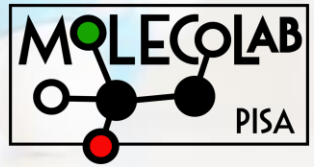
Dalle reazioni chimiche alle proteine

Chimica in 3D nell'era dell'intelligenza artificiale

Lorenzo Cupellini



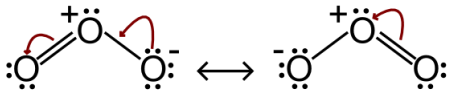


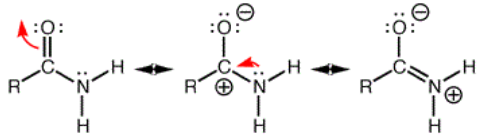

IN SUPREMÆ DIGNITATIS
1343
DCCI
MOLECOLAB
PISA

La struttura delle molecole

Come insegnamo la struttura molecolare?

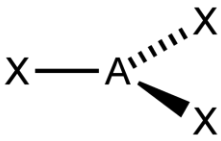
Formule di risonanza

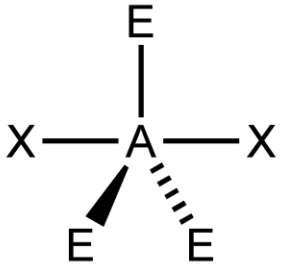


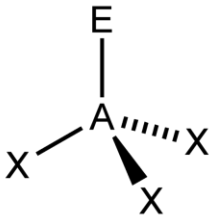




VSEPR

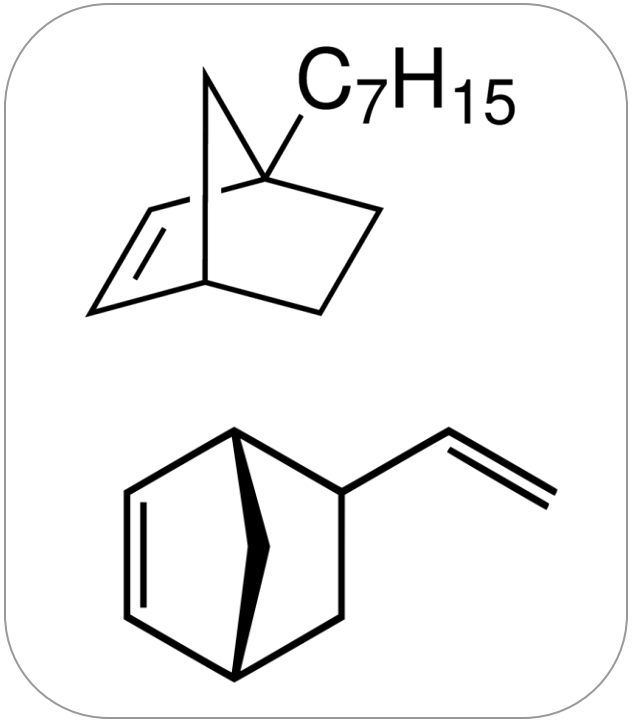


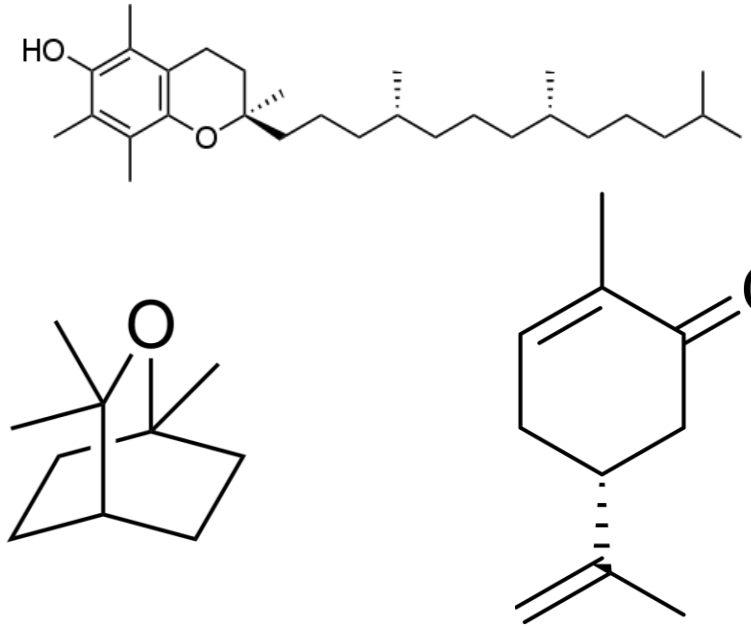




La struttura delle molecole

Visualizzare strutture molecolari diventa **più complesso per molecole organiche**



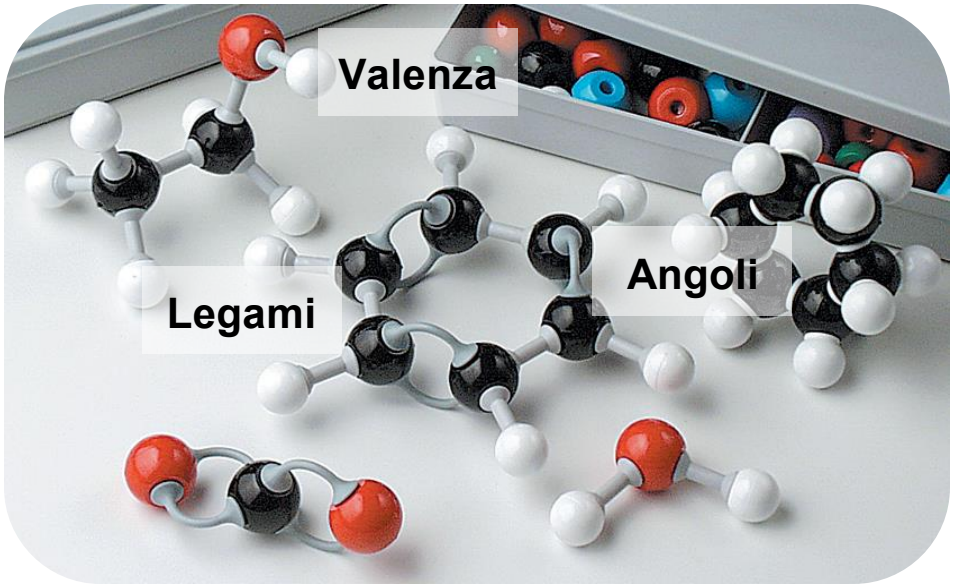


La stereochimica richiede un'idea 3D della molecola!

Cos'è la **modellistica molecolare?**

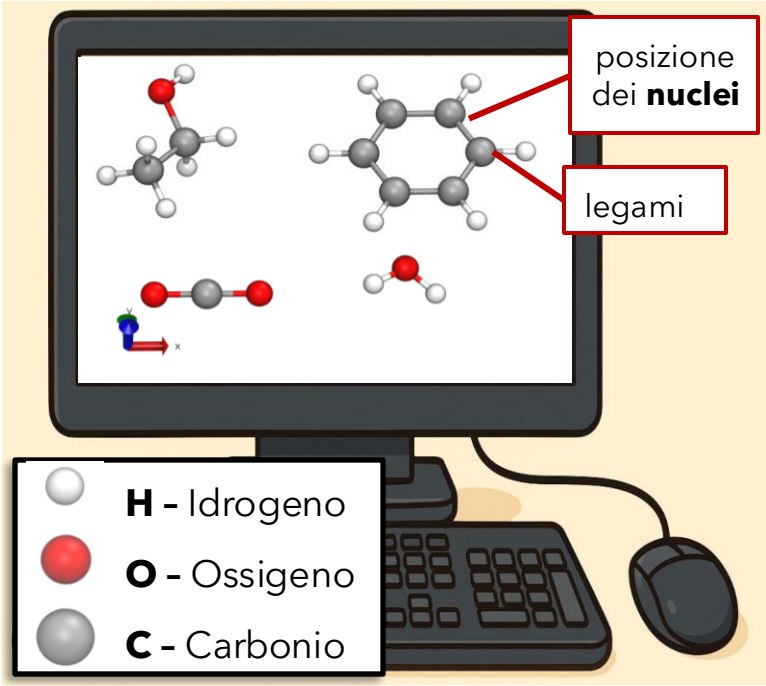
Modelli fisici

Semplici **regole** per costruire la geometria molecolare



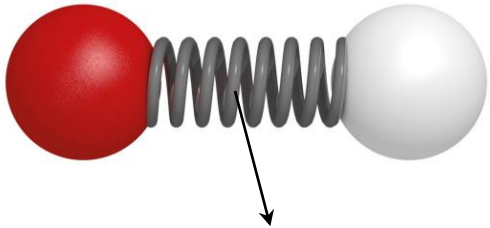
Modelli al computer

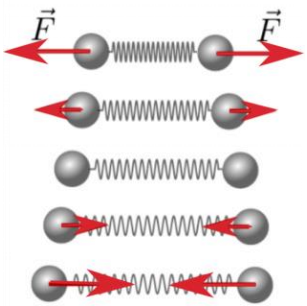
Usano **software, algoritmi, visualizzazione 3D**



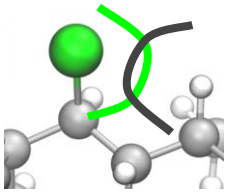
La fisica dei modelli molecolari: meccanica molecolare

Quali **regole** per i modelli al computer?

