

Δομή νουκλεϊκών οξέων (DNA/RNA)

Από εδώ ξεκίνησαν όλα

No. 4358 April 25, 1953

NATURE

737

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

¹ Young, F. B., Gerrard, H., and Fevora, W., *Phil. Mag.*, **40**, 149 (1920).

² Longuet-Higgins, M. S., *Mon. Not. Roy. Soc., Geophys. Surv.*, **2**, 250 (1948).

³ 1306 A.F.E., W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (5) (1950).

⁴ Sjöman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (31) (1956).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey¹. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β -D-deoxyribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There



This figure is partly diagrammatic. The two ribbons symbolize the two phosphate-sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis.

is a residue on each chain every 3.4 Å. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 Å. The distance of a phosphorus atom from the fibre axis is 10 Å. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally^{2,3} that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

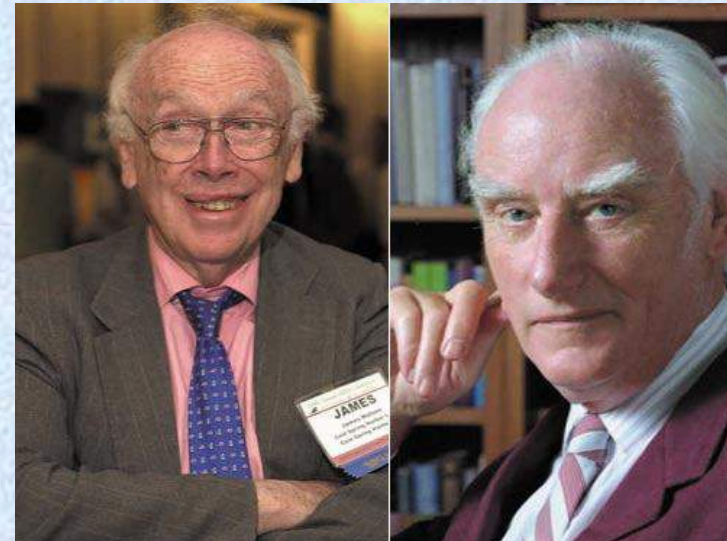
The previously published X-ray data^{4,5} on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at



It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.



Κάποιες παλιές ιστορίες

F. Crick



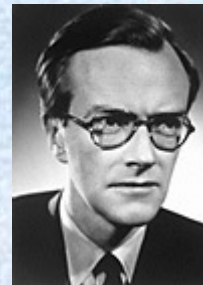
1916-2004

J. Watson



1928-

M. Wilkins



1916-2004

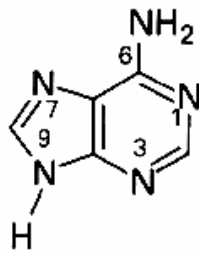
Nobel (Ιατρική), 1962

Παρόλο που το μοντέλο του DNA προτάθηκε το 1953, η πρώτη πειραματική του επιβεβαίωση από κρυστάλλους DNA ήρθε μόλις το 1981 (28 χρόνια μετά!) από το εργαστήριο του **Richard Dickerson** στο Caltech

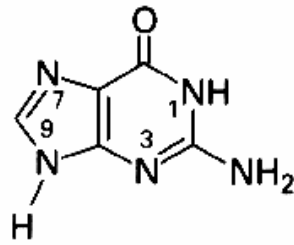
Η πρώτη δομή συμπλόκου DNA/πρωτεΐνης (του αναστολέα του φάγου 434 με ένα κομμάτι DNA του χειριστή) λύθηκε το 1988 στο εργαστήριο του **Stephen Harisson** στο Harvard

Οι πρώτες δομές συμπλόκων RNA/πρωτεΐνης λύθηκαν το 1989 από τις ομάδες των **Stubbs**, **Johnson** και **Steitz**

