



UNIVERSITÀ
DEGLI STUDI
FIRENZE

Scuola di
Scienze Matematiche
Fisiche e Naturali

Corso di Laurea in
Chimica
Curriculum Scienze Chimiche

Investigazione strutturale della modalità di interazione con DNA di leganti azotati di origine sintetica

Structural investigation of interaction modes between DNA and synthetic nitrogen ligands

Relatore
Carla Bazzicalupi

Correlatore
Paola Gratteri

Candidato
Marta Morana

Anno Accademico 2013/2014

Sommario

1	Il DNA e le sue strutture secondarie	4
1.1	Il DNA	4
1.2	Il DNA G-Quadruplex.....	8
1.3	Il DNA telomerico e le sue strutture G-quadruplex	11
2	Leganti per G-quadruplex	14
2.1	Caratteristiche dei leganti.....	14
2.1.1	Esempi di leganti.....	15
2.2	L1, L2 e L3.....	21
3	Obiettivi di questo lavoro di tesi	25
4	Modellistica molecolare	27
4.1	Principi teorici	27
4.1.1	Meccanica molecolare	27
4.1.2	Campo di forza empirico.....	28
4.1.3	Analisi conformazionale.....	33
4.1.4	Minimizzazione	36
4.2	Materiali e metodi	39
4.2.1	Materiali.....	40
4.2.2	Metodi.....	40
5	Cristallografia e cristallizzazione	44
5.1	Principi teorici	44
5.1.1	Cristallografia	44
5.1.2	Cristallografia e cristallizzazione di molecole biologiche.....	46
5.1.3	Tecniche di cristallizzazione	49
5.1.4	Tecniche criogeniche	50
5.2	Materiali e metodi	51
5.2.1	Materiali.....	51

5.2.2	Metodi.....	51
6	Discussione	55
6.1	Prove di cristallizzazione	55
6.2	Modellistica molecolare.....	56
6.2.1	Analisi degli addotti	69
6.2.2	Addotti con L3.....	69
6.2.3	Addotti con L2.....	75
6.2.4	Addotti con L1.....	78
7	Conclusioni	83
8	Bibliografia	85

1 Il DNA e le sue strutture secondarie¹

1.1 1.1 Il DNA

Gli acidi nucleici hanno la funzione di conservare e trasmettere l'informazione genetica. Si distinguono in acidi ribonucleici, RNA, e desossiribonucleici, DNA.

Sono formati da nucleotidi, che sono a loro volta costituiti da tre componenti: una base azotata, un pentosio e dei gruppi fosforici. In assenza di gruppi fosforici si parla di nucleosidi. Le basi azotate derivano da due precursori: la pirimidina e la purina. Sia il DNA sia l'RNA contengono due basi puriniche, l'adenina e la guanina; per quanto riguarda le basi pirimidiniche, in entrambi gli acidi nucleici è presente la citosina, nel DNA si trova poi la timina, sostituita nell'RNA dall'uracile (Figura 1-1). I due acidi nucleici differiscono anche per il pentosio che contengono: nell'RNA si trova un D-ribosio, mentre nel DNA è presente un 2'-deossi-D-ribosio.

¹ Dove non altrimenti specificato la fonte è David L. Nelson, Michael M. Cox, I principi di biochimica di Lehninger, Zanichelli

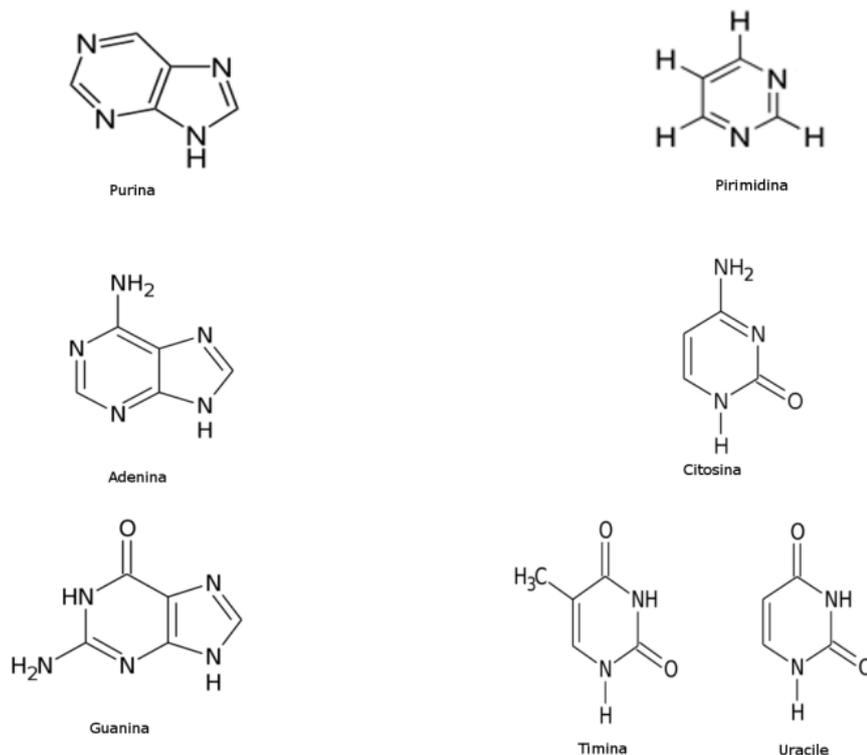


Figura 1-1 Basi azotate

La base di un nucleoside è legata mediante l'atomo N-1 delle pirimidine o l'atomo N-9 delle purine con il carbonio 1' del pentosio; il gruppo fosforico è invece esterificato al carbonio in posizione 5'. I nucleotidi sono uniti tra loro da un legame fosfodiesterico tra il gruppo fosforico in posizione 5' e il gruppo ossidrilico 3' del nucleotide successivo.

Tra i gruppi amminici e i gruppi carbonilici si instaurano legami a idrogeno, che permettono a due, più raramente a tre o quattro, filamenti complementari di acidi nucleici di interagire. Gli appaiamenti più comuni sono quelli descritti da James D. Watson e Francis Crick nel 1953, in cui la guanina si lega alla citosina e l'adenina alla timina o all'uracile; in base al loro modello, il DNA è costituito da due catene elicoidali avvolte intorno allo stesso asse a formare una doppia elica destrorsa; lo scheletro covalente si trova all'esterno, mentre le basi azotate sono impilate all'interno (Figura 1-2). Le due catene dell'elica sono

antiparallele e le basi sono distanziate di 3,4 Å. Per via della relazione spaziale tra le due catene, sulla superficie della doppia elica si formano una scanalatura, o solco, maggiore e una scanalatura, o solco, minore.

La struttura di Watson e Crick viene anche chiamata forma B del DNA ed è la più stabile in condizioni fisiologiche. Sono state caratterizzate due varianti: la forma A e la forma Z. La forma A è sempre costituita da due catene destrorse, ma l'elica più ampia e il piano delle basi inclinato di circa 20°C rispetto all'asse dell'elica rendono il solco maggiore più profondo e il solco minore meno pronunciato. Questa forma è favorita in un ambiente povero d'acqua. La forma Z si distingue invece per la rotazione dell'elica in senso sinistrorso, con un solco maggiore molto piccolo e un solco minore stretto e profondo. Esistono variazioni strutturali legate alla sequenza. Regioni del DNA con sequenze di basi ripetute invertite, ad esempio, possono formare strutture a forcina o a croce. Altre strutture del DNA sono formate da tre o quattro filamenti; in esse sono presenti basi che si legano con appaiamenti diversi da quelli canonici, detti appaiamenti Hoogsteen (Figura 1-3). Le sequenze ricche di guanina possono inoltre formare una struttura tetraplex o G-quadruplex a partire da quattro residui di guanina.

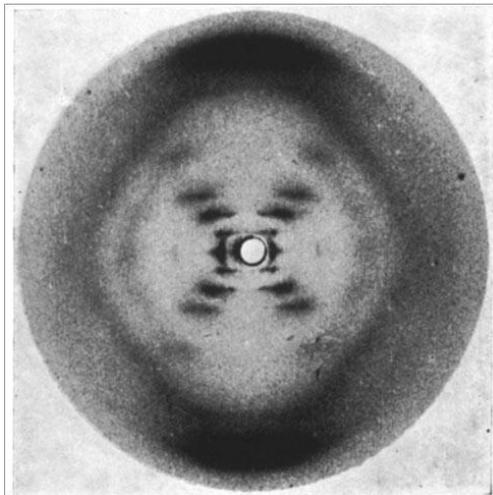


Figura 1-2 Diffrazione di raggi X ottenuta da Rosalind E. Franklin nel 1952

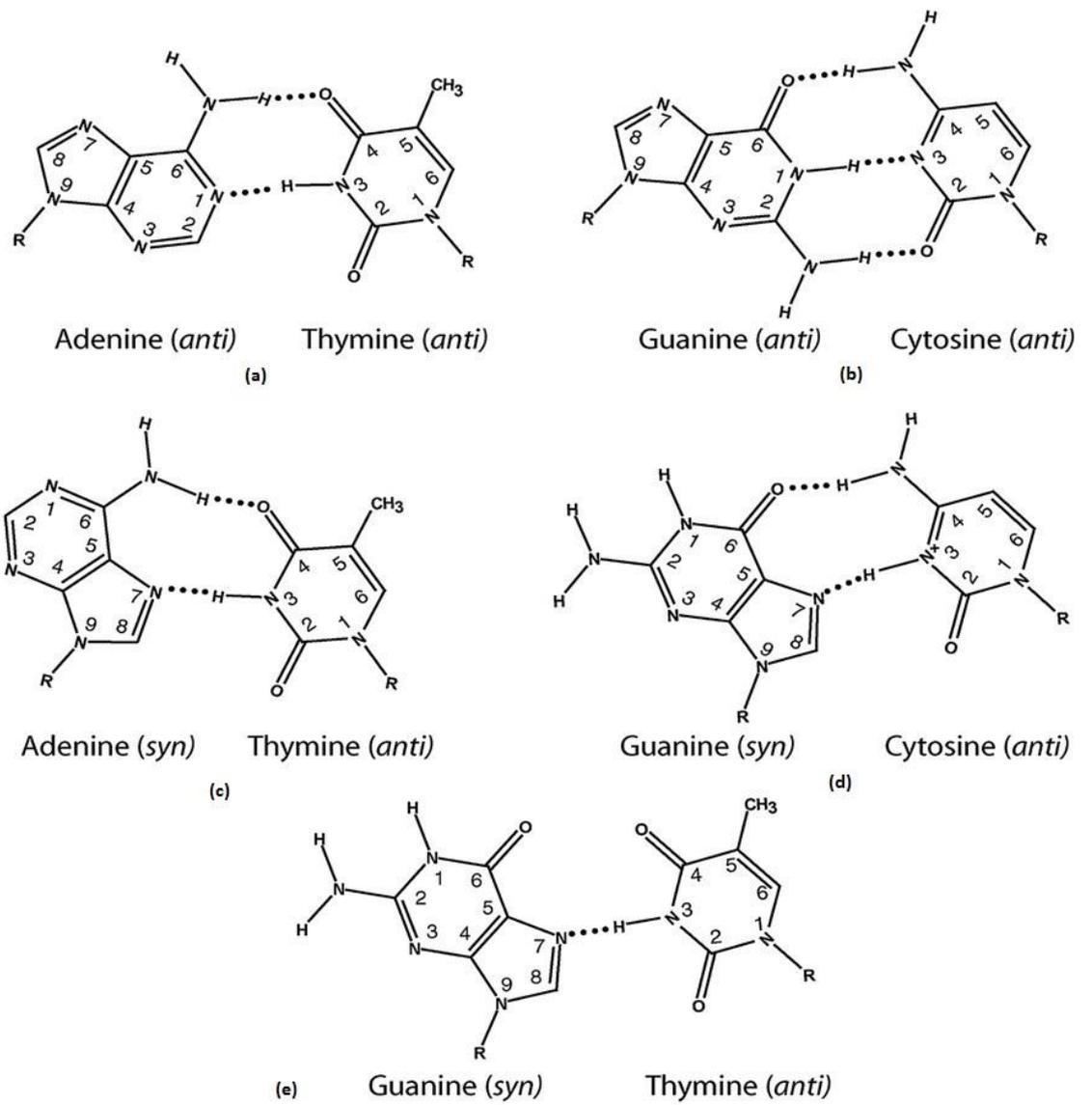


Figura 1-3 Appaiamenti di Watson e Crick, (a) e (b), e di Hoogsteen, (c), (d), (e)

1.2 Il DNA G-Quadruplex

Le prime osservazioni riguardo il particolare comportamento delle sequenze ricche di guanina risalgono al 1910, quando fu riportato che l'acido guanilidico, o guanosina monofosfato, in alte concentrazioni formava un gel². Cinquant'anni più tardi tramite la diffrazione di raggi X fu dimostrato che l'acido guanilidico può formare tetrameri³. In queste strutture quattro guanine formano un G-quartet o G-tetrad, ossia una disposizione planare di forma approssimativamente quadrata in cui ogni guanina è legata tramite legami a idrogeno con altre due guanine (Figura 1-4). Più G-quartet impilati formano una struttura G-quadruplex

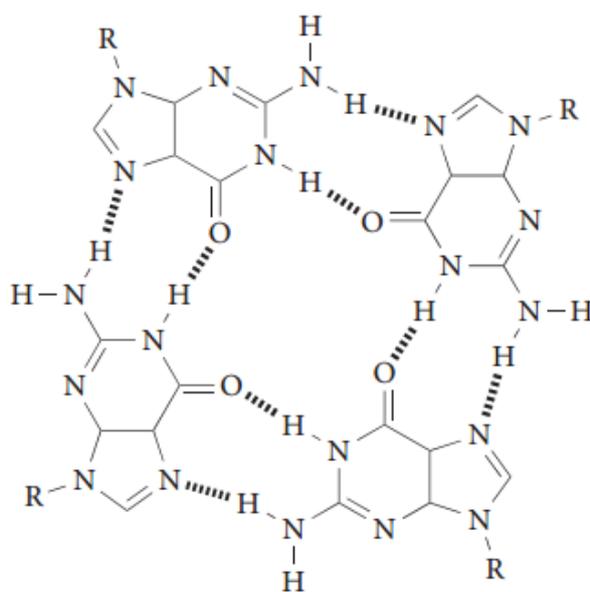


Figura 1-4 G-quartet

² Bang I. Untersuchungen über die Guanylsäure. Biochem. Z.26, 293-231 (1910)

³ Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. Proc. Natl Acad. Sci. USA 48, 2013–2018 (1962).

I piani possono essere collegati da loop a singolo filamento di varia lunghezza, anche se sono generalmente corti, da 1 a 7 nucleotidi. Le strutture con loop più brevi risultano più stabili rispetto a quelle con loop più lunghi⁴. La struttura G-quadruplex è stabilizzata da cationi monovalenti, solitamente Na⁺ o K⁺, che occupano gli spazi tra i piani e neutralizzano la repulsione elettrostatica degli ossigeni delle guanine, che puntano verso l'interno⁵. Le strutture G-quadruplex possono assumere diverse topologie, che sono classificate in base all'orientamento dei filamenti di DNA, in particolare possono essere parallele, antiparallele o ibride. Si possono formare a partire da un unico filamento e quindi essere intramolecolari, o da più filamenti e quindi essere intermolecolari (Figura 1-5). Sono inoltre possibili varie strutture dei loop⁶ (Figura 1-6).

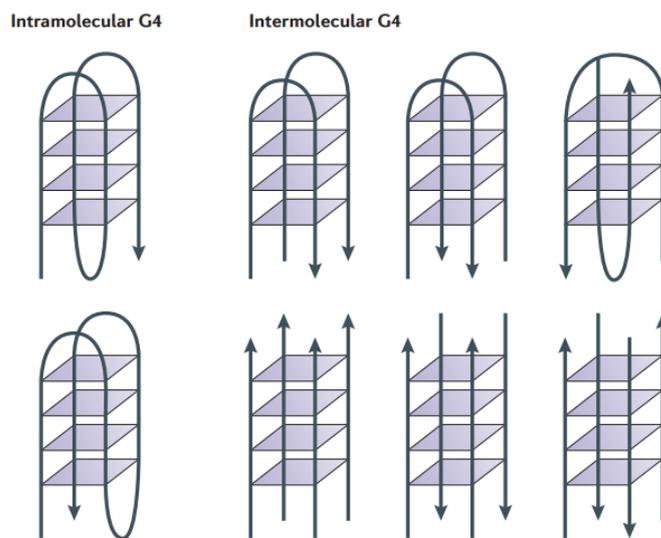


Figura 1-5 Topologie della struttura G-quadruplex

⁴ Huppert, J. L. Structure, location and interactions of G-quadruplexes. FEBS J. 277, 3452–3458 (2010)

⁵ Williamson, J. R., Raghuraman, M. K. & Cech, T. R. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell 59, 871–880 (1989).

⁶ Hazel, P., Parkinson, G. N. & Neidle, S. Predictive modelling of topology and loop variations in dimeric DNA quadruplex structures. Nucleic Acids Res. 34, 2117–2127 (2006)

Un altro importante parametro da considerare è la conformazione anti o sin del legame glicosidico: i quadruplex paralleli hanno tutte le guanine in conformazione sin, mentre negli antiparalleli si trovano entrambe le conformazioni (Figura 1-7).

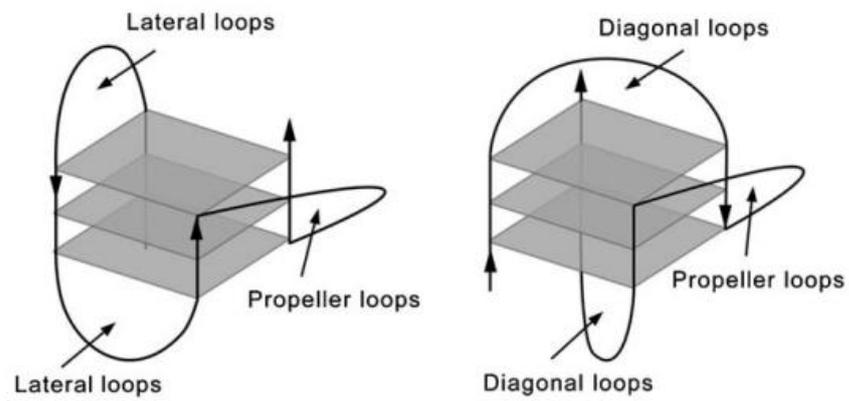


Figura 1-6 Strutture dei loop

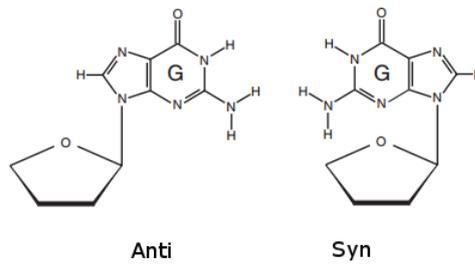


Figura 1-7 Conformazione anti e sin

1.3 Il DNA telomerico e le sue strutture G-quadruplex

Le particolari caratteristiche strutturali del G-quadruplex e il suo polimorfismo sono legati al suo ruolo biologico. Si trovano strutture di questo tipo in importanti regioni, come i promotori di alcuni oncogeni, il DNA ribosomiale e il DNA telomerico. La loro principale funzione è probabilmente attivare o disattivare alcuni processi fisiologici attraverso la regolazione della trascrizione.

I telomeri sono costituiti da DNA non codificante e si trovano all'estremità dei cromosomi. Nell'uomo hanno una lunghezza di decine di migliaia di coppie di basi della sequenza TTAGGG in una catena e AATCCC nella catena complementare. Il filamento TTAGGG è più lungo dell'altro, quindi rimane una regione di DNA a catena singola, lunga poche centinaia di nucleotidi, all'estremità 3'. Hanno una duplice funzione: rendere più stabili i cromosomi e permettere la replicazione dell'intero DNA cromosomico. Per quanto riguarda la stabilità, i telomeri evitano che i cromosomi siano riconosciuti dai meccanismi di riparazione del DNA come rotture della doppia elica e, di conseguenza, degradati e fusi. L'estremità a singolo filamento si ripiega all'indietro e si appaia alla regione complementare nella parte a doppia elica del telomero, formando una struttura chiamata ansa T, o t loop, stabilizzata da alcune proteine, come le TRF1 e TRF2. L'ansa T protegge le estremità 3' dei cromosomi rendendole inaccessibili alle nucleasi e agli enzimi che riparano le interruzioni della doppia elica. Un ulteriore contributo alla stabilità può essere data dalla formazione di strutture G-quadruplex. Nel tempo sono state identificate varie strutture che il DNA telomerico umano può assumere, che si distinguono per l'orientamento del filamento e il tipo di loop presenti (Figura 1-8).