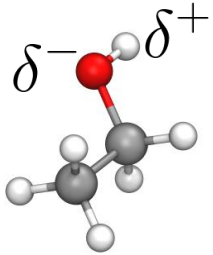




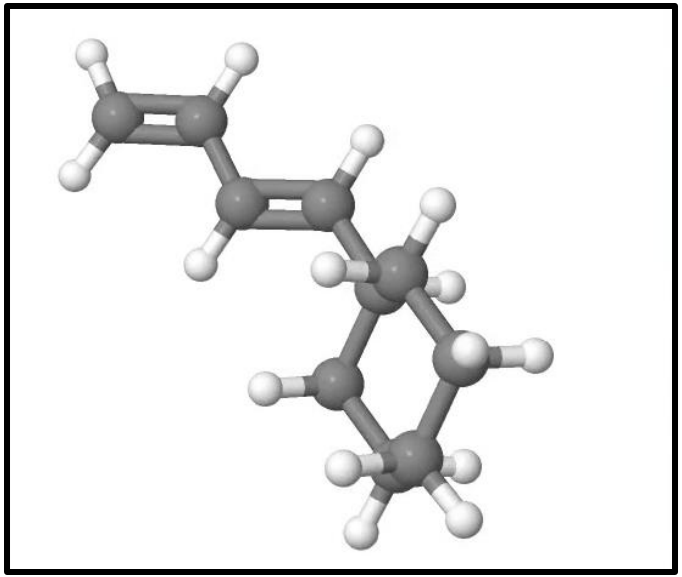
Descriviamo i legami come **molle**



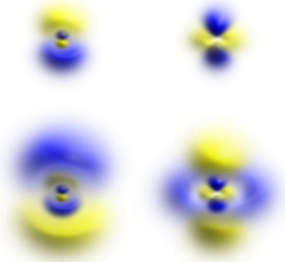
Consideriamo la **repulsione** e le **cariche** atomiche



Cerchiamo l'equilibrio



La fisica dei modelli molecolari: meccanica quantistica



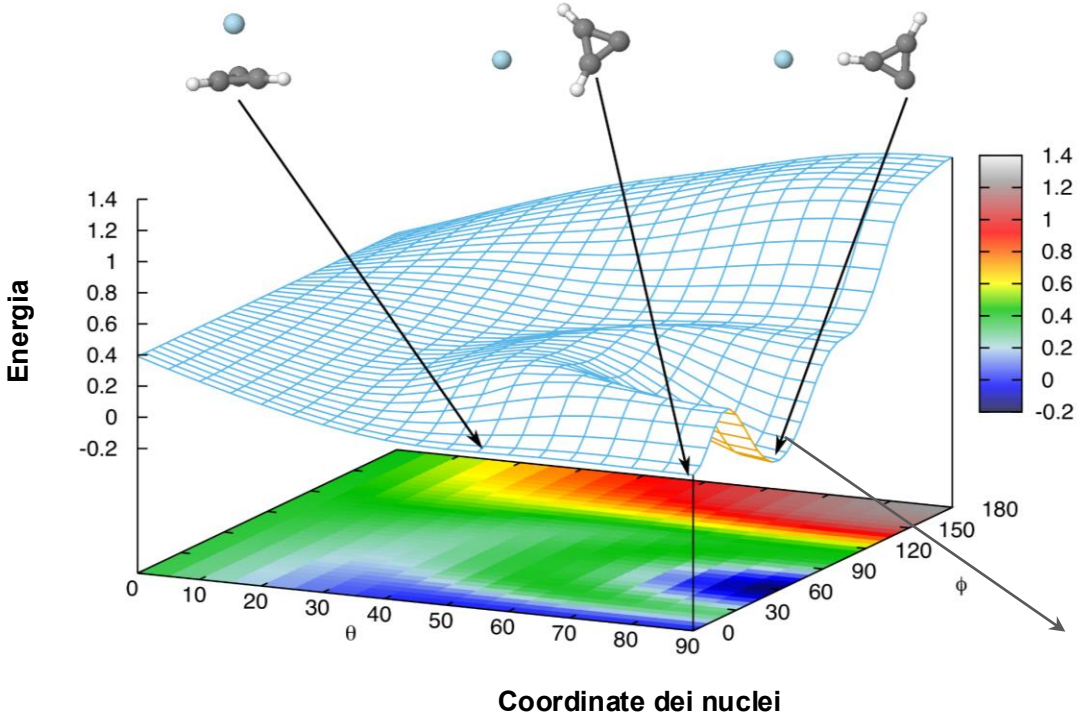
Meccanica Quantistica

$$\hat{\mathcal{H}} \left|\Psi\right\rangle = E \left|\Psi\right\rangle$$



- L'equazione di Schroedinger è **molto** complicata, non si può risolvere esattamente
- Approccio qualitativo → "Teorie di legame" (Valence Bond, MO)
- **Approccio quantitativo → Calcoli numerici**

Geometrie ed energie



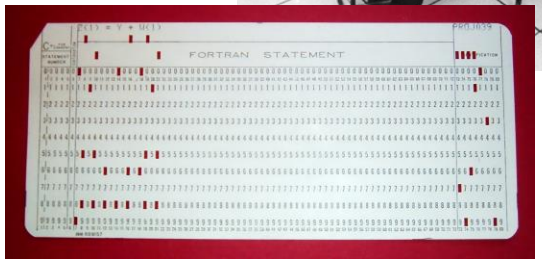
I minimi di energia sono **strutture stabili**

- Isomeri
- Conformeri
- Reagenti/prodotti

Computer e Supercomputer

C’era una volta …





… e c’è adesso!

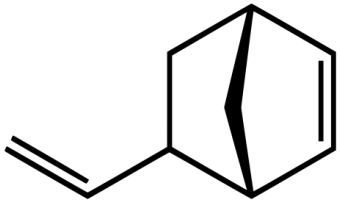


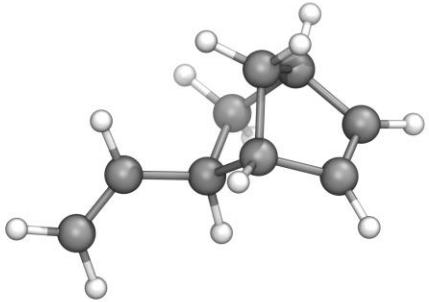


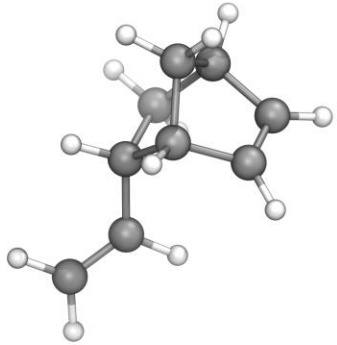


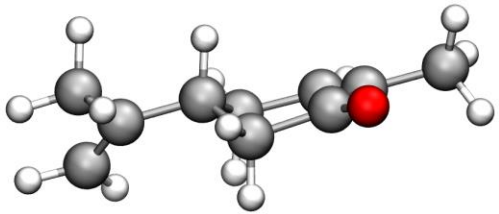


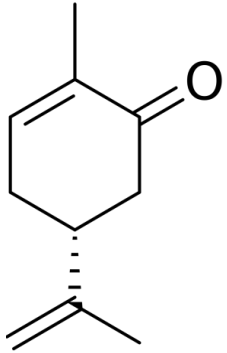
Alcune geometrie 3D





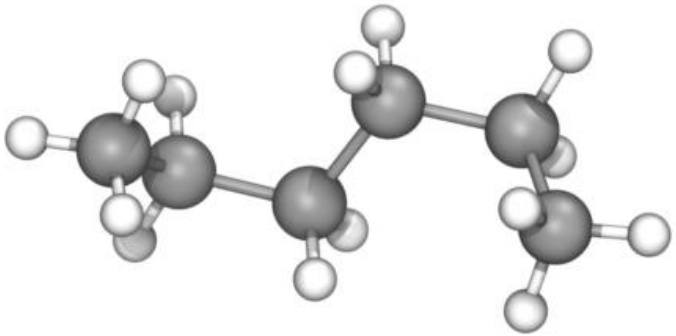




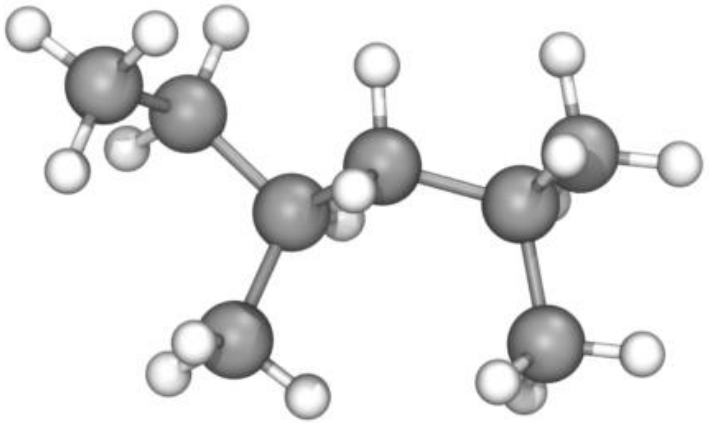


Dinamica e conformazioni

Quando una sola struttura non basta ...



