



# UNIVERSITAT DE BARCELONA

## Síntesis de oligonucleótidos cíclicos y su utilización en el estudio de nuevos motivos estructurales del DNA

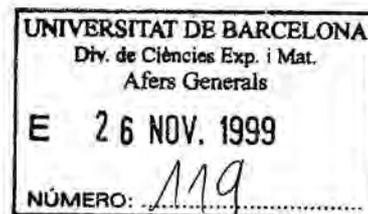
Nuria Escaja Sánchez

**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tdx.cat](http://www.tdx.cat)) i a través del Dipòsit Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tdx.cat](http://www.tdx.cat)) y a través del Repositorio Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tdx.cat](http://www.tdx.cat)) service and by the UB Digital Repository ([diposit.ub.edu](http://diposit.ub.edu)) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Tesis Doctoral



*Síntesis de oligonucleótidos cíclicos y su utilización en el estudio de nuevos motivos estructurales del DNA.*

**Nuria Escaja Sánchez**

Programa de doctorado: Química Fundamental-Química Orgánica  
Bienio: 1996-1998

Departamento de Química Orgánica  
Facultad de Química. División de Ciencias Experimentales y Matemáticas  
Universidad de Barcelona  
1999



Memoria presentada por

Nuria Escaja Sánchez

para optar al grado de Doctora en Ciencias Químicas por la Universidad de Barcelona

Revisada por:



Dr. Enrique Pedroso Müller



Dr. Carlos González Ibáñez

Departamento de Química Orgánica  
Facultad de Química  
Universidad de Barcelona  
Barcelona, noviembre de 1999

*A mis padres*

### *Agradecimientos*

*Quisiera agradecer especialmente a los directores de esta Tesis Enrique Pedroso y Carlos González (y viceversa) la confianza, el apoyo, la dedicación y el entusiasmo que habéis mantenido a lo largo de todos estos años. Me siento muy afortunada de haber tenido la oportunidad de disfrutar de esta colaboración.*

*Quisiera extender este agradecimiento a Anna Grandas, por sus palabras de ánimo, sus consejos y su ayuda en algunos momentos de esta Tesis, y a Manuel Rico, por acogerme a temporadas en el grupo de RMN del CSIC de Madrid.*

*Al grupo de cristalografía de Olga Kennard, del Cambridge Crystallographic Data Center, y especialmente a Steve Salisbury, por el esfuerzo y la tenacidad que hicieron posible resolver la estructura del octámero.*

*A Núria Ferrer, de los Servicios Científico-Técnicos por su ayuda en mis batallas con el controlador de temperatura del DC. A Miguel Feliz, Marga Gairí, M<sup>a</sup> Antonia Molins, Francisco Cárdenas y Carles del Servicio de RMN, por su ayuda y paciencia con esos espectros en fase sólida o muy muy diluidos!. A Irene Fernández y Marta Vilaseca del Servicio de Espectrometría de Masas y a M<sup>a</sup> Pilar Navarro (cuántas horas frente al CARY!) por vuestra ayuda en todo momento.*

*A todas las personas con las que he compartido estos años, mis compañeros de laboratorio. Especialmente a los que he tenido más cerca, "los oligos veteranos" Mostafa, Maite y Jordi, agradeceros vuestra ayuda y consejos, el apoyo, la compañía y, sobretodo, esos tés a media tarde (en el bar!). A Tina, compañera dentro y fuera del laboratorio, por todos esos momentos imborrables (yo guardo los tickets!!), A Cristina, Cris, Josep M<sup>a</sup>, Eva, M<sup>a</sup> Luz, Alberto, Jordi Bacardit, Mons, M<sup>a</sup> Elena...porque habéis sido unos compañeros estupendos. A los oligos que os habéis incorporado después, Miriam, Vicente, Laurent (y que ya tenéis los días contados, ánimo!!). A los "rmneros", Óscar y Miquel Ángel, por vuestra ayuda y paciencia en mis pinitos con las Silicon. Y también recordar a aquellos que, durante estos años, han pasado por el grupo de oligos, Tona, Blanca, Marcus, Begoña, Marta, Yoli, Nausica, Víctor...*

*Al grupo de RMN del CSIC de Madrid, Manolo, Marta, Jorge, Pascal, M<sup>a</sup> Ángeles, Doug, Clara, José Manuel, Eva, Cristina, Apolo....gracias por vuestra acogida, vuestra ayuda en mi descubrimiento de la RMN y porque me habéis hecho sentir como en casa.*

*Esta Tesis Doctoral ha sido posible gracias a la concesión de una beca para la formación de investigadores en el marco del Centre de Referència en Biotecnologia, concedida a través del IRTA.*

## ÍNDICE

### ABREVIATURAS

#### 0.- INTRODUCCIÓN

0.0.- ESTRUCTURAS CÍCLICAS DE ÁCIDOS NUCLEICOS EN LA NATURALEZA .....	1
0.1.- INTERÉS DE LOS OLIGONUCLEÓTIDOS CÍCLICOS DE ORIGEN SINTÉTICO .....	1
0.2.- INTERÉS ESTRUCTURAL .....	1
0.2.1.- APLICACIÓN TERAPÉUTICA .....	3
0.2.2.- OTRAS APLICACIONES BIOLÓGICAS: SUSTRATOS DE POLIMERASAS. COPIAS MULTIMÉRICAS DE DNA Y RNA.....	4
0.2.3.- MÉTODOS DE SÍNTESIS DE OLIGONUCLEÓTIDOS CÍCLICOS .....	6
0.3.- OBJETIVOS.....	9

#### 1.- SÍNTESIS EN FASE SÓLIDA DE OLIGONUCLEÓTIDOS CÍCLICOS. OPTIMIZACIÓN DEL MÉTODO DE SÍNTESIS Y APLICACIÓN A SÍNTESIS A GRAN ESCALA.

1.0.- INTRODUCCIÓN Y OBJETIVOS.....	11
1.1.- OPTIMIZACIÓN DEL MÉTODO DE SÍNTESIS .....	17
1.1.1.- DISEÑO DE UN ESQUEMA DE PROTECCIÓN ORTOGONAL.....	17
1.1.1.1.- Esquema de protección 2: $R_1 = \text{Me}$ , $R_2 = \text{CNE}$ .....	18
1.1.1.1.1.- Eliminación del grupo Me con tiofenol como nucleófilo y distintas bases.....	19
1.1.1.1.2.- Eliminación del grupo Me con otros nucleófilos alternativos al tiofenol .....	23
1.1.1.2.- Esquema de protección 3: $R_1 = \text{Me}$ , $R_2 = \text{CNE}$ , con eliminación de $R_1$ antes de la elongación de la cadena.....	25
1.1.1.3.- Síntesis de oligonucleótidos realizadas con los esquemas de protección 1, 2 y 3 .....	26
1.1.2.- ESTUDIO DE DIFERENTES ETAPAS DEL MÉTODO DE SÍNTESIS .....	34
1.1.2.1.- Ensayo 1: Determinación de grupos 5'-OH libres después de la etapa de ciclación.....	34
1.1.2.2.- Ensayo 2: Bloqueo del grupo OH 5'-terminal?.....	35
1.1.2.3.- Ensayo 3: Análisis de los filtrados procedentes de la reacción de ciclación .....	35
1.1.2.4.- Ensayo 4: Análisis de los filtrados procedentes de la etapa de desprotección de los fosfatos internucleosídicos .....	36
1.1.2.5.- Tratamiento de la resina de síntesis después de la reacción de desanclaje.....	37
1.1.3.- LA REACCIÓN DE CICLACIÓN: UNA ETAPA CLAVE EN LA OBTENCIÓN DE OLIGONUCLEÓTIDOS CÍCLICOS.....	38
1.1.3.1.- Elección de reactivos de acoplamiento.....	42
1.1.3.2.- Elección de la secuencia .....	45
1.1.3.3.- Ensayos con los distintos reactivos .....	46
1.1.3.4.- Caracterización del producto secundario con $\Delta = 187$ umas .....	50
1.1.3.5.- Estudio cinético de la reacción de ciclación con MSNT .....	54

1.2.- SÍNTESIS A GRAN ESCALA DE OLIGONUCLEÓTIDOS CÍCLICOS .....	55
<b>2.- ANÁLISIS MEDIANTE RMN DE d&lt;pCATT&lt;math&gt;CATT&lt;/math&gt;&gt; Y d&lt;pTGCGCGCT&lt;math&gt;TGCGCT&lt;/math&gt;&gt;. EQUILIBRIO MONÓMERO-DÍMERO</b>	
2.0.- INTRODUCCIÓN Y OBJETIVOS .....	59
2.1.- LA RMN COMO HERRAMIENTA PARA ESTUDIAR ESTRUCTURAS DE ÁCIDOS NUCLEICOS .....	60
2.1.1.- ADQUISICIÓN DE DATOS EXPERIMENTALES .....	61
2.1.1.1.- Experimento COSY (CORrelated SpectroscopY) .....	61
2.1.1.2.- Experimento TOCSY (TOrtal Correlation SpectroscopY) .....	62
2.1.1.3.- Experimento NOESY (Nuclear Overhauser Effect SpectroscopY) .....	63
2.1.1.4.- Experimento ROESY (Rotating frame nuclear Overhauser Effect SpectroscopY) .....	64
2.1.2.- ASIGNACIÓN DEL ESPECTRO .....	65
2.1.2.1.- Desplazamientos típicos de protón en DNA .....	65
2.1.2.2.- Generalidades y limitaciones del método de asignación .....	66
2.1.2.3.- Asignación de protones no intercambiables .....	66
2.1.2.3.1.- Identificación de los sistemas de espín .....	66
2.1.2.3.2.- Asignación secuencial .....	66
2.1.2.4.- Asignación de protones intercambiables .....	78
2.1.3.- INFORMACIÓN ESTRUCTURAL .....	79
2.1.3.1.- Distancias .....	79
2.1.3.1.1.- Obtención de distancias .....	79
2.1.3.1.2.- Distancias relevantes .....	71
2.1.3.2.- Constantes de acoplamiento .....	73
2.1.4.- CÁLCULO DE ESTRUCTURAS .....	75
2.1.5.- ANÁLISIS DE ESTRUCTURAS .....	77
2.2.- ASIGNACIÓN DEL ESPECTRO DE RMN DE d<pCATT<math>CATT</math>> .....	77
2.2.1.- IDENTIFICACIÓN DE LOS DIFERENTES SISTEMAS DE ESPÍN .....	77
2.2.2.- ASIGNACIÓN SECUENCIAL .....	79
2.3.- ASIGNACIÓN DEL ESPECTRO DE RMN DE d<pTGCTCGCT<math>TGCTCGCT</math>> .....	81
2.3.1.- IDENTIFICACIÓN DE LOS SISTEMAS DE ESPÍN .....	82
2.3.2.- ASIGNACIÓN SECUENCIAL .....	83
2.4.- EQUILIBRIO MONÓMERO-DÍMERO .....	86
2.5.- INFORMACIÓN ESTRUCTURAL EXPERIMENTAL .....	90
2.5.1.- RASGOS PRINCIPALES DE LAS FORMAS MONOMÉRICAS .....	91
2.5.1.1.- Información de espectros en H <sub>2</sub> O .....	91
2.5.1.2.- Información de espectros en D <sub>2</sub> O .....	92

2.5.2.- RASGOS PRINCIPALES DE LAS FORMAS DIMÉRICAS.....	94
2.5.2.1.- Información de espectros en H <sub>2</sub> O.....	94
2.5.2.2.- Información de espectros en D <sub>2</sub> O.....	97
<b>3.- DETERMINACIÓN ESTRUCTURAL DE d&lt;pCATTTCATT&gt; Y d&lt;pTGCTCGCT&gt;</b>	
3.0.- INTRODUCCIÓN Y OBJETIVOS.....	99
3.1.- ESTRUCTURA CRISTALOGRAFICA DE d<pCATTTCATT>.....	99
3.1.1.- RASGOS PRINCIPALES DE LA ESTRUCTURA CRISTALINA DE d<pCATTTCATT>... 100	
3.1.2.- ESTRUCTURA DE RX DE d<pCAT <sup>Br</sup> UCAT <sup>Br</sup> U>.....	103
3.2.- ESTRUCTURA EN DISOLUCIÓN DE d<pCATTTCATT> Y d<pTGCTCGCT>.....	104
3.2.1.- OBTENCIÓN DE RESTRICCIONES DE DISTANCIAS.....	104
3.2.1.1.- Obtención de distancias en los dímeros.....	104
3.2.1.2.- Obtención de distancias en los monómeros.....	106
3.2.2.- CÁLCULO DE ESTRUCTURAS .....	107
3.2.2.1.- Obtención de las estructuras de los dímeros.....	107
3.2.2.2.- Obtención de las estructuras de los monómeros.....	112
3.2.3.- ANÁLISIS DE LAS ESTRUCTURAS .....	115
3.2.3.1.- Estructuras diméricas.....	115
3.2.3.2.- Comparación con otras estructuras relacionadas.....	124
3.2.3.3.- El motivo bi-loop: importancia y generalidad.....	127
3.2.3.4.- Análisis de las estructuras de los monómeros .....	128
<b>4.- EL EFECTO DE LA SECUENCIA EN LA FORMACIÓN DE ESTRUCTURAS BI-LOOP</b>	
4.0.- INTRODUCCIÓN Y OBJETIVOS.....	131
4.1.- ELECCIÓN DE LAS SECUENCIAS .....	131
4.2.- ESTUDIO POR RMN DE: d<pCGCTCATT>, d<pAGCTAGCT>, d<pTGCATGCA> Y d<pCCGTCCGT>.....	135
4.2.1.- ASIGNACIÓN DE LOS ESPECTROS.....	137
4.2.1.1.- d<pCGCTCATT> .....	138
4.2.1.2.- d<pAGCTAGCT>.....	139
4.2.1.3.- d<pTGCATGCA>.....	140
4.2.1.4.- d<pCCGTCCGT> .....	141
4.2.2.- INFORMACIÓN ESTRUCTURAL.....	142
4.2.2.1.- Estructuras diméricas.....	143
4.2.2.2.- Estructuras monoméricas.....	145
4.3.- CARACTERIZACIÓN POR ABSORCIÓN UV Y DICROÍSMO CIRCULAR DE LAS ESTRUCTURAS DIMÉRICAS Y MONOMÉRICAS.....	146
4.3.1.- ESPECTROS DE UV Y DC DE d<pCATTTCATT> Y d<pTGCTCGCT>.....	148

4.3.1.1.- Fusiones por absorción UV .....	149
4.3.1.2.- Espectros de DC .....	150
4.3.1.3.- Efecto de la sal en el espectro de DC .....	153
4.3.1.4.- Efecto de la temperatura en el espectro de DC.....	155
4.3.2.- EFECTO DE LA SECUENCIA EN LAS CURVAS DE FUSIÓN POR ABSORCIÓN UV..	156
4.3.3.- ESTUDIO POR DC .....	159
<b>5.- ESTUDIO TERMODINÁMICO DEL EQUILIBRIO DE FORMACIÓN DE ESTRUCTURAS BI- LOOP</b>	
5.0.- INTRODUCCIÓN Y OBJETIVOS.....	165
5.1.- DESCRIPCIÓN DEL EQUILIBRIO.....	166
5.2.- EQUILIBRIO <i>DUMBBELL</i> -MONÓMERO DESESTRUCTURADO. ESTABILIDAD DE ESTRUCTURAS <i>DUMBBELL</i> EN FUNCIÓN DE LA SECUENCIA.....	167
5.2.1.- ESTUDIO POR RMN: DETERMINACIÓN DE LAS CONSTANTES DE EQUILIBRIO A PARTIR DE LA VARIACIÓN DE LOS DESPLAZAMIENTOS QUÍMICOS CON LA TEMPERATURA.....	167
5.2.1.1.- Caracterización del equilibrio <i>dumbbell</i> -random coil .....	169
5.2.1.1.1.- d<pCATTTCATT> .....	170
5.2.1.1.2.- d<pTGCTCGCT> .....	172
5.2.1.1.3.- d<pTGCTGCA> .....	173
5.2.1.1.4.- d<pAGCTAGCT> .....	174
5.2.2.- DETERMINACIÓN DE LAS CONSTANTES TERMODINÁMICAS A PARTIR DE CURVAS DE FUSIÓN POR DC .....	176
5.2.3.- ESTABILIDAD DE LAS ESTRUCTURAS <i>DUMBBELL</i> .....	178
5.3.- EQUILIBRIO <i>DUMBBELL</i> -DÍMERO.....	180
5.4.- EQUILIBRIO DÍMERO-FORMA DESESTRUCTURADA .....	183
5.4.1.- ESTUDIO POR RMN. CÁLCULO DE POBLACIONES .....	183
5.4.2.- ESTUDIO POR DC: CURVAS DE FUSIÓN EN MEDIO SALINO.....	187
5.4.3.- ESTUDIO POR DC DE LA INFLUENCIA DE LA SAL EN LA ESTABILIDAD DE LA FORMA DIMÉRICA .....	189
5.4.4.- DEPENDENCIA DE LA T <sub>m</sub> DE LA FORMA DIMÉRICA CON LA CONCENTRACIÓN.	190
5.4.5.- EL EFECTO DE LA SECUENCIA EN LA ESTABILIDAD DEL MOTIVO BI-LOOP.....	191
<b>CONCLUSIONES.....</b>	<b>193</b>

## MATERIALES Y MÉTODOS

### 1.- PROCEDIMIENTOS GENERALES

1.1.- INSTRUMENTACIÓN GENERAL .....	197
1.2.- REACTIVOS Y DISOLVENTES .....	197
1.3.- TAMPONES .....	198
1.4.- CROMATOGRAFÍA .....	199
1.4.1.- CROMATOGRAFÍA EN CAPA FINA .....	199
1.4.2.- CROMATOGRAFÍA EN COLUMNA .....	199
1.4.3.- CROMATOGRAFÍA LÍQUIDA .....	199
1.4.3.1.- Cromatografía líquida analítica y semipreparativa .....	199
1.4.3.2.- Cromatografía líquida preparativa .....	200
1.4.3.3.- Cromatografía de intercambio catiónico .....	200
1.5.- ENSAYO CUALITATIVO DE NINHIDRINA .....	200
1.6.- SÍNTESIS DE OLIGONUCLEÓTIDOS EN FASE SÓLIDA .....	201
1.6.1.- INSTRUMENTACIÓN, REACTIVOS Y DISOLVENTES ESPECÍFICOS PARA LA SÍNTESIS DE OLIGONUCLEÓTIDOS .....	201
1.6.2.- DETERMINACIÓN DE LA FUNCIONALIZACIÓN DE RESINAS .....	201
1.6.3.- CUANTIFICACIÓN DE OLIGONUCLEÓTIDOS POR ABSORCIÓN UV .....	200
1.6.4.- CARACTERIZACIÓN DE OLIGONUCLEÓTIDOS POR DIGESTIÓN ENZIMÁTICA .....	203
1.6.5.- ANÁLISIS DE OLIGONUCLEÓTIDOS POR ELECTROFORESIS EN GEL DE POLIACRILAMIDA .....	204
1.6.5.1.- Gel de poliacrilamida analítico en condiciones desnaturalizantes .....	204
1.6.5.2.- Tinción del gel .....	205
1.7.- PREPARACIÓN DE MUESTRAS PARA ANÁLISIS MALDI-TOF .....	205

### 2.- SÍNTESIS DE OLIGONUCLEÓTIDOS

2.1.- SÍNTESIS DE LOS MODELOS UTILIZADOS PARA EL ESTUDIO DE ORTOGONALIDAD DE LOS GRUPOS PROTECTORES .....	207
2.1.1.- OBTENCIÓN DE 3'-O-DMT-TIMIDINA .....	207
2.1.2.- OBTENCIÓN DE 5'-O-DMT-TIMIDINA 3'-O-METIL- <i>N,N</i> -DIISOPROPILFOSFORAMIDITO .....	208
2.1.3.- OBTENCIÓN DEL FOSFATO DE 3'-O-(5'-O-DMT-TIMIDINIL) 5'-O-(3'-O-DMT-TIMIDINIL) Y METILO .....	209
2.1.4.- OBTENCIÓN DEL FOSFATO DE 3'-O-(5'-O-DMT-TIMIDINIL) 5'-O-(3'-O-DMT-TIMIDINIL) Y CIANOETILO .....	209
2.1.5.- OBTENCIÓN DEL FOSFATO DE BIS-3'-O-(5'-O-DMT-TIMIDINIL) Y METILO .....	209
2.1.6.- OBTENCIÓN DEL FOSFATO DE BIS-3'-O-(5'-O-DMT-TIMIDINIL) Y CIANOETILO .....	210
2.2.- SÍNTESIS EN FASE SÓLIDA DE OLIGONUCLEÓTIDOS CÍCLICOS .....	210

2.2.1.- DERIVATIZACIÓN DE SOPORTES POLIMÉRICOS .....	210
2.2.1.1.- Síntesis de espaciador-resina .....	210
2.2.1.2.- Síntesis de nucleotidil-resina .....	211
2.2.1.3.- Síntesis de nucleotidil-linker .....	211
2.2.2.- ELONGACIÓN DE LA CADENA OLIGONUCLEOTÍDICA .....	213
2.2.3.- OBTENCIÓN DE OLIGONUCLEÓTIDOS CÍCLICOS: REACCIÓN DE CICLACIÓN .....	214
2.2.3.1.- Síntesis de 1-mesitilensulfonilo-3-nitro-1,2,4-triazol (MSNT) .....	214
2.2.4.- DESPROTECCIÓN Y DESANCLAJE .....	215
2.2.5.- DESALADO Y PURIFICACIÓN.....	215
2.2.6.- CUANTIFICACIÓN .....	215
2.2.7.- CARACTERIZACIÓN DE LOS OLIGONUCLEÓTIDOS CÍCLICOS .....	216
<b>3.- RMN DE OLIGONUCLEÓTIDOS CÍCLICOS</b>	
3.1.- <sup>1</sup> H-RMN.....	217
3.1.1.- PREPARACIÓN DE LAS MUESTRAS .....	217
3.1.2.- ADQUISICIÓN DE ESPECTROS .....	217
3.1.3.- PROCESADO DE ESPECTROS.....	218
3.1.4.- ASIGNACIÓN DE ESPECTROS .....	218
3.2.- CÁLCULO DE ESTRUCTURAS.....	219
3.2.1.- OBTENCIÓN DE RESTRICCIONES DE DISTANCIAS.....	219
3.2.2.- DETERMINACIÓN DE ESTRUCTURAS.....	221
3.2.3.- ANÁLISIS DE LAS ESTRUCTURAS .....	223
<b>4.- ESPECTROSCOPIA DE UV Y DICROÍSMO CIRCULAR</b>	
4.1.- INSTRUMENTACIÓN .....	224
4.2.- PREPARACIÓN DE LAS MUESTRAS.....	224
4.3.- ADQUISICIÓN DE LOS ESPECTROS .....	224
4.4.- TRATAMIENTO DE LAS CURVAS DE FUSIÓN .....	224
4.5.- ANÁLISIS DE OLIGONUCLEÓTIDOS CÍCLICOS POR DC .....	225
<b>ANEXOS</b>	
ANEXO A: NOMENCLATURA Y DEFINICIÓN DE PARÁMETROS ESTRUCTURALES EN ÁCIDOS NUCLEICOS.	
ANEXO B: TABLAS DE ASIGNACIÓN.	
ANEXO C: TABLAS DE SUMAS DE CONSTANTES DE ACOPLAMIENTO.	
ANEXO D: DETERMINACIÓN DE LA K PARA LOS EQUILIBRIOS DUMBELL-DÍMERO Y RANDOM COIL-DÍMERO, A PARTIR DEL CÁLCULO DE POBLACIONES EN ESPECTROS 1D DE RMN.	

## ABREVIATURAS

- ACN** - acetonitrilo  
**AC<sub>2</sub>O** - anhídrido acético  
**AcOEt** - acetato de etilo  
**AcOH** - ácido acético  
**AcONH<sub>4</sub>** - acetato amónico  
**anh.** - anhidro  
**AP** - fosfatasa alcalina (EC 3.1.3.1)  
**Ar** - señales de <sup>1</sup>H-RMN correspondientes a anillos aromáticos  
**ATT** - 6-aza-2-tiotimina  
<sup>1</sup>**Bu** - isobutirilo  
<sup>1</sup>**BuOOH** - hidroperóxido de tert-butilo  
**Bz** - benzoílo  
**CAMELSPIN** - Cross-relaxation Appropriate for Mini-molecules Emulated by Locked SPINs  
**C<sub>ar</sub>** - señales de <sup>13</sup>C-RMN correspondientes a anillos aromáticos  
**CNE** - 2-cianoetilo  
**C<sub>q</sub>** - señales de <sup>13</sup>C-RMN correspondientes a carbonos cuaternarios  
**COSY**- COrelated SpectroscopY  
**DC** - dicroísmo circular  
**DCC** - *N,N'*-diciclohexilcarbodiimida  
**DCM** - diclorometano  
**DEAE** - dietilaminoetil  
**DIEA** - *N,N*-diisopropiletilamina  
**DIPC** - *N,N*-diisopropilcarbodiimida  
**DMF** - *N,N*-dimetilformamida  
**DMT** - 4,4'-dimetoxitritilo  
**DNA** - ácido desoxiribonucleico  
**dNTP** - 2'-desoxiribonucleósido trifosfato  
**EDTA** - ácido etilendiaminotetraacético  
**EM** - espectrometría de masas  
**EM-EI** - espectrometría de masas de impacto electrónico  
**eq** - n° de equivalentes  
**EM-ES** - espectrometría de masas por electroespray  
**EM-MALDI-TOF** - espectrometría de masas por desorción láser asistida por matriz y con detección de tiempo de vuelo  
**EtOH** - etanol  
**f** - funcionalización de la resina  
**FA** - fosforamidito  
**Fmoc** - 9-fluorenilmetoxicarbonilo  
**FT-IR** - espectroscopía de infrarojo con transformada de Fourier.  
**HOBt** - 1-hidroxibenzotriazol  
**HOHAHA.**- HOmonuclear HArtmann-HAhn  
**3-HPA** - ácido 3-hidroxicipolínico  
**HPLC** - cromatografía líquida de alta eficacia  
**p-MBHA** - 4-(metilbenzhidrilamina)poliestireno  
**2-MBT** - 2-mercaptobenzotriazol  
**Me** - metilo  
**MeOH** - metanol  
**MPLC** - cromatografía líquida de media presión  
**MSNT** - 1-mesitilensulfonil-3-nitro-1,2,4-triazol  
**NMI** - *N*-metilimidazol  
**NMP** - 1-metil-2-pirrolidona  
**NOE**- Efecto Nuclear Overhauser  
**NOESY**- Nuclear Overhauser Effect SpectroscopY  
**NTP** - ribonucleósido trifosfato  
**OD<sub>260</sub>** - medida de absorción a 260nm de una disolución contenida en 1mL y medida en una cubeta de 1cm de paso  
**Ⓟ** - soporte polimérico  
**PAGE** - electroforesis en gel de poliacrilamida  
**PAS** - persulfato amónico

**PEG-PS** - co-polímero de polietilenglicol-poliestireno

**pf** - punto de fusión (°C)

**<sup>i</sup>PrOH** - isopropanol

**PS** - poliestireno-co-1%-divinilbenceno

**pyr** - piridina

**pip** - piperidina

**Rdt** - rendimiento

**Rf** - factor de retención (cromatografía sobre capa fina)

**RMD**.- Dinámica Molecular Restringida

**RMSD**-Desviación cuadrática media

**RMN** - espectroscopía de resonancia magnética nuclear

**RNA** - ácido ribonucleico

**ROESY**- Rotating frame Nuclear Overhauser Effect Spectroscopy

**SpPD** - fosfodiesterasa de bazo de ternera (EC 3.1.16.1)

**SVPD** - fosfodiesterasa de veneno de serpiente (EC 3.1.4.1)

**TBAF** - fluoruro de tetrabutilamonio

**TBDMS** - tert-butildimetilsililo

**TBE** - tampón tris-borato EDTA (10mM tris, 1M H<sub>3</sub>BO<sub>3</sub>, 20mM EDTA)

**TCA** - ácido tricloroacético

**TEA** - trietilamina

**TEAAc** - acetato de trietilamonio

**TEAB** - bicarbonato de trietilamonio

**TEMED** - N,N,N',N'-tetrametiletildiamina

**TFA** - ácido trifluoroacético

**2,4,6-THAP** - 2,4,6-trihidroxiacetofenona

**THF** - tetrahidrofurano

**TLC** - cromatografía de capa fina

**T<sub>m</sub>** - temperatura de fusión de DNA

**TMG** - 1,1,3,3-tetrametilguanidina

**TMP** - trimetil fosfito

**TOCSY**- T<sup>O</sup>tal Correlation Spectroscopy

**t<sub>R</sub>** - tiempo de retención (cromatografía líquida)

**Tris** - tris(hidroximetil)aminometano

**TSP**.- ácido 3-trimetilsilil(2,2,3,3-D<sub>4</sub>)propanoico

**UV-Vis** - espectroscopía de ultravioleta-visible

## **NUCLEÓSIDOS**

**dA** - 2'-desoxiadenosina

**dC** - 2'-desoxicidina

**dG** - 2'-desoxiguanosina

**T** - timidina

**<sup>Br</sup>dU** - 5-bromo-2'-desoxiuridina

---

*INTRODUCCIÓN Y OBJETIVOS*

---

## 0.- INTRODUCCIÓN.

### 0.0.- ESTRUCTURAS CÍCLICAS DE ÁCIDOS NUCLEICOS EN LA NATURALEZA.

La presencia en la naturaleza de estructuras cíclicas de ácidos nucleicos, así como su participación en diversos procesos biológicos se encuentra ampliamente documentada en la bibliografía. Se conoce que dichas estructuras forman parte del material genético de numerosos virus y bacterias<sup>1</sup> y se ha detectado que la delección de intrones en células eucariotas y en el organismo primitivo *Tetrahymena*<sup>2</sup> implica la presencia de poliribonucleótidos cíclicos.

También es conocida la importancia biológica de algunos oligonucleótidos cíclicos de secuencia muy corta. Entre ellos, se encuentran los derivados cíclicos más pequeños posibles, los mononucleósidos cAMP y cGMP que actúan como reguladores bioquímicos o el diribonucleótido cíclico <GpGp> que activa la síntesis de celulosa en la *Acetobacter Xylinum*<sup>3</sup>. La descripción de los diribonucleótidos cíclicos <ApAp>, <ApUp> y <UpUp> como posibles inhibidores de RNA polimerasas<sup>4</sup>, así como el hecho de que varios desoxi- y ribonucleótidos cíclicos de origen sintético d<GpGp>, <GpGp> y <GpGpGp> hayan demostrado tener actividad biológica<sup>3</sup>, incrementaron el interés por este tipo de estructuras. En la actualidad, la utilización de oligonucleótidos cíclicos de origen sintético se extiende a muy diversos campos de estudio<sup>5,6</sup>.

### 0.1.- INTERÉS DE LOS OLIGONUCLEÓTIDOS CÍCLICOS DE ORIGEN SINTÉTICO.

#### 0.1.1.- INTERÉS ESTRUCTURAL.

El conocimiento estructural detallado de las moléculas de interés biológico, principalmente ácidos nucleicos y proteínas, es fundamental para entender su función en la célula. El desarrollo de métodos sintéticos eficaces para obtener polinucleótidos de secuencia definida, así como el importante desarrollo que han experimentado las técnicas de determinación estructural como la RMN multidimensional o la difracción de rayos X, ha permitido que nuestro conocimiento de la estructura de los ácidos nucleicos aumente enormemente en los últimos años. Si bien la estructura más habitual del DNA, la clásica doble hélice a derechas, es conocida desde hace tiempo<sup>7</sup>, se han observado más recientemente otros motivos estructurales tales como *horquillas*, cruciformes, estructuras de triple cadena (*three-way junctions*, tríplex), estructuras de cadena cuádruple (*Holliday junctions*, *G-quadruplex*, *i-motifs*), etc.

<sup>1</sup> Lewin, B. «*Genes*» 1983, Ed. Wiley, J., New York.

<sup>2</sup> Cech, T.R. *Scientific American* 1986, 255, 76-84.

<sup>3</sup> Ross, P.; Weinhouse, H.; Aloni, Y.; Michaeli, D.; Weinberger-Ohana, P.; Mayer, R.; Braun, S.; De Vroom, E.; Van der Marel, G.A.; Van Boom, J.H. Benziman, M. *Nature* 1987, 325, 279-281.

<sup>4</sup> Hsu, C.J.; Dennis, D. *Nucleic Acids Res.* 1982, 10, 5637-5647.

<sup>5</sup> Kool, E.T. *Acc. Chem. Res.* 1998, 31, 502-510.

<sup>6</sup> Catalanotti, B.; Galeone, A.; Mayol, L.; Pepe, A. Lanzotti, V. *Gazzetta Chimica Italiana* 1997, 127, 231-246.

<sup>7</sup> Watson, J.D.; Crick, F.H.C. *Nature* (London) 1953, 171, 737-738.

El estudio de estas formas no canónicas del DNA constituye un campo científico de gran actividad en la actualidad<sup>8</sup>, ya que la observación de esta elevada diversidad estructural ha venido acompañada del descubrimiento de una gran diversidad funcional. Desde el descubrimiento de la capacidad catalítica del RNA<sup>9</sup> los ácidos nucleicos no son meros transmisores de la información genética, sino que también pueden realizar otras funciones. La reciente determinación de la estructura tridimensional de moléculas de RNA<sup>10</sup> y DNA<sup>11</sup> catalíticas abre aún más el panorama de motivos estructurales en el campo de los ácidos nucleicos.

Entre los diferentes motivos estructurales, las horquillas (*hairpins*) resultan de gran importancia puesto que constituyen un motivo estructural básico, presente en otras formas no canónicas de DNA, como es el caso de estructuras cruciformes. Además, se cree que pueden jugar un papel importante en procesos de recombinación y regulación genética. Secuencias capaces de formar estructuras tipo *hairpin* se han identificado frecuentemente cerca de sitios promotores y de regulación genética en DNA y se sabe que están estabilizadas por la interacción de los *loops* con enzimas específicas<sup>12</sup>. Las estructuras *hairpin* se han observado también en numerosas moléculas de RNA.

Las horquillas de DNA consisten en un fragmento de doble hélice, denominado tallo, con pares de bases canónicos Watson-Crick, y un *loop* constituido por un número variable de nucleótidos no apareados. La estabilidad de dichas estructuras depende del número y tipo de nucleótidos que forman el *loop*, además de otros factores como la simetría impuesta por el esqueleto azúcar-fosfato y las interacciones entre el *loop* y la parte central del *hairpin*, en particular con el par de bases que cierra el *loop*.

El estudio de estructuras *hairpin* de DNA utilizando fragmentos lineales resulta complicado, no sólo en disolución, donde al tratarse generalmente de secuencias repetitivas coexisten en equilibrio con otras formas posibles de apareamiento<sup>13</sup>, sino también mediante métodos cristalográficos. De hecho, hasta el momento, no se ha podido determinar la estructura cristalina de ninguna estructura *hairpin* debido a que tienden a cristalizar como estructuras dúplex<sup>14-15</sup> y toda la información de que se dispone hasta el momento sobre *hairpins* procede de estudios en disolución, realizados mediante la técnica de RMN.

La restricción conformacional impuesta por la ciclación contribuye con un factor entrópico favorable a aumentar la estabilidad térmica de los *hairpins* respecto a los que se forman con cadenas lineales, permitiendo realizar estudios en un amplio rango de temperatura. Este hecho resulta muy conveniente tanto en el estudio mediante técnicas de cristalografía como de RMN.

<sup>8</sup> Lebrun, A.; Lavery, R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 348-354.

<sup>9</sup> Kruger, K.; Grabowski, P.J.; Zaug, A.J.; Sands, J.; Gottshilling, D.E.; Cech, T.R. *Cell* **1982**, *31*, 147-157.

<sup>10</sup> Pley, H.W.; Flaherty, K.M.; McKay, D.B.; *Nature* **1994**, *372*, 68-74.

<sup>11</sup> Nowakowski J., Shim, P.J., Prasad, G.S., Stout, S.D. and Joyce, G.F. *Nature Struct. Biol.* **1999**, *6*, 151-156.

<sup>12</sup> Cheng, K.Y.; Tinoco, I. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8705-8709.

<sup>13</sup> Referencia del artículo donde describen equilibrio *hairpin*-dúplex...

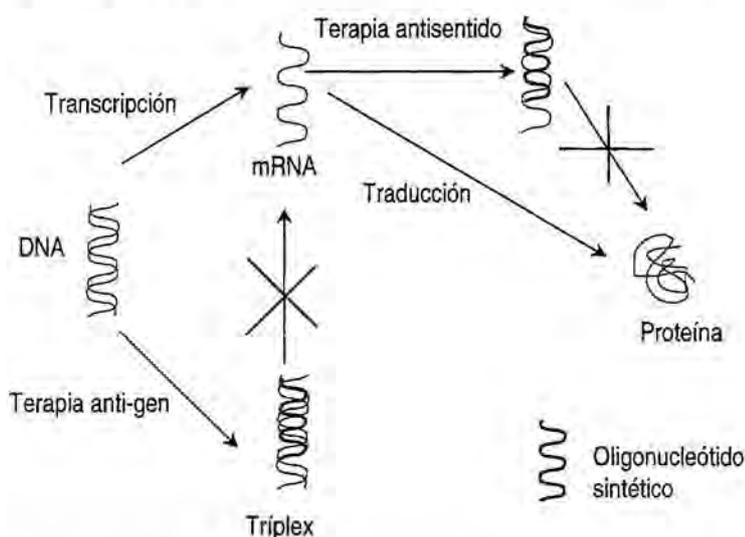
<sup>14</sup> Holbrook, S.R.; Cheong, C.; Tinoco Jr., I. *Nature* **1991**, *353*, 579.

<sup>15</sup> Chattopadhyaya, R.; Grzeskowiak, K.; Dickerson, R.E. *J. Mol. Biol.* **1990**, *211*, 189.

Los oligonucleótidos cíclicos se han revelado, también, como una de las mejores alternativas para estabilizar estructuras tríplex<sup>16,17</sup>. Además de formar complejos más estables, la preorganización del ligando proporciona una mayor afinidad de unión así como una mayor selectividad por la secuencia diana.

### 0.1.2.- APLICACIÓN TERAPÉUTICA.

La utilización de oligonucleótidos como fármacos contra el cáncer, infecciones víricas u otros desórdenes de origen genético, constituye una alternativa muy prometedora para el tratamiento de estas graves enfermedades. La estrategia consiste en el bloqueo de los mecanismos de replicación, transcripción o traducción, mediante las denominadas *terapias antisentido* y *anti-gen*, Figura 0.1.



**Figura 0.1.-** Actuación de oligonucleótidos en terapias antisentido y anti-gen.

Desde la constatación de la utilidad de los oligonucleótidos como fármacos antivirales<sup>18,19</sup>, se han llevado a cabo numerosos esfuerzos para optimizar la aplicación de los oligonucleótidos sintéticos a este tipo de terapias. Una de las principales limitaciones que presenta la utilización de oligonucleótidos no modificados es que son rápidamente degradados por los enzimas nucleasas presentes en el medio celular. Por este motivo, se ha estudiado la posibilidad de introducir modificaciones de tipo químico en la cadena oligonucleotídica que confieran una mayor estabilidad a las moléculas, generándose análogos del tipo fosforotioatos, fosforamidatos o metilfosfonatos<sup>20</sup>.

Sin embargo, no menos interesantes resultan las modificaciones topológicas. La mayor resistencia a las nucleasas que presentan hairpins de DNA o RNA estables, sugiere su uso en la

<sup>16</sup> Prakash, G., Kool, E.T. *J. Chem. Soc. Commun.* **1991**, 1161-1163.

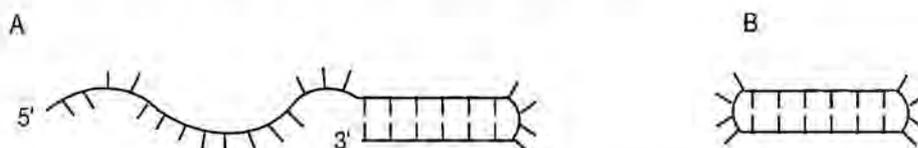
<sup>17</sup> Kool, E.T. *J. Am. Chem. Soc.* **1991**, 113, 6265-6266.

<sup>18</sup> Zamecknik, P.C.; Stephenson, M.L. *Proc. Natl. Acad. Sci. USA* **1978**, 75, 280-284.

<sup>19</sup> Stephenson, M.L.; Zamecknik, P.C. *Proc. Natl. Acad. Sci. USA* **1978**, 75, 285-288.

<sup>20</sup> Uhlman, E.; Peyman, A. *Chem. Rev.* **1990**, 90, 544-584.

estabilización de oligonucleótidos antisentido. Dado que la degradación de oligonucleótidos antisentido se produce predominantemente por la actividad de 3'-exonucleasas<sup>21-22</sup>, los oligonucleótidos pueden ser estabilizados por la formación de estructuras hairpin en el extremo 3' terminal<sup>23-24</sup>, Figura 0.2-A, o incluso en el extremo 5'-terminal<sup>25</sup>. Si este concepto se extiende a ambos extremos, surgen oligonucleótidos cíclicos (Figura 0.2-B) que, al no tener ninguno de los extremos terminales libres, presentarían aún una mayor resistencia a las exonucleasas<sup>26-29</sup>. En todos los casos, en las referencias citadas se demuestra la propiedad reguladora o inhibidora de este tipo de estructuras. Los oligonucleótidos cíclicos, siguen manteniendo los enlaces naturales 3'-5'-fosfodiésteres, reduciendo el riesgo potencial de toxicidad y mutagenicidad asociado a otros tipos de oligonucleótidos modificados químicamente.



**Figura 0.2.-** A) Estabilización de un fragmento oligonucleotídico lineal por formación de una estructura hairpin estable en el extremo 3'-terminal. B) Estructura dumbbell.

Recientemente se ha descrito la síntesis de *dumbbells* constituidos por una secuencia sentido de RNA y la secuencia complementaria antisentido de DNA, unidas por dos *hairpin-loops*<sup>30-31</sup>. En este caso, el oligonucleótido antisentido es liberado a la célula infectada por la acción de la RNasa H sobre el *dumbbell* de RNA/DNA.

### 0.1.3.- OTRAS APLICACIONES BIOLÓGICAS: SUSTRATOS DE POLIMERASAS. COPIAS MULTIMÉRICAS DE DNA Y RNA.

Las DNA polimerasas son enzimas que catalizan la obtención de copias de cadenas de DNA. Para ello, necesitan generalmente una secuencia cebadora y una secuencia de cadena larga que actúe como molde para obtener la cadena complementaria. Cuando el molde es

<sup>21</sup> Rumney, S.; Kool, E.T. *Angew. Chem.* **1992**, 104, 1686.

<sup>22</sup> Gao, H.; Yang, M.; Patel, R.; Cook, A.F. *Nucleic Acids Res.* **1995**, 23, 2025.

<sup>23</sup> Khan, I.M.; Coulson, J.M. *Nucleic Acids Res.* **1993**, 21, 2957-2958.

<sup>24</sup> Poddevin, B.; Megueni, S.; Elias, I.; Vasseur, M.; Blumenfeld, M. *Antisense Res. Rev.* **1994**, 147-154.

<sup>25</sup> Yoshizawa, S.; Ueda, T.; Ishido, Y.; Miura, K.; Watanabe, K.; Hirao, I. *Nucleic Acids Res.* **1994**, 22, 2217-2221.

<sup>26</sup> Clusel, C.; Ugarte, E.; Enjolras, N.; Vasseur, M.; Blumenfeld, M. *Nucleic Acids Res.* **1993**, 21, 3405-3411.

<sup>27</sup> Prakash, G.; Kool, E.T. *J. Am. Chem. Soc.* **1991**, 1161-1163.

<sup>28</sup> Azhayeva, E.; Azhayev, A.; Auriola, S.; Tengvall, U.; Urtti, A.; Lönnberg, H. *Nucleic Acids Res.* **1997**, 25, 4954-4961.

<sup>29</sup> Lim, C.S.; Jabrane, N.; Fontes, J.D.; Okamoto, H.; Garovoy, M.R.; Peterlin, B.M.; Hunt, C.A. *Nucleic Acids Res.* **1997**, 25, 575-581.

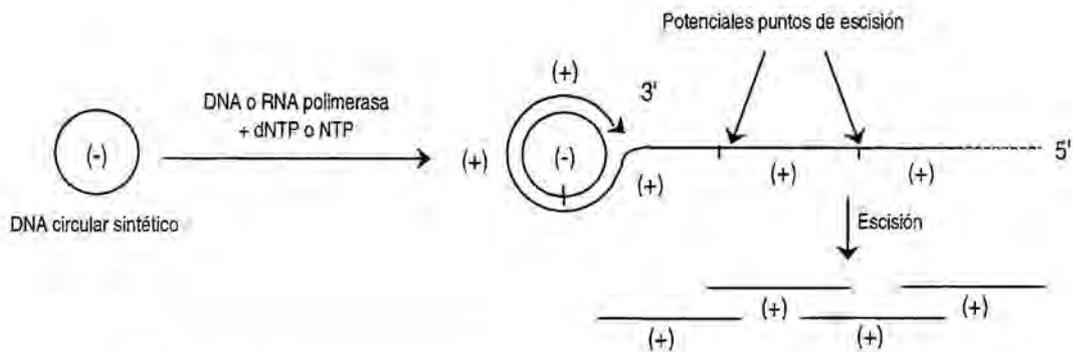
<sup>30</sup> Hamakawa, H.; Takayuki, A.; Takeshi, S.; Kazuyuki, T.; Yamamoto, N.; Hiroshi, T. *Bioorg. & Med. Chem.* **1998**, 1025-1032.

<sup>31</sup> Takayuki, A.; Takai, K.; Nakada, S.; Yokota, T.; Takaku, H. *FEBS Lett.* **1998**, 425, 91-96.

circular, estos enzimas pueden realizar copias siguiendo un mecanismo de replicación de círculo rodante, análogo al utilizado por numerosos plásmidos y virus<sup>32</sup>. Este mecanismo implica dos procesos paralelos, por una parte el enzima debe sintetizar secuencias complementarias al molde circular mientras que, simultáneamente, debe desenrollarse el dúplex entre la cadena sintetizada y el molde, de modo que se permita el avance de la polimerasa. De no existir este último mecanismo, el enzima detendría la replicación después de la primera vuelta<sup>33</sup>. Esto no sucede, sin embargo, cuando se utilizan como moldes ciclos de pequeño y mediano tamaño, tal como se detalla en algunos estudios recientes<sup>34,35</sup>. En estos casos, a medida que avanza la replicación la cadena sintetizada se desenrolla simultáneamente del molde, permitiendo al enzima dar un gran número de vueltas alrededor del molde. Como resultado se obtiene una secuencia repetitiva de DNA cuya longitud puede alcanzar varios miles de nucleótidos, Figura 0.3. Más sorprendente resulta el hecho de que los moldes circulares pueden ser tan pequeños como un 13-mer<sup>36</sup>, cuyo diámetro resulta ser considerablemente inferior a los propios enzimas.

También en el caso de RNA polimerasas, se descubrió que podían utilizar como moldes pequeños DNAs circulares. En este caso la transcripción se inicia en cualquier residuo de citidina presente en la secuencia, sin necesidad de la presencia de una secuencia promotora que dirija el punto de inicio. Análogamente a la síntesis de DNA, se obtienen largas cadenas de RNA repetitivas, Figura 0.3.

En una reciente aplicación, pequeños DNAs circulares se han utilizado como moldes catalíticos para la síntesis de RNAs catalíticos (ribozimas)<sup>37</sup>.



**Figura 0.3.-** Esquema de la utilización de pequeños DNAs circulares como moldes para la obtención de cadenas largas de DNA o RNA repetitivas. Después de la replicación se puede realizar una escisión selectiva de la cadena mediante enzimas de restricción, obteniendo múltiples copias de una misma secuencia.

<sup>32</sup> Baker, T.A.; Kornberg, A. *DNA Replication* 1992, Freeman, New York.

<sup>33</sup> Eisenberg, S.; Scott, J.; Kornberg, A. *Proc. Natl. Acad. Sci. U.S.A.* 1976, 73, 3151.

<sup>34</sup> Fire, A.; Xu, S.-Q. *Proc. Natl. Acad. Sci. U.S.A.* 1995, 92, 4641.

<sup>35</sup> Liu, D.; Daubendiek, S.L.; Zillmann, M.A.; Ryan, K.; Kool, E.T. *J. Am. Chem. Soc.* 1996, 118, 1587.

<sup>36</sup> Frieden, M. Tesis Doctoral en curso, Universidad de Barcelona.

<sup>37</sup> Daubendiek, S.L.; Kool, E.T. *Nature Biotechnol.* 1997, 15, 273.

## 0.2.- MÉTODOS DE SÍNTESIS DE OLIGONUCLEÓTIDOS CÍCLICOS.

La utilización de oligonucleótidos cíclicos en cualquiera de los campos citados anteriormente queda supeditada a la disponibilidad de dichas moléculas. El interés por este tipo de estructuras promovieron, a partir de los años 80 el desarrollo de varias metodologías de síntesis que permitieran la obtención de cadenas cíclicas de DNA de forma rápida y sencilla.

La síntesis de fragmentos cíclicos de ácidos nucleicos se abordó de muy diversas formas. Las primeras tentativas, tanto en disolución<sup>38-40</sup> como en fase sólida<sup>41-44</sup>, se llevaron a cabo aplicando el método del fosfato triéster<sup>45</sup>, consistente en hacer reaccionar derivados fosfato diésteres de los nucleósidos con el grupo hidroxilo 5'-terminal de la cadena en crecimiento. Si bien este método presenta la ventaja de que se manipulan derivados muy estables, las síntesis resultaban lentas y los rendimientos de acoplamiento no cuantitativos. Incluso en el caso de estrategias en fase sólida para las que, al estar la cadena oligonucleotídica en crecimiento unida a un soporte polimérico, la fácil eliminación de productos en disolución por simple filtración permitía el uso de grandes excesos de reactivos, el utilizar el método del fosfato triéster suponía una limitación importante de cara al rendimiento final.

El principal avance en la síntesis de oligonucleótidos cíclicos se produjo con la combinación de estrategias en fase sólida sobre soporte de TentaGel (polietilenglicol-poliestireno) y la utilización del método del fosfito triéster<sup>46-48</sup> para llevar a cabo la elongación de la cadena<sup>49</sup>, Figura 0.4. La elevada efectividad del método del fosfito triéster se debe a que se utilizan derivados fosforamidito de los nucleósidos que, al ser muy reactivos, permiten acoplamientos rápidos y cuantitativos.

El método esquematizado en la Figura 0.4 presenta, sin embargo, dos inconvenientes importantes. Por una parte, el anclaje al soporte polimérico se realiza a través del grupo amino exocíclico de citosina, imponiendo de esta forma un requisito secuencial. Por otra parte, las cadenas oligonucleotídicas, ya sean lineales o cíclicas, presentan el mismo tipo de unión al

<sup>38</sup> Hsu, C.-Y.J.; Dennis, D.; Jones, R.A. *Nucleosides & Nucleotides* **1985**, 4, 377-389.

<sup>39</sup> De Vroom, E.; Broxterman, H.J.G.; Sliedregt, L.A.J.M., Van der Marel, G.A.; Van Boom, J.H. *Nucleic Acids Res.* **1988**, 16, 4607-4627.

<sup>40</sup> Rao, M.V.; Reese, C.B. *Nucleic Acids Res.* **1989**, 17, 8221-8240.

<sup>41</sup> Gait, M.J.; Singh, M.; Sheppard, R.C.; Edge, M.D.; Greene, A.R.; Heathcliffe, G.R.; Atkinson, T.C.; Newton, C.R.; Markham, A.F. *Nucleic Acids Res.* **1980**, 8, 1081-1096.

<sup>42</sup> Barbato, S.; De Napoli, L.; Mayol, L.; Piccialli, G.; Santacroce, C. *Tetrahedron Lett.* **1987**, 28, 5727-5628.

<sup>43</sup> Barbato, S.; De Napoli, L.; Mayol, L.; Piccialli, G.; Santacroce, C. *Tetrahedron Lett.* **1989**, 45, 4523-4536.

<sup>44</sup> De Napoli, L.; Montesarchio, D.; Piccialli, G.; Santacroce, C.; Mayol, L.; Galeone, A.; Messere, A. *Gazz. Chim. Ital.* **1991**, 121, 505-508.

<sup>45</sup> Gait, M.J.; Matthews, H.W.D.; Singh, M.; Sproat, B.S.; Titmas, R.C. «Chemical and Enzymatic Synthesis of Gene fragments» Ed. Gassen, H.G.; Lang, A., Verlag Chemie, Weinheim, **1982**, 1-42.

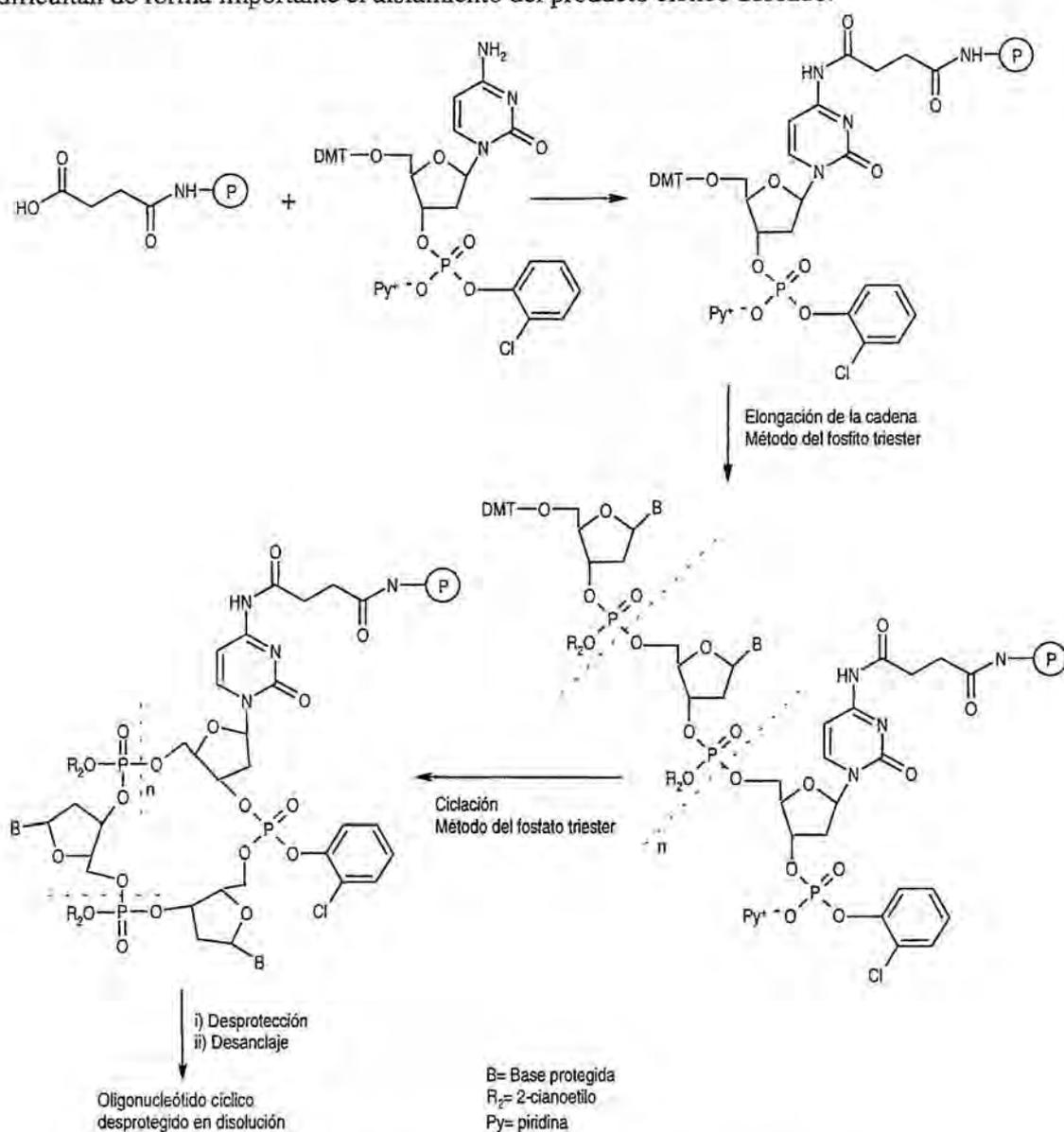
<sup>46</sup> Letsinger, R.L.; Lunsford, W.B. *J. Am. Chem. Soc.* **1976**, 98, 3655.

<sup>47</sup> Caruthers, M.H.; Barone, A.D.; Beaucage, S.L.; Dodds, D.R.; Fisher, E.F.; McBride, L.J.; Matteucci, M.; Stabinsky, Z.; Tang, J.-Y. *Methods in Enzymology* **1987**, 154, 287-313.

<sup>48</sup> Beaucage, S.L. Iyer, R.P. *Tetrahedron* **1992**, 48, 2223-1311.

<sup>49</sup> De Napoli, L.; Messere, A.; Montesarchio, D.; Piccialli, G.; Santacroce, C. *Nucleosides & Nucleotides* **1993**, 12, 21

soporte polimérico, por lo que se obtienen mezclas complejas de muchos productos que dificultan de forma importante el aislamiento del producto cíclico deseado.



**Figura 0.4.-Esquema de síntesis en fase sólida de oligonucleótidos cíclicos propuesto por Santacroce y colaboradores.**

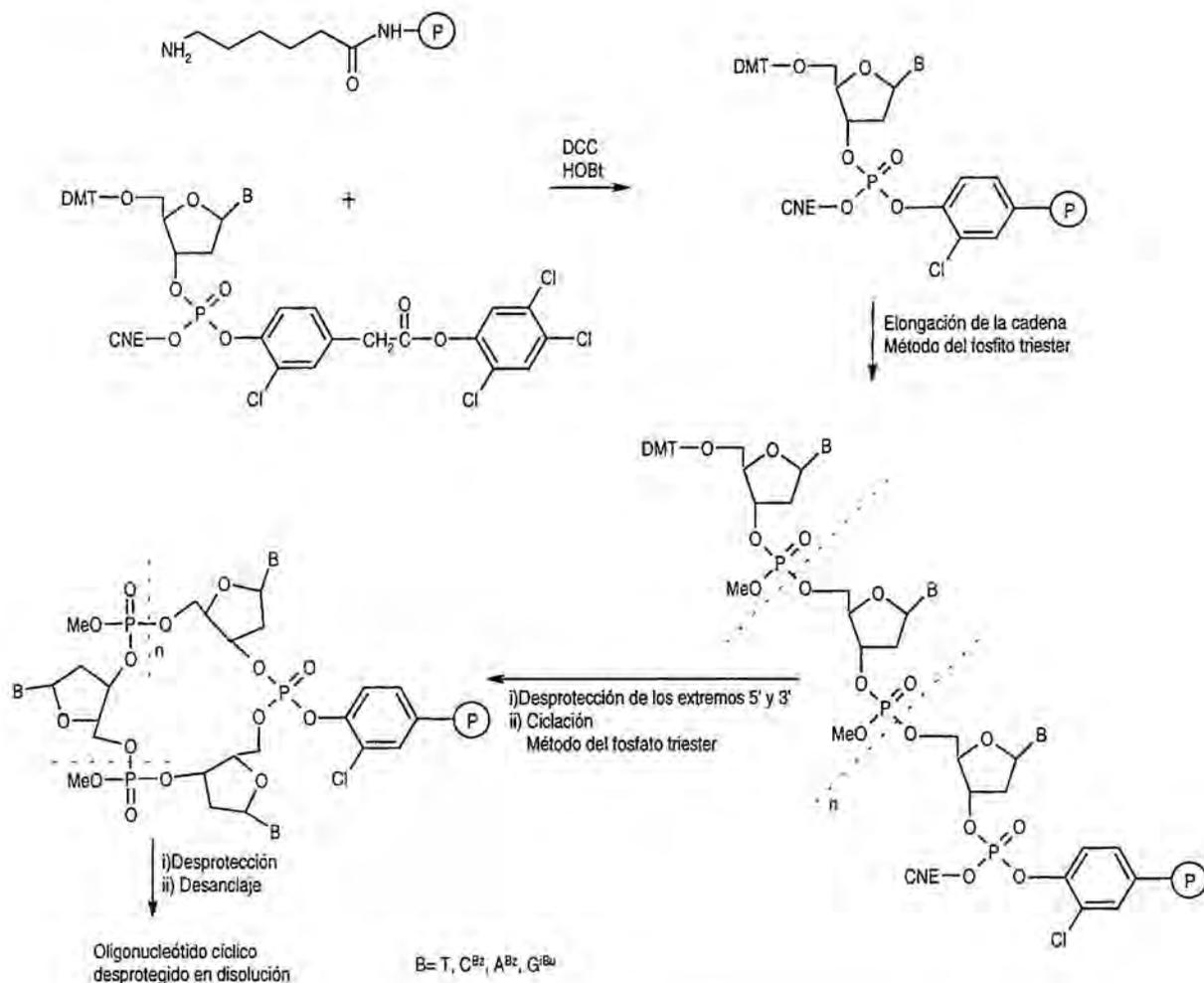
En nuestro grupo se ha desarrollado una estrategia de síntesis de oligonucleótidos cíclicos, también en fase sólida, que permite solventar ambos problemas<sup>50-51</sup>, Figura 0.5.

En este método, el anclaje del oligonucleótido a la resina tiene lugar a través del grupo fosfato del extremo 3'-terminal mediante la inclusión de un espaciador bifuncional, por lo que resulta posible sintetizar cualquier secuencia oligonucleotídica. Además, el hecho de que la unión tenga lugar a través del grupo fosfato implica que la naturaleza de dicho grupo se

<sup>50</sup> Alazzouzi, E. Tesis Doctoral, Universidad de Barcelona, 1996.

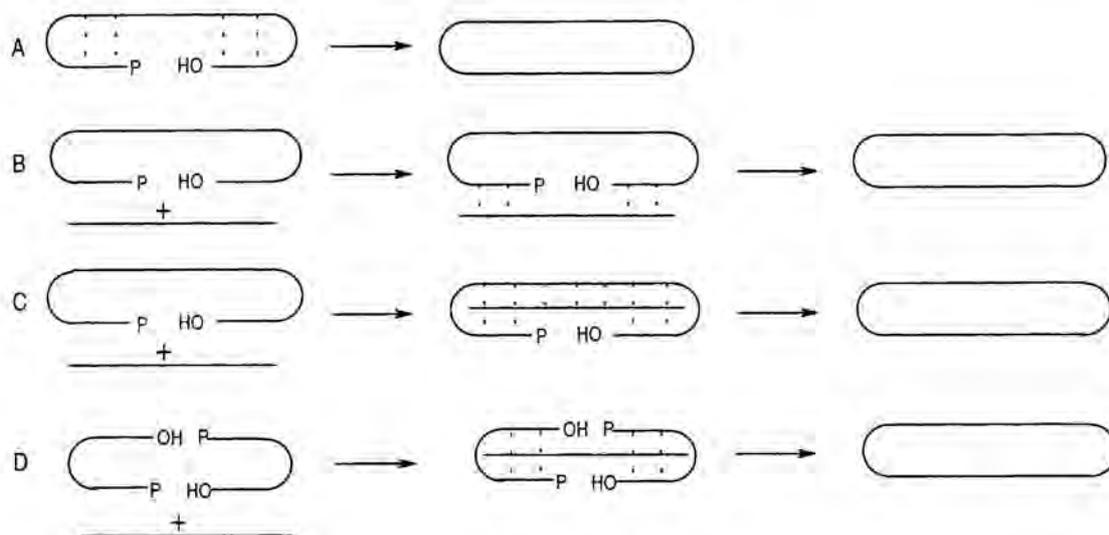
<sup>51</sup> Alazzouzi, E.; Escaja, N.; Grandas, A.; Pedroso, E. *Angew. Chem. Int. Ed. Engl.* 1997, 36, 1506-1508.

modifica cuando se lleva a cabo la reacción de ciclación, pasando de fosfato diéster a fosfato triéster. La existencia de reactivos para la desprotección selectiva de fosfato triésteres permite el desanclaje de únicamente las cadenas cíclicas formadas, lo que proporciona crudos de síntesis de mayor calidad.



**Figura 0.5.-Esquema del método de síntesis de oligonucleótidos cíclicos en fase sólida desarrollado en nuestro grupo.**

Si bien estos métodos permiten la obtención de ciclos de tamaño pequeño-medio con rendimientos aceptables y elevada pureza, pierden efectividad a medida que aumenta la longitud de la cadena. Esto es debido obviamente al factor entrópico desfavorable de la reacción de ciclación. En otras palabras, al aumentar el tamaño del ciclo cada vez resulta más difícil que los extremos 3' y 5' se encuentren y reaccionen. Por este motivo, y con objeto de obtener oligonucleótidos cíclicos de mayor tamaño, se han desarrollado estrategias de síntesis en las que, si bien la elongación de la cadena se puede realizar en fase sólida, la reacción de ciclación se lleva a cabo en disolución y utilizando, como molde para aproximar los grupos que deben reaccionar, un oligonucleótido de secuencia complementaria a los dos extremos de la cadena que se quiere ciclar, Figura 0.6.



**Figura 0.6.-**Diferentes estrategias para llevar a cabo la ciclación del precursor lineal en disolución.

En el caso A, la secuencia está diseñada de forma que sea posible la formación de puentes intramoleculares<sup>52</sup>, mientras que, en los casos B y C se utiliza un molde externo que, por formación de dúplex<sup>53</sup> (caso B) o tríplex<sup>54-55</sup> (casos C y D), aproxima los dos extremos entre los que se quiere que tenga lugar la reacción. La ciclación, en los cuatro casos, está descrita que se puede llevar a cabo indistintamente de forma química<sup>56-57</sup> o enzimática<sup>58</sup>. Siguiendo este tipo de estrategias se han sintetizado oligonucleótidos cíclicos de hasta 74 nucleótidos.

### 0.3.- OBJETIVOS.

En la presente Tesis Doctoral nos hemos planteado seguir dos líneas principales de trabajo. Por una parte, continuar con el trabajo sintético iniciado en el grupo, centrando nuestro esfuerzo en la optimización del procedimiento de síntesis de oligonucleótidos cíclicos y, por otra parte, hemos querido iniciar el estudio estructural de fragmentos cíclicos de DNA utilizando técnicas como la difracción de rayos X y la RMN.

En cuanto a los objetivos sintéticos, nos hemos centrado principalmente en dos aspectos importantes del esquema de síntesis. Por una parte, hemos considerado necesario prestar especial atención a la reacción de ciclación del precursor lineal. Puesto que en nuestro método de síntesis dicha reacción se realiza de forma química, según el método del fosfato triéster, nos hemos planteado explorar la utilización de distintos reactivos de acoplamiento para intentar aumentar los rendimientos de obtención de producto cíclico. Por otra parte, hemos abordado

<sup>52</sup> Luebke, K.J.; Dervan, P.B. *J. Am. Chem. Soc.* **1991**, 113, 7447-7448.

<sup>53</sup> Dolinnaya, N.G.; Blumenfeld, M.; Merenkova, I.N.; Oretskaya, T.S.; Krynetskaya, N.F.; Ivanovskaya, M.G.; Vasseur, M.; Shabarova, Z.A. *Nucleic Acids Res.* **1993**, 21, 5403-5407.

<sup>54</sup> Prakash, G.; Kool, E.T. *J. Am. Chem. Soc.* **1992**, 114, 3523-3527.

<sup>55</sup> Rubin, E.; Rumney, S.; Kool, E.T. *Nucleic Acids Res.* **1995**, 23, 3547-3553.

<sup>56</sup> Kanaya, E.; Yanagawa, H. *Biochemistry* **1986**, 25, 7423-7430.

<sup>57</sup> Ashley, G.A.; Kushlan, D.M. *Biochemistry* **1991**, 30, 2927-2933.

<sup>58</sup> Wemmer, E.; Benight, S. *Nucleic Acids Res.* **1985**, 13, 8611-8621.

también el estudio de otras etapas del método de síntesis, para solucionar los problemas asociados a algunas reacciones secundarias detectadas y que están principalmente relacionadas con el esquema de protección utilizado. En este sentido, hemos considerado conveniente ensayar distintos esquemas de protección de grupos fosfato. (Capítulo 1)

Por último, nos planteamos aplicar el método de síntesis a la obtención de oligonucleótidos cíclicos a gran escala. Esto nos permitiría obtener cantidades importantes de producto cíclico (del orden de miligramos) y, de esta forma, poder llevar a cabo estudios estructurales y conformacionales de distintas secuencias cíclicas, que constituyen la segunda línea principal del presente trabajo. (Capítulo 1)

El estudio estructural de fragmentos cíclicos de DNA se inició con el estudio cristalográfico de la secuencia d<pCATTTCATT>. Esta secuencia no cristalizó como *dumbbell* sino que lo hizo en forma dimérica, revelando la existencia de un nuevo motivo estructural del DNA, tipo tétrada, que se denominó bi-loop. Ante la posibilidad de que dicha estructura constituyera un nuevo motivo de carácter general con posible relevancia biológica, resultaba conveniente observar dicho motivo en disolución, por lo que se decidió llevar a cabo un estudio más amplio mediante RMN. En este sentido, la parte más importante de la presente Tesis se centra en la determinación estructural por RMN de seis secuencias, todas ellas octámeros cíclicos, entre las que se ha incluido la anteriormente mencionada, d<pCATTTCATT>.

El estudio consta de varias fases. En una primera fase se estudiaron las secuencias d<pCATTTCATT> y d<pTGCTCGCT> y se observó la existencia en disolución de un equilibrio de asociación que evidenciaba la formación de una especie dimérica. Tras estos resultados, se procedió a la determinación de la estructura, tanto de la forma dimérica como de la forma monomérica. (Capítulos 2 y 3)

La elevada similitud observada entre las dos estructuras diméricas entre sí y con la estructura determinada en el cristal hizo aumentar nuestro interés por el motivo bi-loop. Para profundizar en su estudio, en una segunda fase, se analizaron otras secuencias de octámeros cíclicos, análogas a las dos estudiadas, con objeto de delimitar posibles requisitos secuenciales para la formación de dicho motivo. No se llegó a determinar la estructura de todas las secuencias pero los datos de RMN evidenciaron en todos los casos la presencia de un producto de asociación con las características estructurales del bi-loop. (Capítulo 4)

Estos resultados y la viabilidad de estudiar el equilibrio de formación de las estructuras diméricas a partir de datos de RMN, hizo que nos planteásemos el llevar a cabo un estudio termodinámico, con el principal objetivo de determinar la estabilidad de las formas diméricas en función de la secuencia. No obstante, puesto que en algunos casos se observó que la forma monomérica presentaba una estructura tipo *dumbbell*, se evaluó también su estabilidad en función de la secuencia. (Capítulo 5)

Por último, de cara a disponer de una técnica alternativa a la RMN para detectar la formación de estructuras bi-loop en disolución se ha estudiado la viabilidad de utilizar el dicroísmo circular (DC). Esta técnica se ha utilizado tanto a nivel analítico, para detectar la presencia de estas estructuras, como para extraer datos termodinámicos que pudieran contrastarse con los obtenidos por RMN. (Capítulos 4 y 5)

---

*CAPÍTULO 1*

*SÍNTESIS EN FASE SÓLIDA DE OLIGONUCLEÓTIDOS CÍCLICOS.  
OPTIMIZACIÓN DEL MÉTODO Y APLICACIÓN A SÍNTESIS A GRAN ESCALA*

---

## 1.-SÍNTESIS EN FASE SÓLIDA DE OLIGONUCLEÓTIDOS CÍCLICOS. OPTIMIZACIÓN DEL MÉTODO DE SÍNTESIS Y APLICACIÓN A SÍNTESIS A GRAN ESCALA.

### 1.0.- INTRODUCCIÓN Y OBJETIVOS.

El procedimiento de síntesis de oligonucleótidos cíclicos en fase sólida desarrollado en nuestro grupo<sup>1</sup> se muestra de forma esquemática en la Figura 1.1.

El método consiste en el anclaje del primer nucleótido al soporte polimérico a través de un espaciador bifuncional, el ácido 3-cloro-4-hidroxifenilacético, que actúa a su vez como protector permanente del grupo fosfato del extremo 3'-terminal. Como grupos protectores semipermanentes se utilizan los grupos metilo (Me), para el fosfato 3'-terminal, y 2-cianoetilo (CNE) para los grupos fosfato internucleosídicos. La elongación de la cadena se lleva a cabo según el método del fosfito triéster<sup>2</sup>, utilizando derivados fosforamidito de los nucleósidos de origen comercial en los que los grupos amino de adenina y citosina están protegidos con el grupo benzoilo (Bz), y los de guanina con el grupo isobutirilo (<sup>i</sup>Bu). Cada acoplamiento consta de las siguientes etapas:

- i) eliminación del protector 4,4'-dimetoxitritilo del grupo hidroxilo primario del extremo 5'-terminal de la cadena nucleotídica en crecimiento anclada sobre resina, mediante un tratamiento con TCA,
- ii) acoplamiento del siguiente nucleótido por reacción del fosforamidito del nucleósido con el grupo hidroxilo libre del extremo 5'-terminal, en presencia de tetrazol como catalizador,
- iii) acetilación de los grupos hidroxilo libres que no han reaccionado durante la etapa de acoplamiento,
- iv) oxidación del enlace fosfito triéster formado a fosfato triéster.

Una vez completada la elongación del oligonucleótido lineal, que se obtiene con el grupo hidroxilo 5' terminal libre, se procede a la desprotección del grupo fosfato del extremo 3' terminal y se lleva a cabo la reacción de ciclación entre los extremos 5' y 3' terminales. Tras la ciclación del precursor lineal se eliminan los protectores de los fosfatos de la cadena y se desancla el producto de la resina mediante un tratamiento con 2-piridinaldoximato de tetrametilguanidinio (TMG), reactivo que rompe selectivamente enlaces fosfatotriéster<sup>3</sup>, tal como se muestra en la Figura 1.2.

Este protocolo de síntesis fue diseñado teniendo en cuenta una serie de criterios, entre los que destacan los siguientes:

- posibilidad de obtener cualquier secuencia oligonucleotídica, sin restricciones de secuencia ni grandes limitaciones en el tamaño del ciclo,
- elongación de la cadena utilizando el método del fosfito triéster<sup>2</sup>, que permite acoplamientos rápidos y sin reacciones secundarias, y, en lo posible, utilizando derivados fosforamidito de los nucleósidos de origen comercial (metil- y 2-cianoetilfosforamiditos),

<sup>1</sup> Alazzouzi, E. Tesis Doctoral, Universidad de Barcelona, 1996.

<sup>2</sup> Caruthers, M.H.; Barone, A.D.; Beaucage, S.L.; Dodds, D.R.; Fisher, E.F.; McBride, L.J. Matteucci, M.; Stabinsky, Z.; Tang, J.-Y. *Methods in Enzymology* 1987, 154, 287-313.

<sup>3</sup> Reese, C.B.; Titmas, R.C.; Yau, L. *Tetrahedron Lett.* 1978, 30, 2727-2730.

- esquema de protección adecuado para que, en la etapa de ciclación del precursor lineal, la formación de la unión internucleosídica tuviera lugar de forma regioselectiva,
- separación selectiva de las cadenas cíclicas del resto de productos de naturaleza oligonucleotídica que puedan generarse durante la síntesis.

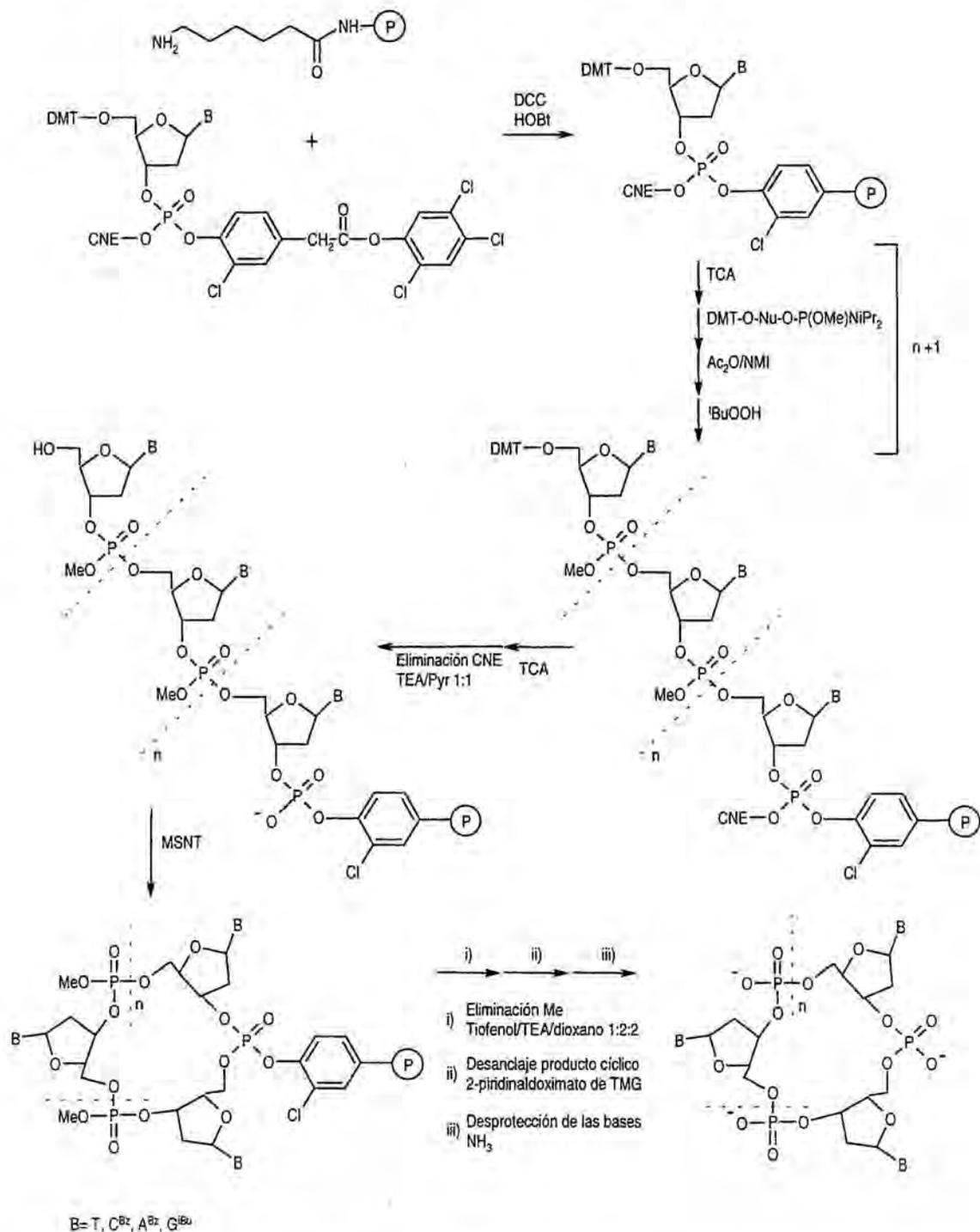
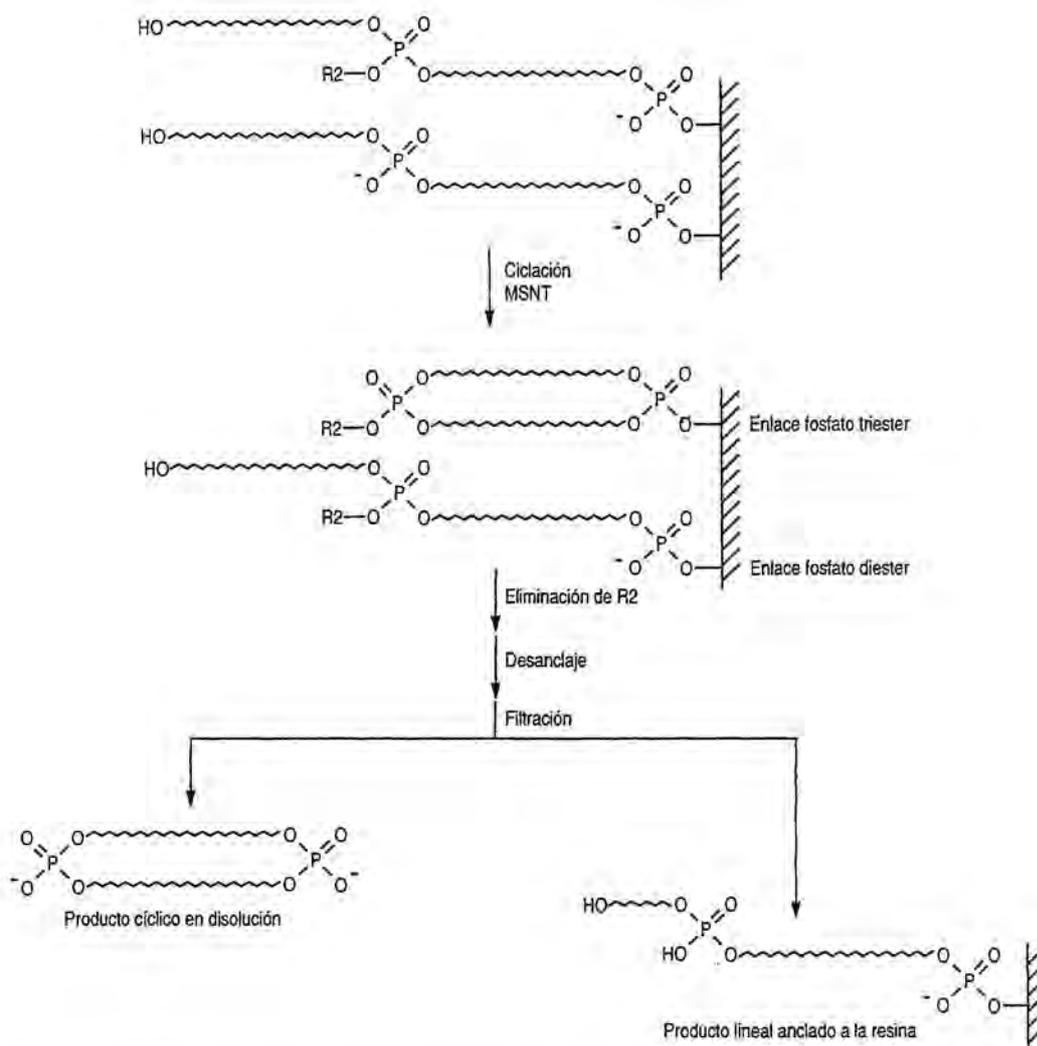


Figura 1.1.-Esquema del método de síntesis en fase sólida para la obtención de oligonucleótidos cíclicos.



**Figura 1.2.-** Esquema sintético del desanclaje selectivo de los productos cíclicos.

Si bien muchos de los requisitos inicialmente propuestos se cumplen satisfactoriamente, y se ha descrito la síntesis de oligonucleótidos cíclicos de tamaño entre 2 y 16-mer<sup>1</sup>, este procedimiento sintético presenta también ciertos inconvenientes. Los dos problemas principales, que se detallan a continuación, derivan del esquema de protección elegido y pueden repercutir de forma negativa en el rendimiento de obtención del producto cíclico.

• **Desprotección no selectiva del fosfato 3'-terminal.**

En este esquema sintético se ha utilizado el grupo 2-cianoetilo como protector semipermanente del fosfato del extremo 3'-terminal y el grupo metilo como protector de los grupos fosfato del resto de la cadena.

El mecanismo y condiciones de eliminación de ambos grupos son distintos, Figura 1.3. El grupo 2-cianoetilo se elimina en presencia de base (TEA o DBU) mediante un mecanismo de  $\beta$ -eliminación, mientras que la desprotección de fosfatos de metilo tiene lugar mediante una reacción de sustitución nucleófila, tipo  $S_N2$ , por ataque de algún nucleófilo blando ( $RS^-$ ,  $Br^-$  o  $I^-$ ) sobre el carbono del grupo metilo. Esta última reacción muestra una marcada preferencia  $Me > RCH_2 > R_2CH$ .

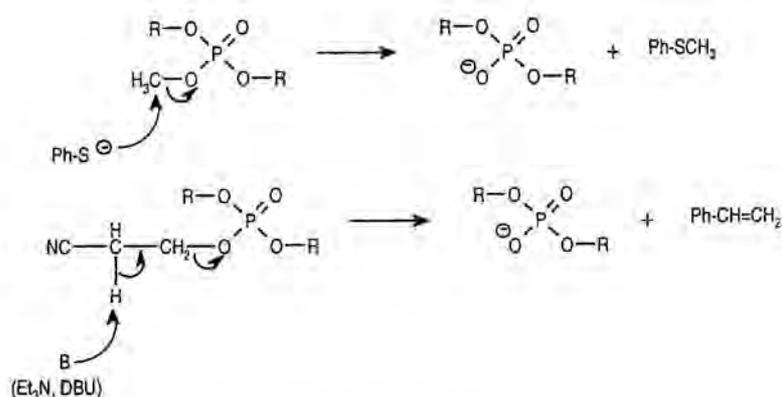


Figura 1.3.- Mecanismo de eliminación de los grupos metilo y 2-cianoetilo.

Estudios realizados en nuestro grupo<sup>1</sup> demuestran, sin embargo, que ambos grupos no son totalmente ortogonales. La TEA, que tiene un cierto carácter nucleófilo, produce la desprotección parcial de metilfosfotriesteres durante la etapa de eliminación del grupo CNE. Aunque se intentó ajustar al máximo el tiempo del tratamiento con TEA/piridina, el tiempo requerido para la total eliminación del grupo CNE podía implicar una pérdida de 6-7% de grupos Me.

Esta eliminación no selectiva del protector del fosfato 3'-terminal conlleva que, durante la etapa de ciclación, los grupos hidroxilo 5'-terminales puedan reaccionar indistintamente con el fosfato 3'-terminal o con aquellos grupos fosfato de la cadena que hayan perdido el metilo, tal como se ilustra en la Figura 1.4.

Los productos formados por condensación del hidroxilo 5'-terminal con fosfatos de la cadena darían lugar a ciclos de menor tamaño que, al estar unidos a la cadena en forma de fosfatotriesteres, podrían ser también, en principio, atacados por el oximato empleado en la reacción de desanclaje. La ausencia de un buen grupo saliente, como en el caso del fosfato triester de *o*-clorofenilo, haría, no obstante, que esta reacción no estuviera favorecida y tendría lugar de forma mucho más lenta y poco selectiva. Tal vez la calidad del crudo de síntesis no se vería muy afectada, pero no hay duda que esta reacción secundaria implicaría una disminución sustancial del rendimiento de obtención del producto cíclico deseado.

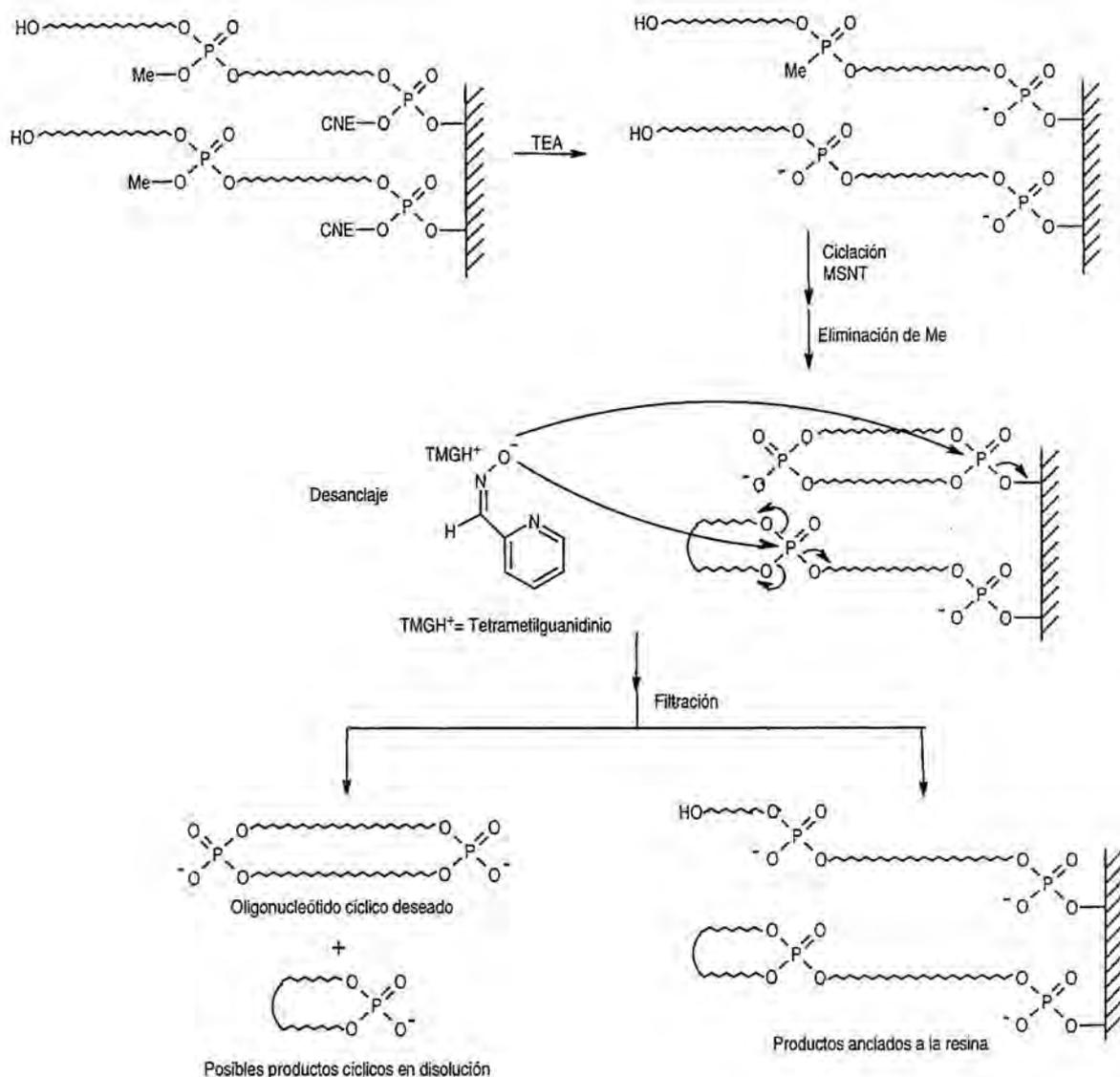
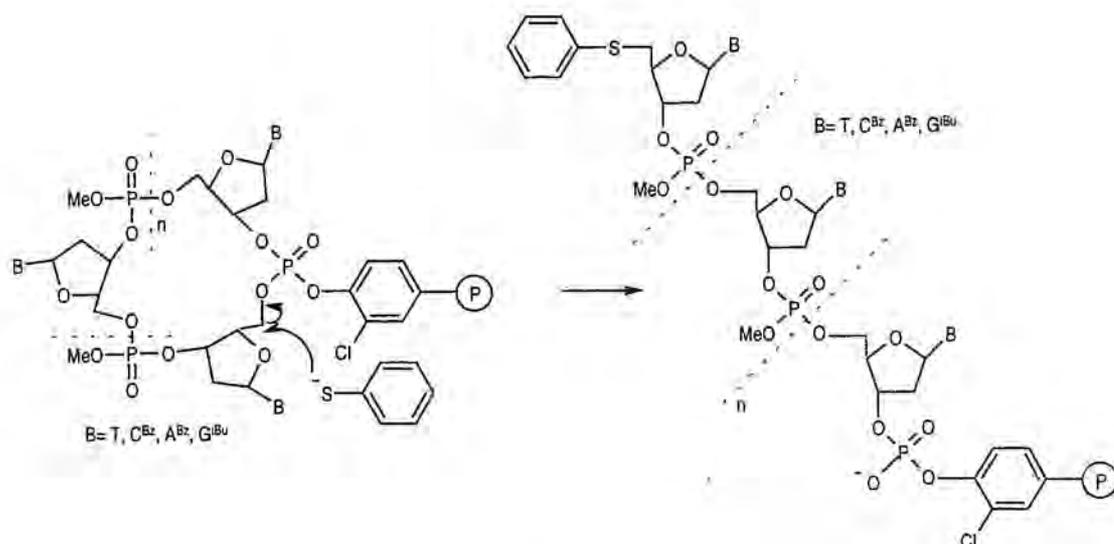


Figura 1.4.- Productos secundarios procedentes de la desprotección parcial de fosfatos de la cadena.

• Pérdida de producto cíclico.

Otro problema importante, íntimamente relacionado también con el esquema de protección utilizado, es una reacción secundaria causada por el tiofenol, reactivo habitualmente empleado para la eliminación de los grupos metilo.

Como ya se ha comentado anteriormente, la desprotección de los fosfatos de la cadena se realiza una vez formado el ciclo. En este estadio, se ha podido observar que, además de los carbonos metílicos, otra posición resulta especialmente sensible al ataque del tiofenolato: el carbono metilénico 5' unido al fosfato de *o*-clorofenilo, Figura 1.5. Este ataque produce la apertura del producto cíclico dando lugar nuevamente a la cadena oligonucleotídica lineal con el grupo feniltioéter en posición 5'.



**Figura 1.5.-** Ataque del tiofenolato sobre el carbono 5' que origina la apertura del producto cíclico.

Esta reacción de ruptura del enlace internucleosídico, que ya había sido descrita con anterioridad en el caso de un dinucleótido con un grupo arilo como protector del fosfato<sup>4</sup>, estaría en este caso favorecida por la presencia del grupo *o*-clorofenilo. Los estudios cinéticos<sup>1</sup> muestran que esta reacción no es de ningún modo despreciable, ya que presenta un  $t_{1/2} = 50$  min y habitualmente para la eliminación de los grupos metilo se realizaban dos tratamientos con tiofenol/TEA de 30 min. Aunque el tiempo de reacción se disminuyó a dos tratamientos de 10 minutos para minimizar los efectos de esta reacción secundaria, era necesario intentar evitar totalmente dicha reacción.

Así pues, se decidió realizar una revisión del esquema de protección de los grupos fosfato en la búsqueda de alternativas, compatibles con el resto de protectores utilizados en la síntesis (DMT y *o*-clorofenilo), que nos permitieran evitar al máximo la existencia de reacciones secundarias como las anteriormente mencionadas.

Este estudio se incluye dentro de los objetivos del presente capítulo, que tiene como objeto final optimizar el método de síntesis y conseguir mayores rendimientos de obtención de producto cíclico, de cara a su posterior aplicación a síntesis a gran escala.

Además de la búsqueda de un esquema de protección más adecuado, nos hemos planteado la necesidad de estudiar una de las etapas clave, como es la reacción de ciclación. Esta reacción, que ya de por sí presenta un factor entrópico desfavorable, puede verse afectada de forma importante por la longitud de la cadena y por efectos de secuencia. Por este motivo, resulta muy importante disponer de datos que nos permitan optimizar las condiciones en las que debe tener lugar la reacción. Básicamente nos hemos centrado en estudiar la viabilidad de otros reactivos de acoplamiento que puedan funcionar mejor que el habitualmente utilizado, el MSNT, y en llevar a cabo estudios cinéticos para optimizar el tiempo de reacción.

Por último, nos planteamos realizar varias síntesis a gran escala de distintas secuencias oligonucleotídicas para evaluar la aplicabilidad del método cuando se trabaja con grandes cantidades de resina y elevada funcionalización.

<sup>4</sup> Reese, C.B.; Titmas, R.C.; Valente, L. *J.C.S. Perkin I* **1981**, 2451-2455.

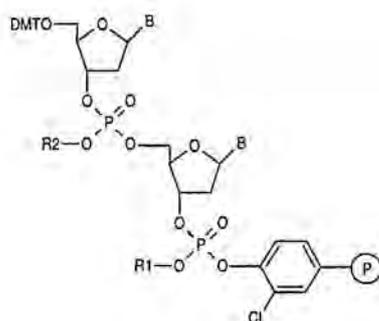
## 1.1.- OPTIMIZACIÓN DEL MÉTODO DE SÍNTESIS.

### 1.1.1.- DISEÑO DE UN ESQUEMA DE PROTECCIÓN ORTOGONAL.

La mayor limitación a la hora de buscar alternativas al esquema de protección descrito radica, además de en la compatibilidad de los distintos grupos protectores de fosfato con el protocolo de síntesis y el resto de grupos protectores, en la disponibilidad comercial de los derivados fosforamiditos de los nucleósidos. Aunque estos derivados son estables y fáciles de obtener<sup>5</sup>, la aplicación del método para llevar a cabo síntesis a gran escala requiere cantidades importantes de reactivos. Por este motivo, hemos intentado utilizar al máximo combinaciones que permitieran el uso de metil- y 2-cianoetilfosforamiditos para la elongación de la cadena y, en caso de considerar un grupo protector distinto, éste debería utilizarse como protector del fosfato 3'-terminal.

En la Tabla 1.1 se muestran los distintos esquemas de protección que hemos ido ensayando a lo largo de este trabajo.

**Tabla 1.1.-** Esquemas de protección propuestos para los distintos grupos fosfato.



Esquema de protección	Protector extremo 3'-terminal: R1	Protector grupos fosfato cadena: R2
1	CNE	Me
2	Me	CNE
3	-	CNE

El esquema 1 es el utilizado hasta el momento en nuestro grupo, por lo que la evaluación de los otros esquemas alternativos se realizará por comparación de los rendimientos obtenidos en cada caso con los obtenidos para el esquema 1.

En el esquema 2 se han invertido los grupos Me y CNE, con el objetivo principal de evitar la presencia de tiofenol una vez formado el producto cíclico. Si bien este esquema nos permite eliminar la reacción secundaria causada por el tiofenol sobre el producto cíclico, será necesario encontrar unas condiciones de desprotección adecuadas, que permitan la eliminación selectiva del grupo Me en presencia de CNE.

<sup>5</sup> Beaucage, S.L., Iyer, R.P. *Tetrahedron* **1992**, 24, 371-391.

El tercer esquema considerado consiste en desproteger el fosfato 3'-terminal antes de la elongación de la cadena. Diversos trabajos descritos en la literatura demuestran que la desprotección ocasional de los fosfatos de la cadena durante la síntesis de oligonucleótidos no impide que la síntesis evolucione de forma normal, sin que se observe ni disminución en el rendimiento de obtención ni merma en la pureza del producto lineal<sup>6-8</sup>. Si esto fuera así, trabajar con el fosfato 3'-terminal desprotegido no supondría un problema para la elongación de la cadena y, sin embargo, podría presentar una serie de ventajas. Por una parte, se eliminaría la etapa de desprotección de dicho grupo una vez completada la síntesis del oligonucleótido lineal, por lo que el resto de fosfatos de la cadena se mantendrían con toda seguridad inalterados. Por otra parte, podríamos utilizar únicamente derivados cianoetilfosforamiditos, evitando por completo la utilización de tiofenol.

#### 1.1.1.1.- Esquema de protección 2: $R_1 = \text{Me}$ , $R_2 = \text{CNE}$ .

El esquema de protección 2, donde se utiliza el grupo Me como protector del fosfato del extremo 3'-terminal y el grupo CNE como protector de los fosfatos internucleosídicos, presenta la ventaja de que se elimina la presencia de tiofenol una vez se ha formado el producto cíclico y se evita, por tanto, la reacción secundaria anteriormente descrita. La eliminación del grupo Me, sin embargo, no se puede llevar a cabo de la forma habitual (tiofenol/TEA/dioxano 1:2:2) ya que la presencia de TEA provocaría, a su vez, la eliminación de los grupos CNE. Por este motivo, resulta necesario encontrar unas nuevas condiciones de eliminación de los grupos metilo que sean compatibles con la presencia de CNE.

En este sentido se han seguido dos estrategias. Por una parte se han ensayado condiciones de eliminación en medio básico más suave, bien utilizando bases menos fuertes que la TEA o bien utilizando un exceso menor de dicha base, manteniéndose en ambos casos el tiofenol como nucleófilo. Por otra parte, se han ensayado otros nucleófilos alternativos, descritos también para la desalquilación de fosfatotriesteres.

Para llevar a cabo el estudio se han utilizado como modelos los dinucleótidos de timidina que se muestran en la Figura 1.6.

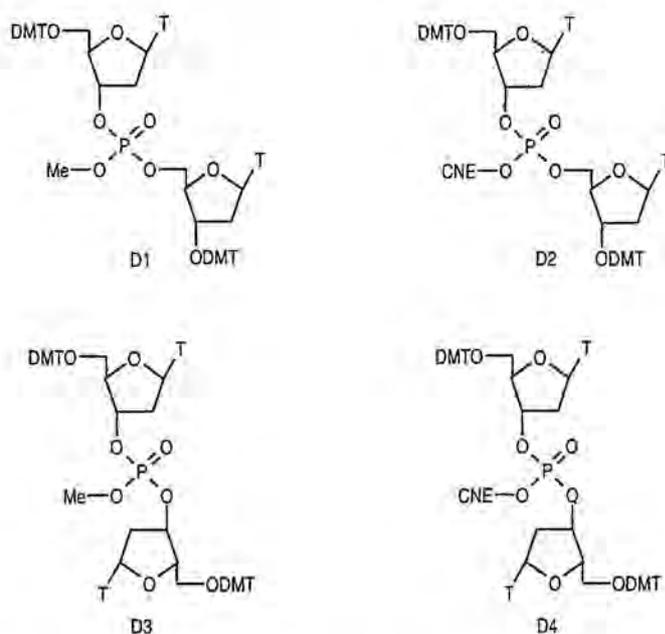
Los modelos que mejor reproducen nuestras moléculas son los dinucleótidos D1 y D2, ya que presentan la unión natural 3'-5'. No obstante, creímos conveniente sintetizar también los correspondientes dinucleótidos simétricos 3'-3' D3 y D4, ya que resultan útiles para llevar a cabo estudios por <sup>31</sup>P-RMN, porque la señal de fosfato es la de un único estereoisómero.

La síntesis de los modelos se encuentra descrita en los apartados 2.1.3, 2.1.4, 2.1.5 y 2.1.6 de Materiales y Métodos.

<sup>6</sup> Rushdi, H.A.; Singman, C.N.; Guangrong, Z.; Letsinger, R.L. *Nucleic Acids Res.* **1991**, 19, 1527-1532

<sup>7</sup> Kuijpers, W. Tesis Doctoral, Rijksuniversiteit de Leiden **1993**, Capítulo I, 2.

<sup>8</sup> Kowalski, D. *Anal. Biochem.* **1979**, 93, 346-354.



**Figura 1.6.-** Dinucleótidos de timidina utilizados como modelos. D1 y D2 corresponden a dinucleótidos 3'-5'-fosfotriester (dínucleótidos asimétricos) y D3 y D4 a dinucleótidos 3'-3'-fosfotriester (dínucleótidos simétricos).

#### 1.1.1.1.1.- Eliminación del grupo Me con tiofenol como nucleófilo y distintas bases.

Para llevar a cabo este estudio se utilizaron como modelos los dinucleótidos asimétricos D1 y D2. Los ensayos se realizaron sobre 5 mg de producto ( $\sim 4.2 \mu\text{mol}$ ).

Con el objeto de evaluar la idoneidad de cada una de las bases, se realizaron unos primeros ensayos en los que la reacción de eliminación de los grupos Me y CNE se siguió por TLC (sistema C), por lo que los valores que se muestran en las Tablas 2 y 4 son aproximados.

- Tiofenol y distintos equivalentes de TEA.

En primer lugar se determinó cómo afectaba la proporción de TEA a la reacción de eliminación del grupo Me. Este estudio nos permitiría fijar el número mínimo de equivalentes necesarios para que la reacción tuviera lugar en un tiempo razonable. Para ello se mantuvo la concentración de tiofenol utilizada habitualmente, tiofenol/dioxano 1:4 v/v y se varió el número de equivalentes de TEA respecto a los equivalentes de dinucleótido, o lo que es lo mismo, respecto a grupos Me. En este sentido, es importante tener en cuenta que, en las condiciones habitualmente utilizadas tiofenol/TEA/dioxano 1:2:2 v/v/v, el exceso de TEA utilizado respecto a grupos Me era  $\sim 150$ -200 eq. Los resultados obtenidos se muestran en la Tabla 1.2.

**Tabla 1.2.- Eliminación del grupo Me utilizando tiofenol como nucleófilo y distintos equivalentes de TEA. Modelo utilizado: dinucleótido D1.**

Condiciones de reacción	TEA eq	Tiempo de reacción h	% Eliminación Me
Tiofenol/dioxano 1:4 v/v	15	0.5	50
		1	50-60
		5	70
		7	80
		24	-100
	50	0.5	60
		1	80
		2	>90
		5	>95
		7	-100
	100	0.5	80
		1	>95
		2	-100

Los valores de porcentaje de eliminación del grupo Me, aunque orientativos, muestran claramente que la disminución del número de equivalentes de TEA aumenta drásticamente el tiempo de reacción necesario para eliminar cuantitativamente los grupos Me. Esto es debido seguramente a la disminución de la concentración de ión tiofenolato en el medio. Este resultado pone de manifiesto que no resulta viable la opción de disminuir la proporción de TEA y que las condiciones óptimas, cuando se utiliza TEA como base, serían las habitualmente utilizadas: tiofenol/TEA/dioxano 1:2:2 v/v/v.

El hecho de que fuera necesaria una presencia importante de TEA en el medio para que la reacción tuviera lugar de forma rápida, no era un buen presagio de cara a la utilización de bases menos fuertes. Sin embargo, decidimos realizar algunos ensayos, que son los que se describen a continuación.

