

## Che cos' è il Modelling Computazionale

Con il termine *modelling* computazionale si designano una varietà di tecniche computazionali sviluppate con lo scopo di simulare, razionalizzare, prevedere, il comportamento e le caratteristiche di sistemi complessi. Essendo praticamente impossibile dare una definizione comprensiva di tutto ciò che con tale termine viene oggi inteso, conviene procedere esemplificando. Si tratta in generale di approcci del tipo *bottom-up* in cui si parte dalla scala atomica per arrivare a sistemi meso e macroscopici. L' esempio tipico è quello delle simulazioni di dinamica molecolare (DM) in cui i potenziali di interazione fra atomi/molecole (*force-fields*), ottenuti generalmente attraverso calcoli di Chimica Quantistica, vengono utilizzati per integrare le equazioni di Newton per un sistema a molte particelle (fino a  $10^{12}$ ) per tempi tipicamente dell'ordine di decine di nanosecondi. Dalla elaborazione dei risultati si possono ricavare funzioni di correlazione utili per l'interpretazione di spettri di diffrazione neutronica, relazioni P-V-T, diagrammi di fase, proprietà di trasporto, strutture di macromolecole sintetiche o biologiche ed informazioni sul riconoscimento molecolare e sull'accessibilità dei siti attivi. Nel metodo Montecarlo (MC) si ottengono direttamente informazioni sul sistema all'equilibrio generando configurazioni *random* che poi sono pesate in accordo con la Meccanica Statistica. Ciò presenta un ovvio vantaggio per sistemi che raggiungono l'equilibrio in tempi superiori a quelli accessibili a simulazioni MD.

Un approccio più corretto, ma molto più limitato per quanto riguarda le dimensioni dei sistemi trattabili, è quello in cui sia il moto degli elettroni che quello dei nuclei è calcolato per via quantomeccanica (QM). Per sistemi più estesi si può ricorrere a metodi QM in cui attraverso il calcolo elettronico si generano *on the fly* le forze necessarie per la dinamica nucleare. Molto popolare è il metodo di Car-Parrinello basato sul funzionale della densità elettronica (DFT).

Una delle caratteristiche salienti di un efficace approccio alla modellizzazione di molecole e materiali è quello di adattare il livello di accuratezza al tipo di problema studiato. Per esempio, nella simulazione della frattura di un materiale è necessario trattare la regione della frattura con metodi QM, poiché lì si rompono e si riformano legami, mentre la regione adiacente può essere efficacemente trattata via MD. Per la regione ancora più esterna si può far ricorso alla trattazione classica del problema elastico (elementi finiti).

Tali approcci multiscala sono indispensabili per procedere dal livello atomico a quello mesoscopico o addirittura macroscopico. Un altro esempio tipico è quello della razionalizzazione del comportamento fotofisico/fotochimico di molecole biologiche che si trovano in vivo in ambiente proteico. In questo caso è necessario accoppiare lo studio quanto-meccanico accurato della dinamica nucleare (effetti non adiabatici compresi) alla trattazione efficace del ruolo del solvente e della proteina. Tipicamente si può, ad esempio, per quanto riguarda il ruolo dell'acqua, includere, se necessario una prima *shell* di idratazione nel calcolo QM (approccio della supermolecola) e poi procedere con un modello di dielettrico continuo. L' ambiente proteico può infine essere simulato via MD.

Gli esempi citati fanno anche intravedere che le problematiche aperte in questo campo spaziano in un ambito molto grande: dallo sviluppo di metodi per il calcolo della struttura elettronica con favorevoli proprietà di scaling (cioè tempo di calcolo in funzione del numero di elettroni), fino allo sviluppo di computer con architetture in grado di supportare le estese parallelizzazioni necessarie per le simulazioni più impegnative. Un cenno particolare meritano gli sforzi per interfacciare in modo efficace i vari metodi operanti in un approccio multiscala e quelli, basati sull'introduzione di potenziali fra frammenti sempre più grandi (*coarse-graining*) indispensabili per simulare sistemi meso- e macroscopici.

