
LA BIOCHIMICA DEL 2000

GIANNI CAPPUGI

Ordinario di Biochimica, Dipartimento di Scienze Biochimiche, Università di Firenze

1. Introduzione

La biochimica è nata e si è sviluppata nel secolo scorso contribuendo in maniera determinante alla comprensione dei fenomeni biologici a livello molecolare. Infatti, agli albori di questa disciplina l'interesse scientifico preminente era quello di comprendere il destino chimico dei nutrienti e i meccanismi di sintesi di tutte le molecole biologiche grandi o piccole che fossero. Tutto questo viene definito come *metabolismo* cellulare e viene convenzionalmente suddiviso in *catabolismo* e *anabolismo*. Nel catabolismo avvengono fondamentalmente reazioni di ossidazione e quindi di distruzione delle biomolecole che hanno il fine ultimo di produrre energia e molecole semplici che rappresentano le condizioni perché si realizzi l'anabolismo, nel quale avvengono le reazioni di sintesi di tutte le molecole che fanno parte del corredo di un organismo vivente.

Il primo passo della ricerca biochimica si è identificato con la decifrazione e l'identificazione delle cosiddette *vie metaboliche*, ove per «via metabolica» si intendono tutte quelle reazioni chimiche concatenate che permettono la trasformazione (per ossidazione o per sintesi) delle singole biomolecole (carboidrati, lipidi, proteine, acidi nucleici). Il secondo passo è stato la identificazione e caratterizzazione dei catalizzatori biologici che permettono alle vie metaboliche di avvenire in tempi velocissimi: questi catalizzatori sono fondamentalmente proteine e prendono il nome generico di *enzimi*. Il terzo passo, o fase, coincide invece con lo studio dei meccanismi di regolazione della velocità delle singole vie metaboliche e questa fase ha dimostrato che di nuovo gli enzimi hanno un ruolo importante: infatti, in ogni via metabolica esiste almeno un enzima cosiddetto «segnapasso» la cui attività catalitica può essere spenta o potenziata a seconda delle esigenze metaboliche cellulari.

Infine, un avanzamento importante alle conoscenze dei meccanismi biochimici si è realizzato grazie alla messa a punto di tecnologie che hanno permesso la determinazione della struttura tridimensionale di macromolecole come le proteine e gli acidi nucleici; gli ultimi anni del secolo scorso hanno visto un'esplosione degli studi sulla struttura delle macromolecole biologiche lasciando una eredità culturale importantissima agli attuali ricercatori.

Qual è il panorama scientifico in cui si muove e si muoverà il biochimico del duemila? Intanto si deve rilevare che quella identificazione dei ruoli nella ricerca

biologica che fino a metà del secolo scorso era abbastanza netta – per cui non si poteva confondere un biochimico con un genetista o con un cosiddetto morfologo –, con la nascita della biologia molecolare e poi della biologia cellulare ha cominciato a farsi via via sempre più sfumata, tanto che oggi è più difficile dare una definizione precisa della biochimica che non coinvolga aspetti scientifici che interessano anche la morfologia, la biologia cellulare, la biologia molecolare o la genetica. Per semplificare un po' questo quadro interdisciplinare si può ancora dire, forzando un po' la cosa, che comunque l'interesse scientifico che accomuna la maggioranza dei biochimici è rappresentato dalle proteine.

Detto questo, vediamo, senza con questo voler esaurire l'argomento, quali sono le linee di ricerca che attualmente ricevono i maggiori consensi: «Studi sul folding delle proteine», «Studi sul meccanismo di azione degli enzimi e progettazione di inibitori specifici», «Proteomica», «Meccanismi di segnalazione cellulare», «Motori molecolari», «Nuove funzioni di RNA».

