

Espressione genica: traduzione (parte 2)

La decifrazione del codice

Agli inizi degli anni Sessanta, i biologi molecolari sono riusciti a «decrittare» il codice genetico. Il problema in cui erano impegnati li lasciava perplessi: come è possibile scrivere 20 «parole» con un alfabeto di sole quattro «lettere»? Ovvero, come fanno le quattro basi (A, U, G, C) a specificare 20 diversi amminoacidi? Una possibilità era un codice a triplete, basato su codoni di tre lettere. Disponendo di sole quattro lettere (A, U, G, C), chiaramente un codice a una sola lettera poteva codificare in modo non ambiguo soltanto quattro amminoacidi, e non 20. Un codice a due lettere ne avrebbe codificati $4 \times 4 = 16$, ancora troppo pochi. Ma un codice a triplete avrebbe potuto render conto di $4 \times 4 \times 4 = 64$ codoni, più che sufficienti per 20 amminoacidi.

Il primo passo verso la decodificazione è stato compiuto nel 1961 dai biochimici Marshall W. Nirenberg e J. Heinrich Matthaei, quando capirono che come messaggero potevano usare un semplice polinucleotide artificiale invece che un mRNA naturale, ben più complesso; riuscirono quindi a identificare il polipeptide codificato da tale messaggero artificiale. I due ricercatori prepararono un mRNA artificiale in cui tutte le basi erano costituite dall'uracile (un mRNA sintetico detto, appunto, poli U). Aggiungendo un poli U in una provetta contenente gli ingredienti necessari alla sintesi proteica, si formò una catena polipeptidica tutta composta da un solo tipo di amminoacido: la fenilalanina. Dunque un poli U codificava la fenilalanina; di conseguenza, UUU era la parola in codice - il codone - per specificare la fenilalanina (Figura 1).

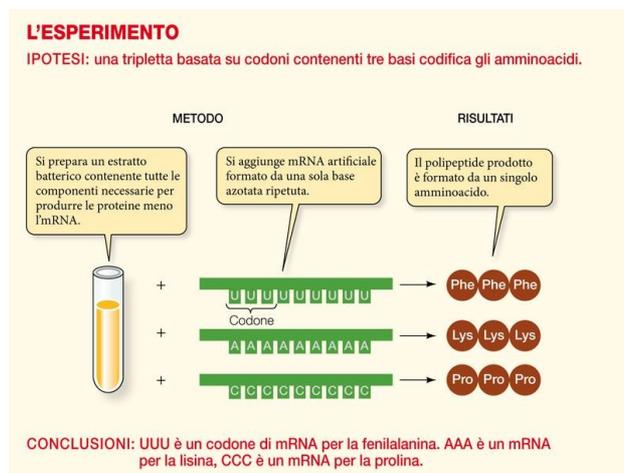


Figura 1. La decifrazione del codice genetico. Nirenberg e Matthaei usarono un sistema di sintesi in vitro per determinare gli amminoacidi specificati da mRNA sintetici di composizione conosciuta.

Sulla scia di questo successo, Nirenberg e Matthaei dimostrarono ben presto che CCC codifica la prolina e AAA codifica la lisina (poli G presentava qualche problema dal punto di vista chimico e inizialmente non fu preso in esame). UUU, CCC e AAA erano tre codoni fra i più facili; per venire a capo degli altri fu necessario modificare l'approccio sperimentale. In seguito, altri scienziati hanno scoperto che era possibile legare a un ribosoma semplici mRNA artificiali lunghi tre sole basi (ciascuno equivalente a un codone) e che il complesso risultante induceva la formazione di un legame fra il tRNA corrispondente e il suo amminoacido specifico. Così, per esempio, un semplice UUU faceva legare al ribosoma il tRNA che trasportava la fenilalanina. Dopo questa scoperta è stato relativamente semplice decifrare l'intero codice genetico. Per scoprire l'amminoacido rappresentato da un certo codone, Nirenberg ha ripetuto il suo esperimento usando un campione di mRNA artificiale con quel codone e ha osservato quale amminoacido si andava a legare.

La sintesi proteica nei procarioti e negli eucarioti

La sintesi proteica è il processo biochimico attraverso il quale l'informazione genetica contenuta nell'mRNA (RNA messaggero), viene convertita in proteine che svolgono nella cellula un'ampia gamma di funzioni. La sintesi proteica inizia da un filamento di mRNA, prodotto a partire da un gene sul DNA attraverso il processo di trascrizione. Questo filamento è usato come stampo per la produzione di una specifica proteina.