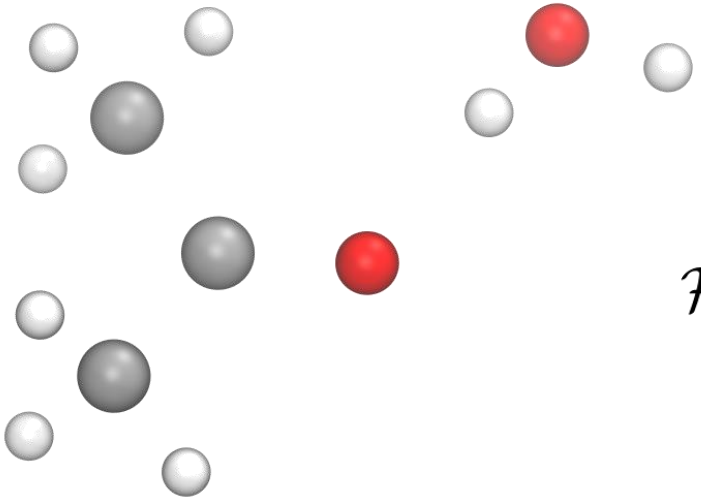
Alcano **lineare**



Alcano **ramificato**

Come funziona?

Posizioni Atomiche



QM

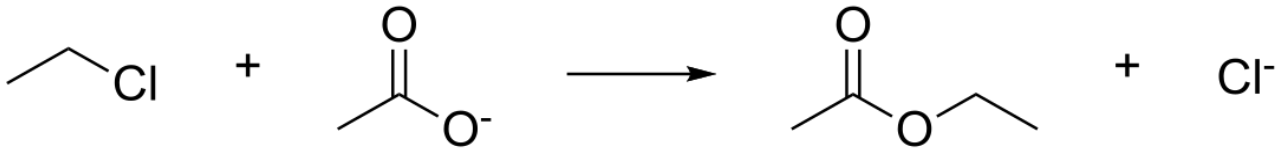
$$\hat{\mathcal{H}}\,|\Psi\rangle = E\,|\Psi\rangle$$

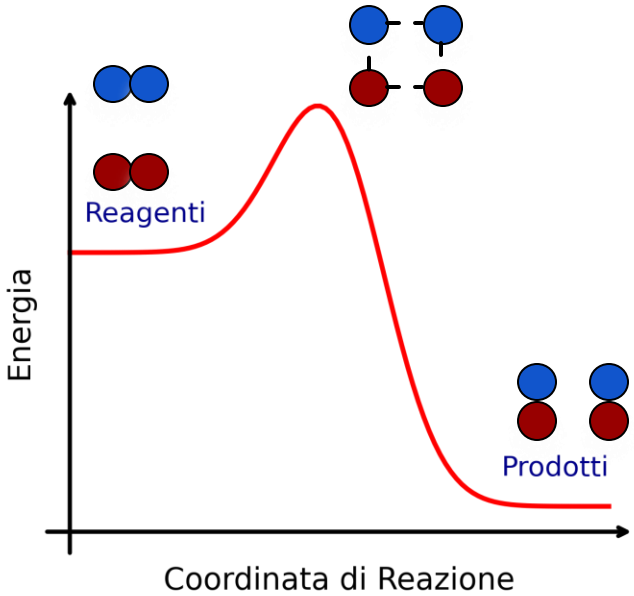
Energia

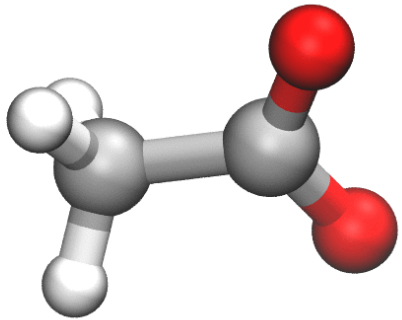
Forze

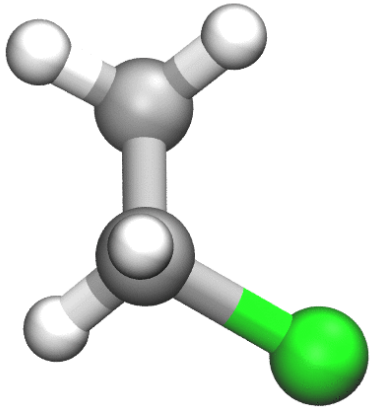
Simulazione

Reazioni chimiche

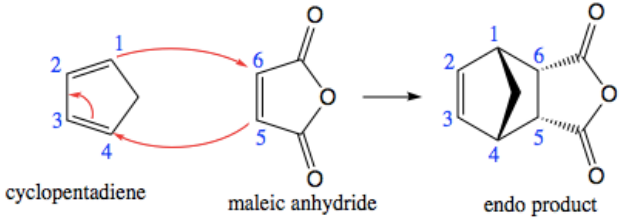




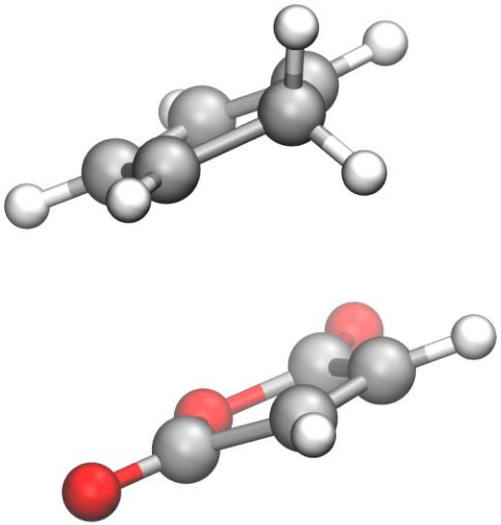


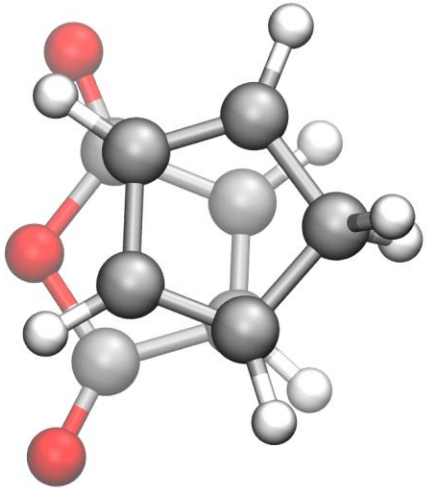


Una reazione più complicata

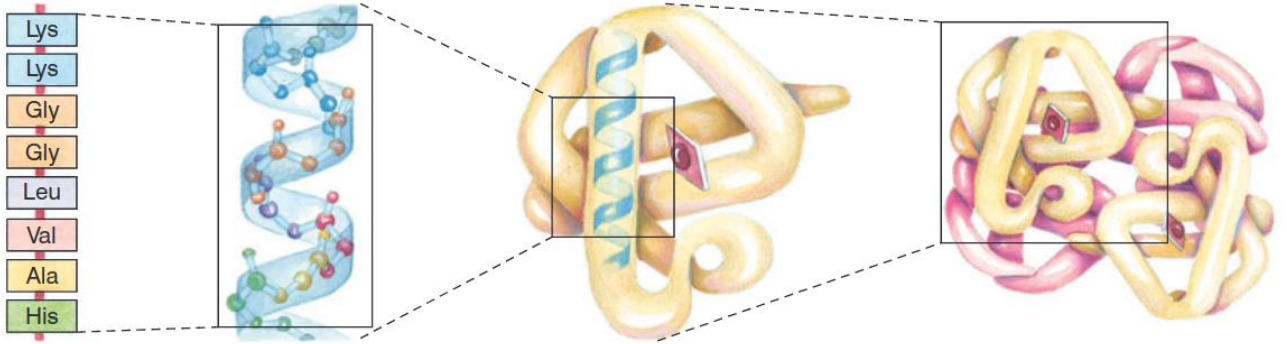


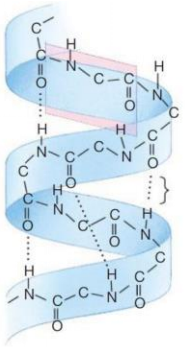
Diels-Alder

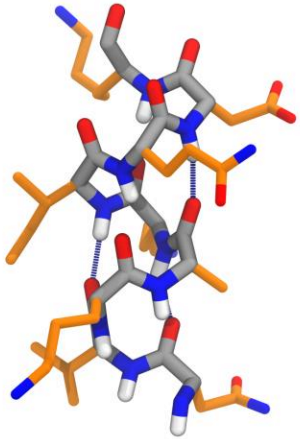


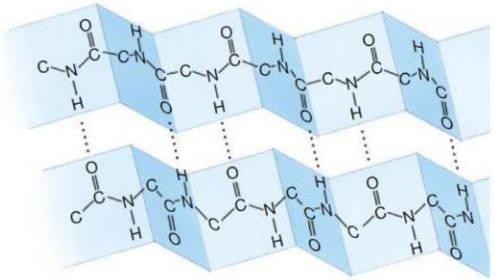


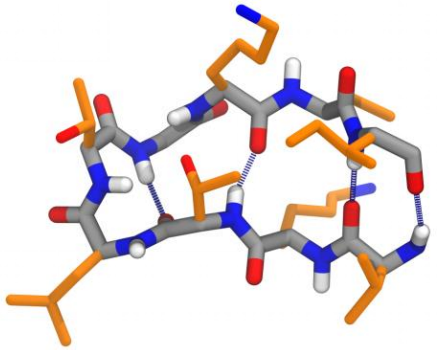
Dalle molecole alle macromolecole: Proteine