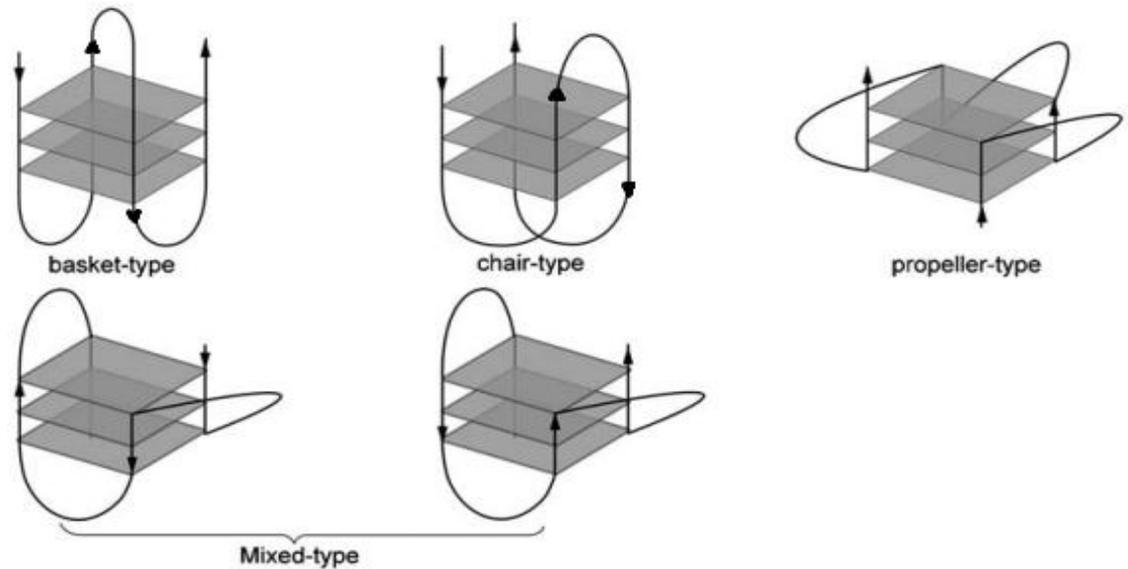


Figura 1-8 Strutture quadruplex del DNA telomerico umano

La seconda funzione dei telomeri è correlata alla replicazione del filamento ritardato della doppia elica. Affinché la replicazione possa cominciare, è necessario che all'altezza della forcella di replicazione siano presenti davanti alla sequenza da copiare un certo numero di nucleotidi che facciano da stampo per un primer di RNA, necessario a sua volta per l'azione della DNA polimerasi. La sequenza che funge da stampo per il primer non viene replicata e quindi è indispensabile che un'altra sequenza aggiuntiva sia presente all'estremità dei cromosomi per consentire la replicazione degli ultimi nucleotidi e quindi dell'intero DNA cromosomico. Queste sequenze aggiuntive sono appunto i telomeri. In seguito a questo meccanismo, i telomeri si accorciano a ogni replicazione, fino a raggiungere una lunghezza di soglia, detta limite di Hayflick, al di sotto della quale la cellula smette di duplicarsi, entra nella fase di senescenza e successivamente muore⁷. Nelle cellule che devono avere un'intensa attività proliferativa, i telomeri vengono costantemente allungati dall'enzima

⁷ C.B. Harley, A.B. Futcher, C.W. Greider, Telomeres shorten during ageing of human fibroblasts, Nature 345 (1990) 458e460.

telomerasi. La telomerasi è costituita da due componenti, uno a base di RNA e uno proteico. Il componente di RNA è lungo circa 150 nucleotidi e contiene circa 1,5 copie della sequenza AATCCC, che serve da stampo per la catena TTAGGG del telomero. Funziona come una trascrittasi inversa, che fornisce il sito attivo per la sintesi del DNA RNA-dipendente. La sintesi richiede l'estremità 3' di un cromosoma come primer e procede nel verso 5'-3'. Dopo aver copiato una prima sequenza, l'enzima si riposiziona per aggiungere le successive e allungare il telomero. Nell'uomo, la telomerasi risulta attivata nelle cellule germinali e staminali, al contrario di quanto avviene nelle cellule somatiche. In oltre l'80% dei tumori, la telomerasi è attivata e mantiene la lunghezza e l'integrità dei telomeri, permettendo alle cellule maligne di proliferare⁸. La formazione di G-quadruplex all'interno del DNA telomerico a singolo filamento inibisce l'attività della telomerasi⁹, di conseguenza la lotta ai tumori si è di recente sviluppata anche nella ricerca di composti che possano stabilizzare queste strutture intramolecolari. In presenza di G-quadruplex, infatti, l'enzima non è in grado di legarsi al substrato, portando avanti il suo meccanismo d'azione¹⁰.

⁸ J. W. Shay, S. Bacchetti, Eur. J. Cancer 1997, 33, 787-791

⁹ A.M. Zahler, J.R. Williamson, T.R. Cech, D.M. Prescott, Inhibition of telomerase by G-quartet DNA structures, Nature 350 (1991) 718-720

¹⁰ Kelland, L. Clinical Cancer Research, 2007, 13, 4960-4963

2 Leganti per G-quadruplex

2.1 Caratteristiche dei leganti

Da quando è stata scoperta la capacità delle strutture G-quadruplex di inibire l'attività della telomerasi, sono stati sintetizzati vari leganti che possono stabilizzare queste strutture. Si tratta generalmente di piccole molecole organiche con alcune caratteristiche comuni. Contengono un'estesa porzione aromatica planare, simile per dimensioni e forma al G-quartet, che può instaurare interazioni π - π con la piattaforma delle guanine. L'interazione con la densità π del G-quartet è favorita, se sono presenti centri di carica positiva. Talvolta i leganti sono dotati di catene laterali con gruppi carichi positivamente, che possono dar luogo a un'interazione coulombiana con lo scheletro fosfodiesterico¹¹. Sono teoricamente possibili tre diverse modalità d'interazione tra una struttura G-quadruplex e un legante: esso può impilarsi su uno dei piani esterni, instaurando delle interazioni π - π , dare fenomeni di intercalazione, posizionandosi tra due quartetti, o interagire con il solco (Figura 2-1.)

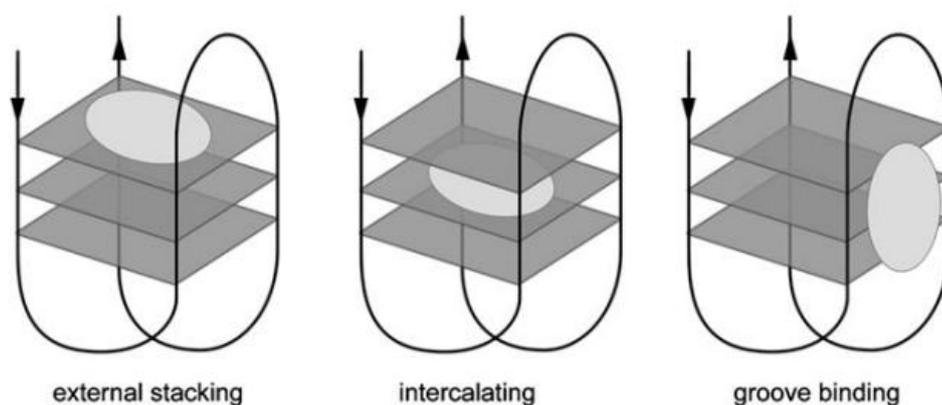


Figura 2-1 Modalità di interazione tra G-quadruplex e legante

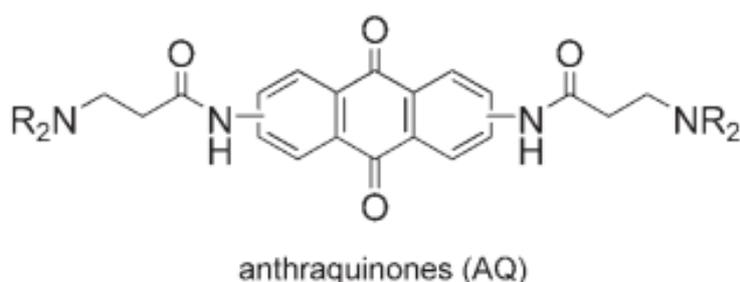
¹¹ G-Quadruplexes: Targets in Anticancer Drug Design, Tian-miao Ou, Yu-jing Lu, Jia-heng Tan, Zhi-shu Huang, Kwok-Yin Wong, Lian-quan Gu

I fenomeni d'intercalazione sono probabilmente poco favoriti, perché la struttura G-quadruplex è estremamente rigida e stabile e la distorsione causata dall'intercalazione avrebbe un elevato costo energetico. Di conseguenza il posizionamento su uno dei piani più esterni, in uno dei solchi laterali o in prossimità dei loop è molto più probabile, anche per quei leganti che possono potenzialmente dare intercalazione o la danno effettivamente con il DNA duplex¹².

2.1.1 Esempi di leganti

2.1.1.1 Leganti planari: antrachinoni e derivati

Uno dei primi leganti sviluppati come inibitori della telomerasi è il 2,6-diamminoalchilammido antrachinone (Figura 2-2)¹³. A partire da questo composto, sono stati sintetizzati molti derivati, variamente sostituiti. Composti planari di questo tipo sono stati sviluppati per interagire con il quartetto terminale, facilmente accessibile; è stato poi dimostrato che anche la forma del legante svolge un ruolo importante. Molecole angolari o a forma di V hanno poi dimostrato una maggiore affinità con il quadruplex¹⁴.



¹² Idem

¹³ D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, L. H. Hurley, *J. Med. Chem.* 1997, 40, 2113–2116

¹⁴ S. Bhattacharya, P. Chaudhuri, A. K. Jain and A. Paul, *Bioconjugate Chem.*, 2010, 21, 1148

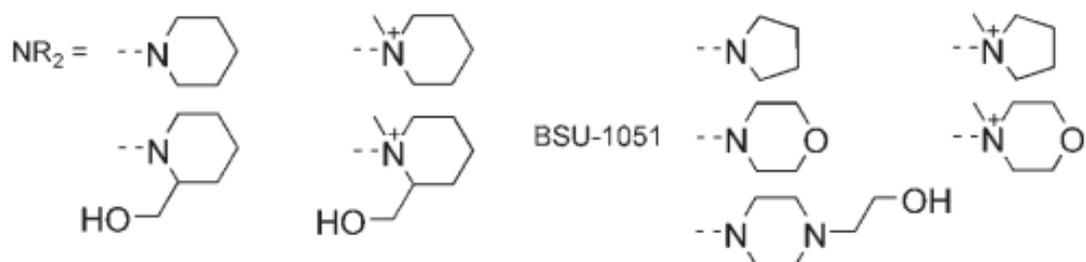


Figura 2-2 Leganti del G-quadruplex: antrachinoni e derivati

L'importanza di poter introdurre una carica positiva nell'anello centrale della parte aromatica ha portato allo sviluppo di derivati dell'acridina, in particolare a una serie di 3,6-diammido acridine (Figura 2-3)¹⁵. Questi composti non hanno dimostrato un'attività d'inibizione molto più marcata degli antrachinoni, probabilmente perché i gruppi ammidici non hanno un elevato effetto elettron-donatore e quindi rendono l'atomo d'azoto dell'anello poco basico. In generale i leganti cationici di questo tipo mostrano una maggiore interazione con il G-quadruplex, ma probabilmente non abbastanza forte affinché questa si manifesti anche in ambienti a elevata forza ionica, come quello fisiologico. Sono spesso poco selettivi per il DNA G-quadruplex rispetto al DNA duplex¹⁶.

¹⁵ M. A. Read, A. A. Wood, J. R. Harrison, S. M. Gowan, L. R. Kelland, H. S. Dosanjh, S. Neidle, *J. Med. Chem.* 1999, 42, 4538–4546

¹⁶ B. Majian, S. Bhattacharya, *Chem. Commun.*, 2014, 50, 6422

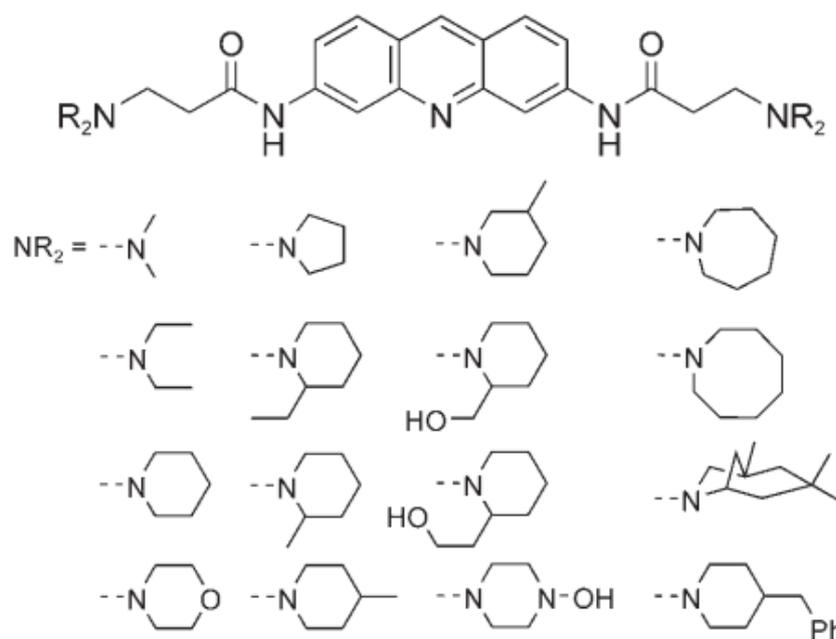


Figura 2-3 Leganti del G-quadruplex: acridine e derivati

Gli alcaloidi naturali sono un altro esempio di leganti planari. Sotto questa definizione rientrano moltissimi composti, alcuni dei quali hanno mostrato una buona affinità sia per il DNA duplex, sia per il DNA G-quadruplex. Tra di essi uno dei più studiati è la berberina, che dà interazione con entrambe le strutture, anche se è un legante migliore per il quadruplex. La berberina è un alcaloide, estratto da alcune piante del genere *Berberis*, *Hydrastis canadensis* e la *Coptis chinensis*. La capacità della berberina di inibire l'attività della telomerasi fu notata per la prima volta nel 1999¹⁷ e da allora ne sono stati sintetizzati e studiati vari derivati (Figura 2-4).

¹⁷ I. Naasani, H. Seimiya, T. Yamori, T. Tsuruo, Cancer Res.1999, 59, 4004–4011

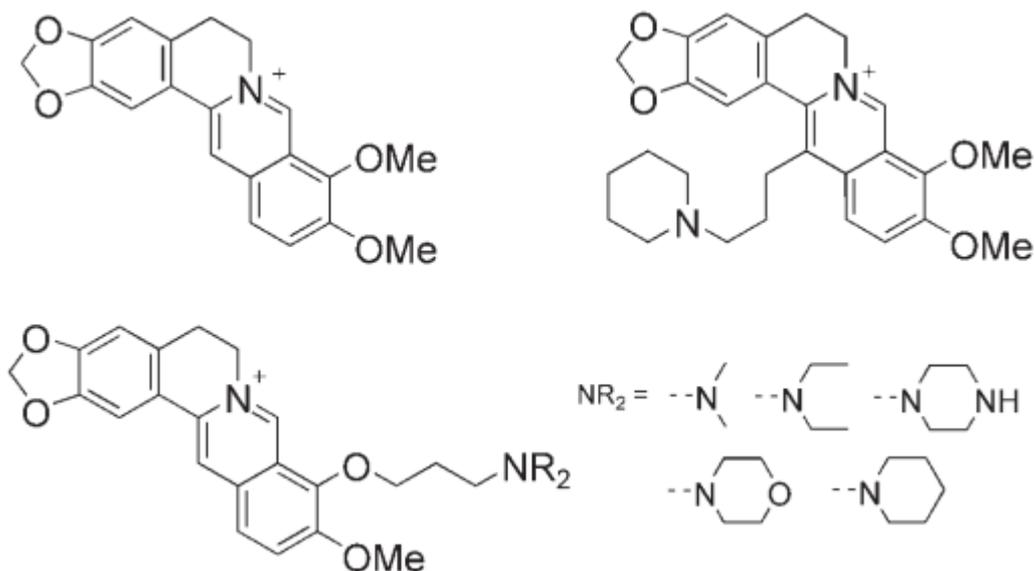


Figura 2-4 Leganti del g-quadruplex: berberina e derivati

2.1.1.2 Leganti macrociclici

Un legante macrociclico è definito come una catena ciclica polidentata, nel cui scheletro sono presenti almeno tre atomi donatori, che possono funzionare da siti di legame per il substrato¹⁸.

I leganti macrociclici costituiscono un'importante e vasta classe di leganti per il DNA G-quadruplex. Fa parte di questi la telomestatina, una sostanza naturale, considerata uno dei migliori leganti per il quadruplex. Fu isolata dal batterio *Streptomyces anulatus* nel 2001 da Shin-ya e collaboratori. Ha due isomeri, tra i quali l'isomero S ha la migliore attività d'inibizione della telomerasi (Figura 2-5).

¹⁸ Pedersen, C.L. *Angew. Chem. Int.* 1988,27,1021-1027

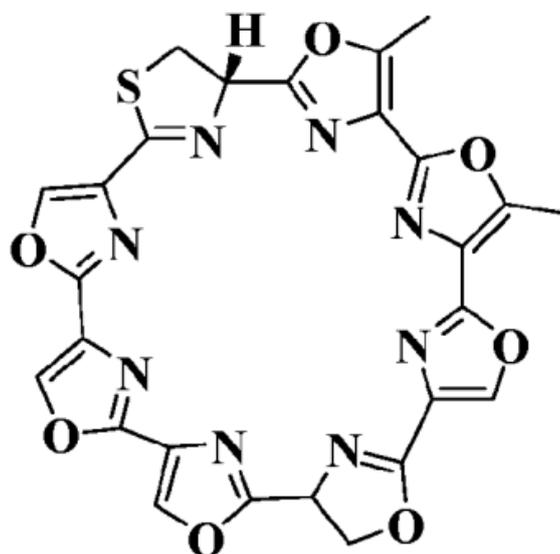


Figura 2-5 Telomestatina

Un altro esempio di macrocicli in grado di legare il DNA sono le porfirine. Le porfirine sono leganti non selettivi, che interagiscono con tutti gli acidi nucleici. Riescono a legare il DNA a doppia elica e si pensa che la disposizione planare degli anelli aromatici possa favorire un'interazione con il G-quartet. Rientrano in questa categoria di composti i leganti TMPyP4 e il suo isomero TMPyP2 (Figura 2-6). Entrambi si legano sia al DNA duplex sia alle strutture G-quadruplex, anche se con una selettività minore per quest'ultime. Il TMPyP4 mostra una migliore inibizione della telomerasi rispetto al suo isomero¹⁹.

¹⁹ R. T. Wheelhouse, D. Sun, H. Han, F. X. Han, L. H. Hurley, J. Am. Chem. Soc. 1998, 120, 3261–3262.

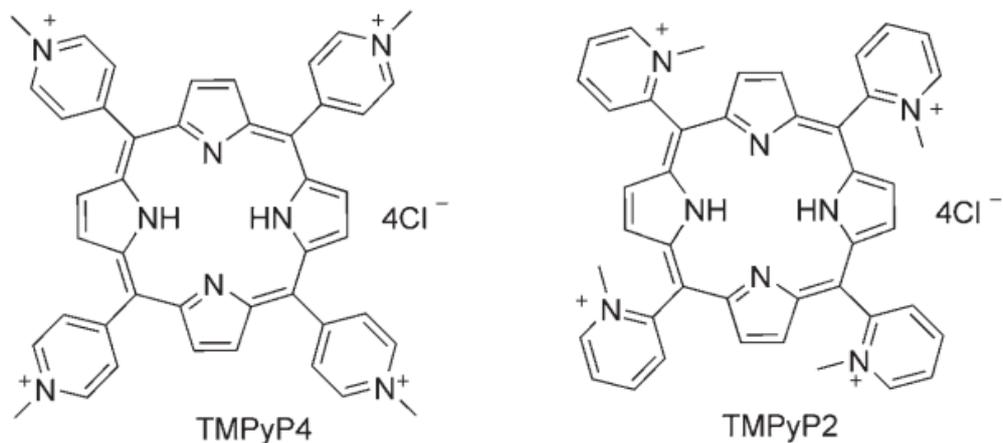


Figura 2-6 TMPyP4 e TMPyP2

2.1.1.3 Complessi metallici che legano il quadruplex

Sono stati sviluppati anche alcuni complessi metallici (Figure 2-7,8), che hanno un grande impatto sulle strutture G-quadruplex. Le modalità di interazione sono legate alla dimensione della molecola, alla sua planarità e alla natura dello ione metallico presente. Quest'ultimo può sostituire gli ioni Na⁺ o K⁺, che normalmente stabilizzano le strutture G-quadruplex, favorendo la formazione delle tetradi. Il metallo è elettron-attrattore e rende così il legame di coordinazione povero di elettroni, facendo così aumentare le interazioni π - π tra il complesso metallico e la tetrade²⁰. A una maggiore superficie che può dare interazioni di questo tipo corrisponde una maggiore inibizione della telomerasi²¹.

²⁰ A. Arola-Arnal, J. Benet-Buchholz, S. Neidle and R. Vilar, *Inorg. Chem.*, 2008, 47, 11910

²¹ Q. Yu, Y. Liu, J. Zhang, F. Yang, D. Sun, D. Liu, Y. Zhou and J. Liu, *Metallomics*, 2013, 5, 222

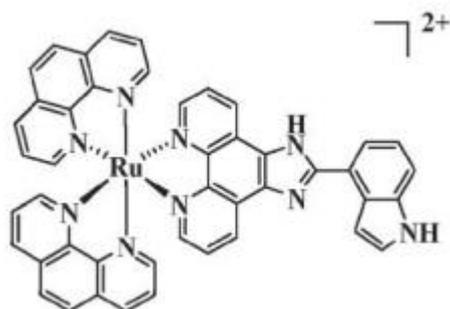


Figura 2-7 Complessi di Ru

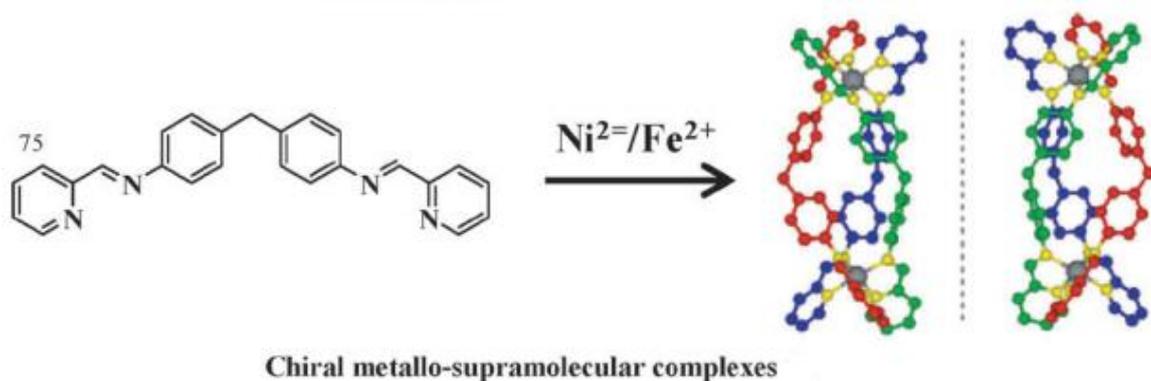


Figura 2-8 Complessi supramolecolari

2.2 L1, L2 e L3

I leganti oggetto di questo lavoro di tesi sono i tre leganti azotati, macrociclici, riportati nelle figure 2-9, 10, 11.

