

Tesi Doctoral

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

Júlia Viladoms Claverol



Dept. de Química Orgànica  
Facultat de Química  
Universitat de Barcelona

## RESULTATS I DISCUSSIÓ



**BLOC 1:**

**A**PROFUNDINT EN L'ESTUDI  
DEL MOTIU BI-LOOP



|   |           |
|---|-----------|
| <b>1 Fonaments</b>  | <b>11</b> |
| 1.1 Estructura d'àcids nucleics   | 11        |
| 1.1.1 Components, nomenclatura i numeració  | 11        |
| 1.1.2 Conformació   | 12        |
| 1.1.3 Interaccions entre bases  | 14        |
| 1.1.4 Paràmetres helicoidals per a àcids nucleics   | 15        |
| 1.2 Síntesi d'oligonucleòtids   | 17        |
| 1.2.1 Síntesi d'oligonucleòtids lineals   | 17        |
| 1.2.2 Síntesi d'oligonucleòtids cíclics   | 18        |
| 1.3 Aplicació de l'espectroscòpia UV i el DC als estudis estructurals d'àcids nucleics                          | 20        |
| 1.3.1 Espectroscòpia d'ultraviolat-visible  | 20        |
| 1.3.2 Dicroisme circular  | 21        |
| 1.3.3 Anàlisi de les corbes de fusió  | 22        |
| 1.3.4 Extracció de paràmetres termodinàmics   | 23        |
| 1.4 Aplicació de la RMN als estudis estructurals d'àcids nucleics   | 24        |
| 1.4.1 Adquisició i interpretació d'espectres  | 25        |
| 1.4.2 Assignació espectral  | 27        |
| 1.4.3 Informació estructural  | 29        |
| 1.4.4 Càlcul d'estructures en dissolució  | 32        |
| 1.4.5 Anàlisi d'estructures   | 34        |
| <br>  |           |
| <b>2 Introducció i objectius</b>  | <b>35</b> |
| 2.1 El motiu <i>bi-loop</i>   | 35        |
| 2.2 Estructura i paràmetres estructurals  | 37        |
| 2.3 Diferents tipus de tètredes observades  | 40        |
| 2.3.1 Observació d'aquest tipus de tètredes en estructures cristal·lines  | 41        |
| 2.3.2 Comparació amb tètredes de solc major   | 43        |
| 2.4 Característiques espectroscòpiques del motiu  | 44        |
| 2.4.1 Espectroscòpia d'ultraviolat i dicroisme circular   | 44        |
| 2.4.2 Ressonància Magnètica Nuclear   | 46        |
| 2.5 Objectius   | 47        |
| <br>  |           |
| <b>3 El motiu <i>bi-loop</i> heterolíneal</b>   | <b>49</b> |
| 3.1 El <i>bi-loop</i> heterodimèric   | 49        |
| 3.2 Elecció de les seqüències   | 51        |
| 3.3 Síntesi dels oligonucleòtids  | 52        |
| 3.3.1 Síntesi d'oligonucleòtids cíclics   | 52        |
| 3.3.2 Síntesi d'oligonucleòtids lineals   | 56        |
| 3.4 Estudis mitjançant espectroscòpia UV i dicroisme circular   | 57        |
| 3.4.1 Corbes de fusió per espectroscòpia d'UV   | 57        |
| 3.4.2 Espectres de dicroisme circular   | 58        |
| 3.5 Estudis mitjançant Ressonància Magnètica Nuclear  | 61        |
| 3.5.1 Complexos formats per d<pCCTTCGGT> i d(GTCCCTCA), r(GUCCCUCA) o d(CTCAGTCC) ...                           | 61        |
| 3.5.2 Complexos formats per d<pCAGTCCCT> i d(CGGTCCTT), d(GTCCTTCG) o d(T <sub>4</sub> GGTTCTT <sub>4</sub> ) . | 64        |

|          |   |            |
|----------|---|------------|
| <b>4</b> | <b><i>Bi-loop</i> estabilitzat per parells no canònics .....</b>        | <b>69</b>  |
| 4.1      | El parell GT .....  | 69         |
| 4.2      | Síntesi dels oligonucleòtids .....                                      | 70         |
| 4.2.1    | Síntesi de d<pCCTTCGGT> .....   | 71         |
| 4.2.2    | Síntesi de d<pTCGTATGT> .....   | 73         |
| 4.3      | Estudi del cicle d<pCCTTCGGT> mitjançant DC i RMN .....                 | 73         |
| 4.3.1    | Dicroisme circular .....  | 74         |
| 4.3.2    | Ressonància Magnètica Nuclear .....                                     | 75         |
| 4.4      | Estudi del cicle d<pTCGTATGT> mitjançant DC i RMN .....                 | 78         |
| 4.4.1    | Dicroisme circular .....  | 78         |
| 4.4.2    | Ressonància Magnètica Nuclear .....                                     | 79         |
| <b>5</b> | <b>El motiu <i>bi-loop</i> amb oligonucleòtids lineals .....</b>        | <b>83</b>  |
| 5.1      | Elecció de les seqüències .....   | 83         |
| 5.2      | Síntesi dels oligonucleòtids .....                                      | 84         |
| 5.3      | Estudis mitjançant dicroisme circular i espectroscòpia UV .....         | 85         |
| 5.3.1    | Escombrat de condicions d'estructuració mitjançant DC .....             | 85         |
| 5.3.2    | Fusions per dicroisme circular .....                                    | 90         |
| 5.3.3    | Espectres de diferència tèrmica (TDS) .....                             | 92         |
| 5.4      | Estudis mitjançant Ressonància Magnètica Nuclear .....                  | 93         |
| 5.4.1    | Anàlisi dels espectres monodimensionals .....                           | 93         |
| 5.4.2    | Assignació espectral i informació estructural de d(TCGTTCGT) (L1) ..... | 96         |
| 5.4.3    | Assignació espectral de d(TGCTTCGT) (L3) i d(TCGTTGCT) (L5) .....       | 100        |
| 5.4.4    | Càlcul de les estructures de d(TGCTTCGT) i d(TCGTTGCT) .....            | 102        |
| 5.4.5    | Anàlisi de les estructures de d(TGCTTCGT) i d(TCGTTGCT) .....           | 106        |
| 5.4.6    | Comparació amb estructures anteriors .....                              | 113        |
| 5.5      | Estudi termodinàmic de d<pCGCTCCGT>, d(TGCTTCGT) i d(TCGTTGCT) .....    | 115        |
| 5.5.1    | Estudi mitjançant corbes de fusió per dicroisme circular .....          | 115        |
| 5.5.2    | Estudi mitjançant calorimetria d'escombratge diferencial (DSC) .....    | 120        |
| <b>6</b> | <b>Generalitat i rellevància del motiu <i>bi-loop</i> .....</b>         | <b>123</b> |
| 6.1      | Resum .....   | 123        |
| 6.2      | Un motiu de caràcter general .....                                      | 124        |
| 6.3      | La seva estabilitat .....   | 126        |
| 6.4      | Rellevància biològica .....   | 127        |
| 6.4.1    | Interacció entre <i>loops</i> .....                                     | 128        |
| 6.4.2    | Implicació en la recombinació homòloga .....                            | 129        |
| 6.4.3    | Implicació en la regulació de la transcripció .....                     | 130        |
| 6.5      | Perspectives de futur .....   | 133        |

# 1 FONAMENTS

En aquest capítol es farà un repàs d'estructura d'àcids nucleics, de síntesi d'oligonucleòtids i de les diferents tècniques emprades per al seu estudi en aquesta tesi. El lector introduït en el tema pot anar directament a la introducció del primer bloc.

## 1.1 Estructura d'àcids nucleics<sup>62</sup>

El gran nombre d'estructures descrites per als àcids nucleics i l'alt polimorfisme que aquestes presenten va fer necessari la definició d'una nomenclatura i de paràmetres per a la descripció i comparació de les estructures. Aquests s'exposaran a continuació.

### 1.1.1 Components, nomenclatura i numeració<sup>63</sup>

Els àcids nucleics estan formats per unitats repetitives anomenades nucleòtids. Cada nucleòtid consta d'un anell de  $\beta$ -D-ribose (en l'àcid ribonucleic o RNA) o  $\beta$ -D-2'-desoxiribosa (en l'àcid desoxiribonucleic o DNA), un fosfat i una base heterocíclica, base nitrogenada plana i aromàtica. Aquestes bases es classifiquen en: purines (Pur, R), adenina (Ade, A) i guanina (Gua, G), i pirimidines (Pyr, Y), citosina (Cyt, C) i timina (Thy, T, en el DNA) o uracil (Ura, U, en l'RNA). La numeració de les bases es mostra en la Figura 1.1; només es numeren els àtoms de l'anell aromàtic i els substituents adopten el mateix número que l'àtom al qual estan units. La unió de la base al sucre té lloc per l'N1 en les pirimidines i per l'N9 en les purines i s'anomena enllaç glicosídic. El conjunt de base i sucre s'anomena nucleòsid i té un nom diferent segons la base: adenosina, guanosina, citidina i timidina (o uridina). L'anell de ribosa es numera emprant primes. Quan hi ha dos hidrògens metilènics es numeren amb primes i segones: H2' o H5' (*pro-S*) i H2'' o H5'' (*pro-R*).

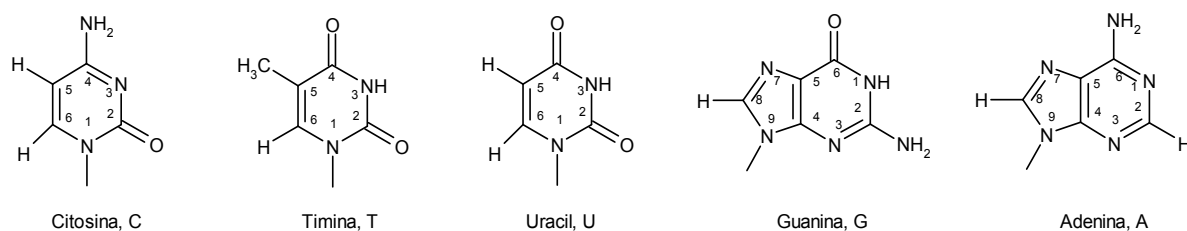


Figura 1.1. Nomenclatura, estructura i numeració de les bases nitrogenades

Les seqüències s'escriuen sempre de l'extrem 5' al 3' i normalment s'ometen els fosfats internucleosídics, però sí que s'indiquen mitjançant una "p" si es presenten en els extrems. Les seqüències de desoxiribonucleòtids s'indiquen amb una "d" i les de ribonucleòtids amb una "r". En aquest treball, per indicar que un oligonucleòtid és cíclic s'empren < > en lloc de ( ). Per tant, la seqüència d'un oligonucleòtid cíclic l'escriuim: d<pNN<sub>n</sub>N>.

### 1.1.2 Conformació

La flexibilitat conformacional dels àcids nucleics ve donada per l'anell de furanosa, la rotació de l'enllaç glicosídic, l'enllaç exocíclic C4'-C5' i els enllaços del grup fosfat. Existeixen 7 angles diedres per nucleòtid, a part dels endocíclics, que s'han d'especificar per caracteritzar-ne la conformació (Taula 1.1 i Figura 1.2). Aquests angles no són totalment independents entre sí, sinó que presenten valors correlacionats i restringits, de manera que els canvis conformacionals en polinucleòtids helicoïdals segueixen moviments concertats.

| Angle de torsió | Àtoms involucrats  | Angle de torsió  | Àtoms involucrats           |
|-----------------|--------------------|------------------|-----------------------------|
| $\alpha$        | (n-1)O3'-P-O5'-C5' | $\chi$ (Pyr/Pur) | O4'-C1'-N1-C2/O4'-C1'-N9-C4 |
| $\beta$         | P-O5'-C5'-C4'      | $\nu_0$          | C4'-O4'-C1'-C2'             |
| $\gamma$        | O5'-C5'-C4'-C3'    | $\nu_1$          | O4'-C1'-C2'-C3'             |
| $\delta$        | C5'-C4'-C3'-O3'    | $\nu_2$          | C1'-C2'-C3'-C4'             |
| $\epsilon$      | C4'-C3'-O3'-P      | $\nu_3$          | C2'-C3'-C4'-P4'             |
| $\zeta$         | C3'-O3'-P-O5'(n+1) | $\nu_4$          | C3'-C4'-O4'-C1'             |

Taula 1.1. Definició dels angles de torsió de l'esquelet sucre-fosfat, angle glicosídic i angles endocíclics

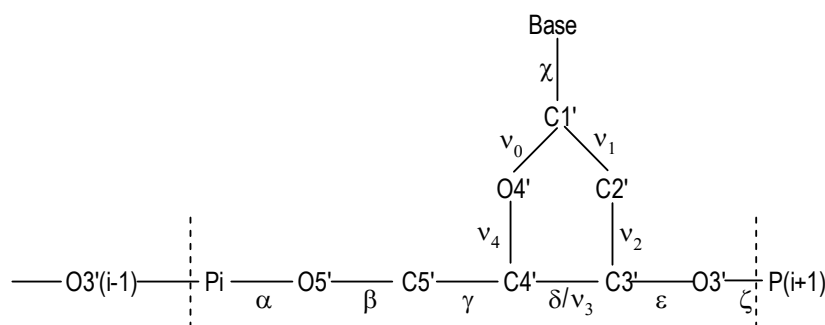


Figura 1.2. Angles de torsió de l'esquelet sucre-fosfat ( $\alpha, \beta, \gamma, \delta, \epsilon$  i  $\zeta$ ), angle glicosídic  $\chi$  i angles endocíclics  $\nu_0, \nu_1, \nu_2, \nu_3$  i  $\nu_4$ .

#### 1.1.2.1 Conformació de l'anell de ribosa

La flexibilitat prové en gran mesura de les diferents conformacions que pot adoptar l'anell de furanosa, que és determinant en la conformació global que adopta l'esquelet de DNA. Aquest anell no és pla, sinó que presenta dues conformacions majoritàries: en forma de sobre o *envelope* (E), amb un àtom fora del pla, i en forma de cadira o *twist* (T), amb 3 àtoms en un pla i els altres dos situats a bandes oposades d'aquest. Els àtoms fora del pla es denominen *endo* o *exo*, segons si estan situats a la mateixa banda que l'àtom C5' o a l'oposada (Figura 1.4), i es descriuen amb superíndex o subíndex, respectivament.

La conformació o *puckering* de l'anell queda definida per les seves cinc torsions endocícliques ( $v_0, v_1, v_2, v_3, v_4$ ), però aquestes torsions no són independents entre si i la conformació es pot descriure amb només dos paràmetres introduint el concepte de pseudorotació: angle de fase de pseudorotació ( $P$ ), que indica el tipus de *puckering*, i amplitud de *puckering* ( $v_m$ ) que n'indica la magnitud màxima<sup>64</sup>. Aquests paràmetres es calculen de la següent manera:

$$\tan P = \frac{(v_4 + v_1) - (v_3 + v_0)}{2 v_2 (\sin 36^\circ + \sin 72^\circ)} \quad v_m = \frac{v_2}{\cos P}$$

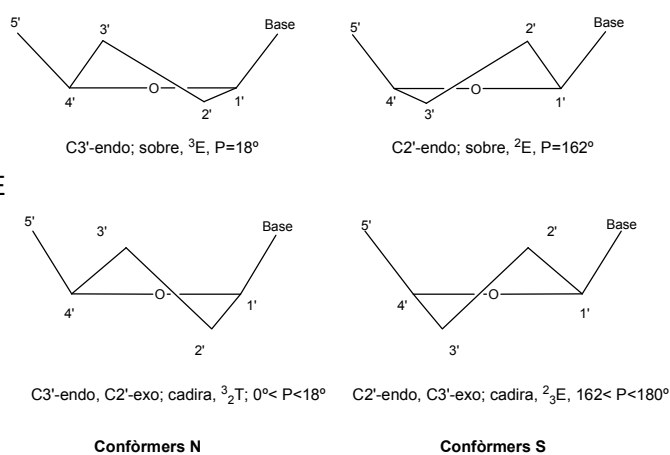
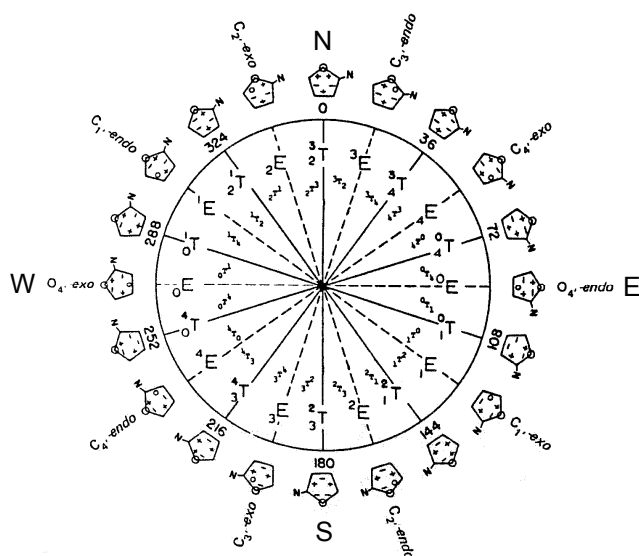


Figura 1.3. Cicle de pseudorotació de l'anell de ribosa, on es mostra la relació entre la fase de l'angle de pseudorotació ( $P$ ), les formes sobre ( $E$ ) i twist ( $T$ ) i les notacions endo i exo.

Figura 1.4. Possibles conformacions de l'anell de ribosa

L'angle de pseudorotació ( $P$ ) pot variar entre  $0^\circ$  i  $360^\circ$  (Figura 1.3), però experimentalment es demostra que la majoria de nucleòtids presenten una de les dues conformacions preferides: la C3'-endo o nord ( $N$ ) ( $0^\circ < P < 36^\circ$ ) i la C2'-endo o sud ( $S$ ) ( $144^\circ < P < 190^\circ$ ). Aquestes conformacions preferides es poden interconvertir en solució a través de les zones  $E$  i  $W$ , que corresponen a conformacions menys estables. Segons el tipus de base o anell de ribosa, trobem una preferència per una conformació o una altra: les purines prefereixen la conformació C2'-endo, mentre que les pirimidines prefereixen la C3'-endo. I els desoxiribonucleòtids presenten preferentment la conformació C2'-endo (B-DNA), mentre que els ribonucleòtids afavoreixen la C3'-endo (A-DNA).

#### 1.1.2.2 Rotació respecte l'enllaç glicosídic ( $\chi$ )

La torsió de l'enllaç glicosídic ( $\chi$ ) està definida pels àtoms O4'-C1'-N9-C4 per a les purines i O4'-C1'-N1-C2 per a les pirimidines. Les bases nitrogenades poden adoptar dues orientacions principals en relació a la posició de l'anell de ribosa: *syn* ( $\chi = 0 \pm 90^\circ$ ) i *anti* ( $\chi = 180 \pm 90^\circ$ ) (Figura 1.5). En la conformació *anti*, la base (l'anell de pirimidina de les purines i l'O2 de les pirimidines) es troba en la banda oposada a l'anell de ribosa i en la conformació *syn* la base es troba sobre l'anell de ribosa (Figura 1.6). Normalment la conformació més estable es la *anti*, tot i que en certs motius estructurals també s'observa la conformació *syn*. En certes condicions en dissolució, les formes *syn* i *anti* es poden interconvertir ràpidament ja que la barrera energètica és petita.



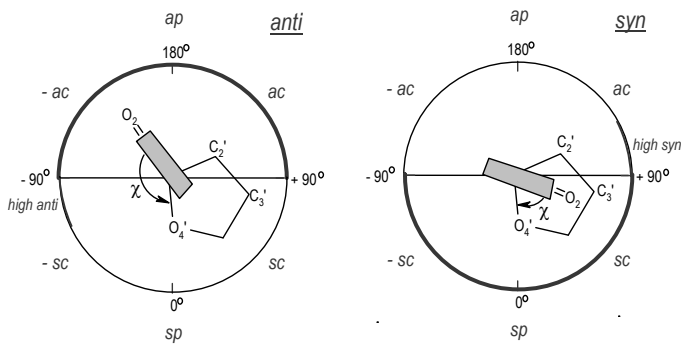


Figura 1.5. Disposició *syn* i *anti* de la base respecte al sucre segons IUPAC

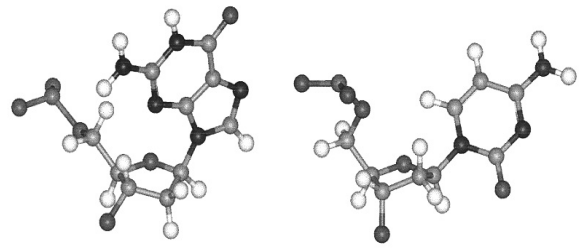


Figura 1.6. Conformació *syn* de la desoxiguanosina i *anti* de la desoxicitidina

### 1.1.3 Interaccions entre bases

De totes les interaccions que afecten l'estructura dels àcids nucleics, l'associació entre bases és la de major importància, tant funcionalment com estructural<sup>65</sup>. Les interaccions entre bases són fonamentalment de dos tipus: interaccions per pont d'hidrogen en el pla de les bases i interaccions d'apilament o *stacking* perpendiculars al pla de les bases.

#### 1.1.3.1 Interaccions per enllaç d'hidrogen

Els enllaços o ponts d'hidrogen són interaccions principalment de caràcter electrostàtic entre un àtom electronegatiu amb parells d'electrons no enllaçants (acceptor) i un altre àtom electronegatiu que té un hidrogen unit (donador). El sistema dels àtoms X-H...Y acostuma a ser lineal, amb una desviació no més gran de 35°, una distància entre els heteroàtoms d'entre 2.5 i 3.4 Å i una estabilitat d'entre 6 i 10 kcal/mol. En els àcids nucleics aquests ponts d'hidrogen són del tipus: N-H...N i N-H...O, on N-H pot correspondre tant a imines com a amines.

Existeixen molts patrons de ponts d'hidrogen possibles<sup>66</sup>, sobretot si es consideren les propietats àcid-base de les bases i les seves diferents formes tautòmeres. En la Figura 1.6 es mostren alguns dels aparellaments més habituals. En el DNA biològic les bases s'aparellen fonamentalment seguint els patrons proposats per Watson i Crick<sup>2</sup>. La característica més important d'aquest aparellament és que els parells A·T i els G·C són isomorfs (la distància entre les cadenes i l'angle d'inclinació és pràcticament el mateix) i donen lloc a dobles hèlixs regulars, que presenten un solc major i un solc menor ben definits. Els parells WC no saturen totes les possibilitats de reconeixement, sinó que deixen posicions susceptibles de formar ponts d'hidrogen addicionals tant en el solc major com en el menor. El parell G·C és més estable (-27.5 kcal/mol) que el parell A·T (-15.0 kcal/mol)<sup>67</sup> degut a la formació de tres enllaços d'hidrogen enfront de dos.

Un altre tipus d'aparellament important és el Hoogsteen<sup>68</sup>, no isomorf amb el WC i que es pot formar entre parells A·T a pH neutre i a pHs àcids entre parells G·C<sup>+</sup>. També s'han observat parells invertits (*reversed*), on una de les bases està rotada 180° i aparellaments tipus *wobble* ("trontollant"), on una de les bases està desplaçada, i que podrien ser l'explicació de la degeneració del codi genètic. A més a més, les bases es poden aparellar de forma incorrecta (*mismatches*): G·A, G·T, A·A, G·G, T·T, C·C, T·C o A·C.

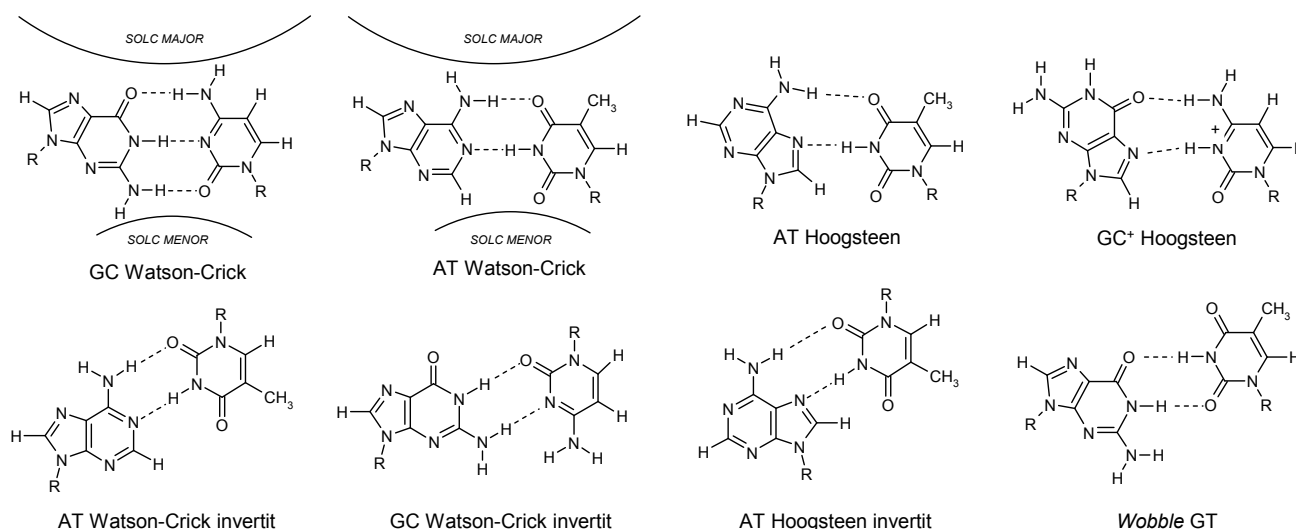


Figura 1.7. Alguns dels possibles aparellaments de les bases més habituals

### 1.1.3.2 Interacció per apilament

Les bases també interaccionen entre elles per forces d'apilament o *stacking*, relacionades amb el solapament dels orbitals  $\pi$  dels anells aromàtics. L'apilament  $\pi$ - $\pi$  ocorre majoritàriament entre bases de la mateixa cadena de la hèlix, però degut al gir també hi ha contribució d'apilament entre bases de cadenes oposades. Aquesta interacció és un fenomen complex en què estan implicats: a) la força de Van der Waals, responsable de l'especificitat de l'apilament, b) la interacció electrostàtica, que origina una disposició de bases apilades que maximitza la interacció favorable i minimitza la desfavorable entre la distribució de càrregues de les bases en contacte i c) en medis aquosos, l'efecte hidrofòbic, és a dir, la contribució entròpica favorable d'excloure l'aigua. Experimentalment s'ha pogut determinar que les purines s'apilen més fortament que les pirimidines, indicatiu que la superfície d'apilament és un factor important. Tot i ser difícil determinar l'estabilitat de l'apilament de forma independent a les contribucions de pont d'hidrogen o les restriccions de l'esquelet de sucre-fosfat, recentment s'ha calculat<sup>69</sup> que aquestes interaccions són de l'ordre de les interaccions de pont d'hidrogen (entre 13.5 i 18.2 kcal/mol) i que són claus per definir els paràmetres de gir (*twist*) i ascensió (*rise*) del DNA (definició dels paràmetres a la Figura 1.8).

