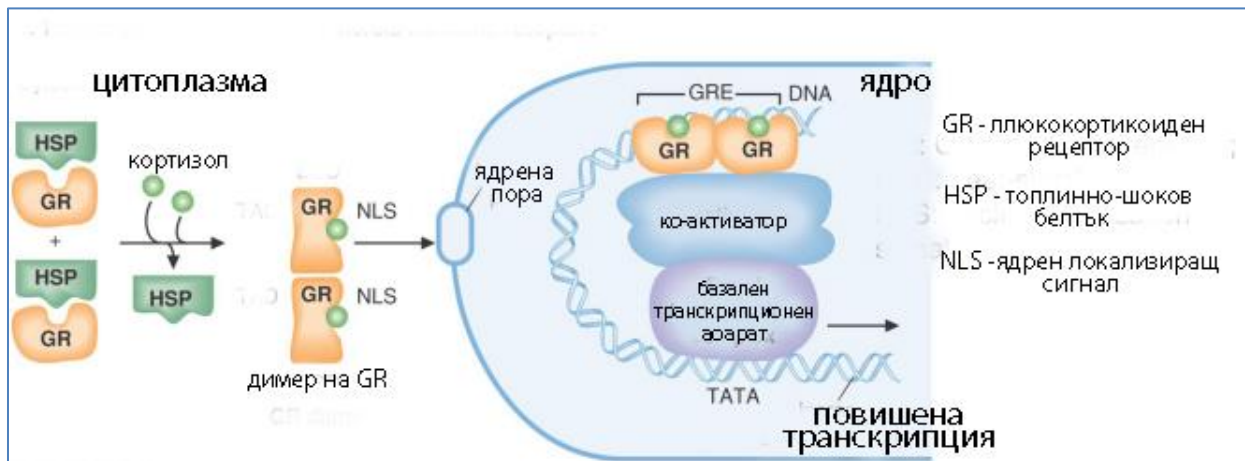
4.2. Регулация на активността на ТФ от лиганди, които навлизат в клетката

При този тип регулация, сигнална молекула, която влиза в клетката, се свързва с ТФ и директно регулира неговата активност, например като индуцира конформационни промени в неговата молекула. В многоклетъчните организми един от най-добре характеризираните примери за този тип регулация са ядрените рецептори, които включват рецепторите за тиреоидния и стероидните хормони.

Рецепторът на тиреоидния хормон в отсъствие на хормона се свързва с ДНК, след което присъединява ко-репресорни молекули и потиска транскрипцията на прицелните гени. В присъствие на хормона, ДНК-свързаният ТФ претърпява конформационни промени, които позволяват по-скоро присъединяването на ко-активатори, отколкото на ко-репресори, и активират транскрипцията на прицелните гени.

За разлика от рецептора на тиреоидния хормон, глюкокортикоидният рецептор нормално е локализиран в цитоплазмата, и само в присъствие на глюкокортикоидния хормон се свързва към GRE в ДНК и активира транскрипцията на съответните прицелни гени. Докато в клетката рецепторът се свързва към ДНК в присъствие на хормона, то *in vitro* рецепторът може да се свърже към ДНК и в отсъствие на хормона. Причината за това е, че глюкокортикоидният рецептор образува комплекс в цитоплазмата с топлинно-индуцируемия белтък HSP90. При свързване с хормона, рецепторът се дисоциира от HSP90 и формира димер. Отделянето на HSP90 е свързано и с излизане на повърхността на скрития до този момент във вътрешността на комплекса ядрен локализиращ сигнал на рецептора, с чиято помощ той навлиза в ядрото (Фигура 30). Установено е, че глюкокортикоидният рецептор е способен не само да активира транскрипцията, а също така и да я потиска. При негативната регулация,

глюкокортикоидният рецептор се свързва към елемент на отговора nGRE, който е сходен на GRE. В случая на POMC гена, присъствието на nGRE води до свързване на рецептора под формата на тример. В тази конформация рецепторните молекули не само не могат да упражняват активиращ ефект, а и блокират свързването на други активиращи фактори, в резултат на което транскрипцията се репресира.



