

EMILIA CHIANCONE\*

## **La biologia strutturale, una chiave di lettura molecolare dei fenomeni vitali\*\***

La tradizione vuole che, nella inaugurazione di un Anno Accademico, uno dei punti centrali sia la «prolusione», una lezione più o meno solenne nella quale il nuovo componente della comunità accademica espone il proprio programma di studi o di insegnamento. Ma il termine prolusione deriva dal latino «procludere» che significa dare l'avvio al discorso o al canto; è una parola composta nella quale è facile riconoscere «ludere». Dunque un gioco, che in antico voleva dire pure festa, rappresentazione teatrale e che assume significati particolari, dal «ludendo discitur» della pedagogia al linguaggio matematico, filosofico, psicologico. Del resto, Aristotele passeggiando insegnava anche fisica e scienze naturali ed alla fine del discorso noi stessi scopriremo complicati giochi di riconoscimento fra molecole che ne permettono l'interazione e quindi lo svolgimento dell'attività biologica.

Ci introduciamo così alla conoscenza della biologia strutturale, una disciplina che diventa occasione d'incontro fra cultori di varie discipline, spinti dalla curiosità di conoscere problemi di un certo respiro anche se lontani dal loro campo di studi, problemi che, per l'approfondimento della conoscenza stessa, richiedono la partecipazione di competenze diverse.

La biologia strutturale è una scienza giovane, ma già ricca di risultati impensabili fino a pochi anni fa, resi possibili dalla disponibilità di metodologie sperimentali e di calcolo sempre più sofisticate. Si può farla nascere nel 1953 con l'articolo in cui Watson e Crick proponevano la struttura a doppia elica del DNA. Questa struttura dalla semplicità meravigliosa, basata sull'accoppiamento specifico fra le basi azotate del DNA, fu descritta in poco più di una paginetta, come spesso è accaduto per le grandi scoperte. L'articolo di Watson e Crick termina con un'affer-

\* Socio dell'Accademia. Centro Studio Biologia Molecolare del C.N.R. Dipartimento di Scienze Biochimiche «A. Rossi Fanelli», Università «La Sapienza», Piazzale Aldo Moro 5, 00185 Roma.

\*\* Prolusione per l'inaugurazione del 217° Anno Accademico. Roma, 22 aprile 1999. Teatro dei Dioscuri, Via Piacenza 1.

mazione, nel tono tipico dell'understatement anglosassone: «non ci è sfuggito che l'accoppiamento specifico da noi postulato suggerisce un possibile meccanismo per copiare il materiale genetico».

Poco dopo, alla fine degli anni cinquanta, le prime strutture tridimensionali di proteine, le strutture della mioglobina e dell'emoglobina risolte da Kendrew e Perutz, diedero una conferma che il ruolo della biologia strutturale sia quello di fornire una chiave di lettura a livello molecolare dei fenomeni biologici. Esse permisero infatti di comprendere i meccanismi con cui l'ossigeno molecolare viene legato dall'emoglobina a livello dei polmoni, trasportato ai tessuti e lì rilasciato alla mioglobina che lo immagazzina pronta a cederlo perché venga utilizzato per il metabolismo.

Come premessa necessaria al nostro discorso può essere opportuno un cenno alle metodologie che permettono di determinare la posizione nello spazio di tutti gli atomi che compongono macromolecole complesse come il DNA e le proteine ed ai principi su cui sono basate. Ancora oggi la gran parte delle strutture viene determinata mediante la diffrazione dei raggi X non tanto da parte di fibre, come fecero Crick e Watson, ma soprattutto da parte di cristalli, come fecero Kendrew e Perutz. Il principio è semplice: il passaggio dei raggi X attraverso una struttura regolare provoca un fenomeno di diffrazione; pertanto la radiazione dispersa dagli elementi ripetitivi della struttura mostra un rafforzamento nelle direzioni in cui le onde sono in fase, mentre viene annullata per interferenza nelle direzioni in cui le onde sono fuori fase. Vengono impiegati i raggi X, che hanno una lunghezza d'onda dello stesso ordine di grandezza dei legami covalenti, perché, per ottenere un quadro di diffrazione nitido, è necessario che la lunghezza d'onda della radiazione sia un po' più breve della spaziatura regolare fra gli elementi della struttura. L'intensità di ogni massimo di diffrazione viene poi utilizzata per ricostruire matematicamente l'immagine tridimensionale della struttura del cristallo tenendo conto della regola che piccole spaziature nella struttura corrispondono a grosse spaziature nel quadro di diffrazione, e viceversa.

