

Immagini, modelli, realtà
La visualizzazione delle molecole delle proteine

Luigi Cerruti, Elena Ghibaudi
Dip. Chimica università di Torino

Nel mondo della vita le proteine svolgono ruoli fondamentali e l'interesse dei chimici nei loro confronti risale ai primi decenni dell'Ottocento. In effetti abbiamo la data di nascita della parola che designa questa classe straordinaria di entità molecolari. Negli anni 1830 lo svedese Jöns Jacob Berzelius era senza dubbio il chimico più noto e per molti aspetti era anche il più importante. Non solo aveva pubblicato una enorme quantità di eccellenti lavori sperimentali, ma era al centro di una rete di rapporti internazionali con i suoi resoconti annuali sulle scienze fisiche (gli *Jahres-berichte*, presenti in tutte le biblioteche universitarie) e con una fitta corrispondenza intrattenuta con gli scienziati di tutta Europa. Nella primavera e prima estate del 1838 l'olandese Gerardus Mulder aveva scritto a Berzelius di aver individuato un 'radicale' presente nella fibrina, nella caseina e nell'albume. Di fronte alla mole crescente di dati provenienti dall'Olanda Berzelius rispose il 10 luglio con una proposta: "si deve dare a questo radicale un nome particolare, ad es. proteina (*protéine*)". Il nome nacque in francese, la lingua franca di allora, e Berzelius lo giustificò con un ottimismo che si dimostrò profetico: "Vorrei derivare da $\pi\rho\omega\tau\epsilon\acute{\iota}\omicron\varsigma$ [*prōteios*, 'che occupa la prima posizione'] il nome proteina [...] perché sembrerebbe essere la sostanza primitiva o principale della nutrizione animale che le piante preparano per gli erbivori e che questi forniscono ai carnivori".¹ Il 'radicale' di Mulder risultò essere un abbaglio,² ma il nome proposto da Berzelius rimase, e l'importanza della classe di sostanze proteiche andò molto oltre quella della catena alimentare citata dal chimico svedese.

La struttura molecolare delle proteine rimase a lungo sconosciuta. Fin dal 1901 Emil Fischer aveva chiarito la natura del legame peptidico, gli amminoacidi presenti nelle proteine erano stati individuati e catalogati, era nota la presenza di ponti disolfuro ma fino agli inizi degli anni 1950 nulla si sapeva sulla reciproca disposizione dei residui di amminoacidi. Nel 1951 e 1952 il chimico inglese Frederick Sanger determinò le sequenze delle due catene presenti nell'insulina di bue (51 residui di amminoacidi in tutto), e nel 1955 portò a conclusione la determinazione delle connessioni interne alla struttura dovute a tre ponti disolfuro. La struttura dell'insulina di Sanger era ancora 'topologica', perché indicava esclusivamente le connessioni fra coppie di atomi – e non distanze e angoli. Il primo 'sguardo' sulla struttura tridimensionale di una molecola di proteina fu dato da John Kendrew e dai suoi collaboratori nel 1958. Si trattava della mioglobina (153 residui di amminoacidi), di cui la prestigiosa rivista inglese *Nature* pubblicò l'immagine di Fig. 1A. Di fronte ad una simile 'struttura' Kendrew non nasconde la propria sorpresa:

“Forse le caratteristiche più notevoli della molecola sono la sua complessità e la sua mancanza di simmetria. La disposizione [arrangement] sembra mancare quasi totalmente di

¹ Johan Erik Jorpes, *Jac. Berzelius: His Life and Work*, Berkeley: University of California Press, 1970, p. 104.

