La struttura degli acidi nucleici (e le loro varianti)

26 ottobre 2014

di Sergio Barocci

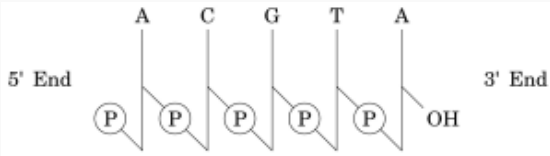
[*prosegue, dell0 stesso Autore, dall'articolo:* ***"Gli acici nucleici: dalla cromatina ai nucleotidi"***]

I nucleotidi del DNA e dell'RNA sono uniti tra loro da ponti covalenti tra gruppi fosforici in cui il gruppo OH 5' di un'unità nucleotidica è unito al gruppo OH 3' di quella successiva, formando un legame fosfodiestereo (reazione di condensazione tra il gruppo fosforico di un nucleotide e l'ossidrile sul carbonio 3' dello zucchero di un altro nucleotide costituendo così un polinucleotide (legami 3' – 5' fosfodiesterici) . Per convenzione, un polinucleotide si scrive con il carbonio terminale 5' a sinistra e quello 3' a destra.
Lo scheletro covalente degli acidi nucleici è quindi costituito da un'alternanza di gruppi fosforici e di residui di pentosio, mentre le basi azotate possono essere considerate come gruppi laterali uniti allo scheletro. Lo scheletro covalente del DNA e dell'RNA è idrofilico.

Noto esperimento didattico per l'estrazione del DNA da frutta

Pertanto, ciò che differenzia un nucleotide dall'altro è appunto la diversa base che lo costituisce e la sequenza in cui sono disposte le basi è quella che determina il codice genetico dell'individuo.3' manca di un nucleotide nella posizione 3'

Le proprietà delle basi dei nucleotidi determinano la struttura tridimensionale degli acidi nucleici. Le basi puriniche e pirimidiniche sono molecole altamente coniugate. Altre e importanti conseguenze sulle proprietà degli acidi nucleici sono la **struttura**, la **distribuzione elettronica**, l'**assorbimento della luce**. Per effetto della risonanza tra gli atomi degli anelli le pirimidine sono planari, le purine sono quasi planari e inoltre tutte le basi assorbono la luce UV. Le basi puriniche e pirimidiniche sono idrofobiche e poco solubili in acqua ai valori di pH della cellula



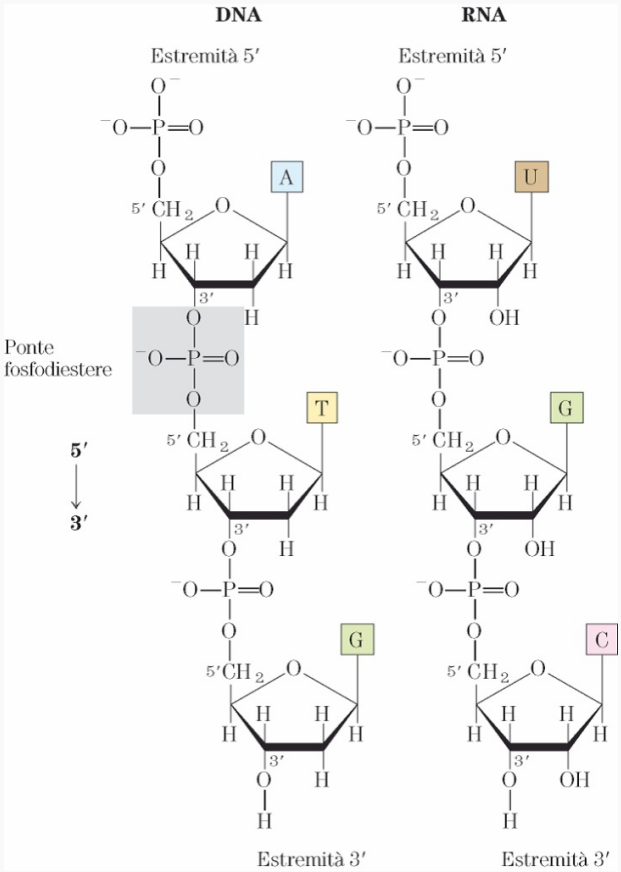
rappresentazione sequenza nucleotidica

La **struttura primaria** di un acido nucleico è la sua **struttura covalente**. Qualsiasi struttura stabile e regolare assunta da una parte o da tutti i nucleotidi di un acido nucleico può essere considerata come **struttura secondaria**. Il complesso ripiegamento di un cromosoma all'interno del nucleotide batterico o della cromatina degli eucarioti viene in genere considerato come **struttura terziaria**.

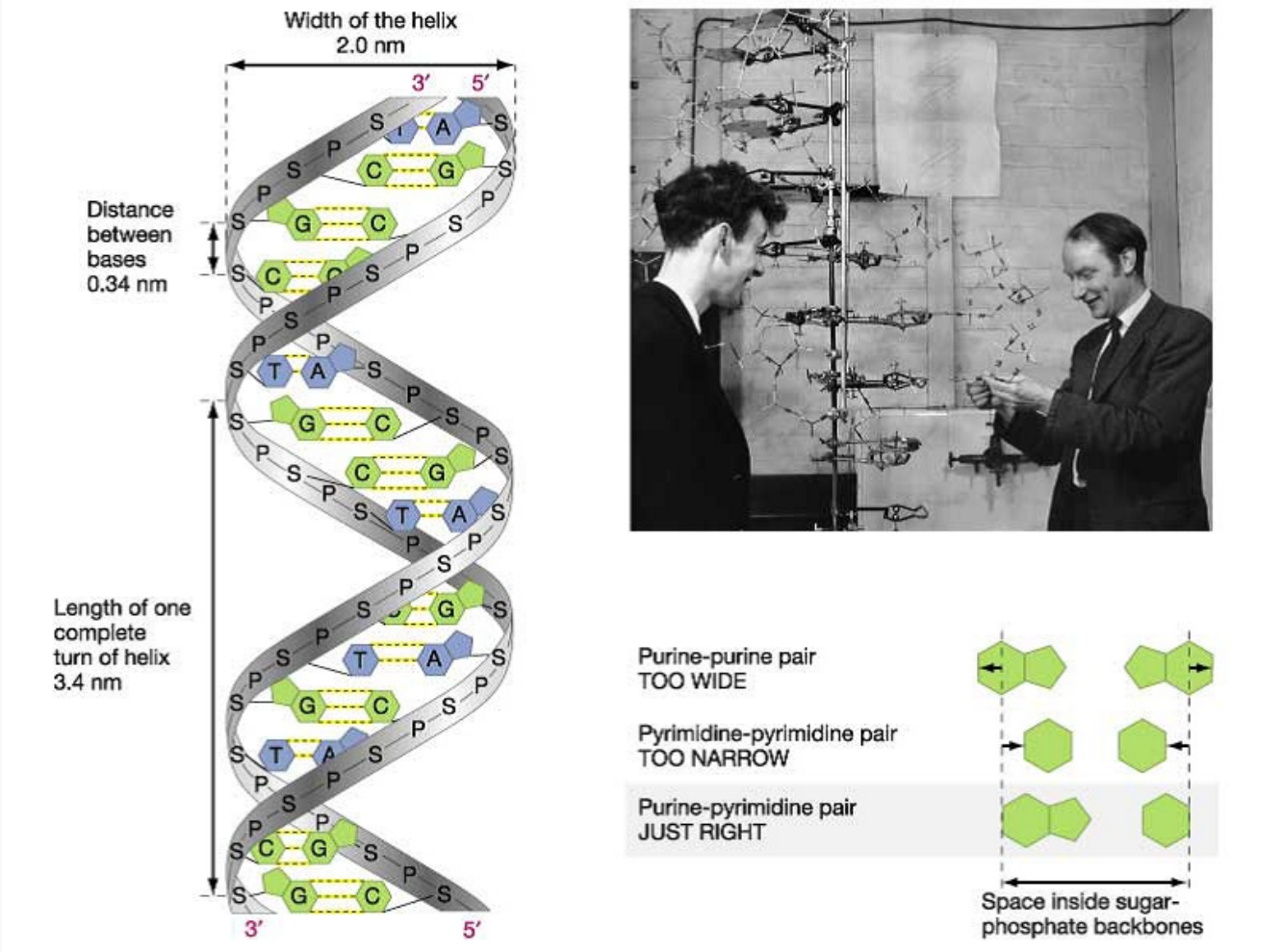
Il DNA

Il DNA risulta di forma circolare nei batteri, mentre nelle altre cellule di organismi superiori è organizzato in forme filamentose chiamate cromosomi. Il DNA di un mammifero e formato da circa 5,5 x 109 coppie di basi. Se tutte le molecole fossero unite linearmente coprirebbero una distanza di circa 2 metri. Ciò è impossibile e la chiave di questo dubbio è stato risolto con la scoperta di proteine chiamate istoni e di piccole dimensioni , ricche di aminoacidi basici in grado di associarsi con i residui di acido fosforico. Tanti istoni possono fungere come un gomitolo dove viene avvolto il filo di lana. I filamenti non sono tutti omogenei, il tratto di DNA che lega due strutture non presenta nessun gene. Ci può essere ancora una seconda struttura dove avviene un altro avvolgimento che accorcia ancora la molecola.

La struttura a doppia elica del DNA dedotta da Watson e Crick ha influenzato enormemente il corso della biologia perché ha indicato come l'informazione genetica viene conservata e replicata.



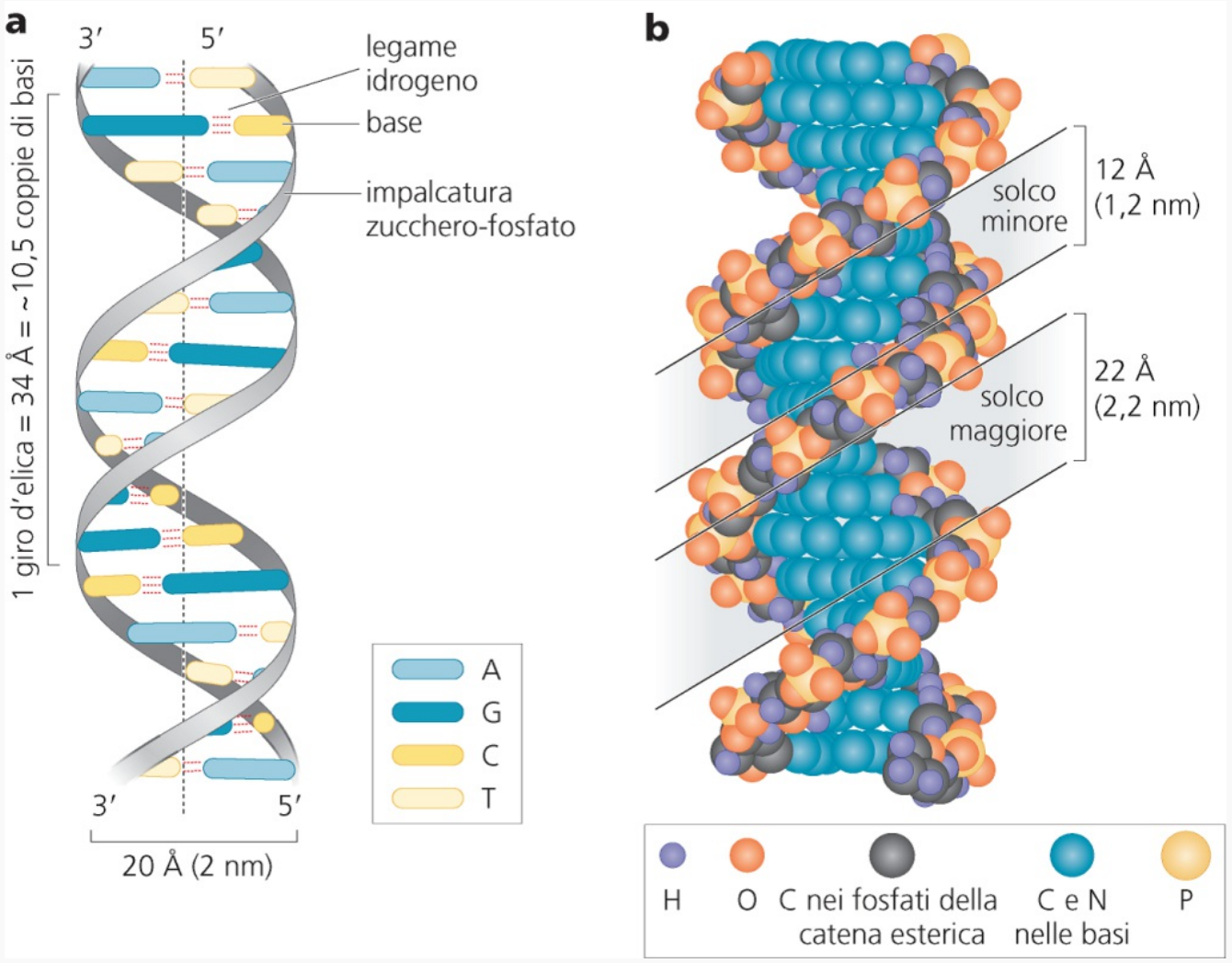
L'estremità 5' manca di un nucleotide nella posizione 5' e l'estremità 3' manca di un nucleotide nella posizione 3'



