

# StuDocu.com

## Esame biologia molecolare

Biologia molecolare (Università degli Studi di Milano-Bicocca)

**TEMA D'ESAME 1:****1-UAS GAL: struttura e definire come è stato determinato.**

Il sistema UAS-GAL4 è costituito principalmente da due parti, il gene GAL4, codificante per la proteina Gal4 in lievito, e l'UAS (upstream activation sequence) una breve sequenza compresa tra le due regioni promotrici di GAL1 e GAL10. La funzione del sistema UAS-GAL4 è quella di regolare il metabolismo del galattosio nel lievito, quindi di regolare la trascrizione dei geni necessari per produrre enzimi utili all'utilizzo del galattosio come fonte energetica.

Si è determinata l'esistenza del sistema UAS-GAL4 studiando i geni strutturali relativi al metabolismo del galattosio mediante l'utilizzo di geni reporter. Tra i promotori di gal1 e gal10 (molto ravvicinati tra loro) c'è una sequenza di 365 paia di basi che è stata isolata e studiata in un plasmide posta davanti un gene reporter, contenente promotore per gene citocromo c1 e il gene di E.coli x la beta-galattosidasi (lacZ).

Il plasmide è stato inserito nelle cellule di lievito, che non sono in grado di produrre da sé la beta-galattosidasi, in presenza di molto glucosio l'attività della beta-gal misurata è molto bassa, in presenza invece del galattosio l'attività della beta-gal è 400 volte maggiore.

La sequenza responsabile di questo aumento dell'espressione genica dovuta alla presenza del galattosio, è stata denominata UAS, che è l'enhancer del lievito. Tramite continue delezioni e altrettante misurazioni dell'attività della beta gal, si è determinato che la vera sequenza consenso è piccola e misura circa 70 nucleotidi.

È costituita da 4 sequenze di 17bp, ciascuna con una simmetria interna palindromica quasi perfetta. Più sono le sequenze ripetute maggiore sarà l'induzione alla trascrizione.

Visto che il prodotto di gal4 è attivatore e il prodotto di gal80 è repressore si è ipotizzato che i prodotti di questi 2 geni interagiscano tra loro; poi in presenza di galattosio, il repressore cambia conformazione e si stacca dal complesso. In presenza di glucosio e non di galattosio, il prodotto di gal80 rimane legato a quello di gal4 e miq1 (prodotto solo in presenza di glucosio) si lega vicino l'UAS, impedendone il legame con il prodotto di gal4. In assenza sia di glucosio che di galattosio, gal80 non si stacca da gal4 che si lega con UAS-GAL, ma a causa del legame con gal80 non riesce a interagire con TATA-BOX (sequenza promotrice dei geni x galattosio) e non avviene trascrizione, in assenza di glucosio il sistema è INDOTTO, gal80 grazie a gal3, si stacca da gal4 i quali interagisce con UAS, con il promotore e recluta TFIIID che dà inizio al meccanismo di trascrizione.

**REGOLAZIONE GAL NEL LIEVITO E NEGLI EUKARIOTI**

LIEVITI: Il sistema GAL è l'analogo nel lievito dell'operone Lac batterico, poiché il lievito usa il glucosio in maniera preferenziale rispetto ad altri zuccheri; in sua assenza e presenza di altri zuccheri vengono indotti gli enzimi necessari al loro metabolismo. A differenza dei batteri il lievito non può utilizzare lattosio ma può metabolizzare il galattosio, il quale necessita di tre enzimi induttivi codificati dai geni strutturali GAL1, GAL7 e GAL10 (che normalmente non vengono prodotti) più un altro enzima (GAL2) che codifica per la Permeasi. Il sistema di regolazione consiste di altri tre geni che si trovano lontano dai geni da essi controllati: GAL3, GAL4, GAL80. GAL3 codifica per un enzima che agisce da vero e proprio induttore, GAL4 codifica per un attivatore (regolatore positivo) che serve per iniziare la trascrizione e GAL80 per un repressore (regolatore negativo) che ferma la trascrizione. In assenza di galattosio GAL80 si lega a GAL4 disattivandone il sito di legame per il DNA.

In presenza di galattosio, questo si lega a GAL80 producendo un cambio conformazionale che ne fa diminuire l'affinità per GAL4 che, libero, va a legarsi l'UAS-gal attivando l'enhancer e la trascrizione dei geni del sistema GAL. Per il buon funzionamento dell'enhancer è necessaria la presenza di una sequenza da 75bp (contenuta in una più ampia da 365bp collocata tra GAL1 e GAL10); queste 75bp sono costituite da 4 ripetizioni di una sequenza consensus quasi palindromica da 17bp. L'attivatore codificato da GAL4 ha 4 siti di legame per le UAS ciascuno dei quali corrisponde a una di queste sequenze da 17bp. GAL4 risulta sempre essere legata alle UAS da cui si distacca solo in presenza di glucosio quando GAL80 viene trascritto maggiormente e si lega ad essa.

EUCARIOTI: Negli eucarioti le sequenze regolative sono chiamate URS e si dividono in Enhancers e Silencers. Gli elementi Silencers svolgono funzione contraria agli enhancers e il loro meccanismo è generalmente di tipo competitivo con l'enhancer del proprio gene (può sia competere per il legame con il DNA, che inibire la sintesi dell'enhancer).

**Esperimento di PTashine:** consiste nel creare un plasmide che porta i primi 147aa di gal 4 fusi con tratti random di coli e un gene marcatore sottoposto al controllo di gal4. Tra le varie possibili culture si vedono quelle che esprimono il marcatore e si va ad analizzare la sequenza random che mostra un'attività trascrizionale. Si vede che i cloni attivi presentano sei ricche di aa acidi al carbox-termin. Facendo delle sequenziature i vari attivatori si nota che le parti attivanti la trascrizione sono intercambiabili. Si richiede solo l'acidità di sequenza ed esposizione mentre la parte specifica è quella che lega il dna. Per verificare se questo sistema è conservato evolutivamente è sufficiente trasfettare celli HeLa con 2 vettori uno che esprime gal 4 e uno che porta il gene reporter CAT sottoposto al controllo delle UASgal- Gal 4 induce la trascrizione anche nei mammiferi con un'elevata attività dei geni reporter.

-promotore mammifero- sv40 enhancer- ecor1 + Tratti gal4 - sv40 polimerasi A.

**2-DNA, come determinare densità, composizione in basi e peso molecolare.**

La densità del DNA è circa 1,7 g/cm. Per misurare la densità precisa del DNA si esegue esperimento mediante gradiente di cloruro di cesio.

Si utilizza il cloruro di cesio perché è inerte nei confronti del DNA e otticamente trasparente, sottoposto a radiazioni UV. Si prende CsCl 5M e lo si mette in una provetta con DNA; riesco a misurare la densità perché il DNA si sposta spontaneamente nella regione dove la soluzione ha lo stesso valore. All'inizio non c'è gradiente perché la soluzione è omogenea, ma metto la provetta in un ultracentrifuga per 48h così gli ioni cesio tendono a sedimentare formando un gradiente di cesio più denso verso il fondo e meno denso verso la parte alta della provetta. Troviamo il DNA nella sua zona di isodensità (in quanto le molecole che si trovavano in alto per forza di gravità sono state spinte in basso e quelle in basso per forza di Archimede si spinte in alto). Nella appendice avrò in fondo l'rna che ha  $\rho = 1,9 \text{ gr/cm}^3$  cubo poi il dna e sopra le proteine. Il cesio è allontanato con tampone a basso sale, riesco a visualizzarlo perché il DNA assorbe l'ultravioletto mentre il CsCl no.

Tuttavia la densità del DNA non è sempre la stessa ma dipende dalla composizione in basi, è più denso in zone ricche di C, G e meno denso in zone ricche di A, T.

DENSITA' =  $1,66 + 0,98 \cdot (\%G+C)$  espressi come frazione molare.

Per determinare la composizione in basi del DNA, contribui Chargoff con i suoi esperimenti nel 1949. Ha purificato DNA da organismi diversi, idrolizzando con acidi forti e separando con cromatografia; quantizzando le diverse basi e facendo le proporzioni si rese conto che la composizione in basi varia da un organismo all'altro. Facendo però i rapporti incrociati tra purine (adenina, guanosina) e pirimidine (timidina, citosina) si rese conto che il rapporto era circa 1 in tutti gli organismi, indipendentemente dalla specie di appartenenza.  $\text{Purine} = \text{pirimidine} = 1$ .

La composizione in basi inoltre influenza in maniera diretta la temperatura di melting, temperatura di fusione in cui la metà del DNA è denaturato. Il DNA e l'RNA assorbono la luce nel campo degli UV, hanno un picco di assorbimento a 260 nm. A determinare questo assorbimento sono i componenti cromofori, cioè le basi azotate. Quando il DNA è nella sua conformazione nativa, quindi stabile, le basi azotate sono schermate dalla doppia elica e quindi assorbono meno la radiazione rispetto ad un filamento denaturato in cui le basi sono maggiormente esposte. Misurando l'assorbimento e la temperatura possiamo studiare la denaturazione termica del DNA. Più aumenta la temperatura, maggiori sono le possibilità che si rompano i legami a idrogeno tra le basi azotate. Prima si denaturano le zone ricche di A, T (perché fanno solo 2 legami a idrogeno) poi le zone ricche di G, C.

La temperatura di melting può dunque permetterci di risalire alla composizione in basi tenendo conto della seguente relazione:  $T_m = 69,3 + 0,41 \cdot (\%G+C)$ .

Metodo di sanger: possibilità di mimare la replicazione del dna in vitro usando tre trucchi:

1) Interruzione sintesi dna dopo aggiunta dntp, chiamati nucleotidi terminator (nucleotidi senza O in posiz 3' indispensabile per il legame ad un altro nucleotide), con la conseguente creazione di lunghezza differente, in soluzione saranno presenti maggiori primer, polimerasi, nucleotidi normali e ovviamente 4 diversi dntp (a- term t-term c-term g-term) ognuno marcato con un fluorocromo differente per permettere il riconoscimento del dntp che ha causato la terminazione della sintesi. 2) separaz frammenti 3) riconoscimento ultimo dntp inserito

Per poter determinare il peso molecolare di un frammento di DNA, utilizzo la tecnica dell'elettroforesi su gel. Il gel utilizzato per elettroforesi dipende dal tipo di macromolecole in esame,

-gel di poliacrilamide è indicato per proteine o per piccoli frammenti di DNA,  
-gel di Agaroso, un polisaccaride naturale, è indicato per separare frammenti di DNA di diverse dimensioni. Quando è molto diluito e si trova in concentrazioni di 0,5 % riesce a separare frammenti anche di migliaia di basi. Quando è concentrato all'1,5% separa frammenti più piccoli, fino a 2000bp circa.

I frammenti di DNA dotati di carica negativa (gruppi fosforici) immersi in un campo elettrico, migrano all'interno del gel verso il polo positivo. La loro velocità di migrazione dipende esclusivamente dalla loro dimensione dal momento che tutti i frammenti di DNA hanno un rapporto carica/massa uniforme. Le molecole più grosse incontrano l'ostacolo della matrice porosa e migrano più lentamente delle molecole piccole.

Si lasciano correre i frammenti di DNA su gel per circa un ora, dopodiché per colorare il risultato si può usare Bromuro di Etidio (emette fluorescenza se irradiato con UV) o un altro colorante per DNA come il blu di metilene. Facendo un confronto elettroforico con le bande su gel di un DNA a lunghezza nota, cioè un marcatore di peso molecolare standard che si utilizza per tracciare la retta di taratura nel grafico con log del peso molecolare (dalton) nelle ordinate e la distanza relativa (distanza percorsa/distanza massima percorribile) nelle ascisse, è possibile risalire al peso molecolare del frammento di DNA. La velocità elettroforica è anche influenzata dalla forma della molecola, infatti una molecola superavvolta avrà, a parità di peso molecolare, una velocità maggiore di una molecola di DNA lineare.

Il peso molecolare di una molecola di dna è possibile scoprirlo con la formula  $x$  calcolare la velocità di sedimentazione  $v = PM \times F \times a$ , dove  $v$  è la velocità di sedimentazione; "F" è il fattore dipendente dalla forma della molecola di DNA ed "a" è l'accelerazione data dal campo gravitazionale in cui si trova la soluzione.

**3-RNA polimerasi e promotori batterici.**

**Biosintesi ma:** la sintesi dell'rna avviene tramite una reaz di PIROFOSFORILASI in cui l'oh 3' attacca il legame fosfodiesterico del filofosfato la reazione è spontanea e usando gli NTP come substrato il deltaG risulta fortemente negativo perché la sintesi diventa spontanea e irreversibile. L'unica cosa che serve è l'enzima ovvero l'rna polimerasi che usa dna a doppia elica come template svolgendola man mano che la sintetizza il dna libero si riavvolge automaticamente quando la ma polimerasi si allontana. Lo svolgimento crea tensioni topologiche che creano superavvolgimenti che possono essere eliminati dalle topoisomerasi per rilasciare il dna e permettere alla sintesi di procedere. L'rna pol polimerizza solo in direz 5' -> 3' e non

necessità di primer perché possiede un sito specifico per i primi due nucleotidi della seq di mma e non possiede attività enoenclaseica, quindi non può fare attività di proofreading, non fa coreazione di bozza perché la vita degli ma è breve e sarebbe una spesa energetica inutile, inoltre non produce danni gravi xk genera una prot la cui durata sarà breve mentre gli errori della dna poi diventano mutazioni permanenti che si tramettono alla prole.

Nei procarioti, la macchina trascrizionale è costituita dall'enzima multimerico RNA polimerasi, dal fattore  $\sigma$  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 conterà la subunità  $\alpha 2$  (alfa2); 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 elcosina-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, che riconosce il promotore e recluta RNAPol.

Esperimento su gradiente di saccarosio dimostra che in assenza di nucleotidi (substrato dell'enzima) se metto l'olenzima radioattivo (olenzima+core enzimatico + fattore sigma) questo sedimenta in una posizione ben precisa, quando poi aggiungo nucleotidi, permettendo alla fase di allungamento di iniziare, noto 2 picchi, uno più grande che rappresenta il complesso ternario enzima, nucleotidi, DNA, e uno più piccolo che rappresenta il fattore sigma che si è distaccato dall'enzima.

Il fattore sigma ha importanza durante la fase di inizio, ed è fondamentale nell'individuazione del promotore.

Una volta che il fattore sigma ha riconosciuto il promotore, RNA polimerasi forma col filamento di DNA interazioni più forti, questo è il complesso d'inizio.

Inizia quindi sintetizzare l'ibrido DNA-RNA lungo 9-10 paia di basi, l'evento che dà il via alla reazione deve essere la denaturazione locale del DNA a formare la cosiddetta bolla di trascrizione, finché non si verifica l'apertura della doppia elica l'RNA poi continua a sintetizzare e a degradare piccoli frammenti di 9-10 paia di basi. Sintesi abortiva. L'inizio vero e proprio della trascrizione è determinato da questo piccolo frammento di RNA-DNA che quando supera le dieci paia di basi a causa dell'ingombro sterico determina il rilascio dei fattori di inizio e quindi l'inizio della fase di allungamento.

**Promotore:** è la regione di DNA coinvolta nel legame con l'RNA polimerasi. Si trovano a monte del gene, e omettono al ma-polimerasi di riconoscere e attaccarsi per dare il via alla trascrizione del DNA.

Un primo esperimento che me ha permesso una grossolana identificazione è quello mediante DNAasi ad alta concentrazione; sapendo che in coli l'RNA pol forma, in assenza di nucleotidi, un complesso stabile con il promotore in cui in pratica il promotore è "coperto" dall'enzima, aggiungo DNAasi ad alta concentrazione. Questo enzima digerisce il DNA non coperto da proteine. L'unica frammento rimasto è evidentemente quello che era legato alla RNAPol (proteina). È un frammento di circa 40 nt detto P-DNA.

Sperimentalmente noto che se dopo il trattamento con DNA asi ad alta concentrazione provo a far riattaccare p-DNA e RNAPol non ottengo nulla perché evidentemente manca qualche fattore che è stato digerito dalle DNAasi.

Mediante foot-printing è possibile individuare delle parti parzialmente coperte che erano state precedentemente digerite.

Trattando il mio DNA con DNAasi a bassa concentrazione ottengo tra gli altri un frammento di circa 200 nt all'interno del quale sono sicuro sia contenuta la sequenza promotore.

1-Questo tratto viene marcato radiativamente con la fosfatasi alcalina (che usa come substrato ATP con p32 su un fosfato) all'estremità di ciascun filamento. Siccome serve una sola delle 2 estremità marcate per poter visualizzare il filamento su gel, una delle 2 viene trattata con enzima di restrizione specifica.

2-Tratto con DNAasi, aspettandomi tante lunghezze di frammenti quanto i nucleotidi perché questi enzimi digeriscono il filamento un nucleotide alla volta. 3-Su gel di poliacrilamide faccio un elettroforesi in condizioni denaturanti, avrà circa 198 frammenti tutti radioattivi.

L'esperimento va ripetuto anche con un DNA a cui è legata RNAPol. Si noterà che nella banda in cui c'è anche l'enzima mancano dei frammenti che non sono stati digeriti perché protetti dall'enzima. Il frammento di DNA protetto è di circa 60 nt.

Con nuove tecniche è stata trovata la sequenza **promotore nei procarioti**, e si è visto che esso è bipartito:

c'è una regione -10 (dall'inizio della trascrizione) TATAAT e una regione -35 TTGACA separate da una regione non conservata. Ogni promotore può poi avere una sequenza diversa da questa ideale, più la sequenza somiglia all'ideale maggiore sarà l'efficienza di trascrizione. Mutazioni che distruggono queste sequenze consensus vengono dette mutazioni DOWN perché abbassano la trascrizione rendendo + debole il promotore quelle invece che alzano la potenza del promotore sono piuttosto rare ma prendono il nome di mutazioni up e aumentano quindi la trascrizione dei geni adiacenti.

**Promotori Eucarioti:** si dividono in 3 classi: classe I, II, III. In base alla ma poi che li riconosce. **Quelli di classe I** = è un unico promotore in quanto quasi tutte le specie hanno un solo tipo di gene riconosciuto dalla pol I: il gene precursore dell'rNa, e una sequenza molto variabile da specie a specie di solito l'unica cosa che viene mantenuta è un iniziatore ricco di AT che circonda il sito di inizio.

**Quelli di classe II** = sono formati da due parti: il **promotore centrale** e un **elemento a monte**. Il promotore centrale può contenere sino a 4 elementi: una TATABOX di solito a 25pb la cui sequenza consenso è TATAAA che è implicata per la determinazione del punto di inizio infatti è sulla tatabox che avviene il legame con la TF2D che contiene la tataBUILDING protein, un **elemento di riconoscimento** per il TF2B che si lega al promotore con la TF2D e alla ma poi che sono posti a monte della tatabox che aiuta la TF2B a legarsi al dna, **degli iniziatori** ovvero sequenze conservate intorno ai siti di inizio della trascrizione ed **altine alcuni elementi promotori a valle chiamati** DPE che compensano la mancanza delle tata box e sono di solito accoppiate a un INR. Gli elementi a monte prevedono la presenza di GC BOX mutazioni in questi siti pregiudicano la potenza della trascrizione.

**Quelli di classe III** = hanno generalmente promotori interni ovvero localizzati internamente al gene. I geni per l'rna 5s presentano un BOX A un piccolo intermedio e un BOX C mentre per i geni dei tRNA sono presenti i boxA e i BOX B

**TERMINATORI:** sono difficili da studiare in quanto non legano l'RNAPol ma tramite comparazioni mRNA-DNA è stato possibile farlo. È stata scoperta una sequenza palindromica ripetuta e invertita poche basi prima della fine dell'mRNA. Essa è una sequenza di 10-15 pb con una zona intermedia variabile che si pensa possa dare origine a una struttura cruciforme. Quest'ultima nell'RNA, emerge spontaneamente, e il terminatore potrebbe dunque essere rappresentato dai formari di questo genere di struttura secondaria nell'mRNA neosintetizzato. Quando l'RNAPol legge la sequenza palindromica avviene un fenomeno di pausa momentanea (pausing) dopo il quale ricomincia la processione veloce. Esistono 2 classi di terminatori: la prima in cui, dopo le sequenze palindromiche, vi è una serie di U ripetute che creano dei legami molto deboli così da rendere spontaneo il distacco dell'RNA. La seconda presenta invece sequenze normali e, impedendo il distacco spontaneo dell'RNA, necessita di una proteina specifica p per il distacco. Queste classi si chiamano terminatori p indipendenti e p dipendenti. La proteina p forma complessi con l'RNA a singola elica, si muove su di esso tramite ATP denaturando le strutture secondarie. Essa raggiunge l'RNAPol quando quest'ultima è in fase di pausing; una volta raggiunto causa il distacco dell'RNA e la fine della trascrizione.

**CONTROLLO TRASCRIZIONALE**

Nei batteri i geni sono raggruppati in più unità costituite da un solo promotore, uno o più terminatori e geni che devono essere espressi in maniera coordinata; questi insieme di geni vengono detti Operoni. Il primo operone studiato fu l'Operone Lac che codifica per  $\beta$ -galattosidasi, galattodidasi permasi e acetilasi; esso è sottoposto a due principali sistemi di regolazione: gli attivatori e i repressori. In un sistema di controllo negativo inducibile entrano in gioco il repressore (proteina che si lega all'operatore) e l'induttore che si lega al repressore: l'affinità del repressore per l'operatore viene ridotta dal legame con l'induttore di circa 1000 volte; con l'intervento di altri fenomeni ausiliari e con l'aumento della concentrazione di induttore il repressore si stacca dall'operatore in quanto aumenta la sua affinità aspecifica, grazie al quale esso può legarsi ai numerosissimi tratti dove si trova il DNA. Infatti, quanto più cresce l'affinità aspecifica, tanto più facilmente il repressore si stacca dall'operatore quando è legato all'induttore. Il repressore funziona in forma di tetramero; due subunità restano libere e due legano il DNA riconoscendo le stesse sequenze palindromiche che costituiscono l'operatore. L'estremità amino terminale del repressore presenta tre c-eliche delle quali la seconda lega il DNA sul solco maggiore mentre la prima copre e stabilizza il legame che induce una curvatura del DNA provocando una sovrapposizione dei siti di riconoscimento, rendendo così improbabile il legame dell'RNA polimerasi al DNA se vi è già legato il repressore. Tuttavia l'RNA polimerasi si può legare anche in presenza del repressore perché le sequenze -35 e -10 sono libere; il repressore agisce quindi come un freno posto a valle della RNAPol, impedendo la polimerizzazione. Alcuni sistemi di controllo positivo dipendono dalla presenza di glucosio. L'operone lac è attivato dalla proteina CAP, codificata dal gene CRP e attivata a sua volta da cAMP. In presenza di glucosio il cAMP è represso, non attiva la proteina CAP e conseguentemente tutto il sistema è represso. In queste condizioni, anche in presenza di lattosio, il sistema è indotto ma comunque represso. Quando il glucosio non è presente ma vi è lattosio, la proteina CAP attiva il sistema e permette la trascrizione.

**La RNA polimerasi II e i promotori di classe II**

La RNA polimerasi di classe II è localizzata nel nucleoplasma e svolge la maggior parte dell'attività sintetica, in quanto responsabile della sintesi dell'RNA eterogeneo nucleare (hnRNA), il precursore dell'mRNA. I promotori riconosciuti da questo enzima presentano generalmente una serie di strutture e di sequenze caratteristiche:

1. Un "core", o promotore minimo, cioè la più breve sequenza di DNA necessaria e sufficiente per un accurato inizio della trascrizione da parte della RNA polimerasi. Gli elementi del promotore minimo meglio caratterizzati sono la TATA-box (sequenza consenso TATAaA/aat) in genere localizzata a circa -25 paia di basi dal sito di inizio della trascrizione e una sequenza ricca di pirimidine (consenso YYAaiaYY) detta iniziatore (Inr) localizzata sempre in prossimità del sito di inizio della trascrizione. Non tutti i promotori possiedono la TATA-box e l'iniziatore ed esistono altresì promotori che diftano di entrambi; per questi promotori è solitamente coinvolta la regione 5' non tradotta (5'UTR) del gene che essi regolano affinché la trascrizione sia efficiente. Tuttavia i meccanismi che regolano un corretto posizionamento dell'apparato trascrizionale a livello del sito di inizio della trascrizione sul DNA è ancora ampiamente dibattuto. E' stata recentemente identificata una proteina, chiamata core-promoter binding protein (CPBP), dotata di tre zinc fingers del tipo C2-H2 (due cisteine e due istidine che coordinano l'atomo centrale di zinco) che sembra partecipare al mantenimento dell'espressione basale di geni privi di TATA-box (Kortischoner et al, 1997).

2. Elementi prossimali localizzati tra le 50 e le 200 paia di basi dal sito di inizio della trascrizione; questi elementi vengono riconosciuti e legati da attivatori trascrizionali, generalmente ubiquitari, con la funzione di stabilizzare l'apparato di trascrizione. I meglio caratterizzati sono la CCAAT-box (generalmente localizzata a -80/-100 paia di basi dal sito di inizio) e la GC-box (consenso GGCGCG, generalmente a -40/-70 paia di basi dal sito di inizio), che sono presenti in molti promotori e possono risultare attivi in entrambi gli orientamenti.