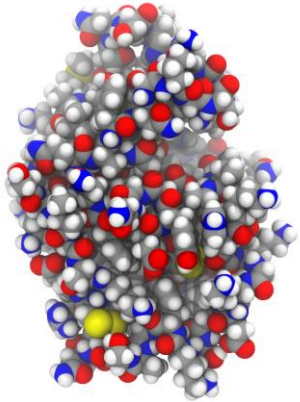


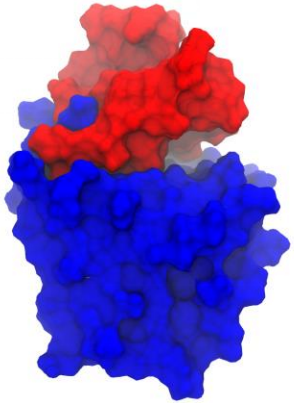


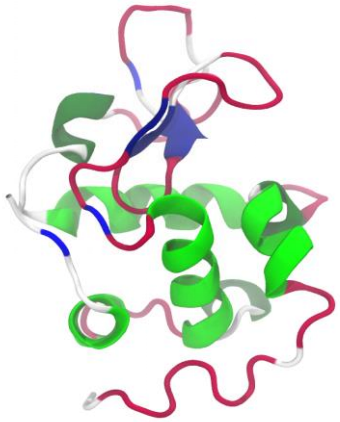


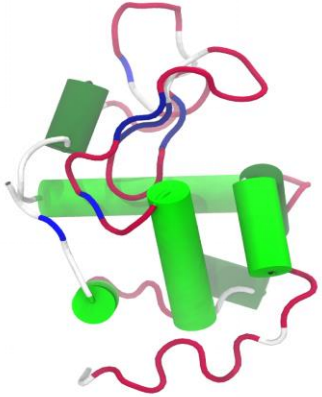


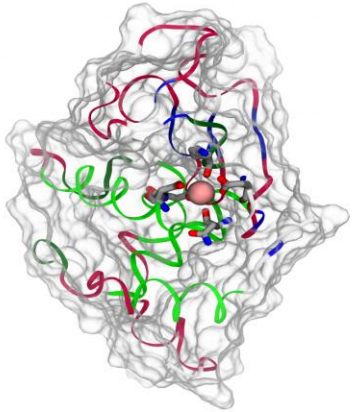
Rappresentazioni delle proteine











Limitazioni per la didattica

- **Software proprietario (€€€)**
- **Esperienza di calcolo molecolare**
- **Conoscenza software di visualizzazione**
- **Database difficili da trovare**

Codici Free/Open

Risorse online

App Browser/mobile

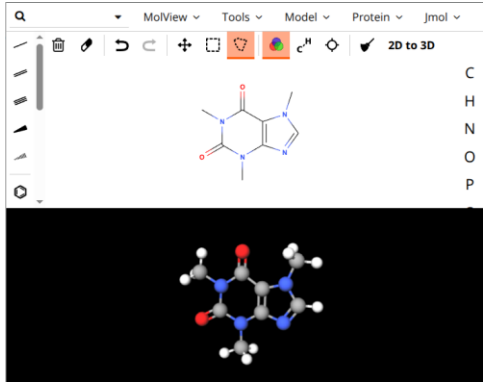
Open data

- **Obiettivi didattici?**
- **“Spazio” nei programmi didattici?**

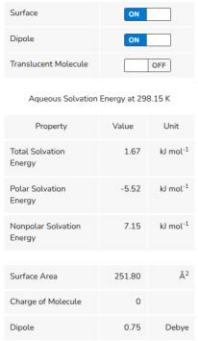
Risorse Online

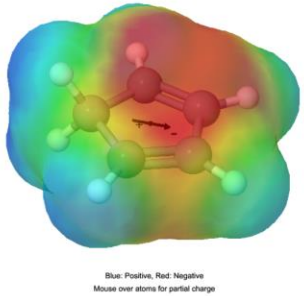
Software online di visualizzazione molecolare

molview.org



molcalc.org





web3dmol.net



Software **open**

Avogadro

Gabedit

ChimeraX

La chimica e l'intelligenza artificiale

Cosa ci viene in mente quando parliamo di AI?



ChatGPT

Generative Pre-trained Transformer

La chimica e l'intelligenza artificiale

Generative Pre-trained Transformer

Generative → AI generativa: **crea** qualcosa a partire da **esempi**

Pre-trained → Pre-allenata con database enormi di testo

Transformer → "Architettura" del modello (l'algoritmo)

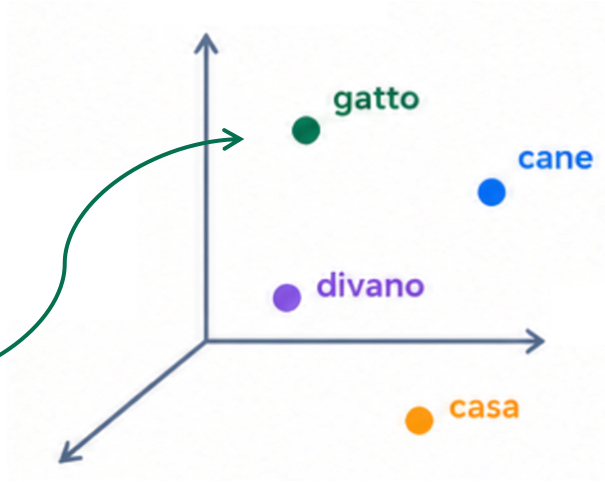
Come funziona un “Transformer”?

Consideriamo una semplice frase



Embedding

Ogni “token” (parola) viene associato alla posizione in uno spazio multi-dimensionale

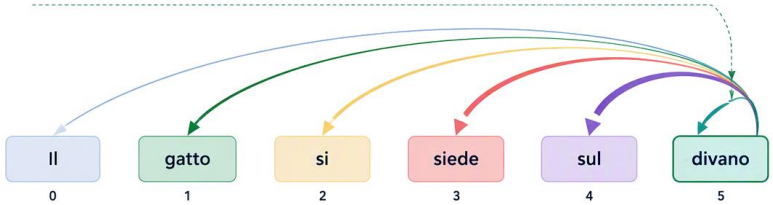


Come funziona un “Transformer”?

Il meccanismo di auto-attenzione

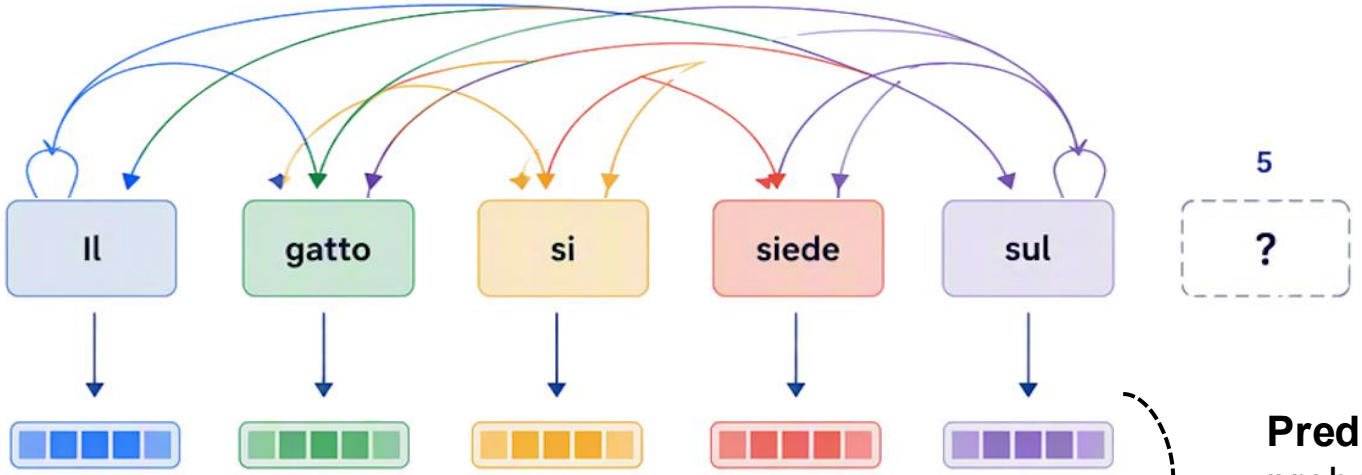
TOKEN CHE PRESTA ATTENZIONE (query)		0 Il	1 gatto	2 si	3 siede	4 sul	5 divano
0	Il	0.15	0.20	0.10	0.20	0.20	0.15
1	gatto	0.05	0.55	0.05	0.15	0.10	0.10
2	si	0.05	0.10	0.30	0.25	0.20	0.10
3	siede	0.05	0.10	0.05	0.60	0.10	0.10
4	sul	0.05	0.05	0.05	0.10	0.60	0.15
5	divano	0.02	0.05	0.03	0.10	0.15	0.65

L’attenzione **cattura le relazioni** tra parole diverse



Genera una rappresentazione della parola **nel contesto**

Come funziona un “Transformer”?