modello di Watson e Crick per il DNA

Gli aspetti essenziali del modello di Watson e Crick sono :

• Lo scheletro di ciascun'elica è composto dai residui di deossiribosio uniti mediante legami fosfodiesterici covalenti in posizione 5' e 3' sull'anello dello zucchero
• i due filamenti (eliche) del DNA sono uniti all'interno della struttura da legami idrogeno che occorrono tra le basi puriniche di una catena e le basi pirimidiniche dell'altra(le basi puriniche e pirimidiniche si trovano infatti all'interno mentre i gruppi fosforici e gli zuccheri sono all'esterno; inoltre il piano delle basi è perpendicolare all'asse dell'elica e il piano degli zuccheri forma un angolo quasi retto con quello delle basi);
• l'adenina è sempre accoppiata alla timina e la guanina è sempre accoppiata alla citosina, in base a fattori sterici e di formazione dei legami idrogeno;
• la sequenza precisa delle basi porta l'informazione genetica;
• le estremità di ciascun filamento di DNA sono chiamate 5'e 3'. Per convenzione, con il 5'si indica la sequenza più vicina all'inizio del gene, con il 3' si indica la sequenza più vicina alla fine del gene;
• i due scheletri corrono in direzioni opposte, sono antiparalleli;
• Il DNA è sempre sintetizzato dal 5'al 3';
• Il diametro dell'elica è di 20 °A, le basi adiacenti sono separate da 3,4°A lungo l'asse dell'elica e formano tra loro un angolo di 36°;
• La struttura dell'elica si ripete quindi 10 residui di ciascuna catena intervalli cioè di 34°A

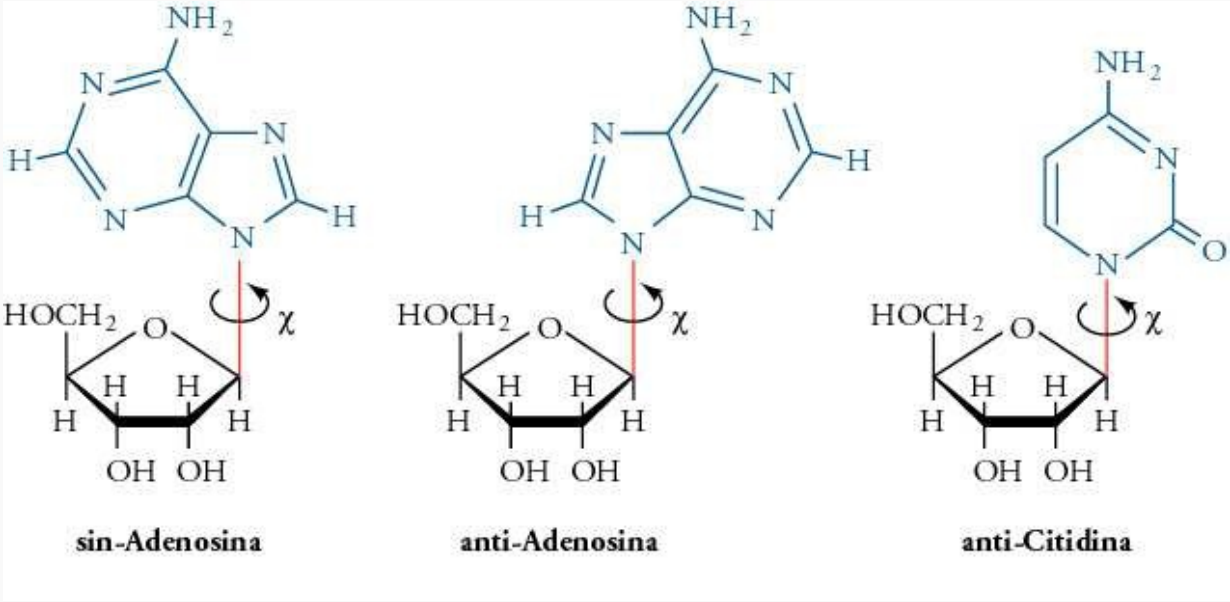


solchi maggiore e minore nella struttura del DNA

Quindi, l'aspetto più importante della doppia elica del DNA consiste nella specificità dell'accoppiamento delle basi. Le restrizioni steriche sono imposte dalla natura regolare ad elica dell'ossatura costituita da zuccheri e gruppi fosforici, di ciascuna catena polinucleotidica. I legami glicosidici di una coppia di basi sono sempre distanti tra loro 10,8°A. Infatti, una coppia di basi purinica- pirimidinica si adatta in maniera perfetta in questo spazio che non è altrettanto sufficiente per due purine mentre due pirimidine sarebbero invece troppo lontane l'una dall'altra per formare legami idrogeno. Pertanto, la coppia di basi nell'elica del DNA deve essere sempre costituita da una purina e da una pirimidina per problemi sterici. L'appaiamento delle basi è ulteriormente ristretto dalle caratteristiche dei legami idrogeno in quanto gli atomi di idrogeno delle basi puriniche e delle basi pirimidiniche hanno posizioni ben definite perché l'adenina non può accoppiarsi con la citosina in quanto ci sarebbero due idrogeni vicini ad una delle posizioni di legame e nessuno vicino all'altra. La guanina non può accoppiarsi alla timina per gli stessi motivi. L'adenina, invece, può formare due legami idrogeno con la timina e la guanina tre con la citosina perché le distanze di tali legami risultano ottimali affinchè si possa stabilire una forte attrazione tra le basi.

A causa delle costrizioni steriche, i nucleotidi purinici e pirimidinici possono assumere, come già riferito, solo due **conformazioni stabili** rispetto al deossiribosio, chiamate ***sin*** e ***anti***. Le pirimidine sono in genere nella conformazione anti a causa di interferenze steriche tra lo zucchero e l'ossigeno carbonilico sul C-2.

La struttura a doppia elica del DNA descritta da Watson e Crick è la **forma B del DNA** o **DNA B**. In realtà, il DNA possiede una **struttura dinamica** e può quindi assumere **diverse forme ad elica**.



conformazioni sin e anti nei nucleotidi

Le attuali conoscenze sul modello tridimensionale della struttura del DNA

Il modello proposto da Watson e Crick , sul noto DNA B, si basa su profili di diffrazione ai raggi X di fibre di DNA che forniscono informazioni sulle proprietà della doppia elica che rappresentano una media delle proprietà dei suoi componenti. Informazioni strutturali più dettagliate sono state poi in seguito ottenute per mezzo della diffrazione ai raggi X di cristalli di DNA a risoluzione atomica che hanno rivelato che il DNA possiede una variabilità e una diversità strutturale maggiore di quanto si pensasse che dipendono da diversi fattori.
Una catena di DNA può ruotare intorno a **6 legami per monomero** (il legame glicosidico tra la base e lo zucchero , il legame C4'- C5' dello zucchero e quattro legami nel ponte fosfodiestereo che unisce il C3' di uno zucchero e il C5' di quello successivo). Infatti, il ripiegamento e il diverso orientamento dell'anello di riboso sono una importante **caratteristica strutturale**.
L'analisi di un **dodecamero cristallizzato di DNA** ai raggi X eseguita da **Richard Dickerson** nel 1985 , ha dimostrato che la sua struttura generale era molto simile a quella della doppia elica di Watson e Crick anche se tale dodecamero differiva dal modello di Watson e Crick, non solo, in quanto non uniforme ma anche per alcune delle deviazioni locali della struttura media, piuttosto consistenti.

a) Le due basi in alcune paia di basi non sono perfettamente complanari ma girate l'una rispetto all'altra come le pale di un'elica. Questa deviazione dalla struttura teorica denominata **torsione ad elica** permette di aumentare l' impilamento delle basi (**stacking** cioè disposizione impilata di molecole aromatiche che ogni base esercita su quelle adiacenti) lungo un filamento. Queste ed altre variazioni locali della doppia elica dipendono dalla sequenza delle basi.

b) Un altro aspetto dell'elica del **DNA B** è che l'elica si può curvare e formare un arco o super avvolgersi senza che si vengano a creare modificazioni rilevanti della struttura locale. Questa facilità di deformazione è molto importante dal punto di vista biologico in quanto permette al DNA circolare di formarsi o di avvolgersi intorno a proteine. Un DNA compatto permette di occupare un volume molto più piccolo grazie alla sua deformabilità. Se il DNA fosse obbligato ad essere lineare non potrebbe trovare posto all'interno di una cellula.

c) Il DNA può anche formare dei ripiegamenti a linea spezzata cioè può essere in grado di cambiare bruscamente di direzione, ripiegamenti indotti da sequenze specifiche di basi come ad esempio una fila di quattro residui di adenina o dal legame di una proteina.

