

Tesi Doctoral

**Formació del motiu quàdruplex *bi-loop* amb oligonucleòtids lineals
i aplicació a la ciclació assistida per motlle**

Júlia Viladoms Claverol



UNIVERSITAT DE BARCELONA



Dept. de Química Orgànica
Facultat de Química
Universitat de Barcelona

Memòria presentada per

Júlia Viladoms Claverol

per optar al grau de Doctora per la Universitat de Barcelona

Dirigida i revisada per

Dr. Enrique Pedroso Muller

Programa de Doctorat en Química Orgànica
Bienni 2002-2004

Dept. de Química Orgànica
Facultat de Química
Universitat de Barcelona

Barcelona, setembre de 2008

ÍNDEX GENERAL

INTRODUCCIÓ I OBJECTIUS

A. Els àcids nucleics: formes no canòniques i relació estructura-funció	1
B. Els oligonucleòtids cíclics com a models per a estudis estructurals	7
C. Síntesi d'oligonucleòtids cíclics	8
D. El motiu <i>bi-loop</i>	9
E. Objectius	10

RESULTATS I DISCUSSIÓ

BLOC 1: Aprofundint en l'estudi del motiu *bi-loop*

1. Fonaments	11
2. Introducció i objectius	35
3. El motiu <i>bi-loop</i> heterolineal	49
4. <i>Bi-loop</i> estabilitzat per parells no canònics	69
5. El motiu <i>bi-loop</i> amb oligonucleòtids lineals	83
6. Generalitat i rellevància del motiu <i>bi-loop</i>	123

BLOC 2: Preparació d'oligonucleòtids cíclics per ciclació assistida per motlle

1. Introducció i objectius	135
2. Síntesi dels oligonucleòtids lineals fosforilats	143
3. Assajos de ciclació amb models de dúplex	151
4. El <i>bi-loop</i> com a motlle per a la ciclació assistida	155
5. Intent d'aplicació de la ciclació assistida per motlle a la fase sòlida	171

CONCLUSIONS	183
--------------------	------------

PART EXPERIMENTAL

1. Materials i Mètodes	187
2. Síntesi d'oligonucleòtids lineals	201
3. Síntesi d'oligonucleòtids cíclics	207
4. Estudis mitjançant UV i CD	215
5. Estudis estructurals d'oligonucleòtids per RMN	219
6. Estudis mitjançant DSC	227
7. Ciclació d'oligonucleòtids assistida per motlle	229

BIBLIOGRAFIA	233
---------------------	------------

ANNEX	261
--------------	------------

INTRODUCCIÓ I OBJECTIUS



INTRODUCCIÓ I OBJECTIUS

A. Els àcids nucleics: formes no canòniques i relació estructura-funció¹

A part de la coneguda doble hèlix descrita per Watson i Crick² ja fa més de 50 anys, els àcids nucleics existeixen a la natura en una varietat de formes estructurals. Totes aquestes formes impliquen un apilament de les bases i la formació de ponts de hidrogen, però presenten importants diferències pel que fa a la seva morfologia: molecularitat, nombre de cadenes que interaccionen, tipus i nombre d'enllaços de hidrogen que es formen i estabilitat, entre altres. La flexibilitat intrínseca del DNA és la que li permet adoptar estructures tan diverses, que van des de dúplexs levogirs fins a estructures multicadena tipus triplex o quàdruplex, així com empaquetar-se en els cromosomes i realitzar la seva funció biològica. Aquesta flexibilitat prové en gran part de la llibertat conformacional de l'anell de ribosa i, en menor grau, de la rotació entorn l'enllaç glicosídic i les rotacions de l'esquelet fosfat.

No obstant, la propietat fonamental del DNA per exercir la seva funció biològica és la seva gran capacitat de reconeixement. En aquesta capacitat es basa l'emmagatzematge de la informació genètica de la qual és portador. Però, contràriament al que es creia anteriorment, el DNA no només és portador d'informació en la seva seqüència, sinó que també codifica informació en la seva estructura. Així doncs, tot i ser una molècula aparentment simple, presenta una enorme versatilitat estructural, clau de la seva importància biològica, ja que les estructures no canòniques afecten importants processos, com la replicació, la transcripció, la recombinació o la reparació del DNA³.

L'RNA també existeix a la natura en una infinitat de formes estructurals; els elements d'estructura secundària tipus cadena senzilla, fragments de doble cadena (tija o *stem*), *hairpin loops*, *bulge loops* i *junctions* construeixen estructures terciàries molt complexes. A més a més, és conegut des dels anys 80⁴ que l'RNA no és un simple transmissor de la informació genètica sinó que posseeix capacitat catalítica. Aquests RNAs amb capacitat per escindir o formar enllaços són els anomenats ribozims. Aquest descobriment va revolucionar la comunitat científica ja que es creia que només les proteïnes tenien capacitat catalítica i va obrir una nova possibilitat per a l'origen de la vida: el món d'RNA. L'RNA podria haver estat una forma de vida precel·lular ja que pot ser tant portador de la informació

genètica com catalitzador.