2. Studi sul *folding* di proteine

Con il termine di *folding* s'intendono tutti quei processi attraverso i quali ogni proteina acquisisce la propria struttura nativa. La struttura nativa è la struttura che ogni proteina presenta nel proprio ambiente naturale permettendole di svolgere una specifica funzione. E' noto da tempo che la struttura tridimensionale di una proteina dipende dalla sua sequenza amminoacidica codificata da uno specifico gene e, in certi casi limite, basta anche la differenza nella posizione in sequenza di un solo amminoacido per avere una struttura tridimensionale diversa (per esempio, l'emoglobina S nell'anemia falciforme). La mole enorme degli studi fatti sulle strutture delle migliaia di proteine note ha permesso di capire quali sono i principi che permettono ad una certa sequenza di strutturarsi localmente (strutture secondarie) ma non hanno ancora dato indicazioni in che modo si possa raggiungere rapidamente la forma finale funzionante.

A questo proposito è utile citare il *paradosso di Cyrus Levinthal* il quale ha preso in considerazione il caso teorico di una ipotetica proteina formata da 100 amminoacidi:

assumendo che ogni amminoacido possa assumere (scegliere) tre conformazioni diverse, il numero di strutture dovrebbero essere 3^{100} e prendendo come tempo minimo per passare da una conformazione ad un'altra il tempo di vibrazione di legame (10^{-13} s), il tempo necessario perché questa proteina raggiunga la propria conformazione nativa è uguale a $1,6 \times 10^{27}$ anni. Un tempo enorme! Ogni proteina dopo la sua sintesi a livello ribosomiale impiega tempi rapidissimi, nell'ordine dei secondi, per raggiungere la struttura nativa.

Molti ricercatori, di provenienza culturale diversa come biochimici, biofisici, biomatematici, stanno attivamente lavorando sui problemi del *folding* e nella maggior parte dei casi l'approccio sperimentale prevede un ribaltamento del fenomeno: si parte infatti da una proteina strutturata, la quale viene denaturata, nel senso che viene eliminato completamente l'avvolgimento tridimensionale, e successivamente si

cerca di far avvenire la rinaturazione in condizioni opportune. Questi studi, insieme a simulazioni portate avanti da complessi programmi *computer-based*, stanno lentamente portando a indicazioni sul meccanismo di folding. Per esempio, è oggi chiaro che non si tratta di un processo di «prova e riprova» quanto invece un processo cooperativo di stabilizzazione progressivo di strutture intermedie per favorire interazioni a distanza fra le varie parti della molecola proteica. Tuttavia, nonostante questo grande impegno internazionale, siamo ancora lontani dall'obiettivo!

Ma perchè sarebbe così importante conoscere le regole che permettono a una sequenza lineare di amminoacidi di organizzarsi in una struttura proteica tridimensionale?

Considerato che una qualsiasi proteina è in grado di portare avanti un certo tipo di funzione (catalitica, di trasporto, di difesa, di segnalazione, motoria etc.) proprio perché ha un certo tipo di struttura, conoscendo le regole che permettono a una specifica sequenza amminoacidica di assumere una specifica struttura tridimensionale, potremmo progettare in laboratorio proteine con sequenze amminoacidiche non presenti in natura in grado di svolgere particolari funzioni. Oltre a questo aspetto clamoroso, sarebbe comunque più semplice l'applicazione di questa forma di ingegneria proteica per il miglioramento funzionale di proteine note. Anche questo secondo aspetto potrebbe avere delle importanti ripercussioni di tipo pratico come l'utilizzazione di enzimi per produzioni industriali o agrarie, oppure per la loro applicazione in campo medico.

Inoltre, in questi anni è nato un grandissimo interesse intorno al cosiddetto *protein misfolding* cioè una non corretta strutturazione alla quale andrebbero incontro certe proteine in condizioni diverse dalla norma. Per fare un esempio che faccia comprendere l'interesse di questo aspetto, basta citare il caso di numerose malattie del sistema nervoso nelle quali si verificano fenomeni di *protein misfolding*.

Nel cosiddetto «morbo della mucca pazza» l'agente di trasmissione è costituito da aggregati di una proteina prionica la quale deriva da una proteina normale che è andata incontro ad un processo di *protein misfolding*. Nella malattia di Alzheimer si è visto che si formano fibre proteiche che derivano dall'aggregazione di un frammento che si genera da un precursore proteico e che contemporaneamente subisce un processo di *misfolding*.