### 1.1.4 Paràmetres helicoïdals per a àcids nucleics

Les bases són planes, però els parells de bases poden presentar certa flexibilitat. Els paràmetres helicoïdals<sup>70</sup> defineixen les posicions i geometries de les bases d'un parell, dels parells de bases entre si i respecte a un eix global o local, i resulten útils a l'hora de descriure una estructura. La definició d'aquests paràmetres es mostra gràficament a la Figura 1.8:

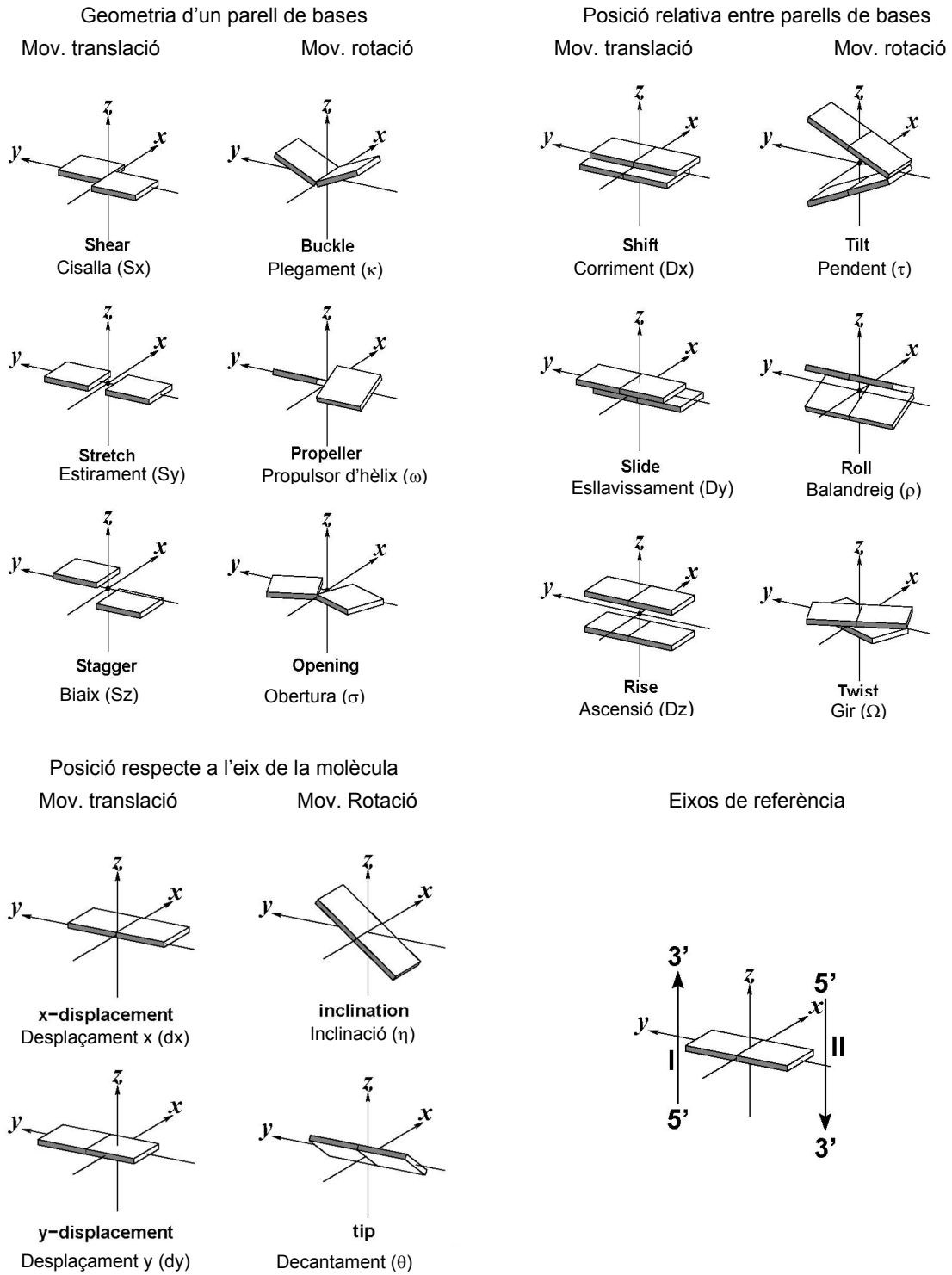


Figura 1.8. Paràmetres helicoidals d'àcids nucleics: definició, noms en anglès, noms en català<sup>71</sup> i símbols

## 1.2 Síntesi d'oligonucleòtids<sup>72</sup>

La síntesi d'oligonucleòtids implica la formació d'enllaços fosfodiester entre el grup hidroxil en 3' d'un nucleòsid i el grup hidroxil en 5' d'un altre, i generalment s'empra un nucleòsid derivatitzat en 3' amb una espècie amb fòsfor que s'activa mitjançant un agent de condensació. Inicialment es van emprar derivats fosforilats dels nucleòtids (mètodes del fosfat diester i fosfat triester) amb agents acoblants tipus DCC o MSNT, respectivament, però no eren prou reactius i els acoblaments no eren quantitius. La gran millora va arribar amb el mètode del fosfit triester o fosforamidit<sup>73</sup>, que usa derivats estables de P(III), molt més reactiu que el P(V), que reaccionen ràpidament en presència d'un catalitzador àcid. El tetrazole és el catalitzador estàndard, suficientment àcid per protonar el nitrogen del fosforamidit i activar-lo sense causar la pèrdua del grup dimetoxitritil (DMT) que protegeix l'hidroxil en 5'. Després de l'acoblament es forma un enllaç fosfit triester que és relativament làbil, i per això cal oxidar-lo en cada cicle a la forma estable de fosfat triester. Abans de l'oxidació, s'incorpora una etapa d'acetilació dels grups hidroxil que no han reaccionat. Aquesta etapa també serveix per revertir adductes que s'hagin pogut formar per condensació de fosforamidit amb les formes enòliques de les bases, adductes molt làbils a la hidròlisi però que si fossin oxidats es tornarien estables i donarien lloc a oligonucleòtids modificats.

### 1.2.1 Síntesi d'oligonucleòtids lineals

Tots els oligonucleòtids d'aquesta tesi s'han sintetitzat seguint la metodologia del fosfit triester o fosforamidit emprant 3'-(2-cianoetil)-*N,N*-diisopropil-fosforamidits de 5'-O-(4,4'-dimetoxitritil)-desoxiribonucleòsids comercials. Les etapes de síntesi són les que es mostren a la Figura 1.9:

- Eliminació del protector 4,4'-dimetoxitritil del grup hidroxil de l'extrem 5'-terminal de la cadena oligonucleotídica en creixement amb àcid tricloroacètic.
- Acoblament d'un nucleòtid en forma de fosforamidit sobre l'hidroxil lliure en 5' amb tetrazole com a catalitzador àcid.
- Acetilació dels grups hidroxil lliures que no han reaccionat durant l'etapa d'acoblament (bloqueig o *capping*).
- Oxidació de l'enllaç fosfit triester a fosfat triester amb <sup>t</sup>BuOOH o I<sub>2</sub>.

Aquesta metodologia s'utilitza en fase sòlida sobre boles de vidre de porus controlat (CPG) com a suport ja que és un polímer inert, rígid i no solvatable, difon bé els reactius i es pot emprar en el sintetitzador automàtic. El suport està derivatitzat amb un llarg espaiador (*long chain alkyl amino, /caa*) i l'hidroxil del nucleòsid 3'-terminal s'hi uneix a través d'un grup succinil. Les resines amb el primer nucleòsid ancorat són assequibles comercialment.

El rendiment de cada etapa d'acoblament es pot determinar quantificant els cations DMT alliberats en la desprotecció del grup hidroxil. Amb reactius estàndard, el rendiment d'acoblament per etapa és superior al 98%.

Les etapes de desancoratge, desprotecció dels fosfats i desprotecció de les bases (A<sup>Bz</sup>, C<sup>Bz</sup>, G<sup>iBu</sup>, T) es realitzen totes en un últim tractament final amb NH<sub>3</sub> aquós concentrat a 55 °C durant 12h.

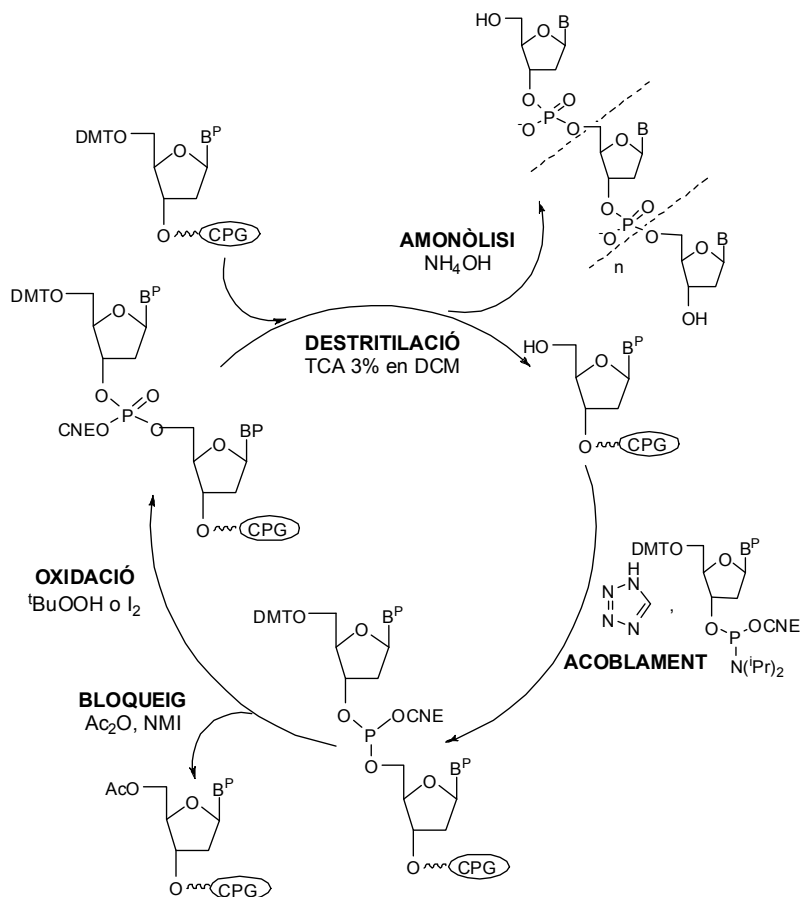


Figura 1.9. Esquema de síntesi d'oligonucleòtids lineals

### 1.2.2 Síntesi d'oligonucleòtids cíclics

La síntesi d'oligonucleòtids cíclics s'ha realitzat en fase sòlida emprant la metodologia en fase sòlida desenvolupada en el nostre grup de treball<sup>57</sup>. L'esquema de síntesi es representa a la Figura 1.10.

Les principals característiques d'aquesta metodologia són:

- Permet la síntesi de qualsevol oligonucleòtid cíclic sense limitació de seqüència ja que la unió a la fase sòlida es pot realitzar a través de qualsevol dels nucleòtids mitjançant la preparació de la nucleotidil-resina corresponent.
- Permet la síntesi d'oligonucleòtids cíclics entre 2 i 30 mer amb rendiments entre bons i moderats i amb crus d'elevada puresa.
- L'elongació de la cadena es realitza de 3' a 5', fent ús de la metodologia estàndard del fosfit triester en fase sòlida, fet que permet treballar amb els esquemes de protecció habituals, reactius assequibles comercialment i sintetitzador automàtic.
- L'etapa clau de ciclació té lloc sobre el precursor lineal unit a la fase sòlida, en unes condicions d'alta dilució que minimitzen les reaccions intercadena. Aquesta etapa es realitza directament sobre el producte lineal obtingut sense necessitat de realitzar cap transvasament de la resina ja que es pot adaptar per realitzar-la al sintetitzador automàtic.

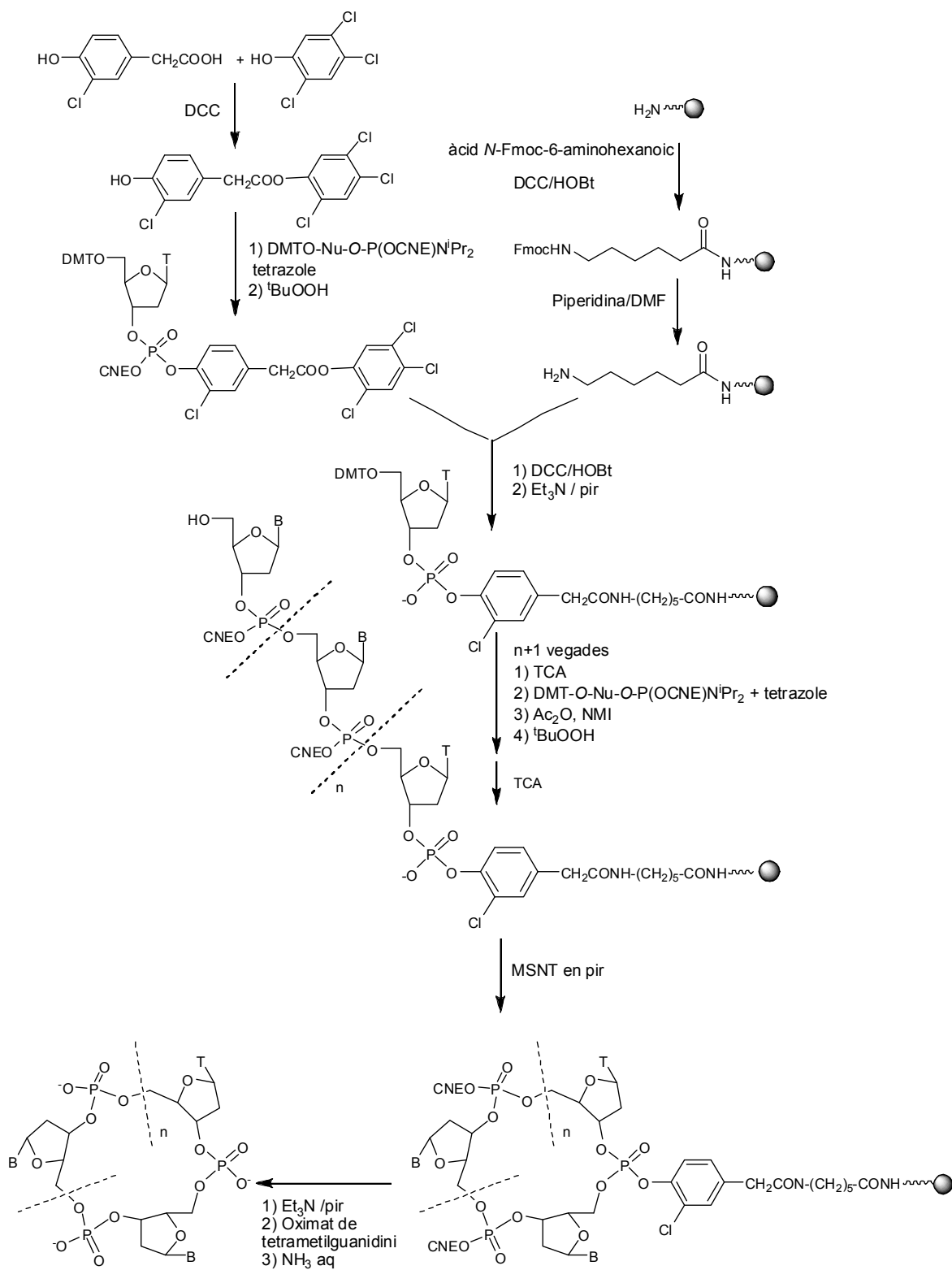


Figura 1.10. Esquema de síntesi d'oligonucleòtids cíclics desenvolupat al nostre grup

- Usa un esquema de protecció adequat que permet tenir desprotegit només el fosfat en 3' perquè la ciclació sigui regioselectiva.
- Es desancoren selectivament els oligonucleòtids cíclics ja que només es desprotegeix el fosfat triester de la unió del cicle a la resina; la resta de grups fosfat s'han desprotegit anteriorment i es troben en forma de fosfat diester.
- Permet la síntesi d'oligonucleòtids tant a petita com a gran escala; interessant per a aplicacions en què es necessiten grans quantitats d'oligonucleòtid, com la RMN.

### **1.3 Aplicació de l'espectroscòpia UV i el dicroisme circular als estudis estructurals d'àcids nucleics**

L'espectroscòpia d'ultraviolat-visible (UV-Vis) i el dicroisme circular (DC) són tècniques espectroscòpiques electròniques que si bé no permeten conèixer l'estructura tridimensional d'un oligonucleòtid, sí que ens poden donar informació sobre l'estructura secundària: existència d'estructuració, tipus i estabilitat.

#### **1.3.1 Espectroscòpia d'ultraviolat-visible<sup>74</sup>**

En els àcids nucleics els cromòfors per excel·lència són les bases nitrogenades, ja que els grups fosfat i els anells de ribosa no presenten absorbància significativa a longituds d'ona majors de 220 nm. Les transicions electròniques en les nucleobases que donen lloc a les bandes d'absorció més intenses són les  $\pi \rightarrow \pi^*$ , que presenten un màxim al voltant de 260 nm amb coeficients d'extinció molar de l'ordre de  $10^4$  ( $M^{-1} \cdot cm^{-1}$ ).

Els espectres d'absorció de nucleòsids i nucleòtids són bàsicament els mateixos que els de les seves bases nitrogenades, però en les estructures ordenades d'oligo- i polinucleòtids les nucleobases s'apilen cara a cara i interaccionen a través dels electrons  $\pi$  produint una disminució de la intensitat de l'absorció en l'espectre d'ultraviolat, coneguda com hipocromicitat<sup>75</sup>, que depèn de l'estructura tridimensional de l'oligonucleòtid. En funció de l'estructura i de les interaccions entre les bases, la hipocromicitat pot variar entre un 1 i un 11% en el cas de desoxiribonucleòtids i arribar fins a un 30% en la majoria de polinucleòtids helicoïdals.

Una de les aplicacions més importants de l'espectroscòpia d'ultraviolat, és el seu ús per a la quantificació d'oligonucleòtids (veure Materials i Mètodes):  $OD_{260} = \epsilon(\mu mol^{-1} \cdot ml \cdot cm^{-1}) \cdot n(\mu mol)$ .

També s'utilitza àmpliament per al control de la desnaturalització o "fusió" dels àcids nucleics ja que en augmentar gradualment la temperatura de la mostra es produeix una desestructuració i, per tant, un desapilament de les bases, fet que es tradueix en un augment de l'absorbància. La representació de la variació de l'absorbància amb la temperatura constitueix el que es coneix com a corba de fusió. Aquest procés de desnaturalització és un fenomen cooperatiu que dona lloc a una corba de tipus sigmoide amb un pendent més pronunciat com més llarg és el fragment d'oligonucleòtid aparellat. La monitorització de la variació de l'absorbància amb la temperatura es realitza tant començant amb els oligonucleòtids estructurats a baixa temperatura i anant augmentant-la (corba de fusió o *melting*) com

al revés, començant a alta temperatura i anant-la disminuint poc a poc per a permetre la renaturalització (corba d'hibridació o *annealing*). Si ens trobem en una situació d'equilibri, aquestes dues corbes són iguals ja que es tracta d'un procés reversible.

El punt mig en la variació de l'absorbància es coneix com temperatura de fusió ( $T_m$ ) i representa la temperatura en què el 50 % de la mostra està estructurada i l'altre 50 % desestructurada i per això és un bon mètode per a la mesura de l'estabilitat tèrmica. La  $T_m$  depèn de la seqüència, de la concentració d'oligonucleòtid i de sals i fins i tot del tipus de contraions presents.

Recentment, s'ha descrit l'ús de l'espectroscòpia d'ultraviolat per a obtenir un perfil característic de cada tipus d'estructura: els espectres de diferència tèrmica (TDS: *thermal difference spectra*)<sup>76</sup>. Aquest mètode consisteix a enregistrar l'espectre d'absorbància a alta temperatura, per sobre de la  $T_m$ , on tenim l'espècie desestructurada, i l'espectre a baixa temperatura, per sota de la  $T_m$ , assegurant que tenim l'espècie estructurada, i restar els dos espectres. L'espectre diferència que se n'obté té un perfil únic per cada tipus d'estructura de DNA ja que reflexa les subtilitats de les interaccions entre les bases.

### 1.3.2 Dicroisme circular<sup>77</sup>

La tècnica del dicroisme circular consisteix en fer interaccionar un molècula òpticament activa amb llum polaritzada linealment i mesurar la diferència en l'absorció de les components polaritzades circularment a l'esquerra i a la dreta a diferents longituds d'ona, ja que el medi òpticament actiu té coeficients d'extinció molar diferents per a la llum polaritzada a l'esquerra i a la dreta:  $\Delta A = A_L - A_R$ . En sortir del medi, les dues components estan desfasades i a més a més tenen diferent amplitud, de forma que la llum polaritzada linealment s'ha convertit en llum polaritzada el·lípticament. Experimentalment es mesura  $\Delta A$ , tot i que per raons històriques es representa l'angle d'el·lipticitat  $\theta$  (en graus):  $\theta = \Delta A \cdot 32.98$ . També es comú representar l'el·lipticitat molar mitjana  $[\Theta]_M$  per compensar els efectes de la concentració (en  $\text{deg} \cdot \text{cm}^2 / \text{dmol}$ ):

$$[\Theta]_M = \frac{\theta_\lambda}{10 \cdot \ell \cdot c} \quad \text{on } \theta_\lambda \text{ és l'el·lipticitat observada en miligrans (mdeg), } \ell \text{ és la longitud del camí recorregut en cm i } c \text{ és la concentració molar (mol/dm}^3\text{)}.$$

En aquesta tesi es representarà l'el·lipticitat molar mitjana per poder comparar espectres a diferents concentracions i entre diferents oligonucleòtids. Per a la representació de corbes de fusió s'emprarà l'el·lipticitat normalitzada, també a efectes comparatius.

Les bases nitrogenades són els cromòfors dels àcids nucleics, però presenten un pla de simetria i, per tant, no són òpticament actives en si mateixes. És la seva unió a un anell de ribosa quiral el que indueix dicroisme circular en les seves bandes d'absorció, encara que de baixa intensitat. Quan tenim un oligonucleòtid amb una estructura secundària helicoidal que manté les bases apilades, aquestes poden interaccionar i donar lloc a espectres de dicroisme circular de gran intensitat (amb bandes molt intenses al voltant de 180 nm i de menor intensitat entre 220 i 300 nm). Com que l'espectre de DC depèn de les interaccions entre les bases, es tracta d'una tècnica extremadament sensible als canvis en l'estructura i per això fa més de 40 anys que s'empra per a l'estudi de les conformacions dels àcids



nucleics en solució<sup>78</sup>. És de destacar, per exemple, el fet que el DC va ser la tècnica amb la qual es va identificar la forma Z del DNA per primera vegada<sup>10</sup> i que actualment s'empra extensament per a avaluar la interacció del DNA amb altres molècules. És una tècnica molt útil perquè és extremadament sensible, requereix molt poca quantitat de mostra, és senzilla experimentalment i no és destructiva.

La informació sobre l'estructura no s'obté a partir de la interpretació directa de les bandes sinó per comparació amb espectres de referència d'estructures conegudes. Així, existeixen patrons característics de les diferents estructures del DNA (veure com a exemple els perfils dels espectres de DC de les formes A, B i Z del DNA, Figura 1.11) que, per analogia, permeten conèixer els trets estructurals bàsics d'una seqüència donat el seu espectre de DC.

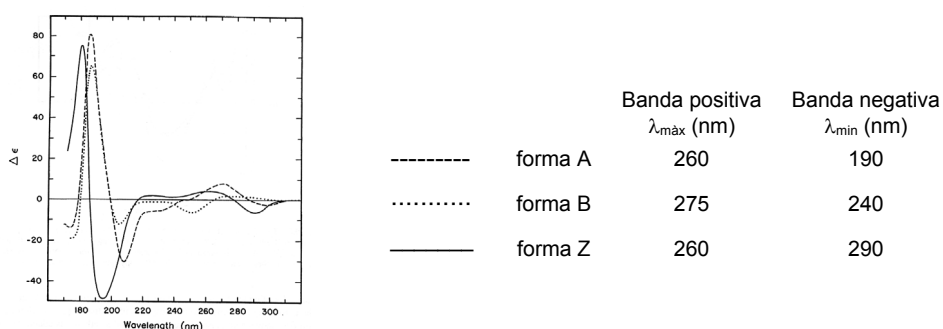


Figura 1.11. Comparació dels espectres de dicromisme circular de les formes A, B i Z del DNA

El dicromisme circular és útil doncs, quan es disposa de patrons per comparar els espectres, però també s'empra extensament per al seguiment de la fusió dels àcids nucleics. En aquest cas el que succeeix és que en augmentar la temperatura l'oligonucleòtid es desestructura i el senyal de dicromisme circular a una determinada longitud d'ona disminueix (si es "monitoritza" la variació d'un màxim) o augmenta (si es "monitoritza" la variació d'un mínim).

Els espectres de dicromisme circular dels àcids nucleics també es poden analitzar d'una forma quantitativa teòrica<sup>79</sup> emprant càlculs de mecànica quàntica o semiempírics, però aquest tractament excedeix els objectius de la present tesi.

### 1.3.3 Anàlisi de les corbes de fusió<sup>80</sup>

Les corbes de fusió es poden obtenir representant la variació de qualsevol propietat enfront de la temperatura: absorbància, dicromisme circular, fluorescència, desplaçament químic d'RMN o altres. El paràmetre directe que se'n pot extreure és la temperatura de fusió,  $T_m$ , que correspon a la temperatura on tenim la meitat de la mostra estructurada. Determinar-la, doncs, implica conèixer la variació de l'absorbància amb la temperatura tant de la forma estructurada com de la desestructurada. Així, primer de tot cal traçar les línies de variació de la propietat d'ambdues espècies (rectes que segueixen els plafons a alta i baixa temperatura) i després calcular-ne la mitjana. La  $T_m$  correspondrà a la temperatura on es creuen la mitjana i la nostra corba de fusió (veure Figura 1.12). El gran inconvenient d'aquest mètode és la subjectivitat a l'hora de triar les línies de base de les espècies estructurada i desestructurada, per això sovint es calcula la  $T_m$  a partir del punt d'inflexió de

la corba sigmoide (extrem de la primera derivada), però aquesta temperatura només correspon realment a la de fusió pel cas d'equilibris unimoleculars. Pel cas d'equilibris bimoleculars, com els estudiats en aquesta tesi, pot existir una variació d'entre 1 i 2 °C. El mètode de la mediana és més subjectiu, però el resultat és més exacte. Tot i així, requereix de plafons a baixa i alta temperatura per poder determinar d'una manera fiable les línies de base.

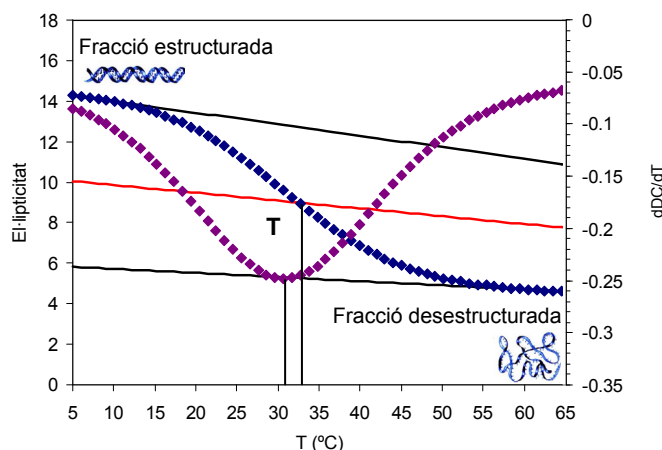


Figura 1.12. Corba de fusió per DC amb la determinació de la  $T_m$  segons el mètode de la mitjana i el mètode de la derivada (♦ corba de fusió, ◆ derivada de la corba)

L'anàlisi més detallada d'una o varies corbes de fusió ens permet obtenir els paràmetres termodinàmics del procés en estudi, com s'explica en el següent apartat.

### 1.3.4 Extracció de paràmetres termodinàmics<sup>81</sup>

Els paràmetres termodinàmics extrets a partir de corbes de fusió depenen d'un model, ja que es basen en l'assumpció que tenim un equilibri de dos estats entre la forma totalment estructurada i la totalment desestructurada (*helix-to-coil theory*), per contraposició amb els paràmetres que es poden obtenir d'estudis calorimètrics. Per això, és recomanable calcular-los amb les dues metodologies que s'expliquen a continuació per comparar-ne els resultats<sup>80</sup>.

#### 1.3.4.1 Anàlisi d'una corba de fusió per extreure'n paràmetres termodinàmics<sup>82</sup>

Assumint que tenim un equilibri de dos estats (model acceptable per oligonucleòtids curts, de menys de 14 parells de bases) i que coneixem la molecularitat de la reacció, només amb l'anàlisi d'una única corba de fusió ja podem obtenir paràmetres termodinàmics per un equilibri. A partir dels valors d'absorbància o el líplicitat a cada temperatura i coneixent la variació de la propietat de l'espècie estructurada i la desestructurada amb la temperatura, podem calcular la fracció molar d'espècie estructurada a cada T:

$$x(t) = \frac{A(t) - A_{des}(t)}{A_{es}(t) - A_{des}(t)} \quad \text{perquè} \quad A(t) = x \cdot A_{es}(t) + (1 - x) \cdot A_{des}$$

La representació de la fracció molar respecte la T ens dona una funció sigmoide. A partir de la fracció molar podem calcular el valor de la constant d'equilibri a cada temperatura amb la fórmula (per un equilibri bimolecular autoassociatiu):

$$K_a = \frac{x}{2c_0(1-x)^2}$$

on  $c_0$  és la concentració total d'oligonucleòtid en mol/l.

Si representem  $\ln K$  enfront de  $1/T$  (*Van't Hoff plot*) podem obtenir directament els paràmetres termodinàmics. Si  $\Delta H^0$  i  $\Delta S^0$  són independents de la temperatura, és a dir  $\Delta C_p^0=0$ , aquesta

representació ha de donar una recta.

