POLITECNICO DI TORINO

Collegio di Ingegneria Chimica e dei Materiali

Corso di Laurea Magistrale
in Ingegneria Chimica e dei Processi Sostenibili

Tesi di Laurea Magistrale

Uso delle ciclodestrine per prevenire la denaturazione interfacciale e l'aggregazione di proteine

Relatori:
Prof. Roberto Pisano
Dott. Andrea Arsiccio

Candidato:
Marcello Rospiccio

Marzo 2020

Ai miei genitori
Ai miei famigliari
Ai miei amici
Ai miei nonni

Indice

1 Introduzione **1**
1.1 Liofilizzazione 1
1.2 Liofilizzazione di proteine 2
1.2.1 Interfacce e superfici 2
1.2.2 Temperatura 3
1.2.3 Crio-concentrazione 4
1.2.4 Aggregazione 4
1.2.5 Reazioni chimiche 5
1.2.6 Variazioni di pH 5
1.2.7 Disidratazione 6
1.3 Meccanismi di protezione, proteine e formulazione 6
1.3.1 Crioprotezione 7
1.3.2 Lioprotezione 7
1.3.3 Protezione da denaturazione superficiale 8
1.3.4 Ciclodestrine 9
1.4 Dinamica Molecolare 11
1.5 Obiettivi della Tesi 13

2 Interfacce acqua-aria ed acqua-ghiaccio **15**
2.1 Interfaccia Acqua-Aria 15
2.1.1 Metodologia 16
2.1.2 Risultati ed osservazioni 22
2.2 Interfaccia Acqua-Ghiaccio 29
2.2.1 Metodologia 29
2.2.2 Risultati ed osservazioni 36
2.3 Conclusioni 39

3 Amminoacidi **41**
3.1 Metodologia 41
3.2 Risultati ed osservazioni 48
3.2.1 Radial Distribution Function 48
3.2.2 Inclusioni 54
3.2.3 Cluster 57

3.3 Conclusioni 59

4 Proteine **61**
4.1 Cenni sulle proteine analizzate 61
4.1.1 Lattato deidrogenasi, LDH 61
4.1.2 Fattore di stimolazione delle colonie di granulociti, GCSF 63
4.2 Metodologia 64
4.3 Risultati ed osservazioni 75
4.3.1 LDH 75
4.3.2 GCSF 81
4.4 Conclusioni 89

5 Considerazioni Finali **93**

Ringraziamenti / Acknowledgements **99**

Capitolo 1

Introduzione

1.1 Liofilizzazione

La liofilizzazione è un processo di essiccamento che si suddivide in tre fasi, ovvero congelamento, essiccamento primario ed essiccamento secondario. E' una tecnologia largamente diffusa nell'industria farmaceutica, in quanto particolarmente adatta per prodotti termolabili. Questi ultimi, infatti, spesso non risultano compatibili con i tradizionali processi di essiccamento quali, per esempio, lo spray-drying che ha elevati carichi termici ed il conseguente rischio di denaturazione [1]. Lo scopo della liofilizzazione è quello di rimuovere l'acqua per ottenere un prodotto a basso contenuto di umidità (tipicamente pari o inferiore a qualche punto percentuale in termini di peso rispetto a quello complessivo), quindi meno soggetto all'insorgenza di meccanismi di degradazione.

Il processo comincia con una fase di raffreddamento (spesso a -40 °C ÷ -50 °C) per ottenere una formulazione completamente congelata. Il ghiaccio formatosi viene successivamente rimosso per sublimazione riducendo la pressione del sistema fino ad un valore opportunamente basso (il vuoto così raggiunto è in genere pari a qualche Pa [2]): questa operazione prende il nome di essiccamento primario. Il vapore acqueo generatosi viene aspirato dalla camera di liofilizzazione e brinato sulle pareti di un apposito condensatore. Risulta difficile e poco praticabile congelare e rimuovere per sublimazione l'intero contenuto d'acqua. Pertanto per raggiungere il tenore di umidità finale desiderato è spesso necessario un essiccamento secondario, in cui l'acqua non congelata viene rimossa per desorbimento. Al termine di questa fase rimane un'esigua quantità d'acqua, legata più o meno fortemente alle molecole di soluto e di cui costituisce lo strato di idratazione [3]. Al termine dell'operazione si può rompere il vuoto e consentire al sistema di riportarsi alle condizioni ambientali, per poi recuperare il prodotto liofilizzato.

I liofilizzatori devono perciò essere costruiti in maniera tale da poter sopportare gli stress meccanici derivanti dagli sbalzi termici e dal vuoto al proprio interno, oltre a possedere opportuno materiale isolante. Ciò è indispensabile per raggiungere e mantenere le basse temperature necessarie al congelamento e per schermare

l'interno dalla radiazione termica proveniente dall'ambiente esterno. Il prodotto è solitamente preparato all'interno di ampolle o flaconi in materiale plastico o vetroso, disposte su ripiani al cui interno scorre un fluido di servizio dedicato alla regolazione della temperatura, gestito da un opportuno sistema di controllo.

Il congelamento avviene abbassando la temperatura della soluzione con un sufficiente grado di sotto-raffreddamento, necessario per indurre il cambiamento di fase. La camera di liofilizzazione è connessa ad un condensatore, dotato di opportuni sistemi di controllo per la regolazione della temperatura e della pressione, ed a sua volta collegato ad una pompa da vuoto per brinare il vapore acqueo.

Si noti che l'essiccamento si verifica già a partire dal congelamento stesso, poiché la progressiva solidificazione dell'acqua implica necessariamente una diminuzione del numero di molecole presenti nella fase amorfa e determina l'aumento della concentrazione dei soluti presenti. In questo contesto, la rampa di raffreddamento scelta ha un ruolo cruciale in quanto determina la velocità di accrescimento dei nuclei cristallini, influenzando di conseguenza la dimensione dei cristalli di ghiaccio, e quindi l'estensione dell'interfaccia acqua-ghiaccio.

1.2 Liofilizzazione di proteine

In questo lavoro di tesi ci si focalizzerà in particolare sul caso di proteine come principio attivo della formulazione. La liofilizzazione di biofarmaci a base proteica si rende necessaria per la corretta conservazione e lo stoccaggio del prodotto finale, poiché le proteine in soluzione acquosa possono essere soggette a fenomeni in grado di causarne la denaturazione. Questi ultimi possono essere di natura fisica, chimica o biologica, e comprendono variazioni di temperatura, pH e concentrazione dei soluti, oltre che reazioni chimiche o biochimiche degradative. Per evitare l'insorgenza di questi fenomeni è di fondamentale importanza la rimozione dell'acqua, poiché essa rappresenta il mezzo attraverso cui numerosi tra i possibili meccanismi di degradazione si possono svolgere. Per fare ciò si può ricorrere alla liofilizzazione che, inoltre, consente di ovviare al rischio di denaturazione a caldo della proteina. Nonostante i vantaggi, la liofilizzazione presenta alcune problematiche che vanno affrontate per garantire e preservare la qualità del prodotto finale, come verrà dettagliato nei paragrafi seguenti.

1.2.1 Interfacce e superfici

La presenza di interfacce può determinare l'adsorbimento delle proteine in loro prossimità e favorirne lo svolgimento, meccanismo che può anche rivelarsi irreversibile compromettendo irrimediabilmente la qualità del farmaco. Vi sono diverse tipologie di superfici presenti durante lo svolgimento della liofilizzazione, e che possono contribuire alla destabilizzazione del prodotto: aria-acqua, acqua-ghiaccio ed eventualmente vetro-acqua (o analoghe), qualora la soluzione da liofilizzare sia contenuta in recipienti vetrosi [4]. Sebbene presentino proprietà chimico-fisiche

differenti, ciascuna di esse possiede la capacità di interagire con le proteine con esiti diversi, a seconda della loro natura e dell'eventuale presenza di tensioattivi.

L'estensione e le proprietà chimico-fisiche dell'interfaccia acqua-vetro possono essere modulate scegliendo opportunamente il contenitore (ad es. selezionandone dimensione e opportuni trattamenti superficiali). La formazione delle altre due superfici interfacciali, per quanto inevitabile, può essere opportunamente limitata durante le varie fasi del processo prestando attenzione alle modalità ed alle condizioni operative. L'interfaccia acqua-aria è necessariamente presente a causa del contatto tra la fase liquida e quella gassosa; tuttavia il contributo più significativo è determinato dalla quantità di aria disciolta nella soluzione stessa durante le operazioni preliminari, in particolare a causa dell'agitazione che ne favorisce la dispersione. Per ovviare a questo problema risulta evidente la necessità di limitare e controllare l'agitazione del sistema.

L'estensione dell'interfaccia acqua-ghiaccio dipende dai parametri operativi che caratterizzano la liofilizzazione, in particolare dalla modalità con cui la fase di congelamento viene effettuata. Maggiore è la rapidità con cui si raffredda la soluzione e minore sarà la dimensione dei cristalli di ghiaccio ottenuti, poiché un brusco calo di temperatura incrementerà il tasso di nucleazione a scapito del tasso di accrescimento. L'estensione dell'interfaccia acqua-ghiaccio è inoltre influenzata dalla presenza di una fase di *annealing* (riscaldamento e mantenimento del prodotto congelato al di sopra della temperatura di transizione vetrosa). L'annealing promuove infatti un accrescimento dei cristalli di ghiaccio tramite il meccanismo della crescita di Ostwald [2].

Infine, la cristallizzazione di eventuali altri soluti facenti parte della soluzione può comportare la formazione di altre interfacce diverse da quelle appena descritte, tuttavia capaci di di destabilizzare la struttura proteica in modo analogo. Pertanto durante la progettazione del processo è altrettanto importante tenere conto delle possibili fasi solide così formatesi.

1.2.2 Temperatura

Poiché la liofilizzazione si svolge a basse temperature la proteina corre il rischio di denaturarsi a freddo. Questo meccanismo di degradazione consiste nella perdita della conformazione naturalmente assunta dalla molecola in esame, ovvero nello svolgimento (*unfolding*) della stessa, e può essere interpretato come un equilibrio termodinamico tra la forma nativa e quella denaturata [1]. A questo equilibrio chimico, e assumendo che tale trasformazione sia reversibile, è possibile associare una variazione di energia libera di Gibbs ΔG_{den}, funzione della temperatura del sistema T_{sist} [2]. In particolare, la variazione di energia libera associata alla denaturazione è legata alla temperatura da una relazione di tipo parabolico, pertanto esistono due valori di T_{sist} per cui ΔG_{den} si annulla, ovvero le temperature di denaturazione a freddo, T_{den}^{cold}, ed a caldo, T_{den}^{hot}, rispettivamente in ordine crescente. In base a questa interpretazione, risulta quindi evidente che la denaturazione termica possa verificarsi sia per temperature inferiori a T_{den}^{cold}, sia superiori a T_{den}^{hot} [5].

1.2.3 Crio-concentrazione

La crio-concentrazione è un fenomeno problematico che consiste nell'aumento della concentrazione dei soluti a causa del congelamento dell'acqua. Ciò può causare sbalzi significativi del pH a seconda del tipo di buffer utilizzato, poiché alcuni di essi possono dare precipitazione durante il processo compromettere quindi la stabilità delle proteine in soluzione [2]. Il pH infatti costituisce un altro parametro capace di causare lo svolgimento della proteina (e di conseguenza la denaturazione), oppure di impedire il corretto svolgimento della sua attività biochimica. Inoltre, l'aumento della concentrazione di alcune specie chimiche, tra cui in particolare l'ossigeno O_2, può far aumentare la cinetica di alcune reazioni chimiche indesiderate in grado di degradare il biofarmaco.

La crio-concentrazione in alcuni casi può anche promuovere la cristallizzazione di alcune specie chimiche eventualmente in soluzione, con conseguente formazione di ulteriori interfacce liquido-solido capaci di interagire con le proteine. Ciò determina un ulteriore rischio di denaturazione a causa dei fenomeni precedentemente discussi nel paragrafo §1.2.1.

1.2.4 Aggregazione

Uno dei fenomeni degradativi più insidiosi per quanto riguarda l'attività biologica delle proteine è rappresentato dall'aggregazione: essa può essere di carattere fisico o chimico, reversibile o irreversibile, a seconda della tipologia di interazioni intermolecolari intervenute nella formazione degli aggregati (o *cluster*). L'insorgenza di questo fenomeno ha diverse possibili conseguenze, dall'inattivazione del biofarmaco all'induzione di risposte immunitarie nel paziente [6]. Quest'ultimo scenario potrebbe rivelarsi decisamente dannoso per lo stato di salute del paziente, che in tal caso verrebbe ulteriormente compromesso con un potenziale aggravamento del quadro clinico.

L'aggregazione è generalmente causata da interazioni di tipo fisico tra le catene laterali dei residui peptidici apolari oppure da veri e propri legami chimici, tipicamente ponti disolfuro formatisi dalla reazione tra due residui di cisteina (*CYS*) appartenenti a proteine diverse. Si fa inoltre notare che i medesimi legami chimici all'interno di uno stesso peptide contribuiscono a stabilizzarne la struttura, pertanto una variazione nel loro contenuto può risultare nella denaturazione della molecola. L'aggiunta alla formulazione di molecole capaci di spezzare ponti disolfuro non è pertanto praticabile, in quanto oltre a limitare fenomeni di aggregazione porterebbe anche a forme non native della proteina.

Gli effetti dell'aggregazione di biofarmaci a base proteica durante la liofilizzazione possono variare: talvolta gli aggregati sono costituiti da molecole irreversibilmente legate e che possono perciò precipitare a seconda della loro solubilità. Tuttavia in altri casi i legami sono di tipo reversibile, e ciò incrementa la stabilità

delle proteine rispetto ad altri meccanismi degradativi, preservandone l'integrità fino alla ricostituzione del prodotto per la somministrazione [2].

1.2.5 Reazioni chimiche

L'insorgenza di reazioni chimiche può compromettere severamente la struttura e/o l'attività biologica delle proteine, a seconda delle specie chimiche coinvolte e della localizzazione del sito reattivo lungo la catena peptidica. In generale non esistono accorgimenti per poter escludere o impedire lo svolgimento di reazioni indesiderate a priori, poiché la loro sussistenza dipende dalle condizioni ambientali e dalle sostanze chimiche presenti nel sistema. Pertanto ogni caso specifico va analizzato con attenzione.
Oltre alla possibile formazione di ponti disolfuro tra cisteine diverse, in grado di causare denaturazione e/o aggregazione come discusso in precedenza, anche l'ossidazione delle catene laterali degli amminoacidi rappresenta un'insidia per l'integrità della proteina. Inoltre, altri plausibili percorsi reattivi sono la racemizzazione, l'idrolisi, la β-eliminazione e la de-ammidazione [1] ed infine la reazione di Maillard, particolarmente insidiosa considerato l'ampio utilizzo di zuccheri nelle formulazioni [2].

Per essere in grado di prevenire o limitare l'insorgenza di questi fenomeni è, quindi, indispensabile un accurato studio sperimentale sugli effetti delle condizioni ambientali e di processo, oltre che sulle caratteristiche di tutte le specie chimiche presenti, affidandosi, quando possibile, ai dati disponibili in letteratura.

1.2.6 Variazioni di pH

Il pH rappresenta uno dei parametri ambientali più rilevanti per la stabilità proteica, in quanto è in grado di influenzare la struttura assunta dalla proteina e di incidere anche sull'attività biologica che essa deve esplicare. Si pensi, per esempio, al delicato equilibrio che si instaura tra gli enzimi ed i ligandi (tra cui i reagenti) coinvolti nei processi biochimici. Per questa ragione anche piccole oscillazioni di pH nella soluzione possono causare inattivazione e/o denaturazione.
Per scongiurare questo pericolo, sia nell'ambito della produzione sia durante la ricostituzione del prodotto, è necessario utilizzare le dovute precauzioni, per esempio, aggiungendo opportuni sali alla soluzione in grado di agire da sistemi tampone (o *buffer*). In particolare, la scelta va effettuata in base all'intervallo di pH di interesse, che coincide con l'intervallo di stabilità della proteina in questione.

Come accennato nella sezione §1.2.3, è anche possibile che il buffer utilizzato precipiti durante il congelamento, a seconda della curva di solubilità e delle modalità operative (temperature iniziali e finali, tasso di raffreddamento). Consci di questo rischio, è necessario optare per un buffer in grado di rimanere in soluzione durante tale fase, esplicando la propria funzione per tutta la durata del processo.
Una tecnica utilizzata per scongiurare oscillazioni di pH eccessive e deleterie per il biofarmaco consiste nell'impiego di più sali in soluzione [2], ad esempio con

solubilità differenti, eventualmente da impiegarsi anche per range di pH diversi, in modo da tenere in considerazione gli effetti della crio-concentrazione. Infatti, la neutralità varia durante il congelamento, poiché diminuendo la temperatura cambia anche la costante di dissociazione dell'acqua.

Infine, la precipitazione del buffer rappresenta un ulteriore fenomeno di cui tenere conto, poiché ciò implica l'introduzione di un ulteriore interfaccia nel sistema, con il conseguente rischio di adsorbimento e denaturazione della proteina. Quest'ultima considerazione è valida per ogni specie chimica presente in soluzione in grado di dare cristallizzazione o ricristallizzazione durante la produzione del biofarmaco.

1.2.7 Disidratazione

Una volta rimosso il contenuto d'acqua il prodotto non può ritenersi privo di rischi per la stabilità: se infatti la rimozione del solvente elimina la possibilità di insorgenza per tutti quei meccanismi di degradazione mediati dalla soluzione, l'assenza di acqua potrebbe avere effetti altrettanto dannosi. Una riduzione eccessiva del contenuto interno di umidità risulterebbe infatti nella perdita dello strato di idratazione delle proteine, le quali necessitano di un sottile strato di molecole di acqua per poter mantenere la propria struttura [2]. A rigore ciò potrebbe anche rappresentare un fenomeno reversibile, pertanto in tal caso il farmaco recupererebbe la forma nativa. Tuttavia, ciò non può essere noto a priori e la formulazione scelta deve perciò essere in grado di far fronte a questo rischio.

1.3 Meccanismi di protezione, proteine e formulazione

Per evitare l'insorgenza di fenomeni degradativi e delle relative problematiche si rende necessario l'utilizzo di eccipienti, ovvero additivi stabilizzanti. Questi hanno l'obiettivo di preservare la qualità del farmaco sia durante la produzione, sia nell'arco della shelf-life dello stesso. Per soddisfare tale esigenza è necessario scegliere opportunamente ciascuna componente in base alla funzione specifica richiesta, ovvero al meccanismo di protezione più adeguato per una corrispondente tipologia di degradazione.
Tra gli eccipienti più comuni vi sono diverse sostanze con proprietà tensioattive (ad esempio amminoacidi, polimeri, Tween 80), ma anche polioli (inositolo, mannitolo, sorbitolo) e zuccheri, tra cui tipicamente monosaccaridi (glucosio, fruttosio), disaccaridi (saccarosio, trealosio, lattosio, maltosio) ed oligosaccaridi [2]. Tra questi ultimi, le ciclodestrine (in particolare β-ciclodestrina, idrossipropil-β-ciclodestrina, metil-β-ciclodestrina, sulfobutileter-β-ciclodestrina) hanno mostrato sperimentalmente la capacità di prevenire l'aggregazione di proteine, anche quando testate a basse concentrazioni. Ciò ha stimolato perciò l'interesse verso tali composti chimici [6], che saranno obiettivo di analisi in questo lavoro.
Ulteriori proprietà di interesse di queste specie chimiche verranno discusse nel pa-

ragrafo §1.3.4.

In generale, in una formulazione possono essere presenti diversi eccipienti, a seconda della loro capacità di porre rimedio ad uno o più dei fenomeni degradativi. Ciascuna di queste componenti va analizzata, sia individualmente che complessivamente, poiché non ci sono regole universali attraverso cui determinare la compatibilità, l'efficacia e l'eventuale interazione tra specie chimiche diverse, con effetti sinergici imprevedibili a priori.

In aggiunta a questi agenti stabilizzanti, ogni formulazione deve possedere anche un composto in grado di fornire il necessario supporto meccanico e strutturale, anche per ragioni estetiche. Le sostanze che soddisfano questi requisiti sono definite agenti di bulking [2]. Tali additivi spesso cristallizzano durante la liofilizzazione formando la matrice solida in cui sono inglobate tutte le altre componenti presenti nella formulazione, che diversamente potrebbero collassare una volta rimossa l'acqua. La compatibilità di ogni agente di bulking va valutata attentamente poiché può influire sulla stabilità proteica.

Per descrivere con maggiore chiarezza il ruolo di ciascuna tipologia di eccipienti si procede quindi nel seguito indicando la problematica ed il corrispondente meccanismo di protezione in atto per contrastarla.

1.3.1 Crioprotezione

Le conseguenze della temperatura, del congelamento, della crio-concentrazione e dell'aggregazione possono essere contrastate ed evitate utilizzando opportuni eccipienti, denominati crioprotettori, i quali esercitano sostanzialmente un meccanismo di stabilizzazione termodinamico, detto esclusione preferenziale. Tali molecole sono infatti respinte dalla superficie proteica in questione, ne impediscono lo svolgimento poiché un aumento dell'area esposta al solvente comporterebbe un incremento della repulsione tra eccipiente e proteina. Ciò corrisponderebbe ad un aumento del contenuto di energia del sistema e, perciò, ad una configurazione termodinamicamente sfavorevole [7].

1.3.2 Lioprotezione

La denaturazione indotta da disidratazione può essere contrastata attraverso la scelta di opportuni eccipienti, in questo contesto denominati lioprotettori, e della loro concentrazione. Il contenuto di umidità residuo deve bilanciare esigenze diverse, ovvero la necessità di ridurre il contenuto di acqua per scongiurare l'insorgenza dei meccanismi di degradazione mediati dalla soluzione, senza tuttavia rimuovere totalmente lo strato di idratazione indispensabile per preservare la corretta struttura proteica.

Si ritiene che esistano sostanzialmente tre differenti meccanismi protettivi: il *water replacement*, il *water entrapment* e la vitrificazione [8]. Sebbene in letteratura vi siano ipotesi diverse riguardo la fenomenologia molecolare di tali meccanismi, è stata anche avanzata un'idea che propone l'unificazione fenomenologica di

tali alternative [9], imputando l'azione stabilizzante alla formazione di una densa rete di legami idrogeno. In base a questa interpretazione, tali meccanismi sono perciò da considerarsi di tipo cinetico.

Il *water replacement* consiste nella capacità dell'eccipiente di rimpiazzare le molecole d'acqua nella formazione di legami idrogeno con la proteina [10]. Si ritiene importante sottolineare che sono fondamentalmente due i fattori in gioco per questo meccanismo: la disponibilità di atomi di ossigeno ed idrogeno in grado di interagire (che ovviamente dipende dalla struttura chimica) ed il volume molecolare, il quale determina necessariamente la distanza minima riscontrabile a causa delle repulsioni steriche e/o elettrostatiche.

Il *water entrapment* si verifica quando gli eccipienti formano una gabbia attorno alla proteina stabilendo numerosi legami idrogeno tra loro, con possibile formulazione di *cluster* e conseguente riduzione della mobilità delle molecole d'acqua confinate all'interno di tale spazio [9]. Queste molecole d'acqua non verranno rimosse durante la liofilizzazione, preservando lo strato di idratazione anche nel caso di essiccamento spinto e consentendo di mantenere la struttura proteica intatta. Lo spessore dello strato di idratazione dipende dalla tipologia di eccipienti utilizzati, poiché in base al loro ingombro sterico, all'affinità verso i residui peptidici ed alla capacità di formare una struttura avvolgente, sufficientemente compatta e regolare, essi si disporranno ad una certa distanza determinando la quantità d'acqua presente.

La vitrificazione si verifica in presenza di concentrazioni decisamente superiori rispetto a quelle sufficienti per i due meccanismi precedenti, ed è agevolata dall'utilizzo di eccipienti a peso molecolare superiore, capaci di determinare un aumento rilevante della viscosità della formulazione anche a bassi dosaggi. Tutto ciò si traduce nella formazione di una matrice compatta ed interconnessa, favorita dall'utilizzo di additivi stabilizzanti in grado di stabilire numerosi legami idrogeno, con ingombri sterici limitati in modo da favorire l'impaccamento [9]. Lo scopo è quello di limitare significativamente la mobilità della proteina, impedendone la denaturazione o lo svolgimento nel tempo [10].
La vitrificazione è agevolata dal fenomeno della crio-concentrazione che è inevitabile durante la liofilizzazione, perciò risulta un meccanismo particolarmente efficace e conveniente in presenza di un adeguato eccipiente.

1.3.3 Protezione da denaturazione superficiale

La denaturazione indotta da interfacce o superfici è un problema insito nella produzione di biofarmaci liofilizzati, poiché la loro presenza è inevitabile. Gli additivi da utilizzare in questo contesto sono i surfattanti, i quali possono sostanzialmente agire secondo due possibili modalità, a seconda dell'affinità chimica delle specie coinvolte. Da un lato, possono disporsi in prossimità della superficie stessa e ricoprirla, supponendo che la concentrazione sia sufficientemente elevata, di conseguenza impedendo alla proteina di adsorbirsi e denaturarsi. Alternativamente, possono interagire direttamente con la proteina, disponendosi attorno agli ammi-

noacidi inclini ad essere adsorbiti o a svolgersi e favorendo la forma nativa.

Individuare surfattanti adeguati è un processo cruciale e delicato, poiché essi possono anche stabilizzare forme parzialmente o completamente svolte della struttura proteica, sia in base alla compatibilità chimica, sia in funzione dell'orientazione assunta da queste molecole [4].

1.3.4 Ciclodestrine

Si ritiene opportuno dedicare un intero paragrafo alle ciclodestrine, in quanto gli eccipienti utilizzati durante lo svolgimento della Tesi appartengono a questa classe di composti chimici.

Le ciclodestrine (da qui in poi abbreviate in *CD*) sono oligosaccaridi composti da più unità di glucosio legate tra loro per formare una struttura circolare chiusa. Tipicamente il numero di unità ripetitive varia tra 6 ed 8, formando rispettivamente le α-*CD*, β-*CD*, γ-*CD*, ognuna di dimensioni progressivamente maggiori e proprietà anche significativamente diverse.
La peculiarità di queste sostanze chimiche consiste nella presenza di una cavità nella propria struttura, che può essere sfrutta per diversi scopi (ad esempio, nel drug delivery per inglobare selettivamente le molecole terapeutiche e rilasciarle in maniera controllata) [6]. Tale caratteristica risulta vantaggiosa se si conoscono le proprietà ed affinità dei composti coinvolti, poiché da esse dipende l'efficacia della somministrazione. In particolare, mentre le α-*CD* possiedono una cavità generalmente troppo piccola per poter ospitare le molecole di interesse, le γ-*CD* spesso risultano non abbastanza selettive verso le specie chimiche da inglobare. Per questa ragione spesso si sceglie di operare con le β-*CD*, le quali tuttavia presentano una solubilità significativamente inferiore rispetto alle altre due *CD* [11], limitandone la possibilità di impiego.
Per ovviare a questa problematica e per fornire alle molecole proprietà *ad hoc* per le applicazioni a cui sono destinate, le *CD* vengono spesso funzionalizzate. Alcuni esempi sono costituiti dalle idrossipropil-β-ciclodestrine, metil-β-ciclodestrine, sulfobutileter-β-ciclodestrine. In generale, il grado di sostituzione dei gruppi ossidrile da parte di altri gruppi funzionali in ciascuna *CD* varia e, di conseguenza, anche le proprietà chimico-fisiche [12]. Infatti, durante il processo di sintesi risulta difficile ottenere molecole esattamente identiche tra loro, aspetto da tenere in considerazione durante la progettazione della formulazione.
In questa trattazione si farà sempre riferimento a molecole nelle quali tutte le unità sono identicamente sostituite, sia in termini di posizione sia di numerosità.

Nell'ambito di questa Tesi l'attenzione sarà focalizzata sulla β-ciclodestrina in forma non sostituita (β-*CD*) e sulla 2-idrossipropil-β-ciclodestrina (da qui in poi chiamata *HP*-β-*CD*). Inoltre, saranno anche utilizzate le stesse β-*CD* nella corrispondente forma a catena aperta, ovvero il maltoeptaosio ed il relativo derivato con funzione idrossipropilica (da qui in poi rispettivamente open-β-*CD* ed open-*HP*-β-*CD*). In particolare, l'aggiunta del gruppo funzionale idrossipropile risulta in una solubilità maggiore rispetto alla molecola nativa [6].

Le ciclodestrine menzionate sono rappresentate in Figura 1.1.

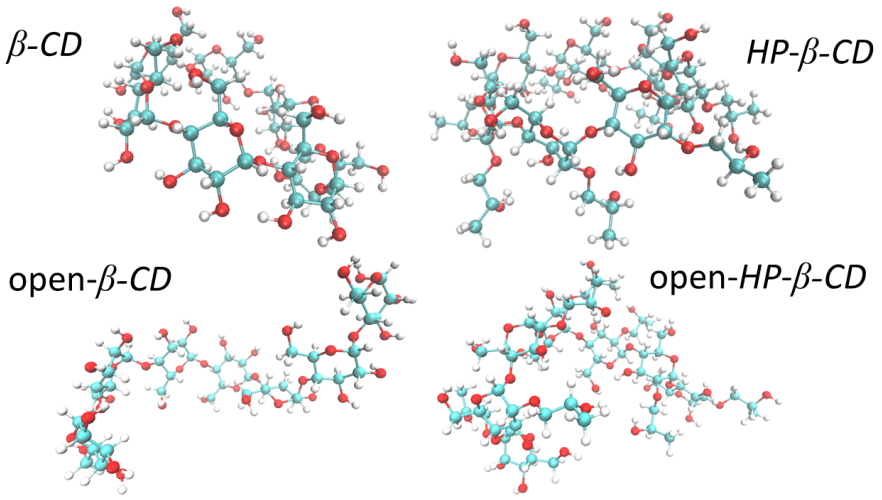

Figura 1.1: Strutture molecolari delle ciclodestrine utilizzate. Immagine generata utilizzando il software VMD [13].

Le β-ciclodestrine a catena chiusa possiedono una struttura molecolare caratteristica, che conferisce loro proprietà peculiari. In particolare si è osservata una significativa idrofobicità all'interno della propria cavità, la quale spiega la sua capacità di interagire significativamente con i residui peptidici [12], eventualmente inglobandoli nella cavità, e superfici apolari [14].
La loro forma geometrica può ricordare quella di un tronco di cono: da questa analogia deriva la nomenclatura per gli atomi di ossigeno corrispondenti alle due "basi", ovvero *primary rim* (base minore) e *secondary rim* (base maggiore), raffigurati in Figura 1.2. Tale peculiarità geometrica sarà utile nei seguenti capitoli per comprendere il comportamento delle *CD* utilizzate. Si fa infine notare che il *secondary rim* delimita idealmente e rappresenta l'estremità della cavità idrofobica [15].

Le β-ciclodestrine sono state testate come eccipienti per diversi sistemi, con risultati variabili. Tuttavia, nel caso della *HP-β-CD*, vi sono stati riscontri positivi nella prevenzione dell'aggregazione tra proteine e dell'adsorbimento (con conseguente denaturazione) in presenza dell'interfaccia acqua-aria [14]. Inoltre questa molecola è stata anche approvata e ritenuta idonea per le formulazioni di biofarmaci destinati alla somministrazione [6].
Ulteriori informazioni riguardanti le proprietà tossicologiche della *HP-β-CD* possono essere reperite in letteratura [16].

Infine, un'altra applicazione delle *CD* consta nell'impiego di queste molecole per sistemi in grado di agevolare il ripiegamento corretto di proteine in parte svolte (più semplicemente noto come *refolding*), particolarmente utile nei casi in cui le alterazioni della struttura proteica siano reversibili [6].

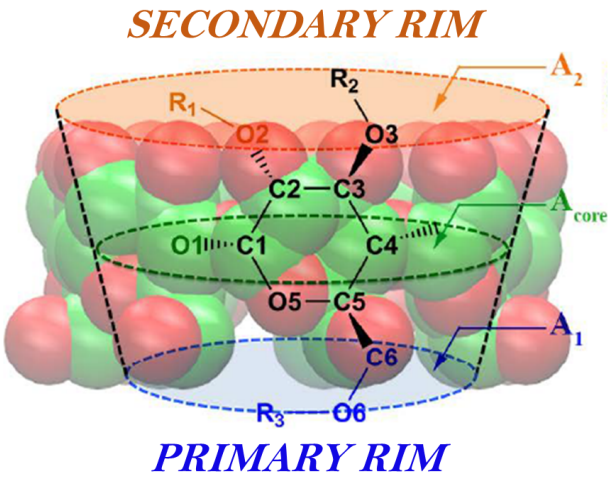

Figura 1.2: Rappresentazione dei *rim* caratteristici delle ciclodestrine. Le ciclodestrine utilizzate sono costituite da unità ripetitive come quella raffigurata, dove R_1 ed R_2 sono atomi di idrogeno mentre R_3 può essere un idrogeno o una funzione idrossipropilica, nel caso si tratti rispettivamente della β-*CD* o della *HP*-β-*CD*. Le varie unità sono collegate da legami α-1,4-glicosidici instaurati tra O1 e C4. Immagine tratta da [15] con modifiche.

1.4 Dinamica Molecolare

In questo lavoro di Tesi si utilizzerà la Dinamica Molecolare per studiare le interazioni tra le ciclodestrine e ad alcune proteine modello.

La Dinamica Molecolare, o Molecular Dynamics (da qui in in poi MD), è una tecnica computazionale che consiste nel simulare l'evoluzione di un sistema di molecole in un arco temporale prefissato. Per fare ciò è necessario descrivere la struttura di ciascuna molecola, indicando la posizione, la tipologia, le distanze e gli angoli dei legami chimici, nonché gli angoli torsionali qualora la molecola non fosse planare.
L'insieme di queste informazioni costituisce la topologia del sistema. In seguito, si sceglie un modello fisico-matematico adeguato per descrivere e stimare i potenziali, che racchiudono i contributi quantistici delle interazioni sia intramolecolari sia intermolecolari. L'insieme di questi modelli costituisce il cosiddetto *force-field*, il campo di forze presenti nel sistema. Una scelta piuttosto comune, sebbene non sempre sufficientemente esaustiva, consiste nell'impiego del potenziale armonico per rappresentare i legami chimici e del potenziale di Lennard-Jones (Eq. 1.1) per le interazioni fisiche.

$$V^{LJ}(r_{ij}) = 4\varepsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right) \tag{1.1}$$

dove r_{ij} è la distanza tra due molecole, atomi o particelle i e j, σ il diametro di

collisione, ε l'energia di interazione.
Oppure si utilizza la formula di Coulomb qualora siano coinvolti degli ioni (Eq. 1.2) [17].

$$V^C(r) = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r} \tag{1.2}$$

dove r è la distanza tra le due cariche Q_1 e Q_2, ed ε_0 la permittività nel vuoto.

Si procede quindi creando un dominio computazionale, comunemente definito "box", di dimensioni prestabilite, al cui interno si collocano le molecole di interesse, per poi solvatare il sistema. Durante questa fase preliminare la posizione delle molecole è stabilita arbitrariamente dal software (per questa Tesi si fa riferimento a GROMACS [18]), perciò il sistema va portato in condizioni di equilibrio dal punto di vista fisico, in modo da rispettare il bilancio tra le interazioni molecolari. Se inizialmente si posizionassero le molecole senza tenere conto delle forze in gioco si potrebbero infatti verificare repulsioni virtualmente infinite, mentre in realtà questo non succede poiché i contributi attrattivi e repulsivi tendono a compensarsi nel tempo.

Infine, è possibile simulare l'evoluzione del sistema nelle condizioni fisiche di interesse, inizializzando la simulazione con un opportuno canone di riferimento *NPT* o *NVT*, ovvero a numero di molecole, temperatura, e rispettivamente pressione o volume costanti [19]. A questo scopo si scelgono gli algoritmi di integrazione (barostati e termostati, di cui si parlerà nei capitoli successivi) ritenuti più adatti per mantenere costanti nell'arco della simulazione le condizioni del canone.

Noti i potenziali di interazione si calcolano le forze corrispondenti e, conoscendo la massa delle molecole, si risale agevolmente all'accelerazione mediante la seconda legge di Newton. Quindi, integrando nel tempo, è possibile calcolare la traiettoria delle stesse. Per fare ciò spesso si ricorre all'algoritmo di Verlet [17], di cui esistono numerose varianti, da utilizzarsi in base alle specifiche esigenze dell'utente e dell'applicazione.
La versione di default utilizzata da GROMACS è la cosiddetta *leap-frog* [20] [18]. Per ulteriori informazioni sulle alternative disponibili si veda il Manuale dell'Utente di GROMACS per la versione in uso, 2018.1 (la documentazione in questione è reperibile all'indirizzo http://manual.gromacs.org/documentation/2018.1/manual-2018.1.pdf).

$$\mathbf{v}_i(t + \frac{1}{2}\Delta t) = \mathbf{v}_i(t - \frac{1}{2}\Delta t) + \frac{\mathbf{F}_i(t)}{m_i}\Delta t \tag{1.3}$$

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\mathbf{F}_i(t)}{m_i}\Delta t^2 + O(\Delta t^4) \tag{1.4}$$

dove $\mathbf{v}_i$, $\mathbf{r}_i$, $\mathbf{F}_i$, m_i sono rispettivamente la velocità, la posizione, la risultante delle forze e la massa dell'entità i (molecola, atomo o particella), mentre Δt è il passo di integrazione relativo alla discretizzazione temporale.

