

LES ESTRUCTURES DEL DNA

LOURDES URPI¹ I FERRAN AZORÍN²

1. *Departament d'Enginyeria Química, ETSEIB. Universitat Politècnica de Catalunya.*
Av. Diagonal, 647. 08028 Barcelona.

2. *Departament de Biologia Molecular i Cel·lular. Institut de Biologia molecular de Barcelona, CSIC.*
Jordi Girona, 18-26. Barcelona.

SUMMARY

The DNA molecule is structurally polymorphic and, depending on the precise environmental conditions, it can exist under a variety of different structural conformations which include different types of double stranded helices, either right handed (B-DNA, A-DNA, etcetera) or left-handed (Z-DNA), but also triple-stranded and tetra-stranded conformations. In this paper we will describe briefly the structural characteristics of the conformations that better reflect the high degree of structural polymorphism of the DNA molecule.

INTRODUCCIÓ

La molècula del DNA és estructuralment polimòrfica. Quan l'any 1953 James Watson i Francis Crick van proposar el famós model de la forma B del DNA, ja es coneixia l'existència d'almenys una altra forma estructural diferent del DNA. La resolució de l'estructura de la forma B del DNA es va fer principalment tenint en compte els diagrames de difracció de raigs X de fibres de DNA obtinguts per Rosalind Franklin i Maurice Wilkins. Aquests estudis mostraven que

l'estructura del DNA depenia del grau d'hidratació de les fibres i, també, del tipus de contraió emprat en la seva preparació. Aquests diagrames de difracció es podien classificar en dos grups: els de tipus B, que corresponien a la forma hidratada del DNA i que van ser els que resolgueren Watson i Crick, i els de tipus A, que s'observaven en disminuir el grau d'hidratació o en presència de determinats cations (Franklin i Gosling, 1953). La resolució estructural d'aquests úl-

TAULA 1. Polimorfisme estructural del DNA

ESTRUCTURES BICATENÀRIES

DÚPLEXS ANTIPARAL·LELS	
FORMES DEXTROGIRES	B-DNA, A-DNA C-DNA, D-DNA,...
FORMES LEVOGIRES	Z-DNA
HOMODÚPLEXS PURÍNICS	$d(GA) \cdot d(GA)$ $d(G) \cdot d(G)$ $d(A^+) \cdot d(A)$
HOMODÚPLEXS PIRIMIDÍNICS	$d(C^+) \cdot d(C)$ $d(C^+T) \cdot d(CT)$
DÚPLEXS PARAL·LELS	
	$d(A) \cdot d(T)$ $d(GA) \cdot d(GA)$

ESTRUCTURES TRICATENÀRIES

TIPUS I, Y * (R·Y)	$d(C^+T)[d(G) \cdot d(C)]$ $d(C^+T)[d(GA) \cdot d(CT)]$ $d(T)[d(A) \cdot d(T)]$
TIPUS II, R * (R·Y)	$d(G)[d(G) \cdot d(C)]$ $d(GA)[d(GA) \cdot d(CT)]$

ESTRUCTURES TETRACATENÀRIES

TÈTRAPLEXS DE GUANINES
TÈTRAPLEXS DE PIRIMIDINES

tims, que corresponen a la forma A del DNA, va trigar encara uns quants anys. Així doncs, la variabilitat conformacional del DNA era ja palesa en els primers experiments adreçats a la determinació de la seva estructura.

Avui sabem que el grau de polimorfisme estructural del DNA és enorme i que, dependent de la seva seqüència nucleotídica i de les condicions experimentals precises, la molècula de DNA pot adoptar *in vitro* un ampli ventall de conformacions estructurals diferents. La taula 1 resumeix les més característiques. En primer lloc, trobem el grup d'estructures bicatenàries que inclou les for-

mes A i B, altres conformacions estretament relacionades amb aquestes (formes C i D), però també conformacions que tenen característiques estructurals molt diferents, com pot ser el sentit de gir de la doble hèlix (forma Z), l'orientació de les cadenes polinucleotídiques (dúplexs paral·lels) o la presència d'aparellaments no canònics (dúplexs homopurínics o homopirimidínics). La molècula del DNA també pot formar estructures tricatenàries i tetracatenàries, resultat de l'associació de tres o quatre cadenes nucleotídiques respectivament. Aquestes estructures multicatenàries també poden ser de dife-

rents tipus, segons la naturalesa química de les cadenes que formen el complex i de la seva orientació relativa.

En aquest treball pretenem descriure breument els trets més característics d'algunes de les conformacions del DNA, les més representatives. No volem fer una anàlisi exhaustiva de totes i cadascuna d'elles; volem simplement destacar aquells aspectes que reflecteixen millor la variabilitat conformacional del DNA. El lector interessat trobarà en aquest treball cites bibliogràfiques que li permetran aprofundir en aquesta temàtica.

ESTRUCTURES BICATENÀRIES DEL DNA

Com ja s'ha indicat a la introducció, a partir dels estudis de difracció de fibres es van determinar les característiques de la forma B (Watson i Crick, 1953), que va ser la primera estructura bicatenària del DNA que es va conèixer. Després, es van poder interpretar els diagrames de difracció de la forma A.

El 1979 va succeir un fet important per a la determinació estructural del DNA. El grup d'Alexander Rich va descobrir una nova estructura del DNA (Wang *et al.*, 1979) que, a diferència de les formes A i B, és una

hèlix levogira. Va ser anomenada forma Z a causa de la ziga-zaga que descriu l'esquelet sucre-fosfat. Fins aquesta data tot el coneixement estructural que es tenia del DNA procedia dels estudis de difracció de fibres. La descoberta de la forma Z es va fer a partir d'un dels primers oligonucleòtids cristal·litzats com a monocristall. Això va ser possible gràcies als avenços en síntesi orgànica d'oligodesoxiribonucleòtids. La difracció de raigs X de monocristall aporta moltes més dades que la de fibres i, per tant, els resultats experimentals són molt més rics. Ara bé, la metodologia de resolució d'estructura és molt més complexa. A partir d'aquesta data es van cristal·litzar altres seqüències que van aportar més informació sobre les diverses conformacions del DNA. Així, el 1980 es va resoldre l'estructura del primer oligonucleòtid en forma B (Wing *et al.*, 1980), i el 1981, la del primer oligonucleòtid en forma A (Shakke *et al.*, 1981). Moltes més estructures d'oligonucleòtids s'han estudiat per difracció de raigs X de monocristall en el període que va des de principis dels vuitanta fins a l'actualitat. Més recentment, la ressonància magnètica nuclear (RMN) també s'ha mostrat molt eficaç per a l'estudi estructural del DNA.

Com altres tipus de molècules, les coordenades de les estructures d'oligonucleò-

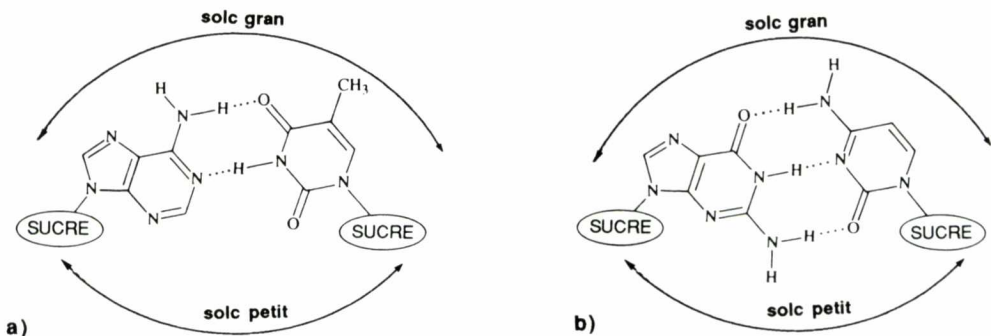


FIGURA 1. Interaccions de pont d'hidrogen de tipus Watson-Crick dels parells de bases A-T (a) i C-G (b). També es mostra la situació dels dos solcs (modificat de Calladine i Drew, 1997).