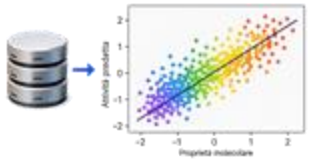
Predizione della probabilità per ogni parola nel dizionario

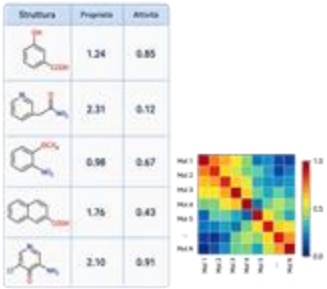


La prossima parola viene generata secondo la probabilità

Cosa possiamo fare *noi chimici* con l'intelligenza artificiale?

Analisi dati

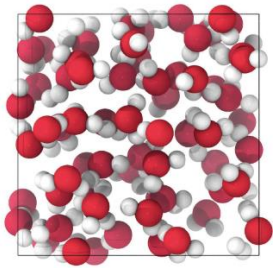




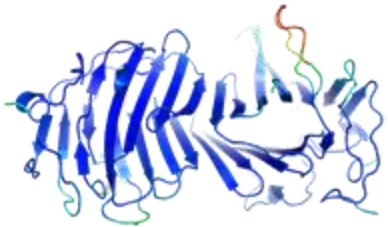
Ottimizzazione di reazioni



Simulazioni molecolari



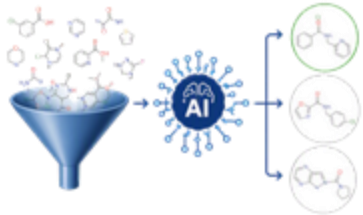
Predizione di strutture



Automazione del laboratorio



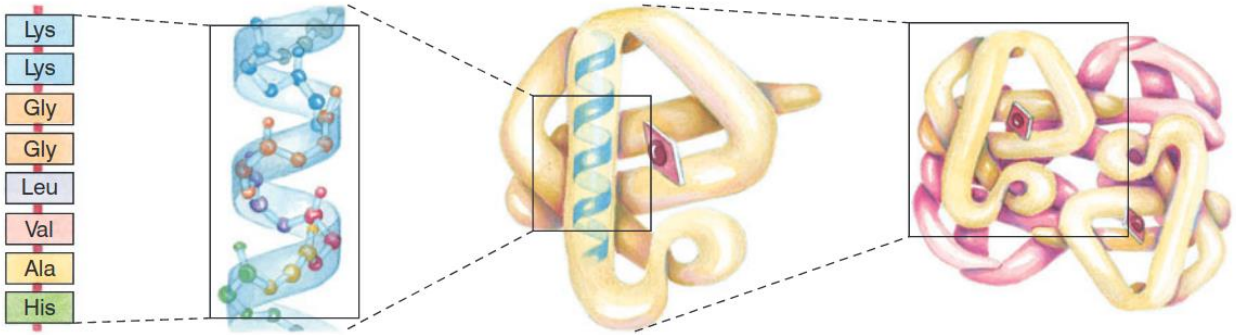
Sviluppo di farmaci



Predizione della struttura di proteine



Predizione della struttura di proteine



Possiamo predire dalla **sequenza primaria** la **struttura terziaria/quaternaria** di una proteina?

M K T A Y I A K Q Q Q I S F V K

La sequenza primaria è come un testo, dove le parole sono gli amminoacidi

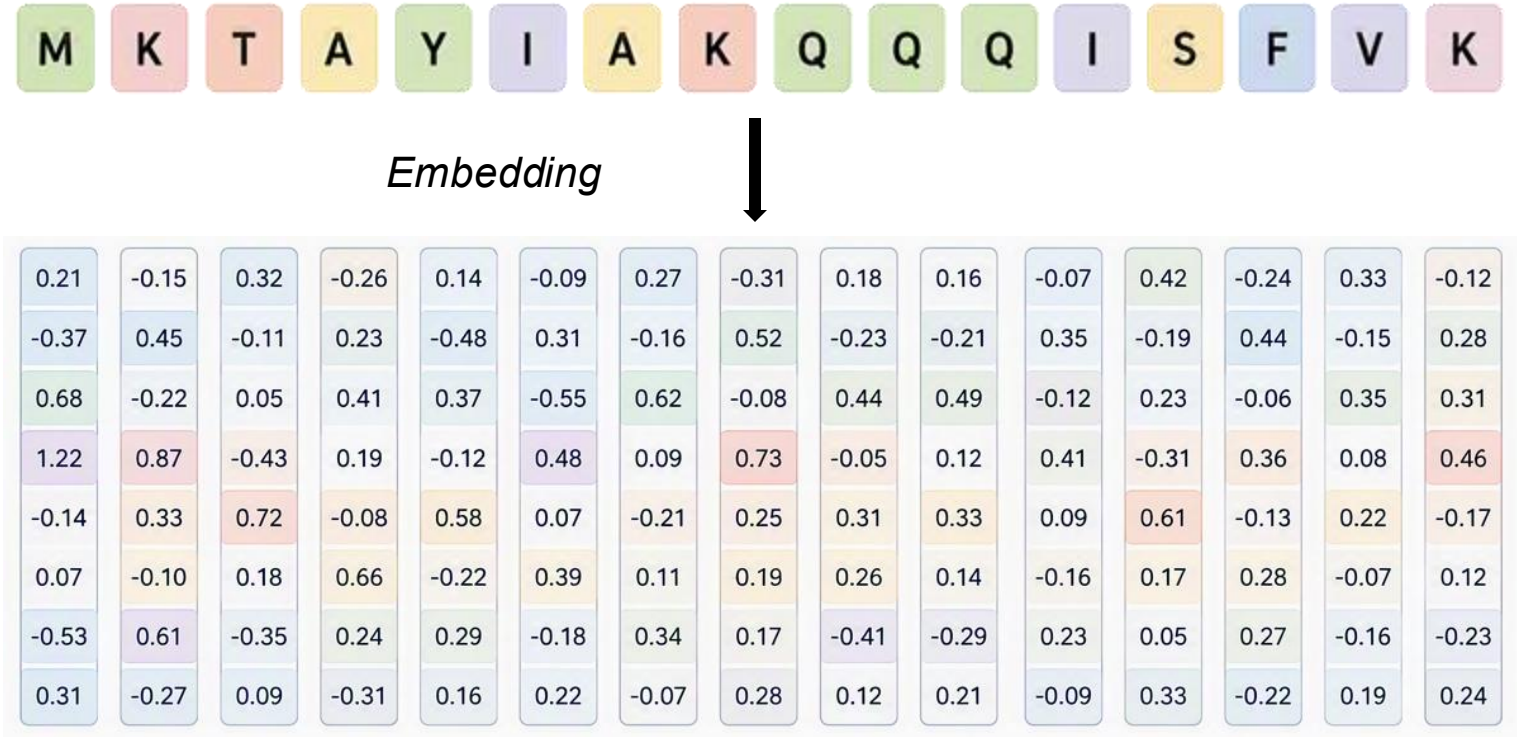
Predizione della struttura di proteine: AlphaFold



Allineamento multi-sequenza

Allinea sequenze simili per avere informazioni sul “contesto evolutivo” di ogni amminoacido

Predizione della struttura di proteine: AlphaFold



Predizione della struttura di proteine: AlphaFold



Allineamento

Meccanismo di attenzione

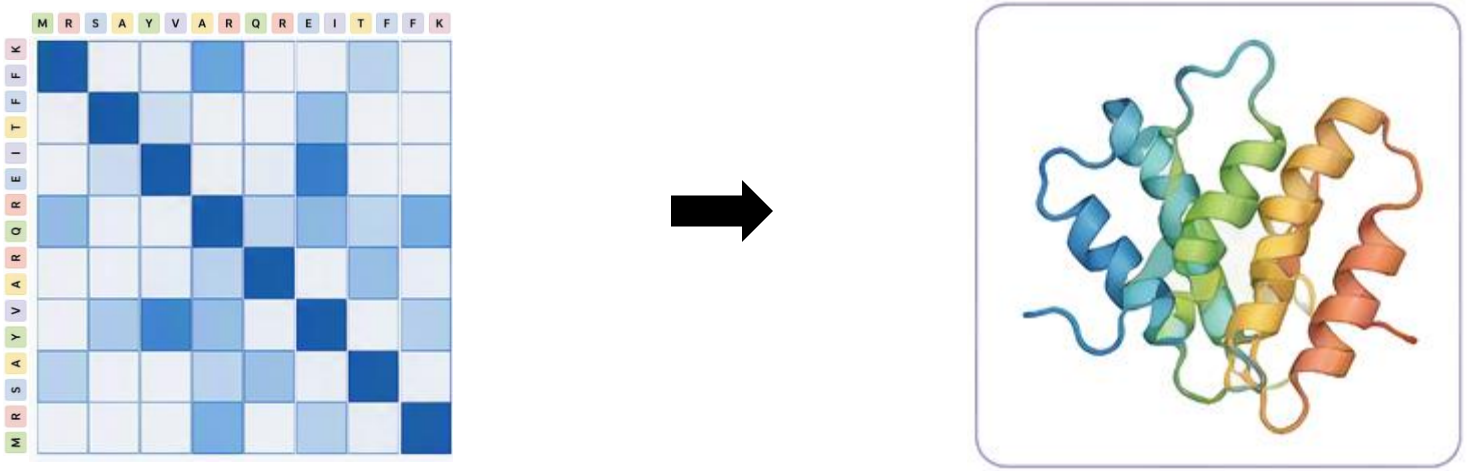


Distanze nella struttura 3D

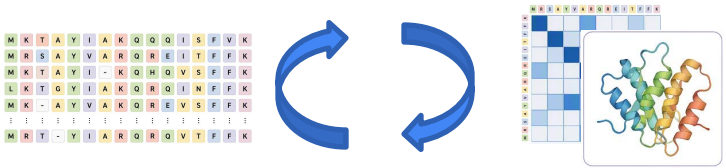
Le “relazioni” tra amminoacidi vengono trasformate in informazione strutturale

