



SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Magistrale in Fisica

ID number

Advisor

Co-Advisor

Academic Year

Sapienza University of Rome

© . All rights reserved

This thesis has been typeset by \LaTeX and the Sapthesis class.

Author's email:

Contents

Introduction

Le proteine rappresentano una classe di biomolecole essenziali alla vita, costituendo la base funzionale e strutturale di tutti gli organismi viventi. Composte da catene di amminoacidi connesse tramite legami peptidici, esse si distinguono per la loro capacità di assumere una straordinaria varietà di conformazioni tridimensionali. Questa caratteristica conferisce alle proteine una flessibilità funzionale senza precedenti, permettendo loro di partecipare a una vasta gamma di processi biologici.

Tra le principali funzioni delle proteine si annoverano:

- **Catalisi enzimatica:** Le proteine enzimatiche accelerano le reazioni chimiche, riducendo l'energia di attivazione e regolando i processi metabolici cellulari.
- **Ruoli strutturali:** Alcune proteine, come il collagene o la cheratina, forniscono supporto meccanico e integrità strutturale ai tessuti.
- **Trasporto:** Molecole come l'emoglobina facilitano il trasporto di ossigeno, mentre altre proteine trasportano nutrienti e ioni attraverso le membrane cellulari.
- **Regolazione e segnalazione:** Le proteine svolgono ruoli chiave nella trasduzione dei segnali cellulari e nella regolazione dell'espressione genica.

Le proteine, dunque, sono sia i costruttori sia i regolatori della vita, operando con precisione nei vari contesti fisiologici.

0.1 Sintesi Proteica

La sintesi proteica è il processo attraverso il quale le cellule costruiscono proteine a partire dall'informazione genetica contenuta nel DNA. Questo meccanismo fondamentale è essenziale per la vita, in quanto le proteine svolgono una varietà di funzioni chiave all'interno degli organismi viventi, tra cui ruoli strutturali, catalitici, di segnalazione e trasporto. La sintesi proteica si suddivide in due fasi principali: *trascrizione* e *traduzione*, ognuna delle quali rappresenta un passo critico nel trasferimento dell'informazione genetica dal DNA alla proteina funzionale.

0.1.1 Fase 1: La Trascrizione

La trascrizione è il primo passo del processo, durante il quale l'informazione genetica contenuta nel DNA viene copiata in una molecola di RNA messaggero (*mRNA*).

Questo processo avviene nel nucleo delle cellule eucariotiche (o nel citoplasma delle cellule procariotiche) e può essere suddiviso nelle seguenti fasi:

1. **Inizio:** Un complesso enzimatico, chiamato RNA polimerasi, si lega a una regione specifica del DNA, detta *promotore*, situata all'inizio del gene da trascrivere.
2. **Allungamento:** L'RNA polimerasi "legge" la sequenza di nucleotidi del DNA stampo e sintetizza una molecola di mRNA complementare, utilizzando basi azotate specifiche: Adenina (A), Uracile (U, al posto della Timina presente nel DNA), Guanina (G) e Citosina (C).
3. **Terminazione:** Quando l'RNA polimerasi raggiunge una sequenza del DNA chiamata *sequenza di terminazione*, il processo si arresta e l'mRNA è rilasciato.

L'mRNA sintetizzato rappresenta una copia temporanea dell'informazione genetica, pronta per essere utilizzata durante la fase successiva.

0.1.2 Fase 2: La Traduzione

La traduzione è il processo attraverso il quale l'mRNA viene decodificato per costruire una proteina. Questo avviene nei ribosomi, complessi molecolari presenti nel citoplasma o associati al reticolo endoplasmatico rugoso. Le principali fasi della traduzione sono:

1. **Inizio:** La traduzione inizia quando il ribosoma si lega a una regione specifica dell'mRNA, individuando il primo codone di inizio (AUG), che codifica per l'amminoacido metionina.
2. **Allungamento:** I *tRNA* (RNA di trasporto) trasportano specifici amminoacidi al ribosoma. Ogni tRNA ha un anticodone complementare al codone dell'mRNA, garantendo che l'amminoacido corretto venga aggiunto alla catena polipeptidica in costruzione. Gli amminoacidi vengono uniti tramite legami peptidici.
3. **Terminazione:** Quando il ribosoma raggiunge uno dei codoni di stop (UAA, UAG, UGA), la sintesi si arresta, e la proteina completa viene rilasciata.

0.1.3 Il Codice Genetico

Il codice genetico è il sistema di corrispondenza tra le triplette di basi nucleotidiche (codoni) dell'mRNA e gli amminoacidi. Esso è:

- **Universale:** Il codice genetico è quasi identico in tutti gli organismi viventi.
- **Ridondante:** Molti amminoacidi sono codificati da più di un codone.
- **Non ambiguo:** Ogni codone codifica per un unico amminoacido o per un segnale di stop.

0.1.4 Importanza della Sintesi Proteica

La sintesi proteica è cruciale per la vita. Attraverso questo processo, le cellule producono enzimi, ormoni, proteine strutturali e altre molecole che regolano o facilitano tutte le funzioni vitali. L'efficienza e la precisione della sintesi proteica sono garantite da un sistema di controllo altamente evoluto, che include la correzione degli errori durante la trascrizione e la traduzione, nonché l'azione di molecole accessorie come i chaperoni, che aiutano le proteine a ripiegarsi nella loro struttura tridimensionale funzionale.

0.1.5 Conclusione

La sintesi proteica è uno dei processi biologici più complessi e affascinanti, che dimostra la straordinaria capacità delle cellule di convertire l'informazione genetica in molecole funzionali. Questo meccanismo non solo garantisce il mantenimento della vita, ma rappresenta anche un punto centrale per molte applicazioni in biotecnologia e medicina, dalla produzione di farmaci alla terapia genica.

Conformazioni proteiche e modi funzionali

La struttura tridimensionale di una proteina è determinata dalla sequenza di amminoacidi (struttura primaria) e dalle interazioni chimiche che stabilizzano la struttura secondaria, terziaria e, in molti casi, quaternaria. Queste interazioni portano alla formazione di strutture funzionali in grado di riconoscere substrati specifici o interagire con altre molecole.

Un concetto fondamentale nello studio delle proteine è quello dei *modi funzionali*, ovvero le configurazioni dinamiche che una proteina può assumere per adattarsi ai cambiamenti ambientali o alle interazioni con partner molecolari. Tali cambiamenti strutturali sono spesso descritti come un insieme di microstati energetici, ciascuno con una probabilità associata.

Distribuzioni di probabilità nelle proteine

Le fluttuazioni strutturali delle proteine possono essere studiate in termini di distribuzioni di probabilità (Probability Density Functions, PDF). Una PDF descrive la probabilità di trovare una proteina in un dato stato conformazionale o con un certo valore di energia. Le proteine non sono statiche, ma vivono in un continuo dinamico tra diversi stati, ciascuno caratterizzato da una differente energia libera.

L'analisi delle PDF trova applicazione in numerosi contesti, tra cui:

- Studio della stabilità proteica e delle transizioni conformazionali.
- Descrizione dell'affinità di legame tra proteina e ligando.
- Modellizzazione dei processi di ripiegamento proteico (*protein folding*).

Allostericità: una proprietà emergente delle proteine

Un fenomeno straordinario legato alla dinamicità proteica è l'*allostericità*, ovvero la capacità di una proteina di trasmettere un segnale o un cambiamento conformazionale da un sito all'altro. Questo meccanismo fondamentale consente a molte proteine di svolgere le loro funzioni biologiche in modo altamente regolato, rappresentando una chiave di volta per la comprensione della biochimica cellulare e per lo sviluppo di nuovi farmaci. L'allostericità gioca un ruolo cruciale nella regolazione dell'attività enzimatica, nella trasduzione del segnale e nella comunicazione intracellulare, fungendo da "interruttore molecolare" che risponde a stimoli ambientali e intracellulari.

Definizione e principi fondamentali dell'allostericità

L'allostericità, dal greco *allos* (altro) e *stereos* (struttura), si riferisce a un fenomeno in cui l'interazione di una molecola (effettore) con un sito specifico di una proteina, noto come *sito allosterico*, provoca una modifica conformazionale che influenza l'attività funzionale di un altro sito, solitamente il *sito attivo*. Questo processo non avviene attraverso interazioni dirette tra i due siti, ma mediante cambiamenti nella rete di interazioni intramolecolari che regolano la struttura e la dinamica della proteina.

Dal punto di vista termodinamico, l'allostericità può essere descritta come una modulazione della distribuzione dei microstati energetici di una proteina. In altre parole, l'interazione con un effettore allosterico altera l'insieme degli stati conformazionali disponibili, facilitando o inibendo l'accesso a configurazioni funzionali.

Matematicamente, il fenomeno può essere rappresentato come una variazione nella popolazione dei microstati $\{s_i\}$ descritta da una distribuzione di Boltzmann ponderata:

$$P(s_i) = \frac{e^{-\Delta G(s_i)/k_B T}}{\sum_j e^{-\Delta G(s_j)/k_B T}}$$

dove $\Delta G(s_i)$ rappresenta l'energia libera associata al microstato s_i , k_B è la costante di Boltzmann e T è la temperatura assoluta.

L'effetto allosterico può essere quantificato anche considerando il cambiamento dell'energia libera associata all'interazione con l'effettore. Se si considera una proteina con due stati principali, R (rilassato) e T (teso), l'equilibrio allosterico può essere espresso in termini di costante di equilibrio:

$$L = \frac{[T]}{[R]}$$

dove $[T]$ e $[R]$ rappresentano le concentrazioni relative dei due stati conformazionali. La presenza di un effettore modifica L , stabilizzando uno stato rispetto all'altro.

Questi modelli forniscono una base quantitativa per comprendere come le interazioni allosteriche modulino le funzioni biologiche delle proteine, come nel caso dell'emoglobina, che sfrutta l'allostericità per modulare il trasporto dell'ossigeno.

Allostericità e dinamica proteica

La moderna comprensione dell'allostericità si basa sempre più sull'idea che il fenomeno sia intrinsecamente legato alla dinamicità proteica. Le proteine non sono strutture rigide, ma ensemble di stati conformazionali che fluttuano su scale temporali e spaziali. Gli effettori allosterici alterano questa rete dinamica, rimodellando i percorsi energetici che collegano i vari stati funzionali.



Figure 0.1. Rappresentazione schematica dell'allostericità: l'interazione con un effettore allosterico (sito A) induce un cambiamento conformazionale che influenza un sito distante (sito B).

Questa visione dinamica si integra con approcci computazionali, come la dinamica molecolare e le reti di interazioni residue, che permettono di identificare percorsi specifici attraverso i quali l'informazione allosterica si propaga all'interno della proteina.

Esempi biologici di allostericità

L'emoglobina, una proteina presente nei globuli rossi, è un classico esempio di allostericità.

Essa trasporta ossigeno dai polmoni ai tessuti e il suo funzionamento dipende da cambiamenti conformazionali che influenzano la capacità di legare l'ossigeno.

L'emoglobina può assumere due stati principali:

- ****Stato T (teso)****: con bassa affinità per l'ossigeno.
- ****Stato R (rilassato)****: con alta affinità per l'ossigeno.

Quando una molecola di ossigeno si lega, l'intera struttura della proteina cambia, passando gradualmente dallo stato T allo stato R.

Questo fenomeno, chiamato ****cooperatività****, rende più facile il legame delle successive molecole di ossigeno.

Nei polmoni, dove l'ossigeno è abbondante, l'emoglobina assume lo stato R per caricarsi.

Nei tessuti, invece, fattori come l'anidride carbonica e i protoni stabilizzano lo stato T, favorendo il rilascio dell'ossigeno.

L'allostericità consente all'emoglobina di adattarsi dinamicamente alle necessità dell'organismo, catturando e rilasciando ossigeno in modo efficiente.

- **Emoglobina**: Uno degli esempi più noti, l'allostericità nell'emoglobina consente il rilascio cooperativo dell'ossigeno ai tessuti. Gli ioni H^+ e il 2,3-bisfosfoglicerato agiscono come effettori negativi, stabilizzando lo stato T e favorendo il rilascio dell'ossigeno.
- **Recettori accoppiati a proteine G (GPCR)**: Questi recettori trasducono segnali extracellulari attraverso meccanismi allosterici che coinvolgono cambiamenti conformazionali tra il dominio extracellulare e quello intracellulare.
- **Enzimi regolatori**: Molti enzimi, come l'aspartato transcarbamilasi, sono regolati allostericamente per coordinare le vie metaboliche in risposta ai bisogni cellulari.