L1	2,5,17,20-tratraaza[12](6,6'')ciclo(2,2':6',2'')-bis-terpiridonofano
L2	2,5,7,17,22,25-esaaza[22](6,6'')ciclo(2,2':6',2'')-bis-terpiridinofano
L3	2,6,10,28,32,36-esaaza[22](6,6'')ciclo(2,2':6',2'')-bis-terpiridinofano

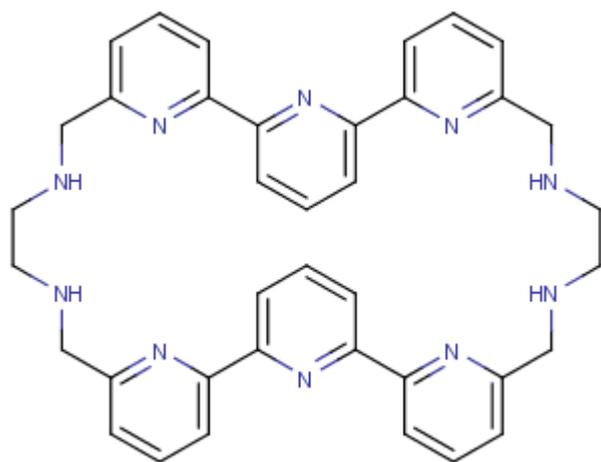


Figura 2-9 L1

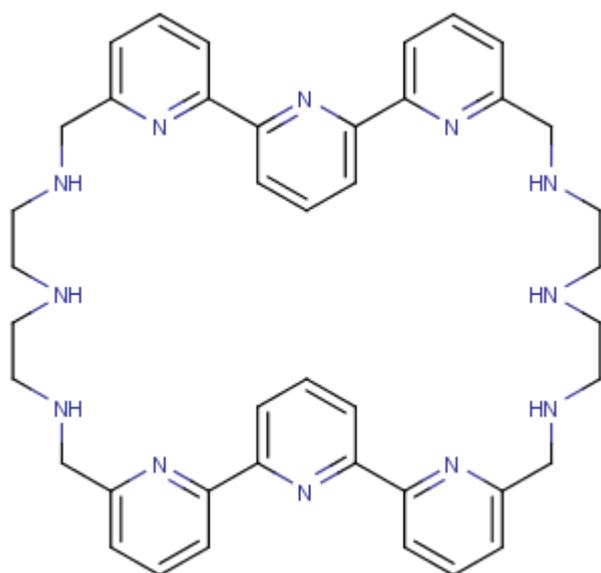


Figura 2-10 L2

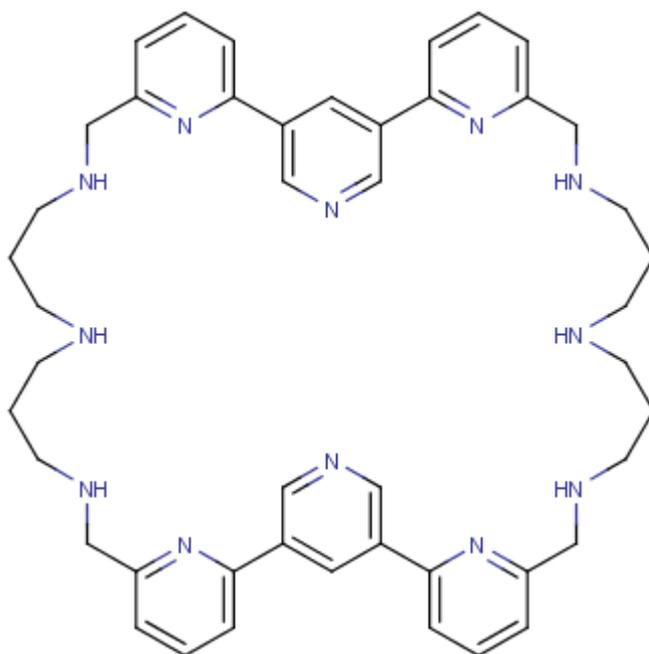


Figura 2-11 L3

Si tratta di composti macrociclici classificati come poliazacicloalcani, molecole cicliche che contengono atomi di azoto come atomi donatori, in grado di legarsi in soluzione acquosa sia a specie cationiche, coordinando cationi metallici, sia a specie anioniche, quando i loro atomi di azoto sono protonati.

L1, L2 e L3 erano stati in precedenza sintetizzati e in parte caratterizzati dal gruppo di ricerca del Prof. Bianchi²², sviluppati come leganti polifunzionali sia verso i cationi sia verso gli anioni; L3 è stato nuovamente sintetizzato per questo lavoro di tesi. Si tratta di leganti politopici simmetrici che presentano zone coordinative con caratteristiche significativamente diverse. In ciascun composto due gruppi terpiridile sono separati da due catene alifatiche poliazotate uguali. Nei tre leganti le catene hanno lunghezze differenti e contengono un diverso numero di atomi donatori, separati da gruppi etilenici, L1 e L2, o

²² Recettori polifunzionali contenenti unità eteroaromatiche per il riconoscimento di substrati, Tesi di Dottorato in Scienze Chimiche del Dott. Samuele Santarelli

propilenici, L3. Queste differenze determinano le proprietà acido-base dei leganti, che si comportano tutti da basi polifunzionali in soluzione acquosa (Tabelle 2-1,2,3)

Equilibri	pKa
$[L1H_2]^{2+} + H^+ = [L1H_3]^{3+}$	7.69
$[L1H_3]^{3+} + H^+ = [L1H_4]^{4+}$	5.44
$[L1H_4]^{4+} + H^+ = [L1H_5]^{5+}$	4.89
$[L1H_5]^{5+} + H^+ = [L1H_6]^{6+}$	2.91

Tabella 2-1 pKa delle costanti di protonazione di L1, in rosso l'equilibrio relativo alla specie più rilevante a pH fisiologico

Equilibri	pKa
$L2H^+ + H^+ = [L2H_2]^{2+}$	8.34
$[L2H_2]^{2+} + H^+ = [L2H_3]^{3+}$	7.68
$[L2H_3]^{3+} + H^+ = [L2H_4]^{4+}$	6.82
$[L2H_4]^{4+} + H^+ = [L2H_5]^{5+}$	3.07
$[L2H_5]^{5+} + H^+ = [L2H_6]^{6+}$	2.53
$[L2H_6]^{6+} + H^+ = [L2H_7]^{7+}$	2.82

Tabella 2-2 pKa delle costanti di protonazione di L2, in rosso l'equilibrio relativo alla specie più rilevante a pH fisiologico

Equilibri	pKa
$[L3H_2]^{2+} + H^+ = [L3H_3]^{3+}$	7.94
$[L3H_3]^{3+} + H^+ = [L3H_4]^{4+}$	7.32
$[L3H_4]^{4+} + H^+ = [L3H_5]^{5+}$	6.57
$[L3H_5]^{5+} + H^+ = [L3H_6]^{6+}$	6.18
$[L3H_6]^{6+} + H^+ = [L3H_7]^{7+}$	4.84
$[L3H_7]^{7+} + H^+ = [L3H_8]^{8+}$	3.08

Tabella 2-3 pKa delle costanti di protonazione di L3, in rosso l'equilibrio relativo alla specie più rilevante a pH fisiologico

3 Obiettivi di questo lavoro di tesi

Le strutture G-quadruplex nel DNA telomerico umano e i leganti in grado di indurle e stabilizzarle sono da tempo studiati nell'ambito della lotta ai tumori. In questo lavoro di tesi sono state studiate le interazioni dei tre leganti macrociclici L1, L2 e L3 con sequenze di DNA telomerico umano.

I tre leganti sono stati scelti considerando che le loro caratteristiche strutturali li rendevano dei potenziali leganti del DNA e in particolare della struttura G-quadruplex. La presenza dei due gruppi terpiridile, infatti, garantisce la disponibilità di una superficie aromatica estesa, requisito fondamentale per l'ottenimento di forti interazioni di stacking con i G-quartet. I gruppi ammoniacali presenti a pH fisiologico sulle catene alifatiche garantiscono inoltre la possibilità di interazione di carica con le tetradi e/o con i gruppi fosfato localizzati nei solchi e nei loop della struttura quadruplex.

Studi preliminari condotti in soluzione hanno dimostrato la capacità dei tre leganti di interagire selettivamente con il folding quadruplex del DNA telomerico. In questo lavoro di tesi è stata sviluppata un'indagine strutturale sugli addotti formati da ciascun legante con sequenze telomeriche umane di lunghezza diversa, in grado di formare sia quadruplex monomolecolari sia bimolecolari (Tabella 3-1).

Tabella 3-1 Sequenze di DNA telomerico umano usate per la tesi

Nome	Sequenza	Tipo di quadruplex
Tel26	TTAGGGTTAGGGTTAGGGTTAGGGTT	Monomolecolare
Tel23	TTAGGGTTAGGGTTAGGGTTAGG	Monomolecolare
Tel22	TTAGGGTTAGGGTTAGGGTTAG	Monomolecolare
Tel12	TTAGGGTTAGGG	Bimolecolare

Lo studio è stato condotto predisponendo screening di cristallizzazione per le varie combinazioni DNA/legante. Le possibili caratteristiche degli addotti sono state inoltre studiate per mezzo di tecniche di modellistica molecolare.

4 Modellistica molecolare

4.1 Principi teorici

4.1.1 Meccanica molecolare

L'equazione di Schrödinger dipendente dal tempo può descrivere in maniera completa il moto di un sistema molecolare:

$$\hat{H}\Psi(\vec{r}, t) = \frac{-i\hbar}{2\pi} \frac{\partial}{\partial t} \Psi(\vec{r}, t)$$

Dove \hat{H} è dato dalla somma di energia cinetica e potenziale.

$$\hat{H} = T + V(\vec{r})$$

$$T = \sum_i \frac{-\hbar^2}{8\pi^2 m_i} \frac{\partial^2}{\partial r_i^2}$$

$$V(\vec{r}) = V(x_1, y_1, z_1, \dots, x_N, y_N, z_N)$$

L'energia potenziale rappresenta l'interazione tra le particelle.

L'utilizzo di una trattazione quantistica di questo tipo per sistemi estesi come quelli biologici è resa difficile dall'elevato numero di atomi che li compongono. Si preferiscono quindi spesso i metodi approssimati come la meccanica molecolare. Questi metodi si basano sull'approssimazione di Born-Oppenheimer, che permette di disaccoppiare i moti elettronici e nucleari. I nuclei, infatti, hanno massa molto maggiore degli elettroni, tanto che si può affermare che i nuclei siano fermi rispetto al moto degli elettroni. Separando le variabili dei due moti, possiamo considerare l'energia del sistema come funzione delle sole coordinate nucleari e trascurare il moto degli elettroni. Il sistema sarà quindi descritto in termini di masse puntiformi che si muovono all'interno del campo generato dagli elettroni. La meccanica molecolare non è in grado di descrivere fenomeni come la formazione e la rottura dei legami, perché non prende in considerazione il moto degli elettroni, mentre è una buona approssimazione per studi strutturali e dinamici.

4.1.2 Campo di forza empirico

Nei metodi di meccanica molecolare, l'energia della molecola, descritta secondo le leggi della fisica classica, è data da due contributi: contributi di legame, dovuti alle deviazioni dalle distanze di legame ideali, dagli angoli di legame ideali e dagli angoli torsionali ideali, e contributi di non legame, dovuti a interazioni di Van der Waals, coulombiane e ai legami a idrogeno. La funzione di potenziale utilizzata per descrivere le interazioni tra atomi è chiamata campo di forza ed è appunto data dalla somma di contributi di legame e di non legame:

$$V = V_{bonded} + V_{non\ bonded}$$

V_{bonded} e $V_{non\ bonded}$ possono essere ulteriormente scomposti:

$$\begin{aligned} V_{bonded} &= V_{stretching} + V_{bending} + V_{torsion} \\ V_{non\ bonded} &= V_{Van\ der\ Waals} + V_{Coulomb} \end{aligned}$$

I campi di forza più utilizzati sono riconducibili a uno schema di questo tipo:

$$\begin{aligned} V(\vec{r}) = & \sum_{legami} \frac{k_{s,i}}{2} (l_i - l_{i,0})^2 + \sum_{angoli} \frac{k_{b,i}}{2} (\vartheta_i - \vartheta_{i,0})^2 \\ & + \sum_{torsioni} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\ & + \sum_{i=1}^N \sum_{j=i+1}^N \left(4 \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \end{aligned}$$

\vec{r} è un vettore di dimensioni 3N e $V(\vec{r})$ indica che l'energia potenziale è una funzione delle tre coordinate $x_1, y_1, z_1, \dots, x_N, y_N, z_N$ degli N atomi del sistema²³.

²³ S.N. Rao, U.C. Singh and P.A. Kollman, Nature 328, 1987,551

Il primo addendo dà conto dell'aumento di energia, dovuto alla deviazione della lunghezza di legame l_i dal valore di riferimento $l_{i,0}$. La curva di energia potenziale $V_{\text{stretching}}$ è data dall'approssimazione armonica, che usa come modello teorico l'oscillatore armonico. Questa approssimazione prevede di schematizzare il legame tra due atomi come due masse collegate da una molla, che ha una forza di richiamo data dalla legge di Hooke:

$$\vec{F} = -k\vec{x}$$

L'energia potenziale, nei pressi di un minimo di energia, è data da uno sviluppo in serie di Taylor in cui ci si ferma al termine in derivata seconda:

$$V(x) = V_0 + \left(\frac{\partial V}{\partial x}\right)_0 x + \frac{1}{2} \left(\frac{\partial^2 V}{\partial x^2}\right)_0 x^2$$

Poiché il potenziale è l'integrale della forza, sarà definito a meno di una costante. Poniamo quindi V_0 costante e nulla. Anche il secondo termine è nullo, perché è la derivata prima in un punto di minimo. Il potenziale si riduce quindi a :

$$V(x) = \frac{1}{2} \left(\frac{\partial^2 V}{\partial x^2}\right)_0 x^2 = \frac{1}{2} kx^2$$

Corrisponde all'energia potenziale di una molla e ha come curva una parabola. Si possono utilizzare anche altri potenziali, perché l'oscillatore armonico si discosta dal caso reale per due motivi: permette alle masse di sovrapporsi, mentre invece il legame tra due atomi oltre un certo valore diventa incompressibile, perché le due cariche positive dei nuclei si respingono, e non prevede la rottura del legame. Per ovviare a questi problemi, bisogna introdurre ulteriori termini nello sviluppo del potenziale, passando a un oscillatore anarmonico. In questo modo s'introducono un asintoto verticale, che tiene conto dell'incompressibilità del legame, e un asintoto orizzontale del valore dell'energia di dissociazione, oltre la quale la molecola si dissocia e i frammenti si muovono come

particelle libere. Sono stati sviluppati vari potenziali per descrivere una curva di questo tipo; tra di essi uno dei più utilizzati è il potenziale di Morse (Figura 4-1).

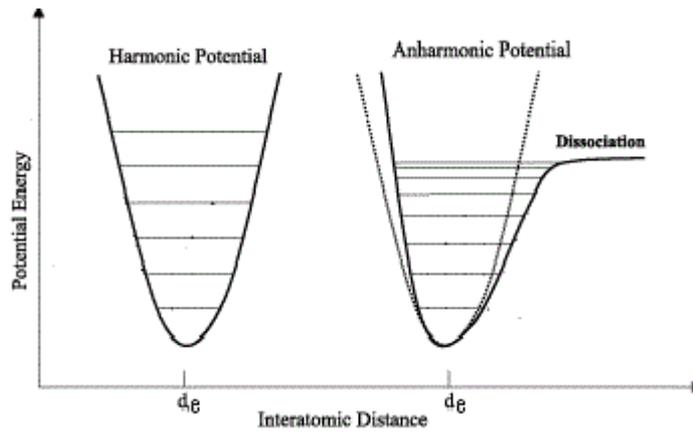


Figura 4-1 Potenziale armonico e anarmonico

L'addendo successivo è analogo al primo, è sempre un termine armonico, e tiene conto delle deviazioni degli angoli di legame ϑ_i (Figura 4-2) dai loro valori di riferimento $\vartheta_{i,0}$. Il contributo fornito da ogni angolo è caratterizzato da due parametri: la costante di forza $k_{b,i}$ e il valore di riferimento $\vartheta_{i,0}$. Le costanti di forza dei bending sono generalmente minori di quelle degli stretching, in quanto distorcere un angolo richiede meno energia che modificare la lunghezza di legame.

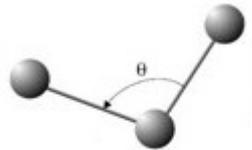


Figura 4-2 Angolo di legame

Il terzo termine mostra come varia l'energia potenziale al variare degli angoli torsionali (Figura 4-3). I potenziali torsionali sono generalmente espressi come espansioni in serie della funzione coseno e vengono comunemente troncati al primo termine. V_n è un parametro che indica la soglia di potenziale da superare per la rotazione, mentre n indica il numero di minimi della funzione V_{torsion} nell'intervallo di ω , che va da 0° a 360° ; γ rappresenta, invece, il fattore di fase e determina dove la funzione potenziale torsionale presenta i minimi. Tutti i parametri inerenti alle interazioni di legame possono essere ricavati da calcoli quantomeccanici ab initio o da dati sperimentali, come strutture cristallografiche, misure NMR, dati spettroscopici e così via.

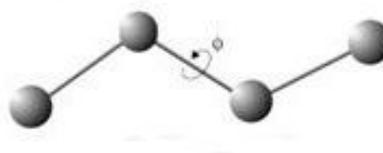


Figura 4-3 Angolo torsionale

L'ultimo addendo, $V_{\text{non bonded}}$, rende conto delle interazioni di non legame; è diviso in due contributi, che dipendono dalle interazioni di Van der Waals e dalle interazioni elettrostatiche. Le interazioni di Van der Waals sono il risultato di un bilanciamento di forze attrattive e repulsive, dovute entrambe a effetti quantistici. Sono determinate dai dipoli istantanei che si generano nella molecola a causa delle fluttuazioni degli elettroni della molecola; sono anche dette forze di dispersione di London. Le forze repulsive si manifestano quando le distanze interatomiche diminuiscono oltre un determinato limite e impediscono che le forze attrattive diventino infinitamente grandi; sono chiamate anche forze di scambio e sono un effetto del principio di esclusione di Pauli. Per esprimere le interazioni di Van der Waals, si possono utilizzare varie funzioni; una delle più utilizzate è la funzione 12-6 di Lennard-Jones (Figura 4-4).

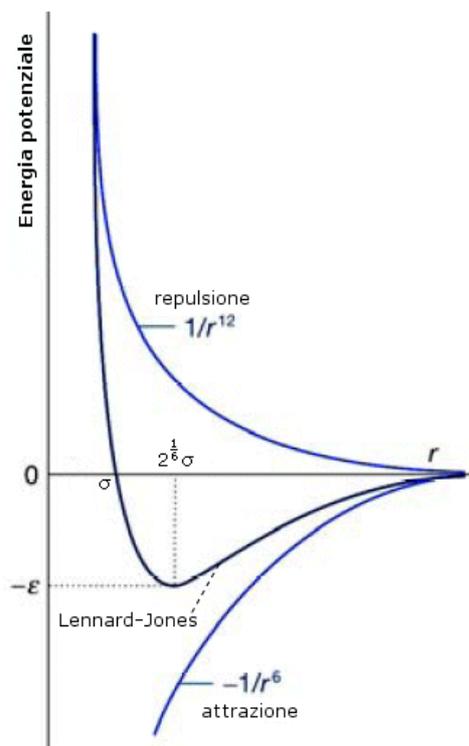


Figura 4-4 Potenziale di Lennard-Jones

La parte attrattiva varia come r^{-6} , mentre la parte repulsiva varia come r^{-12} . Nella funzione sono presenti due parametri ottimizzabili: σ , il diametro di collisione, che è la distanza a cui l'energia di interazione è minima, e ϵ , che è la profondità della buca di potenziale. Si possono ricavare questi parametri da misure sperimentali.

Le interazioni elettrostatiche sono descritte da un termine coulombiano. A ogni atomo del sistema viene attribuita una carica parziale, che può essere ricavata da calcoli quantomeccanici o essere semplicemente un parametro del campo di forza.

Esistono campi di forza più sofisticati, che contengono termini aggiuntivi, per esempio dei termini per considerare separatamente i legami a idrogeno, ma prevedono sempre l'utilizzo dei quattro addendi appena descritti.

I campi di forza usati nelle applicazioni di modellistica molecolare sono studiati per riprodurre proprietà strutturali o predire altre proprietà di una molecola e ogni campo di

forze è progettato, e di conseguenza i suoi parametri ottimizzati, per predire determinate proprietà. Un campo di forza è trasferibile se un certo set di parametri può essere usato per un ampio spettro di tipologie molecolari. In generale si può utilizzare uno stesso campo di forza per molecole della stessa classe. Esistono molti campi di forza, sviluppati e parametrizzati per le diverse classi di molecole, come le proteine o gli acidi nucleici. In questo lavoro di tesi è stato utilizzato un campo di forza OPLS, Optimized Potential for Liquid Simulation, sviluppato da W.L. Jorgensen e J. Tirado-Rives²⁴.

4.1.2.1 Atom type

Un concetto molto importante nelle tecniche che utilizzano i campi di forza empirici è quello di atom type. L'atom type definisce tutte le informazioni riguardo alla specie atomica, il suo stato di ibridazione e l'intorno chimico. La carica associata a un certo atomo in un determinato intorno chimico può, talvolta, far parte della definizione dell'atom type. Ogni parametro del campo di forza è espresso in funzione dei diversi atom type.

4.1.3 Analisi conformazionale

In generale una molecola può assumere diverse conformazioni. Esse influenzano le proprietà chimico-fisiche di una molecola. L'analisi conformazionale consiste appunto nello studio delle conformazioni che una molecola può assumere e di come queste influenzino le sue proprietà. Lo sviluppo dell'analisi conformazionale moderna si può riferire a D. D. R. Burton, che nel 1950 dimostrò che la reattività del cicloesano sostituito era determinata dalla posizione assiale o equatoriale dei suoi sostituenti.