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 = -RT \ln K \quad \boxed{\ln K = -\frac{\Delta H^0}{R} \cdot \frac{1}{T} + \frac{\Delta S^0}{R}}$$

Generalment cal restringir l'anàlisi al rang de temperatura pel qual  $0.03 < x < 0.97$ , però fins i tot es recomana<sup>83</sup> restringir-lo encara més al rang de  $0.15 < x < 0.85$ , on els valors de  $K_a$  són més precisos. Si la representació de  $\ln K$  enfront de  $1/T$  no dona una línia recta, pot ser degut a diferents causes: 1)  $\Delta H^0$  no és independent de la temperatura ( $\Delta C_p^0 \neq 0$ ); 2) la transició no és de dos estats; 3) a l'existència d'errors en escollir les línies de base. Però aquesta no linealitat sovint és difícil d'apreciar a causa dels propis errors experimentals. A més a més, cal dir que l'elecció de les línies de base pot afectar substancialment el perfil d'aquesta gràfica<sup>80</sup>. Per obtenir valors termodinàmics amb confiança, cal repetir aquesta anàlisi amb dades de corbes a diferents concentracions.

#### 1.3.4.2 Dependència de la $T_m$ amb la concentració<sup>84</sup>

Quan tenim equilibris de molecularitat major que u, podem caracteritzar-los de forma més fiable a partir de la dependència de la temperatura de fusió amb la concentració. Per un equilibri bimolecular autoassociatiu, quan  $T=T_m$ :  $x=0.5$  i la constant d'equilibri és igual a la inversa de la concentració total.

Aleshores, podem escriure:  $\Delta H^0 - T_m \Delta S^0 = -RT_m \ln \frac{1}{c_0} \quad \Delta H^0 = T_m (\Delta S^0 + R \ln c_0)$

D'aquí podem extreure l'equació que ens permet obtenir els paràmetres representant  $1/T_m$  enfront de la concentració total inicial ( $c_0$ ). Aquesta representació donarà una recta si és cert que tenim un equilibri entre dos estats i  $\Delta C_p^0=0$ .

$$\boxed{\frac{1}{T_m} = \frac{\Delta S^0}{\Delta H^0} + \frac{R}{\Delta H^0} \cdot \ln c_0}$$

## 1.4 Aplicació de La RMN als estudis estructurals d'àcids nucleics<sup>85</sup>

Des dels treballs inicials de Richard Ernst (Premi Nobel de Química 1991), que va proposar i desenvolupar les bases teòriques de la RMN de polsos mono- i multidimensional, aquesta tècnica ha estat àmpliament estudiada i aplicada a nombrosos camps. Kurt Wütrich (Premi Nobel de Química 2002) va establir les bases metodològiques de les aplicacions biofísiques de la tècnica a les proteïnes i els àcids nucleics.

Les dades experimentals extretes dels experiments de ressonància, en conjunció amb tècniques de càlcul, proporcionen estructures amb un grau de detall comparable a l'obtingut mitjançant difracció de Raigs X de cristalls únics. No obstant, la RMN presenta dos avantatges respecte a la cristal·lografia de Raigs X:

- proporciona informació de l'estructura que adopta la molècula en dissolució i, per tant, en un entorn molt més similar al que es troba naturalment,
- proporciona informació dinàmica, ja que és capaç de detectar una gran varietat de fenòmens en molt diverses escales de temps (de ps a ms), permetent caracteritzar processos de canvis

conformacionals, intercanvi químic, reconeixement molecular, etc . Aquesta informació és important en els àcids nucleics ja que solen presentar regions d'elevada flexibilitat, que alhora en dificulten la cristal·lització.

Tot i aquests avantatges, la major limitació continua essent la mida de les molècules que es poden estudiar. En els últims anys, aquest problema s'ha solucionat parcialment i s'han pogut determinar estructures de fragments d'RNA i DNA de major grandària gràcies als avenços tècnics experimentats: l'ús d'imants superconductors d'alt camp (fins a 18.8-21.1 Tesla, que corresponen a una freqüència de ressonància de 800-900 MHz per l'<sup>1</sup>H), la introducció de l'espectroscòpia multidimensional i els avenços en el marcatge isotòpic de <sup>15</sup>N i <sup>13</sup>C de les molècules de DNA<sup>86</sup> i RNA<sup>87</sup>. Tot i així, l'espectroscòpia d'RMN de protó segueix sent l'eina bàsica per establir la conformació de macromolècules en solució. Per a l'elucidació estructural de molècules de DNA de mida petita (<20 mer), el marcatge isotòpic no resulta imprescindible. El segon nucli més important per a l'estudi dels àcids nucleics és el <sup>31</sup>P, gràcies a les seves bones propietats espectroscòpiques (I=1/2, abundància 100%) i a què dóna informació sobre l'esquelet. El <sup>13</sup>C és possible detectar-lo en mostres no marcades amb tècniques de múltiple quantum de detecció inversa, com l'HSQC. En aquesta tesi s'ha treballat en mostres sense marcar i bàsicament amb espectres de protó.

### 1.4.1 Adquisició i interpretació d'espectres

#### 1.4.1.1 Espectres monodimensionals

El primer pas en la realització d'un estudi estructural d'una molècula de DNA és l'obtenció d'un espectre monodimensional. Amb aquest és impossible realitzar l'assignació, però sí que és possible extreure'n alguna informació estructural.

Amb els espectres realitzats en aigua deuterada (amb presaturació<sup>88</sup> del senyal residual d'H<sub>2</sub>O) podem veure si la dispersió de senyals dels protons no bescanviables és bona, fet indicatiu que l'oligonucleòtid presenta una estructuració que proporciona un entorn químic diferent als diferents residus. A partir del nombre de senyals aromàtics podem conèixer quants residus diferents tenim i, per tant, la simetria de l'estructura en estudi.

Per poder observar els protons bescanviables de les bases cal enregistrar espectres en aigua lleugera (H<sub>2</sub>O/D<sub>2</sub>O 9:1). Amb aquests espectres podem obtenir informació de la formació de parells de bases i, per tant, de l'estructura secundària. Els protons imino i amino aparellats ressonen a camps més baixos que els que no ho estan, i a freqüències característiques segons l'aparellament (veure Taula 1.2). A més a més, només el fet d'observar aquests senyals ja ens indica que són protons aparellats o protegits d'alguna manera, ja que sinó intercanviarien ràpidament amb l'aigua i no els podríem observar. Aquests espectres es poden realitzar en aigua lleugera gràcies a què existeixen maneres d'eliminar el senyal del dissolvent: *watergate*<sup>89</sup> (*water suppression by gradient-tailored excitation*), que empra gradients i polsos d'excitació de l'aigua de manera que el senyal d'aquesta no es reenfoca, i que es pot implementar en experiments 2D, o experiments d'excitació selectiva tipus *jump-and-return*<sup>90</sup>, com el p1331<sup>91</sup>, que consisteixen en l'ús de polsos que no exciten la magnetització de l'aigua i que permeten observar protons que intercanvien ràpidament amb aquesta.

Realitzar espectres monodimensionals a diverses concentracions o canviant la temperatura ens pot donar informació de la molecularitat i l'estabilitat de l'estructura.

Donat el gran solapament, l'anàlisi detallada dels espectres resulta complicat i per poder assignar tots els protons no bescanviabls de la molècula cal recórrer a l'espectroscòpia bidimensional.

#### 1.4.1.2 Espectres bidimensionals

L'espectroscòpia bidimensional es basa en la transferència de magnetització entre dos nuclis durant un temps de mescla  $\tau_m$ . La magnetització transferida s'etiqueta durant el temps d'evolució  $t_1$  amb la freqüència del primer nucli i durant el temps de detecció  $t_2$  evoluciona la freqüència del segon. La doble transformada de Fourier de la freqüència directa  $t_2$  i la dimensió indirecta  $t_1$  (que s'obté enregistrant diferents experiments variant  $t_1$ ) dóna lloc a un espectre bidimensional on els pics de correlació relacionen les freqüències dels dos nuclis. La seqüència general de polsos per a realitzar un espectre bidimensional és la que es mostra de forma esquemàtica en la Figura 1.13. Amb l'espectroscòpia bidimensional podem solucionar els problemes de solapament i, a més a més, obtenir informació estructural valuosa.

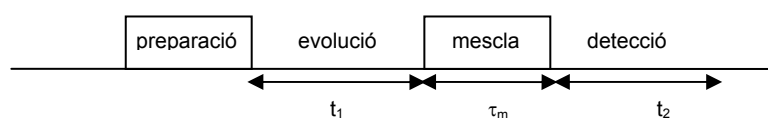


Figura 1.13. Seqüència de polsos general per a un espectre bidimensional

Per a l'assignació i extracció d'informació estructural dels àcids nucleics s'empren experiments bidimensionals estàndards. En la present tesi s'han emprat els següents experiments:

- Experiment COSY (*CORrelated Spectroscopy*)<sup>92</sup>: correlaciona protons acoblats escalarment a dos o tres enllaços, ja que la transferència de magnetització es produeix a través dels electrons dels enllaços. Tots els senyals són en antifase; els de correlació són d'absorció i els de la diagonal en dispersió. Aquesta forma en antifase provoca una cancel·lació parcial del centre del multiplet i no permet determinar de manera correcta la constant d'acoblament J. Existeixen variants de l'experiment COSY dissenyades per a determinar quantitativament aquests valors de les constants d'acoblament (DQF-COSY, E-COSY).

- Experiment TOCSY (*TOTAL Correlation Spectroscopy*)<sup>93</sup> o HOHAHA (*HOmonuclear HARTmann-HAHn*): correlaciona un protó amb tots els del seu sistema de spin. Tots els senyals són d'absorció i en fase.

- Experiment NOESY (*Nuclear Overhauser Effect Spectroscopy*)<sup>94</sup>: es basa en l'efecte nuclear Overhauser i correlaciona protons acoblats dipolarment, propers en l'espai. Els senyals es troben en mode d'absorció pura i en fase. Tot i ser l'experiment clau per a l'extracció de restriccions de distància, l'experiment NOESY presenta alguns problemes:

- 1) No permet distingir si els pics de correlació es deuen a la relaxació creuada entre nuclis propers en l'espai o a l'intercanvi químic (en què el pic de correlació correlaciona els desplaçaments químic dels protons en intercanvi).

2) Poden aparèixer pics de correlació deguts a la difusió de spin, és a dir, que pot correlacionar dos nuclis que no estan propers en l'espai però que s'han transferit la magnetització a través d'un tercer nucli que sí que està proper als dos primers.

Aquests problemes es poden solucionar mitjançant l'experiment ROESY (*Rotating frame nuclear Overhauser Effect Spectroscopy*)<sup>95</sup>, que permet distingir entre pics NOE directes i indirectes (correlació de spin), ja que presenten signes diferents; i també permet distingir entre els pics de correlació dipolar, de signe contrari als de la diagonal, i els que provenen de l'intercanvi químic, del mateix signe que els de la diagonal.

#### 1.4.2 Assignació espectral

La primera etapa per a realitzar un estudi conformacional per RMN és l'assignació de totes, o el major nombre de senyals dels espectres a àtoms concrets. Aquest procés resulta relativament complicat en el cas dels àcids nucleics ja que l'existència de només 4 nucleòtids implica moltes repeticions i dona lloc a solapament de senyals. En la Taula 1.2 es mostren els desplaçaments típics de protó dels diferents nucleòtids, expressats en ppm i referenciats a DSS<sup>96</sup> (sal sòdica de l'àcid 2,2-dimetil-2-silapentà-5-sulfònic). L'assignació es realitza amb l'ajuda d'experiments bidimensionals, seguint el procediment estàndard desenvolupat a la vegada per diversos grups<sup>97</sup>. Aquesta estratègia es basa en la combinació d'experiments de correlació escalar (TOCSY, COSY), per a la identificació dels diferents sistemes de spin, i d'experiments de correlació dipolar (NOESY), que permeten assignar els diferents sistemes de spin a residus específics de la cadena d'àcid nucleic. El protocol es basa en els contactes clau que s'observen en hèlix dextrogires del DNA i per això s'ha d'emprar amb cautela en el cas d'estructures irregulars.

Les ressonàncies corresponents als diferents tipus de protons no intercanviables estan força ben separades en l'escala de desplaçament químic, a excepció dels protons H4', H5' i H5'', que resulten difícils d'assignar. Gràcies a aquest fet, els espectres bidimensionals es poden analitzar per zones.

| Protó                             | Timidina | Citidina | Adenosina | Guanosina |
|-----------------------------------|----------|----------|-----------|-----------|
| H1'                               | 5.0-6.0  | 5.0-6.0  | 5.0-6.0   | 5.0-6.0   |
| H2'                               | 1.7-2.3  | 1.7-2.3  | 2.3-2.9   | 2.3-2.9   |
| H2''                              | 2.1-2.7  | 2.1-2.7  | 2.4-3.1   | 2.4-3.1   |
| H3'                               | 4.4-5.0  | 4.4-5.0  | 4.4-5.2   | 4.4-5.2   |
| H4'                               | 3.8-4.3  | 3.8-4.3  | 3.8-4.3   | 3.8-4.3   |
| H5'                               | 3.8-4.3  | 3.8-4.3  | 3.8-4.3   | 3.8-4.3   |
| H5''                              | 3.8-4.3  | 3.8-4.3  | 3.8-4.3   | 3.8-4.3   |
| H6/H8                             | 6.9-7.9  | 6.9-7.9  | 7.7-8.5   | 7.5-8.3   |
| CH <sub>3</sub> /H5/H2            | 1.0-1.9  | 5.0-6.0  | 7.7-8.5   | -         |
| NH (WC)                           | 13-14.5  | -        | -         | 11.5-13.5 |
| NH (lliure)                       | 10-11    | -        | -         | 10-11     |
| NH (GT)                           | 10-12    | -        | -         | 10-12     |
| NH (C <sup>+</sup> )              | -        | 15-16    | -         | -         |
| NH <sub>2</sub> (WC)              | -        | 6.5-8.5  | 7-8       | 8-9       |
| NH <sub>2</sub> (lliure)          | -        | 6.7-7.0  | 5-6       | 5-6       |
| NH <sub>2</sub> (C <sup>+</sup> ) | -        | 8-9.5    | -         | -         |

Taula 1.2. Valors dels desplaçaments químics de <sup>1</sup>H en DNA, expressats en ppm

Els valors de  $\delta$  són extremadament sensibles a l'entorn químic del nucli i varien amb la conformació de la molècula estudiada, però no existeixen relacions teòriques entre desplaçaments químics i paràmetres geomètrics i estructurals perquè el desplaçament no es pot correlacionar amb un únic paràmetre, sinó que caldria conèixer tota l'estructura. Per això, tot i que inicialment es van emprar els desplaçaments per a extreure'n informació estructural, ja no s'acostumen a utilitzar. De tota manera, sí que existeixen treballs sobre la predicció de desplaçaments químics en àcids nucleics, que poden ser útils per a explicar desplaçaments anòmals o extreure'n informació local<sup>98</sup>.

#### 1.4.2.1 Assignació de protons no bescanviables

Per a l'assignació de protons no bescanviables, primer cal identificar tots els sistemes de spin i després assignar cada sistema de spin a un residu específic. La identificació dels diferents sistemes de spin es duu a terme a partir dels pics de correlació escalar en els espectres TOCSY i COSY. Cada ribosa constitueix un sistema de spin aïllat de la resta de la molècula i podem anar seguint els pics de correlació H1'-H2'/H2'', H2'/H2''-H3', H3'-H4' i H4'-H5'/H5''.

Pel que fa a les bases, només la citosina i la timina tenen protons acoblats escalarment que ens permeten identificar-les. En el cas de la citosina s'observa un pic de correlació escalar entre els protons H5 i H6. Per la timina, s'observa la correlació entre els protons metílics i H6.

L'assignació de cada ribosa a cada base i de la posició de cada nucleòtid en la seqüència es realitza a partir dels pics de correlació dipolar en l'espectre NOESY. Per a l'assignació seqüencial s'utilitzen pics NOE característics d'hèlix dextrogires canòniques (Figura 1.14): els protons aromàtics (H6 en les pirimidines i H8 en les purines) presenten un pic NOE amb els protons H1' i H2'/H2'' del seu propi sucre i amb els mateixos protons de la desoxiribosa del residu anterior (5' a 3').

Per començar, s'ha d'escollir un pic de correlació inequívoc, com per exemple el pic entre el protó H<sub>Ar</sub> del residu 5'-terminal i l'únic protó H1', el seu propi. També es pot començar l'assignació a partir del senyal del protó aromàtic d'una citosina o timina que s'hagi identificat a partir de l'espectre COSY o TOCSY. A continuació, es busca el pic de correlació que connecti el senyal corresponent al protó H1' amb el protó H<sub>Ar</sub> del següent residu en el sentit 5'-3'. Coneixent H<sub>Ar</sub> s'identifica H1' del següent residu i així es va enllaçant un residu amb un altre fins a completar la cadena o fins que s'arriba a alguna discontinuïtat. Aquest procés es pot realitzar de forma anàloga amb els pics de correlació H6/H8-H2'/H2''. La combinació dels dos camins porta a una assignació inequívoca.

L'assignació estereoespecífica d'H2' i H2'' es pot realitzar en base a la intensitat del NOE amb H1': el pic H1'-H2'' (en *cis*) és més intens que el corresponent H1'-H2' (en *trans*). Aquesta apreciació cal

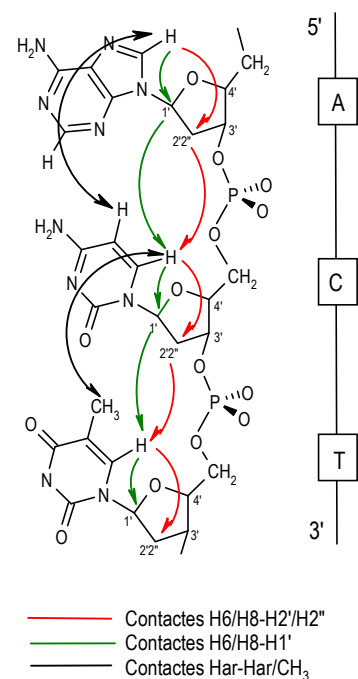


Figura 1.14. Contactes intra i interresidu en el sentit 5'-3' per a realitzar l'assignació seqüencial

realitzar-la en espectres NOESY amb un temps de mescla curt per evitar els efectes de la difusió de spin. En el B-DNA, els protons H2'' generalment ressonen a camps més baixos que els H2' corresponents.

L'assignació estereoespecífica d'H5' i H5'' està relacionada amb l'angle de torsió  $\gamma$  i és més complicada de realitzar. Cal combinar  $J_{4'5'}$  i  $J_{4'5''}$  amb les intensitats dels pics NOE H3'-H5'/H5'', H4'-H5'/H5'' i, a causa del gran solapament en aquesta zona de l'espectre, només és possible en oligonucleòtids de mida petita.

#### 1.4.2.2 Assignació dels protons bescanviables

De l'anàlisi dels espectres NOESY en aigua lleugera s'obté l'assignació dels protons bescanviables en base als pics de correlació intercadena amino-imino i imino-imino (Figura 1.15). En el cas de formació de parells WC, els parells G-C s'assignen seguint la connexió dels protons H5 de la citosina amb els seus protons amino i amb el protó imino de la guanina amb qui s'aparella. Els protons imino H3 de les timines s'identifiquen fàcilment pel pic de correlació intens amb el protó H2 de l'adenina amb qui s'aparellen.

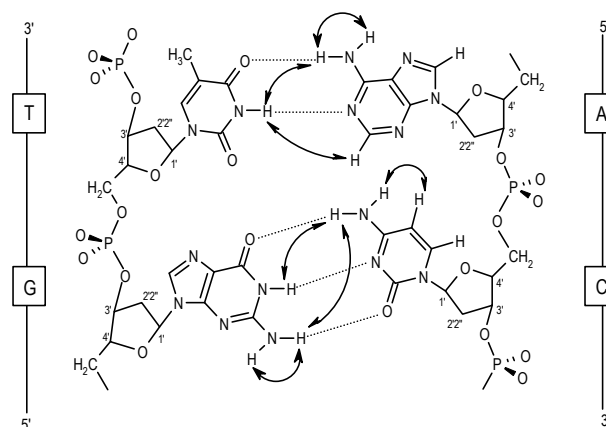


Figura 1.15. Contactes entre els protons bescanviables en la formació de parells de bases WC

També es donen contactes entre protons bescanviables i no bescanviables, tant amb el propi sucre com amb la base aparellada o amb les bases apilades (seqüencial).

#### 1.4.3 Informació estructural

D'entre la informació estructural que es pot extreure dels espectres d'RMN, la més rica i emprada en totes les determinacions estructurals fins ara prové de les restriccions de distància obtingudes de la integració dels pics NOE. Aquesta informació es pot complementar amb restriccions angulars derivades de l'anàlisi de les constants d'acoblament escalar a tres enllaços obtingudes d'experiments COSY i amb restriccions d'orientació obtingudes d'acoblements dipolars residuals en medis anisotròpics (RDCs)<sup>99</sup>.

##### 1.4.3.1 Distàncies obtingudes a partir valors de NOE

Les restriccions de distàncies experimentals s'obtenen a partir de la intensitat dels pics NOE, que depèn no només de la distància entre els protons sinó també del grau de mobilitat de la molècula, del temps de mescla i dels protons veïns. Una primera aproximació per calcular les distàncies interprotòniques és l'*aproximació de dos spins*, que suposa que la molècula es mou com un sòlid rígid i que cada parell de protons es troba aïllat de la resta. En aquesta aproximació la intensitat del NOE és proporcional al temps de mescla de l'experiment NOESY i decreix amb la sexta potència de la distància interprotònica, tal i com s'indica en la expressió:  $NOE \propto \frac{\tau_m}{r^6}$



Per calcular la constant de proporcionalitat s'utilitza la intensitat del pic NOE entre protons separats per una distància fixa, com els protons geminals o aromàtics. Aquesta aproximació presenta dues limitacions principals. En primer lloc, només és vàlida per a temps de mescla curts, en què es pugui menysprear l'efecte de la difusió de spin. En segon lloc, en el cas de macromolècules, la suposició que la molècula es mou isotròpicament amb un únic temps de correlació global no és gaire correcta, ja que generalment presenten un comportament dinàmic complex amb moviments interns en varies escales de temps.

A la pràctica, en molts casos com a estimació preliminar es classifica qualitativament la intensitat del NOE en fort, mig i dèbil: 1.8-3.0, 3.0-4.0 i 4.0-5.0 Å. En el cas de proteïnes i altres molècules globulars, aquesta classificació és suficient ja que la imprecisió es compensa amb el gran nombre de NOEs que s'observen, especialment amb els de llarg abast, entre residus allunyats en la seqüència i que restringiran molt l'espai conformacional. En el cas dels àcids nucleics no apareixen aquests NOEs de llarg abast atesa la naturalesa allargada de les estructures i únicament es disposa de restriccions seqüencials; per aquest motiu es fa necessari obtenir restriccions de distància més precises.

L'obtenció de restriccions de distàncies precises s'aconsegueix aplicant la teoria de la *matriu completa de relaxació*<sup>100</sup>, que considera tots els processos de relaxació d'un sistema de protons amb una estructura determinada, simulant els efectes de difusió de spin. Les intensitats dels pics de correlació NOE es representen mitjançant una matriu A ( $A_{ij}$  representa el pic NOE entre els protons  $i$  i  $j$ ), que es relaciona amb la geometria molecular mitjançant:

$$A(\tau_m) = e^{-R\tau_m} = 1 - R\tau_m + \frac{1}{2}R^2\tau_m^2 - \dots \quad R = \frac{-\ln\left[\frac{A(\tau_m)}{A(0)}\right]}{\tau_m}$$

on  $R$  és la matriu de relaxació ( $R_{ij}$  correspon a la relaxació entre els protons  $i$  i  $j$  i depèn de la distància entre ells  $(r_{ij})^{-6}$  i de la velocitat de reorientació del vector  $ij$  en el camp magnètic) i  $\tau_m$  el temps de mescla.

Aquesta equació es pot resoldre mitjançant la diagonalització de matrius i tècniques d'àlgebra lineal; el problema és que generalment no es disposa experimentalment de la matriu A completa. Aquest problema es soluciona emprant una matriu mixta o híbrida de relaxació, on tots els NOEs no observats es calculen a partir d'una estructura preliminar de la molècula. L'equació amb la matriu mixta de valors experimentals i calculats permet obtenir noves distàncies entre els protons, que s'inclouen en el següent càlcul, i així iterativament fins que es troba la millor solució. Aquest càlcul es realitza amb el programa MARDIGRAS<sup>101</sup>. Per minimitzar els errors, que poden venir de la determinació experimental de les distàncies o de les hipòtesis de coordenades inicials i temps de correlació, és convenient emprar diferents models inicials, tenir valors de NOEs amb diferents temps de mescla i repetir els càlculs amb diferents temps de correlació. La restricció final que s'obté és un rang de distàncies amb una cota inferior i una superior, però emprant la matriu completa de relaxació aquest rang és més estret i les estructures que s'obtenen estan més ben definides.

En el cas de protons bescanviabls, la intensitat del NOE es veu disminuïda pel bescanvi amb el

dissolvent. Per tant, les distàncies obtingudes seran sempre majors que les reals. Es podrien obtenir valors del grau d'intercanvi i introduir-los en el càlcul, però l'aproximació de considerar les distàncies com a cotes superiors és generalment suficient.

A la Taula 1.3 es resumeixen les distàncies interprotòniques més rellevants des d'un punt de vista estructural i la informació que se'n pot extreure de cadascuna d'elles.

| Distància   | Informació estructural   |
|---|--|
| $d_{pi}(T3; A2)$ , $d_{pi}(T3; A61)$<br>$d_{pi}(T3; A62)$ , $d_{pi}(T3; A8)$<br>$d_{pi}(G1; C41)$ , $d_{pi}(G1; C42)$ | Parell A-T: $d(N6-O4)=2.91\text{\AA}$ WC: NOE intens H3-H2, no existeix NOE H3-H8<br>$d(N3-N1)=2.91\text{\AA}$ HT: No existeix NOE H3-H2, NOE intens H3-H8<br>Parell G-C: $d(N4-O6)=2.91\text{\AA}$ WC: NOE intens H1-H42<br>$d(N3-N1)=2.95\text{\AA}$ HT: no existeix NOE H1-H42  |
| $d_i(2''; 4')$<br>$d_i(1'; 4')$   | Conformació S o N de la ribosa.<br>$d_i(2''; 4')=4.2\text{\AA}$ conformació S, ( $144 \leq P \leq 190^\circ$ ). $d_i(2''; 4')=2.8\text{\AA}$ conformació N, ( $0 \leq P \leq 36^\circ$ )<br>$d_i(1'; 4')=3.4\text{\AA}$ per ambdues conformacions però $d_i(1'; 4')=2.6\text{\AA}$ per valors intermedis de P ( $90^\circ$ ).                    |
| $d_i(3'; 5'/5'')$   | Determinació de l'angle de torsió $\gamma$   |
| $d_i(1'; 6/8)$<br>$d_i(2'/2''; 6/8)$  | Determinació del angle de torsió glicosídic, conformació <i>syn</i> o <i>anti</i> .<br>$d_i(1'; 6/8)=2.2\text{\AA}$ 3'-endo, disposició <i>syn</i> . $d_i(1'; 6/8)=3.7\text{\AA}$ 2'-endo, disposició <i>anti</i> .<br>Intensitat NOE H6/8-H2'~H6/8-H2'': disposició <i>syn</i> .<br>Intensitat NOE H6/8-H2'>H6/8-H2'': disposició <i>anti</i> . |
| $d_{cs}(2; 1'/2)3'$ , $d_s(2; 1')$<br>$d_s(2'; 3')$   | Distinció entre A- i B-DNA. Aquests NOEs seqüencials s'observen únicament en hèlix tipus A.<br>$d_s(2'; 3')$ : només mesurable en hèlixes tipus A (~7Å)  |
| $d_s(2'/2''; 5'/5'')$<br>$d_s(2'/2''/3'; 6/8/5)$<br>$d_s(2''; 2'')$ , $d_s(1'; 5'')$                                  | Distinció entre A- i B-DNA. $d_s(2'/2''; 5'/5'')$ petita per a A-DNA i gran per a B-DNA<br>$d_s(2'; 6//)$ petita i $d_s(2''; 6/8/)$ gran en hèlix tipus A, i al revés en hèlix tipus B.<br>$d_s(2''; 2'')$ i $d_s(1'; 5'')$ gran en hèlix A i petita en hèlix B.   |

WC = parells Watson-Crick, HT = parells Hoogsteen. Notació de distàncies:  $d_i(l;r)$  intranucleotídiques,  $d_s(l;r)$  seqüencials, l: nucleòtid 5'; r: nucleòtid 3',  $d_{pi}(l;r)$  entre parells de bases,  $d_{cs}(l;r)$  entre nucleòtids aparellats adjacents

Taula 1.3. Relació de distàncies interprotòniques estructuralment importants

A l'hora de determinar l'estructura tridimensional no totes les distàncies interprotòniques tenen la mateixa rellevància. Moltes d'elles no són útils per determinar l'estructura perquè varien molt poc amb la conformació. Per exemple, la conformació de l'anell de ribosa queda determinada per les distàncies  $d_i(2'';4')$  i  $d_i(1';4')$  (Figura 1.16). La resta de distàncies endocícliques no són rellevants perquè la seva variació queda dins del marge d'error experimental. La distància  $d_i(1'; 6/8)$  resulta crucial per determinar l'angle glicosídic i poder establir la disposició *syn* o *anti*. Les distàncies  $d_i(2'/2''; 6/8)$  també són importants per a determinar aquest angle.

