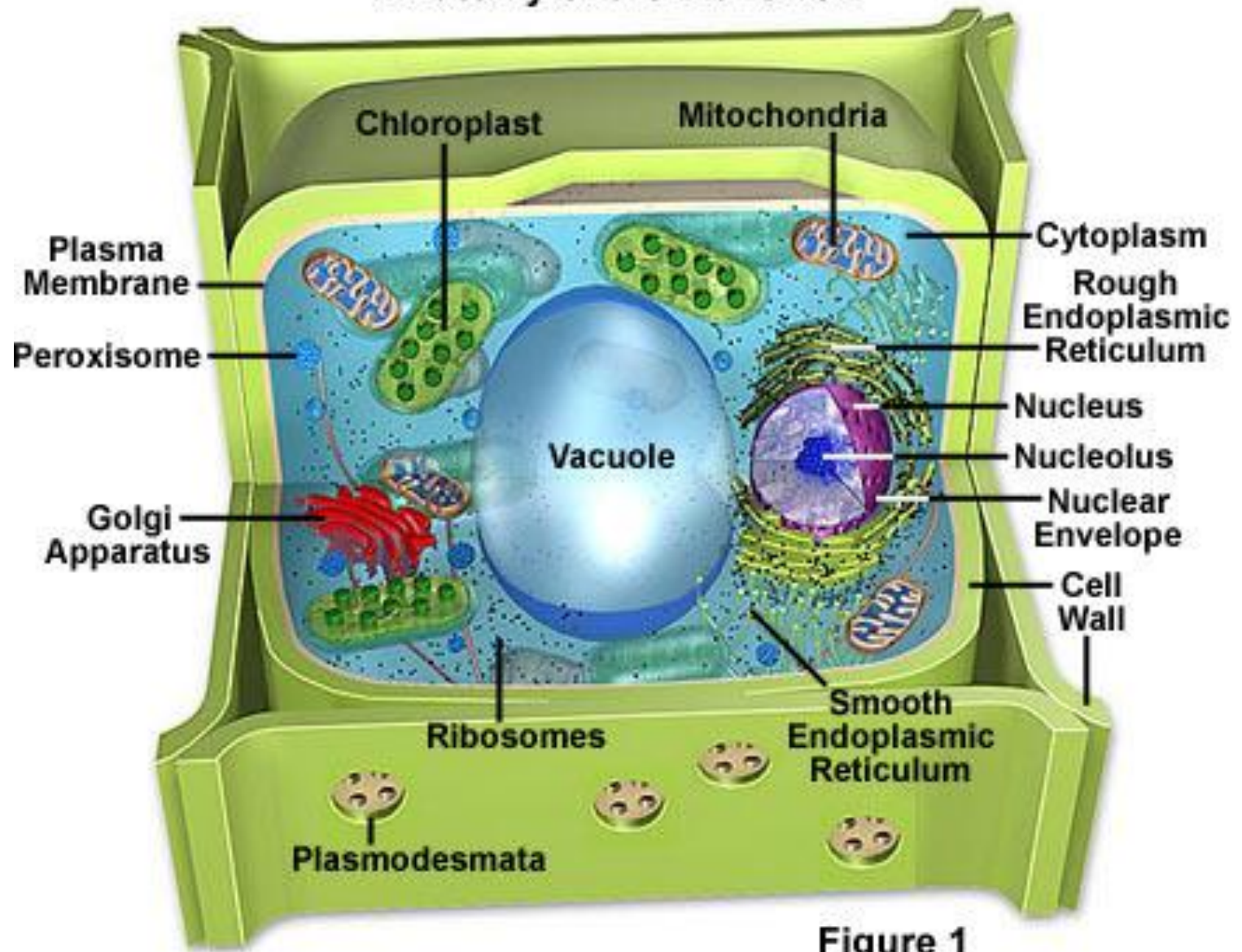


Schema di una cellula animale

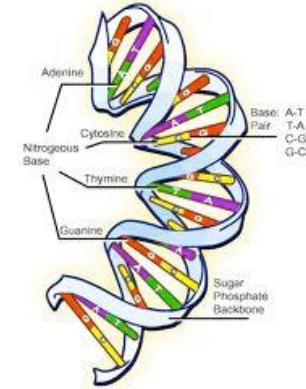
Anatomy of the Plant Cell



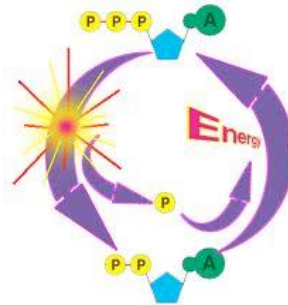
Schema di una cellula vegetale

Funzioni degli acidi nucleici sono conseguenza della loro struttura

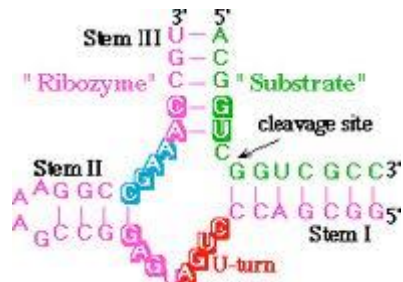
1. Informazione genetica.



2. Trasporto di energia.



3. Attività enzimatica (ribozimi).



Struttura e proprietà chimico-fisiche degli acidi nucleici

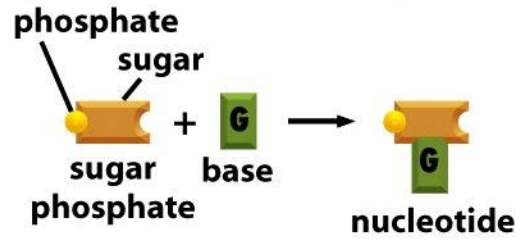
Basi, nucleosidi e nucleotidi

Base	Nucleoside	Nucleotide	RNA	DNA
Adenina	Adenosina	Adenosina-5'-fosfato (Acido adenilico)	AMP	dAMP
Guanina	Guanosina	Guanosina-5'-fosfato (Acido Guanilico)	GMP	dGMP
Citosina	Citidina	Citidina-5'-fosfato (Acido citidilico)	CMP	dCMP
Timina	Timidina	Timidina-5'-fosfato (Acido timidilico)		dTMP
Uracile	Uridina	Uridina-5'-fosfato (Acido uridilico)	UMP	

base azotata + pentoso → nucleoside

nucleoside + 1 o più gruppi fosfati → nucleotide

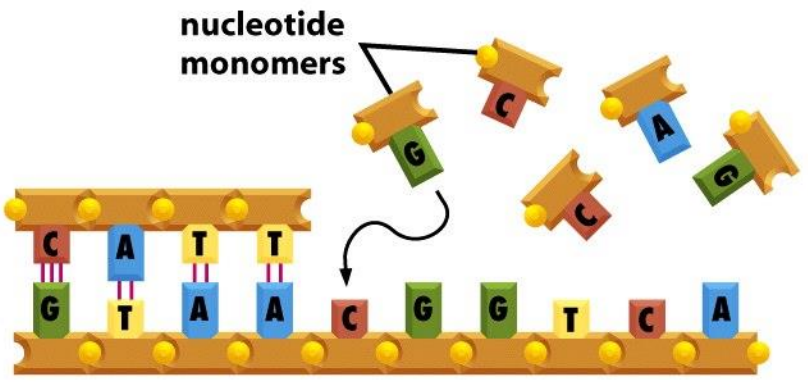
(A) building block of DNA



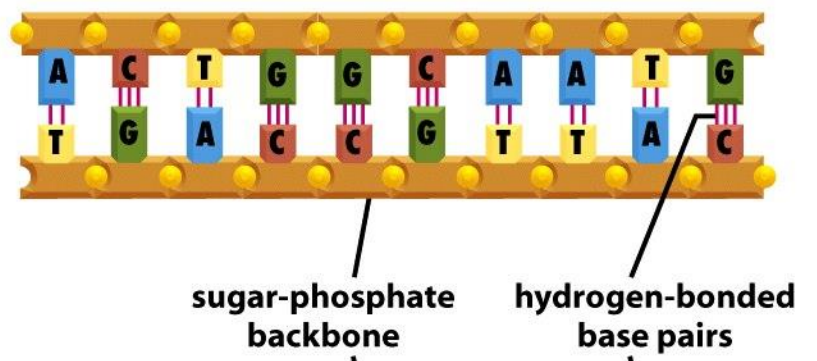
(B) DNA strand



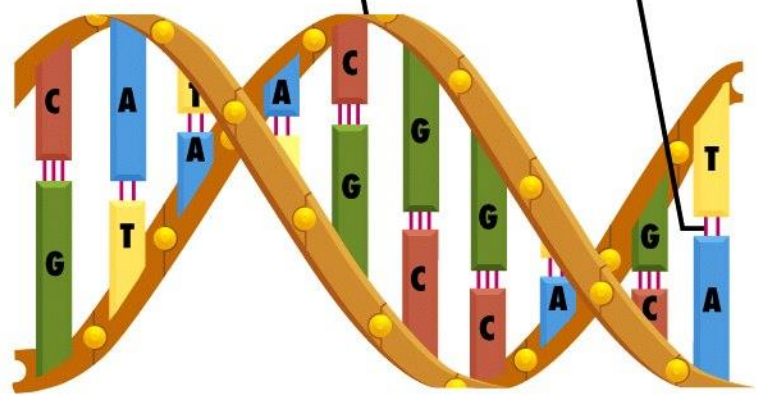
(C) templated polymerization of new strand

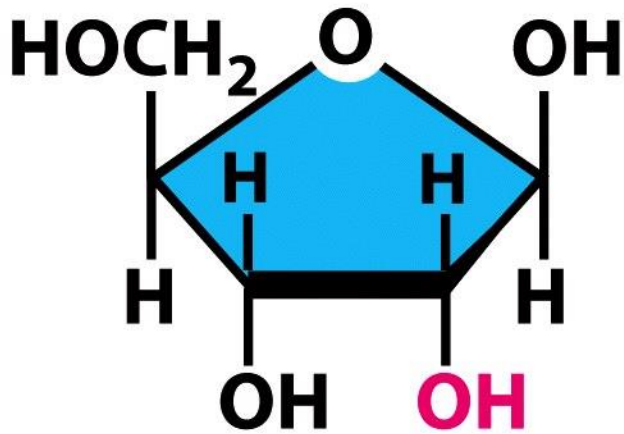


(D) double-stranded DNA



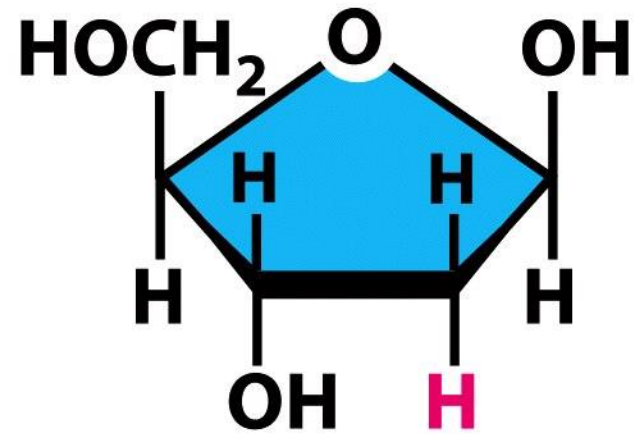
(E) DNA double helix





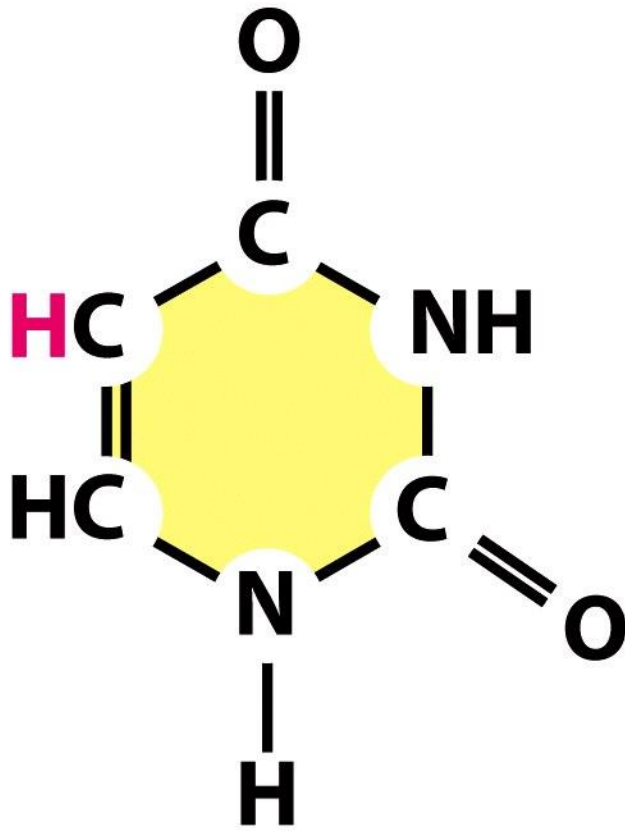
ribose

used in ribonucleic acid (RNA)



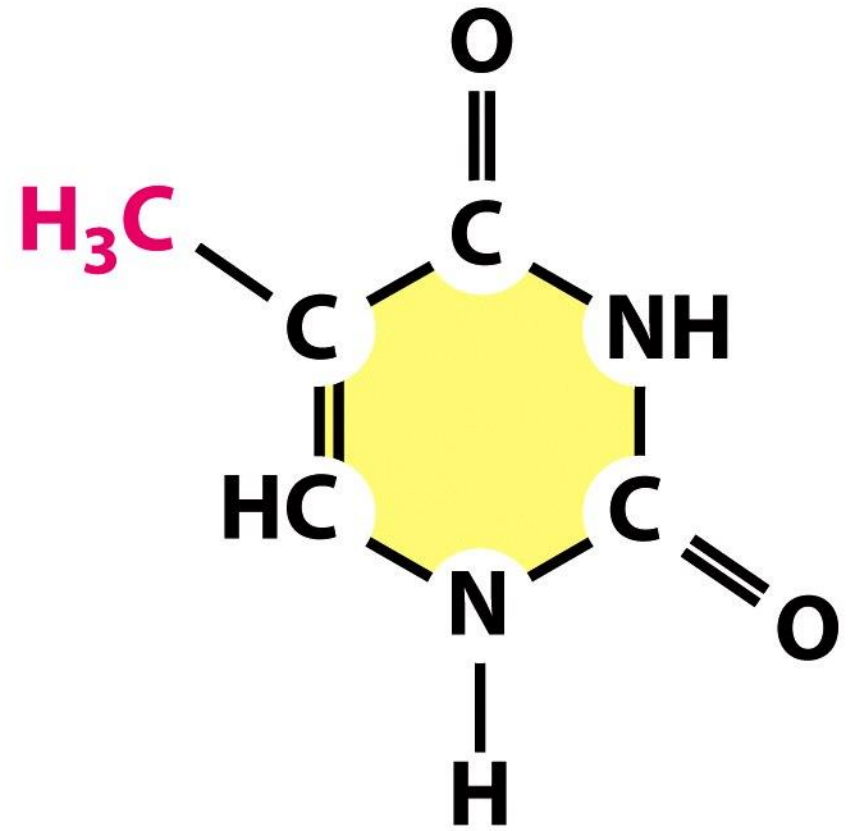
deoxyribose

used in deoxyribonucleic acid (DNA)



uracil

used in RNA



thymine

used in DNA

Simbolo	Base		Nucleoside
A	Adenine		Adenosine
C	Cytosine		Cytidine
G	Guanine		Guanosine
T (U)	Thymine (Uracil)		Thymidine (Uridine)
R	puRine	A o G	
Y	pYrimidine	C o T (U)	
S	Strong	C o G	
W	Weak	A o T (U)	
K	Keto-enolic	G o T (U)	
M	aMino-iminic	A o C	
B	non A	C o G o T (U)	
D	non C	A o G o T (U)	
H	non G	A o C o T (U)	
V	non T (non U)	A o C o G	
N	any	A o C o G o T (U)	
I	Hypoxanthine		Inosine
<i>O*(5mC)Not off.</i>	<i>5-methylcytosine</i>		<i>5-methyl-cytidine</i>

Quanto DNA c'è in una cellula di *Escherichia coli*?