- *Tiofenol y distintas bases.*

Para llevar a cabo este estudio se realizaron ensayos con distintas bases, que tienen un  $pK_a$  inferior al de la TEA pero suficientemente alto como para generar el ión tiofenolato ( $pK_a$  tiofenol = 6.62). Los valores de  $pK_a$  correspondientes a los ácidos conjugados de las distintas bases utilizadas se muestran en la Tabla 1.3.

**Tabla 1.3.- Valores de  $pK_a$  correspondientes a los ácidos conjugados de las bases utilizadas.**

Base	$pK_a$
Lutidina	6.64
Colidina	7.25
N-metilimidazol (NMI)	7.25
N-metilmorfolina (NMM)	8.65
Trietilamina (TEA)	10.85

Como referencia para realizar este estudio se utilizó la combinación tiofenol/TEA en condiciones en las que la eliminación total del grupo Me tenía lugar en ~1 h. Para facilitar la comparación de los resultados se utilizó una relación de reactivos en volumen. Los resultados obtenidos se muestran en la Tabla 1.4.

**Tabla 1.4.- Eliminación del grupo Me utilizando tiofenol como nucleófilo y distintas bases. Modelo utilizado: dinucleótido D1.**

	Condiciones de reacción	Tiempo de reacción h	% Eliminación Me
1	Tiofenol/TEA/dioxano 1 : 0.5 : 3.5 v/v/v	1	100
2	Tiofenol/NMM/dioxano 1 : 0.5 : 3.5 v/v/v	2	<40
3	Tiofenol/NMI/dioxano 1 : 0.5 : 3.5 v/v/v	2	<20
4	Tiofenol/NMM/dioxano 1 : 1 : 3 v/v/v	2	<50
5	Tiofenol/NMI/dioxano 1 : 1 : 3 v/v/v	2	<30
6	Tiofenol/NMM/DMF 1 : 2 : 2 v/v/v	2	-100
7	Tiofenol/colidina/DMF 1 : 2 : 2 v/v/v	2	<50
8	Tiofenol/lutidina/DMF 1 : 2 : 2 v/v/v	2	<40

Estos ensayos ponen de manifiesto que, en efecto, a medida que disminuye la basicidad (comparar condiciones 1, 2 y 3) la eliminación de los grupos Me se hace considerablemente más lenta. Incluso en los casos 4 y 5, en los que se ha doblado la proporción de base respecto a TEA, la eliminación resulta excesivamente lenta. De cara a mejorar estos resultados se pensó en utilizar como disolvente la DMF, ya que en nuestro grupo se contaba con la experiencia de que este tipo de reacciones tenían lugar de forma más rápida en este disolvente. Las tres últimas entradas de la Tabla 1.4 corresponden a los casos en los que se ha sustituido dioxano por DMF. En el caso de la colidina y lutidina, no se obtienen buenos resultados, pero esto puede ser debido a la débil basicidad que presentan ambas bases. Sin embargo, cuando se emplea NMM en DMF la reacción sí que tiene lugar en un tiempo razonable, por lo que se pensó que podría ser una buena opción.

De cara a evaluar la viabilidad de utilizar NMM como base alternativa a la TEA era necesario estudiar la estabilidad del grupo CNE a dicha base. Para ello se realizó un ensayo utilizando tiofenol/NMM/DMF 1 : 2 : 2 (entrada 6) y tomando como referencia la combinación tiofenol/TEA (entrada 1). El ensayo se realizó sobre el dinucleótido D2 y los resultados se muestran en la Tabla 1.5.

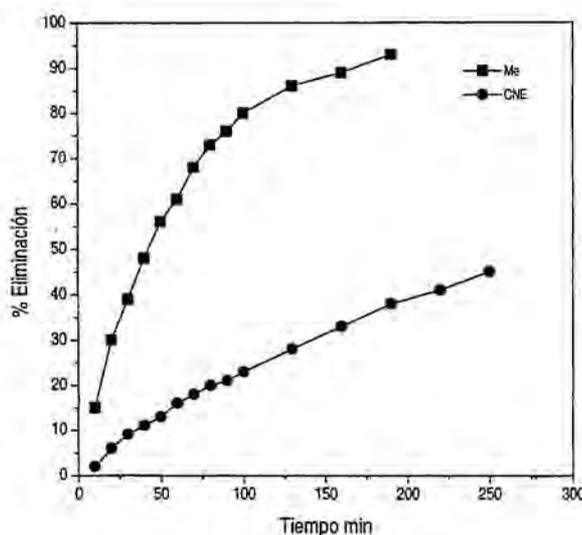
**Tabla 1.5.- Estabilidad del grupo CNE en las mejores condiciones encontradas para la eliminación del grupo Me. Modelo: Dinucleótido D2.**

Condiciones de reacción	Tiempo de reacción h	% Eliminación CNE
Tiofenol/TEA/dioxano 1:0.5:3.5 v/v/v	1	30
Tiofenol/NMM/DMF 1:2:2 v/v/v	1 3.5	<5 <15

Los resultados obtenidos ponen de manifiesto la dificultad de encontrar unas condiciones que permitan la desprotección selectiva de Me en presencia de CNE, al menos si

nos limitamos únicamente a variar la basicidad y número de equivalentes de la base. Sin embargo, se confirma que la NMM podría ser una buena alternativa a la TEA. La eliminación del grupo Me en tiofenol/NMM/DMF 1 : 2 : 2 v/v/v tiene lugar en un tiempo razonable (2h) y la eliminación parcial de grupos CNE es bastante inferior a la que se produce cuando se utiliza TEA.

En estas condiciones, y para cuantificar de forma fiable la extensión de la eliminación de los grupos Me y CNE en las mismas condiciones, se realizó una cinética seguida por  $^{31}\text{P}$ -RMN. En este caso, se utilizaron como modelos los dinucleótidos D3 y D4, ya que presentan un único diastereómero y, por tanto, una única señal. Los resultados obtenidos a partir de la integración de las señales de resonancia se muestran en la Figura 1.7.



**Figura 1.7.-** Representación gráfica de la cinética de eliminación de los grupos Me y CNE en tiofenol/NMM/DMF<sub>a</sub> 1:2:2.

La cinética seguida por  $^{31}\text{P}$ -RMN muestra que, en las condiciones utilizadas, no sólo se confirma el problema de la falta de ortogonalidad entre los grupos Me y CNE, sino que se obtienen valores de eliminación de CNE superiores a los esperados. Aunque a nivel cualitativo, por TLC, parecía que estas condiciones podrían resultar adecuadas, el porcentaje de eliminación de grupos CNE detectado en los experimentos de RMN alcanza un ~30% en 3h. Este porcentaje es, además, superior al porcentaje de eliminación de grupos Me durante la desprotección del fosfato 3'-terminal según el esquema 1, que es del orden del 10%.

A pesar de estos resultados, y teniendo en cuenta que el modelo no reproduce la unión natural 3'-5' ni la situación real en que se encuentra el oligonucleótido lineal protegido y anclado al soporte polimérico, decidimos utilizar este esquema 2, y las condiciones de eliminación del grupo Me descritas para la NMM, para llevar a cabo una síntesis y poder comparar los resultados con los obtenidos para el esquema 1.

### 1.1.1.1.2.- Eliminación del grupo Me con otros nucleófilos alternativos al tiofenol.

Como alternativas a la utilización del tiofenol se han considerado otros nucleófilos blandos descritos como eficaces para la desprotección de fosfatos de alquilo. En este sentido, se ha elegido el 2-mercaptobenzotiazol (2-MBT), que aunque está descrito que la eliminación del grupo metilo es más lenta que con tiofenol<sup>9</sup>, presenta la ventaja de que es inodoro. Las condiciones de reacción óptimas descritas preconizan utilizar una disolución equimolar de 2-MBT y de etildiisopropilamina (DIEA) en 1-metil-2-pirrolidona (NMP). No obstante, puesto que necesitamos que las condiciones de reacción sean compatibles con la presencia de grupos CNE, se han utilizado bases más débiles, Tabla 1.6.

También se han considerado como reactivos alternativos el bromuro de trimetilsililo ((CH<sub>3</sub>)<sub>3</sub>SiBr), descrito para la desalquilación de ésteres de fosfonato<sup>10</sup>, y el ioduro potásico (KI), que presenta la ventaja de que no necesita adición de base.

Los ensayos realizados con los distintos nucleófilos se han realizado también sobre 5 mg de producto (~4.2 μmol), utilizando como modelos los dinucleótidos asimétricos D1 y D2. Los resultados que se muestran son también cualitativos ya que las reacciones se han seguido por TLC.

**Tabla 1.6.- Eliminación del grupo Me con distintos nucleófilos. Ensayos realizados utilizando como modelo el dinucleótido D1.**

Condiciones de reacción	Nucleófilo -eq	Base -eq	Tiempo de reacción h	% Eliminación Me
5% (CH <sub>3</sub> ) <sub>3</sub> SiBr en piridina	25	-	0.5	100
1% (CH <sub>3</sub> ) <sub>3</sub> SiBr en piridina	10	-	0.25	-100
	5	-	0.25	-100
0.1 M KI en acetona/ACN 1:1 v/v	25	-	24	<10
0.1 M KI en DMF	25	-	24	<10
0.1 M KI en MeOH/ACN 1:1 v/v	25	-	24	<10
2-MBT/colidina/NMP 1.5 M	100	100	18	<50
2-MBT/NMM/NMP 1.5 M	100	100	1	<20

Los resultados obtenidos, Tabla 1.6, muestran que los porcentajes de eliminación de grupos Me utilizando KI o 2-MBT como reactivos son excesivamente bajos y se requieren tiempos largos de reacción. Habría sido necesario realizar otros ensayos aumentando el número de equivalentes de KI, o bien, en el caso del 2-MBT, aumentando también la proporción de base, aunque previsiblemente esto conduciría a eliminar simultáneamente grupos CNE. Sin embargo, dados los buenos resultados obtenidos con (CH<sub>3</sub>)<sub>3</sub>SiBr, decidimos probar únicamente este reactivo y no se optimizaron las condiciones para los otros dos.

Como en el apartado anterior, una vez encontradas unas buenas condiciones para la eliminación del Me era necesario evaluar la estabilidad del grupo CNE a dichas condiciones.

<sup>9</sup> Andrus, A.; Beaucage, S.L. *Tetrahedron Lett.* **1988**, 29, 5479-5482.

<sup>10</sup> Fieser, M.; Danheiser, R.L.; Roush, W. «Fieser's Reagents for Organic Synthesis» **1981**, 9, 74.

Para ello se realizaron, sobre el dinucleótido D2, los ensayos que figuran en la Tabla 1.7.

Tabla 1.7.- Estabilidad del grupo CNE a  $(\text{CH}_3)_3\text{SiBr}$ . Modelo utilizado: dinucleótido D2.

Condiciones de reacción	Nucleófilo, eq	Tiempo de reacción, h	% Eliminación CNE
5% $(\text{CH}_3)_3\text{SiBr}$ en piridina	25	0.5	50
1% $(\text{CH}_3)_3\text{SiBr}$ en piridina	10	0.25	0
	5	0.25	0

Los resultados sobre los compuestos modelo muestran que una disolución al 1% de  $(\text{CH}_3)_3\text{SiBr}$  en piridina, no sólo es un reactivo eficaz para la eliminación del grupo Me sino que además garantiza que los grupos CNE se mantengan inalterados. Como en el apartado anterior, se decidió llevar a cabo un ensayo por  $^{31}\text{P}$ -RMN, para confirmar los resultados obtenidos por TLC.

En este caso, sin embargo, optamos por utilizar unos modelos que reprodujeran al máximo la situación real, por lo que éstos debían estar anclados al soporte polimérico que se iba a utilizar para la síntesis (TentaGel, co-polímero polietilenglicol-poliestireno). Los modelos se muestran en la Figura 1.8.

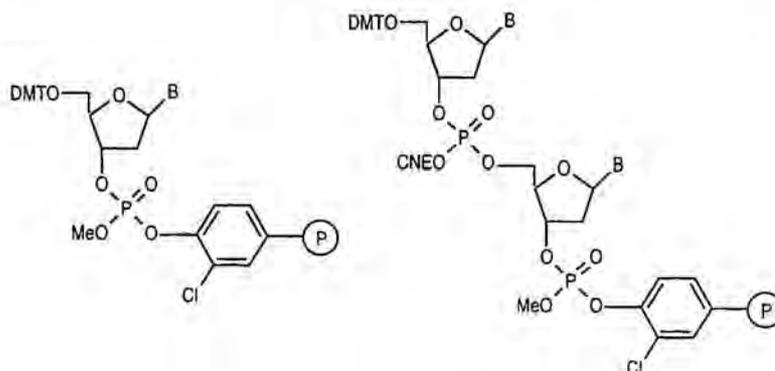


Figura 1.8.- Modelos utilizados para el estudio cinético por  $^{31}\text{P}$ -RMN en fase gel.

El estudio cinético se realizó utilizando como disolvente piridina deuterada. Sorprendentemente, con ninguno de los dos modelos se obtuvieron los resultados esperados. Los espectros de  $^{31}\text{P}$ -RMN no muestran ninguna variación en las señales que indiquen, ni la eliminación total del grupo Me (paso de fosfato triéster a fosfato diéster) ni la estabilidad del grupo CNE. En cambio, sí que se observan cambios en la apariencia del soporte polimérico que, al final del experimento, se encuentra significativamente menos inflado. Se pensó que el  $(\text{CH}_3)_3\text{SiBr}$  podría haber reaccionado con las cadenas de polietilenglicol antes de hacerlo con el fosfato de Me. Un dato a favor de esta hipótesis es el hecho de que, después del tratamiento con este reactivo, se detectó la presencia de grupos DMT en los filtrados procedentes de la reacción. Los grupos DMT en disolución podrían corresponder al nucleótido/dinucleótido (según el modelo) que, por degradación química en algún punto de la cadena de polietilenglicol, se han separado del soporte y han pasado a la disolución. De ser cierto que la incompatibilidad de este reactivo es con la matriz polimérica utilizada, podría pensarse en utilizar una de poliestireno-

divinilbenceno (PS-DVB). Sin embargo, en trabajos anteriores ya se había demostrado que los rendimientos de obtención de oligonucleótidos cíclicos son francamente superiores si se emplea un soporte polimérico de TentaGel<sup>11</sup>.

### 1.1.1.2.- Esquema de protección 3: R1 = CNE, R2 = CNE, con eliminación de R1 antes de la elongación de la cadena.

Como ya se ha comentado anteriormente, se encontraron diversas referencias en la literatura al hecho de que la desprotección parcial de los fosfatos internucleotídicos, durante la elongación de la cadena, no implicaba una disminución ni del rendimiento de obtención ni de la pureza del oligonucleótido lineal<sup>6-8</sup>.

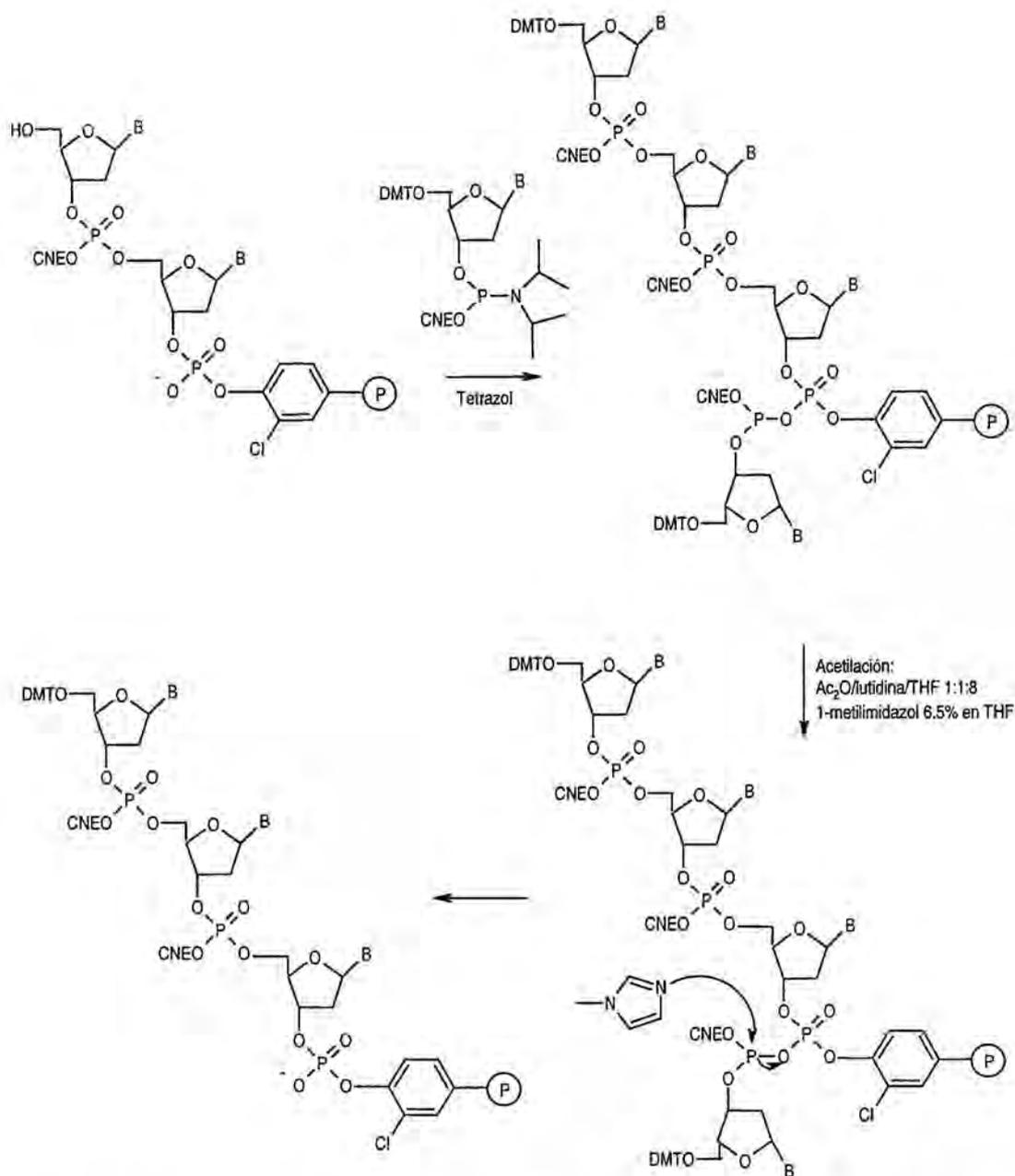
De hecho, la única reacción que podría tener lugar durante la elongación de la cadena, al mantener el fosfato 3'-terminal desprotegido, es un acoplamiento del fosforamidito sobre esta posición. Esta reacción daría lugar a la formación de un anhídrido mixto fosforoso-fosfórico, especie muy reactiva que, en presencia de 1-metilimidazol (reactivo presente en la mezcla utilizada para la reacción de acetilación inmediatamente posterior) revertiría nuevamente al fosfato diéster desprotegido, tal como se muestra en la Figura 1.9. La única consecuencia práctica de esta reacción secundaria sería que se consumiría como máximo un equivalente de fosforamidito, lo cual no resulta grave teniendo en cuenta los grandes excesos (10-20 eq.) que son habitualmente utilizados en la reacción de acoplamiento.

Para aplicar este esquema, se decidió utilizar únicamente cianoetilfosforamiditos, tanto para la síntesis del nucleotidil-linker, como para la elongación de la cadena. Es importante destacar que los controles por <sup>31</sup>P-RMN en fase gel de la nucleotidil-resina realizados para comprobar la total eliminación del grupo CNE del extremo 3'-terminal antes de iniciar la elongación de la cadena, mostraron la necesidad de aumentar el tiempo del tratamiento con TEA/piridina 1:1. En el esquema 1, este tiempo se había reducido a 1 h, para minimizar la eliminación de los grupos Me. Sin embargo, resultan realmente necesarios tres tratamientos de 1 h para asegurar la eliminación completa de todos los grupos CNE.

Con el objeto de comprobar la aplicabilidad de los esquemas de protección 2 y 4 en la síntesis de oligonucleótidos cíclicos, y confirmar los resultados obtenidos con los compuestos modelo, se realizaron varias síntesis que se describen a continuación.

---

<sup>11</sup> De Napoli, L.; Galeone, A.; Mayol, L.; Messere, A.; Montesarchio, D.; Piccialli, G. *Bioorg. & Med. Chem.* **1995**, *10*, 1325.



**Figura 1.9.-** Posible reacción secundaria sobre el fosfato 3'-terminal desprotegido durante la elongación de la cadena.

### 1.1.1.3.- Síntesis de Oligonucleótidos Cíclicos Realizadas con los Esquemas de Protección 1, 2 y 3.

Para probar los esquemas de protección 2 y 4, se llevó a cabo la síntesis de un octámero y un dodecámero cíclicos de secuencia d<pCATT<math>CATT</math>> y d<pTCAT<math>ATCTATAT</math>>, respectivamente. Como referencia, el mismo octámero cíclico se sintetizó también siguiendo el esquema de protección 1.

Aunque el procedimiento general de síntesis es común a todos los esquemas de protección, ver Figura 1.1, y se encuentra ampliamente descrito en Materiales y Métodos, se ha

creído conveniente citar aquí algunos detalles importantes relacionados con cada una de las etapas, así como dar los datos experimentales más relevantes asociados a cada una de las síntesis. En la Figura 1.10 se muestra el proceso de obtención de la nucleotidil-resina y que comprende las etapas 1-3 descritas a continuación.

### 1. Obtención de Fmoc-resina.

El soporte polimérico se hace reaccionar, en presencia de DCC y HOBt, con el ácido N-Fmoc-6-aminohexanoico, cuya función es la de actuar de espaciador entre el soporte polimérico y la cadena oligonucleotídica. Esta reacción se utiliza también para disminuir la funcionalización inicial de la resina en grupos  $\text{NH}_2$  hasta un valor más adecuado para llevar a cabo la síntesis (120-140  $\mu\text{mol/g}$ ). El grupo Fmoc se elimina con piperidina y la cuantificación espectrofotométrica de la N-(9-fluorenilmetil)piperidina permite determinar el grado de sustitución de la resina.

Como soporte polimérico se ha utilizado generalmente TentaGel (polietilenglicol-poliestireno) con una funcionalización en grupos  $\text{NH}_2$  de 240  $\mu\text{mol/g}$ , aunque, ocasionalmente, a lo largo del presente trabajo también se ha utilizado un soporte de poliestireno, *p*-metilbenzohidrilamina (MBHA).

En algún caso se ha optado por no funcionalizar la resina con este espaciador y anclar directamente el nucleotidil-linker. Si bien estudios anteriores muestran que la presencia del espaciador hace que la reacción de desanclaje tenga lugar de forma más rápida<sup>1</sup>, no se observan diferencias en cuanto al rendimiento final de crudo obtenido, por lo que no va a afectar a nuestro análisis de resultados.

### 2. Síntesis del nucleotidil-linker.

El producto se obtiene por reacción del derivado fosforamidito del nucleósido, con el grupo protector adecuado (CNE para los esquemas 1 y 4 y Me para el esquema 2), con el derivado del espaciador bifuncional 3-cloro-4-hidroxifenilacetato de 2, 4, 5-triclorofenilo en presencia de tetrazol.

A continuación se oxida con  $^t\text{BuOOH}$  el fosfito triéster a fosfato triéster.

El principal problema que presenta la obtención de este producto es su purificación. Debido a la elevada reactividad del ester de triclorofenilo, no resulta posible purificar el producto por cromatografía en columna de gel de sílice, al quedar el producto unido irreversiblemente a la sílice. Sin embargo, la principal impureza que presenta el crudo de reacción corresponde al hidrogenofosfonato formado por la hidrólisis del exceso de fosforamidito utilizado en la reacción. La presencia de hidrogenofosfonato, que no reacciona con los grupos amino del soporte polimérico, no interfiere en el posterior anclaje del nucleotidil-linker a la resina. Por este motivo el nucleotidil-linker se hace reaccionar, sin purificación previa, con el soporte polimérico.

### 3. Síntesis de la nucleotidil-resina.

El soporte polimérico, previamente funcionalizado o no con el ácido 6-aminohexanoico, se hace reaccionar con el nucleotidil-linker en presencia de DCC y HOBt. La unión a resina se produce por el ataque del grupo amino del soporte sobre el carbonilo del ester activo de triclorofenilo.

En este punto se determina la funcionalización de la resina mediante la detritilación en medio ácido (TCA) de una alícuota de nucleotidil-resina. Si la funcionalización es la adecuada (120-140  $\mu\text{mol/g}$ ) se procede a la acetilación ( $\text{Ac}_2\text{O}/\text{DIEA}$ ) de las aminas libres que quedan sobre el soporte polimérico.

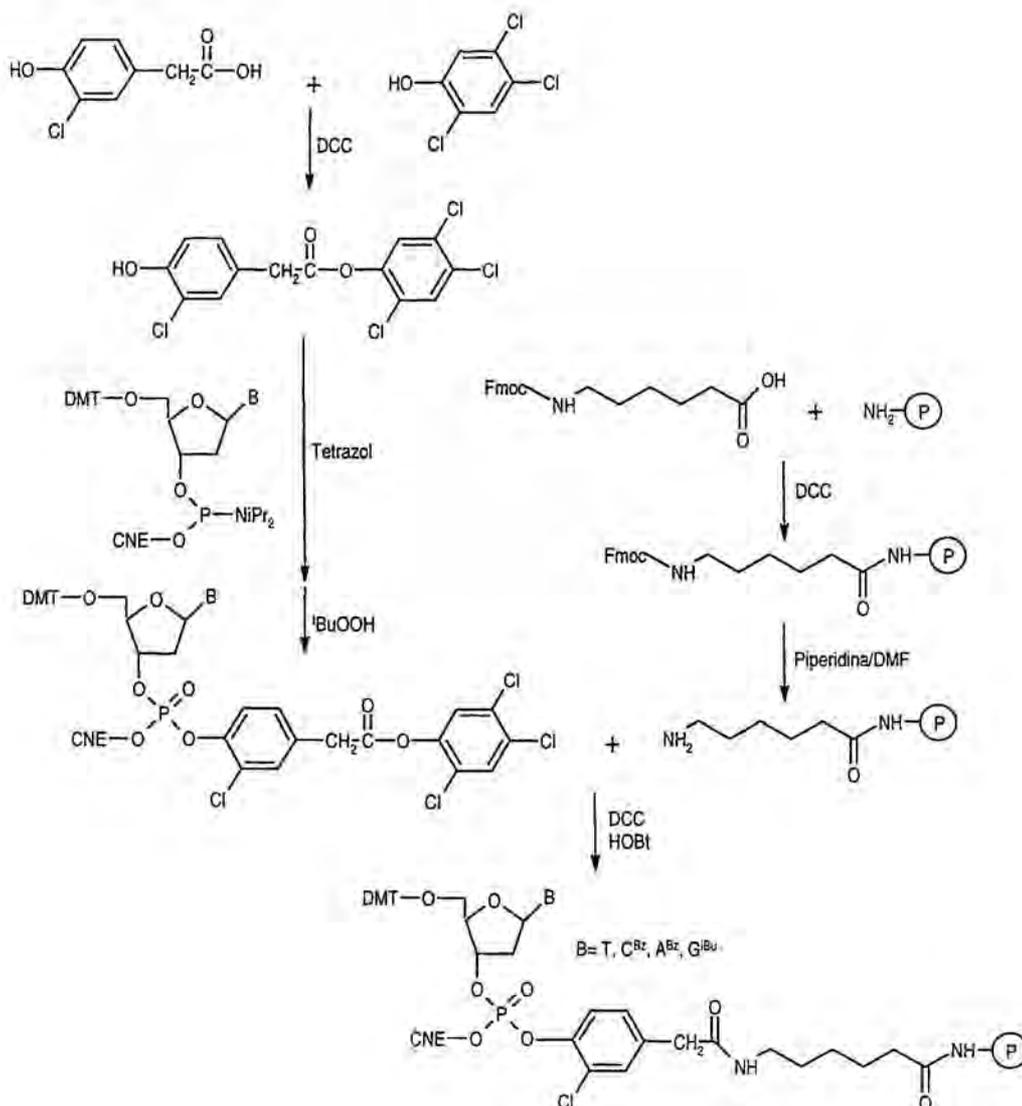


Figura 1.10.- Esquema de obtención de la nucleotidil-resina.

La pureza de la nucleotidil-resina se comprueba antes de cada síntesis de oligonucleótidos por <sup>31</sup>P-RMN en fase gel. En el caso del esquema 4, se realiza además un espectro de <sup>31</sup>P-RMN después de llevar a cabo la desprotección total del fosfato 3'-terminal, antes de iniciar la elongación de la cadena, para confirmar que se ha eliminado la totalidad de los grupos CNE.

En la Tabla 1.8 se muestran los valores de funcionalización obtenidos para las distintas nucleotidil-resinas utilizadas a lo largo de este trabajo, así como el desplazamiento químico de <sup>31</sup>P-RMN. Es importante tener en cuenta que, en el caso de nucleotidil-resinas con el fosfato 3'

protegido con R<sub>1</sub>, no siempre resulta posible observar las dos señales correspondientes a los dos diastereómeros.

**Tabla 1.8.-** Caracterización por <sup>31</sup>P-RMN de los distintos soportes poliméricos (T-resina) utilizados en síntesis. En todos los casos se ha utilizado como disolvente CDCl<sub>3</sub> y referencia externa de trimetilfosfito (TMP).

Soporte polimérico	R1	f μmol/g	δ <sup>31</sup> P-RMN ppm
TentaGel sin espaciador	CNE	83	-7.33
TentaGel sin espaciador	Me	94	-9.75, -9.91
p-MBHA con espaciador	CNE	96	-10.48
TentaGel con espaciador	-	70	-5.23 <sup>§</sup>
TentaGel con espaciador	CNE	47	-7.37 <sup>§</sup>
TentaGel con espaciador	CNE	43	-7.38 <sup>§</sup>
TentaGel con espaciador	CNE	118	-6.34
			-5.12 <sup>§</sup>

<sup>§</sup> corresponde al fosfato desprotegido.

#### 4. Elongación de la cadena.

La elongación de la cadena se realiza de forma automática, siguiendo el método del fosfito triéster, y utilizando los derivados fosforamidito de los nucleósidos adecuados para cada esquema de protección: cianoetilfosforamiditos para los esquemas 2 y 4 y metilfosforamiditos para el esquema 1.

El rendimiento de la síntesis se determina por cuantificación espectrofotométrica del catión DMT<sup>+</sup> liberado en cada etapa de desprotección del hidroxilo 5'-terminal. En la Tabla 1.9, se muestran los rendimientos obtenidos para la síntesis del oligonucleótido lineal en cada caso, indicando el soporte polimérico, escala de trabajo y esquema de protección utilizados.

**Tabla 1.9.-** Resumen de los rendimientos de obtención de los oligonucleótidos lineales anclados a resina para cada una de las síntesis.

Secuencia	Esquema protección	Escala	Soporte polimérico	Rdt. síntesis lineal	Rdt. medio por etapa
HO-CATTCATTp-(P)	1 R1 = CNE R2 = Me	10 μM	TentaGel <sup>a</sup> 83 μmol/g	87	98
HO-CATTCATTp-(P)	2 R1 = Me R2 = CNE	10 μM	TentaGel <sup>a</sup> 94 μmol/g	80	97
HO-CATTCATTp-(P)	4 R1 = CNE R2 = CNE	10 μM	TentaGel <sup>b</sup> 70 μmol/g	80	97
HO-TCATATCTATATp-(P)	2 R1 = Me R2 = CNE	5 μM	TentaGel <sup>a</sup> 94 μmol/g	88	99
HO-TCATATCTATATp-(P)	4 R1 = CNE R2 = CNE	10 μM	TentaGel <sup>b</sup> 96 μmol/g	88	99

<sup>a</sup> Sin espaciador.

<sup>b</sup> Con espaciador.

A continuación se detallan las condiciones de eliminación del grupo protector del fosfato del extremo 3'-terminal, antes de llevar a cabo la ciclación, para cada uno de los esquemas (Tabla 1.10). En la misma tabla se incluyen las condiciones de eliminación de los grupos fosfato de la cadena, aunque esta reacción se lleva a cabo inmediatamente después de la ciclación.

**Tabla 1.10.-** Condiciones de eliminación de los grupos protectores para cada uno de los esquemas utilizados.

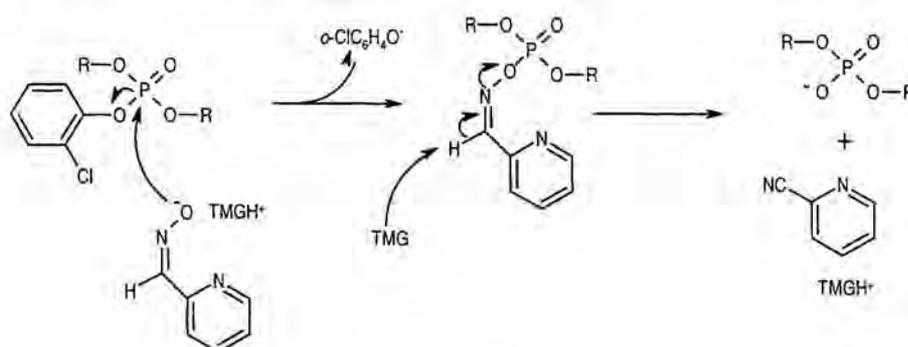
Esquema de protección	R1	R2	Condiciones eliminación CNE	Condiciones eliminación Me
1	CNE	Me	TEA/piridina 1:1 3 x 1 h	Tiofeno/TEA/dioxano 1 : 2 : 2 2 x 10 min
2	Me	CNE	TEA/piridina 1:1 3 x 1 h	Tiofeno/NMM/DMF 1 : 2 : 2 4 x 1 h
4	CNE <sup>s</sup>	CNE	TEA/piridina 1:1 3 x 1 h	-

<sup>s</sup> Se elimina antes de la elongación de la cadena

### 5. Ciclación, desprotección y desanclaje.

De estas tres etapas, la más delicada corresponde a la ciclación, con MSNT, del precursor lineal. Para llevar a cabo esta reacción es necesario que la oligonucleotidil-resina esté bien seca y los reactivos y disolventes que se vayan a utilizar sean anhidros. De no ser así, la pequeña proporción de agua presente podría competir con el grupo hidroxilo del extremo 5'-terminal en su reacción con el fosfato activado.

Las etapas de desprotección de fosfatos (según las condiciones indicadas en la Tabla 1.10), desanclaje del producto cíclico por reacción con el oximato de tetrametilguanidinio (TMG), Figura 1.11, y, por último, eliminación de los protectores de las bases (Bz para adenina y <sup>1</sup>Bu para guanina) mediante un tratamiento con NH<sub>3</sub> conc., se realizan de la forma descrita en la sección 2.2 de Materiales y Métodos. El producto crudo se desala por cromatografía de filtración molecular sobre una columna de Sephadex G-10 y se determina el rendimiento de obtención de oligonucleótido sin purificar (Rdt. crudo). El posterior análisis de la calidad del crudo se lleva a cabo por HPLC en fase reversa.



**Figura 1.11.-**Reacción de desanclaje con oximato de TMG.

### 6. Purificación.

La purificación de oligonucleótidos cíclicos se ha realizado generalmente en nuestro grupo mediante cromatografía de intercambio aniónico sobre resina DEAE-Sephadex, o bien mediante cromatografía líquida de media presión (MPLC). Si bien ambos métodos permiten obtener productos con un alto grado de pureza (>98%), los rendimientos de recuperación, aún y en el caso de crudos de elevada calidad, acostumbran a ser muy bajos. Por este motivo, en este trabajo se ha estudiado la viabilidad de utilizar un método alternativo como es la purificación por HPLC en fase reversa semipreparativo a través de una columna de poli(estireno-divinilbenceno).

### 7. Caracterización.

La determinación de la correcta composición en nucleósidos de los oligonucleótidos cíclicos sintetizados, así como su resistencia a exonucleasas, se ha realizado por ensayos de digestión enzimática con los siguientes enzimas<sup>12-14</sup>:

i) una mezcla de *fosfodiesterasa de veneno de serpiente* (SVPD, EC 3.1.4.1), que escinde el oligonucleótido por el extremo 3'-terminal generando nucleósidos-5'-fosfato, y *fosfatasa alcalina* (AP, EC 3.1.3.1), que hidroliza los nucleósidos 3' o 5'-fosfato obteniéndose los nucleósidos libres,

ii) una disolución de *fosfodiesterasa de bazo de buey* (SpPD, EC 3.1.16.1), que escinde el oligonucleótido por el extremo 5'-terminal generando nucleósidos 3'-fosfato.

Con el ensayo i), y debido a la presencia de fosfatasa alcalina, el producto se degrada totalmente, aún y en el caso de no presentar el extremo 3'-terminal libre, por lo que es posible cuantificar la proporción exacta de los cuatro nucleósidos y, de esta forma, verificar que se ha sintetizado la secuencia correcta. El análisis del crudo de digestión por HPLC permite además, por comparación con patrones, comprobar que no ha tenido lugar ninguna modificación sobre las bases.

La circularidad del producto final se confirma con el ensayo ii). Si el producto es ciertamente cíclico, la inexistencia de un extremo 5'-terminal libre impide que sea degradado por el enzima y el análisis por HPLC del crudo de digestión muestra el producto inalterado.

La caracterización de los productos cíclicos se realiza también por espectrometría de masas (ES<sup>-</sup> y/o MALDI-TOF).

Los resultados obtenidos para cada una de las síntesis, incluyendo rendimientos de crudo, pureza HPLC y rendimientos globales, se resumen en la Tabla 1.11.

<sup>12</sup> Adams, R.L.P.; Knowler, J.T.; Leader, D.P.P. «*The Biochemistry of the Nucleic Acids*» 1992, 11ª Ed., Chapman & Hall, London.

<sup>13</sup> Dixon, M.; Webb, E.C. «*Enzymes*» 1979, 3ª Ed., Longman, London.

<sup>14</sup> «*Biochemica Information*» 1987, Boehringer-Mannheim.

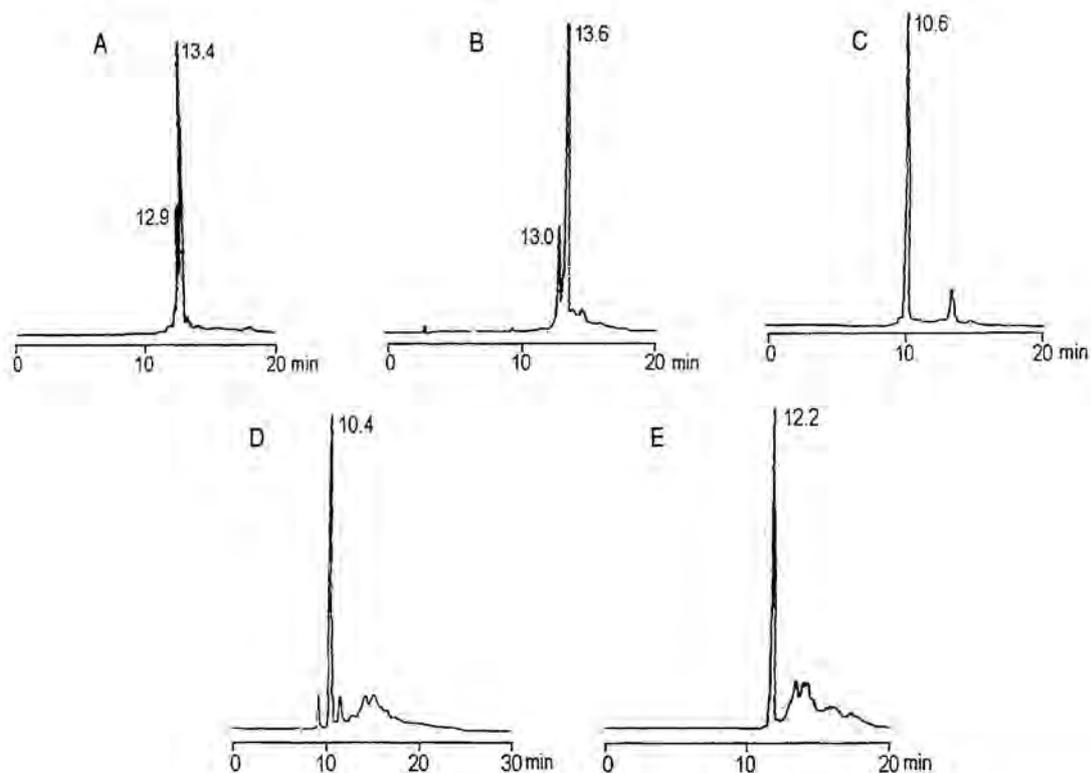
Tabla 1.11.- Resumen de resultados de las síntesis realizadas con los distintos esquemas de protección.

Secuencia	Esquema protección		Rdt. crudo %	Pureza HPLC %	Método purificación	Rdt. recuperación %	Rdt. global %
d<pCATTGATT>	1	R1 = CNE R2 = Me	27 (29) <sup>§</sup>	70	DEAE Sephadex MPLC	27	6
d<pCATTGATT>	2	R1 = Me R2 = CNE	12 (15) <sup>§</sup>	57	DEAE Sephadex	38	2.6
d<pCATTGATT>	4	R1 = CNE R2 = CNE	36 (40) <sup>§</sup>	80	HPLC	75	19.3
d<pTCATATCTATAT>	2	R1 = Me R2 = CNE	10 (11) <sup>§</sup>	51	MPLC	26	1.3
d<pTCATATCTATAT>	4	R1 = CNE R2 = CNE	22 (25) <sup>§</sup>	36	DEAE Sephadex	58	4.6

<sup>§</sup> Entre paréntesis se muestra el rendimiento correspondiente a ciclación, desprotección y desanclaje, excluyendo el rendimiento de la síntesis lineal.

Los rendimientos obtenidos en cada una de las síntesis muestran claramente que el esquema de protección 4 es la mejor de las opciones, pues se obtienen rendimientos muy superiores a los obtenidos en los otros dos esquemas. Estos resultados confirman la viabilidad de llevar a cabo la síntesis en presencia de fosfatos desprotegidos. Además, se pone de manifiesto que la no ortogonalidad de los grupos protectores R<sub>1</sub> y R<sub>2</sub>, que se da tanto en el esquema de protección 1 como en el 2, afecta de forma importante al rendimiento de obtención de producto cíclico. En este sentido, para el esquema 2 se confirman los datos de la cinética por <sup>31</sup>P-RMN realizada sobre los compuestos modelo y que indicaban un elevado porcentaje de eliminación de los grupos CNE en las condiciones de eliminación del grupo Me elegidas.

A nivel de calidad de los crudos de síntesis, cuyo análisis por HPLC se muestra en la Figura 1.12, un rasgo común es el elevado grado de pureza y la presencia de un pico muy mayoritario. Las pruebas de caracterización han confirmado que, en todos los casos, el pico mayoritario corresponde al producto cíclico deseado, Tabla M.9 de Materiales y Métodos. Este hecho, así como el que no se observen en general subproductos a tiempos de retención menores que el producto deseado, es un indicativo de que los posibles subproductos procedentes de ciclaciones entre el extremo 5'-terminal y fosfatos internucleotídicos no se desanclan de la resina, confirmando, por tanto, la elevada selectividad de la reacción de desanclaje.



**Figura 1.12.-** Cromatogramas HPLC de los crudos de síntesis de: A)  $d\langle pCATT\langle CATT \rangle$  esquema 1, gradiente 10-40% B en 20 min, B)  $d\langle pCATT\langle CATT \rangle$  esquema 2, gradiente: 10-40% B en 20 min, C)  $d\langle pCATT\langle CATT \rangle$  esquema 4, gradiente: 10-40% B en 20 min, D)  $d\langle pTCATAT\langle CTATAT \rangle$  esquema 2, gradiente 15-35% B en 20 min, E)  $d\langle pTCATAT\langle CTATAT \rangle$  esquema 4, gradiente: 10-40% B en 20 min.

El subproducto mayoritario ( $t_R = 12.9$  min) de la síntesis de  $d\langle pCATT\langle CATT \rangle$  siguiendo el esquema 1, y cuya proporción se ha estimado en un 25%, se ha podido identificar por espectrometría de masas como el heptámero cíclico  $d\langle pCATT\langle CAT \rangle$  (ES:  $M_{calc} = 2116.3$ ,  $M_{det} = 2116.4$ ). Este producto se podría haber formado como consecuencia de la falta de homogeneidad inicial de la resina utilizada para la síntesis<sup>1</sup>. Puesto que el nucleotidil-linker no se purifica antes de su anclaje al soporte polimérico, si quedan restos de 3-cloro-4-hidroxifenilacetato de 2, 4, 5-triclorofenil, este producto se puede anclar sobre resina y permitir que, sobre su grupo fenólico, crezca también la cadena oligonucleotídica. El producto obtenido tendría la misma secuencia que el oligonucleótido deseado pero sin el primer nucleótido.

Este mismo producto secundario se ha observado también en la síntesis de  $d\langle pCATT\langle CATT \rangle$  con el esquema de protección 2. En este caso no se ha aislado y caracterizado el producto que, sin embargo, coeluye con el producto aislado de la síntesis anterior.

En el caso del dodecámero, la pureza anormalmente baja de los crudos se podría explicar como la consecuencia de un elevado porcentaje de reacciones intercadena, facilitadas por el grado de funcionalización de la resina utilizada. En el caso de la síntesis realizada según el esquema 4, se intentó aislar alguno de los subproductos mayoritarios y caracterizarlos por espectrometría de masas, pero no resultó posible. Sin embargo, el análisis por electroforesis en gel de poliacrilamida, muestra para el subproducto mayoritario una migración que podría

corresponder a un 24-mer, producto de reacciones intercadena. Esta hipótesis se ve reforzada por el hecho de que en la digestión enzimática de este subproducto se obtiene una proporción de bases idéntica a la obtenida para el dodecámero, Tabla xx.

Una de las etapas que resulta crítica de cara al aislamiento final del producto cíclico es sin duda la etapa de purificación. Tal como se ha comentado anteriormente, se han empleado distintos métodos de purificación con el objeto de comparar rendimientos de recuperación y determinar qué método resulta más adecuado. Tal como se deduce de los valores que se muestran en la Tabla 1.11, la purificación por HPLC semipreparativo en fase reversa proporcionó los mejores resultados, significando un notable avance respecto a los dos métodos utilizados habitualmente y permitiendo un aumento del rendimiento de obtención de producto final entre 3-7 veces respecto a la purificación sobre DEAE-Sephadex.

### 1.1.2.- ESTUDIO DE DIFERENTES ETAPAS DEL MÉTODO DE SÍNTESIS.

Las mejoras introducidas en el método de síntesis en cuanto al esquema de protección de los grupos fosfato nos han permitido solventar varios problemas que afectaban de forma importante al rendimiento de obtención del producto cíclico. No obstante, los rendimientos de crudo después del desanclaje siguen siendo moderados, del orden del 40%. Esto significa que, o bien queda una cantidad importante de cadenas oligonucleotídicas ancladas al soporte polimérico, o bien se está produciendo una pérdida de producto, lineal o cíclico, en alguna de las etapas de la síntesis posteriores a la elongación de la cadena.

Se han llevado a cabo distintos ensayos que tienen como objetivo evaluar la estabilidad de la unión de las cadenas oligonucleotídicas a la resina frente a los diferentes tratamientos que se realizan a lo largo de la síntesis, así como algunos intentos de desanclar el producto oligonucleotídico que queda sobre la resina después del tratamiento con oximato.

A continuación se detallan los distintos ensayos realizados.

#### 1.1.2.1.- Ensayo 1: Determinación de grupos 5'-OH libres después de la etapa de ciclación.

Una de las etapas clave para la obtención de oligonucleótidos cíclicos es la etapa de ciclación. Este ensayo tiene como objetivo determinar la eficacia de esta reacción y estimar que porcentaje de grupos hidroxilo 5'-terminales quedan libres sobre resina después del tratamiento con MSNT.

Para ello, tras llevar a cabo la reacción de ciclación en las condiciones habituales, 0.1 M de MSNT en piridina anh. 4 h + 4 h + 12 h, se separó una alícuota de oligonucleotidil-resina y se acopló DMT-T 3'-O-fosforamidito. Se eliminó el grupo DMT y se determinó la funcionalización en grupos hidroxilo de la resina.

El ensayo se realizó sobre la oligonucleotidil-resina procedente de la síntesis del dodecámero d<pTCATCACTATAT> con el esquema de protección 2, Tabla 1.9. La funcionalización de la oligonucleotidil-resina antes de la ciclación se estimó en 85 μmol/g, teniendo en cuenta que la funcionalización inicial era de 94 μmol/g y que la elongación de la cadena había tenido lugar con un 99% de Rdt medio por etapa. Después de la ciclación, la cuantificación de grupos DMT dió un valor de la funcionalización de 6 μmol/g. Este resultado muestra que, después de la ciclación, queda sobre resina un 7% de grupos hidroxilo libres. El 93% restante podría corresponder, en teoría, a producto cíclico. No obstante, dado los bajos rendimientos de desanclaje obtenidos, es probable que una proporción elevada de este 93%

corresponda bien a productos de reacciones intercadena o bien a cadenas resultantes de un bloqueo del grupo hidroxilo. Esta última posibilidad se discute a continuación.

### 1.1.2.2.- Ensayo 2: Bloqueo del grupo OH 5'-terminal?

Una posible reacción secundaria asociada a la utilización de MSNT como reactivo de acoplamiento, y que se encuentra ampliamente descrita en la literatura<sup>15,16</sup>, es la sulfonación del grupo hidroxilo que impediría que tuviera lugar la reacción de ciclación.

Para evaluar la extensión de esta reacción secundaria se llevó a cabo un ensayo consistente en tratar una nucleotidil-resina, donde únicamente el extremo 5'-terminal se encontraba desprotegido, en las condiciones habituales de la reacción de ciclación. De esta forma, o bien el grupo hidroxilo continuaba libre después de los tratamientos con MSNT o bien se había producido sulfonación.

Se trató una timidinil-resina con una funcionalización inicial de 68  $\mu\text{mol/g}$  con TCA para eliminar el grupo DMT, protector del hidroxilo 5'-terminal. Posteriormente se realizaron los tres tratamientos con MSNT 0.1 M en piridina (4 h + 4 h + 12 h) y seguidamente se acopló un DMT-T 3'-O-fosforamidito. La cuantificación de los grupos DMT sobre la resina dió un valor de funcionalización de 70  $\mu\text{mol/g}$ . Este resultado indica que, durante la reacción de ciclación, no se produce bloqueo del grupo hidroxilo y refuerza la idea de que la mayor competencia a la formación de producto cíclico es la existencia de reacciones intercadena.

### 1.1.2.3.- Ensayo 3: Análisis de los filtrados procedentes de la reacción de ciclación.

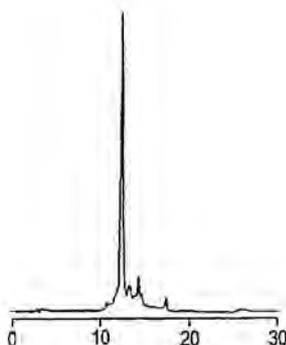
Otra posibilidad, para explicar el bajo rendimiento de obtención de producto cíclico, podría ser la pérdida de producto durante la etapa de ciclación. Por este motivo se decidió coleccionar los filtrados procedentes de los tres tratamientos con MSNT y determinar la posible presencia de producto oligonucleotídico en disolución.

En este caso se utilizó la oligonucleotidil-resina procedente de la síntesis del octámero d<pCATT<CATT>. Los filtrados procedentes de la reacción de ciclación se juntaron y, después de un tratamiento con  $\text{NH}_3$  para desproteger las bases, se desalaron a través de una columna de Sephadex G-10. La cuantificación por UV reveló la presencia de 4.63 OD de producto oligonucleotídico en la disolución, lo que correspondería a una pérdida de un 1%. También en el espectro de  $^{31}\text{P}$ -RMN se observó una señal ancha en la zona de fosfatos, aunque de muy poca intensidad.

Por coinyección con el patrón de d<pCATT<CATT>, el análisis por HPLC demostró que el producto mayoritario (68%) correspondía al producto cíclico deseado, Figura 1.13.

<sup>15</sup> Gait, M.J.; Popov, S.G. *Tetrahedron Lett.* **1980**, 21, 2841-2842.

<sup>16</sup> Katagari, N.; Itakura, K.; Narang, S.A. *J. Am. Chem. Soc.* **1975**, 10, 7332-7337.



**Figura 1.13.-** Cromatograma HPLC de los filtrados procedentes del tratamiento de la oligonucleotidil-resina con MSNT.

Si bien parece que durante la etapa de ciclación se puede desanclar el producto cíclico ya formado, la extensión de esta reacción es tan baja que en ningún caso podría ser la causa principal de los rendimientos finales de crudo obtenidos.

#### **1.1.2.4.- Ensayo 4: Análisis de los filtrados procedentes de la etapa de desprotección de los fosfatos internucleosídicos.**

El siguiente paso, una vez formado el ciclo, es la desprotección de los fosfatos internucleosídicos. Este ensayo se realizó con la finalidad de ver si en dicha etapa se producía una pérdida significativa de producto cíclico.

Puesto que de los distintos esquemas de protección, el más adecuado resultó ser el esquema 4, con el grupo CNE como protector  $R_2$ , se eligió una síntesis realizada según este esquema para llevar a cabo este ensayo, la correspondiente a la secuencia  $d\langle pCAT^{Br}UCAT^{Br}U \rangle$ , descrita más adelante.

Para ello se colectaron y juntaron los filtrados procedentes de cada uno de los tres tratamientos con TEA/piridina 1:1. Los filtrados se evaporaron a sequedad, realizando coevaporaciones con ACN hasta la eliminación total tanto de la TEA como de la piridina y, posteriormente, se disolvió el residuo en  $H_2O$ . El análisis por UV de la disolución resultante no mostraba espectro de oligonucleótido.

En resumen, de los distintos ensayos realizados se puede deducir que no hay pérdida significativa de oligonucleótido a lo largo de las distintas etapas que se llevan a cabo para la obtención del producto cíclico. De acuerdo con los resultados procedentes del ensayo 1, que muestran que casi la totalidad de los grupos hidroxilo 5'-terminales reaccionan durante la etapa de ciclación, parece que la hipótesis más plausible sea que un elevado porcentaje de dichos grupos han podido reaccionar dando lugar a productos intercadena.

A continuación se describen distintos intentos realizados con objeto de desanclar los productos de naturaleza oligonucleotídica que quedan sobre el soporte una vez liberado el producto cíclico.

Es importante destacar que, la pretensión de encontrar unas condiciones capaces de romper el enlace entre el extremo 3'-terminal desprotegido (fosfato diéster) y el espaciador bifuncional, no tiene como único objetivo el de liberar las cadenas oligonucleotídicas que quedan sobre el soporte polimérico. Caso de encontrar unas condiciones adecuadas, ello nos

proporcionaría también un método para obtener oligonucleótidos lineales 3'-fosfato con el mismo método que el utilizado para obtener oligonucleótidos cíclicos.

#### 1.1.2.5.- Tratamiento de la resina de síntesis después de la reacción de desanclaje.

Se han realizado distintos intentos con objeto de liberar los productos de naturaleza oligonucleotídica que aún quedan unidos a la resina después del desanclaje del producto cíclico. Los ensayos se realizaron sobre resinas de síntesis sometidas previamente a cuatro tratamientos con oximato de TMG, para tener la seguridad de que no quedaba producto cíclico anclado al soporte polimérico.

- *Tratamiento con NH<sub>3</sub> conc. 48 h.*

Un primer intento consistió en realizar un tratamiento largo (48 h) con NH<sub>3</sub> conc. El ensayo se realizó sobre la resina procedente de la síntesis de d<pTCATATCTATAT> según el esquema de protección de fosfatos 2. Los filtrados procedentes del tratamiento con NH<sub>3</sub> se pasaron por una columna NAP, para su desalado, y se analizaron por HPLC en fase reversa. Se obtuvo un cromatograma muy sucio, con un gran número de picos, Figura 1.14, y en el que el pico mayoritario ( $t_R = 17.6$  min) parece corresponder a una mezcla de productos con un tiempo de retención muy superior al del producto cíclico cuyo  $t_R \sim 11-12$  min. La cuantificación por UV confirmó la presencia de 8 OD de oligonucleótido en la disolución, que corresponde únicamente a un 2.5% de producto desanclado.

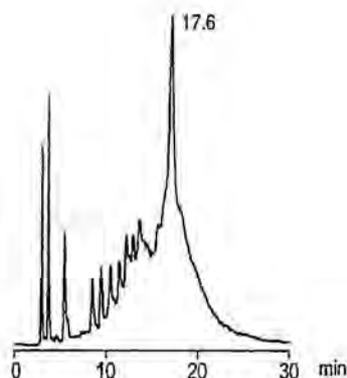


Figura 1.14.- Cromatograma HPLC de los filtrados procedentes del tratamiento con NH<sub>3</sub> conc.

Se puede concluir, por tanto, que el NH<sub>3</sub> no es un reactivo eficaz para desanclar los productos que quedan sobre el soporte.

- *Tratamiento con TBAF.*

Estudios descritos en la bibliografía demuestran que el fluoruro de tetrabutilamonio (TBAF) es un reactivo eficaz para la eliminación de algunos de los protectores utilizados en la síntesis de oligonucleótidos, en particular fenil-, tricloroetil- y cianoetilfosfotriesteres<sup>17,18</sup>.

<sup>17</sup> Ogilvie, K.; Beaucage, S.L. *Tetrahedron Lett.* **1976**, 16, 1255-1256.

<sup>18</sup> Ogilvie, K.; Beaucage, S.L. *Nucleic Acids Res.* **1979**, 7, 805-823.

En nuestro caso, el grupo fosfato del extremo 3'-terminal de la cadena oligonucleotídica tiene como protector permanente el grupo *o*-clorofenilo. El ataque del ión fluoruro sobre dicho fosfato podría estar favorecido por la presencia de un buen grupo saliente y, de esta forma, separar las cadenas oligonucleotídicas del soporte polimérico.

Para llevar a cabo el tratamiento con TBAF se eligió una resina procedente de la síntesis del dímero cíclico d<pGG><sup>19</sup>, obtenido con un rendimiento muy bajo por lo que presumiblemente quedaba una elevada proporción de oligonucleótido anclado al soporte. Se sometió la resina a dos tratamientos con TBAF en THF anh., el primero TBAF 0.1 M (24 h) y el segundo TBAF 0.2 M (60 h), obteniéndose en ambos casos, en el análisis por HPLC de los filtrados, un pico muy ancho y muy poco intenso. Los filtrados correspondientes al segundo tratamiento se desalaron a través de una columna de Sephadex G-10 y se cuantificaron por UV, determinándose un total de 6.87 OD de producto oligonucleotídico, lo que correspondería a un desanclaje del 4%.

Se realizó el mismo ensayo sobre otra oligoresina, pero esta vez en el estadio de oligonucleótido lineal anclado a resina, antes de llevar a cabo las etapas de ciclación y desanclaje. Para ello se utilizó una oligoresina correspondiente a la secuencia 5'-ACGCTATGCGAT-3', realizándose previamente al tratamiento con TBAF la desprotección de los fosfatos de la cadena, protegidos con el grupo CNE. Se realizaron dos tratamientos (2x12 h) con TBAF 0.1M en THF anh., utilizando un exceso de 70 equivalentes para cada uno de los tratamientos. El filtrado se evaporó a sequedad, se trató con NH<sub>3</sub> conc. para eliminar los protectores de las bases y, posteriormente, se desaló a través de una columna de Sephadex G-10. El análisis del crudo por HPLC muestra una serie de picos muy anchos que probablemente corresponden a una mezcla heterogénea de productos. La cuantificación por UV a  $\lambda = 260$  nm reveló un desanclaje de sólo un 4.8%.

Estos resultados corroboran los resultados obtenidos con los otros ensayos y ponen de manifiesto que los productos que quedan sobre resina, una vez se ha liberado el producto cíclico, no son fácilmente recuperables. Es probable que los productos poliméricos, que se puedan formar por reacciones intercadena, queden atrapados dentro del entramado del soporte polimérico. Hay que tener en cuenta que, para que tenga lugar la separación de estos productos poliméricos de la resina deben romperse durante el tratamiento correspondiente varios enlaces de unión al soporte. Por otra parte, y puesto que también se ha realizado un ensayo sobre un oligo lineal anclado a resina, se demuestra que el TBAF no es capaz de desproteger fosfato diesteres de *o*-clorofenilo.

### 1.1.3.- LA REACCIÓN DE CICLACIÓN: UNA ETAPA CLAVE EN LA OBTENCIÓN DE OLIGONUCLEÓTIDOS CÍCLICOS.

Una vez confirmado que no existía pérdida significativa de producto oligonucleotídico a lo largo de todo el proceso de síntesis, nos centramos en intentar optimizar la etapa de ciclación.

La reacción que tiene lugar entre el grupo hidroxilo 5'-terminal y el grupo fosfato 3'-terminal del precursor lineal, para dar lugar al producto cíclico deseado, es una reacción delicada y factores como la secuencia oligonucleotídica o la longitud de la cadena pueden tener una influencia decisiva. En particular, un aumento de la longitud de la secuencia puede propiciar

---

<sup>19</sup> Straub, M. Màster Experimental en Química Orgànica, Universidad de Barcelona, 1997.

un aumento de la formación de productos por reacción intercadena.

En nuestro caso, en el que la ciclación se realiza sobre el soporte polimérico y sin ayuda de un molde externo, la frecuencia de colisión de los extremos 3'- y 5'-terminales debe ser muy pequeña y los factores de secuencia y longitud pueden afectar drásticamente al éxito de la reacción. Por este motivo resulta imprescindible contar con un reactivo de acoplamiento que sea verdaderamente eficaz.

Como ya se ha comentado anteriormente, la reacción de ciclación se lleva a cabo mediante el método del fosfato triéster<sup>20</sup>, en el que se encuentra muy extendida la utilización de reactivos del tipo cloruro de arilsulfonilo como reactivos de acoplamiento<sup>21,22</sup>.

En particular, uno de los reactivos con el que se obtenían mejores resultados cuando la síntesis de oligonucleótidos se llevaba a cabo por este método, era el MSNT<sup>23</sup>. El mecanismo propuesto para este tipo de reactivos, en el caso de síntesis de oligonucleótidos lineales y en condiciones de exceso de monómero, implica la formación de un intermedio del tipo pirofosfato<sup>24,25</sup>, Figura 1.14. En primer lugar se formaría un anhídrido mixto fosfórico-sulfónico que, por ataque de otro grupo fosfato diéster daría lugar al correspondiente pirofosfato. Esta especie reaccionaría rápidamente con el grupo hidroxilo para dar el fosfato triéster correspondiente.

Si bien el pirofosfato es un intermedio comprobado en la síntesis de oligonucleótidos lineales, en nuestro caso es muy poco probable que se pueda formar dicha especie por varios motivos:

- i) los grupos fosfato 3'-terminales de los precursores lineales están unidos al soporte polimérico, por lo que es muy improbable que dichos grupos puedan reaccionar entre ellos,
- ii) no existe un exceso de grupos fosfato respecto a componente hidroxilado,
- iii) en las condiciones de reacción hay un exceso de MSNT, por lo que todos los grupos fosfato 3'-terminales deberían estar activados en forma de anhídrido fosfórico-sulfónico.

Por estos motivos, la reacción de ciclación es más factible que transcurra a través de alguno de los dos mecanismos que se proponen en la Figura 1.15. Es probable que, una vez formado el anhídrido fosfórico-sulfónico, éste reaccione con el 3-nitro-1,2,4-triazol presente en el medio. En este caso el fosfato se encontraría igualmente activado para el posterior ataque del hidroxilo del extremo 5'-terminal (Ruta A). O bien, es probable que el anhídrido mixto fosfórico-sulfónico experimente directamente el ataque nucleófilo del grupo hidroxilo 5'-terminal (Ruta B).

---

<sup>20</sup> Gait, M.J.; Matthes, H.W.D.; Singh, M.; Sproat, B.S.; Titmas, R.C. «*Chemical and Enzymatic Synthesis of Gene Fragments*» **1982**, Ed. Gassen, H.G. & Lang, A., Verlag Chemie, 1-42.

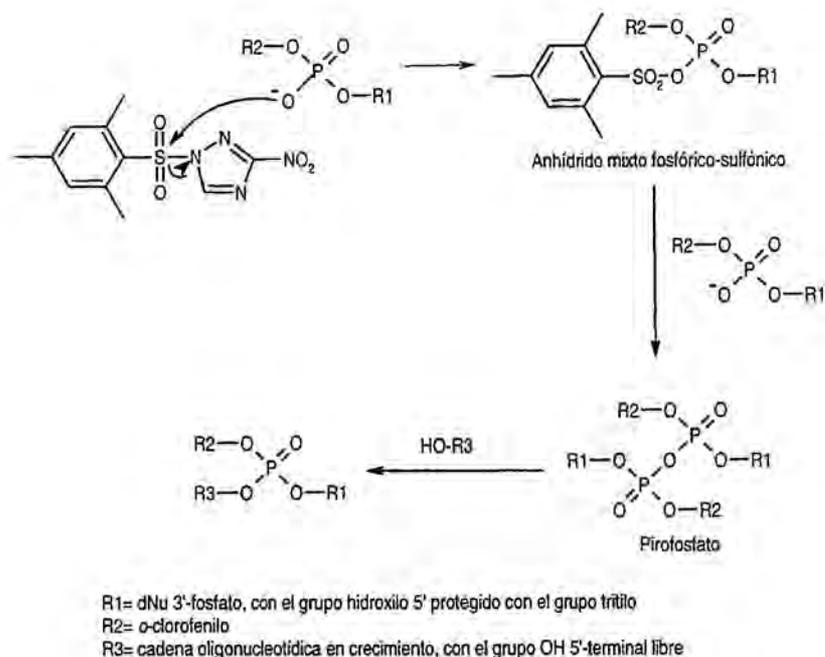
<sup>21</sup> Katagiri, N.; Itakura, K.; Narang, S.A. *J. Am. Chem. Soc.* **1975**, *97*, 7332-7337.

<sup>22</sup> Zarytova, V.F.; Knorre, D.G. *Nucleic Acids Res.* **1984**, *12*, 2091-2110.

<sup>23</sup> Devine, K.G.; Reese, C. *Tetrahedron Lett.* **1986**, *27*, 5529-5532.

<sup>24</sup> Ivanova, E.M.; Khalimskaya, L.M.; Romanenko, V.P.; Zarytova, V.F. *Tetrahedron Lett.* **1982**, *23*, 5447-5450.

<sup>25</sup> Chandrasegaran, S.; Murakami, A.; Kan, L. *J. Org. Chem.* **1984**, *49*, 4951-4957.



**Figura 1.15.-** Mecanismo de reacción con MSNT propuesto para el caso de síntesis de oligonucleótidos lineales.

Si bien con la utilización de MSNT hemos obtenido buenos resultados, se encuentra descrito en la literatura que, éste y otros derivados de arilsulfonilo, son reactivos inestables, requieren tiempos de reacción largos y pueden causar la modificación de las bases guanina y timina<sup>26</sup>, así como el bloqueo del grupo hidroxilo 5'-terminal. Este bloqueo del grupo hidroxilo se traduciría en una disminución de grupos potencialmente reactivos, por lo que el rendimiento de obtención del producto cíclico podría verse afectado de forma importante. Aunque algunos estudios fijan la tasa de sulfonación en un ~1%<sup>27</sup>, estos datos corresponden a acoplamientos durante la elongación de la cadena por el método del fosfato triéster, para los que se utilizan grandes excesos de monómero y tiempos de reacción relativamente cortos. En nuestro caso, dado que la reacción de ciclación requiere tratamientos más largos, la sulfonación podría representar un problema importante.

Para evaluar la extensión de la sulfonación en nuestro caso, se realizó el Ensayo 2, anteriormente descrito, apartado 1.1.2.2. Los resultados obtenidos parecían indicar la ausencia de producto sulfonado, ya que se obtuvieron valores similares de la funcionalización inicial y final de la resina. No obstante, con objeto de buscar alternativas al MSNT, que no implicaran reacciones secundarias y, sobre todo, para intentar encontrar un reactivo que proporcionara mayores rendimientos de obtención de producto cíclico, se han ensayado otros reactivos de acoplamiento.

<sup>26</sup> Reese, C.B., Ubasawa, A. *Tetrahedron Lett.* **1980**, 21, 2265-2268.

<sup>27</sup> Gait, M.J., Popov, S.G. *Tetrahedron Lett.* **1980**, 21, 2841-2842.

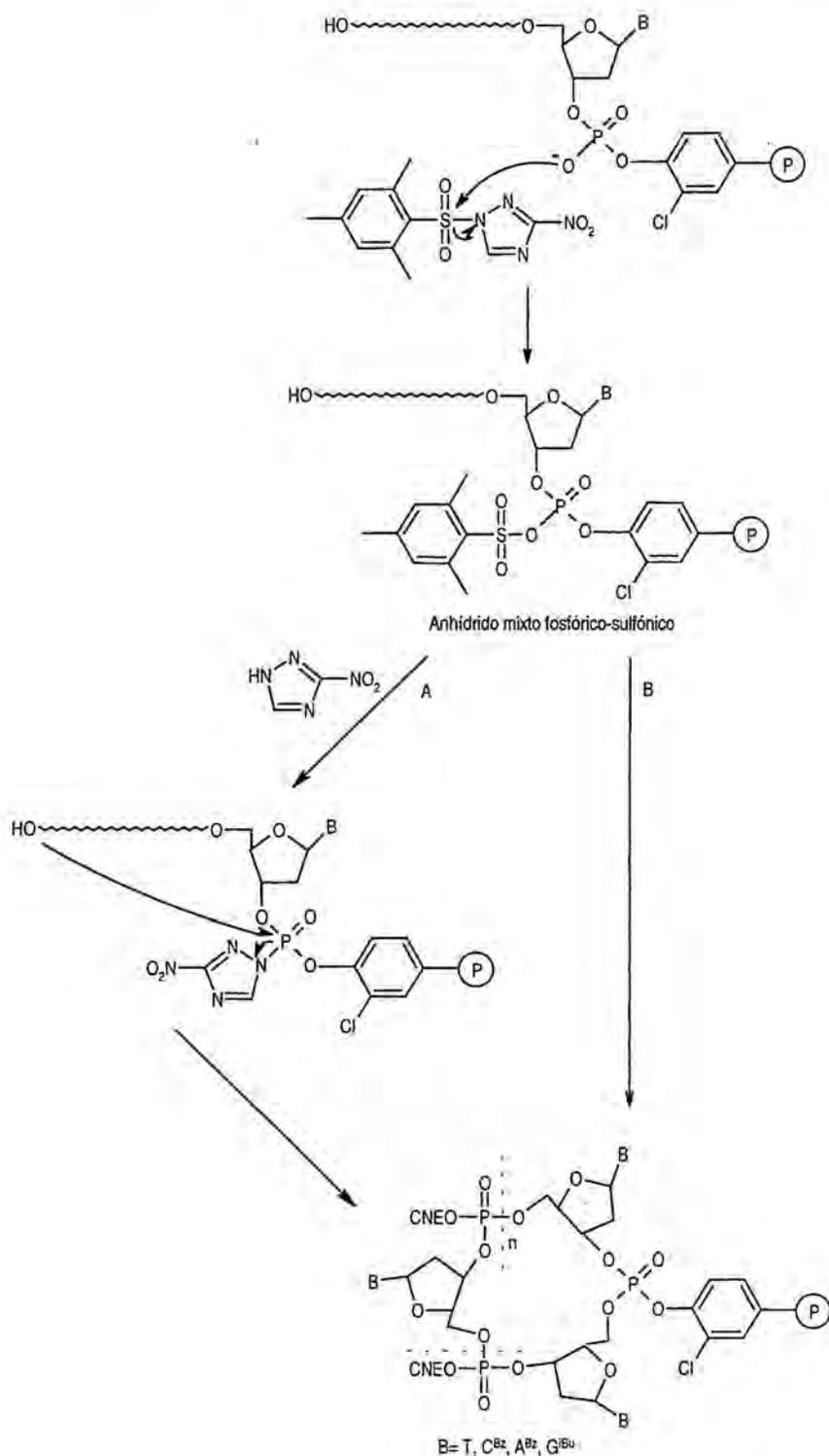


Figura 1.16.- Posibles mecanismos para la reacción de ciclación del precursor lineal anclado a resina con MSNT.

### 1.1.3.1.- Elección de reactivos de acoplamiento.

De entre los muchos reactivos de acoplamiento descritos en la literatura como eficaces para la síntesis de oligonucleótidos, hemos seleccionado aquellos que parecen funcionar mejor en acoplamientos difíciles o con un fuerte impedimento estérico o que, sencillamente, se encuentran descritos como mucho más activos que el MSNT.

- *Cloruro de N,N-bis[2-oxo-3-oxazolidinil]fosfordiamidato (BOPDC).*

Este reactivo fue descrito, por primera vez, para la activación de ácidos carboxílicos y la posterior obtención de los correspondientes anhídridos de ácido<sup>28,29</sup>. Los buenos resultados obtenidos y el paralelismo existente entre la formación de estos anhídridos y la formación de fosfatotriesteres, que se sabe que tiene lugar a través de intermedios pirofosfato<sup>24</sup> propició la aplicación de este reactivo a la síntesis de oligonucleótidos<sup>30</sup>. La obtención de rendimientos elevados, así como la comprobación de que no tenía lugar ninguna reacción secundaria, confirmó que se trataba de una buena alternativa a los reactivos habitualmente utilizados en el método del fosfatotriester. En la Figura 1.17, se muestra el que sería el mecanismo de reacción propuesto aplicado a la obtención de estructuras cíclicas.

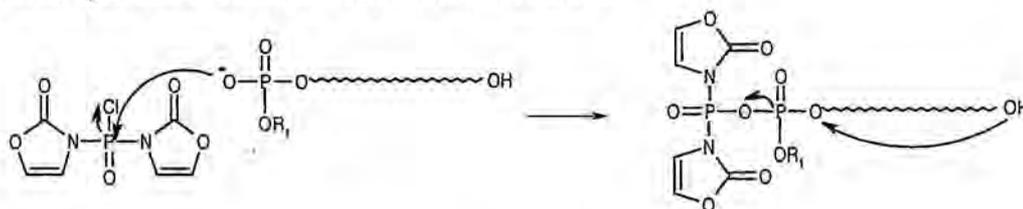


Figura 1.17.- Mecanismo propuesto para la activación de fosfatodiesteres con BOPDC.

- Sales de fosfonio, reactivos BOP:

Los hexafluorofosfatos de benzotriazol-1-iloxi-tris(dimetilamino)fosfonio (BOP)<sup>31</sup>, son reactivos cuya efectividad ha sido probada tanto en disolución como en fase sólida. Aunque su utilización se encuentra mucho más extendida en la síntesis de péptidos, su aplicación a la síntesis de oligonucleótidos ha dado también muy buenos resultados. En particular, se encuentra descrito en la literatura como reactivo muy efectivo el hexafluorofosfato de 6-nitrobenzotriazol-1-iloxi-tris(dimetilamino)fosfonio (NBOP), para el que no se han detectado reacciones secundarias<sup>32</sup>. La mayor efectividad de este reactivo frente a otros derivados BOP se debe al efecto atrayente de electrones que ejerce el grupo NO<sub>2</sub> en posición 6.

El NBOP, sin embargo, no es un reactivo comercial, por lo que decidimos utilizar otro derivado BOP con características similares. Elegimos para ello el PyAOP<sup>33,34</sup>, que es un

<sup>28</sup> Diago, J; Palomo, A.L. *Synthesis* **1980**, 547-551.

<sup>29</sup> Cabré, J.; Palomo, A.; Palomo, A.L. *Synthesis* **1981**, 616-620.

<sup>30</sup> Katti, S.B.; Agarwal, K.L. *Tetrahedron Lett.* **1985**, 26, 2547-2550.

<sup>31</sup> Castro, B.; Dormay, J.R.; Evin, G.; Selve, C. *Tetrahedron Lett.* **1975**, 14, 1219-1223.

<sup>32</sup> Hashmi, S.A.N.; Adiraj, M.; Kumar, A.; Katti, S.B. *Nucleosides & Nucleotides* **1994**, 13, 1059-1067.

<sup>33</sup> Carpino, L.A.; El-Faham, A.; Minor, C.A.; Albericio, F. *J. Chem. Soc. Commun.* **1994**, 201-203.

<sup>34</sup> Albericio, F.; Cases, M.; Alsina, J.; Triolo, S.A.; Carpino, C.A.; Katei, S.A. *Tetrahedron Lett.* **1997**, 30, 4853-4856.

derivado piridinotriazol, en el que el nitrógeno piridínico tendría un efecto atrayente de electrones similar al que ejerce el grupo  $\text{NO}_2$  en posición 6 del NBOP. En la Figura 1.18 se muestran ambos reactivos.

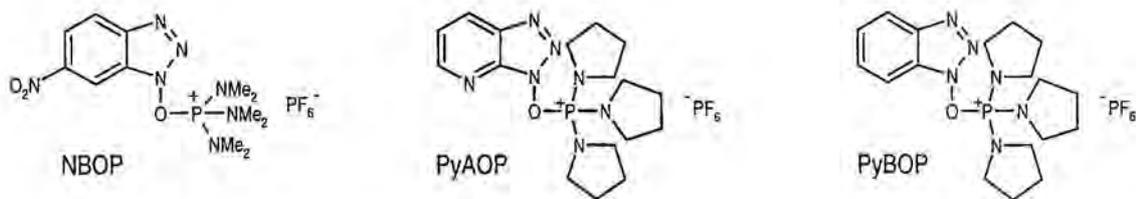


Figura 1.18.- Estructuras de los reactivos NBOP, PyAOP y PyBOP.

Además del PyAOP decidimos probar también el PyBOP<sup>35</sup>. Ambos reactivos presentan la ventaja de que, al tratarse de derivados pirrolidino, durante la reacción de acoplamiento dan lugar a tripirrolidinofosfóricotriamida en lugar de hexametilfosfóricotriamida (HMPA), de elevada cancerogenicidad. En la Figura 1.20 se muestra el mecanismo de reacción propuesto para este tipo de reactivos.

- *Sales de uronio:*

Las sales de uronio se encuentran descritas como alternativa a la utilización de reactivos BOP como agentes acoplantes en síntesis de péptidos<sup>36</sup>. Aunque su efectividad se ha comprobado, por tanto, únicamente para la formación de enlaces amida se cita al TBTU (tetrafluoroborato de 2-(1H-benzotriazol-1-ilo)-1,1,3,3-tetrametiluronio), Figura 1.19, como un reactivo especialmente adecuado en el caso de síntesis en fase sólida. Por este motivo, así como el hecho de que el mecanismo de reacción sea muy similar al de las sales de fosfonio, que sí que han sido descritas en síntesis de oligonucleótidos, hemos creído conveniente incluirlo también en nuestro estudio. El mecanismo de reacción se muestra en la Figura 1.20.

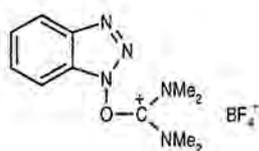


Figura 1.19.- Estructura del TBTU.

<sup>35</sup> Coste, J.; Le-Nguyen, D.; Castro, B. *Tetrahedron Lett.* **1990**, 31, 205-208.

<sup>36</sup> Knorr, R.; Trzeciak, A.; Bannwarth, W.; Gillissen, D. *Tetrahedron Lett.* **1989**, 30, 1927-1930.

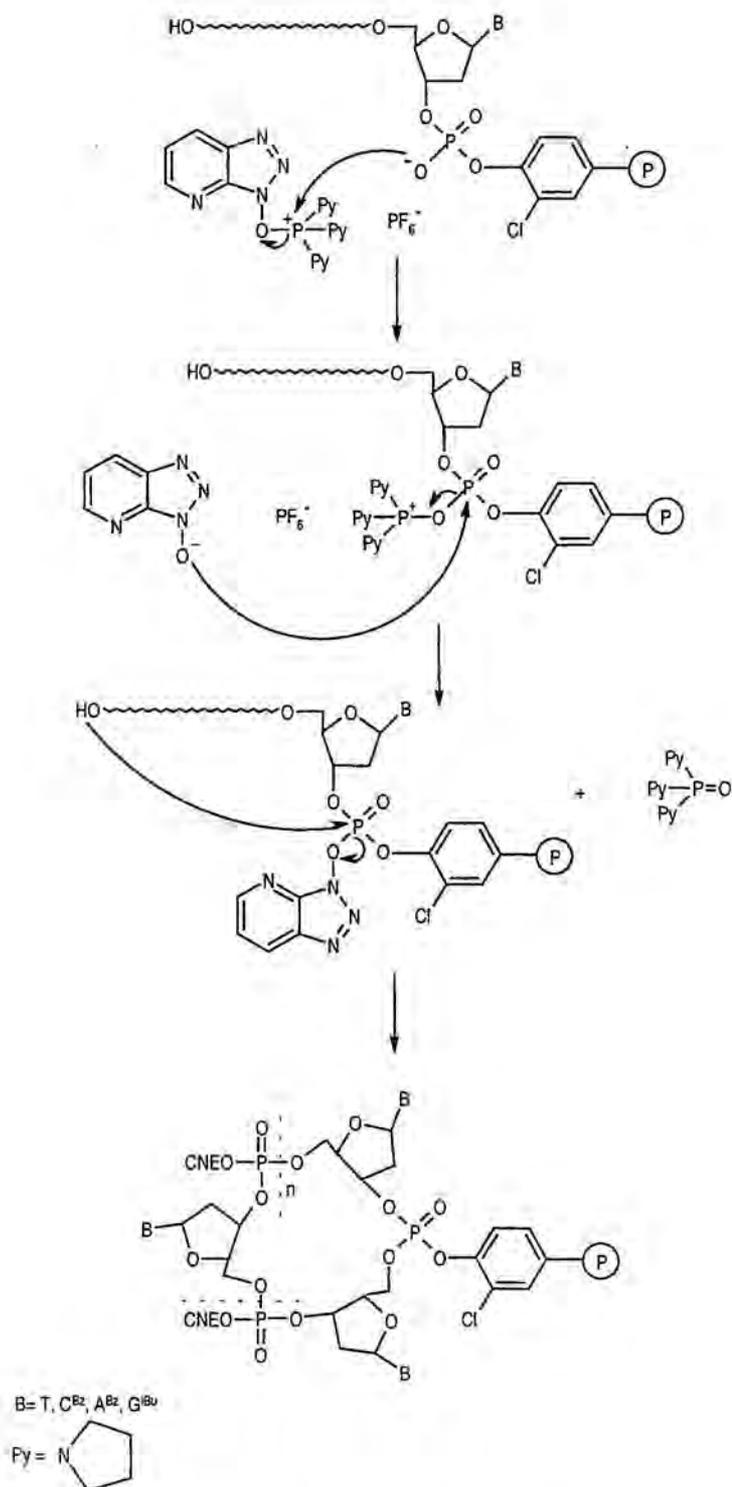
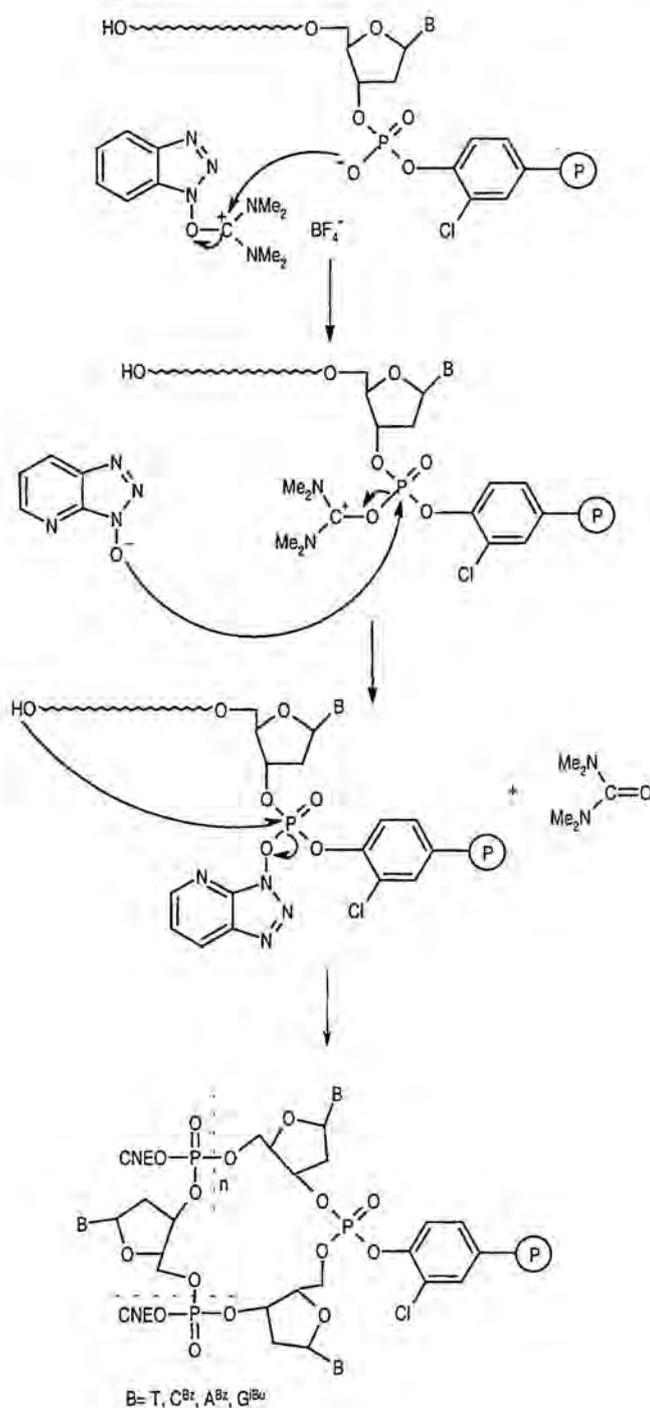


Figura 1.20.- Mecanismo de reacción propuesto para los reactivos BOP.



**Figura 1.21.-** Mecanismo de reacción del TBTU.

### 1.1.3.2.- Elección de la secuencia.

Un punto importante en el estudio de los distintos reactivos de ciclación es la elección de la secuencia. Puesto que la composición nucleotídica, así como la longitud de la cadena, pueden afectar de forma importante a la reacción de ciclación, era necesario escoger una secuencia que fuera lo más representativa posible y que presentara ciertas dificultades para

poder discriminar entre los diferentes reactivos. Para ello se tuvieron en cuenta diferentes factores:

- debería tener un tamaño relativamente grande dentro del rango de oligonucleótidos cíclicos obtenidos en nuestro laboratorio (2 a 16-mer),
- debería contener las cuatro bases,
- que la condensación tuviera lugar purina sobre pirimidina, que suele ser menos favorable que pirimidina sobre pirimidina.

La secuencia elegida resultó ser el dodecámero 5'-ACGCTATGCGAT-3', que podría formar una estructura tipo dumbbell estabilizada por tres pares G-C y en el que los mini-loops estarían constituidos por tres residuos: 5'-TAT-3' y 5'-ATA-3'.

### 1.1.3.3.- Ensayos con los distintos reactivos.

La síntesis del precursor lineal se realizó a escala de 10  $\mu\text{mol}$ , sobre un soporte de TentaGel funcionalizado con el ácido 6-aminohexanoico, y siguiendo el esquema de protección 4: R1 = CNE y R2 = CNE, con eliminación de R1 antes de la elongación de la cadena. Se utilizó una nucleotidil-resina con una funcionalización de 118  $\mu\text{M}$ .

Una vez finalizada la síntesis del precursor lineal, para el que se obtuvo un Rdt global del 90% (Rdt medio por etapa de 99%), se fraccionó la oligonucleotidil-resina en alícuotas de 1  $\mu\text{mol}$ , escala a la que se realizaron los ensayos con los distintos reactivos de ciclación.

Como condiciones de reacción elegidas para probar los distintos reactivos, se mantuvieron las habitualmente utilizadas en el caso del MSNT: disolución 0.1 M del reactivo de acoplamiento en piridina anh. (10 eq de exceso respecto a la escala de trabajo, 1  $\mu\text{mol}$ ). Puesto que en la literatura está descrito que el N-metilimidazol (NMI), juntamente con el reactivo de acoplamiento, actúa como activador de la reacción<sup>37</sup>, y para casi todos los reactivos elegidos las condiciones óptimas de reacción implican la presencia de NMI, se decidió mantener en todos los casos una concentración de N-metilimidazol de 0.1 M, también en piridina anh. En el caso del TBTU, que no es soluble en piridina, se utilizó ACN anh. para llevar a cabo la reacción.

Como ya se ha comentado anteriormente es muy importante que la reacción se lleve a cabo en condiciones estrictamente anhidras. Para ello, los distintos reactivos de síntesis se mantuvieron en el desecador durante 12 h antes de su utilización y se manipularon siempre bajo atmósfera de argón.

Un reactivo que presentó algunos problemas fue el BOPDC. Aunque está descrita su utilización en medio de piridina anh., los distintos intentos realizados para disolver este reactivo en dicho disolvente fueron infructuosos. Se pensó que tal vez el producto podría presentar impurezas correspondientes al producto hidrolizado, así que se disolvió el reactivo en DCM y se realizaron extracciones con H<sub>2</sub>O. La fase orgánica se evaporó a sequedad y se realizaron varias coevaporaciones con piridina anh. antes de disolverlo para su utilización. Aunque el espectro de <sup>31</sup>P-RMN que se realizó confirmaba la pureza del reactivo, la disolución en ningún caso resultó total. En la literatura se hace referencia a la importancia de la pureza y buen estado de este reactivo para su utilización como reactivo de acoplamiento y se siguió el protocolo descrito para la recristalización del reactivo en ACN<sup>38</sup>. Tras dos recristalizaciones en ACN se obtuvo un

<sup>37</sup> Efimov, V.A.; Reverdatto, S.V.; Chakhmakhcheva, O.G. *Tetrahedron Lett.* **1982**, 23, 961-964.

<sup>38</sup> Van der Auwera et al. *Bull. Soc. Chim. Belg.* **1986**, 95, 203-205.

producto de elevada pureza por  $^{31}\text{P}$ -RMN y un  $\text{pf} = 193\text{-}195^\circ\text{C}$ , muy próximo al descrito:  $195\text{-}198^\circ\text{C}$ . Sin embargo, tampoco en este caso resultó totalmente soluble en piridina. Puesto que estaba descrita la utilización del BOPDC conjuntamente con N-metilimidazol, en este punto decidimos probar si la adición de este reactivo era el factor que facilitaba la disolución del BOPDC en piridina. En efecto, pudimos comprobar como al añadir N-metilimidazol sobre el BOPDC en piridina, poco a poco iba desapareciendo el producto insoluble.

Para llevar a cabo las reacciones de ciclación, se preparó una disolución 0.1 M de N-metilimidazol en piridina anh. y en ella se disolvieron los distintos reactivos de acoplamiento. Nuevamente, el BOPDC dió problemas y, después de varias coevaporaciones con piridina anh., al añadir la disolución de NMI no se disolvía totalmente. Este problema lo solventamos adicionando más NMI hasta la disolución total del reactivo, por lo que, en este caso, el exceso de NMI en el medio de reacción es muy superior al utilizado para el resto de reactivos: BOPDC/NMI  $\sim 1:100$ .

Por tanto, las condiciones de reacción finales fueron las siguientes: disolución 0.1 M de reactivo de ciclación y NMI (20 eq de exceso respecto a la escala de trabajo) en piridina anh., excepto en el caso del BOPDC, en que se empleó un exceso de NMI mucho mayor, y del TBTU para el que la reacción se realizó en ACN. Se realizaron dos tratamientos de 4 h, utilizando para cada tratamiento una disolución de reactivo acabada de preparar.

Tal como se había observado alguna vez con el MSNT, la resina se tornó de un color amarillento-naranja durante los tratamientos con PyBOP y PyAOP y, después de la reacción de ciclación, durante la eliminación de los grupos CNE con TEA/piridina la resina tomó una coloración rojo intenso. Esta formación de subproductos coloreados se observa frecuentemente cuando se realiza la reacción en piridina<sup>37</sup>

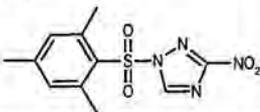
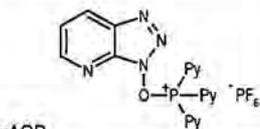
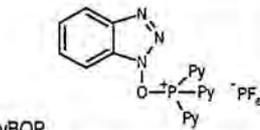
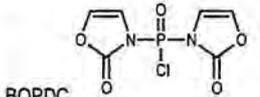
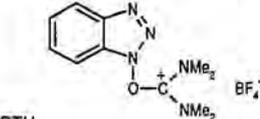
Después de los tratamientos con los distintos reactivos de ciclación, y de la eliminación de los grupos protectores R2, se llevó a cabo el desanclaje del producto, para lo que se realizaron 3 tratamientos con oximato de TMG 0.2 M en dioxano/ $\text{H}_2\text{O}$  1:1. Posteriormente se trataron los filtrados con  $\text{NH}_3$  para eliminar los protectores de las bases y se procedió al desalado a través de una columna de Sephadex G-10.

En la Tabla 1.12 se resumen los datos procedentes del análisis por HPLC de los crudos, Figura 1.21, así como un dato que resulta también muy importante de cara a comparar los distintos reactivos de ciclación, como es el rendimiento de obtención de crudo en cada caso.

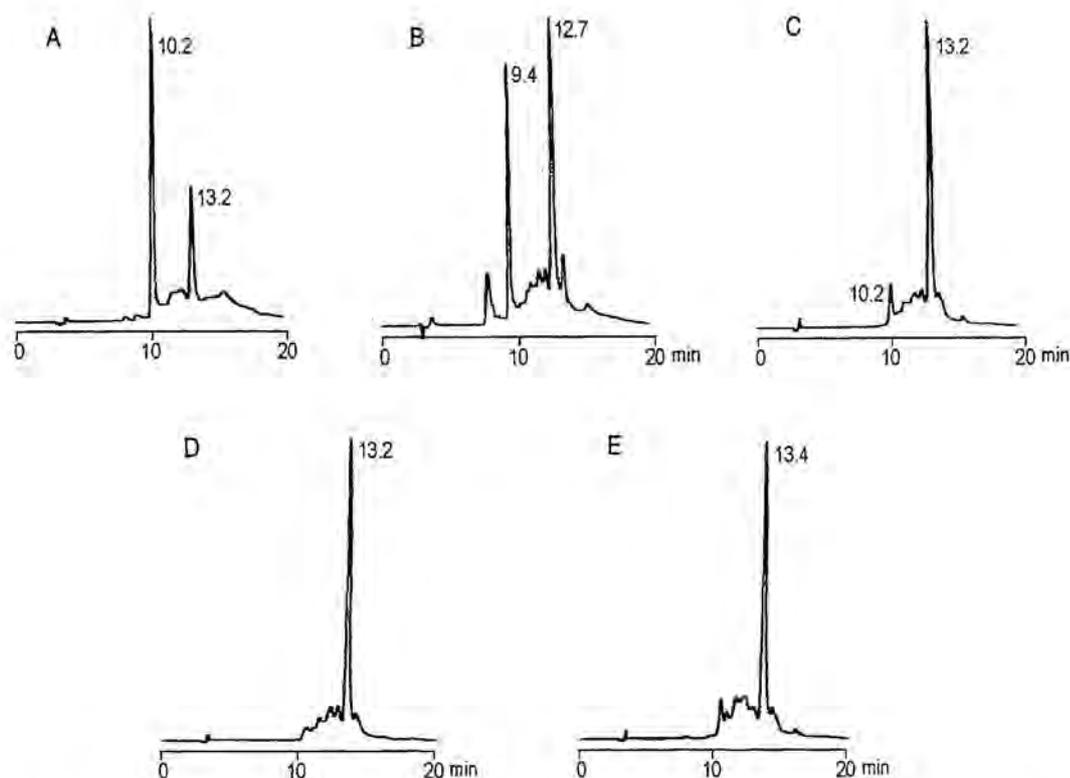
En relación a este último punto, el rendimiento de obtención de crudo, es donde encontramos los primeros resultados negativos. En todos los casos se han obtenido porcentajes de crudo extremadamente bajos, incluso con el reactivo habitualmente utilizado, MSNT, con el que son habituales rendimientos entre 30-40%.

Por otra parte, si bien todos los crudos presentan un elevado grado de pureza (Figura 1.21), existe una diferencia importante y es que en los cromatogramas correspondientes a MSNT y PyAOP se observa la presencia de dos picos mayoritarios, mientras que para los otros tres reactivos (PyBOP, BOPDC y TBTU) se observa un único pico mayoritario. Por coinyección de los distintos crudos, se pudo comprobar que el pico a 10.2 min (cromatograma A) y 9.4 min (cromatograma B) coelúan y que también el pico a  $\sim 13$  min correspondía, en todos los casos, al mismo producto.

**Tabla 1.12.-** Resumen de los rendimientos de obtención de crudo en cada caso y información extraída del análisis por HPLC.

Reactivo de acoplamiento	Rdt crudo %	Análisis por HPLC
MSNT 	12	Se observan dos picos mayoritarios: $t_R = 10.2$ min, 30% $t_R = 13.2$ min, 19%
PyAOP 	9.5	Se observan dos picos mayoritarios: $t_R = 9.4$ min, 18% $t_R = 12.7$ min, 30%
PyBOP 	7	Se distinguen dos picos, uno de ellos muy mayoritario: $t_{Rmin} = 10.2$ min, 8% $t_{Rmay} = 13.2$ min, 52%
BOPDC 	7.5	Se observa un pico muy mayoritario: $t_R = 13.3$ min, 55%
TBTU 	13	Se observa un pico muy mayoritario: $t_R = 13.4$ min, 45%

Tras la separación de ambos productos por HPLC, se llevó a cabo la caracterización por espectrometría de masas MALDI-TOF, así como la coinyección en HPLC con una muestra patrón del producto cíclico. Se confirmó que el pico a menor tiempo de retención, y que únicamente aparece en A y B, correspondía al producto cíclico deseado, mientras que el producto con tiempo de retención  $\sim 13$  min presentaba un incremento de masa de  $\sim 187$  unidades respecto al producto cíclico. Este último resultado, unido al hecho de que dicho producto es mayoritario en cuatro de los cinco crudos, resultaba significativamente importante ya que implicaba el desanclaje de un producto distinto al cíclico deseado. La caracterización detallada de este subproducto se describe en el apartado 1.1.3.4.



**Figura 1.22.-** Cromatogramas HPLC de los crudos obtenidos con cada uno de los reactivos de acoplamiento: A) MSNT, B) PyAOP, C) PyBOP, D) BOPDC y E) TBTU. Gradiente utilizado: 10-40% B en 20 min.

Los resultados obtenidos para los ensayos con distintos reactivos de ciclación ponen de manifiesto varias cosas:

- El mejor reactivo de acoplamiento para llevar a cabo la ciclación parece ser el MSNT, ya que en este caso se obtiene como producto mayoritario el producto deseado.
- La secuencia elegida presenta graves dificultades a la hora de ciclar. Estas dificultades creemos que no son tanto debidas al tamaño, ya que en nuestro grupo se han sintetizado otros dodecámeros con rendimientos aceptables y un elevado grado de pureza, sino que más bien pueden ser debidas a la composición oligonucleotídica y al hecho de que el acoplamiento tenga lugar purina sobre pirimidina.
- El éxito de la reacción de ciclación, además de la secuencia, depende mucho de las condiciones experimentales en las que se lleva a cabo. Es importante destacar que algunos de los ensayos se repitieron varias veces, teóricamente en las mismas condiciones, obteniéndose a veces resultados muy diferentes. En particular, en el caso del MSNT, en un primer intento se obtuvo únicamente el subproducto anteriormente mencionado, mientras que en el ensayo descrito se obtiene mayoritariamente el producto cíclico deseado. Esta dependencia de las condiciones experimentales se observa también, si cabe más claramente, en el caso de los distintos intentos que se han realizado

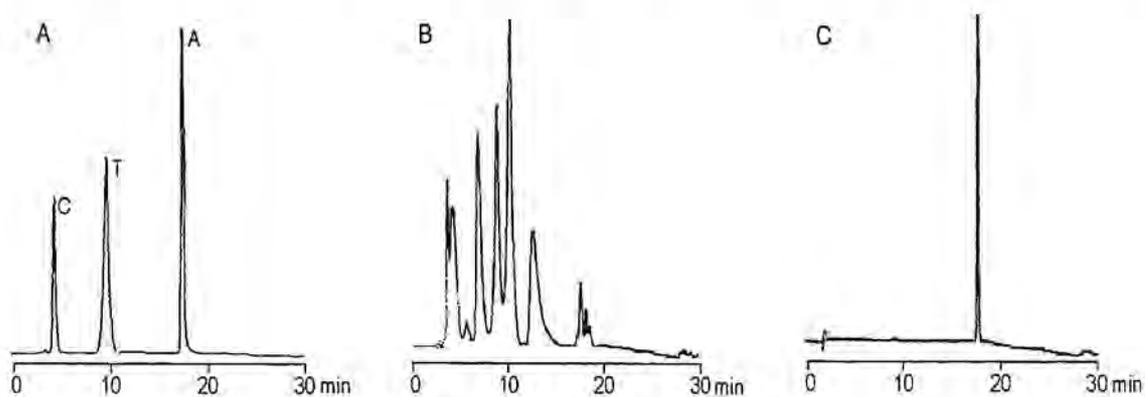
para llevar a cabo un estudio cinético de la reacción de ciclación, que se detallan más adelante.

#### 1.1.3.4.- Caracterización del Producto Secundario con $\Delta = 187$ umas.

Un rasgo común a las síntesis de oligonucleótidos cíclicos, y que además confirmaba que la reacción de desanclaje del producto ciclado tenía lugar de forma selectiva, era la obtención de crudos de síntesis muy limpios donde el producto mayoritario se correspondía siempre con el producto deseado. La aparición de un subproducto en unas síntesis en las que, no solamente era el producto mayoritario, sino que, además, no se apreciaba la presencia del producto cíclico deseado, representaba un problema importante desde diferentes puntos de vista.

En las numerosas síntesis de oligonucleótidos cíclicos realizadas en nuestro grupo, únicamente se había dado el caso de la aparición de un subproducto en una proporción importante de forma muy ocasional y se había atribuido a algún error cometido en las múltiples etapas de que consta el procedimiento. En una síntesis de  $d\langle pCATTTCATT \rangle$  realizada a gran escala se aisló el subproducto mayoritario y se intentó su caracterización por espectrometría de masas y digestión enzimática. Aunque no se pudo determinar de forma precisa de qué producto se trataba sí presentaba dos rasgos muy característicos: presentaba un incremento de masa importante respecto al producto deseado y no era estable a la digestión con fosfodiesterasa de bazo de buey, lo que indicaba que se trataba de un producto lineal. El ensayo enzimático nos proporcionó además otra información importante: se cumplía la proporción de bases esperada para la secuencia ( C/T/A 1.0: 1.9 : 1.1) y no se había producido ninguna modificación en las bases ya que el tiempo de retención de los nucleósidos en HPLC era correcto.

En la Figura 1.23 se muestran los cromatogramas HPLC del crudo de digestión enzimática con fosfodiesterasa de veneno de serpiente (A) y bazo de buey (B) de este producto secundario y se ha incluido también el cromatograma correspondiente a la digestión del producto cíclico deseado  $d\langle pCATTTCATT \rangle$ , en el que se observa el pico de producto inalterado.

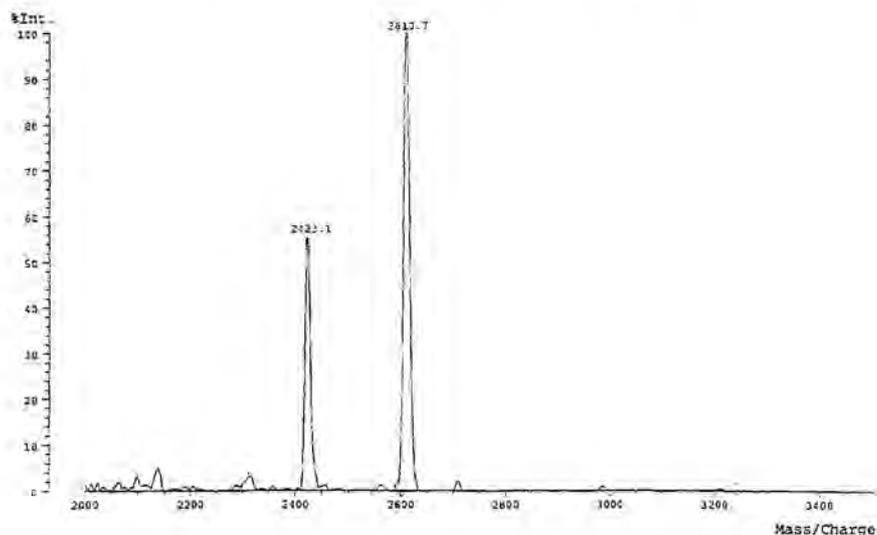


*Figura 1.23.- Cromatograma HPLC del crudo de digestión enzimática del producto secundario con A) fosfodiesterasa de veneno de serpiente y B) fosfodiesterasa de bazo de buey. El cromatograma C corresponde a la digestión del producto cíclico  $d\langle pCATTTCATT \rangle$  con fosfodiesterasa de bazo de buey.*

El hecho de que el producto secundario obtenido en los ensayos con los distintos reactivos de ciclación, presentara características muy similares, nos hizo pensar que se tratara

del mismo tipo de subproducto. Por este motivo se creyó conveniente profundizar en su estudio y caracterización.

