

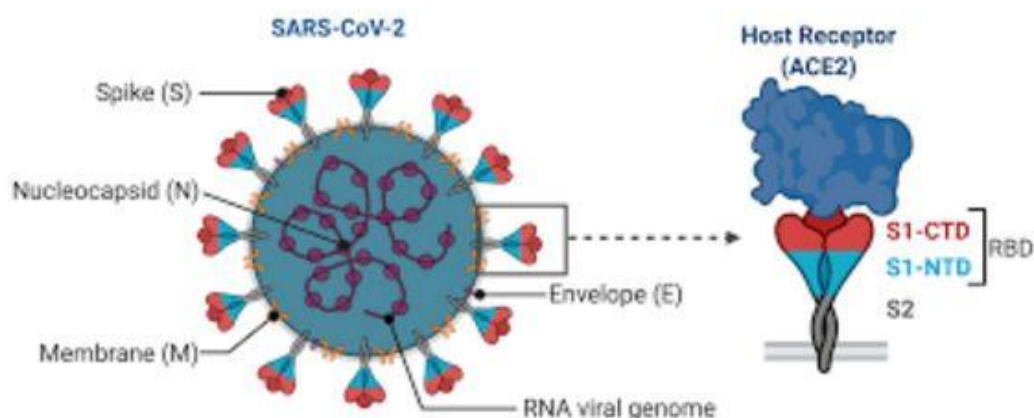
La visione di David Baker sulla struttura 3D delle proteine e di farmaci innovativi (di M. Brunori)

Con l'algoritmo Rosetta sta superando i grandi risultati prodotti da Deep-Mind's Alphafold

Soci dell'Accademia dei Lincei.

(di **Maurizio Brunori**, Professore Emerito, Sapienza Università di Roma; e Presidente Emerito della Classe di Scienze FMN dell'Accademia Nazionale dei Lincei)

Anche il mio fioraio, che incontro in fondo alla via, parlando della pandemia menziona la "spike protein", la macromolecola proteica che il virus utilizza per ancorarsi alla superficie delle nostre cellule, il primo passo per la penetrazione. *Spike* è una grande proteina di membrana formata da ben più di 1.000 amino acidi, che è presente sulla superficie del virione formando la caratteristica *corona esterna* (da cui il nome di questi virus). Il dominio esterno di *Spike* è il sito di legame al recettore di ACE2 (*Angiotensin Converting Enzyme*) presente sulla superficie delle nostre cellule, la porta d'ingresso del patogeno. La penetrazione del virus viene ovviamente contrastata dalla formazione di un legame fra *Spike* e gli anticorpi prodotti dal paziente nel corso della malattia; o anche quelli indotti dal vaccino che utilizza proprio *Spike* come antigene bersaglio.



MAURIZIO BRUNORI

FIGURA 1: (A) il virione con la corona di Spike e a dx (B) il modello della forma di Spike legata ad ACE2.

Fermo restando che il vaccino è lo strumento più efficace per evitare la malattia e innescare la prevenzione nella popolazione, è molto importante ragionare sulla necessità di inventare degli antivirali specifici ed efficaci che possano contrastare la malattia. Anche in questa prospettiva *Spike* è il bersaglio più ovvio: sarebbe essenziale

avere un farmaco specifico che sia ben tollerato ed efficace nel sequestrare il sito di legame con ACE2, il tutto da scoprire presto e a prezzi accessibili. Per tentare questa strategia si deve partire dalla conoscenza dettagliata della struttura atomica in 3 dimensioni (3D) di *Spike* e progettare nuove molecole con alta affinità per il sito di legame. Problema formidabile, che diventa aggredibile da parte di scienziati che conoscano a fondo le proteine e possano sviluppare sofisticati metodi di calcolo e simulazione *ad hoc*.