Els estudis estructurals dels àcids nucleics proporcionen informació sobre la seva conformació i les seves interaccions amb altres biomolècules, informació imprescindible per entendre com porten a terme la seva tasca biològica. Com bé diu Ignacio Jr. Tinoco: “Primer cal conèixer les estructures, abans que es puguin entendre les funcions” (“*Structures are required before functions can be understood*”)⁵. En alguns casos, les funcions eren conegudes des de feia temps, abans que se’n pogués determinar l’estructura, com és el cas del ribosoma o el nucleosoma. Però no és fins que es coneix l’estructura a nivell molecular que es pot entendre tota la complexitat de la funció que aconsegueixen. En altres casos, l’existència de l’estructura va arribar molt abans que se’n pogués determinar la rellevància biològica: és el cas del Z-DNA o del quàdruplex de guanina.

El fet que el DNA pugui adoptar tal varietat d’estructures, no només interessa des del punt de vista de les funcions biològiques, sinó també per la seva potencial aplicació terapèutica. A continuació, s’exposaran breument les diferents estructures que poden adoptar els àcids nucleics, des d’un punt de vista històric, dels rols que desenvolupen *in vivo* i del seu potencial terapèutic.

Estructures dúplex

En els organismes el DNA es troba normalment en forma de doble cadena, però existeixen diferents formes de dúplex. Ja els primers experiments de difracció de RX de fibres de Maurice Wilkins⁶ i Rosalind Franklin⁷ van mostrar l’existència de dues formes diferents del DNA: el B-DNA en condicions d’alta humitat i l’A-DNA en condicions d’humitat baixa. La forma B és l’estructura que adopta el DNA en les cèl·lules i presenta 10 nucleòtids per gir, amb conformació C2’-endo i disposició *anti* de les bases respecte a l’enllaç glicosídic*. Els parells de bases es troben pràcticament perpendiculars a l’eix de l’hèlix i la superfície d’aquesta presenta dos solcs: el solc major i el solc menor (Fig. 1). La forma A és adoptada pels fragments de doble cadena d’RNA i essencialment pel híbrids DNA-RNA⁸. També es pot donar en dúplex de DNA en condicions d’alta concentració de cations o baixa hidratació. L’A-DNA presenta 11 nucleòtids per gir amb la conformació C3’-endo i *anti*. Els parells de bases no es troben sobre l’eix de la molècula i estan lleugerament inclinats, de manera que l’estructura és més compacta i els solcs són diferents: el solc major és estret i profund, mentre que el solc menor és ample i pla (Fig.1). Aquestes diferències són crucials per entendre, per exemple, com les proteïnes reconeixen una estructura de DNA bicatenari (dsDNA) enfront d’una de dsRNA. A més a més, conèixer l’estructura dels híbrids DNA-RNA és de vital importància per entendre el mecanisme de reconeixement de l’RNAsa H i poder preparar anàlegs que l’activin per la teràpia antisentit.

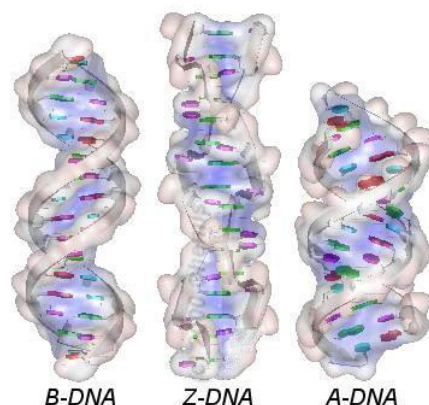


Figura 1. Diferents conformacions dúplex del DNA

* Veure bloc 1, capítol 1, apartat 1.1 sobre estructura d’àcids nucleics per a més informació sobre nomenclatura.

Curiosament, el primer cristall de DNA que es va aconseguir resoldre, va resultar mostrar un dúplex levogir ben diferent de les formes A i B, que per la seva forma en zig-zag va ser anomenat Z-DNA⁹. Aquesta forma resulta de l'alternança de purines amb la conformació C3'-*endo/syn* i pirimidines amb la conformació C2'-*endo/anti*, de manera que la unitat repetitiva és un dinucleòtid. L'existència d'una conformació levogira ja havia estat proposada en base als estudis de dicroisme circular, que havien mostrat perfils quasi inversos per a seqüències de poli(dG-dC) en augmentar la força salina del medi¹⁰. La transició de B-DNA a Z-DNA s'observa especialment amb seqüències de pirimidines i purines alternades, ja que aquestes adopten més fàcilment la conformació *syn*, i en condicions d'elevada força iònica, per apantallar les repulsions entre els fosfats propers. La metilació de C5 de la citosina i N7 de la guanina també afavoreix aquesta transició. Aquest dúplex levogir és allargat i prim (12 nt per volta) i només presenta un solc estret i profund (Fig.1). El Z-DNA i el Z-RNA existeixen *in vivo*, com ho demostra el fet que són antigènics i l'existència de proteïnes que reconeixen amb gran afinitat i especificitat aquestes formes. El Z-DNA està estabilitzat en condicions de superenrotllament negatiu, que es dona durant la transcripció, i es creu que està implicat en la regulació gènica ja que existeix un alt predomini de seqüències que afavoreixen el Z-DNA en zones properes als llocs de transcripció i existeixen indicis de la relació entre l'activitat transcripcional i la presència de Z-DNA¹¹.