Se llevaron a cabo distintos intentos de definir de forma precisa el incremento de masa del subproducto respecto del producto deseado. Espectros de masas (MALDI-TOF) de mezclas del subproducto con patrones del producto cíclico, tanto para el octámero como para el dodecámero, nos llevaron a fijar dicha diferencia en 187 umas, Figura 1.24.



**Figura 1.24.-** Espectro de masa MALDI-TOF correspondiente a la mezcla del subproducto y el producto cíclico de la secuencia d<pCATT<CATT>.

Después de barajar distintas opciones, y teniendo en cuenta que no se habían producido modificaciones en las bases, se dedujo que este incremento de masa podría corresponder al linker de ácido 3-cloro-4-hidroxifenilacético. Esto significaba que el subproducto se formaría como consecuencia del desanclaje del producto lineal, unido al linker, por escisión del enlace amida que lo une al soporte polimérico, Figura 1.25.

Para confirmar esta hipótesis, se disponía de muestra suficiente de subproducto como para intentar su caracterización por  $^1\text{H}$ -RMN, puesto que la síntesis de d<pCATT<CATT> anteriormente mencionada se había realizado a gran escala y se había obtenido mayoritariamente el subproducto. Se registraron espectros monodimensionales en  $\text{D}_2\text{O}$ , que confirmaron que se trataba de un oligonucleótido lineal ya que se observaban señales correspondientes a ocho residuos (en el caso del oligonucleótido cíclico, al ser simétrico, se observarían únicamente cuatro) y para la asignación completa se realizó un experimento TOCSY. En la zona aromática se pudieron identificar, además de las señales correspondientes a los protones aromáticos de las bases, los protones del grupo *o*-clorofenilo, Figura 1.26.

El hecho de que la formación de este subproducto suponga la escisión del enlace amida entre el espaciador y el soporte, así como que las síntesis del octámero cíclico y del dodecámero se realizaran con la misma nucleotidil-resina, nos hizo pensar en un primer momento, que podría tratarse de un problema asociado al soporte polimérico o bien al modo en el que éste se había funcionalizado. Sin embargo, coincidiendo en el tiempo, se observó la aparición del

mismo tipo de subproducto en otras síntesis realizadas en nuestro laboratorio que no guardaban relación alguna ni entre sí ni con las que aquí se describen. En la Tabla 1.13, se muestran algunas de las síntesis en las que ha aparecido este subproducto, así como los datos experimentales más representativos.

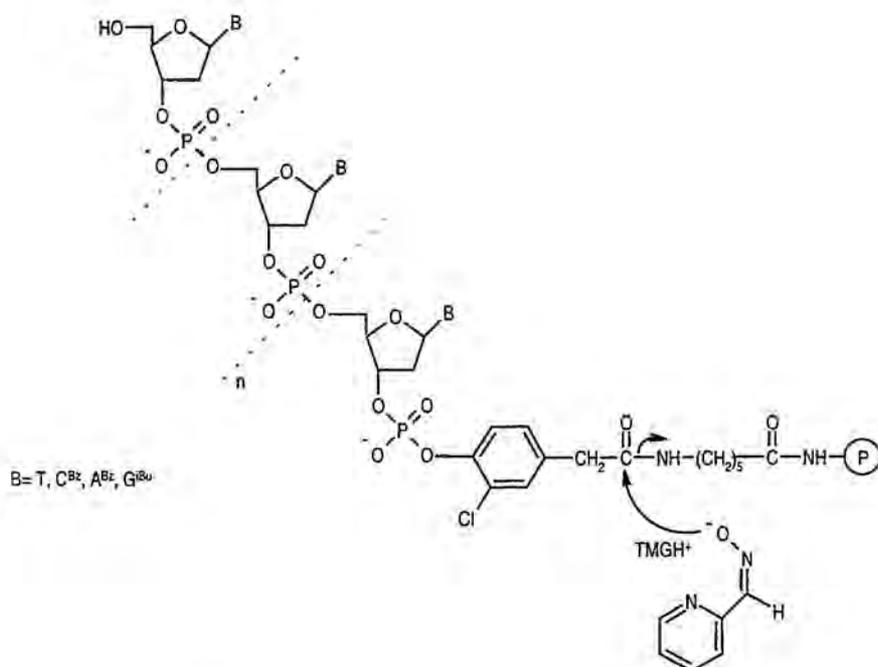


Figura 1.25.- Reacción secundaria que daría lugar al oligonucleótido lineal unido al espaciador.

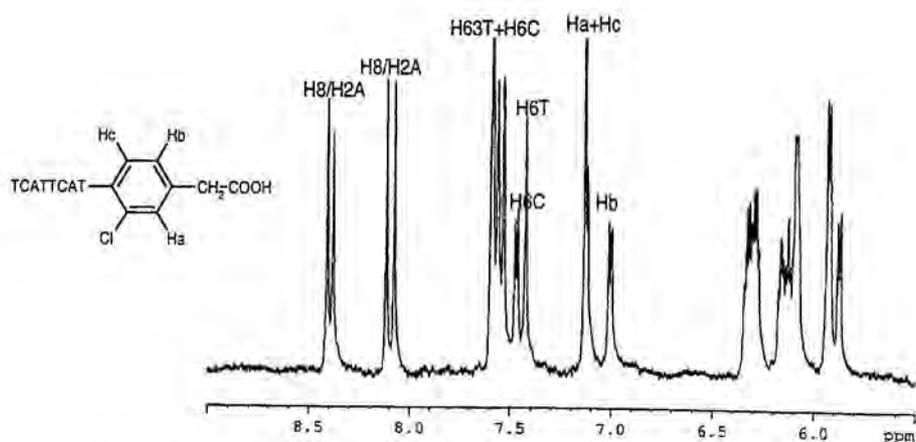


Figura 1.26.- Espectro 1D del subproducto de la síntesis  $d\langle pCATT\langle^d\rangle CATT\rangle$  en  $D_2O$ .

**Tabla 1.13.-** Resumen de distintas síntesis en las que se ha observado el subproducto con un incremento de masa de 187 unidades respecto al producto deseado.

Secuencia	f nucleotidil-resina μmol/g	Rdt. crudo %	HPLC % subproducto	Δmasa umas	Condiciones de ciclación
5'-AGGTTGGTATGGTTGGAT-3' <sup>19</sup>	152.4	12	87	187	MSNT/NMI 1:1 4 eq 3 + 12 + 3 + 4 h
5'-AGGTTGGTATGGTTGGAT-3' <sup>19</sup>	89	11	80	187	MSNT/NMI 1:1 4 eq 3 + 12 + 3 + 4 h
5'-AGGTTGGTATGGTTGGAT-3' <sup>19</sup>	35	3.3	75	187	MSNT/NMI 1:1 4 eq 2 + 4 + 3 + 12 h
5'-UUUUUT-3' <sup>39</sup>	68	17	60	186	BOPDC/NMI 1:1 10 eq 4 + 4 + 12 h
5'-UUUUUU <sub>OMe</sub> -3' <sup>39</sup>	90	15	70	186	BOPDC/NMI 1:1 10 eq 4 + 4 + 12 h
5'-CATT CATT-3'	74	12	81	187	MSNT/NMI 1:1 10 eq 4 + 4 h
5'-CATT CATT-3'	74	10	57	187	MSNT/NMI 1:1 10 eq 4 + 4 h
5'-ACGCTATGCGAT-3'	74	17	60	187	Distintos reactivos de ciclación + NMI 4 + 4 h

Las síntesis en las que se ha observado el subproducto presentan una serie de rasgos comunes:

- se trata de síntesis en las que la elongación de la cadena del oligonucleótido lineal funciona bien, con rendimientos globales entre 71-93%, en función de la longitud de la secuencia, y, sin embargo, se obtienen rendimientos de crudo muy bajos, del orden de 10-15%,
- se obtienen crudos de elevada pureza donde el producto mayoritario corresponde al producto secundario y, en caso de aparecer otro pico en proporción importante, éste corresponde al producto cíclico deseado,
- todos los subproductos presentan el ya comentado incremento de masa de 187 umas.

Puesto que la síntesis del oligonucleótido lineal tiene lugar correctamente, parece que la causa de la aparición de este subproducto puede estar asociada a un bajo rendimiento de la reacción de ciclación. Cuando la reacción de ciclación funciona bien, durante el tratamiento con oximato se separan del soporte principalmente las cadenas cíclicas de oligonucleótido, por escisión del fosfato triéster de *o*-clorofenilo. Si, por el contrario, la reacción de ciclación no funciona bien, el oximato puede atacar otras posiciones más o menos lábiles, como el enlace amida que une el espaciador a la resina, liberando cadenas lineales unidas al espaciador. En este sentido, está descrito en la literatura que el oximato de TMG puede dar lugar a rotura parcial de enlaces amida formados entre las bases y sus grupos protectores.

Un estudio más profundo para tratar de evaluar si, en efecto, es esto lo que ocurre y intentar minimizar la aparición de este subproducto se está llevando a cabo en nuestro laboratorio<sup>39</sup>.

<sup>39</sup> Frieden, M. Tesis Doctoral en curso, Universidad de Barcelona.

### 1.1.3.5.- Estudio cinético de la reacción de ciclación con MSNT.

Ante los resultados obtenidos en el intento de utilizar otros reactivos de acoplamiento alternativos al MSNT, decidimos intentar optimizar el tiempo de reacción cuando se utiliza MSNT/NMI. Este estudio se realizó sobre la misma secuencia que la utilizada para ensayar los distintos reactivos de ciclación.

Para ello llevamos a cabo un estudio cinético consistente en tratar distintas alícuotas de oligonucleotidilresina (escala 0.5  $\mu$ mol) con una disolución 0.1M de MSNT/NMI en piridina anh. (10 eq), durante distintos periodos de tiempo: 30 min, 1h, 2h, 4h y 12h. A continuación, se llevó a cabo el acoplamiento de un fosforamidito sobre cada una de las fracciones y se determinó la funcionalización en cada caso. Además, sobre cada una de las alícuotas se llevaron a cabo los tratamientos de desanclaje, desprotección y desalado y se cuantificó la fracción de crudo obtenido, para disponer de un dato adicional sobre el transcurso de reacción de ciclación en cada caso. Los resultados se muestran en la Tabla 1.14.

Tabla 1.14.- Datos de la cinética de ciclación con MSNT.

Tiempo de reacción	% grupos 5'-OH reaccionados	Rdt. crudo %
30min	8	14
1h	24	14
2h	17	12
4h	57	27
12h	31	12

A la hora de analizar los valores de ambas columnas es importante tener en cuenta varios factores:

- i) los grupos 5'-OH pueden haber reaccionado intra- o intermolecularmente, por lo que es posible obtener valores de reacción de hidroxilos superiores a los de desanclaje,
- ii) en el apartado anterior se ha demostrado que es posible que, además de producto cíclico, se desanclen parte de producto lineal + linker, por lo que el valor de rendimiento de obtención de crudo también puede verse afectado.

Sin embargo, lo que sí que debería observarse en ambas columnas es una misma tendencia de los valores. En este sentido, los datos procedentes de la cuantificación de grupos DMT, aunque no muestran una tendencia regular, sí que indican una disminución de grupos hidroxilo libres al aumentar el tiempo de reacción. No obstante, a partir de los valores de desanclaje con oximato, a excepción del valor a 4h de reacción, para el resto de tiempos se obtiene prácticamente el mismo porcentaje de obtención de crudo, entre 12-14%.

Estas incoherencias se deben en gran medida a que los datos están sujetos a numerosos errores. En el primer caso influyen de forma importante tanto la eficacia del acoplamiento del fosforamidito como la posterior cuantificación de cationes DMT. En el segundo caso, el valor se obtiene después de una serie de etapas (desanclaje, desprotección y desalado) por lo que el rendimiento final está modulado por los rendimientos de cada una de ellas.

El hecho de que los mejores resultados se hayan obtenido en el caso de  $t = 4h$ , pone de manifiesto la fuerte dependencia de la reacción de ciclación con las variables experimentales. Factores como el estado del reactivo utilizado para cada alícuota, así como el carácter más o menos anhídrico de la disolución de piridina, pueden afectar drásticamente a la reacción y hacer

que se obtengan valores totalmente aleatorios. Por otra parte, los malos resultados obtenidos, en general, ponen de manifiesto que la secuencia utilizada para llevar a cabo el estudio no era quizás la más adecuada. Como ya se discutió en el apartado 1.1.3.3 la secuencia elegida parece presentar graves dificultades a la hora de ciclar.

A la vista de estos resultados se puede concluir que, aunque se pudiera mejorar la regularidad de los datos y poder llegar a obtener unos datos cinéticos coherentes, éstos serían únicamente válidos para la secuencia estudiada. En la práctica hay que forzar las condiciones de ciclación realizando varios tratamientos con reactivo recién preparado y utilizando tiempos largos de reacción, más largos cuanto mayor grado de dificultad parezca presentar la secuencia, que es lo que habitualmente se venía realizando

## 1.2.- SÍNTESIS A GRAN ESCALA DE OLIGONUCLEÓTIDOS CÍCLICOS.

Uno de los objetivos que nos planteamos en la presente Tesis era la aplicación del método de síntesis a la obtención de oligonucleótidos cíclicos a gran escala. Nuestro objetivo era poder llevar a cabo, posteriormente, estudios estructurales de algunas de las secuencias y, para ello, resultaba necesario de disponer de cantidad de producto puro del orden de miligramos.

Ya en los estudios que se realizaron utilizando distintos esquemas de protección se trabajó a gran escala,  $\sim 10 \mu\text{mol}$ . En este apartado se describirán síntesis realizadas entre  $\sim 6.5$ - $18.8 \mu\text{mol}$ . El principal problema que podría presentarse al trabajar a gran escala es que, al resultar necesario utilizar soportes con funcionalizaciones elevadas se estaría favoreciendo las reacciones intercadena. Sin embargo, estudios sistemáticos realizados en nuestro grupo<sup>19</sup>, demuestran que tal correlación no es observable, al menos en los rangos de funcionalización estudiados y que, incluso, los mejores resultados se obtienen para los soportes más funcionalizados.

En todos los casos se ha aplicado el método de síntesis general, utilizando el esquema de protección 4: R1 = CNE y R2 = CNE, con eliminación de R1 antes de la elongación de la cadena. Para las reacciones de ciclación se ha utilizado como reactivo de acoplamiento el MSNT, que se ha confirmado como el reactivo más adecuado, y se han llevado a cabo con y sin NMI como catalizador.

En la Tabla 1.15 se resumen las distintas síntesis realizadas a gran escala, así como las condiciones en las que se han llevado a cabo cada una de ellas. La caracterización de las T-resinas utilizadas por  $^{31}\text{P}$ -RMN se muestra en la Tabla 1.8.

**Tabla 1.15.-** Resumen de las condiciones en las que se han llevado a cabo las distintas síntesis, así como los valores de rendimiento de obtención de crudo y pureza por HPLC. Entre paréntesis se muestra el rendimiento correspondiente a ciclación, desprotección y desanclaje, excluyendo el rendimiento de la síntesis lineal.

Secuencia	Escala $\mu\text{mol}$	f nucleotidil-resina $\mu\text{mol/g}$	Rdt. síntesis lineal %	Condiciones de ciclación	Rdt. crudo %	Pureza HPLC %
d<pCATT<math>\text{CATT}>	10	70 (1)	86	MSNT 0.14 M en piridina anh., 10 eq, 4 + 4 + 12 h	36 (41)	80
d<pCATT<math>\text{CATT}>	11	70 (1)	86	MSNT 0.14 M en piridina anh., 10 eq, 4 + 4 + 12 h		
d<pCAT<math>^{\text{B}}\text{UCAT}^{\text{B}}\text{U}>	10	96 (2)	90	MSNT 0.14 M en piridina anh., 10 eq, 4 + 4 + 12 h	25 (28)	78
d<pCGCT<math>\text{CGCT}>	7	47 (1)	86	MSNT 0.14 M en piridina anh., 10 eq, 4 + 4 + 12 h	32 (37)	86
d<pCTAT<math>\text{CTAT}>	6.5	43 (1)	92	MSNT 0.14 M en piridina anh., 10 eq, 4 + 4 + 12 h	30 (35)	90
d<pTGCT<math>\text{CGCT}>	7	47 (1)	93	MSNT 0.14 M en piridina anh., 10 eq, 4 + 4 + 12 h	40 (46)	76
d<pCCGT<math>\text{CCGT}>	18.8	118 (1)	86	MSNT/NMI 1:1, 0.1 M en piridina anh., 10 eq, 4 + 4 h	33 (38)	80
d<pTGCAT<math>\text{GCA}>	18.8	118 (1)	92	MSNT/NMI 1:1, 0.1 M en piridina anh., 10 eq, 4 + 4 h	31 (36)	77
d<pAGCT<math>\text{AGCT}>	12.3	118 (1)	94	MSNT/NMI 1:1, 0.1 M en piridina anh., 10 eq, 4 + 4 h	27 (30)	90
d<pCGCT<math>\text{CATT}>	12.3	118 (1)	93	MSNT/NMI 1:1, 0.1 M en piridina anh., 10 eq, 4 + 4 h	32 (35)	61

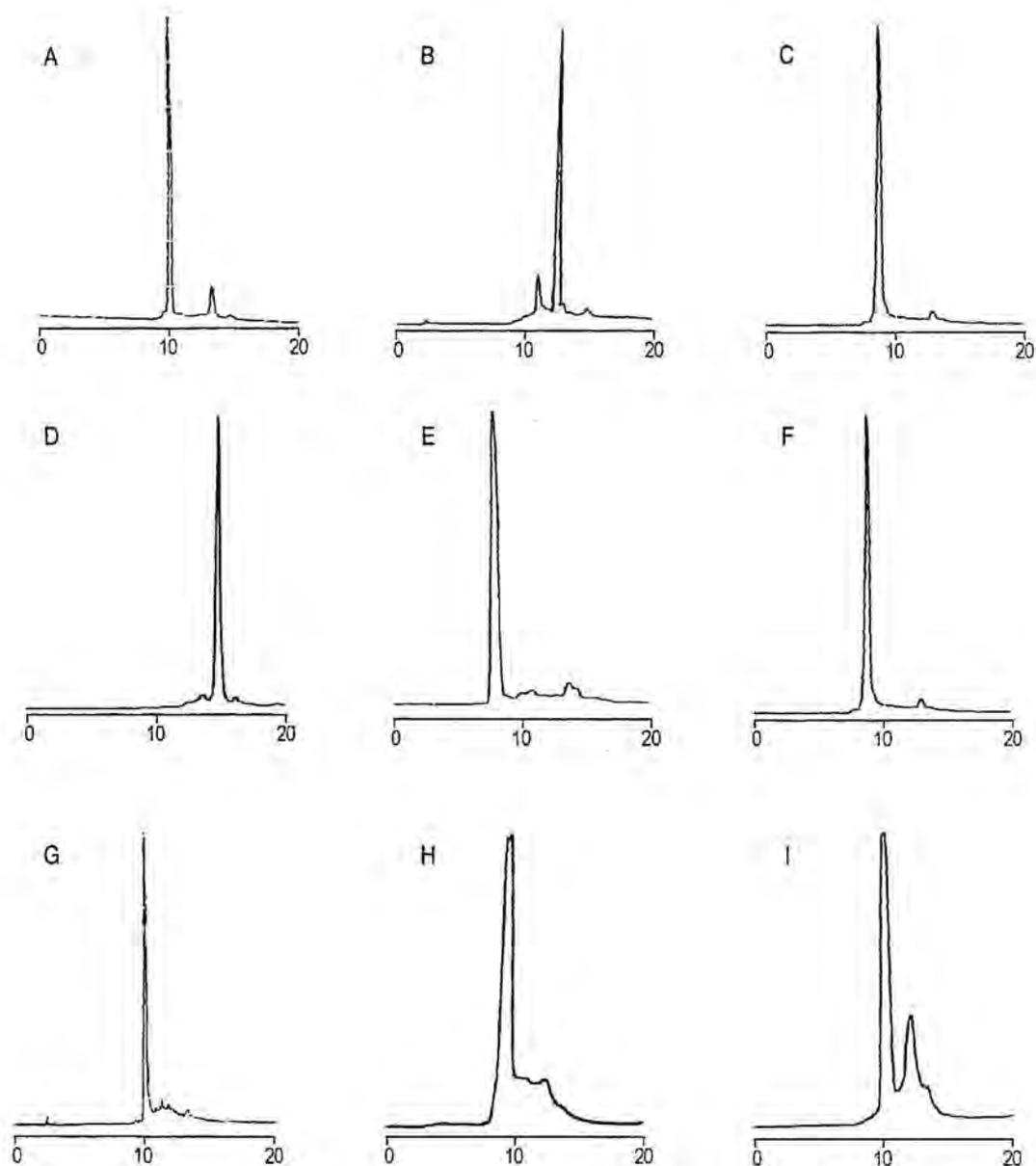
(1) Soporte utilizado: TentaGel con espaciador.

(2) Soporte utilizado: p-MBHA con espaciador.

A partir de los datos que figuran en la Tabla 1.15, se pueden concluir varias cosas. La más importante, remarcar que se mantienen rendimientos de obtención de crudo entre el 30-40% y que, por análisis HPLC, tal como se muestra en la Figura 1.26, se confirma que se trata de crudos de elevado grado de pureza. Puesto que se ha trabajado con resinas con muy distintos valores de funcionalización (43-118  $\mu\text{mol/g}$ ), estos resultados constatan, además, que no existe una correlación clara entre el grado de funcionalización de la resina y el rendimiento y pureza de crudo obtenidos.

Donde tampoco se observan diferencias importantes es en cuanto a qué reactivos se han utilizado para llevar a cabo la ciclación. Parece que la utilización conjunta de MSNT/NMI no supone ninguna mejora respecto a la utilización de únicamente MSNT. De hecho, los mejores rendimientos se han obtenido para las síntesis de d<pCATT<math>\text{CATT}> y d<pTGCT<math>\text{CGCT}>, en las que la reacción de ciclación se ha realizado utilizando como único reactivo de acoplamiento el MSNT.

Todos estos resultados demuestran la viabilidad de aplicar la estrategia de síntesis descrita a la obtención de oligonucleótidos cíclicos a gran escala. Sin embargo, resultaba necesario verificar si también la etapa de purificación permitía el escalado. La purificación se realizó, en todos los casos, mediante HPLC semipreparativo utilizando una columna de poli(estireno-divinilbenceno). En la Tabla 1.16 se muestra el rendimiento de recuperación de producto obtenido para cada una de las secuencias, así como el rendimiento global de síntesis.



**Figura 1.27.-** Cromatogramas HPLC de los crudos de síntesis correspondientes a cada una de las secuencias. A)  $d\langle p\text{CATT CATT}\rangle$ , B)  $d\langle p^{\text{Br}}\text{UCAT}^{\text{Br}}\text{UCAT}\rangle$ , C)  $d\langle p\text{CGCTCGCT}\rangle$ , D)  $d\langle p\text{CTATCTAT}\rangle$ , E)  $d\langle p\text{TGCTCGCT}\rangle$ , F)  $d\langle p\text{CCGTCCGT}\rangle$ , G)  $d\langle p\text{TGCATGCA}\rangle$ , H)  $d\langle p\text{AGCTAGCT}\rangle$  y I)  $d\langle p\text{CGCTCATT}\rangle$ . Gradiente: 10-40% B en 20 min.

**Tabla 1.16.-** Valores de rendimiento de recuperación de producto tras la purificación por HPLC semipreparativo y rendimiento global de síntesis.

Secuencia	Rdt. recuperación %	Rdt. global %
d<pCATT<sup>R</sup>CATT>	71	17
d<pCAT<sup>R</sup>UCAT<sup>R</sup>U>	41	8
d<pCGCTCGCT>	53	16
d<pCTATCTAT>	54	22
d<pTGCTCGCT>	48	22
d<pCCGTCCGT>	71	13
d<pTGCATGCA>	50	13
d<pAGCTAGCT>	53	12
d<pCGCTCATT>	70	16

En efecto, los valores muestran porcentajes elevados de recuperación de producto en todos los casos. Incluso cuando la purificación parece que no ha ido muy bien los valores son siempre superiores al 40%. Esta técnica se confirma, por tanto, como la apropiada para purificar este tipo de productos, ya que se obtienen rendimientos que, aunque en algún caso puedan resultar moderados, son muy superiores a los obtenidos por otros métodos de purificación.

La caracterización de los productos se ha llevado a cabo mediante espectrometría de masas ES<sup>-</sup> y digestión enzimática, Tala M.9 de Materiales y Métodos.

Como conclusión podemos entonces afirmar que el método descrito para la obtención de oligonucleótidos cíclicos, permite llevar a cabo síntesis a gran escala manteniéndose, en general, los buenos rendimientos de obtención de producto final así como la elevada pureza de los crudos.

Sobre los productos descritos en la Tabla 1.16 se han llevado a cabo estudios estructurales mediante difracción de RX y <sup>1</sup>H-RMN de campo alto, tal como se describe en los capítulos siguientes. Estas técnicas, si bien tienen como finalidad la determinación de la estructura tridimensional de las moléculas cíclicas, constituyen también un método adicional de caracterización de las distintas secuencias.

---

*CAPÍTULO 2*

*ANÁLISIS MEDIANTE RMN DE  $d$ -pCATTATT> Y  $d$ -pTGCTCGCT>.  
EQUILIBRIO MONÓMERO-DÍMERO.*

---

## 2.- ANÁLISIS MEDIANTE RMN DE d<pCATT<math>CATT</math>> Y d<pTGCT<math>CGCT</math>>. EQUILIBRIO MONÓMERO-DÍMERO.

### 2.0.- INTRODUCCIÓN Y OBJETIVOS.

Estudios cristalográficos del octámero cíclico d<pCATT<math>CATT</math>> evidenciaron la existencia de una estructura dimérica, tipo tétrada, por asociación intermolecular de dos moléculas de cíclico, tal como se describe en el Capítulo 3. Dada la singularidad de dicha estructura, nos planteamos el estudio en disolución, mediante RMN, de d<pCATT<math>CATT</math>> y una secuencia análoga con pares G-C, d<pTGCT<math>CGCT</math>>. El interés principal se centra en determinar la existencia en disolución de un equilibrio de asociación que pueda dar lugar a una especie dimérica. Sin embargo, en ambas secuencias resulta posible también la formación de dos pares de bases intramoleculares A-T y G-C, respectivamente, por lo que podrían dar lugar a estructuras tipo *dumbbell*.

Como se mencionó en la introducción general a esta Tesis, las estructuras tipo *dumbbell* son buenos análogos de *hairpins* de DNA, muy convenientes para su estudio en disolución. Puesto que una estructura *dumbbell* se puede considerar formada por dos *hairpin-loops* y un tallo de doble cadena de un número variable de residuos, su formación y estabilidad depende de forma importante del número y tipo de residuos que forman los *loops*. Los primeros estudios de *hairpins* de DNA indicaban que *loops* constituidos por cuatro o cinco nucleótidos eran los más estables, mientras que horquillas con únicamente dos residuos en el *loop* eran estéricamente imposibles<sup>1-10</sup>. Estudios más recientes revelaron, sin embargo, que *loops* constituidos por dos

<sup>1</sup> Hilbers, C.W.; Haasnoot, C.A.G.; de Bruin, S.H.; Joordens, J.J.M.; van der Marel, G.A.; van Boom, J.H. *Biochemie* **1985**, 67, 685-695.

<sup>2</sup> Wemmer, D.E.; Chou, S.H.; Hare, D.R.; Reid, B.R. *Nucleic Acids Res.* **1985**, 13, 3755-3772.

<sup>3</sup> Summers, M.F.; Byrd, R.A.; Gallo, K.A.; Samson, C.J.; Zon, G.; Eagen, W. *Nucleic Acids Res.* **1985**, 13, 6375-6386.

<sup>4</sup> Nadeau, J.G.; Gilham, P.T. *Nucleic Acids Res.* **1985**, 13, 8259-8274.

<sup>5</sup> Ikuta, S.; Chatopadhyaya, R.; Hirataka, I.; Dickerson, R.E.; Kearns, D.R. *Biochemistry* **1986**, 25, 4840-4849.

<sup>6</sup> Roy, S.; Weinstein, S.; Borah, B.; Nickol, J.; Apella, E.; Sussman, J.L.; Miller, M.; Sindo, H.; Cohen, J.S. *Biochemistry* **1986**, 25, 7417-7422.

<sup>7</sup> Haasnoot, C.A.G.; Hilbers, C.W.; van der Marel, G.A.; van Boom, J.H.; Singh, U.C.; Pattabiraman, N.; Kollman, P.A. *J. Biomol. Struct. Dynam.* **1986**, 3, 843-857.

<sup>8</sup> Williamson, J.R.; Boxer, S.J. *Biochemistry* **1989**, 28, 2819-2831.

<sup>9</sup> Williamson, J.R.; Boxer, S.J. *Biochemistry* **1989**, 28, 2831-2836.

<sup>10</sup> Germann, M.W.; Kalisch, B.W.; Lundberg, P.; Vogel, H.J.; van de Sande, J.H. *Nucleic Acids Res.* **1990**, 18, 1489-1498.

nucleótidos no sólo podían formar *mini-hairpin loops* muy estables<sup>11-14</sup>, sino que, además, secuencias que forman loops de dos bases *in vitro*, se forman también *in vivo*<sup>15</sup>.

Otro factor importante en la estabilidad de estructuras *dumbbell* es sin duda el tipo de residuos que constituyen el *loop*. De los estudios anteriormente mencionados se ha podido concluir que los loops de dos residuos más estables son aquellos formados por residuos pirimidínicos, en particular aquellos que se encuentran flanqueados por 5'-pirimidina-loop-purina-3'. Para las dos secuencias estudiadas se cumplen ambos requisitos.

Puesto que a lo largo del presente trabajo se ha utilizado de forma importante la RMN, se ha creído conveniente incluir una introducción metodológica de cómo dicha técnica se aplica a la elucidación estructural de fragmentos de DNA.

## 2.1.- LA RMN COMO HERRAMIENTA PARA ESTUDIAR ESTRUCTURAS DE ÁCIDOS NUCLEICOS.

El grado de detalle que proporciona la RMN a nivel estructural es comparable al que se obtiene mediante difracción de rayos X, con la ventaja de que la RMN da información de la estructura que la molécula adopta en disolución. Al no estar esta estructura sujeta a fuerzas de empaquetamiento en el cristal, como a veces puede ocurrir en RX, la RMN permite mimetizar condiciones nativas y resulta de gran importancia en el estudio de biomoléculas. La RMN proporciona además información dinámica. Esto resulta especialmente importante en el caso de ácidos nucleicos que suelen presentar regiones de elevada flexibilidad conformacional. Para entender la especificidad en el reconocimiento de moléculas de DNA y RNA hay que tener en cuenta los efectos de flexibilidad conformacional en la interfase. La RMN resulta por tanto una herramienta ideal para caracterizar procesos dinámicos que tienen una gran importancia en el reconocimiento molecular de ácidos nucleicos.

Los avances técnicos que ha experimentado la RMN en los últimos años<sup>16</sup> han hecho posible la determinación de estructuras de fragmentos de DNA y RNA de mayor tamaño y complejidad. En este sentido hay que destacar los avances en el marcaje isotópico tanto de moléculas de RNA<sup>17-18</sup> como de DNA<sup>19</sup>. El contar con muestras uniformemente marcadas en <sup>15</sup>N y <sup>13</sup>C permite la utilización de experimentos de RMN heteronuclear, que facilitan el proceso de asignación. No obstante, para moléculas de pequeño tamaño, es posible determinar la estructura en disolución utilizando experimentos convencionales y moléculas no marcadas, tal como se detalla a continuación.

<sup>11</sup> Orbons, L.P.M.; van der Marel, G.A.; van Boom, J.H.; Altona, C. *Nucleic Acids Res.* **1986**, *14*, 4187-4195.

<sup>12</sup> Orbons, L.P.M.; van Beuzekon, A.A.; Altona, C. *J. Biomol. Struct. Dynam.* **1987**, *4*, 965-987.

<sup>13</sup> Pieters, J.M.L.; de Vroom, E.; van der Marel, G.A.; van Boom, J.H.; Koning, T.M.G.; Kaptien, R.; Altona, C. *Biochemistry* **1990**, *29*, 788-799.

<sup>14</sup> Ippel, J.H.; Lanzotti, V.; Galeone, A.; Mayol, L.; van den Boogaart, J.E.; Pikkemaat, J.A.; Altona, C. *J. Biomol. Struct. Dynam.* **1991**, *9*, 821-836.

<sup>15</sup> Davison, A.; Leach, D.R.F. *Nucleic Acids Res.* **1994**, *22*, 4361-4363.

<sup>16</sup> Varani, G.; Aboul-ella, F.; Allain, F.H.-T. *Progress NMR Spectr.* **1996**, *29*, 51-127.

<sup>17</sup> Nikonowicz, E.P.; Pardi, A. *Nature* **1992**, *355*, 184-186.

<sup>18</sup> Batey, R.T.; Battiste, J.L.; Williamson, J.R. *Meth. Enz.* **1995**, *261*, 300-322.

<sup>19</sup> Zimmer, D.P.; Crothers, D.M. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3091-3095.

### 2.1.1.- Adquisición de datos experimentales.

En los espectros monodimensionales de RMN de macromoléculas existe generalmente un alto grado de solapamiento entre las distintas señales, debido al gran número de protones presentes. El análisis de los espectros resulta complicado y, para poder asignar por completo todas las resonancias de la molécula, se hace necesario recurrir a la RMN bidimensional.

La secuencia general de pulsos para un espectro bidimensional es la que se muestra de forma esquemática en la Figura 2.1. La variación de la señal durante el tiempo de adquisición ( $t_2$ ) es lo que se conoce como FID y su transformada de Fourier da lugar a una señal de frecuencia característica. Si se realizan numerosos experimentos del mismo tipo variando para cada uno de ellos el tiempo de evolución ( $t_1$ ), la variación de la magnetización con el tiempo  $t_1$ , para cada valor de tiempo  $t_2$ , origina una curva similar a la FID sobre la que también se puede realizar una transformada de Fourier. El resultado de un experimento de estas características consiste en una matriz de datos donde las filas contienen información variable con  $t_2$  a un valor fijo de  $t_1$ , y las columnas información variable con  $t_1$  a un valor fijo de  $t_2$ . La transformada de Fourier en ambos dominios de tiempo da lugar a un espectro bidimensional, donde resonancias a frecuencias en una dimensión se relacionan con resonancias a frecuencias en la otra dimensión. Dependiendo del tipo concreto de experimento bidimensional, puede eliminarse parcialmente el solapamiento y además se puede obtener multitud de información que dependerá del fenómeno que relacione los valores de frecuencias en ambas dimensiones.

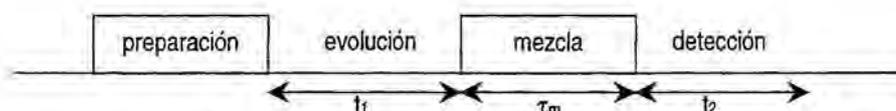


Figura 2.1.- Secuencia general de pulsos para la obtención de un espectro bidimensional.

La secuencia de pulsos consta, en primer lugar, de un periodo de preparación que permite al sistema llegar al equilibrio y convertir la magnetización longitudinal en magnetización transversal. Posteriormente hay un periodo de evolución, en el que se registra información sobre una frecuencia que evoluciona durante  $t_1$ . Durante el tiempo de mezcla, el sistema parte de un estado de no equilibrio y se produce una transferencia de magnetización del núcleo cuya frecuencia evolucionó con  $t_1$  a un núcleo con el que esté conectado, escalar o dipolarmente, y cuya frecuencia evolucionará durante el tiempo de detección  $t_2$ . En función de como sea este periodo de mezcla se obtendrán los diferentes tipos de experimentos bidimensionales. El resultado será un pico de correlación en el espectro bidimensional entre dos protones de distinta frecuencia.

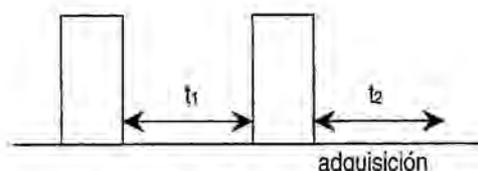
Los diferentes experimentos bidimensionales utilizados en el presente trabajo son los que se describen a continuación.

#### 2.1.1.1.- Experimento COSY (CORrelated Spectroscopy)<sup>20</sup>.

El experimento COSY permite realizar una transferencia de magnetización entre protones que se encuentran acoplados escalarmente. Para que la correlación sea efectiva ha de

<sup>20</sup> Aue, W.P.; Bartholdi, E.; Ernst, R.R. *J. Chem. Phys* **1975**, *64*, 2229-2246.

tratarse de constantes de acoplamiento a dos y tres enlaces. La secuencia de pulsos, que consiste en dos pulsos no selectivos de radiofrecuencia separados por un tiempo variable  $t_1$ , está representada en la Figura 2.2. La adquisición se realiza después del segundo pulso durante el período  $t_2$ .

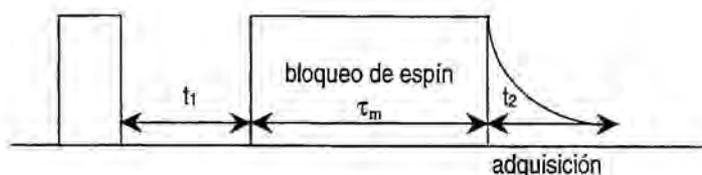


**Figura 2.2.-** Secuencia de pulsos del experimento COSY. Los rectángulos simbolizan pulsos de  $90^\circ$ .

En el espectro obtenido, las señales de correlación están en modo de absorción y las de la diagonal en dispersión, pero todas ellas son señales en antifase. Esta forma antifase de las señales provoca una cancelación parcial del centro del multiplete, que incrementa la separación de las componentes dando lugar a una constante de acoplamiento aparentemente mayor de lo que en realidad es. Para solucionar este problema existen variantes del experimento COSY diseñados para determinar cuantitativamente valores de la constante de acoplamiento (DQF-COSY, E. COSY)<sup>21</sup>.

### 2.1.1.2.- Experimento TOCSY (Total Correlation Spectroscopy)<sup>22,23</sup>.

El experimento TOCSY, también denominado HOHAHA (Homonuclear Hartmann-Hahn) se basa en la condición de Hartmann-Hahn que establece que existe transferencia de magnetización entre dos espines acoplados escalarmemente si las frecuencias de precesión efectivas de ambos son iguales. En sentido estricto, esta condición es muy difícil de conseguir porque para que esto ocurra los espines acoplados escalarmemente deben situarse equidistantes del transmisor. No obstante, es posible transferir magnetización entre todos los espines escalarmemente acoplados aplicando secuencias de bloqueo de radiofrecuencia adecuada, capaces de eliminar los términos de desplazamiento químico en anchuras de varios KHz<sup>24</sup>. En la Figura 2.3 se representa la secuencia de pulsos de un experimento TOCSY.



**Figura 2.3.-** Secuencia de pulsos de un experimento TOCSY. El rectángulo pequeño simboliza un pulso de  $90^\circ$  y el grande un tren de pulsos para conseguir bloquear la magnetización.

<sup>21</sup>Bruix, M. *Curso Avanzado de Resonancia Magnética Nuclear* 1997, 179-222. Escuela de Verano. Universidad de Zaragoza. Sociedad Española de Óptica.

<sup>22</sup>Brüschweiler, L.; Ernst, R. R. *J. Magn. Reson.* 1983, 53, 521.

<sup>23</sup>Bax, A.; Davies, D. G. *J. Magn. Reson.* 1985, 65, 355-360.

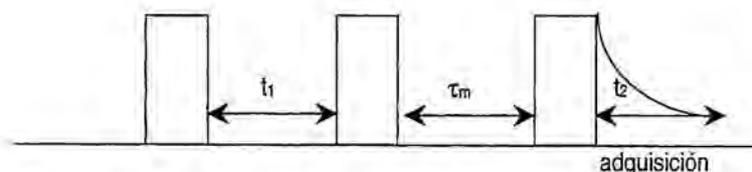
<sup>24</sup>Jiménez-Barbero, J. *Curso Avanzado de Resonancia Magnética Nuclear* 1993, 171-195. Escuela de Verano. Universidad de Zaragoza. Sociedad Española de Óptica.

En un espectro TOCSY es posible correlacionar un protón con todos los pertenecientes al mismo sistema de espín. Tanto las señales de la diagonal como los picos de correlación son señales de absorción y se encuentran en fase. Por contra, este tipo de experimento no permite distinguir entre correlaciones directas e indirectas y su análisis debe realizarse conjuntamente con el del espectro COSY.

### 2.1.1.3.- Experimento NOESY (Nuclear Overhauser Effect Spectroscopy)<sup>25,26</sup>.

El efecto nuclear Overhauser (NOE) es un aspecto de la relajación nuclear que se traduce en un cambio de la intensidad de una señal de resonancia cuando las transiciones de otra han sido perturbadas de alguna manera. El efecto NOE depende de la interacción entre dipolos magnéticos y, al ser esta interacción tanto más efectiva cuanto menor es la distancia entre ellos, la distancia es un factor importante del que depende la velocidad de crecimiento del NOE<sup>27</sup>. Si se utilizan tiempos de mezcla cortos es posible aplicar una aproximación a partir de la cual se puede relacionar la intensidad del NOE con la inversa de la sexta potencia de la distancia, esta fuerte dependencia con la distancia hace que sólo se observen NOEs entre protones cuya separación sea inferior a 5Å.

En la práctica, la generación del efecto NOE se lleva a cabo en la versión bidimensional del NOE transitorio<sup>27</sup>. La secuencia de pulsos del experimento NOESY es la que se muestra en la Figura 2.4.



**Figura 2.4.-** Secuencia de pulsos del experimento NOESY. Los rectángulos simbolizan pulsos de 90°.

El pulso inicial crea magnetización transversal que evoluciona libremente durante  $t_1$ . El segundo pulso convierte parte de esta magnetización en magnetización longitudinal de no equilibrio. Durante el tiempo de mezcla se trata de recuperar la situación de equilibrio mediante relajación dipolar, dándose transferencia de magnetización entre espines cercanos en el espacio, es decir, acoplados dipolarmente. El último pulso transforma la magnetización transferida al plano  $xy$  que se detecta durante  $t_2$ . Una vez realizada la doble transformada de Fourier se obtiene un espectro bidimensional, donde frecuencias de distintos núcleos cercanos en el espacio están correlacionadas.

Las señales de un espectro NOESY se encuentran en forma de absorción pura y en fase. Su intensidad está relacionada con la duración del tiempo de mezcla, la velocidad del proceso de relajación y el tiempo de correlación molecular ( $\tau_c$ ).

<sup>25</sup>Macura, S.; Ernst, R. R. *Molec. Phys.* **1980**, 41, 95.

<sup>26</sup>Kumar, A.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1980**, 95, 1-6.

<sup>27</sup>Neuhaus, D.; Williamson, M. «The Nuclear Overhauser Effect in Structural and Conformational Analysis» **1989**, VCH Publishers, Inc., New York.

El experimento NOESY presenta una serie de características que, en ocasiones, pueden representar un problema para el análisis de los espectros:

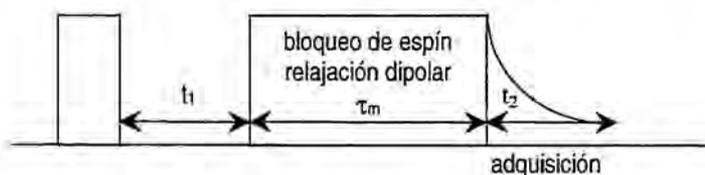
- La transferencia de magnetización puede darse a través de dos mecanismos: relajación cruzada o intercambio químico. Cuando ambos fenómenos son posibles, un experimento NOESY no permite distinguir el origen de un pico de correlación.

- Pueden aparecer picos de cruce por difusión de espín. Este término hace referencia a correlaciones entre dos núcleos que pueden no encontrarse próximos en el espacio, aunque sí cerca de un tercero. Considerando tres núcleos A, B y C, puede aparecer un pico de cruce entre A y C debido a la relajación cruzada que experimenta B con C, seguida ésta de la relajación de B con A.

En ambos casos, la solución es utilizar un experimento CAMELSPIN (Cross-relaxation Appropriate for Mini-molecules Emulated by Locked SPINs), también denominado ROESY.

#### 2.1.1.4.- Experimento ROESY (Rotating frame nuclear Overhauser Effect Spectroscopy)<sup>28</sup>.

En este experimento el efecto NOE se da entre magnetizaciones transversales bloqueadas durante el periodo de mezcla. La secuencia de pulsos se representa en la Figura 2.5.



**Figura 2.5.-** Secuencia de pulsos del experimento ROESY. El rectángulo pequeño corresponde a un pulso de  $90^\circ$  y el grande a un tren de pulsos para bloquear la magnetización en el plano.

En este caso, tras el periodo  $t_1$  se aplica un campo de radiofrecuencia de bloqueo de espín. Los espines, dependiendo de la precesión sufrida durante el período de evolución, presentan distintas cantidades de magnetización alineadas con el campo de radiofrecuencia. En estas condiciones puede tener lugar transferencia de magnetización entre espines espacialmente próximos, lo que generará efectos NOE (ROE) en dirección perpendicular al campo magnético externo.

El experimento ROESY permite diferenciar, por tanto, los NOEs directos de los indirectos (difusión de espín) porque son de signo opuesto. En casos donde pueda existir también equilibrio químico, los NOEs debidos a correlación dipolar son de signo contrario a los debidos a intercambio. Por este motivo el experimento ROESY resulta muy útil tanto para discriminar picos debidos a difusión de espín como para analizar espectros en los que coexisten varias especies en equilibrio.

<sup>28</sup>Bax, A.; Davis, D. J. *J. Magn. Reson.* **1985**, *63*, 207.

## 2.1.2.- Asignación del espectro.

### 2.1.2.1.- Desplazamientos típicos de protón en DNA.

Para realizar un estudio conformacional por RMN, el primer paso consiste en la asignación de las resonancias de la molécula. Este proceso resulta complicado en el caso de ácidos nucleicos debido a que la existencia de únicamente cuatro nucleótidos implica numerosas repeticiones de residuos en la secuencia, dando lugar a solapamiento. En la Tabla 2.1 se muestran los intervalos de desplazamiento químico habituales para los distintos protones que constituyen los nucleótidos.

Tabla 2.1.- Rango de desplazamiento químico de  $^1\text{H}$  en DNA, expresados en ppm<sup>29</sup>.

Protón	Timidina	Citidina	Adenosina	Guanosina
H1'	5-6	5-6	5-6	5-6
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'	4.45.0	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>2</sub> /H <sub>2</sub> /Me	1.0-1.9	5.0-6.0	7.7-8.5	-
NH <sup>a</sup>	13-14	-	-	12-13.6
NH <sup>b</sup>	10-11	-	-	10-11
NH <sub>2</sub> <sup>a</sup>	-	6.5-8.5	7-8	8-9
NH <sub>2</sub> <sup>b</sup>	-	6.7-7.0	5-6	5-6

<sup>a</sup>protón imino o amino formando puente de hidrógeno en pares de bases tipo Watson-Crick

<sup>b</sup>protón imino o amino desapareado

Tal como se puede observar en la Tabla 2.1, las resonancias correspondientes a protones imino se desplazan hacia campos más bajos cuando se forma puente de hidrógeno. En la tabla aparecen los desplazamientos típicos en el caso de que se formen pares de bases Watson-Crick, pero pueden existir además otros casos de apareamiento, como es la formación de un par de bases TT o GT (wobble). En estos casos el protón imino forma enlace de hidrógeno con el oxígeno de un grupo carbonilo y la señal aparece entre 10 y 12 ppm. Las resonancias que aparecen a campo más bajo, alrededor de 15 ppm, son las correspondientes a los protones imino de citosinas protonadas, formando pares de bases tipo Hoogsteen. El comportamiento de los protones amino en este tipo de pares de bases es similar, desplazándose hacia valores de desplazamiento químico entre 9.5-10ppm.

Las resonancias correspondientes a los diferentes tipos de protones no intercambiables están razonablemente bien separadas en la escala de desplazamiento químico, a excepción de los protones H4', H5' y H5'', cuya asignación resulta muy complicada salvo en moléculas de pequeño tamaño.

### 2.1.2.2.- Generalidades y limitaciones del método de asignación.

La estrategia para la asignación secuencial de las resonancias de protón en ácidos nucleicos está basada en la combinación de experimentos bidimensionales de correlación escalar, para la identificación de los diferentes sistemas de espín, y de experimentos de correlación dipolar, que permiten asignar los diferentes sistemas de espín a residuos específicos de la cadena de ácido nucleico.

El protocolo de asignación<sup>29,30</sup>, que se utiliza de forma general, presenta una serie de limitaciones que hay que tener en cuenta. Existen dos problemas principales. El primero de ellos es el solapamiento de señales, que es especialmente importante cuando se utilizan únicamente experimentos bidimensionales de correlación homonuclear de protón, ya que puede resultar imposible completar los diferentes sistemas de espín. El segundo problema está asociado al hecho de que los NOEs clave para la asignación solamente se observan en hélices de DNA a derechas (A- y B-DNA). En el caso de estructuras irregulares los contactos secuenciales pueden ser muy diferentes y el protocolo de asignación descrito puede conducir a resultados erróneos.

### 2.1.2.3.- Asignación de protones no intercambiabiles.

La asignación completa de los protones no intercambiabiles consta de dos etapas. En primer lugar se lleva a cabo la identificación de los diferentes sistemas de espín y, posteriormente se realiza la asignación secuencial.

#### 2.1.2.3.1.- Identificación de los sistemas de espín.

La identificación de los distintos sistemas de espín se realiza en base a los picos de correlación escalar de espectros COSY y TOCSY. En estos espectros resulta fácil identificar las señales que corresponden a protones de una misma ribosa ya que se encuentran acoplados escalarmenete a menos de tres enlaces y cada ribosa constituye un sistema de espín aislado del resto de la molécula. Para identificar cada sistema de espín basta con seguir en estos espectros los picos de correlación H1'-H2'/H2'', H2'/H2''-H3' y así sucesivamente hasta completar todos los protones de las ribosas.

En cuanto a las bases, en experimentos de correlación escalar sólo es posible identificar los sistemas de espín correspondientes a citosina y timina. Para la citosina se observa el pico de correlación escalar entre los protones H5 y H6. En el caso de la timina se observa el pico de correlación entre el grupo metilo en la posición 5 y el protón H6, aunque al tratarse en este caso de un contacto a cuatro enlaces la intensidad del pico de correlación es menor.

#### 2.1.2.3.2.- Asignación secuencial.

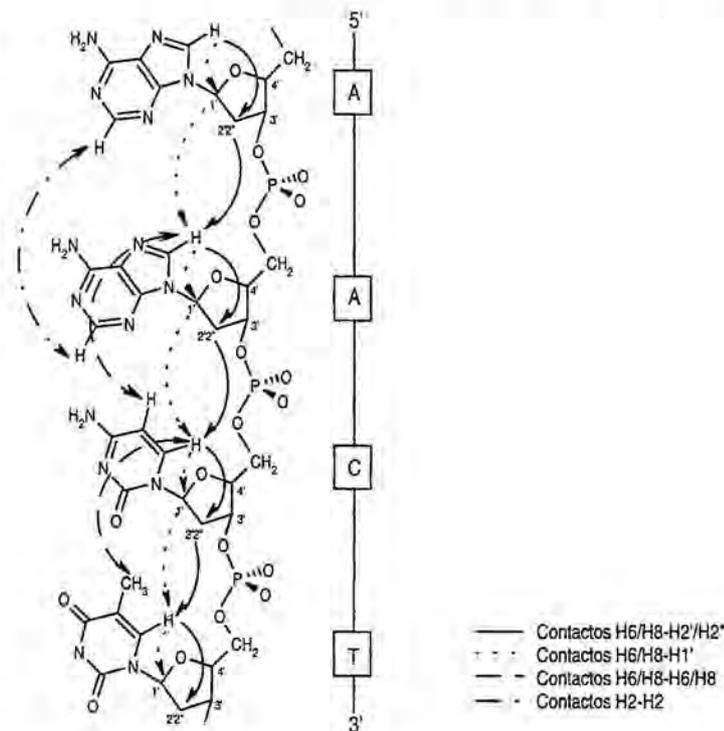
La asignación secuencial de los distintos sistemas de espín se realiza a partir de picos de correlación dipolar. Para confirmar que, efectivamente, los picos observados en el espectro NOESY corresponden a picos de correlación dipolar, y no tienen su origen en procesos de intercambio o difusión de espín, es conveniente analizar conjuntamente el espectro ROESY.

Para la asignación secuencial se utilizan picos NOE que son característicos de hélices dextrógiras canónicas, tal como se muestra en la Figura 2.6. En este tipo de hélices, el protón

<sup>29</sup>Wijmenga, S. S.; Mooren, W.; Hilbers, C. W. "NMR of Macromolecules, A Practical Approach" 1993, Oxford University Press, New York, 217-288.

<sup>30</sup>Wijmenga, S. S.; van Buuren, B. N. M. *Progress in Nuclear Magnetic Resonance Spectroscopy* 1998, 32, 287-387

H6/H8 de un residuo presenta un pico NOE con los protones H1' y H2'/H2'' de su propio azúcar y con los protones H1' y H2'/H2'' de la ribosa del residuo anterior, en la dirección 5'-3'.



**Figura 2.6.-** Fragmento de DNA donde se muestran algunos contactos inter-residuo en la dirección 5'-3' que permiten realizar la asignación secuencial.

A continuación se detalla el protocolo de asignación:

1.- Elección de un pico de correlación inequívoco. En cadenas lineales de DNA, la señal correspondiente al protón H6/H8 del residuo 5'-terminal presenta pico de correlación con un único H1', el de su propio azúcar, y es por tanto un buen punto de inicio. También resulta apropiado iniciar la asignación a partir de la señal del protón aromático H6 de un residuo que se haya identificado como timina o citosina a partir del espectro COSY o TOCSY.

2.- Búsqueda de un pico de correlación que conecte la señal correspondiente al protón H1' con el protón H6/H8 del siguiente residuo en la secuencia, en el sentido 5'-3'.

3.- Desde la posición del protón H6/H8 del siguiente residuo buscar, en la dirección de  $\omega_1$ , la frecuencia de resonancia correspondiente a su H1'.

4.- Continuar con este procedimiento, buscando el pico de correlación entre el protón H1' con el protón H6/H8 del siguiente residuo, y así sucesivamente.

5.- Si por algún motivo se llega a un punto de discontinuidad, bien por la existencia de solapamiento, o bien porque se llegue a un punto de irregularidad estructural donde se pierdan los contactos secuenciales, se deberá elegir un nuevo punto de inicio.

De forma análoga al análisis de los picos de correlación H6/H8-H1', se puede seguir la asignación secuencial en base a los picos de correlación H6/H8-H2'/H2''. La combinación de ambos análisis debe conducir a un resultado inequívoco.

- Asignación estereoespecífica de los protones H2'/H2'' y H5'/H5''.

Las señales correspondientes a los protones H2' y H2'' se pueden diferenciar en base a la intensidad del NOE con H1'. El pico de correlación H1'-H2'' es más intenso que el correspondiente a H1'-H2' a lo largo de todo el ciclo de pseudo-rotación del azúcar (Anexo A). Puesto que estas distancias pueden verse afectadas por efectos de difusión de espín, es muy conveniente usar espectros NOESYs con tiempo de mezcla corto para realizar estas asignaciones. En el apartado 2.1.3.1.2 se muestra como varían las distintas distancias interprotónicas con el ángulo de pseudorotación dentro de una unidad de desoxiribosa.

La asignación estereoespecífica de H5' y H5'' está relacionada con la determinación del ángulo de torsión  $\gamma$ . Se puede conseguir combinando  $J_{4'5'}$  y  $J_{4'5''}$  con las intensidades de los NOEs H3'-H5'/H5'', H4'-H5'/H5''. Si  $J_{4'5'}$  y  $J_{4'5''}$  son pequeñas (2-3 Hz) y se asume que el ángulo de torsión  $\gamma$  se encuentra en la conformación  $g^+$ , que es la más habitual, ( $d_i(3', 5'') \sim 2.4 \text{ \AA}$ ,  $d_i(3', 5') \sim 3.8 \text{ \AA}$ ), entonces el NOE H3'-H5'' será mucho más intenso que el NOE H3'-H5'. En caso de que el ángulo de torsión se encuentre en las conformaciones  $\gamma'$  o  $\gamma$  el valor de  $J_{4'5'}$  y  $J_{4'5''}$  es mucho mayor:

$\gamma'$  ( $d_i(3', 5'') \sim d_i(3', 5') \sim 2.6 \text{ \AA}$ ): el NOE es intenso en ambos casos, pero  $J_{4'5''} > J_{4'5'}$ .

$\gamma$ : el NOE más intenso corresponde al pico de correlación H3'-H5'.

Debido al fuerte solapamiento de señales en esta zona del espectro la asignación estereoespecífica de H5' y H5'' sólo es posible en oligonucleótidos de pequeño tamaño.

#### 2.1.2.4.- Asignación de protones intercambiables.

La asignación de protones intercambiables se realiza en base a los NOEs intercadena imino-imino y amino-imino que se observan cuando se forman los enlaces de hidrógeno entre bases complementarias. En la Figura 2.7 se pueden observar dichos contactos, así como otros NOEs del tipo protón intercambiable-protón no intercambiable (propio o secuencial), que son característicos para cada par de bases.

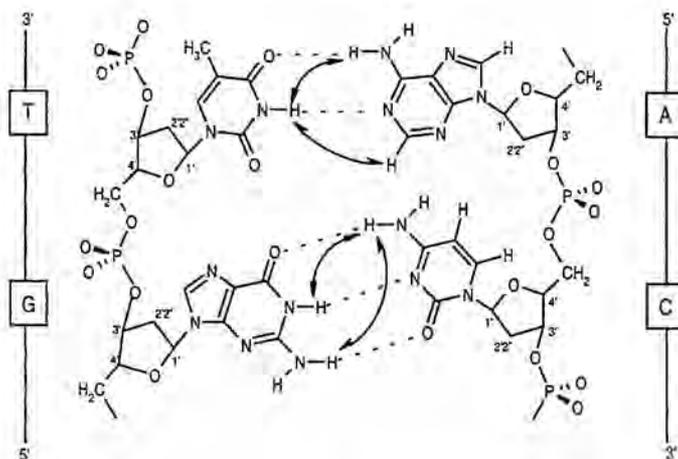


Figura 2.7.- Contactos entre protones intercambiables y no intercambiables en la formación de pares de bases Watson-Crick.

Como ocurre para los protones no intercambiables, también en este caso la asignación sólo es fiable cuando se trata de estructuras donde el apareamiento entre las bases se encuentra bien definido.

### 2.1.3.- Información estructural.

Una vez se ha conseguido asignar totalmente la molécula, el siguiente paso consiste en extraer de los espectros bidimensionales la información estructural necesaria para determinar la estructura tridimensional en disolución. La información más rica proviene de restricciones de distancias interprotónicas derivadas de los picos de correlación dipolar (NOE), aunque esta información se complementa con restricciones angulares derivadas del análisis de las constantes de acoplamiento escalar a tres enlaces, homo- y heteronucleares, obtenidas a partir de experimentos tipo COSY<sup>21</sup>.

La nomenclatura utilizada, así como la definición de los principales parámetros estructurales del DNA se muestran en el Anexo A.

#### 2.1.3.1.- Distancias.

##### 2.1.3.1.1.- Obtención de distancias.

La intensidad del NOE depende de la distancia entre protones y del grado de movilidad de la molécula. Si se supone que la molécula se mueve lentamente como un sólido rígido y que cada par de protones se encuentra aislado del resto, lo que se conoce como *aproximación de dos espines*, la intensidad del NOE es proporcional al tiempo de mezcla del experimento NOESY y decrece con la sexta potencia de la distancia interprotónica, tal como se indica en la expresión:

$$NOE \propto \frac{\tau_m}{r^6} \quad (1)$$

Para calcular la constante de proporcionalidad se utiliza la intensidad de NOEs entre protones cuya distancia es independiente de la conformación, como es el caso de protones geminales o aromáticos. Esta expresión presenta varias limitaciones, en primer lugar sólo es válida para tiempos de mezcla cortos, para los que se pueda considerar despreciable el efecto de difusión de espín. En segundo lugar, la suposición de movimiento isotrópico, en el que la molécula se mueve como un sólido rígido, con un único tiempo de correlación global, no es exacta en el caso de macromoléculas, ya que generalmente presentan un comportamiento dinámico complejo donde tienen lugar movimientos internos en varias escalas de tiempo. En el caso de que estos movimientos internos sean rápidos en comparación con el movimiento global de la molécula, se puede utilizar el formalismo de Lipari y Szabo para correlacionar el efecto NOE con la geometría molecular y un parámetro, llamado parámetro de orden, que contiene la información dinámica local<sup>31</sup>. Un caso especial de movimiento interno rápido es la rotación del los tres protones de los grupos metilo. En este caso particular, se puede llegar a una expresión similar a (1), pero donde la intensidad del NOE es proporcional a la inversa de la tercera potencia de la distancia<sup>27</sup>:

$$NOE \propto \frac{\tau_m}{r^3} \quad (2)$$

En la práctica resulta difícil establecer la validez de las expresiones (1) y (2) para cada caso concreto, así que en muchos casos se tiende a hacer una interpretación cualitativa de la intensidad del NOE. En general, se suele clasificar la intensidad del NOE en fuerte, medio y débil, a la que corresponde una restricción de distancia de 1.8-3.0, 1.8-4.0 y 1.8-5.0 Å respectivamente. Esta clasificación, aunque arbitraria, resulta útil en el caso de proteínas y otras moléculas globulares, donde la imprecisión se compensa con el gran número de NOEs que se suelen observar y, especialmente, con la presencia de los llamados NOEs de larga distancia que restringen mucho las posibles conformaciones de la molécula entre los dos residuos involucrados.

En el caso de ácidos nucleicos no se cuenta con NOEs a larga distancia y se dispone únicamente de NOEs secuenciales, por este motivo resulta imprescindible el uso de restricciones precisas para definir la estructura. La obtención de distancias precisas se consigue aplicando la teoría de la *matriz completa de relajación*<sup>32</sup>, donde las intensidades de los NOEs se representan mediante una matriz A, cuyos elementos  $a_{ij}$  son las intensidades de los NOEs entre los protones i y j. Estas intensidades se relacionan con la geometría molecular mediante la expresión (3):

$$A = \exp[-R\tau_m] \quad (3)$$

donde  $\tau_m$  es el tiempo de mezcla del experimento y R representa la matriz relajación cuyos elementos  $\rho_{ij}$  están relacionados con la geometría de la molécula a través de las expresiones (4) y (5):

$$\rho_{ij} = k \sum_{\substack{j=1 \\ j \neq i}}^N \left( \frac{1}{r_{ij}^6} \right) [6J_2(\omega) + 3J_1(\omega) + J_0(\omega)] \quad (4)$$

$$\sigma_{ij} = k \left( \frac{1}{r_{ij}^6} \right) [6J_2(\omega) - J_0(\omega)] \quad (5)$$

donde  $r_{ij}$  representa la distancia interprotónica y  $J(\omega)$  son las denominadas densidades espectrales, que dependen de la dinámica de la molécula. En el caso de movimiento isotrópico lento,  $J(\omega)$  se expresa como:

$$J_n(\omega) = \frac{\tau_c}{1 + n^2\omega^2\tau_c^2} \quad (6)$$

Al tratarse A de una matriz simétrica, los métodos de álgebra matricial permitirían invertir la expresión (3) y, a partir de los elementos de la matriz relajación, obtener las distancias entre protones. En la práctica y debido principalmente a problemas de solapamiento, no se puede disponer de la matriz experimental A completa y no se puede, por tanto diagonalizar e invertir. El procedimiento más usado para resolver este problema fue

<sup>31</sup>Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, 104, 4546-4559.

<sup>32</sup>James, T.L. *Curr. Opin. Struct. Biol.* **1991**, 1, 1042.

desarrollado simultáneamente por varios grupos<sup>33-37</sup> y consiste en calcular una matriz de NOEs teórica a partir de un modelo de molécula. En esta matriz se sustituyen aquellos elementos para los que se conocen valores experimentales de la intensidad. Generalmente, esta matriz mixta puede invertirse y dar lugar a una matriz de relajación a partir de la cual se pueden obtener distancias entre protones mediante la expresión (5). Las distancias entre protones para los que se ha incluido información experimental se emplean como restricciones para calcular una nueva estructura, que se puede utilizar nuevamente como modelo inicial para obtener la matriz de relajación y repetir el proceso de forma iterativa.

El uso de la *matriz completa de relajación* supone una mejora considerable para el cálculo de distancias, aunque los valores finales no están exentos de errores. Estos errores provienen, por una parte, del propio error experimental en la determinación de las intensidades de los NOEs y, por otra, de las hipótesis utilizadas en el cálculo de la matriz de NOEs teórica: coordenadas iniciales y tiempo de correlación global aproximado. La manera más apropiada de tener en cuenta estos errores es repetir los cálculos bajo diferentes condiciones y hacer un análisis estadístico riguroso de las distancias resultantes. En general, es conveniente realizar cálculos con intensidades provenientes de experimentos a diferentes tiempos de mezcla y utilizando varios modelos iniciales. También es conveniente repetir los cálculos con varios  $\tau_c$ , dentro de un rango amplio, para prevenir posibles efectos de movilidad local. Al igual que al hacer una interpretación cualitativa del NOE, o utilizar la aproximación de dos espines, la restricción final que se obtiene no es una distancia determinada entre dos protones, sino un rango de distancias con una cota inferior y superior, correspondientes a los errores en la determinación de dicha distancia. Este rango resulta, no obstante, más estrecho y las estructuras resultantes estarán mejor definidas.

En el caso de protones intercambiables, la intensidad del NOE queda afectada por la tasa de intercambio con el solvente. Puesto que el efecto del intercambio es siempre disminuir la intensidad del NOE, las distancias resultantes de un cálculo de matriz completa de relajación son siempre mayores que las distancias reales. Aunque es posible obtener valores cuantitativos de las tasas de intercambio con el solvente e introducirlos en el cálculo, la aproximación de considerar las distancias entre protones intercambiables como cotas superiores es generalmente suficiente.

### 2.1.3.1.2.- Distancias relevantes.

No todas las distancias interprotónicas tienen la misma relevancia a la hora de determinar la estructura tridimensional de la molécula. Muchas distancias intranucleotídicas apenas varían con la conformación y no resultan útiles para determinar la estructura. En la Figura 2.8 se muestra la variación de la distancia entre los diferentes protones de una desoxiribosa en función del ángulo de pseudorotación  $P$ . Se puede ver claramente que las distancias  $d_i(2''; 4')$  y  $d_i(1'; 4')$  son las que van a determinar la conformación del azúcar.

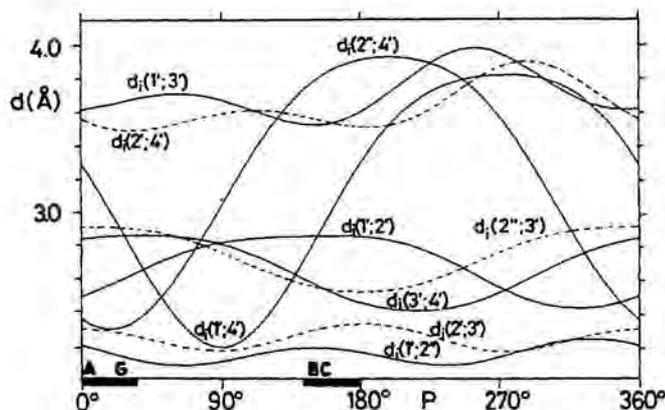
<sup>33</sup> Boelens, R.; Koning, T. M. G.; van der Marel, G. A.; van Boom, J. H.; Kaptein, R. *J. Magn. Res.* **1989**, *82*, 290-308.

<sup>34</sup> Koning, T. M. G.; Boelens, R.; Kaptein, R. *J. Magn. Res.* **1990**, *90*, 111-123.

<sup>35</sup> Olejniczak, E. T. *J. Magn. Res.* **1989**, *81*, 392-394.

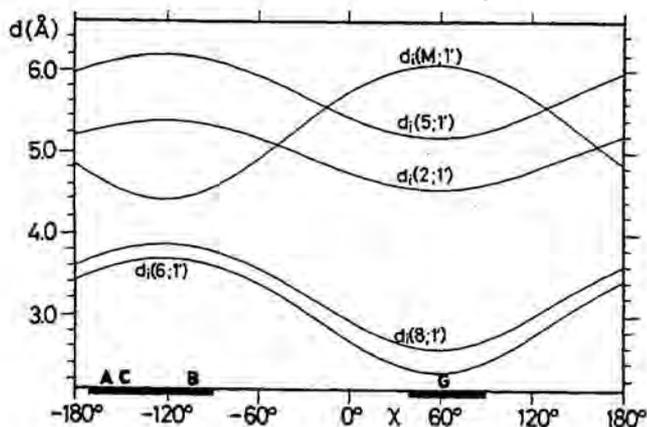
<sup>36</sup> Borgias, B. A.; James, T. L. *J. Magn. Res.* **1988**, *79*, 493-512.

<sup>37</sup> Borgias, B. A.; James, T. L. *J. Magn. Res.* **1990**, *87*, 475-487.



**Figura 2.8.-** Variación de las distancias intranucleotídicas entre diferentes protones del anillo de desoxiribosa en función del ángulo de pseudorotación  $P$ . Sobre el eje se ha indicado el ángulo de pseudorotación preferido para las formas A- y B-DNA, así como para los nucleótidos C y G en la forma Z-DNA en el oligonucleótido  $d(\text{CGCGCG})_2$ .

Las distancias  $d_i(1'; 6/8)$  resultan cruciales para determinar el ángulo glicosídico y, por tanto, poder establecer la disposición de la base respecto del azúcar, *syn* o *anti*, tal como se puede apreciar en la Figura 2.9. Aunque no se encuentran representadas en la gráfica, las distancias  $d_i(2'/2''; 6/8)$  resultan también muy importantes para la determinación de dicho ángulo.



**Figura 2.9.-** Variación de la distancia entre los protones de la base y el protón  $H1'$  del azúcar en función del ángulo glicosídico. Sobre el eje se han indicado los intervalos correspondientes a las disposiciones *syn* ( $40^\circ$  a  $90^\circ$ ) y *anti* ( $-90^\circ$  a  $-170^\circ$ ) y las letras A, B, C y G indican los valores preferidos para las formas A- y B-DNA y los nucleótidos C y G en la forma Z-DNA para el  $d(\text{CGCGCG})_2$ .

Las distancias secuenciales e intercadena, tanto entre protones intercambiables como entre protones no intercambiables, son todas ellas importantes ya que dan información tanto sobre el grado de estructuración de la molécula como del tipo de estructura que se forma. La existencia de NOEs intercadena entre protones intercambiables es indicativo de formación de pares de bases. El análisis de estos NOEs permite determinar el tipo de apareamiento entre las

bases: Watson-Crick, Hoogsteen, etc.(ver Anexo A). Las distancias secuenciales entre protones intercambiables y no intercambiables, además de ser indicativas de apilamiento entre las bases varían con el tipo de hélice, así que resultan muy útiles para diferenciar hélices A- y B-DNA.

En la Tabla 2.2. se da una relación de las distancias más relevantes desde el punto de vista estructural y la información que se extrae de cada una de ellas.

**Tabla 2.2.- Relación de distancias interprotónicas estructuralmente importantes.**  
WC=Formación de pares Watson-Crick, HT=Formación de pares Hoogsteen.

Distancia	Información estructural
$d_{\beta}(T3; A2)$	Par A:T: $d(N6-O4)=2.91\text{Å}$ WC: NOE intenso H3-H2, no existe NOE H3-H8 $d(N3-N1)=2.91\text{Å}$ HT: No existe NOE H3-H2, NOE intenso H3-H8
$d_{\alpha}(T3; A61)$	
$d_{\beta}(T3; A62)$	Par G:C: $d(N4-O6)=2.91\text{Å}$ WC: NOE intenso H1-H42 $d(N3-N1)=2.95\text{Å}$ HT: no existe NOE H1-H42 $d(O2-N2)=2.86\text{Å}$
$d_{\beta}(G1; C41)$	
$d_{\beta}(G1; C42)$	
$d(2'; 4')$ $d(1'; 4')$	Conformación S o N de la ribosa. $d(2'; 4')=4.2\text{Å}$ conformación S, ( $144 \leq P \leq 190^\circ$ ) $d(2'; 4')=2.8\text{Å}$ conformación N, ( $0 \leq P \leq 36^\circ$ ) $d(1'; 4')=3.4\text{Å}$ para ambas conformaciones pero $d(1'; 4')=2.6\text{Å}$ para valores del ángulo de pseudorotación intermedios, $P=90^\circ$ .
$d(3'; 5'/5'')$	Determinación del ángulo de torsión y
$d(1'; 6/8)$ $d(2'/2''; 6/8)$	Determinación del ángulo de torsión glicosídico, conformación <i>syn</i> o <i>anti</i> . $d(1'; 6/8)=2.2\text{Å}$ 3'-endo, disposición <i>syn</i> . $d(1'; 6/8)=3.7\text{Å}$ 2'-endo, disposición <i>anti</i> . Intensidad NOE H6/8-H2'-H6/8-H2'': disposición <i>syn</i> . Intensidad NOE H6/8-H2'>H6/8-H2'': disposición <i>anti</i> .
$d_{\alpha}(2'; 1'/2)_{\beta}$ $d_{\alpha}(2'; 1')$	Distinción entre A- y B-DNA. Estos NOEs secuenciales se observan únicamente en hélices tipo A.
$d_{\alpha}(2'/2''/3'; 6/8/5)$	Distinción entre A- y B-DNA. $d_{\alpha}(2'; 6/8)$ pequeña y $d_{\alpha}(2'; 6/8/5)$ grande en caso de hélices tipo A, y al revés en caso de hélices tipo B.
$d_{\alpha}(2'/2''; 5'/5'')$	Distinción entre A- y B-DNA. $d_{\alpha}(2'/2''; 5'/5'')$ pequeña para A-DNA y grande para B-DNA
$d_{\alpha}(2'; 3')$	Distinción entre A- y B-DNA. Sólo medible en hélices tipo A ( $\sim 7\text{Å}$ )
$d_{\alpha}(2'; 2')$ $d_{\alpha}(1'; 5')$	Distinción entre A- y B-DNA. $d_{\alpha}(2'; 2')$ y $d_{\alpha}(1'; 5')$ grande en hélices A y pequeña en hélices B.

### 2.1.3.2.- Constantes de acoplamiento.

Si bien a partir de determinadas distancias intranucleotídicas es posible predecir la conformación del anillo de ribosa, la forma más adecuada y que permite identificar equilibrio dinámico entre varias conformaciones, es realizar un análisis de las constantes de acoplamiento protón-protón vecinales. Estas constantes se correlacionan con el ángulo diedro de torsión, definido por los tres enlaces que separan los protones considerados, a través de la ecuación de Karplus<sup>38</sup>:

$${}^3J(\theta) = A \cos^2 \theta + B \cos \theta + C \quad (7)$$

donde  $\theta$  es el ángulo de torsión.

Puesto que los distintos ángulos de torsión están relacionados con los ángulos de torsión endocíclicos, se puede establecer una relación entre las constantes de acoplamiento individuales

<sup>38</sup> Karplus, M. J. Chem. Phys. 1959, 30, 11.

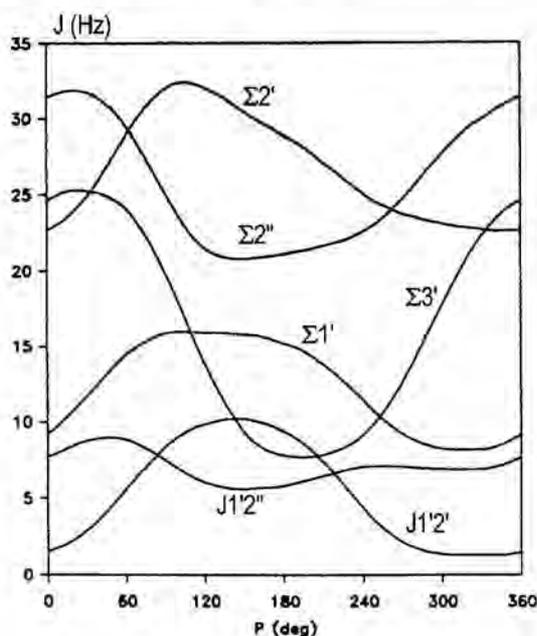
y la fase  $P$  y la amplitud  $\Phi_m$  del ángulo de pseudorotación<sup>29</sup>. De esta forma se obtiene un conjunto de constantes de acoplamiento para diferentes valores de  $P$  y  $\Phi_m$  (Anexo A). A estos efectos son especialmente importantes las constantes de acoplamiento  $J_{1'2}$ ,  $J_{2'3}$  y  $J_{3'4}$ , cuyos valores varían drásticamente al pasar de una conformación  $N$  ( $0^\circ \leq P \leq 36^\circ$ ) a una conformación  $S$  ( $144^\circ \leq P \leq 190^\circ$ ).

Es posible que no todas las constantes se puedan ajustar a un solo valor de  $\Phi_m$  y  $P$ . En este caso se debe asumir un equilibrio entre al menos dos estados para reproducir toda la información estructural, siendo el valor de la constante de acoplamiento la suma de las contribuciones de ambos conformémeros:

$$J_{ij} = (1 - f_s) \cdot J_{ij}(N) + f_s \cdot J_{ij}(S)$$

donde  $f_s$  es la fracción molar de conformémero  $S$  y  $J_{ij}(N)$  y  $J_{ij}(S)$  son las constantes de acoplamiento correspondientes a los conformémeros  $N$  y  $S$  respectivamente.

En muchas ocasiones la anchura de línea en muestras de macromoléculas es superior al valor de la mayor constante de acoplamiento vecinal. Este hecho, unido a problemas de solapamiento, hace que no se puedan obtener los valores de las constantes individuales mediante la medida directa a partir de espectros 2D. En este caso, se suele calcular el valor de la suma de constantes:  $\Sigma 1'$ ,  $\Sigma 2'$ ,  $\Sigma 2''$ ,  $\Sigma 3'$  y  $\Sigma 4'$ , cuya variación, tal como se observa en la Figura 2.10, también permite determinar la conformación del azúcar.



**Figura 2.10.-** Variación de constantes y sumas de constantes de acoplamiento calculadas entre diferentes protones del anillo de desoxiribosa en B-DNA en función del ángulo de pseudorotación  $P$ .  $\Sigma 1' = J_{1'2} + J_{1'2''}$ ,  $\Sigma 2' = J_{1'2} + J_{2'3} + J_{2'2''}$ ,  $\Sigma 2'' = J_{1'2''} + J_{2'3'} + J_{2'2''}$ ,  $\Sigma 3' = J_{2'3'} + J_{2'3''} + J_{3'4}$ .

Generalmente en el caso de oligonucleótidos de pequeño tamaño, si se dispone del valor de las sumas de constantes, se puede predecir la conformación del anillo de ribosa con bastante

fiabilidad<sup>39</sup>. En caso de necesitar valores de las constantes de acoplamiento muy precisos, se debe recurrir a métodos que permiten la determinación indirecta de constantes de acoplamiento<sup>40</sup>. Estos métodos consisten en la simulación de la estructura fina de los picos de correlación de un espectro tipo COSY.

#### 2.1.4.- Cálculo de estructuras.

La determinación de la estructura tridimensional en disolución consiste en obtener coordenadas atómicas compatibles con las restricciones experimentales de ángulos y distancias, obtenidas a partir del espectro de RMN. En realidad no se trata de encontrar la estructura que mejor cumpla las restricciones experimentales, sino todo un conjunto de estructuras compatible con toda la información experimental y que constituya una muestra representativa del espacio conformacional. Para ello se dispone de diferentes metodologías de cálculo que se pueden clasificar en función de la aproximación matemática que utilizan<sup>41</sup>. Los métodos denominados de *Geometría de Distancias*<sup>42</sup> hacen uso de la matriz métrica y operan en el espacio de distancias. Los métodos de *función de blanco variable*<sup>43</sup> consisten en la minimización de una función blanco. Se define una función positiva, cuyo valor sea cero cuando se cumplan las restricciones de distancias y diferente de cero en cualquier otro caso. Si se define dicha función con las propiedades matemáticas apropiadas, el problema de obtener una estructura compatible con las restricciones experimentales se convierte en un problema matemático de minimización de una función. Estos métodos operan generalmente en el espacio de torsiones. Por último están los métodos de *Dinámica Molecular Restringida* (RMD)<sup>44</sup> que consisten en integrar numéricamente las ecuaciones de Newton de un sistema de partículas. Estos últimos son los métodos de cálculo de estructuras más utilizados actualmente y los principales paquetes de cálculo disponen además de subrutinas para los cálculos específicos asociados a la determinación de estructuras a partir de datos de RMN.

La gran ventaja de la RMD respecto a otros métodos radica en su capacidad de saltar barreras de potencial, superando de esta forma mínimos locales y permitiendo explorar eficientemente el espacio conformacional. Estos métodos permiten además introducir de forma bastante directa información experimental como constantes de acoplamiento, desplazamientos químicos o ajustes directos de las intensidades de NOEs.

Como ya se ha mencionado, el método de Dinámica Molecular consiste en integrar las ecuaciones de Newton de un sistema de partículas:

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

<sup>39</sup> Rinkel, L.J.; Altona, C. *J. Biomol. Struct. Dyn.* **1987**, 4, 621.

<sup>40</sup> Widmer, H.; Wüthrich, K. *J. Magn. Reson.* **1987**, 74, 316-336.

<sup>41</sup> González, C. *Curso Avanzado de Resonancia Magnética Nuclear 1997*, 371-405. Escuela de Verano. Universidad de Zaragoza. Sociedad Española de Óptica.

<sup>42</sup> Havel, T.F.; Kuntz, I.D.; Crippen, G.M. *Bull. Math. Biol.* **1983**, 45, 665.

<sup>43</sup> Braun, W. *Q. Rev. Biophys.* **1987**, 19, 115-157.

<sup>44</sup> Scheek, R.M.; van Gunsteren, W.F.; Kaptein, R. *Methods in Enzymology* **1989**, 177, 204-218.

Las posiciones iniciales provienen de una estructura de partida, normalmente generada mediante algún método de Geometría de Distancias o Función Blanco y las velocidades iniciales se toman de una distribución de Maxwell-Boltzman a una temperatura dada.

En las aplicaciones de RMD se utiliza un potencial que consta de dos términos:

$$V_{tot} = V_{teo} + V_{exp}$$

El primer término es el potencial teórico habitual en simulaciones de biopolímeros y el segundo es un potencial que contiene la información estructural que se pretende ajustar. Los diferentes potenciales teóricos difieren entre sí esencialmente en la parametrización aunque todos ellos cuentan generalmente con dos términos de tipo armónico correspondientes a distancias y ángulos de enlace y un término periódico que mantiene los ángulos de torsión cerca del valor experimental de equilibrio. Es conveniente utilizar el algoritmo SHAKE<sup>45</sup> para mantener constantes las distancias y ángulos de enlace que involucran átomos de hidrógeno. Para simular las interacciones Van der Waals se utiliza un potencial Lennart-Jones, mientras que las interacciones electrostáticas se representan por un término Coulomb. También se puede introducir un término similar al de torsión para asegurar la planaridad de dobles enlaces o la quiralidad de centros asimétricos.

En cuanto al potencial que contiene la información experimental, se puede desglosar en los términos correspondientes a ángulos entre protones que forman puente de hidrógeno, distancias NOE, ángulos de torsión y constantes de acoplamiento. Los potenciales para estas restricciones se definen como una función cuyo valor es cero cuando el valor está comprendido entre las cotas superior e inferior de la restricción correspondiente y es de tipo armónico en cualquier otro caso:

$$V \begin{cases} = k(d-u)^2 & \text{si } d > u \\ = 0 & \text{si } l < d < u \\ = k(d-l)^2 & \text{si } d < l \end{cases}$$

Una vez definida la función potencial para las restricciones experimentales el problema reside en como pesar dichas restricciones con el resto de términos del potencial. Si se da un peso excesivo al potencial experimental, la dinámica tenderá a satisfacer las restricciones experimentales a costa de los términos teóricos del potencial, produciéndose distorsiones en la estructura covalente de la molécula. Si por el contrario, el peso del potencial es excesivamente bajo, la dinámica vendrá guiada exclusivamente por términos teóricos y no se llegarán a satisfacer las restricciones experimentales. Por este motivo es conveniente realizar varios cálculos preliminares con diferentes valores de las constantes de forma que se consiga un punto de compromiso, con pesos que permitan ajustar las restricciones experimentales sin distorsionar gravemente la molécula.

Los protocolos utilizados para las dinámicas consisten en calentar la estructura hasta temperaturas relativamente altas, para saltar barreras de potencial y superar mínimos locales, y enfriar suavemente después. Este tipo de estrategias se denominan de *annealing* simulado. El objetivo es obtener valores de energías suficientemente bajas, tanto de la parte teórica como de la parte experimental del potencial. Un valor de energía bajo del potencial teórico indica que la

<sup>45</sup> Ryckaert, J.P.; Cicotti, G.; Berendsen, H.J.C. *J. Comput. Phys.* **1977**, *23*, 327.

estructura tiene una estructura covalente correcta. Por su parte un valor bajo del potencial experimental indica que la estructura cumple satisfactoriamente las restricciones experimentales.

### 2.1.5.- Análisis de estructuras.

El análisis de los resultados de la determinación estructural implica por una parte, determinar la calidad del ajuste de la información experimental, esto es, determinar hasta qué punto las estructuras finales reproducen los datos experimentales y, por otra parte, analizar el grado de definición de las estructuras resultantes, lo que significa determinar la precisión de los resultados finales.

Un indicador de la calidad del ajuste de los datos experimentales es el valor final del potencial experimental, cuanto menor sea este valor mejor ajustadas estarán las restricciones experimentales. No obstante este valor del potencial depende no sólo del grado de ajuste sino también de la forma concreta de la función y de los factores de peso utilizados. Por esta razón se tiende a utilizar como indicador del grado de ajuste la suma de violaciones residuales de las restricciones en las estructuras finales, ya que no depende del potencial utilizado. Este indicador, aunque útil, es un indicador indirecto y hay que tener en cuenta que la verdadera información estructural no son las distancias sino las intensidades de los picos NOE. Las violaciones residuales de distancias dependerán, a su vez, no sólo del grado de ajuste sino de la aproximación que se haya utilizado para convertir las intensidades NOE en distancias.

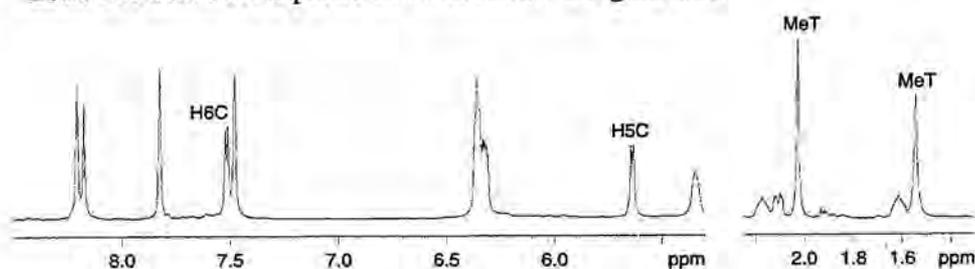
## 2.2.- ASIGNACIÓN DEL ESPECTRO DE RMN DE d<pCATTTCATT>.

El primer paso para el estudio en disolución de d<pCATTTCATT>, así como del resto de oligonucleótidos cíclicos estudiados a lo largo de este trabajo, consistió en la realización de espectros monodimensionales en D<sub>2</sub>O a una concentración de oligonucleótido de ~2.2 mM. De estos espectros se puede extraer cierta información estructural preliminar. El análisis de la zona entre 7-9 ppm del espectro permite determinar, a partir del número de señales observadas, el número de residuos que constituyen el fragmento de DNA: una señal aromática por base, en el caso de timina, citosina y guanina (H6(T), H6(C), H8(G)), y dos para adenina (H8(A), H2(A)). También es posible identificar señales correspondientes protones H6 de citosina puesto que, debido al acoplamiento con el protón H5, son las únicas señales doblete en esta zona del espectro. Las señales correspondientes a los protones H5 de citosina aparecen en la misma zona que los protones H1' de la desoxiribosa pero, al ser también señales doblete, se pueden distinguir claramente.

En este caso se observan cinco señales en la zona entre 7-9 ppm, una de ellas doblete. Al tratarse de una secuencia cíclica y simétrica, donde los nucleótidos son equivalentes dos a dos, este número se corresponde con el número de señales esperado. La secuencia lineal equivalente no sería simétrica y no se observaría una degeneración completa de las señales, por lo que el espectro monodimensional proporciona también un criterio para confirmar el carácter cíclico de la estructura.

Otra zona importante del espectro monodimensional es la zona entre 1-2 ppm. En esta zona se puede determinar el número de residuos de timidina en base al número de señales metilo. En el caso de d<pCATTTCATT> se observan dos señales que corresponden a los dos grupos metilo no equivalentes.

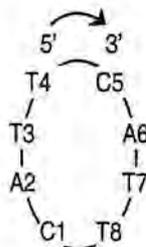
Estas dos zonas del espectro se muestran en la Figura 2.11.



**Figura 2.11.-** Espectro 1D de  $d\langle pCATTTCATT \rangle$  en  $D_2O$  a  $T=25^\circ C$ . En la zona aromática es posible asignar directamente las señales correspondientes a los protones H5 y H6 de citosina. En la zona entre 1-2 ppm se distinguen claramente las señales correspondientes a los grupos metilo de las dos timinas no equivalentes.

Aunque el espectro monodimensional nos indica que las condiciones de la muestra, en cuanto a cantidad, nivel de pureza y carácter cíclico son las adecuadas, es necesario recurrir a la espectroscopía bidimensional para obtener la asignación secuencial completa del espectro.

Para la asignación del espectro se ha utilizado la numeración de residuos que se muestra en la Figura 2.12. Puesto que todas las señales aparecen degeneradas, a lo largo de la discusión se han utilizado paréntesis para indicar el número del residuo equivalente al que se hace referencia, esto es: A2(6), T3(7), etc.



**Figura 2.12.-** Esquema de la secuencia  $d\langle pCATTTCATT \rangle$ , donde se indica la numeración utilizada en el texto. El residuo inicial para la numeración se ha elegido de forma tal que cada mitad de la molécula contiene como residuos centrales el dinucleótido autocomplementario. Dada la simetría de la secuencia, los siguientes residuos son equivalentes entre sí: C1 y C5, A2 y A6, T3 y T7, T4 y T8.

### 2.2.1.- Identificación de los diferentes sistemas de espín.

Para la identificación de los cuatro sistemas de espín se realizó un experimento DQF-COSY<sup>6</sup> y se siguió la metodología descrita en el apartado 2.1.2.3.1.

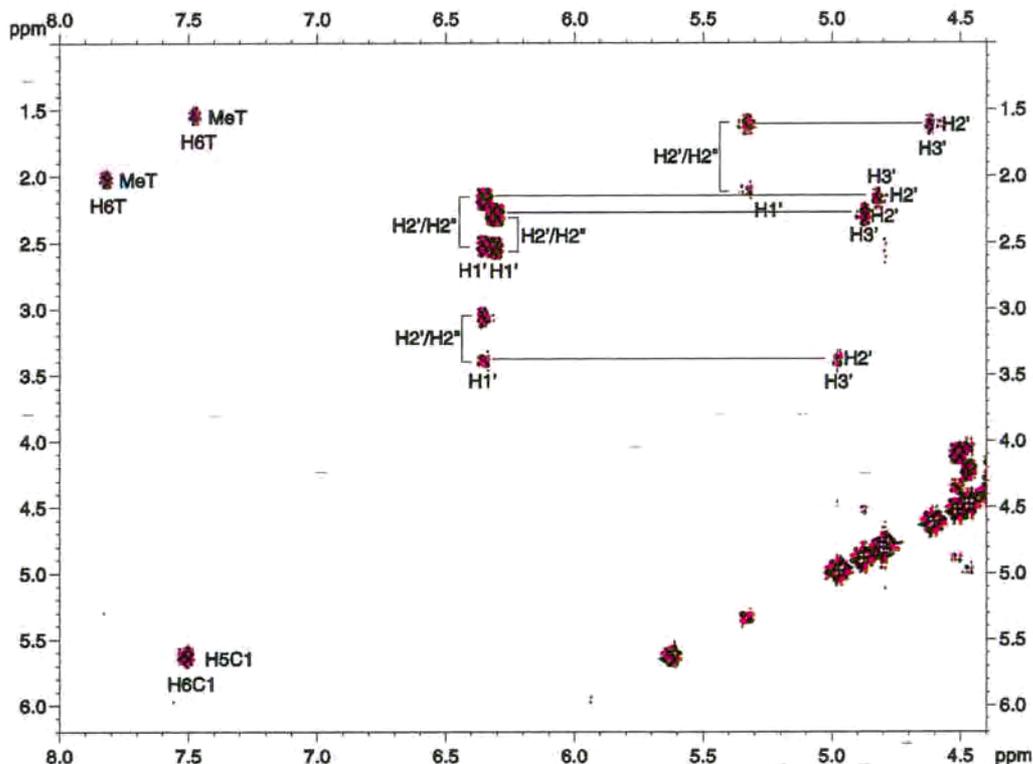
En primer lugar, se buscaron los picos de correlación H5-H6 de citosina y Me-H6 de timina, lo que permitió determinar el desplazamiento químico de los protones aromáticos de las dos timinas y de la citosina, este último identificado ya en el espectro monodimensional.

Por eliminación, las dos señales aromáticas restantes correspondían a los protones H8 y H2 de adenina. La distinción entre ambos protones se realiza a partir de experimentos inversión-recuperación<sup>46</sup>, debido al mayor valor de  $T_1$  para el protón H2.

<sup>46</sup> Freeman, R.; Kempell, S.P.; Levitt, M.H. *J. Magn. Reson.* 1980, 38, 453-479.

En segundo lugar se procedió a la identificación de los cuatro sistemas de espín correspondientes a las cuatro desoxirribosas. Se buscaron las cuatro parejas de protones  $H2'/H2''$  pertenecientes a cada desoxirribosa y, una vez identificados, se buscaron en  $\omega_1$  los picos de correlación  $H2'/H2''-H1'$  y  $H2'/H2''-H3'$ . Los sistemas de espín se completaron con los picos  $H3'-H4'$  y  $H4'-H5'/H5''$ .

En la Figura 2.13 se muestra una expansión del espectro COSY en la que se observan los picos de correlación entre protones acoplados escalarmente.



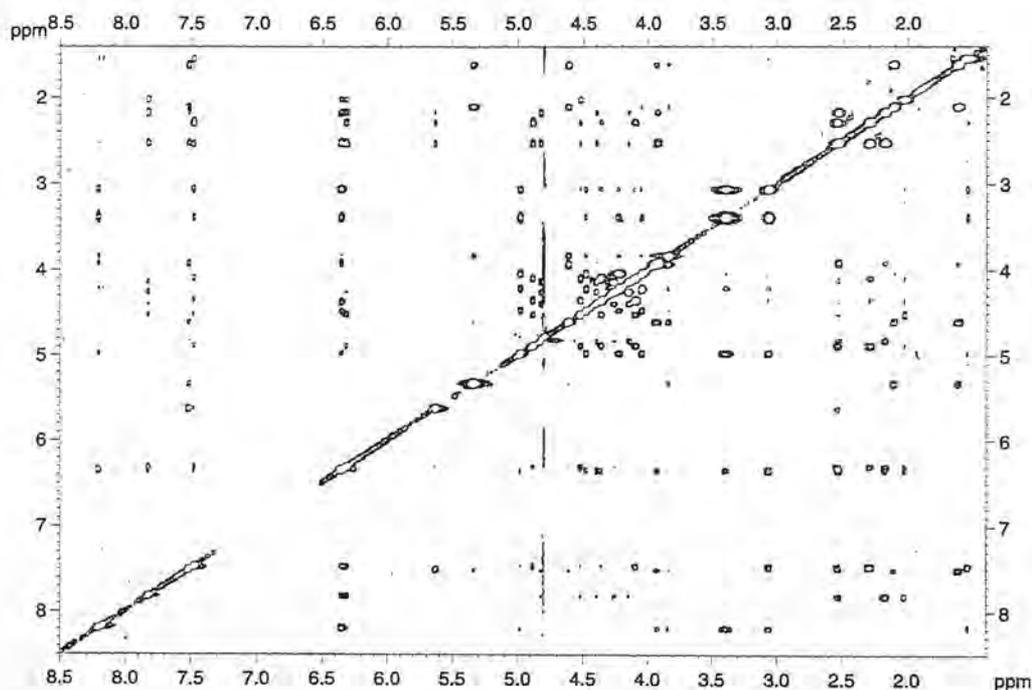
**Figura 2.13.-** Fragmento de un espectro COSY de  $d\langle pCATTTCATT \rangle$  a 2.2mM en  $D_2O$ , donde se identifican los sistemas de espín. Se indican los picos de correlación H6-Me correspondientes a las dos timinas y H6-H5 correspondiente a citosina. Siguiendo el desplazamiento de los protones  $H2'/H2''$  en la dirección de  $\omega_2$  se pueden observar los picos de correlación de dichos protones con los protones  $H1'$  y  $H3'$ . Debido a que  ${}^3J_{2'3'} > {}^3J_{2''3'}$ , se observan únicamente picos de correlación  $H2'-H3'$ .

Para conectar los sistemas de espín de las cuatro desoxirribosas con las cuatro bases, se llevó a cabo la asignación secuencial de los picos de correlación dipolar del espectro NOESY, siguiendo el protocolo descrito en el apartado 2.1.2.3.2.

### 2.2.2.- Asignación secuencial.

La primera información que se obtiene a partir del espectro NOESY es el grado de estructuración del fragmento estudiado. Si el fragmento de DNA no está estructurado, no se observan picos de correlación dipolar entre residuos adyacentes y no se puede, por tanto, llevar a cabo la asignación secuencial. En la Figura 2.14, donde se muestra el espectro NOESY de

$d\langle pCATTTCATT \rangle$ , se puede observar una gran cantidad de NOEs secuenciales, indicativos de un elevado grado de estructuración.



**Figura 2.14.-** Espectro NOESY de  $d\langle pCATTTCATT \rangle$  en  $D_2O$  a una concentración de oligonucleótido 2.2mM.

Para llevar a cabo el análisis secuencial, se tomó como punto de partida el pico de correlación H5-H6 correspondiente a citosina. En este caso establecer las conexiones entre residuos adyacentes resultó sencillo, tanto por la simplicidad de la secuencia como por la riqueza de NOEs. En la Figura 2.15 se representa el camino de asignación seguido, tanto en la zona Ar-H2'/H2'' como en la zona Ar-H1'.

Tal como se puede observar en la Figura 2.15A, de los dos protones aromáticos de adenina sólo el protón H8 da picos NOE con los protones H2'/H2'' y H1' de su propio azúcar. El protón H2 de adenina presenta en esta zona únicamente un NOE muy débil con el metilo de T3(7). Las dos timidinas se pueden diferenciar por la presencia de NOEs secuenciales H2'/H2''-H8 y H1'-H8 entre A2(6) y T3(7). Se observan además otros NOEs secuenciales: H1'-H6 entre los residuos T3(7) y T4(8) y H2'/H2''-H6 entre los residuos T4(8) y C1(5).

La asignación estereoespecífica de los protones H2'/H2'' se llevó a cabo a partir de la intensidad de los NOEs con H1' a tiempos de mezcla cortos, como se explicó en el apartado 2.1.2.3.2. En el caso de los residuos A2(6) los desplazamientos químicos de los protones H2' y H2'' están invertidos respecto a lo que suele ser habitual<sup>47</sup>, presentando el protón H2'' un desplazamiento a campo más alto que el H2'.

La Tabla con la asignación completa de la molécula figura en el Anexo B.

<sup>47</sup> Wijmenga, S.S.; Kruithof, M.; Hilbers, C.W. *J. Biomol. NMR* 1997, 10, 337-350.

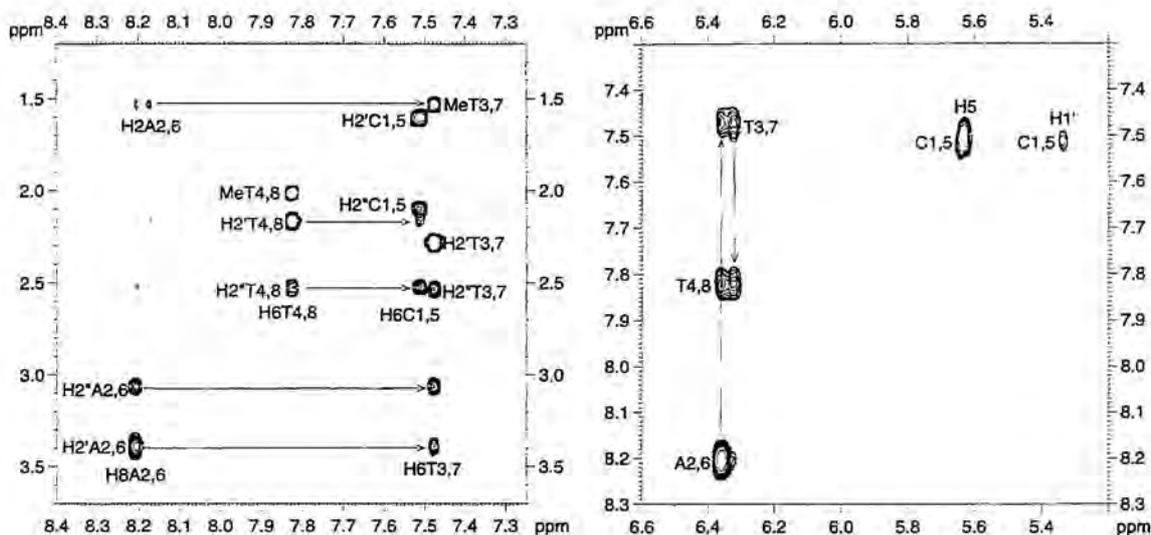


Figura 2.15.- Fragmentos del espectro NOESY en  $D_2O$  de  $d\langle pCATTcATT \rangle$  2.2mM, donde se indica el camino de asignación: A) zona Ar-H2'/H2'' y B) zona Ar-H1'. En ambas zonas se observan los NOEs secuenciales A2(6)-T3(7) que permiten asignar secuencialmente los dos residuos de timidina. Se observan además NOEs que conectan los residuos T3(7)-T4(8) y T4(8)-C1(5).

### 2.3.- ASIGNACIÓN DEL ESPECTRO DE RMN DE $d\langle pTGCTcGCT \rangle$ .

El espectro monodimensional en  $D_2O$  de  $d\langle pTGCTcGCT \rangle$  presenta una mayor complejidad, observándose en la zona aromática (7-9 ppm) más señales de las ocho esperadas teniendo en cuenta que la estructura no es simétrica. En el espectro realizado a  $35^\circ C$ , que se muestra en la Figura 2.16, se puede observar que no todas las señales tienen la misma intensidad, sino que parece haber dos grupos de señales diferentes.

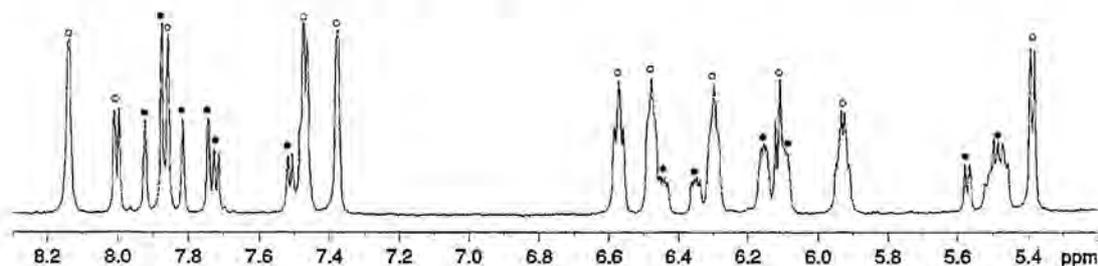
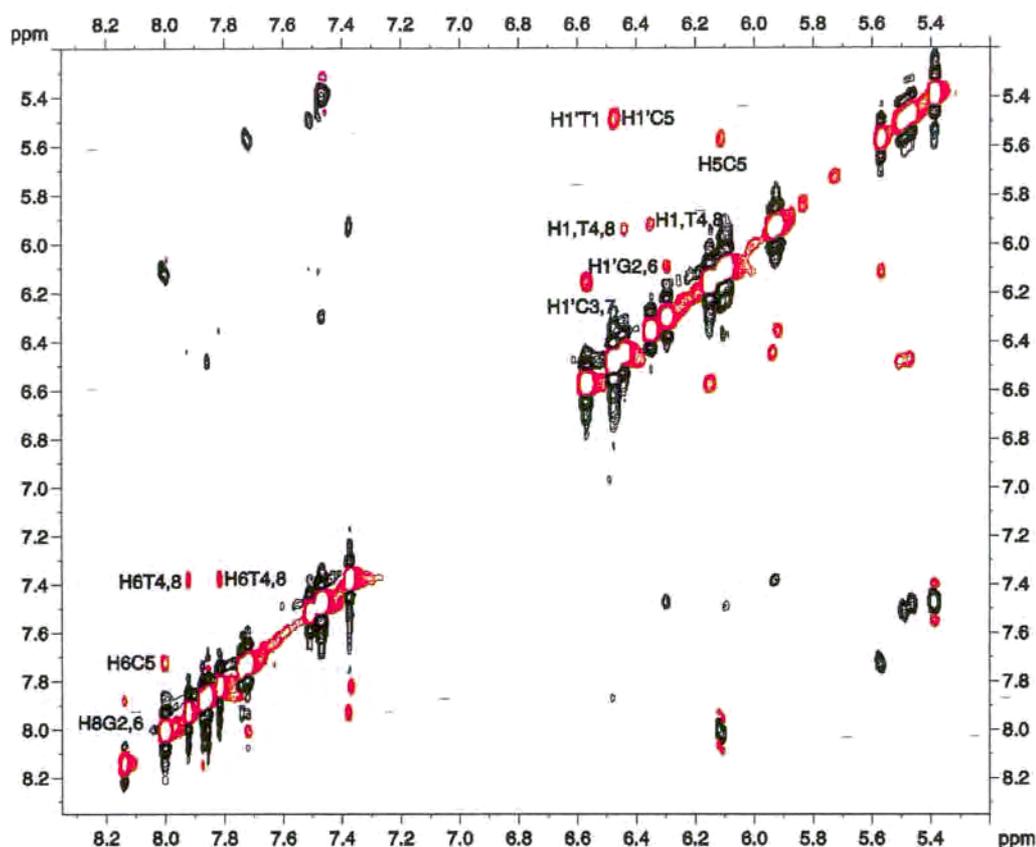


Figura 2.16.- Zona aromática del espectro 1D de  $d\langle pTGCTcGCT \rangle$  a 2.4mM,  $T=35^\circ C$ . Para diferenciar las señales correspondientes a cada especie se han utilizado los símbolos  $\bullet$  y  $\circ$ .

