

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Corso di Laurea Magistrale in Fisica

La spettroscopia di forza basata sull'AFM
nello studio dello spazio conformazionale e
dei processi aggregativi di proteine
prioniche.

Relatore:

Prof. Bruno Samorì

Presentata da:

Anna Murello

Correlatore:

Dott. Giampaolo Zuccheri

Dott. Luca Pasquini

Sessione I

Anno Accademico 2014/2015

Abstract

Le malattie neurodegenerative sono caratterizzate da aggregazione proteica, dipendente dalla perdita della usuale struttura fisiologica funzionale delle proteine coinvolte, a favore di conformazioni tossiche (patologiche). Il modello corrente descrive questi cambiamenti conformazionali come eventi rari e ritiene che non esista una sola conformazione patogena, ma che tali possibili conformazioni siano piuttosto eterogenee. La caratterizzazione di queste strutture è, di conseguenza, difficile con le tradizionali tecniche in *bulk* che permettono di studiare solo la conformazione media e non rendono possibile il riconoscimento delle caratteristiche dei conformeri individuali.

Lo sviluppo delle tecniche di singola molecola ha permesso di studiare in modo approfondito le conformazioni possibili. In questo lavoro la spettroscopia di forza di singola molecola basata sull'AFM viene applicata a PrP (proteina responsabile delle encefalopatie spongiformi trasmissibili). Si studiano gli equilibri conformazionali del monomero e quelli di costrutti oligomerici, allo scopo di caratterizzare gli *step* iniziali dei processi aggregativi.

Nel corso di questo lavoro di tesi è stato, in particolare, sviluppato un sistema di analisi dati, al fine di studiare in modo quantitativo le distribuzioni di eventi ottenute. Grazie a tale strumento è stato possibile riconoscere i segnali di *unfolding* della conformazione nativa del monomero e notare come essa sia presente anche in costrutti oligomerici, ad indicare come questo ripiegamento sia stabile anche in presenza di più monomeri ravvicinati. Si è osservato l'effetto del pH sulla stabilità di tale struttura, notando come pH acidi destabilizzino il ripiegamento nativo. Inoltre si è studiato il ruolo dell'orientazione dei monomeri nella formazione di strutture dimeriche.

Monomeri e oligomeri di PrP sono stati descritti come proteine parzialmente strutturate il cui panorama energetico contiene molti minimi locali, dando origine a parecchie conformazioni transienti.

Questo lavoro di ricerca è stato svolto presso il laboratorio di nano-biotecnologie del dipartimento di farmacia e biotecnologie dell'università di Bologna e si inserisce all'interno di una collaborazione del gruppo dei Prof. Samorì e Zuccheri con il gruppo del Prof. Legname (SISSA di Trieste), il gruppo del Prof. Bustamante (UC Berkeley) ed il gruppo del Prof. Tanaka (*Riken Brain Science Institute*, Tokyo), finanziata dallo *Human Frontiers Science Program* e finalizzata alla caratterizzazione della stabilità meccanica delle proteine prioniche.

Indice

I	Introduzione	1
1	Il problema biologico	2
1.1	Le malattie prioniche come modello per le malattie neurodegenerative	2
1.2	Struttura di PrP	3
2	Spettroscopia di forza di singola molecola (SMFS)	5
2.1	Tecniche di singola molecola	5
2.2	<i>Unfolding</i> meccanico	5
2.2.1	Modello di Bell	5
2.2.2	Distribuzione delle forze di <i>unfolding</i> in stiramenti delle proteine a velocità costante	7
2.2.3	Modelli di fit	7
2.3	Tecnica sperimentale: SMFS con l'utilizzo di un AFM	11
2.3.1	Microscopio a forza atomica	11
2.3.2	Spettroscopia di forza	12
2.3.3	Esperimenti con il metodo del doppio impulso	14
2.3.4	La tecnica dei costrutti multimodulari	16
2.3.5	Energia dissipata nella transizione	16
3	Lo scopo del progetto	18
II	Materiali e metodi	19
4	Il campione biologico	20
4.1	I costrutti utilizzati	20
4.2	La sintesi dei costrutti	20
4.3	<i>Refolding</i> in condizioni riducenti	20
5	SMFS	22
5.1	Esperimenti di <i>unfolding</i>	22
5.2	Esperimenti con il metodo del doppio impulso	23
6	Analisi dati	24
6.1	Analisi delle curve	24
6.2	Confronto tra distribuzioni	25
6.2.1	Analisi monodimensionale	26
6.2.2	Analisi bidimensionale	27

III	Risultati e discussione	29
7	Caratterizzazione di GB1	30
7.1	Dati sperimentali	30
7.2	Analisi statistica dei dati	30
7.3	Analisi degli eventi con $\Delta_{Lc} < 15$ nm	31
7.3.1	Coppie di picchi corti	33
7.3.2	Picchi corti singoli	36
7.3.3	Curve prive di interazione aspecifica	37
7.4	Curve con il metodo del doppio impulso	37
7.5	Struttura di GB1	37
7.6	Sommario	40
8	Caratterizzazione di PrP	41
8.1	Esperimenti a pH neutro	42
8.1.1	Monomero in ambiente ossidante e riducente: riconoscimento della struttura nativa.	42
8.1.2	Dimeri: il ruolo dell'orientazione reciproca di MoPrP.	46
8.1.3	Trimero: presenza di struttura nativa e conformazioni associative.	48
8.1.4	Tetramero: identificazione di nuove strutture.	50
8.1.5	Fit preliminare della distribuzione sulla coda destra di GB1	52
8.2	Esperimenti a pH acidi	57
8.2.1	Monomero, condizioni ossidanti	57
8.2.2	Dimero, orientazione H-T	57
8.2.3	Trimero	57
8.2.4	Tetramero	67
8.3	Sommario	67
9	Conclusioni e sviluppi	73
A	Appendice: codice del <i>software</i> per il confronto delle distribuzioni	75
A.1	Analisi monodimensionale	75
A.2	Analisi bidimensionale	77
	Bibliografia	81

Lista delle abbreviazioni

a.a.: amminoacido

AFM: Microscopio a Forza Atomica

DTT: Ditiotreitolo

FJC: *Freely Jointed Chain*

GB1: dominio B1 della proteina G

IDP: Proteina Intrinsecamente Disordinata

MoPrP: Proteina prionica murina

PDB: *Protein Data Bank*

PrP: Proteina prionica

PrP^C: Proteina prionica cellulare

PrP^{Sc}: Proteina prionica scrapie

PrP^{Tr}: Proteina prionica nella forma troncata (a.a. 89-230)

SMFS: Spettroscopia di Forza di Singola Molecola

TSE: Encefalopatie Trasmissibili Spongiformi

Parte I

Introduzione

Capitolo 1

Il problema biologico

Le proteine vengono sintetizzate nei ribosomi come catene amminoacidiche, negli istanti immediatamente successivi avviene il *folding* proteico, ovvero la catena amminoacidica si ripiega su se stessa in modo che i legami che si formano tra gli amminoacidi non consecutivi portino ad una minimizzazione dell'energia interna dell'intera catena.

Molte proteine riescono a raggiungere la loro conformazione stabile spontaneamente e in tempi rapidi (dell'ordine dei μs o ms a seconda della lunghezza e della complessità della proteina). Esistono, però, proteine che non raggiungono mai una conformazione stabile definitiva, rimanendo intrinsecamente disordinate anche nel loro stato biologicamente attivo, senza strutture secondaria e terziaria definite. Per altre proteine il ripiegamento non avviene in modo spontaneo, ma è mediato da proteine ausiliarie, chiamate *chaperon*.

Generalmente il ripiegamento delle proteine non avviene passando dallo stato non strutturato direttamente a quello nativo, ma piuttosto attraverso conformazioni parzialmente strutturate (intermedi), corrispondenti a minimi locali del panorama energetico.

Nonostante il processo di *folding* delle proteine sia altamente controllato all'interno della cellula, è possibile che alcune di queste molecole sfuggano al controllo qualità e si ripieghino in strutture fisiologicamente non attive. Alcune di queste strutture possono essere anche tossiche per la cellula, con conseguenze anche fatali.

1.1 Le malattie prioniche come modello per le malattie neurodegenerative

L'amiloidosi è un gruppo di malattie caratterizzate, dal punto di vista anatomopatologico, da processi degenerativi dei tessuti con infiltrazione di amiloidi (molecole fibrillari insolubili di natura prevalentemente proteica) in quantità sufficienti da comprometterne la funzione normale. Negli ultimi vent'anni si è scoperto che molte di queste malattie sono associate ad aggregazione proteica e *misfolding*. Quest'ultimo è inteso come un ripiegamento non corretto della proteina. Al momento sono circa 50 le malattie, accompagnate da sintomi molto diversi, associabili al ripiegamento errato di peptidi e proteine normalmente solubili e presenti nell'organismo. In tabella 1.1 sono riassunte le caratteristiche principali delle amiloidosi, come riportato in [18]. Circa metà delle proteine coinvolte in queste malattie sono classificate come intrinsecamente disordinate (IDP), altre sono solo parzialmente strutturate, ovvero presentano un dominio foldato ed uno non strutturato.

Le malattie prioniche sono malattie neurodegenerative note anche come encefalopatie spongiformi trasmissibili (TSE) e possono essere presenti sia negli animali che nell'uomo. Possono manifestarsi come malattie infettive, ereditarie o sporadiche. Presentano fenotipi variegati, accompagnati da sintomi motori e cognitivi differenti e sono tuttora incurabili e letali. Alcuni esempi sono la malattia di Creutzfeldt-Jacob e l'insonnia fatale familiare nell'uomo, la *scrapie* negli ovini e l'encefalopatia spongiforme bovina (morbo della mucca pazza) nei bovini.

Disease	Aggregating protein or peptide	Polypeptide length (number of residues)	Structure of protein or peptide
Neurodegenerative diseases			
Alzheimer's disease	Amyloid- β peptide	37–43	Intrinsically disordered
Spongiform encephalopathies	Prion protein or its fragments	230	Intrinsically disordered and α -helical
Parkinson's disease	α -synuclein	140	Intrinsically disordered
Amyotrophic lateral sclerosis	Superoxide dismutase 1	153	β -sheet and Ig-like
Huntington's disease	Huntingtin fragments	Variable	Mostly intrinsically disordered
Familial amyloidotic polyneuropathy	Transthyretin mutants	127	β -sheet
Non-neuropathic systemic amyloidosis			
Amyloid light chain (AL) amyloidosis	Immunoglobulin (Ig) light chains or its fragments	~90	β -sheet and Ig-like
Amyloid A (AA) amyloidosis	Serum amyloid A1 protein fragments	76–104	α -helical and unknown fold
Senile systemic amyloidosis	Wild-type transthyretin	127	β -sheet
Haemodialysis-related amyloidosis	β_2 -microglobulin	99	β -sheet and Ig-like
Lysozyme amyloidosis	Lysozyme mutants	130	α -helical and β -sheet
Non-neuropathic localized amyloidosis			
Apolipoprotein A1 (Apo A-1) amyloidosis	Apo A-1 fragments	80–93	Intrinsically disordered
Type II diabetes	Amylin	37	Intrinsically disordered
Injection-localized amyloidosis	Insulin	21 and 30	α -helical and insulin-like

Tabella 1.1: Riassunto delle caratteristiche principali delle amiloidosi, come riportato in [18].

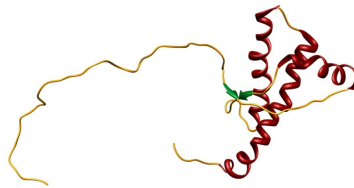


Figura 1.1: Struttura di PrP^C (dall'amminoacido 23 all'aminocido 230): il dominio N-terminale, non strutturato, è indicato in giallo; il dominio C-terminale, strutturato, è prevalentemente composto da eliche α .

La patogenesi è attribuita a cambiamenti conformazionali nella proteina prionica cellulare PrP^C , che è presente nei neuroni, prevalentemente in corrispondenza delle sinapsi, e la cui funzione biologica non è tuttora chiara. PrP^C è altamente conservata tra i mammiferi (ovvero la sequenza amminoacidica varia poco da una specie all'altra) ed è stata identificata anche negli uccelli.

1.2 Struttura di PrP

Nel 1996 Riek *et al.* hanno caratterizzato la struttura di PrP^C utilizzando la risonanza magnetica nucleare [26], la loro descrizione prevede che questa proteina sia composta da due parti (figura 1.1): una prima parte (N-terminale) non strutturata, altamente conservata tra le specie (questo lascia presupporre che abbia una funzione biologica rilevante), e una seconda parte (C-terminale) strutturata e ricca in eliche α . Un ponte disolfuro fra le cisteine 178 e 213 collega due delle eliche α dando stabilità alla struttura. Il dominio C-terminale correttamente ripiegato viene anche indicato come struttura nativa.

La struttura della forma infettiva PrP^{Sc} (*scrapie*, che prende il nome dalla omonima malattia) al contrario non è nota, tuttavia analisi tramite FTIR (*Fourier Transform InfraRed spectroscopy*) hanno mostrato che essa è ricca in foglietti β . Inoltre PrP^{Sc} agisce come

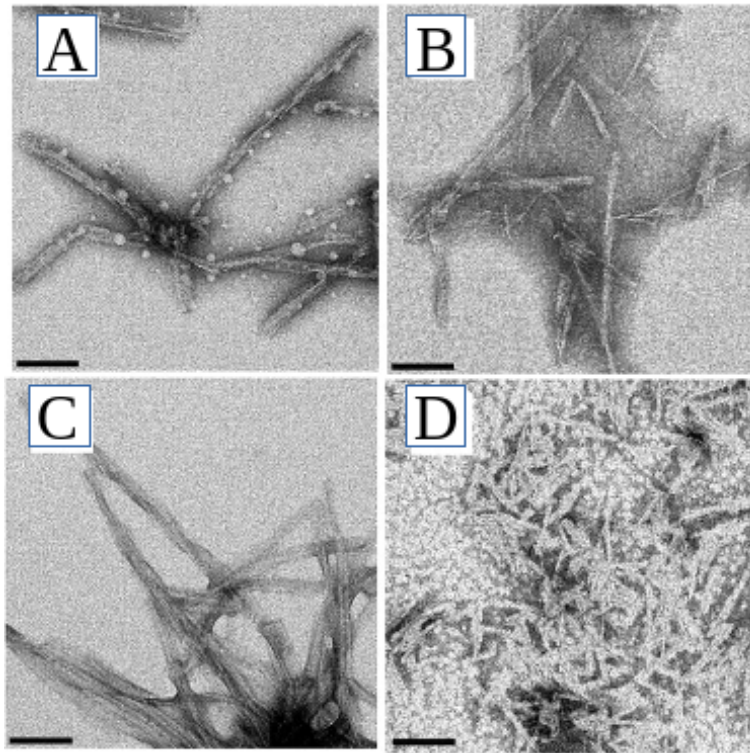


Figura 1.2: Studio di stabilità di fibrille ottenute variando la temperatura e la concentrazione di urea presenti in soluzione (A, B), o utilizzando costrutti contenenti solo alcune parti della proteina prionica (invece che l'intera catena amminoacidica) (C, D). È evidente anche a livello qualitativo la differenza tra le fibrille formate [9].

templato per la conversione di PrP^C in PrP^{Sc} , in altre parole quando una PrP^C viene a contatto con una PrP^{Sc} la seconda induce un cambiamento conformazionale nella prima, trasformandola in PrP^{Sc} . La propagazione della transizione strutturale è fondamentale ai fini dello sviluppo della malattia.

PrP^{Sc} inoltre è resistente all'attacco delle proteasi (enzimi in grado di catalizzare la rottura del legame peptidico), mantenendo inalterata la sequenza degli aminoacidi dal residuo 89 fino al 230 (PrP^{Tr}). Questa resistenza alle proteasi è anche sintomo del fatto che il *core* della struttura fibrillare è fortemente impacchettato.

Non è tuttora chiaro come avvenga l'aggregazione di PrP^{Sc} e quale sia il meccanismo di replicazione. I modelli presenti in letteratura sono vari e tra loro discordi [11], ma in generale prevedono che la formazione di strutture fibrillari avvenga tramite l'interazione dei foglietti β dei monomeri di PrP^{Sc} .

La stabilità della struttura nativa di PrP è molto influenzata dal pH. Diversi studi [5] [16], hanno messo in evidenza come a pH acidi la proteina sia meno stabile e più prona all'aggregazione. Queste evidenze sperimentali sono in accordo con quanto ottenuto da simulazioni di dinamica molecolare [10].

Il ponte disolfuro che si forma nella struttura nativa tra le cisteine 178 e 213 è di fondamentale importanza per la stabilità della proteina prionica. Studi di dicroismo circolare [22] [21] [28] hanno mostrato come la sua assenza sia correlata alla perdita delle strutture secondarie e terziarie della conformazione nativa.

In letteratura sono presenti diversi studi volti a caratterizzare le fibrille di PrP (a titolo esemplificativo si vedano [9] e [31]). Si è osservato come al variare delle condizioni di formazione di queste fibrille vari anche la stabilità delle strutture ottenute (figura 1.2).

Capitolo 2

Spettroscopia di forza di singola molecola (SMFS)

2.1 Tecniche di singola molecola

A partire dagli anni '50 del secolo scorso, utilizzando la cristallografia a raggi X, è stato possibile caratterizzare la struttura delle molecole biologiche a livello atomico giungendo al riconoscimento della doppia elica del DNA. Successivamente, negli anni '80, grazie alla risonanza magnetica nucleare, si è iniziata a studiare la struttura di diverse proteine. Nel 2014 più di 100000 strutture di biomolecole sono state caricate su *Protein Data Bank*. Queste tecniche tradizionali (incluso il dicroismo circolare) permettono di caratterizzare il comportamento medio di un insieme di molecole; le proteine tuttavia sono oggetti dinamici, continuamente fluttuanti tra diversi stati conformazionali (di cui uno può essere favorito e maggiormente popolato), di conseguenza tecniche innovative in grado di studiare i singoli conformeri sono di fondamentale importanza.

L'invenzione del microscopio a forza atomica nel 1986 ha permesso la manipolazione delle singole molecole portando alla nascita di tecniche che hanno il vantaggio di permettere lo studio dei singoli conformeri.

Nel seguito ci si limita all'analisi delle tecniche di spettroscopia di forza di singola molecola (SMFS), in particolare agli esperimenti di trazione a velocità costante effettuati con l'ausilio del microscopio a forza atomica (AFM).

2.2 *Unfolding* meccanico

Negli esperimenti di SMFS, la molecola di interesse è manipolata in modo diretto, applicando una forza meccanica ai suoi estremi, inducendo una transizione strutturale, come schematizzato in figura 2.1. Informazioni strutturali, dinamiche e cinetiche possono essere ricavate dalla forza necessaria ad indurre la transizione.

Le grandezze fisiche misurabili da questi esperimenti sono due: la forza associata all'*unfolding* e la differenza tra la lunghezza del dominio strutturato e quella del dominio denaturato.

Le due tecniche maggiormente utilizzate per svolgere questo tipo di esperimenti sono quelle basate sull'uso dell'AFM e quelle basate sull'uso dell'*optical tweezer*, entrambe applicabili sia allo studio di DNA che a quello delle proteine.

2.2.1 Modello di Bell

La trattazione dell'*unfolding* dovuto all'applicazione di forze esterne risale alla formulazione di Bell del 1978 [3].

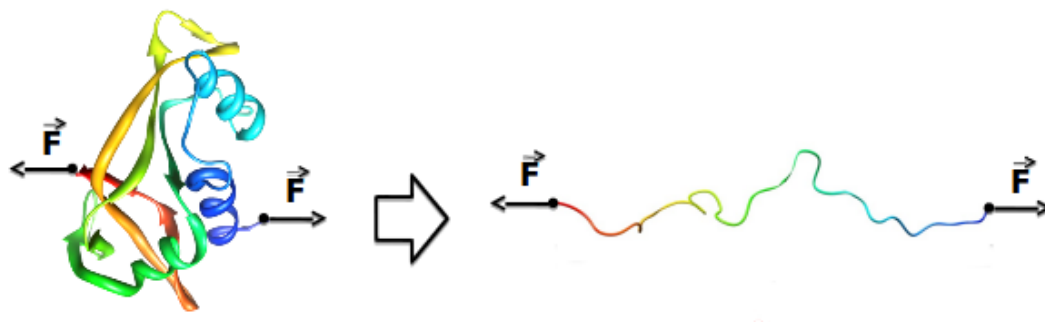


Figura 2.1: Schema di un esperimento di SMFS: una forza viene applicata agli estremi di una molecola inducendone l'*unfolding* meccanico.

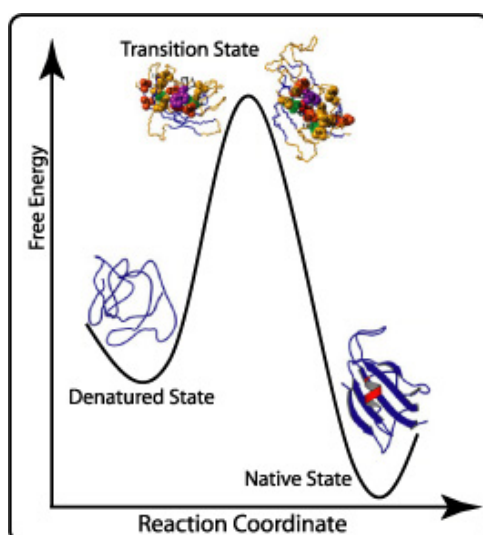


Figura 2.2: Schematizzazione del panorama energetico di una proteina a due stati. (Immagine tratta da <http://www-jackson.ch.cam.ac.uk/> il 14 giugno 2015)

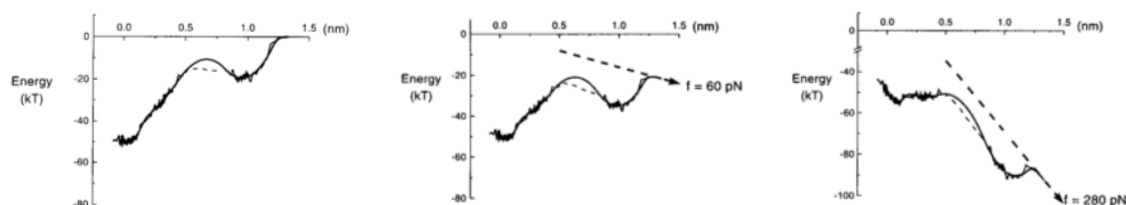


Figura 2.3: Esempio di inclinazione del panorama energetico dovuto all'applicazione di una forza esterna (risultato di simulazioni numeriche). All'aumentare della forza applicata aumenta l'inclinazione del panorama energetico, con conseguente variazione dei minimi dell'energia [13].

La formulazione può essere riassunta come segue: consideriamo una proteina a due stati: foldato o non foldato. Il suo panorama energetico sarà descrivibile tramite due buche di potenziale (figura 2.2), separate da una barriera $\Delta F(0)$ e la probabilità di transizione sarà proporzionale a $e^{-\frac{\Delta F(0)}{k_B T}}$.

Supponiamo, ora, di applicare una forza f agli estremi della proteina. L'applicazione di tale forza deve essere presa in considerazione nel calcolo dell'energia libera della proteina, sottraendo a quest'ultima il termine fx , dove x è la lunghezza di cui è stata estesa la proteina. Il *rate* di transizione dallo stato foldato a quello unfoldato potrà quindi essere riscritto come:

$$k_u(f) = k_0 e^{-\frac{\Delta F(0) - fx}{k_B T}} = k_u(0) e^{\frac{fx}{k_B T}}$$

Questa equazione mostra come il *rate* di transizione aumenti all'aumentare della forza, proprio perché essa riduce di fx la barriera di potenziale. L'*unfolding* proteico è un evento probabilistico associato all'attraversamento di una barriera di potenziale, che, in assenza di forza applicata, è troppo elevata per permettere un *unfolding* frequente della proteina. L'applicazione di una forza, riducendo la barriera di potenziale, aumenta la probabilità di transizione nel corso dell'esperimento.

Il risultato netto dell'applicazione di una forza costante è di fatto un'inclinazione del panorama energetico, come visibile in figura 2.3.

2.2.2 Distribuzione delle forze di *unfolding* in stiramenti delle proteine a velocità costante

In un esperimento in cui la proteina viene stirata a velocità costante la forza esercitata su di essa aumenta nel tempo, si definisce il *loading rate* come:

$$r_f = \frac{df}{dt}$$

La distribuzione delle forze di *unfolding* è piccata e il suo massimo può essere ricavato tramite la convoluzione di due fattori: da un lato la probabilità che un dominio rimanga strutturato diminuisce con il tempo (e quindi con la forza), dall'altro la probabilità che avvenga un evento di *unfolding* aumenta con la forza. Il risultato è che il valore corrispondente al massimo della distribuzione delle forze aumenta all'aumentare del *loading rate*.

La forma della distribuzione varia con il *loading rate*, in figura 2.4 sono riportate una serie di distribuzioni ottenute tramite simulazioni numeriche.

2.2.3 Modelli di fit

Le curve forza-allungamento ottenute con gli esperimenti di *unfolding* meccanico di proteine vengono interpretate ricorrendo alle teorie sviluppate per i polimeri. Nelle seguenti sezioni

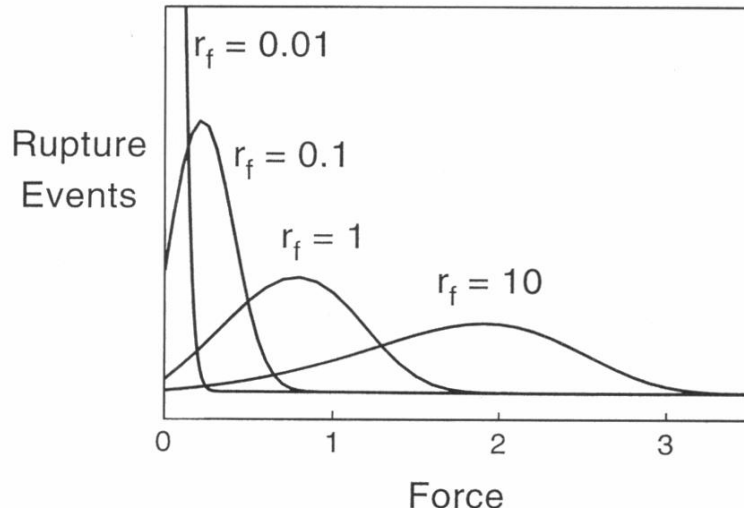


Figura 2.4: Distribuzioni delle forze di *unfolding* ottenute tramite simulazioni Monte Carlo, è evidente la variazione della forma della distribuzione e lo spostamento del picco massimo al variare del *loading rate* [13].

vengono presentati i modelli più comuni per la descrizione dei biopolimeri ed il comportamento meccanico (forza *vs* estensione) che essi prevedono. I tre modelli presentati differiscono tra loro per le assunzioni e sono presentati in ordine di complessità crescente [17].

Catena ideale

La catena ideale è un modello che non tiene in considerazione i cambiamenti in energia interna del polimero ed è utilizzato spesso per modellizzare polimeri flessibili il cui comportamento è dominato dagli effetti entropici. Supponiamo di avere n segmenti di lunghezza b connessi da $(n - 1)$ giunzioni libere di ruotare. Orientiamo i segmenti ed associamo ad ognuno di essi un vettore. Sia \mathbf{R} il vettore congiungente inizio e fine della catena (dato dalla somma dei vettori componenti). La lunghezza della catena nel caso in cui essa sia completamente distesa è chiamata lunghezza al contorno (*contour length*) e d'ora in avanti sarà indicata come L_c . Esiste una sola configurazione nella quale $R = nb = L_c$ (ovvero la catena è completamente stesa) ed essa risulta di conseguenza entropicamente sfavorita, mentre risultano favorite le situazioni in cui tale distanza è minore. La probabilità che la distanza tra i due estremi di una catena composta da n segmenti sia \mathbf{R} può essere ottenuta tramite il metodo del cammino casuale, ottenendo una distribuzione approssimabile con la seguente gaussiana tridimensionale:

$$p_{3d}(n, \mathbf{R}) = \left(\frac{3}{2\pi nb^2}\right)^{\frac{3}{2}} e^{-\frac{3\mathbf{R}^2}{2nb^2}}$$

da tale distribuzione di probabilità si ottiene l'energia libera e quindi la forza come gradiente di quest'ultima, giungendo a:

$$\mathbf{F} = \frac{3k_B T}{nb^2} \mathbf{R}$$

Da queste relazioni si possono trarre due interessanti conclusioni:

- In ogni caso in cui sia la temperatura che R sono diversi da zero, anche la forza non è nulla. Ciò si verifica perché lo stato entropicamente favorito è quello con $R = 0$ (corrispondente al maggior numero di conformazioni possibili), di conseguenza è sempre necessaria una forza per allontanarsi da tale stato.

- La forza dipende linearmente da R , di conseguenza ci si può riferire a questi polimeri come a molle entropiche, in cui la resistenza alla deformazione è completamente derivabile dall'entropia e dipende dalla temperatura. La costante elastica di questa molla è direttamente proporzionale alla temperatura contrariamente a quanto accade per molti materiali di uso comune, ad esempio gli acciai.

Nel caso di polimeri reali esistono interazioni tra atomi vicini (per esempio di tipo elettromagnetico) che si possono manifestare in variazioni della distanza o nella creazione di specifici angoli nella struttura, è quindi lecito chiedersi se questo modello sia effettivamente valido. La risposta risulta affermativa, purchè risulti $R \ll L_c$ e il polimero sia sufficientemente lungo, ovvero purchè la sua lunghezza totale sia maggiore di una lunghezza chiamata lunghezza di persistenza (*persistence length*) e d'ora in poi indicata con l_p . Questa lunghezza rappresenta la lunghezza minima tra due punti di un polimero sottoposto a oscillazioni termiche a partire dalla quale sparisce la correlazione tra le loro posizioni (e quindi gli effetti entropici risultano dominanti). Un polimero sufficientemente lungo può essere pensato come formato da tanti pezzi di lunghezza l_p , ognuno fluttuante in modo indipendente dagli altri, e quindi l'approssimazione di catena ideale risulta valida (è possibile dimostrare che vale la relazione $b = 2l_p$).

Questo modello presenta alcune limitazioni, la principale è che la costante elastica della molla entropica non dipende dall'allungamento e rimane invariata anche per allungamenti superiori a L_c , portando a risultati privi di significato fisico.

Freely Jointed Chain

Il modello della *Freely Jointed Chain* permette di occuparsi anche di polimeri con allungamenti dell'ordine della L_c , ponendo una limitazione alla lunghezza massima di \mathbf{R} e quindi risolvendo il problema principale del modello della catena ideale.

La modellizzazione è molto simile a quella della catena ideale, ovvero si rappresenta il polimero come successione di segmenti connessi da giunzioni libere di ruotare. L'approccio risolutivo tuttavia è diverso: mentre nella catena ideale si considera la variazione di entropia associata all'allungamento, in questo caso si eseguono i conti nell'*ensemble* canonico considerando un bagno termico a temperatura costante in cui è immerso il polimero di interesse. Si definisce un sistema ad energia potenziale nota in modo da ricavare, tramite la distribuzione di Boltzmann, la probabilità di ottenere una determinata conformazione all'equilibrio.

Si considera, dunque, il polimero immerso in un bagno termico, con la posizione di uno dei due estremi fissata e con un peso legato all'altro estremo. La presenza di questo peso comporta un allungamento della catena, con conseguente variazione di energia potenziale del sistema (in quanto essa dipende dalla posizione del peso). Con la distribuzione di Boltzmann è possibile calcolare la probabilità di ogni microstato e quindi risalire all'estensione media del polimero:

$$\langle R \rangle = bn(\coth(k) - \frac{1}{k})$$

$$k = \frac{2Fl_p}{k_B T}$$

Osservando i due casi limite per le soluzioni di questa equazione implicita si ottiene che:

- qualunque sia la forza applicata l'allungamento non può mai essere superiore alla lunghezza al contorno nb ;
- per allungamenti (e forze) piccoli questo modello predice lo stesso andamento forza *vs.* distanza della catena ideale.

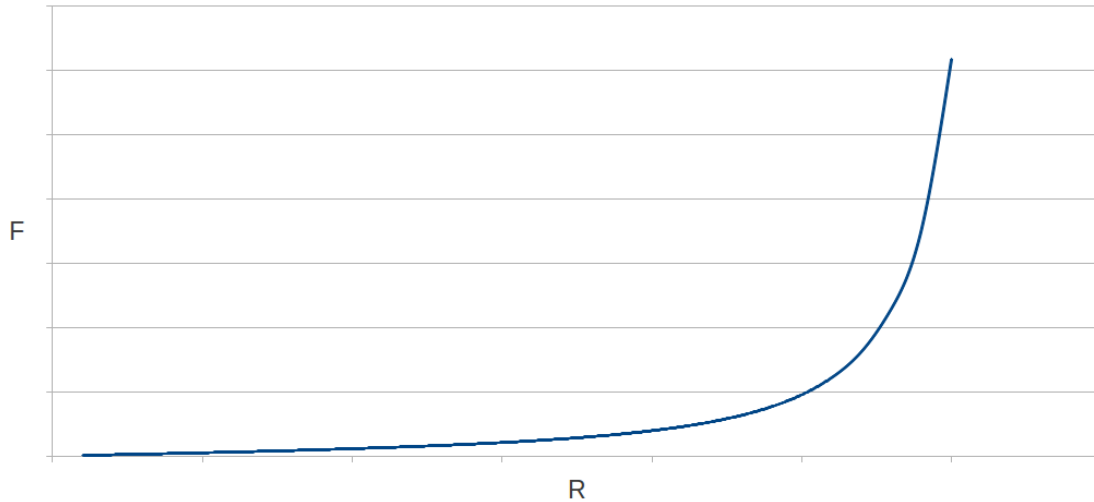


Figura 2.5: Andamento forza *vs.* allungamento nel modello WLC. All'aumento della forza la lunghezza tende asintoticamente a L_c , mentre per $R \ll L_c$ si ritrova l'andamento lineare previsto dal modello della catena ideale.

Worm-Like Chain

Entrambi i modelli precedenti prendono in considerazione solo ed esclusivamente fattori entropici, ignorando il costo energetico dei cambi di orientazione tra segmenti. Il seguente modello, invece, tiene conto di entrambi gli effetti e passa da una modellizzazione discretizzata tramite segmenti della catena polimerica ad una descrizione continua.

Per semplicità il modello assume che la WLC sia inestensibile, ovvero L_c non possa variare; la catena resiste alle deformazioni tramite aumento dell'energia potenziale assunto paragonabile a quello di una trave elastica (per cui l'energia potenziale risulta minima nel caso della proteina completamente estesa e ad una flessione della proteina corrisponde un aumento dell'energia potenziale). La rigidità di questa trave è legata alla l_p .

