



Dipartimento di Matematica e Fisica

Corso di Laurea Magistrale in Fisica

Fisica dei Biosistemi

Transizioni di fase in un sistema di cristalli liquidi di DNA

Candidato:

Simone Milano

Relatore:

Prof. Cristiano De Michele

Anno Accademico 2020/2021

Indice

1	Introduzione	12
1.1	Panoramica sulla struttura del DNA	14
1.1.1	Approcci top-down e bottom-up nella costruzione di nanostrutture	14
1.1.2	L'acido desossiribonucleico e le sue proprietà	16
1.1.3	Il punto di contatto tra genetica e nanotecnologie	22
2	Panoramica sui cristalli liquidi	29
2.1	Cristalli liquidi di DNA	29
2.2	Fase isotropa	32
2.3	Fase Nematica	33
2.4	Fase Smettica	34
3	Teoria di Onsager	36
3.1	Passaggi specifici della Teoria di Onsager	47
3.1.1	Il procedimento per ricavare l'energia libera di Helmholtz per sistemi di particelle dalla simmetria cilindrica	47
3.1.2	Normalizzazione della gaussiana di Onsager	51
3.1.3	Entropia orientazionale e secondo coefficiente del viriale in fase nematica	52
4	Monte Carlo	55
4.1	Campionamento Casuale	57
4.2	Importance sampling	58

4.3	Metropolis	61
4.3.1	Ensemble canonico	65
4.3.2	Ensemble Isobaro-isotermo	66
5	Modellazione Coarse-grained	69
5.1	Struttura del modello Gapped-Duplex	70
6	Modelli teorici	73
6.1	Parametro d'ordine nematico	73
6.2	Parametro d'ordine smettico	76
6.3	Convenienza energetica dello stato folded	77
7	Risultati delle misure	80
7.1	Panoramica sul funzionamento delle misure e sulle grandezze analizzate	80
7.2	La ricerca della K perfetta: come la percentuale di folding ed i parametri d'ordine ci aiutano a determinare le fasi	83
8	Conclusioni	89
9	Parentesi sul funzionamento dei programmi informatici	91
	Bibliografia	96

Elenco delle figure

1.6	Rappresentazione tridimensionale e formula di struttura delle basi azotate pirimidiniche	18
1.7	Formula di struttura di un gruppo fosfato legato ad un generico gruppo organico R: questo residuo fosforico costituisce una parte integrante nella struttura dei nucleotidi. In questo caso, un generico nucleotide che interagirebbe col gruppo in figura prenderebbe la denominazione di <i>nucleotide monofosfato</i>	18
1.8	Formula di struttura del DNA: da notare l'alternanza tra zuccheri e gruppi fosfato.	18
1.9	Modello di Watson-Crick: qui possiamo vedere esplicitamente i legami A-T e C-G	20
1.10	Accoppiamento di base: come si può notare, quattro monomeri di nucleotidi formano due coppie di basi: la guanina si accoppia con la citosina attraverso tre legami idrogeno, mentre l'adenina si accoppia con la timina (o nel caso della qui presente immagine, con l'uracile) attraverso due singoli legami idrogeno.	21
1.11	<i>Profondità</i> : l'opera di Escher che ispirò la nascita della nanotecnologia strutturale del DNA. Come si può notare, la disposizione delle creature ricorda un reticolo cristallino di Bravais, tipico modello nella fisica dello stato solido	22
1.12	Questa immagine apparsa per la prima volta sul Nature del 16 marzo 2006 ha fatto il giro del mondo: si tratta di uno "smile" grande poco meno di 100 nm costruito attraverso la tecnica <i>DNA origami</i> . La sequenza genetica con cui hanno costruito questa faccina proviene da un virus batteriofago ed è lunga 7 kB, ovvero 7000 paia di basi (dove l'unità di misura della lunghezza è il <i>pB</i> , ovvero la distanza tipica tra un paio di basi azotate pari circa a 0,34 nm)	23

1.13 A sinistra, un esempio di "casella" di nano-DNA. Questo modello è stato usato per costruire un reticolo periodico bidimensionale, a destra si può notare la struttura visualizzata attraverso un microscopio a forza atomica.	24
1.14 Una rappresentazione del modello <i>DNA bricks</i> presa direttamente dall'articolo originale su cui la tecnica fu inventata. In una maniera piuttosto intuitiva, i singoli mattoncini di DNA vengono qui paragonati a dei pezzi di LEGO [18].	25
1.15 Una molecola (DX) a doppio crossover. Questa molecola è costituita di cinque singoli elementi di DNA i quali formano due domini doppio-elicoidali, a sinistra e a destra in questa immagine. Ci sono due punti crossover dove i componenti attraversano da un dominio all'altro. [20]	25
1.16 Origami di DNA visualizzato attraverso un microscopio a scansione di sonda. Notare le forme triangolari che si iterano periodicamente lungo la superficie	26
1.17 Strutture di DNA origami generate proceduralmente da un algoritmo informatico [26]. Dai tempi di Seeman, la branca delle nanotecnologie legate al DNA si è evoluta incessantemente e le potenzialità di questi studi offrono grandi speranze per un ulteriore legame tra fisica e biologia.	27
 2.1 Struttura a <i>doppia elica</i> del DNA.	30
2.2 Fase isotropa delle gapped-duplex DNA. Come si può notare facilmente da questa immagine catturata col software <i>MOLGL</i> , ogni direzione è equivalente ed il comportamento del sistema è molto simile a quello di un liquido.	32
2.3 Modello di struttura della fase smettica <i>A</i> . La "direttoria" n è orientata parallelamente alla normale k degli strati smettici, mentre all'interno di ogni singolo strato smettico i centri di massa delle molecole sono distribuiti isotropicamente.	33

2.4	Modello di struttura della fase nematico-chirale. La p si riferisce al passo chirale.	33
2.5	Modello di struttura della fase smettica A. La "direttoria" \mathbf{n} è orientata parallelamente alla normale \mathbf{k} degli strati smettici, mentre all'interno di ogni singolo strato smettico i centri di massa delle molecole sono distribuiti isotropicamente.	34
3.1	Rappresentazione di un cristallo liquido nematico. La stabilità della fase nematica fu proposta da Onsager attraverso la suddivisione dell'entropia di un sistema in due parti: un'entropia <i>orientazionale</i> che diminuisce durante la transizione di fase isotropo-nematica ed una <i>traslazionale</i> che invece aumenta stabilizzando il sistema. . . .	37
3.2	Equazione di stato per un sistema di sfere dure, ovvero un grafico della pressione in funzione della frazione di volume del sistema. La linea continua rappresenta la branca stabile, mentre la tratteggiata quella metastabile. La ricerca di una transizione di fase liquido-solida stabile in un sistema di sfere dure è durata più di un decennio, portando con sé grandi innovazioni sul campo della nascente fisica computazionale.	38
4.1	Modello di passeggiata aleatoria in tre dimensioni. I metodi Monte Carlo si basano sul campionamento casuale, che porta all'approssimazione di osservabili attraverso un'esplorazione completa lo spazio delle configurazioni.	56
4.2	Tre diverse funzioni di importance sampling (linee tratteggiate) usate per integrare la densità della distribuzione normale standard (linea continua) da -50 a 50. I pannelli superiori corrispondono alle curve di densità, mentre i pannelli inferiori sono gli istogrammi di 5000 stime Monte Carlo dell'area (che è esattamente 1). . . .	58
4.3	Illustrazione di un sistema di interesse sospeso in un bagno di calore. Il sistema di interesse è piccolo se comparato al bagno di calore.	65

4.4 Il sistema (di volume \mathbf{V}) è immerso in un bagno termico molto più grande a temperatura costante, chiuso in modo tale che il numero di particelle rimane fissato. Il sistema è separato dal bagno termico attraverso un pistone libero di muoversi, tale che il volume può variare.	68
5.1 Il modello per le molecole G_{20T} duplex usato nelle simulazioni. Le parti rosse sono cilindri rigidi duri. I centri A delle piccole sfere verdi, che modellano l’attrazione tra code dei DNA-Nunchaku, interagiscono attraverso un potenziale <i>square well</i> u_{sw} , mentre i centri B delle sfere arancioni grandi appartengono alle gapped-duplex stesse, che modellano la catena flessibile e interagiscono col potenziale u_p . Infine, il diametro delle sfere indica il raggio d’interazione, mentre u_0 (ovvero la profondità del potenziale <i>square well</i>) indica l’energia di legame.	70
6.1 Differenze di energia libera di Helmholtz tra le fasi $Sm-fA$ ed $Sm-uA$ in funzione della pressione a temperatura $T = 0.125$. La linea solida (ΔA) è stimata dall’equazione $A = G - PV$, dove G è l’energia libera di Gibbs. La linea tratteggiata, invece è l’energia libera di Helmholtz (ΔA_{ex}) basata soltanto sugli effetti di volume escluso	79
7.1 Andamento dell’energia adimensionale. Come si può notare dal grafico, il sistema ha termalizzato dopo i nostri venti milioni di passi.	81
7.2 configurazione iniziale della simulazione originaria	81
7.3 Rilevamento della fase isotropo-nematica attraverso il software <i>MOLGL</i> . Come si può evincere dalla figura, il sistema tende ad allinearsi lungo un unico asse d’orientazione, ovvero quello <i>nematico</i> . .	82

7.4	Raffigurazione dei parametri d'ordine nematico e smettico. Le linee rappresentano i diversi parametri misurati durante le simulazioni e come si può notare il parametro d'ordine nematico è sempre poco sopra 0.4, mentre quello smettico è quasi nullo.	84
7.5	Rilevamento della fase isotropa attraverso il software <i>MOLGL</i> . Come suggerisce il nome, ogni direzione in questa fase è puramente equiprobabile e casuale: ci troviamo di fronte ad un sistema con il comportamento più vicino a quello di un liquido. In questa foto la pressione è circa $\beta Pv_0 = 1.05$, mentre la frazione di folding è pari a 60/840 <i>Gapped-Duplexes</i> foldate.	86
7.6	Rilevamento della fase nematica attraverso il software <i>MOLGL</i> . Possiamo vedere come il sistema tende ad allinearsi lungo un unico asse d'orientazione, ovvero quello <i>nematico</i> . Con l'aumento della pressione, inoltre, sta cominciando l'effetto di folding degli elementi costituenti. In questa foto la pressione è circa $\beta Pv_0 = 1.20$, mentre la frazione di folding è pari a 161/840 <i>Gapped-Duplexes</i> foldate.	86
7.7	Rilevamento della fase smettica attraverso il software <i>MOLGL</i> . Qui l'effetto della stratificazione ha configurato finalmente il sistema in modo da creare dei piani impilati l'uno sopra l'altro. In questa foto la pressione è circa $\beta Pv_0 = 1.35$, mentre la frazione di folding è pari a 213/840 <i>Gapped-Duplexes</i> foldate.	87

7.8	Equazione di stato del sistema misurata al variare di K lungo l'intervallo di pressioni ridotte $\beta Pv_0 \in [1.05, 1.40]$. Soffermiamoci su $K = 0.14$: abbiamo una prima "gobba" in cui c'è un esorbitante aumento di pressione. Trattasi non altro che della fase nematica, in cui il sistema sta repentinamente cambiando di configurazione prima di stabilizzarsi lungo la fase smettica. Come si può vedere, infatti, dopo questa gobba ne sussegue un'altra, dopodiché la curva tende ad avere una pendenza fissa e qui abbiamo finalmente la fase smettica. Questo grafico è facilmente comparabile con quello presente sul <i>Nat.Comm.</i> del 2016, che apparirà nel capitolo successivo. Fino circa ai $165\text{mg}/\text{ml}$ di concentrazione abbiamo un comportamento quasi-liquido, mentre a partire dai $200\text{mg}/\text{ml}$ l'effetto della stratificazione si fa sentire lungo tutto il sistema. Eventuali differenze di concentrazione e pressione sono più che altro dovute al diametro ridotto del nostro sistema rispetto a quello originale, che tende ad abbassare i valori misurati rispetto al modello di riferimento a causa di una modifica dell'aspect ratio. Anche la curva raffigurante $K = 0.15$ presenta le due "gobbe", ma sono molto meno pronunciate. Per K maggiori, infine, non assistiamo ad entrambe le transizioni perché il sistema tende a rimanere in un'unica fase e osserviamo delle linee quasi continue.	88
8.1	Andamento dell'equazione di stato rispetto alla concentrazione (mg/ml) nell'esperimento originale sul quale si basa la maggior parte di questa tesi. Particolare di un'immagine apparsa sul <i>Natural Communications</i> 7 (2016).	89
9.1	Programma in <i>Python3</i> sulla costruzione dell'equazione di stato. L'ambiente su cui è stato implementato il programma è <i>Jupyter Notebook</i>	91
9.2	Programma in <i>Python3</i> sul calcolo della frazione di folding . . .	94

Capitolo 1

Introduzione

Per molto tempo si è pensato che il DNA non presentasse una fase smettica, prima a causa dei lavori a livello teorico di Livolant [1] e successivamente grazie alle simulazioni sulle sfere dure liotropiche  semicilindriche [2]. Tuttavia, più recentemente [3] si è visto come inserendo dei filamenti di 20 timine tra un filamento e l'altro di DNA, si riscontra la presenza di una fase smettica (o a cristalli liquidi). Nell'esperimento si è utilizzato un cosiddetto Gapped Duplex (o G_{20T} Duplex) con delle proprietà fisiche ben fissate, per esempio la lunghezza del filamento timinico era ben superiore alla lunghezza di persistenza del DNA. Ciò che vogliamo sperimentare nella tesi è l'introduzione nel modello di un potenziale di *bending* che regola il legame tra gli elementi costituenti del sistema (chiamato k)  ed i suoi effetti sul modello, chiamato anche *DNA-Nunchaku*. Al variare del parametro k , vogliamo osservare come emerge una fase nematica  nelle nostre Gapped Duplex ed il loro comportamento nelle transizioni di fase *isotropo-nematiche* e *nematico-smettiche*. Ci aspettiamo, dunque, che al variare di k compaiano o spariscano determinate transizioni di fase, ~~poiché il modello è molto sensibile alle configurazioni "trattenute" dal potenziale.~~

Quando è stata osservata questa fase nel DNA tipica dei cristalli liquidi si è subito potuta immaginare la potenzialità dell'acido desossiribonucleico nella costruzione ed ingegneria di nuove strutture, in particolare come mattoncino fondamentale per svariate costruzioni formate da cristalli liquidi liotropici.



Figura 1.1: Tre filamenti di DNA parzialmente complementari che formano una G-Duplex di DNA. La G-duplex consiste quindi in una regione centrale flessibile ~~di basi timiniche (tra 1 e 20) scoppiate~~, con un pezzo rigido lungo attaccato ad entrambe le estremità. La "catena" timinica ha una lunghezza molto più grande di quella di *persistenza* pari ad $l_p = 2 \text{ nm}$.

La tesi verrà suddivisa nei seguenti capitoli:



- Panoramica sui cristalli liquidi di DNA e sulle loro fasi.
- Illustrazione dei modelli coarse-grained e sul metodo Monte Carlo attraverso il quale sono state effettuate le misure.
- Spiegazione della fisica teorica sulla quale si basano le nostre misurazioni, principalmente si parlerà dei parametri d'ordine delle fasi dei cristalli liquidi.
- Discussione sui risultati delle simulazioni

Concluderemo infine con una sintesi sugli argomenti trattati e con un'esposizione delle future applicazioni dei risultati qui esposti.

1.1 Panoramica sulla struttura del DNA

1.1.1 Approcci top-down e bottom-up nella costruzione di nanostrutture

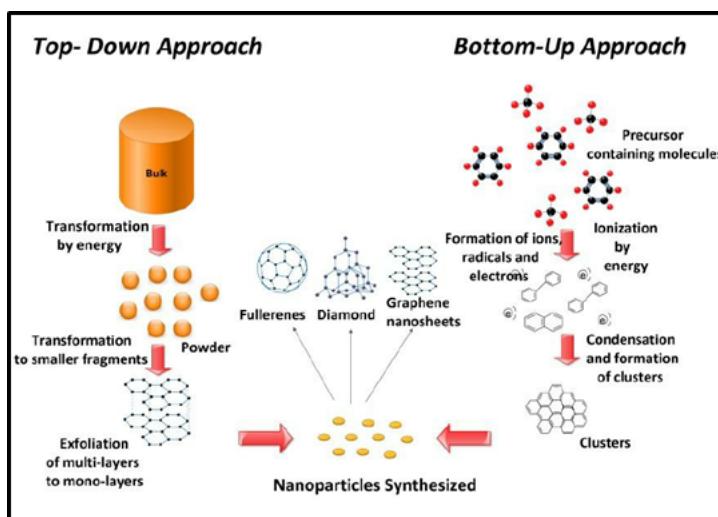


Figura 1.2: Vediamo qui uno schema dei due possibili approcci nella costruzione di nanostrutture: il *top-down* ed il *bottom-up*. L'approccio *top-down* richiede l'intervento di uno sperimentatore, attraverso tecniche come la *fotolitografia*: più in generale, viene richiesto un dispendio energetico per la produzione di frammenti nanometrici, sui quali vengono poi effettuate le misure vere e proprie. Nel *bottom-up*, invece, l'edilizia di queste strutture consiste interamente da processi d'interazione chimica, dunque indipendenti dall'osservazione esterna dello sperimentatore.

Quando parleremo del nostro sistema in esame, ci riferiremo spesso al termine di *nanostrutture* di DNA. La costruzione di nanostrutture è un campo relativamente moderno della fisica, che è stato dominato nel tempo da due approcci complementari. Il primo è il *top-down*, nel quale un sistema viene "sminuzzato" in parti più piccole. Di solito, questo sminuzzamento viene effettuato attraverso la tecnica della *fotolitografia*: ovvero, come indica il nome stesso della tecnica, la produzione di minuscoli tagli distribuiti lungo strati di film sottile [4] attraverso

l'utilizzo di fonti luminose. Questa operazione, purtroppo, necessita di condizioni molto complesse da ottenere in un normale laboratorio, come l'alto vuoto (una pressione nell'ambiente circostante di circa $10^{-7} Pa$). Inoltre, la risoluzione massima in cui si può osservare il sistema spesso non è soddisfacente, arrivando ad un massimo di $10 nm$.

Più recentemente, ha preso piede nella fisica delle nanostrutture l'approccio *bottom-up*, che consiste in una tecnica di auto-assemblaggio delle strutture in esame, indipendente dall'intervento di un operatore in laboratorio. In questi sistemi, le molecole sono in grado di impilarsi le une sulle altre ricreando le più variegate forme geometriche. Quest'ultimi, regolando alcuni parametri come la temperatura e la densità, tendono ad allinearsi tra loro lungo una direzione cardinale [5]. Un vantaggio importantissimo ricavato dall'utilizzo di questa tecnica è la *sintesi parallela*, che indica la produzione di agglomerati chimici (detti anche "librerie" o "biblioteche") a partire da singoli composti [6].

In questo panorama, il DNA è stato senza dubbio uno degli oggetti più interessanti col quale sperimentare la tecnica dell'auto-assemblaggio.

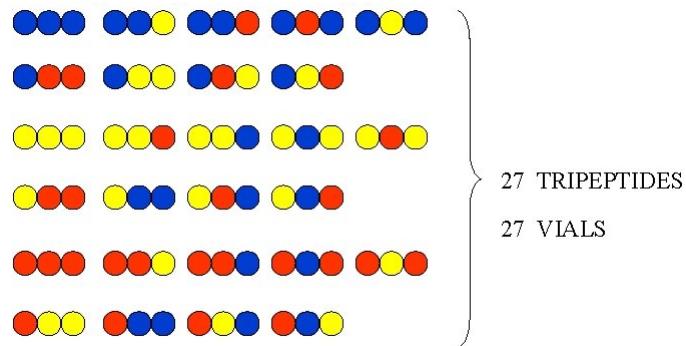


Figura 1.3: Un esempio teorico di *sintesi parallela*. Qui vediamo come da 3 singole viale di amminoacidi siano stati prodotti 27 tripeptidi

1.1.2 L'acido desossiribonucleico e le sue proprietà

Le proprietà chimiche con cui è scritta l'informazione del DNA fanno sì che questo possa essere copiato in un'altra molecola identica, essere trascritto in una molecola di RNA ed essere quindi interpretato per produrre le proteine necessarie alle funzioni della vita.

Il DNA (*acido desossiribonucleico*) e l'RNA (*acido ribonucleico*) sono acidi nucleici, si tratta di una catena di polimeri le cui singole unità sono dette *nucleotidi*. Gli ingredienti [7] fondamentali che costituiscono un nucleotide sono essenzialmente tre:

- Uno zucchero *pentoso*, ovvero a 5 atomi di carbonio
- Una base azotata *purinica*, formata cioè da due anelli azotati, o *pirimidinica*, che ne contengono uno solo.
- Uno o più gruppi fosfato, detti anche *residui fosforici*

La prima grande differenza tra il DNA e l'RNA è lo zucchero pentoso, che difatti viene indicato nella prima lettera delle loro sigle. Nel caso del DNA abbiamo il *desossiribosio*, mentre per l'RNA abbiamo il *ribosio*. In particolare, il ribosio manca di un gruppo ossidrile attacco al suo secondo atomo di carbonio (come si vede in figura).

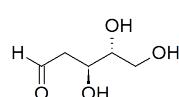
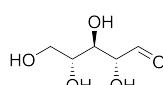


Figura 1.4: Comparazione tra le strutture dello zucchero *ribosio* e quello *desossiribosio*.

Come si può vedere, nel ribosio (figura soprastante) ritroviamo un gruppo OH nella posizione 2', che invece manca nello zucchero desossiribosio sottostante.

Come anticipato, abbiamo due classi di basi azotate. Queste sono le *purine*, di cui fanno parte l'adenina e la guanina (indicate rispettivamente con le lettere A e G) e le *pirimidine*, costituite dalla citosina (C); l'uracile (U) e infine la timina (T), che come vedremo sarà di grande importanza nella nostra tesi. La timina e l'uracile, inoltre, differiscono per un singolo gruppo metilico (CH_3) sulla posizione 5' dell'anello carbonico. Ritroviamo rispettivamente nel DNA la timina e nell'RNA l'uracile.