Fortunatamente le conoscenze sulla struttura e funzione delle proteine sono molto avanzate, e i metodi sperimentali per determinarne la struttura 3D a livello atomico sono molto sofisticati. Un caso relativamente recente è la microscopia elettronica a bassissima temperatura seguita da analisi delle immagini di singole molecole, tecnologia premiata con il Nobel per la Chimica nel 2017. Per la sua carica innovativa non si può non ricordare il fondamentale esperimento sulla denaturazione delle proteine pubblicato il 15 settembre 1961 (quasi 60 anni fa!) da Chris Anfinsen, biochimico danese in forza al NIH, in coll. con E. Haber, M. Sela e FH White. Perché questo esperimento compiuto agli albori della grande Biochimica è stato così importante?

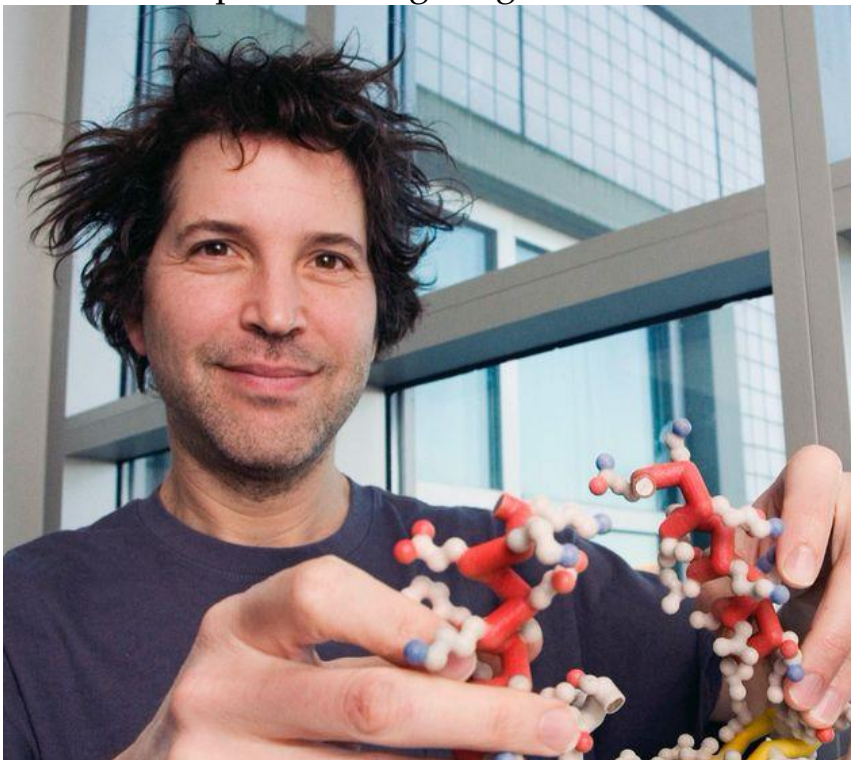
La teoria di Anfinsen si è basata sulla convinzione che la sequenza di amino acidi di una proteina contiene l'informazione necessaria e sufficiente per determinare (in acqua) la sua struttura tridimensionale caratteristica che è idonea a svolgere nella cellula la sua funzione specifica. Questa scoperta ha aperto un campo di ricerca fondamentale per la biologia e la medicina, e ha definito il cosiddetto Protein Folding Problem, in italiano il problema della morfogenesi di una proteina. La sequenza degli amino acidi di una proteina è geneticamente determinata e corrisponde -a mio avviso- alla definizione di *Sostanza* secondo Aristotele: *Substantia causa sui*: come sinolo di materia e forma, la sostanza ha in sé tutto ciò che le consente di essere quello che è. Ne consegue che la conoscenza della sequenza amino acidica dovrebbe permettere di prevedere la struttura tridimensionale di ogni proteina e quindi rivelarne la funzione. *Un vaste programme* avrebbe detto il Generale De Gaulle. Prevedere la struttura 3D di un oggetto costituito da decine di migliaia di atomi è obiettivo molto complesso ma assolutamente fondamentale, che ha attratto straordinari talenti (scienziati teorici e sperimentali) impegnati in una competizione serrata, ad altissimo livello.