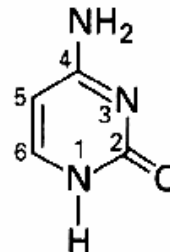
Οι βάσεις του DNA και RNA



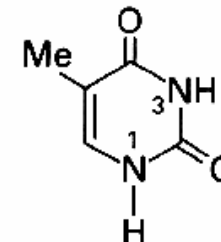
adenine



guanine



cytosine

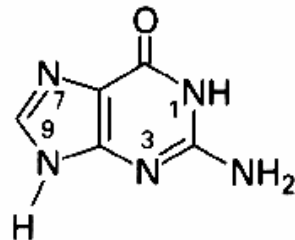


thymine

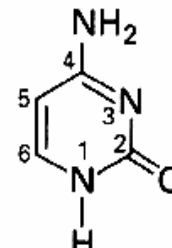
DNA



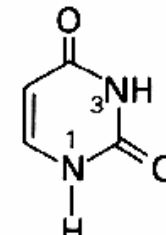
adenine



guanine



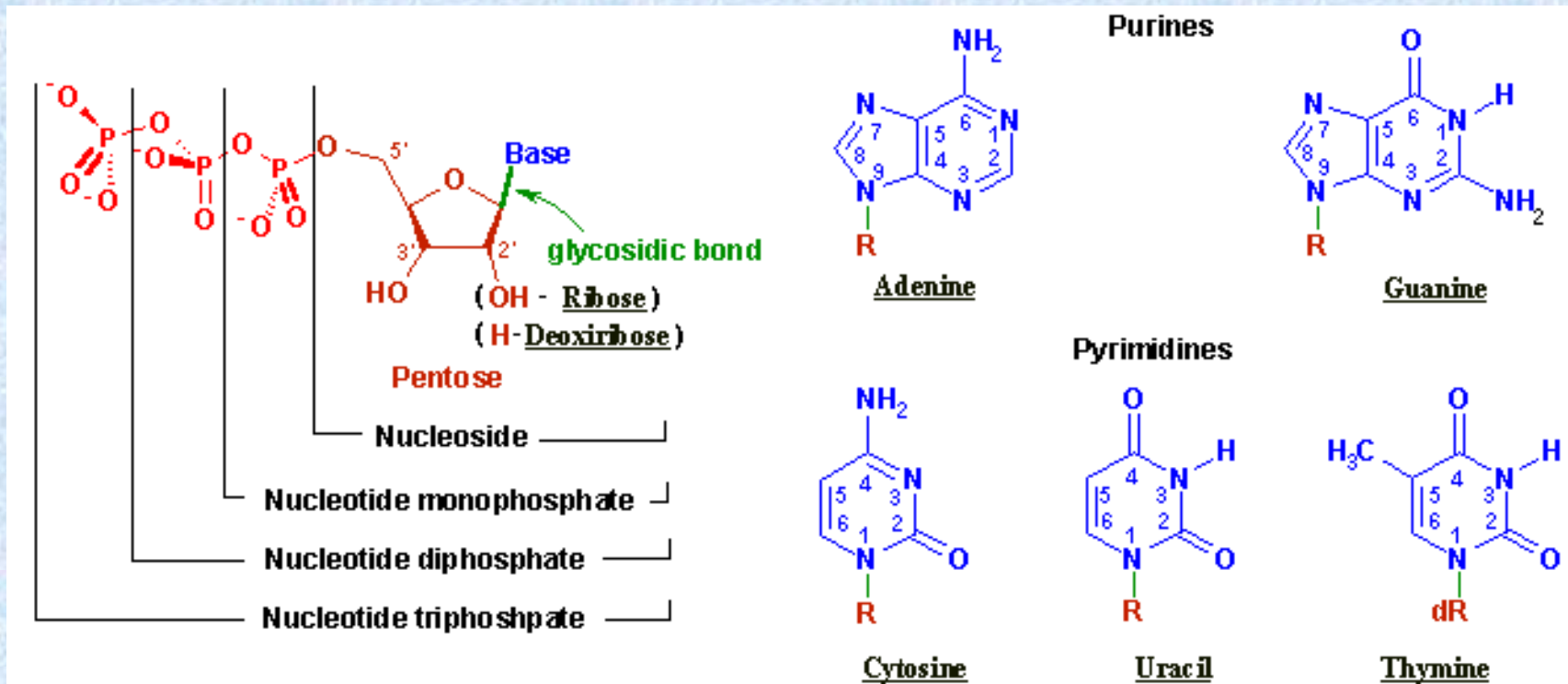
cytosine



uracil

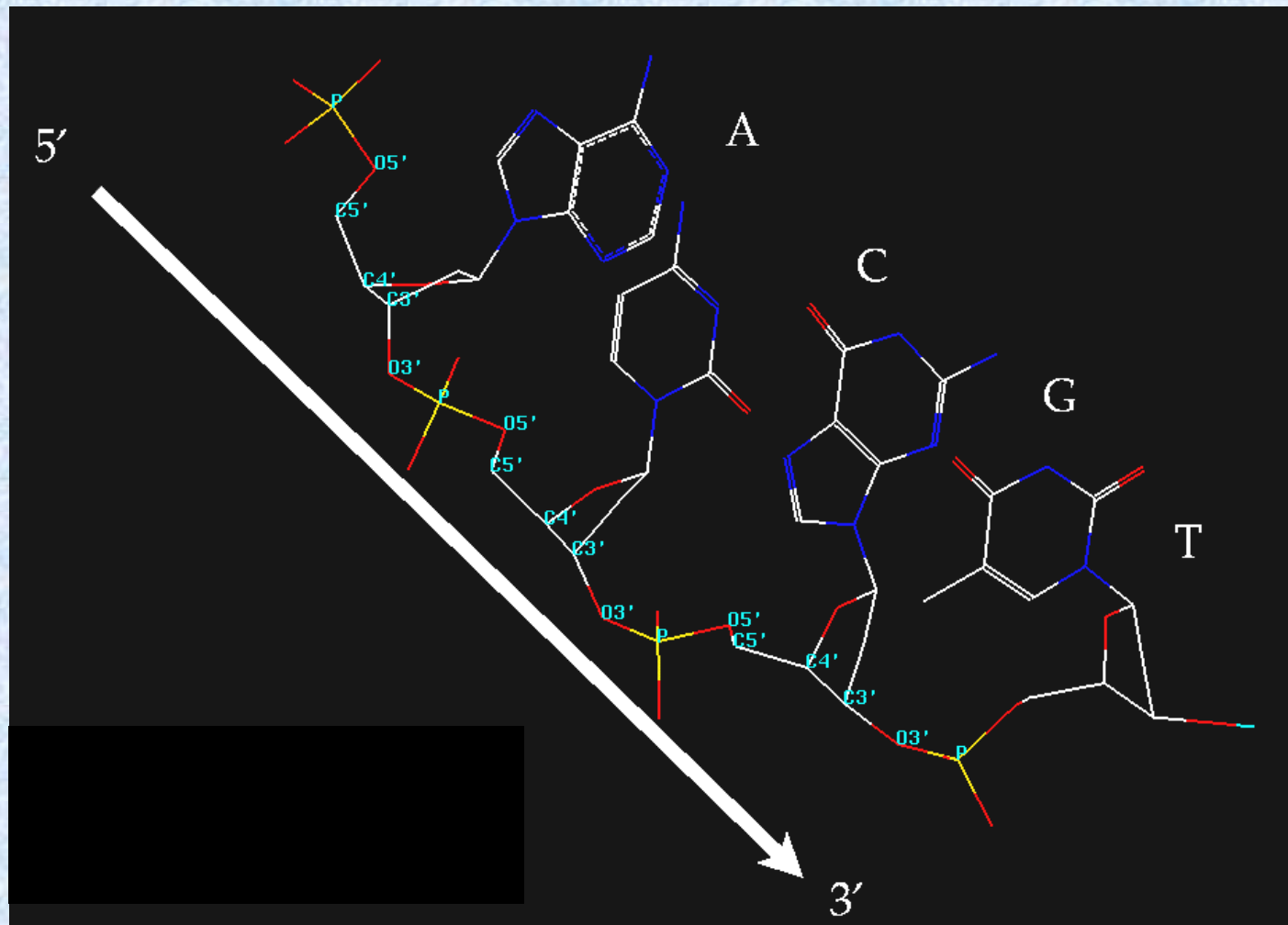
RNA

Οι δομικές μονάδες

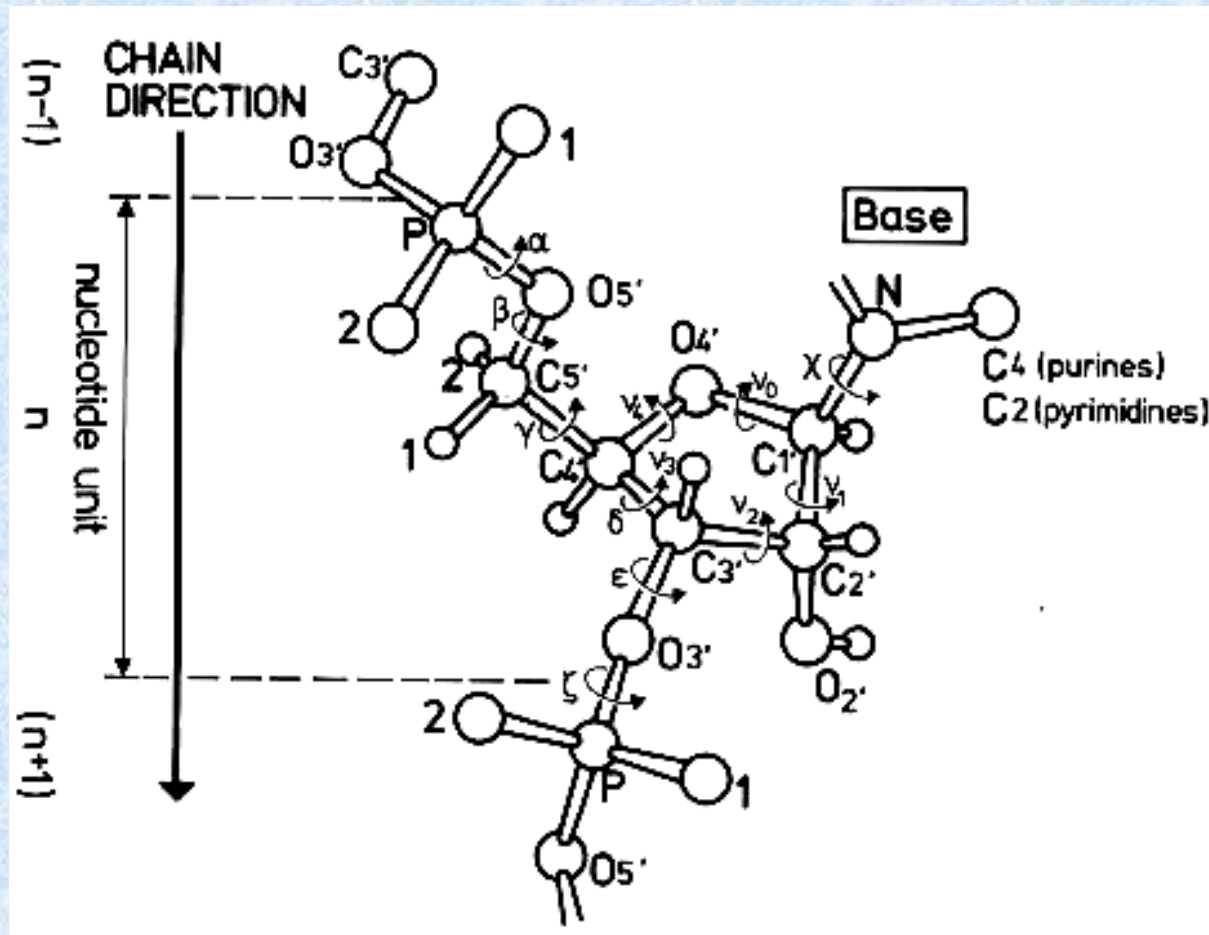


Οι βάσεις ενώνονται με ριβόζη ή δεοξυριβόζη, μέσω ενός β-γλυκοσιδικού δεσμού, φτιάχνοντας τους **νουκλεοζίτες**. Το 5'-OH των νουκλεοζιτών αντιδρά με φωσφορικό οξύ και