Nel cristallo peraltro le molecole di proteine o di acidi nucleici vengono per così dire fissate in un'unica conformazione dalle forze del reticolo cristallino. Anche se gli atomi possono spostarsi intorno alla loro posizione media, le molecole perdono la mobilità strutturale che permette loro lo svolgimento della funzione biologica. Non sorprende quindi che in questi ultimi anni gli sforzi di molti ricercatori siano stati rivolti proprio allo sviluppo di nuove strategie per lo studio degli aspetti dinamici della struttura e della loro correlazione con la funzione. Ad esempio, da circa vent'anni per macromolecole relativamente piccole, cioè con pesi molecolari intorno a 20000 Dalton, è possibile determinare la struttura non nel cristallo ma in soluzione mediante la risonanza magnetica nucleare bidimensionale che fornisce la distanza fra protoni specifici delle macromolecole quando questa sia inferiore a circa 0,5 nm.

Vorrei non soffermarmi oltre sugli aspetti metodologici, ma tornare alla mioglobina ed alla emoglobina e discutere le caratteristiche strutturali che rendono la mioglobina adatta al deposito dell'ossigeno e l'emoglobina al suo trasporto. Nella mioglobina e nella emoglobina, rispettivamente un monomero e un tetramero for-



Fig. 1. Struttura tridimensionale di mioglobina (A), emoglobina umana (B) ed emoglobina dimerica di *Scopharca inaequalis* (C). Nella mioglobina e nell'emoglobina umana il gruppo eme, contenuto in una tasca della catena globinica, è esposto al solvente; nell'emoglobina di *Scopharca* i due gruppi eme sono praticamente a contatto diretto a livello dell'interfaccia fra i due monomeri.

mato da due catene  $\alpha$  e due catene  $\beta$ , tutte le catene polipeptidiche hanno una struttura tridimensionale simile, che forma una tasca in cui è contenuto il gruppo eme; al centro dell'eme un atomo di ferro cui si lega reversibilmente l'ossigeno (Fig. 1). L'affinità per l'ossigeno peraltro è molto diversa nelle due proteine, perché diversi sono i requisiti che devono possedere una proteina di deposito e una di trasporto. Quest'ultima in particolare deve svolgere un compito sofisticato; deve essere in grado di legare l'ossigeno alle pressioni parziali tipiche dei polmoni (circa 100 mm Hg) e di cederne successivamente una frazione apprezzabile ai tessuti, dove la pressione parziale del gas è minore (circa 30 mm Hg). La proteina di deposito invece ha un compito più semplice, deve solo essere saturata di ossigeno alle basse pressioni parziali dei tessuti. Gli organismi peraltro hanno anche un'altra necessità, quella di allontanare l'anidride carbonica prodotta dal metabolismo.

Questi problemi sono stati risolti brillantemente dalla natura, come appare dall'analisi delle curve di legame dell'ossigeno della mioglobina e dell'emoglobina (Fig. 2). La mioglobina ha alta affinità per l'ossigeno e quindi ne è saturata anche alle basse pressioni parziali caratteristiche dei tessuti. L'emoglobina invece è saturata di ossigeno alle pressioni parziali del sangue arterioso, ma alla pressione di ossigeno del sangue venoso ha una affinità per il gas molto inferiore a quella della mioglobina e quindi lo rilascia. L'efficienza del trasporto è aumentata dall'andamento sigmoide della curva di saturazione: a basse concentrazioni di ossigeno l'emoglobina lega l'ossigeno debolmente, mentre, via via che aumenta l'ossigeno legato, anche l'affinità aumenta, ad indicare che i quattro gruppi eme che legano l'ossigeno interagiscono, cooperano in modo che il legame dei primi aumenta l'affinità per l'ossigeno degli altri. Questi fenomeni erano noti ai fisiologi fin dai primi anni del secolo. Christian Bohr e Haldane, in particolare, avevano notato anche altre proprietà dell'emoglobina, importantissime per l'organismo: a livello dei tessuti, protoni e ani-

