



UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE CHIMICHE

CORSO DI LAUREA IN CHIMICA

Trasferimento di entropia nella plexina monomero e dimero da simulazioni di dinamica molecolare nella scala del microsecondo

Relatore: Prof. Mirco Zerbetto

Laureanda: Cristina Chemello
1216564

Anno Accademico 2021/2022

Indice

1	Introduzione	1
1.1	Effetti allosterici	1
1.2	Scopo della tesi	6
2	Metodi	7
2.1	Simulazioni di dinamica molecolare	7
2.2	Trasferimento di entropia	9
3	Risultati	13
4	Conclusioni	19
5	Bibliografia	21

1 Introduzione

1.1 Effetti allosterici

Genesi del termine e definizione

Il termine “effetti allosterici” è stato introdotto per la prima volta da Monod, Jacob e Changeux [1] nel 1961 per descrivere il fenomeno secondo il quale un metabolita è in grado di bloccare l’attività catalitica di un enzima senza mostrare evidenti somiglianze strutturali con il substrato. Nel 1963 gli stessi specificarono che l’aggettivo allosterico si riferisce al sito di legame per il metabolita (effettore allosterico) che legandosi è in grado di portare ad alterazioni reversibili della struttura della proteina (transizione allosterica), che possono comportare variazioni del sito attivo e quindi compromettere l’attività catalitica. Vennero in seguito sviluppati due modelli [2], quello concertato di Monod-Wyman-Changeux (MWC) e quello sequenziale di Koshland-Nemethy-Filmer (KNF). Entrambi considerano solamente due configurazioni principali della proteina: nel modello MWC la proteina è presente in equilibrio tra le due configurazioni, ma la presenza del substrato stabilizza la configurazione con cui esso ha maggiore affinità spostando verso di essa l’equilibrio (*cooperative ligand binding*), nel modello KNF invece il cambiamento di configurazione è indotto dal legame con il substrato (*induced fit*). Nel 1984 Cooper e Dryden [3] avanzarono l’ipotesi che l’allosteria non si manifestasse solamente attraverso una variazione della conformazione media della proteina, ma anche attraverso una variazione della dinamica, cioè nelle frequenze e ampiezze delle fluttuazioni termiche degli atomi della proteina a seguito dell’attacco del legante, effetti essenzialmente di natura entropica. In seguito, Hilser et al. [4, 5], proposero l’*ensemble allostery model* (EAM) secondo il quale, siccome nello spazio conformazionale di una proteina sono presenti configurazioni più probabili e meno probabili a seconda della loro energia, la coordinazione di un legante cambierebbe il peso statistico relativo delle conformazioni della proteina e quindi avrebbe l’effetto di rimodellare la superficie di energia potenziale. Questa è caratterizzata generalmente da numerosi minimi locali ad energia molto simile tra loro. Il legante può indurre perturbazioni differenti legandosi a conformazioni diverse della proteina. Conformazioni con energia paragonabile sono similmente probabili e questo assicura la pluralità di percorsi allosterici. In conclusione, gli effetti allosterici si possono manifestare sotto forma di variazioni conformazionali o della dinamica della proteina.

Inizialmente sono stati studiati solamente gli effetti allosterici in oligomeri simmetrici con almeno due conformazioni, una ad alta affinità con il legante e una a bassa affinità. Con l’allargarsi della definizione di effetto

allosterico a tutte quelle perturbazioni che producono una variazione della distribuzione conformazionale della proteina o della sua dinamica, si è aperta anche la possibilità che tutte le proteine dinamiche [6, 7] (non fibrose) manifestino effetti allosterici. Esempi di eventi perturbativi sono il legame con un legante, una modifica post-traduzionale, una mutazione, un'interazione lipide-proteina. La possibilità di trattare come perturbazioni [8] questi effetti allosterici deriva dal fatto che sono energeticamente modesti, comparati all'energia degli oggetti che regolano (proteine).

Importanza biologica

La trasmissione di segnali è essenziale nella cellula e in essa giocano un ruolo fondamentale gli effetti allosterici nelle proteine di membrana [9]. Le GPCR (*G protein-coupled receptors*) sono per esempio una grande famiglia di proteine di membrana che si occupano di trasmettere i segnali all'interno della cellula da una serie di messaggeri molecolari esterni, attraverso l'attivazione di proteine G e altre molecole segnale. Le GPCR sono caratterizzate da una pluralità di percorsi allosterici perché vi sono diverse conformazioni attive, stabilizzate da diversi effettori allosterici (*ligand or agonist bias*). L'importanza degli effetti allosterici non si riduce alla trasmissione di segnali, ma influenza in maniera importante anche la catalisi [10]. La distribuzione conformazionale di un enzima comprende conformazioni in cui esso è attivo e altre in cui non lo è. L'effettore allosterico è in grado di variare la popolazione dei vari stati conformazionali e quindi promuovere o inibire la catalisi. Un esempio è la famiglia di proteine tirosin-fosfatasi (PTP) che catalizza la defosforilazione post traduzionale. Tra queste è stata studiata la PTP1B che rimuove il fosfato tramite due strutture, il P-loop e l'*acid loop*. L'*acid loop* può assumere due conformazioni, una aperta che permette al substrato di accedere al sito attivo e una chiusa che permette all'acido aspartico di posizionarsi correttamente perché avvenga la catalisi. Un inibitore allosterico, quindi, deve agire in modo da bloccare l'*acid loop* nella sua conformazione aperta e lo fa variando e indebolendo le interazioni con i siti circostanti. La conferma di ciò è data non solo dall'analisi cristallografica [10] ma anche da simulazioni di dinamica molecolare [11]. Se da un lato gli effetti allosterici sono fondamentali per il funzionamento dei sistemi biologici, dall'altro la mutazione di residui chiave può sconvolgere il loro effetto provocando quelle che vengono definite malattie allosteriche. Un campo molto importante di studio è quello dei farmaci allosterici che hanno il vantaggio di essere più specifici dei farmaci ortosterici (che legano il sito attivo), di agire a distanza e di modulare l'attività della proteina senza eliminarla completamente [12].

Studio dei fenomeni allosterici

I fenomeni allosterici vengono studiati con un approccio multidisciplinare che comprende tecniche sperimentali e metodi computazionali. Tra le tecniche sperimentali [2] più utilizzate ci sono la cristallografia ai raggi X, che permette di studiare le conformazioni principali, e l’NMR (*Nuclear Magnetic Resonance*), che permette di avere informazioni anche sulla dinamica delle transizioni conformazionali. Le tecniche sperimentali sono fondamentali anche per ottenere i dati necessari a migliorare i metodi computazionali, per esempio i dati NMR sono spesso utilizzati per parametrizzare il campo di forze utilizzato per le simulazioni di dinamica molecolare delle proteine. Le simulazioni di dinamica molecolare permettono di ottenere informazioni a livello atomistico o molecolare sulla struttura e sulla dinamica della proteina, anche a seguito di perturbazioni.