φτιάχνεται το **(δεοξυ)νουκλεοτίδιο**, η επαναλαμβανόμενη μονάδα των νουκλεϊκών οξέων. Σύνδεση των φωσφορικών ομάδων με φωσφοδιεστερικούς δεσμούς σχηματίζει την πολυνουκλεοτιδική αλυσίδα.

Ένας κλώνος DNA



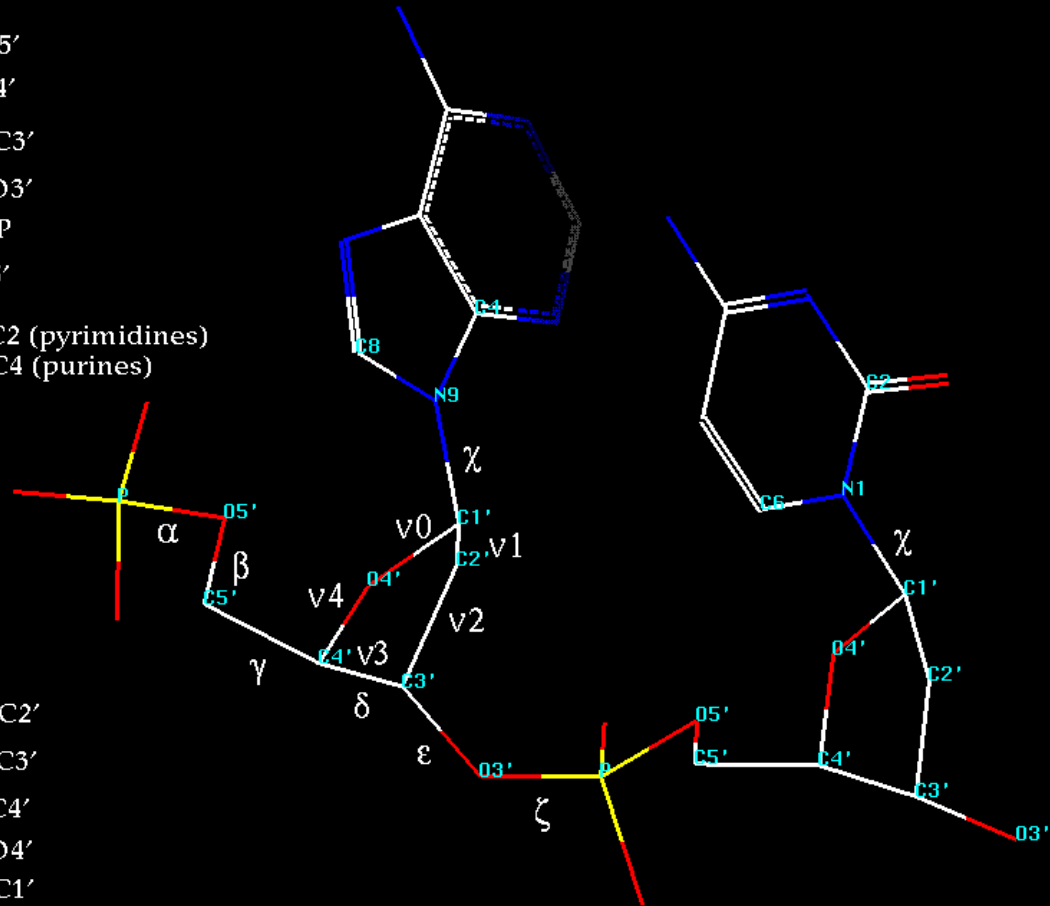
Ευελιξία του DNA



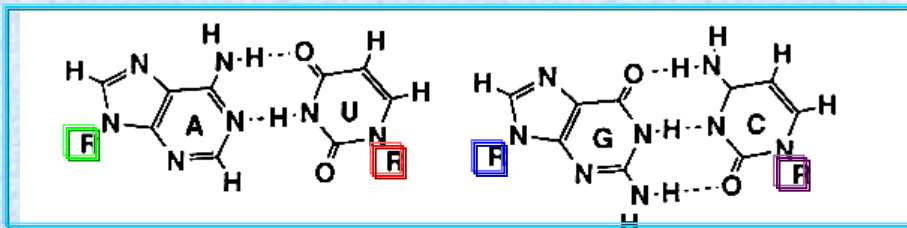
Ευελιξία του DNA

- α O3'-P-O5'-C5'
- β P-O5'-C5'-C4'
- γ O5'-C5'-C4'-C3'
- δ C5'-C4'-C3'-O3'
- ϵ C4'-C3'-O3'-P
- ζ C3'-O3'-P-O5'
- χ O4'-C1'-N1-C2 (pyrimidines)
O4'-C1'-N9-C4 (purines)

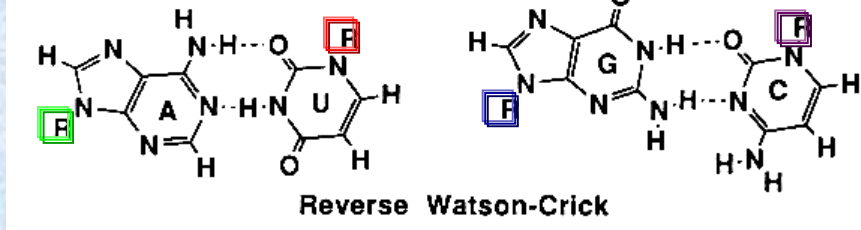
- v0 C4'-O4'-C1'-C2'
- v1 O4'-C1'-C2'-C3'
- v2 C1'-C2'-C3'-C4'
- v3 C2'-C3'-C4'-O4'
- v4 C3'-C4'-O4'-C1'