El DNA també pot formar dúplex dextrogirs però paral·lels, de manera que les dues cadenes es troben en la mateixa orientació. Aquesta doble cadena es troba estabilitzada per parells WC *reversed* i es pot observar si s'escull una seqüència que no pugui formar un dúplex antiparal·lel¹² o es prepara un *hairpin* unint les dues cadenes pel mateix extrem: 5'-*loop*-5' o 3'-*loop*-3'¹³. També es poden formar dúplex dextrogirs paral·lels estabilitzats per parells Hoogsteen¹⁴. A més a més, existeix una altra forma de DNA dextrogir paral·lel que es dona a pHs baixos¹⁵ i està estabilitzat per parells entre bases iguals, del tipus G·G, A·A, C·C⁺ i T·T. Aquest motiu ha estat anomenat Π -DNA i està afavorit per la seqüència 5'-CGA-3'¹⁶. Sorprenentment l'estabilitat de totes aquestes formes de DNA paral·lel només és lleugerament inferior al B-DNA. No se'n coneix cap funció biològica, però podria estar implicat en el reconeixement entre cadenes senzilles de DNA o, més probablement, d'RNA.

Estructures tríplex

L'any 1957 ja es va demostrar que una barreja 1:2 de poli(dA)-poli(dU) podia formar una estructura tipus tríplex¹⁷, però no va ser fins als primers estudis mitjançant ressonància magnètica nuclear a finals dels anys 80¹⁸ que es va confirmar l'aparellament proposat per a la tercera cadena en els tríplexs pirimidínic o paral·lels: T-A·T i C⁺-G·C (Y-R·Y, aparellament tipus Hoogsteen). Aquest tipus de tríplex s'ha anomenat pirimidínic o paral·lel atesa la natura de la tercera cadena i la seva orientació respecte la component purínica del dúplex al qual s'uneix. També existeix un altre tipus de tríplex, descobert posteriorment i anomenat purínic o antiparal·lel (R-Y·R), en què les tríades formades són G-G·C i A-A·T o T-A·T (enllaços *reversed* Hoogsteen)¹⁹. En ambdós casos, la tercera cadena es disposa en el solc major del dúplex, que manté la seva estructura però es troba lleugerament desenrotllat, i la seva conformació és C2'-*endo*²⁰ (Figura 2.A).

En el DNA biològic es poden formar triples hèlixs locals, anomenades H-DNA. Es tracta d'una estructura secundària que es forma en seqüències polipirimidíniques simètriques sota situacions

d'estrès superhelicoidal o a baix pH; la cadena polipirimidínica es separa de la complementària, es replega sobre ella mateixa i s'hibrida a la cadena polipurínica veïna formant una triple hèlix²¹ (Figura 4.A). Aquesta estructura també es pot formar segons el motiu tríplex purínic²² (*H-DNA). La seva funció biològica no està clara, però l'abundància de seqüències formadores d'aquest motiu és molt major en eucariotes que en procariotes i en general es troben en introns o regions no traduïdes. Per això s'ha hipotetitzat que l'H-DNA pot tenir un paper important com a senyal regulador de la replicació o la transcripció en eucariotes²³. A més a més, també s'ha observat que la formació d'H-DNA estimula la recombinació homòloga entre repeticions situades a banda i banda de l'estructura²⁴ i sembla que pot estar implicat en l'expansió de certes repeticions de triplets²⁵.

L'interès en les hèlix triples o tríplex també prové de la seva potencial aplicació per al control de la regulació genètica en la teràpia antigen, en què un oligonucleòtid sintètic es dirigeix al DNA cromosòmic per formació d'una estructura tríplex. L'avantatge d'aquesta estratègia enfront de l'antisentit, dirigit a l'mRNA, rau en la seva major efectivitat, ja que es bloqueja la informació a l'origen, i la seva gran especificitat.

