

COMPITI E SUDDIVISIONE FONDI TRA LE UNITÀ DI RICERCA
prot. 2005025257

| | |
|---------------------------------|---|
| Coordinatore Scientifico | Attilio STELLA |
| Ateneo | Università degli Studi di PADOVA |
| Titolo della Ricerca | Statica e dinamica di polimeri soggetti a vincoli topologici. |
| Finanziamento assegnato | Euro 65.000 |
| Durata | 24 Mesi |

Obiettivo della Ricerca

I tre gruppi proponenti questo progetto hanno esperienze nell'applicazione dei metodi della meccanica statistica alla soluzione di vari problemi riguardanti polimeri e biopolimeri. Queste esperienze hanno anche un alto grado di complementarità. Un primo obiettivo di questo progetto è connesso con la scelta di guardare in una prospettiva unificata ai diversi modi in cui i concetti e le proprietà topologiche possono entrare nella descrizione fisica di macromolecole. I due casi principali sono la complessità topologica associata alle conformazioni spaziali delle catene, e quella determinata dalla struttura dei contatti delle loro specifiche sequenze di monomeri. Come spiegato sotto, i problemi posti da queste due diverse nozioni topologiche sono, per molti aspetti, simili. Il successo del presente progetto si fonda per larga misura sull'abilità dei proponenti di beneficiare del potenziale di conoscenza incrociata sussistente per queste due categorie di problemi. In alcuni casi affronteremo anche problemi appartenenti ad entrambe le categorie.

I nostri obiettivi principali possono essere riassunti come segue:

- 1) Determinare la frequenza e la complessità dei nodi in lunghe catene polimeriche all'equilibrio, specialmente nei regimi compatti.*
- 2) Discutere le proprietà di localizzazione dei nodi negli stessi anelli e dimostrare il fenomeno della delocalizzazione nel regime compatto.*
- 3) Impostare approcci alle proprietà dinamiche dei nodi nei polimeri, come i tempi di persistenza in catene aperte.*
- 4) Eseguire una ricerca sistematica di proteine con stato nativo annodato e determinare il ruolo giocato dai loro nodi nel processo di folding.*
- 5) Dare una giustificazione microscopica a modelli di folding di proteine basati sulla diffusione e sulla coalescenza di elementi delle strutture secondarie.*
- 6) Combinare il modello di Go con l'approccio a foldoni introducendo dei vincoli termodinamici e di mappa di contatto. Lo scopo è di complementare la natura più microscopica del modello di Go con la specificità di sequenza dell'approccio a foldoni del ripiegamento delle proteine.*
- 7) Esplorare le possibili interconnessioni tra misure di complessità configurazionale, misure di frustrazione, misure di complessità di ripiegamento e altri indicatori di complessità strutturale presenti in letteratura.*
- 8) Descrivere la termodinamica all'equilibrio, e il "free-energy landscape", dei modelli a la' Go per l'RNA che tengano conto dell'energia di stacking, entropia di anelli e pseudonodi, nell'ambito dell'approccio sviluppato per la soluzione esatta del modello di folding di Munoz-Eaton.*
- 9) Descrivere l'influsso reciproco tra l'energia di stacking e il volume escluso nella denaturazione dell'RNA, mediante le soluzioni esatte sui reticoli di Bethe e Husimi.*
- 10) Sviluppare un modello Munoz-Eaton per il ripiegamento di proteine descritte in termini di foldoni, e studiare la sua dinamica mediante l'approssimazione di equilibrio locale.*
- 11) Concepire descrizioni più accurate degli eventi elementari di ripiegamento attraverso l'utilizzo di tecniche di reti neurali. Questo è per esempio il caso di transizioni elica-coil che sono attualmente descritte da una approssimazione cruda, di tipo adiabatico, nel modello FDC. Inoltre, modelli tipo Munoz-eaton per la formazione di beta-hairpin saranno utilizzati per descrivere la formazione di foglietti beta e per generalizzare il modello FDC a proteine che contengano anche strutture tipo beta.*