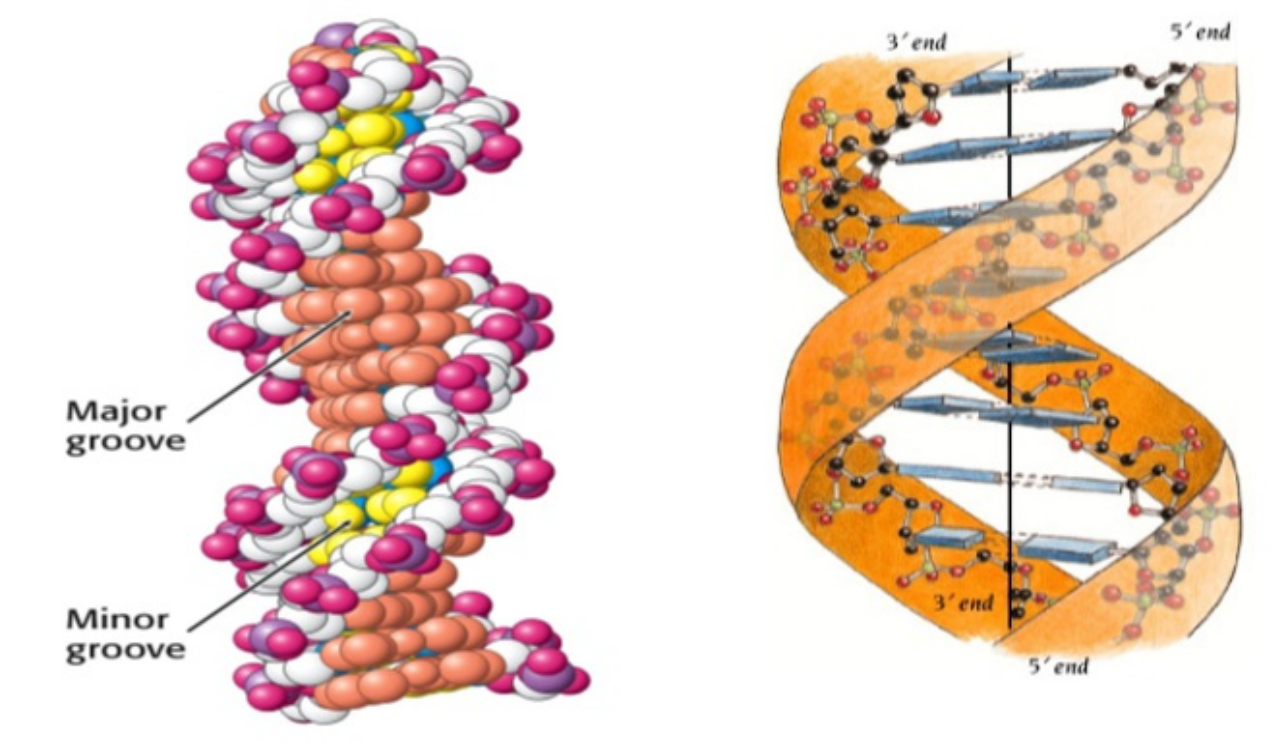
Altre misure strutturali:
a) ad ogni giro di elica vi sono 10,5 coppie di basi a intervalli di 36 Å°. Questa frequenza di rotazione, nota come

passo dell'elica (distanza tra un elica ed un' altra ed un numero di coppie di basi per passo pari), dipende largamente dalle forze di stacking.

b) Il diametro dell'elica è 23,7 A° .

c) Ogni unità nucleotidica è lunga 3,4 Å°.

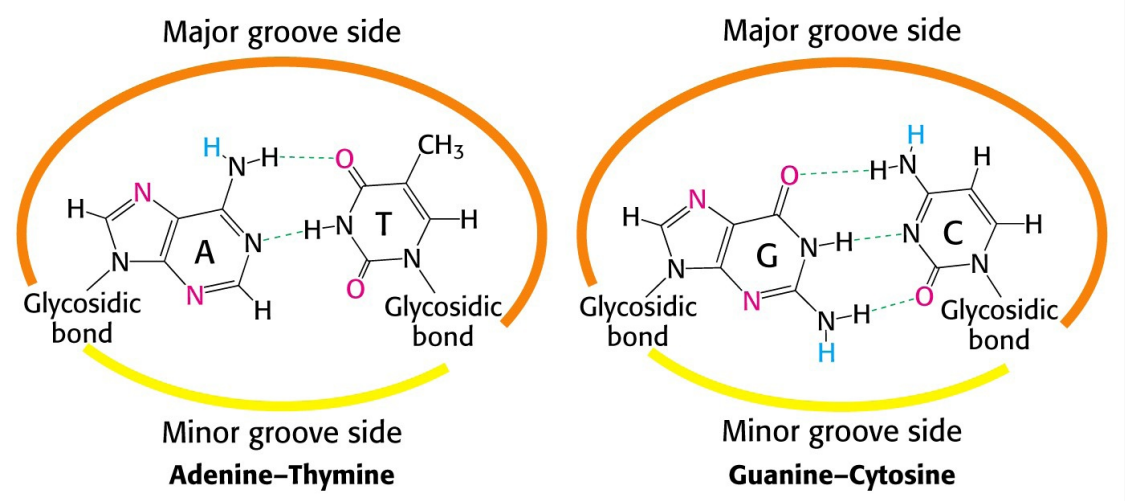
d) Il DNA B possiede due tipi di scanalature (**grooves**), denominate **scanalatura principale** o **maggiore** (major groove) (larghezza 12 °A) e **scanalatura secondaria** o **minore** (minor groove) (larghezza 6 °A) che si formano in quanto i legami glicosidici di una coppia di basi non sono diametralmente opposti l'uno all'altro.



Nella doppia elica di DNA sono presenti due solchi di diversa ampiezza

e) La **scanalatura secondaria** contiene l'O-2 della pirimidina e l'N-3 della purina di una coppia di basi mentre quella principale si trova sul lato opposto della coppia. La scanalatura principale è leggermente più profonda di quella secondaria

f) Le coppie di basi A-T e G-C del DNA contengono atomi che possono formare ulteriori legami idrogeno . Sia la scanalatura principale che quella secondaria sono rivestite da atomi che possono essere donatori o accettori in un legame idrogeno



"lati" nelle coppie di nucleotidi che si affacciano sul solco maggiore e su quello minore nella doppia elica del DNA

g) Nella **scanalatura secondaria** l'N-3 dell'adenina e della guanina e l'O-2 della timina e della citosina possono servire da accettori di idrogeno e il gruppo aminico legato al C-2 della guanina può essere un donatore di idrogeno.

h) Nella **scanalatura principale** l'N-7 della guanina e dell'adenina è un potenziale accettore così come l'O-4 della timina e l'O-6 della guanina.

i) I gruppi aminici legati al C-6 dell'adenina e al C-4 della citosina possono servire da donatori di idrogeno.

j) Inoltre, la scanalatura principale ha più caratteristiche distintive di quella secondaria e che le sue maggiori dimensioni la rendono più accessibile alle interazioni con proteine che riconoscono sequenze specifiche di DNA.

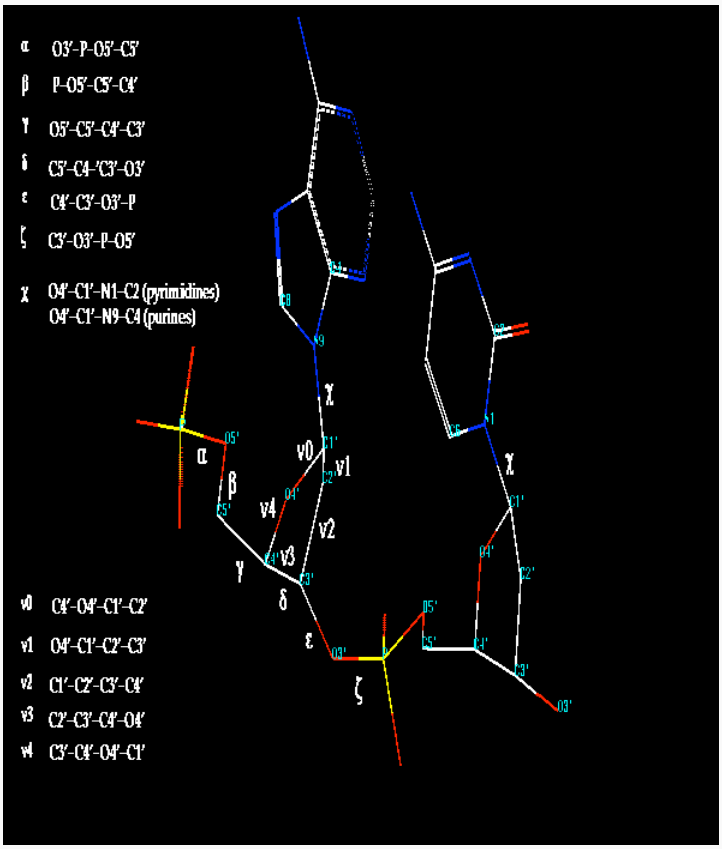
LE DIVERSE ELICHE DEL DNA

Il DNA ha una struttura dinamica molto flessibile per l'assenza di doppi legami e ciò gli consente di assumere diverse conformazioni . E' possibile indurre la transizione da una conformazione di DNA ad un'altra variando le condizioni ambientali.
La forma ad elica del DNA non è soggetta a grandi discussioni, ma esistono vari tipi di elica che possono essere formati. Le tipologie di elica più facilmente riscontrabili prendono il nome di forma A, B e Z. Le prime due tipologie sono legate a differenze dei ripiegamenti dell'anello dello zucchero a 5 atomi di Carbonio : quattro dei cinque atomi dell'anello sono coplanari, mentre il quinto è esterno. Se il quinto atomo è dallo stesso lato di C5 la conformazione è endo, se è dal lato opposto è eso: la forma B possiede la conformazione endo su C2', la forma A su C3'.

I parametri che determinano la geometria e la conformazione della doppia elica sono rappresentati da:
1. Angoli torsionali
2. Movimenti translazionali e rotazionali delle basi azotate

Gli angoli di torsione che descrivono la conformazione dei legami della catena fosfodiesterea sono sufficienti per descrivere in maniera dettagliata la struttura di una catena polinucleotidica. La struttura tridimensionale del DNA viene caratterizzata dalla lunghezza dei legami, dall'angolo di legame e dalla rotazione di gruppi di atomi attorno ai legami covalenti.
L'angolo torsionale rappresenta la rotazione attorno ad un legame centrale e coinvolge quattro atomi in sequenza.
Gli angoli torsionali definiscono la geometria dell'impalcatura Z – P , e l'orientamento delle basi azotate rispetto all'impalcatura. Sono preferiti valori di angoli torsionali che portano alla formazione di conformazioni stericamente possibili. Limitazioni alle libere rotazioni sono determinate dalla presenza dello zucchero e delle basi.

L'anello furanosico a 5 atomi di Carbonio come quello piranosico a 6 atomi di carbonio non è ripiegato in modo tale da avere 4 atomi di C quasi sullo stesso piano e il quinto fuori dal piano. Non ha pertanto una struttura planare per la presenza di sostituenti.



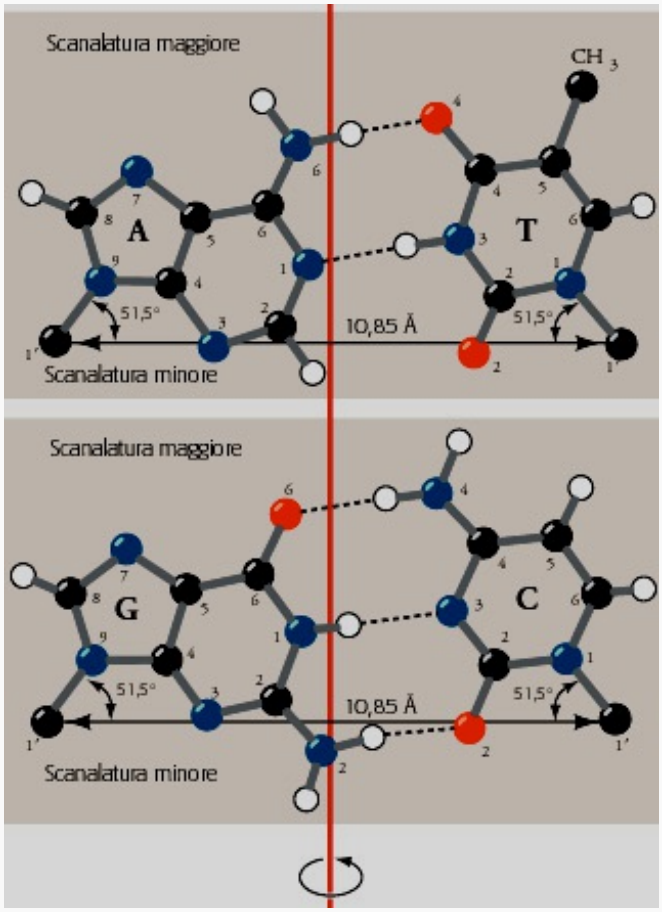
come la conformazione anello furanosidico del desossiribosio influenza la struttura dell'elica del DNA

I sostituenti degli atomi di C dell'anello possono essere di due tipi :

a) **Assiali**: sono circa perpendicolari al piano medio dell'anello
b) **Equatoriali**: sono circa paralleli a questo piano.