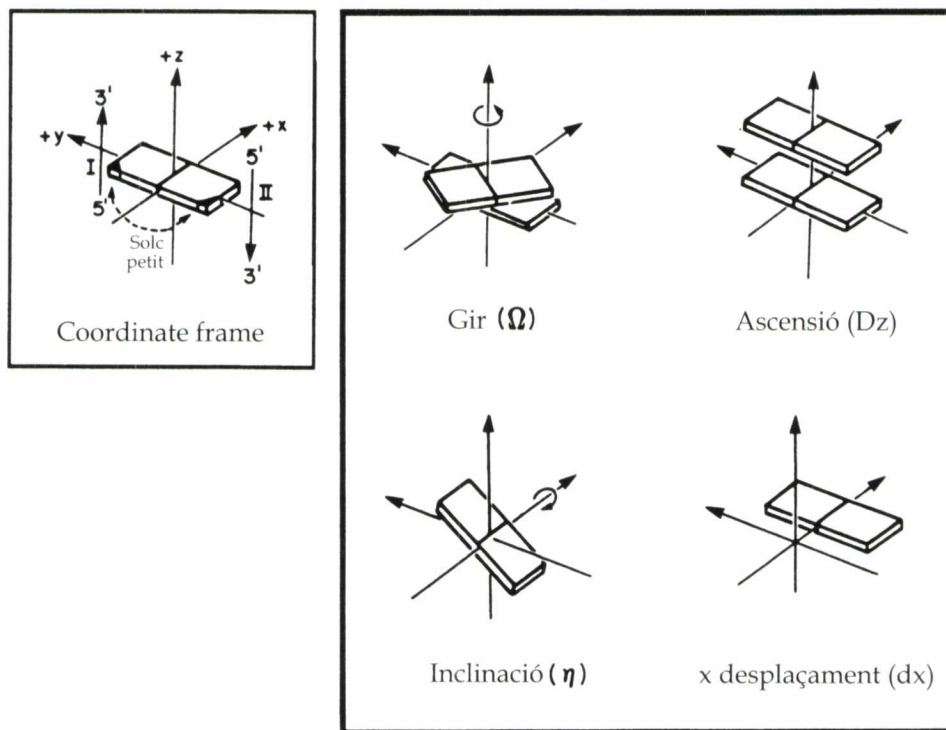


FIGURA 2. Representació esquemàtica d'una doble hèlix per definir alguns dels paràmetres de rotació i de translació que descriuen la conformació del DNA (adaptat de l'EMBO Workshop, 1989).

tids resoltes per cristal·lografia o per RMN, es disposen en un banc de dades. Aquest banc s'anomena «Nucleic Acid Database» (NDB), i és públicament i lliurement accessible per Internet a l'adreça <http://ndbserver.rutgers.edu>. D'aquesta manera es pot disposar dels detalls estructurals d'aquestes macromolècules. També així s'obre la possibilitat de comparar estructures relacionades, com també de fer estudis estadístics, donat l'augment important d'estructures noves.

Paràmetres conformacionals de les estructures bicatenàries

El fet comú de les estructures bicatenàries és que les dues cadenes nucleotídiques descriuen una hèlix. La forma de la molècu-

la és cilíndrica amb les bases nitrogenades a l'interior i l'esquelet sucre-fosfat a la perifèria. Les dues cadenes de DNA interaccionen a través de ponts d'hidrogen específics entre les bases nitrogenades. A la majoria de conformacions del DNA bicatenari, l'especificitat entre bases consisteix en el fet que una determinada base purínica interacciona amb una determinada base pirimidínica (l'adenina amb la timina, i la guanina amb la citosina), com mostra la figura 1. Aquests aparellaments de bases van ser enunciats per primera vegada per Watson i Crick, en el seu model de la forma B, però els presenten també la forma A, la forma Z i altres conformacions. Un altre tret comú de les estructures bicatenàries és que la immensa majoria té les dues cadenes en orientació $5' \rightarrow 3'$ antiparal·lela. I també, que presenten

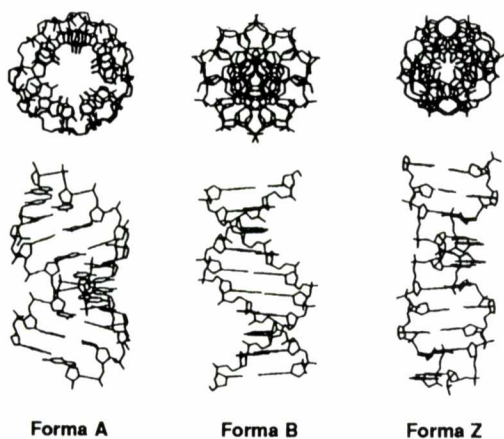


FIGURA 3. Dues imatges perpendiculars de les tres conformacions bicatenàries més conegudes del DNA: A, B i Z (modificat de Chandrasekaran *et al.*, 1989).

dos solcs característics (solc gran i solc petit) perquè els enllaços glicosídics de cada parell de bases es troben tots dos a la mateixa banda (figura 1).

Els paràmetres més bàsics per descriure la conformació del DNA són: els angles de torsió de l'esquelet sucre-fosfat, els possibles plegaments de la 2'-desoxiribosa i la conformació de l'enllaç glicosídic (enllaç entre la base i el sucre). La definició d'aquests paràmetres es pot trobar en obres més extenses (Saenger, 1984; Blackburn i Gait, 1996; Calladine i Drew, 1997).

Pel fet de ser una hèlix, les estructures bicatenàries del DNA també es poden descriure amb els paràmetres típics de les estructures helicoidals: el pas de rosca (P), el nombre de nucleòtids per pas de rosca (n) i l'angle de gir d'un nucleòtid al següent (Ω). $\Omega = 360^\circ/n$.

Hi ha una sèrie de detalls de les estructures bicatenàries del DNA que fan necessària la definició de més paràmetres (figura 2). Per exemple, els parells de bases no estan centrats a l'eix de l'hèlix, sinó que estan desplaçats una certa distància; els plans dels parells de bases no són exactament perpen-

diculars a l'eix de l'hèlix; etc. Vegeu la taula 2 i les figures 2 i 3.

Característiques estructurals de les formes dextrogires: A, B, C i D

A la conformació B del DNA, les bases s'apilen en una disposició perpendicular a l'eix de l'hèlix, el qual travessa les bases (figura 3). El plegament dels sucres és C_2' -endo, cosa que estableix la distància entre fosfats consecutius d'una mateixa cadena en 7,0 Å (més gran que a la forma A). Una conseqüència d'aquesta conformació global és que els dos solcs tenen una profunditat semblant i una amplada diferent (vegeu la taula 2). La forma B té 10 parells de bases per volta d'hèlix, amb un pas de rosca de 34 Å i, per tant, una ascensió per parell de bases de 3,4 Å. Ja en els estudis de difracció de fibres es van posar de manifest altres conformacions anomenades C i D, que tradicionalment s'han considerat relacionades amb la forma B; de manera que amb el terme de «família B» es fa referència a les tres formes (B, C i D). Les formes C i D, però, presenten paràmetres conformacionals molt diferents de la forma B. Les formes C i D tenen menys importància que la forma B i només s'obtenen en condicions especials. Tota la informació d'aquestes conformacions prové dels estudis de difracció de fibres, ja que no s'ha cristal·litzat cap oligonucleòtid en aquestes conformacions, ni tampoc no es tenen dades de RMN. Això fa que la informació sigui més pobre. La forma C s'obté amb sals de liti a humitat relativament baixa. Aquesta conformació té 28 parells de bases en tres voltes d'hèlix. La forma D es presenta en seqüències A-T alternants; la característica principal és el sobrecargolament que presenta l'hèlix: 8 parells de bases per volta.

Les característiques estructurals de la forma A són molt diferents de les de la família B. Les bases estan desplaçades 4,5 Å de

TAULA 2. Paràmetres característics de les principals estructures bicatenàries del DNA

PARÀMETRES	FORMA A	FORMA B	FORMA Z
sentit de l'hèlix	dextrogira	dextrogira	levogira
nre. nucleòtids /volta	11	10	12
angle de gir / p.b. (°)	32,7	36	alternant -9 / -51
desplaçament/pb (Å)	4,5	-0,2 a -1,8	-2 a -3
ascensió / p.b. (°)	2,56	3,3 a 3,4	3,7
inclinació / p.b. (°)	20	-6	-7
plegament de la 2'-desoxiribosa	C _{3'} -endo	C _{2'} -endo	alternant C _{2'} -endo / C _{3'} -endo
ampl. solc gran (Å)	2,7	11,7	8,8
ampl. solc petit (Å)	11,0	5,7	2,0
fond. solc gran (Å)	13,5	8,8	3,7
fond. solc petit (Å)	2,8	7,5	13,8
conformació de l'enllaç glicosídic	<i>anti</i>	<i>anti</i>	alternant <i>anti</i> / <i>syn</i>

Nota de la taula 2

p.b., de parell de bases

ampl.: amplada

fond.: fondària

La denominació «solc gran» i «solc petit» és d'acord amb les mides dels dos solcs a la forma B. En canviar de forma es manté la denominació, però els adjectius «petit» i «gran» no concorden sempre amb les mides relatives dels solcs.

l'eix de l'hèlix i això crea un forat al centre de l'hèlix de 3 Å de diàmetre (figura 3). Una volta d'hèlix conté 11 parells de bases i el pas de rosca és de 28 Å, cosa que proporciona una ascensió per parell de bases de 2,56 Å. Per tal de mantenir una separació normal entre parells de bases consecutius amb una bona interacció de van der Waals (3,4 Å de separació), les bases s'apilen amb un angle d'inclinació de 20°. El plegament dels sucres és C_{3'}-endo, la qual cosa dóna lloc a una separació de 5,9 Å entre dos fosfats consecutius al llarg d'una cadena. Un resultat d'aquestes característiques estructurals és que els dos solcs són molt diferents en amplada i fondària (taula 2).

