

## Espressione genica: trascrizione (parte 2)

### Eventi associati all'inizio della trascrizione nei procarioti e negli eucarioti

In biologia molecolare, la trascrizione è il processo mediante il quale le informazioni contenute nel DNA vengono trascritte enzimaticamente in una molecola complementare di RNA. Concettualmente, si tratta del trasferimento dell'informazione genetica dal DNA all'RNA. Nel caso in cui il DNA codifichi una proteina, la trascrizione è l'inizio del processo che porta, attraverso la produzione intermedia di un mRNA, alla sintesi di peptidi o proteine funzionali. La trascrizione presenta un meccanismo di controllo della fedeltà (o proofreading), ma esso è molto meno efficace di quelli legati alla replicazione del DNA; lo stesso meccanismo di sintesi dell'RNA, nel quale l'enzima RNA polimerasi ha un ruolo centrale, comporta un numero di errori decisamente maggiore.

Come per la replicazione del DNA, la trascrizione procede in direzione  $5' \rightarrow 3'$ . Più precisamente, il filamento lungo il quale il DNA viene scandito, detto filamento stampo, è percorso dall'enzima in direzione  $3' \rightarrow 5'$ . Il nuovo filamento di RNA, identico al filamento senso viene sintetizzato a partire dal suo  $5'$ . La trascrizione avviene attraverso particolari enzimi detti genericamente RNA polimerasi. Tali proteine sono spesso denominate RNA polimerasi DNA-dipendenti, dal momento che producono una molecola di RNA a partire da una di DNA. Tali enzimi utilizzano nucleotidi trifosfati (ovvero nucleotidi con tre gruppi fosfato) per formare l'RNA. Durante il processo, dai nucleosidi trifosfati vengono rimossi due gruppi fosfato per formare un legame covalente tra un nucleotide e quello successivo. La RNA polimerasi si lega al DNA solo presso particolari sequenze, dette promotori, che non sono trascritte.

Dal promotore iniziano a inserirsi i nucleosidi trifosfato per formare una sequenza di nucleotidi che sarà complementare al filamento di DNA in via di trascrizione. Dopo l'individuazione del promotore, la RNA polimerasi rende il DNA adatto alla trascrizione. Il filamento di RNA inizia quindi ad allungarsi, attraverso l'aggiunta di un nucleotide per volta. Il primo nucleotide del neofilamento di RNA trattiene i tre gruppi fosfato mentre quelli successivi vengono privati di due gruppi fosfato attraverso una reazione esoergonica. Quando, durante la trascrizione nel DNA, si incontrano particolari sequenze di basi (poste solitamente alla fine di ogni gene), la trascrizione termina. Nei procarioti, la macchina trascrizionale è costituita dall'enzima multimerico RNA polimerasi, dal  $\sigma$  indispensabile per il riconoscimento del promotore e da eventuali altri fattori di trascrizione. L'assemblaggio di tutto il complesso di trascrizione basale segue delle fasi ben precise: tutto inizia dalla subunità  $\alpha$  (alfa) che contatterà la subunità  $\alpha 2$  (alfa 2); tale dimerico (o più correttamente eterodimerico) ha già la capacità di legare il DNA anche se in maniera aspecifica; insieme promuovono l'assemblaggio del complesso. A questo punto si legano le subunità catalitiche  $\beta$  e  $\beta'$ ; la subunità  $\beta'$  partecipa al legame aspecifico della

polimerasi con il DNA, mentre la subunità  $\beta$  è la responsabile della vera e propria sintesi della molecola di RNA, che si formerà secondo lo schema d'appaiamento citosina-guanina (C-G), e viceversa e adenina-uracile (A-U); il filamento di RNA infatti incorpora l'uracile al posto della timina.

Una volta formata, l'RNA polimerasi non è capace di riconoscere direttamente il promotore; per questo è necessaria la subunità  $\sigma$  (sigma), che prende parte nei più importanti processi di regolazione della trascrizione nei procarioti. Altro fondamentale fattore per l'assemblaggio è una proteina detta Catabolite Activating Protein (abbreviata come CAP o CRP), che legandosi nella sequenza consenso 35 nucleotidi a monte del sito di inizio (detta "sequenza -35", formata dalla sequenza TTGACA), fa assumere al promotore la corretta topologia affinché l'RNA polimerasi possa creare un legame stabile col DNA. La CAP infatti interagisce con l'oloenzima sia a livello di  $\sigma$  sia a livello del dominio carbossi terminale di  $\alpha$ . La CAP è legata alla sequenza consenso posta a -35 e ricca in AT e distorce il DNA. Anche grazie a questa distorsione  $\delta$  contatta un'ampia regione del promotore in un modo tale da reclutare l'RNA polimerasi, che a questo punto comincia a fare le necessarie connessioni per la lettura del filamento senso (5'→3') e quindi sintetizzare l'ibrido DNA-RNA lungo 9-10 paia di basi. Ma l'evento che dà il via alla reazione deve essere la denaturazione locale del DNA a formare la cosiddetta bolla di trascrizione. In passato si pensava che ciò fosse indotto dalla particolare distorsione generata da tutti i fattori con cui il DNA interagisce, primo fra tutti CAP.