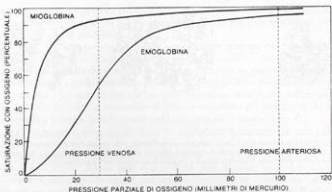


Fig. 2. Curva di ossigenazione della mioglobina e dell'emoglobina umana.

dride carbonica, prodotti dal metabolismo cellulare, favoriscono il rilascio dell'ossigeno dalla proteina, mentre a livello dei polmoni il legame dell'ossigeno favorisce il rilascio dei protoni e quello dell'anidride carbonica e quindi il suo allontanamento.

La base strutturale di tutti questi fenomeni è stata individuata da Perutz; egli osservò che nei cristalli di emoglobina ossigenata e desossigenata cambiano i rapporti spaziali fra le catene polipeptidiche, cambia cioè la cosiddetta struttura quaternaria. Quando si lega l'ossigeno, il dimero detto  $\alpha_2\beta_2$  ruota e scivola rispetto all'altro in modo da portare le catene  $\beta$  più vicine fra loro (Fig. 3). La molecola di emoglobina quindi è dotata di due possibili strutture quaternarie, una caratteristica della forma desossigenata e l'altra favorita nella forma ossigenata. La struttura ossigenata possiede una più alta affinità per l'ossigeno ed è la transizione a questo stato che spiega la cooperatività del legame. Lo scivolamento delle subunità fa sì che vengano rotti alcuni legami critici fra le subunità a livello dell'interfaccia detta  $\alpha_1\beta_2$ , e che diminuisca sia l'affinità per i protoni di alcuni residui amminoacidici sia quella per l'anidride carbonica di altri.

Descritte dunque tutte le proprietà fisiologiche dell'emoglobina in base alla variazione strutturale complessiva prodotta dal legame dell'ossigeno, Perutz ha proposto, alla fine degli anni sessanta, un meccanismo stereochimico per spiegare come l'energia del legame dell'ossigeno possa venire trasmessa dal ferro dell'eme alle interfacce fra le subunità per produrre la transizione molecolare dalla struttura desossigenata a quella ossigenata. Nell'emoglobina umana infatti i gruppi eme sono rivolti verso il solvente, lontani fra loro e dalle interfacce fra le subunità (Fig. 1).

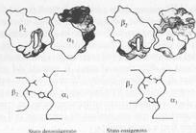


Fig. 3. Dimero  $\alpha_1\beta_2$  dell'emoglobina umana nello stato desossigenato ed in quello ossigenato. Sezione a livello del gruppo eme della catena  $\beta$  c, in basso, rappresentazione schematica delle differenze nei contatti fra i due monomeri.

