

CHIMICA COMPUTAZIONALE: PROGETTAZIONE DI MOLECOLE AL CALCOLATORE

La chimica computazionale ha compiuto progressi enormi negli ultimi due decenni grazie allo sviluppo di metodi avanzati per la determinazione della struttura elettronica di solidi, liquidi e molecole, di algoritmi di calcolo parallelo efficienti nonché all'aumentata potenza di elaborazione. Oggi, la chimica computazionale si affianca a pieno titolo ad altre discipline per l'interpretazione e la comprensione di dati sperimentali e per la progettazione di nuovi sistemi con proprietà ben definite.

1. INTRODUZIONE

La chimica è sicuramente una scienza antichissima e il suo sviluppo è intimamente connesso a quello del genere umano, anche se per molto tempo la possibilità di controllare la trasformazione di sostanze in altre sostanze non è stata percepita in modo chiaro dall'uomo.

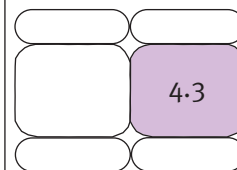
Per molti secoli la chimica, o meglio l'alchimia, è stata vista come una sorta di magia in grado di trasformare gli elementi gli uni negli altri, come dimostrato dai tanti cercatori della pietra filosofale che si sono cimentati in questa improbabile impresa. Comunque la si veda, non c'è alcun dubbio che da quando l'*homo sapiens* si è affermato come specie dominante, tutte le attività connesse con la chimica (dalla cottura dei cibi alla fusione dei metalli e delle leghe, dalla messa a punto di medicinali e tisane, alla tintura dei tessuti ecc.) sono state di natura esclusivamente empirica, basate su tentativi ripetuti più volte, che oggi si definiscono sperimentali. Al punto che nel suo trattato "Corso di Filosofia Positiva" del 1830, il filosofo francese Augu-

ste Comte scriveva che "ogni tentativo di utilizzare metodi matematici nello studio di problemi chimici deve essere considerato profondamente irrazionale e contrario allo spirito della chimica. Se l'analisi matematica dovesse un giorno assumere un ruolo prominente nella chimica – una aberrazione che fortunatamente è quasi impossibile – ciò produrrebbe una rapida e generale degenerazione di questa scienza". Questo la dice lunga su come era visto l'uso di strumenti matematici per lo studio di problemi chimici in un secolo, l'Ottocento, che pure ha prodotto profonde e importantissime rivoluzioni in questo campo, basti pensare alle leggi di Faraday sull'elettrolisi, alla formulazione della termodinamica, al sistema periodico di Mendeléeu o ai composti organici otticamente attivi di Pasteur ecc..

In effetti, molti di questi sviluppi, su cui si basano le conoscenze attuali, sono stati raggiunti senza sapere che esistessero gli atomi e le molecole né come queste fossero fatte in realtà. Questa, infatti, è una conquista relativamente recente nella storia della scienza,



Gianfranco Pacchioni
Cristiana Di Valentin



che risale ai primi tre decenni del 1900, quelli che George Gamow definì con una felice espressione “i trent’anni che sconvolsero la fisica”. Nei tre decenni che vanno dal 1900 al 1930, infatti, vennero poste le basi matematiche della teoria dei *quanti* su cui si basa l’attuale comprensione dei fenomeni atomici e molecolari, ossia della chimica e della fisica della materia. Al punto che il grande fisico inglese Paul A. M. Dirac nel 1929, ossia un secolo dopo la frase di Compton, riferendosi alle leggi della meccanica quantistica recentemente scoperte, poteva dichiarare che “*le leggi fisiche necessarie per una teoria matematica di tutta la chimica e di parte della fisica sono note completamente, e l’unica difficoltà è che l’applicazione esatta di queste leggi porta a equazioni troppo complicate per poter essere risolte*”. È interessante notare che in questa famosa affermazione Dirac parlava consapevolmente del fatto che le leggi della teoria dei quanti spiegano solo parte della fisica, in quanto esistono fenomeni che possono essere interpretati solo con la teoria della relatività. Dirac riteneva, erroneamente, che i fenomeni relativistici non avessero alcun effetto sui processi chimici, da cui l’affermazione per cui la meccanica quantistica spiegherebbe “tutta la chimica” e solo parte della fisica. In realtà, oggi si sa che anche molti fenomeni chimici sono strettamente legati a effetti relativistici. Un aspetto interessante della affermazione di Dirac resta, comunque, che la teoria matematica che è alla base delle trasformazioni chimiche, la teoria dei quanti, è sì completamente formulata, ma anche talmente complessa che le equazioni che andrebbero risolte per arrivare a dei risultati pratici sono di fatto troppo complicate. Se nel corso del 1900 non ci fosse stata la scoperta del *transistor*¹ e il successivo sviluppo della elaborazione elettronica, la frase di Dirac avrebbe costituito una sorta di suggello definitivo allo sviluppo di una teoria formale per la spiegazione dei fenomeni naturali atomici e molecolari elegante, ma di scarsa rilevanza pratica.

Le cose non sono andate così e i circa settant’anni che sono seguiti alla formulazione

della meccanica quantistica hanno visto da una parte, lo sviluppo vorticoso e continuo della potenza di elaborazione dei calcolatori elettronici, dall’altra la messa a punto di metodi e algoritmi sempre più avanzati ed efficienti volti alla soluzione delle equazioni di cui parlava Dirac. La combinazione di questi due fattori ha portato alla nascita di una nuova importante disciplina, la chimica computazionale. La definitiva consacrazione di questo campo è avvenuta verso la fine del secolo appena trascorso con l’assegnazione del Premio Nobel per la chimica 1998 a due “padri” di questa disciplina, John Pople e Walter Kohn. Da notare che Pople di formazione è un matematico e Kohn un fisico, a dimostrare una volta di più la natura interdisciplinare della scienza moderna e il ruolo fondamentale che la matematica e la fisica hanno nei fondamenti della chimica.

Cos’è allora la chimica computazionale o chimica al *computer*? È semplicemente una branca di un settore più ampio in rapida evoluzione che va sotto il nome di simulazioni al calcolatore o *scientific computing*. Questa nuova disciplina sta affiancando sempre di più l’approccio tradizionale alla soluzione di problemi scientifici e tecnologici basati sulla sperimentazione diretta (senza comunque sostituirla). La simulazione al calcolatore e il calcolo scientifico non vanno confusi con le aree più generali della chimica o della fisica teoriche. Gli esperimenti producono nuovi fatti e nuove scoperte e dischiudono all’uomo i segreti della natura. Il ruolo della teoria è quello di fornire un quadro di riferimento generale di spiegazione dei fenomeni osservati tramite un insieme di leggi matematiche. Il celebre Lord Kelvin diceva nel lontano 1900 che “*quando si può misurare ciò di cui si parla ed esprimerlo in numeri se ne sa qualcosa; ma quando non si può esprimerlo in formule o numeri allora la nostra conoscenza è scarna e insoddisfacente*”. Il calcolo scientifico è qualcosa a metà tra teoria ed esperimento: è basato su teorie e formalismi ben definiti e sviluppati ma è usato per produrre nuovi fatti e risultati, in un modo sempre più simile a quello in cui vengono svolti gli esperimenti. L’uso combinato di *software* e algoritmi avanzati e di potenza di calcolo elevata permette, oggi, di simulare un esperimento al computer

¹ Avvenuta nel 1947.

prima di effettuarlo, con costi più bassi e risposte più rapide. Ovviamente, la capacità predittiva di questi calcoli è molto legata agli algoritmi utilizzati e alle approssimazioni introdotte e, quindi, al livello di affidabilità della simulazione. Dopo alcuni decenni in cui i progressi in questa direzione sono stati piuttosto lenti, si è assistito nell'ultima decade del secolo scorso a una definitiva affermazione della chimica computazionale, al punto che oggi una parte consistente dei lavori originali che appaiono sulla letteratura scientifica internazionale si basano su questo metodo. Questo articolo intende fornire alcuni elementi per meglio comprendere come si è arrivati a ciò e quali sono i problemi che si pensa di poter risolvere in questo modo.