In realtà l'apertura del DNA non avviene a -35 come ci si aspetterebbe vista la minore stabilità locale del DNA (ricco in quel punto di AT), ma più a valle a circa -10 (sequenza TATAAT detta Pribnow box o TATA box). Il segnale di apertura (melting del promotore) è indotto da un attivatore esterno capace di fosforilare la subunità  $\alpha 2$  nel dominio carbossiterminale ( $\alpha$ -CTD). Tale dominio si comporta come una sorta di pannello di controllo di tutta la trascrizione: il suo stato più o meno fosforilato infatti regola non solo la fase iniziale ma anche quella di allungamento e di terminazione. Chiamato anche complesso Binario, binario perché formato dal DNA ed RNA polimerasi. Esso può essere chiuso o aperto. Chiuso : processo reversibile, la RNA polimerasi può staccarsi dal promotore ,ma nei casi in cui ciò non avviene esso subisce una modifica conformazionale ovvero una isomerizzazione. Il legame ad H esistente fra A e T (adenina-timina ) si spezza. La bolla di replicazione si estende dalla sequenza -10 al +4 . Aperto: Per la formazione della bolla di replicazione. A Questo punto la Rna polimerasi presenta nel sito attivo il primo nucleotide da trascrivere e da qui l'avvio della trascrizione. A questo punto l'RNA polimerasi può cominciare a leggere il DNA e ad appaiare le corrette basi azotate al fine di sintetizzare il trascritto primario (RNA non ancora processato). Questa fase della "messa a registro" è la più critica al punto che l'inizio della sintesi segue diversi inizi abortivi in cui l'RNA polimerasi non riesce a sintetizzare un filamento di RNA tanto lungo da far cambiare conformazione a tutto il processo. È indispensabile infatti che avvenga un cambiamento conformazionale di tutto il complesso, il fattore  $\sigma$  che è servito a far riconoscere il promotore deve staccarsi, i legami che l'RNA polimerasi faceva col DNA devono diventare aspecifici, così come CAP non è più necessaria. Probabilmente il tutto viene indotto dall'allungamento

dell'ibrido RNA-DNA che nel momento in cui supera quelle 9-10 paia di basi crea un ingombro sterico tale da indurre il rilascio di tutti gli altri fattori. Perché l'allungamento possa verificarsi il fattore  $\sigma$  deve lasciare l'RNA polimerasi, poiché aumentandone l'affinità per il promotore gli impedisce di allontanarsi da questo e di proseguire il suo cammino lungo il DNA da trascrivere.

Questo è un momento delicato per tutto il processo e non sono rari inizi abortivi che si riferiscono all'abbandono della catena da parte dell'enzima entro l'aggiunta dei primi nove nucleotidi all'acido nascente. Terminati gli eventi abortivi, l'enzima perde i contatti dal -55 a -35 così l'enzima copre soltanto 50 basi. Quando la catena raggiunge 15-20 nucleotidi l'enzima compie un'altra transizione e forma il complesso di allungamento coprendo 30-40 basi. Man mano che l'enzima procede nella trascrizione il DNA si deve svolgere più a valle e deve riavvolgersi a monte del sito di trascrizione, si forma così un tratto di circa 12 basi ibrido DNA-RNA. L'allungamento è in direzione 5'→3' grazie all'attacco nucleofilo da parte dell'estremità 3'-OH della catena neoformata verso il ribonucleotide trifosfato precursore che deve essere aggiunto, e più precisamente sul P in  $\alpha$  liberando pirofosfato. I segnali di terminazione sono nella sequenza di DNA, ma espletano la loro funzione solo quando sono trascritti in mRNA. Essi possono indurre l'RNA di nuova sintesi ad assumere una struttura secondaria (generalmente delle forcine di terminazione) tale da destabilizzare e far staccare la Polimerasi, nel caso dei terminatori intragenici. La struttura a forcina si forma grazie all'elevata presenza di guanine e citosine che instaurano tra loro tre legami ad idrogeno.

