



Funzioni delle proteine

La funzione di una proteina, strettamente legata alla sua conformazione tridimensionale, dipende dalla interazione con altre molecole: per esempio gli anticorpi si legano a virus o batteri, l'esochinasi lega il glucosio e l'ATP, le molecole di tropocollagene si aggregano in fibre di maggior diametro.

La sostanza che si lega alla proteina è detta ligando e la regione specifica della proteina che aderisce ad esso si chiama sito di legame.

L'interazione sito di legame/ligando dipende da legami deboli (legami idrogeno, forze di Van der Waals).

È evidente che variazioni, anche minime, nella struttura primaria di una proteina alterano la sua forma tridimensionale (in quanto modificano le interazioni tra gli amminoacidi) compromettendone la funzionalità.

L'attività biologica di molte proteine (enzimi, recettori, proteine del citoscheletro) è regolata in risposta a stimoli differenti, così da adattarla alle esigenze della cellula.

In genere, sono la struttura terziaria e quaternaria a conferire alle proteine le loro caratteristiche funzionali, ma queste strutture possono essere facilmente disorganizzate da condizioni che prevalgono sulle deboli forze da cui dipende l'avvolgimento dei polipeptidi, come avviene ad esempio ad opera di calore, pH acido o basico, detergenti.

L'azione di questi agenti antagonisti provoca così un fenomeno che va sotto il nome di **denaturazione della proteina**, che è sempre accompagnato dalla perdita della sua normale funzione biologica, come ad esempio l'attività enzimatica.



Denaturazione reversibile ed Irreversibile

- Denaturazione: perdita della struttura tridimensionale causata da variazioni anche modeste di temperatura e pH, presenza di solventi organici (alcol o acetone)
- Reversibile: la rimozione dell'agente denaturante permette di riprendere spontaneamente la struttura nativa
- Irreversibile: la proteina non ritorna più a la condizione originale
- Talvolta la proteina può assumere la sua struttura tridimensionale corretta solo nell'ambiente intracellulare, per la presenza di altre proteine chiamate chaperoni o chaperonine



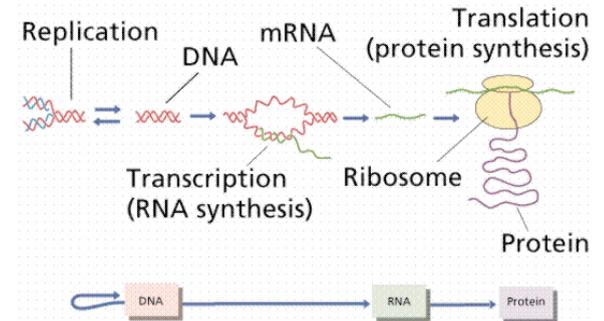
La sintesi proteica



Per fare le proteine nel modo in cui vanno fatte occorrono precise informazioni e queste sono annotate, per ciascuna proteina, nel materiale genetico, ossia nel DNA e precisamente sono codificate nelle sequenze di nucleotidi delle molecole di DNA.

Quando si duplica, il DNA trasmette queste istruzioni dalla cellula madre alle cellule figlie e in questo modo ogni nuova cellula e ogni nuovo organismo possono disporre delle informazioni necessarie alla sintesi di tutte le proteine responsabili delle loro strutture e delle loro funzioni. Nel DNA le informazioni sono organizzate in unità fondamentali costituite da sequenze di nucleotidi dette geni.

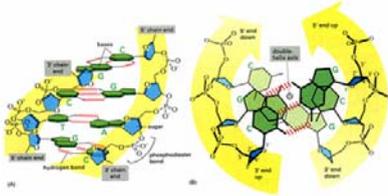
Ciascun gene interviene nella determinazione di una o più proteine.



Il codice genetico ed i nucleotidi

Un codice è un insieme di simboli usato per immagazzinare informazioni; se si conosce la natura del codice è possibile tradurre un'informazione da una forma in un'altra.