Per completare il modello matematico è necessario imporre delle condizioni al contorno. Utilizzando canoni a numero di molecole costante si impongono

le cosiddette *periodic boundary conditions* [17] [18] (si veda Fig. 1.3), le quali consentono di estendere virtualmente la simulazione a repliche identiche ed equivalenti dello stesso sistema affiancate in parallelo, motivo per cui ad esempio per ogni molecola uscente dal sistema ve ne sarà un'altra con le stesse caratteristiche in ingresso.

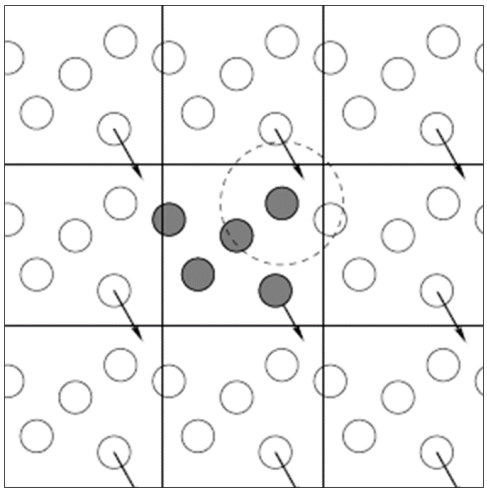

Figura 1.3: Schematizzazione delle *periodic boundary conditions*. Immagine tratta da [17] con modifiche.

1.5 Obiettivi della Tesi

In questa Tesi si propone un approccio alternativo a quello sperimentale per investigare le proprietà degli eccipienti, sfruttando metodi computazionali per la valutazione delle interazioni tra le molecole in esame ed alcune proteine destinate ad uso terapeutico. Come è stato evidenziato nei paragrafi precedenti, sono molteplici le sostanze necessarie per ottenere una formulazione efficace, in grado di preservare la qualità della proteina durante la produzione e la successiva *shelf-life*. Pertanto, testare l'efficacia di ciascuna di esse individualmente e, successivamente, di tutte le componenti, richiede un notevole sforzo a livello di prove sperimentali, quindi in termini di tempo, denaro e risorse. Inoltre, cambiando la proteina in esame, potrebbe essere necessario cambiare formulazione, richiedendo ulteriori esperimenti e perciò investimenti.

La Dinamica Molecolare risulta particolarmente utile quando non esistono tecniche sperimentali in grado di fornire, in maniera più o meno diretta, le informazioni a cui si è interessati, o qualora siano troppo complesse o costose, oppure nell'eventualità in cui le condizioni da studiare risultino proibitive [17]. Inoltre, essa rappresenta anche uno strumento utile per stimare parametri chimico-fisici di interesse per le scienze applicate, ad esempio diffusività, viscosità, conducibilità,

calori specifici, potenziali chimici e costanti cinetiche.

Lo sviluppo dell'informatica e dell'elettronica nell'ultimo decennio ha reso sempre più conveniente l'utilizzo di metodi computazionali come supporto alle attività di laboratorio tradizionali, apportando metodologie caratterizzate da maggiore versatilità e flessibilità in termini di condizioni operative, soprattutto nell'ambito di ricerca e sviluppo dei processi. In quest'ottica l'utilizzo di MD risulta decisamente vantaggioso, considerata la facilità con cui è possibile variare condizioni del sistema quali temperatura e concentrazione dei soluti, consentendo di simulare numerose configurazioni differenti in tempi relativamente ridotti e di ottenere risultati altamente riproducibili.

D'altra parte, per effettuare simulazioni che corrispondano alla realtà è indispensabile avere a disposizione topologie e *force-field* accurati, non sempre disponibili, nonostante lo sviluppo e la progressiva diffusione di banche dati dedicate. Una problematica rilevante è costituita dalla compatibilità tra i *force-field* ed il sistema da simulare [21], poiché non esistono, almeno ad oggi, modelli universali adatti ad ogni evenienza.

Fortunatamente, nell'ambito dei sistemi considerati in questa Tesi esistono modelli adatti per effettuare simulazioni di MD, i quali sono stati convalidati sperimentalmente e perciò consentono di ottenere risultati realistici, come ad esempio è stato osservato in [7]. Ciò rende possibile lo studio dei comportamenti molecolari alla base dei meccanismi di protezione in atto nelle formulazioni destinate alla liofilizzazione.

I prossimi capitoli si focalizzeranno pertanto sulla simulazione di alcuni sistemi a solvente acquoso contenenti eccipienti, nello specifico le β-ciclodestrine, in presenza di interfacce, singoli amminoacidi e proteine di interesse, cercando di estrarre dati utili riguardanti le modalità di interazione e, se possibile, di confermare o integrare le informazioni sperimentali. In particolare, si intende verificare se le *CD* possiedano proprietà simili a quelle dei surfattanti in prossimità delle interfacce tipiche della liofilizzazione (acqua-aria ed acqua-ghiaccio), e siano quindi capaci di prevenire l'adsorbimento e la relativa denaturazione.
Inoltre, se ne studierà il comportamento in termini di esclusione o inclusione preferenziale per la prevenzione dell'aggregazione delle proteine in soluzione.

Capitolo 2

Interfacce acqua-aria ed acqua-ghiaccio

La presenza delle interfacce gioca un ruolo cruciale nella stabilità dei farmaci a base proteica, a causa dei possibili meccanismi di destabilizzazione associati, come discusso nel paragrafo §1.2.1. Per scongiurare tale problema si fa affidamento ad eccipienti che, se scelti opportunamente, consentono di contrastare i danni legati alle superfici. In questo contesto si vuole pertanto usare la Dinamica Molecolare come strumento per investigare le proprietà delle ciclodestrine in presenza di interfacce. In particolare, le superfici analizzate sono l'interfaccia acqua-aria ed acqua-ghiaccio, i cui contributi sono cruciali durante la preparazione della soluzione liquida e la successiva liofilizzazione dei biofarmaci.

2.1 Interfaccia Acqua-Aria

L'estensione dell'interfaccia acqua-aria può essere limitata durante la manipolazione dell'ingrediente attivo, tuttavia risulta difficile o poco praticabile rimuoverla completamente; è pertanto necessario l'impiego di opportuni eccipienti. Le proprietà delle ciclodestrine in presenza di tale superficie sono state studiate dettagliatamente, con diversi riscontri positivi riguardo alle loro potenzialità, sia come eccipienti per formulazioni [14] sia come surfattanti generici, anche in alcune forme derivate [22]. In particolare, nel caso delle ciclodestrine a catena chiusa, la cavità che le contraddistingue ha implicazioni cruciali sul carattere idrofobico atteso e di conseguenza sull'affinità nei confronti dell'interfaccia in questione, vista la natura apolare dell'aria. Per questa ragione si ritiene importante verificare la presenza di un'eventuale orientamento preferenziale. Infine, la formazione di *cluster* è un ulteriore aspetto che, seppur di importanza marginale, contribuisce a chiarire il quadro delle proprietà delle ciclodestrine come eccipienti e perciò merita di essere investigato adeguatamente. Nel paragrafo seguente si descriverà la procedura di simulazione utilizzata per studiare questa interfaccia.

2.1.1 Metodologia

Scelta del Force-Field

Per poter simulare un qualsiasi sistema mediante Dinamica Molecolare sono necessari alcuni passaggi preliminari, come è stato discusso nel paragrafo §1.4. In tutte le simulazioni trattate nell'ambito di questa Tesi si fa riferimento alle ciclodestrine, pertanto è indispensabile avere a disposizione la topologia di tali molecole. In generale le topologie possono essere reperite da opportune banche dati oppure essere definite dagli utenti stessi, qualora non risultino disponibili. Le topologie delle *CD* sono state modellate tenendo conto di alcuni risultati resi noti in letteratura [23] [15] [24]. Ciò è particolarmente importante poiché, per simulare nel modo più accurato possibile, si devono utilizzare topologie che rispecchino l'effettiva struttura molecolare, tenendo conto di tutti gli aspetti fisici che caratterizzano e definiscono univocamente ogni sostanza chimica di interesse.
Maggiore è la complessità strutturale di una molecola, maggiore la difficoltà nel riprodurla accuratamente, basti pensare a sostanze quali le proteine, spesso caratterizzate da un considerevole numero di atomi e la cui struttura ha un ruolo determinante nello svolgimento dell'attività biologica.

Le topologie delle ciclodestrine a catena chiusa sono state opportunamente modificate per ottenere le corrispondenti *CD* a catena aperta. Nel caso specifico delle *CD* in presenza dell'interfaccia acqua-aria le topologie richieste sono fondamentalmente quelle degli eccipienti stessi e dell'acqua, la quale può essere descritta in modi diversi, in base alle esigenze, a seconda dalla scelta del *force-field*.
L'approccio adottato in questa Tesi è detto all-atom, ovvero tutti gli atomi di ogni molecola (e le relative proprietà) vengono trattati in modo esplicito; tuttavia è anche possibile scegliere di rappresentare gruppi di atomi come uniche particelle, le cui proprietà e caratteristiche vengono determinate secondo opportuni criteri. Tale scelta, effettuata allo scopo di ridurre il carico computazionale, è adatta per la simulazione di sistemi contenenti elevati numeri di molecole, talvolta anche con strutture complesse, eventualmente utilizzando geometrie semplificate qualora esse non risultino di interesse primario. Quest'ultimo approccio viene chiamato *coarse-grained* [25].

Una volta raccolte le topologie delle sostanze chimiche di interesse, si procede alla scelta del *force-field*. Nel paragrafo §1.4 sono state elencate brevemente le informazioni racchiuse in questi modelli fisici. La scelta del *force-field* è di cruciale importanza poiché da esso dipende la traiettoria risultante: le simulazioni di Dinamica Molecolare sono basate sulle forze agenti su ogni componente del sistema, da cui si ottiene l'accelerazione e quindi la posizione tramite l'algoritmo di Verlet [20]. Risulta pertanto evidente come l'accuratezza nella descrizione del campo di forze influenzi l'evoluzione della traiettoria nel tempo, perciò è necessario scegliere un *force-field* adeguato al tipo di applicazione o al tipo di sostanze di interesse.
Nell'ambito di questa Tesi il *force-field* utilizzato è il CHARMM [26], di cui esi-

stono diverse varianti: in particolare si adotta il CHARMM36, adatto per simulare carboidrati costituiti da più unità di esopiranosio legate da legami glicosidici, quali le ciclodestrine in esame [27]. Inoltre, tale *force-field* è stato sviluppato anche per la simulazione di proteine [28], eventualmente anche in presenza di carboidrati o loro derivati [29], pertanto risulta adeguato nel contesto delle analisi svolte in questa Tesi.
L'impiego del CHARMM è stato convalidato per ciclodestrine in soluzione [21].

Procedura di Simulazione

Una volta scelto il force-field, si procede con la composizione del dominio computazionale, inserendo le molecole di interesse. Per fare ciò si utilizza il seguente comando:

gmx insert-molecules -ci CD.gro -nmol N -box X Y Z -o CDbox.gro

dove *CD.gro* rappresenta il file che contiene la configurazione della ciclodestrina in questione, *N* indica il numero di molecole da inserire, *X*, *Y* e *Z* costituiscono le dimensioni del box, *CDbox.gro* il file di output contenente le molecole nel volume di dimensioni prestabilite. Per studiare l'interfaccia acqua-aria il volume di liquido simulato è composto da un cubo di dimensioni 8 nm x 8 nm x 8 nm. Il numero di *CD* presenti corrisponde ad una concentrazione pari a 50mM, valore riscontrato in alcune formulazioni impiegate in ambito farmaceutico con risultati positivi in termini di stabilizzazione [6]. Per il sistema in questione *N* è pari a 15. A questo punto, l'unica specie chimica mancante è l'acqua, perciò si prosegue con la solvatazione del box utilizzando il seguente comando:

gmx solvate -cp CDbox.gro -cs spc216.gro -p topol.top -o CDwbox.gro

dove *CDbox.gro* è il file di input, il file *spc216.gro* indica la configurazione dell'acqua, in questo caso esplicita, *single point charge* [30], *topol.top* racchiude informazioni riguardanti topologie, *force-field* e numero di specie chimiche nel sistema, e infine *CDwbox.gro* il box in output solvatato. La topologia di riferimento per l'acqua in questa Tesi è la *TIP3P* [31].

Come accennato nel paragrafo §1.4, il box computazionale viene riempito posizionando le molecole senza tenere conto di attrazioni e repulsioni, pertanto è necessario equilibrare il sistema, ovvero modificarne la configurazione affinché l'energia sia minimizzata. Tale operazione viene tipicamente effettuata mediante l'algoritmo di massima discesa del gradiente (o *steepest descent algorithm*).
Le interazioni elettrostatiche a lungo raggio sono stimate mediante il metodo *Particle Mesh Ewald* o, più comunemente, *PME* [32].
La minimizzazione energetica si effettua in due passi:

gmx grompp -f enermin.mdp -c CDwbox.gro -o em.tpr -p topol.top

un passo preliminare di preparazione, dove *CDwbox.gro* è il file di input, *enermin.mdp* racchiude i parametri per l'algoritmo di minimizzazione, *topol.top* racchiude le informazioni riguardanti topologie, *force-field* e numero di specie chimiche nel sistema, ed *em.tpr* è il file di output.
Il secondo passo è:

gmx mdrun -v -deffnm em

durante il quale si ha l'effettiva generazione di una traiettoria, ovvero la simulazione di un transitorio al termine del quale si ottiene un file di output, *em.gro*, contenente una configurazione del sistema la cui energia è stata minimizzata.

Prima di poter simulare il sistema è necessario portare il sistema alle condizioni fisiche di interesse, in modo tale che eventuali fluttuazioni delle stesse nel tempo siano sufficientemente ridotte da poter essere trascurate. Durante questa operazione, comunemente detta equilibratura, si fa riferimento al canone *NPT* [19], pertanto il sistema attraversa una fase transitoria per portarsi a temperatura e pressione desiderate.
Per fare ciò si ricorre ad alcuni algoritmi dedicati, rispettivamente denominati termostati e barostati. Il numero di molecole rimane inalterato, in modo da soddisfare le *periodic boundary conditions*. La distribuzione delle velocità delle molecole è di tipo Maxwell-Boltzmann e viene generata in modo casuale.
L'equilibratura per i sistemi acqua-aria è stata svolta in due fasi, ognuna delle quali è articolata in due passi, analogamente alla minimizzazione:

gmx grompp -f npt.mdp -c em.gro -o npt.tpr -p topol.top
gmx mdrun -v -deffnm npt

In questa fase si può ragionevolmente supporre che il sistema sia lontano dalle condizioni di pressione e temperatura desiderate, pertanto si utilizzano algoritmi meno aggressivi, ovvero con velocità di convergenza minore. L'equilibratura iniziale è stata svolta per 10 ns, con un time-step pari a 2 fs, utilizzando l'accoppiamento tra termostato e barostato di Berendsen [33], ad 1 bar e 300 K. La distribuzione delle velocità viene generata in modo causale.
La seconda fase, avviata tramite i comandi:

gmx grompp -f npt2.mdp -c npt.gro -o npt.tpr -p topol.top -t npt.cpt
gmx mdrun -v -deffnm npt2

é invece svolta utilizzando algoritmi più aggressivi, ovvero con velocità di convergenza maggiore, poiché il sistema risulta più vicino alle condizioni fisiche desiderate. L'equilibratura in questione è stata svolta per 5 ns, con un time-step pari a 2 fs, utilizzando il termostato di Nosé [34] [35] e Hoover [36], ed il barostato di Parrinello e Rahman [37], ad 1 bar e 300 K. In questo caso la distribuzione delle velocità è la stessa generata durante la fase precedente, fornita in input mediante

l'opzione *-t npt.cpt*. Il risultato finale è il file *npt2.gro*, contenente la configurazione del sistema equilibrata a 300 K ed 1 bar.

Dopo aver effettuato l'equilibratura è possibile monitorare l'evoluzione di temperatura, pressione ed energia del sistema, nell'arco di tempo simulato, per verificare che le fluttuazioni delle grandezze in questione siano state ridotte a sufficienza. Ciò è possibile utilizzando i seguenti comandi:

gmx energy -f npt2.edr -o T.xvg
gmx energy -f npt2.edr -o P.xvg
gmx energy -f npt2.edr -o energy.xvg

dove il file con estensione *.edr* contiene le informazioni relative all'energia del sistema simulato, mentre l'opzione *-o* consente di generare un output contenente il profilo temporale della grandezza desiderata.
Qualora i risultati non fossero soddisfacenti, è possibile estendere la durata dell'equilibratura, aumentando il numero di step nel file *.mdp*, per poi riprendere la simulazione con il seguente comando:

gmx mdrun -v -deffnm npt2 -cpi

dove l'opzione *-cpi* consente di far ripartire la simulazione dall'ultimo step, senza doverla riavviare da capo.

Una volta equilibrata la configurazione del sistema è possibile aggiungere uno strato di aria, di spessore a piacere, in modo da poter simulare propriamente l'interfaccia acqua-aria. In particolare, per fare ciò è necessario modificare le dimensioni del box, inserendo uno spazio vuoto per rappresentare l'aria. Convenzionalmente si estende la dimensione Z del box, mediante il comando:

gmx editconf -f npt2.gro -box 8 8 10 -c 4 4 5 -o airbox.gro

dove l'opzione *-box* consente di estendere la coordinata Z da 8 nm a 10 nm (introducendo uno spessore di aria pari a 2nm), mentre l'opzione *-c* permette di impostare le coordinate del centro del box nel nuovo centro geometrico. Si ottiene quindi in output il file desiderato, *airbox.gro*. E' consuetudine verificare che tale operazione sia stata svolta con successo visualizzando il sistema mediante software quali VMD [13] o PyMOL [38].

Per simulare l'interfaccia acqua-aria non è possibile utilizzare la stessa modalità vista finora, infatti lo strato di vuoto che corrisponde all'aria verrebbe occupato dall'acqua presente nel box se si utilizzasse nuovamente il canone *NPT*. Pertanto, si fa riferimento ad un altro canone, ovvero l'*NVT* [19], il quale mantiene il volume costante anziché la pressione, quindi consente variazioni di pressione per preservare la presenza dell'aria.
Cambiare il canone di riferimento significa operare in condizioni fisiche differenti; perciò, prima di avviare la simulazione finale, è preferibile condurre una breve equilibratura a 300 K, a volume costante, per 1 ns, con un time-step pari a 2 fs.
Si usano i seguenti comandi:

gmx grompp -f md-init.mdp -c airbox.gro -o md-init.tpr -p topol.top
gmx mdrun -v -deffnm md-init

dove il file *md-init.mdp* specifica il nuovo algoritmo utilizzato per la termostatazione, il V-rescale [39], mentre il barostato è disattivato, essendo il canone utilizzato l'*NVT*. Inoltre, è necessario generare nuovamente la distribuzione di velocità, poiché quella precedente corrisponde ad un sistema in condizioni fisiche differenti. Si ritiene opportuno verificare l'evoluzione temporale del volume durante tale transitorio, in modo da stabilire se il sistema si sia portato in condizioni isocore. Analogamente a quanto visto in precedenza, si procede nel modo seguente:

gmx energy -f md-init.edr -o V.xvg

dove il file *V.xvg* contiene il profilo temporale del volume.

Infine, è possibile effettuare la simulazione dell'interfaccia acqua-aria nelle condizioni desiderate. La durata della simulazione è di 60 ns, con un time-step pari a 2 fs, a 300 K.
Si procede mediante i seguenti comandi:

gmx grompp -f md.mdp -c md-init.gro -o md.tpr -p topol.top -t md-init.cpt
gmx mdrun -v -deffnm md

dove l'opzione *-t md-init.cpt* fornisce la distribuzione di velocità generata durante l'equilibratura *NVT*, mentre i file di output *md.gro* e *md.xtc* contengono, rispettivamente, la configurazione del sistema nell'ultimo istante della simulazione e la traiettoria delle molecole.
I file *.gro* e *.xtc* possono essere visualizzati tramite software quali VMD.

Analisi dei Risultati

Per studiare e comprendere le proprietà delle ciclodestrine in presenza dell'interfaccia acqua-aria si analizza il profilo di densità nel dominio computazionale. Questo tipo di informazione può essere estratto dalla traiettoria mediante il seguente comando:

gmx density -f md.xtc -s md.tpr -center -symm -b 20000 -sl 500 -o
density-profile.xvg

dove le opzioni *-center* e *-symm* consentono di discretizzare il sistema in maniera simmetrica rispetto al centro geometrico del box, utilizzando un numero di intervalli stabilito mediante l'opzione *-sl*. Infine, ai fini della stima del profilo in output, *density-profiles.xvg*, si trascurano i primi 20 ns della simulazione (*-b 20000*).

Per confrontare i profili di specie chimiche differenti è necessario un ulteriore passaggio, ovvero normalizzare i valori dividendoli per la concentrazione massica

del composto di interesse nel sistema, ottenendo un profilo di densità adimensionata. Ciò è particolarmente importante quando si confrontano sostanze chimiche aventi masse molari significativamente diverse. Infine, per studiare eventuali orientazioni preferenziali da parte della ciclodestrine, si identificano gli atomi di ossigeno che costituiscono i *rim* caratteristici delle *CD*, definendoli come gruppi in un file indice. I file indice contengono il numero sequenziale di ciascun atomo presente nel sistema, raggruppati per specie chimica oppure in base ad altri criteri scelti dall'utente. Il comando corrispondente è il seguente:

gmx $make_ndx$ *-f md.gro -o index.ndx*

dove *index.ndx* è il file indice in output.

Un ulteriore elemento di interesse è costituito dalla formazione di *cluster* da parte delle ciclodestrine, i quali contribuiscono a chiarire le proprietà delle *CD* come eccipienti. Inizialmente, si procede estraendo dalla traiettoria completa un'altra traiettoria contenente solo le *CD*:

gmx trjconv -f md.xtc -s md.tpr -b 20000 -dt 10 -n index.ndx -o traj-CD.xtc -pbc whole
gmx trjconv -f md.xtc -s md.tpr -b 60000 -e 60000 -n index.ndx -o traj-CD.gro

dove *traj-CD.xtc* è la traiettoria delle *CD* (esclusi i primi 20 ns), mentre *traj-CD.gro* è la configurazione del sistema nell'ultimo istante. In seguito, si genera un nuovo file indice *index-CD.ndx* corrispondente al nuovo sistema in analisi, usando il comando *gmx* $make_ndx$.

gmx $make_ndx$ *-f traj-CD.gro -o index-CD.ndx*

Vi è un ultimo passo preliminare che consiste nel ricavare un file di topologia *CD.tpr*, corrispondente al file *traj-CD.gro*, usando il comando *gmx grompp*, come visto in precedenza.

gmx grompp -f enermin.mdp -c traj-CD.gro -o CD.tpr -p CD.top

A questo punto è infine possibile utilizzare il comando dedicato per l'analisi dei *cluster*:

gmx clustsize -f traj-CD.xtc -s CD.tpr -n index-CD.ndx -mol -hc clust-CD.xvg -nc nclust.xvg

dove l'opzione *-hc clust-CD.xvg* genera un file di output contenente la distribuzione delle ciclodestrine facenti parte di un *cluster*, in funzione delle dimensioni dello stesso, mentre *-nc nclust.xvg* genera in output un file contenente il numero di *cluster* presenti per ogni istante della traiettoria. L'opzione *-mol* è di particolare importanza poiché consente di valutare la formazione di *cluster* su base molecolare anziché atomica, agevolando l'interpretazione e l'analisi dei risultati.

2.1.2 Risultati ed osservazioni

Una volta generati i profili di densità adimensionata secondo la procedura indicata nel paragrafo §2.1.1, è possibile raggruppare tali risultati in un grafico riassuntivo come in Figura 2.1. Da questo grafico risulta evidente una netta distinzione tra le proprietà esibite dalle *HP-β-CD* e dalle *β-CD*: le prime mostrano infatti caratteristiche simili ai surfattanti, adsorbendosi all'interfaccia acqua-aria (si noti il salto di densità per distanze che tendono a zero), mentre le seconde tendono ad accumularsi in bulk. In particolare, il comportamento delle *HP-β-CD* a catena chiusa è già stato osservato sperimentalmente [14]. Inoltre, a sostegno di tale ipotesi, esistono ulteriori conferme derivanti da misure di tensione superficiale, secondo cui soluzioni contenenti *HP-β-CD* esposte all'interfaccia acqua-aria manifestano una diminuzione della tensione all'aumentare della concentrazione e del grado di sostituzione della *CD* in esame [6].

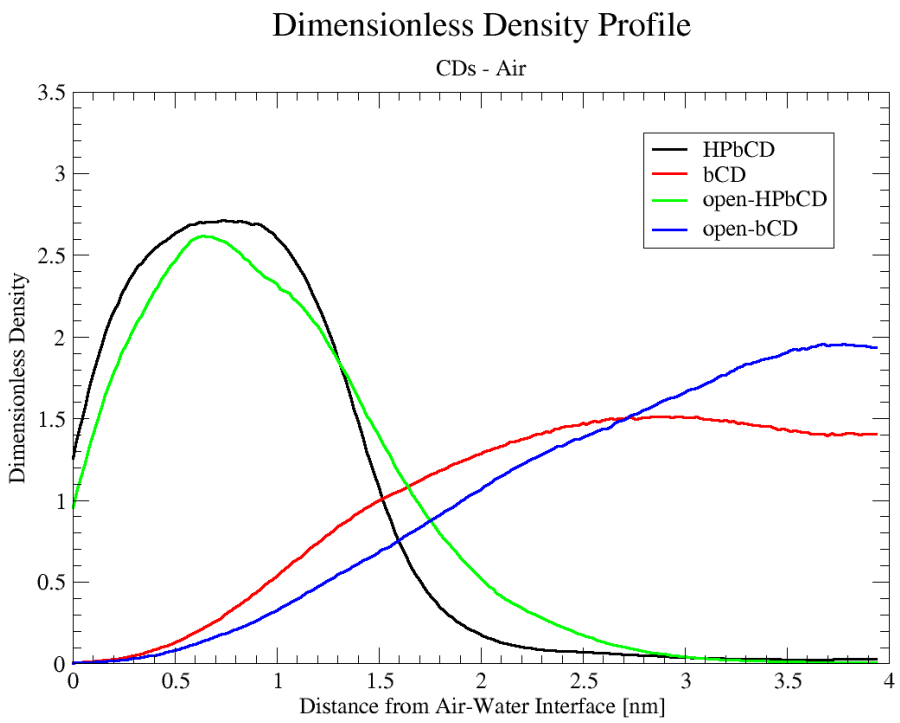

Figura 2.1: Profilo di densità adimensionata delle ciclodestrine in funzione della distanza dall'interfaccia acqua-aria.

Peraltro, questi risultati confermano l'idrofobicità attesa da tali sostanze chimiche.

L'Autore non è a conoscenza di studi sperimentali relativi all'impiego della corrispondente molecola a catena aperta, pertanto ciò costituisce un aspetto che

necessita ulteriori analisi sperimentali.
Per quanto riguarda le β-*CD*, nonostante anch'esse presentino un carattere marcatamente idrofobico, specialmente nei confronti di residui peptidici aromatici (di cui si parlerà approfonditamente nel capitolo successivo) [12], non vi sono interazioni nei confronti dell'interfaccia acqua-aria. Tale comportamento può essere giustificato dalla bassa solubilità della β-*CD*, la quale tende a formare aggregati [11], pertanto esibisce una maggiore affinità verso le *CD* stesse piuttosto che per l'interfaccia. Tale ipotesi risulta pertanto in accordo con il profilo di densità ottenuto. Analogamente a quanto detto in precedenza, l'Autore non è a conoscenza di studi sperimentali incentrati sulle *CD* a catena aperta, pertanto ogni conclusione in merito è prettamente ipotetica.

Nel caso delle *CD* a catena chiusa un aspetto di interesse è costituito dall'orientamento degli eccipienti, specificamente dei *rim*. Per verificare la presenza di eventuali orientazioni preferenziali, si procede generando i profili di densità adimensionata dei *rim*, sovrapponendoli a quelli delle ciclodestrine di appartenenza. L'analisi in questione non viene svolta per le molecole a catena aperta, poiché considerare gli atomi corrispondenti ai *rim* non fornirebbe alcun risultato riguardo all'orientazione. In Figura 2.2 è riportato il profilo di densità della *HP*-β-*CD*, dal quale si può notare un netto orientamento da parte del *secondary rim* verso l'interfaccia acqua-aria. Infatti, il profilo ottenuto per il *secondary rim* mostra un picco di densità a distanza esigua dalla superficie (circa 0.3 nm). Tale risultato è in accordo con le considerazioni espresse in precedenza, poiché questa disposizione indica una chiara interazione tra la cavità idrofobica e l'interfaccia in esame. Diversamente, in Figura 2.3, si può osservare come la β-*CD* si accumuli nel bulk, senza presentare alcuna orientazione preferenziale, in coerenza con le considerazioni precedenti.

I risultati ottenuti per le *HP*-β-*CD* hanno implicazioni significative, innanzitutto perché tali molecole manifestano proprietà analoghe a quelle dei surfattanti, risultando eccipienti di potenziale e concreto interesse per contrastare la denaturazione indotta dall'interfaccia acqua-aria. Inoltre, in altri studi relativi all'impiego di surfattanti come agenti stabilizzanti, è stata evidenziata la relazione tra le capacità di stabilizzazione delle sostanze in questione ed il loro orientamento spaziale [4] [40]. Infine, la tendenza delle *HP*-β-*CD* ad accumularsi in prossimità dell'interfaccia rivolgendo preferenzialmente il *primary rim* verso il bulk è stata riscontrata anche in studi sperimentali focalizzati su surfattanti derivanti dalla β-*CD* a catena chiusa, alla quale sono state legate catene carboniose alifatiche, di lunghezza variabile, in corrispondenza del *secondary rim* [22]. In tal caso le molecole si accumulano all'interfaccia formando uno strato monomolecolare, mentre per le *CD* in esame questo comportamento è stato osservato solo parzialmente.

A scopo illustrativo, si riportano di seguito le raffigurazioni dei sistemi analizzati, (Figure 2.4, 2.5, 2.6 e 2.7) generate dalle traiettorie mediante il software VMD. Ciascun sistema liquido rappresentato è confinato sia superiormente sia inferiormente da uno strato d'aria.

Si considerano, infine, i *cluster* formati dalle ciclodestrine, i quali vengono analizzati estraendo due tipologie di informazioni per ciascun sistema: il numero

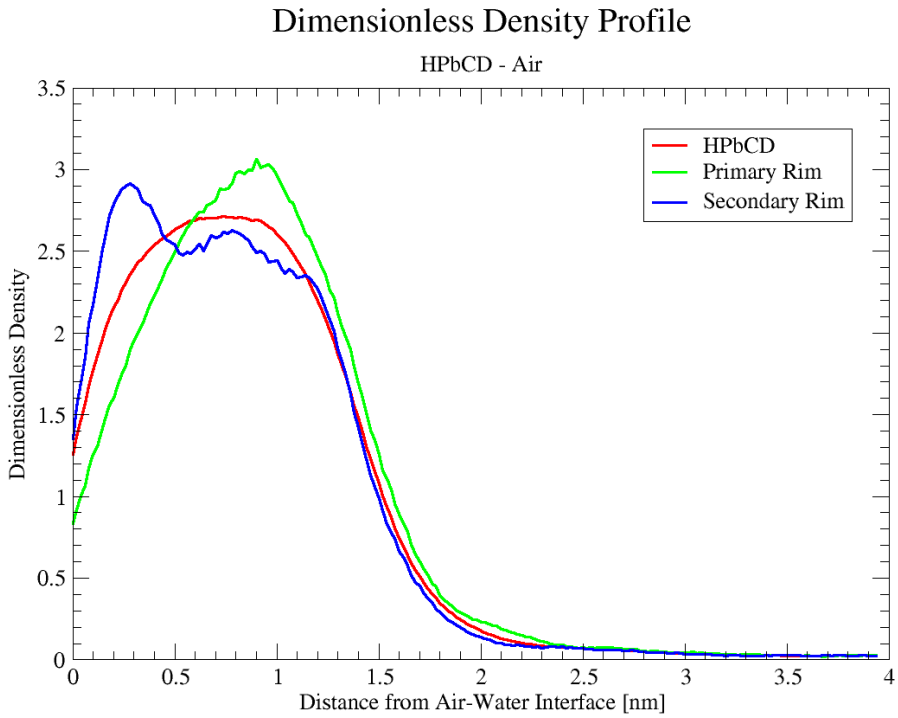

Figura 2.2: Profilo di densità adimensionata della *HP-β-CD* e dei *rim* in funzione della distanza dall'interfaccia acqua-aria.

di *CD* occupate in un *cluster*, in funzione della sua stessa dimensione, ed il numero medio di *cluster* osservati nel tempo.
La tendenza delle *CD* a formare aggregati è notevole e può essere giustificata osservando attentamente la loro struttura chimica: la presenza di gruppi ossidrile in ogni unità di glucopiranosio (e nei gruppi idrossipropile per le *HP-β-CD*) consente l'instaurazione di legami idrogeno intermolecolari, con intensità diverse a seconda dell'affinità tra le specie chimiche coinvolte. Tale caratteristica spiega perché questi eccipienti siano ritenuti ottimi candidati per la stabilizzazione di formulazioni, soprattutto durante le fasi di essiccamento della liofilizzazione, mediante i meccanismi di *water replacement*, *water entrapment* o eventualmente *vitrificazione* [9]. A tale proposito è opportuno distinguere nuovamente il comportamento della *β-CD* da quello della *HP-β-CD*, entrambe a catena chiusa. Infatti, mentre la prima tende a formare più legami idrogeno intramolecolari, soprattutto per il *primary rim* [24], sebbene in presenza di acqua questo comportamento sia attenuato, la seconda instaura interazioni intermolecolari più agevolmente, sia con altre molecole della stessa specie sia con sostanze chimiche differenti, tra cui chiaramente l'acqua [15]. Perciò, la *HP-β-CD* risulta un candidato migliore per i meccanismi

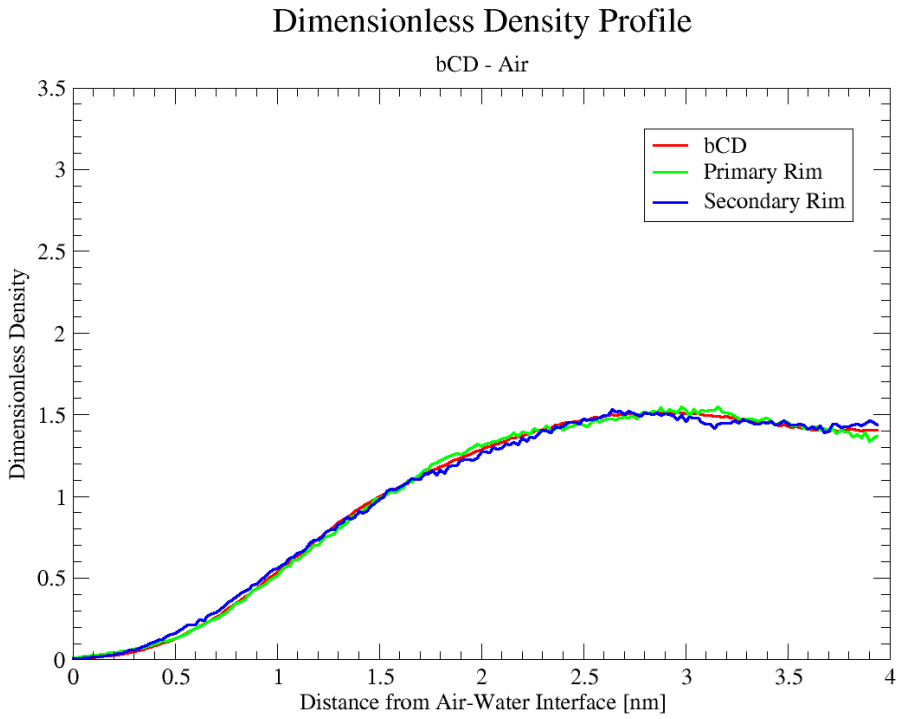

Figura 2.3: Profilo di densità adimensionata della β-*CD* e dei *rim* in funzione della distanza dall'interfaccia acqua-aria.

sopra menzionati [6]. Si precisa nuovamente la mancanza di dati relativi alle *CD* a catena aperta, anche in questo ambito, pertanto le osservazioni riportate di seguito sono prettamente ipotetiche.

Il primo dato che si riporta è la distribuzione del numero di ciclodestrine tra i vari *cluster*, ovvero il numero di *CD* che mediamente fa parte di un *cluster*, in funzione della dimensione di quest'ultimo. Tali informazioni sono racchiuse in Figura 2.8. Risulta difficile trarre conclusioni dal grafico in questione, poiché la popolazione è in continuo cambiamento, perciò il numero di aggregati presenti cambia istante per istante, così come la sua composizione.

Per meglio chiarire le proprietà delle *CD* si è calcolata la dimensione media (in termini di molecole) di un *cluster*, mediante il seguente comando:

gmx analyze -f clust-CD.xvg

dove *clust-CD.xvg* è il file di input, da cui si ottiene la dimensione media di un *cluster* in termini di molecole. I risultati sono riportati in Tabella 2.1. Si può quindi notare che la dimensione media del *cluster* non corrisponde necessariamente alla dimensione avente il massimo di numero di *CD* impegnate.

