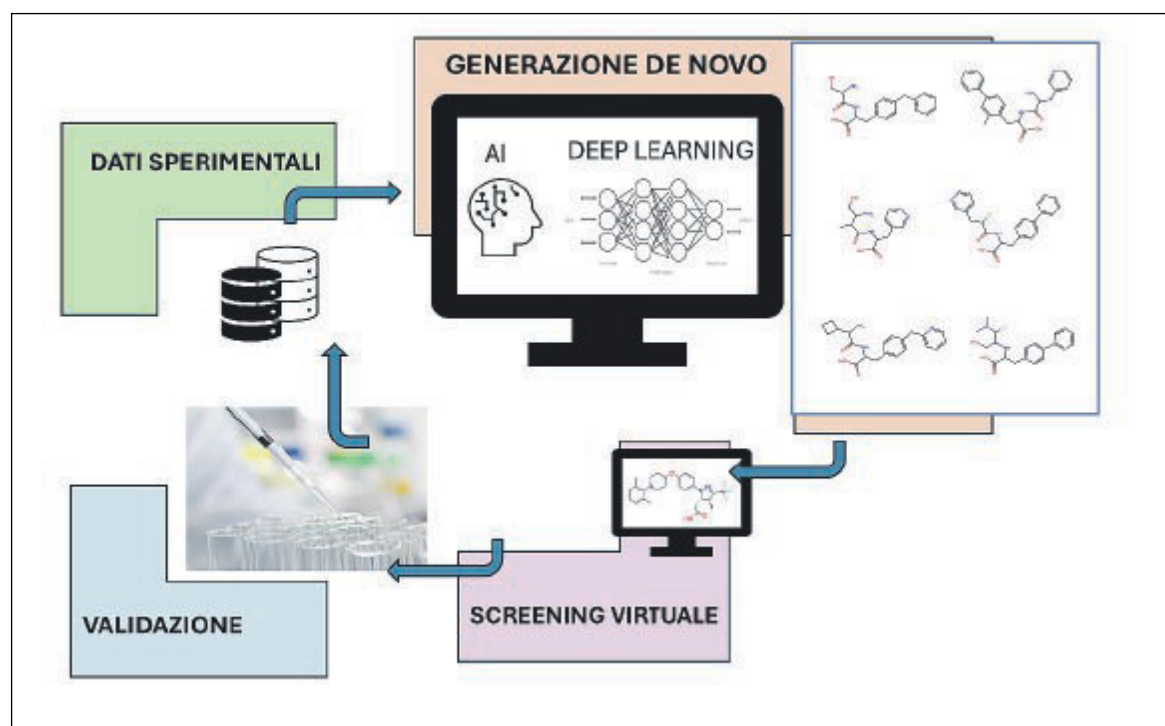




# Consiglio Nazionale delle Ricerche



**INTELLIGENZA ARTIFICIALE**  
Il processo per l'individuazione di nuovi candidati farmaci sviluppato da un gruppo di ricerca che include ricercatori dell'Istituto di Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato (Stima) e dell'Istituto di Cristallografia (Ic) del Consiglio Nazionale delle Ricerche

## Il connubio Gazzetta-Cnr

● La collaborazione e la sinergia fra Gazzetta e Consiglio Nazionale delle Ricerche ha compiuto tre anni. Finora, in ben 54 puntate il nostro giornale ha ospitato le pubblicazioni dei lavori di ricerca del Cnr barese, spaziando tra discipline e istituti.