In una simulazione di dinamica molecolare [2, 13], si descrive l’evoluzione classica dei nuclei di un sistema molecolare, i quali interagiscono con un potenziale internucleare detto campo di forze (*force field*). La traiettoria si ottiene risolvendo in maniera numerica le equazioni di Newton, note le posizioni e le velocità ad un dato istante di tutti i nuclei. Sono molto utili perché da esse si possono ottenere le posizioni degli atomi nel tempo e le condizioni possono essere controllate per simulare diverse perturbazioni del sistema ed osservarne gli effetti sulla struttura e sulla dinamica della proteina. Le perturbazioni applicabili possono essere forze meccaniche, potenziali di membrana, protonazione/deprotonazione o variazione di residui (mutagenesi). Si ricavano informazioni sui siti di legame dei leganti e come gli stimoli in una zona della proteina vengano propagati ad altre zone. Per fare ciò si osservano i movimenti indipendenti di residui tra e in corrispondenza dei siti di interesse, misurando le distanze, gli angoli diedri o utilizzando altri metodi di analisi come il trasferimento di entropia.

Lo studio dell’allosteria nelle proteine rappresenta una sfida per le simulazioni di dinamica molecolare perché gli eventi interessanti avvengono nella scala dei nanosecondi, microsecondi e addirittura tempi più lunghi, che rendono le simulazioni computazionalmente costose da un punto di vista temporale e di spazio di archiviazione. A questo scopo si sono però studiate approssimazioni in grado di ridurre il tempo richiesto, come per esempio le simulazioni *coarse grained*, la scelta di considerare solamente le coordinate ritenute importanti per il processo oppure i metodi di *enhanced sampling* (*metadynamics*, *GaMD*, etc...). Altri metodi computazionali [2] comprendono i *Markov State Models* (MSM), il NMA (*Normal Mode Analysis*), PCA (*Principal Component Analysis*) ed *elasticity-based methods* implementati in una serie di software [14], utilizzabili per esempio per lo studio di possibili farmaci allosterici. Vengono utilizzati su dati ottenuti sperimentalmente o

da simulazioni di dinamica molecolare.

Il trasferimento di entropia applicato allo studio degli effetti allosterici

Il trasferimento di entropia [15, 16], permette di identificare i residui che guidano le fluttuazioni delle posizioni degli altri e che quindi svolgono un ruolo fondamentale per gli effetti allosterici. Definendo $p(\Delta R_i(t))$ e $p(\Delta R_j(t))$ le densità di probabilità marginali di osservare fluttuazioni delle posizioni degli atomi i e j , se i due processi sono indipendenti allora le fluttuazioni di j al tempo $t_0 = 0$ non influenzeranno quelle di i ad un tempo successivo, t , e sarà valida la relazione

$$p(\Delta R_i(t)|\Delta R_i(0)) = p(\Delta R_i(t)|\Delta R_i(0), \Delta R_j(0)) \quad (1)$$

Si definisce il trasferimento di entropia come il confronto, tramite l'entropia relativa (o divergenza di Kullback-Leibler), tra la densità di probabilità condizionale che si manifesti $\Delta R_i(t)$ dati gli stati precedenti di i e j e la densità di probabilità condizionale che si manifesti $\Delta R_i(t)$ dati solamente gli stati precedenti di i . Il trasferimento di entropia non è simmetrico, essendo basato sulle densità di probabilità condizionali, e può essere utilizzato per trovare relazioni di causa-effetto. L'esistenza di una relazione causa-effetto è cruciale nello sviluppo di farmaci allosterici perché permette di sapere quali residui manipolare per controllare l'attività della proteina. Il trasferimento di entropia può essere calcolato a partire dalla definizione di Boltzmann come

$$T_{i \rightarrow j}(t) = \sum_{i,j=1}^N p(\Delta R_i(t), \Delta R_i(0), \Delta R_j(0)) \ln \left(\frac{p(\Delta R_i(t)|\Delta R_i(0), \Delta R_j(0))}{p(\Delta R_i(t)|\Delta R_i(0))} \right) \quad (2)$$

Passando a coordinate continue e utilizzando il teorema di Bayes la (2) può essere scritta come

$$T_{i \rightarrow j}(t) = - \langle \ln p(\Delta R_j(0), \Delta R_j(t)) \rangle + \langle \ln p(\Delta R_i(0), \Delta R_j(0), \Delta R_j(t)) \rangle + \langle \ln p(\Delta R_j(0)) \rangle - \langle \ln p(\Delta R_i(0), \Delta R_j(0)) \rangle \quad (3)$$

Le fluttuazioni dei residui si possono quantificare tramite funzioni di correlazione C_{ij} . Il tempo di correlazione delle fluttuazioni, definito come il tempo necessario alla funzione di correlazione per raggiungere $1/e$ del suo valore iniziale, è dell'ordine degli 1-10 ns [17]. Si dice che il residuo j guida i se il tempo di decadimento di C_{ji} è maggiore di C_{ij} .

Il calcolo del trasferimento di entropia viene eseguito utilizzando i dati sulle traiettorie ricavati da simulazioni di dinamica molecolare. Un esempio

del suo utilizzo è l'analisi dei percorsi allosterici nella chinasi ERK2 [18] (enzima in grado di trasferire gruppi fosfato). Sono stati individuati i residui coinvolti maggiormente nel trasferimento di entropia nelle due conformazioni principali, attiva e inattiva, si sono osservati come cambiano i percorsi allosterici a seguito di perturbazioni (mutazioni/legame con ATP) e si è osservato che la regione N-terminale è accoppiata ai siti di legame del substrato da molti di questi. Anche l'Ubiquitina [19] (proteina che svolge un ruolo fondamentale nella degradazione di altre proteine) è stata studiata in questo modo per identificare quali residui si comportano prevalentemente da donatori (*source*) o accettori (*sink*) di entropia, la presenza di percorsi allosterici e l'effetto della formazione di complessi su di essi.

Plexina-B1

La Plexina B1 [20] fa parte della famiglia delle Plexine (A1, B1, C1 e D1), recettori per le semaforine che, insieme ad altre molecole, svolgono un ruolo fondamentale nel guidare la crescita degli assoni verso i propri target durante lo sviluppo del sistema nervoso. La Plexina B1 possiede anche un sito di legame per le Rho GTPasi [21], interruttori molecolari importanti per la riorganizzazione del citoscheletro, necessarie allo sviluppo degli assoni. Il sito di legame per le Rho GTPasi (RBD) è anche in grado di legarsi al suo omologo di un'altra Plexina B1 e formare un omodimero. La dinamica del sito RBD [21] della Plexina B1 è stata molto studiata in passato [22, 23]. Gli studi suggeriscono un irrigidimento globale a seguito della dimerizzazione soprattutto per quanto riguarda il loop L_4 , anche se localmente ci sono zone che diventano più flessibili. Vi è poi una riduzione della correlazione tra i moti dei residui a seguito della dimerizzazione, individuata tramite il calcolo della matrice di correlazione degli angoli diedri del *backbone*, che indica anche la presenza di *signal pathways* specifici, confermati dalla *Dynamical Network Analysis*. Il calcolo del trasferimento di entropia individua invece il motore del processo di dimerizzazione nel loop L_4 e nell'estremità C terminale, mentre il moto del dimero è guidato prevalentemente dall'estremità N terminale (con il contributo di alcuni residui di L_1 , L_2 e della regione C terminale).