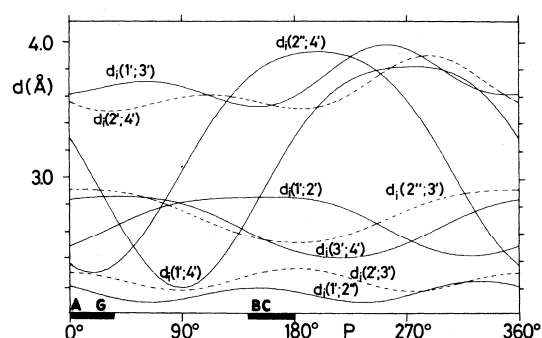


Figura 1.16. Variació de les distàncies entre els diferents protons de l'anell de desoxiribosa en funció de l'angle de pseudorotació  $P^{85a}$

#### 1.4.3.2 Constants d'acoblament

A part del desplaçament químic i l'acoblament dipolar, un altre fenomen observable per RMN és l'acoblament escalar o acoblament spin-spin, que és una interacció magnètica que es transmet a través dels electrons d'enllaç. Aquest acoblament dóna lloc a un desdoblament dels senyals i decreix amb el nombre d'enllaços. El paràmetre que el quantifica és la constant d'acoblament escalar ( $J$ , en Hz).

La conformació de l'anell de ribosa es pot predir a partir de les distàncies  $H1'-H4'$  i  $H2''-H4'$ , però aquestes no són suficients per descriure de forma precisa el plegament de l'anell. Per limitar la conformació de l'anell a una regió estreta del cicle de pseudorotació es necessitarien un gran nombre de distàncies intra i interresidu base-2'-desoxiribosa, amb la dificultat que podem tenir diferents conformacions en equilibri lent en l'escala de temps d'RMN. La forma més adequada per caracteritzar la conformació, i que permet alhora identificar un equilibri dinàmic entre vàries, és realitzar una anàlisi de les constants d'acoblament veïnals. Aquestes constants es poden obtenir de forma precisa a partir d'experiments derivats del COSY i es correlacionen amb l'angle diedre a través de l'equació de Karplus.

#### 1.4.4 Càlcul d'estructures en dissolució

La determinació de l'estructura tridimensional en dissolució consisteix en l'obtenció de coordenades atòmiques compatibles amb les restriccions experimentals. Es tracta de trobar un conjunt d'estructures compatibles amb la informació experimental que sigui una mostra representativa de tot l'espai conformacional. Per això, les estructures determinades per RMN es presenten com una superposició de varies conformacions, o se'n representa la seva estructura promig.

Existeixen una gran varietat de metodologies de càlcul, que es classifiquen en funció de l'aproximació matemàtica que utilitzen: mètodes de geometria de distàncies, que empren la matriu mètrica i operen en l'espai de distàncies, mètodes de minimització de funció blanc, que treballen en l'espai dels angles de torsió i mètodes de dinàmica molecular restringida (RMD).

El mètode de minimització de funció blanc consisteix en definir una funció positiva, que sigui zero quan es compleixin totes les restriccions de distància i diferent de zero en la resta dels casos. Aleshores el problema d'obtenir una estructura compatible amb les restriccions experimentals es redueix a minimitzar la funció esmentada. El mètode més emprat és l'anomenat "*steepest descent*", que consisteix en buscar, en cada iteració el mínim de la funció en la direcció indicada pel gradient. Aquests mètodes operen generalment en l'espai dels angles de torsió. En aquesta tesi s'ha emprat el programa DYANA<sup>102</sup>, que usa algorismes de dinàmica molecular en l'espai dels angles de torsió, per al càlcul de les estructures preliminars.

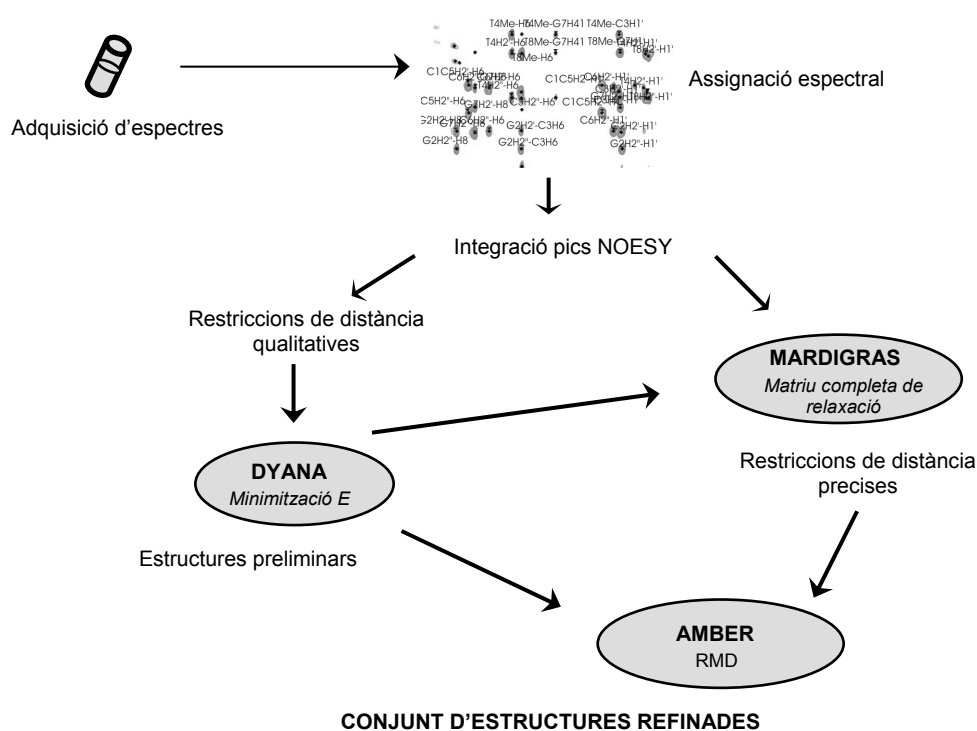


Figura 1.17. Esquema del procés de determinació estructural a partir d'RMN

El mètode de càlcul més emprat és el de Dinàmica Molecular Restringida (RMD), que consisteix en integrar numèricament les equacions de Newton d'un sistema de partícules:

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = -\nabla V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$$

Les coordenades de partida provenen de l'estructura preliminar (en el nostre cas calculada amb DYANA) i les velocitats es prenen d'una distribució de Maxwell-Boltzmann a una temperatura determinada. L'avantatge més gran de la RMD és la seva capacitat d'explorar tot l'espai conformacional i superar mínims locals. En la RMD el camp de força empíric es modifica per incorporar-hi un terme de penalització per incomplir les restriccions experimentals:

$$V_{total} = V_{empíric} + V_{experimental} = (V_{bond} + V_{angle} + V_{dihedral} + V_{VandeWaals} + V_{Coulomb}) + (V_{NOE} + V_{Jcoupling})$$

El potencial empíric inclou els termes clàssics: termes d'enllaç (distància d'enllaç, angle d'enllaç i torsions) i termes de no enllaç (interaccions de Van de Waals i electrostàtiques). La informació experimental s'incorpora com un terme de potencial, sense sentit físic, que penalitza les violacions de les restriccions obtingudes per RMN. El potencial experimental pot incloure les restriccions de distàncies i d'angles de torsió i es defineix com una funció que val zero quan el valor està comprès entre les cotes inferior i superior, i de tipus harmònic en els altres casos.

La dificultat està en el pes que se li dóna a les restriccions experimentals: si se li dóna massa pes, la dinàmica voldrà satisfer aquestes restriccions a costa dels termes teòrics, distorsionant l'estructura; per contra, si se li dóna massa poc pes, la dinàmica estarà guiada només pels termes teòrics i no es veuran satisfetes les restriccions experimentals. Per aquest motiu és convenient realitzar varis càlculs

preliminars amb diferents valors de les constants de manera que s'aconsegueixi un compromís i es puguin ajustar les restriccions experimentals sense distorsionar excessivament l'estructura.

Els protocols utilitzats s'anomenen d'*annealing* simulat i consisteixen en escalfar la molècula fins a temperatures relativament altes, per poder saltar barreres de potencial i superar mínims locals, i refredar suaument després. L'objectiu és obtenir valors d'energies baixos, tant de la part teòrica -indicatiu que l'estructura covalent és correcta- com de la part experimental -indicatiu que l'estructura compleix les restriccions experimentals-.

#### 1.4.5 Anàlisi d'estructures<sup>103</sup>

L'habilitat d'obtenir bones estructures depèn de la quantitat, exactitud i distribució de les restriccions. Una estructura estarà ben definida si tenim un mínim de 15 restriccions de distància per nucleòtid. Si tenim moltes menys restriccions, hauríem de parlar de modelat més que de determinació estructural.

L'anàlisi de les estructures resultants dels càlculs de RMD inclou diferents aspectes:

- Determinar la qualitat de l'ajust de la informació experimental, és a dir, determinar fins a quin punt les estructures finals reproduïxen les dades experimentals. Un indicador n'és el valor final del potencial experimental: com menor sigui, més ajustades estaran les restriccions. Però aquest valor també depèn de la forma concreta de la funció i dels factors de pes emprats. Per aquest motiu s'acostuma a utilitzar la suma de violacions residuals de les restriccions en les estructures finals, que no depenen del potencial utilitzat. Aquest indicador és força útil, però no deixa de ser un indicador indirecte, ja que la veritable informació no són les distàncies sinó les intensitats dels pics NOE, i pot dependre de l'aproximació que s'hagi emprat per convertir les intensitats en distàncies. També es pot emprar el factor R (factor de fiabilitat, *reliability factor*)<sup>104</sup>, que compara l'estructura final amb les dades experimentals d'RMN, anàlogament al que es fa en cristal·lografia de Raigs X.

- Analitzar el grau de definició de les estructures resultants. S'indica a partir del valor de RMSD (*Root Mean Square Deviation*) del conjunt d'estructures resultants del càlcul:

$$RMSD = \sqrt{\frac{\sum_{i=1}^n \|v_i - w_i\|^2}{n}}$$

S'ha de tenir cura, però, ja que és possible fer-lo petit emprant restriccions massa estrictes que donin estructures incorrectes. Per tant, és més un indicador de la precisió dels resultats que de la seva exactitud. El valor de RMSD també és útil per comparar dues estructures diferents.

- Analitzar les estructures resultants: càlcul d'angles de torsió i de paràmetres helicoidals per comparar-los amb els valors canònics o d'altres estructures i poder avaluar la validesa de les estructures calculades. Valors inusuals d'angles de torsió o conformació del sucre poden indicar inconsistències en les restriccions o en el càlcul. El programa CURVES<sup>105</sup> permet calcular els paràmetres estructurals definint eixos d'hèlix locals.

## 2 INTRODUCCIÓ I OBJECTIUS

---

### 2.1 El motiu *bi-loop*

Com ja s'ha dit a la introducció general de la tesi, al nostre grup de recerca, en estudiar l'estructura per cristal·lografia de RX de l'oligonucleòtid cíclic d<pCATTCAAT> es va observar un motiu estructural que ja s'havia vist en el cristall de l'heptàmer lineal d(GCATGCT)<sup>58</sup>, i que s'anomenà *bi-loop*<sup>59</sup>, atesa la seva forma. La possibilitat que es tractés d'un artefacte cristal·logràfic es va rebatre mitjançant la determinació de l'estructura en solució per RMN del mateix octàmer cíclic d<pCATTCAAT> i també de d<pTGCTCGCT><sup>60</sup>, que dimeritzen en solució formant un *bi-loop*. Ambdós dímers existeixen en equilibri conformacional amb la forma monomèrica *dumbbell*<sup>55c</sup> (Figura 2.1).

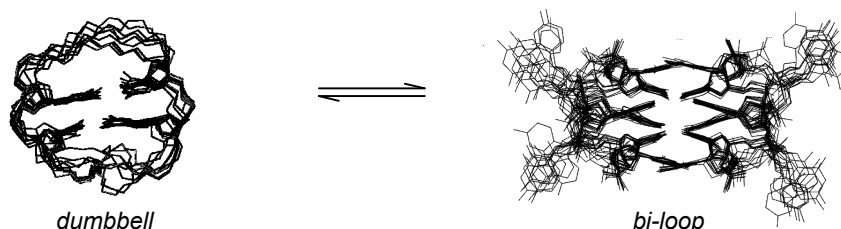


Figura 2.1. Equilibri en solució entre les estructures *dumbbell* i *bi-loop* de d<pCATTCAAT>

L'estudi es va prosseguir amb la determinació mitjançant RMN de l'estructura en solució d'altres octàmers cíclics (veure resum en la Taula 2.1) per intentar trobar els requeriments seqüencials per a la formació del *bi-loop* i estudiar l'efecte de canviar la naturalesa i l'ordre de les bases aparellades o del llaç en la seva formació i estabilitat.

Per una banda, es va estudiar l'efecte de canviar les bases del llaç, amb les seqüències d<pAGCTAGCT> i d<pTGATGCA>. La base que ocupa la primera posició del llaç és important perquè forma part de l'apilament, mentre que la segona queda exposada al dissolvent. La presència d'un residu d'adenina en la primera posició del *loop* estabilitza la forma dimèrica, presumiblement per la formació d'un parell A-A, com l'observat en l'estructura de RX de l'heptàmer d(GCATGCT)<sup>58</sup>, de seqüència molt similar, que conferiria una major estabilitat que la interacció hidrofòbica dels Me quan

tenim T en la primera posició del *loop*. Aquesta major estabilització també podria ser deguda a l'apilament més favorable d'una purina que d'una pirimidina. Quan el residu d'adenina es troba en la segona posició del *loop*, desestabilitza enormement, ja que està totalment exposada al dissolvent.

També es va estudiar l'efecte de canviar les bases que formen els parells intermoleculars. En l'estudi de l'octàmer d<pCGCTCATT><sup>106</sup>, es va observar que es poden formar tètredes mixtes amb parells G·C i A·T, però amb una estabilitat força baixa, similar a la de la seqüència estabilitzada només per parells A·T. Amb l'estudi dels octàmers d<pCGTCCGT><sup>107</sup> i d<pCGCTCCGT><sup>108</sup>, on s'ha canviat l'ordre d'apilament de les bases, es va advertir que també formen un *bi-loop* estabilitzat per parells G·C però donant lloc a tètredes marcadament diferents. Aquests tres últims oligonucleòtids estudiats, a diferència dels anteriors, no s'autoestructuren en forma de *dumbbell* a concentracions baixes d'oligonucleòtid. En els casos de d<pCGCTCATT> i d<pCGTCCGT> és evident, donada la seqüència, que no poden formar parells de bases intramoleculars. Però en el cas de d<pCGTCCGT>, sí que podria autoestructurar-se en forma de *dumbbell* i, en canvi, no ho fa, sinó que a baixes concentracions només s'observa el monòmer desestructurat<sup>55c</sup>. Aquesta diferència amb les seqüències anteriors, especialment amb l'anàloga d<pTGCTCGCT>, reflecteix la tendència observada en altres *hairpins*<sup>109</sup> o *dumbbells*<sup>55b</sup> de DNA: la seqüència 5'-CXXG-3' forma *loops* de dos residus, mentre que la seqüència 5'-GXXC-3' prefereix adoptar *loops* de quatre residus.

Posteriorment es va voler explorar la possibilitat de formar el *bi-loop* amb dos oligonucleòtids de seqüència diferent, és a dir, la possibilitat de formar un *bi-loop* heterodimèric, ja que fins al moment només havia estat observat en homodímers. El *bi-loop* heterodimèric es va poder caracteritzar pel cas de les seqüències: d<pCAGTCCCT> + d<pCCTTCGGT><sup>108b</sup> (veure capítol 3).

| Seqüència                     | Tètredes                        | Bases del <i>loop</i> | Dumbbell?             | Estabilitat ( $\Delta G_{298}^0$ ) <sup>110</sup> |
|-------------------------------|---------------------------------|-----------------------|-----------------------|---|
| d<pCATTCAATT>                 | A:T:A:T                         | TC                    | Sí (baixes conc)      | - 13 kJ/mol                                       |
| d<pTGCTCGCT>                  | G:C:G:C                         | TC/TT                 | Sí (molt baixes conc) | - 32 kJ/mol                                       |
| d<pAGCTAGCT>                  | G:C:G:C                         | TA                    | Sí (molt baixes conc) | - 25 kJ/mol                                       |
| d<pTGCATGCA>                  | G:C:G:C                         | AT                    | Sí (molt baixes conc) | - 38 kJ/mol                                       |
| d<pCGCTCATT>                  | G:C:A:T                         | TC                    | No es pot formar      | - 9 kJ/mol  |
| d<pCGTCCGT>                   | G:C:G:C desplaçada              | TC                    | No es forma           | - 35 kJ/mol                                       |
| d<pCGCTCCGT>                  | G:C:C:G                         | TC                    | No es pot formar      | ---   |
| d<pCAGTCCCT>+<br>d<pCCTTCGGT> | G:C:A:T +<br>G:C:G:C desplaçada | TC                    | No es pot formar      | ---   |

Taula 2.1. Resum de les seqüències estudiades fins al moment

Pel que fa a l'estabilitat del motiu *bi-loop*, l'efecte més important és la naturalesa de les bases que formen les tètredes, essent les de tipus G·C les més estables, però sense que l'ordre de les bases sigui especialment important. Els residus que formen els *loops* també influeixen en l'estabilitat: RY > YY > YR.

## 2.2 Estructura i paràmetres estructurals

El *bi-loop* és un motiu dimèric simètric en què els dos oligonucleòtids es disposen antiparal·lelament i es manté estabilitzat per quatre parells de bases WC intermoleculares que formen dos apilaments. Els apilaments presenten una inclinació relativa d'entre 30 i 40° i enfronten els parells de bases pel solc menor formant dues tètredes no planes. L'estructura global és molt similar en tots els casos, però les tètredes que l'estabilitzen poden ser de diferents tipus (veure apartat 2.3). Les anomenem “tètredes de solc menor” en contraposició a les tètredes observades en el context de quàdruplex de guanina, que enfronten els parells WC pel solc major i són planes (veure apartat 2.3.2).

Els parells de bases estan connectats per dos minillaços de dos nucleòtids cadascun. La primera de les bases del llaç (sentit 5'-3') 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. La primera base del llaç interacciona per apilament amb els parells de bases i també per contactes hidrofòbics amb la base apilada de l'altra unitat, quan són timines.

A banda de les interaccions per enllaç d'hidrogen i apilament de les bases, l'estructura també està estabilitzada per contactes hidrofòbics entre els anells de ribosa. Per contra, la gran proximitat dels esquelets fa que s'observin distàncies P-P molt curtes, d'entre 5.5 i 6 Å. Les repulsions entre els grups fosfat es veuen altament compensades per la presència de cations, especialment divalents, com s'observa en les estructures de RX de d(GCATGCT) ( $Mg^{2+}$  i  $Co^{2+}$ )<sup>58</sup> i d<pCATTcATT> ( $Ba^{2+}$ )<sup>59</sup>.

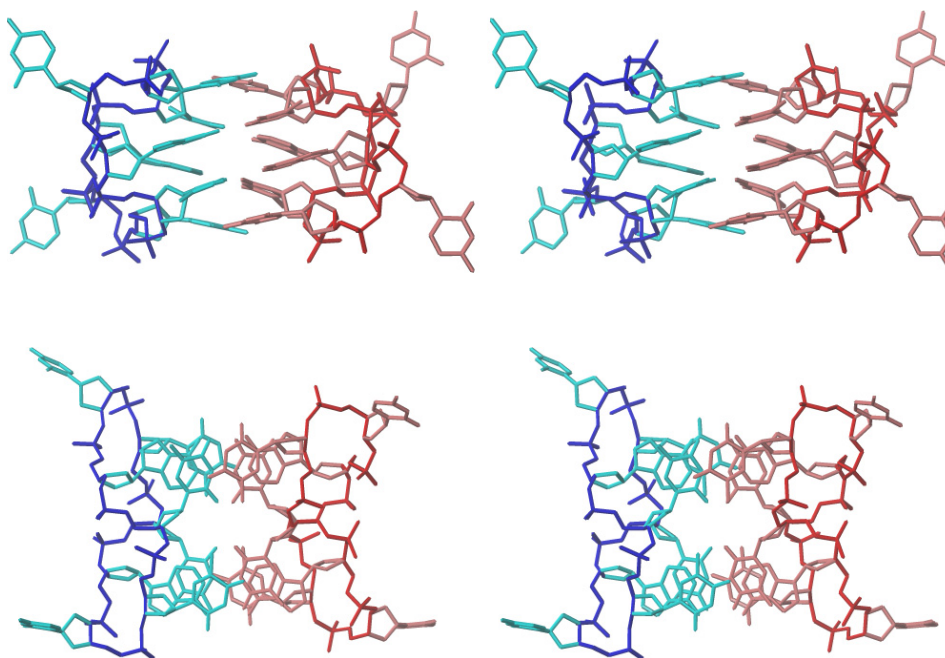


Figura 2.2. Vistes estereoscòpiques (*lateral i superior*) de l'estructura promig de d<pCATTcATT>. Cada unitat es mostra en un color diferent; l'esquelet en color més fosc

El dímer és totalment simètric; presenta dos eixos de simetria, que es poden observar en la Figura 2.2: eixos C2 perpendiculars al pla en les dues vistes mostrades.



Els enllaços glicosídics es troben en la conformació *anti* i els anells de ribosa presenten la conformació C2'-*endo*. El nucli central del *bi-loop* és essencialment un fragment de B-DNA.

A la Taula 2.2 es mostren els paràmetres helicoidals locals que descriuen la posició relativa entre parells de bases apilats per al nucli central del *bi-loop*. Aquests paràmetres han estat calculats per cadascun dels apilaments i es mostren els valors absoluts promig. Els paràmetres més importants relacionats amb l'apilament de les bases són l'ascensió (*rise*) i el gir (*twist*). Es pot observar que pel cas de tenir apilaments AT, l'ascensió entre els parells de bases centrals és major que en el cas de parells GC, en què l'apilament és més fort que l'estàndard en dúplex de B-DNA. Pel que respecta al gir, aquest és molt més pronunciat en el cas de la seqüència amb l'apilament 5'-CG-3' que en la resta. Aquest fet no és estrany ja que en dobles hèlixs ocorre que els passos YR presenten menor solapament dels anells de les bases que els passos RY o RR<sup>111</sup>. Aquesta diferència en l'apilament es pot veure en la Figura 2.10 i és clar que afectarà a les propietats espectroscòpiques del *bi-loop*, com s'explicarà a l'apartat 2.4.

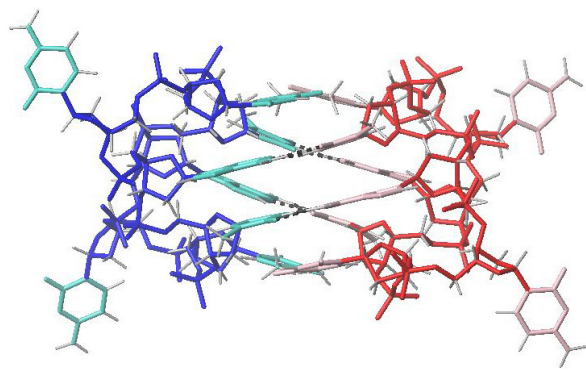
| Seqüència                               | Apilament | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
|---|-----------|-----------|-----------|----------|----------|----------|-----------|
| d<pCATTCAATT>                           | AT        | 0.1       | 1.5       | 3.2      | 1.0      | 0.6      | 25        |
| d<pTGCTCGCT>                            | GC        | 0.3       | 0.4       | 2.8      | 10.5     | 3.5      | 26        |
| d<pCGCTCAATT>                           | GC        | 0.1       | 1.5       | 3.1      | 2.0      | 3.0      | 28        |
|   | AT        | 0.0       | 0.7       | 4.3      | 8.0      | 21.0     | 29        |
| d<pCCGTCCGT>                            | CG        | 0.8       | 0.3       | 2.8      | 10.6     | 12.0     | 42        |
| d<pCGCTCCGT>                            | GC        | 0.1       | 1.1       | 3.2      | 3.6      | 3.3      | 25        |
|   | CG        | 0.4       | 0.7       | 3.4      | 2.6      | 8.1      | 32        |
| B-DNA estàndard <sup>111</sup> (promig) |           | 0.0       | 0.4       | 3.4      | 0.0      | 0.6      | 36        |

Taula 2.2. Resum dels paràmetres helicoidals promig en els passos centrals de bases apilades (2-3, 6-7)

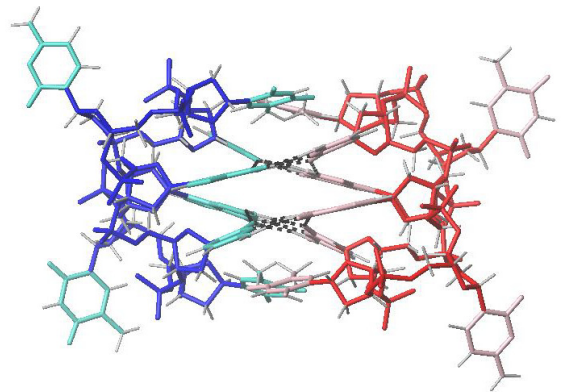
Globalment totes les estructures són molt similars, com es pot observar en la Figura 2.3 i pels valors de RMSD calculats sobre els C1' de les diferents estructures (Taula 2.3), tot i que les tètredes que les estableixen són força diferents. Els RMSDs calculats sobre els C1' dels nucleòtids implicats en el nucli de parells de bases són molt petits, i es situen dins de l'ordre dels valors obtinguts per al conjunt d'estructures d'RMN per a cada seqüència.

| Seqüència     | d<pTGCTCGCT> | d<pCATTCAATT> | d<pCGCTCAATT> | d<pCCGTCCGT> | d<pCGCTCCGT> |
|---------------|--------------|---------------|---------------|--------------|--------------|
| d<pTGCTCGCT>  | ----         | 1.2           | 2.4           | 1.2          | 0.8          |
| d<pCATTCAATT> | 0.8          | ----          | 2.6           | 1.8          | 0.5          |
| d<pCGCTCAATT> | 1.1          | 0.6           | ----          | 1.7          | 0.7          |
| d<pCCGTCCGT>  | 0.8          | 0.9           | 1.0           | ----         | 0.8          |
| d<pCGCTCCGT>  | 1.2          | 1.9           | 1.9           | 0.9          | ----         |
| d(GCATGCT)    | 0.5          | 0.8           | 1.0           | 0.9          | 0.8          |

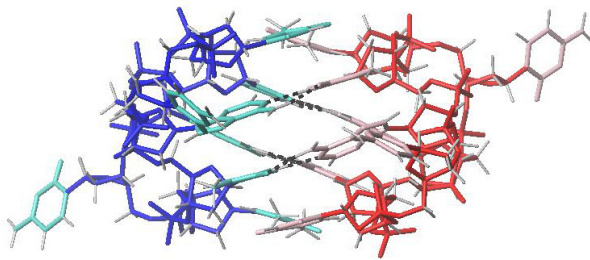
Taula 2.3. RMSDs (Å) entre parelles d'estructures dels diferents *bi-loops*. En la part superior de la diagonal els calculats entre tots els C1', i en la part inferior només entre els C1' dels nucleòtids implicats en les tètredes



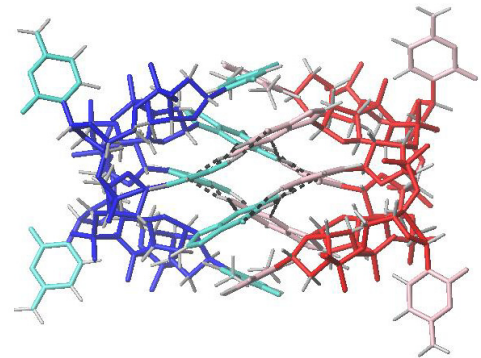
d<pCATT<math>CATT</math>>



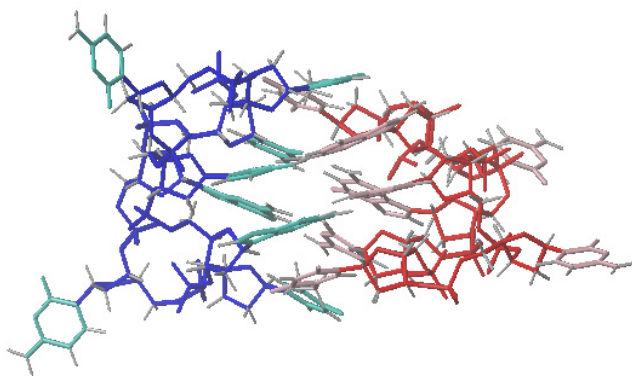
d<pTGCT<math>CGCT</math>>



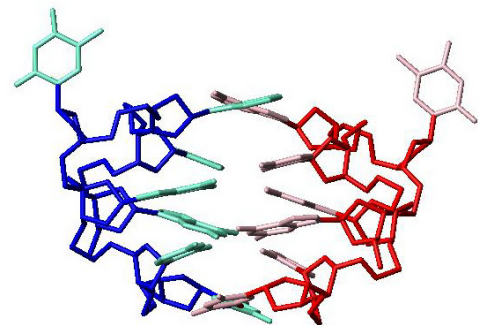
d<pCGCT<math>CATT</math>>



d<pCCGT<math>CCGT</math>>



d<pCGCT<math>CCGT</math>>



d(GCAT<math>GCT</math>)

Figura 2.3. Vistes laterals de les estructures promig dels bi-loops homodimèrics observats

### 2.3 Diferents tipus de tètredes observades

En la Figura 2.4 es mostren les diferents tètredes observades fins al moment en els diferents *bi-loops*.