## Potenzialità scientifiche e tecnologiche

La Chimica Computazionale ed il *molecular modelling* hanno avuto negli ultimi anni uno sviluppo impetuoso, sostenuto dal continuo aumento della capacità di calcolo. Grazie ai numerosi codici disponibili commercialmente, caratterizzati da facilità di input e di visualizzazione dei risultati (orbitali occupati e virtuali, modi normali di vibrazione,...) è oggi molto comune il ricorso al calcolo anche da parte di non specialisti, per l'interpretazione della reattività chimica e fotochimica, delle proprietà termodinamiche e spettroscopiche,... La Chimica Computazionale si rivela uno strumento sempre più affidabile e capace di trattare efficacemente anche problemi complessi, come quelli in cui l'interazione fra la molecola ed il solvente hanno un ruolo importante o quelli in cui è necessaria un'accurata determinazione della conformazione molecolare (riconoscimento molecolare, catalisi enzimatica, supporto all'interpretazione dei diffrattogrammi X di proteine). Progressi significativi sono stati ottenuti anche nel campo dello studio di nanoparticelle, strutture bidimensionali (film sottili), *quantum dots*, *quantum wells*, oltre che nello studio delle bande elettroniche e fononiche di solidi. Uno sviluppo forse anche maggiore si è avuto nella modellizzazione e simulazione di strutture meso e macroscopiche attraverso approcci multiscala.

È da notare come l'uso delle tecniche di *molecular/materials modelling* si siano notevolmente diffuse anche a livello industriale. All'inizio (una ventina di anni fa) le uniche applicazioni riguardavano il *drug design*, avendo verificato che era molto più comodo ed economico investigare ad esempio il ruolo di sostituenti su una certa molecola per via teorica (sulla base ad esempio di correlazioni fra distribuzioni di carica e proprietà farmacologiche) anziché seguire la strada di lunghe e costose sintesi. Poi progressivamente si è registrato un crescente ricorso a tecniche di *modelling* per lo studio di quasi tutti i prodotti dell'industria chimica, dai polimeri ai nuovi materiali, dai cosmetici ai lubrificanti. Oggi le tecniche di *molecular modelling* sono uno degli strumenti principali a disposizione delle sezioni R&D delle principali industrie chimiche, manifatturiere, elettroniche e di molti laboratori governativi coinvolti ad esempio in ricerche militari (es. Livermore e Sandia negli USA). Molti dei codici di calcolo usati hanno avuto origine in ambito accademico (o di enti di ricerca). In uno studio del 1996 il settore della Chimica Computazionale è stato indicato come strategico per lo sviluppo industriale degli Stati Uniti, dovendo svolgere un ruolo primario per il design non solo di nuovi prodotti ma anche dei modi per produrli con efficienza energetica e in modo sostenibile dal punto di vista degli effetti sull'ambiente. Ci sono in pratica varie tipologie di uso dello strumento computazionale, che richiedono tutte una stretta collaborazione con l'attività di laboratorio. Uno dei più diffusi è la ricerca di relazioni fra l'attività di una molecola e le sua struttura (QSAPR = *quantitative structure-activity relationships*) o fra le proprietà di una molecola e la sua struttura (QSPR = *quantitative structure-properties relationships*). In entrambi casi i dati relativi alla struttura sono spesso accessibili solo attraverso calcoli QM (es. distribuzioni di carica, potenziali elettrostatici,...). Gli studi sul *docking* di leganti a molecole biologiche e quelli di dinamica molecolare per la risoluzione della struttura di una proteina ad integrazione dei dati ottenuti per diffrazione dei raggi X stanno avendo grande successo e sono ormai numerosi i brevetti di farmaci concepiti facendo largo uso di tecniche di *modelling* computazionale.

Anche nel campo della catalisi omogenea di processi di polimerizzazione e nella acquisizione di dati termodinamici e cinetici per l'ottimizzazione di processi industriali le tecniche di *modelling* si stanno rivelando estremamente utili e vantaggiose.

Nel campo dei materiali le applicazioni sono numerose, dai calcoli quantomeccanici di proprietà elettroniche, ottiche e magnetiche di solidi, *quantum dots*, *quantum wells*, *clusters*, nanoaggregati e nanotubi, alle simulazioni di crescita di cristalli da fasci molecolari o per deposizione dal vapore. E ancora la sinterizzazione, l'ossidazione, la frattura, le transizioni di fase indotte dalla pressione e dalla temperatura.

## Il Modelling Computazionale nel CNR

È possibile classificare coloro che operano nel campo del *modelling* computazionale all'interno del CNR essenzialmente in due gruppi, pur con significative sovrapposizioni.