Implicazioni biotecnologiche e farmacologiche

L'allostericità rappresenta un obiettivo interessante per lo sviluppo di nuovi farmaci. Gli *allosteric modulators* offrono diversi vantaggi rispetto agli inibitori o agonisti tradizionali:

- **Selettività**: Gli effettori allosterici interagiscono con siti diversi da quelli conservati, riducendo il rischio di effetti collaterali.
- **Regolazione fine**: Possono modulare l'attività proteica in modo più sottile, senza bloccarla completamente.
- **Riduzione della resistenza**: Poiché il sito allosterico è meno soggetto a pressioni evolutive rispetto al sito attivo, è meno probabile che emergano mutazioni resistenti.

Esempi di farmaci allosterici già in uso includono modulatori per recettori GPCR, come i farmaci per il trattamento della schizofrenia e dell'asma.

Conclusioni

L'allostericità è una proprietà emergente delle proteine che illustra la complessità e l'eleganza della regolazione molecolare nei sistemi biologici. Grazie alla combinazione

di approcci sperimentali e computazionali, stiamo iniziando a comprendere meglio i meccanismi alla base di questo fenomeno, con promettenti applicazioni in campo medico e biotecnologico. Continuare a esplorare l'allostericità non solo ci aiuterà a rispondere a domande fondamentali sulla biologia molecolare, ma aprirà anche nuove strade per l'innovazione terapeutica.

La proteina PDZ: una panoramica

La proteina PDZ rappresenta una delle classi più studiate di domini proteici modulari, essenziali per il mantenimento dell'organizzazione spaziale e funzionale nelle cellule. Il nome "PDZ" deriva dai primi tre membri della famiglia in cui è stato identificato: *Post-synaptic density protein 95 (PSD-95)*, *Discs large tumor suppressor (Dlg1)* e *Zona occludens-1 (ZO-1)*. I domini PDZ sono fondamentali per la formazione di complessi multiproteici e per l'ancoraggio di proteine bersaglio a regioni specifiche della membrana cellulare.

Struttura e caratteristiche principali del dominio PDZ

Il dominio PDZ è composto da circa 80-90 residui amminoacidici, che si organizzano in una struttura a barile β (β -sandwich) formata da sei filamenti β alternati e due α -eliche. Questa conformazione crea una tasca idrofobica specifica per il legame con i ligandi, tipicamente localizzati nelle regioni C-terminali delle proteine target.

Le caratteristiche principali del dominio PDZ includono:

- **Specificità di legame:** Il dominio PDZ riconosce sequenze specifiche di 4-5 amminoacidi, note come *motivi PDZ-binding*, situate alla fine delle proteine bersaglio.
- **Modularità:** I domini PDZ possono essere combinati in proteine più grandi, consentendo una vasta gamma di interazioni con altre proteine.
- **Versatilità funzionale:** Grazie alla loro struttura conservata e al riconoscimento sequenza-specifico, i domini PDZ partecipano a processi chiave come la segnalazione cellulare, l'organizzazione sinaptica e la polarità cellulare.

Funzioni biologiche del dominio PDZ

I domini PDZ giocano un ruolo centrale in diversi processi biologici. Tra le loro funzioni principali si annoverano:

1. Organizzazione della segnalazione cellulare

Il dominio PDZ funge da piattaforma di ancoraggio per i complessi proteici coinvolti nella trasduzione del segnale. Ad esempio, proteine PDZ come PSD-95 si associano ai recettori del glutammato nei neuroni, contribuendo alla formazione di sinapsi funzionali e alla modulazione della plasticità sinaptica.

2. Mantenimento della polarità cellulare

Le proteine contenenti domini PDZ, come ZO-1 e Dlg1, sono fondamentali per mantenere la polarità apico-basolaterale nelle cellule epiteliali. Attraverso l'interazione con le giunzioni strette, queste proteine assicurano una corretta compartimentalizzazione delle membrane cellulari.

3. Traffico di proteine e membrana

I domini PDZ regolano il trasporto intracellulare di proteine bersaglio verso compartimenti specifici. Un esempio è rappresentato dalla proteina NHERF1, che interagisce con il recettore della paratiroide per regolare il riassorbimento del calcio nei reni.

4. Sinapsi e plasticità neuronale

Nei neuroni, i domini PDZ sono coinvolti nell'organizzazione delle proteine postsinaptiche. La formazione di complessi proteici attraverso interazioni PDZ-ligando è essenziale per la stabilizzazione dei recettori di membrana e per la trasmissione sinaptica efficace.

Il dominio PDZ e la dinamica allosterica

Un aspetto emergente nello studio delle proteine PDZ è la loro capacità di esibire dinamiche allosteriche. Sebbene il dominio PDZ sia relativamente piccolo, presenta una notevole plasticità conformazionale che gli consente di rispondere a stimoli molecolari attraverso cambiamenti strutturali.

Meccanismo di allostericità

L'allostericità nel dominio PDZ si manifesta quando il legame di un ligando nella tasca canonica induce modifiche conformazionali che influenzano la capacità di legare altri partner molecolari. Questo fenomeno è particolarmente evidente nei domini PDZ delle proteine multi-PDZ, dove l'interazione con un ligando può modulare la funzione di altri domini.

Implicazioni biologiche

L'allostericità nei domini PDZ è cruciale per processi come:

- *Coordinazione sinaptica*: Le modifiche conformazionali nei domini PDZ di PSD-95 regolano l'associazione con recettori postsinaptici.
- *Risposta a stimoli esterni*: Il dominio PDZ può agire come sensore molecolare, adattando la sua funzione a variazioni ambientali o fisiologiche.

Metodologie di studio del dominio PDZ

Lo studio delle proteine PDZ ha beneficiato di un'ampia gamma di approcci sperimentali e computazionali:

1. Cristallografia a raggi X

Le strutture cristallografiche hanno fornito informazioni dettagliate sulla tasca di legame e sulle interazioni ligando-dominio. Ad esempio, le strutture di PSD-95 e ZO-1 hanno rivelato l'importanza delle interazioni idrofobiche nel riconoscimento del ligando.

2. Risonanza magnetica nucleare (NMR)

L'NMR ha permesso di osservare le dinamiche conformazionali del dominio PDZ in soluzione, evidenziando la presenza di microstati energetici correlati all'allostericità.

3. Simulazioni molecolari

Le simulazioni molecolari hanno fornito una visione dettagliata dei processi dinamici del dominio PDZ, inclusi i meccanismi di legame e i cambiamenti conformazionali indotti.

4. Approcci sperimentali funzionali

Mutagenesi sito-specifica e saggi funzionali sono stati utilizzati per identificare residui chiave per il legame del ligando e per testare l'impatto dell'allostericità sulla funzione proteica.

Implicazioni biomediche delle proteine PDZ

Le proteine contenenti domini PDZ sono coinvolte in numerose patologie, tra cui:

- **Tumori:** Alterazioni nelle interazioni PDZ-ligando possono contribuire alla perdita di polarità cellulare e alla progressione tumorale.
- **Malattie neurodegenerative:** Le disfunzioni nei complessi PDZ sono associate a patologie come la malattia di Alzheimer, dove la segnalazione sinaptica è compromessa.
- **Malattie cardiovascolari:** Domini PDZ regolano i canali ionici e i recettori coinvolti nel controllo della pressione sanguigna e nella contrattilità cardiaca.

Conclusioni

Il dominio PDZ rappresenta un modello ideale per studiare le interazioni proteina-proteina e i meccanismi allosterici alla base delle funzioni cellulari. La sua modularità, specificità di legame e versatilità funzionale lo rendono un bersaglio fondamentale per lo sviluppo di terapie innovative volte a modulare l'attività proteica in contesti patologici.

Chapter 1

Modi normali nelle proteine

Introduzione ai modi normali

La teoria dei modi normali (Normal Mode Analysis, NMA) è una tecnica fondamentale per analizzare le fluttuazioni dinamiche delle proteine a livello atomico. Questa metodologia consente di studiare le oscillazioni collettive delle strutture tridimensionali delle proteine attorno a uno stato di equilibrio, modellando il sistema come una rete di oscillatori armonici. I modi normali rappresentano i movimenti collettivi indipendenti che compongono la dinamica della proteina.

Matematicamente, la teoria dei modi normali si basa sull'approssimazione di un minimo locale della superficie di energia potenziale. In questa sezione, verrà introdotto il formalismo matematico alla base dell'NMA, con esempi applicati alla dinamica delle proteine.

1.1 Modello matematico dei modi normali

1. Energia potenziale della proteina

L'energia potenziale di una proteina può essere descritta come una funzione $V(\mathbf{x})$, dove $\mathbf{x} = (x_1, x_2, \dots, x_{3N})$ rappresenta le coordinate dei N atomi nella struttura tridimensionale. In prossimità di un minimo locale, l'energia può essere approssimata tramite uno sviluppo di Taylor di secondo ordine:

$$V(\mathbf{x}) \approx V(\mathbf{x}_0) + \frac{1}{2} \sum_{i,j} \frac{\partial^2 V}{\partial x_i \partial x_j} \Delta x_i \Delta x_j,$$

dove \mathbf{x}_0 è la configurazione di equilibrio e $\Delta \mathbf{x} = \mathbf{x} - \mathbf{x}_0$ rappresenta la deviazione rispetto all'equilibrio. Il termine $\frac{\partial^2 V}{\partial x_i \partial x_j}$ definisce la matrice hessiana \mathbf{H} dell'energia

potenziale:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 V}{\partial x_1^2} & \frac{\partial^2 V}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 V}{\partial x_1 \partial x_{3N}} \\ \frac{\partial^2 V}{\partial x_2 \partial x_1} & \frac{\partial^2 V}{\partial x_2^2} & \cdots & \frac{\partial^2 V}{\partial x_2 \partial x_{3N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial x_{3N} \partial x_1} & \frac{\partial^2 V}{\partial x_{3N} \partial x_2} & \cdots & \frac{\partial^2 V}{\partial x_{3N}^2} \end{bmatrix}.$$

2. Risoluzione del problema agli autovalori

Il movimento della proteina può essere descritto tramite la seconda legge di Newton:

$$M \frac{d^2 \mathbf{x}}{dt^2} = -\mathbf{H} \Delta \mathbf{x},$$

dove M è una matrice diagonale che contiene le masse atomiche. Dividendo per M e trasformando il sistema in coordinate ridotte, si ottiene l'equazione del moto:

$$\frac{d^2 \mathbf{q}}{dt^2} = -\mathbf{H}' \mathbf{q},$$

dove $\mathbf{H}' = M^{-1/2} \mathbf{H} M^{-1/2}$ è la matrice hessiana ridotta. La soluzione di questo sistema si ottiene diagonalizzando \mathbf{H}' :

$$\mathbf{H}' \mathbf{v}_i = \lambda_i \mathbf{v}_i,$$

dove λ_i sono gli autovalori (frequenze quadratiche dei modi normali) e \mathbf{v}_i sono gli autovettori (i modi normali).

1.2 Interpretazione fisica dei modi normali

Gli autovalori λ_i rappresentano le frequenze quadratiche associate ai modi normali, mentre gli autovettori \mathbf{v}_i descrivono la direzione del movimento per ciascun modo. I modi normali con autovalori piccoli (basse frequenze) corrispondono a movimenti collettivi lenti e di grande ampiezza, spesso associati a processi funzionali delle proteine, come l'apertura e chiusura di tasche enzimatiche o il movimento di domini.

1. Modi rigidi e vibrazionali

I primi sei modi normali, corrispondenti agli spostamenti traslazionali e rotazionali rigidi della proteina, hanno autovalori nulli o prossimi allo zero. Gli altri modi descrivono oscillazioni vibrazionali interne della proteina.

2. Fluttuazioni termiche

Le fluttuazioni termiche di una proteina possono essere descritte utilizzando i modi normali. La deviazione media quadratica di un atomo i in equilibrio è proporzionale all'inverso degli autovalori:

$$\langle (\Delta x_i)^2 \rangle \propto \sum_k \frac{(v_k)_i^2}{\lambda_k},$$

dove $(v_k)_i$ è la componente del k -esimo autovettore associata all'atomo i .

1.3 Applicazioni della teoria dei modi normali alle proteine

1. Analisi della dinamica funzionale

La teoria dei modi normali è stata utilizzata per comprendere i movimenti collettivi delle proteine, come l'apertura e la chiusura delle tasche enzimatiche. Ad esempio, per l'enzima adenilato ciclasi, i modi normali a basse frequenze mostrano oscillazioni che facilitano l'accesso del substrato al sito attivo.