Característiques estructurals de les formes levogires: Z-DNA

La característica principal de la forma Z és que el sentit de gir de l'hèlix és cap a l'esquerra (levogira). Al llarg de cada cadena nucleotídica hi ha una alternança de la conformació de l'enllaç glicosídic dels nucleòtids que se succeeixen: *anti* – *syn*. Una base en conformació *anti* (normalment una pirimidina) s'aparella per ponts d'hidrogen amb una base en conformació *syn* (normalment una purina). Per aquest motiu els dos sucres d'un parell de bases tenen una orientació a l'espai molt diferent, com es mostra a la figura 4. A causa de l'alternança d'en-

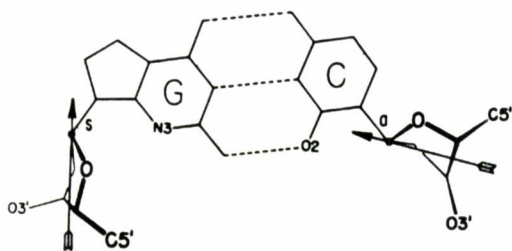


FIGURA 4. Conformació del parell de bases G-C a la forma Z. Les fletxes indiquen la diferent orientació dels sucres, que és deguda a la diferent conformació dels dos enllaços glicosídics: *s* (*syn*) per a la 2'-desoxiguanosina, *a* (*anti*) per a la 2'-desoxicitidina (modificat de Saenger, 1984).

llaços glicosídics, que produeix l'alternança d'orientacions dels sucres al llarg de la cadena nucleotídica, la línia que uneix els fosfats descriu una ziga-zaga, d'on prové el nom de Z. L'alternança en la conformació de l'enllaç glicosídic va unida a l'alternança en el plegament del sucre: un enllaç glicosídic *anti* correspon a un sucre C_2' -*endo* i un enllaç glicosídic *syn*, a un sucre C_3' -*endo*. Com que la forma Z suposa una alternança d'enllaços glicosídics *syn* i *anti*, i les purines poden adoptar més fàcilment la conformació *syn* que les pirimidines, la forma Z es presenta bàsicament en seqüències alternants pirimidina-purina (i amb més contingut de parells C-G que A-T).

Un factor destacable de la forma Z és el diferent valor de l'angle de gir per a la seqüència d(CpG) (-9°) i per a la seqüència d(GpC) (-51°); aquesta diferència no es produeix a les hèlixs dextrogires. Una conseqüència d'aquest fet és que la distància que separa dos fosfats consecutius al llarg d'una cadena depèn del nucleòsid que hi hagi al mig; si té l'enllaç glicosídic *syn* és 7,5 Å i si el té *anti* és 6,2 Å, en contraposició amb les formes dextrogires on tots els fosfats són equivalents i estan separats per la mateixa distància.

Les particularitats de la forma Z respecte a les formes dextrogires del DNA donen lloc

a una mida global diferent: el DNA Z és una hèlix més prima, de tal manera que el diàmetre de la base fa 18 Å en comparació dels 20 Å de les formes dextrogires (figura 3). També els fosfats de les dues cadenes són més a prop: 8 Å, en comparació amb els 11,7 Å que separen els fosfats de la forma B. Això explica que la forma Z requereixi la presència de contraions (Mg^{2+} , Na^+ , $[Co(NH_3)_6]^{3+}$).

Factors que afecten la conformació de les estructures bicatenàries del DNA

Els factors que determinen la conformació del DNA són la seqüència de bases i les condicions del medi en el qual es troba. Per Raman, ressonància magnètica nuclear (RMN), infraroig i dicroisme circular, s'ha posat de manifest que tots els oligòmers de DNA tenen forma B en solució aquosa diluïda a pH 7 i una concentració de 0,1 – 1,0 M de NaCl. Per tal que un oligonucleòtid adopti una conformació diferent de la B cal fer créixer l'activitat de l'aigua que conté. Això s'aconsegueix augmentant-ne la concentració de sal, addicionant-hi cations de metalls pesants o addicionant-hi alcohol.

Inicialment els estudis de difracció de fibres de DNA, i posteriorment la difracció de monocristalls d'oligonucleòtid, han posat de manifest la influència de la seqüència de nucleòtids en la conformació. Les correlacions més clares entre seqüència i conformació són:

1. Gs seguides (i Cs seguides) afavoreixen la forma A.
2. L'alternança de bases CG afavoreix la forma Z.
3. L'abundància de parells A-T afavoreix la forma B.
4. Les seqüències generals que no compleixen els punts anteriors 1 i 2 presenten normalment la forma B.

Cal mencionar un fet curiós observat en

dues seqüències relacionades, que és la coexistència de les formes A i B en un mateix cristall. O sigui, que una mateixa seqüència en un medi determinat pot existir en dues conformacions diferents. En concret, això s'ha observat per les seqüències d(GG^{Br}UA^{Br}UACC) i d(GGTATACC) (Kennard i Hunter, 1989). En aquests casos l'explicació cal buscar-la en el tipus d'empaquetament de les molècules al cristall, que és un cas particular de com les condicions del medi afecten la conformació. La xarxa cristal·lina és formada per molècules de forma A que es disposen al voltant dels eixos senaris helicoidals de la xarxa (grup espacial P6₁) i originen canals de 26 Å de diàmetre. En aquests canals se situen les molècules de forma B. La coexistència de dues conformacions de la mateixa seqüència en el mateix cristall indica que les dues conformacions són properes en energia.

La hidratació és un aspecte de les condicions del medi i, per tant, afecta la conformació del DNA. Com ja s'ha esmentat, la forma B s'obté a alta hidratació i les formes A i Z, a hidratació més baixa. En la resolució de moltes estructures d'oligonucleòtids per cristal·lografia s'ha posat de manifest que hi ha molècules d'aigua, ions i molècules petites, que interaccionen amb el DNA i, per tant, són factors importants en l'estructura del DNA. Si el solvent (aigua majoritàriament) no està ordenat, no apareixen pics clars en els mapes de densitat electrònica. L'absència de pics de solvent generalment indica que l'aigua està desordenada. La falta d'aigua prop dels àtoms carregats o electronegatius seria una situació desfavorable energèticament. En les tres conformacions principals del DNA s'han localitzat molècules d'aigua en posicions característiques. Així, els dodecàmers en forma B que tenen a la zona central la seqüència AATT presenten en aquesta zona el solc petit estret i ocupat per una fila de molècules d'aigua que interaccio-

nen d'una manera específica amb el DNA. En canvi, altres seqüències en forma B presenten el solc petit més ample, de manera que allotja dues files de molècules d'aigua, també amb interaccions específiques. La forma A té poques aigües ordenades al solc petit, contràriament a la forma B. Aquest fet ha de ser degut a la diferència de mida entre els solcs de les dues conformacions. De manera que ambdues formes presenten hidratat el solc que és més estret, o sigui, el solc petit a la forma B i el solc gran a la forma A. A la forma Z també s'ha trobat una hidratació característica, formada per dos tipus d'aigües situades al solc petit. Una aigua estabilitza la conformació *syn* de la 2'-desoxiguanosina, ja que forma un pont d'hidrogen amb el N2 de la guanina i un altre amb un oxigen del fosfat del mateix nucleòtid. L'altra aigua forma una espina al llarg del solc perquè interacciona amb l'O2 de citosines de parells de bases consecutius. Aquest esquema d'hidratació està limitat als parells C-G, ja que en els parells T-A no s'ha observat. L'adenina no té el grup N2, per tant, la interacció en què participa aquest grup no és possible. O sigui, que la hidratació estabilitza més un parell C-G en forma Z que un parell T-A en forma Z (vegeu les figures 1 i 4).

Variabilitat dels paràmetres conformationals de les estructures bicatenàries del DNA

La resolució d'estructures per difracció de raigs X de monocristall ha posat de manifest l'existència d'un important grau de diversitat en els valors precisos de les variables conformationals de les estructures bicatenàries del DNA, en el sentit que un fragment de DNA no presenta les mateixes variables conformationals per cada nucleòtid. De les dades de difracció de fibres només s'obté la mitjana de les variables; així doncs, la diversitat conformational no es

manifesta. En aquest sentit, les dades de la taula 2 són valors que provenen d'estudis de fibres i que normalment estan d'acord amb els valors mitjans de les variables dels oligonucleòtids. Tanmateix, s'ha observat una clara discrepància entre els valors mitjans de les variables en octanucleòtids cristal·litzats en forma A (Verdaguer *et al.*, 1991) i els valors obtinguts de fibres.