2. TEMPO E TEMPERATURA

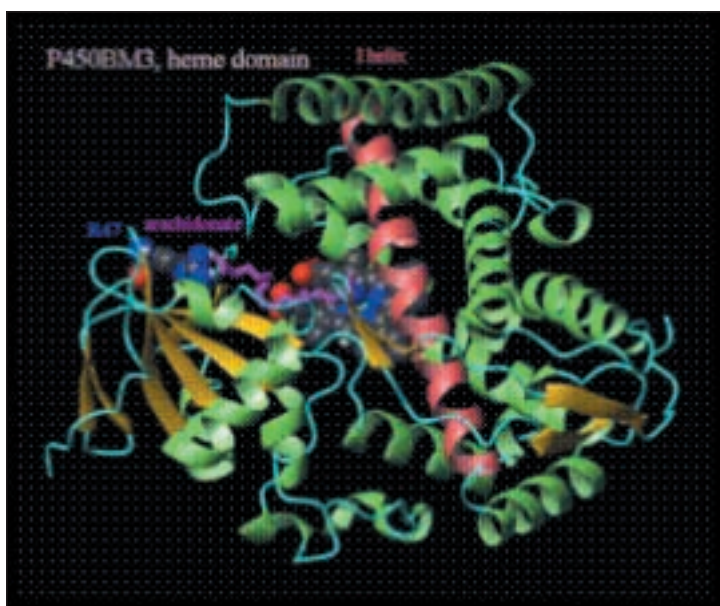
Simulare i processi chimici significa, di fatto, riprodurre in un computer fenomeni estremamente complessi che vanno dalla combinazione di due o più atomi isolati a formare una molecola gassosa, sino a trasformazioni in molti stadi che coinvolgono molecole di grandi dimensioni in soluzioni a molti componenti come i processi biochimici. In tutti i casi, il livello di complessità cresce via via che si introducono nuove variabili da cui dipendono le trasformazioni chimiche. Tra queste, due sono particolarmente importanti, il tempo e la temperatura. La totalità dei processi chimici avviene su scale temporali che vanno da frazioni infinitesimali di secondo sino a tempi geologici; si è di fronte a dei processi dinamici, in cui l'evoluzione temporale gioca un ruolo fondamentale. Più sono lunghi i tempi del processo, più sono lunghe le simulazioni e i relativi calcoli. L'altra variabile fondamentale dei processi dinamici è la temperatura. La temperatura misura l'energia termica di un sistema, a sua volta legata al moto molecolare. La materia a temperature bassissime tende a bloccare il moto dei nuclei atomici e il limite fisico dello zero assoluto ($-273,16\text{ }^{\circ}\text{C}$) corrisponde, in pratica, ad assenza di movimento nucleare. Per considerare questi fattori è necessario fare delle simulazioni dinamiche, in cui si tiene conto della temperatura e l'evoluzione del sistema viene seguita per tempi sufficientemente lunghi. È questo l'ambito della

dinamica molecolare, di fatto l'obiettivo finale di tutte le simulazioni chimiche.

Per arrivare a una buona descrizione dei fenomeni dinamici è però necessario avere come base una conoscenza accurata dei campi di forze che governano le interazioni tra atomi e molecole, ossia dei processi elementari in cui legami chimici vengono rotti per formarne dei nuovi. È questa la vera base della teoria quantistica della chimica, ossia la capacità di prevedere e calcolare a priori la stabilità di nuovi composti, il costo energetico necessario per la loro preparazione, la loro reattività. Per fare ciò non è necessario studiare il fenomeno su scale temporali o introdurre effetti di temperatura. È, invece, necessario risolvere le equazioni fondamentali della quanto-meccanica (quelle di cui parlava Dirac) possibilmente con il minor numero possibile di approssimazioni, senza cioè introdurre nessun "ingrediente" dedotto dai dati sperimentali nei calcoli e partendo solo da costanti universali come massa e carica di nuclei ed elettroni.

Esiste, quindi, una gerarchia di livelli di simulazione, a complessità crescente, dove si passa da un esame delle proprietà fondamentali di una singola molecola al suo modo di interagire con altre molecole anche in situazioni complesse. Lo studio delle proprietà fondamentali senza uso di parametri empirici e di semplificazioni prende il nome di metodi *ab initio* o anche da primi principi, proprio per indicare l'assenza di assunzioni di sorta. Questi calcoli considerano le proprietà elettroniche e strutturali delle molecole alla temperatura dello zero assoluto e senza considerarne l'evoluzione nel tempo. Il fatto però di risolvere le complesse equazioni che descrivono la struttura elettronica in modo "esatto" rende questi metodi assai pesanti dal punto di vista computazionale, tanto che oggi essi sono ancora limitati a sistemi che contengono al massimo alcune centinaia di atomi (sino a pochi anni fa si parlava al massimo di poche decine di atomi!). Solo verso la metà degli anni '80 due ricercatori italiani, Roberto Car e Michele Parrinello, riuscirono a provare la possibilità di estendere l'uso dei metodi *ab initio* allo studio di problemi dinamici, dando così l'avvio al fertile campo della dinamica molecolare in cui evoluzione temporale e temperatura vengono inclusi senza approssimazioni empiriche. Og-

gi, questo tipo di simulazioni rappresenta una realtà, anche se ristretta ancora a tempi molto brevi e a insiemi di poche decine di atomi per via degli enormi costi computazionali. Quando si rende necessario simulare situazioni molto complesse, l'uso dei metodi *ab initio* si rivela, allo stato attuale, impraticabile. Il problema può venire affrontato mediante l'introduzione di approssimazioni, come ad esempio l'uso di parametri empirici derivati da misure sperimentali in calcoli che vengono, quindi, detti semi-empirici o più drasticamente abbandonando la trattazione quantistica del problema ricorrendo a delle descrizioni approssimate delle energie in gioco nella rottura e formazione di legami. Essendo basate sulle leggi della fisica classica, queste simulazioni vanno sotto il nome di dinamica molecolare classica. In questo modo è possibile effettuare simulazioni su scale temporali significative in tempi accettabili, al prezzo di rinunciare alla massima accuratezza dei risultati. Quando poi si debbono descrivere sistemi ancora più complessi (e questo è il caso reale della maggioranza dei processi chimici) la simulazione diviene di tipo sostanzialmente statistico e rinuncia in parte all'ambizioso progetto di partire dalla leggi fondamentali che descrivono la stabilità degli atomi per giungere a capire fenomeni chimici complessi.