La relazione forza-allungamento viene ottenuta, analogamente al caso della FJC, assumendo di porre il polimero in un bagno termico, con un estremo fissato e l'altro legato ad un peso, ottenendo però degli integrali sullo spazio delle configurazioni non risolvibili analiticamente. Utilizzando metodi computazionali si giunge al seguente andamento della forza in funzione dello spostamento medio:

$$F = \frac{k_B T}{l_p} \left(\frac{1}{4} \left(1 - \frac{\langle R \rangle}{L_c} \right)^{-2} - \frac{1}{4} + \frac{\langle R \rangle}{L_c} \right)$$

In figura 2.5 è riportato l'andamento della forza in funzione dell'allungamento previsto dal modello WLC. All'aumento della forza, la lunghezza tende asintoticamente a L_c , mentre per $R \ll L$ si ritrova l'andamento lineare previsto dal modello della catena ideale.

Il modello WLC prevede che una forza maggiore sia richiesta per allungare un polimero fino ad una determinata lunghezza rispetto al modello FJC. L'origine di questo fenomeno può essere compresa considerando le fluttuazioni termiche nel polimero come onde di diversa lunghezza d'onda sovrapposte: nel caso di una FJC la lunghezza d'onda minima è b (in quanto non si possono avere onde di lunghezza inferiore alla lunghezza del *link*), viceversa, questo vincolo non è presente per la WLC. Una forza aggiuntiva è, dunque, richiesta per eliminare queste fluttuazioni di piccola lunghezza d'onda dalla WLC.

In figura 2.6 è riportata la sovrapposizione di dati sperimentali e fit con WLC e FJC per l'estensione in funzione della forza per il DNA (anch'esso modellizzabile come catena

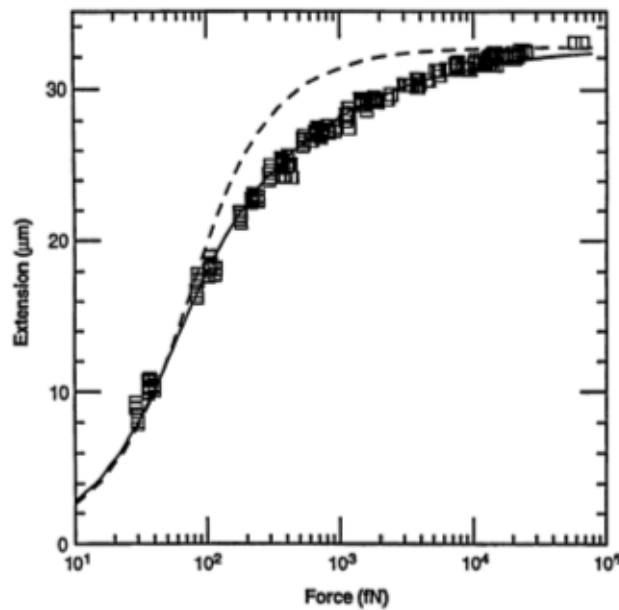


Figura 2.6: Sovrapposizione di dati (quadrantini) ottenuti dallo stiramento di DNA tramite optical tweezer e dei relativi fit tramite WLC (linea continua) e FJC (linea tratteggiata). I fit sono stati effettuati assumendo $l_p=106$ nm e $L_c=33$ μm . Per alti valori di forza l'estensione raggiunge il valore asintotico della L_c . Il fit è migliore con il modello WLC [4].

polimerica), è visibile il fatto che il fit secondo il modello WLC sia migliore rispetto al modello FJC. Inoltre si nota quanto descritto precedentemente, ovvero che a parità di allungamento la forza richiesta da una WLC è maggiore di quella della FJC.

2.3 Tecnica sperimentale: SMFS con l'utilizzo di un AFM

2.3.1 Microscopio a forza atomica

Il microscopio a forza atomica è uno strumento opto-meccanico, sviluppato per la microscopia. La sua componente principale è il *cantilever* che viene spostato tramite elementi piezo-elettrici (si veda la figura 2.7). Le piccole deflessioni del *cantilever* vengono misurate amplificandole con un artificio geometrico: un laser focalizzato sul *cantilever* stesso viene riflesso su un fotodiodo posto ad una certa distanza dalla punta (maggiore sarà la distanza e maggiore risulterà l'amplificazione). Ogni deflessione del *cantilever* produrrà quindi uno spostamento del punto luminoso sul fotodiodo, misurando tale spostamento sarà possibile risalire alla deflessione del *cantilever* stesso.

Meccanicamente il *cantilever* è descrivibile come una trave con un estremo fisso. L'equazione che determina lo spostamento trasverso di una trave di lunghezza L con estremo fisso e forza applicata sull'altro estremo, perpendicolarmente alla trave stessa, è:

$$w = -\frac{F}{EI} \left(\frac{x^3}{6} - \frac{Lx^2}{2} \right) \quad (2.1)$$

Da quest'ultima si ottiene che la deflessione della punta risulta proporzionale alla forza applicata trasversalmente all'estremo libero e dipende dalle sue proprietà meccaniche. (w è lo spostamento trasverso, E il modulo di Young del *cantilever*, I il suo momento d'inerzia, F la forza applicata e x la distanza del punto del *cantilever* considerato dall'estremo fisso di quest'ultimo).

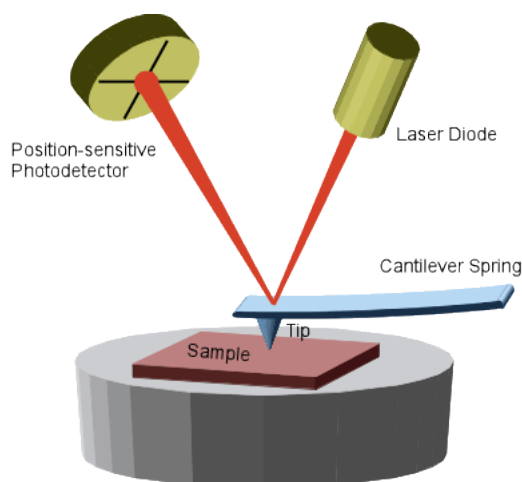


Figura 2.7: Schema di funzionamento di un AFM: un fascio laser viene riflesso dal *cantilever*, dal suo punto di incidenza sul fotodiodo viene ricavata la deflessione di quest'ultimo. Figura scaricata da <https://undergrad.research.ucsb.edu/2014/03/atomic-force-microscope-tiny-record-player/> il 9 giugno 2015.

2.3.2 Spettroscopia di forza

L'AFM, vista la sua natura meccanica, può essere utilizzato anche come strumento per effettuare la spettroscopia di forza, tecnica che permette di studiare le proprietà meccaniche di campioni organici o inorganici su scala nanometrica. Per far questo, il movimento relativo tra *cantilever* e substrato, una volta effettuato il posizionamento del campione, avviene solo in direzione z (mentre nella microscopia i posizionatori piezoelettrici muovono il *cantilever* orizzontalmente).

Sono due gli aspetti principali da considerare, meccanicamente collegati allo spostamento verticale del *cantilever*:

- il *cantilever* viene abbassato fino al momento in cui inizia la deflessione. È possibile misurare la deflessione del *cantilever* conseguente ad un prefissato spostamento verticale del piezo o viceversa l'ampiezza dello spostamento verticale conseguente all'applicazione di una forza nota (controllata tramite la deflessione del *cantilever* stesso). Questo processo, noto anche come indentazione, consente di misurare le proprietà meccaniche locali di un materiale.
- il *cantilever* viene ritratto dopo un determinato tempo di contatto e viene misurata la forza che è necessario applicare per la rottura dei legami che si sono formati tra *cantilever* e campione.

Questo secondo aspetto è sfruttato negli esperimenti di spettroscopia di forza di singola molecola: la molecola è stirata tra la punta del *cantilever* e un substrato piatto (un vetrino portacampione). Uno strato della proteina di interesse è adsorbito sul substrato. Quando la punta e il substrato vengono avvicinati una o più molecole possono venir adsorbite sulla superficie della punta. Durante l'allontanamento della punta dal substrato, all'aumento della distanza relativa corrisponde un allungamento della molecola, con generazione di una forza di richiamo che provoca il ripiegamento del *cantilever*. Tale ripiegamento è misurato tramite lo spostamento del laser sul fotorivelatore. Il sistema sperimentale con il tipo di curva forza-estensione ottenuto è riportato in figura 2.8, mentre lo schema del ciclo di avvicinamento e allontanamento è riportato in figura 2.9.

La curva forza-estensione che si ottiene stirando una proteina multimodulare a velocità costante ha una caratteristica forma a dente di sega, nella quale ogni picco corrisponde ad un evento di *unfolding*. L'andamento ottenuto è ben fittabile con il modello WLC (sezione 2.2.3).

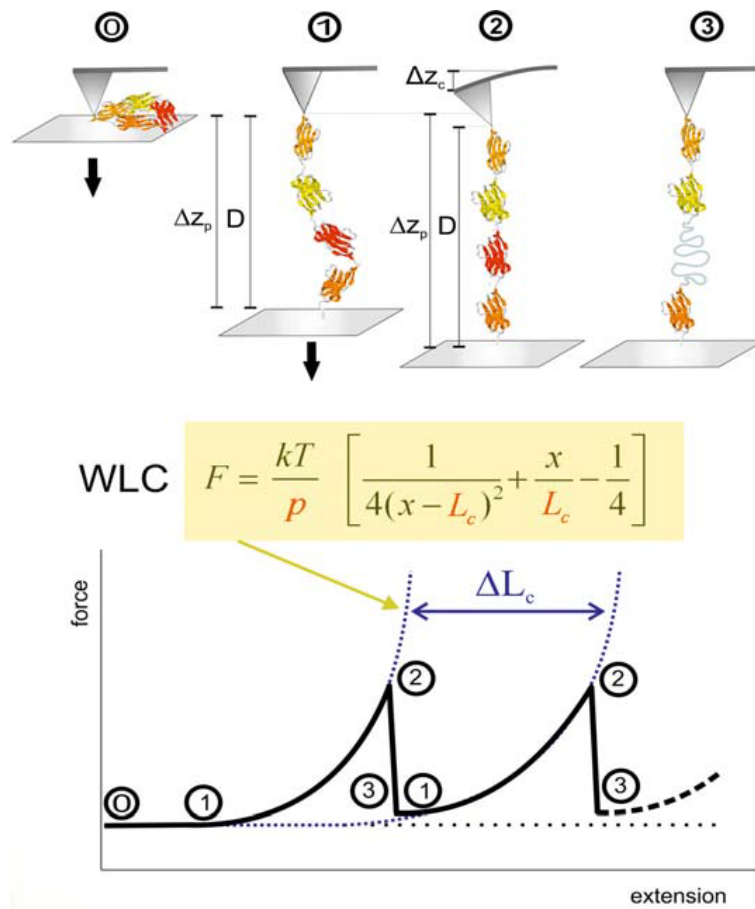


Figura 2.8: Rappresentazione schematica degli esperimenti di SMFS. Durante il tempo di contatto tra substrato e punta la proteina può aderire alla punta (0), nella fase di allontanamento essa esercita una forza di richiamo (1), che aumenta fino al raggiungimento dell'estensione massima (2); in seguito ad un evento di *unfolding* (3) la forza decresce rapidamente ed è possibile allungare ulteriormente la proteina riprendendo il ciclo. I picchi ottenuti sono ben fittabili con il modello WLC [8].

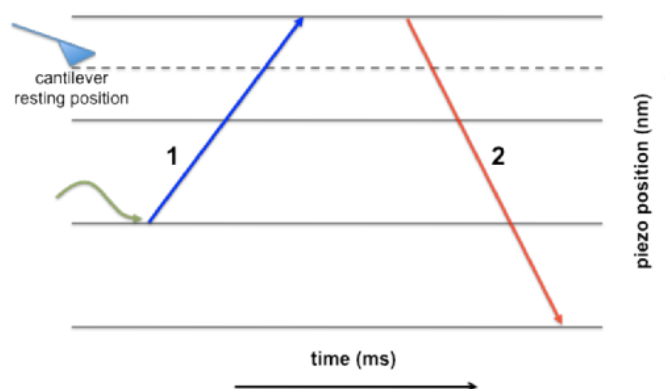


Figura 2.9: Rappresentazione schematica del ciclo di avvicinamento ed allontanamento relativi del vetrino portacampione dal *cantilever*: quest'ultimo rimane fermo nella sua posizione originale, mentre, grazie all'azione del posizionatore piezoelettrico, si ha uno spostamento del campione, che inizialmente si trova distanziato al di sotto del *cantilever*, in seguito sale (1) ed entra in contatto con quest'ultimo, rimanendo in tale posizione per un certo intervallo di tempo. Infine viene ritratto (2) e viene registrata la curva di allungamento [25].

I *cantilever* commerciali utilizzati per esperimenti di spettroscopia di forza hanno valori di costante elastica dell'ordine di 50-100 pN/nm, che, per temperature prossime alla temperatura ambiente, producono oscillazioni termiche dell'ordine dei 15-25 pN. Questo sistema permette, dunque, manipolazioni spaziali dell'ordine del nanometro e misurazione di forze di poche decine di piconewton.

2.3.3 Esperimenti con il metodo del doppio impulso

Nel corso degli esperimenti di SMFS sopra descritti è generalmente riconoscibile il segnale relativo ad interazione non specifica nella prima parte della curva di ritrazione, dovuto all'interazione della punta con il substrato e con le molecole su di esso presenti.

Per ottenere curve più pulite, in cui tale interazione non è presente, sono stati sviluppati metodi sperimentali che prevedono cicli più complessi di avvicinamenti e allontanamenti tra substrato e punta [7].

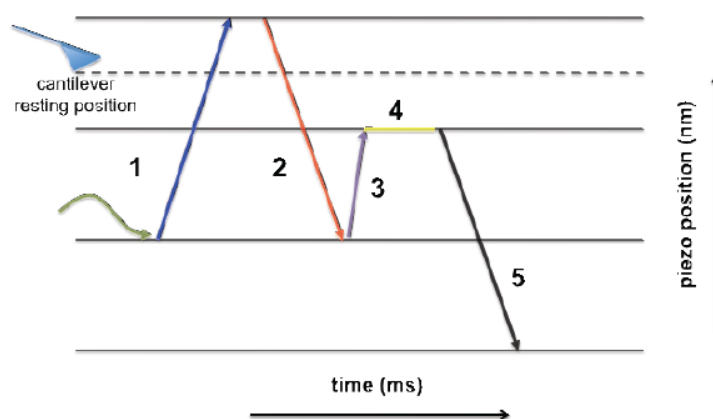


Figura 2.10: Schema del ciclo di avvicinamento e allontanamento del substrato dal *cantilever* nel corso di un esperimento con doppia pulsazione: 1. avvicinamento 2. prima ritrazione, l'allontanamento viene fermato prima di ottenere il distacco della proteina 3. riavvicinamento parziale (senza arrivare al contatto con la superficie) 4. tempo in cui la distanza rimane costante per permettere il *refolding* della proteina 5. seconda ritrazione (questa volta completa di distacco della proteina) [25].

In figura 2.10 è schematizzato il più semplice di questi cicli, ovvero quello costituito da due pulsazioni. Come nel caso del semplice esperimento di *unfolding* la punta e il substrato sono portati in contatto e, successivamente, vengono allontanati solo di una piccola lunghezza, compatibile con l'*unfolding* di alcuni moduli proteici, senza staccare completamente la molecola. Punta e substrato vengono quindi riavvicinati, senza arrivare al contatto. Nel tempo in cui punta e substrato sono vicini ma non in contatto, alcuni moduli proteici hanno il tempo di ripiegarsi nuovamente; in seguito punta e substrato vengono allontanati in modo definitivo fino al completo distacco della molecola. Nella seconda curva di allontanamento registrata è possibile vedere l'*unfolding* dei moduli proteici che non si erano sfoldati nel corso del primo allontanamento e l'*unfolding* dei moduli che si sono ripiegati nel tempo in cui la punta e il substrato sono stati vicini, ma non in contatto; inoltre la seconda curva risulta priva di interazione non specifica, per cui tutti i picchi misurati possono essere associati alla proteina di interesse. Una curva esemplificativa è riportata in figura 2.11.

Un esperimento simile può essere effettuato ripetendo molti cicli di stiramenti e allontanamenti di *cantilever* e substrato prima di effettuare l'allontanamento definitivo (con il distacco della proteina).

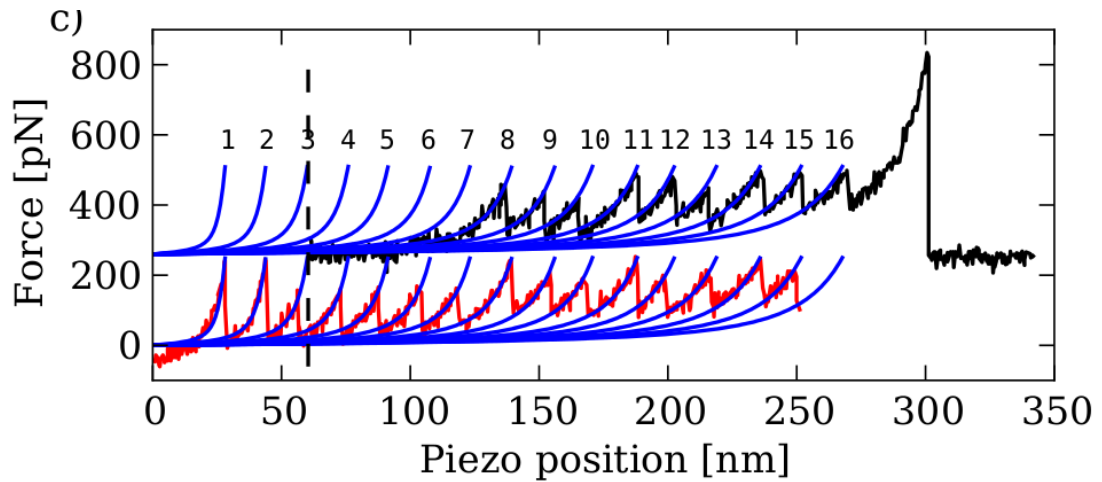


Figura 2.11: Curva esemplificativa di esperimento con doppio impulso: la curva in basso corrisponde al primo stiramento, prima di giungere al distacco si è effettuato il *refolding* (avvicinando substrato e *cantilever* fino al punto indicato dalla linea tratteggiata), la curva in alto corrisponde al secondo stiramento, con distacco finale. Solo una parte dei moduli proteici sfoldati nel primo allontanamento si è rifoldato ed è quindi nuovamente visibile nel secondo stiramento [1].

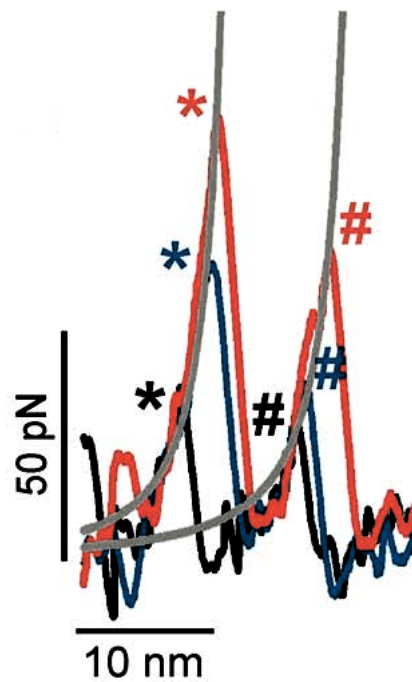


Figura 2.12: Sovrapposizione delle curve ottenute per tre stiramenti della fibronectina (verde, rosso e nero) [20]. Le tre curve sono ben sovrapponibili, le forze di *unfolding* sono diverse, ma il fit tramite WLC restituisce un uguale valore di Δ_{Lc}

Si può studiare la sovrapposizione tra i segnali ottenuti dai differenti stiramenti 2.12, il fit delle WLC si dovrebbe sovrapporre, ma l'altezza del picco non deve essere necessariamente la stessa. Infatti, essendo l'*unfolding* un processo stocastico, non è possibile determinare a priori quale sarà la forza necessaria per la rottura. Un identico fit dei picchi permetterà di ottenere identiche misure per le lunghezze dei domini denaturati (tale effetto è ben visibile anche in figura 2.11).

2.3.4 La tecnica dei costrutti multimodulari

La concentrazione della proteina depositata sul substrato è valutata in modo da essere sufficientemente bassa da ridurre la probabilità che più proteine si attacchino contemporaneamente alla punta, dando luogo a curve non facilmente interpretabili. L'avvicinamento del *cantilever* alla punta avviene in modo casuale, senza sapere in quali punti del substrato la proteina sia effettivamente presente. Il rapporto tra il numero di curve contenenti un segnale valido e quelle prive di segnali è molto basso ($< 0.1\%$) e diviene, quindi, di fondamentale importanza la definizione di un metodo di discriminazione tra curve contenenti eventi di interesse e curve prive di tali eventi.

Il metodo generalmente utilizzato consiste nel sintetizzare costrutti proteici multimodulari, in cui la proteina di interesse viene circondata da proteine note e che restituiscono segnali di *unfolding* facilmente riconoscibili.

Ad esempio si può costruire un costrutto in cui si dispongono in serie quattro proteine note identiche, la proteina di interesse e altre quattro proteine note identiche. Una curva contiene sicuramente l'*unfolding* della proteina di interesse se contiene almeno cinque eventi relativi alle proteine note.

In questo modo è possibile selezionare tra tutte le curve raccolte nel corso dell'esperimento quelle che effettivamente sono utili alla descrizione della proteina che si vuole studiare. Inoltre la presenza di queste proteine note aumenta la lunghezza del costrutto proteico facilitando il processo di adesione e stiramento del costrutto stesso.

Fra le proteine utilizzate per questo scopo vi sono GB1 [30], I27 [20], ubiquitina [15] e IG8 [12].

2.3.5 Energia dissipata nella transizione

Durante un esperimento di SMFS, dopo aver registrato il segnale relativo all'*unfolding* di un modulo proteico è possibile invertire la direzione del moto relativo ed ottenere la curva di *refolding* (due esempi di tali curve sono riportati in figura 2.13). Si è osservato che alcune proteine hanno un comportamento elastico e le curve di *unfolding* e *refolding* sono sovrapposte [27]; altre volte *folding* e *refolding* seguono due diversi cammini energetici, ovvero sono processi che avvengono fuori equilibrio. Quest'ultimo è il comportamento ottenuto per proteine multimodulari, in cui ad una curva di *unfolding* a dente di sega corrisponde una curva di *refolding* priva di strutture riconoscibili. Un'evidente isteresi è presente tra le due curve, la cui area è associabile al calore dissipato nel processo di *unfolding*.

La transizione all'equilibrio si verifica per proteine con struttura semplice. Nelle proteine con struttura complessa si formano legami tra residui che sono molto distanti nella struttura primaria e il tempo necessario per provare tutte le possibili combinazioni (variando i residui che entrano in contatto) e trovare quella ad energia minima diventa maggiore: il processo avviene fuori equilibrio tutte le volte in cui la velocità di stiramento è talmente rapida che il tempo in cui avviene lo stiramento è minore di questo tempo di rilassamento e il sistema, di conseguenza, non riesce a testare tutte le possibili conformazioni durante il processo di *unfolding* meccanico. Si può affermare, dunque, che è necessario ridurre la velocità di stiramento per effettuare le misurazioni in condizioni di quasi-equilibrio. Nel caso di transizioni fuori equilibrio la velocità elevata di stiramento non permette di distinguere la rottura dei

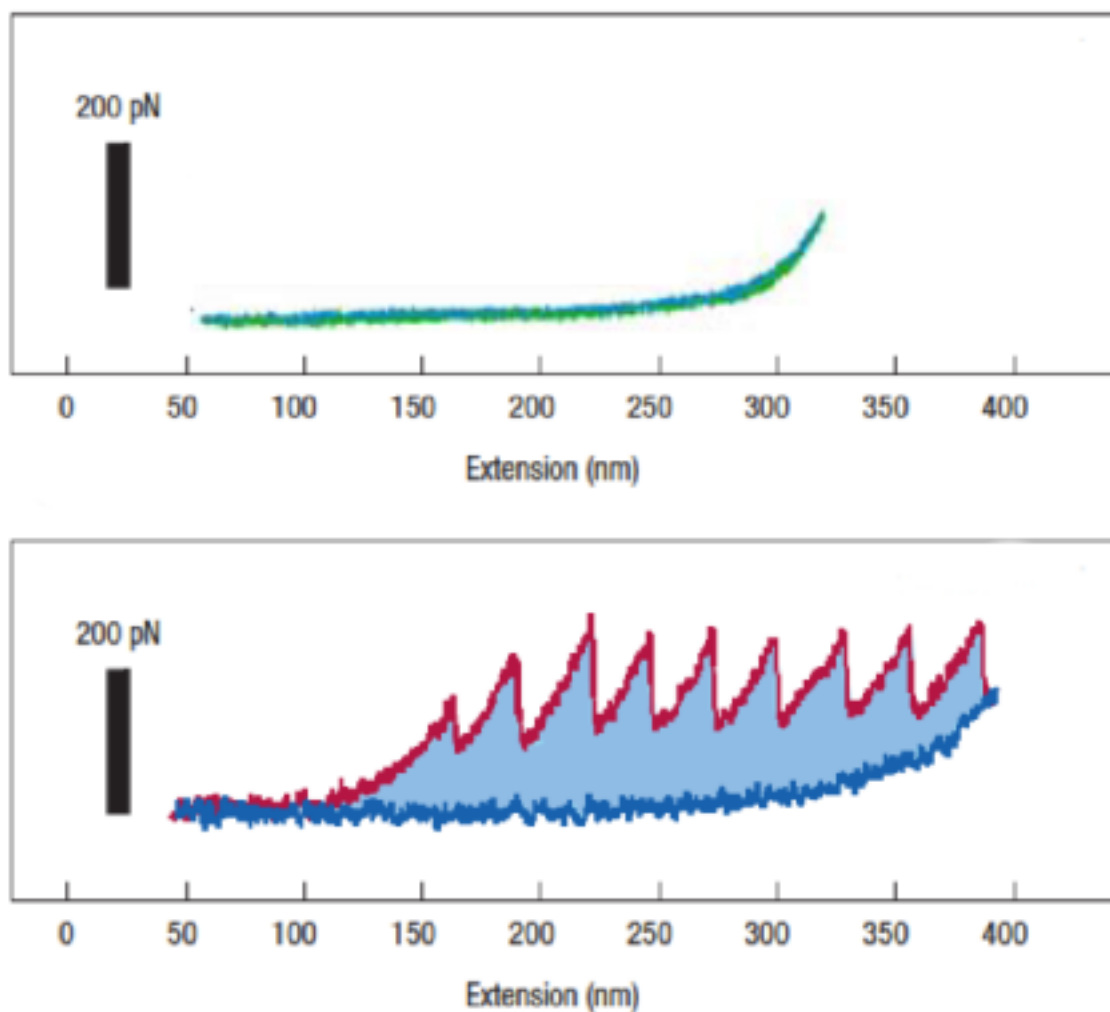


Figura 2.13: Due esempi di comportamento di proteine sottoposte a stiramento. In alto: comportamento elastico dell'elastina, la curva di *unfolding* (verde) e quella di *refolding* sono perfettamente sovrapponibili.[8] In basso: tipico andamento a dente di sega ottenuto per un costrutto contenente ripetizioni di I27: in rosso la tipica curva di *unfolding*, in blu la relativa curva di *refolding*; è evidente l'isteresi tra le due curve (area azzurra) [29].

legami della struttura terziaria seguita del collasso della struttura secondaria: l'*unfolding* viene dunque detto cooperativo ed è descrivibile come la rottura di un singolo legame, la cui barriera energetica è quella che separa lo stato nativo da quello denaturato.

Capitolo 3

Lo scopo del progetto

L'eterogeneità delle malattie prioniche esistenti può essere associata alla eterogeneità strutturale del monomero di *PrP*; queste diverse conformazioni monomeriche possono dare origine a differenti stati aggregativi e a diverse strutture fibrillari, con conseguente sviluppo di patologie caratterizzate da sintomi differenti. Lo studio degli equilibri conformazionali di *PrP* è, dunque, di fondamentale importanza. Nonostante un'ampia letteratura abbia recentemente portato a delineare gli aspetti strutturali di fibrille ed oligomeri, non sono ancora noti i meccanismi all'origine della loro formazione, della loro struttura e degli *step* aggregativi iniziali.

Le tecniche di singola molecola possono permettere l'analisi dei singoli conformeri e risultano quindi particolarmente promettenti in questa prospettiva.

Gli scopi di questo progetto di ricerca sono:

- lo studio, tramite spettroscopia di forza basata su AFM, degli equilibri conformazionali di MoPrPTr, per analizzarne il polimorfismo strutturale;
- la caratterizzazione degli *step* aggregativi iniziali, studiando l'effetto di diverse orientazioni relative dei monomeri e della loro diversa concentrazione;
- l'analisi dell'effetto del pH sulla stabilità delle differenti conformazioni.

Alcuni passaggi intermedi sono necessari per il raggiungimento di tale obiettivo:

- caratterizzare il costrutto proteico utilizzato come riferimento nel corso degli esperimenti: solo una descrizione accurata del suo comportamento può permettere di distinguere gli eventi associati a MoPrP;
- sviluppare un metodo efficace per l'analisi statistica dei risultati ottenuti, che permetta di riconoscere differenze significative ed effettuare un confronto quantitativo tra sistema di riferimento e costrutto di interesse.

Parte II

Materiali e metodi

Capitolo 4

Il campione biologico

4.1 I costrutti utilizzati

Si è scelto di utilizzare come sistema di riferimento la proteina GB1 (dominio B1 della proteina G) di streptococco, in quanto è una proteina stabile, il suo utilizzo come sistema di riferimento è frequente in letteratura [30] e il suo *unfolding* è descritto come un processo a due stati [6], [7]. Il vantaggio di una transizione di *unfolding* a due stati è che verrà generato un solo tipo di segnale (in termini di lunghezza e forza) dalla proteina medesima, consentendo di osservare fenomeni derivanti da PrP al di fuori di una sola distribuzione di eventi. Al fine di caratterizzarne il comportamento si sono effettuati innanzitutto esperimenti con il costrutto $GB1_8$.

Si è scelto di studiare il comportamento di MoPrP nella forma troncata (a.a. 89-230) in quanto è l'unità minima in grado di generare fibrille.

In figura 4.1 è riportato uno schema dei costrutti utilizzati. In generale la molecola o le molecole di PrP sono affiancate ad ogni estremità da quattro moduli di GB1; solo nel caso del tetramero il numero di GB1 che fiancheggiano i monomeri di PrP è due.

Il dimerico è stato realizzato con due geometrie: legando il C-terminale di un monomero con l'N-terminale dell'altro (*Head-to-Tail* o H-T) oppure legando tra loro i due C-terminali (*Head-to-Head* o H-H). Nel caso di costrutti multimerici con tre e quattro monomeri di PrP, l'N-terminale del primo monomero è stato legato al C-terminale del successivo (orientazione H-T).

4.2 La sintesi dei costrutti

I costrutti proteici sono stati sintetizzati nel laboratorio del Prof. Legname della SISSA di Trieste tramite la tecnica del DNA ricombinante, esprimendoli in *E. Coli* e purificati [25].

4.3 *Refolding* in condizioni riducenti

Il monomero è stato utilizzato negli esperimenti sia nella sua versione nativa, che con il ponte disolfuro ridotto.

La preparazione del monomero in condizioni riducenti è stata effettuata ponendo $GB1_4PrPTrGB1_4$ in condizioni denaturanti, in seguito esso è stato rifoldato in presenza di agenti riducenti (DTT 1 mM). Prima di effettuare l'esperimento il campione è stato dializzato per rimuovere il denaturante.

Questa procedura non influisce sulla struttura di GB1 in quanto la proteina non possiede alcun ponte disolfuro.

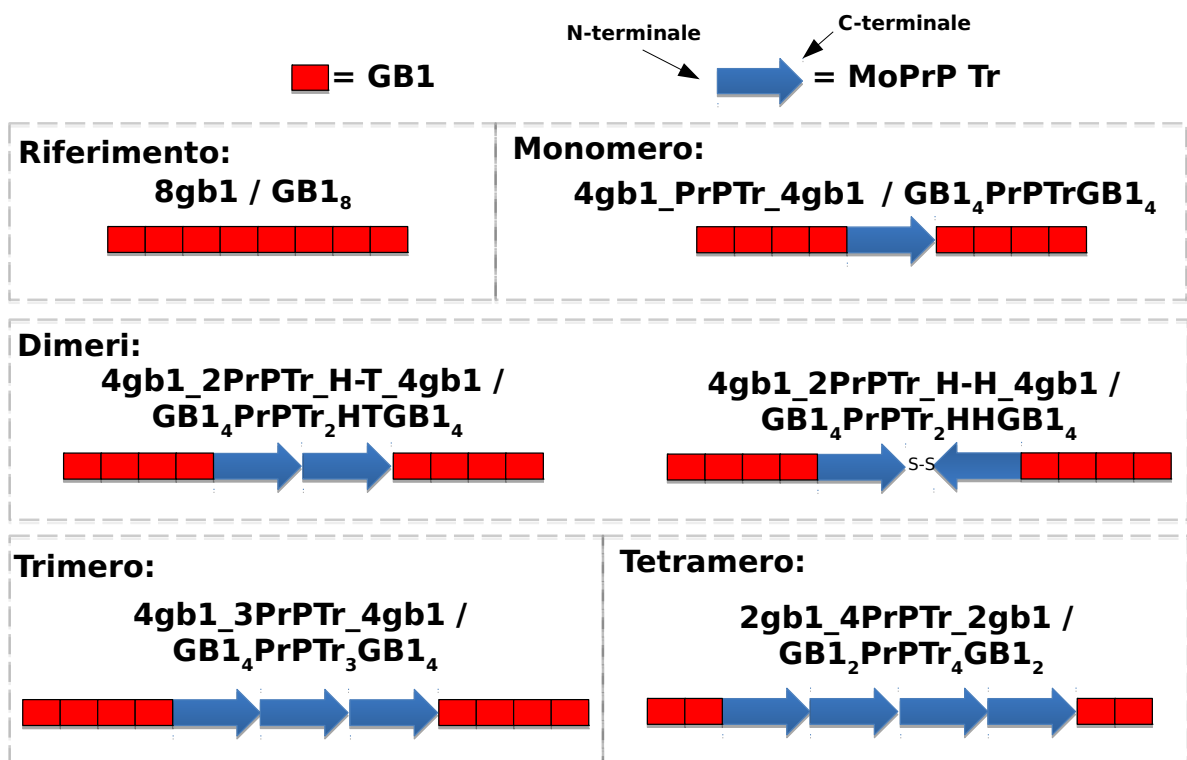


Figura 4.1: Schema riassuntivo dei costrutti utilizzati, sono riportate le due nomenclature utilizzate nel corso della tesi.