Ruolo dell'asimmetria nelle proteine dimeriche

In alcuni omodimeri, come per esempio FAcD [24] (un enzima che catalizza la rottura dei legami carbonio-alogeno), solamente uno dei due monomeri mantiene l'attività catalitica. L'asimmetria si osserva da un punto di vista conformazionale sotto forma di due conformazioni medie diverse dell'interfaccia tra i monomeri e il sito attivo, investigate tramite FTX (*freeze-trapping x-ray crystallography*) visto che la velocità della catalisi enzimatica è lenta. Si

è inoltre osservata un'asimmetria nella dinamica dei due monomeri che promuove la catalisi compensando la perdita di entropia nel monomero che lega il substrato, con un aumento della dinamica dell'altro e l'espulsione di molecole di acqua coordinate. Quest'asimmetria è però mediata in soluzione grazie ad un rapido scambio di ruoli dei due monomeri, evidenziato dall'assenza di differenze di *chemical shift*. Un tipo di reattività simile, detta *half-of-the-sites*, è propria anche dell'enzima 5-amminolevulinato sintasi (ALAS) [25], importante nella biosintesi dell'eme. Esso forma un omodimero che presenta differenze nel numero di residui che formano β strands a temperature diverse e una conseguente asimmetria, confermata dalle simulazioni MD anche in assenza di leganti.

1.2 Scopo della tesi

L'obiettivo è calcolare il trasferimento di entropia tra i residui del sito RBD della Plexina B1 monomero e dimero sfruttando le simulazioni di dinamica molecolare della lunghezza del microsecondo calcolate dal gruppo del Prof. M. Buck, del dipartimento di fisiologia e biofisica della Case Western Reserve University, Cleveland, OH, USA [23]. Si è considerato l'effetto della lunghezza della traiettoria, nello specifico 100% e 50% della lunghezza totale, sui risultati ottenuti e il raggiungimento o meno della convergenza. In secondo luogo, si è analizzato il comportamento dei monomeri del dimero, allo scopo di evidenziare un'asimmetria "transiente" tra le due unità RBD, dovuta a variazioni conformazionali lente nell'ordine dei millisecondi, e studiarne il meccanismo. Precedenti analisi delle traiettorie, difatti, hanno evidenziato come le due unità siano caratterizzate da regimi strutturali e dinamici differenti. Data la simmetria del legame tra le due unità, tale asimmetria è inaspettata e quindi è attribuita alla breve durata della traiettoria rispetto alla scala dei tempi delle variazioni conformazionali sopra citate. Il calcolo del trasferimento di entropia può far luce sul meccanismo di questo transiente.

2 Metodi

2.1 Simulazioni di dinamica molecolare

Le simulazioni di dinamica molecolare [13] considerano solamente i nuclei, i quali vengono trattati in maniera classica. Sono in grado di fornire la traiettoria seguita dagli atomi in esame, a partire dai momenti e dalle coordinate iniziali, risolvendo le equazioni di Newton. Inizialmente è necessario specificare la topologia della proteina in esame (la sua connettività) e gli *atom types* che identificano gli atomi che condividono numero atomico e proprietà energetiche legate all'ibridizzazione, al numero di legami, nel caso dei metalli anche lo stato di ossidazione. I parametri associati a ciascun *atom type* e determinati sperimentalmente o da calcoli quantomeccanici, permettono di implementare i vari contributi al campo di forze. Il campo di forze è costituito da tre contributi che descrivono le interazioni covalenti (*stretching*, *bending* e torsione) e due che descrivono le interazioni non covalenti (potenziale Coulombiano e di Lennard-Jones). Un ulteriore contributo (*improper dihedrals*) è presente se la molecola deve mantenere determinate caratteristiche (es. planarità).

Specificato il campo di forze e i suoi parametri si procede alla scelta dell'*ensemble* a seconda delle informazioni che si vogliono ricavare dalla simulazione. I tipi di *ensemble* sono il canonico (NVT), il microcanonico (NVE) e il grancanonico (μVT) con N numero di particelle, V volume, T temperatura assoluta, E energia totale e μ potenziale chimico del sistema. Per ridurre gli effetti di superficie si scelgono le condizioni periodiche al contorno che consistono nel circondare il sistema con una serie di repliche di sé stesso, come se esso fosse un'unità di un sistema periodico. Si stabiliscono i parametri della simulazione, come per esempio i *cutoff* per il calcolo delle interazioni non di legame e il *time step* di integrazione. Il *time step* viene scelto sulla base dei tempi caratteristici dei fenomeni da osservare, generalmente due ordini di grandezza minore dei moti più veloci. Infine, si stabiliscono le velocità, per esempio campionandole dalla distribuzione di Maxwell.

Il protocollo che viene generalmente eseguito è il seguente: in primo luogo si minimizza l'energia, successivamente si esegue un riscaldamento controllato per portare il sistema alla temperatura di lavoro. Durante questo processo è bene verificare che anche la proteina abbia subito il riscaldamento e non solo il solvente (acqua) presente in eccesso, questa porzione della traiettoria è detta periodo di equilibrizzazione e viene scartata (1-10 ns). Applicati questi accorgimenti il sistema è pronto e si può calcolare la traiettoria. La lunghezza della stessa e la frequenza di campionamento vanno stabilite sulla base dei tempi caratteristici dei processi che si vogliono osservare.

Un problema fondamentale in dinamica molecolare è verificare la convergenza di una simulazione, cioè verificare di aver ottenuto abbastanza punti dello spazio delle fasi da poter dire di averlo esplorato in maniera ergodica. È impossibile da dimostrare, ma da un punto di vista statistico si può dire di aver raggiunto la convergenza quando le proprietà statistiche rimangono pressoché costanti al proseguire della simulazione, quindi all'aumentare del numero di punti acquisiti nello spazio delle fasi. Ovviamente, siccome la lunghezza delle simulazioni è limitata, è possibile che il sistema rimanga bloccato in uno stato metastabile con un tempo di vita più lungo e che quindi la convergenza sia solo apparente. Il problema viene generalmente risolto facendo diverse simulazioni a partire da strutture differenti, anche se da un punto di vista computazionale è un metodo costoso e comunque non vi è certezza di aver raggiunto la convergenza. Un problema analogo si presenta quando la molecola possiede più di una conformazione stabile e vi è una barriera di energia tra queste che è molto poco probabile venga superata durante una simulazione. Oltre ad eseguire una simulazione per ogni conformazione come visto sopra, si può per esempio scaldare periodicamente il sistema a temperatura molto alta per rendere più frequente il cambio di conformazione e successivamente raffreddarlo e raccogliere i dati (chiamato *simulated annealing* quando si sta cercando il minimo assoluto) o abbassare artificialmente la barriera (*accelerated MD*).

L'utilità delle simulazioni di dinamica molecolare nello studio degli effetti allosterici [26], ma anche nella ricerca di candidati farmaci, sta nel determinare i siti di legame, anche se sono nascosti inizialmente, trovare le interazioni tra il legante e il sito attivo e le variazioni di conformazione che queste comportano, determinare come gli stimoli vengono trasmessi tra zone distanti della molecola attraverso lo studio della sua dinamica e l'applicazione di perturbazioni controllate. Nell'analizzare i risultati non bisogna dimenticare che le simulazioni si basano su una serie di approssimazioni [27]. In primo luogo, i *force fields* (AMBER, CHARMM, OPLS...) sono approssimazioni anche se stanno diventando sempre più accurati. In secondo luogo, anche la disponibilità di buone strutture sperimentali influenza la simulazione. Il fatto che non si possano rompere i legami covalenti comporta, invece, che si debbano fare simulazioni diverse per tener conto della protonazione per esempio. In ultimo, il tempo limitato influenza il tipo di processi osservabili con una singola simulazione. A questo scopo si utilizzano i metodi di *enhanced sampling* [28], alcuni già citati prima, come *steered MD* [29], *targeted MD* [30], *biased MD* [31], *metadynamics* [32]. In alternativa si utilizzano simulazioni *coarse-grained* [33] al posto di quelle *all-atom*, raggruppando atomi adiacenti in *beads* e considerando le interazioni tra di essi.