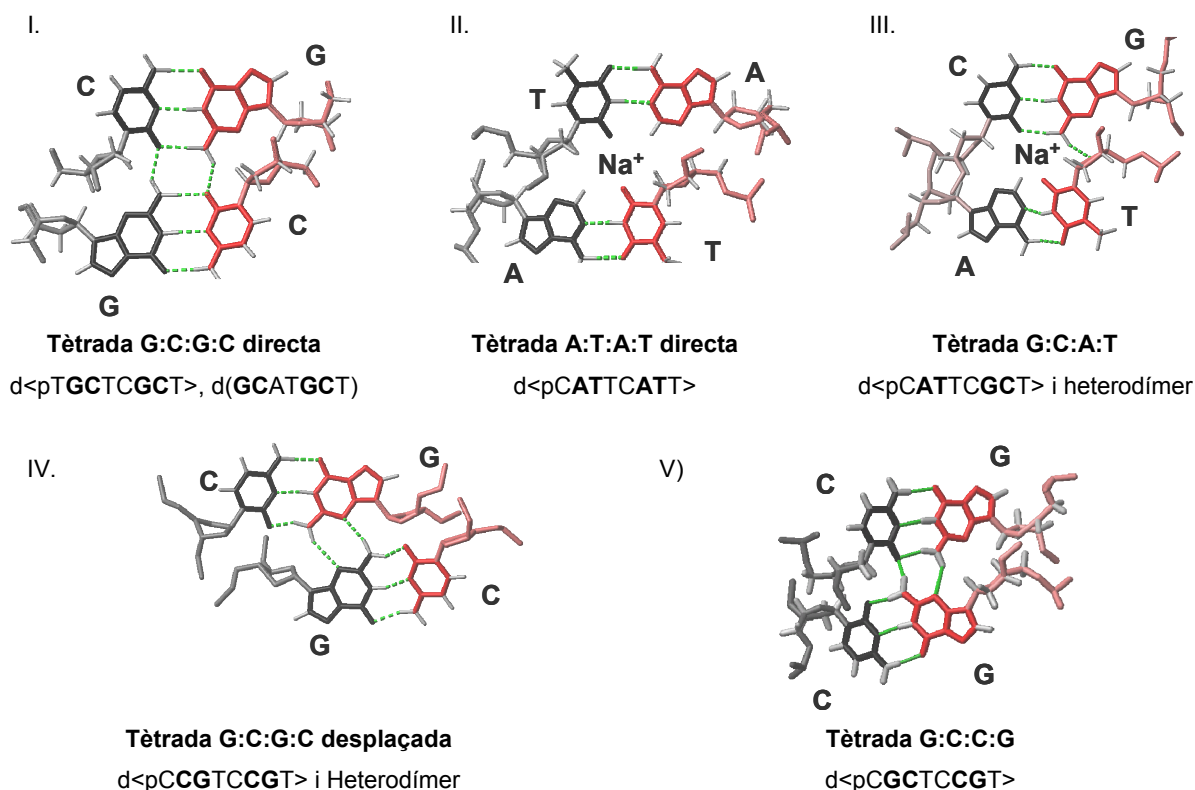


Figura 2.4. Tètredes observades i seqüències en què s'han observat

El tipus d'interaccions que estableixen cada tètreda és diferent. En el cas de les tètredes ATAT i GCAT, l'estabilització prové de la presència d'un catió al centre de la molècula, probablement coordinat als oxígens de les timines, tal i com es va observar en l'estructura de RX<sup>58</sup> i com es va calcular per a la tètreda mixta<sup>106</sup>. En les tètredes amb parells G-C, l'estabilització prové de ponts d'hidrogen addicionals entre els dos parells WC. Aquestes interaccions addicionals depenen de la disposició relativa dels parells G-C. En la tètreda directa (I) els enllaços d'hidrogen addicionals tenen lloc entre el protó H21 de la guanina i el carbonil (O2) de la citosina, mentre que en la tètreda desplaçada (IV) la interacció ocorre entre l'H21 de la G i l'N3 de la guanina d'enfront. En la tètreda GCCG (V) existeixen dues interaccions no simètriques: H21(Ga)-O2(Cb) i H21(Gb)-N3(Ga). Així doncs, les tètredes amb parells G-C estan estabilitzades per 8 ponts d'hidrogen, en contraposició amb els 4 que estableixen la tètreda ATAT i els 5 que estableixen la GCAT, motiu que podria explicar la marcada diferència en estabilitat dels *bi-loops* que les contenen.

Cal remarcar la correlació entre l'ordre d'apilament de les bases i el tipus de tètreda formada. Quan l'ordre d'apilament és 5'-pur-pyr-3', com en les seqüències inicialment estudiades, es formen tètredes que enfronten els dos parells de bases i que anomenem directes (tètredes I, II i III)\*. En canvi, quan

\* La tètreda GCGC és la única realment directa, ja que en les tètredes ATAT i GCAT els parells de bases estan correguts ~1Å.

canviem l'ordre d'apilament per 5'-pyr-pur-3', la tètrada formada és desplaçada. En aquest cas només es disposa de l'exemple de la tètrada GCGC desplaçada o *slipped* (IV). El cinquè tipus de tètrada s'observa quan la seqüència no és repetitiva i presenta un apilament 5'-pur-pyr-3' i un altre 5'-pyr-pur-3', donant lloc a una tètrada no simètrica on els parells de bases G·C estan disposats en el mateix sentit.

### 2.3.1 Observació d'aquest tipus de tètrades en estructures cristal·lines

La majoria dels oligonucleòtids cristal·litzats tenen parells G·C als extrems, que en molts casos interaccionen amb les hèlixs veïnes a través del solc menor<sup>112</sup>. Aquestes interaccions contribueixen a estabilitzar l'organització tridimensional del cristall i generalment impliquen un contacte de les guanines a través del solc menor, formant ponts d'hidrogen entre l'N3 d'una guanina i un protó de l'amina (N2) de la guanina del dúplex veí. En la majoria dels casos, els dos parells G·C dels extrems es troben intactes i es formen dues tètrades amb una estructura molt similar a la GCGC desplaçada. Aquesta interacció es dona entre dúplexs apilats formant columnes i és molt comuna en els dodecàmers acabats amb passos CG, com el dodecàmer de Dickerson<sup>113</sup>, i per això ha estat anomenada "interacció de Dickerson". Aquesta interacció no es troba en el cas dels decàmers, que senzillament apilen els parells de bases, formant com una hèlix continua però tallada, però sí que s'ha observat també en octàmers<sup>114</sup>. En la Figura 2.5 es mostra la superposició de les columnes en el cristall i el detall de la interacció, que dona lloc a una estructura extremadament semblant al nostre *bi-loop*.

En altres casos, les guanines terminals no estan aparellades i envaeixen el solc menor del dúplex adjacent, formant el que podríem anomenar tríades o triplets<sup>115</sup>. També es dona el cas que les citosines es troben majoritàriament desordenades<sup>116</sup> i només es conserven els contactes G·G.

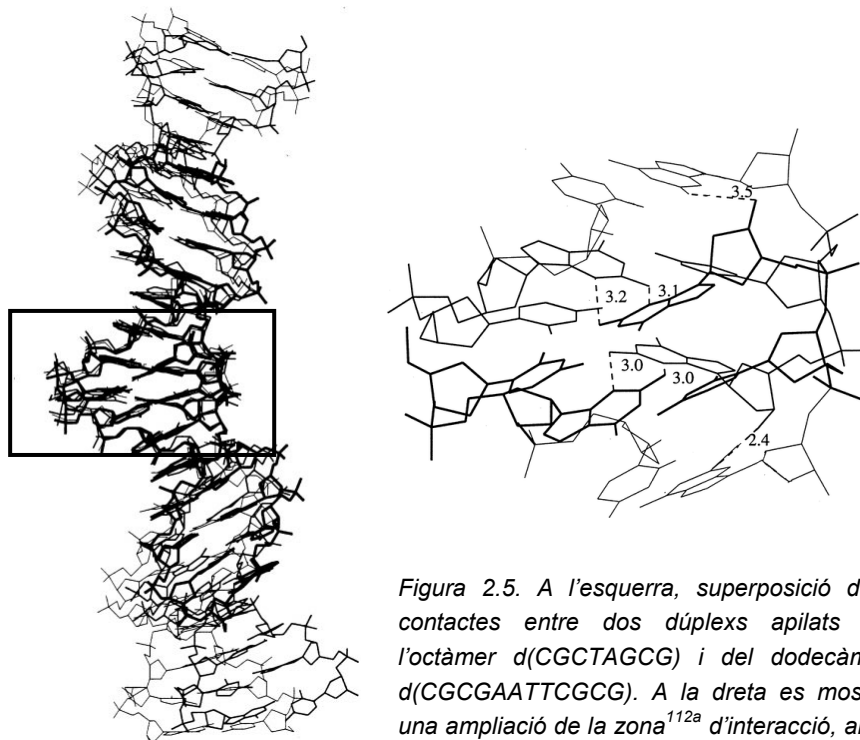


Figura 2.5. A l'esquerra, superposició dels contactes entre dos dúplexs apilats de l'octàmer d(CGCTAGCG) i del dodecàmer d(CGCGAATTCGCG). A la dreta es mostra una ampliació de la zona<sup>112a</sup> d'interacció, amb les distàncies de pont d'hidrogen addicionals entre les guanines

Les tètrades GCGC desplaçades han estat observades en cristalls de dúplex en forma B del DNA. Per contra, en cristalls en forma A s'han observat<sup>117</sup> contactes a través del solc menor que donen lloc a la formació de tètrades GCGC directes<sup>117b</sup> o del tipus GCCG<sup>117c</sup> (Figura 2.6). Aquest fet es pot

explicar ja que en la forma A el solc menor és més ample i poc profund i pot facilitar aquest tipus d'interaccions. En aquests casos es tracta de decàmers i els contactes es donen entre un parell G-C a l'extrem de la seqüència que es col·loca en el solc menor d'un dúplex veí amb el que es relaciona per simetria, interaccionant amb un altre parell G-C situat al centre de la seqüència. L'angle observat entre aquests parells de bases és de 110-120°. En el cas del que nosaltres anomenem tètrada GCCG, els autors l'anomenen tríada extrahelicoïdal ja que només és la guanina d'un parell la que interacciona amb el parell G-C de l'extrem. En alguns casos, una única guanina de l'extrem pot participar en dos d'aquests triplets. Cal remarcar que aquests tipus de contactes també s'han observat en una seqüència que conté un ribonucleòtid<sup>117d</sup> de guanina [d(CCGGC)r(G)d(CCGG)] i que aquest participa en el contacte quadruplet (Figura 2.6 inferior); o sigui, sí que seria possible formar aquest tipus de tètredes amb almenys un ribonucleòtid. En seqüències d'octàmers que cristal·litzen en forma d'A-DNA s'han observat altres tipus de triplets per formació d'un únic pont d'hidrogen a través del solc menor de parells G-C<sup>118</sup>.

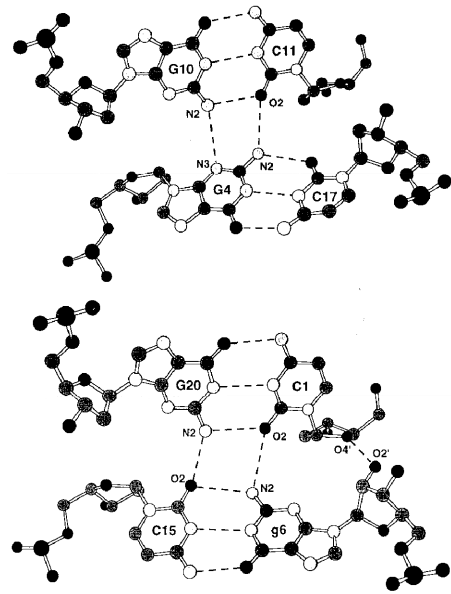


Figura 2.6. Contactes observats en cristalls d'A-DNA<sup>117a</sup>

Per acabar, també cal comentar que recentment s'han publicat set estructures de complexos de DNA amb agents intercalants<sup>119</sup>, en què els parells G-C dels extrems no formen parell WC amb el seu propi dúplex sinó amb un dúplex adjacent amb el qual es relacionen per simetria. Aquests complexos s'han cristal·litzat en tots els casos amb l'hexàmer d(CGTAACG) i en presència de Co<sup>2+</sup>. Sempre s'observa que el nucli central de d(TA) forma un B-DNA, mentre que les bases dels extrems es troben cap a fora del dúplex i interaccionen amb altres bases d'altres dúplex formant una intersecció entre 4 cadenes diferents amb l'agent intercalant al centre de la cavitat. Aquesta intersecció està formada per dues tètredes de solc menor GCGC desplaçades, amb l'agent intercalant desordenat apilat entre les dues tètredes (Figura 2.7), i s'assembla molt al *bi-loop* format per d<pCCGTCCGT>.

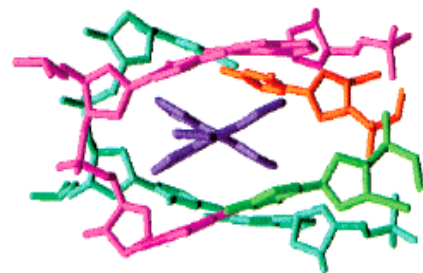


Figura 2.7. Detall de la intersecció quàdruplex amb l'agent intercalant en color lila i cada cadena d'un color diferent<sup>119e</sup>

Cal remarcar que la mateixa seqüència d(CGTAACG) s'ha cristal·litzat amb altres agents intercalants en absència de Co<sup>2+</sup> i no s'ha observat aquest encreuament tipus quàdruplex, sinó la intercalació esperada del lligand<sup>120</sup>. Per tant, sembla que aquesta estructura quàdruplex està estabilitzada per interaccions DNA-DNA, interaccions N7(G) amb Co<sup>2+</sup> i forces d'empaquetament, però no per una interacció especial amb els lligands. Aquests només serien elements passius que mantindrien les dues tètredes a la distància adequada per a l'empaquetament (3.4 Å) i, per tant, aquestes estructures

no serien rellevants per entendre el mecanisme amb què alguns dels agents intercalants inhibeixen la topoisomerasa.

### 2.3.2 Comparació amb tètredes de solc major

En el context de les estructures de quàdruplex de guanina s'han pogut observar altres tipus de tètredes, diferents de la tètreda de guanina i apilades sobre o entre aquestes. S'han observat tètredes formades per un únic tipus de base en estructures de quàdruplex de G paral·lels: tètreda d'A<sup>121</sup>, tètreda de C<sup>122</sup> o tètreda de T<sup>123</sup>. Tètredes formades per parells WC han estat observades tant en quàdruplex paral·lels com en estructures replegades. En tots els casos els parells WC s'associen pel solc major formant tètredes planes. Aquests parells es poden enfrontar formant tètredes directes del tipus ATAT<sup>124</sup> o GCGC<sup>125</sup>, o formant tètredes desplaçades, tant ATAT<sup>126</sup> com GCGC<sup>127</sup> (Figura 2.8). Aquesta darrera ha estat observada en la mateixa seqüència en què es va observar una tètreda GCGC directa<sup>125b</sup>, en canviar el catió Na<sup>+</sup> per K<sup>+</sup>. La gran diferència entre aquestes tètredes i les de solc menor és que aquestes són planes, mentre que les de solc menor presenten una inclinació de 30-40° entre els parells de bases. A més a més, cal destacar que les tètredes de solc major només han estat observades en el context de les estructures de quàdruplex de guanina i mai establint una estructura elles soles. Generalment els parells de bases d'aquestes tètredes de solc major són intramoleculares, mentre que en el cas del *bi-loop* són intermoleculares.

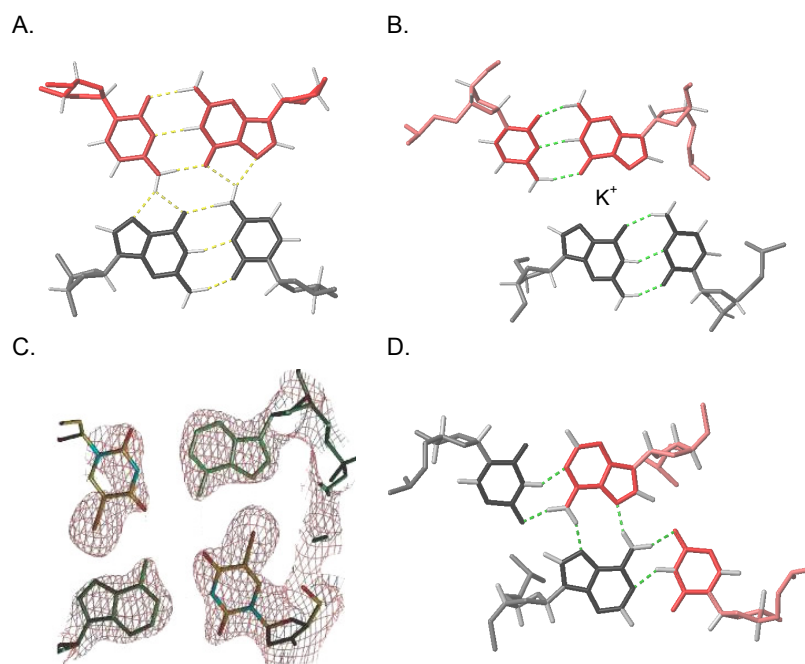


Figura 2.8. Tètredes de solc major observades en el context d'estructures de quàdruplex de guanina. A. Tètreda GCGC directa<sup>127b</sup>. B. Tètreda GCGC desplaçada<sup>127</sup>. C. Tètreda ATAT directa<sup>124</sup>. D. Tètreda ATAT desplaçada<sup>126</sup>. Cada cadena s'indica d'un color diferent.

També s'ha observat una tètreda GTGT directa amb parells no canònics<sup>128</sup> apilada sobre tètredes de G i altres tipus de contactes tipus tríades<sup>129</sup>, hèxades<sup>130</sup> i fins i tot hèptades<sup>131</sup>.

## 2.4 Característiques espectroscòpiques del motiu

Les característiques espectroscòpiques aquí recopilades seran útils per a la detecció i la caracterització del *bi-loop* en solució en els estudis que durem a terme en aquesta tesi.

### 2.4.1 Espectroscòpia d'ultraviolat i dicroisme circular

L'espectroscòpia d'ultraviolat i el dicroisme circular s'han emprat per a la detecció, caracterització i estudi del motiu *bi-loop* i han demostrat ser útils tot i ser difícil extreure'n patrons de comportament general<sup>110,108a</sup>. A continuació s'intenten resumir els trets característics de les corbes de fusió per UV-Vis i dels espectres de dicroisme circular de les diferents seqüències estudiades fins al moment que presenten estructura *bi-loop*, junt amb les seves característiques estructurals (Taula 2.4).

| Seqüència                           | Característiques estructurals      |                                     |                                 | Corba de fusió per UV ( $\lambda=260$ nm) | Dicroisme circular    |                       |
|-------------------------------------|------------------------------------|-------------------------------------|---------------------------------|---|-----------------------|-----------------------|
|                                     | tètrades                           | apilament de bases                  | bases del <i>loop</i> (5' a 3') |   | $\lambda_{\max}$ (nm) | $\lambda_{\min}$ (nm) |
| 1 d<pCATTCAATT>*                    | A:T:A:T                            | 5'-pur-pyr-3'                       | T C                             | --  | --                    | --                    |
| 2 d<pTGCTCGCT>                      |                                    |                                     | T C<br>TT                       | hipocròmica                               | 290                   | 260                   |
| 3 d<pTGCATGCA>                      | G:C:G:C                            | 5'-pur-pyr-3'                       | A T                             | hipocròmica                               | 291                   | 255                   |
| 4 d<pAGCTAGCT>                      |                                    |                                     | T A                             | hipocròmica                               | 290                   | 260                   |
| 5 d<pCGCTCAATT>*                    | G:C:A:T                            | 5'-pur-pyr-3'                       | T C                             | --  | --                    | --                    |
| 6 d<pCCGTCCGT>                      | G:C:G:C desplaçada                 | 5'-pyr-pur-3'                       | T C                             | hipercròmica                              | 265                   | 243                   |
| 7 d<pCGCTCCGT>                      | G:C:C:G                            | 5'-pur-pyr-3'<br>+<br>5'-pyr-pur-3' | T C                             | hipercròmica                              | 274                   | 254                   |
| 8 d<pCAGTCCCT><br>+<br>d<pCCTTCGGT> | G:C:A:T<br>+<br>G:C:G:C desplaçada | 5'-pur-pur-3'<br>+<br>5'-pyr-pyr-3' | TC                              | hipercròmica                              | 273                   | 240                   |

\* només s'observa el *bi-loop* a concentracions molt elevades mitjançant RMN

Taula 2.4. Resum de les característiques estructurals i espectroscòpiques de les seqüències estudiades anteriorment que s'estructuren en forma de *bi-loop* (10 mM Na<sub>2</sub>PIPES pH 7, 100 mM NaCl, 10 mM MgCl<sub>2</sub>)

Inicialment es va pensar que la hipercromicitat del *bi-loop* respecte a la forma desestructurada, responsable de la fusió hipocròmica observada, era un tret característic d'aquests dímers. En aquesta corba de fusió "anòmala" s'observa una disminució de l'absorbància al passar de la forma *bi-loop* a *dumbbell* i un augment en la transició de *dumbbell* a monòmer desestructurat o *random coil*; o senzillament una disminució de l'absorbància amb la temperatura a causa de la transició de *bi-loop* a *random coil*, si no és possible la formació del *dumbbell* (Figura 2.9). Aquesta generalització es va qüestionar en observar les fusions hipercròmiques ("normals") en els casos de les seqüències d<pCCGTCCGT> i d<pCGCTCCGT>, seqüències que presenten un apilament de bases diferent al de la resta, que és en definitiva el màxim responsable de l'efecte que s'observa durant la fusió.

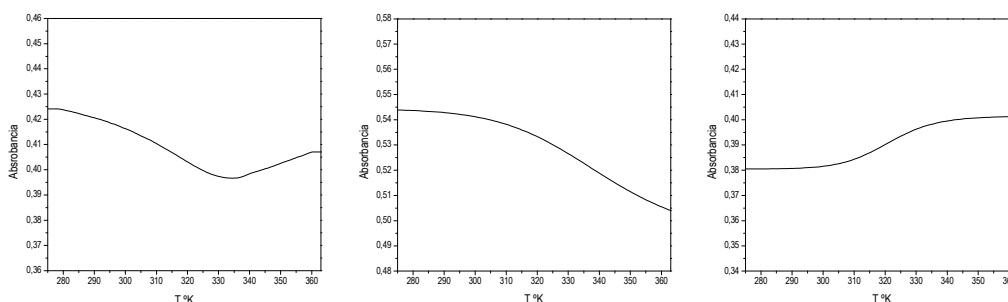


Figura 2.9. Diferents fusions observades del *bi-loop*<sup>110</sup>: *d*<pTGCATGCA>, fusió anòmla amb formació de dumbbell, *d*<pCGCTCATT>, fusió anòmla sense formació de dumbbell i *d*<pCCGTCCGT>, fusió normal

Podem concloure, doncs, que en el cas de tenir un ordre de les bases 5'-pur-pyr-3', que dona lloc a tètredes directes, el *bi-loop* és hipercròmic respecte a l'espècie desestructurada, contràriament al que succeeix en les dobles hèlixs de DNA. Per contra, quan l'ordre d'apilament és 5'-pyr-pur-3' (tètreda desplaçada) o mixta (5'-pur-pyr-3'+5'-pyr-pur-3'), la fusió és normal.

Com que l'estructura es manté només per l'apilament de dues tètredes, els canvis d'absorbància amb la temperatura són petits (sobre un 5% i fins a un 10%), i per això no se'n pot extreure una temperatura de fusió gaire fiable. Per aquest tipus de mostres és molt més fiable la temperatura de fusió determinada a partir de corbes obtingudes per dicromisme circular.

El mateix comportament que s'observa en les corbes de fusió per ultraviolat es dona en els espectres de dicromisme circular. El patró de desplaçament batocròmic de  $\lambda_{\max}$  i hipsocròmic de  $\lambda_{\min}$  si comparem l'espectre en medi salí respecte l'espectre en aigua, es compleix per la majoria de les seqüències excepte per les que tenen canviat l'ordre d'apilament i donen fusions normals: *d*<pCCGTCCGT> i *d*<pCGCTCCGT>, indicant una vegada més que l'apilament de les bases, màxim responsable de la hipocromicitat i dels espectres de dicromisme circular, és marcadament diferent (Figura 2.10).

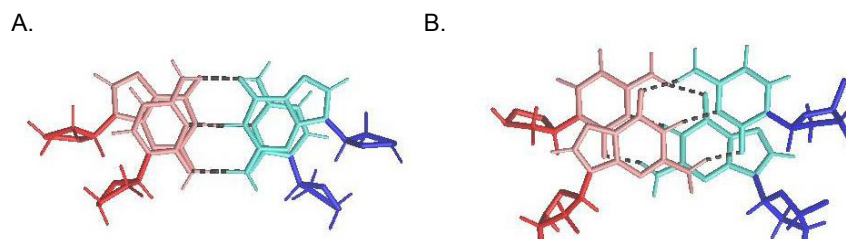


Figura 2.10. Detall de l'apilament dels parells centrals en les estructures de:  
A. *d*<pTGCTCGCT> i B. *d*<pCCGTCCGT>

Sembla, doncs, que podem concloure que les seqüències en què l'ordre d'apilament és 5'-pur-pyr-3' donen lloc a fusions anòmlas per UV i desplaçament batocròmic de  $\lambda_{\max}$  de l'espectre de DC a valors al voltant de 290 nm. En canvi, en el cas de l'apilament 5'-pyr-pur-3', la fusió per UV és normal i  $\lambda_{\max}$  es desplaça a 265 nm (Figura 2.11). En el cas de tenir altres apilaments (5'-pur-pyr-3'+5'-pyr-pur-3' o 5'-pur-pur-3'+5'-pyr-pyr-3', com en l'heterodímer) les fusions també són normals i l'espectre de dicromisme circular presenta un màxim al voltant de 275 nm.