I simboli del codice genetico per tutti gli esseri viventi sono soltanto quattro e si trovano nelle molecole di DNA: si tratta delle quattro basi azotate (**Adenina A**, **Citosina C**, **Timina T**, **Guanina G**) che caratterizzano ciascuno dei quattro nucleotidi.



5' base	4' base				3' base
T	U	C	A	G	T
UUU	UUU	UUU	UUU	UUU	UUU
UUC	UUC	UUC	UUC	UUC	UUC
UUA	UUA	UUA	UUA	UUA	UUA
UUG	UUG	UUG	UUG	UUG	UUG
UUC	UUC	UUC	UUC	UUC	UUC
UUA	UUA	UUA	UUA	UUA	UUA
UUG	UUG	UUG	UUG	UUG	UUG
UUU	UUU	UUU	UUU	UUU	UUU
UUU	UUU	UUU	UUU	UUU	UUU
UUU	UUU	UUU	UUU	UUU	UUU
UUU	UUU	UUU	UUU	UUU	UUU

Come può un alfabeto così limitato codificare i tanti caratteri propri di ciascuna specie vivente, vegetale ed animale?

Le parole del codice genetico, corrispondenti a ciascun aminoacido sono formate da sequenze di tre basi azotate. L'insieme di tre basi in un certo ordine è chiamato tripletta. Poiché le basi sono 4, sono possibili $4^3 = 64$ triplette. In natura esistono 20 aminoacidi disponibili per formare proteine, e ciascun aminoacido corrisponde ad una tripletta.

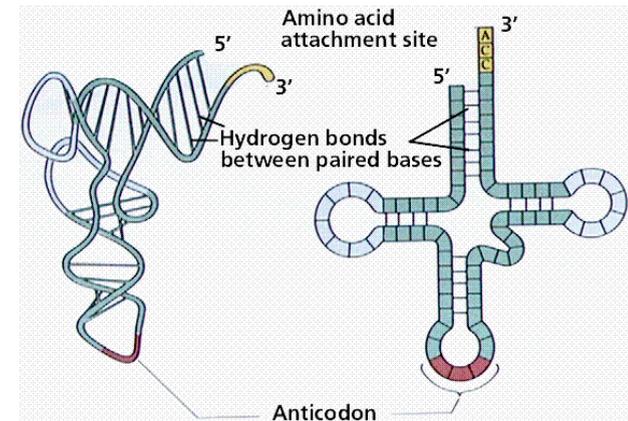
Degenerazione del codice genetico: più triplette diverse codificano per uno stesso aminoacido.

Non solo: determinate triplette funzionano anche come segnali che indicano al DNA dove iniziare e dove terminare la sintesi di una proteina. Ciascun segmento del DNA nel quale è codificata l'informazione per una particolare proteina strutturale o per un enzima prende il nome di **gene**.

La molecola di DNA, a causa delle sue dimensioni non può uscire dal nucleo della cellula, mentre la fabbricazione delle proteine avviene nel citoplasma e precisamente nei **ribosomi**.

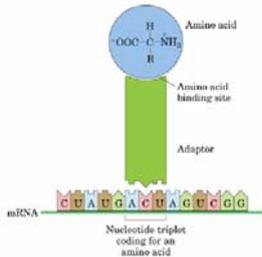
C'è quindi necessità di un intermediario, o per meglio dire un messaggero, che prelevi l'informazione dal DNA all'interno del nucleo e la trasferisca nel citoplasma. Per poter uscire dal nucleo questo messaggero deve essere formato da una molecola di dimensioni inferiori a quelle del DNA: questa molecola è l'acido ribonucleico, o **mRNA**, che differisce dal DNA non soltanto perché è più piccolo ma anche perché lo zucchero è rappresentato dal ribosio e non dal desossiribosio e perché una delle sue quattro basi è l'uracile (U) al posto della timina.

Questo migra dal nucleo al citoplasma dove un'altra forma di RNA, detta RNA transfer (**tRNA**), ha il compito di tradurre le triplette nel linguaggio degli aminoacidi.