Negli studi di modellistica molecolare è utile ricercare le conformazioni che sono più stabili dal punto di vista energetico, che corrispondono cioè a punti di minimo.

La superficie di energia potenziale è solitamente molto corrugata, cioè contiene vari punti di minimo e di massimo. I punti che hanno energia minima o massima sono detti punti

²⁴ W.L. Jorgensen e J. Tirado-Rives, J. Am. Chem. Soc. 110, 1988, 1657-1666

stazionari; un punto di sella corrisponde invece a una struttura di transizione²⁵. Il punto che ha energia più bassa di tutti gli altri è detto minimo globale.

La procedura di conformational search consiste nell'identificazione delle conformazioni preferite dalla molecola ed è il cardine dell'analisi conformazionale. Permette di localizzare le conformazioni in corrispondenza delle quali la superficie di energia potenziale presenta dei minimi. La ricerca conformazionale viene condotta attraverso vari metodi, in primo luogo si può esplorare lo spazio conformazionale in maniera sistematica o casuale. Con un approccio sistematico, si ruotano di un valore fissato tutti i legami della molecola che possono farlo, lasciando però inalterati le lunghezze e gli angoli di legame. Con un approccio random, invece, si esplora la superficie del potenziale in maniera non prevedibile. Uno dei metodi più utilizzati per questo tipo di calcolo è il metodo Montecarlo, un metodo numerico basato su procedimenti probabilistici, utile per la risoluzione di problemi di varia natura, che non possono essere affrontati per via analitica. Il metodo si basa sulla possibilità di eseguire, utilizzando numeri estratti a caso, un campionamento di una distribuzione di probabilità assegnata, $F(\mathbf{x})$; ossia sulla possibilità di generare una sequenza di eventi x_1, x_2, \dots, x_n , distribuiti secondo la $F(\mathbf{x})$. Non si serve di un campione di numeri effettivamente estratti a caso, ma fa ricorso a una sequenza di numeri ottenuti con un processo iterativo ben determinato; tali numeri vengono detti pseudo-casuali perché, pur non essendo casuali, hanno proprietà statistiche analoghe a quelle dei veri numeri casuali. Non esiste un solo metodo Montecarlo, ma una serie di approcci che seguono un determinato schema. Per prima cosa si definisce un dominio di possibili dati di input e si generano input casuali dal dominio con una certa distribuzione di probabilità determinate. Si esegue poi un calcolo deterministico in base ai dati ottenuti e si aggregano i risultati dei calcoli singoli nel risultato finale.

²⁵ Un punto di sella di una funzione reale di più variabili reali $f: \mathbb{R}^n \rightarrow \mathbb{R}$ è un punto critico P del dominio f in cui la matrice hessiana risulta indefinita.

Lo spazio conformazionale può essere simulato anche per mezzo della dinamica molecolare. In una simulazione di dinamica molecolare si generano configurazioni successive del sistema per integrazione della seconda legge di Newton:

$$F_{x_i} = -\frac{\partial V}{\partial x_i} = m_i \frac{d^2 x_i}{d^2 t}$$

Integrando questa equazione per ogni atomo, si può simulare l'evoluzione nel tempo del sistema di atomi descritto dalla funzione energia potenziale. Poiché non è possibile risolvere per via analitica queste equazioni differenziali, si ricorre a dei metodi numerici, come l'algoritmo di Verlet, che somma opportunamente gli sviluppi in serie di Taylor delle coordinate $x_i(t + \Delta t)$ e $x_i(t - \Delta t)$:

$$x_i(t + \Delta t) \approx 2x_i(t) - x_i(t - \Delta t) + \frac{F_{x_i}(t)}{m_i} \Delta t^2$$

Δt è il passo di integrazione e F_x si ricava dalla seconda legge di Newton.

Per prima cosa un programma di dinamica molecolare legge le coordinate iniziali degli atomi e calcola il potenziale di interazione V , in base al campo di forza scelto. Una volta fissata la temperatura a cui eseguire la simulazione, assegna in modo casuale le velocità iniziali degli atomi secondo l'equazione di Maxwell-Boltzmann:

$$p(v_{ix}) = \left(\frac{m_i}{2\pi k_B T} \right)^{\frac{1}{2}} \exp \left[-\frac{m_i v_{ix}^2}{2k_B T} \right]$$

Dove $p(v_{ix})$ è la probabilità che l'atomo i di massa m_i abbia la velocità v_{ix} nella direzione x della temperatura T .

Si avvia poi un ciclo di calcolo, nel quale vengono calcolate le forze agenti su ciascun atomo date dalla legge di Newton, che vengono integrate con l'algoritmo di Verlet per ottenere le posizioni e le velocità al tempo successivo.

4.1.3.1 Cluster

Gli algoritmi di ricerca conformazionali generano spesso un insieme di conformazioni molto simili, se non identiche, tra loro. Può essere utile selezionare un sottoinsieme di conformazioni rappresentative a partire dall'insieme iniziale. L'analisi permette di suddividere l'insieme di conformazioni in raggruppamenti di elementi simili, i cluster, tra cui scegliere il membro più rappresentativo.

Per la suddivisione di un insieme di conformazioni in cluster, sono stati sviluppati molti algoritmi e non esiste ancora un unico metodo corretto. Può ad esempio essere valutata la somiglianza tra coppie di conformazioni diverse in termini di distanza quadratica media (RMSD) su coppie di atomi corrispondenti o su angoli torsionali.

4.1.4 Minimizzazione

Il campo di forza, così come è stato descritto, viene utilizzato in meccanica molecolare per costruire la superficie di energia potenziale della molecola.

Se si escludono i sistemi più semplici, in generale non è possibile visualizzare la superficie di energia potenziale, perché la funzione energia potenziale è una funzione multidimensionale delle coordinate di tutti gli atomi della molecola. Se la molecola è composta da N atomi, V è una funzione di $3N$ coordinate cartesiane o $3N-6$ coordinate interne:

$$V = V(\vec{r}) = V(x_1, y_1, z_1, \dots, x_N, y_N, z_N)$$

Identificare le conformazioni più stabili equivale a individuare i punti di minimo sulla superficie di potenziale. Una configurazione molecolare si trova in un minimo locale quando sono verificate le condizioni:

$$\sum_{i=1}^N \frac{\partial V}{\partial x_i} = 0$$

$$\sum_{i=1}^N \frac{\partial^2 V}{\partial x_i^2} = 0$$

ossia quando le derivate prima e seconda sono nulle.

Quando si ha a che fare con un N grande e una funzione V complessa, non è possibile trovare i minimi per via analitica e bisogna ricorrere a dei metodi numerici. Questi metodi prevedono di variare gradualmente le coordinate molecolari, ottenendo conformazioni con energie via via più basse, fino a raggiungere il minimo. Tutti gli algoritmi di minimizzazione richiedono un set di coordinate di partenza. Per identificare questi minimi iniziali, è necessario generare delle conformazioni di partenza tramite metodi di dinamica molecolare o simulazione Montecarlo; le conformazioni così generate vengono poi sottoposte alla vera e propria minimizzazione. Non esiste ancora un algoritmo che possa garantire la localizzazione del minimo globale da un punto iniziale arbitrario.

In generale i metodi di minimizzazione basati sulle derivate sono i più utilizzati. La derivata prima rappresenta il gradiente di energia, ha come direzione quella che indica la localizzazione del minimo e ha un'intensità proporzionale alla profondità di quest'ultimo. La derivata seconda indica, invece, la curvatura della funzione.

I metodi possono essere classificati in base all'ordine più alto della derivata usato; è, infatti, possibile espandere in serie di Taylor la funzione $V(\vec{r})$ intorno al punto di coordinate \vec{r}_k :

$$V(\vec{r}) = V(\vec{r}_k) + (\vec{r} - \vec{r}_k)V' + (\vec{r} - \vec{r}_k)^T \frac{V''(\vec{r}_k)(\vec{r} - \vec{r}_k)}{2} + \dots$$

I metodi di minimizzazione più utilizzati nella modellistica molecolare sono metodi del primo ordine e usano solamente il gradiente di potenziale; tra questi gli algoritmi più utilizzati sono lo steepest descent e il conjugate gradient.

In entrambi i metodi, le coordinate degli atomi sono variate in maniera graduale e iterativa fino al raggiungimento di un punto di minimo. La prima iterazione si fa a partire da una conformazione fornita da un altro metodo o da un dato sperimentale, poi le successive

iterazioni vengono condotte sulla base delle coordinate ottenute dall'iterazione precedente.

Nel metodo dello steepest descent, a partire da una geometria molecolare iniziale, definita dalle coordinate \vec{r}_1 , si calcola il gradiente:

$$\vec{f}_1 = \nabla V(\vec{r}_1)$$

Il gradiente definisce la direzione lungo la quale l'energia varia più rapidamente. Le coordinate degli atomi vengono poi mosse in questa direzione di un passo λ_1 , ottenendo così le nuove coordinate del sistema:

$$\vec{r}_2 = \vec{r}_1 - \vec{f}_1 \lambda_1$$

A partire dalla nuova geometria si calcola il nuovo gradiente:

$$\vec{f}_2 = \nabla V(\vec{r}_2)$$

Si muovono poi di un passo λ_2 le coordinate del sistema fino a raggiungere il punto:

$$\vec{r}_3 = \vec{r}_2 - \vec{f}_2 \lambda_2$$

Nella maggior parte delle applicazioni di modellistica molecolare, l'ampiezza del passo i -esimo, λ_i , è predefinita. Se l' i -esima iterazione porta a una riduzione dell'energia, l'ampiezza del passo viene aumentata di un fattore moltiplicativo $a > 1$ e $\lambda_{i+1} > \lambda_i$. Se invece si riscontra un aumento di energia, il passo viene ridotto e $\lambda_{i+1} < \lambda_i$. L'ampiezza del passo dipende in generale dalla natura della superficie di energia: si preferisce un passo ampio per superficie con pendenze lievi, mentre è preferibile un passo più corto se la superficie è corrugata.

Il metodo steepest descent è veloce e robusto, ma non è molto efficiente, avrà in particolare un andamento oscillatorio intorno al minimo, quando la valle della superficie di potenziale è molto stretta.

Il metodo del conjugate gradient converge più velocemente del metodo steepest descent, perché tiene maggiormente conto del percorso di minimizzazione già compiuto e guida quindi la minimizzazione utilizzando sia il gradiente attuale sia la direzione di ricerca pregressa. I suoi primi passaggi sono identici a quelli del metodo steepest descent, ma il punto \vec{r}_2 viene individuato lungo la direzione data da:

$$\vec{s}_2 = -\vec{f}_2 - \beta_2 \vec{s}_1$$
$$\beta_2 = \frac{f_2 \cdot f_2}{f_1 \cdot f_1}$$

La direzione di ricerca \vec{s}_k nel punto \vec{r}_k non è la direzione del gradiente \vec{f}_k , bensì la direzione data dalla combinazione lineare del nuovo gradiente e di quello che indicava la direzione di ricerca:

$$\vec{s}_k = -\vec{f}_k - \beta_k \vec{s}_{k-1}$$
$$\beta_k = \frac{f_k \cdot f_k}{f_{k-1} \cdot f_{k-1}}$$

con $k > 1$.

Anche in questo metodo gli atomi vengono spostati di un'ampiezza λ_k variabile.

4.2 Materiali e metodi

Le simulazioni di modellistica molecolare di questo lavoro di tesi sono state svolte presso il laboratorio di Molecular Modeling, Chemoinformatics & QSAR del Dipartimento Neuroscienze, Psicologia, Area del Farmaco e Salute del Bambino (NEUROFARBA)-sezione di farmaceutica e nutraceutica.

4.2.1 Materiali

Per lo svolgimento di questo lavoro di tesi, è stato utilizzato il seguente hardware:

- Workstation mm14 SO Linux (distribution: LINUX 2.6.9-42.0.10.ELsmp), Dual Core AMD Opteron 4 processori, OS:, CPU: i686.

È stato utilizzato il software

- Schrödinger package Suite2011: Maestro²⁶, MacroModel²⁷, Jaguar²⁸

4.2.2 Metodi

4.2.2.1 Studio dei leganti

I leganti sono stati considerati come specie triprotonate, per L1 e L2, e tetraprotonate, per L3.

Poiché i calcoli sono stati effettuati su i tre leganti, è stata definita una procedura uniformemente a ciascuno di essi.

I leganti sono stati inizialmente disegnati utilizzando il programma Builder, contenuto all'interno del software Maestro, e quindi minimizzati utilizzando il campo di forza OPLS-2005 e un metodo conjugate gradient come algoritmo di minimizzazione. Le cariche atomiche sono state assegnate dal campo di forza tramite la definizione dell'atom type.

Dopo la minimizzazione, sono state condotte delle analisi conformazionali, utilizzando MacroModel. I tre leganti contengono due gruppi terpiridina. Questo gruppo può assumere tre diverse conformazioni, dette trans/trans, cis/trans e cis/cis, in funzione del valore assunto dai due angoli torsionali N-C-C-N (Figura 4-5)

²⁶ <http://www.schrodinger.com/Maestro>

²⁷ <http://www.schrodinger.com/MacroModel>

²⁸ Bochevarov, A.D.; Harder, E.; Hughes, T.F.; Greenwood, J.R.; Braden, D.A.; Philipp, D.M.; Rinaldo, D.; Halls, M.D.; Zhang, J.; Friesner, R.A., "Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences," *Int. J. Quantum Chem.*, **2013**, *113(18)*, 2110-2142

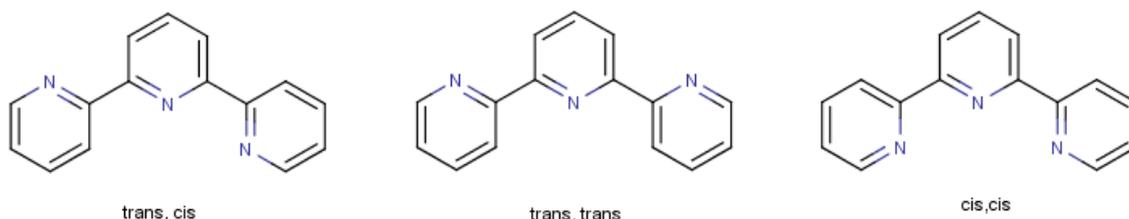


Figura 4-5 Conformazioni della terpiridina

Per questo motivo le analisi conformazionali sono state condotte a partire da tre configurazioni di partenza per ogni legante, corrispondenti alle tre conformazioni possibili per la terpiridina (Tabella 4-1). L'analisi conformazionale è stata condotta utilizzando la procedura Conformational Search, implementata nel programma Macromodel, che sfrutta il metodo Montecarlo.

Tabella 4-1

Legante	Conformazione della terpiridina	Nome
L1	cis/cis	L1_3
L1	trans/cis	L1_21
L1	trans/trans	L1_111
L2	cis/cis	L2_3
L2	trans/cis	L2_21
L2	trans/trans	L2_111
L3	cis/cis	L3_3
L3	cis/trans	L3_21
L3	trans/trans	L3_111

Dopo l'analisi conformazionale, le strutture risultanti sono state raccolte in famiglie utilizzando la routine X-Cluster della suite Schrodinger, presente all'interno dello stesso

Maestro. Sono state selezionate le conformazioni più rappresentative, sulle quali è stata condotta un'ottimizzazione quantomeccanica sulle frequenze vibrazionali, utilizzando il programma Jaguar.

Per ogni legante è stata scelta la conformazione a minore energia, che è stata utilizzata per gli studi delle modalità di interazione con i recettori.

4.2.2.2 Studio delle modalità di interazione del legante con il DNA telomerico umano

Gli addotti DNA/legante sono stati costruiti usando come coordinate iniziali quelle disponibili nella banca dati Protein Databank per le sequenze Tel26 (PDB refcode: 2HY9) e wtTel26 (PDB refcode:2JPZ), che corrispondono alle conformazioni cosiddette ibride (3+1) e Tel12 (PDB refcode: 4P1D) che corrisponde alla conformazione cosiddetta propeller. Come coordinate iniziali dei leganti sono state utilizzate quelle ottenute per mezzo dell'analisi conformazionale sopra descritta.

Su tutti i recettori è stata condotta una procedura di minimizzazione, prima di passare allo studio delle modalità di interazione con i leganti. Le condizioni di minimizzazione prevedevano di utilizzare il metodo conjugate gradient. Sono stati imposti dei vincoli sugli atomi dei residui di guanina delle tetradi e sugli ioni K⁺, mentre i tratti TTA e i residui timina e adenina iniziali e terminali sono stati lasciati liberi di muoversi. La costante di forza per i vincoli aggiuntivi è stata posta pari a 100.

Su ogni recettore sono state individuate sei possibili regioni di interazione per il legante: due zone corrispondenti rispettivamente al quartetto superiore e inferiore, e quattro zone laterali, che corrispondono ai quattro lati del quadruplex.

Per ogni recettore e ogni legante sono stati così definiti sei addotti iniziali, in cui il legante occupa una delle varie zone, per un totale di cinquantaquattro addotti iniziali, diciotto per legante.

Su ogni addotto sono stati eseguiti due passaggi: prima una minimizzazione con metodo conjugate gradient, poi una simulazione di dinamica molecolare con MacroModel. La minimizzazione è stata svolta con costante dielettrica dipendente dalla distanza ($\epsilon=4r$).

Le simulazioni di dinamica molecolare condotte sugli addotti minimizzati sono state svolte alla temperatura di 500 K usando gli stessi vincoli e trattamento elettrostatico della minimizzazione.

Il sistema è stato lasciato equilibrare per 5 ps, mentre la simulazione è stata condotta per 500 ps, campionando una conformazione ogni 10 ps per un totale di 50 conformazioni. L'intervallo di integrazione è stato di 1.5 fs. Le strutture risultanti sono state automaticamente minimizzate con le stesse impostazioni usate per la minimizzazione preliminare.

Per ogni addotto è stata quindi stimata l'energia di interazione tra DNA telomerico e leganti, secondo la formula:

$$\Delta E = E_{addotto} - (E_{legante} + E_{DNA})$$

5 Cristallografia e cristallizzazione

5.1 Principi teorici

5.1.1 Cristallografia

Un cristallo è definito come un corpo solido in cui gli atomi sono disposti in modo ordinato e periodico, oscillando intorno a definite posizioni nello spazio. Questa periodicità definisce le proprietà fisiche macroscopiche del cristallo e gli permette di diffrangere i raggi X.

Si ha diffrazione quando una radiazione incontra ostacoli di grandezza comparabile con la sua lunghezza d'onda; per il principio di Huygens l'ostacolo diventa una nuova sorgente di radiazione. I raggi X hanno una lunghezza d'onda dell'ordine dell'Åström, ossia della stessa entità delle distanze interatomiche (Figura. 5-1). Ogni atomo del reticolo diventa così una nuova sorgente puntuale di radiazione.

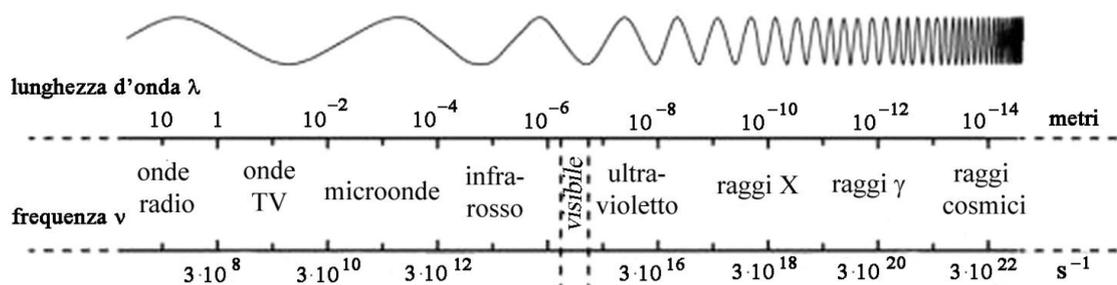


Figura 5-1 Spettro elettromagnetico

In un reticolo cristallino, gli atomi omologhi formano dei piani paralleli, che si trovano a una certa distanza interplanare l'uno dall'altro.

La legge di Bragg

$$n\lambda = 2d\sin\theta$$

afferma che si avrà interferenza positiva massima quando la differenza di cammino ottico tra i raggi riflessi da due diversi piani è pari a un numero intero n di lunghezze d'onda della radiazione incidente (Figura 5-2).

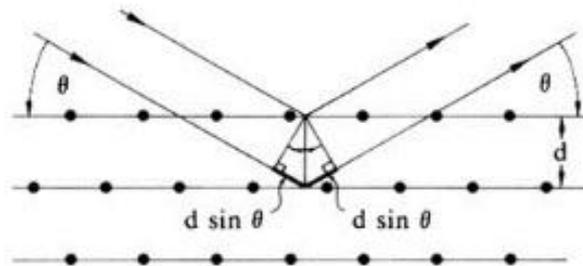


Figura 5-2 Rappresentazione dei parametri che definiscono la condizione di diffrazione secondo Bragg

I risultati diretti di una misura di diffrazione sono: i parametri di cella, informazioni sul gruppo spaziale, che descrive la simmetria del reticolo cristallino, e le intensità dei riflessi, cioè il quadrato dell'ampiezza dell'onda diffratta, dal quale sarà estratta l'informazione strutturale, sotto forma di fattore di struttura.

Il fattore di struttura è un numero complesso e la misura di diffrazione ne fornisce il modulo, ma non la fase, da qui il cosiddetto problema legato alla risoluzione vera e propria della struttura, detto problema della fase. Nel caso delle macromolecole biologiche viene spesso applicato il metodo di sostituzione molecolare²⁹. Si può applicare questo metodo solo quando si hanno a disposizione delle coordinate atomiche di una molecola avente una struttura analoga a quella in esame; in particolare è necessario che la molecola con coordinate note abbia una sequenza di residui omologa almeno al 40% a quella della molecola da studiare. Il modello strutturale a coordinate note viene posizionato e orientato all'interno della cella elementare del cristallo della molecola incognita, calcolando una funzione di rotazione e una di traslazione in cui si sovrappongono la funzione di Patterson,

²⁹ Drenth J. Principles of protein X-ray crystallography, Ed Springer-Verlag, New York, 213, 1994

una sorta di mappa vettoriale di tutti i vettori interatomici, calcolata per il modello e quella ottenuta dalle misure. Più sono simili le strutture, più è probabile che il problema della fase sia risolto.