Dado que se disponía de datos experimentales que garantizaban la pureza y circularidad del producto (ver sección 2.2.7 de Materiales y Métodos), el número adicional de señales sugería la presencia de dos especies en equilibrio lento en la escala de tiempos de RMN. Para confirmar esta hipótesis se realizó un experimento ROESY. En este espectro se observaron picos de signo opuesto a los picos de correlación dipolar, lo que permitió confirmar la existencia

de intercambio químico e identificar, por tanto, los dos grupos de señales correspondientes a cada una de las dos especies en equilibrio. En la Figura 2.17 se muestran dos zonas del espectro ROESY donde se observan claramente las señales debidas a intercambio entre las dos especies.



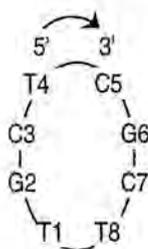
**Figura 2.17.-** Espectro ROESY de  $d\text{-pTGCTCGCT}$  a 2.4mM en el que se muestra la zona de la diagonal correspondiente a protones aromáticos y  $H1'$ . En rojo se muestran los picos de intercambio  $H6/H8\text{-}H6/H8$  y  $H1'\text{-}H1'$  entre ambas especies y en negro los picos de correlación dipolar (NOE), que son de signo contrario.

Una vez identificados los dos grupos de señales se pudo llevar a cabo la asignación de cada una de las especies, de forma análoga a la descrita para la secuencia  $d\text{-pCATTTCATT}$ .

### 2.3.1.- Identificación de los sistemas de espín.

La identificación de los diferentes sistemas de espín resultó en este caso más complicada debido a solapamiento de algunas de las señales y también a la existencia de señales degeneradas, puesto que en la zona aromática del espectro 1D no se observan las 16 señales esperadas para el equilibrio entre dos formas sino únicamente 12.

El problema de solapamiento se solucionó gracias al diferente comportamiento que presentan los dos grupos de señales con la temperatura, resultando ser muy mayoritaria una de las dos especies a baja temperatura. De los 12 protones aromáticos observados en el espectro 1D, 5 corresponden a la especie mayoritaria a baja temperatura, que se ha denominado especie **1**, y 7 a la especie **2**, poniéndose de manifiesto la existencia de degeneración de señales en ambas especies. Para la asignación se utilizó la numeración que se muestra en la Figura 2.18.



**Figura 2.18.-** Numeración de residuos utilizada para la asignación de  $d\langle pTGCTcGCT \rangle$ , tanto para la especie 1 como para la especie 2.

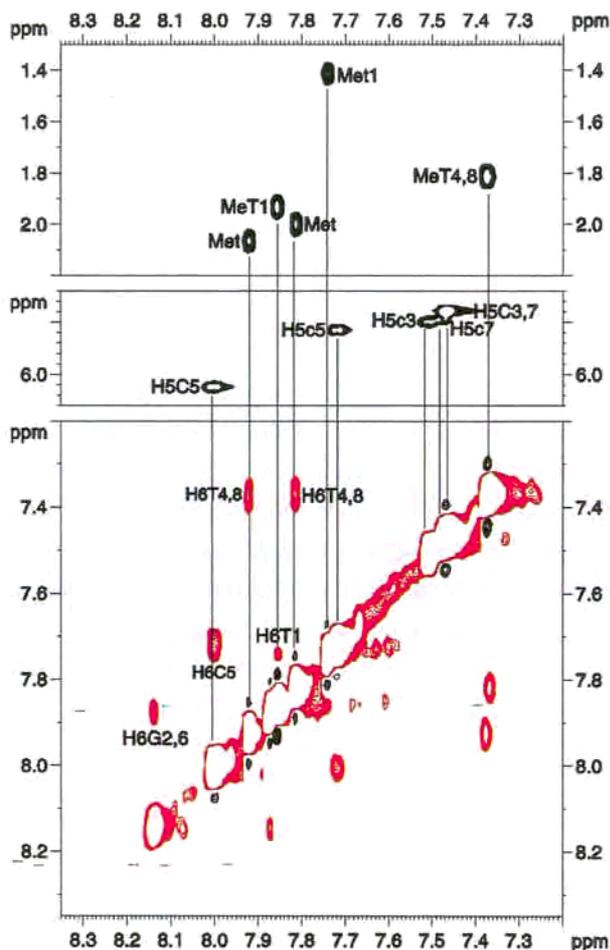
Se realizó un experimento TOCSY a 5°C que permitió la identificación de los cinco sistemas de espín correspondientes a desoxirribosas para la especie 1. En este espectro se observan solamente dos picos de correlación H6-H5 y dos picos Me-H6, por tanto dos de las citidinas y dos de las timidinas presentan señales degeneradas. Los otros dos residuos con señales degeneradas son las dos guanosinas, por la posición casi equivalente que ocupan en la secuencia. Por tanto para esta especie resulta que G2=G6, C3=C7 y T4=T8.

Para la asignación de los distintos sistemas de espín de la especie 2 se realizó un experimento TOCSY a 35°C, por existir a esta temperatura una buena dispersión de las señales de ambas especies. Para la especie 2 se observan tres picos Me-H6 y tres picos H6, por tanto las señales degeneradas corresponden a los dos residuos de guanosina. En la Figura 2.19 se muestran dos zonas del espectro TOCSY (Figura 2.19A y Figura 2.19B) y la zona Ar-Ar del espectro ROESY (Figura 2.19C), donde se pueden observar los picos Me-H6 y H6-H5 correspondientes a las timidinas y citidinas de ambas especies y los correspondientes picos de intercambio H6/H8-H6/H8 entre protones aromáticos. Gracias a la buena dispersión de señales, todos los sistemas de espín pudieron asignarse, salvo uno (C7) en que no pudieron completarse los protones H4' y H5'/H5''.

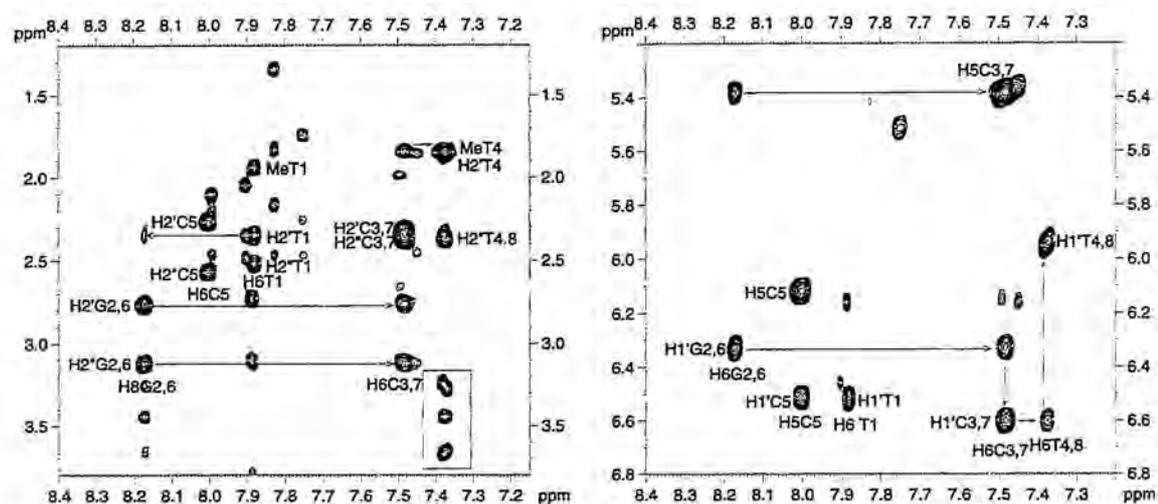
### 2.3.2.- Asignación secuencial.

Para llevar a cabo la asignación secuencial de las especies 1 y 2 se realizaron experimentos NOESY a diferentes temperaturas. Como en el caso de  $d\langle pCATTcATT \rangle$ , los espectros NOESY presentan numerosos picos secuenciales indicando también un alto grado de estructuración en las dos especies.

Para la asignación de la especie 1 se realizó un experimento NOESY a 5°C. La conexión de los protones aromáticos con los diferentes sistemas de espín de desoxirribosas se pudo realizar fácilmente gracias a los NOEs H2'/H2''-H6/H8 secuenciales que conectaban los residuos G2(6), C3(7) y T4(8). En la Figura 2.20 se muestra el camino de asignación.



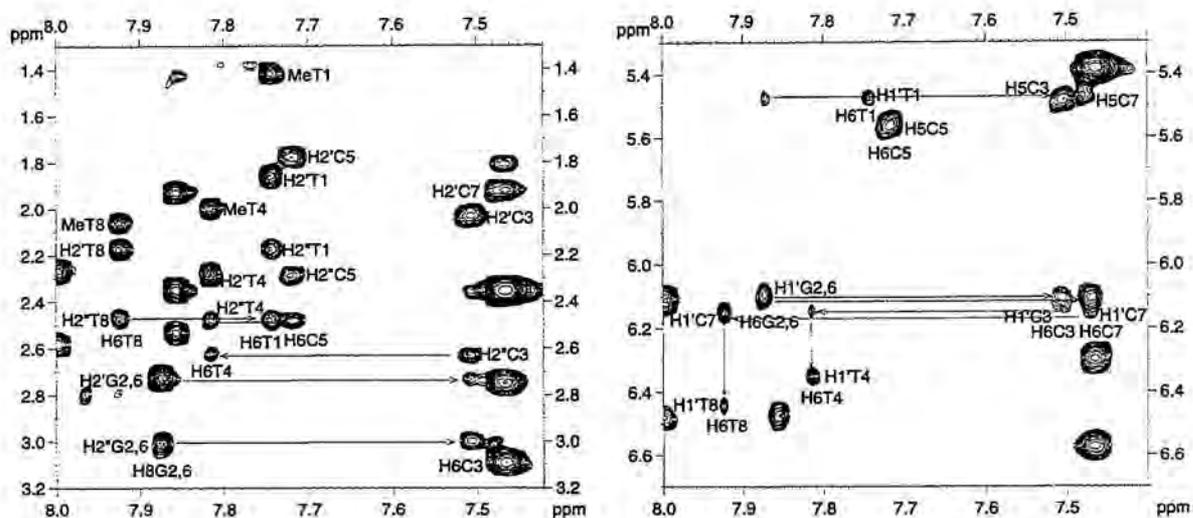
**Figura 2.19.-** A) Espectro TOCSY de  $d\langle p\text{TGCTCGCT} \rangle$ , zona H6/H8-Me. B) Espectro ROESY, zona H6/H8-H6/H8. Para cada H6 de timina, siguiendo la dirección  $\omega_1$  se observan los cinco picos de correlación H6-Me correspondientes a las cinco timinas, dos de la especie 1 y tres de la especie 2. Los picos de intercambio permiten diferenciar qué timinas corresponden a cada una de las especies: la especie 1 se indica con mayúsculas (T) y la especie 2 con minúsculas (t).



**Figura 2.20.-** Fragmento de espectro NOESY en  $D_2O$  de  $d\langle p\text{TGCTCGCT} \rangle$  a 2.4mM,  $T=5^\circ\text{C}$ , donde se indica el camino de asignación para la especie 1: A) zona H6/H8-H2'/H2'' y B) zona H6/H8-H1'. Se observan NOEs H2'/H2''-H6 y H1'-H6 entre los residuos G2(6) y C3(7). Entre los residuos C3(7) y T4(8) no se observan los NOEs H2'/H2''-H6 debido a que el H2'' de T4(8) se solapa con los protones H2' y H2'' de C3(7) que están degenerados. Sí se observan, en cambio, los NOEs H6-Me (A) y H1'-H6 (B) entre C3(7) y T4(8).

De la asignación de la especie 1 hay que destacar los desplazamientos químicos a campos anormalmente altos que presentan los protones H4' y H5'/H5'' de los residuos T4(8), que en la zona del espectro NOESY que se muestra en la Figura 10A se han encerrado dentro de un rectángulo.

La asignación secuencial de la especie 2 se realizó a partir de un espectro NOESY a  $35^\circ\text{C}$ . Como para la especie 1, la existencia de NOEs secuenciales entre los residuos G2(6), C3, C7, T4 y T8, juntamente con la existencia de picos de intercambio entre ambas especies, permitió una asignación inequívoca. El camino de asignación para la especie 2 se muestra en la Figura 2.21.



**Figura 2.21.-** Fragmento de espectro NOESY en  $D_2O$  de  $d\langle pTGCTCGCT \rangle$  a 2.4mM,  $T=35^\circ C$ , donde se indica el camino de asignación para la especie 2: A) zona H6/H8-H2'/H2'' y B) zona H6/H8-H1'. Se observan NOEs H2'/H2''-H6 y H1'-H6 entre los residuos G2(6)-C3, G2(6)-C7, C3-T4, T4-C5 y T8-T1

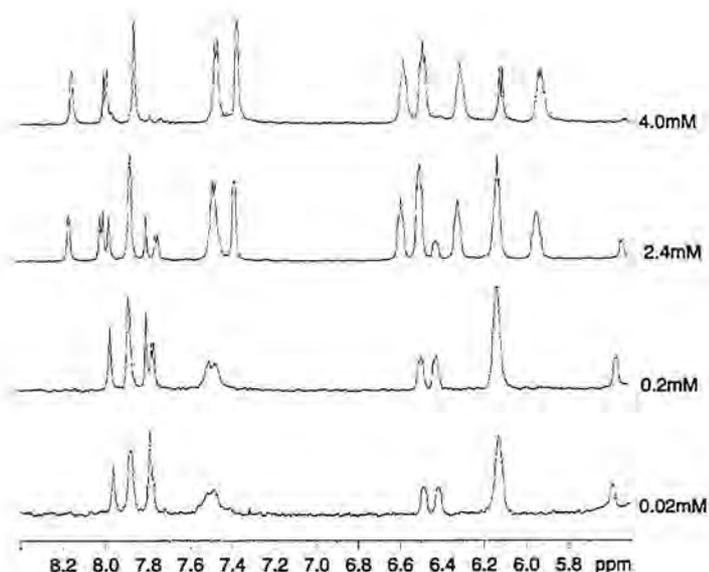
Resulta importante resaltar las diferencias existentes entre el paso del residuo 3 al residuo 4 y el paso equivalente, del 7 al 8. Mientras que entre los residuos C3 y T4 se observan NOEs secuenciales H2'/H2''-H6, éstos no se observan para los residuos C7 y T8. Existe también diferencia en cuanto a la intensidad de los NOEs H2''T4-H6C5 y H2''T8-H6T1, siendo este último significativamente más intenso. Estos datos ponen de manifiesto que los dos mini-loops en la especie 2 no son completamente simétricos.

Las tablas de asignación de las especies 1 y 2 figuran en el Anexo B.

#### 2.4.- EQUILIBRIO MONÓMERO-DÍMERO.

La presencia de dos conjuntos de señales en los espectros de  $d\langle pTGCTCGCT \rangle$  así como la existencia de picos de intercambio entre las señales de ambos conjuntos, indican que existe un equilibrio entre dos especies y que dicho equilibrio es lento dentro de la escala de tiempos de RMN. Podría tratarse de un equilibrio conformacional o bien de un equilibrio de asociación. Para distinguir entre los dos fenómenos, se realizaron espectros monodimensionales a diferente concentración de oligonucleótido.

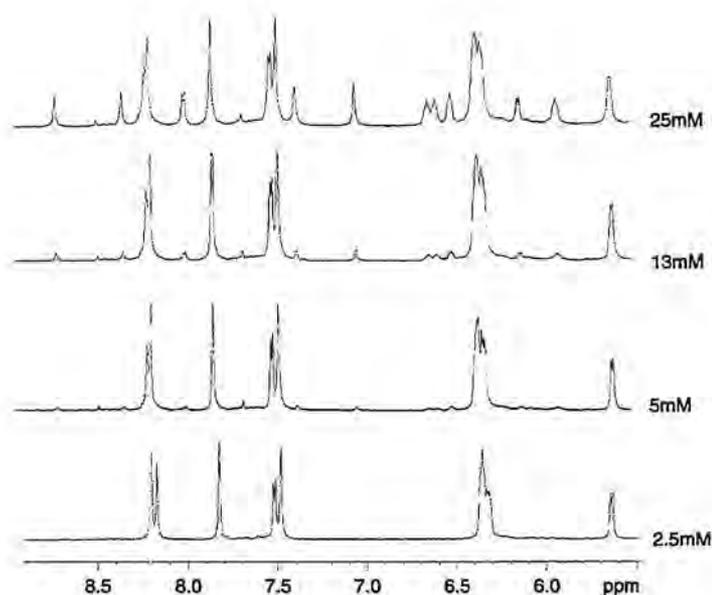
Tal como se puede observar en la Figura 2.22, a baja concentración se observa solamente uno de los dos grupos de señales, el correspondiente a la especie 2, mientras que a alta concentración se observan únicamente las señales correspondientes a la especie 1. El cambio drástico de las señales en el espectro al variar la concentración de oligonucleótido, pone de manifiesto que se trata de un equilibrio de asociación.



**Figura 2.22.-** Serie de espectros de  $d\langle pTGCTcGCT \rangle$  a diferente concentración de oligonucleótido,  $T=20^{\circ}C$ .

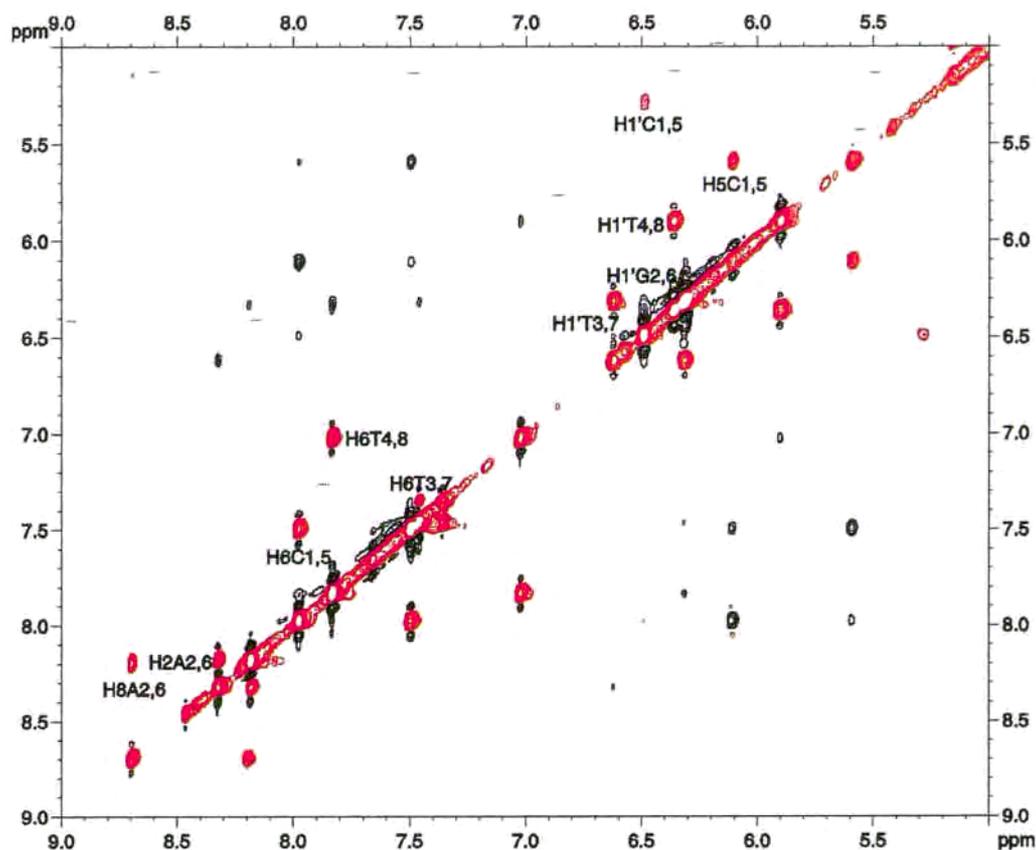
A partir de estos espectros, se pudo deducir que en el caso de la especie **2** se trata de una estructura monomérica, ya que al aumentar la dilución de 0.2 a 0.02 mM no se observan cambios en el espectro. En el caso de **1**, dada la simetría observada en los espectros de RMN y, el que se tratara de un producto de asociación, se llegó a la conclusión de que debía tratarse de una forma dimérica. La estructura de ambas especies se discutirá más adelante en base a los NOEs secuenciales observados.

Una vez se tuvo la evidencia de la existencia de equilibrio de asociación para el  $d\langle pTGCTcGCT \rangle$ , cabía preguntarse si este equilibrio se daba también en el caso del oligonucleótido  $d\langle pCATTcATT \rangle$  pero a una concentración diferente a la estudiada. Se realizaron espectros monodimensionales a diferente concentración de oligonucleótido, no observándose ningún cambio en el espectro al diluir la muestra. En cambio, se observó la aparición de un segundo grupo de señales al aumentar la concentración por encima de 10 mM, tal como se muestra en la Figura 2.23.



**Figura 2.23.-** Serie de espectros de  $d\langle pCATTcATT \rangle$  a diferente concentración de oligonucleótido,  $T=20^{\circ}\text{C}$ .

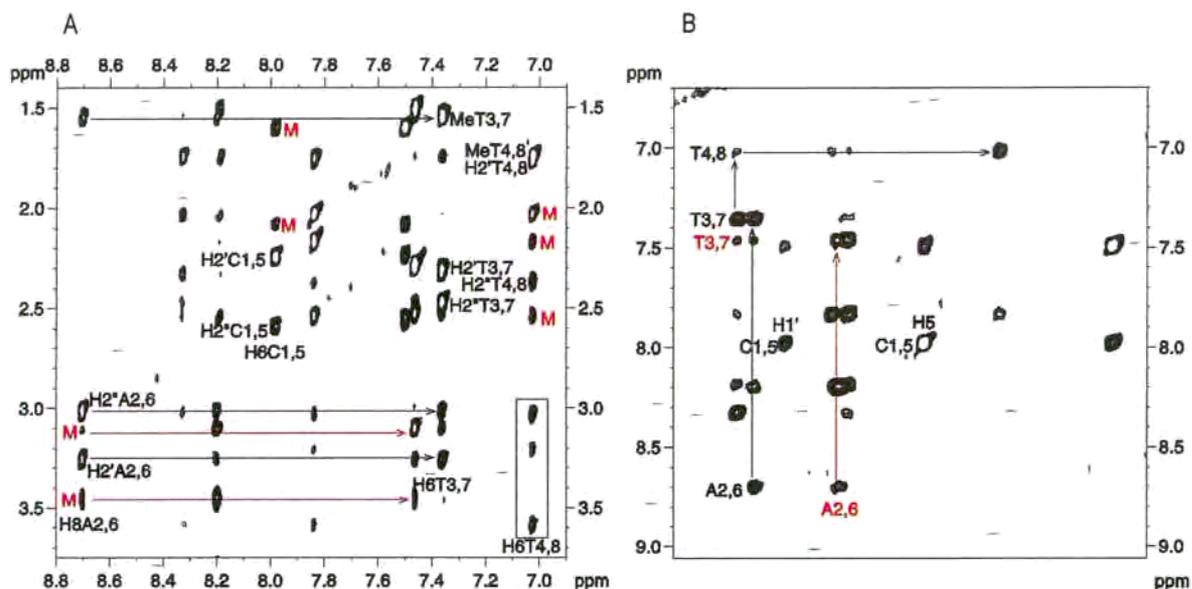
Para confirmar que en este caso se trataba también de un fenómeno de equilibrio químico, se realizó un espectro ROESY a una concentración de oligonucleótido de  $\sim 20$  mM, a la que ambas especies coexistían. Como en el caso de  $d\langle pTGCTcGCT \rangle$ , se pueden observar los picos debido al intercambio químico entre ambas especies, tal como se muestra en la Figura 2.24.



**Figura 2.24.-** Espectro ROESY de las zonas Ar-Ar y H1'-H1' de  $d\langle p\text{CATTCATT} \rangle$  a  $\sim 20\text{mM}$ ,  $T=20^\circ\text{C}$ . Se muestran en rojo los picos debido a intercambio químico (H6/H8-H6/H8 y H1'-H1') y en negro los picos NOE, de signo contrario.

Para la asignación de esta forma dimérica aparecida a alta concentración, se realizaron experimentos TOCSY y NOESY a  $\sim 20\text{ mM}$ , observándose también para esta especie un gran número de picos NOE. En este caso la dispersión de señales es aún mayor y, tal como se puede observar en la Figura 2.25, la asignación de ambas especies podría realizarse simultáneamente. Un detalle importante en este caso es la presencia de numerosos picos NOE+intercambio, observándose en  $\omega_1$ , para un determinado desplazamiento de protón aromático, picos NOE a desplazamientos correspondientes a protones H1', H2'/H2'' de ambas especies.

Como en el caso de  $d\langle p\text{TGCTCGCT} \rangle$ , los protones H4', H5'/H5'' de los residuos T4(8) presentan también desplazamientos a campos anormalmente altos.



**Figura 2.25.-** Fragmento de espectro NOESY en  $D_2O$  de  $d\langle pCATTTCATT \rangle$  a 20 mM,  $T=5^\circ C$ , donde se indica el camino de asignación. A) zona H6/H8-H2'/H2'' y B) zona H6/H8-H1'. Se observan NOEs H2'/H2''-H6 y H1'-H6 entre los residuos A2(6)-T3(7). Entre los residuos T3(7)-T4(8) se observan NOEs H1'-H6. Con una M en rojo se han indicado picos que corresponden a NOE+intercambio y que corresponden a desplazamientos químicos de protón en la especie monomérica. Se han encuadrado los protones H4' y H5'/H5'' de los residuos T4(8) que presentan un desplazamiento químico anormal.

Las señales correspondientes a la forma dimérica indican también una completa simetría en la estructura.

## 2.5.- INFORMACIÓN ESTRUCTURAL EXPERIMENTAL.

Como ya se ha comentado anteriormente, el elevado número de picos NOE observados en los espectros, es una indicación clara de que las dos especies en equilibrio presentan un alto grado de estructuración. A continuación se detalla la información estructural que se extrae de los espectros para ambas formas, tratándose de manera conjunta los dos monómeros y los dos dímeros, para las dos secuencias estudiadas.

En este sentido, es importante remarcar las diferentes condiciones en las que se han obtenido los espectros de ambas secuencias. En el caso de  $d\langle pCATTTCATT \rangle$  se pueden realizar espectros de 100% monómero a concentraciones óptimas de resonancia, obteniéndose espectros de muy buena calidad en los que se observan un gran número de picos de correlación para esta forma. En el caso de  $d\langle pTGCTCGCT \rangle$ , a la misma concentración, coexisten la forma monomérica y dimérica, siendo el dímero la especie mayoritaria. Esto hace que para la forma monomérica se observen menos picos de correlación y de menor intensidad.

En los experimentos en  $H_2O$  de  $d\langle pTGCTCGCT \rangle$  se observan también las señales correspondientes a protones intercambiables de ambas especies. Estas señales se comentarán de forma separada en los apartados correspondientes.

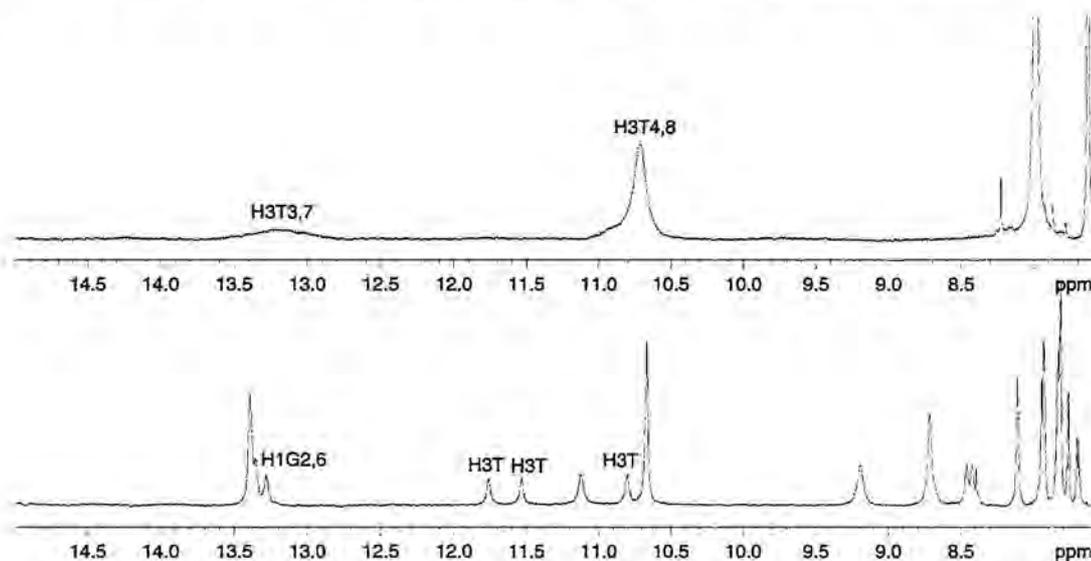
## 2.5.1.- Rasgos principales de las formas monoméricas.

### 2.5.1.1.- Información de espectros en $H_2O$ .

La información estructural más importante que se extrae de experimentos en  $H_2O$  es la formación de puentes de hidrógeno entre bases complementarias. En este caso, al tratarse de estructuras monoméricas únicamente es posible la formación de pares de bases intramoleculares. Para determinar su posible existencia se realizaron espectros monodimensionales en  $H_2O$  a  $5^\circ C$  y a una concentración de oligonucleótido 2.2-2.4mM, Figura 2.26.

En el caso de  $d\langle pCATTTCATT \rangle$ , se observa una señal en la zona de protones imino apareados, entre 13-14 ppm, y una señal en la zona de protones imino no apareados, entre 10-11 ppm, correspondientes a las dos timinas no equivalentes T3(7) y T4(8). La señal en la zona de iminos apareados implica la formación de pares A·T intramoleculares (A2-T7, T3-A6).

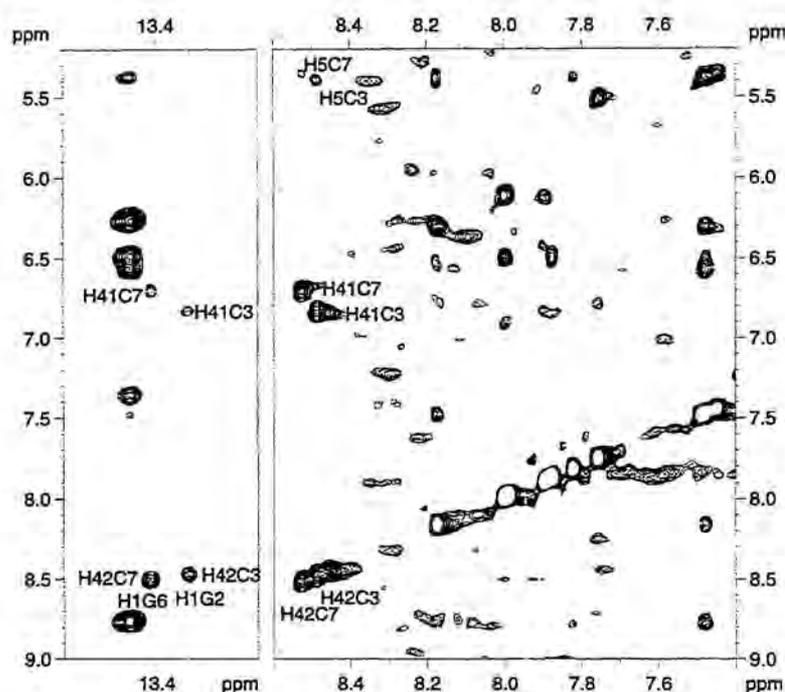
En el caso de  $d\langle pTGCTCGCT \rangle$ , se observan para el monómero dos señales en la zona de iminos apareados y tres en la zona de iminos no apareados. La presencia de las dos primeras señales pone de manifiesto la no equivalencia total de los residuos de guanosina, aunque para el resto de protones de estos residuos se hayan observado señales degeneradas. En este caso se confirma también la formación de pares de bases intramoleculares (G2-C7, C3-G6).



**Figura 2.26.-** Espectro 1D en  $H_2O$  de  $d\langle pCATTTCATT \rangle$ , 2.2mM, y  $d\langle pTGCTCGCT \rangle$ , 2.4mM, a  $T=5^\circ C$  y  $pH=7$ . En el caso del  $d\langle pTGCTCGCT \rangle$  las señales que no aparecen etiquetadas corresponden a protones intercambiables correspondientes a la especie dimérica y que se comentarán en el apartado correspondiente.

En ambos casos las señales de los protones imino son anchas. Esto es indicativo de que el protón está expuesto al disolvente y se encuentra en intercambio rápido con él. Se realizaron experimentos NOESY en  $H_2O$  a  $5^\circ C$  para poder observar los patrones de NOEs característicos de pares de bases A·T y G·C y determinar el tipo de apareamiento que estabiliza los respectivos monómeros. En el caso de  $d\langle pTGCTCGCT \rangle$ , las señales a baja temperatura son suficientemente estrechas y, a pesar de tratarse de picos muy débiles, se pudieron observar los

picos de correlación H1G2(6)-H42C3(7), H42C3-H5C3 y H42C7-H5C7 que permiten asignar secuencialmente estos protones, Figura 2.27. En el caso de los grupos amino de G y C, se ha designado como HX2 el protón amino que está formando puente de hidrógeno y HX1 el protón amino que no forma puente de hidrógeno. La presencia de los picos de correlación observados es indicativa de apareamiento G-C tipo Watson-Crick.



**Figura 2.27.-** Fragmento del espectro NOESY en  $H_2O$   $d\langle pTGCTcGCT \rangle$ , 2.4mM,  $T=5^\circ C$ ,  $pH=7$ . Se han etiquetado los picos de correlación H1G2-H42C3, H1G6-H42C7, H1G2-H41C3 y H1G6-H41C7 de la forma monomérica que permiten establecer que se trata de un apareamiento tipo Watson-Crick.

En el caso de  $d\langle pCATTcATT \rangle$  la anchura de la señal imino de la timina 3 impide la observación de ningún pico de correlación, lo que impide determinar el tipo de apareamiento (Watson-Crick o Hoogstein). Es interesante observar la evolución de la anchura de esta señal al variar la temperatura. Habitualmente, las señales de los protones intercambiables se ensanchan al aumentar la temperatura debido al incremento de la tasa de intercambio con el disolvente. Sin embargo, en esta señal se observa un ligero estrechamiento al pasar de 5 a 15°C, para luego ensancharse definitivamente a temperaturas más altas (véase Figura 2.30). Este comportamiento podría indicar que el protón imino de esta timina se encuentra involucrado en algún fenómeno de equilibrio (distinto del equilibrio monómero-dímero).

### 2.5.1.2.- Información de espectros en $D_2O$ .

La existencia de picos H2'/H2"-H6/H8 secuenciales en el espectro NOESY, característicos de una conformación apilada de las bases, constituyen un indicio adicional de la existencia de pares de bases, previamente confirmada por los espectros en  $H_2O$ . En el caso de una doble hélice de DNA, donde no existe interrupción en el apareamiento de las bases, estos

NOEs secuenciales se observan a lo largo de todo el fragmento. Tanto en el caso de d<pCATT<CATT> como en el de d<pTGCT<CGCT>, esto no es así y sólo se observan NOEs H2'/H2''-H6/H8 intensos entre los residuos 2-3 y 6-7, que son los que están formando los dos pares intramoleculares.

La intensidad de los NOEs H1'-H6/H8 y H2'/H2''-H6/H8 permite deducir la disposición de la base respecto del azúcar (Tabla 2.2). En el caso de d<pTGCT<CGCT>, todos los picos NOE H1'-H6/H8 son poco intensos y el protón H6/H8 da un NOE más intenso con H2' que con H2''. Esto indica que todos los nucleósidos se encuentran en disposición *anti*. En el caso de d<pCATT<CATT>, en cambio, el pico H1'-H8 correspondiente a la adenosina presenta un intensidad muy superior al resto de NOEs H1'-H6/H8 de los tres nucleósidos restantes (como se puede apreciar en la Figura 2.15). Esto podría indicar una conformación *syn* para este residuo, sin embargo dicha conformación es incompatible con el NOE intenso observado entre los protones H2' con H8 de la adenina.

Estos NOEs, aparentemente contradictorios, podrían explicarse suponiendo que las adeninas se encuentran en un equilibrio dinámico entre los isómeros *syn* y *anti*. Dicho equilibrio, al contrario del equilibrio monómero-dímero, sería rápido en la escala de tiempos de la RMN y daría lugar, por tanto, a un solo juego de señales. El isómero *syn* sería el responsable del NOE intenso H1'-H8, mientras que el isómero *anti* daría lugar a los NOEs entre H8 y los restantes protones de su azúcar (H2'/H2'' y H3').

Además de los NOEs intraresiduales y secuenciales entre los nucleótidos centrales del monómero, se observan algunos contactos entre protones de los residuos 1, 4, 5 y 8. Los NOEs entre los protones de las bases y sus respectivos azúcares son indicativos de conformaciones *anti* en todos los residuos. Los NOEs secuenciales no existen o son muy débiles, lo que indica que estos residuos se encuentran globalmente desordenados.

Para determinar la conformación del azúcar (puckering), se estimó el valor de las sumas de constantes de acoplamiento  $\Sigma 1'$ ,  $\Sigma 2'$  y  $\Sigma 2''$ , a partir de espectros DQF-COSY. Para los dos oligonucleótidos estudiados, se obtuvieron valores para dichas sumas consistentes con una conformación S de todas las ribosas (véase Anexo C). La conformación S de las ribosas se puede deducir también directamente del espectro DQF-COSY, ya que para una conformación S  ${}^3J_{2'3} \gg {}^3J_{2''3}$  y no se observan los picos de correlación H2''-H3' (ver Figura 2.13).

Se observan además algunos contactos entre protones de los residuos 3(7)-4(8) y 4(8)-5(1). En particular existen una serie de NOEs que, aunque débiles, podrían estar indicando una cierta preferencia de los residuos 4 y 8 a colocarse con la base apilada sobre los pares de bases del fragmento central: NOEs H1'(3,7)-H6(4,8), Me(4,8)-H6(3,7).

Toda esta información experimental es consistente con una estructura tipo *dumbbell* de las formas monoméricas, con un fragmento central tipo B-DNA constituido por los residuos 2, 3, 6 y 7, cuyas bases se encuentran apiladas. Los cuatro residuos restantes forman dos mini *loops* de dos residuos cada uno. Todas las desoxirribosas se encuentran mayoritariamente en conformación S y los ángulos glicosídicos en *anti*, salvo en el caso de las adeninas de d<pCATT<CATT>, donde las intensidades de los NOEs parecen indicar que existe un equilibrio conformacional entre las disposiciones *syn* y *anti* de la purina.

## 2.5.2.- Rasgos principales de las formas diméricas.

### 2.5.2.1.- Información de espectros en $H_2O$ .

En el caso del dímero de  $d\langle pTGCTCGCT \rangle$ , de las tres señales de protón imino esperadas (G2(6), T4(8) y T1), se observa una señal en la zona de imino apareado y dos señales en la zona de iminos no apareados. A diferencia de las señales observadas para el monómero, la señal del protón imino apareado G2(6) es muy estrecha, y se puede observar a temperaturas tan altas como  $60^\circ C$ , Figura 2.28. Este hecho pone de manifiesto que los protones imino de las guaninas apareadas están muy protegidos e intercambian muy lentamente con el solvente. Esto es indicativo de una estructura muy compacta.

También se observan diferencias entre los protones imino de las timinas. La señal correspondiente al imino de T4(8) es también una señal estrecha y se mantiene hasta casi  $50^\circ C$ , a diferencia de la señal imino de T1, que es ancha y desaparece a baja temperatura. Este último comportamiento es similar al de las señales correspondientes a protones imino no apareados de la forma monomérica. Esto estaría indicando que las timinas 4 y 8 en la forma dimérica no se encuentran totalmente expuestas al disolvente.

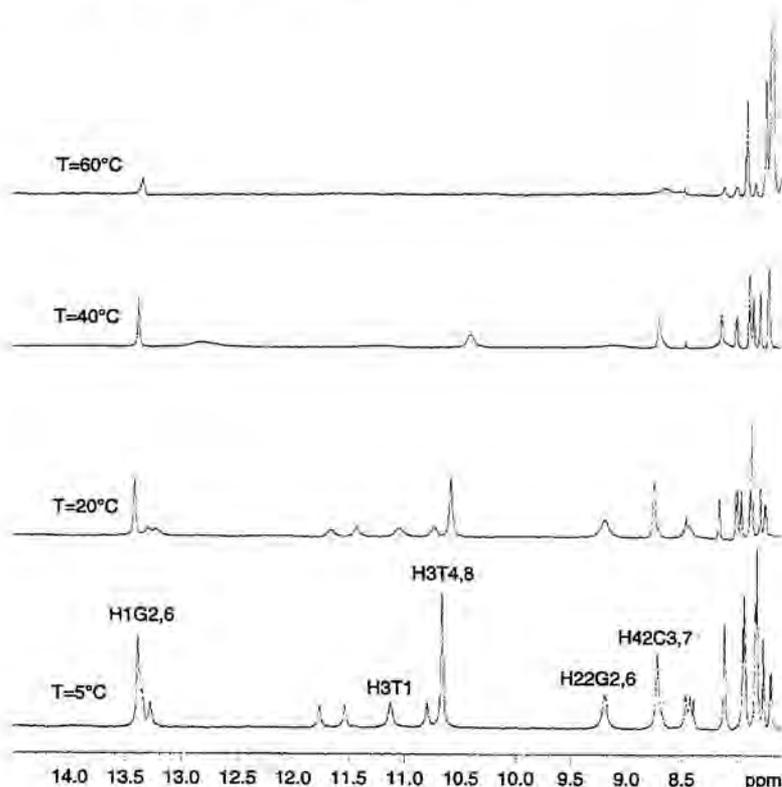
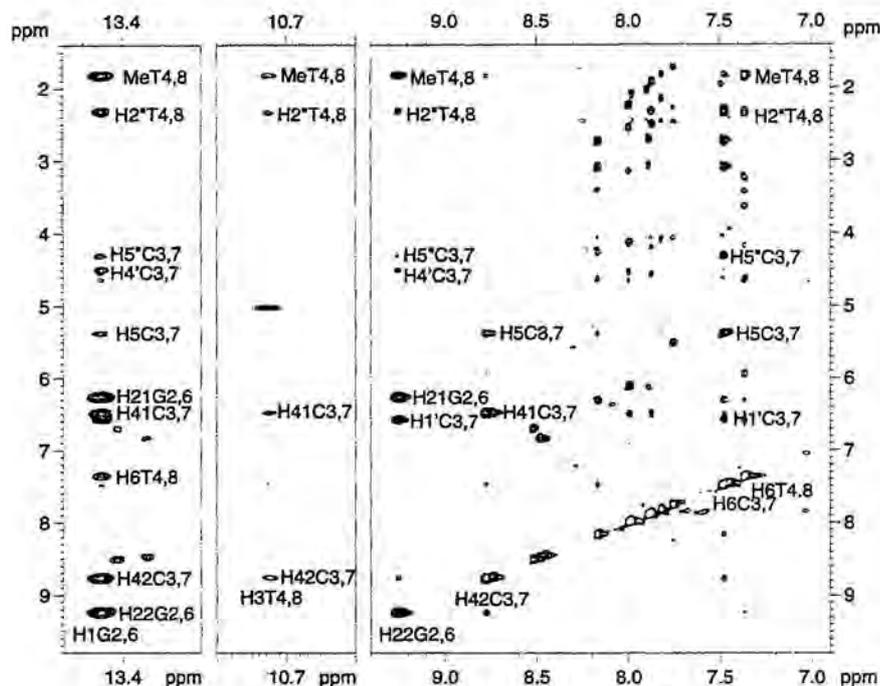


Figura 2.28.- Serie de espectros 1D en  $H_2O$  de  $d\langle pTGCTCGCT \rangle$ , 2.4mM, pH=7. Se indican únicamente las señales de los protones intercambiables de la forma dimérica.

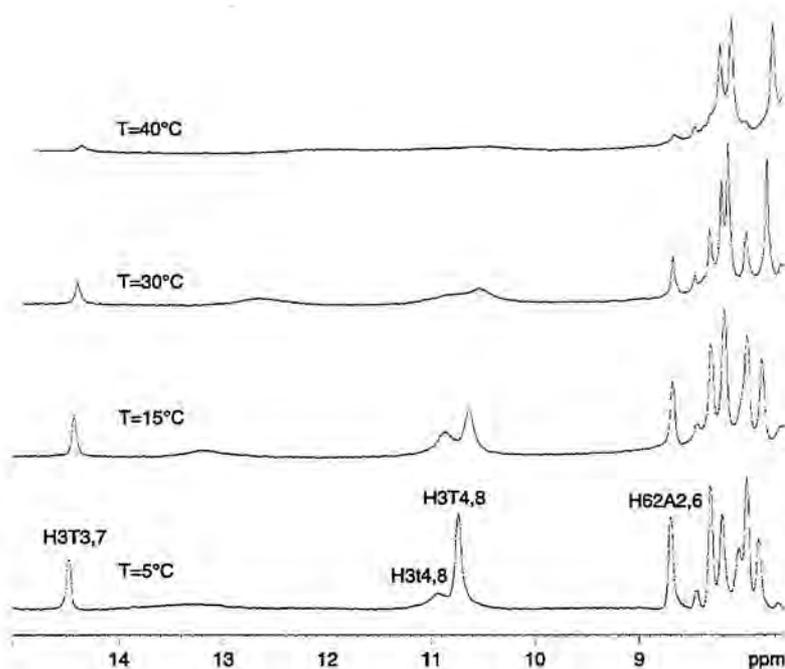
En el espectro NOESY en  $H_2O$  es posible asignar todas las señales imino y amino del dímero. Como se puede observar en la Figura 2.29, existen NOEs H1G2(6)-H4C3(7), H1G2(6)-H41C3(7), H1G2(6)-H5C3(7), que son indicativos de apareamiento tipo Watson-Crick. La existencia de picos de correlación H3T4(8)-H4C3(7), H3T4(8)-H41C3(7) y

$H3T4(8)-H6C3(7)$  pone de manifiesto que los residuos 4 y 8 se encuentran también interaccionando con las bases apiladas. De cara a la posterior determinación de la estructura resultan también muy importantes los NOEs que se observan entre protones intercambiables y protones no intercambiables.



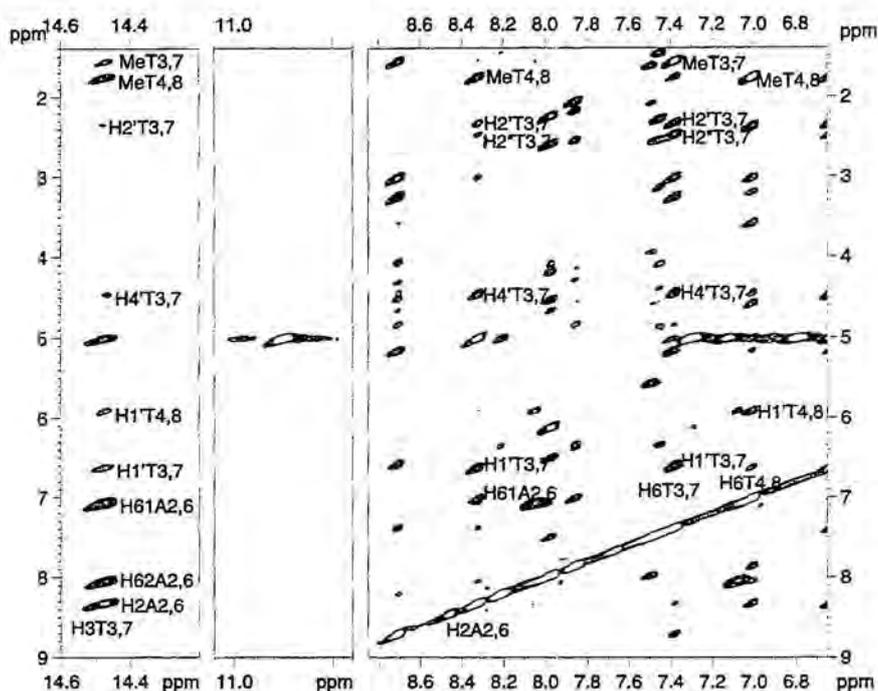
**Figura 2.29.-** Fragmento del espectro NOESY en  $H_2O$   $d\langle pTGCTcGCT \rangle$ , 2.4mM,  $T=5^\circ C$ ,  $pH=7$ . En este caso además de los picos de correlación  $H1G2(6)-H42C3(7)$ ,  $H1G2(6)-H41C3(7)$  y  $H1G2(6)-H5C3(7)$ , que permiten establecer que se trata de un apareamiento tipo Watson-Crick, se observan NOEs entre el protón imino de  $G2(6)$  y protones no intercambiables de los residuos  $C3(7)$  y  $T4(8)$ . La presencia de NOEs  $H3T4(8)-H42C3(7)$  y  $H3T4(8)-H41C3(7)$  ponen de manifiesto la implicación de los residuos  $T4(8)$  en el apilamiento de las bases.

En el caso de  $d\langle pCATTcATT \rangle$  también se observaron señales muy estrechas: una señal de imino apareado y una señal de imino no apareado, correspondientes a las dos timinas no equivalentes, aunque estas señales desaparecen a una temperatura inferior al del caso anterior. Tal como se observa en la Figura 2.30, a medida que se aumenta la temperatura aumenta la intensidad de las señales de protones imino correspondientes a la forma monomérica, pero las de la forma dimérica se mantienen en todo momento mucho más estrechas, indicando que la forma dimérica de  $d\langle pCATTcATT \rangle$  es también una estructura compacta con los pares de bases bien protegidos.



**Figura 2.30.-** Serie de espectros 1D en  $H_2O$  de  $d\langle pCATTcATT \rangle$ , 20 mM, pH=7. Se ha utilizado mayúsculas para designar las señales correspondientes a la forma dimerica y minúsculas para la forma monomérica.

En el espectro NOESY en  $H_2O$  de  $d\langle pCATTcATT \rangle$  se observan los NOEs H3T3(7)-H2A2(6), H3T3(7)-H62A2(6) y H3T3(7)-H61A2(6), lo que confirma el apareamiento tipo Watson-Crick entre las bases 2(6) y 3(7). En este caso no se observan NOEs para el protón imino no apareado T4(8), debido probablemente a la anchura de la señal. En cambio, sí se observan los NOEs H3T3(7)-MeT4(8) y H3T3(7)-H1 T4(8), lo que estaría indicando que, también en este caso, los residuos T4(8) interaccionan con los pares de bases. En la Figura 2.31 se detalla la asignación de los distintos picos de correlación observados.



**Figura 2.31.-** Fragmento del espectro NOESY en  $H_2O$   $d\langle pCATTcATT \rangle$ , 20 mM,  $T=5^\circ C$ ,  $pH=7$ . Se observan los picos de correlación  $H3T3(7)-H62A2(6)$ ,  $H3T3(7)-H61A2(6)$  y  $H3T3(7)-H2A2(6)$  indicativo de que se trata de un apareamiento tipo Watson-Crick. En este caso no se observan NOEs para el protón imino de  $T4(8)$ .

Los datos obtenidos de los espectros en  $H_2O$  para ambas estructuras diméricas, ponen de manifiesto la formación de cuatro pares de bases Watson-Crick intermoleculares, entre los residuos 2-3 (6-7). Las timinas  $T4(8)$  no están apareadas pero interaccionan con el resto de la molécula.

### 2.5.2.2.- Información de espectros en $D_2O$ .

La intensidad de los NOEs  $H1'-H6/H8$  intraresiduales, así como el hecho de que para todos los residuos el NOE  $H2'-H6/H8$  es más intenso que el  $H2''-H6/H8$ , confirma una disposición *anti* de todos los nucleótidos en ambos dímeros. En cuanto a la conformación del azúcar, la suma de constantes de acoplamiento indica una conformación mayoritariamente S para todas las ribosas.

Los espectros en  $D_2O$  confirman la información obtenida de los espectros en  $H_2O$ . Para ambos dímeros se observan NOEs secuenciales  $H2'/H2''-H6/H8$  y  $H1'-H6/H8$  que conectan los residuos 2 y 3 (y los respectivos 6 y 7), indicando que estas bases se encuentran apiladas. Los residuos 1 y 5, sin embargo, presentan muy pocos NOEs secuenciales por lo que parecen estar en una posición en la que interaccionan muy poco con el resto de la molécula. Por otra parte el comportamiento que presenta el protón imino de  $T1$  indica que este protón se halla totalmente expuesto al disolvente. Existen además un gran número de NOEs entre protones de los residuos 4 y 8 con protones de los residuos 2, 3, 6 y 7, que no se pueden asignar de forma inequívoca sin hacer suposiciones estructurales. Volveremos sobre ellos en el apartado 3.2.1.1. En relación a los residuos 4 y 8, hay que destacar también el desplazamiento a campo anormalmente alto de los protones  $H4'$ ,  $H5'$  y  $H5''$ .

La simetría observada en los espectros de las formas diméricas, juntamente con la información obtenida a partir de los espectros NOESY, así como la existencia de cuatro pares de bases intermoleculares, resultan únicamente compatibles con una disposición antiparalela de las dos moléculas que forman el dímero. En este sentido, es importante destacar que, a pesar de que la secuencia d<pTGCTCGCT> no es completamente repetitiva, la elevada simetría estructural del dímero produce una casi completa degeneración de las señales que ocupan posiciones equivalentes.

La determinación de la estructura tanto de las formas diméricas como de las formas monoméricas se describe en el capítulo siguiente.

---

*CAPÍTULO 3*

*DETERMINACIÓN ESTRUCTURAL DE  $d\langle p\text{CATTCAAT} \rangle$  y  $d\langle p\text{TGCTCGCT} \rangle$ .*

---

### 3.- DETERMINACIÓN ESTRUCTURAL DE d<pCATTTCATT> Y d<pTGCTCGCT>.

#### 3.0.- INTRODUCCIÓN Y OBJETIVOS.

El objetivo de este capítulo es describir la determinación de la estructura tridimensional de los octámeros cíclicos d<pCATTTCATT> y d<pTGCTCGCT>. En la primera parte del capítulo nos centraremos en la estructura de d<pCATTTCATT> obtenida por métodos cristalográficos para pasar, más adelante, a la determinación de la estructura en disolución de ambas moléculas.

Como se describió en el capítulo anterior, estos dos octámeros coexisten en dos formas diferentes bien estructuradas en disolución. A lo largo de este capítulo se determinan las estructuras, tanto de las formas monoméricas (tipo *dumbbell*), como de las diméricas (bi-loop).

#### 3.1.- ESTRUCTURA CRISTALOGRÁFICA DE d<pCATTTCATT>.

La primera incursión en el estudio estructural de moléculas cíclicas de DNA en nuestro grupo se realizó con el octámero cíclico de secuencia d<pCATTTCATT>. En los distintos estudios realizados utilizando moléculas cíclicas para el estudio de formación y estabilidad de estructuras *dumbbell* no se habían utilizado hasta entonces oligonucleótidos de menos de 10 residuos. Para la secuencia d<pCATTTCATT> resulta posible, sobre el papel, la formación de pares A·T intramoleculares que, en caso de confirmarse, darían lugar a una estructura tipo *dumbbell* estabilizada únicamente por dos pares de bases y cerrada por dos mini-loops de dinucleótidos.

Estos primeros estudios se realizaron mediante la técnica de difracción de RX. Hay que señalar que la única estructura cristalina descrita de una molécula cíclica de DNA hasta el momento era el dímero d<pAA><sup>1</sup>.

La determinación de la estructura cristalina de d<pCATTTCATT><sup>2</sup> se llevó a cabo en el Cambridge Crystallographic Data Centre (CCDC), por el grupo dirigido por Olga Kennard. Para la resolución de la estructura fue necesaria la aplicación de varios métodos. Al no resultar posible la resolución de la estructura utilizando métodos de *sustitución molecular*<sup>3,4</sup>, se recurrió al método de *sustitución isomórfica*<sup>5</sup>. Para ello fue necesaria la obtención del derivado bromado d<pCAT<sup>Br</sup>UCAT<sup>Br</sup>U>, donde las dos timinas que ocupan la primera posición de los mini-loops (T4 y T8) en la secuencia original se sustituyeron por dos 5-bromo-uracilos. Tampoco con este método se obtuvieron buenos resultados, al no resultar la estructura del derivado bromado isomórfica con la estructura nativa. La estructura pudo finalmente determinarse, con una resolución de 1.1Å, mediante la aplicación de *métodos directos*<sup>6</sup>. Se utilizaron datos obtenidos utilizando radiación sincrotrón ( $\lambda=0.934\text{Å}$ ).

<sup>1</sup> Frederick, C.A.; Coll, M.; van der Marel, G.A.; van Boom, J.H.; Wang, A. H.-J. *Biochemistry* **1988**, 27, 8350-8361.

<sup>2</sup> Salisbury, S.A.; Wilson, S.E.; Powell, H.R.; Kennard, O.; Lubini, P.; Sheldrick, G.M.; Escaja, N.; Alazzouzi, E.; Grandas, A.; Pedroso, E. *Proc. Natl. Acad. Sci. USA* **1997**, 94, 5515-5518.

<sup>3</sup> Rossmann, M.G.; Blow, D.M. *Acta Cryst.* **1962**, 15, 24-31.

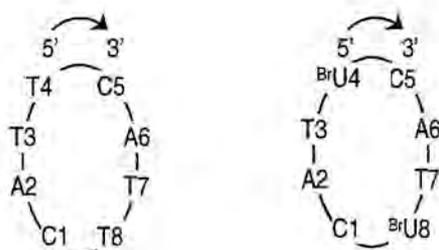
<sup>4</sup> Rossmann, M.G. *Acta Cryst.* **1990**, A-46, 73-82.

<sup>5</sup> Lattman, E. *Methods in Enzymology* **1985**, 115, 55-76.

<sup>6</sup> Stout, G.H.; Jensen, L.H. «X-ray Structure Determination. A Practical Guide.» **1989**, John Wiley & Sons.

La descripción detallada de todo el proceso de cristalización y resolución de la estructura, tanto de  $d\langle pCATTTCATT \rangle$  como del derivado bromado  $d\langle pCAT^{Br}UCAT^{Br}U \rangle$  se encuentra en la Tesis Doctoral de Sara E. Wilson<sup>7</sup>. Las coordenadas atómicas y las funciones de estructura están depositadas en el Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (referencia UDH052).

En la Figura 3.1 se muestra la numeración de residuos utilizada en esta Tesis, tanto para la secuencia original como para el derivado bromado. Esta numeración, aunque difiere de la que aparece en las referencias citadas, se corresponde con la que se ha utilizado en el presente trabajo en los estudios por RMN de distintas secuencias de octámeros cíclicos.



**Figura 3.1.-** Esquema de las secuencias  $d\langle pCATTTCATT \rangle$  y  $d\langle pCAT^{Br}UCAT^{Br}U \rangle$ , donde se indica la numeración utilizada en el texto

### 3.1.1.- RASGOS PRINCIPALES DE LA ESTRUCTURA CRISTALINA DE $d\langle pCATTTCATT \rangle$ .

El motivo por el que los métodos de *sustitución molecular* no permitieron la elucidación de la estructura resultó obvio una vez que ésta pudo ser resuelta: la estructura de  $d\langle pCATTTCATT \rangle$  constituye un nuevo motivo estructural de DNA, para el que los modelos en los que se basan los métodos de *sustitución molecular*, principalmente estructuras dúplex de DNA, resultaban poco adecuados.

La molécula no adopta una conformación tipo *dumbbell* con formación de puentes intramoleculares, que en un principio se pensaba que sería la estructura más probable, sino que cristaliza en forma dimérica. Todas las bases, excepto las citosinas se proyectan hacia fuera del esqueleto azúcar-fosfato en la misma dirección. Dos moléculas de octámero se disponen, en sentido antiparalelo, una frente a la otra permitiendo la formación de puentes intermoleculares.

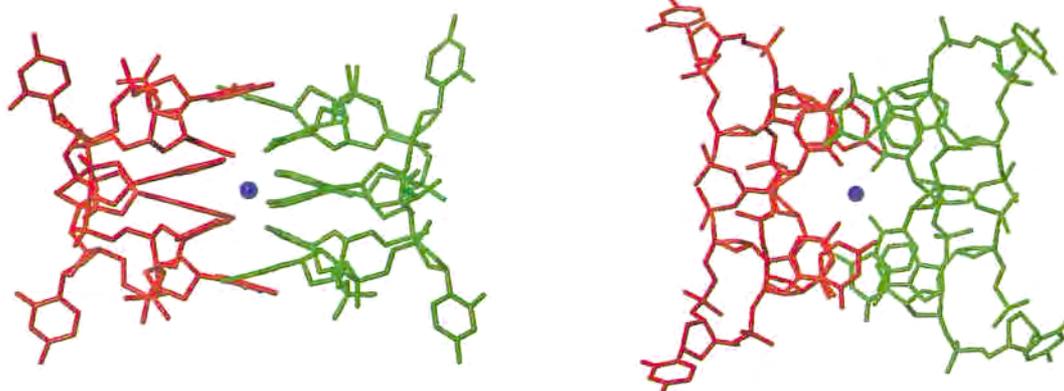
La unidad asimétrica, constituida por una molécula de dímero, presenta simetría no cristalográfica. Las dos moléculas cíclicas están relacionadas por dos ejes de segundo orden no cristalográficos ( $d_1, d_2$ ) y tienen conformaciones similares ( $rms=1.18\text{Å}$  o  $1.15\text{Å}$ , para las dos superposiciones químicamente equivalentes). Las dos partes equivalentes de cada molécula de octámero están relacionadas por un tercer eje  $d_3$  ( $rms=0.83\text{Å}$ ). Los residuos de citidina de cada una de las moléculas (relacionadas por el eje  $d_3$ ) presentan conformaciones similares, pero difieren de la conformación que presentan las dos citidinas de la otra molécula (relacionadas por  $d_1$  y  $d_2$ ).

El corazón de la estructura está formado por cuatro pares A·T intermoleculares sobre los cuales las timinas de los mini-loops, T4 y T8 (y las equivalentes en la otra subunidad), se

<sup>7</sup> Wilson, S.E. Tesis Doctoral, Cambridge University, 1997

encuentran apiladas. Los cuatro pares de bases A·T están dispuestos formando dos apilamientos con dos pares de bases cada uno. La disposición de los apilamientos es tal que los residuos pertenecientes a una misma cadena de octámero se enfrentan vía el surco menor, con una inclinación relativa de  $\sim 37.4^\circ$ .

Los pares de bases son planos mostrando un *propeller twist* y un *buckle* menores que los observados en otras estructuras cristalinas de DNA. La causa puede ser el hecho de que las timinas del *loop* estén también formando parte del apilamiento, ya que el *propeller twist* observado para pares A·T en B-DNA no permitiría que las timinas no apareadas, localizadas sobre los puentes de hidrógeno de los pares de bases A·T, presentaran interacciones de apilamiento favorables con los pares A·T. Entre los grupos metilo de estas timinas apiladas existe una interacción hidrofóbica favorable.



**Figura 3.2.-** Estructura RX desde dos perspectivas diferentes: A) eje *d*1 y B) eje *d*2. En el centro de la estructura se encuentra un ión de sodio coordinado.

La diferente conformación observada para los residuos de citidina se debe al empaquetamiento en el cristal. Las dos citosinas de una misma molécula, (C1, C5) están cerca del grupo Me de una timina (T4, T8) de otra unidad asimétrica adyacente. En otras palabras, el plano del anillo heterocíclico de C1 y C5 se encara con el grupo Me de T4 y T8 respectivamente. Los N4 de C1 y C5 forman puentes de hidrógeno ( $\sim 2.7\text{\AA}$ ) con los O1P de A2 y A6 respectivamente. Las citosinas de la otra molécula, C11 y C15, en cambio, están involucradas en el apilamiento entre dos unidades de dímero, dando lugar a un apilamiento continuo que involucra a cuatro unidades de dímero, Figura 3.3.

En la región de los pares de bases el apilamiento es similar al B-DNA. En este sentido resulta sorprendente, considerando las restricciones impuestas por el hecho de que las moléculas sean cíclicas, el gran parecido de los pares A·T del dímero con el B-DNA.

En el esqueleto azúcar-fosfato, constituido por dos secciones antiparalelas y un *hairpin* loop de dos residuos a cada banda, existen dos puntos de máxima proximidad:

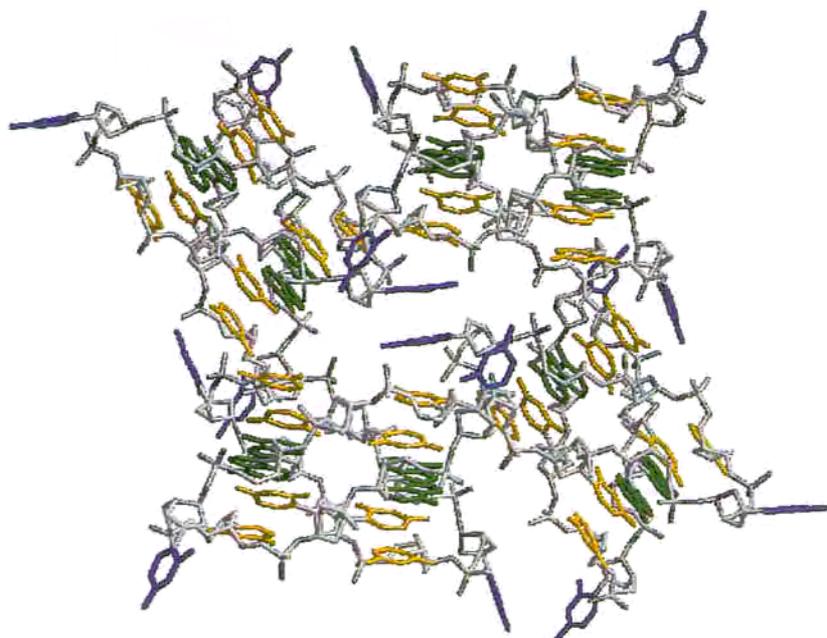
O2PT3-O2PT7: 4.8Å

O2PT13-O2PT17: 4.7Å

Los efectos repulsivos se ven contrarrestados por la presencia de varios iones bario en la estructura. El centro de la estructura dimerica está ocupado por un ión de sodio formando

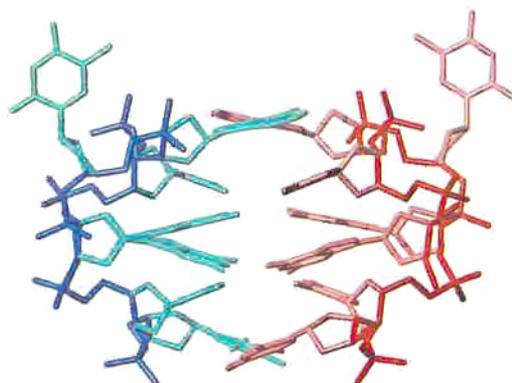
una coordinación octahédrica con los O2 de las cuatro timinas apareadas y dos moléculas de H<sub>2</sub>O. Este ión sodio reduce la repulsión electrostática generada por la proximidad de los cuatro grupos CO de las timinas.

La estructura cristalina del *d*<pCATTcATT>, que en principio podría pensarse que se trata de una curiosidad cristalográfica producto de fuerzas de empaquetamiento en el cristal, no constituye un ejemplo aislado. Anteriormente se había observado el mismo tipo de estructura cristalográfica para un heptámero lineal de secuencia *d*(GCATGCT)<sup>8</sup>, Figura 3.4. En este caso la estructura está estabilizada por cuatro puentes G·C Watson-Crick intermoleculares (O6G-H42C, H1G-N3C, H22G-O2C) con la formación adicional de cuatro puentes G·C intramoleculares (H21G-O2C). Los pares de bases forman también dos apilamientos con una inclinación relativa de ~30°. Las dos adeninas que ocupan la primera posición del loop se encuentran apiladas sobre los pares de bases G·C en uno de los extremos de la estructura. Entre ambas adeninas existe la formación de un par A·A (N62A-N7A), que ayuda a la estabilización global de la estructura. En el otro extremo de la estructura se encuentran apiladas las dos timinas terminales, entre cuyos grupos metilo existe, como en el caso del octámero cíclico, interacción hidrofóbica.



**Figura 3.3.-** Stacking continuo en el cristal, en el que se encuentran implicadas cuatro unidades de dímero.

<sup>8</sup> Leonard, G.A.; Zhang, S.; Peterson, M.R.; Harrop, S.J.; Helliwell, J.R.; Cruse, W.B.T., Langlois d'Estaintot, B.; Kennard, O.; Brown, T.; Hunter, W.N. *Structure* **1995**, 3, 335-340.



**Figura 3.4.-** Estructura del heptámero lineal  $d\langle p\text{GCATGCT}\rangle$  desde la perspectiva del eje  $d1$ .

A pesar de las diferencias existentes tanto a nivel de secuencia, como por el hecho de que una secuencia sea lineal y la otra cíclica, resulta sorprendente el elevado grado de coincidencia entre ambas estructuras. Las similitudes entre estas estructuras se discutirán más adelante, conjuntamente con las estructuras en disolución obtenidas para los dos octámeros cíclicos estudiados en el capítulo anterior.

Dada esta elevada similitud estructural se pensó que podría tratarse de un nuevo motivo estructural de carácter general. A esta nueva forma de estructura cuádruplex constituida por dos *hairpin-loops* dispuestos en sentido antiparalelo se la denominó estructura **bi-loop**.

### 3.1.2.- ESTRUCTURA DE RX DE $d\langle p\text{CAT}^{\text{Br}}\text{UCAT}^{\text{Br}}\text{U}\rangle$ .

Aunque el derivado bromado no pudo ser utilizado para la elucidación estructural de  $d\langle p\text{CATTCATT}\rangle$ , también se determinó su estructura cristalina. Este derivado cristaliza también en forma dimérica y presenta una estructura muy similar a la obtenida para  $d\langle p\text{CATTCATT}\rangle$  y  $d\langle \text{GCATGCT}\rangle$ , por lo que contribuye a reforzar la idea de que el motivo **bi-loop** es un motivo estructural bien definido.

La estructura de  $d\langle p\text{CAT}^{\text{Br}}\text{UCAT}^{\text{Br}}\text{U}\rangle$  presenta, sin embargo, un serie de diferencias con la estructura nativa. En la unidad asimétrica existen tres ejes de simetría de segundo orden, pero en este caso uno de ellos ( $d3$ ) sí tiene simetría cristalográfica. La simetría alrededor de los otros dos ejes ( $d1$  y  $d2$ ), en cambio, es menor que en el caso de la estructura nativa, observándose una menor similitud entre las dos moléculas cíclicas que forman el dímero.

Como en el caso de la estructura nativa, los pares A·T forman dos *apilamientos* con una inclinación relativa de aproximadamente  $32^\circ$ . La diferencia radica en los residuos que ocupan la primera posición del *loop* y que se encuentran también formando parte del apilamiento. En el caso de la estructura nativa los residuos T4, T8, T14 y T18 se encuentran apilados sobre los puentes de hidrógeno de los pares A·T. En el derivado bromado, en cambio, los residuos de  $^{\text{Br}}\text{U}$ , se encuentran apilados sobre la T del par A·T. Esta diferente posición de la  $^{\text{Br}}\text{U}$  permite un mayor *propeller twist* del par A·T subyacente.

El centro de la estructura dimérica del derivado bromado está ocupado también por un ión sodio, pero en este caso la geometría es ligeramente diferente. El ión  $Na^+$  está coordinado por los cuatro O2 de las timinas apareadas y una única molécula de  $H_2O$ .

La determinación estructural por métodos cristalográficos resulta complementaria a su principal alternativa: la RMN multidimensional. Las estructuras determinadas por cristalografía están sujetas a posibles deformaciones debido a fuerzas de empaquetamiento en el cristal. Aunque el hecho de que se haya observado el mismo motivo estructural en secuencias distintas apoya la hipótesis de que se trata de un nuevo motivo bien definido, resulta conveniente ver si se forma también en disolución. A continuación se describe la determinación de la estructura en disolución de las formas diméricas y monoméricas correspondientes a d<pCATT<math>CATT</math>> y d<pTGCT<math>CGCT</math>>.

### 3.2.- ESTRUCTURA EN DISOLUCIÓN DE d<pCATT<math>CATT</math>> Y d<pTGCT<math>CGCT</math>>.

A partir de la información estructural obtenida de los espectros de RMN correspondientes a las secuencias d<pCATT<math>CATT</math>> y d<pTGCT<math>CGCT</math>> se ha llevado a cabo la determinación de la estructura en disolución. Si bien el principal interés se centra en las formas diméricas, se ha determinado también la estructura de las formas monoméricas que, como se discutió en el apartado 2.5.1., se trata de estructuras tipo *dumbbell*.

#### 3.2.1.- OBTENCIÓN DE RESTRICCIONES DE DISTANCIAS.

##### 3.2.1.1.- Obtención de distancias en los dímeros.

El éxito en la determinación de estructuras de ácidos nucleicos mediante RMN depende enormemente del número de restricciones experimentales y de la calidad de las mismas. Para obtener estructuras altamente refinadas, es necesario asignar el mayor número posible de picos de correlación a pares concretos de protones y obtener distancias precisas entre éstos a partir de la intensidad del NOE. La conversión de los NOEs en restricciones de distancia resulta más complicada en el caso de especies diméricas. Esto es debido a la ambigüedad a la hora de diferenciar si un determinado pico de correlación corresponde a un contacto intra- o intermolecular. Aunque recientemente se han desarrollado metodologías para solucionar este problema con carácter general<sup>9</sup>, en casos favorables esta distinción se puede llevar a cabo de manera bastante directa haciendo asignaciones tentativas y realizando con éstas un cálculo de estructura preliminar. Examinando las violaciones residuales en las estructuras resultantes para cada restricción de distancia, es posible determinar las restricciones inconsistentes y probar otra asignación alternativa. Al cabo de un cierto número de intentos, se puede obtener un conjunto de distancias consistentes entre sí.

En nuestro caso, la situación resulta aún más compleja puesto que, por un lado la especie es dimérica y, por otro, la secuencia es repetitiva (o casi repetitiva en el caso de d<pTGCT<math>CGCT</math>>). Como se comprobó al asignar el espectro de RMN de las formas diméricas (secciones 2.2 y 2.3), todas las señales de los residuos 1 a 4 y 5 a 8 para d<pCATT<math>CATT</math>> y de los residuos 2 a 4 y 6 a 8 para d<pTGCT<math>CGCT</math>> están degeneradas. Una consecuencia de esta degeneración de señales es que, por cada par de protones que presentan un pico de correlación, son posibles cuatro restricciones de distancia diferentes. Por

<sup>9</sup> Nilges, M. *Proteins* 1993, 17, 297-309.

ejemplo, un contacto entre los residuos 2 y 4 puede tratarse también de un contacto 2-8 intra-subunidad o bien 2-12 o 2-16 inter-subunidad (se ha utilizado la numeración de residuos siguiente: 1-8 para una de las cadenas de octámero que forman el dímero y 9-16 para la otra cadena de octámero). Afortunadamente, al tratarse de moléculas de pequeño tamaño se pueden realizar cálculos rápidos de estructura utilizando diferentes combinaciones de restricciones. En muchos casos, como en el de correlaciones intraresiduales o de algunos NOEs entre residuos apareados, la asignación se pudo realizar mediante la simple inspección de un modelo preliminar de la molécula. Para el resto de picos, se realizó una primera asignación tentativa y las restricciones resultantes se utilizaron para calcular estructuras preliminares con el programa DYANA<sup>10</sup> (Véase la sección 3.2.2 de Materiales y Métodos para los detalles de estos cálculos). Cuando una restricción de distancia resultaba violada significativamente ( $>$  de 1Å) en todas las estructuras resultantes, se revisaba y cambiaba la asignación para ese pico de correlación, repitiéndose nuevamente el proceso de cálculo. Después de varios ciclos de asignación y cálculo de estructura se obtuvo un conjunto de restricciones consistente. Todos estos cálculos se llevaron a cabo con restricciones de distancias cualitativas. En este punto aún existían algunos picos de correlación ambiguos, para los que resultaban posibles dos o más asignaciones compatibles con la estructura tridimensional. Estos picos de correlación no se tuvieron en cuenta en los cálculos posteriores.

Como resultado de los cálculos con el programa DYANA, se obtuvieron una serie de estructuras preliminares consistentes con un conjunto de restricciones de distancias cualitativas asignadas de forma inequívoca. Con objeto de refinar estas estructuras, se procedió a obtener un conjunto de restricciones precisas mediante el uso de métodos de matriz completa de relajación<sup>11</sup>. Estos cálculos se llevaron a cabo con el programa MARDIGRAS<sup>12-13</sup>. Los detalles de estos cálculos de distancias precisas se describen en la sección 3.2.1 de Materiales y Métodos. Solamente se utilizaron las intensidades experimentales correspondientes a picos NOE asignados de forma inequívoca. Para estimar adecuadamente los errores en las distancias resultantes, se repitieron los cálculos con MARDIGRAS partiendo de intensidades obtenidas a partir de espectros NOESY adquiridos con diferentes tiempos de mezcla, así como utilizando varios modelos iniciales y diversos tiempos de correlación molecular. Con los resultados de todos estos cálculos, se llevó a cabo un análisis de errores del que se obtuvieron límites superiores e inferiores para las restricciones de distancia finales. No en todos los casos se pudieron obtener distancias precisas. En el caso de picos muy débiles o con fuerte solapamiento, así como en aquellos NOEs en los que están involucrados protones intercambiables, se utilizaron restricciones de distancias cualitativas. También se utilizaron distancias cualitativas en aquellas restricciones donde están involucrados protones H5'/H5" no asignados estereoespecíficamente, así como con los protones H2'/H2" de C3(7) en  $d\langle p\text{TGCTCGCT}\rangle$ , que están degenerados. En estos casos se añadió una corrección de 0.9Å a la cota superior de distancias.

El hecho de que la estructura que estamos estudiando sea simétrica complica la asignación de los picos de correlación, pero tiene la ventaja de que una vez asignados éstos, el

<sup>10</sup> Güntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, 273, 283-298.

<sup>11</sup> James, T.L. *Curr. Opin. Struct. Biol.* **1991**, 1, 1042.

<sup>12</sup> Borgias, B.A.; James, T.L. *J. Magn. Reson.* **1988**, 79, 493-512.

<sup>13</sup> Borgias, B.A.; James, T.L. *J. Magn. Reson.* **1990**, 87, 475-487.

número de restricciones que se obtienen es mucho mayor. Por ejemplo, una vez asignado un determinado NOE a un contacto entre los residuos 2 y 4, inmediatamente obtenemos cuatro restricciones de distancias: entre los residuos 2-4, 6-8 y los correspondientes a la otra subunidad, 10-12 y 14-16. Esto hace que el número total de restricciones de distancias sea inusualmente grande, dado el tamaño de las moléculas estudiadas. Para el caso de  $d\langle p\text{CATTCA}\rangle$ , se obtuvieron un total de 392 restricciones de distancia experimentales, incluyendo tanto protones intercambiables como no intercambiables. Este conjunto de restricciones está constituido por 220 restricciones intraresiduales, 102 secuenciales y 70 interesiduales entre protones situados a más de un residuo de distancia en la secuencia. De estas 392 restricciones de distancias, 340 son intra-subunidad y 52 corresponden a restricciones entre las dos subunidades. Para el caso de  $d\langle p\text{TGCTCGCT}\rangle$ , se obtuvieron 486 restricciones, incluyendo 252 intraresiduales, 144 secuenciales y 90 interesiduales no secuenciales. De ellas, 430 son intra-subunidad y 56 inter-subunidad. La distribución de estas restricciones a lo largo de la secuencia, para las dos secuencias estudiadas se muestran en las Figuras 3.5 y 3.6.

Además de las restricciones de distancias derivadas del efecto NOE, se introdujeron restricciones angulares para ciertos ángulos de torsión de los anillos de desoxiribosa y para los ángulos  $\gamma$  del esqueleto de los residuos 2 y 4. Tanto en los cálculos preliminares como en el refinamiento posterior, en los residuos apareados se incluyeron restricciones de distancia y angulares correspondientes a los puentes de hidrógeno característicos de apareamiento tipo Watson-Crick.

### 3.2.1.2.- Obtención de Distancias en los Monómeros.

En el caso de las formas monoméricas no ha sido posible obtener un conjunto de restricciones experimentales tan completo y preciso como para las formas diméricas. Por un lado, en el caso de  $d\langle p\text{TGCTCGCT}\rangle$ , la estructura monomérica es minoritaria cuando se trabaja en las condiciones óptimas para obtener los espectros bidimensionales. Esto hace que se observen menos NOEs y que éstos sean débiles. El problema en el caso de  $d\langle p\text{CATTCA}\rangle$  es diferente. Aunque podemos obtener espectros bidimensionales de gran calidad de la forma monomérica, la anchura de las señales de protones intercambiables impide determinar de forma inequívoca el tipo de apareamiento. Además, tanto la anchura de estas señales, como los NOEs observados entre protones no intercambiables, parecen indicar la existencia de un equilibrio conformacional en el propio monómero.

Estas dificultades experimentales para obtener restricciones van a hacer imposible la obtención de estructuras altamente refinadas de las formas monoméricas de estas moléculas. No obstante se ha extraído un conjunto de restricciones de distancias basado en una interpretación cualitativa de los NOEs observados. Estas restricciones nos permitirán calcular un modelo aproximado de estas moléculas que se discutirá en el apartado 3.2.2.2.

En el caso de  $d\langle p\text{TGCTCGCT}\rangle$ , se obtuvieron un total de 75 restricciones de distancias informativas (no se incluyen distancias entre protones del mismo azúcar salvo para H1'-H4' y H2"-H4'), de las cuales 45 son intraresiduales y 30 secuenciales. En el caso de  $d\langle p\text{CATTCA}\rangle$ , el número total de restricciones de distancia experimentales fue de 92, siendo 52 de ellas intraresiduales y 40 secuenciales. Además de estas restricciones experimentales, se incluyeron restricciones de puente de hidrógeno, típicas de apareamientos

Watson-Crick. En el caso de d<pCATT<math>CATT</math>>, la inclusión de estas restricciones no está del todo justificada experimentalmente y la estructura resultante debe considerarse como un modelo hipotético, como se discutirá en el apartado correspondiente.

Además de estas restricciones de distancias, se incluyeron restricciones angulares correspondientes a la conformación de las desoxiribosas tipo S.

### 3.2.2.- CÁLCULO DE ESTRUCTURAS.

#### 3.2.2.1.- Obtención de las Estructuras en los Dímeros.

Como se comentó en el apartado anterior, las estructuras preliminares se calcularon con el programa DYANA 1.4. Dicho programa trabaja en el espacio de ángulos de torsión y, por lo tanto, las posiciones relativas de cada átomo están definidas por los valores de los ángulos diedros a lo largo de la cadena polinucleotídica. Esta aproximación es muy útil a la hora de acelerar el proceso de cálculo, pero tiene ciertas implicaciones que deben tenerse en cuenta al tratar moléculas diméricas. Para definir la posición relativa de cada subunidad, es necesario introducir una serie de residuos, denominados "linkers". Estos "linkers" no interaccionan con el resto de la molécula, pero permiten unir las dos subunidades en una misma cadena. La posición de la segunda subunidad queda definida por los ángulos de torsión de los "linkers". Es necesario introducir suficiente número de tales "linkers" para asegurarnos de que ambas subunidades puedan estar en cualquier posición relativa.

Otra consecuencia de trabajar en el espacio de ángulos de torsión es que no se pueden definir moléculas con cadenas cíclicas. Para solventar este problema, es necesario introducir restricciones de distancia adicionales que mantengan próximos los extremo 3' y 5' de cada subunidad. También es necesario introducir restricciones para "cerrar" los anillos de desoxiribosa.

Una vez definida la topología de la molécula y con las restricciones de distancia adicionales para asegurarnos una estructura cíclica, se procedió al cálculo de estructuras preliminares. Los cálculos se realizaron utilizando restricciones de distancias cualitativas y restricciones de puentes de hidrógeno. No se utilizaron restricciones angulares. En todos los casos, los cálculos se comenzaron a partir de estructuras generadas al azar. El protocolo utilizado en estos cálculos se detalla en la sección 3.2.2 de Materiales y Métodos.

Las 10 estructuras resultantes con el valor más bajo de la función blanco se utilizaron como modelos iniciales para un refinamiento de dinámica molecular restringida en el espacio de coordenadas cartesianas. Este segundo cálculo se llevó a cabo con el módulo SANDER del paquete de dinámica molecular AMBER, versión 4.1<sup>14</sup>. Como se indica en el apartado anterior, este refinamiento se realizó utilizando restricciones de distancias precisas y con las asignaciones obtenidas con ayuda de los cálculos preliminares con DYANA. Puesto que este cálculo se realizó en el espacio de coordenadas cartesianas, no es necesario introducir ningún "linker". Asimismo, el hecho de que la molécula sea cíclica queda directamente asegurado al definir la topología molecular sin necesidad de introducir ninguna restricción adicional.

---

<sup>14</sup> Pearlman, D.A.; Case, D.A.; Caldwell, J.C.; Seibel, G.L.; Singh, U.C.; Weiner, P.; Kollman, P.A. AMBER 4.0 1991, University of California, San Francisco.

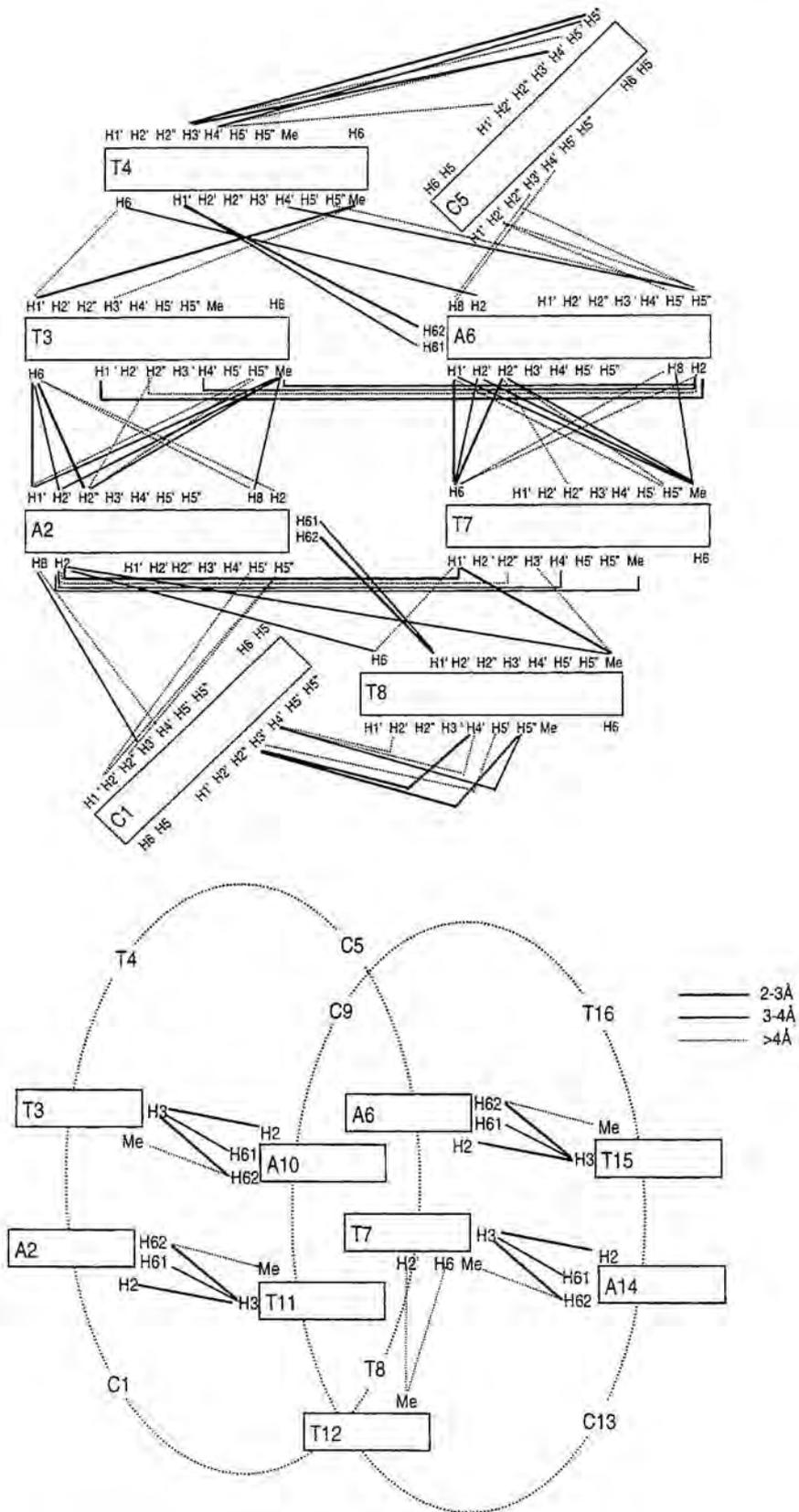


Figura 3.5.- Restricciones intra- e intersubunidad para  $d\langle pCATTTCATT \rangle$ .

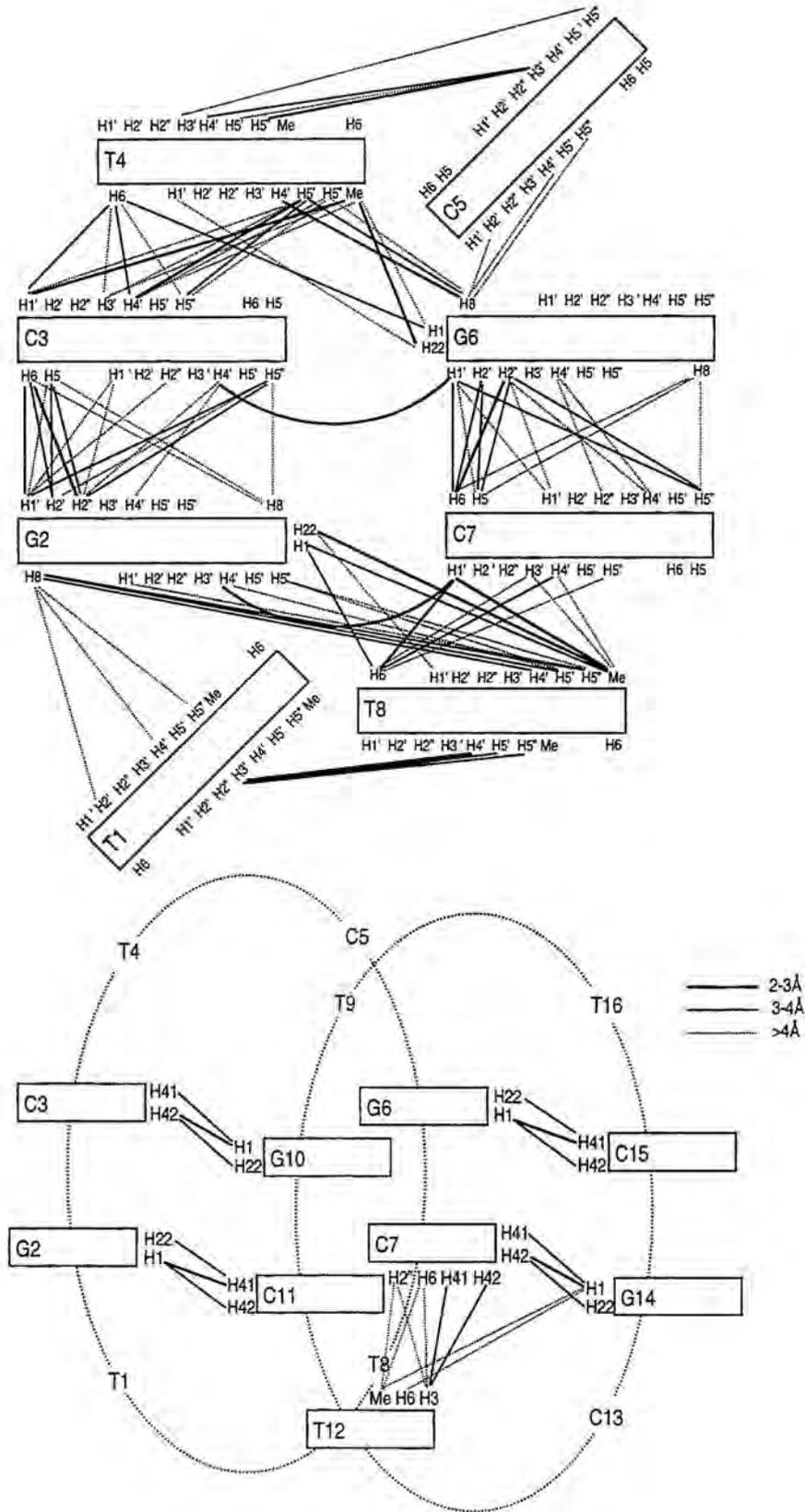


Figura 3.6.- Restricciones intra- e intersubunidad para  $d\langle pTGCTCGCT \rangle$ .

Las 10 trayectorias de dinámica molecular se corrieron durante 40 ps siguiendo un protocolo de calentamiento y enfriamiento *annealing*, detallado en la sección 3.2.2 de Materiales y Métodos. Como se muestra en la Tablas 3.1 y 3.2, las estructuras resultantes presentan pocas violaciones de distancias y factores R pequeños, lo cual indica que el grado de ajuste de la información experimental es satisfactorio. Asimismo, la parte no experimental de las energías finales presentan valores bajos en todas las estructuras y similares entre sí. Esto indica que el potencial de AMBER es compatible con las restricciones de distancias y que no es necesario inducir distorsiones importantes en la estructura para satisfacer todas las restricciones experimentales simultáneamente.

**Tabla 3.1.-** Valores de energía y violaciones de distancia finales para las diez estructuras resultantes de la forma dimérica de  $d\langle p\text{CATTATT} \rangle$ .

Estructura	$E_{\text{tot}}$ kcal/mol	$E_{\text{vdw}}$ kcal/mol	$E_{\text{el}}$ kcal/mol	$E_{\text{exp}}$ kcal/mol	$\Sigma_{\text{viol}}$ Å	Viol. máx. Å	Factor R <sup>§</sup> x100
1	-2193	110	-2651	76	16.6	0.55	6.91
2	-2220	80	-2632	44	13.9	0.32	6.78
3	-2241	107	-2713	81	16.9	0.49	6.86
4	-2231	84	-2646	50	14.2	0.49	7.18
5	-2253	84	-2665	49	14.4	0.47	7.07
6	-2221	106	-2683	79	17.2	0.56	6.88
7	-2249	90	-2667	46	14.7	0.33	6.70
8	-2222	91	-2652	66	15.0	0.54	6.87
9	-2249	97	-2681	66	14.9	0.51	6.81
10	-2268	105	-2729	85	16.7	0.54	7.21
Media	-2235	87	-2672	64	15.5	0.48	6.93

<sup>§</sup> Factor R dependiente de la sexta potencia de la intensidad del NOE<sup>15,16</sup>.

**Tabla 3.2.-** Valores de energía y violaciones de distancia finales para las diez estructuras resultantes de la forma dimérica de  $d\langle p\text{TGCTCGCT} \rangle$ .

Estructura	$E_{\text{tot}}$ kcal/mol	$E_{\text{vdw}}$ kcal/mol	$E_{\text{el}}$ kcal/mol	$E_{\text{exp}}$ kcal/mol	$\Sigma_{\text{viol}}$ Å	Viol. máx. Å	Factor R x100
1	-2481	76	-2877	77	11.5	0.36	7.93
2	-2490	75	-2879	80	12.0	0.30	7.72
3	-2484	76	-2876	78	12.3	0.33	6.96
4	-2477	76	-2873	109	13.1	0.62	7.23
5	-2461	65	-2838	98	12.9	0.49	7.15
6	-2391	70	-2789	92	12.9	0.39	7.88
7	-2436	80	-2836	84	12.7	0.33	7.50
8	-2468	90	-2898	117	14.7	0.41	6.95
9	-2462	70	-2853	82	12.5	0.33	6.79
10	-2467	69	-2855	89	12.6	0.34	7.32
Media	-2462	75	-2857	91	12.7	0.39	7.24

<sup>§</sup> Factor R dependiente de la sexta potencia de la intensidad del NOE<sup>15,16</sup>.

<sup>15</sup> González, C.; Rullman, J.A.C.; Bonvin, A.M.J.J.; Boelens, R.; Kaptein, R. *J. Magn. Reson.* **1991**, 91, 659-664.

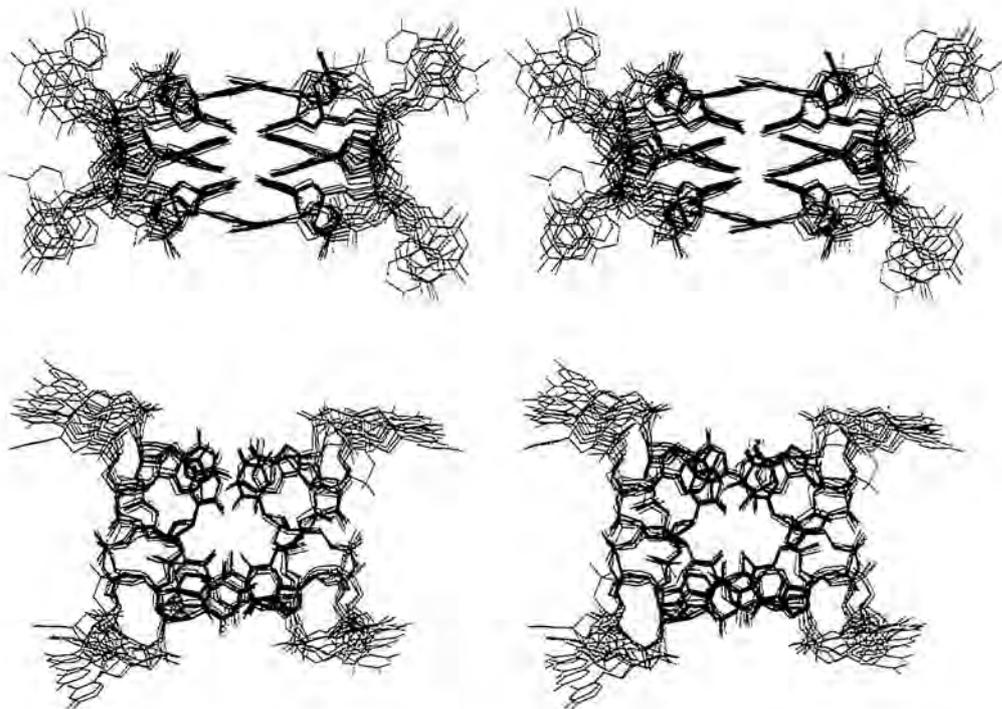
<sup>16</sup> Thomas, P.D.; Basus, V.J.; James, T.L. *Proc. Natl. Acad. Sci.* **1991**, 88, 1237-1241.

En las Figuras 3.7 y 3.8 se muestra la superposición de las 10 estructuras finales de los dos dímeros. Gracias a la gran cantidad de restricciones experimentales y a la precisión con la que se han obtenido las mismas, las estructuras resultantes están muy bien definidas. Sólo los residuos en posiciones 1 y 5 de cada subunidad aparecen desordenados. Como es habitual en estructuras de ácidos nucleicos obtenidas mediante RMN, las bases están mejor definidas que la región del esqueleto. Los valores medios de RMSD para las bases de los residuos bien definidos (todos salvo 1,5,9 y 13 están en ambos casos por debajo de los 0.4 Å). Estos valores aumentan, lógicamente, si incluimos los átomos de los azúcares y del esqueleto de estas regiones, como se observa en la Tabla 3.3.

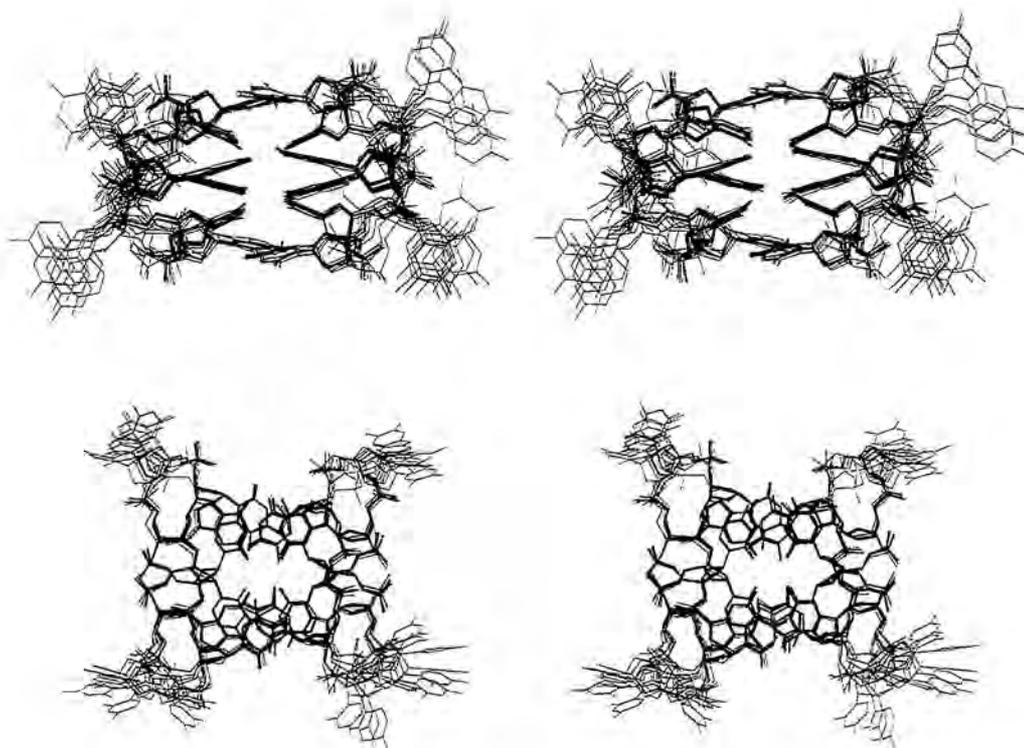
**Tabla 3.3.-** Valor medio de la RMSD entre las diez estructuras diméricas finales<sup>§</sup>.

Átomos considerados	<i>d</i> <pCATTcATT>	<i>d</i> <pTGCTcGCT>
Bases, excepto 1,5,9 y 13.	0.33 ± 0.07	0.25 ± 0.06
Todos los átomos, excepto residuos 1, 5, 9 y 13 <sup>§</sup> .	0.90 ± 0.22	0.50 ± 0.08
Esqueleto azúcar-fosfato.	1.05 ± 0.2	0.71 ± 0.11
Todas las bases.	1.41 ± 0.3	2.7 ± 0.8
Todos los átomos.	1.33 ± 0.3	2.0 ± 1.2

<sup>§</sup> En ningún caso se incluyen los átomos de hidrógeno en el cálculo de la RMSD.



**Figura 3.7.-** Vista estereoscópica, desde los dos ejes que relacionan las dos subunidades del dímero, de la superposición de las diez estructuras resultantes para *d*<pTGCTcGCT>.