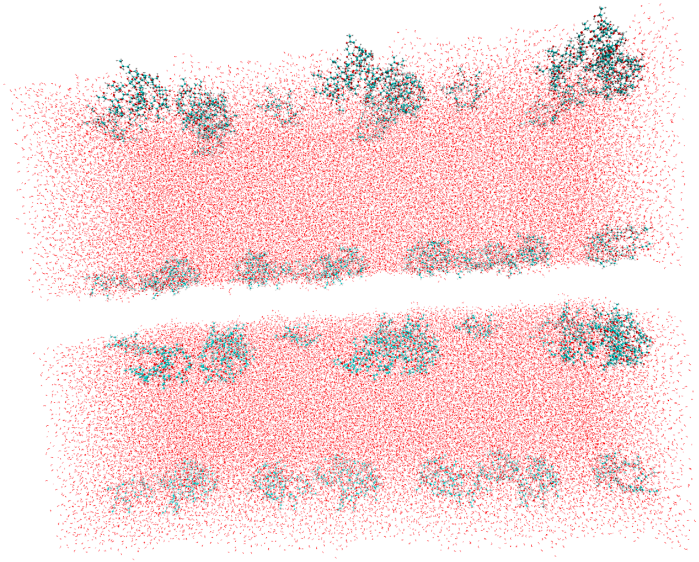

Figura 2.4: Rappresentazione delle *HP-β-CD* nel sistema acqua-aria. Si noti l'accumulo delle CD all'interfaccia. Immagine generata con VMD.

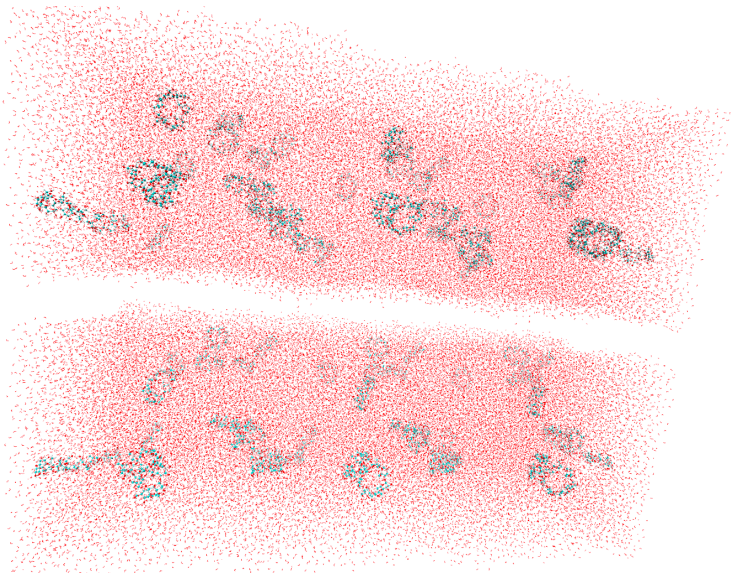

Figura 2.5: Rappresentazione delle *β-CD* nel sistema acqua-aria. Si noti l'accumulo delle CD nel bulk. Immagine generata con VMD.

Infine, per completare il quadro delle informazioni, si stima il numero medio di *cluster* formati nel tempo mediante il comando:

gmx analyze -f nclust.xvg

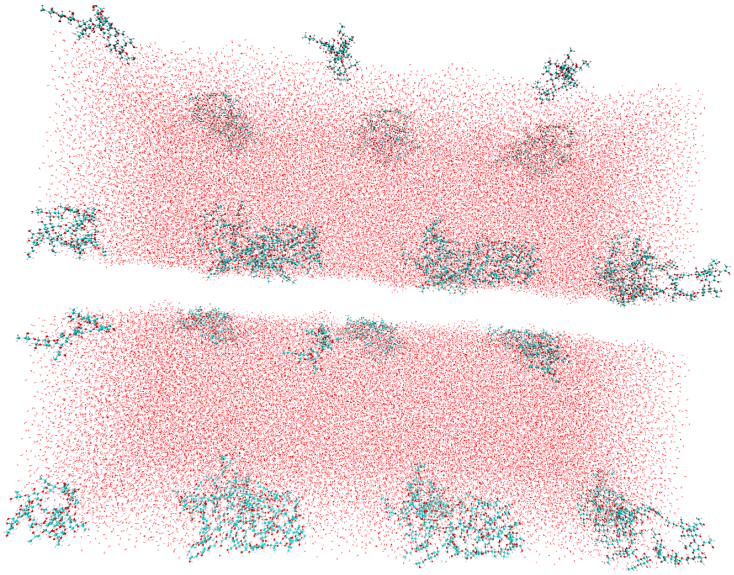

Figura 2.6: Rappresentazione delle open-*HP*-β-*CD* nel sistema acqua-aria. Si noti l'accumulo delle CD all'interfaccia. Immagine generata con VMD.

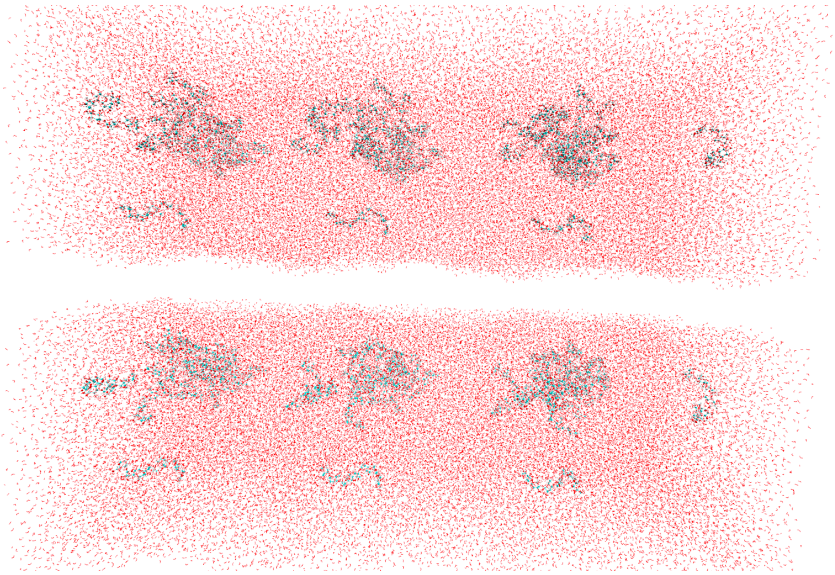

Figura 2.7: Rappresentazione delle open-β-*CD* nel sistema acqua-aria. Si noti l'accumulo delle CD nel bulk. Immagine generata con VMD.

che, a partire dal file *nclust.xvg*, stampa in output il valore medio stimato in funzione del tempo. I risultati così ottenuti sono raggruppati in Tabella 2.2.

Combinando le informazioni a disposizione, è possibile osservare che ai sistemi in cui i *cluster* hanno una dimensione maggiore corrisponde un numero medio di aggregati inferiore e viceversa. In particolare, entrambe le *HP*-β-*CD* sembrano

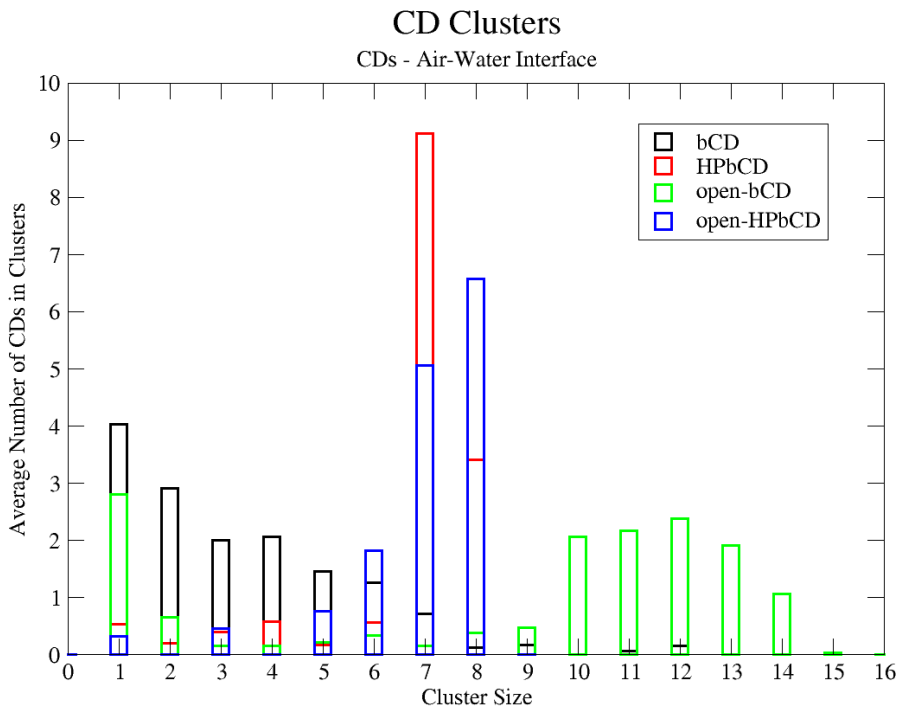

Figura 2.8: Distribuzione della dimensione dei *cluster* per le CD nel sistema acqua-aria.

Tabella 2.1: N° medio di CD in un *cluster* per i sistemi acqua-aria.

HP-β-CD	*β-CD*	open-*HP-β-CD*	open-*β-CD*
6.68	3.65	6.97	9.39

Tabella 2.2: N° medio di *cluster* nel tempo per i sistemi acqua-aria.

HP-β-CD	*β-CD*	open-*HP-β-CD*	open-*β-CD*
2.77	7.33	2.48	4.28

dare un numero di *cluster* inferiore, composti da un numero maggiore di *CD*. Invece, entrambe le *β-CD* tendono a formare un numero di *cluster* maggiore, tuttavia con comportamenti opposti in merito alla dimensione degli aggregati stessi.

Bisogna inoltre considerare il ruolo della struttura delle *CD*, capace di influenzare la propensione a formare legami idrogeno con altre ciclodestrine e/o con l'acqua, effetto particolarmente marcato nel caso della *β-CD* a catena chiusa. Infatti, la presenza di legami intramolecolari ne pregiudica la capacità di interagire con altre molecole, risultando in un numero minore di *cluster*. Nonostante la composizione chimica sia la stessa, le molecole di *β-CD* a catena aperta sembrano in grado

di formare meno *cluster*, tuttavia di dimensioni maggiori, esibendo un comportamento più simile a quello delle *HP-β-CD*. Ciò sottolinea l'impatto della struttura molecolare sulle proprietà della *β-CD* a catena chiusa, la quale sembra mostrare proprietà opposte rispetto alle altre *CD*. Si può infatti escludere il ruolo del gruppo funzionale idrossipropile, in quanto la sua assenza nella open-*β-CD* non ne altera il comportamento qualitativo, pertanto tale differenza sembra essere imputata alla struttura chimica.

Le osservazioni presentate in merito ai *cluster* saranno riprese e discusse nei capitoli successivi, in modo da confrontarne le caratteristiche e cercando di trarre conclusioni più generali, basate su un numero di sistemi più ampio.

2.2 Interfaccia Acqua-Ghiaccio

L'interfaccia acqua-ghiaccio è un'interfaccia tipica nell'ambito della liofilizzazione, la cui presenza è inevitabile vista la natura del processo. Lo studio di tale superficie, delle sue proprietà, nonché dei possibili effetti sulla stabilità del prodotto risulta particolarmente difficile, in quanto la sua estensione cambia continuamente, specialmente durante la fase di congelamento. Sono stati infatti tentati sia approcci sperimentali sia computazionali, tuttavia con risultati limitati, nel primo caso a causa della difficoltà materiale nell'effettuare misure [41], nel secondo per via della mancanza di modelli in grado di tenere conto delle caratteristiche chimico-fisiche e della complessità del sistema.
Si ipotizza che la natura destabilizzante dell'interfaccia acqua-ghiaccio possa imputarsi alla capacità di adsorbire, quindi a favorire l'*unfolding*, di proteine la cui superficie sia prevalentemente polare [42]. Tuttavia, non è possibile escludere a priori l'esistenza di altri fenomeni destabilizzanti causati o promossi dall'interfaccia in esame, pertanto sono necessari ulteriori studi per comprenderne meglio il ruolo nella stabilità di farmaci a base proteica.

In questa sezione si propone pertanto l'utilizzo della MD come mezzo per investigare i sistemi acqua-ghiaccio in presenza delle ciclodestrine. In particolare si tenterà di verificare la presenza di eventuali interazioni tra gli eccipienti e la superficie, allo scopo di capire se le *CD* possano rappresentare stabilizzanti adatti a contrastare l'azione denaturante dovuta all'adsorbimento.

2.2.1 Metodologia

Procedura di Simulazione

La metodologia usata per simulare l'interfaccia acqua-ghiaccio presenta alcune differenze sostanziali rispetto alle modalità impiegate per lo studio del sistema acqua-aria, pertanto si ritiene opportuno riprendere i passaggi mettendo in risalto i cambiamenti necessari. Le informazioni relative alla scelta della topologia delle ciclodestrine e dell'acqua, nonché del *force-field*, sono le stesse indicate nel paragrafo §2.1.1 e saranno le medesime nei capitoli seguenti.

La prima sostanziale differenza è data dalla necessità di modellare la struttura del ghiaccio, indispensabile per poter simulare l'interfaccia. La topologia del ghiaccio è stata ottenuta mediante GenIce [43], un algoritmo utilizzabile da linea di comando nel seguente modo:

genice - -water tip3p - -rep X Y Z 1h

dove l'opzione *- -water tip3p* consente di scegliere la topologia dell'acqua, in questo caso la *TIP3P* (si veda il paragrafo §2.1.1), l'opzione *- -rep X Y Z* permette di indicare il numero di ripetizioni del reticolo cristallino lungo le tre rispettive direzioni; infine *1h* indica la struttura cristallina del ghiaccio desiderata. Si ottiene quindi un file di configurazione del ghiaccio denominato, per esempio, *1h-tip3p.gro*.
A questo punto si hanno a disposizione tutte le topologie necessarie, pertanto è possibile procedere alla definizione del box. Si parte dall'inserimento della topologia del ghiaccio in un box, lasciando dello spazio libero per la fase liquida, con il comando:

gmx editconf -f 1h-tip3p.gro -box X Y Z -center a b c -o ibox.gro

da cui si ottiene un box di dimensioni *X x Y x Z* nm contenente uno strato di ghiaccio di spessore prefissato (di cui si fornisce la topologia in input con l'opzione *-f 1h-tip3p.gro*), il cui centro geometrico è impostato nelle coordinate *a, b, c*, del file di output *ibox.gro*. Le dimensioni del box *X, Y e Z* sono rispettivamente 7.824 nm x 8.09 nm x 9 nm; in particolare, lungo l'asse *Z* è collocato uno strato di ghiaccio di spessore 2.72 nm. Si noti che le dimensioni lungo gli assi *X* e *Y* sono fissate dal numero di ripetizioni del reticolo cristallino del ghiaccio.
Noto il volume destinato alla fase liquida, è possibile calcolare il numero di molecole di *CD* necessarie per ottenere una concentrazione pari a 50 mM, analogamente al sistema precedente, in questo caso pari a 12. Si inseriscono quindi le ciclodestrine mediante il comando:

gmx insert-molecules -f ibox.gro -ci CD.gro -nmol N -o iCDbox.gro

dove l'opzione *-f ibox.gro* prende in input il box precedente, *CD.gro* rappresenta il file che contiene la configurazione della ciclodestrina, *-nmol N* indica il numero di *CD* da inserire, infine *iCDbox.gro* è il file di output. Successivamente si provvede alla solvatazione del box usando il seguente comando:

gmx solvate -cp iCDbox.gro -cs spc216.gro -p topol.top -o iCDwbox.gro

dove *spc216.gro* indica la configurazione dell'acqua, *topol.top* raggruppa topologie, *force-field* e specie chimiche presenti, *iCDwbox.gro* è il file di output. Tuttavia, per questo sistema, la solvatazione comporta la presenza di molecole d'acqua allo stato liquido all'interno dei cristalli di ghiaccio (si veda Figura 2.9), situazione

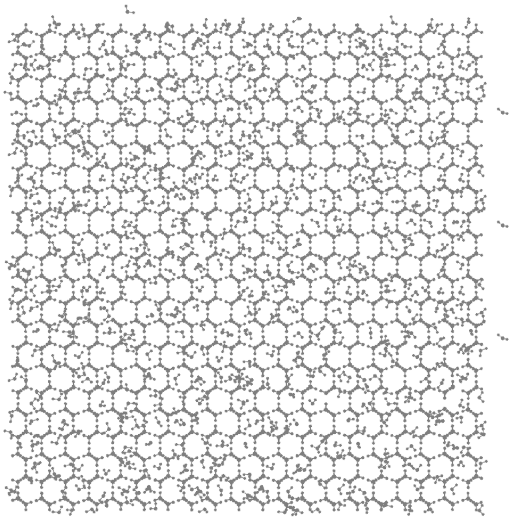

Figura 2.9: Raffigurazione delle molecole di acqua liquida indesiderate presenti nella matrice di ghiaccio. Vista ortogonale dall'alto. Immagine generata con VMD.

che si tradurrebbe in un modello erroneo e che porterebbe all'instaurarsi di forze intermolecolari troppo intense, con conseguente instabilità della simulazione.
Per porre rimedio a questa situazione è necessario individuare e rimuovere tali molecole, utilizzando il comando:

gmx select -f iCDwbox.gro -s iCDwbox.gro -on undesired-water.ndx

che usando la configurazione e la topologia del file *iCDwbox.gro* crea un file indice, *undesired-water.ndx*, contenente i riferimenti di tutte le molecole d'acqua presenti nello spessore del ghiaccio, incluse quelle indesiderate. Si modifica pertanto tale file eliminando dalla sequenza gli indici degli atomi facenti parte della struttura solida, ottenendo infine un indice delle sole molecole d'acqua indesiderate. Successivamente si crea un file indice escludendo tali molecole, *index.ndx*, da cui si ottiene la configurazione corretta del sistema, *box.gro*:

gmx $make_ndx$ *-f iCDwbox.gro -n undesired-water.ndx -o index.ndx*
gmx trjconv -f iCDwbox.gro -s iCDwbox.gro -n index.ndx -o box.gro

L'esito di questa operazione può essere facilmente verificato visualizzando il file *box.gro* mediante VMD o PyMOL.
A questo punto è necessario aggiornare il numero di molecole d'acqua presenti nel file di topologia *topol.top*, poiché esso deve sempre corrispondere allo stato attuale del sistema.

Si può successivamente procedere con la minimizzazione energetica, tuttavia in questo caso si deve "congelare" la posizione delle molecole di ghiaccio, poiché la

minimizzazione consiste nello spostare le molecole affinché l'energia del sistema sia minimizzata. Dovendo simulare l'interfaccia tra un liquido ed un solido, è necessario imporre che quest'ultimo sia indeformabile e fisso nello spazio, per evitare il rischio di spostamenti o deformazioni a causa della minimizzazione. Tale disposizione può essere inserita nel file *enermin.mdp*. Analogamente all'interfaccia acqua-aria, i comandi usati sono i seguenti:

gmx grompp -f enermin.mdp -c box.gro -p topol.top -n index.ndx -o em.tpr
gmx mdrun -v -deffnm em

dove il file *index.ndx* è necessario per identificare il gruppo di atomi costituenti il ghiaccio. Si ottiene infine la configurazione del sistema la cui energia è minimizzata, *em.gro*. Successivamente, si effettua l'equilibratura per portare il sistema alle condizioni fisiche di interesse. Il canone a cui si fa riferimento per tutte le simulazioni riportate in questa Tesi, fatta eccezione per l'interfaccia acqua-aria, è l'*NPT* [19]. L'equilibratura a 260 K e 1 bar in questo caso avviene in un'unica fase, suddivisa nuovamente in due stadi, utilizzando l'accoppiamento del termostato e del barostato di Berendsen [33]. Questa prima simulazione viene condotta per 1 ns, con time-step pari a 2 fs, bloccando la posizione del ghiaccio e generando una distribuzione di *Maxwell-Boltzmann* delle velocità casuale. I comandi impiegati sono:

gmx grompp -f npt.mdp -c em.gro -o npt.tpr -p topol.top -n index.ndx
gmx mdrun -v -deffnm npt

dove il file *index.ndx* è nuovamente necessario per identificare le molecole di ghiaccio. Si verifica quindi l'esito dell'equilibratura visualizzando l'andamento di temperatura, pressione ed energia nel tempo simulato:

gmx energy -f npt.edr -o T.xvg
gmx energy -f npt.edr -o P.xvg
gmx energy -f npt.edr -o energy.xvg

Qualora le fluttuazioni ai tempi finali fossero troppo elevate, si può estendere la durata dell'operazione aumentando il numero di step nel file *npt.mdp*, riprendendo la simulazione con il comando:

gmx mdrun -v -deffnm npt -cpi

Una volta completata l'equilibratura con successo si può passare alla simulazione vera e propria dell'interfaccia acqua-ghiaccio, svolta a 260 K e 1 bar. Tale simulazione viene estesa per 60 ns, con un time-step pari a 2 fs, riprendendo la distribuzione delle velocità generata nella fase precedente, mediante l'opzione *-t npt.cpt*. Si utilizzano a tal scopo i comandi:

gmx grompp -f md.mdp -c npt.gro -o md.tpr -p topol.top -n index.ndx -t npt.cpt
gmx mdrun -v -deffnm md

da cui si ottengono la configurazione dello stato finale del sistema *md.gro* e la traiettoria delle molecole *md.xtc*.

Analisi dei Risultati

Per procedere con lo studio delle proprietà delle ciclodestrine in presenza dell'interfaccia acqua-ghiaccio si estraggono dalle traiettorie i profili di densità, analogamente ai sistemi acqua-aria. Il comando utilizzato è nuovamente:

gmx density -f md.xtc -s md.tpr -n index.ndx -center -symm -b 20000 -sl 500 -o density-profile.xvg

che fornisce in output i profili di densità delle ciclodestrine (ed eventualmente dei *rim*), i quali devono essere adimensionati, secondo lo stesso criterio descritto nel paragrafo §2.1.1.

Per completare lo studio del sistema si analizzano le proprietà relative ai *cluster*, pertanto si isola nuovamente la traiettoria delle sole *CD*, per poi valutare la formazione, la numerosità e la dimensione degli aggregati. I comandi impiegati sono:

gmx trjconv -f md.xtc -s md.tpr -b 20000 -dt 10 -n index.ndx -o traj-CD.xtc -pbc whole

gmx trjconv -f md.xtc -s md.tpr -b 60000 -e 60000 -n index.ndx -o traj-CD.gro

gmx $make_ndx$ *-f traj-CD.gro -o index-CD.ndx*

gmx grompp -f enermin.mdp -c traj-CD.gro -o CD.tpr -p CD.top

gmx clustsize -f traj-CD.xtc -s CD.tpr -n index-CD.ndx -mol -hc clust-CD.xvg -nc nclust.xvg

da cui si ottengono la distribuzione delle *CD* impegnate in *cluster*, contenuta in *clust-CD.xvg*, ed il numero di aggregati presenti in funzione del tempo, nel file *nclust.xvg*.

Di seguito si riportano le raffigurazioni dei sistemi acqua-ghiaccio (Figure 2.10, 2.11, 2.12 e 2.13), da cui sono state omesse le molecole di acqua della fase liquida per motivi di chiarezza. Il sistema liquido si intende compreso tra due strati di ghiaccio, posti alle estremità superiori ed inferiori del box.

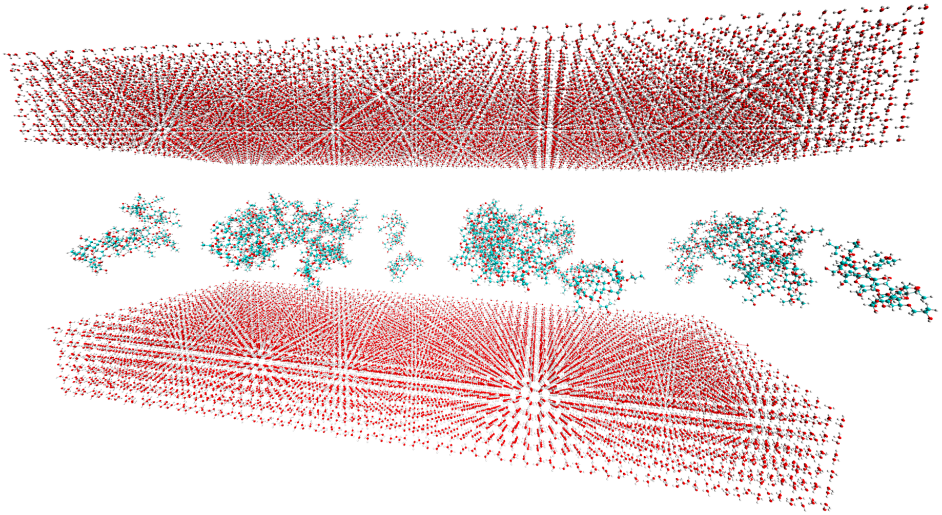

Figura 2.10: Rappresentazione delle *HP-β-CD* nel sistema acqua-ghiaccio. Si noti l'accumulo delle CD nel bulk. Immagine generata con VMD.

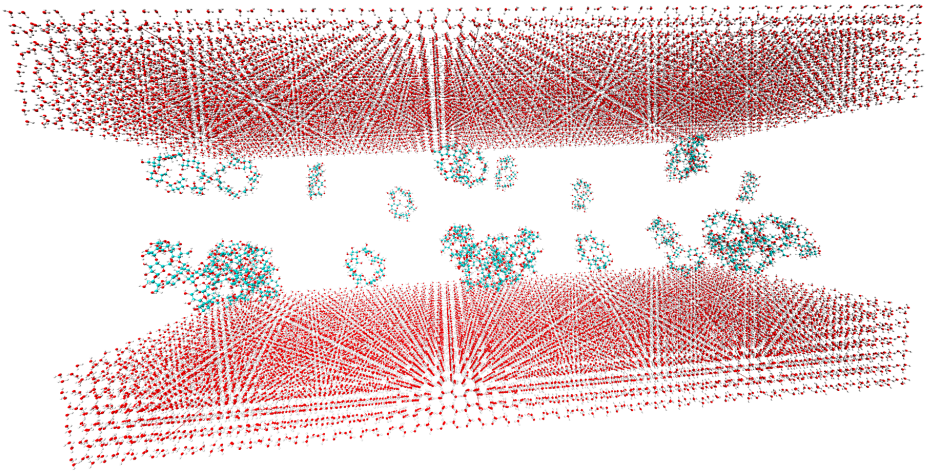

Figura 2.11: Rappresentazione delle *β-CD* nel sistema acqua-ghiaccio. Si noti l'accumulo delle CD nel bulk. Immagine generata con VMD.

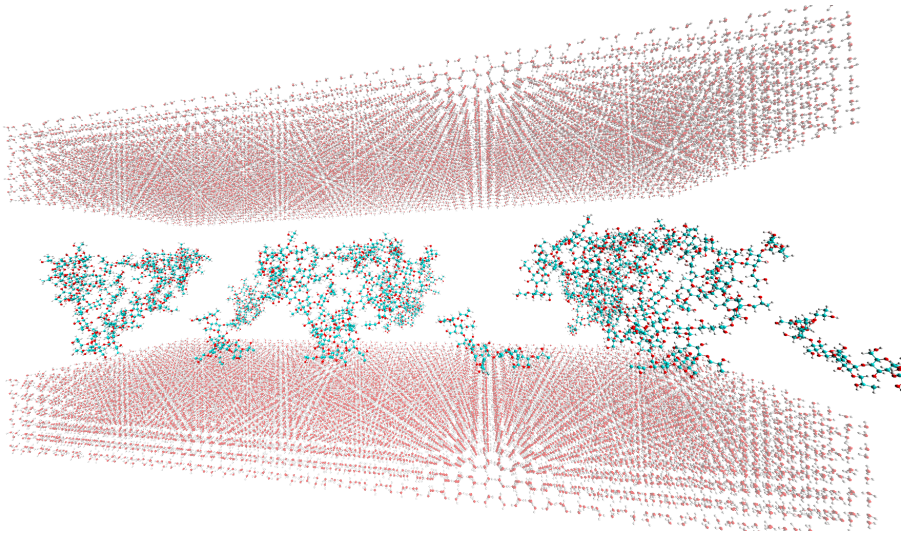

Figura 2.12: Rappresentazione delle open-*HP*-β-*CD* nel sistema acqua-ghiaccio. Si noti l'accumulo delle CD nel bulk. Immagine generata con VMD.

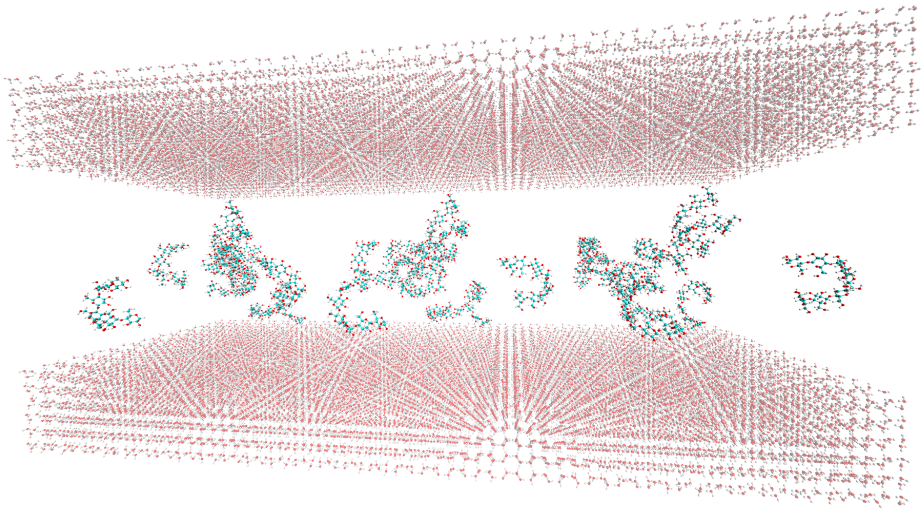

Figura 2.13: Rappresentazione delle open-β-*CD* nel sistema acqua-ghiaccio. Si noti l'accumulo delle CD nel bulk. Immagine generata con VMD.

2.2.2 Risultati ed osservazioni

I primi dati da analizzare, per stabilire il comportamento delle ciclodestrine in presenza dell'interfaccia acqua-ghiaccio, sono i profili di densità adimensionata delle *CD*, in funzione della distanza dalla superficie stessa, riportati in Figura 2.14.

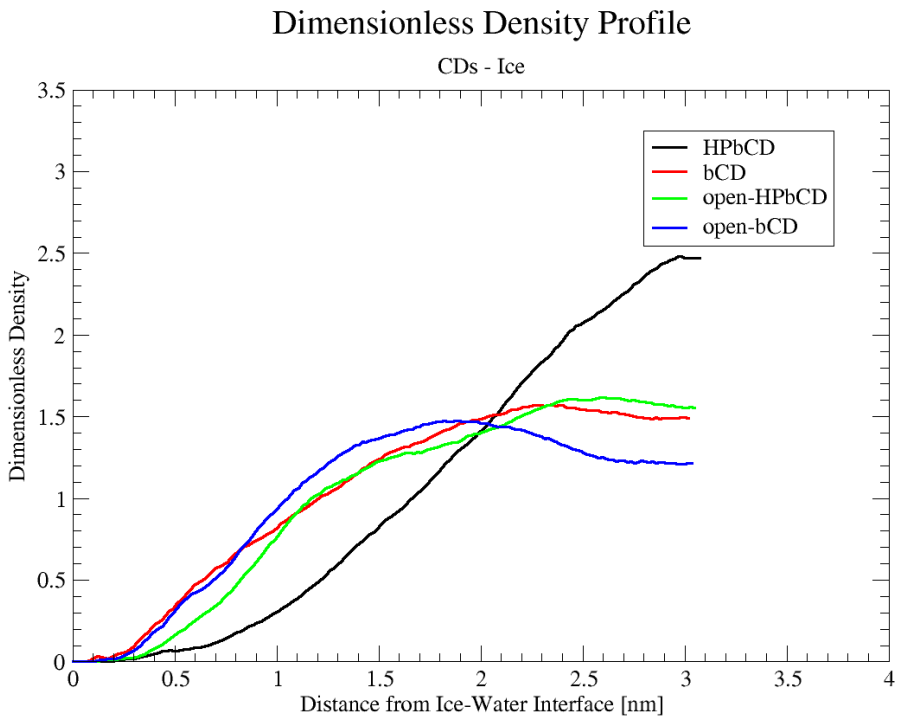

Figura 2.14: Profilo di densità adimensionata delle ciclodestrine in funzione della distanza dall'interfaccia acqua-ghiaccio.

A differenza del sistema acqua-aria, in questo caso tutte le ciclodestrine risultano escluse dall'interfaccia, accumulandosi nel bulk. Il comportamento è in linea con l'idrofobicità delle *CD*, tuttavia tale tendenza sembra essere accentuata per la *HP-β-CD* a catena chiusa, nonostante le sue caratteristiche più polari rispetto a quelle della *β-CD*. Ciò potrebbe essere legato alla formazione di *cluster* e alla temperatura ridotta rispetto al caso precedente, aspetti di cui si tratterà in seguito.

Questi primi risultati suggeriscono l'assenza di interazioni tra le ciclodestrine e l'interfaccia, poiché non vi sono indicazioni relative ad un comportamento paragonabile a quello dei surfattanti, contrariamente a quanto visto per le *HP-β-CD* in prossimità dell'aria. Per supportare tale affermazione si possono osservare i profili di densità dei *rim* sovrapposti a quelli delle *CD* corrispondenti, mostrati in Figura 2.15 ed in Figura 2.16.

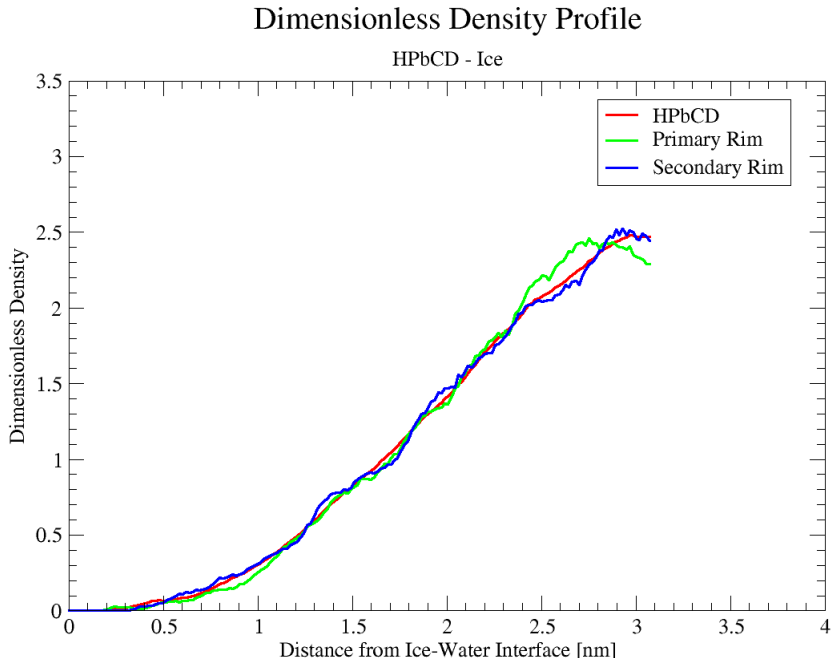

Figura 2.15: Profilo di densità adimensionata della *HP-β-CD* e dei *rim* in funzione della distanza dall'interfaccia acqua-ghiaccio.

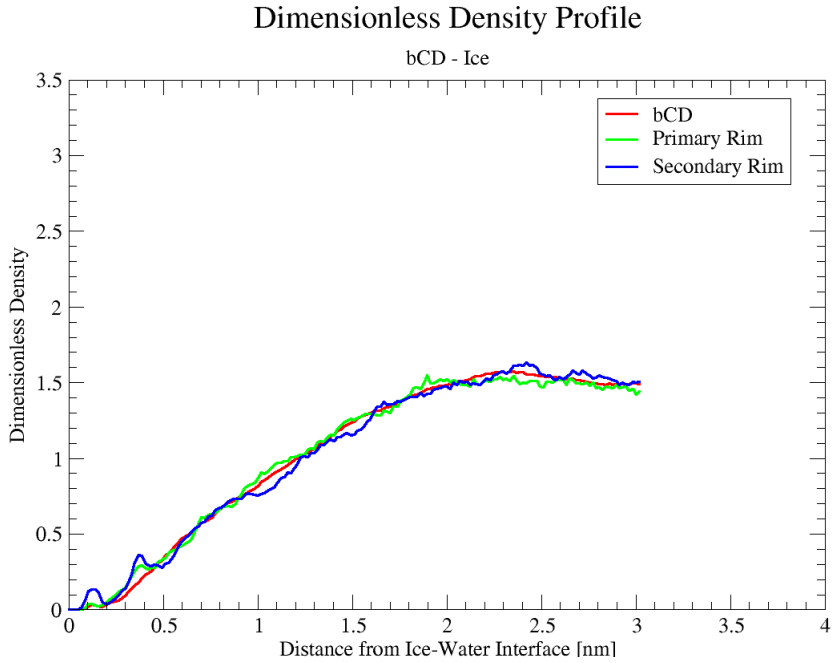

Figura 2.16: Profilo di densità adimensionata della *β-CD* e dei *rim* in funzione della distanza dall'interfaccia acqua-ghiaccio.