3. Studi sul meccanismo di azione degli enzimi e progettazione di inibitori specifici

Questo settore della ricerca biochimica ha sempre appassionato i biochimici e ha rappresentato anche in passato il fiore all'occhiello dell'enzimologia.

Per stabilire il funzionamento a livello molecolare di un enzima occorre avere a disposizione la struttura tridimensionale dettagliata della proteina e in maniera particolare del cosiddetto «sito attivo» che costituisce quella parte di struttura dove i substrati (le molecole che devono subire una trasformazione) entrano in contatto con le catene laterali di specifici amminoacidi o con i gruppi chimici di specifici coenzimi per essere trasformati rapidamente in prodotti di reazione. Fortunatamente tutto ciò è

reso possibile dalla grande potenza e raffinatezza dei metodi di indagine che si basano fondamentalmente sulla diffrazione dei raggi X e sulla Risonanza Magnetica Nucleare. Questi siti di legame sono modellati per poter accogliere il ligando e quindi la loro struttura dipende dalla forma e dalla grandezza dei ligandi; i siti di legame per le macromolecole sono generalmente localizzati sulla superficie dell'enzima e possono essere concavi, convessi o piatti mentre i siti di legame per piccole molecole sono delle fenditure, o una sorta di tasca o cavità, poste più profondamente nella struttura dell'enzima.

Per la individuazione dei siti di legame vengono impiegati metodi computazionali o metodi sperimentali. I metodi computazionali utilizzano una molecola sonda e i siti di legame sono identificati come regioni della struttura in cui l'energia di interazione calcolata tra la sonda e la proteina è favorevole al legame. I metodi sperimentali sono attualmente più precisi rispetto a quelli computazionali e si basano soprattutto su una tecnica cristallografica che prevede l'immersione dei cristalli proteici in un solvente organico che simuli il gruppo funzionale del ligando; in maniera simile si possono utilizzare anche metodi basati sulla Risonanza Magnetica Nucleare.

Una volta identificato il sito di legame, si passa alla ricerca dei gruppi funzionali che partecipano chimicamente al meccanismo di catalisi; i metodi tradizionali prevedevano una serie di analisi cinetiche, anche assai complicate, in presenza o assenza di inibitori o attivatori, mentre un approccio più moderno è rappresentato dalla mutagenesi sito-specifica.

La mutagenesi sito-specifica è una tecnica di ingegneria genetica che permette di ottenere in laboratorio strutture enzimatiche dove è stato sostituito uno specifico amminoacido per evidenziare la sua eventuale funzione nel meccanismo di catalisi.

Partendo da queste informazioni è oggi possibile, con sofisticati programmi computazionali, simulare e quindi progettare delle molecole che siano in grado di interagire con i siti di legame e con le strutture catalitiche bloccando l'attività catalitica dell'enzima. Questa operazione non può tuttavia prescindere da un approccio sperimentale, perché non sempre questa interazione è prevedibile esattamente. Infatti le proteine, e quindi anche gli enzimi, non sono strutture rigide bensì parzialmente flessibili e spesso questa flessibilità è essenziale per la funzione biologica dell'enzima. Il sito attivo, ad esempio, non ha sempre una struttura complementare ai substrati ma invece assume una struttura complementare al cosiddetto stato di transizione. Trovare una molecola che sia complementare a questa conformazione e che blocchi l'attività dell'enzima significa disporre di un forte inibitore specifico dell'enzima.

Tutto questo può avere una importante ricaduta ad esempio nel campo della salute pubblica (*Drug Design*). Per ricordare qualche esempio, il captopril è una molecola organica di sintesi che è in grado di legarsi e di bloccare il metalloenzima ACE (*angiotensin-converting enzyme*) regolando così la pressione sanguigna. Un altro esempio da citare sono gli inibitori di una particolare aspartil-proteasi come è la proteasi dell'HIV; fra questi inibitori c'è l'aloiperidolo e il crivivan che vengono usati come

farmaci nella terapia dell'AIDS. Il crivivan somiglia al peptide substrato della proteasi HIV ed è costruito intorno ad una struttura alcolica che simula lo stato di transizione; quando si lega all'enzima induce in questo il movimento di due ali flessibili dell'enzima che comporta la chiusura e quindi il blocco all'accesso dell'acqua nel sito attivo.