2. Calcolo dei fattori B

I fattori B, osservati nella cristallografia a raggi X, rappresentano la flessibilità atomica in una struttura proteica. Tramite NMA, è possibile prevedere i fattori B teorici:

$$B_i = \frac{8\pi^2}{3} \langle (\Delta x_i)^2 \rangle.$$

3. Disegno di farmaci

La dinamica delle proteine, studiata tramite NMA, è essenziale per il design di farmaci. I movimenti lenti e collettivi identificati tramite modi normali possono rivelare siti allosterici potenziali per l'intervento farmacologico.

4. Stabilità e ripiegamento proteico

I modi normali possono essere utilizzati per studiare il ripiegamento proteico e le transizioni conformazionali. Ad esempio, analizzando le frequenze dei modi, è possibile identificare i movimenti correlati alla destabilizzazione di strutture secondarie.

1.4 Modelli semplificati: elastic network model (ENM)

Per ridurre la complessità computazionale, sono stati sviluppati modelli semplificati come l'Elastic Network Model (ENM). In questi modelli, gli atomi della proteina sono rappresentati come nodi connessi da molle, e l'energia potenziale è descritta come:

$$V = \frac{1}{2} \sum_{i < j} k_{ij} (d_{ij} - d_{ij}^0)^2,$$

dove d_{ij} è la distanza tra gli atomi i e j , d_{ij}^0 è la distanza di equilibrio, e k_{ij} è la costante elastica.

L'ENM semplifica il calcolo della matrice hessiana, rendendo l'analisi dei modi normali computazionalmente più efficiente.

1.5 Conclusioni

La teoria dei modi normali offre una descrizione dettagliata delle dinamiche proteiche, fornendo informazioni preziose sui movimenti collettivi e sulle fluttuazioni termiche. Applicata a contesti biologici, come il design di farmaci e lo studio della funzione enzimatica, l'NMA rappresenta uno strumento potente per comprendere il comportamento dinamico delle proteine.

Chapter 2

Gaussian Network Models (GNM) nelle proteine

Introduzione ai modelli di rete gaussiani

Il Gaussian Network Model (GNM) è un approccio semplificato ma efficace per analizzare le dinamiche collettive delle proteine basandosi sulla loro struttura tridimensionale. Questo modello si colloca nella categoria degli *elastic network models* (ENM) e rappresenta una potente astrazione per descrivere le fluttuazioni attorno alla configurazione nativa di una proteina. Per costruire una descrizione rigorosa, è essenziale passare da una rappresentazione generale della dinamica molecolare, basata sull'Hamiltoniana, a un sistema equivalente che utilizza molle armoniche per modellare le interazioni tra i residui.

2.1 Dall'Hamiltoniana alla descrizione elastica

La dinamica completa di una proteina può essere rappresentata da un'Hamiltoniana generica, che descrive l'energia totale del sistema come somma dell'energia cinetica e dell'energia potenziale:

$$H = T + V, \quad (2.1)$$

dove T rappresenta l'energia cinetica e V l'energia potenziale. Nello specifico, l'energia cinetica è espressa come:

$$T = \frac{1}{2} \sum_{i=1}^N m_i \|\dot{\mathbf{r}}_i\|^2, \quad (2.2)$$

con m_i massa del residuo i , $\dot{\mathbf{r}}_i$ velocità del residuo i , e N il numero totale di residui. L'energia potenziale V descrive le interazioni tra i residui e, in linea di principio, include termini complessi che considerano legami covalenti, interazioni non covalenti e forze di esclusione sterica.

Semplificazione tramite approccio armonico

Per studiare le fluttuazioni attorno alla configurazione nativa, si assume che le deviazioni dalle posizioni di equilibrio siano piccole. In questo regime, l'energia potenziale può essere approssimata tramite un'espansione di Taylor al secondo ordine attorno alla configurazione nativa $\{\mathbf{r}_i^0\}$:

$$V(\{\mathbf{r}_i\}) \approx V_0 + \frac{1}{2} \sum_{i,j=1}^N \frac{\partial^2 V}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j, \quad (2.3)$$

dove V_0 è il valore dell'energia potenziale nella configurazione nativa e $\Delta \mathbf{r}_i = \mathbf{r}_i - \mathbf{r}_i^0$ rappresenta la deviazione della posizione del residuo i rispetto all'equilibrio. Il termine di ordine zero può essere ignorato poiché non contribuisce alla dinamica, e il termine di ordine uno si annulla poiché la forza è nulla all'equilibrio. Pertanto, l'energia potenziale è dominata dal termine quadratico.

La matrice delle seconde derivate, detta matrice Hessiana \mathbf{H} , cattura la curvatura locale dell'energia potenziale:

$$H_{ij} = \frac{\partial^2 V}{\partial \mathbf{r}_i \partial \mathbf{r}_j}. \quad (2.4)$$

Modello a molle armoniche

In un GNM, l'Hamiltoniana viene ulteriormente semplificata assumendo che i residui siano collegati da molle armoniche, con una costante elastica uniforme γ . In questo modello, l'energia potenziale è scritta come:

$$V(\{\Delta \mathbf{r}_i\}) = \frac{\gamma}{2} \sum_{i=1}^N \sum_{j=1}^N \Gamma_{ij} (\Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j), \quad (2.5)$$

dove $\mathbf{\Gamma}$ è la matrice di Kirchoff, definita come:

$$\Gamma_{ij} = \begin{cases} -1 & \text{se } r_{ij} \leq R_c \text{ e } i \neq j, \\ \sum_{k \neq i} A_{ik} & \text{se } i = j, \end{cases} \quad (2.6)$$

e A_{ij} è la matrice di adiacenza della rete, che rappresenta le connessioni elastiche tra i residui basate su un cutoff di distanza R_c :

$$A_{ij} = \begin{cases} 1 & \text{se } r_{ij} \leq R_c \text{ e } i \neq j, \\ 0 & \text{altrimenti.} \end{cases} \quad (2.7)$$

In questo contesto, γ rappresenta una costante elastica uniforme che caratterizza la rigidità delle molle.

Energia totale del sistema

L'Hamiltoniana finale, che combina l'energia cinetica e quella potenziale armonica, è quindi data da:

$$H = \frac{1}{2} \sum_{i=1}^N m_i \|\dot{\mathbf{r}}_i\|^2 + \frac{\gamma}{2} \sum_{i=1}^N \sum_{j=1}^N \Gamma_{ij} (\Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j). \quad (2.8)$$

Questa forma descrive completamente il sistema in termini di una rete elastica di nodi (i residui) connessi da molle (le interazioni elastiche). Le equazioni del moto derivate da questa Hamiltoniana permettono di calcolare le fluttuazioni dinamiche e le correlazioni tra i residui.

2.2 Fluttuazioni e correlazioni dinamiche

Le fluttuazioni medie quadratiche $\langle (\Delta R_i)^2 \rangle$ di ciascun residuo e le correlazioni $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle$ tra coppie di residui sono direttamente ottenibili risolvendo le equazioni di equilibrio. La matrice inversa di $\mathbf{\Gamma}$, detta *matrice delle correlazioni inverse*, è fondamentale per queste analisi:

$$\langle \Delta \mathbf{R} \Delta \mathbf{R}^T \rangle \propto \mathbf{\Gamma}^{-1}. \quad (2.9)$$

Le fluttuazioni medie quadratiche per un residuo i sono date da:

$$\langle (\Delta R_i)^2 \rangle = \left[\mathbf{\Gamma}^{-1} \right]_{ii}, \quad (2.10)$$

mentre le correlazioni tra i residui i e j sono:

$$\langle \Delta R_i \cdot \Delta R_j \rangle = - \left[\mathbf{\Gamma}^{-1} \right]_{ij}. \quad (2.11)$$

2.3 Conclusioni

Gli indicatori dinamici, come la correlazione, la risposta lineare e l'entropia di produzione, forniscono strumenti matematici potenti per analizzare i meccanismi allosterici delle proteine. Combinando questi approcci, è possibile ottenere una comprensione dettagliata delle reti dinamiche che governano la funzione proteica, aprendo nuove prospettive per la progettazione di farmaci e per l'ingegneria delle proteine.

Chapter 3

Indicatori per l'analisi delle dinamiche proteiche e allosteriche

3.0.1 Il concetto di indicatori dinamici

Gli indicatori dinamici rappresentano strumenti matematici e fisici fondamentali per analizzare e caratterizzare le proprietà dinamiche di sistemi complessi, come le proteine, le reti molecolari o altri sistemi biologici. Essi si basano su una descrizione approfondita delle fluttuazioni, delle interazioni e dei processi stocastici che governano il comportamento collettivo dei componenti del sistema. L'importanza degli indicatori dinamici risiede nella loro capacità di estrarre informazioni quantitative dalle dinamiche sottostanti, fornendo un quadro dettagliato delle relazioni tra struttura e funzione.

Definizione generale e motivazioni Gli indicatori dinamici sono definiti come quantità matematiche o fisiche che catturano aspetti specifici delle dinamiche di un sistema. A differenza delle analisi statiche, che considerano una singola configurazione della proteina o una struttura cristallografica, questi strumenti forniscono informazioni sulle fluttuazioni temporali, sui moti collettivi e sulle interazioni spaziali tra diverse parti del sistema.

In particolare, essi sono utili per:

- **Quantificare le fluttuazioni locali e globali:** Le proteine non sono strutture rigide, ma oscillano costantemente attorno alla loro configurazione nativa a causa delle fluttuazioni termiche. Gli indicatori dinamici permettono di misurare e caratterizzare l'ampiezza di queste fluttuazioni, evidenziando regioni flessibili o rigidamente vincolate.
- **Esplorare le correlazioni dinamiche:** I residui all'interno di una proteina possono muoversi in modo sincronizzato o indipendente. Gli indicatori dinamici aiutano a determinare il grado di cooperazione tra questi movimenti, identificando moti collettivi che possono essere cruciali per la funzione biologica.

- **Valutare la trasmissione di informazioni:** Le proteine spesso trasmettono segnali conformazionali da una regione all'altra, un fenomeno noto come allostericità. Gli indicatori dinamici forniscono strumenti per quantificare il trasferimento di informazioni dinamiche attraverso la struttura proteica.

Fluttuazioni locali e globali: una panoramica Le fluttuazioni dinamiche di una proteina possono essere analizzate a diversi livelli di dettaglio:

- **Fluttuazioni locali:** Queste descrivono i moti di piccoli gruppi di atomi o residui. Indicatori come le fluttuazioni medie quadratiche ($\langle(\Delta R_i)^2\rangle$) forniscono una misura della mobilità di ciascun residuo, mettendo in evidenza le regioni flessibili come loop o terminali liberi.
- **Fluttuazioni globali:** Questi moti coinvolgono interi domini o subunità della proteina. Gli indicatori dinamici, come i modi normali a basse frequenze, permettono di identificare movimenti collettivi che supportano funzioni biologiche chiave, come l'apertura e chiusura di tasche enzimatiche.

Correlazioni dinamiche e interazioni tra residui Le correlazioni dinamiche rappresentano un altro aspetto critico delle dinamiche proteiche. Queste misurano il grado di sincronia nei movimenti di due residui o regioni distinte. Indicatori come la matrice delle correlazioni dinamiche $C_{ij}(t)$ sono fondamentali per:

- Identificare reti di residui cooperativi, che spesso formano il cuore dei meccanismi allosterici.
- Distinguere tra movimenti sincroni (correlazioni positive) e movimenti opposti (anticorrelazioni negative).

Questi concetti sono particolarmente rilevanti per analizzare proteine multidominio, dove il coordinamento tra diversi segmenti è essenziale per la funzione.

Trasferimento di informazione e comunicazione dinamica Un ulteriore elemento chiave degli indicatori dinamici è la loro capacità di quantificare il trasferimento di informazioni tra diverse parti della proteina. Questo è cruciale per comprendere i meccanismi allosterici, in cui eventi locali, come il legame di un ligando, causano cambiamenti conformazionali che si propagano attraverso la struttura. Indicatori come la *transfer entropy* (TE) permettono di:

- Quantificare il flusso di informazione da un residuo a un altro.
- Identificare i "nodi" critici della rete dinamica che mediano la comunicazione tra regioni distanti.

Connessione con i processi stocastici Il comportamento dinamico di una proteina è governato da processi stocastici, che combinano forze deterministiche, come quelle elastiche, e fluttuazioni casuali termiche. Gli indicatori dinamici sono strettamente legati a questa descrizione probabilistica, poiché derivano da equazioni del moto che incorporano:

- **Fluttuazioni termiche:** Modellate come rumore bianco gaussiano.
- **Reti elastiche:** Che descrivono le interazioni tra i residui vicini.