Struttura tridimensionale cristallina del centro emico dell'enzima P450 BM3
Questa e le figure successive dell'articolo presentano esempi di molecole progettate al computer

A seconda del problema che si vuole descrivere esistono, quindi, strumenti computazionali specifici, dai più complessi e accurati, ma anche molto "costosi", metodi *ab initio* statici o dinamici, sino ai più semplici metodi approssimati classici o statistici. Nel seguito del presente articolo, si cercherà di illustrare gli aspetti fondamentali di questi approcci alle simulazioni in chimica.

3. LA TEORIA QUANTISTICA

I primi esempi di calcoli di chimica quantistica risalgono agli anni '30, al tempo in cui la meccanica quantistica era stata ormai formulata nelle sue grandi linee. Come noto, la meccanica quantistica affonda le sue radici nella teoria dei quanti enunciata da Max Planck all'inizio del secolo. Fu solo tra il 1920 e il 1930 che la nuova e rivoluzionaria teoria, secondo cui l'energia viene emessa o assorbita dai sistemi atomici o molecolari secondo quantità ben definite dette "quanti", ebbe un suo pieno riconoscimento nonché un proprio apparato matematico grazie ai contributi di scienziati quali de Broglie, Pauli, Schödinger, Heisenberg, Dirac.

Le leggi della meccanica quantistica si distinguono da quelle della fisica classica per il loro carattere fondamentale: infatti, mentre le leggi quantistiche sono in grado di descrivere altrettanto bene il moto di un elettrone attorno a un nucleo atomico quanto quello di un pianeta attorno al sole, ciò non è più vero per le leggi classiche. Per mezzo delle leggi classiche si può descrivere il comportamento (moto) di un meccanismo composto da molle, leve ecc., se si conoscono alcune costanti di materiali quali la densità, il calore specifico, l'elasticità. Però, se ci si chiede perché le costanti elastiche hanno i valori che hanno, perché una barra si spezza se la tensione cui è sottoposta supera un certo limite, e così via, la fisica classica non fornisce risposta. Essa non dice perché il rame fonde a 1083 °C, perché il vapore di sodio emette luce gialla, perché l'idrogeno ha certe proprietà chimiche, perché il sole brilla, perché il nucleo dell'atomo di uranio si disintegra spontaneamente, perché l'argento conduce l'elettricità, perché lo zolfo è un isolante. A queste domande può dare una risposta la meccanica quantistica.



Sulla base delle elementari interazioni tra le particelle costituenti la materia si può via via risalire ai comportamenti più complessi che danno vita ai fenomeni macroscopici che ogni giorno si osservano. Occorre, però, fare una precisazione: il fatto che la teoria dei quanti possa dare una risposta in teoria a tali domande non sempre significa che la possa dare anche in pratica. Qui entrano in gioco, infatti, la complessità del problema e la necessità di servirsi dei calcolatori elettronici.

4. L'EQUAZIONE DI SCHRÖDINGER

Sempre negli anni '30, la formulazione matematica della fisica quantistica proposta da Schrödinger con la sua funzione d'onda si prestava particolarmente per lo studio di sistemi atomici o molecolari. Alcune grandezze fisiche nei sistemi microscopici sono in relazione con l'energia totale del sistema, ossia con il lavoro che occorre compiere per distruggere il sistema stesso. Se si parla di una molecola, ad esempio, tale quantità corrisponde al lavoro necessario per "scollare" le une dalle altre tutte le particelle presenti (nuclei ed elettroni) e portarle a distanza infinita privandole della loro energia cinetica. Ad ogni sistema atomico o molecolare è possibile associare un operatore hamiltoniano, indicato con H , che descrive l'energia del sistema. Applicare l'operatore hamiltoniano alla funzione d'onda Ψ porta alla celebre equazione di Schrödinger che nella forma più generale è nota come:

$$H \Psi = E \Psi$$

Più in dettaglio si può scrivere:

$$[-\nabla^2 + U(x, y, z)] \Psi(x, y, z) = E \Psi(x, y, z)$$

omettendo alcune costanti moltiplicative. Il termine in parentesi quadra è l'operatore hamiltoniano ed E è l'energia totale del sistema. Schrödinger dimostrò che tale equazione è risolvibile solo per valori quantizzati di E . Nell'operatore hamiltoniano sono contenuti due termini che rappresentano l'energia cinetica e quella potenziale delle particelle presenti nel sistema. Nel caso di un atomo, composto

da un nucleo carico positivamente e da un certo numero di elettroni, l'energia totale del sistema è data dalla somma di alcuni contributi. Innanzitutto, va considerata l'energia cinetica degli elettroni che, essendo in continua rotazione attorno al nucleo, possiedono una propria energia di movimento; poi va considerata l'attrazione elettrostatica nucleo-elettrone nonché la repulsione che ogni elettrone esercita sugli altri elettroni. Quindi l'operatore hamiltoniano conterrà tre termini, ciascuno dei quali rappresenta uno di questi contributi all'energia totale del sistema. L'operatore ∇^2 , descrive l'energia cinetica, mentre le interazioni elettrostatiche, attrattive o repulsive, sono racchiuse nel potenziale $U(x, y, z)$. Le informazioni relative al numero di elettroni e alla loro maggiore o minore distanza dal nucleo, alla loro disposizione spaziale sono, invece, contenute nella funzione d'onda Ψ . Il significato fisico di tale funzione è che essa, o meglio, il suo quadrato, descrive la probabilità di trovare un elettrone in una certa regione di spazio a un dato istante.

Da questi "ingredienti" di partenza è possibile ricavare *ab initio* tutta una serie di informazioni sulla struttura dell'atomo e sulle sue proprietà. Perfettamente simile è il caso molecolare, dove l'operatore hamiltoniano con-

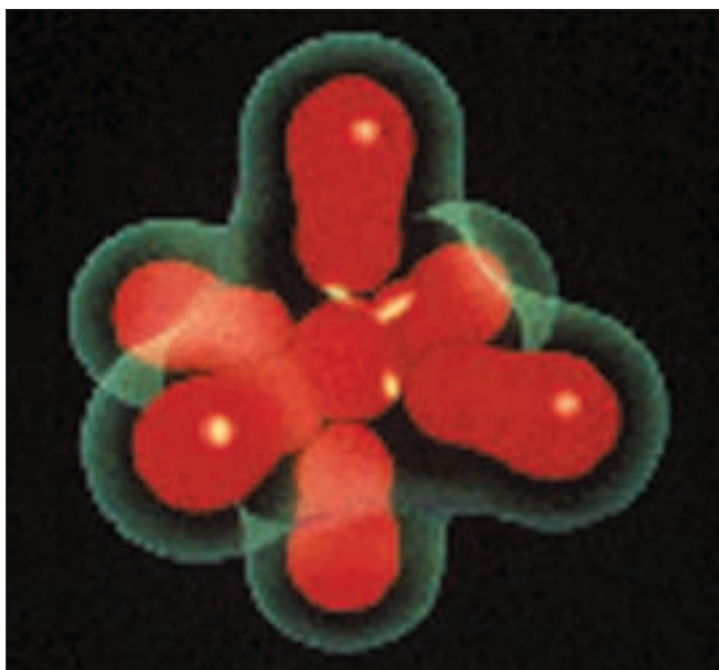
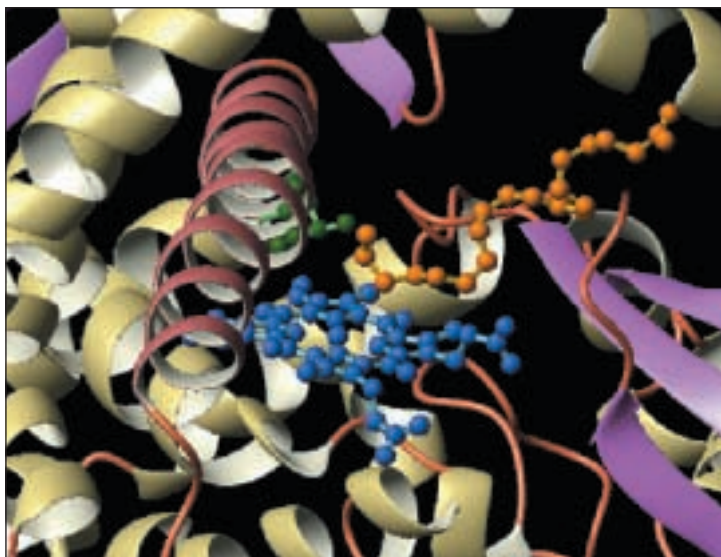