Questo tipo di studi può essere molto utile naturalmente anche in altri campi applicativi come il settore agrario o alimentare ma può avere anche una grande importanza nella ricerca pura quando si deve assegnare una funzione precisa ad una proteina: disporre di un inibitore altamente specifico per questa proteina significa poterne bloccare l'effetto in una cellula o in organismo e studiarne gli effetti.

4. Proteomica

La proteomica studia il cosiddetto proteoma, cioè, il corredo proteico di una cellula. Questi termini sono un riflesso di termini conati per definire il patrimonio genetico: il genoma rappresenta infatti l'intera collezione di geni di un organismo. Attualmente questa terminologia si è allargata a tutti quei settori della biologia nei quali, invece di studiare la singola molecola di un certo tipo, si prendono in considerazione contemporaneamente tutte le molecole di quel tipo, per cui oggi si parla anche di trascrittoma, di interattoma e così via.

Per capire meglio possiamo vedere alcune differenze fra la proteomica e la chimica delle proteine: mentre la chimica delle proteine prende in considerazione di volta in volta la singola proteina, la proteomica studia miscele complesse di proteine; se nel primo caso si studia la sequenza amminoacidica completa della proteina, nel secondo è sufficiente una informazione parziale; nella chimica delle proteine si mira alla definizione della struttura e della funzione, mentre, nella proteomica, alla identificazione delle varie proteine attraverso una ricerca nei *database*.

Studiare il proteoma di una cellula non è la stessa cosa che studiare il genoma: infatti, mentre ogni nostra cellula presenta un genoma completo, non tutte le cellule hanno lo stesso proteoma perché in ogni singola cellula verranno espressi solo i geni che codificano per le proteine necessarie alle funzioni essenziali e per le proteine che assicurano funzioni cellula-specifiche.

Questo non significa che il proteoma sia più semplice del genoma, anzi è proprio il contrario. Ogni proteina infatti, anche se rappresenta il prodotto di un singolo gene può esistere in forme multiple in una stessa cellula o in cellule diverse e questo è dovuto essenzialmente a processi di modificazione della proteina (la cosiddetta maturazione) che incidono poi sulla struttura, sulla funzione, sulla localizzazione e sulla vita della proteina stessa.

La proteomica ha fondamentalmente quattro tipi di applicazioni:

1. identificazione e classificazione di tutte o quasi tutte le proteine presenti in un campione;
2. profilo di espressione proteica; l'identificazione di proteine in un particolare stato della cellula o organismo (es. differenziazione, sviluppo, malattia);

3. mappatura dell'interazione fra proteine; determinazione di come le proteine interagiscono l'una con l'altra nei sistemi viventi;
4. mappatura delle modificazioni di proteine; identificazione di come e dove le proteine vengono modificate.

5. Meccanismi di segnalazione cellulare

Le cellule sono sensibili a segnali che provengono dall'esterno, sotto forma di fenomeni chimici o fisici, e sono in grado di convertire questi segnali in una attivazione o disattivazione del proprio metabolismo, oppure sono capaci di modificare l'espressione di specifici geni in modo da modificare la concentrazione cellulare di particolari proteine.

Negli organismi multicellulari, segnali chimici provenienti da cellule diverse permettono la coordinazione di meccanismi fisiologici come risposta al segnale stesso. Ad esempio, quando un individuo si trova in una situazione di pericolo, le ghiandole surrenali rilasciano nel sangue l'adrenalina, un ormone che favorisce l'utilizzazione delle sostanze di riserva per consentire a certi tessuti come i muscoli di aumentare la propria funzionalità; per la mobilitazione delle sostanze di riserva, le cellule che le contengono devono sentire il segnale chimico (adrenalina) e trasferire questo segnale al proprio interno con un meccanismo di amplificazione del segnale stesso (*signal transduction*).