3. Elementi distali detti "enhancer" o "silencer" a seconda che svolgano funzioni attivatorie o inibitorie sull'espressione genica. Gli elementi distali possono trovarsi anche a notevole distanza dal sito di inizio della trascrizione (anche migliaia di basi); sono in grado di funzionare indipendentemente dall'orientamento, posizione e distanza e contengono sequenze riconosciute da fattori trascrizionali tessuto-specifici e/o inducibili. Gli enhancer ed i silencer possono agire sul promotore minimo, aumentandone o inibendone l'attività attraverso interazioni con il complesso basale della trascrizione, oppure possono modificare la struttura della cromatina, così da facilitare a sfavore l'accesso del complesso di inizio sul DNA. I fattori di trascrizione.

I fattori di trascrizione vengono classificati in base all'interazione con uno dei tre diversi tipi di moduli descritti in precedenza:

- Il primo gruppo è costituito dai "general transcription factors" (GTFs), richiesti per il reclutamento ed il corretto posizionamento della RNA polimerasi al sito di inizio della trascrizione e per la sintesi dell'RNA. I GTFs formano un complesso multiproteico con la polimerasi, che costituisce l'apparato trascrizionale di base.

- Gli "upstream factors", che riconoscono gli elementi prossimali attraverso un legame sequenza-specifico al DNA; la loro attività non è regolata ed essi agiscono su qualsiasi promotore che contenga la sequenza consenso, accrescendo l'efficienza di trascrizione e garantendo un adeguato livello funzionale del promotore mediante la stabilizzazione del complesso di pre-inizio.

- L'ultimo gruppo è rappresentato dai fattori inducibili da stimoli extracellulari e/o tessuto-specifici, che si differenziano dalla classe precedente grazie al loro ruolo regolatorio; essi, infatti, vengono sintetizzati o attivati in un determinato momento del ciclo cellulare, in uno specifico tessuto, o in funzione ad attivatori molecolari, controllando finemente l'espressione del gene bersaglio.

Un promotore che contiene solo elementi di riconoscimento appartenenti alle prime due classi di fattori è attivo in tutti i tipi cellulari, ed è responsabile della trascrizione costitutiva di geni ad espressione ubiquitaria (housekeeping); tali geni codificano in molti casi per proteine strutturali della cellula. L'espressione di proteine a carattere regolatorio risente dell'induzione da parte di fattori trascrizionali inducibili. L'apparato di base per la trascrizione.

L'inizio della trascrizione richiede un'ordinata e definita associazione tra diversi

GTFs e la RNA polimerasi II, così da costituire un complesso multiproteico. Tutti i fattori associati all'azione della RNA polimerasi II prendono il nome di transcription factors II (TFII); fino ad oggi sono noti otto fattori che partecipano alla formazione dell'apparato trascrizionale di base (TFII A, TFII B, TFII D, TFII E, TFII F, TFII H, TFII J).

Soltanto il primo evento nella formazione del complesso è il legame alla TATA-box di TFII D, il quale è a sua volta un complesso multiproteico formato dalla TATA-box binding protein (TBP) e da numerosi altri fattori detti TATA-box binding protein associated factors (TAFs). Esso copre una regione di DNA che si estende da circa -42 a -17 paia di basi dal sito di trascrizione (+1).

Gli altri GTFs vengono reclutati nell'ordine seguente: TFII A e TFII B, che stabiliscono interazioni con TFII D e con sequenze a monte ed a valle della TATA-box; TFII F, A e C che è associata la RNA polimerasi II; TFII E, TFII H, TFII J (figura 3). Questi ultimi posseggono attività elicastica, in grado di determinare l'apertura e lo srotolamento della doppia elica di DNA da sequenze. Mutazioni nella TATA-box determinano una stabilizzazione del complesso di pre-inizio, situazione che comporta una trascrizione inefficiente; questa constatazione pone interrogativi riguardo la trascrizione di geni che sono sprovvisti totalmente di TATA-box, affinché la trascrizione di questi geni sia efficiente si ritiene indispensabile il ruolo regolatorio di specifiche sequenze contenute nella regione 5'UTR.

Il complesso di pre-inizio interagisce sia con attivatori che repressori trascrizionali, responsabili di un'azione rispettivamente stabilizzatrice e destabilizzatrice e quindi in grado di regolare finemente i livelli di trascrizione del gene bersaglio.

#### Promotori RNA POL III

I promotori dei geni dell'RNA ribosomale 5S e del tRNA sono interni

- per il gene dell'RNA ribosomale 5S il promotore si estende da +55 a +80.

- Questo è stato determinato mediante analisi di delezione.

- Successive analisi mutazionali hanno identificato due sottotipi

- tipo 1, contenente box A e box C;

- tipo 2, contenente box A e box B.

- Il promotore del piccolo RNA nucleare U6

ha una struttura più canonica

- tipo 3, contenente tre elementi conservati, Oct,

PSE e TATA nella regione a monte.

Il promotore 1 della RNA polimerasi III

- TFIIIA si lega per primo

- membro della classe di proteine a "dito di Zinco"

- TFIIIA aiuta TFIIIC a legare box C.

- La contemporanea presenza di TFIIIA e TFIIIC permette a TFIIIB di legarsi al sito di inizio.

- TFIIIB recluta la RNA polimerasi III.

- Una volta che TFIIIB si è legato, TFIIIA e TFIIIC possono anche essere rimossi

- fattori di assemblaggio

- TFIIIB resta legato al sito di inizio e la sua presenza è sufficiente per permettere alla polimerasi III di legarsi.

- I promotori di tipo 1 si trovano solo nei geni per l'RNA ribosomale 5S.

Gli promotori 2 e 3 della RNA polimerasi III

- TFIIIC lega sia box A che box B

- TFIIIC è un complesso di 6 subunità, di dimensioni (> 500 kD) comparabili alla RNA polimerasi III

- Questo permette a TFIIIB di legarsi al sito di inizio e, in seguito, di reclutare la polimerasi

- Il fattore di posizionamento TFIIIB è formato da 3 subunità:

- TBP, presente anche in TFIIID (vedi dopo);

- Bfr, omologo di TFIIIB (vedi dopo);

- B' contribuisce alla bolla di trascrizione

I promotori di tipo 3 assemblano un complesso di pre-inizio che recluta la pol III.

L'elemento di sequenza TATA (l'unico essenziale) serve per il posizionamento della polimerasi III.

#### SEQUENZE UAS E ENHANCER

Le sequenze UAS e le sequenze enhancer sono entrambe sequenze il cui ruolo è quello di aumentare l'efficienza della trascrizione.

Le UAS sono costituite da 50-100 bp e si trovano negli eucarioti semplici, collocate obbligatoriamente a monte del promotore. Se ne possono trovare più copie sullo stesso gene.

Le enhancer sono costituite da circa 200 bp e si trovano negli organismi eucarioti superiori, non obbligatoriamente collocate a monte del promotore. Possiedono tre proprietà fondamentali: agiscono a distanze relativamente grandi e la loro espressione è indipendente sia dall'orientamento che dalla posizione. Inoltre contengono, a volte, sequenze ripetute che hanno parziali attività regolative (ad es. l'enhancer del virus SV40 possiede due ripetizioni dirette di 72bp, ognuna delle quali è sufficiente all'attività regolativa).

Le enhancer possono funzionare in modo indipendente dalla distanza grazie ad una particolare sequenza chiamata MPE (Main Promoter Element) che, sfruttando la flessibilità del DNA, ripiega l'enhancer sul promotore assemblando i complessi di inizio della sintesi.

Infine, sulle sequenze enhancer si legano proteine chiamate AP 1, 2, 3, 4.

Per inattivare l'enhancer è necessario introdurre mutazioni contemporanee in più siti, poiché queste sequenze sono molto ricche di siti di legame con le proteine.

#### 4-fattori di allungamento e reazione.

**Reazione di allungamento** : il primo aminoacil-RNA viene portato nella posizione corretta, sul sito P, da eIF2 legato a GTP. Il movimento della subunità minore sull'mRNA è facilitato dall'idrolisi di ATP in ADP+P. Il RNA con l'anticodone

complementare alla seconda tripletta sull'mRNA, legato al fattore di allungamento EF-Tu, si attacca nel sito A; il fattore di allungamento esegue anche un controllo di qualità per verificare il corretto appaiamento codone-anticodone.

Una volta riconosciuto l'appaiamento corretto, l'rRNA sulla subunità minore idrolizza il GTP legato a EF-Tu e si stacca dal tRNA.

A questo punto la peptidil transferasi contenuta nella subunità maggiore del ribosoma, catalizza la formazione del legame tra i due aminoacidi presenti sui tRNA nei siti P ed A, con la formazione di un peptidil-RNA allungato nel sito A.

A questo punto entra in gioco un'altra proteina G ausiliaria, una translocasi, EF-G codificata dal gene FUS, è molto simile come struttura al complesso ternario EF-Tu/aa-tRNA/GTP perché deve poter entrare nel ribosoma in corrispondenza del sito A, è una proteina che ha attività GTPasica, l'idrolisi del GTP a GDP, determina una variazione conformazionale in EF-G, il che gli permette di insinuarsi nel sito A e di causare TRASLOCAZIONE:

- spostamento di tRNA occupato da sito A a sito P.

- spostamento di tRNA scarico da sito P a sito E.

-traslazione dell'mRNA di 3pb.

EF-TU serve a far entrare nel ribosoma l'aminoacil tRNA, complementare con il codone presente nel ribosoma, a cui è legato assieme all'GTP, formando un complesso ternario (EF-TU/GTP/aa-tRNA). Questo fattore di allungamento funziona solo quando è legato a GTP, nel momento in cui l'aminoacil tRNA viene posizionato in maniera corretta sul ribosoma, una nuova idrolisi del GTP a GDP determina il distacco di EF-TU dal complesso, e fa sì che possa avvenire la formazione del legame peptidico.

Il sito A è occupato dalla catena nascente + tRNA occupato, mentre sul sito P si trova il tRNA scarico.

EF-TS ha la funzione di liberare EF-TU dal GDP, liberando il sito x un nuovo GTP, così da poter iniziare un nuovo ciclo. A questo punto entra in gioco un'altra proteina G ausiliaria, una translocasi, EF-G codificata dal gene FUS, è molto simile come struttura al complesso ternario EF-Tu/aa-tRNA/GTP perché deve poter entrare nel ribosoma in corrispondenza del sito A, è una proteina che ha attività GTPasica, l'idrolisi del GTP a GDP, determina una variazione conformazionale in EF-G, il che gli permette di insinuarsi nel sito A e di causare:

- spostamento di tRNA occupato da sito A a sito P.

- spostamento di tRNA scarico da sito P a sito E.

- traslazione dell'mRNA di 3pb.

EF-TU ha 3 domini, organizzati in beta foglietto ed alfa eliche, il legame a GDP o con GTP determina nel fattore di allungamento dei cambiamenti conformazionali che la espongono o meno al sito di legame con il tRNA. L'attività GTPasica è stimolata da interazioni con il ribosoma.

**Aminoacil tRNA sintetasi**: enzima che catalizza l'esterificazione di uno specifico aminoacido ad uno dei possibili tRNA corrispondenti a formare un aminoacil-tRNA, ne esistono 20 specifico uno x ogni aa. L'attivazione degli aminoacidi da parte della sintetasi richiede due reazioni distinte, la prima è la reazione dell'aa con una molecola di ATP, che porta alla formazione di un aminoacido adenilato e di un pirofosfato la cui eliminazione sposta l'equilibrio della reazione verso i prodotti.

La seconda reazione è il legame dell'aa adenilato all'estremità 3' del tRNA appropriato dove ho un gruppo OH legato a una sequenza ACC.

#### REAZIONE DI TERMINAZIONE

L'allungamento si ripete più volte finché non si incontra un codone di stop che segnala il momento di terminare la traduzione. I codoni di stop sono 3: UAA UAG UGA. Quando uno di questi tre codoni arriva nel sito A del ribosoma si ha normalmente che nessun tRNA lo riconosce. Il riconoscimento avviene a carico di proteine chiamate FATTORI DI RILASCIO (RF).

RF1 riconosce i codoni di stop UAA eUAG

RF2 riconosce UAA e UAG

RF3 è una proteina che lega GTP facilitando il legame di RF1 e RF2 al ribosoma.

Gli eucarioti hanno invece due fattori di rilascio:

eRF1 che riconosce tutt'e tre i codoni di terminazione

eRF2 che è un GTPasi ribosoma dipendente che aiuta eRF1 nel rilascio del polipeptide completo.

Dopo la terminazione è necessario che i ribosomi rilascino l'mRNA e tale rilascio avviene grazie ad un fattore chiamato FATTORE DI RILASCIO RIBOSOMIALE (RRF). Il meccanismo di rilascio ribosomiale è tutt'ora sconosciuto.

#### 5-esperimento su quanto cambia il passo dell'elica con Et-Br.

Il Bromuro di Etilio è un intercalante, cioè una molecola che si infila tra le coppie di basi del DNA allontanandole e diminuendo di conseguenza l'elical twist (angolo di rotazione tra le coppie di basi) e alterando la struttura secondaria perché varia il passo dell'elica. Per fare un giro completo infatti servono più basi.

Quando lo ione etidio si intercala fra le due coppie di basi, causa al DNA uno srotolamento della doppia elica di 26 gradi, passa da 36 gradi (DNA fisiologico) a circa 10 gradi.

In genere gli intercalanti sono costruiti da 3 anelli aromatici, struttura simile a quella delle basi azotate planari e da una carica + delocalizzata su N, fanno interazioni elettrostatiche con la carica - portata dal DNA.

Il numero max di molecole che posso intercalare è di 1 ogni 2 pb in quello circolare (si creano superavvolgimenti) e di 1 ogni 1 pb in quello lineare.

Se l'intercalazione avviene all'interno di DNA lineare che presenta estremità libere di ruotare non si creano tensione e il filamento di DNA si srotola spontaneamente; nel DNA circolare è chiuso. Il linking number non cambia perché non rompo nessun legame, ma il twist diminuisce di 26 gradi per ogni molecola di Etilio intercalata nel DNA. Poiché Lk=Tw+Wr la diminuzione del Tw è controbilanciata da un aumento del Wr, in altre parole l'aggiunta dell'intercalante porta a un DNA più rilassato. Se la concentrazione di Etilio aumenta, il numero dei superavvolgimenti potrà raggiungere lo zero; aumentando ancora la quantità di Etilio, Wr diventerà maggiore di zero e il DNA superavvolto positivamente.

#### Gli Intercalanti:

Gli agenti intercalanti sono sostanze coloranti, a volte antitumorali, che si inseriscono tra le coppie di basi mimandone la struttura mediante interazioni di van der Waals. Sono formati da tre anelli condensati e presentano cariche positive.

L'intercalazione avviene in due tempi: si stabilisce un complesso salino fra i fosfati e l'intercalante e poi si completa l'inserimento.

Gli agenti intercalanti irrigidiscono il DNA e ciò causa quindi un aumento (proporzionale al grado di intercalazione) della Viscosità e una diminuzione della Velocità di sedimentazione, ciò vuol dire che si ha anche una diminuzione della densità. La presenza degli intercalanti inibisce l'attività di alcuni enzimi e da ciò ne deriva la tossicità; essi inoltre possono anche portare a saturazione, cioè quando ve ne è uno per ogni coppia di basi. Gli agenti intercalanti inoltre variano l'angolo di Twist producendo un avvolgimento meno stretto del successivo svolgersi dell'elica; il numero di basi per giro passa da 10 a 12 aumentando quindi il passo dell'elica.

L'intercalazione è quindi una variazione strutturale che colpisce sia la struttura secondaria che quelle superiori. All'aumentare degli agenti intercalanti si ha un accumulo di tensioni che in un DNA lineare vengono scaricate alle estremità, mentre in un DNA circolare possono essere eliminate solo da superavvolgimenti. All'aumentare dell'intercalazione il superavvolgimento si espanderà e non è più possibile aggiungerne altri. Se misuriamo il passo dell'elica è possibile determinare il grado di intercalazione.

Un famoso agente intercalante è il bromuro di etidio. Esso è un colorante (arancione) che entrando nel DNA ha l'effetto di diminuire la densità. Il bromuro di etidio varia la sua attività quando è legato al DNA ed è di interesse per l'uso di laboratorio per distinguere il DNA lineare da quello circolare. Il DNA lineare, che presenta un maggior grado di intercalazione, è individuabile quindi da una maggiore fluorescenza rispetto a quella del DNA circolare. Altri tipi di intercalanti sono i coloranti ad acridina, ioduro di propidio e coloranti specifici per il DNA che interagiscono con AT.

L'inserzione di molecole intercalanti riduce l'angolo di rotazione della porzione di DNA in cui avviene da 36 a 10°; l'intercalazione è favorita termodinamicamente in quanto si formano più legami che con le basi normali.

L'intercalazione ha anche una certa importanza sia dal punto di vista biologico che tecnologico, in quanto può essere utilizzata per modificare le basi. La reazione DNA+ dimetilsolfato da mutazione specifica delle purine. Dal punto di vista biologico gli agenti mutageni sono mutageni: la base mutata diviene più facilmente idrolizzabile e si perde facilmente, indebolendo il legame fosfodiesterico e favorendo la possibile rottura del filamento. Dal punto di vista tecnologico è possibile utilizzare, ad esempio, agenti alchilanti per tagliare il DNA a livello delle G: proprio su questo tipo di tecniche si basa il sequenziamento del DNA. Un secondo trattamento, coinvolge l'etilnitrosourea, che attacca direttamente il legame fosfodiesterico favorendo la rottura della catena.

#### TEMA D'ESAME 2:

##### 1-disegnare ribosio, deossiribosio, doppio filamento DNA

##### 2-replisoma + frammenti di okazaki

Si definisce replisoma il complesso di tutte le proteine che lavorano durante la replicazione del DNA di livello della forza replicativa, appartenendo ad esso le elicas, enzimi che aprono la doppia elica del DNA rompendo legami a idrogeno tra le basi azotate, le topoisomerasi che risolvono le tensioni a seguito dell'apertura del doppio filamento, le SSB (SINGLE STRAND BINDING PROTEIN) che legano il DNA a singolo filamento e lo stabilizzano, le primasi che inseriscono i primer ad RNA, le ligasi che saldano i nick tra frammenti appaiati.

La polimerizzazione del nuovo filamento avviene sempre in direzione 5' → 3', quindi i filamenti presenti nella forcella non potranno essere replicati allo stesso modo perché sono anti-paralleli.

Il filamento in cui la direzione di apertura della forcella è opposta alla direzione di polimerizzazione della DNApolimerasi è detto filamento LAGGING, la replicazione su questo filamento è discontinua, il DNA si replica in tanto piccoli frammenti, i frammenti di Okazaki che in un secondo momento saranno uniti dalla DNAligasi a formare un unico filamento. Ogni frammento di Okazaki ha il suo primer ad RNA che sarà successivamente rimosso. Scoperti negli anni 40 dopo essere marcati con T<sup>3</sup> per tempi brevissimi, i frammenti sono di 1000-2000 pb entrambi i filamenti leggeri e lagging si allungano in maniera frammentaria ma in realtà c'è anche un altro meccanismo che genera frammenti ovvero l'inserimento di uracile errato che viene tolto dall'uracil-n-glicosidasi.

Un primo esperimento che mostra l'esistenza dei frammenti di Okazaki, è quello in cui il DNA di mutanti lig7<sup>-</sup> (che hanno attività ligasica ridotta così sono sicuro di vedere i frammenti) viene marcato per breve tempo (pulse) con timidina tritizzata. Estraggo il DNA dalle cellule uccise e separo i vari frammenti sul gradiente di saccarosio, misurando la radioattività mi rendo conto che ho 2 picchi (rappresentanti uno il filamento leading, più pesante e sul fondo della provetta, e l'altro meno pesante rappresentante i frammenti di Okazaki). La distribuzione della radioattività è però disomogenea e va contro le aspettative, si nota una maggior radioattività nel filamento più leggero; questo riflette una maggior presenza di DNA frammentato. Infatti può capitare che per errore venga inserito nel frammento leading del DNA d'UTP al posto di dTTP. Interviene l'enzima uracil N-glicosilasi rievla la presenza dell'errore e rimuove la base sbagliata lasciando un gap e quindi una rottura nel DNA.

In organismi che possiedono la uracil N-glicosilasi non funzionante infatti la distribuzione della radioattività è uguale sul filamento pesante e sui filamenti leggeri.

Un altro esperimento prova la presenza di primer ad RNA sui frammenti di Okazaki. Primer lunghi 6/10 nucleotidi e gli enz sono le ma polimerasi quando trovano dei promotori inseriscono i nucleotidi dopo aver tagliato l'elica, quando c'è 3'oh libera può iniziare a funz la dnapol. L'rna pol è fondamentale per aprire le forche e iniziare la sintesi ma non per mantenere aperte quelle per esistenti. Alla fine del frammento di oka rimane un nick da chiudere e un primer di ma da rimuovere che viene rimosso dalla dna pol I tramite il nick traslation e viene poi chiuso il nick dalla dna ligasi.

Utilizzo mutanti poli A- (il gene poli A codifica x la DNApoli I enzima che si occupa della rimozione dei primer) e l'enzima enzoonucleasi di milza, in grado di idrolizzare acidi nucleici solo a partire da un'estremità 5' libera.

Tratto i frammenti di Okazaki con NaOH, (solo l'RNA è suscettibile all'idrolisi in ambiente alcalino) e successivamente lo tratto con l'esucleasi di milza, ottengo l'idrolisi totale. Se invece tratto i frammenti di Okazaki direttamente con l'esucleasi di milza non ottengo nulla. Ho così dimostrato la presenza dei primer ad RNA sui frammenti di Okazaki.

### 3-RNAPolimerasi III + fattori di trascrizione 3

L'enzima RNAPolimerasi 3 è un enzima relativo alla trascrizione negli eucarioti per i geni codificanti tRNA e rRNA 5S piccola componente del ribosoma. Per poter svolgere la sua funzione l'RNA polimerasi 3 necessita di **3 fattori** di trascrizione detti TF3A (transcription factor A) TF3B, TF3C.

Sono stati identificati i promotori della RNApol3 (gli unici che si mantengono grossomodo conservati) utilizzando le uova di xenopus, all'interno delle quali sono stati iniettati un plasmide contenente DNA noto + nucleotidi marcati se avviene la trascrizione significa che l'RNApol3 riconosce il promotore, tramite delle delezioni e cambiamenti di basi nella sequenza del DNA noto, si è risaliti alla sequenza promotore. Grazie a questa tecnica si è riuscito a determinare che sono dei promotori interni alla regione codificante.

I promotori per l'RNA sono 2 regioni abbastanza conservate dette BOXA e BOXB di 10 nt ciascuna, il TF3C si lega a queste due regioni, questo legame permette il reclutamento del TF3B e successivamente della RNAPolimerasi3 che dà inizio alla trascrizione. Per la trascrizione del tma sono suff TF3B e C mentre per l'5s è necessario l'intervento di TF3A, per creare un tma: tfc è il primo fattore a legarsi (esso è un grosso complesso formato da 6 polipeptidi diversi che lega contemporaneamente box a e b, poi si lega tfo-b composto da TBP BRE e BA. T3b porta il sito di attacco per la pol3 la quale si lega per effetto di interazione proteo-prot con tfc3)

Nei geni x5sRNA l'unico promotore è il BOXB che viene riconosciuto da TF3A, che recluta il TF3C, che lega il TF3B, questo legame richiama l'RNA pol 3 che dà inizio alla trascrizione.

Il TF3A è una piccola metallo-proteina (lega solo il core del BOX C), che contiene dai 7-10 atomi di Zinco che sono fondamentali nel suo legame col DNA. La proteina contiene 9 sequenze ripetute ricche di Cys e His, che coordinano 1 atomo di zinco. Questo tipo di struttura è detto zinc-finger, ed è fondamentale nel legame con il DNA.

### 4-spliceosoma e hmrRNA

#### SPlicing e SPOICEOSOMI

La discrepanza fra l'organizzazione interrotta del gene e l'organizzazione non interrotta del suo mRNA richiede una modificazione del trascritto primario, che ha la stessa organizzazione del gene ed è talvolta chiamato pre-mRNA. La rimozione degli introni da questo precursore lascia un tipico messaggero di circa 2,2 kb, il processo per cui gli introni sono rimossi si chiama splicing dell'RNA, ed è una parte importante della produzione di RNA in tutti gli eucarioti. Uno dei primi indizi sulla natura della discrepanza negli eucarioti superiori è stato fornito dalle proprietà dell'RNA nucleare, che ha dimensioni medie molto maggiori dell'mRNA, è instabile ed ha una complessità di sequenza molto più grande: è stato chiamato hnRNA (nucleare eterogeneo) e comprende pre-mRNA e altri trascritti. La sua forma fisica è di una particella ribonucleoproteica, hnRNP; le proteine più abbondanti sono quelle del core della particella, ma ce ne sono altre presenti in stechiometria minore per un totale di circa 20; alcune possono avere un ruolo strutturale, altre vanno avanti e indietro tra nucleo e citoplasma e hanno un ruolo nell'esportazione dell'RNA o nel controllo delle sue attività.