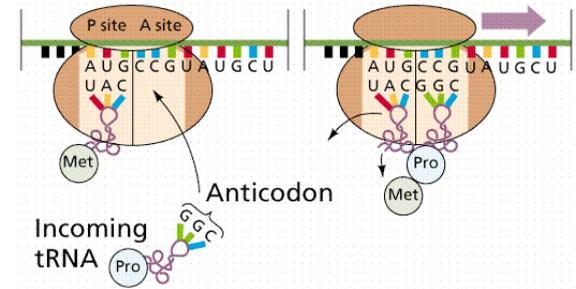


- Un piccolo acido nucleico (presumibilmente l'RNA) deve servire da "adattatore" e deve essere così strutturato:
 - Una porzione deve legare un amminoacido specifico
 - Un'altra porzione che riconosce un breve segmento di mRNA codificante per lo specifico amminoacido.

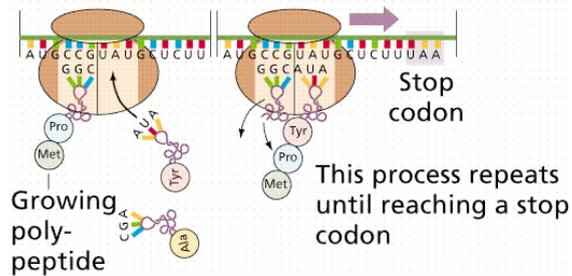
Il processo della sintesi proteica guidata dall'mRNA viene definito **TRADUZIONE**.



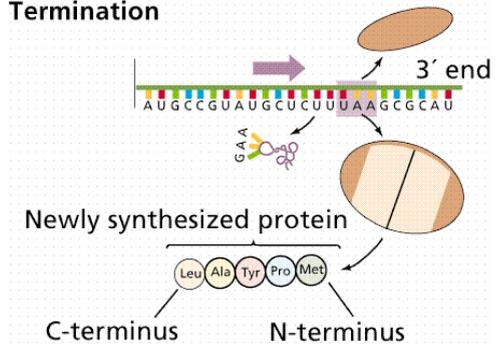
Elongation (translation)



Elongation continues



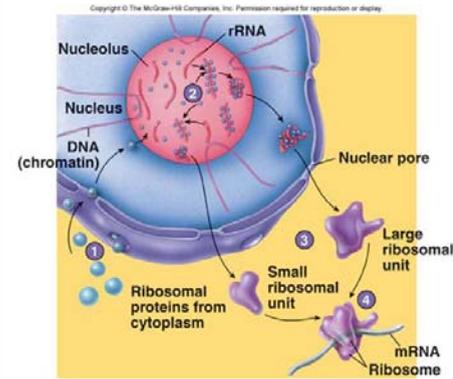
Termination





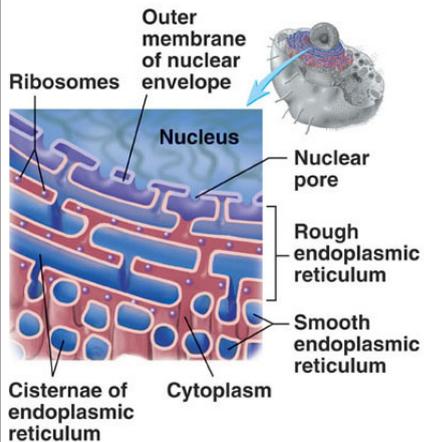
Codoni con funzioni specifiche

- **Codone d'inizio (AUG):** segnala l'inizio delle catene polipeptidiche in tutte le cellule; codifica per l'AA **METIONINA**
- **Codoni di terminazione:** sono tre (UAA, UAG, UGA) e non codificano per alcun AA. Essi segnalano la fine della catena polipeptidica
- **Open Reading Frame (ORF)** denominata anche **Quadro di Lettura aperto**. Si ha se un quadro non presenta un codone di terminazione per più di 50 nucleotidi consecutivi
- **Degenerazione del codice:** è, forse, la caratteristica più sorprendente del codice genetico. Ad ogni singolo AA può corrispondere più di un codone. La differenza si realizza di solito a carico del terzo nucleotide.