Capitolo 5

SMFS

5.1 Esperimenti di *unfolding*

Gli esperimenti di *unfolding* meccanico a velocità costante sono stati eseguiti con un AFM Veeco Picoforce (Multimode Nanoscope IIIa, Bruker).

Si sono utilizzati *cantilever* triangolari di nitruro di silicio ricorperiti d'oro su entrambi i lati (NPG, Bruker) con costante elastica nominale di 0.06 N/m. La costante elastica effettiva è stata calcolata tramite il metodo del rumore termico, basato sul teorema di equipartizione dell'energia [14].

Tutti gli esperimenti sono stati eseguiti a temperatura ambiente (si è stimata una temperatura interna alla cella di circa 28°C).

Le proteine vengono depositate per almeno 20 minuti su un vetrino porta-campione pulito sul bunsen e asciugato con azoto, in seguito tale vetrino viene inserito in una cella per fluidi.

Le soluzioni tampone utilizzate sono: Tris 20 mM pH 7.4, NaOAc 20 mM pH 5.5 e NaOAc 20 mM pH 4.0. Nel caso di esperimenti a pH acido si è inizialmente montato il sistema con il tampone a pH 7.4 e dopo circa un'ora di esecuzione delle misure si è effettuato il cambio *buffer* sfruttando i due canali di accesso alla cella (si veda, per chiarezza, figura 5.1). In questo modo è possibile avere, per ogni esperimento, un controllo sulla bontà del campione, avendo il confronto con i dati a pH 7.4.

Gli esperimenti sono stati eseguiti tramite cicli di avvicinamento e allontanamento del *cantilever* dalla superficie, come descritto nella sezione 2.3, impostando i seguenti parametri:

- distanza fra posizione iniziale del *cantilever* dal punto di contatto: 500 nm
- spinta sulla superficie corrispondente ad una deflessione di 10 nm (leggibile dalla curva forza-distanza)
- velocità di allontanamento della punta dal substrato: 2180 nm/s
- tempo di contatto tra punta e superficie: variabile da esperimento a esperimento, ma compreso tra 0 e 200 ms

Nel corso degli esperimenti la posizione del *cantilever* nel piano parallelo al substrato viene periodicamente modificata, sia in modo automatico, che manualmente, in modo da modificare il punto di contatto e quindi essere certi di pescare proteine diverse.

Generalmente in un singolo esperimento si raccoglie un numero di curve compreso tra 30000 e 150000, ma il filtro di queste curve in base all'effettiva presenza dei segnali di GB1 necessari fa sì che di queste ne vengano considerate valide meno dello 0.1 %.

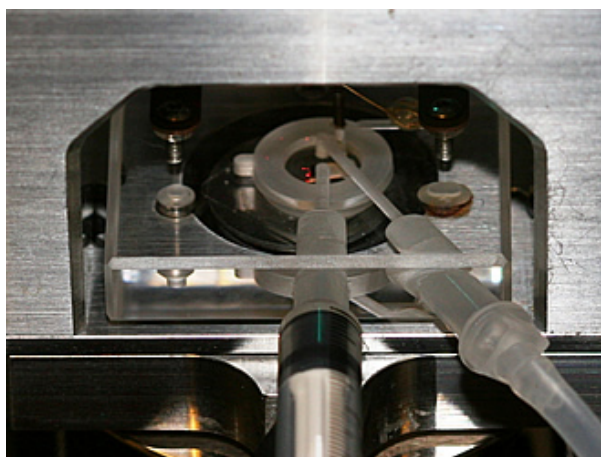


Figura 5.1: Fotografia del sistema sperimentale (scaricata da <https://wiki.ch.ic.ac.uk> il 26/06/2015). Sono visibili i due canali di accesso alla cella, inserendo in uno la siringa e nell'altro un tubicino per lo scarico è possibile effettuare il cambio *buffer* senza aprire la cella e mantenendo costante il volume di liquido in essa contenuto.

5.2 Esperimenti con il metodo del doppio impulso

Gli esperimenti sono stati eseguiti tramite cicli di avvicinamento e allontanamento del *cantilever* dalla superficie, come descritto nella sezione 2.3.

La preparazione del campione per gli esperimenti con la tecnica del doppio stiramento è esattamente analoga a quella per gli esperimenti con un singolo stiramento (sezione 5.1).

La velocità di allontanamento è, durante entrambi gli stiramenti, pari a 1596 nm/s. Gli altri parametri (distanza a cui avviene il *refolding*, lunghezza della prima ritrazione, tempo lasciato per il *refolding*) variano da curva a curva e sono stati aggiustati manualmente nel corso dell'esperimento, in modo da ottimizzare il numero di eventi visibili nel secondo stiramento della proteina. Il loro valore non influisce sulla lunghezza delle Δ_{Lc} misurate.

Anche in questo caso la posizione del *cantilever* nel piano parallelo al vetrino porta campione viene frequentemente modificata.

Capitolo 6

Analisi dati

6.1 Analisi delle curve

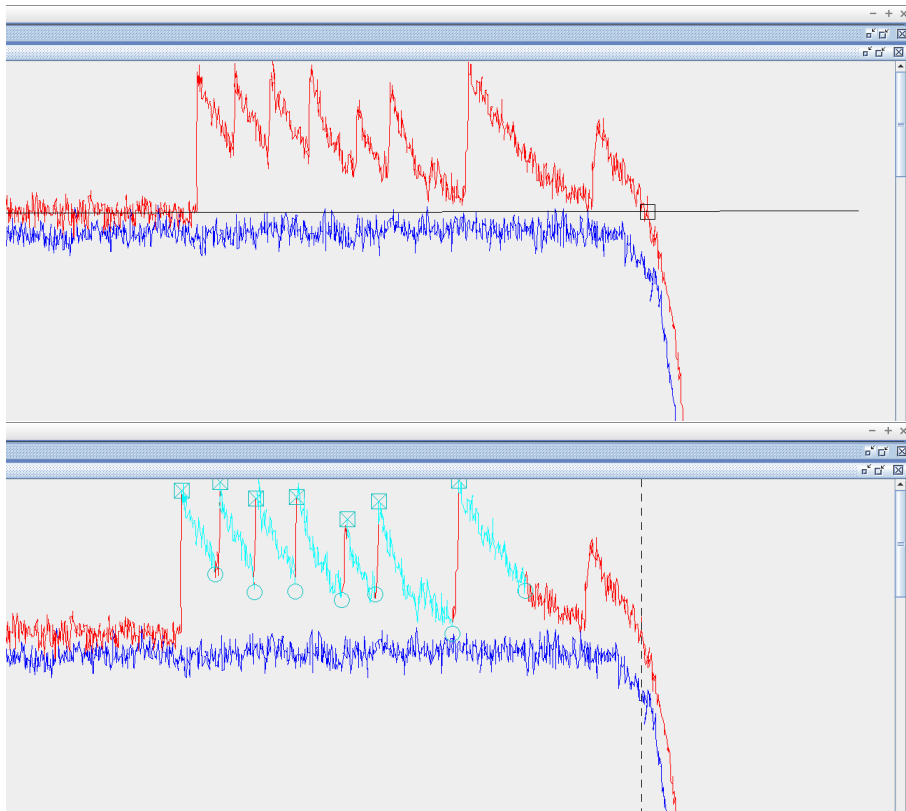


Figura 6.1: Schermata dell'interfaccia di analisi delle curve. IN ALTO: correzione della posizione dello zero. È possibile spostare interattivamente la linea nera. IN BASSO: fit dei picchi, le sezioni azzurre possono essere modificate dall'utente e rappresentano i dati utilizzati per l'ottimizzazione del fit.

Tra tutte le curve sperimentali ottenute si sono selezionate le curve contenenti eventi significativi in base al numero di eventi associabili all' *unfolding* di GB1 presenti, il metodo di riconoscimento di tali eventi è stato definito in seguito all'analisi statistica dei dati ottenuto per il costrutto $GB1_8$, descritta nel capitolo 7. Nell'analisi di $GB1_8$ si sono considerate valide tutte le curve contenenti almeno 3 eventi di *unfolding* con Δ_{Lc} compresa tra 15 e 23 nm (basandosi sulle descrizioni di GB1 presenti in letteratura [7] [33]); in seguito, grazie ai

dati raccolti, è stato possibile fissare la finestra per il riconoscimento di GB1 nelle proteine contenenti moduli prionici, nell'intervallo 15.8-21.2 nm.

Il numero minimo di eventi di GB1 necessari per considerare valida una curva contenente un modulo prionico è posto pari a 5, per avere la certezza che la curva contenga anche eventi associabili alla proteina di interesse, per tutti i costrutti eccetto il tetramero. Infatti esso contiene solo 4 GB1 (due ad ogni estremità), il numero di eventi necessario per considerare una curva come significativa è stato posto, per questo costrutto, pari a 3.

Tramite un *software* di analisi dati sviluppato nel nostro laboratorio [2] si sono filtrate le curve in modo da considerare solo quelle contenenti un numero idoneo di eventi associabili a GB1 ed una lunghezza compatibile con quella totale dei nostri costrutti. I vari picchi sono stati fittati tramite il modello WLC (sezione 2.2.3). In seguito si sono aggiustati manualmente tutti i fit principalmente tramite tre operazioni:

1. eliminazione delle curve contenenti un eccesso di interazione aspecifica (visibile fino ad una distanza del *cantilever* dal substrato maggiore di 50 nm) che avrebbe potuto mascherare il comportamento del costrutto reale;
2. eventuale correzione, curva per curva, dell'inclinazione della linea dello zero;
3. eventuale correzione, picco per picco, del fit: è possibile effettuare questa operazione modificando la sezione di curva appartenente al picco (da utilizzare nella procedura di *fitting*).

In figura 6.1 sono riportate le schermate dell'interfaccia grafica del *software* di fit in due diverse fasi della procedura.

Le curve ottenute sono state filtrate in base a lunghezza minima e massima in modo da essere compatibili con la lunghezza delle proteine in esame, calcolata a partire dal numero di amminoacidi contenuti in ogni costrutto (un numero minoritario di curve è più lungo della lunghezza massima e potrebbe essere associato a dimerizzazioni tra proteine). La seguente tabella riporta le lunghezze minime e massime utilizzate nella selezione delle curve, per ciascun costrutto:

costrutto	lunghezza minima (nm)	lunghezza massima (nm)
$GB1_8$	50	190
$GB1_4PrPTrGB1_4ox$	115	225
$GB1_4PrPTrGB1_4rid$	115	237
$GB1_4PrPTr_2GB1_4$	150	262
$GB1_4PrPTr_3GB1_4$	180	301
$GB1_2PrPTr_4GB1_2$	190	256

6.2 Confronto tra distribuzioni

Allo scopo di riconoscere gli eventi associabili a PrP è necessario confrontare le distribuzioni degli eventi ottenuti con costrutti che lo contengono con le distribuzioni degli eventi ottenuti con il costrutto multimodulare $GB1_8$. In figura 6.2 sono riportati, a titolo esemplificativo, i grafici a dispersione degli eventi ottenuti nel caso di $GB1_8$ e $GB1_4PrPTrGB1_4$ (forza *vs* Δ_{Lc}), ogni punto corrisponde ad un singolo evento di *unfolding*.

Per confrontare due distribuzioni occorre, prima di tutto, filtrare le curve dei due costrutti di interesse, in modo che contengano lo stesso numero minimo di eventi di *unfolding* associabili a GB1. Non è utile, per paragonare due costrutti, un confronto diretto tra gli istogrammi della frequenza degli eventi (in base a forza e Δ_{Lc}), in quanto tale istogramma sarà sempre normalizzato a 1, di conseguenza l'aumento di eventi con Δ_{Lc} lunga provocherà un'inevitabile

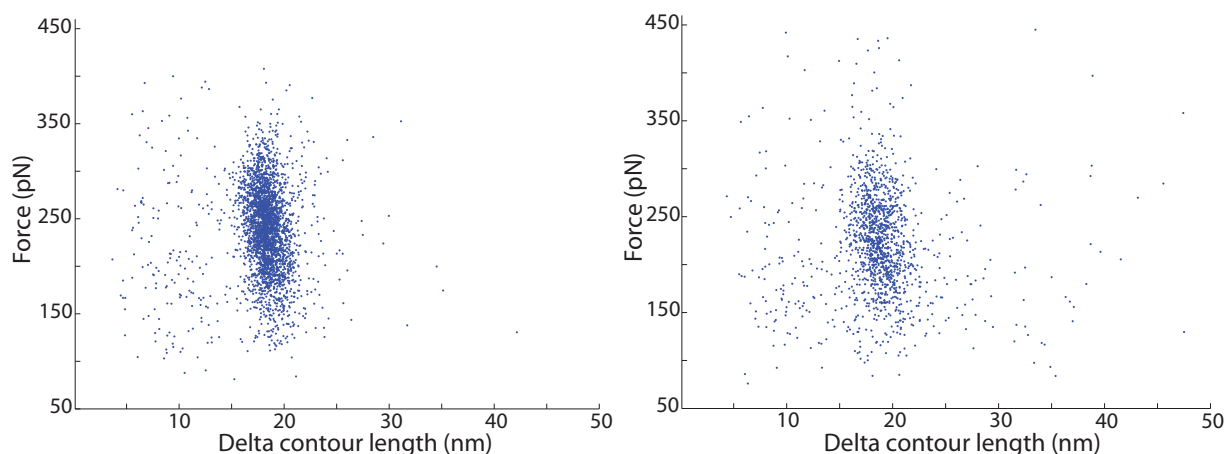


Figura 6.2: Grafici a dispersione (forza vs Δ_{Lc}) ottenuti per $GB1_8$ e $GB1_4PrPTrGB1_4$: ogni punto corrisponde ad un singolo evento di *unfolding*.

diminuzione dell'altezza del picco degli eventi associabili a GB1, rendendo, di fatto, inutile la sovrapposizione dei due istogrammi.

Allo scopo di poter confrontare i dati ottenuti con due costrutti si è sviluppato uno strumento basato sull'analisi del numero medio di eventi per curva: tale sistema opera sia sugli istogrammi monodimensionali (forza o Δ_{Lc}), che sugli istogrammi bidimensionali (che permettono di studiare anche le correlazioni tra la forza di *unfolding* e la lunghezza del dominio unfoldato). Tutto il codice è stato scritto in MATLAB 2013a, ed è riportato in appendice.

Per ogni campione viene calcolato l'istogramma (monodimensionale e/o bidimensionale) per ogni esperimento, dividendo il numero di conteggi in ogni intervallo per il numero totale di curve valide ottenute nell'esperimento. Per ogni intervallo viene calcolata la media e la deviazione standard della media dei conteggi ottenuti sulla singola curva.

Una volta scelti i due campioni che si vogliono confrontare vengono anche calcolate, in ogni intervallo, la differenza dei valori registrati per i due campioni, il valor medio delle deviazioni standard dei due campioni e il rapporto tra i due precedenti valori. Nel caso in cui uno dei due campioni abbia, in un determinato intervallo, media nulla e deviazione standard nulla, mentre l'altro campione abbia deviazione standard diversa da zero, come media delle deviazioni standard dei due campioni viene utilizzata quella del campione che ce l'ha diversa da zero.

Il rapporto tra la differenza e la deviazione standard media viene utilizzato come metrica per riconoscere la significatività delle differenze: se esso è maggiore di 3 le differenze sono considerate significative.

6.2.1 Analisi monodimensionale

Nella parte superiore si osserva la sovrapposizione degli istogrammi, ogni classe è corredata, oltre che dal valore medio, dalla deviazione standard inter-esperimento. Poichè lo scopo di questo tipo di analisi è il confronto tra due distribuzioni viene riportato l'istogramma limitatamente alla parte contenente entrambe le distribuzioni. Le linee spezzate uniscono i valori medi delle varie classi e sono tracciate al puro scopo di aiutare la lettura visiva del grafico.

Nella parte inferiore è riportato il rapporto tra la differenza tra i valori ottenuti per ogni singola classe e la deviazione standard media tra i due campioni in quella medesima classe. In questo modo si ha una lettura immediata della significatività della differenza, tramite l'altezza dei punti. La posizione delle linee rosse è impostabile dall'utente, nei grafici presenti

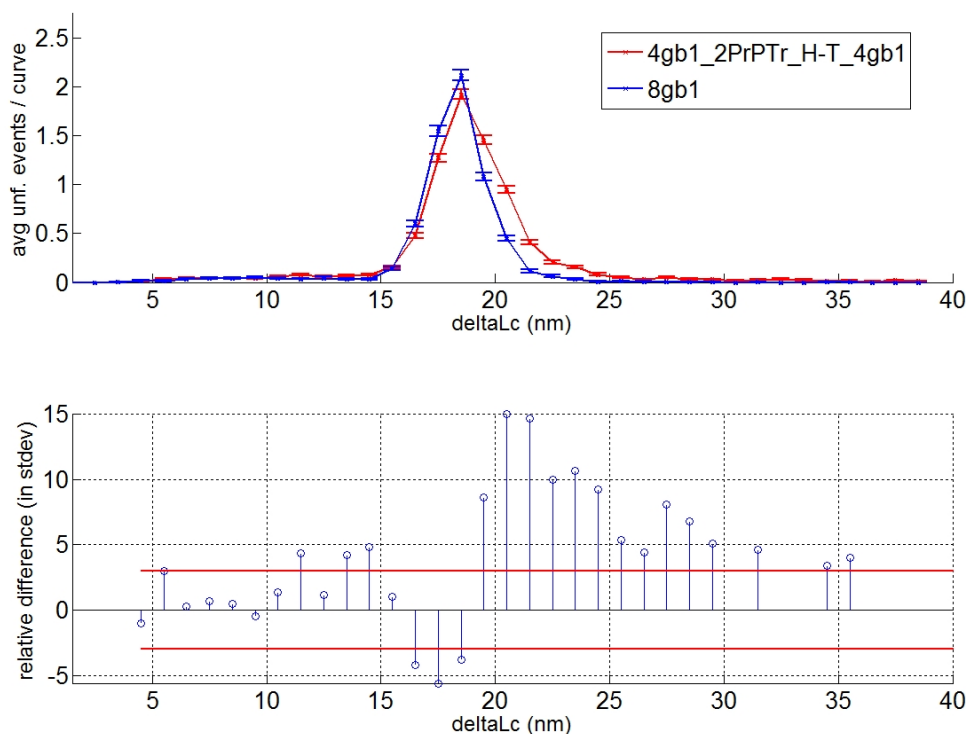


Figura 6.3: Sistema di analisi monodimensionale. IN ALTO: la sovrapposizione degli istogrammi riportanti il numero medie di eventi per curva, corredati da deviazione standard inter-esperimento (sia la media che la deviazione standard sono pesate in base al numero di curve presenti nell'esperimento). IN BASSO: l'altezza dei punti rappresenta il rapporto tra la differenza tra i valori misurati per i due costrutti nelle classi di interesse e la deviazione standard media dei due costrutti. La riga rossa è posta in corrispondenza a tre deviazioni standard.

in questa tesi punti esterni alle linee rosse corrispondono a classi in cui i conteggi per i due costrutti distano più di 3 deviazioni standard medie. In classi in cui uno dei due costrutti non è presente non viene riportata l'altezza della differenza delle distribuzioni. Per confrontare due costrutti viene prodotto un grafico come quello riportato in figura 6.3.

6.2.2 Analisi bidimensionale

Per confrontare due costrutti considerando anche le correlazioni tra forza e Δ_{Lc} si osservano gli istogrammi bidimensionali. Vengono prodotti quattro grafici, i cui assi vengono impostati automaticamente identici per i due campioni da confrontare e vengono scelti in modo che siano visibili per entrambi i campioni tutte le classi contenenti un valore diverso da 0. Lo schema di rappresentazione è riportato in figura 6.4:

- In alto: istogramma bidimensionale dei due costrutti. La scala di colori è la medesima per entrambi i grafici in modo da poter facilmente fare un confronto qualitativo visivo. Supponiamo, per semplicità, che il costrutto a destra sia il campione, mentre quello a sinistra rappresenti il riferimento (la loro trattazione è comunque simmetrica).
- In basso a sinistra: differenza tra i due istogrammi in unità di deviazioni standard medie dei due costrutti, corredato da una scala cromatica. Valori positivi sono indicativi di classi in cui il campione è maggiormente presente del riferimento, valori negativi

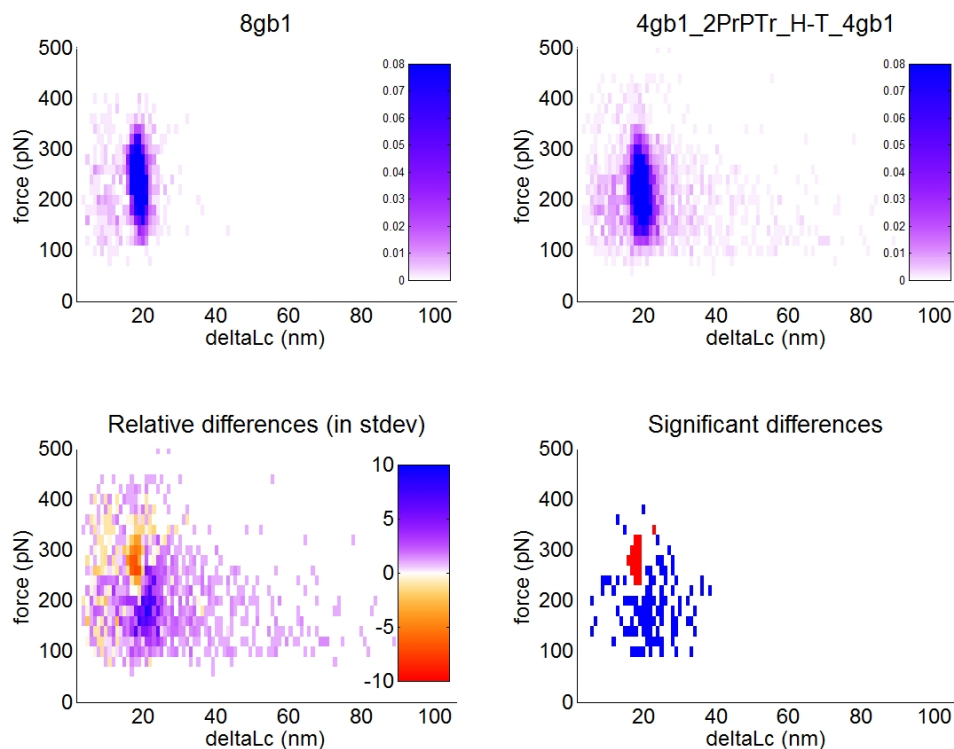


Figura 6.4: Sistema di analisi bidimensionale. IN ALTO: i due istogrammi bidimensionali affiancati, riportanti il numero medie di eventi per curva (forza *vs* Δ_{Lc}). Indichiamo come riferimento il costruito sulla sinistra e come campione quello sulla destra. IN BASSO A SINISTRA: l'altezza dei punti rappresenta il rapporto tra la differenza dei valori misurati per i due costrutti nella classe di interesse e la deviazione standard dei due costrutti. Valori positivi sono indicativi di classi in cui il campione è maggiormente presente del riferimento, valori negativi indicano classi in cui il riferimento è maggiormente presente rispetto al campione. IN BASSO A DESTRA: in blu sono indicate le classi in cui il riferimento presenta un numero di eventi significativamente maggiori, in rosso le classi in cui il campione presenta un numero di eventi significativamente maggiore. Le differenze sono considerate significative se il rapporto raffigurato in basso a sinistra è maggiore di 3.

indicano classi in cui il riferimento è maggiormente presente rispetto al campione. L'interpretazione di questo grafico è analoga a quella delle altezze dei punto nella parte bassa della figura 6.3, con la sola eccezione che in questo caso vengono colorate tutte le classi, anche quelle in cui uno dei due campioni non è presente.

- In basso a destra: differenze significative tra i due istogrammi, determinate in base alla deviazioni standard media dei costrutti, (calcolata nella singola classe). Il numero di deviazioni standard utilizzate per definire la significatività è impostabile dall'utente, ma in questa tesi è posto uguale a tre. Colori diversi indicano se la differenza significativa è dovuta a eccesso o a mancanza del campione rispetto al costruito di riferimento: il rosso indica eccesso di campione, il blu eccesso di riferimento. L'interpretazione di queste classi colorate è analoga all'analisi dei punti che si trovano all'esterno della riga rossa nella parte bassa del grafico in figura 6.3. Lo scopo di questo grafico è analizzare la significatività di differenze in regioni in cui gli eventi sono molto frequenti, non sono visualizzati gli eventi sporadici.

Parte III

Risultati e discussione

Capitolo 7

Caratterizzazione di GB1

7.1 Dati sperimentali

Si sono effettuati esperimenti di *unfolding* a tre diversi pH (4.0, 5.5 e 7.4), ottenendo per ogni valore di pH circa 1300 curve valide (considerando valide le curve contenenti almeno tre eventi di *unfolding* di lunghezza compresa tra 15 e 23 nm e lunghezza totale della molecola compresa tra 50 e 190 nm). I campioni utilizzati a tutti i pH sono misti, sia freschi che congelati. In figura 7.1 è riportata una curva tipica di un esperimento di *unfolding* di GB1₈, mentre in figura 7.2 sono riportati i grafici a dispersione ottenuti ai tre pH: è subito evidente come a pH 4.0 gli eventi siano distribuiti a valori di forza minori.

7.2 Analisi statistica dei dati

Si sono prodotti gli istogrammi delle forze degli eventi di *unfolding* e delle relative Δ_{Lc} , normalizzandoli rispetto al numero di curve, in modo da leggere sull'asse delle ordinate il numero medio di eventi per ogni classe nella singola curva.

Tale istogramma è stato ottenuto anche per i singoli esperimenti, in modo da poter calcolare la deviazione standard del valor medio della singola classe tra i vari esperimenti (pesata in base la numero di curve raccolte per ogni esperimento). Il numero di esperimenti effettuati è pari a cinque per pH 7.4, quattro per pH 5.5 e sei per pH 4.0. I due grafici riportati in figura 7.3 raffigurano la sovrapposizione dei tre istogrammi, con la relative deviazioni standard.

Per quanto riguarda le forze si nota quanto già precedentemente sottolineato: a pH 4.0 la distribuzione dei singoli eventi di *unfolding* è piccata su valori di forza inferiore rispetto a quelle a pH più alto. Tale andamento è compatibile con quanto presentato in letteratura [33]: Hong Bin Li e collaboratori hanno misurato l'effetto del pH sulle forze di *unfolding* di una variante (mutante) di GB1 utilizzando la stessa tecnica qui presentata giungendo alla conclusione che esse diminuiscono al diminuire del pH. Il nostro lavoro permette di concludere che questo è valido anche per la versione non mutata di GB1.

Dalla sovrapposizione delle distribuzioni delle Δ_{Lc} ai diversi pH si nota come la lunghezza del dominio sfoldato sia indipendente dall'acidità della soluzione in cui è immersa la proteina. Un fit gaussiano di tali distribuzioni restituisce i valori riportati in tabella:

pH	μ	σ
7.4	18.3	0.9
5.5	18.7	1.0
4.0	18.5	0.9

Tabella 7.1: Valori di μ e σ ottenuti con un fit gaussiano sulla distribuzione delle Δ_{Lc}

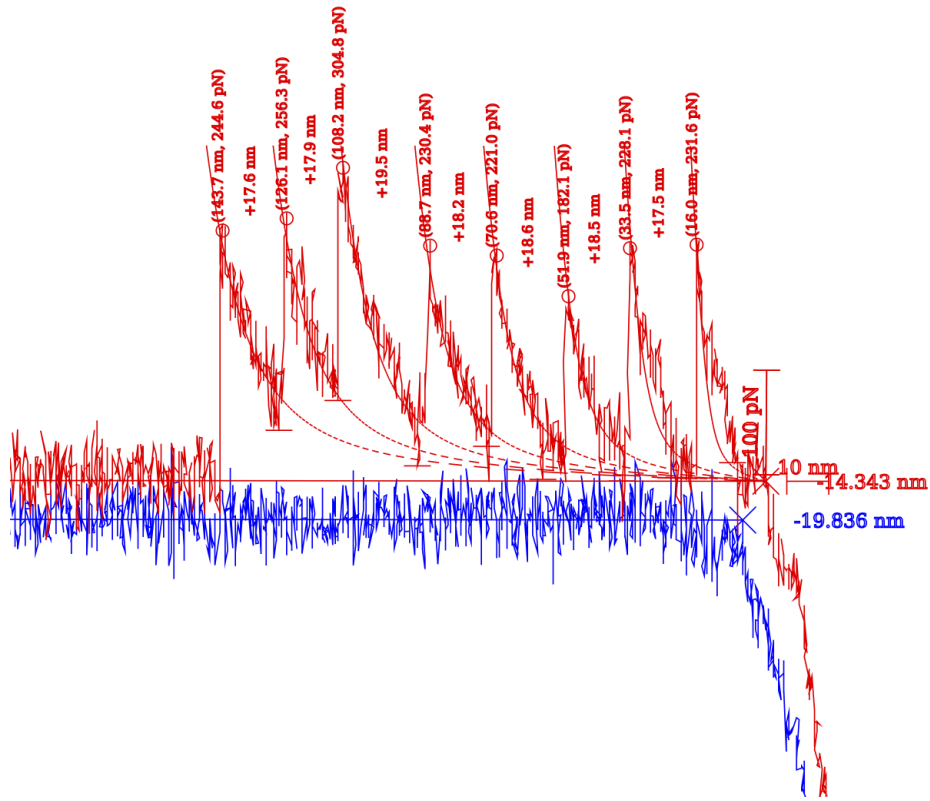


Figura 7.1: Curva tipica di un esperimento di *unfolding* di $GB1_8$. Sono ben riconoscibili sette eventi di *unfolding*.

Questi valori sono in buon accordo con quanto già presente in letteratura ([19] riporta un valore di Δ_{Lc} di 18.3 ± 0.5 nm, ottenuta a pH 7).

La finestra per il riconoscimento degli eventi di GB1 nell'analisi delle curve ottenuti con costrutti contenenti uno o più PrP viene fissata in base al fit gaussiano della distribuzione, nell'intervallo $\mu \pm 3\sigma$ (considerando la media dei tre fit): di conseguenza in queste curve vengono valutati gli eventi con $15.8 \text{ nm} < \Delta_{Lc} < 21.2 \text{ nm}$ come appartenenti con buona probabilità a GB1.

7.3 Analisi degli eventi con $\Delta_{Lc} < 15$ nm

Dall'osservazione della distribuzione delle Δ_{Lc} si evince come la coda sinistra abbia un peso maggiore di quella destra e tale fatto non è discusso in letteratura. Eventi con $\Delta_{Lc} < 15$ nm (d'ora in avanti indicati anche come picchi corti) sono presenti a tutti i pH in quantità comparabili e significativamente diverse da zero. La seguente tabella contiene le percentuali di curve che contengono n picchi con $\Delta_{Lc} < 15$ nm, mostrando come non sia presente una dipendenza dal pH.

pH	$n \geq 1$	$n=1$	$n=2$	$n \geq 3$
7.4	25.5	14.3	9.5	1.4
5.5	21.8	14.0	6.2	1.5
4.0	25.7	17.7	6.8	1.2

Tabella 7.2: Percentuale di curve contenenti n picchi con $\Delta_{Lc} < 15$ nm al variare del pH.

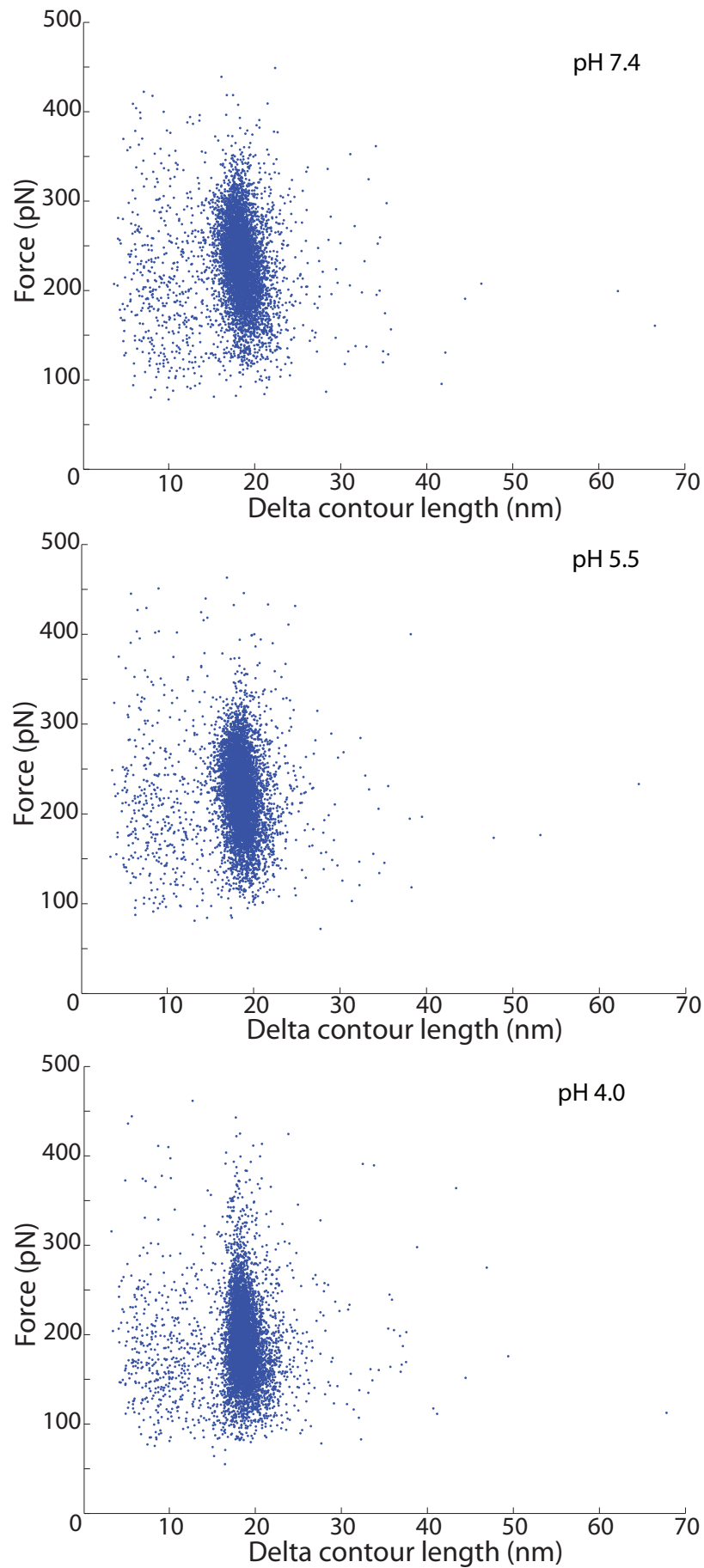


Figura 7.2: Grafici a dispersione ottenuti ai tre diversi pH per *GB1*₈: si nota come a pH 4.0 la distribuzione degli eventi sia spostata verso valori di forza minori.