Predizione della struttura di proteine: AlphaFold

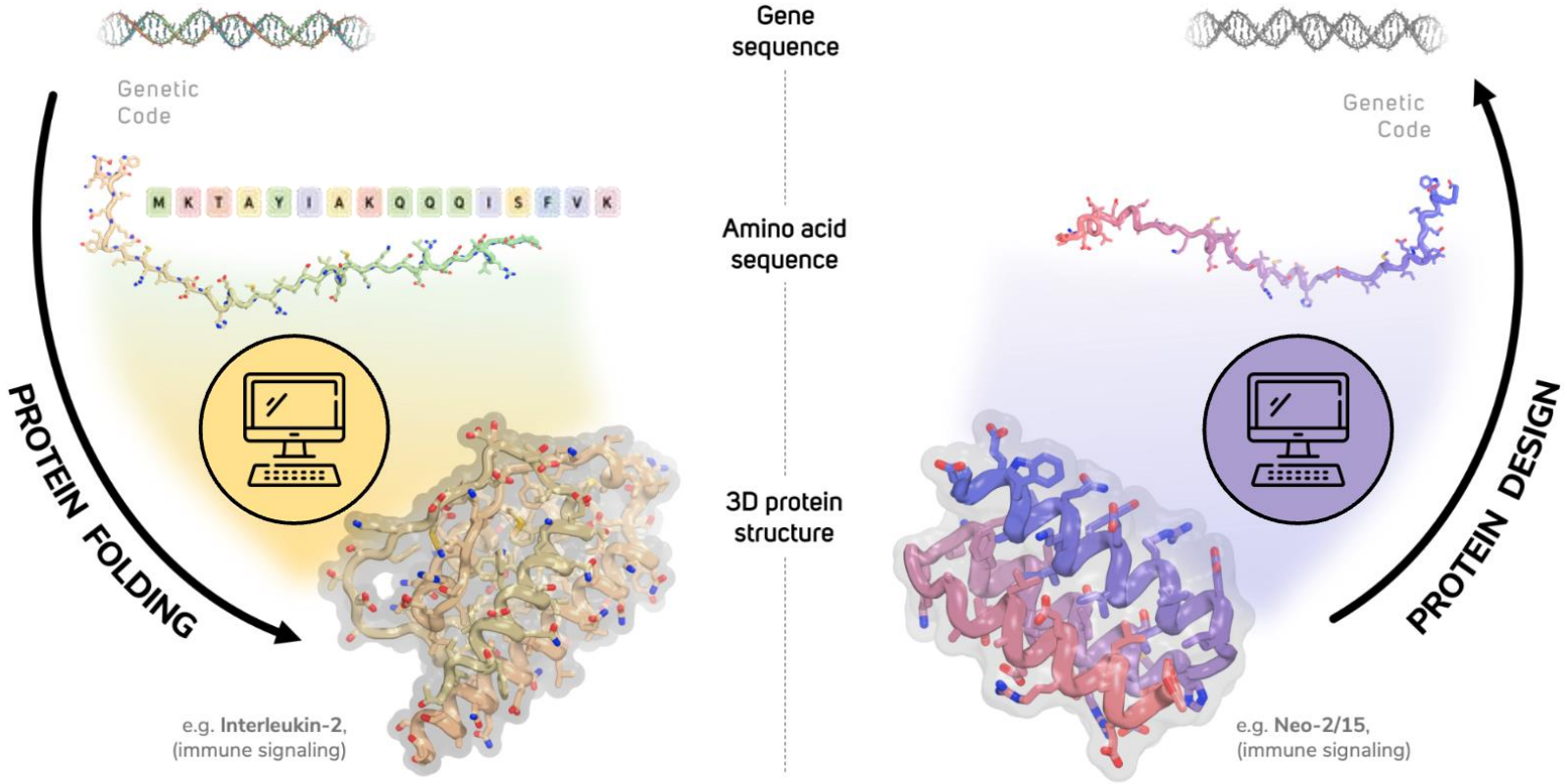
Le informazioni sulle distanze vengono usate per predire la struttura



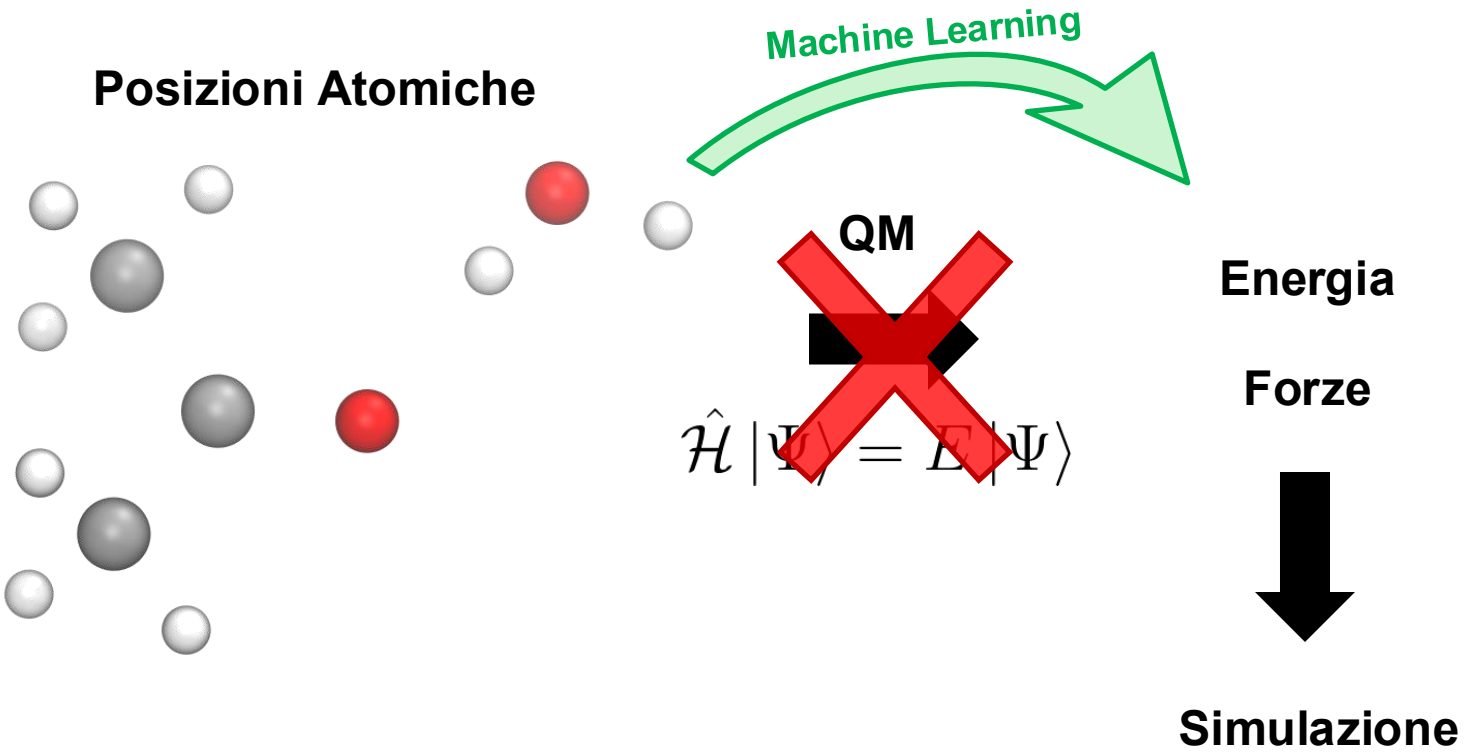
Il processo viene iterato per rafforzare la relazione sequenza-struttura



