



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Corso di Laurea Triennale in Fisica

Tesi di Laurea

Caratterizzazione cinetica di biosensori di DNA in presenza di effettori allosterici tramite esperimenti di singola molecola

Relatrice

Dr.ssa Annamaria Zaltron

Correlatore

Dr. Xavier Viader Godoy

Laureando

Riccardo Tancredi

Matricola

1226214

Anno Accademico 2021/2022

*A mio nonno Giuseppe
Ai miei genitori
Alla mia famiglia*

*“È veramente bello battersi con persuasione,
abbracciare la vita e vivere con passione.
Perdere con classe e vincere osando,
perché il mondo appartiene a chi osa!
La vita è troppo bella per essere
insignificante.”*

— *Charlie Chaplin*

Indice

Indice	i
Introduzione	iii
1 Pinzette ottiche	1
1.1 Trappola ottica	1
1.2 Manipolazione di biosensori	3
2 Metodi sperimentali	5
2.1 Optical-mini tweezers	5
2.2 Sistema biologico e modello WLC	8
3 Esperimenti di singola molecola	11
3.1 Misure di hopping	12
3.2 Tempo di residenza totale	16
3.3 Stime conclusive	18
3.4 Hidden Markov Model - HMM	19
3.5 Legame tra molecola e oligo-nucleotide	21
Conclusioni	23
Bibliografia	25

Introduzione

Negli ultimi 20 anni il settore delle pinzette ottiche è cresciuto enormemente, da quando Arthur Ashkin brevettò il primo strumento in grado di sfruttare il momento lineare e angolare della luce per esercitare forze sulla materia [1]. Le pinzette ottiche sono ad oggi uno degli strumenti più utilizzati per investigare biomolecole e processi biofisici che governano una vastissima gamma di polimeri. La peculiarità di questo strumento risiede nella possibilità di poter condurre esperimenti di singola-molecola, in cui le proprietà dei polimeri oggetto di studio vengono studiate individualmente; rappresenta pertanto un grande passo in avanti rispetto ad esperimenti di *ensemble* con un numero elevato di molecole, dell'ordine del numero di Avogadro, dove i comportamenti individuali delle singole molecole non si possono contraddistinguere e si misurano solo proprietà medie, d'insieme.

Attraverso le *Optical-Tweezers* si riescono a fissare le estremità di una biomolecola a due opportune superfici: il movimento relativo di queste superfici genera una forza ed un allungamento del sistema in esame, consentendo di realizzare forze e spostamenti rispettivamente dell'ordine del piconewton (pN) e del nanometro (nm).

Nel presente lavoro di tesi si è impiegato un apparato strumentale, che prende il nome di *Optical mini-Tweezers*, per effettuare misure di *hopping*: adoperando due fasci laser contro-propaganti nel vicino infrarosso ($\lambda = 805\text{ nm}$), si intrappola una perlina di polistirene con un diametro di circa $3\text{ }\mu m$ in un'apposita camera detta *camera microfluidica*. Questa è costituita da tre canali: nel canale superiore vengono introdotte in soluzione alcune biglie che sono intrappolate otticamente nel canale centrale; qui avviene anche la suzione di una seconda biglia, proveniente dal canale inferiore, del diametro di circa $2\text{ }\mu m$, attraverso una micropipetta. L'ancoraggio del sistema biologico in studio avviene tra le due biglie che possono essere mosse proprio grazie alle pinzette ottiche che applicano così uno stimolo meccanico esterno alla biomolecola.

Il sistema studiato in questo lavoro è un *hairpin* di DNA: l'acido desossiribonucleico assume generalmente una configurazione a doppia elica, ma è possibile ottenerlo a singolo filamento (*single-stranded DNA*, ssDNA) sotto forma di *hairpin*, con una struttura a stem-loop. In particolare, la molecola in esame presenta nel loop tre siti di riconoscimento per un *oligo-nucleotide*, anche detto *target* o *affettore allosterico*, che una volta legato all'*hairpin* ne modifica le proprietà cinetiche ed energetiche. Il legame con il *target* permette infatti di modificare la struttura tridimensionale della molecola di DNA, facendola passare dallo stato nativo chiuso (*folded*) a quello aperto (*unfolded*), e consente quindi di controllare il rilascio di una eventuale terza molecola (detta *cargo*) legata allo stem. Quanto appena descritto rappresenta il progetto più ampio in cui si inserisce il presente lavoro di tesi, che si è focalizzato sulla caratterizzazione cinetica ed energetica della sola molecola di DNA e su un primo studio sperimentale degli effetti indotti dal legame di quest'ultima con il *target*.

Nello specifico, da misure di forza e allungamento effettuate sulla biomolecola si ricavano parametri elastici tramite il modello teorico *Worm-Like Chain* (WLC): a seconda della forza

applicata, la molecola può trovarsi nella configurazione aperta, *unfolded*, oppure nativa (chiusa), *native (folded)*. Si definisce *forza di coesistenza* la forza alla quale si ha uguale probabilità di trovare il sistema in uno dei due possibili stati; lavorando in un intorno di questa forza si effettuano misure di *hopping*, in cui cioè il sistema performa un grande numero di passaggi di stato tra le due configurazioni di equilibrio. Tramite il modello di Bell-Evans è poi possibile risalire ad altri parametri fisici d'interesse, come il *tempo di residenza* dello stato aperto o chiuso, o delle energie di formazione in gioco.

La tesi si articola complessivamente in tre capitoli:

- Nel primo capitolo è spiegato il principio fisico dietro il funzionamento delle pinzette ottiche;
- Nel secondo capitolo si descrive nel dettaglio la strumentazione utilizzata presso il *Dipartimento di Fisica e Astronomia* dell'*Università di Padova* e si discute nello specifico il biosensore impiegato, descrivendo anche il modello *WLC*;
- Nel terzo capitolo, infine, si presentano e si commentano le tipiche curve di *hopping* e si esegue un'analisi dati per stimare la *forza di coesistenza* (f_c), la *lunghezza del ssDNA* (x_{NU}) e l'*energia di formazione* (ΔG_0), a forza zero, dell'*hairpin* di DNA. Si esegue infine un'interpolazione dei dati sperimentali ottenuti attraverso un *Hidden Markov Model* e si mostra inoltre un esempio di legame di un *oligo-nucleotide* ad uno dei siti di attacco della catena singola di DNA.

Capitolo 1

Pinzette ottiche

Le pinzette ottiche sono uno strumento che sfrutta la luce per intrappolare piccoli oggetti di materia in una buca di potenziale, per poter manipolare biomolecole e studiarne le proprietà biofisiche. I loro impieghi sono molteplici: dall'applicazione e la misurazione di forze controllate, dell'ordine del pN , su sistemi microscopici, a spostamenti dell'ordine del nm , o ancora per investigare comportamenti di non equilibrio termodinamico.

La biologia molecolare ha mostrato come una quantità smisurata di eventi della vita cellulare è regolata da forze che governano meccanismi di trasporto cellulare o di deformazione indotta su certi enzimi durante processi di catalisi. La possibilità di applicare forze su singole molecole e misurarle permette pertanto di esplorare tali trasformazioni mecano-chimiche.

1.1 Trappola ottica

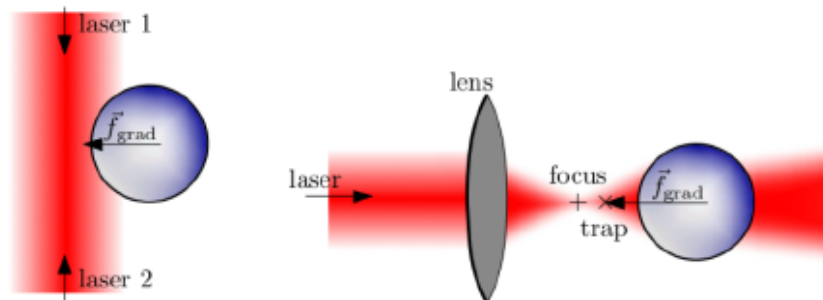


Figura 1.1: Optical Tweezers

Una trappola ottica comporta il bilanciamento di due tipi di forze: una di *scattering* che spinge gli oggetti lungo la direzione di propagazione della luce, e una di *gradiente* che invece attira gli oggetti lungo il gradiente spaziale dell'intensità luminosa [1]. Quando la forza di *gradiente* eccede su quelle di *scattering*, un oggetto è attratto nel punto di massima intensità luminosa ed è qui intrappolato nelle tre dimensioni spaziali. All'atto pratico si adoperano uno o più laser focalizzati da obiettivi da microscopio, per creare un alto gradiente di intensità luminosa, necessario per ottenere una trappola stabile. Nelle vicinanze del punto di fuoco degli obiettivi, la trappola si comporta come una molla lineare, generando forze sull'oggetto proporzionali al suo spostamento dal centro della trappola stessa.

Ciò che si intrappola sono generalmente delle biglie in polistirene di diametro di $2 - 3 \mu m$: va-

riando la posizione relativa delle due, una volta attaccato a queste una biomolecola, si possono eseguire misure di forza e di spostamento.

La pressione di radiazione è definita [2] come la forza media per unità di area esercitata dal campo elettromagnetico su una superficie. Dalla conservazione della quantità di moto, il momento trasportato dalla luce è modificato quando il fascio luminoso è diffuso dalla biglia intrappolata; come conseguenza della conservazione della quantità di moto, un momento uguale ed opposto è esercitato sulla particella intrappolata, comportando l'azione di una forza su di essa.

A seconda del rapporto tra la lunghezza d'onda della luce λ e il diametro d della particella intrappolata, si seguono approcci teorici differenti per comprendere meglio il principio fisico dietro le *optical tweezers*. Per $d \gg \lambda$, è possibile adoperare l'ottica geometrica per descrivere il fascio luminoso, rappresentandolo come un insieme di raggi luminosi ciascuno caratterizzato da una certa intensità e direzione. La quantità di moto del raggio incidente \vec{p}_i è scomposta nella componenti \vec{p}_r riflessa e \vec{p}_t trasmessa: la biglia investita dal fascio acquisisce allora un momento lineare dato dalla conservazione della quantità di moto e pari a:

$$\vec{p}_b = \vec{p}_i - (\vec{p}_r + \vec{p}_t)$$

In questo regime, detto di *Mie*, il contributo dato da \vec{p}_r viene trascurato, poiché molto più debole rispetto a \vec{p}_t . La forza risultante applicata alla biglia è allora data dalla variazione nel tempo di \vec{p}_b :

$$\vec{f}_b = \frac{d\vec{p}_b}{dt}$$

La forza risultante viene scomposta nelle due componenti: quella di *scattering* \vec{f}_{scatt} e quella di *gradiente* \vec{f}_{grad} . Poiché la forza di gradiente è proporzionale al gradiente d'intensità del fascio [3], nel caso di un fascio con profilo gaussiano fortemente focalizzato, è possibile intrappolare la particella nella posizione in cui la luce ha intensità maggiore, ossia nel fuoco della lente, poiché la forza di *gradiente* sarà dominante su quella di *scattering*.

Qualora invece $d \ll \lambda$, bisognerebbe fare riferimento al regime di *Raylight*, in cui si tratta la biglia come un piccolo dipolo sferico immerso in campo elettromagnetico uniforme. Si riescono così ad ottenere delle espressioni separate per le due componenti della forza: la forza di *scattering* di un dipolo sferico con indice di rifrazione n_m è data da

$$\vec{f}_{scatt} = n_m \frac{\sigma I_0}{c} \hat{k}$$

con I_0 l'intensità del fascio, c la velocità della luce, \hat{k} il versore che indica la direzione di propagazione del fascio e σ la sezione d'urto di scattering di una particella di raggio r ; la forza di gradiente si ricava invece a partire dalla forza di Lorentz che agisce sul dipolo ed è pari a:

$$\vec{f}_{grad} = \frac{2\pi\alpha}{cn_m^2} \vec{\nabla} I_0$$

con α la polarizzabilità della particella.