Questa sequenza viene anche stabilizzata da legami ad idrogeno dovuti ad appaiamenti non convenzionali. A valle della zona a forcina che si è formata si trova una sequenza di poliU: questo significa che sul DNA è presente una sequenza di poliA, e poiché l'appaiamento tra A ed U è il più debole che si possa avere, il distacco dell'mRNA del DNA, e la conseguente terminazione della trascrizione viene favorita. Se il terminatore non è sufficientemente forte da promuovere da solo la terminazione, e questo può accadere quando non è presente la zona poliU, risulta essere necessaria una sequenza che ha una alta affinità per la proteina RHO, un fattore di terminazione intergenica che stacca letteralmente dal complesso di trascrizione il trascritto primario. RHO si attacca all'mRNA in determinati punti che prendono il nome di RUT, RHO UTILization. Raggiunge la polimerasi e promuove il distacco del DNA quando la polimerasi si sofferma su un terminatore o quando i ribosomi si staccano per la presenza di un codone di terminazione e lasciano lo spazio necessario a RHO per legarsi all'mRNA.

## Trascrizione negli eucarioti

Dato l'elevato numero di geni presenti in un organismo eucariote, la trascrizione è un meccanismo molto più complesso perché necessita di un sistema di regolazione che agisca su più livelli. Una prima notevole differenza sta proprio nelle RNA Polimerasi, esse infatti vengono suddivise in tre categorie in funzione dell'ambito genetico in cui lavorano. L' RNA Polimerasi I presente nel nucleolo è capace di trascrivere i geni per tre dei quattro rRNA ovvero 5.8S, 18S e 28S. L'RNA Polimerasi II che si trova nel nucleoplasma (parte di nucleo che non contiene il

nucleolo) trascrive gli mRNA, i miRNA, i siRNA, alcuni snoRNA e snRNA. Tra gli snRNA figurano U1, U2, U4, U5 e U6 facenti parte di ribozimi coinvolti nell'assemblaggio dello spliceosoma per la maturazione dei trascritti primari. Infine l'RNA Polimerasi III che si trova nel nucleoplasma trascrive l'rRNA 5S, tutti i geni dei tRNA e alcuni snRNA. Tutto ciò ci dà delle significative indicazioni della complessità del genoma eucariotico tale da giustificare la presenza di tre RNA Polimerasi e di conseguenza come queste tre classi abbiano dei sistemi di regolazione molto eterogenei.

Prima fase:

legame con la TBP. Negli eucarioti, a differenza dei procarioti, il DNA è fortemente ripiegato nella cromatina che a sua volta è strutturata in modo complesso e renderebbe difficile qualunque intervento sul DNA senza un meccanismo apposito di svolgimento. Nel complesso costituito da circa un centinaio di subunità proteiche responsabile della trascrizione eucariotica sono compresi alcuni complessi di rimodellamento della cromatina e svariati enzimi che modificano gli istoni agendo sull'epigenoma. I complessi di rimodellamento della cromatina agiscono svolgendo la fibra di cromatina da 30nm costituita da un solenoide di nucleosomi per poi agire sul nucleosoma stesso con l'aiuto di chaperonine degli istoni rimuovendo temporaneamente i due dimeri degli istoni H2A-H2B dall'ottamero e facendovi quindi rimanere solo il tetramero H3-H4. Altri complessi di rimodellamento della cromatina svolgono temporaneamente il DNA dal corrispondente nucleosoma rendendolo disponibile per la trascrizione.

Le operazioni svolte dai vari complessi di rimodellamento della cromatina consumano ATP. Eventuali modificazioni agli istoni (in particolare alle loro code) possono essere effettuate dagli enzimi modificatori degli istoni (HAT, HDAC, DNMT e altri). È arduo stabilire come si susseguono gli eventi che portano alla trascrizione di un gene, dal momento che sono possibili variazioni nei tempi d'assemblaggio dei vari fattori proteici in base alle sequenze di DNA che si devono trascrivere, alle sequenze consenso e ad altri fattori. Convenzionalmente la trascrizione genica inizia quando a livello del promotore si assembla una classe di proteine nota come fattori generali di trascrizione della Polimerasi II (TFII). Il compito di queste proteine è di segnalare alla RNA Polimerasi II la collocazione del promotore su un dato gene e le sue sequenze consenso, stabilizzarla in loco, aprire la doppia elica di DNA, attivare la RNA Polimerasi II e svolgere tutte quelle funzioni che garantiscono un corretto inizio del processo di trascrizione. Ci sono cinque fattori generali di trascrizione essenziali: TFIIB, TFIID, TFIIE, TFIIIF e TFIIH. TFIID è descritto come il primo ad entrare in gioco, si tratta di un grosso complesso proteico formato da 12 subunità, tra queste particolarmente importante è la subunità TBP (TATA binding protein). La TBP si lega a livello del TATA-box, una sequenza consenso presente a monte di molti geni generalmente ad una distanza di 25-30 nucleotidi dal punto di inizio della trascrizione. La sua funzione è quella di produrre due forti piegature al DNA (o kink) a che inducono in questo punto un angolo di circa 90°. Questo forte angolo non sarà la causa dell'apertura del promotore (o melting) ma ha la funzione di allargare i solchi maggiori al fine di favorire una migliore interazione delle altre

proteine del complesso di attivazione basale. Tutto ciò vale solo nei geni con l'elemento TATA e solo se non ci sono meccanismi di attivazione a monte di tutto ciò. Le restanti subunità TAF di TFIID riconoscono altre sequenze consenso minori presso il punto d'inizio della trascrizione e regolano il legame alla TATA-box da parte di TBP.