- 1) Utenti di software destinato al *modelling*, con competenze in particolari settori della Chimica e Fisica dei materiali, delle biomolecole, dei farmaci, della *softmatter*, della catalisi, ... Si tratta di persone che, avendo acquisito un'elevata preparazione unita ad una grande sensibilità per problemi specifici, operano in genere a stretto contatto con gruppi sperimentali integrandone i risultati con simulazioni al computer via software disponibile commercialmente.
- 2) Sviluppatori di metodi di calcolo ai vari livelli previsti da un approccio multiscala. I campi di competenza possono spaziare in ambito molto vasto. Solo a titolo di esempio citiamo la ricerca di nuovi funzionali per migliorare la qualità di calcoli DFT, lo sviluppo di metodi efficaci per trattare la dinamica nucleare, l'implementazione di metodi per la trattazione del ruolo del solvente nel determinare la struttura elettronica, il calcolo, con i metodi più raffinati della Chimica computazionale, di *force fields* per la MD.

Come indicato in precedenza, molti ricercatori si collocano in una posizione intermedia, avendo sia la capacità di contribuire al miglioramento delle metodologie di calcolo che di usare efficacemente il software commerciale esistente, di cui talvolta sono stati essi stessi degli sviluppatori.

Le specifiche tematiche di ricerca a cui vengono applicati i metodi della Chimica computazionale e del *modelling* sono vari. Ecco un elenco sicuramente incompleto ma significativo.

- Calcolo di stati elettronici e di superfici di energia potenziale con metodi dipendenti dalle dimensioni dei sistemi studiati (post Hartree-Fock, *coupled cluster* fino alle triple eccitazioni, DFT e TDDFT, semiempirici) con applicazioni allo studio di sistemi di interesse per la chimica dell'atmosfera, per la catalisi omogenea ed eterogenea, per la chimica dei complessi metallici, per una migliore comprensione delle relazioni struttura-proprietà e struttura-reattività di specie coinvolte in processi di interesse.
- Studi con metodi di dinamica quantistica ed ibridi (quantistici-classici) del moto nucleare in stati eccitati, per l'interpretazione di dati spettroscopici anche risolti nel tempo, per la femtochimica, per il controllo di processi chimici e fisici attraverso impulsi laser, anche molto intensi, opportunamente modulati.
- Struttura elettronica di specie contenenti atomi pesanti (effetti relativistici).
- Studio con metodi DFT di proprietà strutturali e spettroscopiche di grandi molecole (peptidi, proteine, acidi nucleici, farmaci) con l'inclusione di effetti di solvente via PCM (*polarizable continuum model*) di grande aiuto per l'assegnazione, l'interpretazione e la razionalizzazione di molti dati sperimentali.

- Studi con metodi AIMD (*ab-initio molecular dynamics*) di sistemi complessi in cui gli effetti quantistici sono importanti, condotti accoppiando calcoli elettronici accurati con metodi di dinamica molecolare *on the fly* (per esempio Carr-Parrinello).
- Studio di specie adsorbite su metalli e semiconduttori, con applicazioni alla fotosensibilizzazione, alla catalisi eterogenea, alla progettazione di bio-sensori.
- Calcolo di proprietà strutturali, elettriche ed ottiche lineari e non-lineari (anche in soluzione) di molecole ed aggregati supramolecolari, *molecular wires*, nanotubi, nanoparticelle, *nanoalloys*.
- Applicazione di MD e MC allo studio di conformazioni di biomolecole e polimeri, alla catalisi enzimatica, al riconoscimento molecolare, al *docking* molecolare, allo studio della *softmatter*, alla determinazione della struttura di biomolecole e materiali, come complemento dell'analisi cristallografica a raggi X.
- Modellizzazione di processi in condizioni estreme (plasmi chimici) di interesse per lo studio dell'interazione laser-solido, della catalisi eterogenea, di problematiche aerospaziali, di reattori a plasma. Modellizzazione di antenne ed apparati di rivelazione di onde elettromagnetiche.