Фигура 30. Активиране на глюкокортикоидния рецептор чрез свързване с кортизол

4.3. Регулация на активността на ТФ чрез фосфорилиране, индуцирано от извънклетъчни сигнални

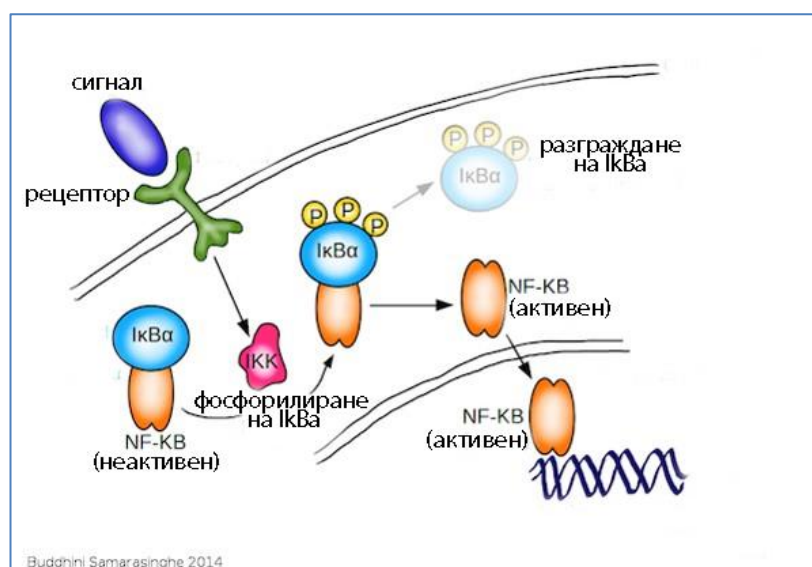
Молекули като стероидните хормони и тиреоидния хормон мога да преминават през клетъчната мембрана и да регулират ТФ като се свързват с тях директно вътре в клетката. Този механизъм не е възможен за много сигнални молекули с белтъчна природа, които действат на повърхността на клетката и не могат да проникнат в нея. Такива извънклетъчни сигнални молекули могат да оказват ефект като се свързват към повърхностно-клетъчни рецептори, които на свой ред изпращат сигнали вътре в клетката, които в крайна сметка модифицират специфични ТФ.

- **JAK/STAT пътеката** - при тази пътека, извънклетъчни сигнални молекули на цитокините се свързват към специфични мембранни рецептори. След свързването на цитокините към извънклетъчната част на рецептора, сигналът се предава към JAK белтъка, който в свързан с извънклетъчната част на рецептора. Сигналът активира киназната активност на JAK и му позволява да фосфорилира членове на фамилията ТФ - **STAT** (signal transducers and activators of transcription, англ. - сигнални трансдюсери и активатори на транскрипцията). В нестимулирани клетки, STAT факторите се намират в цитоплазмата. При стимулиране на клетката и последващо фосфорилиране на STAT, те образуват димери, които се придвижват до ядрото, където се свързват към специфични ДНК секвенции и включват експресията на прицелни гени. Тази сигнална пътека е известна като JAK/STAT пътека.

Цитокини - малки сигнални белтъчни молекули, които се секретират от специфични клетки на имунната система; играят важна роля в регулацията на растежа и диференциацията на специфични клетъчни типове; участват в регулацията на имунитета и възпалението.

- **NFκB/IκB системата** предоставя пример за ТФ – **NFκB** (nuclear factor kappa-light-chain-enhancer of activated B cells, англ. – ядрен фактор на енхансера на карпа-