mRNA: è il trascritto di un gene nel linguaggio dell'RNA, che prevede quattro basi, identiche al DNA fatta eccezione per l'uracile che sostituisce la timina. Un mRNA maturo appena esportato dalla nucleoplasma attraverso la membrana nucleare è stato inoltre modificato mediante aggiunta di una 5'-metilguanina all'estremità 5' (cappuccio dell'mRNA), mediante una coda poliadenilata (AAAAA...) all'estremità 3' e associato a diverse tipologie di proteine utili a far riconoscere al resto della cellula l'avvenuta maturazione ed esportazione di questo fuori dal nucleo. Il ribosoma legge le basi dell'mRNA a triplette (dette codoni) e a ciascuna tripletta (per un totale di 64, che si ottiene elevando il numero delle basi dell'mRNA, cioè 4, al numero di cifre di una tripletta, 3) fa corrispondere un amminoacido rispettando il codice genetico. In realtà, essendo il codice ridondante, alcune triplette codificano per lo stesso amminoacido e solo due amminoacidi sono specificati da una sola tripletta (AUG per la metionina e UGG per il triptofano), esistono inoltre triplette che specificano la fine della traduzione, dette codoni di stop, sono UAA, UAG e UGA, esse non codificano per nessun amminoacido. Il codice genetico è identico in tutti gli organismi viventi con poche eccezioni, come *Candida albicans* e il DNA mitocondriale di alcuni organismi, che si è sviluppato in un certo qual modo indipendentemente da quello contenuto nel nucleo all'inizio della storia evolutiva. La lunghezza di un mRNA varia in funzione della lunghezza degli esoni del gene trascritto.

tRNA: è una piccola molecola di RNA composta mediamente da 80 nucleotidi. Oltre ai nucleotidi convenzionali per l'RNA vi sono anche nucleotidi modificati (61 tipi differenti) tra cui i più abbondanti sono l'inosina, la pseudouridina, la diidrouridina, la 4-tiouridina e varie forme dimetilate della guanina. Ha una caratteristica forma simile a quella di un quadrifoglio (la forma è data dal fatto che i nucleotidi che lo compongono sono appaiati in modo tale da originare una molecola con quattro sporgenze; due di queste sono più importanti: la prima è costituita da due nucleotidi non appaiati che sono sempre gli stessi per tutte le molecole di tRNA, cioè A C C; la seconda molecola è opposta alla prima, è costituita da una tripletta di basi anch'esse non appaiate che sono diverse per ciascuno dei 61 diversi tipi di tRNA) (o più realisticamente ad una L), che assume per la formazione di legami idrogeno tra le doppie eliche ripiegate in alcuni tratti della sua struttura.

Di norma i nucleotidi modificati tendono a non prendere parte a questi legami. Vi sono quattro aree di particolare interesse in ciascuna molecola di tRNA, in particolare due anse, dette ansa D e ansa T (spesso qui si concentrano i nucleotidi modificati), l'anticodone, specifico per ciascun tRNA e complementare ad un codone del mRNA (per esempio AGC letta da 5' a 3' sull'mRNA codifica per serina, dunque il tRNA corrispondente avrà un anticodone GCU, letto da 3' a 5' come si conviene) e l'attacco dell'amminoacido corrispondente al codone all'estremità 3'. Esistono più tRNA per uno stesso amminoacido (infatti vi sono solo 48 anticodoni diversi codificati da circa 500 geni) e nel contempo un tRNA può associarsi a più di un codone.

Questo accade perché ciascun anticodone di tRNA si associa saldamente solo alle prime due basi di un codone, mentre è poco specifico e più tollerante per la terza base, tanto che non raramente si verificano appaiamenti sbagliati, causando il fenomeno dello wobbling cioè il tentennamento della terza posizione. Questo spiega anche perché la serie di codoni che specifica uno stesso amminoacido è identica nelle prime due basi e differisce solo nella terza, con poche eccezioni. I tRNA che intervengono nella sintesi proteica sono il prodotto di pre-tRNA più lunghi che vengono modificati nel nucleo grazie ad uno speciale splicing che segue un meccanismo "taglia e cuci", diverso da quello comune che prevede la formazione di una sorta di cappio per rimuovere le sequenze introniche. Successivamente un amminoacil-tRNA-sintetasi specifica per ciascun amminoacido (circa 20 tipi diversi) accoppia covalentemente l'amminoacido corrispondente all'estremità 3' di ciascun tRNA. Prima di essere legato all'estremità 3' di un tRNA ciascun amminoacido è attivato da ATP che vi lega un AMP formando un amminoacido adenilato e liberando pirofosfato. Successivamente l'amminoacido adenilato è legato dall'amminoacil-tRNA-sintetasi al tRNA con fuoriuscita di AMP. L'amminoacil-tRNA-sintetasi sceglie l'amminoacido corretto da attaccare al tRNA in parte perché ha un'affinità maggiore per quell'amminoacido rispetto a tutti gli altri e in parte perché ne è facilitata dato che il suo sito attivo esclude tutti gli altri amminoacidi più grandi. Successivamente, quando anche tRNA si lega all'enzima, l'amminoacido si sposta in un secondo sito attivo ancora più specifico che funge da meccanismo di controllo. Un meccanismo simile avviene con la correzione esonucleolitica delle bozze da parte della DNA polimerasi. In questo modo la amminoacil-tRNA-sintetasi raggiunge un'accuratezza di 1 errore ogni 40.000 accoppiamenti.