I **sostituenti assiali** sporgono sopra e sotto il piano medio dell'anello mentre quelli equatoriali si trovano alla periferia dell'anello. Questi sostituenti possono generare ostacoli sterici ad altri gruppi se emergono ad esempio sullo stesso lato dell'anello mentre quelli equatoriali hanno molto più spazio a disposizione.
A causa della repulsione tra i sostituenti degli atomi di carbonio, l'anello a cinque termini dello zucchero non è planare, ma può assumere varie forme nello spazio: le più comuni sono quelle che presentano un atomo fuori dal piano determinato dagli altri quattro atomi dell'anello (forme " a busta") o due atomi adiacenti spostati su lati opposti rispetto al piano determinato dagli altri tre atomi (forme 'a mezza sedia' o twist). Nel primo caso le denominazioni Y- endo e Y – eso indicano che l'atomo Y è fuori dal piano e spostato, rispettivamente, nello stesso verso dell'atomo C5' o nel verso opposto. I dati sperimentali stanno ad indicare che le forme più stabili sono le due forme "a busta" C3'- endo e C2′- endo. Lo scheletro degli acidi nucleici è formato dalla ripetizione di sei atomi (−P−O5′−C5′−C4′−C3′−O3′−)n, due dei quali, gli atomi C4′ e C3′, fanno parte anche della struttura ciclica del pentosio . Assumendo l'equivalenza conformazionale dei nucleotidi, la conformazione della catena è definita da cinque angoli di torsione che vengono contrassegnati da lettere greche , ai quali va aggiunto l'angolo χ che definisce l'orientazione della base rispetto allo zucchero. L'angolo di torsione δ intorno al legame C4′−C3′ può variare entro intervalli ristretti e i valori che può assumere sono strettamente legati alla conformazione dello zucchero (~ 80° per la conformazione C3′- endo del ribosio e ~ 140° per quella C2′- endo). Le interazioni attrattive e repulsive, tra atomi non direttamente legati, limitano notevolmente i valori assunti dagli altri angoli di torsione.

Inoltre, i dati strutturali su oligonucleotidi indicano che essi sono correlati. L'angolo χ intorno al legame glucosidico può assumere valori vicini a 0° (conformazione sin) o a 180° (conformazione anti): le basi puriniche possono adottare entrambe le conformazioni, mentre le basi pirimidiniche di norma prediligono la conformazione anti. Un altro fattore che diminuisce drasticamente il numero di stati conformazionali energeticamente accessibili agli acidi nucleici è l'accoppiamento antiparallelo di due catene con sequenze complementari, cioè di sequenze di nucleotidi tali per cui a una base purinica G o A su una catena corrisponde rispettivamente una base pirimidinica C o T (U) sull'altra, e viceversa; la complementarità delle basi permette la formazione di legami idrogeno specifici tra A e T (U) e tra C e G. L'unicità e l'importanza di questo accoppiamento, rispetto ad altri possibili, risiedono nei parametri geometrici che collegano le coppie di basi all'atomo C1′ delle rispettive catene: le coppie A/T, T/A, C/G e G/C sono da questo punto di vista sufficientemente isomorfe (forma e dimensioni molto simili) da poter essere interscambiate senza perturbare significativamente la struttura delle catene; pertanto, nella condizione di complementarità delle due catene, la mancanza di regolarità nella successione delle basi non invalida sostanzialmente il principio di equivalenza conformazionale.



Isomorfismo geometrico delle coppie di basi. I legami β –glicosidici delle 4 coppie di basi sono equivalenti

La comprensione di tutti questi fattori, che limitano drasticamente le possibili conformazioni dell'unità nucleotidica, congiuntamente alle caratteristiche del diagramma di diffrazione da fibra, hanno portato nel 1953 J.

D. Watson e F. H. C. Crick a proporre la struttura a doppia elica per il DNA. Oltre alla formazione dei legami idrogeno tra le coppie di basi, uno dei fattori più importanti di stabilizzazione per questa struttura è la forte sovrapposizione tra le coppie di basi successive lungo l'asse dell'elica: questo effetto dipende dalla geometria della sovrapposizione ed è maggiore, per geometrie simili, per la coppia C/G rispetto alla coppia A/T. Per questo motivo, i DNA con un contenuto più alto di basi C e G sono più stabili e fondono a temperature più alte. Il volume di ingombro della doppia elica è assimilabile a quello di un cilindro, lungo la cui superficie laterale è possibile individuare due solchi elicoidali di dimensioni diverse: in entrambi i solchi ognuna delle quattro possibili coppie di basi è esposta al solvente e può essere riconosciuta in maniera univoca per la diversità degli atomi esposti. Molecole di dimensioni opportune possono interagire specificamente con le basi senza alterare sostanzialmente la struttura a doppia elica. Fino all'inizio degli anni ottanta, le informazioni strutturali sugli acidi nucleici sono state ottenute essenzialmente dai dati di diffrazione su fibre orientate di DNA, RNA e polinucleotidi sintetici, oltre che dai dati di diffrazione su cristallo singolo di mono- e dinucleotidi.
Con lo sviluppo di metodi nuovi per la sintesi di frammenti di DNA, aventi lunghezza e sequenza controllata, è stato possibile produrre oligonucleotidi autocomplementari in quantità e purezza sufficienti per la preparazione di cristalli singoli; gli studi cristallografici ad alta risoluzione di questi nucleotidi, oltre a fornire informazioni dettagliate sulle strutture già note del DNA, hanno anche permesso a Wang e Rich nel 1979 a scoprire un nuovo tipo di struttura sinistrorsa, chiamata struttura Z .

Alexander Rich

Alexander Rich (1924 –) biologo e fisico statunitense. Nel 1963 scopre i polisomi, gruppi di ribosomi che leggono contemporaneamente un filamento di RNA e nel 1979 insieme a Wang esaminando un picollo oligonucleotide sintetico di DNA contenente solo coppie di Basi C-G scopre un'ulteriore forma di DNA, detta forma Z, la quale assume la configurazione di una doppia elica sinistrorsa.
R. Dickerson nel 1985 evidenzia che in condizioni di bassa forza ionica la forma B è stabile, caratterizzata da una simmetria 10/1 e da una conformazione C2'- endo del desossiribosio. Le basi giacciono in piani praticamente normali all'asse dell'elica e lo spostamento assiale h, tra coppie di basi successive, è di 3,38 Å. Invece nella forma A, stabile per bassi valori di umidità relativa, il desossiribosio adotta la conformazione C 3'- endo e ciò porta a un leggero svitamento della struttura, che assume la simmetria 11/1 e a una modifica della orientazione delle basi, il cui asse risulta inclinato di circa 20° rispetto all'asse dell'elica; la struttura diventa più larga e più corta (h = 2,56 Å) e mostra anche notevoli variazioni nella larghezza e nella profondità dei solchi sulla superficie laterale della doppia elica.

Queste variazioni sono anche responsabili del diverso grado di idratazione delle due strutture: nella forma B il solco maggiore è rivestito da un monostrato di acqua di idratazione, mentre il solco minore, più profondo e stretto, contiene una catena di molecole di acqua ben ordinate, legate l'una all'altra mediante legami idrogeno. È stata avanzata sempre da Dickerson l'ipotesi che questa struttura di molecole di acqua svolga un ruolo importante per la stabilizzazione della forma B in condizioni di elevata umidità relativa (~ 95%) e che la sua rottura sia un primo passo necessario per la transizione alla forma A .

Richard E. Dickerson (1931 –): chimico statunitense, diventato famoso per aver effettuato per primo l'analisi della struttura del DNA su un dodecamero cristallizzato (dodecamero di Dickerson : CGCGAATTCGCG) attraverso cristallografia ai raggi X.
La scoperta della forma Z sinistrorsa è strettamente legata alla risoluzione della struttura cristallina di oligonucleotidi con sequenze alternanti delle basi C e G. Anche se strutture a doppia elica sinistrorse furono prese in considerazione nel periodo in cui fu proposta la struttura B, il buon accordo di questa con i dati sperimentali portò al convincimento generale che essa fosse la forma dominante nei sistemi biologici. Invero, la forma Z può essere ottenuta da soluzioni con elevata forza ionica e solo con sequenze alternanti di nucleotidi purinici e pirimidinici, tra le quali la sequenza CG è la più favorevole.
L'unità strutturale indipendente della forma Z è una coppia di nucleotidi: il nucleotide purinico adotta una conformazione C3'- endo per il desossiribosio e sin per la base, quello pirimidinico preferisce una conformazione C2'- endo per lo zucchero e anti per la base. La doppia elica sinistrorsa ha una simmetria 6/1 (con 12 nucleotidi

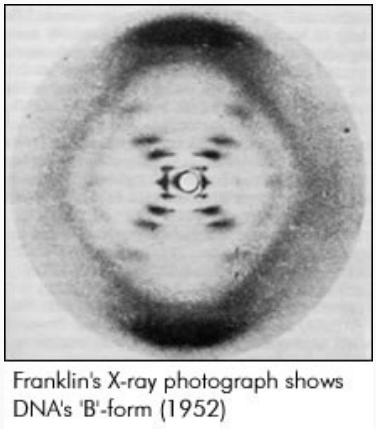
per spira) e un passo h per nucleotide di 3,70 Å essa è molto più sottile e più alta della forma B. L'alternarsi dei parametri conformazionali lungo la catena conferisce un caratteristico andamento a zig-zag, da cui deriva il nome di forma Z. Gran parte del DNA nelle cellule è nella forma B; vi sono tuttavia indicazioni che la forma Z esista in vivo e contribuisca alla regolazione dell'espressione genica attraverso l'interazione con proteine specifiche.

Richard E. Dickerson

L'acido ribonucleico (RNA) differisce dal DNA per la presenza del ribosio al posto del 2′-desossiribosio e della base uracile (U) al posto della timina (T). Le strutture a doppia elica dell'RNA sono simili alla forma A del DNA; la forma B, mai osservata nell'RNA, è destabilizzata da interazioni troppo strette tra il gruppo ossidrilico in posizione 2′ con la base e i gruppi fosfati vicini.
Ibridi RNA-DNA assumono la conformazione A in quanto comune a entrambe le molecole. Normalmente, le molecole di RNA sono molto più corte delle molecole di DNA e non si aggregano. Le molecole di RNA transfer (tRNA) sono formate da ~ 80 nucleotidi e assumono una conformazione ripiegata tenuta insieme dallo stesso tipo di interazioni non covalenti che stabilizzano la struttura a doppia elica del DNA.
Nel tRNA le coppie di basi complementari, legate da legami idrogeno, fanno parte di nucleotidi che appartengono alla stessa catena; quattro segmenti corti della molecola hanno una struttura a doppia elica, lasciando due triplette di basi non accoppiate che svolgono un ruolo essenziale nel riconoscimento di una molecola di RNA messaggero e di una molecola di amminoacido. Per il modo specifico di ripiegamento della catena polinucleotidica, la molecola di tRNA ricorda il comportamento delle proteine globulari.

