UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Computational Tools for Bioinformatics

2627-1-F0803Q045

Obiettivi

L'insegnamento si propone di esplorare le principali strategie computazionali utilizzate nel *molecular modeling* e nella *bioinformatica strutturale*, con un focus sulla caratterizzazione delle macromolecole biologiche. . Verranno affrontati sia gli aspetti strutturali statici sia il comportamento dinamico delle biomolecole, insieme alla simulazione delle interazioni molecolari, incluse quelle tra piccole molecole e bersagli biologici, e allo studio dei meccanismi che ne determinano la reattività. Il corso offrirà inoltre una panoramica sullo stato dell'arte dell'impiego pratico di queste metodologie nei principali ambienti di calcolo e software di modellistica molecolare attualmente in uso, includendo i più recenti approcci basati su Intelligenza Artificiale e Machine Learning per lo studio della relazione struttura-funzione e delle simulazioni molecolari.

Conoscenza e capacità di comprensione
Al termine dell'insegnamento, lo studente saprà:

- conoscere le strategie computazionali finalizzate allo studio delle relazioni struttura-funzionalità di marcomolecole biologiche
- comprendere i principali algoritmi su cui si fondano tali metodologie computazionali e le loro più moderne alternative basate sull'AI
- comprendere le differenze tra i vari metodi e tools computazionali, in termini teorici e di applicabilità, finalizzati allo studio di molecole di interesse biologico/industriale
- utilizzare in autonomia software e web-server per simulazioni di riconoscimento molecolare (proteina-ligando, proteina-peptide, proteina-proteina) e di supporto al protein design.

Capacità di applicare conoscenza e comprensione
Lo studente saprà, al termine dell'insegnamento, saper applicare le conoscenze acquisite, sapendo riconoscere, da un punto di vista pratico e teorico, potenzialità ed eventuali limiti delle metodologie computazionali trattate.

Autonomia di giudizio
Al termine di questa attività formativa, lo studente sarà in grado di scegliere l'approccio computazionale più idoneo per affrontare problematiche a carattere molecolare di rilevanza biologica o industriale. Saprà inoltre valutare con

spirito critico i risultati di simulazioni computazionali e fornirne un'interpretazione autonoma. Tali competenze saranno sviluppate attraverso esercitazioni guidate, discussioni interattive in aula sui casi studio proposti e l'analisi critica dei dati ottenuti tramite gli strumenti computazionali presentati nel materiale didattico fornito.

Abilità comunicative
Questa attività formativa consentirà allo studente di esporre in modo idoneo gli argomenti trattati e i concetti appresi con opportuno linguaggio scientifico. Lo sviluppo di tali abilità è supportato dalla preparazione e presentazione di relazioni scritte su attività svolte a lezione, basate sull'analisi e discussione di letteratura scientifica, che costituiscono parte integrante dell'esame finale.

Capacità di apprendimento
Al termine dell'insegnamento lo studente avrà gli strumenti necessari per applicare conoscenze e abilità acquisite nel trattare problematiche differenti da quelle affrontate a lezione. Disporrà inoltre delle basi sufficienti alla consultazione autonoma di riviste scientifiche riguardanti studi computazionali.

Contenuti sintetici

- Modelling computazionale della struttura di macromolecole biologiche.
- Relazione tra struttura molecolare ed energia nell'ambito della meccanica molecolare (MM).
- Cenni sulla relazione tra struttura molecolare ed energia in meccanica quantistica (QM) e tecniche ibride per lo studio della reattività di macromolecole (QM-MM).
- Ottimizzatori locali e globali, nella teoria e nella pratica.
- Fondamenti teorici del molecular docking e principali applicazioni: docking proteina-ligando, proteina-peptide, proteina-proteina e covalente.
- Approfondimento di tecniche di simulazione (dinamica molecolare, simulazioni Coarse Grained).
- Cenni di protein design computazionale.
- Applicazioni di intelligenza artificiale (AI) e machine learning (ML) a supporto delle simulazioni, della predizione di struttura di proteine e dei loro complessi e del protein design (*de novo* e rational).

Programma esteso

Fondamenti di bioinformatica strutturale e modellistica molecolare:

- Intriduzione al molecular modeling, contesto ed ambiti di applicazione;
- Concetto di computazionabilità e requisiti di base per il molecular modeling;
- Valutazione della bontà di una struttura molecolare e la sua preparazione al modelling;
- Approcci AI-based per la predizione della struttura proteica.