Immagine di una molecola di $Cr(CO)_6$. Le superfici rappresentate sono a densità elettronica costante



Particolare della struttura cristallina del centro eme (complesso ferro-porfirina) dell'enzima P450 BM3

terrà nuovi termini relativi alle interazioni non presenti nell'atomo come, per esempio, la repulsione reciproca dei nuclei atomici o l'attrazione che il nucleo A esercita sugli elettroni del nucleo B e viceversa.

Purtroppo l'equazione di Schrödinger, esattamente risolvibile nel caso dell'atomo di idrogeno e, con grandi sforzi, di altri sistemi molto semplici, si trasforma in un complicato insieme di equazioni differenziali man mano che il sistema in esame aumenta le sue dimensioni. Questo comporta a sua volta la risoluzione di un certo numero di integrali il cui numero cresce grosso modo come N^4 , dove N è il numero di funzioni matematiche usate per descrivere il sistema (di fatto N cresce con il numero di elettroni e , quindi, con la complessità della molecola). Sino alla fine degli anni '50, quindi, la possibilità di studiare teoricamente il comportamento chimico-fisico della materia è stato limitato dalla scarsa diffusione di calcolatori sufficientemente potenti. Per dare un'idea di come l'analisi della struttura elettronica di una molecola richieda uno sforzo computazionale enorme basta dire che gli integrali da calcolare nel caso di una molecola non particolarmente grande come il glucosio sono circa tre milioni e mezzo. Essi possono però diventare molti di più quando si vogliono ottenere dei risultati particolarmente affidati.

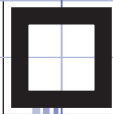
L'aumento progressivo della potenza di cal-

colo, dagli anni '80 a oggi, ha permesso di lavorare con sistemi molecolari di dimensioni via via maggiori. Grazie al calcolo parallelo e allo sviluppo di software in grado di sfruttare al massimo il parallelismo è possibile raggiungere grande efficienza di calcolo e risultati accurati per sistemi molecolari di dimensioni piuttosto grandi. La sfida però non è ancora conclusa: per esempio, lo studio *ab initio* di sistemi proteici può risultare impraticabile anche per i più moderni computer. Tanto per fare un esempio, lo studio di una molecola di interesse biologico come la ferro-porfirina, un'unità di base dell'emoglobina, richiede come minimo il calcolo di circa 220 milioni di integrali, ognuno dei quali viene valutato mediante un buon numero di operazioni.

5. DAI PRIMI CALCOLI QUANTOMECCANICI AI PROGRAMMI USER-FRIENDLY

La complessità delle equazioni da risolvere e del formalismo quantistico ha limitato notevolmente lo sviluppo della chimica teorica almeno sin verso la metà degli anni '60, quando la diffusione dei primi programmi di calcolo di una certa efficienza nonché di elaboratori più potenti ha permesso a questa disciplina di staccarsi dalla fase puramente teorico-matematica per entrare in quella di interpretazione, verifica e infine previsione di dati sperimentali.

Il tipo di calcolo prevede come unici dati di partenza le costanti fisiche del sistema, come la carica e la massa di elettroni e nuclei, e, ovviamente, il numero di elettroni e di nuclei presenti nella molecola nonché le posizioni spaziali dei nuclei stessi. Fatta l'approssimazione che i nuclei siano fissi, dato che la loro velocità di movimento risulta estremamente più bassa rispetto a quella degli elettroni, il calcolo viene effettuato, senza introdurre cioè ulteriori approssimazioni. Secondo la teoria MO (*Molecular Orbital*), un orbitale molecolare è espresso come combinazione lineare di orbitali atomici e la forma, nonché l'energia, di tali orbitali molecolari viene determinata risolvendo un'equazione matriciale ad autovalori-autovettori in cui compaiono matrici di ordine n , dove n è il numero di orbitali di base del siste-



ma. È questo l'approccio noto come metodo di Hartree-Fock, dal nome dei due scienziati che lo hanno sviluppato.

Tra i primi calcoli eseguiti in questo modo vi sono quelli che l'italiano Enrico Clementi fece dal 1960 al 1964 negli Stati Uniti. Il programma di calcolo molecolare IBMOL, che venne messo a punto per questo genere di studi, è uno dei primi che sia circolato negli ambienti scientifici, ed è stato utilizzato per almeno i successivi 10-15 anni. Questo e altri codici basati sul metodo di Hartree-Fock sono stati al centro dell'attività in chimica quantistica con metodi *ab initio* negli anni '60 e '70. Va detto, però, che la soluzione dell'equazione di Schrödinger e la determinazione della funzione d'onda del sistema mediante l'approssimazione di Hartree-Fock soffre di un problema. Il metodo di Hartree-Fock non include, infatti, nel calcolo gli effetti noti come correlazione elettronica, una correzione all'energia totale del sistema piccola in termini relativi ma importantissima ai fini della descrizione accurata del legame chimico. Per fare un esempio, l'energia di legame di una molecola semplice come l'ossigeno molecolare, O₂, è circa un quinto di quella sperimentale quando determinata mediante l'approssimazione di Hartree-Fock. L'introduzione della correlazione elettronica porta, invece, a un risultato praticamente identico a quello sperimentale. Purtroppo l'introduzione della correlazione elettronica a partire dalla funzione d'onda Hartree-Fock rappresenta un problema complesso che ha impegnato i chimici quantistici per oltre trent'anni nel tentativo di mettere a punto algoritmi e metodi efficienti per introdurre questo termine. Va detto che questi sforzi hanno condotto a metodi avanzati per risolvere l'equazione di Schrödinger (teoria delle perturbazioni, interazione di configurazione ecc.) che portano praticamente alla soluzione esatta al punto di dover rivedere e re-interpretare in alcuni casi particolari gli stessi dati sperimentali.