² H. B. Vickery, "The Origin of the Word Protein", in: *Yale Journal of Biology and Medicine*. 1950 May; 22(5) pp. 387–393. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2598953/>

quel genere di regolarità che istintivamente ci si aspetterebbe, ed è più complicato di quanto sia stato predetto da qualsiasi teoria sulla struttura delle proteine” (T.d.A).³

Nel 1958 le informazioni ottenute dagli scienziati inglesi non erano ancora sufficienti per determinare la posizione di ogni singolo atomo: di qui l’aspetto poco raffinato dei modelli di Fig. 1. Tuttavia l’andamento contorto ed involuto delle catene di amminoacidi era evidente, e nessun miglioramento della precisione dei dati avrebbe potuto introdurre una ‘simmetria’ che era palesemente assente. I risultati della risoluzione a livello atomico furono pubblicati nel febbraio del 1960, ancora su *Nature*.⁴ Nel titolo dell’articolo compare la parola *structure* e non più il termine *model*, presente nel titolo dell’articolo precedente, e tuttavia ciò che Kendrew e i suoi colleghi possono proporre come rappresentazione della struttura molecolare della mioglobina è ancora un modello tridimensionale costruito in laboratorio (Fig. 1B).

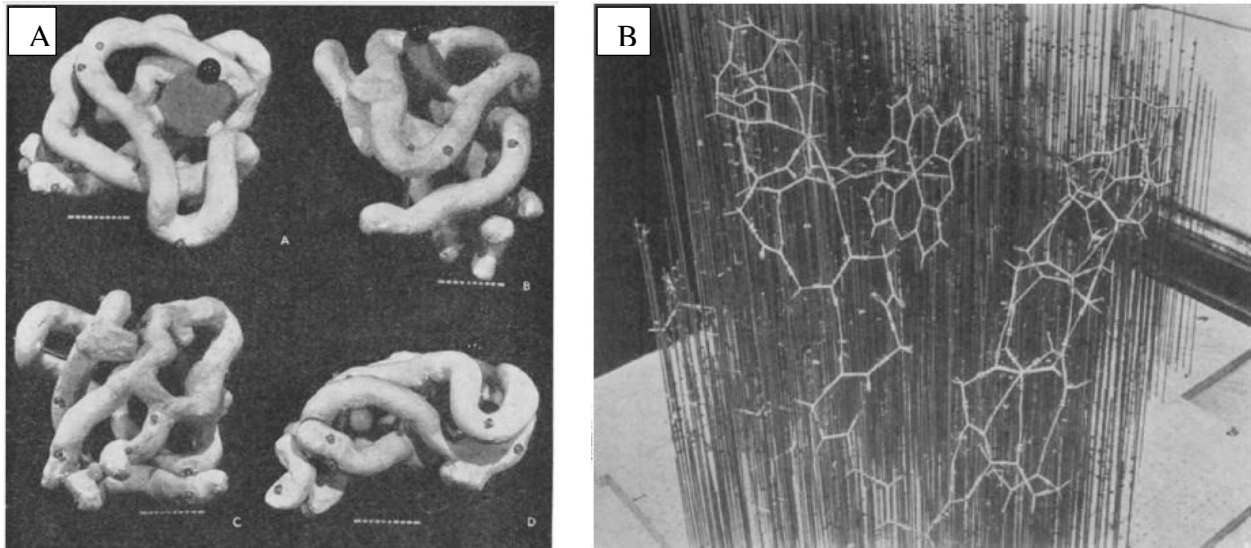


Fig. 1. Modelli strutturali della molecola di mioglobina: A) modello in plastilina della struttura secondaria a 6 Å di risoluzione, del 1958; B) modello a 2 Å di risoluzione, costruito nel 1960 e denominato ‘foresta di bastoncini’ (figure tratte da rif. 3 e 4; riprodotte su permesso di Macmillan Publishers Ltd).