Il ribosoma è un grande ribozima e rappresenta la macchina esecutrice della sintesi proteica. Possiede due subunità, la maggiore (60 S) e la minore (40 S), la maggiore contiene gli rRNA 28 S, 5.8 S e 5 S, mentre la minore contiene l'rRNA 18 S. Gli rRNA così come i ribosomi sono sintetizzati nel nucleolo all'interno del nucleo, i primi dalla RNA polimerasi III. Oltre ai quattro tipi di rRNA specificati, ciascun ribosoma contiene circa 50 proteine diverse; è dunque composto per i 2/3 da RNA e per 1/3 da proteine. La subunità minore funge da sostegno e da sito di ingresso dell'mRNA da tradurre, mentre la subunità maggiore è la principale macchina catalizzatrice del complesso. Quando sono unite le due subunità possiedono quattro siti d'attacco, uno per l'mRNA e i siti E, P, A per i tRNA in arrivo o in uscita. I primi due sono i più voluminosi e contengono tRNA i cui anticodoni sono legati al codone dell'mRNA mentre il sito E contiene il tRNA che si sta per staccare dal ribosoma poiché ha già aggiunto il suo amminoacido al polipeptide in formazione. Un ribosoma umano aggiunge circa 2 amminoacidi al secondo. Ciascuna cellula contiene alcuni milioni di ribosomi sparsi nel citoplasma con le subunità distaccate fra loro, oppure in attiva sintesi con le subunità unite, aggregati in polisomi (gruppi di decine di ribosomi in sintesi proteica) oppure attaccati alla membrana del reticolo endoplasmatico rugoso (RER). Le cellule in attiva sintesi ne contengono di più rispetto ad altre meno attive.

Sintesi proteica nei procarioti

Nei procarioti, il corretto legame tra il ribosoma e l'mRNA è facilitato dall'accoppiamento di una serie di basi nota come sequenza di Shine-Dalgarno, che si trova tra 5 e 10 nucleotidi prima del codone di avvio. Il tRNA, sia che rechi metionina o N-formil-metionina, accoppia le sue basi con quelle del codone di avvio coadiuvato da dei fattori di inizio (IF) e si lega al sito P (peptide) del ribosoma formando un ponte tra la subunità minore e la subunità maggiore. La sub-unità maggiore forma quindi un complesso con quella minore e i fattori di inizio vengono liberati. A questo punto avviene l'allungamento. Un nuovo aminoacil-tRNA complessato con il fattore di allungamento (una proteina GTP dipendente, EF-Tu nei procarioti, α EF1 negli eucarioti, più semplicemente Tu) entra sul sito A del ribosoma ed accoppia le sue basi con quelle dell'mRNA. La subunità ribosomiale maggiore possiede azione peptidil-transferasica, grazie alla quale crea un legame peptidico tra gli amminoacidi vicini. Appena questo accade, l'amminoacido sul sito P si stacca dal suo tRNA (a questo punto definiamo tRNA Scarico, ossia privo di amminoacido) e la catena peptidica in crescita si lega al tRNA sul sito A formando un dipeptide .

Il ribosoma quindi si muove lungo l'mRNA Avviene così il processo di traslocazione in cui ,sapendo che la lettura dei codoni avviene a triplette ,il dipeptide posto nel sito A torna va in P ed il deacilato in P migra sul sito E.A questo punto si libera il tRNA Scarico con il suo anticodone . Introduciamo l'ultimo tRNA Avente come amminoacido la serina (Ser) e l' unico sito che ha complementare il codone complementare è il sito A .Sempre ad opera del Ef-Tu avviene il secondo legame peptidico con la già formata Met-Fen che magrano nel sito A formando il tripeptide. Per la seconda volta avviene la traslocazione al quarto condone con

rilascio del tRNA scarico o deacilato in P. Questo processo è noto come traduzione o espressione genica.

Sintesi proteica negli eucarioti

La sintesi proteica inizia sempre dall'N-terminale di una proteina in formazione verso il suo C-terminale. Il primo amminoacil-tRNA ad essere aggiunto è invariabilmente quello con legata metionina, detto tRNA iniziatore, poiché la sequenza d'inizio di ciascun mRNA è AUG. Generalmente la metionina è sempre rimossa alla fine della traduzione da una proteasi specifica. La subunità minore del ribosoma si attacca all'estremità 5' dell'mRNA che viene riconosciuta poiché possiede un cappuccio di 7-metilguanosa e i fattori di inizio eIF4G e eIF4E, poi, una volta trovata, vi si attacca anche la subunità maggiore e quindi il ribozima cerca la prima tripletta AUG sull'mRNA che è il suo segnale d'inizio della sintesi proteica. In qualche caso la prima AUG può essere saltata e il ribosoma può iniziare a tradurre dalla seconda o dalla terza, questo permette in certi casi la produzione di proteine alternative a partire da uno stesso mRNA maturo. Tale fenomeno è comunemente noto come leaky scanning.