Affinché la biglia possa essere intrappolata, la forza di gradiente deve essere molto più forte di quella di scattering, condizione che, come nel regime di *Mie*, risulta soddisfatta per fasci di luce fortemente focalizzati.

1.2 Manipolazione di biosensori

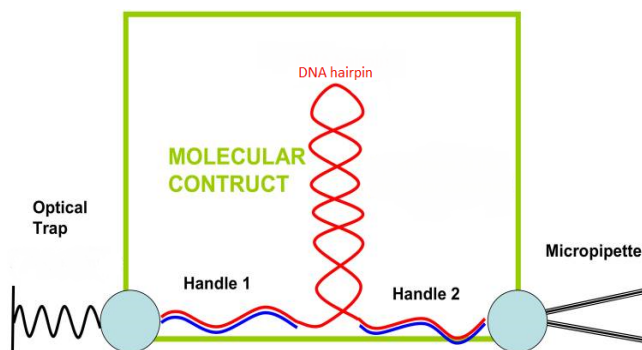


Figura 1.2: Disposizione di un *hairpin* di DNA in una trappola ottica

Solitamente per evitare fenomeni di danneggiamento ottico, le pinzette ottiche vengono usate per manipolare direttamente oggetti dielettrici macroscopici come perle di polistirene, ai quali si legano le biomolecole d'interesse. Come mostrato in Figura [1.2], queste perle sono adoperate come maniglie per il sistema biologico, permettendo la sua manipolazione.

La biomolecola esaminata nel presente lavoro è un singolo filamento di DNA (*hairpin*), alle cui estremità sono attaccate due sequenze identiche di DNA a doppio filamento:

queste fungono da maniglie (*handles*), poiché sono a loro volta attaccate a due biglie di polistirene, tramite legami biochimici di Streptavidina-Biotina sulla biglia più piccola con un diametro di circa $2\ \mu\text{m}$ (SB) e Digoxigenina-Anti-digoxigenina sulla biglia più grande, con un diametro di $3\ \mu\text{m}$ approssimativamente (AD).

Il legame non covalente della Biotina alla Streptavidina è uno dei legami più forti conosciuti in natura, motivo per il quale questo complesso è ampiamente usato in biologia molecolare e in biotecnologie per legare proteine o, come in questo caso, filamenti di DNA. La Digossigenina è invece uno steroide che viene riconosciuto da anticorpi Anti-digossigenina, dando pertanto origine ad un legame difficile da rompere alla temperatura alla quale si lavora, che è la temperatura ambiente del laboratorio, tra i 20°C e i 25°C .

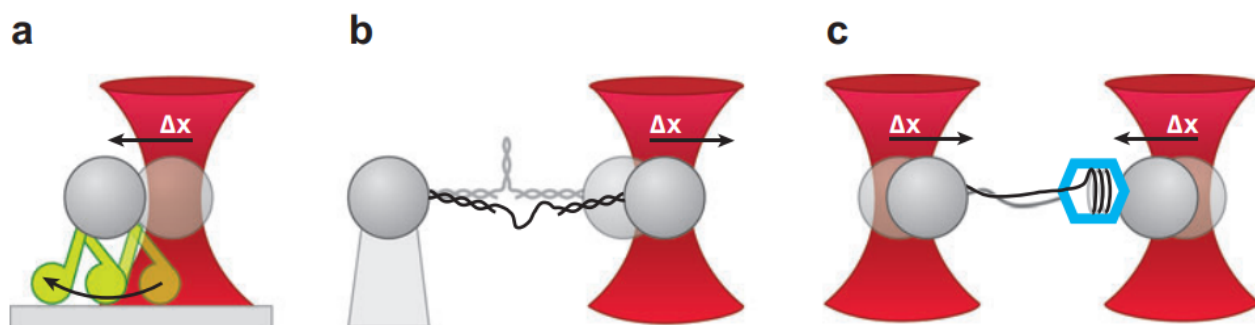


Figura 1.3: Geometrie di intrappolamento ottico possibili per pinzette ottiche

Una volta intrappolata otticamente una di queste perline, l'altra estremità del sistema viene attaccata tramite legami biochimici ad una seconda superficie, che può essere un vetrino (Figura [1.3-a]) o una seconda biglia tenuta fissa per suzione sulla punta di una micropipetta (Figura [1.3-b]), o intrappolata in un'ulteriore trappola ottica (Figura [1.3-c]). In questo modo le proprietà cinetiche del sistema possono essere studiate muovendo la trappola ottica relativamente al secondo punto di attacco del sistema biologico.

Nel presente lavoro si è adoperata una geometria come quella in Figura [1.3-b].

Capitolo 2

Metodi sperimentali

2.1 Optical-mini tweezers

L'apparato sperimentale adoperato è noto con il nome di *Optical mini-Tweezers* ed è stato sviluppato dal gruppo di S. Smith e C. Bustamante [4] della University of California nei primi anni 2000.

Il nome deriva dal fatto che la strumentazione in uso è particolarmente compatta, con maggiore stabilità e migliore risoluzione se confrontata con le convenzionali *Optical Tweezers* [5].

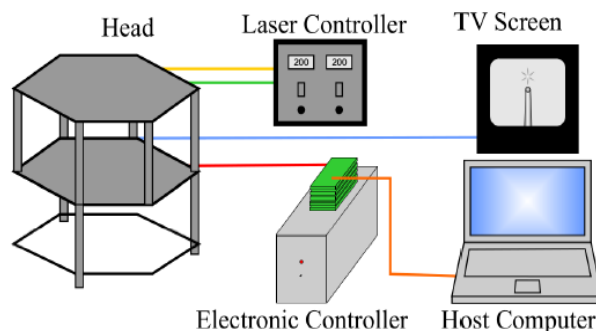


Figura 2.1: Optical mini-Tweezers

2.1.1 Laser e percorso ottico

La trappola ottica è generata a partire da due fasci laser contro-propaganti di lunghezza d'onda pari a $\lambda = 805 \text{ nm}$ e potenza variabile fino a 200 mW .

Il fascio luminoso prodotto ha un profilo gaussiano ed è polarizzato linearmente: i due fasci generati dai due laser a diodo (Lumics SN0834770) entrano in una fibra ottica monomodale che fornisce la componente trasversale del campo elettromagnetico. L'intensità e la temperatura dei laser sono monitorati da appositi sistemi di controllo.

Come mostrato in [6] è possibile realizzare una trappola a partire da un singolo laser, ma è necessario disporre di obiettivi ad alta apertura numerica ($\text{NA} = 1.2\text{-}1.4$), così da far prevalere la forza di gradiente su quella di scattering. In queste condizioni è però difficile raccogliere tutta la luce deflessa dal corpo intrappolato, fondamentale per il calcolo della forza applicata a partire dalla variazione del momento dei raggi luminosi.

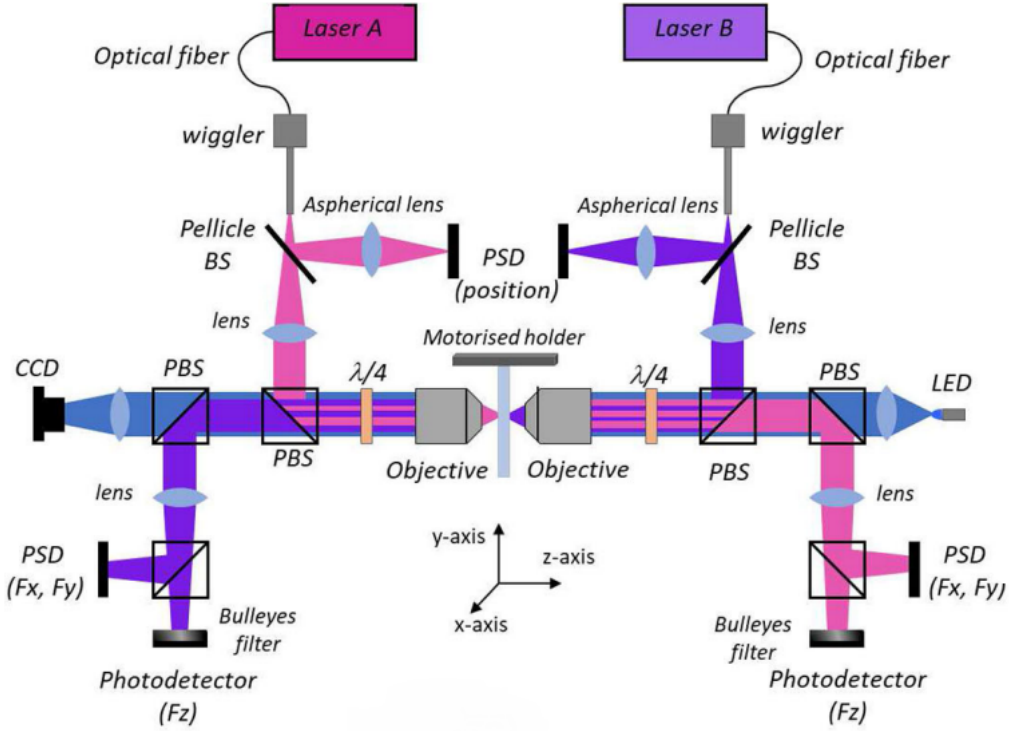


Figura 2.2: Schema del percorso ottico

Nel set-up impiegato in laboratorio, come quello in Figura [2.2], la simmetria del sistema permette la completa compensazione delle forze di scattering lungo la direzione di propagazione dei due fasci laser. Poiché allora la forza di gradiente è l'unica forza risultante sulla biglia, non è più necessario impiegare obiettivi ad elevata apertura numerica, ed è possibile raccogliere più del 95% della luce deflessa, che viene mandata ad appositi rivelatori (*detectors*) per monitorare nel tempo le forze applicate al sistema in esame.

I due fasci laser seguono un identico cammino ottico [7]; ogni laser è accoppiato ad una fibra ottica che è ancorata ad un *wiggler* che la muove delicatamente tramite piccoli motori piezoelettrici, variando così la posizione della trappola ottica. Quasi il 4% della luce è diretta in un PSD (*positive-sensitive-detector*) da un BS, un *beam-splitter* a pellicola. I due PSD monitorano lo spostamento della trappola in due dimensioni spaziali, ossia nel piano di lavoro.

La luce rimanente passa attraverso una lente, che collima il fascio; da qui, un *polarizing-beam-splitter* (PBS), una lamina a quarto d'onda ($\lambda/4$) e un obiettivo, focalizzano la luce dentro una camera microfluidica, polarizzandola circolarmente. In questo modo la forza agente sulla biglia non dipende dalla polarizzazione della luce incidente.

La luce che va incontro a scattering è raccolta da un secondo obiettivo, identico al precedente, e diretta in un PSD e un fotodiodo che convertono ciascuno spostamento del fascio in segnali di forza F_x , F_y , F_z . L'esperimento è poi di fatto condotto all'interno di una camera microfluidica che è incastrata tra i due obiettivi, ed è sostenuta da un supporto motorizzato equipaggiato con dei motori piezoelettrici che permettono la traslazione dell'intera area dell'esperimento. La visualizzazione dell'area dell'esperimento sullo schermo è infine possibile grazie ad un LED e una camera CCD. La variazione di polarizzazione del fascio garantisce la non interazione dei due fasci nei tratti in cui si trovano su percorsi sovrapposti.