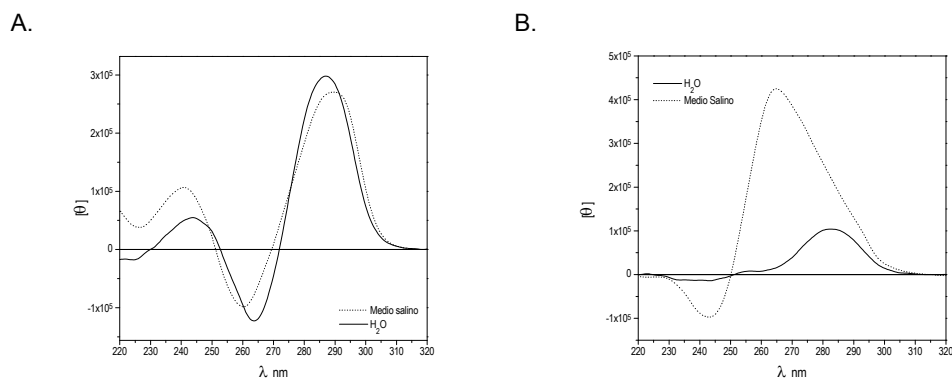


Figura 2.11. Espectres de dicroïsmo circular<sup>110</sup> de A. d<pTGCTCGCT> (espectre en H<sub>2</sub>O correspon a la forma *dumbbell*) i B. d<pCCGTCCGT> (espectre en H<sub>2</sub>O correspon a la forma desestructurada)

### 2.4.2 Ressonància Magnètica Nuclear

Els espectres de ressonància magnètica nuclear dels oligonucleòtids que s'estructuren en forma de *bi-loop* també presenten alguns trets característics.

Primer de tot, cal remarcar que els espectres monodimensionals de protó presenten una forta dependència amb la concentració d'oligonucleòtid, concentració de sals i temperatura. A baixes concentracions d'oligonucleòtid, observem majoritàriament l'oligonucleòtid monòmer, estructurat en forma de *dumbbell* o desestructurat, segons la seqüència, mentre que a altes concentracions es forma principalment el motiu dimèric. Ambdues espècies es troben en intercanvi lent en l'escala de temps d'RMN i, per tant, se'n poden observar les ressonàncies simultàniament.

Pel que fa als protons bescanviables, cal remarcar que estan ben protegits perquè donen senyals estrets en un ampli rang de temperatures. A més a més, es poden seguir els contactes intramoleculars típics dels parells WC. Quan es formen ponts d'hidrogen addicionals, s'observa que els dos protons amino de la guanina es desplacen a camps més baixos.

A partir del nombre de senyals de protons aromàtics, podem veure que es tracta d'estructures dimèriques simètriques, ja que tot i ser octàmers, només presenten 4 sistemes de spin (en el cas de d<pCATTATT> i d<pCCGTCCGT> o 5 en el cas de d<pTGCTCGCT>), o 8 sistemes en els cas de les seqüències no simètriques: d<pCGCTCATT> i d<pCGCTCCGT>. El fet de tenir aquest elevat grau de simetria facilita l'assignació seqüencial perquè simplifica l'espectre, però dificulta l'assignació específica a causa de l'ambigüitat intrínseca per determinar quins contactes són intramoleculars i quins intermoleculars.

En el cas de la tètrada GCGC desplaçada i la tètrada GCCG, es veuen alguns contactes clau entre els protons bescanviables d'una base i els protons del sucre del nucleòtid enfrontat amb ella de l'altra cadena. Per exemple, en el cas de la tètrada desplaçada, es veu el contacte H22 de la guanina amb l'H1' de la guanina d'enfront, que per simetria és ella mateixa, però que no pot correspondre de cap manera a un contacte intraresidu. Aquest és un contacte especialment important ja que només pot ser

intermolecular i indica que els parells de bases s'enfronten pel solc menor.

Els pics de correlació NOE H1'-H<sub>Ar</sub> intern són poc intensos, indicant una disposició *anti* de la base respecte al sucre.

Els contactes seqüencials típics d'hèlix dextrogires (veure fonaments, Figura 1.14, pàg. 28) només s'observen en alguns passos. L'apilament dels residus aparellats (2-3/6-7 si comencem a numerar per la segona base del *loop*) dóna lloc als contactes típics B<sub>n</sub>H<sub>1</sub>-B<sub>n+1</sub>H<sub>Ar</sub> i B<sub>n</sub>H<sub>2</sub>/H<sub>2'</sub>-B<sub>n+1</sub>H<sub>Ar</sub>, però l'apilament de la primera base del *loop* (contactes 3-4/7-8) només dóna lloc als contactes del protó aromàtic amb H1'. Entre els residus 4-5/8-1 (primera i segona base del *loop*) i entre els residus 5-6/1-2, no s'observen contactes seqüencials típics, però sí alguns contactes entre protons del sucre.

La primera base del *loop* es troba apilada sobre els parells de bases donant lloc a contactes seqüencials B<sub>n</sub>H<sub>1</sub>-B<sub>n+1</sub>H<sub>Ar</sub>, com ja s'ha comentat, a contactes intersubunitat entre els Me de T4/T8 amb els protons aromàtics i bescanviables del parell de bases sobre el qual s'apila (i en alguns casos també amb H1') i a un desplaçament anòmal d'H4', H5' i H5'' de la primera T del *loop*. Els desplaçaments químics d'aquests protons estan influïts per l'anell aromàtic de la nucleobase del residu que es troba dues posicions més enllà en el sentit 5'-3', sobre la qual s'apila l'anell de ribosa (Figura 2.12). Quan aquesta base és purina, els desplaçaments químics estan desplaçats a camps anormalment alts, entre 0.5 i 1 ppm, mentre que quan és pirimidina no ho estan tant, entre 0.2 i 0.4 ppm respecte el valor teòric (3.8-4.3 ppm).

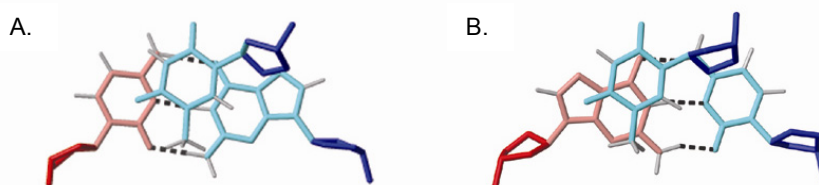


Figura 2.12. Detall de l'apilament de la primera T del *loop* sobre el parell G·C en el cas de: A. d<pTGCTCGCT> i B. d<pCCGTCCGT>

## 2.5 Objectius

Així doncs, i tal com s'ha comentat a la introducció general, l'objectiu d'aquest bloc és prosseguir amb l'estudi del motiu *bi-loop* en solució, per poder precisar els requeriments seqüencials per a la seva formació i intentar observar-lo en solució amb oligonucleòtids lineals.

Aquests objectius generals s'aborden en els propers capítols de la següent manera:

- Estudi de la formació del motiu entre un oligonucleòtid cíclic i un altre lineal, que hem anomenat *bi-loop* heterolineal (capítol 3).
- Estudi de la formació del motiu amb parells no canònics, en aquest cas amb parells G·T (capítol 4).
- Estudi de la formació del motiu en solució amb oligonucleòtids lineals, fins ara no observat, per tal de reforçar la seva potencial rellevància biològica (capítol 5).

En aquest bloc s'hi inclou un capítol més, en què s'intentarà recopilar tota la informació de què es disposa per determinar fins a quin punt el motiu *bi-loop* podria ser general i rellevant biològicament (capítol 6).

### 3 EL MOTIU BI-LOOP HETEROLINEAL

Un dels objectius d'aquesta tesi era estudiar la formació del motiu *bi-loop* amb oligonucleòtids lineals i vam iniciar l'estudi partint de l'heterodímer anteriorment dissenyat amb els *bi-loops* que anomenem heterolineals. Entenem com a *bi-loop* heterolineal aquell que està format per dos oligonucleòtids diferents, un dels quals és cíclic i l'altre lineal. L'interès d'aquestes estructures és divers:

- L'observació de la formació del motiu amb oligonucleòtids lineals en solució.
- La possibilitat d'utilitzar oligonucleòtids cíclics per dirigir-los a seqüències lineals mitjançant la formació d'aquest motiu.
- La possibilitat d'utilitzar aquest motiu per desenvolupar una metodologia de ciclació assistida per motlle (veure bloc 2, apartat 4.1).

#### 3.1 EL *bi-Loop* heterodimèric

Per aprofundir en l'estudi dels requeriments per a la formació del motiu *bi-loop*, s'havia investigat si es podia formar el motiu per interacció de dos cicles diferents. Amb aquesta finalitat s'havien dissenyat dues parelles de cicles que, sobre el paper, podien formar el motiu i, en canvi, no podien autoestructurar-se ni per formació d'un *bi-loop* homodimèric ni d'un *dumbbell*, evitant així la presència de diverses espècies en solució, fet que en dificultaria l'estudi. Les seqüències escollides havien estat: d<pCAGTCCCT> + d<pCCTTCGGT> i d<pCGTATGTT> + d<pCACATACT> (Figura 3.1).

Aquestes parelles d'oligonucleòtids van ser estudiades i es va poder determinar mitjançant RMN que la primera d'elles s'estructurava en forma d'un *bi-loop* heterodimèric<sup>108b</sup>, mentre que en el cas de la segona, no se'n va poder resoldre l'estructura donada la presència de diverses espècies en solució, com evidenciava el fet de tenir nous senyals de protons imínics.

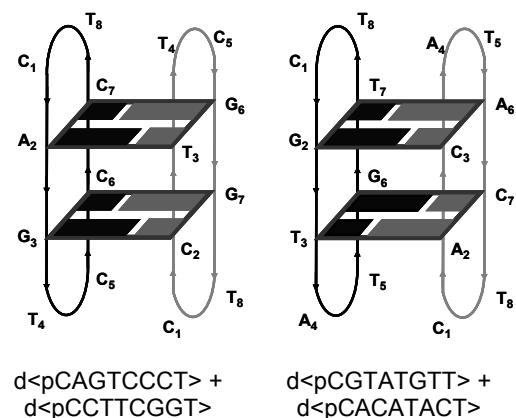
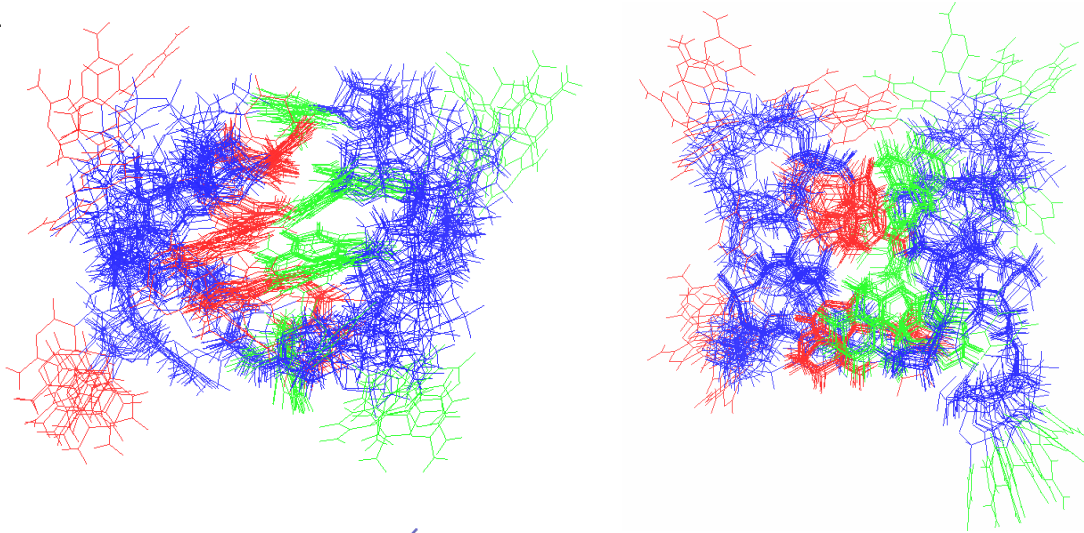


Figura 3.1. Esquemes dels possibles *bi-loops* heterodimèrics

A continuació es presenta l'estructura obtinguda per l'heterodímer d<pCAGTCCCT> + d<pCCTTCGGT> mitjançant RMN i dinàmica molecular restringida.

A.



B.

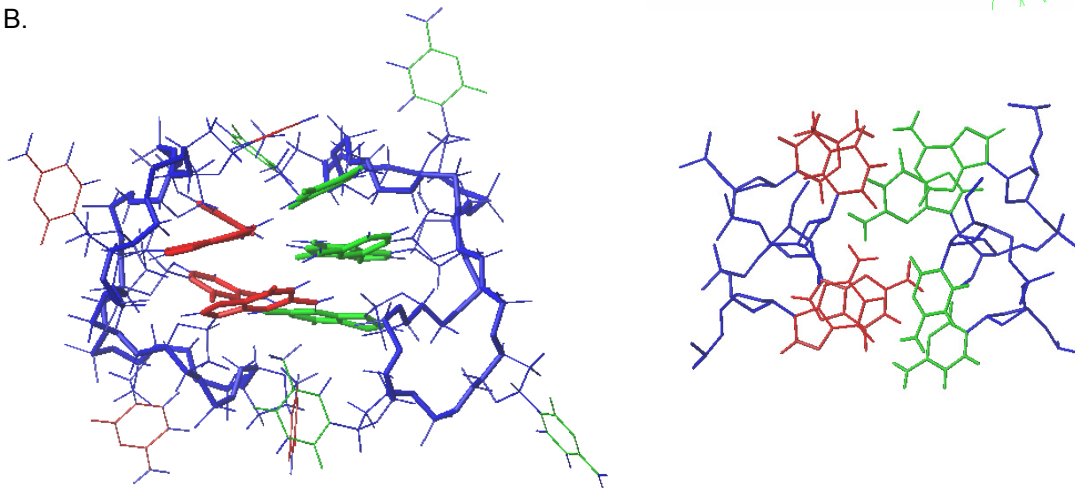


Figura 3.2. Estructura de l'heterodímer d<pCAGTCCCT> + d<pCCTTCGGT>. A. Vista lateral i superior de les 10 estructures superposades. B. Estructura promig del complex i vista superior de l'apilament

Cal remarcar diverses diferències d'aquesta estructura en comparació amb les estudiades anteriorment. En aquest cas, al tractar-se d'un heterodímer, en l'espectre d'RMN s'observaven 16 sistemes de spin i un gran solapament de senyals. Tot i així, es va poder realitzar una assignació quasi completa, però el nombre de restriccions experimentals que se'n van poder extreure va ser petit. En aquest cas l'avantatge és que no existeix el problema de l'ambigüitat de contactes intra i intermoleculars. A causa de les poques restriccions experimentals, les estructures estan pobrament definides. Les timines en les primeres posicions dels *loops* estan especialment desordenades. Aquesta manca de definició no és deguda a una major flexibilitat, sinó al baix nombre de restriccions. Pel que fa a l'estructura en si, val a dir que s'ha pogut demostrar que es pot formar entre oligonucleòtids diferents, i que és el primer cas en què es troba estabilitzada per dues tètades diferents (una GCAT i una GCGC desplaçada) i també l'únic cas en què els apilaments són 5'-pur-pur-3' i 5'-pyr-pyr-3'.

### 3.2 Elecció de Les seqüències

Com ja s'ha dit, per a l'estudi del motiu entre un oligonucleòtid cíclic i un altre lineal es va prendre com a model el *bi-loop* heterodimèric d<pCAGTCCCT> (cicle A) + d<pCCTTCGGT> (cicle B).

Partint de l'estructura de l'heterodímer, havíem de decidir per on tallàvem els cicles per tenir possibles oligonucleòtids lineals formadors del motiu *bi-loop*, evitant, per una banda, que els oligonucleòtids lineals es puguin autoestructurar formant forquetes (*hairpins*) o segments de doble cadena, i, per l'altra, considerant la possibilitat de tallar per un dels llaços o tallar "pel motiu" (pel mig del dinucleòtid que forma els parells de bases). Respecte a l'estabilitat de l'estructura obtinguda és difícil de predir si seria més favorable l'estructura formada per un oligonucleòtid cíclic i un lineal tallat pel motiu o pel llaç. Respecte a la posterior possibilitat de ciclació per unió dels extrems propers en l'espai, creiem que tallant pel motiu la proximitat dels extrems és major, però alhora potser la conformació és massa restringida i en dificulta la ciclació.

Considerant els aspectes mencionats, es van escollir tres seqüències lineals a assajar: una seqüència tallada pel motiu i dues tallades per cadascun dels *loops* en cicles diferents (Figura 3.3), que podrien eventualment estructurar-se formant un *bi-loop* heterodimèric entre dos oligonucleòtids lineals.

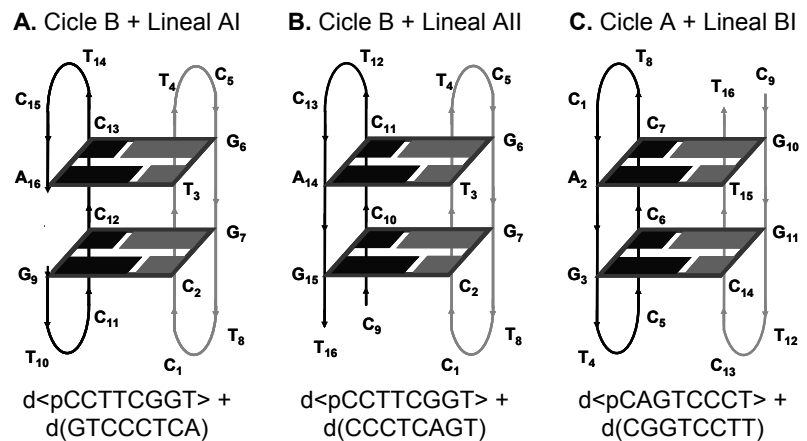


Figura 3.3. Seqüències lineals escollides i esquemes dels possibles *bi-loops*

Posteriorment, basant-nos amb resultats que s'explicaran més endavant, es va decidir assajar també unes altres seqüències, que es mostren en la Figura 3.4.

En les Figures 3.3 i 3.4, a part de mostrar les seqüències i els esquemes dels possibles *bi-loops*, també s'hi inclou la nomenclatura que s'emprarà per designar les diferents seqüències: lineal Ax per als oligonucleòtids lineals derivats del cicle A, i lineal Bx, per als derivats del cicle B.

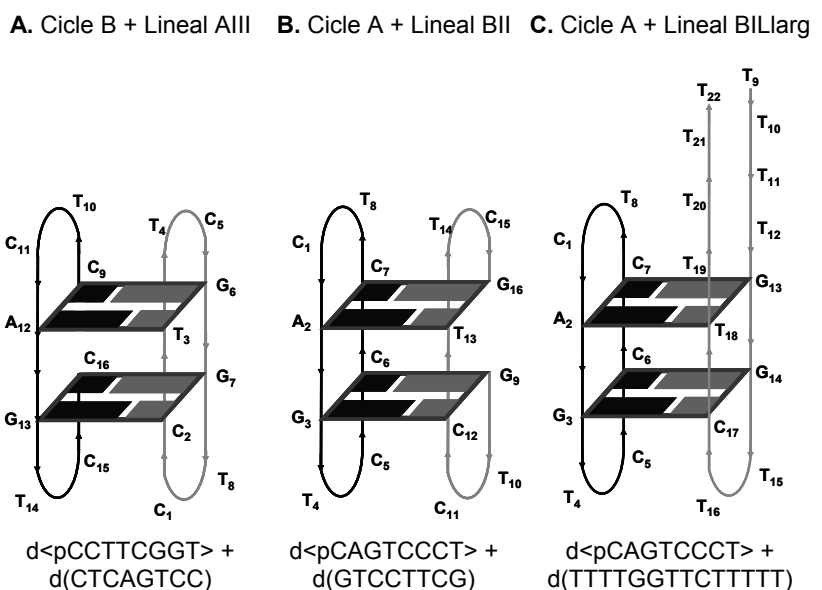


Figura 3.4. Altres seqüències escollides i els seus possibles *bi-loops*

### 3.3 Síntesi dels oligonucleòtids

#### 3.3.1 Síntesi d'oligonucleòtids cíclics

Els oligonucleòtids cíclics d<pCAGTCCCT> i d<pCCTTCGGT> es van sintetitzar seguint la metodologia desenvolupada al nostre grup<sup>57</sup>, emprant el sintetitzador ABI 380B i la nucleotidil-resina preparada de funcionalització de 130 µmol/g (veure part experimental, apartat 3.1). Es va realitzar una primera síntesi a escala 2 µmol i, posteriorment, una segona síntesi a escala 10 µmol per tenir suficient oligonucleòtid pels estudis per espectroscòpia d'UV, CD, RMN i ciclació assistida per motlle.

##### 3.3.1.1 Preparació de la nucleotidil-resina

La unió del primer nucleòtid a la fase sòlida es realitza a través del grup fosfat en 3', que s'ancora a la resina a través d'un *linker* bifuncional, l'àcid 3-cloro-4-hidroxifenilacètic, el qual actua alhora de protector permanent durant tota la síntesi (protector típic emprat en el mètode del fosfat triester). Aquest *linker* bifuncional no s'uneix directament a la resina sinó que s'hi intercala un espaiador, l'àcid 6-aminohexanoic, amb el propòsit d'augmentar la flexibilitat i separar l'oligonucleòtid de la resina facilitant així les etapes de ciclació i desancoratge<sup>110</sup>. A més a més, la incorporació d'aquest espaiador permet controlar i disminuir, si cal, el grau de funcionalització de la resina, que es determina a partir de la quantificació de la *N*-(9-fluorenilmetil)piperidina alliberada en l'eliminació del grup Fmoc de l'espaiador.

La preparació de la nucleotidil-resina es realitza derivatitzant el suport polimèric amb l'espaiador, per una banda, preparant el nucleotidil-*linker* en forma d'ester actiu, per l'altra, i acoblant-los posteriorment (Figura 3.5). L'ester serveix de protector del grup àcid del *linker* i alhora l'activa per a l'acoblament sobre l'espaiador-resina. El nucleotidil-*linker* s'utilitza directament ja que, atesa l'elevada reactivitat de l'ester de triclorofenil, no és possible purificar-lo per cromatografia en gel de sílice perquè hi queda irreversiblement unit. La caracterització per <sup>31</sup>P-RMN mostra que la impuresa més important és l'hidrogenfosfonat format per hidròlisi del fosforamidit, impuresa que no interfereix en l'acoblament perquè no reacciona amb els grups amino de la resina.

La funcionalització després de l'ancoratge del nucleotidil-*linker* es determina per quantificació dels cations DMT alliberats de la desprotecció amb àcid d'una petita alíquota de la resina i ha d'estar entre 100 i 150 µmol/g (resines de funcionalització menor donen resultats equiparables de rendiments i puresa<sup>132</sup> però dificulten l'aplicació en síntesis a gran escala). La puresa de la resina es comprova per <sup>31</sup>P-RMN en fase gel. Aleshores es desprotegeix el grup fosfat en 3', que es manté desprotegit durant tota l'elongació de la cadena perquè posteriorment pugui reaccionar en l'etapa de ciclació. D'aquesta manera s'evita l'elecció d'un protector semipermanent per a aquest grup que sigui ortogonal al de la resta de fosfats i les reaccions secundàries que es podrien produir en la seva eliminació.

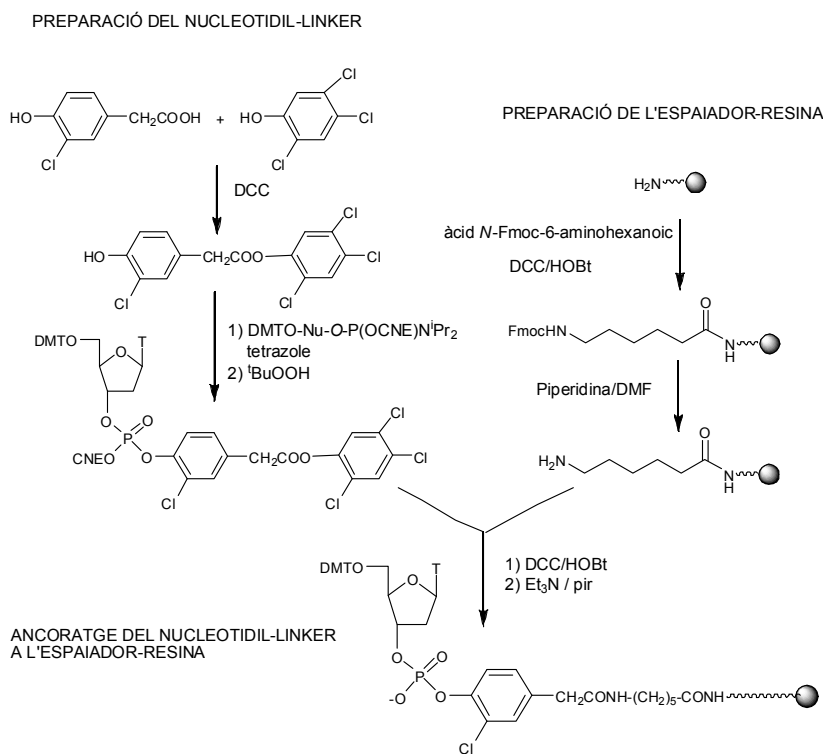


Figura 3.5. Esquema sintètic per a l'obtenció de la nucleotidil-resina

### 3.3.1.2 Elongació de la cadena oligonucleotídica

L'elongació de la cadena té lloc al sintetitzador automàtic emprant 3'-(2-cianoetil)-*N,N*-diisopropilfosforamidits de 5'-*O*-(4,4'-dimetoxitritil)-desoxiribonucleòsids i emprant un cicle de síntesi estàndard lleugerament modificat, amb temps de reacció allargats i la incorporació de rentats amb DMF (veure part experimental 3.2 i cicles de síntesi a l'annex). Els suports polimèrics que s'han emprat amb aquest mètode de síntesi han estat tant el copolímer de polietilenglicol-poliestirè (TentaGel-NH<sub>2</sub>) com l'aminometil-poliestirè-co-1 %-divinilbenzè (PS-NH<sub>2</sub>), ja que tots dos són compatibles amb el mètode del fosfit triester i al no ser rígids faciliten la ciclació; però s'obtenen millors rendiments en el cas de la resina TentaGel, que ha estat l'emprada en aquest treball.

Els grups 4,4'-dimetoxitritil alliberats en cadascuna de les etapes són col·lectats en tubs d'assaigs i a partir de la seva quantificació es pot calcular el rendiment de síntesi del precursor lineal i el rendiment promig per etapa (Materials i Mètodes, apartat 1.3.2). Els rendiments promig per etapa han estat de l'ordre del 96%, amb rendiments globals de precursor lineal sobre el 80%.

Per a la síntesi d'oligonucleòtids cíclics cal triar una seqüència lineal a sintetitzar que serà la que posteriorment es ciclarà. Els criteris que s'utilitzen per realitzar aquesta elecció es basen en el nucleòtid a ancorar a la resina, que normalment és timidina, i el nucleòtid de l'extrem 5' terminal que en la ciclació s'unirà al 3' terminal i que acostuma a ser una pirimidina, ja que funciona millor la condensació pirimidina sobre pirimidina. En base a aquests criteris per a la síntesi de les seqüències cícliques dissenyades es van triar les següents seqüències lineals: 5'-dCAGTCCCT-3' i 5'-dCCTTCGGT-3'.



### 3.3.1.3 Ciclació

La ciclació és l'etapa clau en l'esquema de síntesi d'oligonucleòtids cíclics. En aquest cas, el reactiu emprat per unir els extrems 5'-OH i 3'-fosfat és l'1-mesitilen-sulfonyl-3-nitro-1,2,4-triazole (MSNT), reactiu d'acoblament típic emprat en el mètode del fosfat triester. Aquesta reacció es força mitjançant la realització de tres tractaments en què es renova l'MSNT (4+4+12h). El mecanisme és el mostrat en la Figura 3.6.

Aquest reactiu és molt sensible a la humitat i per això és molt important comprovar que es troba en bon estat realitzant-ne el punt de fusió i extreure les precaucions per treballar en condicions molt anhidres, ja que sinó el reactiu es podria hidrolitzar i, a més a més, la petita proporció d'aigua podria competir amb l'OH en 5' en la reacció amb el fosfat activat.