Con lo sviluppo di nuovi formalismi e metodi, e lo scambio nell'ambito scientifico accademico dei primi **codici di calcolo** (si veda a tal proposito il riquadro a p. 38) si è assistito così allo sviluppo via via più rapido anche delle applicazioni della chimica quantistica. In questo senso, un ruolo importante lo ha avu-

to una istituzione nota come QCPE (*Quantum Chemistry Program Exchange*)², presso l'Università dell'Indiana, che per almeno due decenni ha costituito un punto di riferimento per mettere a disposizione della comunità scientifica programmi e software di chimica quantistica, in un'era dominata dalla mancanza di standard, se si eccettua il linguaggio di programmazione, il *Fortran*. Da alcuni di questi codici e *subroutine*, poi inglobati in programmi via via più ricchi, versatili e complessi sono nati alcuni dei programmi oggi disponibili a chiunque voglia fare del calcolo quantistico. Tra questi il più noto è certamente il programma Gaussian, nelle sue molteplici versioni, sviluppato a partire dai primi anni '70 dal gruppo di John Pople alla Carnegie-Mellon University a Pittsburgh. Questo programma è divenuto poi di ampio utilizzo anche per non-esperti e il premio Nobel per la chimica del 1998 a Pople ha, di fatto, sancito l'importante ruolo di questo codice nel rendere popolare e accessibile la chimica quantistica anche a molti ricercatori di formazione sperimentale.

Si può dire, quindi, che l'obiettivo iniziale che ci si è posti nell'affrontare lo studio delle interazioni chimiche per via computazionale, il riprodurre i dati sperimentali esistenti al fine di verificare l'affidabilità del metodo, è stato pienamente raggiunto già a metà degli anni '80. Da allora in poi, la chimica computazionale viene comunemente utilizzata come strumento di modellizzazione, interpretazione, razionalizzazione e previsione e ha, quindi, assunto, a tutti gli effetti, un ruolo attivo e insostituibile nello sviluppo della chimica moderna. La maggior parte delle reazioni di interesse chimico, fisico o biologico, coinvolge però sistemi di grandi dimensioni. Basti pensare ai processi di sintesi di composti organici naturali, alla struttura dello stato solido o alle reazioni del DNA. In questo senso, i metodi basati sul formalismo di Hartree-Fock e sulla introduzione della correlazione elettronica si sono rivelati molto costosi, richiedendo tempi di elaborazione che crescono

² Quantum Chemistry Program Exchange: http://www.chem.indiana.edu/facilities/qcpe_frontend.htm

I codici di calcolo di simulazione quantistica

GAUSSIANO3 (<http://www.gaussian.com/>)

Gaussian 03 è la versione più recente della serie *Gaussian* di programmi per il calcolo della struttura elettronica. *Gaussian 03* è usato da chimici, ingegneri chimici, biochimici, fisici e altri scienziati per la ricerca in ambiti consolidati ed emergenti di interesse chimico. Partendo dalle leggi fondamentali della meccanica quantistica, *Gaussian* è in grado di prevedere le energie, le strutture molecolari, e le frequenze di vibrazione di sistemi molecolari, insieme a numerose altre proprietà molecolari derivanti da questo tipo di conti di base. Può essere applicato allo studio di molecole e reazioni in un'ampia finestra di condizioni, che include sia le specie stabili sia i composti che non possono o possono solo con difficoltà essere osservati sperimentalmente, come, ad esempio intermedi a vita breve e strutture di transizione.

GAMESS-UK (<http://www.cse.clcr.ac.uk/qcg/gamess-uk/>)

GAMESS-UK è un codice per il calcolo *ab initio* della struttura elettronica. È in grado di svolgere calcoli Hartree-Fock, DFT e multi-configurazionali, oltre che tutta una serie di calcoli utilizzando le varie tecniche post Hartree Fock. *GAMESS-UK* è basato sul software originale *GAMESS* (1981) e il suo sviluppo è stato coordinato negli ultimi dieci anni dal Daresbury Laboratory.

ADF (<http://www.scm.com/>)

ADF è un pacchetto software basato sulla Teoria del Funzionale della Densità sviluppato presso l'Università di Amsterdam. Permette di svolgere calcoli di struttura elettronica a principi primi ed è usato con profitto da chimici teorici e computazionali di tutto il mondo, sia nell'industria che nell'accademia. *ADF* è usato in effetti in tutti i campi della chimica, ed è particolarmente popolare in catalisi, chimica inorganica, in quella dei metalli pesanti, come in spettroscopia. Può essere applicato a tutti i sistemi, proteine, solventi, polimeri, superfici, e solidi, alle semplici molecole in fase gassosa.

NWCHEM (<http://www.emsl.pnl.gov:2080/docs/nwchem/>)

NWChem è un pacchetto software di chimica computazionale, progettato per essere eseguito sia su supercomputer paralleli ad alta prestazione, sia su convenzionali *cluster* di personal computer. L'obiettivo della progettazione è quello di creare un codice in grado di scalare al meglio sia nell'abilità di trattare efficientemente problemi di grosse dimensioni, sia nell'uso di risorse di calcolo parallele. *NWChem* è stato sviluppato presso la *Pacific Northwest National Laboratory* (PNNL).

TURBOMOLE (<http://www.turbomole.com/>)

TURBOMOLE è stato sviluppato presso l'Università di Karlsruhe ed è progettato con speciale attenzione per stazioni di lavoro UNIX e personal computer, sfruttando in modo efficiente le caratteristiche e capacità di questo tipo di hardware. Una caratteristica eccellente di *TURBOMOLE* è costituita dagli algoritmi semi-diretti adattabili alla memoria e alle richieste di spazio-disco.

Q-CHEM (<http://www.q-chem.com>)

Q-Chem 2.0 è in grado di condurre calcoli a principi primi sia sullo stato fondamentale che sugli stati eccitati delle molecole. Include tutta una serie di metodi che non sono disponibili in altri pacchetti di chimica quantistica computazionale.

CPMD (<http://www.cpmc.org>)

CPMD (Dinamica Molecolare Car-Parrinello) è un codice a onde piane/pseudopotenziali basato sulla Teoria del Funzionale della Densità, progettato specialmente per condurre calcoli di dinamica molecolare a principi primi. È applicato principalmente a problemi di stato solido.

DeMon (http://www.demon-software.com/public_html)

Il programma *deMon-KS*, sviluppato presso l'Università di Montreal, permette di condurre calcoli *ab initio*, basati sulla Teoria del Funzionale della Densità, su sistemi di grosse dimensioni, caratterizzati anche dalla presenza di metalli di transizione, ottenendo risultati di alta precisione in tempi relativamente brevi.

CRYSTAL (<http://www.cse.clcr.ac.uk/cm/g/CRYSTAL>)

Il programma *CRYSTAL*, sviluppato presso l'Università di Torino, calcola la struttura elettronica di sistemi periodici con i metodi Hartree Fock e DFT. Il codice può essere utilizzato per condurre studi approfonditi su struttura e proprietà fisiche, elettroniche e magnetiche di molecole, polimeri, superfici e solidi cristallini.

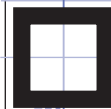
VASP (<http://cms.mpi.univie.ac.at/vasp/>)

VASP è stato sviluppato presso l'Università di Vienna e progettato per condurre calcoli di dinamica molecolare (MD) quantistica da primi principi utilizzando pseudopotenziali e funzioni di base costituite da onde piane. È particolarmente indicato per lo studio e la descrizione di solidi cristallini.