Lo splicing avviene nel nucleo, insieme alle altre modificazioni che vengono apportate agli RNA appena sintetizzati. Possiamo identificare diversi tipi di splicing:

- gli introni sono rimossi dai pre-mRNA nucleari degli eucarioti superiori da un sistema che riconosce soltanto brevi sequenze consenso conservate ai confini esone-introne e all'interno dell'introne. Questa reazione richiede un grande apparato di splicing, che prende la forma di una serie di proteine e di ribonucleoproteine che funzionano come un grande complesso particolato: lo spliceosoma. Il meccanismo di splicing comporta transesterificazioni e il centro catalitico comprende RNA oltre che proteine.
- Certi RNA hanno la capacità di eliminare i loro introni autonomamente. Gli introni di questo tipo si dividono in due gruppi, distinti dalla struttura secondaria e terziaria; entrambi i gruppi usano reazioni di transesterificazione in cui l'RNA è l'agente catalitico.
- La rimozione degli introni dai precursori nucleari dell'RNA del lievito coinvolge attività enzimatiche che agiscono sul substrato come fanno gli enzimi che modificano l'RNA: queste reazioni sono eseguite da enzimi che eseguono taglio e legatura.

I siti di splicing sono i due confini esone-introne che comprendono tutti i siti di rottura e riunione. Fra le due estremità di un introne non c'è estesa omologia o complementarità, ma le giunzioni possiedono sequenze consenso brevi e ben conservate; si parla di DGA GT AG (inizio e fine): ciò riflette il fatto che le sequenze sono state analizzate originariamente nel DNA ma naturalmente GT nella sequenza del filamento codificante del DNA diventa GU nell'RNA. I due siti hanno anche sequenze diverse, quindi definiscono direzionalmente le estremità dell'introne, e sono chiamate sito di splicing 5' (o donatore) e sito di splicing 3' (o accettore). I siti di splicing sono generici, cioè non hanno alcuna specificità per singoli RNA precursori, e questi ultimi non portano informazioni specifiche necessarie per lo splicing. Inoltre l'apparato di splicing non è tessuto-specifico: un RNA può solitamente essere sottoposto a splicing in modo appropriato in qualunque cellula, anche se non è stato da essa sintetizzato.

In linea di principio quindi qualunque sito di splicing 5' può essere capace di reagire con qualsiasi sito di splicing 3', anche se in circostanze normali esso avviene tra siti di splicing dello stesso introne.

Lo splicing segue una via definita, ma non sembra esservi una via unica, in quanto si possono trovare intermedi in cui sono state rimosse combinazioni diverse di introni, il che testimonia l'esistenza di una o più vie preferenziali. Generalmente: quando è andato perduto soltanto un introne si tratta praticamente sempre del 5' o del 6, che possono anche essere persi entrambi per primi; quando sono stati persi due introni, di nuovo il 5 e il 6 sono i più frequenti, ma ci sono anche altre combinazioni; l'introne 3 non viene mai perso, se non raramente, nei primi tre passaggi; la via preferenziale è 5/6 7/4 2/1 3. La conformazione dell'RNA influenza l'accessibilità ai siti di splicing; quando vengono rimossi introni particolari, la conformazione cambia e nuove coppie di siti di splicing diventano disponibili; la capacità del precursore di rimuovere i suoi introni in più di un ordine indica che ad ogni stadio sono disponibili conformazioni alternative (la reazione non procede sequenzialmente lungo il precursore).

#### Gli snRNA nello splicing

Per **small nuclear RNA** (piccolo RNA nucleare, abbreviato come **snRNA**) si intende una piccola molecola di acido ribonucleico, solitamente molto ricca di uracile, che partecipa alla maturazione dell'mRNA. La localizzazione degli snRNA è tipicamente nel nucleo eucariote.

Gli snRNA, trascritti dalla **RNA polimerasi I** o dalla **RNA polimerasi III**, sono coinvolti in una serie di importanti processi come lo **splicing** (la rimozione degli introni dai pre-mRNA), la regolazione dei **fattori di trascrizione** (7SK RNA) o della stessa RNA polimerasi II (B2 RNA) o il mantenimento dei telomeri.

Le snRNP formano lo **spliceosoma**, cioè quell'insieme di particelle che contribuiscono allo splicing di un tratto della catena di mRNA.

I siti di splicing 3' e 5' e la sequenza di ramificazione sono riconosciuti da componenti dell'apparato di splicing che si assemblano per formare un grande complesso che unisce i siti stessi prima che avvenga qualunque reazione. Il complesso si assembla sequenzialmente sul pre-mRNA e lo splicing avviene solo dopo che tutti i componenti si sono assemblati. L'apparato di splicing contiene sia proteine che RNA; gli RNA prendono la forma di piccole molecole che esistono come particelle ribonucleoproteiche, sia il nucleo che il citoplasma delle cellule eucariotiche contengono molte specie distinte di piccoli RNA che variano in dimensioni da 100 a 300 bp negli eucarioti superiori e si estendono fino a 1000 basi nel lievito. L'abbondanza di questi piccoli RNA è molto variabile, da 10000 a 100000 fino a concentrazioni troppo piccole per essere rilevate. Quelli che si trovano solo nel nucleo si chiamano snRNA, mentre quelli presenti nel citoplasma si chiamano scRNA; allo stato naturale esistono come particelle ribonucleoproteiche, denominate snRNP e scRNP. Esiste anche una classe di piccoli RNA presenti nel nucleo, chiamati snRNA, che sono coinvolti nelle modificazioni dell'RNA. Le snRNP sono coinvolte nello splicing, formano insieme a molte proteine un grande complesso articolato chiamato spliceosoma, che costituisce una particella da 50-60 S; questo può formarsi in stadi successivi man mano che si aggiungono snRNP, procedendo attraverso diversi complessi di pre-splicing. Lo spliceosoma è molto grande, con una massa maggiore del ribosoma.

I 5 snRNA sono associati con le loro 45 proteine circa; oltre 70 proteine presenti nello spliceosoma comprendono proteine necessarie per l'assemblaggio, proteine per i legami con l'RNA substrato e proteine coinvolte nel processo catalitico; altre 30 proteine inoltre, sono implicate in alcuni stadi dell'espressione genica, il che indica che lo spliceosoma può servire da apparato di coordinamento.

Lo spliceosoma si forma sull'RNA precursore intatto e passa attraverso uno stato intermedio in cui contiene l'esone 5' lineare e il cappio introneo destro. Nel complesso si trova poco prodotto sottoposto a splicing, il che indica che questo viene di solito rilasciato immediatamente dopo il taglio del sito 3' e la legatura degli esoni.

Si può considerare che le particelle snRNP siano coinvolte nella costruzione della struttura dello spliceosoma; esso dipende dalle interazioni RNA-RNA oltre che proteina-RNA e proteina-proteina; alcune delle reazioni che coinvolgono le snRNP richiedono che i loro RNA formino coppie di basi direttamente con le sequenze dell'RNA sottoposto a splicing, mentre altre interazioni richiedono il riconoscimento fra le snRNP e o fra le loro proteine e altre componenti dello spliceosoma.

EXP: producendo mutazioni nei geni degli snRNA del lievito si è visto che le mutazioni in 5 geni sono letali e impediscono lo splicing. Tutti gli snRNA coinvolti nello splicing possono essere riconosciuti in forma conservata nelle cellule di animali, insetti e uccelli. Gli RNA del lievito sono decisamente più grandi, ma regioni conservate comprendono caratteristiche che sono simili agli snRNA degli eucarioti superiori.

Le snRNP coinvolte nello splicing sono U1, U2, U5, U4 e U6 che prendono il nome dagli snRNA che contengono. Ciascuna di esse contiene un singolo snRNA e meno di 20 proteine; le snRNP U4 e U6 si trovano di solito sotto forma di un'unica particella.

Un nucleo strutturale comune a tutte le snRNP consiste di un nucleo formato da 8 proteine, mentre le altre proteine sono caratteristiche.

Tutti i componenti dello splicing si assemblano e assicurano che i siti di splicing siano disponibili prima che all'RNA venga apportato un cambiamento irreversibile. Il riconoscimento delle sequenze consenso coinvolge sia RNA che proteine; l'attacco della snRNP U1 al sito di splicing 5' è il primo passaggio che avviene: esso forma coppie di basi con il sito 5' per mezzo di una regione a singolo filamento di 4-6 basi complementare al sito di splicing. Questa complementarità è di fondamentale importanza per il processo generale, e questa operazione è stabilizzata dalle proteine contenute nella ribonucleoproteina U1. Il primo complesso che si forma durante il processo di splicing è il complesso E, che contiene snRNP U1, il fattore U2AF e membri di una famiglia chiamata proteine SR, che costituiscono un gruppo importante di fattori e regolatori nello splicing e

sono caratterizzate dalla presenza di una regione ricca di Arg e Ser di lunghezza variabile; anche questi fattori legano il DNA e sono un componente essenziale dello spliceosoma, in quanto formano una struttura sull'RNA substrato U2AF a U1. Nel complesso E, U2AF è legato alla regione fra il sito di ramificazione e il sito di splicing 3' (U2AF riflette il fatto che è stato isolato come U2 auxiliary factor; nella maggior parte degli organismi è formata da una grande subunità U2AF65, che prende contatto con il tratto di pirimidine a valle del sito di ramificazione, e di una subunità U2AF35, che contatta il dinucleotide di termine del sito in 3', cioè AG).

C'è una differenza tra *S.cerevisiae* e tutti gli altri organismi rispetto ai primi stadi del processo di splicing: nel primo il sito di splicing 3' non viene coinvolto mentre negli altri è essenziale.

Un secondo fattore di splicing, chiamato SF1 nei mammiferi o BBP nel lievito, connette U2AF (o Mu2d del lievito) alla snRNP U1 legata al sito di splicing 5'. Il complesso E è convertito a complesso A quando la snRNP U2 si lega al sito di ramificazione grazie alla sequenza di basi complementari ad esso che possiede sia snRNP U1 che U2AF sono necessari per questo evento. Inoltre, parecchie proteine di snRNP U2 si legano all'RNA substrato appena a monte del sito di ramificazione. Il legame della snRNP U2 richiede l'idrolisi di ATP e impiega un pre-mRNA nella via di splicing.

Dopo la formazione del complesso A, quando un trigno contenente le snRNP U5, e U6/U4, si forma il complesso B1; questo è considerato uno spliceosoma, in quanto contiene tutti i componenti necessari per la reazione, ed è convertito in B2 quando viene rilasciato U1. La dissociazione di questo snRNA è necessaria per permettere agli altri componenti di giustapporre al sito di splicing 5', soprattutto all'snRNA U6; a questo punto l'snRNA U5 si sposta dalla sua posizione iniziale vicino alle sequenze esoniche al sito di splicing 5' nelle vicinanze delle sequenze introniche.

La reazione catalitica è innescata dal rilascio di U4, che richiede idrolisi di ATP; il ruolo dell'snRNA U4 può essere quello di sequestrare l'snRNA U6 fino a quando è necessario. Nella snRNP U4/U6, un tratto continuo di 26 basi di U6 è appaiato a due regioni di U4; quando U4 si dissocia, la regione di U6 che era prima legata diventa ora libera di assumere un'altra struttura. La prima parte di questa struttura si appaia a U2, mentre la seconda forma una forcina intramolecolare. L'interazione fra U4 e U6 è reciprocamente incompatibile con l'interazione tra U2 e U6: il rilascio di U4 che controlla la capacità di procedere dello spliceosoma.

Il sito di splicing 5' è vicino alla sequenza di U6 immediatamente a monte del tratto legato a U2; questa sequenza si appaia a sequenze dell'introne appena a valle di GU del sito 5'.

Nel corso dello splicing avvengono parecchie reazioni di appaiamento fra gli snRNA e l'RNA substrato per mezzo di interazioni di appaiamento di basi; queste stesse interazioni permettono cambiamenti di struttura che possono avvicinare i gruppi che devono reagire e possono anche creare reazioni catalitiche; naturalmente, i cambiamenti di conformazione negli snRNA sono reversibili.

#### INTRONI

Tra hnRNA e mRNA maturo si ha una perdita di circa l'80% del materiale (da 100000 a 200000). Gli introni, cioè le sequenze non codificanti all'interno dei geni eucariotici, sono stati scoperti negli studi fatti sugli mRNA, in particolare tramite ibridazione su globine. Si è preso il DNA per le globine e lo si è ibridato con mRNA per la stessa proteina, pensando di trovare sequenze extra agli estremi; in realtà i dati sperimentali furono dicevano che le sequenze mancanti provenivano dall'interno del gene (da cui il termine introne). Vi sono geni che presentano introni multipli, che possono anche arrivare ad essere 20 in un singolo gene con solo il 5% delle sequenze che codificano (mammiferi).

In generale, per quanto riguarda il loro significato biologico, potrebbero fornire un vantaggio selettivo per l'incremento della velocità di riassorbimento delle sequenze esoniche tramite ricombinazione, velocizzando di conseguenza il processo evolutivo. In altri casi potrebbero, tramite splicing alternativi (mettendo insieme esoni in ordine diverso), produrre una famiglia di proteine correlate, codificare domini, funzioni o più proteine dallo stesso gene.

Gli introni, sebbene abbiano un costo aggiuntivo di biosintesi in quanto è necessario lo splicing, danno un vantaggio evolutivo che lo ripaga.

Possiamo classificare gli introni in tre differenti classi; infatti, oltre agli introni del pre-mRNA, che necessitano di snRNA per essere excisi, ne esistono altri due tipi che si excidono per autosplicing.

**CLASSE 1** è la più conosciuta delle due classi ed è stata scoperta in E. Coli. In E. Coli, infatti, non si hanno introni, si può però prendere un ceppo del tipo LacZ- e trasferirlo con un plasmide che ha incorporato il gene funzionante (LacZ) infrazionato da un introne. Se l'introne non viene exciso, verrà espresso come LacZ- (non funzionante); al contrario, se avviene l'autosplicing, verrà espresso il gene funzionante LacZ. Il risultato dell'esperimento sono colonie che esprimono il gene LacZ e che, quindi, indicano la capacità dell'introne di fare autosplicing.

La struttura dell'introne è molto complessa e ricca di siti indispensabili al suo funzionamento: tra queste si trovano le IGS, sequenze ricche di piridine (A e G) che possono apparirsi in modo errato consentendo di individuare i siti di taglio. C'è, inoltre, una cavità che riconosce il cofattore G (guanosina) e una tasca che lo lega specificamente. La guanosina entra nel sito di legame, viene attivata, il suo ossidrilio attacca il legame in 5', stacca l'introne G e rimane legata.

Vi sono molte altre molecole di questo tipo in organismi diversi, sono dette tutti introni di classe 1 e usano la guanosina come cofattore, avvalendosi di strutture secondarie molto conservate. Tutti eseguono autosplicing tramite una reazione di transesterificazione.

**CLASSE 2** sono stati scoperti da poco, sono presenti solo nei cloroplasti e non necessitano di cofattori; generano un intermedio a laccio grazie alla presenza di una A interna e sono aiutati da proteine.

Splicing di introni II: autosplicing formante un cappio

Gli introni di geni che codificano per proteine si possono dividere in tre classi generali; gli introni del pre-mRNA nucleare sono identificati solo dal possesso dei dinucleotidi GU AG e dal tratto di pirimidine/sito di ramificazione vicino al sito di splicing 3', non hanno nessuna caratteristica comune di struttura secondaria. Gli introni I e II si trovano negli orfanelli e nei batteri, e sono classificati secondo la loro organizzazione interna, che riflette il tipo caratteristico di struttura secondaria. Essi hanno notevole capacità di eliminare se stessi da un RNA con un meccanismo chiamato autosplicing. Gli introni di gruppo I sono più comuni di quelli di gruppo II, e fra le due classi ci sono poche relazioni. In entrambe i casi l'RNA può eseguire da solo la reazione di splicing in vitro, senza richiedere attività enzimatiche fornite da proteine, che sono tuttavia quasi certamente necessarie in vivo per assistere il ripiegamento.

Gli introni sono excisi da due transesterificazioni successive; nel primo passaggio la giunzione esone introne 5' è attaccata da un gruppo ossidrilio libero (fornito da una posizione 2'OH interna di introni nucleari e I) e da un nucleotide guaninico per I); nella seconda reazione il 3'OH libero all'estremità dell'esone rilasciato attacca a sua volta la giunzione introne esone 3'.

Gli introni mitocondriali di gruppo II sono excisi dallo stesso meccanismo utilizzato dai pre-mRNA nucleari tramite un cappio che è tenuto insieme da un legami 2'; quando è incubato in vitro in assenza di altri componenti, un RNA isolato di gruppo II è capace di eseguire lo splicing, quindi le due transesterificazioni possono essere eseguite dalla sequenza stessa dell'RNA. Poiché il numero dei legami fosfodiesterici è conservato, non è necessario un apporto interno di energia.

Un introne di gruppo II si ripiega in una struttura secondaria contenente parecchi domini formati da steli appaiati ed anse a singolo filamento; il dominio 5 è separato da 2 basi dal dominio 6, che contiene un residuo di A e dona il gruppo 2'OH per la prima transesterificazione e costituisce un dominio catalitico dell'RNA.

La capacità di rimuovere se stessi in forte contrasto con la necessità di un apparato di splicing degli introni nucleari; si può considerare che gli snRNA dello spliceosoma compensino la mancanza di informazioni di sequenza dell'introne e forniscono l'informazione necessaria per formare particolari strutture di RNA. Le funzioni degli snRNA potrebbero essersi evolute dal sistema autototalitico originario. Questi snRNA agiscono in trans sul pre-mRNA substrato, e si potrebbe immaginare che la capacità di U1 di legarsi al sito 5' o di U2 di appaiarsi alla sequenza di ramificazione, abbia sostituito una reazione simile che richiedeva che la sequenza rilevante fosse portata dall'introne; quindi gli snRNA possono partecipare a reazioni fra loro e con il pre-mRNA substrato che hanno sostituito la serie di cambiamenti di conformazione che avvengono negli RNA che utilizzano meccanismi di splicing di gruppo II.

Quando un gene interrotto è trascritto in un RNA che dà origine a un singolo tipo di mRNA sottoposto a splicing, non c'è ambiguità nell'assegnazione di esoni e introni, ma gli RNA di alcuni geni seguono schemi di splicing alternativo, in cui un singolo gene dà origine a più di una sequenza di RNA. In alcuni casi lo schema finale di espressione è dettato dal trascritto primario, perché l'uso di punti di inizio diversi o la generazione di estremità 3' alternative altera lo schema di splicing, ma in alcuni casi un singolo trascritto primario è sottoposto a splicing in più di un modo ed esoni interni sono scelti, aggiunti o meno. In alcuni casi tutti i prodotti multipli sono generati nella stessa cellula, ma in altri casi il processo è regolato in modo che particolari schemi di splicing si verifichino soltanto in particolari condizioni. Ciò che controlla l'uso di queste vie alternative è un gruppo di proteine, che intervengono per favorire l'uso di siti alternativi di splicing. I doli molecolari di questi fattori di controllo non sono ancora noti, ma in termini generali lo splicing alternativo che coinvolge siti 5' diversi può essere influenzato da proteine coinvolte nell'assemblaggio dello spliceosoma.

Lo splicing alternativo può anche essere influenzato dalla espressione di un sito; ad esempio, gli esoni 2 e 3 del gene della tropoina nel topo si escludono a vicenda: il 2 è usato nel muscolo liscio e il 3 in altri tessuti. Il muscolo liscio contiene proteine che si legano a ripetuti posti sui due lati dell'esone 3, che impediscono l'uso di tutti e 3 i siti 5' e 5' che sono necessari per includerlo.

#### Autosplicing degli introni di gruppo I: transesterificazione

Gli introni di gruppo I si trovano in posizioni diverse. Sono presenti nei geni che codificano i tRNA nei nuclei degli eucarioti inferiori, sono comuni nei geni dei mitocondri dei funghi, sono presenti in 3 geni del fago T4 e in geni batterici. Hanno la capacità intrinseca di rimuovere se stessi in una reazione chiamata autosplicing.

Quando l'RNA precursore 35S è incubato in vitro, lo splicing avviene come reazione autonoma; l'introne è exciso dal precursore e si accumula sotto forma di un frammento lineare di 400 basi, che è successivamente convertito in un RNA circolare.

La reazione richiede soltanto un catione monovalente e un nucleotide guaninico come cofattore; nessun'altra base può sostituire G, ma non è necessario un trifosfato (si può usare GTP, GDP o GMP). Il nucleotide guaninico deve avere un gruppo 3'OH.

Il residuo viene unito all'estremità 5' dell'introne da un normale legame fosfodiesterico; si verificano tre reazioni di trasferimento.

Nei primi trasferimenti, il nucleotide guaninico si comporta come un cofattore che fornisce un gruppo libero 3'OH che attacca l'estremità 5' dell'introne, questa reazione crea il legame G-introne e genera un gruppo 3'OH all'estremità dell'esone. Il secondo trasferimento comporta una reazione chimica simile, in cui il 3'OH attacca il secondo esone, i due trasferimenti sono connessi, non sono stati osservati esoni liberi e quindi la loro legatura potrebbe avvenire come parte della stessa reazione che rilascia l'introne sotto forma di una molecola lineare, ma il terzo trasferimento lo converte in un cerchio. Tutte gli stadi della reazione di autosplicing avvengono per transesterificazione, in cui un estere fosfato è convertito direttamente in un altro, senza un'idrolisi intermedia. I legami sono scambiati direttamente e l'energia è conservata, quindi la reazione non richiede apporto di energia derivata dall'idrolisi di ATP o GTP (la concentrazione di GTP è alta rispetto a quella di RNA e perciò spinge la reazione in avanti e un cambiamento della struttura del DNA impedisce la reazione opposta).

Il sistema in vitro non comprende proteine e quindi la capacità di splicing è intrinseca all'RNA. L'RNA forma una struttura secondaria/terziaria specifica in cui i gruppi rilevanti sono portati in giustapposizione in modo che un nucleotide guaninico possa essere legato a un sito specifico e quindi possono avvenire le reazioni di rottura e riunione di un legame. In vivo la reazione è assistita da proteine che stabilizzano la struttura dell'RNA.

L'introne continua a essere reattivo dopo la sua escissione come molecola lineare; può ricircularizzare quando la G al terminale 3' attacca una di due posizioni vicine al terminale 5'. Il legame interno viene rotto e l'estremità 5' è trasferita all'estremità 3' OH dell'introne.

L'attività autocatalitica riflette una capacità generalizzata della molecola di RNA di formare un centro attivo che può legare cofattori guaninici, riconoscere oligonucleotidi e unire gruppi reattivi in una conformazione che permette di rompere e riunire legami.

In vivo la reazione è assistita da un complesso proteico.

**HmRNA** è una lunga molecola di RNA a singolo filamento, che costituisce il trascritto primario della trascrizione e si trova all'interno del nucleo delle cellule eucariotiche. Il trascritto primario prima di passare nel citoplasma va incontro a maturazione, processo che comprende diverse modificazioni chimico-strutturali:

- aggiunta all'estremità 5' di un CAP, cappuccio di 7 metili guanosina legata con legame 5'-5' al RNA in via di maturazione, questa modificazione è importante per il riconoscimento e l'aggancio appropriato dell'mRNA al ribosoma, oltre che rende più stabile l'mRNA.

- rimozione degli introni, splicing, grazie a spliceosoma, -aggiunta all'estremità 3' di una coda poli A, aiuta la stabilità dell'mRNA proteggendolo dall'esonucleasi, ed è importante nel processo di uscita dell'RNA dal nucleo.

Per via di queste modificazioni il trascritto primario viene trasformato in mRNA maturo. È possibile isolare hmRNA e mRNA dalla restante parte di RNA effettuando una cromatografia per affinità in cui la resina della fase stazionaria è formata da code di T, in maniera tale che vi sia interazione tra la resina e gli mRNA. Successivamente si recupera l'mRNA, lavando con un tampone a bassa forza ionica e alzando la temperatura.

## 5- E REAZIONE DELLA PEPTIDIL PUROMICINA. (disegnare formula)

La puromicina è un antibiotico inibitore della sintesi proteica in quanto assomiglia molto ad un AA aromatico legato ad un'unità di zucchero-base e potrebbe quindi essere scambiato per la parte terminale di un aminoacido t-RNA. È in grado di inibire tutti i tipi di ribosomi; entra nel sito A del ribosoma, ma essendo molto più piccolo rispetto al tRNA, non riesce a compiere il legame peptidico, e il peptide viene rilasciato determinando l'interruzione della traduzione (la peptidilpuromicina si distacca).

Per studiare il meccanismo di azione si prepara una miscela in vitro con ribosomi, mRNA, GTP, phe-tRNA e i fattori di allungamento della trascrizione (EF-TU, EF-TS, EF-G).

Dal profilo dell'assorbimento su gradiente di saccarosio noto la presenza della subunità minore (30S) della subunità maggiore (50S) e del ribosoma completo (70S); misurando la radioattività, che mi dà informazioni riguardo la posizione della catena peptidica nascente, noto che il picco coincide esattamente con il picco dell'assorbimento del ribosoma completo, dal momento che non ho segnali di terminazione tutti gli aa rimangono associati al ribosoma.