## Il Progetto “Modelling Computazionale” del DPM

Il Dipartimento Progettazione Molecolare (DPM) del CNR ha riconosciuto che le tecniche di modellizzazione di molecole e materiali possono contribuire in modo significativo allo sviluppo in settori della Ricerca ritenuti strategici per il loro grande impatto socio-economico: la salute (design di nuovi farmaci e di nanoveicolanti), la sostenibilità (design di nuovi materiali e processi industriali ottimali sia dal punto di vista energetico che dell'impatto ambientale), lo sviluppo delle cosiddette tecnologie abilitanti (Enabling o General Purpose Technologies) come quelle concernenti i nuovi materiali (in particolare nanoscienze e nanosensoristica) e la produzione di energia (fotosensibilizzazione di semiconduttori, sistemi supramolecolari biomimetici per la cattura e l'immagazzinamento di luce solare). Per promuovere e valorizzare le potenzialità crescenti del *modelling* computazionale a tutti i livelli (dalla ricerca di base alle applicazioni industriali) il DPM ha istituito nel 2006 un nuovo Progetto (il N. 7) ad esso dedicato. Una delle caratteristiche peculiari del Progetto 7 è costituito dalla sua natura spiccatamente interdisciplinare, come ben esemplificato dal fatto che, in molti casi, i materiali di interesse possono essere classificati come *softmatter*, e sono quindi studiati/caratterizzati con metodologie al confine fra Chimica, Fisica e Biologia.

## Idee per una roadmap di Progetto

Per una piena valorizzazione dei metodi di *modelling* computazionale, nell'ambito della strategia generale del DPM , si ritiene necessario operare in varie direzioni, sintetizzate nei punti seguenti.

È indispensabile una ricognizione approfondita delle competenze presenti all'interno del CNR, con ovvio riferimento anzitutto al DPM, ma con ampie aperture ad altri Dipartimenti, in particolare al DMD. La collaborazione fra i ricercatori e la direzione del progetto è ritenuta di fondamentale importanza e va molto sviluppata. La creazione di un gruppo di lavoro si sta rivelando molto utile per la raccolta di idee e proposte riguardo alle strade da seguire. Essa va affiancata da incontri semestrali (primavera ed autunno) fra tutti i ricercatori coinvolti come occasioni di discussione scientifica ed organizzativa

I vari gruppi operanti devono coordinarsi efficacemente in modo da poter offrire le competenze necessarie ad affrontare le richieste di *expertise* che molto facilmente superano gli ambiti ristretti in cui operano i singoli gruppi. Questo richiede un certo sforzo sia dal punto di vista scientifico che computazionale. Dal punto di vista scientifico il problema si articola in varie fasi: 1) ideazione delle linee generali del modello che si pensa possa portare ad un'efficace simulazione del sistema investigato; 2) determinazione dei parametri usati nella formulazione del modello (via esperimento e/o calcolo; 3) *benchmarking* e validazione attraverso il confronto con i dati sperimentali; 4) test delle capacità predittive del modello. Procedere attraverso queste 4 fasi richiede spesso la stretta collaborazione fra persone in possesso di competenze diverse. È quindi necessario adoperarsi per superare gli steccati impropriamente creati dall'appartenenza ad Istituti (e Dipartimenti) diversi. Ciò è stato in parte già realizzato con commesse miste ICTP-IBB e ICCOM-IPCF. Nel futuro lo sforzo in tale direzione va ulteriormente intensificato. Per quanto riguarda l'aspetto informatico/computazionale e quello delle infrastrutture per il calcolo la discussione è rimandata ai punti che seguono.

I metodi di Chimica Computazionale e Modelling di molecole e materiali possono richiedere notevoli potenza di calcolo e memoria. È quindi necessario perseguire accordi con le grandi infrastrutture di calcolo. Il DPM ha già avviato una stretta collaborazione con il CASPUR di Roma, che ha portato all'assegnazione di un totale di 83000 ore di calcolo a gruppi operanti per lo più nel DPM (con significative presenze di gruppi di altri dipartimenti CNR per il periodo 1 Luglio-31 Dicembre 2008. In tal modo si sono offerte grandi e nuove potenzialità a tutte le unità operanti nel progetto creando le premesse per l'ulteriore sviluppo del settore. Tali iniziative vanno potenziate ed estese ad altre infrastrutture di calcolo a livello italiano (primo fra tutti il CINECA) ed europeo. Nell'impostare tali nuovi rapporti è importante presentarsi e negoziare avendo raggiunto il peso specifico necessario per far valere le esigenze peculiari del *modelling* computazionale