La simulazione MD utilizzata per il calcolo del trasferimento di entropia

si è svolta nelle seguenti condizioni [23]. Il sito RBD della Plexina-B1 è costituito da 121 residui. Il sistema è stato preparato con il programma CHARMM: il sito RBD è stato posto ad almeno 12 Å dal bordo della box rettangolare, il solvente (acqua) è stato trattato in maniera esplicita (TIP3P) e si sono aggiunti controioni per neutralizzare e portare la concentrazione di NaCl a 0.15 M. La simulazione è stata preceduta da una breve fase di minimizzazione e un riscaldamento controllato con 20 ns di equilibrizzazione alla temperatura di 300 K e alla pressione di 1 atm svolti con NAMD [34]. Sono state applicate le condizioni periodiche al contorno e si è utilizzato il metodo di Ewald [34] per calcolare le interazioni Coulombiane a lungo raggio. Si sono mantenuti fissi i legami con H tramite l'algoritmo SHAKE, si è impostato il *time step* di integrazione a 2 fs, si è impostato il *cutoff* per le interazioni di non legame a 12 Å e infine è stato usato il campo di forze CHARMM 22 con le correzioni CMAP [35]. La simulazione di produzione, lunga 1 μ s, è stata calcolata con il computer dedicato "Anton", che si trova presso la Columbia University di New York [36].

2.2 Trasferimento di entropia

Descrizione della dinamica

Al fine di calcolare il trasferimento di entropia è necessario stabilire un modello per la descrizione della dinamica della proteina. Assumendo che la dinamica lenta sia sostanzialmente dovuta al *backbone*, in una visione *coarse-grained*, si descrive la sua dinamica come quella di siti virtuali convenzionalmente posizionati in corrispondenza dei C_α di ogni amminoacido. Per semplicità si parla di dinamica dei C_α , riferendosi però in realtà ai siti virtuali definiti sopra. Si è definita $\Delta\mathbf{R} = [\Delta R_1, \Delta R_2, \dots, \Delta R_N]^{tr}$ la matrice colonna che raccoglie gli spostamenti degli N atomi C_α rispetto alle posizioni di equilibrio.

$$\Delta R_j = \sqrt{(X_j - X_j^{(eq)})^2 + (Y_j - Y_j^{(eq)})^2 + (Z_j - Z_j^{(eq)})^2} \quad (4)$$

Sulla base di questa approssimazione la dinamica dei C_α è stocastica e soggetta a un rumore bianco. L'equazione

$$\Delta\dot{\mathbf{R}} = \boldsymbol{\xi}^{-1} \frac{\partial U(\Delta\mathbf{R})}{\partial \Delta\mathbf{R}} + \mathbf{n}(t) \quad (5)$$

è la traiettoria browniana, dove $U(\Delta\mathbf{R})$ è il potenziale di campo medio (frutto dell'integrazione rispetto alle coordinate non considerate), $\boldsymbol{\xi}$ è il tensore di attrito e $\mathbf{n}(t)$ è il rumore bianco (media nulla e varianza $\langle n_i(t)n_j(t') \rangle =$

$\delta(t - t')2mk_{\text{B}}T\xi_{i,j}$ dal teorema di fluttuazione e dissipazione di Einstein). A questo punto si applica l'approssimazione quasi-armonica per cui

$$U(\Delta\mathbf{R}) \approx \frac{k_{\text{B}}T}{2} \Delta\mathbf{R}^{\text{tr}} \mathbf{C}^{-1} \Delta\mathbf{R} \quad (6)$$

dove k_{B} è la costante di Boltzmann, T è la temperatura assoluta e \mathbf{C} la matrice di varianza-covarianza $C_{i,j} = \langle \Delta R_i \Delta R_j \rangle$. Si assume, inoltre, che il tensore di attrito sia diagonale e che tutti gli elementi siano uguali all'attrito idrodinamico traslazionale per una sfera di raggio $R = 2 \text{ \AA}$ con condizioni al contorno *stick* ($C = 6$). Si definisce quindi $\boldsymbol{\xi} = 6\pi R\eta(T)\mathbf{1}_N = \xi_0\mathbf{1}_N$, con $\eta(T)$ viscosità dell'acqua $\mathbf{1}_N$ matrice identità $N \times N$. La traiettoria browniana (5) descrive traiettorie stocastiche, cioè che partendo dallo stesso punto $\Delta\mathbf{R}(0)$ evolvono casualmente in maniera diversa nel tempo. La probabilità di trovare il sistema nella configurazione $\Delta\mathbf{R}(t)$ al tempo t è data da $p(\Delta\mathbf{R}(t), t | \Delta\mathbf{R}(0), 0)$. Nelle condizioni in cui è valida l'equazione (5), il sistema è markoviano (stazionario) e gaussiano (data la forma del potenziale di campo medio). All'equilibrio termico, per qualsiasi tempo, la densità di probabilità vale

$$p(\Delta\mathbf{R}(t)) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C}(0)|}} e^{-\frac{1}{2} \Delta\mathbf{R}(t)^{\text{tr}} \mathbf{C}(0)^{-1} \Delta\mathbf{R}(t)} = p_{\text{eq}}(\Delta\mathbf{R}) \quad (7)$$

con $\mathbf{C}(0) = \langle \Delta\mathbf{R}(0) \otimes \Delta\mathbf{R}(0) \rangle$. La probabilità congiunta a due tempi è invece data come

$$p(\Delta\mathbf{R}(t), \Delta\mathbf{R}(0)) = \frac{e^{-\frac{1}{2} \begin{bmatrix} \Delta\mathbf{R}(t) \\ \Delta\mathbf{R}(0) \end{bmatrix}^{\text{tr}} \begin{bmatrix} \mathbf{C}(0) & \mathbf{C}(t) \\ \mathbf{C}(t) & \mathbf{C}(0) \end{bmatrix}^{-1} \begin{bmatrix} \Delta\mathbf{R}(t) \\ \Delta\mathbf{R}(0) \end{bmatrix}}}{\sqrt{(2\pi)^{2N} \begin{vmatrix} \mathbf{C}(0) & \mathbf{C}(t) \\ \mathbf{C}(t) & \mathbf{C}(0) \end{vmatrix}}} \quad (8)$$

con $\mathbf{C}(t) = \langle \Delta\mathbf{R}(0) \otimes \Delta\mathbf{R}(t) \rangle$. Le funzioni di correlazione $C_{i,j}(t)$ possono essere facilmente calcolate se si diagonalizza la matrice

$$\mathbf{C}(0) = \mathbf{V} \boldsymbol{\Lambda} \mathbf{V}^{\text{tr}} \quad (9)$$

con \mathbf{V} la matrice unitaria degli autovettori, e $\boldsymbol{\Lambda}$ la matrice degli autovalori. Per un generico elemento della matrice di correlazione [17] si ha

$$C_{i,j}(t) = \sum_{n=1}^N \lambda_n V_{i,n} V_{j,n} e^{-\frac{k_{\text{B}}T}{\lambda_n \xi_0} t} \quad (10)$$

La matrice $\mathbf{C}(0)$ viene calcolata come matrice di varianza-covarianza dalla traiettoria di dinamica molecolare della proteina in acqua. A tale scopo, la traiettoria viene riferita rispetto a un sistema di riferimento solidale con la proteina e scelto in maniera tale che, ad ogni istante, abbia l'origine collocata sul centro di massa della proteina e sia orientato lungo agli assi principali d'inerzia.