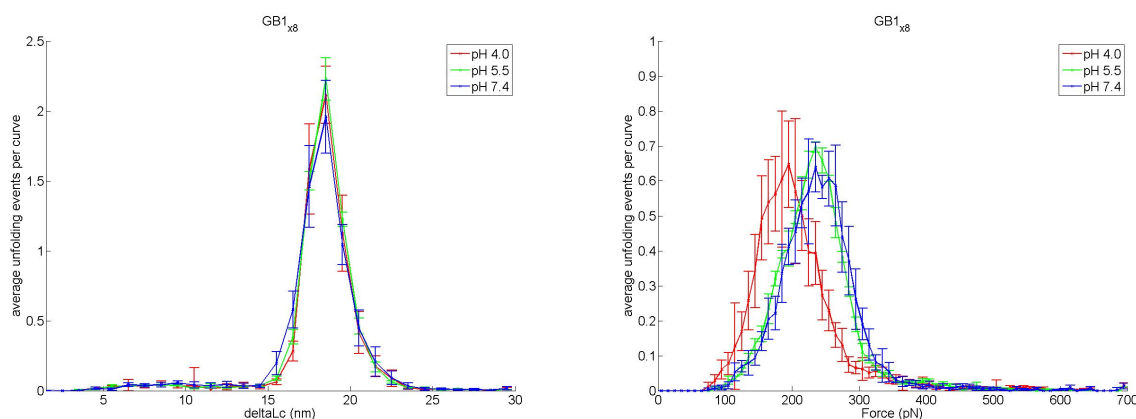


Figura 7.3: Sovrapposizione degli istogrammi riportanti il numero medio di eventi per curva a diversi pH. A sinistra è visibile l'istogramma delle Δ_{Lc} , mentre a destra è riportato quello delle forze. Le barre di errore sulla singola classe corrispondono alla deviazione standard interesperimento, pesata in base al numero di curve contenute in ogni esperimento. E' ben visibile l'ottima sovrapposizione delle distribuzioni delle Δ_{Lc} , mentre si nota una diminuzione della forza media a pH 4.0.

Si è dunque fatta una classificazione, dividendo i picchi “corti” in tre categorie:

- picchi singoli (quelli che compaiono in curve con solo un picco “corto”, o in curve in cui siano presenti 3 picchi “corti” di cui 2 appaiati ed uno isolato, quest'ultimo è classificato come singolo);
- picchi doppi (quelli che compaiono appaiati in curve che presentano 2 o 3 picchi “corti”);
- altri picchi (quelli che non rientrano nelle categorie precedenti).

In questo modo si osserva che i picchi sono distribuiti nelle tre categorie secondo le seguenti percentuali:

pH	singoli (%)	doppi (%)	altro (%)
7.4	36.6	45.1	18.3
5.5	44.2	34.4	21.4
4.0	49.9	27.8	22.3

La tabella mette in evidenza come al diminuire del pH aumenta la percentuale di picchi classificati come singoli, mentre diminuiscono i picchi classificati doppi.

A tutti i pH molti picchi singoli corrispondono al primo evento di *unfolding* (il 43 % dei picchi singoli misurati a pH 7.4 e 4.0 e il 54 % di quelli misurati a pH 5.5).

7.3.1 Coppie di picchi corti

Di particolare interesse è lo studio dei picchi doppi in quanto la loro presenza potrebbe essere indicativa di un intermedio di *unfolding*.

Studiando le Δ_{Lc} di questi picchi si nota come l'ordine di comparsa non sia direttamente correlato alla lunghezza del picco, infatti al primo dei due picchi corrisponde una Δ_{Lc} maggiore nel 56 % degli eventi doppi a pH 7.4, nel 51 % di quelli a pH 5.5 e nel 55 % di quelli a pH 4.0. Generalmente nel caso di intermedi di *unfolding* l'ordine in cui si presentano gli eventi è riproducibile in quanto una parte della proteina risulta più instabile e si denatura sempre per prima ([20]). In questo caso, al contrario, non è possibile riconoscere un ordine nettamente preferenziale.

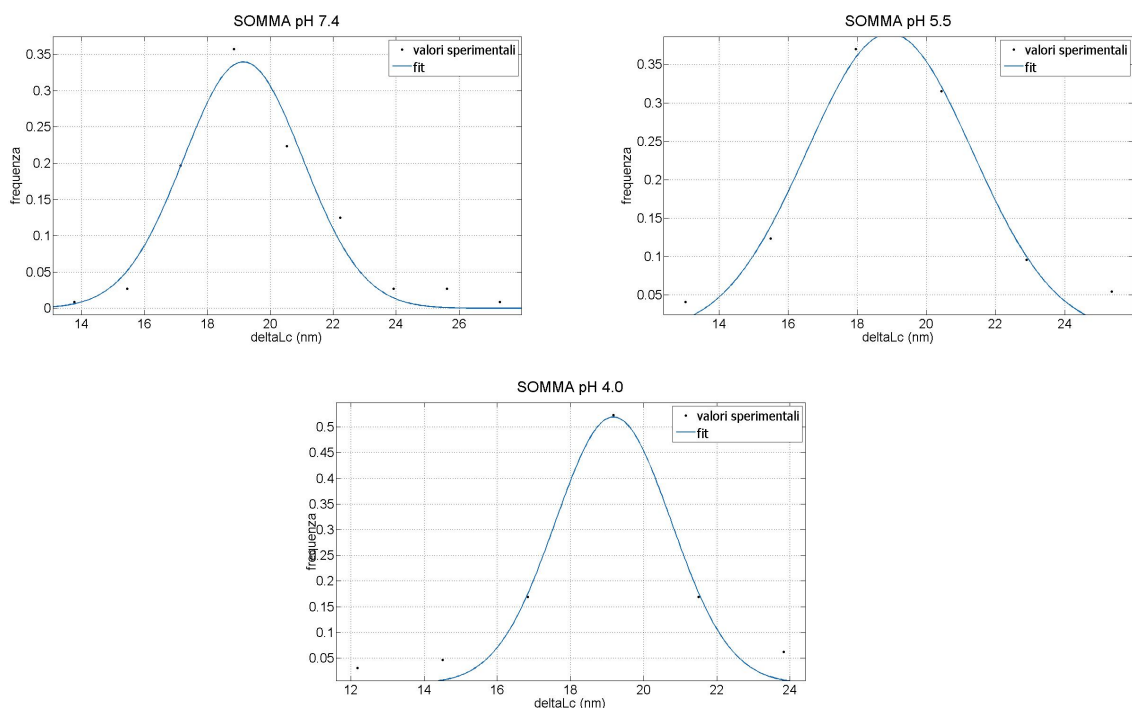


Figura 7.4: Distribuzioni delle somme dei picchi doppi: la distribuzione ottenuta è compatibile con quella del picco principale.

Si nota come le forze del primo picco corto sono, a tutti i pH, mediamente più elevate di quelle del secondo picco della coppia:

pH	Forza primo picco (media \pm dev std) (pN)	Forza secondo picco (media \pm dev std) (pN)
7.4	237 \pm 66	187 \pm 70
5.5	236 \pm 62	192 \pm 87
4.0	201 \pm 57	161 \pm 40

Nuovamente è evidente come, per quanto riguarda la forza, il comportamento a pH 5.5 e quello a pH 7.4 siano sovrapponibili, mentre a pH 4.0 le forze tendano a diminuire. Se questi eventi rappresentassero veramente un intermedio di *unfolding* la loro somma dovrebbe essere distribuita come quella degli eventi associabili alla transizione principale. In figura 7.4 è riportata la distribuzione di tale somma a tutti i valori di pH, con sovrapposto un fit gaussiano, i valori di μ e σ per tali fit sono riportati nella seguente tabella:

pH	μ	σ
7.4	19.1	1.2
5.5	18.9	1.3
4.0	19.2	1.1

Confrontando questi dati con quelli presentati in tabella 7.1 si nota come essi siano effettivamente compatibili con il picco principale di GB1, le deviazioni standard sono maggiori rispetto a quelle della distribuzione originale, è necessario comunque tener presente che questi dati derivano da una somma e quindi l'errore dei due valori addizionati si propaga, contribuendo all'allargamento della gaussiana.

Non è possibile riconoscere delle distribuzioni note negli istogrammi delle Δ_{Lc} del primo e del secondo picco, che sono piuttosto simili. Nell'ipotesi in cui questi picchi "corti" rappresentino effettivamente degli intermedi di *unfolding* la loro lunghezza dovrebbe essere ben

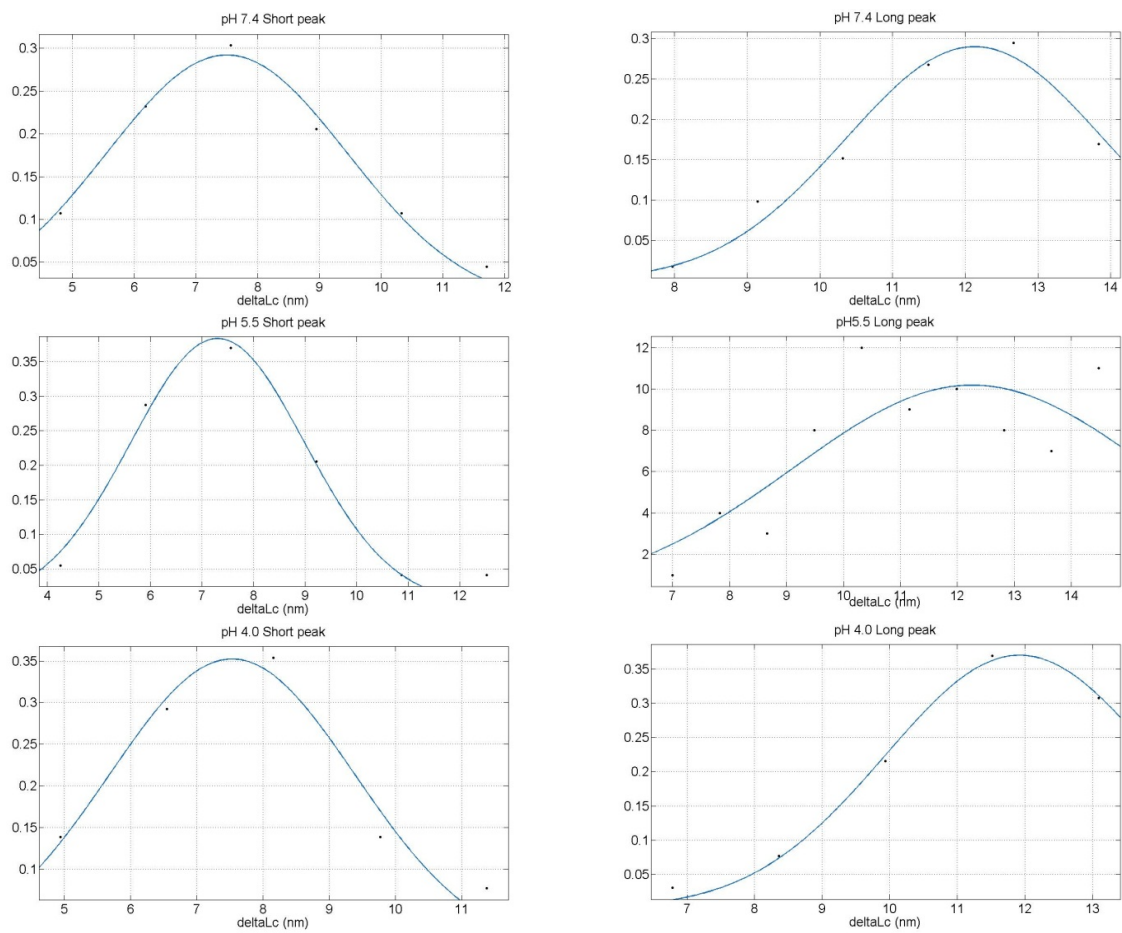


Figura 7.5: Fit delle distribuzioni di Δ_{Lc} per i due picchi della coppia, divisi in base alla lunghezza relativa: i punti indicano l'altezza della classe, mentre la linea continua indica il fit gaussiano, sull'asse delle ordinate è riportata la frequenza, normalizzata al numero di eventi.

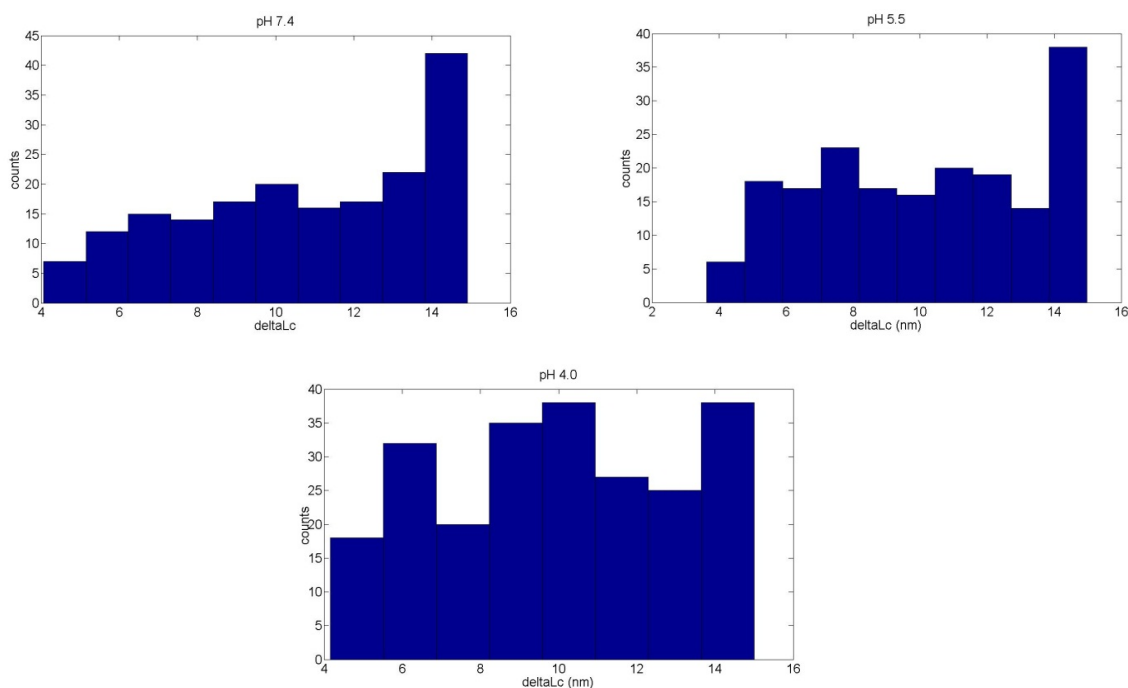


Figura 7.6: Istogramma della Δ_{Lc} degli eventi classificati come singoli.

determinata, si guarda dunque la distribuzione del picco corto della coppia e di quello lungo, ottenendo i grafici riportati in figura 7.5. Nella seguente tabella sono riportati i risultati della procedura di fit gaussiano:

pH	picco con Δ_{Lc} minore		picco con Δ_{Lc} maggiore	
	μ (nm)	σ (nm)	μ (nm)	σ (nm)
7.4	7.5	2.0	12.1	1.8
5.5	7.3	1.7	12.3	3.1
4.0	7.5	1.9	11.9	2.0

Dividendo i picchi in base alla loro lunghezza si ottiene necessariamente una distribuzione di valori più corti per il picco corto e più lunghi per il picco lungo, in ogni caso però si nota che per quanto riguarda il picco corto le tre distribuzioni ottenute ai tre diversi pH sono molto simili e il fit gaussiano è piuttosto buono ($R^2 > 0.94$ a tutti i pH). Anche le distribuzioni del picco lungo a pH 4.0 e 7.4 sono molto simili ($R^2 > 0.96$). Al contrario il fit del picco lungo a pH 5.5 non è altrettanto buono ($R^2 = 0.65$). Le deviazioni standard di queste distribuzioni sono, in generale, troppo larghe per trattarsi dell'*unfolding* di una struttura ben definita.

7.3.2 Picchi corti singoli

Gli eventi classificati come singoli presentano a tutti i pH distribuzioni senza picchi accentuati (si veda figura 7.6). L'innalzamento del numero di conteggi in corrispondenza dell'ultima classe, visibile sia a pH 7.4 che a pH 5.5, può essere attribuito alla coda sinistra del picco principale centrato su 18 nm. Questa distribuzione degli eventi singoli potrebbe essere dovuta al modo casuale in cui la proteina si attacca alla superficie o alla punta; ovvero non è possibile essere certi che l'adesione avvenga tra un dominio e l'altro, l'eventuale adesione in mezzo ad un dominio potrebbe provocare la presenza di questi eventi, associabili alla rottura della parte del dominio libera dai legami.

7.3.3 Curve prive di interazione aspecifica

Per verificare che il gran numero di picchi singoli corrispondenti al primo evento di *unfolding* non fossero dovuti ad un errore sistematico nella procedura di fit, con conseguente scambio di un segnale associabile all'interazione aspecifica con un evento di *unfolding*, si è osservata la frequenza di comparsa dei picchi corti limitandosi alle curve completamente prive di interazione aspecifica a pH 7.4 (in quanto abbassando il pH l'interazione aspecifica tende ad aumentare). La percentuale di curve contenenti picchi corti è riportata nella seguente tabella e, confrontandola con la tabella 7.2 si nota come l'assenza di aspecifico non diminuisca drasticamente la presenza di questi picchi.

pH	n \geq 1	n=1	n=2	n \geq 3
7.4	21.1	11.0	9.3	0.9

Inoltre in queste curve il 56 % dei picchi classificati come singoli corrispondono al primo evento di *unfolding* e questo permette di escludere che la loro presenza sia dovuta ad un errore sistematico di fit.

7.4 Curve con il metodo del doppio impulso

Si sono effettuati esperimenti con il metodo del doppio stiramento a tutti tre i pH, ottenendo in tutti i casi circa 120 curve significative. In figura 7.7 è riportata una tipica curva ottenuta stirando due volte il costrutto GB1₈.

Analizzando i dati ottenuti nel secondo stiramento (in quanto privi di interazione aspecifica, si ottengono, per le Δ_{Lc} gli istogrammi riportati in figura 7.8.

È nuovamente visibile il picco principale di GB1, centrato a circa 18.5 nm, affiancato sulla destra da un numero minoritario di eventi caratterizzati da Δ_{Lc} corta. Analizzando tali picchi si ottengono i valori riportati nella seguente tabella:

pH	n \geq 1	n=1	n=2	n \geq 3
7.4	17.3	14.0	0.8	1.6
5.5	22.4	19.0	3.4	0
4.0	24.6	18.8	2.9	2.9

Non è possibile un confronto diretto e quantitativo con i dati presentati in tabella 7.2 in quanto, come spiegato in sezione 5.2, la velocità di ritrazione è diversa. Ciò nonostante si può osservare qualitativamente che la riduzione degli eventi doppi è notevole rispetto alle curve di semplice *unfolding*.

7.5 Struttura di GB1

In figura 7.9 sono riportate la sequenza amminoacidica e la struttura secondaria di GB1, come descritte in PDB. I primi due foglietti β sono composti da venti amminoacidi, per cui, considerando la lunghezza di un legame peptidico pari 0.36 nm [24], la lunghezza di questo tratto di proteina è pari a $20 \cdot 0.36 = 7.2$ nm. La restante parte della proteina è lunga $36 \cdot 0.36 = 12.9$ nm. Questa coppia di lunghezze è confrontabile con i due picchi delle gaussiane corrispondenti alle Δ_{Lc} delle coppie di picchi corti e si potrebbe quindi ipotizzare che queste possano essere le strutture associate a tali eventi; anche gli ultimi due foglietti β possono restituire un analogo risultato. I dati da noi raccolti non sono sufficienti ad associare con certezza gli eventi di *unfolding* corti a sezioni della proteina.

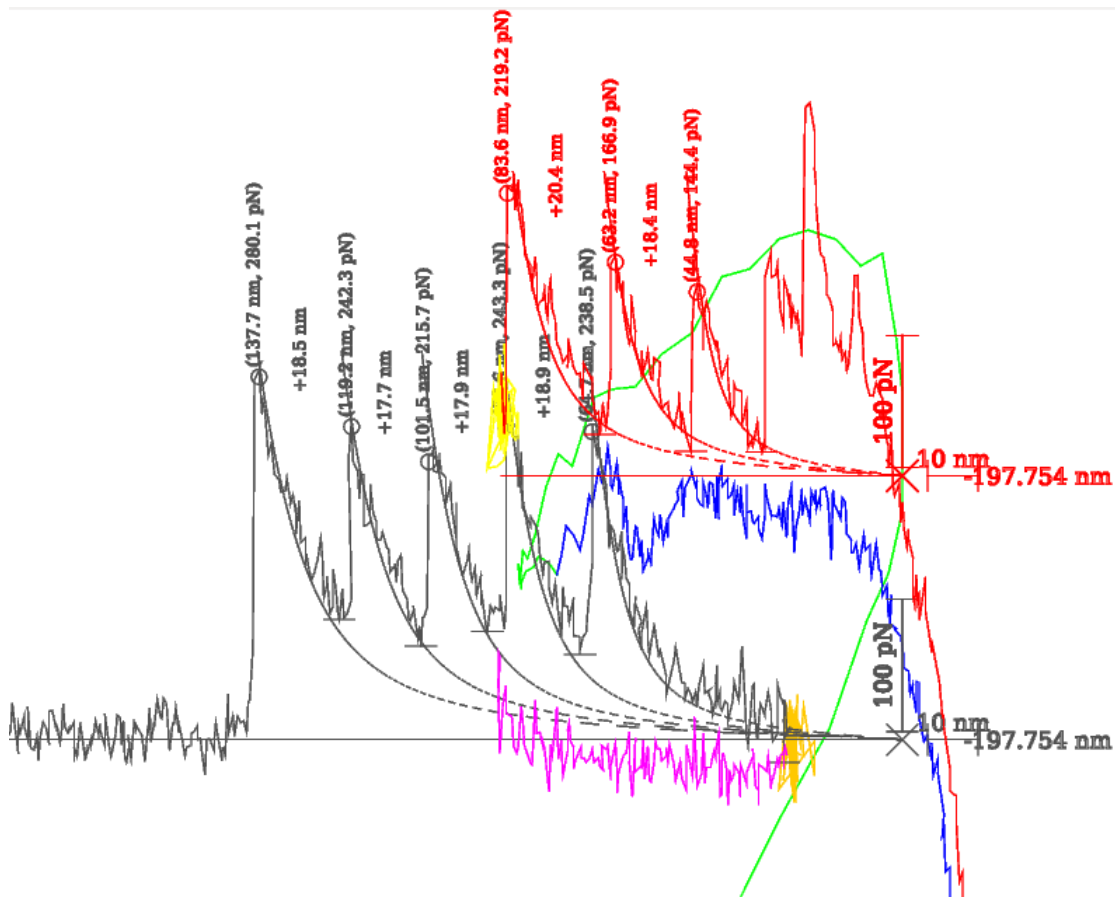


Figura 7.7: Sequenza di stiramenti di GB1, i colori indicano le varie fasi della misura: in blu il primo avvicinamento, in rosso il primo stiramento, sono riconoscibili i segnali di 3 eventi di *unfolding*, in magenta il secondo avvicinamento, l'arancione corrisponde al tempo in cui il sistema è rimasto fermo per permettere il *refolding*, il nero al secondo stiramento, sono riconoscibili 4 eventi di *unfolding* e il distacco finale. Nel secondo stiramento non è visibile interazione aspecifica.

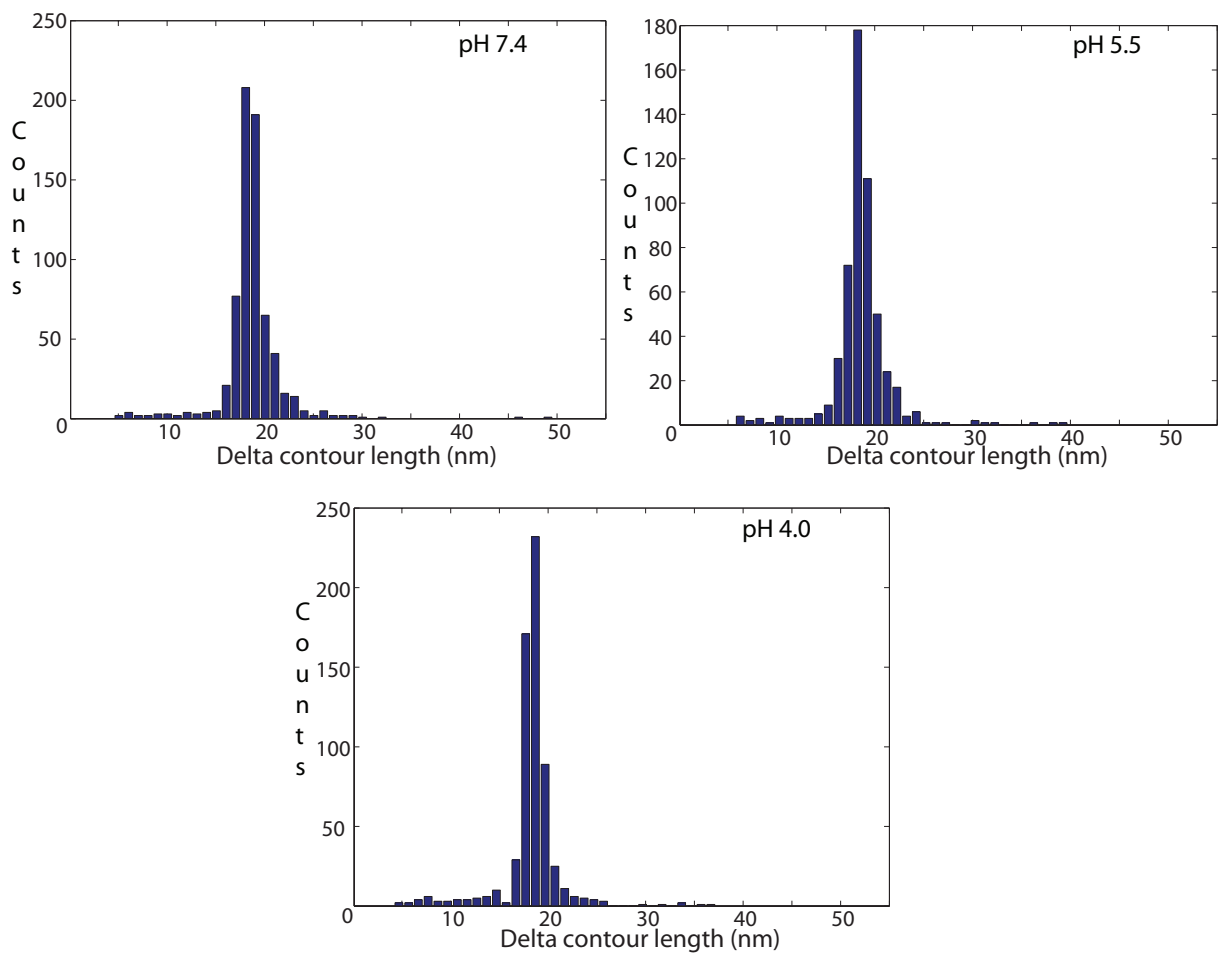


Figura 7.8: Istogrammi delle Δ_{Lc} ottenute nel corso del secondo stiramento di $GB1_8$.

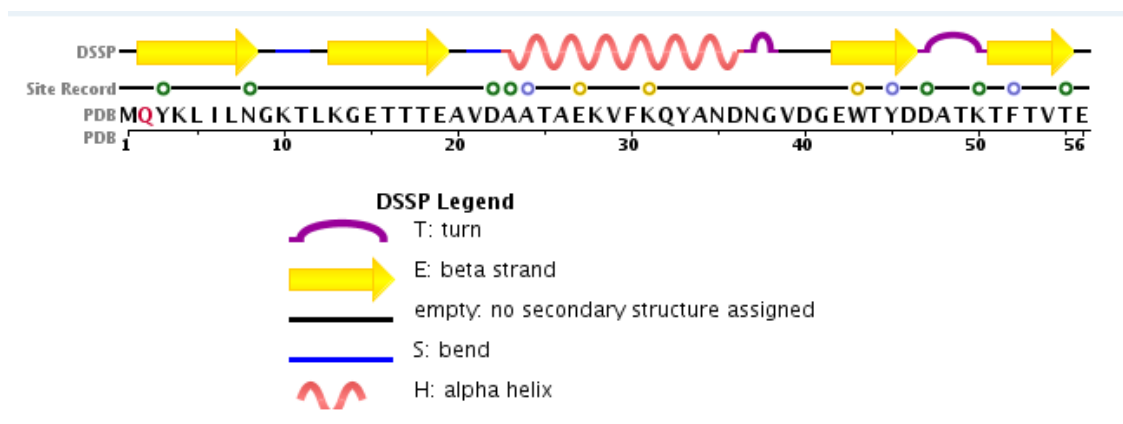


Figura 7.9: Struttura di GB1, come riportata da PDB

7.6 Sommario

Dall'analisi delle curve ottenute per il costrutto $GB1_8$ si sono tratte le seguenti conclusioni:

- è possibile confermare quanto presentato in letteratura riguardo al meccanismo di *unfolding*, descrivendolo come un processo a due stati (gli eventi corti sono in ogni caso assolutamente minoritari);
- la forza di *unfolding* è dipendente dal pH. Le sue distribuzioni a pH 7.4 e 5.5 sono sovrapponibili, mentre si sposta significativamente a valori di forza inferiori a pH 4.0;
- la finestra per il riconoscimento degli eventi di GB1 nei costrutti multimodulari è stata fissata nell'intervallo di Δ_{Lc} comprese tra 15.8 e 21.2 nm;
- la coda sinistra della distribuzione pesa notevolmente di più di quella destra, al suo interno sono riconoscibili sia eventi che si presentano individualmente, che eventi ravvicinati a coppie. La somma delle Δ_{Lc} degli eventi che si presentano a coppie è ben fittabile con una gaussiana simile alla gaussiana principale degli eventi. L'origine di questi eventi non è ancora stata indagata e si necessita di ulteriori esperimenti per caratterizzarli completamente; in ogni caso la loro presenza va considerata nel confronto delle distribuzioni ottenute con costrutti multimodulari contenenti moduli di cui la struttura è ignota. Al fine di individuare l'origine di questi eventi di *unfolding* si potrebbe effettuare esperimenti con il metodo degli stiramenti multipli, infatti nel caso in cui questi eventi corrispondano all'*unfolding* di intermedi stabili la loro presenza in percentuale deve essere la medesima in tutte gli stiramenti.

Capitolo 8

Caratterizzazione di PrP

Il numero totale di esperimenti effettuati e di curve contenenti il numero minimo di GB1 per poter essere considerate valide sono riportati nella seguente tabella.

costrutto	pH	num. esperimenti	num. curve selezionate	
monomero (ossidato)	7.4	17	189	
	5.5	2	83	
	4.0	7	90	
monomero(ridotto)	7.4	11	184	
	dimero H-T	7.4	15	682
		5.5	3	107
dimero H-H	4.0	5	69	
	7.4	11	285	
	trimero	7.4	30	129
5.5		5	136	
4.0		12	88	
tetramero	7.4	15	58	
	5.5	7	65	
	4.0	7	46	

A pH 7.4 in tutti i costrutti, eccetto il tetramero, si sono ottenute molte curve. A pH acidi, invece, per alcuni costrutti, il numero di curve raccolte è ancora inferiore al centinaio. Per questo motivo l'analisi e la caratterizzazione dei costrutti sono state effettuate prevalentemente a pH 7.4.

Per effettuare il confronto dei dati ottenuti da questi costrutti con quelli ottenuti per $GB1_8$ si sono filtrate le curve di quest'ultimo in modo che contenessero un numero di eventi nell'intervallo 15.8-21.2 nm paragonabile a quello del costrutto di interesse (cioè minimo 5 per tutti i costrutti eccetto il tetramero che può avere solo 3 o 4 eventi associati a GB1). Nella seguente tabella è riportato il numero di curve contenenti n eventi con Δ_{Lc} compresa tra 15.8 e 21.2 nm, per il costrutto $GB1_8$ ai vari pH. Solo queste curve sono state utilizzate nei confronti con i costrutti contenenti una o più unità di MoPrP.

pH	$n > 5$	$3 \leq n \leq 4$
7.4	471	642
5.5	618	592
4.0	498	642

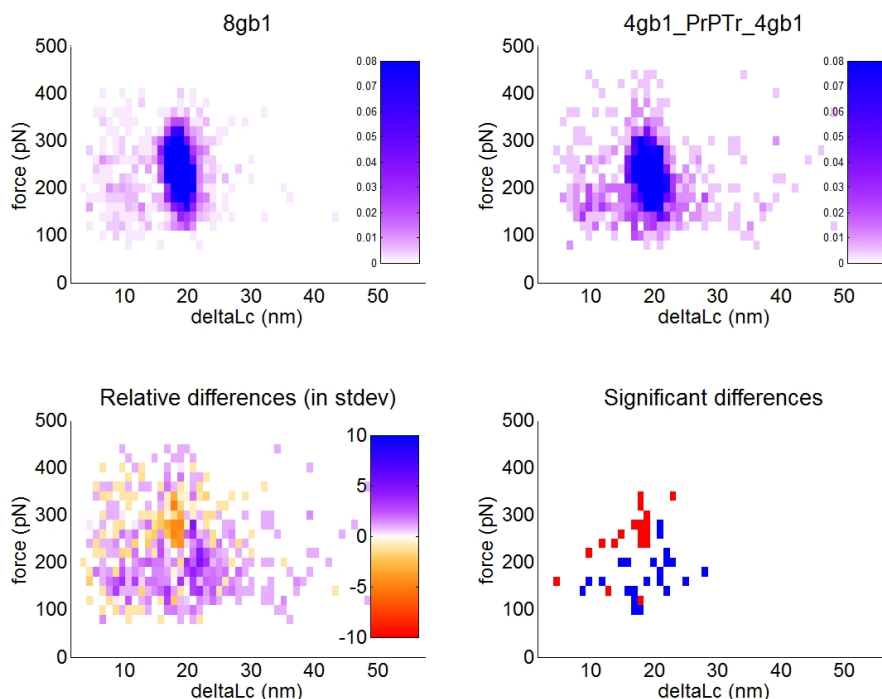


Figura 8.1: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTrGB1_4$ e di $GB1_8$; il metodo utilizzato è descritto in sezione 6.2.2

8.1 Esperimenti a pH neutro

8.1.1 Monomero in ambiente ossidante e riducente: riconoscimento della struttura nativa.

Esperimenti con proteina in condizioni ossidanti

Osservando gli istogrammi bidimensionali riportati nella prima riga di figura 8.1 si nota come non siano riconoscibili due distribuzioni di eventi nel grafico di $GB1_4PrPTrGB1_4$. La posizione della distribuzione principale degli eventi è la medesima della distribuzione visibile in $GB1_8$, ma la dispersione dei dati appare maggiore.