2.1.2 Sensori e imaging

I PSD rivelano la posizione x e y della trappola e i valori di forza F_x e F_y applicata alla molecola. Essi forniscono un segnale in corrente proporzionale allo spostamento del centroide del fascio che colpisce l'area attiva.

Un PSD è un diodo PIN: quando un fascio di fotoni incide su di esso produce coppie elettrone-lacuna che vengono raccolte da una serie di elettrodi come segnale di corrente. Quest'ultimo è proporzionale alla distanza tra il punto inciso dai fotoni e l'elettrodo, consentendo di rivelare la posizione del fascio sul detector.

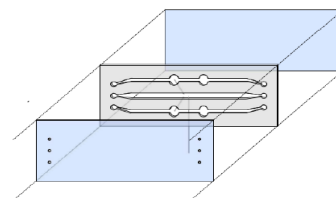
La posizione lungo z della biglia nella trappola ottica viene invece ricavata antepponendo un filtro *bulleye* a un fotodiodo: il detector rileva una potenza che dipende dalla dimensione dell'area colpita, funzione della posizione z della biglia. La forza F_z si stima a partire dal diametro del fascio che varia rispetto ad una situazione di assenza di forze.

Il sistema di imaging è invece costituito da una luce LED blu ($\lambda = 470\text{nm}$) e una camera CCD. L'immagine è proiettata tramite una lente sul CCD della camera, che rileva luce nel visibile e nel vicino infrarosso, permettendo di vedere all'occorrenza sia i laser che l'esperimento.

I sei sensori di cui è dotato lo strumento permettono quindi di monitorare in tempo reale la distanza tra le due biglie e dunque di determinare sia l'allungamento del costrutto molecolare che la forza applicata al sistema.

2.1.3 Camera microfluidica e Pipet-Puller

La camera microfluidica è costituita da due vetrini delle dimensioni di $24\text{mm} \times 60\text{mm} \times 150\mu\text{m}$ ognuno, tra i quali sono posti due strati di parafilm opportunamente tagliato, così da realizzare tre canali paralleli. Con una sabbiatrice si realizzano su uno dei due vetrini dei fori, del diametro di $500\mu\text{m}$, in corrispondenza di quelli presenti sui parafilm. Con una colla a UV, si attaccano i due parafilm sui vetrini: tra i due strati si inseriscono due tubi dispensatori del diametro di $100\mu\text{m}$ per mettere in comunicazione i canali inferiore e superiore con il canale centrale.



Al centro del canale di mezzo si colloca invece la micropipetta, realizzata attraverso uno strumento chiamato *Pipet-Puller* a partire da tubicini in vetro del diametro di $80\mu\text{m}$. I due vetrini, con i tubi dispensatori e la micropipetta, sono poi uniti attraverso la fusione del parafilm su una piastra a circa 90°C . I fori presenti alle estremità di ogni canale servono per inserire e far uscire del buffer (soluzione in cui si immerge il materiale biologico), i biosensori di DNA e le biglie. Le biglie AD, per esempio, vengono immerse nella camera attraverso il canale superiore: qui, una di esse, viene intrappolata otticamente [6].

Una volta realizzata, la camera microfluidica è fissata su un supporto tramite dei blocchi in plexiglas, ancorati con delle viti, come in Figura [2.3]. Tale supporto presenta dei fori in cui vengono fissati i connettori in plastica in cui sono inseriti i tubi di entrata e di uscita. Il centro dei connettori è posto in corrispondenza dei fori realizzati sul vetrino, cosicché il materiale (buffer, biglie e biosensore) inserito nei tubicini, attraverso apposite siringhe, possa fluire senza ostacoli nella camera e fuoriuscire nel contenitore di scarti.

La micropipetta è invece inserita in un ulteriore tubicino, collegato ad una siringa contenente aria con cui si effettua la suzione della biglia SB, precedentemente immersa nella camera mi-

crofluidica dal canale inferiore

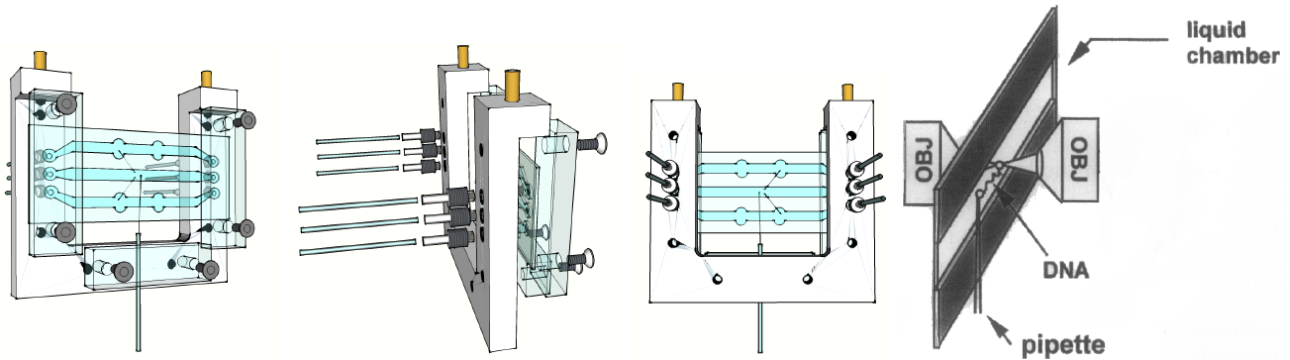
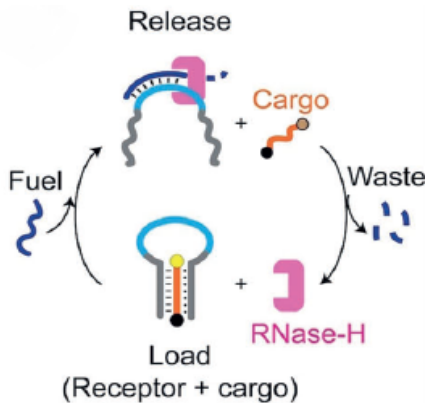


Figura 2.3: Montaggio della cameretta e posizionamento della camera tra i due obiettivi

Il *Piper-Puller*, sopra menzionato, permette la realizzazione di pipette con punta micrometrica, sfruttando un meccanismo molto ingegnoso: si infila il tubicino di diametro di $80\ \mu\text{m}$ tra due blocchi, uno superiore fisso ed uno inferiore appeso, ed è fatto passare per un filamento di tungsteno, arrotolato in due spire. Applicando una differenza di potenziale al filamento ($\approx 1.5\ \text{V}$), per effetto Joule, la parte di vetro vicino al tungsteno si deforma e si allunga; per gravità, la parte inferiore cede ed è trascinata verso il basso dal blocco inferiore. Si forma così una micropipetta, in corrispondenza del doppio avvolgimento di tungsteno, con un diametro interno medio di circa $1\ \mu\text{m}$.

2.2 Sistema biologico e modello WLC



Lo scopo complessivo del progetto [8] in cui si sviluppa la presente tesi è quello di adoperare un *hairpin* di DNA come biosensore, cioè legare un filamento di DNA ad un *target* capace di modificarne la struttura tridimensionale e quindi di controllare l'eventuale rilascio di una molecola cargo. In questo lavoro ci si è focalizzati sullo studio delle proprietà cinetiche ed energetiche del solo *hairpin* tramite misure di hopping e si è fatto un primo studio sull'interazione del biosensore con il *target*.

2.2.1 Hairpin di DNA

Il DNA è solitamente costituito da due catene di nucleotidi legate tra loro da legami a idrogeno. Le due catene che formano l'acido desossiribonucleico sono complementari e dunque, nota una delle due, è possibile risalire immediatamente all'altra.

Invece, il DNA impiegato nel presente lavoro è a singolo filamento (*single-stranded DNA*, ssDNA) e forma un *hairpin* con tre possibili strutture, riportate in Figura [2.4]. Tutte le strutture presentano un doppio loop: sperimentalmente i due loop sono indistinguibili, poiché separati da una sola coppia di basi azotate con energia di legame $\delta G = -2.58\ \text{kcal/mol}$, molto inferiore

a quella delle cinque coppie di basi del valore di $\delta G = -10.25 \text{ kcal/mol}$.

Le strutture fanno riferimento alle condizioni sperimentali in cui l'*hairpin* è analizzato, cioè a temperatura $T \approx 25^\circ\text{C}$, da cui $1/\beta = k_B T = 0.593 \text{ kcal/mol} = 4.11 \text{ pN} \cdot \text{nm}$ e con una concentrazione di NaCl nel buffer pari a 1 M .

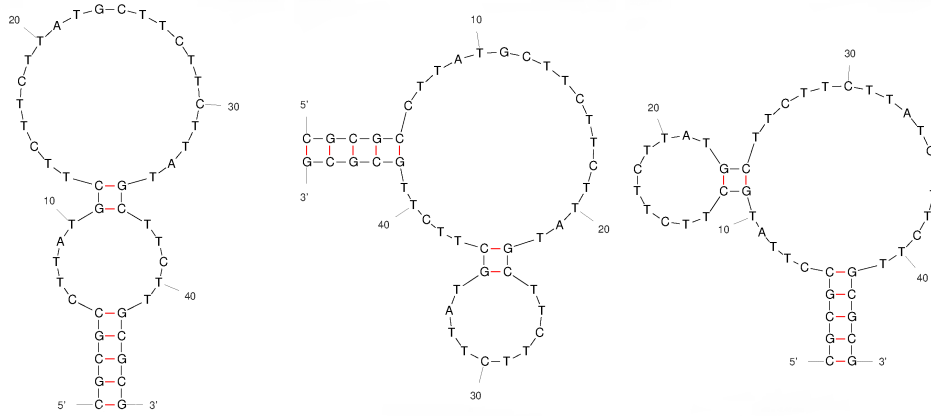


Figura 2.4: Strutture possibili assunte dalla sequenza di DNA, ricavate tramite la piattaforma di simulazione online Mfold

2.2.2 Modello WLC

Il DNA può essere anche descritto come una catena di polimeri costituita da unità fondamentali dette monomeri: le interazioni e i gradi di libertà di questi ultimi determinano la struttura e le possibilità di movimento della catena.

Il *Worm-Like Chain* (WLC) non è altro che un modello a catena lineare utilizzato per schematizzare l'*hairpin* di DNA impiegato. Questo modello permette di studiare le proprietà elastiche delle biomolecole, pensando al DNA come ad un polimero sul quale è possibile effettuare tre tipologie diverse di deformazione [9]. A partire da un piccolo segmento ds posizionato ad una distanza s da un'estremità della catena si possono introdurre le deformazioni di:

- (a) *bending*: misura come varia il vettore unitario \hat{t} tangente alla curva nel punto s , durante la deformazione;
- (b) *stretching*: descrive la variazione di lunghezza della catena;
- (c) *twisting*: delinea la rotazione attorno all'asse della catena.

