

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE CHIMICHE

CORSO DI LAUREA IN CHIMICA

**Calcolo di spettri vibrazionali IR di basi del DNA con metodo
PMM: caso della Timidina 5' monofosfato**

Relatore: Prof. Stefano Corni

Laureando: Francesco Ciavarella

Matricola: 2042811

Anno accademico 2023/24

Indice

1 Introduzione	1
1.1 Contesto	1
1.2 Obiettivo della tesi	2
2 Metodi	4
2.1 Perturbed Matrix Method.....	4
2.1.1 Fondamenti teorici.....	4
2.1.2 Procedura di calcolo con metodo PMM.....	5
2.1.3 Simulazione e dettagli computazionali	6
2.2 Molecular Dynamics.....	7
2.2.1 Aspetti generali	7
2.2.2 Simulazione e dettagli computazionali	9
2.3 TD-DFT	10
2.3.1 Aspetti generali	10
2.3.2 Calcoli e dettagli computazionali.....	12
3 Risultati	13
3.1 Gas phase.....	13
3.2 Risultati del calcolo PMM.....	14
3.3 Confronto tra dati calcolati e sperimentali della TMP.....	18
3.4 Confronto tra i dati calcolati per la TMP e quelli di MeT	20
3.5 Confronto tra i dati calcolati per la TMP e quelli di ssT10.....	21
4 Conclusioni	23

Capitolo 1

Introduzione

1.1 Contesto

Le funzioni biologiche delle biomolecole, come il DNA, hanno ultimamente suscitato interesse crescente grazie ai progressi nella genetica e proteomica. La ricerca si sta concentrando sulla possibilità di creare biosensori estremamente sensibili utilizzando *aptameri*^[1] di DNA (ovvero acidi nucleici artificiali), sfruttando le loro proprietà di interazione dinamica altamente selettiva e specifica con le proteine. Questo tipo di analisi potrebbe fornire importanti informazioni sulla relazione tra struttura e funzione biologica di biomolecole poco note, ma particolarmente importanti, ad esempio nell'insorgere di malattie quali il morbo di Alzheimer^[2] o il morbo di Parkinson^[3]. L'analisi permetterebbe l'individuazione di cambiamenti conformazionali in proteine target, rese più evidenti dall'interazione con l'aptamero, facilitandone la rilevazione e promuovendo lo sviluppo di nuove tecniche di diagnostica medica^[4] estremamente sensibili.

Questa tesi è legata al progetto europeo *iSenseDNA*^[5], che si inserisce nell'ambito di cui sopra discusso. Il progetto si pone infatti l'obiettivo di comprendere la funzione delle biomolecole durante processi biologici complessi, sviluppando una tecnologia, basata sull'interazione specifica di aptameri con proteine target, per identificare ciò che un cambiamento nella struttura molecolare implica riguardo alla sua funzione.

Tra i diversi modi per poter analizzare una variazione strutturale molecolare, in questo progetto, in accordo con gli altri partner, si è deciso di fare affidamento alla spettroscopia IR; questo perché si può andare a sfruttare la riproducibilità della tecnica e la sua elevata sensibilità che contraddistingue in particolare le molecole biologiche, (quali DNA e proteine), data la presenza di gruppi di atomi che restituiscono segnali molti intensi e specifici.

Solitamente per molecole più complesse si ottengono spettri più congestionati, con picchi sovrapposti e difficili da interpretare. Un'analisi computazionale, focalizzata solamente sull'analisi di alcune bande particolarmente significative, può rivelarsi un ottimo strumento per comprendere meglio lo spettro IR del campione, permettendo di distinguere contributi diversi a zone particolarmente complesse, fornendo inoltre importanti informazioni su come interpretare variazioni spettrali, ricollegandole a modifiche strutturali della molecola in analisi.

Nell'ambito del progetto *iSenseDNA* questo tipo di studi sono stati affidati al partner UNIPD (coordinato dal Prof. Stefano Corni e dalla Dr. Laura Zanetti Polzi, *CNR NANO Modena*). Per il calcolo degli spettri vibrazionali, è stato adottato un metodo ibrido QM/MM (*quantum mechanics/molecular mechanics*), noto come *Perturbed Matrix Method* (PMM), rivelatosi efficace in studi simili^[6].

Tale metodo permette di mantenere la complessità configurazionale dell'intero sistema fornendo una trattazione adeguata dei gradi quantistici di libertà di una porzione specifica del sistema, da trattare esplicitamente a livello elettronico^[6].

Il PMM si basa sull'idea che la parte di sistema trattata a livello classico induca una perturbazione di tipo elettrostatico sul *quantum center* (QC), modificandone le frequenze

vibrazionali e restituendo dunque un valore più corretto di quest'ultime, precedentemente riferite al solo QC in un sistema isolato (per una trattazione più completa si veda la [sezione 2.1](#)).

La fase iniziale del progetto prevede lo studio computazionale, mediante il metodo sopraindicato, delle bande vibrazionali di sistemi semplici, come singoli acidi nucleici o oligomeri degli stessi in soluzioni acquose o analoghe (come D₂O), al fine di ricavare informazioni fondamentali per il calcolo di strutture più complesse.

Questo lavoro di tesi si inserisce in questo ambito specifico.

1.2 Obiettivo della tesi

Il gruppo di lavoro di UNIPD ha già precedentemente analizzato lo spettro vibrazionale associato alla *1-metiltimina* (MeT) (figura 1.1), molecola derivante dalla timina (base azotata pirimidinica naturalmente presente nel DNA) che presenta però un gruppo metile in posizione 1 invece che un H. La MeT può essere considerata un modello per la *timidina* (nucleoside formato dal 2-desossi-D-ribosio e dalla timina), in quanto il gruppo metilico adiacente all'atomo N1 simula il desossiribosio, il quale non dovrebbe influenzare in modo rilevante i modi vibrazionali della singola base azotata in analisi^[7].

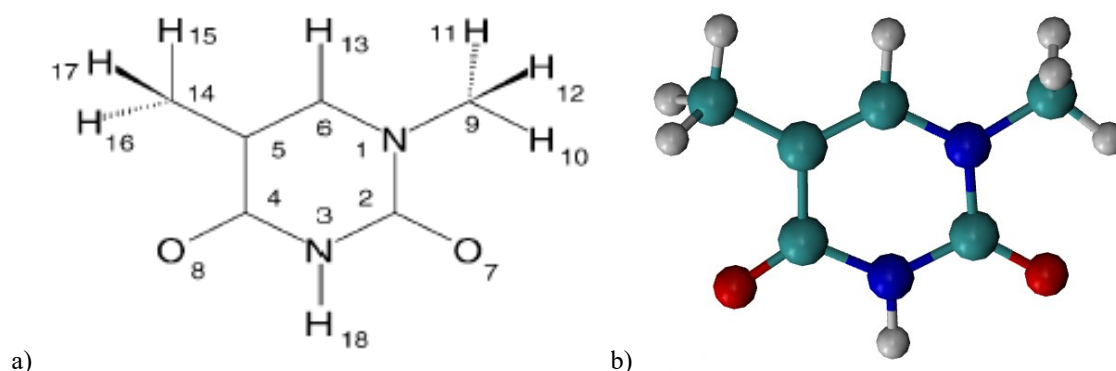


Figura 1.1 a) struttura 2D della *metiltimina*^[7] b) struttura 3D della metiltimina, realizzata con VMD^[8]

In realtà non si calcolano tutte le bande vibrazionali associate all'intera molecola, ma solo le bande associate ad un singolo modo normale di vibrazione della molecola, ovvero quello relativo allo stretching del gruppo carbonilico C=O. Quest'ultimo è stato scelto in quanto, essendo dotato di un momento di dipolo elettrico μ elevato, presenta anche variazioni significative dello stesso (dovute all'allungamento o accorciamento del legame carbonilico) a cui sono direttamente proporzionali le intensità delle bande dello spettro IR. Di conseguenza (come riportato anche in altri studi^[9]) i picchi principali dello spettro della *metiltimina* e di molecole analoghe sono principalmente dovuti allo stretching del gruppo C=O (ed in parte alla deformazione dell'anello pirimidinico o purinico delle basi azotate).

Questo lavoro di tesi ha l'obiettivo di ampliare le informazioni ottenute dagli studi precedenti, andando ad analizzare un sistema simile a quello della MeT ma "più completo", la *timidina 5' monofosfato* (TMP) (figura 1.2), ovvero l'acido nucleico completo, con fosfato annesso. Principalmente tale studio viene eseguito in quanto sono disponibili le informazioni ricavate dallo studio precedentemente condotto sono in realtà

basati sull'approssimazione che la parte di struttura che riguarda lo zucchero ed il fosfato sia ininfluente sui modi vibrazionali della base azotata.

Eseguendo il calcolo con la struttura completa si otterranno dati su un sistema privo di approssimazioni strutturali e si potrà verificare l'efficacia dell'approssimazione precedentemente utilizzata.