Estructures quàdruplex

Les estructures quàdruplex també es coneixen des de fa molt temps. Ja l'any 1958 es va veure que les fibres d'àcid poliinosínic donaven lloc a patrons de difracció que es podien explicar per formació de cadenes tríplex o quàdruplex. Originalment es van atribuir a una estructura tríplex²⁶, però posteriorment, l'any 1974, emprant també difracció de Raigs X de fibres, ja es va proposar una estructura quàdruplex²⁷, estabilitzada per tètredes de guanina. A finals dels anys 80 es va observar que aquestes estructures podien formar-se en l'element ric en guanines de la regió de canvi de la immunoglobulina (*immunoglobuline switch region*)²⁸ i en els telòmers²⁹. Però les primeres estructures de quàdruplex de G d'alta resolució daten del 1992³⁰; el mateix any es va publicar l'estructura de la seqüència telomèrica d'*Oxytricha* tant per cristal·lografia de RX^{30a} com per ressonància magnètica nuclear i mètodes de càlcul^{30b}. Des d'aleshores, les múltiples estructures determinades han mostrat que el quàdruplex de guanina és altament polimòrfic, amb diferents topologies respecte al nombre i l'orientació de les cadenes, conformació *syn* o *anti* de l'enllaç glicosídic de les guanines i disposició dels llaços. A la Figura 2.B es mostra l'estructura d'un quàdruplex de guanina unimolecular. A més a més, dins d'estructures de quàdruplex de G, s'han observat altres tipus de tètredes formades per parells WC, com ja es comentarà amb detall més endavant (introducció del primer bloc, apartat 2.3.2). Així doncs, l'existència dels quàdruplex de guanina era coneguda des dels anys 70, però es creia que era merament un "artefacte" i no era rellevant biològicament. En canvi, actualment és un dels camps més actius en biologia estructural per dos motius. En primer lloc, s'ha demostrat que els telòmers de diversos organismes es poden estructurar en forma de quàdruplex de G^{30,31}, i aquest fet té implicacions en la regulació de la llargada dels telòmers i l'envelliment cel·lular. Els agents estabilitzadors de quàdruplex poden ser inhibidors de la telomerasa, la qual requereix l'extrem 3' dels telòmers en forma de cadena senzilla, i poden tenir aplicació en teràpia del càncer³². En segon lloc, existeixen seqüències riques en guanines en altres regions del genoma humà, incloent regions promotores d'oncogèns, centròmers, llocs de recombinació, repeticions implicades en l'expansió de

triplet repeats o dominis de dimerització d'RNA³³. Per últim, cal dir que existeixen proteïnes que reconeixen selectivament aquests quàdruplexs i n'afavoreixen la seva formació o el seu desenrotllament³⁴, i que s'ha demostrat que l'RNA també pot formar estructures de quàdruplex de guanina³⁵.

Les seqüències riques en guanines van acompanyades de la seva cadena complementària, que és rica en citosines i també pot formar una estructura tetracatenària especial. Inicialment, als anys 60, es pensava que l'àcid policitílic formava un dúplex paral·lel amb parells C·C⁺ a pHs baixos, tant pels estudis de difracció de RX de fibres³⁶ com per estudis de les seves propietats en solució³⁷. Quina va ser la sorpresa, 30 anys més tard, quan es va descriure l'estructura en solució de la primera seqüència rica en C i es va veure que formava una estructura tipus quàdruplex³⁸ (Figura 2.C). Aquesta estructura es va anomenar *i-motif* ja que està formada per dos dúplexs paral·lels amb parells C·C⁺ intercalats, disposats antiparal·lelament un respecte l'altre. Aquesta estructura, obtinguda mitjançant ressonància magnètica nuclear i dinàmica molecular restringida, va ser confirmada mitjançant cristal·lografia de RX³⁹. Posteriorment, es va observar que l'*i-motif* podia ser unimolecular i que la seqüència complementària a la telomèrica pot formar aquest motiu en diferents organismes⁴⁰. Els genomes eucariòtics també contenen seqüències riques en C en altres zones importants, com regions promotores, centròmers o en els *triplet repeats*. Recentment s'ha vist que l'*i-motif* pot presentar diferents topologies⁴¹, que l'RNA també pot formar aquest motiu⁴² i, a més a més, que existeixen proteïnes que interaccionen selectivament amb regions riques en citosines⁴³, fet que en reforça la seva implicació biològica.

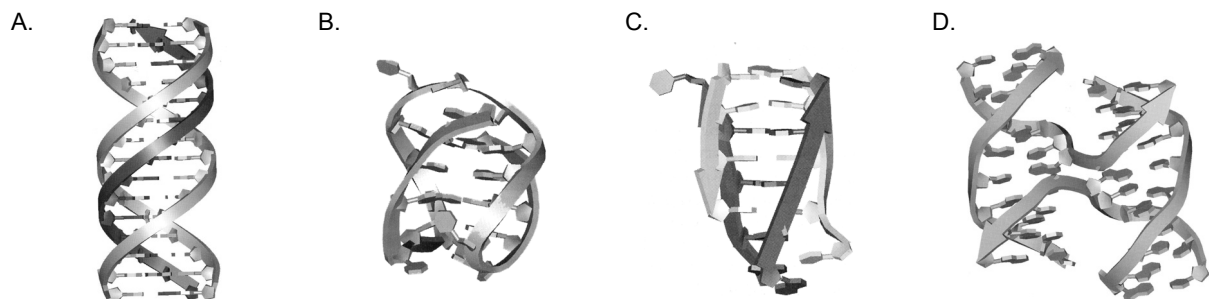


Figura 2. Algunes estructures no canòniques del DNA⁴⁴. A. Triple hèlix. B. Quàdruplex de guanina. C. Quàdruplex *i-motif*. D. Holliday junction

Encreuaments de quatre cadenes (*Four-way junctions*)