Gli esiti della ricerca di Kendrew erano tecnicamente straordinari, e gli meritavano il premio Nobel per la chimica nel 1962, ma il limite conoscitivo della sua struttura era ancora difficile da colmare: "Nessun serio tentativo è stato fatto di identificare le catene laterali, e in realtà si può dubitare che sia possibile farlo sistematicamente alla attuale risoluzione". Risulta così che la ‘struttura’ determinata da Kendrew consisteva nella catena di legami polipeptidici e nella posizione del gruppo eme all’interno della struttura terziaria, mentre rimaneva parzialmente oscura la natura degli amminoacidi presenti nella catena stessa. Riflettendo sulla Fig. 1B, nel nostro contesto, emerge ancora un altro aspetto, e cioè che il gruppo di ricerca diretto da Kendrew non aveva a disposizione alcun metodo grafico di rappresentazione di una molecola complessa come quella della mioglobina.

³ "A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-ray Analysis Offsite Link" (with G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff,) *Nature* 181 (1958) 662-666.

⁴ "Structure of Myoglobin: A Three-Dimensional Fourier synthesis at 2 Å Resolution" (with R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, V. C. Shore). *Nature* 185 (1960) 422-27.

L'aumento del numero di strutture proteiche disponibili rese più urgente la necessità di una rappresentazione grafica che fosse efficace sia nella ricerca, sia nella comunicazione scientifica. Una proposta di enorme successo venne da Jane Shelby Richardson, una biofisica americana. Secondo lo standard degli scienziati americani Jane Shelby fece studi irregolari. Il suo unico titolo accademico è un Master of Arts in filosofia, conseguito ad Harvard nel 1966, ma dopo un periodo di insegnamento nella scuola secondaria Jane si associò alla ricerca del marito David Richardson, che allora stava completando il suo PhD al MIT. David era impegnato in una ricerca di cristallografia con i raggi x applicata alla determinazione di una proteina piuttosto complessa, e Jane, priva di una preparazione specifica, lavorava come tecnica di laboratorio. David faceva le misure con il diffrattometro e scriveva i programmi per l'elaborazione elettronica, mentre Jane 'accudiva' i dati (su nastri di carta e schede perforate), batteva a macchina i programmi di Dave e disegnava i contorni delle mappe elettroniche.⁵ Nel 1971 l'articolo su quella struttura fu pubblicato con la firma dei due Richardson, ma intanto essi si erano trasferiti alla Duke University, dove Jane lavorò per parecchio tempo in una semi-clandestinità, sia perché non aveva conseguito il PhD, sia perché secondo regolamento due coniugi non potevano lavorare nello stesso dipartimento. L'accudimento dei dati da parte di Jane Richardson ebbe un esito straordinario. In una intervista del 2004 Jane dichiarò: "spesi due anni per imparare come fare dei disegni che esprimessero la semplicità e l'eleganza che trovavo in quelle strutture".⁶

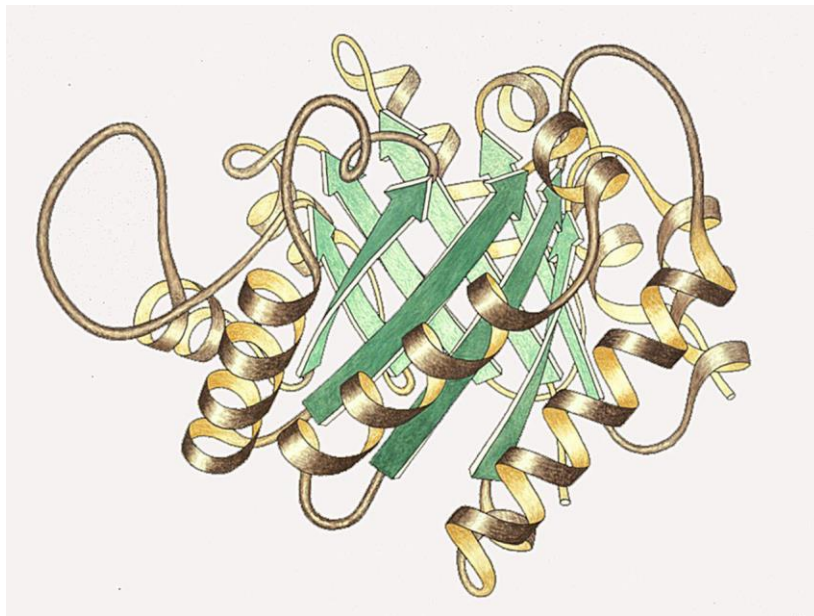


Fig. 2. Pastello di Jane Richardson, 1981. Rappresentazione della trioso fosfato isomerasi.
(https://it.wikipedia.org/wiki/Trioso_fosfato_isomerasi)