Massa molecolare delle basi:

-Adenina	135 Da
-Citosina	111 Da
-Guanina	151 Da
-Timina	126 Da

Massa molecolare media delle basi 131 Da

Massa molecolare media di un deossinucleotide libero 327 Da

Massa molecolare media di un deossinucleotide in catena 309 Da

Massa molecolare media di un paio di deossinucleotidi in catena 618 Da

Il cromosoma di *E. coli* contiene circa 4.7×10^6 paia di basi,

quindi ha una massa di circa 2.9×10^9 Da, corrispondenti a 4.8×10^{-15} g ($1 \text{ Da} \approx 1,661 \times 10^{-27}$ kg)



Cenorhabdits elegans è un nematode di circa 1 mm. Ha un genoma di 100,3 Mb con 20443 geni per proteine e 1270 per diversi RNA. Il numero dei geni è comparabile a quello umano



Saccharomyces cerevisiae ha 16 cromosomi (~12 Mb) portanti 6607 geni di cui solo 233 presentano introni.

Oryza sativa possiede un genoma piccolo (389 Mb) con un numero molto elevato di geni (56000) superiore a quello umano.



Mus musculus possiede un genoma molto simile a quello dell'uomo (2700 Mb) con i suoi 22000 geni per proteine e 3200 geni che codificano per diversi tipi di RNA (rRNA, tRNA, snRNA e microRNA).

E in una cellula umana?

Un corredo aploide contiene circa 3.2×10^9 paia di basi, una cellula diploide il doppio, quindi una massa di circa 2.0×10^{12} Da, corrispondenti a circa 3.3×10^{-12} g, a cui andrebbe aggiunta una piccola quantità di DNA mitocondriale.

Essendo il numero stimato di cellule in un adulto dell'ordine di 10^{14} , per la gran parte diploidi, ne consegue un contenuto complessivo di DNA dell'ordine di qualche ettogrammo.

Genoma nucleare, che comprende il 99,9995% dell'informazione genetica
Genoma mitocondriale, che copre il rimanente 0,0005%.

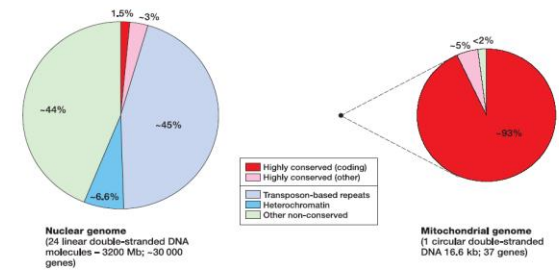
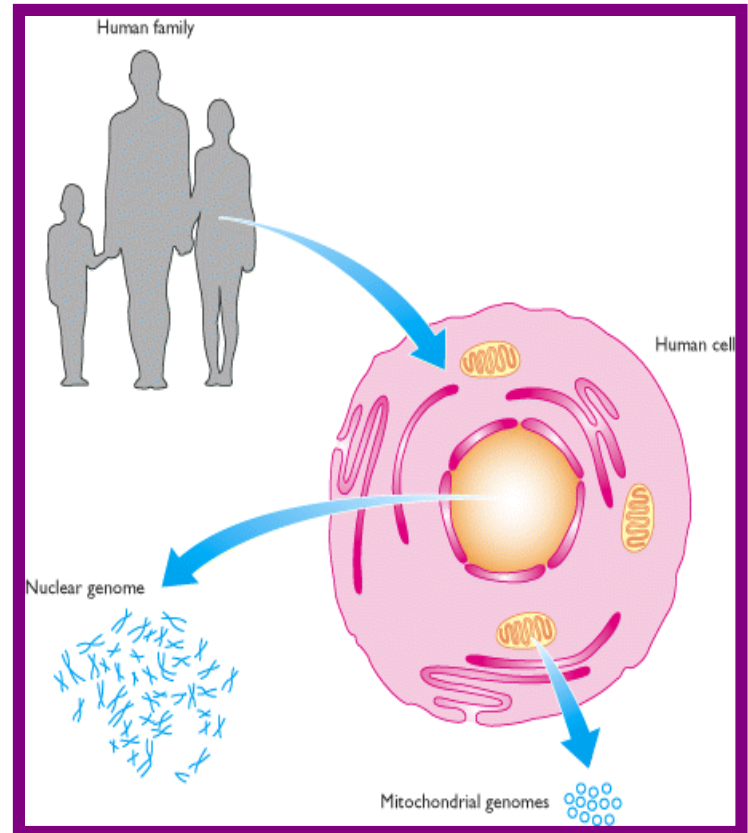
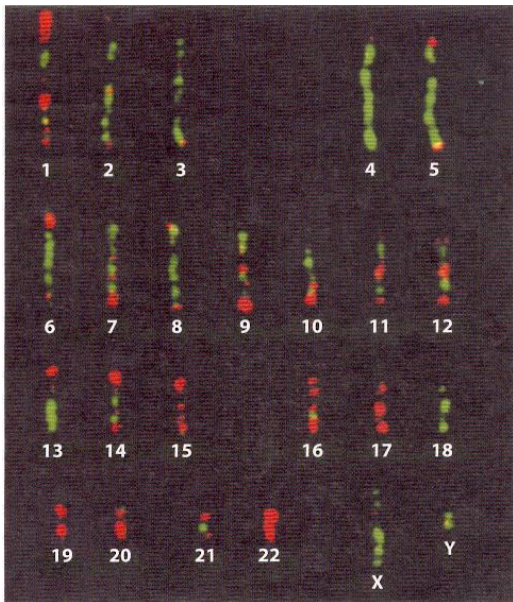
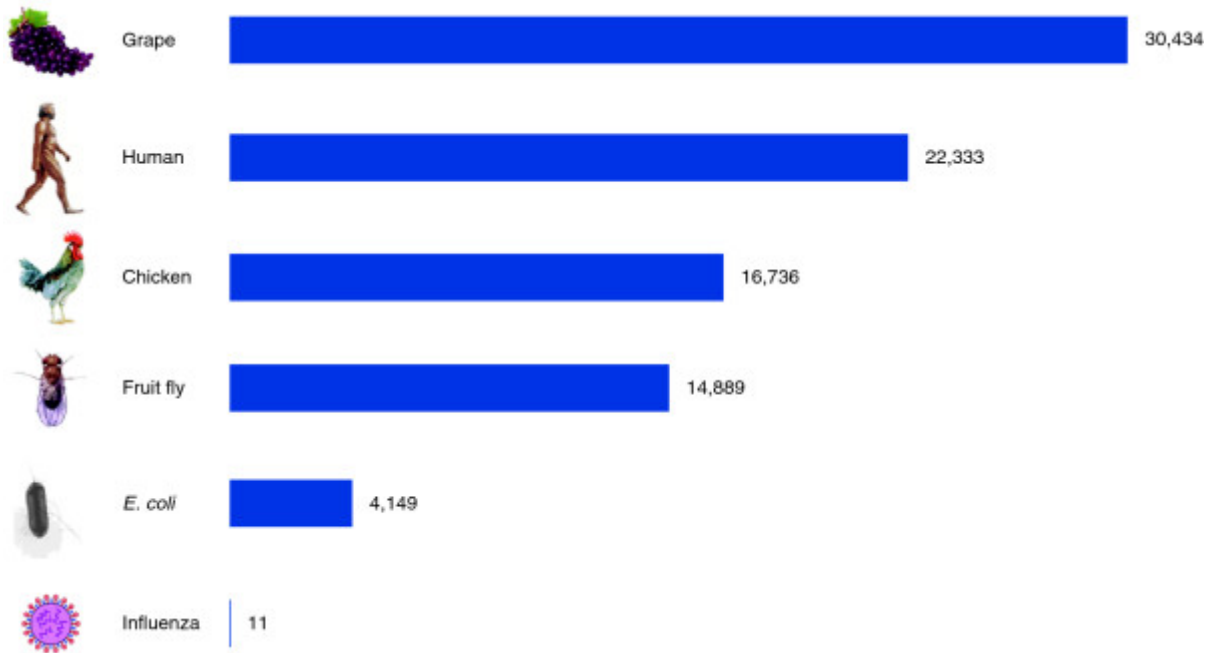


Figure 9.1: Organization of the human genome.

The dot in the middle represents the relative size of the mitochondrial genome using the same scale as for the nuclear genome. Note also the profound difference between the two genomes in the extent of highly conserved DNA (coding sequence, regulatory sequence etc.) and the fraction of highly repetitive noncoding DNA.

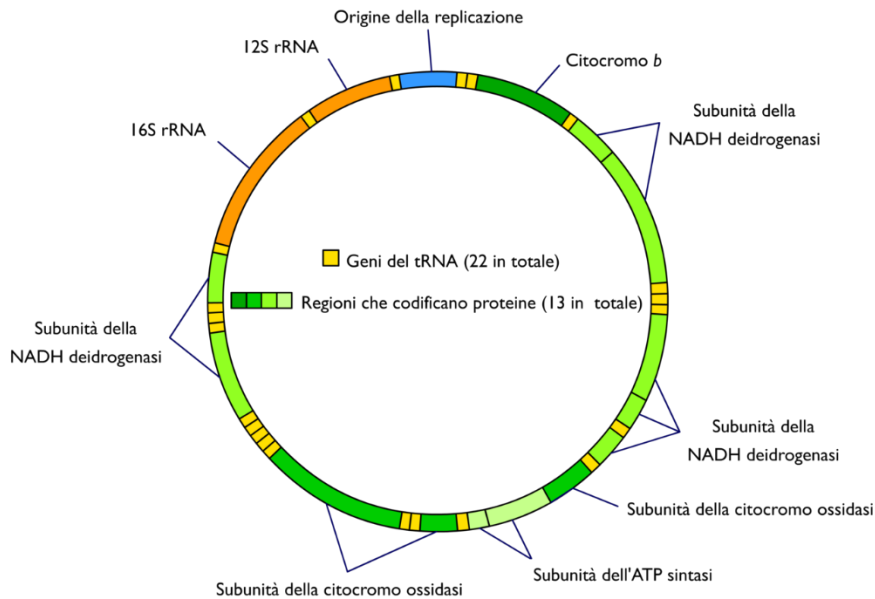


Mappa genica umana

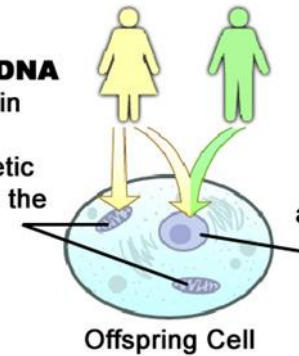
Le regioni in rosso indicano porzioni dei cromosomi ad alta densità genica (ad esempio i cromosomi 15, 16, 17, 19, 20 e 22).

Altri cromosomi come 4, 18, X e Y mostrano una colorazione rossa molto debole e sono poveri di geni.

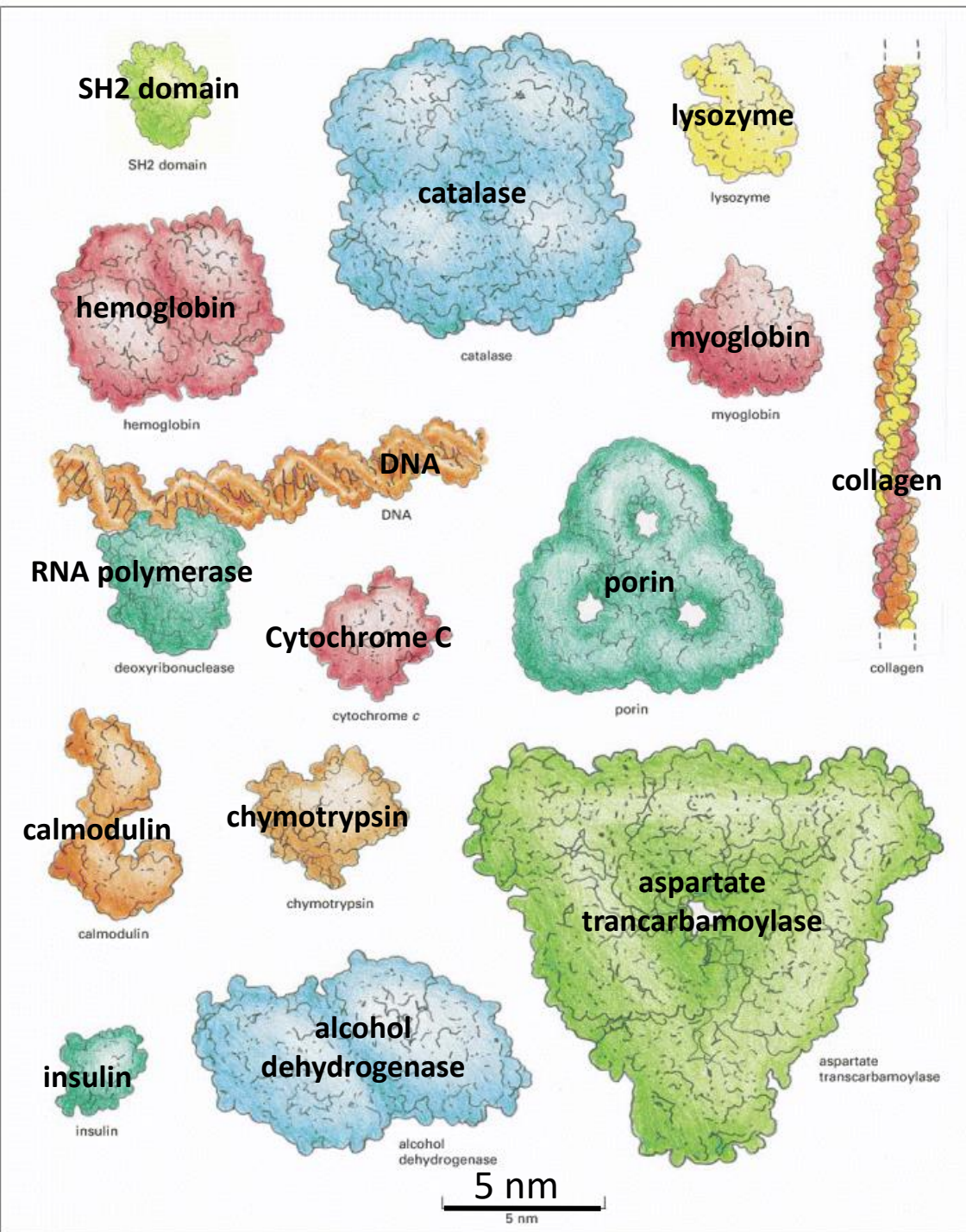
DNA mitocondriale dell'uomo: 16569 paia di basi e 37 geni (codificano per 13 polipeptidi sintetizzati dal ribosoma mitocondriale, 22 tRNA e 2 rRNA), coinvolti nella produzione di proteine necessarie alla respirazione cellulare.