Lo zucchero pentoso, quando interagisce con la base azotata, forma un complesso detto *nucleoside*. La costituzione di un nucleoside avviene quando una base azotata si lega alla posizione 1' dello zucchero pentoso tramite un legame covalente, detto *glycosidico*.

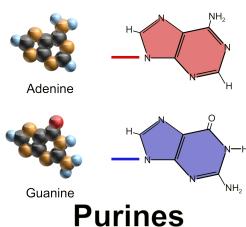


Figura 1.5: Rappresentazione tridimensionale e formula di struttura delle basi azotate puriniche

Finalmente, entra in gioco il gruppo fosfato (PO_4^{3-}). Quando questo si lega col nucleoside, otteniamo il *nucleotide*: il mattoncino unitario sulla cui concatenazione si basa la struttura del DNA. Per la polimerizzazione del DNA, ma in generale di tutti gli acidi nucleici, i nucleotidi hanno tre gruppi fosfato nella posizione 5'.

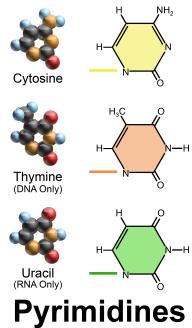


Figura 1.6: Rappresentazione tridimensionale e formula di struttura delle basi azotate pirimidiniche

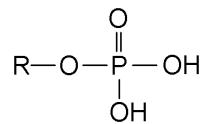


Figura 1.7: Formula di struttura di un gruppo fosfato legato ad un generico gruppo organico R: questo residuo fosforico costituisce una parte integrante nella struttura dei nucleotidi. In questo caso, un generico nucleotide che interagirebbe col gruppo in figura prenderebbe la denominazione di *nucleotide monofosfato*

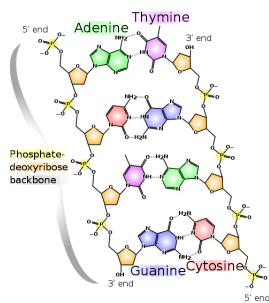


Figura 1.8: Formula di struttura del DNA: da notare l’alternanza tra zuccheri e gruppi fosfato.

A questo punto, la struttura molecolare del DNA viene a costituirsi attraverso un’alternanza tra zuccheri e gruppi fosfato. Iterando il processo, viene a formarsi una catena composta da nucleositi monofosfato. Quest’alternanza è generata dalla perdita di un pirofosfato durante la concatenazione, ovvero di due gruppi fosfato legati tra loro [8]. Più precisamente, ogni gruppo fosfato forma un legame, detto *fosfodiesterico*, tra il carbonio 5’ dello zucchero successivo e il 3’ di quello precedente. Questo legame è molto importante, perché determina una polarità intrinseca della catena polinucleotidica. Le due estremità, infatti, sono contrassegnate sia da un carbonio 3’ libero sia da un carbonio 5’ libero [9].

Già dalla metà dello scorso secolo era stato individuato un legame tra le basi pirimidiniche e le basi puriniche: la regola di Chargaff, infatti, egualava il numero di A con quello di T e il numero di C con le G già dalla fine degli anni ’40 (in termini di formula matematica: $A + G = T + C$). Il problema era trovare un modello che giustificasse questa ed altre osservazioni sperimentali. Attraverso la diffrazione a raggi X ed altre tecniche sviluppate per l’osservazione delle proteine, si riuscirono a determinare molte proprietà chimico-fisiche dell’acido desossiribonucleico. Il DNA, senza dubbio, è stato da sempre un oggetto di grande interesse. Tra i grandi scienziati che si occuparono dello studio del DNA figura anche il nome del premio nobel Linus Pauling, che nel 1952 elaborò un modello a tre eliche (formate da zucchero e fosfato) intrecciate con le basi sporgenti verso l’esterno. Bisognerà aspettare l’anno successivo per dimostrare l’erroneità del modello di Pauling, quando Franklin, Watson e Crick svilupparono il celeberrimo modello a doppia elica (questa volta a basi rivolte verso l’interno). Qui, l’adenina si lega con la timina, mentre la citosina si lega con la guanina. Oltre alla regola di Chargaff, si è potuto spiegare anche il diametro costante del DNA. La doppia elica, che ha un diametro di circa 2 nm , compie un giro completo ogni $3,4\text{ nm}$: questo fa sì che ci siano 10 coppie di basi presenti per ogni giro, visto che la distanza media tra una coppia di base e l’altra sia di circa $0,34\text{ nm}$. A questo punto, nella morfologia del DNA vengono a crearsi due solchi, dovuti all’asimmetria tra i legami delle basi e la concatenazione di zuccheri: il solco maggiore ($2,2\text{ nm}$)

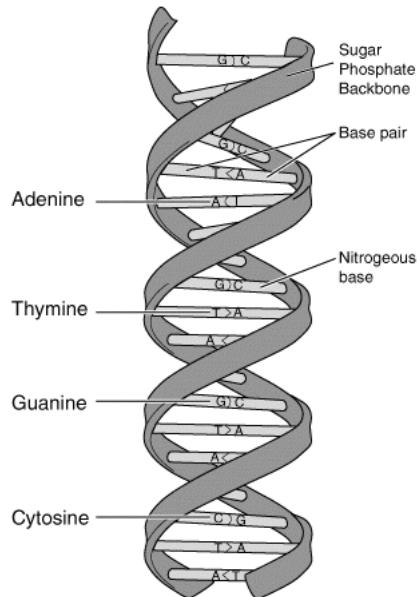


Figura 1.9: Modello di Watson-Crick: qui possiamo vedere esplicitamente i legami A-T e C-G

ed il solco minore ($1,2\text{ nm}$). Nel frattempo, le basi azotate interagiscono fra loro tramite legami idrogeno (un’attrazione fra un atomo d’idrogeno ed un qualsiasi altro atomo più elettronegativo)[10]. La geometria delle basi azotate fa sì che G-C formino tra loro tre legami, mentre A-T ne formino solo due. All’interno del DNA operano insieme legami covalenti e forze deboli, che cooperano per mantenere stabile la struttura dell’acido. I nucleotidi sono legati tra loro in maniera molto forte, perché interagiscono attraverso legami covalenti. Le due catene di cui è composto il DNA, tuttavia, devono permettere il movimento della doppia elica (essenzialmente, l’apertura e la chiusura) durante la fase di trascrizione e replicazione dell’informazione genetica. Per questo, le catene sono legate attraverso due forze “deboli”: queste sono l’accoppiamento tra basi (detto anche *pairing*) e l’impilamento (*stacking*). Abbiamo già discusso il *base pairing* poc’anzi, si tratta semplicemente dei legami a idrogeno tra A-T e C-G [11]. L’impilamento delle basi, invece, è dovuto ad un effetto di solvatazione causato dalla loro idrofobicità: le basi, per minimizzare il contatto con l’acqua, tendono a sovrapporsi le une

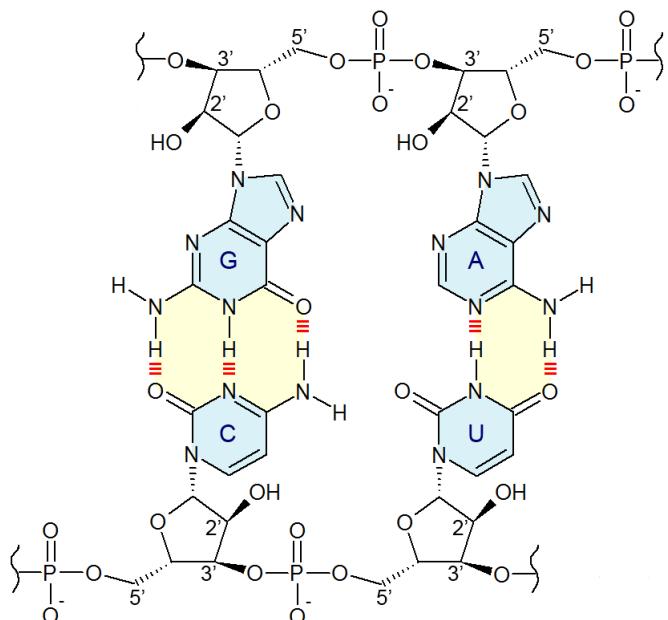


Figura 1.10: Accoppiamento di base: come si può notare, quattro monomeri di nucleotidi formano due coppie di basi: la guanina si accoppia con la citosina attraverso tre legami idrogeno, mentre l'adenina si accoppia con la timina (o nel caso della qui presente immagine, con l'uracile) attraverso due singoli legami idrogeno.

sulle altre creando appunto degli "strati" [12]. Una caratteristica interessante di queste forze deboli, però, si rivela solo analizzando l'energia libera di Gibbs dei processi di accoppiamento ed impilamento. L'energia libera di Gibbs ΔG è una funzione di stato termodinamica, che ci permette di osservare quando un processo fisico è *energicamente* favorevole ($\Delta G < 0$) o sfavorevole ($\Delta G > 0$). Osservando prima il fenomeno di *base pairing*, ci accorgiamo che il suo contributo alla stabilità del sistema è pressoché nullo: l'accoppiamento C-G registra una $\Delta G \approx 0$, mentre quello A-T risulta addirittura energicamente sfavorevole. La doppia elica, lasciata così, risulterebbe addirittura instabile. La forza debole responsabile per la stabilità della doppia elica, quindi, risulta essere l'impilamento, che favorisce la formazione della doppia elica.

1.1.3 Il punto di contatto tra genetica e nanotecnologie



Figura 1.11: Profondità: l'opera di Escher che ispirò la nascita della nanotecnologia strutturale del DNA. Come si può notare, la disposizione delle creature ricorda un reticolo cristallino di Bravais, tipico modello nella fisica dello stato solido

Svolta questa panoramica sulla struttura del DNA, verrebbe quindi da chiedersi: dove inizia il legame tra le nanostrutture e lo studio dell'acido desossiribonucleico? Tutto iniziò nel 1982, quando il biochimico Nadrian Seeman inventò il campo della *nanotecnologia strutturale del DNA* [13], ispirandosi all'opera del pittore olandese Maurits Cornelis Escher "Profondità" per proporre la costruzione di reticolni tridimensionali basati sulla struttura del DNA [14]. La necessità di questa tecnica fiorì quando si accorse, per i suoi studi di cristallografia, che l'orientazione di molecole pre-selezionate era molto più semplice dell'ottenimento di un vero cristallo su cui effettuare le misure. Similmente al modello che verrà utilizzato in questa tesi, la struttura di Seeman era caratterizzata da estremità adesive che chiamò *sticky-ends*, formate da bracci le cui estremità terminavano in un singolo filamento, formando un sito di legame. Utilizzando questi oggetti, Seeman ipotizzò di poter costruire oggetti complessi a partire da questi mattoncini fondamentali, cristallizzando il sistema con l'inserzione di macromolecole all'interno del *Bravais lattice* [15] [16].

Per fortuna, il DNA è un ottimo materiale sulle cui fondamenta ci si può poggiare.

re per la costruzione di modelli più complessi. La sua grandezza ci permette di lavorare lungo la scala nanometrica e il modello Franklin-Watson-Crick riesce a prevedere le interazioni tra base accoppiate di diverso tipo. Questo permette di creare i siti che desideriamo, come per esempio le estremità attrattive teorizzate da Seeman. L'altra caratteristica peculiare del DNA è la sua malleabilità, in particolare come possa risultare sia flessibile sia rigido a seconda delle condizioni in cui si trova. All'estremità delle giunzioni abbiamo infatti dei ripiegamenti che resistono anche a molte micro-deformazioni, permettendo addirittura di adattarsi a strutture dalle geometrie più differenti. Lungo il tratto che lega le estremità, invece, il comportamento è quello di un'asta rigida lunga circa 50 nm. Queste catene di nucleotidi, che formano la base delle nanostrutture di DNA, sono realizzabili in maniera completamente automatizzata [17]. Attraverso l'utilizzo di specifici enzimi, la modifica del DNA può avvenire persino lungo i singoli siti, permettendo una manipolazione dell'acido desossiribonucleico nei particolari più certosini.

I metodi di costruzione del DNA sono essenzialmente due: il primo è il *tiles*, che come indica il nome è basato sulla costruzione di mattonelle, mentre il secondo è *origami*, che si basa sul ripiegamento di un filamento più grande.



Figura 1.12: Questa immagine apparsa per la prima volta sul Nature del 16 marzo 2006 ha fatto il giro del mondo: si tratta di uno "smile" grande poco meno di 100 nm costruito attraverso la tecnica *DNA origami*. La sequenza genetica con cui hanno costruito questa faccina proviene da un virus batteriofago ed è lunga 7 kB, ovvero 7000 paia di basi (dove l'unità di misura della lunghezza è il pB, ovvero la distanza tipica tra un paio di basi azotate pari circa a 0,34 nm)

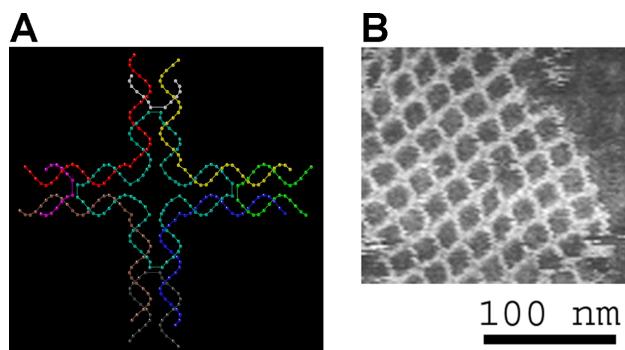


Figura 1.13: A sinistra, un esempio di "casella" di nano-DNA. Questo modello è stato usato per costruire un reticolo periodico bidimensionale, a destra si può notare la struttura visualizzata attraverso un microscopio a forza atomica.

Analizziamo prima la tecnica del *DNA tiles*. Si tratta della primissima tecnologia nanostrutturale del DNA, inventata da Seeman stesso nel 1982. In questo caso, i singoli filamenti (denominati *ssDNA*, dove "ss" sta per *single stranded*) vengono costruiti appositamente per formare un "mattoncino". Questo non è altro che il blocco di partenza con cui facciamo partire l'auto-assemblaggio di strutture più complicate. Via via che continuavano gli esperimenti [19], i bracci delle strutture di Seeman aumentarono creando complessi sempre più ramificati. Inizialmente, tuttavia, gli esperimenti non procedettero in maniera spedita: l'elevata flessibilità strutturale impediva la costruzione di oggetti particolarmente complessi. Per superare questo ostacolo bisognerà aspettare il 1996, quando Seeman progettò una tecnica di costruzione di nanostrutture basata sui crossover doppi (DX) e tripli (TX) del DNA, che ha reso più rigide le componenti delle "caselle". Il nuovo metodo spalancò le porte della tecnologia nanostrutturale, permettendo la costruzione di una vasta gamma di geometrie: reticolati, cristalli e strutture di DNA ancora più complesse. L'approccio smise di essere a singolo filamento per evolvere in una tecnica utilizzante gli *msDNA*, ovvero "*multistranded*". Rispetto alla tecnologia originaria, però, questo nuovo approccio comportava delle complicazioni considerevoli [21], come lunghi tempi di calcolo (che conducevano alla possibilità di errori durante l'assemblaggio) e la necessità di una regolazione minuziosa della stechiometria e del filtraggio dalle impurità dei singoli mattoncini di base.

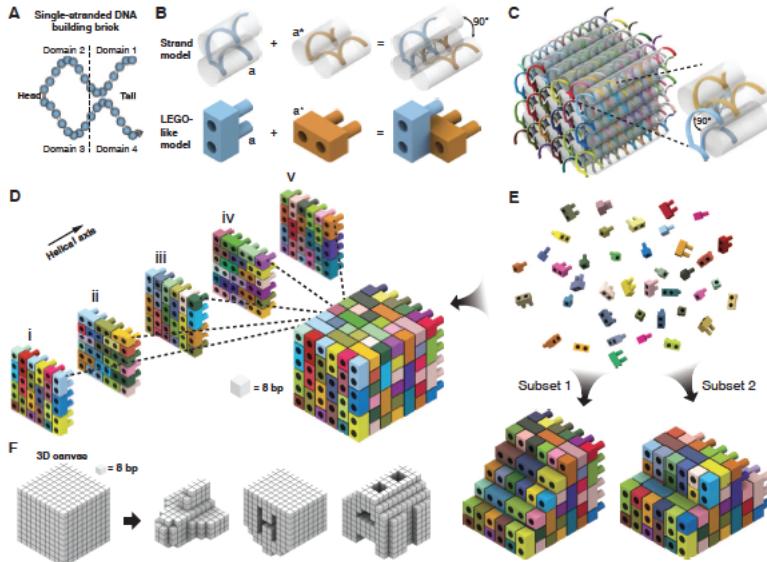


Fig. 1.14: Design of DNA brick structures analogous to structures built of LEGO® bricks. (A) A 32-nt four-domain single-stranded DNA brick. Each domain is 8 nt in length. The connected domains 2 and 3 are “head” domains; domains 1 and 4 are “tail” domains. (B) Each two-brick assembly forms a 90° dihedral angle via hybridization of two complementary 8-nt domains “a” and “a*”. (C) A molecular model that shows the helical structure of a 6H by 6H by 48B cuboid 3D DNA structure. Each strand has a particular sequence, indicated by a distinct color. The inset shows a pair of bricks. (D) A LEGO-like model of

Figura 1.14: Una rappresentazione del modello *DNA bricks* presa direttamente dall’articolo originale su cui la tecnica fu inventata. In una maniera piuttosto intuitiva, i singoli mattoncini di DNA vengono qui paragonati a dei pezzi di LEGO [18].

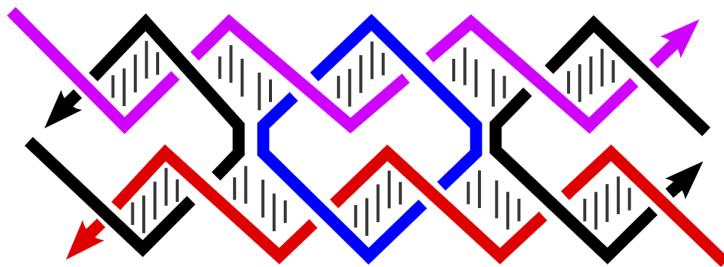


Figura 1.15: Una molecola (DX) a doppio crossover. Questa molecola è costituita di cinque singoli elementi di DNA i quali formano due domini doppio-elicoidali, a sinistra e a destra in questa immagine. Ci sono due punti crossover dove i componenti attraversano da un dominio all’altro. [20]

Per molto tempo, queste complicazioni hanno rallentato lo sviluppo delle nanostrutture di DNA fino al 2012 [22], quando Wei, Dai e Yin utilizzarono degli *ssDNA* per costruire forme molto più complesse di quanto si fosse mai fatto in precedenza. Mentre le caselle multistranded formavano delle strutture compatte attraverso dei siti d'interazione, la nuova tecnica si è basata sulla concatenazione degli *sticky ends* inventati da Seeman fino a formare un filamento da 42 basi. Infine, nello stesso anno questa stessa tecnica fu perfezionata [18] attraverso il *DNA brick*. Trattasi di una catena da 32 nucleotidi (definita *NT*), divisa in 4 domini consecutivi formati da 8 *NT*. La morfologia acquisita è quella di due eliche antiparallele lunghe 16 *NT*, unita da un legame fosfato. I due domini adiacenti al legame fosfato sono molto importanti, perché quando si legano con un dominio di un altro "mattoncino" *NON* adiacente al legame fosfato formano un angolo di 90°. Per questa ragione, i due domini adiacenti al legame fosfato vengono detti *testa*, mentre i restanti due vengono chiamati *coda*. Parliamo dunque della tecnica

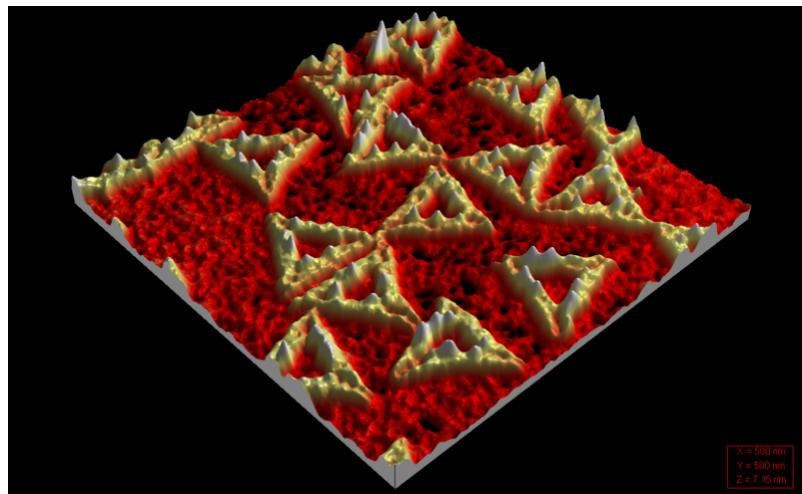


Figura 1.16: Origami di DNA visualizzato attraverso un microscopio a scansione di sonda. Notare le forme triangolari che si iterano periodicamente lungo la superficie

del *DNA origami*, resa celebre dall'immagine dello "smiley" apparso sulla rivista *Nature* nel 2006. Si capisce subito, quindi, che si tratta di un metodo molto più moderno delle "caselle" di Seeman. Il primo passo di questa tecnica avvenne

quando Paul Rothemund [23] ripiegò su sé stessa una singola lunga catena di DNA (chiamata *scaffold*) utilizzando diverse centinaia di oligonucleotidi (che denominò *graffette*). Metodi di ripiegamento non erano sconosciuti alla scienza [24], ma l'origami fu di gran lunga più efficiente delle tecniche utilizzate in precedenza. Utilizzando più di 200 "graffette", aveva infatti ripiegato un filamento lungo 7 kilobasi (come illustrato nella figura dello "smiley"). Il metodo origami, inoltre, non si ferma alle strutture planari, ma grazie ad esso è possibile costruire delle morfologie tridimensionali con una cavità interna [25], come fu dimostrato nel 2009 con la costruzione di una scatola dal coperchio semovibile. Tuttavia, nessuna tecnica è priva di svantaggi: quello principale dell'origami a "singolo strato" era la forte suscettibilità agli sforzi meccanici, mentre la resistenza e la malleabilità alle deformazioni era una caratteristica tipica dei *DNA tiles*. Per ovviare a questo problema,

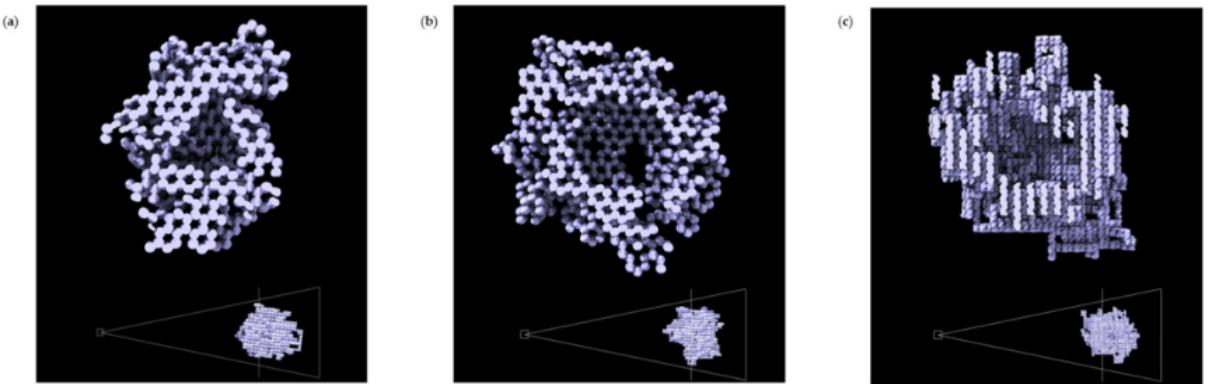


Figura 1.17: Strutture di DNA origami generate proceduralmente da un algoritmo informatico [26]. Dai tempi di Seeman, la branca delle nanotecnologie legate al DNA si è evoluta incessantemente e le potenzialità di questi studi offrono grandi speranze per un ulteriore legame tra fisica e biologia.

analogamente a ciò che avvenne per le "caselle" di DNA, furono inventate delle tecniche origami a strati multipli. Riprendendo da un'idea originale di Seeman del 1991 [27], fu ideato il metodo *wireframe*: ciò consiste nel ripiegamento di un singolo filamento, a partire da una struttura bidimensionale, per costruire una figura in 3D. Questo metodo ha creato strutture molto complesse, con un volume e una superficie molto elevata rispetto alla planimetria di partenza. Più recentemente,

nel 2009 Shawn Douglas [28] utilizzò diverse "graffette" per ripiegare un singolo filamento. Da questo singolo mattoncino, formò delle eliche antiparallele densamente impacchettate, unite fra loro con un cross-over dalla morfologia esagonale. L'assemblaggio, in questo caso, è decisamente più complesso e quindi comporta ulteriori tempi di calcolo, ma la struttura finale è di gran lunga più robusta. Come pubblicato su un articolo uscito recentemente ~~il mese scorso~~ [26], le tecniche di nanostruttura del DNA stanno continuando ad evolvere verso nuovi orizzonti.