In Fig 2 abbiamo ripreso un disegno a pastello di Jane Richardson, un'immagine importante nella storia dell'iconografia scientifica. L'immagine è sfacciatamente artistica, e tuttavia con codici di lettura quasi autoevidenti (per chi sa qualcosa sulle proteine) trasmette una quantità di informazioni strutturali. Sono rappresentati otto foglietti beta (i nastri verdi) e otto eliche alfa. L'artista-scienziata è ricorsa ad alcune convenzioni visuali per comunicare la tridimensionalità della molecola proteica: sulle spirali delle

⁵ Nel cenno autobiografico da cui sono tratte queste notizie è usata la frase "coddled the data" Richardson JS, Richardson DC (2012). "Studying and Polishing the PDB's Macromolecules". *Biopolymers*. 99: 170–182 URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3535681/>.

⁶ S. Bahar, *Ribbon Diagrams and Protein Taxonomy: A Profile of Jane S. Richardson*, *The Biological Physicist*, Vol 4 No 3 August 2004, pp. 5-8.

eliche vi sono delle ombreggiature, i foglietti hanno uno spessore che mette in luce la loro torsione, gli elementi strutturali più 'lontani' dall'osservatore sono meno carichi di colore. Nell'immagine è anche accolta la convenzione per cui la catena polipeptidica inizia al terminale N e finisce al terminale C; questa convenzione (biochimica) strutturale si trasforma in una convenzione grafica attraverso la 'punta a freccia' con cui 'terminano' i foglietti. In questo modo è pure introdotto nell'immagine un certo dinamismo. Dal punto di vista della 'manipolazione' e selezione delle informazioni Jane Richardson fece una scelta radicale: gli elementi che caratterizzano la struttura di una proteina sono le forme fondamentali della struttura secondaria, le eliche e i foglietti, e la loro reciproca posizione nello 'spazio' tridimensionale occupato dalla molecola proteica; i lunghi tratti della catena polipeptidica che non si strutturano in eliche e foglietti sono rappresentati come semplici fili. Non sono prese in considerazione tutte le informazioni riguardanti i residui degli amminoacidi, comprese quelle sui gruppi laterali di evidente rilevanza spaziale.

Nell'attuale contesto scientifico e didattico, la visualizzazione delle molecole delle proteine è ottenuta quasi esclusivamente mediante grafica computazionale ed è a questo tipo di immagini che rivolgiamo ora la nostra attenzione.