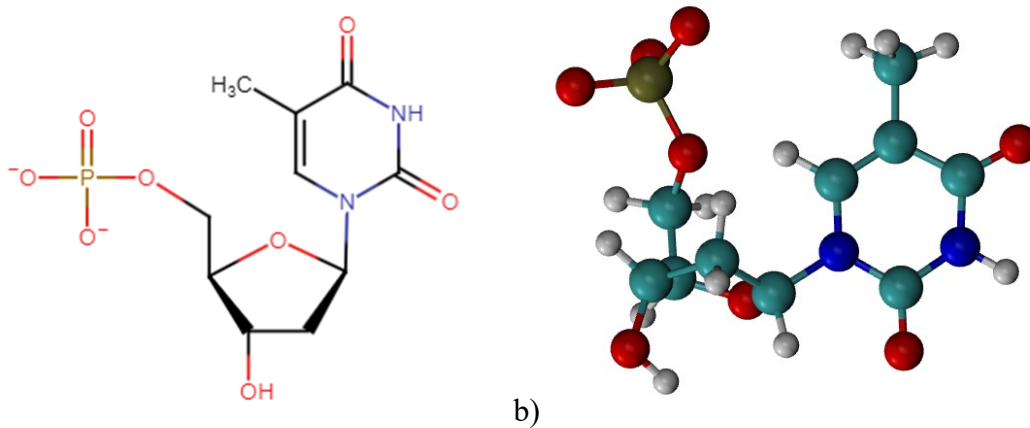


Figura 1.2 a) struttura 2D della TMP b) struttura 3D della TMP realizzata con VMD^[8]

Nello specifico dunque, in questo elaborato di tesi verranno eseguiti calcoli QM/MM con metodo PMM, per la determinazione delle bande vibrazionali associate allo stretching del gruppo C=O della timidina 5' monofosfato; i risultati ottenuti verranno poi confrontati con:

- Spettri IR sperimentali della medesima molecola;
- Spettri IR della *metiltimina*, calcolati con il medesimo metodo;
- Spettro IR di un oligomero *ssT10* composto da 10 nucleotidi aventi come base azotata la timina, calcolato con il medesimo metodo.

Capitolo 2

Metodi

2.1 Perturbed Matrix Method

Per la stesura di questa sezione si è fatto riferimento ai testi [10] e [11].

2.1.1 Fondamenti teorici

Approcci ibridi, che combinano il trattamento classico e quantistico di un sistema complesso, noti anche come *quantum mechanics/molecular mechanics (QM/MM) methods*, sono strumenti ben consolidati e ampiamente utilizzati per l'indagine di diversi processi. In particolare, questi metodi affiancano una descrizione classica del sistema, spesso relegata all'ambiente di reazione o a parti secondarie della molecola in analisi, ad una dettagliata descrizione quantistica di una porzione limitata di quest'ultimo, il *quantum center (QC)*.

Il *Perturbed Matrix Method (PMM)* rientra in questa classe di metodi computazionali; esso può essere utilizzato per simulazioni di osservabili sperimentali, come ad esempio: potenziali di riduzione, costanti cinetiche e, nel caso specifico del progetto iSenseDNA, segnali spettroscopici. Questo metodo ha dunque come obiettivo principale quello di mantenere la complessità configurazionale del sistema, con un appropriato trattamento dei gradi di libertà quantistici di una porzione dello stesso, trattata a livello quantomeccanico; un tale approccio permette l'analisi di sistemi estesi utilizzando di un numero molto elevato di configurazioni atomiche dello stesso, tipicamente fornite da simulazioni di dinamica molecolare.

Il vantaggio principale di questo metodo è la capacità di fornire un enorme campionamento statistico delle proprietà quantistiche perturbate, considerando un ambiente atomico-molecolare di grandi dimensioni, con costi computazionali contenuti.

L'idea di base del PMM è considerare il QC come incorporato in un campo esterno, generato dal resto del sistema, il quale esercita una perturbazione elettrostatica sugli stati elettronici del primo. Tale condizione può essere rigorosamente derivata^[12], all'interno di condizioni e presupposti generali, tramite un'approssimazione accurata dell'operatore Hamiltoniano perturbato \hat{H} e degli autostati corrispondenti.

Il termine perturbativo dell'Hamiltoniano viene dunque espresso come un operatore perturbazione, funzione del campo elettrico $\vec{\mathbf{E}}$ associato alla parte di sistema non quantistica, al quale corrisponde un potenziale elettrico V , che agisce sul QC.

L'espressione analitica dell'Hamiltoniano perturbato è dunque:

$$\hat{H} \cong \hat{H}^0 + \hat{\mathbf{I}}q_T V + \hat{\mathbf{Z}} + \Delta V \hat{\mathbf{I}} \quad \text{eq. 2.1}$$

$$[\hat{\mathbf{Z}}]_{k,k'} = -\vec{\mathbf{E}} \cdot \langle \Phi_k^0 | \hat{\boldsymbol{\mu}} | \Phi_{k'}^0 \rangle \quad \text{eq. 2.2}$$

dove q_T rappresenta la carica totale del QC, $\hat{\boldsymbol{\mu}}$ è l'operatore momento di dipolo, $\hat{\mathbf{I}}$ invece è la matrice identità, ΔV approssima eventuali termini di ordine più elevato rispetto ad un semplice potenziale a corto raggio; nell'eq. 2.2 le parentesi angolari (*notazione di Dirac*) indicano una integrazione lungo coordinate elettroniche. Per quanto riguarda invece \hat{H}^0 , esso rappresenta l'hamiltoniano imperturbato, riferito dunque al solo QC

isolato in assenza del campo elettrico esterno indotto dal sistema; rappresenta dunque il punto di partenza del calcolo da eseguire. Risolvendo l'eq. 2.3, ovvero diagonalizzando \hat{H} si ottengono un set di autovettori Φ_k^0 e autovalori U rappresentanti gli autostati e le energie del QC nel sistema completo.

$$\hat{H}\Phi_k^0 = U\Phi_k^0 \quad \text{eq. 2.3}$$

Quando si trattano perturbazioni di piccola entità, può essere introdotta un'approssimazione: il modo vibrazionale del QC in analisi non accoppia con altri modi vibrazionali dello stesso; ciò implica che la perturbazione non modifichi il modo vibrazionale in modo significativo, ma solo gli autovalori della matrice hessiana.

2.1.2 Procedura di calcolo con metodo PMM

Per determinare stati vibrazionali perturbati viene utilizzata la seguente procedura:

- Vengono eseguiti calcoli quanto meccanici sul QC isolato in fase gas e dunque imperturbato, per ottenere autovalori ed autostati riferiti ad \hat{H}^0 , da cui poter ottenere gli autovalori della matrice delle derivate seconde dell'energia rispetto alle coordinate nucleari del sistema non perturbato, ponderati sulla massa del QC (*mass-weighted Hessian matrix*).
- Dopo aver selezionato il modo di vibrazione di interesse, vengono generate una serie di configurazioni lungo la coordinata q ; per ognuna di queste configurazioni viene calcolato, mediante procedure QM standard, un set di autofunzioni elettroniche ortonormali, ognuna delle quali rappresenta uno stato imperturbato.
- Viene eseguita una simulazione di dinamica molecolare (MD) sull'intero sistema (QC più parte trattata a livello classico), per ottenere una traiettoria dalla quale poter determinare la componente perturbativa di \hat{H} .
- Successivamente tramite la risoluzione dell' eq. 2.3 per ogni configurazione generata, sfruttando le eq. 2.1 e 2.2 ed implementando il campo elettrico perturbativo derivato dalla simulazione di MD, si calcolano gli autostati e gli autovalori dell'Hamiltoniano perturbato ad ogni frame della traiettoria ottenuta. L'energia elettronica, ovvero l'autovalore della matrice rappresentativa dell'Hamiltoniano perturbato \hat{H} , lungo la coordinata vibrazionale q , può essere espressa ad ogni frame della simulazione MD, come:

$$\varepsilon'_0(q, t) = \varepsilon_0(q, t) + q_T V(t) + \Delta V(q, t) \quad \text{eq. 2.4}$$

dove ε_0 è l'autovalore di stato fondamentale della matrice $\hat{H}^0 + \hat{Z}$.

- Infine, viene determinata la frequenza di vibrazione armonica $\left(\frac{\omega}{2\pi}\right)$ del sistema perturbato, calcolata ad ogni istante della simulazione mediante un fitting polinomiale della curva di energia, fornendo la derivata seconda dell'energia nella posizione di minimo.

2.1.3 Simulazione e dettagli computazionali

Nell'ambito specifico di questo lavoro di tesi, per determinare lo spettro vibrazionale IR della *timidina 5' monofosfato* sono stati eseguiti calcoli di tipo QM con metodo *TD-DFT*, livello di teoria B3LYP/6-31G+(d), andando ad utilizzare come QC la *metiltimina* (sezione 2.3).

Da questo calcolo sono state estratte in totale 51 diverse configurazioni relative al modo vibrazionale associato al gruppo C=O, nello specifico 25 configurazioni lungo il versante positivo della coordinata vibrazionale q e altre 25 per il versante negativo, più la configurazione iniziale. Per ognuna di queste è stato poi possibile calcolare gli autostati imperturbati del sistema, necessari per il calcolo PMM.

Per quanto riguarda la determinazione della perturbazione elettrostatica, è stata eseguita una simulazione di MD su un sistema composto da *timidina 5' monofosfato*, posta all'interno di una box riempita utilizzando come solvente H₂O (sezione 2.2).

Prima di poter eseguire il calcolo vero e proprio che porta all'ottenimento delle frequenze del sistema perturbato è stato necessario "unire" i dati ottenuti dal calcolo QM con quelli del calcolo di MD; il problema riguarda il fatto che essi sono stati eseguiti su sistemi diversi, ovvero il primo è stato eseguito sulla *metiltimina* (molecola utilizzata per rappresentare il QC), mentre il secondo ha usato la *timidina 5' monofosfato*. Di conseguenza è stato necessario far coincidere gli indici del QC con quelli della struttura completa della molecola in analisi e ricalcolare la carica del QC e della parte di acido nucleico rimanente, per poter poi andare a neutralizzare il sistema.

A questo punto è stato possibile costruire e successivamente diagonalizzare la matrice rappresentativa dell'Hamiltoniano del sistema perturbato \hat{H} , calcolando infine la banda vibrazionale, con la sua frequenza caratteristica.

2.2 Molecular Dynamics

2.2.1 Aspetti generali

La *Dinamica Molecolare* (MD) è un tipo di analisi computazionale che rientra a far parte dei metodi classici. Questi vengono utilizzati per lo studio di sistemi molto estesi, con un numero di particelle molto elevato, non trattabili con metodi quantomeccanici, in quanto troppo costosi computazionalmente.