Il meccanismo di Perutz prevede che possono presentare cooperatività nel legame dell'ossigeno solo le emoglobine che posseggono due requisiti: essere tetrameriche ed essere formate da due catene polipeptidiche diverse. Un dogma senza eccezioni fino alla scoperta di un'emoglobina che lega l'ossigeno in maniera cooperativa, pur essendo dimerica e non tetrameriche e formata da due catene non diverse, ma identiche. Questa emoglobina particolare, che ho avuto la fortuna di studiare con il mio gruppo, proviene da un mollusco bivalve, la *Scapharca inaequivalvis*, originario del Pacifico e che si è insediato nell'Adriatico negli anni 70. È stato subito chiaro che l'insolita presenza di cooperatività in un omodimero dovesse avere come base strutturale un assemblaggio delle subunità diverso da quello delle emoglobine tetrameriche dei vertebrati. Infatti la struttura ai raggi X ha rivelato che nel dimero di *Scapharca* i gruppi eme sono praticamente a contatto a livello dell'interfaccia fra le subunità (Fig. 1) e che l'ossigenazione non provoca grosse variazioni nella struttura quaternaria come nell'emoglobina umana, ma variazioni strutturali, limitate alla tasca dell'eme, che lasciano inalterati i rapporti spaziali fra le subunità stesse. Proprio per questa ragione, l'affinità per l'ossigeno dell'emoglobina dimerica di *Scapharca* non è sensibile a variazioni della composizione del mezzo; è una macchina molecolare molto più semplice dell'emoglobina dei vertebrati, troppo semplice per soddisfare le loro complesse esigenze fisiologiche.

Le strutture cristallografiche delle mioglobine ed emoglobine appena descritte non rivelano peraltro come l'ossigeno entri ed esca dalla tasca in cui è legato il gruppo eme che, quando vengono rappresentati tutti gli atomi, appare circondato dalla proteina. Si è persa dunque l'informazione sulla dinamica molecolare, sulla mobilità della struttura che è legata intimamente alla funzione. Movimenti molto rapidi degli amminoacidi nella tasca dell'eme, che permettono e controllano l'ac-

cesso e la fuoriuscita dell'ossigeno, sono predetti da studi teorici di simulazione al computer e possono anche essere messi in evidenza con esperimenti di cinetica rapida, con misure che si spingono nell'ambito dei nano- e dei picosecondi. Un ruolo importante hanno ad esempio le fluttuazioni dell'istidina distale che permettono la fuoriuscita dell'ossigeno. In questo contesto non posso non ricordare alcuni esperimenti pionieristici di diffrazione ai raggi X eseguiti dal gruppo di Keith Moffat utilizzando luce di sincrotrone pulsata. Sfruttando la proprietà del monossido di carbonio di essere facilmente fotodissociabile, ne è stato seguito il percorso nella tasca dell'eme dopo averlo fotolizzato direttamente in un cristallo di mioglobina. Esperimenti pionieristici, con una risoluzione temporale dell'ordine dei nanosecondi, irraggiungibile fino a pochi anni fa.

Un'altra sfida affrontata dai cristallografi con successo crescente è rappresentata dalla determinazione di strutture sempre più grandi, costituite da proteine giganti oppure da complessi di proteine diverse o di proteine con acidi nucleici. La ferritina, la proteina ubiquitaria di deposito del ferro, che ha la stessa struttura nei batteri, nelle piante, fino all'uomo è un esempio di proteina complessa. Ventiquattro subunità, ognuna di dimensioni paragonabili a quelle della mioglobina, formano una sfera cava (Fig. 4) che permette di immagazzinare fino a 4500 atomi di ferro, impedendo così a questo metallo, indispensabile per tutte le forme di vita, di formare precipitati di idrossido ferrico e di partecipare a reazioni che producono radicali estremamente tossici. Come tutte le strutture complesse, quella della ferritina è caratterizzata da diversi elementi di simmetria, assi di simmetria binaria, ternaria, quaternaria. A livello degli assi ternari e quaternari le subunità formano dei canali che permettono al ferro di entrare e uscire dal guscio proteico, a seconda delle necessità dell'organismo. La struttura della ferritina si adatta quindi perfettamente al ruolo della proteina nel metabolismo del ferro. Non sorprende pertanto che durante l'evoluzione ne sia stata conservata l'architettura caratteristica a 24 subunità.

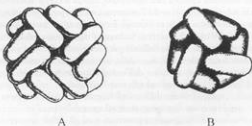


Fig. 4. Disposizione spaziale delle subunità nella ferritina. In A) il polimero canonico formato da 24 subunità con simmetria 432; in B) il dodecamero della ferritina di *Listeria innocua* con simmetria 23.