Se metto nella miscela la puromicina, osservo che il profilo dell'assorbimento rimane immutato, mentre quello della radioattività varia, la catena peptidica nascente risulta essere staccata dal ribosoma e quindi molto più leggera.

Per vedere se la puromicina può bloccare la sintesi proteica necessita dei fattori di allungamento, rieseguo l'esperimento senza i fattori ed osservo che la puromicina inibisce la sintesi circa nel 50% dei casi. Questo perché la puromicina per inibire ha bisogno di legarsi al sito A del ribosoma, se questo è occupato da un aminoacido tRNA l'antibiotico non può funzionare. In presenza dei fattori di allungamento (Tu e Ts che insieme fanno T) e specificatamente della EF-G (TRANSLOCAS) permette lo spostamento del ribosoma e quindi la liberazione del sito A inibito i ribosomi a essere sensibili alla puromicina) che libera il sito A la puromicina inibisce sempre. Questo esperimento è anche una prova della presenza di 2 siti diversi nel ribosoma.

Il fattore G permette la traslocazione e quindi la liberazione del sito A che può essere occupato dalla puromicina. Se aggiungiamo T e G la sintesi si ferma, se solo T e GTP non si hanno variazioni se aggiungo G e GTP si ha un rilascio completo.

Esperimento "reazione del frammento" per vedere quando si forma il legame pep tra aa-tRNA e puromicina, si crea leg covalente tra aa e puromicina con l'uso di sub unità 50s anche senza aggiunta di energia che è già presente all'interno dell'aa-tRNA.

## TEMA D'ESAME 3

### 1-CROMATINA E NUCLEOSOMA

La **cromatina** è composta da DNA (50%) + proteine (50%). Si può suddividere in:

-eterocromatina, appare più condensata e più colorata al microscopio, contiene sequenze non codificanti, sono difficilmente accessibili dagli enzimi per esempio RNAPolimerasi.  
-eucromatina, meno condensata, più accessibile, contiene per lo più sequenze codificanti.

Le proteine che si trovano insieme alla cromatina sono state denominate ISTONI, sono delle proteine molto conservate nell'evoluzione. Esse sono: H1, H2, H2B, H3, H4, contengono sia aminoacidi basici che acidi. Gli istoni si staccano dal DNA all'aumentare della forza ionica della soluzione in cui si trovano dimostrando che l'interazione fra loro è solo di natura elettrostatica.

Esperimento con una soluzione di NaCl dimostra il tipo di interazione che gli istoni hanno con il DNA, trattando la cromatina con una soluzione di NaCl a forza ionica crescente osservo che gli istoni si staccano in successione, H1, H2, H3, H4 ne deduco che il legame istone-DNA è di tipo elettrostatico e che non tutti gli istoni hanno la stessa forza di legame. La struttura fondamentale della cromatina è il nucleosoma, la maggior parte del DNA è organizzato nel seguente modo.

Nucleasi micrococcalica enzima restrizione che ha la caratteristica di tagliare solo il dna non legato ad alcuna proteina, enzima taglia solo frammenti di 200pb. Quindi deduciamo che il dna deve essere associato a prot che coprono regioni di 200pb. Gli istoni presentano un solco con cariche + che attira il dna.

Il **nucleosoma** è costituito da 8 proteine istoniche (2 H2A, 2 H2B, 2 H3, 2 H4) che formano un core proteico, a cui si lega il DNA. Ha un'interazione di tipo elettrostatico solo su un lato della doppia elica con le proteine basiche (+), questo fa sì che sul lato opposto vi sia repulsione tra le cariche negative portate dai gruppi fosfato. Questo determina che il DNA si avvolga facendo 1,7 giri (147pb) intorno al core istonico. Il DNA compie anche dei legami a idrogeno (circa 40) con il core istonico, la gran parte di questi sono fra le proteine e l'ossigeno del legame fosfodiesterico, solo 7 sono tra le proteine e le basi e tutti questi sono nel solco minore del DNA.

Inoltre c'è anche una regione del DNA che lega i **nucleosomi**, DNA linker lungo 20/60 pb. Ogni istone è composto principalmente da due regioni:

-istone FOLD: zona molto conservata, contiene 3/4a eliche ha la funzione di dirigere l'assemblamento del core proteico del nucleosoma,  
-code N-terminali, sono delle zone molto ricche di aa basici che fuoriescono dal core proteico, e hanno la funzione di creare interazioni con il DNA.

Il nucleosoma permette di impaccare in DNA di circa 7 volte, il passaggio successivo che permette di determinare un ulteriore impaccamento è determinato dall'istone H1. H1 è una piccola proteina con carica netta positiva, che interagisce con il DNA linker restringendo il DNA ad una maggiore adesione al core istonico. L'aggiunta di questa proteina determina una struttura con una sezione trasversale di circa 30nm: la fibra 30nm, in cui il DNA è impaccato di circa 40volte. Accorpa circa 8 nucleosomi uno vicino all'altro.

La fibra 30nm può ulteriormente contorcersi su se stessa formando anse e contro anse contenenti 40-90 kb di DNA. Questa anse sono bloccate alla loro base da una struttura proteica indicata come scaffold nucleare (proteina non istonica). Questa struttura forma il cromosoma, dove il DNA è avvolto di circa 10laquarta volta l'inizio.

**Esperimento:** proteine troviamo 8 proteine istoniche associate; da questo dato emerge un problema abbastanza interessante, in quanto le 8 proteine non sono lunghe a sufficienza per coprire un tratto così lungo di nucleotidi, e dunque il DNA non può essere contenuto all'interno delle stesse come in un involucro protettivo. Un primo indizio dei meccanismi implicati lo si ha tagliando il DNA con DNAsI (pancreatica) che agisce sulla doppia elica e taglia un legame solo generando dei nick. Se il DNA fosse coperto da proteine la DNAsI non potrebbe digerirlo per niente, mentre sperimentalmente si vede la comparsa dei nick. I frammenti sono visibili in modo chiaro sul gel dopo aver denaturato il DNA (altrimenti rimangono impaccati con l'elica integra e non si vede alcuna differenza). La serie di bande che otteniamo è regolare, con una periodicità di 10 pb, pari ad un giro d'elica. Il DNA dunque non è all'interno delle proteine, ma vie è avvolto intorno.

### 2-data la seguente sequenza individua la ORF.

### 3-ENHANCHER DI SV40

Gli enhancer sono sequenze di DNA che svolgono il loro ruolo di amplificatori della trascrizione attraverso l'associazione con diverse proteine. Sono sequenze nucleotidiche distanti anche migliaia di basi dal promotore ma comunque cis, cioè sono situate sullo stesso cromosoma del gene da trascrivere.

Hanno una lunghezza media pari a 100bp. **Senza seq prossimali non si ha trascrizione mentre se si trasporta l'intero promotore con un gene reporter in celli di mammifero si ha un'espressione altissima.** Sono stati studiati nel genoma di sv40, un virus, il cui patrimonio genetico è circolare, a doppio filamento ed è molto simile ad un grosso plasmide.

Ci sono 2 sequenze promotore:  
-EARLY (E) serve per far trascrivere il gene per l'antigene T.  
-LATE (L) che inizia a funzionare dopo, perché necessita dell'antigene T prodotto a partire da EARLY; codifica per il proteina del capside del virus.

SV40 ha 6CGBOX (21pb), sono delle regioni a monte del promotore che legano SP (stimolator protein), tale legame permette il reclutamento di fattori della trascrizione; e 2 TATABOX, formato da 6nt TATA/TAA la trascrizione inizia circa 20/30 nt dopo e così come gli altri segnali del promotore serve a reclutare fattori di trascrizione.

A monte del promotore ha 2 sequenze di 72 pb ognuna, ricche di C, G che sono molto importanti per l'efficienza del promotore, di cui non fanno parte ma influenzano notevolmente sulla sua efficienza. Infatti mutazioni in queste 2 regioni riducono molto l'efficienza di trascrizione, se non subiscono mutazioni aumenta notevolmente indipendentemente da quale sia il promotore. Per inattivarle sono necessarie diverse essere introdotte molte mutazioni contemporaneamente. Gli enhancer dal punto di vista funzionale sono molto simili all'UAS, degli eucarioti minori (UAS = upstream activator system. Nei lieviti). Influiscono sull'efficienza del promotore trovandosi a monte di esso sono lunghe dalle 50 alle 100pb esse regolano la trascrizione. Sono presenti anche in più di una copia per gene.

Tramite footprinting si è visto che l'enhancer è sempre ricoperto di specifiche proteine a specifiche sequenze. Queste proteine sono dette AP (activator protein), non hanno solo il sito di legame con l'enhancer ma permettono il ripiegamento dello stesso sul promotore. Esse sono:

-AP1, riconosce la regione TGA/G/C TCA,  
-AP2, lega il core dell'enhancer cioè la sequenza centrale, su cui mutazioni farebbero perdere la funzione all'enhancer. Hanno strutture Basic Zip mediante le quali interagiscono con il DNA. In sv40 ad AP2 può legarsi l'antigene T, che ne blocca il funzionamento (meccanismo a feedback negativo)  
-AP3 lega anch'esso il core dell'enhancer con basic zip,  
-AP4, si legano ai margini dell'enhancer.  
Le seq enancher presenti in sv40 sono identiche a quelle di mammifero, le ap sono circa 40 diverse.

### 4-RIBOSOMI

I ribosomi sono strutture presenti in tutti gli organismi, il loro numero all'interno della cellula dipende dalla velocità di crescita. Sono costituiti al 85% da ma e al 35% da proteine e contengono l'80% dell'ma tot della cell.

Sono strutture costituite da proteine ed RNA, nei procarioti contengono il 10% del totale delle proteine presenti e l'80% dell'RNA totale, sono associati all'mRNA che a sua volta è legato al DNA (nei batteri); questo perché nei procarioti la traduzione avviene contemporaneamente alla trascrizione.

Negli eucarioti invece i ribosomi sono associati al citoscheletro e alle membrane del reticolo endoplasmatico, nei procarioti sono liberi. Principalmente le funzioni del ribosoma sono:

-riunire i componenti della sintesi proteica e metterli nella giusta posizione,  
-offrire un supporto strutturale, sceglie il codone per l'mRNA,  
-ha 3 siti di movimento per il tRNA.

I ribosomi sono stati isolati a partire da un lisato grezzo di cellule di coli, in cui dopo 2 centrifugazioni (una di 30 min e una da 60 min) si è prelevato il supernatante contenente tutte le componenti citoplasmatiche, nel pellet avevo dna, parete, membrane. Dopo un ulteriore centrifugazione del supernatante si è prelevato il pellet, in cui sono contenuti i ribosomi (pellet di colore bianco gelatinoso), nel surna ci sono enzimi solubili.

Questo preparato è stato analizzato su gradiente di saccarosio e tramite misurazioni dell'assorbimento a 260nm si sono osservati 3 picchi più altri piccoli picchi di materiale pesante rappresentato dai poliribosomi (ribosomi associati che attuano una codifica di proteine in serie, sono ribosomi mantenuti saldi da un mma con maASI si eliminano i polisomi e aumenta il picco dei ribosomi 70s); un picco 30s subunità minore, uno 50s subunità maggiore e uno 70s ribosoma completo.

Questi valori sono un po' diversi per i ribosomi eucarioti che sono più grossi, 40s, 60s, 80s, il valore di sedimentazione del ribosoma completo non coincide con la somma dei valori delle 2 subunità poiché la velocità di sedimentazione (misurata in Svedberg) non dipende solo dal peso ma anche dalla forma della molecola. Per togliere il piccolo relativo ai materiali pesanti è sufficiente aggiungere RNasi che taglia l'mRNA che tiene uniti i ribosomi, mentre per togliere il picco relativo al ribosoma completo aggiungo EDTA che elimina Mg<sup>2+</sup> che tiene unita la subunità minore con la maggiore.

Dal materiale che costituisce ogni picco posso separare RNA da proteine con metodi di purificazione; estraggo rRNA purificato centrifugando su gradiente di saccarosio e misuro l'assorbimento.

-il ribosoma procariotico ha una massa di 2700 kDa ed un coefficiente di sedimentazione di 70S è costituito da 2 subunità, 50S (ha una parte piatta una + tondeggiate ed alcune preteruberanze cui si va a posiz la sub 30s) +30S (ha una forma allungata, una tondeggiate, una più allungata (corpo) e una a piattaforma). La sub maggiore contiene rRNA 23S+5S, la sub minore contiene rRNA 16S.

-il ribosoma eucariotico ha una massa di circa 4000 kDa ed un coefficiente di sedimentazione di circa 80S, è costituito da 2 sub 60S (ma 5s + ma 5.8s + ma 28s e 45 prot) + 40S (ma 18s + 33 prot), la prima è costituita da 3 rRNA 28S, 5S, 5.8S la seconda è composta da una sola catena di rRNA 18S.

Le proteine ribosomiali invece sono più difficili da caratterizzare, queste sono state analizzate mediante elettroforesi bidimensionale. (è una separazione ortogonale nella quale i componenti della miscela proteica sono separati sulla base del peso molecolare e del peso molecolare). Si è visto che la prima subunità del ribosoma batterico ci sono 21 proteine, chiamate s1/s21, nella seconda subunità inoltre ci sono 32 proteine chiamate h1/h32. Le proteine non hanno attività catalitica ma solo funzione strutturale. Sui rbo esistono 2 siti Attivi A e P: A dove c'è il sito per nuovi aa, P è il sito dove si colloca la catena nascente che si trova sulla sub 50s.

Dal punto di vista morfologico, notiamo delle somiglianze tra i ribosomi dei Eucarioti e i ribosomi degli Archea, per esempio il beco nella subunità minore, mentre molti diversificati dal punto di vista strutturale sono i ribosomi degli eubatteri.

rRNA e ribosomi possono essere usati come orologi molecolari, per costruire un albero evolutivo. Tutti gli organismi esistenti si sono evoluti da un progenitore comune che doveva avere ribosomi perché eubatteri, eucarioti e archea presentano ribosomi. Il ribosoma ancestrale era probabilmente grosso e dotato di becco (sub minore), quasi come quello eucariotico e ancora abbastanza simile a quello degli archibatteri, poi è stato ottimizzato per essere più comodo e avere comunque alte prestazioni per gli organismi (archea) che necessitano di una crescita rapida.

### SINTESI PROTEICA

Le proteine cellulari vengono sintetizzate nei ribosomi che sono la componente principale della cellula e costituiscono il 10% delle proteine totali. Il meccanismo di traduzione viene diviso in tre fasi: inizio, allungamento e terminazione. INIZIO: il ribosoma si lega all'mRNA e contemporaneamente viene attaccato al primo aa attaccato al suo tRNA; ALLUNGAMENTO: il ribosoma aggiunge un aa alla volta alla catena polipeptidica nascente; TERMINAZIONE: il ribosoma rilascia l'mRNA e il polipeptide completo.

Prima che l'inizio della traduzione abbia inizio devono avvenire due cose:

- caricamento del tRNA: i tRNA devono essere caricati con i corrispondenti aa, formando l'aa-tRNA;
- dissociazione dei ribosomi: i ribosomi si devono scindere nelle due subunità costituenti, affinché si formi il complesso d'inizio sulla subunità ribosomiale piccola.

### ATTIVAZIONE DEGLI AMMINOACIDI

Per attivazione degli aa si intende l'attacco dell'aa con il tRNA corretto. Tutti i tRNA all'estremità 3' presentano le stesse tre basi: CCA, con l'adenosina che rappresenta il bersaglio per il caricamento. L'aa viene aggiunto al tRNA mediante legame esteri tra il suo gruppo COOH e il gruppo OH dell'adenosina. Il legame esteri è un legame energeticamente sfavante, per evitare a questo problema esiste l'amminoacil-tRNA-sintetasi che, essendo un enzima, riesce a scindere ATP per permettere questo legame che, altrimenti, sarebbe impossibile. Il caricamento avviene in due passaggi:

1. aa+enzima+ATP → enzima-aa-AMP+2P<sub>i</sub>
2. enzima-aa-AMP+tRNA → aa-tRNA+enzima+AMP

dove la reazione 1 spiega come l'aa viene attivato utilizzando l'energia dell'ATP, a creare un aa-AMP più 2 fosfati. L'energia rilasciata dal distacco di questi due gruppi fosfati viene immagazzinata nell'aa-AMP che così si definisce attivato. La reazione due, invece, mostra come l'energia immagazzinata nell'aa-AMP viene usata per trasferirla l'aa al tRNA a formare l'aa-tRNA, che andrà a legarsi al sito A del ribosoma.

L'aa-tRNA-sintetasi è un enzima importantissimo poiché determina la specificità della reazione.

### REAZIONE DI INIZIO NEI PROCARIOTI

Il ribosoma procariotico è formato da due subunità: 50S e 30S che unite originano il complesso ribosomiale 70S. La subunità 30S lega l'mRNA e l'anticodone del tRNA. La subunità 50S, invece, lega l'estremità del tRNA caricata con l'aa e possiede l'attività peptidil-transferasica della fase di allungamento.

La dissociazione dei ribosomi batterici è parte integrante della reazione d'inizio e dipende da tre fattori: IF1, IF2 e IF3. In particolare i fattori IF1 e IF3 prendono parte alla dissociazione. Effettuata la dissociazione, la cellula costituisce un complesso sulla subunità ribosomiale 30S, che include l'mRNA, l'aa-tRNA e i fattori di inizio, questo complesso prende il nome di complesso di inizio 30S.

La formazione del complesso di inizio 30S implica le seguenti tappe:  
Dissociazione del ribosoma 70S nelle due subunità 30S e 50S sotto l'influenza di IF1;  
Legame di IF3 alla subunità 30S, che non permette la riasociazione tra le subunità ribosomiali;  
Legami di mIF1 e IF2-GTP a fianco di IF3;  
Legame dell'mRNA e di fMet-tRNA a formare il complesso di inizio 30S. In particolare IF2 promuove il legame di fMet-tRNA, mentre IF3 promuove il legame dell'mRNA. Dopo il caricamento sul tRNA la Metionina viene fosforilata dando origine a N-formil-Metionina che è il primo aa incorporato in un polipeptide, ma che viene frequentemente rimosso dalla proteina durante la sua maturazione. Usando fMet-tRNA, le Metionine sono incorporate solo nella prima posizione del polipeptide, fMet-tRNA serve, quindi, da aa-tRNA iniziatore. Il corretto legame tra la subunità ribosomiale 30S e il sito di inizio del messaggero dipende da una breve sequenza dette di Shine-Dalgarno che si trova a monte del codone d'inizio.

La formazione del complesso 70S (complesso finale) implica, invece, le seguenti tappe:  
La subunità 30S è legata al filamento di mRNA, fMet e ai fattori IF1, IF3 IF2-GTP;  
Legame della subunità 50S con perdita di IF1 e IF3 da parte della subunità minore;  
Dissociazione anche di IF2 dalla subunità minore con simultanea idrolisi di GTP.  
Il prodotto è il complesso d'inizio 70S pronto per iniziare l'allungamento.

### REAZIONE DI INIZIO NEGLI EUKARIOTI

Il ribosoma eucariotico è formato da due subunità: la subunità maggiore 60S e quella minore 40S che si associano a formare il complesso ribosomiale 80S. A differenza di quello procariotico, l'inizio della traduzione negli eucarioti usa la metionina e non la N-formil-metionina. Inoltre gli mRNA eucarioti non contengono sequenze di SHINE DALGARNO per indicare ai ribosomi da dove iniziare a tradurre, però la maggior parte contiene un cap in 5' che dirige i fattori a legarsi e a cercare un codone d'inizio. Le cellule riconoscono il cap in 5' e successivamente fanno la scansione del trascritto finché non incontrano il codone d'inizio. Ma allora cosa riconosce l'AUG giusto da quello sbagliato? Si è scoperto che nell'interno dell'AUG d'inizio esiste una sequenza conservata altamente conservata.

CCRCCAUGG dove R è una purina (A o G)  
E si trovò inoltre che l'inizio ottimale si ha quando è presente una purina in posizione -3 e una G in posizione +4, considerando la A dell'AUG il punto +1.

ACCAUGG  
Esistono inoltre delle strutture secondarie che possono avere effetti positivi o negativi sulla traduzione:  
1- uno stelo ad ansa a soli 12nt a valle del cap inibisce fortemente la traduzione, poiché impedisce la formazione dello spazio sufficiente per assemblare il complesso d'inizio  
2- se lo stelo ad ansa si trova a 52nt a valle del cap non interferisce con la traduzione, poiché nonostante AUG sia nascosto dentro una forcina, i fattori d'inizio riescono a svolgere il tratto di RNA a doppio filamento.  
3- Se lo stelo ad ansa è posto a 71nt a valle del cap blocca completamente la scansione, impedendo così la traduzione.

Così come nei procarioti, anche negli eucarioti esistono dei fattori d'inizio chiamati eIF.

La formazione del complesso 80S implica i seguenti passaggi:

- Il fattore eIF3 converte la subunità ribosomiale 40S a 40Sn che resiste alla riasociazione con la particella ribosomiale 60S ed è pronta ad accettare l'amminocilic-IRNA iniziatore.
- Con l'aiuto di eIF2 legato a GTP, l'Met-IRNA si lega alla particella 40Sn, formando il complesso 43S.
- Supportato da eIF4f l'mRNA si lega al complesso 43S, formando il complesso 48S.
- I fattori eIF1 e eIF1a promuovono la scansione fino al codone d'inizio
- Il fattore eIF5 promuove l'idrolisi del GTP legato a eIF2 per poter riunire le due subunità, e produrre così il complesso 80S che è pronto per iniziare la traduzione dell'mRNA.

#### REAZIONE DI ALLUNGAMENTO (MODELLO A TRE SITI)

L'allungamento della sintesi proteica può essere spiegata mediante il modello a 3siti, secondo il quale il ribosoma contiene mRNA, e tre siti di legame per l'amminocilic-IRNA: SITO P (peptidico) SITO A (amminocilico) e SITO E (di uscita).

Considerando che il sito P è occupato da Met-IRNA

1 FASE: per cominciare l'allungamento, serve un altro amminocilico che si unisce al primo. Questo secondo aa arriva legato a un tRNA e si posiziona sul sito A. Il legame del secondo aa-IRNA al sito A richiede un fattore di allungamento detto EF-Tu e GTP.

Il fattore specifico di allungamento EF-Tu si associa al ribosoma solo durante il processo di ingresso dell'amminocilic-IRNA e una volta che l'aa-IRNA è al suo posto, lascia il ribosoma per ripetere la reazione. EF-Tu è una G proteina. Quando è presente il GTP il fattore è nel suo stato attivo, quando il GTP è idrolizzato a GDP il fattore diventa inattivo e l'attività viene ripristinata quando il GDP è sostituito da GTP. Il complesso binario EF-Tu-GTP lega aa-IRNA per formare un complesso ternario aa-IRNA-EF-Tu-GTP che si lega al sito A dei ribosomi quando il sito P è già occupato dal peptidil-IRNA. Questa è la reazione critica che assicura che aa-IRNA e peptidil-IRNA siano posizionati correttamente per la formazione del legame peptidico. Il riconoscimento codone-anticodone scatena l'idrolisi del GTP da parte di EF-Tu e viene rilasciato il complesso binario EF-Tu-GDP. Questa forma è inattiva e non è in grado di legare un altro aa-IRNA in modo efficace perché interviene l'altro **fattore EF-G** che media la rigenerazione della forma attiva EF-Tu-GTP, mediante la sua azione di catalisi. Questo ciclo avviene migliaia di volte al secondo durante la sintesi proteica. Il rilascio di EF-Tu GDP è necessario per la formazione del legame peptidico fra i due aa.

L'interazione di EF-Tu ha anche un ruolo nel controllo di qualità. [ ] Gli aa-IRNA sono portati al sito A senza sapere se i loro anticodoni si adatteranno ai codoni. L'idrolisi di EF-Tu-GTP è relativamente lenta poiché richiede un tempo più lungo di quello necessario per la dissociazione di un aa-IRNA non corretto dal sito A. La maggior parte delle specie non correte viene rimossa a questo stadio. Anche il rilascio di EF-Tu-GDP dopo l'idrolisi è lento e quindi qualsiasi aa-IRNA può dissociarsi a questo stadio. Il principio fondamentale è che le reazioni che coinvolgono EF-Tu avvengono abbastanza lentamente da permettere agli aa-IRNA di dissociarsi prima di rimanere intrappolati nella sintesi proteica. L'equilibrio fra la velocità e l'accuratezza della traduzione è quindi delicato. Se la formazione del legame peptidico è troppo rapida, gli aa-IRNA non corretti non hanno abbastanza tempo per rilasciare il ribosoma, perciò i loro aa verranno incorporati nella proteina. Se la traduzione è invece troppo lenta non verranno prodotte proteine abbastanza velocemente per permettere il successo della crescita.

2 FASE: Si forma il primo legame peptidico. Un enzima detto PEPTIDIL TRASFERASI che fa parte della subunità ribosomiale grande, trasferisce l'F-Met dal suo tRNA nel sito P all'aa-IRNA nel sito A: si forma così un dipeptidil tRNA. Nel sito P resta il tRNA deacilato (cioè privo del suo aa). I componenti minimi necessari per l'attività della peptidil-transferasi in vitro sono l'RNA 23S e le proteine L2 e L3. Il legame peptidico parte solo quando EF-Tu rilascia il suo tRNA.