Si può notare che in entrambi i casi le *CD* si accumulano nel bulk, senza mostrare alcun segno di orientamenti preferenziali. Ciononostante, non si può escludere a priori la capacità delle ciclodestrine di stabilizzare le proteine in presenza dell'interfaccia acqua-ghiaccio. Sono pertanto necessarie analisi di tipo sperimentale, in modo da chiarire le proprietà delle *CD* in questo contesto.

Per completare l'analisi del sistema acqua-ghiaccio si studia la formazione dei *cluster*. In questo caso, la temperatura pari a 260 K, anziché 300 K come nel caso dell'interfaccia acqua-aria, implica una diminuzione della solubilità. Si considera inizialmente la distribuzione delle ciclodestrine impiegate in *cluster*, mostrata in Figura 2.17.

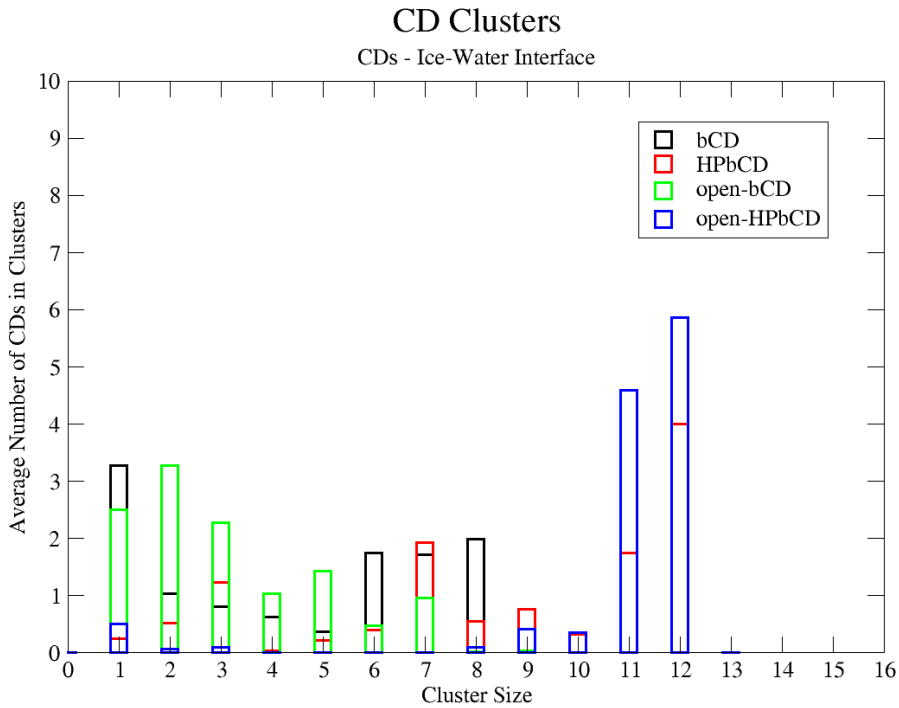

Figura 2.17: Distribuzione della dimensione dei *cluster* per le CD nel sistema acqua-ghiaccio.

Si può notare che le *HP-β-CD* tendano ad essere mediamente occupate nella formazione di *cluster* di dimensioni elevate, i quali comprendono la quasi totalità delle *CD* funzionalizzate, mentre gli aggregati di dimensioni medie e basse risultano scarsamente popolati. Diversamente, le *β-CD* mostrano una distribuzione più ampia, ovvero sono mediamente occupate in più *cluster* di dimensioni medie e basse, a scapito degli aggregati più massivi. Tale comportamento è qualitativamente in linea con quello osservato per l'interfaccia acqua-aria, ovvero le *HP-β-CD*

sembrerebbero più propense a formare pochi *cluster* composti da molte molecole, mentre le β-*CD* formerebbero più aggregati di poche molecole. In particolare, confrontando il numero medio di *CD* in un *cluster* (si veda Tabella 2.3), ed il numero medio di *cluster* (si veda Tabella 2.4) con i valori del sistema precedente si può notare come la diversità nel comportamento tra β-*CD* a catena aperta ed a catena chiusa risulti smorzata. Inoltre, si può osservare un aumento nel numero di *HP*-β-*CD*, in entrambe le forme, facenti parte di *cluster* rispetto al sistema acqua-aria. Ciò può essere spiegato dalla significativa affinità tra le *CD* in questione e l'interfaccia acqua-aria, la quale interagisce ed attrae gli eccipienti stessi, limitandone pertanto la possibilità di interagire e formare aggregati. Contrariamente, l'assenza di interazioni tra le due specie di *HP*-β-*CD* e la superficie acqua-ghiaccio consente e promuove l'instaurazione di interazioni intermolecolari e, di conseguenza, la formazione di *cluster* più massivi.

Tabella 2.3: N° medio di CD in un *cluster* per i sistemi acqua-ghiaccio.

HP-β-*CD*	β-*CD*	open-*HP*-β-*CD*	open-β-*CD*
8.27	5.27	11.1	3.15

Tabella 2.4: N° medio di *cluster* nel tempo per i sistemi acqua-ghiaccio.

HP-β-*CD*	β-*CD*	open-*HP*-β-*CD*	open-β-*CD*
2.00	5.13	1.58	5.66

2.3 Conclusioni

In conclusione, sono stati riscontrati buoni risultati nell'impiego di *HP*-β-*CD* in presenza dell'interfaccia acqua-aria; diversamente, nell'ambito della superficie acqua-ghiaccio, l'assenza di interazioni evidenti da parte delle ciclodestrine suggerisce la necessità di investigazioni sperimentali.

Infine, si precisa che lo studio sulle caratteristiche denaturanti dell'interfaccia acqua-ghiaccio e delle proprietà degli eccipienti in presenza di tale superficie è ancora limitato, soprattutto se paragonato ai sistemi acqua-aria. Questa problematica deriva dalla difficoltà materiale nella conduzione di esperimenti e nell'acquisizione di dati durante tali prove, a causa della complessità del sistema in esame e della dinamicità con cui esso evolve. Analogamente, anche utilizzando approcci computazionali lo studio rimane difficoltoso, a causa della mancanza di modelli adatti a tenere conto di questi aspetti, in parte derivante dalla scarsità di dati sperimentali. L'Autore non è a conoscenza di fonti bibliografiche relative all'impiego della Dinamica Molecolare per l'analisi delle proprietà delle ciclodestrine in presenza dell'interfaccia acqua-ghiaccio, pertanto le osservazioni proposte hanno carattere prevalentemente ipotetico.

Capitolo 3

Amminoacidi

Dopo aver esaminato le proprietà delle ciclodestrine relative alle interfacce acqua-aria ed acqua-ghiaccio, valutate allo scopo di saggiare l'adeguatezza di questi eccipienti a contrastare la denaturazione indotta da superfici, si procede con lo studio delle interazioni tra *CD* ed amminoacidi (in seguito indicati come *AA*). Questa analisi preliminare è volta a comprendere meglio i meccanismi di interazione tra proteine e ciclodestrine, e consentirà di paragonare il comportamento delle *CD* nei confronti dei singoli amminoacidi e dei residui peptidici. In particolare, questo confronto è necessario per comprendere sia l'influenza dell'affinità delle ciclodestrine verso le varie tipologie di *AA*, sia il ruolo della struttura proteica, la quale influisce inevitabilmente sull'accessibilità dei residui e, di conseguenza, sul rischio di aggregazione. Nell'ambito di questo studio si impiegheranno solamente le *CD* a catena chiusa, tenendo in considerazione anche eventuali orientamenti preferenziali, da parte di entrambe le specie chimiche.

In generale, è nota l'affinità delle ciclodestrine nei confronti degli amminoacidi apolari, specialmente quelli aromatici [6], in merito ai quali si hanno riscontri sperimentali [12], pertanto tale proprietà costituisce un importante criterio per stabilire la validità delle analisi svolte.

3.1 Metodologia

La metodologia seguita per la simulazione dei sistemi ciclodestrine-amminoacidi (*CD-AA*) differisce rispetto a quella descritta per le interfacce, poiché i sistemi in esame sono studiati in bulk. In particolare, è necessario l'impiego di analisi differenti, poiché i profili di densità non risulterebbero efficaci al fine di valutare le interazioni tra eccipienti ed *AA*. Pertanto, si ricorre alla *radial distribution function*, o più comunemente *rdf*, di cui si parlerà in dettaglio nei paragrafi successivi.
Nonostante le differenze nella procedura e nelle modalità di analisi, le scelte modellistiche rimangono le stesse (si veda il paragrafo §2.1.1), ovvero il *force-field* impiegato è nuovamente il CHARMM36 [26] [27] [28] [29], mentre la topologia dell'acqua rimane la *TIP3P* [31], trattata in modo esplicito [30].

Procedura di Simulazione

Noti i modelli da adottare per la simulazione, è necessario procurarsi le topologie delle specie chimiche da studiare. Per quanto riguarda le ciclodestrine si utilizzano le stesse indicate in precedenza, mentre bisogna reperire i file topologici per gli amminoacidi. In particolare, le analisi effettuate sono volte a studiare gli effetti dei residui peptidici quindi, invece di simulare gli *AA* nella loro forma nativa, si è scelto di impiegare una forma modificata. Le modifiche consistono nell'applicazione di un gruppo acetile (CH_3CO-) ed un gruppo amminico (NH_2-), rispettivamente alle estremità *N* e *C* di ciascun amminoacido, allo scopo di ottenere una struttura molecolare più simile a quella di un residuo peptidico. Questa operazione si svolge in due passi: inizialmente, mediante il software PyMOL [38], si aggiunge un gruppo acetile al terminale *N* dell'*AA* di interesse, salvando il residuo ottenuto su file in formato *.pdb*. Per completare la molecola si ricorre ad un comando di GROMACS, il quale consente di generare la configurazione opportuna, aggiungendo il gruppo amminico all'estremità *C*, tenendo conto del *force-field* e della topologia dell'acqua adottati:

gmx pdb2gmx -f AA.pdb -o AA.gro -water tip3p -ter -ignh

dove *AA.pdb* è il file di input, *AA.gro* è la configurazione finale desiderata in output, l'opzione *-water* consente la scelta della topologia per l'acqua. Inoltre, l'opzione *-ter* permette di indicare i gruppi terminali da applicare alla molecola in input, con il relativo stato di carica, mentre *-ignh* è un flag utilizzato per la gestione corretta degli atomi di idrogeno. Ulteriori informazioni sono reperibili sul Manuale dell'Utente di GROMACS [18] (http://manual.gromacs.org/documentation/2018.1/manual-2018.1.pdf). A questo punto si hanno a disposizione tutti i file necessari per la definizione del dominio computazionale, perciò si procede in maniera analoga a quanto visto per le simulazioni precedenti (si vedano i paragrafi §2.1.1 e §2.2.1). Si riportano quindi i comandi da utilizzare:

gmx insert-molecules -ci AA.gro -o pbox.gro -nmol 25 -box 8 8 8
gmx insert-molecules -ci CD.gro -f pbox.gro -nmol 15 -o pCDbox.gro
gmx solvate -cp pCDbox.gro -cs spc216.gro -o pCDwbox.gro -p topol.top

da cui si ottiene un box solvatato, *pCDwbox.gro*, di dimensioni 8 x 8 x 8 nm, contenente 15 molecole di *CD* e 25 molecole di *AA*. Si può pertanto procedere alla minimizzazione energetica, sebbene sia necessario applicare alcune modifiche alla modalità descritta nel paragrafo §2.1.1. Infatti, se gli amminoacidi non presentano cariche a pH 7, la procedura che si impiega è la medesima, ovvero si usa il comando:

gmx grompp -f enermin.mdp -c pCDwbox.gro -o em.tpr -p topol.top

Tuttavia, nel caso in cui si tratti di *AA* carichi è necessario introdurre degli ioni nel box, allo scopo di mantenere la neutralità complessiva. I comandi da impiegare per effettuare tale operazione sono i seguenti:

gmx grompp -f enermin.mdp -c pCDwbox.gro -o ions.tpr -p topol.top
gmx genion -s ions.tpr -pname NA -nname CL -np N_P *-nn* N_N *-p topol.top -o box.gro*
gmx grompp -f enermin.mdp -c box.gro -o em.tpr -p topol.top

dove *ions.tpr* rappresenta la topologia del sistema le cui cariche sono da bilanciare. A seconda del segno e del valore di carica del box si opta per l'aggiunta di ioni Na^+ oppure Cl^-, rispettivamente aggiunti mediante le opzioni *-pname* e *-nname*. La numerosità degli ioni da aggiungere per neutralizzare la carica del box viene fornita in input mediante l'opzione *-np* per i cationi o *-nn* per gli anioni. Si noti che per fare ciò si rimuovono delle molecole d'acqua, le quali vengono sostituite dagli ioni in questione, perciò è necessario aggiornare la topologia del sistema mediante l'opzione *-p topol.top*. Infine, si ottiene un box elettricamente neutro, con il quale si può procedere alla minimizzazione energetica, mediante il comando *gmx grompp*. Si conclude con il comando *gmx mdrun*, da cui si ottiene il box energeticamente minimizzato, *em.gro*.

gmx mdrun -v -deffnm em

Prima di simulare i sistemi *CD-AA* si effettua la fase di equilibratura a 300 K ed 1 bar, utilizzando l'accoppiamento del termostato e del barostato di Berendsen [33], per una durata complessiva di 1 ns, con un time-step di 2 fs. Si genera inoltre la distribuzione delle velocità in modo casuale. Si ricorda che il canone di riferimento adottato per tutte le simulazioni, fatta eccezione per i sistema acqua-aria, è l'*NPT* [19].

gmx grompp -f npt.mdp -c em.gro -p topol.top -o npt.tpr
gmx mdrun -v -deffnm npt

Una volta ottenuto il box equilibrato, *npt.gro*, si può verificare l'esito dell'operazione mediante il comando *gmx energy* ed eventualmente estendere la durata dell'equilibratura in base alle necessità, come suggerito nel paragrafo §2.1.1. Successivamente, si procede alla simulazione vera e propria, la cui durata è di 60 ns (time-step 2 fs), impiegando i comandi riportati di seguito. Il termostato utilizzato è il Nosé-Hoover [34] [35] [36], mentre il barostato è il Parrinello-Rahman [37]. Si noti che la distribuzione delle velocità viene ripresa dall'equilibratura, mediante l'opzione *-t npt.cpt*.

gmx grompp -f md.mdp -c npt.gro -p topol.top -o md.tpr -t npt.cpt
gmx mdrun -v -deffnm md

Si ottengono pertanto la traiettoria *md.xtc* e la configurazione dell'istante finale *md.gro*.
Gli amminoacidi in uso, nella loro forma modificata, sono rappresentati in Figura 3.1.

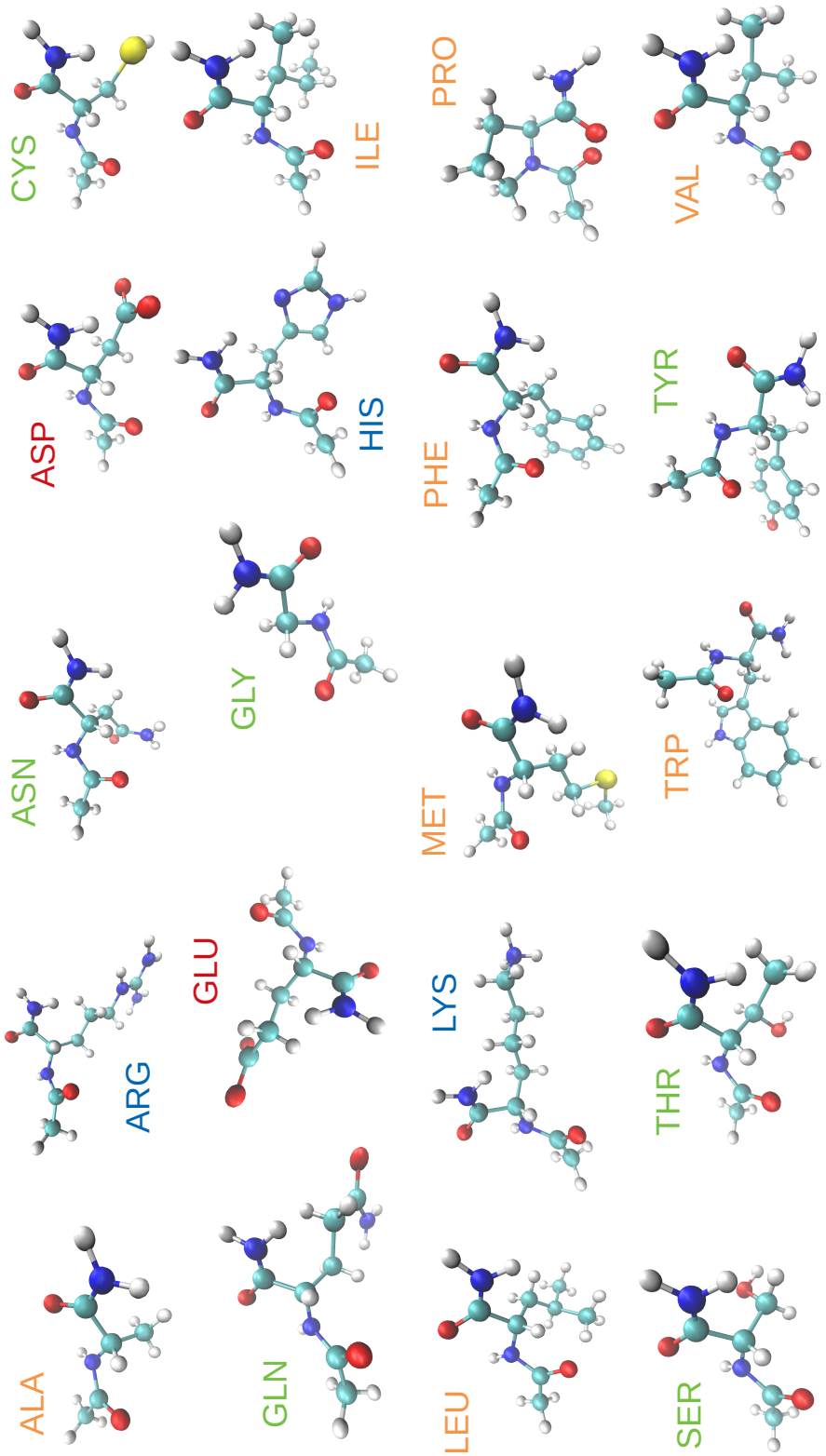

Figura 3.1: Rappresentazione degli amminoacidi modificati. Legenda delle etichette: nomi degli AA carichi positivamente in BLU, degli AA carichi negativamente in ROSSO, degli AA polari in VERDE, degli AA apolari in ARANCIONE. Immagini generate con VMD [13].

Analisi dei Risultati

Come accennato nel paragrafo precedente, per i sistemi *CD-AA* è necessario utilizzare una metodologia di analisi differente, poiché i profili di densità non sono adatti per lo studio delle interazioni tra le due specie chimiche di interesse. Innanzitutto, si procede a creare un file indice del sistema, generato a partire dall'ultimo istante della traiettoria, avendo cura di definire i *rim* delle ciclodestrine, in modo da poter valutare la presenza di eventuali orientamenti preferenziali, con le relative implicazioni. Si ricorre quindi al comando:

$$gmx\ make_ndx\ \text{-}f\ md.gro\ \text{-}o\ index.ndx$$

Lo strumento da utilizzare è la *radial distribution function* (in seguito *rdf*), la quale permette di stabilire se vi sia accumulo o dispersione tra *CD* e *AA*, in funzione della distanza tra le specie stesse. Lo stesso tipo di analisi può essere condotto equivalentemente anche tra porzioni delle molecole in esame, ovvero i *rim* per le ciclodestrine, la catena carboniosa (da qui in poi *backbone*) e la catena laterale (in seguito *sidechain*) per gli amminoacidi. In questo modo si potranno valutare orientamenti ed interazioni preferenziali. Una rappresentazione generica degli amminoacidi e delle due componenti è riportata in Figura 3.2.

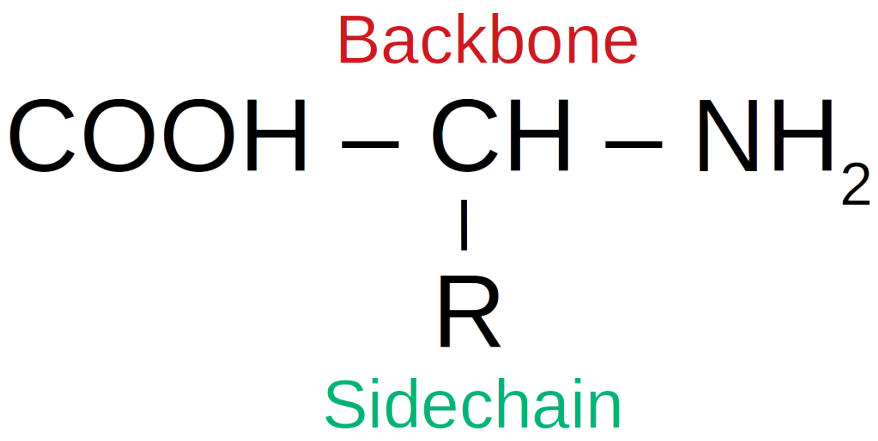

Figura 3.2: Rappresentazione generica di un amminoacido, con *backbone* e *sidechain* messi in evidenza.

L'espressione semplificata della *rdf*, indicata con $g_{AA-CD}(r)$, valutata per le *CD* rispetto agli *AA*, è riportata nell'equazione 3.1. Al numeratore vi è la densità media di ciclodestrine, espressa in funzione della distanza *r* dal centro di massa degli amminoacidi, mentre al denominatore vi è la densità media di *CD* nel box.

$$g_{AA-CD}(r) = \frac{\langle \rho_{CD}(r) \rangle}{\langle \rho_{CD}^{locale} \rangle} \tag{3.1}$$

Un'espressione più rigorosa ed estesa è invece proposta nell'equazione 3.2:

$$g_{AA-CD}(r) = \frac{1}{\langle \rho_{CD}^{locale} \rangle} \frac{1}{N_{AA}} \sum_{i \in AA}^{N_{AA}} \sum_{j \in CD}^{N_{CD}} \frac{\delta(r_{ij} - r)}{4\pi r^2} \tag{3.2}$$

dove l'*rdf* è calcolata come la densità media di *CD* in un intorno sferico di raggio *r*, originato dal centro di massa di ciascun *AA*, mediata su tutti gli amminoacidi contenuti nel box, divisa dalla densità media di ciclodestrine presenti in un intorno sferico di diametro pari all'ampiezza del dominio per ogni *AA*. Inoltre, tali stime sono da intendersi come mediate nel tempo.
Il comando impiegato per il calcolo delle *rdf* è il seguente:

gmx rdf -f md.xtc -s md.tpr -n index.ndx -o rdf-CD-AA.xvg -b 20000 -selrpos mol_com

dove l'opzione *-selrpos* mol_com è utilizzata per impostare il centro di massa delle molecole come riferimenti per le posizioni. L'opzione *-b 20000* è usata per trascurare la prima parte della simulazione, mentre il file di indice è fornito in input qualora si volessero considerare i *rim*, il *backbone* o la *sidechain*. Il file in output, contenente la *rdf* per il sistema in questione, è rappresentato da *rdf-CD-AA.xvg*. Grazie a questi dati, è possibile stabilire l'entità e la tipologia delle interazioni tra le varie specie chimiche esaminate, in base all'affinità reciproca. In particolare, tra le possibili combinazioni, si è deciso di focalizzarsi sul paragone tra *rdf* degli amminoacidi interi, *backbone* e *sidechain*, per comprendere quale delle due porzioni degli *AA* sia la più interagente. Infine, allo stesso modo si valuta l'orientamento preferenziale delle ciclodestrine nei confronti degli amminoacidi, confrontando le *rdf* di *CD*, *primary rim* e *secondary rim*.

Analogamente a quanto visto per le interfacce, si procede allo studio della formazione dei *cluster* tra le ciclodestrine e tra *CD* ed *AA*. In particolare, si vuole cercare corrispondenza tra le affinità rilevate mediante le *rdf* e la numerosità dei complessi ciclodestrina-amminoacido.
Lo studio dei *cluster* delle sole *CD* si svolge secondo le stesse modalità indicate nei paragrafi §2.1.1 e §2.2.1.
I comandi impiegati sono i seguenti:

gmx trjconv -f md.xtc -s md.tpr -b 20000 -dt 10 -n index.ndx -o traj-CD.xtc -pbc whole

gmx trjconv -f md.xtc -s md.tpr -b 60000 -e 60000 -n index.ndx -o traj-CD.gro

gmx $make_ndx$ *-f traj-CD.gro -o index-CD.ndx*

gmx grompp -f enermin.mdp -c traj-CD.gro -o CD.tpr -p CD.top

gmx clustsize -f traj-CD.xtc -s CD.tpr -n index-CD.ndx -mol -hc clust-CD.xvg -nc nclust-CD.xvg

Per valutare la formazione dei *cluster* tra *CD* e *AA* sono tuttavia necessarie alcune modifiche:

gmx trjconv -f md.xtc -s md.tpr -b 20000 -dt 10 -n index.ndx -o traj-CD-AA.xtc -pbc whole

gmx trjconv -f md.xtc -s md.tpr -b 60000 -e 60000 -n index.ndx -o traj-CD-AA.gro

gmx $make_ndx$ *-f traj-CD-AA.gro -o index-CD-AA.ndx*

gmx grompp -f enermin.mdp -c traj-CD-AA.gro -o CD-AA.tpr -p CD-AA.top

gmx clustsize -f traj-CD-AA.xtc -s CD-AA.tpr -n index-CD-AA.ndx -mol -hc clust-CD-AA.xvg -nc nclust-CD-AA.xvg

con cui si genera una traiettoria contenente solamente ciclodestrine ed amminoacidi *traj-CD-AA.xtc*, il corrispondente stato finale *traj-CD-AA.gro*, soffermandosi su dimensioni e numerosità medie dei *cluster CD-AA*. Si fa notare che non è possibile distinguere la specie chimica di appartenenza per le molecole facenti parte degli aggregati misti, salvo mediante confronti indiretti. Per questa ragione, si è utilizzato un approccio più selettivo, al fine di valutare più accuratamente la presenza di inclusioni di amminoacidi da parte delle ciclodestrine.

Il principio alla base dell'analisi delle inclusioni è quello di verificare la presenza di atomi appartenenti agli *AA*, in un intorno sferico di raggio 0.5 nm, tracciato rispetto al centro di massa delle *CD*. Pertanto, per ogni istante della simulazione, facendo riferimento al file indice del sistema, si prende nota dell'indice numerico corrispondente a tutti gli atomi compresi nel volume di controllo, centrato in ogni molecola di eccipiente, nonché della loro numerosità. Infine, si effettua una media temporale del numero di occorrenze rispetto alla quantità totale di ciclodestrine.
I comandi impiegati per questa analisi sono i seguenti:

gmx select -s md.tpr -f md.gro -n index.ndx -on index-temp-N.ndx -select 'group 'AA'; "CD-N" atomnr a to b'

gmx select -s md.tpr -f md.xtc -b 20000 -n index-temp-N.ndx -os incl-atom-N.xvg -select 'group Protein and within 0.5 of com of group CD-N'

dove le voci contenenti *N* si riferiscono alla numerazione crescente della ciclodestrina considerata (da 1 a 15). Nel dettaglio, ciascuno dei file indice *index-temp-N.ndx* contiene il numero sequenziale corrispondente a tutti gli atomi di tutti gli *AA*, unitamente a quelli della *CD* in considerazione. Tali file fungono da input per il secondo comando, con il quale si valuta effettivamente la quantità di atomi appartenenti agli amminoacidi inclusi da ogni singola ciclodestrina, per ogni istante. Alla fine di queste operazioni, si procede effettuando la media temporale del numero di *CD* che includono almeno un atomo appartenente agli *AA*.

3.2 Risultati ed osservazioni

Vista la numerosità dei sistemi analizzati in questo capitolo, nonché la varietà dei dati da essi ottenuti, si ritiene opportuno procedere suddividendo i dati in base alla tipologia degli stessi, lasciando le conclusioni di carattere generale al paragrafo successivo.

3.2.1 Radial Distribution Function

L'analisi delle *rdf* è la prima tipologia di risultati da considerare. Da esse, si possono ottenere informazioni fondamentali riguardo alle caratteristiche delle interazioni tra *CD* ed *AA*, quali l'affinità delle ciclodestrine nei confronti delle differenti classi di amminoacidi, nonché gli eventuali orientamenti preferenziali reciprocamente assunti dalle due specie chimiche in esame.

Come detto nel paragrafo §3.1, le *rdf* possono essere valutate per qualsiasi gruppo di atomi specificato nel file indice, pertanto inizialmente l'analisi è stata incentrata sull'orientamento degli eccipienti. Per fare ciò, si sono confrontate le *rdf* per tre coppie di gruppi atomici, ovvero *CD-AA*, *primary rim-AA* e *secondary rim-AA*, al fine di stabilire come le ciclodestrine si orientino nei confronti di ciascun amminoacido. Si mostrano due esempi di *rdf* nelle Figure 3.3 e 3.4.

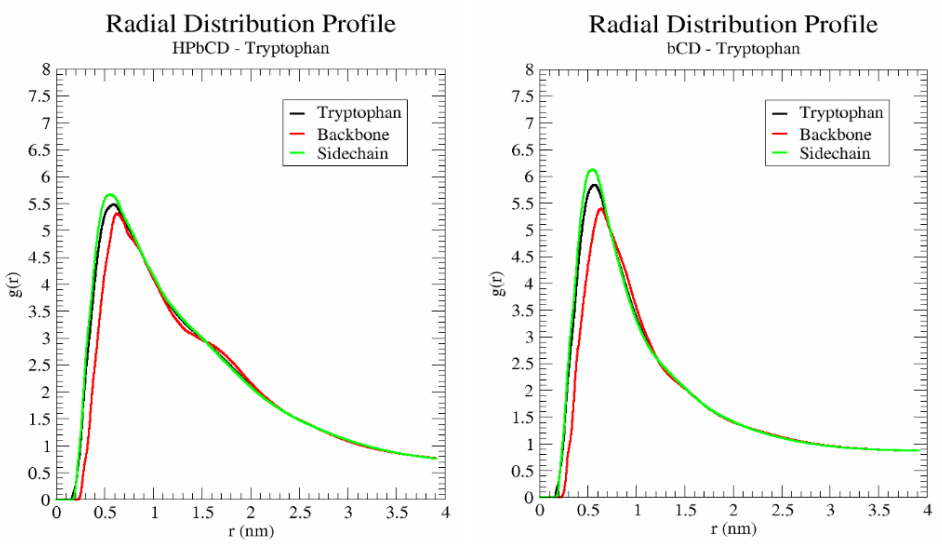

Figura 3.3: Esempio di *rdf* per entrambe le ciclodestrine. Confronto tra i profili ottenuti per le coppie di gruppi atomici *CD-TRP*, *CD-backbone* e *CD-sidechain*.

Risulta evidente come amminoacidi, appartenenti a categorie differenti, possano esibire profili significativamente diversi. Questo comportamento testimonia come le proprietà delle coppie *CD-AA* varino in base alle specie chimiche considerate. Pertanto, è necessario stabilire ed adottare un criterio adatto a descrivere le caratteristiche di ciascun sistema. A tale scopo, si è optato per il confronto tra i valori

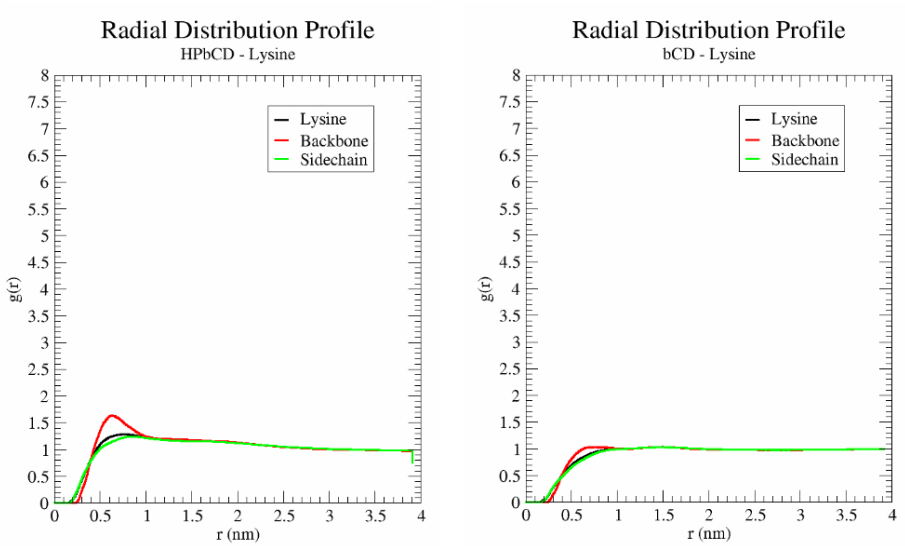

Figura 3.4: Esempio di *rdf* per entrambe le ciclodestrine. Confronto tra i profili ottenuti per le coppie di gruppi atomici *CD-LYS*, *CD-backbone* e *CD-sidechain*.

massimi di ciascuna *rdf*, i quali consentono di stabilire l'intensità di ciascuna interazione. In particolare, in Tabella 3.1 sono riportati i dati relativi a tutti i sistemi *CD-AA* analizzati. Da essi si può concludere che le ciclodestrine si orientino sistematicamente con il *secondary rim* verso gli amminoacidi, ad eccezione della β*-CD* in presenza di lisina (*LYS*) ed arginina (*ARG*). Infatti, in questi casi non c'è una netta prevalenza da parte di uno dei due *rim*, pertanto, a rigore, non è possibile stabilire un orientamento preferenziale. Questa valutazione è stata effettuata calcolando lo scarto tra i massimi dei dati dei due *rim*, per ciascun sistema, e rapportandolo al valore maggiore tra i due. Si ottiene pertanto un dato paragonabile ad un errore relativo, espresso per comodità in termini percentuali, mediante il quale discriminare orientazioni ben definite da situazioni limite, quali quelle riportate. A questo proposito si è stabilita una soglia di sbarramento, ovvero un *cut-off*, pari al 5%. Ne consegue che, soltanto nei casi in cui lo scarto tra le coppie di massimi superi tale valore, si potrà definire un'orientazione preferenziale.

Una volta stabilito l'orientamento preferenziale assunto da parte delle ciclodestrine, si procede analogamente per quanto riguarda la disposizione degli amminoacidi. La procedura adottata è sostanzialmente la medesima, applicata in questo caso alle *rdf* relative alle seguenti coppie di gruppi atomici: *CD-AA*, *CD-backbone* e *CD-sidechain*. Si potrebbe obiettare la decisione di considerare le intere ciclodestrine, anziché i soli *secondary rim*, avendo riscontrato l'esistenza di un orientamento preferenziale. Tuttavia, tale scelta deriva dalla necessità di mediare le interazioni con gli amminoacidi su tutti gli atomi delle *CD*.
In questo modo, il risultato ottenuto è indipendente dall'orientazione assunta dagli eccipienti.

Tabella 3.1: Confronto dei massimi delle *rdf* di *CD*, *primary rim* (*P.R.*) e *secondary rim* (*S.R.*), calcolate rispetto agli amminoacidi.

AA	*HP-β-CD*	*P.R.*	*S.R.*	*β-CD*	*P.R.*	*S.R.*
LYS	1.288	1.310	1.491	1.032	1.017	1.044
ARG	2.146	2.052	2.666	1.359	1.380	1.416
HIS	2.601	2.540	3.040	2.033	1.775	2.285
ASP	1.211	1.212	1.436	2.110	1.714	2.453
GLU	1.331	1.309	1.619	1.162	1.094	1.269
CYS	4.174	3.888	5.223	4.396	3.367	4.772
ASN	1.999	1.944	2.470	1.731	1.618	1.905
GLY	2.514	2.246	3.246	1.853	1.587	1.904
GLN	2.277	2.175	2.745	1.908	1.693	2.087
SER	2.651	2.536	3.439	1.712	1.626	1.857
THR	2.624	2.525	3.066	1.559	1.481	1.636
TYR	5.318	4.375	6.627	4.976	3.648	5.102
ALA	3.084	2.767	3.991	1.734	1.565	1.975
LEU	4.988	4.151	6.292	4.807	3.280	4.929
ILE	4.533	4.033	5.962	6.149	4.324	6.863
MET	4.976	4.288	6.266	6.434	4.377	6.653
VAL	3.729	3.440	4.792	4.792	3.598	5.291
PHE	4.122	3.323	5.207	7.017	5.178	7.031
PRO	3.893	3.589	4.968	2.883	2.398	3.231
TRP	5.491	5.227	7.150	5.845	4.300	6.606

Le stesse analisi sono state ripetute scegliendo *AA*, *backbone* e *sidechain*, come riferimenti rispetto a *primary rim* e *secondary rim*, con risultati sostanzialmente identici. Infatti, le uniche differenze sono state riscontrate solo quando i massimi, relativi a *backbone* e *sidechain*, differiscono di un valore inferiore a quello di *cut-off*, pertanto, in base al criterio adottato in precedenza, risultano trascurabili. In generale, si è osservato come le ciclodestrine interagiscano maggiormente con il *backbone*, sebbene vi siano alcuni casi privi di una netta distinzione tra le due componenti degli amminoacidi. Ciò si verifica per i seguenti sistemi:

- *HP-β-CD* con tirosina (*TYR*), leucina (*LEU*), metionina (*MET*) e fenilalanina (*PHE*);
- *β-CD* con lisina (*LYS*), tirosina (*TYR*) ed isoleucina (*ILE*).