Struttura molecolare ed energia:

- Relazione teorica tra struttura e energia: gradi di libertà, coordinate interne (Z-matrix), simmetria molecolare;
- Analisi topologica della superficie di energia potenziale (PES);
- L'energia in meccanica molecolare: componenti di un force field (FF), confronto tra diversi FF (anche Coarse Grained) e le loro applicazioni;
- L'energia in meccanica quantistica (cenni) e panoramica sulle tecniche ibride (QM/MM, ONIOM) per lo studio della reattività di sistemi proteici (con applicazioni in catalisi e drug design);
- TUTORIAL:
  -Esempio di calcolo QM e QM/MM.

Algoritmi di ottimizzazione:

- Ottimizzatori locali: algoritmi di ordine zero, primo e secondo (simplex, grid search, univariate search, steepest descent, gradienti coniugati, Newton-Raphson (NR) e quasi-NR); algoritmi di ricerca di stati di transizione per la stima delle costanti di velocità di processi chimico-biologici;
- Ottimizzatori globali: algoritmi deterministici e stocastici a confronto (es. multi-start/Global Optimization Algorithm, Monte Carlo Metropolis, Simulated Annealing, algoritmi evolutivi);
- ESERCITAZIONI:
  -Utilizzo di diversi algoritmi per la ricerca di minimo locale, in modo da razionalizzarne le performance sulla base della teoria;
  -Utilizzo di algoritmi stocastici per la ricerca di minimo globale e per il ripiegamento di un piccolo peptide.

Molecular Docking

- Algoritmi di Docking Molecolare: search algorithms (stocastici e sistematici) e scoring functions implementate nei software più popolari;
- La griglia di energia potenziale e concetti chiave nel pre-processing;
- Il docking covalente e il docking macromolecola-macromolecola e le sue applicazioni;
- Survey sui software e web server più popolari ed indicazioni pratiche su utilizzo dei programmi;
- ML nel docking ed AI per drug design;
- ESERCITAZIONI:
  -Docking semiflessibile con diverse teorie a confronto, per comprendere limiti e potenzialità della tecnica;
  -Docking covalente per diverse applicazioni (drug design/industriale);
  -Docking proteina-proteina e mutagenesi in silico usando software in locale e web-servers.

La Dinamica Molecolare

- Cenni di termodinamica statistica e insiemi canonici;
- Algoritmi di integrazione, termostati e barostati;
- Analisi delle traiettorie.
- Applicazioni della MD
- ESERCITAZIONE/TUTORIAL:
  -Simulazione di MD con GROMACS.

Cenni di Protein Design:

- Concetti fondamentali del protein design *de novo*: approcci physics-based e AI-based;
- Strumenti e webserver disponibili per il protein design;
- Esempi di applicazioni.

Prerequisiti

Prerequisiti: Non sono strettamente necessarie conoscenze specifiche. **Tutta la teoria necessaria ad affrontare le varie tematiche verrà ripresa da zero**. E' auspicabile l'**interesse** e la **curiosità** nel voler approfondire in silico i dettagli molecolari alla base dei fenomeni chimico-bilogici.

Propedeuticità. Nessuna

Modalità didattica

- 15 lezioni da due ore svolte in modalità didattica prevalentemente erogativa in presenza;
- 2 lezioni da due ore svolte in modalità didattica interattiva in presenza;
- 6 esercitazioni da due ore svolte in modalità didattica interattiva in presenza.

L'insegnamento si svolgerà in lingua italiana.

Materiale didattico

Saranno disponibili sulla piattaforma e-learning dell'insegnamento:

- Videoregistrazioni di tutte le lezioni
- Videotutorial delle esercitazioni pratiche
- Slides
- Dispense di supporto
- Bibliografia consigliata, tra cui una selezione di articoli scientifici e monografie a complemento/approfondimento delle tematiche viste a lezione.
- Link a seminari utili come approfondimento.

Libri di testo per (eventuale) supporto:
“Bioinformatica” , Stefano Pascarella, Alessandro Paiardini;
"Molecular Modelling", Andrew Leach.
"Introduction to Computational Chemistry", Frank Jensen (chapter 2).

Periodo di erogazione dell'insegnamento

Secondo semestre

Modalità di verifica del profitto e valutazione

L'esame consiste nella preparazione di un elaborato su alcune delle esercitazioni / dimostrazioni pratiche svolte a lezione. Tale elaborato verrà poi discusso in modalità orale.

Orario di ricevimento

Ricevimento: su appuntamento tramite richiesta via email al docente

Sustainable Development Goals

CONSUMO E PRODUZIONE RESPONSABILI