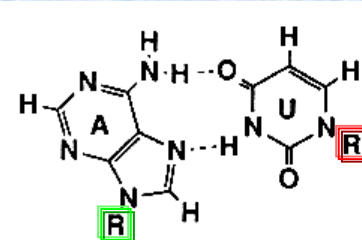
Συμπληρωματικότητα βάσεων



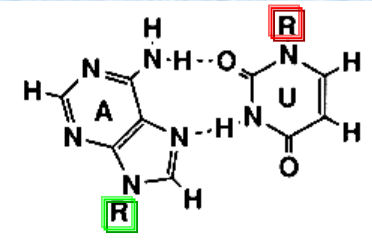
Watson-Crick



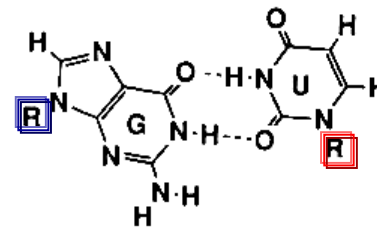
Reverse Watson-Crick



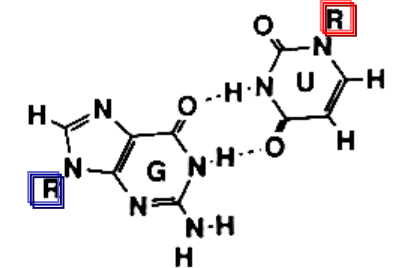
A-U Hoogsteen



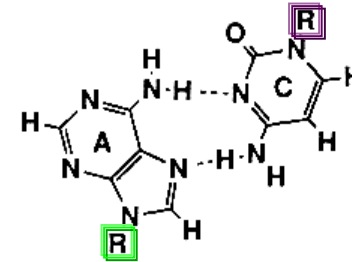
A-U Reverse Hoogsteen



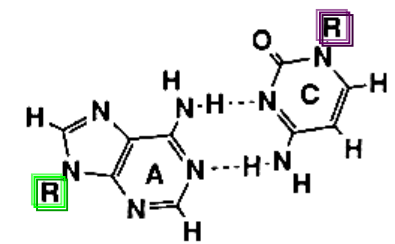
G-U Wobble



G-U Reverse Wobble

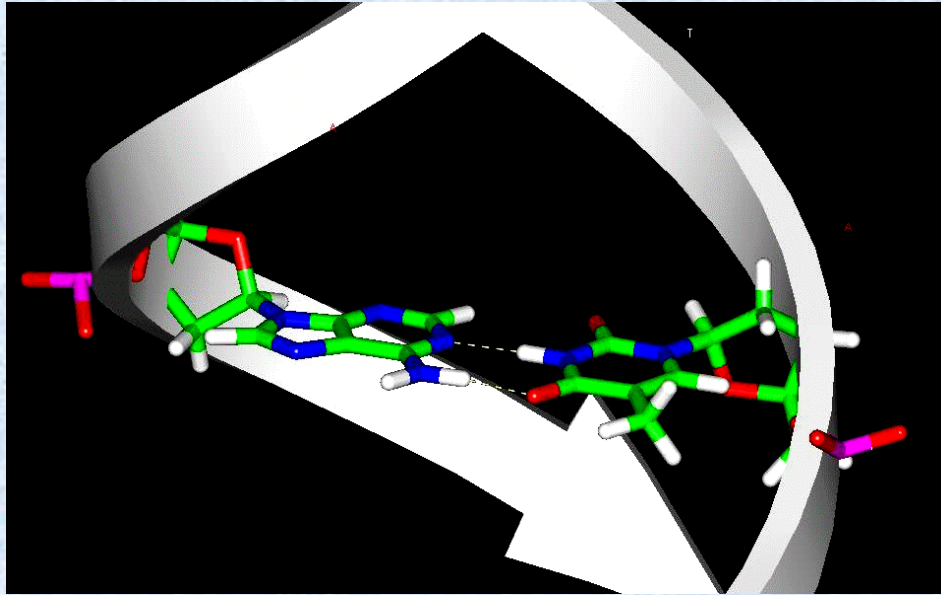