In generale, nei metodi classici non si parla di atomi con la solita definizione quantistica, ma di *atom types*, ovvero di punti materiali dotati di dimensioni e cariche specifiche per simulare atomi di elementi diversi o che si trovino in intorni chimici diversi. Questi punti vengono rappresentati secondo il modello *balls and springs* (figura 2.1), ovvero collegati tra di loro da “molle” che vanno a simulare il legame chimico tra due atomi (basandosi su una approssimazione armonica per poterlo rappresentare); ciò consente la simulazione dei moti vibrazionali, torsionali e delle deformazioni molecolare in generale.

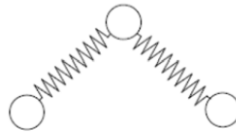


Figura 2.1 Modello *Balls and Springs*

È inoltre necessario definire come le particelle interagiscono tra di loro e quali tipi di energia potenziale governano il loro comportamento; ciò viene fatto mediante l'utilizzo di un *Force Field* (FF), ovvero un modello computazionale costituito da un insieme di funzioni matematiche e parametri utilizzati per descrivere le forze che agiscono sugli atomi e le molecole in un sistema simulato. Esistono diversi tipi di FF e la scelta di quello più adatto è fondamentale per ottenere una simulazione più corretta possibile del sistema in analisi, calcolando dunque correttamente le proprietà richieste.

Le simulazioni di MD restituiscono come output una cosiddetta *traiettoria*, ovvero un insieme di singole “istantanee”, dette *snapshot*, raccolte ad intervalli temporali specifici e caratterizzate da posizioni e velocità diverse dei singoli atom types; la successione degli snapshot mostra l'evoluzione temporale naturale del sistema e consente la previsione delle proprietà statiche e dinamiche del sistema stesso direttamente dalle interazioni tra gli atom types, descritte dal FF.

Le configurazioni successive vengono generate attraverso l'integrazione numerica di equazioni del moto per ogni atomo *i-esimo*; si possono utilizzare diverse equazioni del moto, solitamente quelle di Newton:

$$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i} \quad \text{Eq. 2.5}$$

La forza dell'*i-esima* particella F_i viene espressa come gradiente dell'energia potenziale U rispetto alla posizione r_i (x_i, y_i, z_i):

$$F_i = -\Delta_{r_i} U(r_1, r_2, \dots, r_N) \quad \text{Eq. 2.6}$$

andando ad esplicitare la componente *i-esima* della forza lungo la coordinata k :

$$F_{i,k} = -\frac{\partial U}{\partial r_{i,k}} \quad \text{Eq. 2.7}$$

con $r_{i,k}$ che indica la componente *k-esima* della particella *i-esima*

La risoluzione delle equazioni di Newton avviene mediante utilizzo di algoritmi implementativi specifici.

Il risultato di una simulazione di MD è una *traiettoria*, ovvero un insieme di “snapshot” (configurazioni calcolate ad uno specifico tempo t della simulazione) che si susseguono nel tempo. Da questa traiettoria è possibile ricavare le proprietà del sistema come una media aritmetica su tutte le configurazioni del sistema esplorate durante la simulazione, ottenendo così delle *variabili dinamiche*, che possono essere funzioni della velocità e/o del tempo.

Un altro parametro molto importante è la *funzione di distribuzione radiale* $g(r)$, ovvero una funzione della distanza che descrive la probabilità di trovare una particella a una certa distanza r da un'altra particella scelta come riferimento. Solitamente la $g(r)$ viene calcolata su tutta la durata della simulazione e rappresenta quindi la distribuzione media nel tempo delle specie in analisi e non una distribuzione istantanea.

A questa funzione è legato anche il *numero di coordinazione cumulativo* (CN), il quale, in una simulazione molecolare, rappresenta la quantità totale di particelle che si trovano entro una certa distanza r da una particella di riferimento.

Fondamentale infine risulta essere l'utilizzo di una struttura iniziale da cui poter far partire la simulazione, ricavata tramite metodi di analisi strutturale specifici, come la cristallografia a raggi X, oppure NMR o ancora tramite calcoli quanto-meccanici. La struttura che si ottiene è però riferita ad uno stato non naturale della molecola, che potrebbe presentarsi con differenze strutturali in una soluzione acquosa, a temperatura e pressione ambiente. Prima di ogni simulazione risulta necessario eseguire una serie di procedure volte a aggiornare la struttura iniziale a quella effettivamente presente in tali condizioni.



Figura 2.2 Schema degli step da eseguire per una simulazione di MD^[13]

In generale dunque si segue lo schema riportato in figura 2.1:

- Si genera una *box* con forma e dimensioni specifiche, ovvero uno spazio attorno alla molecola in analisi “riempito” con il solvente. Risulta inoltre importante definire delle *periodic boundary conditions*, ovvero regole che determinano come le particelle interagiscono con il resto del sistema in prossimità dei confini della *box*.
- Si aggiungono, sostituendoli a molecole di solvente, ioni per controbilanciare la carica della molecola in analisi, neutralizzando il sistema.
- Viene eseguita una breve simulazione in cui si minimizza l'energia del sistema complessivo appena creato.
- Vengono eseguite due fasi di equilibrizzazione, in funzione delle condizioni in cui si vuol far avvenire la simulazione:
 - *NVT*, in cui viene raggiunta la temperatura di simulazione desiderata e stabilizzata la struttura del sistema
 - *NPT*, in cui viene raggiunta anche la pressione di simulazione desiderata.
- Infine, viene eseguita la simulazione di dinamica molecolare vera e propria.

2.2.2 Simulazione e dettagli computazionali

Per quanto riguarda questo lavoro di tesi, la simulazione è stata condotta utilizzando il force field CHARMM36^[14], andando ad utilizzare come algoritmo implementativo l'*algoritmo di Verlet*^[15], combinato con l'utilizzo dell'algoritmo *LINCS*^[16] per vincolare le posizioni degli atomi di H. Per svolgere il calcolo è stato utilizzato il programma GROMACS^[17] nella versione 2024.2. L'input della simulazione, ovvero la struttura molecolare della TMP, è stata ottenuta mediante l'utilizzo del software Avogadro^[18].

È stata generata una *box* cubica, di dimensioni 3.50082 nm, riempita con 1364 molecole di H₂O, utilizzando un modello esplicito, ovvero il TIP3P^[19]; sono poi state sostituite due molecole di acqua con due ioni Na⁺, per portare ad elettroneutralità il sistema.

Dopo aver eseguito la minimizzazione di energia, sono state eseguite due fasi di equilibrizzazione:

- equilibrizzazione *NVT*, per raggiungere e mantenere costante la temperatura di 298.15K, attraverso il *canonical velocity-rescale algorithm*^[20]; per questa operazione è stata condotta una simulazione di 200 ps, con *timestep* di 1 fs.
- equilibrazione *NPT* per raggiungere e mantenere la pressione di 1 bar, attraverso un *barostato di Parrinello-Rahman*^[21]; per questa operazione è stata condotta una simulazione di 500 ps, con *timestep* di 2 fs.

Infine, è stata condotta la simulazione di MD vera e propria, che ha avuto una durata di 100 ns, con un *timestep* di 2 fs, andando a raccogliere dunque 5×10^7 *snapshot* che hanno costituito la traiettoria finale.

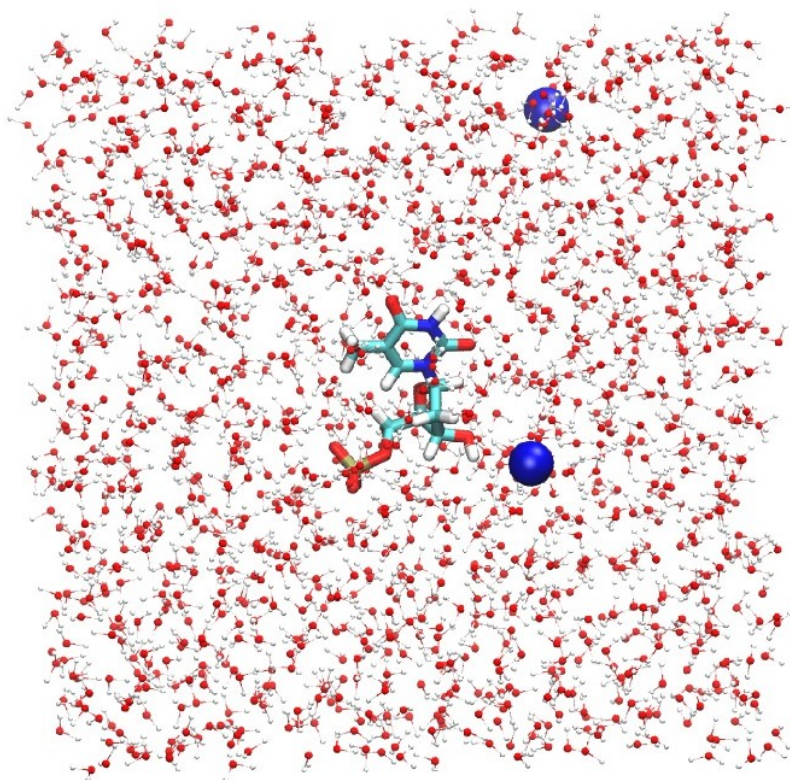


Figura 2.3 snapshot finale della simulazione di MD condotta sulla TMP in H₂O, ottenuto tramite VMD^[8].

2.3 TD-DFT

Questa sezione è stata sviluppata sulla base della review [22] e del testo [23].

2.3.1 Aspetti generali

Il metodo TD-DFT (*Time-Dependent Density Functional Theory*) è una tecnica computazionale quanto meccanica, utilizzata per studiare le proprietà elettroniche e ottiche dei sistemi quantistici, come molecole e materiali solidi. Si basa sull'estensione della *Teoria del funzionale della densità* (DFT), basata sull'idea di Hohenberg e Kohn, che afferma che per poter descrivere un sistema elettronico stazionario sia sufficiente conoscere la sua densità elettronica^[24]. Di conseguenza, tutti gli osservabili di tale sistema, anche la funzione d'onda, possono essere ricavati come funzionali della densità elettronica, così descritta:

$$\rho(r) = N \int |(\psi(r, r_2, \dots, r_N))|^2 dr_2 \dots dr_N \quad \text{Eq. 2.9}$$

Dove $\rho(r) dr$ rappresenta la probabilità di trovare un qualsivoglia elettrone nell'elemento di volume dr centrato nel punto r dello spazio.