L'encreuament de quatre cadenes més estudiat és l'estructura proposada el 1964 per Robin Holliday com a intermedi clau en els processos de recombinació homòloga⁴⁵ i anomenada en nom seu "*Holliday junction*". Es forma entre dúplexs homòlegs i es tracta de quatre cadenes que interaccionen entre elles i cadascuna forma part de dos dels quatre braços dels que consta l'estructura, que té forma d'X (veure Figura 2.D). Aquesta estructura pot estar estesa o més compactada segons el medi i ja havia estat modelada a partir de dades obtingudes mitjançant tècniques biofísiques abans d'obtenir-se, de forma casual, les primeres estructures d'alta resolució al tombant del segle XXI⁴⁶. Aquest encreuament de quatre cadenes (*four-way junction*, 4H) pot desplaçar-se al llarg de la cadena

de DNA (migració de cadenes) o pot ser transformat per enzims per tal de completar el procés d'intercanvi (resolució).

Un altre tipus d'encreuament de quatre cadenes són les estructures cruciformes, que es formen en fragments de DNA amb seqüències palindròmiques (repeticions invertides) sota situacions d'estrès superhelicoidal, extrudint els parells de bases de la seqüència palindròmica i formant dues forquetes⁴⁷ (veure Figura 4.B). Aquestes estructures existeixen *in vivo* i es creu que estan implicades en la regulació de la replicació i la transcripció, ja que aquestes repeticions invertides es troben sovint properes a regions promotores.

Regions amb bases desaparellades

A banda dels llaços de bases desaparellades que existeixen en les cruciformes, també podem trobar bases desaparellades en una de les cadenes d'un dúplex (protuberància o *bulge*), en les dues cadenes d'un dúplex (*loop* intern), en un encreuament o en el llaç d'una forqueta (*hairpin*). En el DNA, les protuberàncies poden provocar mutacions a causa d'errors en la replicació o la transcripció i en l'RNA formen part d'elements de reconeixement importants. Les forquetes, que contenen una regió de doble hèlix i un llaç de bases desaparellades, es poden formar *in vivo* quan existeixen repeticions directes en el DNA (Figura 4.C).

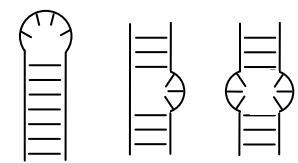


Figura 3. Estructures amb bases desaparellades: llaç, protuberància i bucle intern

Només una petita part del genoma humà codifica per proteïnes; segurament la resta del genoma "codifica" informació en la seva estructura per a la regulació de la primera part. En les regions no codificants existeixen moltes seqüències de DNA simètriques: repeticions invertides o palindròmiques, repeticions mirall i repeticions directes. Aquestes repeticions, sota situacions d'estrès superhelicoidal o per efecte de proteïnes, poden donar lloc a diferents tipus d'estructures no canòniques (Figura 4), estructures que semblen estar implicades en processos de reorganització genòmica, base de diverses malalties genètiques⁴⁸.

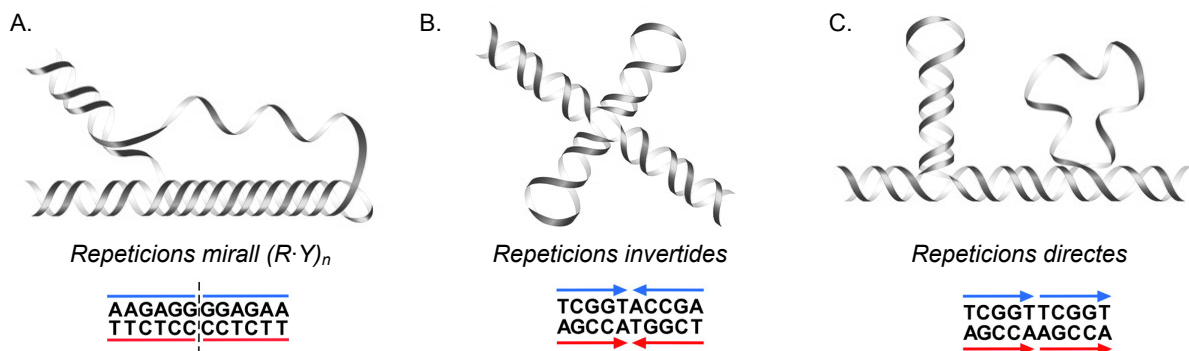


Figura 4. Exemples d'estructures que es poden formar *in vivo* segons el tipus de repeticions presents⁴⁸:
A. Tríplex (H-DNA). B. Forqueta. C. Estructura cruciforme

Alguns dels exemples citats reforcen la premissa que les formes no canòniques del DNA poden tenir important rellevància biològica pel funcionament i la regulació dels gens. És poc probable que existeixin estructures alternatives al B-DNA sense cap funció biològica. Com bé va dir Alexander Rich:

“La natura és oportunista i aprofitarà tot el que tingui al seu abast”. (“*Nature is opportunistic. It will use what is there*”)⁴⁹. El descobriment de noves estructures del DNA representa un repte per entendre el rol que poden desenvolupar *in vivo*. (“*There is a constant challenge to relate conformational changes which are discovered in vitro to changes that occur in vivo*”)⁵⁰.