Questo quadro matematico consente di collegare gli indicatori dinamici a grandezze fisiche misurabili, come i fattori di Debye-Waller (*B*-factors) ottenuti dalla cristallografia a raggi X.

Relazione con le proprietà spettrali della matrice di Kirchhoff Un aspetto cruciale degli indicatori dinamici è la loro relazione con le proprietà spettrali della matrice di Kirchhoff, che descrive la connettività elastica del sistema. Gli autovalori λ_k e gli autovettori \mathbf{u}_k della matrice di Kirchhoff permettono di decomporre i movimenti dinamici in *modi normali*, ciascuno caratterizzato da una frequenza e un'ampiezza specifica. Gli indicatori dinamici possono quindi essere interpretati come combinazioni ponderate di questi modi normali, fornendo una comprensione più profonda delle dinamiche collettive.

Applicazioni pratiche degli indicatori dinamici Gli indicatori dinamici sono stati applicati con successo in numerosi studi per:

- Identificare regioni flessibili o rigide in enzimi, che sono spesso correlate a siti attivi o di legame.
- Analizzare i meccanismi di apertura e chiusura di canali ionici.
- Progettare proteine ingegnerizzate con proprietà dinamiche desiderate.
- Sviluppare farmaci mirati contro regioni funzionali identificate dinamicamente.

Limiti e sfide Nonostante il loro successo, gli indicatori dinamici presentano alcune limitazioni:

- La loro interpretazione dipende fortemente dal modello utilizzato per descrivere la dinamica proteica (es. rete elastica vs dinamica molecolare).
- Alcuni indicatori richiedono una grande quantità di dati o calcoli intensivi, specialmente in sistemi complessi o di grandi dimensioni.
- La scelta dei parametri (es. cutoff di distanza, costanti elastiche) può influenzare significativamente i risultati.

Conclusioni Gli indicatori dinamici offrono un potente set di strumenti per analizzare le dinamiche proteiche, fornendo informazioni critiche sulle interazioni e sui meccanismi funzionali. La loro combinazione con modelli computazionali avanzati e dati sperimentali rappresenta una direzione promettente per future ricerche in biologia strutturale e biofisica.

3.0.2 Tipologie di indicatori dinamici

Nel contesto delle dinamiche proteiche, gli indicatori dinamici rappresentano strumenti essenziali per quantificare e comprendere i moti collettivi e locali di una proteina, analizzando non solo le interazioni statiche ma anche quelle temporali e causali. Questi indicatori consentono di esplorare le relazioni tra struttura, dinamica e funzione, fornendo informazioni fondamentali su come le proteine rispondono a stimoli interni o esterni e su come comunicano all'interno della loro rete molecolare.

Il ruolo degli indicatori dinamici nelle proteine Gli indicatori dinamici sono progettati per catturare diversi aspetti delle fluttuazioni e delle interazioni proteiche:

- **Aspetto locale:** caratterizzano il comportamento di singoli residui o piccoli gruppi di residui.
- **Aspetto collettivo:** descrivono i moti che coinvolgono regioni più ampie della proteina o addirittura l'intera struttura.
- **Causalità:** determinano in che misura il movimento di una parte della proteina influenzi direttamente un'altra.
- **Allostericità:** identificano come le modifiche locali (ad esempio, il legame con un ligando) si traducano in cambiamenti funzionali globali.

La classificazione degli indicatori dinamici è utile per separare le informazioni relative ai diversi livelli di organizzazione della dinamica proteica. Nel contesto delle reti elastiche e delle descrizioni stocastiche, possiamo identificare tre categorie principali di indicatori dinamici: **correlazioni dinamiche**, **risposta lineare** e **trasferimento di informazione**. Ognuno di essi ha un ruolo specifico nella comprensione della dinamica proteica.

1. Correlazioni dinamiche

Le correlazioni dinamiche rappresentano la sincronia nei movimenti di coppie di residui e sono descritte matematicamente dalla matrice delle correlazioni temporali $C_{ij}(t)$:

$$C_{ij}(t) = \langle x_i(0)x_j(t) \rangle, \quad (3.1)$$

dove $x_i(t)$ e $x_j(t)$ rappresentano i movimenti dei residui i e j nel tempo, e $\langle \cdot \rangle$ denota la media sull'ensemble. Le correlazioni dinamiche hanno diverse interpretazioni:

- **Correlazioni positive:** indicano che i residui si muovono in modo cooperativo nella stessa direzione.
- **Correlazioni negative:** rappresentano movimenti opposti, spesso associati a cambiamenti conformazionali significativi.
- **Correlazioni nulle:** suggeriscono che i movimenti di due residui sono indipendenti.

Le correlazioni dinamiche sono particolarmente utili per identificare *reti di residui cooperativi*, che sono spesso coinvolte in meccanismi funzionali chiave come l'allostericità. Ad esempio, le correlazioni elevate tra un sito attivo e un sito distante possono indicare la presenza di un canale di comunicazione dinamico che trasmette segnali conformazionali attraverso la proteina.

Esempi applicativi delle correlazioni dinamiche Le correlazioni dinamiche sono state utilizzate per analizzare enzimi come l'HIV-1 proteasi, dove le correlazioni tra i flap e il core enzimatico hanno rivelato meccanismi di apertura e chiusura essenziali per la catalisi. Analogamente, nei canali ionici, le correlazioni dinamiche hanno permesso di identificare residui critici per la regolazione del flusso ionico.

2. Risposta lineare

La risposta lineare misura come un residuo risponde a perturbazioni esterne applicate ad altri residui. Formalmente, la risposta dinamica è definita come:

$$R_{ij}(t) = \frac{\delta \langle x_i(t) \rangle}{\delta F_j(0)}, \quad (3.2)$$

dove $F_j(0)$ rappresenta una forza esterna applicata al residuo j al tempo iniziale.

La risposta lineare è un indicatore di *sensibilità dinamica*: descrive la misura in cui un residuo è influenzato dai cambiamenti in altre parti della proteina. In particolare, permette di identificare residui che agiscono come nodi critici nella rete di comunicazione interna della proteina.

Interpretazione fisica della risposta lineare La risposta $R_{ij}(t)$ è strettamente legata alla matrice di Kirchhoff \mathbf{K} e ai suoi autovalori/autovettori. I residui che mostrano una risposta elevata a perturbazioni esterne sono spesso coinvolti in meccanismi di trasmissione del segnale. Inoltre:

- Una risposta alta suggerisce una connessione dinamica forte tra due residui.
- Una risposta bassa implica una connessione debole o un'assenza di comunicazione diretta.

Utilizzo della risposta lineare nell'allostericità Nei sistemi allosterici, la risposta lineare può essere utilizzata per prevedere quali residui trasmettono segnali conformazionali dal sito di legame all'intera proteina. Ad esempio, in una proteina come l'emoglobina, la risposta lineare permette di mappare il percorso dinamico attraverso cui il legame dell'ossigeno in un sito attivo induce cambiamenti strutturali cooperativi.

3. Trasferimento di informazione

Il trasferimento di informazione è un indicatore fondamentale per quantificare la causalità nelle dinamiche proteiche. Formalmente, si basa sulla misura di *transfer entropy* (TE), definita come:

$$TE_{j \rightarrow i}(t) = \langle \log ($$

$\overline{)P[x_i(t)|x_i(0), x_j(0)]P[x_i(t)|x_i(0)], (3.3)}$ dove $P[x_i(t)|x_i(0), x_j(0)]$ è la probabilità condizionata del movimento del residuo i dato lo stato iniziale di i e j , e $P[x_i(t)|x_i(0)]$ esclude l'influenza di j .

Causalità e comunicazione dinamica A differenza delle correlazioni, che descrivono solo sincronia, il trasferimento di informazione identifica relazioni causali tra residui. Un valore elevato di $TE_{j \rightarrow i}(t)$ implica che i movimenti del residuo j influenzano direttamente i movimenti del residuo i . Questo è cruciale per comprendere come le proteine trasmettono segnali conformazionali attraverso reti di interazioni dinamiche.

Importanza nell'allostericità Nel contesto dell'allostericità, il trasferimento di informazione permette di tracciare i percorsi dinamici che collegano un evento locale (come il legame di un ligando) a cambiamenti globali nella struttura proteica. Ad esempio:

- Nelle chinasi, il trasferimento di informazione può rivelare come il legame di ATP regoli l'attività catalitica attraverso moti cooperativi.
- Nei recettori accoppiati a proteine G (GPCR), l'analisi del trasferimento di informazione evidenzia i residui chiave coinvolti nella trasduzione del segnale attraverso la membrana.

Calcolo e applicazione pratica Il calcolo del trasferimento di informazione si basa sull'analisi di traiettorie stocastiche e richiede una buona conoscenza delle proprietà dinamiche del sistema. È stato applicato con successo per identificare regioni allosteriche in numerosi sistemi proteici, aiutando a progettare farmaci che mirano a residui specifici per modulare la funzione proteica.

Conclusioni Le tre tipologie principali di indicatori dinamici – correlazioni dinamiche, risposta lineare e trasferimento di informazione – forniscono una gamma completa di strumenti per analizzare le dinamiche proteiche. Ciascuno di essi contribuisce a un aspetto specifico della comprensione delle relazioni dinamiche e causali nelle proteine, con applicazioni dirette nella biologia strutturale, nell'ingegneria proteica e nel design di farmaci. La loro combinazione offre una visione dettagliata e multilivello delle interazioni che governano i processi biologici.

3.0.3 Gli indicatori dinamici come strumenti per lo studio delle proteine

Gli indicatori dinamici sono definiti come quantità matematiche o fisiche che descrivono le proprietà dinamiche di un sistema complesso. Nel contesto delle proteine, questi indicatori aiutano a:

- Quantificare le fluttuazioni dei residui attorno alla loro configurazione di equilibrio.
- Esplorare le correlazioni nei movimenti collettivi di residui distanti.

- Valutare la trasmissione di segnali conformazionali lungo la catena proteica.

Tali informazioni sono essenziali per comprendere fenomeni complessi come l'allostericità, in cui eventi locali, come il legame di un ligando, influenzano la dinamica e la funzione globale della proteina.

3.0.4 Contesto generale: reti elastiche e modelli stocastici

La modellazione dinamica delle proteine avviene solitamente tramite una rappresentazione semplificata: le reti elastiche. In questi modelli, una proteina è rappresentata come un grafo non orientato, in cui:

- I nodi corrispondono a residui (o atomi specifici, come i C_α).
- Gli spigoli rappresentano le connessioni elastiche tra residui, che possono essere determinate da un criterio di distanza (cutoff) o da interazioni chimiche specifiche.

La dinamica del sistema è poi descritta utilizzando equazioni differenziali stocastiche, che rappresentano la combinazione di forze deterministiche (elastiche) e fluttuazioni casuali (termiche). Queste equazioni consentono di calcolare quantità osservabili come:

- **Correlazioni dinamiche:** la sincronia nei movimenti di coppie di residui.
- **Risposta lineare:** la sensibilità del sistema a piccole perturbazioni.
- **Trasferimento di informazione:** il flusso di dati dinamici tra residui distanti.

3.0.5 Importanza degli indicatori dinamici nello studio delle proteine

Lo studio delle dinamiche proteiche con indicatori dinamici ha portato a importanti scoperte in biofisica e biologia strutturale, come:

- Identificazione di regioni altamente flessibili, che spesso corrispondono a loop o a terminali liberi.
- Individuazione di siti chiave responsabili della comunicazione allosterica.
- Interpretazione dei movimenti collettivi che supportano funzioni biologiche specifiche, come l'apertura e la chiusura di siti catalitici.

La capacità di quantificare tali proprietà è fondamentale per applicazioni come il design di farmaci e la progettazione di proteine.

3.1 Modello matematico delle dinamiche proteiche

3.1.1 Rappresentazione di una proteina come rete elastica

Un approccio comune per modellare le dinamiche proteiche è rappresentare la proteina come un grafo in cui i nodi corrispondono ai residui e gli spigoli descrivono

interazioni elastiche. Questa rappresentazione semplificata cattura efficacemente i moti collettivi che sono rilevanti per la funzione.

La rete è costruita considerando un cutoff di distanza R_c :

$$r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\| \leq R_c, \quad (3.4)$$

dove r_{ij} è la distanza tra i residui i e j . Questo criterio crea una matrice di adiacenza \mathbf{A} , con elementi:

$$A_{ij} = \begin{cases} 1 & \text{se } r_{ij} \leq R_c \text{ e } i \neq j, \\ 0 & \text{altrimenti.} \end{cases} \quad (3.5)$$

3.1.2 La matrice di Kirchhoff

La matrice di Kirchhoff \mathbf{K} , detta anche matrice Laplaciana, è costruita a partire dalla matrice di adiacenza. Gli elementi di \mathbf{K} sono:

$$K_{ij} = \begin{cases} -\mathbf{A}_{ij}, & \text{se } i \neq j, \\ \sum_{k \neq i} \mathbf{A}_{ik}, & \text{se } i = j. \end{cases} \quad (3.6)$$

Questa matrice descrive la connettività globale della rete e gioca un ruolo centrale nella dinamica del sistema.

