

La Didattica in 3D: come studiare le proteine al computer

GIOVANNI CASAVECCHIA*

Riassunto

La modellistica molecolare e le simulazioni al computer hanno contribuito allo studio di molti sistemi molecolari complessi. E' opportuno introdurre questi strumenti nella didattica delle scienze in una scuola media secondaria? In questo articolo viene presentata un'esperienza in classe che ha portato all'approfondimento di alcuni concetti fondanti la biochimica.

Abstract

Molecular modeling and computer simulations have contributed to the study of many complex molecular systems. Is it useful to introduce these tools in a secondary school to teach science? In the following paper we present, as an example, a lecture focused on some basic biochemistry concepts.

Introduzione.

L'interesse verso la modellizzazione molecolare è aumentato in modo considerevole. Negli ultimi anni, la disponibilità di computer sempre più potenti ha permesso di considerare sistemi molecolari via via più complessi. Le elaborazioni al computer hanno permesso di creare e diffondere una serie di programmi *user friendly*, caratterizzati cioè da interfacce grafiche facilmente utilizzabili. Questi programmi permettono agli studenti e ai docenti di fare nuovi esperimenti e di esplorare il mondo molecolare attraverso i modelli virtuali. Inoltre, in questi ultimi anni, la Rete si è arricchita di materiale "multimediale", tanto da rendere indispensabile l'utilizzo dei computer anche all'interno della normale gestione della didattica. I docenti di chimica e di scienze, così come gli studenti, possono trarre enormi vantaggi da questa evoluzione, grazie ai programmi di modellistica molecolare e all'enorme quantità di materiale disponibile in rete.

Con questo intervento, cercherò di spiegare come la modellizzazione molecolare nell'insegnamento della chimica e della biochimica nasce dalla tradizione stessa della disciplina, che ha sempre fatto uso di immagini e modelli per descrivere il mondo atomico-molecolare. Nello specifico, ho affrontato con la classe un percorso di biochimica che, altrimenti, sarebbe stato di difficile "lettura" e comprensione per gli studenti.

Percorso Didattico.

Il percorso svolto con gli studenti è legato ad alcuni argomenti di biochimica che reputo difficili da apprendere se affrontati con un approccio "frontale". Quello che più mi

interessa è chiarire ai ragazzi alcuni concetti legati alla natura molecolare degli aminoacidi e, soprattutto, alla struttura molecolare delle proteine. A grandi linee, gli argomenti che ho affrontato anche attraverso l'uso dei programmi di modellistica molecolare sono:

- Struttura Aminoacidi;
- Classificazione Aminoacidi;
- Stereochimica Aminoacidi;
- Legame Peptidico;
- Classificazione & Struttura delle Proteine.

Confesso subito che la stereochimica degli aminoacidi è un capitolo troppo complesso e, sotto certi aspetti, poco significativo per dei ragazzi che si avvicinano per la prima volta a simili tematiche.

Il Laboratorio Virtuale.

Il docente che decidesse di lavorare attraverso l'uso dei programmi di modellistica molecolare, dovrebbe cercare di strutturare ed organizzare un laboratorio caratterizzato da una macchina ogni due, massimo tre, studenti. Oltre al browser, gli elementi necessari per strutturare un ambiente di questo tipo sono:

1. **Plug-in:** programmi che visualizzano le immagini in Internet in 3D;
2. **Programmi:** applicativi che permettono la rappresentazione e, in alcuni casi, la costruzione dei modelli tridimensionali delle molecole.

In rete c'è l'imbarazzo della scelta e decidere quale programma scaricare e far utilizzare ai ragazzi può non essere semplice. Dopo un'attenta analisi ho scelto un programma che solo apparentemente può sembrare più complesso. In realtà, si è rivelato decisamente utile per il raggiungimento dei nostri (miei e dei miei studenti) obiettivi.

Perché RasMol.