La Fig. 3 riporta quattro distinte rappresentazioni della struttura della solfito reduttasi, un enzima che catalizza un'importante reazione metabolica e contiene due cofattori non proteici: un siroeme e una gabbia ferro-zolfo. E' importante sottolineare che le quattro immagini si riferiscono allo stesso 'oggetto molecolare'; ciò rende ancora più sorprendente la loro dissimiglianza. Ci si potrebbe infatti domandare quale di esse è una rappresentazione fedele dell'ente microscopico cui si riferiscono. La risposta è, inevitabilmente, tutte e nessuna. Ciascuna immagine è frutto di una scelta, operata dall'autore della rappresentazione in funzione delle caratteristiche che desidera evidenziare. La rappresentazione cosiddetta CPK (dai nomi dei suoi ideatori Corey, Pauling e Koltun) in Fig. 3A è quella che consente di visualizzare il massimo numero di informazioni sulla struttura, in quanto riporta le posizioni di tutti i singoli atomi presenti nella proteina (con la sola esclusione degli atomi di idrogeno). Questa rappresentazione non comporta alcuna selezione delle informazioni: tutto è considerato ugualmente rilevante e non vi è alcuna organizzazione gerarchica dell'informazione. Ne risulta l'immagine di un groviglio complesso che, paradossalmente, è poco utile allo sperimentatore se considerata nel suo insieme. L'immagine di Fig. 3B è costruita a partire da una logica opposta alla precedente: l'informazione che essa contiene è rigidamente selezionata. Come abbiamo già accennato in precedenza, qui si è operata una scelta drastica: tutti i residui laterali degli aminoacidi sono stati ignorati, con il significativo risultato di evidenziare i livelli di organizzazione della struttura proteica, ossia le strutture secondarie e quella terziaria. Il risultato è rilevante poiché consente una lettura molto più approfondita rispetto all'immagine del pannello A. Ad es., l'ispezione di questa immagine da parte di un occhio esperto consente di individuare zone caratterizzate da una maggior flessibilità strutturale (le α -eliche) rispetto a zone più rigide (i foglietti β), zone più variabili (i filamenti disordinati) da zone strutturalmente più conservate durante l'evoluzione molecolare, ecc. Inoltre l'immagine ci dice quale sia la relazione spaziale tra i due cofattori e la catena peptidica, fornendo suggerimenti utili per comprendere le modalità di funzionamento di questo complesso sistema molecolare. La Fig. 3C ci restituisce un'immagine del tutto diversa: il codice scelto per la rappresentazione è radicalmente differente dai precedenti. Qui ogni atomo è rappresentato come una sfera di van der Waals ed è anche presente un codice di colore che consente di differenziare alcuni atomi tra loro. L'impressione che ne ricaviamo è quasi opposta alla precedente: laddove c'erano dei 'vuoti', qui troviamo dei 'pieni'; laddove l'occhio penetrava all'interno del 'corpo proteico', qui si ferma alla superficie. Le scelte operate per la costruzione delle immagini 3B e 3C sono nettamente distinte, ma evidentemente non si tratta di mere scelte estetiche. Esse rispondono a interessi di ricerca differenti: in Fig. 3B ci si focalizza sui rapporti strutturali interni alla molecola proteica, mentre la Fig. 3C ci restituisce informazioni utili

ad indagare i rapporti della molecola con il suo intorno. Ciò è altrettanto chiaro nella Fig. 3D, che mette perfino in evidenza la presenza di superfici, rugosità, anfratti, della cui esistenza le immagini precedenti non lasciavano sospettare. Lo scopo è, ancora una volta, fornire informazioni utili allo studio delle interazioni tra questa e altre molecole. E tuttavia, risulta quasi impossibile non chiedersi quale sia lo statuto di realtà di questa accidentata superficie, data la relazione con i dati sperimentali a partire dai quali è stata generata. Si torna allora alla domanda posta poco sopra: quale di queste rappresentazioni è vera o, almeno, veridica?