Capitolo 2

Panoramica sui cristalli liquidi

In questo capitolo faremo un rapido riepilogo delle proprietà dei cristalli liquidi e delle loro fasi.

I cristalli liquidi non sono altro che fluidi anisotropi che condividono le proprietà sia dei solidi sia dei liquidi [29]. Sono detti anche *mesofasi*, ovvero fasi con un ordinamento intermedio tra quello cristallino e liquido. In particolare, si tratta di uno stato intermedio in cui non abbiamo né il totale ordinamento  dei centri di massa in un reticolo cristallino (come in un cristallo) né la totale isotropia di un liquido. Ciò che abbiamo, invece, è un ordine anisotropo almeno in una direzione, nella quale il sistema possiede un ordine simile a quello liquido, ovvero le particelle sono libere di muoversi in certi limiti con fluidità  Ciò che ha permesso ad oggi la diffusione in larga scala di questa classe di composti sono le loro caratteristiche peculiari, come la dipendenza del loro ordine (e quindi delle loro fasi) dai campi elettromagnetici e i loro flussi caratteristici determinati dalla loro anisotropia. Per distinguere le fasi, ciò che facciamo è controllare il numero di simmetrie del nostro sistema, in particolare l'ordine di orientazione e posizionale.

2.1 Cristalli liquidi di DNA

Nella sua forma più comune, il DNA si presenta come una doppia elica destrorsa di due catene antiparallele (quindi di polarità opposte) con le basi azotate

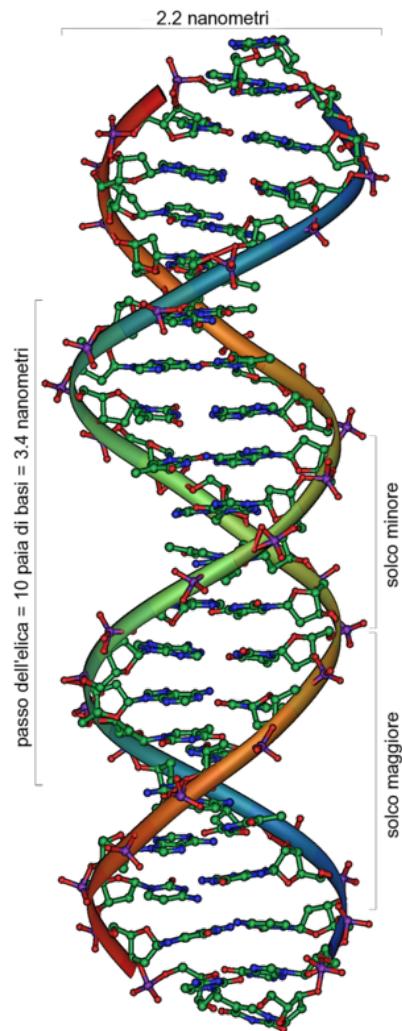


Figura 2.1: Struttura a *doppia elica* del DNA.

inserite perpendicolarmente ad esse. In particolare, si vengono a creare due solchi di dimensioni diverse, a causa dell’asimmetria tra i legami delle basi e i costituenti dell’impalcatura della catena (principalmente zuccheri). Le suddette basi interagiscono tra loro tramite legami a idrogeno [30], quindi un’interazione attrattiva. Le doppie eliche del DNA presentano una fase liquido-cristallina quando immerse in una soluzione, a causa della loro anisotropia. Inoltre, le loro dimensioni mesoscopiche li rendono per definizione dei *materiali colloidali* [31]. Difatti, i colloidì (che ruotano attorno alla grandezza di 1nm ed $1\mu\text{m}$) sono molto più grandi del

solvente in cui sono immersi, ma sufficientemente piccoli da disperdersi nel fluido. Insieme ai cristalli liquidi ed altri polimeri o materiali biologici in soluzione, i colloidì fanno parte di una classe di materiali detti *soffici* [32] poiché l'energia necessaria alla loro deformazione è molto minore rispetto a quella di un cristallo. La stabilità dei filamenti è governata essenzialmente da due forze. Per quanto i legami tra i nucleotidi siano molto forti, perché covalenti, si è visto [34] che le due catene sono tenute assieme da *forze deboli* per permettere l'apertura della doppia elica. La prima di queste è l'appaiamento tra basi [35], che corrisponde ai legami idrogeno tra basi azotate e non contribuisce di per sé alla stabilità del sistema. La vera forza principale è l'impilamento delle basi (detto *interazione di stacking*), che mantiene unite le catene polinucleotidiche. Dato che le basi sono idrofobiche, tendono a sovrapporsi per minimizzare il contatto con l'acqua e quindi la forza è dovuta principalmente [36] ad effetti di solvatazione.

2.2 Fase isotropa

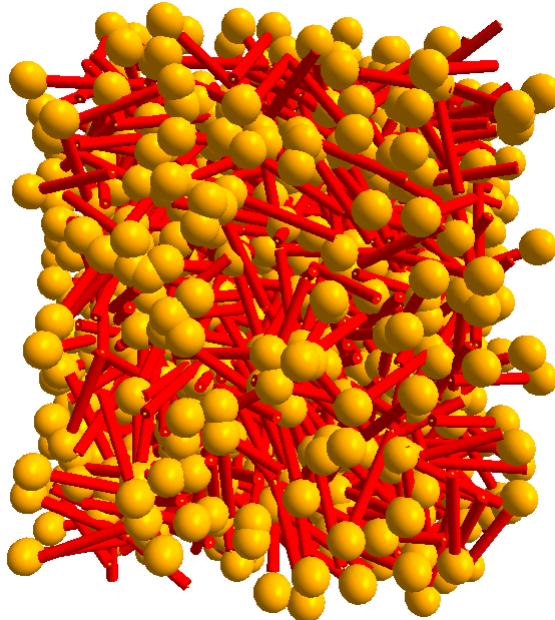


Figura 2.2: Fase isotropa delle gapped-duplex DNA. Come si può notare facilmente da questa immagine catturata col software *MOLGL*, ogni direzione è equivalente ed il comportamento del sistema è molto simile a quello di un liquido.

La fase in cui il nostro sistema conserva di più le caratteristiche di un liquido è quella isotropa. Nella fase isotropa, infatti, ogni direzione è equivalente e quindi le particelle possono trovarsi in qualunque posizione e direzione. Osservando il sistema non troviamo alcun tipo di strutture particolari e nessun tipo di allineamento. La somiglianza coi liquidi è tale che su qualunque direzione abbiamo una lunghezza di correlazione oltre la quale ogni particella si comporterà indipendentemente dalle altre. Questa distanza non sarà molto lunga e ci aspetteremo di trovare soltanto correlazioni a corto raggio dovute alle interazioni tra primi vicini.

2.3 Fase Nematica

Nella fase nematica [37], riscontriamo un allineamento delle particelle attorno a un singolo asse, detto asse nematico, ma all'interno del sistema stesso non si denota alcun tipo di ordine posizionale. Questo vuol dire che il sistema ha ancora molte somiglianze coi liquidi, visto che il centro di massa delle sue molecole è distribuito isotropicamente nello spazio.

Se vogliamo parlare in forma più generale di questa fase, possiamo vedere il caso in cui l'asse nematico non è fisso nello spazio. In questo modo, quindi, l'asse forma un'elica di passo p e si parla in questo caso di *fase colesterica*, che è tipica delle particelle chirali (motivo per cui questa fase è detta anche *nematico chirale*). A questo punto diventa intuitivo interpretare la fase nematica come una fase chirale il cui passo dell'elica tende all'infinito.

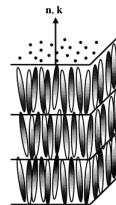


Figura 2.3: Modello di struttura della fase smettica A. La "direttoria" \mathbf{n} è orientata parallelamente alla normale \mathbf{k} degli strati smettici, mentre all'interno di ogni singolo strato smettico i centri di massa delle molecole sono distribuiti isotropicamente.

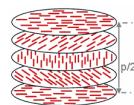


Figura 2.4: Modello di struttura della fase nematico-chirale. La p si riferisce al passo chirale.

2.4 Fase Smettica

Nella fase smettica, invece, non abbiamo soltanto un allineamento orientazionale lungo un unico asse di simmetria, ma ritroviamo anche una ben definita disposizione delle molecole. Abbiamo diversi tipi di fasi smettiche nei cristalli liquidi, ma ciò che le accomuna più di ogni altra caratteristica è la presenza di strutture a strati all'interno del sistema. In particolare, le particelle si dispongono su diversi piani che possono ~~dirigere~~ con l'asse nematico una generica angolazione (o esserne addirittura perpendicolari). I piani sono disposti ciascuno ad una distanza costante dall'altro, ma al loro interno il sistema si comporta come un liquido bidimensionale, ovvero non vi è alcun ordine posizionale ~~o orientazionale~~. La fase smettica è stata sempre piuttosto elusiva nello studio dei cristalli liquidi

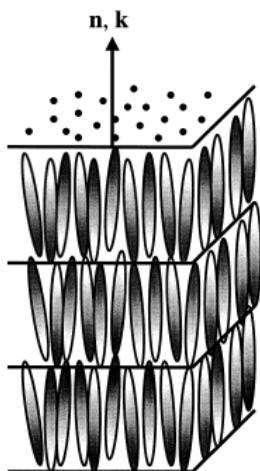


Figura 2.5: Modello di struttura della fase smettica A. La "direttoria" \mathbf{n} è orientata parallelamente alla normale \mathbf{k} degli strati smettici, mentre all'interno di ogni singolo strato smettico i centri di massa delle molecole sono distribuiti isotropicamente.

di DNA, a differenza di altre mesofasi che emergono a concentrazioni sufficientemente elevate [1]. Difatti, la presenza di una fase smettica è stata in passato proposta [39] per poi essere prontamente smentita [40], fino ai più recenti studi sulle gapped-duplex. Nel nostro lavoro potremmo riferirci alla fase del sistema

usando la nomenclatura $Sm-fA$ che indica la fase smettica SmA (in cui i piani smettici sono perpendicolari all'asse nematico) di molecole per la maggior parte in stato "*folded*". Sullo stato *folded* o *unfolded* degli elementi del nostro sistema discuteremo più avanti nel capitolo dedicato alla fisica teorica dietro i nostri modelli.

Capitolo 3

Teoria di Onsager

Nella teoria di Lars Onsager proposta nel 1949 [41], la transizione isotropo-nematica ha l'importanza di essere la prima transizione di fase puramente entropica. Con questo s'intende che la transizione è emersa esclusivamente attraverso l'interazione entropica  tra particelle, perciò ~~regolate semplicemente dalla temperatura e dalla configurazione del sistema~~ . Prima di Onsager, le transizioni di fase venivano interpretate come la manifestazione di una competizione tra contributi entalpici ed entropici all'energia del sistema, di conseguenza una transizione di fase puramente entropica non era prevista da un modello del genere. La dimostrazione della natura entropica della transizione di fase isotropo-nematica ha una certa eleganza logica.

Essenzialmente, all'epoca di Onsager veniva proposto un esempio astratto sulla natura della transizione. Assumiamo di avere un generico numero di cilindri dalle caratteristiche identiche: come sappiamo l'energia libera di Helmholtz del sistema è data dalla celebre formula $F = U - TS$, dove U sta per l'energia interna (che manterremo costante) ed S indica l'entropia. Ora, quando il sistema passa dalla fase isotropa a quella nematica le particelle tendono ad allinearsi lungo l'asse nematico. Per questa ragione, il sistema tende ad essere più ordinato e di conseguenza la sua entropia deve diminuire. Tenendo conto del fatto che abbiamo mantenuto costante l'energia interna, questo implica anche che l'energia libera è aumentata. Uno sbalzo energetico del genere avrebbe dunque dovuto dimostrare

l'instabilità di una transizione di fase puramente entropica.

La genialità dell'osservazione di Onsager si trova di certo nella sua disarmante

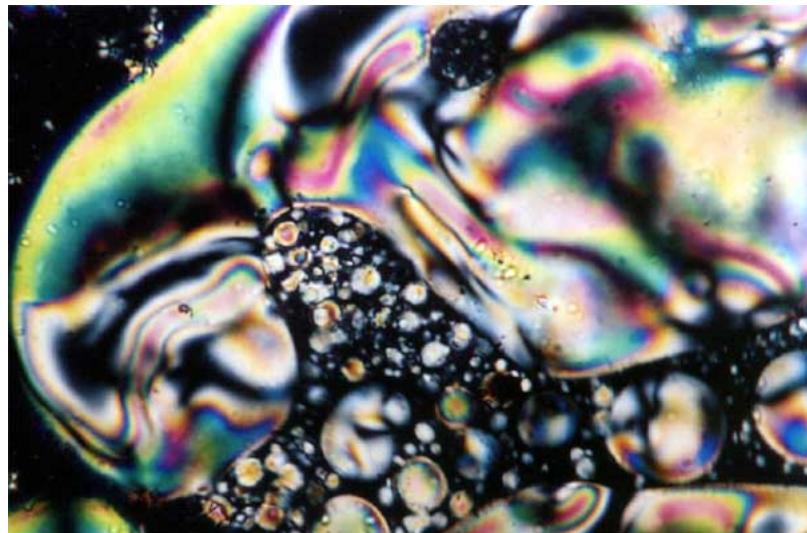


Figura 3.1: Rappresentazione di un cristallo liquido nematico. La stabilità della fase nematica fu proposta da Onsager attraverso la suddivisione dell'entropia di un sistema in due parti: un'entropia *orientazionale* che diminuisce durante la transizione di fase isotropo-nematica ed una *traslazionale* che invece aumenta stabilizzando il sistema.

semplicità: fece notare, infatti, che due tipi di entropie giocavano un ruolo in quel sistema. A diminuire è l'*entropia orientazionale*, ma allineandosi lungo l'asse nematico le particelle usufruiscono di un maggior spazio a causa della diminuzione del volume escluso. Quindi esiste un'altra funzione di stato termodinamica che sta aumentando di valore: l'*entropia traslazionale*. L'aumento di entropia traslazionale controbilancia la diminuzione di entropia orientazionale e ciò sta a significare che l'energia libera complessiva sta diminuendo, rendendo perciò stabile la fase nematica.

Ci volle molto tempo prima che la teoria di Onsager fosse pienamente accettata dal mondo accademico. In particolare, nel corso degli anni '50, gli scienziati William Wood e James Adler ricercarono la transizione dallo stato liquido al c:

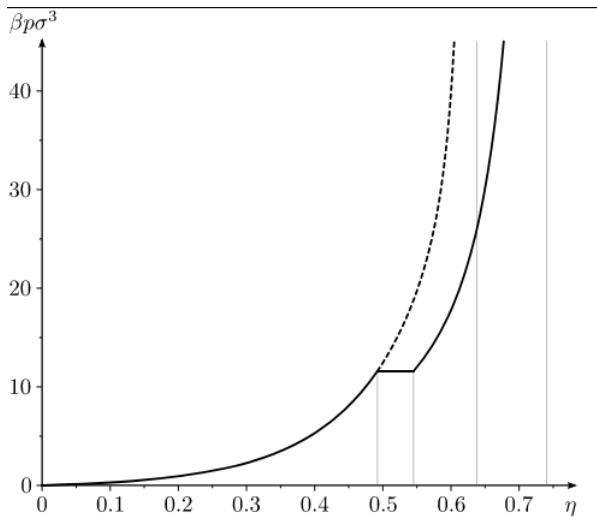


Figura 3.2: Equazione di stato per un sistema di sfere dure, ovvero un grafico della pressione in funzione della frazione di volume del sistema. La linea continua rappresenta la branca stabile, mentre la tratteggiata quella metastabile. La ricerca di una transizione di fase liquido-solida stabile in un sistema di sfere dure è durata più di un decennio, portando con sé grandi innovazioni sul campo della nascente fisica computazionale.

stallino in un sistema di sfere dure utilizzando dei metodi analitici che all'epoca erano estremamente innovativi. Si trattò di uno dei primi studi di fisica teorica che usufruirono della potenza di calcolo dei computer. William Wood [42] si servì del metodo che oggi chiamiamo Metropolis-Montecarlo , mentre James Adler studiò lo stesso sistema con delle simulazioni di dinamica molecolare. Si dovette aspettare le prime rappresentazioni grafiche di queste simulazioni, nel corso degli anni '60, per avere una prova *visiva* dell'esistenza di transizioni di fase puramente entropiche.

Illustriamo di seguito la teoria di Onsager [43] [44] riprendendo in qualche modo l'esempio precedente, immaginiamo di avere un sistema di N molecole dotate di simmetria cilindrica immerse in un volume V . Abbiamo già visto che un tale sistema va descritto non solo con le coordinate dei baricentri, ma anche con una

variabile orientazionale. A ciascuna particella, dunque, va assegnato un insieme di 6 coordinate: 3 traslazionali che ne individuano la posizione nello spazio, che per la particella i -esima corrisponde ad un vettore $\mathbf{r}_i = (x_i, y_i, z_i)$, più 3 orientazionali che indicano la posizione dell'allineamento, ovvero $\hat{\mathbf{u}}_i = (u_{ix}, u_{iy}, u_{iz})$ con $|\hat{\mathbf{u}}_i| = 1$. Ai gradi di libertà orientazionali si associa la distribuzione di singola particella $f(\hat{\mathbf{u}}_i)$, definita in modo che $f(\hat{\mathbf{u}}_i)d\Omega_i$ sia pari alla probabilità che la particella i -esima abbia un'orientazione contenuta nell'angolo solido $d\Omega_i$ centrato nella direzione $\hat{\mathbf{u}}_i$.

Le molecole interagiscono mediante il seguente potenziale:

$$U(\{\mathbf{r}\}, \{\hat{\mathbf{u}}\}) = \sum_{i < j} u_{ij} = \sum_{i < j} u(\mathbf{r}_i, \mathbf{r}_j, \hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j)$$

che è dato dalla somma delle singole interazioni di coppia. Se il sistema è abbastanza diluito, si può scrivere l'energia libera di Helmholtz usando il coefficiente del viriale (di cui parleremo più a fondo successivamente) troncato al prim'ordine nella densità $\rho = \frac{N}{V}$:

$$\frac{F}{Nk_B T} = \ln \rho - 1 - \Delta S_{or} + B_2 \rho,$$

dove ΔS_{or} è l'entropia orientazionale:

$$\Delta S_{or}[f(\hat{\mathbf{u}})] = - \int f(\hat{\mathbf{u}}) \ln(4\pi f(\hat{\mathbf{u}})) d\Omega \equiv -\sigma$$

e B_2 è il secondo coefficiente del viriale:

$$B_2 = -\frac{1}{2V} \int \Phi_{12} f(\hat{\mathbf{u}}_1) f(\hat{\mathbf{u}}_2) d\mathbf{r}_1 d\mathbf{r}_2 d\Omega_1 d\Omega_2$$

in cui compare la funzione di Mayer $\Phi_{ij} = e^{-\frac{u_{ij}}{k_B T}} - 1$.