Queste vie di segnalazione sono spesso molto complesse e prevedono l'intervento di più molecole e reazioni organizzate in maniera lineare ma anche secondo vie ramificate non sempre chiare nella loro effettiva costruzione. Nonostante che il numero di queste vie aumenti continuamente grazie agli studi portati avanti da tantissimi laboratori di tutto il mondo, e nonostante che ogni via sia organizzata in maniera più o meno specifica, è possibile evidenziare alcuni principi e strategie comuni.

Uno stimolo provoca il rilascio di una molecola-segnale che prende il nome di «primo messaggero»; nella maggior parte dei casi il primo messaggero, per la sua natura chimica, non riesce a penetrare all'interno della cellula, per cui si lega a specifiche proteine localizzate sulla membrana plasmatica dette «recettori»; il legame del primo messaggero con il recettore induce in quest'ultimo una variazione di conformazione che mette in moto un processo biochimico che porta al rilascio all'interno della cellula di un «secondo messaggero»; alcuni importanti secondi messaggeri sono l'AMP ciclico, il GMP ciclico, lo ione calcio, l'inositolo 1,4,5-trifosfato e il diacilglicerolo; in questa fase inizia il processo di amplificazione del segnale: infatti, ogni singolo recettore attivato dal legame con il ligando è in grado di attivare un enzima che produce nell'unità di tempo un gran numero di molecole di secondo messaggero; il secondo messaggero direttamente, ma più spesso indirettamente attraverso strade complesse, attiva o inibisce specifici enzimi, specifiche pompe (proteine di trasporto) o specifici fattori di trascrizione di geni; dopo che la cellula ha risposto al segnale il processo di segnalazione viene opportunamente bloccato per permettere alla cellula di rispondere a nuovi segnali.

L'enorme interesse che esiste verso queste ricerche deriva dal fatto che oggi sappiamo che molti tipi di cancro sono associati a processi alterati di segnalazione dove ad esempio non si realizza una corretta terminazione del processo di segnalazione stessa.

Come esempio si può ricordare una via di segnalazione innescata dall'insulina.

L'insulina si lega ad uno specifico recettore il quale si autoattiva attraverso un processo catalitico di autofosforilazione; il recettore attivato catalizza la fosforilazione di specifiche proteine dette IRS (insulin-receptor substrates); le IRS fosforilate interagiscono con altre proteine come la fosfoinositide 3-kinasi che causa un aumento della concentrazione di fosfadil-inositolo 3,4,5 trifosfato (PIP3); il PIP3 attiva un altro enzima PDK1 che a sua volta attiva, sempre per fosforilazione un'altra proteina kinasi detta Akt; l'Akt attivata catalizza la fosforilazione di componenti che portano ad un aumento sulla membrana plasmatica di proteine trasportatrici di glucosio (GLUT4) oppure all'attivazione di enzimi che stimolano la sintesi di glicogeno; lo spegnimento del segnale avviene poi ad opera di enzimi (proteina-fosfatasi) che sono al loro volta finemente regolati.

6. Motori molecolari

I movimenti che possiamo osservare negli organismi viventi, dai più semplici come i batteri ai più complessi come gli animali, hanno tutti una origine molecolare in quanto prendono origine da molecole proteiche che funzionano come dei veri e propri motori sfruttando una fonte di energia per attuare delle modificazioni strutturali sincronizzate.

Questi motori molecolari possono utilizzare l'energia di idrolisi di molecole di ATP per ottenere lo scivolamento fra proteine fibrose, come succede nel caso di miosina, kinesina e dineina, oppure sfruttano la formazione di gradienti protonici per fare ruotare strutture complesse come i flagelli o come il più piccolo motore rotatorio esistente che ha la funzione di produrre molecole di ATP (ATP sintasi). Gli studi che sono stato condotti in questi ultimi anni sull'ATP sintasi hanno avuto uno sviluppo entusiasmante e hanno portato all'assegnazione di più di un premio Nobel l'ultimo dei quali è stato assegnato a John E. Walker per la struttura tridimensionale dettagliata.