*Figura 3.8.- Vista estereoscópica, desde los dos ejes que relacionan las dos subunidades del dímero, de la superposición de las diez estructuras resultantes para  $d\langle pATTcatt \rangle$ .*

### 3.2.2.2.- Obtención de las estructuras de los monómeros.

Las estructuras de las formas monoméricas de  $d\langle pCATTcatt \rangle$  y  $d\langle pTGCTcgct \rangle$  se calcularon de manera enteramente análoga a las de las formas diméricas, descritas en el apartado anterior. En ambos casos las estructuras resultantes están mucho peor definidas que en los correspondientes dímeros. Esto es debido al menor número de restricciones de distancias obtenido y a que éstas no pudieron calcularse con el mismo nivel de precisión, ya que la baja calidad de los espectros y el alto grado de flexibilidad de las moléculas no permitieron llevar a cabo los cálculos de matriz completa de relajación. En las tablas 6 y 7 se muestran los valores de energías y violaciones residuales de las 10 estructuras calculadas para cada secuencia. En ambos casos, las energías finales son pequeñas y similares en todas las estructuras. Las violaciones residuales son mucho menores que en el caso de los dímeros. Esto no es sorprendente si tenemos en cuenta que tenemos muchas menos restricciones que satisfacer en el proceso de cálculo. Hay que señalar que en el caso de  $d\langle pCATTcatt \rangle$ , el cálculo de estructura se llevó a cabo sin incluir restricciones H1'-H8 intraresiduales de las adeninas, cuyo valor estimado es de 2.5Å. La inclusión de estas restricciones hace que las estructuras resultantes del cálculo aparezcan muy distorsionadas y con valores de energías y violaciones mucho mayores.

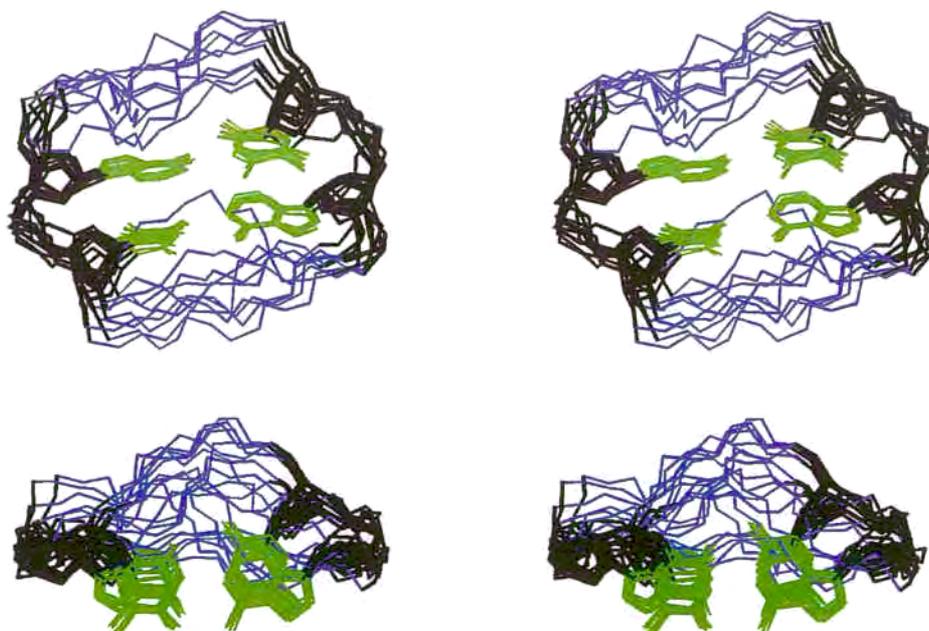
**Tabla 3.4.-** Valores de energía y violaciones de distancia finales para las diez estructuras resultantes de la forma monomérica de  $d\langle pCATTcATT \rangle$ .

Estructura	$E_{tot}$ kcal/mol	$E_{vdw}$ kcal/mol	$E_{el}$ kcal/mol	$E_{exp}$ kcal/mol	$\Sigma_{viol}$ Å	Viol. máx. Å
1	-968	-29	-1096	10	2.2	0.17
2	-927	-11	-1068	7	1.7	0.18
3	-961	-11	-1106	9	2.1	0.15
4	-957	-13	-1096	8	2.1	0.14
5	-972	-67	-1092	10	2.1	0.20
6	-996	-12	-1139	10	2.5	0.12
7	-995	-23	-1122	12	2.4	0.17
8	-962	-12	-1101	9	2.1	0.15
9	-1024	-30	-1149	11	2.5	0.15
10	-1020	-58	-1160	12	2.4	0.17
Media	-978	-27	-1113	10	2.2	0.16

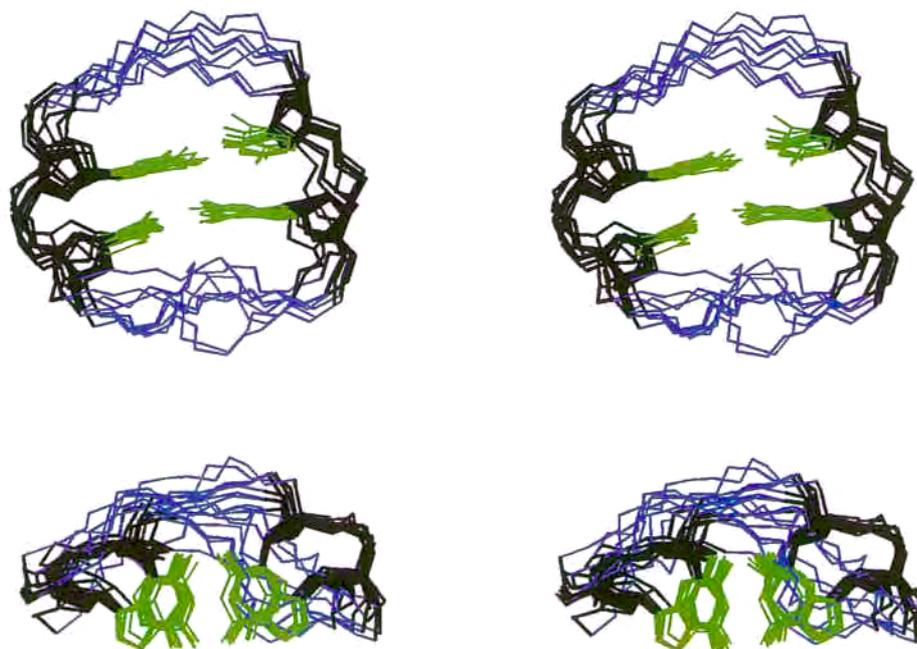
**Tabla 3.5.-** Valores de energía y violaciones de distancia finales para las diez estructuras resultantes de la forma monomérica de  $d\langle pTGCTCGCT \rangle$ .

Estructura	$E_{tot}$ kcal/mol	$E_{vdw}$ kcal/mol	$E_{el}$ kcal/mol	$E_{exp}$ kcal/mol	$\Sigma_{viol}$ Å	Viol. máx. Å
1	-592	-43	-701	1	0.5	0.04
2	-617	-59	-716	1	0.6	0.08
3	-572	-56	-693	2	0.9	0.12
4	-504	-44	-691	3	0.9	0.14
5	-557	-52	-689	2	1.1	0.11
6	-598	-56	-700	1	0.7	0.10
7	-552	-39	-690	1	0.7	0.11
8	-576	-46	-693	1	0.7	0.11
9	-599	-53	-699	1	0.6	0.08
10	-593	-53	-699	1	0.6	0.09
Media	-576	-50	-697	1	0.7	0.10

La superposición de las 10 estructuras finales se muestra en las Figuras 3.9 y 3.10. Por claridad no se han incluido las bases de los loops en la figura, ya que aparecen completamente desordenados (RMSD de casi 3Å). En la Tabla 3.6 aparecen los valores de RMSD para diferentes regiones de las moléculas. Sólo las bases de los residuos apareados están relativamente bien definidas. Aunque se observaron algunos NOEs entre los residuos apareados y las bases de los loops, como se comentó en el apartado 3.3.1.2, éstos son débiles y no parecen ser suficientes para restringir la conformación de estas regiones.



**Figura 3.9.-** Superposición de las 10 estructuras finales de la forma monomérica de  $d\langle pCATTCATT \rangle$ . Por claridad, sólo se muestra el esqueleto de la molécula y los cuatro nucleótidos apareados.



**Figura 3.10.-** Superposición de las 10 estructuras finales de la forma monomérica de  $d\langle pTGCTCGCT \rangle$ . Por claridad, sólo se muestra el esqueleto de la molécula y los cuatro nucleótidos apareados.

**Tabla 3.6.-** Valores promedio de las RMSD ( $\text{\AA}$ ) entre las diez estructuras monoméricas finales<sup>§</sup>.

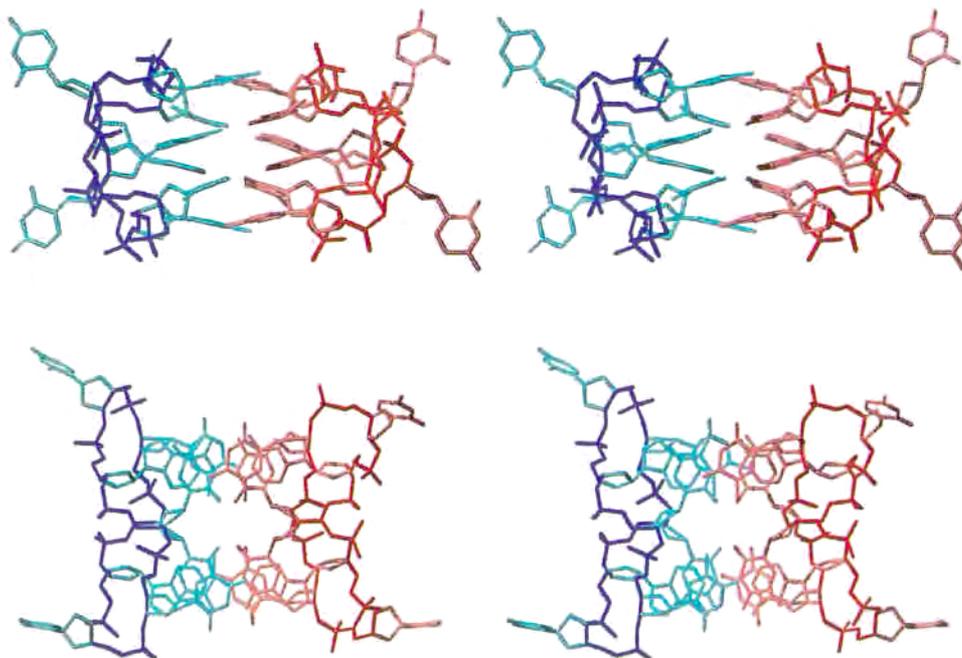
Átomos considerados	$d\langle pCATTCAAT \rangle$	$d\langle pTGCTCGCT \rangle$
Bases apareadas (2,3,6, y 7).	$0.50 \pm 0.16$	$0.48 \pm 0.16$
Nucleótidos apareados (2,3,6, y 7). <sup>§</sup>	$1.13 \pm 0.20$	$0.98 \pm 0.18$
Esqueleto azucar-fosfato.	$1.83 \pm 0.46$	$1.38 \pm 0.31$
Todas las bases.	$2.80 \pm 0.70$	$2.85 \pm 1.2$
Todos los átomos.	$2.67 \pm 0.6$	$2.65 \pm 1.0$

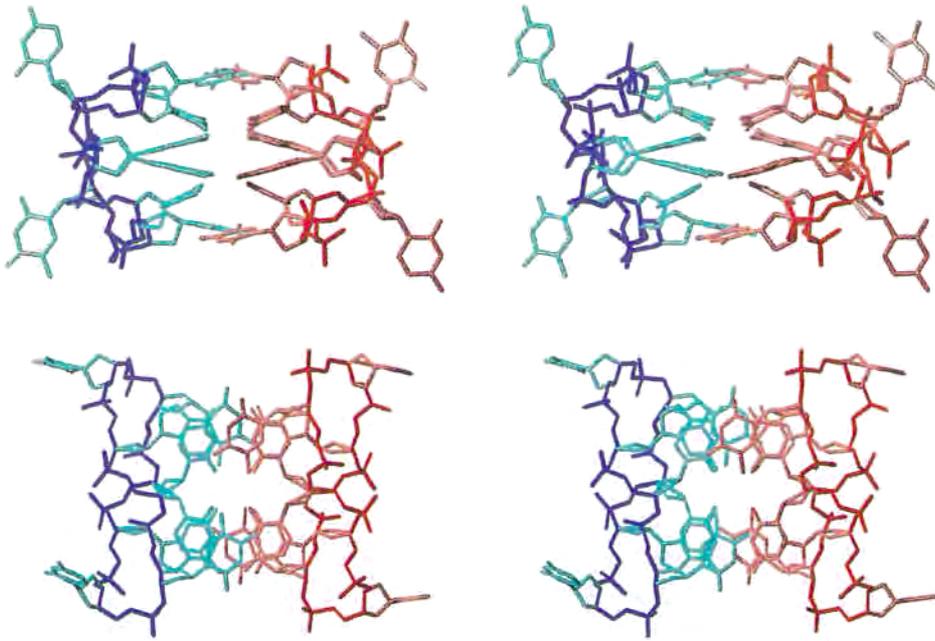
<sup>§</sup> En ningún caso se incluyen los átomos de hidrógeno en el cálculo de la RMSD.

### 3.2.3.- ANÁLISIS DE LAS ESTRUCTURAS.

#### 3.2.3.1.- Estructuras diméricas.

Las estructuras diméricas obtenidas para las dos secuencias estudiadas presentan rasgos estructurales muy similares. En las Figuras 3.11 y 3.12 se muestran las estructuras promedio obtenidas, para cada dímero, a partir de las diez estructuras resultantes. Resulta sorprendente no sólo la enorme similitud entre ambas estructuras, teniendo en cuenta que las secuencias son muy diferentes, sino que se trata de la misma estructura observada en estado sólido: el motivo bi-loop.

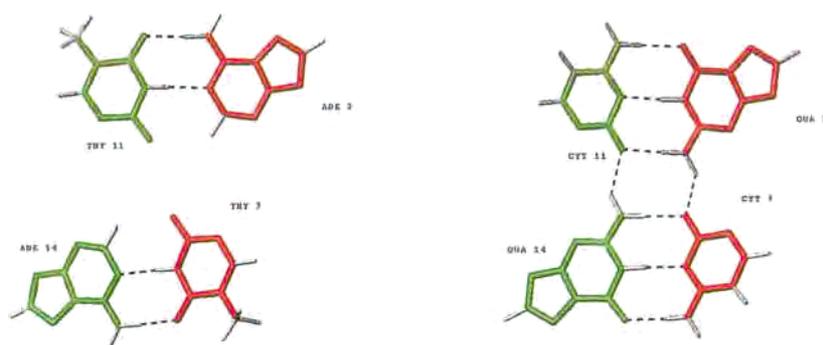
**Figura 3.11.-** Vista estereoscópica de la estructura promedio para el dímero de  $d\langle pCATTCAAT \rangle$ .



**Figura 3.12.-** Vista estereoscópica de la estructura promedio para el dímero de *d*<pTGCTCGCT>.

Para ambas secuencias, las estructuras diméricas están constituidas por dos moléculas de octámero cíclico, dispuestas en sentido antiparalelo, entre las que se forman cuatro pares de bases intermoleculares. Estos pares de bases se encuentran dispuestos formando dos apilamientos, con una inclinación relativa de 37° para *d*<pCATTATT> y de 40° en el caso de *d*<pTGCTCGCT>, que interaccionan via el surco menor formando una tétrada. En el caso de *d*<pTGCTCGCT> esta interacción entre dinucleótidos apareados está además estabilizada por la formación de puentes de hidrógeno H21G-O2C intramoleculares (entre G2-C7, G6-C3, G10-C15 y G14-C11), lo que supone la formación de un total de 16 puentes de hidrógeno. Estos puentes intramoleculares extra que se forman entre guaninas y citosinas ya se habían observado en la estructura dímica del heptámero lineal<sup>8</sup> y en el caso de una tétrada G:C:G:C<sup>17</sup>. En el caso de *d*<pCATTATT>, al no ser posible la formación adicional de estos puentes intramoleculares, la estructura está estabilizada únicamente por los 8 puentes de hidrógeno intermoleculares, Figura 3.13.

<sup>17</sup> Kettani, A.; Kumar, R.A.; Patel, D.J. *J. Mol. Biol.* **1995**, 254, 368.



**Figura 3.13.-** A) Tétrada A:T:A:T con 4 puentes de hidrógeno y B) tétrada G:C:G:C con 8 puentes de hidrógeno.

Las bases que ocupan la primera posición del *loop*, timinas en los dos casos, se encuentran apiladas sobre los pares de bases haciendo de "tapa" en los extremos 3' de cada apilamiento. Este hecho explica el desplazamiento a campos anormalmente altos observado para los protones H4' y H5'/H5'' de los residuos 4 y 8, ya que se encuentran localizados sobre la zona de influencia de corriente de anillo de las purinas apareadas. Entre estas timinas existe una interacción hidrofóbica favorable entre los grupos metilo (T4-T16 y T8-T12), que contribuye a la estabilización global de la estructura, Figura 3.14.



**Figura 3.14.-** Vista estereoscópica de la parte superior de la estructura dimérica de  $d\langle pTGCTCGCT \rangle$  donde se muestra la disposición de las timinas 8 y 12 haciendo de "tapa" sobre el apilamiento formado por las bases C11-G2 y G14-C7, respectivamente.

Las bases que ocupan la segunda posición del *loop* (residuos 1 y 5 en la secuencia) se disponen hacia el exterior de la estructura y se encuentran totalmente expuestas al disolvente, sin presentar ningún tipo de interacción con el resto de la estructura.

Estas estructuras diméricas constituyen un nuevo tipo de motivo tetrada. No obstante, a diferencia de lo que ocurre en otras estructuras tipo cuádruplex, entre las bases complementarias se forman pares tipo Watson Crick, al igual que en el DNA de doble cadena. En la parte central de la estructura, donde se produce el apareamiento de las bases, los parámetros de hélice indican una gran similitud con el B-DNA. En la Tabla 3.7 se muestran los parámetros de hélice locales, correspondientes a residuos apilados, en ambas estructuras diméricas.

Tabla 3.7.- Parámetros de hélice locales "inter-base" para cada subunidad.

Residuo	Shift Dx	Slide Dy	Rise Dz	Tilt $\tau$	Roll $\rho$	Twist $\Omega$	
d<pCATTcATT>	A2/T3	0.06	-1.46	3.17	0.57	-0.46	25.56
	T3/T4	-10.83	-1.36	4.68	-29.89	-2.32	-40.89
	A6/T7	-0.11	-1.40	3.15	0.82	-0.27	24.51
	T7/T8	-10.82	-1.40	4.69	-30.21	-2.01	-40.33
	T12/T11	10.80	-1.39	4.75	30.34	-3.66	-40.92
	T11A10	0.12	-1.47	3.17	-0.64	-1.78	24.49
	T15/T16	10.84	-1.28	4.70	30.37	-1.00	-40.83
	T14/A15	-0.07	-1.37	3.11	-1.29	0.11	25.81
d<pTGCTcGCT>	G2/C3	0.25	-0.42	2.81	9.81	-4.49	26.00
	C3/T4	-10.16	-1.94	4.73	-38.38	7.83	-35.30
	G6/C7	0.31	-0.34	2.74	10.87	-2.96	25.67
	C7/T8	-9.72	-1.96	5.13	-41.13	12.49	-31.45
	T12/C11	10.11	-1.91	4.79	38.17	7.11	-34.84
	C11/G10	-0.28	-0.43	2.81	-10.27	-4.19	26.05
	T16/C15	9.79	-1.94	5.03	39.73	10.62	-32.29
	C15/G14	-0.35	-0.37	2.74	-11.25	-2.32	25.57
B-DNA <sup>18</sup>	0.0	0.0	3.38	0.0	0.0	36.0	

Para los pasos 2-3 y equivalentes, se observan valores similares a los obtenidos en dúplex de DNA, pero con valores de *twist* y de *rise* significativamente inferiores. En particular se observa un valor de *rise* muy bajo para el dímero de  $d\langle pTGCTcGCT \rangle$ , que presenta un valor más parecido al que presentan estructuras de A-DNA,  $(2.6\text{Å})^{18,19}$ .

Para el paso 3-4 y equivalentes se observan diferencias importantes con el B-DNA. Los residuos que ocupan la primera posición del loop presentan, respecto al residuo anterior en la secuencia, valor de *twist* de  $\sim -40^\circ$ , *shift* de  $\sim 11\text{Å}$  y *tilt* de entre  $30-40^\circ$ . Estos parámetros tan inusuales hacen que las bases de los residuos 4 y 8 se aparten de su propia hebra y formen una especie de "tapa" sobre el apilamiento de la otra hebra de la subunidad (residuo 4 sobre apilamiento 6-15 y 7-14 y residuo 8 sobre apilamiento 2-11 y 3-10). El valor de *tilt* observado responde, por tanto, a la inclinación observada entre los dos apilamientos ya que la primera base del loop y el par de bases sobre el que se sitúa se encuentran en planos prácticamente paralelos.

En la Figura 3.15 se muestra la parte de la estructura correspondiente a los pares de bases y se compara el apilamiento observado para los pasos 2-3 y 3-4 con el típico en un dúplex de B-DNA. En el caso de la estructura dimérica se observa un mayor solapamiento entre las bases apiladas que forman el par.

<sup>18</sup> Saenger, W. «Principles of Nucleic Acid Structure» 1984, Springer-Verlag, New York Inc., 266-267.

<sup>19</sup> Sinden, R.R. «DNA Structure and Function» 1994, Academic Press, Inc, 21-30.



**Figura 3.15.-** Apilamiento 5'-purina-pirimidina-3' en: A) pares A-T de la estructura dimérica de  $d\langle pCATTCAAT \rangle$  y B) pares AT en B-DNA.

La similitud con B-DNA observada para el apilamiento entre pares de bases, se puede extender también al par de bases. En la Tabla 3.8 se muestran los valores obtenidos para los parámetros de hélice base-base para cada par de bases.

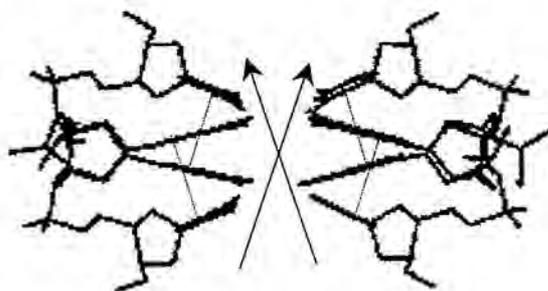
**Tabla 3.8.-** Parámetros de hélice base-base para cada par de bases<sup>§</sup> en ambas estructuras diméricas respecto al eje de cada apilamiento.

Base pair	Shear $S_x$	Stretch $S_y$	Stagger $S_z$	Buckle K	Propeller twist $\omega$	Opening $\sigma$
A2-T11	0.17	-0.11	-0.28	1.47	-6.06	3.91
T3-A10	-0.25	-0.13	-0.27	0.29	-5.09	5.03
T7-A14	0.24	-0.12	-0.24	1.02	-6.53	5.98
A6-T15	-0.22	-0.13	-0.30	0.16	-6.53	5.98
G2-C11	-0.23	-0.04	-0.30	-9.25	-3.64	2.55
C3-G10	0.21	-0.04	-0.30	9.55	-4.02	2.23
C7-G14	-0.18	-0.04	-0.39	-10.14	-8.54	2.49
G6-C15	0.19	-0.04	-0.37	8.85	-7.58	3.31
B-DNA	0.0	0.0	0.1	0.2	3.7	-4.1

§ Para determinar los parámetros de hélice globales se ha definido un eje formado por las bases 2, 3, 10 y 11, para los pares 2-10 y 3-11, y un segundo eje formado por las bases 6, 7, 14 y 15 para los pares 6-15 y 7-14.

De acuerdo con los valores obtenidos para los diferentes parámetros, los pares de bases para ambas estructuras diméricas presentan también una gran similitud con los pares observados en dúplex de DNA. Únicamente cabe destacar ligeras variaciones en cuanto a los valores de *propeller twist*, *buckle* y *opening*, pero que se encuentran dentro del margen de valores observados para distintos oligonucleótidos que presentan una estructura de B-DNA.

La definición de un eje para cada uno de los apilamientos nos permitió calcular de forma precisa la inclinación relativa entre ellos. Esta inclinación se determinó calculando el producto escalar de los vectores unitarios que definen cada eje, obteniéndose un valor de  $37^\circ$  en el caso de  $d\langle pCATTCAAT \rangle$ , y de  $40^\circ$  en el caso de  $d\langle pTGCTCGCT \rangle$ . En la Figura 3.16 se muestra una molécula de dímero donde se han indicado los ejes definidos por cada apilamiento.



**Figura 3.16.-** Disposición de los ejes definidos por cada uno de los apilamientos.

El valor de desplazamiento de las bases respecto al eje de cada apilamiento en la dirección  $x$  ( $dx$ ) permite estimar la distancia entre los dinucleótidos apareados de una misma tétrada. Esta distancia nos puede dar una idea del área en el interior de cada tétrada. En la Tabla 3.9 se muestran estos valores teniendo en cuenta cada uno de los ejes.

**Tabla 3.9.-** Valores de  $dx$  de las diferentes bases respecto al eje de cada apilamiento.

Base	x displacement dx (1)	x displacement dx (2)	Base	x displacement dx (1)	x displacement dx (2)
A2	-0.19	7.54	G2	-0.16	6.60
T3	-0.13	7.55	C3	0.07	6.92
A10	0.12	-8.06	G10	-0.14	-6.89
T11	-0.36	-7.16	C11	0.06	-6.55
A6	7.69	-0.09	G6	6.78	-0.19
T7	7.49	-0.20	C7	6.76	0.11
A14	-7.97	0.04	G14	-6.71	-0.07
T15	-7.09	-0.32	C15	-6.81	0.00

(1) Eje definido por los residuos 2, 3, 10 y 11.

(2) Eje definido por los residuos 6, 7, 14 y 15.

En base a los valores obtenidos, la cavidad en el interior de la estructura bi-loop parece ser mayor en el caso del dímero con pares A-T. Este hecho podría implicar diferencias en cuanto a interacción de dichas estructuras con cationes metálicos y dar una explicación a por qué en la estructura de RX de  $d\langle pCATTcATT \rangle^2$  se observó un ión  $Na^+$  en el centro de la estructura mientras que en el caso del heptámero lineal  $d(GCATGCT)^8$ , que presenta rasgos estructurales muy similares, no fue así. Esta hipótesis deberá comprobarse en disolución estudiando la influencia de diferentes cationes para ambas estructuras.

Para cada residuo de la estructura dimérica se obtuvieron los valores promedio<sup>20</sup> para los diferentes ángulos de torsión, así como la conformación del azúcar en cada caso, Tablas 10 y 11. Como es habitual en estructuras determinadas mediante RMN, los ángulos del azúcar y el glicosídico están mejor definidos que el resto. Como indicador del grado de definición de cada ángulo se ha utilizado el parámetro de orden<sup>21</sup>. Un valor de parámetro de orden de 1 indica que

<sup>20</sup>Los valores promedio de los ángulos de torsión y glicosídicos se han obtenido calculando la media angular.

<sup>21</sup>Hyberts, S.G.; Golberg, M.S.; Havel, T.F.; Wagner, G. *Protein Science* **1992**, 1, 736-751.

el valor del ángulo está muy bien definido, encontrándose para las diferentes estructuras resultantes un valor muy similar. Un valor de 0 indica que el ángulo no está definido, obteniéndose valores muy dispares en las diferentes estructuras resultantes.

**Tabla 3.10.-** Valores promedio de los ángulos de torsión para cada residuo de  $d\langle pCATTCAAT \rangle$  en la estructura dimérica. En la última columna de la tabla se ha indicado la conformación del azúcar.

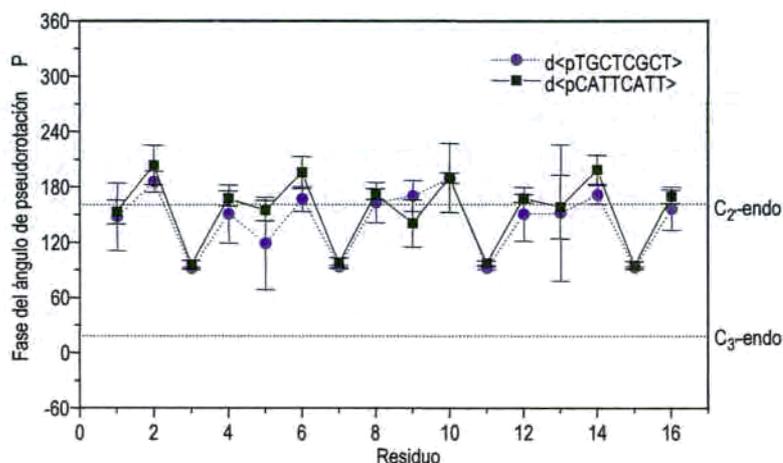
	$\alpha$		$\beta$		$\gamma$		$\delta$		$\epsilon$		$\zeta$		$\chi$		Conformación azúcar
	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	
1C	107.4	0.67	-151.5	0.80	40.7	0.81	138.7	0.99	-112.3	0.85	100.9	0.40	-125.3	0.99	C <sub>2'</sub> -endo
2A	-8.4	0.57	178.9	0.59	99.0	0.73	135.9	1.00	114.4	0.46	-32.0	0.42	-99.2	0.99	C <sub>3'</sub> -exo
3T	-116.0	0.60	176.9	0.99	51.6	0.98	102.0	1.00	-101.0	0.97	-13.9	0.27	-134.6	1.00	O <sub>4'</sub> -endo
4T	-88.0	0.98	171.6	0.98	53.0	0.99	151.1	0.98	-146.3	0.29	130.7	0.99	-151.3	1.00	C <sub>2'</sub> -endo
5C	171.5	0.16	176.5	0.83	1.1	0.49	138.8	0.99	-103.1	0.88	110.1	0.42	-129.4	1.00	C <sub>2'</sub> -endo
6A	-45.7	0.84	-141.1	0.85	69.2	0.90	137.3	1.00	152.4	0.58	-49.6	0.74	-97.4	1.00	C <sub>3'</sub> -exo
7T	-87.2	0.67	172.6	0.99	50.0	0.99	102.9	1.00	-98.6	0.99	-71.6	0.45	-136.2	1.00	O <sub>4'</sub> -endo
8T	-83.4	0.99	170.4	0.99	52.6	0.99	153.6	0.98	-163.8	0.86	131.9	0.99	-148.0	1.00	C <sub>2'</sub> -endo
9C	70.7	0.46	-158.5	0.80	33.8	0.73	132.5	0.97	-117.9	0.77	101.0	0.68	-128.9	1.00	C <sub>2'</sub> -endo
10A	-29.8	0.49	-151.9	0.86	77.7	0.80	132.8	0.98	115.7	0.45	-67.6	0.36	-100.0	0.99	C <sub>3'</sub> -exo
11T	-106.9	0.67	177.3	0.98	49.9	0.98	102.5	1.00	-99.2	0.99	-18.5	0.29	-135.2	0.99	O <sub>4'</sub> -endo
12T	-88.3	0.99	167.2	1.00	56.9	1.00	149.6	0.99	152.6	0.57	132.5	0.99	-151.8	1.00	C <sub>2'</sub> -endo
13C	18.0	0.30	173.8	0.51	-25.2	0.29	141.6	0.96	-95.3	0.91	142.0	0.83	-134.0	0.88	C <sub>2'</sub> -endo
14A	-44.8	0.82	-143.3	0.96	67.9	0.96	136.1	0.99	133.8	0.53	-71.3	0.83	-98.4	0.99	C <sub>3'</sub> -exo
15T	-100.7	0.69	174.8	0.99	46.8	0.99	101.6	1.00	-95.3	0.98	-44.8	0.34	-133.9	0.99	O <sub>4'</sub> -endo
16T	-96.0	0.92	175.3	0.93	56.3	0.97	157.1	1.00	-139.9	0.73	131.3	0.96	-149.3	0.99	C <sub>2'</sub> -endo

**Tabla 3.11.-** Valores promedio de los ángulos de torsión para cada residuo de  $d\langle pTGCTCGCT \rangle$  en la estructura dimérica. En la última columna de la tabla se ha indicado la conformación del azúcar.

	$\alpha$		$\beta$		$\gamma$		$\delta$		$\epsilon$		$\zeta$		$\chi$		Conformación azúcar
	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	
1T	175.6	0.67	178.2	0.90	34.7	0.95	154.5	0.95	177.2	0.62	-66.4	0.78	-137.4	0.99	C <sub>2'</sub> -endo
2G	-129.7	0.90	-121.0	0.99	39.0	1.00	138.3	1.00	-177.8	1.00	31.9	0.91	-99.2	1.00	C <sub>3'</sub> -exo
3C	-54.4	1.00	169.7	1.00	41.7	1.00	106.1	1.00	-89.6	0.99	-86.8	1.00	-120.2	1.00	O <sub>4'</sub> -endo
4T	-82.0	0.99	175.7	0.99	33.1	1.00	143.0	0.95	-65.9	0.74	143.4	0.98	-149.8	1.00	C <sub>2'</sub> -endo
5C	-132.8	0.13	-151.8	0.77	60.3	0.78	139.9	0.86	-117.8	0.55	-82.9	0.10	-140.1	0.86	C <sub>2'</sub> -endo
6G	-68.1	0.69	-140.9	0.93	53.3	0.98	131.4	1.00	-178.4	1.00	-41.4	0.59	-100.8	1.00	C <sub>2'</sub> -endo
7C	-55.9	1.00	169.3	1.00	43.9	1.00	108.5	1.00	-90.6	1.00	-85.2	1.00	-120.7	1.00	O <sub>4'</sub> -endo
8T	-86.9	1.00	174.3	1.00	39.4	1.00	138.1	0.99	5.0	0.85	150.5	1.00	-146.4	1.00	C <sub>2'</sub> -endo
9T	70.7	0.46	178.9	0.88	48.5	0.89	140.7	0.93	-176.9	1.00	150.1	0.72	-142.4	0.87	C <sub>2'</sub> -endo
10G	-125.5	0.98	-120.8	1.00	35.8	1.00	139.1	1.00	-179.0	1.00	36.2	1.00	-99.2	1.00	C <sub>3'</sub> -exo
11C	-54.1	1.00	170.3	1.00	43.2	1.00	106.5	1.00	-91.2	1.00	-87.5	1.00	-120.3	1.00	O <sub>4'</sub> -endo
12T	-79.5	1.00	172.8	1.00	32.3	1.00	140.3	0.94	-52.5	0.66	144.8	1.00	-151	1.00	C <sub>2'</sub> -endo
13C	106.9	0.55	154.9	0.68	63.2	0.21	160.8	0.86	-123.5	0.32	32.8	0.08	-170	0.55	C <sub>2'</sub> -endo
14G	-87.0	0.70	-127.7	1.00	54.4	1.00	132.3	1.00	-177.8	1.00	-17.8	0.66	-101.2	1.00	C <sub>2'</sub> -endo
15C	-53.7	1.00	170.2	1.00	44.1	1.00	108.2	1.00	-91.4	1.00	-86.0	1.00	-121.3	1.00	O <sub>4'</sub> -endo
16T	-86.5	1.00	173.9	1.00	38.6	1.00	134.6	0.98	13.6	0.70	149.9	1.00	-149.9	0.99	C <sub>2'</sub> -endo

Los parámetros de orden indican que el esqueleto azúcar-fosfato se encuentra muy bien definido en la zona correspondiente a residuos apilados, mientras que en el caso de los residuos 1, 5, 9 y 13 se observa una menor definición. Esto es debido a que estos residuos se encuentran totalmente expuestos al disolvente y presentan una mayor flexibilidad. Si se comparan ambas tablas, el esqueleto azúcar-fosfato resulta estar mejor definido en el caso de  $d\langle p\text{TGCTCGCT} \rangle$ . Esto es una consecuencia directa del mayor número de restricciones experimentales (ver sección 3.2.1.1 del presente capítulo).

En ambos dímeros, la conformación del azúcar es típicamente  $C_2'$ -endo, excepto en el caso de los residuos 3, 7, 11 y 15, que adoptan una conformación  $O_4'$ -endo. Este tipo de conformaciones también se han observado en fragmentos de B-DNA donde, para mantener constante la distancia fosfato-fosfato, los residuos de pirimidina, adoptan una conformación  $O_4'$ -endo, favoreciendo valores bajos de  $\delta$  ( $\sim 110^\circ$ ), mientras que los residuos de purina, con una conformación  $C_3'$ -exo, favorecen valores altos de  $\delta$  ( $\sim 135^\circ$ )<sup>22</sup>. Tal como se muestra en la Figura 3.17 este efecto se observa para los residuos 2, 3, 6, 7 y equivalentes, que son los que están formando la parte central de las estructuras diméricas.

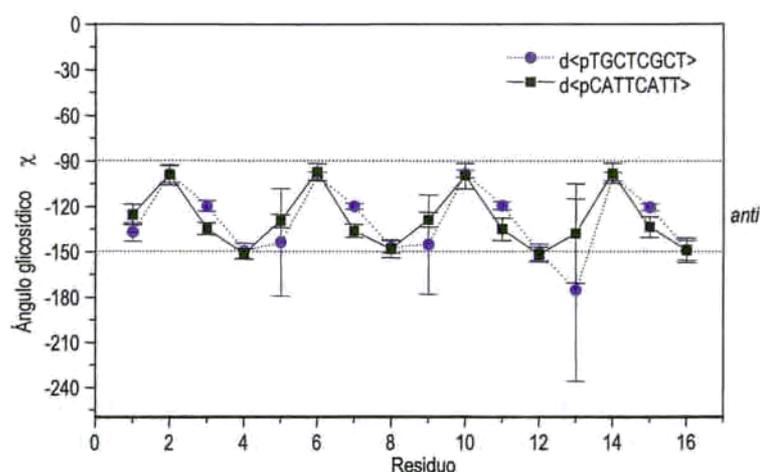


**Figura 3.17.-** Variación de la fase del ángulo de pseudorotación  $P$  a lo largo de la secuencia.

Todos los ángulos glicosídicos presentan una conformación *anti*, no obstante se observan variaciones importantes a lo largo de la secuencia. El ángulo  $\chi$  está fuertemente correlacionado con la conformación del azúcar. Los residuos 2, 3, 6, 7 y equivalentes difieren en el valor de  $\chi$ :  $\sim 100^\circ$  para los residuos 2, 6, 10 y 14 y de  $-120$  a  $-130^\circ$  para los residuos 3, 7, 11 y 15, como era de esperar en base al principio de anticorrelación<sup>23</sup>, que relaciona el ángulo glicosídico con el ángulo  $\delta$ . En los residuos 4, 8, 12 y 16 este principio no se cumple y se observan valores de  $\chi$  de  $\sim 150^\circ$  para una conformación del azúcar  $C_2'$ -endo. En los residuos 1, 5, 9 y 13 estos parámetros se encuentran muy poco definidos. En la Figura 3.18 se muestra como varía el ángulo glicosídico a lo largo de la secuencia nucleotídica.

<sup>22</sup> Saenger, W. «Principles of Nucleic Acid Structure», Springer-Verlag, New York Inc., 1984, 263.

<sup>23</sup> Saenger, W. «Principles of Nucleic Acid Structure», Springer-Verlag, New York Inc., 1984, 262.



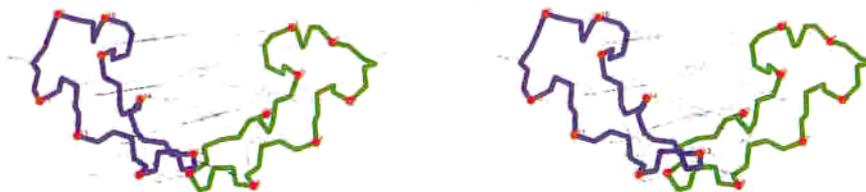
**Figura 3.18.-** Variación del ángulo glicosídico  $\chi$  a lo largo de la secuencia.

Como se observa en la Tablas 3.10 y 3.11, los ángulos de torsión del esqueleto azúcar-fosfato en la zona central de la molécula (residuos 2, 3 y equivalentes) presentan valores típicos de B-DNA: conformación  $g^-$  de  $\alpha$ ,  $g^+$  de  $\gamma$  y trans para  $\beta$ ,  $\epsilon$  y  $\zeta$ . En los residuos del loop, en cambio, estos ángulos deben apartarse de sus valores estándar para que se produzca el giro de la cadena. Como se puede observar en las tablas, los parámetros de orden son bajos en estas zonas, lo que indica que el esqueleto en la zona de los loops está esencialmente desordenado en disolución.

A pesar de la similitud de la parte central de la estructura con el B-DNA, al tratarse de estructuras irregulares en las que se producen giros cerrados en el esqueleto, se observan distancias interfosfato diferentes en función de la localización de los residuos en la estructura. En ambos dímeros los valores más bajos se han obtenido entre los residuos 3-4 (5.3-5.8 Å) y 7-8 (5.4-5.9 Å), que son justamente puntos donde se producen giros en el esqueleto. Entre los residuos 2-3 y 6-7 se han observado distancias entre 5.8-6.7 Å, para el  $d\langle pCATTCAATT \rangle$ , y entre (5.6-5.9 Å) para el  $d\langle pTGCTCGCT \rangle$ , observándose para este último una distancia media significativamente inferior a la habitual en B-DNA,  $\sim 7.0$  Å<sup>24</sup>.

El esqueleto azúcar-fosfato de cada subunidad presenta, además, un punto de máxima proximidad entre las dos hebras antiparalelas, observándose una distancia entre 5.4-5.8 Å entre el grupo fosfato del residuo 3 y el del residuo 7. En este sentido, hay que subrayar el hecho de que para las estructuras dimericas estabilizadas con pares AT, donde no se pueden formar puentes de hidrógeno intramoleculares a través del surco menor, se observe también esta proximidad. Este hecho pone de manifiesto que la proximidad de las hebras responde principalmente a interacciones Van der Waals favorables entre las desoxirribosas y no sólo a las interacciones entre las bases para completar cada cuarteto. Esta interacción entre desoxirribosas posiblemente contribuye a estabilizar los loops en cada subunidad. En la Figura 3.19 se muestra el esqueleto de una de las subunidades del dímero  $d\langle pTGCTCGCT \rangle$ , donde se han señalado las distancias interfosfato entre las dos hebras del ciclo.

<sup>24</sup> Saenger, W. «Principles of Nucleic Acid Structure», Springer-Verlag, New York Inc., 1984, p.237.

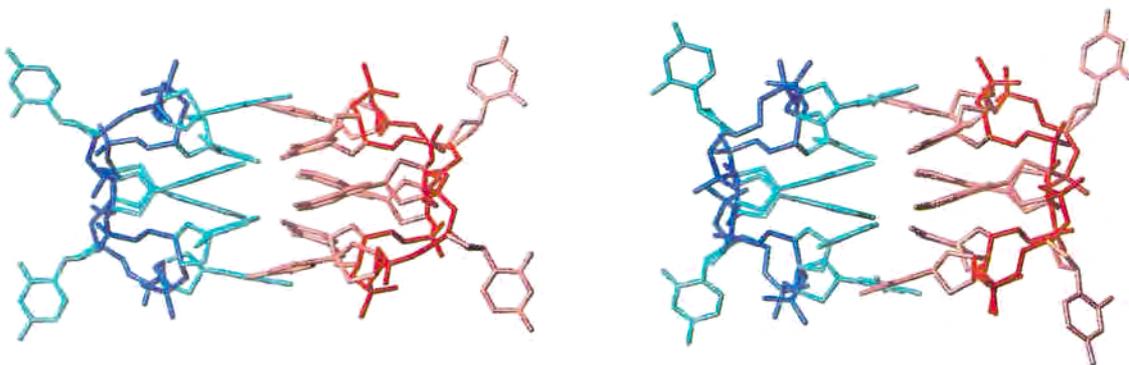


**Figura 3.19.-** Vista estereoscópica del esqueleto azúcar-fosfato de una de las subunidades de dímero para  $d\langle pTGCTCGCT \rangle$ .

Como consecuencia de la formación de una estructura de cuatro hebras, se observan cuatro surcos, dos anchos y dos muy estrechos. En ambos octámeros, los dos primeros presentan una anchura de 14.5 Å (estimado a través de la distancia entre los carbonos C4' enfrentados en cada hebra). En el caso de  $d\langle pCATTCAAT \rangle$ , los surcos estrechos presentan una anchura de 5.5 Å, que es incluso menor (4.5 Å) en el caso de  $d\langle pTGCTCGCT \rangle$ .

### 3.2.3.2.- Comparación con otras estructuras relacionadas.

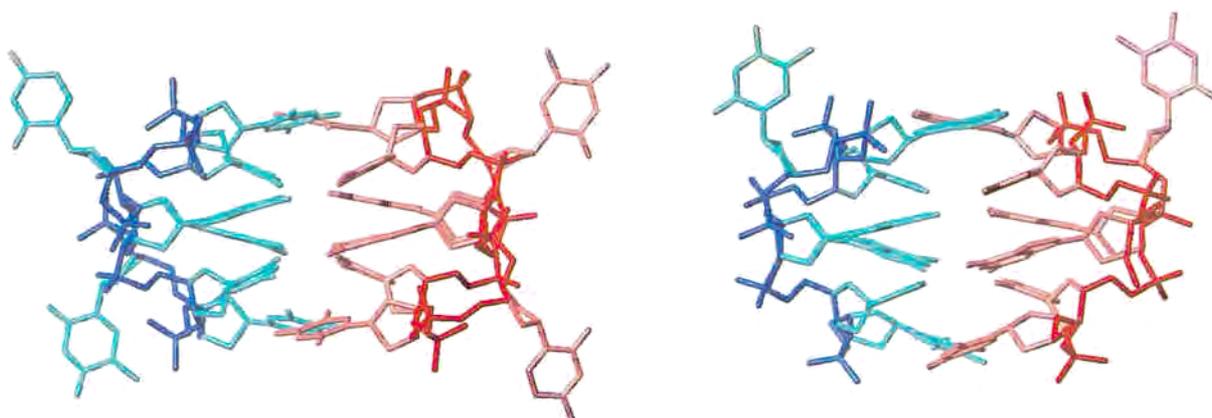
Los dos octámeros estudiados en disolución no sólo tienen una estructura muy similar entre sí, sino que presentan una gran similitud con la estructura cristalina del octámero cíclico  $d\langle pCATTCAAT \rangle$ . En la Figura 3.20 se muestra la estructura cristalográfica de esta molécula y el promedio de las 10 estructuras obtenidas mediante cálculos de dinámica molecular en una orientación de superposición. Se puede observar que ambas estructuras se superponen perfectamente. Aunque la RMSD calculada para toda la molécula es de 1.30 Å, este valor baja a 0.93 Å si no se incluyen los residuos 1,5,9 y 13 en la comparación. Si se superponen únicamente las bases de los residuos bien ordenados (2,3,4 y equivalentes), la RMSD baja hasta 0.64 Å, lo que indica que las estructuras son prácticamente idénticas en esta zona.



**Figura 3.20.-** A) Estructura promedio en disolución y B) estructura cristalina de  $d\langle pCATTCAAT \rangle$ .

El gran parecido entre estas estructuras pone de manifiesto que la estructura cristalina no es producto de fuerzas de empaquetamiento en el cristal. Este hecho apoya fuertemente la hipótesis de que el bi-loop podría ser un motivo general.

Aún más interesante es la similitud entre la estructura de *d*<pTGCTCGCT> y la del heptámero lineal *d*(GCATGCT). En la Figura 3.21 se muestran ambas estructuras, también en una orientación de superposición. Puesto que, en este caso, los nucleótidos que forman el loop no son los mismos, sólo se pueden superponer los nucleótidos apareados. La RMSD para los nucleótidos que forman la tetrada es de 0.8 Å, si se incluyen todos los átomos pesados, y baja hasta 0.4 Å, si sólo se superponen las bases.



**Figura 3.21.-** A) Estructura promedio en disolución de *d*<pTGCTCGCT> y B) estructura cristalina de *d*(pGCATGCT).

Ambas estructuras están estabilizadas por 4 pares GC Watson-Crick y la formación de puentes de hidrógeno GC intramoleculares, que implican 8 puentes de hidrógeno por cuarteto. Sin embargo, la diferente composición de los mini-loops, así como el hecho de que una de las secuencias sea un octámero cíclico y la otra un heptámero lineal, implica importantes diferencias en cuanto a las interacciones en las que se encuentran involucradas las bases que ocupan la primera posición de los mini-loops. En el octámero cíclico, las bases que ocupan la primera posición son timinas, con los grupos metilo muy próximos entre sí. La interacción favorable es, por tanto, de tipo hidrofóbico. En el heptámero lineal, en cambio, esta posición está ocupada por dos adeninas enfrentadas, que se encuentran formando un par A·A. Este par A·A implica la formación de dos puentes de hidrógeno adicionales entre H6 y N7. En el otro extremo de la molécula, las dos timinas situadas en posición 7 presentan una interacción hidrofóbica entre sus grupos metilo análoga a la del octámero cíclico. En la Figura 3.22 se muestran de forma esquemática las dos estructuras y los puentes de hidrógeno que estabilizan cada una de ellas.

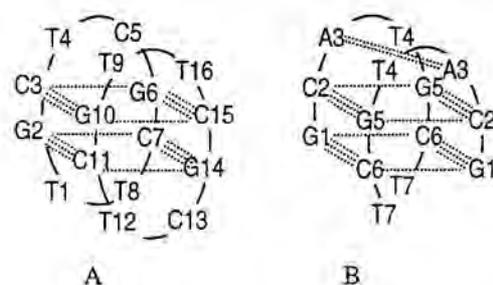


Figura 3.22.- Interacciones de puente de hidrógeno que estabilizan las estructuras dimericas de A)  $d\langle pTGCTCGCT \rangle$  y B)  $d\langle GCATGCT \rangle$ .

La elevada similitud estructural, a pesar de las diferencias en la secuencia, unido al hecho de que se haya observado la misma estructura para secuencias cíclicas y lineales, es lo que proporciona al motivo bi-loop un carácter general.

El motivo bi-loop presenta ciertas similitudes con otras formas irregulares de DNA. Al tratarse de un motivo tipo cuádruplex, las estructuras con las que guarda un mayor parecido son las tétradas de guaninas<sup>25-26</sup>. Este tipo de estructuras se ha observado en distintos oligonucleótidos, en muchos casos fragmentos de secuencias de telómeros<sup>27-28</sup>. El núcleo central de este motivo estructural es una tétrada G:G:G:G con las bases en conformaciones alternantes *syn* y *anti*. Cada tétrada está estabilizada por 8 puentes de hidrógeno. Al igual que en nuestras estructuras de octámeros cíclicos, las hebras de los cuádruplex de guaninas se encuentran muy próximas, con distancias entre los oxígenos de grupos fosfato pertenecientes a hebras antiparalelas contiguas del orden de 5-5.5 Å. La principal diferencia existente entre ambas estructuras es la inclinación entre las bases que interaccionan vía el surco menor ya que, para los dímeros, los dos dinucleótidos que constituyen un cuarteto presentan una inclinación relativa de  $\sim 37^\circ$ . En el caso de las tétradas de guaninas, las cuatro bases que forman el cuarteto suelen estar en el mismo plano, sin embargo existen casos donde se ha observado un buckle importante en alguna de las guaninas<sup>29-30</sup>. Existe también en la literatura el caso de una tétrada de guaninas en la que, a ambos extremos de la estructura, dos timinas se colocan apiladas actuando como tapa<sup>31</sup>, tal como sucede en el bi-loop. Recientemente se han observado dos estructuras de cuádruplex en las que no todas las tétradas están formadas por guaninas sino que incluyen una tétrada de timinas<sup>32</sup> o de adeninas<sup>33</sup>.

Si bien el caso de  $d\langle pCATTCATT \rangle$  es el único en el que se ha observado una tétrada A:T:A:T, tétradas del tipo G:C:G:C se han observado con anterioridad, no sólo en la estructura del motivo bi-loop, sino en la estructura de dos cuádruplex de guaninas en los que una de las

<sup>25</sup> Smith, F.W.; Feigon, J. *Nature* **1992**, 356, 164-168.

<sup>26</sup> Schultze, P.; Smith, F.W.; Feigon, J. *Structure* **1994**, 2, 221-233.

<sup>27</sup> Williamson, J.R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, 23, 703-730.

<sup>28</sup> Rhodes, D.; Giraldo, R. *Curr. Opin. Struct. Biol.* **1995**, 5, 311-322.

<sup>29</sup> Wang, Y.; Patel, D. *Structure* **1993**, 1, 263-282.

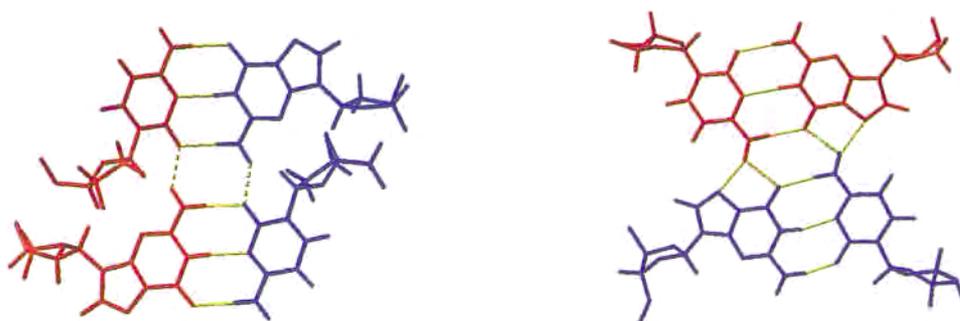
<sup>30</sup> Kang, C.H.; Zhang, X.; Ratcliff, R.; Moyzis, R.; Rich, A. *Nature* **1992**, 356, 126-131.

<sup>31</sup> Kettani, A.; Bouaziz, S.; Wang, W.; Jones, R.A.; Patel, D. *Nat. Struct. Biol.* **1997**, 4, 382-389.

<sup>32</sup> Patel, P.K.; Hosur, R.V. *Nucleic Acids Res.* **1999**, 27, 2457-2464.

<sup>33</sup> Patel, P.K.; Hoti, A.S.R.; Hosur, R.V. *Nucleic Acids Res.* **1999**, 27, 3836-3843.

tétradas no está formada únicamente por guaninas sino por guaninas y citosinas, todas ellas en conformación anti<sup>17,31</sup>. En estos casos también se forman estructuras diméricas con pares de bases tipo Watson-Crick pero, al contrario de lo que ocurre en el motivo bi-loop, el apareamiento es intra-molecular y la tétrada se forma a través del surco mayor, Figura 3.23. Otra diferencia importante es que estas tétradas son coplanares, al igual que las de guaninas, y no presentan la inclinación de  $\sim 37\text{-}40^\circ$  observada en los bi-loops.



**Figura 3.23.-** A) Tétrada G:C:G:C de  $d\langle pTGCTCGCT \rangle$ , donde se observa la asociación de los pares Watson-Crick a través del surco menor. B) Tétrada G:C:G:C de la estructura en disolución de  $d(GCG_2T_3GCG_2)^{31}$ , donde se observa la asociación de los pares G-C a través del surco mayor.

### 3.2.3.3.- El motivo bi-loop: importancia y generalidad.

El hecho de haber observado el motivo bi-loop en disolución, y que la estructura, obtenida mediante cálculos de dinámica molecular, presente un grado de coincidencia tan elevado con la obtenida mediante técnicas cristalográficas, constituye una razón suficiente para considerar la estructura bi-loop como un nuevo motivo del DNA bien definido. Además, el que este motivo se haya observado para tres secuencias distintas, una de ellas lineal, contribuye a reforzar la idea de que, en efecto, podría tratarse de un motivo de carácter general.

Este motivo constituye una nueva forma de reconocimiento molecular entre fragmentos de DNA. Se puede especular que este tipo de reconocimiento molecular puede estar involucrado en procesos de recombinación homóloga. La formación del bi-loop podría servir como reconocimiento del punto de iniciación del proceso de intercambio de cadenas<sup>2,34</sup>. No obstante, para poder valorar la importancia de este motivo y entrar a considerar posibles implicaciones biológicas con mayor fundamento, es necesario ampliar el estudio y establecer posibles requisitos secuenciales para su formación.

Las tres secuencias en las que se ha observado el motivo bi-loop, aunque difieren en la composición, presentan ciertas similitudes.

<sup>34</sup> Patel, D.J.; Bouaz, S.B.; Kettani, A.; Wang, Y. «*Nucleic Acid Structure*» 1999, 389, Ed. Neidle, S., Oxford University Press.



Todas ellas contienen dos dinucleótidos autocomplementarios, que son purina-pirimidina en la dirección 5'-3'. Estos dinucleótidos están separados por dos residuos, que son los que van a formar los dos mini-loops. Por los resultados obtenidos, el primer residuo del mini-loop puede ser tanto purina como pirimidina, mientras que el segundo residuo ha resultado ser una pirimidina en los tres casos estudiados. A partir de estas similitudes podría extraerse a priori una secuencia consenso ...RYNYRYN....., donde R= purina, Y= pirimidina y N= cualquier base.

Para analizar los requisitos secuenciales necesarios para la formación del bi-loop y comprobar la validez de esta secuencia consenso, es necesario llevar a cabo el estudio de otras secuencias análogas en las que se hayan introducido cambios en las distintas posiciones. La descripción de estos estudios se detalla en el Capítulo 4.

### 3.2.3.4.- Análisis de las estructuras de los monómeros.

Las estructuras de las formas monoméricas, descritas en el apartado 3.2.2.2, no son tan refinadas como las obtenidas para los dímeros. Como ya se comentó anteriormente, en ambos casos surgen dificultades experimentales que impiden obtener un conjunto de restricciones tan completo como en el caso de las formas diméricas. Dada la indeterminación de estas estructuras, no tiene sentido llevar a cabo un análisis completo en términos de parámetros generalizados de hélice. No obstante, del resultado de los cálculos pueden concluirse un buen número de características estructurales.

En las Tablas 12 y 13 se muestran los valores promedio de los diversos ángulos de torsión de los dos monómeros. La baja definición general de las estructuras se ve reflejada en los parámetros de orden correspondientes, que son mucho más bajos que en el caso de los dímeros. La excepción más notable es la de los ángulos glicosídicos, que aparecen bastante bien definidos. La conformación de los azúcares, aunque variable, siempre se sitúa en la zona S del diagrama de ángulos de pseudorotación. Salvo el ángulo  $\delta$ , que queda restringido por el ciclo de la desoxiribosa, todos los ángulos de esqueleto están pobremente definidos.

*Tabla 3.12.- Valores promedio de los ángulos de torsión para cada residuo de la estructura monomérica de  $d\langle pCATTcATT \rangle$ . En la última columna de la tabla se ha indicado la conformación del azúcar.*

	$\alpha$		$\beta$		$\gamma$		$\delta$		$\epsilon$		$\zeta$		$\chi$		Conformación azúcar
	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	
1C	168.0	0.47	-165.2	0.50	102.6	0.56	145.8	0.98	-30.9	0.21	-106.0	0.41	-93.3	0.87	C <sub>2</sub> -endo/C <sub>3</sub> -exo
2A	-55.6	0.45	-45.9	0.48	-179.0	0.79	130.0	0.97	101.0	0.59	-170.0	0.70	-105.2	0.99	C <sub>1</sub> -exo/C <sub>3</sub> -exo
3T	-176.4	0.49	109.0	0.44	166.7	0.82	149.0	1.00	-174.2	0.98	55.1	0.19	-138.9	0.99	C <sub>2</sub> -endo
4T	-145.3	0.52	-166.0	0.69	66.0	0.67	134.8	0.94	-146.0	0.75	-118.7	0.89	-121.5	0.91	C <sub>1</sub> -exo/C <sub>2</sub> -endo
5C	-157.3	0.39	-161.9	0.62	93.0	0.55	144.0	0.97	-92.2	0.57	-139.1	0.25	-113.5	0.95	C <sub>2</sub> -endo/C <sub>3</sub> -exo
6 <sup>A</sup>	-169.8	0.44	92.4	0.53	162.7	0.86	128.6	0.96	83.6	0.30	87.6	0.52	-104.4	0.99	C <sub>1</sub> -exo/C <sub>3</sub> -exo
7T	-17.2	0.17	161.6	0.59	159.7	0.56	137.2	0.97	-174.7	0.99	51.7	0.19	-133.9	0.99	C <sub>2</sub> -endo
8T	-168.8	0.29	-178.8	0.77	76.9	0.54	135.5	0.95	-130.2	0.67	123.1	0.90	-120.3	0.99	C <sub>1</sub> -exo/C <sub>2</sub> -endo

**Tabla 3.13.-** Valores promedio de los ángulos de torsión para cada residuo de  $d\langle p\text{TGCTCGCT} \rangle$  en la estructura monomérica. En la última columna de la tabla se ha indicado la conformación del azúcar

	$\alpha$		$\beta$		$\gamma$		$\delta$		$\epsilon$		$\zeta$		$\chi$		Conformación azúcar
	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	Media	Par. orden	
1T	-123.1	0.12	-121.3	0.30	140.9	0.64	116.7	0.98	170.5	0.98	-142.4	0.49	-130.6	0.86	C <sub>1'</sub> -exo
2G	163.9	0.73	-163.0	0.73	170.4	0.86	142.0	0.99	132.9	0.37	-41.8	0.74	-87.4	0.98	C <sub>3'</sub> -exo
3C	116.5	0.63	-168.2	0.87	178.5	0.98	109.4	1.00	146.4	0.53	-128.0	0.20	-124.4	0.99	C <sub>1'</sub> -exo
4T	148.5	0.58	-178.2	0.87	148.9	0.59	143.6	0.99	-175.1	0.86	-62.6	0.46	-115.6	0.92	C <sub>2'</sub> -endo
5C	-173.2	0.39	179.8	0.99	145.3	0.62	142.4	0.98	-168.4	0.88	-110.0	0.90	-132.7	0.97	C <sub>2'</sub> -endo
6G	-104.7	0.66	-82.2	0.69	-170.8	0.99	113.3	0.97	-170.0	0.99	173.1	0.61	-90.2	0.98	C <sub>1'</sub> -exo/C <sub>3'</sub> -exo
7C	117.5	0.72	-138.3	0.85	174.8	0.86	109.3	1.00	174.2	0.60	-135.0	0.95	-135.8	0.98	C <sub>1'</sub> -exo
8T	144.5	0.91	-177.9	0.97	109.3	0.70	127.5	0.95	-126.6	0.40	-52.5	0.70	-115.6	0.91	C <sub>1'</sub> -exo/C <sub>2'</sub> -endo

La estructura de  $d\langle p\text{TGCTCGCT} \rangle$  es, en rasgos generales, similar a la del único dumbbell cuya estructura ha sido determinada con anterioridad<sup>35</sup>. El dumbbell  $d\langle p\text{CGC-TT-GCG-TT} \rangle$  contiene un tramo de tres pares de bases y un mini-loop, a cada extremo, de dos timinas. Los pares de bases son también tipo Watson-Crick, pero se observa cierta población de conformero *syn* en la G del loop G-TT-C. En nuestro caso, no hemos observado ningún indicio de este conformero, pero hay que tener en cuenta que los dos mini-loops son, en nuestro caso, C-PyPy-G. En el dumbbell  $d\langle p\text{CGC-TT-GCG-TT} \rangle$  tampoco se observa equilibrio conformacional en el loop C-TT-G. De acuerdo con estudios termodinámicos sistemáticos de diferentes horquillas de DNA y RNA<sup>36,37</sup>, los loops compuestos de pirimidinas y flanqueados por 5'-C y 3'-G (5'-C-(Py)<sub>n</sub>-G-3') presentan una elevada estabilidad. No obstante, no deja de resultar sorprendente que sólo dos pares de bases sean suficientes para estabilizar la estructura dumbbell en el octámero cíclico.

Al contrario que en el caso de  $d\langle p\text{TGCTCGCT} \rangle$ , contamos con varios indicios de que puede existir un equilibrio conformacional en la estructura de  $d\langle p\text{CATTTCATT} \rangle$ . Además de la anchura de línea del protón amino y de los NOEs característicos de conformación *syn* en la adenina, aparece un detalle importante en las estructuras calculadas. En todas las estructuras se observa un ángulo de *buckle* muy grande en los pares AT (como se puede observar en la Figura 3.9. No debemos olvidar que esta estructura se ha calculado forzando el apareamiento Watson-Crick y eliminando los NOEs que provienen claramente de la forma *syn* de la adenina. No obstante, la intensidad de otros NOEs podría estar, en parte, afectada por la presencia de este conformero. El ajuste de estas distancias puede dar lugar a este *buckle* tan grande. Un par A(*syn*)-T(*anti*), similar a éste, ha sido observado anteriormente en horquillas de DNA<sup>37,38</sup>.

También tiene interés comparar las estructuras de los dumbbells con las de las formas diméricas descritas en el apartado 3.2.2.1. Aunque las estructuras son globalmente bastante distintas, la conformación de los azúcares y de los ángulos glicosídicos no son muy diferentes. La diferencia más notable se halla en el esqueleto de fosfatos. En los dumbbells, los fosfatos se

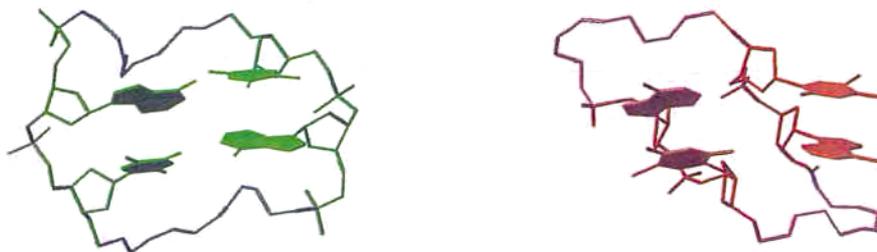
<sup>35</sup> Ippel, J.H.; Lanzotti, V.; Galeone, A.; Mayol, L.; Van den Boogaart, J.E.; Pikkemaat, J.A.; Altona, C. *J. Biomol NMR* **1995**, 6, 403-422.

<sup>36</sup> Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, 24, 379-404.

<sup>37</sup> Blommers, M.; Van den Ven, G.; Van der Marel, J.; Van Boom, J.; Hilbers, C. *Eur. J. Biochem* **1991**, 201, 33-51.

<sup>38</sup> Kallick, D.A.; Wemmer, D.E. *Nucleic Acids Res.* **1991**, 19, 6041-6046.

sítúan de manera que se maximiza la distancia entre ellos (entre 6.5 y 7.5 Å). En cambio, en los dímeros, las cadenas de fosfato se encuentran mucho más próximas, con distancias P-P de hasta 5.5 Å, Figura 3.24. El efecto destabilizante debido a la proximidad de grupos cargados en las formas diméricas debe verse compensado por las interacciones de apilamiento entre las bases y, probablemente, por contactos de tipo hidrofóbico entre los anillos de desoxiribosa.



**Figura 3.24.-** Esqueleto azúcar-fosfato correspondiente a A) monómero y B) dímero.

---

*CAPÍTULO 4*

*ELEFECTO DE LA SECUENCIA EN LA FORMACIÓN DE ESTRUCTURAS  
BI-LOOP.*

---

## 4.- EL EFECTO DE LA SECUENCIA EN LA FORMACIÓN DE ESTRUCTURAS BI-LOOP.

### 4.0.- INTRODUCCIÓN Y OBJETIVOS.

Los resultados expuestos en el capítulo anterior sugieren que el bi-loop podría ser un motivo estructural de carácter general del DNA. Por una parte, se ha encontrado tanto en estado sólido como en disolución, lo que indica que no se trata de un mero artefacto producido por las fuerzas de empaquetamiento en el cristal. Por otra parte, se ha observado en oligonucleótidos tanto cíclicos como lineales, por lo que queda claro que dicho motivo no está estabilizado únicamente por la ciclación de la cadena y que podría darse en el DNA celular. Finalmente, se ha observado en varias secuencias, tanto con pares de bases A·T como G·C y con diversos nucleótidos en los loops. Aunque el número de secuencias estudiado hasta el momento es muy escaso, en el capítulo anterior se sugirió una posible secuencia consenso basada en las similitudes observadas para las tres secuencias en las que se observó el bi-loop.

El objetivo de este capítulo consiste en explorar hasta qué punto el motivo bi-loop puede ser considerado un motivo de carácter general y qué requisitos secuenciales son necesarios para su aparición. En este sentido, el estudio de diversas secuencias puede ayudar a determinar qué tipo de interacciones resultan fundamentales para estabilizar este tipo de estructura. En última instancia, el entender las fuerzas que estabilizan el bi-loop puede ayudarnos a diseñar una secuencia en la que dicho motivo sea lo suficientemente estable como para ser observable en oligonucleótidos lineales en disolución.

El estudio de octámeros cíclicos en disolución, descrito en el capítulo anterior, se llevó a cabo íntegramente mediante RMN. Dicha técnica resulta muy útil para determinar la estructura tridimensional pero es muy costosa, tanto por el tipo de instrumentación (necesidad de realizar experimentos largos en espectrómetros de alto campo), como por la gran cantidad de producto que se requiere. Para profundizar en el estudio del motivo bi-loop sería muy útil contar con otras técnicas que nos permitan detectar dicho motivo con facilidad. En este capítulo se ha explorado la utilización de técnicas como la espectroscopía UV o el dicroísmo circular (DC) para detectar la formación de este nuevo motivo estructural. A diferencia del UV, el DC no se ha utilizado de forma tan extensa en el estudio de ácidos nucleicos, aunque su sensibilidad para detectar diferentes estructuras del DNA es conocida<sup>1</sup>. En este capítulo, nos planteamos, por tanto, el estudio de cuatro nuevas secuencias cíclicas en disolución, mediante RMN, y la caracterización, tanto de éstas como de las dos secuencias estudiadas en el capítulo anterior, mediante la técnica de DC.

### 4.1.- ELECCIÓN DE SECUENCIAS.

En el estudio por RMN de las secuencias d<pCATT<math>CATT</math>> y d<pTGCT<math>CGCT</math>>, se observaron diferencias importantes en cuanto a la estabilidad de las formas diméricas. Aunque las estructuras tridimensionales de ambas moléculas presentan una elevada similitud, las condiciones experimentales en las que se observó la formación del bi-loop en cada caso resultaron muy

<sup>1</sup> Gray, D.M.; Ratliff, R.L.; Vaughan, M.R. *Methods Enzymol.* **1992**, 211, 389-406.

diferentes. Mientras que la estructura dimérica con cuatro pares G-C se observa a concentraciones de 2 mM, la sustitución de pares G-C por pares A-T hace que sea necesaria una concentración cinco veces superior, del orden de 10 mM, para empezar a observar el dímero. El cambio drástico que supone la sustitución de pares G-C por pares A-T, nos induce a pensar que la secuencia de nucleótidos es un factor determinante en la formación y estabilidad del motivo. En este sentido, son muchas las variaciones que, a priori, podrían introducirse en la secuencia de cara a determinar el grado de incidencia de cada una de ellas. En el presente trabajo se han diseñado secuencias destinadas a estudiar tres de los factores que, en base a los resultados del capítulo anterior, hemos creído que podrían tener una mayor importancia.

**Efecto de la naturaleza de los pares de bases.** Una de las características fundamentales del motivo bi-loop es la formación de cuatro pares de bases Watson-Crick intermoleculares. En el capítulo anterior se comprobó que estos pares pueden ser tanto G-C como A-T, siendo la estructura tridimensional muy similar en ambos casos. Una pregunta que se plantea es si es posible la existencia de bi-loops mixtos, estabilizados por la formación de dos pares G-C y dos pares A-T. Una característica importante de un bi-loop mixto es que, al no ser complementarios los dinucleótidos de cada subunidad, no resulta posible la formación de la estructura dumbbell correspondiente.

Para estudiar la existencia de bi-loops mixtos, nos planteamos estudiar un octámero cíclico de secuencia d<pCGCTCATT>. Los *loops* de esta molécula son iguales a los de d<pCATTTCATT> y a uno de los de d<pTGCTCGCT>, por lo que la comparación entre estas tres moléculas nos ayudará a entender el efecto del tipo de par de bases en la formación del bi-loop.

**Efecto del orden en los pares de bases.** En las tres secuencias estudiadas hasta el momento, el orden de los dinucleótidos apareados en el núcleo del bi-loop es 5'-purina-pirimidina-3'. El estudio de un octámero con esta secuencia invertida nos permitiría comprobar la posible secuencia consenso expuesta en el apartado 3.2.3.3 del capítulo anterior. Puesto que la parte central del bi-loop es muy similar al DNA tipo B, sería de esperar que la contribución a la estabilidad de la molécula, debida al apilamiento de estas bases, fuese similar a la que se observa en de dobles hélices de DNA. Para hélices de DNA tipo B, se ha encontrado que la energía de apilamiento es más favorable en el caso de apilamiento purina-3',5'-pirimidina, que en el caso de apilamiento pirimidina-3',5'-purina<sup>2</sup>. La comparación de las secuencias d<pTGCTCGCT> y d<pCCGTCCGT> permitirá determinar cómo afectan las diferencias en el apilamiento de las bases a la formación y estabilización de las estructuras diméricas. De todos modos, hay que ser conscientes que, al invertir la secuencia de los dinucleótidos apareados, se cambia también la de la mini-horquilla que contiene el *loop*, que pasa de ser 5'-pirimidina-loop-purina-3' a 5'-purina-loop-pirimidina-3'.

**Efecto de los residuos en el loop.** En numerosos estudios realizados sobre la estabilidad de diferentes horquillas de DNA<sup>3,4</sup>, se concluye que éstas son más estables cuando los residuos en el *loop* son pirimidinas en lugar de purinas<sup>5-7</sup>. Esto parece ser debido a que el apilamiento de una

<sup>2</sup> Saenger, W. "Principles of Nucleic Acid Structure" 1984, Springer-Verlag New York Inc.

<sup>3</sup> Blomers, M.J.J.; Walters, J.A.L.I.; Haasnoot, C.A.G.; Aelen, J.M.A.; van der Marel, G.A.; Van Boom, J.H.; Hilbers, C.W. *Biochemistry* 1989, 28, 7491-7498.

<sup>4</sup> Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* 1995, 24, 379-404.

<sup>5</sup> Senior, M.M.; Jones, R.A.; Breslauer, K.J. *Proc. Natl. Acad. Sci. U.S.A.* 1988, 85, 6242-6246.

pirimidina en el *loop* presenta un ángulo de *helical twist* muy similar al que presenta en DNA tipo B, mientras que cuando la base es una purina, el ángulo de *helical twist* aumenta mucho<sup>8,9</sup>. Cuando se trata de residuos del *loop* que no se encuentran apilados puede influir el elevado coste energético que supone tener un residuo de purina totalmente expuesto al disolvente.

Para estudiar el efecto de la sustitución de algún residuo del *loop* por una purina, hay que tener en cuenta que la primera base del mini-loop está implicada en el apilamiento de las bases de la estructura dimerica, mientras que la segunda se encuentra completamente expuesta al disolvente. Por esta razón, es previsible que el efecto de la sustitución sea muy diferente si ésta se produce en la primera o en la segunda posición del mini-loop.

Se sintetizaron, por tanto, dos moléculas de secuencia d<pTGCAATGCA> y d<pAGCTAGCT>, con purinas en diferentes posiciones del *loop*. En el primer caso, la adenina del *loop* ocupa la misma posición que en el heptámero lineal d(GCATGCT), cuya estructura cristalográfica presenta un par A·A que contribuye a estabilizar el bi-loop<sup>10</sup>. En el segundo caso, la sustitución se produce en la segunda base del *loop*, que se encuentra totalmente expuesta al disolvente. La comparación de estas dos secuencias con los resultados obtenidos previamente para d<pTGCTCGCT> permitirá determinar la importancia de cada base en las diferentes posiciones del *loop* para la estabilidad del dímero.

Estas secuencias son también interesantes porque, a priori, pueden dar lugar a cuatro estructuras diferentes. Al estar compuestas por cuatro dinucleótidos complementarios es posible la formación de: a) una estructura *dumbbell* con dos pares G·C y dos mini-loops AT (o TA), b) una estructura *dumbbell* con dos pares A·T(T·A) y dos mini-loops GC, c) una estructura bi-loop con cuatro pares G·C y cuatro mini-loops AT(TA) y, por último, d) una estructura bi-loop con cuatro pares A·T(T·A) y cuatro mini-loops GC.

En la Figura 4.1 se muestran de forma esquemática las cuatro secuencias propuestas y las estructuras posibles para cada una de ellas.

El análisis conjunto de cómo afectan cada una de las modificaciones propuestas a la formación y estabilidad de la estructura dimerica, permitirá delimitar posibles requisitos secuenciales para su formación. Puesto que, a excepción de la secuencia mixta, para el resto de secuencias es posible la formación de una estructura *dumbbell*, este estudio nos permitirá también, por una parte, extraer conclusiones en cuanto a la formación de este tipo de estructuras y, por otra parte, determinar si la secuencia afecta de la misma forma a estructuras monoméricas y dimericas. En este sentido resulta especialmente interesante ver cómo afectan variaciones en los mini-loops a

<sup>6</sup> Xodo, L.E.; Manzini, G.; Quadrifoglio, F.; Yathindra, N.; van der Marel, G.A.; van Boom, J.H. *Biochemistry* **1988**, *27*, 6321-6326.

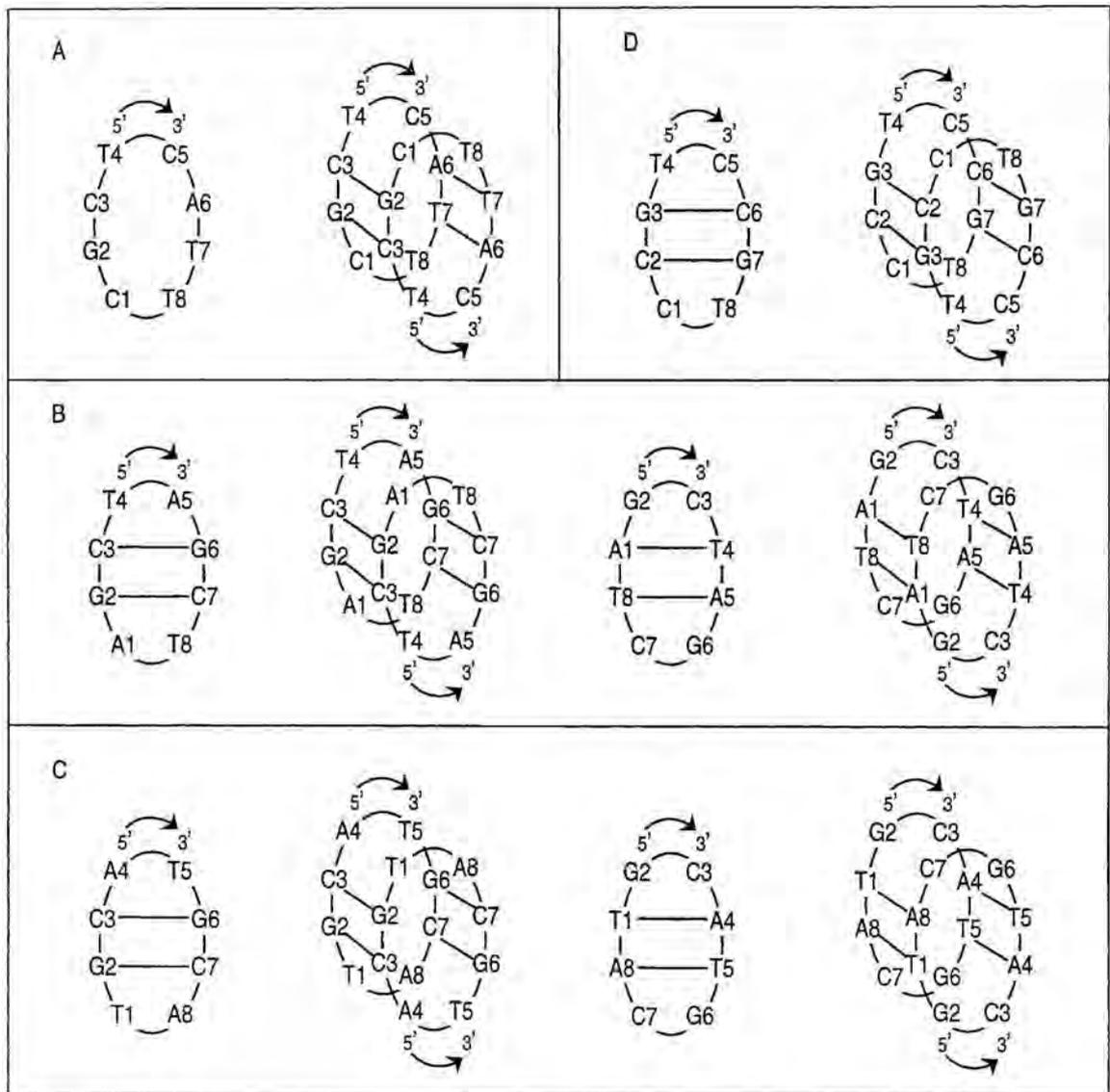
<sup>7</sup> Xodo, L.E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G.A.; van Boom, J.H. *Biochimie* **1989**, *71*, 793-806.

<sup>8</sup> Avizonis, D.Z.; Kearns, D.R. *Biopolymers* **1995**, *35*, 187-200.

<sup>9</sup> Avizonis, D.Z.; Kearns, D.R. *Nucleic Acids Res.* **1995**, *23*, 1260-1268.

<sup>10</sup> Leonard, G.A.; Zhang, S.; Peterson, M.R.; Harrop, S.J.; Helliwell, J.R.; Cruse, W.B.T.; Langlois d'Estaintot, B.; Kennard, O.; Brown, T.; Hunter, W.N. *Structure* **1995**, *3*, 335-340.

la formación de la forma dimérica, ya que su composición resulta crítica en el caso de estructuras tipo dumbbell<sup>3,4</sup>.



**Figura 4.1.-** Posibles estructuras para cada una de las secuencias; A)  $d\langle pCGCTCATT \rangle$ , B)  $d\langle pAGCTAGCT \rangle$ , C)  $d\langle pTGCATGCA \rangle$  y D)  $d\langle pCCGTCCGT \rangle$ .

#### 4.2.- ESTUDIO POR RMN DE: d<pCGCTCATT>, d<pAGCTAGCT>, d<pTGCATGCA> Y d<pCCGTCCGT>.

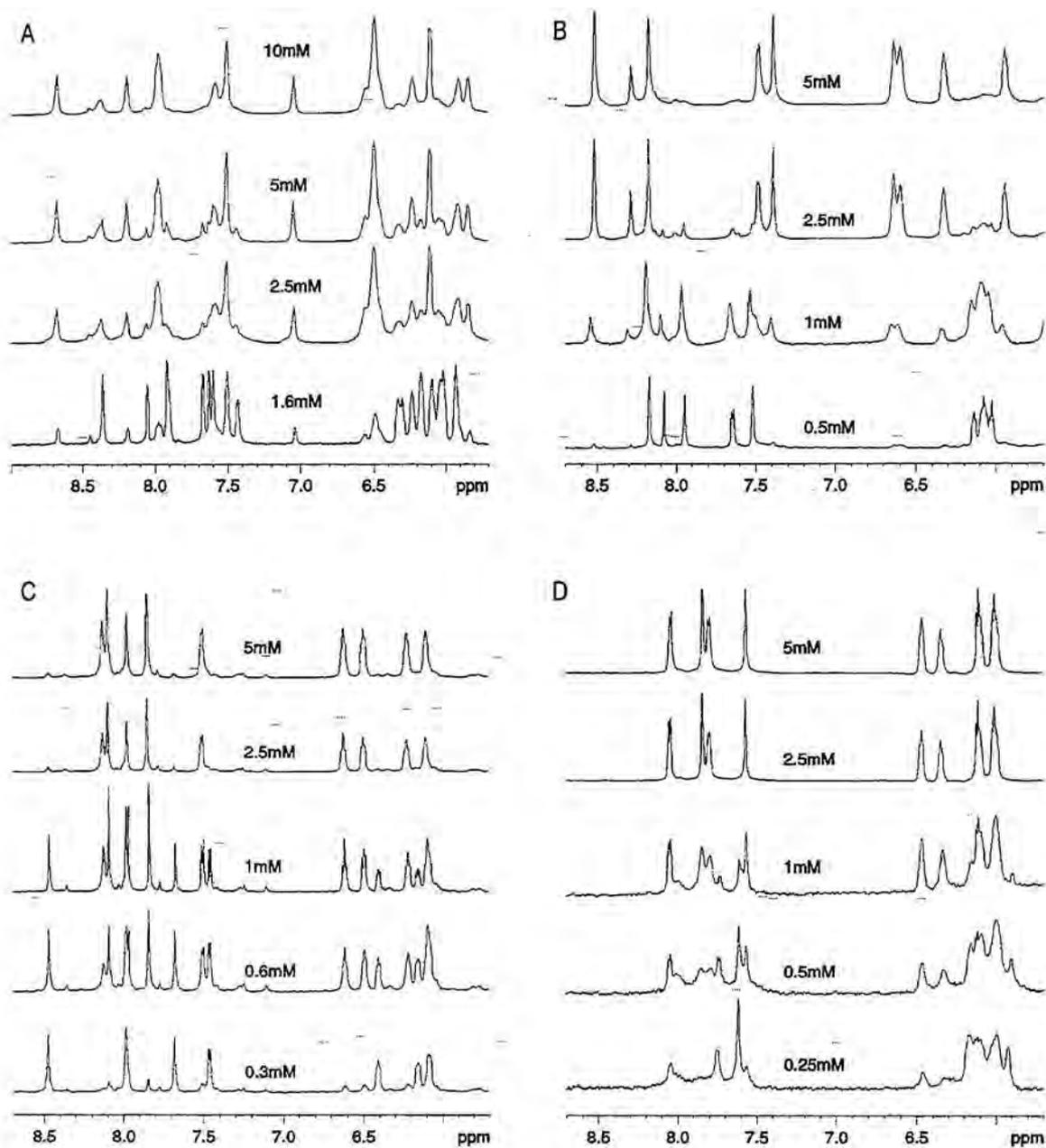
Como se discutió en el capítulo anterior, una de las evidencias más claras de la formación del bi-loop es la variación del espectro de RMN frente a la concentración de oligonucleótido. Por esta razón, y con el objeto de ver si en todos los casos se observaba la aparición de un producto de asociación, el primer experimento que se llevó a cabo con las nuevas secuencias fue la adquisición de espectros monodimensionales a diferente concentración de oligonucleótido.

En la Figura 4.2 se muestran las series de espectros 1D realizados en D<sub>2</sub>O a T=25°C para cada una de las secuencias. Las muestras se disolvieron en el volumen de D<sub>2</sub>O adecuado para obtener disoluciones 15 mM de oligonucleótido y se realizaron series de espectros monodimensionales variando la concentración de 15 a 0.3 mM. En el caso de la secuencia d<pCGCTCATT>, para la que se observó un importante ensanchamiento de las señales a 25°C, dicha serie se realizó a T=5°C.

Tal como se puede apreciar en las series de espectros, en todos los casos se observan dos grupos de señales cuya proporción relativa varía en función de la concentración de oligonucleótido, poniéndose de manifiesto la formación de un producto de asociación en todos ellos. Como ya se observó para las dos secuencias estudiadas anteriormente, no existe desdoblamiento de señales en la especie que aparece a mayor concentración de oligonucleótido, observándose el mismo número de señales que para la forma monomérica. Este hecho indica que se trata de una especie totalmente simétrica y es, por tanto, consistente con que el producto de asociación sea, también en estos casos, una estructura dimérica.

Por otro lado, que para todas las secuencias se hayan observado únicamente dos grupos de señales indica que en las secuencias d<pTGCATGCA> y d<pAGCTAGCT> se forman una única estructura monomérica y una única estructura dimérica aunque, a priori, resultaban posibles un total de cuatro estructuras (ver Figura 4.1).

La diferencia más importante observada entre las cuatro series es la concentración de oligonucleótido a la que aparece la forma dimérica en cada una de ellas. En la Tabla 4.1 se muestra este valor de concentración, así como el valor para el que la forma dimérica es la muy mayoritaria (>90%), para cada una de las secuencias. Para realizar un estudio comparativo completo se han incluido en esta tabla las dos secuencias d<pCATTTCATT> y d<pTGCTCGCT>, en las que también se había observado una marcada diferencia.



**Figura 4.2.-** Series de espectros 1D en  $D_2O$  a diferente concentración de: A)  $d\langle pCGCTCATT \rangle$ , B)  $d\langle pAGCTAGCT \rangle$ , C)  $d\langle pTGCATGCA \rangle$  y D)  $d\langle pCCGTCCGT \rangle$ . Las series se han realizado a  $T=25^\circ C$ , excepto en el caso A, que se realizó a  $T=5^\circ C$ .

**Tabla 4.1.-** Concentración de oligonucleótido a la que aparece la forma dimérica y a la que dicha forma es muy mayoritaria.

Secuencia	Aparición dímero	Dímero > 90%
cd(CATTCATT)	10 mM	>20 mM
cd(TGCTCGCT)	-	2.4-4.0 mM
cd(CGCTCATT)	5 mM	>20 mM
cd(AGCTAGCT)	0.5-1.0 mM	2.4 mM
cd(TGCATGCA)	0.3-0.6 mM	2.4 mM
cd(CCGTCCGT)	0.25-0.50 mM	2.4 mM

Los datos corresponden a valores de concentración a T=25°C.

No se observan diferencias significativas en la dependencia con la concentración entre las dos secuencias para las que la forma dimérica está estabilizada por cuatro pares G·C, d<pTGCTCGCT> y d<pCCGTCCGT>. Aunque en el caso de d<pTGCTCGCT> no se llevó a cabo una dilución progresiva de la muestra, y no se pudo determinar la concentración de aparición de la forma dimérica, el comportamiento es muy similar al observado para d<pCCGTCCGT>. En ambos casos la especie dimérica aparece a concentraciones relativamente bajas, resultando ser la especie mayoritaria a concentraciones habituales de RMN (~2mM). El hecho de que las dos secuencias d<pTGCATGCA> y d<pAGCTAGCT> presenten un comportamiento muy similar a las dos anteriores constituyó un primer indicio de que las estructuras diméricas correspondientes estaban también estabilizadas por cuatro pares G·C.

Así pues, los valores que se muestran en la Tabla 4.1 confirman la gran diferencia que existe entre las secuencias con cuatro pares G·C y aquellas en las que se han sustituido pares G·C por pares A·T. Ello pone de manifiesto que existe una relación directa entre la naturaleza de los pares de bases y la concentración de oligonucleótido a la que aparece la forma dimérica, observándose una tendencia:

$$4 \text{ pares G·C} \ll 2 \text{ pares G·C} + 2 \text{ pares A·T} \ll 4 \text{ pares A·T}$$

-  $\xrightarrow{\text{[oligonucleótido]}}$  ++

Una vez confirmada la formación de un producto de asociación para todas las secuencias, el siguiente paso consistió en determinar, mediante la asignación y análisis de los espectros de RMN, si dichos productos, presumiblemente diméricos, presentan rasgos característicos de estructura bi-loop.

#### 4.2.1.- ASIGNACIÓN DE LOS ESPECTROS.

La asignación de las cuatro nuevas secuencias se llevó a cabo de manera análoga a la descrita en el capítulo anterior para las secuencias d<pCATTCATT> y d<pTGCTCGCT>. En esta ocasión, y puesto que las condiciones de aparición del dímero así lo permiten, la asignación de las distintas formas monoméricas y diméricas se realizó a partir de espectros NOESY y TOCSY, en condiciones de -100% monómero y -100% dímero respectivamente. En condiciones en que

coexisten las dos especies, se observan numerosos picos de intercambio en los espectros NOESY, que fueron confirmados con experimentos ROESY. Estos picos de intercambio permiten confirmar la asignación de las respectivas señales.

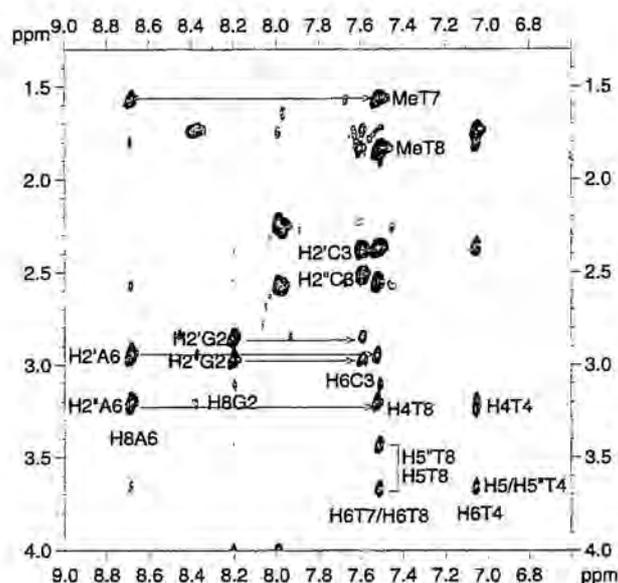
A continuación se detalla para cada secuencia los datos más relevantes que se derivan del proceso de asignación. La numeración utilizada para la asignación es la misma que se muestra en la Figura 4.1. Las tablas de asignación para las formas monoméricas y díméricas correspondientes a cada secuencia, y especie, se muestran en el Anexo B.

#### 4.2.1.1.- d<pCGCTCATT>.

En el espectro 1D realizado a una concentración de oligonucleótido de 1.6 mM, que se muestra en la Figura 2A, se observa una buena dispersión entre las señales correspondientes a las dos especies. Los espectros para esta secuencia se han realizado a baja temperatura ya que, tal como se ha comentado anteriormente, se ha observado un ensanchamiento importante de las señales al aumentar la temperatura. Este aumento de la anchura de las señales es indicativo de un intercambio rápido entre ambas especies a temperaturas a las que, en todas las demás secuencias, el equilibrio es lento en la escala de tiempos de RMN.

En este caso, el monómero no presenta estructuración, dado que esta secuencia no puede formar estructuras tipo *dumbbell* (ver Figura 4.1). La falta de estructura hace que no se observen NOEs secuenciales. No obstante, resulta posible completar la asignación de todos los sistemas de espín a partir del TOCSY, a excepción de los protones H5'/H5'' para los residuos T4 y T8. La asignación secuencial se puede realizar en base a los picos de intercambio monómero-dímero.

En el caso del dímero, éste presenta mucha similitud con los estudiados en el capítulo anterior. Se observan NOEs típicos de apilamiento entre los residuos G2-C3 y A6-T7: H2'/H2''-H6/H8 y H1'-H6/H8 y desplazamiento a campos anormalmente altos de los protones H4', H5' y H5'' de los residuos T4 y T8. También se observan los NOEs H1'-H6 entre C3-T4 y T7-T8 que indican la participación de los residuos 4 y 8 en el apilamiento. En la Figura 4.3 se muestra el camino de asignación en la zona H2'/H2''-H6/H8 para la especie dímérica.



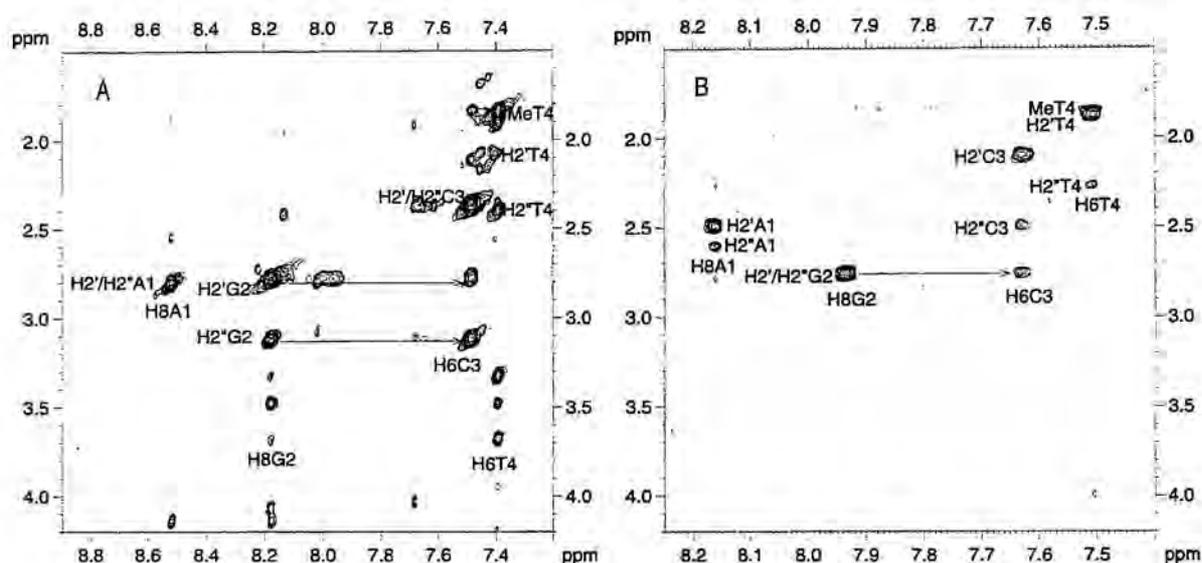
**Figura 4.3.-** Fragmento del espectro NOESY en  $D_2O$  de la especie dimérica de  $d\langle pCGCTCATT \rangle$ , 5 mM,  $T=5^\circ C$ , donde se indica el camino de asignación. Se observan los NOEs secuenciales G2-C3 y A6-T7. También se han indicado los picos de correlación H6-H4' y H6-H5'/H5'' para los residuos T4 y T8, para los que se puede observar el desplazamiento a campos anormalmente altos de los protones H4' y H5'/H5''.

#### 4.2.1.2.- $d\langle pAGCTAGCT \rangle$ .

En este caso se observa una buena dispersión de las señales correspondientes a las dos especies y se distinguen en la zona aromática las cinco señales esperadas para cada una de ellas (Figura 4.2B). Las señales con mayor dispersión corresponden, como para las dos secuencias estudiadas en el capítulo anterior, a la forma dimérica.

En los espectros NOESY se observa que ambas especies presentan estructuración, aunque en el caso del monómero, los NOEs secuenciales H2'/H2''-H6/H8 y H1'-H6/H8 propios de apilamiento no son demasiado intensos. Este hecho podría estar indicando que para esta secuencia se forma una estructura dumbbell poco estable.

En el caso del dímero, también los protones H4', H5' y H5'' de los residuos T4(8) presentan un desplazamiento a campos anormalmente altos, por lo que parece que también en este caso se forma el dímero con pares G-C. El camino de asignación para ambas especies se muestra en la Figura 4.4.



**Figura 4.4.-** Fragmento del espectro NOESY en  $D_2O$  de  $d\langle pAGCTAGCT \rangle$ : A) 5 mM,  $T=25^\circ C$  y B) 0.5 mM,  $T=25^\circ C$ . Se ha indicado el camino de asignación en la zona H6/H8-H2'/H2'', seguido para cada una de las especies.

#### 4.2.1.3.- $d\langle pTGCATGCA \rangle$ .

En este caso existe también una buena dispersión de señales de ambas especies en el espectro monodimensional. No obstante, a diferencia de lo observado para otras de las secuencias estudiadas, las que aparecen más dispersas son las correspondientes a la forma monomérica, Figura 4.2C.

Tanto la forma monomérica como la dimérica presentan espectros NOESY típicos de estructuración, con NOEs secuenciales H2'/H2''-H6/H8 y H1'-H6/H8 intensos entre los residuos G2(6) y C3(7) y, en ambos casos, se ha podido completar la asignación.

Para el dímero se observan también NOEs secuenciales entre los residuos C3(7) y A4(8), aunque menos intensos. Los protones H4', H5' y H5'' de los residuos A4(8) del dímero muestran un claro desplazamiento a campos altos. Este efecto supone un nuevo y decisivo indicio de que en la estructura dimérica se forman pares G-C y que los residuos de adenosina están efectivamente ocupando la primera posición del mini-loop.

En la Figura 4.5 se muestra el camino de asignación en la zona H6/H8-H2'/H2'', tanto para la forma dimérica como para la forma monomérica.



monomérica. Este hecho constituye un claro indicio de no estructuración y pone de manifiesto que para esta secuencia la forma monomérica no adopta una estructura tipo *dumbbell*. Además, puesto que la forma dimérica aparece a muy baja concentración, para la asignación del monómero se utilizó un espectro TOCSY realizado a una concentración 0.25mM de oligonucleótido, obteniéndose un espectro de muy baja calidad en el que apenas se observaron picos de intercambio entre las dos especies. Por este motivo únicamente se pudieron asignar los protones H6/H8, H5/H2/Me, H1' y H2'/H2'' correspondientes a cada uno de los residuos.

En el caso de la estructura dimérica el espectro NOESY presenta un gran número de picos de correlación, confirmando un alto grado de estructuración. Se observan NOEs secuenciales intensos H2'/H2''-H6/H8 y H1'-H6 H8 entre los residuos C2(6)-G3(7) y también, aunque más débiles, entre los residuos G3(7) - T4(8), que permitieron llevar a cabo la asignación secuencial.

La diferencia principal observada en estos espectros es que los protones H4', H5' y H5'' de los residuos T4(8) aparecen degenerados, hecho que no se ha observado para ninguna de las otras secuencias, y no presentan un desplazamiento tan marcado a campos altos. En la Figura 4.6 se muestra el fragmento del espectro NOESY correspondiente a la zona Ar-H2'/H2'', en la que se han incluido también los protones H4' y H5'/H5''.

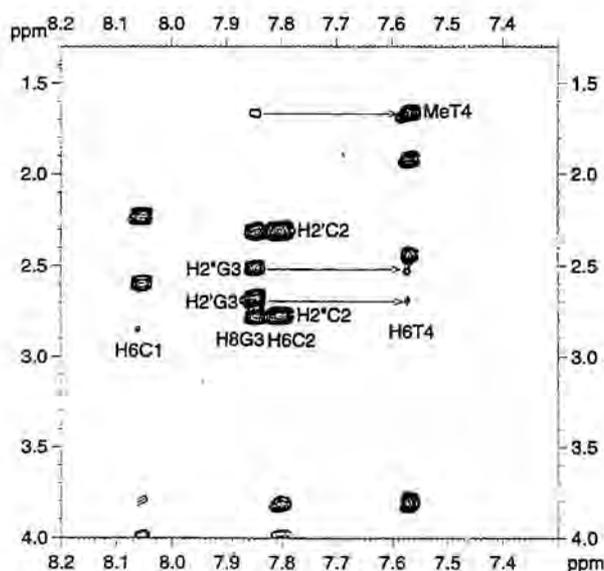


Figura 4.6.- Fragmento del espectro NOESY en D<sub>2</sub>O de *d*-pCCGTCCGT>, 2.2 mM, T=25°C, donde se ha indicado parte del camino de asignación en la zona H6/H8-H2'/H2''.

#### 4.2.2.- INFORMACIÓN ESTRUCTURAL.

Como se deriva de los datos obtenidos de los espectros de RMN, en todas las secuencias se observa una forma dimérica perfectamente estructurada. En todos los casos, tanto por los NOEs observados, como por el desplazamiento anómalo a campos altos de los protones H4' y H5'/H5'' de los residuos 4 y 8, los dímeros de estas secuencias parecen presentar rasgos estructurales muy

similares a los obtenidos en las dos secuencias del capítulo anterior. A continuación se detallan otros datos que refuerzan esta hipótesis.

En cuanto a las formas monoméricas, la situación es más dispar. Si bien está claro que los monómeros de los octámeros d<pCGCTCATT> y d<pCCGTCCGT> no están estructurados, las secuencias d<pTGCATGCA> y d<pAGCTAGCT> parecen presentar cierta estructura.

#### 4.2.2.1.- Estructuras diméricas.

La formación de pares de bases en las estructuras diméricas se confirmó con experimentos en H<sub>2</sub>O. En todos los casos aparecen señales muy estrechas de protones imino apareados. Tal como se había observado en los dos primeros octámeros, descritos en el capítulo anterior, estas señales se mantienen hasta temperaturas elevadas para los tres dímeros en los que se forman pares G-C (~60°C) y hasta una temperatura inferior para el dímero de pares mixtos (~40°C). Este comportamiento de los protones imino constituye otro indicio de la similitud entre estas estructuras y las obtenidas para d<pCATTCAATT> y d<pTGCTCGCT>.

En los espectros NOESY en H<sub>2</sub>O correspondientes a todos los dímeros se observan claramente los NOEs típicos de apareamiento Watson-Crick: H1(G2)-H41(C3), H1(G2)-H42(C3), H1(G2)-H5(C3), correspondientes a la formación de pares G-C. En el dímero de pares mixtos, d<pCGCTCATT>, se observan además los picos H3(T7)-H2(A6), H3(T7)-H61(A6) y H3(T7)-H62(A6), confirmándose la formación de dos pares A-T y dos pares G-C. En la Figura 4.7 se muestra el espectro NOESY en H<sub>2</sub>O correspondiente a esta última secuencia donde se han etiquetado los diferentes picos de correlación observados.