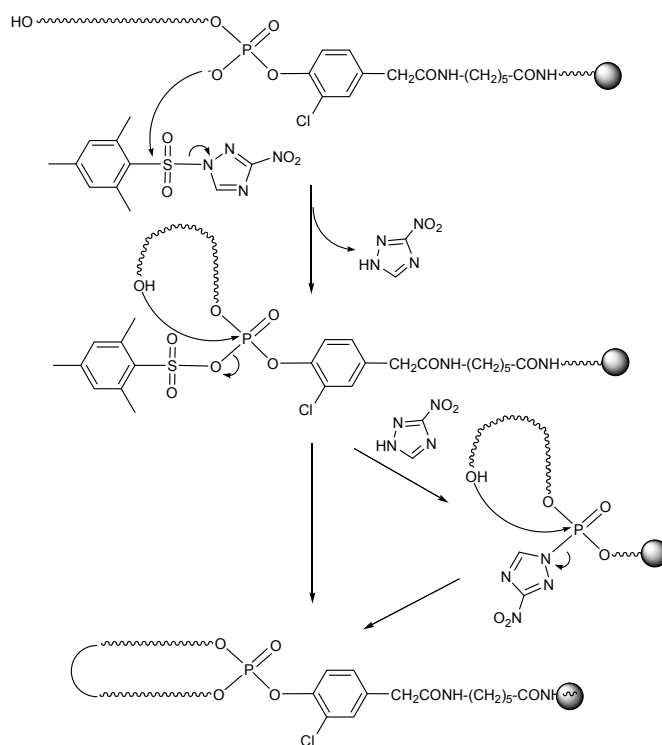


Figura 3.6. Mecanisme de ciclació amb MSNT en piridina

### 3.3.1.4 Desprotecció, desanclatge i dessalatge

Una vegada tenim l'oligonucleòtid ciclat es procedeix a realitzar les etapes finals de desprotecció i desanclatge. En primer lloc es realitza la desprotecció dels grups fosfat internucleosídics per  $\beta$ -eliminació dels grups 2-cianoetil amb trietilamina/piridina, de manera que queden com a fosfat diester.

La separació de l'oligonucleòtid cíclic de la resina té lloc mitjançant un atac nucleòfil sobre el fosfat en 3' amb *syn*-piridin-2-aldoximat de tetrametilguanidini (Figura 3.7), que trenca selectivament enllaços fosfat triester<sup>133</sup> (és el reactiu emprat en la desprotecció dels fosfats internucleosídics en el mètode del fosfat triester).

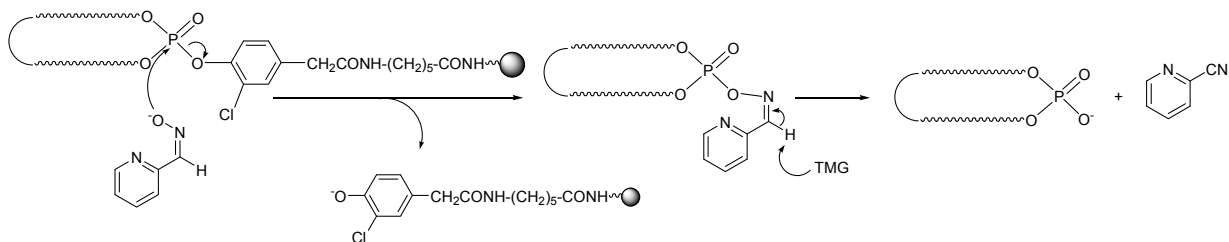


Figura 3.7. Mecanisme de desanclatge amb *syn*-piridin-2-aldoximat de tetrametilguanidini

El precursor lineal i els possibles productes de reacció intercadena estan units a la resina a través d'un grup fosfat diester i la seva cinètica de desancoratge és molt més lenta que la dels fosfats triester, de forma que existeix una bona discriminació i les cadenes no ciclades romanen unides a la resina.

Per últim, el tractament amb amoníac aquós permet desprotegir les nucleobases del producte ja desancorat. El producte desancorat i desprotegit es sotmet a una cromatografia d'exclusió molecular sobre Sephadex G-10 per tal d'eliminar l'oximat en excés i altres productes de baix pes molecular. Gràcies al desancoratge selectiu i a aquesta filtració molecular els crus obtinguts són de gran puresa.

### 3.3.1.5 Resultats, anàlisi, purificació i caracterització

Els productes obtinguts es quantifiquen per espectroscòpia d'ultraviolat, se'n comprova la massa per MS-MALDI-TOF i s'analitzen per cromatografia líquida d'alta eficàcia (HPLC) sobre una fase estacionària inversa d'octadecil sílica, obtenint en tots els casos perfils semblants i d'elevada puresa (Figura 3.8). A la Taula 3.1 es resumeixen els resultats de les síntesis.

Els rendiments obtinguts (14 a 20 %) no són tan bons com els que havien estat descrits per cicles d'aquesta mida (30 a 40 %) <sup>134</sup> i pensem que es deu a què el reactiu de ciclació no es trobava en bon estat o no era prou pur, tot i que el seu punt de fusió era prou correcte ( $P_{f \text{ exp.}} = 136.8\text{-}138.3 \text{ }^\circ\text{C}$ ;  $P_{f \text{ ref}} = 137.7\text{-}138.3 \text{ }^\circ\text{C}$  <sup>134</sup> i  $135.5\text{-}137 \text{ }^\circ\text{C}$  <sup>72b</sup>).

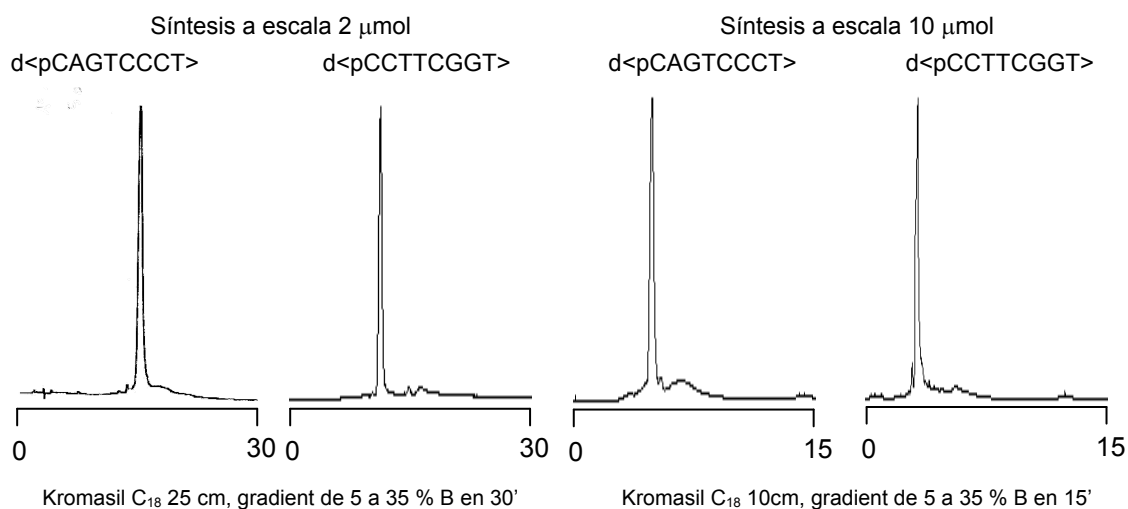


Figura 3.8. Anàlisi per HPLC dels crus de síntesi dels oligonucleòtids cíclics

Els productes es van purificar per HPLC a escala analítica o semipreparativa segons la quantitat de cru. Els rendiments de purificació de les primeres síntesis realitzades a escala 2 µmol van ser força baixos, potser perquè la purificació es va realitzar per HPLC a escala analítica i es van requerir moltes injeccions. Les purificacions de les síntesis a escala 10 µmol van tenir lloc amb bons rendiments. En aquest cas es van emprar condicions semipreparatives i sembla que la columna PRP-I (rebliment de poli(estirè-divinilbenzè), 305 × 7 mm) dóna millors recuperacions que la X-Terra (C<sub>18</sub>, 50 × 10 mm).

| Seqüència          | Síntesi a escala 2 µmol |                       | Síntesi a escala 10 µmol |                         |
|--------------------|-------------------------|-----------------------|--------------------------|-------------------------|
|                    | d<pCAGTCCCT>            | d<pCCTTCGGT>          | d<pCAGTCCCT>             | d<pCCTTCGGT>            |
| Rend. cru          | <b>14 %</b> (21.4 OD)   | <b>17 %</b> (25.7 OD) | <b>19 %</b> (~141 OD)    | <b>15 %</b> (~108 OD)   |
| Puresa cru (HPLC)  | 95 %                    | 92 %                  | 72 %                     | 81 %                    |
| Rend. purificació  | 20 %                    | 17 %                  | 69 % (X-Terra)           | 90 % (PRP)              |
| Rend. producte pur | <b>2.7 %</b> (4.3 OD)   | <b>2.9 %</b> (4.4 OD) | <b>13.3 %</b> (96.8 OD)  | <b>13.9 %</b> (98.4 OD) |

Taula 3.1. Resum de les síntesis realitzades dels oligonucleòtids cíclics

Els productes purificats es van caracteritzar per MS-MALDI-TOF i per digestions enzimàtiques amb fosfodiesterasa de verí de serp (SVPD) i fosfatasa alcalina (AP) (Materials i Mètodes, 1.3.7), confirmant que teníem els productes desitjats.

### 3.3.2 Síntesi d'oligonucleòtids lineals

Les seqüències lineals inicialment escollides (d(GTCCCTCA), d(CCCTCAGT) i d(CGGTCCTT)) es van sintetitzar segons la metodologia estàndard al sintetitzador automàtic ABI380B. Es van realitzar dues síntesis de cada seqüència a escala 1 µmol (2x1 µmol). Els crus obtinguts van ser de gran puresa (veure com a exemple la Figura 3.9) i els rendiments de síntesi, força bons, en el rang dels esperats en la síntesi d'oligonucleòtids amb el mètode del fosforamidit (veure la Taula 3.2). Els productes purificats mitjançant HPLC es van caracteritzar per MS-MALDI-TOF, confirmant que eren els desitjats.

| Seqüència   | Síntesi        | Rend. etapa (%) | Rend. total cru (%) | Rend. prod. pur (%) |
|-------------|----------------|-----------------|---------------------|---------------------|
| d(GTCCCTCA) | 1 <sup>a</sup> | 94.4            | 66.7                | 10.8 (7.9 OD)       |
|             | 2 <sup>a</sup> | 92.4            | 57.4                | 13.2 (9.7 OD)       |
| d(CCCTCAGT) | 1 <sup>a</sup> | 97.4            | 81.5                | 20.6 (15.1 OD)      |
|             | 2 <sup>a</sup> | 97.5            | 84.2                | 35.6 (26.1 OD)      |
| d(CGGTCCTT) | 1 <sup>a</sup> | 96.8            | 81.5                | 22.5 (17.4 OD)      |
|             | 2 <sup>a</sup> | 96.6            | 82.4                | 39.7 (29.4 OD)      |

Taula 3.2. Dades de les síntesis d'oligonucleòtids lineals a escala 1 µmol

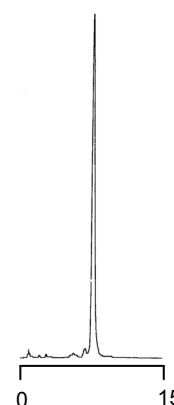


Figura 3.9. Cromatograma del cru de d(CCCTCAGT)

De l'oligonucleòtid d(GTCCCTCA) es va realitzar posteriorment una altra síntesi (2x1 µmol) per als estudis mitjançant RMN al sintetitzador Expedite que va tenir lloc amb rendiments moderats (rend. cru=59%, que correspon a un rendiment per etapa del 93%, rend. total=31 %).

Posteriorment, es van sintetitzar els oligonucleòtids dissenyats en una segona etapa: d(CTCAGTCC), d(GTCCTTCG) i d(TTTTGGTTCTTTT) en el sintetitzador automàtic Expedite. Es van realitzar diverses síntesis segons el cas per obtenir suficient producte per als estudis mitjançant RMN. En la Taula 3.3 es resumeixen els resultats d'aquestes síntesis que en general van ser bons ja que els crus eren de gran puresa (>90 %). Tot i així, es van purificar mitjançant MPLC i posteriorment dessalat per cromatografia d'exclusió molecular ja que per als estudis mitjançant RMN es requereix la màxima puresa. Els productes es van caracteritzar per MS-MALDI-TOF o MS-ESI.

| Seqüència        | Síntesi  | Rend. total cru (%) | Rend. prod. pur (%) |
|------------------|----------|---------------------|---------------------|
| d(CTCAGTCC)      | 3x1 µmol | --                  | 19 % (35 OD)        |
|                  | 2x1 µmol | 56 % (78 OD)        | 25 % (35 OD)        |
| d(GTCCTTCG)      | 2x1 µmol | --                  | 55 % (85 OD)        |
| d(TTTTGGTTCTTTT) | 3x1 µmol | 63 % (224 OD)       | 33 % (116 OD)       |

Taula 3.3. Resum de les síntesis realitzades dels nous oligonucleòtids lineals

### 3.4 Estudis mitjançant espectroscòpia UV i dicroisme circular

Els experiments es van dur a terme per intentar observar la formació d'un motiu *bi-loop* heterodimèric entre una seqüència cíclica i una lineal o entre dues lineals. Per això, es van estudiar totes les possibles combinacions de les seqüències oligonucleotídiques inicialment sintetitzades que es podrien estructurar en forma de *bi-loop*: A+B en medi aquós i en medi salí, A+BI, Al+B, All+B, Al+BI i All+BI en medi salí. Per cadascuna de les mostres es van realitzar les corbes de fusió per espectroscòpia d'ultraviolat i se'n va enregistrar l'espectre de dicroisme circular a baixa temperatura.

Les condicions experimentals emprades van ser 10 µM de cada oligonucleòtid i un medi salí de 10 mM de Na<sub>2</sub>PIPES, 10 mM de MgCl<sub>2</sub> i 100 mM de NaCl. Aquestes condicions salines han demostrat ser les adequades per a la formació del *bi-loop* estabilitzat per quatre parells de bases G:C<sup>110</sup> ja que no s'observa una estabilització addicional en augmentar més la concentració de sals. Per dímers estabilitzats per parells A:T s'ha demostrat que és millor treballar a concentració 500 mM de MgCl<sub>2</sub> ja que així l'estabilitat del dímer és major. En el present treball s'ha decidit emprar les primeres condicions perquè les possibles estructures dimèriques estan estabilitzades per 3 parells G:C i 1 parell A:T i s'ha considerat que seria suficient, sobretot treballant a una major concentració d'oligonucleòtid: 10 µM de cada espècie enfront de 5 µM de monòmer en els casos de *bi-loops* homodimèrics.

#### 3.4.1 Corbes de fusió per espectroscòpia d'UV

Les corbes de fusió per UV no van donar resultats gaire fiables ni definitius ja que van ser molt poc reproduïbles. En tots els casos, però, s'observen corbes de fusió normals en què l'absorbància augmenta amb la temperatura i de perfil no gaire sigmoide. Si ho comparem amb les dades de què es disposa fins al moment no és d'estranyar que la fusió sigui normal ja que en aquest cas els apilaments de les bases també són diferents (5'-pur-pur-3' i 5'-pyr-pyr-3'). Els pendents de les corbes són molt petits i no presenten plafons definits a baixa i alta temperatura, sinó que la variació d'absorbància, tot i que petita, és bastant contínua. Aquestes fusions no permeten extreure conclusions definitives. En la Figura 3.10 es presenten algunes de les corbes realitzades amb els valors aproximats de T<sub>m</sub> que se'n poden extreure a partir de la primera derivada de la corba.

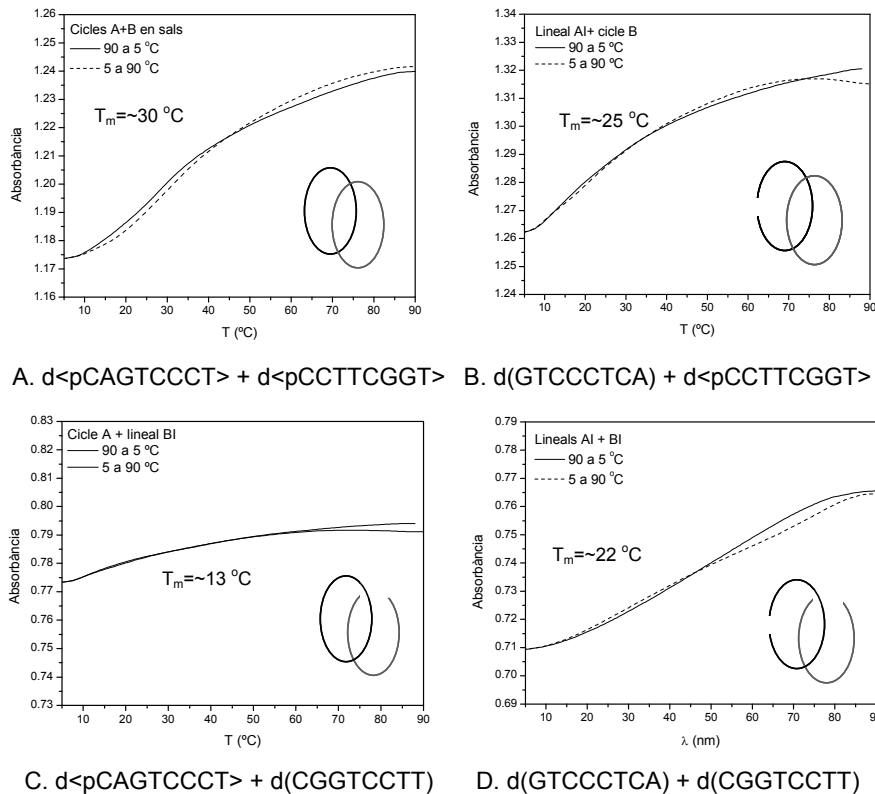


Figura 3.10. Corbes de fusió per espectroscòpia d'ultraviolat

Tot i que són poc concloents, de les corbes de fusió per UV sí que en podem extreure alguna informació ja que s'observa una relació entre  $\Delta Abs$  i la  $T_m$ , indicadors ambdós de l'existència d'una certa estructuració. Així, l'ordre de major a menor estructuració seria: cicle A+B > lineal AI + cicle B > lineals AI+BI > cicle A + lineal BI. Era d'esperar que la mescla dels dos cicles presentés la major estructuració, però no pas que una de les mescles de dos oligonucleòtids lineals (AI+BI) s'estructurés més que la mescla cicle A + lineal BI.

### 3.4.2 Espectres de dicroisme circular

Els espectres que es mostren a continuació es van enregistrar amb l'espectropolarímetre Jasco J-720. A causa de problemes tècnics amb el controlador de temperatura del bany, no es van poder realitzar corbes de fusió. Abans de realitzar els espectres de les mescles, es van enregistrar els espectres dels oligonucleòtids per separat en les condicions salines descrites i a concentració 10  $\mu M$  com a control negatiu de la manca d'estructuració específica i per poder veure si les mescles d'oligonucleòtid presentaven una estructuració addicional a la suma dels espectres control.

La comparació dels espectres de dicroisme circular de la mescla de  $d\langle pCAGTCCCT \rangle$  (cicle A) i  $d\langle pCCTTCGGT \rangle$  (cicle B) en  $H_2O$  i en medi salí mostra canvis significatius en

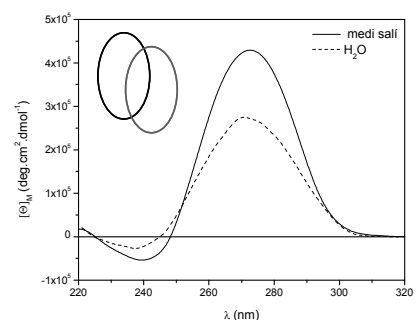


Figura 3.11. Comparació dels espectres de dicroisme circular de la mescla  $d\langle pCAGTCCCT \rangle$  i  $d\langle pCCTTCGGT \rangle$  en  $H_2O$  i en medi salí a 5 °C

l'el·lipticitat però no en el perfil de les corbes (Figura 3.11). Tot i així, l'el·lipticitat observada en H<sub>2</sub>O és anormalment elevada. Pensem que potser la mostra oligonucleotídica no estava ben dessalada i que, sigui com sigui, l'estructuració, tant en aigua com en medi salí, ha de correspondre a la formació d'un *bi-loop* heterodimèric, menys estructurat en el medi aquós. Aquesta és la única mescla que s'ha analitzat en medi aquós ja que l'estructuració en H<sub>2</sub>O és molt més feble que en condicions salines.

També cal destacar que, si bé no es van poder realitzar corbes de fusió per dicroisme circular per problemes tècnics, sí que es van enregistrar els espectres a 80°C, en els quals no s'observa més que l'el·lipticitat residual, corroborant que aquestes estructures fonen amb la temperatura.

De les combinacions d'oligonucleòtids cíclic i lineal, la que presenta una el·lipticitat més elevada que indica clarament existència d'estructuració és la mescla de cicle B i lineal AI: d<pCCTTCGGT> i d(GTCCCTCA) (Figura 3.12.A). Si comparem l'espectre de la mescla AI+B amb el de la mescla dels cicles A+B (Figura 3.12.B) observem que no és del tot coincident, però sí molt semblant. L'el·lipticitat és una mica menor però el perfil és similar, indicant que es pot tractar de la mateixa estructura.

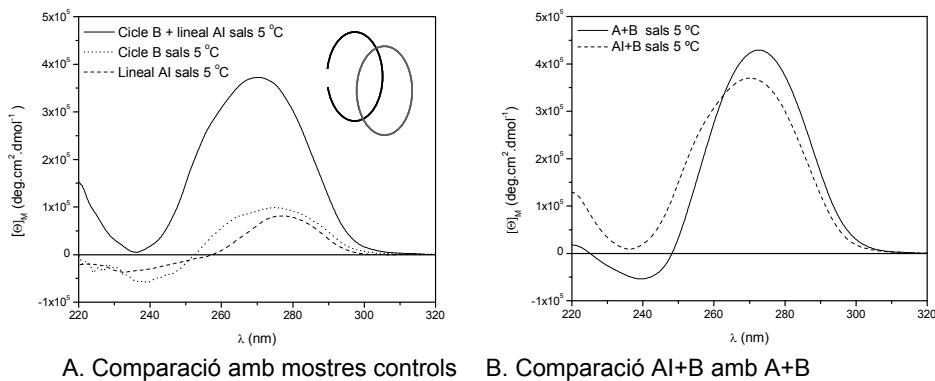


Figura 3.12. Espectres de dicroisme circular de la mescla de d(CAGTCCCT) i d<pCCTTCGGT>

A primera vista, els espectres de les mescles A+BI i AII+B no difereixen significativament de la suma dels espectres control, indicant que no existeix una estructuració especial (Figura 3.13), però si ens hi fixem bé, veiem que la combinació A+BI (Fig. 3.13.B) sí que mostra un desplaçament hipsocròmic del màxim que podria ser indicatiu de la formació del *bi-loop*, encara que poc estabilitzat.

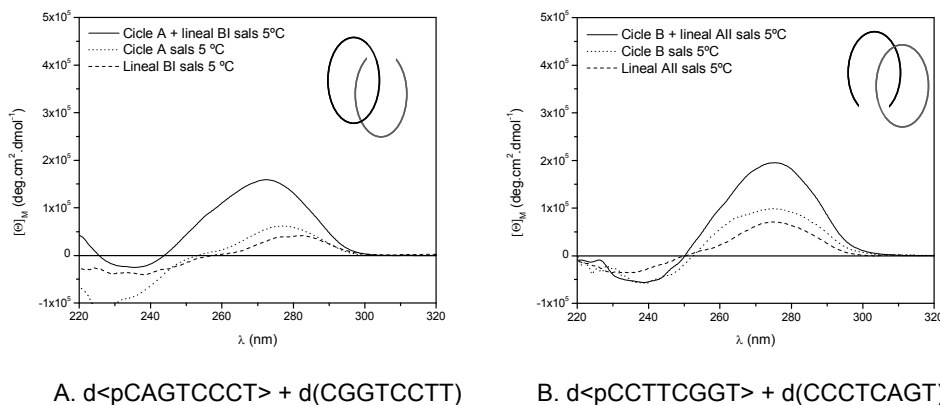


Figura 3.13. Espectres de dicroisme circular de les mescles cicle A+lineal BI (A) i cicle B+lineal AII (B)

De les dues mescles de seqüències lineal + lineal (Figura 3.14), la de d(GTCCCTCA) (lineal AI) + d(CGGCCTT) (lineal BI) presenta un espectre de dicroisme circular amb una el·lipticitat més elevada de l'esperada i un desplaçament de  $\lambda_{m\grave{a}x}$  si la comparem amb els espectres de les mescles control, indicant que existeix una certa estructuració. La mescla AII+BI, en canvi, només presenta una el·lipticitat residual.

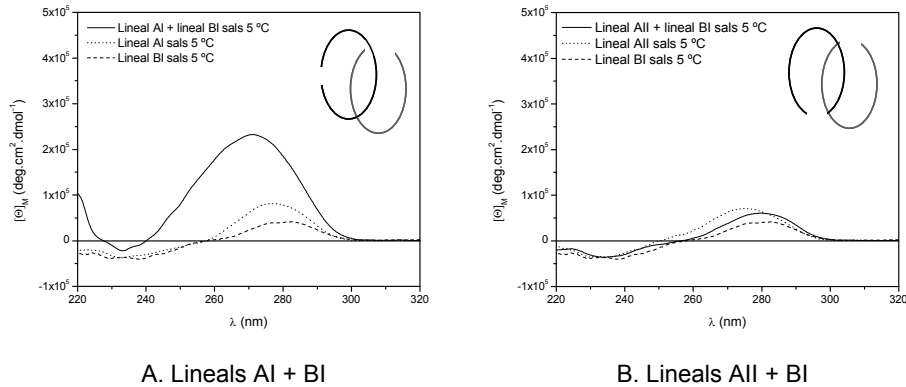
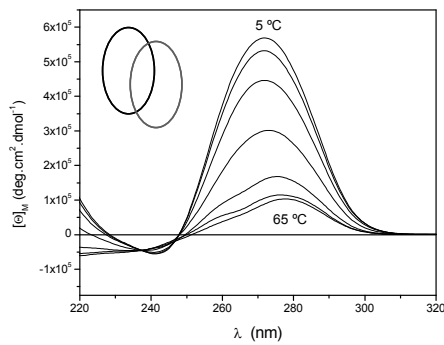


Figura 3.14. Espectres de dicroisme circular de les mescles de seqüències lineals

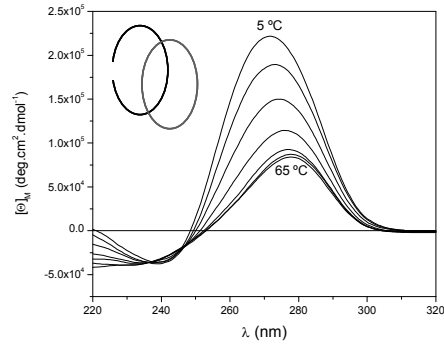
Si comparem els espectres de les mescles AI+B (Figura 3.12.A) i AI+BI (Figura 3.14.A) (la superposició no es mostra) observem grans similituds en el perfil tot i que l'el·lipticitat en el segon dels casos és menor. Segurament es tracta d'estructures molt similars però de diferent estabilitat, de manera que el grau d'estructuració en la mescla AI+BI és menor.

Clarament, doncs, la seqüència lineal AI és la més prometedora per a la formació d'una estructura heterodimèrica amb un altre oligonucleòtid, ja sigui cíclic o fins i tot lineal. D'acord amb aquestes dades es va procedir a l'estudi mitjançant RMN de l'heterodímer format per d<pCCTTCGGT> (cicle B) i d(GTCCCTCA) (lineal AI). Si bé es va poder observar que es forma el complex, com s'explicarà en el proper apartat, es va veure que l'extrem 3' era força flexible i el parell A-T només estava parcialment format. Per això es va decidir estudiar la formació del motiu entre el cicle d<pCCTTCGGT> (B) i l'oligonucleòtid lineal d(CTCAGTCC) (AIII) (Figura 3.4.A), que també està tallat pel motiu, però presentaria parells G-C en els dos extrems, que són més estables que els A-T i podrien resoldre el problema de la flexibilitat en l'extrem 3'.