Inoltre, tra i casi analizzati, si nota come per alcune coppie *CD-AA* prevalga l'interazione tra ciclodestrine e *sidechain*:

- *HP-β-CD* con triptofano (*TRP*);

- *β-CD* con leucina (*LEU*), metionina (*MET*) fenilalanina (*PHE*) e triptofano (*TRP*).

Ciascuna delle osservazioni precedenti si basa sui dati riportati nelle Figure 3.5 e 3.6.

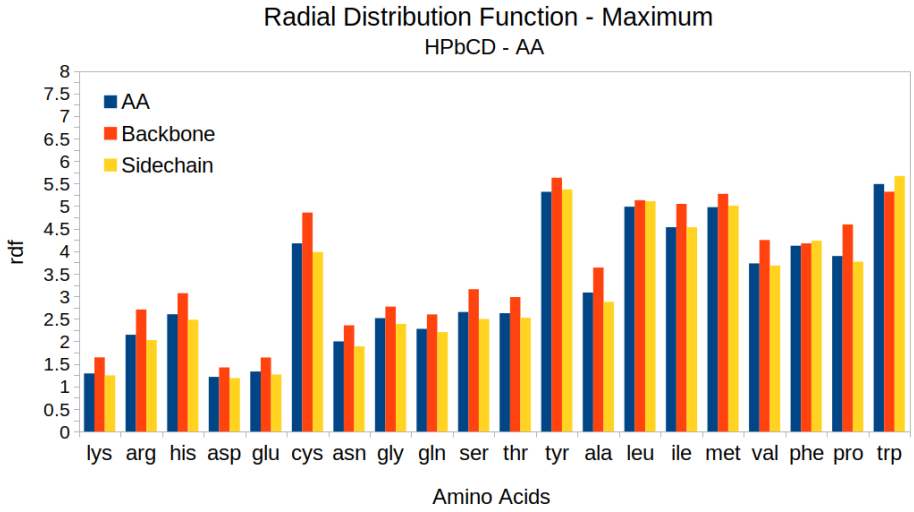

Figura 3.5: Confronto dei massimi delle *rdf* relativi alla *HP-β-CD*, calcolate rispetto ad *AA*, *backbone* e *sidechain*.

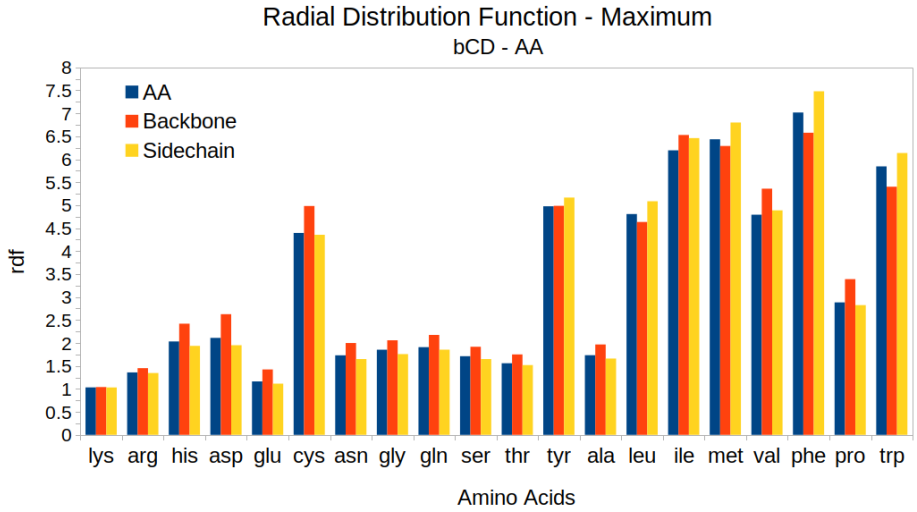

Figura 3.6: Confronto dei massimi delle *rdf* relativi alla *β-CD*, calcolate rispetto ad *AA*, *backbone* e *sidechain*.

I dati appena mostrati, consentono di visualizzare ed evidenziare più agevolmente le differenze tra le intensità di interazione, a seconda della tipologia di amminoacido considerato. Infatti, risulta evidente come non esista un comportamento gene-

ralizzabile, valido per tutti i sistemi *CD-AA*. Per questa ragione, si ritiene opportuno commentare separatamente i risultati, suddividendoli tra le quattro tipologie di amminoacidi.

Amminoacidi carichi positivamente

Gli amminoacidi carichi positivamente, ovvero la lisina (*LYS*), l'arginina (*ARG*) e l'istidina (*HIS*), interagiscono, con entrambe le ciclodestrine, in maniera poco significativa. La lisina è tra gli *AA* che interagiscono meno con gli eccipienti in esame, infatti sembra quasi essere esclusa dalle *CD*, come si può notare dai profili mostrati precedentemente in Figura 3.4.

La generalizzazione dei risultati ottenuti è resa ostacolata da differenze nel comportamento degli amminoacidi, soprattutto nell'ambito di una stessa categoria. Infatti, per quanto riguarda gli *AA* carichi positivamente, le proprietà di istidina ed arginina si discostano da quelle della lisina: è evidente che, il massimo valore della *rdf* di quest'ultimo amminoacido sia considerevolmente più basso rispetto alle altre specie con cariche positive, soprattutto in presenza della *HP-β-CD*. Per giustificare tale differenza si deve fare riferimento alle strutture molecolari delle specie chimiche in esame. Si può quindi notare che, mentre per la lisina vi sia una carica netta positiva, posizionata al termine della *sidechain* sul gruppo $-NH_3^+$, nel caso dell'arginina sia impossibile collocare esattamente la carica, poiché tale molecola dà origine a degli ibridi di risonanza. Infine, l'istidina non presenta cariche nette a pH 7. La presenza di un gruppo eterociclico nella catena laterale dell'*HIS*, conferisce a questa specie chimica proprietà elettrostatiche più blande, in quanto capace di delocalizzare meglio le cariche e formare, all'occorrenza, un maggior numero di ibridi di risonanza. Pertanto, la possibilità di delocalizzare elettroni conferisce all'arginina, e soprattutto all'istidina, caratteristiche più simili agli amminoacidi aromatici, nei confronti dei quali le *CD* mostrano grande affinità [44].

Amminoacidi carichi negativamente

L'acido aspartico (*ASP*) e l'acido glutammico (*GLU*) sono gli amminoacidi carichi negativamente. Le interazioni tra ciclodestrine e tali *AA* sono paragonabili a quelle del gruppo precedente, in quanto poco significative. Infatti, i valori massimi delle *rdf* sono simili a quelli registrati per la lisina, eccetto per il sistema *β-CD-ASP*, nel qual caso l'intensità è analoga a quella dell'istidina.
La struttura molecolare degli amminoacidi con cariche negative non suggerisce la presenza di fattori tali da poter spiegare questa differenza, la quale ha una valenza più quantitativa piuttosto che qualitativa. Pertanto, si può concludere che le interazioni tra questa categoria di *AA* e le ciclodestrine siano generalmente deboli, poiché l'effetto delle cariche non può essere smorzato da fenomeni di risonanza, perciò le proprietà risultanti corrispondono a quelle osservate per la lisina.

Amminoacidi polari

Gli amminoacidi polari sono costituiti dalla cisteina (*CYS*), dall'asparagina (*ASN*), dalla glicina (*GLY*), dalla glutammina (*GLN*), dalla serina (*SER*), dalla treonina (*THR*) e dalla tirosina (*TYR*). Per i sistemi appartenenti a questa categoria, si osservano sostanzialmente due comportamenti differenti: attrazioni significativamente intense, riscontrate per cisteina e tirosina, ed attrazioni piuttosto deboli per gli *AA* rimanenti.

In generale, si può affermare che per gli amminoacidi polari le interazioni con la *HP-β-CD* siano più forti rispetto alla ciclodestrina nativa, fatta eccezione per la cisteina. Tale risultato è giustificato dalla presenza di gruppi idrossipropilici nella *CD* sostituita, i quali forniscono a questo eccipiente un carattere maggiormente polare, unitamente ad una maggiore solubilità. Ne consegue che l'affinità di queste ciclodestrine, nei confronti degli *AA* in esame, sia superiore rispetto a quella riscontrata per la *β-CD*. Facendo riferimento alla struttura molecolare, si può notare che gli amminoacidi che interagiscono più debolmente posseggono gruppi più polari (ossidrile, $-OH$, ed amminico, $-NH_2$) rispetto ai rimanenti. Infatti, la *sidechain* della cisteina termina con un gruppo $-SH$, il quale ha una polarità analoga a quella dei legami tra carbonio ed idrogeno (è sufficiente confrontare l'elettronegatività degli elementi chimici in questione). Ciò spiegherebbe quindi l'elevata affinità che caratterizza i sistemi *CD-CYS*. Si potrebbe pertanto notare, ed obiettare, che la tirosina possegga una funzione ossidrilica, tuttavia facente parte di un gruppo fenolico, ovvero contenente un anello aromatico, il quale giustifica la notevole intensità di interazione.

Infine, si può concludere che le ciclodestrine interagiscano significativamente con cisteina e tirosina mentre, per i restanti amminoacidi, l'affinità è paragonabile a quella riscontrata per l'istidina, perciò complessivamente piuttosto debole.

Amminoacidi apolari

Gli amminoacidi apolari, ovvero alanina (*ALA*), leucina (*LEU*), isoleucina (*ILE*), metionina (*MET*), valina (*VAL*), fenilalanina (*PHE*), prolina (*PRO*) e triptofano (*TRP*), rappresentano la categoria che interagisce maggiormente con le ciclodestrine. In generale, gli *AA* appartenenti a questa categoria instaurano interazioni forti con gli eccipienti, salvo per i sistemi *CD-ALA* e *β-CD-PRO*, in cui l'intensità è più moderata.

Queste proprietà possono essere spiegate facendo nuovamente riferimento alle strutture molecolari degli amminoacidi apolari, in particolare alle *sidechain*. Esse sono costituite da catene alifatiche, eventualmente ramificate, o da gruppi aromatici, che pertanto manifestano caratteristiche tipicamente idrofobiche, quindi affini alle ciclodestrine. Le differenze riscontrate per alanina e prolina possono essere giustificate mediante un confronto con i restanti *AA* di questa tipologia. Infatti, l'*ALA* possiede una catena laterale carboniosa più breve rispetto a quella delle altre specie chimiche in esame, pertanto con un carattere apolare relativamente meno

pronunciato. Analogamente, per il sistema *β-CD-PRO*, la presenza di un gruppo eterociclico, che comprende l'estremità *N* della molecola, non conferisce un carattere apolare paragonabile a quello delle specie aromatiche, pertanto il massimo della *rdf*, indice dell'affinità, risulta inferiore. Tuttavia, si ritiene necessario osservare e notare come la ciclodestrina nativa interagisca più intensamente con gli altri amminoacidi in esame, soprattutto nei confronti delle *sidechain* di leucina, metionina, fenilalanina e triptofano, anche rispetto ai corrispondenti sistemi della *HP-β-CD*. Questo comportamento evidenzia una differenza cruciale tra i due eccipienti: la *β-CD* tende infatti ad interagire prevalentemente con gli amminoacidi apolari, esibendo scarse affinità nei confronti delle altre specie. Diversamente, la *HP-β-CD* interagisce più facilmente, rispetto alla molecola nativa, con tutte le tipologie di *AA*, mostrando escursioni di affinità più moderate tra le varie categorie.

Infine, si fa notare che, nonostante in generale sia il *backbone* ad interagire più spontaneamente, la *sidechain* contribuisce ugualmente a determinare le proprietà e l'affinità di ogni amminoacido. Ciò è particolarmente evidente per gli *AA* carichi, per i quali le interazioni non coinvolgono la catena laterale, tuttavia ne risentono ugualmente.

3.2.2 Inclusioni

Le interazioni che si manifestano tra ciclodestrine ed amminoacidi consistono nelle inclusioni, ovvero l'introduzione di atomi o porzioni più estese di una molecola all'interno della cavità degli eccipienti. Le *CD* sono infatti note per la loro capacità di inglobare diverse sostanze chimiche all'interno della propria cavità, in relazione alla reciproca affinità, nonché in base alla compatibilità sterica. Quest'ultimo aspetto è di fondamentale importanza, poiché concorre a determinare la stabilità delle inclusioni, assieme ad altri fattori termodinamici, legati all'energia di interazione. Tale capacità delle ciclodestrine può essere sfruttata opportunamente per vari scopi, tra cui, per esempio, applicazioni di drug delivery, sistemi di sequestro e separazione selettiva di specie molecolari, *refolding* di proteine denaturate.

L'analisi delle *rdf*, in particolare l'individuazione ed il confronto dei massimi, rappresenta una fase preliminare e necessaria per la valutazione delle inclusioni, poiché fornisce un'idea generale delle affinità tra le specie chimiche in esame, tuttavia senza fornire informazioni riguardo l'effettivo avvenimento del fenomeno. Pertanto, per fare ciò, si ricorre alla metodologia esplicata nel paragrafo §3.1, da cui si ottiene la percentuale di ciclodestrine che mediamente formano inclusioni. Si potrebbe obiettare la scelta delle *CD* come riferimento, tuttavia, qualora si fossero adottati gli amminoacidi, sarebbe stato necessario normalizzare il numero di atomi inclusi per la quantità totale di atomi in ciascun *AA*. Tale scelta fornirebbe la frazione di molecole di amminoacidi inclusi in media dalle ciclodestrine, senza però tenere conto delle possibili differenze di orientamento tra le due specie chimiche, con il possibile rischio di sfavorire, a rigore, le *sidechain* delle molecole stericamente più ingombranti. Inoltre, vista la numerosità degli *AA* rispetto agli eccipienti, una quota di molecole non può necessariamente essere inclusa, poiché

la stechiometria di interazione è 1:1.

Una volta chiarita la scelta del riferimento, risulta opportuno confrontare i dati relativi alle inclusioni, riportati in Figura 3.7, con i massimi delle *rdf*, in Figura 3.8, per tutti i sistemi *CD-AA*.

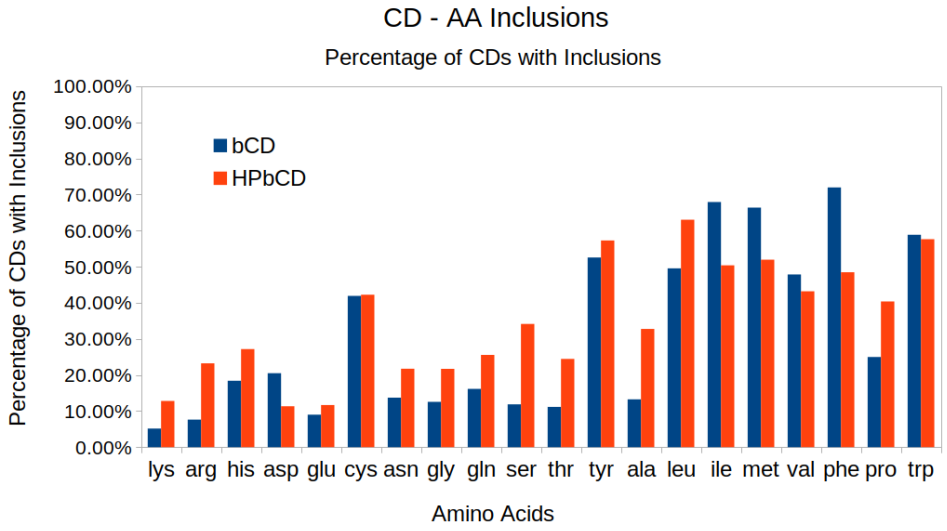

Figura 3.7: Confronto della percentuale di ciclodestrine che formano inclusioni per i sistemi *CD-AA*.

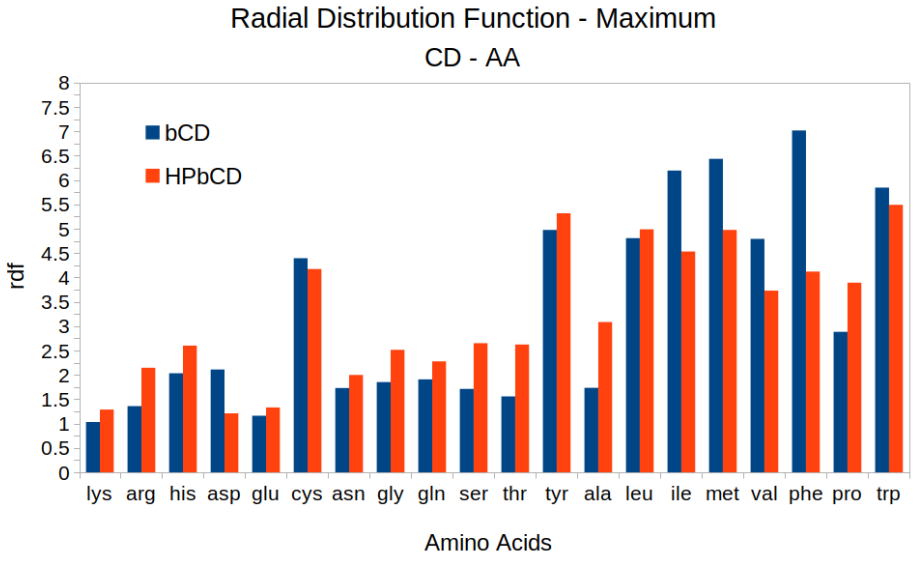

Figura 3.8: Confronto dei massimi delle *rdf* per i sistemi *CD-AA*.

Si può notare che la tendenza delle ciclodestrine a formare complessi con gli am-

minoacidi corrisponde sostanzialmente a quella dei massimi registrati per le *rdf*. Questo risultato conferma pertanto la validità di queste ultime come strumento preliminare per l'analisi delle inclusioni.
Si riportano alcune rappresentazioni delle inclusioni nelle Figure 3.9 e 3.10.

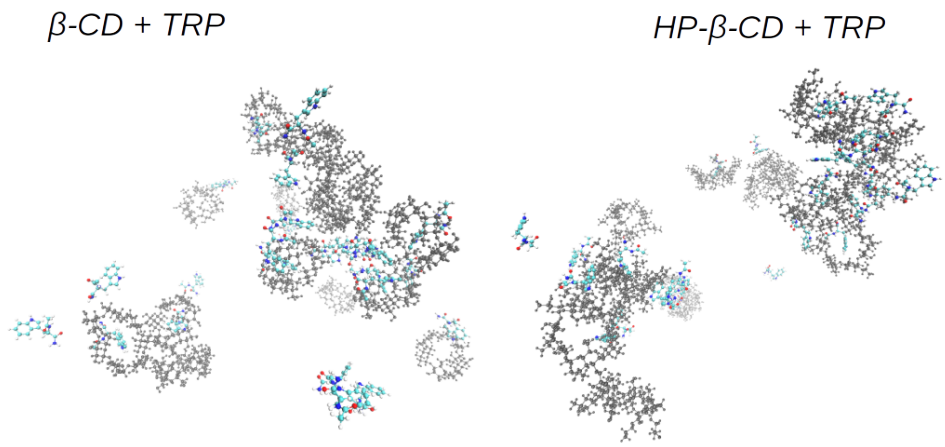

Figura 3.9: Rappresentazione delle inclusioni per i sistemi *CD-TRP*. Immagini generate con VMD [13].

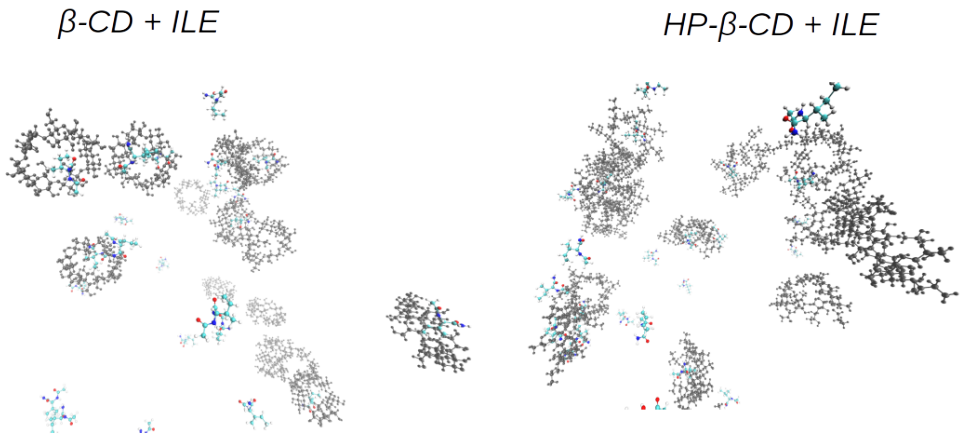

Figura 3.10: Rappresentazione delle inclusioni per i sistemi *CD-ILE*. Immagini generate con VMD [13].

3.2.3 Cluster

L'analisi dei *cluster* completa il quadro delle informazioni in merito alle interazioni per i sistemi *CD-AA*. A differenza dal capitolo precedente, in questo caso non ci si limiterà allo studio degli aggregati delle sole ciclodestrine, bensì si tratterà anche dei *cluster* formati da entrambe le specie chimiche. In particolare, si fa notare che per gli aggregati "misti" *CD-AA*, non è possibile distinguere le due tipologie di molecole, pertanto, sarà possibile valutare la presenza di amminoacidi soltanto mediante confronto indiretto con i dati dei soli eccipienti. Ciò risulta importante, non solo per la completezza delle analisi, ma anche e soprattutto per verificare la coerenza tra i diversi tipi di risultati ottenuti e, in particolare, sulle conclusioni che da essi si possono trarre.

Si procede con le dimensioni medie per entrambe le tipologie di *cluster*, riportate in Tabella 3.2.

Tabella 3.2: Confronto della dimensione media dei *cluster CD* e *CD-AA*.

AA	*HP-β-CD-AA*	*β-CD-AA*	*HP-β-CD*	*β-CD*
LYS	11.13	3.82	11.13	4.05
ARG	7.96	3.74	7.85	3.26
HIS	13.27	4.54	10.70	4.42
ASP	9.37	4.70	8.61	4.93
GLU	6.28	4.67	5.45	5.41
CYS	12.54	4.49	8.98	3.99
ASN	9.81	3.83	8.75	3.34
GLY	10.27	3.61	8.72	3.28
GLN	9.61	3.71	10.90	3.10
SER	12.02	4.33	10.46	4.16
THR	11.72	3.92	12.68	3.71
TYR	14.05	6.65	9.88	6.22
ALA	8.65	6.36	6.41	6.23
LEU	11.55	3.99	7.53	3.12
ILE	10.75	4.82	9.34	3.91
MET	11.56	4.79	9.76	3.56
VAL	9.27	7.00	6.90	6.66
PHE	13.56	5.28	10.70	4.43
PRO	11.16	4.15	10.46	3.59
TRP	20.11	6.93	9.96	5.14

Come osservato in precedenza, le *HP-β-CD* formano *cluster* di dimensioni maggiori, ovvero composti da un numero più elevato di molecole, rispetto alle *β-CD*. Si noti, inoltre, come lo stesso comportamento sia osservato per entrambe le tipologie di aggregati. Se si considerano i *cluster CD-AA*, le dimensioni medie sono generalmente maggiori rispetto al caso delle sole ciclodestrine: ciò può essere imputato al numero più elevato di amminoacidi presenti nel box, più abbondanti degli

eccipienti. In questo modo, le possibilità di interazione tra le due specie chimiche aumentano, giustificando il risultato ottenuto. Si ritiene opportuno precisare che la formazione di *cluster CD-AA* non implichi necessariamente l'instaurazione di inclusioni, poiché esse si considerano tali solo qualora le molecole siano incapsulate nella cavità.

L'analisi del numero medio di *cluster* formati presenta risultati analoghi a quelli ottenuti in precedenza, ovvero le *β-CD* formano quantità più elevate di aggregati rispetto alle *HP-β-CD*, necessariamente di dimensioni inferiori. I dati sono riportati in Tabella 3.3.

Tabella 3.3: Confronto del numero medio di *cluster CD* e *CD-AA*.

AA	*HP-β-CD-AA*	*β-CD-AA*	*HP-β-CD*	*β-CD*
LYS	18.36	25.61	2.59	8.00
ARG	14.79	22.54	4.52	8.57
HIS	14.90	22.20	2.08	7.34
ASP	20.78	24.64	2.84	6.60
GLU	20.57	24.72	4.06	6.13
CYS	12.68	20.35	2.70	8.09
ASN	17.03	23.96	3.13	8.51
GLY	17.96	25.75	3.30	8.60
GLN	15.85	23.27	3.45	8.56
SER	16.96	24.06	2.28	7.39
THR	16.63	24.65	2.32	8.09
TYR	8.23	15.17	2.68	6.66
ALA	16.38	22.50	4.13	4.79
LEU	11.31	20.37	3.59	9.52
ILE	11.77	17.66	3.46	7.23
MET	9.65	15.60	3.71	7.68
VAL	14.88	18.29	4.16	5.00
PHE	10.68	14.51	2.35	7.40
PRO	13.35	22.17	2.90	7.86
TRP	5.92	11.78	2.51	6.17

Pertanto, si può osservare come la presenza degli amminoacidi non alteri il comportamento qualitativo delle ciclodestrine. Inoltre, limitatamente per i casi esaminati, non ci sono evidenze di effetti antagonistici tra la formazione di inclusioni e di aggregati, in quanto l'andamento osservato è qualitativamente il medesimo. A sostegno di tale affermazione, è possibile confrontare i dati qui riportati con quelli relativi alle interfacce (Tabelle 2.1, 2.2, 2.3 e 2.4); si ricordi che per i sistemi acqua-ghiaccio il numero di ciclodestrine è inferiore (12 anziché 15).

3.3 Conclusioni

Lo studio delle interazioni tra ciclodestrine ed amminoacidi rappresenta un passaggio intermedio fondamentale al fine di investigare la proprietà stabilizzanti nei confronti delle proteine. In questo capitolo, le analisi sono state incentrate sulle *CD* a catena chiusa, focalizzandosi sulla loro capacità di includere molecole (o eventualmente porzioni delle stesse), nel caso specifico gli *AA*. Ciò costituisce una premessa fondamentale per l'applicazione degli stessi metodi a residui peptidici facenti parte di oligopeptidi e polipeptidi.
Sono state raccolte tre tipologie di dati dalle simulazioni di MD, ovvero:

- le *radial distribution function*, per valutare variazioni di densità relativa (rispetto a quella media dei singoli box) in funzione della distanza tra ciclodestrine ed amminoacidi. In particolare, queste analisi possono essere sfruttate non solo per molecole intere, ma anche per parti delle stesse, come *primary rim* e *secondary rim* per le *CD*, *backbone* e *sidechain* per gli *AA*. Questo approccio ha permesso di stabilire gli orientamenti preferenziali di ciascuna coppia ciclodestrina-amminoacido, da cui si è concluso che, generalmente, le interazioni avvengano tra *secondary rim* e *backbone*. Tuttavia, vi sono alcuni casi in cui l'affinità nei confronti della catena laterale prevale, altri in cui non vi è una netta distinzione. Infine, il confronto dei massimi valori registrati per le *rdf* ha permesso di identificare e distinguere le intensità di interazione nei confronti delle diverse categorie di amminoacidi;

- le inclusioni, ovvero l'incapsulamento di porzioni degli *AA* all'interno della cavità delle ciclodestrine. La tendenza ad includere varia sostanzialmente da caso a caso, tuttavia sono state osservate alcune proprietà generali, peraltro in accordo con le conclusioni tratte dalle *rdf*. In generale, le *HP-β-CD* interagiscono ed includono più agevolmente gli amminoacidi carichi e polari, salvo alcuni casi in cui vi siano interazioni altrettanto forti da parte della *β-CD*, come in presenza della cisteina e della tirosina. Infine, gli *AA* apolari, tra cui in particolare gli aromatici, sono le specie chimiche che interagiscono più facilmente ed intensamente con entrambe le ciclodestrine, ciascuna con intensità diverse ma pur sempre notevoli;

- la formazione e le caratteristiche dei *cluster*, che in questo contesto differiscono dalle analisi precedenti, a causa della presenza di aggregati "misti", ovvero ottenuti dall'interazione di entrambe le specie chimiche. Il comportamento qualitativo rimane lo stesso, sia per i *cluster* composti da *CD* ed *AA* sia per quelli di soli eccipienti. Inoltre, le proprietà riscontrate per le due ciclodestrine sono le stesse osservate in precedenza, ovvero la formazione di aggregati di dimensioni maggiori per le *HP-β-CD*, a scapito della numerosità, e viceversa per le *β-CD*, ovvero agglomerati più numerosi, tuttavia meno massivi.

A fronte di queste considerazioni, si può affermare che i risultati siano sostanzialmente coerenti tra di loro, evidenziando chiaramente le differenze nei comportamenti delle due *CD* in esame, soprattutto in relazione alle categorie di amminoacidi.
Vi sono tuttavia alcuni aspetti che necessitano di ulteriori ed approfondite investigazioni, ad esempio la distinzione tra *backbone* e *sidechain* per quanto concerne le inclusioni, in modo da capire quale porzione della molecola sia più inclusa oppure, ancora più specificamente, quali atomi siano quelli più vicini al centro della cavità delle ciclodestrine. Questa tipologia di informazioni può essere facilmente estratta dalle simulazioni sfruttando i file indice, ovvero definendo i gruppi atomici di interesse rispetto ai quali valutare affinità ed interazioni. Quindi appare chiaro come la versatilità e la relativa semplicità di impiego di queste metodologie rappresenti un considerevole vantaggio rispetto ai vari approcci sperimentali, utilizzati tradizionalmente per lo studio di questi stessi aspetti. Inoltre, l'impiego di software di visualizzazione consente di osservare "direttamente" l'evoluzione dei fenomeni studiati, mentre ciò risulta spesso difficile, se non impossibile, dal punto di vista sperimentale.
L'approccio presentato in questo capitolo vuole pertanto servire da base per studi futuri e più dettagliati, in materia di interazioni tra eccipienti e principi attivi o altre biomolecole di interesse.

Nel capitolo successivo si impiegheranno metodi analoghi a quelli qui presentati per lo studio delle interazioni tra proteine e ciclodestrine, focalizzandosi anche sul ruolo dei singoli residui peptidici, cercando di verificare l'eventuale corrispondenza tra le proprietà esibite dagli amminoacidi, sia come molecole a sé stanti sia come componenti di polipeptidi.

Capitolo 4

Proteine

Il passo successivo allo studio delle interazioni tra amminoacidi e ciclodestrine consiste nell'analisi del comportamento degli eccipienti in presenza di proteine. L'obiettivo è valutare, e possibilmente validare, le informazioni precedentemente ottenute. In particolare, si vogliono testare le capacità stabilizzanti delle *CD*, in modo da stabilire se tali sostanze siano adatte all'uso per formulazioni di farmaci a base proteica, ed eventualmente mediante quale meccanismo di stabilizzazione agiscano. Per fare ciò, si intende prima verificare la validità dell'approccio computazionale adottato in questa Tesi, studiando il caso di una proteina modello, ovvero la *lattato deidrogenasi* (o *lactate dehydrogenase*, indicata da ora in poi come *LDH*). Successivamente, si intende applicare la stessa metodologia ad un'altra proteina di interesse, ovvero il *fattore di stimolazione delle colonie di granulociti* (o *granulocyte-colony stimulating factor*, a cui si farà riferimento come *GCSF* da qui in seguito).

Lo studio delle interazioni tra proteine e ciclodestrine si focalizzerà su più aspetti, i quali verranno dettagliati nei paragrafi successivi, allo scopo di chiarire il meccanismo di stabilizzazione in atto e di correlare quanto osservato ai risultati riscontrati con gli amminoacidi. A questo scopo si impiegheranno la β-*CD* e la *HP*-β-*CD*, sia in forma a catena aperta sia chiusa.

4.1 Cenni sulle proteine analizzate

Prima di procedere alla descrizione delle metodologie di simulazione ed analisi si ritiene opportuno dedicare uno spazio alle proteine scelte per questo studio, cercando di descriverne alcune caratteristiche ed applicazioni.

4.1.1 Lattato deidrogenasi, LDH

La *lattato deidrogenasi* è un enzima globulare tetramerico, le cui proprietà son ben note, motivo per cui tale proteina viene spesso impiegata nei più svariati tipi di studi sperimentali. L'*LDH* è di grande interesse nel campo medico, poiché un aumento

della sua concentrazione nel nostro organismo, oltre un valore di soglia, può essere associata all'insorgenza di tumori [45] [46]. Ciò è dovuto all'attività metabolica anomala ed incontrollata delle cellule tumorali. Esse incrementano in maniera smisurata il consumo di glucosio, fonte energetica principale delle cellule, che viene metabolizzato portando alla formazione di acido lattico. Per fare ciò è necessaria la presenza e la produzione a livello intracellulare di *LDH*. L'intensa attività delle cellule tumorali può quindi provocare scompensi tra l'approvvigionamento di nutrienti ed il loro consumo, il quale è molto più rapido, causando danni cellulari anche fino alla necrosi, con conseguente liberazione dell'enzima in questione. Per questa ragione, il contenuto di *LDH* nel flusso sanguigno viene monitorato, al fine di individuare precocemente, ove possibile, attività tumorale [47].

Questo enzima può venire utilizzato anche per altri scopi, per esempio nell'ambito del controllo di qualità industriale, per quantificare il contenuto di glucosio presente nei prodotti, direttamente proporzionale all'attività enzimatica registrata.

Un'applicazione, attualmente in via di sviluppo nel contesto industriale, è costituita dall'impiego dell'*LDH* nei processi di produzione di esteri dell'acido lattico, da parte di cellule microbiche [48]. In questo contesto, lo scopo è quello di ottenere solventi ecosostenibili, prodotti a partire da risorse rinnovabili (biomassa). Alcuni di questi prodotti trovano impiego anche nel campo farmaceutico. Tale approccio è di considerevole interesse, poiché attualmente alcuni di questi composti chimici vengono prodotti mediante processi industriali ad alte temperature, che coinvolgono sostanze corrosive ed inquinanti, quali l'acido solforico e l'acido cloridrico. Invece, la possibilità di impiegare metodi produttivi basati sull'attività microbica, in condizioni ambientali, consente di evitare queste problematiche.
L'*LDH* è mostrata in Figura 4.1.

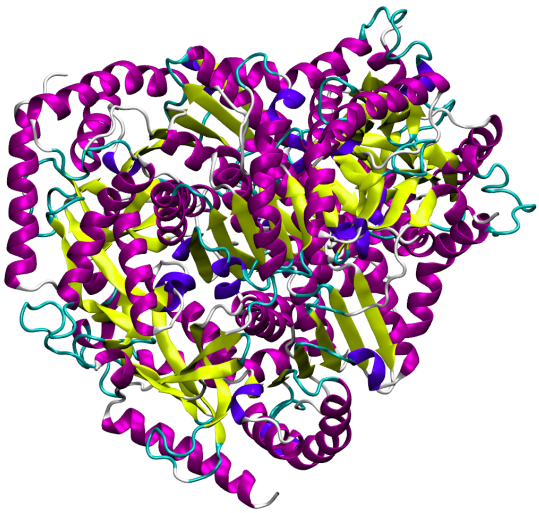

Figura 4.1: Rappresentazione dell'*LDH*. Immagine generata con VMD [13].

4.1.2 Fattore di stimolazione delle colonie di granulociti, GCSF

Il *fattore di stimolazione delle colonie di granulociti* è una proteina terapeutica di grande interesse, coinvolta nella differenziazione e proliferazione di cellule staminali, provenienti dal midollo osseo e convertite in cellule ematopoietiche. Misure di cristallografia a raggi X hanno evidenziato la presenza di α-eliche nella struttura di tale molecola [49]. La *GCSF* è impiegata per il trattamento di diverse patologie rilevanti, tra cui si annoverano:

- malattie cardiovascolari, contrastate migliorando il funzionamento del miocardio e la neovascolarizzazione [50];
- tumori, come farmaco somministrato parallelamente ai cicli chemioterapici, per rafforzare il sistema immunitario [49] [51];
- morbo di Parkinson, mitigandone i danni e lo sviluppo nelle fasi iniziali della malattia [52].

Sono stati riportati anche studi nell'ambito della rigenerazione di midollo spinale, in seguito al verificarsi di traumi, che prevedono il supporto della *GCSF* nel trattamento a base di cellule staminali neuronali, derivate da denti del giudizio dell'uomo [53].
Tali esempi spiegano e giustificano l'interesse nella produzione di questo farmaco; vista l'importanza delle tematiche trattate è perciò necessario l'impiego di formulazioni efficaci. La *GCSF* è mostrata in Figura 4.2.

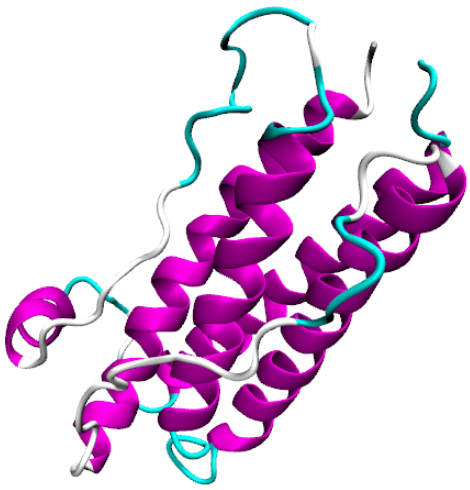

Figura 4.2: Rappresentazione della *GCSF*. Immagine generata con VMD [13].