A-C Reverse Hoogsteen



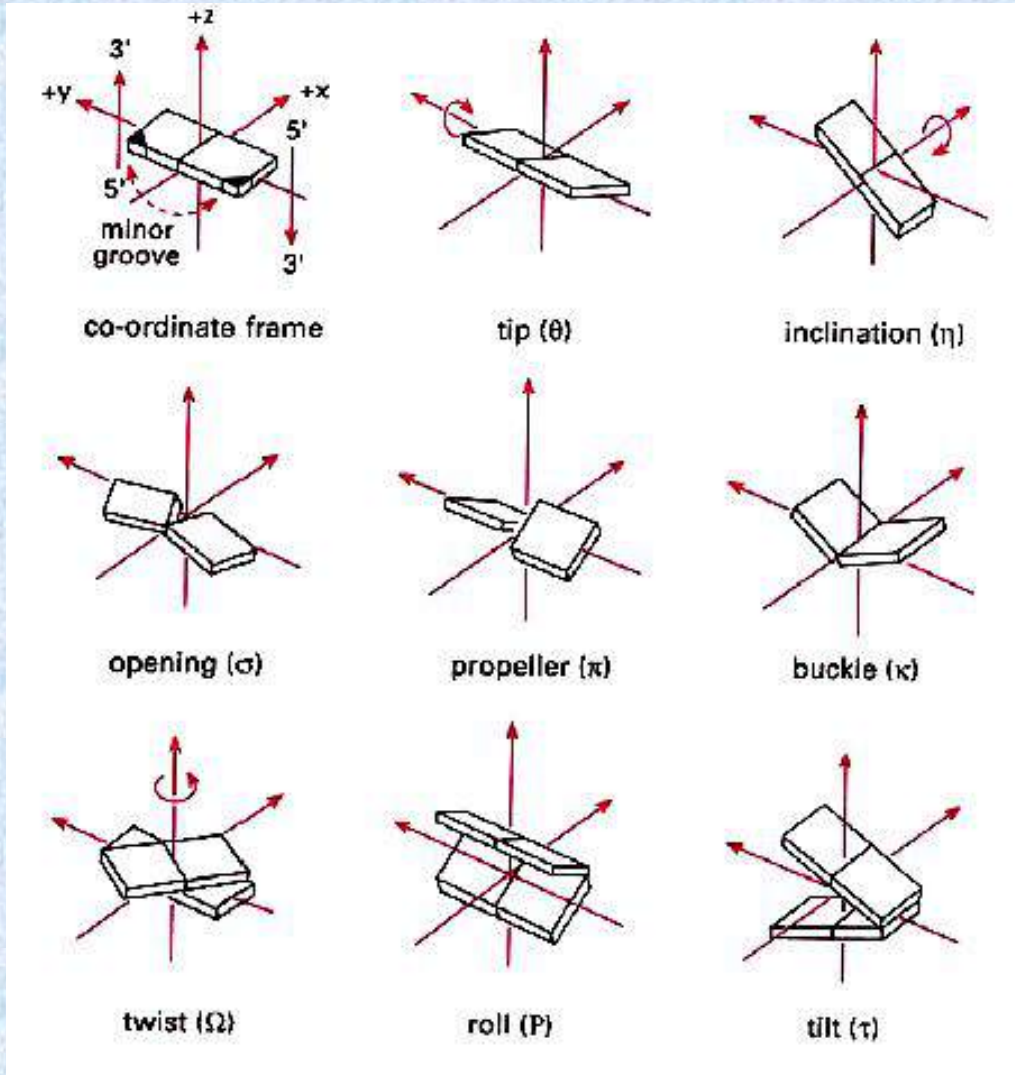
A-C Reverse Wobble

Ένα ζεύγος βάσεων



A-T

Διευθέτηση ζευγών βάσεων

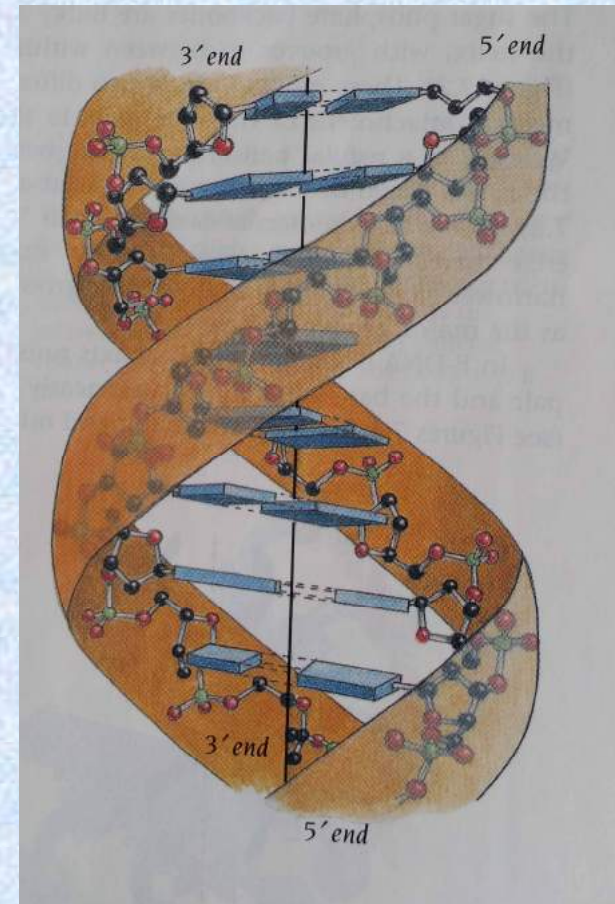
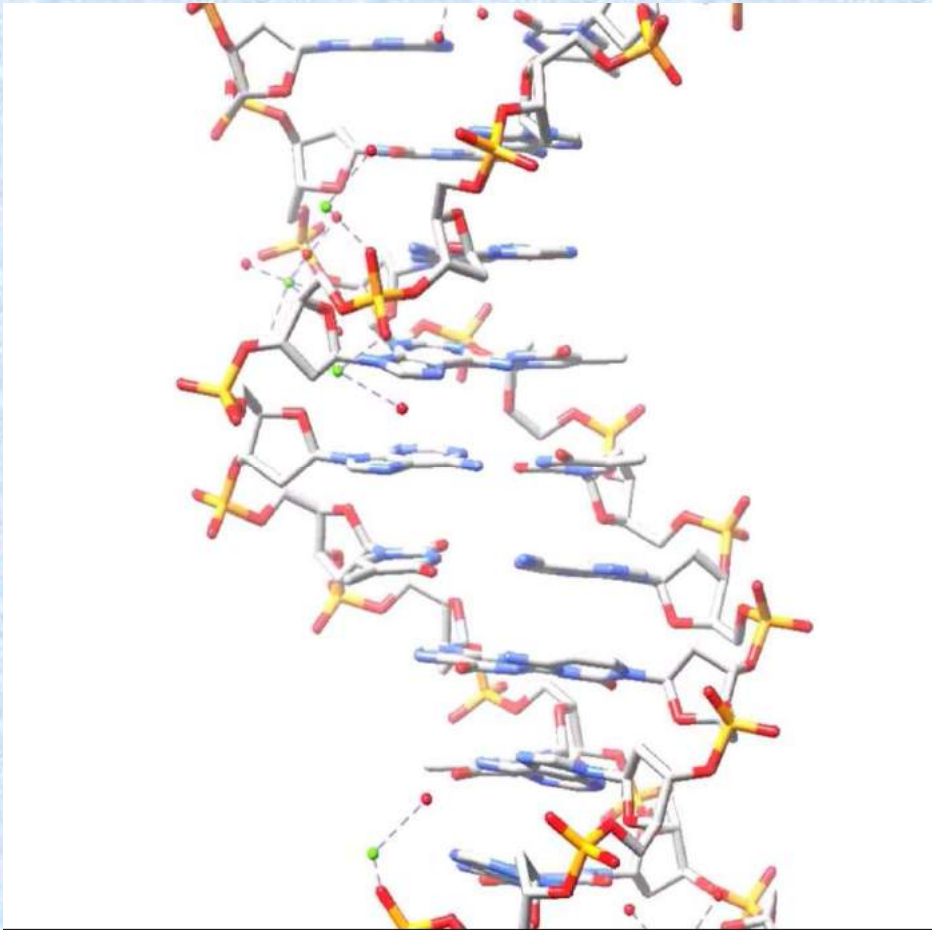


Ολική: διευθέτηση ως προς τον άξονα της έλικας

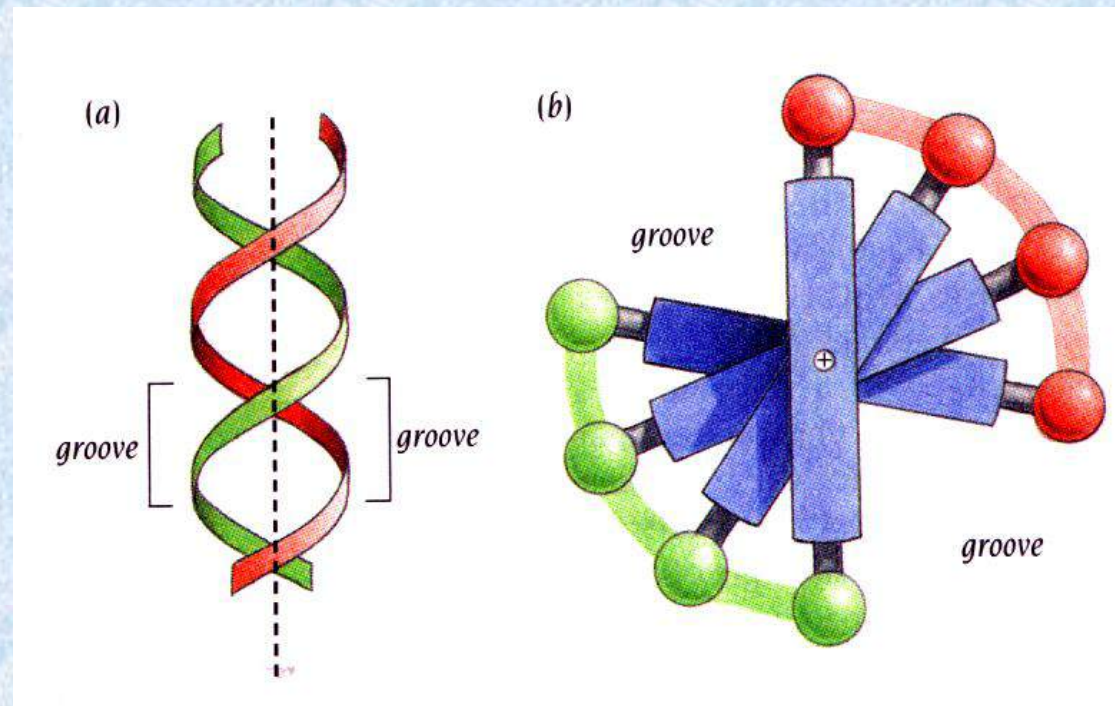
Εσωτερική: προσανατολισμός της μίας βάσης ως προς την συζυγή της στο ίδιο ζεύγος