Innovazione rispetto allo stato dell'arte nel campo

La topologia pone domande fondamentali nella statistica dei polimeri, ispirate dalla necessità di una comprensione più profonda delle proprietà strutturali, dinamiche e funzionali delle biomolecole. Esempio importante è il ruolo centrale che è stato ed è tuttora giocato dalla topologia nella conoscenza del DNA circolare, del supercoiling, e degli enzimi topoisomerasi. Nozioni basilari di topologia sono anche quelle che si possono arguire dalla struttura dei contatti (contact map) dello stato nativo. Tali strutture mostrano la possibile formazione di strutture secondarie come alfa-eliche, forcine e pseudonodi, che sono elementi chiave per la descrizione della dinamica del ripiegamento di proteine e RNA.

Il presente progetto affronta problemi di frontiera nella statistica e dinamica conformazionale di polimeri e biopolimeri. In questi problemi le proprietà topologiche rappresentano o l'obiettivo finale o un importante strumento di investigazione. Un denominatore comune è dato dal fatto che la topologia entra nel problema sotto forma di vincoli tipicamente difficili da imporre. Con i metodi della meccanica statistica si studieranno modelli come i poligoni autoevitantesi su reticolo, le catene di sfere, o gli insiemi di elementi di strutture secondarie che diffondono e coalescono come particelle. Tali modelli, seppur relativamente semplici, saranno adeguati a rappresentare le caratteristiche essenziali del sistema, come il volume escluso o la struttura primaria. Gli strumenti principali di lavoro saranno i metodi Monte Carlo e metodi analitici quali l'approssimazione variazionale o le equazioni di Fokker-Planck.

Si mirerà ad una caratterizzazione completa della complessità topologica e delle proprietà di localizzazione dei nodi per lunghi

polimeri circolari nei regimi di equilibrio esteso, adsorbito, e collassato. Cio' sara' prezioso per problemi come l'interpretazione di recenti esperimenti sul DNA in capsidi virali, la validita' delle previsioni di teoria dei campi per lo scaling dei polimeri circolari con nodo fissato, o l'identificazione di nodi nella struttura nativa delle proteine.

Si determinera' lo scaling del tempo di persistenza dei nodi nei polimeri lineari, e gli effetti della topologia nel ripiegamento delle proteine. Nel primo caso ci si occupera' prima di nodi piatti, che possono essere un modello prezioso per l'interpretazione di esperimenti su catene granulari vibrante. Per le proteine useremo modelli di Go per identificare il percorso del folding e i tempi di formazione dei nodi. Modelli di Go saranno usati anche nell'approccio alla dinamica delle proteine basato sui foldoni, cioe' elementi delle strutture secondarie che si formano all'inizio, come precursori, e controllano l'intero processo del folding. Lo scopo e' di ottenere una giustificazione microscopica della descrizione a diffusione-collisione della dinamica dei foldoni implementando le informazioni sui foldoni in modelli di Go che tengano traccia della struttura a catena della proteina. Questo ci consentira' anche di vedere i possibili effetti dei nodi nel folding. Si cerchera' inoltre di generalizzare l'approccio basato sui foldoni, includendo, ad es., la possibilita' di formazione di foglietti beta e considerando la risposta a campi elettrici.

Per l'RNA considereremo gli pseudonodi, che sono forme topologiche piu' semplici rispetto ai nodi, e possono essere identificate gia' dalla mappa dei contatti. Affronteremo la termodinamica delle lunghe catene di RNA e la dinamica del loro ripiegamento con metodi e modelli enfaticanti a volte la descrizione tridimensionale del polimero, altre volte la necessita' di avere di una sequenza specifica. Si spera in tal modo di elucidare la natura della transizione, da fase denaturata a quella compatta, dell'RNA, e di ottenere una descrizione del suo folding, identificando il ruolo degli pseudonodi. Nel secondo caso useremo formulazioni come il modello di Munoz-Eaton gia' proposto per le proteine.