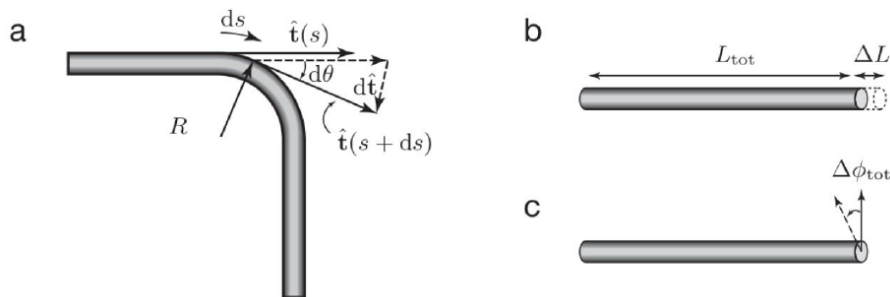


Figura 2.5: Possibili deformazioni di una catena: (a) *bending*, (b) *stretching*, (c) *twisting*

Per forze dell'ordine dei $10 - 15 \text{ pN}$, il termine che domina è quello di *bending*, dovuto al solo piegamento della catena polipeptidica. La relazione che lega la forza f applicata al sistema alla sua estensione z è data da:

$$f(z) = \frac{k_B T}{P} \left[\frac{1}{4(1 - z/L)^2} - \frac{1}{4} + \frac{z}{L} + \sum_{i=2}^{i \leq 7} \alpha_i \left(\frac{z}{L} \right)^i \right] \quad (2.1)$$

Nell'equazione soprastante, P rappresenta la lunghezza di persistenza e L la lunghezza a riposo della molecola, definita come $L = N \cdot d_{aa}$, cioè il prodotto tra il numero di nucleotidi che compongono la molecola e la distanza d_{aa} tra nucleotidi consecutivi. I coefficienti α_i sono termini correttivi che possono essere posti nulli oppure possono essere calcolati come in [10].

Nel corso della seguente trattazione sarà imposto $z = x_{ssDNA}$ e la forza f in gioco sarà quella di coesistenza, f_c . Questa in realtà verrà ricavata in un altro modo (equazioni [3.5] e [3.14]) e si invertirà l'equazione [2.1] per stimare f_c a partire da x_{ssDNA} , per calcolare l'energia libera di formazione, ΔG_0 . Per il DNA a singolo filamento [11] è noto che $P = (1.35 \pm 0.05) \text{ nm}$ e che $d_{aa} = (0.58 \pm 0.02) \text{ nm/base}$.

Capitolo 3

Esperimenti di singola molecola

Attraverso le pinzette ottiche ora descritte è possibile condurre esperimenti di singola-molecola. Nel set-up sperimentale qui adoperato, ricordiamo che è soltanto una la biglia che si muove con la trappola ottica, mentre l'altra è mantenuta ben ferma alla micropipetta.

Aumentando la forza, quindi l'allungamento dell'*hairpin*, si promuove la rottura dei legami idrogeno tra le basi azotate, fino alla completa apertura del DNA, che raggiunge una configurazione di stato aperto detta *unfolded* (U). Quando invece le biglie vengono progressivamente avvicinate, ad una certa forza di transizione, il sistema ritorna nel suo stato nativo ripiegato *folded* (F), qui indicato come *native* (N). Le transizioni tra i due stati N ed U si rivelano come un improvviso salto nella forza registrata dai sensori, che indica il cambiamento di configurazione della molecola. Questo è infatti ciò che si osserva in un grafico $\lambda - f$ di allungamento-forza in cui si vede chiaramente la presenza di un'irregolarità dall'andamento lineare previsto, in corrispondenza di due valori f_N ed f_U , corrispondenti alla stessa distanza tra la biglia nella micropipetta e la trappola ottica.

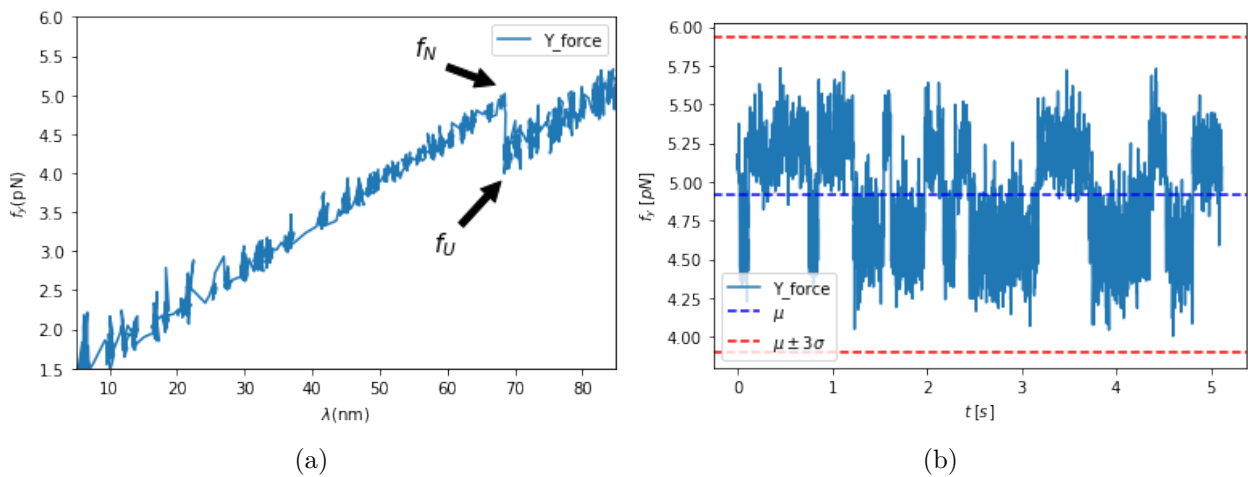


Figura 3.1: (a) Esempio di *pulling* (unfolding). (b) Esempio di *hopping*

Quelle evidenziate nel grafico in Figura [3.1-a] sono le due forze di transizione: f_N rappresenta la forza oltre la quale il sistema dalla configurazione *native* passa a quella *unfolded* e f_U il contrario, ed è uno dei tipi di esperimenti di singola molecola che si può condurre, detto di *pulling*, in cui si monitora l'allungamento della biomolecola al variare della forza, allontanando tra loro le biglie.

Nel corso del presente lavoro si è però condotto un altro tipo di esperimento, detto di *hopping*, in cui cioè la micropipetta e la distanza relativa con la trappola ottica, λ , è mantenuta fissa, e l'*hairpin* esegue transizioni spontanee, tra gli stati *native* e *unfolded* [12]. Dalle informazioni ottenute in precedenti esperimenti di *pulling*, si ha una stima di quella che è la cosiddetta *forza di coesistenza* f_c , definita come la forza alla quale la molecola ha la stessa probabilità di trovarsi in uno dei due stati possibili: questo è utile poiché, lavorando in un intorno di f_c , l'*hairpin* di DNA performa nel tempo un numero notevole di transizioni tra le configurazioni.

Nella Figura [3.1-b] sovrastante si osserva come in un tempo di poco più di 5 secondi, il sistema passa da uno stato all'altro più di una decina di volte. La forza nello stato *unfolded* è più bassa di quella nello stato *native*: questo è dovuto al fatto che la rigidità del sistema diminuisce quando la molecola è aperta e questo provoca uno spostamento della biglia verso il centro della trappola ottica, che corrisponde ad una forza inferiore applicata all'intero sistema molecolare. Poiché le forze con cui si opera, seppur dell'ordine di grandezza del pN , cioè mille miliardi più piccole di quelle che si sperimentano su scala antropica nella vita di tutti i giorni, sono fortunatamente più grandi del rumore termico dovuto all'ambiente, è possibile osservare e registrare le variazioni nei valori di forza che agiscono sul sistema, collezionando pertanto informazioni sulla configurazione assunta dalla biomolecola.

Si può dunque, anche attraverso esperimenti di *hopping*, ottenere una stima della forza di coesistenza, nonché dei rate cinetici di *unfolding*, k_U , e di *folding*, k_F , e, dai loro reciproci, una stima del *tempo di residenza* e di *coesistenza*, t_c , cioè della vita media di coesistenza dei due stati. Quest'ultimo si stima a partire da considerazioni sulla probabilità di osservare il sistema in uno stato piuttosto che in un altro.

Nel corso della presente analisi, si è stimata anche la variabile termodinamica ΔG_0 che rappresenta l'energia libera di formazione, a forza nulla, tra le configurazioni N ed U .

3.1 Misure di hopping

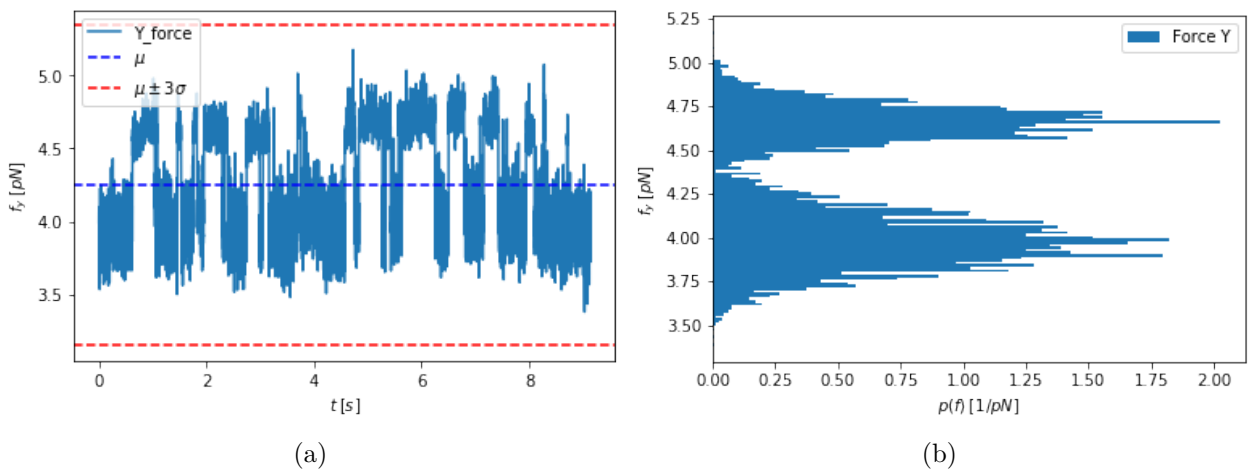


Figura 3.2: Misura di (a) *hopping* e (b) istogramma

Una volta catturato un *hairpin* di DNA tra le due biglie, si sono eseguite misure di hopping, con la metodologia descritta in precedenza. Sono state campionate e analizzate 13 molecole, adoperando due procedure diverse: *passive mode* (*pm*) e *constant position mode* (*cp*) [13]. Nella

prima modalità la distanza tra il centro della trappola e la punta della micropipetta è mantenuta fissa, con conseguente salto nella forza e nell'estensione della molecola al passaggio tra gli stati N e U . Nel *cp* mode invece la forza è mantenuta costante, attraverso un meccanismo elettronico detto di *force-feedback*, in cui dei motori piezoelettrici riposizionano la trappola ottica facendo in modo che rimanga costante la posizione della biglia rispetto al centro della trappola stessa. Il cambiamento di stato può essere pertanto identificato osservando l'improvvisa variazione nel parametro λ e quindi nell'estensione della molecola.

Dove non indicato si farà riferimento alla modalità *pm*; in alcuni casi la stessa molecola è stata studiata in entrambi i modi, così da poter poi anche confrontare i risultati.

Scelta la modalità, si sono eseguiti, per la stessa molecola, misure di *hopping* a diversi valori di forza. Si ottengono allora diversi grafici come quello in Figura [3.2-a].