RasMol è un visualizzatore molecolare che nasce alla fine del 1989 da un'idea di Roger Sayle [1], il quale sviluppa e porta a termine l'elaborazione del *software* nel 1993. RasMol è un *helper*: "questo significa che, mentre con MDL Chime potevate visualizzare le molecole all'interno del browser, per fare lo stesso con RasMol dovete prima scaricare il file.pdb relativo alla molecola sul vostro computer" [2]. Come dicevo, è possibile scegliere differenti programmi e, nel mio caso, la scelta è ricaduta su un programma che è compatibile con molti sistemi operativi d'uso comune, Windows, Apple Macintosh, GNU/Linux ed è compatibile con altri programmi di modellistica molecolare che sono in grado di creare una molecola o più molecole, che saranno lette e rappresentate da RasMol. Inoltre, l'applicativo è caratterizzato da una struttura a duplice interfaccia: una grafica ed una testuale. Questo aspetto è molto importante, infatti, nella prima interfaccia le molecole sono visualizzate e rappresentate in diversi modi e, come per MDL Chime, possono essere modificate, salvate ed espor-

* Dipartimento di Chimica Generale ed Organica applicata, Università di Torino, Liceo Vittorio di Torino.
gio.casavecchia@fastwebnet.it

tate in differenti formati (GIF, PPM, PMP), mentre, nella seconda interfaccia, chiamata *Command Line*, è data una descrizione testuale di tutto ciò che è raffigurato o che si desidera raffigurare nella prima. Un altro significativo vantaggio è la capacità che la *Command Line* ha di memorizzare il percorso fatto, che può essere sempre richiamato.

Per operare in modo corretto, è sufficiente imparare una serie di comandi (sono pochi e piuttosto semplici). In definitiva, RasMol è un programma decisamente più flessibile di MDL Chime e si può liberamente scaricare dalla rete. [3]

Le Proteine & La tradizione.

Quando si inizia a svolgere il programma di biologia, ci si trova di fronte la sezione dedicata alla trattazione delle biomolecole: molte di queste sostanze sono dei polimeri, cioè macromolecole costituite dalla successione di unità fondamentali o gruppi atomici, detti monomeri.

I libri di testo introducono l'argomento specificando che, questi sistemi molecolari sono caratterizzati da pochi elementi della tavola periodica: il carbonio, l'azoto, l'ossigeno, l'idrogeno, il fosforo e lo zolfo e, ogni volta, gli studenti tendono a rilassarsi. In realtà, le difficoltà emergono subito. Infatti, è innegabile che all'interno di questi sistemi molecolari ci siano "solo" sei elementi, ma è altrettanto vero che queste macromolecole sono strutturalmente complesse. In particolare, le proteine sono molecole la cui complessità influisce sulla trattazione teorica a tal punto che, nella maggior parte dei casi, il docente si limita ad una descrizione superficiale dell'argomento. Infatti, nella quasi totalità dei libri di testo adottati nelle scuole medie secondarie, le proteine sono presentate come: molecole compatte e tondeggianti o di forma allungata. Nel primo caso, siamo stiamo parlando delle proteine globulari (es. emoglobina); nel secondo, delle fibrose (es. cheratina). Questa prima classificazione ne include una seconda che identifica quattro livelli di struttura:

1. la **struttura primaria**: *semplice sequenza lineare degli aminoacidi*;
2. la **struttura secondaria**: *disposizione tridimensionale delle catene che compongono le proteina*;
3. la **struttura terziaria**: *ripiegamento di un'intera catena.pensate a cosa accade quando arricciate un nastro....ne risulta un groviglio irregolare, simile ad una struttura terziaria*;
4. la **struttura quaternaria**: *interazione tra due o più catene ripiegate*.

Trovandoci a dover leggere e capire una serie di definizioni di questo tipo, anche nel caso in cui si conoscesse molto bene l'argomento, le difficoltà cognitive legate al semplice apprendimento delle tematiche affrontate non sarebbero poche. Per cercare di non ridurre e semplificare troppo la spiegazione, ho cercato di favorire l'acquisizione di questi concetti anche attraverso la mediazione dei programmi di modellizzazione molecolare. L'adozione di questi strumenti e di queste metodologie didattiche ha permesso al docente ed allo studente di interagire in modo diretto e costruttivo con le immagini dei sistemi molecolari presi in considerazione.