In particolare, osservando l'istogramma delle differenze riportato nel quadrante in basso a sinistra, si nota la presenza di una vasta area viola, caratterizzata da eventi con forza minore di 250 pN e Δ_{Lc} compresa tra 20 e 30 nm. Il numero medio di eventi appartenenti a questa regione nelle curve ottenute per il costrutto contenente MoPrP è dunque maggiore rispetto al numero medio di eventi presenti nelle curve ottenute per il costrutto contenente solo GB1. Si presenta in viola anche una regione a forza bassa e lunghezza inferiore a 18 nm.

È possibile osservare una regione, centrata su circa 18 nm e 280 pN, che appare gialla, questo indica che in questa zona il numero di eventi misurati con il costrutto $GB1_8$ tende ad essere più alto di quelli di $GB1_4PrPTrGB1_4$.

Il minor numero di eventi nel picco principale di GB1 può essere correlato a due fattori:

1. in fase di selezione delle curve alcuni eventi di MoPrP possono essere stati scambiati con eventi di GB1, essendo essi nella stessa regione di Δ_{Lc} . Tuttavia gli eventi di MoPrP hanno mediamente lunghezza maggiore e forza minore e quindi l'effetto netto visibile nei dati sono queste regioni di abbondanza di uno rispetto all'altro;

2. la probabilità di adesione di una proteina alla punta è legata alla sua lunghezza; aumentando la lunghezza della proteina la probabilità che essa aderisca alle sue estremità è minore. In presenza di MoPrP ci si può, quindi, attendere, una lieve diminuzione del numero di eventi associabili a GB1.

Osservando il quadrante in basso a destra si nota che alcune delle differenze sopra menzionate sono significative: in zone di Δ_{Lc} comparabile si nota un eccesso di eventi misurati con il costrutto contenente solo GB1 a forze più elevate e un eccesso di eventi misurati con il costrutto contenente MoPrP a forze meno elevate.

Sono presenti nella distribuzione di $GB1_4PrPTrGB1_4$ eventi sporadici a lunghezze maggiori di 30 nm.

Concentrandosi sugli istogrammi monodimensionali riportati in figura 8.2 si nota come, per quanto riguarda le Δ_{Lc} , le curve siano generalmente ben sovrapposte per lunghezze inferiori a 20 nm, per lunghezze maggiori le distribuzioni tendono invece a distanziarsi in modo significativo.

Per quanto riguarda l'istogramma delle forze è visibile come l'altezza dell'istogramma relativo al costrutto $GB1_4PrPTrGB1_4$ tenda ad essere significativamente maggiore di quella di $GB1_8$ per eventi di forza compresa tra 110 e 200 pN, questi eventi sono associabili agli eventi di lunghezza maggiore di 20 nm, infatti, osservando le differenze per gli istogrammi bidimensionali riportati in figura 8.1 (in basso a sinistra) si nota come tutti i punti viola di lunghezza maggiore a 20 nm avvengano in corrispondenza di forze inferiori a 300 pN.

Essendo la lunghezza teorica di MoPrP uguale a 39.7 nm, la presenza di eventi sporadici con Δ_{Lc} superiore a 40 nm è riconducibile a imprecisioni sperimentali, ad esempio erranea selezione delle curve contenenti eventi significativi.

L'interpretazione strutturale

La lunghezza del dominio C-terminale di MoPrP (riportato in figura 8.3) può essere calcolata in base al numero di amminoacidi che contiene. Da tale valore bisogna escludere gli amminoacidi compresi tra le cisteine 178 e 213 del ponte disolfuro, in quanto un legame covalente non potrebbe rompersi se non a forze superiori a 1000 pN, nascondendo così tali residui durante l'estensione della parte strutturata. Si ottiene dunque un valore di lunghezza teorico di 22.48 nm. L'*unfolding* del dominio C-terminale, nella struttura nativa, sarebbe, quindi, compatibile con gli eventi visualizzati in corrispondenza della coda destra della distribuzione delle Δ_{Lc} di GB1.

Eventi con Δ_{Lc} di lunghezza maggiore della lunghezza del dominio C-terminale e minore della lunghezza di MoPrPTr sono associabili a conformazioni in cui i domini C-terminale ed N-terminale sono entrambi strutturati, almeno parzialmente.

Esperimenti con proteina rifoldata in condizioni riducenti

Allo scopo di verificare la corretta associazione dei dati presenti attorno a 22 nm con la struttura nativa di MoPrP, sono stati effettuati esperimenti ripiegando il monomero in condizioni riducenti, in modo da essere certi che non si formi il ponte disolfuro tra le cisteine 178 e 213. L'assenza di tale ponte può avere due effetti sul risultato degli esperimenti:

- spostamento della distribuzione degli eventi della transizione nativa ad una lunghezza maggiore rispetto a quella osservata in condizioni riducenti. L'allungamento dovrebbe essere pari alla lunghezza del *loop* racchiuso dal ponte disolfuro, ovvero 45 a.a., corrispondenti a circa 12 nm. Questo risultato sarebbe in accordo con [32].
- Impossibilità di ripiegamento del dominio C-terminale nella struttura nativa con conseguente assenza di segnale ad essa associato. Tale risultato sarebbe in accordo con [22].

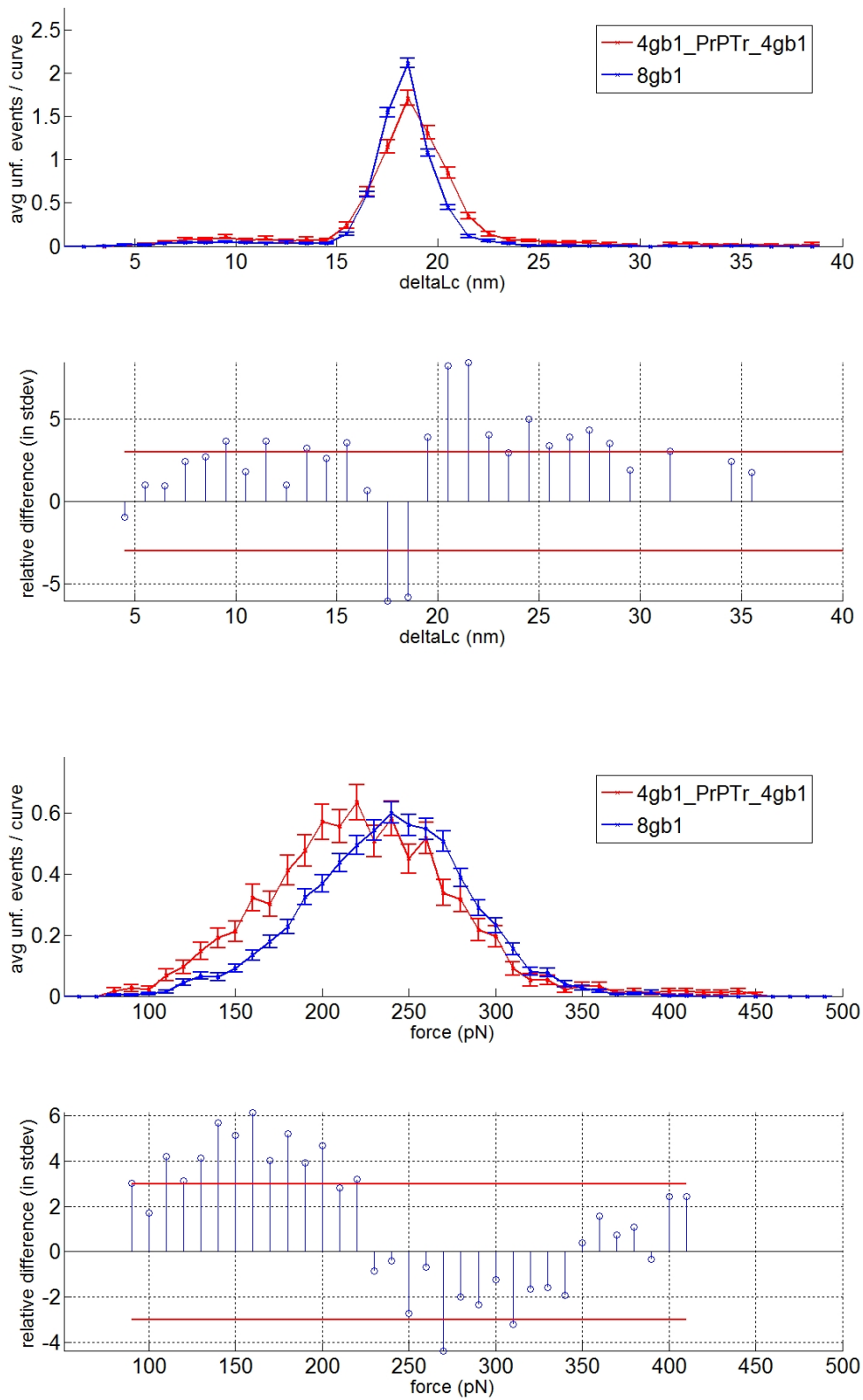


Figura 8.2: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB1_4PrPTrGB1_4$ e di $GB1_8$. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_4PrPTrGB1_4$ e di $GB1_8$. Il metodo utilizzato è descritto in sezione 6.2.1 .

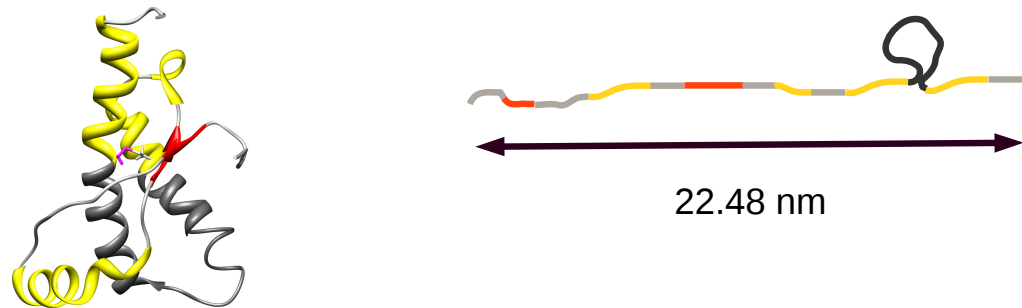


Figura 8.3: SINISTRA: rappresentazione della struttura del dominio C-terminale di MoPrP. DESTRA: il dominio C-terminale dopo la denaturazione meccanica; i colori hanno lo scopo di riprodurre solo qualitativamente le sezioni di catena amminoacidica. La parte grigia scura corrisponde al *loop* chiuso dal ponte disolfuro che si forma tra le cisteine 178 e 213 e che non viene rotto nel corso dell'*unfolding* meccanico. Le sezioni di eliche *alpha* interne al *loop* rimangono inalterate dopo la denaturazione e non sono state raffigurate per semplicità di rappresentazione.

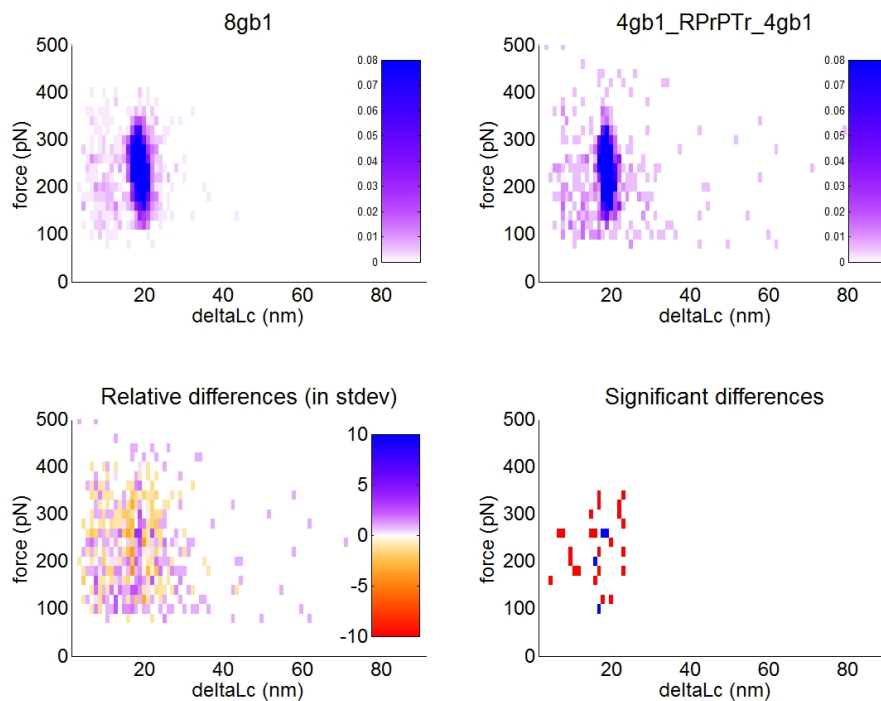


Figura 8.4: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4RPrPTrGB1_4$ e di $GB1_8$ (monomero ripiegato in condizioni riducenti); il metodo utilizzato è descritto in sezione 6.2.2.

Gli istogrammi riportati nella prima riga di figura 8.4, mostrano come il costrutto contenente un monomero di MoPrP ripiegato in condizioni riducenti sia più simile al costrutto contenente solo GB1, rispetto al costrutto monomero rifoldato in condizioni ossidanti (riportato in figura 8.1).

Sempre in figura 8.4, in basso a sinistra, si nota come non siano presenti zone caratterizzate da colori accesi (e quindi molto positive o molto negative), ad indicare che non esistono zone continue in cui i due costrutti sono distinguibili, infatti nel grafico in basso a destra il numero di punti blu è molto inferiore a quello ottenuto nel caso del monomero rifoldato in condizioni ossidanti.

Continuano, comunque, ad essere presenti eventi di lunghezza associabile a conformazioni comprendenti sia il dominio C-terminale che il dominio N-terminale almeno parzialmente strutturati. La presenza di queste conformazioni è scorrelata dalla stabilità del ripiegamento nativo del dominio C-terminale.

In figura 8.5 è visibile come non ci siano più differenze significative attorno a Δ_{Lc} di circa 22 nm, inoltre non si vede nemmeno una transizione della distribuzione di questi eventi a lunghezze maggiori. Le lunghezze misurate per il costrutto contenente MoPrP in condizioni riducenti e per il costrutto $GB1_8$ sono estremamente simili.

Anche la distribuzione delle forze, risulta perfettamente compatibile con quella degli eventi di *unfolding* di GB1.

Questi risultati suggeriscono come in assenza di ponte disolfuro il dominio C-terminale non sia in grado di raggiungere una conformazione stabile, analogamente a quanto discusso in [22].

8.1.2 Dimeri: il ruolo dell'orientazione reciproca di MoPrP.

Orientazione H-T

Dal confronto degli istogrammi bidimensionali riportati in figura 8.6 si nota immediatamente come siano presenti molti più eventi nella distribuzione di $GB1_4PrPTR_2HTGB1_4$ rispetto a quella di $GB1_8$.

Sono presenti eventi con $39 \text{ nm} < \Delta_{Lc} < 78 \text{ nm}$, tali eventi sono certamente associabili ad eventi associativi del dimero, in quanto hanno una lunghezza maggiore della lunghezza del monomero completamente esteso. Il fatto che tra essi ce ne siano di più lunghi di 61 nm significa che in queste conformazioni associative sono coinvolti (e strutturati) anche i domini N-terminali. Questo dato suggerirebbe per la prima volta che la porzione non strutturata di MoPrP in realtà potrebbe ripiegarsi fino a formare una struttura stabile durante le prime fasi dell'aggregazione.

Si nota inoltre come in questo costrutto ad un aumento di Δ_{Lc} la forza di *unfolding* massima raggiunta tenda a diminuire, suggerendo una bassa stabilità meccanica delle strutture associative.

Osservando l'istogramma delle differenze si nota la presenza di due regioni scure: una regione positiva centrata a circa 22 nm e 180 pN e una regione negativa centrata a 18 nm e 280 pN. Tali differenze sono anche indicate come significative. Dall'osservazione degli istogrammi monodimensionali riportati in figura 8.7 si nota come la presenza di eventi in corrispondenza alla coda destra della distribuzione di Δ_{Lc} del picco di GB1 sia maggiormente visibile rispetto al caso del monomero (figura 8.2).

Inoltre aumenta la differenza nella coda sinistra della distribuzione delle forze del picco di GB1. Tali differenze sono associabili al fatto che, come precedentemente descritto e visibile dagli istogrammi bidimensionali, eventi con Δ_{Lc} maggiore hanno forza media minore.

Queste evidenze sperimentali suggeriscono che anche in presenza di due monomeri ravvicinati in geometria H-T il dominio C-terminale possa adottare una struttura nativa in modo stabile e indipendente.

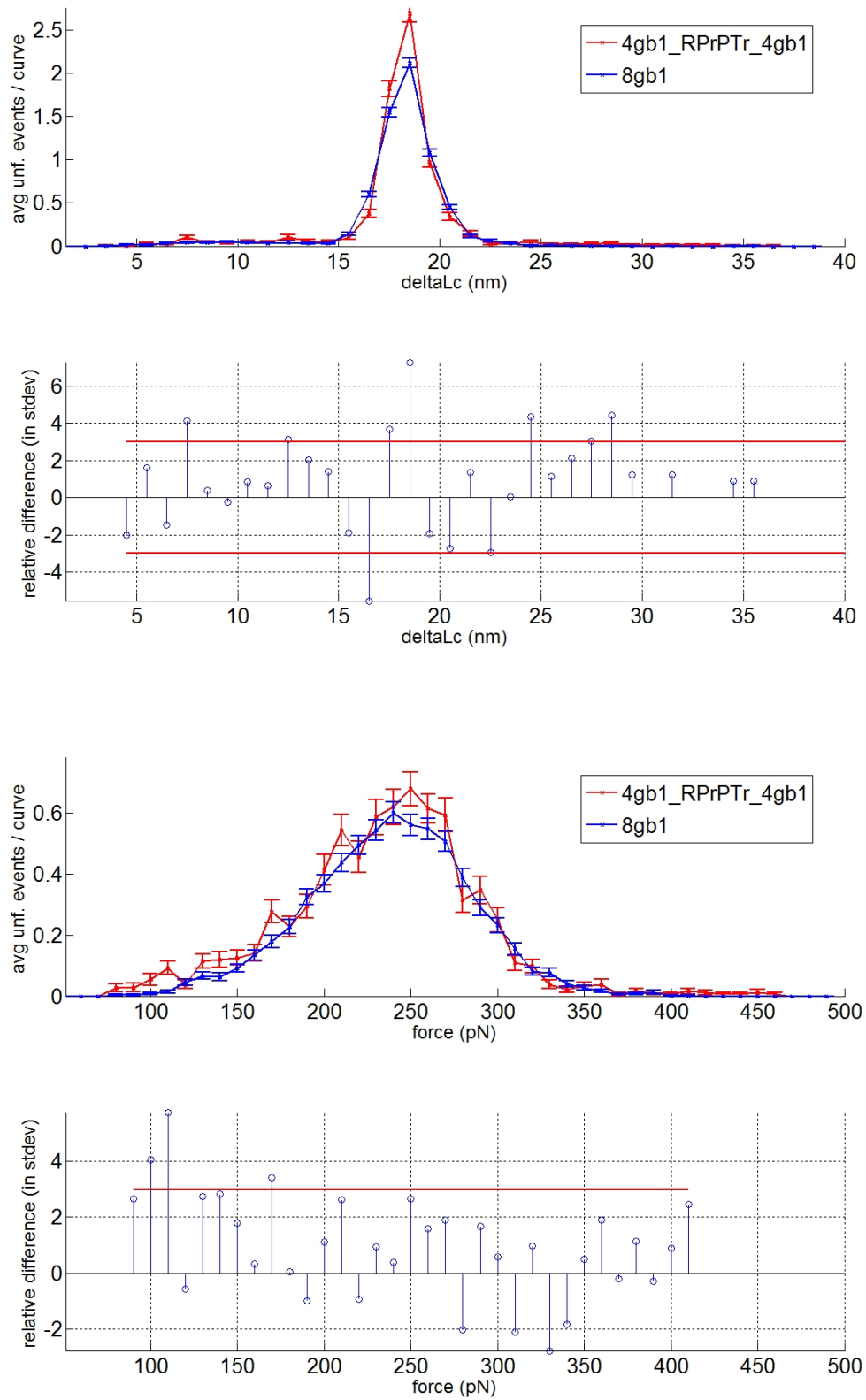


Figura 8.5: IN ALTO: confronto tra gli istogrammi delle ΔL_c di $GB1_4RPrPTrGB1_4$ (monomero ripiegato in condizioni riducenti) e di $GB1_8$. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_4RPrPTrGB1_4$ e di $GB1_8$.

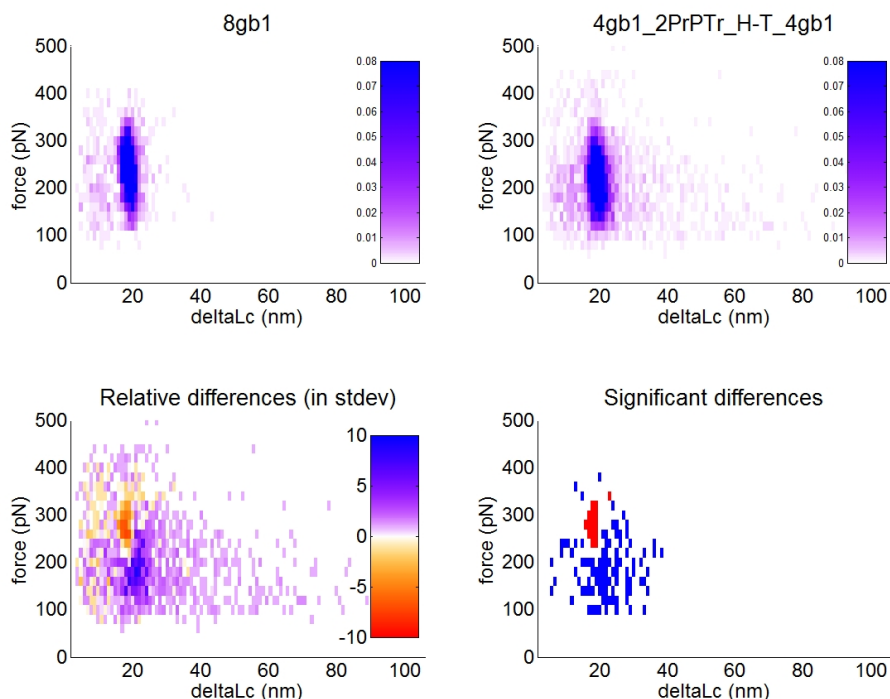


Figura 8.6: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTr_2HTGB1_4$ e di $GB1_8$; il metodo utilizzato è descritto in sezione 6.2.2

Orientazione H-H

In figura 8.8 è riportato il confronto tra gli istogrammi bidimensionali di $GB1_4PrPTr_2HHGB1_4$ e $GB1_8$. Esistono differenze tra i due costrutti, ma le differenze tra questi due costrutti sono molto minori di quelle osservate nel caso del dimero in orientazione H-T (figura 8.6).

In particolare, se si osservano gli istogrammi monodimensionali riportati in figura 8.9 si nota come le forze per $GB1_4PrPTr_2HHGB1_4$ e $GB1_8$ siano molto simili, mentre le Δ_{Lc} siano significativamente diverse solo al di sopra di 25 nm.

Questo significa che non c'è sovrapposizione significativa tra gli eventi di GB1 e quelli del dimero, invece sono presenti eventi più lunghi del C-terminale.

L'orientazione H-H, dunque, non è compatibile con il *folding* nativo del dominio C-terminale.

8.1.3 Trimero: presenza di struttura nativa e conformazioni associative.

In figura 8.10 è riportato il confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTr_3GB1_4$ e di $GB1_8$. È nuovamente ben visibile un'ampia regione viola nell'istogramma in basso a sinistra per Δ_{Lc} comprese tra 18 e 35 nm e forza inferiore di 300 pN. La regione corrispondente a Δ_{Lc} di circa 18 nm e forze di circa 300 pN è gialla, ovvero in questa regione l'altezza dell'istogramma di $GB1_8$ è maggiore di quella dell'istogramma di $GB1_4PrPTr_3GB1_4$. Queste due regioni sono anche indicate come significativamente differenti.

Non sono presenti eventi con Δ_{Lc} maggiore di 80 nm, non esistono dunque eventi di *unfolding* associabili con certezza ad un'associazione tra tre monomeri che si denatura in un'unica transizione. Ciò nonostante non è possibile escludere che transizioni con Δ_{Lc} compresa tra 41 e 80 nm possano essere dovute ad associazione di tre monomeri, in quanto queste lunghezze

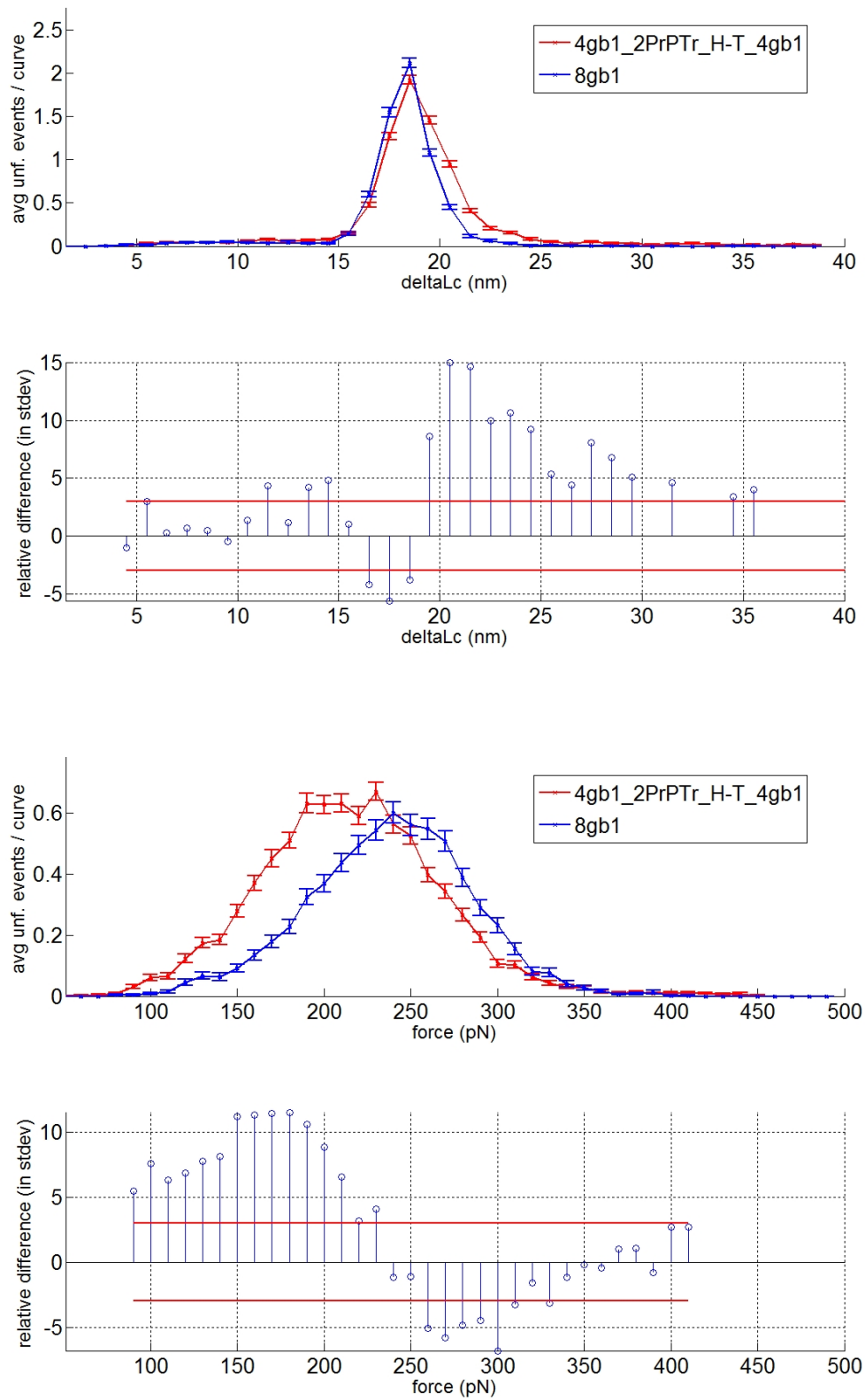


Figura 8.7: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB1_4PrPTr_2HTGB1_4$ e di $GB1_8$. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_4PrPTr_2HTGB1_4$ e di $GB1_8$. Il metodo utilizzato è descritto in sezione 6.2.1

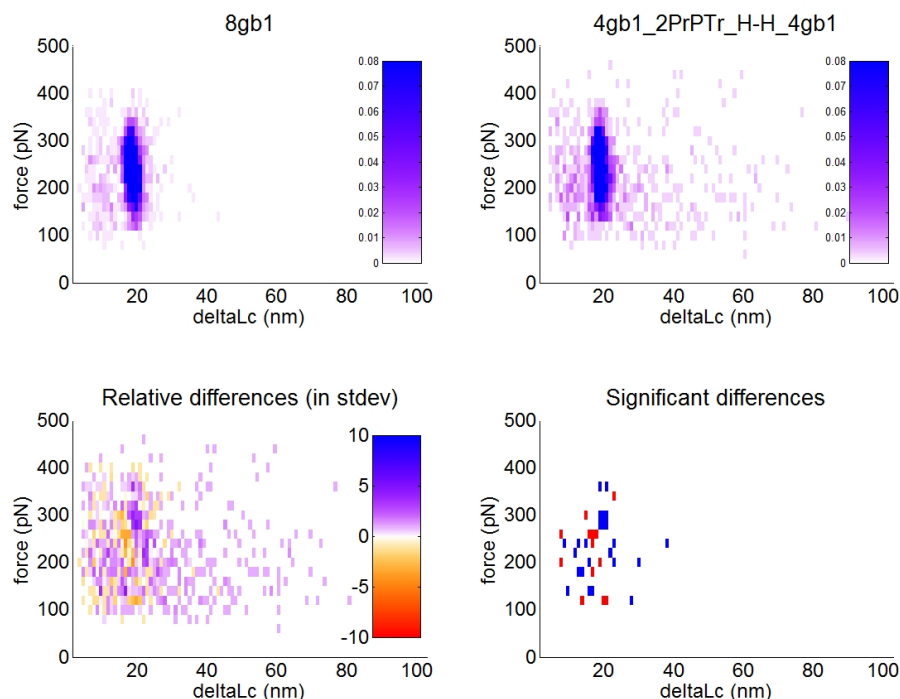


Figura 8.8: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTr_2HHGB1_4$ e di $GB1_8$; il metodo utilizzato è descritto in sezione 6.2.2

sono compatibili non solo con i dimeri, ma anche con strutture comprendenti la molecola centrale di MoPrP e i primi amminoacidi delle due molecole che la fiancheggiano.

Osservando gli istogrammi monodimensionali riportati in figura 8.11 si nota come la spalla destra sia molto evidente nella distribuzione di GB1. La maggior visibilità di questa spalla è associabile al maggior numero di monomeri presenti in questo costrutto che continuano ad avere un dominio C-terminale strutturato nella conformazione nativa.

Anche l'istogramma delle forze mette in evidenza, nuovamente, come per questo costrutto aumentino gli eventi a forza bassa. Come visibile dagli istogrammi bidimensionali, per gli eventi associativi un aumento della lunghezza del dominio sfoldato corrisponde ad una diminuzione della forza media di *unfolding*.

8.1.4 Tetramero: identificazione di nuove strutture.

In figura 8.12 è riportato il confronto tra gli istogrammi bidimensionali di $GB1_2PrPTr_4GB1_2$ e $GB1_8$ (filtrato in modo da avere un numero di eventi di *unfolding* compreso tra tre e quattro, per essere confrontabile con il costrutto contenente il tetramero). Si può osservare come gli eventi misurati in corrispondenza del massimo del picco principale di GB1 siano di meno nel caso del costrutto con MoPrP (regione giallo/rossa nei grafici delle differenze e delle differenze significative), al contrario il numero di eventi misurati per il costrutto contenente il tetramero risulta essere maggiore per lunghezze appena superiori e per forze minori.

Nella regione di Δ_{Lc} compresa tra 40 e 55 nm e forza compresa tra 100 e 300 pN sono visibili una serie di eventi che potrebbero rappresentare un *cluster*, in questo caso sarebbero associabili ad una transizione tipica del tetramero.