Mitochondrial DNA (mtDNA) is found in cell mitochondria and contains genetic material only from the **mother**.



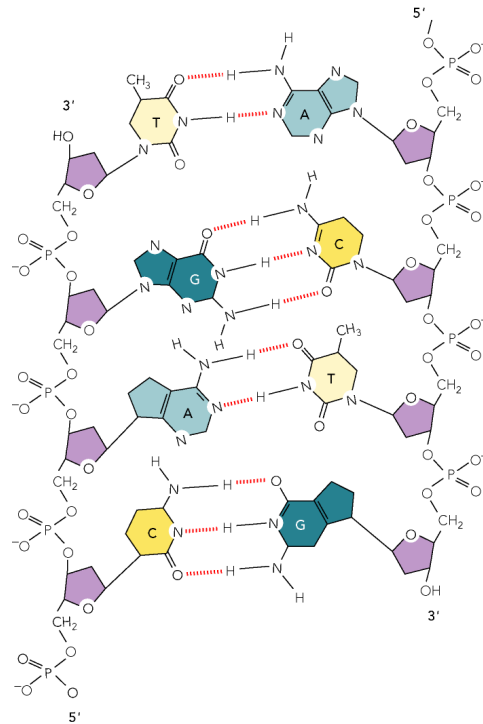
Nuclear DNA (nuDNA) is found in the cell nucleus and contains genetic material from **both parents**.



Strutture delle Proteine

A collection of protein molecules, shown at the same scale.

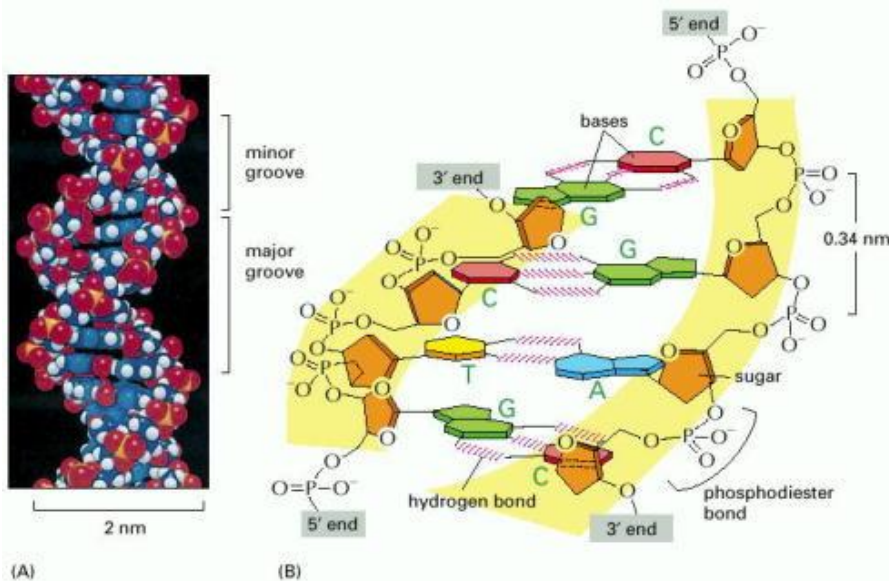
Strutture del DNA



Gli acidi nucleici, come le proteine, sono essenzialmente costituiti dalla successione di unità monomeriche covalentemente legate in una catena non ramificata, e come tali sono biosintetizzati.

Struttura secondaria DNA

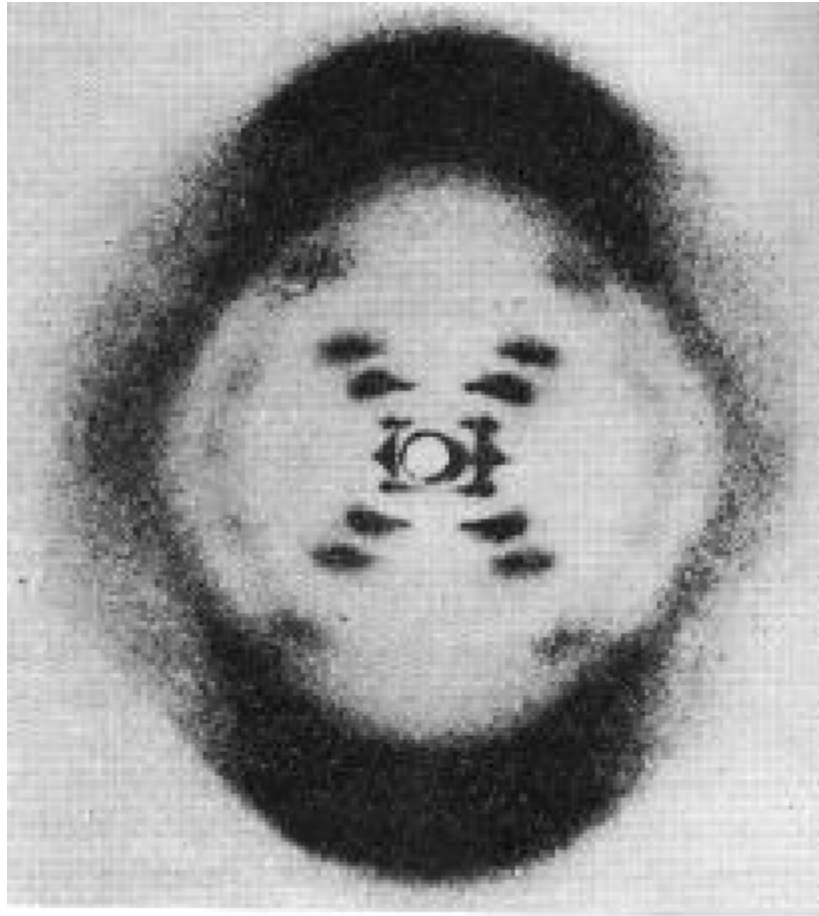
- La struttura secondaria del DNA è la ben conosciuta doppia elica.



E' data dall'avvolgimento tra loro di due catene polinucleotidiche antiparallele collegate da legami idrogeno tra basi complementari (A:T e C:G).
I legami H e le forze idrofobiche di stacking dovute all'impilamento delle basi disposte sostanzialmente piatte una sull'altra conferisce la forte stabilità alla doppia elica.

LA RICERCA SUL DNA

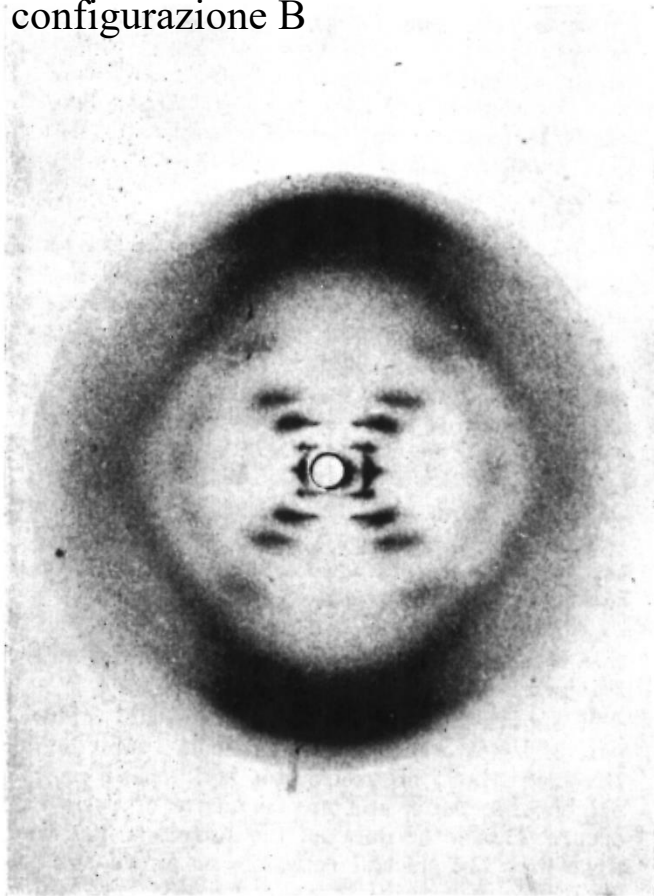
Nel King's College erano iniziate le ricerche sul DNA, acido desossiribonucleico, la componente principale dei cromosomi e quindi dei geni. In poco tempo, Rosalind mise a punto una tecnica innovativa che utilizzava i raggi X per fotografare i costituenti di tutti i materiali viventi e non viventi.



Il dispositivo consisteva in una microcamera capace di produrre fotografie ad alta definizione dei singoli filamenti del DNA. La Franklin riuscì dunque a fare la prima fotografia dello scheletro del Dna...