Oggi pubblichiamo la 55ª puntata. Le precedenti uscite hanno riguardato altrettanti lavori di ricerca realizzati da: Istituto per i Processi Chimico-Fisici (Ipcf), Istituto di Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato (Stima), Istituto di Cristallografia (IC), Istituto ISPA (Istituto di Scienze delle Produzioni Alimentari), Istituto di Ricerca per la Protezione Idrogeologica (Irpi), Istituti Nanotec e Processi chimico-fisici, Istituto di Biomembrane, Bioenergetica e Biotecnologie Molecolari, Istituto di Bioscienze e Biorisorse (IBBR), Istituto di chimica dei composti organometallici (Iccom), Istituto di Ricerca sulle Acque, Istituto per il Rilevamento Elettromagnetico dell'Ambiente (Irea) dell'Istituto per la Scienza e Tecnologia dei Plasmi (Istp), Istituto di Tecnologie Biomediche (ITB), dell'Istituto per le Tecnologie della Costruzione (Itc) e «Matematica per l'Ambiente» dell'Istituto per Applicazioni del Calcolo di «Mario Picone» (Iac-Cnr), dell'Istituto sui Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato (Stima) con l'Isipa di Foggia e Isp-Cnr, di Irpi-Cnr e Uniba, Istituto per la Scienza e tecnologia dei plasmi (Istp), dell'Istituto di fotonica e nanotecnologie (Ifn), dell'Istituto Cnr Nanotec, dell'Istituto di Cristallografia e dell'Istituto di Scienze delle Produzioni Alimentari (Isipa), dell'Istituto di Biomembrane, Bioenergetica e Biotecnologie Molecolari-Cnr, dell'Istituto per il Rilevamento elettromagnetico dell'ambiente (Irea), del gruppo Osservazione della Terra dell'Istituto sull'Inquinamento atmosferico (Iia) e infine dell'Istituto di chimica dei composti organometallici (Iccom).

# Da proteine, molecole e AI le nuove frontiere dei farmaci

Col modello Prot2Drug è possibile progettare nuove cure ed evoluzioni delle attuali

● Il processo di scoperta di nuovi farmaci è lungo e costoso e funestato da alti tassi di fallimento soprattutto nelle fasi di sperimentazione clinica. Contribuire a individuare nuovi potenziali farmaci in maniera più veloce, economica e precisa è una delle sfide dell'intelligenza artificiale (IA) che permette di automatizzare varie fasi del processo. Una delle tecnologie chiave è l'applicazione ai linguaggi della chimica e della biologia della «Elaborazione del linguaggio naturale» (Nlp) che permette ai computer di leggere, interpretare e generare testi; proprio come farebbe un essere umano.

Le proteine, componenti essenziali della vita, hanno infatti un loro linguaggio, che comprende un alfabeto costituito da 20 aminoacidi. Anche le strutture delle piccole molecole organiche, come i farmaci, possono essere rappresentate come sequenze di lettere e simboli. Grazie alla nlp, oggi possiamo ana-

cidi e candidati farmaci.

A differenza di approcci più tradizionali che si basano sulla conoscenza di noti ligandi della proteina o della sua struttura tridimensionale, Prot2Drug si basa sulla rappresentazione fornita da un modello linguistico della proteina per analizzare la sua sequenza e capire le sue proprietà biochimiche, proprio come se interpretasse un testo scritto in un linguaggio biologico.

I ricercatori del CNR hanno addestrato Prot2Drug su un enorme dataset di oltre 300mila coppie di proteine e molecole sperimentalmente validate, ottenendo un'IA capace di «tradurre» una proteina nella sua molecola ideale, permettendo così la progettazione di farmaci con alta affinità per il bersaglio. Possiamo immaginare che ogni proteina contenga la ricetta da seguire per progettare molecole che possano legarla e Prot2Drug utilizza tali istruzioni per proporre nuove molecole utilizzando la sintassi del linguaggio chimico.