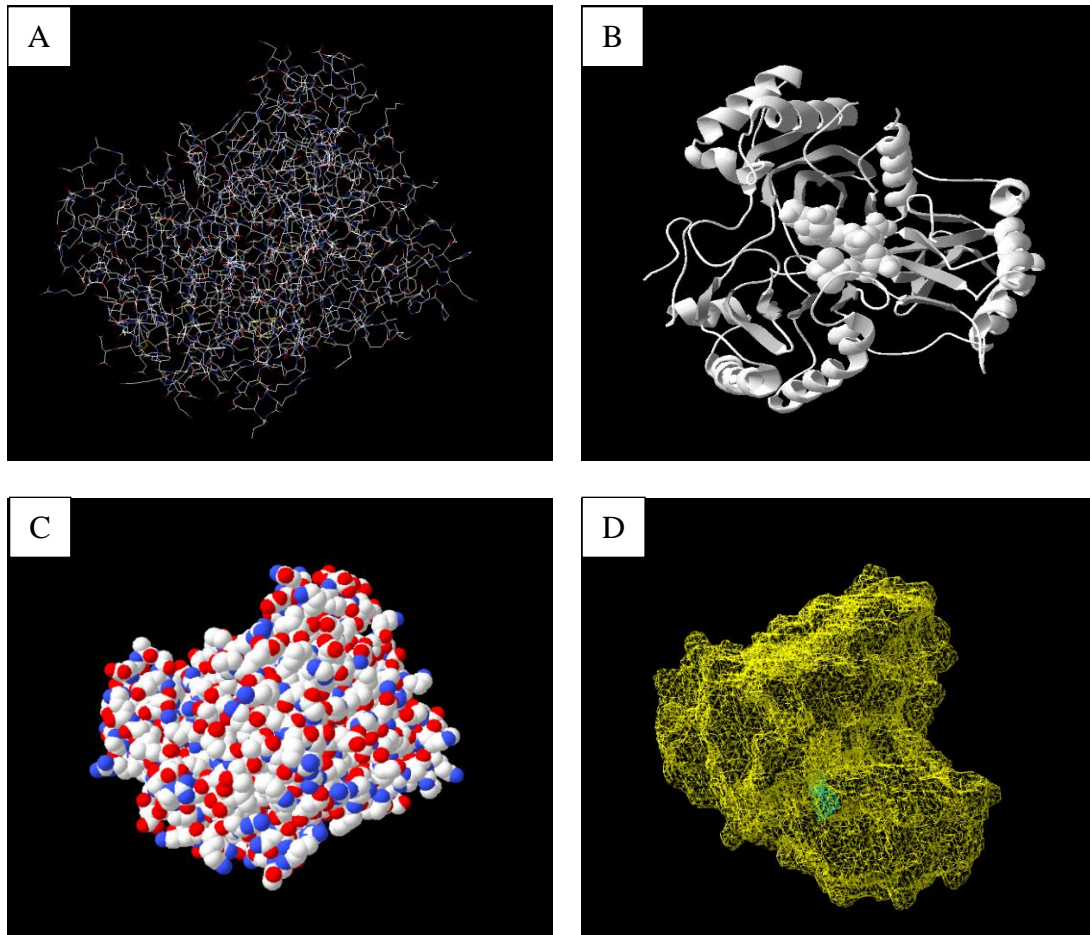


Fig. 3. Rappresentazioni della struttura molecolare dell'enzima sulfito reduttasi (immagini costruite dagli autori con il programma di grafica molecolare PDB Viewer 4.1.0 a partire dalle coordinate atomiche depositate nel Protein data Bank, 1AOP): A) rappresentazione CPK; B) rappresentazione a nastro; C) rappresentazione a raggi di van der Waals, con atomi di ossigeno e azoto rappresentati in rosso e blu, rispettivamente; D) rappresentazione della superficie di potenziale della proteina.

La risposta va cercata ragionando sul piano epistemologico. Nel commentare i due lavori di Kendrew, abbiamo sottolineato il passaggio dal termine *model* (nel 1958) al termine *structure* (nel 1960). La tentazione di considerare un modello di struttura molecolare ottenuto per diffrazione di raggi X come una riproduzione fedele della realtà microscopica è molto forte ed è alimentato dalla potenza di questa tecnica, che ha consentito progressi impressionanti nel campo della strutturistica biochimica e nella comprensione del mondo biologico. Tuttavia, qualunque rappresentazione grafica di una struttura microscopica costituisce un tentativo di rendere visibile un mondo che, nei fatti, non è visualizzabile. Questo è forse l'aspetto più sorprendente di qualsiasi immagine riferita al mondo atomico-molecolare:

la capacità di rappresentare ciò che è intrinsecamente impossibile visualizzare. In questo senso, le immagini di proteine che abbiamo appena esaminato sono tentativi di rendere visibili delle astrazioni, quali le proprietà chimico-fisiche (il potenziale elettrostatico, il carattere acido o basico di un sito, ecc.) e topologiche (le relazioni spaziali tra atomi, la presenza di strutture secondarie, ecc.) del sistema in oggetto. Consideriamo l'esempio della superficie di potenziale di una proteina: quella superficie non esiste materialmente; non è altro che la trasposizione grafica di una serie di dati teorici, calcolati a partire da dati sperimentali. E tuttavia, se questi dati fossero presentati in forma di tabella, la loro utilità sarebbe minima in quanto non saremmo in grado di definirne l'orizzonte di senso. Grazie alla loro rappresentazione visiva, si rende disponibile uno scenario all'interno del quale il fruitore dell'immagine può muoversi. Ad esempio, diventa possibile formulare ipotesi riguardo alle possibilità di interazione di una specifica porzione della superficie proteica con substrati o effettori di vario genere, con i conseguenti effetti sul funzionamento del sistema enzimatico. Detto in altri termini, le rappresentazioni grafiche del mondo microscopico hanno un potente valore euristico, che le configura come strumenti atti a suggerire ipotesi che saranno poi verificate sperimentalmente. Ciascuna delle quattro rappresentazioni di Fig. 3 è in grado di suggerire ipotesi distinte e risponde ad esigenze diverse, a seconda delle intenzioni di ricerca del loro autore e fruitore. Nessuna di esse è una fotografia⁷ della realtà: si tratta di modelli, ciascuno dei quali partecipa della realtà stessa in modo peculiare, facendosi strumento ermeneutico della stessa. Del resto, qualsiasi modello scientifico è un tentativo di interpretare la realtà materiale e uno strumento che ci consente di anticiparne il comportamento. Non si pone dunque il problema di stabilire se un modello sia vero, ché non potrà mai esserlo: un modello può solo essere più o meno efficace nell'assolvere un ruolo interpretativo, predittivo ed euristico. In questo senso, le rappresentazioni di enti atomico-molecolari non fanno eccezione.

Un ulteriore spunto di riflessione ci è offerto dalla Fig. 4, che raffigura ciò che - nel gergo tecnico - prende il nome di *esperimento di docking*, ossia un esperimento virtuale che simula la formazione di un addotto proteina-effettore. Una prima osservazione, conseguente a quanto esposto finora, riguarda proprio l'utilizzo dei modelli strutturali: essi rendono possibile di verificare (dapprima virtualmente e poi in laboratorio) la possibilità di interazioni specifiche del sistema con varie molecole. Ma ci interessa qui soffermarci su un aspetto interpretativo che ci pare peculiare del linguaggio iconico applicato al mondo atomico-molecolare. La rappresentazione iconica di set di dati sperimentali implica la possibilità di attribuire a questi dati un significato. Uno degli strumenti di questa attribuzione di significato è l'analogia. Accade così di riferirsi a canali, solchi, cavità, tasche all'interno di una proteina, prendendo a prestito termini del linguaggio naturale che sono abitualmente riferiti al mondo macroscopico e che, se presi alla lettera, non potrebbero in nessun modo essere applicati all'universo atomico-molecolare. Eppure è proprio grazie a queste analogie che riusciamo ad attribuire un senso alle osservazioni sperimentali. La Fig. 4 diventa così la rappresentazione dell'accomodamento di una molecola di inibitore all'interno di una *cavità* della proteina, il sito attivo. Grazie a questa immagine riusciamo a ricavare informazioni riguardo al meccanismo attraverso il quale l'inibitore esplica la sua azione sulla neuraminidasi: esso stabilisce contatti con specifici aminoacidi, generando un complesso stabile che non è cataliticamente attivo. Similmente, tutti siamo familiari con il *solco maggiore* e il *solco minore* del DNA o con i *setacci molecolari*: esempi dell'utilizzo dell'analogia come strumento ermeneutico e descrittivo.