L'attacco del tRNA iniziatore al sito P della subunità minore del ribosoma in assenza di quella maggiore è aiutato dal fattore di inizio eIF2 che lega GTP, il quale è idrolizzato quando la subunità maggiore si associa alla minore. Il movimento della subunità minore è ulteriormente facilitato dall'idrolisi di ATP in ADP+P. Il tRNA si lega al sito P del ribosoma, mentre il tRNA con l'anticodone corrispondente al codone successivo nell'mRNA legato al fattore d'allungamento EF1 si attacca nel sito A. EF1 unisce il tRNA al sito A ed effettua un controllo di qualità per non venire aggiunto se la corrispondenza codone-anticodone non è corretta; probabilmente questo avviene per una maggiore affinità tra le molecole, anche se non è stato ancora chiarito precisamente il meccanismo. EF1 inoltre lega il tRNA sull'mRNA in una conformazione curva che impedisce l'immediato legame dell'amminoacido legato al tRNA cui è attaccato al resto del polipeptide in crescita. Una volta riconosciuto un appaiamento corretto, l'rRNA sulla subunità minore idrolizza il GTP legato a EF1 a GDP e si stacca dal tRNA. Sembra che una corrispondenza corretta tra codone ed anticodone sia ulteriormente facilitata dalla formazione di legami idrogeno tra la subunità minore del ribosoma e la stessa coppia codone-anticodone. A questo punto vi sono quindi sul ribosoma due tRNA adiacenti uniti ciascuno al proprio codone sull'mRNA. A questo punto la peptidil transferasi, contenuta nella subunità maggiore del ribosoma, catalizza lo spostamento del legame che unisce l'amminoacido al suo tRNA nel sito P formando un legame peptidico tra l'amminoacido del tRNA presso il sito P e quello presso il sito A.

La subunità maggiore si sposta quindi di tre nucleotidi verso l'estremità 3' dell'mRNA e così fa anche la subunità minore così che, alla fine degli spostamenti, il primo tRNA aggiunto si trova nel sito E e il secondo nel sito P mentre nel sito A si lega il fattore di allungamento EF2 con legato GTP. Il primo tRNA ormai privo di amminoacido si stacca dal sito E ed esce dal ribosoma,

mentre un nuovo amminoacil-tRNA si attacca al sito A, il GTP di EF2 è idrolizzato a GDP e EF2 si stacca dal sito A. Quindi il legame tra amminoacido e tRNA del sito P è trasferito tra gli amminoacidi dei tRNA tra A e P per formare un nuovo legame peptidico e così la sintesi va avanti sino a che il ribosoma non trova un codone di stop. Quando l' Rna si duplica permette alla cellula di far avvenire un processo di splicing. Quando il codone di stop è raggiunto nella fase di terminazione il ribosoma cattura una molecola d'acqua che idrolizza il polipeptide ormai completo il quale si distacca dal ribosoma. Il processo è coadiuvato da fattori di rilascio (RF1, RF2, RF3 nei batteri, RF1 ed RF3 negli eucarioti), proteine che simulano l'azione del tRNA, si legano al sito A e liberano la proteina nel citoplasma (fenomeno conosciuto come mimetismo molecolare). La sintesi delle proteine può avvenire molto rapidamente perché più ribosomi possono legarsi ad uno stesso filamento di mRNA consentendo quindi la costruzione simultanea di più catene proteiche. Un filamento di mRNA con più ribosomi è chiamata polisoma.

Inibitori della sintesi proteica

È possibile bloccare specificamente la sintesi proteica facendo usare inibitori specifici quali l'anisomicina e la cicloesimide. La traduzione può anche essere bloccata per effetto di mutazioni genetiche come ad esempio le mutazioni con slittamento di fase (frameshift) le quali, possono essere ottenute con la delezione (o l'inserzione) di un singolo paio di basi. Un ulteriore sistema di protezione della cellula da mutazioni, in particolare da mutazioni puntiformi in grado di originare un codone di stop prematuro è rappresentato dal cosiddetto complesso di giunzione esonica o EJC. Questo agglomerato di proteine in caso di normale traduzione viene rimosso dal ribosoma nel suo normale avanzamento, ma resta sull'mRNA la cui traduzione sia stata interrotta prematuramente, portandolo alla degradazione grazie al suo marchio.