Un sito in cui è possibile trovare un elenco dettagliato di pacchetti software per simulazioni in aree di interesse chimico è il seguente:

SAL (<http://sal.kachinatech.com/Z/2>)

SAL (Applicazioni Scientifiche su Linux) è una collezione di informazioni e collegamenti virtuali a pacchetti software di interesse in campo scientifico e ingegneristico.



come N^7 dove N è il numero funzioni usate per descrivere il sistema. È a questo punto che nel panorama delle simulazioni quantistiche in chimica si è affacciata, in modo decisivo, una diversa metodologia per affrontare il problema.

6. LA TEORIA DEL FUNZIONALE DELLA DENSITÀ

Dall'inizio degli anni '90, la teoria del funzionale della densità, che sino ad allora occupava una posizione periferica in ambito chimico essendo stata sviluppata per lo studio di solidi cristallini, ha svolto un ruolo da protagonista nell'ambito della modellistica chimica. Lo sviluppo è dovuto essenzialmente all'eccellente accuratezza raggiunta dai metodi basati su questa teoria, ma anche al loro costo computazionale, decisamente inferiore rispetto ai metodi tipo Hartree-Fock. La teoria del funzionale della densità ha trovato inizialmente applicazione nell'ambito della fisica dello stato solido, e solo in un secondo momento, anche in quello della chimica. La transizione non è stata semplice in quanto il linguaggio della comunità dei fisici dello stato solido era completamente estraneo a quello ormai consolidato nella comunità dei chimici quantistici.

Alla base della teoria del funzionale della densità c'è il teorema di Hohenberg e Kohn del 1964 secondo cui l'energia elettronica dello stato fondamentale di un solido o di una molecola, E , è determinata completamente dalla densità elettronica ρ , una grandezza che descrive la distribuzione degli elettroni attorno ai nuclei atomici. In altre parole, nota la densità elettronica dello stato fondamentale, l'energia totale del sistema (come anche le altre proprietà elettroniche) è determinata in modo univoco. La conseguenza di questo teorema può essere meglio illustrata per confronto con la teoria degli orbitali molecolari. Mentre la complessità della funzione d'onda molecolare Ψ aumenta con il numero degli elettroni, la densità elettronica, che dipende solo dalle tre coordinate spaziali, ha lo stesso numero di variabili, indipendentemente dalle dimensioni del sistema.

Nonostante sia stato dimostrato che a cia-

scuna differente densità elettronica corrisponde una differente energia dello stato fondamentale, esiste ancora un problema non risolto per questa teoria: il funzionale che mette in relazione le due quantità, E e ρ , non è noto. L'obiettivo dei vari metodi basati sulla teoria del funzionale della densità, la DFT (*Density Functional Theory*), è la formulazione di funzionali in grado di mettere in relazione la densità elettronica con l'energia.

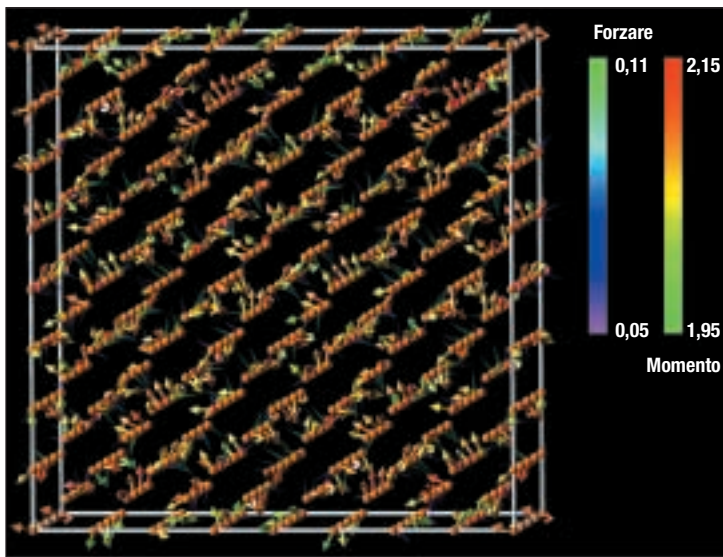
I metodi DFT sono stati adottati dalla comunità dei chimici computazionali anche grazie all'introduzione degli orbitali di Kohn e Sham (1965). Alla base della teoria di Kohn-Sham vi è il calcolo dell'energia cinetica E_c nell'ipotesi che gli elettroni non interagiscano. Poiché nel sistema reale gli elettroni interagiscono, l'energia cinetica calcolata in questa approssimazione non è quella totale. Tuttavia, la differenza tra il valore esatto e quello calcolato è piccola e viene assorbita in un termine definito di scambio-correlazione.

Un'espressione generale per l'energia DFT può essere scritta come:

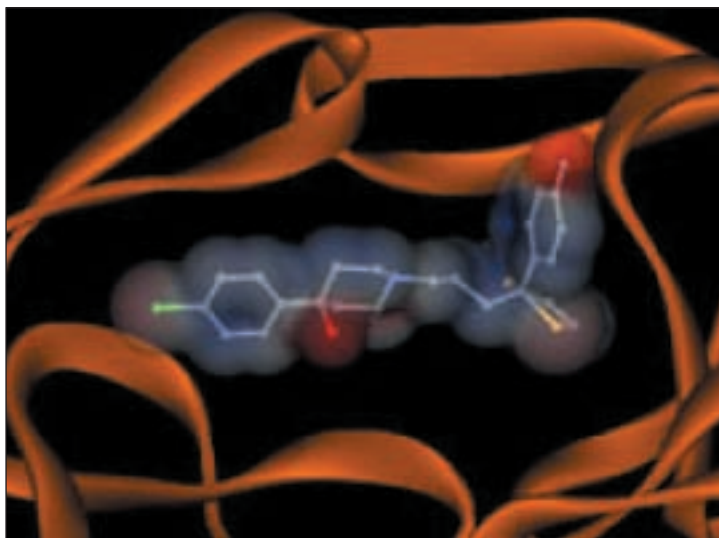
$$E_{\text{DFT}} = E_c(\rho) + E_{\text{ne}}(\rho) + J(\rho) + E_{\text{xc}}(\rho)$$

dove E_c è l'energia cinetica nell'ipotesi di elettroni non-interagenti, E_{ne} è l'energia di attrazione nucleo-elettrone, J è l'energia di interazione coulombiana e E_{xc} è l'energia di scambio-correlazione. In questo modo, la correlazione elettronica, che con i metodi basati sulla determinazione della funzione d'onda tipo Hartree-Fock richiede calcoli estremamente onerosi, viene introdotta in modo molto diretto e senza sostanziale aggravio. Il risultato è una accuratezza elevata a un ragionevole costo di elaborazione.

Il maggior problema nell'ambito della teoria DFT è ottenere funzionali in grado di descrivere bene il termine di scambio-correlazione. Una volta determinato il funzionale, il problema risulta molto simile a quello incontrato per la teoria degli orbitali molecolari. Nonostante si possano riscontrare molte somiglianze, esistono anche notevoli differenze rispetto alla teoria basata sugli orbitali molecolari. La forma dei funzionali è progettata in modo che essi presentino un



Il risultato del calcolo dei momenti magnetici per una cella unitaria di 512 atomi di ferro mediante calcoli di tipo DFT. Il momento magnetico di ciascun atomo è descritto da un vettore



Modello a nastro per l'enzima proteasi HIV-1 complessato a un inibitore non peptidico

determinato comportamento limite, e che riproducano determinati parametri noti. Nella *Local Density Approximation* (LDA) si assume che la densità possa essere trattata localmente come un gas uniforme di elettroni. Gli sviluppi dell'approccio LDA hanno portato a considerare un gas di elettroni non-uniforme. Ciò comporta una dipendenza dell'energia del sistema anche dalla derivata prima della densità. Questi metodi sono noti come *Gradient Corrected or Generalized Gradient Approximation* (GGA) e sono

quelli che hanno consentito l'applicazione con successo della teoria DFT ai problemi di interesse chimico dove sono richieste accuratissime molto elevate. Oggi, la quasi totalità delle simulazioni da primi principi fatte in ambito chimico si basano sul metodo DFT, mentre solo 10-15 anni fa non più del 10-15% degli studi si basavano su questo approccio.