леката верига в активирани В клетки), чиято активност може да бъде регулирана чрез фосфорилиране на инхибиращия го белтък – **IκB**. NFκB е ТФ, който се активира от множество стимули като възпаление или стресови сигнали. Такива стимули предизвикват освобождаване на възпалителни цитокини като тумор-некротичния фактор α (TNF-α) или интерлевкин-1, които се свързват към специфични клетъчни рецептори. В нестимулирани клетки, двете субединици на NFκB (p50 и p65) са асоциирани с инхибиторния белтък – IκB. Това свързване маскира ядрените локализиращи сигнали на p50 и p65 и NFκB остава в цитоплазмата. При излагане на активиращи стимули и следващо активиране на рецепторите, се извършва активиране на IκB киназите, които на свой ред фосфорилират IκB. Това фосфорилиране бележи IκB за понататъшни модификации като убикитиниране и разграждане от специфични протеолитични ензими. Разграждането на инхибиторния белтък е свързано и с излизане на повърхността на скритите до този момент във вътрешността на комплекса NFκB/IκB ядрени локализиращи сигнали на NFκB, което позволява неговото навлизане в ядрото и активиране на транскрипцията на прицелните гени (Фигура 31).



Фигура 31. Активиране на NFκB чрез фосфорилиране и разграждане на неговия инхибитор IκB.

IκB киназите фосфорилират и друг регулаторен белтък – CBP. NFκB е един от множеството ТФ, които използват като CBP ко-активатор. Нещо повече, NFκB се конкурира за ограничените количества CBP в клетката. Всъщност, фосфорилирането на CBP от IκB киназите засилва неговата способност за свързване на NFκB и намалява афинитета му към други ТФ като p53. В този случай фосфорилирането се използва за засилване на активността на NFκB пътеката като я прави по-ефективен конкурент за ограничените количества на CBP в клетката.

4.4. Регулация на активността на ТФ чрез други пост-транслационни модификации

- **Ацетилиране** – въпреки, че ацетилирането на лизинови остатъци е най-често използваната пост-транслационна модификация за регулиране на хистоновата активност, по-късно е установено, че то играе важна роля за регулиране активността на множество различни ТФ. Напр. в NFκB/IκB системата, в допълнение на фосфорилирането на IκB, се извършва ацетилиране на NFκB, което потиска

взаимодействието I κ B и NF κ B. Ацетилирането на NF κ B се извършва от CBP (и свързания с него p300). Следователно, освен че ацетилира хистоните и отваря хроматиновата структура след свързване към специфичен транскрипционен активатор, CBP/p300 може също да модифицира активността на ТФ чрез ацетилиране.

- **Метилиране** – освен, че могат да бъдат фосфорилирани, CBP и p300 могат да бъдат метилирани при специфични аргининови остатъци. Подобно на ацетилирането, метилирането променя афинитета на тези ко-активатори за различни ТФ. Така метилирането премахва способността на CBP/p300 да се свързват към CREB, но няма ефект върху афинитета им към ядрените рецептори.
- **Убикитиниране** - по-горе беше дискутирано, че фосфорилирането на I κ B се следва от убикитиниране, което води до разграждане на I κ B и освобождаване на активен NF κ B. Друг пример за този тип модификация е индуцируемият от хипоксия фактор **HIF-1**, който опосредства генното активиране в отговор на ниски кислородни нива. В условия на високи нива на кислород, HIF-1 α се модифицира чрез прибавяне на хидроксилна (OH) група към специфичен пролинов остатък в молекулата. Хидроксилирането бележи HIF-1 α за последващо убикитиниране и разграждане на белтъка. При ниски нива на кислород, хидроксилиране на HIF-1 α не се извършва. HIF-1 α е стабилен и се свързва с HIF-1 β за да образува хетеродимер, който активира транскрипцията на прицелните му гени.
- **Сумоилиране** - Подобно на убикитина, SUMO модифицира лизинови остатъци, поради което SUMO и убикитина могат да се конкурират за специфични места върху молекулата на ТФ. Така напр. в случая на I κ B, SUMO модификацията защитава I κ B от убикитиниране и води до неговото стабилизиране. Други прицелни молекули за SUMO са ко-репресорите N-CoR и RIP140 на ядрените рецептори, при които сумоилирането засилва техният репресиращ ефект върху транскрипцията.

