



A CURA DI SILVIA CAUTERUCCIO E MONICA CIVERA
 DIPARTIMENTO DI CHIMICA
 UNIVERSITÀ DI MILANO
 SILVIA.CAUTERUCCIO@UNIMI.IT
 MONICA.CIVERA@UNIMI.IT

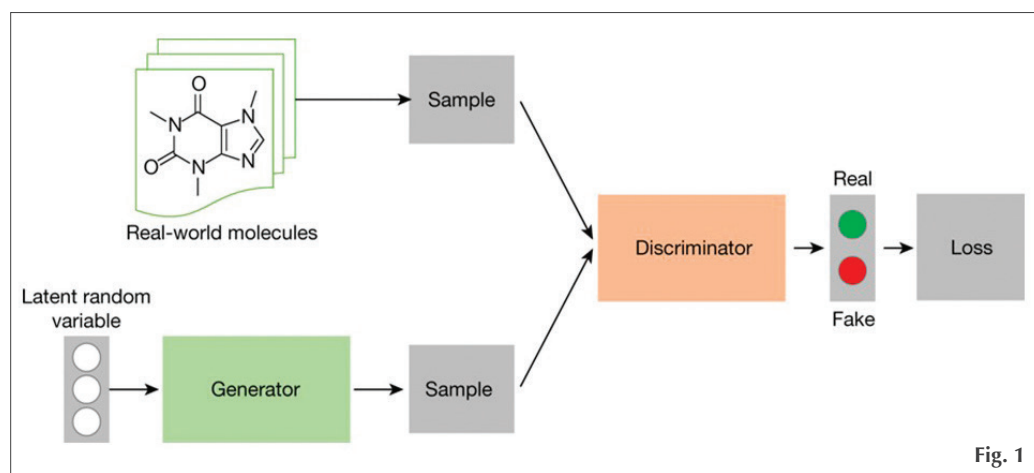


Fig. 1

cialmente nel campo della progettazione di farmaci o composti biologicamente attivi. Nel metodo *generative adversarial network* (GAN, Fig. 1), un tipo di apprendimento *unsupervised*, due modelli sono generati contemporaneamente: uno generativo (*generator*), che descrive la distribuzione dei

Intelligenza artificiale: i metodi *machine-learning* e le scienze molecolari

La combinazione di intelligenza artificiale e *big data*, nota come il quarto paradigma della scienza o come la quarta rivoluzione industriale, sta modificando il ruolo e l'utilizzo dei computer in diversi settori scientifici. Recentemente i metodi *machine-learning*, avendo la potenzialità di risolvere problemi molto complessi spesso non trattabili con le classiche metodologie, stanno avendo molto successo in diversi campi delle scienze molecolari e materiali [K.T. Butler, *Nature*, 2018, **559**, 547]. Rispetto ai metodi tradizionali in cui il computer viene usato come un calcolatore potenziato che esegue un algoritmo fornito da un utente, negli approcci *machine-learning* le regole o connessioni di un *dataset* sono apprese valutando una porzione di dati (*training dataset*) da cui il modello impara a fare predizioni senza bisogno di altri *input* forniti dall'utente. Il cuore di queste tecniche è rappresentato da algoritmi statistici la cui performance migliora con la pratica. Se ai dati in *input* sono associati dei valori di *output*, il modello *machine-learning* è detto *supervised* e fornirà una predizione la cui accuratezza dipende dai dati forniti (*training set*), nel caso siano disponibili solo dati di *input* si parla di apprendimento *unsupervised* e il modello organizza i dati secondo caratteristiche e *trends*. I metodi *machine-learning* hanno una lunga storia nel campo dello sviluppo delle relazioni struttura-attività, spe-

cialmente nel campo della progettazione di farmaci o composti biologicamente attivi. Nel metodo *generative adversarial network* (GAN, Fig. 1), un tipo di apprendimento *unsupervised*, due modelli sono generati contemporaneamente: uno generativo (*generator*), che descrive la distribuzione dei dati, e uno discriminativo (*discriminator*), che stima la probabilità che un campione venga dal *training set* o dal *generator*. Per il *generator*, la procedura tende a rendere massima la probabilità che il *discriminator* faccia un errore.