L'energia del sistema, definita in funzione della densità elettronica, può essere scomposta in tre termini:

$$E_{DFT} = E[\rho(r)] = T[\rho(r)] + V_{Ne}[\rho(r)] + V_{ee}[\rho(r)] \quad \text{Eq. 2.10}$$

con T funzionale dell'energia cinetica, $V_{Ne}[\rho(r)]$ energia potenziale dell'interazione tra nuclei ed elettroni e $V_{ee}[\rho(r)]$ energia potenziale dell'interazione inter-elettronica.

$E[\rho(r)]$ presenta un minimo per la densità elettronica "vera" del sistema elettronico, posto che $\rho(r)$ sia ottenibile dall'integrazione di una funzione d'onda. La forma del funzionale esatto è sconosciuta, ma ne esistono varie approssimazioni che riguardano in particolare il funzionale dell'energia cinetica e l'interazione elettrone-elettrone.

L'implementazione più comune ed efficiente della DFT deriva dal *metodo di Kohn and Sham*^[25], che si basa essenzialmente sullo studio di un *sistema non interagente*, ovvero in cui non vi sono interazioni elettroniche specifiche, ma un potenziale effettivo che ne tiene conto, il quale ha la caratteristica di avere la medesima densità elettronica del sistema reale.

La *DFT* è in grado solamente di descrivere lo stato fondamentale di un sistema elettronico; per poter fornire una descrizione accurata degli stati elettronici eccitati, consentendo di calcolare le energie di transizione e gli spettri di assorbimento, è stato necessario includere anche l'evoluzione temporale del sistema; fu così che Runge e Gross^[26] svilupparono un nuovo metodo computazionale, ovvero la *TD-DFT*.

Tale metodo permette infatti di studiare le proprietà e la dinamica dei sistemi a più corpi in presenza di potenziali dipendenti dal tempo, come campi elettrici o magnetici, in grado di indurre transizioni tra gli stati elettronici del sistema.

La *TD-DFT* si basa sugli stessi principi della *DFT*, andando però a risolvere l'*equazione di Schrödinger dipendente dal tempo* (Eq. 2.22), posto che si conosca lo stato de sistema a $t_0=0$.

$$i \frac{\partial}{\partial t} \psi(\{r\}, t) = \hat{H}(\{r\}, t) \psi(\{r\}, t) \quad \text{Eq. 2.11}$$

con $\{r\}$ le coordinate spaziali degli N elettroni del sistema.

L'hamiltoniano può essere generalmente suddiviso in due termini:

$$\hat{H}_{rg} = \hat{T}(\{r\}) + \hat{V}_{ext}(\{r\}, t) = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_i^N v_{ext}(r_i, t) \quad \text{Eq. 2.12}$$

dove $T(\{r\})$ rappresenta l'energia cinetica elettronica, mentre $V(\{r\})$ è dato dalla sommatoria di un potenziale esterno, dipendente dal tempo, risentito da ciascun elettrone i -esimo; tale potenziale, oltre ad includere l'interazione Coulombiana tra elettrone-elettrone, elettrone-nucleo ed il potenziale di scambio e correlazione, tiene conto anche dell'eventuale interazione degli elettroni con campi esterni dipendenti dal tempo, in grado di provocare transizioni tra gli stati del sistema.

Dalla risoluzione dell'*equazione di Schrödinger dipendente dal tempo* si ottengono gli autostati del sistema, da cui poter ricavare la densità elettronica secondo l'equazione:

$$\rho(r, t) = N \int |\psi(r, r_2, \dots, r_N, t)|^2 dr_2 \dots dr_N \quad \text{Eq. 2.13}$$

la quale definisce la probabilità di trovare un elettrone nella posizione r al tempo t .

Il teorema centrale della *TD-DFT* (chiamato anche *teorema di Runge-Gross*) dimostra che esiste una corrispondenza biunivoca tra il potenziale esterno v_{ext} (dipendente dal tempo) e la densità elettronica $\rho(r, t)$, per un sistema a molti corpi che evolve da uno stato iniziale fisso e noto. Ciò implica che tutte le proprietà di un sistema possano dunque essere determinate in funzione della sola densità.

In possesso del teorema di Runge-Gross, è possibile sviluppare uno schema dipendente dal tempo simile a quello di Kohn-Sham, usato in *DFT*.

Andando a definire il *potenziale effettivo* dipendente dal tempo come:

$$v_{ks,t}(r, t) = v_{eN}(r, t) + v_{Coul}(r, t) + v_{xc}(r, t) \quad \text{Eq. 2.14}$$

- $v_{eN}(r,t)$ potenziale di interazione elettrone nucleo, calcolato come interazione elettrostatica coulombiana;
- $v_{Coul}(r,t)$ potenziale di repulsione elettrostatica tra gli elettroni, rappresentato anch'esso da un'interazione coulombiana;
- $v_{xc}(r,t)$ potenziale di correlazione e scambio, che rappresenta appunto la parte di *correlazione elettronica* non calcolabile esattamente e che di conseguenza è necessario approssimare con funzionali.

Utilizzando un funzionale adatto per la parte di correlazione e scambio è dunque possibile costruire il sistema elettronico non interagente a cui è associata la medesima densità elettronica, dipendente dal tempo, del sistema interagente, calcolata a seguito della risoluzione delle *equazioni di Kohn-Sham dipendenti dal tempo*:

$$\hat{H}_{rg} \varphi_i(r, t) = \varepsilon_i(r, t) \varphi_i(r, t) \quad \text{Eq. 2.15}$$

$$\hat{H}_{rg} = \left(\frac{1}{2} \hat{\nabla}^2 + v_{ext}(r, t) \right) \quad \text{Eq. 2.16}$$

da cui poter ricavare φ_i e dunque la densità elettronica associata:

$$\rho(r, t) = \sum_i n_i |\varphi_i(r, t)|^2 \quad \text{Eq. 2.17}$$

Una trattazione più completa ed approfondita del metodo non rientra negli obiettivi di questo lavoro di tesi.

Il calcolo delle frequenze dei modi normali di vibrazione avviene mediante costruzione della *mass-weighted Hessian matrix*, ovvero una matrice composta dalle derivate seconde dell'energia rispetto alle posizioni degli atomi, pesate sulle masse degli atomi; tale matrice ha dimensione $3N \times 3N$, con N numero degli atomi del sistema, ed ogni elemento è del tipo:

$$\epsilon_{i,j} = \frac{1}{\sqrt{M_i M_j}} \frac{\partial^2 E}{\partial x_i \partial x_j} \quad \text{Eq. 2.18}$$

con M_i massa e x_i posizione dell'atomo i -esimo.

Gli autovalori e autovettori della matrice rappresentano rispettivamente le frequenze di vibrazione ν ed i corrispondenti modi normali.

2.3.2 Calcoli e dettagli computazionali

L'input del calcolo, cioè la struttura molecolare del QC (*metiltimina*), è stata ottenuta mediante l'utilizzo del software Avogadro^[18]. Sono state simulate le frequenze vibrazionali del sistema tale e quale e del sistema idealmente posto in una soluzione di acqua deuterata D₂O, la quale può "scambiare" atomi di D, andando a "sostituire" l'atomo di H in posizione 18 della TMP ([figura 1.1](#)), che presenta caratteristiche acide; in questo caso è stato necessario variare la topologia del sistema prima di eseguire il calcolo.

Tramite il programma Gaussian 16^[27] sono state calcolate le frequenze di vibrazione dei modi normali relativi allo stato fondamentale.

Dapprima è stata eseguita una ottimizzazione di geometria e un calcolo TDDFT dei primi tre modi eccitati, per ottenere struttura ed energia effettive degli stati in analisi. Successivamente sono state calcolate le frequenze relative ad i modi normali di tutta la molecola, andando poi ad isolare solamente le frequenze relative allo stretching dei due gruppi carbonile C=O e C-O.

In entrambi i calcoli il livello di teoria utilizzato è stato B3LYP/6-31G+(d).

Capitolo 3

Risultati

In questa sezione vengono riportati dapprima i risultati ottenuti con il calcolo DFT in *gas-phase*, ovvero eseguito sul QC isolato e dunque imperturbato.

Successivamente vengono riportati i risultati dal calcolo PMM eseguito utilizzando la *timidina 5' monofosfato* in soluzione di H₂O e D₂O, ovvero nello specifico le frequenze di vibrazione relative allo stretching dei due gruppi carbonile (C₂O e C₄O) della molecola in analisi. Questi dati vengono poi confrontati con dati sperimentali e ottenuti da calcoli precedenti, eseguiti utilizzando la *metiltimina*.

3.1 Gas phase

Si riportano in Tabella 3.1 i valori sperimentali delle frequenze dei due gruppi carbonile C₂O e C₄O riferiti agli spettri IR della MeT registrati in Ar ed N₂:

		ν sperimentale (cm ⁻¹)
Ar	<u>C₂O</u>	1734
N ₂	<u>C₄O</u>	1705

Tabella 3.1 Valori sperimentali delle frequenze di vibrazione associate ai gruppi carbonilici C₂O e C₄O della MeT, misurati in atmosfera inerte di Ar.

Si riportano in Tabella 3.2 i risultati ottenuti dal calcolo DFT eseguito sul *quantum center* del sistema, ovvero la *metiltimina* (MeT), isolato e dunque non perturbato dal resto del sistema trattato a livello classico, per la *metiltimina* e la *metiltimina deuterata* (MeTD), rilevanti per le soluzioni in H₂O e D₂O. Questi dati vengono confrontati inoltre con i valori sperimentali riportati in Tabella 3.1.