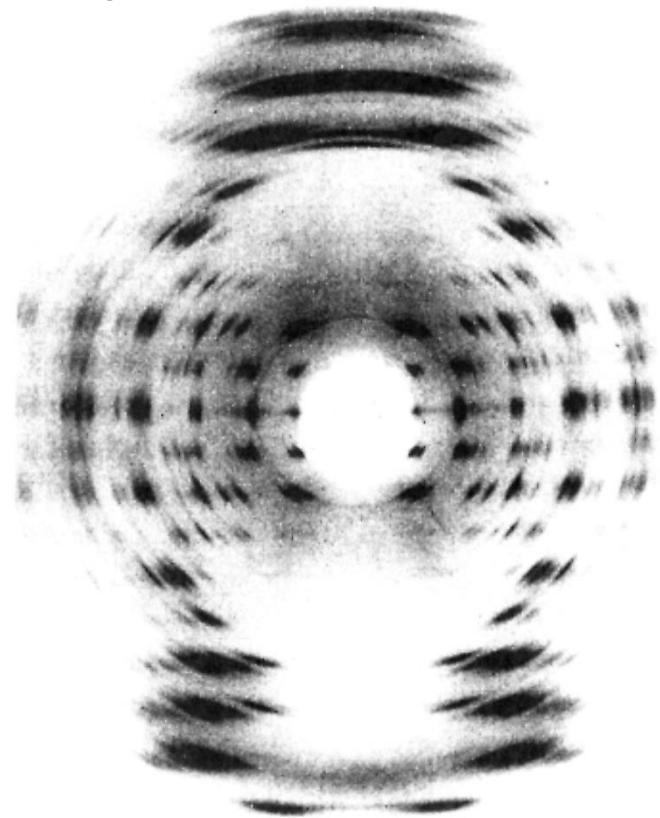
Diffrazione dei raggi X di fibre di DNA

configurazione B



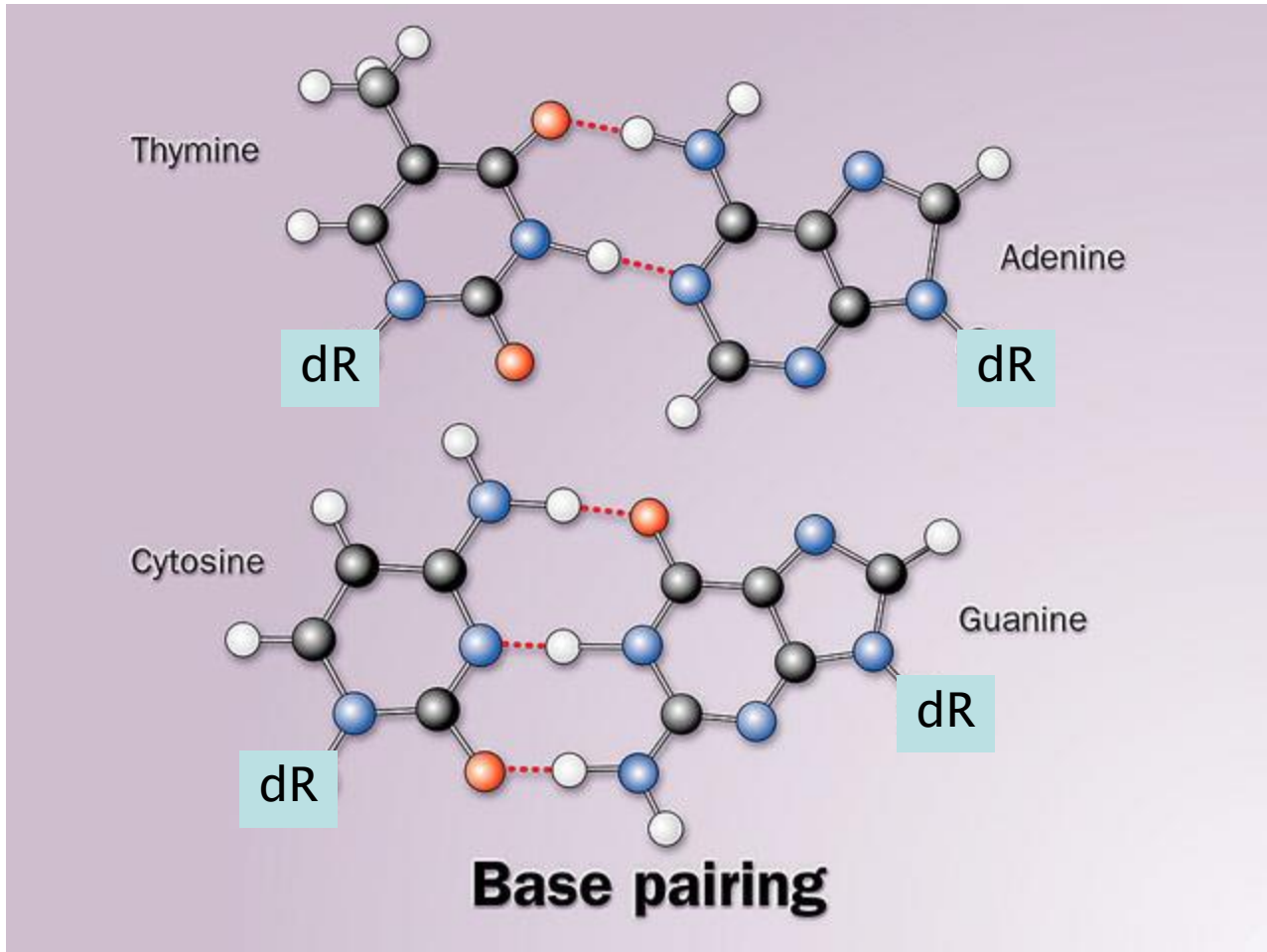
Franklin R.E. & Gosling R., 1953

configurazione A



Wilkins M.H.F., 1956

Appaiamenti di Watson e Crick: coppie isomorfe

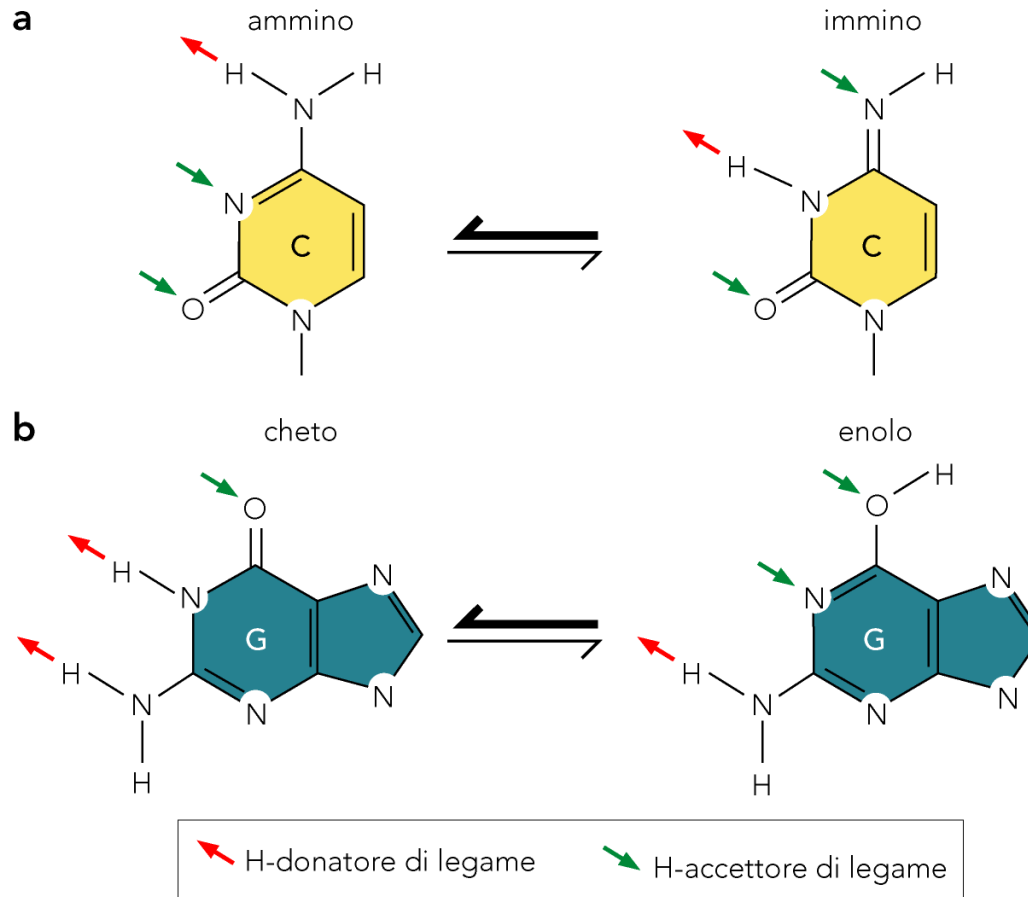


dR = deoxyRibose

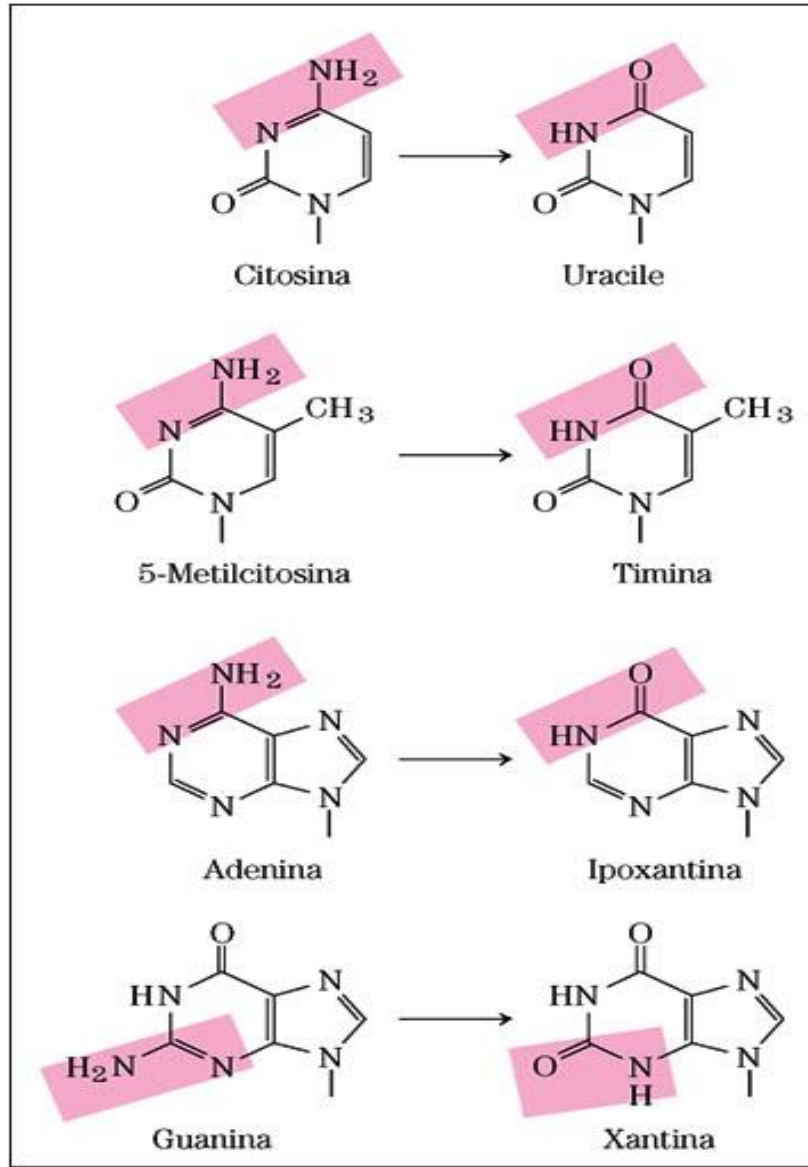
Conformazioni tautomeriche delle basi

- a) Equilibrio ammino = immino nella citosina;
- b) equilibrio cheto = enolo nella guanina.

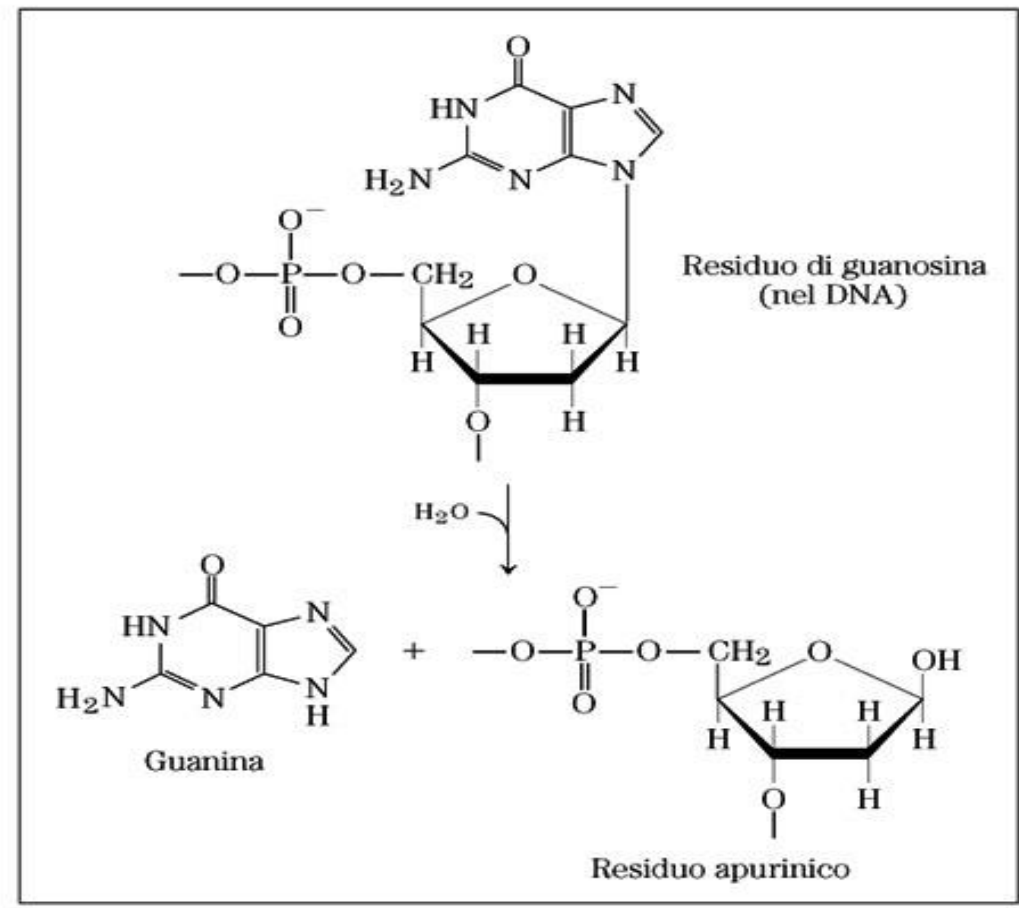
Le forme più stabili amminiche e chetoniche sono quelle prevalenti nella doppia elica.



Trasformazioni non enzimatiche dei nucleotidi



(a) Deaminazione

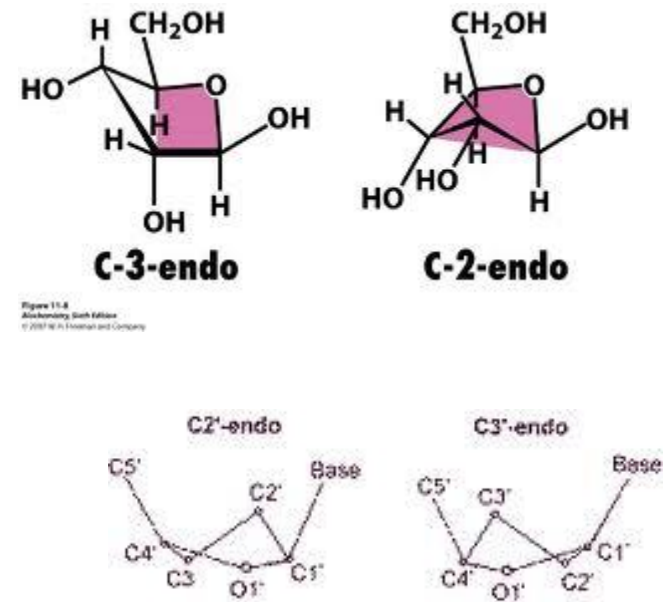
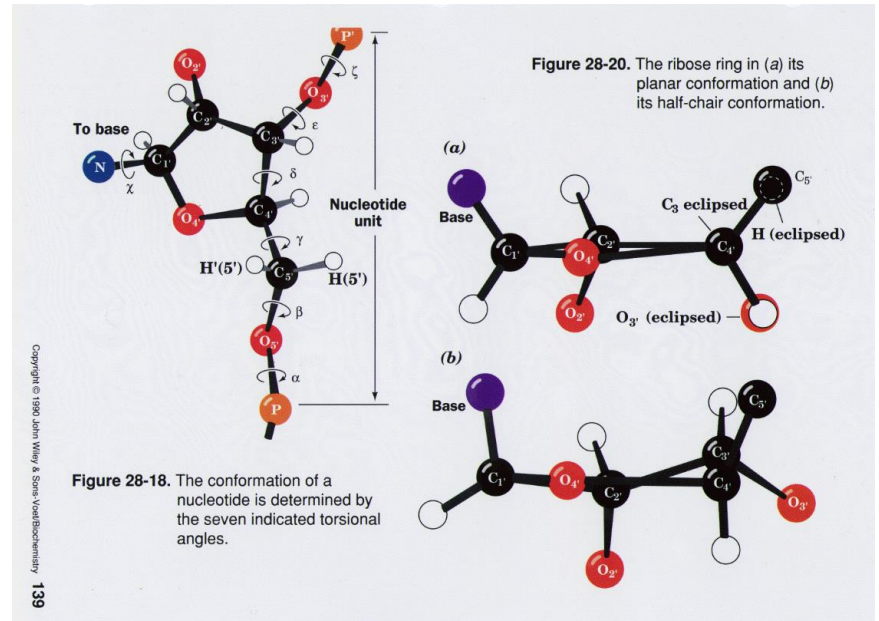
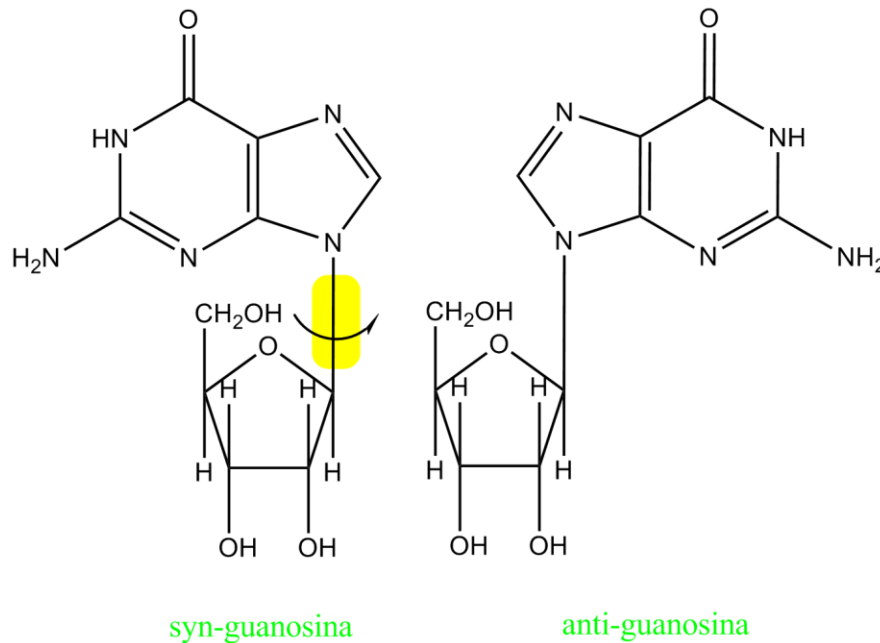


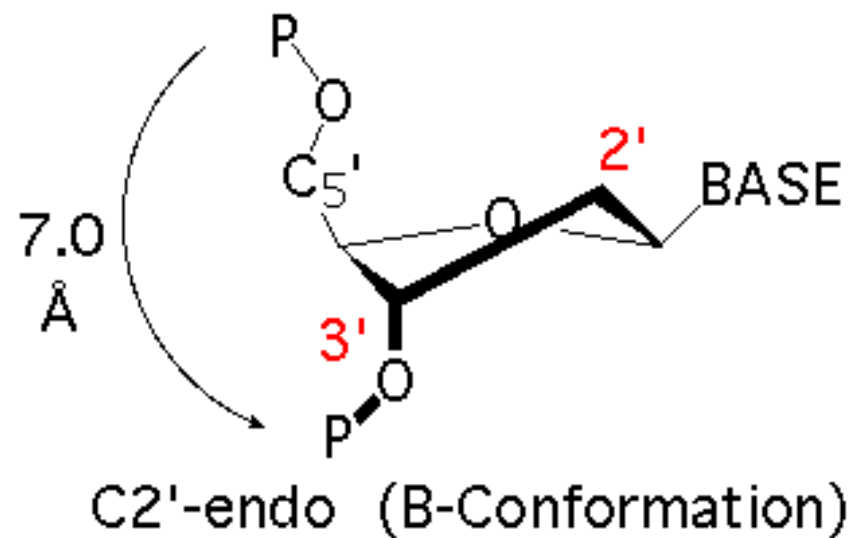
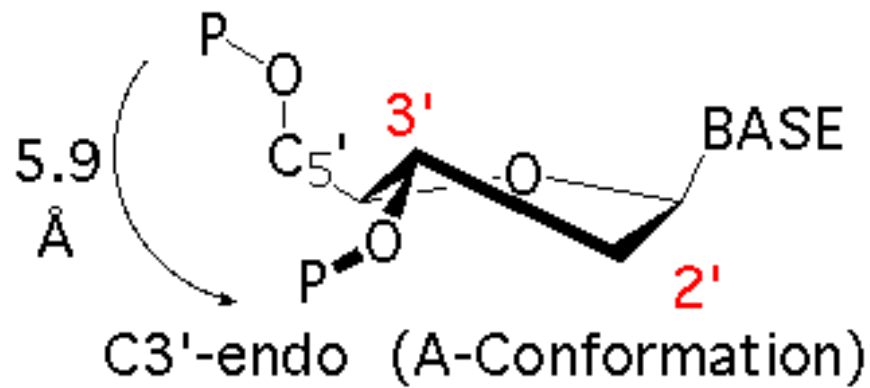
(b) Depurinazione

A differenza dei legami peptidici delle proteine che sono planari, i legami fosfodiesteri negli acidi nucleici possono assumere una grande varietà di conformazioni, definite dagli angoli torsionali α , β , e γ . Ulteriore complessità deriva dalle possibili conformazioni degli anelli dello zucchero. Nel DNA l'atomo C2' è endo rispetto a C5'

Inoltre, le basi possono assumere conformazioni sin e anti rispetto al legame glicosidico. Infatti, le nucleobasi possono adottare conformazioni sin (dalla stessa parte) o anti (parte opposta) rispetto al ribosio.