Τοπική: διευθέτηση ως προς το επόμενο ή προηγούμενο ζεύγος

Αύλακες

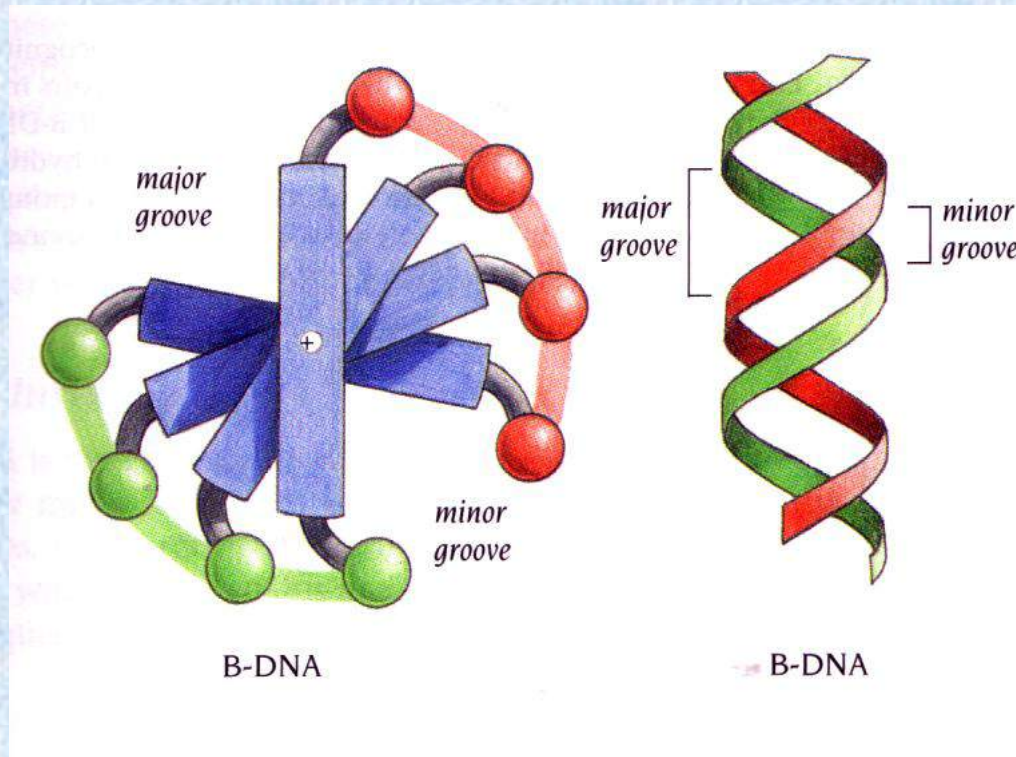


Αύλακες



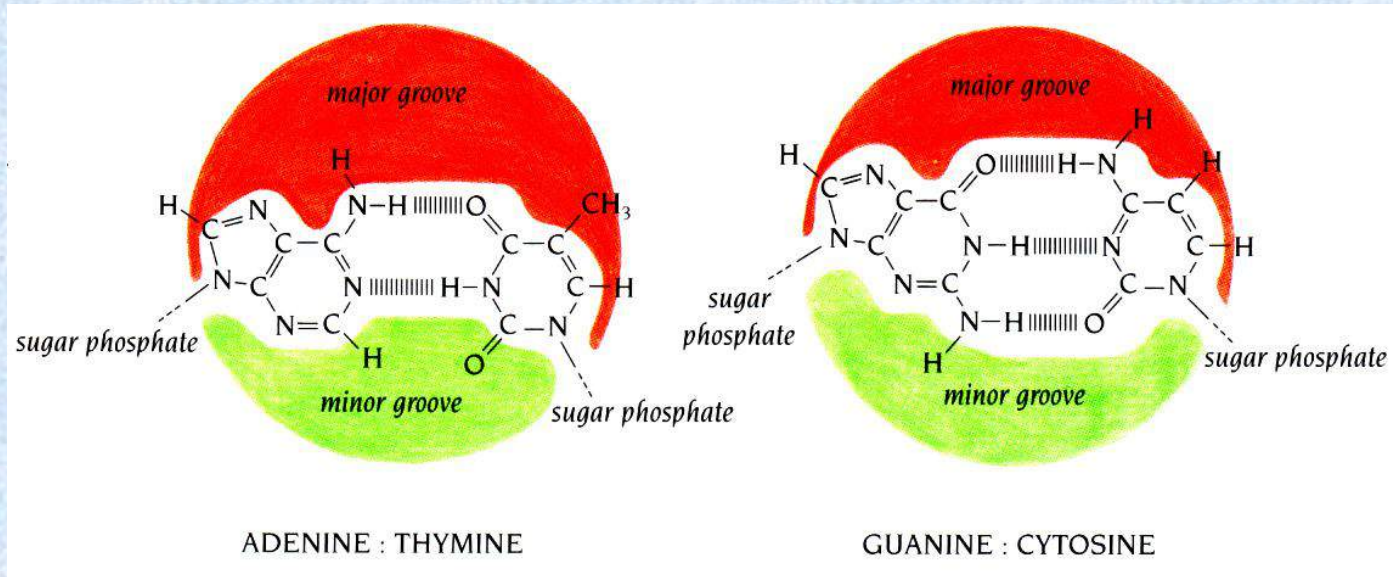
Σε μια τέλεια έλικα οι αύλακες που σχηματίζονται είναι ισοδύναμες

Μικρή και μεγάλη αύλακα



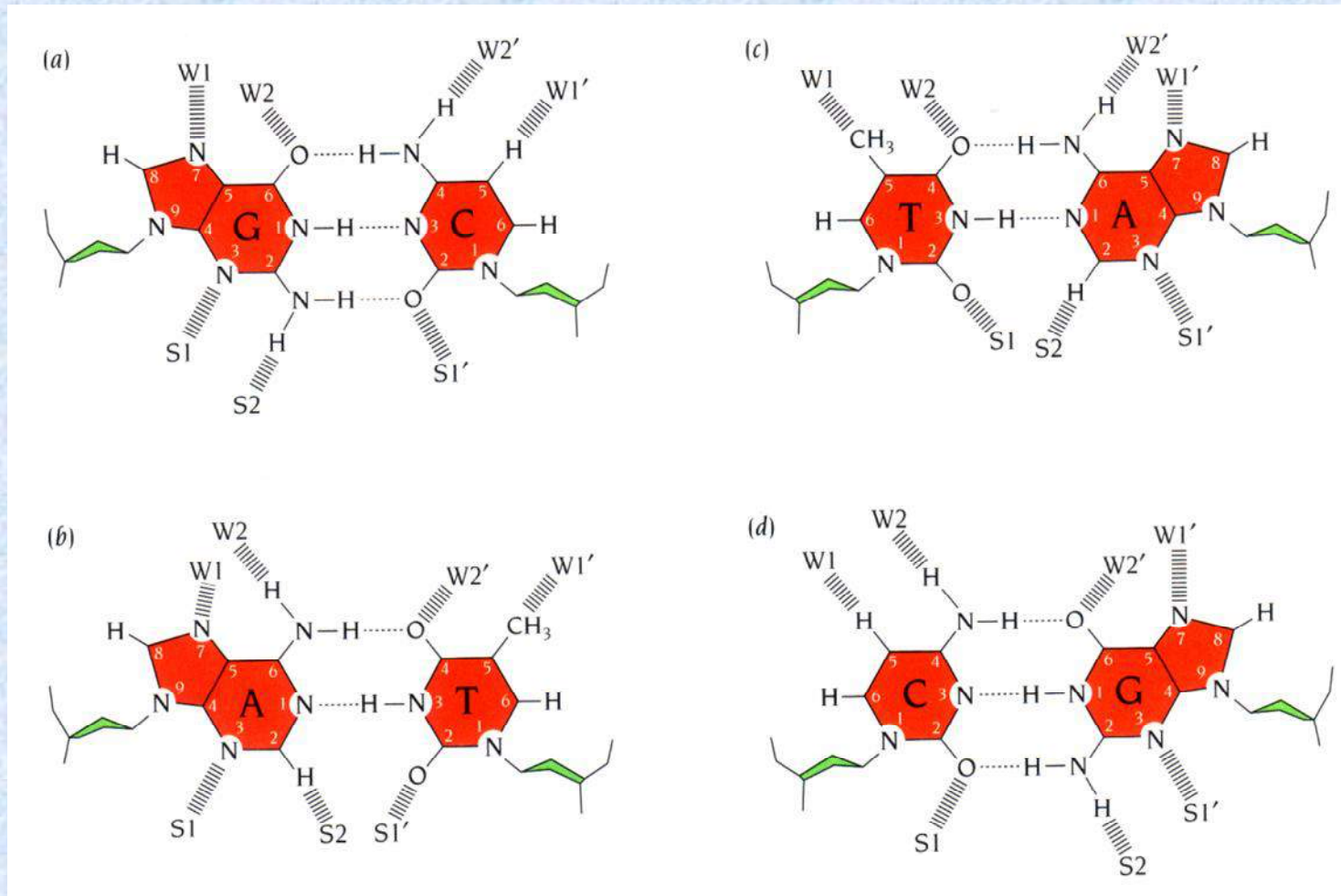
Στο B-DNA λόγω της ασύμμετρης σύνδεσης των βάσεων στον φωσφοσακχαρικό σκελετό σχηματίζονται δύο άνισες αύλακες