QC	Gruppo	ν calcolata (cm ⁻¹)	$\Delta\nu$ (cm ⁻¹)
MeTD	<u>C₂O</u>	1777	43
	<u>C₄O</u>	1746	41
MeT	<u>C₂O</u>	1780	46
	<u>C₄O</u>	1760	55

Tabella 1.2 Valori delle frequenze di vibrazione associate ai gruppi carbonilici C₂O e C₄O della MeT, calcolati mediante metodo DFT livello di teoria B3LYP/6-31G+(d), confrontati poi con i dati sperimentali, riportando le differenze nei valori delle frequenze.

Discrepanze tra valore calcolate in *gas phase* e valore sperimentale sono dovute:

- all'errore intrinseco del calcolo QM (che deriva ad esempio dall'utilizzo di un funzionale dell'energia approssimato, oppure un set di base finito);
- dall'utilizzo dell' approssimazione armonica per descrivere i modi normali di vibrazione dei gruppi carbonilici.

3.2 Risultati del calcolo PMM

Si riportano in Figura 3.1 gli spettri calcolati in questo lavoro di tesi tramite il PMM, raffiguranti le bande vibrazionali, relative allo stretching dei gruppi carbonilici della *timidina 5' monofosfato* **A)** D₂O e **B)** H₂O.

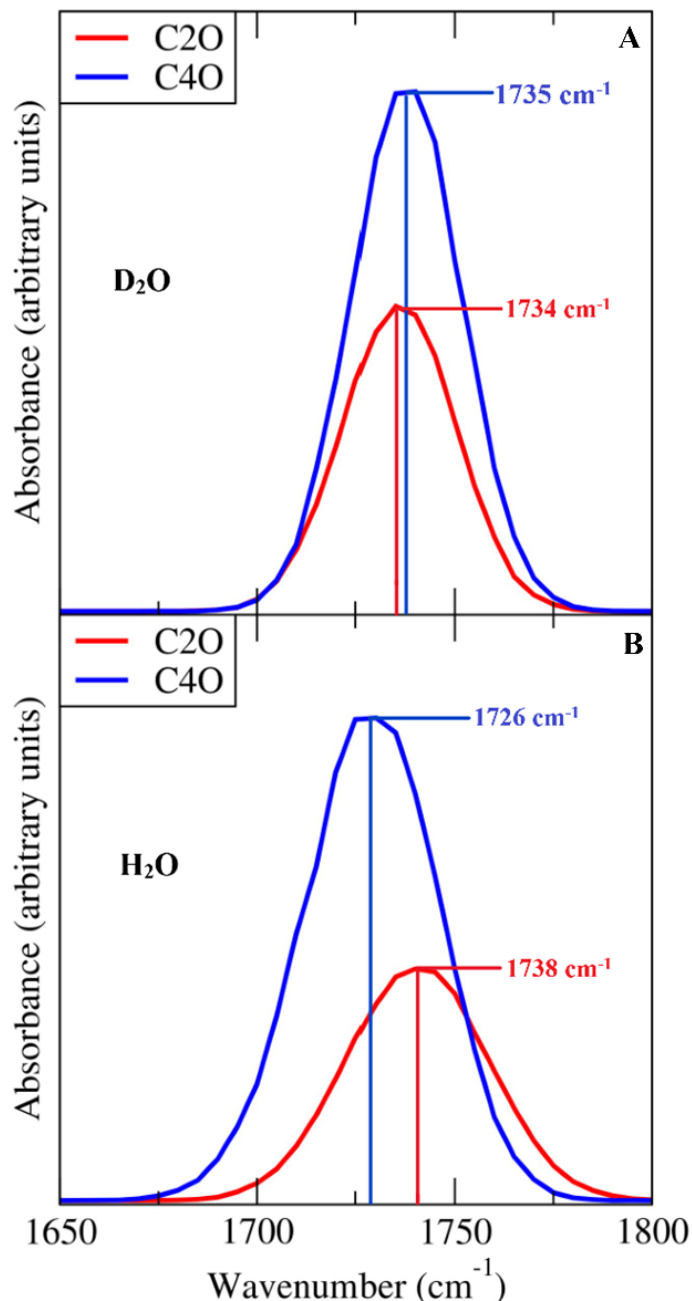


Figura 3.1 Bande vibrazionali relative ai modi normali di stretching dei gruppi carbonile della *timidina 5' monofosfato*, C₂O (rosso) e C₄O (blu), calcolate con il metodo PMM, in soluzioni di **A)** D₂O e **B)** H₂O. Entrambe le bande sono state normalizzate e di conseguenza non viene riportata l'intensità assoluta dei picchi; l'intensità relativa risulta essere però significativa.

I valori ottenuti necessitano di alcune correzioni, che permettono di tenere conto di errori dovuti all'accoppiamento dei modi di vibrazione (approssimati in questo calcolo come indipendenti) e dell'errore sperimentale del metodo utilizzato per il calcolo.

Le modifiche ai valori delle frequenze calcolate vengono riportati nella seguente tabella:

Solv.	Gruppo	ν calcolata (cm ⁻¹)	Correzione (cm ⁻¹)	ν corretta (cm ⁻¹)
D ₂ O	<u>C₂O</u>	1734	-41	<u>1693</u>
	<u>C₄O</u>	1735	-52	<u>1683</u>
H ₂ O	<u>C₂O</u>	1738	-46	<u>1692</u>
	<u>C₄O</u>	1726	-55	<u>1671</u>

Tabella 3.3 Correzioni applicate alle frequenze calcolate con il metodo PMM

A seguito delle correzioni applicate si può notare che sia in H₂O che in D₂O lo stretching del C₂O cade a frequenze più elevate rispetto a quello associato al C₄O. Questo in quanto, come riportato in letteratura^[7], la distanza di legame del gruppo C₂O è inferiore a quella del C₄O, dovuto anche al fatto che la coniugazione tra il doppio legame degli atomi C₅=C₆ e il gruppo C₄O comporta un allungamento del legame carbonilico. Questo fattore influenza dunque la costante di forza del legame e dunque la frequenza vibrazionale (proporzionale alla radice quadrata della prima), giustificando le posizioni rispettive delle bande.

Per approfondire queste affermazioni sono stati calcolati dei parametri relativi alla simulazione di MD, che ci permettono di analizzare meglio le interazioni che avvengono tra i due gruppi carbonilici C₂O e C₄O e il solvente H₂O, ovvero principalmente legami idrogeno. Nello specifico sono state calcolate:

- la funzione di distribuzione radiale $g(r)$ delle molecole di solvente rispetto ai due gruppi carbonilici:

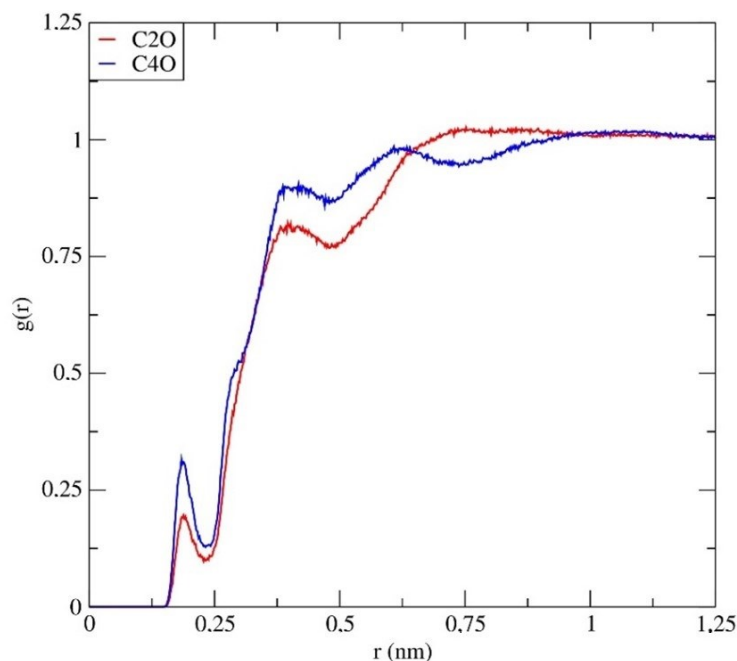


Figura 3.2 Grafico raffigurante la distribuzione radiale delle molecole di solvente rispetto ai due gruppi carbonilici della TMP C₂O (rosso) e C₄O (blu).

Dall'analisi di $g(r)$ si evince che il numero di molecole di H₂O coordinate al gruppo C₄O è maggiore rispetto invece a quelle legato al gruppo C₂O; dunque il primo risulta essere maggiormente esposto al solvente, con cui instaura un numero maggiore di legami idrogeno.

- il numero di coordinazione cumulativo (CN) delle molecole di solvente rispetto ai due gruppi carbonilici:

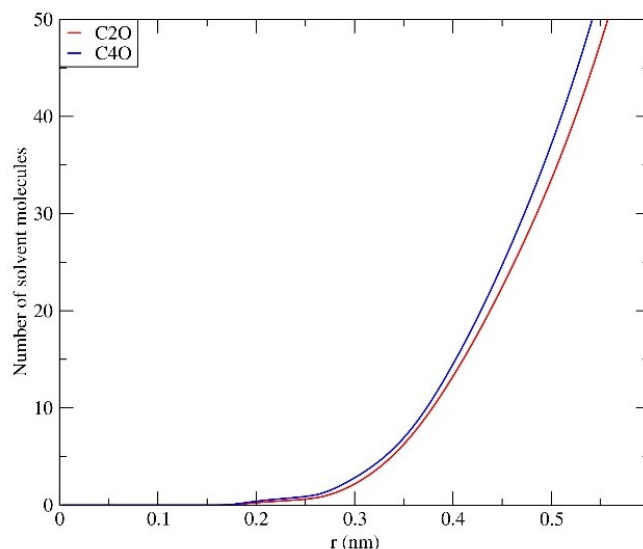


Figura 3.3 Grafico rappresentante il numero di coordinazione cumulativo delle molecole di H₂O presenti nell'intorno dei gruppi carbonilici della TMP C₂O (rosso) e C₄O (blu) in funzione della distanza da essi.