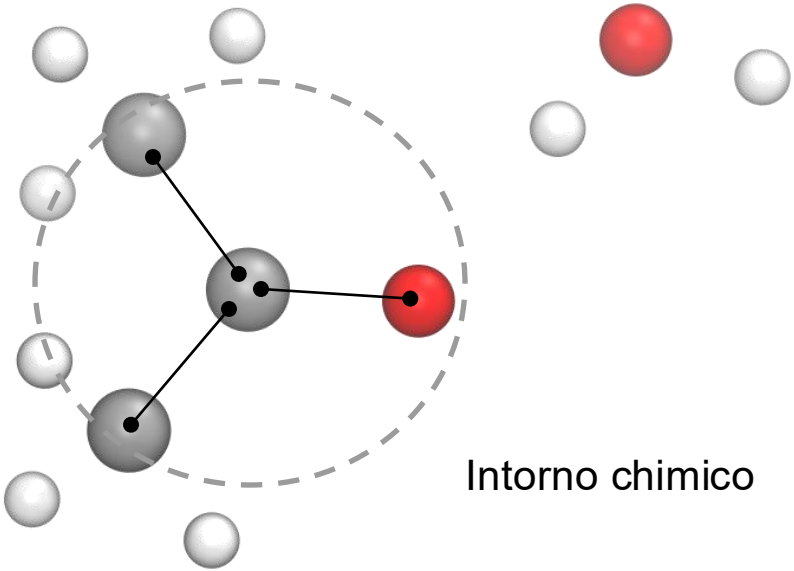
Dalla predizione al disegno di proteine



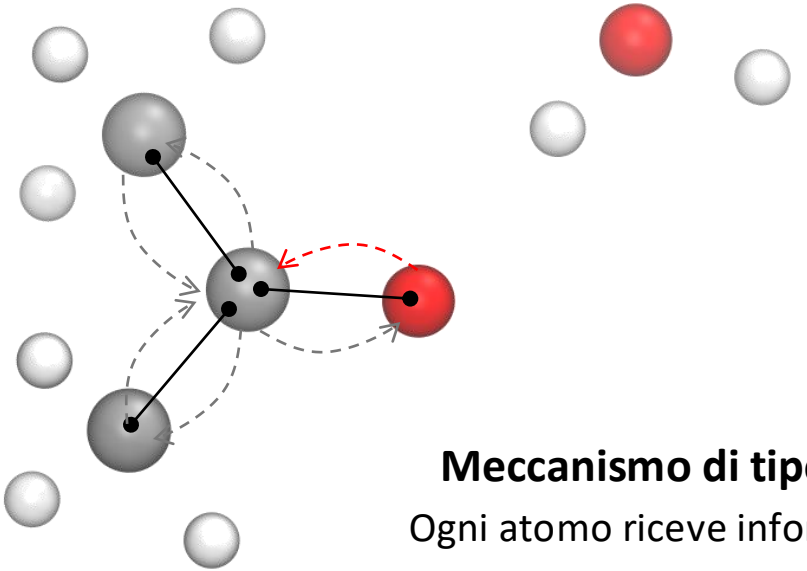
Un’AI per le simulazioni molecolari?



Un’AI per le simulazioni molecolari?



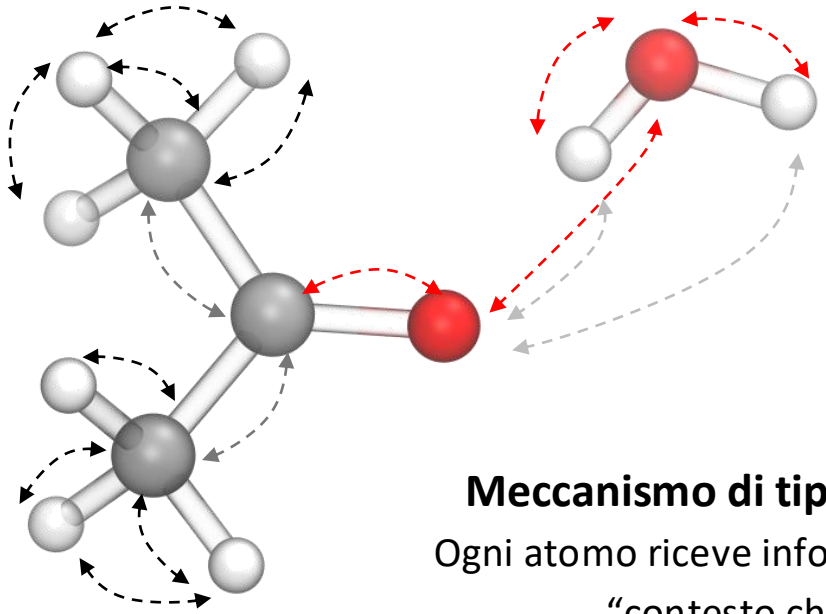
Un’AI per le simulazioni molecolari?



Meccanismo di tipo attenzione

Ogni atomo riceve informazioni sul suo “contesto chimico”

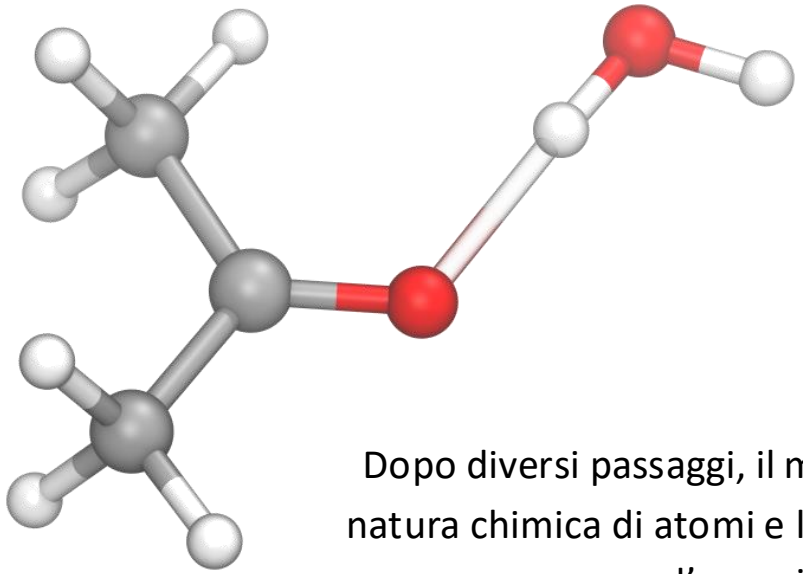
Un’AI per le simulazioni molecolari?



Meccanismo di tipo attenzione

Ogni atomo riceve informazioni sul suo “contesto chimico”

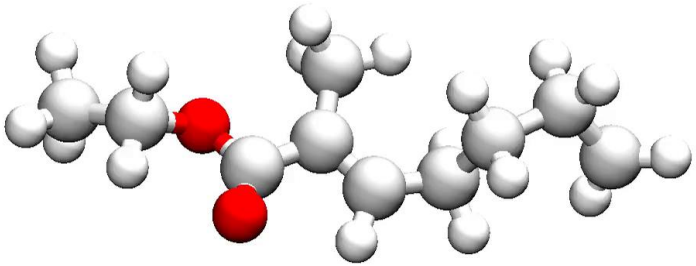
Un’AI per le simulazioni molecolari?



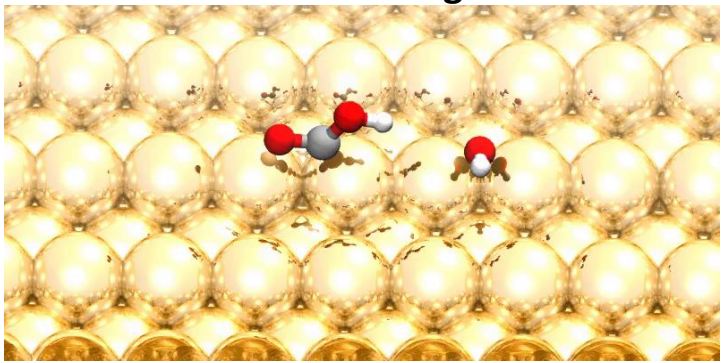
Dopo diversi passaggi, il modello “capisce” la natura chimica di atomi e legami e può predire l’energia

Simulazioni accelerate da modelli AI

Reazioni organiche



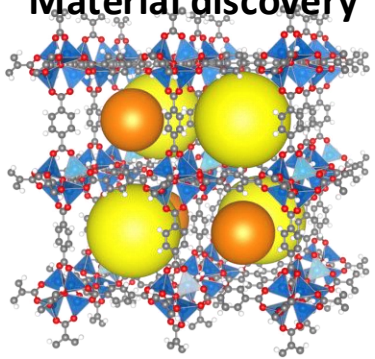
Catalisi eterogenea



Credits: Dr. Miguel de la Puente, MoLEcoLab, UNIPI

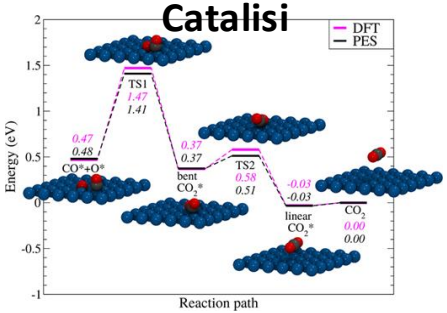
Cosa possono fare le simulazioni con modelli AI?

Material discovery

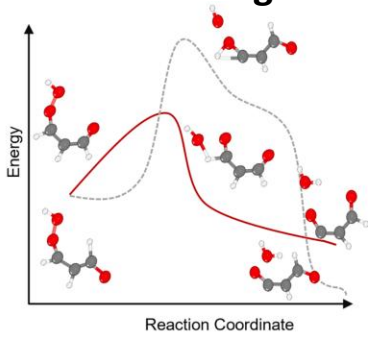


Nat. Materials 20, 750–761

Catalisi

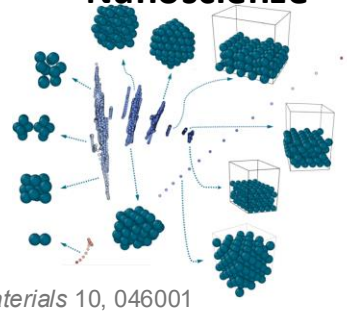


Reattività organica



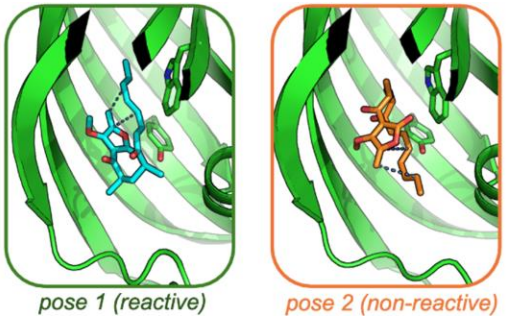
Chem Cat 5, 101458

Nanoscienze



Phys. Rev. Materials 10, 046001

Enzimi



Conseguenze per la didattica?