Negli anni Novanta fu lanciata una interessante iniziativa internazionale denominata CASP (*Computer Assessment Structure Prediction*). Una sorta di gara aperta ai gruppi di eccellenza della biofisica e della chimica computazionale cui viene chiesto di impegnarsi nel calcolare e presentare un modello tridimensionale di proteine selezionate dal Comitato organizzatore, proteine di cui è nota la sequenza degli amino acidi ma ovviamente ignota la struttura 3D. Il Comitato invita gli scienziati partecipanti a sottoporre il risultato delle loro previsioni della struttura 3D per valutazione "in doppio ceco" da parte di un comitato di esperti. Questa gara, nata per

iniziativa del Prof Jonh Moulton (Univ Maryland, USA), ha visto per molti anni la nostra cara amica Anna Tramontano, Professoressa di Biochimica della Sapienza di Roma, impegnata come organizzatrice e “arbitra” nella valutazione dei risultati volti a risolvere un problema di estrema difficoltà.

A novembre 2020 si è concluso CASP14, che ha visto in campo circa 100 gruppi di ricerca. Alla fine della elaborata valutazione è stato rivelato un risultato eclatante: gli esperti hanno sancito la vittoria ottenuta con un algoritmo di Google detto *DeepMind's AlphaFold* che ha prodotto risultati assolutamente straordinari: il Prof John Moulton ha dichiarato senza riserve: “*This is a big deal, ... and in some sense the problem is solved*”. Come commentato in un precedente blog su HuffPost, la performance di *DeepMind's AlphaFold* sembra aver contribuito in maniera determinante alla soluzione del problema della morfogenesi di una proteina (*protein folding problem*).

Tutto risolto? Non proprio se si considerano i nuovi straordinari risultati ottenuti molto recentemente da David Baker, Professore di Biochimica e Biofisica e Direttore del *Institute for Protein Design* della Washington University a Seattle, USA (Fig. 2). Baker ha un'eccellente reputazione nel campo del *protein folding and design* dimostrata nel corso degli ultimi 20 anni da molti risultati originali e di grande impatto. Le regole fondamentali e i percorsi chimici coinvolti nella morfogenesi di una proteina sono stati inclusi in un programma di simulazione del tutto innovativo che ha chiamato *Rosetta*. Baker è un ricercatore brillante e visionario che ha affrontato con fantasia e coraggio il complesso problema del *folding* e della progettazione di proteine del tutto nuove e mai prodotte dagli organismi viventi.



MAURIZIO BRUNORI

FIGURA 2: David Baker, Professore di Biochimica alla Washington University e Direttore del Institute for Protein Design a Seattle, USA

Appena la sequenza degli amino acidi di *Spike* fu pubblicata e quindi resa disponibile alla comunità scientifica tutta, Baker e i suoi collaboratori hanno iniziato un lavoro frenetico per disegnare *de novo* piccole molecole capaci di legarsi con altissima affinità al sito di *Spike* che riconosce il recettore di ACE2 (vedi sopra). Questa strategia, supportata da una versione evoluta di *Rosetta*, ha dato ottimi risultati e permesso di sintetizzare molecole inesistenti in natura che si legano a *Spike* con altissima affinità, e che Baker ha chiamato "*hyperstable minibinders*." (Fig. 3).

Per questi risultati gli è stato assegnato nel settembre 2020 il *Breakthrough Prize*, premio molto prestigioso di 3 milioni di dollari, con la seguente motivazione: "... per aver sviluppato una tecnologia che ha consentito di disegnare nuove proteine mai prodotte in Natura, comprese nuove molecole con potenziale attività terapeutica in alcune malattie dell'uomo". (Cao L, Goreshnik I, ... Baker D. *Science*. 2020 Oct, 370: 426-431).