Criteri di verificabilità

Essendo un progetto di carattere esclusivamente teorico la verificabilità del raggiungimento dei vari obiettivi dovra' basarsi su pubblicazioni su riviste internazionali (con referee) e su comunicazioni a congressi con le quali le unita' presenteranno i vari risultati ottenuti.

Elenco delle Unità di Ricerca

| | |
|---------------------------------|----------------------------------|
| Sede dell'Unità | Università degli Studi di PADOVA |
| Responsabile Scientifico | Attilio STELLA |
| Finanziamento assegnato | Euro 28.000 |

Compito dell'Unità

In questo progetto l'Unita' di Padova fara' uso della sua esperienza nei campi della statistica conformazionale dei polimeri e biopolimeri, nei metodi di simulazione Monte Carlo e nell'analisi di fenomeni che presentano invarianza di scala. Questa esperienza sara' ovviamente condivisa con le altre Unità allo scopo di ottenere parte degli obiettivi prefissati in questo progetto.

In particolare l'Unita' di Padova sviluppera' metodi di simulazione e di analisi statistica allo scopo di accertare l'esistenza e i tipi di nodi in vari modelli di polimeri nonché di caratterizzarne la loro complessita' ed il grado di localizzazione.

Si costruirà un data base per le proteine globulari che contengono nodi e cio' sara' anche un punto di partenza per lo studio del ruolo che possono avere i nodi nella dinamica di ripiegamento di queste proteine. A questo proposito si utilizzerà un modello a la Go in collaborazione con l'Unita' di Camerino che investighera' il ruolo dei foldoni nella formazione dei nodi.

Si costruirà inoltre un algoritmo dinamico stocastico tipo Monte Carlo per studiare fenomeni come il tempo di persistenza dei nodi in polimeri lineari, partendo da modelli semplificati come i nodi piatti.

Con l'Unita' di Camerino si proverà a implementare l'informazione relativa ai foldoni in modelli a la Go per dare una caratterizzazione piu' microscopica al modello FDC.

Con l'unita' di Torino si utilizzeranno versioni opportune del modello di Munoz eaton per studiare, con metodi approssimati, le strutture secondarie e terziarie (pseudonodi) di modelli di RNA.

| | |
|---------------------------------|-----------------------|
| Sede dell'Unità | Politecnico di TORINO |
| Responsabile Scientifico | Alessandro PELIZZOLA |
| Finanziamento assegnato | Euro 24.000 |

Compito dell'Unità

Riteniamo che l'unità di ricerca del Politecnico di Torino possa contribuire a questo progetto per mezzo delle sue competenze sui modelli alla Munoz-Eaton e sulle approssimazioni semi-analitiche, generalizzazioni della teoria di campo medio, in meccanica statistica.

In particolare, l'approccio sviluppato per risolvere esattamente la termodinamica del modello di Munoz-Eaton per il ripiegamento delle proteine, potrà essere

- applicato alla descrizione del ripiegamento delle proteine in termini di foldoni, sviluppata dall'unità di ricerca di Camerino;

- esteso alla dinamica dello stesso modello, realizzando approssimazioni per la dinamica che abbiano come stato di equilibrio quello ottenuto dalla soluzione esatta della termodinamica;

- applicato a modelli per la struttura secondaria e terziaria dell'RNA che tengono conto dell'energia di stacking, dell'entropia di loop e degli pseudonodi (in collaborazione con l'unità di ricerca di Padova); in questo caso non ci aspettiamo di ottenere soluzioni esatte, ma solamente delle buone approssimazioni, per mezzo del metodo variazionale a cluster.