Questo motore è formato da un rotore che è costituito da 10 proteine tutte uguali, inserite parallelamente nella membrana, e viene attivato dal passaggio di protoni. Il movimento rotatorio è trasmesso tramite una proteina di collegamento ad uno statore. Lo statore è a sua volta formato da diverse subunità e fra queste esistono tre subunità beta uguali le quali toccate con movimento rotatorio dalla proteina di collegamento sono in grado prima di legare ADP e fosfato inorganico, poi di formare ATP e quindi di liberarlo nel mezzo.

7. Nuove funzioni di RNA

Il grande interesse che attualmente suscitano le ricerche sulla funzione di RNA è nato con la scoperta che questa macromolecola non è solamente un messaggero al servizio del DNA ma è in grado di agire anche da catalizzatore, come gli enzimi, intervenendo

in importanti meccanismi cellulari come l'*auto-splicing*, il taglio di precursori, la modificazione sito-specifica, la traduzione e la regolazione dei geni.

Considerata la limitatezza dei gruppi funzionali presenti nell'RNA rispetto alla diversità delle catene laterali degli amminoacidi, grande fu la sorpresa quando Cech scoprì che alcune molecole di RNA possono comportarsi come gli enzimi aumentando la velocità di reazioni chimiche (ribozimi). Questa scoperta ha suggerito che i ribozimi possono aver avuto un ruolo fondamentale nei primi stadi dell'evoluzione della vita dove l'RNA poteva funzionare come materiale genetico e come catalizzatore per le reazioni primordiali del metabolismo.

Nella moderna biosfera la maggior parte delle funzioni enzimatiche appartengono alle proteine, tuttavia sono rimaste come fossili viventi alcune molecole di RNA con attività catalitica: gli introni che subiscono l'*autosplicing*, la Ribonucleasi P che processa i tRNA, i ribosomi che catalizzano la formazione del legame peptidico nella biosintesi delle proteine.

Un'altra scoperta recente riguarda la capacità dell'RNA di interferire con l'espressione di specifici geni (RNA *interference*).

Il meccanismo di silenziamento genico prevede il taglio di molecole di RNA a doppio filamento da parte di un particolare tipo di ribonucleasi, detta *Dicer*, così da formare frammenti a doppio filamento di 21-28 nucleotidi; questi frammenti vengono poi separati (probabilmente da una elicasi) in strutture a singolo filamento dette siRNA (*small interfering RNA*). Gli siRNA vengono poi incorporati in un sistema enzimatico complesso detto RISC (*RNA-induced silencing complex*), dove la sequenza nucleotidica di questi RNA serve come guida per legare e distruggere molecole di RNA messaggero che contengono una sequenza complementare.

Oltre al fatto che il silenziamento da RNA sta rivoluzionando la genomica funzionale, ci si aspetta per il futuro un impatto ancora maggiore se questo meccanismo di silenziamento genico specifico potrà essere applicato alla prevenzione di patologie umane.

BIBLIOGRAFIA

- Breaker R.R., Natural and engineered nucleic acids as tools to explore biology, *Nature* 432, 838-845, 2004.
- Capaldi R. A. and Aggeler R., Mechanism of the F_1F_0 -type ATP synthase, a biological rotatory motor, *Trends in Biochemical Sciences*, 27, 154-160, 2002
- Dobson C.M., Principles of protein folding, misfolding and aggregation Seminars in Cell and Developmental biology, 15, 3-16, 2004.
- Liebler D.C., *Introduction to Proteomics*, Humana Press, Totowa, New Jersey 2002.
- Scott J.D. and Pawson T. 'Cell communication: the inside story', *Scientific American*, 282, 72-79, 2000
- Vega S., Kang L.W., Velazquez-Campoy A., Kiso Y., Amzel L.M. and Freire E., 'A structural and thermodynamic escape mechanism form a drug resistant mutation of the HIV-1 protease', *Proteins*, 55, 594-602, 2004.