Contemporaneamente, per un efficace e razionale utilizzo delle varie competenze sparse sul territorio nazionale, è vitale sviluppare ed utilizzare razionalmente una griglia computazionale ad alte prestazioni, basata sui protocolli INFN e dedicata al calcolo nel campo di materiali, biosistemi, softmatter e scienze ambientali. Tale laboratorio virtuale (CNR- VILLAGE) dovrebbe essere parte integrante di un Italian Virtual Integrated Laboratori for Large-scale Applications in a Geographically distributed Environment (I-VILLAGE) comprendente anche l' INSTM- VILLAGE ed il CASPUR. Lo scopo della griglia è quello di facilitare l'accesso a tecniche di *modelling* computazionale ottimizzando ed integrando i diversi livelli di hardware, database e software per il calcolo, per l'analisi e la visualizzazione dei risultati. In una prima fase sono stati identificati come nodi CNR l'IPCF di Pisa, l'NNL (INSTM-CNR ) di Lecce ed il laboratorio integrato napoletano formato da IBB,ICTP, ICMB del DPM.. Per una efficace gestione di un tale Laboratorio virtuale si ritiene necessario coinvolgere le competenze informatiche e di gestione delle reti degli istituti informatici di Pisa (IIT ed ISTI) come fase iniziale di un auspicato maggiore coinvolgimento del Dipartimento ICT.

Il Progetto 7 del DPM intende porsi come punto di riferimento, all'interno del CNR, per chiunque decida di avvalersi delle competenze nel campo del *modelling* computazionale. Per ottenere questo scopo è necessaria un'efficace azione di promozione delle capacità raggiunte dalle tecniche di indagine "in silico" (Chimica computazionale, modellazione e simulazione di molecole, strutture supramolecolari e materiali) e delle modalità con cui tali strumenti possono essere usati con maggiori probabilità di successo. In particolare bisogna diffondere l'idea che si tratta di un potente strumento di indagine il cui uso richiede non solo competenze tecniche specifiche ma anche un'approfondita conoscenza delle problematiche a cui viene applicato, da acquisire anche attraverso la stretta collaborazione con chi usa tecniche sperimentali. Ogni occasione (incontri di gruppo, workshop, etc.) dovrebbe essere utilizzata anche a questo scopo. Una particolare cura deve essere rivolta al mondo della piccola e media industria, che spesso ignora o sottovaluta l'utilità di queste tecniche ma, talvolta, pur ritenendole potenzialmente utili, rinuncia a valersene semplicemente perché non può permettersi di avere personale dedicato a tale scopo. C'è quindi molto spazio per un efficace azione prima di promozione e poi di trasferimento tecnologico, secondo varie modalità da studiare caso per caso. Non va esclusa neanche la via del trasferimento temporaneo di personale dal CNR all'industria (se è vero che uno dei mezzi più efficaci per trasferire know-how è quello di trasferire la persona o le persone che sono in possesso di tale know-how), nelle forme appropriate e con i giusti incentivi.

## Considerazioni finali

La Direzione del Progetto, operando nel quadro di riferimento del DPM e coinvolgendo eventualmente altri Dipartimenti CNR e, se ritenuto necessario, gli altri soggetti presenti nell' I-Village, (oltre che partners industriali), deve creare le condizioni per aumentare la propria competitività nell'accedere ai fondi stanziati nell'ambito di Progetti Nazionali ed Europei nei settori strategici indicati in precedenza. In definitiva, questo offrirà la misura del successo delle linee di sviluppo delineate nel presente documento. Il potenziamento della struttura del Progetto sulla base dei punti elencati è un ingrediente indispensabile per il raggiungimento di tale obiettivo. Altri ingredienti fondamentali sono la creatività dei singoli Ricercatori operanti nel campo del *modelling* computazionale e la loro capacità di coinvolgere innanzitutto i propri Colleghi, anche attraverso l'interfaccia dei Gruppi di Lavoro creati recentemente per ogni progetto del DPM. Compito della Direzione del Progetto e del DPM deve essere quello di indirizzare tale creatività lungo i binari appropriati per massimizzare la probabilità che essa si riveli fruttifera. Come ben sappiamo, ciò presuppone innanzitutto la capacità di acquisire le risorse necessarie per sviluppare le proprie idee.