Un exemple de la variabilitat de la forma B és que alguns nucleòtids presenten els sucres en conformació C_3' -endo, quan a la forma B el plegament característic dels sucres és C_2' -endo. Els nucleòtids de pirimidines (C i T) tenen una tendència més acusada a adoptar la conformació C_3' -endo que els nucleòtids de purines (G i A); i també influeix el tipus de nucleòtids contigus (Subirana *et al.*, 1995).

Amb les dades experimentals de què es disposa, es manifesta que no hi ha cap parell de nucleòtids consecutius que sigui totalment rígid. Seqüències similars poden adoptar diferents conformacions segons les interaccions que estableixin amb molècules veïnes. Tanmateix, estudis estadístics realitzats amb estructures en forma B (Subirana i Faria, 1997) mostren que la variabilitat intrínseca de cada parell de nucleòtids és diferent, i que les seqüències flanquejants poden tenir només una petita influència. Només per al pas de nucleòtids YR (pirimidina-purina), sobretot per als passos CG i CA/TG, hi ha una influència acusada dels nucleòtids flanquejants. La qual cosa equival a dir que els passos CG i CA/TG tenen una gran variabilitat conformacional, i la resta de passos de dinucleòtids són més rígids. Encara es disposa de relativament poques dades per fer un estudi estadístic complet de les tendències conformacionals de les seqüències. Hi ha passos de dinucleòtids dels quals no es té prou informació per establir les seves característiques. El pas TA n'és un exemple, té la característica de tenir un

angle de gir més elevat de la mitjana de la forma B, però en contraposició hi ha una estructura amb un valor acusadament més petit (Urpí *et al.*, 1996).

La forma A també mostra una considerable variabilitat conformacional. Un estudi recent (Fernandez *et al.*, 1997) posa de manifest que una mateixa seqüència pot variar molt la seva conformació segons les condicions del medi. Aquest estudi s'ha realitzat amb l'oligonucleòtid d(pCCCGCGGG), que ha estat cristal·litzat en cinc condicions diferents, i en cada cas la conformació varia, tot i mantenir-se sempre dins la forma A. S'observa un canvi gradual d'alguns paràmetres en les cinc estructures, des de l'estructura més hidratada o d'empaquetament més obert a l'estructura menys hidratada o d'empaquetament més compacte. Així, s'observa una variació molt important en l'amplada del solc gran; si es considera la distància P-P a través del solc gran, per a l'estructura més oberta és 5,7 Å i per a la més compacta és 0,7 Å. També la curvatura de l'eix de l'hèlix canvia (de 8,4° a 20,8°) de l'estructura més oberta a la més compacta.

En alguns casos la variabilitat conformacional és tan gran que un dúplex en forma B pot presentar en alguna zona paràmetres característics de la forma A. Aquest és el cas de l'hexàmer d(CTCGAG) (Wahl *et al.*, 1996), que té el plegament de les quatre 2'-desoxiriboses terminals en C_3' -endo, fet que els autors interpreten que està relacionat amb la facilitat de transició a la forma A.

De la forma Z, s'ha dit habitualment que és més rígida (menys variable) que les formes A i B. Aquesta afirmació prové del fet que la majoria d'oligonucleòtids cristal·litzats en forma Z presenten els mateixos tipus de contactes entre molècules, cosa que afecta els paràmetres conformacionals. Tanmateix, un estudi recent d'una mateixa seqüència cristal·litzada en dues condicions diferents en forma Z (Malinina *et al.*, 1998) i compara-

da amb la mateixa seqüència cristal·litzada anteriorment també en forma Z però en condicions diferents, mostra que la forma Z és tan variable com les formes A o B.

Estructures amb cadenes nucleotídiques que formen diversos dúplexs

Els últims anys s'han determinat estructures d'una sèrie d'oligonucleòtids que han posat de manifest l'existència de conformacions diferents de les esmentades anteriorment. En diversos casos s'han relacionat les noves estructures amb la seva funcionalitat, però de moment són només hipòtesis. Les seves seqüències són: d(CCGCGG), d(AGGCATGCCT) i d(GCGCGCG). En cap dels casos no era previsible a partir de la seqüència l'estructura que s'ha observat. El fet comú en aquestes estructures és que la part central presenta una conformació coneguda (A o Z), però als extrems de les seqüències la conformació és diferent de les conformacions ja establertes. Un altre fet característic d'aquestes estructures és que la mateixa cadena nucleotídica forma més d'un dúplex (dos o tres dúplexs). Per aquest motiu s'han relacionat aquestes estructures amb funcions com la replicació i la transcripció.

L'oligonucleòtid d(CCGCGG) forma un dúplex a la zona central que s'obre pels extrems (Malinina *et al.*, 1994). El tetràmer central (CGCG) forma un dúplex en forma Z, i les citosines 5' de les dues cadenes del dúplex no interaccionen amb les guanines corresponents de la cadena complementària, sinó que s'orienten cap a l'exterior del dúplex i interaccionen amb les guanines corresponents d'un altre dúplex. De manera semblant, les guanines terminals 3' interaccionen amb les citosines corresponents del mateix dúplex veí (figura 5 de l'annex 5). Es produeix, doncs, una associació de dos dúplexs en conformació Z al centre que bescan-

vien les seves citosines terminals. Totes les interaccions dels parells de bases C·G són del tipus Watson-Crick. Aquesta estructura inusual als extrems de la seqüència respon al fet que la seqüència no és alternant C·G als extrems, cosa que dificulta l'adopció de la conformació Z. La macroestructura formada pels dos dúplexs associats, a més de mantenir-se per interaccions de ponts d'hidrogen dels quatre parells de bases dels extrems, també està estabilitzada per un ió sodi que se situa a la part central entre els dos dúplexs i que es coordina a l'O6 d'una guanina de cada dúplex. (Aquest ió no s'ha representat a la figura 5 de l'annex 5).

Una altra estructura que té similituds amb l'anterior ha estat determinada posteriorment (Nunn i Neidle, 1996). Es tracta de la seqüència d(AGGCATGCCT), que presenta l'octàmer central en forma A i els nucleòtids 5'-A i 3'-T orientats cap a l'exterior del dúplex i interaccionant amb els nucleòtids complementaris d'un dúplex veí. Malgrat la diferent conformació de la zona central [Z per d(CCGCGG) i A per d(AGGCATGCCT)], el tipus d'interacció dels nucleòtids terminals és anàleg en aquestes dues estructures, com també ho és la macroestructura de quatre cadenes nucleotídiques que es forma.

La seqüència d(GCGCGCG) també té una estructura amb analogies amb les estructures anteriors, però té particularitats que la diferencien, que provenen de la longitud de la seqüència. Aquest oligonucleòtid no és totalment autocomplementari, sinó que té una base extra, només sis de les set bases poden formar un dúplex amb aparellaments complementaris C·G. L'estructura determinada (Pan *et al.*, 1997) mostra que la 5'-G és la base que no forma aparellaments C·G i que la resta de la seqüència forma un dúplex en forma Z. Les guanines 5' de les dues cadenes del dúplex (G1 i G8) adopten la conformació *anti* / C_{2'}-*endo*, que s'allunya