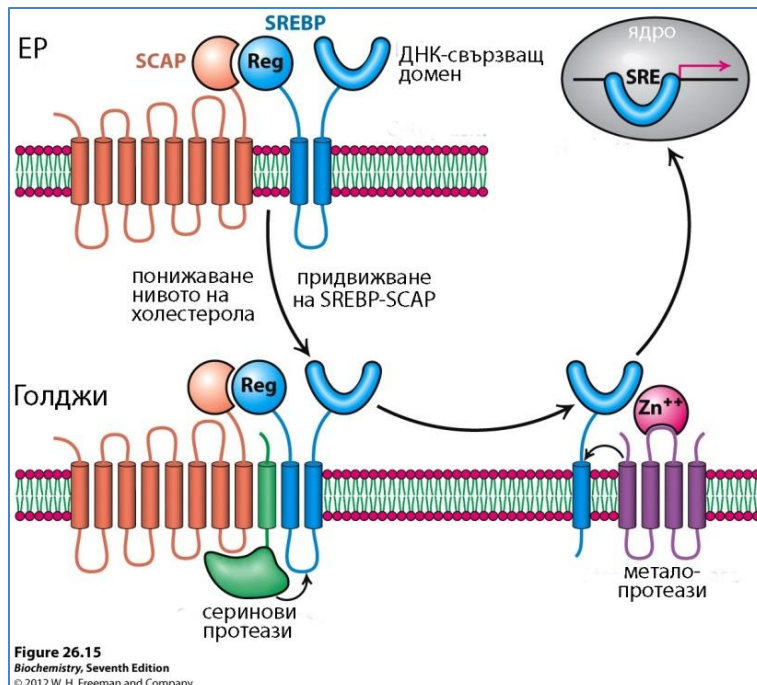
Убикитин - малък регулаторен белтък (76 аминокиселинни остатъка), силно консервативен, открива се в почти всички тъкани на еукариотните организми (от лат. *ubique* - въездесъщ); основна функция - да бележи други белтъци за протеолитично разграждане; той може също така да повлиява тяхната активност, локализация и взаимодействие с други белтъци.

SUMO (*small ubiquitin-related modifier*) - подобен на убикитина белтък; модифицира други белтъци; участва в контрола на ядрения транспорт, репликацията и поправката на ДНК, сигналната трансдукция и др.

4.5. Регулация на активността на ТФ чрез сигнали, които регулират процесинга на техните предшественици

- Активиране чрез разрязване на мембрано-свързан предшественик - типичен пример е **SREBP** (sterol regulatory element-binding proteins, англ. – белтъци свързващи стероловия регулаторен елемент), кои-то активират генната експресия в отговор на понижени нива на холестерол. При високи нива на холестерола, SREBP е представен от неактивен предшественик, който е свързан с холестерол-свързващия белтък **SCAP** (*SREBP cleavage-activating protein*, англ. – белтък, активиращ разрязването на *SREBP*) и е закотвен към мембраната на ендоплазматичния ретикулум. Кога-то нивата на холестерола в клетката паднат, SCAP не е свързан повече с холестерола и комплекса SREBP-SCAP се придвижва

от ендоплазматичния ретикулум към апарата на Голджи, където SREBP претърпява протеолитично разрязване и освобождава N-терминалния си домен. Последният представлява ДНК-свързващ фактор, който се придвижва към ядрото за да се свърже с прицелните си гени и да активира тяхната експресия (Фигура 32).



Фигура 32. Активиране на SREBP чрез разрязване на неговия предшественик при понижаване на нивото на холестерола

- Разрязването на ТФ може да го превърне от активатор в репресор - свързването на Hedgehog белтъците към техните специфични рецептори инхибира фосфорилирането на Ci (cubitus interruptus) – ТФ, който участва в контрола на сегментационния поляритет при *Drosophila*. Дефосфорилирането стабилизира молекулата на Ci, което прави възможно нейното свързване към ДНК, както и присъединяването на ко-активатори, в резултат на което се активират прицелните гени. Обратно, когато Hh липсват, Ci се фосфорилира и претърпява протеолитично разрязване, което води до образуване на малък фрагмент. Този белтъчен фрагмент все още е свързан с ДНК, но към него се присъединяват ко-репресори, в резултат на което се потиска транскрипцията.