Queste piccole differenze hanno molta importanza nel definire la struttura secondaria e terziaria dell'acido nucleico.

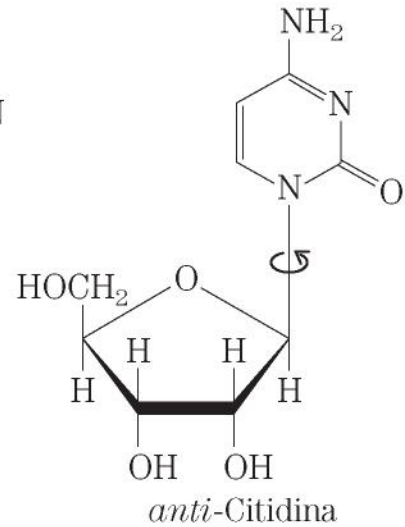
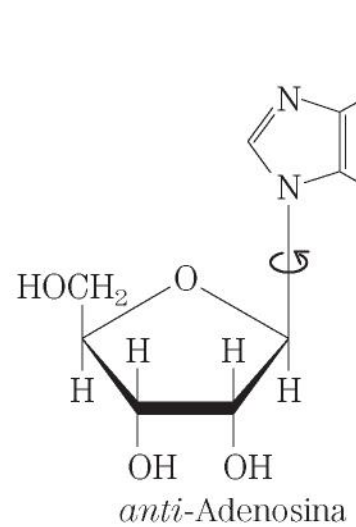
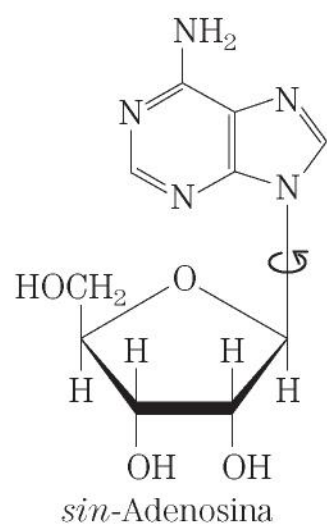
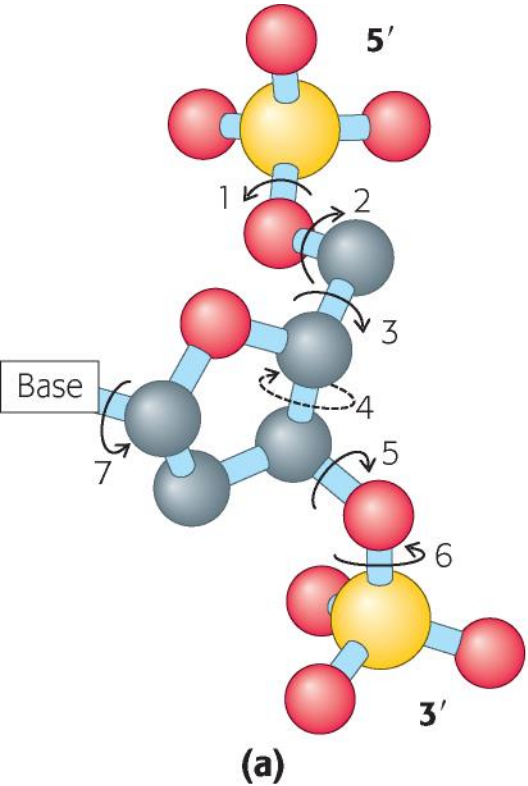




Principali conformazioni degli zuccheri furanosidici

Flessibilità dello scheletro zucchero-fosfato

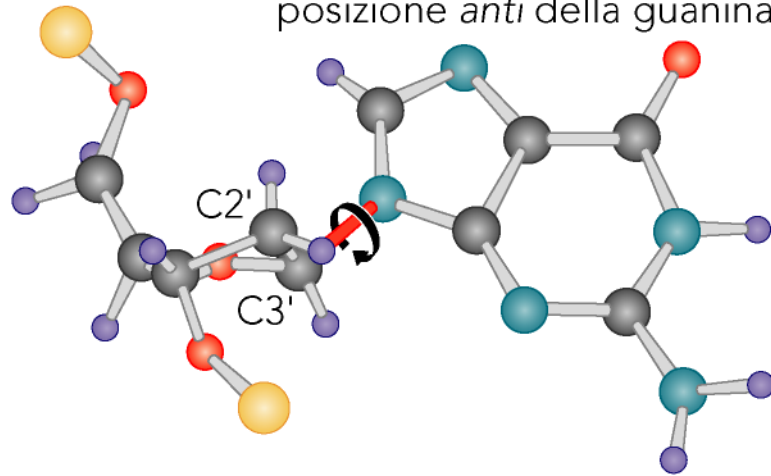
E' grazie alle rotazioni attorno ai legami covalenti semplici dello scheletro che gli acidi nucleici possono assumere diverse conformazioni. Queste dipendono da 3 fattori:
a) conformazione endo-eso degli zuccheri, b) rotazioni attorno ai legami dello scheletro, c) rotazione attorno al legame glicosidico (base-zucchero) in C1'.



(b)



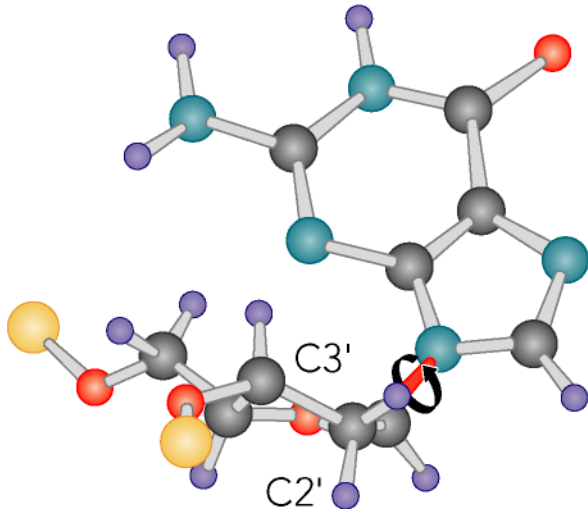
posizione *anti* della guanina



deossiguanosina nel DNA in forma B



posizione *sin* della guanina



deossiguanosina nel DNA in forma Z

Conformazioni anti e sin delle purine

Per rotazione del legame glicosidico, le purine possono assumere due conformazioni, dette anti e sin.

Ciò non è possibile con le pirimidine, che restano sempre in forma anti.

DNA can adopt different conformations

B-DNA: Most common, 10 bp per helical turn.

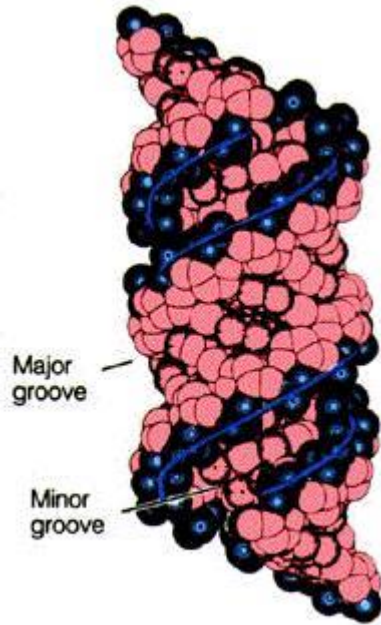
A-DNA: 11 bp per turn, RNA/DNA and RNA/RNA helices.

Z-DNA: Left-handed helix.

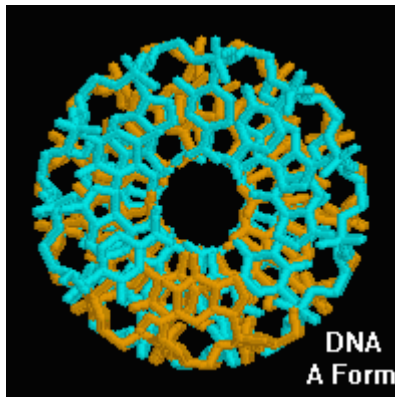
Triple-helix: Poly Pu/poly Py/poly Py

All can exist in a chromosome simultaneously.

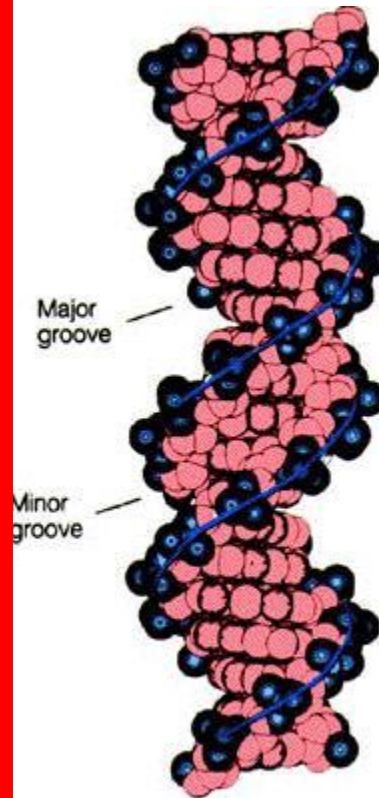
Per il DNA in condizioni non fisiologiche (fibra anidra) ma fisiologiche nel caso di dsRNA o ds DNA/RNA



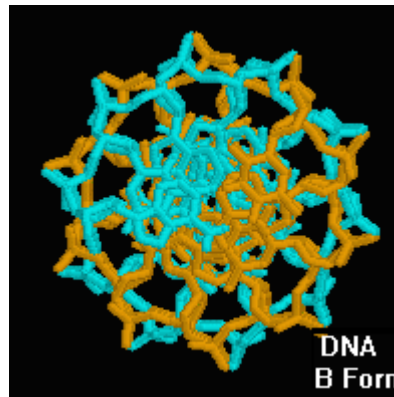
A-DNA



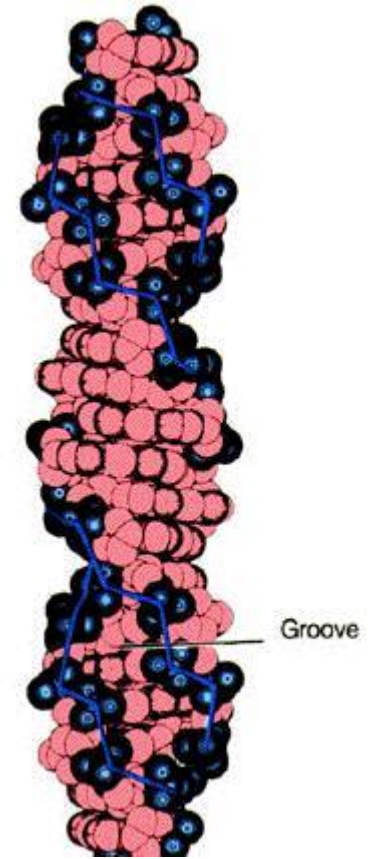
DNA
A Form



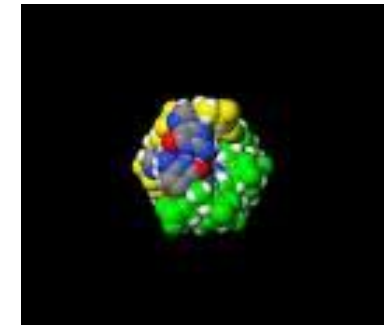
B-DNA



DNA
B Form



Z-DNA



Solo per sequenze alterne $(YR)_n$ in condizioni particolari

Le conformazioni A-DNA, B-DNA e Z-DNA sono le principali osservate nei sistemi biologici naturali.

La conformazione *B* è la più frequente nelle condizioni standard delle cellule.

La forma *A* è una spirale destrorsa (il solco minore è largo ma poco profondo, quello maggiore è più stretto e profondo), con un passo di 2,9 nm (circa 11bp) ed un diametro di 2,5 nm. Tale conformazione è presente in condizioni non fisiologiche, quando il DNA viene disidratato.

La conformazione *Z* è tipica invece delle sequenze che presentano modificazioni chimiche come la metilazione (tratti di DNA ricchi di basi C e G). Essa assume un andamento sinistrorso, opposto rispetto alla conformazione *B*.

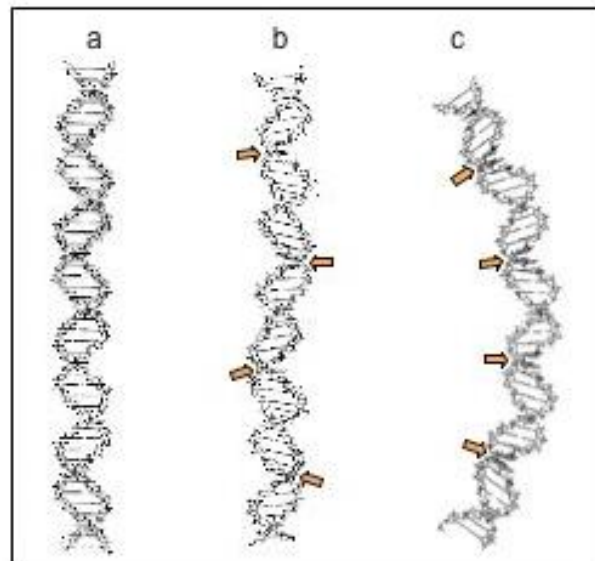
I parametri della struttura B del DNA a doppia elica, come descritti da Crick e Watson, descrivono una doppia elica molto regolare (FIGURA a).