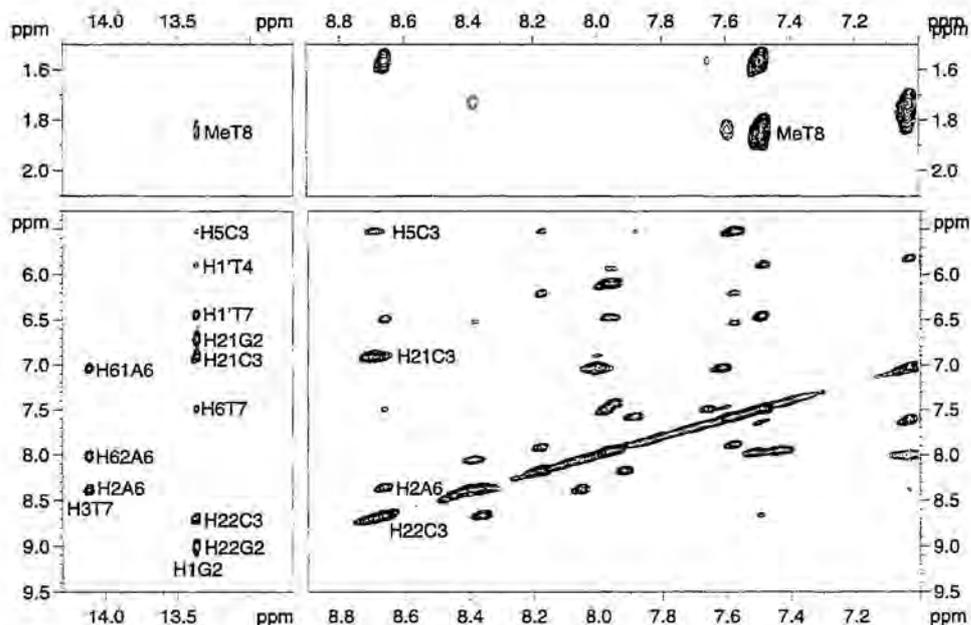


Figura 4.7.- Fragmento del espectro NOESY en H<sub>2</sub>O correspondiente a la secuencia d<pCGCTCATT>, 10mM, T=5°C.

El hecho de haberse observado únicamente apareamiento G-C tipo Watson-Crick para las secuencias d<pTGCATGCA> y d<pAGCTAGCT>, confirma definitivamente la formación de una única estructura dimérica en ambos casos y pone de manifiesto que no se forma la posible estructura estabilizada por pares A·T.

Además de este comportamiento tan similar de los protones intercambiables de todas las formas diméricas, se observan para los protones no intercambiables el mismo patrón de picos de correlación: NOEs H2'/H2"-H6/H8 y H1'-H6/H8 intensos entre los residuos 2,3 y 6,7, propios de bases apiladas, y NOEs H1'-H6/H8 entre residuos 3, 4 y 7, 8, indicativos de la implicación de los residuos 4 y 8 en el apilamiento de las bases. En las secuencias donde la primera posición del loop está ocupada por una timidina, se observan NOEs MeT4(8)-H6C3(7) y MeT4(8)-H1'C3(7) que, además de confirmar el apilamiento entre estos residuos, resultan claves en la estructura del bi-loop, ya que, a partir de las estructuras obtenidas para las dos secuencias estudiadas en el capítulo anterior, se pudo deducir que no corresponden a NOEs secuenciales sino a NOEs intersubunidad T4-C15 y T8-C11.

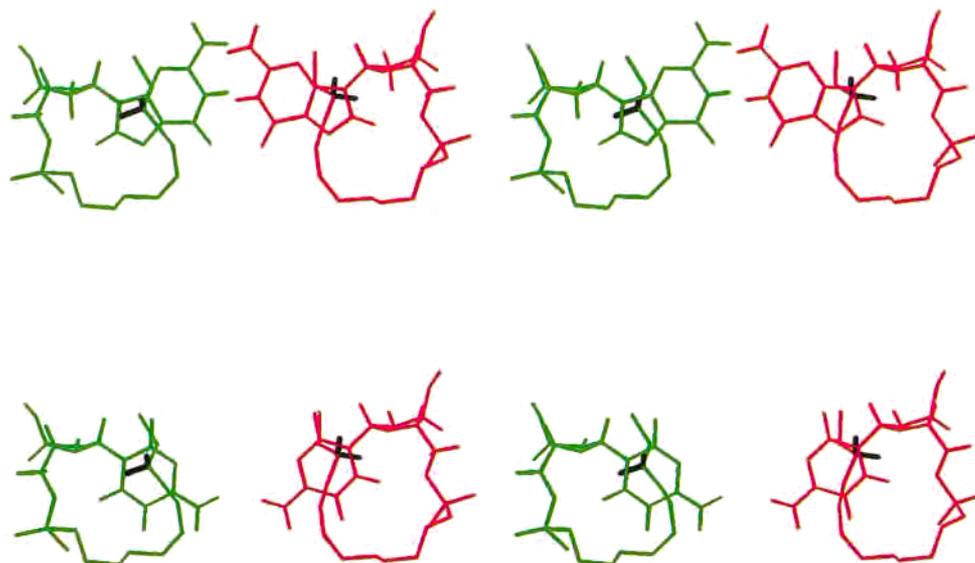
En el caso de la secuencia d<pTGCATGCA>, la presencia de NOEs H2A4(8)-H5C3(7), H2A4(8)-H6C3(7) y H8A4(8)-H6C3(7) pone de manifiesto que las bases de los residuos A4(8) están implicadas en interacciones de apilamiento, análogas a las observadas en la T4(8) de las secuencias estudiadas anteriormente. En este caso no se ha podido confirmar la existencia de un par A·A, observado en la estructura cristalina del heptámero lineal d(GCATGCT)<sup>10</sup>, ya que no ha sido posible asignar los protones amino correspondientes a estas adeninas.

La intensidad de los NOEs H1'-H6/H8 indica, en todos los casos, una disposición *anti* de la base respecto del azúcar. Este hecho, unido a los datos expuestos anteriormente, refuerza la hipótesis de una gran similitud estructural de las formas diméricas. Esta similitud se ve también reflejada en el esqueleto azúcar fosfato, ya que el análisis de las constantes de acoplamiento es consistente con una conformación tipo S de todos los anillos de ribosa, como ya se había observado para las dos secuencias estudiadas en el capítulo anterior.

Para corroborar la similitud de todas las formas diméricas, se han obtenido modelos de estos cuatro dímeros mediante simple sustitución de las bases correspondientes en la estructura promedio de d<pTGCTCGCT>. La sustitución de estas bases se ha llevado a cabo con el paquete SYBYL<sup>11</sup>, conservando los ángulos glicosídicos de la estructura original. Aunque estos modelos constituyen únicamente una aproximación, prácticamente todos los NOEs predecibles a partir de estas estructuras se observan en los espectros NOESY de cada una de ellas. Resulta interesante resaltar que estos modelos pueden incluso explicar algunas de las pequeñas diferencias observadas en los espectros de RMN. Por ejemplo, el menor apantallamiento de los protones H4' y H5'/H5" de los residuos que ocupan la primera posición del mini-loop en la secuencia d<pCCGTCCGT>. Este efecto parece estar relacionado con el hecho de haber invertido el orden de apilamiento de los dinucleótidos autocomplementarios. En este caso los protones H4' y H5'/H5" estarían localizados sobre un anillo de pirimidina en lugar de un anillo de purina, Figura 4.8. Cálculos del efecto de la

<sup>11</sup> Versión 6.0, Tripos Inc., St. Louis, MO.

corriente de anillo producidos por estas bases<sup>12</sup> predicen un desplazamiento hacia campo alto de entre 1.0 y 0.5 ppm para los protones H5' y H5'' respectivamente en el caso de d<pTGCTCGCT> y de sólo 0.4 y 0.2 ppm en el caso de d<pCCGTCCGT>.



**Figura 4.8.-** Pares estereoscópicos de las estructuras de: A) forma dimérica de d<pTGCTCGCT> y B) modelo para la forma dimérica de d<pCCGTCCGT> confeccionado a partir de la estructura anterior. Se muestran (en trazo negro grueso) los protones H5'/H5'' de los residuos que ocupan la primera posición del mini-loop (T4 y T8) junto con los anillos aromáticos de las bases 6 y 14, G en A) y C en B).

El que todas las secuencias estudiadas formen dímeros con el mismo motivo estructural tiene importantes consecuencias. Por un lado, se refuerza aún más la hipótesis de que el bi-loop puede ser un motivo estructural general en el DNA, puesto que, con estas cuatro secuencias, son ya siete las secuencias en las que se ha observado. Por otra parte, la observación de las formas diméricas en d<pCCGTCCGT> y d<pAGCTAGCT> contradice la posible secuencia consenso para la formación del bi-loop, NRYNYRYN, que había sido postulada inicialmente<sup>13</sup> (apartado 3.2.3.3, capítulo anterior).

#### 4.2.2.2.- Estructuras monoméricas.

En los espectros en H<sub>2</sub>O de las formas monoméricas, tanto de d<pCGCTCATT> como de d<pCCGTCCGT>, se observa una total ausencia de señales imino en la zona de 12-14ppm, lo que indica que no están formando pares de bases. Este dato, junto con la falta de NOEs secuenciales en los espectros de D<sub>2</sub>O, confirman la no estructuración de las formas monoméricas correspondientes a

<sup>12</sup> Case, D.A.J. *Biomolecular NMR* **1995**, 6, 341-346.

<sup>13</sup> Salisbury, S.A.; Wilson, S.E.; Powell, H.R.; Kennard, O.; Lubini, P.; Sheldrick, G.M.; Escaja, N.; Alazzouzi, E.; Grandas, A.; Pedroso, E. *Proc. Natl. Acad. Sci. USA* **1997**, 94, 5515-5518.

estas secuencias. En el caso de la secuencia con pares mixtos, d<pCGCTCATT> este resultado era de esperar, puesto que fue diseñada para que no se formasen pares de bases intramoleculares. Más interesante es el resultado para la molécula d<pCCGTCCGT>, donde el cambio de orden en la secuencia respecto a d<pTGCTCGCT>, parece alterar drásticamente la estabilidad del dumbbell.

En el caso de las secuencias d<pTGCATGCA> y d<pAGCTAGCT>, los espectros en D<sub>2</sub>O indican que estas moléculas están más estructuradas. Por un lado, se observan NOEs secuenciales entre los nucleótidos centrales de la secuencia, 2-3 y 6-7, y, por otro lado, las intensidades de los NOEs intraresiduales entre los protones H1'-H6/H8 y H2'/H2''-H6/H8 de la base indican que todos los nucleótidos se encuentran en conformación predominantemente *anti*. Un dato adicional a favor de la estructuración de estas moléculas es la variación de los desplazamientos químicos con la temperatura, que se discutirá en el Capítulo 5.

Lamentablemente, los espectros en H<sub>2</sub>O a baja concentración no son concluyentes respecto al apareamiento de bases que tiene lugar en los monómeros. Para detectar los protones imino, los espectros deben hacerse a baja temperatura (5°C), condiciones en las que, en ambos casos, existe una población importante de forma dimerica hasta concentraciones muy bajas. Puesto que, en el caso de los dímeros, la señal correspondiente a protón imino apareado es muy intensa a baja temperatura, probablemente solapa la señal del imino correspondiente a la forma monomérica.

#### 4.3.- CARACTERIZACIÓN POR ABSORCIÓN UV Y DICROÍSMO CIRCULAR DE LAS ESTRUCTURAS DIMÉRICAS Y MONOMÉRICAS.

Para detectar la formación de estructuras bi-loop de forma rápida y sin necesidad de disponer de estructuras refinadas a partir de datos de RMN, se decidió explorar la utilización de otras técnicas alternativas, como la absorción UV o el dicroísmo circular (DC). Estas técnicas, aunque no proporcionan información estructural detallada sí son extremadamente sensibles a cambios conformacionales, por lo que constituyen herramientas muy útiles en la caracterización de las diferentes estructuras que pueden adoptar los ácidos nucleicos. Ambas técnicas permiten detectar si un fragmento presenta o no estructuración pero el DC, como se podrá comprobar más adelante, permite distinguir además, en base a la posición y intensidad de las bandas, entre distintas conformaciones.

**Absorción UV.**- Durante muchos años el método utilizado para detectar cambios conformacionales en fragmentos de ácidos nucleicos ha sido el denominado de hipocromía, basado en el descenso de la absorbancia que acompaña la formación de una estructura más ordenada. Los oligonucleótidos presentan un máximo de absorción,  $\lambda_{\text{máx}}$ , a aproximadamente 260 nm y un coeficiente de extinción molar,  $\epsilon$ , del orden de  $10^4$  (dm<sup>3</sup> mol<sup>-1</sup> cm<sup>-1</sup>). Esta absorción es debida casi exclusivamente a transiciones electrónicas en las bases púricas y pirimidínicas. La intensidad y la posición de la  $\lambda_{\text{máx}}$  es función de la composición en bases del fragmento, de las interacciones de apareamiento, de la concentración salina y del pH. La existencia de apilamiento entre las bases produce un descenso de la  $\epsilon$ , conocido como *hipocromicidad*. Este fenómeno es debido a interacciones dipolo-dipolo entre las bases y depende de la estructura tridimensional del oligonucleótido. En función del tipo de estructura, la hipocromicidad puede variar en magnitud

entre un 1-11%, en el caso de desoxiribonucleótidos fosfato, y hasta un 30% para la mayoría de oligonucleótidos helicoidales.

El efecto de la estructura en la absorción UV de oligonucleótidos es muy complejo y, normalmente, sólo se hacen interpretaciones cualitativas en términos de estructuración o no estructuración de la molécula. Esta técnica se utiliza principalmente para monitorizar la fusión de fragmentos de ácidos nucleicos. A medida que la molécula pierde su estructura al aumentar la temperatura y se produce el desapilamiento de las bases, la variación de la absorción UV en función de la temperatura da lugar a una curva de fusión en la que se observa un aumento de la absorbancia. El punto medio en el aumento de la absorción al aumentar la temperatura es lo que se conoce como temperatura de fusión del DNA,  $T_m$ . Este valor de  $T_m$  indica la temperatura a la que hay un 50% de fragmento estructurado y un 50% de fragmento desestructurado, por lo que resulta un buen método para determinar la estabilidad de la estructura.

**Dicroísmo circular.**- El DC es una técnica muy sensible a cambios de conformación. Cualquier variación conformacional que se produce en el fragmento se ve reflejada en cambios en el espectro de DC. Generalmente estos cambios, en el caso de ácidos nucleicos, suelen ser enormes comparados con los que se observan en la absorbancia<sup>1-15</sup>, por lo que resulta una técnica más adecuada cuando se quieren estudiar transiciones entre diferentes conformaciones. De hecho, en la actualidad una de las aplicaciones más importantes del DC en el campo de los ácidos nucleicos es su utilización en el estudio de la interacción de ácidos nucleicos con otras moléculas y de cómo esta interacción modifica la estructura del DNA<sup>16-17</sup>. No obstante, resulta muy difícil hacer una interpretación directa de las bandas y para obtener información absoluta sobre la estructura es necesario disponer de compuestos estándar cuyas estructuras tridimensionales se conozcan mediante otros métodos.

En el caso de los ácidos nucleicos los grupos ópticamente activos observables son las purinas y las pirimidinas. Estas bases son claros ejemplos de cromóforos simétricos que se hacen ópticamente activos al enlazarse con un azúcar mediante el enlace N-glicosídico. Su actividad óptica aumenta sustancialmente cuando se convierten en parte de una estructura helicoidal.

El DC constituye una herramienta muy útil para la caracterización de transiciones entre distintas formas del DNA de doble cadena, ya que es el apilamiento helicoidal vertical de los cromóforos de las bases lo que induce cambios apreciables en las propiedades quirópticas.

La técnica del DC consiste en hacer interaccionar una molécula ópticamente activa con luz polarizada y cuantificar el cambio que se produce en la absorción para luz polarizada circularmente a la izquierda y luz polarizada circularmente a la derecha:  $\epsilon_l - \epsilon_D = \Delta\epsilon$ . Como la absorción diferencial de las ondas D e I significa que las amplitudes de las ondas transmitidas serán distintas entre sí, el resultado es luz polarizada elípticamente. Experimentalmente se suele medir  $\Delta\epsilon$ , aunque por razones

<sup>14</sup> Durand, M.; Peloille, S.; Thuong, N.T.; Maurizot, J.C. *Biochemistry* **1992**, 31, 9197-9204.

<sup>15</sup> Freifelder, «Técnicas de Bioquímica y Biología Molecular», capítulo 16, 499-529.

<sup>16</sup> Jiménez, E.; Portugal, J. *Biochemistry* **1992**, 31, 11641-11646.

<sup>17</sup> Barceló, F.; Portugal, J. *J. Biomol. Struct. & Dyn.* **1994**, 12, 203-216.

históricas se representa gráficamente la elipticidad  $\theta$ , que se relaciona con  $\Delta\epsilon$  mediante la expresión:

$$\theta = 3300\Delta\epsilon$$

Se denomina espectro de DC a la representación de la dependencia de  $\theta$  con la longitud de onda. Generalmente, se representa la elipticidad molar media:

$$[\theta]_{\lambda} = \frac{\theta_{\lambda}}{10 \cdot l \cdot c}$$

donde  $\theta_{\lambda}$  es la elipticidad observada en grados,  $l$  es la longitud del camino recorrido en centímetros y  $c$  es la concentración molar. A lo largo de este trabajo se ha representado  $[\theta]_{\lambda}$  en aquellos casos en los que se comparan espectros de DC en diferentes condiciones o correspondientes a distintas especies y se ha representado  $\Delta\epsilon$  en el caso de curvas de fusión por DC.

En este capítulo los objetivos se han centrado en utilizar estas técnicas a nivel cualitativo, únicamente para detectar los cambios conformacionales que conllevan la formación de la estructura dimérica, tanto respecto a la forma totalmente desestructurada como respecto a la forma dumbbell. No obstante, las curvas de fusiones obtenidas por DC permiten además extraer parámetros termodinámicos, lo que posibilitará analizar la estabilidad de los distintos dumbbells y dímeros tal como se describe en el capítulo siguiente.

#### 4.3.1.- ESPECTROS DE UV Y DC DE d<pCATT<CATT> Y d<pTGCT<CGCT>.

Uno de los principales motivos para incluir técnicas como la absorción UV o el DC en el presente trabajo es disponer de un criterio rápido, y que no implique la utilización de una gran cantidad de muestra, para detectar la formación de estructuras diméricas. Como ya se ha comentado anteriormente, la dificultad que supone interpretar las bandas de DC, o los cambios en el espectro UV, hace que para ello sea necesario disponer de espectros de referencia, tanto de la estructura dumbbell como de la forma bi-loop. Por este motivo, y puesto que es preciso trabajar a concentraciones muy inferiores a las de RMN, necesitábamos encontrar condiciones que nos permitieran observar la formación de la estructura dimérica a baja concentración de oligonucleótido. Para ello se utilizaron las secuencias d<pTGCT<CGCT> y d<pCATT<CATT>.

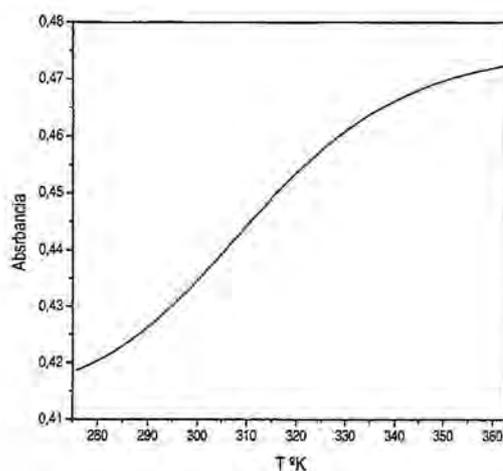
Por los resultados de RMN sabemos que, en H<sub>2</sub>O y a baja concentración de oligonucleótido existe únicamente la forma dumbbell en ambos casos. No obstante, a partir de la estructura del esqueleto de fosfatos de las formas diméricas, cabe esperar que la dimerización se vea favorecida en medios de alta fuerza iónica (ver Capítulo 3). Esto permitiría observar la forma dimérica a concentraciones bajas mediante la adición de sales en el medio. Los espectros en H<sub>2</sub>O constituirán, por una parte, espectros de referencia de la forma *dumbbell* y, por otra parte, nos servirán para observar el efecto de la adición de sales.

En la búsqueda de unas condiciones óptimas de trabajo se decidió realizar unos primeros experimentos en H<sub>2</sub>O y en un medio salino estándar en el estudio de ácidos nucleicos, que estabiliza

la formación de estructuras dúplex<sup>18</sup>: medio tamponado a pH=7, 10mM MgCl<sub>2</sub> y 100mM NaCl. Estas condiciones salinas se eligieron por la gran similitud que presenta la parte central de la forma dimerica con el ADN tipo B. Como tampón se utilizó el Na<sub>2</sub>PIPES que se encuentra descrito en la bibliografía como uno de los tampones más adecuados para la realización de experimentos con variación de temperatura, ya que su pK<sub>a</sub> presenta una menor dependencia con la temperatura que la observada para otros tampones.

#### 4.3.1.1.- Fusiones por absorción UV.

Los resultados obtenidos en las fusiones por UV de las dos secuencias se muestran en la Figura 4.9 y en la Figura 4.10. En ambos casos, las fusiones evidencian estructuración aunque se han observado diferencias importantes. Para d<pCATTTCATT> se obtuvo, tanto en H<sub>2</sub>O como en medio salino, una curva de fusión con el perfil sigmoide habitual en desnaturalizaciones de ácidos nucleicos, con un aumento de la absorbancia al aumentar la temperatura, Figura 4.9, observándose en medio salino una T<sub>m</sub> ligeramente superior a la observada en H<sub>2</sub>O.



**Figura 4.9.-** Curva de fusión por absorción UV para d<pCATTTCATT> 5.6μM en H<sub>2</sub>O (negro) y en medio salino (rojo).

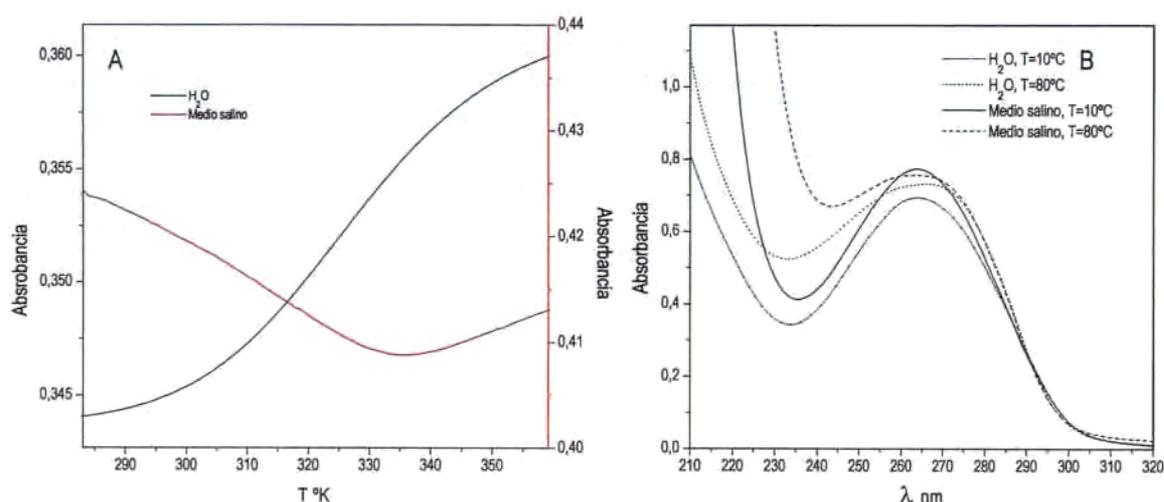
En el caso de d<pTGCTCGCT>, sin embargo, se obtuvo un resultado sorprendente. Mientras que en H<sub>2</sub>O la curva de fusión también presenta un perfil que podríamos llamar normal, en condiciones salinas se obtuvo una curva de fusión anómala en la que se produce un descenso inicial de la absorbancia para luego aumentar ligeramente a partir de ~60°C, Figura 4.10A.

Hay que señalar que esta fusión anómala no viene acompañada de ningún desplazamiento en la posición del máximo de absorción, tal como se puede observar en los espectros realizados en H<sub>2</sub>O y en medio salino, a baja y alta temperatura, que se muestran en la Figura 10B. De todos

<sup>18</sup> Kool, E.T. *Nucleic Acids Res.* **1995**, 23, 3547-3553.

modos, sí se confirma que a 260 nm, la intensidad de la absorción en agua aumenta con la temperatura mientras que en medio salino sucede lo contrario, así como que la mayor cromicidad corresponde a la especie presente en medio salino a baja temperatura. Todo ello indica que la fusión anómala en medio salino obedece a un cambio conformacional que implica hipercromicidad al aumentar la temperatura de 10 a 60°C.

Tan importante como la forma anómala de la curva de fusión en medio salino, es la baja hipocromicidad/hipercromicidad observada. Mientras que para el d<pCATTCAATT> en H<sub>2</sub>O se observa un 11% de hipocromicidad, valor que podría considerarse dentro del margen esperado, en el caso de d<pTGCTCGCT> este valor es de solamente un 4.5%. En medio salino, el cambio de absorbancia con la temperatura que presenta el d<pTGCTCGCT> es aún menor y la hipercromicidad observada es de apenas un 2.5%. Este hecho nos hizo dudar del resultado obtenido para esta secuencia, por lo que la fusión en medio salino de este producto se realizó varias veces utilizando muestras independientes y diferentes tampones, obteniéndose como resultado siempre el mismo perfil.

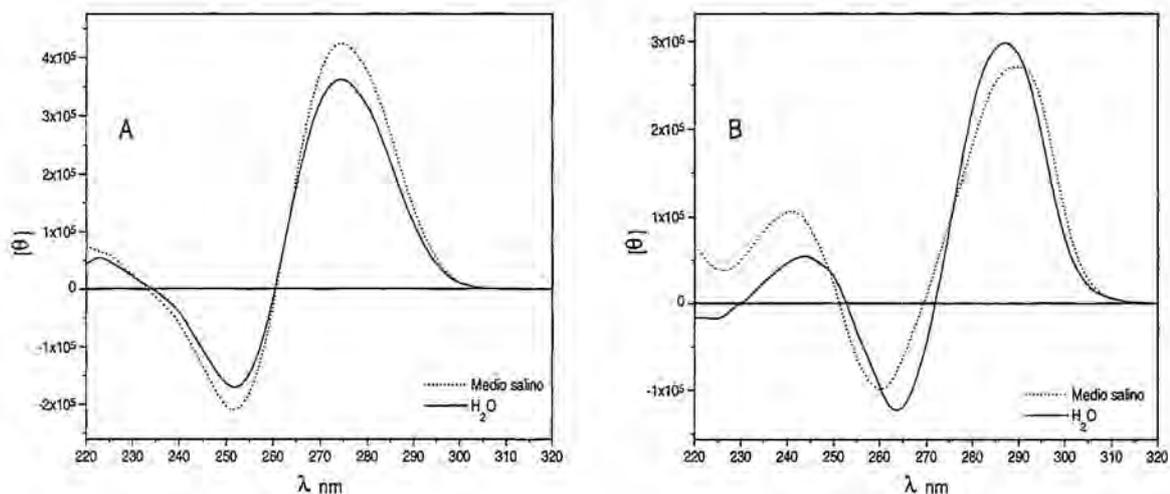


**Figura 4.10.-** A) Curvas de fusión por absorción UV de d<pTGCTCGCT> en H<sub>2</sub>O (negro) y Na<sub>2</sub>PIPES 10 mM pH=7, MgCl<sub>2</sub> 10 mM y NaCl 100 mM.(rojo). En ambos casos la concentración de oligonucleótido es de 5.8 μM. B) Espectros de UV de d<pTGCTCGCT> 5 μM en H<sub>2</sub>O y en medio salino, a baja y alta temperatura.

#### 4.3.1.2.- Espectros de DC.

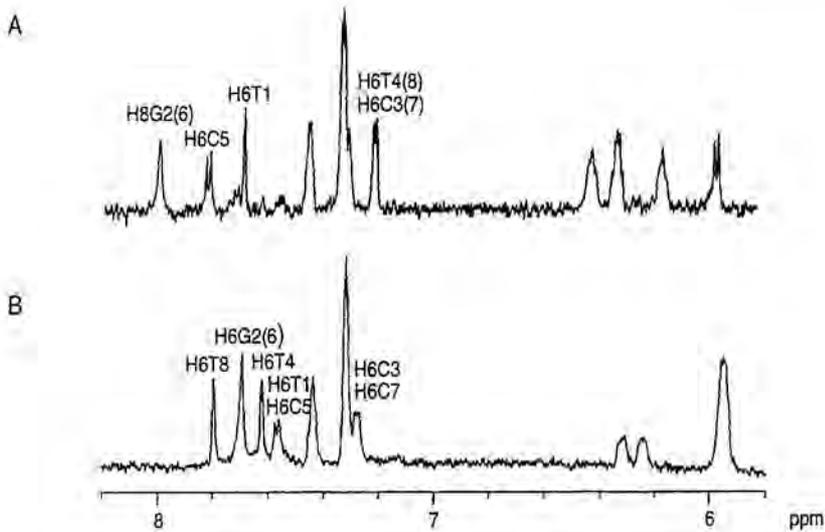
Los resultados obtenidos por DC muestran también un comportamiento diferente de ambas secuencias, Figura 4.11. Para la secuencia d<pCATTCAATT> tampoco se observaron cambios relevantes entre el espectro de DC realizado en H<sub>2</sub>O y el realizado en medio salino, observándose únicamente un aumento de la intensidad de las bandas. Sin embargo, en el caso de d<pTGCTCGCT> se observaron cambios significativos en cuanto a la posición de éstas. El

espectro de  $d\langle pTGCTCGCT \rangle$  en  $H_2O$  presenta una banda positiva muy intensa a 287 nm, una banda negativa a 263 nm y una segunda banda positiva de menor intensidad a 243.5 nm. En medio salino se produce un desplazamiento batocrómico de la banda de 287 nm a 290 nm y un corrimiento hacia longitudes de onda más bajas de las otras dos bandas hasta 260 y 241 nm, respectivamente. Este cambio en la posición de las bandas viene acompañado de un descenso de la intensidad de las dos bandas a longitudes de onda mayores y un aumento de la banda a longitud de onda menor. Puesto que, en el caso de  $d\langle pCATTCATT \rangle$ , estos cambios en la posición de las bandas no se observaron, no son en principio atribuibles a simples efectos salinos y deben corresponder a un cambio conformacional. En el caso de  $d\langle pCATTCATT \rangle$ , el aumento de la intensidad de las bandas podría estar indicando una cierta estabilización en la estructura *dumbbell*, mientras que en el caso de  $d\langle pTGCTCGCT \rangle$  todos los datos apuntaban a que las estructuras existentes en agua y medio salino eran diferentes.



**Figura 4.11.-** Espectro de DC de A)  $d\langle pCATTCAAT \rangle$  5.6  $\mu M$  y B)  $d\langle pTGCTCGCT \rangle$  5.8  $\mu M$ , en  $H_2O$  y en medio salino:  $Na_2PIPES$  10mM pH=7, 10mM  $MgCl_2$  y 100mM  $NaCl$ ,  $T=20^\circ C$ .

Con objeto de correlacionar los cambios observados, tanto por UV como por DC, con la presencia de la estructura *dumbbell* o del bi-loop, se repitieron los experimentos anteriores a mayor concentración de oligonucleótido, de manera que nos permitiera también obtener un espectro 1D de RMN. Como ya se ha indicado en el capítulo anterior, los espectros 1D de RMN correspondientes a monómero y dímero presentan señales muy características y con una buena dispersión, que permiten detectar la presencia de ambas formas. Se prepararon muestras en  $D_2O$  y medio salino deuterado, utilizando esta vez como tampón Tris-DCl a pH=7, a una concentración de oligonucleótido de 20  $\mu M$ . Los espectros 1D de RMN se muestran en la Figura 4.12.



**Figura 4.12.-** Espectros 1D de  $d\langle pTGCTCGCT \rangle$ ,  $20 \mu M$ ,  $T=25^\circ C$ , realizados en: A) medio salino:  $10 \text{ mM Tris-DCl pH}=7$ ,  $10 \text{ mM MgCl}_2$  y  $100 \text{ mM NaCl}$  y B)  $D_2O$ . Los espectros se referenciaron respecto a TSP.

Tal como se puede apreciar en la Figura 4.12, si comparamos ambos espectros con el espectro que se muestra en la Figura 2.16, las señales de la zona aromática en agua corresponden a una población de  $\sim 100\%$  monómero, mientras que en medio salino las señales corresponden a  $\sim 100\%$  dímero. Los espectros de DC y las curvas de fusión realizadas en estas mismas condiciones reproducen exactamente los resultados anteriores.

Por tanto, se pone de manifiesto que los cambios observados, tanto por UV como por DC, en el caso de la secuencia  $d\langle pTGCTCGCT \rangle$  se deben a que en  $H_2O$  se observa la estructura dumbbell, mientras que en medio salino se encuentra presente la estructura dimerica. En el caso de  $d\langle pCATTCAATT \rangle$ , sin embargo, el que no se haya apreciado ningún cambio significativo ni por UV ni por DC, indica que estamos observando la forma dumbbell tanto en  $H_2O$  como en medio salino. La adición de sal en estas condiciones, únicamente supone una estabilización de la forma *dumbbell*, que explica un ligero aumento de la  $T_m$  en estas condiciones.

Centrándonos en la secuencia  $d\langle pTGCTCGCT \rangle$ , este resultado da explicación al diferente espectro de DC obtenido en  $H_2O$  y en medio salino en cuanto a la posición y intensidad en las bandas de DC, puesto que se trata de dos estructuras diferentes.

Ahora bien, la curva de fusión por UV continúa resultando sorprendente. Esta curva de fusión, en la que se produce una disminución inicial de la absorbancia, seguida posteriormente de un ligero aumento, podría explicarse en base a la existencia de equilibrio entre las tres formas: dímero-monómero-random coil. Puesto que el monómero estructurado presenta una menor absorción que el dímero, el descenso inicial de la absorbancia al aumentar la temperatura podría ser debido al desplazamiento del equilibrio monómero-dímero hacia la formación de monómero. Posteriormente el aumento de absorbancia respondería a la fusión de la forma monomérica. El estudio del equilibrio monómero-dímero, que se detalla en el capítulo siguiente, permitirá

determinar si, en efecto, el aumento de temperatura desplaza el equilibrio hacia la forma *dumbbell*. No obstante, de ser cierta esta hipótesis y si esta fusión 'anómala' resulta ser un rasgo común a todos los dímeros, en el caso de  $d\langle pCGCTCATT \rangle$  y  $d\langle pCCGTCCGT \rangle$ , cuyas formas monoméricas no presentan estructuración, debería observarse únicamente una disminución de la absorbancia al aumentar la temperatura.

#### 4.3.1.3.- Efecto de la sal en el espectro de DC.

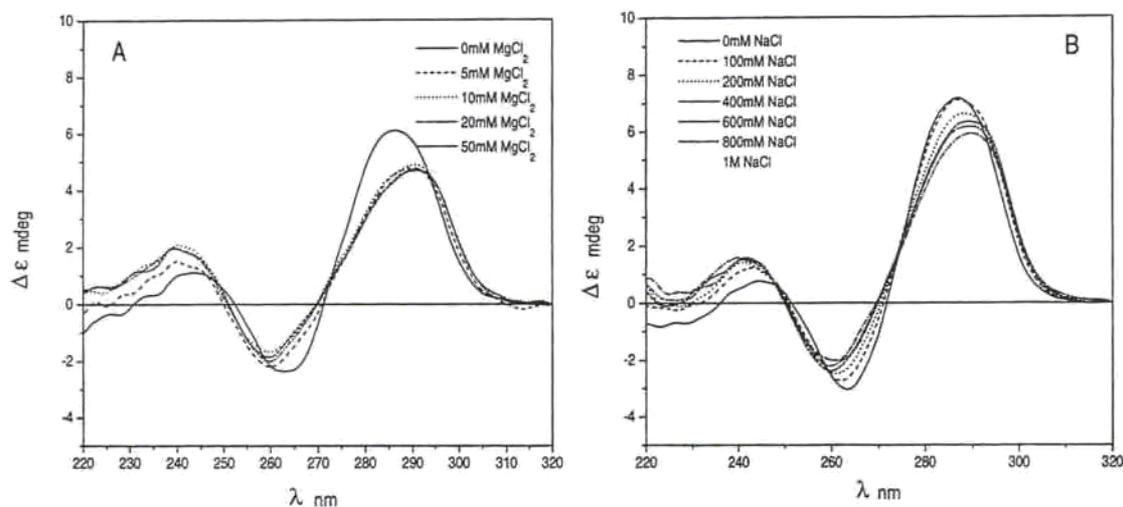
De estos primeros experimentos podemos concluir que, en efecto, mediante la adición de sal al medio se puede inducir la formación de la forma dimérica. Ahora bien, esto sólo ha resultado cierto en el caso del  $d\langle pTGCTCGCT \rangle$ . El diferente comportamiento de las secuencias  $d\langle pCATTCAATT \rangle$  y  $d\langle pTGCTCGCT \rangle$  con la concentración salina puede estar relacionado con el hecho de que la estructura dimérica aparezca en ambos casos a concentraciones muy diferentes. Si esto es así las condiciones anteriormente descritas resultan únicamente válidas para aquellas secuencias en las que se formen pares G-C. En el caso de las secuencias con algún par A-T,  $d\langle pCATTCAATT \rangle$  y  $d\langle pCGCTCATT \rangle$ , sería necesario encontrar unas condiciones más adecuadas en las que realizar el estudio.

Puesto que los primeros experimentos se habían realizado utilizando un medio salino descrito para estabilizar dúplex, decidimos estudiar primero qué efecto tiene sobre la formación del dímero cada una de las sales por separado. Para ello realizamos una valoración de  $d\langle pTGCTCGCT \rangle$  con  $MgCl_2$  y  $NaCl$ . Estas valoraciones permitirían, por una parte, ver el grado de influencia de cada catión en la formación de la estructura dimérica y, por otra parte, determinar las condiciones óptimas de trabajo para realizar los espectros de DC. Las series de espectros a diferente concentración salina, para cada una de las sales se muestran en la Figura 4.13.

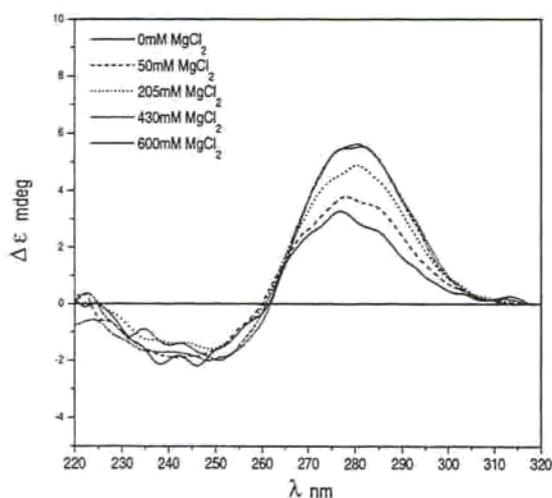
En ambos casos se observan puntos isodicroicos que indican la existencia del cambio conformacional inducido por la adición de sal en el medio. No obstante, la presencia de  $MgCl_2$  parece ser más decisiva en la formación de estructuras diméricas. La adición de pequeñas cantidades de esta sal induce la formación del dímero a concentraciones de oligonucleótido tan baja como  $5\mu M$ . Por el contrario, para llegar a obtener el mismo efecto adicionando  $NaCl$  resulta necesario llegar a concentraciones de  $\sim 800mM$ , por lo que no parece que la presencia de esta sal resulte crítica para la formación de la forma dimérica. Al no observarse diferencias entre los espectros realizados a  $10mM$  de  $MgCl_2$  y los realizados a una concentración superior de esta sal, se decidió mantener, como concentración para el estudio de las distintas secuencias con pares G-C una concentración de  $10mM$   $MgCl_2$  en el medio.

El hecho de que el  $MgCl_2$  sea más decisivo para la formación del dímero resulta curioso teniendo en cuenta que en el centro de la estructura de RX de  $d\langle pCATTCAATT \rangle$  se encuentra coordinado un ión  $Na^+$ . Así pues, sería necesario profundizar en el estudio de cómo y por qué diferentes cationes inducen o no la formación de estas estructuras.

Dada la mayor influencia de la presencia de magnesio en el medio, para encontrar unas condiciones que nos permitieran observar la forma dimérica de las secuencias con pares A-T, se realizó únicamente una valoración del producto  $d\langle pCGCTCATT \rangle$  con  $MgCl_2$ , Figura 4.14.



**Figura 4.13.-** Serie de espectros de DC de  $d\langle pTGCTCGCT \rangle 5\mu M$  a diferente concentración salina de A)  $MgCl_2$  y B)  $NaCl$ . En ambos casos los espectros se han realizado en medio tamponado  $Na_2PIPES$ ,  $pH=7$  y  $T=20^\circ C$ .



**Figura 4.14.-** Serie de espectros de DC de  $d\langle pCGCTCATT \rangle 5\mu M$  a diferente concentración de  $MgCl_2$ .  $Na_2PIPES$  10 mM,  $pH=7$ ,  $T=5^\circ C$ .

La serie de espectros a distinta concentración de  $MgCl_2$  muestra que, en efecto, la adición de la sal induce también en este caso la formación de una forma estructurada que, dado que no resulta posible la formación de una estructura dumbbell, únicamente puede corresponder a la forma dimérica. En la valoración no se observan cambios entre el espectro realizado a 430 mM de  $MgCl_2$

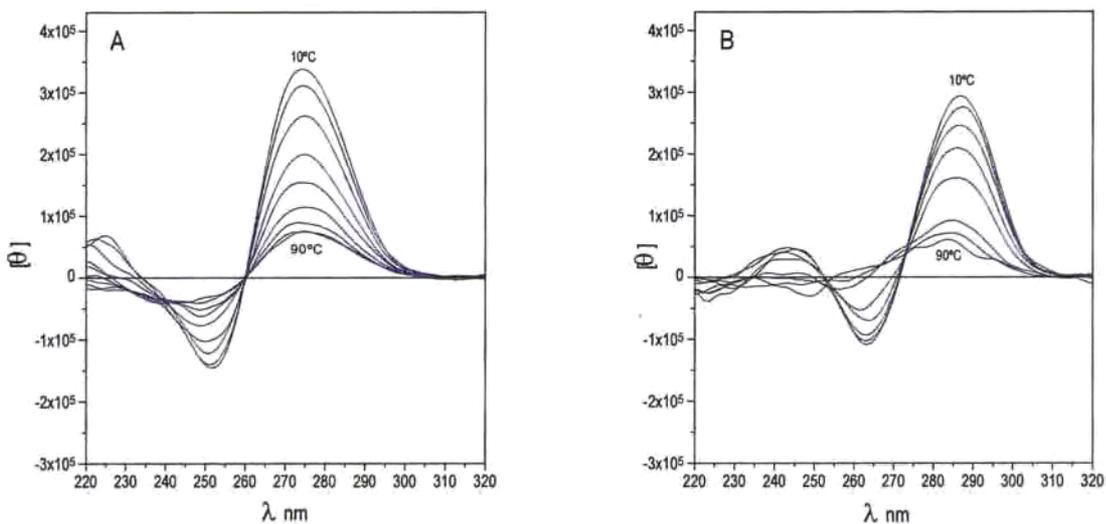
y el realizado a 600 mM, por lo que se decidió utilizar una concentración de 500 mM de  $MgCl_2$  para los productos que presentaban pares A·T.

Aunque la formación de la estructura dimérica quedaba garantizada por la presencia de magnesio, se decidió mantener igualmente una concentración de 100 mM de sodio en el medio en todos los casos. Así pues el posterior estudio por UV y DC de todas las secuencias se realizó en condiciones: 10 mM  $Na_2PIPES$  pH=7, 10 mM  $MgCl_2$  y 100 mM NaCl, para las secuencias en las que se forman pares G·C, y 10 mM  $Na_2PIPES$  pH=7, 500 mM  $MgCl_2$  y 100 mM NaCl para las secuencias estabilizadas con pares A·T.

Para que los resultados fueran totalmente comparables, los estudios mediante ambas técnicas se realizaron sobre las mismas muestras a una concentración de oligonucleótido aproximadamente 5  $\mu M$ .

#### 4.3.1.4.- Efecto de la temperatura en el espectro de DC.

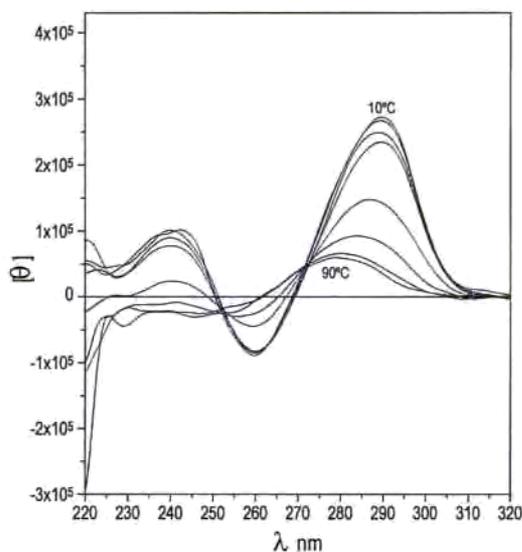
En la sección anterior se demuestra que la variación del espectro de DC permite observar un cambio conformacional inducido por la presencia de sal en el medio. Es de esperar que el DC nos permita también detectar cambios conformacionales inducidos por la temperatura. Esto es particularmente importante en las formas diméricas, para las que las curvas de fusión de UV son anómalas y resultan difíciles de interpretar. Para ello se realizaron espectros de DC en  $H_2O$  y en medio salino a diferente temperatura.



**Figura 4.15.-** Serie de espectros de DC a diferentes temperaturas en  $H_2O$  para A)  $d\langle pCATT\langle math>CATT\langle math>\rangle\rangle$  y B)  $d\langle pTGCT\langle math>CGCT\langle math>\rangle\rangle$ .

En la Figura 4.15 se muestran los espectros de DC a diferentes temperaturas para  $d\langle pCATT\langle math>CATT\langle math>\rangle\rangle$  y  $d\langle pTGCT\langle math>CGCT\langle math>\rangle\rangle$  en  $H_2O$ . En ambas secuencias se puede observar como las

bandas de DC disminuyen drásticamente con la temperatura. Más interesante aún es la variación del espectro de DC en presencia de sales. En la Figura 4.16 se puede observar como la banda con máximo a 290nm disminuye y se desplaza hacia longitudes de onda más cortas a medida que aumenta la temperatura. A altas temperaturas, los espectros de DC de  $d\langle pTGCTCGCT \rangle$ , en presencia o ausencia de sales, son prácticamente idénticos.



**Figura 4.16.-** Serie de espectros de DC a diferentes temperaturas en medio salino para la secuencia  $d\langle pTGCTCGCT \rangle$ .

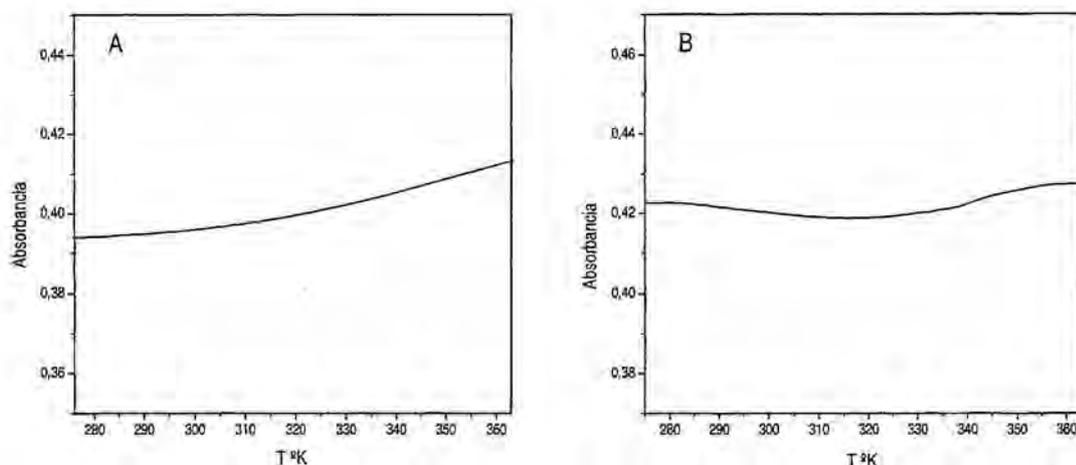
Estas series, en las que se puede apreciar claramente la variación del espectro de DC con la temperatura, constituyen una evidencia más de que tanto en  $H_2O$  como en medio salino estamos observando formas perfectamente estructuradas. El hecho de que, para una misma secuencia, se obtenga el mismo espectro de DC a alta temperatura en  $H_2O$  y en medio salino partiendo de espectros de DC distintos, pone de manifiesto que se trata de dos formas estructuradas diferentes, dumbbell y bi-loop respectivamente, que al aumentar la temperatura convergen a una misma forma desestructurada.

Midiendo la disminución con la temperatura de la elipticidad a  $\lambda_{m\acute{a}x}$  podemos seguir las fusiones tanto de las formas dimericas como monomericas con mucha más fiabilidad en nuestro caso que con las curvas de UV. En el próximo capítulo utilizaremos este tipo de curvas para obtener parámetros termodinámicos, pero antes debemos comprobar estos resultados en las otras secuencias estudiadas.

#### 4.3.2.- EFECTO DE LA SECUENCIA EN LAS CURVAS DE FUSIÓN POR ABSORCIÓN UV.

De cara a comprobar si todas las secuencias presentaban el mismo comportamiento por absorción UV, es decir si todos los dumbbells presentaban fusiones normales y todos los dímeros

fusiones anómalas, se completaron las curvas de fusión en  $H_2O$  y en medio salino de todas las secuencias. En la Figura 4.17 se muestran las fusiones correspondientes a las formas monoméricas de  $d\langle pAGCTAGCT \rangle$  y  $d\langle pTGCATGCA \rangle$  realizadas en  $H_2O$ . Para las secuencias  $d\langle pCCGTCCGT \rangle$  y  $d\langle pCGCTCATT \rangle$ , puesto que no tienen una forma monomérica estructurada, no se observó ningún cambio de absorbancia con la temperatura.

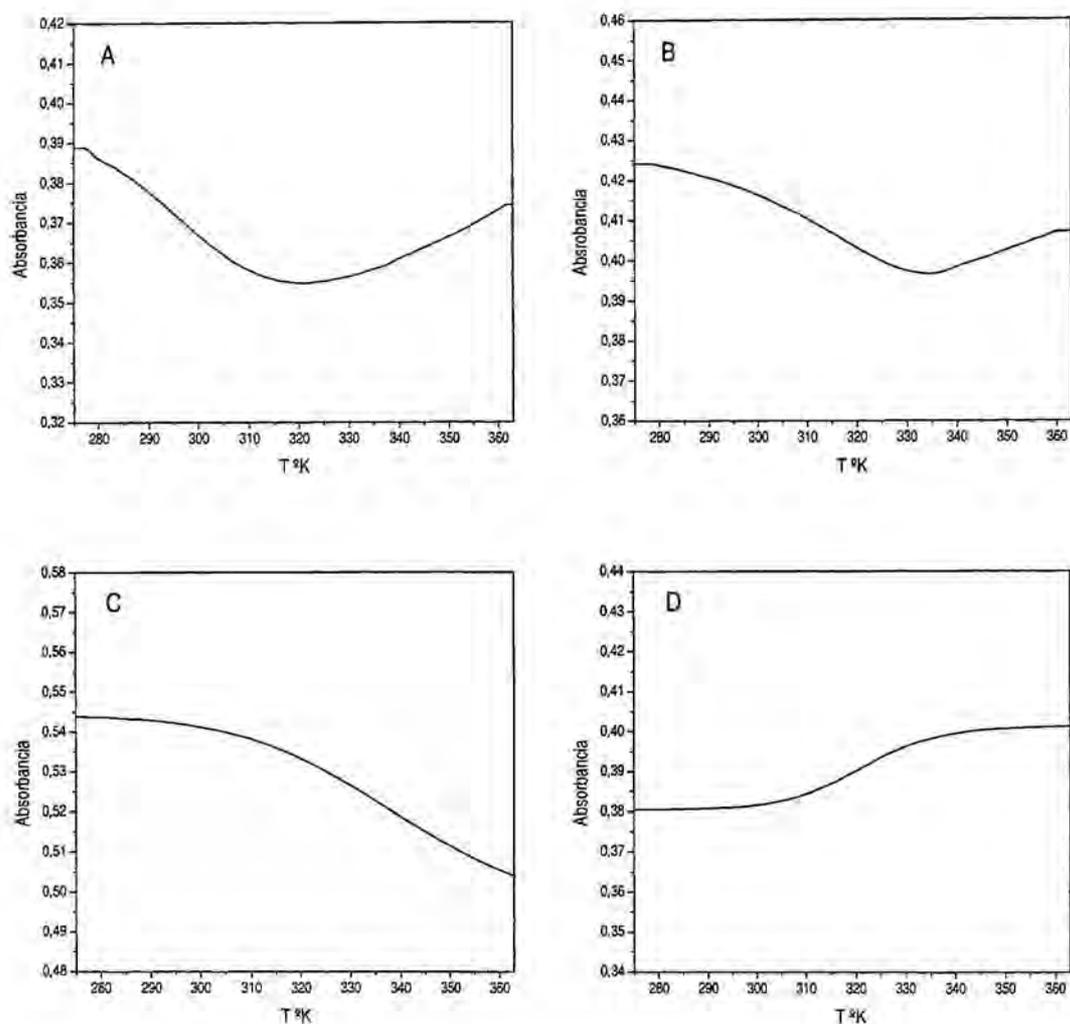


**Figura 4.17.-** Fusión por UV de A)  $d\langle pAGCTAGCT \rangle$   $5.5\mu M$  y B)  $d\langle pTGCATGCA \rangle$ ,  $5.8\mu M$  en  $H_2O$ .

En el caso de  $d\langle pAGCTAGCT \rangle$ , se observa una ligera hipocromicidad a baja temperatura y un perfil no sigmoide, que sugieren que la forma dumbbell no es muy estable. Este resultado es consistente con los espectros de RMN, en los que si bien se habían observado picos secuenciales propios de estructuración, se trataba de picos muy poco intensos.

Resulta, sin embargo, sorprendente la fusión obtenida para  $d\langle pTGCATGCA \rangle$  en  $H_2O$ . De tratarse del monómero estructurado se debería observar una curva de fusión normal, como la obtenida para  $d\langle pTGCTCGCT \rangle$  y  $d\langle pCATT CATT \rangle$ . Si por el contrario se trata de un monómero poco estable debería obtenerse un resultado similar al obtenido para  $d\langle pAGCTAGCT \rangle$ . No obstante, en la curva apenas existe variación de la absorbancia, observándose un ligero descenso hasta  $\sim 320^\circ K$  para aumentar, también ligeramente, a partir de esta temperatura. Este resultado recuerda el obtenido para  $d\langle pTGCTCGCT \rangle$  en medio salino y no resulta fácilmente interpretable si se excluye la presencia de dímero a baja temperatura.

A continuación se muestran las fusiones realizadas para cada una de las secuencias en medio salino, Figura 4.18.



**Figura 4.18.-** Curva de fusión por UV de A)  $d<pAGCTAGCT>$ , B)  $d<pTGCATGCA>$ , C)  $d<pCGCTCATT>$  y D)  $d<pCCGTCCGT>$  en medio salino.

De las fusiones en medio salino se podría deducir que la hiper Cromicidad parece ser un rasgo común de la mayoría de las formas díméricas y que la disminución de absorbancia observada responde al equilibrio con la forma monomérica estructurada. A favor de esta hipótesis se encuentra el hecho de que para la secuencia  $d<pCGCTCATT>$ , para la que no es posible la formación de una estructura *dumbbell*, se registra una curva donde únicamente se observa un descenso de la absorbancia con la temperatura y no la subida final observada en aquellas secuencias para las que sí que se forma la estructura *dumbbell*. Pero nuevamente nos encontramos con una excepción. Para la secuencia  $d<pCCGTCCGT>$ , la forma dímica presenta una curva de fusión "normal", es decir, que presenta un aumento de cromicidad con la temperatura. De los datos de RMN parecía deducirse que, si bien podían existir pequeñas diferencias, la estructura de todos los dímeros era básicamente

la misma. No obstante esta gran diferencia mostrada por el dímero en el que se ha invertido el orden de los dinucleótidos, parece implicar, por lo menos en este caso, diferencias importantes en el apilamiento de las bases, que es en definitiva el máximo responsable del efecto que se observa durante la fusión.

Este último resultado pone de manifiesto que el hecho de obtener curvas de fusión 'anómalas', aunque puede suponer un indicio, no puede considerarse un criterio general para detectar la formación de estructuras diméricas. Aunque este criterio fuera válido, nos encontramos con el problema de que la variación observada por UV entre las distintas formas posibles, dímero, dumbbell y monómero desestructurado, reflejan cambios en el espectro UV de muy pequeña magnitud y no permiten extraer conclusiones claras sobre la estabilidad de las formas estructuradas. En este sentido, el DC ha resultado en nuestro caso una herramienta más útil para caracterizar la forma dumbbell y la forma dimérica.

#### 4.3.3.- ESTUDIO POR DC.

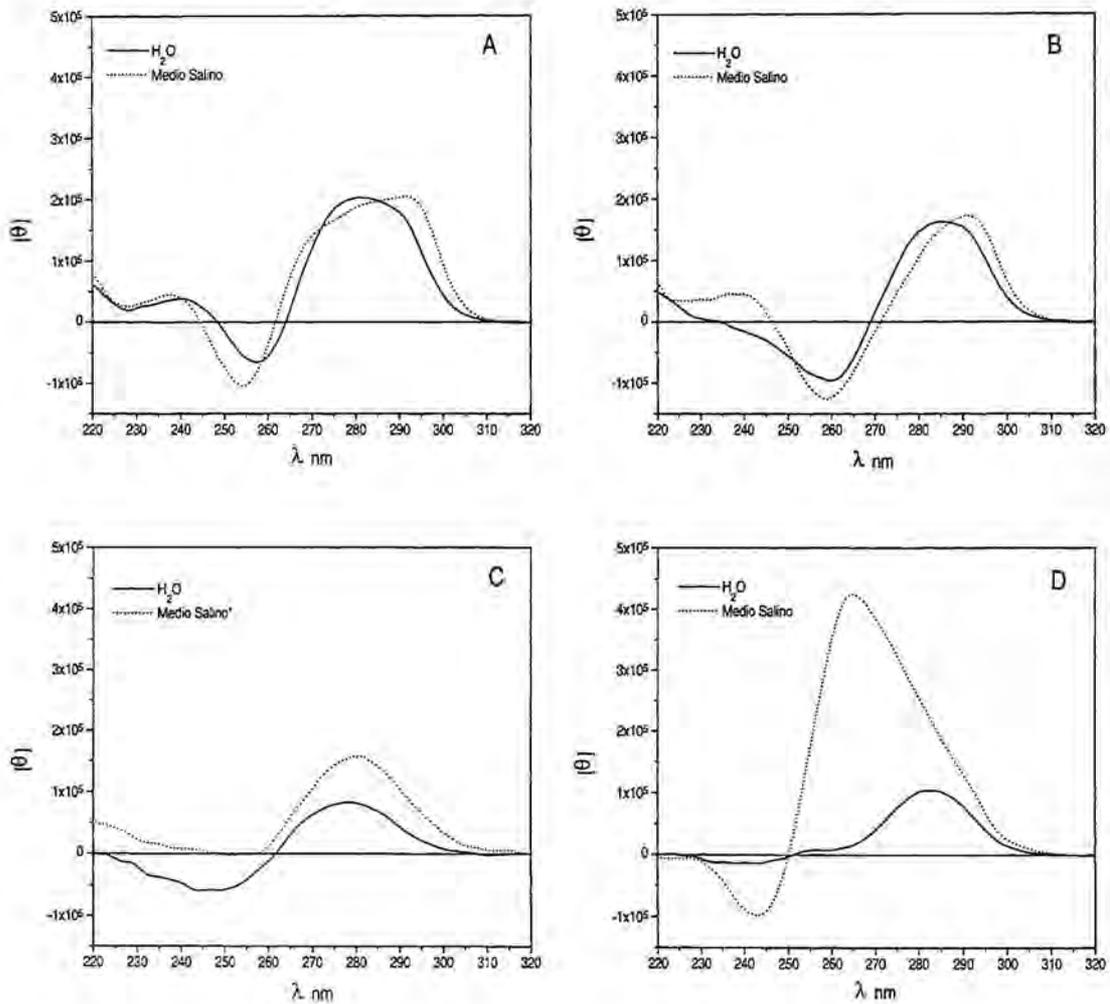
En la Figura 4.19 se muestran los espectros de DC en H<sub>2</sub>O y en medio salino para las secuencias d<pTGCATGCA>, d<pAGCTAGCT>, d<pCGCTCATT> y d<pCCGTCCGT>. En ellos se pueden observar diferencias importantes entre las distintas secuencias. Aunque aparecen cambios significativos entre el espectro realizado en H<sub>2</sub>O y en medio salino, estos cambios no son generalizables y no resulta posible establecer qué cambios deben producirse en el espectro de DC al formarse la estructura dimérica.

En el caso de d<pCGCTCATT>, Figura 19C, aunque se había realizado previamente una valoración con MgCl<sub>2</sub> y parecía que habíamos llegado a observar la forma dimérica a una concentración de sal ~500 mM, el espectro de DC no parece reflejar la aparición de una forma bien estructurada. En este caso, puesto que no se observaron cambios en el espectro a mayores concentraciones de sal, ver Figura 4.14, tal vez sería necesario trabajar a una concentración mayor de oligonucleótido para aumentar la estabilidad de la forma dimérica.

Para las tres secuencias en las que se forman pares GC y para las cuales se habían utilizado las condiciones ensayadas con d<pTGCTCGCT>, decidimos comprobar que la adición de sal garantizaba la formación de la estructura dimérica y que una mayor adición de magnesio no inducía a cambios en el espectro de DC. Para ello, una vez realizados los espectros de DC en condiciones de 10mM mgCl<sub>2</sub> se añadió MgCl<sub>2</sub> en el medio hasta obtener una concentración 50mM de dicha sal. En los tres casos el espectro realizado a 10mM de magnesio es totalmente superponible con el realizado a 50mM por lo que asumimos que las condiciones: 10mM Na<sub>2</sub>PIPES pH=7, 10mM MgCl<sub>2</sub> y 100mM NaCl son apropiadas para el estudio y que, en efecto, estamos en condiciones de máxima estructuración.

La secuencia d<pAGCTAGCT> podría decirse que presenta un comportamiento que coincide con el anteriormente observado para la secuencia d<pTGCTCGCT>, con el desplazamiento de la banda a  $\lambda_{m\acute{a}x}$  hacia longitudes de onda mayores como dato más característico. En el caso de d<pTGCATGCA>, el espectro es parecido pero las bandas, tanto en H<sub>2</sub>O como en medio salino, son más anchas.

En el caso de  $d\langle p\text{CCGTCCGT} \rangle$ , en cambio, se observa un comportamiento completamente distinto. Sorprendentemente la adición de sal induce la aparición de una forma estructurada, presumiblemente el dímero, para la que la  $\lambda_{\text{máx}}$  aparece a 265nm. El hecho de que se observe un comportamiento tan diferente en esta secuencia impide obtener conclusiones generales sobre los cambios que deberían observarse en las bandas de DC al formarse la estructura dimérica. Recordemos que el comportamiento de esta estructura, "excepcional" respecto a las demás estudiadas, se observa también en la curva de fusión por UV.



**Figura 4.19.-** Espectros de DC en  $\text{H}_2\text{O}$  y en medio salino de A)  $d\langle p\text{TGCATGCA} \rangle$  B)  $d\langle p\text{AGCTAGCT} \rangle$  C)  $d\langle p\text{CGCTCATT} \rangle$  y D)  $d\langle p\text{CCGTCCGT} \rangle$

Si bien los espectros de DC muestran que, en efecto, la adición de sal induce un cambio conformacional en las secuencias, era necesario evaluar si el DC, también para estas secuencias, nos permitiría estudiar el comportamiento de las formas diméricas y monoméricas con la temperatura y, poder extraer conclusiones en cuanto a su estabilidad. Para ello, se realizaron espectros en  $\text{H}_2\text{O}$  y en

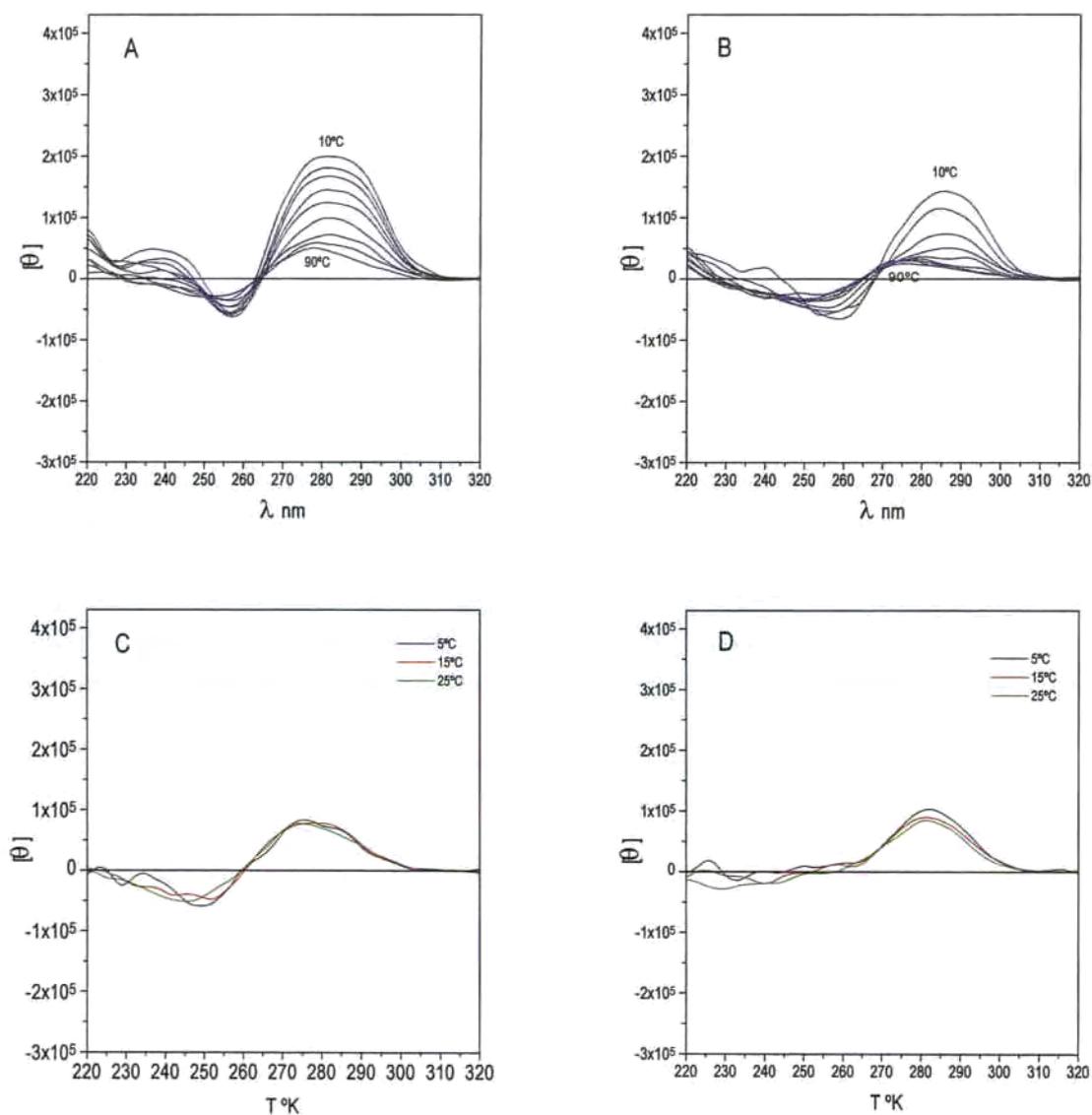
medio salino a diferente temperatura. En el caso de las secuencias d<pCCGTCCGT> y d<pCGCTCATT>, para las que no tiene lugar la formación de *dumbbell*, el espectro de DC en H<sub>2</sub>O no presenta variación con la temperatura, Figura 4.20.

En la Figura 4.20 se muestran las series de espectros en H<sub>2</sub>O a diferente temperatura para las secuencias d<pTGCATGCA> y d<pAGCTAGCT>. En ambos casos los espectros reflejan que se trata de formas monoméricas perfectamente estructuradas. El aumento de la temperatura produce además del esperado descenso de la elipticidad, un ligero desplazamiento de la banda a  $\lambda_{\text{máx}}$  hacia longitudes de onda menores. Se observa, sin embargo, una diferencia importante entre ambas series en cuanto a estabilidad térmica de la forma monomérica, resultando ser mucho más estable la correspondiente a la secuencia d<pTGCATGCA>. Este hecho confirma la importancia que tiene, para la estabilidad de las formas *dumbbell*, la composición de los mini *loops* y la posición que ocupan cada uno de los residuos.

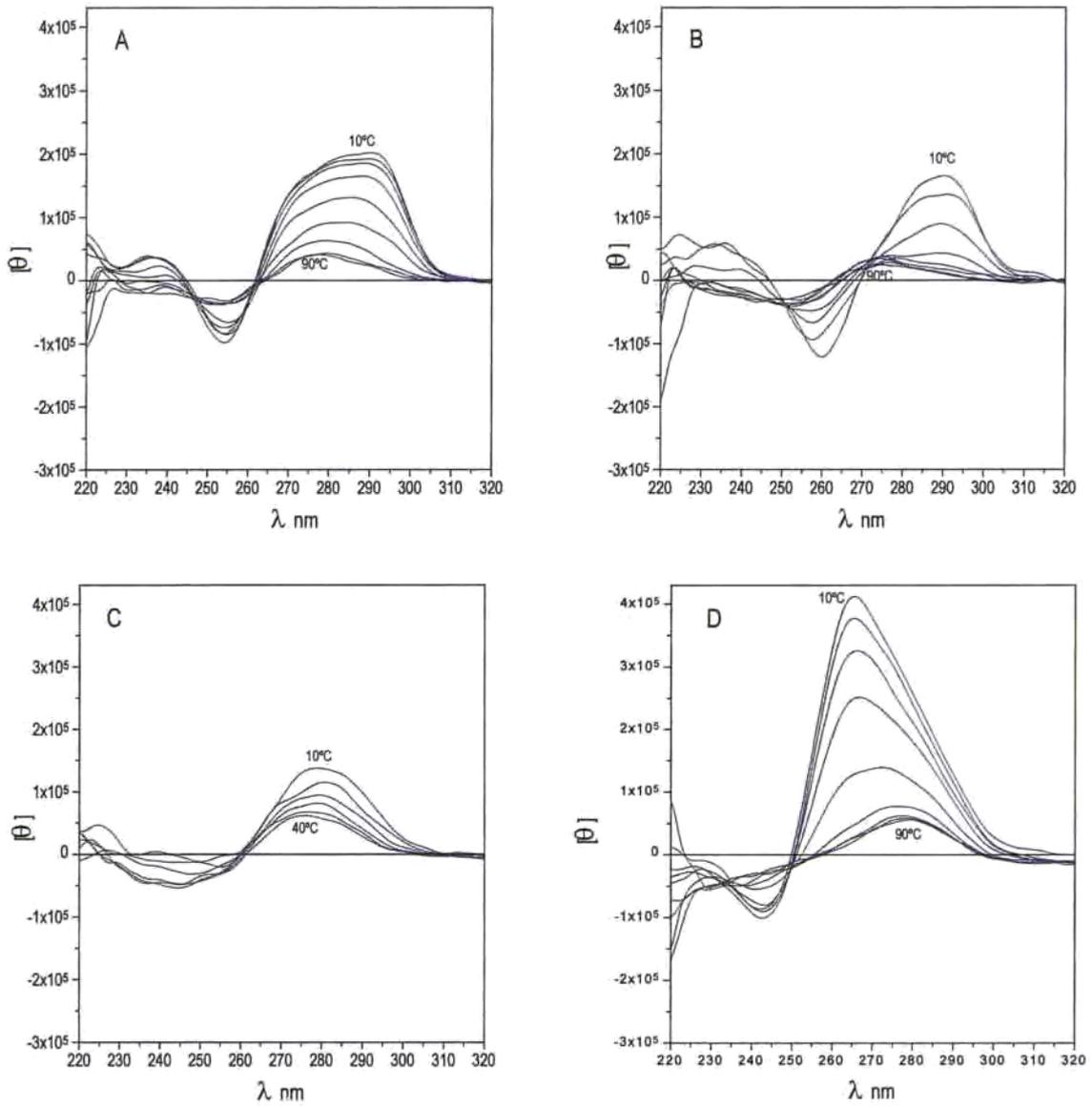
En la Figura 4.21 se muestran los espectros de DC en medio salino a diferentes temperaturas. A pesar de que el espectro de DC de las formas diméricas parece depender de la secuencia, resulta interesante observar que, en todos los casos, converge hacia un único tipo de espectro a medida que aumenta la temperatura y pierde su estructura.

En el caso de d<pCGCTCATT> se confirma que se trata de una forma dimérica poco estable, ya que a 40°C se obtiene ya el espectro correspondiente a la forma totalmente desestructurada, y que las condiciones elegidas no son seguramente las más adecuadas para el estudio de esta secuencia.

A la vista de estos resultados, la conclusión del apartado 4.3.1.4.- parece ser válida para todas las secuencias y, por tanto, hace de la variación del espectro de DC con la temperatura la herramienta más útil para estudiar la formación de la estructura bi-loop. A diferencia de lo observado por absorción UV, las bandas de DC presentan una importante variación con la temperatura ( $\Delta\theta \sim 1.5-3.5 \times 10^5$ ), con lo que la realización de curvas de fusión por DC nos permitirá determinar parámetros termodinámicos, tal como se describe en el capítulo siguiente.



**Figura 4.20.-** Serie de espectros de DC a diferentes temperaturas para las formas monoméricas de A)  $d\langle pTGCATGCA \rangle$ , B)  $d\langle pAGCTAGCT \rangle$ , C)  $d\langle pCGTCATT \rangle$  y D)  $d\langle pCCGTCCGT \rangle$ . Espectros realizados en  $H_2O$ .



**Figura 4.21.-** Serie de espectros de DC frente a temperatura para las formas diméricas de: A)  $d\langle pTGCATGCA \rangle$ , B)  $d\langle pAGCTAGCT \rangle$ , C)  $d\langle pCGCTCATT \rangle$  y D)  $d\langle pCCGTCCGT \rangle$ . Espectros realizados en medio salino.

---

**CAPÍTULO 5**

***ESTUDIO TERMODINÁMICO DEL EQUILIBRIO DE FORMACIÓN DE ESTRUCTURAS  
BI-LOOP.***

---

## 5.- ESTUDIO TERMODINÁMICO DEL EQUILIBRIO DE FORMACIÓN DE ESTRUCTURAS BI-LOOP.

### 5.0.- INTRODUCCIÓN Y OBJETIVOS.

Dado que la formación de estructuras bi-loop no constituye un ejemplo aislado sino que parece tratarse de un motivo de carácter general, consideramos importante disponer de datos termodinámicos que permitan evaluar la estabilidad de dicho motivo y la incidencia que en ella tiene la secuencia oligonucleotídica.

Los resultados descritos en el capítulo anterior, muestran que la secuencia no parece ser crítica de cara a la formación de las estructuras diméricas, pero afecta de forma drástica a su estabilidad, tal y como parece indicar el hecho de que, para las distintas secuencias estudiadas, el dímero se observe a concentraciones muy diferentes. En la bibliografía existen numerosos trabajos sobre la estabilidad de estructuras tipo *hairpin* y *dumbbell*, a partir de los cuales se han podido delimitar ciertos requisitos secuenciales para la composición de los mini-loops<sup>1-4</sup>. En este sentido, y puesto que en algunas de las secuencias se observa la formación de las dos formas, *dumbbell* y bi-loop, también resultaría interesante comprobar si la secuencia oligonucleotídica afecta a las estructuras bi-loop de la misma forma que a estructuras tipo *dumbbell*.

La forma más usual de obtener parámetros termodinámicos correspondientes a un equilibrio de desestructuración de un fragmento de DNA es a partir de curvas de fusión de absorción UV o de DC<sup>5-6</sup>. Como se comprobó en el Capítulo 4, las curvas de fusión por absorción UV no son apropiadas, en este caso, para extraer parámetros termodinámicos, pero sí resulta útil, en cambio, la técnica de DC ya que se ha observado una fuerte dependencia de la elipticidad con la temperatura, tanto en el caso de estructuras *dumbbell* como en el caso de los dímeros. Existe además la posibilidad de llevar a cabo un estudio termodinámico a partir de experimentos de RMN. La variación del espectro monodimensional con la temperatura no sólo contiene la misma información global que la obtenida por métodos ópticos, sino que además proporciona información local que en ocasiones puede resultar de gran utilidad<sup>7-12</sup>.

Los objetivos de este capítulo se centran, por tanto, en determinar los parámetros

<sup>1</sup> Hilbers, C.W.; Heus, H.A.; Van Dongen, M.J.P.; Wijmenga, S.S. *Nucleic Acids and Molecular Biology* **1994**, 8, 56-104 (Eds. Eckstein, F.; Lilley, D.M.J.), Springer, Berlín.

<sup>2</sup> Blommers, M.J.J.; Walters, J.A.L.I.; Haasnoot, C.A.G.; Aelen, J.M.A.; Van der Marel, G.A.; Van Boom, J.H.; Hilbers, C.W. *Biochemistry* **1989**, 28, 7491-7498.

<sup>3</sup> Senior, M.M.; Jones, R.A.; Breslauer, K.J. *Proc. Natl. Acad. Sci. USA* **1988**, 85, 6242-6246.

<sup>4</sup> Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, 24, 379-404.

<sup>5</sup> Marky, L.A.; Breslauer, K.J. *Biopolymers* **1987**, 26, 1601-1620.

<sup>6</sup> Breslauer, K.J. *Methods in Enzymology* **1995**, 259, 221-242.

<sup>7</sup> Hartel, A.; Lankhorst, P.P.L.; Altona, C. *Eur. J. Biochem.* **1982**, 129, 343-357.

<sup>8</sup> Olsthoorn, C.M.S.; Doornbos, J.; De Leeuw, H.P.M.; Altona, C. *Eur. J. Biochem.* **1982**, 125, 367-382.

<sup>9</sup> Mellema, J.-R.; Pieters, J.M.L.; Van der Marel, G.A.; Van Boom, J.H.; Altona, C. *Eur. J. Biochem.* **1984**, 143, 285-301.

<sup>10</sup> Mellema, J.-R.; Jellema, A.K.; Haasnoot, C.A.G.; Van Boom, J.H.; Altona, C. *Eur. J. Biochem.* **1984**, 141, 165-175.

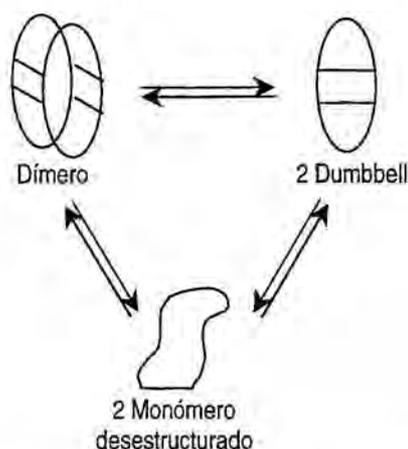
<sup>11</sup> Rinkel, L.J.; Van der Marel, G.A.; Van Boom, J.H.; Altona, C. *Eur. J. Biochem.* **1987**, 166, 87-101.

<sup>12</sup> Pieters, J.M.L.; Mellema, J.-R.; Van den Elst, H.; Van der Marel, G.A.; Van Boom, J.H.; Altona, C. *Biopolymers* **1989**, 28, 717-740.

termodinámicos asociados al equilibrio entre las distintas especies observadas en disolución utilizando, por una parte, los datos obtenidos por RMN y, por otra parte, los datos obtenidos a partir de curvas de fusión por DC. Los valores obtenidos para los distintos equilibrios, aunque se trata de un estudio preliminar, nos permitirán evaluar la estabilidad de las distintas formas para las distintas secuencias.

### 5.1.- DESCRIPCIÓN DEL EQUILIBRIO.

En el presente trabajo pretendemos estudiar el equilibrio que tiene lugar en disolución entre las dos formas estructuradas posibles, *dumbbell* y dímero, y el equilibrio que, a su vez, ambas formas presentan con la forma totalmente desestructurada. Se trata, por tanto, de un equilibrio entre tres estados tal como se representa en la Figura 5.1.



**Figura 5.1.-** Equilibrio en disolución entre las formas *dumbbell*, bi-loop y la forma totalmente desestructurada.

En todos los casos estudiados tenemos evidencias de que tanto el equilibrio dímero-*dumbbell* como el equilibrio dímero-random coil son lentos en la escala de tiempos de RMN, pues se observa para cada una de las especies un conjunto de señales diferenciado. En cambio, el equilibrio *dumbbell*-random coil es un equilibrio rápido y se observa únicamente un grupo de señales que aparecen al desplazamiento químico promedio del protón en cada una de las especies. Este desplazamiento, al ser función de la proporción entre ellas, presenta una fuerte dependencia con la temperatura.

Teniendo en cuenta la distinta naturaleza de los tres equilibrios, éstos deberán estudiarse utilizando distintas metodologías. El equilibrio *dumbbell*-random coil se puede estudiar utilizando la variación de los desplazamientos químicos con la temperatura, así como a partir de curvas de fusión por DC en H<sub>2</sub>O. En el caso de los equilibrios dímero-*dumbbell* y dímero-random coil, puesto que observamos señales independientes para las dos formas, podemos calcular directamente las poblaciones de cada una de las especies mediante una simple integración de señales. A partir de estas poblaciones, podemos obtener la constante de equilibrio, *K*, y, estudiando la variación de ésta con la temperatura, podemos obtener los parámetros termodinámicos.

Del diagrama termodinámico de la Figura 5.1, el único caso que involucra solamente procesos mono-moleculares es el equilibrio *dumbbell*-random-coil. Este equilibrio, por tanto, no depende de la concentración y puede estudiarse de manera independiente trabajando a bajas

concentraciones, en las que no aparece la forma dimérica. Esto permite comparar directamente los resultados obtenidos mediante RMN y mediante curvas de fusión de DC, aunque se hayan obtenido a diferente concentración de oligonucleótido.

El equilibrio dímero-*dumbbell* puede estudiarse mediante RMN trabajando en condiciones de concentración en las que ambas especies coexistan y a temperaturas suficientemente por debajo de la temperatura de fusión de las formas estructuradas. Análogamente, podemos estudiar el equilibrio dímero-forma desestructurada trabajando a concentraciones suficientemente altas, que garanticen una población de *dumbbell* muy minoritaria.

En principio, el equilibrio dímero-random coil podría estudiarse también mediante curvas de fusión de DC. Desgraciadamente, no ha sido posible realizar estos experimentos a las concentraciones de oligonucleótido necesarias para garantizar un 100% de forma dimérica en agua. No obstante, como se comprobó en la sección 4.3.1.3, la presencia de sales en el medio favorece la formación de las formas diméricas a concentraciones menores. Este efecto se ha aprovechado para observar la fusión de las formas diméricas mediante DC. Sin embargo, a diferencia del equilibrio *dumbbell*-random coil, los datos obtenidos mediante RMN y DC del equilibrio dímero-random coil no serán directamente comparables debido a la diferente fuerza iónica del medio.

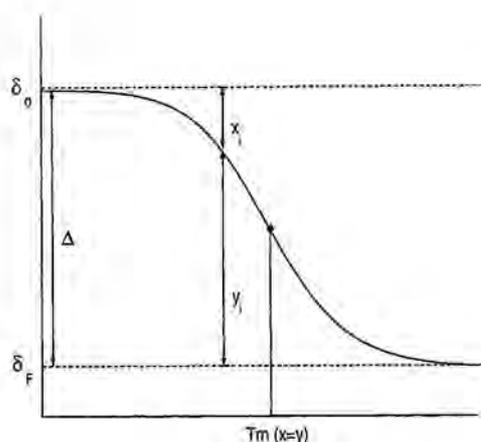
El diagrama termodinámico de la Figura 5.1 se simplifica notablemente para las secuencias d<pCGCTCATT> y d<pCCGTCCGT>, en las que la forma monomérica no presenta estructura. El diagrama se reduce a un equilibrio de dos estados entre la forma bi-loop y el monómero desestructurado.

## **5.2.- EQUILIBRIO DUMBBELL-MONÓMERO DESESTRUCTURADO. ESTABILIDAD DE ESTRUCTURAS DUMBBELL EN FUNCIÓN DE LA SECUENCIA.**

### **5.2.1.- ESTUDIO POR RMN: DETERMINACIÓN DE LAS CONSTANTES DE EQUILIBRIO A PARTIR DE LA VARIACIÓN DE LOS DESPLAZAMIENTOS QUÍMICOS CON LA TEMPERATURA.**

En el caso del equilibrio *dumbbell*-monómero desestructurado la señal de resonancia es una señal promedio de la correspondiente a forma estructurada y de la forma no estructurada. La representación de la variación del desplazamiento químico ( $\delta$ ) con la temperatura origina una curva sigmoide que corresponde a la curva de fusión de la forma estructurada. A partir de esta curva, y aplicando el tratamiento de equilibrio entre dos estados, se pueden extraer las constantes termodinámicas asociadas a este equilibrio, así como la  $T_m$  de la forma estructurada.

En la Figura 5.2 se muestra el perfil de cómo sería la variación del desplazamiento químico con la temperatura para un equilibrio rápido entre dos estados:



**Figura 5.2.-**Representación de la variación del desplazamiento químico con la temperatura para un equilibrio rápido entre dos estados.  $\delta_0$  y  $\delta_F$  representan el valor inicial (100% de forma estructurada) y final (100% de forma desestructurada) del desplazamiento químico, respectivamente,  $\Delta$  es la diferencia entre el valor final e inicial del desplazamiento y " $x$ " e " $y$ " la fracción molar de forma desestructurada y forma dumbbell, respectivamente. Puede suceder que  $\delta_0 > \delta_F$ , como en la gráfica, o al revés obteniéndose entonces la curva invertida.

Puesto que en cada punto de la curva el desplazamiento químico es función de la fracción molar de cada una de las especies, éste se puede expresar de la forma:

$$\delta_i = \delta_0 \cdot y_i + \delta_F \cdot x_i \quad (1)$$

Utilizando las expresiones:  $y_i = 1 - x_i$  y  $\Delta = \delta_F - \delta_0$ , la expresión (1) queda de la forma:

$$\delta_i = \delta_0 + \Delta \cdot x_i \quad (2)$$

Teniendo en cuenta de que para un equilibrio entre dos estados la  $K$  se expresa como:

$$K = \frac{x_i}{(1 - x_i)} \quad (3)$$

y que, a su vez, la  $keq$  se relaciona con el  $\Delta G^0$  conforme a la expresión:

$$K = e^{\frac{-\Delta G^0}{RT}} \quad (4)$$

podemos obtener una expresión donde el desplazamiento químico se expresa únicamente en función de los parámetros termodinámicos y de la temperatura.

$$\delta_i = \delta_0 + \Delta \cdot \frac{e^{\frac{\Delta S^0}{R}} \cdot e^{\frac{-\Delta H^0}{RT}}}{1 + e^{\frac{\Delta S^0}{R}} \cdot e^{\frac{-\Delta H^0}{RT}}} \quad (5)$$

En esta expresión se supone que tanto  $\Delta H^0$  como  $\Delta S^0$  son independientes de la temperatura. Esto es equivalente a suponer que la variación de la capacidad calorífica a presión constante,  $\Delta C_p^0$ , a lo largo de la transición entre los dos estados es cero. Al contrario de lo

habitual en el caso de proteínas, esta aproximación parece ser de aplicación general en moléculas de DNA<sup>13</sup>.

El ajuste de la representación  $\delta_i$  vs T según la expresión (5) permite la obtención de las constantes termodinámicas correspondientes al equilibrio *dumbbell*-random coil. La calidad del ajuste nos permitirá, además, conocer si la aproximación realizada es correcta y si los puntos experimentales se ajustan a un equilibrio de dos estados, sin necesidad de incluir un tercer estado.

Una vez obtenidos los valores de  $\Delta H^0$  y  $\Delta S^0$  correspondientes al equilibrio, se podría calcular directamente los valores de  $T_m$  y  $\Delta G^0$  a una temperatura dada. No obstante, el valor de  $T_m$  se obtiene con mayor precisión a partir de la curva de variación de  $\delta$  con la temperatura, bien calculando  $\delta_i$  y  $\delta_0$ , o bien calculando el punto de inflexión. Así mismo, el valor de  $\Delta G^0$  a una temperatura dada se obtiene a partir de la expresión siguiente, que es válida para equilibrios monomoleculares<sup>5</sup>:

$$\Delta G^0 = \Delta H^0 \cdot \left(1 - \frac{T}{T_m}\right)$$

### 5.2.1.1.- Caracterización del equilibrio *dumbbell*-random coil.

Para llevar a cabo el estudio del equilibrio entre la forma *dumbbell* y la forma desestructurada se realizaron espectros de RMN 1D en D<sub>2</sub>O a diferente temperatura, de 2 a 80°C, para las cuatro secuencias que presentan una forma monomérica estructurada: d<pCATTCAATT>, d<pTGCTCGCT>, d<pTGCATGCA> y d<pAGCTAGCT>.

En la Figura 5.3 se muestran las series de espectros correspondientes a las secuencias d<pTGCTCGCT> y d<pCATTCAATT>, donde se puede apreciar como, en ambos casos, el conjunto de señales que corresponden a la forma monomérica estructurada varían con la temperatura.

En el caso de d<pTGCTCGCT>, se adquirieron los espectros a una concentración de oligonucleótido de 2.4 mM, por lo que aparecen tanto la especie monomérica como la dimérica (Figura 3B). En dichos espectros se puede observar como, a diferencia de las señales correspondientes a la forma *dumbbell*, las señales de la forma dimérica disminuyen en intensidad pero se mantienen prácticamente en el mismo desplazamiento químico.

Para las cuatro secuencias anteriormente citadas se ha representado la variación del desplazamiento químico de distintos protones con la temperatura. Se han escogido principalmente protones aromáticos y protones metílicos ya que, en ambas zonas, existe un menor solapamiento y resulta más fácil seguir la variación del desplazamiento con la temperatura. Las curvas  $\delta$  vs T se han ajustado a la expresión (5) descrita en el apartado anterior correspondiente a un equilibrio de dos estados.

Los errores correspondientes a cada uno de los parámetros termodinámicos obtenidos a partir del ajuste, se han calculado aplicando el método de propagación de errores, tal como se describe en el apartado 3.3 de Materiales y Métodos.

A continuación se muestran los ajustes para cada una de las secuencias. En las tablas donde se muestran los parámetros termodinámicos se ha incluido también el valor de  $T_m$  que se obtiene como el punto medio de la curva ajustada y así como el que se obtiene a partir de calcular la derivada de la curva ajustada en cada caso.

<sup>13</sup> Breslauer, K.J.; Frank, R.; Blocker, H.; Marky, L.A. *Proc. Natl. Acad. Sci. USA* 1986, 83, 3746.

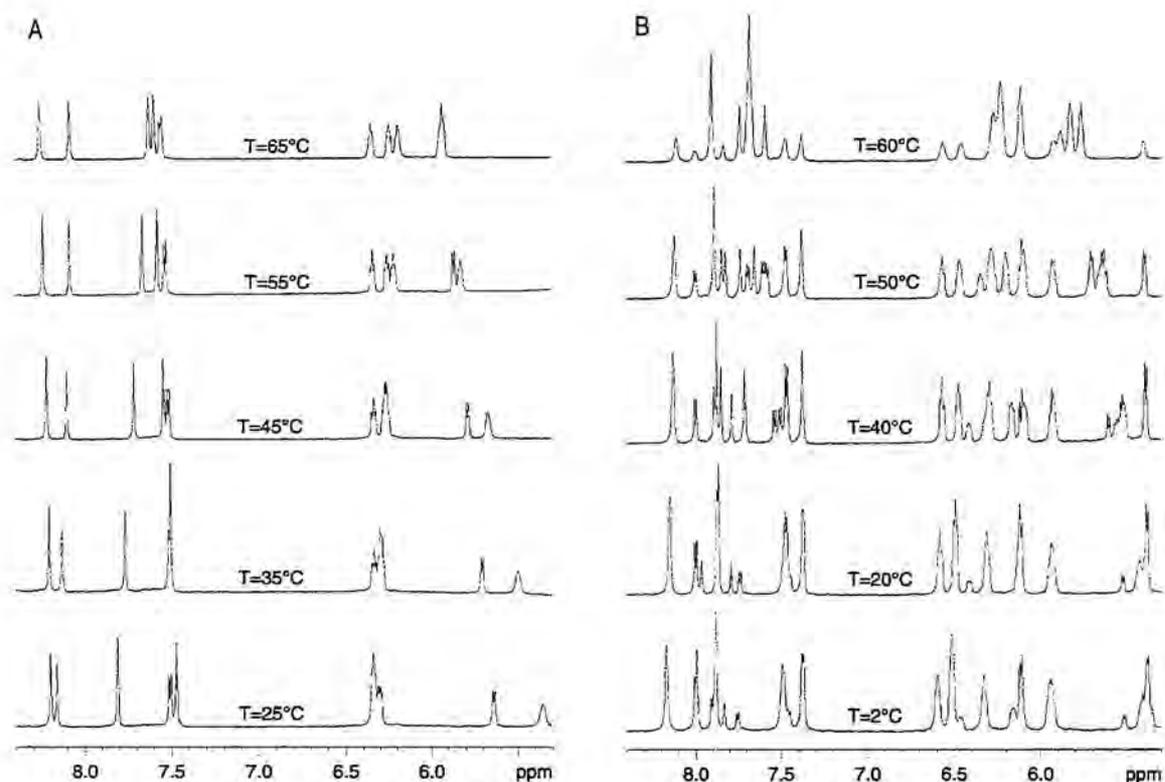
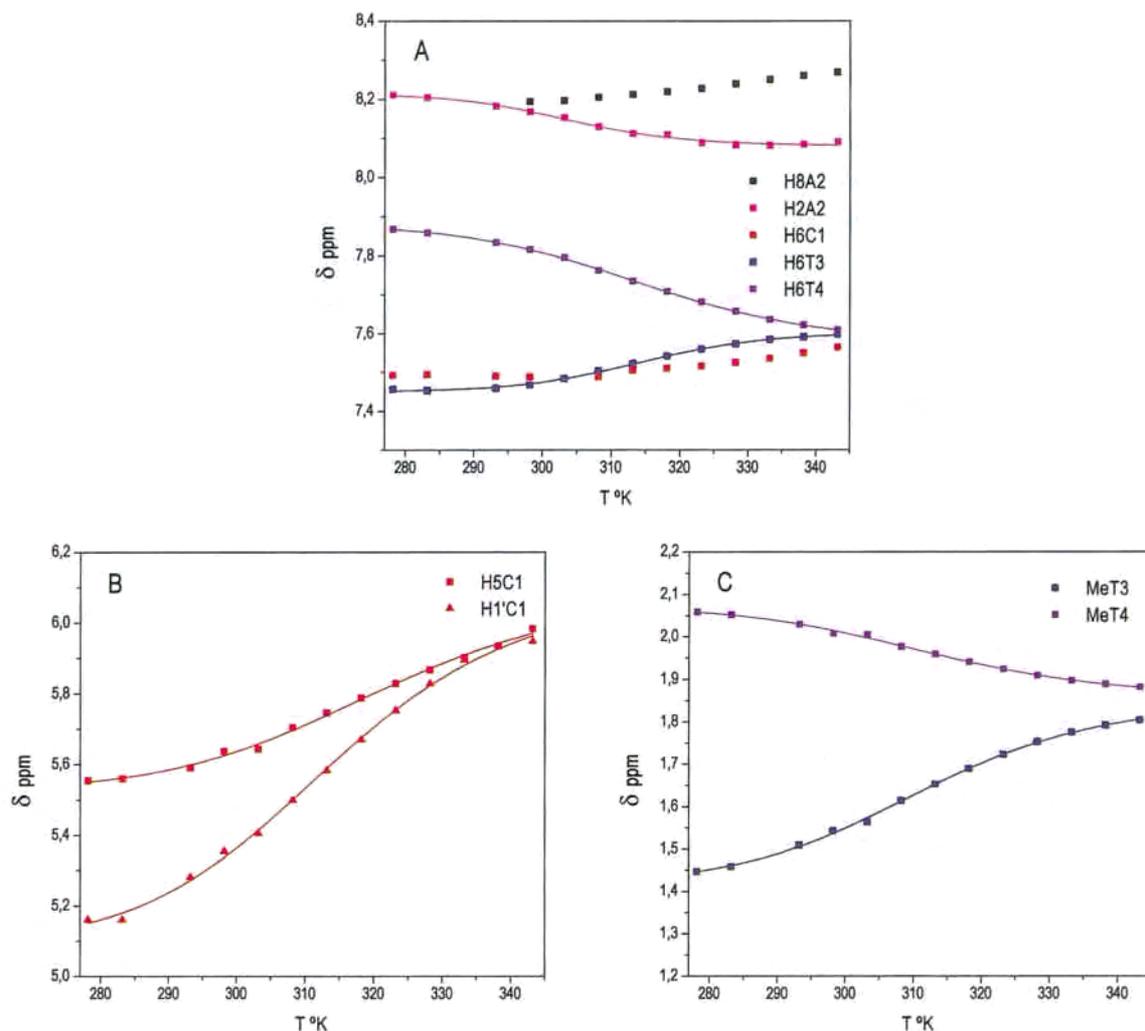


Figura 5.3.- Series de espectros a diferente temperatura para la forma dumbbell de A)  $d\langle pCATTcATT \rangle$  y B)  $d\langle pTGCTcGCT \rangle$ .

#### 5.2.1.1.1.- $d\langle pCATTcATT \rangle$ .

En la Figura 5.4 se muestra la representación  $\delta$  vs T para distintos protones de la forma dumbbell de  $d\langle pCATTcATT \rangle$ . En el caso de los protones aromáticos no ha sido posible ajustar las curvas correspondientes a los protones H8A2 y H6C1 ya que, además de presentar una variación muy pequeña, el perfil de la curva no corresponde al de una transición bien definida. Los protones para los que se observa una mayor variación del desplazamiento químico con la temperatura son los protones metílicos y los protones H5C1 y H1'C1. En la Tabla 5.1 se muestran los parámetros termodinámicos calculados a partir de los ajustes.



**Figura 5.4.-** Ajustes de las curvas  $\delta$  vs  $T$  para distintos protones correspondientes a la forma monomérica de  $d\langle pCATTcATT \rangle$ . A) protones H6/H8, B) protones H5 y H1' y C) protones metílicos. Las curvas para H8A2 y H6C1 no se ajustan correctamente y no aparecen dibujadas.

**Tabla 5.1.-** Parámetros termodinámicos correspondientes al equilibrio de formación de la forma dumbbell de  $d\langle pCATTcATT \rangle$ .  $\Delta G^0$  se ha calculado como  $\Delta G^0 = \Delta H^0 (1 - T/T_m)$ .

	$\Delta H^0$ kJ/mol	$\Delta S^0$ J/mol	$\Delta G^0_{298}$ kJ/mol	$T_m$ °K (a)	$T_m$ °K (b)
H2A2(6)	-90±5	-297±17	-2±0.1	304	303
H6T3(7)	-91±2	-288±7	-5±0.1	315	314
H6T4(8)	-62±1	-196±3	-3±0.05	315	313
H5C1(5)	-55±4	-170±15	-4±0.3	321	318
H1'C1(5)	-59±3	-185±9	-4±0.2	319	311
MeT3(7)	-59±2	-191±7	-2±0.1	311	309
MeT4(8)	-57±3	-184±11	-3±0.1	312	309
Media	-68±6	-216±20	-3±1	314	311

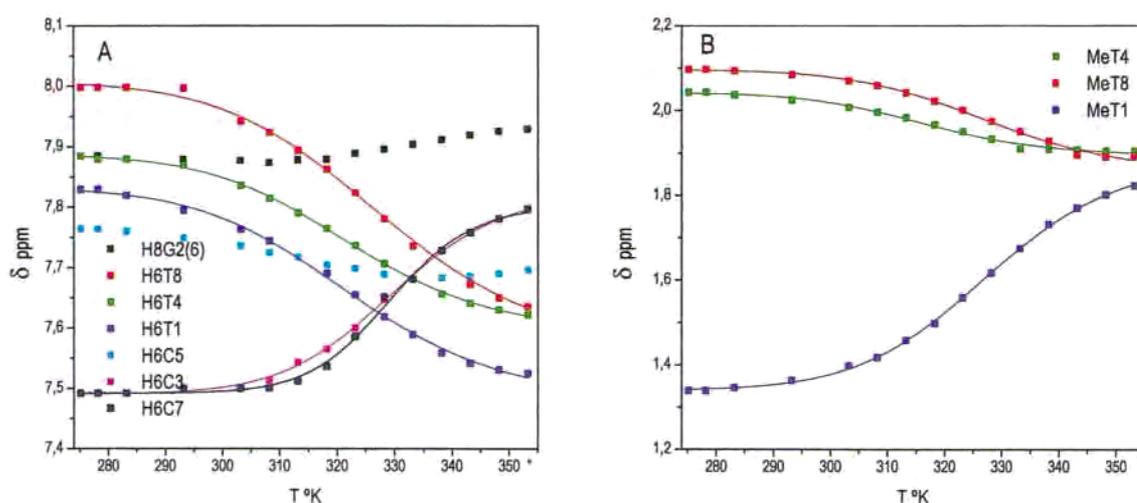
(a)  $T_m$  correspondiente al punto medio de la curva ajustada.

(b)  $T_m$  calculada a partir de la derivada de la curva  $\delta$  vs  $T$ .

Si nos fijamos en el valor de  $T_m$ , que es un indicativo de la estabilidad del dumbbell, se obtienen valores similares para todos los protones excepto para los protones correspondientes a los residuos A2(6) y C1(5), para los que se obtienen valores significativamente inferiores y superiores a la media, respectivamente. Probablemente este hecho se debe a los errores que se cometen en la determinación de los parámetros termodinámicos, ya que los valores correspondientes a estos protones son los que tienen asociados errores mayores en el ajuste. Los residuos C1(5), además, ocupan la segunda posición en el mini-loop y están totalmente expuestos al disolvente, por lo que no existe una justificación estructural para que presenten una  $T_m$  superior a la que se obtiene para el resto de residuos. Si calculamos la  $T_m$  sin tener en cuenta estos protones, se obtiene un valor de 313.1°K.

#### 5.2.1.1.2.- $d\langle pTGCTCGCT \rangle$ .

En la Figura 5.5 se muestra la variación  $\delta$  vs T para distintos protones de la forma *dumbbell* de  $d\langle pTGCTCGCT \rangle$ .



**Figura 5.5.-** Representación  $\delta$  vs T para distintos protones correspondientes a la forma *dumbbell* de  $d\langle pTGCTCGCT \rangle$ . A) Protones aromáticos y B) protones metílicos. Los desplazamientos químicos proceden de una serie de espectros a distinta temperatura realizada a 2.4mM de oligonucleótido. Las curvas para H8G2 y H6C5 no se ajustan correctamente y no aparecen dibujadas.

Como en la estructura anterior, algunas curvas no dieron lugar a ajustes satisfactorios y los parámetros termodinámicos que se extraen no son fiables. Este es el caso de los protones H8G2(6) y H6C5 para los que, tal como se puede comprobar en la Figura 5.5, la variación del desplazamiento químico con la temperatura es muy pequeña en comparación a la observada para el resto de protones. Por este motivo, para el cálculo de los valores medios no se han tenido en cuenta los ajustes de las curvas correspondientes a estos dos protones.

**Tabla 5.2.-** Parámetros termodinámicos correspondientes al equilibrio de formación de la estructura *dumbbell* para la secuencia  $d<pTGCTCGCT>$ .

	$\Delta H^{\circ}$ kJ/mol	$\Delta S^{\circ}$ J/mol	$\Delta G^{\circ}_{298}$ kJ/mol	$T_m$ °K (a)	$T_m$ °K (b)
H6T1	-63±2	-196±8	-5±0.1	322	320
H6C3	-92±8	-279±6	-9±1	330	329
H6T4	-68±1	-211±4	-5±0.1	322	320
H6C7	-125±3	-379±8	-6±0.3	329	329
H6T8	-63±3	-191±12	-7±0.3	328	325
MeT1	-76±1	-232±4	-7±0.1	328	326
MeT4	-78±4	-245±14	-4±0.2	316	315
MeT8	-74±3	-225±11	-7±0.3	328	326
Media	-80±7	-245±22	-6±1	325	324

(a)  $T_m$  correspondiente al punto medio de la curva ajustada.(b)  $T_m$  calculada a partir de la derivada de la curva  $\delta$  vs T.

En este caso se obtiene un *dumbbell* bastante más estable que para la secuencia  $d<pCATTATT>$  con una  $T_m \sim 325^{\circ}\text{K}$ . Esta diferencia se discutirá más adelante, cuando se analice la estabilidad de los distintos *dumbbells* en función de la secuencia.