Riportando in un istogramma il numero di punti sperimentali in funzione della corrispondente forza, si ottiene un grafico come quello in Figura [3.2-b], dal quale si possono ottenere importanti informazioni sulle proprietà cinetiche del sistema. I dati sono stati analizzati sotto l'ipotesi di una distribuzione gaussiana bidimensionale della probabilità di trovare la molecola nello stato nativo o in quello denaturato (*unfolded*) al variare della forza applicata [14]:

$$P(f) = \frac{w_U}{\sqrt{2\pi\sigma_U^2}} \exp\left(-\frac{1}{2}\left(\frac{f-f_U}{\sigma_U}\right)^2\right) + \frac{w_N}{\sqrt{2\pi\sigma_N^2}} \exp\left(-\frac{1}{2}\left(\frac{f-f_N}{\sigma_N}\right)^2\right) \quad (3.1)$$

dove f_U ed f_N rappresentano le forze medie degli stati *unfolded* e *folded*, σ_U e σ_N le rispettive deviazioni standard e, infine, w_U e w_N le probabilità che la molecola si trovi in uno dei due stati. Poiché l'equazione [3.1] è una probabilità, e dunque è normalizzata a 1, deve allora valere che $w_U + w_N = 1$; si ricorre allora ad un fit a 5 incognite, in cui il programma stima, al variare della forza, i parametri f_U , f_N , σ_N , σ_U e w_N (o w_U), insieme alle loro incertezze, che meglio minimizzano il χ^2 .

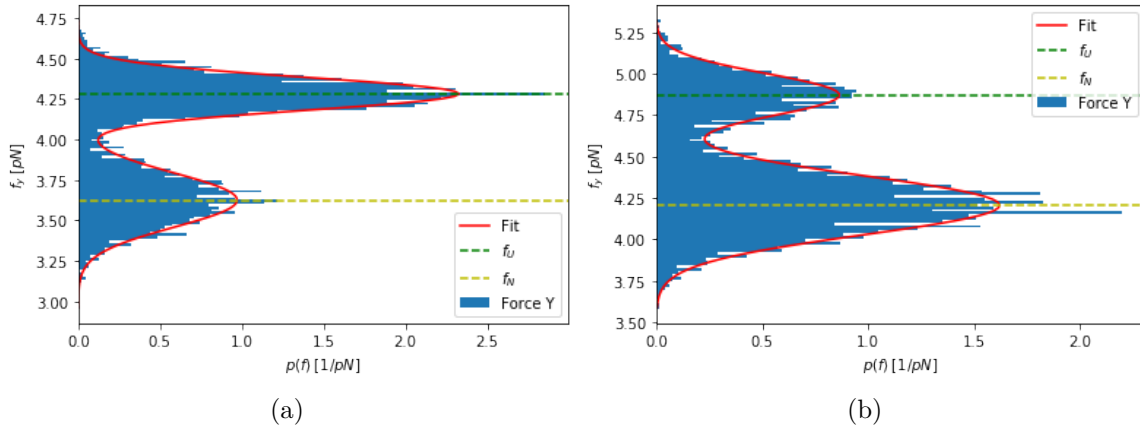


Figura 3.3: Esempi di istogrammi con fit gaussiano: (a) $f > f_c$. (b) $f < f_c$

I grafici in Figura [3.3] si riferiscono a due misure di *hopping* realizzate sulla stessa molecola ad una forza maggiore (sinistra) e inferiore (destra) rispetto a quella di coesistenza. Risulta evidente come la probabilità di trovare la molecola nella stato nativo o denaturato vari sensibilmente al variare della forza, traducendosi in una diversa altezza delle due gaussiane che compongono l'istogramma di forze. In rosso è evidenziata la curva di interpolazione degli istogrammi, con

i valori medi f_N e f_U trovati. Per ottenere informazioni sull'energia libera di formazione del biosensore, si possono adoperare modelli cinetici, primo fra tutti il modello Bell-Evans [14]. Sotto l'ipotesi di un sistema a due stati, il modello ha lo scopo di caratterizzare i fenomeni di *folding* e *unfolding* di un *hairpin* di DNA. Indicata con B_{TOT} l'ampiezza complessiva della barriera energetica esistente tra le due configurazioni, si dimostra che essa dipenda linearmente dalla forza, secondo la relazione

$$B_{TOT} = B_{NU} - f \cdot x_{N-TS}$$

con B_{NU} altezza della barriera cinetica a forza nulla e x_{N-TS} la distanza tra lo stato di transizione (TS), dove è collocata la barriera B_{TOT} , e lo stato nativo N del sistema (stato *folded*).

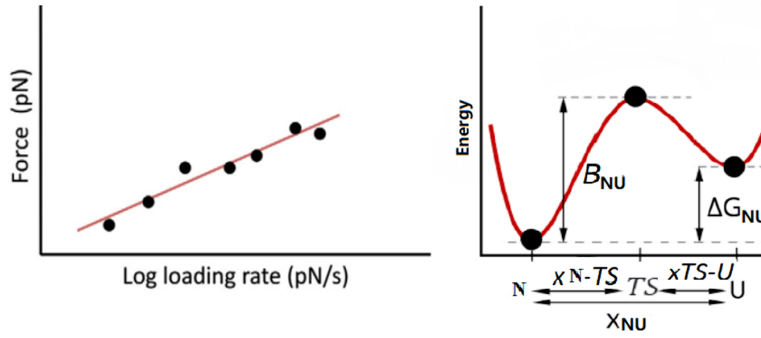


Figura 3.4: Modello Bell-Evans [15]

Il modello di Bell-Evans prevede che per le costanti cinetiche vi è un andamento esponenziale dato dalle equazioni:

$$\begin{aligned} k_{N \rightarrow U}(f) \equiv k^+(f) &= k_0 \exp[-\beta(B_{NU} - f \cdot x_{N \rightarrow TS})] \\ &= k_m \exp[\beta(f \cdot x_{N \rightarrow TS})] \end{aligned} \quad (3.2)$$

$$\begin{aligned} k_{U \rightarrow N}(f) \equiv k^-(f) &= k_0 \exp[-\beta(B_{NU} - \Delta G_{NU} + f \cdot x_{TS \rightarrow U})] \\ &= k_m \exp[-\beta(f \cdot x_{TS \rightarrow U} - \Delta G_{NU})] \end{aligned} \quad (3.3)$$

Le distanze tra lo stato di transizione e gli stati *native* e *unfolded* sono rappresentate da $x_{N \rightarrow TS} = x_{TS} - x_N$ e $x_{TS \rightarrow U} = x_U - x_{TS}$; la loro somma restituisce l'estensione complessiva della molecola: $x_{NU} = x_{N \rightarrow TS} + x_{TS \rightarrow U} = x_U - x_N$. ΔG_{NU} , come mostrato in Figura [3.4] è la differenza di energia tra le due configurazioni e il termine B_{NU} è stato incluso nell'ultimo passaggio nel fattore $k_m = k_0 \exp(-\beta B_{NU})$.

Alla forza di coesistenza f_c si ha che

$$k^+(f_c) \equiv k^-(f_c) \implies \Delta G_{NU} = f_c(x_{N \rightarrow TS} + x_{TS \rightarrow U}) = f_c \cdot x_{NU}. \quad (3.4)$$

3.1.1 Analisi dati

A partire dalle equazioni [3.2] e [3.3] si può esprimere il rapporto tra le costanti cinetiche come

$$\begin{aligned} \frac{1}{\beta} \ln \left(\frac{k^+(f)}{k^-(f)} \right) &= k_B T \ln \left(\frac{w_U(f)}{w_N(f)} \right) = \\ -\Delta G_{NU} + f \cdot (x_{N \rightarrow TS} + x_{TS \rightarrow U}) &= (f - f_c) \cdot x_{NU} \end{aligned} \quad (3.5)$$

dove si è usata la relazione [3.4] nell'ultimo passaggio.

La variabile f che compare nell'equazione [3.5] è data dalla media delle forze f_U e f_N stimate in precedenza tramite il fit gaussiano degli istogrammi ottenuti a partire dalle misure di *hopping*. Un esempio di interpolazione è mostrato nella figura sottostante.

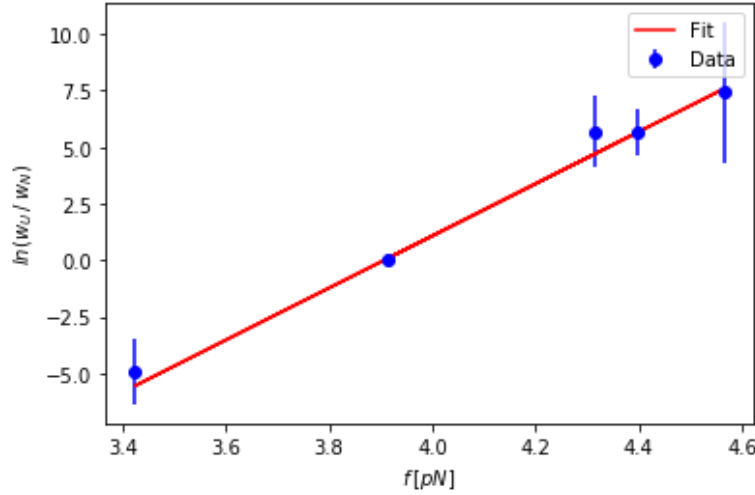


Figura 3.5: Fit lineare $\ln(\frac{w_U}{w_N})$

Quella in [3.5] è una relazione lineare del tipo $y = ax + b$, in cui è facile trovare il coefficiente angolare $a = x_{NU}$ e l'intercetta $b = -f_c \cdot x_{NU}$; è immediato allora ricavare la variabile d'interesse $f_c = -\frac{b}{a}$, da cui è possibile ottenere l'energia libera di formazione ΔG_{NU} usando la [3.4]. Tutte le incertezze sono state ricavate attraverso opportune propagazioni degli errori, eseguendo sviluppi in serie di Taylor, a partire dagli errori sui parametri stimati dal fit.

In Tabella [3.1] si riportano i diversi valori stimati per le molecole analizzate, dove il numero che distingue un biosensore da un altro fa semplicemente riferimento ad un ordinamento cronologico al momento della presa dati: le biomolecole seppur diverse, sono tra loro identiche, cioè caratterizzate dalla stessa sequenza nucleotidica.

L'energia libera di formazione ottenuta usando le equazioni [3.4] e [3.5] risulta però un'approssimazione, dal momento che essa viene calcolata come estrapolazione a forza nulla partendo da valori cinetici calcolati in prossimità della forza di coesistenza f_c . Infatti, la linearità tra $\ln\left(\frac{k^+(f)}{k^-(f)}\right)$ e $\ln\left(\frac{w_U(f)}{w_N(f)}\right)$ con la barriera energetica esistente tra gli stati N e U della molecola prevista dal modello di Bell-Evans è sicuramente valida solo in prossimità della forza di coesistenza. A forze distanti da questa (quindi anche in prossimità di $f = 0$ pN) si potrebbero però osservare significative deviazioni, compromettendo una corretta stima di ΔG_{NU} .

Come già visto nella formula [3.5], l'energia libera tra gli stati nativo (N) e unfolded (U) della molecola dipende dalla forza applicata al sistema, ossia:

$$\frac{1}{\beta} \ln\left(\frac{k^+(f)}{k^-(f)}\right) = k_B T \ln\left(\frac{w_U(f)}{w_N(f)}\right) = -\Delta G_{NU}(f) \quad (3.6)$$

dove il termine $\Delta G_{NU}(f)$ tiene conto dell'energia libera di formazione della molecola a forza nulla e del lavoro fatto per allineare il dipolo di DNA e per allineare l'intera catena nucleotidica fino alla forza di coesistenza f_c . Nel seguito l'energia di formazione a forza nulla verrà indicata con ΔG_0 , per differenziarla dal corrispondente termine ΔG_{NU} calcolato tramite il modello di

Bell-Evans.