3.1.3 Dinamica stocastica del sistema

Il movimento dei nodi nella rete elastica è governato dall'equazione differenziale stocastica:

$$\gamma \dot{x}_i = -g \sum_{j=1}^N K_{ij} x_j + \sqrt{2\gamma k_B T} \xi_i(t), \quad (3.7)$$

dove:

- γ è il coefficiente di attrito.
- g è una costante di forza elastica.
- K_{ij} è la matrice di Kirchhoff.
- k_B è la costante di Boltzmann.
- $\xi_i(t)$ è un rumore bianco gaussiano con media nulla e correlazione:

$$\langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij} \delta(t - t'). \quad (3.8)$$

3.1.4 Soluzione dell'equazione differenziale

La soluzione dell'equazione differenziale può essere scritta in forma compatta utilizzando la matrice di Kirchhoff:

$$\mathbf{x}(t) = e^{-\mu \mathbf{K} t} \mathbf{x}(0) + \sqrt{\frac{2k_B T}{\gamma}} \int_0^t e^{-\mu \mathbf{K}(t-s)} \boldsymbol{\xi}(s) ds. \quad (3.9)$$

In questa espressione, $\mu = g/\gamma$ è il rapporto tra costante elastica e coefficiente di attrito.

3.2 Calcolo della correlazione dinamica

La correlazione tra i movimenti di due nodi è definita come:

$$C_{ij}(t) = \langle x_i(0)x_j(t) \rangle. \quad (3.10)$$

Sfruttando la soluzione esplicita di $\mathbf{x}(t)$ e le proprietà del rumore stocastico, si ottiene:

$$C(t) = e^{-\mu \mathbf{K}t} C(0), \quad (3.11)$$

dove $C(0)$ è la matrice di correlazione iniziale. Espandendo $C(t)$ in termini di autovalori λ_k e autovettori \mathbf{u}_k della matrice \mathbf{K} :

$$C_{ij}(t) = \frac{3k_B T}{g} \sum_{k=2}^N \frac{u_i(k)u_j(k)}{\lambda_k} e^{-\lambda_k t}. \quad (3.12)$$

3.3 Calcolo della risposta lineare

La risposta lineare è definita come:

$$R_{ij}(t) = \frac{\delta \langle x_i(t) \rangle}{\delta F_j(0)}. \quad (3.13)$$

Per un sistema armonico, si dimostra che:

$$R_{ij}(t) = \sum_{k=1}^N u_i(k)u_j(k)e^{-\lambda_k t}. \quad (3.14)$$

3.4 Calcolo del trasferimento di informazione

La Cross-Entropy misura il flusso di informazione direzionale:

$$TE_{j \rightarrow i}(t) = \langle \log ($$

$\frac{P[x_i(t)|x_i(0), x_j(0)]P[x_i(t)|x_i(0)]}{P[x_i(t)|x_i(0), x_j(0)]P[x_i(t)|x_i(0)]} \rangle$ (3.15) Per processi gaussiani, si esprime come:

$$TE_{j \rightarrow i}(t) = -\frac{1}{2} \ln \left(1 - \frac{\alpha_{ij}(t)}{\beta_{ij}(t)} \right), \quad (3.16)$$

con:

$$\alpha_{ij}(t) = [C_{ii}(0)C_{ij}(t) - C_{ij}(0)C_{ii}(t)]^2, \quad (3.17)$$

$$\beta_{ij}(t) = C_{ii}(0)^2 C_{jj}(0). \quad (3.18)$$

3.5 Modello matematico delle dinamiche proteiche

3.5.1 La rappresentazione della proteina come rete elastica

Le proteine sono entità complesse che possono essere modellate come reti elastiche per descriverne la dinamica. Questa rappresentazione si basa sull'idea che i residui della proteina interagiscono tra loro attraverso connessioni elastiche, che possono essere rappresentate come spigoli di un grafo. Tale approccio riduce la complessità atomistica di una proteina a un sistema discretizzato, mantenendo comunque le informazioni essenziali per descrivere le fluttuazioni collettive.

Definizione della rete elastica La rete elastica di una proteina è definita come segue:

- **I nodi** rappresentano i residui della proteina. Spesso, per semplicità, si utilizzano gli atomi C_α di ciascun residuo, poiché questi catturano il centro geometrico delle unità amminoacidiche.
- **Gli spigoli** descrivono le connessioni elastiche tra i nodi. Due nodi sono considerati collegati se la distanza euclidea tra loro soddisfa un criterio di cutoff R_c , generalmente scelto nell'intervallo 7-10 Å.

Ruolo del cutoff R_c La scelta del cutoff R_c è cruciale per costruire una rete accurata:

- **Cutoff troppo basso:** Se R_c è troppo piccolo, molte interazioni rilevanti vengono ignorate, portando a una rete frammentata in cui i movimenti collettivi non possono essere catturati.
- **Cutoff troppo alto:** Un cutoff elevato può introdurre connessioni irrealistiche, che non riflettono le proprietà fisiche della proteina, sovrastimando le correlazioni.

Per determinare il valore ottimale di R_c , si utilizza una procedura di ottimizzazione basata sul confronto tra i risultati del modello e i fattori B ottenuti sperimentalmente dalla cristallografia a raggi X. L'obiettivo è minimizzare l'errore medio assoluto (MAE) tra le fluttuazioni teoriche e i dati sperimentali.

Applicazioni della rete elastica La rappresentazione della proteina come rete elastica è ampiamente utilizzata in:

- **Modelli Gaussian Network (GNM):** per calcolare le fluttuazioni medie quadratiche e le correlazioni dinamiche.
- **Modelli Anisotropic Network (ANM):** per analizzare i modi normali di vibrazione anisotropici.
- **Previsioni strutturali:** per identificare regioni flessibili o rigide in proteine e complessi proteici.

3.5.2 Costruzione della matrice di Kirchhoff

La matrice di Kirchhoff \mathbf{K} è una rappresentazione matematica della connettività globale della rete elastica. Essa codifica le relazioni elastiche tra i residui e fornisce la base per calcolare proprietà dinamiche come le fluttuazioni e le correlazioni.

Definizione e proprietà di \mathbf{K} Gli elementi della matrice di Kirchhoff sono definiti come:

$$K_{ij} = \begin{cases} -\gamma_{ij}, & \text{se } i \neq j, \\ \sum_{k \neq i} \gamma_{ik}, & \text{se } i = j, \end{cases} \quad (3.19)$$

dove γ_{ij} rappresenta la costante elastica tra i nodi i e j . In genere, si assume che γ_{ij} sia una costante uniforme γ per tutte le connessioni valide, rendendo la matrice \mathbf{K} simmetrica e semidefinita positiva.

Caratteristiche principali di \mathbf{K}

- **Simmetria:** \mathbf{K} è simmetrica, il che garantisce che gli autovalori siano reali.
- **Autovalore nullo:** Il primo autovalore $\lambda_0 = 0$ corrisponde al modo rigido traslazionale della rete.
- **Diagonale dominante:** Gli elementi diagonali superano la somma dei valori assoluti degli elementi fuori diagonale, garantendo la stabilità della rete.

Costruzione pratica La matrice \mathbf{K} è costruita a partire dalla matrice di adiacenza \mathbf{A} , che identifica le connessioni tra i nodi:

$$A_{ij} = \begin{cases} 1, & \text{se } r_{ij} \leq R_c, \\ 0, & \text{altrimenti.} \end{cases} \quad (3.20)$$

Successivamente, la matrice \mathbf{K} è ottenuta combinando \mathbf{A} con le costanti elastiche γ_{ij} .

3.5.3 Autovalori e autovettori della matrice di Kirchhoff

Gli autovalori λ_k e gli autovettori \mathbf{u}_k della matrice \mathbf{K} sono fondamentali per descrivere la dinamica del sistema. Essi forniscono una rappresentazione spettrale delle vibrazioni collettive nella rete elastica.

Interpretazione degli autovalori

- **Autovalori piccoli:** Corrispondono a modi normali a bassa frequenza, associati a movimenti collettivi su larga scala.
- **Autovalori grandi:** Rappresentano vibrazioni locali ad alta frequenza.
- **Autovalore nullo:** Rappresenta il modo traslazionale rigido e non contribuisce alle fluttuazioni interne.

Autovettori come modi normali Ogni autovettore \mathbf{u}_k descrive un modo normale di vibrazione della rete. Le componenti dell'autovettore rappresentano le ampiezze relative dei movimenti di ciascun nodo.

Decomposizione spettrale La matrice \mathbf{K} può essere diagonalizzata come:

$$\mathbf{K} = \mathbf{U}\mathbf{A}\mathbf{U}^\dagger, \quad (3.21)$$

dove \mathbf{U} è la matrice degli autovettori e \mathbf{A} è la matrice diagonale degli autovalori.

3.5.4 Pseudo-invertibilità della matrice di Kirchhoff

Poiché \mathbf{K} non è invertibile a causa dell'autovalore nullo, si utilizza la pseudo-inversa \mathbf{K}^+ per calcolare proprietà dinamiche come fluttuazioni e correlazioni. La pseudo-inversa è definita come:

$$\mathbf{K}^+ = \mathbf{U}\mathbf{A}^+\mathbf{U}^\dagger, \quad (3.22)$$

dove \mathbf{A}^+ è ottenuta invertendo solo gli autovalori non nulli di \mathbf{A} .

3.6 Indicatori dinamici: correlazione, risposta e trasferimento di informazione

3.6.1 Descrizione generale degli indicatori dinamici

Gli indicatori dinamici si basano sulla matrice di Kirchhoff e sulla sua pseudo-inversa per calcolare quantità che descrivono le fluttuazioni e le interazioni dinamiche tra i residui. Essi includono:

- **Correlazioni dinamiche:** Quantificano la sincronia nei movimenti tra residui.
- **Risposta lineare:** Misurano la sensibilità di un residuo a perturbazioni esterne.
- **Trasferimento di informazione:** Identificano la causalità nei moti dei residui.

Questi indicatori offrono una comprensione dettagliata dei processi allosterici e delle dinamiche funzionali nelle proteine.

3.7 GRAFO

Allora la scelta per il grafo, per semplicità di scrittura è stato quello di fare dei legami solo e solo se il raggio di interazione è minore di un certo valore. Come si fa a trovare la threshold migliore per il raggio in questo modello? Si runna il modello al fine di minimizzare i MAE sui residui ogni volta con un raggio diverso. Il modello che ottiene il mae minimo, è quello che meglio descrive il tutto. Inoltre in questo modo è possibile ottenere anche il miglior grafo senza alcun tipo di vincolo. Prima di predire i beta factor per davvero (con autovalori e autovettori) avrei bisogno di utilizzare i parametri corretti per temperatura Kb ecc.