DNA-A, DNA-B, DNA-Z

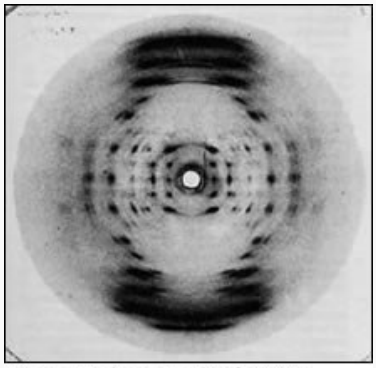
La forma B (il **DNA-B** si trova in una soluzione ad altro grado di umidità o a bassa forza ionica), contiene 10 paia di basi per giro d'elica e ha un ampio solco maggiore e un solco minore più chiuso), quella originariamente descritta da James D. Watson e Francis Crick è ritenuta essere quella predominante nelle cellule , mentre condizioni di bassa idratazione (che si verificano spesso nei campioni cristallizzati) favoriscono la forma DNA – A a causa di una diversa elettrostatica che modifica la forza di repulsione tra i gruppi fosfato e quindi la geometria dei due grooves (i solchi o scanalature che separano i bordi dell'elica), la cui ampiezza risulta infatti diversa nei due tipi di elica.



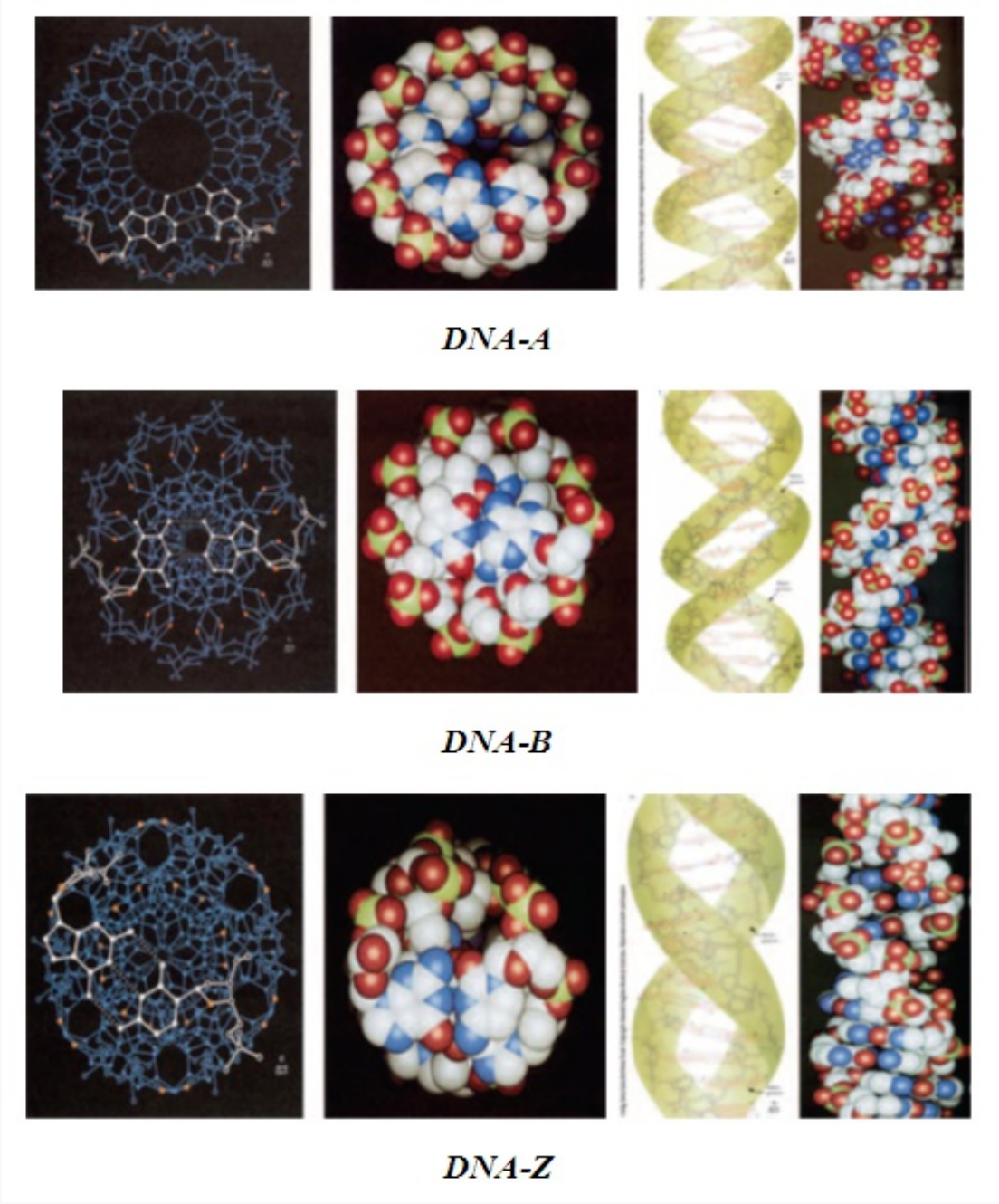
diffrattogramma a raggi X di B-DNA (fibre idratate)

Il **DNA-A** è presente in condizioni non fisiologiche quando il DNA viene disidratato (si ottiene da una soluzione a più basso contenuto d'acqua, è più compatta in quanto ogni giro d'elica è formato da 11 bp . E' anche la forma favorita in lunghi tratti omopurinici (o pirimidinici), e può quindi comparire anche all'interno di un'elica altrimenti in conformazione B. Il suo solco maggiore è più stretto rispetto alla forma B mentre il solco minore è più largo ma poco profondo) è un'ampia spirale destrorsa, con un diametro di 25,5 A°. In condizioni fisiologiche, questa conformazione caratterizza gli eteroduplex di DNA e RNA e i complessi formati dalle associazioni DNA-proteina.

Globalmente, la conformazione A risulta meccanicamente piú rigida rispetto alla curvatura (per un effetto di decentramento dell'asse rispetto all'impilazione delle basi), ma meno stabile nel DNA, che preferisce quando possibile la forma B.
Tuttavia questa forma metastabile sembra avere un ruolo importante nel conferire al DNA una maggiore versatilità strutturale: compare infatti di preferenza quando il DNA partecipa ad una qualche funzione e quindi interagisce con proteine o RNA (ad esempio nelle eliche ibride DNA-RNA).
Inoltre la conformazione B è quella preferita del DNA nelle triple eliche in cui una terza catena si sistema nel **major groove** stabilizzata da accoppiamenti di Hoogsten o nei tratti di quadrupla elica.



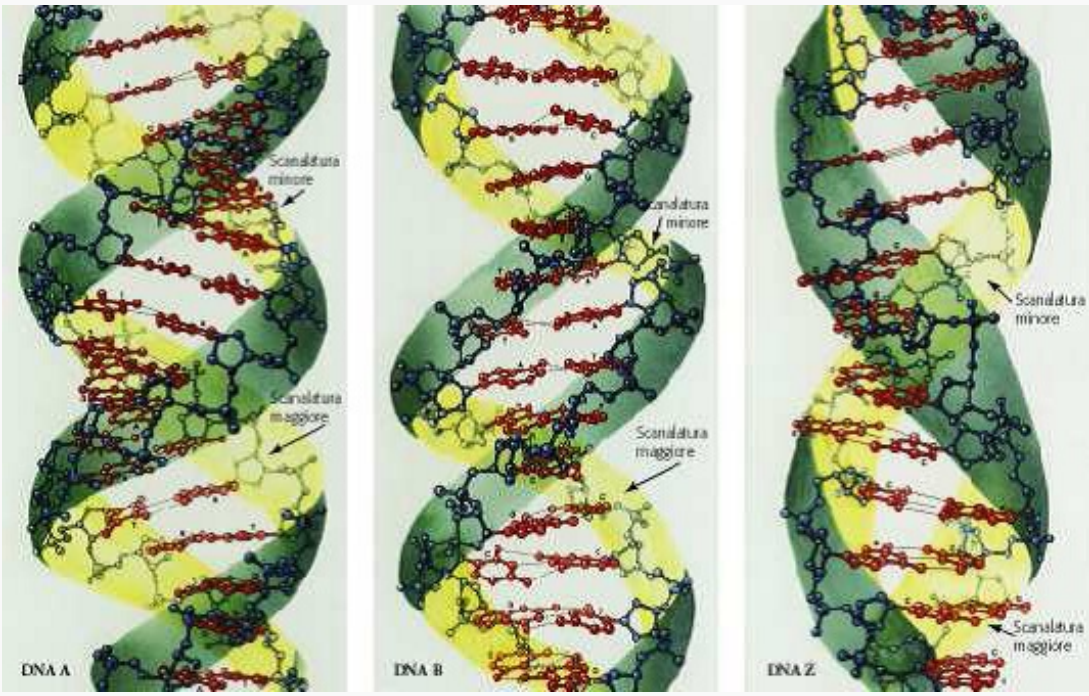
difrattogramma a raggi X di A-DNA (disidratato)



conformazione nel senso radiale ed assiale di DNA-A, DNA-B e DNA-Z

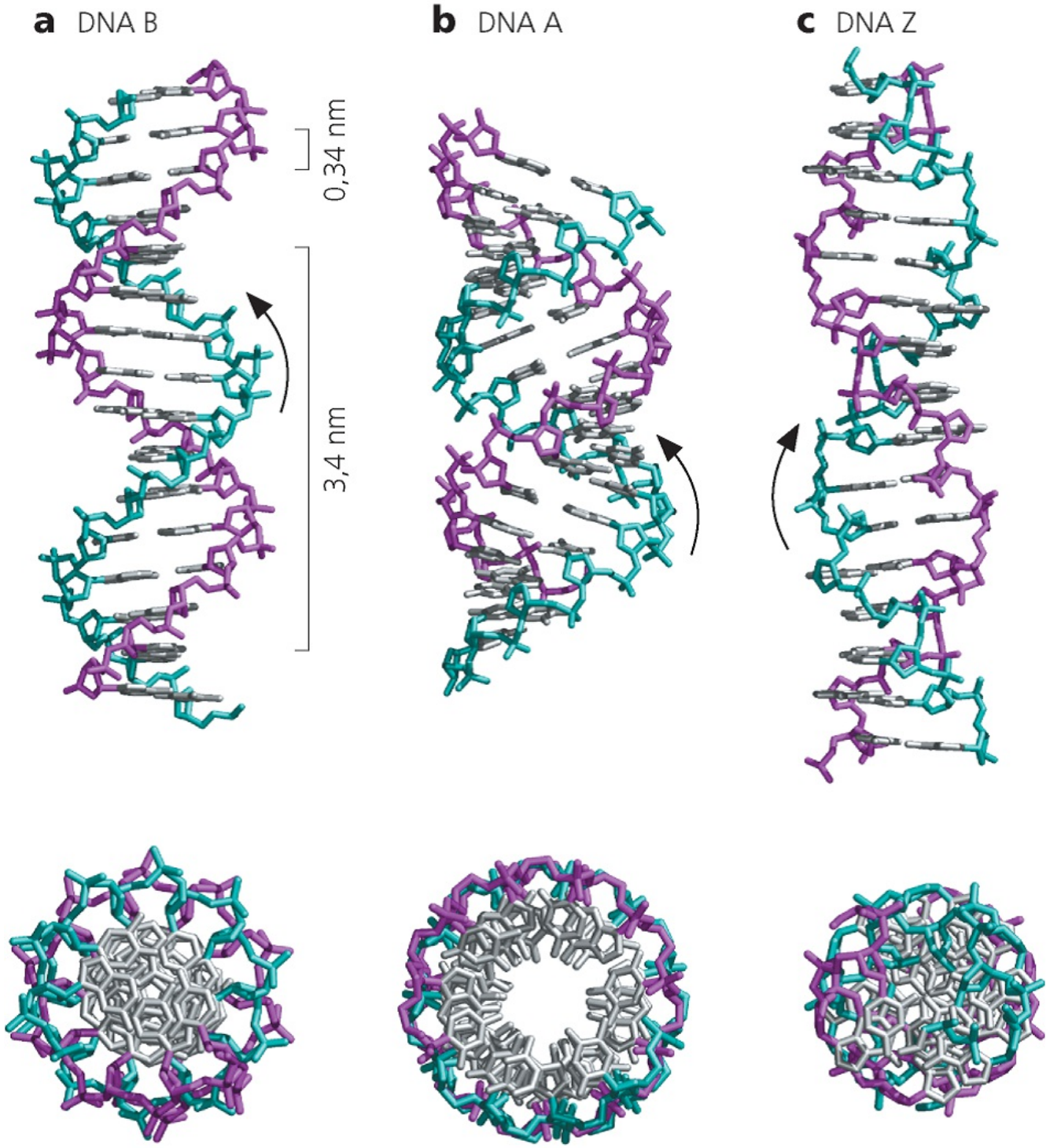
Le conformazioni A e B sono però in ultima analisi associate a due distinte conformazioni degli zuccheri,

rispettivamente C'3- endo e C'2- endo che determinano diverse conformazioni dello scheletro della catena e in alcuni casi diverse orientazioni delle basi e quindi forme diverse dell'elica.
La conformazione Z è tipica invece delle sequenze che presentano modificazioni chimiche come la metilazione e dei tratti di DNA ricchi di basi C e G. Ha un passo di 45,6 A° nel quale trovano posto 12 bp ed un diametro di 18,4 A°. Il **DNA-Z** assume una configurazione sinistrorsa anziché destrorsa .