3 FASE: Avviene la TRASLOCAZIONE: l'mRNA con il suo peptidil tRNA attaccato nel sito A si muove di un codone in avanti verso sinistra.

- 1) il tRNA scarico nel sito P lascia il ribosoma attraverso il sito E
- 2) il dipeptidil tRNA nel sito A insieme con il suo codone corrispondente, si muove nel sito P.
- 3) il codone che era in attesa sulla destra si muove dentro al sito A, pronto per interagire con un aa-IRNA.

Affinché avvenga la traslocazione serve un fattore di allungamento detto EF-G ed anche GTP. I ribosomi non possono legare simultaneamente EF-Tu e EF-G. Quindi EF-Tu-GDP deve essere rilasciato prima che si possa attaccare EF-G, che a sua volta deve staccarsi prima che si possa legare il nuovo complesso aa-IRNA-EF-Tu-GTP. Le strutture di EF-Tu e EF-G fanno pensare che i due fattori competano per lo stesso sito di legame (sito A). Entrambi i fattori di allungamento sono attivi se legate a GTP e inattive se legate a GDP quindi la forma trifosfato è necessaria per l'attacco al ribosoma, solo in compagnia di GTP di cui ha bisogno per svolgere la sua funzione. EF-G si lega al ribosoma per promuovere la traslocazione e viene quindi rilasciato dopo lo spostamento del ribosoma: EF-G si lega al ribosoma, viene idrolizzato il GTP e il ribosoma si sposta di tre nucleotidi. EF-G e GDP vengono rilasciati.

#### REAZIONI DI TERMINAZIONE

L'allungamento si ripete più volte finché non si incontra un codone di stop che segnala il momento di terminare la traduzione. I codoni di stop sono 3: UAA UAG UGA. Quando uno di questi tre codoni arriva nel sito A del ribosoma si ha normalmente che nessun tRNA lo riconosce. Il riconoscimento avviene a carico di proteine chiamate FATTORI DI RILASCIO (RF).

RF1 riconosce i codoni di stop UAA eUAG

RF2 riconosce UAA e UAG

RF3 è una proteina che lega GTP facilitando il legame di RF1 e RF2 al ribosoma.

Gli eucarioti hanno invece due fattori di rilascio:

eRF1 che riconosce tutt e tre i codoni di terminazione

eRF2 che è una GTPasi ribosoma dipendente che aiuta eRF1 nel rilascio del polipeptide completo.

Dopo la terminazione è necessario che i ribosomi rilascino l'mRNA e tale rilascio avviene grazie ad un fattore chiamato FATTORE DI RILASCIO RIBOSOMIALE (RRF). Il meccanismo di rilascio ribosomiale è tutt'ora sconosciuto.

#### INIBITORI SINTESI PROTEICA

Sono specifici, inibiscono sintesi batteri ma innocui per eucarioti, usati come antibiotici uccidono batteri ma nessun danno per l'ospite. Batteri: legano sub 30s - aminoglicosidi con amminozuccheri. Streptomina: causa errata lettura codice. Tetraciclina: bloccano sito A impedendo legame amminocilic che arrivano - eucarioti: usati come antifungini (tosina difterica) agiscono localmente → 40S = kanamicina, 60S: glutaramici - emetina. Inibitori universali: per 50 e 60 s promicina e sparsomicina. Per 30 e 40 s gromicina. Streptomina lega 30 s, effetto disastroso bocca redox complesso inizio quando arriva 50 s, i 70s si disfa, aumenta errori lettura, miscoding induce incorporazione AA sbagliati (a basse [I] ad alte concentrazioni inibisce inizio sintesi.

#### TEMA D'ESAME 4

##### 2-regolazione genica procarioti

Una prima approssimativa regolazione dell'espressione genica dei procarioti la si ha a livello delle sequenze promotore. Negli procarioti i promotori per l'RNAPolimerasi sono abbastanza conservati, sono composti da 2 regioni esonucleotidiche -35 TTGACA e -10 TATAAT, questi rappresentano i promotori ideali, quelli cioè che meglio riconosce RNAPolimerasi (fattore sigma), più le sequenze si somigliano a queste ideali maggiore sarà l'efficienza di trascrizione. In col per esempio i geni che non devono essere particolarmente espressi hanno le sequenze promotori differenti da queste. L'RNAPolimerasi tramite la subunità  $\alpha$  è in grado di riconoscere altre sequenze, se presenti, ricche di AT, dette sequenze UP, il cui riconoscimento aumenta ulteriormente l'efficienza di trascrizione.

Un'altra modalità di regolazione è legata al fattore sigma, subunità proteica molto importante per l'RNA polimerasi procariotica dal momento che stabilisce un'interazione stabile col filamento di DNA e riconosce il promotore.

Esistono diversi fattori sigma,  $\sigma_{70}$  che si occupa dei geni che codificano per la gran parte delle proteine,  $\sigma_{32}$  che si occupa dei geni che codificano per proteine utili in uno shock termico,  $\sigma_{24}$  riconosce i promotori dei geni per le proteine utili nella chemiotassi,  $\sigma_{54}$  riconosce tutti i promotori dei geni x il metabolismo dell'azoto.

La vera e propria regolazione genica nei procarioti è controllata da geni regolatori, geni che codificano per le proteine la cui funzione è regolare l'espressione di altri geni.

La regolazione può essere:

-positiva, utilizzata dai batteri ma solitamente è tipica degli eucarioti, nello stato basale il gene è inattivo; affinché possa iniziare la trascrizione sono necessarie delle proteine che si legano a sequenze a monte del promotore, non è sufficiente l'interazione RNAPol-DNA.

-negativa, tipica dei batteri, nello stato basale il gene è attivo, cioè codifica per il proprio prodotto proteico, vi è però una proteina repressore, che ha come bersaglio l'operatore vicino al promotore che impedisce all'RNAPolimerasi di iniziare la trascrizione.

Un esempio di regolazione della trascrizione nei batteri è l'operon LAC:

LacZ .codifica per la betalgalattosidasi, scinde galattosio in lattosio e glucosio.

LacY .permeasi del lattosio, permette l'ingresso del lattosio dentro la cellula se presente nel terreno.

LacA .codifica per la transacetilasi.

P1 :sito CAP, punto in cui si lega la proteina CAP.

O : operatore, sito in cui si lega il repressore.

LacI : sintetizza il repressore, è il gene regolatore, non fa parte dell'operone Lac.

-In presenza di glucosio, poco lattosio viene bloccata la trascrizione dei geni dal repressore che si lega all'operatore perché le cellule preferiscono il glucosio come fonte energetica, perché è più efficiente. lacI produce il repressore che si lega all'operone e blocca la trascrizione.

-In presenza sia di glucosio che lattosio, non c'è il repressore LAC perché viene prodotto da LacI solo in assenza di lattosio, anche l'attivatore CAP è inattiva perché ha bisogno di legarsi a C-amp (indice di poco glucosio) per attivarsi. La trascrizione dei geni avviene con una bassa efficienza di trascrizione.

-In mancanza di glucosio, in presenza di lattosio, l'allolattosio si lega al repressore e lo stacca dall'operatore, C-amp interagisce con CAP, cambia la sua conformazione, si lega al sito CAP e attiva la trascrizione dei geni per l'utilizzo del lattosio.

##### 3-sintesi e maturazione del tRNA

**TRNA** = molecole piccole conservate che presentano molte basi modificate nella zona adiacente all'anticodone. La struttura del tna è formata da 3 bracci che terminano in anse, il braccio della pseudouridina, il braccio della diidrouidina, e il braccio dell'anticodone + un piccolo braccio extra formato da 10pb e dal sito accettore a cui si legherà l'aa corretto.

La sintesi e la maturazione del tRNA è diversa tra procarioti ed eucarioti.

**-PROCARIOTI:** i geni che codificano per il tRNA si trovano all'interno di operoni (sequenze codificanti per più segmenti genici che sono controllati da un unico promotore). Questi operoni vengono chiamati rrm(A-G), la lettera indica quali dei 7 operoni per il tRNA stiamo parlando. Ad esempio in coli rrmC rappresenta l'unico operone che porta il tRNA per il triptofano. Una mutazione su questo gene potrebbe essere infatti deleteria.

Questi geni possono essere di 2 tipi:

1-se hanno già la sequenza CCA in 3' (espressi dal gene di tipo 1)

2-se non hanno la sequenza CCA in 3' (espressi dal gene di tipo 2) tutti i geni per i tRNA umani sono di questo tipo.

Il sistema di maturazione è composto da ribonucleasi, enzimi che tagliano l'RNA all'interno della sequenza codificante. Il trascritto primario dell'operone è tRNA 30S, il primo enzima che interviene è la ribonucleasi P, ha attività endonucleasica e genera per tutti i tRNA la giusta estremità in 5'. E' stato il primo ribosoma ad essere stato studiato, composto da RNA e proteine, la funzione enzimatica è svolta dall'RNA stesso che riconosce la struttura secondaria (a trifoglio) del tRNA che si forma già nel trascritto primario, la proteina svolge la funzione di cofattore. I tRNA sono ora tutti quindi con la giusta estremità in 5' ma con una coda in 3', interviene dunque la ribonucleasi D che diversamente dal primo ha attività esonucleasica e idrolizza in prossimità di 3'. Se ho un RNA di tipo 1 lascia la sequenza CCA, se ho un RNA di tipo 2 taglia dove si dovrebbe attaccare CCA che poi viene successivamente attaccato da tRNA nucleotidil transferasi.

Tuttavia mutanti di coli difettivi per il gene che codifica per il tRNA nucleotidil transferasi, non riescono a sopravvivere anche se il suo RNA è di tipo 1, perché questo enzima ha anche diverse funzioni strutturali. Sintesi in cui i geni sono contenuti in operoni e all'interno degli stessi operoni i geni per tna possono essere organizzati in cluster.

**-EUCARIOTI:** i geni che codificano per tRNA negli eucarioti non sono organizzati in operoni, ci sono introni nei geni, tutti i geni sono di tipo 2 e sono presenti in più copie all'interno del genoma, trascritti dalla pol III con un promotore interno e un prodotto iniziale più grande di quello maturo infatti oltre al gene x i tna ci sono regioni brevi aggiuntive in 5' e in 3' e introni seq che subiranno uno splicing ovvero un'eliminazione in fase di maturazione.

Le fasi di sintesi sono analoghe a quelle negli procarioti e anche gli enzimi sono gli stessi. La maturazione però è differente poiché negli eucarioti deve avvenire lo splicing (citocologico, in due fasi: l'introine viene exciso da una endonucleasi e la seconda fase viene scaldato il terminale 3' con quello 5' con consumo di ATP) degli introni (sequenze presenti nel trascritto primario ma non nel tRNA maturo). In questo processo intervengono 5 diversi enzimi: endonucleasi (enzima1) che taglia via l'introine lineare nel citoplasma, dopo l'eliminazione degli introni:

-3' ha un fosfato ciclico 3'/2

-3' ha un OH,

la ligasi (enzima2) per poter unire i tratti degli esoni richiede estremità 3'-OH e 5'-P servono quindi enzimi che ristabiliscono queste estremità: intervengono fosfodiesterasi (enzima3) al 3' che rompe il ciclo e lo trasforma in P-3'-OH e chinasi (enzima4) che defosforila ATP per fosforilare l'estremità al 5'. In queste condizioni può intervenire l'RNAligasi (enzima2) usando come cofattore ATP per unire le 2 estremità, il fosfato in 3' è eliminato dalla fosfatasi (enzima5).

##### 4-topoisomeri e topoisomerasi

I topoisomeri, sono molecole di uguale formula chimica, con lo stesso tipo di legami ma differenti topologie. Nelle molecole lineari di DNA, poiché le terminazioni sono libere, il numero di avvolgimenti di un filamento intorno all'altro filamento può essere variato mediante la rotazione reciproca. Ma se le estremità della molecola sono legate covalentemente l'una all'altra a formare un struttura circolare, il numero di volte che un filamento gira intorno all'altro non può cambiare se non si rompono i legami fosfodiesteri tra i nucleotidi.

**-Lk:** linking number, rappresenta il numero di volte che un filamento deve essere passato attraverso l'altro affinché le due catene possano separarsi l'una dall'altra. È sempre un numero intero, è una proprietà topologica invariante del cccDNA (DNA circolare covalentemente chiuso).

Il linking number è la somma di 2 componenti geometriche, il twist e il writhe;

**-Tw:** numero di avvolgimenti di un filamento di DNA intorno all'altro, lo si trova facendo (lunghezza del DNA in pb)/(passo dell'elica in pb). Se consideriamo una molecola di cccDNA planare, il twist number coincide con il linking number, però i cccDNA di norma non hanno una conformazione planare, ma presentano generalmente delle tensioni torsionali che impongono all'asse longitudinale della doppia elica di incrociarsi su se stesso. Questo incrocio tridimensionale viene definito writhe.

**-Wr:** numero di writhe, numero numero di incroci dell'asse longitudinale su se stesso. Rappresenta la superelicità, cioè il grado di superavvolgimento. Se Wr < 0 allora si dice che il DNA è superavvolto negativamente, se Wr > 0 allora si dice che il DNA è superavvolto positivamente.

Tw e Wr possono variare ma al loro somma è costante nei cccDNA, **Lk=Tw+Wr**.

È più semplice e significativo esprimere il superavvolgimento della molecola (Wr) come densità di superelicità cui viene assegnato il simbolo  $\sigma$ ,  $\sigma = \text{Lk} - \text{Tw} / \text{Lk}$  e nei DNA circolari è circa  $\sigma = -0,06$  il che significa che ho un giro di superavvolgimento negativo ogni 200pb.

Il superavvolgimento rappresenta una riserva di energia, circa 8 kcal/mol, il DNA superavvolto è posto a un livello energetico superiore rispetto al DNA rilassato, questa energia viene utilizzata per svolgere l'elica e facilitare quei processi in cui è richiesto l'apertura del doppio filamento come replicazione e trascrizione. Le regioni di DNA superavvolto negativamente hanno la tendenza a svolgersi localmente.

Viene favorito il rilassamento dell'elica perché il superavvolgimento genera delle tensioni locali, l'apertura e il rilassamento avvengono nelle zone ricche di A,T.

Il superavvolgimento positivo invece si oppone all'apertura dell'elica; per questo il DNA fisiologico è sempre superavvolto negativamente, mentre per esempio quello dei batteri che vivono in condizioni estreme (termofili) è superavvolto positivamente per far sì che si possa opporre alla denaturazione.

Le **topoisomerasi** sono una categoria di enzimi appartenenti alla classe delle isomerasi che determinano una diminuzione o un aumento del grado di superavvolgimento, fanno variare Lk, mediante la rottura del legame fosfodiesteri. Esistono due tipi di isomerasi:

**-topoisomerasi di tipo1:** rompono transitoriamente una sola delle catene di DNA, la ruotano attorno a quella integra un unità per volta e infine riuniscono le estremità interrotte, non usa ATP, perché tende sempre a ridurre il numero di superavvolgimento e quindi è un'azione favorita termodinamicamente. Il DNA sostanzialmente si rilassa attraverso una serie di stati intermedi ogni intervento fa variare di 1 il num di superavvolgimento.

La sua attività viene detta di Nick-closing (rottura-riparo) perché l'enzima rompe il legame fosfodiesteri tra 3'OH e 5'P, un residuo di Tyr nel sito attivo si lega all'estremità in 5' e conserva l'energia del legame fosfodiesteri che viene successivamente utilizzata per riformare il legame fosfodiesteri.

La topoisomerasi 1 ha 3 tipi di attività :

-svolge il DNA dai superavvolgimenti,

-svolge i nodi del DNA a singolo filamento,

-unisce due filamenti di DNA a singolo filamento con superavvolgimento opposti.

È presente sia nei batteri che negli eucarioti, ma in sottoclassi diverse:

-sottoclasse 1A: nei batteri, svolge solo DNA superavvolto negativamente,

-sottoclasse 1B: negli eucarioti, sia DNA + che -.

-

**-topoisomerasi di tipo 2:** dette anche DNA Girasi, hanno anch'esse una doppia attività di NICKING CLOSING ma usano ATP come fonte di energia, agiscono su DNA a doppio filamento e ad ogni intervento aumentano di due il numero di superavvolgimenti negativi. È un'enzima formato da 2 subunità A e due B: la A è un topo e potrebbe funzionare senza ATP, ma accoppiata alla B produce superavvolgimenti negativi. Si attacca al DNA creando 2 nick, rompendo l'elica ma mantenendo i lembi vicini tra loro. Il DNA viene fatto passare attraverso i filamenti interrotti facendo variare o (numero di superavvolgimenti totali) di due unità.

Una molecola chiusa di DNA può essere caratterizzata dal suo numero di legame, cioè il numero di volte che un filamento incrocia l'altro nello spazio. Molecole chiuse di DNA con la stessa sequenza possono avere numero di legame diverso, il che riflette diversi gradi di superavvolgimento. Queste molecole identiche ma con numero di legame differente prendono il nome di isomeri topologici o topoisomeri. Il numero di legame, L, è composto da un numero di avvolgimento, W, cioè la rotazione dell'asse di duplo nello spazio (numero di superavvolgimenti), e dal numero di avvitemento, T, che rappresenta la rotazione di un filamento intorno all'altro (numero totale di giri del duplex).

Le DNA topoisomerasi sono enzimi che catalizzano cambiamenti nella topologia del DNA rompendo transitoriamente uno o entrambi i filamenti del DNA, facendo passare il filamento intatto attraverso l'interruzione e quindi risalendo quest'ultima. Le estremità generate dalla rottura non sono mai libere, ma sono manipolate esclusivamente entro i confini dell'enzima (sono legate ad esso covalentemente). Le topoisomerasi agiscono sul DNA indipendentemente dalla sua sequenza.

##### 5-Dna polimerasi batteriche

È un enzima specifico per la sintesi del dna usa template di DNA e seguendo le regole di complementarità delle basi riesce a catalizzare la reazione di sintesi riconoscendo e scegliendo le corrette basi usando desossinucleotidi fosfato, polimerizzando progressivamente con attacco nucleofilo al 3' oh ad alpha P del NTP con polarsità chiaramente inversa rispetto al template in direzione sempre 5'-3' necessita di un primer che le dia l'avvio (di 5/10 nucleotidi) con l'oh in 3' pronto al legame.

Nei batteri sono presenti 3 tipi di polimerasi:

**Dna pol I** è costituita da una singola catena di 109 kDa si presenta come un enzima notevolmente versatile oltre ad avere un'attività dna polimerasica infatti possiede 2 attività esonucleasiche, separabili con tripsina. Il frammento di Klenow di circa 70 kDa dotato di attività di polimerizz e di pirfosfatasi con attività esonucleasica (correzione bozze, togli e nucleotidi sbagliati se così accade la pol I si ferma e usa l'attività 3'→5' per tornare indietro togliere il nucleotide sbagliato e continuare l'attività pol I)

(l'attività 5'→3' permette alla pol I di degradare il filamento di fronte alla pol in avanzamento così facendo può rimuovere e sostituire il filamento in un unico passaggio) e da un frammento piccolo. Può polimerizzare solamente frammenti di dna con nicks non filamenti intatti. Si lega all'estremità 3' del nick e inizia a polimerizz in direz 5'. Esegue attività di nick translation, questo enzima per quanto riguarda attività esonucleasica ( per evitare errori di appaiamento) discrimina purine e pirimidine. Mutanti pol I- non è indispensabile anche se estremamente utile.

**Dna pol II** (pesa 80 kda) funziona bene nei dna con gap. È + processiva rimane legata per un tratto di basi > rispetto alla I, non è in grado di eseguire nick translation e non lavora su dna con nicks. Ha attività di gap-filling ovvero chiude i buchi se di lunghezza inferiore alle 100 pb. Non è indispensabile per replicazione

**Dna pol III** indispensabile per la sopravvivenza. È un coenzima formato da due diverse sub. La piu' impo è la subunità  $\alpha$ , è la piu' grande ed essa è la sede dell'attività polimerasica. La sub A fzione male in vitro se come substrato viene usato DNA con gaps e non funziona del tutto se come substrato si usa SS DNA e primer, ma in condizioni fisiologiche ovvero legato alle altre sub minori lavora 10 volte + velocemente rispetto alla pol I. Le altre subunità minori sono epsilon (ε) con attività di proofreading, theta (θ) che ha attività di assemblaggio alfa+epsilon+θ per dare un core polimerasico, beta' fattore che

permette di utilizzare DDNA come sub. fattori ausiliari gamma, delta. Essa sintetizza solo in modo continuo in direzione 5'→3'. Il filamento 3'→5' è detto lagging e su esso la pol III genera i frammenti di okazaki. si è scoperto invece che essa è indispensabile per la sopravvivenza della cellula. Essa è costituita da 4/8 proteine; la subunità  $\alpha$  è la più grande ed è la sede dell'attività polimerasica. Tuttavia per avere una corretta velocità di sintesi sono necessarie anche le altre subunità. La DNApol III, una volta attaccata al DNA, è molto processiva e prosegue la sua attività per milioni di basi; essa forma sul DNA una specie di struttura a doppia C invertita. In ciascuna forca replicativa sono presenti due diverse polimerasi III che allungano ciascuna dei due filamenti. Siccome la DNApol III sintetizza soltanto in direzione 5'→3', un filamento sarà continuo (filamento leading) mentre l'altro sarà sintetizzato per filamenti successivi detti frammenti di Okazaki (circa 1000-1200 pb) (filamento lagging). Su questi ultimi filamenti torna a giocare un ruolo molto importante la DNApol che con il suo Nick Translation rimuove i primers di RNA permettendo alla DNAligasi di saldare le estremità di DNA.

#### Dna polimerasi negli eucarioti

**Esistono** diverse Dna polimerasi eucariotiche: **α, β, γ**. Queste sono presenti a diverse concentrazioni nelle varie fasi della vita cellulare, la  $\alpha$  pol è legata alla proliferazione dei tessuti, la  $\beta$  nei processi di riparo e la  $\gamma$  è presente soprattutto nei mitocondri.

La prima ad essere scoperta è stata la  $\alpha$  che copre l'80/90% di tutta l'attività polimerasica tot di solito è associata ad altre proteine che hanno attività diversa come l'attività primasica o' esonucleasica 3'→5' e 5'→3' con un complesso formato da 4 polipeptidi infatti da sola presenta attività esonucleasica. La dna polimerasi beta pesa 40 Kdal e la sua presenza aumenta in caso di stress o di esposizioni ai mutageni questo perché ripara il dna danneggiato, la dna polimerasi gamma non si trova nel nucleo ma nei mitocondri che hanno dna derivante da batteri. L'ultima a essere stata scoperta è la dna pol  $\gamma$  negli anni 70 che è presente a bassa concentrazione cell ed è l'unica polimerasica ad avere attività polimerasica in vitro su dna a singolo filamento con primer, funziona xò solo in presenza di ciclina o di PCNA perché è presente solo quando le celli stanno proliferando. La ciclina è un fattore di processività e forma un anello attorno al dna rendendo la poli delta molto processiva. Il replisoma è composto da due polimerasi, la poli delta e la poli alpha, la prima sintetizza la leading stand, la seconda i frammenti di okazaki perché è dotata di attività primasica.

#### Dna A, Dna B, Dna Z

Il dna è un polimero di nucleotidi, ognuno dei quali è formato da uno zucchero, un estere fosforico e una base azotata. L'insieme zucchero + base è chiamato nucleoside. Essi sono saldati tra loro dai gruppi fosfato che si legano in posizione 5' e 3'. Il complesso nucleoside + gruppo P è il nucleotide. Le purine sono A e G le pirimidine sono C e T. Il dna è composto da due catene antiparallele a forma di elica. Il fosfato si lega contemporaneamente allo zucchero in 5' e 3'. A condizioni fisiologiche si ha una struttura elicoidale con periodicità di 34 Armstrong e i rapporti purine/pirimidine tendono a uno. **Il B-DNA** esiste in condizioni fisiologiche (se lo guardi dall'alto si nota che non sono perfettamente centrale e l'elica non è simmetrica. Esiste un solo maggiore e uno minore, le basi sono tenute insieme da forze di vanderWaals. È connessa solo al DNA che in posizione 2' del deossiribosio ha solo l'H, è un acido ad andamento destrorso. Lo scheletro di zuccheri fosfati è molto flessibile e la struttura ciclica dello zucchero può assumere varie conformazioni. La posizione anti nel legame N-Glicosidico è favorita l'angolo di tilt è 6° e il passo 10 basi. **La A-DNA** ha l'elica con maggiore diametro perché assume non taglia le basi ma corre nel solo maggiore, la struttura è cava e le basi appaiono impilate + strette, l'angolo di tilt è accentuato (15/20°) il passo è maggiore (11 basi) e poco evidente la differenza fra solo maggiore e solo minore ed è + stabile in ambiente disidratato. **Lo Z-DNA** è favorito ad alte concentrazioni saline e superavvolgimenti andamento sinistrorso passo di 12 basi (elica + stretta) l'andamento dei legami glicosidici dello scheletro sono si e provocano andamento zig-zag, l'elica è molto compatta senza foro al centro e l'asse passa nel solo minore dell'elica. Questa forma ha sequenze G-C-G molto lunghe. Angolo di twist rotazione tra due coppie di basi sovrapposte, angolo di tilt è l'angolo tra i piani delle basi e la perpendicolare all'asse dell'elica, angolo di roll è l'angolo di rotazione rispetto all'asse del legame base-zucchero.