4.2 Metodologia

Procedura di Simulazione

La metodologia impiegata per le simulazioni delle proteine, in presenza di ciclodestrine in bulk, è sostanzialmente analoga a quella vista per gli amminoacidi. Pertanto, si riporteranno brevemente i comandi già introdotti, soffermandosi sulle differenze. Le scelte modellistiche sono le stesse indicate in precedenza, ovvero la topologia dell'acqua *TIP3P* [31], trattata esplicitamente [30] ed il *force-field* CHARMM36 [26] [27] [28] [29]. Quindi, è necessario procurarsi la topologia delle proteine da simulare, mentre per quelle degli eccipienti sono valide le stesse indicazioni date in precedenza [23] [15] [24]. La scelta della topologia delle proteine è tutt'altro che banale poiché, vista la complessità di tali molecole, è indispensabile avere a disposizione modelli accurati della struttura molecolare. Ciò deriva dalla nota dipendenza dell'attività biologica dall'aspetto strutturale. Ne consegue che, evidentemente, non sia possibile assemblare le biomolecole di interesse come si è fatto per gli amminoacidi, ad esempio mediante PyMOL [38]. Sarebbe infatti necessario conoscere l'intera sequenza di residui peptidici, nonché i dati topologici. Fortunatamente, esistono banche dati, presso le quali è possibile reperire le topologie di svariate proteine di interesse scientifico. Per entrambe le molecole in esame si è fatto riferimento alla banca dati RCSB PDB [54] (http://www.rcsb.org/).

Le strutture vengono tipicamente ottenute mediante misure sperimentali di cristallografia a raggi X, spettroscopia NMR e microscopia crioelettronica, con il supporto della modellazione teorica. Un ulteriore aspetto da considerare è l'organismo di provenienza del modello topologico, in quanto la stessa tipologia di proteina può presentare delle mutazioni, in base alla specie. A rigore, le mutazioni possono potenzialmente alterare le caratteristiche e le proprietà osservate. Ciò è particolarmente importante per garantire una maggiore riproducibilità dei dati, specialmente qualora si desideri accoppiare studi sperimentali e computazionali.

Per le simulazioni in questa Tesi si è deciso di adottare la topologia della *L-lattato deidrogenasi* estratta dai muscoli del coniglio per l'*LDH*, reperita dalla RCSB Protein Data Bank, PDB 5KKC [55]. Peraltro, essa è già stata impiegata in uno studio sulla formulazione di tale proteina, combinando gli approcci sperimentali e computazionali [7]. Diversamente, per la *GCSF* è stata adottata la topologia della proteina presente nell'uomo, ottenuta nuovamente dalla RCSB Protein Data Bank, PDB 1CD9 [56].

Una volta reperiti i file topologici necessari in formato *.pdb*, è necessario convertirli in formato *.gro*, per poter essere processati. Si usa il comando:

gmx pdb2gmx -f protein.pdb -o protein.gro -water tip3p -ignh

mediante il quale si converte il file topologico iniziale *protein.pdb* nel formato desiderato *protein.gro*, selezionando il *force-field* impiegato e la topologia dell'acqua, come visto nel paragrafo §3.1. A questo punto è possibile proseguire con la composizione del dominio computazionale. Inizialmente si inserisce la proteina in un

box, centrandone la posizione in corrispondenza del centro geometrico del volume. Il comando corrispondente è:

gmx editconf -f protein.gro -o protein-box.gro -c -d 1.0 -bt cubic

dove *protein-box.gro* è il box in output, di forma cubica (opzione *-bt cubic*), al cui centro vi è la proteina, posizionata mediante l'opzione *-c* a distanza di 1 nm da tutte le facce (opzione *-d 1.0*). Essendo l'*LDH* e la *GCSF* proteine di dimensioni diverse, le dimensioni del dominio computazionale saranno necessariamente diverse. Allo stesso modo, il numero di ciclodestrine presenti nel box varia, al fine di avere la stessa concentrazione, pari a 50 mM, per entrambe le tipologie di sistemi. I dati in questione sono riportati in Tabella 4.1.

Tabella 4.1: Caratteristiche dei sistemi simulati.

Proteina	*Lato box, nm*	*N^o CD*
LDH	11.72232	47
GCSF	7.29206	11

Si ricorda che il numero di ciclodestrine viene ottenuto per arrotondamento all'intero più vicino in quanto, evidentemente, non è possibile inserire porzioni di molecole per raggiungere la concentrazione desiderata nel dominio computazionale.
Successivamente si procede in maniera analoga a quanto visto per gli amminoacidi. Si aggiunge un numero *N* di ciclodestrine in ogni box:

gmx insert-molecules -ci CD.gro -f protein-box.gro -nmol N -o protein-CD-box.gro

da cui si ottiene in output il file *protein-CD-box.gro*. Si procede quindi alla solvatazione, mediante il comando:

gmx solvate -cp protein-CD-box.gro -cs spc216.gro -o protein-CD-w-box.gro -p topol.top

da cui si ottiene il box solvatato *protein-CD-w-box.gro*, assieme al corrispondente file di topologia aggiornato. Tuttavia, le proteine in questione presentano delle cariche, pertanto è necessario aggiungere degli ioni per garantire la neutralità del dominio computazionale:

gmx grompp -f enermin.mdp -c protein-CD-w-box.gro -o ions.tpr -p topol.top
gmx genion -s ions.tpr -pname NA -nname CL -np N_P -nn N_N -p topol.top -o box.gro

Tabella 4.2: Cariche e pH dei sistemi simulati.

Proteina	*pH*	*Ioni aggiunti*	N^o *ioni*
LDH	7	Cl^-	13
GCSF	4.5	Cl^-	7

da cui si ottiene il file *box.gro*, contenente ioni di tipologia e numerosità adatti a garantire la neutralità del sistema. Si aggiorna di conseguenza la topologia, poiché gli ioni aggiunti sostituiscono un numero corrispondente di molecole d'acqua. Le informazioni riguardanti le cariche dei sistemi sono riportate in Tabella 4.2.
A questo punto è possibile procedere con la minimizzazione energetica del box, impiegando i comandi:

gmx grompp -f enermin.mdp -c box.gro -o em.tpr -p topol.top
gmx mdrun -v -deffnm em

dove il file *em.gro* contiene la configurazione del sistema minimizzato. Come visto in precedenza, si prosegue con l'equilibratura, prima di poter effettuare la simulazione finale. Il canone di riferimento usato è l'*NPT* [19]. Per l'equilibratura si accoppiano il barostato ed il termostato di Berendsen [33], a 300 K ed 1 bar, per una durata di 1 ns (time-step 2 fs), generando casualmente la distribuzione di *Maxwell-Boltzmann* per le velocità.

gmx grompp -f npt.mdp -c em.gro -p topol.top -o npt.tpr
gmx mdrun -v -deffnm npt

Si verifica che l'energia, la temperatura e la pressione raggiungano, a meno di oscillazioni trascurabili, un valore stazionario, mediante il comando *gmx energy*, eventualmente estendendo la durata dell'equilibratura, secondo quanto visto nel paragrafo §2.1.1.

Infine, si procede alla simulazione, per una durata di 100 ns (time-step 2 fs), a 300 K ed 1 bar, impiegando il termostato di Nosé-Hoover [34] [35] [36] ed il barostato di Parrinello-Rahman [37]. Si riprende la distribuzione di velocità generata al passo precedente, mediante l'opzione *-t npt.cpt*.

gmx grompp -f md.mdp -c npt.gro -p topol.top -o md.tpr -t npt.cpt
gmx mdrun -v -deffnm md

Si ottengono, infine, la traiettoria e la configurazione dell'ultimo stato del sistema, rispettivamente *md.xtc* e *md.gro*.
Per concludere, si riportano le immagini relative ai sistemi *CD-LDH* (Figure 4.3, 4.4, 4.5 e 4.6) e *CD-GCSF* (Figure 4.7, 4.8, 4.9 e 4.10).

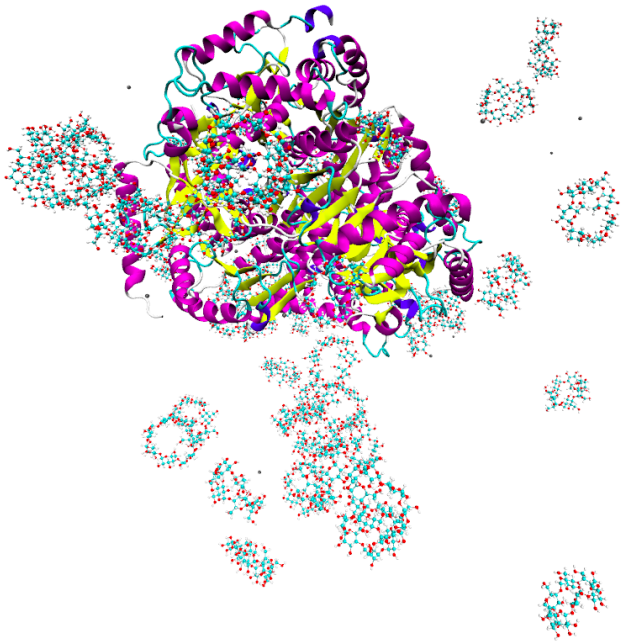

Figura 4.3: Rappresentazione del sistema β-*CD-LDH*. Immagine generata con VMD [13].

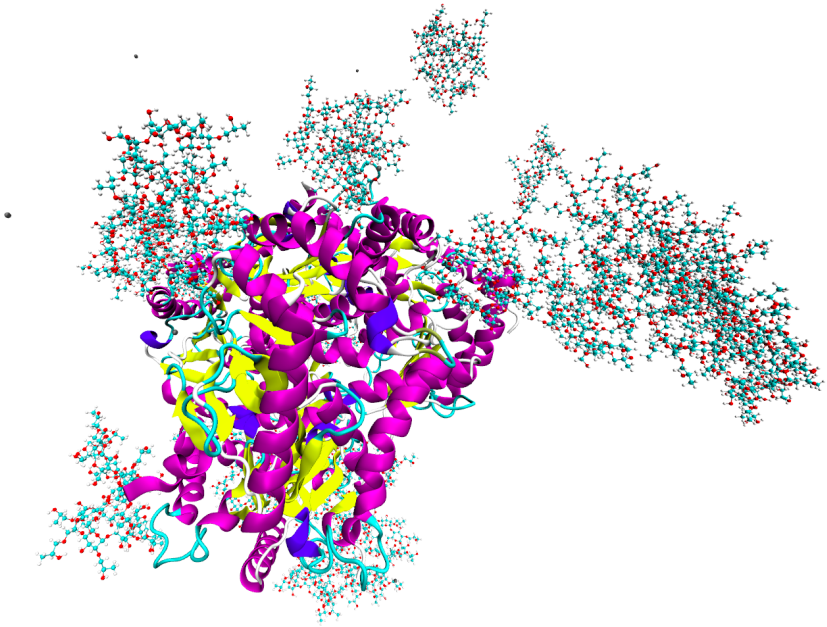

Figura 4.4: Rappresentazione del sistema *HP-*β*-CD-LDH*. Immagine generata con VMD [13].

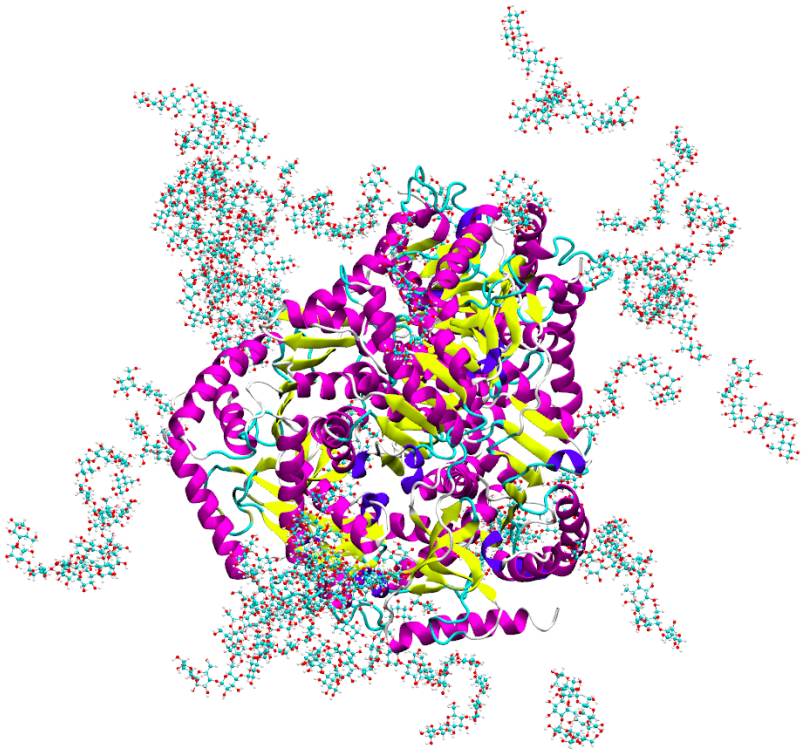

Figura 4.5: Rappresentazione del sistema open-β-*CD-LDH*. Immagine generata con VMD [13].

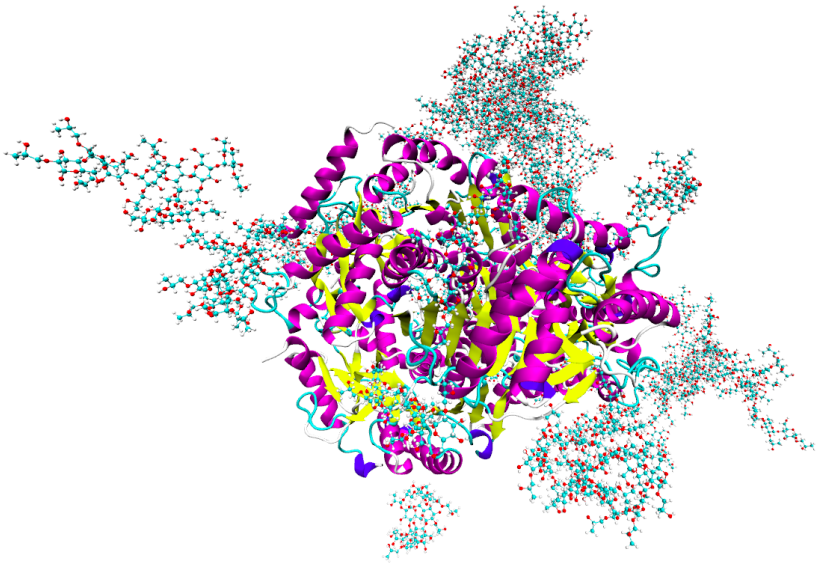

Figura 4.6: Rappresentazione del sistema open-*HP*-β-*CD-LDH*. Immagine generata con VMD [13].

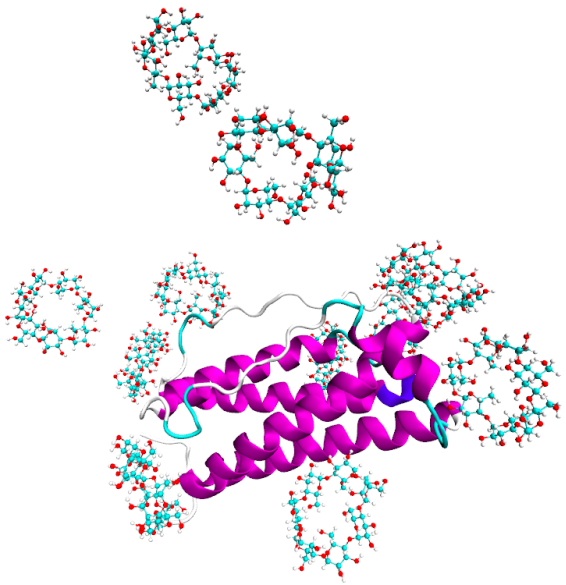

Figura 4.7: Rappresentazione del sistema *β-CD-GCSF*. Immagine generata con VMD [13].

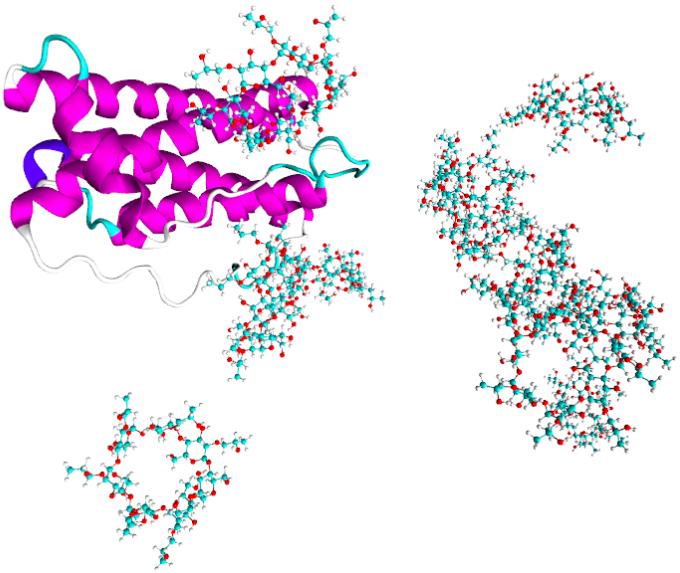

Figura 4.8: Rappresentazione del sistema *HP-β-CD-GCSF*. Immagine generata con VMD [13].

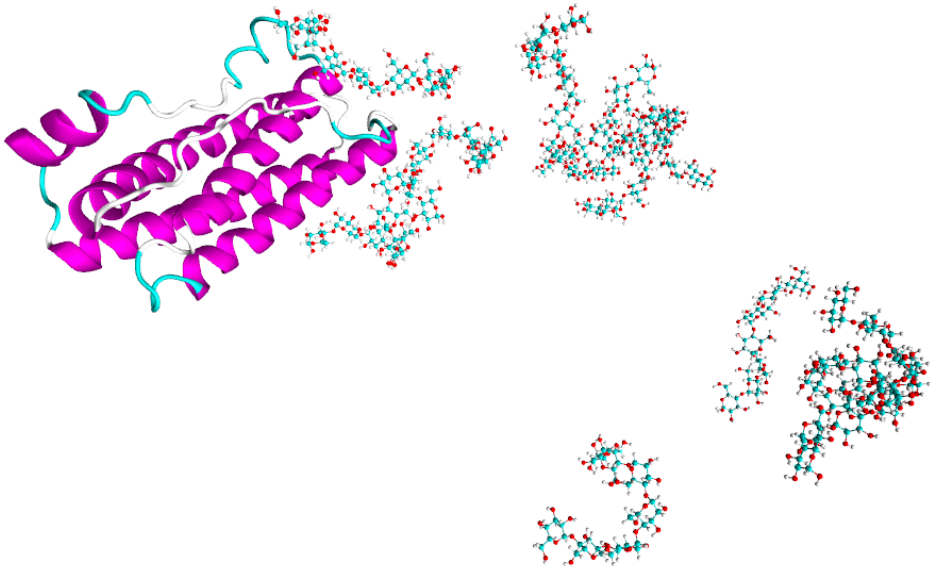

Figura 4.9: Rappresentazione del sistema open-β-*CD-GCSF*. Immagine generata con VMD [13].

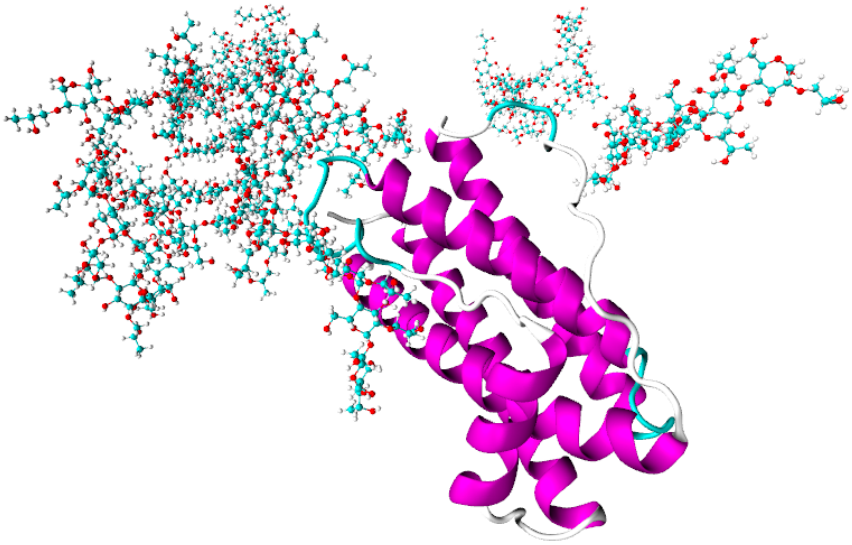

Figura 4.10: Rappresentazione del sistema open-*HP*-β-*CD-GCSF*. Immagine generata con VMD [13].

Analisi dei Risultati

L'analisi dei sistemi ciclodestrine-proteina è più complessa ed articolata rispetto a quanto visto in precedenza, pertanto è necessario introdurre le diverse tipologie di risultati ed i comandi corrispondenti.

La prima tipologia di analisi è la *radial distribution function* cumulativa, differente rispetto a quella vista nel paragrafo §3.1. Infatti, invece di ottenere il profilo di densità di una specie chimica nei confronti di un'altra, normalizzato rispetto al valore in bulk, essa viene espressa come numero totale di atomi in funzione della distanza tra due gruppi, entrambi appartenenti all'indice del sistema. La *rdf* cumulativa viene calcolata per due coppie, ovvero ciclodestrine-proteina e proteina-acqua, usando i comandi:

gmx rdf -f md.xtc -s md.tpr -selrpos mol_com *-surf mol -cn rdf-protein-CD-cn.xvg -o rdf-protein-CD.xvg -b 20000 -dt 10 -n index.ndx*

gmx rdf -f md.xtc -s md.tpr -selrpos mol_com *-surf mol -cn rdf-protein-water-cn.xvg -o rdf-protein-water.xvg -b 20000 -dt 10 -n index.ndx*

dove, in aggiunta a quanto visto nel paragrafo §3.1, si ottengono le *rdf* cumulative in output per le due coppie precedentemente indicate, *rdf-protein-CD-cn.xvg* e *rdf-protein-water-cn.xvg*.
Alcune opzioni aggiuntive sono:

- *-surf mol*, con la quale si prende la superficie delle molecole appartenenti al gruppo di riferimento per il calcolo della distanza della seconda specie chimica selezionata;
- *-b 20000*, per trascurare il transitorio dei primi 20 ns;
- *-dt 10*, per ridurre il carico computazionale;
- *-n index.ndx*, per la selezione dei gruppi atomici di interesse.

Le *rdf* cumulative ottenute sono utilizzate diversamente da quanto visto nel caso degli amminoacidi, poiché vengono impiegate per calcolare il profilo di un parametro chiamato β*-parameter*, il quale indica il grado di esclusione preferenziale. Tale approccio è stato adottato in quanto proposto in altri studi affini a questa Tesi, ottenendo una buona corrispondenza con i dati sperimentali [7]. Il β*-parameter* può essere espresso come:

$$\beta(r) = \frac{n_{CD}(r)/n_W(r)}{n_{CD}(\infty)/n_W(\infty)} \tag{4.1}$$

dove il valore di tale parametro è in funzione della distanza dalla superficie della proteina, presa come riferimento. A numeratore si trova il rapporto tra il numero medio di molecole di ciclodestrine e di acqua a distanza *r* dalla superficie proteica, mentre a denominatore si trova lo stesso rapporto stimato a distanza virtualmente infinita. Essendo il sistema analizzato finito, la massima distanza dalla proteina evidentemente non può superare i confini del dominio computazionale. Ciò implica che il valore al denominatore sia dato semplicemente dal rapporto tra il numero di *CD* e di acqua presenti nel box, pertanto può essere facilmente stimato avendo a disposizione il file di topologia *topol.top*.

Il parametro β fornisce quindi un'indicazione del grado di esclusione di un eccipiente, rispetto ad una proteina. Esso ha valenza relativa, ovvero viene utilizzato per comparare la capacità stabilizzante di eccipienti diversi. In particolare, se il parametro, ad un valore di distanza fissato, è maggiore di 1 si ha inclusione preferenziale, diversamente si verifica esclusione. Per ulteriori dettagli su questo meccanismo di stabilizzazione si rimanda al paragrafo §1.3.1.

Un dato di interesse per i sistemi ciclodestrine-proteina è la distanza minima degli eccipienti dai residui peptidici. Questo tipo di informazione, infatti, può fornire un grado di dettaglio maggiore rispetto al β*-parameter*. Per fare ciò, si calcola la distanza minima di tutti gli amminoacidi facenti parte del polipeptide, rispetto alle molecole di *CD*. Il comando utilizzato è:

gmx mindist -f md.xtc -s md.tpr -b 20000 -dt 10 -od mindist.xvg -or mindist-res.xvg -n index.ndx

dove l'opzione *-od mindist.xvg* fornisce la distanza minima, calcolata per ogni istante temporale, tra gli atomi appartenenti alla proteina e quelli appartenenti alle ciclodestrine, mentre *-or mindist-res.xvg* fornisce la distanza minima registrata tra ogni residuo peptidico e gli eccipienti. In particolare, utilizzando opportunamente il file indice *index.ndx*, è possibile fare una distinzione per ciascuna categoria o specie di *AA*. Queste informazioni sono utili per studiare l'affinità tra le *CD* e gli amminoacidi che costituiscono la biomolecola, pertanto saranno affiancate ai dati relativi alle inclusioni, per verificare un'eventuale corrispondenza con le proprietà osservate nel capitolo precedente. Per procedere in questo senso è necessario valutare la formazione di inclusioni. La procedura è sostanzialmente simile a quella vista nel paragrafo §3.1, tuttavia con alcune differenze che verranno qui messe in evidenza:

gmx select -s md.tpr -f md.gro -n index.ndx -on index-temp-N.ndx -select 'group 'Protein'; "CD-N" atomnr a to b'

gmx select -s md.tpr -f md.xtc -b 20000 -n index-temp-N.ndx -oi incl-atoms-N.dat -resnr number -os incl-atom-N.xvg -select 'group Protein and within 0.5 of com of group CD-N'

dove i termini contenenti *N* si riferiscono alla numerazione progressiva della ciclodestrina considerata. I file indice *index-temp-N.ndx* contengono il numero sequenziale degli atomi della proteina, assieme a quelli della *CD* in considerazione. Tali file fungono da input per il secondo comando, il quale valuta, mediante l'opzione *-os incl-atom-N.xvg*, la quantità di atomi appartenenti ai residui peptidici inclusi da ogni singola ciclodestrina, per ogni istante analizzato. In particolare, ciò avviene quando un atomo appartenente alla proteina intercetta un intorno sferico, di raggio 0.5 nm, centrato nel centro di massa delle *CD*. L'opzione *-oi incl-atoms-N.dat* genera un file in output che contiene, per ogni istante, la quantità di inclusioni ed il numero sequenziale di ogni atomo incluso, mentre la voce *-resnr number* consente di adottare la stessa numerazione dei file di indice in input, *index-temp-N.ndx*. Ciò viene fatto in modo da poter successivamente identificare i residui di appartenenza, quantificando le occorrenze per ciascuno di essi. Infine, si ottiene la lista di amminoacidi, con il numero corrispondente nella sequenza peptidica della proteina, assieme alla quantità di inclusioni occorse in termini di istanti temporali. Si individuano quindi i residui inclusi almeno per l'1% del tempo analizzato. Tuttavia, a rigore, ciascuna ciclodestrina può formare un numero di inclusioni diverse, pertanto, per poter confrontare i valori per i vari eccipienti, è necessario utilizzare un criterio comune per la normalizzazione delle occorrenze. Tale risultato si ottiene dividendo ciascun valore per il prodotto tra la quantità di *CD* presenti ed il numero totale di istanti analizzati. Infine, come per gli amminoacidi, si calcola la percentuale di ciclodestrine a catena chiusa che forma inclusioni con i residui peptidici.

Vi sono altri due aspetti di interesse per quanto concerne le proprietà delle proteine, ovvero il raggio di giro ed il rapporto tra la superficie proteica idrofobica e quella totale esposte al solvente. Queste informazioni, in particolare la loro evoluzione nel tempo, forniscono indicazioni sullo stato della biomolecola che consentono di rilevare l'eventuale insorgenza di fenomeni di denaturazione. Per poter effettuare tali valutazioni, si calcolano le medie temporali per il raggio di giro e per il rapporto delle superfici, confrontandoli ai valori massimi e minimi.
Il raggio di giro R_g può essere espresso secondo quanto segue:

$$R_g = \left(\frac{\sum_{i \in prot}^{N_{AP}} \|r_i\|^2 m_i}{\sum_{i \in prot}^{N_{AP}} m_i} \right)^{\frac{1}{2}} \tag{4.2}$$

dove a numeratore si ha la sommatoria del prodotto tra la massa di ciascun atomo (m_i) appartenente alla proteina, ed il quadrato della posizione corrispondente (r_i), definita rispetto al centro di massa della molecola stessa, mentre a denominatore si ha la sommatoria delle masse di tutti gli atomi, ovvero la massa totale. Il comando utilizzato è:

gmx gyrate -f md.xtc -s md.tpr -o gyrate.xvg -b 20000 -n index.ndx

gmx analyze -f gyrate.xvg

dove il file di output *gyrate.xvg* contiene il raggio di giro e le sue proiezioni nelle direzioni *x*, *y* e *z*, in funzione del tempo simulato. Analogamente a quanto visto per altre analisi, si escludono i primi 20 ns mediante l'opzione *-b 20000*. Infine, si calcola la media temporale mediante il comando *gmx analyze*, confrontando il risultato con i valori massimi e minimi registrati per ogni sistema.

Il rapporto tra la superficie idrofobica e quella totale della proteina si ottiene dal seguente comando, che sfrutta un algoritmo dedicato [57]:

gmx sasa -f md.xtc -s md.tpr -surface 'group "Protein"' -output '"Hydrophobic" group "Protein" and charge -0.2 to 0.2; "Hydrophilic" group "Protein" and not charge -0.2 to 0.2; "Total" group "Protein"' -o sasa.xvg -n index.ndx -b 20000

gmx analyze -f sasa.xvg

con il quale si analizzano gli atomi facenti parte della proteina, impostando un intervallo di cariche per definire i limiti che identificano e distinguono i gruppi idrofobici da quelli idrofili. Si ottiene un file di output *sasa.xvg*, contenente la superficie totale esposta al solvente, la superficie idrofobica ed idrofila, per ogni istante analizzato. Si calcola quindi facilmente il rapporto desiderato, in funzione del tempo simulato. Analogamente a quanto visto per il raggio di giro, si stima il valore medio mediante *gmx analyze* e lo si confronta con i valori di massimo e minimo. L'uso del comando *gmx sasa* è stato già effettuato in altri studi affini, per valutare cambiamenti nella conformazione delle proteine in presenza o assenza di eccipienti e/o surfattanti [40].

Vi è un ultimo aspetto da considerare per completare le analisi , ovvero la formazione dei *cluster* di ciclodestrine. La procedura impiegata è la stessa descritta nei paragrafi §2.1.1, §2.2.1 e §3.1, pertanto si riportano in breve i comandi necessari e la tipologia di dati di interesse.

gmx trjconv -f md.xtc -s md.tpr -b 20000 -dt 10 -n index.ndx -o traj-CD.xtc -pbc whole

gmx trjconv -f md.xtc -s md.tpr -b 100000 -e 100000 -n index.ndx -o traj-CD.gro

gmx $make_ndx$ *-f traj-CD.gro -o index-CD.ndx*

gmx grompp -f enermin.mdp -c traj-CD.gro -o CD.tpr -p CD.top

gmx clustsize -f traj-CD.xtc -s CD.tpr -n index-CD.ndx -mol -hc clust-CD.xvg -nc nclust.xvg

da cui si ottengono la distribuzione delle *CD* impegnate in *cluster*, contenuta nel file *clust-CD.xvg*, dal quale si calcola anche la dimensione media, e, infine, il numero di aggregati presenti in funzione del tempo, dal file *nclust.xvg*.

4.3 Risultati ed osservazioni

Vista la numerosità dei risultati da presentare, nonché la diversa natura delle due proteine in esame, si è deciso di presentare i dati separatamente, rimandando le conclusioni di carattere generale al paragrafo finale di questo capitolo.

4.3.1 LDH

Il β-*parameter* costituisce la prima tipologia di dati presentati per lo studio delle interazioni tra ciclodestrine e proteine, allo scopo di valutare e confrontare il grado di esclusione degli eccipienti dalla superficie proteica. Il grafico per l'*LDH* è riportato in Figura 4.11.

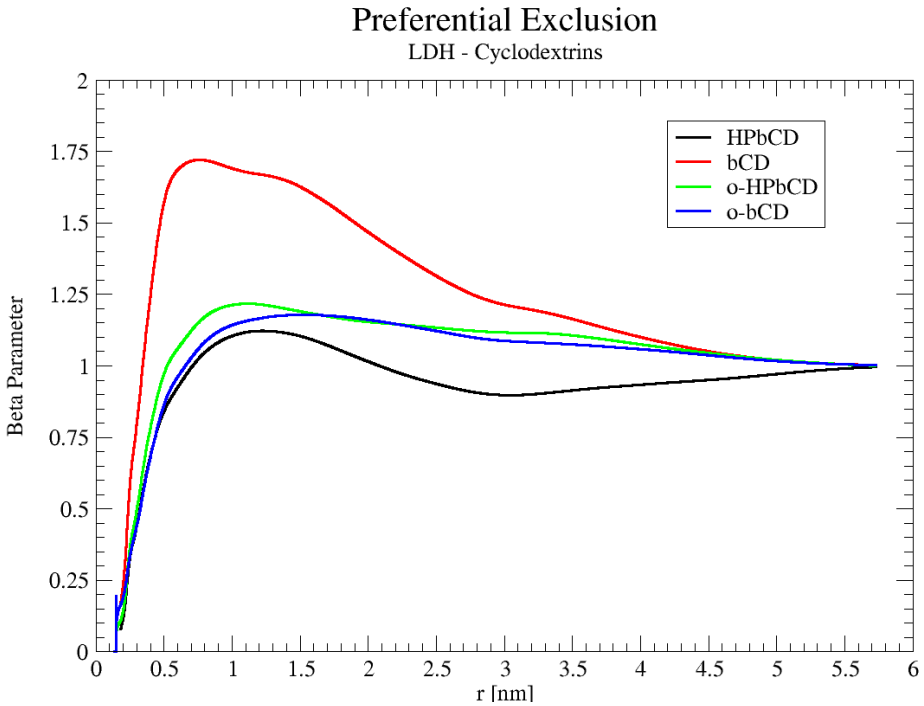

Figura 4.11: Profilo del β-*parameter* per i sistemi *CD-LDH*.

Osservando la Figura 4.11 si nota immediatamente che la β-*CD* abbia una maggiore propensione ad essere inclusa dalla proteina, mentre le restanti ciclodestrine sono scarsamente incluse a brevi distanze. Tale risultato può essere giustificato dalla formazione di inclusioni tra la β-*CD* ed i residui peptidici esposti al solvente. Pertanto, le altre *CD* potrebbero essere impiegate come eccipienti, esercitando esclusione preferenziale ed ostacolando la denaturazione della proteina per *unfolding*. Inoltre, è stato ipotizzato che le ciclodestrine, in particolare quelle idrossipropilate, possano essere più efficaci come lioprotettori [6]. Infatti, durante il congelamento

la crio-concentrazione favorirebbe la formazione di legami idrogeno tra gli eccipienti e la proteina, stabilizzandola mediante almeno uno dei possibili meccanismi descritti nel paragrafo §1.3.2 [2] [58]. Diversamente, la β-*CD* potrebbe rivelarsi più efficace nell'impedire l'aggregazione, qualora essa coinvolga residui idrofobici, specialmente se aromatici, verso cui tale ciclodestrina mostra grande affinità come visto nel paragrafo §3.2 [59]. Tuttavia tale affermazione necessita di ulteriori dati a sostegno, per cui si procede ad analizzare la distanza minima tra le *CD* e la proteina stessa, facendo distinzione tra i vari tipi di amminoacidi presenti, in modo da riscontrare l'eventuale presenza di interazioni superficiali.

Il comando *gmx mindist* fornisce la minima distanza riscontrata tra due gruppi atomici di riferimento. In particolare, si è scelto di focalizzarsi su uno dei due file di output disponibili, *mindist-res.xvg* (si veda il paragrafo §4.2). Si ricorda che questo file contiene, per ogni residuo peptidico, la distanza minima riscontrata rispetto alle ciclodestrine. Si è quindi scelto di riportare la media delle distanze minime, calcolata su tutti gli amminoacidi della catena peptidica. Infatti, vista la numerosità degli *AA* presenti, sarebbe poco pratico raffigurare i risultati distinguendo ogni unità. Tuttavia, per tenere conto della natura dei residui, si è scelto di distinguere e confrontare i risultati ottenuti per la proteina intera e per le classi di amminoacidi, in particolare discriminando le specie aromatiche da quelle alifatiche per il gruppo idrofobico. Inoltre, si è scelto di includere la tirosina tra gli *AA* aromatici, sebbene convenzionalmente faccia parte di quelli polari, poiché l'interesse è posto verso la tipologia di interazione con le ciclodestrine. Infatti, dai risultati presentati nel paragrafo §3.2, assieme alle relative considerazioni sul ruolo della struttura molecolare, si può ragionevolmente supporre che l'interazione sia da imputarsi all'affinità idrofobica tra la cavità delle *CD* e la *sidechain* della *TYR*. Tale affermazione è sostenuta dai massimi riscontrati tra le *rdf*, i quali evidenziano un'affinità significativamente più spiccata nei confronti della catena laterale.
I dati sono rappresentati in Figura 4.12.
Si può osservare che la media delle distanze minime complessiva per l'*LDH* sia sostanzialmente analoga a quella dei residui polari. Ciò suggerirebbe una prevalenza di tale gruppo in superficie, tuttavia tali considerazioni saranno fatte dopo aver presentato i dati relativi alle inclusioni ed alle proprietà superficiali. Complessivamente, le specie di amminoacidi che mediamente si trovano a distanza inferiore sono quelle cariche, mentre per quelle idrofobiche il comportamento sembrerebbe dettato dagli *AA* alifatici, sebbene gli aromatici risultino più vicini. Infine, si può notare come la *HP*-β-*CD* sia, in tutti i casi proposti, la ciclodestrina più distante. Ciò risulta in accordo con i risultati della Figura 4.11, giustificando la maggiore tendenza all'esclusione di questo eccipiente.