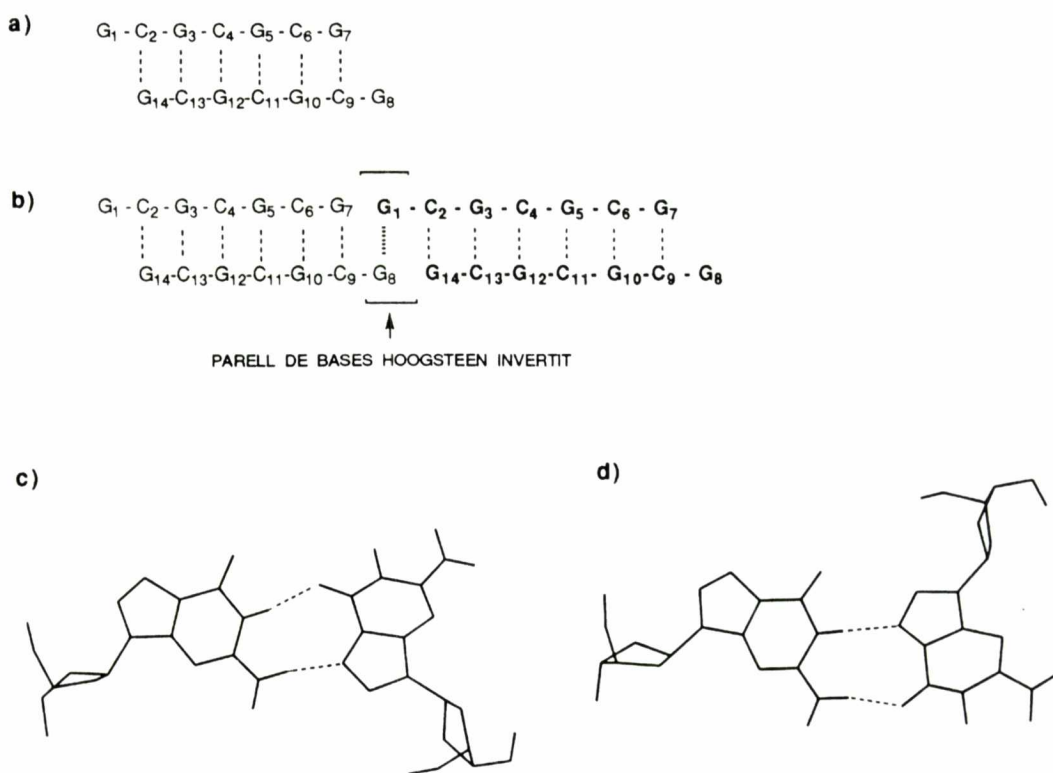


FIGURA 6. a) Numeració dels nucleòtids que formen un dúplex a l'estructura de l'heptàmer d(GCGCGCG); b) interaccions entre dos dúplex veïns (un dúplex representat en gruix normal i l'altre, en negreta) a través d'un parell G-G; c) i d) parell de bases G-G amb interaccions de pont d'hidrogen de tipus Hoogsteen i de tipus Hoogsteen invertit, respectivament.

totalment de la que tindrien si adoptessin la forma Z. Aquestes guanines (G1 i G8) s'aparellen amb altres guanines de l'extrem 5' d'un dúplex veï mitjançant interaccions del tipus Hoogsteen invertit (figura 6). En aquest cas es forma una macroestructura infinita de dúplex mitjançant ponts d'hidrogen intermoleculars G-G.

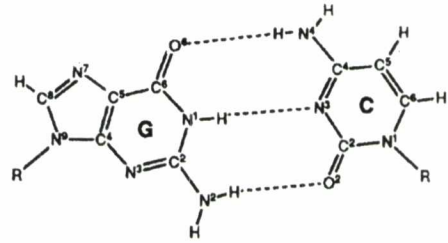
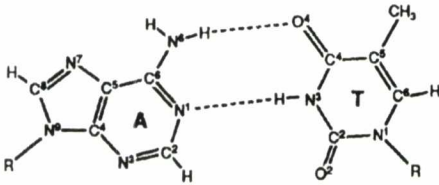
ESTRUCTURES TRICATENÀRIES

Les estructures tricatenàries o tríplex resulten de l'associació d'un fragment de DNA bicatenari amb una tercera cadena de DNA (figura 7 de l'annex 5). Tant la tercera cadena com el fragment bicatenari poden

ser de naturalesa ribonucleotídica (RNA) i la formació de tríplex de RNA o de tríplex híbrids DNA-RNA està àmpliament demostrada. De fet, la primera estructura tricatenària descrita, l'any 1957, corresponia a un tríplex de RNA (Felsenfeld *et al.*, 1957). En aquest treball tractarem només el cas dels tríplex de DNA que ha estat motiu d'excel·lents revisions recentment (Cheng i Pettitt, 1992; Radhakrishnan i Patel, 1994; Mirkin i Frank-Kamenetskii, 1994; Bernués i Azorín, 1995; Frank-Kamenetskii i Mirkin, 1995).

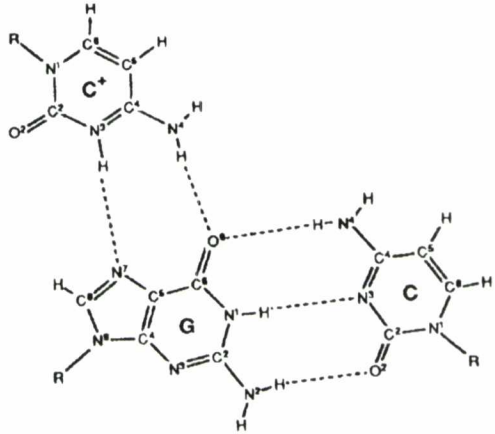
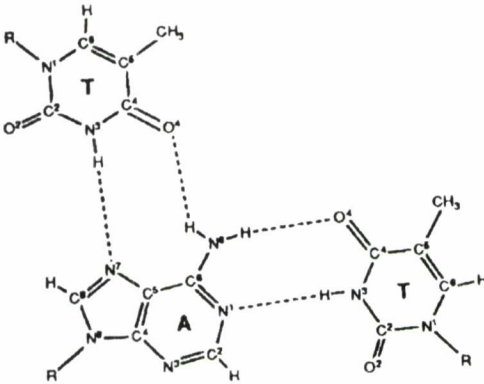
El nostre coneixement de les característiques estructurals dels tríplex prové principalment d'estudis de RMN que han permès la determinació estructural de diversos ti-

A



B

TIPUS I



TIPUS II

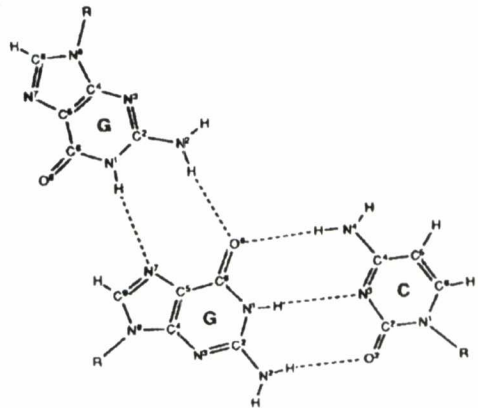
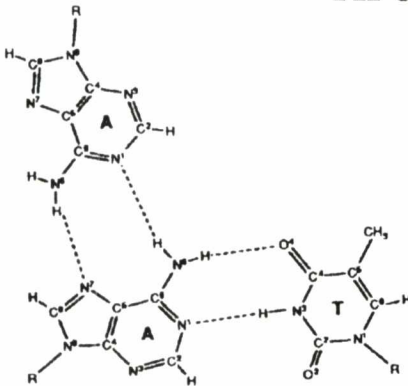


FIGURA 8. a) Representació esquemàtica dels dos aparellaments Watson-Crick canònics i b) de les quatre triades més estables.

pus de tríplexs. En aquest cas, la cristal·lografia de raigs X ha resultat menys informativa donat que, fins ara, només ha estat possible cristal·litzar complexos que, ja sigui per la seva naturalesa química (Betts *et al.*, 1995) o per la seva mateixa estructura (Van Meervelt *et al.*, 1995), representen casos molt especials.

En un tríplex la tercera cadena recorre el solc gran del dúplex (figura 7 de l'annex 5) establint interaccions de ponts d'hidrogen amb les bases del dúplex. Aquestes interaccions donen lloc a la formació de tríades. Als aparellaments canònics G-C i A-T, les bases puríniques mantenen lliure un major nombre de grups capaços d'establir interaccions per ponts d'hidrogen, en comparació de les bases pirimidíniques (figura 8A). Aquest fet determina una de les característiques més importants dels tríplexs: la tercera cadena s'associa preferentment amb les purines del dúplex. L'estabilitat dels tríplexs depèn directament de la de llurs tríades constituents. Les tríades més estables són les T(A-T), C+(G-C), A(A-T) i G(G-C) (figura 8B) que poden ser considerades com a canòniques. En determinats casos s'ha pogut demostrar la formació d'altres tipus de tríades menys estables, com les A+(G-C), A(G-C) o T(C-G), però la seva formació sempre té un efecte inestabilitzador, en comparació de tríplexs formats només per tríades canòniques. La mateixa estructura química de les tríades determina que la formació de tríplexs estigui afavorida en seqüències que presentin una marcada asimetria en la distribució de purines i pirimidines entre ambdues cadenes. Seqüències que presenten alternança de purines i pirimidines a la mateixa cadena formen tríplexs amb dificultat, atès que aquests tríplexs hauran d'incloure tríades menys estables o que, com veurem més endavant, l'associació de la tercera cadena amb el dúplex haurà de canviar de cadena, i de motiu de reconeixement, a cada alternança purina-pirimidina.