3.8 Matrice di kirchoff

La matrice di Kirchhoff (o laplaciana) rappresenta un'analogia con una rete elastica in cui le connessioni tra i nodi (atomi) descrivono le interazioni elastiche. Questa matrice codifica il modo in cui ogni nodo è collegato agli altri, e attraverso i suoi autovalori e autovettori, si può studiare come le vibrazioni collettive (modi normali) si propagano attraverso il sistema. Autovalori e autovettori della matrice di Kirchhoff: Gli autovalori della matrice di Kirchhoff descrivono le frequenze naturali di vibrazione del sistema. Gli autovettori rappresentano i corrispondenti modi normali di vibrazione, cioè come ogni nodo (atomo) si muove in un determinato modo di vibrazione. Quelli a bassa frequenza corrispondono alle vibrazioni collettive del sistema, quelli ad alta frequenza sono fluttuazioni locali

3.9 Calcolo correlazione

Risolvi l'equazione differenziale:

$$\gamma \dot{x}_i = -g \sum_j K_{ij} x_j + \sqrt{2\gamma k_B T} \xi_i(t) \quad (3.23)$$

$$\mathbf{x}(t) = e^{-\mu K t} \left\{ \mathbf{x}(0) + \sqrt{\frac{2k_B T}{\gamma}} \int_0^t ds e^{-\mu K s} \boldsymbol{\xi}(s) \right\} \quad (3.24)$$

$$C(t) = \langle \mathbf{x}(0) \mathbf{x}^\top(t) \rangle \quad (3.25)$$

$$C(t) = e^{-\mu \mathbf{K} t} C(0) \quad (3.26)$$

$$C(0) = \langle \mathbf{x}(0) \mathbf{x}^\top(0) \rangle \quad (3.27)$$

$$\mathbf{K} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^\dagger \quad (3.28)$$

$$C_{ij}(t) = \frac{3k_B T}{g} \sum_{k=2}^N \frac{u_i(k) u_j(k)}{\lambda(k)} e^{-\lambda(k)t} \quad (3.29)$$

3.10 Calcolo risposta

$$R(t) = \frac{C(t)}{C(0)} \quad (3.30)$$

$$\mathbf{R}(t) = e^{-\mu \mathbf{K} t} \quad (3.31)$$

$$R_{ij}(t) = - \left\langle \frac{\partial \ln(P)_s(x)}{\partial x_j(t)} x_i(0) \right\rangle \quad (3.32)$$

$$R_{ij}(t) = \sum_{k=1}^N u_i(k) u_j(k) e^{-\lambda(k)t} \quad (3.33)$$

3.11 Cross-Entropy

$$TE_{j \rightarrow i}(t) = \langle \log ($$

$$\frac{P[x_i(t)|x_i(0), x_j(0)]P[x_i(t)|x_i(0)]}{P[x_i(t)|x_i(0), x_j(0)]P[x_i(t)|x_i(0)]} \rangle \quad (3.34)$$

$$T_{j \rightarrow i}(t) = -\frac{1}{2} \ln \left(1 - \frac{\alpha_{ij}(t)}{\beta_{ij}(t)} \right) \quad (3.35)$$