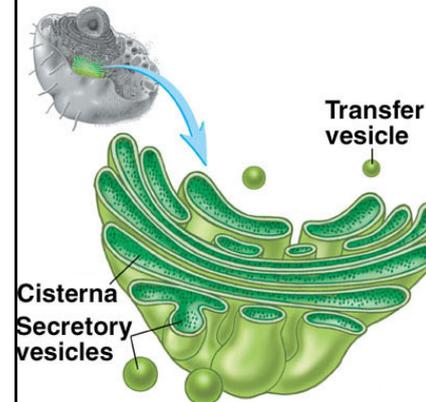
- Sites of protein synthesis
- Composed of a **large** and **small** subunit
- Types
 - Free
 - Attached to endoplasmic reticulum

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



- Types
 - **Rough**
 - Attached ribosomes
 - Proteins produced and modified
 - **Smooth**
 - Not attached ribosomes
 - Manufacture lipids
- **Cisternae:** Interior spaces isolated from rest of cytoplasm

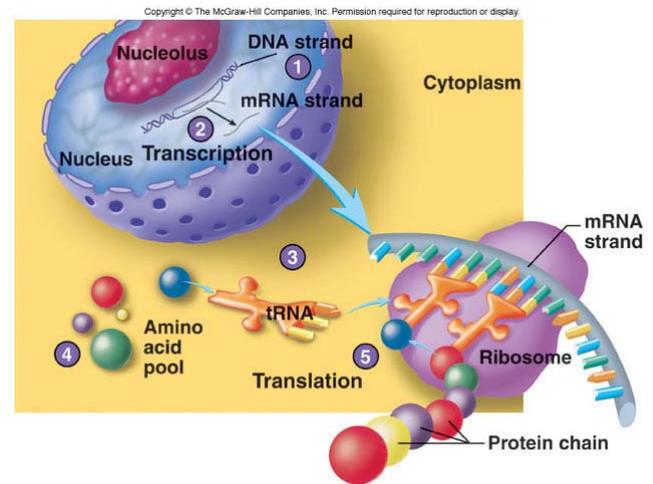
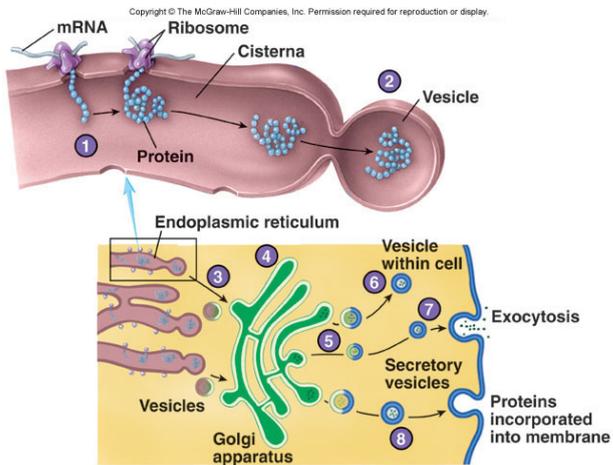
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Modification, packaging, distribution of **proteins and lipids** for secretion or internal use

Diversi polipeptidi sono destinati a luoghi diversi all'interno o all'esterno della cellula

I polipeptidi hanno una sequenza aa chiamata **signal sequence** che li dirige nella loro corretta posizione e che viene rimossa quando la proteina diviene attiva



Elettroforesi e Proteine

Le proteine, polimeri di aminoacidi, sono estratte dalle cellule con vari metodi. Le proteine hanno un peso molecolare ed una Carica (dipendente dal tipo di aminoacidi della sequenza primaria). Se poste in un campo elettrico, le proteine migrano verso il polo opposto ad una rapidità direttamente proporzionale alla carica ed inversamente proporzionale al peso molecolare.

Un particolare tipo di corsa elettroforetica è detta SDS-PAGE (elettroforesi su gel di poliaccrilammide in presenza di SodioDodecilSolfato). Nella SDS-PAGE tutte le proteine sono trattate con SDS che ha una forte carica negativa ed è in grado di legarsi alle stesse; in seguito a questo fenomeno se fatte correre su un gel di acrilammide le proteine migrano tutte verso il catodo (polo +) ad una velocità inversamente proporzionale alla massa.