Nella realtà dei fatti lungo la molecola del DNA ci sono piccole variazioni strutturali locali che in alcuni casi possono influenzare la struttura complessiva della molecola. Un caso ben studiato è quello del DNA intrinsecamente "curvo".

Due coppie di basi **A:T** adiacenti hanno una intrinseca tendenza a piegarsi dalla parte del **solco minore**, mentre due paia **G:C** hanno una tendenza a ripiegarsi dalla parte **del solco maggiore**. Ne risulta che la doppia elica non è perfettamente dritta ma presenta piccoli piegamenti. Questi, essendo la sequenza di nucleotidi eterogenea e casuale, tendono a elidersi a vicenda cosicché nell'insieme la struttura della molecola, pur con qualche deformazione, resta più o meno dritta (FIGURA b) .

Se però un tratto di DNA contiene, per esempio, due o tre paia di basi A:T con una periodicità di circa 10 paia di basi (cioè ripetute ogni giro di elica), i piccoli angoli si sommeranno tra loro producendo così una apprezzabile curvatura dell'asse della molecola di DNA (Figura c).

Queste curvature intrinseche del DNA si trovano alle volte in punti specifici del genoma, per esempio a monte di vari promotori, e in alcuni casi possono avere una rilevanza funzionale.



Confronto fra le forme A, B e Z

Caratteristiche generali comuni:

1. Le due catene sono antiparallele
2. L'appaiamento delle basi è a geometria Watson-Crick

	Forma A	Forma B	Forma Z
Senso dell'elica	Destrorsa	Destrorsa	Sinistrorsa
Diametro	26 Å	20 Å	18 Å
Coppie basi/giro	11	10.5	12
Distanza fra le basi	2.6 Å	3.4 Å	3.7 Å
Piegamento basi rispetto alla normale all'asse	20°	6°	7°
Solco maggiore	stretto, profondo	largo, medio	assente
Solco minore	largo, poco profondo	stretto, medio	stretto, profondo
Conformazione	C3'-endo	C2'-endo	C2'-endo (Pyr)
Conformazione legame N-glicosidico	Anti	Anti	Syn (Pur) Anti (Pyr)

Le strutture A e B possono sussistere per qualsiasi sequenza di basi purchè vi sia complementarità tra le due catene, la struttura Z solo per sequenze alterne pirimidina-purina in condizioni particolari (alta conc. salina o stress torsionale)

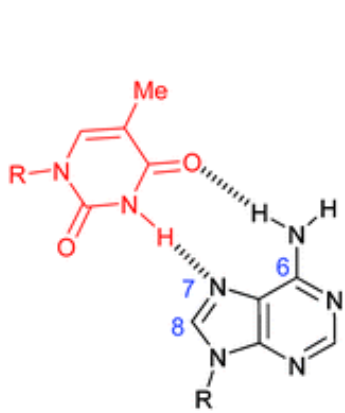
Disease associations with Z-DNA

- Rheumatoid arthritis
- Multiple sclerosis
- Type 1 Diabetes
- Infectious disease susceptibility
- Chronic hepatitis C
- Leukemia
- Lymphoma

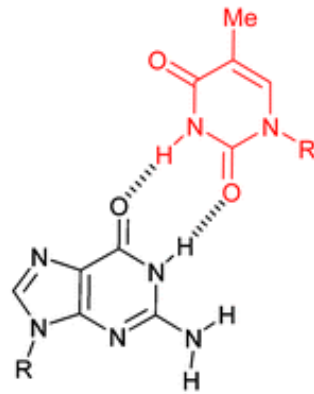
Z- DNA and disease

Alzheimer's disease

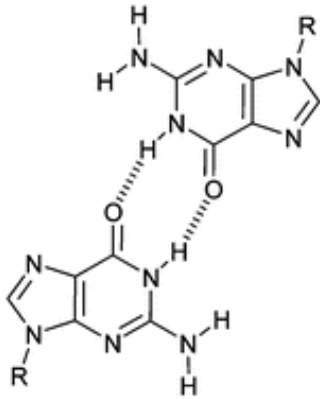
- Severely infected – Z DNA conformation
- Moderate - probable B-Z intermediate form
- Normal - B-DNA conformation
- Amyloid Beta and aluminum (an etiologic factor) can modulate helical alterations *in vivo*
- Alternate purine pyrimidine sequences found in promoter regions of AD specific genes (amyloid precursor proteins, presenilin and ApoE)



Hoogsteen (TA)



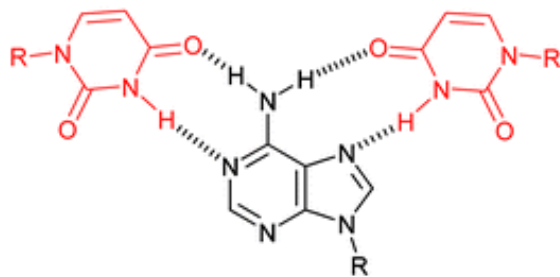
Wobble (GT)



Self-dimerization (GG)



Reverse Watson-Crick (AT)

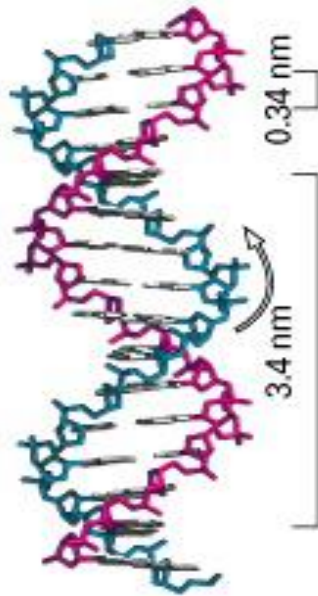


UAU base triplet (Watson-Crick and Hoogsteen)

Molto raramente nel DNA, più spesso nell'RNA si possono trovare appaiamenti di legame idrogeno tra basi con geometria diversa da quella di Watson e Crick. Ecco alcuni esempi.

DNA forms

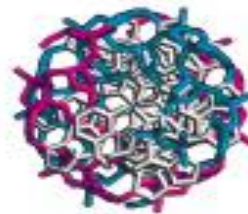
(a) B DNA



(b) A DNA

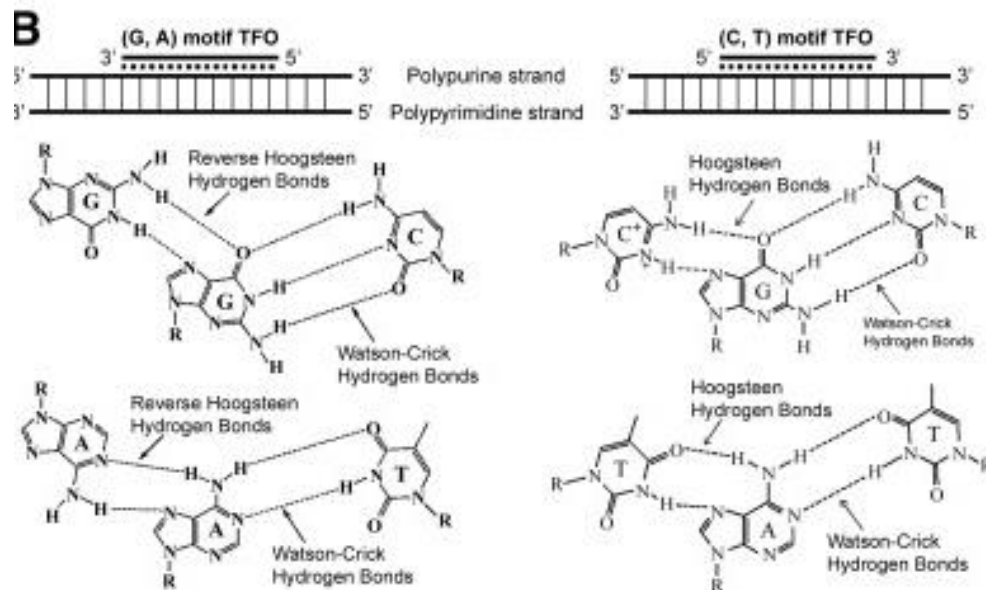
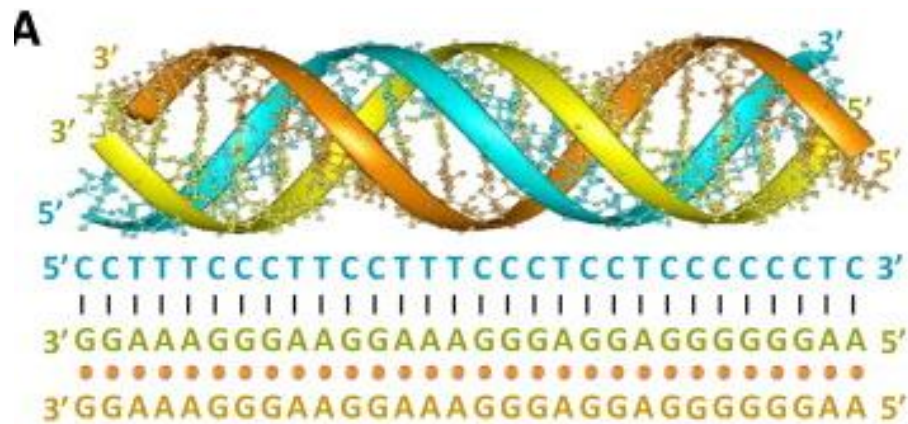


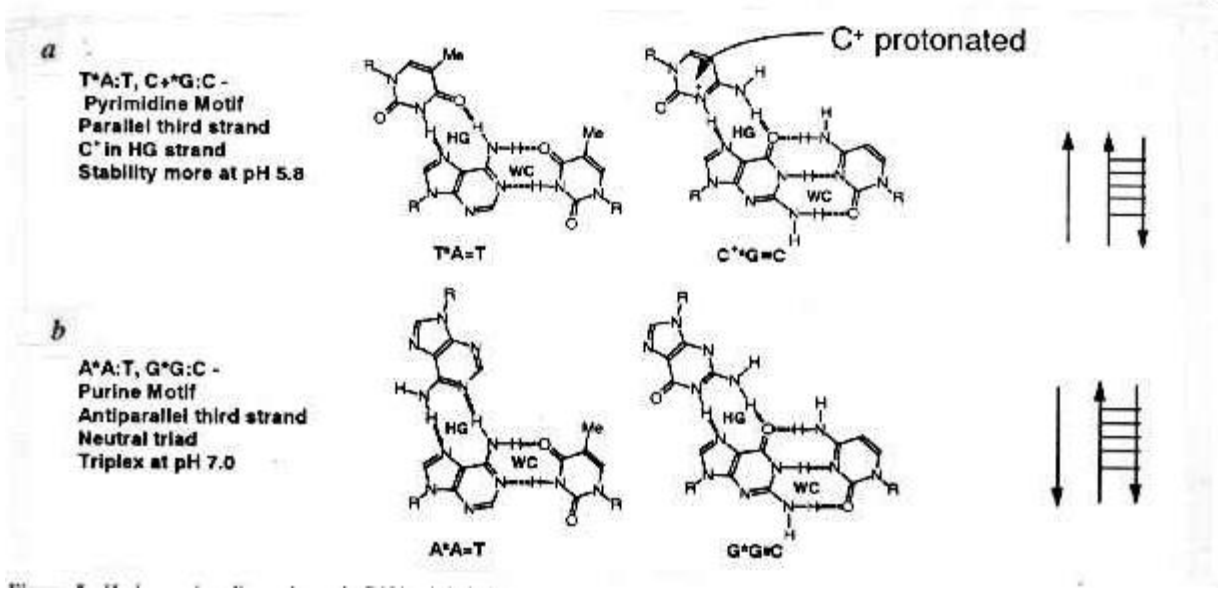
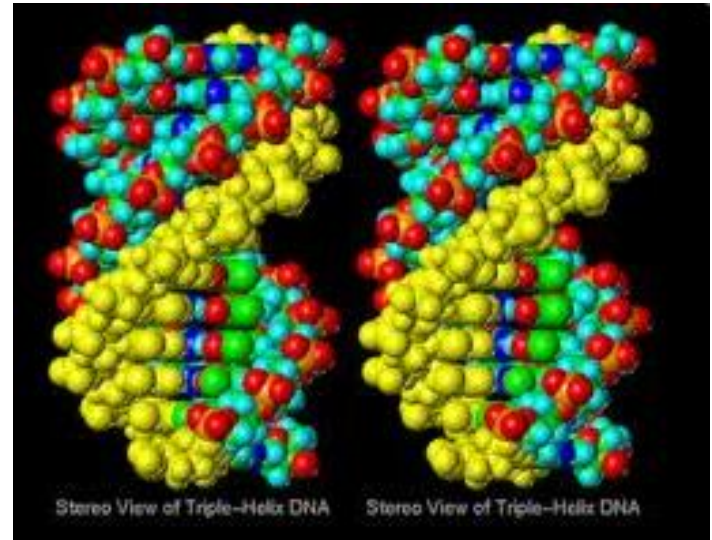
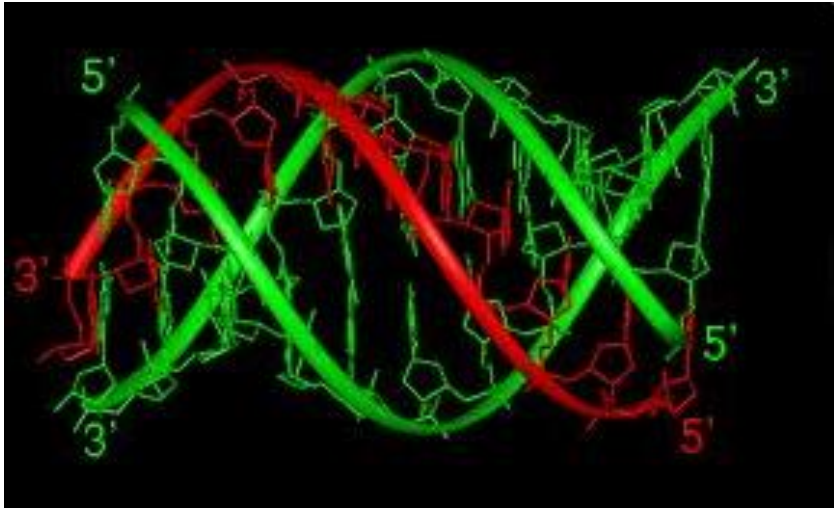
(c) Z DNA