5.1.2 Cristallografia e cristallizzazione di molecole biologiche

La cristallografia di molecole biologiche è molto importante, poiché permette di conoscere la struttura di proteine e di complessi legante-recettore.

I cristalli di macromolecole biologiche si differenziano dai cristalli di piccole molecole sotto molti aspetti. Hanno morfologia e dimensioni variabili tanto quanto i cristalli di piccole molecole, ma sono in generale più piccoli e raramente superano il volume di 1mm^3 . Gli acidi nucleici e le proteine, essendo costituiti rispettivamente da D-zuccheri e L-amminoacidi, cristallizzano sempre in gruppi spaziali chirali, cioè che non hanno né piani di riflessione né centro di inversione; inoltre le celle cristalline sono molto grandi e contengono spesso più entità non equivalenti nell'unità asimmetrica. Presentano fragilità ed elevata sensibilità alle condizioni esterne. Hanno un elevato contenuto di solvente e devono, perciò, essere conservati in un ambiente saturo di soluzione madre, per impedire la disidratazione che può portare alla rottura del cristallo.

La cristallizzazione è il processo di formazione di cristalli solidi a partire da una soluzione e comporta quindi un cambiamento di fase, dalla distribuzione disordinata delle particelle nella soluzione alla struttura ordinata e periodica del reticolo cristallino. Il passaggio da una situazione di disordine a una più ordinata causa una variazione negativa di entropia, che è però compensata da un contributo entalpico favorevole, perché nel cristallo si formano interazioni e legami chimici stabili. Si può razionalizzare il processo di cristallizzazione, assumendo che sia dovuto a due diversi contributi: la solubilità e la sovrasaturazione. A una data temperatura, un soluto può restare in soluzione solo se si trova al di sotto della concentrazione che corrisponde al suo limite di solubilità. La formazione di un precipitato, che sia o meno cristallino, avviene dopo che si è prodotto uno stato metastabile di

sovrasaturazione, nel quale la concentrazione del composto supera il limite di solubilità. Questa situazione è instabile e il sistema evolve cercando di ristabilire l'equilibrio tramite la precipitazione della sostanza. Il tempo necessario per stabilire l'equilibrio e il grado di sovrasaturazione possono variare molto da caso a caso, ma in generale si nota che le dimensioni delle particelle diminuiscono all'aumentare della sovrasaturazione media. Ciò avviene, perché la formazione del precipitato è dovuta a due processi competitivi tra loro: la nucleazione e l'accrescimento. Le dimensioni delle particelle sono determinate da quanto un processo prevale in velocità sull'altro. Il processo di nucleazione consiste nell'aggregazione di un numero minimo di molecole, che formano dei nuclei o germi cristallini. Durante la fase di accrescimento, invece, strati successivi di molecole si depositano sui nuclei precedentemente formati e contribuiscono ad aumentare le dimensioni del cristallo. L'accrescimento può essere suddiviso in due diverse fasi: inizialmente si ha trasporto delle molecole dalla massa della soluzione alla superficie del cristallo in base al gradiente di concentrazione, poi le molecole s'inseriscono ordinatamente all'interno del reticolo cristallino. L'accrescimento del cristallo può interrompersi per varie ragioni, come l'eccessiva nucleazione, brusche variazioni delle condizioni di cristallizzazione o l'inclusione di impurezze nel reticolo cristallino.

Per poter utilizzare dei cristalli in un'analisi cristallografica, essi devono essere singoli, ordinati e con dimensioni almeno dell'ordine del decimo di mm. La cristallizzazione di molecole biologiche è influenzata da un numero molto più elevato di variabili rispetto alle molecole più piccole, variabili che è necessario controllare con cura per poter ottenere cristalli con le caratteristiche che li rendono adatti a un'analisi cristallografica. Riveste quindi particolare importanza la determinazione delle condizioni di cristallizzazione. Per via delle caratteristiche chimico-fisiche e della flessibilità conformazionale delle molecole biologiche, il processo di cristallizzazione è molto sensibile alle variazioni di concentrazione, pH, temperatura e simili parametri.

A livello pratico i parametri che si possono variare sono:

1. Agenti precipitanti

a) Sali inorganici

I sali inorganici diminuiscono la solubilità della macromolecola e influenzano la forza ionica della soluzione. Secondo la teoria di Debye-Hückel, la presenza di un elettrolita nella soluzione acquosa di cristallizzazione fa sì che intorno a ogni specie carica si formi un'atmosfera ionica di carica opposta a quella dello ione centrale. L'effetto di quest'atmosfera dipende dalla concentrazione dell'elettrolita e quindi dalla variazione della forza ionica:

$$I = \frac{1}{2} \sum c_j z_j^2$$

c e z sono la concentrazione e la carica.

Quando la concentrazione dell'elettrolita supera un certo valore, l'atmosfera ionica diminuisce la solubilità della macromolecola, perché quest'ultima e l'elettrolita competono per le molecole d'acqua; si parla di effetto salting out. Un altro fattore da considerare è la solubilità in acqua del sale, che deve essere abbastanza elevata da impedirgli di precipitare prima della cristallizzazione della macromolecola. Tra i sali più utilizzati ci sono NaCl e KCl.

b) Solventi organici

Un solvente organico abbassa la costante dielettrica del solvente, facendo di conseguenza aumentare l'attrazione elettrostatica tra macromolecole. I solventi più utilizzati sono il 2-metil-2,4-pentadiolo, o MPD, e i polietilenglicoli, PEG; che hanno formula chimica $H-(O-CH_2-CH_2)_n-OH$ e peso molecolare tra 2000 e 6000 g/mol.

2. Condizioni chimico-fisiche

a) Concentrazione della macromolecola

La concentrazione della macromolecola dovrebbe essere tale da determinare una precipitazione abbastanza lenta, per evitare di ottenere un precipitato amorfo. D'altro canto una precipitazione molto lenta potrebbe portare alla formazione di quantità troppo esigue di precipitato.

In generale, non disponendo di una concentrazione teorica ideale, la si mantiene più alta possibile.

b) pH

Le macromolecole biologiche sono stabili in un intervallo limitato di pH intorno alla neutralità e quindi sono da evitare pH molto alti o molto bassi.

c) Temperatura

La temperatura influisce sulla solubilità delle macromolecole biologiche in maniera molto variabile e poco prevedibile, tanto che in letteratura sono riportate temperature di cristallizzazione tra -20°C e 40°C.

d) Ioni metallici

È stato notato che gli ioni metallici possono indurre o favorire la cristallizzazione di alcune macromolecole. Possono influire sulla forza ionica o aiutare a mantenere compatta la struttura cristallina.

5.1.3 Tecniche di cristallizzazione

Per la cristallizzazione di macromolecole biologiche, sono stati sviluppati vari metodi, che si basano principalmente sul portare la soluzione della macromolecola in uno stato di sovrasaturazione. La tecnica più comune è la diffusione di vapore, che permette di utilizzare piccole quantità di campione; sono sufficienti, infatti, pochi μL della soluzione della macromolecola. Si basa sull'equilibrio in fase vapore, che si stabilisce tra due soluzioni della stessa sostanza a diversa concentrazione. Si prepara una goccia contenente la molecola da cristallizzare in una soluzione tampone, insieme all'agente precipitante e altri additivi. Viene poi lasciata equilibrare con un serbatoio, reservoir, che contiene una soluzione dell'agente precipitante a una concentrazione superiore a quella della goccia. La specie più volatile, l'acqua o un solvente organico, diffonde fino a quando la sua pressione di vapore nella goccia uguaglia quella nel serbatoio. Se l'equilibrio viene raggiunto per diffusione di molecole d'acqua, queste si sposteranno dalla goccia al serbatoio. La goccia diminuisce in

volume, facendo così aumentare la concentrazione dei vari componenti. In questo modo si può raggiungere la condizione di sovrasaturazione.

La diffusione di vapore può essere effettuata con diversi assetti sperimentali (Figura 5-3).

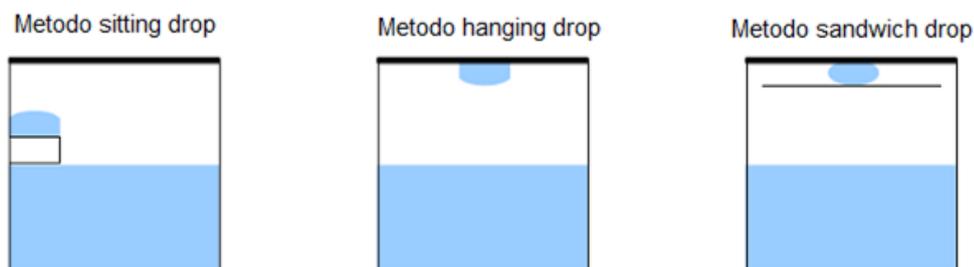


Figura 5-3 Schema dei setting sperimentali utilizzati nei metodi sitting drop, hanging drop e sandwich drop

Il metodo più semplice è il cosiddetto sitting drop; si deposita la gocciola su un supporto all'interno del serbatoio. Il metodo sandwich drop prevede che la gocciola venga sistemata in un supporto progettato in maniera tale che, una volta chiuso il pozzetto con un vetrino, questa si trovi a contatto sia con il supporto sia con il vetrino. Nel metodo hanging drop la gocciola viene posta su un vetrino opportunamente trattato per avere una superficie idrofobica; il vetrino è poi sospeso sul serbatoio.

5.1.4 Tecniche criogeniche

La raccolta di dati cristallografici da molecole biologiche viene eseguita generalmente a basse temperature. Questi cristalli hanno un basso potere di diffrazione e lavorare a bassa temperatura fa diminuire l'agitazione termica degli atomi e quindi aumenta l'intensità e la risoluzione. Contengono, inoltre, un'elevata percentuale di solvente, che può formare dei canali attraverso i quali radicali e ioni prodotti dalla radiazione X possono diffondere nel cristallo e danneggiarlo. A bassa temperatura, la diffusione di radicali è ridotta.

La bassa temperatura (100 K) è ottenuta prelevando azoto liquido da un dewar e flussandolo direttamente sul cristallo.

Per poter lavorare a bassa temperatura, quando il cristallo viene montato sullo strumento, deve essere immerso in una soluzione criogenica, che impedisce la disidratazione e il conseguente danneggiamento del cristallo ed evita la formazione di ghiaccio, solidificando in una fase vetrosa trasparente ai raggi X.

La soluzione criogenica contiene generalmente tutti i componenti presenti nella goccia, ad eccezione della macromolecola, e una certa percentuale in volume di un agente crioprotettente, come MPD, glicerolo o PEG.

5.2 Materiali e metodi

Le prove di cristallizzazione sono state effettuate presso il laboratorio di Bioinorganica/Biocristallografia del Dipartimento di Chimica.

5.2.1 Materiali

I filamenti di DNA utilizzati, Tel12 e Tel23, sono stati acquistati dall'azienda Jena Bioscience, sotto forma di campioni liofilizzati e purificati per HPLC.

5.2.2 Metodi

Come primo passo, sono state preparate delle soluzioni di DNA con una concentrazione pari a 3 mM, sciogliendo i vari filamenti in una soluzione di KCl 200 mM e tamponata a pH 6,5 con cacodilato di sodio 50 mM. La soluzione è stata poi sottoposta al processo di annealing, ossia è stata riscaldata a bagnomaria per 15 minuti a circa 90°C ed è stata lasciata gradualmente tornare a temperatura ambiente. Questa procedura consente al DNA di assumere la corretta struttura.

Sono state preparate soluzioni dei leganti: per L1 è stata preparata una soluzione 2,5 mM in una miscela acqua: DMSO, in rapporto 1:4; L2 e L3 sono invece stati sciolti in solo DMSO per giungere a una concentrazione di 20 mM.

Sono stati poi preparati gli addotti in rapporto 1:1 in soluzioni di concentrazione 1 o 2 mM, in particolare: Tel23-L1 1mM, Tel23-L2 2mM, Tel12-L2 1mM, Tel12-L3 1 mM.

Per la deposizione sono state utilizzate le piastre Grenier 96 Well Protein Crystallization Plates, anch'esse prodotte dalla ditta Jena Bioscience, di dimensioni 127,76 mm x 85,84 mm. Le piastre contengono 96 pozzetti, divisi in un serbatoio di capacità 200 µL e tre compartimenti allineati di dimensione minore.

Il protocollo di cristallizzazione, che utilizza il metodo sitting drop, prevede di depositare nel compartimento centrale tra i tre più piccoli una goccia contenente 1 µL di complesso DNA-legante e 1 µL di soluzione precipitante, detta anche condizione, mentre il serbatoio doveva essere riempito con un coprecipitante. Sono stati utilizzati quattro coprecipitanti e per ogni condizione sono state effettuate tre deposizioni con concentrazioni crescente di coprecipitante nel serbatoio. Lo schema delle condizioni utilizzate e delle soluzioni deposte nei serbatoi è riportato nella seguente tabella (Tabella 5-1).

Tabella 5-1 Protocollo di cristallizzazione

	Goccia			Reservoir
	Tampone	Sali	Coprecipitante	Coprecipitante
1	50 mM Na Caco, pH 6,5	150 mM NaCl, 50 mM Li ₂ SO ₄	10% MPD	MPD 20%, 35%, 50%
2	50 mM Na Caco, pH 6,5	300 mM KCl	15% MPD	
3	50 mM Na Caco, pH 6,5	150 mM NaCl	15% MPD	
4	50 mM Na Caco, pH 6,5	150 mM NaCl, 50 mM MgCl ₂	20% MPD	

5	50 mM Na Caco, pH 6,5	50 mM KCl, 50 mM LiCl	20% MPD	
6	50 mM Na Caco, pH 6,5	300 mM Li ₂ SO ₄	25% MPD	
7	50 mM Na Caco, pH 6,5	200 mM NaBr, 50 mM MgCl ₂	25% MPD	
8	50 mM Na Caco, pH 6,5	150 mM NaCl, 100 mM Li ₂ SO ₄	30% MPD	
9	50 mM Na Caco, pH 6,5	300 mM NaCl	30% MPD	
10	50 mM Na Caco, pH 6,5	200 mM KCl, 50 mM KI	25% PEG400	PEG 400 15%, 30%, 45%
11	50 mM Na Caco, pH 6,5	100 mM KCl, 50 mM Li ₂ SO ₄	10% PEG400	
12	50 mM Na Caco, pH 6,5	100 mM Li ₂ SO ₄	20% PEG400	
13	50 mM Na Caco, pH 6,5	300 mM KI	15% PEG400	
14	50 mM Na Caco, pH 6,5	300 mM NaI	15% PEG400	
15	50 mM Na Caco, pH 6,5	100 mM KCl, 30 mM CaCl ₂	1.3 M AS	AS 1.2M, 1.8 M, 2.4 M
16	50 mM Na Caco, pH 6,5	200 mM KCl, 50 mM Li ₂ SO ₄	1.8 M AS	
17	50 mM Na Caco, pH 6,5	100 mM NaCl, 50 mM MgCl ₂	1.0 M AS	
18	50 mM Na Caco, pH 6,5	50 mM Li ₂ SO ₄	1.0 M AS	
19	50 mM Na Caco, pH 6,5	150 mM NaCl, 100 mM Li ₂ SO ₄	10% PEG3000	PEG 3000 20%, 35, %, 50%

20	50 mM Na Caco, pH 6,5	200 mM NaBr, 50 mM LiCl	15% PEG3000
21	50 mM Na Caco, pH 6,5	300 mM KI	20% PEG3000
22	50 mM Na Caco, pH 6,5	150 mM NaCl, 50 mM MgCl ₂	20% PEG3000
23	50 mM Na Caco, pH 6,5	150 KCl	25% PEG3000
24	50 mM Na Caco, pH 6,5	300 mM NaCl	30% PEG3000

Dopo aver completato la deposizione, le piastre sono state chiuse con un apposito nastro adesivo trasparente, prodotto dalla Hampton Research, e poste in incubatore alla temperatura di 23°C. Sono poi state controllate periodicamente per monitorare l'eventuale crescita di cristalli.

6 Discussione

6.1 Prove di cristallizzazione

Le prove di cristallizzazione sono state condotte seguendo un protocollo di cristallizzazione basato su dati ottenuti dalla letteratura e da precedenti prove di cristallizzazione del DNA telomeric.

Nella cristallizzazione di campioni di DNA con il metodo sitting drop la procedura più comune consiste nel porre nella goccia il complesso DNA:legante insieme a un tampone, uno o più sali inorganici e il precipitante. Nel reservoir viene introdotto lo stesso precipitante presente nella goccia, ma in concentrazione maggiore.

Lo screening usato in questo lavoro di tesi prevedeva di utilizzare per ogni condizione tre diverse concentrazioni di precipitante nel reservoir, con percentuali crescenti. Dalla concentrazione del precipitante nel reservoir, infatti, dipende la velocità di evaporazione del solvente dalla goccia ed è quindi uno dei parametri che si variano comunemente nei protocolli di cristallizzazione per testare le condizioni di cristallizzazione.

Durante le prove di cristallizzazione sono stati ottenuti alcuni risultati classificati in seguito come falsi positivi, in particolare la condizione 15, contenente KCl, CaCl_2 e solfato di ammonio come precipitante, dava sistematicamente origine a cristalli di natura inorganica, probabilmente CaSO_4 , ed è per questo stata eliminata nelle prove successive.

I risultati in generale non sono stati sufficienti. Ciò può essere dovuto sia alla tendenza del DNA ad avere tempi di cristallizzazione molto lunghi, tendenza già riscontrata in precedenti esperimenti, sia alla natura polifunzionale dei leganti, che contengono nel loro scheletro sia gruppi aromatici sia gruppi amminici e ammoniaci. Nel legante, infatti, sono presenti gruppi funzionali con caratteristiche molto diverse, che possono dar luogo a una interazione aspecifica con il target: mentre gli anelli aromatici sono adatti a interagire con i quartetti di

guanine, i gruppi amminici e ammoniacali tendono a dare interazioni carica-carica e a ponte d'idrogeno con i gruppi fosfato dei solchi.

In generale le formazioni presenti hanno scarso carattere cristallino e si presentano come addotti non regolari.

L'unico legante che ha dato formazioni a maggiore carattere di cristallinità è L3 con la sequenza Tel23 e la condizione 21, che contiene KI e PEG 3000 come precipitante (Figura 6-1).

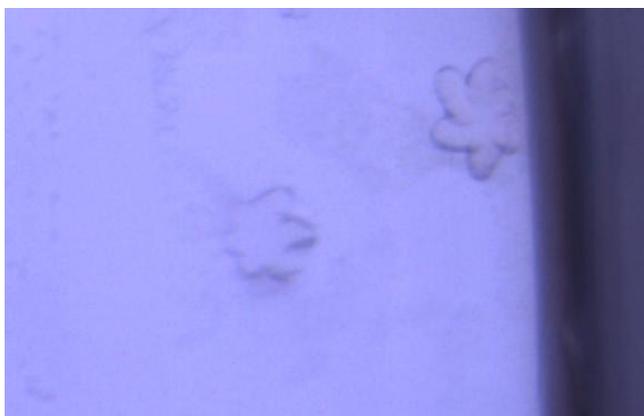


Figura 6-1 aggregati cristallini ottenuti per il sistema L3+Tel23; condizione 21: 50 mM NaCaco, 300 mM KI, 20% PEG3000

6.2 Modellistica molecolare

La prima parte del lavoro di modellistica molecolare è stata dedicata alla ricerca delle conformazioni iniziali del legante e del DNA. Per quanto riguarda il legante, è necessario considerare le caratteristiche strutturali del gruppo terpiridile. Esso, infatti, può assumere tre diverse conformazioni, dette trans/trans, cis/trans e cis/cis, in funzione del valore assunto dai due angoli torsionali N-C-C-N (Figura 6-2).

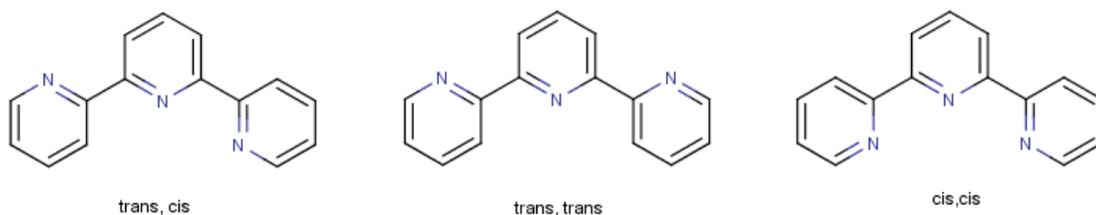


Figura 6-2 Conformazioni della terpiridina

Calcoli quantomeccanici, strutture cristalline e strutture ricavate da studi di risonanza magnetica dimostrano che la conformazione più frequente nel terpiridile non protonato e non coordinato a centri metallici è quella trans/trans, sebbene la presenza di solventi polari possa favorire la formazione di addotti in cui la terpiridina è nella conformazione cis/cis³⁰.

Durante i precedenti studi effettuati su questi leganti dal gruppo del Prof. Bianchi, è stata risolta la struttura cristallina della forma tetraprotonata di L1 in addotto con l'anione HTTP, nella quale il legante assume una conformazione trans/trans (Figura6-3)³¹.

