

PREMIO NOBEL PER LA CHIMICA 2024

PROTEINE IN 3D

di Giorgio Colombo *

Il Nobel per la Chimica 2024 premia scoperte rivoluzionarie nella previsione e progettazione delle strutture proteiche. È stato assegnato per metà a David Baker "per la progettazione computazionale delle proteine" e per l'altra metà a Demis Hassabis e John Jumper "per i sistemi di predizione delle proteine". Baker è direttore dell'Institute for Protein Design alla Università di Washington a Seattle; Hassabis e Jumper sono gli ideatori di AlphaFold, il programma di IA sviluppato per predire la struttura tridimensionale delle proteine.



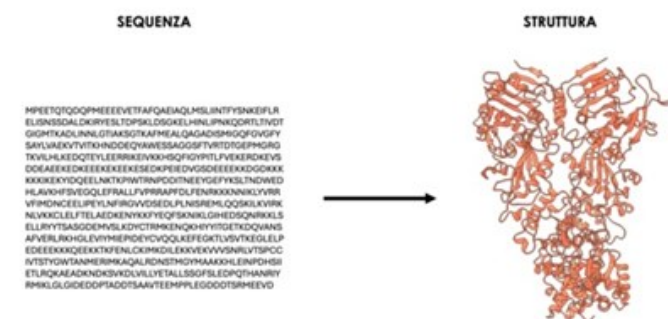
* Dipartimento di Chimica,
Università di Pavia



Le proteine sono le macchine molecolari fondamentali per la vita delle nostre cellule. Proprio come nel caso degli strumenti che usiamo nella vita di tutti i giorni, la funzione delle proteine dipende strettamente dalla loro forma tridimensionale.

Da più di cinquant'anni, chimici, fisici e biologi hanno cercato di capire come si determina la struttura a partire dalla sequenza di amminoacidi che compone ciascuna proteina.

La sequenza si può intendere come l'elenco dei pezzi che costituiscono un oggetto, con qualche indicazione generale e vaga su come certi pezzi dovranno essere vicini nello spazio ad altri, ma non fornisce nessuna informazione precisa su come montarli in una forma tridimensionale funzionante.



Da sinistra: David Baker (Seattle 1962), John Jumper (Little Rock 1985), Demis Hassabis (Londra 1976).
(© The Nobel Foundation)

Il problema della previsione della struttura

Christian Anfinsen dimostrò che la struttura tridimensionale di una proteina è determinata dalla sua sequenza di amminoacidi.

Tuttavia, prevedere questa struttura si rivelò un compito complesso. Nonostante gli sforzi con tecniche come la cristallografia a raggi X, il progresso fu lento e richiese decenni di lavoro. Decifrare e capire le regole per il ripiegamento di una certa sequenza di amminoacidi in una specifica struttura proteica rappresenta quindi una delle sfide più profonde per la ricerca con implicazioni sia nel campo fondamentale che applicativo.

In questo contesto, il Premio Nobel per la Chimica 2024 è stato assegnato a due scienziati dell'azienda *Google DeepMind*, Demis Hassabis e John Jumper, e al biochimico americano David Baker, dell'Università di Washington a Seattle, USA, per le loro scoperte rivoluzionarie nella previsione e progettazione delle strutture proteiche.

Il premio dell'Accademia Reale delle Scienze di Svezia è un riconoscimento per risultati che potrebbero trasformare la scienza e la medicina.

Metà del premio è stata assegnata a David Baker, dell'Università di Washington, per il suo lavoro pionieristico nella creazione di proteine completamente nuove, con applicazioni che spaziano dai vaccini ai nanomateriali e ai sensori avanzati. Il contributo intellettuale originalissimo di Baker, dai primi anni duemila, è stato determinante per la comprensione dei meccanismi e delle interazioni fondamentali nella determinazione della struttura delle proteine. Questo livello di conoscenza ha permesso al gruppo di Baker di generare regole per la progettazione razionale di nuove molecole proteiche non esistenti in natura.

Hassabis e Jumper, i creatori di *AlphaFold*, hanno condiviso l'altra metà del premio per aver sviluppato un modello di intelligenza artificiale capace di prevedere la struttura delle proteine a partire dalla loro sequenza chimica.