Sono visibili eventi di lunghezza maggiore di 80 nm: non essendo presenti nelle curve misurate con il trimero essi potrebbero essere associati a transizioni del tetramero. Vista la

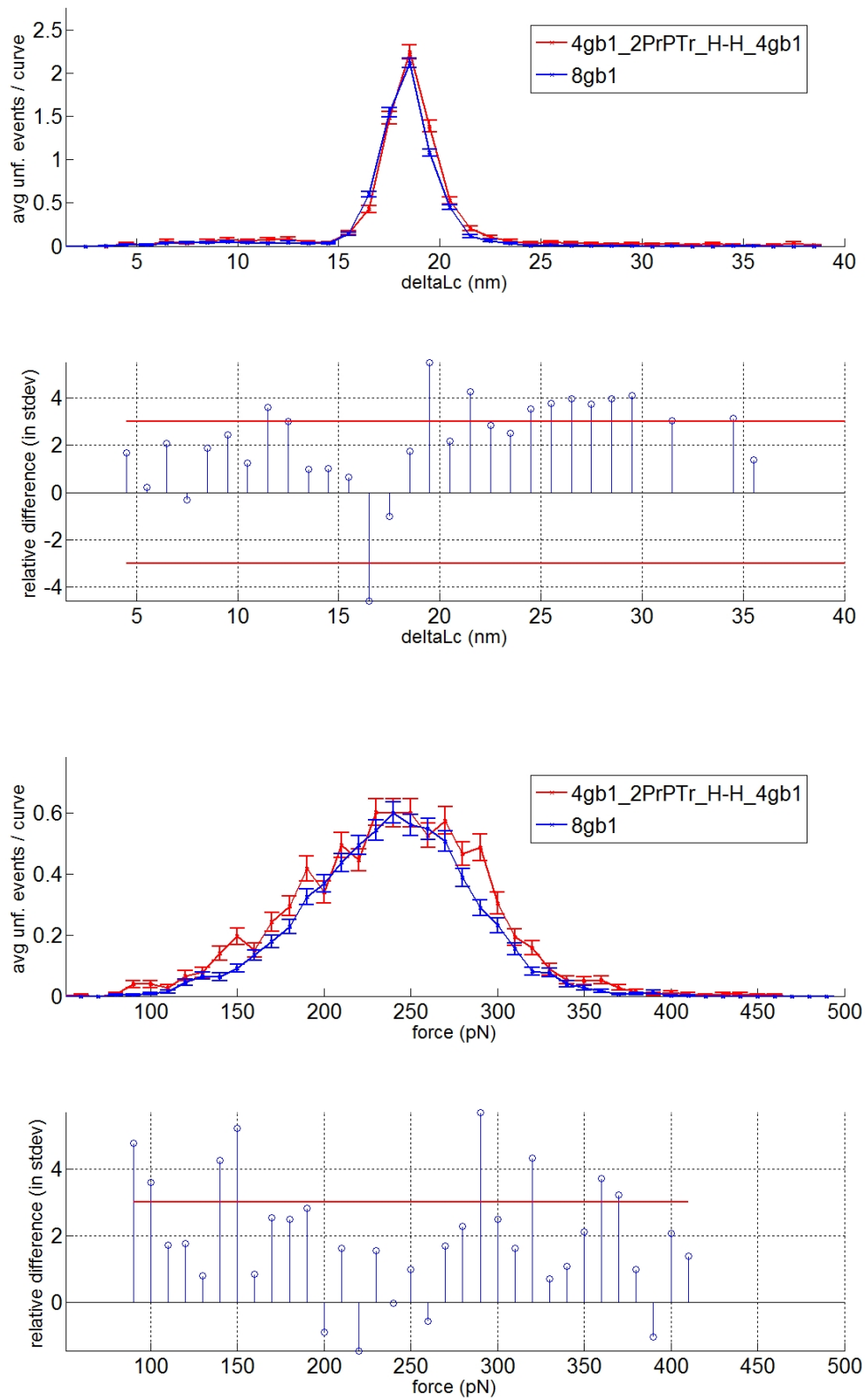


Figura 8.9: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB_{14}PrPT_r_2HHGB_{14}$ e di GB_{18} . IN BASSO: confronto tra gli istogrammi delle forze di $GB_{14}PrPT_r_2HHGB_{14}$ e di GB_{18} . Il metodo utilizzato è descritto in sezione 6.2.1

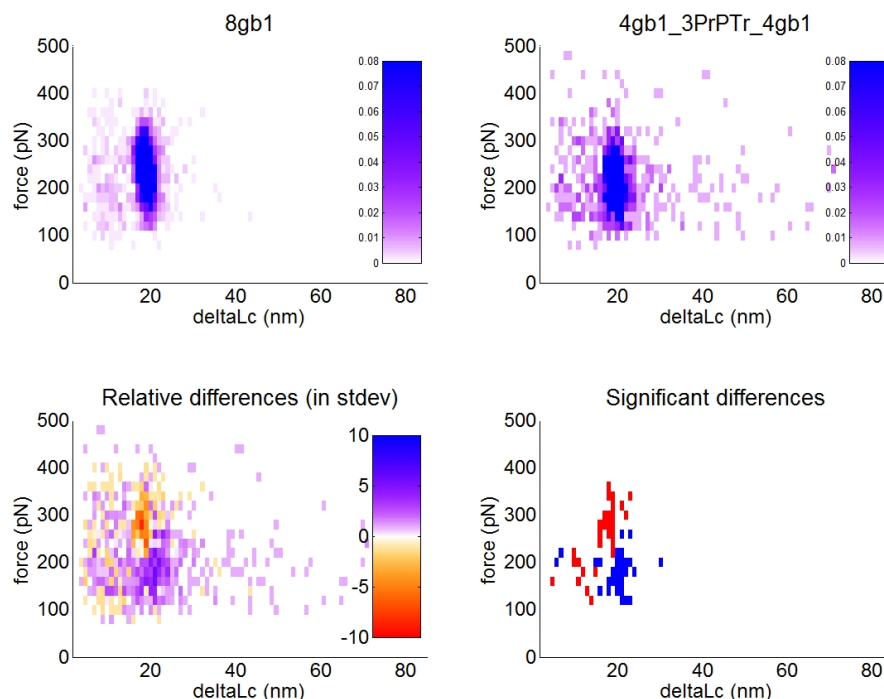


Figura 8.10: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTR_3GB1_4$ e di $GB1_8$; il metodo utilizzato è descritto in sezione 6.2.2

loro lunghezza inferiore a 116 nm non è comunque possibile associare con certezza queste transizioni a conformazioni del tetramero.

Osservando il confronto tra istogrammi monodimensionali riportato in figura 8.13 si nota, per quanto riguarda le Δ_{Lc} , uno spostamento verso destra del picco principale. Questo fenomeno è dovuto al fatto che, come descritto in sezione 8.1.1, nella fase di selezione delle curve non è stato possibile discriminare i picchi di GB1 dai picchi di MoPrP caratterizzati da lunghezze di Δ_{Lc} confrontabili. L'effetto è maggiormente visibile in questo costrutto rispetto agli altri perché in questo caso sono presenti solo due GB1 ad ogni estremo del tetramero e il rapporto tra molecole di GB1 e di MoPrP presenti è pari a uno (negli altri casi era sempre maggiore di due).

Sempre nell'istogramma delle Δ_{Lc} è evidente una distribuzione di eventi centrata su circa 10 nm.

Per quanto riguarda le forze si nota un effetto sempre maggiore della presenza di moduli di MoPrP dall'aumento del numero di eventi a forza più bassa rispetto a quella media dell'*unfolding* di GB1.

8.1.5 Fit preliminare della distribuzione sulla coda destra di GB1

Al fine di caratterizzare completamente la distribuzione della distribuzione di eventi associabili alla transizione nativa del dominio C-terminale sarà necessario sviluppare strategie di fit delle distribuzioni bidimensionali.

Per il momento è stato effettuato un fit monodimensionale sulla distribuzione delle Δ_{Lc} che, sebbene non tenga conto della correlazione tra lunghezza e forza, può essere di aiuto nella caratterizzazione dei costrutti.

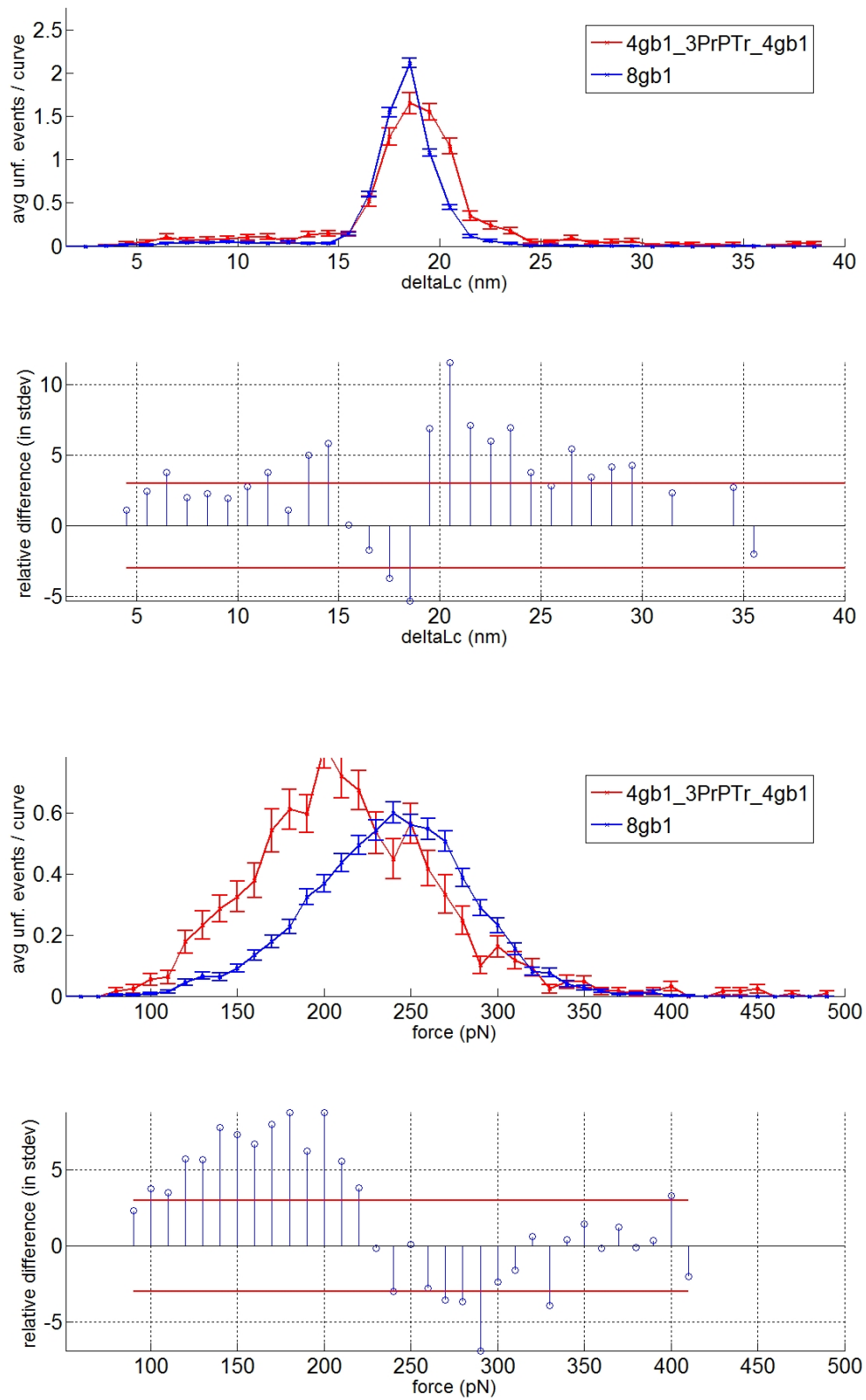


Figura 8.11: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB1_4PrPTR_3GB1_4$ e di $GB1_8$. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_4PrPTR_3GB1_4$ e di $GB1_8$. Il metodo utilizzato è descritto in sezione 6.2.1

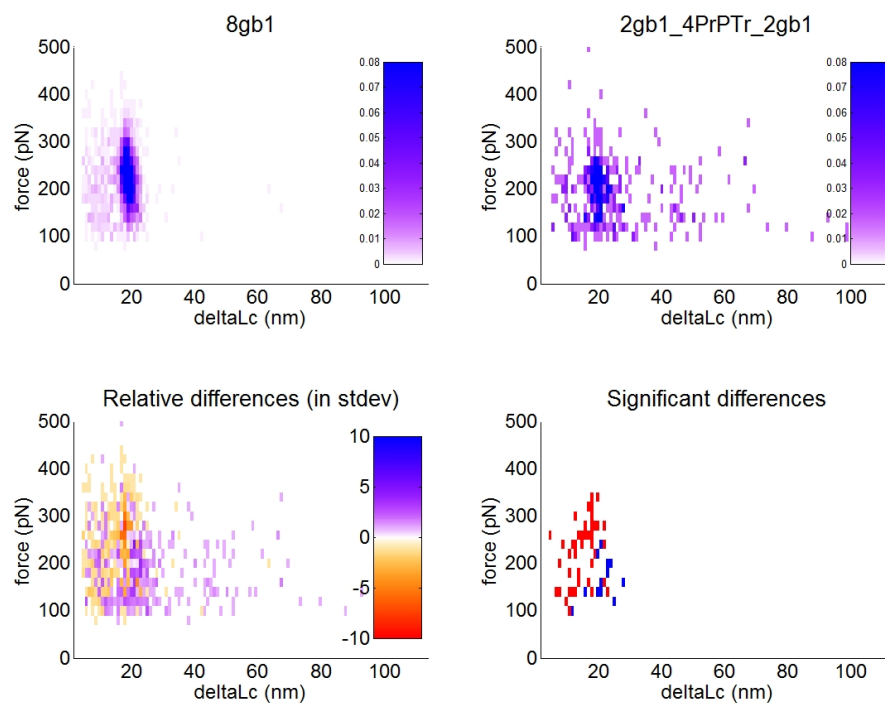


Figura 8.12: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_2PrPTr_4GB1_2$ e di $GB1_8$ (filtrato in modo da avere un numero di eventi di *unfolding* compreso tra 3 e 4); il metodo utilizzato è descritto in sezione 6.2.2

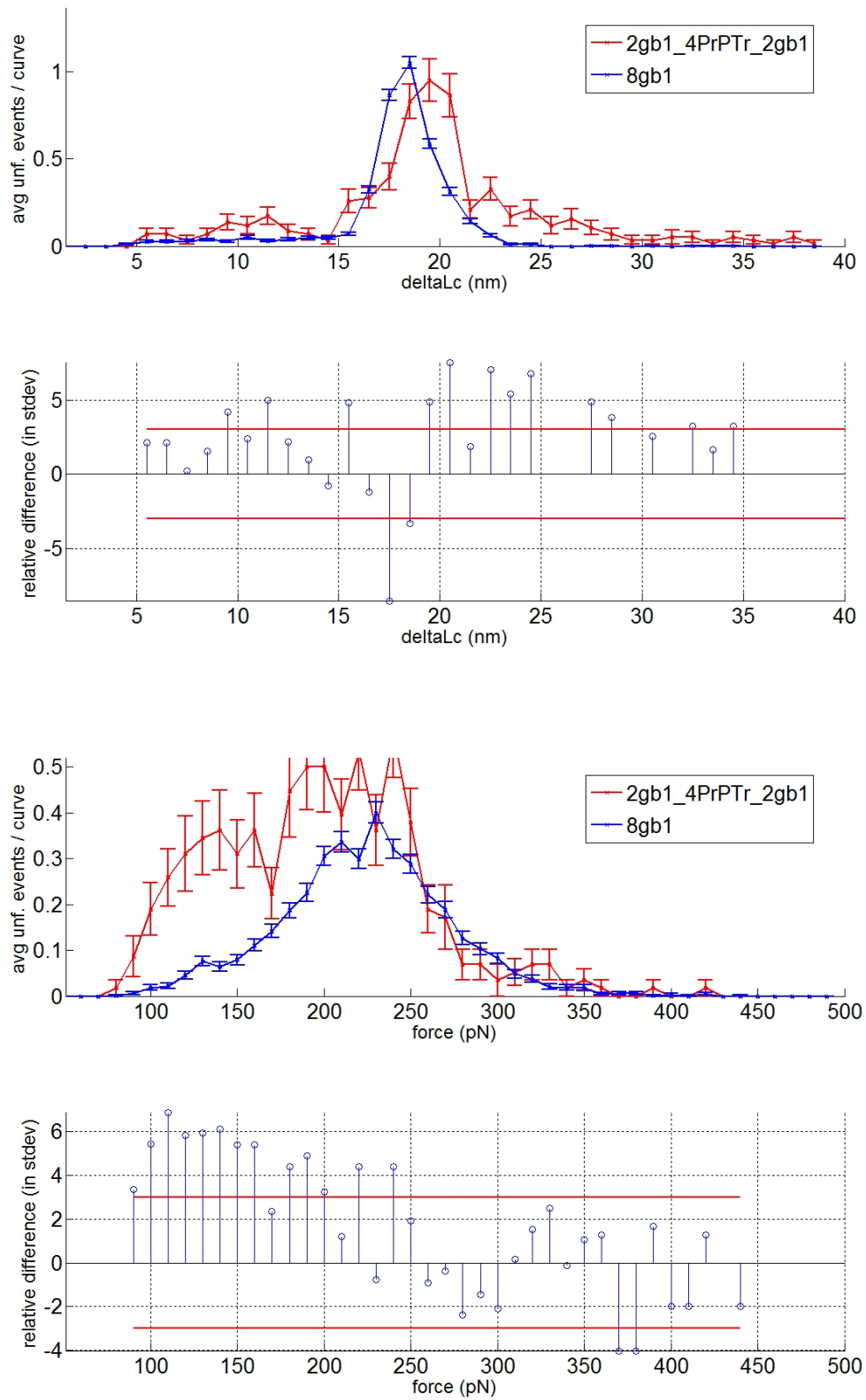


Figura 8.13: IN ALTO: confronto tra gli istogrammi delle ΔL_c di $GB_{12}PrPTr_4GB_{12}$ e di GB_{18} . IN BASSO: confronto tra gli istogrammi delle forze di $GB_{12}PrPTr_4GB_{12}$ e di GB_{18} . Il metodo utilizzato è descritto in sezione 6.2.1.

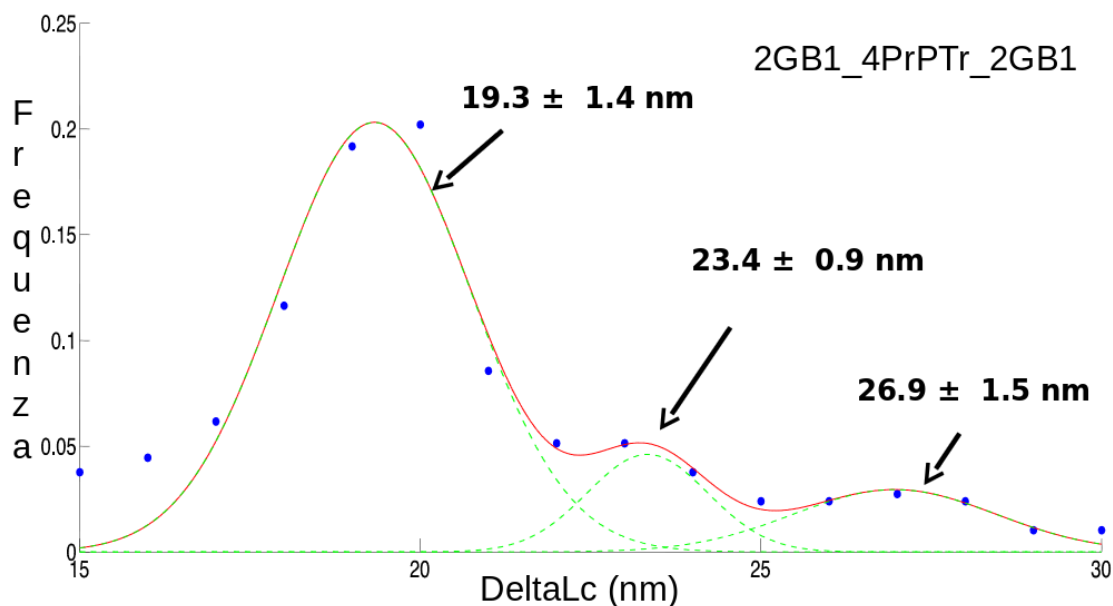


Figura 8.14: Fit ottenuto per la distribuzione di eventi del tetramero a pH 7.4, sono visibili tre distribuzioni. In blu è riportata l'altezza delle classi, in rosso il fit, in verde, le tre gaussiane ottenute. Sull'asse delle ordinate è leggibile la frequenza degli eventi. I numeri riportati corrispondono a μ e σ delle distribuzioni gaussiane fittate.

A livello bidimensionale è possibile vedere, per la maggioranza dei costrutti, un aumento di eventi a forze inferiori dentro la finestra di lunghezza scelta per GB1 a discapito di quelli a lunghezze superiori. A livello monodimensionale questa informazione viene persa, restituendo spesso, per la distribuzione di Δ_{LC} , un picco di altezza comparabile.

I fit presentati in seguito sono affetti, dunque, da questo errore sistematico e vanno considerati in modo puramente indicativo. Un metodo di fit bidimensionale verrà sviluppato in futuro per ovviare a questo problema.

Lo scopo della procedura di fit è ottenere informazioni sulla distribuzione a 22 nm; per questo motivo sono state considerate, nella procedura di fit, solo le classi comprese tra 15 e 30 nm.

La prima distribuzione ad essere stata fittata è quella del tetramero (figura 8.14) in quanto è quella che mostra un maggior numero di eventi corrispondenti alla transizione nativa del dominio C-terminale. Essa sarebbe anche la distribuzione maggiormente affetta dell'errore sistematico sopra discusso, a causa del rapporto tra numero di GB1 e di MoPrP presenti.

Il fit ottimale ottenuto è a tre distribuzioni. La gaussiana principale è centrata su 19.33 nm con una deviazione standard di 1.4 nm. Essa è dunque più larga e spostata a valori maggiori di quelli riportati per il fit di GB1 in tabella 6.1. Questo fenomeno è dovuto all'errore sistematico sopra descritto, ovvero questa gaussiana non è dovuta solo alla presenza di GB1, ma incorpora in sé anche alcuni dati relativi a MoPrP.

La seconda gaussiana è quella che tiene maggiormente in considerazione gli eventi associabili alla conformazione nativa del dominio C-terminale; è centrata su 23.4 nm con una deviazione standard di 0.9 nm, dunque è compatibile con la lunghezza teorica di questo dominio (circa 22.5 nm).

La terza gaussiana è centrata su 27 nm con una deviazione standard di 1.4 nm e può essere associata a conformazioni che coinvolgono parzialmente sia il dominio C-terminale che il dominio N-terminale.

Per costrutti di lunghezza inferiore il fit è stato effettuato con sole due gaussiane perché la terza distribuzione di eventi non è riconoscibile (figura 8.15).

Anche in questi tre costrutti la gaussiana di GB1 è più larga di quella riportata in tabella 1.6 (deviazione standard di 1.4 nm) ed è centrata su valori leggermente più alti (18.6 nm per monomero e dimeri, 18.7 per il trimero). La seconda gaussiana è stata posizionata su 22 nm con una deviazione standard di 1.4 nm.

8.2 Esperimenti a pH acidi

8.2.1 Monomero, condizioni ossidanti

Nelle figure 8.16 e 8.17 sono riportate le analisi degli istogrammi bidimensionali degli eventi di *unfolding* del monomero a pH acidi. Nelle figure 8.18 e 8.19 sono riportate le analisi degli istogrammi monodimensionali degli eventi di *unfolding* del monomero a pH acidi.

Ad entrambi i pH riportati si nota come, osservando gli istogrammi monodimensionali, non sia più riconoscibile l'*unfolding* della struttura nativa del dominio C-terminale, in altre parole si può affermare che l'abbassamento del pH destabilizza il *fold* di MoPrP.

Osservando gli istogrammi bidimensionali delle differenze (quadrante in basso a sinistra) si nota che esse non sono estremamente significative, anche se permane una certa divisione tra zone leggermente positive e leggermente negative.

Un accurato sistema di fit bidimensionale deve essere applicato per discriminare e caratterizzare questi eventi.

8.2.2 Dimero, orientazione H-T

Anche per il dimeri l'abbassamento del pH risulta in una perdita di stabilità del *fold* nativo, come visibile negli istogrammi bidimensionali (figure 8.20 e 8.21) e monodimensionali (figure 8.22 e 8.23). Anche in questo caso un metodo di fit bidimensionale potrebbe essere di aiuto per la discriminazione degli eventi.

8.2.3 Trimero

La destabilizzazione della conformazione nativa operata dall'abbassamento pH è meno evidente nel caso del trimero. Dall'analisi degli istogrammi bidimensionali ottenuti a pH 5.5 (figura 8.24) si nota la presenza di un'ampia regione con abbondanza di eventi del costrutto contenente MoPrP (centrata su 20 nm, con forze inferiori a 200 pN) ed una regione con carenza di eventi del costrutto con MoPrP (corrispondente al massimo del picco di GB1). La regione ricca di eventi corrisponde alla regione degli eventi associati alla denaturazione del dominio C-terminale, la carenza di eventi sul picco principale di GB1 è già stata discussa per i precedenti costrutti (sezione 8.1.1).

Dall'analisi degli istogrammi monodimensionali riportati in figura 8.26 si nota nuovamente l'innalzamento significativo della spalla destra nella distribuzione di GB1 in caso di presenza di MoPrP e lo spostamento della distribuzione delle forze a valori inferiori, tale spostamento è con buona probabilità associabile allo scambio di eventi di MoPrP con eventi di GB1 in fase di selezione delle curve.

In figura 8.25 è riportata l'analisi dell'istogramma bidimensionale misurato a pH 4.0.

Si nota come l'abbassamento ulteriore del pH abbia portato ad una ulteriore destabilizzazione della struttura nativa di MoPrP. Ciò nonostante è ancora visibile una piccola regione blu nel grafico delle differenze significative, ad indicare che la struttura nativa può ancora essere presente.

La destabilizzazione è ben evidente negli istogrammi monodimensionali riportati in figura 8.27, in cui la distribuzione a circa 22 nm è ancora presente (visibile nell'innalzamento signifi-

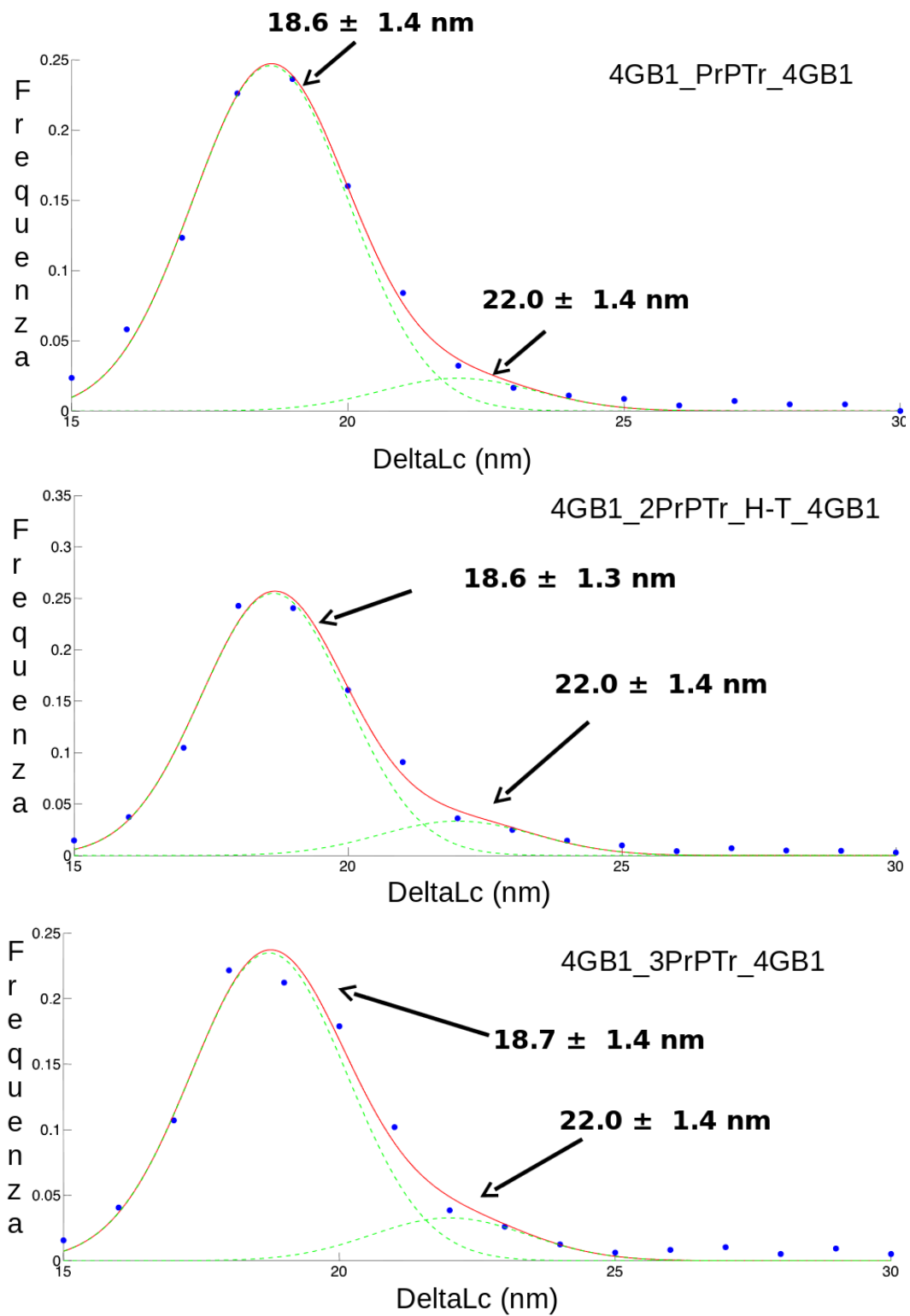


Figura 8.15: Fit ottenuto per la distribuzione di eventi, sono visibili tre distribuzioni. In blu è riportata l'altezza della classe, in rosso il fit, in verde, le tre gaussiane ottenute. I numeri riportati corrispondono a μ e σ delle distribuzioni fittate. IN ALTO: monomero, IN CENTRO: dimero H-T, IN BASSO: trimero. Sull'asse delle ordinate è leggibile la frequenza relativa degli eventi.

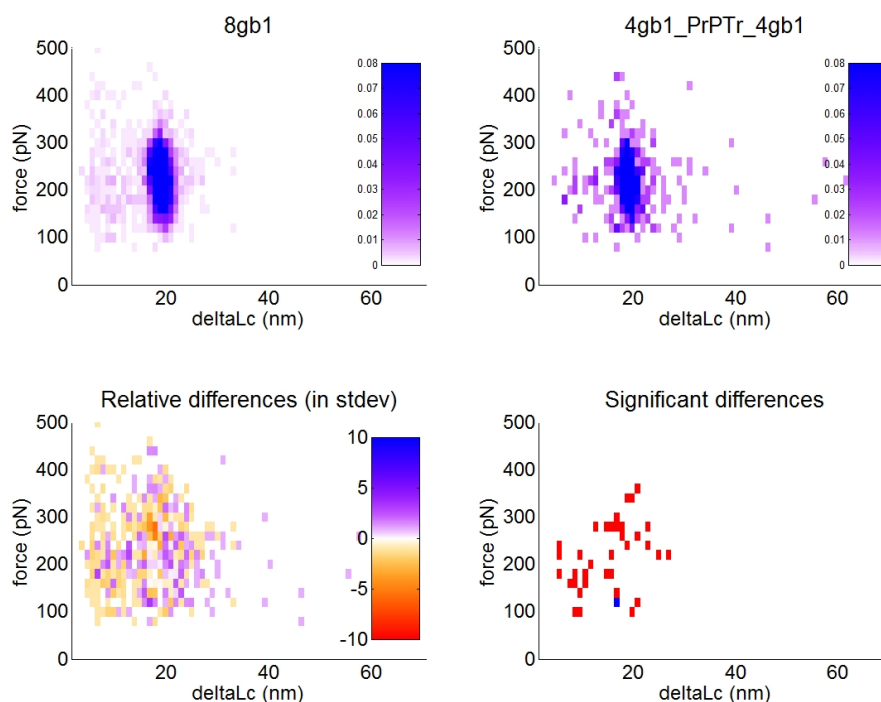


Figura 8.16: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTRGB1_4$ e di $GB1_8$, misurati a pH 5.5; il metodo utilizzato è descritto in sezione 6.2.2

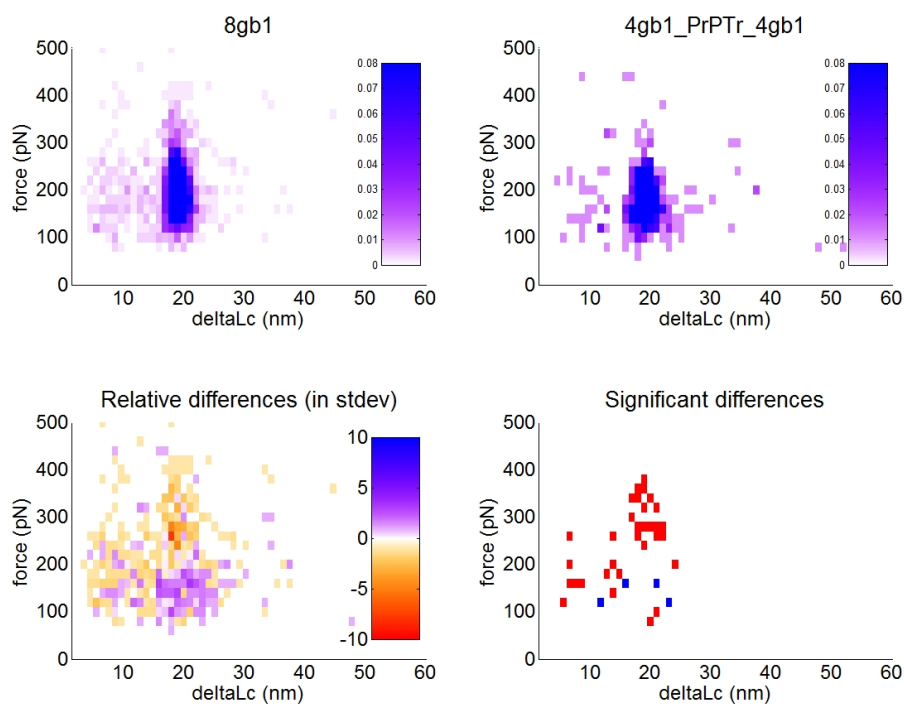


Figura 8.17: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTRGB1_4$ e di $GB1_8$, misurati a pH 4.0; il metodo utilizzato è descritto in sezione 6.2.2

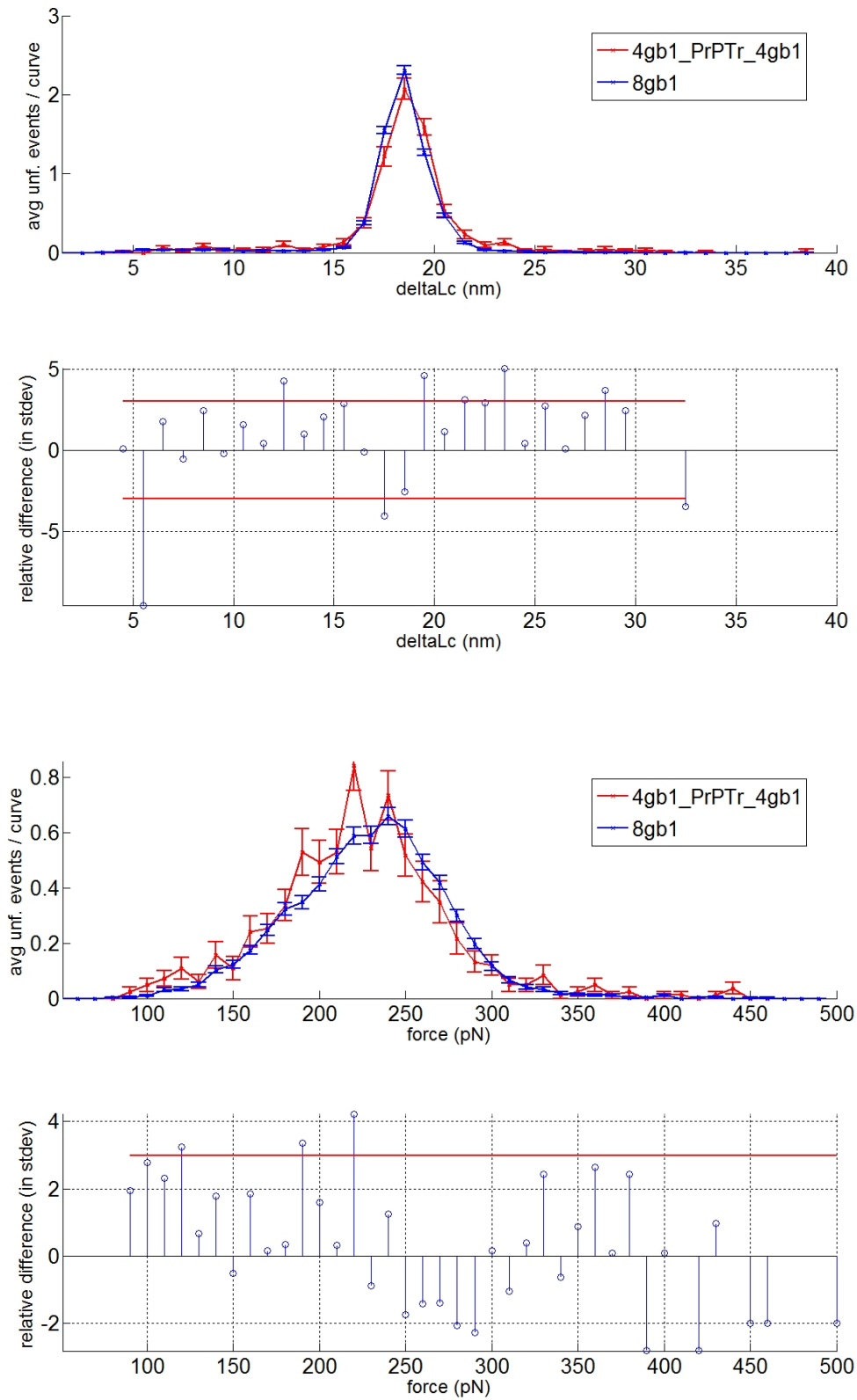


Figura 8.18: IN ALTO: confronto tra gli istogrammi delle ΔL_c di $GB1_4PrPTrGB1_4$ e di $GB1_8$, misurati a pH 5.5. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_4PrPTrGB1_4$ e di $GB1_8$, misurati a pH 5.5.