Hedgehog (Hh) фамилия - мощни сигнални белтъчни молекули, които действат като морфогени в хода на индивидуалното развитие на гръбначни

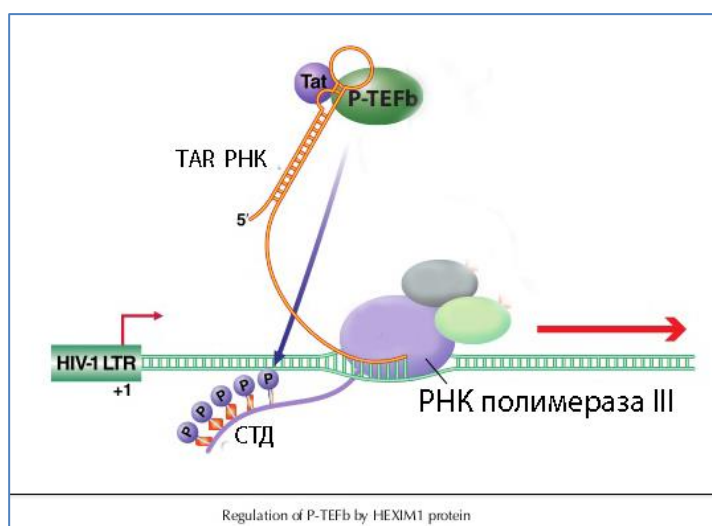
5. Регулация на транскрипционната елонгация на РНК полимераза II

Макар, че болшинството случаи на транскрипционна регулация се извършват на ниво инициация, установената способност на РНК полимераза II да спира след инициране на транскрипцията разкрива възможността за контролиране на транскрипцията и през етапа на елонгация.

Такава форма на регулация е отговорна например за почти десетократното намаляване на нивата на мРНК, кодираща клетъчния онкоген *c-тус*, което е установено при индуцирана диференциация на човешката про-миелоидна клетъчна линия HL-60 до гранулоцитен клетъчен тип. Генът *c-тус* се състои от 3 екзона, които са представени в мРНК, която се получава след изрязване на интроните от първичния транскрипт чрез сплайсинг. Сравняването на нивата на транскрипция на първия и втория екзон в недиференцирани и диференцирани клетки показва, че регулацията се извършва по-скоро през етапа на елонгация, отколкото при инициация. Въпреки, че подобен брой полимеразни молекули иницират транскрипцията на *c-тус* гена в двата клетъчни типа, болшинството молекули терминират в края на първия екзон, без да транскрибират целия ген и без да продуцират функционална мРНК. Няколко дни след началото на диференциация, бързото инхибиране на транскрипционната елонгация се допълва и инхибиране на транскрипцията още на етапа на инициация. Подобни ефекти върху транскрипционната елонгация са установени и при други клетъчни онкогени – *c-myb*, *c-fos* и *c-mos*.

Регулация на ниво транскрипционна елонгация е идентифициран при ретровируса HIV-1, който причинява синдрома на придобитата имунна недостатъчност (СПИН). В ранните етапи на инфекция се иницират много ниски нива на транскрипция от HIV-1 промотора, а много от транскриптите терминират в близост до промотора, при което се произвеждат много къси РНК молекули. По-късно, вирусният белтък Tat се свързва към нарастващите HIV-1 РНК молекули в област, известна като Tar, която е локализирана от +19 до +42 относно стартовата точка на транскрипция +1. Това свързване има два ефекта. Първо, силно се увеличава транскрипционната инициация, което води до иницирането на много РНК транскрипти. В допълнение, Tat превъзможва блокирането на елонгацията, в резултат на което се увеличава силно делът на дългите транскрипти, които могат да кодират вирусни белтъци (Фигура 33).