Come tutte le regole, anche questa ha un'eccezione: l'abbiamo scoperta, ancora una volta con una certa dose di fortuna, studiando la ferritina di *Listeria innocua*, un batterio Gram-positivo. Del resto, dicono i tedeschi, per fare ricerca servono tre G: Glück (fortuna appunto) Geduld (pazienza) e Geld (denaro)! Il guscio proteico di questa ferritina insolita è costituito da sole 12 e non da 24 subunità (Fig. 4). Sono peraltro conservati i canali con simmetria ternaria che permettono l'ingresso del ferro all'interno del guscio proteico. Uno degli aspetti più intriganti di questa ricerca è che la struttura della molecola, sia in termini di architettura che di sequenza amminoacidica delle subunità, somiglia a quella delle proteine Dps (DNA binding proteins from starved cells), una famiglia di proteine espresse dai batteri in condizioni di stress nutrizionale e ossidativo, che legano il DNA in maniera aspecifica, cioè senza specificità per una determinata sequenza di basi azotate. Sono stati già individuati alcuni degli elementi strutturali che rendono una proteina capace di incorporare ferro e l'altra di legare il DNA: una ricerca che si prospetta affascinante anche per le implicazioni in termini dell'evoluzione di nuove funzioni biologiche nelle proteine.

Le proteine Dps legano il DNA in maniera aspecifica. Più interessanti per il tema di oggi sono invece le proteine che interagiscono con il DNA in maniera specifica, che riconoscono cioè sequenze determinate delle basi azotate lungo la doppia elica. Le proteine che possiedono questa proprietà sono dette 'regolatrici' in quanto permettono ad esempio la replicazione e la trascrizione del DNA e regolano l'attività dei geni, selezionando quelli che devono essere trascritti in proteine e quelli che devono rimanere inattivi. Per comprendere in che modo si svolgono queste funzioni così importanti per la vita, ma estremamente complesse, la biologia strutturale si è combinata in maniera molto efficace con la genetica molecolare.

Bisogna ritornare quindi alla doppia elica di Crick e Watson per capire come, almeno in teoria, potrebbero essere riconosciute sequenze specifiche delle basi azotate. I margini delle basi, che presentano gli accoppiamenti adenina - timina, guanina - citosina, sono esposti al solvente sul fondo sia del solco maggiore che del solco minore della doppia elica. Essi presentano atomi di ossigeno e di azoto, in posizioni diverse nelle diverse coppie (Fig. 5), che possono formare legami idrogeno con residui amminoacidici di proteine, come accettori o come donatori. Possiamo dire quindi che a ogni coppia di basi corrisponde un 'codice di riconoscimento specifico' basato sul susseguirsi di atomi in grado di accettare o donare legami idrogeno, un codice che in linea di principio può essere letto direttamente da proteine che si leghino al solco maggiore del DNA. Per conferire specificità all'interazione vengono usati i legami idrogeno perché sono gli unici fra i legami deboli ad essere altamente direzionali.

È stato emozionante vedere confermate queste idee da un numero sempre crescente di strutture risolte negli ultimi 15 anni, e scoprire che proteine che svolgono la stessa funzione in organismi molto diversi usano gli stessi motivi strutturali per riconoscere il DNA e legarsi ad esso. Farò solo un esempio: il motivo strutturale noto come elica-ansa-elica presente in proteine che attivano o reprimono l'espres-

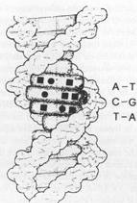


Fig. 5. Rappresentazione della doppia elica del DNA e del 'codice di riconoscimento' delle coppie di basi azotate adenina-timina, A-T, e citosina-guanina, C-G, da parte delle proteine. Sono indicati a livello del solco maggiore gli atomi in grado di donare (●) o accettare (■) legami idrogeno. La loro disposizione, diversa per ogni coppia di basi, ne permette il riconoscimento.

sione genica dai batteri fino all'uomo. Nel repressore *cro* del fago lambda, una delle due eliche, quella di riconoscimento, forma legami idrogeno specifici con le basi del DNA nel solco maggiore; l'altra elica, di posizionamento, si adagia lungo il DNA. Lo stesso motivo elica-ansa-elica si ritrova nelle proteine di insetti e vertebrati che contengono sequenze particolari dette omeodomini e che si legano al DNA per controllare i processi di differenziamento cellulare.