Nel caso di particelle rigide, il potenziale di coppia è

$$u_{ij} \begin{cases} \infty & \text{se si presenta una sovrapposizione} \\ 0 & \text{altrimenti} \end{cases}$$

Se ne deduce che il coefficiente B_2 non è altro che la metà del volume escluso tra due particelle mediato sulle orientazioni. Il volume escluso tra due particelle, infatti, è dato da:

$$v_{exc}(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) = -\frac{1}{V} \int \Phi_{12}(\mathbf{r}_1, \mathbf{r}_2, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) d\mathbf{r}_1 d\mathbf{r}_2$$

perciò B_2 può essere riscritto come:

$$B_2 = \frac{1}{2} \int v_{exc}(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) f(\hat{\mathbf{u}}_1) f(\hat{\mathbf{u}}_2) d\Omega_1 d\Omega_2.$$

Nota l'energia libera, si possono trovare i valori delle concentrazioni alla coesistenza imponendo l'uguaglianza della pressione e del potenziale chimico nelle due fasi:

$$\begin{cases} P_{iso} = P_n \\ \mu_{iso} = \mu_n \end{cases}$$

L'esistenza di soluzioni accettabili del qui presente sistema implica l'esistenza di una transizione di fase.

Anche se la strategia appena delineata appare semplice in linea di principio, all'atto pratico la questione è abbastanza complicata. Infatti, nella nostra definizione dell'energia libera di Helmholtz del sistema abbiamo i due parametri σ e B_2 che sono in realtà dei funzionali della distribuzione $f(\hat{\mathbf{u}})$ e dipendono dunque dalla fase considerata. Il caso nematico, poi, complica ulteriormente le cose.

Per mostrare come si risolve la situazione, bisogna assumere che le particelle siano dei cilindri di diametro D ed altezza L e partire dalla fase più semplice. Nello stato isotropo, tutte le orientazioni sono equivalenti per definizione e dunque casuali: la loro distribuzione è allora un semplicissimo $f(\hat{\mathbf{u}}) = \frac{1}{4\pi}$ ed i parametri prendono i valori di:

$$\begin{cases} \Delta S_{or}^{iso} = \sigma_{iso} = 0 \\ B_2^{iso} \equiv b = \frac{\pi}{4} DL^2 \end{cases}$$

dove B_2^{iso} viene determinato approssimando che, se l'elongazione dei cilindri è $X_0 = L/D \gg 1$, il contributo del volume escluso dovuto all'estremità degli elementi del sistema è pressoché trascurabile. Infatti, per i cilindri molto lunghi vale la seguente approssimazione analiticamente dimostrata dallo stesso Onsager:

$$v_{exc} \simeq 2DL^2 |\sin \gamma|,$$

dove γ è l'angolo formato dai versori $\hat{\mathbf{u}}_1$ ed $\hat{\mathbf{u}}_2$, ovvero $\hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2 = \cos \gamma$.

Integrando l'approssimazione di Onsager sulle orientazioni, si ottiene proprio la nostra definizione di B_2^{iso} .

A questo punto ci conviene riscrivere i nostri risultati, utilizzando le variabili adimensionali:

$$\begin{cases} c = b\rho \\ B = \frac{B_2}{b} = \frac{4}{\pi} \int |\sin \gamma(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)| f(\hat{\mathbf{u}}_1) f(\hat{\mathbf{u}}_2) d\Omega_1 d\Omega_2 \end{cases}$$

di modo che l'energia libera possa essere riscritta come

$$\frac{\Delta F}{Nk_B T} = cost. + \ln c + \sigma + Bc,$$

in questa notazione nella fase isotropa si ottiene

$$\begin{cases} \sigma_{iso} = 0 \\ B_{iso} = 1 \end{cases}$$

mentre in quella nematica

$$\begin{cases} \sigma_n > 0 \\ B_n < 1 \end{cases}$$

Quindi per ora i parametri nematici restano indeterminati: per conoscerli bisogna scoprire la forma della $f(\hat{\mathbf{u}})$.

Per risolvere questo problema cercheremo di minimizzare l'energia libera. Ciò sta a dire che troveremo quella f tale che:

$$\frac{\delta}{\delta f(\hat{\mathbf{u}})} \left(\frac{\Delta F[f(\hat{\mathbf{u}})]}{Nk_B T} \right) = 0.$$

Da questa condizione si ottiene la seguente equazione integrale:

$$\ln(4\pi f(\hat{\mathbf{u}})) + 1 + \frac{8}{\pi} \int |\sin \gamma(\hat{\mathbf{u}}, \hat{\mathbf{u}}')| f(\hat{\mathbf{u}}') d\Omega' = 0.$$

Risolvendo tale equazione si troverà la forma funzionale ottimale per la f , ma bisogna ammettere che è piuttosto difficile trovare una soluzione per un'equazione del genere. In alternativa, Onsager propose una riformulazione attraverso una funzione di prova $f_\alpha(\theta)$, in modo che la minimizzazione possa essere effettuata sul parametro α senza per questo toccare il campo delle derivate funzionali e delle equazioni integrali. Nella definizione della funzione di prova richiediamo che soddisfi solamente due condizioni, suggerite dalla geometria del sistema:

$$\begin{cases} f_\alpha(\hat{\mathbf{u}}) = f_\alpha(\theta) \\ f_\alpha(\theta) = f_\alpha(\pi - \theta) \end{cases}$$

dove $\hat{\mathbf{u}} \cdot \hat{\mathbf{n}} = \cos \theta$.

La forma funzionale di Onsager è la seguente:

$$f_\alpha(\theta) = \frac{\alpha}{4\pi \sinh \alpha} \cosh(\alpha \cos \theta).$$

Si verifica facilmente che, se $\alpha \rightarrow 0$ (ovvero ci troviamo nello stato isotropo), risulta $f_\alpha(\theta) \approx \frac{1}{4\pi}$, mentre se $\alpha \rightarrow \infty$ (dunque siamo nello stato nematico) si ottiene un andamento gaussiano della forma:

$$f_\alpha(\theta) \begin{cases} \frac{\alpha}{4\pi} e^{-\alpha \frac{\theta^2}{2}} & \text{se } \theta \in [0, \pi/2] \\ \frac{\alpha}{4\pi} e^{-\alpha \frac{(\pi-\theta)^2}{2}} & \text{se } \theta \in [\pi/2, \pi] \end{cases}$$

La normalizzazione di questa gaussiana è corretta a meno di termini trascurabili

per grandi valori di α :

$$\int_0^{2\pi} \int_0^\pi f_\alpha(\theta) \sin \theta d\theta d\phi \simeq 1 - \frac{1}{3\alpha}.$$

Per comodità, si consideri uno stato nematico molto ordinato in cui la nostra $f_\alpha(\theta)$ segua l'andamento gaussiano che abbiamo ritrovato.

Il problema della minimizzazione diventa quindi:

$$\frac{d}{d\alpha} \left(\frac{\Delta F[\alpha]}{Nk_B T} \right) = \frac{d}{d\alpha} (\sigma(\alpha) + B(\alpha)c) = 0.$$

Tramite opportune manipolazioni possiamo riscrivere i parametri nematici di modo che

$$\begin{cases} \sigma_n \simeq \ln \alpha - 1 \\ B_n \simeq A\alpha^{-1/2} \end{cases}$$

dove A è una costante. A seguita di questa riscrittura, l'energia libera varrà

$$\frac{\Delta F}{Nk_B T} = -1 - \ln b + \ln c + \ln \alpha - 1 + A\alpha^{-1/2}$$

e la condizione di minimizzazione si traduce in:

$$\frac{1}{\alpha} - \frac{1}{2} A c_n \alpha^{-3/2} = 0.$$

In questa maniera potremo approssimare i parametri nematici come:

$$\begin{cases} \sigma_n \simeq \ln \left(\frac{A^2 c_n^2}{4} \right) - 1 \\ B_n \simeq \frac{2}{c_n} \end{cases}$$

Una volta determinati i parametri σ e B per entrambe le fasi possiamo sfrutare le condizioni di coesistenza sulla pressione e sul potenziale chimico. Questi, sono legati all'energia libera tramite le celeberrime relazioni

$$\begin{cases} P_n = - \left(\frac{\partial F}{\partial V} \right)_{N,T} \\ \mu_n = \left(\frac{\partial F}{\partial N} \right)_{V,T} \end{cases}$$

e quindi, riprendendo la nostra definizione originaria dell'energia libera di Helmholtz ed esplicitando in funzione di N e V , otterremo

$$F = k_B T \left(-N - N \ln b + N \ln \left(b \frac{N}{V} \right) + N\sigma + Bb \frac{N^2}{V} \right).$$

Da cui si ottiene, svolgendo le singole derivate:

$$\begin{cases} P = k_B T \frac{c}{b} (1 + Bc) \\ \mu = k_B T \left(\ln \frac{c}{b} + \sigma + 2Bc \right) \end{cases}$$

e sostituendo i parametri nei due diversi casi

$$\begin{cases} P_{iso} = k_B T c_{iso} (1 + c_{iso}) / b \\ \mu_{iso} = k_B T \left(\ln \frac{c_{iso}}{b} + \sigma + 2c_{iso} \right) \\ P_n = 3k_B T c_n / b \\ \mu_n = k_B T \left(\ln \frac{A^2}{4b} + 3 \ln c_n \right) \end{cases}$$

Ora è possibile risolvere numericamente il sistema delle condizioni di coesistenza utilizzando le incognite c_{iso} e c_n , che sono le concentrazioni di coesistenza. Ciò che si ottiene sono due insieme di soluzioni $\{c_{iso}, c_n, \alpha\}$:

$$\begin{cases} c_{iso,1} = 0.983 \\ c_{n,1} = 0.65 \\ \alpha_1 = 0.538 \end{cases}$$

e

$$\begin{cases} c_{iso,2} = 3.45 \\ c_{n,2} = 5.12 \\ \alpha_2 = 33.40 \end{cases}$$

dove il primo insieme è da scartare, perché la concentrazione della fase isotrope non può essere maggiore della fase nematica. Il valore di α , inoltre, è troppo basso per essere compatibile con una fase nematica (stiamo utilizzando l'approssimazione gaussiana). Il secondo insieme di soluzioni, dunque, è quello che ci fornisce i giusti valori di densità alla coesistenza.

Si deve tener presente che il risultato ottenuto è quello ricavato partendo un'approssimazione ($\alpha >> 1$) della funzione di prova di Onsager. Un procedimento più rigoroso avrebbe richiesto la minimizzazione dell'equazione integrale originaria rispetto alla $f(\hat{\mathbf{u}})$. Uno studio più completo è stato eseguito da Kayser e Raveché [45], i quali hanno dimostrato che per concentrazioni sufficientemente alte esiste una fase nematica. Inoltre Lekkerkerker [46], partendo sempre dall'equazioni integrale, ha ricavato i seguenti valori di coesistenza:

$$\begin{cases} c_{iso} = 3.29 \\ c_n = 4.19 \\ \alpha = 22.36 \end{cases}$$

e confrontando questi valori con quelli ottenuti tramite l'approssimazione gaussiana si nota la somiglianza tra gli uni e gli altri. Dunque, la semplificazione così operata risulta un espediente ragionevolmente corretto. Infine, levando ogni tipo di approssimazione, possiamo usare la funzione di Onsager nella sua forma più completa per ottenere

$$\begin{cases} c_{iso} = 3.34 \\ c_n = 4.49 \\ \alpha = 18.58 \end{cases}$$

che è molto più vicino ai valori veri rispetto alla versione approssimata. Di fat-

to, la funzione di prova di Onsager è un'ottima approssimazione che viene usata abitualmente per descrivere la distribuzione delle orientazioni nell'ambito dello studio di sistemi nematici.

3.1 Passaggi specifici della Teoria di Onsager

Per far scorrere di più l'esposizione della Teoria di Onsager, sono stati omessi alcuni passaggi della discussione. In particolare, abbiamo saltato

- Il procedimento per ricavare l'espressione dell'energia libera di Helmholtz
- La normalizzazione della funzione di prova di Onsager nell'approssimazione gaussiana.
- Infine, il calcolo effettivo dell'entropia orientazionale σ_n e del secondo coefficiente del viriale nematico adimensionale B_n .

Nelle prossime sottosezioni ci concentreremo sugli specifici passaggi elencati [44].

3.1.1 Il procedimento per ricavare l'energia libera di Helmholtz per sistemi di particelle dalla simmetria cilindrica

Nel caso generale, l' energia libera di Helmholtz F in un ensemble canonico è data dall'equazione:

$$F = -k_B T \ln Q,$$

dove Q è la funzione di partizione del nostro sistema, composto da N particelle dai gradi di libertà traslazionali:

$$Q = \frac{1}{N! h^{3N}} \int e^{-\beta H(\{\mathbf{r}\}, \{\mathbf{p}\})} d^N \mathbf{r} d^N \mathbf{p}.$$

dove h è la costante di Planck, \mathbf{r}_i e \mathbf{p}_i sono rispettivamente la posizione e l' impulso della particella i -esima ed H è l'hamiltoniana del sistema. Se quest'ultima si può scrivere come somma dell'energia cinetica K e di quella potenziale U , ovvero se

$$H(\{\mathbf{r}\}, \{\mathbf{p}\}) = K(\{\mathbf{p}\}) + U(\{\mathbf{r}\}),$$

allora la definizione della funzione di partizione si semplificherebbe in modo tale che

$$Q = \frac{1}{N! \lambda^{3N}} \int e^{-\beta U(\{\mathbf{r}\})} d^N \mathbf{r},$$

in cui compare la lunghezza di De Broglie

$$\lambda = \left(\frac{h^2}{2\pi m k_B T} \right)^{1/2}.$$

Sia assunta che l'energia potenziale sia data dalle somme d'interazione di coppia

$$U(\{\mathbf{r}\}) = \sum_{i < j} u_{ij} = \sum_{i < j} u(\mathbf{r}_i, \mathbf{r}_j),$$

allora l'integrandi della funzione di partizione può essere riespresso in termini della funzione di Mayer $\Phi_{ij} = e^{-\beta u_{ij}} - 1$

$$e^{-\beta U(\{\mathbf{r}\})} = \prod_{i < j} e^{-\beta u_{ij}} = \prod_{i < j} +\Phi_{ij}$$

e approssimando al prim'ordine in Φ_{ij}

$$e^{-\beta U(\{\mathbf{r}\})} \simeq 1 + \sum_{ij} \Phi_{ij}.$$

A bassa densità quest'approssimazione è sufficientemente ragionevole. In tal caso, infatti, le particelle sono mediamente molto distanziate tra loro e risentono poco del potenziale d'interazione, che di conseguenza sarà piccolo in valore assoluto per tutte le coppie. Tuttavia, se $u_{ij} \ll 1$ allora $\Phi_{ij} \ll 1$.

Sostituendo nella definizione semplificata della funzione di partizione, otteniamo

$$\begin{aligned} Q &\simeq \frac{1}{N!\lambda^{3N}} \int \left(1 + \sum_{ij} \Phi_{ij} \right) d^N \mathbf{r} = \frac{1}{N!\lambda^{3N}} \int d^N \mathbf{r} + \frac{1}{N!\lambda^{3N}} \sum_{i \neq j} \int \Phi_{ij} d^N \mathbf{r} = \\ &= \frac{V}{N!\lambda^{3N}} \left(1 + \frac{N(N-1)}{2V^2} \int \Phi_{12} d\mathbf{r}_1 d\mathbf{r}_2 \right) \simeq \frac{V}{N!\lambda^{3N}} \left(1 + \frac{\rho^2}{2} \int \Phi_{12} d\mathbf{r}_1 d\mathbf{r}_2 \right), \end{aligned}$$

dove nell'ultimo passaggio si è usata l'approssimazione $N(N-1) \simeq N^2$, valida se $N \gg 1$.

A seguito dell'approssimazione di Q che abbiamo appena visto, l'energia libera di Helmholtz diventa

$$\begin{aligned} \frac{F}{k_B T} &= -\ln \left[\frac{V^N}{N!\lambda^{3N}} \left(1 + \frac{\rho^2}{2} \int \Phi_{12} d\mathbf{r}_1 d\mathbf{r}_2 \right) \right] = -\ln \left(\frac{V^N}{N!\lambda^{3N}} \right) - \ln \left(1 + \frac{\rho^2}{2} \int \Phi_{12} d\mathbf{r}_1 d\mathbf{r}_2 \right) \simeq \\ &\simeq - \left[\ln \left(\frac{V^N}{N!\lambda^{3N}} \right) + \frac{\rho^2}{2} \int \Phi_{12} d\mathbf{r}_1 d\mathbf{r}_2 \right] \end{aligned}$$

e facendo uso dell'approssimazione di Stirling per il fattoriale ($N! \simeq N \ln N - N$) con $N \gg 1$

$$\frac{F}{Nk_B T} = \ln(\rho\lambda^3) - 1 - \frac{\rho}{2V} \int \Phi_{12} d\mathbf{r}_1 d\mathbf{r}_2 = \ln(\rho\lambda^3) - 1 + \rho\beta_2.$$

Qui, il termine $\ln(\rho\lambda^3) - 1$ rappresenta naturalmente il contributo del gas ideale, mentre β_2 è il coefficiente del primo termine del viriale.

Se l'interazione è del corpo rigido, β_2 non è altro che la metà del volume escluso calcolato per una coppia di particelle

$$v_{exc} = -\frac{1}{V} \int \Phi_{12} d\mathbf{r}_1 d\mathbf{r}_2.$$

Adesso il risultato ottenuto può essere generalizzato al caso di particelle dalla simmetria cilindrica. Si devono aggiungere tre gradi di libertà per particella, che identificano l'orientazione: $\hat{\mathbf{u}}_i = (u_{ix}, u_{iy}, u_{iz})$ con $|\hat{\mathbf{u}}_i| = 1$. L'introduzione di questi gradi di libertà cambia sia la parte cinetica sia la parte relativa al potenziale d'interazione. Poiché la parte interessante è proprio quest'ultima, riscriviamo la

densità d'energia libera nel caso puramente traslazionale. Tralasciando la dipendenza da λ , che difatti indica la parte cinetica relativa agli impulsi, ritroviamo

$$\frac{F}{Nk_B T} = \ln \rho - 1 + \rho B_2.$$

Definiamo ora una nuova quantità :

$$\begin{cases} \Delta N_i = N f(\hat{\mathbf{u}}_i) \Delta \Omega_i \\ \sum_i \Delta N_i = N \end{cases}$$

Questa ΔN_i non è altro che il numero di particelle la cui orientazione si trova in un intorno $\Delta \Omega_i$ centrato nella direzione $\hat{\mathbf{u}}_i$. Particelle con diversa orientazione possono essere considerate come se appartenessero a specie chimiche differenti, quindi l'energia libera di Helmholtz può essere riscritta come

$$\frac{F}{k_B T} = \sum_i \Delta N_i \left(1 + \ln \frac{V \Delta \Omega_i}{4\pi \Delta N_i} \right) + \sum_{i,j} \beta_2(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j) \Delta N_i \Delta N_j$$

ed integrando il tutto otteniamo finalmente

$$\frac{F}{Nk_B T} = \ln \rho - 1 - \int f(\hat{\mathbf{u}}) \ln (4\pi f(\hat{\mathbf{u}})) d\Omega + \rho \int \beta_2(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j) d\Omega_1 d\Omega_2,$$

ovvero la F che cercavamo.

3.1.2 Normalizzazione della gaussiana di Onsager

Nel limite in cui il grado orientazionale è molto elevato, cioè $\alpha \rightarrow \infty$, la distribuzione di Onsager può essere approssimata da un andamento gaussiano:

$$f_\alpha(\theta) = \begin{cases} \frac{\alpha}{4\pi} e^{-\alpha \frac{\theta^2}{2}} & \text{se } \theta \in [0, \pi/2] \\ \frac{\alpha}{4\pi} e^{-\alpha \frac{(\pi-\theta)^2}{2}} & \text{se } \theta \in [\pi/2, \pi] \end{cases}$$

dove θ è l'angolo formato dal versore di una particella con l'asse nematico. Si vuole quindi verificare che nel limite considerato la normalizzazione sia corretta e di conseguenza unitaria.

$$\begin{aligned} & \int_0^{2\pi} \int_0^\pi f(\theta) \sin(\theta) d\theta d\phi = 2\pi \int_0^\pi f(\theta) \sin(\theta) d\theta = 4\pi \int_0^{\pi/2} f(\theta) \sin(\theta) d\theta = \\ & = \alpha \int_0^{\pi/2} e^{-\alpha \frac{\theta^2}{2}} \sin(\theta) d\theta \simeq \alpha \int_0^\infty e^{-\alpha \frac{\theta^2}{2}} \sin(\theta) d\theta \simeq \alpha \int_0^\infty e^{-\alpha \frac{\theta^2}{2}} \left(\theta - \frac{\theta^3}{6} \right) d\theta = \\ & = 1 - \frac{3}{\alpha}. \end{aligned}$$

Alla prima approssimazione il dominio è stato esteso all'infinito, poiché l'integrandino è un decadimento esponenziale, mentre la seconda approssimazione è invece legata allo sviluppo di Taylor del seno. Quest'ultima approssimazione è giustificata dal fatto che, a causa della presenza dell'esponenziale, gli angoli che contribuiscono significativamente alla somma sono quelli infinitesimi.

Pertanto, la funzione di Onsager è ben normalizzata nel limite di α grandi.

3.1.3 Entropia orientazionale e secondo coefficiente del viriale in fase nematica

Nell'espressione dell'energia libera di Helmholtz suggerita da Onsager compaiono due parametri che dipendono dalla fase considerata: σ e B . Essi rappresentano rispettivamente l'entropia associata alla distribuzione delle orientazioni e il secondo coefficiente del viriale riscalato rispetto al caso isotropo, che è quello più semplice.