Dall'analisi di questi dati possiamo confermare quanto riportato nel punto precedente, ovvero che il gruppo C₄O coordina un numero maggiore di molecole di solvente.

- Distribuzione dei legami idrogeno, ovvero una funzione $h(n)$ che valuta la percentuale di configurazioni della simulazione MD che presentano un numero di legami idrogeno n . La distribuzione viene calcolata sulla sommatoria di tutte le configurazioni valutate.

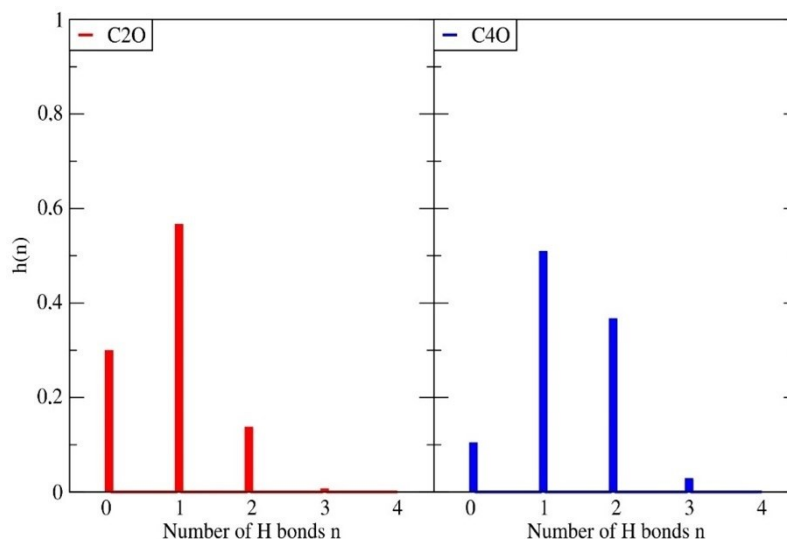


Figura 3.4 Grafico rappresentante la frazione degli snapshot della simulazione di MD in cui i gruppi carbonilici C₂O (rosso) e C₄O (blu) presentano un numero di legami idrogeno n .

Dall'analisi di questi dati possiamo affermare che il numero di legami idrogeno che coinvolgono il gruppo C_4O è maggiore; dunque mediamente, nel tempo, questo gruppo riesce ad instaurare e mantenere un numero maggiore di legami idrogeno.

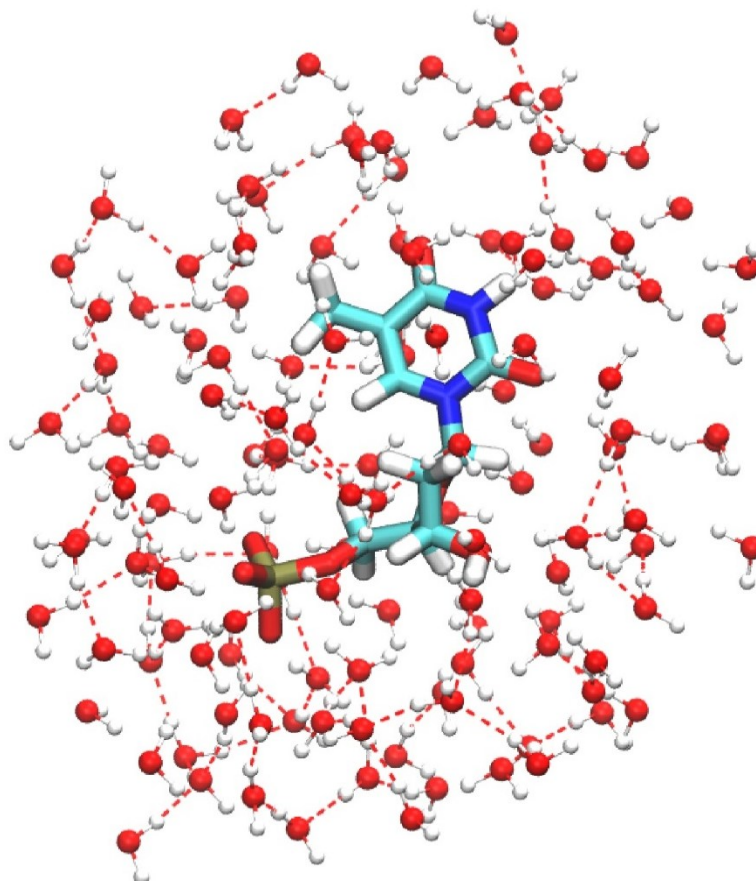


Figura 3.5 Immagine realizzata con VMD^[8] raffigurante una sfera di 5Å di H₂O intorno alla molecola di TMP, mettendo in evidenza i legami idrogeno.

In generale dunque possiamo affermare che il gruppo carbonilico C_4O ha maggiori interazioni con il solvente (tramite legami idrogeno) che comportano un aumento della massa ridotta del gruppo riflettendosi in una banda vibrazionale a frequenza minore, rispetto invece alla banda del C_2O .

3.3 Confronto tra dati calcolati e sperimentali della TMP

Si riportano in Figura 3.6 le bande vibrazionali relative alla *timidina 5' monofosfato* in soluzione in D₂O e H₂O calcolate con PMM, poste a confronto con lo spettro sperimentale^[30] della stessa molecola in soluzione di D₂O.

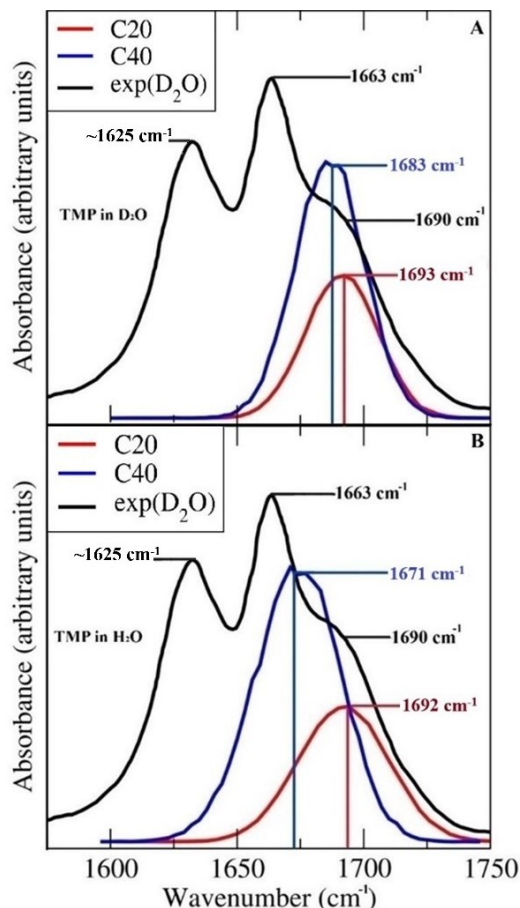


Figura 3.6 Confronto tra lo spettro sperimentale della TMP in D₂O (nero)^[29] (dove la banda vibrazionale a ~1625 cm⁻¹ non si origina da uno stretching di un legame C=O) e le bande vibrazionali di TMP in H₂O relative agli stretching dei gruppi C₂O (rosso) e C₄O (blu), calcolate in soluzione di **A**) D₂O e **B**) H₂O.

Per quanto riguarda la frequenza della banda del C₂O possiamo affermare che i dati ottenuti riproducono bene il valore sperimentale (1693 cm⁻¹ vs 1690 cm⁻¹, rispettivamente, per il risultato computazionale in D₂O; 1692 cm⁻¹ vs 1690 cm⁻¹, rispettivamente in H₂O); inoltre le frequenze non differiscono di molto l'una dall'altra. Non possiamo dire la stessa cosa per la banda del C₄O che, oltre a presentare valori delle frequenze diversi per D₂O e H₂O, restituisce in entrambi i casi un valore della frequenza di vibrazione sovrastimata rispetto a quello sperimentale.

Nello specifico la differenza rispetto al valore sperimentale per la banda C₄O è:

Solv.	valore calcolato (cm ⁻¹)	valore sperimentale (in D ₂ O) (cm ⁻¹)	$\Delta\nu$ (cm ⁻¹)
D ₂ O	1683	1663	20
H ₂ O	1671	1663	8

Tabella 3.4 Confronto tra i valori della frequenza della banda vibrazionale associata allo stretching del gruppo C₄O della TMP, calcolate utilizzando i modelli per TMP in D₂O e in H₂O, con il dato sperimentale della medesima banda vibrazionale in D₂O.

Per risolvere questo problema sono state applicate ulteriori correzioni ai valori delle frequenze del gruppo C_4O , pari ai valori di $\Delta\nu$, per poter ottenere un valore correlabile a quello sperimentale:

Solv.	valore calcolato (cm^{-1})	correzione (cm^{-1})	valore corretto (cm^{-1})
D ₂ O	1683	20	1663
H ₂ O	1671	8	1663

Tabella 3.5 Correzioni applicate alle frequenze del gruppo C_4O calcolate con il metodo PMM

Viene utilizzato come riferimento il valore della frequenza dello spettro sperimentale registrato in D₂O, nonostante il valore calcolato faccia riferimento alla TMP in H₂O in quanto si può supporre che non vi siano differenze sostanziali tra gli spettri sperimentali registrati in H₂O o in D₂O. Questo perché, data la ridotta acidità del protone legato all'azoto in posizione 6, questo non tenderà molto ad essere scambiato con atomi di D del solvente; di conseguenza la struttura della TMP in D₂O non differirà molto da quella in H₂O e conseguentemente il suo spettro vibrazionale.

Una possibile spiegazione dell'errore commesso nella determinazione delle bande vibrazionali potrebbe essere dovuta al fatto, che nel caso specifico dei calcoli eseguiti su TMP non si tiene conto di un possibile coupling eccitonico tra le bande dei gruppi C₂O e C₄O; tale effetto può modificare sia la frequenza di vibrazione, come nel caso del C₄O, che la forma e l'ampiezza della stessa. Questo tipo di test va però oltre il lavoro di tesi.