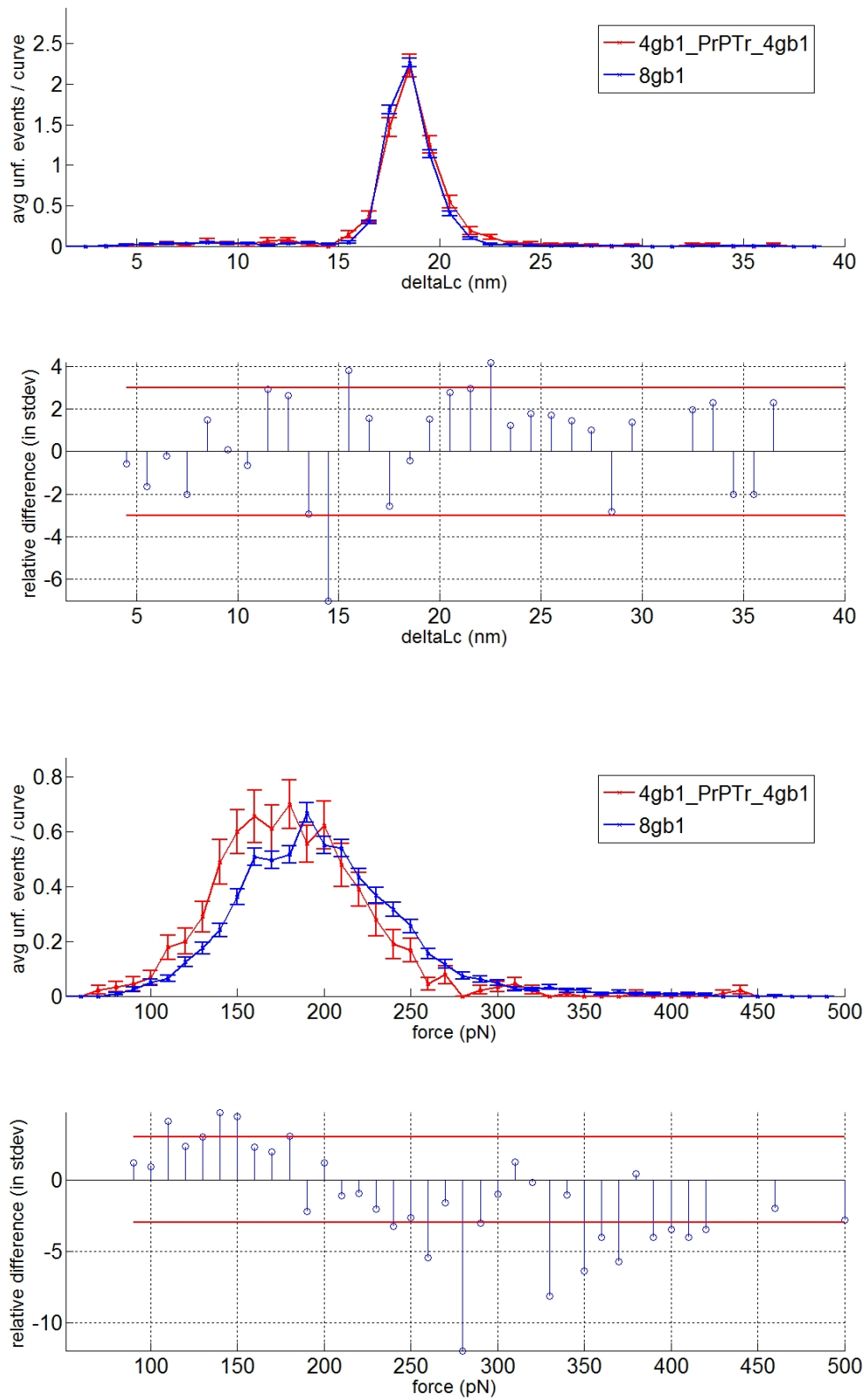


Figura 8.19: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB_{14}PrPTrGB_{14}$ e di GB_{18} , misurati a pH 4.0. IN BASSO: confronto tra gli istogrammi delle forze di $GB_{14}PrPTrGB_{14}$ e di GB_{18} , misurati a pH 4.0.

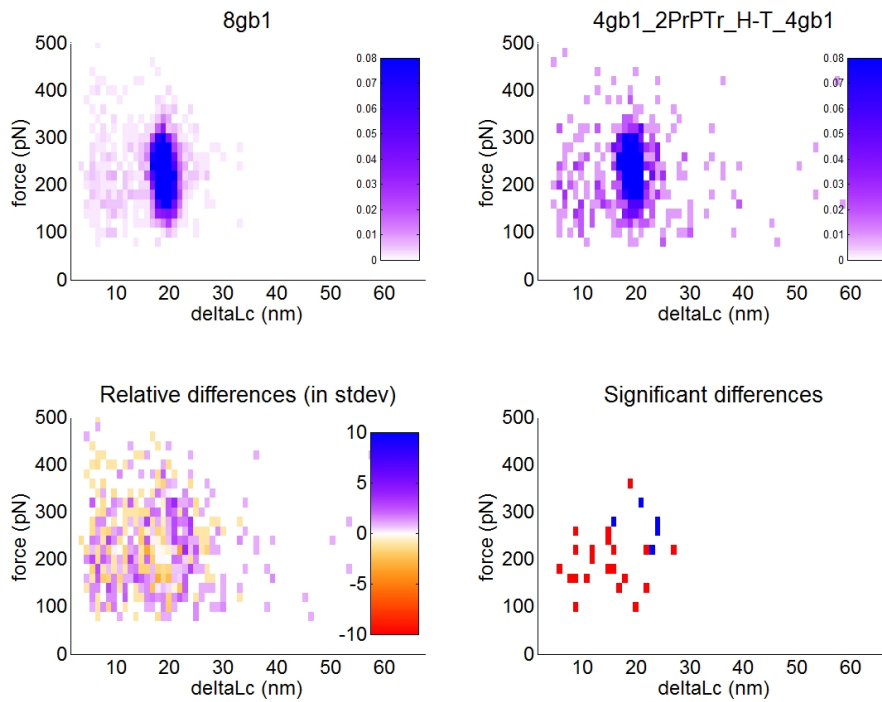


Figura 8.20: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPT_r2HTGB1_4$ e di $GB1_8$, misurati a pH 5.5; il metodo utilizzato è descritto in sezione 6.2.2

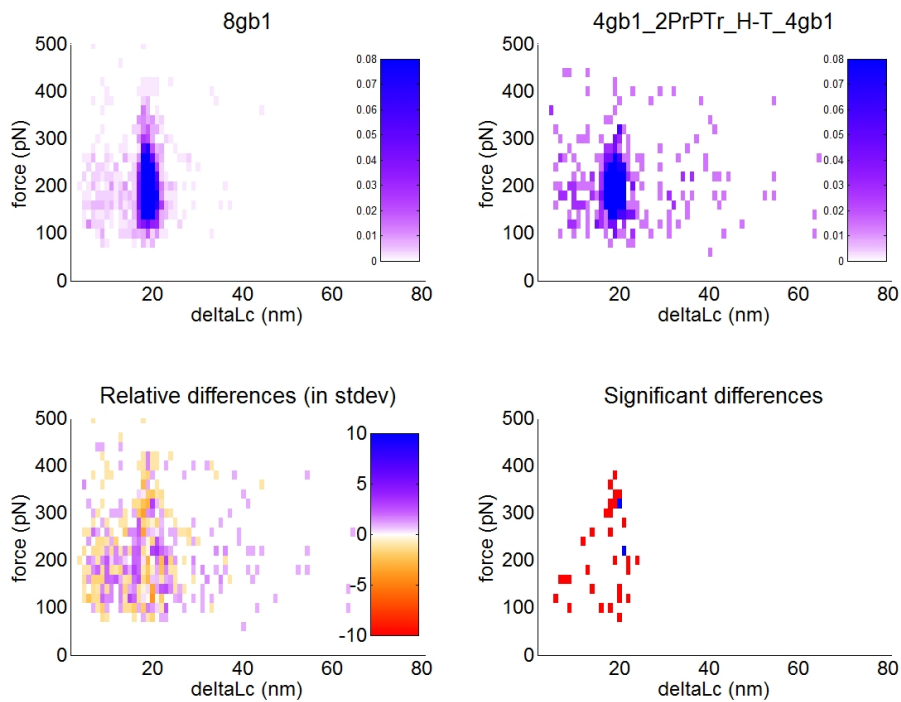


Figura 8.21: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPT_r2HTGB1_4$ e di $GB1_8$, misurati a pH 4.0; il metodo utilizzato è descritto in sezione 6.2.2

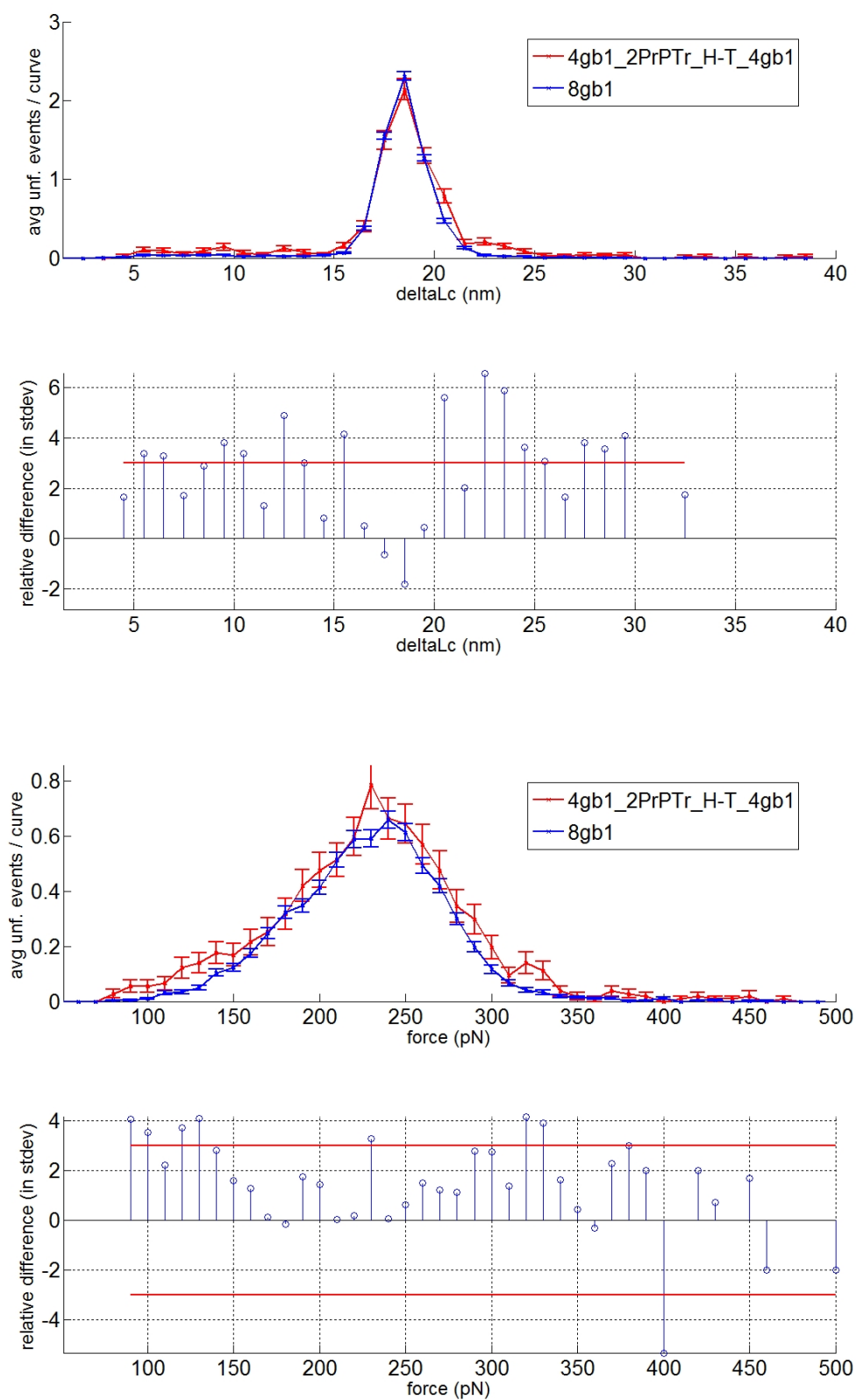


Figura 8.22: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB_{14}PrPTR_2HTGB_{14}$ e di GB_{18} , misurati a pH 5.5. IN BASSO: confronto tra gli istogrammi delle forze di $GB_{14}PrPTR_2HTGB_{14}$ e di GB_{18} , misurati a pH 5.5.

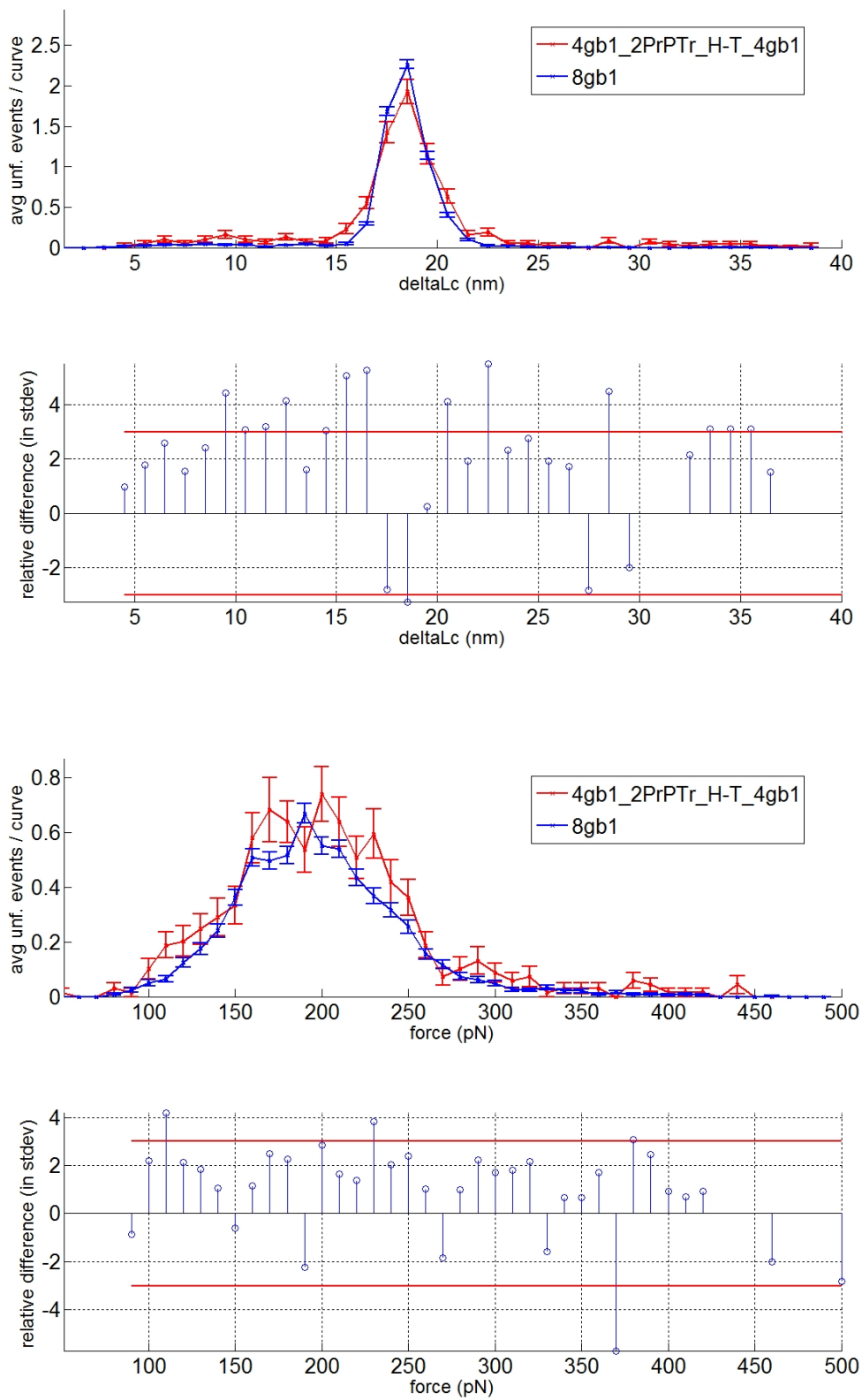


Figura 8.23: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB1_4PrPTr_2HTGB1_4$ e di $GB1_8$, misurati a pH 4.0. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_4PrPTr_2HTGB1_4$ e di $GB1_8$, misurati a pH 4.0.

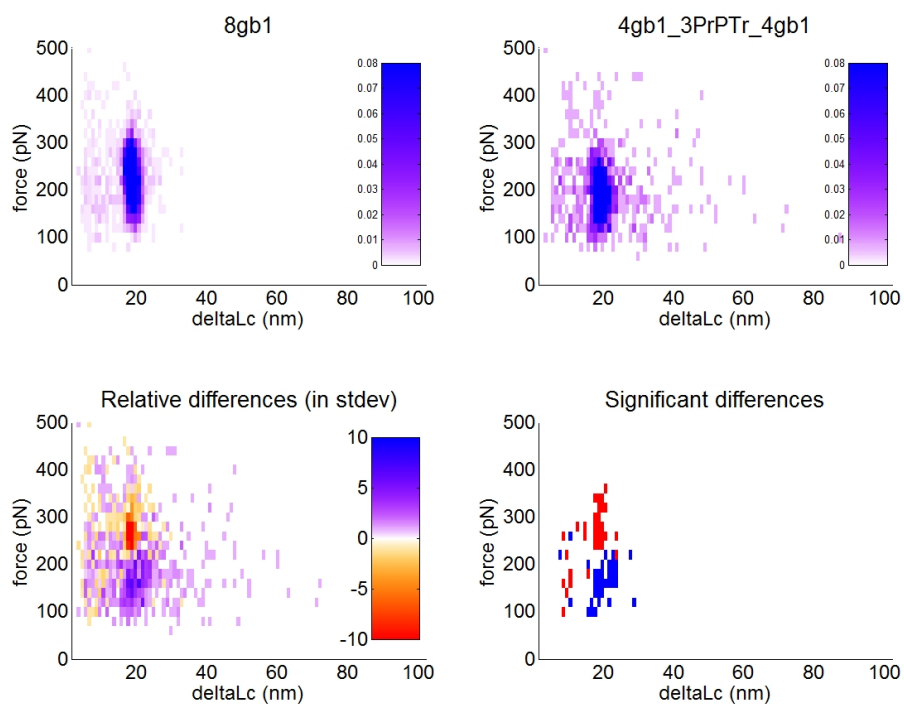


Figura 8.24: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTR_3GB1_4$ e di $GB1_8$, misurati a pH 5.5; il metodo utilizzato è descritto in sezione 6.2.2

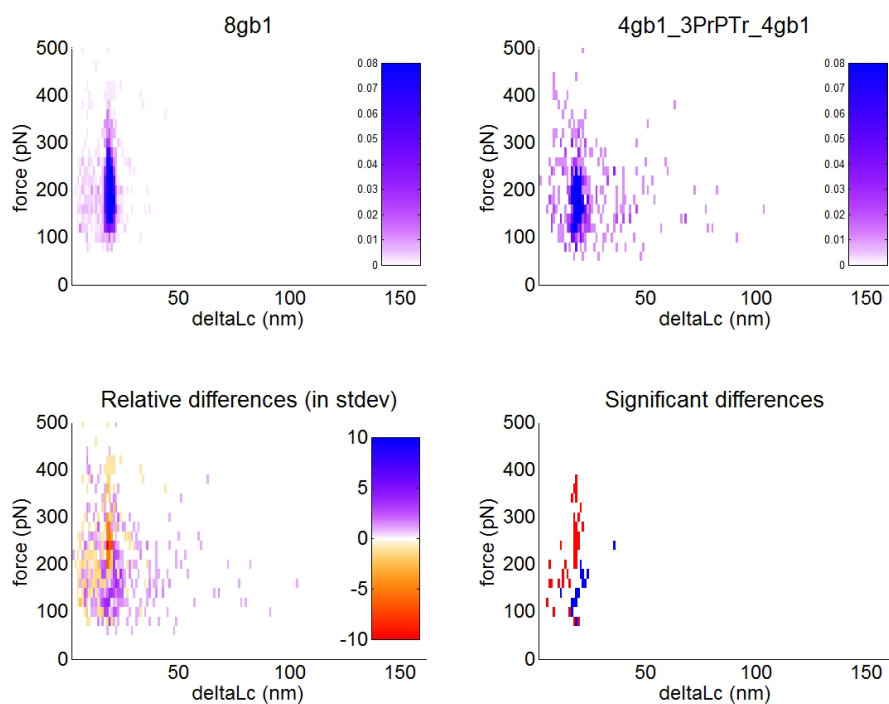


Figura 8.25: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_4PrPTR_3GB1_4$ e di $GB1_8$, misurati a pH 4.0; il metodo utilizzato è descritto in sezione 6.2.2

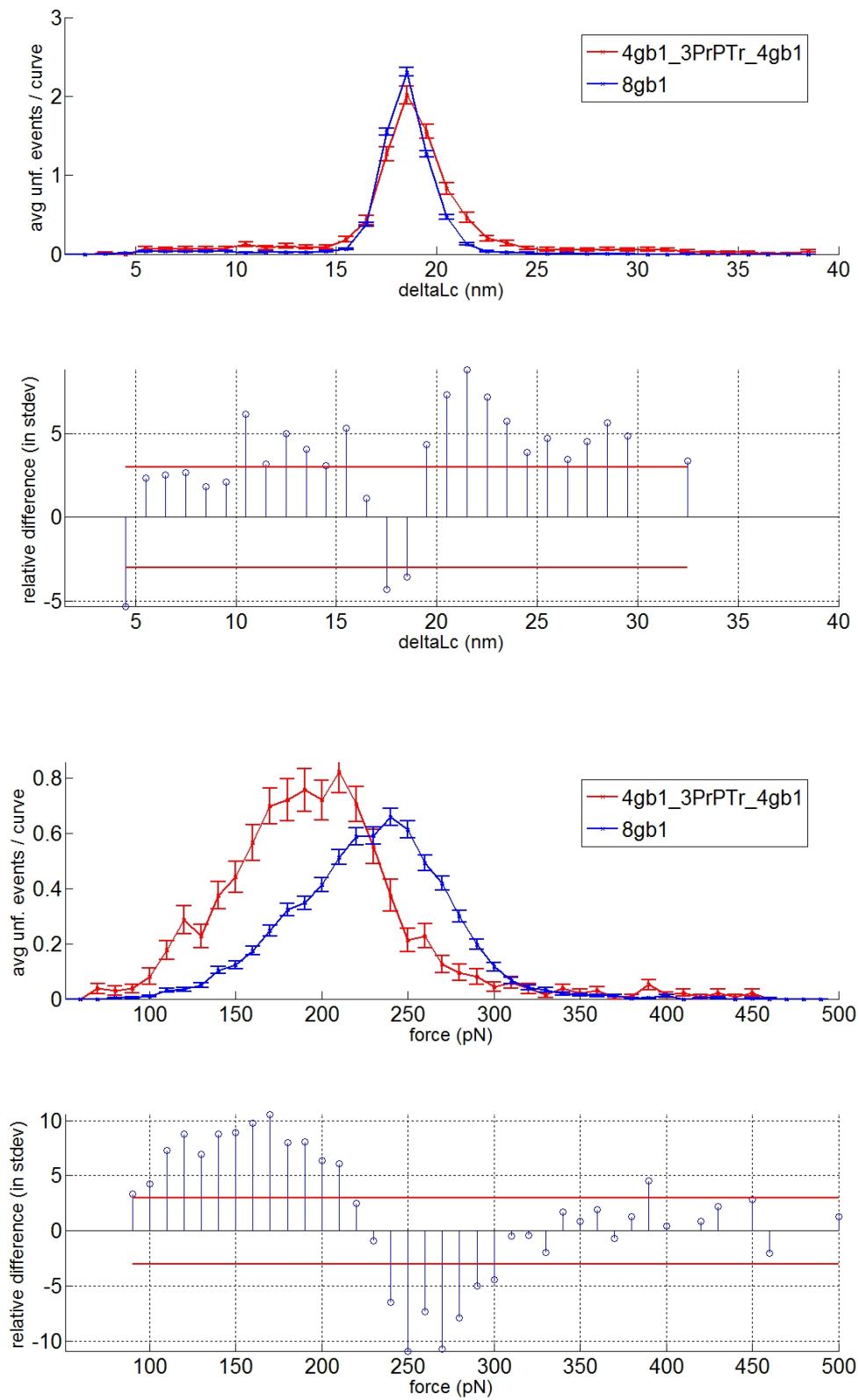


Figura 8.26: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB14PrPTR_3GB1_4$ e di $GB1_8$, misurati a pH 5.5. IN BASSO: confronto tra gli istogrammi delle forze di $GB14PrPTR_3GB1_4$ e di $GB1_8$, misurati a pH 5.5. Il metodo utilizzato è descritto in sezione 6.2.1

ficativo della spalla destra di GB1), ma meno visibile rispetto a misure effettuate in ambiente meno acido. Nella distribuzione delle forze si nota l'innalzamento significativo della coda sinistra della distribuzione, ma meno evidente rispetto a quello osservato a pH più elevati.

8.2.4 Tetramero

Nel caso del tetramero l'abbassamento del pH ha un effetto diverso dai casi precedenti, in quanto comporta un progressivo allargamento della distribuzione con eventi, apparentemente, sempre più sparsi. Questo fenomeno è ben visibile sia negli istogrammi monodimensionali (figure 8.28 e 8.29) che in quelli bidimensionali (figure 8.30 e 8.31). Negli istogrammi monodimensionali si nota in particolare un innalzamento generico della distribuzione delle Δ_{Lc} , senza che sia evidente una lunghezza preferenziale. Di conseguenza il picco principale di GB1 si abbassa per il già discusso fenomeno dovuto alla selezione delle curve.

Per quanto riguarda le forze si nota come la parte bassa della distribuzione sembri rimanere invariata al variare del pH, questo andamento potrebbe suggerire un'indipendenza della forze di *unfolding* per gli eventi del tetramero.

Anche in questo caso un fit bidimensionale delle distribuzioni sarà necessario per caratterizzarle completamente.

8.3 Sommario

Lo studio degli eventi di *unfolding* di monomeri e oligomeri di MoPrPTr (a.a. 89-230) misurati tramite SMFS ha permesso di caratterizzare diverse transizioni strutturali. L'utilizzo di costrutti diversi e di condizioni ambientali diverse ha permesso di delucidare gli equilibri conformazioni della proteina. Più in particolare:

- Si è ottenuta una distribuzione di Δ_{Lc} centrata a circa 22 nm in tutti i costrutti, eccetto nel monomero rifoldato in condizioni riducenti e nel dimero in orientazione H-H. Tale transizione è stata associata all'*unfolding* dello stato nativo del dominio C-terminale. La sua presenza permane, ed è sempre più visibile, all'aumentare del numero di monomeri presenti, suggerendo che le conformazioni stabili in presenza di più monomeri non siano necessariamente associative.
- L'orientazione relativa dei monomeri è di fondamentale importanza per il *foldng*: infatti, mentre nel dimero in orientazione H-T permane la transizione associabile allo stato nativo, essa non è visibile nel dimero in orientazione H-H.
- Per quanto riguarda monomero, dimero e trimero l'abbassamento del pH comporta una destabilizzazione dello stato nativo.
- Il tetramero ha un comportamento diverso dagli altri costrutti per quanto riguarda il pH: un abbassamento del pH in questo caso comporta una distribuzione sempre meno piccata degli eventi.
- Nel tetramero sono visibili più distribuzioni di eventi: sembrano essere visibili una distribuzione a 10 nm, una a 27 nm e una tra 40 e 50 nm. Sono necessari ulteriori esperimenti per caratterizzare strutturalmente queste transizioni.
- Esistono eventi associabili a MoPrP all'interno dell'intervallo di Δ_{Lc} associabili a GB1, tali eventi influiscono sulla selezione delle curve, in particolare sono rilevanti per il tetramero, poichè in questo costrutto il rapporto tra numero di GB1 e numero di monomeri di MoPrP presenti è minore rispetto agli altri costrutti.

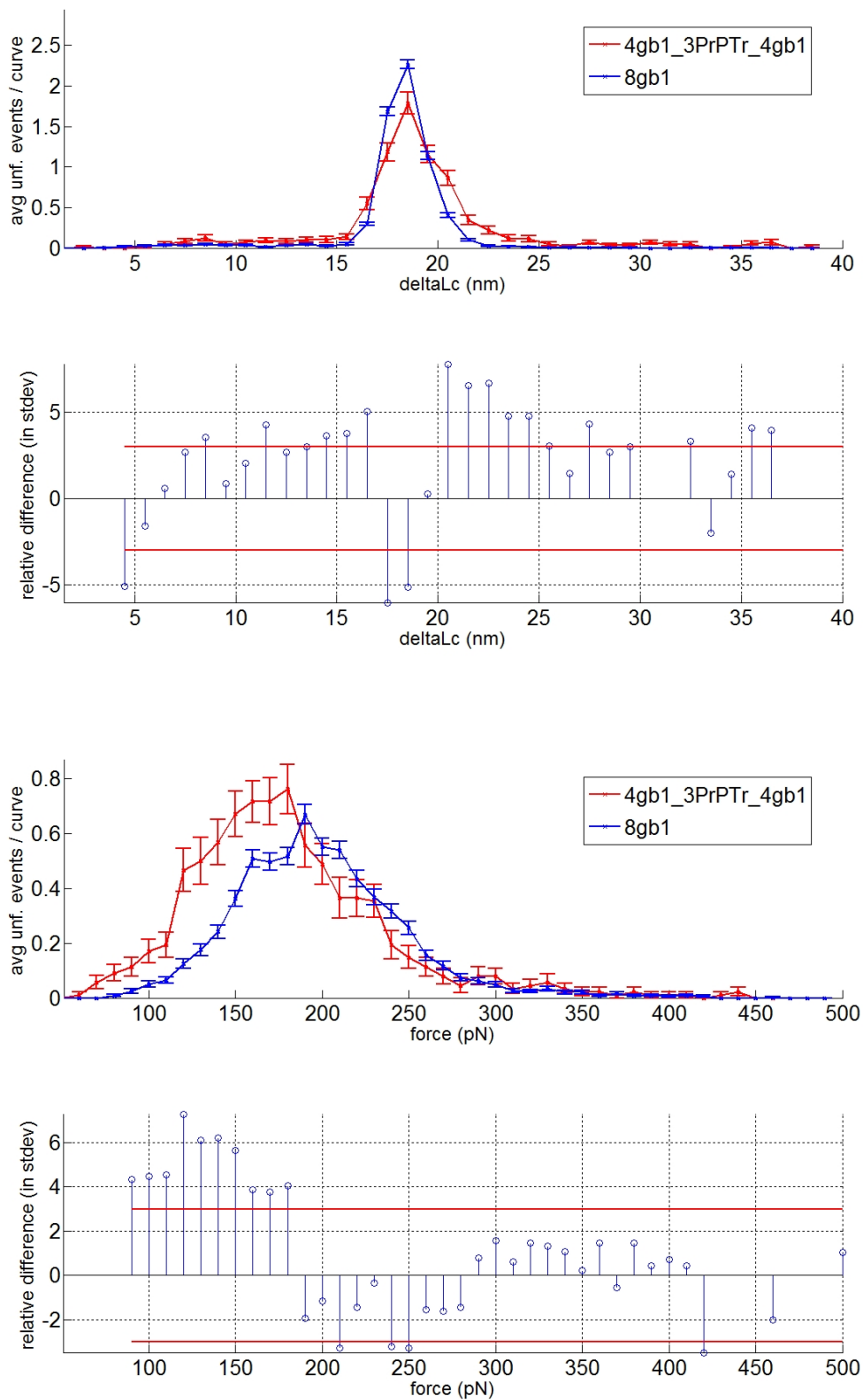


Figura 8.27: IN ALTO: confronto tra gli istogrammi delle Δ_{Lc} di $GB1_4PrPTr_3GB1_4$ e di $GB1_8$, misurati a pH 4.0. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_4PrPTr_3GB1_4$ e di $GB1_8$, misurati a pH 4.0. Il metodo utilizzato è descritto in sezione 6.2.1

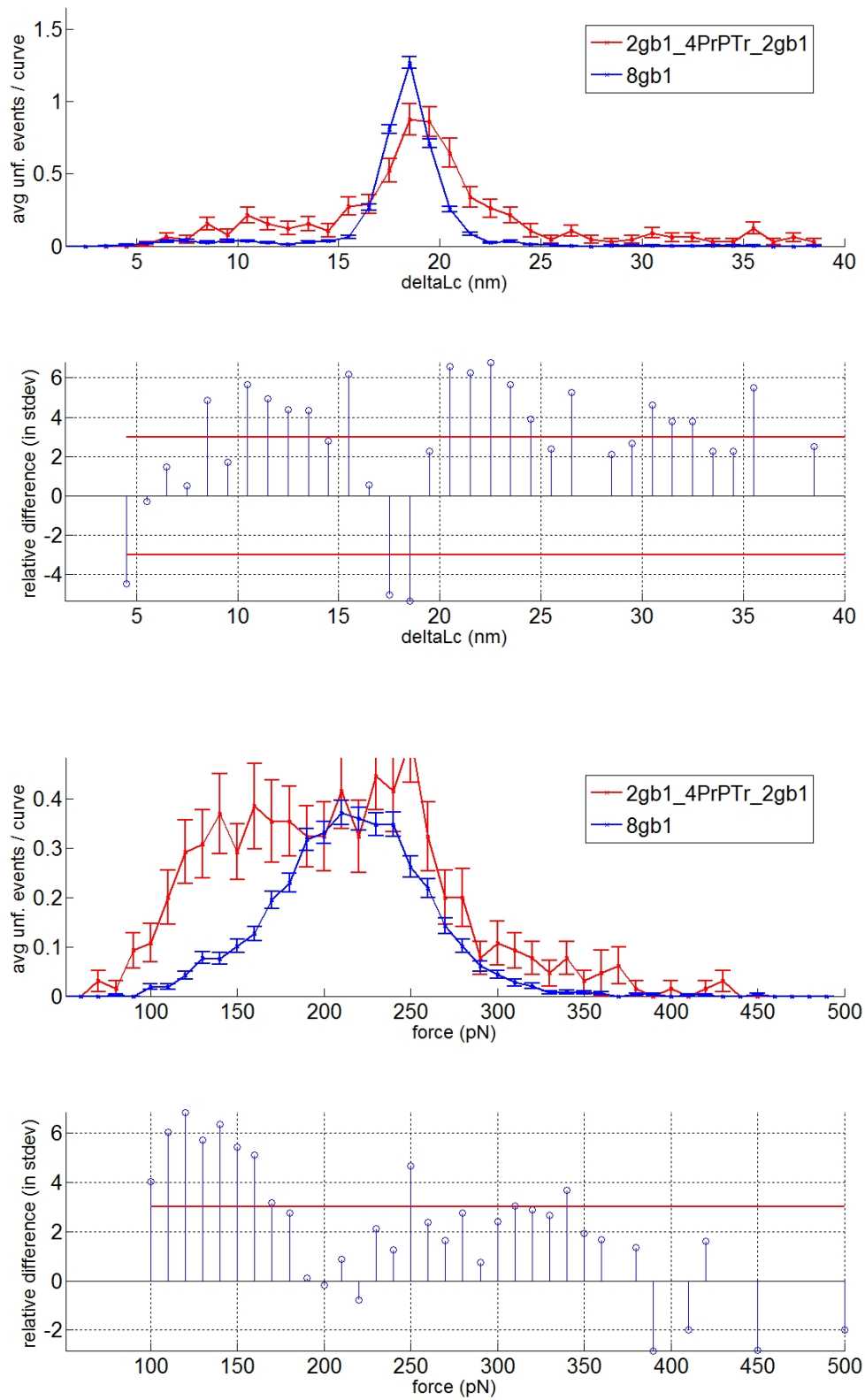


Figura 8.28: IN ALTO: confronto tra gli istogrammi delle ΔL_c di $GB_{12}PrPTr_4GB_{12}$ e di GB_{18} , misurati a pH 5.5. IN BASSO: confronto tra gli istogrammi delle forze di $GB_{12}PrPTr_4GB_{12}$ e di GB_{18} , misurati a pH 5.5. Il metodo utilizzato è descritto in sezione 6.2.1.