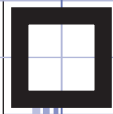
7. I METODI SEMIEMPIRICI

Dalla seconda metà degli anni '60, hanno trovato largo impiego i metodi semiempirici, così detti in quanto oltre alle informazioni di base già presenti nei metodi *ab initio* si introducono altri dati di carattere sperimentale (da cui il nome di metodi semiempirici). In questo modo, molti degli integrali da calcolare, corrispondenti ad altrettante interazioni all'interno della molecola vengono trascurati o parametrizzati in modo da ridurre drasticamente le dimensioni del calcolo. Inoltre, questi metodi tendono a considerare espressamente solo alcuni degli elettroni presenti nella molecola e in particolare quelli che determinano le proprietà chimiche.

Grazie ai metodi semiempirici è stato possibile per la prima volta gettare un ponte tra teoria e sperimentazione negli anni in cui la complessità dei conti *ab initio* era fuori portata, sia per la mancanza di algoritmi efficienti sia per la ridotta capacità dei processori (si parla degli anni '60 e '70). Oggi, questi metodi vengono spesso usati per descrivere la porzione più esterna rispetto al sito reattivo di una molecola o di un solido. In pratica, la molecola viene suddivisa in due regioni trattate a livello differente. La porzione più interna viene trattata ad un elevato livello di teoria (*ab initio*), mentre la porzione più esterna viene trattata con un metodo semiempirico.

8. DINAMICA MOLECOLARE

Fino ad ora, si è analizzato in dettaglio come calcolare l'energia totale e le proprietà di una molecola, senza considerare la sua evoluzione nel tempo. Si è parlato, infatti, dei metodi "statici", in cui le posizioni nucleari del siste-



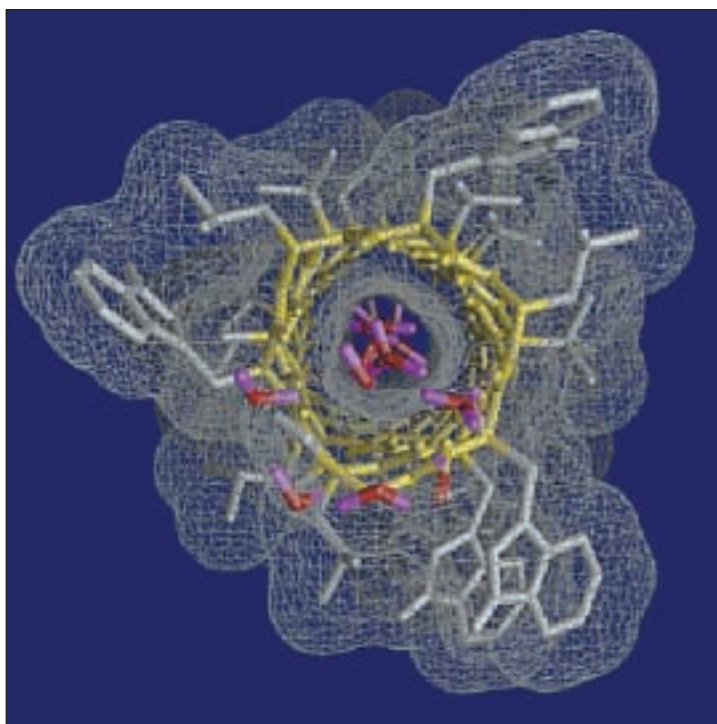
ma sono fisse. Si è però detto all'inizio che il maggior interesse sta nella trattazione di fenomeni dinamici.

I metodi di simulazione tradizionale possono essere suddivisi in due classi:

1. simulazioni stocastiche;
2. simulazioni deterministiche, che sono ampiamente rappresentate dal metodo di Monte Carlo (MC) e dal metodo di Dinamica Molecolare (MD), rispettivamente.

Le simulazioni Monte Carlo si basano su un approccio statistico. Varie configurazioni del sistema vengono generate per spostamenti casuali delle particelle e un algoritmo stabilisce se accettare o rifiutare una nuova configurazione sulla base della variazione di energia. Cammini verso energie inferiori sono sempre accettati, quelli a più alte energie sono accettati con una probabilità determinata dalla statistica di Boltzmann. In questo modo, le proprietà del sistema possono essere calcolate come media su tutti gli spostamenti. Al contrario, i metodi di dinamica molecolare si basano sulle equazioni del moto definite dall'hamiltoniana del sistema. Il metodo consiste nell'integrazione di queste equazioni per ottenere le nuove posizioni e velocità delle particelle. Le simulazioni richiedono sempre la definizione di un modello (campo di forze) per le interazioni tra le varie componenti del sistema. Tale modello deve rispettare alcune leggi fondamentali e va testato attraverso un confronto con i dati sperimentali disponibili. Gli ingredienti indispensabili per una simulazione di dinamica molecolare sono: il modello, un integratore per propagare le posizioni e le velocità in un intervallo di tempo prestabilito e un *ensemble statistico*. Il risultato della simulazione è corretto solo rispetto al modello scelto e va, quindi, analizzato sulla base delle previsioni teoriche e delle osservazioni sperimentali.

Una questione importante relativa alle simulazioni è la scala dei tempi e delle lunghezze percorribili. Più sofisticata è la tecnica di simulazione, più breve è l'intervallo di tempo e delle lunghezze accessibili. La *performance* di una simulazione di dinamica di particelle dipende fortemente dalle attrezzature computazionali disponibili. Metodi di dinamica molecolare classica sono oggi applicati a una vasta gam-



La gramicidina costituisce il canale proteico più studiato. La struttura riportata è stata ottenuta attraverso un calcolo di dinamica molecolare

ma di problemi, come, ad esempio, proprietà dei liquidi, difetti nei solidi, proprietà di superficie, grandi aggregati molecolari, biomolecole ecc.. Negli ultimi anni molti programmi sono stati progettati anche per il calcolo su computer paralleli, accelerando, quindi, in modo considerevole, i tempi di calcolo.

Il regno dei metodi di dinamica molecolare tradizionale e di struttura elettronica si è esteso ampiamente grazie allo sviluppo della famiglia di tecniche chiamata dinamica molecolare *ab initio* di cui si diceva al paragrafo 2. L'idea alla base di questi metodi è di calcolare le forze che agiscono sui nuclei determinando la struttura elettronica del sistema da primi principi. Il problema, in questo caso, non è più selezionare il campo di forze ma selezionare il livello di approssimazione a cui risolvere l'equazione di Schrödinger. Questo permette di studiare anche sistemi chimicamente complessi. Ovviamente, i notevoli vantaggi sono accompagnati da un alto prezzo da pagare a causa del collocamento della dinamica molecolare in una cornice *ab initio*: i tempi e le lunghezze accessibili sono molto più piccoli. Tuttavia, quest'ultima tecnica, non influenzata da modelli predisposti, per-

mette di riscontrare eventuali fenomeni inattesi e, quindi, presenta una potenzialità di previsione molto elevata.