Assumendo che per $\alpha \rightarrow \infty$ la distribuzione di Onsager abbia la forma gaussiana, si calcolare i parametri σ_n e B_n dello stato nematico in funzione di α . Cominciamo dall'entropia orientazionale.

$$\begin{aligned}\sigma_n &= \int_0^{2\pi} \int_0^\pi f(\theta) \ln(4\pi f(\theta)) \sin(\theta) d\theta d\phi = 2\pi \int_0^\pi f(\theta) \ln\left(\alpha e^{-\alpha \frac{\theta^2}{2}}\right) \sin(\theta) d\theta = \\ &= \ln \alpha \left(2\pi \int_0^\pi f(\theta) \sin(\theta) d\theta \right) - \alpha \pi \int_0^\pi f(\theta) \theta^2 \sin(\theta) d\theta = \ln \alpha - \frac{\alpha^2}{2} \int_0^{\pi/2} \theta^2 e^{-\alpha \frac{\theta^2}{2}} \sin(\theta) d\theta \simeq \\ &\simeq \ln \alpha - \frac{\alpha^2}{2} \int_0^\infty \theta^2 e^{-\alpha \frac{\theta^2}{2}} \sin(\theta) d\theta \simeq \ln \alpha - \frac{\alpha^2}{2} \int_0^\infty \theta^2 e^{-\alpha \frac{\theta^2}{2}} \left(\theta - \frac{\theta^3}{6} \right) d\theta = \\ &= \ln \alpha - 1 + \frac{2}{3\alpha}.\end{aligned}$$

quindi per $\alpha \gg 1$ si ha $\sigma_n(\alpha) \simeq \ln \alpha - 1$. Le approssimazioni sono le stesse viste per la normalizzazione della distribuzione di Onsager nel caso gaussiano. Possiamo ora passare al calcolo del coefficiente del viriale.

L'espressione del secondo coefficiente del viriale in fase nematica normalizzato rispetto al caso isotropo è la seguente:

$$B_n = \frac{4}{\pi} \int |\sin \gamma(\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2)| f(\hat{\mathbf{u}}_1) f(\hat{\mathbf{u}}_2) d\Omega_1 d\Omega_2,$$

dove f è la distribuzione di Onsager in approssimazione gaussiana e l'angolo γ è quello compreso tra i due versori $\hat{\mathbf{u}}_1$ ed $\hat{\mathbf{u}}_2$. Esplicitando le variabili d'integrazione si ottiene

$$B_n = \frac{4}{\pi} \int_0^{2\pi} \int_0^\pi \int_0^{2\pi} \int_0^\pi |\sin \gamma(\theta_1, \theta_2)| f(\theta_1) f(\theta_2) \sin(\theta_1) \sin(\theta_2) d\theta_1 d\theta_2 d\phi_1 d\phi_2$$

ed il legame tra γ , θ_1 e θ_2 è fornito dalla relazione

$$\cos \gamma = \cos(\theta_1) \cos(\theta_2) + \sin(\theta_1) \sin(\theta_2) \cos(\phi_1).$$

Dato che le funzioni di Onsager sono dei decadimenti esponenziali, i valori del dominio di θ_1 e θ_2 che danno un contributo rilevante all'integrale sono quelli infinitesimali. Si può dunque focalizzare l'attenzione sugli angoli $\theta_i \ll 1$ rendendo il dominio di interesse quello per cui $\gamma \ll 1$. Le relazioni divengono dunque, rispettivamente

$$B_n = \frac{4}{\pi} \int_0^{2\pi} \int_0^\pi \int_0^{2\pi} \int_0^\pi |\gamma(\theta_1, \theta_2)| f(\theta_1) f(\theta_2) \theta_1 \theta_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2$$

e

$$\gamma^2 = \theta_1^2 + \theta_2^2 - 2 \theta_1 \theta_2 \cos(\phi_1).$$

che unite danno

$$\begin{aligned} B_n &\simeq \frac{4}{\pi} \int_0^{2\pi} \int_0^\pi \int_0^{2\pi} \int_0^\pi \sqrt{\theta_1^2 + \theta_2^2 - 2 \theta_1 \theta_2 \cos(\phi_1)} f(\theta_1) f(\theta_2) \theta_1 \theta_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2 = \\ &= 8 \int_0^{2\pi} \int_0^\pi \int_0^\pi \sqrt{\theta_1^2 + \theta_2^2 - 2 \theta_1 \theta_2 \cos(\phi_1)} f(\theta_1) f(\theta_2) \theta_1 \theta_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2. \end{aligned}$$

Per comodità poniamo $g(\theta_1, \theta_2, \phi_1) = \sqrt{\theta_1^2 + \theta_2^2 - 2 \theta_1 \theta_2 \cos(\phi_1)}$. Estendiamo adesso per motivi già spiegati il dominio degli angoli all'infinito ed esplicitando la funzione di Onsager si arriva alla seguente espressione:

$$B_n \simeq \frac{\alpha^2}{2\pi} \int_0^{2\pi} \int_0^\infty \int_0^\infty g(\theta_1, \theta_2, \phi_1) e^{-\frac{\theta_1^2 + \theta_2^2}{2}} \theta_1 \theta_2 d\theta_1 d\theta_2 d\phi_1 d\phi_2.$$

Ciò che serve di quest'ultima equazione è soltanto la parte dipendente da α . Oltre che nel fattore quadratico iniziale, tale parametro compare anche nell'argomento dell'esponenziale integrando. Il vero problema è proprio questo esponenziale. Per isolare α si può eseguire un opportuno cambio di variabili. La forma delle nuove coordinate è suggerita dalla somma quadratica degli angoli. Seguendo questa traccia, il cambio di variabili è della forma:

$$\begin{cases} \theta_1 = \xi \cos \beta \\ \theta_2 = \xi \sin \beta \end{cases}$$

dove $\xi \in \mathbb{R}^+$ e $\beta \in [0, 2\pi)$.

In tal modo otteniamo

$$\begin{aligned} B_n &= \frac{\alpha^2}{2\pi} \int_0^{2\pi} \int_0^{2\pi} \int_0^\infty \sqrt{\xi^2 - 2\xi^2 \cos \beta \sin \beta \cos \phi_1} e^{-\alpha \frac{\xi^2}{2}} \xi^3 \cos \beta \sin^2 \beta d\beta d\xi d\phi_1 = \\ &= \frac{\alpha^2}{2\pi} \left(\int_0^\infty e^{-\alpha \frac{\xi^2}{2}} \xi^4 d\xi \right) \left(\int_0^{2\pi} \int_0^{2\pi} \sqrt{1 - 2\cos \beta \sin \beta \cos \phi_1} \cos \beta \sin^2 \beta d\beta d\phi_1 \right). \end{aligned}$$

Il primo integrale è facilmente calcolabile:

$$\int_0^\infty e^{-\alpha \frac{\xi^2}{2}} \xi^4 d\xi = a\alpha^{-5/2},$$

dove a è una costante. In conclusione si ha

$$B_n = A\alpha^{-1/2},$$

dove A contiene tutti gli altri fattori numerici.

Capitolo 4

Monte Carlo

Coi metodi Monte Carlo ci si riferisce ad una classe di procedimenti che servono a stimare le probabilità di eventi che presentano difficoltà analitica. Concettualmente il metodo si basa sulla possibilità di eseguire, utilizzando numeri estratti a caso (numeri casuali), un campionamento di una distribuzione di probabilità assegnata, $\pi(x)$; ossia sulla possibilità di generare una sequenza di eventi $\vec{X} = \{x_1, \dots, x_N\}$ distribuiti secondo la $\pi(x)$. In pratica, invece di servirsi di un campione di numeri effettivamente estratti a caso, si ricorre a una sequenza di numeri ottenuti con un processo iterativo ben determinato; tali numeri vengono detti pseudo-casuali giacché, pur non essendo casuali, hanno proprietà statistiche analoghe a quelle dei veri numeri casuali. Benché si fosse già a conoscenza di metodi di campionamento casuale dai tempi dell'illuminismo (la cosiddetta *passeggiata aleatoria* del Conte di Buffon potrebbe costituire il primo esempio di metodo Monte Carlo) si deve la sua formalizzazione agli scienziati legati al progetto Manhattan. In particolare, Nicholas Metropolis non solo diede un nome a questa classe di metodi [47], ma formulò l'algoritmo forse più celebre all'interno di questa serie di metodi numerici: per l'appunto, il *Metropolis Monte Carlo*.

~~Ma andiamo con ordine.~~

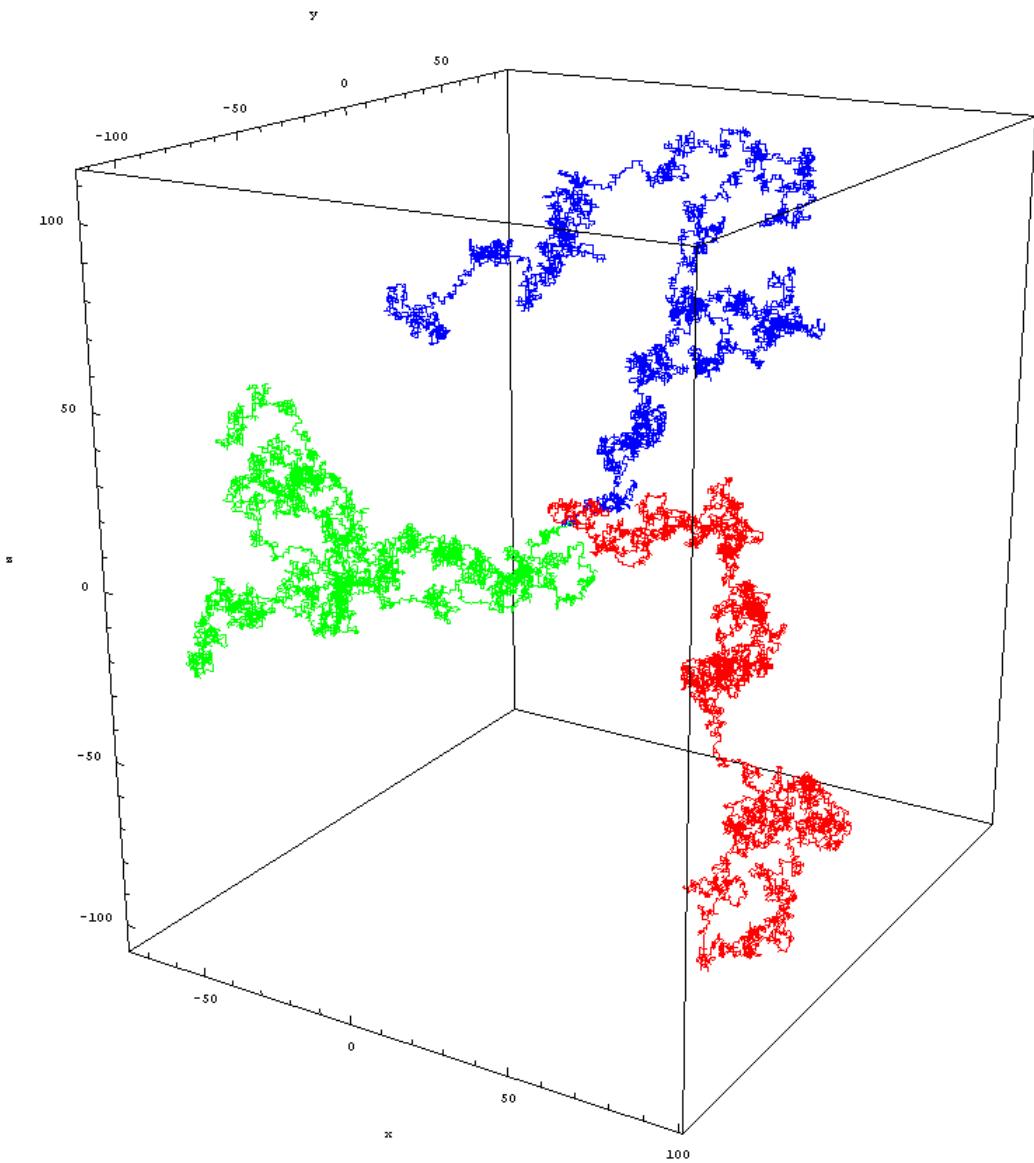


Figura 4.1: Modello di passeggiata aleatoria in tre dimensioni. I metodi Monte Carlo si basano sul campionamento casuale, che porta all'approssimazione di osservabili attraverso un'esplorazione completa lo spazio delle configurazioni.

4.1 Campionamento Casuale

Una delle ragioni per cui i metodi Monte Carlo hanno avuto così tanto successo è la loro versatilità nel calcolo integrale. Difatti, la strategia Monte Carlo per trovare il valore di un integrale non è altro che riscriverlo come una media calcolata su una distribuzione di probabilità a nostra scelta. Per esempio se definiamo una generica [48]

$$\pi(x) = \frac{1}{L}$$

con L pari alla lunghezza dell'intervallo nel quale avviene l'integrazione, non dobbiamo fare altro che soddisfare le proprietà

- $\pi(x) \geq 0$
- $\int_0^L \pi(x) dx = 1$

per ottenere una densità di probabilità utilizzabile nel Monte Carlo. Con questa definizione potremo calcolare il valore medio di un'osservabile rispetto alla distribuzione $\pi(x)$

$$\int_0^L f(x) dx = L \int_0^L f(x) \frac{1}{L} dx = L \int_0^L \pi(x) f(x) dx = L \langle f \rangle_{\pi}.$$

La vera difficoltà è generare i numeri x_i con probabilità $\pi(x)$, ma gli algoritmi Monte Carlo esistono proprio per risolvere questo problema.

A livello informatico, dove si incontra più spesso il Monte Carlo, tutti gli algoritmi si basano sulla routine $RAN()$, che genera uniformemente numeri pseudo-casuali nell'intervallo $[0,1]$. In particolare, questa routine possiede due importanti proprietà:

- Genera numeri x_i tali che la probabilità $p(x)$
 $p(a < x_i < a + \delta) = \delta \quad \forall 0 \leq a < a + \delta \leq 1.$

- Cosa ancora più importante, ci sono bassissime correlazioni tra i diversi x_i ed x_j
 $p(a < x_i < a + \delta \wedge b < x_j < b + \delta) = \delta^2$.

Un ultimo importante commento sulla routine *RAN()* è come la generazione di numeri casuali avvenga in realtà nell'intervallo $[0,1)$ e non $[0,1]$, perciò la generazione $RAN() = 1$ non occorre mai. Ciò è importante in molti contesti, perché nel caso il numero 1 venisse generato il programma potrebbe crashare.

4.2 Importance sampling

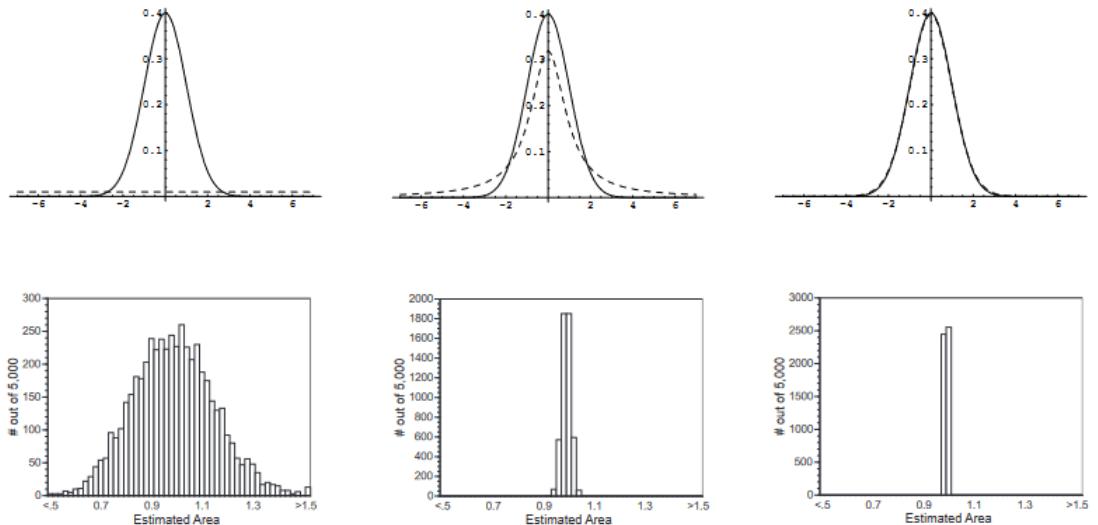


Figura 4.2: Tre diverse funzioni di importance sampling (linee tratteggiate) usate per integrare la densità della distribuzione normale standard (linea continua) da -50 a 50. I pannelli superiori corrispondono alle curve di densità, mentre i pannelli inferiori sono gli istogrammi di 5000 stime Monte Carlo dell'area (che è esattamente 1).

Spesso ci si può ritrovare con una statistica dominata da dei dati che non sono tuttavia caratteristici del sistema in esame. Il Random Sampling Montecarlo utilizza variabili casuali uniformemente distribuite in un intervallo per il calcolo di un

integrale. Tuttavia, in moltissimi casi, la maggior parte del tempo computazionale viene utilizzato per calcolare punti nei quali la funzione integranda è trascurabile. La soluzione consiste nel campionare più punti dove quest'ultima porta ad un contributo importante mentre estrarne di meno da regioni nelle quali si avrebbe un esito trascurabile. Questa soluzione per rimediare alla perdita di tempo computazionale si chiama *importance sampling* [49]

Mentre nel campionamento casuale abbiamo usato una distribuzione uniforme che genera numeri tra 0 ed 1, qui scriviamo una nuova densità di probabilità definita come $\alpha(x)$ tale che

- $\alpha(x) \geq 0$
- $\int \alpha(x) dx = 1$

e così facendo potremo riscrivere il nostro integrale come

$$I = \int_0^L \pi(x)f(x) dx = \int_0^1 \alpha(x) \frac{f(x)}{\alpha(x)} dx$$

e con un cambio di variabili possiamo riscriverlo come

$$I = \int_0^1 \frac{f(x(\epsilon))}{\alpha(x(\epsilon))} d\epsilon.$$

Assumendo che $\alpha(x)$ non è altro che la derivata di una funzione positiva o nulla e non decrescente $\epsilon(x)$ tale che $\epsilon(0) = 0$ e $\epsilon(1) = 1$. In questo modo, generando L variabili casuali di ϵ nell'intervallo $[0, 1]$, otterremo la stima:

$$I \simeq I_L = \frac{1}{L} \sum_{i=1}^L \frac{f[x(\epsilon_i)]}{\alpha[x(\epsilon_i)]}.$$

Il vantaggio di questo approccio si può evincere andando a guardare la varianza di I_L , che è pari a:

$$\sigma_I^2 = \frac{1}{L} [\langle (f/\alpha)^2 \rangle - \langle f/\alpha \rangle^2],$$

Infatti, anche se la varianza ha ancora l’andamento $1/L$, il fattore moltiplicativo può essere ridotto tramite un’opportuna scelta della distribuzione di campionamento $\alpha(x)$. Ponendo $\alpha(x)$ costante si riottiene il Random Sampling Montecarlo, con $\sigma_I^2 = 1/L$. Idealmente, la distribuzione di campionamento $\alpha(x)$ dovrebbe essere un multiplo della $f(x)$ di nostro interesse, in modo tale da ottenere una varianza nulla. Bisogna però prestare molta attenzione alla scelta di $\alpha(x)$, ad esempio che non rappresenta appieno la funzione integranda, può portare a risultati peggiori rispetto al Random Sampling [50].

Con una scelta di $\alpha(x)$ opportuna possiamo dunque campionare nello spazio delle probabilità in un modo tale che rispecchi chiaramente l’andamento della nostra funzione integranda. Purtroppo questo metodo è soggetto a molte limitazioni, derivanti soprattutto dal fatto che non si è a conoscenza di una trasformazione come quella necessaria al funzionamento dell’algoritmo [51].



4.3 Metropolis

Di tutti gli algoritmi Monte Carlo, il più celebre è senza dubbio quello che prende il nome da Nicholas Metropolis. Nel Metropolis Monte Carlo si usano le catene di Markov per generare stati con probabilità proporzionale ad una distribuzione data, realizzando quindi l'importance sampling nello spazio delle configurazioni [55]. Le *catene di Markov* sono una concatenazione di processi casuali indipendenti in cui la *memoria* di un sistema a un qualunque stato dipende soltanto dallo stato precedente. Più formalmente, una *catena di Markov* è definita su di uno spazio delle fasi S e da una matrice di transizione P_{xy} con $x, y \in S$ tale che

$$P_{xy} \geq 0, \quad \sum_{y \in S} P_{xy} = 1.$$

La seconda condizione ci assicura che, per ogni x fissato, P_{xy} è una probabilità in y .

Le caratteristiche base delle catene di Markov sono dunque

- La probabilità di essere in y al tempo $t + 1$ dipende solo dalla posizione x al tempo t .
- Le probabilità sono indipendenti dal tempo.

In pratica, la dinamica procede nel seguente modo:

- al tempo $t = 0$ il sistema è nel punto x_0 .
- P_{x_0y} ad x_0 fissato è una probabilità p_y . Le due condizioni di cui sopra implicano dunque che

$$P_{xy} \geq 0, \quad \sum_{y \in S} P_{xy} = 1.$$

Scegliamo un punto y con una probabilità p_y ed imponiamo $x_1 = y$. A questo punto x_1 è lo stato del sistema al tempo $t = 1$

- Fatto ciò, iteriamo la procedura definendo P_{x_1y} ad x_1 fissato come una probabilità p_y . Si sceglie un punto y con una probabilità p_y ed imponiamo $x_2 = y$. A questo punto x_2 è lo stato del sistema al tempo $t = 2$ e così via...

Pasno ora all'oggetto principale della nostra analisi, il Monte Carlo Metropolis. Ora che conosciamo il formalismo delle catene markoviane, ricordiamoci il nostro obiettivo: data una distribuzione di probabilità π definita su uno spazio delle fasi S , vogliamo determinare una matrice di transizione P tale che π è la sua distribuzione di equilibrio.

L'algoritmo di Metropolis viene suddiviso essenzialmente in due passi:

- Il sistema è attualmente al tempo i , nello stato $x_i = x$. Proponiamo una configurazione $y \neq x$ con una matrice di transizione $P^{(0)}$, quest'ultima è arbitraria e non ha relazioni con π_x .
- Ora stabiliamo se la nuova configurazione y ~~dovrebbe~~ essere accettata o rifiutata. La accettiamo con una probabilità A_{xy} . Se la configurazione y viene accettata, poniamo $x_{i+1} = y$, altrimenti $x_{i+1} = x_i$. La matrice d'accettazione dipende dalla matrice di transizione e dalla distribuzione π_x . La matrice A_{xy} è una probabilità e perciò soddisfa la condizione $A_{xy} \in [0, 1]$.