Durante il lavoro in classe, siamo partiti dagli aminoacidi essenziali, evidenziando alcuni aspetti chimico-fisici che regolano il comportamento dei sistemi più complessi, e siamo arrivati alle proteine, definendo e chiarendo il concetto di struttura primaria, secondaria, terziaria e quaternaria.

Le Proteine & RasMol: come divertirsi.

Dove reperire il materiale: i File .pdb.

RasMol legge file di coordinate molecolari e raffigura interattivamente la molecola sulla finestra grafica. I file di entrata sono i formati: Brookhaven Protein Databank (PDB), Tripos Associates Alchemy y Sybyl Mol2, Molecular Design Limited s (MDL) Mol, Minnesota Supercomputer Centre s (MSC) XYZ (XMol), CHARMm, formato CIF e archivi mmCIF. Nel nostro specifico caso, quelli più utilizzati sono i file **PDB** (.pdb). Per ottenerli, è necessario interrogare la banca dati Protein Data Bank, accessibile all'indirizzo <http://www.rcsb.org/pdb/>, oppure ricercare il file direttamente in rete. La home page della PDB evidenzia una finestra, *Search the Archive*, predisposta alla ricerca di macromolecole con estensione .pdb. Il database, molto ricco, rallenta la visualizzazione del file, quindi, almeno in alcuni casi, è consigliabile il link: *Search Lite*. Proviamo a cercare la mioglobina (Fig.1): il database evidenzia un numero elevato di strutture! Senza entrare troppo nel dettaglio, cerchiamo di capire alcune informazioni: **101M**, indica la versione del file ed il nome della molecola; i tre simboli alla destra rappresentano i formati scaricabili; **explore** permette di accedere alla pagina SUMMARY INFORMATION. Questa pagina ricca di link ci permette di scaricare sulla nostra macchina il file di interesse: selezionare **Download Display**, quindi, nella sezione **Download the Structure File**, scaricare il file compresso (versione: Unix compressed o GNU zipped) formato .pdb.

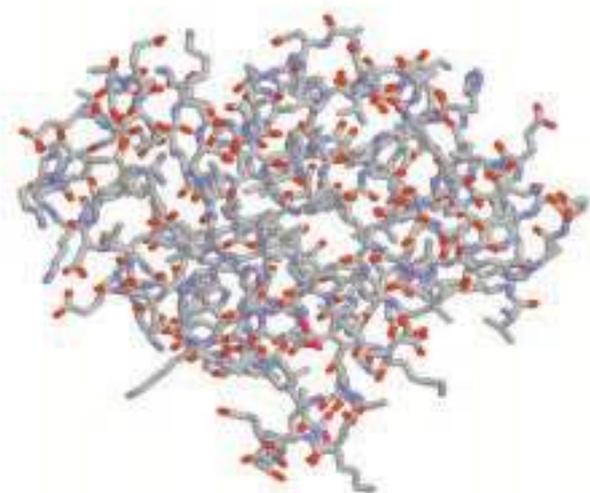


Fig1. Molecola di mioglobina in formato Sticks

Sezioniamo ed esploriamo la mioglobina. Proviamo a trarre alcune informazioni con RasMol digitando alcuni semplici comandi. E' chiaro che per operare con il software è necessario aver trovato e scaricato dalla rete la molecola. Nel nostro caso, lavoreremo con un file chiamato: **1mbn.pdb**.

Il file è apribile (quindi visualizzabile da RasMol) operando con il **Menù Principale** alla voce **File, Open: 1mbn.pdb**. Per coloro che utilizzano sistemi operativi diversi da Windows, quali ad esempio Linux, è indispensabile specificare il percorso, altrimenti il programma non trova il file richiesto.