**5.2.1.1.3.-**  $d<pTGCATGCA>$ .

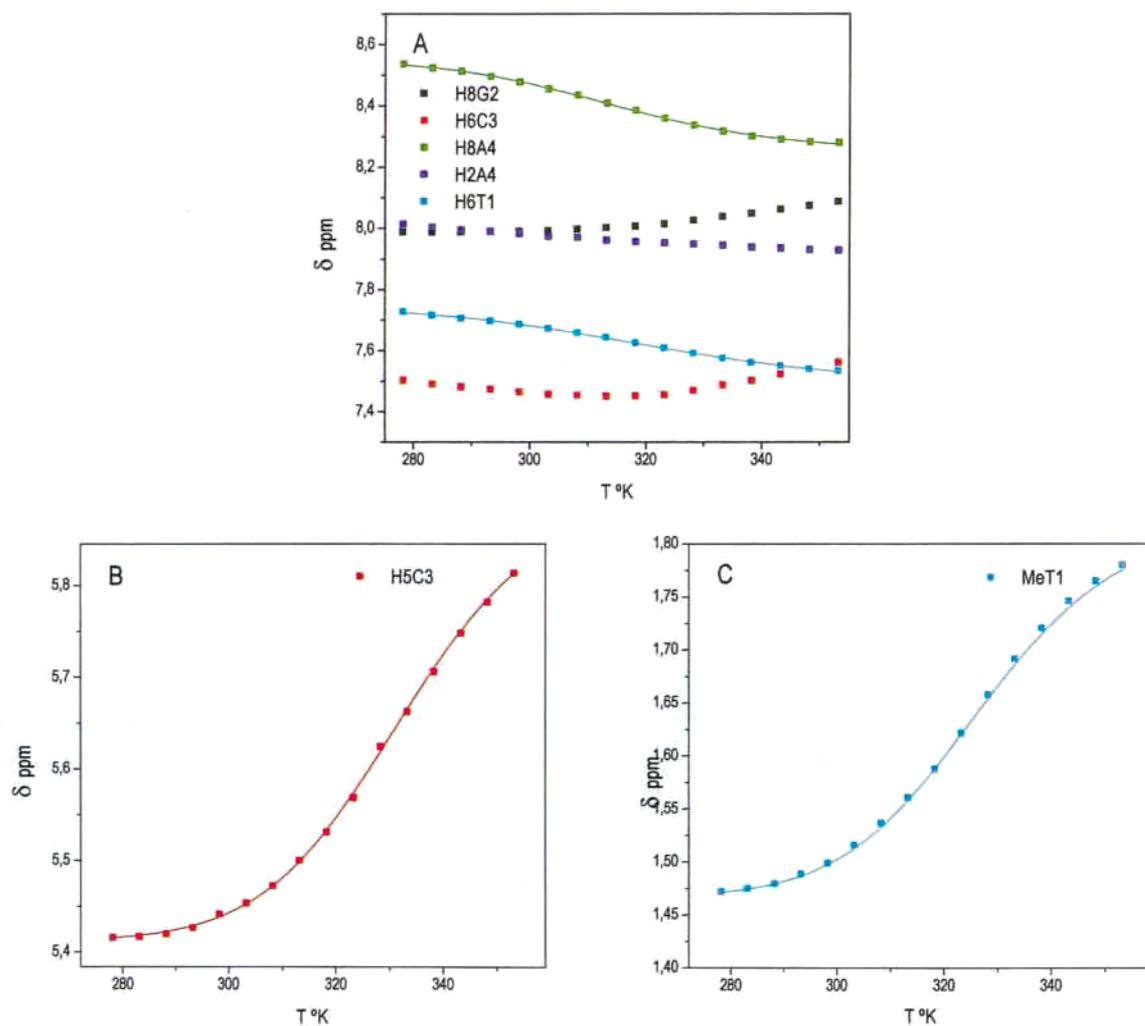
En el caso de  $d<pTGCATGCA>$  se observa, en general, una menor variación de los desplazamientos químicos con la temperatura. Aún y así resulta posible ajustar todas las curvas excepto la correspondiente al protón H6C3(7), para cuya representación no se obtiene una curva sigmoide, Figura 5.6. Para los protones H2A4(8) y H8G2(6), aunque sí que ha resultado posible realizar el ajuste, la variación del desplazamiento químico es demasiado pequeña para considerar fiables los parámetros termodinámicos extraídos a partir de sus correspondientes curvas. En el caso del H2A4(8), además, no resulta posible extraer, mediante la derivada, una  $T_m$  ya que no se observa un mínimo en la derivada de la curva. Si consideramos el resto de protones representados, se obtiene una  $T_m$  muy similar a la obtenida para  $d<pTGCTCGCT>$ ,  $323.8^{\circ}\text{K}$  ( $50.6^{\circ}\text{C}$ ), lo que resulta sorprendente teniendo en cuenta la presencia de una purina en el loop.

En la Tabla 5.3 se muestran los parámetros termodinámicos obtenidos a partir del ajuste.

**Tabla 5.3.-** Parámetros termodinámicos correspondientes al equilibrio de formación de la estructura *dumbbell* para la secuencia  $d<pTGCATGCA>$ . Los valores correspondientes a los protones H8G2 y H2A4 no se han tenido en cuenta en el cálculo de la media.

	$\Delta H^{\circ}$ kJ/mol	$\Delta S^{\circ}$ J/mol	$\Delta G^{\circ}_{298}$ kJ/mol	$T_m$ °K (a)	$T_m$ °K (b)
H6T1	-44±2	-136±7	-3±0.1	322	316
MeT1	-66±1	-202±2	-6±0.1	327	325
H5C3	-69±1	-207±4	-7±0.1	332	330
H8A4	-55±2	-174±5	-3±0.1	315	312
Media	-58±6	-180±16	-5±1	324	321

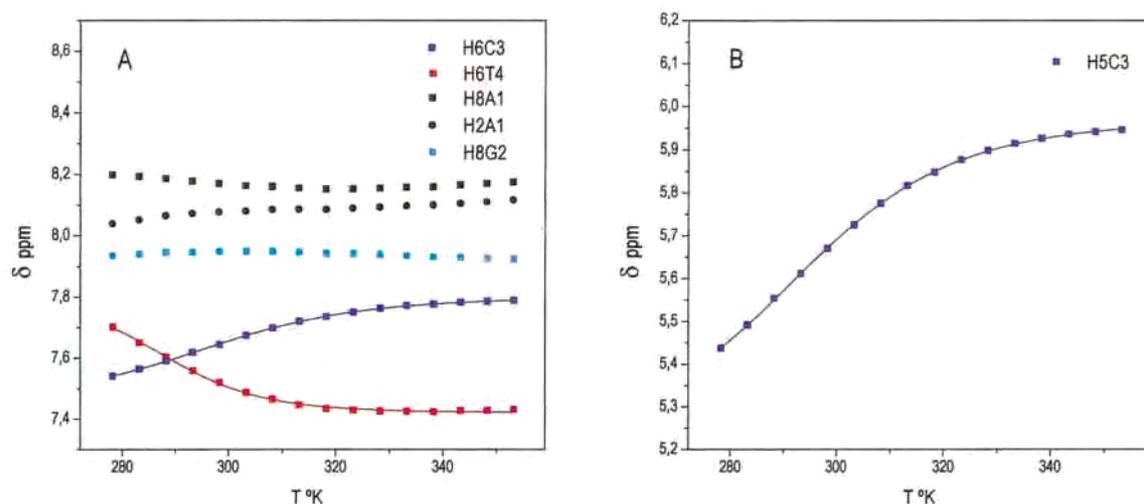
(a)  $T_m$  correspondiente al punto medio de la curva ajustada.(b)  $T_m$  calculada a partir de la derivada de la curva  $\delta$  vs T.



**Figura 5.6.-** Representación  $\delta$  vs  $T$  para distintos protones correspondientes a la forma monomérica de  $d\langle pTGCATGCA \rangle$ . Espectros realizados a 0.15mM de oligonucleótido.

#### 5.2.1.1.4.- $d\langle pAGCTAGCT \rangle$ .

En el caso de la forma *dumbbell* correspondiente a  $d\langle pAGCTAGCT \rangle$ , aunque se observa variación de los desplazamientos químicos con la temperatura, únicamente ha sido posible ajustar las curvas correspondientes a los protones H6C3, H5C3 y H6T4, Figura 5.7.



**Figura 5.7.-** Ajustes de las curvas  $\delta$  vs  $T$  para distintos protones del dumbbell de  $d\langle pAGCTAGCT \rangle$ . A) protones H6/H8/H2 y B) protón H5.

Los parámetros termodinámicos obtenidos partir de estos ajustes, Tabla 5.4, constatan que se trata de un *dumbbell* muy poco estable ( $T_m \sim 290^\circ\text{K}$ ) y que a  $T_{amb}$  es mayoritaria la forma totalmente desestructurada. Probablemente sea éste el motivo por el que la mayoría de curvas no se pueden ajustar convenientemente y sería necesario disponer de datos de desplazamiento químico a temperaturas inferiores a  $278.15^\circ\text{K}$ .

**Tabla 5.4.-** Parámetros termodinámicos correspondientes a la formación de la estructura *dumbbell* de  $d\langle pAGCTAGCT \rangle$ .

	$\Delta H^0$ kJ/mol	$\Delta S^0$ J/mol	$\Delta G^0_{298}$ kJ/mol	$T_m$ °K (a)	$T_m$ °K (b)
H6C3	$-51 \pm 1$	$-175 \pm 2$	$0.2 \pm 0.01$	297	294
H6T4	$-75 \pm 3$	$-260 \pm 10$	$3 \pm 0.1$	288	287
H5C3	$-50.4 \pm 1$	$-171 \pm 2$	$0.7 \pm 0.01$	294	291
Media	$-59 \pm 8$	$-202 \pm 29$	$1.3 \pm 1$	293	291

(a)  $T_m$  correspondiente al punto medio de la curva ajustada.

(b)  $T_m$  calculada a partir de la derivada de la curva  $\delta$  vs  $T$ .

Un rasgo común a todos los *dumbbells* es que los protones de las bases pertenecientes a los residuos centrales, residuos 2, 3, 6 y 7, experimentan un desapantallamiento al aumentar la temperatura mientras que los correspondientes a residuos de los mini-loops experimentan, por lo general, el efecto contrario, desplazándose las señales hacia campos más altos. Esto es razonable teniendo en cuenta que los residuos centrales se encuentran apilados y los de los mini-loops totalmente expuestos al solvente. Este hecho podría estar indicando una cierta similitud estructural entre todos los *dumbbells*.

De cara a confirmar la validez de los parámetros termodinámicos obtenidos a partir de los ajustes de las curvas  $\delta$  vs  $T$ , y poder utilizar dichos valores para discutir la estabilidad de los distintos *dumbbells*, resultaba importante disponer de datos procedentes de otra técnica. A continuación se muestran las curvas de fusión por DC.

### 5.2.2.- DETERMINACIÓN DE LAS CONSTANTES TERMODINÁMICAS A PARTIR DE CURVAS DE FUSIÓN POR DC.

El equilibrio dumbbell-random coil, lo hemos estudiado también a partir de curvas de fusión por DC realizadas en H<sub>2</sub>O, para poder confirmar los parámetros obtenidos a partir de experimentos de RMN. Si bien las curvas de fusión se han realizado a una concentración de oligonucleótido muy inferior a la utilizada en RMN, al tratarse de un equilibrio independiente de la concentración, los parámetros termodinámicos obtenidos por son directamente comparables a los obtenidos a partir del análisis de la variación de los desplazamientos químicos con la temperatura.

En todos los casos se ha ajustado la curva de fusión por DC a una expresión análoga a la descrita en el apartado 5.2.1.-, donde ahora la propiedad que varía con la temperatura es la elipticidad de la muestra<sup>14</sup>:

$$\Delta\varepsilon_i = P_1 + (P_2 - P_1) \cdot \left[ \frac{P_3 \cdot e^{\frac{P_4}{T}}}{1 + P_3 \cdot e^{\frac{P_4}{T}}} \right]$$

donde  $P_1 = \Delta\varepsilon_0$ ,  $P_2 = \Delta\varepsilon_f$ ,  $P_3 = \exp(\Delta S^0/R)$  y  $P_4 = -\Delta H^0/R$ .

Para realizar las fusiones por DC se ha seguido en cada caso la variación de la elipticidad a  $\lambda_{\text{máx}}$ . Las condiciones experimentales se encuentran descritas en el apartado 3.2 de Materiales y Métodos.

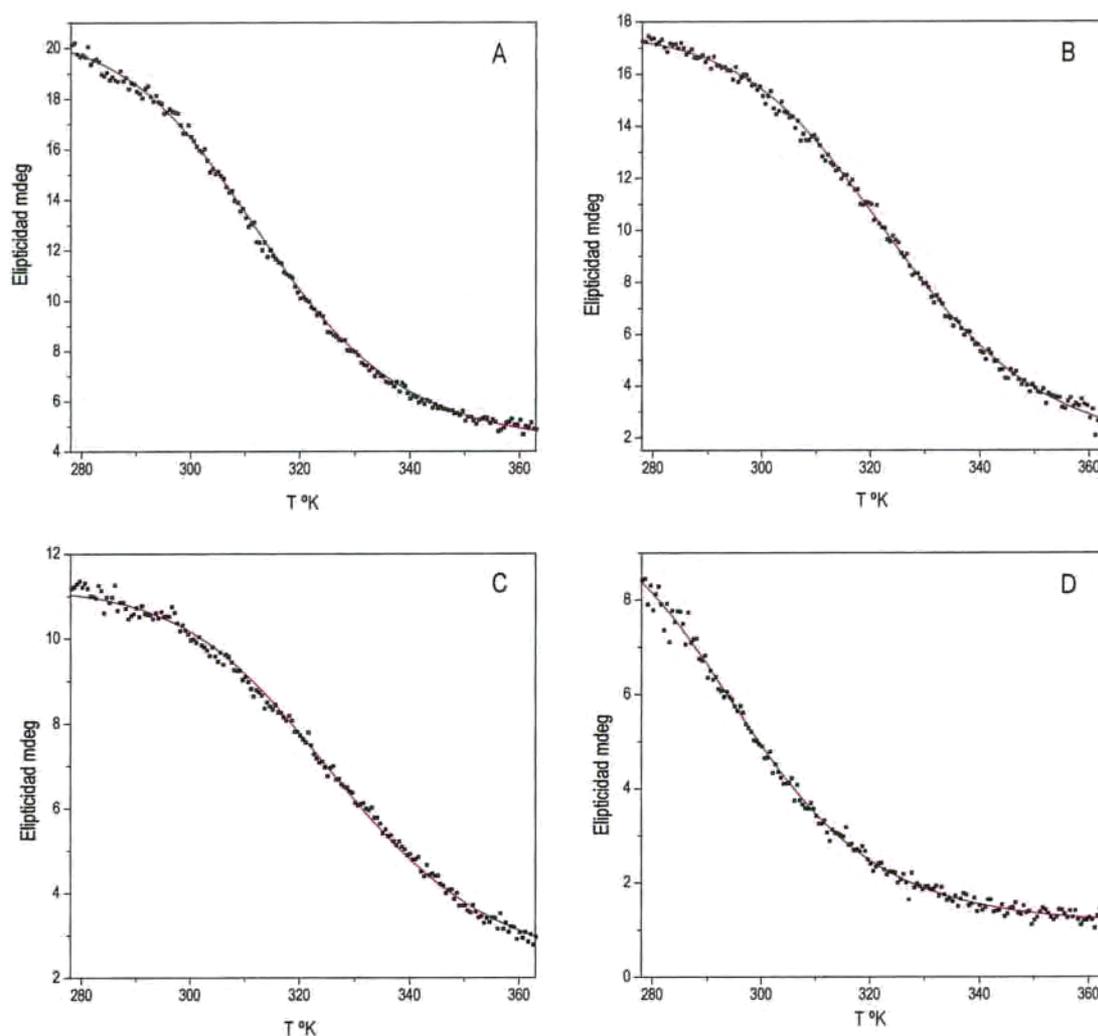
En la Figura 5.8 se muestran los ajustes de las curvas de fusión por DC en H<sub>2</sub>O para las cuatro secuencias que presentan forma *dumbbell*: d<pCATTCAATT>, d<pTGCTCGCT>, d<pTGCATGCA> y d<pAGCTAGCT>.

A partir de estos ajustes, y de forma análoga a la descrita en el apartado anterior, se han extraído los parámetros termodinámicos correspondientes al equilibrio *dumbbell*-random coil. Los resultados obtenidos para cada una de las secuencias se muestran en la Tabla 5.5.

**Tabla 5.5.-** Parámetros termodinámicos correspondientes al equilibrio *dumbbell*-random coil para las cuatro secuencias que presentan una forma monomérica estructurada. El valor de  $T_m$  es el calculado a partir de la derivada de la curva ajustada.

Secuencia	$\Delta H^0$ kJ/mol	$\Delta S^0$ J/mol·°K	$\Delta G_{298}^0$ kJ/mol	$T_m$ °K
d<pCATTCAATT>	-63±0.5	-201±2	-3±0.01	313
d<pTGCTCGCT>	-62±1	-191±2	-5±0.01	324
d<pTGCATGCA>	-62±1	-189±3	-5±0.01	327
d<pAGCTAGCT>	-57±1	-192±3	0.3±0.01	296

<sup>14</sup> Ramsing, N.B.; Rippe, K.; Jovin, T.M. *Biochemistry* **1989**, 28, 9528-9535.



**Figura 5.8.-** Ajustes de las curvas de fusión por DC de A)  $d\langle pCATT \rangle$ ,  $\lambda_{\text{máx}}=274.5\text{nm}$  B)  $d\langle pTGCTCGC \rangle$ ,  $\lambda_{\text{máx}}=287\text{nm}$  C)  $d\langle pTGCATGCA \rangle$ ,  $\lambda_{\text{máx}}=281.5\text{nm}$  y D)  $d\langle pAGCTAGCT \rangle$ ,  $\lambda_{\text{máx}}=285.5\text{nm}$ . Fusiones realizadas en  $\text{H}_2\text{O}$  a una concentración de  $5\ \mu\text{M}$ .

El hecho de que los valores de  $T_m$  obtenidos a partir de los datos de RMN y los obtenidos a partir de curvas de fusión por DC sean tan similares, nos da confianza para pensar que los parámetros termodinámicos para los cuatro *dumbbells* presentan una gran fiabilidad. En la siguiente sección utilizaremos estos parámetros para discutir la estabilidad de los distintos *dumbbells*.

Hay que destacar que, aunque en el caso del DC se obtienen parámetros globales correspondientes a toda la molécula, éstos no difieren de forma significativa de los encontrados por RMN a partir de las curvas de  $\delta$  vs  $T$ , donde los valores obtenidos pueden depender de la localización del protón considerado dentro de la estructura. En este sentido, es importante destacar que, a diferencia de lo descrito en otros trabajos en general no hemos observado diferencias muy significativas entre los valores de  $T_m$  obtenidos para diferentes protones. En nuestro caso, puesto que los dos *hairpins* que forman el *dumbbell* son equivalentes, los dos mini-loops deben presentar la misma estabilidad. Únicamente en el caso de la secuencia

d<pTGCTCGCT> podría haberse observado una mayor estabilidad de uno de los mini-loops. El *dumbbell* correspondiente a esta secuencia estaría formado por un loop 5'-CTTG-3', el loop más estable descrito en la bibliografía<sup>2</sup>, y un loop 5'-CTCG-3', pero la estabilidad de ambos parece ser similar según se deduce de los parámetros termodinámicos que se muestran en la Tabla 5.2.

### 5.2.3.- ESTABILIDAD DE LAS ESTRUCTURAS DUMBBELL.

Aunque disponemos de datos termodinámicos correspondientes únicamente a cuatro secuencias, el diseño de las mismas nos puede permitir realizar un análisis de la estabilidad de este tipo de estructuras en función de la composición de los mini-loops y la naturaleza de los pares de bases. En este sentido, no debemos olvidar que, tan importante como disponer de parámetros termodinámicos, es la evidencia de que, para una determinada secuencia, no se observe la formación de *dumbbell*. Este es el caso de d<pCCGTCCGT> que, a priori, también podría adoptar una estructura tipo *dumbbell*. Por este motivo, se ha incluido también esta última secuencia en la discusión.

En la Tabla 5.6 se muestran los valores de  $T_m$  y  $\Delta G^0$  para cada secuencia y se indica, en cada caso la, composición de los dos *hairpin* que forman el *dumbbell*. Puesto que, en muchas de las curvas de fusión, tanto por RMN como por DC, no se alcanza un plató final y la fusión parece inacabada, se ha considerado más correcto utilizar los valores de  $T_m$  que se extraen del cálculo de la derivada.

**Tabla 5.6.-** Valores de  $T_m$  y  $\Delta G^0$  para los *dumbbells* de las cuatro secuencias, donde  $T_m$  es el valor calculado a partir de la derivada de la curva ajustada.

Secuencia	Hairpin	$T_m$ °K RMN	$\Delta G^0_{298}$ RMN	$T_m$ °K DC	$\Delta G^0_{298}$ DC
d<pCATTGATT>	5'-T-TC-A-3'	311	-3±1	313	-3.0±0.01
d<pTGCTCGCT>	5'-C-TC-G-3' 5'-C-TT-G-3'	324	-6±1	324	-5.0±0.01
d<pTGCATGCA>	5'-C-AT-G-3'	321	-5±1	327	-5.0±0.01
d<pAGCTAGCT>	5'-C-TA-G-3'	291	1.3±1	296	0.3±0.01
d<pCCGTCCGT>	5'-G-TC-C-3'	No estructurado			

A la hora de discutir la estabilidad de los distintos *dumbbells*, es importante tener en cuenta no sólo qué residuos componen el mini-loop, sino también qué par de bases cierra el *hairpin*. Se encuentra descrito en la bibliografía que secuencias 5'-piridina-TT-purina-3' pueden formar loops de dos residuos, mientras que secuencias 5'-purina-TT-pirimidina-3' dan lugar a loops de cuatro residuos<sup>2-15</sup>. Los mismos autores concluyen que, en el primer caso, la sustitución de los residuos centrales de timidina por residuos de adenosina, más voluminosos, conlleva una desestabilización que induce la formación de loops de cuatro residuos. No obstante, esta última afirmación no parece ser generalizable ya que posteriormente distintos trabajos demuestran que es posible la presencia de purinas en loops de dos residuos sin que se produzca una desestabilización significativa<sup>16-17</sup>.

<sup>15</sup> Haasnoot, C.A.G.; Blommers, M.J.J.; Hilbers, C.W.; Structure, Dynamics and Function of Biomolecules 1987, Spring Series in Biophysics, 1, 212-216, Spring-Verlag, Berlín. (Editores: Ehrenberg, A.; Rigler, R.; Gräslund, A.; Nilsson, L.).

<sup>16</sup> Howard, F.B.; Chen, C.; Ross, P.D.; Miles, H.T. *Biochemistry* 1991, 30, 779-782.

En nuestro caso las estructuras *dumbbell* están formadas únicamente por dos *hairpins* de cuatro residuos, donde los dos pares de bases que estabilizan la estructura son, a su vez, los pares que cierran cada uno de los *hairpins*. Por este motivo, al no tener un tallo que establezca la estructura, la naturaleza de los residuos que constituyen dichos pares puede afectar dramáticamente a la estabilidad del *dumbbell*. Para todas las secuencias estudiadas, excepto para d<pCCGTCCGT>, se cumple que la secuencia del *hairpin* es 5'-pirimidina-loop-purina-3', descrita en la bibliografía como la más favorable. El hecho de que para la secuencia d<pCCGTCCGT> no se haya observado la formación del *dumbbell* coincide, por tanto, con lo descrito en la bibliografía y se debe a la preferencia que muestra la secuencia 5'-G-loop-C-3' a formar *loops* de 4 residuos.

Entre los *dumbbells* correspondientes a las secuencias d<pCATTCATT> y d<pTGCTCGCT>, cuyos mini-loops están constituidos únicamente por residuos pirimidínicos, la diferente estabilidad observada es debida principalmente a la distinta naturaleza de los pares de bases.

De la comparación de las secuencias d<pTGCTCGCT>, d<pAGCTAGCT> y d<pTGCATGCA>, podemos conocer el efecto que tiene la presencia de un residuo de purina en el mini-loop. En los tres casos, los *dumbbells* están formados por pares GC y los *hairpins* son siempre 5'-C-loop-G-3'. En el caso de la secuencia d<pAGCTAGCT>, en el que el residuo de purina ocupa la segunda posición del mini-loop, se observa una  $T_m$  muy inferior, tanto que a temperatura ambiente es mayoritaria la forma totalmente desestructurada. Esto indica que la localización de una purina en esta posición ejerce un efecto altamente desestabilizante. Este resultado, si bien demuestra que la presencia de una purina en la segunda posición del mini-loop no obliga que el loop formado sea de cuatro residuos, tal como se deduce de los datos de RMN, constata una menor estabilidad de estos loops respecto a los constituidos únicamente por pirimidinas.

Sin embargo, resulta sorprendente el resultado obtenido en el caso de d<pTGCATGCA>, cuyo *dumbbell* tiene un valor de  $T_m$  muy similar al obtenido para la secuencia d<pTGCTCGCT>. Teniendo en cuenta las estructuras obtenidas a partir de los datos de RMN para los *dumbbells* de d<pCATTCATT> y d<pTGCTCGCT>, en las que las dos bases de los mini-loops parecen estar totalmente expuestas al disolvente, no se explica la gran diferencia de estabilidad observada según que la purina ocupe la primera o la segunda posición del loop. No obstante, hay que señalar que las estructuras obtenidas para los *dumbbells* anteriores no son estructuras altamente refinadas y la estructura del loop podría ser diferente en el caso de d<pTGCATGCA>.

En resumen, de los datos de la Tabla 5.6 se deduce que el efecto más importante en la estabilidad de las estructuras tipo *dumbbell* en octámeros cíclicos es el orden de nucleótidos que cierran el loop, siendo el 5'-C-loop-G-3' el más estable. El segundo factor que afecta a la estabilidad es la presencia o no de purinas en el loop, aunque este efecto parece tener importancia sólo si la purina ocupa la segunda posición del mismo. Finalmente, nuestros resultados parecen indicar que la naturaleza de los pares de bases (A-T o G-C) tiene un efecto menor que los dos anteriores en la estabilidad de estas moléculas.

---

<sup>17</sup> Raghunathan, G.; Jernigan, R.L.; Miles, H.T.; Sasisekharan, V. *Biochemistry* **1991**, 30, 782-788.

### 5.3.- EQUILIBRIO DUMBBELL-DÍMERO.

En este apartado nos planteamos estudiar el equilibrio *dumbbell*-bi-loop para las cuatro secuencias que presentan una forma *dumbbell* bien estructurada. Para ello es necesario trabajar en condiciones en las que ambas especies coexistan y poder determinar a cada temperatura la población correspondiente a cada una de ellas. Por este motivo, el estudio pudo llevarse a cabo únicamente mediante RMN.

Como ya se ha comentado anteriormente, el equilibrio entre la forma dimérica y la forma *dumbbell* es lento en la escala de tiempo de RMN, observándose un conjunto de señales correspondiente a cada una de las especies, Figura 5.3B. Este hecho nos permite determinar, por integración de las señales en el espectro 1D, la población correspondiente a forma dimérica y forma monomérica a una temperatura dada. Si realizamos una serie de espectros 1D a diferente temperatura podremos saber la población de dímero y *dumbbell* a cada temperatura y determinar, de esta forma, la K correspondiente al equilibrio. Para ello utilizamos la metodología descrita para equilibrios de dimerización de cadenas autocomplementarias de DNA<sup>18</sup>.

Para un equilibrio de este tipo, la K se expresa como:

$$K = \frac{[D]}{[d]^2} = \frac{\alpha}{2c_0(1-\alpha)^2}$$

Donde D=dímero, d=*dumbbell* y  $\alpha$  representa la fracción del número total de cadenas que están formando dímero y  $c_0$  la concentración inicial de oligonucleótido. En nuestro caso, y puesto que trabajamos con la información obtenida a partir de los espectros de RMN, nos interesa poner esta expresión en función de las integraciones de las señales correspondientes a forma monomérica y dimérica. Para ello definimos una nueva variable x:

$$x = \frac{I_D}{I_d}$$

Donde  $I_D$  representa la integración de una señal correspondiente a un protón en la forma dimérica y  $I_d$  la integración de una señal correspondiente a un protón en la forma *dumbbell*.

Puesto que la señal de un protón en la estructura dimérica integra el doble (corresponde a cuatro protones,  $4N_D$ ) que la del mismo protón en la forma *dumbbell* (corresponde a 2 protones,  $2N_d$ ), esto nos va a servir para relacionar concentraciones con intensidades:

$$x = \frac{I_D}{I_d} = \frac{4N_D}{2N_d} = 2 \frac{[D]}{[d]}$$

En cada momento  $c_0 = [d] + 2[D]$ , por lo que podemos extraer expresiones para la concentración de dímero y *dumbbell* en función de x y  $c_0$  y, de esta forma, expresar la K únicamente en función de x y  $c_0$ :

<sup>18</sup> Marky, L.A.; Breslauer, K.J. *Biopolymers* **1987**, 26, 1601-1620.

$$[d] = \frac{c_0}{(1+x)}$$

$$[D] = \frac{x}{2} \cdot \frac{c_0}{(1+x)}$$

$$K = \frac{x \cdot (1+x)}{2 \cdot c_0}$$

Determinando, a partir de los espectros 1D, la relación  $I_D/I_d$  a diferentes temperaturas podemos calcular el valor de la  $k_{eq}$  a cada una de las temperaturas. La representación de la K en función de la temperatura, o el  $\ln K$  en función del inverso de la temperatura, nos permitirá obtener los parámetros termodinámicos correspondientes al equilibrio:

$$K = e^{\frac{\Delta G^0}{RT}}$$

$$K = e^{\frac{\Delta S^0}{R}} \cdot e^{\frac{-\Delta H^0}{RT}}$$

Para realizar este estudio es necesario determinar el margen de temperaturas en el que podemos trabajar, suponiendo únicamente un equilibrio de dos estados (*dumbbell* y dímero), sin considerar la presencia de la forma monomérica desestructurada. A partir del valor de  $\Delta G^0$  obtenido para el equilibrio monómero desestructurado-*dumbbell* calculamos la K correspondiente a dicho equilibrio y determinamos la temperatura máxima en la que la presencia de forma desestructurada es despreciable. Para las cuatro secuencias estudiadas en el apartado anterior obtuvimos los valores de K a varias temperaturas inferiores a la  $T_m$  de cada *dumbbell*. Estos valores se muestran en la Tabla 5.7.

**Tabla 5.7.-** Valores de la K a diferentes temperaturas para el equilibrio random-*dumbbell* y el correspondiente % esperado de forma totalmente desestructurada.

Secuencia	$K_{278}$	%random	$K_{296}$	%random	$K_{305}$	%random	$K_{305}$	%random	$K_{315}$	%random
d<pCATTTCATT>	22.0	4.3	10.2	8.9	4.2	19.2	1.9	34.5	0.9	52.6
d<pTGCTCGCT>	45.8	2.1	21.6	4.4	9.0	10	4.3	18.9	1.96	33.8
d<pTGCATGCA>	55.4	1.8	26.1	3.7	10.9	8.4	5.2	16.1	2.4	29.4
d<pAGCTAGCT>	4.5	18.2	2.3	30.3	1.0	50	0.52	65.8	0.25	80

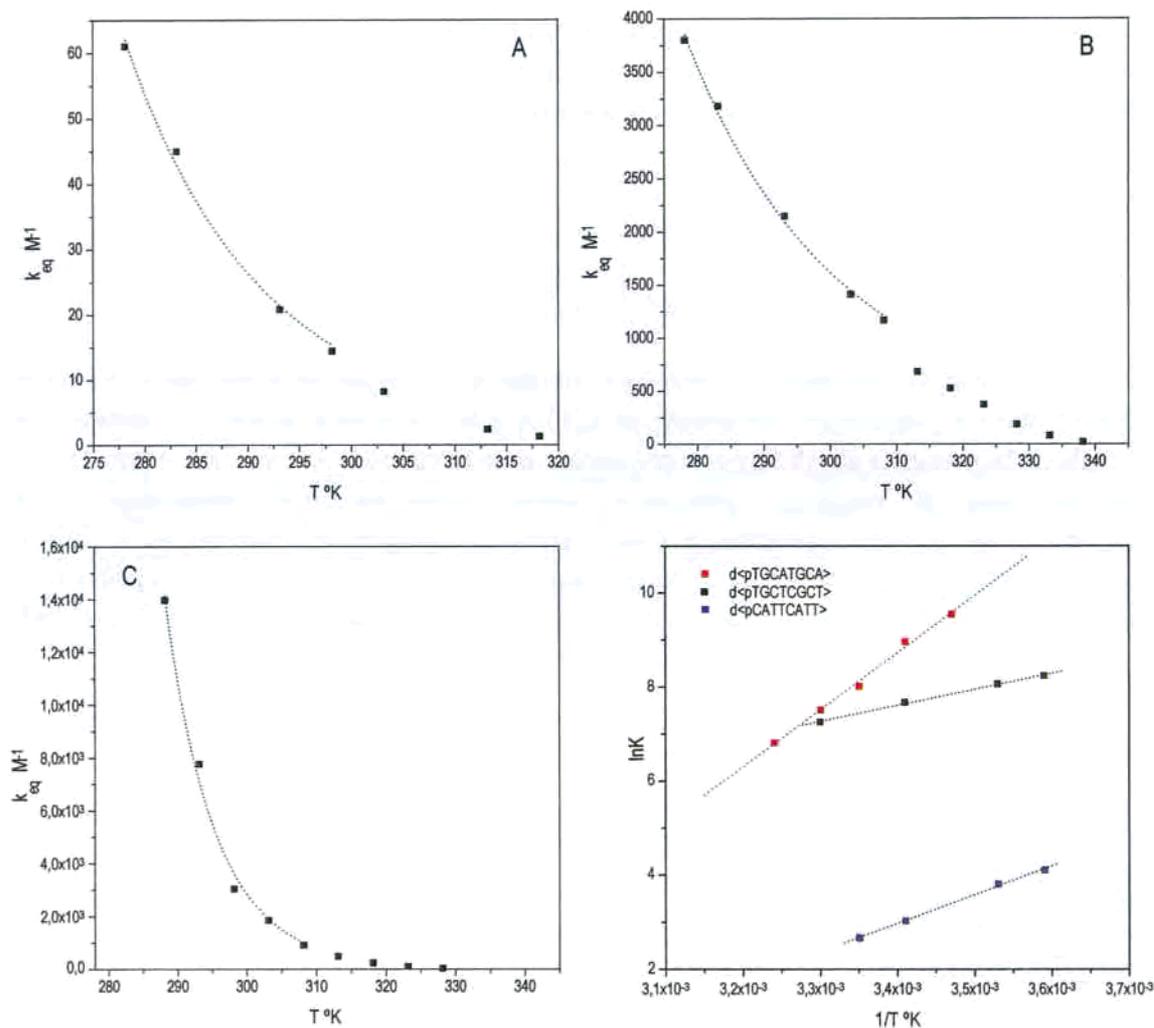
En el caso de d<pAGCTAGCT>, cuyo *dumbbell* presenta una baja estabilidad, estos valores muestran que incluso a 278°K la población de forma desestructurada es de casi un 20%, por lo que no resulta posible estudiar el equilibrio *dumbbell*-dímero para esta secuencia.

Para las otras tres secuencias hemos tomado como aproximación el considerar aquellos puntos situados hasta 20°C por debajo de la  $T_m$  de la forma *dumbbell*. Estrictamente esto supone que en algunos casos la proporción de forma desestructurada está entre 10-20%, sin embargo la representación de los puntos  $\ln K$  vs T presenta una buena relación lineal (Figura 5.12D) por lo que creímos que podríamos obtener valores suficientemente fiables. Para la obtención de los parámetros termodinámicos los puntos los hemos ajustado a la función:

$$K = P_1 \cdot e^{\frac{P_2}{T}}$$

donde  $P_1 = \exp(\Delta S^0/R)$  y  $P_2 = -\Delta H^0/R$ .

Los ajustes para cada una de las tres secuencias se muestran en la Figura 5.9. Los valores de la K a las distintas temperaturas se muestran en las tablas que figuran en el Anexo D.



**Figura 5.9.-** Ajustes de la representación  $K$  vs  $T$  para las tres secuencias que tienen una forma *dumbbell* estable. A)  $d\langle pCATT\langle CATT \rangle$  20mM, B)  $d\langle pTGCT\langle CGCT \rangle$  2.4mM y C)  $d\langle pTGCAT\langle GCA \rangle$  1mM. Las gráficas comprenden únicamente los puntos ajustados.

A partir de los ajustes hemos determinado los parámetros termodinámicos que corresponderían al equilibrio entre la forma *dumbbell* y la forma dimérica, Tabla 5.8. Puesto que no se trata de un proceso de desestructuración, sino de un equilibrio entre dos especies estructuradas, se ha utilizado  $T_{1/2}$  para designar la temperatura en la que existe un 50% de cada una de las especies. En este caso, al tratarse de un equilibrio de asociación intermolecular, la  $T_{1/2}$  no es independiente de la concentración y debe calcularse utilizando la expresión<sup>18</sup>:

$$T_{1/2} = \frac{\Delta H^0}{(R \cdot \ln c_0 + \Delta S^0)}$$

**Tabla 5.8.-** Parámetros termodinámicos del equilibrio dumbbell-dímero a partir del ajuste  $K$  vs  $T$ . La  $T_{1/2}$  está calculada para una concentración de oligonucleótido 5mM.

Secuencia	$\Delta H^0$ kJ/mol	$\Delta S^0$ J/mol·°K	$\Delta G^0_{298}$ kJ/mol	$T_{1/2}$ °K
d<pCATTTCATT>	-48±3	-139±10	-6±1	244
d<pTGCTCGCT>	-27±1	-30±3	-18.0±0.01	308
d<pTGCATGCA>	-97±6	-259±21	-20±0.1	306

El hecho de que para la secuencia d<pCATTTCATT> a una concentración de 5 mM se haya calculado un valor de  $T_{1/2}$  de 244°K, muestra la baja estabilidad de la correspondiente forma dimérica y pone de manifiesto la necesidad de alcanzar concentraciones muy superiores (por encima de 20 mM) para poder llegar a observar dicha forma incluso a baja temperatura. Los valores para las otras dos secuencias están de acuerdo con la observación experimental de que por debajo de 20°C las formas diméricas son mayoritarias.

#### 5.4.- ESTUDIO DEL EQUILIBRIO DÍMERO-FORMA DESESTRUCTURADA.

Este equilibrio, al igual que el equilibrio *dumbbell*-random coil, lo hemos estudiado tanto a partir de datos de RMN como a partir de curvas de fusión por DC. No obstante, en este caso los parámetros obtenidos no son directamente comparables ya que, mientras en RMN los experimentos se realizaron a elevada concentración de oligonucleótido y en H<sub>2</sub>O, para las curvas de DC se utilizaron condiciones salinas que garantizaran la formación de dímero a baja concentración de oligonucleótido. Para poder comparar los parámetros termodinámicos en ambas condiciones sería necesario calcular el efecto que la sal ejerce sobre ellos. Aunque se ha demostrado en este trabajo que una mayor fuerza iónica del medio tiende a estabilizar las formas diméricas, el estudio cuantitativo de este efecto estabilizador excede los objetivos de la presente Tesis Doctoral.

##### 5.4.1.- ESTUDIO POR RMN. CÁLCULO DE POBLACIONES.

Al igual que en el caso del equilibrio *dumbbell*-dímero, debemos calcular la población de cada una de las especies a diferente temperatura y ajustar curvas del tipo  $K$  vs  $T$ . En este caso las poblaciones se han calculado a partir de series de espectros a diferente temperatura y elevada concentración de oligonucleótido, tal que no se observen las señales correspondientes a la forma *dumbbell*. En la Figura 5.10 se muestra la serie correspondiente a la secuencia d<pCGCTTCATT> que, recordémoslo no forma *dumbbell*.

En el caso de las secuencias d<pCATTTCATT> y d<pTGCTCGCT> no se realizaron experimentos de RMN en condiciones de 100% dímero por lo que no ha sido posible llevar a cabo este estudio.

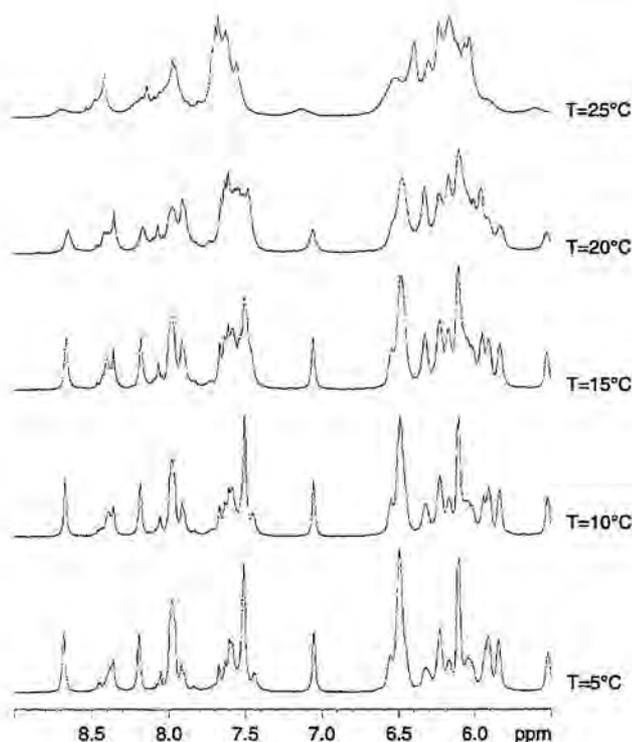


Figura 5.10.- Serie de espectros  $d\langle pCGCTCATT \rangle$  5 mM a diferente temperatura.

De forma análoga a la descrita en el apartado 5.3.- se llevaron a cabo los ajustes para las secuencias  $d\langle pCGCTCATT \rangle$ ,  $d\langle pTGCATGCA \rangle$ ,  $d\langle pAGCTAGCT \rangle$  y  $d\langle pCCGTCCGT \rangle$ .

En el caso de  $d\langle pCGCTCATT \rangle$ , puesto que se disponía de series de espectros a distinta temperatura realizadas a diferente concentración de oligonucleótido, el valor de la  $K$  para cada valor de temperatura se ha podido obtener representando la variación de la variable  $x = I_D/I_M$  (donde  $I_M$  corresponde a la integración de una señal de la forma monomérica desestructurada) con la concentración, tal como se describe en la bibliografía<sup>19,20</sup>:

La  $K$  en nuestro caso se expresa como:

$$K = \frac{x \cdot (1+x)}{2 \cdot c_0}$$

De este modo, la representación de  $x \cdot (1+x)$  vs  $c_0$ , nos permite obtener la  $K$  como la pendiente de la recta:  $x \cdot (1+x) = 2Kc_0$ . Representando el valor de  $x \cdot (1+x)$  frente a tres concentraciones de oligonucleótido: 2.4, 5 y 10 mM hemos obtenido el valor de la  $K$  para cuatro temperaturas. Tal como se puede observar en los gráficos, los puntos se ajustan bien para  $T = 283$ , 288 y 293°K, sin embargo, los puntos correspondientes a 278°K presentan una mayor dispersión. La baja estabilidad que presenta esta forma dimerica hace que no se disponga de puntos a una temperatura superior a 293°K, ya que a partir de esta temperatura se observan únicamente las señales correspondientes a la forma totalmente desestructurada. Los valores de la  $K$  obtenida en cada caso se muestran en la Tabla 5.9.

<sup>19</sup> Poland, D. «Cooperative Equilibria in Physical Biochemistry». Monographs on Physical Biochemistry 1978. Clarendon Press, Oxford.

<sup>20</sup> Biopolymers 1974, 13, 687.

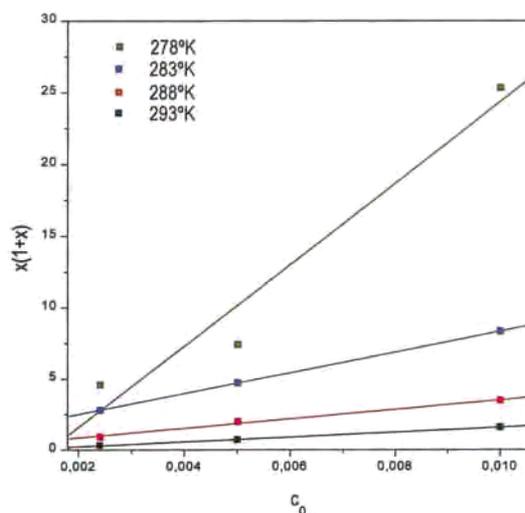


Figura 5.11.- Ajuste de las curvas  $x(1+x)$  vs  $c_0$  para concentraciones de 2.4, 5 y 10 mM de  $d\langle pCGCTCATT \rangle$ , para cuatro temperaturas: 278, 283, 288 y 293 K.

Tabla 5.9.- Valores de la K para  $d\langle pCGCTCATT \rangle$  obtenidos a partir de los ajustes  $x(1+x)$  vs T.

T °K	$k_{eq}$ M <sup>-1</sup>	R
278	1423.1	0.9765
283	362.9	0.9999
288	165.6	0.9951
293	84.7	0.9946

Para obtener los parámetros termodinámicos correspondientes al equilibrio se han ajustado los valores de K vs T según la expresión descrita en el apartado 5.3.-. En el ajuste, Figura 5.12, no se ha tenido en cuenta el punto correspondiente a T=278 K ya que su inclusión disminuye considerablemente la calidad del ajuste.

En los otros tres casos, disponemos únicamente de espectros a distinta temperatura realizados a una única concentración de oligonucleótido, 5 mM. En el caso de  $d\langle pCCGTCCGT \rangle$ , al no existir una buena dispersión de las señales, se calculó el % de cada una de las especies por la relación entre la integración de señales suma de random coil y dímero y la K se determinó como función de  $\alpha^{18}$ . Las tablas con los valores de K para cada una de las secuencias se muestran en el Anexo D. Las representaciones correspondientes de K vs T aparecen en la Figura 5.12.

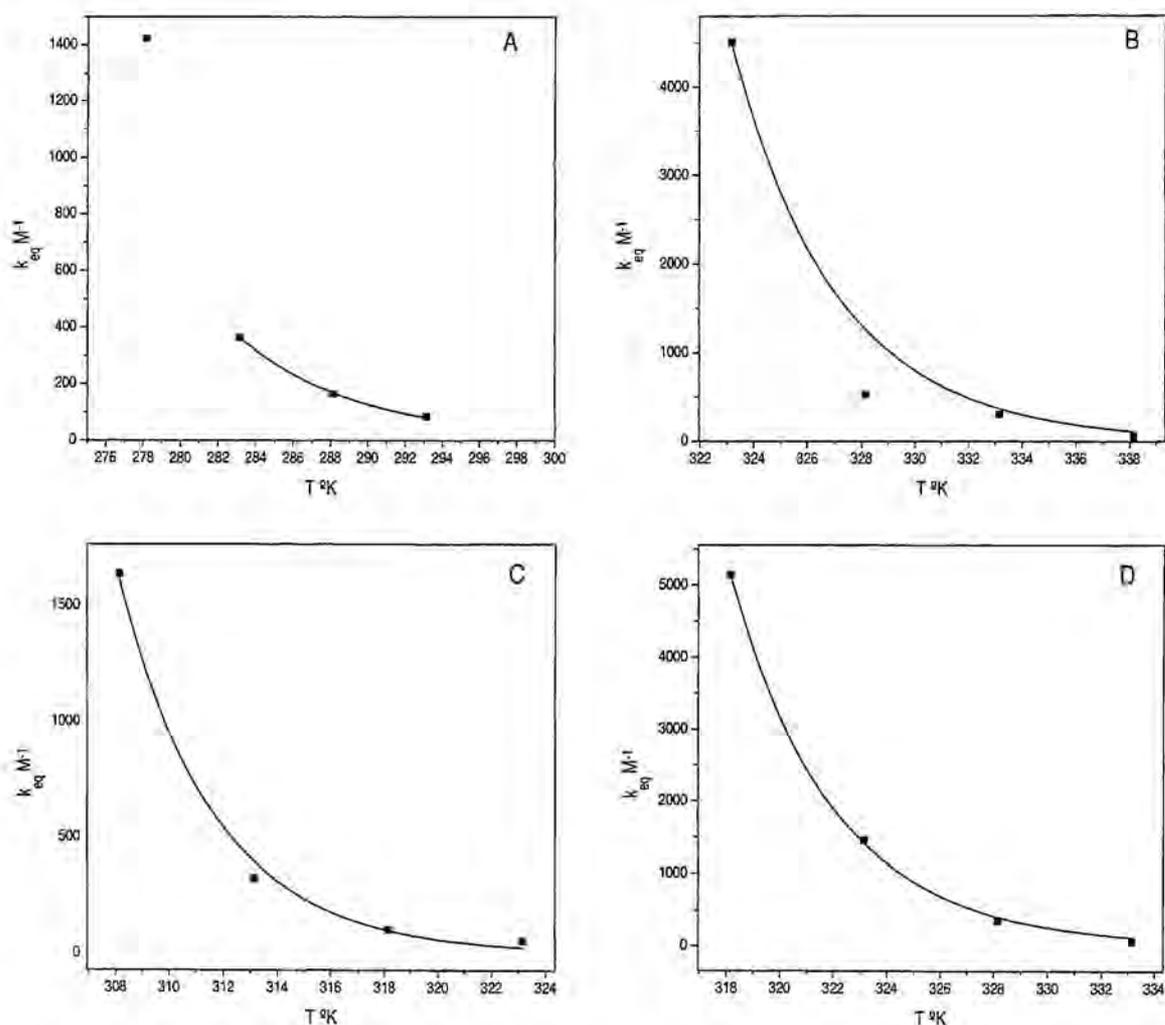


Figura 5.12.- Ajuste de los puntos  $k$  vs  $T$  para el equilibrio dímero-random de A)  $d\langle pCGCTCATT \rangle$ , B)  $d\langle pTGCATGCA \rangle$ , C)  $d\langle pAGCTAGCT \rangle$  y D)  $d\langle pCCGTCCGT \rangle$ .

En el caso de  $d\langle pTGCATGCA \rangle$ , como se puede observar en la figura, hay un punto que se aparta mucho de los otros tres, por lo que no ha sido tenido en cuenta en el ajuste. Los parámetros termodinámicos obtenidos a partir de estos ajustes figuran en la Tabla 5.10.

Es interesante recalcar que los distintos equilibrios termodinámicos que estamos estudiando en este capítulo no son independientes. A la vista del diagrama de la Figura 5.1 de este capítulo, queda claro que deben cumplirse las relaciones:  $\Sigma\Delta H^0=0$  y  $\Sigma\Delta S^0=0$ <sup>13</sup> o, lo que es lo mismo:

$$\begin{aligned}\Delta H^0_{\text{dímero-dumbbell}} + 2\Delta H^0_{\text{dumbbell-random}} + \Delta H^0_{\text{random-dímero}} &= 0 \\ \Delta S^0_{\text{dímero-dumbbell}} + 2\Delta S^0_{\text{dumbbell-random}} + \Delta S^0_{\text{random-dímero}} &= 0\end{aligned}$$

Puesto que en el caso de  $d\langle pTGCATGCA \rangle$ , los parámetros termodinámicos correspondientes a los tres equilibrios se han obtenido de forma totalmente independiente, estas relaciones pueden utilizarse para comprobar los resultados obtenidos. Si consideramos los

valores experimentales correspondientes a los equilibrios dímereo-*dumbbell* y *dumbbell*-random coil tenemos que  $\Delta H^0_{\text{random-dímereo}} = -214.3 \text{ kJ/mol}$  y  $\Delta S^0_{\text{random-dímereo}} = -618.8 \text{ J/mol}\cdot\text{K}$ . A partir del ajuste de la  $K$  vs  $T$  hemos obtenido los valores  $\Delta H^0_{\text{random-dímereo}} = -224.2 \text{ kJ/mol}$  y  $\Delta S^0_{\text{random-dímereo}} = -623.8 \text{ J/mol}\cdot\text{K}$ , por lo que parece que, aunque los ajustes llevan implícitos una serie de errores, los valores son suficientemente fiables como para permitirnos el análisis de la estabilidad de las formas díméricas en función de la secuencia.

En las secuencias  $d\langle pTGCTCGCT \rangle$  y  $d\langle pCATTTCATT \rangle$  se pueden estimar los parámetros de la misma forma, obteniendo los valores que aparecen en la Tabla 5.10. En el caso particular de  $d\langle pTGCTCGCT \rangle$ , en el que las señales de la forma dímérica se mantienen hasta temperaturas muy elevadas, se dispone de puntos correspondientes a la  $K$  entre la forma dímérica y la monomérica a temperaturas superiores a la  $T_m$  del *dumbbell*. Estos puntos corresponden únicamente al equilibrio entre la forma dímérica y la forma desestructurada. Por este motivo decidimos ajustar los tres últimos puntos  $K$  vs  $T$ , que se muestran en la gráfica que aparece en la Figura 5.9B. El resultado de dicho ajuste es:  $\Delta H^0_{\text{random-dímereo}} = -176.6 \text{ kJ/mol}$  y  $\Delta S^0_{\text{random-dímereo}} = -494.7 \text{ J/mol}\cdot\text{K}$ . Estos valores, aunque ligeramente inferiores, se corresponden bastante bien con los indicados en la Tabla 5.10, lo que está indicando, una vez más, que la aproximación realizada permite obtener valores suficientemente aceptables.

**Tabla 5.10.-** Parámetros termodinámicos asociados a la formación de la forma dímérica para la secuencia  $d\langle pCGCTCATT \rangle$ ,  $d\langle pTGCTCGCT \rangle$ ,  $d\langle pAGCTAGCT \rangle$  y  $d\langle pCCGTCCGT \rangle$ .

Secuencia	$\Delta H^0 \text{ kJ/mol}$	$\Delta S^0 \text{ J/mol}\cdot\text{K}$	$\Delta G^0 \text{ kJ/mol}$	$T_m \text{ }^\circ\text{K}^a$
$d\langle pCGCTCATT \rangle$	$-104 \pm 3$	$-320 \pm 11$	$-9 \pm 6$	$287 \pm 3$
$d\langle pTGCTCGCT \rangle$	$-224 \pm 12$	$-624 \pm 47$	$-38 \pm 26$	$335 \pm 8$
$d\langle pAGCTAGCT \rangle$	$-226 \pm 13$	$-673 \pm 54$	$-25 \pm 29$	$315 \pm 8$
$d\langle pCCGTCCGT \rangle$	$-221 \pm 4$	$-625 \pm 17$	$-35 \pm 10$	$330 \pm 3$
$d\langle pCATTTCATT \rangle^b$	-183	-570	-13	298
$d\langle pTGCTCGCT \rangle^b$	-187	-520	-32	331

<sup>a</sup>  $T_m$  calculada para  $[\text{oligonucleótido}] = 5 \text{ mM}$ .

<sup>b</sup> Valores calculados a partir de las Tablas 1, 2 y 8.

Como se puede observar en los valores de  $\Delta G^0$ , las formas díméricas son siempre más estables que los *dumbbells* correspondientes. Una discusión comparativa de las estabildades de los diferentes dímeros se da en la sección 5.4.5.

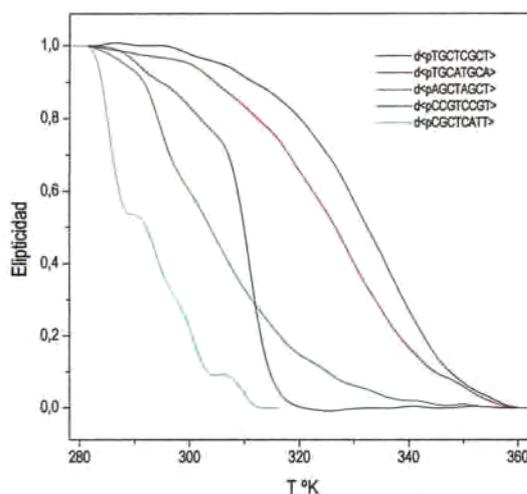
#### 5.4.2.- ESTUDIO POR DC: CURVAS DE FUSIÓN EN MEDIO SALINO.

Antes de proceder a la discusión de los parámetros termodinámicos obtenidos para los diferentes dímeros, es interesante comprobar estos resultados, aunque sea de manera cualitativa, mediante una técnica independiente. Como en el caso de las formas *dumbbell*, esta técnica será el DC.

Como ya se demostró en el Capítulo 4, existen condiciones salinas que nos permiten obtener la forma dímérica de las distintas secuencias, excepto en el caso de  $d\langle pCATTTCATT \rangle$  para el que no se encontraron unas condiciones adecuadas, a baja concentración de oligonucleótido. Esto va a permitir realizar curvas de fusión por DC en dichas condiciones y obtener la temperatura de fusión de la forma dímérica en medio salino. El valor obtenido no va a ser comparable con el obtenido a partir de RMN, ya que, además de tratarse de un equilibrio dependiente de la concentración, la presencia de sal en el medio ejerce un efecto estabilizador

que debería tenerse en cuenta. No obstante, esperamos observar la misma correlación, en cuanto a la estabilidad de los distintos dímeros, que la observada a partir de las curvas K vs T.

A continuación se muestran las curvas de fusión en medio salino correspondientes a las secuencias d<pTGCTCGCT>, d<pTGCATGCA>, d<pAGCTAGCT>, d<pCCGTCCGT> y d<pCGCTCATT>.



**Figura 5.13.-** Curvas de fusión por DC normalizadas en medio salino:  $\text{Na}_2\text{PIPES}$  10mM pH=7,  $\text{MgCl}_2$  10mM y  $\text{NaCl}$  100mM, para las secuencias: d<pTGCTCGCT> 5.8  $\mu\text{M}$ , d<pTGCATGCA> 5.5  $\mu\text{M}$ , d<pAGCTAGCT> 5.1  $\mu\text{M}$ , d<pCCGTCCGT> 6.2  $\mu\text{M}$  y d<pCGCTCATT> 5.0  $\mu\text{M}$ . En el caso de la secuencia d<pCGCTCATT> se ha utilizado una concentración de  $\text{MgCl}_2$  de 500mM.

A partir de estas curvas se han obtenido, para las distintas formas diméricas, los siguientes valores de  $T_m$ :

**Tabla 5.11.-** Valores de  $T_m$  para los distintos dímeros obtenidos a partir de la derivada primera de la curva de fusión y comparación con los valores de  $T_m$  obtenidos por RMN.

Secuencia	$T_m$ °K DC <sup>a</sup>	$T_m$ °K RMN <sup>b</sup>
d<pTGCTCGCT>	332	-
d<pTGCATGCA>	326	309
d<pAGCTAGCT>	303	292
d<pCCGTCCGT>	310	304
d<pCGCTCATT>	292	246

<sup>a</sup> concentración de oligonucleótido ~5-6 $\mu\text{M}$ .

<sup>b</sup>  $T_m$  calculada para una concentración de oligonucleótido 5 $\mu\text{M}$ .

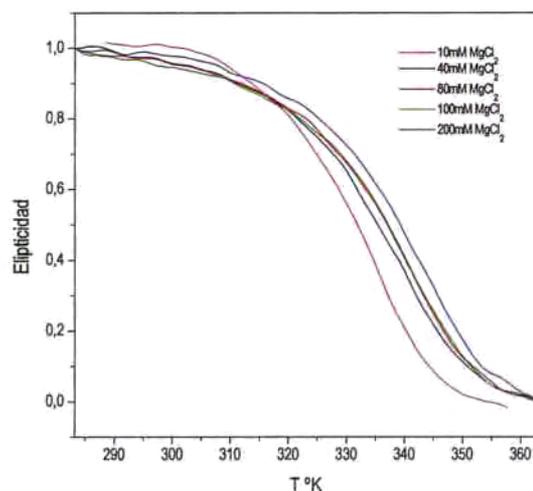
Las temperaturas de fusión de los dímeros obtenidas a partir de las curvas de fusión de DC siguen, en efecto, la misma gradación que la observada a partir de los datos de RMN que figuran en la Tabla 5.10. Si calculamos, a partir de los parámetros termodinámicos obtenidos por RMN, el valor de  $T_m$  a una concentración de oligonucleótido de 5  $\mu\text{M}$ , Tabla 5.11, se obtienen valores significativamente inferiores a los obtenidos por DC. Esta diferencia es explicable en base al enorme efecto estabilizador que ejerce la presencia de sal en el medio. Resulta especialmente destacable el caso de d<pCCGTCCGT>, en el que se observa la mayor

diferencia en el valor de la  $T_m$ , así como el hecho de que para la secuencia de pares mixtos d<pCGCTCATT> se obtenga un valor superior de temperatura de fusión a baja concentración de oligonucleótido y en medio salino (292°K), que a alta concentración de oligonucleótido (287°K), comparar Tablas 11 y 12.

Para poder comparar los valores de  $T_m$  en presencia y ausencia de sal, sería necesario, por una parte, cuantificar de forma precisa el efecto que la sal ejerce sobre la estabilidad de las distintas formas diméricas. Por otra parte, necesitaríamos conocer como varía la temperatura de fusión con la concentración de oligonucleótido. Este estudio no ha sido posible llevarlo a cabo en el presente trabajo, aunque sí se han realizado algunos estudios preliminares que se muestran a continuación.

#### 5.4.3.- ESTUDIO POR DC DE LA INFLUENCIA DE LA SAL EN LA ESTABILIDAD DE LA FORMA DIMÉRICA.

El hecho de que a baja concentración de oligonucleótido sea posible inducir la formación de la forma dimérica mediante la adición de sal en el medio, es una prueba del efecto estabilizador que la sal ejerce sobre este tipo de estructuras. En particular, es sorprendente la capacidad inductora observada para el  $MgCl_2$ , ya que pequeñas cantidades de esta sal convierten a la forma dimérica en la única especie presente. Por este motivo, creímos importante estimar el efecto de dicha sal en la  $T_m$  de la forma dimérica. Para ello realizamos fusiones por DC de la secuencia d<pTGCTCGCT> en condiciones  $Na_2PIPES$  10mM, pH=7 y distintas concentraciones de  $MgCl_2$ .



**Figura 5.14.-** Curvas de fusión por DC normalizadas de d<pTGCTCGCT> en  $Na_2PIPES$  10 mM pH=7 y distintas concentraciones de  $MgCl_2$ , de 10-200 mM. En todos los casos la concentración de oligonucleótido es de 7.2  $\mu M$ .

Los valores de  $T_m$  para cada concentración de  $MgCl_2$  obtenidos a partir de las curvas de fusión se muestran en la Tabla 5.12.

**Tabla 5.12.-** Valores de  $T_m$  para el dímero de  $d\langle pTGCTCGCT \rangle$  a distintas concentraciones de  $MgCl_2$ .

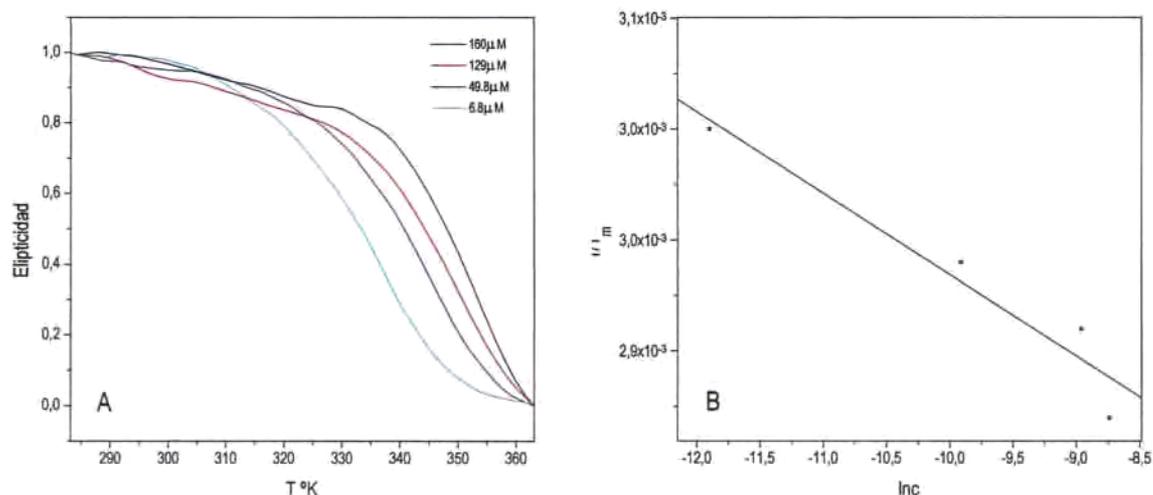
[MgCl <sub>2</sub> ] mM	$T_m$ °K
10	333
40	336
80	337
100	337
200	339

Efectivamente, se observa un aumento del valor de  $T_m$  debido a la presencia de cantidades crecientes de  $MgCl_2$ . De cara a cuantificar dicho efecto sería necesario realizar este mismo experimento a diferentes concentraciones de oligonucleótido para cada concentración de  $MgCl_2$ .

#### 5.4.4.- DEPENDENCIA DE LA $T_M$ DE LA FORMA DIMÉRICA CON LA CONCENTRACIÓN.

Mediante el estudio de la variación de  $T_m$  con la concentración de oligonucleótido pueden extraerse parámetros termodinámicos del equilibrio de asociación bimolecular<sup>18</sup> Aunque no ha sido posible llevar a cabo un estudio sistemático para todas las secuencias, en esta sección ilustraremos esta metodología mediante la estimación de los parámetros termodinámicos de la transición dímero-random coil de  $d\langle pTGCTCGCT \rangle$ .

Se realizaron curvas de fusión por DC para  $d\langle pTGCGCGCT \rangle$  en medio salino ( $Na_2PIPES$  10 mM, pH=7, 10 mM  $MgCl_2$ ) a diferente concentración de oligonucleótido, Figura 5.15A.



**Figura 5.15.-** A) Curvas de fusión por DC de la forma dimérica de  $d\langle pTGCTCGCT \rangle$  a distintas concentraciones de oligonucleótido. B) Representación  $1/T_m$  vs  $\ln c_0$  para la forma dimérica de  $d\langle pTGCTCGCT \rangle$ .

Tal como se puede observar en las curvas de fusión, los valores de  $T_m$  muestran una fuerte dependencia con la concentración, observándose un desplazamiento del punto de inflexión de las curvas hacia temperaturas más elevadas a medida que aumenta la concentración. En la Tabla 5.13 se muestran los valores de  $T_m$  obtenidos en cada caso:

**Tabla 5.13.-** Valores de  $T_m$  para la forma dimérica de  $d\langle pTGCTCGCT \rangle$  a diferente concentración de oligonucleótido.

$[d\langle pTGCTCGCT \rangle] \mu M$	$T_m \text{ } ^\circ K$
6.8	333
49.8	341
129.0	344
160.0	348

La representación de  $1/T_m$  vs  $\ln c_0$  permite obtener los parámetros termodinámicos correspondientes al equilibrio dímero-random coil. Esta representación se muestra en la Figura 5.15B.

A partir de este ajuste se obtienen, para el equilibrio de formación en presencia de sales de la forma dimérica los siguientes parámetros termodinámicos:  $\Delta H^0 = -226.58$  kJ/mol,  $\Delta S^0 = -582.31$  J/mol $\cdot^\circ K$ ,  $\Delta G^0 = -53.05$  kJ/mol. Comparando este último valor con el obtenido en  $H_2O$  mediante RMN ( $\Delta G^0 = -32$  kJ/mol), se comprueba, una vez más, el fuerte efecto estabilizador que produce el  $MgCl_2$  en las formas diméricas.

#### 5.4.5.- EL EFECTO DE LA SECUENCIA EN LA ESTABILIDAD DEL MOTIVO BI-LOOP.

Con los diferentes datos termodinámicos del equilibrio dímero-random coil recopilados en este capítulo, se puede intentar dar respuesta a las cuestiones planteadas en el apartado 4.1 sobre el efecto de la secuencia de nucleótidos en la estabilidad del motivo bi-loop.

En primer lugar, a partir de los valores de  $\Delta G^0_{r-D 298}$  que se muestran en la Tabla 5.10, se observa claramente que hay dos grupos de secuencias: las que forman dos pares de bases G-C y las que no, siendo las primeras significativamente más estables. Es interesante observar que, para que el efecto estabilizante de los pares G-C tenga lugar, deben formarse los cuatro pares, puesto que el octámero mixto, de secuencia  $d\langle pCGCTCATT \rangle$ , con un solo par G-C por tétrada, tiene una estabilidad similar al  $d\langle pCATTTCATT \rangle$ , que sólo forma pares A-T. En principio, la baja estabilidad del bi-loop  $d\langle pCGCTCATT \rangle$  podría resultar sorprendente puesto que la diferencia en estabilidad de los pares Watson-Crick G-C y A-T no es tan acusada en dúplex de DNA<sup>21</sup>. Además, era de esperar que este octámero tuviese una estabilidad intermedia entre los valores de  $\Delta G^0_{r-D 298}$  de  $d\langle pTGCTCGCT \rangle$  y  $d\langle pCATTTCATT \rangle$ . Las estructuras tridimensionales determinadas en el Capítulo 2 permiten dar una explicación a este hecho. Como se puede observar en la Figura 3.13, cada tétrada de  $d\langle pTGCTCGCT \rangle$  está estabilizada por seis puentes de hidrógeno intermoleculares (tres por cada par G-C) más otros dos puentes intra-moleculares. Estos puentes intramoleculares no están presentes en las tétradas de  $d\langle pCATTTCATT \rangle$  y tampoco pueden formarse en el caso de  $d\langle pCGCTCATT \rangle$ . Por tanto, el número de puentes pasa de 8 en una tétrada G:C:G:C a 5 en G:C:A:T, frente a los 4 presentes en A:T:A:T. Esto explica que la estabilidad de  $d\langle pCGCTCATT \rangle$  sea más parecida a la de  $d\langle pCATTTCATT \rangle$  que a la de  $d\langle pTGCTCGCT \rangle$ .

La diferencia de estabildades entre  $d\langle pCGCTCATT \rangle$  y  $d\langle pCATTTCATT \rangle$  es pequeña y puede estar relacionada con la pérdida de simetría en la estructura de  $d\langle pCGCTCATT \rangle$ . Aunque esta estructura no se ha calculado, no sería extraño que el apilamiento entre las bases no fuese tan completo como en los otros casos.

<sup>21</sup> Breslauer, K.J.; Frank, R.; Blöcker, H.; Marky, L.A. *Proc. Natl. Acad. Sci. USA* **1988**, 83, 3746-3750.

El segundo factor en importancia en la estabilidad del bi-loop es la presencia de purinas en el loop. Como era de esperar a partir de las estructuras de los bi-loops calculadas en esta tesis, el efecto desestabilizante sólo se produce si la purina se halla situada en la segunda posición del loop, puesto que es la única totalmente expuesta al disolvente. Es interesante comparar este efecto con el observado en los *dumbbells*. Si comparamos la diferencia de  $\Delta G^0_{298}$  entre d<pAGCTAGCT> y d<pTGCTCGCT> en las Tablas 5.5 y 5.10, observamos que la purina expuesta al disolvente desestabiliza la estructura de una manera muy similar (aprox. 7 kJ/mol). Obviamente, este efecto resulta más dramático en el caso de los *dumbbells*, que son menos estables que los dímeros.

Es muy interesante observar que la secuencia con una purina en la primera posición del loop, d<pTGCATGCA>, forma el bi-loop más estable. Esta secuencia es muy similar a la del único oligonucleótido lineal donde se ha observado este motivo estructural, d(GCATGCA). Puesto que en la estructura de rayos X de esta molécula se observan dos pares A·A, los valores de  $\Delta G^0_{298}$  aquí obtenidos inducen a pensar que la estructura en disolución de d<pTGCATGCA> también puede estar estabilizada por dos pares AA. De ser así, esta interacción sería más eficaz (aprox. 5kJ/mol) en la estabilización del bi-loop que la interacción hidrofóbica entre los metilos de timinas, observada en las estructuras de d<pTGCTCGCT> y d<pCATTTCATT>. Además de la formación de dichos pares, la mayor estabilidad del bi-loop con purina en la primera posición del loop podría ser debida a fuerzas de apilamiento más favorables que las que se obtienen con pirimidina en esta posición.

Finalmente, también es interesante mencionar que los valores de  $\Delta G^0_{298}$  para la secuencia d<pCCGTCCGT> son muy similares a los de d<pTGCTCGCT>. Este resultado contrasta completamente con el obtenido para los *dumbbells*. En estos últimos, la elevada inestabilidad de los loops 5'-G-loop-C-3' impide la formación del *dumbbell* correspondiente. Sin embargo, este hecho no parece tener influencia en la estabilidad del bi-loop. Hay que tener en cuenta que la orientación de las bases que flanquean el dinucleótido del loop en un *dumbbell* es muy diferente de la que se da en una estructura bi-loop.

A modo de respuesta a las preguntas suscitadas al comienzo del Capítulo 4, podemos resumir los resultados aquí obtenidos diciendo que el efecto más importante en la estabilidad del bi-loop es la naturaleza de las bases que forman las tétradas, siendo las del tipo G:C:G:C las más estables. Los residuos que forman los loops también influyen, siendo el orden de estabildades RY > YY > YR (R= purina, Y= pirimidina). Finalmente, el orden de las bases que forman los pares Watson-Crick (5'-GC-3' o 5'-CG-3'), no parece tener una influencia significativa en la estabilidad.

---

*CONCLUSIONES*

---

- La eliminación del grupo protector del fosfato 3'-terminal antes de la elongación de la cadena oligonucleotídica permite evitar las reacciones secundarias detectadas durante las etapas de desprotección de los grupos fosfato y mejorar los rendimientos de obtención de oligonucleótidos cíclicos. Estos resultados confirman que el ensamblaje del oligonucleótido por el método del fosfito trister se puede llevar a cabo sin problemas en presencia de fosfatos desprotegidos.

- Los controles realizados demuestran que en las distintas etapas de las que consta el procedimiento de síntesis no se produce pérdida de cadenas oligonucleotídicas, por lo que el rendimiento de obtención de producto cíclico depende en gran medida de la efectividad de la reacción de ciclación. La ciclación del precursor lineal es una reacción delicada, en la que además de la secuencia influyen drásticamente los factores experimentales. El reactivo con el que se obtienen mejores resultados es el habitualmente utilizado, MSNT, habiendo resultado infructuosos los intentos de encontrar un reactivo más efectivo.

- La aplicación del método a la preparación de distintos octámeros cíclicos a gran escala confirma que la estrategia de síntesis permite el escalado, obteniéndose rendimientos aceptables y crudos de elevado grado de pureza.

La purificación mediante HPLC, utilizando una columna de poli(estireno-divinilbenceno) semipreparativa, se confirma como el método más adecuado para el aislamiento de oligonucleótidos cíclicos de este tamaño.

- La estructura cristalográfica de  $d\langle pCATTTCATT \rangle$  revela la existencia de un nuevo motivo estructural del DNA, tipo tétrada, que hemos denominado bi-loop. Se trata de una estructura dimérica en la que las dos unidades de octámero se disponen en sentido antiparalelo. La estructura está estabilizada por cuatro pares A-T intermoleculares que forman dos apilamientos, de dos pares cada uno, enfrentados por el surco menor y en los que participan también las timinas que ocupan la primera posición de los mini-loops. Entre ambos apilamientos existe una inclinación relativa de  $\sim 37^\circ$ . El centro de la estructura está ocupado por un ión de sodio que establece coordinación octaédrica con los O2 de las timinas apareadas y dos moléculas de  $H_2O$ .

- El estudio en disolución por RMN de distintos octámeros cíclicos,  $d\langle pCATTTCATT \rangle$ ,  $d\langle pTGCTCGCT \rangle$ ,  $d\langle pCGCTCATT \rangle$ ,  $d\langle pCCGTCCGT \rangle$ ,  $d\langle pTGCATGCA \rangle$  y  $d\langle pAGCTAGCT \rangle$ , pone de manifiesto la existencia de un equilibrio entre una forma monomérica y un producto de asociación que, por la elevada simetría observada en los espectros, se concluye que corresponde a un dímero en el que los dos ciclos se disponen en forma antiparalela. La forma dimérica es, para todas las secuencias estudiadas, una forma perfectamente estructurada estabilizada por puentes intermoleculares.

- La determinación de la estructura en disolución de las formas diméricas de  $d\langle pCATTTCATT \rangle$  y  $d\langle pTGCTCGCT \rangle$  revela que se trata del mismo motivo estructural observado en estado sólido, el motivo bi-loop. En el segundo caso, se forman tétradas G:C:G:C y se ha podido confirmar la existencia adicional de puentes intramoleculares H21G-O2C.

- La elevada similitud observada en los espectros de RMN entre estas dos secuencias y el resto de secuencias estudiadas, indica que las estructuras diméricas correspondientes a d<pTGCATGCA>, d<pAGCTAGCT>, d<pCGCTCATT> y d<pCCGTCCGT> son también estructuras bi-loop.
- La observación en disolución del mismo motivo estructural en todos los octámeros estudiados constituye un progreso en el conocimiento de las estructuras tetraméricas del DNA, y aporta nuevos ejemplos de tétradas formadas por purinas y pirimidinas del tipo A:T:A:T, G:C:G:C y G:C:A:T.
- Las secuencias que dan lugar a la formación de bi-loop están constituidas por dos dinucleótidos autocomplementarios separados por mini-loops de dos residuos. En la mayoría de las secuencias estudiadas el dinucleótido es 5'-purina-pirimidina-3', excepto en la secuencia d<pCCGTCCGT> en que el orden es el inverso. Fuera de estos rasgos generales, hasta el momento no se han encontrado otros requisitos secuenciales para la formación de dichas estructuras. El motivo bi-loop, lejos de ser una curiosidad cristalográfica, parece tener un carácter muy general que incita a plantearse su posible relevancia biológica.
- El estudio termodinámico del equilibrio de formación de las estructuras diméricas por RMN y DC indica que existen dos factores que afectan de forma decisiva a su estabilidad. El más importante lo constituye la naturaleza de los pares de bases, siendo significativamente más estables las estructuras estabilizadas por cuatro pares G-C que las estabilizadas por cuatro pares A-T o por pares mixtos, 2 pares A-T y 2 pares G-C. El segundo factor en importancia es la presencia de un residuo de purina en el mini-loop, que cuando ocupa la segunda posición ejerce un efecto altamente desestabilizador.
- Como alternativa a la RMN, el DC resulta una técnica apropiada para detectar la formación de estructuras bi-loop de forma rápida. Dicha técnica permite, además, seguir de forma precisa tanto el proceso de desestructuración como la influencia de la sal en la formación de este tipo de estructuras. Estudios preliminares muestran que la presencia de sal en el medio, y muy en particular de pequeñas cantidades de MgCl<sub>2</sub>, ejerce un efecto altamente estabilizador de la estructura bi-loop.  
La espectroscopia UV, sin embargo, resulta menos útil, ya que se observan variaciones de cromicidad muy pequeñas al aumentar la temperatura.
- En algunas de las secuencias estudiadas en disolución, d<pCATTTCATT>, d<pTGCTCGCT>, d<pTGCATGCA> y d<pAGCTAGCT>, el monómero, que se encuentra en equilibrio con la forma dimérica, adopta una estructura tipo *dumbbell* con formación de puentes intramoleculares.
- La estabilidad de los *dumbbells* también se ve afectada de forma importante por la naturaleza de los pares de bases y por la composición de los mini-loops. En los cuatro *dumbbells* mencionados se ha observado un orden de estabilidad d<pTGCATGCA> > d<pTGCTCGCT> > d<pCATTTCATT> > d<pAGCTAGCT>. En el caso de los *dumbbells*, la localización de una purina en la segunda posición del mini-loop afecta drásticamente a la

estabilidad de la estructura, siendo el efecto desestabilizador mayor al observado para las formas diméricas.

- El hecho de no haberse observado formación de *dumbbell* en el oligonucleótido d<pCCGTCCGT> confirma que, en horquillas de DNA de secuencia 5'-G-pirimidina-pirimidina-C-3', no hay apareamiento G-C y se forma preferentemente un loop de cuatro residuos en vez de un *hairpin-loop* de dos.

---

*MATERIALES Y MÉTODOS*

---

## 1.- PROCEDIMIENTOS GENERALES.

### 1.1.- INSTRUMENTACIÓN GENERAL.

Los espectros de resonancia magnética nuclear de  $^1\text{H}$ ,  $^{13}\text{C}$  y  $^{31}\text{P}$  se han realizado en aparatos Bruker Avance DXR-500MHz, Varian Unity-300MHz y Gemini-200MHz, ubicados en la Unidad de RMN de Campo Alto de la UB. Los desplazamientos químicos de  $^1\text{H}$  y  $^{13}\text{C}$  se expresan en ppm respecto a tetrametilsilano. En el caso del  $^{31}\text{P}$  se han utilizado  $\text{H}_3\text{PO}_4$  85% en  $\text{D}_2\text{O}$  ( $\delta$ : 0 ppm) y trimetil fosfito (TMP) en  $\text{CDCl}_3$  ( $\delta$ : 141.6 ppm) como patrones externos. Los experimentos de resonancia de  $^{31}\text{P}$  en fase gel se han realizado suspendiendo la resina en  $\text{CDCl}_3$  en tubos especiales de fondo plano que, con la ayuda de un émbolo, permiten concentrar la muestra en la zona de irradiación (tubos SHIGEMI Co, LTD, Japón).

Los espectros de masas se han realizado en el Servicio de Espectrometría de Masas de la División III de la UB. Los espectros de impacto electrónico (EM IE) se han realizado en un aparato Hewlett-Packard 5988. Los espectros por ionización en electroespray (EM ES) se han realizado en un aparato VG Quattro (Fison Instruments), utilizando un voltaje del capilar de 3.5kV. Los espectros de masas adquiridos mediante desorción iónica por láser asistida por matriz, con detección de tiempo de vuelo, (EM MALDI-TOF) se han obtenido en un aparato Bruker Biflex™ III con un láser de  $\text{N}_2$  de 337nm y pulsos de 3ns.

Los espectros de infrarojo se han adquirido en un espectrofotómetro Nicolet 510 FT-IR.

Los espectros de absorción de UV-Vis se han realizado indistintamente en un espectrofotómetro Perkin Elmer Lambda 5 UV-Vis o bien en un Varian Cary 5E UV-Vis-NIR.

Los puntos de fusión se han obtenido en aparatos Gallenkamp.

Las medidas de pH se han realizado en un pHmetro Crison micropH 2002. En las medidas en  $\text{D}_2\text{O}$  no se ha realizado ninguna corrección.

La liofilización de las disoluciones acuosas se ha realizado en aparatos Virtis, modelos Freezemobile QD6 o Freezemobile 12EL.

### 1.2.- REACTIVOS Y DISOLVENTES.

El ACN anh. se obtiene por destilación del acetonitrilo de calidad HPLC (Scharlau) sobre  $\text{CaH}_2$  en polvo y se guarda sobre piedras de  $\text{CaH}_2$  bajo atmósfera inerte de argón hasta su utilización.

El THF anh. se obtiene directamente por destilación sobre sodio metal en presencia de benzofenona y bajo atmósfera de nitrógeno.

El diclorometano (DCM, Scharlau, calidad Normasolv) libre de trazas de ácido se obtiene haciéndolo pasar por una columna de  $\text{Al}_2\text{O}_3$  básica. En el caso de necesitar utilizarlo en condiciones anhidras, se seca y se guarda sobre piedras de  $\text{CaH}_2$  bajo atmósfera de argón y protegido de la luz.

El dioxano (pro análisis) y el amoníaco (32%, extra puro) son de la casa Merck. El dioxano anh. y libre de peróxidos se obtiene por destilación sobre KOH del dioxano previamente pasado por una columna de alúmina básica y se guarda sobre piedras de  $\text{CaH}_2$ .

La piridina se ha obtenido anhidra de la calidad Karl-Fisher (Panreac) y se han añadido piedras de  $\text{CaH}_2$  para evitar la contaminación de humedad.

La N,N-dimetilformamida (DMF, Scharlau o Panreac, calidad síntesis de péptidos) se burbujea con nitrógeno antes de su utilización para eliminar contaminantes volátiles y se guarda sobre tamiz molecular activado de 4Å.

El agua utilizada es desionizada y filtrada con un sistema Milli-Q (Millipore) de forma que su resistividad sea superior a  $18\text{M}\Omega\text{cm}^{-1}$ .

La N,N'-diciclohexilcarbodiimida (DCC), la N,N'-diisopropilcarbodiimida (DIPC) y el 1-hidroxibenzotriazol (HOBT) son Fluka Chemie p.s. La piperidina utilizada es de la casa Aldrich y la N,N-diisopropiletilamina (DIEA) es de Merck.

Todos los disolventes utilizados en cromatografía líquida analítica son de calidad HPLC de la casa Panreac, Scharlau o Merck. Otros reactivos empleados, como el fluoruro de tetrabutilamonio (TBAF) y las matrices empleadas en la técnica de espectrometría de masas MALDI-TOF se han adquirido en Fluka o bien Aldrich.

### 13.- TAMPONES.

- *Tampón acetato de trietilamonio (TEAAc) 2M, pH=7.*

Preparación de 500mL de tampón TEAAc 2M (pH=7): se adicionan 140mL (1mol) de trietilamina (TEA) y 58mL de ácido acético glacial (AcOH) en un erlenmeyer con 250mL de agua y se mantiene en agitación vigorosa. Una vez atemperado, se ajusta el pH añadiendo unas gotas de TEA o de AcOH, según convenga. Finalmente se enrasa a un volumen de 500mL con  $\text{H}_2\text{O}$  MilliQ en un matraz aforado.

Para utilizar este tampón como eluyente de HPLC, se debe filtrar previamente a través de filtros de  $45\mu\text{m}$  de tamaño de poro (Gelman Sciences Inc, Michigan).

- *Tampón bicarbonato de trietilamonio (TEAB) 2M, pH=8.*

Preparación de 1000mL de TEAB 2M (pH=8): en un erlenmeyer de 2L se prepara una disolución acuosa de TEA 2M (277.6mL TEA) sobre la que se hace burbujear  $\text{CO}_2$  durante el tiempo necesario para que la disolución tenga el pH adecuado (aproximadamente 8-10h).

- *Tampón acetato de amonio ( $\text{AcONH}_4$ ) 2M, pH=5.5.*

Preparación de 250mL de  $\text{AcONH}_4$  2M (pH=5.5): se prepara una disolución acuosa de  $\text{AcONH}_4$  2M y se agita vigorosamente. El pH se ajusta añadiendo unas gotas de AcOH glacial.

- *Tampón Tris-HCl 500mM, pH=7.*

Preparación de 100mL de Tris-HCl 0.5M (pH=7): se prepara una disolución acuosa de Tris-HCl (6.05g de Tris-HCl en 80mL de agua) y se agita vigorosamente. Se ajusta el pH añadiendo unas gotas de HCl y finalmente se enrasa la disolución a un volumen de 100mL con  $\text{H}_2\text{O}$  MilliQ en un matraz aforado.

## 1.4.- CROMATOGRAFÍA.

### 1.4.1.- CROMATOGRAFÍA EN CAPA FINA.

La cromatografía en capa fina (TLC) se ha realizado sobre cromatofolios de gel de sílice con indicador de fluorescencia sobre soporte de aluminio (60F, 0.2mm, Merck) utilizando las mezclas de elución que se indican a continuación: sistema A: hexano/AcOEt 70:30, sistema B: DCM/AcOEt/TEA 45:45:10 y sistema C: hexano/DCM 30:70.

Para el análisis de nucleósidos, el revelado de las capas finas se ha llevado a cabo utilizando un reactivo general de azúcares. Este reactivo consiste en una mezcla de 180mL de EtOH, 1g de difenilamina, 100mL de AcOH glacial y 3mL de H<sub>2</sub>SO<sub>4</sub> conc.. Se rocía la placa con el reactivo y se calienta a 110°C durante 3min. Transcurrido este tiempo los nucleósidos muestran una coloración azulada. En el caso de que los nucleósidos contengan el grupo protector DMT, se observa primero una coloración naranja debida a la liberación del catión DMT<sup>+</sup> en medio ácido.

### 1.4.2.- CROMATOGRAFÍA EN COLUMNA.

La cromatografía en columna se ha realizado sobre gel de sílice (Chromatogel 60Å CC, 230-400 mesh, SDS)

### 1.4.3.- CROMATOGRAFÍA LÍQUIDA.

#### 1.4.3.1.- Cromatografía líquida analítica y semipreparativa.

La cromatografía líquida de alta eficacia (HPLC) a escala analítica y semipreparativa se ha realizado en cromatógrafos Shimadzu con las siguientes características:

- dos bombas LC-10AS con mezclador de alta presión, autoinyector SIL-9A, detector UV-Vis de longitud de onda variable SPD-10A y registrador integrador Chromatopac C-R5A,
- dos bombas LC-10AD con mezclador de alta presión, autoinyector SIL-10Axl, detector UV-Vis de longitud de onda variable SPD-M10AVP6A (Diode-Array), impresora Hewlett-Packard, módulo de comunicación CBM-10A y ordenador COMPAQ DESKPRO.

Condiciones analíticas:

Para llevar a cabo la cromatografía en fase reversa en condiciones analíticas, se han utilizado columnas tipo Nucleosil C<sub>18</sub> (Sharlau, 250x4 mm, 10 μm, flujo 1 mL/min). Los productos de las digestiones enzimáticas de oligonucleótidos se han analizado con una columna Spherisorb ODS 2 (250x4.6mm, 10μm, flujo 1mL/min).

Condiciones semipreparativas:

Se ha utilizado una columna Hamilton PRP-1 de poli(estireno-divinilbenceno) (305x7 mm, 10 μm, flujo 2 mL/min).

En todos los casos se ha utilizado acetato de trietilamonio (TEAAc) 0.01M, como eluyente A, y ACN/H<sub>2</sub>O 1:1 v/v como eluyente B.

### 1.4.3.2.- Cromatografía líquida preparativa

El desalado de los oligonucleótidos se ha realizado mediante *filtración molecular* a través de Sephadex G-10 (Pharmacia, 80x2cm), en sistemas compuestos por una bomba peristáltica Pharmacia P-1 (LKB) o 2132 Microperpex (LKB), detector de UV-Vis de longitud de onda fija Uvicord 2158 SD (LKB), colector de fracciones automático Ultrorac (LKB) y registrador Xpectra, Omniscrite (Houston Instruments), Sevoscrite 1S (Philips) o Pharmacia-LKB REC 101. La elución de los productos se ha realizado con tampón TEAB 0.05M a un flujo de 0.8mL/min.

La purificación por *intercambio iónico* de oligonucleótidos se ha realizado sobre columnas de DEAE-Sephadex (Pharmacia, 15x2 cm), en sistemas análogos a los descritos en el caso de la filtración molecular. La elución de los productos se ha llevado a cabo a un flujo de 1.2 mL/min y utilizando gradientes de tampón acuoso de TEAB.

La purificación por *cromatografía líquida de media presión* (MPLC) se ha realizado sobre columnas de fase reversa de sílice-C<sub>18</sub> (Vydac, 10x2 cm, 40-50 µm de diámetro de partícula) utilizando como sistema una bomba de pistón Milton-Roy (flujo 1.2mL/min).

### 1.4.3.3.- Cromatografía de intercambio catiónico.

El intercambio catiónico, para obtener la sal sódica o amónica de oligonucleótidos, se ha realizado sobre una resina Dowex 50Wx4 (200-400- mesh, 4.8 meq Na<sup>+</sup>/g de resina seca, Fluka). La resina se deposita en una jeringa de polipropileno, provista de un disco filtrante de polietileno poroso, y se sigue el proceso que se describe a continuación:

- i) lavar la resina con agua,
- ii) eluir por la columna 200mL de HCl al 10% a flujo lento,
- iii) lavar la resina con agua hasta que el pH sea el del agua,
- iv) eluir por la columna 200mL de NaOH 1M (o bien NH<sub>4</sub>OH 1M) a flujo lento,
- v) repetir el paso 3.

El producto se disuelve en la mínima cantidad de agua (1-2mL) y se pasa por la resina, colectándose fracciones de 0.5mL. Se analizan las diferentes fracciones por UV y se liofilizan las correspondientes a producto.

Tras su utilización, la resina debe regenerarse realizando los pasos i)- iii).

## 1.5.- ENSAYO CUALITATIVO DE NINHIDRINA.

El ensayo de ninhidrina<sup>1</sup> permite detectar la presencia de aminas primarias libres sobre soportes poliméricos. Se trata una alícuota de resina lavada y seca (0.5-2mg) con 3 gotas de reactivo A y una gota de reactivo B y se calienta a 110° durante 3min. Una coloración azul indica la presencia de aminas primarias libres, mientras que una coloración amarilla indicaría un ensayo negativo. Es conveniente realizar simultáneamente un ensayo con un blanco sin resina.

Reactivo A: Para la preparación de este reactivo se emplean dos disoluciones: una

<sup>1</sup> Kaiser, E.; Bosinger, C.D.; Cook, P.I. *Anal. Biochem.*, 1970, 34, 595-598.

disolución de 40g de fenol en 10mL de EtOH absoluto y una disolución de 2mL de KCN 10mM en 100mL de piridina anh. recién destilada. Ambas disoluciones se agitan por separado con 4g de resina Amberlite MB-3 (Merck) durante 45min, se filtran y se mezclan.

Reactivo B: Se disuelven 2.5g de ninhidrina en 50mL de EtOH absoluto. La disolución obtenida se conserva en un recipiente cerrado protegida de la luz.

## 1.6.- SÍNTESIS DE OLIGONUCLEÓTIDOS EN FASE SÓLIDA.

### 1.6.1.- INSTRUMENTACIÓN, REACTIVOS Y DISOLVENTES ESPECÍFICOS PARA LA SÍNTESIS DE OLIGONUCLEÓTIDOS.

La síntesis en fase sólida de oligonucleótidos se ha realizado automáticamente en un sintetizador Applied Biosystems 308B. Como soporte polimérico se ha utilizado un copolímero de polietilenglicol-poliestireno (TentaGel NNH<sub>2</sub>, RAPP POLYMERE) con una funcionalización en grupos amino de 0.24mmol/g. Ocasionalmente se ha utilizado un soporte de 4-(metilbenzhidrilamina)poliestireno (MBHA-PS, Novabiochem) con una funcionalización en grupos amino de 0.56mmol/g.

Los derivados fosforamidito y algunos de los reactivos habituales (disolución activadora: tetrazol 0.5M en ACN, reactivos de acetilación: Ac<sub>2</sub>O/lutidina/THF 1:1:8 + 1-metilimidazol 6.5% en THF, disolución oxidante: I<sub>2</sub> 0.1M en THF/Pyr/H<sub>2</sub>O y reactivo de destrilación: TCA 3% en DCM) se han adquirido a Applied Biosystems, Cruachem o Glenn Research. El resto de disolventes utilizados se preparan en las condiciones que se describen en Reactivos y Disolventes, y se filtran antes de su utilización (filtros Pro-X<sup>TM</sup>, nylon, 0.45mm, 25mm, Teknokroma). Los disolventes y reactivos directamente relacionados con la reacción de acoplamiento deben ser completamente anhidros y manipularse bajo atmósfera de argón. Todo el material de vidrio o metálico a utilizar debe manejarse también en condiciones anhidras (mantener en estufa a 110°C como mínimo dos horas y dejar atemperar en el interior de un desecador al vacío; abrir con atmósfera inerte de argón).

Como grupo espaciador entre el soporte polimérico y la cadena oligonucleotídica se ha utilizado el ácido N-Fmoc-6-aminohexanoico (Bachem). El reactivo de ciclación 1-mesitilensulfonil-3-nitro-1,2,4-triazole (MSNT) se ha sintetizado tal como se describe en el apartado 2.2.3.1 o bien se ha adquirido a Peninsula Laboratories. El N-metilimidazol, utilizado como activador durante la etapa de ciclación, es de Fluka (calidad tratamiento de genes). Para el desanclaje de los oligonucleótidos cíclicos, se ha utilizado el 2-piridinaldoximato de tetrametilguanidinio (2-piridincarbaldéhído oxima, Merck; 1,1,3,3-tetrametilguanidina (TMG), Fluka).

### 1.6.2.- DETERMINACIÓN DE LA FUNCIONALIZACIÓN DE RESINAS

- *Método I: Funcionalización de Fmoc-resinas.*

Se trata una alícuota de resina seca (~15mg) con piperidina 20% en DMF, en las condiciones indicadas en la Tabla M.1. Se recogen los filtrados de las etapas 3-4 y los lavados posteriores con DCM en un matraz aforado de volumen adecuado, se enrasa con DCM y se mide la absorbancia a  $\lambda=300\text{nm}$ . debida a N-(9-fluorenilmetil)piperidina (Fmp,  $\epsilon_{300}=7800$ )

formada durante la desprotección del grupo amino del espaciador. Aplicando la fórmula que se indica a continuación se determina la funcionalización de la resina:

$$f(\mu\text{mol/g}) = \frac{A_{300} \cdot V \cdot 10^6}{\epsilon_{300} \cdot l \cdot m}$$

donde A: absorbancia, V: volumen (mL),  $\epsilon$ : coeficiente de extinción molar ( $\text{Lmol}^{-1}\text{cm}^{-1}$ ), l: longitud de la cubeta (cm) y m: mg de resina.

**Tabla M.1.- Tratamiento para la determinación de la funcionalización de Fmoc-resinas.**

Etapa	Disolvente/Reactivo	NºTratamientos	Tiempo (min)
1.- Lavado	DCM	4	1
2.- Lavado	DMF	4	1
3.- Desprotección	Pip 20% en DMF (200 $\mu$ L)	3	3
4.- Lavado	DMF (100 $\mu$ L)	1	1

- **Método II: Funcionalización de nucleotidil-resinas.**

Se trata una alícuota de resina seca (~15mg) con TCA 3% en DCM (3x300 $\mu$ l). Los filtrados se evaporan a sequedad y se enrasa en un matraz aforado de volumen adecuado con  $\text{HClO}_4/\text{EtOH}$  3:2. Se mide la absorbancia a  $\lambda=498\text{nm}$ , debida al catión 4,4'-dimetoxitritilo liberado durante la desprotección ( $\text{DMT}^+$ ,  $\epsilon=71700$ ). Aplicando la fórmula que aparece en el apartado anterior se determina la funcionalización en grupos DMT.

### 1.6.3.- CUANTIFICACIÓN DE OLIGONUCLEÓTIDOS POR ABSORCIÓN UV.

La cuantificación de oligonucleótidos se realiza midiendo la absorbancia en UV alrededor del máximo que presentan a 260nm y se indica en unidades de absorbancia  $\text{OD}_{260}^{2,3}$ . Se puede determinar la equivalencia molar de estas unidades teniendo en cuenta la aproximación de que la absorción del oligonucleótido se debe al efecto acumulado de la absorción individual de cada base, corregido por un factor ligado al fenómeno de apilamiento. De esta manera, es posible conocer el coeficiente de extinción molar del oligonucleótido  $\epsilon_{\text{oligo}}$ :

$$\epsilon_{\text{oligo}} = \left( \sum \epsilon_b \right) \cdot f_a$$

donde  $\epsilon_b$  son los coeficientes individuales de cada base y  $f_a$  el factor de corrección (0.9 para oligonucleótidos de cadena única, 0.8 para oligonucleótidos autocomplementarios o de doble cadena). A continuación se muestran los valores de  $\epsilon_{260}$  en  $\text{H}_2\text{O}$  correspondientes a cada una de las bases:

<sup>2</sup> «Oligonucleotides and Analogues: a Practical Approach», Ed. Eckstein, F.; IRL Press, Oxford, 1991, 20.

<sup>3</sup> «Oligonucleotide Synthesis: a Practical Approach», Ed. Gait, M.J.; IRL Press, Oxford, 1984, 126.

Tabla M.2.- Valores de  $\epsilon_{260}$  en  $H_2O$  para los distintos nucleósidos.

Nucleósido	$\epsilon_{260}$
T	8830
dA	15200
dG	11500
dC	7700

#### 1.6.4.- CARACTERIZACIÓN DE OLIGONUCLEÓTIDOS POR DIGESTIÓN ENZIMÁTICA.

La determinación de la correcta composición en nucleósidos de los oligonucleótidos cíclicos sintetizados, así como su resistencia a exonucleasas, se ha realizado por digestión enzimática. Para ello se han utilizado los siguientes enzimas<sup>4-6</sup>:

- *Fosfodiesterasa de veneno de serpiente (SVPD, EC 3.1.4.1)*: Escinde el oligonucleótido por el extremo 3'-terminal generando nucleósidos-5'-fosfato. Se ha utilizado fosfodiesterasa de *Crotalus durissus* de la casa Sigma (SVPD-A) o de la casa Boehringer-Mannheim (1.5u/500mL, SVPD-B).

- *Fosfodiesterasa de bazo de ternera (SpPD, EC 3.1.16.1)*: Escinde el oligonucleótido por el extremo 5'-terminal generando nucleósidos 3'-fosfato. Se ha utilizado fosfodiesterasa de bazo de buey de la casa Sigma (0.0337u/mL).

- *Fosfatasa alcalina (AP, EC 3.1.3.1)*: Hidroliza los nucleósidos 3' o 5'-fosfato, obteniéndose los nucleósidos libres. Se ha utilizado fosfatasa alcalina de la casa Sigma (0.23u/mL).

Determinación de la composición de los oligonucleótidos:

Se ha utilizado una mezcla de fosfodiesterasa de veneno de serpiente y fosfatasa alcalina preparada tal como se describe a continuación: se prepara una disolución tamponada que contiene 50 $\mu$ L de  $MgCl_2$  0.1M, 250 $\mu$ L de Tris-HCl 0.1M pH=8 y 200 $\mu$ L de agua. La disolución enzimática se obtiene disolviendo 5 $\mu$ L de la disolución comercial de AP y una punta de espátula de SVPD-A en 100 $\mu$ L de la disolución anterior. Finalmente se añaden 50 $\mu$ L/OD<sub>260</sub> de disolución enzimática sobre la muestra. En el caso de utilizar SVPD-B la disolución enzimática se prepara mezclando 1 $\mu$ L de SVPD-B, 5 $\mu$ L de AP, 50 $\mu$ L de Tris-HCl 0.1M pH=8, 10 $\mu$ L de  $MgCl_2$  0.1M y 34 $\mu$ L de agua (volumen total = 100 $\mu$ L).

Resistencia de los oligonucleótidos cíclicos a exonucleasas:

Se utiliza una disolución de fosfodiesterasa de bazo de ternera, sin fosfatasa alcalina. La disolución enzimática se prepara disolviendo 30 $\mu$ L de la disolución comercial de SpPD en 100 $\mu$ L de tampón  $AcONH_4$  0.2M, preparado a partir de una disolución madre 2M y pH= 5.4.

En todos los casos, la digestión se ha realizado tratando una alícuota de oligonucleótido liofilizado (0.5-1OD<sub>260</sub>) en un eppendorf, durante toda una noche a 37°C, con 50 $\mu$ L de la disolución enzimática correspondiente.

<sup>4</sup> Adams, R.L.P.; Knowler, J.T.; Leader, D.P.P.; «*The Biochemistry of the Nucleic Acids*», 11ª-ed, Chapman & Hall, London, 1992, p. 87-108

<sup>5</sup> Dixon, M.; Webb, E.C.; «*Enzymes*», 3ª-ed, Longman, London, 1979, p. 255-316.

<sup>6</sup> «*Biochemica Information*», Boehringer Mannheim, 1ª-ed, USA, 1987.

Las disoluciones resultantes de los tratamientos enzimáticos se han analizado por HPLC en fase reversa (apartado 0.4.3.1), utilizando los gradientes indicados a continuación:

I: 5-8%B 10min, isocrático 8%B 10min, 8-80%B 5min, 80-100%B 1min.

II: isocrático 5%B 10min, 5-60%B 10min, 60-100%B 5min, isocrático 100%B 5min.

Los diferentes nucleósidos se han identificado por comparación con los tiempos de retención para una disolución patrón de nucleósidos, Tabla M.3.

**Tabla M.3.- Tiempos de retención HPLC de los nucleósidos para los gradientes I y II.**

Gradiente	tsdN min			
	dC	dG	T	dA
I	4.8	11.3	11.9	22.5
II	8.9	13.8	14.4	16.2

La proporción relativa de cada uno de los nucleósidos se ha obtenido por integración de los picos de HPLC, teniendo en cuenta los valores de los coeficientes de extinción molar a 260 nm, Tabla M.2, y se ha calculado tal como se indica a continuación:

(a) cálculo del valor medio de la relación área/coeficiente de extinción molar:

$$\frac{\sum (\text{Área} / \epsilon)_i}{N} = X$$

donde: i: nucleósido, Área: área del pico correspondiente al nucleósido i,  $\epsilon$ : coeficiente de extinción molar del nucleósido i y N: número de bases esperables.

(b) cálculo de la proporción de nucleósidos:

$$\frac{(\text{Área} / \epsilon)_i}{X} = n_i$$

donde  $n_i$  es el valor experimental calculado para el nucleósido i.

### 1.6.5.- ANÁLISIS DE OLIGONUCLEÓTIDOS POR ELECTROFORESIS EN GEL DE POLIACRILAMIDA.<sup>7-9</sup>

Para la electroforesis en fase gel se ha utilizado un aparato Pharmacia Biotech modelo SE 400. Se han realizado únicamente geles analíticos de 0.75mm de espesor para los que se han utilizado placas de vidrio de tamaño 180x245mm. La cantidad de muestra cargada oscila entre 0.1-0.2OD<sub>260</sub>, dependiendo de la pureza de la muestra.

#### 1.6.5.1.- Gel de poliacrilamida analítico en condiciones desnaturalizantes

Se han utilizado geles de acrilamida al 20% con una concentración de urea 7M.

*Preparación del gel:* se prepara una disolución de 12.6g de urea (Serva), 15mL de

<sup>7</sup> Sambrook, J.; Frisch, E.F.; Maniatis, T.; «*Molecular cloning: A laboratory manual*», 2ª-ed, Cold Spung Harbor, 1989, 2, p. 11.21-11.30. (b) ref. 3(b), p. 70-72. (c) ref. 3(a), p. 42-44.

<sup>8</sup> «Evaluating and isolating synthetic oligonucleotides», Applied Biosystems, 1987, 3.1-3.19.