Possiamo ora calcolare la matrice di transizione del processo markoviano. La probabilità di andare da x ad y è il prodotto di probabilità di proporre y volte l'accettazione della mossa eseguita

$$P_{xy} = P_{xy}^{(0)} A_{xy} \quad ; \quad x \neq y \quad ,$$

mentre la probabilità di rimanere in x è ricavata dalla legge di conservazione della probabilità

$$P_{xx} = 1 - \sum_{y \neq x} P_{xy} = P_{xy}^{(0)} + \sum_{y \neq x} P_{xy}^{(0)} (1 - A_{xy}).$$

I due termini nell'ultima equazione rappresentano la probabilità di non proporre alcun cambiamento ($P_{xy}^{(0)}$) e la somma delle probabilità che la mossa proposta non venga accettata: abbiamo usato

$$1 = \text{sum}_y P_{xy}^{(0)} = P_{xx}^{(0)} + \text{sum}_{y \neq x} P_{xy}^{(0)}.$$

La condizione di ergodicità  del sistema viene soddisfatta quando la probabilità P del nostro processo markoviano soddisfa la condizione di *bilancio dettagliato* [56], ovvero

$$\pi_x P_{xy} = \pi_y P_{yx},$$

che per ogni coppia di stati (x, y) diventa

$$\pi_x P_{xy}^{(0)} A_{xy} = \pi_y P_{yx}^{(0)} A_{yx},$$

dove $P_{xy}^{(0)}$ e $P_{yx}^{(0)}$ sono ergodiche. Si vede subito che quando almeno una tra $P_{xy}^{(0)}$ o $P_{yx}^{(0)}$ è nulla, la condizione è soddisfatta automaticamente. Più interessante è il caso in cui entrambe le probabilità sono positive e la condizione di bilancio dettagliato può essere riscritta come

$$\frac{A_{xy}}{A_{yx}} = \frac{\pi_y P_{yx}^{(0)}}{\pi_x P_{xy}^{(0)}}$$

Introducendo la definizione

$$R_{xy} = \frac{\pi_y P_{yx}^{(0)}}{\pi_x P_{xy}^{(0)}}$$

si vede subito come

$$R_{xy} = \frac{1}{R_{yx}}.$$

Ora, tutto ciò che ci resta da fare, è determinare la matrice d'accettazione A_{xy} in modo tale che soddisfi l'equazione

$$\frac{A_{xy}}{A_{yx}} = R_{xy}.$$

La scelta di Metropolis è definire la matrice d'accettazione come

$$A_{xy} = \min(1, R_{xy}).$$

La scelta di Metropolis non solo soddisfa l'equazione, ma è quella ottimale, infatti ci fornisce la maggiore probabilità di accettazione 

$$A_{xy} = R_{xy}A_{yx} \leq R_{xy},$$

poiché $A_{yx} \leq 1$. Inoltre $A_{xy} \leq 1$ e dunque

$$A_{xy} \leq \min(1, R_{xy}),$$

che dimostra l'ottimalità dell'algoritmo di Metropolis.

4.3.1 Ensemble canonico



Figura 4.3: Illustrazione di un sistema di interesse sospeso in un bagno di calore. Il sistema di interesse è piccolo se comparato al bagno di calore.

All'interno dell'ensemble canonico (NVT), il Metropolis seguendo la distribuzione canonica

$$\pi(\vec{r}_1, \dots, \vec{r}_N) = \frac{1}{Z} e^{-\beta U(\vec{r}_1, \dots, \vec{r}_N)}.$$

Ciò che viene fatto è definire una proposta, il sistema si trova in una configurazione $x = (r_1, \dots, r_N)$ e dobbiamo scegliere una nuova configurazione y . Si procede in questo modo:

- Scegliamo tra N particelle una generica $i = 1 + N \cdot RAN()$.
- Dato Δ parametro da fissare, poniamo $r'_i = (x_i, y_i, z_i)$, con
$$\begin{cases} x_i = r_{ix} + \Delta(RAN() - 0.5) \\ y_i = r_{iy} + \Delta(RAN() - 0.5) \\ z_i = r_{iz} + \Delta(RAN() - 0.5) \end{cases}$$

La nuova configurazione è dunque

$$y = (r_1, \dots, r_{i-1}, r'_i, r_{i+1}, \dots, r_N)$$

In pratica, muoviamo la particella all'interno di un cubo di dimensioni Δ centrato sulla sua vecchia posizioni. La densità di probabilità di compiere mosse è dunque $(\frac{1}{N} \cdot \frac{1}{\Delta^3})$, questo perché scegliamo i tra N particelle diverse all'interno un cubo di dimensioni Δ^3 . La densità di probabilità è perciò indipendente da x e dunque la matrice di transizione è *simmetrica*.

Possiamo dunque scrivere la matrice di accettazione come

$$A_{xy} = \min(1, \frac{\pi_y}{\pi_x}) = \min(1, e^{-\beta(U(y) - U(x))}).$$

Assumendo di osservare il caso di potenziale per interazione a due corpi

$$U(r_1, \dots, r_N) = \sum_{i < j} V(|r_i - r_j|)$$

e vogliamo calcolare $\Delta U = U(y) - U(x)$. Ciò non è un problema, perché nella differenza tutti gli elementi che non contengono i si cancellano, ma resta comunque la questione pratica della scelta di Δ . Questo parametro andrebbe scelto in modo tale da ottenere il più piccolo errore possibile per un certo set di interazioni dell'algoritmo. Esso è dunque un parametro di efficienza e andrebbe selezionato in modo che l'accettazione

$$f = \frac{\text{numero di mosse accettate}}{\text{numero di mosse totali}}$$

ruoti attorno al 30-50%.

4.3.2 Ensemble Isobaro-isotermo

In questo ensemble dobbiamo aggiornare sia le coordinate delle particelle sia il volume, anche se quest'ultimo appare indirettamente all'interno delle coordina-

te \vec{q} . Scriviamo la funzione di partizione dell'ensemble, con un piccolo cambio di variabili che denota come il sistema sia una scatola cubica $V = L^3$ e le coordinate si possano scrivere come $\vec{r} = \vec{s}L$:

$$\Delta(N, P, T) = \frac{3}{\lambda^{3N} N!} \int_0^\infty dL L^{2+3N} e^{-\beta PL^3} \int ds^{3N} e^{\beta U(\vec{s}_1 L, \dots, \vec{s}_N L)}.$$

Quindi dobbiamo effettuare un aggiornamento sulle variabili s_i e sulla dimensione L della scatola.

- Aggiornamento le variabili s_i come nel caso canonico.
- Aggiornamento delle variabili L .

La proposta è dunque

$$\begin{cases} x = L \\ y = L + \Delta(RAN() - 0.5) = L_n \end{cases} \quad \text{che come si può vedere è perfettamente simmetrica.}$$

Possiamo infine scrivere la nostra matrice d'accettazione

$$A_{xy} = \min\left(1, \frac{\pi_y}{\pi_x}\right) = \min\left[1, \left(\frac{L_n}{L}\right)^{2+3N} e^{-\beta P(L_n^3 - L^3)} \cdot e^{-\beta[U(s_1 L_n, \dots, s_N L_n) - U(s_1 L, \dots, s_N L)]}\right]$$

notando che non dipende soltanto dalla differenza di energie.

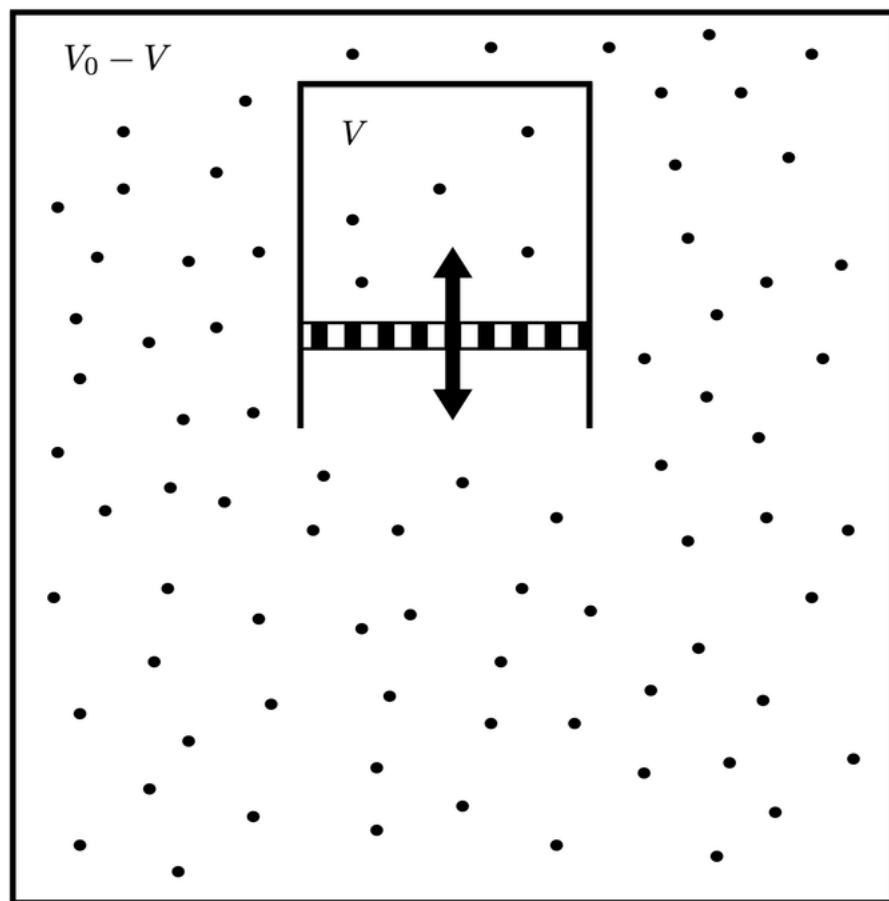


Figura 4.4: Il sistema (di volume V) è immerso in un bagno termico molto più grande a temperatura costante, chiuso in modo tale che il numero di particelle rimane fissato. Il sistema è separato dal bagno termico attraverso un pistone libero di muoversi, tale che il volume può variare.

Capitolo 5

Modellazione Coarse-grained

Per un sistema complesso come ~~una co~~ntenzione di DNA siamo costretti per forza di cose ad utilizzare un modello che approssimi l'oggetto in esame a qualcosa di più semplice. In nostro aiuto viene una serie di modellazioni che prendono il nome di *coarse-grained*, che si occupano esattamente di questo. La filosofia del *coarse-grained* non è altro che la rappresentazione semplificata del comportamento di sistemi complessi, in particolare della loro visualizzazione a diversi livelli di *granularità*. Con *granularità* s'intende la capacità di un sistema di poter essere suddiviso in pezzi distinguibili, caratteristica fondamentale per i sistemi da studiare con questo tipo di modelli. Nel *coarse-grained*, infatti, ciò che facciamo è suddividere un sistema in componenti macroscopiche e discrete da estrarre dal sistema originale. Questa tipologia di modelli è attualmente molto in uso nella biofisica e più specificatamente nella dinamica molecolare, il suo vantaggio in questo campo è la possibilità di sostituire una descrizione a livello atomistico e microscopico di un sistema biologico con un modello che descrive il suo comportamento "medio" e trascura il comportamento "minuzioso". Come si potrà intuire, questo tipo di modello può essere applicato solo sacrificando una descrizione dettagliata dell'oggetto in esame. Il *coarse-grained* funziona, per l'appunto, rimuovendo gradi di libertà al sistema (per esempio visualizzando l'interazione fra due particelle come una singola particella di caratteristiche distinte dalle due originarie) e quindi peggiorando la sua descrizione più particolareggiata, ma resta

comunque un enorme vantaggio: la scelta del numero di gradi di libertà da rimuovere spetta allo sperimentatore ed è quindi completamente nelle sue mani il grado di precisione della misura. Questo è ciò che ha reso il *coarse-grained* un modello di grande successo all'interno della biofisica, visto che permette la descrizione di fenomeni fisici che avvengono per "tempi lunghi", che sono molto importanti in diversi processi biologici.

5.1 Struttura del modello Gapped-Duplex

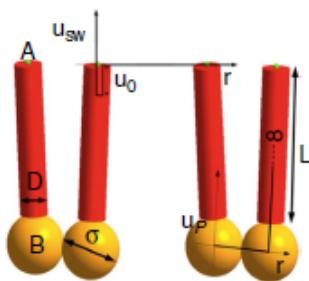


Figura 5.1: Il modello per le molecole G_{20T} duplex usato nelle simulazioni. Le parti rosse sono cilindri rigidi duri. I centri A delle piccole sfere verdi, che modellano l'attrazione tra code dei DNA-Nunchaku, interagiscono attraverso un potenziale *square well* u_{sw} , mentre i centri B delle sfere arancioni grandi appartengono alle gapped-duplex stesse, che modellano la catena flessibile e interagiscono col potenziale u_p . Infine, il diametro delle sfere indica il raggio d'interazione, mentre u_0 (ovvero la profondità del potenziale *square well*) indica l'energia di legame.

Nel nostro modello, abbiamo approssimato la struttura di due filamenti di DNA attaccati tra l'oro da una catena composta da 20 timine come due cilindri ~~attaccati ad un cilindretto più piccolo~~. Questa struttura è stata definita G_{20T} gapped-duplex e nella nostra simulazione ha delle caratteristiche fisiche ben definite. Il G_{20T} gapped-duplex è composto da due filamenti della lunghezza di 16 nm legati da una catena timinica di lunghezza 12,6 nm (assumendo che ogni base azotata abbia una lunghezza di circa 0.63 nm). Da notare come questa catena abbia una

lunghezza molto superiore alla sua *lunghezza di persistenza* pari ad $l_p = 2\text{ nm}$. Questa lunghezza ci permette di caratterizzare la rigidità della catena timinica, essendo definita come la distanza oltre la quale i monomeri che formano un generico polimero non sono più correlati tra loro. Questo vuol dire che se il polimero è più corto della sua l_p allora si comporterà come un'asta rigida. In alternativa, come nel nostro caso, per polimeri molto più grandi della loro l_p i loro pezzi costituenti possono essere descritti statisticamente come in una *passeggiata aleatoria*. Questa proprietà è molto importante, dal momento che vogliamo studiare i nostri G_{20T} attraverso le simulazioni Monte Carlo. Nel nostro modello, in pieno stile *coarse-grained*, rappresentiamo il sistema come due cilindri alle cui basi sono attaccate delle sfere dure. Mentre le dimensioni dei cilindri sono identiche tra loro (una lunghezza di $L = 16\text{ nm}$ ed uno spessore pari a $D = 3\text{ nm}$), le due sfere hanno caratteristiche contraddistinte. Ad un'estremità abbiamo una sfera di diametro $\sigma = 6,3\text{ nm}$, che poggia sul sito A , ad una distanza dal centro di massa del cilindro pari ad $L/2 + 0.15D/2$. Il sito B , invece, che è localizzato lungo l'asse di simmetria ad una distanza di $L/2 + 0.15\sigma/2$ dal centro di massa del cilindro con un diametro $\delta = 0.25D$. Ora, quando due coppie di cilindri sono vicine tra loro, possiamo studiare dei potenziali d'interazione per osservare il comportamento dei siti A e B quando i due sistemi s'influenzano a vicenda. Come ultimo appunto notiamo come il nostro tipo di modello è un sistema di tipo *patchy*, ovvero [57] una classe di colloidì in cui le particelle sono dotate di almeno un sito attrattivo, mediante il quale interagisce anisotropicamente (quindi in maniera direzionale) con l'ambiente circostante. Per cominciare occupiamoci di quando due siti B interagiscono tra loro: definiamo un potenziale

$$u_p \begin{cases} 0 & r < \sigma \\ \infty & \text{altrimenti} \end{cases}$$

dove r è la distanza tra due siti. Il raggio d'interazione σ , ovvero il diametro della sfera associata ai siti attrattivi B , è preso come la metà della lunghezza della catena timinica (da cui i $6,3\text{ nm}$ prima definiti). La lunghezza della catena timinica è stata approssimata dalla lunghezza di ogni base timinica che circa pari a $0,63\text{ nm}$ (da cui i nostri $12,6\text{ nm}$). In questo modo, se i due siti appartengono a

due gapped-duplex diverse, il loro potenziale sarà nullo.

I due siti A vengono invece trattati con un potenziale *square well*

$$\beta u_{sw} \begin{cases} \beta u_0 & r < \delta \\ 0 & altrimenti \end{cases}$$

dove $\delta = 0.25D$ è il raggio d'interazione, ovvero il diametro della sfera associata ai siti attrattivi A .

Le scelte sulla geometria del sistema e sul potenziale d'interazione del sito B sono state effettuate per assicurare la totale flessibilità dei gapped-duplex senza nessun costo energetico associato al loro ripiegamento. Il diametro D , per esempio, è scelto come maggiore del diametro sterico del DNA pari a 2 nm per tenere conto della repulsione elettrostatica tra particelle. I siti A , invece, tengono conto dell'interazione idrofobica tra le punte dei duplex. Il potenziale *square well*, per esempio, è $\beta u_0 = 8.06$ per far sì che l'energia libera sia in linea con le proprietà colesteriche e di fase dei duplex-DNA.

Per studiare la frazione di gapped-duplex foldate nel nostro modello di DNA-Nunchaku G_{20T} , abbiamo introdotto alcuni accorgimenti. In particolare, abbiamo calcolato la frazione di folding

$$\eta_f = \frac{\langle N_f^{45} \rangle}{N},$$

dove $\langle N_f^{45} \rangle$ corrisponde al numero medio di gapped-duplex il cui asse di simmetria forma un angolo di folding $\theta_f < 45^\circ$ (con $\theta_f = 0$ corrispondente ad uno stato *fully folded*), mentre N altro non è che il numero totale di particelle.

Maggiori informazioni su cosa voglia dire *folding* ed *unfolding* nella biofisica seguiranno nel prossimo capitolo.

Capitolo 6

Modelli teorici

Per studiare il nostro modello abbiamo preso in prestito alcuni concetti dalla fisica teorica ed in particolare dalla fisica delle transizioni di fase e dello stato critico. Una parte rilevante del lavoro è stato effettuato dallo studio dei parametri d'ordine, che ci hanno permesso di controllare quando le nostre G_{20T} gapped-duplex si trovavano in una fase o in un'altra. Di particolare interesse sono stati i parametri d'ordine della fase nematica e smettica.

6.1 Parametro d'ordine nematico

La struttura nematica del nostro modello è il fulcro della discussione della nostra tesi. Quando il sistema si trova ancora in uno stato semi-liquido non possiamo sapere con certezza se il suo comportamento è tipico di una fase isotropa o nematica, è per questo che facciamo uso di un parametro d'ordine nematico che descrive con accuratezza lo stato del nostro sistema. La fase nematica è caratterizzata unicamente da un ordine "angolare" delle particelle, dunque ciò di cui ha bisogno il nostro parametro d'ordine non è altro che una buona descrizione dell'orientazione media del sistema [37].

$$Q_{\alpha\beta} = \langle a_\alpha a_\beta \rangle - \frac{1}{3} \delta_{\alpha\beta}.$$

In questa formula, per ciascuna delle N particelle generiche, il versore a corrisponde al sistema di riferimento fisso nel quale si osservano le molecole in posizione $a = (\mathbf{r}, \alpha, \beta) = (x, y, z)$. A questi versori dobbiamo sottrarre un tensore di Kronecker $\delta_{\alpha\beta}$: questo risulta nullo per ogni valore di α e β eccetto quello per cui $\alpha = \beta$, in tal caso $\delta_{\alpha\beta} = 1$. Dato che le parentesi $\langle \cdot \rangle$ indicano una media spazio-temporale, vediamo come nel caso in cui $\delta_{\alpha\beta} = 1$ la media sui versori è pari ad $\langle a^2 \rangle = 1$. Ciò non rappresenta altro che il caso in cui il sistema si trova nella fase isotropa, per la quale il nostro parametro d'ordine nematico risulta nullo. Un altro caso particolare è quello in cui il parametro d'ordine è unitario, in cui il sistema si trova nella fase ~~biassiale~~. Per scoprire in quale fase si trova il sistema dobbiamo dunque diagonalizzare questo tensore. La diagonalizzabilità di Q , però, ci è già garantita dalla sua simmetria e dal fatto che la sua traccia sia nulla. In questo modo, possiamo calcolare gli autovalori, coi loro coi rispettivi autovettori, che rappresentano i singoli parametri d'ordine relativi agli assi orientazionali delle particelle. In particolare, l'autovalore più grande corrisponde, quando esiste, all'asse di allineamento del sistema.

Il parametro d'ordine discussso finora è un tensore simmetrico orientazionale a traccia nulla, che può essere utilizzato in una vasta gamma di sistemi. Quando però ci troviamo in una fase uniassiale, come quella nematica, possiamo riscrivere il nostro parametro d'ordine limitandoci al caso di particelle dalla simmetria cilindrica con libertà rotazionale lungo l'asse nematico:

$$Q_{\alpha\beta} = S_2 (\mathbf{n}_\alpha \mathbf{n}_\beta - \frac{1}{3} \delta_{\alpha\beta}),$$

dove

$$S_2 = \frac{1}{2} \langle 3\cos^2\theta_i - 1 \rangle.$$

Questo è il parametro d'ordine di Hermans [52], che fu introdotto come parametro d'ordine scalare per i cristalli liquidi da Tsvetkov [53]. Scegliendo il direttore \mathbf{n} lungo l'asse z, gli elementi diagonali diventano $S_{xx} = S_{yy} = -\frac{1}{3}S_2$ ed $S_{zz} = \frac{2}{3}S_2$.

Le parentesi del parametro S_2 sono le medie spazio-temporali su un ensemble di molecole dall'asse di simmetria deviante di un angolo θ_i dal direttore **n**. S_2 è il parametro d'ordine scalare che descrive la transizione dalla fase isotropa a quella nematica: $S_2 = 0$ indica un ordine orientazionale casuale (isotropia), mentre $S_2 = 1$ indica molecole perfettamente allineate lungo l'asse nematico.