confronto fra DNA-A (sinistra), DNA-B (centro) e DNA-Z (destra) in rapporto alla forma della doppia elica e relativi solchi

In altre parole, mentre la forma DNA-B è quella del DNA in condizioni "statiche", la sua forma metastabile DNA- A compare in condizioni non standard quando il DNA partecipa a qualche processo di trasformazione o di interazione con altre molecole. Infine, la forma DNA- Z è tipica di tratti di catena a composizione purinica pirimidinica alternata.
La sua unità ripetuta è il dimero purina – pirimidina e le sue caratteristiche strutturali (elicità sinistrorsa, piccolo diametro, scheletro a zig-zag e orientazione syn delle purine sono piuttosto peculiari. Il loro ruolo non è ancora del tutto chiarito, anche se è noto che DNA – Z compare solo in locazioni specifiche del genoma e viene riconosciuto da specifiche proteine.



L'orientamento delle basi rispetto agli anelli di deossiribosio.

1. Nelle forme A e B tutte le basi sono in orientamento anti. Nel DNA-Z tutte le pirimidine sono in orientamento anti mentre le purine sono sin.
2. Il legame della base al deossiribosio può presentarsi nella conformazione sin e anti. Nella forma destrogira del DNA le basi sono sempre in conformazione anti.
3. Nella struttura Z del DNA si osserva un'alternanza di purine e pirimidine dove le purine (G) assumono una conformazione sin mentre le pirimidine (C) hanno la conformazione canonica anti provocando un pattern a zig-zag dei gruppi fosfato da cui il nome Z-DNA

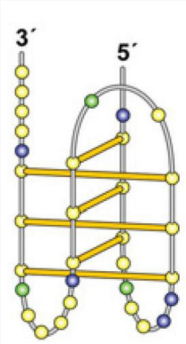
Sono state osservate inoltre anche altre forme di DNA, come quelle dette C e D, che si presentano come fibre. Le fibre sono organizzazioni così chiamate per via della struttura tubolare avvolta, simile a quella di un tessuto. Esse sono ottenute tramite azione meccanica sugli estratti di DNA da cellule viventi, con la conseguente perdita dell'informazione biologica della macromolecola. Si formano strutture completamente distorte da quelle di partenza ma potenzialmente utili, quanto meno in passato, per uno studio strutturale su questi sistemi

macromolecolari.

La forma C si forma in condizioni di relativa bassa umidità per una fibra di DNA ed è molto simile alla forma B della macromolecola, avente però 9.3 residui per giro dell'elica. Il DNA D è invece caratterizzato da un'alternanza di purine (solitamente adenosina ed inosina) e pirimidine (generalmente timina) e presenta una struttura piuttosto complessa. Diverse altre forme di DNA a fibre sono state identificate.

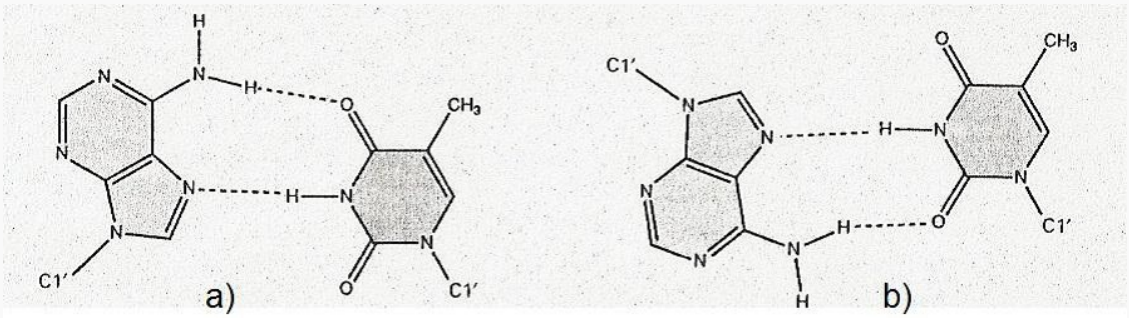
Il modello Watson-Crick del DNA non esaurisce tuttavia le possibili organizzazioni che questa polimorfica macromolecola può assumere. Un primo esempio si ha in ambiente leggermente più acido di quello cellulare con la formazione dell' **i-motif** . In esso si ha la formazione di accoppiamenti C≡G◦C+ con un residuo di citosina protonato. È interessante notare come la tripla C≡G◦C+ si formi soltanto quando la citosina è protonata, per cui queste strutture sono più stabili in ambiente acido. Misure sperimentali hanno fornito il valore della pKa della citosina coinvolta nell' i-motif pari a 7.5 o maggiore, decisamente più elevato di una citosina non impegnata in questo tipo di legami, avente pK_a = 4.2.

*Struttura dell'**organizzazione i-motif**. I filamenti nucleotidici sono schematizzati come strutture tubolari bianche, mentre le basi sono colorate in modo diverso (giallo deossicitidilato, blu deossiadenilato e verde timidilato).*



i-motif

Le tre forme di DNA A, B e Z sono dette duplex, in quanto coinvolgono due filamenti nucleotidici. Sono state osservate però strutture di tipo triplex e quadruplex, nelle quali rispettivamente tre e quattro tratti deossiribosio-fosfodiestere (strands) interagiscono tra loro. Alla base di queste vi è la formazione di legami ad idrogeno diversi rispetto al caso del DNA duplex, in particolare con i gruppi funzionali localizzati nel solco maggiore delle strutture a due filamenti. Queste interazioni deboli, non previste nel modello di Watson e Crick, sono note col nome di appaiamenti di Hoogsteen , descritti per la prima volta nel 1963.



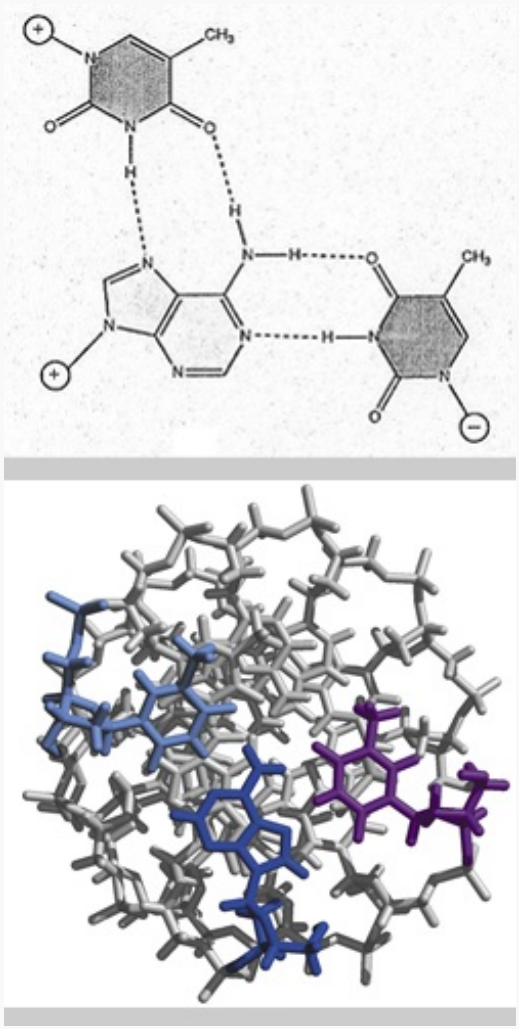
Accoppiamenti di legame ad idrogeno per le coppie di basi adenina e timina di tipo Hoogsteen (a) ed Hoogsteen inverso (b

Considerando il DNA non duplex possiamo quindi individuare accoppiamenti diversi da quelli classici A≡T e G=C, ad esempio gli elementi ternari T=A◦T e del DNA triplex.
Arrangiamento di legame ad idrogeno per gli elementi ternari T=A◦T del DNA triplex. I segni indicano la direzione dei filamenti delle basi

Infatti, in alcune situazioni, è possibile che alla canonica doppia elica di DNA si affianchi , per brevi tratti, un ulteriore filamento. In questo caso viene a configurarsi un frammento di DNA a tripla elica o DNA triplex. Normalmente il DNA a tripla elica o DNA triplex si osserva solamente nei grandi solchi di DNA ricchi di G e di A che si appaiano in maniera anomala mediante i legami o appaiamenti di Hoogsteen (legami che si formano quando un singolo filamento di DNA si associa ad un doppio filamento di B-DNA)

Karst Hoogsteen nel 1963 formula una nuova forma di accoppiamento delle basi nel DNA triplex, (**legami di Hoogsteen**) che consente la formazione della tripla elica in tratti omopurinici e omopirimidinici (AG o TC). In questo modo, due basi azotate, possono essere tenute insieme da legami idrogeno nel solco maggiore. L'**appaiamento di Hoogsteen** per A-T prevede la base purinica ribaltata con l'-NH_2 in posizione 6 e l'N in

posizione 7 dell'adenina che formano legami idrogeno rispettivamente con l'O in posizione 4 e l'H in posizione 1 della timina. (L'N-7 dell'adenina diventa accettore e l'NH2 – 6 donatore per la formazione del legame idrogeno, portando così ad una struttura contorta con disposizione non lineare). Allo stesso modo, il ribaltamento si può verificare anche per G–C. In questa coppia di basi l'appaiamento di Hoogsteen prevede l'O in posizione 6 e l'N in posizione 7 della guanina a formare dei legami idrogeno rispettivamente con l'-NH_2 in posizione 4 e l'N protonato in posizione 1 della citosina. Inoltre, mentre nella coppia di basi Watson–Crick il legame glicosidico di entrambi i nucleotidi ha la conformazione anti, nella coppia di basi Hoogsteen il legame glicosidico del nucleotide purinico ha la conformazione sin.