Trasferimento di entropia

Il trasferimento di entropia $T_{i \rightarrow j}(t)$ è una misura di quanto viene ridotta l'incertezza sulla fluttuazione di j conoscendo quella di i ad un tempo precedente [16]. In termini probabilistici:

$$T_{i \rightarrow j}(t) = - \langle \ln p(\Delta R_j(0), \Delta R_j(t)) \rangle + \langle \ln p(\Delta R_i(0), \Delta R_j(0), \Delta R_j(t)) \rangle + \langle \ln p(\Delta R_j(0)) \rangle - \langle \ln p(\Delta R_i(0), \Delta R_j(0)) \rangle \quad (11)$$

Il primo termine è la media della densità di probabilità marginale congiunta di osservare una fluttuazione della posizione del j -esimo atomo al tempo iniziale, $\Delta R_j(0)$, e al tempo t , $\Delta R_j(t)$. La densità marginale è data da

$$p(\Delta R_j(0), \Delta R_j(t)) = \int_{-\infty}^{\infty} d\Delta R_1(0) \dots d\Delta R_{j-1}(0) d\Delta R_{j+1}(0) \dots d\Delta R_N(0) \times \int_{-\infty}^{\infty} d\Delta R_1(t) \dots d\Delta R_{j-1}(t) d\Delta R_{j+1}(t) \dots d\Delta R_N(t) p(\Delta \mathbf{R}(t), \Delta \mathbf{R}(0)) \quad (12)$$

L'integrale dell'equazione (12) è analitico per cui la densità probabilità diventa

$$p(\Delta R_j(0), \Delta R_j(t)) = \frac{e^{-\frac{1}{2} \begin{bmatrix} \Delta R_j(0) \\ \Delta R_j(t) \end{bmatrix} \begin{bmatrix} C_{j,j}(0) & C_{j,j}(t) \\ C_{j,j}(t) & C_{j,j}(0) \end{bmatrix}^{-1} \begin{bmatrix} \Delta R_j(0) \\ \Delta R_j(t) \end{bmatrix}}}{\sqrt{(2\pi)^2 |C_{j,j}^2(0) - C_{j,j}^2(t)|}} \quad (13)$$

il primo addendo dell'equazione (11) risulta

$$\begin{aligned} \langle \ln p(\Delta R_j(0), \Delta R_j(t)) \rangle &= \ln \left(\frac{1}{2\pi} \right) + \ln \left(\frac{1}{\sqrt{C_{j,j}^2(0) - C_{j,j}^2(t)}} \right) \\ &\quad - \frac{1}{2} \sum_{n=1}^N |V_{j,n}|^2 \left[\frac{\langle z_n^2(0) \rangle}{\lambda_n} + \frac{\langle z_n^2(t) \rangle}{\lambda_n} \right] \\ &= -\frac{1}{2} \ln [\langle \Delta R_j^2(0) \rangle^2 - \langle \Delta R_j(0) \Delta R_j(t) \rangle^2] - \ln(2\pi e) \quad (14) \end{aligned}$$

dove si è fatto uso della normalizzazione degli autovettori e del fatto che ad ogni tempo $\langle z_n^2 \rangle = \lambda_n$. Analogamente, gli altri addendi dell'equazione (11) possono essere espressi in funzione degli elementi della matrice di varianza-covarianza.

$$\langle \ln p(\Delta R_j(0)) \rangle = -\frac{1}{2} \ln [\langle \Delta R_j^2(0) \rangle] - \frac{1}{2} \ln (2\pi e) \quad (15)$$

$$\begin{aligned} \langle \ln p(\Delta R_i(0), \Delta R_j(0)) \rangle = \\ -\frac{1}{2} \ln [\langle \Delta R_i(0) \rangle \langle \Delta R_j(0) \rangle - \langle \Delta R_i(0) \Delta R_j(0) \rangle^2] - \ln (2\pi e) \end{aligned} \quad (16)$$

$$\begin{aligned} \langle \ln p(\Delta R_i(0), \Delta R_j(0), \Delta R_j(t)) \rangle = \\ -\frac{1}{2} \ln \{ \langle \Delta R_i(0)^2 \rangle [\langle \Delta R_j(0)^2 \rangle^2 - \langle \Delta R_j(0) \Delta R_j(t) \rangle^2] - \\ \langle \Delta R_i(0) \Delta R_j(t) \rangle [\langle \Delta R_i(0) \Delta R_j(0) \rangle \langle \Delta R_j(0)^2 \rangle - \langle \Delta R_i(0) \Delta R_j(t) \rangle \langle \Delta R_j(0) \Delta R_j(t) \rangle] + \\ \langle \Delta R_i(0) \Delta R_j(t) \rangle [\langle \Delta R_i(0) \Delta R_j(0) \rangle \langle \Delta R_j(0) \Delta R_j(t) \rangle - \langle \Delta R_j(0)^2 \rangle \langle \Delta R_i(0) \Delta R_j(t) \rangle] \\ - \frac{3}{2} \ln (2\pi e) \end{aligned} \quad (17)$$

3 Risultati

Al fine di comparare i risultati ottenuti per il $T_{i \rightarrow j}$ di monomero e dimero, si sono calcolate le seguenti quantità:

$$T_{\text{out},i} = \sum_{j=1}^N T_{i \rightarrow j} \quad (18)$$

$$NET = T_{i \rightarrow j} - T_{j \rightarrow i} \quad (19)$$

$T_{\text{out},i}$ rappresenta la somma del trasferimento di entropia dall'atomo i a tutti gli altri. Ha un valore negativo se l'atomo i si comporta complessivamente da accettore di entropia (*sink*), mentre ha un valore positivo se si comporta complessivamente da donatore (*source*). NET invece rappresenta lo scambio netto di entropia tra le coppie di atomi i e j , un valore negativo identifica j come *source* e i come *sink*, viceversa un valore positivo.

Per prima cosa si è rappresentato in grafico T_{out} a valori di tempo diversi per trovare il massimo trasferimento di entropia per il monomero e il dimero. Si è osservato che esso raggiunge un massimo a tempi minori per il monomero (0.1 ns circa) rispetto al dimero (> 5 ns). Il tempo di correlazione delle fluttuazioni delle posizioni degli atomi del monomero è quindi in media minore di quello del dimero. Questo suggerisce che la dinamica del dimero non sia la somma di dinamiche di monomeri indipendenti, ma vi siano interazioni tra i due. Nei grafici successivi si è utilizzato un tempo di 0.1 ns per il monomero e 5 ns per il dimero se non specificato diversamente.