³⁰ C. Bazzicalupi, A. Bencini, A. Bianchi, E. Faggi, C. Giorgi, S. Santarelli, B. Valtancoli, *Coord. Chem. Rev.* 2007

³¹ C. Bazzicalupi, A. Bencini, A. Bianchi, E. Faggi, C. Giorgi, S. Santarelli, B. Valtancoli, *J. AM. CHEM. SOC.* 2008,130, 2440-2441

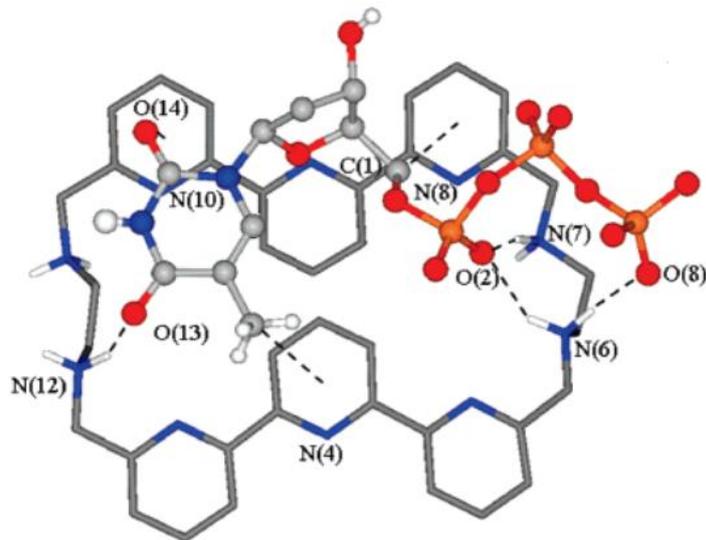


Figura 6-3 Struttura cristallina di $(H_4L1)HTTP$

Bisogna sottolineare che i due gruppi terpiridile assumono tale conformazione nonostante siano inseriti in uno scheletro ciclico e che tale scheletro è in L1 più piccolo rispetto a L2 e L3. Era quindi presumibile che la conformazione trans-trans potesse essere assunta anche dagli altri due leganti, anche se per essi non erano disponibili informazioni strutturali preliminari.

Per questo motivo sono state scelte tre diverse conformazioni iniziali per ogni legante, corrispondenti alle tre diverse configurazioni possibili per la terpiridina.

In ciascun legante i due gruppi terpiridile sono separati da due catene alifatiche poliazotate di lunghezze differenti e contenenti un diverso numero di atomi donatori, separati da gruppi etilenici, nel caso di L1 e L2, o propilenici, per L3 (Figure 2-9,10,11).

Queste differenze determinano le proprietà acido-base dei leganti, che si comportano tutti da basi polifunzionali in soluzione acquosa, come individuato con gli studi di caratterizzazione effettuati durante la già citata tesi di dottorato.

Di seguito sono riportati alcuni risultati di tali studi rilevanti per la presente tesi: le pKa degli equilibri di protonazione (Tabelle 6-1,2,3) e i diagrammi di distribuzione delle specie

protonate in soluzione acquosa (Figure 6-4,5). Anche la distribuzione dei protoni nei vari step di protonazione è stata ricavata dai già citati studi precedenti.

Equilibri	pKa
$[L1H_2]^{2+} + H^+ = [L1H_3]^{3+}$	7.69
$[L1H_3]^{3+} + H^+ = [L1H_4]^{4+}$	5.44
$[L1H_4]^{4+} + H^+ = [L1H_5]^{5+}$	4.89
$[L1H_5]^{5+} + H^+ = [L1H_6]^{6+}$	2.91

Tabella 6-1 pKa degli equilibri di protonazione di L1 – in rosso l'equilibrio di relativo alla specie più abbondante a pH fisiologico.

Equilibri	pKa
$L2H^+ + H^+ = [L2H_2]^{2+}$	8.34
$[L2H_2]^{2+} + H^+ = [L2H_3]^{3+}$	7.68
$[L2H_3]^{3+} + H^+ = [L2H_4]^{4+}$	6.82
$[L2H_4]^{4+} + H^+ = [L2H_5]^{5+}$	3.07
$[L2H_5]^{5+} + H^+ = [L2H_6]^{6+}$	2.53
$[L2H_6]^{6+} + H^+ = [L2H_7]^{7+}$	2.82

Tabella 6-2 pKa degli equilibri di protonazione di L2 – in rosso l'equilibrio di relativo alla specie più abbondante a pH fisiologico.

Equilibri	pKa
$[L3H_2]^{2+} + H^+ = [L3H_3]^{3+}$	7.94
$[L3H_3]^{3+} + H^+ = [L3H_4]^{4+}$	7.32
$[L3H_4]^{4+} + H^+ = [L3H_5]^{5+}$	6.57
$[L3H_5]^{5+} + H^+ = [L3H_6]^{6+}$	6.18
$[L3H_6]^{6+} + H^+ = [L3H_7]^{7+}$	4.84
$[L3H_7]^{7+} + H^+ = [L3H_8]^{8+}$	3.08

Tabella 6-3 pKa degli equilibri di protonazione di L3 – in rosso l'equilibrio di relativo alla specie più abbondante a pH fisiologico.

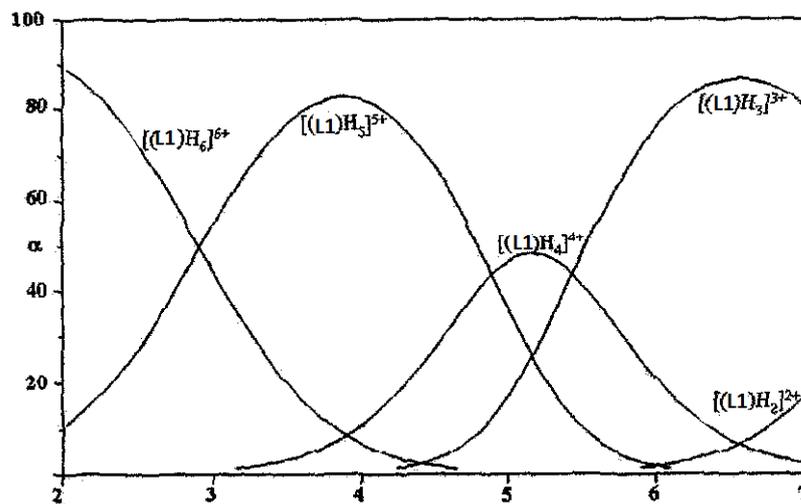


Figura 6-4 Diagramma di distribuzione delle specie per la protonazione di L1

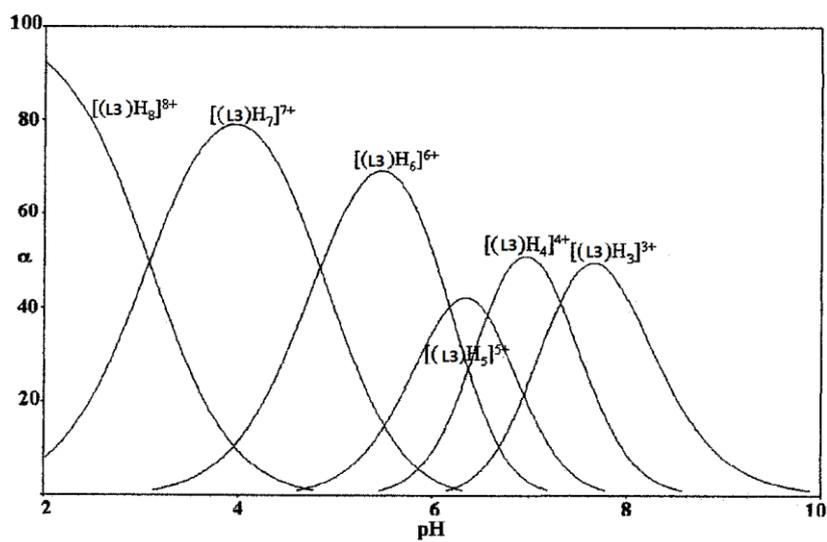


Figura 6-5 Diagramma di distribuzione delle specie per la protonazione di L3

Sulla base di questi dati per i leganti L1, L2 sono state utilizzate nei calcoli le specie triprotonate, mentre per L3 la specie tetraprotonata.

Per quanto riguarda la struttura quadruplex del DNA, anche in questo caso sono possibili varie conformazioni. Poiché, come già sottolineato, l'effettiva struttura assunta in vivo dal quadruplex telomerico umano non è stata ancora definitivamente accertata, in questo lavoro di tesi sono state valutate le conformazioni prevalentemente individuate in presenza di potassio: le conformazioni ibride e la conformazione propeller. Grazie alla spettroscopia NMR, infatti, è stato dimostrato che le sequenze AAAGGG[TTAGGG]₃AA (Tel26) e TTAGGG[TTAGGG]₃TT (wtTel26) assumono in soluzioni contenenti ioni potassio le conformazioni ibride, rispettivamente la hybrid 1 e la hybrid 2³²; sempre in presenza di ioni potassio nello stato solido è, invece, stata riscontrata la struttura propeller (Figura 6-6).

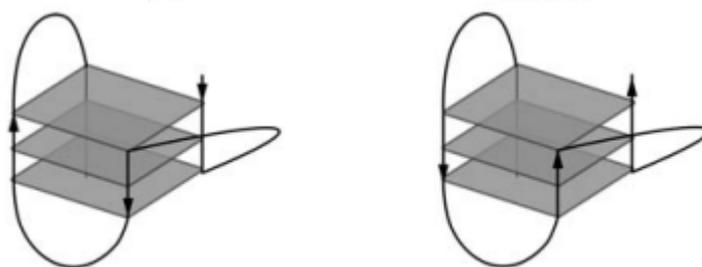


Figura 6-6 Strutture hybrid 1, a sinistra, e hybrid 2, a destra

Uno studio preliminare delle interazioni in soluzione di L1, L2 e L3 con il DNA telomerico umano è stato condotto dalla professoressa Claudia Sissi presso il Dipartimento di Scienze del Farmaco dell'Università degli Studi di Padova. Si è valutata capacità di legare e stabilizzare strutture G-quadruplex, attraverso esperimenti di Fluorescence Melting. Questa tecnica prevede di registrare lo spettro di fluorescenza di oligonucleotidi marcati con un fluoruro e un quencher. Se una soluzione contenente un acido nucleico strutturato viene

³² I. Bessi, C. Bazzicalupi, C. Richter, H. R. A. Jonker, K. Saxena, C. Sissi, M. Chioccioli, S. Bianco, A. Rita Bilia, H. Schwalbe, P. Gratteri *ACS Chem. Biol.*, 2012, 7 (6), pp 1109–1119

sottoposta a riscaldamento, può raggiungere la temperatura in cui i ponti a idrogeno si rompono; si può poi proseguire il riscaldamento fino a che l'acido nucleico si troverà in forma denaturata. La Temperatura di Melting (T_m) è la temperatura alla quale il 50% delle catene dell'oligonucleotide si trova in forma strutturata e il restante 50% in forma denaturata e rappresenta un importante indice della stabilità della struttura assunta dall'oligonucleotide in determinate condizioni di pH e forza ionica.

L'analisi, per ogni legante, è stata condotta confrontando la T_m dell'oligonucleotide in presenza di concentrazioni crescenti di legante con quella dell'oligonucleotide da solo (T_{m_0}), definendo così un ΔT_m , dove $\Delta T_m = T_m - T_{m_0}$.

Dai risultati, emerge che L3 ha una maggiore affinità per il DNA telomerico umano rispetto a L1 e L2, che hanno efficacia comparabile (Figura 6-7).

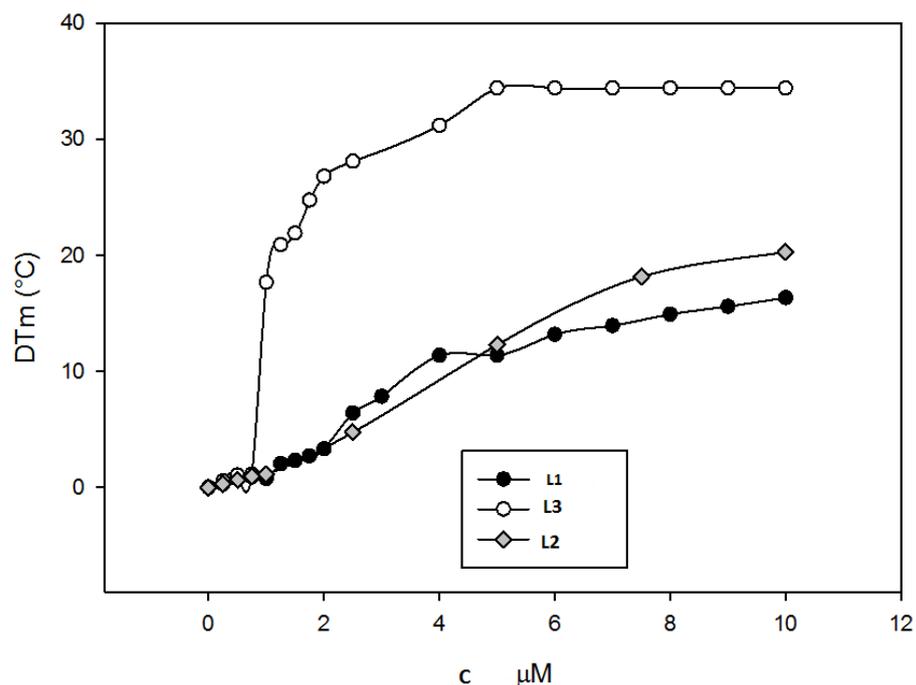


Figura 6-7 Variazione di ΔT_m all'aumentare della concentrazione di L1, L2, L3

Utilizzando per questo tipo di misure anche DNA a doppio filamento, si è potuto notare che i tre leganti sono selettivi per il DNA G-quadruplex e non per quello a doppia elica. Questo è un dato molto interessante, soprattutto per un eventuale uso di questi leganti in vivo, dove il DNA duplex è decisamente più abbondante del DNA G-quadruplex all'interno delle cellule ed è quindi necessaria un'elevata selettività a favore del quadruplex, affinché i leganti possano effettivamente interagire con quest'ultimo (Figure 6-8,9,10.)

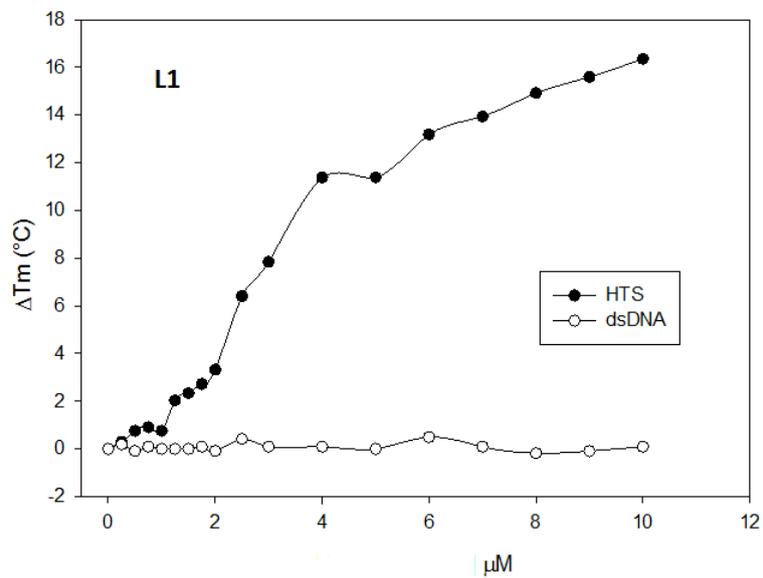


Figura 6-8 Variazione di ΔT_m all'aumentare della concentrazione di L1 con DNA telomerico in avvolgimento G-quadruplex e DNA a doppio filamento

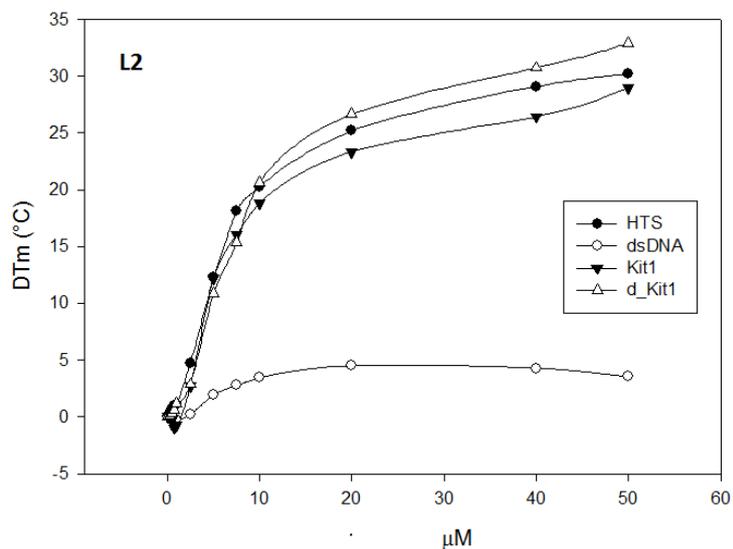


Figura 6-9 Variazione di ΔT_m all'aumentare della concentrazione di L2 con DNA a doppio filamento e DNA telomerico in avvolgimento G-quadruplex, . In figura sono riportati anche i dati relativi all'interazione del legante con G-quadruplex formati da altre sequenze ricche in guanina

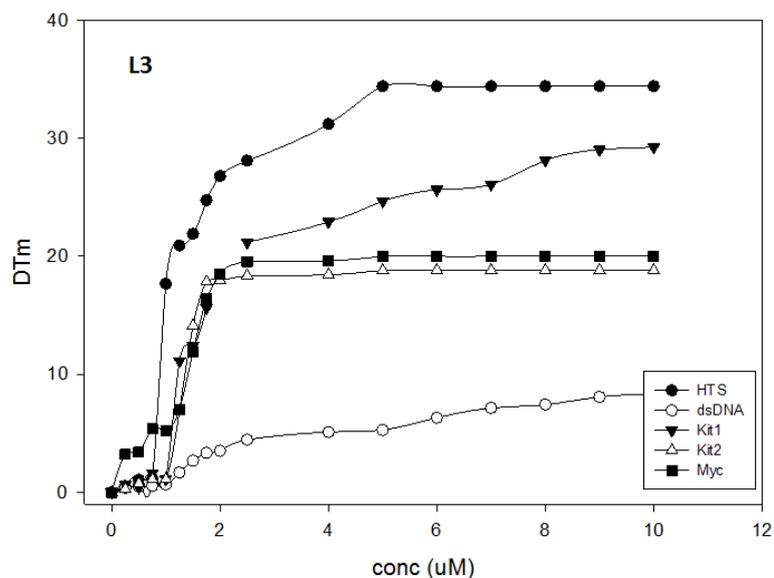


Figura 6-10 Variazione di ΔT_m all'aumentare della concentrazione di L3 con DNA a

doppio filamento e DNA telomerico in avvolgimento G-quadruplex, DNA a doppio filamento. In figura sono riportati anche i dati relativi all'interazione del legante con G-quadruplex formati da altre sequenze ricche in guanina

Come già accennato, l'analisi conformazionale è stata condotta su tre diverse conformazioni iniziali per ogni legante, corrispondenti alle tre diverse conformazioni possibili per la terpiridina. Su queste conformazioni è stata inizialmente eseguita una minimizzazione, un'analisi conformazionale e infine un'ottimizzazione utilizzando il metodo DFT, per identificare la conformazione a minore energia per ogni legante. Per ogni legante è stata selezionata la struttura di minima energia individuata con la procedura di analisi conformazionale. Le tre strutture così scelte sono state utilizzate per costruire le conformazioni iniziali degli addotti DNA:legante. In mancanza di informazioni dettagliate, DNA e legante sono stati considerati in rapporto stechiometrico 1:1.

Su ogni recettore sono state individuate sei zone, che corrispondono rispettivamente ai quartetti superiore e inferiore della struttura e approssimativamente ai quattro lati del G-tetrad. I leganti sono stati posizionati manualmente nelle varie zone e le strutture ottenute sono state minimizzate. Il sistema è stato quindi studiato tramite dinamica molecolare. Per il trattamento degli effetti del solvente si è optato per il metodo con costante dielettrica dipendente dalla distanza. Quest'approssimazione, per quanto poco raffinata, si è resa necessaria perché le prove svolte con approssimazioni migliori hanno evidenziato una mobilità molto scarsa dei loop e, di conseguenza, l'investigazione della superficie di energia potenziale avrebbe richiesto tempi molto lunghi, non compatibili con le tempistiche della tesi. Anche la scelta della dinamica rispetto a una procedura di docking, spesso utilizzata in questo tipo di studi, è legata alla presenza dei loop, che nelle conformazioni ibride bloccano l'accesso al quartetto. Si è perciò deciso di ottenere le coordinate iniziali tramite posizionamento manuale e successiva minimizzazione degli addotti ottenuti. La lunghezza delle dinamiche è stata scelta facendo prove con tempi diversi.

Per ciascun sistema DNA:legante studiato sono state selezionate le conformazioni più stabili, caratterizzate da un valore di ΔE minore in un intervallo di circa 40 kJ/mol dal minimo.

I valori di ΔE (kJ/mol) riportati nella tabella successiva sono relativi al processo di formazione dell'addotto secondo la reazione $L + \text{DNA} = L:\text{DNA}$. In rosso sono indicate le energie corrispondenti agli addotti più stabili per ogni legante (Tabella 6-4).

Tabella 6-4 Valori di ΔE (kJ/mol) relativi al processo di formazione dell'addotto

Legante	Regione	Hybrid 1 (H1)	Hybrid 2 (H2)	Propeller (S1)
L1	R1		-482.852	
	R2			
	R3			
	R4			-488.735
	R5	-478.089		-506.498
	R6	-468.791		
L2	R1			
	R2	-531.388		
	R3			
	R4	-527.604	-491.933	
	R5			-485.710
	R6			
L3	R1		-590.851	-593.039
	R2			
	R3			
	R4		-567.790	
	R5	-567.792		
	R6			

L'analisi dei risultati ottenuti evidenzia che tutti e tre i leganti possono interagire con ciascuna delle strutture quadruplex considerate per il DNA telomeric. Si nota, in particolare, il seguente trend di interazione: $L1 < L2 < L3$. Tale andamento può però essere in parte dovuto al diverso grado di protonazione dei leganti. A pH fisiologico, infatti, L3 è stabile in soluzione come specie tetraprotonata, mentre L1 ed L2 possono legare solo tre protoni ciascuno (Tabelle 6-1,2,3). La carica positiva maggiore determina, ovviamente, una più favorevole interazione elettrostatica con il recettore polianionico.

Questo risultato è in buon accordo con le stabilità relative degli addotti in soluzione. Le misure in soluzione evidenziano come a pH fisiologico L3 determini una variazione della temperatura di melting decisamente maggiore rispetto a L1 e L2 (Figura 6-7), che invece inducono un'analoga stabilizzazione della struttura quadruplex.