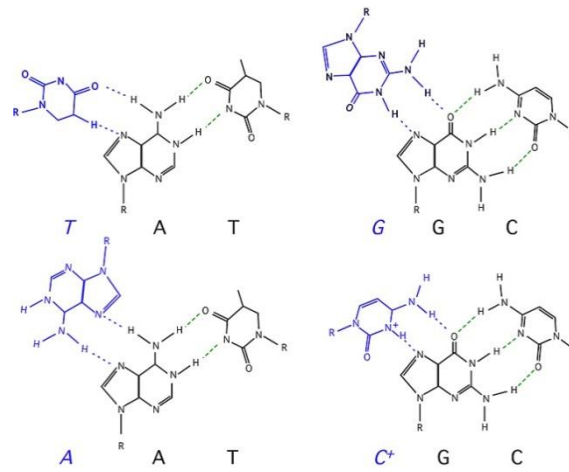


(d) Triple-Helical DNA

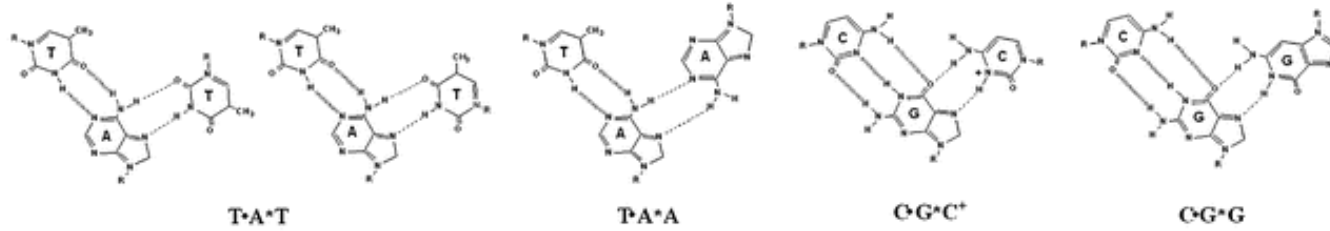




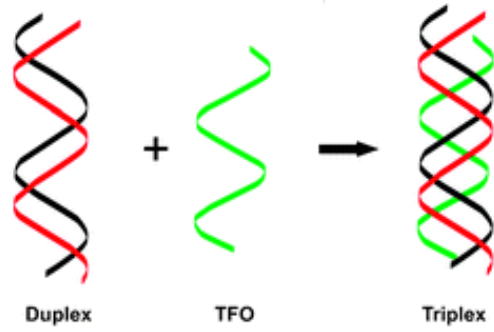




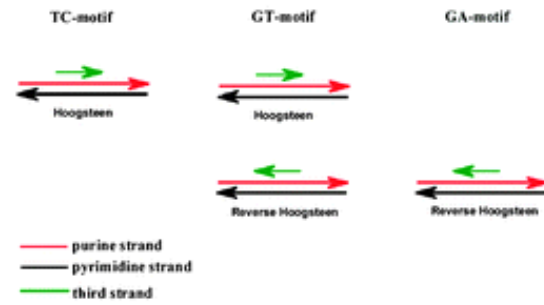
A



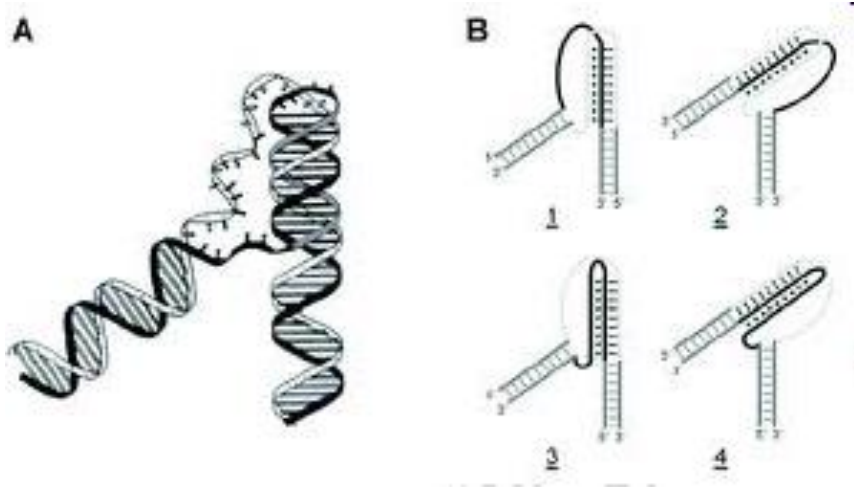
B



C



Tripla elica di DNA (DNA-H)



Triplex structures may form in vivo. However, their presence need not always be advantageous.

Intramolecular triplexes could promote undesirable recombination events or interfere with normal transcriptional controls. Thus, another aspect supporting the existence of triplex species might be the presence of **proteins that destabilize these structures.**

The absence of these proteins has been found to increase genomic instability, especially at regions having the potential to form triplex structures



©IRB Barcelona



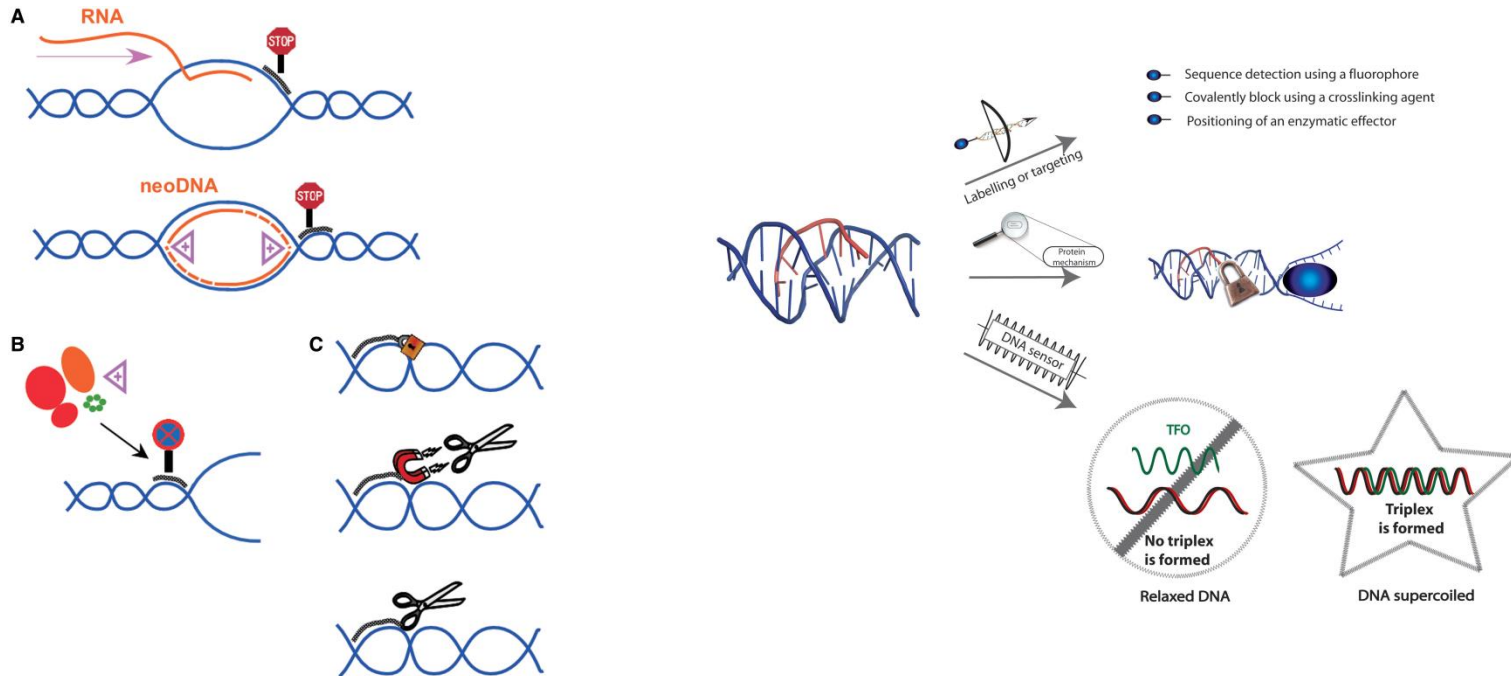
tripla elica del DNA in fase gassosa

Triple-helix target sites (TTS) are over-represented in the human genome and especially at promoter regions

It has been reported the H-DNA structure (an intramolecular triplex) modulates transcription in the human c-MYC promoter

The d(G) content is nonrandom in eukaryote genomes, and many G tetrads promote deletions unless peculiar helicases unwind them. d(G) tracts are nonrandomly scattered within the genome but display a skewed representation in oncogene and tumor suppressor genes

Anti-gene approach based on the triplex strategy by using TFO has been developed for therapeutic and biotechnologic aspects



Examples of use of triplexes to modulate gene expression:

- (A) physical block of transcription or replication elongation;
- (B) Blockage of transcription and replication initiation
- (C) targeting of DNA modifying agents such as cross-linking and cleaving agents.

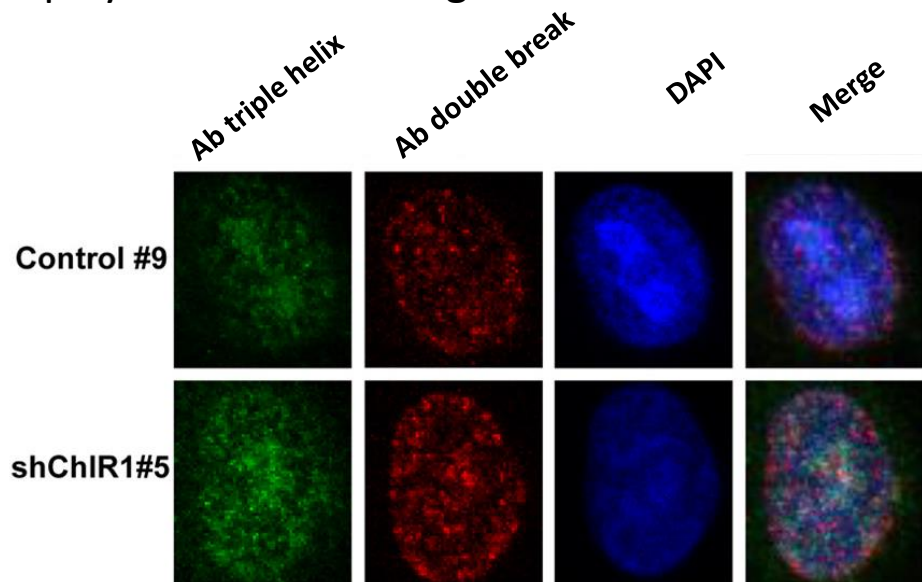
Triplexes as biotechnological tools: to target a molecule of interest on a specific DNA sequence, such as fluorophores, enzymatic effectors (transcription factors, restriction enzymes), cross-linking agents; to study the mechanisms of proteins that act on DNA, such as translocases; to sense DNA topology: using immobilized TFOs, it is possible to trap negatively supercoiled plasmids versus relaxed plasmids.

A Distinct Triplex DNA Unwinding Activity of ChIR1 Helicase

❖ **Background:** DNA triplex helix structures are alternate DNA structures that can be a source of genomic instability.

❖ **Results:** ChIR1 helicase has a novel and distinct triplex DNA unwinding activity. Conclusion: ChIR1 defends genome integrity by resolving triplex DNA structures.

❖ **Significance:** The processing of triplex DNA substrates by proteins such as ChIR1 plays critical roles in genome maintenance



H-DNA e ATASSIA di FRIEDREICH

Espansione della tripletta GAA nel primo introne del gene per l'atassia di Friedreich che si trova sul cromosoma 9 (normale 8-30; patologico >30 fino a 1000)

Gravità direttamente correlata al numero di espansioni

INIBIZIONE DELLA TRASCRIZIONE DEL GENE della frataxina (proteina mitocondriale) PER LA FORMAZIONE DELLA TRIPLA ELICA

Prevalenza: 1/30-50.000, eredità autosomica recessiva

Patogenesi: Atrofia cellule a T dei gangli spinali, atrofia-degenerazione cordoni post. + degeneraz. tratto spino-cerebellare dorsale e ventrale

Gene: FRDA, mappato su cromosoma 9q13-21, per la frataxina (proteina mitocondriale)

Clinica: Esordio < 25 aa., progressiva atassia della marcia senza remissione



Recentemente, la tecnologia antigene (antigene technology , AT) ha attratto notevole interesse per lo sviluppo di nuove strategie basate su TFO (Triple helix-Forming Oligonucleotide). Ci sono infatti crescenti evidenze che i TFOs possono essere importanti nella regolazione dell'espressione genica e per la modifica di una regione di un gene specifico, sia in vitro che in vivo. Inoltre la AT può controllare e/o modulare le funzioni cellulari a livello del DNA genomico. È stato dimostrato che un TFO anti-IGF-I può efficacemente sopprimere lo sviluppo di tumori in modelli animali e nelle neoplasie umane.

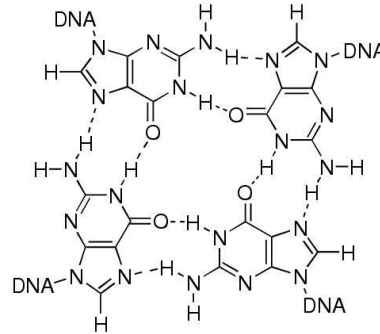
Approcci terapeutici basati su TFO sono stati sviluppati per modulare l'espressione genica, per promuovere la mutagenesi sito-specifica e quindi modificare specifiche regioni di DNA genomici.

Alcuni recenti risultati sostengono che il TFO può bloccare il macchinario della trascrizione non solo di bloccando il sito di inizio e/o alterando il riconoscimento del DNA da parte dei fattori di trascrizione, ma anche bloccando/rallentando la strada alla RNA polimerasi mediante stabilizzazione delle coppie di basi nel sito di trascrizione.

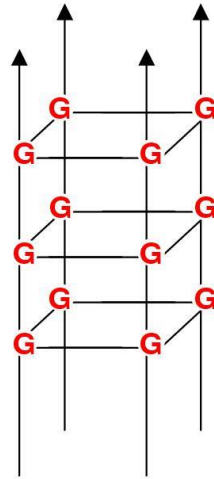
Recenti studi sostengono che i miRNA possono esercitare la loro regolazione sulle regioni di promotori attraverso la formazione della tripla elica miRNA – DNA•DNA che sarebbe stabilizzata dalle proteine AGO. Questi studi rappresentano un modello plausibile silenziamento genico trascrizionale (TGS) mediato da miRNA.

G-quadruplex (nelle estremità dei telomeri ed in alcune regioni introniche)

Appaiamenti Hoogsteen

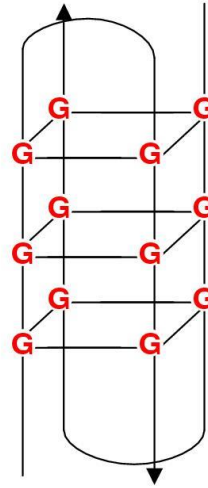


A. G-tetrad structure



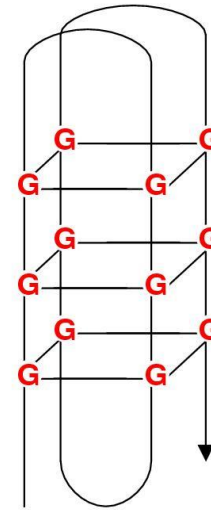
B.