<sup>9</sup> «Acrylamide gel casting handbook», Pharmacia Biotech, 1994.

acrilamida:bisacrilamida (38:2) en agua<sup>§</sup> y 1.5mL de TBE (10x, BioRad) y, si es necesario, esta disolución se enrasa hasta 30mL con agua MilliQ. Sobre la disolución anterior se añaden, por este orden, 300mL de una disolución de PAS 10% en agua (1 centésima del volumen total) y 30mL de TEMED (1 milésima del volumen total). Se agita y rápidamente se introduce la mezcla entre los dos vidrios donde debe polimerizar (~ 30min). Una vez polimerizado, se limpian los pocillos con TBE (0.5x) y se somete el gel a un *pre-running* durante 1h a un voltaje de 750V, con el fin de homogeneizar el gel. Como medio conductor se utiliza TBE 0.5x, preparado por dilución de la disolución comercial (25mL de TBE 10x en 500mL).

*Preparación de las muestras:* las muestras (0.1-0.2OD<sub>260</sub>), previamente liofilizadas, se disuelven en 20mL de formamida(Aldrich) (o bien 20mL de formamida/H<sub>2</sub>O 1:1 v/v) y se las somete a un tratamiento térmico a 95°C durante 1min. Posteriormente se mantienen en un baño de hielo hasta el momento de cargarlas en el gel.

*Indicadores:* como indicadores se han utilizado el cianoxilenol (migración equivalente a un 22-mer lineal) y el azul de bromofenol (migración equivalente a un 6-mer lineal). Normalmente se cargan 5mL de una mezcla de ambos indicadores: 0.25% azul de bromofenol, 0.25% cianoxilenol, 30% glicerol en agua; hay que diluir la mezcla 5 veces antes de su utilización.

La electroforesis se ha llevado a cabo aplicando un voltaje inicial de 750V, que se varía en función de la velocidad de migración de los indicadores. La duración de la electroforesis oscila normalmente entre 3-4h.

### 1.6.5.2.- Tinción del gel

La tinción de los geles se ha realizado con el reactivo Stains-all (Sigma). Se prepara una disolución madre al 10% en formamida, disolviendo 100mg del reactivo en 100mL de formamida y se ajusta el pH a un valor de 7.3-7.4 con HCl conc.. La disolución debe guardarse en la nevera y protegida de la luz.

*Tinción:* se mezclan 10mL de la disolución madre de Stains-all al 10% en formamida, 10mL de formamida, 50mL de isopropanol (Fluka), 1mL de Tris-HCl 3M (pH= 8.8) y 129 mL de H<sub>2</sub>O MilliQ. Rápidamente se sumerge el gel en esta mezcla y se deja reaccionar durante toda una noche (mínimo 6h) protegido de la luz. Para desteñir, se introduce el gel en agua suficiente para cubrirlo y se espera hasta que no presente coloración rosada (1-2h). Los oligonucleótidos presentan coloración azul.

## 1.7.- PREPARACIÓN DE MUESTRAS PARA ANÁLISIS POR MALDI-TOF.

Para el análisis de los oligonucleótidos por EM MALDI-TOF, se adiciona 1μL de citrato amónico 0.1M en agua sobre 1μL de muestra 25-50mM en agua. Se deja durante unos segundos y se añade 1μL de matriz. De la mezcla anterior se deposita 1μL en el portamuestra.

Matrices utilizadas: 2,4,6-trihidroxiacetofenona (2,6,6-THAP)<sup>10,11</sup> 0.5M en EtOH, 6-

<sup>§</sup> La disolución se prepara con 38g de acrilamida y 2g de bisacrilamida en 60mL de agua. Si es necesario, se enrasa hasta 100mL con agua. Una vez preparada, se guarda en la nevera en una botella opaca.

<sup>10</sup> Pieles, U.; Zürcher, W.; Moser, H.E. *Nucleic Acids Res.*, **1993**, 21, 3191-3196.

aza-2-tiotimina (ATT)<sup>12</sup> 0.5M en ACN:H<sub>2</sub>O 1:1 (v/v) o bien ácido 3-hidroxicolínico (3-HPA)<sup>13-15</sup> 0.5M en ACN:H<sub>2</sub>O 1:1 (v/v).

---

<sup>11</sup> Zhu, Y.F.; Chung, C.N.; Taranenko, N.I.; Allman, S.L.; Martin, S.A.; Haff, L.; Chen, C.H. *Rapid Commun. Mass. Spectrom.*, **1996**, 10, 383-388.

<sup>12</sup> Lecchi, P.; Le, H.M.T.; Pannell, L.K. *Nucleic Acids Res.*, **1995**, 23, 1276-1277.

<sup>13</sup> Wu, K.J.; Shaler, T.A.; Becker, C.H. *Anal. Chem.*, **1994**, 66, 1637-1645.

<sup>14</sup> Wang, B.H.; Biemann, K. *Anal. Chem.*, **1994**, 66, 1918-1924.

<sup>15</sup> Tang, K.; Fu, D.; Köter, S.; Cotter, R.J.; Cantor, C.R.; Köster, H. *Nucleic Acids Res.*, **1995**, 23, 3126-3231.

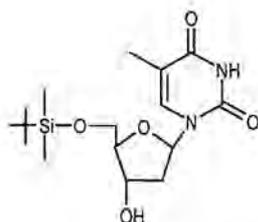
## 2.- SÍNTESIS DE OLIGONUCLEÓTIDOS CÍCLICOS.

### 2.1.- SÍNTESIS DE LOS MODELOS UTILIZADOS PARA EL ESTUDIO DE ORTOGONALIDAD DE LOS GRUPOS PROTECTORES.

#### 2.1.1.- OBTENCIÓN DE 3'-O-DMT-TIMIDINA.

##### i) Obtención de 5'-O-TBDMS-timidina.

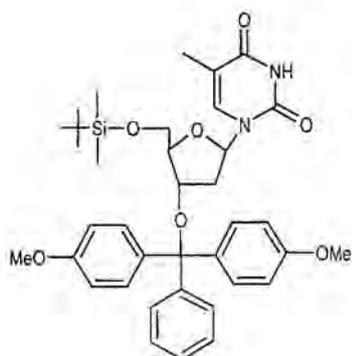
Se parte del nucleósido timidina (1.0 g, 4.1 mmol), que se seca por coevaporación con piridina anh. (x2). Se disuelve, bajo atmósfera de argón en 1.5 mL del mismo disolvente y se añade, gota a gota, una disolución de cloruro de *t*-butildimetilsililo (0.93g, 6.1 mmol), también en piridina (0.5 mL). Se deja reaccionar durante 12 h. Una vez finalizada la reacción se adiciona MeOH para destruir el exceso de cloruro de sililo y se evapora el disolvente. El residuo se purifica por columna de gel de sílice utilizando como eluyente DCM y cantidades crecientes de MeOH (0-8%). El producto se obtiene por evaporación directa de las fracciones correspondientes. Se obtiene un Rdt = 80%.



**Caracterización:** sólido blanco; pf: 166-168°C; Rf (sistema A): 0.40; <sup>1</sup>H-RMN (CDCl<sub>3</sub>, 200 MHz) δ (ppm): 8.98 (1H, m, NH-3), 7.52 (1H, s, H6), 6.40 (1H, m, H1'), 4.45 (1H, m, H3'), 4.11 (1H, m, H4'), 3.86 (2H, m, H5'/H5''), 2.40, 2.10 (2H, m, H2'/H2''), 1.92 (3H, s, CH<sub>3</sub>-5), 0.92 (9H, s, CH<sub>3</sub>-<sup>t</sup>Bu), 0.11 (6H, s, SiCH<sub>3</sub>); <sup>13</sup>C-RMN (CDCl<sub>3</sub>, 50 MHz) δ (ppm): 163.8 (C4), 150.7 (C2), 135.4 (Carom, C6), 110.9 (C5), 87.3, 85.0 (C4', C1'), 72.6 (C3'), 63.6 (C5'), 41.1 (C2'), 25.9 (CH<sub>3</sub>-<sup>t</sup>Bu), 16.3 (Cq Si), 12.5 (CH<sub>3</sub>-5), -5.5, -5.9 (CH<sub>3</sub>-Si).

##### ii) Obtención de 5'-O-TBDMS-3'-O-DMT-timidina.

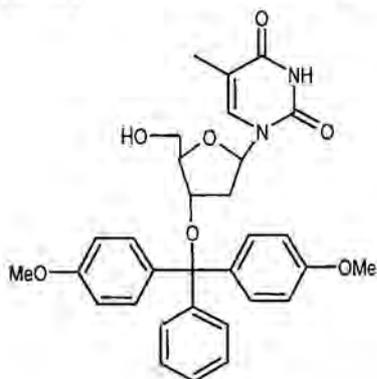
Se seca 5'-O-TBDMS-timidina (1.0 g, 2.8 mmol) por coevaporación con piridina anh. (x2). Se disuelve, bajo atmósfera de argón, en este mismo disolvente (1.5 mL) en un matraz protegido de la luz. Se añade cloruro de 4, 4'-dimetoxitritilo sólido (1.42 g, 4.2 mmol) y se deja reaccionar durante 12 h. Una vez finalizada la reacción se añade MeOH para destruir el exceso de cloruro de tritilo y se evapora a sequedad. El residuo se purifica por columna de gel de sílice utilizando como eluyente DCM/TEA 98:2 y cantidades crecientes de MeOH (0-3%). El producto se obtiene por evaporación directa de las fracciones correspondientes y posterior precipitación sobre hexano. Se obtiene un Rdt = 68%.



**Caracterización:** Sólido amarillento; pf: 96-98°C; Rf (sistema A): 0.70; <sup>1</sup>H-RMN (CDCl<sub>3</sub>, 200 MHz) δ (ppm): 8.80 (1H, m, NH-3), 7.60-7.20 (10H, m, ArDMT, H6), 6.92 (4H, d, J= 8.4 Hz, ArDMT), 6.50 (1H, m, H1'), 4.38 (1H, m, H3'), 4.16 (1H, m, H4'), 3.80, 3.20 (2H, m, H5'/H5''), 3.89 (6H, s, OCH<sub>3</sub>), 2.10, 1.80 (2H, m, H2'/H2''), 1.95 (3H, s, CH<sub>3</sub>-5), 0.90 (9H, s, <sup>t</sup>Bu), 0.10, 0.05 (6H, s, SiCH<sub>3</sub>); <sup>13</sup>C-RMN (CDCl<sub>3</sub>, 50 MHz) δ (ppm): 163.7 (4), 158.6 (Carom), 150.3 (C2), 145.0 (Carom), 136.4, 135.4 (Carom, C6), 130.2, 130.1, 128.2, 127.9, 127.0, 113.2 (Carom-H), 110.8 (OCH<sub>3</sub>), 39.9 (C2'), 25.7 (CH<sub>3</sub>-<sup>t</sup>Bu), 19.0 (Cq Si), 12.4 (CH<sub>3</sub>-5), -5.5, -5.8 (SiCH<sub>3</sub>).

### iii) Obtención de 3'-O-DMT-timidina.

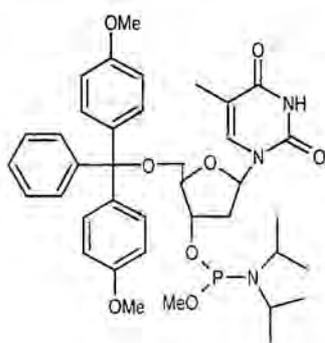
Se seca 5'-O-TBDMS-3'-O-DMT-timidina (1.25 g, 1.9 mmol) por coevaporación con ACN anh. (x2). Se disuelve, bajo atmósfera de argón, en 1.5 mL de THF anh. y se añade TBAF·3H<sub>2</sub>O sólido (1.5 eq). Se deja reaccionar durante 2 h. Al finalizar se adiciona MeOH para neutralizar el exceso de ión fluoruro y se evapora a sequedad. El residuo se purifica por columna de gel de sílice eluyendo con DCM/TEA 98:2 y cantidades crecientes de MeOH (0-5%). El producto se obtiene por evaporación de las fracciones correspondientes y posterior precipitación sobre hexano. Se obtiene un Rdt = 77%.



**Caracterización:** Sólido blanco; pf: 116-118°C; Rf (sistema A): 0.47; <sup>1</sup>H-RMN (CDCl<sub>3</sub>, 200 MHz) δ (ppm): 7.40-7.15 (9H, m, ArDMT, H6), 6.83 (4H, d, J=8.4 Hz, ArDMT), 6.14 (1H, m, H1'), 4.35 (1H, m, H3'), 3.96 (1H, m, H4'), 3.65, 3.30 (2H, m, H5'/H5''), 3.78 (6H, s, OCH<sub>3</sub>), 1.90, 1.60 (2H, m, H2'/H2''), 1.84 (3H, s, CH<sub>3</sub>-5); <sup>13</sup>C-RMN (CDCl<sub>3</sub>, 50 MHz) δ (ppm): 162.7 (C4), 157.7 (Carom), 149.4 (C2), 144.1 (Carom), 136.2, 135.3 (Carom, C6), 129.2, 127.2, 126.9, 126.1, 112.3 (Carom-H), 109.9 (C5), 86.3, 86.2 (C4', C1'), 85.6 (Cq), 73.3 (C3'), 61.5 (C5'), 54.2 (OCH<sub>3</sub>), 37.7 (C2'), 11.4 (CH<sub>3</sub>-5).

### 2.1.2.- OBTENCIÓN DE 5'-O-DMT-TIMIDINA 3'-O-METIL-N,N-DIISOPROPILFOSFORAMIDITO.

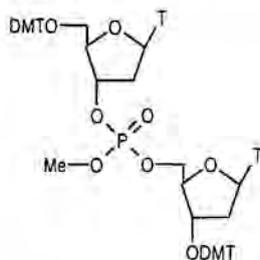
Se seca 5'-O-DMT-timidina (544.6 g, 1 mmol) por coevaporación con ACN anh. (x3) y se disuelve en 5 mL de DCM anh. y TEA anh. (4.5 eq). Se añade la cloro(diisopropil)aminometoxifosfina (529.9 mg, 613 μL, 3 mmol) y se deja reaccionar, bajo atmósfera de argón, durante 1 h a temperatura ambiente. Se observa la formación de un precipitado blanco que corresponde a la formación de cloruro de trietilamonio. Una vez transcurrida 1 h se añade MeOH para parar la reacción. La mezcla de reacción se diluye con DCM y se realizan extracciones con NaCl sat. (2x200 mL). La fase orgánica se seca con MgSO<sub>4</sub> anh. y se evapora a sequedad. El residuo se purifica por cromatografía de gel de sílice eluyendo con hexano/DCM/TEA 59:39:2 y aumentando la proporción de DCM (39-98%). El producto se obtiene por evaporación directa de las fracciones correspondientes y posterior precipitación sobre hexano. Se obtiene un Rdt = 80%.



**Caracterización:** Sólido amarillento; Rf (sistema G): 0.51; <sup>31</sup>P-RMN (CDCl<sub>3</sub>, 120 MHz): δ (ppm): 146.1, 145.6 (2 diastereoisómeros).

### 2.1.3.- OBTENCIÓN DEL FOSFATO DE 3'-O-(5'-O-DMT-TIMIDINIL) 5'-O-(3'-O-DMT-TIMIDINIL) Y METILO.

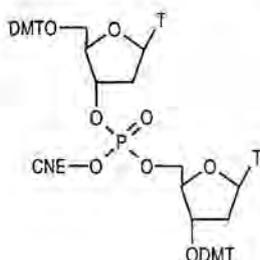
Se seca 3'-O-DMT-timidina (272.3 mg, 0.5 mmol) por coevaporación con ACN anh. (x2). Se añade tetrazol (34.5 mg, 0.5 mmol) y la mezcla se disuelve, bajo atmósfera de argón, en 2 mL de ACN anh.. Se adiciona, con una jeringuilla y en condiciones anhidras, una disolución de 5'-O-DMT-timidina 3'-O-metil-N,N'-diisopropilfosforamidito (317.6 mg, 0.45 mmol), previamente secado por coevaporación con ACN anh. (x2), en 2 mL de este disolvente. Se deja reaccionar a temperatura ambiente durante 1 h. Una vez acabada la reacción se añade una disolución de <sup>t</sup>BuOOH 3M en tolueno (1.5 mL, 5mmol) y se deja reaccionar durante 10 min. La mezcla de reacción se diluye con DCM y se extrae con TEAB 2 M (2x100 mL) y H<sub>2</sub>O (2x 200 mL). La fase orgánica se seca con MgSO<sub>4</sub> y se evapora a sequedad. El residuo se purifica por cromatografía en columna de gel de sílice eluyendo con hexano/acetona/TEA 49:49:2 y cantidades crecientes de MeOH (0-2%). El producto se obtiene por evaporación directa de las fracciones correspondientes y posterior precipitación sobre hexano. Se obtiene un Rdt = 65%.



**Caracterización:** Sólido blanco; R<sub>f</sub> (sistema F): 0.14; <sup>31</sup>P-RMN (CDCl<sub>3</sub>, 120 MHz) δ (ppm): -3.7, -3.9 (2 diastereoisómeros).

### 2.1.4.- OBTENCIÓN DEL FOSFATO DE 3'-O-(5'-O-DMT-TIMIDINIL) 5'-O-(3'-O-DMT-TIMIDINIL) Y 2-CIANOETILO.

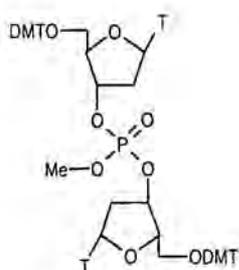
Se prepara de la forma descrita en el apartado anterior, utilizando 5'-O-DMT-timidina 3'-O-(2-cianoetilo)-N,N'-diisopropilfosforamidito. En este caso el producto se purifica por cromatografía de gel de sílice utilizando como eluyente hexano/acetona/piridina 49:49:2, ya que la presencia de TEA produce la eliminación del grupo 2-cianoetilo, y cantidades crecientes de MeOH (0-5%). El producto se obtiene por evaporación directa de las fracciones correspondientes y posterior precipitación sobre hexano. Se obtiene un Rdt = 30%.



**Caracterización:** Sólido amarillento; R<sub>f</sub> (sistema F): 0.11; <sup>31</sup>P-RMN (CDCl<sub>3</sub>, 120 MHz) δ (ppm): -1.3, -1.4 (2 diastereoisómeros).

### 2.1.5.- OBTENCIÓN DEL FOSFATO DE BIS-3'-O-(5'-O-DMT-TIMIDINIL) Y METILO.

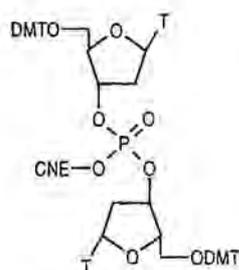
Se prepara de la forma indicada en el apartado 2.1.3.-, utilizando 5'-O-DMT-timidina en lugar de 3'-O-DMT-timidina. Se obtiene un Rdt = 85%.



**Caracterización:** Sólido blanco; Rf (sistema A): 0.54;  $^{31}\text{P}$ -RMN ( $\text{DMF}_d/\text{NMM}$  1:1, 120 MHz)  $\delta$  (ppm): -0.9.

### 2.1.6.- OBTENCIÓN DEL FOSFATO DE BIS-3'-O-(5'-O-DMT-TIMIDÍNIL) Y 2-CIANOETILO.

Se prepara de la forma indicada en el apartado 2.1.4.-, utilizando 5'-O-DMT-timidina en lugar de 3'-O-DMT-timidina. Se obtiene un Rdt = 70%.



**Caracterización:** Sólido blanco; Rf (sistema A): 0.52;  $^{31}\text{P}$ -RMN ( $\text{DMF}_d/\text{NMM}$  1:1, 120 MHz)  $\delta$  (ppm): -2.5.

## 2.2.- SÍNTESIS EN FASE SÓLIDA DE OLIGONUCLEÓTIDOS CÍCLICOS.

### 2.2.1.- DERIVATIZACIÓN DE SOPORTES POLIMÉRICOS.

#### 2.2.1.1.- Síntesis de espaciador-resina

Se coloca la cantidad de resina (TentaGel  $f \approx 0.24 \text{ mmol/g}$ ) que se quiera funcionalizar en una jeringa de polipropileno, provista de un disco filtrante de polietileno poroso, y se somete a una serie de lavados con los disolventes y reactivos que posteriormente se utilizarán en las diferentes etapas de derivatización del soporte, Tabla M.4.

Se incorpora el ácido N-Fmoc-6-aminohexanoico (1 eq respecto a la funcionalización inicial de la resina), utilizando indistintamente como agente acoplante DCC o DIPC (1eq). Tras el acoplamiento se comprueba la funcionalización obtenida tal como se describe en el apartado 1.6.2 de Materiales y Métodos. Si la funcionalización es correcta (0.1-0.15 mmol/g), se procede a la acetilación de los grupos amino que no han reaccionado, mediante un tratamiento con  $\text{Ac}_2\text{O}/\text{DIEA}$  1:1 (20 eq respecto de las aminas libres sobre resina), y se deja reaccionar durante 20 min (2x10). Se comprueba la total acetilación de las aminas libres con el ensayo de ninhidrina (apartado 1.5 de Materiales y Métodos). Finalmente se desprotege el grupo amino del espaciador con piperidina 20% en DMF.

**Tabla M.4.-** Obtención de espaciador-resina en el caso de utilizar como soporte TentaGel<sup>\*</sup>.

Etapa	Disolvente/Reactivo	NºTratamientos	Tiempo (min)
1.- Lavado	DCM	2	2
2.- Lavado	20%TCA en DCM	2	2
3.- Lavado	DCM	2	2
4.- Lavado	Pip/DMF 1:1	2	2
5.- Lavado	DMF	2	2
6.- Lavado	MeOH	2	2
7.- Acoplamiento	ácido N-Fmoc-6-aminohexanoico (1eq) DCC(o DIPC) (1eq) en DCM/DMF	1	120-300
8.- Lavado	DMF	3	2
9.- Lavado	DCM	3	2
10.-Acetilación <sup>§</sup>	Ac2O/DIEA 1:1 en DMF (20eq)	2	10+10
11.- Lavado	DMF	3	2
12.- Lavado	DCM	3	2
13.- Desprotección	Pip 20% en DMF	3	3+3+5
14.- Lavado	DMF	3	2
15.- Lavado	DCM	3	2
16.- Lavado	MeOH	3	2

<sup>\*</sup> En el caso de utilizar el soporte *p*-MBHA-PS, en las etapas 2 y 4 se utiliza TFA 5% en DCM y DIEA 5% en DCM respectivamente. Para el acoplamiento del espaciador se utilizan únicamente 0.5 eq de ácido N-Fmoc-6-aminohexanoico y 0.5eq de agente acoplante, ya que este soporte presenta una funcionalización en grupos amino muy superior: 0.56 mmol/g.

<sup>§</sup> Previa determinación de la funcionalización.

### 2.2.1.2.- Síntesis de nucleotidil-resina.

El primer nucleótido se incorpora al soporte polimérico, derivatizado con el espaciador, en forma de nucleotidil-linker (1.2-2eq), cuya síntesis se describe en el apartado 2.2.1.2.1. El nucleotidil-linker se seca por coevaporación con ACN anh. (x2) y se mantiene en un desecador sobre P<sub>2</sub>O<sub>5</sub> y KOH durante una noche. Se disuelve en la mínima cantidad de DCM, se adiciona sobre la resina previamente inflada con DCM, y se deja reaccionar durante 10-12h con agitación mecánica. Como agentes acoplantes se utiliza DCC y HOBt (1.2-2eq), éste último disuelto en ~100 µL de DMF.

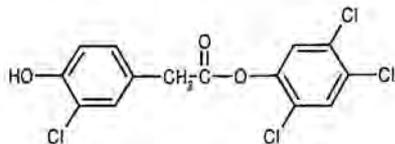
Como en el apartado anterior, tras el acoplamiento se comprueba la funcionalización de la resina tal como se describe en el apartado 1.6.2.- de Procedimientos Generales. Finalmente se elimina el grupo protector del fosfato que, tratándose del grupo cianoetilo, se realiza con tres tratamientos de 1h con TEA/Pyr 1:1. La total eliminación del grupo cianoetilo se comprueba mediante <sup>31</sup>P-RMN en fase gel.

#### 2.2.1.2.1.- Síntesis de nucleotidil-linker.

##### i) Obtención de 3-cloro-4-hidroxifenilacetato de 2,4,5-triclorofenilo.

Se disuelven 4.7 g de 2,4,5-triclorofenol (23.6 mmol) y 4.9 g de DCC (23.6 mmol) en 22 mL de DCM y se adiciona lentamente una disolución de 4 g de ácido 3-cloro-4-hidroxifenilacético (21.6 mmol) en 130 mL de AcOEt. La mezcla de reacción se deja en agitación magnética durante 12 h y se separa por filtración la N,N-diciclohexilurea formada durante la reacción. Se enfría la mezcla de reacción en un baño a 0°C con el objeto de precipitar la N,N-diciclohexilurea que pueda quedar y se vuelve a filtrar. Se lava el filtrado con una disolución acuosa de NaHCO<sub>3</sub> al 5% (x3), se seca la fase orgánica con MgSO<sub>4</sub> anh. y se evapora

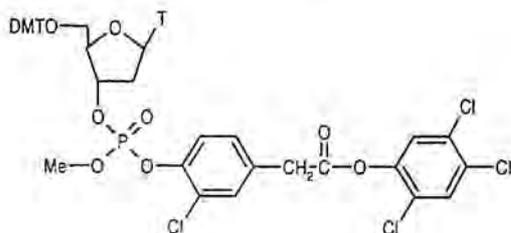
a sequedad. Se obtiene un aceite de color anaranjado que se purifica por cromatografía en columna de gel de sílice utilizando como eluyente hexano/DCM 80:20 y aumentando la proporción de DCM (20-80%). El producto se obtiene por evaporación directa de las fracciones correspondientes y posterior precipitación sobre hexano. Se obtiene un Rdt = 34%.



**Caracterización:** Sólido blanco; pf: 77-78°C; Rf (sistema D): 0.33;  $^1\text{H-RMN}$  ( $\text{CDCl}_3$ , 200 MHz)  $\delta$  (ppm): 7.53 (1H, s, Ar-triclorofenilo), 7.34 (1H, d,  $J= 2.19$  Hz, Ar-clorofenilo), 7.25 (1H, s, Ar-triclorofenilo), 7.18 (1H, dd,  $J'= 8.39$  Hz,  $J= 2.19$  Hz, Ar-clorofenilo), 7.01 (1H, d,  $J'= 8.39$  Hz, Ar-clorofenilo), 5.60 (1H, s, OH), 3.82 (2H, s,  $\text{CH}_2$ );  $^{13}\text{C-RMN}$  ( $\text{CDCl}_3$ , 50 MHz)  $\delta$  (ppm): 167.3 (CO), 149.9 (Carom triclorofenilo), 144.6 (Carom clorofenilo), 130.0 (CH, triclorofenilo), 129.7 (Carom triclorofenilo), 128.9 (CH triclorofenilo), 128.5 (CH clorofenilo), 124.5 (Carom clorofenilo), 124.2 (CH clorofenilo), 118.9 (Carom clorofenilo), 115.5 (CH clorofenilo); IE:  $M^+$  = 366,  $m/z$  = 198, 141, 97, 77.

ii) *Obtención del fosfato de 3'-O-(5'-O-DMT-timidinil) 2-cloro-4-(2,4,5-triclorofenoxycarbonilmetil)fenilo y metilo.*

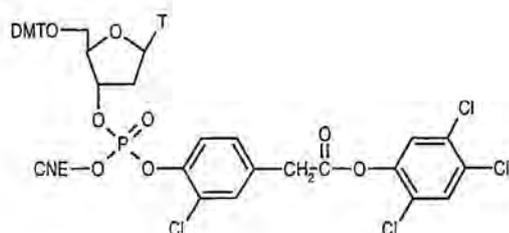
En un matraz bien seco (previamente secado en la estufa a 120°C y atemperado en un desecador) se introducen 5'-O-DMT-timidina 3'-O-metil-N,N-diisopropilfosforamidito (141.1 mg, 0.2 mmol) y 3-cloro-4-hidroxifenilacetato de 2,4,5-triclorofenilo (65.9 mg, 0.18 mmol), previamente secados en un desecador, y se disuelven, bajo atmósfera de argón, en 0.8 mL de DCM anh. Se adiciona tetrazol (13.8 mg, 0.2 mmol), disuelto en 0.5 mL de ACN anh., con ayuda de una cánula y con presión de argón. Se deja reaccionar en atmósfera inerte durante 1 h. A continuación se añade una disolución 3M de  $^t\text{BuOOH}$  en tolueno (0.5 mL, 2mmol) y se deja reaccionar durante 10 min. Posteriormente se diluye la mezcla de reacción con DCM y se lava con  $\text{H}_2\text{O}$  (2x50 mL) y una disolución acuosa de NaCl sat. (2x50 mL). La fase orgánica se seca con  $\text{Na}_2\text{SO}_4$  anh. y se evapora a sequedad. El producto se precipita directamente sobre hexano y se utiliza sin purificación previa. Generalmente se observa una impureza que corresponde al fosforamidito, inicialmente en exceso, hidrolizado. Este hidrógenofosfonato no interfiere en la posterior reacción de anclaje a la resina. Se obtiene un Rdt = 95%.



**Caracterización:** Sólido blanco; Rf (sistema A): 0.51;  $^{31}\text{P-RMN}$  ( $\text{CDCl}_3$ , 120 MHz)  $\delta$  (ppm): -10.5 (2 diastereoisómeros), 5.6, 5.2 (impureza: hidrógenofosfonato)

iii) *Obtención del fosfato de 3'-O-(5'-O-DMT-timidinil) 2-cloro-4-(2,4,5-triclorofenoxycarbonilmetil)fenilo y 2-cianoetilo.*

Se obtiene de la forma descrita en el apartado anterior, utilizando 5'-O-DMT-timidina 3'-O-(2-cianoetilo)-N,N-diisopropilfosforamidito. Se obtiene un Rdt = 91%.



**Caracterización:** Sólido blanco; Rf (sistema B): 0.33;  $^{31}\text{P}$ -RMN ( $\text{CDCl}_3$ , 120 MHz)  $\delta$  (ppm): -7.7, -7.8 (2 diastereoisómeros), 8.4, 8.1 (impureza: hidrógenofosfonato).

## 2.2.2.- ELONGACIÓN DE LA CADENA OLIGONUCLEOTÍDICA.

La elongación de la cadena oligonucleotídica, a partir de la nucleotidil-resina en la que se ha eliminado el protector del grupo fosfato 3'-terminal, se realiza de forma automática utilizando la metodología del fosfito triéster, en la que se utilizan derivados nucleosídicos del tipo fosforamidito. Las síntesis se realizan en columnas de teflón del tamaño adecuado a la escala de trabajo. Se han utilizado diferentes ciclos de síntesis en función tanto de la escala como del soporte polimérico elegido.

El rendimiento individual de cada acoplamiento, así como el rendimiento medio de síntesis, se determina a partir de la cuantificación en UV-Vis de los filtrados procedentes de cada una de las etapas de destrilación. La cuantificación se efectúa de forma análoga a la descrita para la determinación de la funcionalización de nucleotidil-resinas (apartado 0.6.2.3). Opcionalmente, se puede realizar un análisis cualitativo enrasando los filtrados procedentes de cada desprotección a un mismo volumen con una disolución de ácido *p*-toluensulfónico 0.1M en ACN y midiendo la absorbancia debida al catión  $\text{DMT}^+$  a una longitud de onda cercana a 498nm.

En la Tabla M.5 se resume el protocolo de síntesis a gran escala para los dos tipos de soportes utilizados a lo largo de este trabajo. Los ciclos completos se incluyen al final de esta sección.

**Tabla M.5.- Protocolo estándar para síntesis de oligonucleótidos a gran escala**

Etapas	TentaGel		MBHA-PS	
	Disolvente/Reactivo	Tiempo (s)	Disolvente/Reactivo	Tiempo (s)
1.- Lavados	ACN	65	DCM	100
2.- Destrilación	TCA 3% en DCM	2x120+1x90	TCA 3% en DCM	1x120+2x100
3.- Lavados	ACN	120	DCM	240
	DMF	210	DMF	120
	ACN	270	DCM	270
	ACN anh	300	ACN anh.	150
4.- Secado	Argón	90	Argón	100
5.- Acoplamiento	FA 0.1M en DCM tetrazol 0.5M en ACN	60 +1200 <sup>§</sup>	FA 0.2M en DCM tetrazol 0.5M en THF	80+1800 <sup>§</sup>
6.- Lavados	ACN	60	THF	120
7.- Acetilación	Ac <sub>2</sub> O/NMI <sup>a</sup>	100+180 <sup>§</sup>	Ac <sub>2</sub> O/NMI <sup>a</sup>	100+180 <sup>§</sup>
8.- Oxidación	I <sub>2</sub> aq. <sup>b</sup>	120+150 <sup>§</sup>	I <sub>2</sub> aq. <sup>b</sup>	150+120 <sup>§</sup>
9.- Lavados	DCM	90	THF	120
	DMF	120	DMF	180
	ACN	180	DCM	220

<sup>a</sup> Ac<sub>2</sub>O/lutidina/THF 1:18+1-metilimidazol 6.5% en THF

<sup>b</sup> I<sub>2</sub> 0.1M en THF/Pyr/H<sub>2</sub>O 2:20:80

<sup>§</sup> corresponden a tiempos de espera una vez la columna se ha llenado de reactivo

### 2.2.3.- OBTENCIÓN DE OLIGONUCLEÓTIDOS CÍCLICOS: REACCIÓN DE CICLACIÓN.

La reacción de ciclación se realiza sobre el precursor lineal con los extremos 3' y 5'-terminales desprotegidos. Como reactivo de ciclación se utiliza una disolución de MSNT 0.1M en piridina anh. (10eq respecto escala de trabajo) y se realizan tres tratamientos: 2x4h + 12h, utilizando para cada tratamiento una disolución de MSNT acabada de preparar. La reacción de ciclación debe realizarse en condiciones anhidras y bajo atmósfera inerte de argón, para ello se ha diseñado un protocolo que permite realizar esta reacción en el sintetizador automático:

*Tabla M.6.- Protocolo de ciclación de oligonucleótidos.*

Etapa	Disolvente/Reactivo	Tiempo (s)
1.- Secado	Argón	90
2.- Lavados	Pyr anh.	180
	ACN anh	240
3.- Secado	Argón	90
4.- Ciclación	MSNT 0.1M en Pyr anh.	2x4h + 1x12h

Los ensayos de ciclación utilizando distintos reactivos de ciclación se han llevado a cabo de forma análoga, en las condiciones que se detallan en la Tabla M.7.

Es muy importante que antes de llevar a cabo la reacción de ciclación se compruebe el buen estado del reactivo determinando el punto de fusión o bien realizando espectros de RMN. En el caso de BOPDC es necesario recrystalizar el reactivo en ACN antes de su utilización.

*Tabla M.7.- Condiciones de reacción para los distintos reactivos de acoplamiento utilizados.*

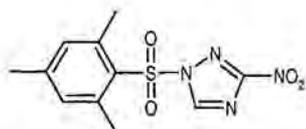
Reactivo	Condiciones de reacción
MSNT	MSNT 0.1 M en piridina anh.
	MSNT/NMI 0.1 M de cada reactivo en piridina anh.
PyAOP	PyAOP/NMI 0.1 M de cada reactivo en piridina anh.
PyBOP	PyBOP/NMI 0.1 M de cada reactivo en piridina anh.
BOPDC	BOPDC 0.1 M en piridina anh.
	NMI -10 M en piridina anh.
TBTU	TBTU 0.1 M en ACN anh.
	NMI 0.1 M en piridina anh.

La forma más apropiada de preparar las distintas disoluciones consiste en preparar una disolución 0.1 M de NMI en piridina anh. y, bajo atmósfera de argón, disolver el reactivo de ciclación con el volumen adecuado de dicha disolución, de forma que la concentración final de reactivo sea también 0.1 M.

#### 2.2.3.1.- Síntesis de 1-mesitilensulfonil-3-nitro-1,2,4-triazol (MSNT).

Se disuelven el cloruro de mesitilensulfonilo (2.5 g, 11.4 mmol), previamente purificado<sup>5</sup>, y el 3-nitro-1,2,4- triazol (1.2 g, 11.4 mmol) en 75 mL de THF anh., obteniéndose una suspensión marronosa. Se adiciona TEA (1.1 g, 11.4 mmol) y se observa la aparición de un precipitado blanco, que corresponde al hidrocioruro de trietilamonio. Se deja reaccionar con agitación magnética durante 3 h. La mezcla de reacción se filtra a través de lana de vidrio y el filtrado, de color anaranjado, se evapora a sequedad, obteniéndose un sólido de color naranja. Se disuelve este producto en DCM y se lava con H<sub>2</sub>O (3x200 mL). La fase orgánica se seca con

$\text{Na}_2\text{SO}_4$  anh. y se evapora a sequedad. El residuo se purifica por cromatografía en columna de gel de sílice eluyendo con DCM. El producto se obtiene por evaporación directa de las fracciones correspondientes y posterior precipitación sobre hexano. Se obtiene un Rdt = 53%.



**Caracterización:** Sólido blanco; pf: 137.7-138.3°C; Rf (sistema E): 0.13;  $^1\text{H}$ -RMN ( $\text{CDCl}_3$ , 200 MHz)  $\delta$  (ppm): 8.83 (1H, s, H5-triazol), 7.09 (2H, s, H3/H5 mesitileno), 2.71 (6H, s,  $\text{CH}_3$  mesitileno), 2.36 (3H, s,  $\text{CH}_3$  mesitileno);  $^{13}\text{C}$ -RMN ( $\text{CDCl}_3$ , 50 MHz)  $\delta$  (ppm): 147.6 (C4 mesitileno), 145.2 (C2/C6 mesitileno), 142.6 (2 CH mesitileno), 133.1 (2 C=N triazol), 128.0 (C-S mesitileno), 23.2 (2  $\text{CH}_3$  mesitileno), 21.4 ( $\text{CH}_3$  mesitileno); IR ( $\text{cm}^{-1}$ ): 3141.27 (st N-H), 1599.58 (st C=C), 1503.10 (st asim.  $\text{NO}_2$ ), 1383.47 (st sim.  $\text{NO}_2$ ), 1121.06 (st S=O).

#### 2.2.4.- DESPROTECCIÓN Y DESANCLAJE.

La eliminación los grupos protectores de los fosfatos internucleosídicos se realiza tal como se ha descrito anteriormente: TEA/Pyr 1:1, 3x1h, cuando el protector es el 2-cianoetilo, y con un tratamiento con tiofenol/TEA/dioxano 1:2:2, 2x10 min, cuando se trata de metilo.

Una vez se han desprotegido los grupos fosfatos, se desancla el oligonucleótido de la resina mediante una serie de tratamientos (4h + 4h + 12h) con una disolución de oximato de TMG 0.2M en dioxano/ $\text{H}_2\text{O}$  1:1, en el caso de utilizar como soporte TentaGel, y en dioxano/ $\text{H}_2\text{O}$  2:1 en el caso del MBHA-PS. Se recogen los filtrados correspondientes a los tres tratamientos y se evaporan a sequedad. El residuo obtenido se trata con una disolución acuosa de  $\text{NH}_3$  35% a 55°C durante 12 h, para eliminar los protectores de las bases (benzofilo en el caso de dA y dC, isobutirilo para dG). Finalmente se evapora a sequedad.

#### 2.2.5.- DESALADO Y PURIFICACIÓN.

El desalado de los crudos de síntesis se realiza mediante filtración molecular a través de una columna de Sephadex G-10.

Para llevar a cabo la purificación de los productos se han utilizado diferentes métodos: cromatografía de intercambio aniónico sobre resina DEAE-Sephadex, cromatografía líquida de media presión (MPLC) y cromatografía líquida de alta resolución (HPLC) en fase reversa semipreparativa.

Las condiciones utilizadas en cada caso se describen en la sección 1.4.3 de Procedimientos Generales.

#### 2.2.6.- CUANTIFICACIÓN.

La cuantificación de los oligonucleótidos cíclicos se ha realizado de la forma descrita en el apartado 1.6.3 de Procedimientos Generales. A continuación se dan los valores de  $\epsilon_{\text{oligo}}$  calculados para las distintas secuencias sintetizadas.

**Tabla M.8.-** Valores del coeficiente de extinción micromolar calculado para las distintas secuencias sintetizadas.

Secuencia	$\epsilon_{260}$ (OD/ $\mu$ mol·mL)
d<pCATT<CATT>	72.5
d<pTCATATCTATAT>	
d<p <sup>Br</sup> UCAT <sup>Br</sup> UCAT>	
d<pCGCTCGCT>	63.2
d<pCTATCTAT>	72.5
d<pTGCTCGCT>	64.5
d<pTGCATGCA>	77.7
d<pAGCTAGCT>	77.7
d<pCCGTCCGT>	63.2
d<pCGCTCATT>	67.8

### 2.2.7.- CARACTERIZACIÓN DE LOS OLIGONUCLEÓTIDOS CÍCLICOS.

La caracterización de los oligonucleótidos cíclicos se ha realizado mediante espectrometría de masas ES<sup>-</sup> o MALDI-TOF y, en algunos casos, digestiones enzimáticas, tal como se muestra en la Tabla M.9.

Para las secuencias d<pCATT<CATT>, d<pTGCTCGCT>, d<pTGCATGCA>, d<pAGCTAGCT>, d<pCCGTCCGT> y d<pCGCTCATT>, se dispone además de la caracterización por RMN. Las tablas de asignación figuran en el Apéndice B.

**Tabla M.9.-** Caracterización de los oligonucleótidos cíclicos sintetizados.

Secuencia	EM					Digestión SVPD			Resistencia a SpPD
	M calc.	M exp. ES <sup>-</sup>				M exp. MALDI-TOF	T:dC:dA:dG teór.	T:dC:dA:dG exp.	
		M-3H <sup>3</sup>	M-4H <sup>4</sup>	M-5H <sup>5</sup>	M				
d<pCATT<CATT>	2421.56	806.15	604.44	-	2421.63	2421.3	2 : 1 : 1 : 0	1.9 : 1.0 : 1.1 : 0	Sí
d<pTCATATCTATAT>	3656.40	-	913.46	730.62	3656.25	-	3 : 1 : 2	3.0 : 1.0 : 1.9 : 0	Sí
d<p <sup>Br</sup> UCAT <sup>Br</sup> UCAT>	2551.29	849.70	637.10	-	2551.06	2552.3	1 : 1 : 1 : 1 <sup>§</sup>	1.2 : 1.0 : 1.1 : 0.7	Sí
d<pCGCTCGCT>	2423.53	807.00	605.00	-	2424.10	-	-	-	Sí
d<pCTATCTAT>	2421.56	806.50	604.45	-	2421.80	-	-	-	Sí
d<pTGCTCGCT>	2438.56	811.80	608.70	-	2438.84	-	-	-	Sí
d<pTGCATGCA>	2471.58	822.50	616.60	-	2470.69	2471.39	-	-	Sí
d<pAGCTAGCT>	2471.58	822.60	616.60	-	2470.66	2471.80	-	-	Sí
d<pCCGTCCGT>	2423.53	806.40	604.60	483.50	2422.71	2423.35	-	-	Sí
d<pCGCTCATT>	2422.55	806.10	604.30	483.10	2422.11	2422.42	-	-	Sí

<sup>§</sup> El último valor corresponde a proporción de <sup>Br</sup>U.

### 3.- RMN DE OLIGONUCLEÓTIDOS CÍCLICOS.

#### 3.1.- $^1\text{H}$ -RMN.

##### 3.1.1.- PREPARACIÓN DE LAS MUESTRAS.

Los productos, en forma de sal sódica, se han disueltos en 0.5 mL de  $\text{D}_2\text{O}$  o bien  $\text{D}_2\text{O}/\text{H}_2\text{O}$  9:1, obteniéndose disoluciones de oligonucleótido de entre 2.2 y 2.4 mM.

Para realizar los espectros a elevada concentración de oligonucleótido se han utilizado tubos SHIGEMI, que permiten trabajar con volúmenes muy pequeños de muestra (~200-250  $\mu\text{L}$ ) en buenas condiciones de homogeneidad de campo magnético.

El pH de las muestras se ha ajustado a 7 con una disolución concentrada de  $\text{DCl}$ .

##### 3.1.2.- ADQUISICIÓN DE ESPECTROS.

Los espectros de RMN de  $^1\text{H}$  de campo alto se han realizado en un espectrómetro Bruker AMX-600 MHz, ubicado en el Instituto de Estructura de la Materia (CSIC, Madrid).

Como referencia interna se ha utilizado el ácido 3-trimetilsilil-(2,2,3,3- $\text{D}_4$ ) propanoico (TSP), a cuya señal se le asignó el valor de 0.00 ppm.

- Espectros en  $\text{D}_2\text{O}$ .

Los espectros bidimensionales en  $\text{D}_2\text{O}$  se han adquirido mayoritariamente con 2K en la dimensión  $t_2$  y 400 incrementos de  $t_1$ , excepto en el caso de experimentos COSY, que se han adquirido con 4K en la dimensión  $t_2$  y 512 puntos en  $t_1$ . Se ha utilizado el método de presaturación para eliminar la señal de agua residual y, en todos los casos, la anchura espectral ha sido de 13 ppm. El número de acumulaciones ha dependido en cada caso del tipo de experimento y de la concentración de oligonucleótido. Todos los espectros se han adquirido en modo sensible a la fase según el método TPPI (Time Proportional Phase Incrementation)<sup>16</sup>.

Los espectros TOCSY y ROESY se han registrado utilizando la secuencia de bloqueo de espín estándar MLEV-17<sup>17</sup>. La potencia de campo magnético para el bloqueo de espines ha sido de 10 dB para experimentos TOCSY y de entre 20-30 dB para experimentos ROESY, los tiempos de mezcla utilizados han sido de 80 y 300 ms respectivamente.

Los espectros NOESY, en  $\text{D}_2\text{O}$ , se han realizado con tiempos de mezcla de 100, 200 y 300 ms. El tiempo de repetición entre acumulación ha sido de al menos 3s para permitir la mayor relajación posible del sistema.

- Espectros en  $\text{H}_2\text{O}$ .

Los espectros monodimensionales se han registrado utilizando la secuencia de pulsos *jump-return*<sup>18</sup>, que permite observar protones que intercambian muy rápidamente con el solvente. La anchura espectral en este caso ha sido de 20 ppm.

<sup>16</sup>Redfield, A. G.; Kuntz, S. D. *J. Magn. Reson.* **1975**, 19, 250.

<sup>17</sup>Bax, A.; Davis, D.G. *J. Magn. Reson.* **1985**, 65, 355-360.

<sup>18</sup>Plateau, P.; Güeron, M. *J. Am. Chem. Soc.* **1982**, 104, 7310.

En experimentos NOESY en H<sub>2</sub>O, la supresión de agua se ha conseguido incluyendo un módulo WATERGATE<sup>19</sup> en la secuencia de pulsos antes de la adquisición. El tiempo de mezcla para estos experimentos ha sido de 200 ms.

### 3.1.3.- PROCESADO DE ESPECTROS.

Para el procesado de los espectros se ha utilizado el programa UXNMR de Bruker. La matriz de datos se ha rellenado de ceros hasta obtener, después de la doble transformada de Fourier, matrices de 2Kx2K. Se han utilizado diferentes funciones de apodización en función del tipo de experimento, pero en la mayoría de casos se han empleado ventanas tipo seno cuadrado desplazadas en  $\pi/3$  y en  $\pi/4$  para  $\omega_1$  y  $\omega_2$  respectivamente.

En experimentos en H<sub>2</sub>O, la señal residual del agua fue eliminada mediante el procesado de espectros con el método de Marion et al<sup>20</sup>.

En todos los casos se ha realizado una corrección de la línea de base. Esta corrección ha consistido en el ajuste del plano base a un polinomio de potencia 5. En el caso de espectros con una importante señal residual de H<sub>2</sub>O, dicha corrección se llevó a cabo por zonas, de 20 ppm hasta la señal de H<sub>2</sub>O y de la señal de H<sub>2</sub>O a 0 ppm.

### 3.1.4.- ASIGNACIÓN DE ESPECTROS.

La asignación de los espectros se ha realizado utilizando la metodología descrita en el apartado 2.0.2 de la Parte Teórica. Dado que todas las secuencias estudiadas contenían residuos de citidina, en todos los casos, la asignación se ha podido realizar tomando como punto de inicio uno de dichos residuos.

Para la asignación de las formas dimericas y monomericas se ha utilizado el espectro ROESY, que permite la identificación de picos debidos al equilibrio entre las dos formas. En los casos en los que ha sido posible, la asignación se ha realizado por separado, en condiciones de 100% dímero o 100% monómero, aunque debido a la gran dispersión entre señales de dímero y monómero la asignación de las dos especies se ha podido seguir también en espectros donde ambas especies estaban presentes.

La asignación total de los espectros se llevó a cabo con ayuda del programa de análisis espectral XEASY<sup>21</sup>, que permite la asignación semiautomática de los picos de correlación. La evaluación cuantitativa de las intensidades de los picos NOE se llevó a cabo automáticamente utilizando las rutinas de integración incluidas en dicho paquete. En el caso de picos de correlación con mucho solapamiento la selección de la región de integración se efectuó manualmente. De esta forma se obtuvieron listas de picos, con las correspondientes integrales, para cada tiempo de mezcla.

<sup>19</sup> Piotto, P.; Saudek, V.; Sklenar, V. *J. Mol. Biol.* **1992**, *2*, 661-665.

<sup>20</sup> Marion, D.; Kura, T.; Bax, J. *Magn. Res.* **1989**, *84*, 425-430.

<sup>21</sup> Bartels, C.; Xia, T.; Billeter, M.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1995**, *6*, 1-10.

## 3.2.- CÁLCULO DE ESTRUCTURAS.

### 3.2.1.- OBTENCIÓN DE RESTRICCIONES DE DISTANCIAS.

- *Obtención de distancias cualitativas:*

Las restricciones de distancias cualitativas se obtuvieron a partir de los espectros NOESY, asignando a cada pico de correlación una cota superior de distancia de 3.0, 4.0 o 5.0 Å. Las restricciones de distancias se calcularon primero a partir de las listas de picos integradas procedentes del paquete XEASY, utilizando el programa CALIBA<sup>22</sup>. Este programa hace uso de la aproximación de dos espines y calcula las distancias a partir de expresiones del tipo

$$\text{NOE} \propto \tau m/r^6 \text{ o } \text{NOE} \propto \tau m/r^3 \text{ (para el caso de metilos)}$$

en donde la constante de proporcionalidad se calcula a partir de las intensidades NOE entre protones cuya distancia en la estructura es conocida (H5-H6 de C, Me-H6 de T, etc.). Las distancias resultantes fueron redondeadas a los valores superiores más cercanos (3.0, 4.0 o 5.0 Å). Las listas de picos obtenidas se manipularon teniendo en cuenta las siguientes consideraciones:

i) Se duplicaron los NOEs correspondientes a protones de residuos degenerados teniendo en cuenta la simetría del monómero o del dímero. En el caso de d<pCATTTCATT>, se duplicaron todos los NOEs una vez para el monómero y dos veces para el dímero, teniendo en cuenta la equivalencia de residuos 1-5, 2-6, 3-7 y 4-8 y la presencia de dos subunidades idénticas en el caso del dímero. En este segundo caso, los residuos de la segunda subunidad se numeraron del 9 al 16 y la equivalencia de residuos pasa a ser de 1-5-9-13, 2-6-10-14, 3-7-11-15 y 4-8-12-16. En total se obtuvieron dos restricciones de distancia por cada pico de correlación para la forma monomérica y cuatro para la forma dimérica.

En el caso de d<pTGCTCGCT>, se procedió de manera similar pero teniendo en cuenta la diferencia entre los residuos situados en posición 1 y 5.

ii) Si bien, en el caso de d<pCATTTCATT>, la degeneración de señales afecta por igual a todos los protones de la molécula, en el caso de d<pTGCTCGCT> la intensidad de los NOES con protones de los residuos 1 y 5 es la mitad que en el resto. Para tener en consideración este efecto, se dividieron por un factor de 1.12 todas las distancias correspondientes a estos protones. Este factor se obtiene a partir de la expresión que relaciona la intensidad del NOE con la distancia interprotónica. Dada la expresión  $\text{NOE} \propto \tau m/r^6$ , un factor 2 en la intensidad se traduce en un factor  $2^{1/6}$  en términos de distancia.

iii) En aquellas restricciones donde están involucrados protones H5'/H5" no asignados estereoespecíficamente, se introdujo el pseudoátomo correspondiente y se añadió una corrección de 0.9 Å a la cota superior de distancias. De manera análoga se trataron los protones H2'/H2" en caso de estar degenerados.

iv) Se obtuvieron listas de restricciones para monómero y dímero a partir de experimentos NOESY realizados a varios tiempos de mezcla. En caso de conflicto entre diferentes experimentos, se dio prioridad a la distancia resultante del espectro realizado con

<sup>22</sup> Güntert, P.; Braun, W. & Wüthrich, K. *J. Mol. Biol.* **1991**, 217,517-530.

menor tiempo de mezcla.

- *Obtención de distancias precisas:*

Para obtener restricciones de distancias precisas se utilizaron las intensidades de los picos NOE obtenidas con XEASY en cálculos de matriz completa de relajación, utilizando para ello el programa MARDIGRAS<sup>23,24</sup>. Antes de introducir las intensidades experimentales en el cálculo de MARDIGRAS, se duplicaron las listas de intensidades según los criterios expuestos en el punto i) del apartado anterior. En el caso de d<pTGCTCGCT>, se dividieron por dos las intensidades correspondientes a picos de correlación entre protones de residuos degenerados (todos menos el 1 y 5). En el caso de d<pCATTATT> esta corrección no es necesaria porque todos los NOEs se ven afectados por un mismo factor global.

Con estas listas de intensidades modificadas, y considerando únicamente protones no intercambiables, se llevaron a cabo los cálculos de matriz completa de relajación. En los cálculos con MARDIGRAS no se tuvieron en consideración movimientos internos, excepto en el caso de los grupos metilo, para los que se utilizó un modelo de *three-state jump*<sup>25</sup>. Se llevaron a cabo diversos cálculos utilizando diferentes modelos iniciales, intensidades procedentes de experimentos con diferente tiempo de mezcla y varios tiempos de correlación, para evitar errores procedentes de utilizar unas condiciones iniciales concretas. Como modelos iniciales se utilizaron las tres estructuras para las que se obtuvieron los valores más bajos de la función blanco en los cálculos con DYANA. Los tiempos de correlación utilizados fueron de 1.0, 2.0 y 4.0 ns, asumiendo en todos los casos un movimiento isotrópico de la molécula. Las intensidades experimentales se obtuvieron de espectros realizados con tiempos de mezcla de 100, 200 y 300 ms. Las distancias finales se obtuvieron promediando las 27 salidas de MARDIGRAS (3 tiempos de mezcla x 3 estructuras iniciales x 3 tiempos de correlación). Las cotas superiores e inferiores para cada distancia se calcularon a partir del valor medio de cada distancia sumando o sustrayendo la desviación cuadrática media. En casos donde esta desviación era demasiado pequeña (normalmente para distancias muy cortas), se utilizó un valor de  $\pm 0.25\text{Å}$  como desviación mínima.

Puesto que tanto el dímero como el monómero son moléculas simétricas, las restricciones de distancias para pares de protones equivalentes deben ser idénticas. Aunque en estos casos las restricciones resultantes de MARDIGRAS son muy similares entre sí, la lista final de restricciones utilizada en el cálculo de la estructura se confeccionó utilizando los valores promedio para las cotas superiores e inferiores en cada grupo de restricciones equivalentes.

A la lista de restricciones de distancia obtenida mediante MARDIGRAS para los protones no intercambiables, se le añadieron restricciones cualitativas para los protones intercambiables. En el análisis de las intensidades de picos NOE de experimentos en H<sub>2</sub>O no se tuvo en cuenta el efecto de intercambio con el solvente, por lo que en el caso de protones intercambiables se utilizaron únicamente cotas superiores para las restricciones de distancias.

El tratamiento de los protones que no pudieron ser asignados estereoespecíficamente se

<sup>23</sup> Borgias, B.A.; James, T.L. *J. Magn. Reson.* **1988**, 79, 493-512.

<sup>24</sup> Borgias, B.A.; James, T.L. *J. Magn. Reson.* **1990**, 87, 475-487.

<sup>25</sup> Liu, H.; Thomas, P.D.; James, T.L. *J. Magn. Reson.* **1992**, 98, 163-175.

llevó a cabo de manera análoga al usado con las restricciones cualitativas.

- *Restricciones de ángulos de torsión:*

En los casos en los que el análisis de las constantes de acoplamiento indicaban que los anillos de desoxiribosa se encontraban mayoritariamente en conformación tipo S, se introdujeron restricciones en los ángulos de torsión para penalizar las conformaciones tipo N. Estas restricciones fueron:  $\gamma_1$  entre 5 y 65,  $\gamma_2$  entre -65 y -5, y  $\delta$  entre 110 y 170.

En algunos casos, se pudieron analizar los NOEs H4'-H5'/5'', H3'-H5'/5'' así como las constantes  $^3J_{4'5'}$  y  $^3J_{4'5''}$  obteniéndose valores consistentes con una conformación  $g^+$  para el ángulo  $\gamma$  del esqueleto. Es estos casos, se introdujo una restricción angular para dicho ángulo entre 30 y 90.

- *Restricciones de puente de hidrógeno:*

Tanto en los cálculos con restricciones cualitativas como cuantitativas, además de las restricciones experimentales, se introdujeron restricciones de distancia y angulares correspondientes a los puentes de hidrógeno característicos de apareamientos tipo Watson-Crick. Los valores exactos de estas restricciones se muestran en la siguiente tabla:

**Tabla M.10.- Restricciones de ángulos y distancias correspondientes a puente de hidrógeno en un apareamiento tipo Watson-Crick**

Ángulos		Distancias	
Pares GC			
N2-H21-O2	170-190°	N2-O2	2.72-2.92Å
N1-H1-N3	170-190°	N1-N3	2.85-3.05Å
O6-H41-N4	170-190°	O6-N4	2.85-3.05Å
Pares AT			
N6-H62-O4	170-190°	N6-O4	2.85-3.05Å
N1-H3-N3	170-190°	N1-N3	2.85-3.05Å

### 3.2.2.- DETERMINACIÓN DE ESTRUCTURAS.

- *Cálculos de estructura con DYANA.*

Los modelos iniciales de estructura se calcularon con el programa DYANA (versión 1.4). En el cálculo de las estructuras diméricas, se introdujeron 50 "linkers" entre las dos subunidades. Además de las restricciones experimentales, se introdujeron restricciones entre C4' y O4' para "cerrar" los anillos de desoxiribosa, así como restricciones entre O3' del residuo 8(16) y el átomo de fósforo del residuo 1(9) para "cerrar" los ciclos de cada subunidad. En todos los cálculos se utilizó la librería de residuos estándar del programa: "dyana.lib".

En cada cálculo se generaron 200 estructuras al azar y se sometieron a 4000 pasos de dinámica molecular restringida en el espacio de ángulos de torsión. La dinámica se llevó a cabo siguiendo un protocolo tipo "annealing" implementado en el programa con ligeras modificaciones: la temperatura se aumentó paulatinamente entre 0 y 8 durante los primeros 800 pasos, para descender después de nuevo hasta cero durante los 3200 pasos restantes. Los valores de temperatura corresponden a unidades de función blanco por grado de libertad de la

molécula<sup>26</sup>. Al final de la dinámica, las estructuras se someten a 1000 pasos de minimización. De las 200 estructuras finales, se eligieron las 50 con valores menores de función blanco y se sometieron nuevamente a un proceso de calentamiento y enfriado similar. En este caso, 6000 pasos calentando sólo hasta un valor de temperatura de 4 durante los primeros 1200 pasos. Las 10 estructuras con valores más pequeños de función objetivo se tomaron como estructuras resultantes.

- *Refinamiento de la estructura con AMBER.*

Para el refinamiento de las estructuras tridimensionales de las formas diméricas se realizaron cálculos de dinámica molecular restringida en espacio de coordenadas cartesianas con el paquete AMBER (versión 4.1).

Al potencial habitual de programa se le añadió un término correspondiente a las restricciones de distancia experimentales. Dicho término tiene la forma:

$$E_{\text{NOE}} = \begin{cases} = 2k_{\text{NOE}}(r_2-r_1)(r_1-r) + k_{\text{NOE}}(r_2-r_1)^2 & \text{si } r < r_1 \\ = k_{\text{NOE}}(r_2-r)^2 & \text{si } r_1 < r < r_2 \\ = 0 & \text{si } r_2 < r < r_3 \\ = k_{\text{NOE}}(r-r_3)^2 & \text{si } r_3 < r < r_4 \\ = 2k_{\text{NOE}}(r_4-r_3)(r-r_4) + k_{\text{NOE}}(r_4-r_3)^2 & \text{si } r_4 < r \end{cases}$$

donde  $r_3$  y  $r_2$  son las cotas superior e inferior de la restricción de distancia. La forma de esta función de penalización es parabólica entre  $r_3$  y  $r_4$  y entre  $r_1$  y  $r_2$ . Cuando la distancia es mayor que  $r_4$  o menor que  $r_2$ , la función de penalización es lineal. La anchura de las dos regiones parabólicas es de 1 Å. El valor de la constante  $k_{\text{NOE}}$  varía a lo largo del cálculo como se describe más adelante. Se utilizaron términos similares a éste para las restricciones de ángulos de torsión derivadas de constantes de acoplamiento, así como para las restricciones de puentes de hidrógeno.

Todos los cálculos con AMBER se realizaron en el vacío, considerando todos los átomos de la molécula (no se utilizó la aproximación de átomo unido). Los parámetros utilizados para el campo de fuerza fueron los correspondientes a la parametrización de 1991 ("parm91.dat")<sup>27</sup>. Se añadieron contraiones  $\text{Na}^+$  hexahidratados a una distancia de 5 Å de todos los átomos de fósforo para neutralizar las cargas negativas de los grupos fosfato. El término electrostático se calculó utilizando una constante dieléctrica dependiente de la distancia. El valor de corte para las interacciones de no enlace se fijó en 10 Å. La temperatura y los pesos relativos de las restricciones experimentales se variaron durante las simulaciones, de acuerdo con la siguiente estrategia: Cada una de las estructuras de partida fue primero minimizada y después sometida a 40 ps de dinámica molecular. Durante los primeros 4 ps la temperatura se aumento linealmente de 300 a 1500°K, se mantuvo a este valor durante 16 ps y se descendió a

<sup>26</sup> La función blanco tiene unidades arbitrarias y el número de grados de libertad de la molécula es el número de ángulos de torsión.

<sup>27</sup> Weiner, P.; Kollman, P.A.; Nguyen, D.T.; Case, D.J. *Comput. Chem.* **1986**, *7*, 230-252.

300°K en 10 ps. El valor final se mantuvo constante durante los restantes 10 ps. Las constantes de fuerza para las restricciones de distancia se variaron simultáneamente con la temperatura. El valor de  $k_{\text{NOE}}$  se aumentó de  $10 \text{ kcal}/(\text{mol})^2$  a  $100 \text{ kcal}/(\text{mol})^2$ , se mantuvo este valor durante el período de alta temperatura y posteriormente se bajó a  $50 \text{ kcal}/(\text{mol})^2$ , manteniéndose este valor durante el resto de la dinámica. La estructura resultante se sometió a 1000 pasos de minimización de energía restringida con una constante  $k_{\text{NOE}}$  de  $10 \text{ kcal}/(\text{mol})^2$ . Los valores de  $E_{\text{exp}}$ , presentados en las Tablas 2.1 y 2.2, corresponden a este valor de la constante.

La temperatura se mantuvo al valor requerido re-escalando las velocidades cuando las fluctuaciones superaban los 20°K. El algoritmo SHAKE se usó para mantener fijas las longitudes de enlace, permitiendo una constante de 2 fs en la integración de las ecuaciones de movimiento.

### 3.2.3.- ANÁLISIS DE LAS ESTRUCTURAS.

El análisis de las estructuras finales, así como de las trayectorias de dinámica molecular, se llevó a cabo con los programas Curves<sup>28</sup>, Dials and Windows<sup>29</sup> y MOLMOL (versión 2.6 )<sup>30</sup>. Las figuras de modelos moleculares se han realizado con el programa MOLMOL.

El paquete SYBYL versión 6.1 (Tripos Inc., St. Louis, MO) se usó para el modelado teórico de los dímeros del Capítulo 3.

---

<sup>28</sup> Lavery, R.; Sklenar, H. CURVES 1990. Curves 3.0, helical analysis of irregular nucleic acids, Laboratory of Theoretical Biochemistry CNRS, Paris.

<sup>29</sup> Ravishanker, G.; Swaminathan, D.L.; Beveridge, D.L.; Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669-699.

<sup>30</sup> Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51.

## 4.- ESPECTROSCOPIA DE UV Y DICROÍSMO CIRCULAR.

### 4.1.- INSTRUMENTACIÓN.

UV: Los espectros y curvas de fusión de absorción UV se han realizado en un espectrofotómetro Varian Cary 5E UV-Vis-NIR.

DC: Los espectros y curvas de fusión de dicroísmo circular (DC) se han realizado en un espectropolarímetro Jasco J-720, provisto de un accesorio de cubetas termostatzable conectado a un baño refrigerante Neslab RP-100. Se han utilizado cubetas de 1cm de paso y de 400 $\mu$ L de capacidad (HELLMA).

### 4.2.- PREPARACIÓN DE LAS MUESTRAS.

Las muestras se han obtenido disolviendo los productos, en forma de sal Na<sup>+</sup>, en H<sub>2</sub>O o en el medio salino indicado en cada caso. Se ha trabajado utilizando concentraciones de oligonucleótido entre 5-50 $\mu$ M.

En todos los casos se ha utilizado como tampón Na<sub>2</sub>PIPES 10mM a pH=7, obtenido por dilución de una disolución madre 100mM preparada tal como se indica en el apartado 0.3 de Materiales y Métodos. El pH que se indica es el de la disolución madre.

Para las curvas de fusión, previamente se han calentado las muestras a 90°C durante 5min y se han dejado enfriar lentamente dentro del mismo baño hasta temperatura ambiente. Posteriormente se han guardado en la nevera a 4°C hasta su utilización.

### 4.3.- ADQUISICIÓN DE LOS ESPECTROS.

UV: Para los espectros de UV se ha realizado un barrido entre 200-320nm. Las curvas de fusión se han realizado monitorizando la variación de la absorbancia a  $\lambda=260$ nm, ya que todos los productos presentan el máximo de absorción cerca de este valor. Se ha utilizado un rango de temperatura de 10-90°C, variando la temperatura a 0.5°C/min. En las curvas de fusión se ha mantenido purga de N<sub>2</sub> constante en el compartimento de muestra.

DC: Para los espectros de DC se ha realizado un barrido entre 220-320nm. Las curvas de fusión por DC se han realizado monitorizando la variación de  $\Delta\epsilon$  a la longitud de onda correspondiente al máximo de la banda más intensa. Las curvas de fusión se han realizado entre 5-90°C variando la temperatura a 20°C/h.

En todos los espectros se ha realizado una corrección de la línea de base por eliminación del blanco. En el caso del DC se ha realizado además un suavizado del espectro.

### 4.4.- TRATAMIENTO DE LAS CURVAS DE FUSIÓN.

Las curvas de fusión se han ajustado utilizando la aproximación de equilibrio entre dos estados, tal como se describe en el apartado 4.2.2. El ajuste se ha realizado utilizando el programa Origin, versión 3.73, aplicando una variación fraccional de  $\chi^2 < 0.05$  o bien una rutina de mínimos cuadrados no lineal ML.

El error asociado a la determinación de los distintos parámetros termodinámicos a partir de las curvas de fusión, se ha calculado aplicando el método de propagación de errores<sup>31</sup>. Para cada uno de los parámetros se han obtenido las siguientes expresiones:

Ajuste de las curvas según la expresión

$$y = P_1 + (P_2 - P_1) \cdot \left[ \frac{P_3 \cdot \exp\left(\frac{P_4}{T}\right)}{\left(1 + P_3 \cdot \exp\left(\frac{P_4}{T}\right)\right)} \right]$$

donde  $P_1=y_0$ ,  $P_2=y_F$ ,  $P_3=\exp(\Delta S^0/R)$  y  $P_4=-\Delta H^0/R$ :

$$\delta\Delta H^0 = R\delta P_4$$

$$\delta\Delta S^0 = \frac{R}{P_3} \cdot \delta P_3$$

$$\delta\Delta G^0 = R\delta P_4 \cdot \left(1 - \frac{T}{T_m}\right)$$

Ajuste de las curvas según la expresión:

$$y = P_1 \cdot \exp\left(\frac{P_2}{T}\right)$$

donde  $P_1=\exp(\Delta S^0/R)$  y  $P_2=-\Delta H^0/R$ .

$$\delta\Delta H^0 = R\delta P_2$$

$$\delta\Delta S^0 = \frac{R}{P_1} \cdot \delta P_1$$

$$\delta\Delta G^0 = R\delta P_2 + T \left( \frac{R}{P_1} \cdot \delta P_1 \right)$$

$$\delta T_m = \frac{\delta P_2 \cdot (\ln c_0 + \ln P_1) + \left( \frac{\delta P_1}{P_1} \cdot P_2 \right)}{(\ln c_0 + \ln P_1)^2}$$

#### 4.5.- ANÁLISIS DE OLIGONUCLEÓTIDOS CÍCLICOS POR DC.

Aunque el DC no se ha utilizado como técnica de caracterización, las diferencias observadas entre los espectros correspondientes a distintas secuencias, así como entre las formas monomérica y dimérica, convierten en dicha técnica en un buen complemento para la caracterización de oligonucleótidos cíclicos. En la Tabla M.11 se muestran los valores de  $\lambda_{\text{máx}}$ ,  $\lambda_{\text{mín}}$  y elipticidad molar para cada una de las secuencias estudiadas.

<sup>31</sup> Spiegel, M.R.; Abellanas, L. «Fórmulas y Tablas de Matemática Aplicada» 1989, 207-208, McGRAW-HILL, Inc., U.S.A.

**Tabla M.II.- Análisis por DC de las formas monomérica y dimérica de distintos octámeros cíclicos.**

Secuencia	Forma monomérica (H <sub>2</sub> O)				Forma dimérica (Medio Salino) <sup>§</sup>			
	λ <sub>máx</sub> nm	[θ]	λ <sub>mín</sub> nm	[θ]	λ <sub>máx</sub> nm	[θ]	λ <sub>mín</sub> nm	[θ]
d<pCATT<CATT>	274.5	3.59E5	252.0	-1.62E5	-	-	-	-
d<pTGCTCGCT>	287.0	2.97E5	264.0	-1.22E5	290.0	2.70E5	260	-9.84E4
	243.5	5.44E5			241.0	1.06E5		
d<pTGCATGCA>	281.5	2.02E5	257.5	-6.71E4	291.0	2.06E5	254.5	-1.02E5
	240.5	3.69E4			237.5	4.69E4		
d<pAGCTAGCT>	285.5	1.61E5	259.5	-9.83E4	290.0	1.67E5	260	-1.19E5
					233.5	5.98E4		
d<pCCGTCCGT> No estructurado	282.5	1.1E5	239.5	-1.3E4	265.0	4.24E5	243	-9.70E4
	255.5	8.17E3						
d<pCGCTCATT> No estructurado	278.5	8.27E4	247.0	5.87E4	-	-	-	-

<sup>§</sup> Na<sub>2</sub>PIPES 10 mM, pH=7, NaCl 100 mM, MgCl<sub>2</sub> 10 mM.

---

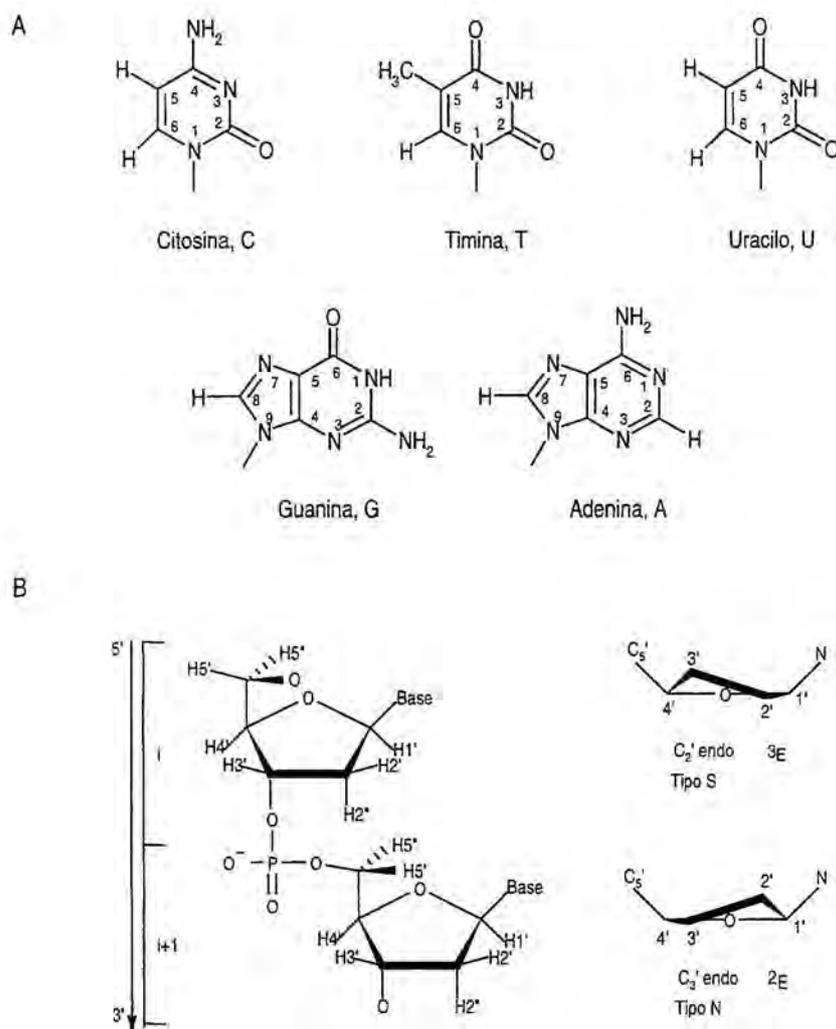
*ANEXOS*

---

## ANEXO A

### NOMENCLATURA Y DEFINICIÓN DE PARÁMETROS ESTRUCTURALES EN ÁCIDOS NUCLEICOS.

#### ❖ NOMENCLATURA, ESTRUCTURA Y NUMERACIÓN DE β-D-RIBOSAS Y DE LAS BASES DEL DNA.

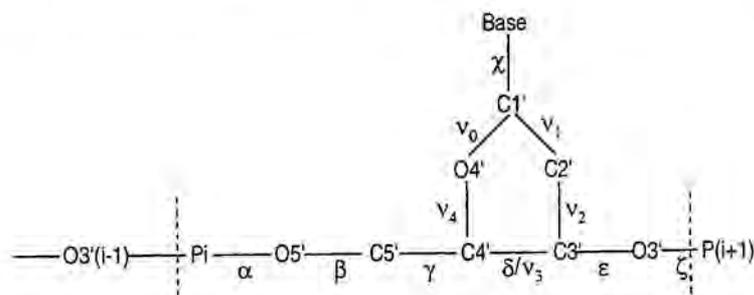


**Figura A.1-** Nomenclatura, estructura y numeración estándar de acuerdo con la IUPAC/IUB<sup>1</sup> de: A) las cinco bases habituales (pirimidinas C, T y U; purinas G y A) y B) β-D-desoxirribosas y puckering de las dos conformaciones más habituales en β-D-desoxirribosas: C2' endo (tipo S) y C3' endo (tipo N).

<sup>1</sup> IUPAC/IUB Joint Commission Biochemical Nomenclature, *Eur. J. Biochem.* 1983, 131, 9.

❖ CONFORMACIÓN DEL ANILLO DE β-D-DESOXIRIBOSA.

La conformación del anillo de ribosa se puede describir en base a los ángulos de torsión, definidos tal como se muestra a continuación:



**Figura A.2-** Ángulos de torsión del esqueleto azúcar-fosfato ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\varepsilon$  y  $\zeta$ ), ángulo de torsión glicosídico  $\chi$  y ángulos de torsión endocíclicos  $v_0$ ,  $v_1$ ,  $v_2$ ,  $v_3$  y  $v_4$ .

**Tabla A.1-** Definición de los ángulos de torsión

Ángulo de torsión	Átomos involucrados
$\alpha$	(n-1)O3'-P-O5'-C5'
$\beta$	P-O5'-C5'-C4'
$\gamma$	O5'-C5'-C4'-C3'
$\delta$	C5'-C4'-C3'-O3'
$\varepsilon$	C4'-C3'-O3'-P
$\zeta$	C3'-O3'-P-O5'(n+1)
$\chi$ (Py)	O4'-C1'-N1-C2
$\chi$ (Pu)	O4'-C1'-N9-C4
$v_0$	C4'-O4'-C1'-C2'
$v_1$	O4'-C1'-C2'-C3'
$v_2$	C1'-C2'-C3'-C4'
$v_3$	C2'-C3'-C4'-P4'
$v_4$	C3'-C4'-O4'-C1'

Puesto que los ángulos de torsión endocíclicos no son independientes entre sí, la conformación del anillo puede describirse únicamente con dos parámetros introduciendo el concepto de pseudorotación<sup>2</sup>. La fase P y la amplitud  $\Phi_m$  del ángulo de pseudorotación se relacionan con los ángulos de torsión endocíclicos mediante la expresión:

$$v_j = \Phi_m \cdot \cos(P + 144 \cdot (j - 2))$$

donde  $j = 0-4$ .

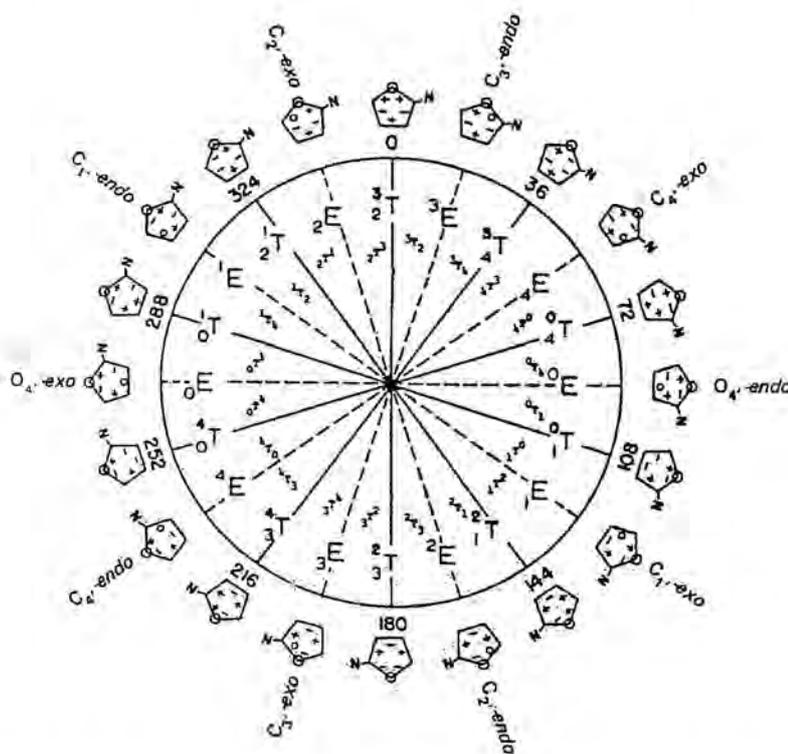
Esta ecuación es en realidad únicamente aplicable en el caso de anillos de cinco miembros equilaterales. En el caso de anillos heterocíclicos se han de introducir correcciones de no idealidad,

<sup>2</sup> Altona, C; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, 94, 8205.

obteniéndose para los ángulos de torsión entre protones de un anillo de ribosa las siguientes expresiones:

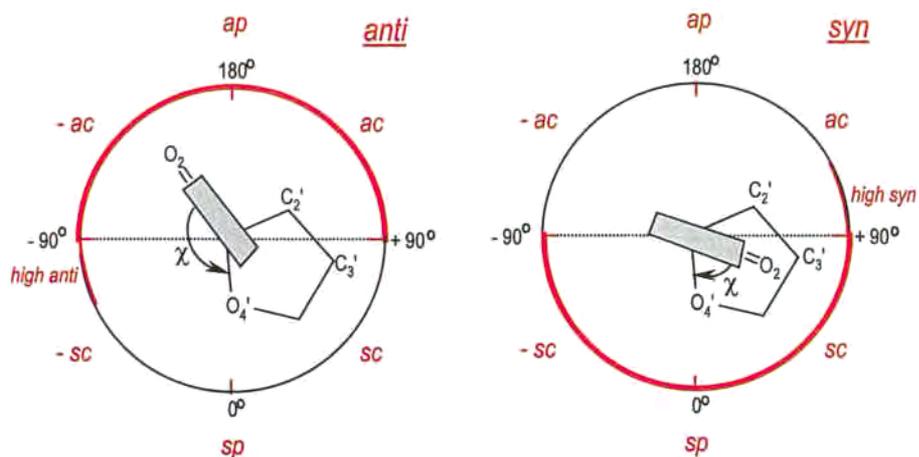
$$\begin{aligned}\phi_{1'2'} &= 121.4 + 1.03\Phi_m \cos(P - 144) \\ \phi_{1'2''} &= 0.9 + 1.02\Phi_m \cos(P - 144) \\ \phi_{2'3'} &= 2.4 + 1.06\Phi_m \cos(P) \\ \phi_{2'3''} &= 122.9 + 1.06\Phi_m \cos(P) \\ \phi_{3'4'} &= -124.0 + 1.09\Phi_m \cos(P - 144)\end{aligned}$$

El valor de  $\Phi_m$  varía entre  $34^\circ$  y  $42^\circ$ . En la Fig. I.III se muestra como varía la conformación del azúcar en función de la fase del ángulo de pseudorotación,  $P$ , siguiendo un ciclo completo de  $360^\circ$ .



**Figura A.3.-** Ciclo de pseudorotación del anillo de ribosa, donde se muestra la relación entre la fase del ángulo de pseudorotación ( $P$ ) y las formas sobre ( $E$ ) y twist ( $T$ ) y las notaciones endo y exo. Los valores de  $P$  preferidos se encuentran en dos partes de la rueda:  $P=0^\circ-36^\circ$  ( $C3'$ endo; tipo  $N$ ) y  $P=144^\circ-180^\circ$  ( $C2'$ endo; tipo  $S$ ).

❖ DISPOSICIÓN DE LA BASE RESPECTO DEL AZÚCAR.



**Figura A.4.-** Disposición *syn* y *anti* de la base respecto del azúcar de acuerdo con la normativa IUPAC <sup>1</sup>.

❖ NOTACIÓN DE DISTANCIAS.

**d<sub>i</sub>(l;r)** distancias intranucleotídicas

**d<sub>s</sub>(l;r)** distancias secuenciales, l: protón del nucleótido 5'; r: protón del nucleótido 3'

**d<sub>ci</sub>(l;r)** distancias entre pares de bases

**d<sub>cs</sub>(l;r)** distancias entre nucleótidos apareados adyacentes, p: indica si es entre nucleótidos 3' o 5'

**d(l;r)** distancias entre protones alejados en la secuencia

❖ ALGUNAS POSIBLES FORMAS DE APAREAMIENTO ENTRE LAS BASES.

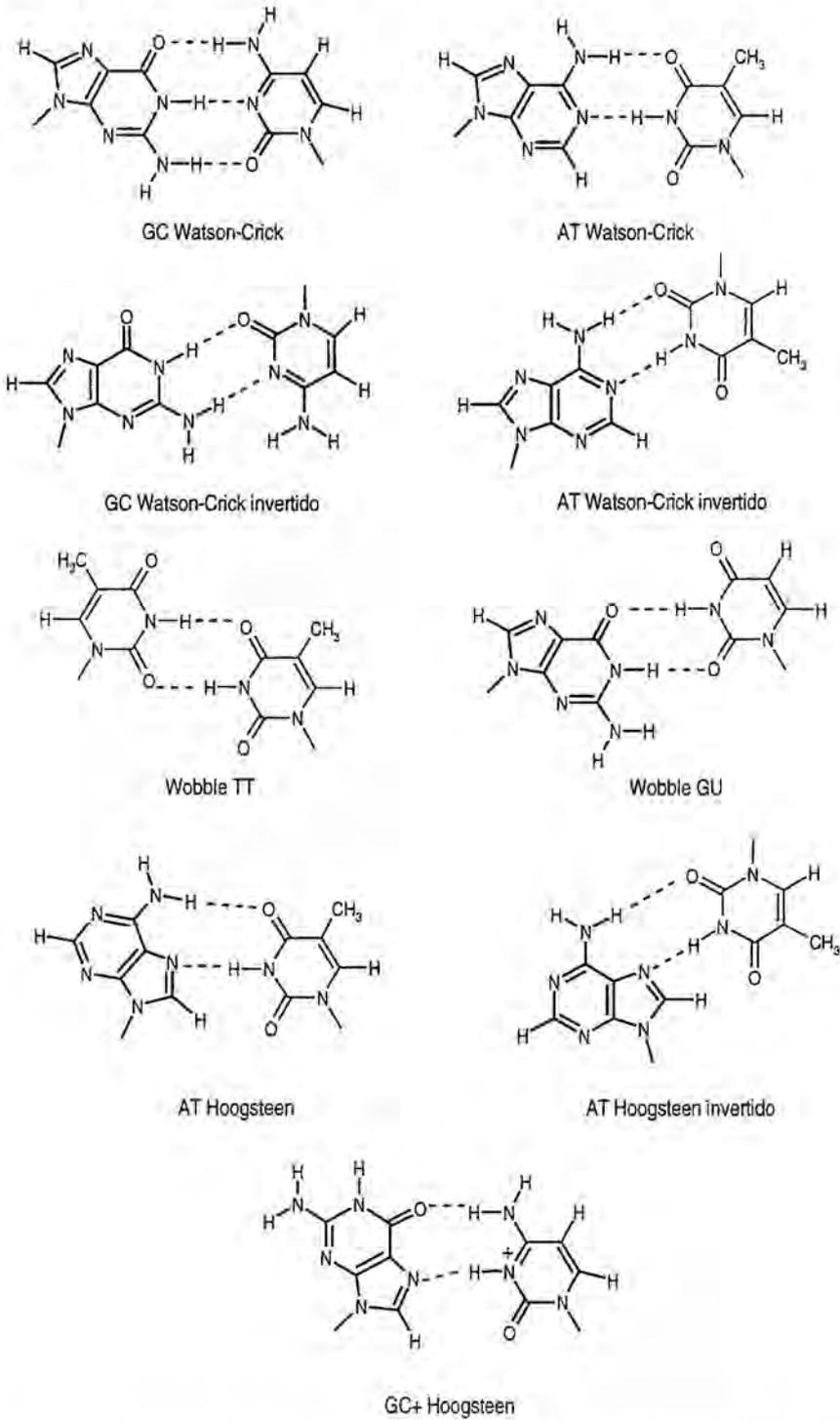


Figura A.5.- Algunas de las posibles formas de apareamiento de bases en DNA y RNA.

❖ PARÁMETROS DE HÉLICE.

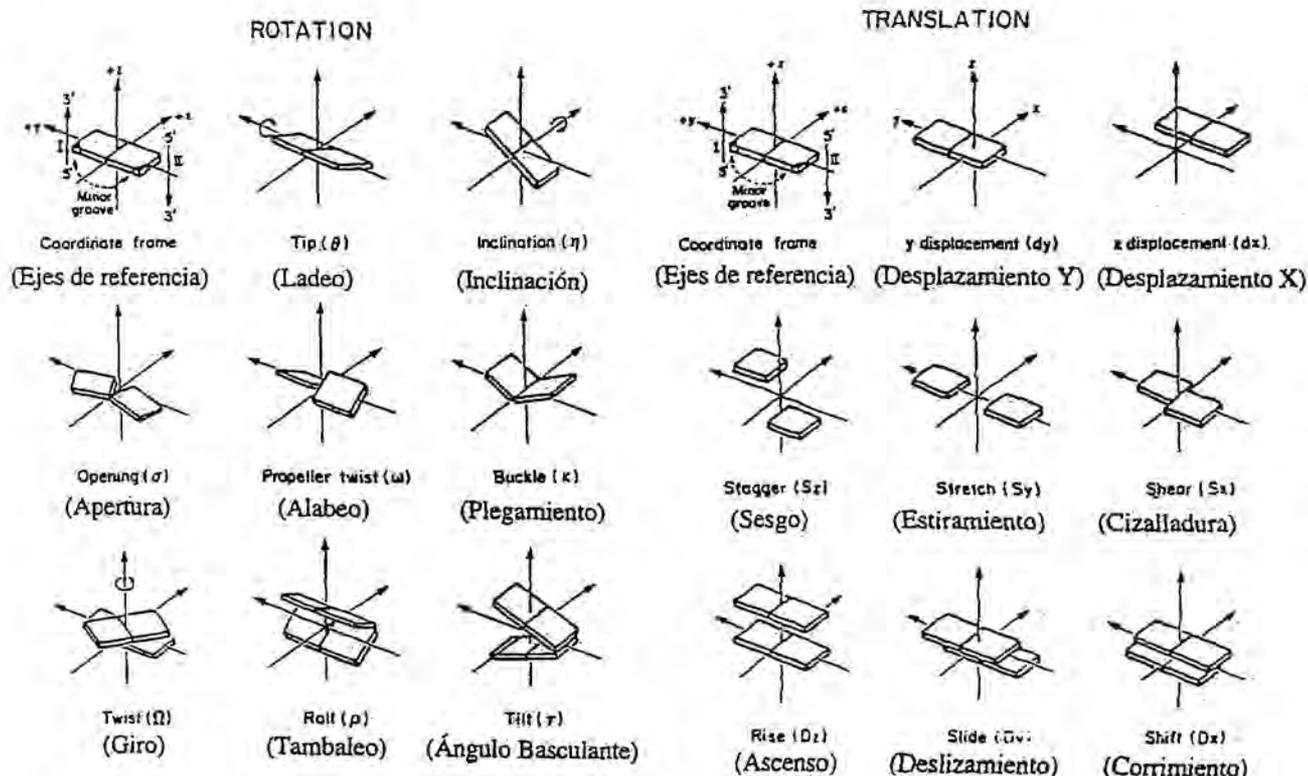


Figura A.6.-Definición de los parámetros de estructura de ácidos nucleicos<sup>3</sup>. En las primeras tres columnas se definen parámetros de rotación, que involucran o bien dos bases de un par (dos líneas superiores) o bien dos pares de bases sucesivos (línea inferior). En las últimas tres columnas se definen los parámetros de traslación de dos bases de un par (dos líneas superiores) o bien de dos pares de bases sucesivos (línea inferior).

<sup>3</sup> Diekman, S. EMBO J. 1989, 8, 1.

## ANEXO B

### TABLAS DE ASIGNACIÓN.

Los desplazamientos se expresan en ppm.

*Tabla B.1.- Desplazamientos químicos de d<pCATTcATT), T=20°C. Forma dimérica.*

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
C1(5)	-	7.29	6.74	7.97	6.11	6.49	2.24	2.59	4.53	4.65	4.09	4.19
A2(6)	-	8.03	7.07	8.69	8.32	6.57	3.00	3.25	5.15	4.42	4.04	4.31
T3(7)	14.46	-	-	7.36	1.55	6.62	2.46	2.32	5.15	4.47	4.08	
T4(8)	10.74	-	-	7.02	1.74	5.90	1.78	2.37	4.57	3.19	3.02	3.57

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para T=5°C.

*Tabla B.2.- Desplazamientos químicos de d<pCATTcATT), T=20°C. Forma monomérica.*

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
C1(5)	-	n.o.	n.o.	7.49	5.59	5.28	1.60	2.08	4.59	3.75	3.93	
A2(6)	-	n.o.	n.o.	8.19	8.18	6.34	3.44	3.09	4.97	4.47	4.04	4.21
T3(7)	13.30	-	-	7.46	1.51	6.32	2.28	2.53	4.87	4.53	4.09	4.37
T4(8)	10.90	-	-	7.83	2.03	6.35	2.17	2.54	4.82	4.40	4.14	4.26

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para T=5°C.

*Tabla B.3.- Desplazamientos químicos de d<pTGCTcGCT>, T=20°C. Forma dimérica.*

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
T1	11.18	-	-	7.85	1.92	6.47	2.34	2.52	4.59	4.62	4.10	4.12
G2(6)	13.45	9.25	6.27	8.13	-	6.29	2.74	3.08	5.04	4.59	4.02	4.25
C3(7)	-	8.77	6.49	7.46	5.38	6.57	2.34		5.03	4.52	3.99	4.32
T4(8)	10.72	-	-	7.37	1.81	5.92	1.88	2.36	4.64	3.42	3.27	3.62
C5	-	n.o.	n.o.	7.99	6.11	6.47	2.26	2.57	4.57	4.59	4.12	

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para T=5°C.

*Tabla B.4.- Desplazamientos químicos de d<pTGCTcGCT>, T=20°C. Forma monomérica.*

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
T1	b	-	-	7.74	1.41	5.47	1.85	2.17	4.73	4.56	4.08	4.20
G2(6)	c	n.o.	n.o.	7.87	-	6.09	2.71	2.99	4.93	4.38	3.81	4.05
C3	-	8.48	6.85	7.50	5.48	6.15	2.03	2.61	4.66	5.03	4.01	4.35
T4	b	-	-	7.81	1.99	6.35	2.27	2.46	4.77	4.37	4.08	4.24
C5	-	n.o.	n.o.	7.72	5.56	5.50	1.77	2.28	4.71	4.55	4.02	4.17
C7	-	8.53	6.71	7.47	5.47	6.14	1.92	2.45	4.66			
T8	b	-	-	7.92	2.05	6.44	2.16	2.47	4.82	4.42	4.09	4.26

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para T=5°C.

b Los H3 de las tres timinas aparecen a 10.86, 11.59 y 11.82, pero no se pudieron asignar secuencialmente.

c Solamente en el caso del protón H1 se rompe la degeneración observada para las dos guanosinas: H1G2: 13.34, H1G6: 13.41.

**Tabla B.5.- Desplazamientos químicos de  $d<pCGCTCATT>$ ,  $T=25^{\circ}C$ . Forma dimérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
C1	-	n.o.	n.o.	7.99	6.11	6.47	2.23	2.56	4.54	4.62	3.99	4.17
G2	13.37	9.08	6.70	8.19	-	6.22	2.85	2.97	5.02	4.69	4.05	4.23
C3	-	8.69	6.90	7.57	5.51	6.55	2.38	2.51	5.17	4.69	4.33	
T4	n.o.	-	-	7.04	1.72	5.83	1.80	2.36	4.57	3.24	3.65	
C5	-	n.o.	n.o.	7.96	6.11	6.49	2.27	2.58	4.58	4.66	4.08	4.16
A6	-	8.00	7.04	8.68	8.37	6.50	2.93	3.20	5.11	4.72	4.06	4.31
T7	14.11	-	-	7.51	1.57	6.44	2.36	2.55	5.18	4.60	4.38	
T8	n.o.	-	-	7.50	1.85	5.89	1.87	2.38	4.60	3.11	3.41	3.66

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para  $T=5^{\circ}C$ .

**Tabla B.6.- Desplazamientos químicos de  $d<pCGCTCATT>$ ,  $T=25^{\circ}C$ . Forma monomérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
C1	-	n.o.	n.o.	7.52	5.96	6.08	1.76	2.28	4.67	4.11	3.90	
G2	n.o.	n.o.	n.o.	7.89	-	6.05	2.84	2.67	4.97	4.35	4.07	
C3	n.o.	n.o.	n.o.	7.86	5.98	6.30	2.29	2.58	4.90	4.45	4.12	4.22
T4	n.o.	-	-	7.60	1.82	6.14	2.16	2.38	4.77			
C5	-	n.o.	n.o.	7.45	5.95	6.02	1.65	2.20	4.68	4.11	3.90	
A6	-	n.o.	n.o.	8.33	8.06	6.31	2.89	2.81	5.01	4.42	4.08	
T7	n.o.	-	-	7.63	1.79	6.22	2.32	2.50	4.90	4.35	4.15	4.21
T8	n.o.	-	-	7.58	1.85	6.16	2.20	2.40	4.77	n.a.	n.a.	n.a.

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para  $T=5^{\circ}C$ .

**Tabla B.7.- Desplazamientos químicos de  $d<pAGCTAGCT>$ ,  $T=25^{\circ}C$ . Forma dimérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
A1(5)	-	n.o.	n.o.	8.52	8.29	6.65	2.79		4.68	4.73	4.13	4.28
G2(6)	13.47	9.28	6.31	8.17	-	6.35	2.77	3.11	5.06	4.65	4.06	4.31
C3(7)	-	8.79	6.52	7.48	5.41	6.59	2.37		5.06	4.53	4.20	4.35
T4(8)	n.o.	-	-	7.39	1.84	5.94	1.92	2.41	4.70	3.48	3.32	3.67

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para  $T=5^{\circ}C$ .

**Tabla B.8.- Desplazamientos químicos de  $d<pAGCTAGCT>$ ,  $T=25^{\circ}C$ . Forma monomérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
A1(5)	-	n.o.	n.o.	8.17	8.06	6.07	2.49	2.61	4.86	4.24	3.90	4.03
G2(6)	n.o.	n.o.	n.o.	7.94	-	6.01	2.75		4.91	4.41	4.05	4.18
C3(7)	-	n.o.	n.o.	7.62	5.66	6.12	2.09	2.48	4.72	4.23	4.02	
T4(8)	n.o.	-	-	7.51	1.87	6.05	1.87	2.27	4.70	4.21	3.91	3.98

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para  $T=5^{\circ}C$ .

**Tabla B.9.- Desplazamientos químicos de  $d<pTGCA TGCA>$ ,  $T=25^{\circ}C$ . Forma dimérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
T1(5)	n.o.	-	-	7.84	1.92	6.49	2.34	2.57	4.55	4.68	4.09	4.23
G2(6)	12.85	8.84	6.08	8.13	-	6.22	2.75	3.01	5.02	4.55	4.01	4.27
C3(7)	-	8.23	6.32	7.50	5.26	6.62	2.56	2.73	5.06	4.59	4.33	4.44
A4(8)	-	n.o.	n.o.	8.09	7.97	6.09	1.97	2.55	4.74	3.47	3.69	3.81

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para  $T=5^{\circ}C$ .

**Tabla B.10.- Desplazamientos químicos de d<pTGCATGCA>, T=25°C. Forma monomérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
T1(5)	n.o.	-	-	7.68	1.50	5.55	1.90	2.20	4.76	4.30	4.14	
G2(6)	n.o.	n.o.	n.o.	7.99	-	6.15	2.76	2.98	4.83	4.42	3.84	4.09
C3(7)	-	n.o.	n.o.	7.46	5.43	6.09	1.96	2.33	4.56	4.12	3.96	
A4(8)	-	n.o.	n.o.	8.47	7.99	6.42	2.97	2.88	4.87	4.46	3.88	4.09

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para T=5°C.

**Tabla B.11.- Desplazamientos químicos de d<pCCGTCCGT>, T=25°C. Forma dimérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
C1(5)	-	n.o.	n.o.	8.04	6.11	6.48	2.23	2.60	4.59	4.59	3.98	4.20
C2(6)	-	9.06	7.03	7.79	6.01	6.34	2.31	2.78	4.94	4.42	4.07	4.18
G3(7)	13.76	9.37	8.02	7.84	-	6.09	2.68	2.51	5.10	4.50	4.20	
T4(8)	10.51	-	-	7.56	1.67	5.99	1.94	2.45	4.70	3.81	3.81	3.74

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para T=5°C.

**Tabla B.12.- Desplazamientos químicos de d<pCCGTCCGT>, T=25°C. Forma monomérica.**

Protón	NH <sup>a</sup>	NH <sub>2</sub> (2) <sup>a</sup>	NH <sub>2</sub> (1) <sup>a</sup>	H6/H8	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	H5'/H5''	
C1(5)	-	n.o.	n.o.	7.72	5.89	5.99	2.10	2.45	n.a.	n.a.	n.a.	n.a.
C2(6)	-	n.o.	n.o.	8.02	6.08		2.23	2.61	n.a.	n.a.	n.a.	n.a.
G3(7)	n.o.	n.o.	n.o.	7.53	-	6.08	1.69		n.a.	n.a.	n.a.	n.a.
T4(8)	n.o.	-	-	7.58	1.83	6.19	2.16	2.47	4.81	n.a.	n.a.	n.a.

<sup>a</sup> El desplazamiento químico de los protones intercambiables se da para T=5°C.

## ANEXO C

### TABLAS DE SUMA DE CONSTANTES DE ACOPLAMIENTO.

Las constantes de acoplamiento se expresan en Hz.

*Tabla C.1.- Suma de constantes de acoplamiento para la forma dimérica de d<pCATTTCATT>.*

Residuo	$\Sigma 1'$	$\Sigma 2'$	$\Sigma 2''$
C1	14	25	21
A2	15	24	28
T3	16	27	26
T4	15	24	23

*Tabla C.2.- Suma de constantes de acoplamiento para la forma monomérica de d<pCATTTCATT>.*

Residuo	$\Sigma 1'$	$\Sigma 2'$	$\Sigma 2''$
C1	14	28	22
A2	16	29	23
T3	15	28	20
T4	15	28	20

*Tabla C.3.- Suma de constantes de acoplamiento para la forma dimérica de d<pTGCTCGCT>.*

Residuo	$\Sigma 1'$	$\Sigma 2'$	$\Sigma 2''$
T1	15	26	-
G2	16	26	22
C3	15	-	-
T4	20	-	29
C5	15	27	-

*Tabla C.4.- Suma de constantes de acoplamiento para la forma dimérica de d<pCATTTCATT>.*

Residuo	$\Sigma 1'$	$\Sigma 2'$	$\Sigma 2''$
T1	14	-	-
G2	14	26	21
C3	13	26	23
T4	14	26	24
C5	13	-	-
C7	14	-	23
T8	14	24	-

## ANEXO D

DETERMINACIÓN DE LA K PARA LOS EQUILIBRIOS *DUMBELL*-DÍMERO Y *RANDOM COIL*-DÍMERO, A PARTIR DEL CÁLCULO DE POBLACIONES EN ESPECTROS 1D DE RMN.

❖ EQUILIBRIO *DUMBELL*-DÍMERO.

*Tabla D.1.- Equilibrio dumbbell-dímero para d<pCATTCAAT> a 20mM.*

	T °K	I <sub>b</sub> /I <sub>a</sub>	K M <sup>-1</sup>
d<pCATTCAAT>	278	1.14	61
	283	0.96	47.0
	293	0.54	20.8
	298	0.41	14.4
	303	0.26	8.19
	313	0.09	2.4
	318	0.05	1.3

*Tabla D.2.- Equilibrio dumbbell-dímero para d<pTGCTCGCT> a 2.4mM.*

	T °K	I <sub>b</sub> /I <sub>a</sub>	K M <sup>-1</sup>
d<pTGCTCGCT>	278	3.80	3800.0
	283	3.44	3182.0
	293	2.75	2148.4
	303	2.15	1411.0
	308	1.92	1168.0
	313	1.38	684.2
	318	1.17	529.0
	323	0.93	374.0
	328	0.57	186.4
	333	0.29	78.0
	338	0.09	20.4

*Tabla D.3.- Equilibrio dumbbell-dímero para d<pTGCAATGCA> a 1mM.*

	T °K	I <sub>b</sub> /I <sub>a</sub>	K M <sup>-1</sup>
d<pTGCAATGCA>	283	6.44	23956.8
	288	4.81	13973.1
	293	3.47	7755.4
	298	2.01	3025.1
	303	1.48	1835.2
	308	0.94	911.8
	313	0.60	480
	318	0.37	235.45
	323	0.18	106.2
	328	0.069	36.9

❖ EQUILIBRIO RANDOM COIL-DÍMERO.

*Tabla D.4.- Equilibrio random coil-dímero para d<pCGCTCATT> a 20mM.*

	T °K	Id/Ir 2.4 mM	Id/Ir 5 mM	Id/Ir 10 mM	K M <sup>-1</sup>
d-pCATTTCATT>	278	1.70	2.27	4.56	1423.1
	283	1.25	1.73	2.43	362.9
	288	0.58	1.00	1.43	165.6
	293	0.25	0.48	0.86	84.7

*Tabla D.5.- Equilibrio random coil-dímero para d<pAGCTAGCT> a 5mM.*

	T °K	Id/Ir	K M <sup>-1</sup>
d-pAGCTAGCT>	308	3.58	1639.6
	313	1.37	324.7
	318	0.64	105.0
	323	0.4	56.0

*Tabla D.6.- Equilibrio random coil-dímero para d<pTGCATGCA> a 5 mM.*

	T °K	Id/Ir	K M <sup>-1</sup>
d-pTGCATGCA>	323	6.23	4504.3
	328	1.85	527.2
	333	1.32	306.2
	338	0.40	56.0

*Tabla D.7.- Equilibrio random coil-dímero para d<pCCGTCCGT> a 5mM. En este caso no existe una buena dispersión de señales por lo que en lugar de utilizar la relación Id/Ir se calculó el % de cada una de las especies.*

	T °K	% dímero	K M <sup>-1</sup>
d-pCCGTCCGT>	318	87	5147.9
	323	77	1455.6
	328	59	351.0
	333	30	61.2