Прицелното място за ТФ, които регулират транскрипционната елонгация, е СТД на най-голямата субединица на РНК полимераза II. Фосфорилирането на серинов остатък 2 в копията на хептапептида на СТД (виж гл. Базална транскрипция) е ключово събитие,



Фигура 33. Регулация на транскрипцията на РНК полимераза III при HIV-1.

необходимо за преминаване на РНК полимераза II от инициация към елонгация на транскрипцията. Установено е, че HIV-1 кодираният белтък Tat е способен да взаимодейства с киназата pTEF-b, която фосфорилира STD при серин 2, и по този начин прави възможно протичането на транскрипционната елонгация.

6. Регулация на транскрипцията на РНК полимераза I и РНК полимераза III

- **Транскрипцията на РНК полимераза I и РНК полимераза III може да се регулира чрез промяна в нивото на експресия или на активността на техните главни ТФ.**

Това е възможно, тъй като двете полимерази транскрибират гени, които участват в образуването на рибозомите. По тази причина двете полимерази могат да бъдат контролирани по-глобално в сравнение с РНК полимераза II, която транскрибира голям набор от гени, които вземат участие в разнообразни функции.

Така например главните ТФ променят активността си по време на митозата, когато транскрипцията на рибозомните гени е инхибирана в следствие на понижена активност на главните ТФ UBF и SL1/TIF-1B. Когато клетката влиза в митоза, двата фактора се фосфорилират, което понижава тяхната активност и блокира транскрипцията на рибозомните гени от РНК полимераза I. След завършване на митозата, двата фактора се дефосфорилират и това води до активиране на SL1. Въпреки това обаче, транскрипцията на рибозомните гени се инициира едва, когато клетката навлезе в G1 фазата на следващия клетъчен цикъл, защото това изисква фосфорилиране на UBF в друга област на молекулата, което е необходимо за активиране на фактора.

Растящите клетки, в сравнение с почиващите клетки, се нуждаят от засилена активност на РНК полимераза I и РНК полимераза III. Анти-онкогенните белтъци ограничават клетъчния растеж чрез инхибиране на специфични за всяка от двете полимерази главни ТФ. Обратно, продуктът на онкогена c-Myc стимулира активността на тези фактори за да повиши транскрипцията, осъществявана от двете полимерази, и да стимулира клетъчния растеж.

Активността на РНК полимераза III може да бъде променена и чрез изменения в нивата на експресия на нейните главни ТФ. Така понижението на нива на TFIIIC са причина за понижената транскрипционна активност на полимеразата, установена по време на диференциацията на ембрионните клетки.

- **Транскрипцията на РНК полимераза III може да се регулира чрез специфични ТФ, които се свързват или към ДНК, или към РНК**

Присъствието на регулаторен елемент в транскрибирания от РНК полимераза III регион, и от тук както в ДНК, така и в РНК продукта, предоставя друг потенциален механизъм на генна регулация. Такъв регулаторен механизъм е описан при 5S рРНК гени, които притежават вътрешен регулаторен елемент, към който се свързва TFIIIA.

Така в генома на жабата *Xenopus laevis* са открити два типа 5S рРНК гени: ооцитни гени, които се експресират само при развитието на ооцита преди оплождането, и соматични гени, които се транскрибират в клетките на ембриона и възрастните индивиди. И двата типа гени съдържат вътрешен регулаторен елемент, който свързва TFIIIA, необходим за започване на транскрипцията. Тъй като секвенциите на елемента в двата типа гени имат малки различия, TFIIIA има по-голям афинитет за свързване към соматичните гени, при сравнение с ооцитните гени. По тази причина ооцитните гени се транскрибират само в ооцитите, където TFIIIA е в много големи количества, но не и в другите типове клетки, където TFIIIA е в лимитиращи количества, достатъчни да активират соматичните гени.

В развиващите се ооцити, TFIIIA се синтезира в големи количества и транскрипцията на ооцитните гени се инициира. Колкото повече 5S рРНК молекули, съдържащи TFIIIA-свързващ регион, се натрупват в зреещия ооцит, толкова повече молекули на TFIIIA биват свързани. Свързаният от рРНК молекулите TFIIIA се складира в специални структури и става недостъпен за транскрипция на ооцитните гени. Така нивата на TFIIIA падат под необходимите нива за транскрипция на ооцитните гени и транскрипцията им спира, докато транскрипцията на високо-афинитетните соматични гени остава незасегната.