Risulta quindi che l'energia libera tra i due stati della molecola ad una generica forza f può essere ottenuta come

$$\Delta G_{NU}(f) = \Delta G_0 - \int_0^f x_{ssDNA}(f')df' + \int_0^f x_d(f')df' \quad (3.7)$$

dove x_d e x_{ssDNA} sono le estensioni del dipolo e della catena di nucleotidi, come descritto dal modello *WLC* discusso precedentemente nella sezione [2.2.2].

Poiché alla forza di coesistenza f_c lo stato *native* ed *unfolded* hanno la stessa energia libera, si avrà $\Delta G_{NU}(f_c) = G_U(f_c) - G_N(f_c) = 0$, da cui risulta che

$$\Delta G_0 = \int_0^{f_c} x_{ssDNA}(f')df' - \int_0^{f_c} x_d(f')df' \quad (3.8)$$

L'incertezza è stata questa volta stimata a partire dall'equazione [3.8] considerando una distribuzione triangolare tra $\Delta G_0(f_c + \sigma_{f_c})$ e $\Delta G_0(f_c - \sigma_{f_c})$.

Tabella 3.1: Stime di f_c attraverso w_N , w_U

N° molecola	$f_c \pm \sigma_{f_c} [pN]$	$x_{NU} \pm \sigma_{x_{NU}} [nm]$	$\Delta G_{NU} \pm \sigma_{\Delta G_{NU}} [pN \cdot nm]$	$\Delta G_0 \pm \sigma_{\Delta G_0} [pN \cdot nm]$
2	4.3 ± 0.5	10.5 ± 0.9	45 ± 34	35 ± 3
2 (cp)	3 ± 1	6 ± 2	20 ± 5	24 ± 6
5 (cp)	3.9 ± 0.7	12 ± 2	45 ± 6	30 ± 4
6	4.4 ± 0.9	9 ± 1	39 ± 6	36 ± 6
6 (cp)	4 ± 2	15 ± 4	60 ± 18	33 ± 9
14	5 ± 3	14 ± 6	70 ± 30	43 ± 16
14 (cp)	5 ± 1	23 ± 4	110 ± 20	41 ± 6
15	4.0 ± 0.7	11 ± 1	42 ± 5	31 ± 3
18	3.9 ± 0.6	13 ± 1	49 ± 5	30 ± 3
19	4.2 ± 0.5	5.5 ± 0.5	23 ± 2	33 ± 3
19 (cp)	4.2 ± 0.5	9.5 ± 0.8	39 ± 3	33 ± 3
20	3.5 ± 0.7	8 ± 1	30 ± 5	25 ± 4
20 (cp)	4 ± 1	12 ± 2	41 ± 9	25 ± 5
21	4.3 ± 0.6	3.2 ± 0.3	14 ± 2	35 ± 3
22	4.5 ± 0.6	7.5 ± 0.7	34 ± 3	37 ± 3
22 (cp)	4 ± 3	13 ± 6	60 ± 30	36 ± 15

Si evidenzia come le ΔG_0 siano tutte tra loro compatibili e si nota inoltre come in un solo caso su sette la modalità *cp* produca gli stessi risultati della modalità *pm*, e che, invece, in quattro casi fornisca stime più in accordo con le altre misure in modalità *pm*. Non si è pertanto in grado di stabilire, da questo set di misure, quale sia la modalità più consona per eseguire le misure.

3.2 Tempo di residenza totale

L'analisi cinetica presentata nel paragrafo precedente può anche essere svolta prendendo come riferimento il tempo di residenza totale della molecola nello stato aperto o nello stato chiuso ad una data forza. In particolare, si può definire $t^\pm(f) = \frac{1}{k^\pm(f)}$. Allora:

$$t^+(f) \equiv t_U = \frac{1}{k_m} \exp[-\beta(f \cdot x_{N \rightarrow TS})] \quad (3.9)$$

$$t^-(f) \equiv t_N = \frac{1}{k_m} \exp[\beta(f \cdot x_{TS \rightarrow U} - \Delta G_{NU})] \quad (3.10)$$

Si nota immediatamente come, all'aumentare della forza, il tempo nello stato nativo, t_N , decresce esponenzialmente, a differenza invece del tempo nello stato aperto, t_U , che invece segue una crescita esponenziale. Questi due nuovi parametri prendono il nome di *tempi di residenza totali*, rispettivamente dello stato *native* e *unfolded*.

Si ricava allora:

$$\begin{aligned} \ln t_U(f) &= -\beta(f \cdot x_{N \rightarrow TS}) - \ln k_m \\ &=: m_1 f + q_1 \\ \ln t_N(f) &= \beta(f \cdot x_{TS \rightarrow U} - \Delta G_{NU}) - \ln k_m \\ &= (\beta \cdot x_{TS \rightarrow U})f - \ln k_m - \beta \Delta G_{NU} \\ &=: m_2 f + q_2 \end{aligned} \quad (3.11)$$

con $q_1 = -\ln k_m = -\beta B_{NU} - \ln k_0$, $m_1 = -\beta \cdot x_{N \rightarrow TS}$, $q_2 = -\ln k_m - \beta \Delta G_{NU} = q_1 - \beta \Delta G_{NU}$ e $m_2 = \beta \cdot x_{TS \rightarrow U}$.

Da $(m_2 - m_1) = \beta \cdot x_{NU}$ si ricava $\beta = \frac{m_2 - m_1}{x_{NU}}$ utile per ottenere una stima di

$$\Delta G_{NU} = \frac{q_1 - q_2}{\beta} \quad (3.12)$$

Alla forza di coesistenza, si ricava il *tempo di coesistenza*, invertendo le formule [3.11] in funzione della variabile indipendente f , ricavando:

$$t_c = \exp\left(\frac{q_4 - q_3}{m_3 - m_4}\right) \quad (3.13)$$

dove $q_3 = -\frac{q_1}{m_1}$, $m_3 = \frac{1}{m_1}$, $q_4 = -\frac{q_2}{m_2}$ e $m_4 = \frac{1}{m_2}$ sono ricavati partire dalle precedenti equazioni.

La forza di coesistenza è data risolvendo il sistema $\begin{cases} \frac{(f_c - q_4)}{m_4} = \ln t_c \\ \frac{(f_c - q_3)}{m_3} = \ln t_c \end{cases}$ trovando:

$$f_c = \frac{q_4 m_3 - q_3 m_4}{m_3 - m_4} = \frac{q_1 m_2 - q_2 m_1}{m_2 - m_1} \quad (3.14)$$

3.2.1 Analisi dati

Le variabili t_N e t_U sono state ricavate dalle misure di *hopping* contando i secondi in cui la molecola si trova nello stato aperto e quelli in cui si trova nello stato chiuso. Per farlo, si sono considerate come indipendenti le gaussiane che costituiscono l'equazione [3.1]: $P(f) = p_U(f) + p_N(f)$. Per ogni valore di forza si è calcolata la probabilità $p_i(f)$, $i = U, N$ della molecola di trovarsi in uno dei due possibili stati e si è interpretato il valore maggiore di probabilità tra $p_N(f)$ e $p_U(f)$ come fattore discriminante per stabilire la configurazione della molecola. Contando tutte le volte in cui il biosensore era nello stadio N si è ottenuto t_N e in modo perfettamente analogo t_U . Questa procedura è stata successivamente migliorata adoperando un algoritmo di *Hidden Markov Model* discusso in sezione [3.4]

In Figura [3.6], a titolo di esempio, si mostra l'andamento del logaritmo di t_U e t_N al variare della forza media applicata al sistema.

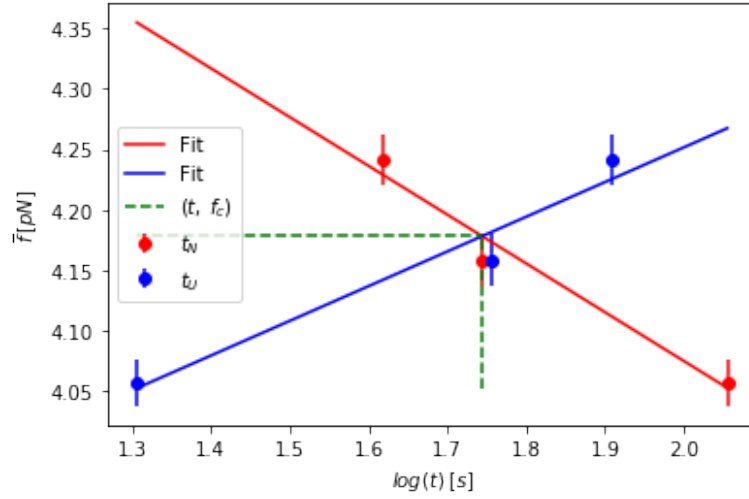


Figura 3.6: Fit lineare dei tempi di residenza

Le rette tratteggiate in verde in figura rappresentano il punto d'intersezione delle due rette, che individuano il punto (t, f_c) , con $t = \ln(t_c)$: dalle equazioni [3.12] e [3.14] si ottiene una nuova stima per la *forza di coesistenza* f_c e la variazione di energia libera tra le configurazioni aperta e chiusa ΔG_{NU} ; nella Tabella [3.2] sottostante si riportano le stime ottenute eseguendo interpolazioni lineari delle equazioni [3.11] e, nelle ultime due colonne, le compatibilità r dei parametri d'interesse con i risultati esposti nella precedente Tabella [3.1].

Tabella 3.2: Stime di f_c attraverso t_c

N° molecola	$f_c \pm \sigma_{f_c}$ [pN]	$t_c \pm \sigma_{t_c}$ [s]	$\Delta G_{NU} \pm \sigma_{\Delta G_{NU}}$ [pN · nm]	r_{f_c}	$r_{\Delta G_{NU}}$
2	4.306 ± 0.008	5.06 ± 0.07	45 ± 4	0.0	0.0
2 (cp)	3.48 ± 0.01	4.73 ± 0.05	21 ± 5	0.5	0.1
5 (cp)	3.90 ± 0.01	4.83 ± 0.07	45 ± 6	0.0	0.0
6	4.44 ± 0.07	5.7 ± 0.4	39 ± 6	0.0	0.0
6 (cp)	4.21 ± 0.08	5.6 ± 0.9	60 ± 20	0.1	0.0
14	4.918 ± 0.009	2.29 ± 0.03	70 ± 30	0.0	0.0
14 (cp)	4.78 ± 0.02	3.9 ± 0.2	110 ± 20	0.2	0.0
15	4.01 ± 0.01	4.14 ± 0.05	43 ± 5	0.0	0.0
18	3.914 ± 0.006	3.03 ± 0.03	49 ± 5	0.0	0.0
19	4.153 ± 0.005	3.64 ± 0.03	23 ± 2	0.1	0.0
19 (cp)	4.185 ± 0.009	3.61 ± 0.06	40 ± 4	0.0	0.0
20	3.587 ± 0.009	3.25 ± 0.04	30 ± 5	0.1	0.0
20 (cp)	3.56 ± 0.01	4.05 ± 0.05	41 ± 8	0.4	0.0
21	4.37 ± 0.01	3.08 ± 0.01	14 ± 1	0.1	0.0
22	4.446 ± 0.005	2.47 ± 0.02	33 ± 3	0.1	0.2
22 (cp)	4.50 ± 0.03	3.1 ± 0.2	60 ± 30	0.2	0.0

3.3 Stime conclusive

Entrambi i metodi di analisi producono gli stessi risultati con compatibilità tutte ottime.