G4-DNA intermolecolare

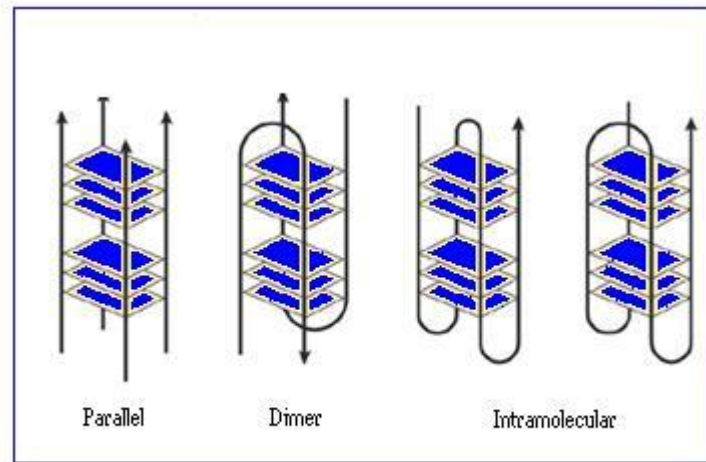
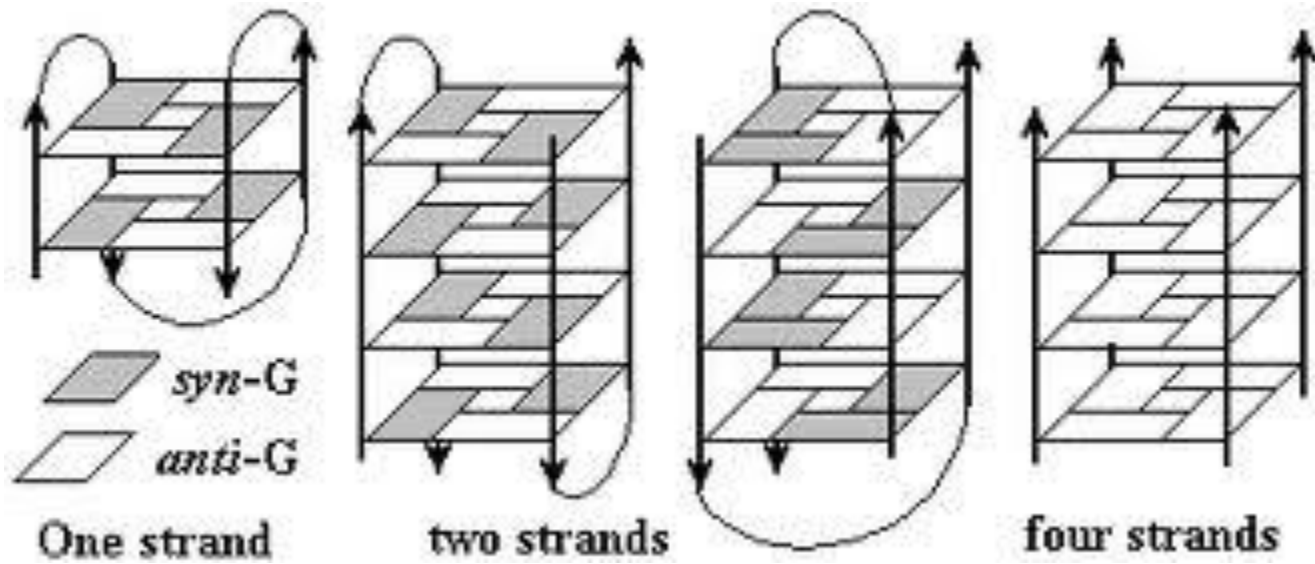


C.

G4-DNA intramolecolare
Tipo basket

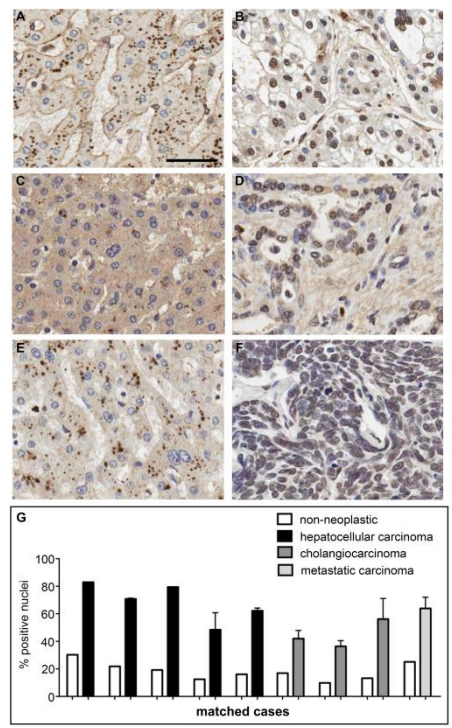
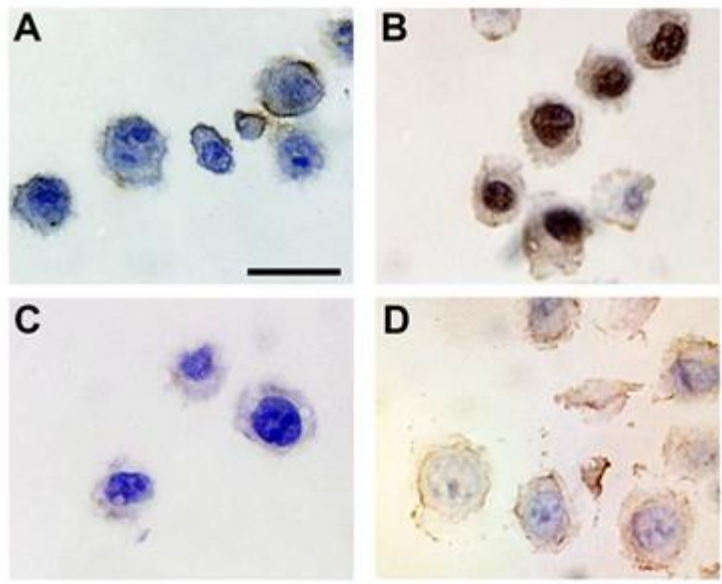


D.



“We observe a significantly elevated number of G-quadruplex-positive nuclei in human cancers of the liver and stomach as compared to background non-neoplastic tissue.”

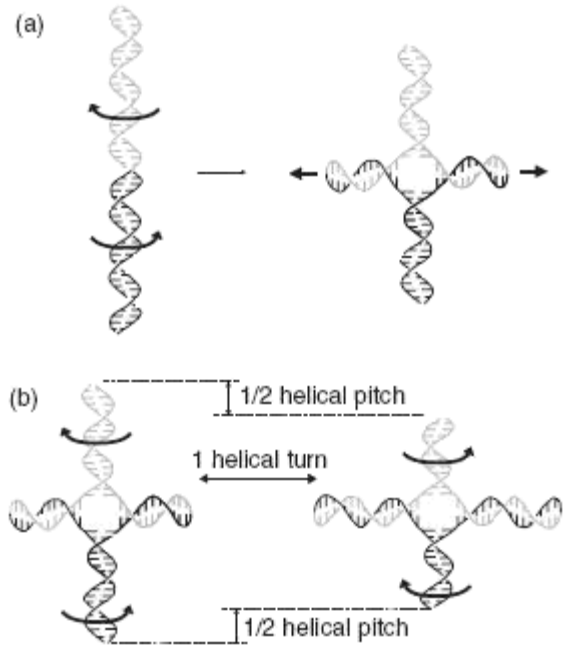
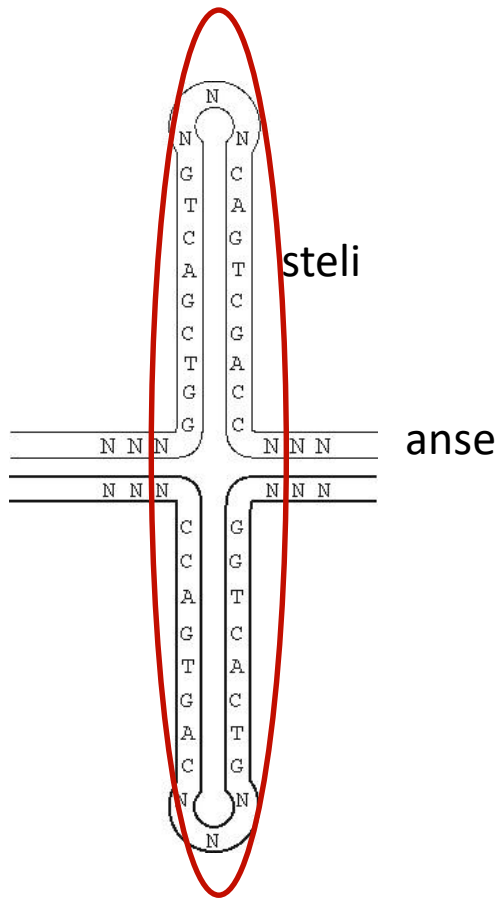
Figure 1. Validation of G-quadruplex nuclear staining by immunohistochemistry.



Biffi G, Tannahill D, Miller J, Howat WJ, Balasubramanian S (2014) Elevated Levels of G-Quadruplex Formation in Human Stomach and Liver Cancer Tissues. PLOS ONE 9(7): e102711. doi:10.1371/journal.pone.0102711
<http://dx.doi.org/10.1371/journal.pone.0102711>

CRUCIFORMI

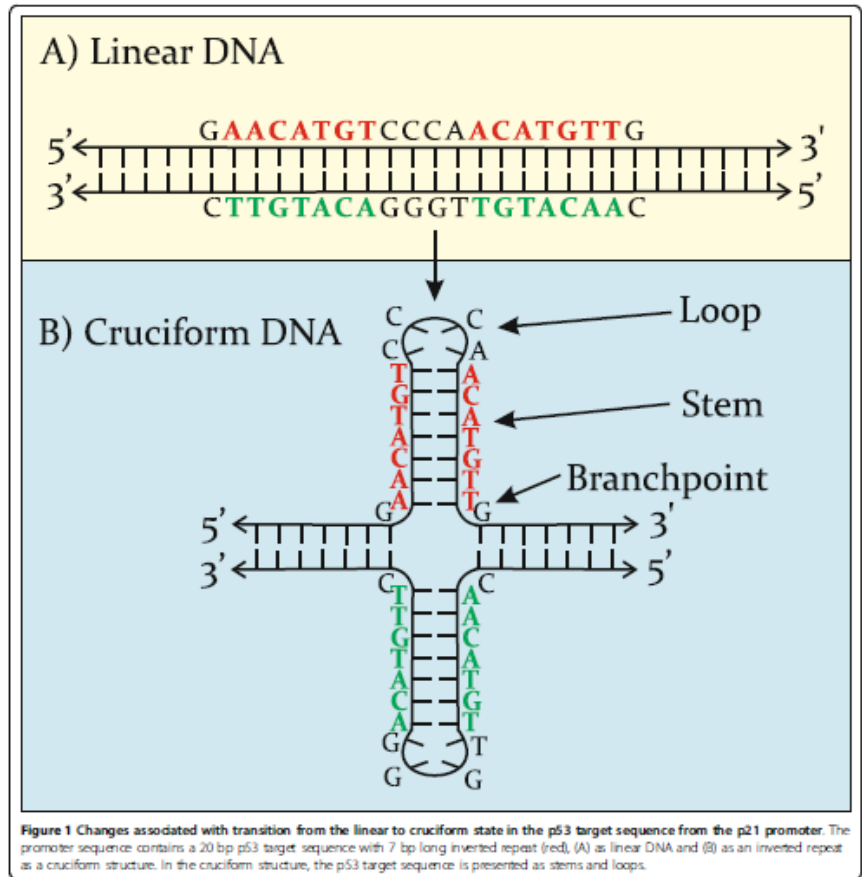
(quando sono presenti a breve distanza due coppie della stessa sequenza in orientamento opposto-palindromi-)




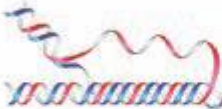
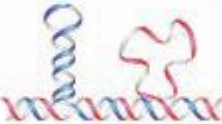
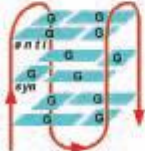

Cruciform structures may potentiate intrachromosomal recombination causing diseases like cancer.

Table 1 Proteins involved in interactions with cruciform structures

Protein	Source	Reference
Junction-resolving enzymes		
Integrase family		
RuvC	E.coli	[133-135]
Cce1	yeast	[136]
Ydc2	Spombe	[134]
A22	Coccinia virus	[137]
Integrases	all	[119,138]
Restriction nuclease family		
Endonuclease I	Phage T7	[139-141]
RedJ	G+ bacteria	[134,142]
Hjc, Hje	archea	[134,143]
MutH	Eukaryotes	[25,144]
Other		
Endonuclease VII	phage T4	[25,145]
RusA	E.coli	[146]
MSH2	S. cerevisiae	[147,148]
Mus81-Erm1	Eukaryotes	[42,149-151]
TRF2	H. sapiens	[52,152]
XPB, XPG protein families	Eukaryotes	[56,153,154]
Transcription, Transcription factors and DNA repair		
PAP-1	H. sapiens and others	[5,163]
BRC1	H. sapiens and others	[49,50,91,93]
P53	H. sapiens and others	[69,73,75,76,132,155,156]
Bmh1	Sacerevisiae	[35]
14-3-3	H. sapiens, Sacerevisiae	[34,110]
Rmi-1	Yeast	[157]
Gp-1	S. cerevisiae	[158]
HMG protein family	all	[47,159-161]
Smc	S. cerevisiae	[118,162]
Hop1	S. cerevisiae	[163,164]
ER estrogen receptor	mammals	[58]
Chromatin-associated proteins		
DEK	mammals	[84,85]
BRC1	mammals	[49,50,91,93]
HMG protein family	Eukaryotes	[47,159-161]
Rad54	Eukaryotes	[48]
Rad51ap	Eukaryotes	[81]
Topoisomerase I	Eukaryotes	[101,165]
Replication		
S16	E.coli	[113]
GF14, homolog of 14-3-3	plants	[35]
MLL (leukemia)	H. sapiens	[125,126]
WRN (Werner syndrome)	H. sapiens	[129]
AF10	H. sapiens	[114]
14-3-3	Eukaryotes	[34,110]
DEK	mammals	[84,85]
DNA-PK	Eukaryotes	[166]
VF1	Baculoviruses	[119]
HU	E. coli	[105,162,164]
Helicases (59, 44, and others)	all	[55]



Non-B DNA structures

Name	Conformation	General Seq. Requirements	Sequence
Cruciform		Inverted Repeats	$\begin{array}{c} \text{TCGGTACCGA} \\ \text{AGCCATGGCT} \end{array}$
Triplex		$(R \cdot Y)_n$ Mirror Repeats	$\begin{array}{c} \text{AAGAGG} \cdot \text{GGAGAA} \\ \text{TTCTCC} \cdot \text{CCTCTT} \end{array}$
Slipped (Hairpin) Structure		Direct Repeats	$\begin{array}{c} \text{TCGGTTCGGT} \\ \text{AGCCAAGCCA} \end{array}$
Tetraplex		Oligo (G) _n Tracts	$\text{AG}_3(\text{T}_2\text{AG}_3)_3$ single strand
Left-handed Z-DNA		$(YR \cdot YR)_n$	$\begin{array}{c} \text{CGCGTGCGTGTG} \\ \text{GCGCACGCACAC} \end{array}$