Proviamo a trarre le prime informazioni. Digitiamo nella shell (finestra) testuale i seguenti comandi (in rosso):

· **show information**, la risposta (output) della Command Line (finestra testuale) è decisamente convincente. Il programma ci informa che siamo di fronte ad una struttura che

è caratterizzata da:

Molecule name MYOGLOBIN (FERRIC IRON - METMYOGLOBIN)

Classification OXYGEN STORAGE

Secondary Structure ... PDB Data Records

Database Code 1MBN

Number of Chains 1

Number of Groups 153 (1)

Number of Atoms 1216 (44)

Number of Bonds 1292

Number of Helices 8

Number of Strands 0

Number of Turns 3

· **show sequence**, anche in questo caso, le informazioni date sono piuttosto interessanti:

VAL1 LEU2 SER3 GLU4 GLY5 GLU6 TRP7 GLN8 LEU9 VAL10

LEU11 HIS12 VAL13 TRP14 ALA15 LYS16 VAL17 GLU18 ALA19 ASP20

VAL21 ALA22 GLY23 HIS24 GLY25 GLN26 ASP27 ILE28 LEU29 ILE30

La Command Line ci dice che la molecola che stiamo analizzando è costituita da catene caratterizzate da una precisa sequenza di aminoacidi. (La sequenza è molto lunga e per ragioni di spazio non è copiata integralmente).

Dai libri di testo, sappiamo che:

1. La mioglobina è una proteina globulare relativamente piccola che contiene una sola catena polipeptidica. Essa contiene una ferroporfirina o gruppo eme;
2. La molecola è molto compatta e nel suo interno ci sono poche molecole d'acqua;
3. Quasi tutti i gruppi -R polari degli aminoacidi sono posti sulla superficie esterna della molecola;
4. Quasi tutti i gruppi -R non polari degli aminoacidi sono posti all'interno della molecola.

Per prima cosa proviamo ad individuare la struttura primaria della molecola. Digitiamo nella Command Line i seguenti comandi:

· RasMol> **select backbone**
612 atoms selected!

Come risposta, il programma ci informa che la struttura primaria è costituita da 612 atomi. Se uno desiderasse sapere quali atomi entrano a far parte della struttura primaria, dovrebbe semplicemente digitare:

```
· RasMol> show selected
Group: VAL 1 (C) (4/7) atoms
Group: LEU 2 (C) (4/8) atoms
Group: SER 3 (H) (4/6) atoms
Group: GLU 4 (H) (4/9) atoms
Group: GLY 5 (H) (4/4) atoms
Group: GLU 6 (H) (4/9) atoms
Group: TRP 7 (H) (4/14) atoms
Group: GLN 8 (H) (4/9) atoms
Group: LEU 9 (H) (4/8) atoms
Group: VAL 10 (H) (4/7) atoms
Group: LEU 11 (H) (4/8) atoms
Group: HIS 12 (H) (4/10) atoms
Group: VAL 13 (H) (4/7) atoms
Group: TRP 14 (H) (4/14) atoms
```

(Anche in questo caso, la sequenza è molto più lunga). Le informazioni presenti ci dicono il gruppo di appartenenza, l'atomo presente ed il numero seriale.

Novembre - Dicembre 2005

A questo punto per evidenziare la struttura primaria è necessario selezionare dal Menù Principale: Display e Backbone e dovreste ottenere l'immagine di Fig.2:

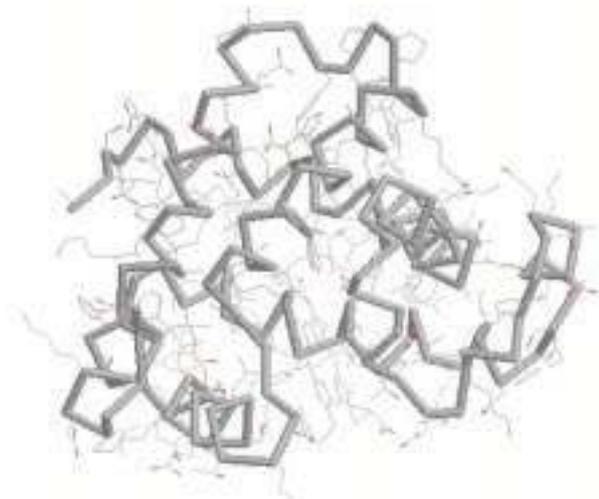


Fig.2 Struttura primaria della molecola di Mioglobina.