Per verificare la convergenza si sono confrontati T_{out} calcolati usando tutta o solo metà traiettoria per il monomero e il dimero (Fig.1). Si è osservato che, a parte qualche atomo isolato, il comportamento complessivo di accettore o donatore rimane costante. La durata della simulazione (1 μs) sembra sufficiente a far convergere il calcolo dell'entropia trasferita. Anche in questo caso dal grafico del dimero è evidente come atomi omologhi su monomeri diversi si comportino in maniera differente e quindi vi sia un'asimmetria nel trasferimento di entropia, indice di interazioni inter-monomero.

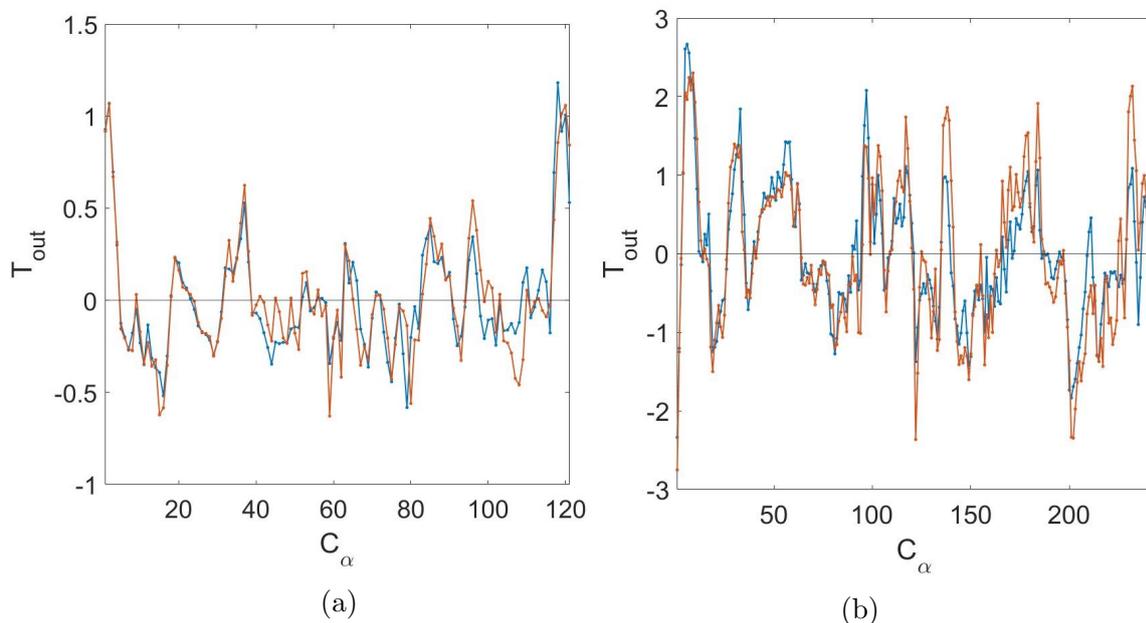


Figura 1: T_{out} del monomero (1a) e del dimero (1b) calcolato utilizzando tutta la traiettoria (in blu) o solo metà (in rosso).

Informazioni più precise sull'entità del trasferimento di entropia si possono ottenere dai grafici di *NET*. Osservando il grafico del monomero (Fig.2a) si nota innanzitutto un trasferimento significativo tra atomi (C_α) adiacenti nel *backbone*, ma soprattutto il forte carattere donatore delle due estremità C (T_2) e N (T_1) terminali. L'estremità T_1 dona soprattutto a β_1 , β_2 , α_1 e leggermente verso la zona L_4 , α_2 . L'estremità T_2 dona soprattutto a L_1 , L_2 , β_3 , β_4 , β_5 e leggermente a α_1 . Tra gli accettori più importanti si trovano invece β_1 e β_4 , che ricevono non solo dalle estremità, ma anche da altre strutture secondarie. Nel grafico del monomero a 5 ns (Fig.2b) è ancora evidente il carattere donatore di T_1 , T_2 ed emerge il carattere accettore di L_4 . Nel complesso l'intensità del trasferimento diminuisce e molti atomi non evidenziano più comportamenti specifici a riprova del fatto che il tempo è troppo lungo per osservare la correlazione tra di essi.

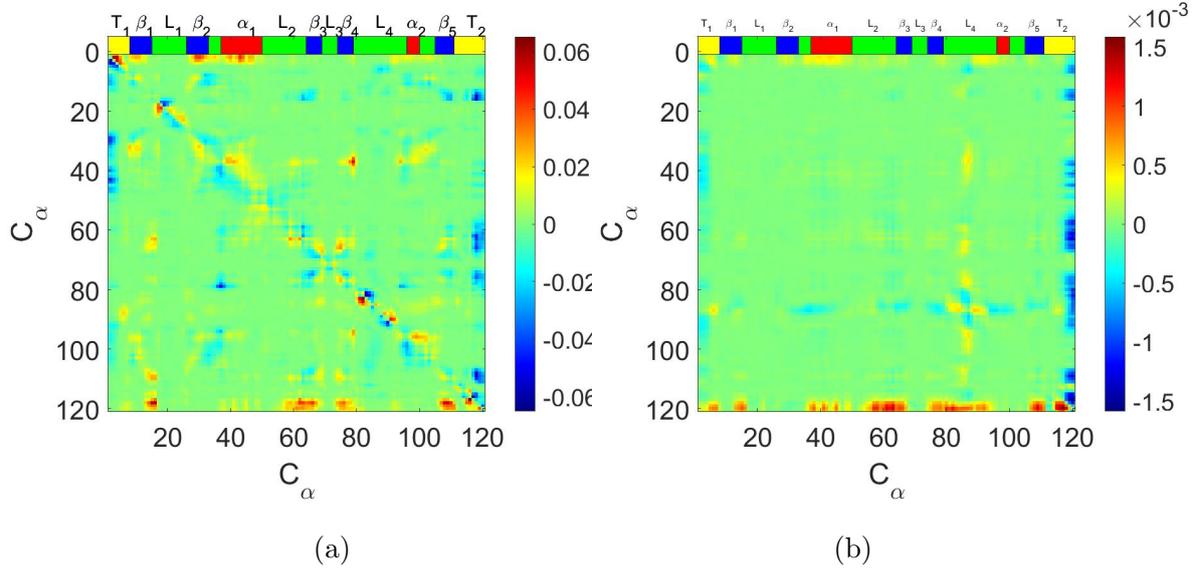


Figura 2: *NET* del monomero a 0.1 ns (2a) e 5 ns (2b).

Per quanto riguarda il dimer, il grafico *NET* con $t = 0.1$ ns (Fig.3a) non è particolarmente significativo perché il trasferimento di entropia nel dimer ha tempi più lunghi. Il trasferimento avviene in prevalenza lungo la diagonale, quindi tra atomi adiacenti nel *backbone*, ma si osserva anche un principio di trasferimento inter-monomero tra alcuni atomi del monomero A (A: 1-121) e la zona L_4 di B (B: 121-242).