Seconda fase:

assemblaggio del complesso DAB. In realtà la TBP intercetta la sequenza TATA spesso legata ad altri fattori chiamati TAF's, e in questa forma viene chiamata TFIID ("TBP+TAF's" = TFIID) per formare il complesso di inizio preliminare. A questo punto la topologia del DNA è tale da poter reclutare altri fattori di trascrizione ovvero il TFIIA e il TFIIB. Tale complesso trimerico TBP TFIIA TFIIB è chiamato anche DAB. Il complesso DAB così assemblato è capace di legare un quarto fattore di trascrizione che porta con sé la RNA Polimerasi II che è TFIIF, questo è formato da 3 proteine chiamate RAP 30, RAP 74 e RAP 38 (che fa parte in realtà anche di TFIIS) alcune di esse hanno delle forti omologie con il fattore *s* procariotico e questo ci fa capire come la funzione di TFIIF è quella di permettere il riconoscimento del promotore alla RNA Polimerasi II anche se in questo caso il quadro è ben più complesso.

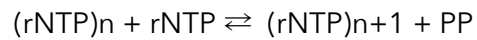
Terza fase:

superamento dello stallo. Quando TFIIF porta la RNA Polimerasi II sul promotore essa interagisce con delle  $\alpha$ - eliche in posizione carbossi terminali alla TBP. Tale interazione richiama sulla TBP altri 2 fattori di trascrizione la TFIIE e TFIIH, la prima ha attività elicastica mentre la seconda attività chinastica sulla RNA Polimerasi II, queste reazioni causano rispettivamente un allargamento della bolla di trascrizione e un cambiamento conformazionale alla RNA Polimerasi II che se opportunamente fosforilata nella regione C-Terminale supera la fase di stallo e inizia a diventare processiva.

Quarta fase:

pulizia del promotore e allungamento. Nonostante ciò nelle prime fasi la velocità di trascrizione non può essere elevata a causa delle ancora forti interazioni con tutti gli altri fattori ancora legati alla RNA Polimerasi II. La fase di pulizia del promotore avviene grazie al rilascio del DAB e di TFIIF che avevano inizialmente il solo ruolo di stabilizzare il complesso di inizio. Tale rilascio avviene grazie a opportune fosforilazioni/defosforilazioni della RNA Polimerasi II a livello del dominio C-Terminale che potremmo definire una sorta di pannello di controllo su cui accedono

molti altri regolatori della trascrizioni sia in fasi precoci che tardive. Nell'operazione di allungamento del filamento di RNA, le RNA polimerasi catalizzano la seguente reazione di polimerizzazione:



Trattandosi di una reazione all'equilibrio, le polimerasi possono fare una pausa cinetica, non dovuto ad una vera e propria attività di proofreading ma ad al semplice fatto che esse catalizzano una reazione reversibile, termodinamicamente favorita in ambedue i sensi, quindi con  $\Delta G=0$ . La reazione viene spinta verso destra, mediante l'azione congiunta della pirofosfatasi che degrada il pirofosfato inorganico: attraverso il principio di Le Chatelier (o equilibrio mobile), infatti, diminuendo la concentrazione dei prodotti (in questo caso del pirofosfato inorganico), il sistema tenderà a consumare di più i reagenti per trasformarli in prodotti e ripristinare l'equilibrio chimico. Nella fattispecie, la diminuzione del pirofosfato inorganico spinge il sistema a consumare rNTP per liberare di nuovo il pirofosfato spingendo la reazione verso destra. In conclusione la pausa cinetica della RNA polimerasi non è dovuta ad attività di proofreading ma ad una mancata sincronia tra l'azione di polimerizzazione e quella di degradazione della pirofosfatasi.

Quinta fase:

terminazione. Il terminale 3' dell'RNA eucariotico è definito da modifiche post-trascrizionali, in pratica l'RNA trascritto viene accorciato in seguito al taglio a livello di una specifica sequenza che si trova 20-30 nucleotidi a valle del sito di poliadenilazione (AAUAAA); immediatamente dopo il taglio una specifica polimerasi aggiunge al 3' del nuovo trascritto una coda di poli-A (ca. 200 nucleotidi). Funzioni della coda poli-A:

- esportare l'mRNA maturo dal nucleo;
- influenzare la stabilità degli mRNA nel citoplasma.