B. Els oligonucleòtids cíclics com a models per a estudis estructurals

Així doncs, el coneixement de les diferents estructures en què els àcids nucleics es presenten a la natura és fonamental per poder entendre a escala molecular la seva funció. Però el seu estudi pot resultar complicat per la competència amb altres estructures i/o la baixa estabilitat de les formes que es volen estudiar. En aquest sentit, els oligonucleòtids cíclics representen un bon model per als estudis estructurals i conformacionals gràcies a la restricció intrínseca de la seva flexibilitat respecte als seus homòlegs lineals. Aquesta restricció conformacional induïda per la ciclació dels extrems pot augmentar l'estabilitat de l'estructura a estudiar i, encara més important, impedir la formació d'estructures alternatives, fet que en facilita l'estudi tant en estat sòlid, mitjançant cristal·lografia de RX, com en solució, mitjançant RMN.

En els primers estudis sobre l'estabilitat i estructura en solució de dúplexs petits autocomplementaris existien problemes ja que les corbes de fusió presentaven una doble transició, de dúplex a forqueta (*hairpin*) i posterior fusió a estructura desordenada (cabdell estadístic o *random coil*). A més a més, la temperatura i la forma de fusió dels dúplexs petits depenien de la concentració, i els extrems dels dúplexs no es mostraven totalment aparellats. Sovint era difícil comparar els resultats dels experiments d'RMN amb els estudis de fusions per tècniques espectroscòpiques a causa de les grans diferències de concentració en què es realitzaven aquests experiments. Una manera d'eliminar aquesta dependència de la concentració, la doble fusió i la desorganització dels extrems era lligant-los i fent l'oligonucleòtid cíclic⁵¹. Aquests models cíclics es van anomenar *dumbbell* (que es traduiria per “mancuerna” en castellà i “coble” en català) i contenen un fragment central de doble hèlix i dos *loops* o llaços amb bases no aparellades (Figura 5). D'aquesta manera es podia estudiar la fusió dels dúplexs a *random coil* de forma independent a la concentració. Dels primers estudis se'n va derivar tota una col·lecció⁵² de treballs dedicats a les propietats dels *dumbbells* de diferents mides i en diferents medis, per extreure dades d'estabilitat en funció de la seqüència (teoria del veí més pròxim, *nearest-neighbour theory*) i comparar-la amb la dels seus anàlegs lineals. Altres autors també van emprar els oligonucleòtids cíclics per als primers estudis termodinàmics d'àcids nucleics⁵³.

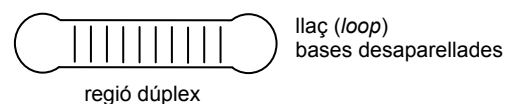


Figura 5. Estructura dumbbell

Els *dumbbells* també són bons models per als estudis estructurals de forquetes, llaços i cruciformes, ja que eviten la formació del dúplex bimolecular corresponent. Des que van ser proposats com a models de petites forquetes⁵⁴, s'han emprat amb aquesta finalitat en diversos estudis⁵⁵, per determinar la influència de la mida i seqüència en l'estabilitat dels llaços. Per exemple, amb el

decàmer cíclic de seqüència d<pCGC-TT-GCG-TT>, s'ha pogut estudiar mitjançant RMN la diferència entre els *loops* flanquejats pel mateix parell però amb diferent direccionalitat: 5'-CTTG-3', que forma llaços de dos residus, *versus* 5'-GTTC-3', que adopta *loops* de quatre residus^{55b}. Els oligonucleòtids cíclics també s'han emprat com a models per estudiar la formació d'estructures cruciformes en seqüències de repeticions invertides⁵⁶.

C. Síntesi d'oligonucleòtids cíclics

Per poder emprar els oligonucleòtids cíclics com a models es van haver de desenvolupar metodologies per a la seva síntesi. Actualment existeixen diferents metodologies per a la preparació d'aquests oligonucleòtids: metodologies de ciclació en solució, metodologies de ciclació en fase sòlida i metodologies de ciclació assistida per motlle (*template-directed ligation*). En les dues primeres, la ciclació té lloc per unió dels extrems hidroxil i fosfat desprotegits, mentre que la resta de grups funcionals es troben protegits convenientment. Per contra, en la ciclació assistida per motlle, el precursor lineal, totalment desprotegit, es preorganitza en solució, amb l'ajuda o no d'un motlle extern, de manera que els extrems que han de reaccionar es troben propers en l'espai i es pot produir la seva lligació en presència d'un agent condensant. Com a motlles per a la ciclació assistida s'han emprat pràcticament totes les estructures conegudes, basades tant en dúplex com en tríplex de diferents tipus i en estructures quàdruplex de guanina o *i-motif*. Aquesta metodologia, a banda de ser un mètode de preparació d'oligonucleòtids cíclics més o menys versàtil segons l'estructura emprada, constitueix una evidència de la formació de l'estructura en solució. Els tres tipus d'aproximacions per a la preparació d'oligonucleòtids cíclics (Figura 6) s'explicaran amb major detall en la introducció del segon bloc d'aquesta tesi.