Osservando invece il grafico a 5 ns (Fig.3b) emerge una fitta rete di scambi di entropia non solo intra-monomero ma anche inter-monomero. Il grafico è suddiviso in quattro quadranti, quello in alto a sinistra relativo al trasferimento intra-monomero AA, quello in basso a destra è invece l'analogo BB, quello in alto a destra è l'interazione inter-monomero AB e quello in basso a sinistra BA. Concentrandosi sui quadranti AA e BB si può osservare la variazione del comportamento delle estremità tra di loro e rispetto al monomero. Nel monomero T_1 e T_2 si comportano prevalentemente da donatori, mentre nel dimer T_1 di AA ha un comportamento misto, T_1 di BB si comporta da accettore, T_2 di AA dona solo verso L_4 e T_2 di BB ha un comportamento misto, anche se in prevalenza donatore. Ci sono zone che mantengono il comportamento in entrambi i monomeri, come per esempio α_1 , L_2 che rimangono donatori, zone che variano parzialmente il comportamento come L_1 che passa da accettore ad un comportamento misto e infine zone che cambiano completamente funzione come α_2 che passa da donatore ad accettore. Per quanto riguarda il trasferimento inter-monomero, è evidente che il monome-

ro A si comporta prevalentemente da donatore con i residui appartenenti a T_1 , β_2 , L_2 , α_2 , T_2 , mentre il monomero B si comporta prevalentemente da accettore con i residui di β_1 , β_2 , L_4 .

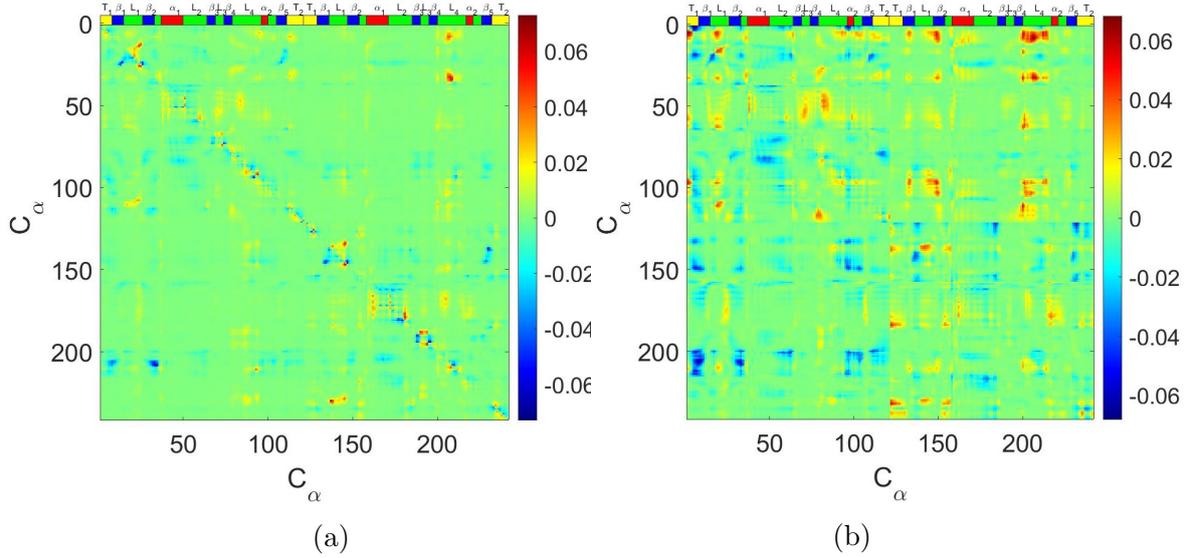


Figura 3: *NET* del dimero a 0.1 ns (3a) e 5 ns (3b).

I dati sul trasferimento di entropia sono stati confrontati con i risultati ottenuti dal gruppo del Prof. M. Buck, del dipartimento di fisiologia e biofisica della Case Western Reserve University, Cleveland, OH, USA [21], i quali sono stati calcolati utilizzando una traiettoria di 55 ns. Per quanto riguarda il monomero, si osserva che anche utilizzando una traiettoria più corta l'estremità C terminale (T_2) si comporta da donatore verso il resto della proteina e gli atomi individuati come accettori coincidono. Le due differenze importanti sono il comportamento di L_4 che utilizzando la traiettoria da 55 ns si comporta da donatore, mentre utilizzando quella da 1 μs si comporta da accettore, e l'estremità N terminale che si comporta da donatore utilizzando la traiettoria da 1 μs , ma non con quella da 55 ns. Il calcolo svolto con la traiettoria di 55 ns individua, quindi, l'estremità C terminale (T_2) e L_4 come residui guida del moto del dominio RBD, mentre il calcolo con la traiettoria da 1 μs individua come tali T_2 e T_1 . Il confronto sul dimero si concentra su un unico quadrante di tipo AA. Utilizzando la traiettoria da 55 ns l'estremità N terminale T_1 , L_1 , la fine di L_2 e l'inizio dell'estremità C terminale T_2 si comportano da donatori, mentre β_2 , α_1 , l'inizio di L_4 e β_5 si comportano da accettori. Utilizzando invece la traiettoria da 1 μs la fine di T_1 , β_1 , α_1 , L_2 , α_2 e T_2 (solo verso L_4) si comportano da donatori, mentre L_1 , L_4 e T_1 si com-

portano da accettori. Emerge quindi il comportamento diverso dell'estremità T_2 che non dona più a tutti gli altri residui ma solo ad L_4 e una generale variazione di ruoli accettore/donatore. Si conferma però L_4 come accettore nella forma dimerica, comportamento analogo si riscontra nel monomero con le traiettorie di $1 \mu s$, mentre con quelle di 55 ns si comporta da donatore.

L'asimmetria nel comportamento dei monomeri nel dimero è evidente e non deriva da un'asimmetria nel legame tra le due subunità, ma da variazioni conformazionali che avvengono nella scala dei millisecondi e che quindi non sono apprezzabili nel loro intero tramite simulazioni così corte. Si può osservare che la dinamica delle subunità che contribuisce alla variazione conformazionale completa è guidata da atomi differenti a seconda del monomero e a seconda della lunghezza della traiettoria considerata. Traiettorie di lunghezza diversa permettono di esplorare momenti diversi della variazione conformazionale in cui entrano in gioco atomi differenti e variano i loro comportamenti. I singoli atomi hanno poi tempi di correlazione diversi come evidenziato dai grafici a 0.1 e 5 ns del monomero e del dimero.

4 Conclusioni

In questo lavoro di tesi si è utilizzato il calcolo del trasferimento di entropia per descrivere la dinamica del sito RBD della Plexina B1 e del suo dimero allo scopo di osservare l'asimmetria transiente tra i due monomeri e l'effetto della lunghezza della traiettoria su di essa. In primo luogo, si sono calcolate le matrici di varianza-covarianza dalle simulazioni di dinamica molecolare [23] utilizzando CHARMM [37] e considerando la dinamica dei soli C_α . Il trasferimento di entropia, definito come la divergenza di Kullback-Leibler, è stato poi calcolato con MATLAB [38] facendo uso della matrice di varianza-covarianza nel calcolo delle densità di probabilità marginali congiunte. A questo punto per estrarre informazioni si è calcolato il trasferimento di entropia da un atomo verso tutti gli altri ($T_{\text{out},i}$), che rappresenta il comportamento globale, e la differenza tra il trasferimento dall'atomo i all'atomo j e quello da j a i che rappresenta il comportamento relativo di due atomi tra di loro (NET). $T_{\text{out},i}$ permette di osservare che la lunghezza della traiettoria è sufficiente a far convergere il calcolo del trasferimento di entropia. Si può inoltre osservare la presenza dell'asimmetria già dal grafico di $T_{\text{out},i}$ del dimero, in cui atomi analoghi su monomeri diversi hanno comportamenti differenti. NET rappresenta bene la fitta rete di scambi inter-monomero in cui il monomero A si comporta prevalentemente da donatore mentre B si comporta prevalentemente da accettore. Il calcolo del trasferimento di entropia è quindi ottimo per evidenziare dinamiche differenti tra i due monomeri che si traducono in un'asimmetria transiente. Lo studio della correlazione tra la dinamica degli atomi è importante per l'individuazione di percorsi allosterici che svolgono un ruolo fondamentale nei sistemi biologici.