Si riportano in Figura 3.7 le bande vibrazionali con le correzioni applicate, posto a confronto con lo spettro sperimentale di TMP:

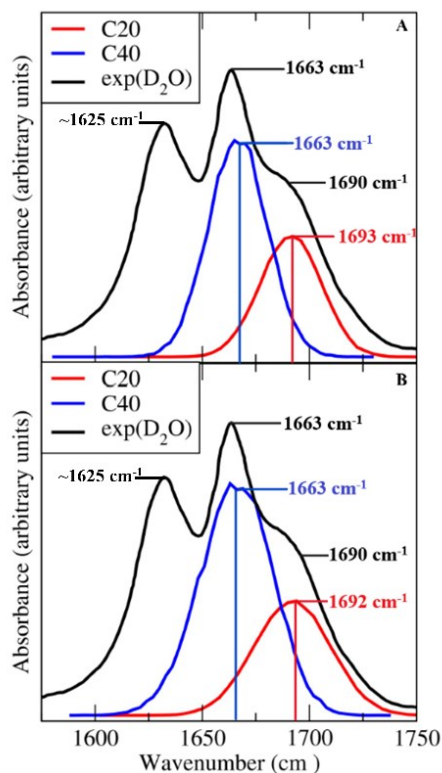


Figura 3.7 Confronto tra lo spettro sperimentale della TMP in D₂O (nero)^[29] (dove la banda vibrazionale a $\sim 1625\text{ cm}^{-1}$ non è riferita ad uno stretching di un legame C=O) e le bande vibrazionali corrette di TMP in H₂O relative agli stretching dei gruppi carbonilici C_2O (rosso) e C_4O (blu), calcolate in **A**) D₂O e **B**) H₂O.

3.4 Confronto tra i dati calcolati per la TMP e quelli di MeT

Si riporta in Figura 3.8 un confronto tra i dati calcolati in questo lavoro di tesi della TMP in D₂O, ed i dati ottenuti precedentemente dal gruppo di ricerca utilizzando come sistema di riferimento la MeT in soluzione di H₂O; in quest'ultimo caso sono state utilizzate le medesime approssimazioni per il calcolo PMM e di conseguenza è stato necessario includere ulteriori correzioni specifiche per questo sistema, per ricondurre i valori delle frequenze calcolate a quelli sperimentali.

Tali correzioni coinvolgono solamente il gruppo C₄O e comportano una variazione di 20 cm⁻¹; trascurando questa variazione si otterrebbe una frequenza pari a 1688 cm⁻¹, comparabile anch'essa con il valore calcolato per la TMP in D₂O privo di correzione, pari a 1683 cm⁻¹.

Si riportano sempre in Figura 3.8 gli spettri sperimentali di MeT in H₂O e TMP in D₂O.

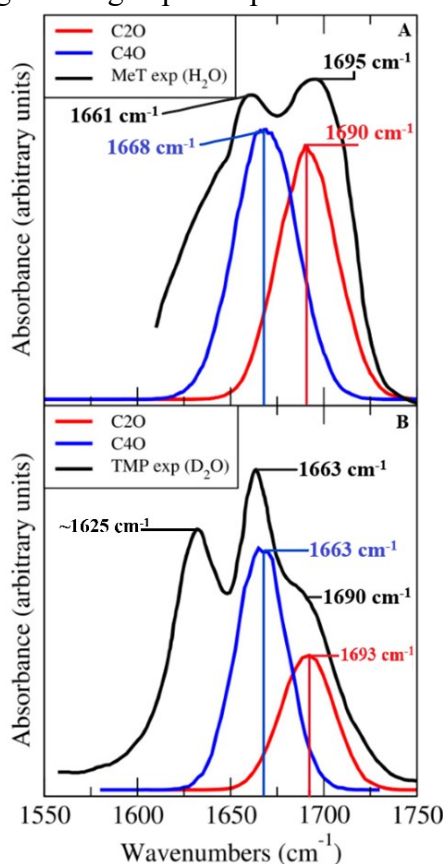


Figura 3.8 **A**) Bande vibrazionali di MeT in H₂O relative agli stretching dei gruppi carbonilici C₂O (**rosso**) e C₄O (**blu**), precedentemente calcolate dal gruppo di ricerca, poste a confronto con lo spettro sperimentale della MeT in H₂O (**nero**). **B**) Bande vibrazionali corrette di TMP in D₂O, relative agli stretching dei gruppi carbonilici C₂O (**rosso**) e C₄O (**blu**), poste a confronto con lo spettro sperimentale della TMP in D₂O (**nero**) (dove la banda vibrazionale a ~1625 cm⁻¹ non è riferita ad uno stretching di un legame C=O).

Dal confronto tra gli spettri **A** e **B** si può notare che le bande relative allo stretching dei gruppi C₂O e C₄O calcolate utilizzando come sistema di riferimento rispettivamente la TMP e la MeT sono pressoché identiche, a meno di una differenza trascurabile (1668 cm⁻¹ vs 1663 cm⁻¹ per C₄O della MeT e della TMP rispettivamente; 1690 cm⁻¹ vs 1693 cm⁻¹ per C₂O della MeT e della TMP rispettivamente).

Si può dunque affermare che l'utilizzo di un sistema semplificato come la MeT, risulti essere una ottima approssimazione del sistema più complesso della TMP.

3.5 Confronto tra i dati calcolati per la TMP e quelli di ssT10

Si riporta in Figura 3.9 il confronto tra gli spettri vibrazionali della TMP e di un oligomero ssT10 costituito da 10 nucleotidi aventi come base azotata la timina. Nello specifico si rappresentano :

- gli spettri sperimentali di entrambe le molecole, centrati sulle bande di vibrazione riferite agli stretching dei gruppi C₂O e C₄O, poste a confronto;
- gli spettri calcolati tramite il metodo PMM, andando a rappresentare:
 - una somma delle bande dei gruppi C₂O e C₄O per la TMP in D₂O;
 - la banda vibrazionale di ssT10 calcolata tenendo conto del coupling eccitonico sia intra-modo (tra modi normali di gruppi appartenenti allo stesso nucleotide) che inter-modo (tra modi normali di gruppi appartenenti nucleotidi diversi).

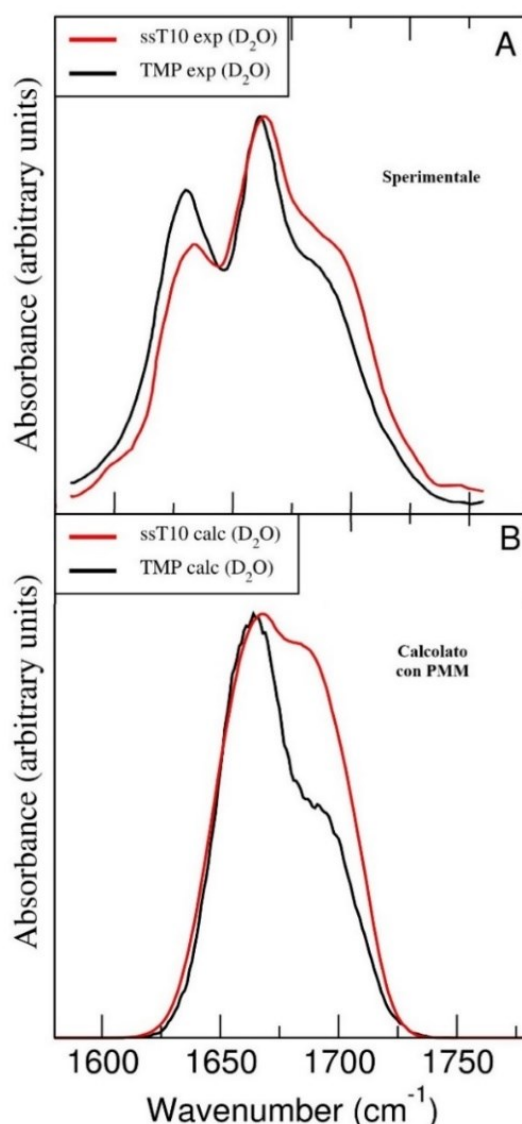


Figura 3.9 **A**) Spettri sperimentali di TMP (nero) e ssT10 (rosso) in D₂O relative agli stretching dei gruppi carbonilici C₂O e C₄O della base azotata posti a confronto l'uno con l'altro. **B**) Spettri vibrazionali calcolati con il metodo PMM per ssT10 (rosso) e TMP (nero), quest'ultimo raffigurante la somma delle bande vibrazionali dei gruppi C₂O e C₄O.

Ciò che si può notare dal confronto sia degli spettri sperimentali che di quelli calcolati è che la banda derivante dal gruppo C₂O si sposta a frequenze maggiori in ssT10, rispetto alla frequenza in TMP. Questo effetto è dovuto al *coupling eccitonico* intra-modo ed inter-modo tra i gruppi C₂O e C₄O delle basi azotate di ogni singolo acido nucleico componente l'oligomero. L'accoppiamento dipende dalla *frazione di stacking* delle basi azotate, che dipende a sua volta dall'elongazione della struttura di ssT10: più la struttura è allungata e meno le singole basi interagiscono; al contrario più la struttura tende a "compattarsi", maggiori sono le interazioni. In questo caso nello spettro calcolato si fa riferimento all'oligomero nella sua "conformazione naturale", ovvero non allungata.

L'effetto dovuto all'accoppiamento si riflette inoltre in una maggior sovrapposizione delle bande vibrazionali ed un aumento dell'intensità, come si può ben notare in figura 3.9 B; vi è infatti una netta differenza nella forma della banda associata alla TMP, rispetto a quella derivante da ssT10, in quanto in quest'ultimo caso si è tenuto conto in modo esplicito degli effetti di coupling. L'aumento di intensità relativa della banda C₄O passando da TMP a ssT10 è evidente sia nel dato sperimentale, sia nel calcolo, ma sembra sovrastimato in quest'ultimo.