Si prosegue con l'analisi delle inclusioni, che verrà presentata solo per le ciclodestrine a catena chiusa. Sebbene sia possibile effettuare lo stesso tipo di valutazione anche per le *CD* a catena aperta, utilizzando i medesimi comandi, ciò non avrebbe senso fisico per tali specie chimiche.
Per queste analisi si presentano due tipologie di dati per ogni eccipiente considerato:

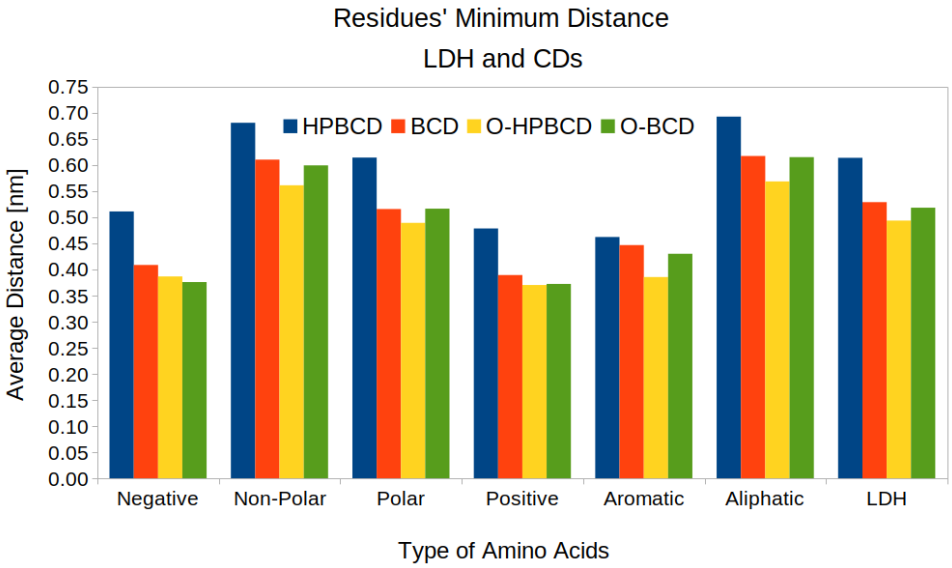

Figura 4.12: Media delle distanze minime tra ciclodestrine e residui peptidici per i sistemi *CD-LDH*.

- una tabella riassuntiva contenente la specie di residuo, il numero sequenziale corrispondente nella catena polipeptidica e la durata dell'inclusione, espressa come percentuale temporale, secondo il criterio descritto nel paragrafo §4.2;

- un'immagine della proteina in cui sono messi in evidenza i residui inclusi, colorati a seconda della categoria di appartenenza.

I dati relativi alla *β-CD* sono riportati in Tabella 4.3 ed in Figura 4.13.
Si può notare come le interazioni più significative si instaurino con amminoacidi polari, apolari e carichi positivamente.
I dati relativi alla *HP-β-CD* sono riportati in Tabella 4.4 ed in Figura 4.14.
Si osserva che le interazioni, in questo caso, siano stabilite prevalentemente con i residui carichi positivamente e polari.

Confrontando i risultati ottenuti per le due ciclodestrine, si nota immediatamente come la *β-CD* interagisca maggiormente, formando un numero maggiore di inclusioni rispetto a quelle dell'altro eccipiente, e più durature. Ciò risulta in accordo con il profilo del *β-parameter* mostrato precedentemente (Figura 4.11), secondo cui la *HP-β-CD* sarebbe più esclusa. Inoltre, una conferma ulteriore deriva dal grafico delle distanze minime, in base al quale, per tutti i gruppi analizzati, la *CD* funzionalizzata sarebbe mediamente più distante. Entrambe queste considerazioni giustificherebbero le proprietà osservate per le ciclodestrine a catena chiusa. Infine, a sostegno di tali affermazioni si propone una tabella riassuntiva, Tabella 4.5, nella quale si riporta la percentuale di eccipienti che formano inclusioni,

Tabella 4.3: Residui inclusi per il sistema β-*CD-LDH*.

Residuo	*N^o del residuo*	*Durata inclusione, %*
PRO	128	2.50
HIS	129	1.96
LEU	12	1.68
LYS	13	0.30
ILE	241	1.03
VAL	52	3.69
MET	53	4.18
LYS	317	0.47
PRO	215	0.89
VAL	109	0.27
TYR	238	1.39
ASN	114	1.19

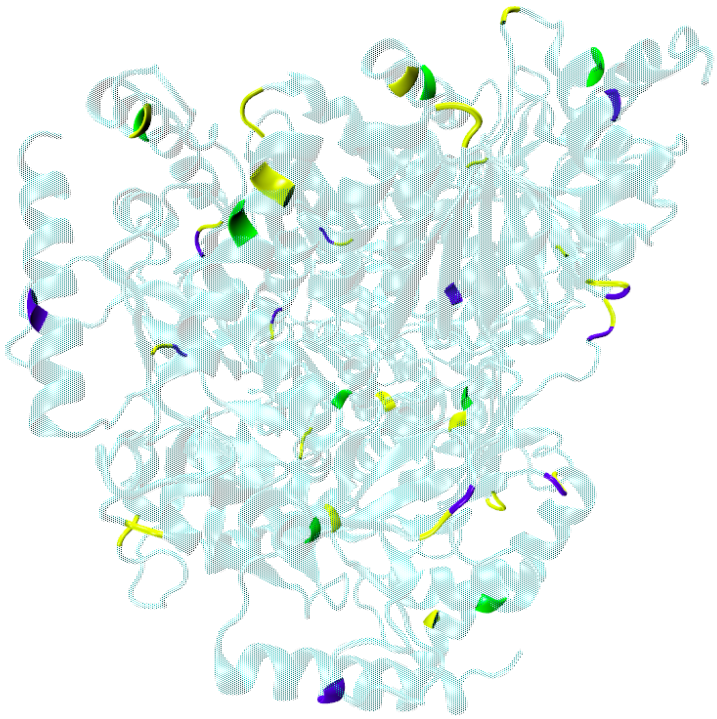

Figura 4.13: Rappresentazione dei residui peptidici inclusi per il sistema β-*CD-LDH*. Legenda: *AA* carichi positivamente in blu, *AA* carichi negativamente in rosso, *AA* polari in verde, *AA* apolari in giallo. Immagine generata con VMD [13].

Tabella 4.4: Residui inclusi per il sistema *HP-β-CD-LDH*.

Residuo	*N^o del residuo*	*Durata inclusione, %*
HIS	129	0.97
GLY	151	0.05
LEU	108	0.03
ARG	98	0.02
GLN	110	0.16
ARG	176	0.24
ASN	114	0.20

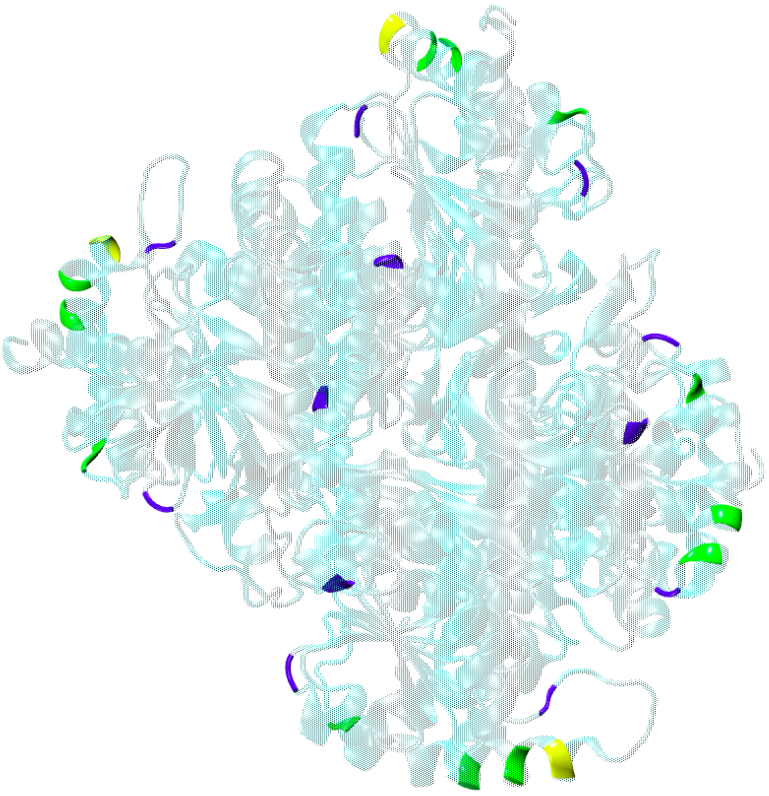

Figura 4.14: Rappresentazione dei residui peptidici inclusi per il sistema *HP-β-CD-LDH*. Legenda: *AA* carichi positivamente in blu, *AA* carichi negativamente in rosso, *AA* polari in verde, *AA* apolari in giallo. Immagine generata con VMD [13].

analogamente a quanto visto per gli amminoacidi nel paragrafo §3.2.2. Si conclude quindi che la *β-CD* formi più agevolmente delle inclusioni, grazie alla minore distanza dalla proteina.

Il raggio di giro rappresenta un dato di fondamentale importanza per capire se la

Tabella 4.5: Confronto della percentuale di ciclodestrine che formano inclusioni per i sistemi *CD-LDH*.

CD	*CD* con inclusioni, %
β-CD	14.04
HP-β-CD	1.74

proteina sia soggetta a denaturazione, poiché durante lo svolgimento parziale o completo della molecola si riscontrano grandi variazioni di questa variabile. Alcuni eccipienti possono interagire con le catene laterali dei residui peptidici, favorendo il dispiegamento della catena polipeptidica in prossimità di questi ultimi, specialmente quando le interazioni coinvolgono gruppi idrofobici, pertanto poco affini all'acqua. Per questa ragione si valuta il lo scostamento del valore medio del raggio di giro rispetto ai valori di massimo e minimo. I dati sono riportati in Tabella 4.6.

Tabella 4.6: Confronto del raggio di giro della proteina per i sistemi *CD-LDH*.

CD	Minimo, nm	Media, nm	Massimo, nm
β-CD	3.099	3.122	3.132
HP-β-CD	3.092	3.121	3.151
open-*β-CD*	3.104	3.128	3.146
open-*HP-β-CD*	3.100	3.131	3.154

Osservando i valori si evince che le variazioni del raggio di giro per i sistemi *CD-LDH* siano poco significative, ovvero non mostrino segni evidenti di denaturazione. Ciò è di fondamentale importanza per valutare la compatibilità tra eccipienti e proteina, poiché, qualora essi favorissero conformazioni proteiche diverse da quelle nativa, risulterebbero non solo inefficaci come stabilizzanti, bensì dannosi per la preservazione del biofarmaco. Analogamente, le stesse considerazioni sono valide per la superficie esposta al solvente, in quanto variazioni significative dell'estensione, oppure della polarità, rappresentano segnali di denaturazione. Seguendo lo stesso criterio, si riportano in Tabella 4.7 i dati relativi al rapporto tra la superficie idrofobica della proteina e quella totale, esposte al solvente.

Tabella 4.7: Confronto del rapporto tra superficie proteica idrofobica e totale, esposte al solvente, per i sistemi *CD-LDH*.

CD	Minimo, %	Media, %	Massimo, %
β-CD	47.29	48.93	50.63
HP-β-CD	47.11	48.68	50.53
open-*β-CD*	47.18	48.62	50.07
open-*HP-β-CD*	47.13	48.68	50.03

I risultati presentati suggeriscono che l'*LDH* possieda una superficie prevalente-

mente polare, in accordo con quanto osservato per le distanze minime tra ciclodestrine e residui peptidici, in base ai quali il comportamento medio della proteina corrisponderebbe a quello degli amminoacidi polari presenti nella stessa. Inoltre, in precedenza si è fatto riferimento alla capacità delle ciclodestrine di impedire la denaturazione secondo diversi meccanismi, pertanto alla luce delle analisi svolte si può affermare che:

- la debole propensione all'inclusione preferenziale è giustificata dalla polarità della superficie, poiché tra i residui inclusi prevalgono sostanzialmente quelli apolari e carichi positivamente, mentre per quelli polari le interazioni sono piuttosto scarse, ad eccezione della tirosina. L'assenza di inclusioni significative tra la *HP-β-CD* e le specie idrofile, in aggiunta alla maggiore distanza media dalla proteina, favorirebbero questa ciclodestrina come candidato per il meccanismo di stabilizzazione menzionato;

- nonostante l'*LDH* presenti una superficie prevalentemente polare, la *β-CD* forma inclusioni più significative e risulta mediamente più vicina alla biomolecola rispetto alla *CD* funzionalizzata, pertanto si ritiene di poter scartare l'esclusione preferenziale come meccanismo di stabilizzazione esercitato dalla ciclodestrina nativa. Infatti, in questo caso risulta più plausibile ipotizzare che l'eccipiente in esame stabilizzi i residui superficiali tramite inclusioni, ostacolando lo svolgimento e l'interazione tra amminoacidi idrofobici, tipicamente capaci di causare aggregazione [60].

Infine, l'ultimo aspetto che completa lo studio delle proprietà delle ciclodestrine è l'analisi dei *cluster*. Come in precedenza, queste informazioni integrano ed aiutano a spiegare i risultati osservati, sebbene non influenzino le caratteristiche delle interazioni tra eccipienti e proteina.
Il primo aspetto da considerare è la distribuzione delle *CD* nei *cluster*, in base alla loro dimensione, riportata in Figura 4.15.

Dal grafico si evince che il comportamento qualitativo sia essenzialmente lo stesso riscontrato in precedenza (si vedano i paragrafi §2.1.2, §2.2.2 e §3.2). Le due *HP-β-CD* formano aggregati di dimensioni maggiori, necessariamente a scapito della numerosità, mentre le *β-CD* privilegiano *cluster* più numerosi ma meno massivi. Questi risultati sono infatti coerenti con le interazioni ciclodestrine-*LDH* osservate, in particolare per gli eccipienti a catena chiusa. La *β-CD* in questo modo è più libera di interagire con i residui peptidici, mentre la *CD* idrossipropilata predilige interazioni con le molecole della stessa specie, distanziandosi dalla proteina. A sostegno di queste affermazioni si propongono i dati riportati nelle Tabelle 4.8 e 4.9.

I risultati sono in accordo con quelli delle tabelle dei paragrafi §2.1.2, §2.2.2 e §3.2.

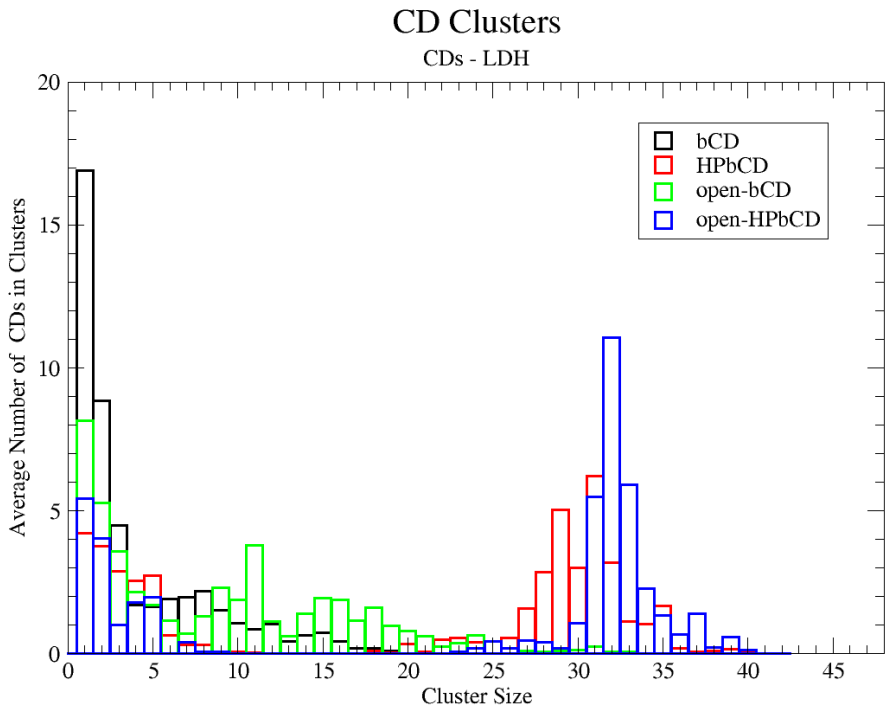

Figura 4.15: Distribuzione della dimensione dei *cluster* per le CD nei sistemi *CD-LDH*.

Tabella 4.8: N° medio di CD in un *cluster* per i sistemi *CD-LDH*.

HP-β-CD	*β-CD*	open-*HP-β-CD*	open-*β-CD*
8.63	3.76	10.02	5.98

Tabella 4.9: N° medio di *cluster* nel tempo per i sistemi *CD-LDH*.

HP-β-CD	*β-CD*	open-*HP-β-CD*	open-*β-CD*
9.45	25.07	9.71	15.00

4.3.2 GCSF

Lo studio delle interazioni tra proteine e ciclodestrine prosegue secondo le stesse modalità con la *GCSF*.

La prima tipologia di dati da esaminare è il profilo del *β-parameter*, la cui espressione è riportata nell'Eq. 4.2, mostrato in Figura 4.16.

Osservando il grafico si può notare una sostanziale differenza nel comportamento delle due tipologie di ciclodestrine, ovvero tra le *HP-β-CD* e le *β-CD*. Infatti, queste ultime risultano significativamente più incluse. Inoltre, si osserva che sia per la ciclodestrina nativa sia per quella funzionalizzata l'inclusione preferenziale è più spiccata per le forme a catena chiusa. Si può pertanto ipotizzare che l'esclusione preferenziale sia da scartare come possibile meccanismo di stabilizzazione delle

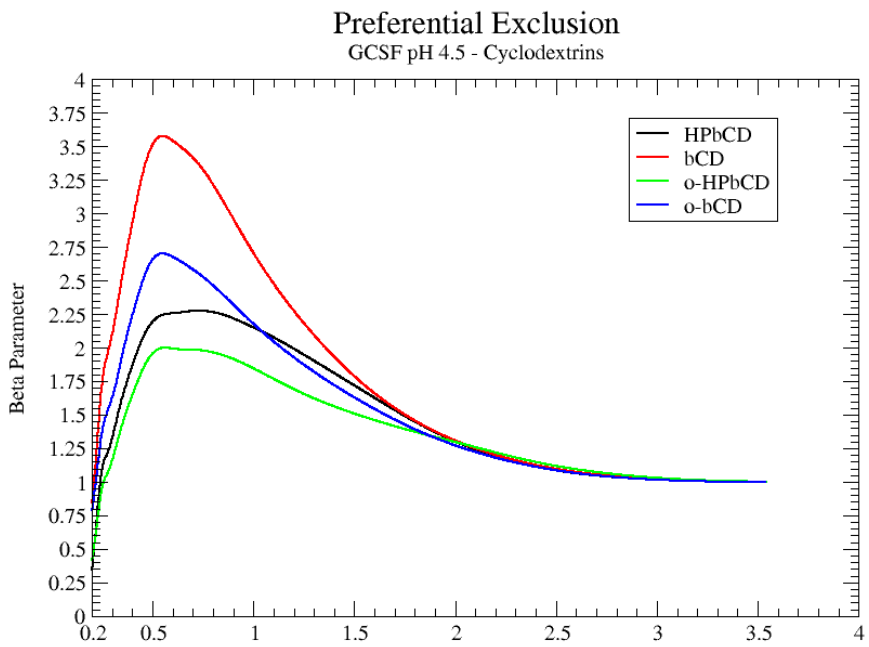

Figura 4.16: Profilo del β-*parameter* per i sistemi *CD-GCSF*.

CD in presenza della *GCSF*. Infatti, sembra più plausibile supporre che l'azione stabilizzante sia da imputarsi o ai fenomeni descritti nel paragrafo §1.3.2, oppure alla formazione di inclusioni superficiali da parte degli eccipienti, volte ad impedire l'*unfolding* ed il rischio di aggregazione annesso. In questo caso le proprietà degli eccipienti sembrano ben definite, perciò con le analisi seguenti si vogliono cercare ulteriori prove a sostegno del comportamento riscontrato.

Si prosegue con lo studio delle distanze minime tra ciclodestrine e residui peptidici. In particolare, si applicano gli stessi criteri discussi nel paragrafo §4.3.1, ovvero si presenta la media delle distanze minime per ogni amminoacido appartenente ad uno dei gruppi atomici in esame. I dati in questione sono rappresentati in Figura 4.17.

In questo caso le distanze sono nettamente inferiori a quelle riscontrate per l'*LDH* (Figura 4.12) e le differenze tra una ciclodestrina e l'altra sono meno marcate. Inoltre, si nota che il comportamento dei residui apolari corrisponde a quello generale della proteina, suggerendo quindi un carattere prevalentemente idrofobico, di cui si verificherà l'esistenza con le successive analisi relative alla superficie. Questa ipotesi è ulteriormente avvalorata dalla nota affinità tra le ciclodestrine ed i gruppi idrofobici [6], che giustificherebbe la notevole propensione all'inclusione superficiale degli eccipienti e, necessariamente, le distanze ridotte appena riscontrate. Tali considerazioni sono in accordo con il profilo del β-*parameter* in Figura 4.16. Per poter ottenere altre conferme è necessario studiare la formazione di inclusioni tra *CD* e *GCSF*. I dati vengono nuovamente riportati solo per le ciclodestrine a ca-

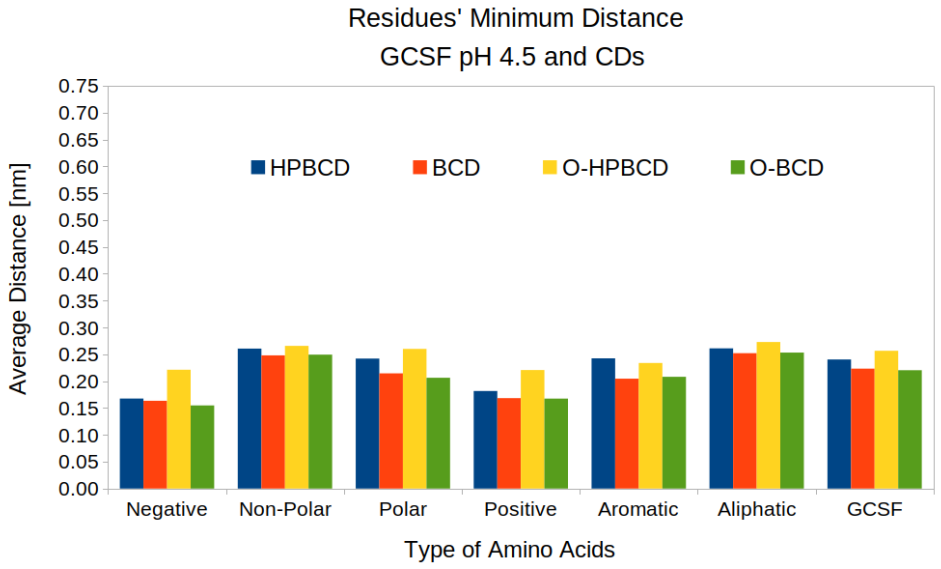

Figura 4.17: Media delle distanze minime tra ciclodestrine e residui peptidici per i sistemi *CD-GCSF*.

tena chiusa, poiché solo queste ultime sono in grado di inglobare residui peptidici all'interno della propria cavità. Analogamente al paragrafo precedente, i risultati sono costituiti da:

- una tabella riassuntiva contenente la specie di residuo, il numero sequenziale corrispondente nella catena polipeptidica e la durata dell'inclusione, espressa come percentuale temporale;
- un'immagine della proteina in cui sono messi in evidenza i residui inclusi.

Si riportano i dati relativi al sistema *β-CD-GCSF*, in Tabella 4.10 e Figura 4.18.
I risultati indicano che la *β-CD* formi inclusioni significative con residui appartenenti a tutte le categorie, sebbene le più numerose ed intense siano quelle con gli amminoacidi alifatici. Degne di menzione sono anche le interazioni con l'acido glutammico.
Si prosegue con i dati relativi al sistema *HP-β-CD-GCSF*, in Tabella 4.11 e Figura 4.19.
Per la *HP-β-CD* le inclusioni coinvolgono unicamente i residui polari ed apolari, in particolare gli alifatici. Come osservato per l'*LDH* nel paragrafo precedente, per la ciclodestrina funzionalizzata le interazioni sono meno numerose e meno intense, in accordo con il profilo del *β-parameter*, secondo il quale risulterebbe meno inclusa. Inoltre, la Figura 4.17 suggerisce una distanza minima maggiore rispetto a quella della *β-CD*, a supporto di tali affermazioni.
Per concludere lo studio delle inclusioni si è stimata la percentuale di ciclodestrine che formano inclusioni. I risultati sono riportati in Tabella 4.12.

Tabella 4.10: Residui inclusi per il sistema *β-CD-GCSF*.

Residuo	*N^o del residuo*	*Durata inclusione, %*
ALA	129	4.05
LEU	130	4.82
THR	133	0.76
GLN	145	0.83
ARG	146	4.89
GLU	46	2.37
CYS	36	0.38
LEU	50	2.49
LEU	61	0.40
SER	66	0.57
GLN	67	1.04
GLU	93	3.54
GLU	98	0.38
GLU	122	1.91
LEU	124	0.39
GLY	125	4.48

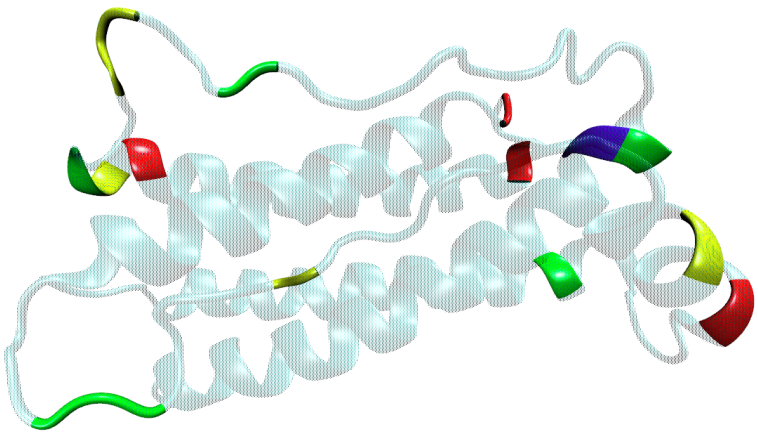

Figura 4.18: Rappresentazione dei residui peptidici inclusi per il sistema *β-CD-GCSF*. Legenda: *AA* carichi positivamente in blu, *AA* carichi negativamente in rosso, *AA* polari in verde, *AA* apolari in giallo. Immagine generata con VMD [13].

La quantità di *β-CD* impegnate in inclusioni è significativamente maggiore rispetto alla *HP-β-CD*, sebbene la differenza tra le due ciclodestrine sia ridotta rispetto ai sistemi *CD-LDH* (si veda la Tabella 4.5). I risultati ottenuti finora confermano

Tabella 4.11: Residui inclusi per il sistema *HP-β-CD-GCSF*.

Residuo	*N^o del residuo*	*Durata inclusione, %*
PRO	132	0.62
THR	133	1.12
GLN	134	2.86
GLY	135	2.08
SER	62	0.81
SER	65	0.42
PRO	97	1.28
PRO	101	0.92

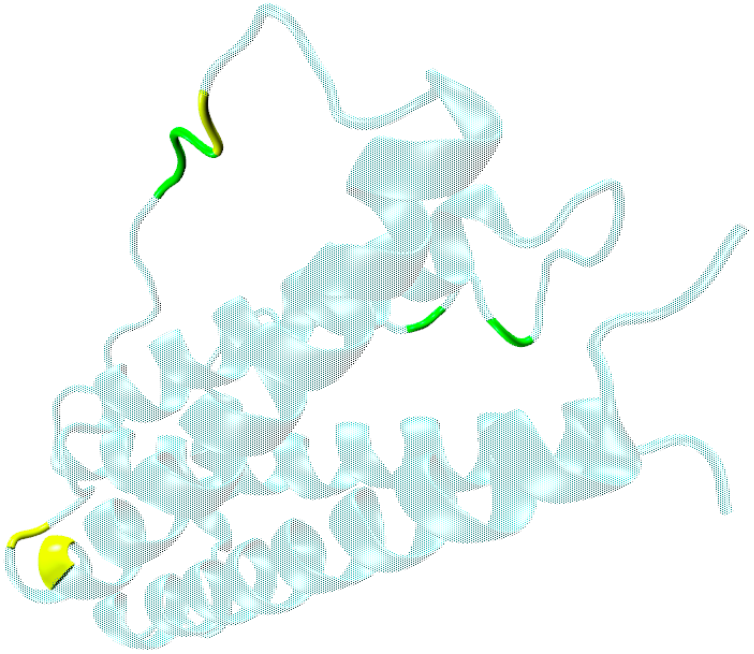

Figura 4.19: Rappresentazione dei residui peptidici inclusi per il sistema *HP-β-CD-GCSF*. Legenda: *AA* carichi positivamente in blu, *AA* carichi negativamente in rosso, *AA* polari in verde, *AA* apolari in giallo. Immagine generata con VMD [13].

la maggiore propensione della *CD* nativa ad includere significativamente residui peptidici, mentre per la forma funzionalizzata le interazioni con gli amminoacidi sono più deboli.
Si prosegue con lo studio degli aspetti strutturali della *GCSF*, in particolare degli effetti che gli eccipienti possono avere sulla proteina. Per fare ciò si analizzano nuovamente il raggio di giro ed il rapporto tra superficie idrofobica e totale esposte,

Tabella 4.12: Confronto della percentuale di ciclodestrine che formano inclusioni per i sistemi *CD-GCSF*.

CD	*CD* con inclusioni, %
β-CD	25.32
HP-β-CD	6.87

confrontando i valori medi con i massimi ed i minimi. La Tabella 4.13 contiene i dati relativi al raggio di giro.

Tabella 4.13: Confronto del raggio di giro della proteina per i sistemi *CD-LDH*.

CD	Minimo, nm	Media, nm	Massimo, nm
β-CD	1.610	1.659	1.698
HP-β-CD	1.639	1.671	1.706
open-*β-CD*	1.607	1.652	1.686
open-*HP-β-CD*	1.624	1.651	1.680

Le variazioni del raggio di giro sono piuttosto limitate, dello stesso ordine di grandezza di quelle osservate per l'*LDH*, sebbene abbiano un peso relativo maggiore, essendo i valori più bassi. Ciononostante esse non costituiscono segni evidenti di denaturazione. Si procede pertanto con l'analisi del rapporto tra la superficie idrofobica e quella totale, esposte al solvente, i cui dati sono riportati in Tabella 4.14.

Tabella 4.14: Confronto del rapporto tra superficie proteica idrofobica e totale, esposte al solvente, per i sistemi *CD-LDH*.

CD	Minimo, %	Media, %	Massimo, %
β-CD	50.78	54.26	57.59
HP-β-CD	51.39	54.71	58.44
open-*β-CD*	49.50	53.55	57.34
open-*HP-β-CD*	51.43	55.34	59.06

Si può notare che le variazioni relative al rapporto in esame siano più significative di quelle osservate per l'*LDH*. Inoltre, sembrerebbe che con entrambe le *HP-β-CD* aumenti lievemente il carattere apolare della superficie rispetto alle *β-CD*. Questi risultati suggeriscono che le proprietà superficiali siano prevalentemente idrofobiche. Ciò giustificherebbe la propensione di tutte le ciclodestrine a manifestare inclusione preferenziale, particolarmente più spiccata per la *β-CD* a catena chiusa, in linea con la nota affinità nei confronti dei residui alifatici, riscontrata anche per gli amminoacidi nel paragrafo §3.2.
In sintesi, tutti i dati raccolti sulle interazioni tra proteine e ciclodestrine forniscono le stesse indicazioni in merito alle proprietà degli eccipienti analizzati. Le *CD* in presenza della *GCSF* esibiscono inclusione preferenziale, sebbene con in-

tensità variabili. In particolare, i meccanismi di stabilizzazione plausibili sono quelli di lioprotezione (si veda il paragrafo §1.3.2) e di prevenzione dell'aggregazione mediante la formazione di inclusioni, in modo particolare nei confronti dei residui idrofobici. Quest'ultima affermazione è sostenuta dalle osservazioni su β-*parameter*, distanze minime, inclusioni e frazioni di superficie idrofobica. Nel caso della β-*CD* si può anche notare l'inclusione della cisteina, la quale può reagire con altri residui della stessa specie per formare ponti disolfuro, causando aggregazione e quindi denaturazione [60].

Infine, si completa lo studio dei sistemi *CD-GCSF* con l'analisi dei *cluster*. Si procede inizialmente con la distribuzione delle ciclodestrine tra i vari aggregati, in funzione della dimensione di questi ultimi. I dati sono rappresentati in Figura 4.20.

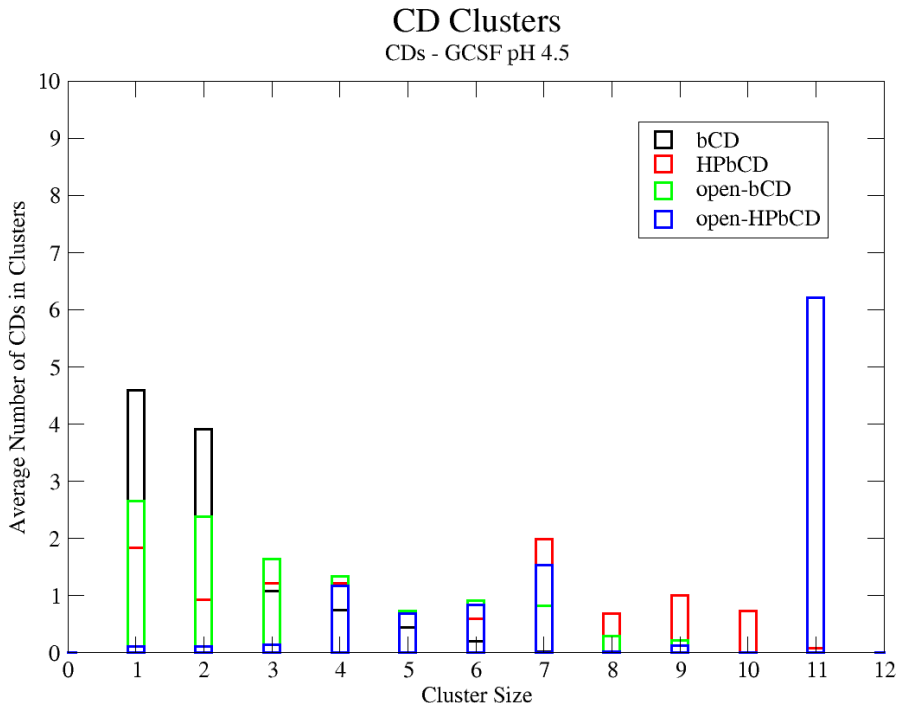

Figura 4.20: Distribuzione della dimensione dei *cluster* per le CD nei sistemi *CD-GCSF*.

Come visto in precedenza, si osserva una netta distinzione tra il comportamento esibito dalle due β-*CD* e dalle *HP*-β-*CD*. Le prime infatti popolano preferibilmente cluster di dimensioni minori mentre, al contrario, le seconde si dispongono in *cluster* più massivi. A supporto di tali osservazioni si propongono i dati relativi a dimensioni e numero medio degli aggregati, riportati rispettivamente in Tabella 4.15 e 4.16.
Si possono trarre alcune conclusioni riguardo alla formazione di *cluster* per i sistemi *CD-GCSF*:

- il comportamento esibito dalle due tipologie di ciclodestrine (forme native e

Tabella 4.15: N° medio di CD in un *cluster* per i sistemi *CD-LDH*.

HP-β-CD	*β-CD*	open-*HP-β-CD*	open-*β-CD*
5.39	2.54	8.47	3.71

Tabella 4.16: N° medio di *cluster* nel tempo per i sistemi *CD-LDH*.

HP-β-CD	*β-CD*	open-*HP-β-CD*	open-*β-CD*
3.82	7.22	1.60	5.20

funzionalizzate) è in accordo con tutti i sistemi analizzati precedentemente, pertanto le *HP-β-CD* formano aggregati di dimensioni maggiori in quantità minori, mentre le *β-CD* manifestano proprietà opposte, formando numerosi *cluster* necessariamente meno massivi;

- le *CD* a catena aperta formano *cluster* di dimensioni superiori, quindi in quantità inferiori, mentre gli eccipienti a catena chiusa si comportano in maniera opposta. Tale risultato è coerente con tutte le altre tipologie di dati esaminate, poiché le ciclodestrine a catena chiusa sono capaci di formare inclusioni con i residui peptidici. Pertanto, interagendo maggiormente con la proteina, il numero di molecole a disposizione per la formazione di aggregati è necessariamente minore.