Abbiamo così un ottimo strumento per riconoscere la fase in cui si trova il nostro sistema.

6.2 Parametro d'ordine smettico

Come abbiamo visto nei capitoli precedenti, la fase smettica è caratterizzata da una morfologia a strati. In particolare, nella fase *SmA*, l'asse direttore **n** e quindi gli assi ottici sono perpendicolari ai piani di strati smettici. Ciò che ci interessa dunque è avere un parametro d'ordine che ci dia informazione sul grado di ordine traslazionale con periodicità d tra i piani smettici. Effettuando la trasformata di Fourier della densità lungo l'asse nematico **n**, otteniamo

$$\hat{\rho}(\mathbf{n}) = \rho(1 + \Gamma \cos(\frac{2\pi\mathbf{n}}{d} + \phi)).$$

All'interno dell'equazione, oltre alla densità ρ del sistema, abbiamo la presenza di tre parametri liberi che ci costringono a trovare la nostra $\hat{\rho}(\mathbf{n})$ attraverso un fit. Questi sono il coefficiente di proporzionalità Γ , che c'informa di quanto risulta smettica la configurazione; ϕ , che fornisce lo sfasamento tra i piani smettici e per finire il sopracitato d , che indica la distanza periodica tra un piano e l'altro.

6.3 Convenienza energetica dello stato *folded*

Osservando il nostro DNA-Nunchaku, notiamo come una parte fondamentale del suo comportamento sia dato dalla fase *Sm-fA*, quella in cui gli elementi del sistema sono per la maggior parte in stato *folded*. Nel mondo della biofisica, lo stato "ripiegato" (*folded*) o "disteso" (*unfolded*) di una proteina, di un enzima o in generale di qualsiasi struttura biologica assume una vitale importanza sulla corretta interpretazione delle sue caratteristiche e soprattutto del suo funzionamento all'interno di un sistema. Per esempio, una cellula può trovarsi in uno stato *folded* o *unfolded* in base alla sua necessità di mantenere una struttura idrofoba o idrofila. Difatti, in base al ripiegamento, la struttura cellulare può subire un certo grado d'influenza dall'ambiente liquido circostante. Nello stato *unfolded*, la membrana cellulare non è in grado di mantenere una struttura solida quando è esposta troppo tempo all'ambiente circostante e corre quindi il rischio di sfaldare l'intero ordine cellulare. Se però una cellula diventa troppo idrofoba, invece, potrebbe denaturarsi e dunque smettere di funzionare.

Nel nostro caso, la presenza di una fase *Sm-fA* non implica l'inesistenza di uno stato *unfolded* delle gapped-duplex (e quindi di una fase *Sm-uA*). Più semplicemente, lo stato *folded* è più termodinamicamente stabile (e perciò energeticamente favorito) rispetto all'*unfolded*. Sfruttando il fatto che le due fasi *Sm-fA* ed *Sm-uA* non effettuano transizioni l'una verso l'altra [58], possiamo ottenere due equazioni di stato in base alla nostra condizione iniziale (cioè o la maggior parte delle gapped-duplex è ripiegata o è distesa). In questo modo possiamo calcolare la differenza d'energia libera di Helmholtz tra le due fasi, punto per punto lungo la simulazione, senza tenere conto di transizioni di fase del prim'ordine. Si vede che quando la fase *Sm-uA* è ancora nella sua metastabilità, la fase *Sm-fA* è già stabile. A causa delle attrazioni idrofobiche sulle estremità dei gapped-duplex, le conformazioni *unfolded* rendono l'ordinamento di tipo smettico sfavorevole, dato che producono la formazione di uno strato di agglomerati semi-flessibili, aumentando perciò il costo entropico necessario alla stratificazione planare. Una conformazione *folded*, al contrario, sopprime la formazione di aggregati e riduce il volume escluso tra le gapped-duplex, rendendo più conveniente la formazione degli strati smettici.

Per calcolare la differenza di energia libera possiamo usare una sorta di coefficiente del viriale. Utilizzando il fatto che la maggior parte del costo energetico smettico proviene dall'effetto di volume escluso, calcoliamo l'energia libera di Helmholtz considerando unicamente gli effetti di volume escluso. Questa sarà pari ad

$$\frac{\beta A_{excl}}{N} = \frac{\beta A_{id}}{N} + \eta(\phi) \frac{\phi}{4v_{hc}} \bar{v}_{excl} .$$

Sono qui dovute una serie di definizioni: come si vede dall'equazione, la nostra scrittura di A_{excl} non è altro che l'energia libera di un gas ideale A_{id} alla quale viene aggiunta una correzione *viriale* dovuta al volume escluso medio \bar{v}_{excl} . Ad esso vi sono associati diversi coefficienti di proporzionalità come il volume di ogni cilindro duro che compone le gapped-duplex v_{hc} e la frazione di volume $\phi = 2\rho v_{hc}$, con $\rho = N/V$ pari alla densità del sistema. Per tenere conto dei fattori di ordine maggiore nell'espansione del viriale, abbiamo infine il *fattore di Parsons-Lee* definito come

$$\eta(\phi) = \frac{4 - 3\phi}{4(1 - \phi)^2} .$$

Assumendo che nella fase *Sm-uA* tutte le gapped-duplex sono *unfolded* e in *Sm-fA* sono tutte *folded*, abbiamo che nello stato disteso $\bar{v}_{excl} = v_{excl}^u$ e in quello ripiegato $\bar{v}_{excl} = v_{excl}^f$ (con v_{excl}^f ed v_{excl}^u i volumi esclusi nella fase smettica rispettivamente delle configurazioni perfettamente *folded* ed *unfolded*). A questo punto, sottraendo il contributo dei gas ideali, abbiamo un'espressione per la differenza d'energia libera.

$$\frac{\beta \Delta A_{excl}}{N} = \eta(\phi) \frac{\phi}{4v_{hc}} (v_{excl}^u - v_{excl}^f) .$$

Nonostante il *fattore di Parsons-Lee* sia considerato piuttosto approssimativo per alte concentrazioni [59], i valori di ΔA_{excl} osservati indicano un contributo fondamentale degli effetti di volume escluso alla stabilità della fase *Sm-fA*.

L'inesistenza di una transizione spontanea da *Sm-fA* ad *Sm-uA* è ascrivibile alla

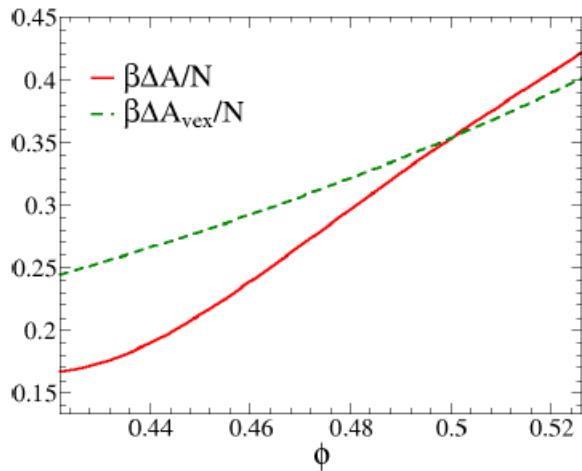


Figura 6.1: Differenze di energia libera di Helmholtz tra le fasi *Sm-fA* ed *Sm-uA* in funzione della pressione a temperatura $T = 0.125$. La linea solida (ΔA) è stimata dall'equazione $A = G - PV$, dove G è l'energia libera di Gibbs. La linea tratteggiata, invece è l'energia libera di Helmholtz (ΔA_{ex}) basata soltanto sugli effetti di volume escluso

loro differenza d'energia libera troppo piccola



Capitolo 7

Risultati delle misure

7.1 Panoramica sul funzionamento delle misure e sulle grandezze analizzate

Utilizzando un cluster mosix, abbiamo potuto osservare l'evoluzione di un sistema *in silico* avviando molteplici simulazioni in contemporanea, ciascuna partendo da condizioni iniziali differenti. Come spesso succede nelle simulazioni di termodinamica, siamo partiti da un blocco "congelato" nella sua posizione iniziale che poi viene "sciolto" nel corso dell'esperimento, per vedere come si evolve via via il sistema (7.3).

Una volta che è passato un tempo sufficiente a far equilibrare il sistema, abbiamo ricavato l'equazione di stato. Per spaziare in un intervallo ragionevole le nostre misure, abbiamo calcolato diverse pressioni e in base ad esse abbiamo calcolato l'energia e la densità del sistema, attraverso le quali abbiamo osservato l'equazione di stato del nostro "cristallo" *in silico*.

Per scrivere queste equazioni di stato abbiamo usato un volume di "riferimento" v_0 , in base al diametro ~~ella catena timonica~~. Lasciando invariata la lunghezza della catena di $16nm$, abbiamo usato la formula

$$v_0 = 16\pi * (1.2)^2,$$

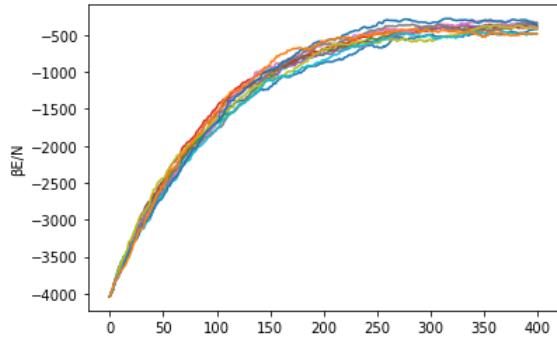


Figura 7.1: Andamento dell’energia adimensionale. Come si può notare dal grafico, il sistema ha termalizzato dopo i nostri venti milioni di passi.

Come si può vedere dalle figure catturate col software *molgl* (7.2), partendo da una configurazione smettica gli elementi sono allineati in strati.

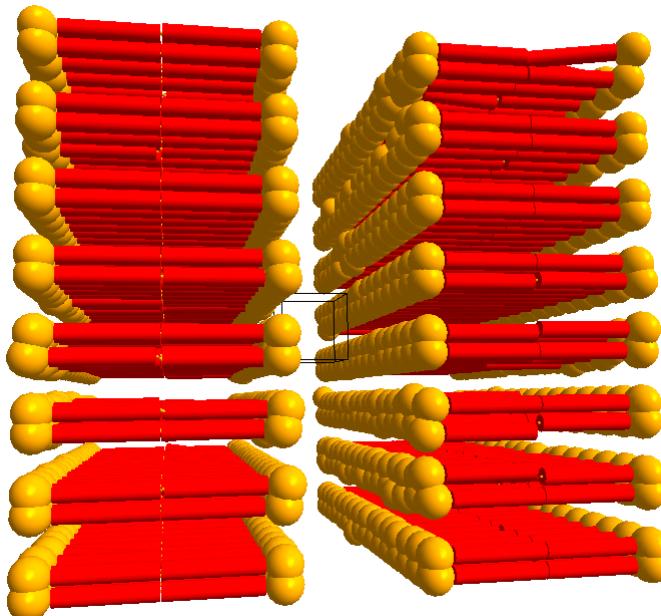


Figura 7.2: configurazione iniziale della simulazione originaria

Dopo aver osservato l’equazione di stato per i primi 500’000 passi delle simulazioni, abbiamo preso la configurazione finale delle misurazioni e l’abbiamo definita come la *nuova* configurazione iniziale e abbiamo continuato le misure per altri

20'000'000 passi, che come si può osservare dal grafico dell’andamento dell’energia sono sufficienti a far equilibrare il sistema (figura 7.1).

Ciò dimostra l’indipendenza dalle dimensioni fisiche della catena timinica delle transizioni di fase isotropo-nematiche e smettiche dei Gapped-Duplex di DNA.

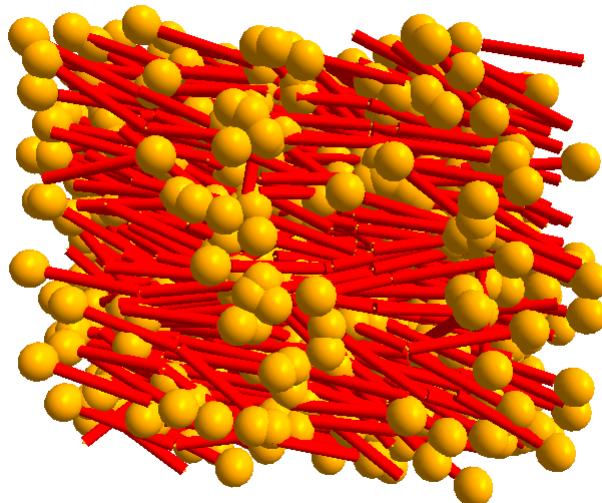


Figura 7.3: Rilevamento della fase isotropo-nematica attraverso il software *MOLGL*. Come si può evincere dalla figura, il sistema tende ad allinearsi lungo un unico asse d’orientazione, ovvero quello *nematico*.

7.2 La ricerca della K perfetta: come la percentuale di folding ed i parametri d'ordine ci aiutano a determinare le fasi

Ora incomincia la parte più interessante delle nostre misure. Dopo esserci appurati del corretto funzionamento del sistema, abbiamo inserito un parametro K (chiamato "*potenziale di bending*"). Questo parametro è costruito in modo tale che, all'aumentare della K , stiamo favorendo un "blocco" sulla fase nematica e più il potenziale di bending è alto più sarà difficile per gli elementi del sistema ripiegarsi (o "foldarsi") ed avere così una fase smettica.

l'obiettivo è riprodurre i dati sperimentali che mostrano sia una fase nematica che smettica, mentre le simulazioni originarie mostravano solo una fase smettica. Per far sì che il sistema replicasse quanto più possibile il comportamento di quello presentato sul *Natural Communications* del 2016, abbiamo cercato la K *perfetta*: ovvero quella lungo la quale si presentano tutte le tre fasi di nostro interesse: la fase isotropa, la fase nematica e la fase smettica.

Queste misure hanno preso la maggior parte del lavoro di tesi, non solo perché abbiamo cercato le K , ma anche l'intervallo di pressioni lungo cui misurare l'equazione di stato. All'aumentare della pressione, infatti, si è riscontrata una tendenza del sistema a "foldare" e quindi a transire direttamente dalla fase isotropa a quella smettica. La nostra K "perfetta" doveva dunque svolgere un duplice compito: anzitutto volevamo che il sistema a tutti gli effetti transisse dalla fase isotropa a quella smettica, di modo tale che vedessimo prima un comportamento *quasi-liquido* e poi far emergere una fase nematica.

Il secondo compito, tuttavia, è stato ancora più arduo del precedente. La fase nematica, nella sua fase più stabile e caratteristica, tende a durare molto di meno rispetto a quella smettica (a meno che non si aumenti in modo spropositato il valore di K). Persino osservando le equazioni di stato, dove la linea continua centrale corrisponde alla fase nematica, sembra quasi di osservare più una *fase metastabile*,

piuttosto che una vera e propria fase stabile del sistema.

Per fortuna ci vengono in soccorso i parametri d'ordine nematico e smettico. Quando il parametro d'ordine nematico è all'incirca 0.4 possiamo essere sicuri del fatto che il sistema ha effettuato una transizione isotropo-nematica, mentre quando il parametro d'ordine smettico è pressoché nullo il sistema si trova nella fase smettica (7.4).

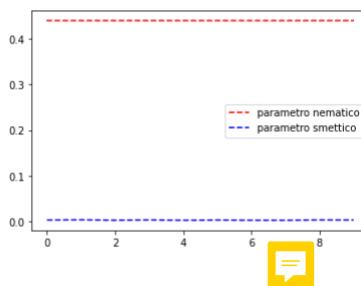


Figura 7.4: Raffigurazione dei parametri d'ordine nematico e smettico. Le linee rappresentano i diversi parametri misurati durante le simulazioni e come si può notare il parametro d'ordine nematico è sempre poco sopra 0.4, mentre quello smettico è quasi nullo.

Nel nostro esperimento, abbiamo simulato il comportamento di 840 *Gapped-Duplexes* di DNA. A seconda di dove ci trovavamo all'interno della simulazione, sia controllando il parametro K sia le misure di pressione, abbiamo osservato che in certe istanze il sistema tendeva a foldare, mentre in altre i costituenti del nostro sistema in esame rimanevano "distesi" (7.6). Come anticipato, abbiamo osservato che all'aumentare del folding il sistema transisce verso una fase smettica. La ragione di questo comportamento è chiara: basta osservare una qualsiasi "fotografia" del sistema in fase smettica e notiamo subito la struttura *stratificata* (7.7) che esso assume. La fase smettica è caratterizzata, come abbiamo anche descritto in precedenza, da "piani" ripiegati che s'impilano l'uno sopra all'altro e la presenza di un grande numero di elementi foldati è quindi un marchio di fabbrica della fase smettica.

Discuteremo, in particolare, dell'equazione di stato che figura in questo capitulo-

lo (7.8). Soffermiamoci su $K = 0.14$: abbiamo una prima "gobba" in cui c'è un esorbitante aumento di pressione. Trattasi non altro che della fase nematica, in cui il sistema sta repentinamente cambiando di configurazione prima di stabilizzarsi lungo la fase smettica. Come si può vedere, infatti, dopo questa gobba ne sussegue un'altra, dopodiché la curva tende ad avere una pendenza fissa e qui abbiamo finalmente la fase smettica. Questo grafico è facilmente comparabile con quello presente sul *Nat.Comm.* del 2016, che apparirà nel capitolo successivo. Fino circa ai 165mg/ml di concentrazione abbiamo un comportamento quasi-liquido, mentre a partire dai 200mg/ml l'effetto della stratificazione si fa sentire lungo tutto il sistema. Eventuali differenze di concentrazione e pressione sono più che altro dovute al diametro ridotto del nostro sistema rispetto a quello originale, che tende ad abbassare i valori misurati rispetto al modello di riferimento a causa di una modifica dell'aspect ratio. Anche la curva raffigurante $K = 0.15$ presenta le due "gobbe", ma sono molto meno pronunciate. Per K maggiori, infine, non assistiamo ad entrambe le transizioni perché il sistema tende a rimanere in un'unica fase e osserviamo delle linee quasi continue.

Se proprio vogliamo descrivere qualitativamente il processo di transizione dalla fase isotropa a quella smettica, possiamo sfruttare le immagini tratte dal software *MolGL* (ovvero, *Molecular Graphic Library*): anzitutto, possiamo vedere come nella fase isotropa ogni orientazione delle *Gapped-Duplexes* sia equivalente ed equiprobabile, qui abbiamo dunque il comportamento più simile a quello di un liquido (7.5). Pian piano che si passa verso la fase nematica, le nostre *Gapped-Duplexes* tendono ad allinearsi lungo una direzione cardinale (che è appunto quella dell'*asse nematico*) e in questa fase incomincia il processo di folding che si fa via via sempre più diffuso tra gli elementi del sistema. Specialmente a grandi valori di K (a partire da $K = 0.18$) possiamo tenere "ferme" le *Gapped-Duplexes* in questa fase ed assistiamo ad una qual certa stabilità di questa transizione ben descritta dalla *Teoria di Onsager* discussa nei capitoli precedenti. Infine, all'aumentare del folding, incominciamo ad assistere alla formazione di più strati impilati l'uno sopra all'altro ed otteniamo finalmente la fase smettica.

Dopo alcuni mesi di misurazioni, possiamo dire di aver trovato un ottimo candidato per il titolo di potenziale perfetto nel valore $K = 0.14$ lungo l'intervallo di pressioni ridotte $\beta Pv_0 \in [1.05, 1.40]$. Oltre alle immagini relative alla presente simulazione, presentiamo in questa tabella un'indicazione orientativa delle transizioni di fase in funzione del numero di *Gapped-Duplexes* foldate (rispetto alle 840 totali):

Fase isotropa	Fase nematica	Fase smettica
<120	130-170	>190

Come si vede dalla tabella, l'intervallo in cui osserviamo la fase nematica è molto più breve rispetto a quello in cui vediamo le altre due fasi di nostro interesse.

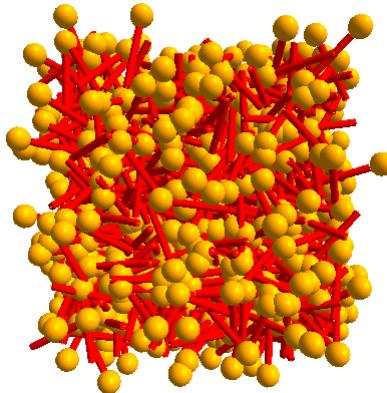


Figura 7.5: Rilevamento della fase isotropa attraverso il software *MOLGL*. Come suggerisce il nome, ogni direzione in questa fase è puramente equiprobabile e casuale: ci troviamo di fronte ad un sistema con il comportamento più vicino a quello di un liquido. In questa foto la pressione è circa $\beta Pv_0 = 1.05$, mentre la frazione di folding è pari a 60/840 *Gapped-Duplexes* foldate.

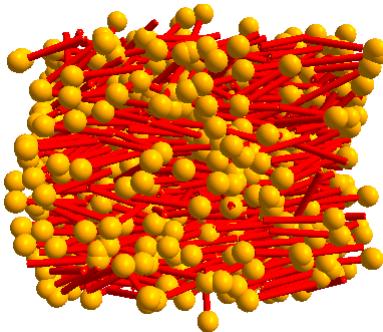


Figura 7.6: Rilevamento della fase nematica attraverso il software *MOLGL*. Possiamo vedere come il sistema tende ad allinearsi lungo un unico asse d'orientazione, ovvero quello *nematico*. Con l'aumento della pressione, inoltre, sta cominciando l'effetto di folding degli elementi costituenti. In questa foto la pressione è circa $\beta Pv_0 = 1.20$, mentre la frazione di folding è pari a 161/840 *Gapped-Duplexes* foldate.