Possiamo usare questi modelli per dimostrazioni in classe?

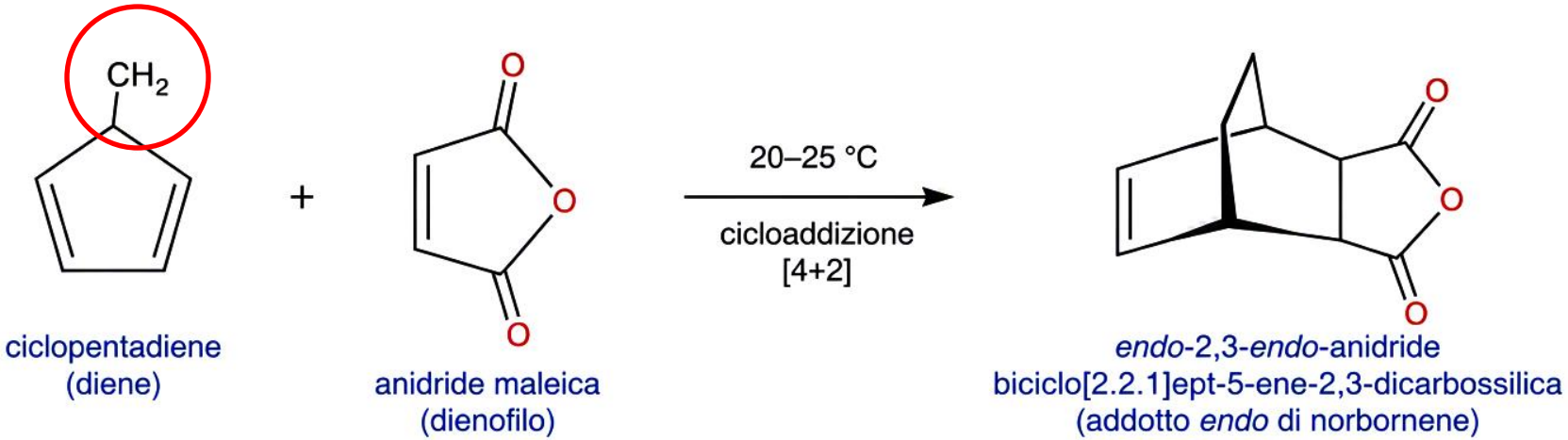
... quasi! Richiedono comunque di integrare diversi software...

- Costruzione di molecole
- Simulazione
- Visualizzazione

Nel futuro, con AI ⟶ soluzioni integrate / automatizzate, anche del tutto online

Rischi dell'AI in didattica

Illustra la reazione di Diels-Alder tra ciclopentadiene e anidride maleica



ciclopentadiene
(diene)

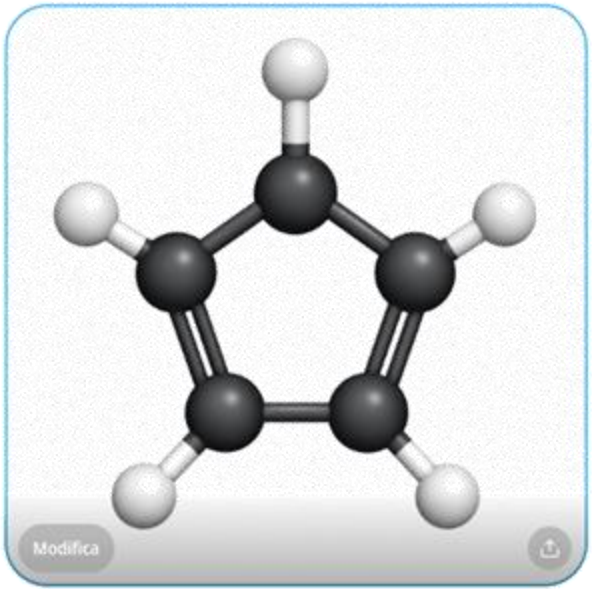
+

anidride maleica
(dienofilo)

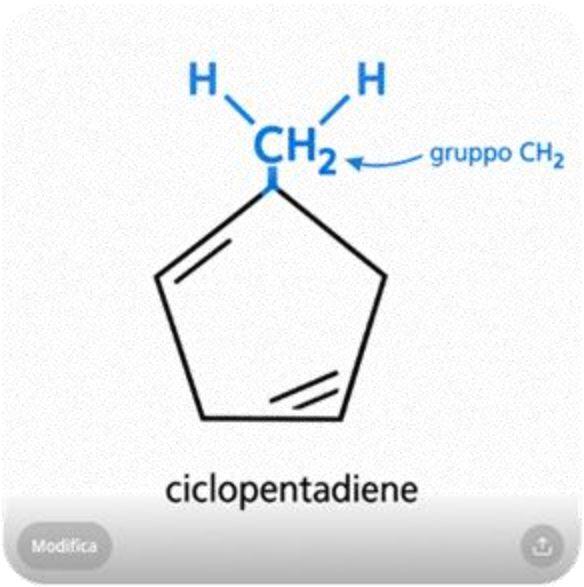
endo-2,3-*endo*-anidride
biciclo[2.2.1]ept-5-ene-2,3-dicarbossilica
(addotto *endo* di norbornene)

Rischi dell'AI in didattica

Disegna una molecola di ciclopentadiene



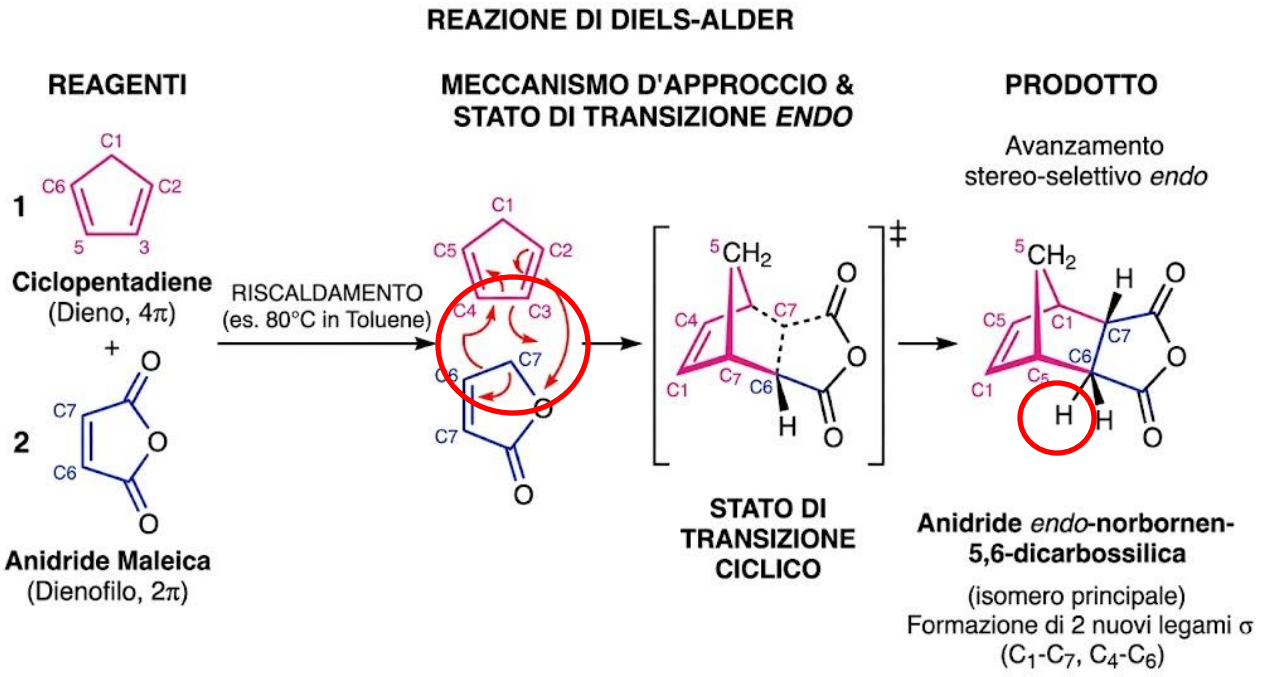
Disegna una molecola di ciclopentadiene, evidenziando il gruppo CH2



Rischi dell'AI in didattica

Illustra la reazione di Diels-Alder tra ciclopentadiene e anidride maleica

Gemini



Rischi dell'AI in didattica

- Non sempre le risposte sono **corrette**...
- ... anche se a prima vista sembra tutto giusto!
- L'AI è **molto convincente**: come capire se le risposte sono giuste?
- Non sempre le risposte sono **consistenti** o **adeguate**
- È difficile ottenere risposta consistente con il **livello didattico**

Conclusioni

- Modelli 3D al computer aiutano a comprendere meglio struttura e reattività di molecole, proteine, materiali, ...
- I modelli AI accelerano questi calcoli permettendo di esplorare la diversità chimica e biochimica
- Risorse online per calcolo/visualizzazione
- Prossimamente?
  - Database ad uso didattico
  - Calcoli AI-powered per scopi didattici

Grazie per l’attenzione!