4.4 Conclusioni

In questo capitolo si sono studiate le proprietà delle ciclodestrine nei confronti delle proteine, in particolare facendo riferimento a due casi studio, la *lattato deidrogenasi* (*LDH*) ed il *fattore di stimolazione delle colonie di granulociti* (*GCSF*).

Lo scopo di questo studio è quello di testare i metodi computazionali della Dinamica Molecolare come potenziali strumenti deputati alla scelta degli eccipienti da destinare alla formulazione di farmaci a base proteica. Nel perseguire tale obiettivo si è cercato di fare riferimento a studi analoghi, tuttavia l'Autore non è a conoscenza di altri esempi, tra quelli reperibili in letteratura, oltre a quelli finora indicati. Per questa ragione è stato necessario riferirsi a dati di origine sperimentale, laddove disponibili, a sostegno dei risultati ottenuti. Tuttavia, i mezzi e le metodologie impiegati negli esperimenti difficilmente consentono di discriminare i possibili meccanismi di stabilizzazione, pertanto le simulazioni molecolari rappresentano l'unica alternativa praticabile per meglio comprendere tali aspetti. In altri termini, è necessario intensificare l'utilizzo di tecniche computazionali per avere a disposizione più dati, in modo da agevolare la convalida di modelli e procedure adottati.

Lo studio dei sistemi *CD*-proteine si colloca successivamente a quello delle coppie *CD-AA*, pertanto le analisi non sono state finalizzate semplicemente all'investigazione delle proprietà degli eccipienti, bensì anche a correlare i comportamenti dei

residui peptidici a quelli dei singoli amminoacidi. Inizialmente si è valutata la tendenza delle ciclodestrine ad esibire esclusione preferenziale, tuttavia con scarsi risultati, in quanto per entrambi i polipeptidi si verifica inclusione. In particolare, la *β-CD* a catena chiusa è l'eccipiente più incluso. In base a questi risultati è difficile ipotizzare il successo, da parte delle ciclodestrine, nell'attuare l'esclusione preferenziale per stabilizzare i principi attivi in esame. L'unico scenario più plausibile è costituito dal sistema *HP-β-CD-LDH*, per il quale l'inclusione è piuttosto ridotta.

A supporto delle informazioni ottenute dai profili del *β-parameter* sono state valutate le distanze minime delle ciclodestrine dai residui peptidici e la formazione di inclusioni per le ciclodestrine a catena chiusa, riscontrando corrispondenza tra le varie tipologie di dati ottenuti. Le inclusioni hanno coinvolto principalmente gli amminoacidi apolari e polari, sebbene in alcuni casi anche quelli carichi abbiano contribuito. Questo risultato, apparentemente sorprendente, sembrerebbe contrastare con le osservazioni discusse nel paragrafo §3.2, tuttavia testimonia la complessità delle interazioni tra le *CD* ed i peptidi che fanno parte della proteina. Infatti, mentre per i sistemi ciclodestrina-amminoacidi tutte le specie chimiche sono libere di muoversi ed interagire liberamente nel box, per i sistemi proteici ciò non è possibile, pertanto i residui inclusi non sono necessariamente i più affini, bensì quelli più accessibili dal punto di vista sterico. Perciò, si evince che questo comportamento in realtà non sia in contrasto con i risultati precedenti, anzi conferma che l'affinità prevalente sia nei confronti del *backbone*, consentendo anche agli amminoacidi meno affini di interagire significativamente.

Successivamente si sono considerati gli aspetti strutturali delle proteine, ovvero le variazioni del raggio di giro e della frazione di superficie idrofobica rispetto alla totale, esposte al solvente, allo scopo di rilevare potenziali segni di denaturazione. Le escursioni di tali valori rispetto alla media temporale sono apprezzabili, sebbene non sufficientemente elevate da poter suggerire l'insorgenza di *unfolding*.

Infine, la formazione di *cluster* rappresenta un aspetto di secondaria importanza, che tuttavia rafforza le affermazioni presentate sinora. In particolare, le ciclodestrine a catena chiusa formano un numero di aggregati mediamente inferiori alle corrispettive specie a catena aperta, poiché interagiscono maggiormente con la proteina, pertanto esibiscono una minore tendenza ad aggregare.

Si può quindi concludere che, per quanto riguarda le ciclodestrine a catena chiusa, vi siano dati che giustifichino l'azione stabilizzante di questi eccipienti, in accordo con alcuni studi sperimentali. Il meccanismo di stabilizzazione che si può ipotizzare, basandosi sui risultati ottenuti, è uno oppure una combinazione tra quelli proposti nel paragrafo §1.3.2 [61] [62]. Alternativamente, l'inclusione di residui peptidici coinvolti in fenomeni di svolgimento parziale ed aggregazione, impedisce loro di interagire e, conseguentemente, di causare denaturazione [59]. L'Autore non è a conoscenza di studi riguardanti le ciclodestrine a catena aperta, pertanto non è possibile fare riferimento a meccanismi specifici. Si può ragionevolmente supporre che il comportamento delle open-*CD* si discosti poco da quello delle analoghe specie a catena chiusa, poiché i risultati sono qualitativamente si-

mili, ad eccezione delle inclusioni che non sono fisicamente realizzabili da tali molecole.

Da queste valutazioni si intuisce la necessità di ulteriori studi riguardo alle proprietà di questi sistemi, focalizzandosi ad esempio sui fenomeni di aggregazione, simulando almeno due proteine nello stesso box. Infine, anche l'influenza delle interfacce è un aspetto estremamente importante che può influire sulla stabilità proteica. Sarebbe pertanto opportuno tentare l'utilizzo della Dinamica Molecolare come strumento per l'analisi di tali aspetti, focalizzandosi sugli aspetti meccanicistici di questi fenomeni che vengono tipicamente investigati sperimentalmente.

Capitolo 5

Considerazioni Finali

Questa Tesi ha trattato alcune delle principali problematiche riscontrate nell'ambito della produzione di farmaci a base proteica, specificamente relative all'impiego della liofilizzazione come processo di essiccamento. Infatti, in questo contesto, sono numerosi i fattori che possono influire sulla stabilità del biofarmaco, rischiando di comprometterne la qualità. Alcuni dei fenomeni alla base della denaturazione delle proteine sono:

- l'adsorbimento ed *unfolding* delle stesse in prossimità di interfacce e superfici;
- l'aumento di concentrazione dei soluti a causa delle fasi di congelamento ed essiccamento, noto come crio-concentrazione;
- l'aggregazione;
- l'insorgenza di reazioni chimiche degradanti;
- gli stress dovuti alla disidratazione.

Inoltre, oscillazioni di temperatura e pH al di fuori degli intervalli di stabilità possono rivelarsi altrettanto deleteri nei confronti dell'integrità del farmaco. Tali fenomeni possono pertanto indurre meccanismi di destabilizzazione, potenzialmente capaci di causare l'inattivazione della proteina. Come è noto, la perdita della struttura proteica nativa, qualsivoglia il meccanismo di degradazione, può comportare la perdita dell'attività biologica [2] che costituisce la funzione del farmaco. Ciò può causare risposte immunogeniche [63] dovute, per esempio, ad aggregati proteici, con conseguenze potenzialmente gravi sulla salute del paziente. Risulta evidente la necessità di far fronte a tali problematiche, perciò nell'ambito della formulazione dei biofarmaci si aggiungono degli eccipienti. Essi sono sostanze chimiche dotate di proprietà adeguate a limitare o possibilmente evitare l'insorgenza dei meccanismi di degradazione. Vi sono varie tipologie di additivi stabilizzanti, tra cui saccaridi (monosaccaridi, disaccaridi, oligosaccaridi), amminoacidi, polimeri, surfattanti e talvolta proteine [2]. Per questa Tesi si è scelto di fare riferimento

ad una promettente classe di oligosaccaridi, ovvero le ciclodestrine. In particolare si sono impiegate le β-ciclodestrine, in forma nativa e funzionalizzata (gruppi idrossipropilici), sia a catena aperta sia a catena chiusa. Considerata la varietà dei meccanismi di degradazione si è scelto di focalizzare l'attenzione su due aspetti cruciali: la denaturazione indotta da interfacce e l'aggregazione delle proteine.

Per investigare l'adeguatezza delle ciclodestrine come eccipienti, si è utilizzato un approccio alternativo a quello sperimentale più tradizionale, ovvero la Dinamica Molecolare. Essa è una tecnica computazionale che consiste nella simulazione di molecole. Una volta in possesso dei modelli delle specie chimiche da esaminare, nonché delle interazioni intra ed intermolecolari, si possono simulare i sistemi di interesse. Ciò costituisce un grande vantaggio poiché, generalmente, gli approcci sperimentali non riescono a fornire indicazioni precise sulla tipologia di meccanismi di stabilizzazione esercitati dagli eccipienti, i quali hanno origini molecolari. Tra essi si ricordano due tipologie, quelli termodinamici, come l'esclusione preferenziale tipica della crioprotezione, e quelli cinetici, tra cui *water replacement*, *water entrapment* e vitrificazione, caratteristici della lioprotezione [2] [9] [10].

Lo scopo di questa Tesi è pertanto stato quello di sfruttare la Dinamica Molecolare per studiare le proprietà delle ciclodestrine in merito alla denaturazione interfacciale e all'aggregazione. Per fare ciò si sono esaminate alcune tipologie di sistemi e sostanze, ovvero interfaccia acqua-aria, interfaccia acqua-ghiaccio, amminoacidi e proteine.

Procedendo per ordine, le uniche ciclodestrine ad esibire proprietà nei confronti dell'interfaccia acqua-aria sono state le *HP-β-CD*. Tale risultato è in accordo con quanto osservato sperimentalmente [14]. In particolare, la forma a catena chiusa si orienta preferenzialmente con il *secondary rim* verso la fase gassosa. Diversamente, per l'interfaccia acqua-ghiaccio non sono state riscontrate interazioni tra eccipienti e superficie. Tuttavia, ciò non implica necessariamente che le *CD* non siano adatte come stabilizzanti nei confronti dell'azione del ghiaccio. Successivamente, si è passati allo studio delle proprietà dei sistemi *CD-AA*, come passaggio preliminare all'esame delle proteine, solo per le ciclodestrine a catena chiusa. Sono state valutate principalmente due tipologie di dati, ovvero i massimi dei profili delle *rdf* e le inclusioni formate tra le due specie chimiche. Inoltre, si è cercato di capire se esistessero orientamenti preferenziali da parte di *CD* ed *AA*, ed eventualmente quali. I risultati ottenuti variano a seconda della categoria di amminoacido e dipendono dalla struttura molecolare. In particolare, è stato stabilito che le interazioni avvengano tra il *secondary rim* ed il *backbone*. In generale:

- gli amminoacidi carichi positivamente e negativamente interagiscono piuttosto debolmente con gli eccipienti, prevalentemente con il *backbone*;

- gli amminoacidi polari sono chiaramente più affini alla *HP-β-CD*, tuttavia nel caso della cisteina e della tirosina le interazioni sono decisamente significative per entrambe le *CD*. In particolare, per la *β-CD* prevale l'affinità verso la *sidechain* di questi residui;

- gli amminoacidi apolari interagiscono intensamente con le ciclodestrine, eccetto in qualche caso per la *CD* nativa. Le interazioni prediligono nuovamente il *backbone* sebbene, vista la grande affinità degli eccipienti in esame nei confronti di questa classe di *AA*, per alcune coppie prevalga la catena laterale. Per i dovuti dettagli si rimanda al paragrafo §3.2;

- la *HP-β-CD* interagisce più agevolmente con tutte le categorie di amminoacidi, sebbene con intensità talvolta inferiore a quella esibita dalla *β-CD*;

- le inclusioni tra amminoacidi e ciclodestrine rispecchiano le affinità riscontrate mediante le *rdf*.

Una volta studiate le proprietà in presenza degli amminoacidi, si è cercato di riscontrare la validità dei risultati ottenuti esaminando le caratteristiche delle proteine. Si è fatto riferimento a due casi di studio, la *lattato deidrogenasi* (*LDH*) ed il *fattore di stimolazione delle colonie di granulociti* (*GCSF*). In particolare, la prima proteina è ben nota in letteratura, pertanto si è scelta come riferimento per la convalida delle procedure analitiche, nonché dei risultati ottenuti, facendo riferimento, ove possibile, a studi computazionali [7].

Per valutare l'adeguatezza delle ciclodestrine come eccipienti si sono esaminate le caratteristiche delle interazioni con i residui peptidici. Inizialmente, si è adottato uno strumento, il *β-parameter*, per valutare e confrontare il grado di esclusione preferenziale per le varie *CD*. Tutte le ciclodestrine risultano preferenzialmente incluse, sebbene con intensità diverse. Infatti, per l'*LDH* l'inclusione è meno spiccata, mentre per i sistemi con *GCSF* l'interazione è significativamente più forte. In particolare, la *β-CD* a catena chiusa è fortemente inclusa in entrambi i casi. Pertanto, l'esclusione preferenziale non risulta essere un meccanismo di stabilizzazione plausibile per questi sistemi. L'unica eccezione è costituita dal sistema *HP-β-CD-LDH*, nel quale la cicldoestrina è inclusa molto debolmente, pertanto potrebbe risultare un candidato efficace. Successivamente, si è cercato di giustificare tali proprietà studiando in maniera più mirata le interazioni tra *CD* e residui peptidici, confrontando i risultati con quelli ottenuti per gli amminoacidi. Si è quindi osservato che le inclusioni non avvengano necessariamente con gli *AA* più affini poiché, facendo parte di una proteina, non sempre risultano accessibili al solvente e, di conseguenza, neanche agli eccipienti. Inoltre, ciò conferma il ruolo del *backbone* nelle interazioni, infatti alcune delle inclusioni più significative sono avvenute con amminoacidi che, se presi individualmente, interagiscono debolmente con le ciclodestrine. Infine, si è valutata la presenza di eventuali segni di denaturazione, analizzando l'evoluzione del raggio di giro e della frazione di superficie idrofobica delle proteine, tuttavia senza riscontrare anomalie.

Un ultimo aspetto preso in considerazione è la formazione di *cluster*, che sostanzialmente rimane inalterato a prescindere dalla tipologia di sistemi studiati. Infatti, il comportamento qualitativo ha generalmente confermato la formazione di aggregati di dimensioni maggiori, quindi con numerosità inferiore, per le *HP-β-CD*, e l'opposto per le *β-CD*, propense alla formazione di *cluster* meno massivi

ma in quantità superiori.

Dopo aver effettuato tali valutazioni, si è potuto concludere che le ciclodestrine possano essere efficaci stabilizzanti nell'ambito della lioprotezione. Inoltre, si ipotizza ragionevolmente che le *CD* a catena chiusa possano includere residui peptidici, soprattutto apolari, potenzialmente coinvolti in fenomeni di aggregazione, impedendone lo sviluppo [59]. Infine, per le *HP-β-CD* si conferma la possibilità di utilizzo come surfattanti efficaci nei confronti della denaturazione indotta dall'interfaccia acqua-aria, come è già stato osservato sperimentalmente per la molecola a catena chiusa [14].

L'approccio utilizzato si è quindi dimostrato utile per investigare gli aspetti molecolari di fenomeni ben noti nell'ambito della stabilità proteica e della liofilizzazione, chiarendo aspetti che difficilmente possono essere investigati sperimentalmente. Tuttavia, la scarsità di riferimenti bibliografici a sostegno delle procedure utilizzate e, conseguentemente dei risultati così ottenuti, attualmente non consente di sostituire gli esperimenti con simulazioni molecolari. Si auspica che la diffusione di mezzi di calcolo sempre più potenti ed accessibili favorisca l'adozione dei metodi computazionali come strumenti di supporto per lo studio della stabilità proteica e la progettazione di formulazioni efficienti. Parallelamente, è necessario intensificare l'attività di validazione dei modelli impiegati con prove sperimentali, allo scopo di incrementare progressivamente l'accuratezza dei risultati ottenuti. Inoltre, una strategia di rilevante importanza e recente impiego per la realizzazione di tale obiettivo, è l'utilizzo del Machine Learning, per agevolare e migliorare lo sviluppo di *force-field* [64] [65].

Per quanto riguarda lo studio delle proprietà stabilizzanti delle ciclodestrine, con gli annessi meccanismi, vi sono alcuni aspetti che meritano maggiore attenzione, pertanto potrebbero essere oggetto di ulteriori studi. Si propongono in seguito alcune idee:

- simulazione dei sistemi open-*CD-AA*, per verificare l'effetto della struttura sulle interazioni con gli amminoacidi, allo scopo di chiarire meglio il comportamento in presenza di proteine, analogamente a quanto svolto per le molecole a catena chiusa;
- per comprendere meglio le modalità di aggregazione si potrebbero simulare box con due o più molecole di una stessa proteina, sia in presenza sia in assenza di ciclodestrine, in modo da poterne valutare l'effetto;
- simulazioni dei sistemi proteici alle interfacce acqua-aria ed acqua-ghiaccio, con e senza eccipienti;
- combinare gli approcci dei due punti precedenti, per verificare l'effetto delle interfacce sulla tendenza ad aggregare.

Lo studio di questi sistemi chiarirebbe alcuni aspetti cruciali, tuttavia vi è un ulteriore problema di cui occuparsi, ovvero quello della scala temporale. Il tempo

simulato nelle simulazioni di Dinamica Molecolare tradizionale è generalmente limitato alla scala dei microsecondi, soprattutto a causa dell'elevato costo computazionale. Ciò ha una conseguenza cruciale, ovvero le simulazioni di MD potrebbero non dare segni di denaturazione a causa della breve estensione temporale. Essa non consentirebbe di osservare configurazioni del sistema a maggiore contenuto energetico. Infatti, è noto dalla Meccanica Statistica che ad una qualsiasi configurazione sia associata una probabilità, la quale dipende strettamente dall'energia corrispondente [19]. Di conseguenza, essendo le simulazioni in regime dinamico, alcuni fenomeni potrebbero richiedere tempi significativamente maggiori rispetto a quelli simulati, per poter essere osservati. Fortunatamente esiste un rimedio a questa problematica, che consiste nell'impiego di una tecnica computazionale più avanzata, la Metadinamica (in seguito indicata come MetaD [66]. Questo potente strumento consente di migliorare drasticamente il cosiddetto *sampling* o campionamento di dati [67] [68], riducendo l'entità delle barriere energetiche [69], permettendo al sistema di evolvere più agevolmente [70], esplorando più stati. In questo modo i dati ottenuti dalla traiettoria presentano un quadro più completo delle caratteristiche del sistema, aumentando la possibilità di osservare eventuali meccanismi di denaturazione. Infatti essi sono spesso legati all'*unfolding*, fenomeno che può richiedere tempi anche significativamente più lunghi rispetto ai 100 ns simulati [71]. L'adozione di questo approccio per studiare la stabilità proteica in presenza di interfacce e surfattanti è già stata proposta in letteratura, con buoni risultati [4].

I dati ottenuti da queste simulazioni migliorerebbero significativamente la comprensione dei meccanismi molecolari responsabili per la denaturazione e la destabilizzazione. Ciò consentirebbe una scelta degli eccipienti più razionale ed efficiente, riducendo potenzialmente i costi di ricerca e sviluppo, tradizionalmente basati su campagne sperimentali. Per concludere, un ulteriore possibile sviluppo di questa strategia consiste nell'impiego di ciclodestrine funzionalizzate *ad hoc*, in base alle esigenze specifiche di ogni farmaco. Vi sono infatti numerose *CD* derivate impiegate in formulazioni, alcune delle quali attualmente impiegate in prodotti farmaceutici in commercio, tra cui la ormai nota *HP-β-CD* e la sulfobutileter-β-ciclodestrina (*SBE-β-CD*) [6]. Risulta quindi evidente che l'utilizzo della Dinamica Molecolare, ed eventualmente della Metadinamica, siano di grande aiuto per studiare rapidamente le proprietà degli eccipienti, in modo da restringere il numero dei plausibili candidati e concentrare gli indispensabili test sperimentali esclusivamente su formulazioni di concreto interesse.

L'Autore si augura di aver mostrato i vantaggi nell'impiego delle tecniche computazionali trattate, senza tuttavia tralasciarne le limitazioni o i difetti, auspicando che questo studio possa servire da traccia per ulteriori sviluppi, basati sulla coniugazione di simulazioni molecolari e della modellazione con dati di natura sperimentale.

Ringraziamenti / Acknowledgements

Approfitto di questo spazio per ringraziare tutti coloro che hanno contribuito alla mia crescita personale e alla mia istruzione, durante il mio percorso universitario ed il lavoro di Tesi.
Innanzitutto desidero ringraziare Maria Grazia e Roberto, i miei genitori, per avermi dato la possibilità di istruirmi senza dovermi preoccupare del costo dell'Università e delle spese annesse. Li ringrazio inoltre per il loro supporto, indispensabile in questi anni. Ringrazio allo stesso modo anche i miei parenti, per aver creduto in me ed avermi incoraggiato a dare il meglio in questo percorso.
Ringrazio i miei amici più cari, di cui non farò i nomi (se state leggendo sapete chi siete), per avermi sostenuto, per avermi trasmesso la loro fiducia nei miei confronti, per i loro sinceri auguri per la buona riuscita dei miei progetti personali, delle mie sfide e soprattutto per la mia realizzazione personale. A tutto Voi dico grazie e, nel ricambiare gli stessi auguri, spero di avervi resi fieri di me in questi anni, nonostante gli impegni e, francamente, la mia presenza a volte un po' troppo sporadica.
Ringrazio tutte le persone che hanno contribuito, in qualsiasi maniera, alla mia istruzione, per avermi sempre incoraggiato a studiare, imparare e migliorare, poiché mi hanno permesso di raggiungere questo traguardo e di realizzare gli obiettivi e, perché no, i sogni che mi attendono. Sono profondamente grato a tutti coloro che quindi hanno avuto un ruolo nella mia crescita intellettuale. In particolare, un sentito ringraziamento va agli insegnanti del Liceo Amaldi di Novi Ligure, che mi hanno fornito tutte le basi necessarie per affrontare al meglio il percorso universitario. Allo stesso modo, ringrazio i docenti del Politecnico di Torino che hanno saputo ispirarmi e motivarmi allo studio, trasmettermi la loro passione e valorizzare le mie qualità.
Un sentito ringraziamento va al Prof. Roberto Pisano ed al Dott. Andrea Arsiccio, i quali hanno supervisionato questa Tesi, dandomi la possibilità di affrontare un tema verso cui ho sviluppato un grande interesse. Li ringrazio per avere creduto nelle mie capacità ed avermi assegnato questo progetto, permettendomi di avvicinarmi ed approcciarmi al mondo della ricerca.
Ringrazio i colleghi che mi hanno accompagnato in questo percorso, poiché hanno

contribuito alla mia crescita personale, ognuno a proprio modo.
Ringrazio chiunque mi abbia ispirato a perseguire i miei obiettivi, senza demordere nei momenti di difficoltà.
Ringrazio infine i miei nonni, per avermi ispirato a conseguire gli studi ingegneristici, spero di avervi resi fieri di me e di continuare a farlo in futuro.

Risorse di calcolo fornite da HPC@POLITO, progetto di Academic Computing del Dipartimento di Automatica e Informatica presso il Politecnico di Torino (http://www.hpc.polito.it).

The Author wishes to acknowledge and thank the developers of the following software, for providing valuable tools and resources, necessary to the development of this project: GROMACS, PyMOL, VMD, GenIce, PLUMED, Grace, Gnuplot, LibreOffice, GIMP, VLC, MiKTeX Console, TeXworks and Ubuntu 18.04 LTS.

Computational resources were provided by HPC@POLITO, a project of Academic Computing within the Department of Control and Computer Engineering at the Politecnico di Torino (http://hpc.polito.it).

Bibliografia

[1] F. Franks, T. Auffret, Freeze-drying of Pharmaceuticals and Biopharmaceuticals, The Royal Society of Chemistry, 2007.

[2] W. Wang, Lyophilization and development of solid protein pharmaceuticals, International Journal of Pharmaceutics 203 (2000) 1–60.

[3] T. Arakawa, S. Prestrelski, W. Kenney, J. Carpenter, Factors affecting short-term and long-term stabilities of proteins, Advanced Drug Delivery Reviews 46 (2001) 307–326.

[4] A. Arsiccio, J. McCarty, R. Pisano, J. E. Shea, Effect of Surfactants on Surface-Induced Denaturation of Proteins: Evidence of an Orientation-Dependent Mechanism, The Journal of Physical Chemistry B 122 (2018) 11390–11399.

[5] M. Kristjànsson, J. Kinsella, Protein and Enzyme Stability: Structural, Thermodynamic, and Experimental Aspects, Vol. 35 of Advances in Food and Nutrition Research, Academic Press, 1991.

[6] T. Serno, R. Geidobler, G. Winter, Protein stabilization by cyclodextrins in the liquid and dried state, Advanced Drug Delivery Reviews 63 (2011) 1086–1106.

[7] A. Arsiccio, A. Paladini, F. Pattarino, R. Pisano, Designing the Optimal Formulation for Biopharmaceuticals: A New Approach Combining Molecular Dynamics and Experiments, Journal of Pharmaceutical Sciences 108 (2019) 431–438.

[8] A. Arsiccio, R. Pisano, Design of the formulation for therapeutic proteins - How to improve stability of drugs during freezing and in the dried state, Chimica Oggi/Chemistry Today 36 (2018) 16–18.

[9] A. Arsiccio, R. Pisano, Water entrapment and structure ordering as protection mechanisms for protein structural preservation, The Journal of Chemical Physics 148 (2018) 055102.

[10] A. Arsiccio, R. Pisano, Stability of Proteins in Carbohydrates and Other Additives during Freezing: The Human Growth Hormone as a Case Study, The Journal of Physical Chemistry B 121 (2017) 8652–8660.

[11] A. W. Coleman, I. Nicolis, N. Keller, J. P. Dalbiez, Aggregation of cyclodextrins: An explanation of the abnormal solubility of β-cyclodextrin, Journal of Inclusion Phenomena and Molecular Recognition in Chemistry 13 (1992) 139–143.

[12] J. Horský, J. Pitha, Inclusion complexes of proteins: Interaction of cyclodextrins with peptides containing aromatic amino acids studied by competitive spectrophotometry, Journal of Inclusion Phenomena and Molecular Recognition in Chemistry 18 (1994) 291–300.

[13] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, Journal of Molecular Graphics 14 (1996) 33–38.

[14] T. Serno, J. F. Carpenter, T. W. Randolph, G. Winter, Inhibition of Agitation-Induced Aggregation of an IgG-Antibody by Hydroxypropyl-β-Cyclodextrin, Journal of Pharmaceutical Sciences 99 (2010) 1193–1206.

[15] W. Khuntawee, M. Karttunen, J. Wong-ekkabut, A molecular dynamics study of conformations of beta-cyclodextrin and its eight derivatives in four different solvents, Physical Chemistry Chemical Physics 19 (2017) 24219–24229.

[16] S. Gould, R. Scott, 2-Hydroxypropyl-β-Cyclodextrin (HP-β-CD): A toxicology review, Food and Chemical Toxicology 43 (2005) 1451–1459.

[17] N. Attig, Computational Soft Matter: from Synthetic Polymers to Proteins: Lecture notes, NIC series, NIC, 2004.

[18] M. J. Abraham, T. Murtola, R. Schulz, S. Pall, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers, SoftwareX 1-2 (2015) 19–25.

[19] I. Levine, Physical Chemistry, McGraw-Hill, 2008.

[20] S. Páll, B. Hess, A flexible algorithm for calculating pair interactions on SIMD architectures, Computer Physics Communications 184 (2013) 2641–2650.

[21] J. Gebhardt, C. Kleist, S. Jakobtorweihen, N. Hansen, Validation and Comparison of Force Fields for Native Cyclodextrins in Aqueous Solution, The Journal of Physical Chemistry B 122 (2018) 1608–1626.

[22] P. C. Tchoreloff, M. M. Boissonnade, A. W. Coleman, A. Baszkin, Amphiphilic Monolayers of Insoluble Cyclodextrins at the Water/Air Interface. Surface Pressure and Surface Potential Studies, Langmuir 11 (1995) 191–196.

[23] E. Mixcoha, R. Rosende, R. Garcia-Fandino, A. Piñeiro, Cyclo-lib: a database of computational molecular dynamics simulations of cyclodextrins, Bioinformatics 32 (2016) 3371–3373.

[24] W. Snor, E. Liedl, P. Weiss-Greiler, A. Karpfen, H. Viernstein, On the structure of anhydrous β-cyclodextrin, Chemical Physics Letters 441 (2007) 159–162.

[25] M. Saunders, G. Voth, Coarse-graining methods for computational biology, Annual Review of Biophysics 42 (2013) 73–93.

[26] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, CHARMM: The biomolecular simulation program, Journal of Computational Chemistry 30 (2009) 1545–1614.

[27] O. Guvench, E. Hatcher, R. M. Venable, R. W. Pastor, A. D. MacKerell, CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses, Journal of Chemical Theory and Computation 5 (2009) 2353–2370.

[28] P. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess, E. Lindahl, Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models, Journal of Chemical Theory and Computation 6 (2010) 459–466.

[29] O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. Jamison, A. D. MacKerell, CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate–Protein Modeling, Journal of Chemical Theory and Computation 7 (2011) 3162–3180.

[30] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, The missing term in effective pair potentials, The Journal of Physical Chemistry 91 (1987) 6269–6271.

[31] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, M. Klein, Comparison of Simple Potential Functions for Simulating Liquid Water, The Journal of Chemical Physics 79 (1983) 926–935.

[32] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh Ewald method, The Journal of Chemical Physics 103 (1995) 8577–8593.

[33] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath, The Journal of Chemical Physics 81 (1984) 3684–3690.

[34] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, The Journal of Chemical Physics 81 (1984) 511–519.

[35] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, Molecular Physics 52 (1984) 255–268.

[36] W. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Physical Review A 31 (1985) 1695–1697.

[37] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, Journal of Applied Physics 52 (1981) 7182–7190.

[38] Schrödinger, LLC, The PyMOL Molecular Graphics System, Version 1.8 (November 2015).

[39] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, The Journal of Chemical Physics 126 (2007) 014101.

[40] A. Arsiccio, R. Pisano, Surfactants as stabilizers for biopharmaceuticals: An insight into the molecular mechanisms for inhibition of protein aggregation, European Journal of Pharmaceutics and Biopharmaceutics 128 (2018) 98–106.

[41] J. Schwegman, J. Carpenter, S. Nail, Evidence of partial unfolding of proteins at the ice/freeze-concentrate interface by infrared microscopy, Journal of Pharmaceutical Sciences 98 (2009) 3239–3246.

[42] G. Strambini, E. Gabellieri, Proteins in frozen solutions: Evidence of ice-induced partial unfolding, Biophysical Journal 70 (1996) 971–976.

[43] M. Matsumoto, T. Yagasaki, H. Tanaka, GenIce: Hydrogen-Disordered Ice Generator, Journal of Computational Chemistry 39 (1) 61–64.

[44] F. Aachmann, K. Larsen, R. Wimmer, Interactions of cyclodextrins with aromatic amino acids: A basis for protein interactions, Journal of Inclusion Phenomena and Macrocyclic Chemistry 73 (2012) 349–357.

[45] L. Fiume, M. Manerba, M. Vettraino, G. Di Stefano, Inhibition of lactate dehydrogenase activity as an approach to cancer therapy, Future Medicinal Chemistry 6 (2014) 429–445.

[46] C. Valvona, H. Fillmore, P. Nunn, G. Pilkington, The Regulation and Function of Lactate Dehydrogenase A: Therapeutic Potential in Brain Tumor, Brain Pathology 26 (2016) 3–17.

[47] A. Livesey, F. Garty, A. R. Shipman, K. E. Shipman, Lactate dehydrogenase in dermatology practice, Clinical and Experimental Dermatology 1 (2019) 1–5.

[48] J.-W. Lee, C. Trinh, Microbial biosynthesis of lactate esters, Biotechnology for Biofuels 12 (2019) 1–20.

[49] A. Herman, T. Boone, H. Lu, Characterization, formulation, and stability of Neupogen (Filgrastim), a recombinant human granulocyte-colony stimulating factor., Pharmaceutical biotechnology 9 (1996) 303–328.

[50] D. D'Amario, A. Leone, J. Borovac, F. Cannata, A. Siracusano, G. Niccoli, F. Crea, Granulocyte colony-stimulating factor for the treatment of cardiovascular diseases: An update with a critical appraisal, Pharmacological Research 127 (2018) 67–76.

[51] Y. Wang, L. Chen, F. Liu, N. Zhao, L. Xu, B. Fu, Y. Li, Efficacy and tolerability of granulocyte colony-stimulating factors in cancer patients after chemotherapy: A systematic review and Bayesian network meta-analysis, Scientific Reports 9 (2019) 1–12.

[52] S.-T. Tsai, S.-C. Chu, S.-H. Liu, C.-Y. Pang, T.-W. Hou, S.-Z. Lin, S.-Y. Chen, Neuroprotection of granulocyte colony-stimulating factor for early stage Parkinson's disease, Cell Transplantation 26 (2017) 409–416.

[53] C.-Y. Pang, K.-L. Yang, C.-H. Fu, L.-Y. Sun, S.-Y. Chen, C.-H. Liao, G-CSF enhances the therapeutic potency of stem cells transplantation in spinal cord-injured rats, Regenerative Medicine 14 (2019) 571–583.

[54] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank, Nucleic Acids Research 28 (2000) 235–242.

[55] B. A. Beaupre, J. V. Roman, M. R. Hoag, K. M. Meneely, N. R. Silvaggi, A. L. Lamb, G. R. Moran, Ligand binding phenomena that pertain to the metabolic function of renalase, Archives of Biochemistry and Biophysics 612 (2016) 46–56.

[56] M. Aritomi, N. Kunishima, T. Okamoto, R. Kuroki, Y. Ota, K. Morikawa, Atomic structure of the GCSF-receptor complex showing a new cytokine-receptor recognition scheme, Nature 401 (1999) 713–717.

[57] F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander, M. Scharf, The Double Cubic Lattice Method: Efficient Approaches to Numerical Integration of Surface

Area and Volume and to Dot Surface Contouring of Molecular Assemblies, Journal of Computational Chemistry 16 (1995) 273–284.

[58] J. Iwai, N. Ogawa, H. Nagase, T. Endo, T. Loftsson, H. Ueda, Effects of various cyclodextrins on the stability of freeze-dried lactate dehydrogenase, Journal of Pharmaceutical Sciences 96 (2007) 3140–3143.

[59] F. Aachmann, D. Otzen, K. Larsen, R. Wimmer, Structural background of cyclodextrin-protein interactions, Protein Engineering 16 (2003) 905–912.

[60] H. J. Lee, A. McAuley, K. F. Schilke, J. McGuire, Molecular origins of surfactant-mediated stabilization of protein drugs, Advanced Drug Delivery Reviews 63 (2011) 1160–1171.

[61] K.-i. Izutsu, S. Yoshioka, T. Terao, Stabilizing effect of amphiphilic excipients on the freeze-thawing and freeze-drying of lactate dehydrogenase, Biotechnology and Bioengineering 43 (1994) 1102–1107.

[62] K.-i. Izutsu, S. Yoshioka, S. Kojima, Increased Stabilizing Effects of Amphiphilic Excipients on Freeze-Drying of Lactate Dehydrogenase (LDH) by Dispersion into Sugar Matrices, Pharmaceutical Research 12 (1995) 838–843.

[63] E. M. Moussa, J. P. Panchal, B. S. Moorthy, J. S. Blum, M. K. Joubert, L. O. Narhi, E. M. Topp, Immunogenicity of Therapeutic Protein Aggregates, Journal of Pharmaceutical Sciences 105 (2016) 417–430.

[64] J. L. McDonagh, A. Shkurti, D. J. Bray, R. L. Anderson, E. O. Pyzer-Knapp, Utilizing Machine Learning for Efficient Parameterization of Coarse Grained Molecular Force Fields, Journal of Chemical Information and Modeling 59 (2019) 4278–4288.

[65] J. Wang, S. Olsson, C. Wehmeyer, A. Pérez, N. E. Charron, G. de Fabritiis, F. Noé, C. Clementi, Machine Learning of Coarse-Grained Molecular Dynamics Force Fields, ACS Central Science 5 (2019) 755–767.

[66] A. Barducci, M. Bonomi, M. Parrinello, Metadynamics, WIREs Computational Molecular Science 1 (2011) 826–843.

[67] J. Pfaendtner, M. Bonomi, Efficient Sampling of High-Dimensional Free-Energy Landscapes with Parallel Bias Metadynamics, Journal of Chemical Theory and Computation 11 (2015) 5062–5067.

[68] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, M. Parrinello, Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics, The Journal of Physical Chemistry B 110 (2006) 3533–3539.

[69] A. Laio, M. Parrinello, Escaping free-energy minima, Proceedings of the National Academy of Sciences 99 (2002) 12562–12566.

[70] J. F. Dama, M. Parrinello, G. A. Voth, Well-Tempered Metadynamics Converges Asymptotically, Physical Review Letters 112 (2014) 240602.

[71] O. Valsson, P. Tiwary, M. Parrinello, Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint, Annual Review of Physical Chemistry 67 (2016) 159–184.