Μικρή και μεγάλη αύλακα



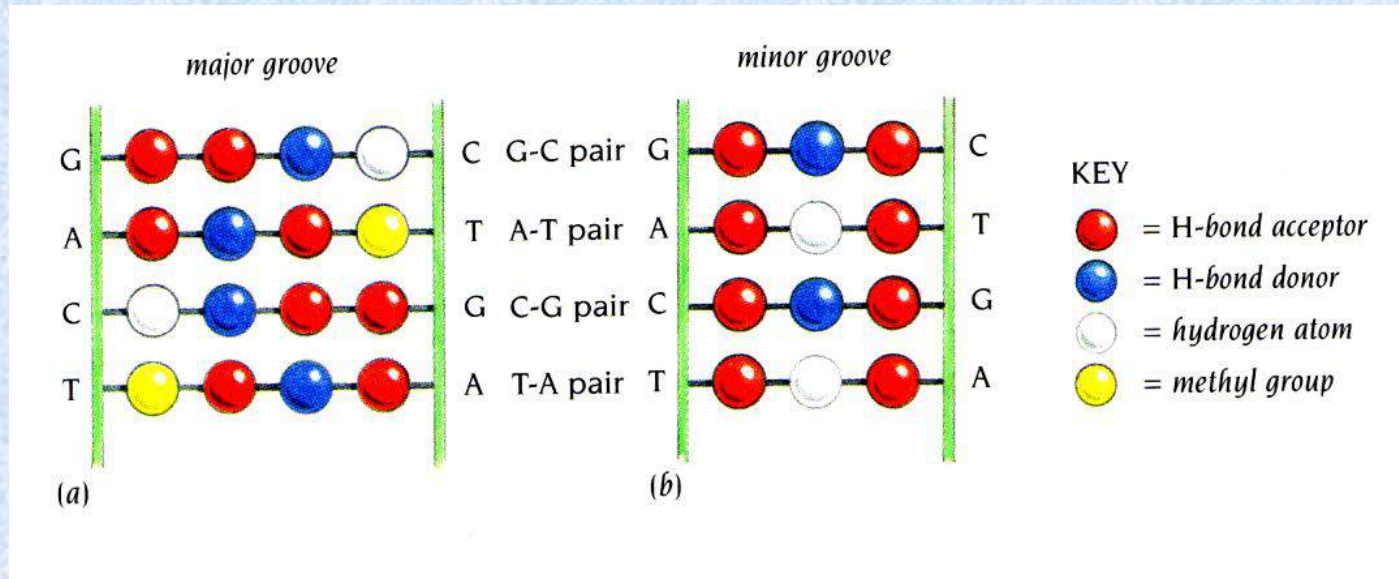
Οι χημικές ομάδες που είναι διαθέσιμες για «αναγνώριση», π.χ. από πρωτεΐνες, είναι διαφορετικές στη μεγάλη και μικρή αύλακα

Πιθανές θέσεις «αναγνώρισης» βάσεων



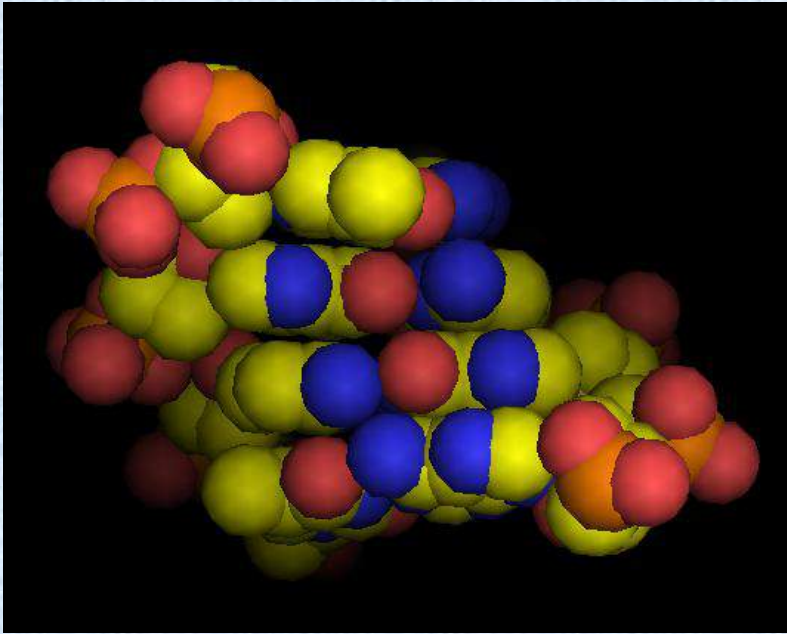
W: θέσεις αναγνώρισης από την πλευρά της μεγάλης αύλακας
S : αναγνώριση από την πλευρά της μικρής αύλακας

Κώδικας «αναγνώρισης» του DNA

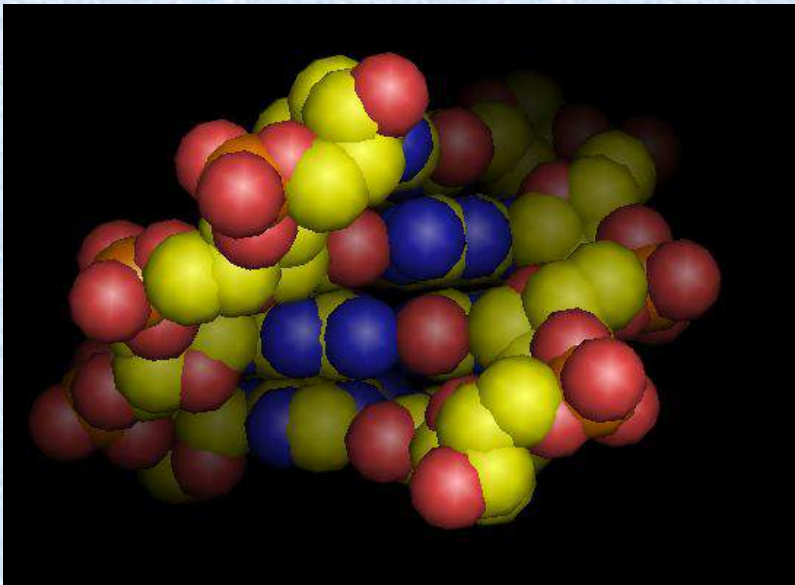


Ο χημικός κώδικας αναγνώρισης είναι πολύ διαφορετικός στη μεγάλη και μικρή αύλακα