Segons la naturalesa química de la tercera cadena, els tríplexs es poden classificar en dos tipus (figura 8B). Els tríplexs de tipus I tenen la tercera cadena de tipus polipirimidínic i estan estabilitzats per la formació de tríades T(A-T) i C+(G-C). En aquests tríplexs, les bases de la tercera cadena adopten la conformació *anti* i els aparellaments que s'estableixen amb les bases del dúplex són de tipus Hoogsten. Això determina que l'orientació de la tercera cadena sigui paral·lela amb relació a la cadena purínica del dúplex. Contràriament, els tríplexs de tipus II tenen la tercera cadena polipurínica i estan estabilitzats per la formació de les tríades A(A-T) i G(G-C). En aquest cas l'orientació de la tercera cadena és antiparal·lela amb relació a la cadena purínica del dúplex, donat que els aparellaments són del tipus Hoogsten invertit. Així doncs, l'orientació de la tercera cadena respecte al dúplex serà diferent depenent de si s'empra el motiu pirimidínic (tipus I) o purínic (tipus II). Finalment cal indicar que els tríplexs de tipus I són més estables a pHs àcids, donat el requeriment de protonació de la citosina ($pK_a = 4,5$) per a la formació de les tríades C+(G-C). En canvi, els tríplexs de tipus II són estables en un ampli marge de pH, i són estabilitzats per cations divalents, principalment magnesi, però, en alguns casos, també per certs cations metàl·lics de transició (zinc, cadmi, cobalt).

La formació de tríplexs constitueix un mecanisme per al reconeixement específic del DNA amb àcids nucleics de cadena senzilla, com oligonucleòtids sintètics. En principi, per a qualsevol seqüència de DNA es pot dissenyar un oligonucleòtid sintètic (TFO) que la reconegui específicament mitjançant la formació d'una estructura tricaenària. El codi de reconeixement ve determinat pels diferents tipus de tríades i té les limitacions abans esmentades, la principal de les quals fa referència a l'efecte inestabi-

litzador de les alternances purina-pirimidina. La figura 9 resumeix esquemàticament aquesta problemàtica. Per tal de reconèixer una seqüència purínica que presenta diverses interrupcions de pirimidines, es poden formar tríades menys estables amb les pirimidines (figura 9a), o canviar de cadena a cada interrupció. Ara bé, aquesta última solució, que és en principi la menys dolenta, requereix també canviar de motiu de reconeixement, donats els requeriments d'orientació de la tercera cadena amb relació al dúplex (figura 9b). De manera alternativa, es pot canviar l'orientació 5' → 3' de la tercera cadena (figura 9c) però la síntesi d'oligonucleòtids amb canvis d'orientació de la cadena no és encara rutinària. Tot i aquestes limitacions, ha estat possible demostrar *in vivo* el reconeixement específic de seqüències concretes de DNA amb oligonucleòtids sintètics formadors de tríplexs o TFO. Aquesta estratègia s'ha demostrat útil en diverses aplicacions. TFO conjugats amb Fe-EDTA s'han emprat *in vitro* com a nucleases amb alta especificitat de seqüència o, conjugats amb psolaré, com a sistema de mutagènesi dirigida *in vivo*. Els TFO també s'han demostrat eficaços com a inhibidors de l'ini-

ci i de l'elongació de la transcripció tant *in vitro* com *in vivo*.

Fins aquí ens hem referit als tríplexs que anomenem intermoleculars, en els quals la tercera cadena i el dúplex pertanyen a molècules de DNA diferents. Els tríplexs també poden ser intramoleculars si la tercera cadena pertany a la mateixa molècula de DNA que el dúplex sobre el qual es forma l'estructura tricatènaria. Les característiques fonamentals de la formació i l'estabilització dels tríplexs intramoleculars són idèntiques a les esmentades abans per als tríplexs intermoleculars. Ara bé, existeixen algunes petites diferències que resumirem a continuació. La formació dels tríplexs intramoleculars requereix la desnaturalització local de la regió de la molècula de DNA que aporta la tercera cadena. Aquest procés de desnaturalització és energèticament desfavorable i, en conseqüència, la formació de tríplexs intramoleculars es veu afavorida per factors que el facilitin, com per exemple increments de la superhelicitat negativa del DNA o presència de determinats cations metàl·lics de transició (zinc, cadmi, cobalt). Un cop format el tríplex intramolecular, la regió complementària a la tercera cadena ha de romandre

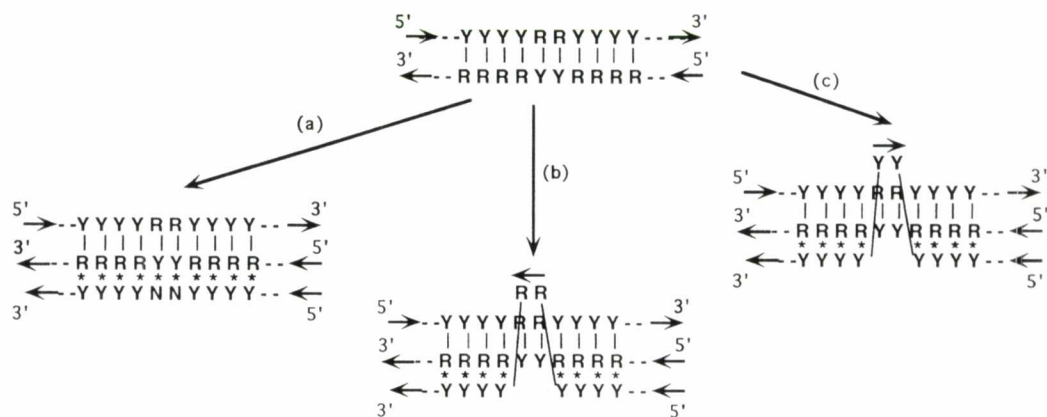


FIGURA 9. Les alternances purina(R)-pirimidina(Y) tenen un efecte inestabilitzador sobre la formació d'estructures tricatènaries. A cada alternança: a) es pot formar una triada menys estable, b) o canviar de cadena i de motiu de reconeixement, c) o canviar l'orientació de la tercera cadena. R = purina, Y = pirimidina, N = qualsevol base.

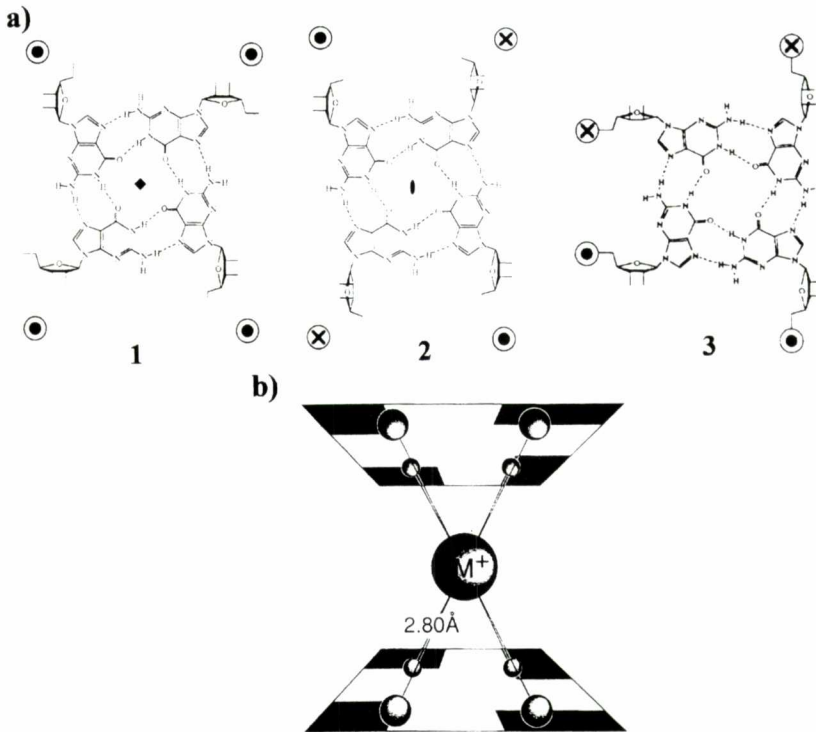


FIGURA 10. *a)* Les quatre cadenes que formen les tètredes de guanines poden ser paral·leles (1) o antiparal·leles (2 i 3). Els cercles amb creus o punts indiquen la diferent orientació de les cadenes; *b)* l'espai entre dues tètredes consecutives té unes dimensions ideals per a la coordinació del potassi.

necessàriament desnaturalitzada, en forma de cadena senzilla.

ESTRUCTURES TETRACATENÀRIES

Les estructures tetracatenàries, o tètreples, són el resultat de l'associació de quatre cadenes de DNA. La formació de tètreples es veu extraordinàriament afavorida en seqüències enriquides en residus de guanina, tot i que també s'ha observat en altres tipus de seqüències (Sundquist, 1991, per a una revisió sobre aquesta temàtica).