Elettroforesi monodimensionale

Anodo (+) e Catodo (-)
Gel di poliaccrilammide: sfrutta le differenze di carica e peso molecolare

Le molecole provviste di carica netta si muovono nel campo elettrico generato dalla ddp creata tra anodo e catodo.

Il gel funziona da setaccio: le molecole con dimensioni più piccole dei pori del gel si muovono più rapidamente mentre quelle molto più grandi restano immobili.

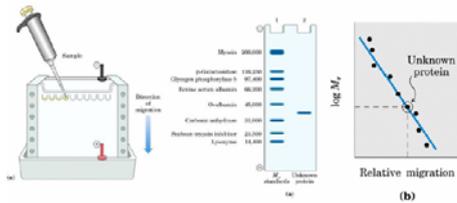
SDS PAGE: separazione delle proteine denaturate in base alla loro massa su gel di poliaccrilammide. La miscela di proteine viene sciolta in una soluzione di SDS (composto apolare con carica -) e Mercaptoetanolo per spezzare le interazioni non covalenti ed i ponti di solfuro (S-S, SH).

Complesso SDS-proteina sottoposto a elettroforesi: tutte le proteine migrano verso il polo positivo e si differenziano solo per la massa.





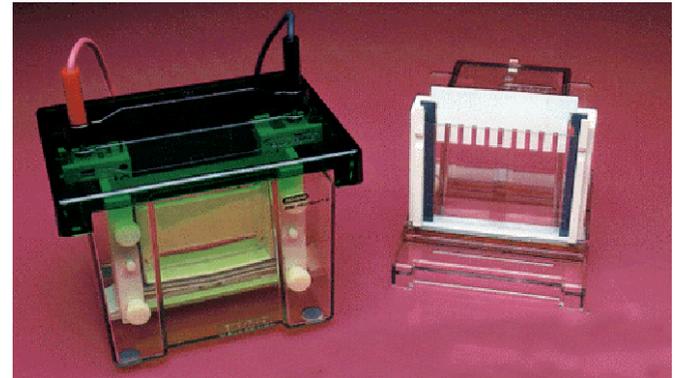
Other methods: SDS-PAGE



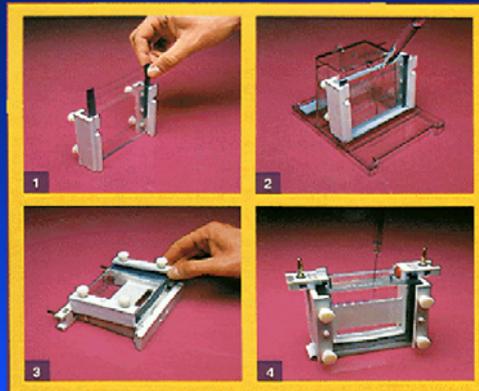
- SDS-PAGE. Separates based on size. SDS binds to proteins then subject to electric field
- Semi-log relationship between size and mobility
- Used to determine purity and also molecular weight



Apparato per elettroforesi

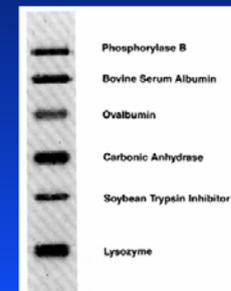


Gel Preparation:



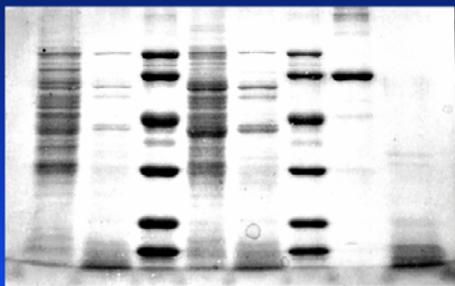
Molecular Weight Standards:

- Proteins of known molecular weight run concomitantly with "unknown" samples.