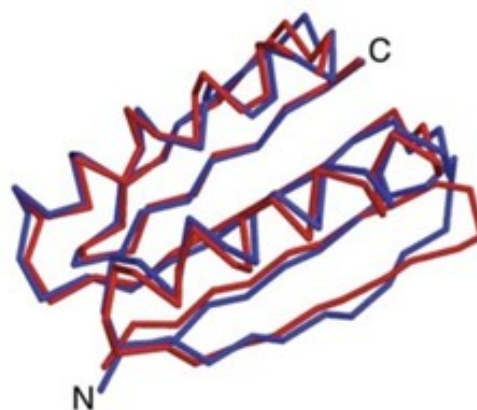
Progettazione di nuove proteine con David Baker

Il lavoro di David Baker, partendo dall'analisi chimico-fisica di come i frammenti che costituiscono le proteine si assemblano in natura, ha permesso di sviluppare modelli generali per la previsione e successivamente la progettazione di proteine. L'idea di assemblare proteine da frammenti preesistenti non è nuova in assoluto (è stata infatti inizialmente introdotta per la risoluzione delle strutture ai raggi X), ma Baker e il suo team sono riusciti a sviluppare metodi efficienti e hanno dimostrato come questa strategia possa essere utilizzata per la progettazione molecolare.

Il loro successo si basa infatti sullo sviluppo del programma informatico Rosetta. Rosetta assembla brevi frammenti strutturali da strutture proteiche non correlate con sequenze locali simili, prese da *Protein Data Bank* (www.pdb.org), una banca dati pubblica che raccoglie e annota le strutture di tutte le proteine note caratterizzate sperimentalmente. Rosetta in seguito modella simultaneamente sequenza e struttura rispetto alla conformazione della struttura finale desiderata, utilizzando algoritmi di ottimizzazione e considerando diversi fattori energetici che ricapitolano le interazioni naturali tra amminoacidi per generare insiemi di possibili strutture proteiche. Il programma genera molte soluzioni possibili e le classifica in termini di energie.

La validità di questa idea è stata dimostrata nel 2003, quando Baker è riuscito per la prima volta a progettare *ex novo* una proteina di 93 amminoacidi, chiamata *Top7*. Da allora il suo gruppo ha sviluppato molte altre con applicazioni significative in medicina e scienza dei materiali. Inizialmente l'obiettivo era creare nuove strutture proteiche ma successivamente la ricerca si è concentrata sullo sviluppo di proteine con effettive funzioni biologiche, come per esempio le attività catalitiche (enzimi), e sulla progettazione di proteine in grado di legare specificamente altre proteine o piccole molecole.

Un risultato particolarmente importante è stato lo sviluppo di enzimi completamente nuovi, progettati dal computer, per svolgere reazioni non naturali. Sebbene le prime versioni di questi enzimi fossero meno efficienti rispetto a quelli naturali, successivi cicli di evoluzione in laboratorio hanno permesso di migliorarne significativamente le prestazioni.



La struttura della molecola Top7: in blu quella prevista, in rosso quella determinata ai raggi X. (Nobel Prize Outreach)

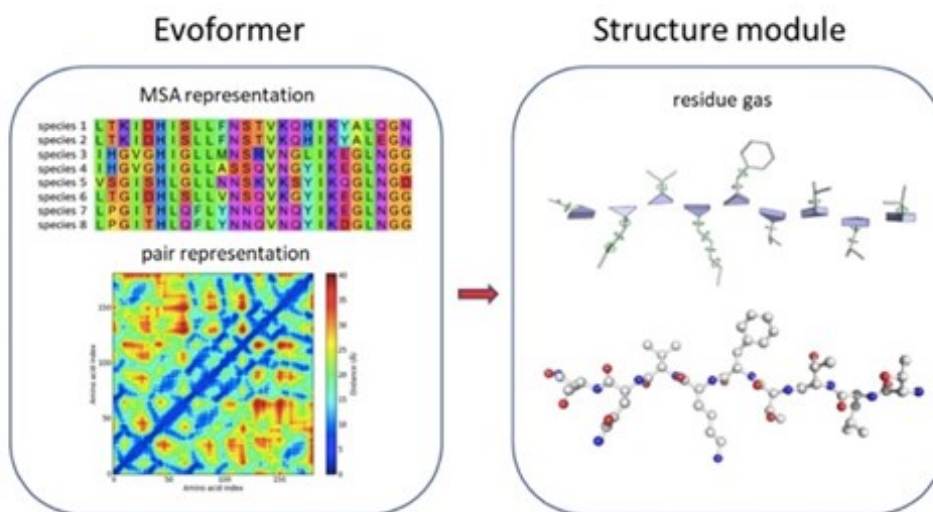
Un'altra area di successo è stata la progettazione di proteine in grado di legare molecole specifiche. Per esempio, durante la pandemia di COVID-19, le proteine progettate dal laboratorio di Baker hanno dimostrato la capacità di avere un'alta affinità per le proteine del virus, bloccandone i meccanismi di infezione. In un altro caso, sono state progettate proteine capaci di legare piccole molecole come gli steroidi, da utilizzare come biosensori. Queste proteine progettate al computer hanno dimostrato un'affinità e una selettività molto elevate per i loro bersagli.