Arrivati a questo punto, proseguiamo l'indagine prendendo in considerazione: *La mioglobina è una proteina globulare Essa contiene una ferroporfirina o gruppo eme.* Fig.3

· RasMol> **select hetero**
44 atoms selected!

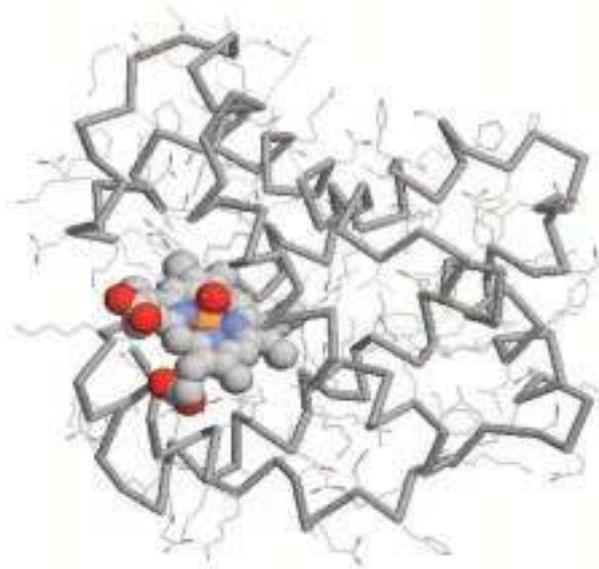


Fig 3. Molecola di Mioglobina dove sono messi in evidenza la struttura primaria ed il gruppo EME.

Per non dilungarci troppo, proviamo ad impartire una serie di comandi:

· RasMol> **select solvent**
No atoms selected!

· RasMol> **select ions**
No atoms selected!

· RasMol> **select water**
No atoms selected!

Tramite questi input, capiamo che nella nostra molecola non sono presenti molecole d'acqua. Proseguiamo l'esplorazione.

razione:

- RasMol> **select polar** (aminoacidi polari)
677 atoms selected!
- RasMol> **cpk 200**
- RasMol> **colour red**
- RasMol> **select hydrophobic** (aminoacidi idrofobici)
539 atoms selected!

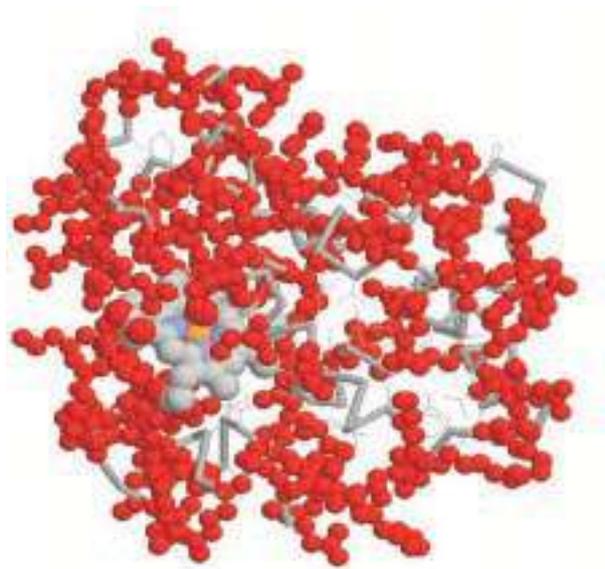


Fig.4 Aminoacidi Polari (rosso)

- RasMol> **cpk 200**
- RasMol> **colour blue**
- RasMol> **select helix**
930 atoms selected!
- RasMol> **colour yellow**

Alla fine, dovrete ottenere una sequenza di questo tipo:

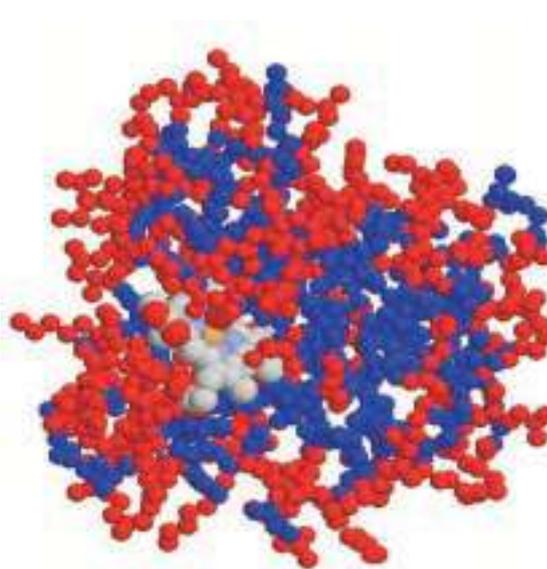


Fig.5 Aminoacidi Idrofobici (blu)