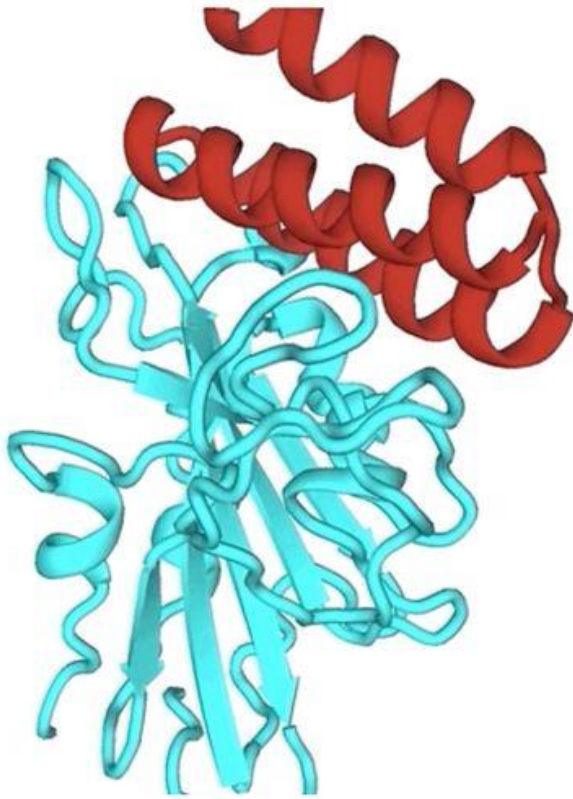
Il trasferimento di questi nuovi farmaci di sintesi alla terapia dovrà attendere i risultati dei canonici procedimenti di validazione e approvazione da parte delle Autorità vigilanti; a tempo debito i *minibinders* saranno, assieme agli anticorpi monoclonali da poco resi disponibili grazie al lavoro di Rino Rappuoli e coll., un arricchimento dell'arsenale di farmaci necessari per battere la malattia da Covid-19.

Alla fine del 2020 quando furono resi pubblici i risultati della gara CASP14, fu evidente che il vincitore era *DeepMind's AlphaFold*; e Baker era secondo. Il primo ad ammettere che il gruppo di Google aveva introdotto innovazioni concettualmente originali e molto rilevanti fu lo stesso Baker. Il gruppo della Washington University guidato da David Baker e Baek Minkyung, si mise al lavoro e in breve tempo sviluppò un sistema di intelligenza artificiale a tre canali, che considerava simultaneamente: la sequenza degli amino acidi (una dimensione), le distanze fra i vari residui (due dimensioni) e le coordinate spaziali (tre dimensioni). Questo sistema evoluto chiamato *RosettaTTAFold*, consente di ottenere modelli 3D di strutture proteiche singole e/o di complessi con un livello di accuratezza uguale a quello dimostrato dal sistema di Google *DeepMind's AlphaFold*; il vantaggio è che ottiene questo risultato molto più rapidamente (minore potenza di calcolo), a costi molto contenuti, e comunque è universalmente disponibile. Un risultato che solo all'inizio della pandemia era impensabile.

Il gruppo di Baker al *Institute for Protein Design* immagina che gli straordinari progressi nella chimica computazionale dovuti *DeepMind's AlphaFold* e a *RosettaTTAFold* hanno aperto nuovi orizzonti nelle Scienze della vita. *In primis* è ben noto che le proteine sono oggetti che hanno una dinamica molto estesa dai pico-secondi ai minuti; e che la struttura di riferimento a riposo è solo una delle conformazioni che possono modulare la funzione; basti pensare agli stati allosterici di Monod, Wyman e Changeux.

Inoltre le necessità della Medicina moderna sono assai vaste: dai vaccini universali multivalenti ai nuovi farmaci per malattie croniche come l'Alzheimer; dalle proteine

come micro-contenitori per trasporto mirato alle proteine intelligenti (*the magic bullet*) che riconoscono efficacemente il bersaglio; etc. Un universo di sfide largamente inesplorato che richiederà molta ricerca *curiosity driven*, come progettato nei Grand Challenges dell'*Institute for Protein Design*.



MAURIZIO BRUNORI

FIGURA 3: Modello stilizzato in 3 dimensioni che mostra il complesso fra il minibinder ad alta affinità (Baker e coll, Breakthrough prize) e il dominio di Spike che interagisce con ACE2.

Articolo pubblicato il 30 luglio 2021 su
<https://www.huffingtonpost.it/author/accademia-dei-lincei/>