La formació de tètreples de guanina es fonamenta en la capacitat d'aquests residus de formar tètredes on quatre guanines s'associen per mitjà d'interaccions de ponts

d'hidrogen (figura 10). Contràriament al cas dels tríplexs de DNA, la informació estructural sobre els tètreples de guanines, obtinguda tant per RMN com per cristal·lografia de raigs X, és molt abundant. La formació d'aquests tètreples requereix la presència de, com a mínim, tres tètredes consecutives. Les tètredes de guanines són estructures planars que, quan són consecutives, estableixen fortes interaccions d'apilament. Aquestes forces d'apilament contribueixen de manera molt important a l'estabilitat dels tètreples, que és alta. Els tètreples de guanines tenen temperatures de fusió similars, o fins i tot superiors, a les corresponents a dúplexs similars en la conformació B. Es poden distingir dos tipus de tètredes segons que les quatre cadenes tinguin la mateixa

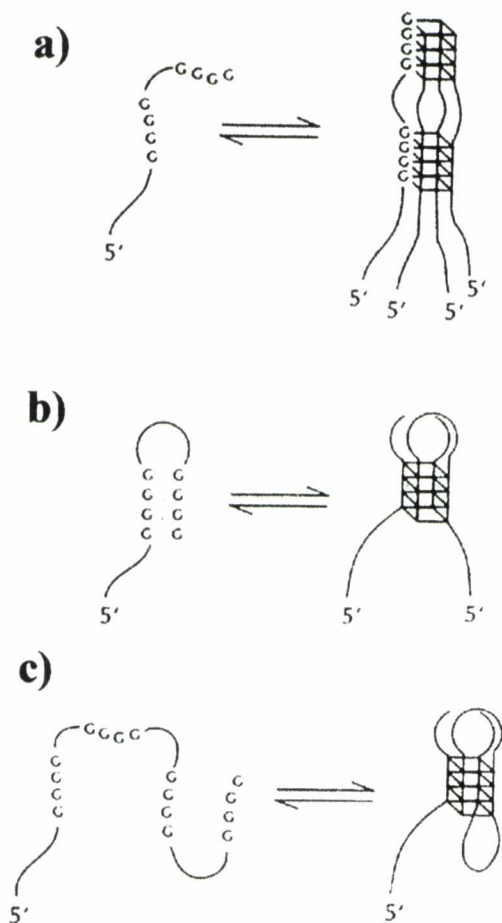


FIGURA 11. Els tètreples poden ser tetramolèculars (a), bi-molèculars (b) o unimolèculars (c).

orientació paral·lela o bé siguin antiparal·leles dues a dues (figura 10a). Aquestes darres poden classificar-se també en dos tipus segons que cadascuna de les cadenes tingui una orientació antiparal·lela respecte a les veïnes (tipus 2) o sigui paral·lela a alguna d'elles (tipus 3). Als tètreples paral·lels les quatre guanines adopten la conformació *anti*, però als antiparal·lels s'alternen les conformacions *anti* i *syn*. En el cas dels tètreples antiparal·lels, les diferents possibilitats d'alternança en l'orientació relativa de les cadenes i la conformació de les bases

dóna lloc a fins a deu possibles formes estructuralment isomèriques.

Els cations tenen un paper molt important en l'estabilització dels tètreples. En general, els cations monovalents tenen un major efecte estabilitzador que els de valència superior i, de tots, el potassi és el que té l'efecte més important. L'efecte estabilitzador del potassi es deu, d'una banda, a la neutralització de les càrregues negatives dels fosfats de les cadenes i, en segon lloc, a la seva coordinació amb els grups O6 de les guanines de tètredes consecutives (figura 10b). De fet, l'espai comprès entre dues tètredes consecutives constitueix un bon lloc d'unió per als cations, especialment per al potassi,

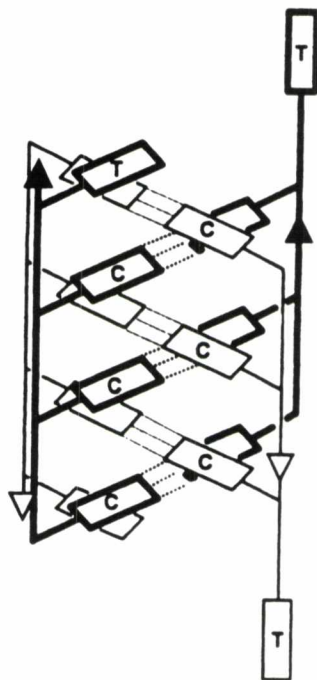


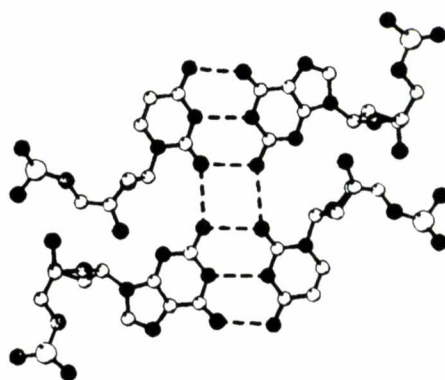
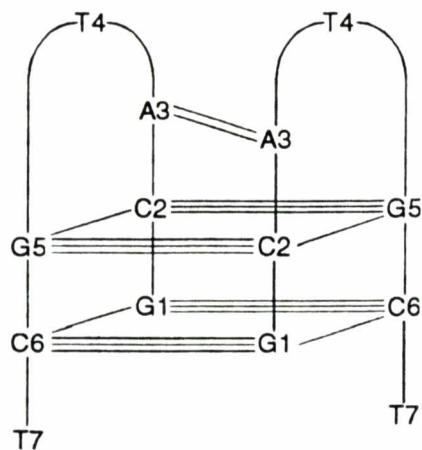
FIGURA 12. Esquema de l'estructura de $d(C_4T)$, formada per dos dúplex intercalats. Les dues cadenes que formen un dúplex són paral·leles; en canvi, els dos dúplex tenen una disposició relativa antiparal·lel. (Les sagetes indiquen l'orientació de les cadenes. Un dúplex s'ha representat amb línia gruixuda i l'altre, amb línia estreta) (adaptat de Wahl i Sundaralingam, 1995).

donat que les seves dimensions són ideals per a la coordinació d'aquest catió. D'acord amb aquest argument, l'efecte estabilitzador dels cations monovalents decreix tant en augmentar ($K > Rb > Cs$) com en disminuir ($K > Na > Li$) el radi iònic.

Els tètplex paral·lels han de ser necessàriament tetramolculars, és a dir, en la seva formació intreenen sempre quatre molècules de DNA independents (figura 11a). En canvi, els tètplex antiparal·lels poden ser també bimolculars (figura 11b) o unimolculars (intramolculars) (figura 11c). L'equilibri entre aquestes diferents formes molculars està influït per diversos factors i molt especialment pels ions. Les formes tetramolculars es veuen afavorides en presència de potassi, mentre que les formes bimolculars i intramolculars es formen també en presència de sodi. Resulta curiós que tots els tètplex tetramolculars descrits fins ara siguin de tipus paral·lel. És probable que el tipus de catió tingui, també en aquest cas, un efecte determinant.

L'interès principal de l'estudi dels tètplex de guanines radica en el fet que el DNA telomèric té una gran tendència a formar estructures tetracatenàries. En general, les seqüències telomèriques contenen grups de tres o més guanines consecutives espaiades per 2-4 residus, normalment timines. Aquestes seqüències s'estenen en forma de cadena senzilla a la fi dels cromosomes, en direcció 3'. D'ací la seva gran tendència a formar estructures tetracatenàries.

La formació d'estructures tetracatenàries no està limitada al cas de les seqüències enriquides en guanina. Altres seqüències també poden formar tètplex encara que amb característiques estructurals molt diferents a les dels tètplex de guanines. Així per exemple, els oligonucleòtids $d(C_4)$, $d(C_3T)$ i $d(TAAGCC)$ (Kang *et al.*, 1995; i algunes referències contingudes en aquest article) donen lloc a estructures tetracatenàries on les interaccions són bàsicament a través de citosines. L'estructura és molt diferent de la dels tètplex de guanines. Els

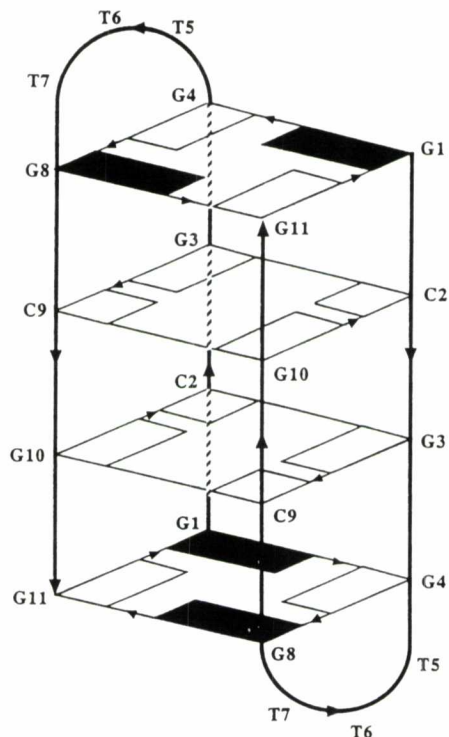


a)

b)