$$\alpha_{ij}(t) = [C_{ii}(0)C_{ij}(t) - C_{ij}(0)C_{ii}(t)]^2 \quad (3.36)$$

3.12 2M0Z

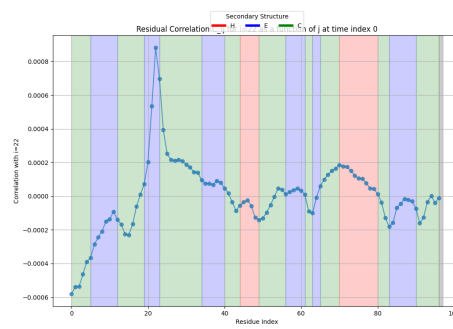


Figure 3.1. Correlazione

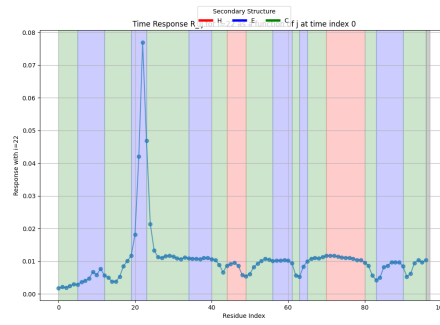


Figure 3.2. Risposta

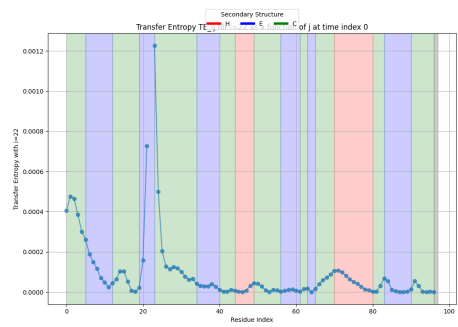


Figure 3.3. Transfer Entropy

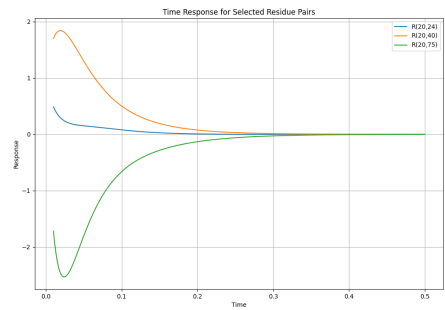


Figure 3.4. Multiple time response

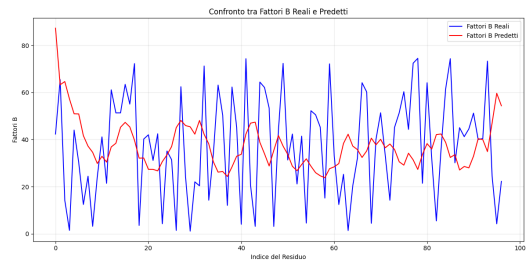


Figure 3.5. B factors

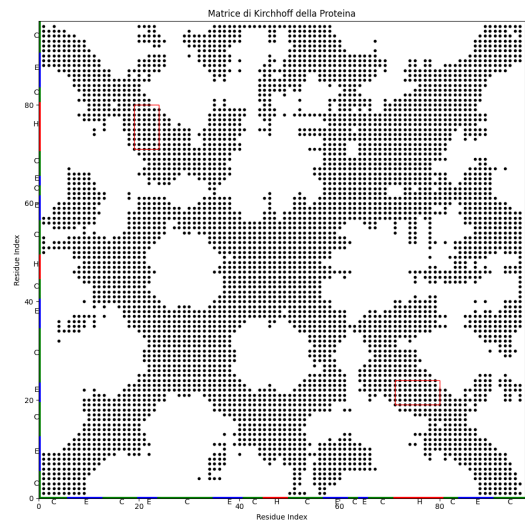


Figure 3.6. Kirchhoff

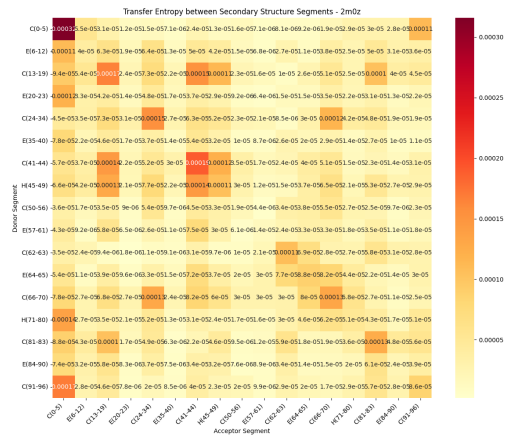


Figure 3.7. Secodnaria Structure

3.13 2M10

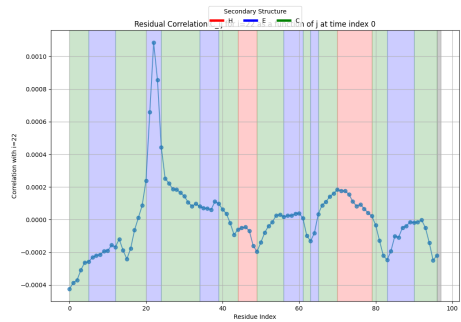


Figure 3.8. Correlazione

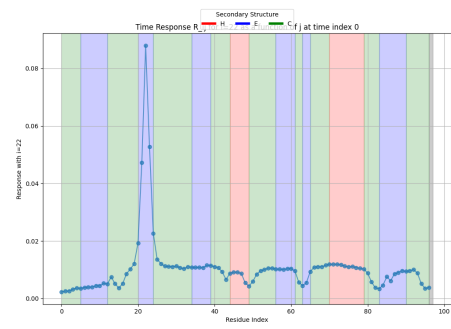


Figure 3.9. Risposta

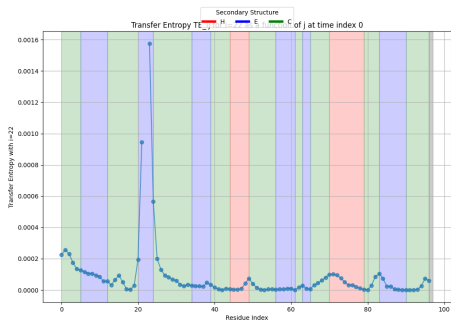


Figure 3.10. Transfer Entropy

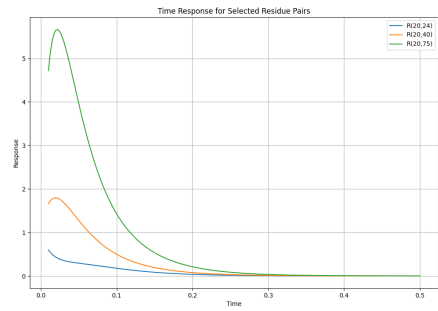


Figure 3.11. Multiple time response

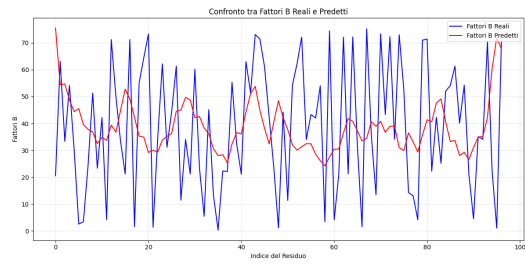


Figure 3.12. B factors

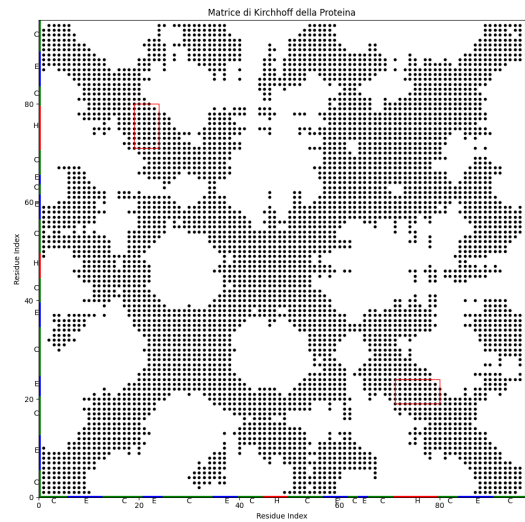


Figure 3.13. Kirchhoff

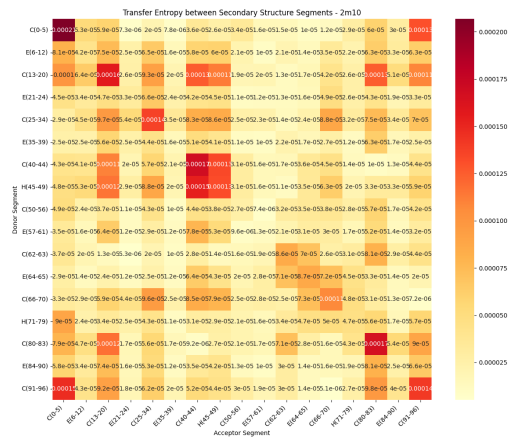


Figure 3.14. Transfer struttura Secondaria

3.14 3LNX

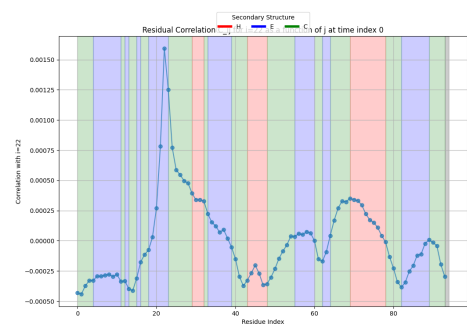


Figure 3.15. Correlazione

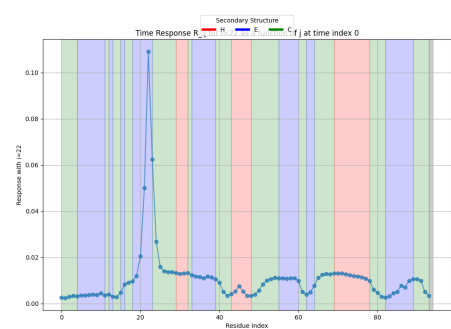


Figure 3.16. Risposta

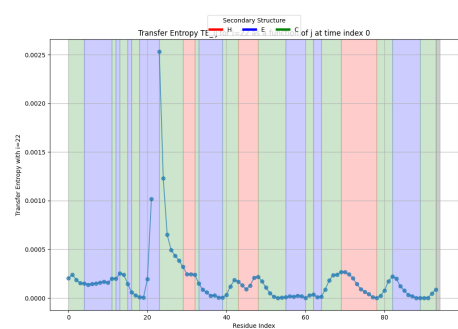


Figure 3.17. Transfer Entropy

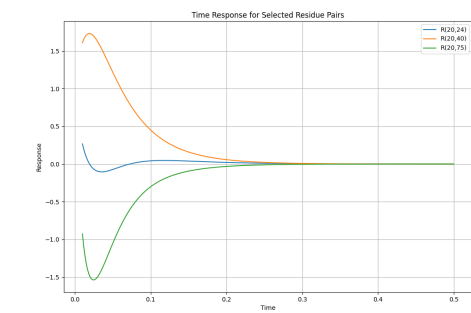


Figure 3.18. Multiple time response

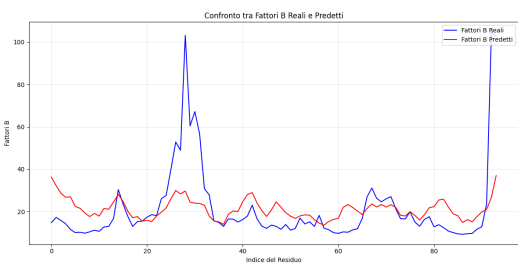


Figure 3.19. B factors

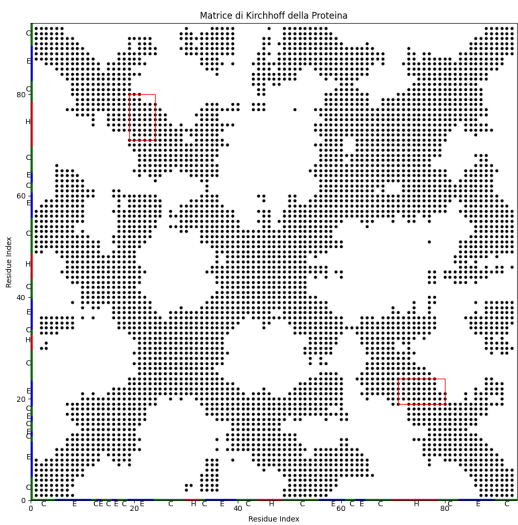


Figure 3.20. Kirchhoff

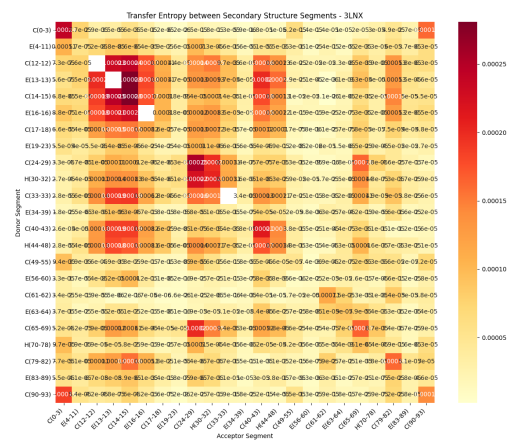


Figure 3.21. Secondaria

3.15 3LNY

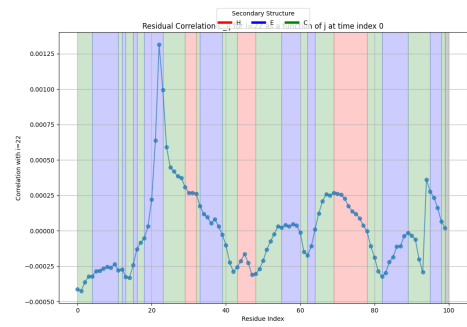


Figure 3.22. Correlazione

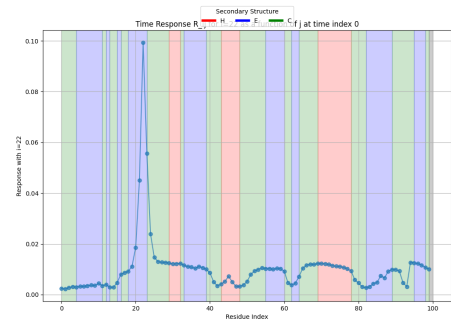


Figure 3.23. Risposta

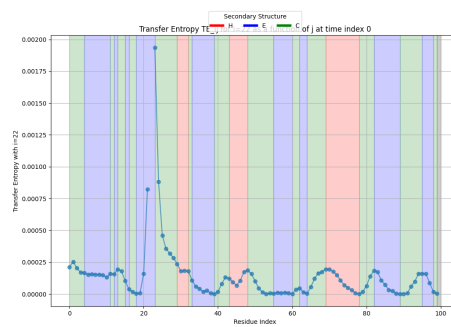


Figure 3.24. Transfer Entropy

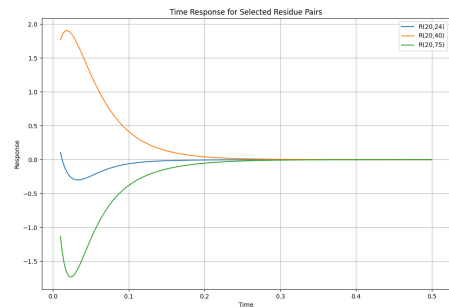


Figure 3.25. Multiple time response

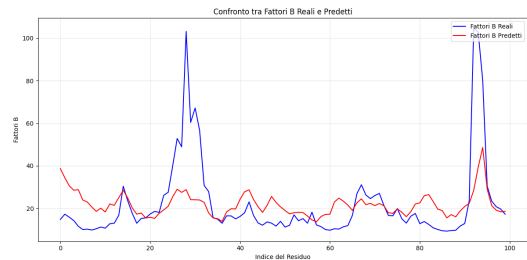


Figure 3.26. B factors

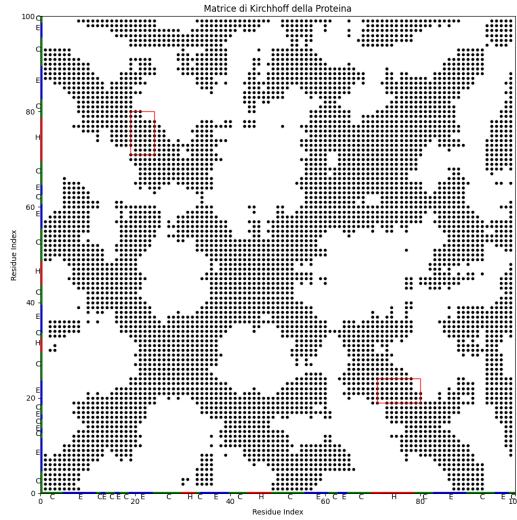


Figure 3.27. Kirchhoff



Figure 3.28. Seocndaria

3.16 Dinamica

$$\gamma \dot{x}_i = -g \sum_j K_{ij} x_j + \epsilon(t)(r_{21} - r_{76})\delta_{i,21} - \epsilon(t)(r_{21} - r_{76})\delta_{i,76} + \sqrt{2\gamma k_B T} \xi_i(t) \quad (3.37)$$

Ora posso simulare il moto Time dependent di ogni atomo e vedere come si propagano le vibrazioni. Gli autovalori sono smere le frequeze naturali del sistema.

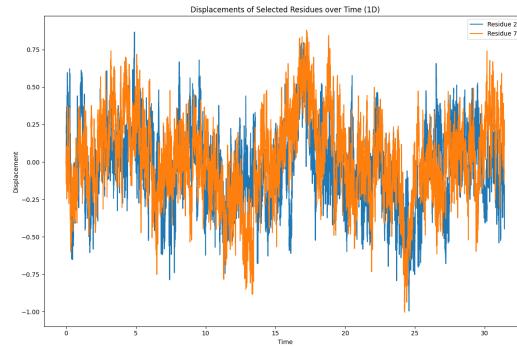


Figure 3.29. Seocndaria

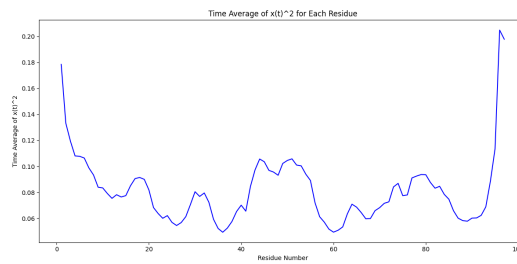


Figure 3.30. Seocndaria

3.17 Gradiente di temperature

$$\gamma_i \dot{x}_i = -g \sum_j K_{ij} x_j + \sqrt{2\gamma_i k_B T} \xi_i(t) \quad (3.38)$$

Gamma e T sono diventati vettori. Se la risolvi ottieni:

$$\sum_u e^{\exp(\cdot) \exp(\cdot) \exp(\cdot) (\sum_j K_{ij} u_j) \cdot \beta \cdot \beta^T}.$$

Ora hai diversi modi di scegliere i valori di T:

3.18 Troncamento Tmeperatura

se sono sotto 5 legami allora ho temperatura bassa. Se sono sopra allora ho temperatura piu' alta. Posso prendere banalmente 2 valori 0,1 .

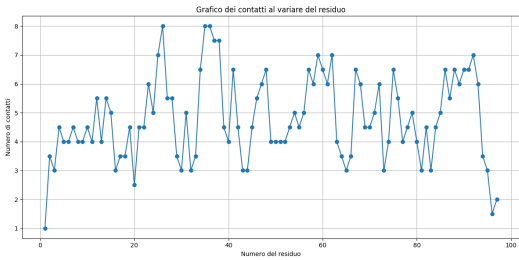


Figure 3.31. Temperatura Troncamento

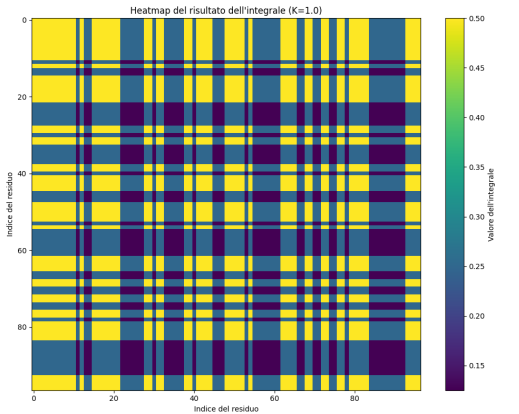


Figure 3.32. Temperatura Troncamento

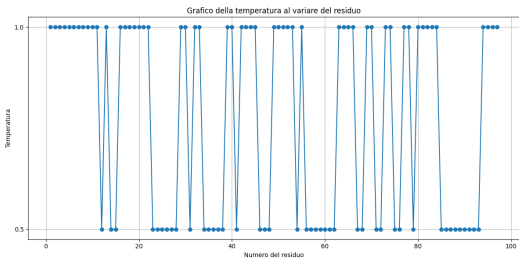


Figure 3.33. Temperatura Troncamento

3.19 Temperatura radiale

$$T(r)=T_0 + (Tb - T0)/R * r$$

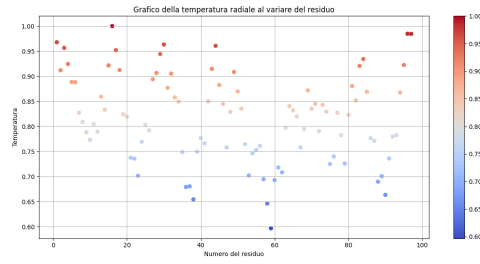


Figure 3.34. Temperatura Radiale

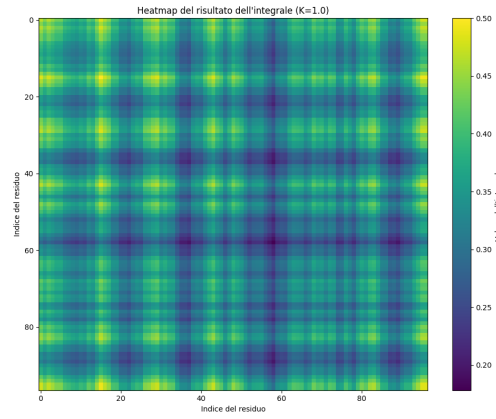


Figure 3.35. Correlazione Radiale

Risoluzione dell'equazione differenziale e calcolo della correlazione

Partiamo dall'equazione differenziale data per $x_i(t)$:

$$\frac{dx_i}{dt} = - \sum_j K_{ij} x_j + \sqrt{2k_B T} \eta_i + (1 - \cos((\omega t)) (\delta_{a_i} - \delta_{b_i}))$$

dove: - K_{ij} rappresenta la matrice di accoppiamento tra i nodi, - η_i è un processo di Wiener (rumore bianco gaussiano) con $\langle \eta_i(t) \eta_j(s) \rangle = \delta_{ij} \delta(t - s)$, - δ_{a_i} e δ_{b_i} sono costanti di offset che definiscono la parte oscillatoria del sistema.