In conclusione, il lavoro di Baker e dei suoi collaboratori ha aperto nuove prospettive nel campo della biologia strutturale e dell'ingegneria delle proteine.

AlphaFold e la previsione delle strutture proteiche

AlphaFold ha rappresentato una svolta nell'applicazione dei metodi di previsione della struttura tridimensionale di una proteina basandosi solo sulla sua sequenza di amminoacidi. *AlphaFold* è un sistema di intelligenza artificiale che prevede la struttura delle proteine e le loro interazioni con un'accuratezza dei risultati mai raggiunta prima.

AlphaFold è basato su una rete neurale con l'aggiunta di un essenziale meccanismo di attenzione per apprendere quali parti dell'*input* sono più importanti per l'obiettivo della rete. *AlphaFold* è stato addestrato su tutte le strutture del *Protein Data Bank* per produrre una mappa di distanza tra residui, o meglio una mappa di distribuzioni di probabilità per le distanze, basata su più allineamenti di sequenze tra loro simili e associate a strutture note. Da questa mappa, utilizzando regole geometriche e potenziali di interazione, si generano le strutture 3D predette.



Rappresentazione schematica dei due moduli principali di AlphaFold.
(Nobel Prize Outreach)

Un ruolo fondamentale qui è stato giocato dall'applicazione delle tecniche di *Deep Learning*. Hassabis, Jumper e il team di *Google Deepmind* sono stati infatti dei pionieri di queste tecniche e ne hanno dimostrato le potenzialità già nel 2018, quando pubblicarono il programma *AlphaZero*, che mostrava prestazioni impensabili fino a quel momento nei giochi degli scacchi o di *Go*.

L'impatto di AlphaFold è straordinario: ha permesso di prevedere la struttura di oltre 200 milioni di proteine conosciute, coprendo quasi tutte le proteine catalogate in biologia, dagli esseri umani ai batteri.

Le sue applicazioni sono già evidenti, come nello studio della resistenza agli antibiotici e nella progettazione di enzimi capaci di degradare la plastica. Inoltre, *AlphaFold* ha accelerato la ricerca su nuove proteine con applicazioni mediche, per esempio nei vaccini innovativi.

Implicazioni più ampie della scoperta

Le proteine sono fondamentali per la vita: controllano le reazioni chimiche, agiscono come ormoni e anticorpi, e costituiscono i mattoni di tessuti e organi. Comprendere e progettare le proteine è quindi essenziale per migliorare la salute e affrontare sfide globali come la resistenza agli antibiotici. Il contributo di Baker, Hassabis e Jumper non solo ha accelerato la ricerca in biologia molecolare, ma ha anche aperto la strada a nuove possibilità nella progettazione di farmaci, vaccini e materiali sostenibili.

La tecnologia AI dietro *AlphaFold* si basa su reti neurali che individuano modelli in grandi quantità di dati, utilizzando informazioni dai database di proteine conosciute e sequenze di aminoacidi. Questa metodologia rappresenta un esempio concreto di come l'intelligenza artificiale possa ampliare le conoscenze scientifiche e avere un impatto positivo sulla società. Tuttavia, Hassabis ha avvertito che l'IA è una tecnologia a doppio uso, con potenziali rischi se utilizzata in modo improprio.

Un futuro promettente con IA

Il Premio Nobel per la Chimica 2024 evidenzia il ruolo crescente dell'intelligenza artificiale nella scienza e nella medicina. Sebbene IA non sia perfetta, come dimostrato da alcune imprecisioni nelle visualizzazioni di *AlphaFold*, il suo miglioramento continuo promette di fornire strumenti ancora più potenti per comprendere e sfruttare le strutture biomolecolari. In futuro, si prevede che IA possa modellare non solo le proteine, ma anche le loro interazioni con altre strutture, come acidi nucleici e farmaci.

Il premio riconosce non solo i progressi scientifici ma anche il potenziale trasformativo della combinazione tra intelligenza artificiale e biologia. Come ha affermato il comitato del Nobel, queste scoperte consentono ai ricercatori di comprendere meglio fenomeni complessi, come la resistenza agli antibiotici, e di sviluppare soluzioni innovative per problemi globali. Questo è solo l'inizio di una nuova era nella scienza computazionale e nelle sue applicazioni pratiche.

Giorgio Colombo (Dipartimento di Chimica, Università di Pavia)