In secondo luogo, le approssimazioni di Bethe e Husimi, equivalenti a soluzioni esatte su grafi e ipergrafi aleatori, permetteranno di risolvere modelli semplificati, con interazioni indipendenti dalla sequenza, per la denaturazione dell'RNA che includono sia l'energia di stacking che, almeno localmente, effetti di volume escluso.

| | |
|---------------------------------|------------------------------------|
| Sede dell'Unità | Università degli Studi di CAMERINO |
| Responsabile Scientifico | Mario COMPIANI |
| Finanziamento assegnato | Euro 13.000 |

Compito dell'Unità

Il ruolo dell'Unità di Camerino nel contesto del presente progetto è quello di fornire competenze e modelli relativi a proteine reali (modello FDC), e predizioni comparabili con dati sperimentali. Tali competenze vanno ad integrare quelle delle altre Unità nella convinzione che approcci di tipo polymer physics possano offrire strumenti di analisi utili a potenziare e integrare la modellistica FDC dell'Unità di Camerino. Questi obiettivi generali dell'Unità di Camerino si articolano come segue:

(1) Sinergia con l'Unità di PADOVA.

Generalizzazione del modello FDC per includere potenziali statistici per modellizzare l'interazione tra foldoni ed includere forze esterne (ad es. elettrostatiche), non contemplate nella presente versione del modello.

(2) Sinergia con: Unità di PADOVA (modelli di Go e proteine annodate) e Unità di TORINO (modelli alla Muñoz-Eaton (ME)).

In questa fase si prevede l'applicazione di modelli di Go modificati a proteine all-a. I modelli saranno corredati con informazione sulla topologia dei foldoni ed un bias sulle energie di contatto foldone-foldone. Successive simulazioni di dinamica MC stocastica consentiranno poi di sondare la cinetica di folding e il pathway relativo. Lo scopo di questa ibridazione è quello di coniugare la specificità a sequenza nelle zone foldoniche, tipica del modello FDC, con la migliore modellizzazione del resto della catena che è propria dei modelli di Go. L'analisi verrà integrata da studi termodinamici usando theta-point models alla Flory-Huggins sempre arricchiti di informazione relativa ai foldoni e alle relative strutture. L'Unità di Camerino indagherà poi l'eventuale ruolo dei foldoni nella formazione dei nodi in molecole di un data base di proteine annodate che verranno identificate con gli algoritmi dell'Unità di Padova.

Infine si propone lo studio di un modello ME in cui l'informazione topologica e termodinamica relativa ai foldoni verrà utilizzata per semplificare e rendere più sequence-specific il modello ME dell'Unità di Torino.

(3) Sinergia con l'Unità di TORINO (modelli ME).

Il ruolo assegnato ai foldoni come siti di nucleazione per le eliche foldoniche, suggerisce la modellizzazione stocastica della transizione elica-coil usando un modello stocastico che rappresenta la crescita dell'elica foldonica come moto su una superficie di energia libera il cui profilo verrebbe desunto dalle uscite di una rete neurale. L'Unità di Camerino si prefigge anche di studiare l'integrazione del modello ME per la formazione di beta-hairpin nel modello FDC per estendere quest'ultimo a proteine con strutture beta.

(4) Sinergia con l'Unità di PADOVA.

Il modello FDC e altri modelli di letteratura enfatizzano il ruolo di interazioni tra residui distanti in sequenza (effetti entropici) nel determinare le caratteristiche e la velocità del folding. Tali interazioni sono descritte da indici (RCO, ECO etc.) che correlano con il folding rate e sono interpretabili come misure di complessità delle strutture native.

L'Unità di Camerino ha recentemente quantificato la cooperatività del meccanismo di folding attraverso un indice di modularità basato sul modello FDC. L'Unità di Camerino è allora interessata a studiare effetti entropici da chiusura di loop in collaborazione con l'Unità di Padova, e a studiare la relazione tra misure di complessità delle strutture native e del folding quali l'indice di modularità, gli indici RCO, ECO etc. e il formalismo basato su grafi proposto dall'Unità di Padova.