Il confronto tra i risultati ottenuti in questo lavoro di tesi e quelli ottenuti dal gruppo del Prof. M. Buck [21] ha permesso di evidenziare come la lunghezza della traiettoria abbia un impatto non trascurabile sui risultati ottenuti. I tempi di correlazione nel monomero e nel dimero sono diversi, ma si aggirano entrambi nell'ordine dei nanosecondi. Le traiettorie di 55 ns non sono sufficienti però a far convergere il calcolo. Questo è probabilmente dovuto al fatto che pochi ns bastano a raggiungere il massimo di trasferimento di entropia ma non a descrivere il processo nel suo intero. Nello specifico è bene sottolineare l'inversione di comportamento del loop L_4 , fondamentale per la formazione del dimero, che utilizzando le traiettorie da 55 ns si comporta da donatore, mentre con quelle da 1 μs si comporta da accettore. Anche il calcolo dell'entropia conformazionale del processo di dimerizzazione presenta risultati opposti con le traiettorie da 1 μs e 55 ns per quasi tutte le zone del sito RBD [39].

Il trasferimento di entropia potrebbe essere un metodo per studiare l'ef-

fetto di mutazioni sull'attività biologica delle proteine. Le mutazioni sono facilmente implementabili nelle simulazioni di dinamica molecolare che forniscono i dati per il calcolo del trasferimento di entropia. Questo permette di osservare se i percorsi allosterici vengono conservati o meno e se questo influenza l'attività della proteina, anche nell'ottica di sviluppo di farmaci allosterici.

5 Bibliografia

- (1) Changeux, J. *Protein Sci.* **2011**, *20*, 1119–1124.
- (2) Ribeiro, A. A. S. T.; Ortiz, V. *Chem. Rev.* **2016**, *116*, 6488–6502.
- (3) Cooper, A.; Dryden, D. *Eur. Biophys. J.* **1984**, *11*, 103–109.
- (4) Hilser, V. J.; Wrabl, J. O.; Motlagh, H. N. *Annu. Rev. Biophys.* **2012**, *41*, 585–609.
- (5) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. *Nature* **2014**, *508*, 331–9.
- (6) Gunasekaran, K.; Ma, B.; Nussinov, R. *Proteins: Struct., Funct. Bioinf.* **2004**, *57*, 433–443.
- (7) *J. Mol. Biol.* **2018**, *430*, 2309–2320.
- (8) Guarnera, E.; Berezovsky, I. *Curr. Op. Struct. Biol.* **2018**, *56*, 18–27.
- (9) Cournia, Z.; Chatzigoulas, A. *Curr. Op. Struct. Biol.* **2020**, *62*, 197–204.
- (10) Lisi, G. P.; Loria, J. P. *Curr. Op. Struct. Bio.* **2017**, *47*, 123–130.
- (11) Li, S.; Zhang, J.; Lu, S.; Huang, W.; Geng, L.; Shen, Q.; Zhang, J. *PloS one* **2014**, *9*, e97668.
- (12) Nussinov, R.; Tsai, C. J. *Cell* **2013**, *153*, 293–305.
- (13) Cramer, C., *Essentials of Computational Chemistry: Theories and Models*; Wiley: 2005.
- (14) Duan, N.; Liu, Y.; Kong, R.; Yu, Z.; Lu, S.; Zhang, J. *Drug Discov.* **2022**, *27*.
- (15) Schreiber, T. *Phys. Rev. Lett.* **2000**, *85*, 461–464.
- (16) Hacısuleyman, A.; Erman, B. *J. Drug Des. Res.* **2017**, *4(5)*, 1054.
- (17) Hacısuleyman, A.; Erman, B. *Proteins* **2017**, *85*, 1056–1064.
- (18) Garcia Michel, L.; Keirns, C.; Ahlbrecht, B.; Barr, D. *J. Chem. Theory Comput.* **2020**, 3168–3177.
- (19) Hacısuleyman, A.; Erman, B. *PLoS Comput. Biol.* **2017**, *13(1)*, e1005319.
- (20) Negishi, M.; Oinuma, I.; Katoh, H. *PLoS Comput. Biol.* **2005**, *62(12)*, 1363–1371.
- (21) Zhang, L.; Centa, T.; Buck, M. *J. Phys. Chem. B.* **2014**, *118*, 7302–7311.

- (22) Zerbetto, M.; Anderson, R.; Bouguet-Bonnet, S.; Rech, M.; Zhang, L.; Meirovitch, E.; Polimeno, A.; Buck, M. *J. Phys. Chem. B.* **2013**, *117*, 174–184.
- (23) Zhang, L.; Buck, M. *J. Phys. Chem. B.* **2017**, *121*, 1485–1498.
- (24) Kim, T. H.; Mehrabi, P.; Ren, Z.; Sljoka, A.; Ing, C.; Bezginov, A.; Ye, L.; Pomès, R.; Prosser, R. S.; Pai, E. F. *Science* **2017**, *355*, eaag2355.
- (25) Na, I.; Catena, D.; Kong, M. J.; Ferreira, G. C.; Uversky, V. N. *Int. J. Mol. Sci.* **2018**, *19*(7), 1899.
- (26) Hertig, S.; Latorraca, N. R.; Dror, R. O. *PLoS Comput. Biol.* **2016**, *12*(6), e1004746.
- (27) Hollingsworth, S.; Dror, R. *Neuron* **2018**, *99*, 1129–1143.
- (28) Yang, Y. I.; Shao, Q.; Zhang, J.; Yang, L.; Gao, Y. Q. *J. Chem. Phys.* **2019**, *151*, 070902.
- (29) Isrlaelewitz, B.; Gao, M.; Schulten, K. *Curr. Op. Struct. Biol.* **2001**, *11*, 224–230.
- (30) Ferrara, P.; Apostolakis, J.; Caffisch, A. *J. Phys. Chem. B.* **2000**, *104*, 4511–4518.
- (31) Paci, E.; Karplus, M. *J. Mol. Biol.* **1999**, *288*, 441–459.
- (32) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.
- (33) Marrink, S. J.; Tieleman, D. P. *Chem. Soc. Rev.* **2013**, *42*(16), 6801–6822.
- (34) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé L, S. K. *J. Comput. Chem.* **2005**, *26*(16), 1781–1802.
- (35) Buck, M.; Bouguet-Bonnet, S.; Pastor, R.; MacKerell, A. *Biophys. J.* **2006**, *90*, L36–8.
- (36) Shaw, D. E. et al. *Commun. ACM* **2008**, *51*, 91–97.
- (37) Brooks, B. R. et al. *J. Comp. Chem.* **2009**, *30*, 1545–1615.
- (38) MATLAB version 9.10.0.1739362 (R2021a) Update 5; The Mathworks, Inc., Natick, Massachusetts, 2021.
- (39) Pshetitsky, Y.; Mendelman, N.; Li, Z.; Zerbetto, M.; Buck, M.; Meirovitch, E. *J. Phys. Chem. B.* **2022**.