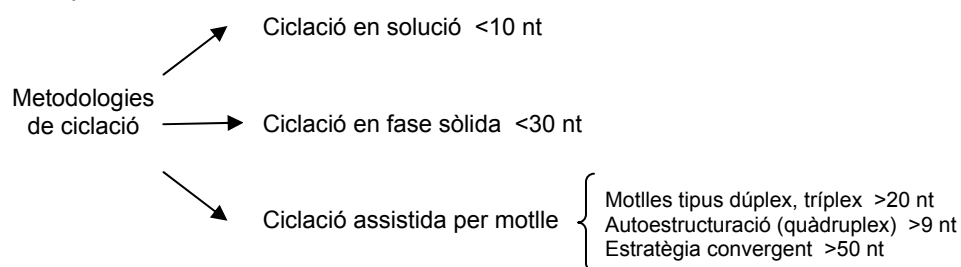


Figura 6. Diferents estratègies de preparació d'oligonucleòtids cíclics en funció de la seva mida

Al nostre grup de treball s'ha desenvolupat un mètode de síntesi d'oligonucleòtids cíclics en fase sòlida⁵⁷ que permet sintetitzar oligonucleòtids cíclics de petita i mitjana grandària (de 2 a 30 mer amb rendiments moderats) sense restricció de seqüència. La unió a la fase sòlida es realitza a través del grup fosfat en 3', l'elongació de la cadena utilitza el mètode del fosfit triester i la ciclació té lloc en fase sòlida en unes condicions de pseudodilució. El gran avantatge del mètode rau en el desancoratge selectiu dels oligonucleòtids ciclats envers els productes lineals o possibles productes de reacció intercadena, ja que es trenquen selectivament els enllaços fosfats triester (unió dels cicles a la resina) envers els diester (unió de les cadenes no ciclades a la resina). Com que aquest és el mètode que s'ha utilitzat en aquest treball, s'explicarà amb major detall en els capítols 1 i 3 del bloc 1. El problema d'aquesta metodologia rau en el cost entròpic de la ciclació, que es fa major a mesura que

s'augmenta la mida del cycle i provoca que els rendiments de ciclació per a oligonucleòtids majors de 15 mer siguin ja força baixos.

D. El motiu *bi-loop*

L'any 1995 es va trobar una estructura curiosa en cristal·litzar un heptàmer gairebé totalment autocomplementari: $d(\text{GCATGCT})^{58}$. En comptes del dúplex que es pretenia estudiar, es va observar la dimerització de l'oligonucleòtid doblegat, com si es tractés d'una interacció entre dos *hairpins* o forquetes, implicant la interacció de quatre cadenes i donant lloc a una estructura tipus quàdruplex. Aquest fet no hauria tingut major rellevància que el de curiositat cristal·logràfica si no s'hagués tornat a observar una estructura molt similar en cristal·litzar un octàmer cíclic de seqüència totalment diferent, $d\langle p\text{CATTTCATT} \rangle^{59}$, quan el que s'intentava era estudiar el *dumbbell* més petit possible. En la Figura 7.A es pot observar l'extraordinària semblança entre ambdues estructures, tot i estar estabilitzades per parells WC diferents: G·C en el cas de l'heptàmer lineal i A·T en el cas de l'octàmer cíclic. Aquest fet, que no es podia deure a una simple coincidència, va fer pensar que potser ens trobàvem davant d'un nou motiu de caràcter general, que va ser anomenat *bi-loop*⁵⁹, atesa la seva forma (Figura 7.B).

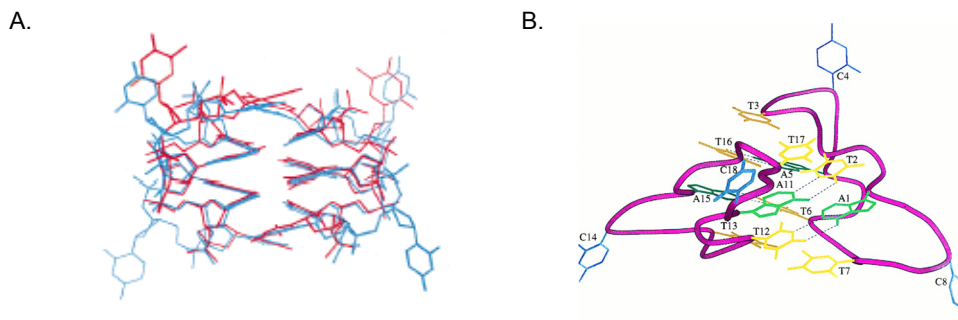


Figura 7. A. Superposició de les estructures cristal·lines de $d(\text{GCATGCT})$ (vermell) i $d\langle p\text{CATTTCATT} \rangle$ (blau). B. Representació esquemàtica del *bi-loop* en què l'esquelet de ribosa-fosfat es dibuixa com un llaç