Diverso è invece se si confrontano tra loro i risultati per le diverse molecole, ottenendo valori anche molto incompatibili tra di loro: consideriamo allora nel prosieguo dell'analisi solamente le stime tra loro compatibili. Non tutte le interpolazioni infatti sono state eseguite considerando

lo stesso numero di punti e molti dati sono stati scartati a causa di effetti di *drift*, dovuti ad un aumento della forza applicata, nell'eseguire misure di *hopping*, all'aumentare del tempo di campionamento. Si evidenzia inoltre come questi due effetti combinati influenzino la variabile t_c portando quindi anche a incompatibilità di risultati ottenuti, per la stessa molecola, a differenza di modalità adoperata (*passive* o *constant position*): questi dati, come ad esempio la molecola 2, sono stati scartati.

Dall'intero campione di dati raccolto, si sceglie allora di lavorare unicamente con un subset, scartando dunque le molecole - facendo riferimento alle Tabelle [3.1] e [3.2] soprastanti - 2 (*cp*), 6(*cp*), 14(*cp*), 19, 20, 21, 22. Si fornisce allora una stima finale per ciascun parametro d'interesse, ottenendo:

$$\begin{aligned} f_c &= (4.1 \pm 0.2) \text{ pN} \\ x_{\text{NU}} &= (10.6 \pm 0.4) \text{ nm} \\ \Delta G_{\text{NU}} &= (44 \pm 2) \text{ pN} \cdot \text{nm} \\ \Delta G_0 &= (32 \pm 1) \text{ pN} \cdot \text{nm} \end{aligned}$$

dove per f_c si sono impiegati i dati riportanti in Tabella [3.1].

3.4 Hidden Markov Model - HMM

Nel corso del presente lavoro si è assunto di poter lavorare con due stati, *folded* e *unfolded*, ma in realtà il biosensore, quando sottoposto ad uno stress esterno, può passare attraverso stati distinti di equilibrio. Inoltre, lavorando a forze molto basse - poche unità di *pN* - il rumore sperimentale influenza sensibilmente le misure, rendendo difficoltosa la corretta catalogazione degli stati.

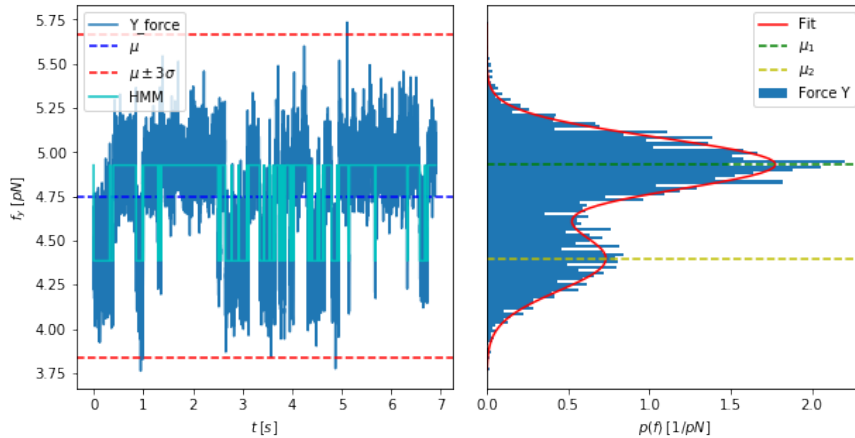


Figura 3.7: Analisi di hopping con HMM a 2 hidden states

Per capire qual è il numero effettivo di stati, detti *hidden states*, si può ricorrere ad un metodo algebrico-matematico, detto *Hidden Markov Model (HMM)*. Il modello implementa tre diversi algoritmi ([16], [17], [18]) atti a stimare quale sia la sequenza più probabile di configurazioni raggiungibili dal sistema e il loro numero, a partire dai dati raccolti in laboratorio:

- il forward-backward algorithm è impiegato per stimare la probabilità $P(O|\Delta)$ della sequenza di dati osservata, O , a partire dai parametri Δ del modello. Questo costituisce il cosiddetto *evaluation problem*;

- il Viterbi algorithm è usato per trovare la corretta sequenza di stati X ($X=\{native, unfolded\}$ nel caso di due stati), in modo tale da massimizzare la probabilità di produrre la sequenza O di osservazioni. Questo rappresenta il cosiddetto *uncovering problem*;
- il Baum-Welch algorithm, o Expectation Maximization (EM) algorithm, risolve invece il *learning problem*, cioè aggiusta i parametri Δ del modello in modo da migliorare la qualità dell'*HMM*.

I parametri Δ sono generalmente una terna $\Delta = \{\Pi, A, B\}$, dove Π rappresenta la probabilità iniziale, al tempo $t_0 = 0$, del sistema di trovarsi in uno degli stati possibili e A e B sono due matrici. In questo caso, per Π , si è scelta una distribuzione uniforme, in modo tale che ci fosse il 50% di probabilità che il sistema si trovasse in uno dei due stati a t_0 . A è denominata *matrice di transizione di probabilità* e rappresenta la probabilità di ogni stato di evolvere in un altro stato o di rimanere in se stesso, nel tempo. Per il primo step, a t_0 , si è usata la matrice esempio fornita in [17]. B , invece, è la *observation probability matrix*: da [19] si evince come in questo caso, lavorando con una sequenza O di numeri reali, non è possibile rappresentare la transizione da un dato ad un altro attraverso una matrice finito-dimensionale, ma è necessario ricorrere ad una *Probability Density Function (PDF)*. Si è fatto allora uso dei risultati riportati nella sezione precedente, impiegando una PDF normale gaussiana bidimensionale. Poiché si estende l'*HMM* a osservazioni in uno spazio continuo invece che discreto, il modello prende propriamente il nome di *Continuous Density Hidden Markov Model (CDHMM)*.

Nella Figura [3.7] è evidenziato a destra il fit gaussiano della forza applicata, mentre a sinistra l'andamento temporale di quest'ultima, con evidenziata la sequenza di stati nascosti che mettono in luce l'*hopping* tra le due configurazioni aperta-chiusa dell'*hairpin* di DNA.

Nella Figura [3.8], invece, si riporta un confronto tra il primo algoritmo introdotto euristica-mente in sezione [3.2.1] per la classificazione degli stati *native* e *unfolded* e la classificazione degli stati che si ottiene implementando un algoritmo *HMM*. Risulta evidente che l'uso di un algoritmo *HMM* porta ad un netto miglioramento nella classificazione degli stati, permettendo di isolare questi ultimi dal solo rumore strumentale, mostrando un netto miglioramento nella computazione di t_N e t_U .

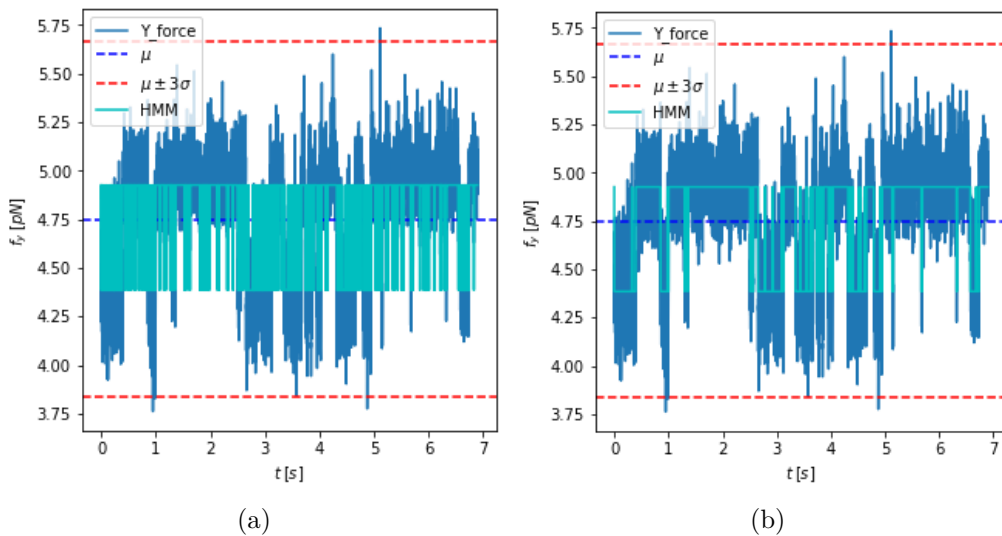


Figura 3.8: Confronto grafico tra (a) l'algoritmo euristico in sezione [3.2.1] e (b) *HMM*

3.5 Legame tra molecola e oligo-nucleotide

Si sono infine condotti esperimenti con *oligo-nucleotidi*, cioè piccole (oligo) sequenze di DNA, aggiunte in soluzione, che fungevano da *target* riconosciuto dal loop dell'*hairpin* studiato.

Nel condurre l'esperimento poteva accadere che una delle sequenze oligo si attaccasse alla molecola di DNA oggetto di studio, e ne modificasse la cinetica. Questo è esattamente ciò che si osserva: l'oligo può attaccarsi in uno dei tre siti disponibili dell'*hairpin*, oppure più di un oligo possono attaccarsi alla stessa molecola [20].

Il legame tra l'*hairpin* e il *target* fa sì che lo stato nativo della prima risulti più debole, in altre parole che l'energia di formazione sia minore, favorendo quindi l'apertura della molecola, che si osserva a forze minori rispetto al caso in cui il *target* non è presente. La cinetica del complesso *hairpin-target* risulta quindi modificata rispetto al caso del solo *hairpin*, come riportato nelle Figure [3.9], [3.10] e [3.11].

Il comportamento ora descritto è esattamente quello che si osserva in laboratorio: nella Figura [3.9] è *fotografato* il momento in cui l'*oligo* si lega al filamento di DNA. Si osserva infatti come il fit lineare non produca risultati sensati, ma è possibile interpretare il grafico come se si stessero analizzando contemporaneamente due molecole distinte.

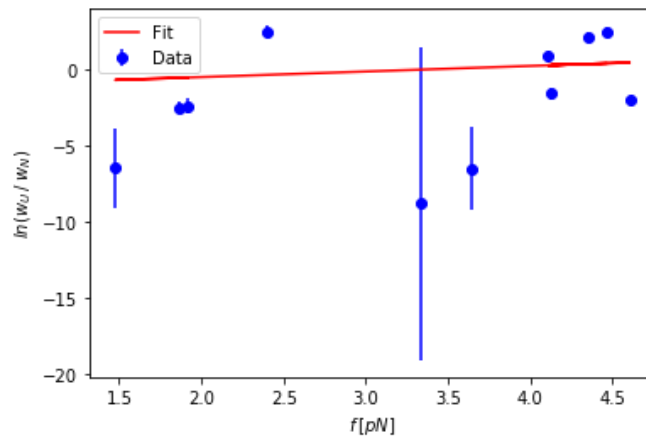


Figura 3.9: Cambiamento della cinetica del biosensore dovuta all'oligo-nucleotide