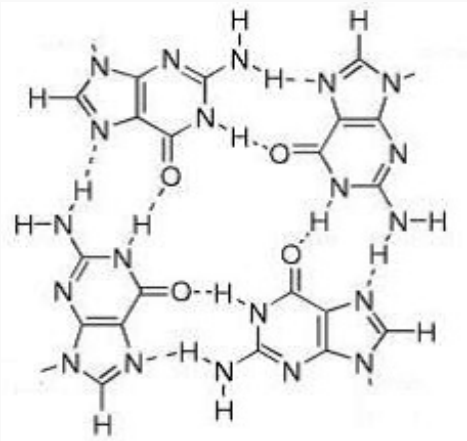
DNA triplex. In alto: legami idrogeno fra 3 basi azotate; in basso: la loro posizione in vista radiale rispetto al filamento

Le prime indicazioni, indirette circa la possibilità di formazione di un tratto di DNA a tripla elica risale a studi sulla densità ottica con miscele di poli U e poli A condotti da **B. Felsenfeld** e **A. Rich** nel 1957. Essi stabilirono che filamenti di acidi poliuridinici [poli(U)] e poliadeninici [poli(A)] erano capaci di formare un complesso stabile a tripla elica in un rapporto 2:1. In seguito, fu evidenziato che anche i filamenti di **acidi polipirimidinici** (dT, dC) e **polipurinici** (dA, dG) erano in grado di dar luogo allo stesso tipo di strutture.

Il DNA triplex si forma più facilmente in presenza di lunghe sequenze costituite solo da purine o soltanto da pirimidine. A questo riguardo sono importanti le combinazioni con due filamenti pirimidinici ed uno purinico, ed un filamento pirimidinico e due purinici. Questa organizzazione secondaria fu la prima scoperta che si distingueva da quella proposta da Watson e Crick. Dopo la sua scoperta, Hoogsteen descrisse dettagliatamente le interazioni deboli specifiche del triplex, evidenziando come queste si sviluppassero a partire dagli atomi non coinvolti negli accoppiamenti del DNA duplex, in particolare per quelli appartenenti alle purine.

Possono esistere anche sequenze ricche in residui guanosinici che possono formare **G-tetraplex**.

Il **DNA quadruplex** si basa anch'esso sulla formazione di interazioni deboli non classiche. L'esistenza di tali sistemi è stata verificata con varie tecniche, tra cui diffrattometria a raggi X, spettroscopia NMR e spettrometria di massa elettrospray, oltre ad essere stata osservata al microscopio a forza atomica (AFM) . L'unità fondamentale è un accoppiamento di 4 residui di guanina e perciò si parla a riguardo di tetraplex o quadruplex G. Solo sequenze ricche in queste basi azotate possono organizzarsi in queste strutture.

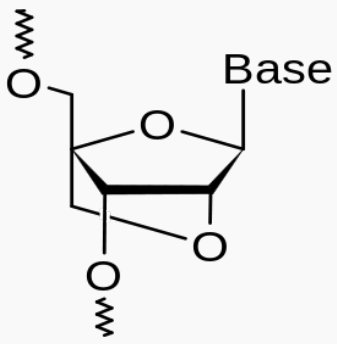


Accoppiamenti, tramite legami ad idrogeno, di 4 residui di guanina che originano quadruplex G

Quindi, solo le sequenze di acidi nucleici ricche in guanina sono in grado di formare strutture a quattro filamenti note come G-quadruplex (o G-tetradi o G4-DNA). Tali strutture sono generate attraverso legami idrogeno tra quattro guanine (che formano la cosiddetta tetrade) e stabilizzate attraverso la presenza di un catione monovalente (in genere potassio) che si dispone al centro della tetrade stessa.

I G-quadruplex possono essere costituiti da DNA, RNA, LNA (locked nucleic acid : nucleotide a RNA modificato utilizzato per incrementare la sensibilità e la specificità dell'espressione genica nei DNA microarray e nella PCR

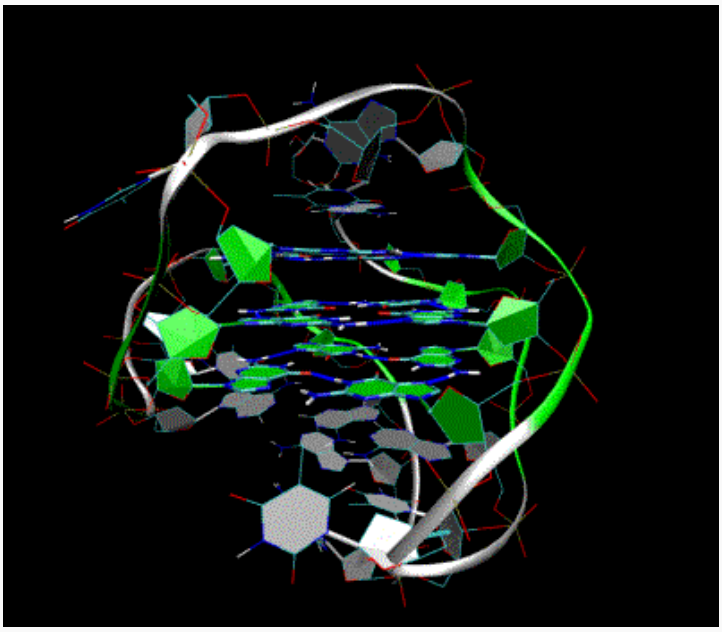
quantitativa. La struttura del ribosio nel nucleotide LNA viene modificata con un ponte supplementare che collega l'Ossigeno in posizione 2 al Carbonio in posizione 4) e PNA o acido peptidonucleico (polimero organico simile al DNA ed all'RNA ma differente da essi nella composizione dello "scheletro".



monomero di LNA

Il PNA non è presente in natura ma viene sintetizzato artificialmente ed usato in ricerche biologiche e trattamenti medici. È stato anche ipotizzato che le prime forme di vita sulla terra utilizzavano il PNA come materiale genetico, grazie alla sua robustezza, passando poi ad un sistema basato su DNA ed RNA) .
Struttura del PNA
Possono essere intramolecolari, bimolecolari o tetramolecolari. A seconda della direzione dei filamenti, essi possono essere definiti paralleli o antiparalleli.

Le ripetizioni telomeriche, presenti in numerosi organismi, si dispongono a G-quadruplex in vitro (e, in alcuni casi, anche in vivo). I telomeri umani e di tutti i vertebrati consistono nella ripetizione della sequenza TTAGGG e la conseguente struttura è stata estensivamente analizzata attraverso cristallografia a raggi X.
Il **tetraplex G** è stabile in una relativamente ampia gamma di condizioni di pH, forza ionica e densità della soluzione acquosa.
Strutturalmente queste organizzazioni si presentano come impilamento (π – π stacking tra le basi) di più piani di quartetti di guanina poste ai vertici di un quadrato ideale, con i lati che schematizzano le interazioni deboli tra le basi di uno stesso piano per via dei legami ad idrogeno. Rispetto al modello del DNA duplex di Watson-Crick si ha quindi un sistema supramolecolare con una superficie di interazione per π – π stacking tra le basi di diversi piani ben maggiore.


struttura tridimensionale di un G-quadruplex intramolecolare telomerico umano in soluzione contenente potassio

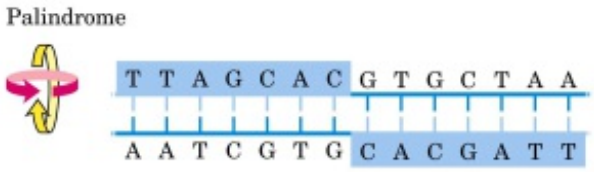
Oggi l'interesse della comunità scientifica è rivolta principalmente ai **G-quadruplex** alternativi ai telomeri, i cosiddetti Quadruplex non telomerici (analisi di quei geni che presentano strutture a G-quadruplex nei rispettivi promotori, come ad esempio quelli della β- globina di pollo, della ubiquitina-ligasi RFP2 umana ed dei protooncogeni c-kit (proto-oncogene che codifica per un recettore transmembranario ad attività tirosin-chinasica) bcl-2 (prototipo di una classe di geni che codificano proteine che governano la permeabilità della membrana mitocondriale esterna) VEGF (gene che codifica il fattore di crescita dell'endotelio vascolare o VEGF coinvolto sia nella vasculogenesi cioè nella genesi ex novo di un sistema circolatorio in età embrionale che nell'angiogenesi ossia la formazione di vasi da strutture già esistenti), H-ras e N-ras (i ras sono una famiglia di geni che codificano una classe di proteine chiamate piccole small GTPase coinvolte nella trasmissione di segnali all'interno delle cellule cioè nella trasduzione del segnale cellulare).

SEQUENZE DI DNA CHE ADOTTANO STRUTTURE INSOLITE

Sequenze particolari di DNA possono originare singolari organizzazioni di doppie eliche. L'esempio tipico è quello dei segmenti nucleotidici palindromi, il cui codice può essere letto indifferentemente sia in un verso che nell'altro,

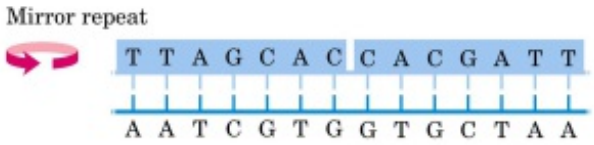
i quali possono formare le cosiddette strutture a croce od a forcina.

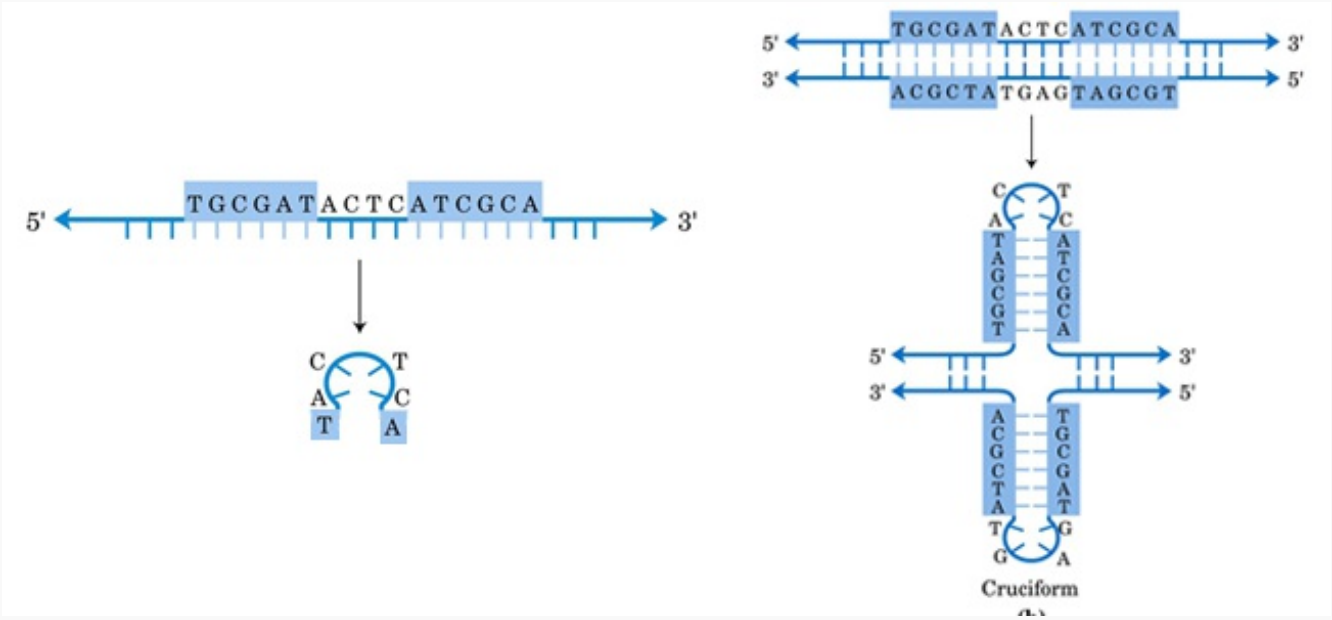
Sequenza palindromica
Un palindromo è una parola che risulta la stessa letta in una direzione o in quella inversa (es.: ossesso). Il termine viene applicato alle regioni di DNA di in cui vi sono ripetuti invertiti di una sequenza di basi con una doppia simmetria presente nelle due catene del DNA



Quando la sequenza è presente in ciascuna catena, la sequenza viene detta **ripetuto speculare**.