#### Dna satellite e seq altamente ripetute

Il DNA satellite è un DNA a sequenza semplice che consiste di molte ripetizioni in tandem di una breve unità fondamentale ripetuta.

Le sequenze ripetute, responsabili dei picchi di DNA satellite, sono costituite da milioni di ripetizioni di un segmento ancestrale più piccolo.

Il DNA satellite crea bande radioattive in varie regioni dei cromosomi. Si trova presente in tutti i cromosomi presso la zona dei centromeri (stabilizzano cromosomi, studiate nei protzoli. Telomeri: dna polimerasi che usa come stampo ma che è parte stessa dell'enzima che è una ribonucleoproteina per sintetizzare telomeri) e del centromero. La trascrizione di queste sequenze è estremamente bassa e la loro funzione è tuttora ignota. Si pensa possano svolgere un ruolo nell'organizzazione del cromosoma (come la condensazione).

Tramite gli enzimi di restrizione è possibile fare una mappa fisica del DNA, studiando il DNA satellite nei DNA brevi ed evidenziandone la presenza in un DNA genomico.

Frazione altamente ripetitiva del menoma consiste di copie multiple in tandem di unità ripetute molto brevi, che hanno spesso proprietà insolite. Una di queste è la possibilità di identificarle come un picco separato in un'analisi del DNA centrifugato in un gradiente di densità, che ha originato il nome di DNA satellite. Queste sequenze sono spesso associate a regioni inerti dei cromosomi e in particolare ai centromeri. Oltre alle sequenze satelliti, esistono tratti più corti di DNA che mostrano un comportamento simile, chiamati minisatelliti, utili in quanto possiedono un alto grado di divergenza fra i singoli genomi che può essere usato per la costruzione di mappe.

Il DNA ripetitivo è definito dalla sua relativamente rapida rinaturazione. Il componente che rigatura più rapidamente in un menoma eucariotico si chiama DNA altamente ripetitivo; questo tipo di componente presente in quasi tutti i genomi degli eucarioti superiori, ma la sua quantità totale è variabile (mammiferi 10%). La ripetizione in tandem di una breve sequenza spesso crea una frazione con proprietà fisiche distinte che possono essere usate per isolarla. In alcuni casi la sequenza ripetitiva ha una composizione in basi diversa da quella della media del menoma, quindi diversa densità di galleggiamento. Ricordiamo che il DNA satellite non è né trascritto né tradotto.

Quando si centrifuga DNA eucariotico su un gradiente di densità si possono distinguere due tipi di materiale: la maggior parte del menoma forma un continuum di frammenti che prende la forma di un largo picco centrato sulla densità di galleggiamento corrispondente al contenuto medio di GC nel menoma, talvolta si vedono uno o più ulteriori picchi più piccoli con valore diverso corrispondenti al DNA satellite. Il DNA satellite si trova in regioni di eterocromatina, che descrive le regioni dei cromosomi sempre strettamente avvolte e inerti a differenza dell'eucromatina che rappresenta la maggior parte del genoma. L'eterocromatina si trova nei centromeri e questo indica che il DNA satellite ha una qualche funzione strutturale nel cromosoma, funzione che potrebbe essere connessa al processo della segregazione dei cromosomi. Esperimento: proteine troviamo 8 proteine istoniche associate, da questo dato emerge un problema abbastanza interessante, in quanto le 8 proteine non sono lunghe a sufficienza per coprire un tratto così lungo di nucleotidi, e dunque il DNA non può essere contenuto all'interno delle stesse come in un involucro protettivo. Un primo indizio dei meccanismi implicati lo si ha tagliando il DNA con DNAsi I (pancreatici) che agisce sulla doppia elica e taglia un legame solo generando dei nicks. Se il DNA fosse coperto da proteine la DNAsi non potrebbe digerirlo per niente, mentre sperimentalmente si vede la comparsa dei nicks. I frammenti sono visibili in modo chiaro sul gel dopo aver denaturato il DNA (altrimenti rimangono impaccati con l'elica integra e non si vede alcuna differenza). La serie di bande che otteniamo è regolare, con una periodicità di 10 pb, pari ad un giro d'elica. Il DNA dunque non è all'interno delle proteine, ma vie è avvolto intorno.

#### Sintesi e maturazione dell'rRNA

Gli rRNA sono componenti essenziali di elementi indispensabili nella sintesi proteica: i ribosomi.

1-RIBOSOMA BATTERICO (70S-2.5Md-RNA 66%); composto da una subunità grande 50S e una piccola 30S. La subunità grande è costituita da rRNA 23S e 5S, la subunità piccola è invece costituita da rRNA 16S; più un numero noto di proteine.

2-RIBOSOMA EUKARIOTICO (80s-4.2Md-RNA 60%); composto da una subunità grande 60S e una piccola 40S. La subunità grande è costituita da rRNA 28S, 5.8S e 5S, la subunità piccola è invece costituita da rRNA 18S; più un numero noto di proteine.

I geni per gli rRNA eucariotici e procariotici sono trascritti come precursori grandi che devono essere processati per poter ottenere rRNA maturi.

#### rRNA PROCARIOTICI

La biosintesi dei rRNA nei batteri avviene tramite operoni dotati di due promotori e due terminatori e intervallati da geni per i tRNA. Il precursore è un unico tratto di RNA 30S che viene maturato in contemporanea alla sua sintesi. La RNAsi I riconosce i geni per i rRNA 16S e 23S e li excide. Le sequenze codificanti sono precedute e seguite da identiche sequenze palindromiche che consentono all'RNA di assumere una struttura a peduncolo lungo e stabile. La maturazione finale avverrà dopo l'associazione con le proteine ribosomiali specifiche che andranno a formare le subunità ribosomiali 30S e 50S mature (procarioti).

#### rRNA EUKARIOTICI

Anche negli eucarioti gli rRNA vengono trascritti insieme ma non si hanno le caratteristiche sequenze di rRNA intervallate tipiche dei procarioti. I geni per gli eucarioti sono ripetuti centinaia di volte e si trovano raggruppati nel nucleo della cellula. Caratteristica di tali geni è di essere separati da regioni chiamati "spaziatori" non trascritti (NTS). Dopo la trascrizione sia gli rRNA appaiono come precursori che presentano nucleotidi in eccesso alle loro estremità e anche nelle regioni codificanti che verranno separate nell'rRNA maturo. Tale accorciamento prende il nome di processamento, è simile allo splicing ma differisce poiché non vi è ricongiungimento delle porzioni di RNA, e avviene nel nucleolo, dove sono prodotti gli rRNA e dove sono assemblati i ribosomi.

#### PROCESSAMENTO:

- A partire da un precursore 45S si ha un taglio dello spaziatore all'estremità 5' con il rilascio di un intermedio da 41S
- Il taglio successivo spezza l'RNA 41S in due frammenti: il 32S e il 20S che contengono rispettivamente le sequenze 28S e 5.8S da una parte e 18S dall'altra
- Il precursore 32S viene tagliato per dare gli rRNA maturi 28S e 5.8S che si appaiano tra loro, mentre il 20S è degradato fino alla taglia matura 18S
- Il processamento è visibile mediante una tecnica di pulse-chase.
- Si prendono cellule di topo e si marciano per breve periodo con uridina-triziata. Una volta fatto ciò si trova che l'rRNA marcato sedimenta con un picco a 45S. Se si aggiunge un eccesso di uridina non marcata per diluire il nucleoside marcato

radioattivamente, si nota che la radioattività nel precursore 45S diminuisce man mano che aumenta la marcatura presente negli rRNA 18S e 28S. Ciò suggerisce che l'rRNA 45S è il precursore degli rRNA 18 e 28S.

Allo stesso modo è stato analizzato un eucariote inferiore: il lievito.

Nel lievito si ha un cluster di geni per gli rRNA che presenta un enhancer, un promotore e geni per gli rRNA 17S, 5.8S e 25S intervallati da sequenze spacer (non introni). Il prodotto di trascrizione è un unico rRNA 35S che dopo circa 5 minuti si scinde in un rRNA 20S e un rRNA 28S che, nel giro di un processo maturativo, perdono gli spacer convertendosi rispettivamente in rRNA 17S maturo e rRNA 5.8S più rRNA 25S maturi. Anche per studiare l'rRNA 30S, molto rapido nel maturare a 30°C, si può portare a 10°C e analizzare mediante tecnica del pulse-chase che sfrutta la mutazione precoce per marcare in modo selettivo il precursore del rRNA maturo.

Nel nucleolo le fasi di processamento sono dirette da una classe di piccoli RNA nucleari: snoRNA, che sono unite a molte proteine a formare piccole ribonucleoproteine nucleari snoRNP. In particolare partecipano al processamento tre snoRNP dette E1, E2, E3. Se tali proteine vengono inibite il processamento dell'rRNA si blocca in fasi delinente. Una caratteristica della struttura primaria dell'rRNA è la presenza di residui metilati. Ci sono circa 10 gruppi metilici nell'rRNA 16S e circa 20 nell'rRNA 23S. Nelle cellule eucariotiche gli rRNA 18 e 28S portano rispettivamente 43 e 74 gruppi metilici e quindi il 2% circa dei nucleotidi sono metilati.

Gli rRNA vengono metilati molto precocemente e questa caratteristica viene utilizzata per studiarli tramite radioattività. La metilazione non risulta casuale ma avviene solo su regioni che poi diventeranno rRNA maturo.

#### rRNA 16S BATTERICO

E' lungo circa 1500 nucleotidi e presenta una struttura secondaria complessa. Risulta essere una molecola ad elevatissima conservazione e le poche regioni in cui la sequenza non risulta conservata presentano comunque conservazione della struttura secondaria. Se viene paragonato con il suo analogo eucariotico, l'rRNA 18S, da cui è separato da 3 miliardi di anni di evoluzione, si nota una notevole omologia di sequenze e struttura. Essendo molecola a lentissima evoluzione, possono essere utilizzati come orologi molecolari evolutivi. L'rRNA 16S svolge importanti funzioni nella sintesi delle proteine: 1- il terminale 3' dell'rRNA 16S interagisce direttamente con il mRNA al momento dell'inizio.

2- regioni specifiche dell'rRNA 16S interagiscono direttamente con le regioni degli anticodoni del rRNA sia nel sito A che nel sito P.

3- l'interazione fra le subunità comporta interazioni fra gli rRNA 16S e 23S.

E' anche importante ricordare che l'attività peptidil transferasica risiede esclusivamente nell'rRNA 23S.

#### Rna pol II

La ma pol II utilizza un singolo filamento come template svolgendo la doppia elica durante la sua azione di polimerizzazione. È composta da un core di due subunità grandi + di due subunità medie e altre proteine ausiliarie tipiche degli eucarioti. È una situazione intermedia fra la RNA pol I e III. Il TATA BOX è una seq ricca di T e A che determina il promotore riconosciuto da questa polimerasi. Queste sequenze possono essere anche presenti in 2 o 3 copie consecutive e garantiscono un livello trascrizionale basale. La sequenza INR (formata da A A circondata da pirimidine) rappresenta l'inizio della trascrizione e insieme al TATA BOX costituisce la struttura basale del promotore. Le GC sono aumentano l'efficienza del promotore trovandosi a monte di esse. Negli eucarioti il gene regolato presenta delle seq CAAT e GC che influenzano il promotore prima della struttura secondaria. Se viene paragonato con il suo analogo eucariotico, l'rRNA 18S, da cui è separato da 3 miliardi di anni di evoluzione, si nota una notevole omologia di sequenze e struttura. Essendo molecola a lentissima evoluzione, possono essere utilizzati come orologi molecolari evolutivi. L'rRNA 16S svolge importanti funzioni nella sintesi delle proteine: 1- il terminale 3' dell'rRNA 16S interagisce direttamente con il mRNA al momento dell'inizio.

2- regioni specifiche dell'rRNA 16S interagiscono direttamente con le regioni degli anticodoni del rRNA sia nel sito A che nel sito P.

3- l'interazione fra le subunità comporta interazioni fra gli rRNA 16S e 23S.

E' anche importante ricordare che l'attività peptidil transferasica risiede esclusivamente nell'rRNA 23S.

#### Il codice genetico

È composto da triplette codificanti ed è universale. È stato determinato dal fatto che gli aa sono 20 e le basi 4, 4<sup>3</sup>=64 numero sufficiente per coprire ampiamente i 20 aa. Proprio per questo motivo, le triplette codificano per il triplo degli aa, il codice è definito dicodizzazione. La tripletta ha cmq un suo significato chimico fisico dato che 3 è il numero di basi necessarie perché avvenga la dicodizzazione.

Esperimento: necessari: ribosomi, triplette singole e 20 trna, uno per aa. La soluzione fu poi fatta filtrare su nitrocellulosa su cui si fermano tutti i ribosomi ma non i trna singoli. I trna invece legati ad un aminoacido si legano anche ai ribosomi. Quindi questi venivano fermati dal filtro e ne si poteva calcolare la radioattività. Sono stati scoperti 3 codoni di stop UAA UAG UGA. Il codice non è degenerato per tutti gli aminoacidi nello stesso modo, alcuni hanno triplette codificanti, altri meno. Inoltre l'aa è spesso determinato dalle prime 2 basi della tripletta, mentre nella terza vi è un'altra degenerazione. In altri casi la specificità è data dalla presenza in terza base di una purina o di una pirimidina. La alta degenerazione nella terza base è molto importante, perché nel 33% dei casi la mutazione avviene proprio in qst posizione, quindi il codice si rivela una buona barriera per rendere silenti le mutazioni puntiformi che avvengono in questa zona. In alcuni organismi sono state riscontrate delle deviazioni dal codice originale esse spesso sono a carico di codoni di stop e in alcuni casi sono sostituiti da trna soppressori che li rimuovono e vi mettono al loro posto un'altra base: es tetrahymena un fungo UAG E UAA codificano per GLU. Oppure nel mitocondrio UGA è il triptofano e non stop. AJA codifica per la metionina e AGG e AGA sono codoni di stop. Nel mitocondrio ci sono una ventina di trna diversi da quelli del citosol che servono a codificare proteine proprie del mitocondrio.

#### LE SSB Proteins – Quick stop

Appartiene alla categoria dei geni quickstop. Si nota che qst molecole si dispongono l'una di seguito all'altra su di un singolo filamento di dna, oppure si legano ai doppi filamenti provocando denaturazione parziale del Dna anche a 37°C. L'associazione è dovuta all'elevata cooperatività dei tetrameri e alla fluttuazione del Dna causata dall'apertura momentanea da zone ricche di A e T. Le SSB una volta associate al DNA tengono stabili le bolle di denaturazione che altrimenti tenderebbero a richiudersi spontaneamente.

Sono stati identificati in celli infettate da FAGI T4 in escherichia Coli le SSB pesano 18 kda e formano dei tetrameri. Mescolando i tetrameri SSB di coli a singoli filamenti di dna si osservano al microscopio elettronico degli impaccamenti dei filamenti in corrispondenza dei siti di attacco. Fenomeno visibile misurando l'assorbenza a 260 nm di un campione di dna a cui si aggiungono le proteine SSB a temperatura costante di 37°C; si nota dopo 10/20 minuti un aumento dell'assorbanza il che ci indica una parziale denaturazione del dna. Questo fenomeno è anomalo dato la temperatura del sistema, se viene aggiunto Mg 2+ la reazione è velocissima. Questo dato ci indica che la denaturazione è parziale questo fenomeno nasce da due fattori, il primo è il che c'è un'elevata cooperatività dei tetrameri, e l'altro è la fluttuazione del DNA in zone ricche di AT, in condizioni normali queste bolle si richiudono con proteine SSB invece le bolle diventano stabili. Ai confini delle bolle le prot si attaccano sempre d+ a favorite da affinità conseguente all'aumento del n° di SSB legato e questa ha un senso, infatti la funzione biologica delle SSB è tenere aperto il dna durante la replicazione.

#### Metodo studio interazione DNA-Proteine

Ci sono vari metodi per studiare le interazione tra dna e prot anche perché esistono diversi sistemi adottati dall'evologia per permettere queste interazioni.

**Zinc finger:** ovvero Zbeta shift antiparalleli e 1 alfa elica dove avviene la vera e propria interazione, lo zinco serve solo ad indicare la conformazione idonea all'interazione con il Dna, ma esistono proteine che non usano lo zinco per legare il DNA, come l'heix-turn-helix con 3 alfa eliche che interagiscono con il solo maggiore di dna, oppure domini omoleici che derivano dai geni omoleici che sono responsabili della mitogenesi e codificano per proteine che interagiscono con il dna controllando l'espressione dei geni.

**ZINC FINGER:** le proteine dotate di Zinc Finger non cristallizzano e pertanto risulta difficile studiarne la conformazione. Recentemente si è riusciti a farlo mediante nuove tecniche quali la Risonanza Magnetica Nucleare (NMR). I dati ottenuti ci indicano una conformazione di due j shift antiparalleli e una  $\alpha$  elica. L'interazione con il DNA avviene tramite la  $\alpha$  elica ed è presente un atomo ZN che serve ad indurre la conformazione strutturale adatta a formare il legame. Gli Zinc Finger più comuni hanno la conformazione C<sub>2</sub>H<sub>2</sub> (C = cys e H = his) ma in alcuni casi possono assumere anche conformazione C<sub>2</sub>C<sub>2</sub>. Tra le istidine e le cisteine si trovano circa 5 o 6 AA.

**Filter binding:** serve per determinare parametri termodinamici. Keq e costanti di velocità. È una procedura rapida e semplice ma non mostra il sito di interazione, però permette di misurare i dati termodinamici e cinetici.

**Footprinting:** è una procedura lunga (7gg) ma è l'unica che ci dà dati precisi sul sito di interaz e ci consente di valutare la specificità. Prendo un filamento di dna e marco il 5' uso fosfasi alcalina per togliere la radioattività da un lato. Quindi ottingo un primo 5' non è radioattivo taglio con enzimi di restrizione faccio un elettroforesi e faccio l'analisi dei gel di poliacrilammide. In questa tecnica si marca solo un'estremità di DNA un solo filamento si inserisce la RNAPol e si aggiunge dopo un po' di DNAsi a basse concentrazioni, e per breve tempo in modo da avere solo un taglio per complesso. I vari complessi saranno tagliati in posizioni casuali tranne nelle zone in cui il dna è associato alle proteine, si ottengono tante molecole di dna con nicks in diverse posizioni a questo punto si denatura il dna e se fa un elettroforesi di trazi a SS-DNA e si esaminano le bande radioattive in cui ci appariranno zone prive di bande che corrisponderanno a zone con DNA e proteine. Con questa tecnica siamo in grado di scoprire che le proteine coprono una zona di 60 pb e si è arrivato a definire una sequenza consensus dei Promotori che risulta essere altamente conservata, una sequenza è -10 (TATAAT) il secondo è a -35 (TTGACA) sono importanti solo qst sequenze ma anche la distanza che intercorre tra esse, la RNAPol deve interagire contemporaneamente con entrambe le seq, se allontaniamo la seq dal consensus il promotore viene riconosciuto lo stesso ma con un'efficacia minore.

**Gel retardation:** misura il ritardo di mobilità su gel che si ha come risultato di un legame di una proteina ad un tratto di DNA. Si separano i complessi di DNA-Proteina su gel di agarosio. In un pozzetto si carica DNA e nell'altro DNAsi+ Proteine. Le quantità sono nell'ordine del ng e senza radioattività non sarebbero visibili. La migrazione è rapida per la ridotta grandezza delle molecole. Se si verifica l'interazione Dna-proteine questo complesso migrerà + lento del solo dna campione. **Esperimento di competizione:** ad un campione di DNA\* si aggiunge dna di qualsiasi tipo; ma non marcato in cui le proteine possono legarsi in maniera aspecifica, ma non in modo specifico facendo permanere in soluzione il complesso DNA\*-PROTEINE. Se ho specificità posso costruire un oligonucleotide non marcato contenente la sequenza su cui avviene il legame e ottenere lo spiazzamento del DNA marcato dalla proteina. Si usa per isolare prot da una miscela di altre prot diverse e per preparare footprinting.

#### Ruolo e funzione della TFIIIA

Uno dei fattori fondamentali per il funzionamento della ma pol II. Insieme ai fattori TF II C e b riconosce i promotori BOX e costituito da una sola proteina che crea un sito di legame per la pol III. È necessario solo per la sintesi di 5S rRNA esso è costituito da una sola proteina che si lega in modo specifico solo ad una sequenza e consiste in una proteina che lega circa 10 atomi di zinco. Oltre a legare il dna è in grado di legare l'rna 5s e questa sua doppia specificità rappresenta una caratteristica peculiare che costituisce un sistema di regolazione. Infatti una volta sintetizzato l'rna 5s è incorporato nei ribosomi, ma se è presente in eccesso il fattore TFIIA A si lega ad esso e ne diminuisce la sintesi ribosomiale.

#### Sintesi e maturazione dell'mRNA

**MATURAZIONE, SINTESI E PROMOTORE:** mRNA  
Quando si parla di maturazione dell'mRNA si parla di eucarioti, in quanto l'mRNA procariotico non viene processato (la sintesi proteica avviene cotrasportivamente). L'mRNA eucariotico è prodotto nel nucleo e sintetizzato da RNA polimerasi II, dopo qualche passaggio di maturazione viene trasportato nel citosol. Esso va incontro, prima di tutto, all'aggiunta del CAP di 7metilguanidina da parte della guaniltransferasi. La funzione del CAP è quella di difendere l'mRNA dall'attacco di esosucclasi, ma soprattutto essa funge per il riconoscimento dell'mRNA da parte del ribosoma. Un'altra caratteristica degli mRNA è la coda di poli Adenina in 3' che viene anch'essa introdotta dopo la trascrizione da poli-A-polymerasi, la quale serve per la maturazione e media il trasporto nel citosol; la sequenza di poliA non è codificata dal DNA ma è aggiunta nel nucleo dopo la trascrizione. Il tratto di poliA dell'RNA nucleare e dell'mRNA è associato ad una proteina chiamata PABP; forme correlate di questa proteina si trovano in molti eucarioti. Un monomero di PABP di circa 70 kDa si lega ogni 10-20 basi della coda di poliA e quindi una caratteristica comune della maggior parte degli eucarioti è che l'estremità 3' dell'mRNA consiste di un tratto di poliA unito ad una grande massa di proteine. L'aggiunta del poliA avviene nel corso di una reazione in cui l'estremità 3' dell'mRNA viene generata e modificata da un complesso di enzimi. Di solito poliA stabilizza l'mRNA e lo protegge dalla degradazione anche grazie alle PABP che vi si sono legate.

Nelle cellule eucariotiche l'mRNA è sintetizzato nel nucleo ma viene tradotto nel citoplasma; esso passa nel citoplasma sotto forma di una particella ribonucleoproteica che è trasportata attraverso il poro nucleare e, una volta nel citosol è libera di associarsi ai ribosomi ed essere tradotta. È quindi probabile che la maggior parte del mRNA sia tradotta in posizioni casuali, determinate dal loro punto d'ingresso nel citosol e dalla distanza a cui possono essersi spostati da questo punto; alcuni mRNA invece, sono tradotti in punti specifici mediante diversi meccanismi. Gli mRNA eucariotici sono di solito stabili per parecchie ore; possono avere sequenze multiple che danno inizio alla degradazione; l'mRNA di lievito è degradato in almeno due vie: entrambe iniziano con la rimozione del poliA che provoca la perdita della proteina che lega poliA che a sua volta porta alla rimozione del cappuccio mutilato dall'estremità 5'. Una via degrada l'mRNA a partire dall'estremità 5' per mezzo di una esosucclasi, mentre l'altra lo degrada dall'estremità 3' per mezzo dell'esosoma, un complesso contenente parecchie esosucclasi.

L'mRNA fornisce un intermedio che porta la copia di una sequenza di DNA che rappresenta una proteina; esso ha la stessa funzione in tutte le cellule, ma esistono differenze importanti nei dettagli della sintesi e della struttura stessa negli eucarioti e nei procarioti. Una differenza importante dipende dal luogo in cui avvengono trascrizione e traduzione. Nei batteri l'mRNA è trascritto e tradotto nell'unico compartimento cellulare e i due processi sono collegati così strettamente da avvenire contemporaneamente. Poiché i ribosomi si attaccano all'mRNA batterico anche prima del completamento della sua trascrizione, è probabile che il poliosoma sia ancora attaccato al DNA. Di solito l'mRNA batterico è instabile e viene perciò tradotto in proteine soltanto per pochi minuti. In una cellula eucariotica la sintesi e la maturazione dell'mRNA avvengono esclusivamente nel nucleo ed è soltanto dopo il completamento di questi eventi che l'mRNA è esportato nel citoplasma, dove viene tradotto dai ribosomi. L'mRNA eucariotico è relativamente stabile e continua ad essere tradotto per parecchie ore.

#### Splicing del mRNA

L'azione dello splicing non è collegata alla trascrizione, in quanto in vitro estratti nucleari possono eseguire lo splicing di RNA precursori, ma le due operazioni avvengono in vivo in maniera coordinata, e altri eventi di modificazione possono influenzare lo splicing.