Референции

- Britten B.J. and Davidson E.H. (1969) Gene regulation for higher cells: a theory. *Science* 165: 349-357.
- Courey, Albert J. "Mechanisms in transcriptional regulation" (2008) Blackwell Publishing, ISBN: 978-1-4051-0370-1
- Fry C.J. and Peterson C.L. (2002) Transcription. Unlocking the gates to gene expression. *Science* 295(5561):1847-8.
- Juven-Gershon T. and James T. Kadonaga (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Developmental Biology* 339: 225–229.
- Kadonaga J. T. (2012) Perspectives on the RNA Polymerase II Core Promoter. *Interdiscip Rev Dev Biol*. 1(1): 40–51.
- Krebs J. E., Goldstein E.S., Kilpatrick S.T. (2011) "Lewin's Genes XI" Jones & Bartlett Learning. ISBN-13: 978-0763766320
- Latchman, David S. "Eukaryotic transcription factors" (2008) Elsevier Publishing, ISBN: 978-0-12-373983-4
- Latchman, David S. „Gene control" (2010) Garland Science, Taylor and Francis Group, ISBN: 978-08153-6513-6
- Lawrence P.A. and Morata G. (1994) Homeobox genes: their function in *Drosophila* segmentation and pattern formation. *Cell*. 78(2):181-9.
- Lemon B. and Tjian R. (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev*. 14(20):2551-69.
- Mangelsdorf D.J., Thummel C., Beato M., Herrlich P., Schütz G., Umesono K., Blumberg B., Kastner P., Mark M., Chambon P., Evans R.M. (1995) The nuclear receptor superfamily: the second decade. *Cell*. 83(6):835-9.
- McKenna N.J. and O'Malley B.W. (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell*. 108(4):465-74.
- Mitchell P.J. and Tjian R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*. 245(4916):371-8
- Mohd-Sarip A. and Verrijzer C.P. (2004) Molecular biology. A higher order of silence. *Science*. 306(5701):1484-5.
- Orphanides G, Lagrange T, Reinberg D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev*. 10(21):2657-83.
- Russell J. and Zomerdijs J. C.B.M (2005) RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem Sci*. 30(2): 1-20.
- Sandelin A., Carninci P., Lenhard B., Ponjavic J., Hayashizaki Y., Hume D.A. (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet*. 8(6):424-36.
- Scully K.M., Jacobson I.E.M, Jepsen K., Lunyak V., Viadiu H., Carriere K., Rose D.W., Hooshmand, F, Aggarwal A.K., Rosenfeld M.G. (2000) Allosteric Effects of Pit-1 DNA Sites on Long-Term Repression in Cell Type Specification. *Science* 290:1127-1131.
- Thomas M. C. and Chiang C.-M. (2006) The General Transcription Machinery and General Cofactors. *Critical Reviews in Biochemistry and Molecular Biology*, 41:105–178.

Thomas M.C. and Chiang C.M. (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol.* 41(3):105-78.

Turner J, Crossley M. (199) Mammalian Krüppel-like transcription factors: more than just a pretty finger. *Trends Biochem Sci.* (6):236-40

Wade J.T. and Struhl K. (2008) The transition from transcriptional initiation to elongation. *Curr Opin Genet Dev.* 18(2): 130–136.

<https://www.boundless.com>

<http://www.learnnext.com>

Най-често срещани съкращения

бд – базова двойка

ДНК – дезоксирибонуклеинова киселина

мРНК – матрична РНК (наричана още информационна РНК, иРНК)

нд - нуклеотид

нкРНК - не-белтък кодираща РНК

ПИК – преинициаторен комплекс

РНК – рибонуклеинова киселина

рРНК – рибозомна РНК

СТД – С-терминален домен

ТФ – транскрипционен фактор (TF)