Capitolo 4

Conclusioni

In questo lavoro di tesi sono state calcolate le bande vibrazionali relative allo stretching dei due gruppi carbonilici (C_2O e C_4O) della *timidina 5' monofosfato* (TMP), utilizzando un metodo ibrido *QM/MM*, ovvero il *perturbed matrix method* (PMM).

Questi calcoli sono stati eseguiti con l'obiettivo di valutare l'efficacia riguardo l'utilizzo di un sistema approssimato, ovvero la *metiltimina* (MeT), per valutare le medesime bande; questa molecola, derivata della base azotata *timina*, presenta un gruppo CH_3 dove dovrebbe invece legarsi il deossiribosio, con annesso il gruppo fosfato, presenti invece nella struttura della TMP. L'approssimazione strutturale si basa essenzialmente sul fatto che lo zucchero ed il gruppo fosfato non inducano variazioni sensibili sulle frequenze di vibrazione dei gruppi carbonilici della base azotata.

Nel calcolo PMM riferito alla TMP, sono stati utilizzati come *quantum center* (QC) la MeT e la *metiltimina deuterata* (MeTD) (che simula il sistema posto in una soluzione di acqua deuterata), mentre il deossiribosio ed il gruppo fosfato rientrano nella parte di sistema trattata a livello classico, andando dunque ad influire sulla componente perturbativa applicata al QC.

Come risultato di questo calcolo sono state ottenute le bande vibrazionali, con le rispettive frequenze, riferite allo stretching dei due gruppi carbonile della base azotata.

Questi risultati sono stati dapprima confrontati con lo spettro vibrazionale sperimentale della TMP in soluzione di D_2O ; a seguito dell'applicazione di opportune correzioni, necessarie per tener conto di errori dovuti all'accoppiamento dei modi di vibrazione (approssimati in questo calcolo come indipendenti) e dell'errore sperimentale associato al metodo utilizzato per il calcolo, i valori delle frequenze vibrazionali calcolate con il PMM risultano essere coerenti e molto simili a quelli sperimentali, a meno di differenze trascurabili. Tramite un'analisi più approfondita della simulazione di dinamica molecolare, ovvero valutando la *funzione di distribuzione radiale* ed il *numero di coordinazione cumulativo* riferiti a i due gruppi carbonilici, è stato possibile giustificare le posizioni relative delle due bande: il gruppo C_4O presenta una frequenza di vibrazione inferiore rispetto al gruppo C_2O , in quanto il primo ha un numero di coordinazione delle molecole di solvente maggiore rispetto al secondo, coerente con il fatto di essere in grado di stabilire un numero medio di legami idrogeno più elevato; ciò comporta un aumento della massa ridotta del gruppo, che si riflette in una banda vibrazionale a frequenza minore.

I risultati del calcolo PMM sono poi stati confrontati con i valori delle frequenze vibrazionali dei gruppi carbonilici C_2O e C_4O , calcolati precedentemente dal gruppo di ricerca, utilizzando come sistema di riferimento la MeT. Dall'analisi dei dati ottenuti è stato possibile affermare che non vi sono sostanziali differenze nei valori delle frequenze calcolate utilizzando come sistema di riferimento la MeT o la TMP. Questo permette di concludere che l'utilizzo della MeT risulta essere una approssimazione efficace del sistema più complesso della TMP, in cui il deossiribosio ed il gruppo fosfato non

costituiscono fonti di variazioni rilevanti nei valori delle frequenze di vibrazione dei gruppi carbonilici. L'utilizzo della MeT come sistema di riferimento permette dunque una analisi computazionalmente meno costosa, perché eseguita su un sistema di dimensioni ridotte, ma che porta all'ottenimento di dati attendibili e coerenti, anche rispetto al confronto con dati sperimentali.

In ultima analisi i dati ottenuti dal calcolo PMM sulla TMP sono stati confrontati con lo spettro vibrazionale di un oligomero ssT10 (composto da 10 acidi nucleici aventi tutti come base azotata la timina) calcolato con il medesimo metodo, tenendo conto però degli effetti dell'accoppiamento tra i gruppi C₂O e C₄O delle basi azotate di ogni singolo nucleotide componente l'oligomero. In questo sistema è presente anche un effetto di *stacking* dovuto alla sovrapposizione delle basi azotate dei diversi nucleotidi; tale effetto modifica sia l'ambiente elettrostatico che l'accoppiamento eccitonico del sistema, portando all'ottenimento di spettri vibrazionali diversi rispetto a quelli di MeT e TMP. Nello specifico si nota una sovrapposizione più marcata delle bande, in funzione della *frazione di stacking* delle basi azotate di ogni singolo nucleotide, valore a sua volta dipendente dalla struttura più o meno allungata (e quindi denaturalizzata) del polimero. Non potendo distinguere a priori gli effetti di *stacking* e di coupling eccitonico, non è possibile valutare in modo dettagliato le variazioni apportate dall'inclusione di quest'ultimo (per fare ciò bisognerebbe valutare due diversi spettri calcolati con l'inclusione degli accoppiamenti eccitonici e senza di essi). Ciò che possiamo concludere è però che il modello completo è in grado di riprodurre in modo appropriato i trend ottenuti nello spettro sperimentale.

In conclusione, i dati ottenuti sono risultati coerenti con le simulazioni precedentemente condotte dal gruppo e appaiono promettenti per applicazioni pratiche; tuttavia, è indispensabile capire l'origine della sovrastima della frequenza della banda C₄O rispetto al dato sperimentale in modo da giustificare la correzione empirica applicata ai dati ottenuti in questo lavoro di tesi, al fine di sviluppare al meglio il metodo utilizzato.

Bibliografia

1. Hernandez, L. I., Machado, I., Schaefer, T. & Hernandez, F. J. Aptamers Overview: Selection, Features and Applications. *Curr. Top. Med. Chem.* 15, 1066–1081(2015).
2. Hegde, M. L. et al. First evidence for helical transitions in supercoiled DNA by amyloid β peptide (1–42) and aluminum. *Journal of Molecular Neuroscience* 22, 19–31(2004).
3. Jiang, K. et al. Alpha-Synuclein Modulates the Physical Properties of DNA. *Chemistry – A European Journal* 24, 15685–15690(2018).
4. Radom, F., Jurek, P., Mazurek, M., Otlewski, J. & Jelen, F. Aptamers: Molecules of great potential. *BIOTECHNOLOGY ADVANCES* 31, 1260–1274(2013).
5. *iSenseDNA*. <https://www.isensedna.eu/w/is>
6. Daidone, I., Aschi, M., Zanetti-Polzi, L., Di Nola, A. & Amadei, A. On the origin of IR spectral changes upon protein folding. *Chemical Physics Letters* 488, 213–218(2010).
7. Morzyk-Ociepa, B., Nowak, M. J. & Michalska, D. Vibrational spectra of 1-methylthymine: matrix isolation, solid state and theoretical studies. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 60, 2113–2123(2004).
8. Spivak, M. et al. VMD as a Platform for Interactive Small Molecule Preparation and Visualization in Quantum and Classical Simulations. *J. Chem. Inf. Model.* 63, 4664–4678(2023).
9. Miyamoto, K. et al. Hydration of single-stranded DNA in water studied by infrared spectroscopy. *Chemical Physics Letters* 436, 233–238(2007).
10. Daidone, I., Aschi, M., Zanetti-Polzi, L., Di Nola, A. & Amadei, A. On the origin of IR spectral changes upon protein folding. *Chemical Physics Letters* 488, 213–218 (2010).
11. Zanetti-Polzi, L. et al. Extending the perturbed matrix method beyond the dipolar approximation: comparison of different levels of theory. *Phys. Chem. Chem. Phys.* 20, 24369–24378(2018).
12. Amadei, A., Daidone, I. & Aschi, M. A general theoretical model for electron transfer reactions in complex systems. *Physical chemistry chemical physics : PCCP* 14, 1360–70 (2011).
13. <http://www.mdtutorials.com/gmx/lysozyme/index.html>
14. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* 14, 71–73(2017).
15. 1. Verlet, L. Computer ‘Experiments’ on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* 159, 98–103(1967).

16. Hess, B., Bekker, H., Berendsen, H. & Fraaije, J. LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry* 18,(1998).
17. Abraham, M. et al. GROMACS 2024.2 Manual.(2024)
18. Hanwell, M. D. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics* 4, 17 (2012).
19. Brice, A. R. & Dominy, B. N. Examining Electrostatic Influences on Base-Flipping: A Comparison of TIP3P and GB Solvent Models. *Commun. comput. phys.* 13, 223–237 (2013).
20. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* 126, 014101(2007).
21. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* 52, 7182–7190(1981).
22. Gross, E. K. U. & Maitra, N. T. Introduction to TDDFT. in *Fundamentals of Time-Dependent Density Functional Theory* (eds. Marques, M. A. L., Maitra, N. T., Nogueira, F. M. S., Gross, E. K. U. & Rubio, A.) vol. 837 53–99 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012).
23. Marques, M. A. L. & Gross, E. K. U. TIME-DEPENDENT DENSITY FUNCTIONAL THEORY. *Annu. Rev. Phys. Chem.* 55, 427–455(2004).
24. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* 136, B864–B871 (1964).
25. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* 140, A1133–A1138 (1965).
26. Runge, E. & Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* 52, 997–1000(1984).
27. Gaussian 16 | Gaussian.com. <https://gaussian.com/gaussian16/>.
28. GaussView, Version 6.1, Roy Dennington, Todd A. Keith, and John M. Millam, Semichem Inc., Shawnee Mission, KS, 2016.
29. Kolling, O. W. Effect of Hydrogen Bonding Solvents on the Infrared Absorption Band for the Fundamental Vibration of the Carbonyl Group in 1,1,3,3-Tetramethylurea. *Transactions of the Kansas Academy of Science* (1903-) 102, 53(1999).