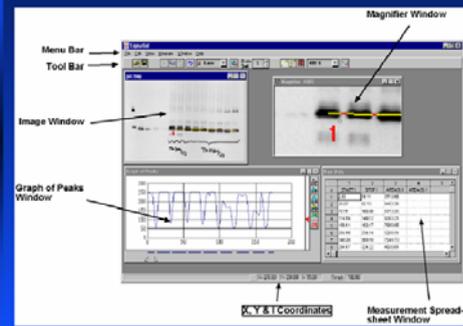




Electrophoresis:



SG Application Window Anatomy:



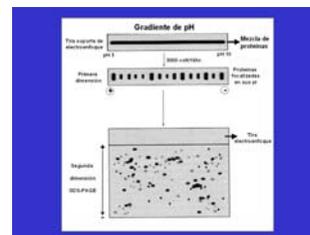
Punto isoelettrico (pI):

È il pH a cui la carica netta della proteina è 0.
A questo pH la mobilità elettroforetica è 0.

- ◊ Resine costituite da polimeri carichi in modo diverso creano un gradiente di pH (acido in alto, alcalino in basso).
- ◊ Ogni proteina si muove nel gel sotto azione di un campo elettrico fino a raggiungere una posizione in cui il pH sarà uguale al suo pI.

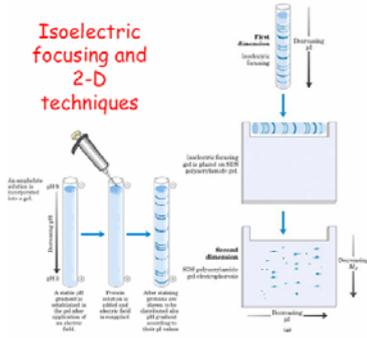
Elettroforesi bidimensionale

- ◊ Campione sottoposto a focalizzazione isoelettrica: separazione in base al pI.
- ◊ Gel posto orizzontalmente su gel di poliacrilamide con SDS, le proteine sono allineate a seconda della migrazione avvenuta nel passaggio precedente.
- ◊ Elettroforesi in direzione perpendicolare alla precedente: separazione in base alla massa.
- ◊ Ottengo una separazione ad altissima risoluzione delle proteine in direzione orizzontale secondo il loro pI e in direzione verticale secondo la loro massa.

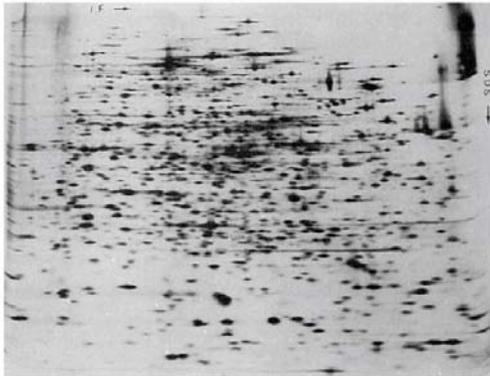
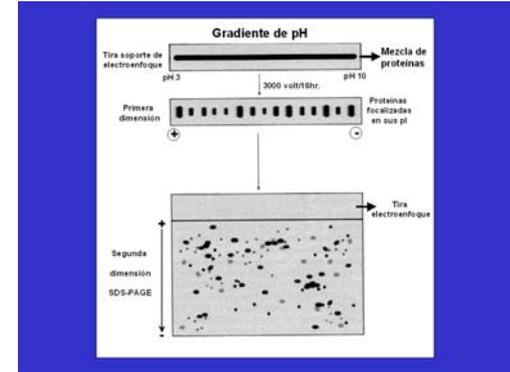




Isoelectric focusing and 2-D techniques



- Use small molecules with different pI values referred to as ampholytes to establish a pH gradient.
- Proteins migrate until they reach pH of pI, then zero charge.
- Put IEF gel on top of SDS-PAGE.
- Second dimension now separates based on size
- See next slide for sample separation of proteins





Adesso comandano le proteine

La proteomica è la grande novità del biotech.

Catalogare le proteine degli organismi e capire come interagiscono.