Nel caso dei metodi *deep reinforcement learning*, come il ReLeaSe [M. Popova, *Sci. Adv.*, 2018, **4**(7), 4:eaap788], le molecole vengono disegnate con le desiderate proprietà chimiche e biologiche risolvendo problemi decisionali di programmazione dinamica. La procedura di ReLeaSE, include due *deep neural networks* (generativo, G, crea nuove molecole, e predittivo, P, ne determina le proprietà) e un processo di *training* a due stadi: nella prima fase i due modelli vengono addestrati separatamente mentre nella seconda fase insieme, utilizzando un approccio di *deep reinforcement learning* che ottimizza le proprietà delle molecole generate assegnando un *reward* numerico. Questo parametro dipende dalla molecola generata dal modello predittivo e il modello generativo è addestrato per massimizzare il *reward* atteso. Le librerie chimiche vengono disegnate a partire dalle stringhe SMILES, sono sinteticamente accessibili e hanno determinate proprietà.

La fotocatalisi nell'attivazione di legami C-H e della CO₂

Continuando il nostro viaggio alla ricerca di metodologie sintetiche efficienti, selettive ed eco-sosteni-



bili per la formazione di legami C-C, un approccio molto valido ed efficace per condurre, ad esempio, reazioni di ossidoriduzione in condizioni blande è rappresentato dalla fotocatalisi. I catalizzatori *fotoredox* (FC), costituiti da complessi organometallici o da coloranti organici, sono in grado di assorbire luce nella regione del visibile, fornendo degli stati fotoeccitati (FC*, Fig. 2a) caratterizzati da tempi di vita lunghi. Questi stati eccitati sono a loro volta coinvolti in processi di trasferimento a singolo elettrone (*single-electron-transfer*, SET) per donare/catturare un elettrone a/da un substrato, convertendolo in specie radical anioniche Sub^{-•} o radical cationiche Sub^{+•} altamente reattive (Fig. 2a). Il fotocatalizzatore sarà poi ripristinato ad opera di un opportuno riducente od ossidante presente nell'ambiente di reazione. Tra le diverse tipologie di reazioni in cui è stata utilizzata la fotocatalisi, notevoli risultati sono stati ottenuti nell'attivazione di legami C-H, soprattutto nella funzionalizzazione di legami C(sp²)-H attraverso reazioni di arilazione diretta, (perfluoro)alchilazione e acilazione [J.-F. Soulé, *Chem. Rev.*, 2018, **118**, 7532]. Rispetto alla funzionalizzazione diretta di legami

C-H di (etero)areni e olefine mediante complessi a base di metalli di transizione, la fotocatalisi permette di operare in condizioni decisamente più blande, spesso a temperatura ambiente, portando alla formazione di legami C-C con elevata chemo, regio e stereoselettività.

Oltre all'attivazione di legami C-H, la catalisi *fotoredox* è una strategia altrettanto promettente anche per la formazione di legami C-C utilizzando direttamente l'anidride carbonica come sintone C1 [C.S. Yeung, *Angew. Chem. Int. Ed.*, DOI 10.1002/anie.201806285]. La CO₂, che è una molecola a basso costo e non tossica, sta suscitando notevole interesse in sintesi organica perché permette di accedere direttamente ad acidi carbossilici, tra cui amminoacidi diversamente funzionalizzati, quali precursori chiave di peptidi, proteine e prodotti naturali. Queste reazioni di carbossilazione richiedono generalmente l'impiego di reagenti altamente reattivi, a causa dell'elevata stabilità cinetica e termodinamica della CO₂. È possibile però accelerare significativamente l'inserzione diretta di CO₂ sfruttando meccanismi di tipo radicalico. In questo contesto la fotocatalisi è

sicuramente tra le metodologie più appropriate, perché permette un facile accesso a specie a singolo elettrone attraverso l'impiego della luce. A titolo esemplificativo menziono una procedura altamente regio e chemoselettiva per la sintesi di α-amminoacidi α,α' disostituiti, ottenuti mediante l'idrocarbossilazione di enammidi ed immine in atmosfera di CO₂ a temperatura ambiente [T. Ju, *Angew. Chem. Int. Ed.*, DOI 10.1002/anie.201806874], utilizzando un derivato dicianobenzenico (4CzIPN, Fig. 2b) quale fotocatalizzatore organico.