D'altronde, il contributo elettrostatico non sembra essere l'unico a guidare l'interazione con il DNA quadruplex. In presenza del solo contributo elettrostatico, infatti, sarebbe lecito aspettarsi come favorite le regioni d'interazione da R3 a R6, cioè quelle caratterizzate dalla presenza dei gruppi fosfato. Questo non si verifica nel caso di L3, per il quale sono maggiormente stabili gli addotti coinvolgenti le tetradi. Il legante L1 dimostra, invece, un'interazione preferenziale per i solchi piuttosto che per il quartetto. Tra i solchi, inoltre, sono interessate più zone e con ΔE simili in folding diversi; si nota, per esempio, una variazione di ΔE di 10 kJ/mol tra la zona 4 della conformazione propeller e la zona 5 della struttura hybrid1. La conformazione migliore risulta essere quella in cui il legante occupa la regione 5 del propeller. In generale sembrano dominare le interazioni dei legami a idrogeno e le interazioni carica-carica rispetto al π -stacking. Per L2 si riscontrano delle interazioni del tutto analoghe a quelle di L1.

Per L3 si ha, invece, un comportamento diverso. Le strutture più stabili sono gli addotti in cui il legante interagisce con un quartetto della struttura propeller o della struttura hybrid2. Le interazioni con i solchi non sono assenti, ma sembrano prevalere quelle con il quartetto.

6.2.1 Analisi degli addotti ³³

Come già spiegato, le modalità di interazione possibili tra un legante e il DNA G-quadruplex sono tre: il legante può impilarsi sul quartetto, interagire con i solchi o dare intercalazione fra due quartetti adiacenti (Capitolo 2, paragrafo 2.1, figura 2-1). L'intercalazione è sfavorita dal punto di vista energetico, quindi i primi due casi sono più comuni.

L'impilamento è dovuto a interazioni π - π tra il legante e il quartetto di guanine ed è per questa ragione che la maggior parte dei leganti per le strutture G-quadruplex del DNA telomerico contengono un'ampia porzione aromatica. Nei solchi, a causa della presenza dei gruppi fosfato, predominano, invece, le interazioni elettrostatiche e le interazioni a ponte d'idrogeno.

In generale, leganti caratterizzati dalla presenza di gruppi carichi positivamente risultano favoriti, in quanto queste possono dar luogo a interazioni favorevoli sia con i gruppi fosfato sia con la densità elettronica π delle guanine.

In molte delle conformazioni selezionate per gli addotti formati fra i leganti studiati e i quadruplex telomerici, sono spesso presenti entrambi i contributi, π - π stacking e interazioni coulombiane.

6.2.2 Addotti con L3

Il primo addotto analizzato è quello che presenta il ΔE minore tra tutti quelli ottenuti, sia per L3 sia per L1 e L2. La zona considerata è quella corrispondente al quartetto di guanine vicino all'estremità 5' (Figura 6-10,11).

³³ Le immagini degli addotti sono state realizzate con il programma VMD: Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics" *J. Molec. Graphics* **1996**, *14.1*, 33-38.

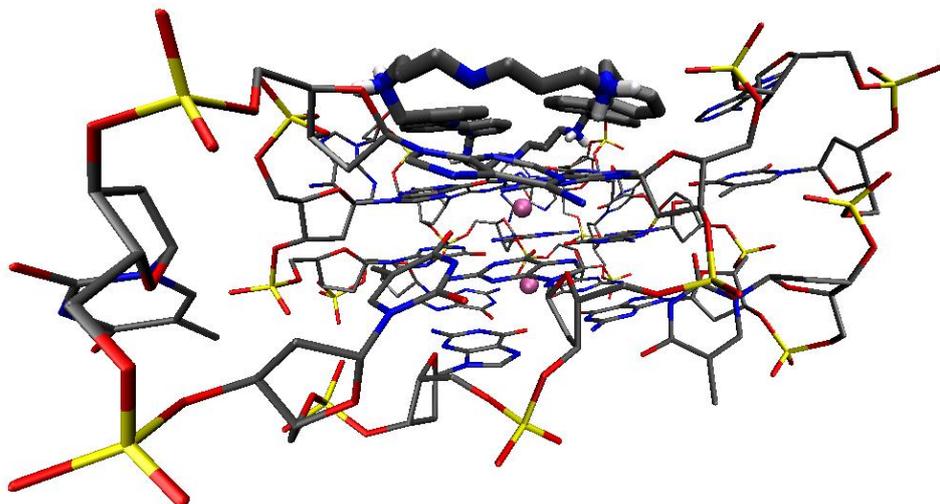


Figura 6-10 Addotto L3_S1_R1

L'interazione predominante è lo stacking π - π . I piani medi, definiti rispettivamente dai residui di guanina e i due gruppi terpiridile, si trovano a una distanza di 4,039 Å. L'interazione con il quartetto è favorita dalla particolare topologia del DNA telomerico, che in questo caso è la propeller. In questa struttura, al contrario di quanto avviene nelle strutture ibride, i loop non ingombrano i quartetti, disponendosi, invece in corrispondenza delle facce laterali del quadruplex. Sono perciò favorite le interazioni di stacking con il quartetto, che è, di fatto, più disponibile che nelle strutture di tipo ibrido.

E' interessante notare (Figura 6-10) che le catene protonate del legante si allontanano dal quartetto, puntando verso i solchi laterali. S'instaura in questo modo una serie di contatti a ponte d'idrogeno coinvolgenti i gruppi ammonio del legante e i gruppi fosfato del DNA, che danno un importante contributo alla stabilità risultante dell'addotto. Le interazioni, mostrate nella figura 6-11, sono generalmente abbastanza forti e dell'ordine dei 2 Å.

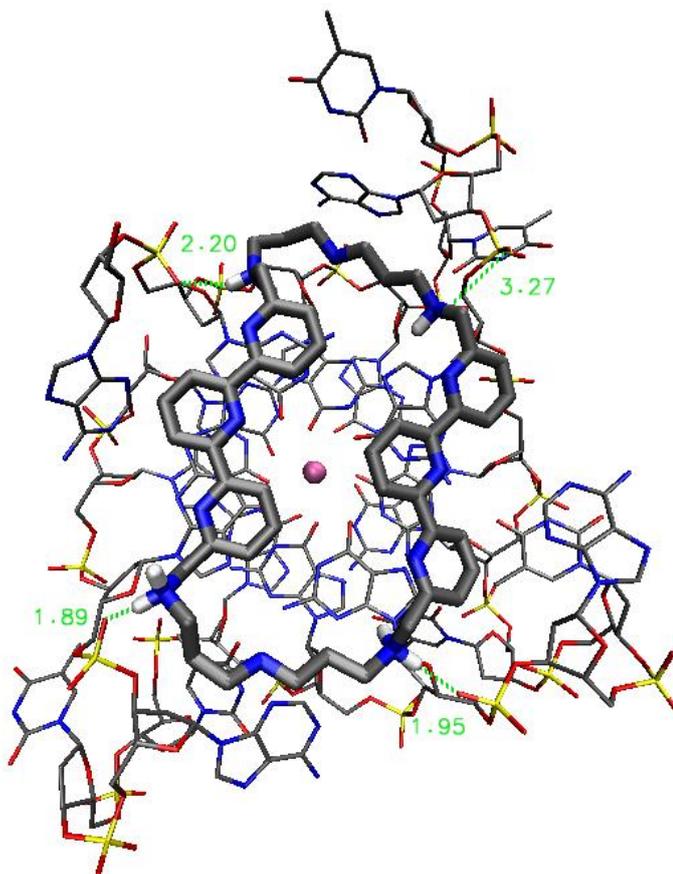


Figura 6-11 Interazioni a ponte a idrogeno tra L3 e il DNA telomerico nella struttura propeller (addotto L3_S1_R1)

La conformazione assunta da L3 in questo addotto presenta delle similitudini con quella assunta da L1 nella struttura cristallina (H₄L1)HTTP (Figura 6-2) nella quale il gruppo timidina è posizionato in prossimità degli anelli aromatici del legante, mentre i gruppi fosfato interagiscono con i gruppi ammonio del legante.

L'addotto che presenta il secondo migliore valore di ΔE è quello formato dal legante L3 e il quartetto all'estremità 5' della struttura hybrid2 del DNA telomerico (Figura 6-12).

La zona in cui si trova il legante è, come nel caso precedente, uno dei quartetti, ma, trattandosi di una topologia ibrida, questo è molto più ingombrato dai loop rispetto alla

conformazione propeller. Il contributo determinante nella formazione dell'addotto sembra essere fornito dai gruppi ammonio, che interagiscono sia con le guanine sia con lo scheletro fosfodiesterico. Poiché il legante L3 è tetraprotonato, si possono avere entrambe le interazioni, anche se quella con i solchi sembra predominante (Figura 6-13). La presenza del legante in prossimità del loop può dare luogo a ulteriori legami a idrogeno con le basi azotate che lo costituiscono. In figura sono evidenziate le misure di alcune distanze selezionate tra il legante e il recettore. Si possono notare, in particolare, distanze di circa 2 Å per le interazioni a ponte d'idrogeno tra gli atomi di azoto protonati e atomi di ossigeno dei gruppi fosfato e distanze più lunghe, di circa 3-4 Å, relative a interazioni di tipo NH... π .

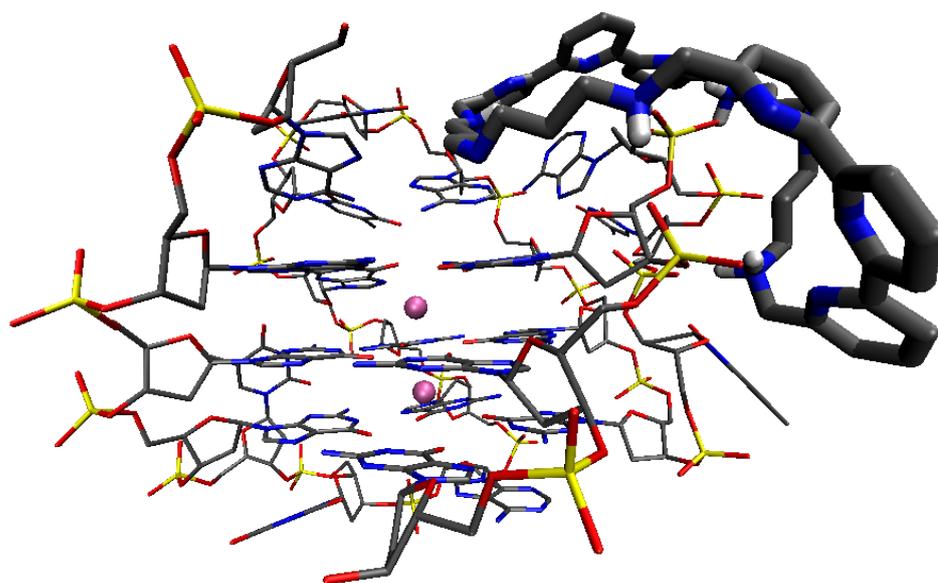


Figura 6-12 Addotto L3_H2_R1

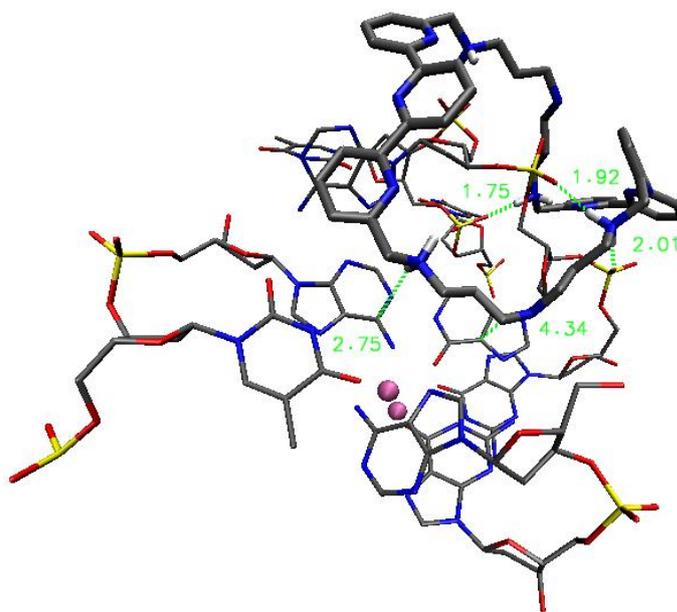


Figura 6-13 Particolare dell'addotto L3_H2_R1

Tra le strutture selezionate ci sono altri due addotti con L3. In entrambi i casi, il legante si trova vicino ai solchi (Figure 6-14,15). In queste regioni, che si trovano in corrispondenza dei lati del G-tetrad, le interazioni con i gruppi fosfato sono più importanti.

Le energie di formazione sono la terza e la quarta migliori in assoluto, in accordo con l'evidenza che L3 è il legante che interagisce meglio con la struttura G-quadruplex. I due addotti hanno energie del tutto comparabili, perché anche le interazioni presenti sono del tutto analoghe. Le zone coinvolte sono due diversi solchi, ma in entrambi i casi si ha interazione sia con i gruppi fosfato, sia, in misura minore con le guanine che formano il quartetto. Ciò è possibile grazie alle dimensioni relativamente grandi di L3.

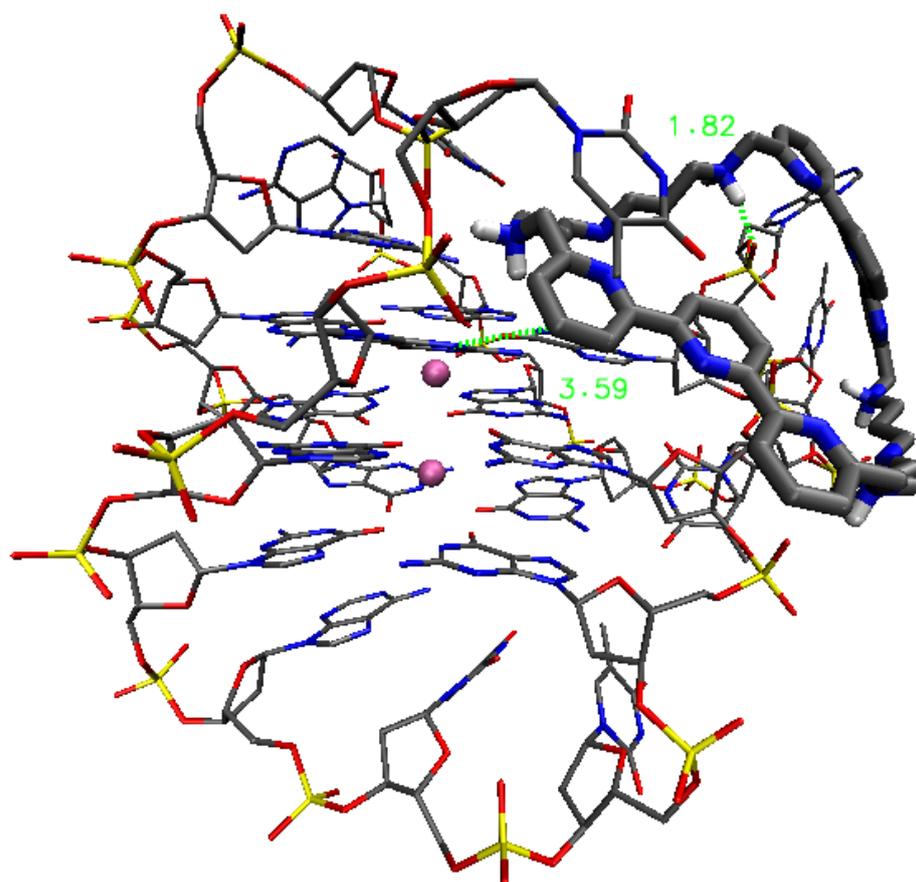


Figura 6-14 Interazioni tra il legante L3 e il DNA in avvolgimento hybrid1 (addotto L3_H1_R5)

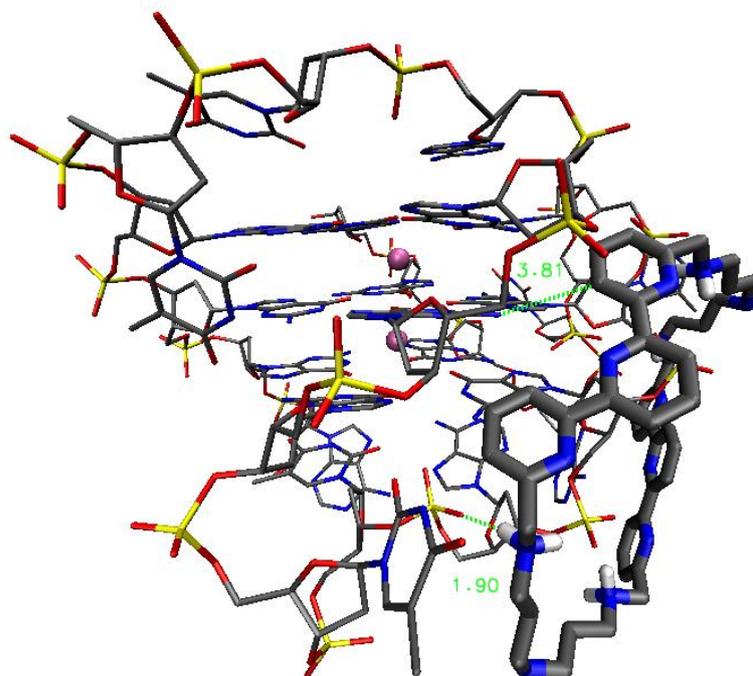


Figura 6-15 Interazioni tra il legante L3 e il DNA in avvolgimento hybrid2 (addotto L3_H2_R4)

6.2.3 Addotti con L2

I due addotti hanno le migliori energie per L2 hanno energie simili tra loro, come simili sono le interazioni a cui danno luogo con il quadruplex (Figura 6-16,17).

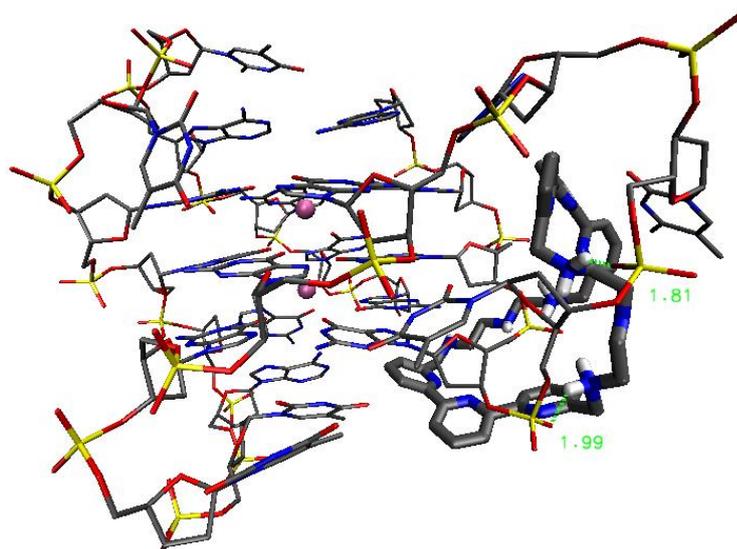


Figura 6-16 Interazioni a ponte a idrogeno tra il legante L2 e il DNA in avvolgimento hybrid1 (addotto L2_H1_R2)

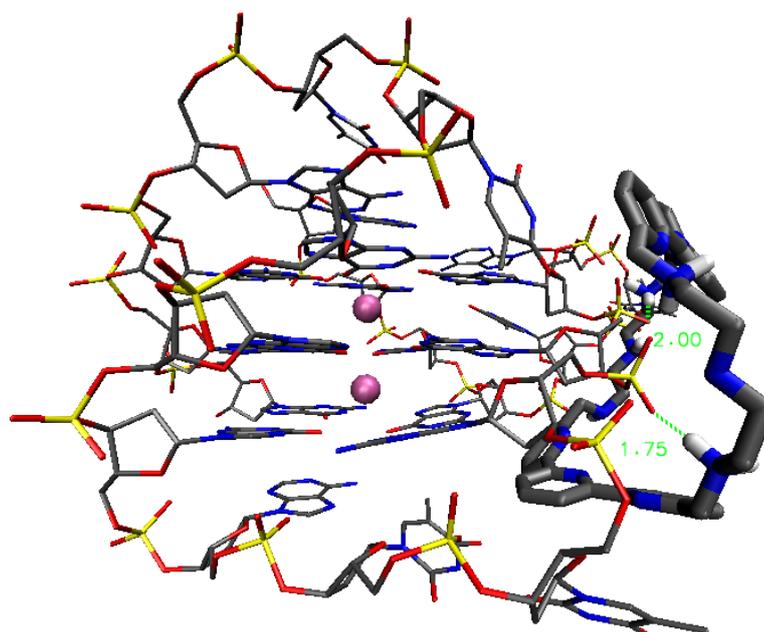


Figura 6-17 Interazioni a ponte a idrogeno tra il legante L2 e il DNA in avvolgimento hybrid1 (addotto L2_H1_R4)

In entrambi i casi prevalgono le interazioni con i solchi a scapito del contributo legato alla parte aromatica.

Anche nel caso del DNA in avvolgimento L2 e propeller (Figura 6-18,19), l'interazione avviene principalmente con i gruppi fosfodiesterici.