FIGURA 13. a) Diagrama esquemàtic de l'estructura tètplex formada per dues cadenes de d(GCATGCT); i b) interaccions de pont d'hidrogen, representades amb línia discontinua, d'una tètrega G-C-G-C (adaptat de Leonard *et al.*, 1995).

a)



b)

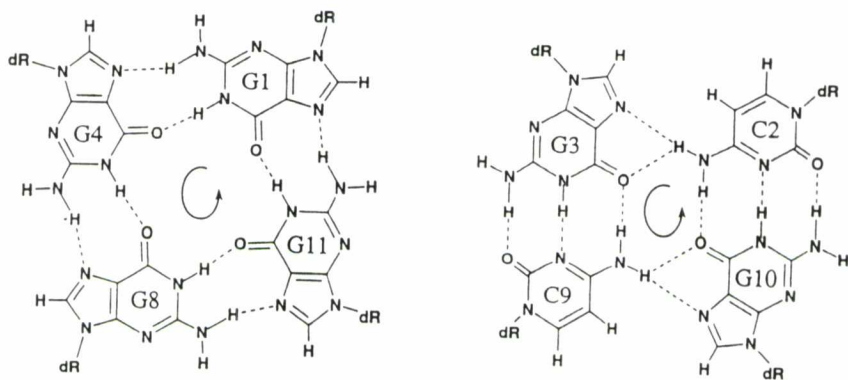


FIGURA 14. a) Esquema del plegament de la seqüència d(GCGGT, GCGG) i de l'associació de dues cadenes per formar un tètraplex. L'esquelet sucre-fosfat s'indica amb línia gruixuda, i el sentit 5'→3' de les cadenes, mitjançant fletxes gruixudes. Les bases C i G es representen amb rectangles, i els ponts d'hidrogen de les tètrades G-G-G-G i els de tipus Watson-Crick de la tètrada C-G-C-G s'indiquen per fletxes primes orientades del donador a l'acceptor. En negre es mostren les guanines en conformació syn. b) Representació de les tètrades: G-G-G-G (a l'esquerra) i C-G-C-G (a la dreta) (modificat de Kettani *et al.*, 1995).

- KETTANI, A.; KUMAR, R. A.; PATEL, D. J. (1995). «Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat». *J. Mol. Biol.*, núm. 254, pàg. 638-656.
- LAUGHLAN, G.; MURCHIE, A. I. H.; NORMAN, D. G.; MOORE, M. H.; MOODY, P. C. E.; LILLEY, D. M. J.; LUISI, B. (1994). «The high-resolution crystal structure of a parallel-stranded guanine tetraplex». *Science*, núm. 265, pàg. 520-524.
- LEONARD, G. A.; ZHANG, S.; PETERSON, M. R.; HARROP, S. J.; HELLIWELL, J. R.; CRUSE, W. B. T.; LANGLOIS D'ESTAINOT, B.; KENNARD, O.; BROWN, T.; HUNTER, W. N. (1995). «Self-association of a DNA loop creates a quadruplex: crystal structure of d(GCATGCT) at 1.8 Å resolution». *Structure*, núm. 3, pàg. 335-340.
- MALININA, L.; TERESHKO, V.; IVANOVA, Z.; SUBIRANA, J. A.; ZARYTOVA, V.; NEKRASOV, Y. (1998). «Structural variability and new intermolecular interactions of Z-DNA in crystals of d(pCpGpCpGpCpG)». [Enviat]
- MALININA, L.; URPI, L.; SALAS, X.; HUYNH-DINH, T.; SUBIRANA, J. A. (1994). «Recombination-like structure of d(CCGCGG)». *J. Mol. Biol.*, núm. 243, pàg. 484-493.
- MIRKIN, S. M.; FRANK-KAMENESTKII, M. D. (1994). «H-DNA and related structures». *Annu. Rev. Biophys. Biomol. Struct.*, núm. 23, pàg. 541-576.
- NUNN, C. M.; NEIDLE, S. (1996). «The high resolution crystal structure of the DNA decamer d(AGGCATGCCT)». *J. Mol. Biol.*, núm. 256, pàg. 340-351.
- PAN, B.; BAN, C.; WAHL, M. C.; SUNDARALINGAM, M. (1997). «Crystal structure of d(CGCGCG) with 5'-overhang G residues». *Biophysical J.*, núm. 73, pàg. 1553-1561.
- RADHAKRISHNAN, I.; PATEL, D. J. (1994). «DNA triplexes: Solution structures, hydration sites, energetics, interactions and function». *Biochem.*, núm. 33, pàg. 11405-11416.
- SAENGER, W. (1984). «Principles of Nucleic Acid Structure», pàg. 1-296. (C. R. Cantor ed.) Springer-Verlag Nova York Inc.
- SALISBURY, S. A.; WILSON, S. E.; POWELL, H. R.; KENNARD, O.; LUBINI, P.; SHEDRICK, G. M.; ESCAJA, N.; ALAZZOUZI, E.; GRANDAS, A.; PEDROSO, E. (1997). «The bi-loop, a new general four-stranded DNA motif». *Proc. Natl. Acad. Sci. USA*, núm. 94, pàg. 5515-5518.
- SHAKKED, Z.; RABINOVICH, D.; CRUSE, W. B. T.; EGERT, E.; KENNARD, O.; SALA, G.; SALISBURY, S. A.; VISWAMITRA, M. A. (1981). «Crystalline A-DNA: the x-ray analysis of the fragment d(GGTATACC)». *Proc. R. Soc. Lond. (Biol.)*, núm. 213, pàg. 479-487.
- SMITH, F. W.; FEIGON, J. (1992). «Quadruplex structure of Oxytricha telomeric DNA oligonucleotides». *Nature*, núm. 356, pàg. 164-168.
- SUBIRANA, J. A.; FARIA, T. (1997). «Influence of sequence on the conformation of the B-DNA helix». *Biophysical J.*, núm. 73, pàg. 333-338.
- SUBIRANA, J. A.; SALAS, X.; URPI, L.; FONT, E.; VERDAGUER, N. (1995). Capítol: «Aportaciones de la cristalografía al estudio del ADN». A: *Cristalografía, col·lecció: «Nuevas Tendencias»*, vol. 26, Madrid, CSIC, pàg. 223-240.
- SUNDQUIST, W. I. «The structures of telomeric DNA». *Nucleic Acids and Molecular Biology*, núm. 5, pàg. 1-24.
- URPI, L.; TERESHKO, V.; MALININA, L.; HUYNH-DINH, T.; SUBIRANA, J. A. (1996). «Structural comparison between the d(CTAG) sequence in oligonucleotides and trp and met repressor-operator complexes». *Nature Struct. Biol.*, núm. 3, pàg. 325-328.
- VAN MEERVELT, L.; VLEGHE, D.; DAUTANT, A.; GALLOIS, B.; PRÉCIGOUX, G.; KENNARD, O. (1995). «High resolution structure of a DNA helix forming (C.G)*G base triplets». *Nature*, núm. 374, pàg. 742-744.
- VERDAGUER, N.; AYMAMI, J.; FERNÁNDEZ-FORNER, D.; FITA, I.; COLL, M.; HUYNH-DINH, T.; IGOLEN, J.; SUBIRANA, J. A. (1991). «Molecular structure of a complete turn of A-DNA». *J. Mol. Biol.*, núm. 221, pàg. 623-635.
- WAHL, M. C.; RAO, S. T.; SUNDARALINGAM, M. (1996). «Crystal structure of the B-DNA hexamer d(CTCGAG): Model for an A-to-B transition». *Biophysical J.*, núm. 70, pàg. 2857-2866.
- WAHL, M. C.; SUNDARALINGAM, M. (1995). «New crystal structures of nucleic acids and their complexes». *Current Opinion in Structural Biology*, núm. 5, pàg. 282-295.
- WANG, A. H.-J.; QUIGLEY, G. J.; KOLPAK, R. J.; CRAWFORD, J. L.; VAN BOOM, J. H.; VAN DER MAREL, G. A.; RICH, A. (1979). «Molecular structure of a left-handed double helical DNA fragment at atomic resolution». *Nature*, núm. 282, pàg. 680-686.
- WATSON, J. D.; CRICK, F. H. (1953). «Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid». *Nature*, núm. 171, pàg. 737-738.
- WING, R.; DREW, H.; TAKANO, T.; BROKA, C.; TANAKA, S.; ITAKURA, K.; DICKERSON, R. E. (1980). «Crystal structure analysis of a complete turn of B-DNA». *Nature*, núm. 287, pàg. 755-758.