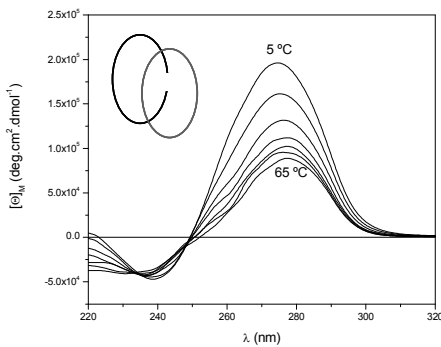
Per poder comparar l'estabilitat dels tres complexos (cicle A + cicle B, lineal AI + cicle B i lineal AIII + cicle B), es van realitzar estudis mitjançant dicroisme circular en condicions de concentració d'oligonucleòtid més elevada, que es mostren a la Figura 3.15. El perfil de fusió és molt similar en els tres casos, però l'heterodímer cicle A + cicle B mostra una major estructuració i és més estable. Els dos complexos heterolineals presenten una estabilitat i estructuració similars.



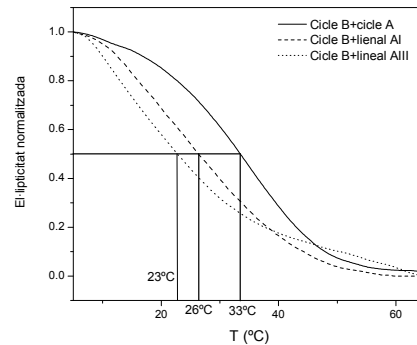
A. Cicle B + cicle A  
 $d\langle pCCTTCGGT \rangle \cdot d\langle pCAGTCCCT \rangle$  0.24 mM  
 (fosfat sòdic 25 mM pH 7, 100 mM NaCl)



B. Cicle B + lineal AI  
 $d\langle pCCTTCGGT \rangle \cdot d\langle GTCCCTCA \rangle$  0.67 mM  
 (Na<sub>2</sub>PIPES 10 mM pH 7, 100 mM NaCl, 10 mM MgCl<sub>2</sub>)



C. Cicle B + lineal AIII  
 $d\langle pCCTTCGGT \rangle \cdot d\langle CTCAGTCC \rangle$  0.26 mM  
 (fosfat sòdic 25 mM pH 7, NaCl 100 mM, 10 mM MgCl<sub>2</sub>)



D. Comparació de les corbes de fusió normalitzades

Figura 3.15. Espectres de dicroisme circular a diferents temperatures i fusions per als complexos: A. Cicle B + cicle A. B. Cicle B + lineal AI. C. Cicle B + lineal AIII. D. Comparació de les fusions

### 3.5 Estudis mitjançant Ressonància Magnètica Nuclear

Els estudis mitjançant ressonància magnètica nuclear es van dur a terme a l'Institut de Química-Física Rocasolano del CSIC, a Madrid, sota la supervisió del Dr. Carlos González.

#### 3.5.1 Complexos formats per $d\langle pCCTTCGGT \rangle$ i $d\langle GTCCCTCA \rangle$ , $r\langle GUCCCUCA \rangle$ o $d\langle CTCAGTCC \rangle$

En aquest apartat abordarem l'estudi mitjançant ressonància magnètica nuclear dels complexos formats pel cicle B i un oligonucleòtid lineal. Els oligonucleòtids lineals escollits han estat: l'octàmer lineal AI,

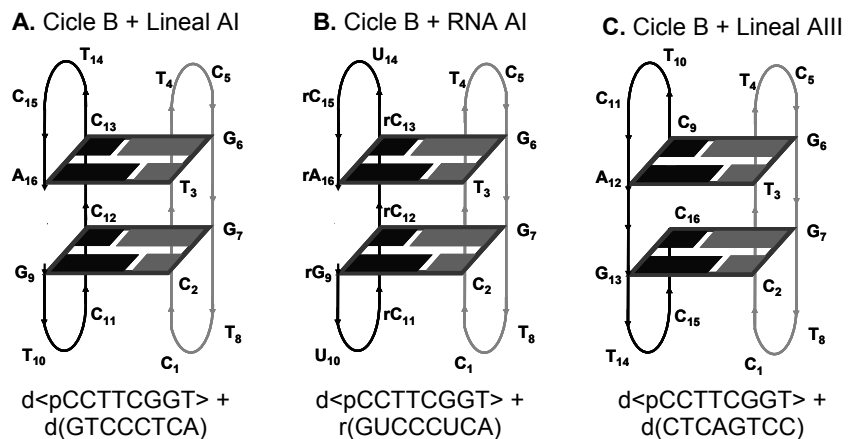


Figura 3.16. Heterodímers estudiats mitjançant RMN amb el cicle  $d\langle pCCTTCGGT \rangle$



d(GTCCCTCA), seqüència més prometedora segons els estudis per UV i CD; el seu anàleg ribonucleotídic r(GUCCCUCA), per determinar si és possible la formació del motiu amb RNA, i la seqüència lineal provinent de tallar per l'altra banda del motiu: AIII, d(CTCAGTCC), amb l'idea d'aconseguir un *bi-loop* més estable, ja que estaria flanquejat en els dos extrems per parells G·C. Els esquemes dels possibles *bi-loops* es mostren en la Figura 3.16. La seqüència d'RNA es va adquirir a Eurogentec (OliGold®, escala 2.5 μmol).

Com ja s'ha comentat en els fonaments, el primer pas per a l'anàlisi d'un oligonucleòtid mitjançant RMN és enregistrar l'espectre monodimensional de protó. Els espectres dels oligonucleòtids per separat en el medi tamponat de treball (25 mM fosfat sòdic pH 7, 100 mM NaCl, 50 mM MgCl<sub>2</sub>) no presenten cap tipus de senyal de protó bescanviable, indicant que no s'autoestructuren, excepte pel cas del cicle B, que sí que mostra autoestructuració (Figura 3.17.A). Podem controlar la formació dels complexos mitjançant els canvis que es produeixen en la regió de protons bescanviables quan adicionem el segon oligonucleòtid sobre el cicle (veure Figura 3.17.B,C,D,E). Pel cas del complex cicle B + cicle A (Figura 3.17.B), l'addició de només un equivalent de d<pCAGTCCCT> fa desaparèixer els senyals a 11.4, 12.2 i 13.7 ppm, corresponents a l'autoestructuració de d<pCCTTCGGT>, i fa aparèixer uns nous senyals entre 13 i 15 ppm, corresponents als protons imino de les guanines (parells G·C, entre 13 i 14 ppm) i al protó imino de la timina (parell A·T, 15 ppm). Pel cas dels oligonucleòtids lineals d(GTCCCTCA) i d(CTCAGTCC) (Figura 3.17.C,D) cal adicionar fins a 5 equivalents d'oligonucleòtid per poder desplaçar l'equilibri cap a la formació de l'heterodímer, indicant una menor afinitat de d<pCCTTCGGT> per al DNA lineal que per a l'anàleg cíclic.

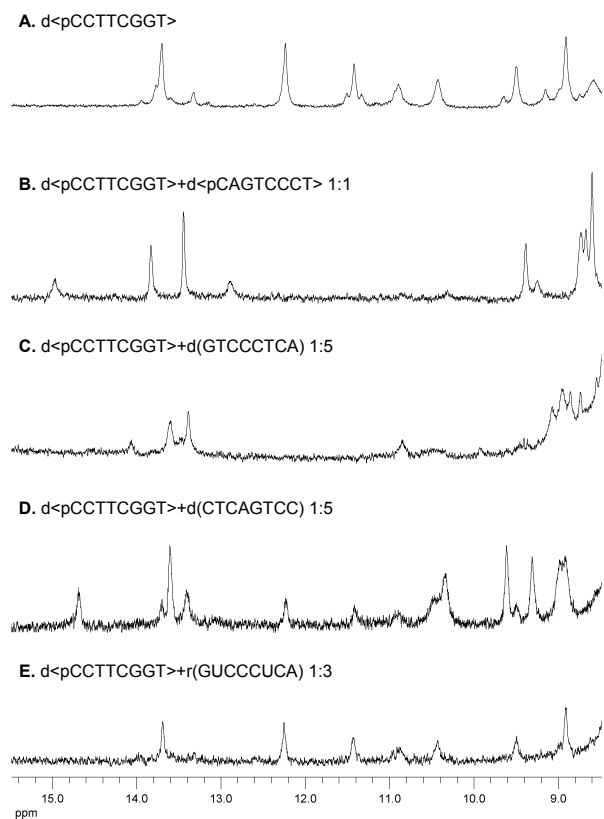


Figura 3.17. Espectres monodimensionals de protó del cicle B i els diversos complexos en medi fosfat sòdic 25 mM pH 7, 100 mM NaCl, 50 mM MgCl<sub>2</sub>, 5 °C. Concentració de cicle: A. 1 mM, B. 1 mM, C. 0.6 mM, D. 0.25 mM, E. 0.7 mM.

A la Figura 3.18 es mostra la valoració pel cas del complex entre d<pCCTTCGGT> i d(CTCAGTCC) (cicle B + lineal AIII), on s'observa clarament com van desapareixent els senyals corresponents al cicle i apareixen els corresponents a l'heterodímer a mesura que adicionem més oligonucleòtid lineal. Aquest efecte no s'observa quan realitzem la valoració amb l'RNA. L'addició de fins a 3

equivalents de r(GUCCUCA) no modifica l'espectre del cicle, com es veu clarament en la Figura 3.17.E. Aquest fet és indicatiu de que ambdós oligonucleòtids no interaccionen.

Dels complexos d<pCCTTCGGT>-d(GTCC-CTCA) (B+AI) i d<pCCTTCGGT>-d(CTCA-GTCC) (B+AIII), se'n van enregistrar espectres bidimensionals (NOESY i TOCSY) tant amb les mescles 1:1 com amb les mescles 1:5. El gran solapament de senyals, junt amb l'existència de més d'una espècie en solució, corresponent a l'estructuració del cicle, no va permetre dur a terme l'assignació total, però es van poder caracteritzar les estructures per comparació amb els espectres de l'heterodímer cicle-cicle i, sobretot, en base als contactes entre protons bescanviables. Així, en el cas del primer complex, B+AI, s'observen clarament els contactes típics de tres parells G·C WC (veure Figura 3.19): H5C → NH<sub>2</sub>C → H1G → NH<sub>2</sub>G. També s'observa el senyal d'un imino de timina aparellada, però no dona cap pic

NOE amb l'H2 de l'adenina amb la qual s'aparella, indicant que aquest parell només està parcialment format. A més a més, els pics H1'-base intraresidu indiquen que els angles glicosídics estan en *anti* i també s'observen els desplaçaments anòmals d'H4', H5' i H5'' per les timines de les primeres posicions dels *loops*, especialment per un parell d'elles, tot i que no han estat assignades específicament. L'espectre és en gran part superposable al de l'heterodímer cicle-cicle.

En el cas del complex d<pCCTTCGGT>-d(CTCAGTCC) (B+AIII), tot i estar flanquejat per parells G·C, també s'observa el mateix efecte que per al complex anterior: l'extrem 3' és flexible i el parell G·C només està parcialment format, com ho demostra el fet que s'observi la ressonància de l'imino de G, però no els pics de correlació amb els seus aminos ni amb els de la C amb la qual s'aparella. La resta de l'espectre presenta un gran solapament, però s'hi observen els trets característics del *bi-loop* i és també en gran mesura superposable al de l'heterodímer cicle-cicle.

Així doncs, tot i que no s'ha realitzat l'assignació total ni el càlcul de les estructures, podem concloure que aquests complexos oligonucleòtid cíclic + oligonucleòtid lineal sí que s'estructuren en forma de *bi-loop*<sup>135</sup>. Podem dir que l'oligonucleòtid cíclic força el lineal a plegar-se sobre si mateix per adoptar la forma corbada que li permet interaccionar amb el cicle, de manera que es tracta d'un cas de reconeixement induït (*induced fit recognition*).

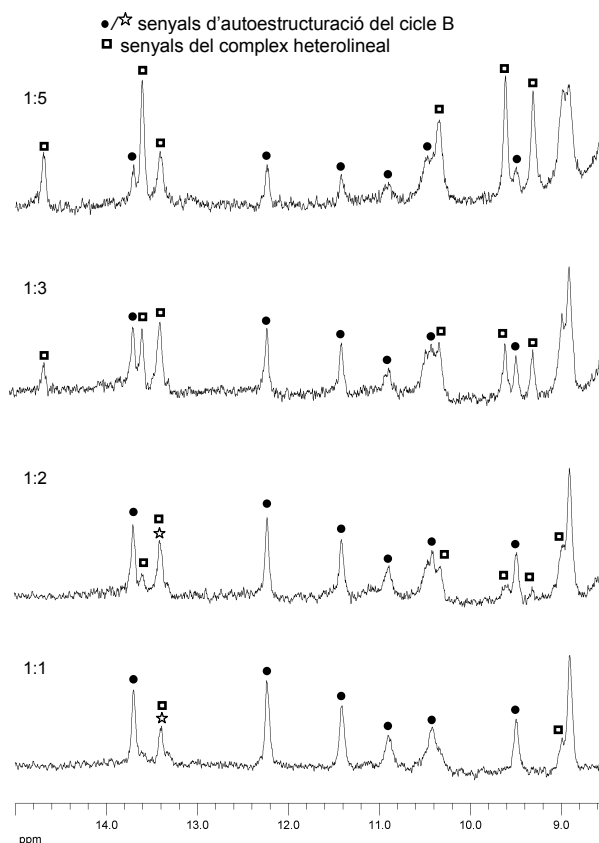


Figura 3.18. Valoració de d<pCCTTCGGT> amb d(CTCAGTCC) (0.25 mM de cicle, fosfat sòdic 25 mM pH 7, 100 mM NaCl, 50 mM MgCl<sub>2</sub>, 5 °C)

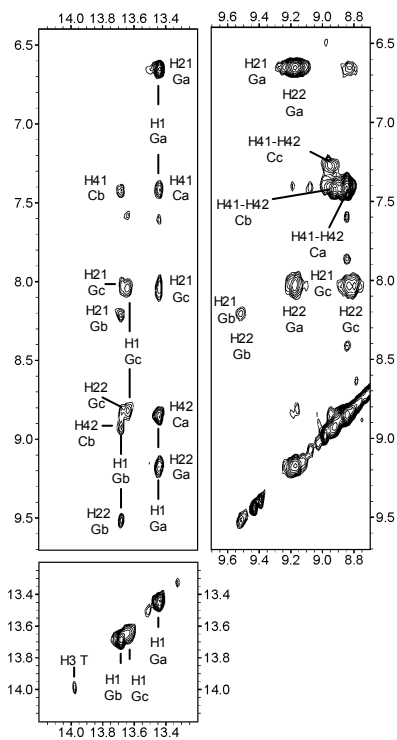


Figura 3.19. Regió de protons bescanviats de l'espectre NOESY (300 ms) de  $d\langle pCCTTCGGT \rangle \cdot d(GTCCCTCA)$  (B+AI, 1:1) en  $H_2O$  (0.6 mM de cada oligonucleòtid, 100 mM NaCl, 50 mM  $MgCl_2$ , pH 7, 5 °C). Els subíndexs a,b,c indiquen els 3 parells de bases G·C observats.

Per contra, en el cas del complex amb RNA, tant en els espectres monodimensionals (Figura 3.17.E) com en els bidimensionals, només s'observa l'autoestructuració del cicle. D'aquí se n'extreuen dues conclusions importants:

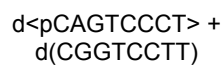
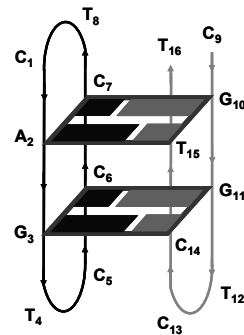
- La formació del *bi-loop* amb una seqüència d'RNA no sembla possible. Aquest resultat és consistent amb el fet que el motiu *bi-loop* homodimèric s'havia intentat estudiar prèviament al grup sense resultats satisfactoris. Probablement es degui a la presència de l'hidroxil en 2', que dona lloc a contactes desfavorables degut a la gran proximitat dels esquelets. Nogensmenys, sí que sabem que és possible formar una tètrada de solc menor on un dels nucleòtids sigui un ribonucleòtid, com ho demostra la tètrada observada per cristal·lografia de RX en contactes de dúplex d'A-DNA<sup>117d</sup>.
- El cicle  $d\langle pCCTTCGGT \rangle$ , tot i haver estat dissenyat per no autoestructurar-se, sí que ho fa. Aquesta autoestructuració, ja sigui monomèrica o dimèrica, sembla que ha d'implicar parells no canònics i per això no s'havia considerat en el moment del disseny. En el seu espectre monodimensional s'observen tres senyals d'iminos clars: 13.7 ppm, senyal d'imino de G d'un parell G·C, 12.2 i 11.4 ppm, senyals imínics que no corresponen a parells WC, sinó als iminos de parells G·T. En els espectres NOESY en aigua lleugera, s'observa que aquests dos iminos donen un pic de correlació entre ells, confirmant la formació d'un parell G·T. L'estudi d'aquesta estructura s'explicarà en el següent capítol.

### 3.5.2 Complexos formats per $d\langle pCAGTCCCT \rangle$ i $d(CGGTCCTT)$ , $d(GTCCTTCG)$ o $d(TTTTGGTTCTTTTT)$

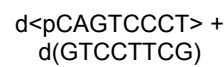
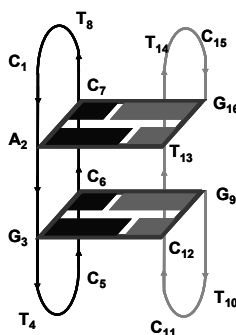
Atesa l'autoestructuració del cicle  $d\langle pCCTTCGGT \rangle$ , que va dificultar clarament l'estudi dels complexos heterolínicals, ens vam plantejar estudiar aquests tipus de complexos partint del mateix heterodímer però emprant l'altre cicle (cicle A,  $d\langle pCAGTCCCT \rangle$ ) com a motlle i diferents oligonucleòtids lineals (Figura 3.20).

Cal recordar que en els estudis preliminars mitjançant dicroisme circular, la combinació cicle A + lineal BI també semblava presentar una certa estructuració que podria correspondre al *bi-loop*.

A. Cicle A + Lineal BI



B. Cicle A + Lineal BII



C. Cicle A + Lineal BIIllarg

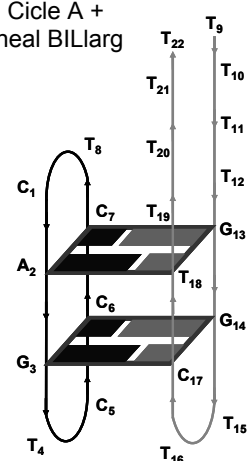


Figura 3.20. Heterodímers estudiats mitjançant RMN amb el cicle  $d\langle pCAGTCCCT \rangle$

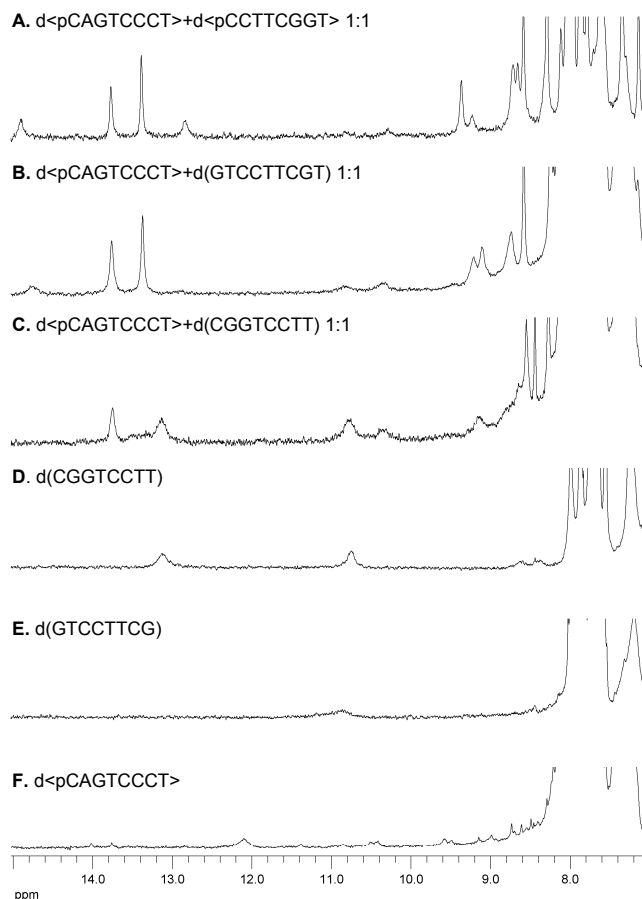


Figura 3.21. Espectres monodimensionals de protó dels oligonucleòtids per separat i les seves combinacions en medi fosfat sòdic 25 mM pH 7, 100 mM NaCl, 5 °C. Concentracions: A: 1 mM complex, B: 0.5 mM complex, C: 0.5 mM complex, D: 0.65 mM, E: 1.7 mM, F: 1.9 mM.

Primer de tot es van enregistrar els espectres monodimensionals dels oligonucleòtids per separat en el medi tamponat de treball (Figura 3.21.D,E,F) per assegurar-nos que no s'autoestructuraven. L'espectre de l'oligonucleòtid lineal BI, d(CGGTCCTT), va mostrar algun senyal de protó bescanviabl, que deu correspondre a la formació de parells G-C d'un dúplex parcial. En la resta de casos, no s'apreciava una estructuració especial. El fet que l'oligonucleòtid cíclic no s'autoestructuri permet veure la formació dels complexos en mostres on la relació és 1:1 (Figura 3.21.A,B,C), contràriament al que succeïa en el cas del cicle B. El complex que presenta una major estructuració, similar a la de l'heterodímer de referència, és la combinació cicle A + lineal BII (Fig.3.21.B).

Per al complex de  $d\langle pCAGTCCCT \rangle + d(GTCCTTCG)$  (A+BII), els espectres bidimensionals van mostrar que es forma clarament el complex de l'oligonucleòtid cíclic amb el lineal, com ja ho evidenciava

l'espectre monodimensional. No es van arribar a assignar tots els senyals a causa del solapament per la presència de 16 sistemes de spin, però sí que es van poder caracteritzar tots els protons bescanviables i els patrons de parells de bases. Amb aquest complex ocorre el mateix que amb els complexos heterolineals anteriors: només s'observen 3 parells WC totalment formats (1 A-T i 2 G-C). L'extrem 3' de l'oligonucleòtid lineal és flexible i el tercer parell G-C només està parcialment format (Figura 3.22).

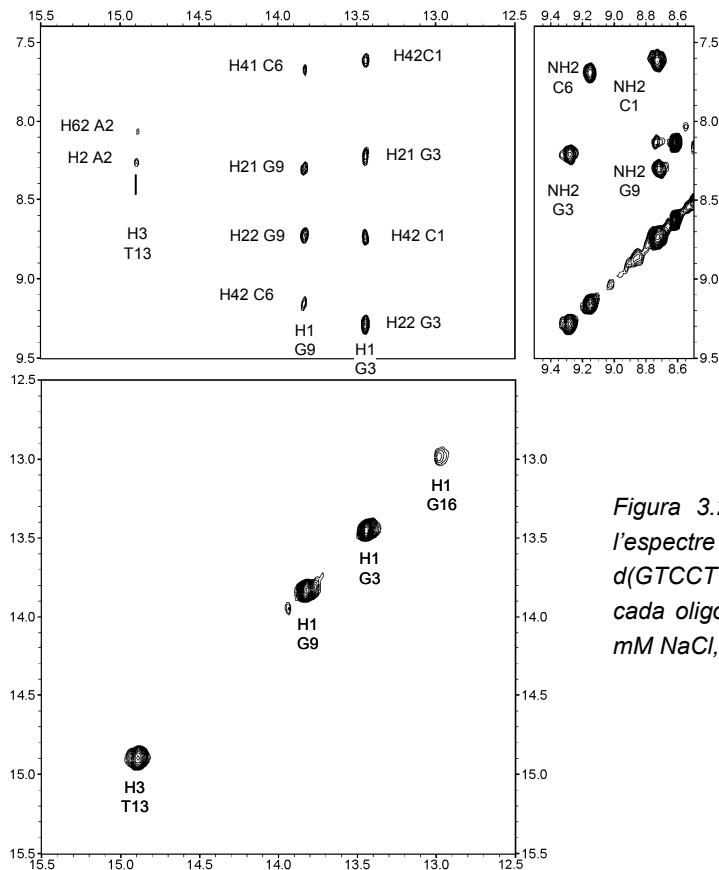


Figura 3.22. Regió de protons bescanviables de l'espectre NOESY (150 ms) de  $d\langle pCAGTCCCT \rangle + d\langle GTCCCTCG \rangle$  (A+BII, 1:1) en  $H_2O$  (0.5 mM de cada oligonucleòtid, fosfat sòdic 25 mM pH 7, 100 mM NaCl, 1 °C).

Per a la mostra  $d\langle pCAGTCCCT \rangle + d\langle CGGTCCTT \rangle$  (A+BI), la situació és una mica més confusa degut a l'estructuració parcial de l'oligonucleòtid lineal. En afegir-li el cicle, sí que s'observa l'aparició d'un nou senyal imino a 13.81 ppm, però l'altre senyal resta invariable. Creiem que aquest senyal a 13.81 ppm correspon a un parell G-C del *bi-loop*, i que l'altre senyal a 13.19 ppm correspondria al segon parell G-C de l'espècie dimèrica, que malauradament coincideix amb el senyal imino d'estructuració parcial de l'oligonucleòtid lineal. Aquests senyals són força amples, indicant que intercanvien amb l'aigua, i no se'n poden observar correlacions amb altres senyals, de manera que és impossible caracteritzar els parells de bases formats. Però la resta de l'espectre bidimensional NOESY sí que presenta característiques que semblen indicar la formació del *bi-loop*, especialment els desplaçaments anòmals dels protons del sucre de les timidines en les primeres posicions del llaç.

Amb la combinació d<pCAGTCCCT> + d(TTTTGGTTCTTTT) (A+BILlarg) es va voler estudiar la possibilitat de formar un *bi-loop* amb un oligonucleòtid lineal més llarg per veure si es podria dirigir un oligonucleòtid cíclic a una seqüència lineal de DNA. Per aconseguir aquest objectiu, l'oligonucleòtid lineal ha d'estar tallat pel llaç, de manera que puguem allargar la seqüència pels extrems, ja que el motiu *bi-loop* no pot formar-se amb més de dues tètades. De les dues dobles cadenes que formen el *bi-loop*, en podríem allargar una fins a l'infinit, formant un llarg dúplex en el solc menor del qual s'hi introduirien dos nucleòtids de l'altra doble cadena, però no es poden allargar les dues cadenes alhora ja que donarien lloc a xocs estèrics.

Com es veu en l'espectre monodimensional de la mostra 1:1 (Figura 3.23), no s'observa cap tipus de senyal de protó bescanviable que impliqui una estructuració de l'oligonucleòtid cíclic amb el lineal. És a dir, que no sembla que es pugui formar el *bi-loop* amb un oligonucleòtid lineal més llarg. Si enregistrem un espectre monodimensional amb la seqüència de polsos p1331, que permet veure els protons que intercanvien ràpidament amb l'aigua, s'observen diversos senyals d'imino, però de molt baixa intensitat, indicant que només existeix una estructuració molt minoritària.

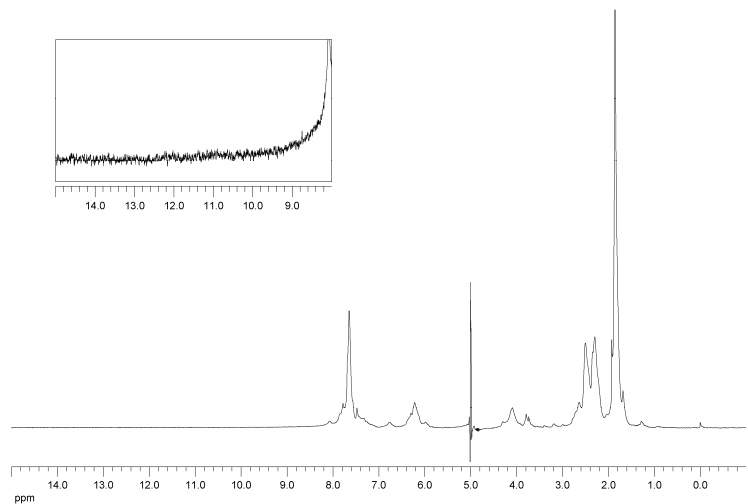


Figura 3.23. Espectre monodimensional watergate de d<pCAGTCCCT> + d(TTTTGGTTCTTTT) (A+BILlarg, 1:1) en H<sub>2</sub>O (0.5 mM de cada oligonucleòtid, fosfat sòdic 25 mM pH 7, 100 mM NaCl, 1 °C).