Proteoma:

Insieme delle proteine che costituiscono le cellule e i tessuti del nostro corpo. Il genoma, patrimonio di informazione genetica del nostro organismo, contiene solamente la ricetta per fabbricare le proteine, ma i mattoni e la calce che costituiscono le nostre cellule e che compiono la maggior parte del lavoro sono proprio le proteine. E sono sempre loro che permettono di distinguere i diversi tipi di cellule; per quanto tutte le cellule abbiano essenzialmente lo stesso genoma, esse si differenziano in base a quali geni sono attivi e a quali proteine vengono prodotte.

Cercare di catalogare tutte le proteine degli esseri viventi e scoprire come interagiscono tra di loro.



Il termine «**proteoma**» fu coniato nel 1994 da Marc R. Wilkins, vicepresidente e direttore della divisione di bioinformatica presso la Proteome Systems di Sydney, in Australia, per definire il complemento proteico codificato da un genoma.



L'esatta definizione di **proteomica** varia a seconda dell'interlocutore, ma la maggior parte dei ricercatori concorda che si possa ridurre a tre caratteristiche:

◊ Identifica tutte le proteine prodotte in una determinata cellula, tessuto o organismo;

◊ definisce come queste proteine si organizzino in reti simili a circuiti elettrici;



Sfortunatamente, il **proteoma** è molto più complicato del genoma. L'«alfabeto» del DNA è costituito da quattro basi chimiche indicate con le loro iniziali: adenina (A), citosina (C), guanina (G) e timina (T). Le proteine, invece, sono costituite da ben 20 unità fondamentali, gli aminoacidi. I geni specificano quali aminoacidi debbano essere combinati per produrre una determinata proteina. Ma, anche quando si conosce la sequenza degli aminoacidi di una proteina, non è ancora possibile dedurre le funzioni e le caratteristiche, né individuare con quali altre proteine essa interagisca. Né è sempre possibile prevedere la struttura tridimensionale con assoluta sicurezza. A differenza dei geni, che hanno una struttura lineare, le proteine si dispongono in conformazioni che, in alcuni casi, sfuggono a ogni previsione. Come se non bastasse, le cellule spesso modificano le proteine aggiungendovi zuccheri o acidi grassi, o entrambi, in modi altrettanto difficilmente prevedibili.



Per quanto gran parte dei ricercatori concordi sul fatto che il genoma umano contenga circa 40.000 geni, una cellula produce centinaia o migliaia di proteine differenti. Per comprendere il funzionamento del proteoma, i ricercatori devono conoscere le caratteristiche di tutte quelle proteine. Applicare semplicemente i dati prodotti dal Progetto Genoma umano - che ha finalmente mandato in soffitta l'obsoleto dogma secondo il quale un gene codifica per una proteina - non risolve il problema. È ormai chiaro, anche se i meccanismi rimangono spesso oscuri, che un gene può produrre molte e diverse proteine.

«Rimane ignota una frazione compresa tra il 30 e il 50 per cento delle proteine umane così come le loro funzioni»



Il lievito *Saccharomyces cerevisiae* è il primo organismo eucariote di cui sia stato sequenziato tutto il genoma nel 1996 nell'ambito di un progetto satellite del più vasto Progetto Genoma Umano e tutte le sequenze sono raccolte in apposite Banche Dati. Esistono anche delle mappe 2D-PAGE di ceppi di *Saccharomyces* con caratteristiche di interesse enologico (K 310) utilizzato nella vinificazione di Sangiovese per la produzione di vini DOCG.



Catalogare le proteine porta fino a un certo punto. Per capire che cosa realmente facciano le proteine nell'organismo dobbiamo sapere come varia la componente proteica da una cellula all'altra e all'interno della cellula con il mutare delle condizioni circostanti. Dobbiamo inoltre sapere come le diverse proteine collaborino per portare a termine le varie attività della cellula. Scoprire come le proteine si legano l'una all'altra fino a formare catene di reazioni chimiche o creare meccanismi molecolari come il fuso mitotico che, al momento della divisione cellulare, separa le due cellule. «Le proteine si combinano per formare reti e se c'è una cosa da scoprire a proposito di una proteina è con quali altre proteine essa interagisca.»