Prot2Drug non solo è in grado di riprodurre interazioni già note tra farmaci e proteine, ma può anche suggerire nuovi bersagli per farmaci esistenti,

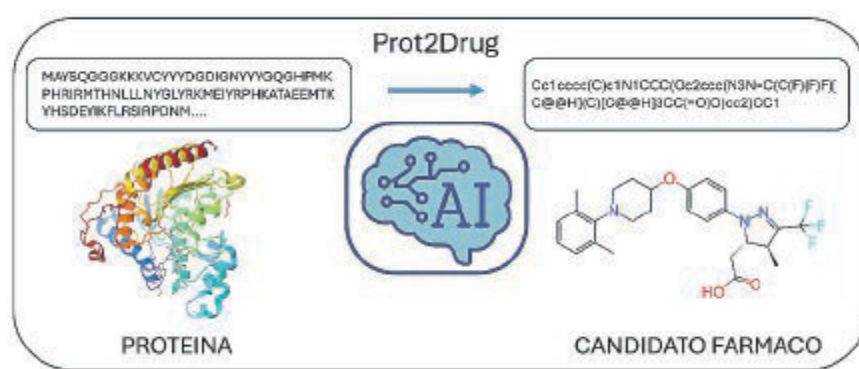
aprendo la strada al riposizionamento di farmaci – una strategia che accelera la scoperta di nuove cure.

Un punto di forza è la sua capacità di progettare promettenti ligandi anche per proteine la cui struttura tridimensionale non è nota e di cui non sono note altre molecole in grado di legarle. Questo lo rende uno strumento promettente per esplorare nuove frontiere nella ricerca farmaceutica.

Lo studio è stato finanziato dall'Unione Europea - Next Generation EU nell'ambito dei progetti di ricerca PRIN2022 «Development of broad-spectrum coronavirus antiviral agents acting as allosteric modulators of the host protein sigma-1» (Codice 2022Z3BBPE); Prin 2022 Pnrr «AID-CARE: Artificial intelligence-aided design of cannabinoid receptor subtype 2 (CB2R) Multitarget ligands: a new strategy for the treatment of inflammation» (Codice P2022TRR3Y) and Mur 3264/2021 PNRR M4/C2/L3.1.1 «Potentiating the Italian Capacity for Structural Biology Services in Instruct Eric (ITACA.SB)» (Project no. IR0000009).

**Teresa Maria Creanza, Domenico Alberga, Cosimo Patruno, Giuseppe Felice Mangiatordi, Nicola Ancona**  
Istituto di Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato (STIIMA); Istituto di Cristallografia (IC)

Il Prot2Drug si basa sulla rappresentazione fornita da un modello linguistico della proteina per analizzare la sua sequenza e capire le sue proprietà biochimiche



lizzare e comprendere la grammatica molecolare di questi linguaggi, ovvero le regole con cui gli atomi si combinano per formare nuove molecole. E non è tutto: i modelli linguistici avanzati possono decifrare i «linguaggi segreti» delle proteine e delle piccole molecole e suggerire nuove strade per progettare farmaci innovativi.

Un esempio di successo è Prot2Drug, un modello di intelligenza artificiale sviluppato da un gruppo di ricerca che include ricercatori dell'Istituto di Sistemi e Tecnologie Industriali Intelligenti per il Manifatturiero Avanzato (STIIMA) e dell'Istituto di Cristallografia (IC) del Consiglio Nazionale delle Ricerche. Prot2Drug è stato addestrato per progettare nuovi candidati farmaci a partire dalla semplice sequenza degli aminoacidi (sequenza primaria) di una proteina bersaglio. I risultati dello studio sono stati pubblicati sulla prestigiosa rivista scientifica Journal of Chemical Information and Modeling. Il modello Prot2Drug sfrutta la potenza dei transformer, modelli di IA ispirati alla traduzione linguistica, per tradurre il «linguaggio» delle proteine in quello dei farmaci in grado di modularne la funzione. In pratica, è come se l'IA traducesse una frase da una lingua all'altra usando, al posto delle parole, amino-



### LO STUDIO DEL CNR BARESE

Sopra il gruppo di ricerca: Nicola Ancona, Domenico Alberga, Giuseppe Felice Mangiatordi, Teresa Maria Creanza, Cosimo Patruno

Al centro lo schema di funzionamento del modello di Intelligenza Artificiale Prot2Drug