Il primo passaggio è un taglio al sito di splicing 5', che separa l'esone di sinistra e la molecola che contiene esone-introne di destra; l'esone di sinistra prende la forma di una molecola lineare, mentre la molecola di destra forma un cappio, in cui il terminale 5' generato all'estremità dell'introne si unisce con un legame 5'2' a una base dentro l'introne; la base bersaglio è una A, contenuta in una sequenza che viene chiamata sito di ramificazione. Il taglio al sito 3' rilascia l'introne libero sotto forma di cappio, e l'esone destro viene legato all'esone sinistro. Il taglio e la legatura avvengono come un unico trasferimento coordinato. Il cappio viene poi deramificato a formare una sequenza libera lineare, che viene poi degradata rapidamente. Come abbiamo detto, le sequenze necessarie per lo splicing sono le brevi sequenze consenso ai siti di splicing 5' e 3' e al sito di ramificazione; questo ha un ruolo importante nell'identificazione del sito di splicing 3', è altamente conservato e ha una sequenza UACUUAAC, ma negli eucarioti superiori non è così ben conservato con una precedenza però di purine o pirimidine in ciascuna posizione mantenendo il nucleotide bersaglio A, si trova 18-40 nct a monte del sito di splicing 3'. Mutazioni o delezione del sito di ramificazione nel lievito impediscono lo splicing; negli eucarioti superiori, le minori limitazioni di sequenza portano alla capacità di usare sequenze correlate. I siti critici, quando la ramificazione autentica è deleta (il sito critico è sempre vicino a quello autentico ed è usato soltanto quando il sito di ramificazione è stato inattivato). Da tutto ciò emerge che il sito di ramificazione ha il compito di identificare il sito di splicing 3' più vicino come bersaglio della connessione al sito di splicing 5', il che può essere spiegato dal fatto che si verifica un'interazione fra complessi proteici che si legano a questi due siti.

Il legame che forma il cappio va dalla posizione 5' della G invariante che si trovava all'estremità 5' dell'introne alla posizione 2' della A invariante nel sito di ramificazione, che corrisponde alla terza A nel box UACUUAAC (del lievito).

Le reazioni chimiche procedono per transesterificazione: un legame viene trasferito da una posizione all'altra; il primo passaggio è un attacco nucleofilo da parte del 2'OH della A invariante del sito di ramificazione su sito di splicing 5'. Nel secondo passaggio, il 3'OH libero dell'esone che era stato rilasciato dalla prima reazione attacca adesso il legame al sito di splicing 3'. Il numero di legami fosfodiesterici è conservato: in origine c'erano due legami 5'3' nei siti di splicing esone introne, uno dei quali è stato sostituito dal legame 5'3' tra esone e l'altro dal legame 5'2' che si forma nel cappio.

#### Come è stato isolato Ori-C e che funza?

Per scoprire Ori C, che è l'unica origine di replicazione del cromosoma batterico, è stata necessaria la tecnica del DNA ricombinante. Essa consiste nel costruire un vettore per espressione genica partendo da un plasmide naturale in cui si inseriscono dei geni che conferiscono resistenza ad alcuni antibiotici (di solito Amp<sup>r</sup>, resistenza all'Ampicillina, e Tet<sup>r</sup>, resistenza alla Tetracicclina). Questi consentono così, dopo la trasformazione e il trasferimento su terreno contenente questi antibiotici, di avere un metodo di selezione (screening). Soltanto i batteri che avranno incorporato il plasmide saranno in grado di generare colonie. I vettori posseggono un sito di restrizione che se tagliato e miscelato con il DNA batterico sottoposto allo stesso taglio, può ricoincaltarsi alle estremità di quest'ultimo ed esservi incorporato. In questo modo si ottengono dei ricombinanti che oltre al proprio DNA endogeno esprimono anche il DNA inserito nel plasmide.

Per costruire un vettore per isolare l'origine di replicazione è quindi necessario partire da un plasmide Pbr323 con marcatori Amp<sup>r</sup> e Tet<sup>r</sup> e Ori C. Utilizzo due enzimi di restrizione che esocidano il tratto in cui è presente Ori C ed ottengo così un tratto di DNA lineare contenente i due marcatori. Tratto poi il Genoma di Coli con gli stessi enzimi di restrizione ottengo vari frammenti tra i quali uno conterrà Ori C. Mescolo quindi il tratto di plasmide con i segmenti di DNA genomico e fornisco DNA ligasi e ATP. Tra le varie possibilità di legame è possibile che il plasmide si richiuda dopo avere incorporato il segmento con Ori C batterico. Il plasmide neofornato avrà caratteristiche nuove che derivano dalla nuova Origine di replicazione. Esso si replicherà una sola volta e non lo si potrà amplificare. Con questo metodo è dunque possibile studiare Ori C batterica. Utilizzando diversi enzimi di restrizione si è poi scoperto che questo tratto risulta essere di circa 250 pb. Si evidenziano su di esso delle sequenze ripetute sia a destra che a sinistra dell'origine di replicazione vera e propria. A destra si trovano 4 sequenze ripetute di 9 pb, riconosciute dal DNA-A, che risulta specifica per l'apertura delle forche (crea tensioni nei superavvolgimenti), essa però richiede l'uso di ATP. A sinistra vi sono invece 3 sequenze ripetute di 13pb particolarmente ricche di A e T e dunque facilmente apribili. Una volta che il DNA-A si è attaccato ed ha cominciato l'apertura della forca, intervengono le altre proteine replicative quali la DNA-B (elicasi), DNA-C (necessaria alla primasi), SSB e Primasi.

Negli eucarioti esistono vari fattori corrispondenti del DNA-B, DNA-C e altri. Il replisoma è composto da 2 polimerasi, pol I per la sintesi continua sul leading strand e la pol II con i suoi fattori per i frammenti di Okazaki. I fattori necessari sono RF-A, analogo alle SSB, RF-B e RF-C. Le origine di replicazione negli eucarioti, a differenza dai procarioti, sono molteplici; se ne trova una ogni 50/100.000 paia di nct. Esse sono costituite da sequenze ripetute prevalentemente di T con la presenza minoritaria di A. Queste sequenze si trovano in 3 copie per 100pb e sono chiamate ARS. A questo punto sorge spontaneo chiedersi, visto che il meccanismo di completamento della replicazione del DNA plasmidico è batterico è basato sull'incontro tra due forche replicative che marciano in direzioni opposte, come nel DNA lineare eucariotico sia garantita una replicazione fedele delle estremità. Si potrebbero infatti perdere gli ultimi tratti del cromosoma. La situazione è stata studiata in Saccharomyces e furono fatte numerose scoperte. Prima di tutto furono scoperti i vettori YIP (trasformano solo per via integrativa) e quelli YRP (capaci di replicazione autonoma). Le cellule che posseggono vettori YRP possono anche possederne diverse copie ma il fenomeno risulta eterogeneo perché, non avendo il centromero, il plasmide non partecipa alla mitosi e rimane sempre nella cellula madre. Anche all'aggiunta del centromero però, si nota che dopo pochi cicli di replicazione cellulare essi diventano comunque incapaci di crescere. Questo ci riconduce al problema della conservazione del tratto terminale del cromosoma; in questo caso infatti, il cromosoma neo-trasfettato perde ad ogni replicazione un segmento terminale di DNA e in capo a pochi cicli risulta incapace di svolgere la sua funzione. Si è così scoperto che i cromosomi autosomici sono dotati nelle loro regioni terminali di una specifica struttura detta Telomero. Esso è formato da sequenze ripetute centinaia di volte (CCCCAA) e protegge così il DNA informativo contenuto nella parte terminale. Da solo però non è sufficiente a svolgere questo compito ed è quindi affiancato da un enzima che ricostruisce le sequenze telomeriche, la Telomerasi. Queste conoscenze sulla struttura dei telomeri sono state utilizzate per creare gli YAC (Yeast Artificial Chromosome) che vengono utilizzati per clonare sequenze molto grandi di DNA eucariotici; essi costituiscono il contenuto delle banche genomiche.

#### Ruolo del TF2D

È il primo elemento che si lega perché il promotore della pol II funziona. È un complesso di 700kDa. La proteina che si lega effettivamente è la TBP (TATA binding protein) è un dimero formato da due parti simili e interagisce con il passo minore dell'elica. Per funzionare in maniera completa la TBP necessita degli elementi TAF che sommati ad essa danno il TFIID. Sono elementi molto eterogenei e formano un grappolo attorno al TBP e sono responsabili dell'interazione con le proteine che legano le UAS.

#### Attivazione degli aminoacidi

La formazione del legame peptidico non è termo dinamicamente favorita, quindi perché l'aa venga attivato è necessaria nel ribosoma la presenza di un enzima l'aminoacil-t-mna-sintetasi che il COO- esterificato con il ribosio (reaz spontanea anche a T ambiente). L'attivazione conferisce anche specificità alla reazione dato che l'esterificazione avviene solo con il t-mna specifico per il codone che codifica per l'aa voluto. La sintesi proteica prevede tre fasi: 1- inizio 2- allungamento 3- termine. Per ogni aa incorporato nella prot sono consumati 4 ATP di cui 2 per l'attivazione dell'aa per 2 per la sintesi.

L'attivazione degli aa prevede due fasi:

- 1- AA + atp + enzima = enzima-aa-amp + PPI
- 2- E-AA-AMP + tma = AA-t-ma + AMP+ E

L'aminoacil-t-mna sintetasi è l'enzima che è specifico per un solo aa e ad esso associato il t-mna corretto.

La frequenza di errore dovrebbe essere di 1/200. In realtà l'enzima possiede una attività di proof-reading grazie a cui riesce a correggere l'errore. Se si dà val e lle all'enzima, all'arrivo del t-ma ci può essere due reazioni in competizione tra loro: o caricamento di valiloleucina-t-ma da evitare o una fase di correzione in cui si rompe il legame tra Va e AMP. Le due reazioni sono in contrasto e sono verificate un aa correttamente allora la reazione di caricamento sarà + breve di quella di idrolisi, se si ha un caricamento errato caricherà il contrario. Quindi di solito non si carica l'aa perché dovrebbero avvenire 2 eventi casuali in sequenza (bassa prob).

#### Come si misura l'attività trascrizionale

La trascrizione dell'mna prevede l'attività dell'rna pol e la presenza di un template di dna da usare come stampo per la sintesi di un filamento complementare di rna. L'unità di trascrizione è composta dalle DPE, ovvero distanti promoter enancher, ovvero sequenze a monte del promotore che possono modificare l'attività con la RNA pol dal promotore appunto che precede la ORF a cui si legano i fattori di trascrizione ed infine il terminator.

Si può misurare dal punto di vista quantitativo e qualitativo: qualitativo serve per calcolare se un gene è espresso o meno (nothern blot) o per stabilire da dove parte la trascrizione (marcatura con nucleasi S1: mi serve per localizzare l'estremità 3' o 5' degli rna e per quantifica la quantità di un dato rna nelle cell in un determinato momento, anche la primer extension serve per mappare l'estremità 5' del trascritto con accuratezza) e (i primer extension), quantitativo permette di misurare i livelli di mna, attività di un promotore, la stabilità di un mna e l'attività di trascrizione (geni reporter, run off, degradazione mna). La tecnica del northern blot si utilizza x studiare quale mna sia presente in un determinato istante nella cellula, si procede estraendo ma dalla cellula, si separa su gel di agarosio e si ibrida su filtro di nitrocellulosa. Il gel viene posto in una soluzione tampone gli viene poggiato sopra il foglio di nitrocellulosa e viene pressato con dei pesi. La carta assorbe il tampone attraverso il gel e la nitrocellulosa che cattura gli acidi nucleici e che verrà usato per essere ibridato con da marcatore complementare al gene di cui si vuole conoscere l'espressione.

Si miscela con una sonda di dna marcato con 32P con un foglio di nitrocellulosa su cui è legato l'mna estratto da norther blot. Si ottiene un etero duplex mna dna marcata, si esegue una autoradiografia del complesso per valutare l'espressione del gene, la grandezza del suo mna e la sua quantità.

La tecnica dei geni reporter si utilizza per studiare l'attività specifica di un promotore, si isola un promotore del gene per poter studiare attraverso modificazioni della sua sequenza e attivarsi a capire quali parti svolgono determinate funzioni. Dopo averlo estratto, modificato e clonato lo si reinserisce nel cromosoma originario e ne si esamina il comportamento. In base alla quantità di sintesi di mna che si ottiene si può conoscere la funzionalità che la regione mutata ha sul ruolo delle espressioni del gene. I + comuni geni reporter sono beta-gal per i batteri e lacZ per i lieviti.

La tecnica del run-off su nuclei isolati permette di vedere i geni espressi in un monomero, si recuperano nuclei da cell lisati ai quali si inserirà nucleotidi marcati con 32P. I nuclei porteranno a termine la trascrizione e l'mna prodotto risulta marcato radio attivamente.

#### Use dei geni reporter

La tecnica dei geni reporter si utilizza per studiare l'attività specifica di un promotore tramite un norther blot si calcola la quantità di mna derivante dalla sua velocità di sintesi e degradazione. Una delle due velocità viene modificata, non siamo però in grado di sapere la causa e il modo, quindi come è regolata l'attività specifica di un gene. Infatti il promotore di un gene può essere molto complesso e può agire influenzando da molti fattori. Per comprendere il funzionamento delle varie componenti di un promotore è quindi necessario isolarlo per poterlo studiare attraverso modificazioni della sua sequenza. Successivamente gli si apportano modificazioni lo si clona e infine lo si reinserisce nel cromosoma originario e ne si esamina il comportamento: se la sintesi dell'mna è simile o identica a quella originale, significa che le modificazioni apportate a quella determinata regione sono state poco influenti, quindi quella zona del promotore ha un'importanza funzionale relativamente bassa (o nulla) al contrario se la quantità varia, significa che ho apportato delle modifiche ad una regione importante nell'espressione del gene. Tra i geni reporter + comunemente utilizzati troviamo il gene beta gal, che codifica per la beta galattosidasi batterica, il gene lac-z che rende le cellule che lo esprimono di colore blu in presenza di X-gal; nei mammiferi si tende ad utilizzare CAT o luciferasi.

#### Separazione RNA cellulari e analisi

Esistono due tecniche di separazione:

-ultracentrifugazione in gradiente di saccarosio, in cui le molecole si separano in base al coefficiente di sedimentazione. La centrifugazione è di breve durata perché l'ma + denso del saccarosio non finisce in fondo alla provetta. Quindi con uno spettrofotometro si misura l'assorbanza. Il processo è utile per grossi quantitativi di rna.

-elettroforesi su gel (gel di agarosio). Si lavora in condizioni denaturanti perché se tutte le molecole denaturano l'unico fatto discriminante sia la lunghezza della molecole. Per fare sì che accada si utilizzano delle sostanze che impediscono la formazione di legami H, inizialmente si usava l'urea 6M ma l'urea tende a cristallizzare quindi ora si usano denaturanti chimici che formano legami con l'ma impedendo la formazione di strutture secondarie, ad esempio gli aldi, che interagiscono con i gruppi OH formando emiacetali. Per quantizzare i diversi tipi di rna sintetizzati nella cell si utilizza il sistema dell'pulse and chase (sintesi rna). Il procedimento in toto prevede: mettere dell'uracile marcato insieme a delle cellule, quindi prelevare un'aliquota e metterla in ghiaccio. Il dna, l'rna e i metoni preciptano. La quantità di ma marcato aumenta esponenzialmente, ma se si fornice U non marcato la rna ed il metonare cala e l'ma sintetizzato non incorpora + U marcato (pulse & chase) questo sistema fa capire il destino di molecole di ma sintetizzate in un certo momento. Il chase è il blocco immediato dell'assunzione di molecole marcate ad opera di molecole non marcate.

L'mna non è visibile ed è presente in piccole quantità (5% ma tot). Metodo ottimo per i batteri perché un piccolo ma non per gli eucarioti dove per bloccare l'assunzione di U è necessario un inibitore della RNA POL come l'actinomicina che inibisce ttt le ma polimerasi svolgendo così l'azione del chase, il problema è che gli inibitori sono tossici che creano problemi metabolici che inducono la morte cellulare.

#### Rna polimerasiche eucariotiche

Mentre nei procarioti esiste una sola ma pol, negli eucarioti esistono 3 tipi che svolgono diverse funzioni con localizzazione nucleare (pol I) e nucleo plasmatiche (II e III). Le ma pol sono enzimi molto complessi ma di una certa regolarità. Presentano 2 subunità grandi simili a quelle alpha di E.coli, costituite da prot di tipo a e costituiscono 2/3 del peso della polimerasi, inoltre sono presenti sub unità + leggere e diverse per ciascun enzima e infine sub unità + leggere.

L'ma pol I trascrive solo geni per gli rna con un solo tipo di promotore. Ci sono + copie del gene alternate a sequenza spaziarizi e organizzate in cluster. Il promotore del gene enzimatico si trova tra -45 e +20 ma la trascrizione è regolata anche da seq upstream. Gli eucarioti presentano gli istoni di cromatina che potrebbero interagire con le prot. Accessorie. Negli eucarioti poi ci sono un' altra RNA POL codificata dal genoma mitocondriale e cloroplastico, quello cloroplastico è simile a quella batterica mentre quello mitocondriale è monomero. Esistono quindi due tipi di ma pol: monomeriche ovvero molto piccole, molto efficienti che leggono pochi mitocondri e le multimeriche: molto grosse e complesse riconoscono proteine diverse.

L'ma pol II utilizza un singolo filamento

#### Denaturazione del DNA

Una tecnica per denaturare il DNA è l'utilizzo del Calore, per evidenziarlo basta porre la soluzione nello Spettrofotometro e ne aumento gradualmente la temperatura. Inizialmente l'assorbanza si mantiene costante perché il DNA non varia la sua struttura, dopo una certa temperatura in poi l'assorbanza inizia ad aumentare fino a stabilirsi a un valore rispetto a quello iniziale qui il DNA si è completamente denaturato.

Per questi studi si utilizzano molecole lunghe al max 500 nucleotidi in quanto se si aumentasse la lunghezza la soluzione diventerebbe troppo viscosa. Un altro fattore che influenza la denatura è la FORZA IONICA del mezzo, infatti la doppia elica è molto + stabile in soluzione con elevata forza ionica, questo accade perché i Sal neutralizzano le cariche negative presenti su DNA diminuendo la repulsione tra i fosfati che è uno dei fattori k destabilizzano la doppia elica.

Anche la concentrazione di CG fa variare la temperatura di fusione, infatti generalmente la denaturazione non avviene in modo omogeneo, le regioni AT si denaturano anche a temperatura ambiente reversibilmente con il calore la denaturazione si estende a dare Bolle di denaturazione fino a raggiungere zone ricche di CG che sono le ultime a denaturarsi grazie alla loro maggiore forza di legame. La curva di denaturazione ha dunque un andamento sigmoidale non regolare.

Se avviene un abbassamento della T si può avere rinaturazione in quanto il DNA possiede solo una struttura 2aria e dunque ha meno possibili conformazioni di rinaturazione, il riappiamento, può avvenire solo se il raffreddamento è lento e graduale, se invece è rapido otteniamo una diminuzione dell'assorbanza senza però mai raggiungere i livelli del dna nativo. Questo fenomeno si chiama nucleazione, perché la nucleazione avvenga basta un allineamento di sole 3pb, quest' ha un DG > 0 x le prime 3 basi, poi diventa estremamente favorito. Nel caso di una corretta rinaturazione questa fase diventa estremamente favorita.

#### Replicazione DNA e Forza replicativa

Il DNA immagazzina, come già detto, il codice genetico nelle cellule viventi. L'informazione genetica è codificata nelle sequenze particolari di basi presenti che determinano, a loro volta, l'esatta sequenza in cui i vari aminoacidi devono comparire nelle proteine sintetizzate come, ad esempio, un enzima. La trasmissione del codice genetico da una cellula all'altra avviene attraverso una duplicazione esatta della catena del DNA. La replicazione del DNA implica la separazione dei due filamenti della doppia elica, ognuno dei quali funziona come un vero e proprio stampo che determina la sequenza di basi dei due nuovi filamenti che si vanno formando. Immediatamente prima della divisione cellulare, la doppia elica di DNA comincia a svolgere una delle due estremità e su ciascuna delle due catene iniziano a formarsi le due catene complementari. Questo tipo di replicazione, schematizzata in Figura 27, è considerata di tipo conservativo poiché in ognuna delle due nuove doppie eliche è presente un filamento proveniente dall'elica originale. In realtà, la replicazione implica una serie piuttosto complessa di reazioni catalizzate da enzimi diversi. A titolo di esempio, l'enzima La sintesi di entrambe le due

nuove catene di DNA avviene nella direzione (estremità 5' → estremità 3'). Ricordando che le due catene originali corrono parallele, ma in direzioni opposte, si ha come conseguenza che al bivio di replicazione (vedi Figura 27) una delle due nuove catene deve presentare libero un OH all'estremità 3' e l'altra nuova catena un OH all'estremità 5'. È interessante notare

che, mentre il complementare del filamento 3' → 5' viene sintetizzato come una catena continua senza interruzioni, il complementare dell'altro filamento viene prodotto come una serie di corte catene. Enzimi specifici provvedono poi a saldare tra di loro questi segmenti e a formare il nuovo filamento completo di DNA.

Il dna si replica a partire dalla doppia elica tramite una replicazione semiconservativa, infatti dopo la sintesi otteniamo due filamenti ibridi. Per ciascuno contiene un filamento nuovo e uno vecchio che ha funzionato da stampo. Questo è possibile xkè la sintesi avviene in 2 fasi, la prima è una fase di denaturazione e la seconda di duplicazione. La prima fase di denaturazione ha inizio nelle origini di replicazione, strutture a occhio nudo la cui esistenza è stata dimostrata tramite **esperimenti di meselson e stahl** i quali marcarono le cellule di Colli affinché il dna si replicasse 1 volta e mezzo. A questo punto la lisarono e stratificarono il dna su di un filtro per effettuare un esame radiografico, notarono che la radioattività dovuta ai marcatori era presente in tutto il dna ma c'erano zone dove vi si concentrava maggiormente a causa della presenza dei doppi filamenti. Sul cromosoma batterico si aprono 2 forche replicative che procedono in senso opposto in maniera contemporanea. Nei cromosomi eucariotici poi si aprono + forche in modo che la duplicazione risulti altrettanto veloce nonostante il genoma sia + grande.

Le varie zone di replicazione si chiamano REPLICONI e la replicazione si interrompe quando 2 repliconi si incontrano. Per verificare ciò si sono presi cellule di Fibroblasti di cricotto e le se sono inibite in fase S, così facendo le si è sincronizzate tutto l'inibitore gli si è fornita TIMINA\* per 10 minuti, estratto il DNA e filtro per autoradiografia. Si sono notati diversi punti di origine nello stesso filamento. È importante ricordare che la replicazione avviene in maniera discontinua sulle 2 eliche, il filamento guida e quello in ritardo dovuto al fatto che le DNAPOLI sintetizzano il dna solo in 5'→3'.

### Sintesi e maturazione del T-Rna

La sintesi dei trna nei procarioti e negli eucarioti è diversa. Nei batteri infatti i geni per i trna sono all'interno di **operoni raggruppati in cluster** dove sono raggruppate anche + copie dello stesso trna e trna diverse. Dopo che vengono sintetizzati tutti i trna devono andare incontro alla maturazione. Negli eucarioti invece la sintesi non prevede la presenza di operoni, che non esistono mai negli eucarioti, ci sono invece geni ripetuti + volte, +tosto corti che vengono trascritti dalla pol III che riconosce un promotore interno e che produce un prodotto piuttosto lungo che dovrà andare incontro alla maturazione e allo splicing in quanto a differenza dei batteri i trna appena prodotti presentano introni ovvero regioni brevi aggiunte in 5' o in 3' che vanno eliminate per avere trna funzionali. Tutti i trna siano essi eucariotici che proca vanno incontro a maturazione ad opera delle ribonucleasi che riconoscono le strutture secondarie dei trna appena vengono sintetizzati, le rmasi di tipo P tagliano il trna al 5' rendendo così un'estremità già a posto, per i procarioti e per gli eucarioti (fibr v' è un enzima molto simile alla masi P) per l'estremità 3' invece nei procarioti esistono geni C a 2 classi, la classe I che accorpa tutti quei geni che presentano nella sequenza il tratto CCA e quindi la loro maturazione avviene da parte di esonucleasi che tagliano sino a trovare tale sequenza, e la classe II che raggruppa quei geni che non presentando la seq CCA vengono digerite l'estremità 3' da endonucleasi blunt che tagliano l'estremità 3' lasciando il blunt end dove una trna nucleotidiltrasferasi attaccherà tale sequenza, di questa classe fanno parte tutti i geni eucariotici. Per gli eucarioti inoltre c'è anche un passaggio ulteriore, ovvero lo splicing, infatti i trna maturato è di circa 70/80 nucleotidi, mentre il precursore 100/120 nct, lo splicing dell'introne del trna eucariotico avviene a livello della punta dell'anticodone subito dopo infatti è presente quest seq complementare a quella dell'anticodone in modo da permettere la struttura a quadrifoglio, avviene quindi uno splicing citosolico in 2 fasi: nella prima fase l'introne viene exciso da una endonucleasi che lo libera nel citosol, nella 2° fase invece avviene la saldatura dei 2 terminali 3'oh e 5'-p03- con consumo di ATP

### Lo splicing del tRNA

Lo splicing dei geni per il tRNA avviene con un meccanismo che si basa su reazioni separate di taglio e legatura. Gli introni presenti nei geni per il tRNA correlati hanno una sequenza correlata, ma gli introni dei geni del tRNA che rappresentano AA diversi non sono correlati: non esiste una sequenza consenso che potrebbe essere riconosciuta dagli enzimi di splicing. Tutti gli introni includono una sequenza che è complementare all'anticodone del trna che crea una conformazione alternativa del braccio dell'anticodone in cui l'anticodone si appia per formare un'estensione del solito braccio. La sequenza esatta e le dimensioni dell'introne non sono importanti e la maggior parte delle mutazioni che lo riguardano non impediscono lo splicing: la reazione dipende quindi principalmente dal riconoscimento di una struttura secondaria comune presente nei tRNA e non da una sequenza comune nell'introne. Quest'ultimo non è però del tutto irrilevante: l'appaiamento fra una base dell'ansa dell'introne e una base non appaiata dello stelo è necessario per lo splicing, in quanto mutazioni che influenzano questo appaiamento influenzano anche lo splicing.