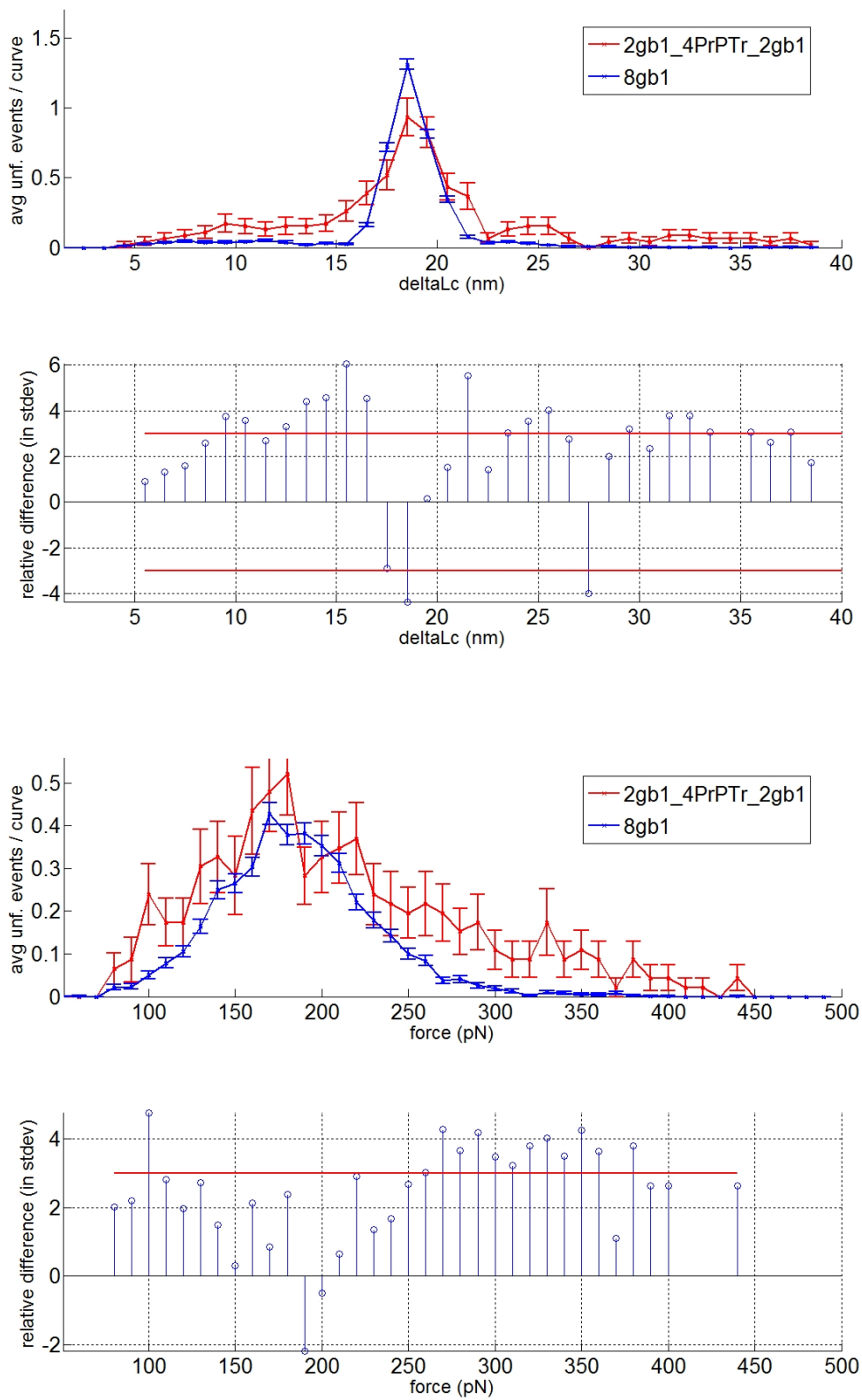


Figura 8.29: IN ALTO: confronto tra gli istogrammi delle ΔLc di $GB1_2PrPTR_4GB1_2$ e di $GB1_8$, misurati a pH 4.0. IN BASSO: confronto tra gli istogrammi delle forze di $GB1_2PrPTR_4GB1_2$ e di $GB1_8$, misurati a pH 4.0. Il metodo utilizzato è descritto in sezione 6.2.1.

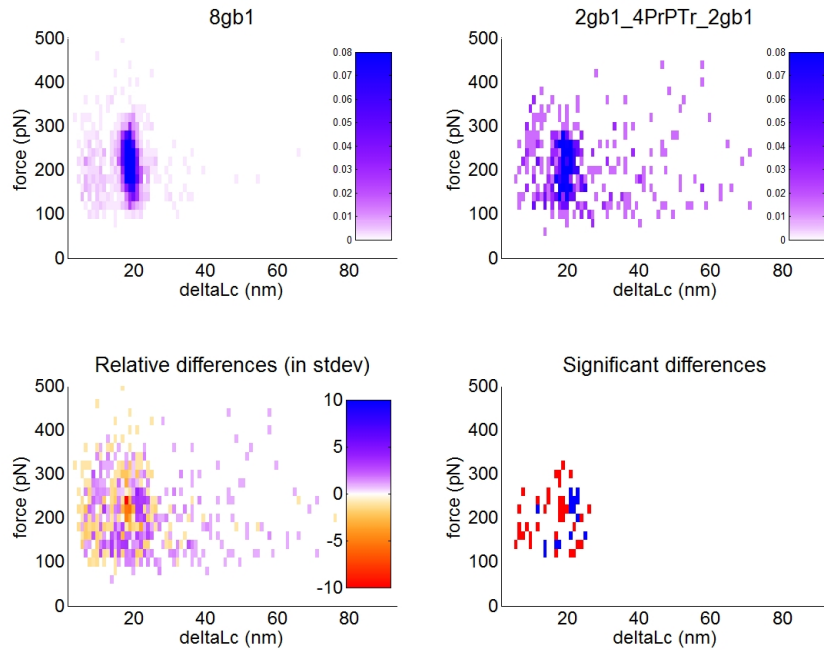


Figura 8.30: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_2PrPTR_4GB1_2$ e di $GB1_8$ (filtrato in modo da avere un numero di eventi di *unfolding* compreso tra 3 e 4), misurati a pH 5.5; il metodo utilizzato è descritto in sezione 6.2.2

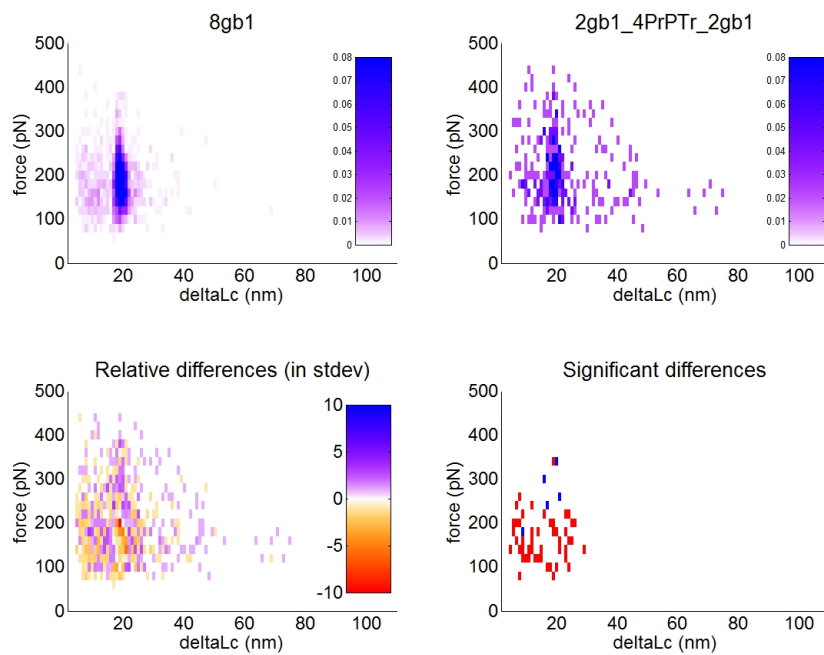


Figura 8.31: Confronto tra gli istogrammi bidimensionali degli eventi di $GB1_2PrPTR_4GB1_2$ e di $GB1_8$ (filtrato in modo da avere un numero di eventi di *unfolding* compreso tra 3 e 4), misurati a pH 4.0; il metodo utilizzato è descritto in sezione 6.2.2

Allo scopo di migliorare la caratterizzazione delle transizioni sono necessarie le seguenti operazioni:

- Sviluppo di un sistema di fit bidimensionale che tenga conto della correlazione tra forza e Δ_{Lc} .
- Esperimenti con un nuovo costrutto $GB1_4PrPTr_4GB1_4$ (attualmente in fase di esecuzione): tali esperimenti permetteranno di aumentare il rapporto tra numero di GB1 e di MoPrP presenti nel tetramero e ridurre quindi gli scambi di eventi di GB1 e di MoPrP, la selezione delle curve valide in questo costrutto sarà più precisa e permetterà di caratterizzare con maggior precisione le transizioni evidenti sul tetramero.
- Esperimenti con un costrutto contenente solo il dominio C-terminale (affiancato da GB1), questo esperimento di controllo permetterà di verificare l'attribuzione delle transizioni che avvengono a circa 22 nm.
- Studio della correlazione tra eventi appartenenti ad una medesima curva (soprattutto nei costrutti polimerici, al fine di identificare eventuali intermedi di *unfolding*).

Capitolo 9

Conclusioni e sviluppi

Nel corso di questo progetto di tesi si sono effettuati esperimenti di spettroscopia di forza di singola molecola per caratterizzare gli equilibri conformazionali della proteina prionica murina (MoPrP). Sono stati inoltre effettuati esperimenti su costrutti oligomerici in modo da studiare i primi eventi aggregativi.

Inizialmente è stato necessario caratterizzare il processo di denaturazione meccanica della proteina GB1, utilizzata come riferimento nel corso degli esperimenti. Esso è stato descritto come un processo a due stati, caratterizzato da una Δ_{Lc} di 18.5 ± 0.9 nm. La distribuzione delle forze di *unfolding* è piuttosto larga e dipende dal pH (a pH 4.0 le forze di *unfolding* diminuiscono). Si sono misurati anche un numero minoritario di eventi con Δ_{Lc} ; 15 nm (e quindi lontani dalla distribuzione principale). Tali eventi sono stati studiati e la loro distribuzione è stata caratterizzata, ma al momento non si è giunti ad una interpretazione definitiva del fenomeno.

Allo scopo di confrontare la distribuzione degli eventi di GB1 con quelle dei costrutti contenenti una o più MoPrP si è sviluppato un sistema di analisi dati funzionante sia su una sola dimensione (forza o Δ_{Lc}) che in due dimensioni (forza *vs.* Δ_{Lc}). Questo sistema confronta gli istogrammi delle distribuzioni di due costrutti, normalizzando il numero di conteggi ottenuti in ogni classe in base al numero totale di curve raccolte per quel determinato costrutto. Vengono dunque confrontati i numeri medi di eventi ottenuti per i due costrutti in un determinato intervallo, permettendo di riconoscere differenze significative.

Utilizzando questo sistema di analisi è stato possibile studiare diversi costrutti.

Si è riconosciuta la transizione del ripiegamento nativo del dominio C-terminale della proteina MoPrP. La distribuzione degli eventi ad esso associata è stata riconosciuta nel costrutto monomero e nei costrutti oligomerici con orientazione H-T. Tale transizione è meccanicamente stabile anche aumentando la concentrazione di monomero presente. Il ponte disolfuro tra le cisteine 178-213 si è dimostrato fondamentale per il corretto ripiegamento del dominio C-terminale, mentre l'abbassamento del pH destabilizza la struttura nativa di questo dominio. Questi risultati sono in accordo con esperimenti in bulk presenti in letteratura.

Si è riscontrato come l'orientazione di due monomeri abbia un ruolo fondamentale nella determinazione della struttura, infatti nel dimero in orientazione H-H il ripiegamento nativo non è più favorito, probabilmente a causa della eccessiva vicinanza dei due domini C-terminali.

Gli eventi non associabili alla transizione nativa sono molto sparsi, ad indicare che non esistono strutture associative uniche, delineando un panorama energetico della proteina prionica caratterizzato dalla presenza di molti minimi locali, corrispondenti a molteplici conformazioni che la proteina può assumere e tra le quali può diffondere.

Solo nel costrutto tetramero sono visibili alcuni *cluster* di eventi associabili a transizioni di PrP, riconducibili ad associazioni tra due o tre monomeri. La loro caratterizzazione ed interpretazione strutturale potranno rivelare importanti caratteristiche dei processi aggregativi. Per questo motivo è in fase di produzione un nuovo costrutto tetramero, che dovrebbe

garantire una maggior sicurezza nella selezione delle curve e quindi nell'interpretazione dei risultati.

Complessivamente questo lavoro ha permesso di descrivere la proteina MoPrP come una proteina strutturata in cui esistono, però, conformazioni transienti. La varietà di conformazioni possibili rende questa proteina molto simile ad un'altra proteina legata alle malattie neurodegenerative, ovvero all' α -sinucleina, una proteina intrinsecamente disordinata responsabile del morbo di Parkinson [23]. Questa interessante analogia contribuisce alla formulazione di modelli unificati per la spiegazione dell'origine delle malattie neurodegenerative.

Appendice A

Appendice: codice del *software* per il confronto delle distribuzioni

Di seguito è riportato il codice utilizzato per la generazione delle immagini di confronto tra distribuzioni. Le funzioni riportate vengono richiamate da un *software* sviluppato appositamente per la visualizzazione dei dati contenuti nei file *.hkp* e *.repr* restituiti dal *software* di analisi delle curve.

I dati relativi a tutte le curve raccolte vengono memorizzati in dei **structured array** (di seguito indicati come *playlist*) contenenti vari campi. L'unico campo di interesse nelle funzioni di seguito riportate è il campo **data**. Tale campo contiene uno **structured array**, in cui sono memorizzati tutti i dati estratti dal fit della curva, salvati in appositi campi. I campi di interesse per le seguenti funzioni sono il campo **deltaLc** e il campo **force** in cui sono memorizzati, rispettivamente, i valori di Δ_{Lc} e di forza misurati per tutti i picchi contenuti nella curva.

A.1 Analisi monodimensionale

Le funzioni di seguito riportate sono utilizzate per l'analisi monodimensionale, gli *input* necessari a *histmonodim* sono i seguenti:

- . molR,molS : le *playlist* relative ai due costrutti da confrontare
- . minx,maxx,maxp,step : valore minimo e massimo delle classi, larghezza delle classi e valore della classe massima raffigurata
- . name : variabile di cui si vuole fare l'istogramma
- . limR,limS :valori di deviazione standard massimi critici per determinare la significatività, il sistema è sviluppato in modo da poter valutare indipendente la significatività delle differenze in base alla deviazione standard di un campione piuttosto che dell'altro. Per ottenere i grafici riportati nel testo è necessario porre limR=1.5, limS=1.5
- . NameRef, NameS : stringhe che vengono riportate nella legenda
- . panel : parametro necessario per la corretta visualizzazione nell'interfaccia grafica.

```
function [ s1,s2,meanValueS, meanValueR,stR,stS ] =histmonodim(molR,molS,...
    minx,maxx,maxp,step,name,limR,limS,NameRef,NameS, panel )
```

```
    s1=subplot(2,1,1, 'Parent',panel);
    hold on;
```

```

bins=[minx:step:maxx];

%compute and plot histograms
[meanValueS,stS]= interexpcomparison1d(molS, bins,name,'r',...
    floor((maxp-minx)/step));
[meanValueR,stR]= interexpcomparison1d(molR, bins,name,'b',...
    floor((maxp-minx)/step));
legend({NameS,NameRef},'Interpreter', 'none');
xlabel(name,'FontSize',18);
ylabel('avg unf. events / curve','FontSize',18);
axis([minx+1 maxp 0 max(meanValueR)+0.3*max(meanValueR)]);
set(gca, 'FontSize',20);
hold off;

%find differences between sample and reference
ind1a = ( meanValueR >= meanValueS);
ind1b = (meanValueR-limR*stR>meanValueS+limS*stS);
ind1=find(ind1a.*ind1b.*stR);

ind2a= meanValueR<=meanValueS;
ind2b= meanValueR+limR*stR<meanValueS-limS*stS;
ind2= find(ind2a.*ind2b.*stR);

j=1;

for i=1: length(stS)
    if stR(i) ~= 0
        diff(1,j)=2*(meanValueS(i)-meanValueR(i))/(stR(i)+stS(i));
        diff(2,j)=bins(i);
        j=j+1;
    end
end

s2=subplot(2,1,2, 'Parent',panel);
hold on;
upconf=zeros(1,length(diff)-1);
lowconf=zeros(1,length(diff)-1);
upconf(:)=limS+limR;
lowconf(:)=-limS-limR;
stem(diff(2,2:end),diff(1,2:end));
plot(diff(2,2:end),upconf(1,:),'r','linewidth',2);
plot(diff(2,2:end),lowconf(1,:),'r','linewidth',2);
xlabel(name,'FontSize',18);
ylabel('relative difference (in stdev)','FontSize',18);
axis([minx+1 maxp min(diff(1,:)) max(diff(1,:))]);
set(gca, 'FontSize',20);
grid on;
hold off;
end

```

```

function [ meanValue,stdev ] = interexpcomparison1d( mol, bins ,name,c,maxp)

    i=1;
    histo=zeros(length(mol),length(bins));

    %create histogram for all curves
    for i=1:length(mol)
        selectedVal=[mol(i).data.(genvarname(name(1:end-5)))(1:end-1)];
        histo(i,:)=hist(selectedVal,bins);
    end

    %compute mean and stdev of histograms
    stdev=std(histo)/sqrt(length(mol));
    meanValue=mean(histo);

    %plot
    errorbar(bins(1:maxp),meanValue(1:maxp),stdev(1:maxp),[c,'-x'],...
        'linewidth',2);
end

```

A.2 Analisi bidimensionale

```

function [ax1, ax2, ax3, ax4,ann]=histbidim(molR,molS,limR,limS,...
    NameRef,NameS, panel)

    dataVector=[molS.data];

    %define data range and bins
    maxdLc=max([dataVector.deltaLc]);
    minF=min([dataVector.force]);
    ctrs{1}=[0.5:1:maxdLc+10];
    ctrs{2}=[0:20:500];
    x=[1.5 maxdLc+10];
    y=[0 500];

    %comparison
    [ meanValueRef,stdev ] = interexpcomparison2d( molR, ctrs );
    [ meanValueS,stdevS ] = interexpcomparison2d( molS, ctrs );

    %3D histogram construction
    fig=(squeeze(meanValueRef))';%reference
    num=(squeeze(meanValueS))';%sample
    maxVal=0.08;

    %colormap definition (to optimize the visibility of low counts)
    ncol=255;
    cmap=ones(ncol,3);
    cmap(1:ncol/2,3)=[0:1/(ncol/2-1):1].^4;
    cmap(ncol/2:ncol,1)=[1:-1/(ncol/2):0];

```

```

cmap(1:ncol/2,2)=[0:1/(ncol/2-1):1];
cmap(ncol/2:ncol,2)=[1:-1/(ncol/2):0].^4;

%plot

ax1=subplot(2,2,2, 'Parent',panel); % sample 3D histogram
hold on;
title(NameS,'FontSize',20,'Interpreter', 'none');
caxis([-maxVal,maxVal]);
image(x,y,(num(:,2:end-1)),'CDataMapping','scaled');
plot(x,[0,0],'k');
plot([x(1),x(1)],y,'k');
axis([x,y]);
set(ax1, 'FontSize',18);
xlabel('deltaLc (nm)','FontSize',18);
ylabel('force (pN)','FontSize',18);
h=colorbar('east');
set(h, 'ylim', [0 maxVal]);
hold off;

ax2=subplot(2,2,1, 'Parent',panel); % reference 3D histogram
hold on;
title(NameRef,'FontSize',20,'Interpreter', 'none');
caxis([-maxVal,maxVal]);
image(x,y,(fig(:,2:end-1)),'CDataMapping','scaled');
plot(x,[0,0],'k');
plot([x(1),x(1)],y,'k');
axis([x,y]);
set(ax2, 'FontSize',18);
xlabel('deltaLc (nm)','FontSize',18);
ylabel('force (pN)','FontSize',18);
h=colorbar('east');
set(h, 'ylim', [0 maxVal]);
hold off;

ax4=subplot(2,2,4, 'Parent',panel); %significant differences
hold on;
title('Significant differences','FontSize',20);
crit=zeros(size(num));
ind1a = ( fig >= num);
ind1b = (fig-limR*(squeeze(stdev))'>num+limS*(squeeze(stdevS))');
ind1=find(ind1a.*ind1b);
ind2a= ( fig <= num);
ind2b= (fig+limR*(squeeze(stdev))'<num-limS*(squeeze(stdevS))');
ind2= find(ind2a.*ind2b);
crit(ind1)=-1;
crit(ind2)=1;
caxis([-1,1]);
image(x,y,(crit(:,2:end-1)),'CDataMapping','scaled');
plot(x,[0,0],'k');
plot([x(1),x(1)],y,'k');

```

```

axis([x,y]);
set(gca, 'FontSize',18);
xlabel('deltaLc (nm)', 'FontSize',18);
ylabel('force (pN)', 'FontSize',18);
strRed = {'sample > Ref'};
strOr = {'sample < Ref'};
ann(1)=annotation('textbox', [.78 .25 .1 .1], 'String', strRed,'Color',...
    cmap(ncol,:), 'Fontsize',12,'LineStyle','none', 'Tag', 'annotate');
ann(2)=annotation('textbox', [.78 .28 .1 .1], 'String', strOr,'Color',...
    cmap(1,:), 'Fontsize',12,'LineStyle','none', 'Tag', 'annotate');
hold off;

ax3=subplot(2,2,3, 'Parent',panel); %relative differences
hold on;
title('Relative differences (in stdev)', 'FontSize',20);
A=find(stdev==0);
stdev(A)=stdevS(A);
B=find(stdevS==0);
stdevS(B)=stdev(B);
div=((squeeze(stdev)+squeeze(stdevS))/2);
diff=((num-(squeeze(meanValueRef)))'./div');
diff(find(isnan(diff)))=0;
caxis([-10,10]);
image(x,y,(diff(:,2:end-1)),'CDataMapping','scaled');
plot(x,[0,0], 'k');
plot([x(1),x(1)],y, 'k');
axis([x,y]);
colorbar('east');
set(ax3, 'FontSize',18);
xlabel('deltaLc (nm)', 'FontSize',18);
ylabel('force (pN)', 'FontSize',18);
hold off;

colormap((cmap));
end

```

```

function [ meanValue,stdev ] = interexpcomparison2d( mol, ctrs )

histo=zeros(length(mol), length(ctrs{1}),length(ctrs{2}));

%create histogram for all curves
for i=1:length(mol)
    selectedMol=mol(i).data;
    histo(i, :, :)=hist3([[selectedMol.deltaLc]', [selectedMol.force]'], ctrs);
end

%compute mean and stdev of histograms
stdev=std(histo)/sqrt(length(mol));

```

```
meanValue=mean(histo);  
end
```

Gli input di *histbidim* aventi lo stesso nome di quelli di *histmonodim* rappresentano le stesse quantità.

In questo caso l'intervallo per l'istogramma delle forze è fissato (in base ai valori ottenibili nel corso degli esperimenti) tra 0 e 500 pN. L'intervallo in cui fare l'istogramma delle lunghezze viene definito in modo da garantire la rappresentazione di tutti i punti esistenti.

Bibliografia

- [1] Daniel Aioanei. *The effect of osmolytes on protein fold stability at the single-molecule level*. PhD thesis, Università di Bologna, 2013.
- [2] Daniel Aioanei, Marco Brucale, and Bruno Samorí. Open source platform for the execution and analysis of mechanical refolding experiments. *Bioinformatics*, 27(3):423–425, 2011.
- [3] George I Bell. Models for the specific adhesion of cells to cells. *Science*, 200(4342):618–627, 1978.
- [4] C Bustamante, J F Marko, E D Siggia, and S Smith. Entropic elasticity of λ -phage dna. *Science*, 265:1599–1600, 1994.
- [5] Luigi Calzolari and Ralph Zahn. Influence of ph on nmr structure and stability of the human prion protein globular domain. *Journal of Biological Chemistry*, 278(37):35592–35596, 2003.
- [6] Yi Cao, Canaan Lam, Meijia Wang, and Hongbin Li. Nonmechanical protein can have significant mechanical stability. *Angewandte Chemie*, 118(4):658–661, 2006.
- [7] Yi Cao and Hongbin Li. Polyprotein of gb1 is an ideal artificial elastomeric protein. *Nature materials*, 6(2):109–114, 2007.
- [8] Mariano Carrión-Vázquez, Andrés F Oberhauser, Héctor Díez, Rubén Hervás, Javier Oroz, Jesús Fernández, and David Martínez-Martín. Protein nanomechanics as studied by afm single-molecule force spectroscopy. In *Advanced Techniques in Biophysics*, pages 163–245. Springer, 2006.
- [9] David W Colby, Kurt Giles, Giuseppe Legname, Holger Wille, Ilia V Baskakov, Stephen J DeArmond, and Stanley B Prusiner. Design and construction of diverse mammalian prion strains. *Proceedings of the National Academy of Sciences*, 106(48):20417–20422, 2009.
- [10] Mari L DeMarco and Valerie Daggett. Molecular mechanism for low ph triggered misfolding of the human prion protein. *Biochemistry*, 46(11):3045–3054, 2007.
- [11] Rodrigo Diaz-Espinoza and Claudio Soto. High-resolution structure of infectious prion protein: the final frontier. *Nature structural & molecular biology*, 19(4):370–377, 2012.
- [12] Hendrik Dietz and Matthias Rief. Exploring the energy landscape of gfp by single-molecule mechanical experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46):16192–16197, 2004.
- [13] Evan Evans and Ken Ritchie. Dynamic strength of molecular adhesion bonds. *Biophysical journal*, 72(4):1541–1555, 1997.

-
- [14] E-L Florin, Matthias Rief, H Lehmann, Markus Ludwig, C Dornmair, Vincent T Moy, and Hermann E Gaub. Sensing specific molecular interactions with the atomic force microscope. *Biosensors and Bioelectronics*, 10(9):895–901, 1995.
- [15] Rubén Hervás, Javier Oroz, Albert Galera-Prat, Oscar Goñi, Alejandro Valbuena, Andrés M Vera, Ángel Gómez-Sicilia, Fernando Losada-Urzáiz, Vladimir N Uversky, Margarita Menéndez, et al. Common features at the start of the neurodegeneration cascade. *PLoS-Biology*, 10(5):1014, 2012.
- [16] Simone Hornemann and Rudi Glockshuber. A scrapie-like unfolding intermediate of the prion protein domain prp (121–231) induced by acidic ph. *Proceedings of the National Academy of Sciences*, 95(11):6010–6014, 1998.
- [17] Christopher R Jacobs, Hayden Huang, and Ronald Y Kwon. *Introduction to cell mechanics and mechanobiology*. Garland Science, 2012.
- [18] Tuomas PJ Knowles, Michele Vendruscolo, and Christopher M Dobson. The amyloid state and its association with protein misfolding diseases. *Nature Reviews Molecular Cell Biology*, 15(6):384–396, 2014.
- [19] Hongbin Li, Hui-Chuan Wang, Yi Cao, Deepak Sharma, and Meijia Wang. Configurational entropy modulates the mechanical stability of protein gb1. *Journal of molecular biology*, 379(4):871–880, 2008.
- [20] Lewyn Li, Hector Han-Li Huang, Carmen L Badilla, and Julio M Fernandez. Mechanical unfolding intermediates observed by single-molecule force spectroscopy in a fibronectin type iii module. *Journal of molecular biology*, 345(4):817–826, 2005.
- [21] Bao-Yuan Lu, Pamela J Beck, and Jui-Yoa Chang. Oxidative folding of murine prion mprp (23–231). *European Journal of Biochemistry*, 268(13):3767–3773, 2001.
- [22] Nilesh Ranjan Maiti and Witold K Surewicz. The role of disulfide bridge in the folding and stability of the recombinant human prion protein. *Journal of Biological Chemistry*, 276(4):2427–2431, 2001.
- [23] Krishna Neupane, Allison Solanki, Iveta Sosova, Miro Belov, and Michael T Woodside. Diverse metastable structures formed by small oligomers of α -synuclein probed by force spectroscopy. *PloS one*, 9(1), 2014.
- [24] Linus Pauling and Robert B Corey. The pleated sheet, a new layer configuration of polypeptide chains. *Proc Natl Acad Sci USA*, 37(5):251–256, 1951.
- [25] Andrea Raspadori. *Characterization of the conformational space of the murine prion protein using single-molecule force spectroscopy techniques*. PhD thesis, SISSA, October 2014.
- [26] Roland Riek, Simone Hornemann, Gerhard Wider, Martin Billeter, Rudi Glockshuber, and Kurt Wüthrich. Nmr structure of the mouse prion protein domain prp(121-231). *Nature*, 382:11, 1996.
- [27] Bruno Samorì, G Zuccheri, and P Baschieri. Protein unfolding and refolding under force: methodologies for nanomechanics. *ChemPhysChem*, 6(1):29–34, 2005.
- [28] Jason C Sang, Chung-Yu Lee, Frederick Y Luh, Ya-Wen Huang, Yun-Wei Chiang, and Rita P-Y Chen. Slow spontaneous α -to- β structural conversion in a non-denaturing neutral condition reveals the intrinsically disordered property of the disulfide-reduced recombinant mouse prion protein. *Prion*, 6(5):489–497, 2012.

-
- [29] Ingo Schwaiger, Clara Sattler, Daniel R Hostetter, and Matthias Rief. The myosin coiled-coil is a truly elastic protein structure. *Nature materials*, 1(4):232–235, 2002.
- [30] Deepak Sharma, Ognjen Perisic, Qing Peng, Yi Cao, Canaan Lam, Hui Lu, and Hongbin Li. Single-molecule force spectroscopy reveals a mechanically stable protein fold and the rational tuning of its mechanical stability. *Proceedings of the National Academy of Sciences*, 104(22):9278–9283, 2007.
- [31] Holger Wille, Wen Bian, Michele McDonald, Amy Kendall, David W Colby, Lillian Bloch, Julian Ollesch, Alexander L Borovinskiy, Fred E Cohen, Stanley B Prusiner, et al. Natural and synthetic prion structure from x-ray fiber diffraction. *Proceedings of the National Academy of Sciences*, 106(40):16990–16995, 2009.
- [32] Hao Yu, Xia Liu, Krishna Neupane, Amar Nath Gupta, Angela M Brigley, Allison Solanki, Iveta Sosova, and Michael T Woodside. Direct observation of multiple misfolding pathways in a single prion protein molecule. *Proceedings of the National Academy of Sciences*, 109(14):5283–5288, 2012.
- [33] Peng Zheng, Yi Cao, Tianjia Bu, Suzana K Straus, and Hongbin Li. Single molecule force spectroscopy reveals that electrostatic interactions affect the mechanical stability of proteins. *Biophysical journal*, 100(6):1534–1541, 2011.