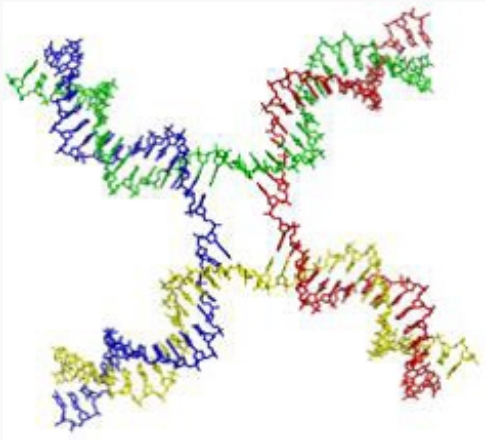
Un palindromo può formare strutture a **forcina** (hairpin) e a forma di **croce**.





struttura a forcina (a sinistra) e a croce (a destra) formata da DNA palindromo

Giunzione di Holliday
Un'altra organizzazione molto interessante è la giunzione di Robin Holliday, presente durante i processi di ricombinazione genetica. Nel 1964 R. Holliday propose un modello di ricombinazione omologa chiamata ricombinazione generale per spiegare i prodotti della meiosi nei funghi



giunzione di Holliday

Robin Holliday *(1932 -2014): biologo molecolare britannico. Ha descritto, nel 1964, una struttura mobile a croce composta da quattro filamenti di DNA nota come " Holliday junction" per spiegare un particolare scambio di materiale genetico che si verifica durante la meiosi nei funghi (Ustilago maydis). La mobilità di tali giunzioni è dovuta al fatto che esse si formano tra sequenze omologhe, che permettono lo scorrimento del DNA (ricombinazione omologa) Per completare l'evento di ricombinazione, le giunzioni devono essere risolte cioè si devono generare almeno due tagli nella struttura per rigenerare due duplex lineari.*

ALTRE FORME ALTERNATIVE DI DNA: TAUTOMERICO E DNA-H

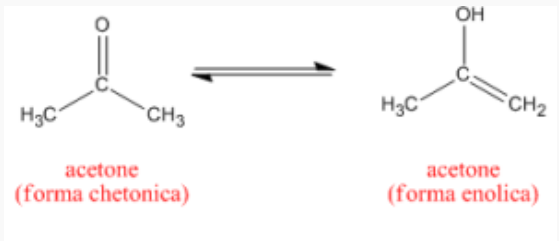
Robin Holliday

DNA tautomerico

Anche se il DNA- B rappresenta la forma predominante di acido nucleico presente nelle cellule eucariotiche, procariotiche e virus. Tuttavia, è possibile anche riscontrare altre forme alternative di DNA che presentano morfologia e proprietà differenti.
La presenza di sequenze più o meno lunghe di DNA alternativo, in qualsiasi sua forma, rappresenta un momento patologico della vita della cellula. I sistemi di controllo riescono, in genere, a correggere queste situazioni in modo efficiente limitando al massimo i danni derivanti da una catena di DNA malformata.
Il DNA tautomerico è formato da nucleotidi alternativi rispetto a quelli “standard”.
La **tautomeria** è una particolare forma di isomeria tra i composti organici per la quale una molecola si presenta sotto due diverse forme pur avendo lo stesso numero di atomi.
Nella maggior parte dei casi la tautomerizzazione comporta un trasferimento di un protone o di atomo di idrogeno, accompagnata dallo scambio di un legame covalente singolo con uno doppio adiacente: si parla in questo caso di tautomeria prototropica; più rari sono i casi di tautomeria aniotropica in cui ad essere scambiato è un gruppo idrossilico. In soluzione, dove è possibile che avvenga la tautomerizzazione, si raggiunge un equilibrio chimicotra i vari tautomeri. L’esatta frazione di ciascun tautomero dipenderà da numerosi fattori quali la temperatura, ilsolvente e il pH.

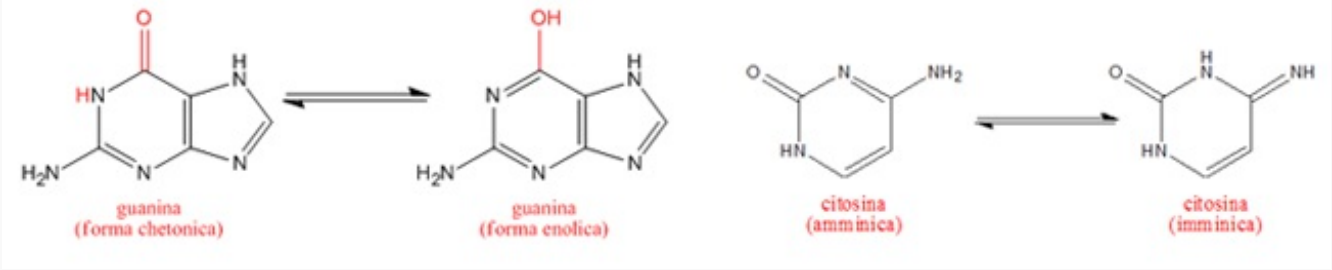
Le principali **coppie tautomeriche** sono:
1. chetone / enolo (tautomeria cheto-enolica)
2. chetene / inolo
3. ammina / immina – tra gli altri casi anche nella base azotata guanina e adenina
4. ammide / acido imminico – che può aver luogo nella citosina o durante la reazione di idrolisi del nitrile
5. lattame / lattime – analogo al precedente che si verifica però in composti eterociclici; ne possono essere soggette le basi azotate guanina, timina e citosina.
6. immina / enammina, che può ad esempio avvenire nell’amminoacido istidina
7. nitroderivati / acidi nitronici, solo per quei nitroderivati con almeno un atomo di idrogeno in alfa al nitrogruppo.

Le aldeidi possiedono un gruppo funzionale -CHO mentre i chetoni -C=O; quando la molecola ha un carbonio in α che lega un atomo di idrogeno, si può verificare il fenomeno della tautomeria. Un esempio molto utilizzato per spiegare questa proprietà molecolare lo possiamo recuperare dallo studio della chimica organica prendendo a modello l’acetone:



tautomeria cheto-enolica acetone

Lo stesso modello di tautomeria può essere applicato alla guanina. Esiste anche un altro tipo di tautomeria che riguarda il gruppo funzionale amminico-imminico che, per quanto riguarda le basi, potrebbe essere illustrata prendendo a modello la citosina.

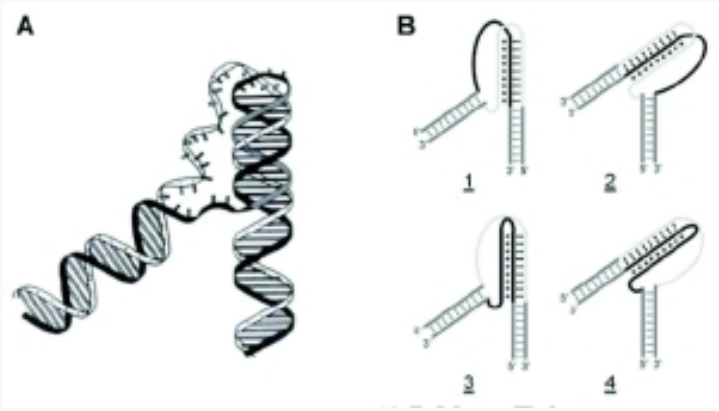


tautomeria sulla guanina e sulla citosina

L’esistenza di forme tautomeriche delle basi rende possibile alcune variazioni, minori, nella forma e nella costituzione del DNA. Alla luce di quanto abbiamo detto prima, riferendoci all’appaiamento tra A-T e G-C, la

presenza di forme tautomere di basi rende possibile varianti che, ad esempio, portano alla formazione di A-C e T-G.

DNA H
Un'organizzazione insolita del DNA, nota con il nome di **DNA H**, si può formare in tratti di polipirimidine/polipurine che contengono anche all'interno della sequenza ripetuti speculari.
Questa struttura produce un brusco ripiegamento del DNA che porta ad un superavvolgimento particolare, a forma di H. Il DNA H viene definito anche **intramolecular triplex**.



DNA-H o intramolecular triplex

LE CLASSI DEL DNA

Nei cromosomi il DNA non presenta un'organizzazione uniforme ma è composto da diverse classi con specifiche caratteristiche:

– **DNA ripetuto in tandem**: Il genoma umano contiene moltissimi cluster (raggruppamenti) di sequenze ripetute di DNA che non codificano per nessuna proteina
A seconda dell'estensione di questi blocchi di DNA ripetuto possiamo distinguere:
a) DNA satellite: da 5 a 171 nucleotidi dell'unità ripetuta;
b) DNA minisatellite: da 6 a 64 nucleotidi dell'unità ripetuta;
c) DNA microsatellite: da 1 a 4 nucleotidi dell'unità ripetuta.

– **DNA ripetuto intersperso**: le unità ripetute non sono concentrate in specifiche regioni bensì disperse lungo i cromosomi. Costituiscono circa un terzo del genoma umano e in base alla lunghezza dell'unità ripetuta possiamo distinguere 2 classi:
1. **Sines** (short interspersed repeated sequences): fino a 300 basi di unità ripetuta;
2. **Lines** (long interspersed repeated sequences): fino a 6400 basi di unità ripetuta.

– **DNA non ripetitivo**: costituisce i geni, le sequenze di DNA che codificano per un mRNA.
a) nelle cellule diploidi esistono due copie di ciascun cromosoma;
b) generalmente è presente una sola copia di ogni gene su ciascuno dei due cromosomi di ogni coppia;
c) gli pseudogeni sono sequenze di DNA che pur mostrando un elevato grado di omologia con un gene funzionale, non sono attivi dal punto di vista trascrizionale.

LE FUNZIONI DEL DNA

A livello biologico il DNA è un componente essenziale della cellula, dal momento che in esso sono contenute tutte le informazioni per generare altre cellule. Dunque una sua funzione importante da considerare nella divisione cellulare è la capacità di replicarsi; infatti la cellula figlia deve avere esattamente lo stesso DNA della madre. Nella divisione cellulare il doppio filamento di DNA si rompe e ognuno dei due filamenti ne sintetizza uno complementare grazie all'enzima DNA polimerasi. Questa sintesi ha una direzione preferenziale, che di solito è da 5′ a 3′.

Per quanto riguarda la determinazione del fenotipo – cioè delle caratteristiche dell'organismo che effettivamente sono espresse e visibili – non tutto il materiale viene codificato, bensì si distingue in sequenze codificanti o esoni, che costituiscono solo il 5% del totale e non codificanti, gli introni. Questi ultimi sembravano non rivestire alcun ruolo rilevante fino a una ventina di anni fa, tanto che venivano definiti "DNA spazzatura", tuttavia si è scoperto che in realtà possiedono una certa importanza in meccanismi di regolazione genica, dal momento che vanno a formare i miRNA. Sequenze di introni ed esoni, mappate precisamente all'interno del codice, vengono dette geni; questi rappresentano le unità ereditarie fondamentali degli organismi e vengono ritrovati anche nell'RNA. Il DNA, mediante un processo di trascrizione viene trasformato in RNA, il quale svolge numerose funzioni tra cui quella di codificare le proteine.