Aquest descobriment va conduir a l'estudi de diversos octàmers cíclics mitjançant RMN. Així, es va poder determinar que l'oligonucleòtid cíclic $d\langle p\text{CATTTCATT} \rangle$, anteriorment estudiat, i l'octàmer $d\langle p\text{TGCTCGCT} \rangle$ que podria dimeritzar per formació de parells G·C, com en el cas de l'heptàmer, s'estructuraven en solució en forma d'aquest nou motiu tipus quàdruplex⁶⁰. Posteriorment, l'estudi de noves seqüències va conduir a l'observació de diversos tipus de tètades, com s'explicarà detalladament en la introducció del primer bloc. També es va estudiar el motiu format per dos cycles diferents, és a dir, el motiu heterodimèric.

El *bi-loop* és un motiu dimèric simètric en què les dues unitats es disposen antiparal·lelament i es manté estabilitzat per quatre parells de bases intermoleculares que formen dos apilaments. Els apilaments presenten una inclinació relativa de $\sim 35^\circ$ i enfronten els parells de bases pel solc menor formant dues tètades. Els parells de bases estan connectats per dos petits llaços de dos nucleòtids cadascun. La primera de les bases del llaç queda apilada sobre els parells de bases fent de tapa, mentre que la segona base es disposa cap a l'exterior i queda totalment exposada al solvent (Fig. 8).

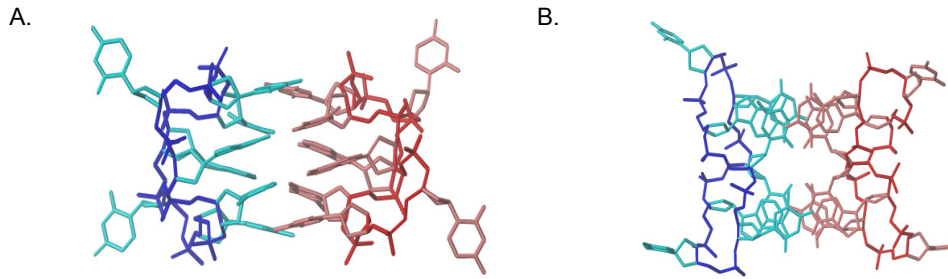


Figura 8. Vistes lateral (A) i superior (B) de l'estructura promig de d<pCATT⁵⁹ determinada mitjançant RMN i Dinàmica Molecular Restringida.

Es va postular que aquest motiu podria estar implicat en processos de recombinació⁵⁹ i per això seria útil la seva observació en altres seqüències, especialment en fragments lineals en solució. L'observació del *bi-loop* en oligonucleòtids lineals i alhora en solució seria d'especial importància per poder afirmar de manera convincent que aquest motiu es podria donar en el DNA cel·lular i podria tenir alguna rellevància biològica, ja que es podria descartar definitivament que es tractés d'un artefacte cristal·logràfic o un efecte de la ciclació. En estat sòlid, les forces d'empaquetament del cristall poden fer adoptar al DNA conformacions diferents a les formades en solució⁶¹. Així mateix, la ciclació d'oligonucleòtids tan petits podria resultar la responsable de la formació d'aquesta estructura.

E. Objectius

Tota la informació de què es disposava en iniciar-se aquesta tesi suggeria que el *bi-loop* podia ser un motiu de caràcter general, ja que s'havia observat tant en estat sòlid com en solució i tant en oligonucleòtids cíclics com en lineals. Mancava només la observació d'aquest motiu en oligonucleòtids lineals en solució per reforçar la seva possible rellevància biològica. Aquesta observació pot resultar complicada per la competència d'altres estructures també estabilitzades per parells WC i per això creïem que no s'havia tingut èxit fins al moment. En aquesta línia de treball, els objectius que ens vam proposar en iniciar aquesta tesi doctoral foren:

- Determinar si és possible la formació d'estructures heterodimèriques del motiu *bi-loop* amb una seqüència cíclica i l'altra lineal.
- Determinar si és possible la formació del motiu homodimèric amb oligonucleòtids lineals en solució.
- De manera general, aportar informació addicional a l'estudi estructural del motiu *bi-loop*.

Aquests objectius i els resultats que se'n deriven s'exposaran en el primer bloc d'aquesta memòria.

Una vegada abordats els primers objectius, i aprofitant la conformació i l'estabilitat del motiu *bi-loop*, ens proposàvem dissenyar i posar a punt un mètode de ciclació d'oligonucleòtids lineals dirigit per motlle. Aprofitaríem l'estructuració del precursor lineal en forma de *bi-loop* per lligar-ne els extrems i evidenciar així la seva formació en solució. En aquesta línia de treball de síntesi d'oligonucleòtids cíclics, també volíem intentar combinar les aproximacions de ciclació assistida per motlle i ciclació en fase sòlida per aprofitar els avantatges de cadascuna d'elles i superar-ne els inconvenients. Aquests objectius sintètics es descriuran en el segon bloc de la memòria.