La reazione in vitro richiede ATP e la caratterizzazione delle reazioni che avvengono con e senza ATP dimostra che due stadi separati della reazione sono catalizzati da due enzimi diversi.

Il primo passaggio non richiede ATP e comporta il taglio di un legame fosfodiesterico da parte di una reazione nucleasica alipica, catalizzata da una endonucleasi, che è responsabile della specificità del riconoscimento degli introni e taglia il precursore a entrambe le estremità dell'introne. Il secondo passaggio richiede invece ATP e comporta la formazione di un legame: si tratta di una reazione di legatura e l'enzima responsabile di questa attività è l'RNA ligasi.

I prodotti di taglio sono un introne lineare e due mezze molecole di tRNA, ciascun terminale 5' termina con un gruppo ossidrilico e ciascun terminale 3' termina in un gruppo 2'P ciclico. I due mezz tRNA formano coppie di basi per produrre una struttura simile al tRNA e quando si aggiunge ATP avviene la seconda reazione, in cui devono essere alterate entrambe le estremità insolite formate dalla endonucleasi.

Il gruppo P ciclico è aperto per formare un terminale 2'P in una reazione che richiede un'attività fosfodiesterasica ciclica: il prodotto ha un gruppo 2'P e uno 3'OH. Il gruppo 5'OH generato dalla nucleasi deve invece essere fosforilato per dare un 5'P; questo genera un sito in cui il 3'OH si trova vicino al 5'P. l'integrità covalente della catena polinucleotidica è ripristinata da una ligasi.

Tutte etre le attività (fosfodiesterasi, polinucleotidasi chinase e adenilato sintetasi) sono disposte in domini funzionali diversi su una singola proteina e agiscono in sequenza.

La molecola sottoposta a splicing non è più interrotta e possiede un legame 3'5'P al sito di splicing e un gruppo 2'P che deve essere rimosso da una fosfatasi.

La reazione di splicing del tRNA usa taglio e sintesi di legami ed è determinata da sequenze che sono esterne all'introne. Altre reazioni di splicing usano transesterificazione, in cui i legami sono trasferiti direttamente e le sequenze necessarie per la reazione si trovano nell'introne.

### SEQUENZE RIPETUTE ED ENZIMI DI RESTRIZIONE

Le sequenze ripetute sono costituite da milioni di ripetizioni di un segmento ancestrale più piccolo. Possiamo utilizzare queste sequenze per vedere in quali regioni dei cromosomi si trovano e per vedere se esistono RNA a loro complementari. I dati sperimentali indicano che la trascrizione di queste sequenze è bassissima, e negli esperimenti di ibridazione RNA-DNA\* non si ha ibridazione e dunque le sequenze non sono dei geni. La funzione di queste sequenze è tutt'ora ignota e si pensa abbiano un ruolo nell'organizzazione del cromosoma.

Un metodo di studio delle sequenze ripetute è l'utilizzo di enzimi di restrizione, i quali sono nucleasi in grado di tagliare il DNA in corrispondenza di sequenze specifiche, rompendo il legame fosfodiesterico all'interno o all'esterno del filamento mediante idrolisi; nei procarioti sono abbinate a Metilasi che metta le sequenze di taglio. Esistono tre diverse classi di enzimi di restrizione: Nucleasi di tipo I che riconoscono la sequenza e tagliano a caso molte basi più a valle (non permette di prevedere la sequenza tagliata); Nucleasi di tipo II che tagliano in corrispondenza della sequenza riconosciuta; Nucleasi di tipo III che riconoscono sequenze e tagliano a distanza di 22-24 basi prima della sequenza (taglio ancora casuale). Tramite gli enzimi di restrizione è possibile fare una mappa fisica del DNA.

Le endonucleasi di restrizione sono uno strumento chiave per la mappatura del DNA; queste tecniche possono essere applicate all'RNA a singolo filamento, facendo una copia di DNA a doppio filamento. Si può ottenere una mappa fisica di qualsiasi molecola di DNA spezzandola in punti definiti la cui distanza può essere determinata accuratamente; questi tagli specifici sono possibili appunto grazie alle endonucleasi, che hanno la capacità di riconoscere come bersaglio per il taglio sequenza di DNA duplex piuttosto brevi, 4-6 bp; enzimi di restrizioni diversi hanno sequenze bersaglio differenti, e oggi è disponibile una vasta gamma di queste attività.

Una mappa di restrizione rappresenta una sequenza lineare dei siti in corrispondenza dei quali enzimi di restrizione particolari trovano i loro bersagli; la distanza su queste mappe si misura in coppie di basi (bp).

Quando viene tagliata, una molecola di DNA si divide in frammenti di dimensioni diverse che possono essere separati mediante elettroforesi su gel; analizzando i frammenti di restrizione del DNA si genera così una mappa della molecola originale, che contiene le posizioni in cui è avvenuto il taglio e le distanze tra i diversi siti di taglio.

**METODI DI PURIFICAZIONE:** bisogna ottenere un lisato e trattare con solventi organici (bisogna includere negli inibitori delle ribonucleasi per evitare la digestione dell'Rna) per togliere le proteine e precipitare i nucleotidi con etanolo. Il precipitato contiene Dna e Rna, bisogna quindi purificare con CsCl<sub>2</sub> (che forma un gradiente di circa 36h), si può purificare più velocemente creando un gradiente discontinuo.

Rna estratto contiene il 5% mRna, 70% rRna, 25% tRna

Metodo del gradiente di saccarosio (basso potere risolutivo):

- 1) Rna Stabili: molecole che durano a lungo e durante la loro crescita non vengono degradate
- 2) Rna Labili (mRna), non visibile poiché non mostra picchi

Metodo dell'elettroforesi: viene utilizzato gel d'agarosio e condotta in condizioni denaturanti per tenere l'Rna a singolo frammento:

- 1) l'Urea 5M o formamide (20%); capaci di formare molti legami idrogeno e mantenere così l'Rna denaturato
- 2) Glicosale o formaldeide (10%): le aldeidi sono denaturanti perché: con OH formano acetati, con NH2 formano le immine, entrambe con processi reversibili

### DOMINI DI LEGAME PER IL DNA

Le proteine che agiscono sul DNA possiedono dei domini che permettono il legame con esso. Esistono vari tipi di questi domini come lo Zinc Finger, Helix-Turner-Helix, il B-Zip e i Domini Omeotici.

**ZINC FINGER:** le proteine dotate di Zinc Finger non cristallizzano e pertanto risulta difficile studiarne la conformazione. Recentemente si è riusciti a farlo mediante nuove tecniche quali la Risonanza Magnetica Nucleare (NMR). I dati ottenuti ci indicano una conformazione di due jg shift antiparalleli e una α elica. L'interazione con il DNA avviene tramite la α elica ed è presente un atomo Zn che serve ad indurre la conformazione strutturale adatta a formare il legame. Gli Zinc Finger più comuni hanno la conformazione C<sub>2</sub>H<sub>2</sub> (C = cys e H = his) ma in alcuni casi possono assumere anche conformazione C<sub>2</sub>C<sub>2</sub>. Tra le istidine e le cisteine si trovano circa 5 o 6 AA.

**DOMINI OMEOTICI E HELIX-TURNER-HELIX:** Riguardo ai domini omeotici essi derivano dai geni omeotici scoperti inizialmente in Drosophila; tutti questi geni presentano regioni di alta omologia e conservazione detti Domini Omeobox. Il motivo è molto simile al sistema Helix-Turner-Helix, con la terza elica che interagisce con il solco maggiore del DNA; l'orientazione corretta è indotta principalmente da interazioni con lo scheletro zucchero-fosfato e non si ha riconoscimento di una specifica sequenza.

### Cinetiche di rinaturazione

Si parla di cinetiche di rinaturazione in quanto, mantenendo costante la temperatura, la rinaturazione è solo in funzione del tempo. Dopo aver frammentato, denaturato e poi rinaturato il DNA a particolari condizioni (25°C sotto Tm, NaCl 0,18 M) si misurano dei parametri:

$$dC/dt = -KC^2 \quad (C = \text{concentrazione DNA denaturato})$$

$$dC/C^2 = -Kdt$$

che integrato per  $t = 0$  e  $C = C_0$  ( $C_0$  = concentrazione iniziale filamento singolo) da

$$C/C_0 = 1/1 + KC_0t$$

Quando  $C = \frac{1}{2}$  anche  $t = \frac{1}{2}$  e abbiamo:

$$C/C_0 = 1/2 = 1/1 + K C_0 t/2 \text{ da cui}$$

$$C_0 t/2 = 1/K$$

ovvero l'inverso della costante di rinaturazione K.

Esistono però sostanziali differenze nella velocità di rinaturazione del DNA nelle varie specie. Infatti diversi DNA differiscono per il  $C_0 t/2$  e ciò a causa non della lunghezza ma della loro composizione in basi; in particolare a causa della loro complessità X (sequenza minima ripetuta/AAAAAAA compl 1, ATCATCATC compl 3). Con questo sistema è possibile definire lunghezze e complessità dei geni di batteri e virus (per gli eucarioti risulta più difficile perché hanno sequenze ripetute vastissime). Di solito per effettuare l'analisi si utilizza DNA campione di riferimento usando la seguente formula:  $C_0 t/2 \text{ (DNA stud)} = \text{compl DNA stud} \times \frac{C_0 t/2 \text{ (DNA campione)}}{\text{compl DNA campione}}$

Da analisi come queste si è giunti infine a scoprire le Sequenze Altamente Ripetute (100.000 copie per genoma in su) e quelle Mediamente Ripetute (100/1000 in su).

Il tempo di rinaturazione varia da specie a specie: infatti il dna di un mammifero è 300 volte + lento rispetto a quello di un batterio che è a sua volta 20 volte + lento di quello di un virus. Il Dna poli AT rinatura istantaneamente, il fatto che diversi DNA differiscano tra loro per  $C_0 t/2$  non dipende dalla loro lunghezza ma dalla loro complessità intesa come lunghezza di seq minima ripetute (poli A è di complessità 1, poli AT è di complessità 2, seq minima 10<sup>5</sup> è di complessità 10<sup>5</sup>). La complessità si indica con X e vale circa  $X \approx C_0 t/2$  ovvero  $X \approx K$ , in condizioni standard A vale  $5 \times 10^5$ . Questo si può utilizzare per stabilire la lunghezza dei geni dei batteri e dei virus, ma per misurare quella di organismi più complessi questo metodo non funziona +. Per effettuare l'analisi si usa un DNA campione (E. Coli  $X = 4.2 \times 10^6$ ) di riferimento utilizzando la seguente formula:

$(C_0 t/2 \text{ del dna studiato}) / (C_0 t/2 \text{ dna campione} = \text{compl. Dna studiato}) / (4.2 \times 10^6)$ . Si ottiene una curva di rinaturazione divisa in almeno 3 porzioni. Il  $C_0 t/2$  della terza curva è quella che si riferisce alla complessità del genoma; le sequenze della curva 1 hanno tempi brevi di rinaturazione: sono le sequenze altamente ripetute; quelle della curva 2 sono le sequenze mediamente ripetute.

Da questo tipo di analisi possiamo ricavare : le dimensioni del genoma ( $C_0 t/2$  sequenze uniche); frazioni; la frazione di sequenze ripetute; la frequenza di ripetizione; la complessità.

Le frazioni genomiche dei diversi tipi di dna sono in genere del tipo : 25% altamente ripetute, 30 % mediamente ripetute, e 45% uniche, ma variano molto in base al genoma considerato.

Dal grafico misuro i valori di  $C_0 t/2$  per ricavare la complessità del genoma ottengo:

- 0,0013 mol x sec / litro altamente rip
- 1,9 mediamente
- 630 uniche.

Da cui ricavo la complessità. 240, 600'000, 300'000'000. E la frequenza di ripetiz: 500'000 veloce (630 0,0013), 350 media , 1 lenta.

Analisi numerica:

-Freq genomiche:

alfa= seq uniche 45%

Beta= med. Ripetute 30%

Gamma = altamente ripetute 25%

-Cot2 del singolo componente puro

$Cot2 = f(\alpha) \times Cot2(\text{miscela}) + f(\alpha) \text{ è la fraz genoma}$

-complessità di ciascun componente

Per le seq non ripetute:  $Nnr = \alpha \times X \times G$  (G = dimensione del genoma)

O

$Nnr = [(4.2 \times 10^6) \times (C_0 t/2 \text{ (E.coli)}) \times C_0 t/2 \text{ (soluzione)}$

Per le seq ripetute:  $Nr = (\beta \times X) / \beta \times F$  = freq di ripetizione

**B-ZIP\*** è una regione costituita da un α elica molto lunga composta da una Leucina intervallata da sei aminoacidi in maniera tale che le Leucine risultino incolonnate nella e sopra le altre dalla stessa parte dell'elica. Queste proteine funzionano solo in forma dimerica per il fatto che si viene a formare una struttura coil-coil a forma di zip stabilizzata da legami di Van der Waals. La sequenza di Leucine è preceduta da regioni basiche. Le due molecole tendono ad elicitare leggermente l'una sull'altra, dato che il passo della struttura è 3 e non 3.6. La B-Zip forma una specie di pinza che blocca il DNA. Sono state isolate proteine che controllano geni per il differenziamento che presentano regioni di legame al DNA detti Bzip (jun, fos, oncogeni).

### TRASCRIZIONE, ORF

Ogni sequenza ha tre possibili modi di lettura, a seconda del punto di partenza; solitamente ne viene usato solamente uno per la traduzione, mentre gli altri possibili due sono bloccati da frequenti segnali di termine.

Un modulo di lettura che consiste esclusivamente di triplette che rappresentano aminoacidi si chiama open reading frame, ORF; il modulo di lettura inizia con uno specifico codone di inizio, AUG e si estende attraverso una serie di triplette che rappresentano AA fino a terminare con uno dei tre codoni di stop, ossia UAG, UAA, UGA.

Un modulo di lettura si dice bloccato quando non può essere letto in proteina perché sono presenti numerosi codoni di termine. Se una sequenza è bloccata in tutti e tre i moduli di lettura, essa non può avere la funzione di codificare per una proteina. È inoltre improbabile che un modulo di lettura esista per caso, in quanto se non fosse tradotto in proteina non ci sarebbe stata una pressione selettiva ad impedire l'accumulo di codoni di termine.

### DIMOSTRARE CHE UN DNA SUPERAVVOLTO NEGATIVAMENTE E' UGUALE AD UNA MOLECOLA CON UNA REGIONE DENATURATA

Per fare questa dimostrazione è necessario servirsi di una relazione che spiega il numero di superavvolgimenti di un tratto di DNA circolare:

$$L = T + W$$

in cui L corrisponde al numero totale di avvolgimenti della molecola di DNA. T corrisponde al numero di avvolgimenti dovuti alla struttura secondaria e, infine, W indica il numero di supereliche (superavvolgimenti) presenti nella molecola. In tutti i genomi è presente un grado di superavvolgimento pari a 1 ogni 200 b.p. circa: questo genera una certa energia (circa 9 Kcal/mol) che viene ceduta quando il DNA torna allo stato rilassato.

Se consideriamo ad esempio un DNA rilassato di 360 b.p. con un passo dell'elica di 10b.p. possiamo scrivere che:  $T = 360/10 = 36$  (numero di superavvolgimenti dovuti alla struttura secondaria). In questo caso W è uguale a zero in quanto ho considerato un DNA rilassato (che, quindi, non presenta superavvolgimenti); da ciò ne deriva che L deve essere uguale a T (cioè il numero di avvolgimenti totali corrisponde al numero di avvolgimenti dovuti alla struttura secondaria).

Se, ora, consideriamo lo stesso DNA, trattato, però, con un enzima che taglia i legami creando 4 superavvolgimenti negativi (per esempio potrebbe essere una topoisomerasi di tipo II, la quale ad ogni azione produce due superavvolgimenti negativi) abbiamo che:

$T = 36$  poiché il DNA è lo stesso di prima (il passo dell'elica e il numero totale di b.p. non cambiano) e  $W = -4$  (numero di supereliche formate dall'enzima). Da ciò ne consegue che

$$L = 36 - 4 = 32 \text{ cioè si è ridotto il numero di avvolgimenti totali, ovvero il numero totale dei legami all'interno del DNA.}$$

Se si annulla W e si considera una denaturazione complessiva di 40b.p. (necessaria per annullare l'effetto di quattro superavvolgimenti negativi) si osserva che T cambia il suo valore da 36 a 32 (si sciogliono i legami tra 40 b.p.), mentre, ora, W risulta essere uguale a 32 (32 - 0).

Vediamo, in questo modo, che il DNA avendo quattro superavvolgimenti negativi corrisponde ad un DNA con 40 b.p. denaturate (poiché nei due casi si ha lo stesso valore di L) e si possono, dunque, definire topologicamente equivalenti. Infatti, il DNA necessita di energia per potersi denaturare localmente e questa energia è fornita proprio dai superavvolgimento negativo.

**STRUTTURA PRIMARIA E SECONDARIA DEL DNA**

Vi è una certa analogia strutturale fra acidi nucleici e proteine. Così come le proteine sono polimeri lineari di aminoacidi, gli acidi nucleici sono polimeri lineari i cui mattoni costituenti sono i *nucleotidi*. Queste unità monomere possono essere ottenute per degradazione in condizioni blande degli acidi nucleici, mentre un'idrolisi completa (acida, basica o enzimatica) ne fornisce i componenti fondamentali: questi sono rappresentati da uno zucchero (il D-ribosio nel caso dell'RNA e il deossiribosio nel caso del DNA, entrambi in forma furanosica), una base eterociclica (purinica o pirimidinica) e uno ione fosfato. La struttura del  $\beta$ -D-ribosio e del  $\beta$ -2-deossiribosio è mostrata in Figura 20 (il termine deossi sta semplicemente a indicare che il gruppo OH in posizione 2 è sostituito da un idrogeno). Esse sono l'*adenina*, la *guanina*, la *citosina* e la *timina*. Le basi che si ottengono dall'RNA sono le stesse eccettuata la *timina* che è sostituita dall'*uracile*, anch'esso mostrato in Figura 21. Un *nucleoside* si ottiene combinando una molecola di D-ribosio (RNA) o di 2-deossiribosio (DNA) con una base eterociclica azotata tramite un legame  $\beta$ -N-glicosidico.

In Figura 22 sono mostrate le strutture di due nucleosidi: uno derivato dall'RNA ed uno dal DNA. Il primo (adenosina) contiene infatti come unità mosaccaridica il D-ribosio legato alla base adenina e il secondo (deossiadenosina) contiene il 2-deossi-D-ribosio legato alla medesima base. Gli altri nucleosidi associati al DNA sono la deossiguanosina (2-deossi-D-ribosio + guanina), la deossicitidina (2-deossi-D-ribosio + citosina), la deossitimidina (2-deossi-D-ribosio + timina). Analogamente, i rimanenti nucleosidi per l'RNA sono la guanosina (D-ribosio + guanina), la citidina (D-ribosio + timina) e l'uridina (D-ribosio + uracile). Un *nucleotide*, che costituisce la vera unità monomera dell'acido nucleico, può essere ottenuto da un nucleoside esterificando una molecola di acido fosforico con un ossidrilico libero della componente monosaccaridica, normalmente l'ossidrilico attaccato al carbonio C3' o C5' (nei nucleotidi gli atomi di carbonio dello zucchero sono indicati con i numeri 1', 2', 3', ...). La struttura generale di un nucleotide del DNA, in cui l'esterificazione è avvenuta sul gruppo OH in posizione 5', è mostrata in Figura 23.

Nel seguito di questa sezione focalizzeremo la nostra attenzione soprattutto sul DNA di cui esamineremo a grandi linee la struttura fondamentale e il meccanismo di replicazione.

**1. La struttura primaria degli acidi nucleici**

Esaminiamo ora la sequenza con cui i vari monomeri si succedono lungo la catena di un acido nucleico a formare lo *scheletro covalente* o *struttura primaria* del polimero. Se consideriamo un filamento di DNA, lo scheletro è costituito da unità alternate di deossiribosio legato ad una base (nucleoside), e di fosfato. L'acido fosforico è doppiamente esterificato e forma un ponte fra il carbonio C3' di un nucleoside e il carbonio C5' di un altro nucleoside. Le basi eterocicliche, legate al carbonio anomero C1', si distaccano dall'ossatura del polimero costituita dall'alternarsi di unità di acido fosforico e di zucchero. Una rappresentazione schematica della struttura primaria di un tratto di filamento di DNA è riportata in Figura 24. La *sequenza delle basi*, ovvero l'ordine con cui le varie basi si succedono lungo la catena, può essere specificato elencando semplicemente le lettere che indicano le singole basi, partendo, a sinistra, dall'estremità della catena con il gruppo OH in posizione 5' libero e procedendo verso destra in direzione della posizione 3' libera. I simboli che vengono utilizzati per le varie basi sono le lettere (maiuscole) iniziali del nome: A=adenina, G=guanina, C=citosina, T=timina, U=uracile.

**2. Struttura secondaria del DNA**

Come nel caso delle proteine anche gli acidi nucleici sono caratterizzati da una *struttura secondaria*. Un modello per la struttura secondaria del DNA fu proposto nel 1953 da J.D. Watson e F.H.C. Crick e fu confermato nel 1962 da M. Wilkins tramite tecniche ai raggi X. Secondo questo modello due lunghe catene di DNA sono avvolte in una *doppia elica* attorno allo stesso asse. Le due catene polimeriche sono disposte parallelamente l'una all'altra, ma corrono in direzione opposta: questo significa che a entrambe le estremità della doppia elica è presente l'estremità 5' di un filamento e l'estremità 3' dell'altro filamento. Le due eliche destrorse sono orientate in modo tale da avere lo scheletro di zucchero e fosfato nella parte esterna della struttura con le basi rivolte verso l'interno. Legami idrogeno fra le basi tengono insieme la doppia elica. L'ampiezza massima della spirale è di 20 Å. Dieci coppie successive di nucleotidi, con dieci coppie di basi, danno luogo ad un giro completo della spirale. Due diverse rappresentazioni molto schematiche dei due filamenti della doppia elica che corrono paralleli, ma in senso opposto, con le varie basi accoppiate fra di loro, sono date in Figura 25.

I legami idrogeno che tengono insieme la doppia elica non possono formarsi in modo casuale, ma devono formarsi solo fra specifiche coppie di basi. Le coppie consentite nel caso del DNA sono: *adenina-timina (A-T)* e *citosina-guanina (C-G)*. Le due basi adenina e timina possono interagire tramite due forti legami idrogeno, mentre citosina e guanina possono formare tre legami idrogeno. L'effetto stabilizzante complessivo dovuto a queste interazioni è di circa 10 kcal/mol nel primo caso e circa 17 kcal/mol nel secondo. Le due coppie di basi considerate con l'indicazione dei legami idrogeno e i valori dei relativi parametri strutturali sono mostrate in Figura 26. Altri accoppiamenti non sono consentiti perché, o si avrebbero coppie di basi troppo voluminose o i legami idrogeno che si verrebbero a formare non sarebbero altrettanto forti e stabilizzanti. Questo fatto fa sì che le due catene della doppia elica siano del tutto complementari: ogni volta che una molecola di adenina (A) compare in una catena, una timina (T) deve essere presente nella corrispondente posizione sulla catena opposta. In modo analogo una guanina (G) da una parte implica la presenza di una citosina (C) dall'altra. Questo carattere complementare delle catene è evidente dai due schemi di Figura 25.

La struttura terziaria invece sono i cosiddetti *superavvolgimenti* del DNA. Il *superavvolgimento* è diverso se si tratta di DNA procariotico, o eucariotico. Nel primo caso il DNA è circolare e possiamo avere un *superavvolgimento positivo* (destrorso) oppure *negativo* (sinistrorso). Insomma il DNA si avvolge su se stesso. Nel secondo caso invece il DNA eucariotico è associato a delle proteine chiamate *ISTONI*. Questa associazione forma la cosiddetta *cromatina* ed è possibile grazie all'interazione tra DNA con carica totale negativa e istoni con carica totale positiva.