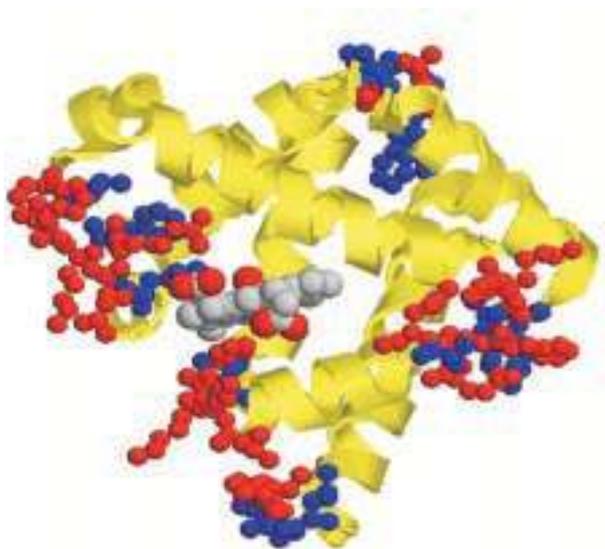


Fig.6 Catene ripiegate

Come si è dimostrato, attraverso l'acquisizione di una serie di semplici comandi, è possibile esplorare e studiare sistemi molecolari di una certa complessità.

Conclusioni.

Imparare con il computer è imparare attraverso una particolare forma di esperienza. Le macchine ci permettono di vedere e di interagire con i modelli che descrivono la realtà. Nel nostro specifico caso, il computer ci permette di arricchire le nostre conoscenze e competenze, facendoci interagire direttamente con i modelli molecolari delle strutture proteiche. Questo aspetto diventa determinante nella didattica delle scienze. Infatti, oltre alla difficoltà intrinseca della materia, bisogna cercare di valutare il livello cognitivo dello studente, in modo da suscitare in lui la curiosità (aspetto motivazionale), che è uno degli elementi trainanti l'intero percorso educativo. Nella crescita intellettuale diventa significativa la partecipazione diretta dello studente e della classe all'azione didattica. La possibilità di *sezionare* una molecola complessa come quella proteica apre orizzonti co-

noscitivi difficilmente esplorabili attraverso la semplice lettura di un libro di testo. Gli sviluppi delle tecnologie informatiche stanno aprendo nuove possibilità a questa forma di apprendimento. In particolare, l'utilizzo della finestra di dialogo (Command Line) impone allo studente un rigore conoscitivo, sia a livello disciplinare sia informatico, non necessariamente aprioristico. Nella trattazione delle discipline scientifiche, diventa difficile spiegare un qualsiasi fenomeno ricorrendo solo alla formulazione di una teoria. Nel nostro specifico caso, arrivare all'identificazione delle "strutture" è una conquista conoscitiva che dipende dalla propria capacità di interagire con il sistema molecolare. Ogni studente può arrivare all'acquisizione dei concetti fondanti la materia seguendo percorsi leggermente differenti. La Command Line memorizza tali fasi ed evidenzia, è opportuno ricordarlo, gli errori commessi! Dare significatività a questi eventi permette di sfruttare altri elementi utili alla crescita dell'individuo/studente.

132 Bibliografia.

- [1] Per avere maggiori informazioni su Roger Sayle, consultare la pagina: <http://www.umass.edu/microbio/rasmol/pershist.txt>.
- [2] Maurizio Mattioli, *Studiare con Internet Chimica*, Milano: Hoepli, 2000.
- [3] RasMol è scaricabile all'indirizzo: <http://www.umass.edu/microbio/rasmol/getras.htm>.