Vorrei finire con un accenno alle ricerche di biologia strutturale che hanno chiarito uno dei misteri dell'immunologia, cioè come un linfocita T del sistema immunitario riesca a distinguere ciò che «è proprio» da ciò che «non è proprio», che è estraneo all'organismo stesso e va distrutto. Le cellule T sono necessarie perché gli anticorpi solubili, le immunoglobuline, il secondo sistema di riconoscimento di sostanze estranee del nostro organismo, sono efficaci contro agenti patogeni extracellulari, ma danno scarsa protezione verso agenti patogeni intracellulari come i virus e i micobatteri. Le cellule T citotossiche (cellule T killer) hanno sulla membrana recettori che esaminano continuamente la superficie di tutte le cellule ed identificano quelle che mostrano marcatori estranei e che pertanto vanno uccise. Il meccanismo, che si può chiamare 'taglia e esponi' è ingegnoso: quasi tutte le cellule dei vertebrati espongono sulla superficie un campione di peptidi derivati dalla digestione di proteine presenti nel citosol; questi peptidi sono legati a proteine inte-



grali di membrana dette MHC (major histocompatibility complex) di classe I. Peptidi estranei, ad esempio peptidi virali, legati a queste proteine segnalano che una cellula è infettata e la marcano affinché venga uccisa dalle cellule T killer. La cellula T killer rilascia infatti una proteina — la perforina — che forma pori nella membrana della cellula infettata, provocando la fuoriuscita di ioni e quindi la morte della cellula. Riuscire a capire come il recettore della cellula T riesca a riconoscere contemporaneamente la molecola MHC e l'antigene virale era considerato il Graal dell'immunologia. Nel 1987 è stato fatto il primo passo; venne pubblicata da Don Wiley e collaboratori un'immagine ormai famosa in cui si vede il peptide virale adagiarsi in un solco profondo della molecola MHC, dove il peptide rimane legato tenacemente perché tutte le sue catene amminoacidiche si adattano perfettamente a piccoli incavi del solco (Fig. 6). Questa immagine più di mille parole spiega perché il peptide virale riesca a rimanere attaccato per giorni alla superficie della cellula infettata senza lasciarla mai; se così facesse sarebbe la fine della risposta immunitaria. Per il passo successivo, per comprendere come il recettore della cellula T riconosca il complesso fra peptide virale e molecola MHC di classe I sono stati necessari a Don Wiley e collaboratori ben nove anni. Il risultato è spettacolare. La struttura del complesso fra tutti e tre gli attori di questo complicato gioco di riconoscimenti mostra che il recettore interagisce, mediante anse a composizione amminoacidica variabile, sia con il peptide virale, seppellendolo quasi, sia con la molecola MHC. È interessante che la struttura tridimensionale della parte della molecola implicata nell'interazione sia nei recettori delle cellule T che nelle molecole MHC sia simile a quella delle immunoglobuline, gli anticorpi solubili, ad indicare la generalità dei principi di riconoscimento in tutte le molecole del nostro sistema immunitario ed una probabile origine evolutiva comune. Sono già molte le applicazioni

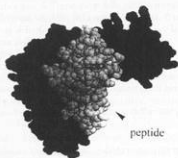


Fig. 6. Struttura del complesso fra una molecola MHC di classe I ed un peptide virale.

di queste ricerche in campo biomedico; vanno dalla produzione di vaccini virali alla comprensione dei meccanismi di rigetto nei trapianti.

In conclusione, la biologia strutturale, grazie alle molteplici interfacce con discipline che spaziano dalla biochimica alla fisiologia, dall'immunologia alla genetica molecolare, alla patologia, dà una chiave di lettura ormai indispensabile per la conoscenza dei fenomeni vitali.