Passaggio agli autovettori

Definiamo il cambiamento di base usando gli autovettori della matrice K , scrivendo:

$$x_i(t) = \sum_k V_{ik} Q_k(t)$$

dove V è la matrice degli autovettori di K , e $Q_k(t)$ rappresenta la dinamica lungo ciascun autovettore. L'equazione differenziale per $Q_k(t)$ diventa:

$$\frac{dQ_k}{dt} = -\lambda_k Q_k + C_k (1 - \cos((\omega t)) + \sqrt{2k_B T} \tilde{\eta}_k$$

con: - λ_k autovalori associati agli autovettori di K , - $C_k = \sum_i V_{ik}(\delta_{a_i} - \delta_{b_i})$, - $\tilde{\eta}_k = \sum_i V_{ik}\eta_i$ è il rumore proiettato sugli autovettori.

Soluzione dell'equazione differenziale per $Q_k(t)$

La soluzione formale per $Q_k(t)$ è data dalla somma di una parte deterministica e di una parte stocastica:

$$Q_k(t) = Q_k^{\text{det}}(t) + Q_k^{\text{rumore}}(t)$$

1. **Parte deterministica***: Integrando la parte deterministica, otteniamo

$$Q_k^{\text{det}}(t) = C_k \left(\frac{1 - e^{-\lambda_k t}}{\lambda_k} - \frac{e^{-\lambda_k t} - \cos((\omega t) + \frac{\lambda_k}{\omega} \sin((\omega t))}{\lambda_k^2 + \omega^2} \right)$$

2. **Parte stocastica***: La soluzione per la parte stocastica dovuta al rumore è

$$Q_k^{\text{rumore}}(t) = \sqrt{2k_B T} \int_0^t e^{-\lambda_k(t-u)} \tilde{\eta}_k(u) du$$

Calcolo della correlazione $\langle x_i(t)x_j(s) \rangle$

La correlazione media $\langle x_i(t)x_j(s) \rangle$ è:

$$\langle x_i(t)x_j(s) \rangle = \sum_{k,l} V_{ik}V_{jl} \langle Q_k(t)Q_l(s) \rangle,$$

dove $\langle Q_k(t)Q_l(s) \rangle$ può essere suddivisa nei contributi deterministico e stocastico:

$$\langle Q_k(t)Q_l(s) \rangle = \langle Q_k^{\text{det}}(t)Q_l^{\text{det}}(s) \rangle + \langle Q_k^{\text{rumore}}(t)Q_l^{\text{rumore}}(s) \rangle.$$

Contributo deterministico

La correlazione tra le parti deterministiche è:

$$\langle Q_k^{\text{det}}(t)Q_k^{\text{det}}(s) \rangle = C_k^2 (f_k(t,s) + g_k(t,s) \cos((\omega(t-s))) + h_k(t,s) \sin((\omega(t-s))),$$

dove: - $f_k(t,s)$ rappresenta il termine di decadimento esponenziale dato da:

$$f_k(t,s) = \frac{(1 - e^{-\lambda_k t})(1 - e^{-\lambda_k s})}{\lambda_k^2},$$

- $g_k(t,s)$ rappresenta il termine di correlazione coseno:

$$g_k(t,s) = \frac{e^{-\lambda_k(t+s)} - \cos((\omega t)) \cos((\omega s))}{\lambda_k^2 + \omega^2},$$

- $h_k(t,s)$ rappresenta il termine di correlazione seno:

$$h_k(t,s) = \frac{\lambda_k(\sin((\omega t)) \cos((\omega s)) + \cos((\omega t)) \sin((\omega s)))}{\lambda_k^2 + \omega^2}.$$

Contributo stocastico

Per la parte stocastica, otteniamo:

$$\langle Q_k^{\text{rumore}}(t) Q_k^{\text{rumore}}(s) \rangle = \frac{k_B T}{\lambda_k} \delta_{kl} e^{-\lambda_k |t-s|}.$$

Risultato finale

Combinando i termini, otteniamo la correlazione media finale per $\langle x_i(t) x_j(s) \rangle$:

$$\langle x_i(t) x_j(s) \rangle = \sum_k V_{ik} V_{jk} \left(C_k^2 (f_k(t, s) + g_k(t, s) \cos((\omega(t-s))) + h_k(t, s) \sin((\omega(t-s)))) + \frac{k_B T}{\lambda_k} e^{-\lambda_k |t-s|} \right)$$

Ogni termine è ora chiaramente identificato: - $C_k^2 f_k(t, s)$: componente di decadimento esponenziale, - $C_k^2 g_k(t, s) \cos((\omega(t-s)))$: componente oscillatoria con coseno, - $C_k^2 h_k(t, s) \sin((\omega(t-s)))$: componente oscillatoria con seno, - $\frac{k_B T}{\lambda_k} e^{-\lambda_k |t-s|}$: contributo del rumore termico.

Produzione di Entropia

La produzione di entropia S può essere calcolata utilizzando l'espressione data per la forza F e l'equazione differenziale associata al sistema.

Equazione Differenziale

Consideriamo l'equazione differenziale:

$$\frac{dx_i}{dt} = - \sum_j K_{ij} x_j + \sqrt{2k_B T} \eta_i + (1 - \cos((\omega t)))(\delta_{a_i} - \delta_{b_i}).$$

Espressione per F

L'espressione per F è data da:

$$F = \sum_j K_{ij} x_j + (1 - \cos((\omega t)))(\delta_{a_i} - \delta_{b_i}).$$

Produzione di Entropia

La produzione di entropia può essere espressa come:

$$S = \frac{F \cdot \frac{dx}{dt}}{T}.$$

Sostituendo F e $\frac{dx}{dt}$, otteniamo:

1. Sostituendo F :

$$F = \sum_j K_{ij} x_j + (1 - \cos((\omega t)))(\delta_{a_i} - \delta_{b_i}).$$

2. Sostituendo $\frac{dx}{dt}$:

$$\frac{dx_i}{dt} = - \sum_k K_{ik} x_k + \sqrt{2k_B T} \eta_i + (1 - \cos((\omega t)))(\delta_{a_i} - \delta_{b_i}).$$

Produzione di Entropia Completa

Pertanto, la produzione di entropia può essere scritta come:

$$S = \frac{1}{T} \left[\sum_j K_{ij} x_j \left(- \sum_k K_{ik} x_k + \sqrt{2k_B T} \eta_i + (1 - \cos((\omega t)))(\delta_{a_i} - \delta_{b_i}) \right) + (1 - \cos((\omega t)))(\delta_{a_i} - \delta_{b_i}) \right]$$

Considerazioni sulle Semplificazioni

Esploriamo ora alcune possibili semplificazioni dell'espressione per la produzione di entropia:

1. **Condizioni di Equilibrio:** Se il sistema è in equilibrio termico, potremmo avere $\langle x_j \rangle$ costante, semplificando il calcolo della somma sui termini di $K_{ij} x_j$.
2. **Media Temporale:** Considerando la media temporale di S , alcuni termini oscillatori (come quelli con $\cos((\omega t))$) potrebbero cancellarsi nel lungo termine, specialmente se le oscillazioni sono simmetriche attorno a zero.
3. **Domini di Oscillazione:** Se i contributi oscillatori non sono dominanti, potremmo trascurarli per un'analisi qualitativa.
4. **Assunzione di Piccole Oscillazioni:** Se le oscillazioni sono piccole, possiamo linearizzare i termini, ma questa assunzione potrebbe non essere sempre valida a seconda dell'ampiezza delle oscillazioni.

Conclusioni

In generale, mentre l'espressione per la produzione di entropia è complessa, alcune semplificazioni possono essere applicabili in condizioni specifiche, come stati di equilibrio o considerando medie temporali. Tuttavia, senza ulteriori informazioni sulle dinamiche del sistema o sul comportamento dei termini, non è facile ottenere una forma significativamente più semplice dell'espressione.

3.20 Fokker planck N temperature

L'equazione di Fokker-Planck associata all'equazione differenziale

$$\frac{dx_i}{dt} = - \sum_j K_{ij} x_j + B \eta_i$$

è data da:

$$\frac{\partial P(x, t)}{\partial t} = \sum_i \frac{\partial}{\partial x_i} \left[\sum_j K_{ij} x_j P \right] + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} [B_{ij} \delta_{ij} P]$$

dove:

- $P(x, t)$ è la densità di probabilità del sistema,
- $A_i = -\sum_j K_{ij}x_j$ è il termine drift,
- $D_{ij} = B_{ii}\delta_{ij}$ è la matrice di diffusione.

Soluzione dell'Equazione Differenziale

Consideriamo l'equazione differenziale iniziale:

$$\frac{\partial P}{\partial t} = \sum_i \left(\sum_j K_{ij}x_j P \right) + 0.5 \sum_{i,j} \left(\frac{\partial^2}{\partial x_i \partial x_j} (B_{ij}P) \right).$$

Dove $P = P(x, t)$, K è la matrice di Kirchhoff e B è una matrice diagonale. Supponiamo una condizione iniziale:

$$P(x, 0) = \delta(x - x_0).$$

Trasformata di Fourier

Applicando la trasformata di Fourier rispetto a x :

$$P(k, t) = \int_{-\infty}^{+\infty} P(x, t) e^{-ikx} dx.$$

La trasformata dell'equazione diventa:

$$\frac{\partial P(k, t)}{\partial t} = \int_{-\infty}^{+\infty} \left(\sum_i \left(\sum_j K_{ij}x_j P \right) \right) e^{-ikx} dx + 0.5 \int_{-\infty}^{+\infty} \left(\sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (B_{ij}P) \right) e^{-ikx} dx.$$

Primo Termine: Derivata e Trasformata

Focalizziamoci sul primo termine:

$$\frac{\partial P(k, t)}{\partial t} = \sum_i (K_{ii}P(k, t)) + \sum_{i,j} K_{ij} \left(\int_{-\infty}^{+\infty} x_j P(x, t) e^{-ikx} dx \right).$$

Per il secondo termine, utilizziamo la proprietà della trasformata di Fourier per la derivata:

$$\int_{-\infty}^{+\infty} \frac{\partial^2 P}{\partial x_i \partial x_j} e^{-ikx} dx = -k_i k_j P(k, t).$$

Quindi, l'equazione diventa:

$$\frac{\partial P(k, t)}{\partial t} = \sum_i K_{ii}P(k, t) + \sum_{i,j} K_{ij} \left(-i \frac{\partial P(k, t)}{\partial k} \right) - 0.5 \sum_{i,j} B_{ij} k_i k_j P(k, t).$$

Semplificazione

Riorganizzando l'equazione, otteniamo:

$$\frac{\partial P(k, t)}{\partial t} = \sum_i K_{ii} P(k, t) + \sum_{i,j} K_{ij} (-ik_j P(k, t)) - 0.5 \sum_{i,j} B_{ij} k_i k_j P(k, t).$$

Definiamo $\Lambda(k)$:

$$\Lambda(k) = \sum_i K_{ii} + \sum_{i,j} K_{ij} (-ik_j) - 0.5 \sum_{i,j} B_{ij} k_i k_j.$$

Quindi, l'equazione diventa:

$$\frac{\partial P(k, t)}{\partial t} = \Lambda(k) P(k, t).$$

Risoluzione dell'Equazione Differenziale

La soluzione dell'equazione differenziale è:

$$P(k, t) = P(k, 0) e^{\Lambda(k)t}.$$

Utilizzando la condizione iniziale $P(k, 0) = e^{-ikx_0}$:

$$P(k, t) = e^{-ikx_0} e^{\Lambda(k)t}.$$

Trasformata Inversa

Per trovare $P(x, t)$, applichiamo la trasformata di Fourier inversa:

$$P(x, t) = \int_{-\infty}^{+\infty} P(k, t) e^{ikx} dk.$$

Sostituendo $P(k, t)$:

$$P(x, t) = \int_{-\infty}^{+\infty} e^{-ikx_0} e^{\Lambda(k)t} e^{ikx} dk.$$

Fattorizzando:

$$P(x, t) = e^{-ix_0 k} \int_{-\infty}^{+\infty} e^{ik(x-x_0)} e^{\Lambda(k)t} dk.$$

Assumiamo una forma specifica per $\Lambda(k)$:

$$\Lambda(k) = -\gamma k^2,$$

dove γ è una costante positiva. Allora, otteniamo:

$$P(x, t) = \int_{-\infty}^{+\infty} e^{ik(x-x_0)} e^{-\gamma k^2 t} dk.$$

Risoluzione dell'Integrale

Utilizziamo la formula dell'integrale gaussiano:

$$P(x, t) = \sqrt{\frac{\pi}{\gamma t}} e^{-\frac{(x - x_0)^2}{4\gamma t}}.$$

Risultato Finale

La soluzione finale è:

$$P(x, t) = \sqrt{\frac{\pi}{\gamma t}} e^{-\frac{(x - x_0)^2}{4\gamma t}}.$$

Chapter 4

Conclusions

Bibliography