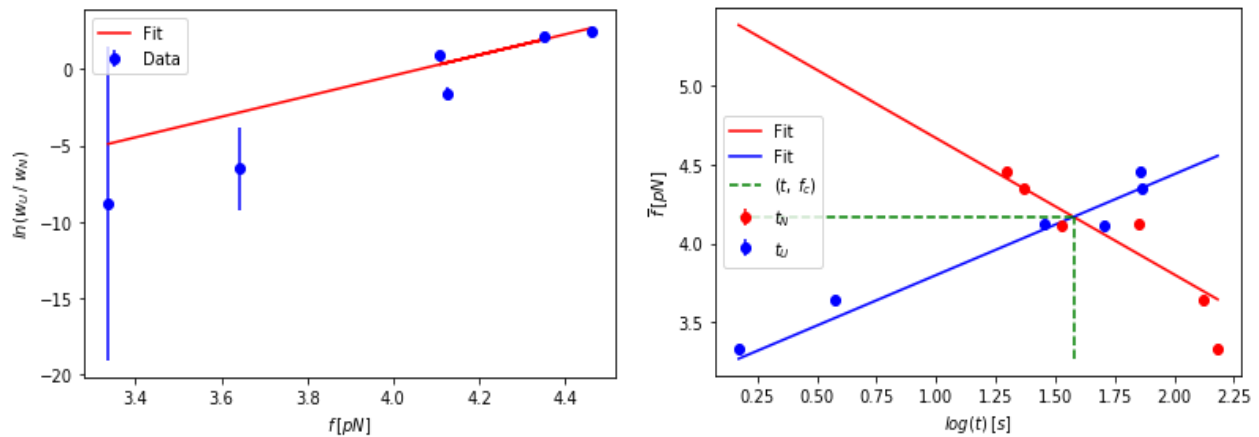
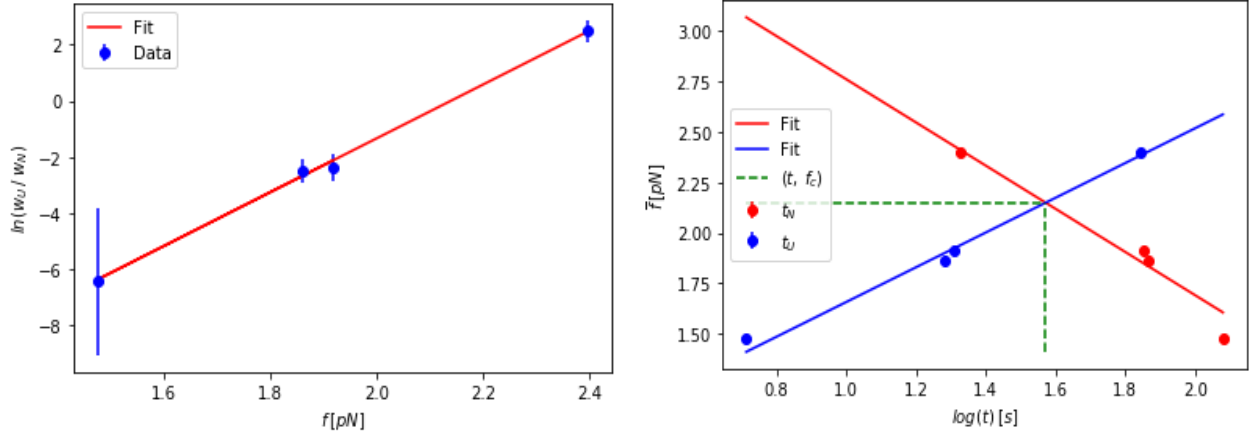


Figura 3.10: Molecola *prima* dell'attaccamento dell'*oligo*

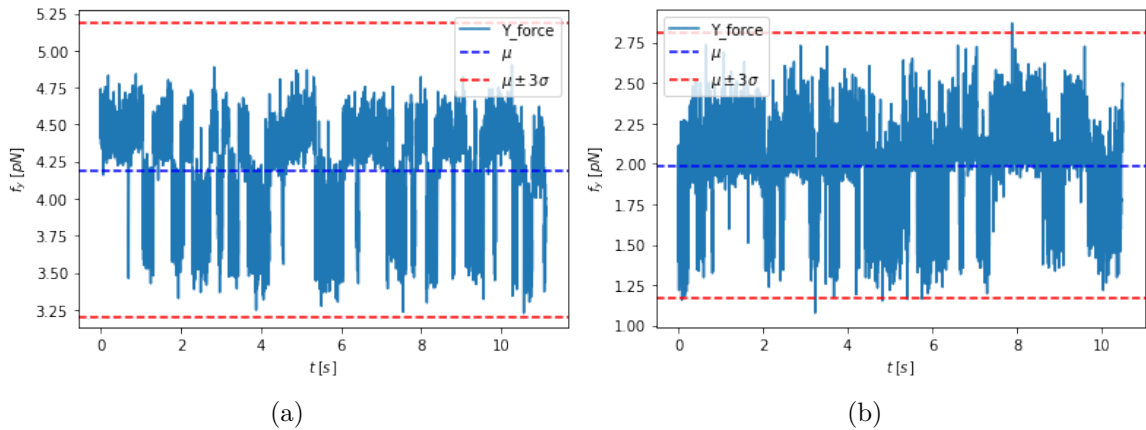
Figura 3.11: Molecola *dopo* l'attaccamento dell'*oligo*

Dalla tabella sottostante si osserva come i risultati stimati dai fit della Figura [3.10] sono in accordo con quelli trovati precedentemente (sezione [3.3]), mentre le stime ottenute a partire dai fit in Figura [3.11] evidenziano un abbassamento della *forza di coesistenza*.

Tabella 3.3: Stime *pre* e *post-oligo*

	$f_c \pm \sigma_{f_c}$ [pN]	$t_c \pm \sigma_{t_c}$ [s]	$\Delta G_{NU} \pm \sigma_{\Delta G_{NU}}$ [pN · nm]	$x_{NU} \pm \sigma_{x_{NU}}$ [nm]	ΔG_0 [pN · nm]
<i>pre-oligo</i>	4.1 ± 0.8	4.8 ± 0.1	27 ± 4	7 ± 1	32 ± 4
<i>post-oligo</i>	2.1 ± 0.3	4.8 ± 0.1	20 ± 2	10 ± 1	11 ± 1

Si nota inoltre l'aumento della lunghezza complessiva del biosensore, x_{NU} e l'abbassamento sostanziale dell'energia libera di formazione a forza zero ΔG_0 . La molecola tende pertanto a preferire lo stato aperto, poiché la doppia catena è più rigida del singolo filamento di *hairpin*. Quest'ultimo è pertanto penalizzato e risulta più semplice portarlo nello stato *unfolded*. Si mostrano infine due tracce di misure di *hopping* compiute vicino alla *forza di coesistenza* nei due casi con e senza l'*oligo-nucleotide*.

Figura 3.12: Confronto fra due tracce misure di hopping: (a) *pre-oligo* e (b) *post-oligo*

Si nota come le cinetiche dei due stati sono simili seppur le forze in gioco sono diverse: la molecola rimane circa lo stesso tempo nei due stati, con una forza due volte più piccola nel caso in cui l'*oligo-nucleotide* si lega all'*hairpin* di DNA.

Conclusioni

Il presente lavoro di tesi ha avuto l'obiettivo di studiare la cinetica di un *hairpin* di DNA quando è applicata, alle sue estremità, una forza, attraverso uno strumento di pinzette ottiche. Si è raggiunto tale obiettivo compiendo misure di *hopping*: applicando alla biomolecola una forza vicina alla *forza di coesistenza*, essa esegue spontaneamente numerosi passaggi tra le due possibili configurazioni *native* e *unfolded*.

Per verificare la presenza di due stati e capire ciascun punto sperimentale a quale delle due configurazioni appartiene, si è eseguito un *Hidden Markov Model*, studiando l'evoluzione delle forze campionate nel tempo.

Applicando modelli come il *Worm-Like Chain* per ottenere una descrizione matematica del biosensore di DNA impiegato, è stato possibile dare una stima di quantità fisiche fondamentali, come la *forza di coesistenza* $f_c = (4.1 \pm 0.2) \text{ pN}$ e l'energia libera di formazione dell'*hairpin* a forza zero $\Delta G_0 = (32 \pm 1) \text{ pN} \cdot \text{nm}$, ricorrendo ulteriormente al modello di Bell-Evans che considera i due possibili stati della molecola come separati da una barriera energetica la cui ampiezza dipende dalla forza applicata al sistema.

Si è inoltre eseguita un'analisi di una molecola di *oligo-nucleotide*, verificando che l'attacco di una piccola sequenza di DNA ad uno dei siti attivi dell'*hairpin* ne influenza la cinetica, rendendo il biosensore più rigido e favorendone la configurazione aperta. Questo si è riscontrato nella stima della *forza di coesistenza* $f_c = (2.1 \pm 0.3) \text{ pN}$, inferiore rispetto al caso precedente, così come dell'energia libera di formazione, pari in questo caso a $\Delta G_0 = (11 \pm 1) \text{ pN} \cdot \text{nm}$.

Una stima migliore delle energie in gioco si potrebbe ottenere aumentando la statistica degli esperimenti o adottando modelli teorici più complessi che tengano conto, per esempio, di più stati di equilibrio, dunque di più barriere energetiche e di più stati di transizione.

Bibliografia

1. Moffitt, J. R., Chemla, Y. R., Smith, S. B. & Bustamante, C. Recent advances in optical tweezers. *Annu. Rev. Biochem.* **77**, 205–228 (2008).
2. Griffiths, D. J. INTRODUCTION TO ELECTRODYNAMICS Fourth Edition (2021).
3. Luise, A. Studio del meccanismo di apertura e ripiegamento di un biosensore di DNA tramite l'uso di pinzette ottiche.
4. Smith, S. B., Cui, Y. & Bustamante, C. in *Methods in enzymology* 134–162 (Elsevier, 2003).
5. Zaccaria, C. Optical mini-Tweezers: un esperimento di singola molecola (2018).
6. Bustamante, C. J., Chemla, Y. R., Liu, S. & Wang, M. D. Optical tweezers in single-molecule biophysics. *Nature Reviews Methods Primers* **1**, 1–29 (2021).
7. Zaltron, A., Merano, M., Mistura, G., Sada, C. & Seno, F. Optical tweezers in single-molecule experiments. *The European Physical Journal Plus* **135**, 896 (2020).
8. Del Grosso, E., Amodio, A., Ragazzon, G., Prins, L. J. & Ricci, F. Dissipative Synthetic DNA-Based Receptors for the Transient Loading and Release of Molecular Cargo. *Angewandte Chemie* **130**, 10649–10653 (2018).
9. Nelson, P. *Biological physics* (WH Freeman New York, 2004).
10. Bouchiat, C. *et al.* Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical journal* **76**, 409–413 (1999).
11. Alemany i Arias, A. Dynamic force spectroscopy and folding kinetics in molecular systems (2014).
12. Landuzzi, F., Viader-Godoy, X., Cleri, F., Pastor, I. & Ritort, F. Detection of single DNA mismatches by force spectroscopy in short DNA hairpins. *The Journal of Chemical Physics* **152**, 074204 (2020).
13. Wen, J.-D. *et al.* Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophysical journal* **92**, 2996–3009 (2007).
14. Evans, E. Probing the relation between force—lifetime—and chemistry in single molecular bonds. *Annual review of biophysics and biomolecular structure* **30**, 105–128 (2001).
15. Whited, A. M. & Park, P. S.-H. Atomic force microscopy: a multifaceted tool to study membrane proteins and their interactions with ligands. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1838**, 56–68 (2014).
16. Jurafsky, D. & Martin, J. H. Hidden markov models. *Speech and Language Processing*, 1024 (2016).

17. McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical journal* **91**, 1941–1951 (2006).
18. Poritz, A. B. *Hidden Markov models: a guided tour*. in *ICASSP* **88** (1988), 7–13.
19. Nguyen, L. Continuous observation hidden Markov model. *Revista Kasma* **44**, 65–149 (2016).
20. Gold, L., Polisky, B., Uhlenbeck, O. & Yarus, M. DIVERSITY OF OLIGONUCLEOTIDE FUNCTIONS. *Annual Review of Biochemistry* **64**. PMID: 7574500, 763–797. doi:10 . 1146/annurev . bi . 64 . 070195 . 003555. eprint: [https://doi.org/10.1146/annurev . bi.64.070195.003555](https://doi.org/10.1146/annurev.bi.64.070195.003555) (1995).