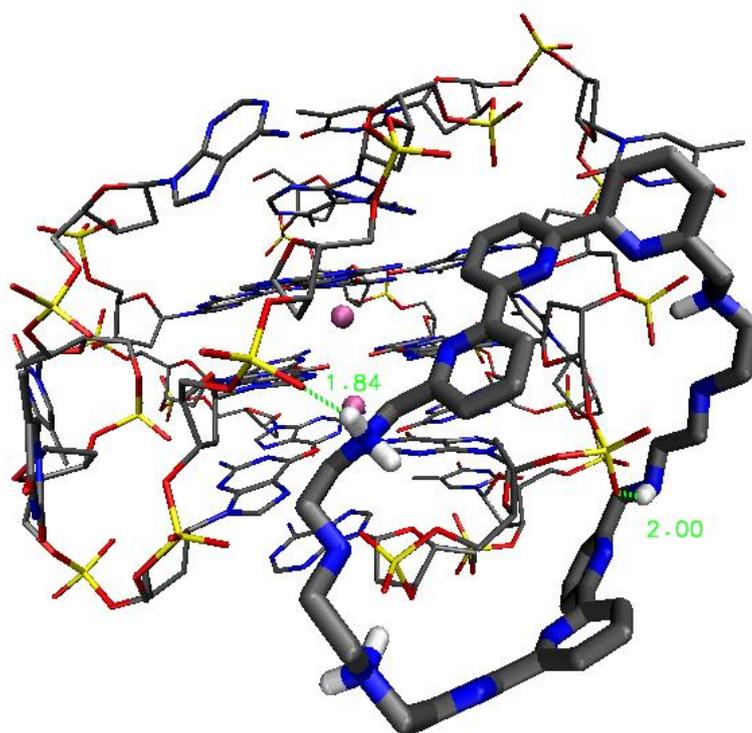


Figura 6-18 Interazioni a ponte a idrogeno tra il legante L2 e il DNA in avvolgimento hybrid2 (addotto L2_H2_R4)

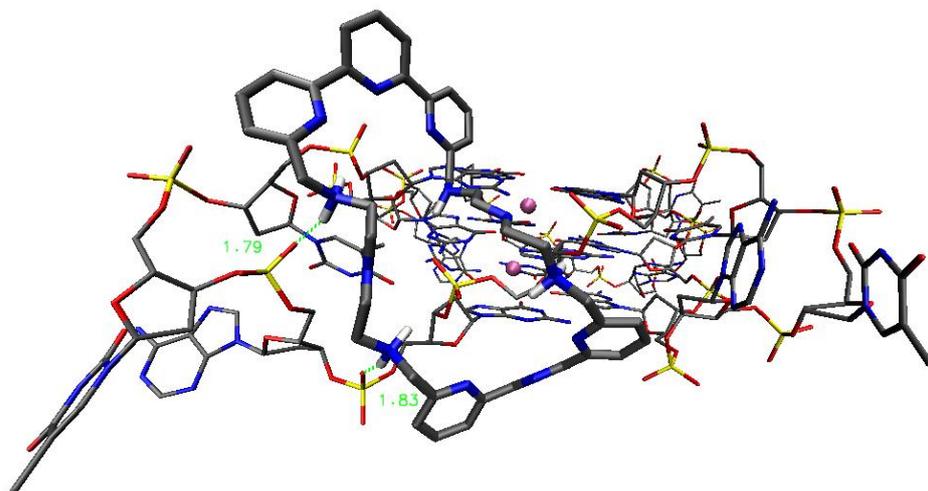


Figura 6-19 Interazioni a ponte a idrogeno tra il legante L2 e il DNA in avvolgimento propeller (addotto L2_S1_R5)

6.2.4 Addotti con L1

I due addotti che presentano le migliori energie per L1 sono mostrati in figura (Figura 6-20,21). Nonostante vi sia coinvolta la topologia propeller, le regioni interessate sono quelle relative ai loop laterali, mentre i quartetti non partecipano alla formazione dell'addotto

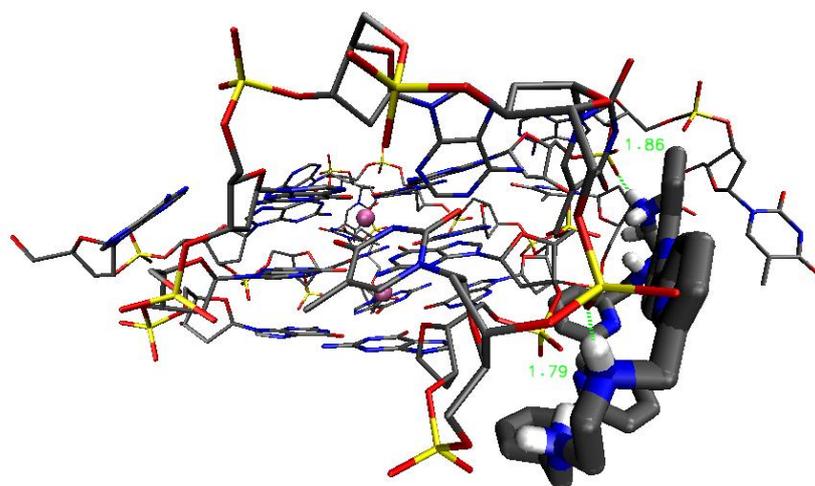


Figura 6-20 Interazioni tra il legante L1 e il DNA in avvolgimento propeller (addotto L2_S1_R4)

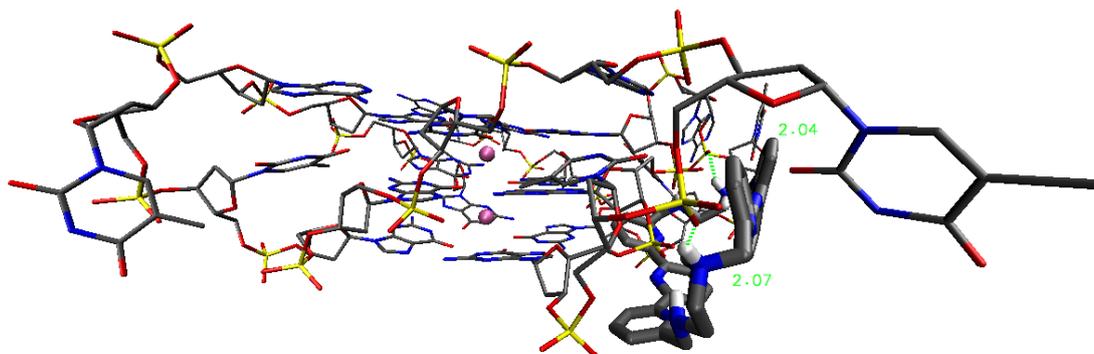


Figura 6-21 Interazioni tra il legante L2 e il DNA in avvolgimento propeller (addotto L2_S1_R5)

Per entrambe gli addotti, le interazioni predominanti sono quelle coulombiane fra il legante triprotonato e le cariche negative dello scheletro fosfodiesterico.

Bisogna però sottolineare che nel caso di L1 è stato ottenuto un addotto con il DNA in conformazione propeller, caratterizzato da un'energia di poco superiore al limite di soglia scelto per l'analisi dei risultati del modeling, ossia 40 kJ/mol. In questo addotto L1 dà stacking con il quartetto vicino all'estremità 5'. Anche in questo caso si possono notare delle similitudini con la struttura (H₄L1)HTTP e di conseguenza con la conformazione che L3 assume nell'analogo addotto. La differenza di energia così marcata fra i due addotti è probabilmente dovuta al fatto che L1, più piccolo e stressato, non riesce a portare ambedue i gruppi terpiridile in contatto con il quartetto. L3, invece, come si è detto, oltre a posizionare correttamente entrambi i terpiridili sul quartetto per dar luogo all'interazione di π stacking, riesce anche a stabilire contatti tra le catene contenenti gruppi ammonici e alcuni gruppi fosfato (Figura 6-22).

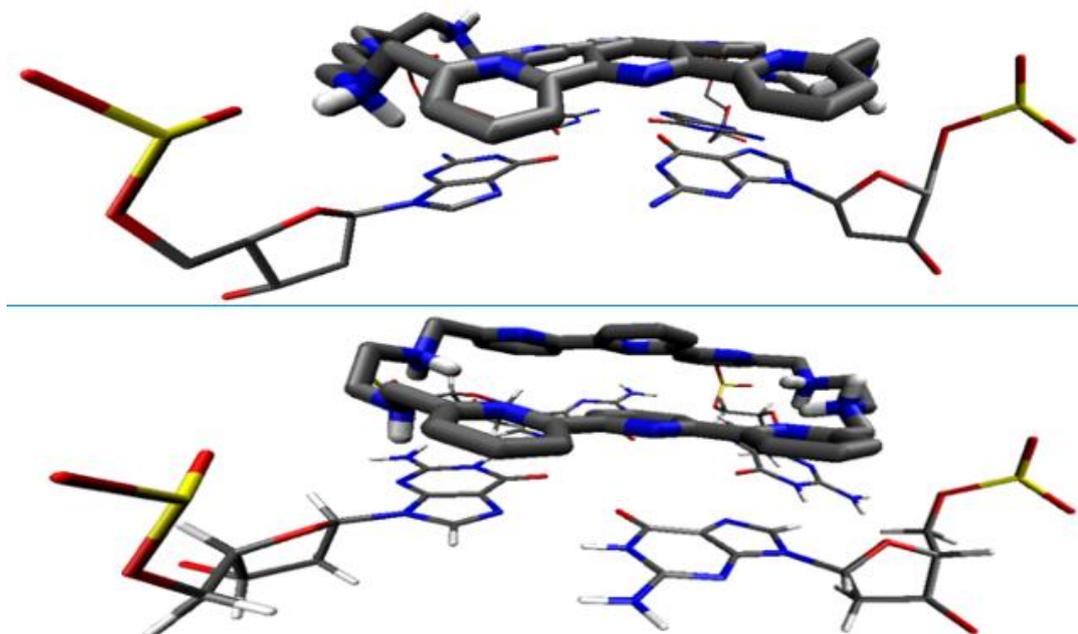


Figura 6-22 Confronto tra gli addotti di L3, sopra, e L1, sotto, con il DNA telomerico in conformazione propeller (particolare)

Tra gli addotti selezionati per L1, ci sono anche altri tre addotti, nei quali il legante non interagisce direttamente con il quartetto, ma si posiziona in prossimità dei solchi laterali, dando luogo a interazione con i gruppi fosfato (Figura 6-23, 25), o in prossimità del loop vicino all'estremità 5' (Figura 6-24).

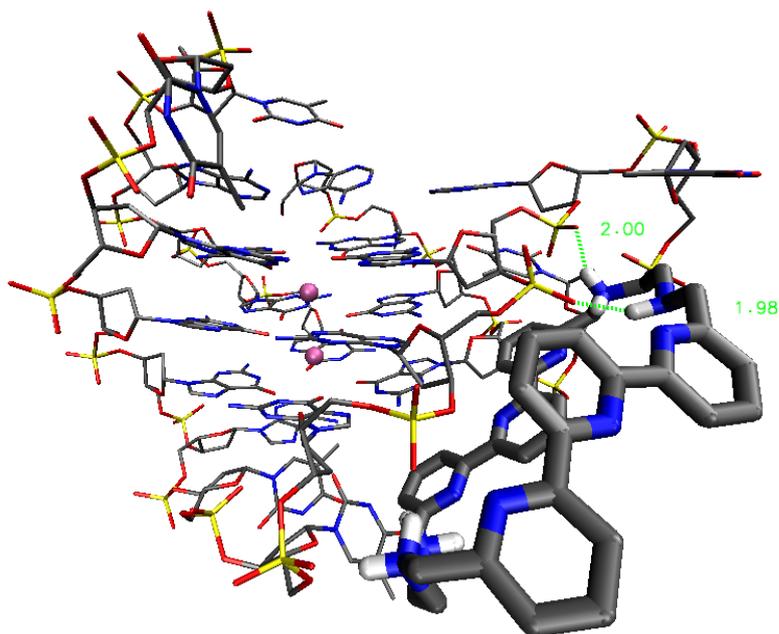


Figura 6-23 Interazioni a ponte a idrogeno tra il legante L e il DNA in avvolgimento hybrid1 (addotto L1_H1_R5)

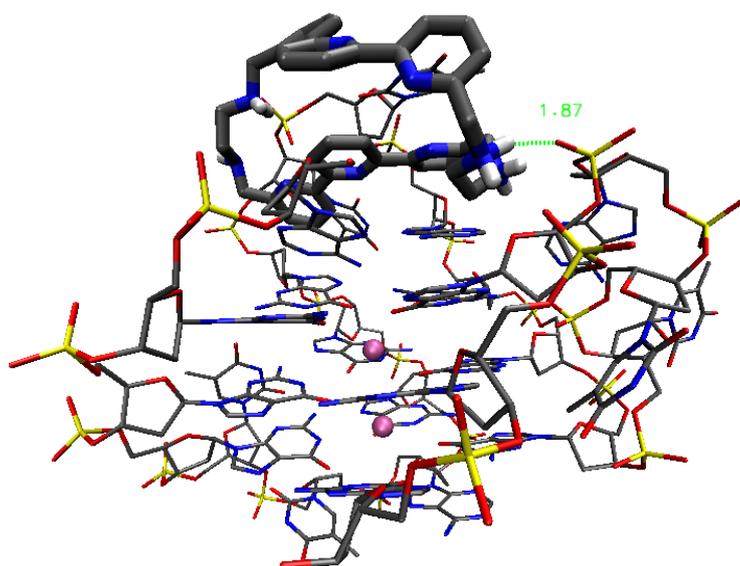


Figura 6-24 Interazioni a ponte a idrogeno tra il legante L1 e il DNA in avvolgimento hybrid2 (addotto L1_H2_R1)

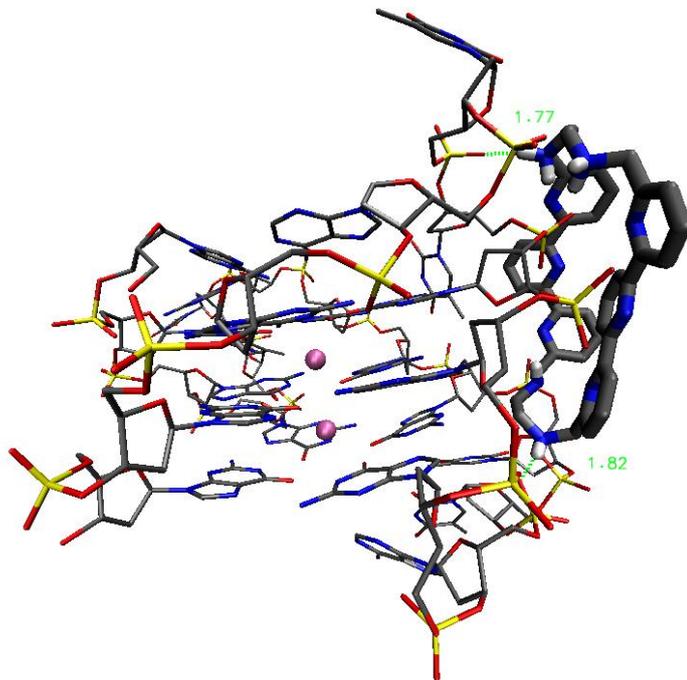


Figura 6-25 Interazioni a ponte a idrogeno tra il legante L1 e il DNA in avvolgimento hybrid1 (addotto L1_H1_R6)

7 Conclusioni e prospettive future

Da quando è stata scoperta la relazione tra telomeri, enzima telomerasi e tempo di vita delle cellule, in particolare di quelle tumorali, sono state svolte numerose ricerche su come inibire l'attività di questo enzima, sovraespresso nelle cellule tumorali. Uno dei metodi possibili prevede di sfruttare la presenza di strutture G-quadruplex nel filamento del DNA telomerico, che impediscono l'interazione tra DNA ed enzima. Per questo motivo, si cerca di sviluppare leganti che possano indurre e stabilizzare queste strutture.

I leganti L1, L2 e L3 sono in grado di legarsi al DNA telomerico umano. In questo lavoro di tesi si è cercato di ottenere informazioni sulla modalità con la quale questi leganti interagiscono con alcune delle conformazioni del DNA telomerico riscontrate sperimentalmente più di frequente. Sono stati seguiti due diversi approcci: sono stati eseguiti esperimenti di cristallizzazione, allo scopo di ottenere cristalli adatti all'indagine diffrattometrica a raggi X, e sono stati effettuati calcoli teorici basati su metodi di meccanica molecolare. Gli studi di cristallizzazione non hanno portato a risultati utili per proseguire l'analisi diffrattometrica, ma hanno messo in evidenza una maggiore propensione a cristallizzare di L3, rispetto agli altri due leganti.

I calcoli teorici hanno dimostrato che tutti e tre i leganti possono interagire con le topologie di DNA telomerico considerate, in particolare le topologie hybrid1, hybrid2 e propeller. Negli addotti ottenuti i leganti interagiscono sia con il quartetto sia con i solchi e i loop. Anche in questo caso, il legante che mostra una maggiore affinità per il DNA telomerico è il legante L3. Questo dato è in accordo con gli studi preliminari d'interazione in soluzione, nei quali L3 con il DNA telomerico mostrava la temperatura di melting più elevata. Per questa ragione, sarebbe interessante proseguire il lavoro di calcolo, in particolare la parte di dinamica molecolare, che al momento è a uno stadio preliminare per via della lunghezza dei tempi di calcolo. Può sicuramente essere migliorata la simulazione del solvente. Utilizzando tempi più lunghi, si potrebbe inoltre valutare meglio le interazioni con il quartetto nelle conformazioni ibride; in queste conformazioni, infatti, il quartetto è molto più ingombrato

dai loop rispetto a una conformazione propeller, ma con una simulazione di durata maggiore, come si è notato durante le varie prove di dinamica, aumenta la mobilità dei loop e, di conseguenza, diminuisce l'ingombro sul quartetto.

È interessante notare che i calcoli teorici mettono in evidenza una particolare propensione di L3 a interagire con il DNA telomerico in conformazione propeller. Poiché è noto che allo stato solido il DNA telomerico è sempre stato trovato in questa conformazione, l'evidenza sopra riportata potrebbe spiegare come mai solo le prove di cristallizzazione effettuate hanno dato risultati parzialmente promettenti solo con questo legante. Un ulteriore sviluppo di questo lavoro di tesi sarà quindi volto a migliorare le prove di cristallizzazione dei tre leganti, e in particolare di L3, allo scopo di proseguire l'indagine strutturale.

8 Bibliografia

1. Dove non altrimenti specificato la fonte è David L. Nelson, Michael M. Cox, I principi di biochimica di Lehninger, Zanichelli
2. Bang I. Untersuchungen über die Guanylsäure. *Biochem. Z.* 26, 293-231 (1910)
3. Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. *Proc. Natl Acad. Sci. USA* 48, 2013–2018 (1962).
4. Huppert, J. L. Structure, location and interactions of G-quadruplexes. *FEBS J.* 277, 3452–3458 (2010)
5. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* 59, 871–880 (1989).
6. Hazel, P., Parkinson, G. N. & Neidle, S. Predictive modelling of topology and loop variations in dimeric DNA quadruplex structures. *Nucleic Acids Res.* 34, 2117–2127 (2006)
7. C.B. Harley, A.B. Futcher, C.W. Greider, Telomeres shorten during ageing of human fibroblasts, *Nature* 345 (1990) 458e460.
8. J. W. Shay, S. Bacchetti, *Eur. J. Cancer* 1997, 33, 787–791
9. A.M. Zahler, J.R. Williamson, T.R. Cech, D.M. Prescott, Inhibition of telomerase by G-quartet DNA structures, *Nature* 350 (1991) 718e720
10. Kelland, L. *Clinical Cancer Research*, 2007, 13, 4960-4963
11. G-Quadruplexes: Targets in Anticancer Drug Design, Tian-miao Ou, Yu-jing Lu, Jiaheng Tan, Zhi-shu Huang, Kwok-Yin Wong, Lian-quan Gu
12. Idem
13. D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, L. H. Hurley, *J. Med. Chem.* 1997, 40, 2113–2116
14. S. Bhattacharya, P. Chaudhuri, A. K. Jain and A. Paul, *Bioconjugate Chem.*, 2010, 21, 1148
15. M. A. Read, A. A. Wood, J. R. Harrison, S. M. Gowan, L. R. Kelland, H. S. Dosanjh, S. Neidle, *J. Med. Chem.* 1999, 42, 4538–4546
16. B. Majian, S. Bhattacharya, *Chem. Commun.*, 2014, 50, 6422
17. I. Naasani, H. Seimiya, T. Yamori, T. Tsuruo, *Cancer Res.* 1999, 59, 4004–4011
18. Pedersen, C.L. *Angew. Chem. Int.* 1988, 27, 1021-1027
19. R. T. Wheelhouse, D. Sun, H. Han, F. X. Han, L. H. Hurley, *J. Am. Chem. Soc.* 1998, 120, 3261–3262
20. A. Arola-Arnal, J. Benet-Buchholz, S. Neidle and R. Vilar, *Inorg. Chem.*, 2008, 47, 11910

21. Q. Yu, Y. Liu, J. Zhang, F. Yang, D. Sun, D. Liu, Y. Zhou and J. Liu, *Metallomics*, 2013, 5, 222
22. Recettori polifunzionali contenenti unità eteroaromatiche per il riconoscimento di substrati, Tesi di Dottorato in Scienze Chimiche del Dott. Samuele Santarelli
23. S.N. Rao, U.C. Singh and P.A. Kollmana, *Nature* 328, 1987, 551
24. W.L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* 110, 1988, 1657-1666
25. Un punto di sella di una funzione reale di più variabili reali $f: \mathbb{R}^n \rightarrow \mathbb{R}$ è un punto critico P del dominio f della in cui la matrice hessiana risulta indefinita
26. www.schrodinger.com/Maestro
27. www.schrodinger.com/MacroModel
28. Bochevarov, A.D.; Harder, E.; Hughes, T.F.; Greenwood, J.R.; Braden, D.A.; Philipp, D.M.; Rinaldo, D.; Halls, M.D.; Zhang, J.; Friesner, R.A., "Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences," *Int. J. Quantum Chem.*, **2013**, 113(18), 2110-2142
29. Drenth J. Principles of protein X-ray crystallography, Ed Springer-Verlag, New York, 213, 1994
30. C. Bazzicalupi, A. Bencini, A. Bianchi, E. Faggi, C. Giorgi, S. Santarelli, B. Valtancoli, *Coord. Chem. Rev.* 2007
31. C. Bazzicalupi, A. Bencini, A. Bianchi, E. Faggi, C. Giorgi, S. Santarelli, B. Valtancoli, *J. AM. CHEM. SOC.* 2008, 130, 2440-2441
32. I. Bessi, C. Bazzicalupi, C. Richter, H. R. A. Jonker, K. Saxena, C. Sissi, M. Chioccioli, S. Bianco, A. Rita Bilia, H. Schwalbe, P. Gratterer *ACS Chem. Biol.*, 2012, 7 (6), pp 1109–1119
33. Le immagini degli addotti sono state realizzate con il programma VMD: Humphrey, W., Dalke, A. and Schulten, K., "VMD - Visual Molecular Dynamics" *J. Molec. Graphics* **1996**, 14.1, 33-38