9. LE RISPOSTE DELLA CHIMICA TEORICA

Chiariti pregi e difetti dei principali metodi di calcolo di cui si avvale oggi la chimica computazionale si può tentare di illustrare le domande a cui l'uso combinato dei principi della meccanica quantistica e dei calcolatori elettronici può dare una risposta adeguata. Per prima cosa è possibile uno studio conformazionale di molecole, ioni, radicali ecc., ossia della disposizione geometrica ottimale dei nuclei atomici all'interno della struttura molecolare. Si tratta di un aspetto piuttosto importante in quanto dalla "forma" delle molecole spesso dipende la reattività. Per determinare la geometria di una molecola si spostano gli atomi dalle loro posizioni valutando le conseguenti variazioni dell'energia totale del sistema. La geometria che corrisponde alla minima energia è quella più favorita.

Una seconda importante risposta riguarda

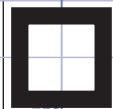
la stabilità dei sistemi chimici. Per fare un semplice esempio, si consideri la molecola di idrogeno, H_2 , la cui energia è stata valutata per via teorica in 31,8 eV (elettronvolt). L'energia dell'atomo di idrogeno è di 13,6 eV, per cui dalla combinazione di due atomi a dare la molecola si ha un guadagno di 4,6 eV. Questa è la forza del legame H-H e corrisponde all'energia che occorre fornire per "rompere" la molecola H_2 e riottenere i due atomi separati. L'importanza della valutazione delle energie di legame è notevole. Ogni reazione chimica, partendo da certi reagenti iniziali, conduce a nuovi prodotti finali coinvolgendo la rottura di alcuni legami e la formazione di altri. Se i legami di partenza sono deboli è evidente che la reazione sarà più facile in quanto tali legami si romperanno senza che si debba fornire troppa energia (per esempio, sotto forma di riscaldamento). Inoltre, se si formano nuovi legami particolarmente forti il prodotto finale sarà stabile e non tenderà a decomporre o ad alterarsi.

Va fatta, però, una distinzione tra gli aspetti termodinamici e quelli cinetici. La variazione di energia nel passaggio da reagenti a prodotti caratterizza l'energetica della reazione ma può accadere che un processo favorito dal punto di vista energetico presenti una velocità di reazione talmente bassa che la reazione non può essere osservata. La chimica computazionale permette di calcolare non solo le quantità termodinamiche ma anche le barriere di attivazione per il passaggio da reagenti a prodotti e, quindi, la cinetica o velocità della reazione.

L'altro grande campo di applicazione è quello della determinazione delle proprietà elettroniche del sistema. Un numero enorme di informazioni è, infatti, racchiuso nella funzione d'onda elettronica del sistema o, il che è di fatto la stessa cosa, nella sua densità totale nel caso si usino metodi tipo DFT. Per esempio, è possibile assegnare delle cariche atomiche, un parametro molto utile per la discussione e la razionalizzazione della struttura e della reattività molecolari. Da queste derivano poi i potenziali elettrostatici delle molecole che determinano molte delle interazioni con quello che sta attorno, per esempio, molecole di solvente o altre molecole.



Copertina del numero di Marzo 2003 della rivista "Science" dedicata allo studio della catalisi sulla base del calcolo quantistico



Basti dire che dalla forma del potenziale elettrostatico dipendono molte delle interazioni biologiche e farmacologiche. Esiste poi la risposta di molecole e solidi alla radiazione elettromagnetica che sta alla base delle moderne spettroscopie. Grazie a queste accurate tecniche di indagine è possibile acquisire un gran numero di informazioni sulle sostanze chimiche, nonché riconoscerle anche quando sono presenti in tracce. A questo proposito, basta accennare a una tecnica spettroscopica sviluppata in origine a puro scopo analitico, la *Risonanza Magnetica Nucleare* (NMR) ma che viene oggi utilizzata comunemente in diagnostica medica. Spettri NMR possono essere simulati mediante il calcolo quantistico e, attualmente, è piuttosto comune affiancare alla misura sperimentale i valori calcolati in modo da poter effettuare assegnazioni precise e prive di ambiguità. Queste risposte riguardano tantissime proprietà, da quelle vibrazionali e quelle ottiche (il che permette di spiegare perché e quando certi composti sono colorati), da quelle di conducibilità elettrica ai comportamenti magnetici ecc.. Si può dire che oggi la previsione delle proprietà elettroniche e spettroscopiche dei composti chimici è forse una delle conquiste più importanti della chimica computazionale.

10. CONCLUSIONI

La teoria dei quanti ha da poco compiuto i 100 anni, essendo stata formulata da Max Planck ai primi del Novecento, ma solo da poco più di un decennio le simulazioni quantitative in chimica hanno assunto un ruolo veramente importante. Questo è dovuto, da una parte, al vorticoso sviluppo della capacità di elaborazione e dall'altra al continuo progresso della teoria nel fornire strumenti adeguati ed efficaci per la soluzione del problema. I grandi progressi ottenuti su questo secondo fronte hanno permesso di raggiungere livelli di affidabilità delle simulazioni tali da promuovere, a tutti gli effetti, la chimica computazionale da un ruolo ancillare a quello di disciplina complementare al metodo sperimentale. Il confronto diretto dei risultati computazionali con i valori sperimentali permette non solo di convalidare il modello teorico proposto, ma anche di fare previsioni e suggerire eventuali altri esperimenti da condurre. Recenti progressi nel campo dei fenomeni dinamici, lo sviluppo di algoritmi sempre più efficienti e accurati, insieme al prevedibile aumento delle capacità di elaborazione garantiscono che il ruolo delle simulazioni in chimica è destinato ad aumentare nei prossimi anni a una velocità maggiore di quella sostenuta sin qui.

GIANFRANCO PACCHIONI è Ordinario di Chimica dello Stato Solido presso il Corso di Laurea in Scienza dei Materiali dell'Università di Milano-Bicocca. Ottenuto il Dottorato di ricerca a Berlino nel 1984, è stato Visiting Professor al centro ricerche IBM di Almaden in California, al Politecnico di Monaco di Baviera e all'Università di Barcellona. È autore di oltre 250 lavori sulla struttura elettronica mediante calcolo quantistico di complessi e materiali inorganici. Dal 1999, dirige la Scuola di Specializzazione in Scienza e Tecnologia dei Materiali dell'Università di Milano-Bicocca.
gianfranco.pacchioni@unimib.it

CRISTIANA DI VALENTIN ha conseguito il Dottorato di ricerca nel 2000 presso l'Università di Pavia, svolgendo una tesi di chimica computazionale in collaborazione con la Technische Universität di Monaco. Nel 2002, si è specializzata in Scienza e Tecnologia dei Materiali presso l'Università di Milano-Bicocca. Da dicembre 2002 è ricercatore universitario presso il Dipartimento di Scienza dei Materiali dell'Università di Milano-Bicocca.
cristiana.divalentin@unimib.it