⁷ Utilizziamo qui il termine *fotografia* nella impropria accezione di 'riproduzione oggettiva della realtà'. Siamo tuttavia ben consci che una fotografia è ben lungi dall'essere oggettiva ed è sempre frutto delle scelte del fotografo. Non per nulla, una stessa scena, fotografata da persone diverse o con distinte intenzioni, può apparire assai difforme.

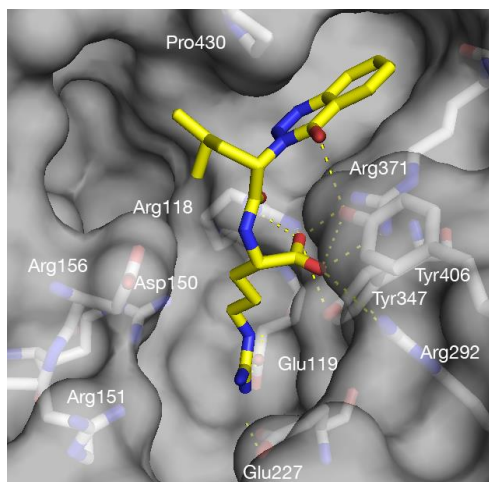


Fig. 4. Rappresentazione del complesso della neuraminidasi con un inibitore
 (<http://hermes.chem.ut.ee/~alfx/research.html>)

In conclusione, con questi esempi abbiamo voluto mettere in evidenza alcune caratteristiche inerenti le rappresentazioni grafiche del mondo microscopico, il loro significato e il loro utilizzo da parte dei ricercatori. Le immagini del mondo atomico-molecolare non sono (e non possono essere) riproduzioni oggettive di un mondo che è intrinsecamente irrepresentabile e inaccessibile agli occhi. Sono dunque delle *invenzioni*, dei prodotti della creatività e delle scelte di coloro che le producono, il cui scopo è quello di rendere accessibile e comprensibile il mondo molecolare. Il termine *invenzione* applicato all'ambito scientifico suscita spesso diffidenza, in quanto giudicato improprio di una pratica – quella scientifica – che agli occhi di molti ha come scopo lo svelamento della realtà *quale essa è*. Non apriremo qui una discussione sulla natura e gli scopi della pratica scientifica. E' tuttavia opportuno chiarire che utilizziamo il termine *invenzione* nell'accezione di prodotto dell'ingegno, ossia il risultato di un esercizio di creatività e capacità interpretativa. Queste immagini non sono frutto dell'arbitrarietà soggettiva: rispondono a regole di codifica ben precise e condivise dalla collettività scientifica e sono prodotte a partire da dati sperimentali che ne sanciscono l'inequivocabile rapporto con la realtà materiale che esse vogliono rappresentare. Tuttavia, sono anche frutto di scelte individuali e di creatività, che rispondono forse ad una ricerca di bellezza, come ben risulta dai testi di Jane Richardson. In questo senso, non sono dissimili dall'opera di un artista.

Ci piace chiudere citando la motivazione che la Reale Accademia delle Scienze svedese ha formulato per l'assegnazione del premio Nobel per la chimica 2017 ai ricercatori che hanno sviluppato la tecnica della microscopia crio-elettronica, la quale consente di ottenere modelli di strutture supramolecolari di grande complessità: “Un'immagine è una chiave di comprensione. Fondamentali passi avanti nella scienza poggiano sulla possibilità di visualizzare oggetti invisibili all'occhio umano”⁸. E' incontestabile che una parte dei progressi realizzati negli ultimi decenni nella comprensione del comportamento delle biomolecole e delle relazioni tra la loro struttura e la loro funzione siano stati resi possibili dal ricorso a rappresentazioni iconiche.

⁸ Royal Swedish Academy of Sciences: Press release, The Nobel Prize in Chemistry 2017. URL: https://www.nobelprize.org/nobel_prizes/chemistry/laureates/2017/press.html (2017).