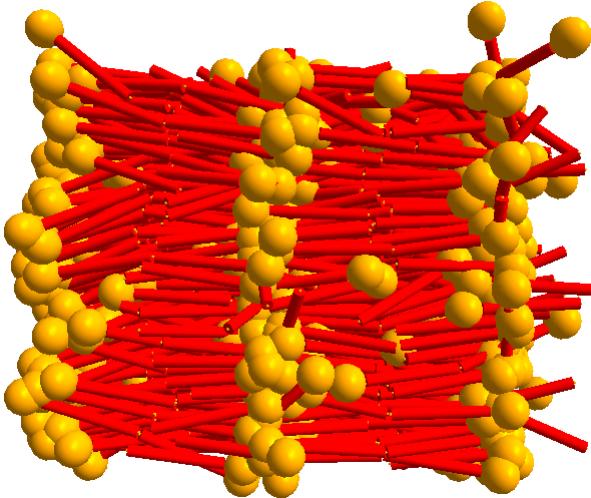


Figura 7.7: Rilevamento della fase smettica attraverso il software *MOLGL*. Qui l'effetto della stratificazione ha configurato finalmente il sistema in modo da creare dei piani impilati l'uno sopra l'altro. In questa foto la pressione è circa $\beta Pv_0 = 1.35$, mentre la frazione di folding è pari a 213/840 *Gapped-Duplexes* foldate.

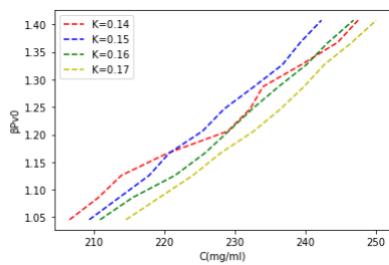


Figura 7.8: Equazione di stato del sistema misurata al variare di K lungo l'intervallo di pressioni ridotte $\beta Pv_0 \in [1.05, 1.40]$.

Capitolo 8

Conclusioni

Dopo attente esaminazioni, abbiamo rilevato una transizione isotropico-nematica ed una fase smettica anche nel sistema di gapped-duplex a diametro ristretto. Nell'immagine (8.1) abbiamo eseguito uno zoom sulle pressioni ridotte in cui appare la transizione isotropo-nematica che c'interessa. Questo è stato anche il principio con cui abbiamo determinato il nostro intervallo di pressioni ridotte su cui effettuare le misure di questa tesi. Se compariamo la nostra equazione di stato con quella nell'esperimento originale possiamo subito vedere le somiglianze col

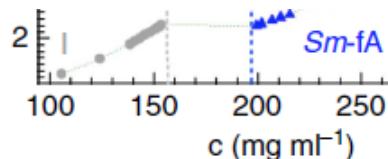


Figura 8.1: Andamento dell'equazione di stato rispetto alla concentrazione (mg/ml) nell'esperimento originale sul quale si basa la maggior parte di questa tesi. Particolare di un'immagine apparsa sul Natural Communications 7 (2016).

nostro grafico. L'unica differenza tra le due strutture è l'aggiunta del potenziale di bending. La transizione di fase resta quindi completamente invariata rispetto all'esperimento originale? In realtà, abbiamo soltanto delle differenze di "proporzionalità". Le nostre transizioni di fase, infatti, avvengono a valori del parametro βPV leggermente inferiori rispetto a quelli del 2016. Se si cambia il diametro cambia anche l'aspect ratio e questo fa sì che si abbiano transizioni a diversi va-

lori della concentrazione.

Già in precedenza, era stato notato come la fase smettica fosse osservabile nei cristalli liquidi di DNA, ma questo studio ci permette di sperare che in futuro la possibilità di studiare lo stato critico dei sistemi in esame sia più agevole rispetto al passato: per osservare queste fasi, infatti, c'è bisogno di creare artificialmente "ad hoc" delle strutture che siano in grado di riprodurre le interazioni e gli ordinamenti che ci interessano. Senza alcun dubbio, la costruzione di tali oggetti diverrà via via sempre più complessa col passare del tempo. Come ultima osservazione si può certo vedere come l'analisi dati necessiti di alcune migliorie, in modo tale da poter caratterizzare meglio il nostro modello.

Capitolo 9

Parentesi sul funzionamento dei programmi informatici

In questo capitolo, ci soffermeremo sugli script che abbiamo utilizzato per il calcolo delle grandezze misurate in questa tesi.

Comincerei da quello con cui abbiamo costruito l'equazione di stato. In questo appendice vorrei, per quanto possibile, mostrare i codici che abbiamo usato e spiegarne il funzionamento riga per riga. Mi sono interessato alla libreria *Pandas*

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as pp
l=[]
x, y = np.loadtxt("rho1.dat", unpack=True, usecols=(0,1))
l.append(x)
for i in range(1,11):
    x, y = np.loadtxt("rho"+str(i)+".dat", unpack=True, usecols=(0,1))
    l.append(y)
mycsv = pd.DataFrame(l)
mycsv.to_csv('my_csv.csv', index=False, header=False)
ciao = pd.read_csv('my_csv.csv')
prova = np.transpose(ciao)
print(prova.mean().sort_values())
Vo = 16*np.pi
B = 8
C = (48*2 +20)*330
BPV = [B*Vo*0.0026, B*Vo*0.0027, B*Vo*0.0028, B*Vo*0.0029, B*Vo*0.0030, B*Vo*0.0031, B*Vo*0.0032, B*Vo*0.0033, B*Vo*0.0034, B*Vo*0.0035]
rho= prova.mean().sort_values()*(48*2 +20)*330
pp.plot(rho, BPV)
pp.xlabel('C(mg/ml)')
pp.ylabel('PV0')
pp.show()
```

Figura 9.1: Programma in *Python3* sulla costruzione dell'equazione di stato. L'ambiente su cui è stato implementato il programma è *Jupyter Notebook*

cercando un metodo per fare statistica sul linguaggio di programmazione *Python*, che da quando l’ho incontrato la prima volta mi è sembrato molto più intuitivo di quelli che avevo studiato in precedenza. Il limite principale di questa libreria è il fatto che manipola soltanto file di tipo *.csv* (ovvero “*comma separated values*”) ed è solo su quest’ultimi che si possono applicare i *metodi* della libreria come l’essenziale *.read()* per leggere i dati e *.DataFrame()* per costruirne tabelle su cui effettuare le operazioni matematiche. Questa necessità dei *.csv* mi ha portato, per la maggior parte dei mesi in cui abbiamo effettuato le misure, a convertire i file originali prodotti dalle misure (ovvero dei semplicissimi *.dat* che si visualizzano tramite un blocco note, contenenti le misure di densità del sistema) in *.csv*. Prima, però, dovevo riempire un foglio di lavoro Excel con i suddetti dati e solo in seguito potevo, misura per misura, convertire il tutto in *.csv* da manipolare e misurare. Questo processo portava via diverse ore, le quali sono via via diventate sempre più difficili da gestire con l’aumentare delle simulazioni.

Qui entra in gioco la prima parte del programma, scritta interamente dal relatore di questa tesi, che si occupa della conversione automatizzata di una serie di file *.dat* in un *DataFrame*, ovvero una tabella di *.csv*.

Attraverso il metodo *.loadtxt()* della famosa libreria matematica *numpy*, estraiamo i dati presenti in ciascun *.dat* prodotto dalle misure. La vera parte automatizzata comincia col seguente ciclo *for*: inseriamo in un’unica cartella i file *rhoi.dat*, dove *i* è un numero che va da 1 a 10 e corrispondente alle misure di densità per un singolo parametro *K*, ordinati per ordine di pressione crescente. I file vengono organizzati in un’unica lista, che viene poi trasformata in tabella attraverso *.DataFrame(l)*, dove *l* è il nome con cui abbiamo identificato la suddetta lista, inizialmente vuota nella prima riga del programma.

A questo punto abbiamo una tabella, che però non può essere manipolata attraverso *Pandas*: ci viene in soccorso il metodo *to*, che converte la tabella in un file *.csv*, tuttavia è molto importante che venga scritto il testo *header = False* o apparirà una prima riga vuota che causerà problemi nei futuri calcoli e grafici.

Restano solo un paio di accorgimenti e il programma potrà calcolare le densità medie su cui disegnare il grafico dell'equazione di stato. Anzitutto, per qualche motivo, ci siamo resi conto che la tabella creata da `.DataFrame()` ha righe e colonne scambiate di posto, quindi dovremo lavorare sulla sua trasposta. Per farlo, basta usare il metodo di `numpy.transpose()`. I conti che verranno attraverso il metodo `.mean()` saranno esattamente le medie delle densità che ci servivano. L'ultimo appunto del calcolo delle medie non è tanto relativo alla computazione delle medie in sé, quanto ad un aggiustamento per ricavare un grafico leggibile. Ciò che grafichiamo sulle ascisse, infatti, non è la lista delle medie così come data in output dal programma, ma ordinata per valore ascendente delle densità. Se lungo il percorso una rho è più grande dell'altra in maniera discontinua, purtroppo, si vengono a creare degli sbalzi nell'equazione di stato che rendono il grafico privo di senso. Questo accorgimento finale appena descritto fa sì che il problema non sussista, potendo finalmente passare alla rappresentazione del grafico tramite il metodo `pyplot()` della celeberrima libreria `matplotlib`.

Sulle ordinate, si può vedere come i valori delle pressioni non siano un semplice numero, ma per ciascun punto ci sia un fattore di conversione $B * Vo$ che rende la nostra pressione adimensionale. $B = 8$ non è altro che la temperatura inversa del sistema, mentre $Vo = 16 * np.pi$ è il volume ridotto (dove `np.pi` è il metodo della libreria `numpy` che richiama il valore di π). Sulle ascisse, le rho sono moltiplicate per il fattore che converte le densità in concentrazioni per unità di mg/ml $C = 330 * (48 * 2 + 20)$, dove 330 è il fattore di conversione in unità *Dalton*, 20 è pari alla lunghezza della catena timinica e 48 sono le coppie di basi, che sono state duplicate per tenere conto che nelle *Gapped Duplexes* abbiamo due filamenti di DNA. Ciò è stato fatto per replicare quanto più possibile il grafico dell'equazione di stato presente sul *Natural Communications*.

Parliamo adesso del programma responsabile del calcolo della frazione di folding, scritto interamente dal prof. Cristiano De Michele

```

#!/usr/local/bin/python3
import numpy as np
import sys
#Usage calc_folding.py <threshold_in_degree>
fn='cnf9'
with open(fn,encoding='utf-8') as f:
    lines=f.readlines()
if len(sys.argv) > 2:
    thr=float(sys.argv[2])
else:
    thr=45.0
lines=lines[1:]
orients=[]
for l in lines:
    ls=l.strip('\n').split()
    orients.append(np.array([float(ls[3]),float(ls[4]),float(ls[5])]))
Np=int(len(orients)/2)
nfolded=0
for i in range(0,Np):
    i1=i*2
    i2=i*2+1
    theta=180.0*np.arccos(np.dot(orients[i1],orients[i2]))/np.pi
    if theta < thr: # if theta is less than 10 degree
        nfolded+=1
print('Number of gapped duplexes is ', Np, ' of which ', nfolded, ' are
folded')

```

Figura 9.2: Programma in *Python3* sul calcolo della frazione di folding

Come si può notare, anche qui sono di vitale importanza i cicli *for*. Lo scopo di questo programma è di contare il numero di *Gapped Duplexes* foldate, per fare ciò viene paragonato il loro angolo di ripiegamento *theta* rispetto ad un valore di soglia *thr* (ovvero *thetarange*), che come sull'articolo del *Nat.Comm* viene impostato attraverso un *if*: ovvero, un'operazione logica che pone *thr* pari a 45° quando non si vede il versore d'orientazione che appare sui file di configurazione generati dalla simulazione (che da qui in avanti chiameremo *cnf*), ovvero quando il *cnf* ha meno di tre argomenti.

L'angolo di folding si ottiene come l'arcocoseno del prodotto scalare delle orientazioni dei due cilindri che costituiscono il duplex, mentre il numero medio di duplex si ottiene calcolando i duplex che hanno un angolo minore di 45° per ogni configurazione e facendone poi la media. Queste operazioni vengono svolte nei due successivi cicli *for* costituenti il corpo centrale del programma. In questi due cicli, oltre al calcolo vero e proprio, vengono effettuati anche dei dimezzamenti dovuti alla natura delle *Gapped Duplexes*. I legami, infatti, sono la metà delle particelle (ovvero i "cilindri" delle simulazioni) e ciò si riscontra nella definizione di *Np* (ovvero il numero di particelle) e nel calcolo dell'arcocoseno, dove gli indici della lista *orients* (cioè le orientazioni) vengono raddoppiati.

Bibliografia

- [1] Liquid crystalline phases given by helical biological polymers (DNA, PBLG and xanthan). Columnar textures F. Livolant et Y. Bouligand J. Phys. France, 47 10 (1986) 1813-1827 DOI: 10.1051/jphys:0198600470100181300
- [2] Strzelecka, T.E., Davidson, M.W. and Rill, R. L. Multiple liquid crystal phases of DNA at high concentrations. Nature 331, 457-460 (1988)
- [3] Smectic phase in suspension of gapped DNA duplexes, Salamonczyk M. et al., Nat. Commun. 7, 13358 (2016)
- [4] Harriott, L. Limits of lithography. Proceedings of the IEEE, 89 (2001), 366. doi:10.1109/5.915379.
- [5] Li, H., Carter, J. D., and LaBean, T. H. Nanofabrication by DNA self-assembly. Materials Today, 12 (2009), 24. doi:10.1016/s1369-7021(09)70157-9.
- [6] R. B. Merrifield, J. Am. Chem. Soc. 85, 2149 (1963)
- [7] Amaldi, F., Benedetti, P., Pesole, G., and Plevani, P. Biologia Molecolare. Casa Editrice Ambrosiana (2014). ISBN 9788808185679.
- [8] Scofield, M. Nucleic Acids. xPharm: The Comprehensive Pharmacology Reference, (2007), 1. doi:10.1016/B978-008055232-3.60059-5.
- [9] Ziad, F. and Agnieszka, S. Clustering algorithms in radiobiology and DNA damage quantification. Data Security, Data Mining and Data Management: Technologies and Challenges (2013).

- [10] Arunan, E., et al. Definition of the hydrogen bond (IUPAC Recommendations 2011). Pure Appl. Chem., 83 (2011), 1637. doi:10.1351/PAC-REC-10-01-02.
- [11] Yakovchuk, P., Protozanova, E., and Frank-kamenetskii, M. D. Basestacking and base-pairing contributions into thermal stability of the DNA double helix. Nucleic Acids Research, 34 (2006), 564. doi:10.1093/nar/gkj454.
- [12] Guckian, K. M., Schweitzer, B. A., Ren, R. X.-F., Sheils, C. J., Tahmassebi, D. C., and Kool, E. T. Factors contributing to aromatic stacking in water: evaluation in the context of DNA. Journal of the American Chemical Society, 122 (2000), 2213. doi:10.1021/ja9934854.
- [13] Seeman, N. C. Structural DNA Nanotechnology. Methods in Molecular Biology, 303 (2007), 143. doi:10.1017/CBO9781139015516.
- [14] Seeman, Nadrian C. (June 2004). "Nanotechnology and the double helix". Scientific American. 290 (6): 64–75. Bibcode:2004SciAm.290f..64S. doi:10.1038/scientificamerican0604-64. PMID 15195395.
- [15] Seeman, N. C. Nucleic acid junctions and lattices. Journal of Theoretical Biology, 99 (1982), 237. doi:10.1016/0022-5193(82)90002-9.
- [16] Seeman, N. C. and Sleiman, H. F. DNA nanotechnology. Nature Reviews Materials, 3 (2017). doi:10.1038/natrevmats.2017.68.
- [17] Lin, C., Liu, Y., Rinker, S., and Yan, H. DNA Tile Based Self-Assembly : Building Complex Nanoarchitectures. ChemPhysChem, 7 (2006), 1641. doi: 10.1002/cphc.200600260.
- [18] Ke, Y., Ong, L. L., Shih, W. M., and Yin, P. Three-Dimensional Structures Self-Assembled from DNA Bricks. Science, 338 (2012), 1177. doi:10.1126/science.1227268.

- [19] Hong, F., Zhang, F., Liu, Y., and Yan, H. DNA origami: Scaffolds for creating higher order structures. *Chemical Reviews*, 117 (2017), 12584. doi:10.1021/acs.chemrev.6b00825.
- [20] Chengde Mao, The Emergence of Complexity: Lessons from DNA, in *PLoS Biology*, vol. 2, n. 12, dicembre 2004, pp. 2036–2038, DOI:10.1371/journal.pbio.0020431
- [21] Saccà, B. and Niemeyer, C. M. DNA Origami : The Art of Folding DNA. *Angew. Chem. Int. Ed.*, 51 (2012), 58. doi:10.1002/anie.201105846.
- [22] Wei, B., Dai, M., and Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature*, 485 (2012), 623. Available from: <http://dx.doi.org/10.1038/nature11075>, doi:10.1038/nature11075.
- [23] Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440 (2006), 297. doi:10.1038/nature04586.
- [24] Shih, W. M., Quispe, J. D., and Joyce, G. F. A 1.7-kilobase singlestranded DNA that folds into a nanoscale octahedron. *Nature*, 427 (2004), 618. doi:10.1038/nature02310.1.
- [25] Andersen, E. S., et al. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature*, 459 (2009), 73. doi:10.1038/nature07971.
- [26] Babatunde, B.; Arias, D.S.; Cagan, J.; Taylor, R.E. Generating DNA Origami Nanostructures through Shape Annealing. *Appl. Sci.* 2021, 11, 2950. <https://doi.org/10.3390/app11072950>
- [27] Seeman, N. C. and Chen, J. Synthesis from DNA of a Molecule with the Connectivity of a Cube. *Nature*, 350 (1991), 631. doi:10.1038/350631a0.
- [28] Douglas, S. M., Dietz, H., Liedl, T., Höglberg, B., Graf, F., and Shih, W. M. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, 459 (2009), 414. doi:10.1038/nature08016.

- [29] De Gennes, P.G. *The Physics of Liquid Cystals*, Clarendon Press Oxford University Press, Oxford New York (1993)
- [30] Arunan, E. et al. Definition of the hydrogen bond (IUPAC Recomendations 2011). *Pure Appl. Chem.*, 83 (2011)
- [31] Fernandez-Nieves, An and Puertas, A.M. *Fluids, Colloids and Soft Materials: an introduction of soft matter physics*. John Wiley and Sons (2016)
- [32] Gompper, G. and Schick, M. *An introduction to Soft Matter*, vol.1: *Polymer Melts and Mixtures*. Whiley VCH Verlag GmbH (2005)
- [33] Hamley, I. W. *Introduction to Soft Matter*. John Wiley and Sons (2007)
- [34] Bellini, T., Cerbino, R. and Zanchetta, G. *DNA-Based Soft Phases*, pp. 225-279. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
- [35] Yakovuchk, P., Protozanova, E. and Frank-Kamenetskii, M.D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acid Researches*, 34 (2006)
- [36] Guckia, K.M., Schweitzer, B.A., Ren, X.-F., Sheils, C.J., Tahmassebi, D.C., and Kool, E.T. Factors contributing to aromatic stacking in water: evaluations in the context of DNA. *Journal of the American Chemical Society*, 122 (2000)
- [37] Textures of liquid Crystals. Ingo Dierkling, WILEY-VCH Verlag GmbH and Co. KGaA, Weinheim, 2003.
- [38] M. Nakata, G.Zanchetta, B.D. Chapman, C.D.Jones, J.O. Cross, R.Pindak, T. Bellini, N.A. Clark, *Science* 318, 1276 (2007)
- [39] T.E. Strzelecka, M.W. Davidson, R.L. Rill, *Nature* 311, 457 (1988)
- [40] F. Liovolant, A.M. Levelut, J.Doucet, J.P. Benoit, *Nature* 339, 724 (1989)

- [41] L. Onsager, The effects of shape on the interaction of colloidal particles, Ann. N.Y. Acad. Sci. 51, 627 (1949).
- [42] W. W. Wood, Early history in computer simulations in statistical mechanics in Molecular dynamics simulation of statistical mechanical systems, Società Italiana di Fisica (eds. G. Ciccotti e W. G. Hoover) (1986)
- [43] T. Odijk, Theory of lyotropic polymer liquid crystals, Macromolecules 19, 2313 (1986).
- [44] M. Proietti, Teoria di Onsager dei cristalli liquidi (2014).
- [45] R. F. Kaiser e H. J. Raveché, Bifurcation in Onsager's model of the isotropic-nematic transition, Phys. Rev. A 17, 2067 (1978).
- [46] H. N. W. Lekkerkerker, P. Coulon, R. van der Haegen e R. Deblieck, On the isotropic liquid crystal phase separation in a solution of rodlike particles of different lengths, J. Chem. Phys. 80, 3427 (1984).
- [47] Metropolis, N. and Ulam, S. The monte carlo Method. Journal of the american statistical association, 44 (1949)
- [48] Frenkel, D. and Smit, B. Understanding Molecular Simulation. Academic Press (2001)
- [49] Eric C. Anderson, "Monte Carlo Methods and Importance Sampling", Lectures Notes for Stats 578C Statistical Genetics, 20 october 1999.
- [50] Tuckerman, M. E. Statistical Mechanics: Theory and Molecular Simulation. Oxford Graduate Texts (2010). ISBN 978-0-19-852526-4.
- [51] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M.N., Teller, A.H. and Teller, E. Equation of state calculations by fast computing machines. The Journal of Chemical Physics, 21 (1953)
- [52] P. H. Hermans, P. Platzek, Kolloid Z., 88, (1939), 68.

- [53] V. Tsvetkov, Acta Physicochim. (USSR), 16, (1942), 132.
- [54] Torrie, J. and Valleau, J.P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella Sampling. Journal of Computational Physics 23 (1977)
- [55] Jayasri, D., Sastry, V.S.S. and Murthy, K.P.N. Wang-Landau Monte Carlo simulation of isotropic-nematic transition in liquid crystals. Physical Review E, 72 (2005)
- [56] De Michele, C. Simulating Hard Rigid Bodies. Journal of Computational Physics, 229 (2010)
- [57] Zhang, Z. and Glotzer, S.C. Self-Assembly of patchy particles. Nano letters, 4(2004)
- [58] Alberto Giacomo Orellana et Cristiano De Michele, Free energy of conformational isomers: The case of Gapped DNA Duplexes, Eur.Phys.J. E (2019) 42:71, 7 June 2019
- [59] C. De Michele, T.Bellini, F.Sciortino, Macromolecules 45, 1090 (2012)