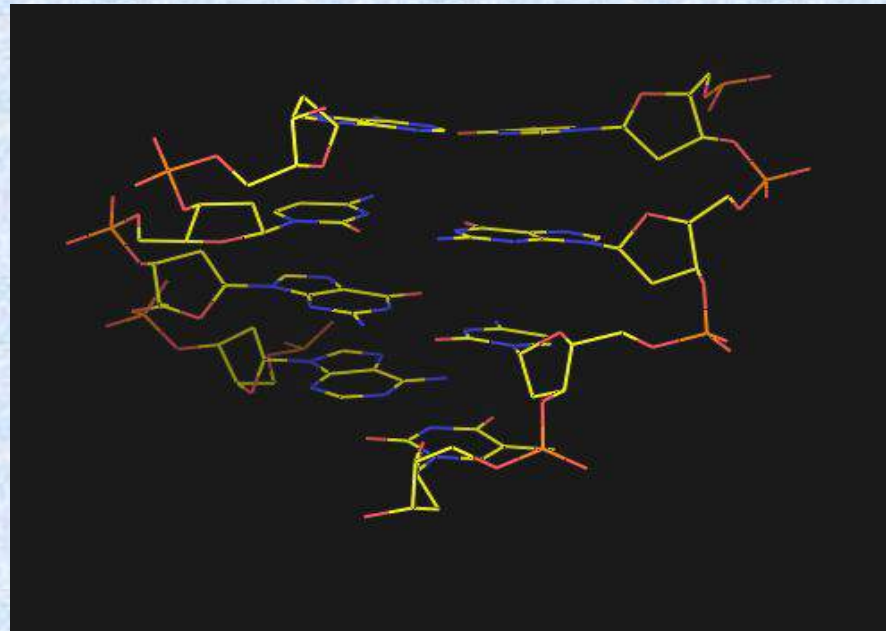
Από την παραπάνω διάταξη μπορούμε να συμπεράνουμε ότι οι πρωτεΐνες θα «αναγνώριζαν» ειδικά (specifically) το DNA καλύτερα από την μεγάλη αύλακα



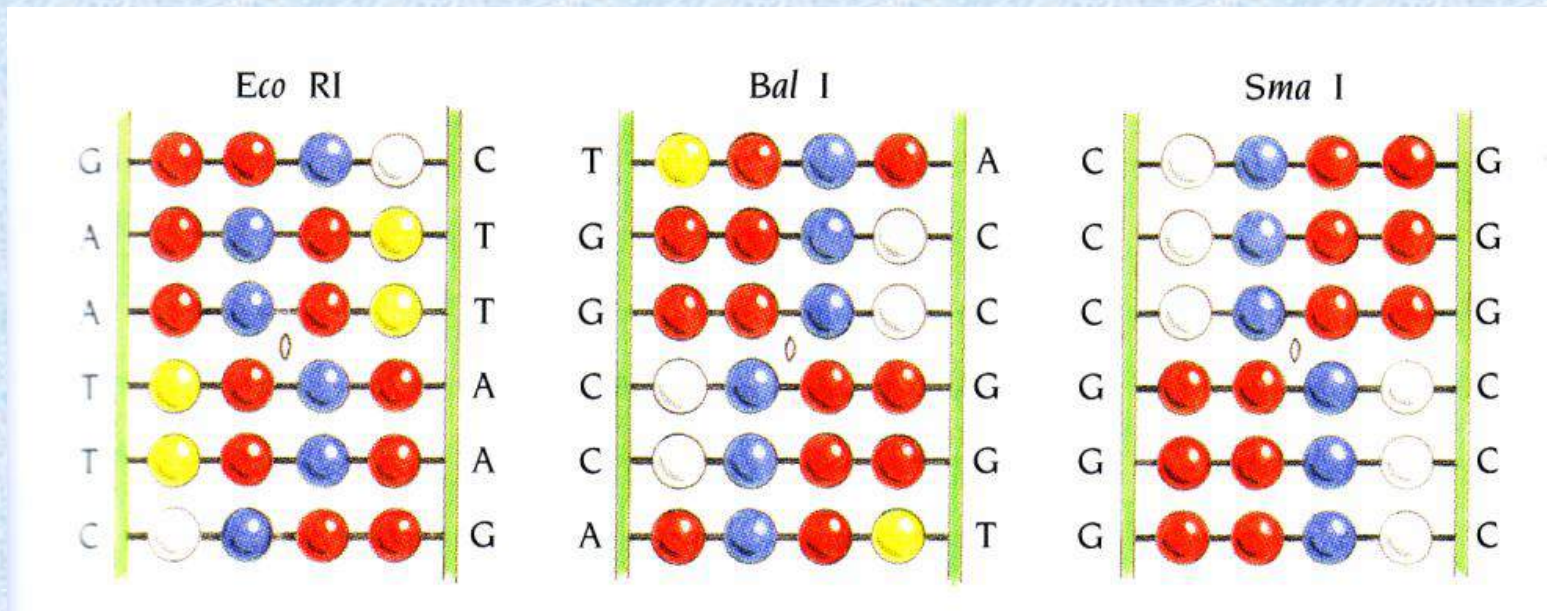
Μεγάλη αύλακα



Μικρή αύλακα



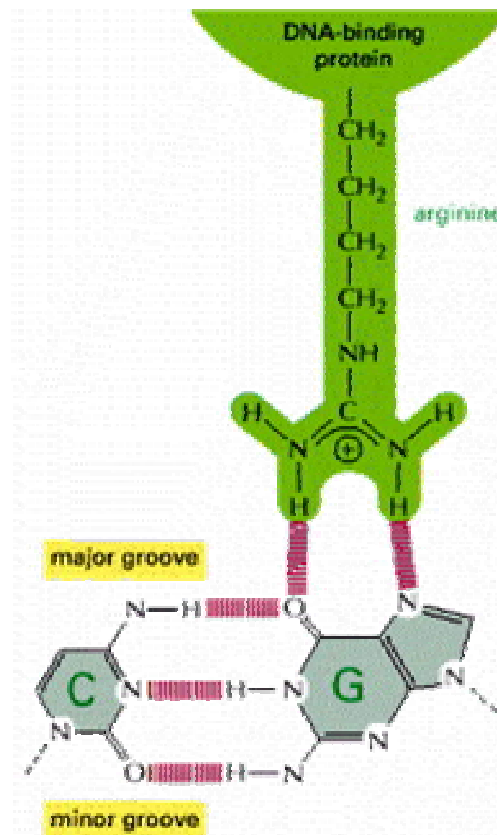
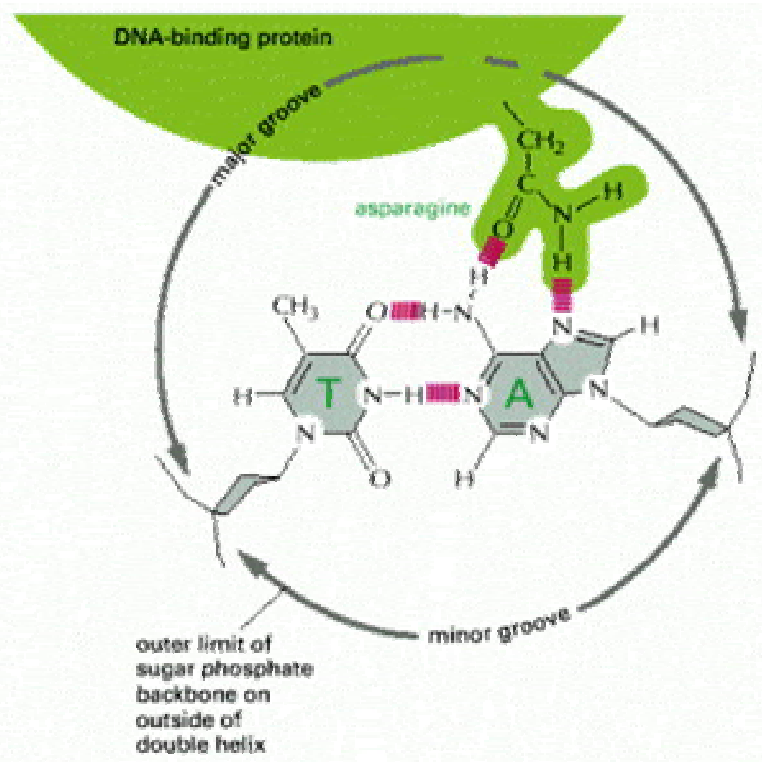
Κώδικας «αναγνώρισης» του DNA



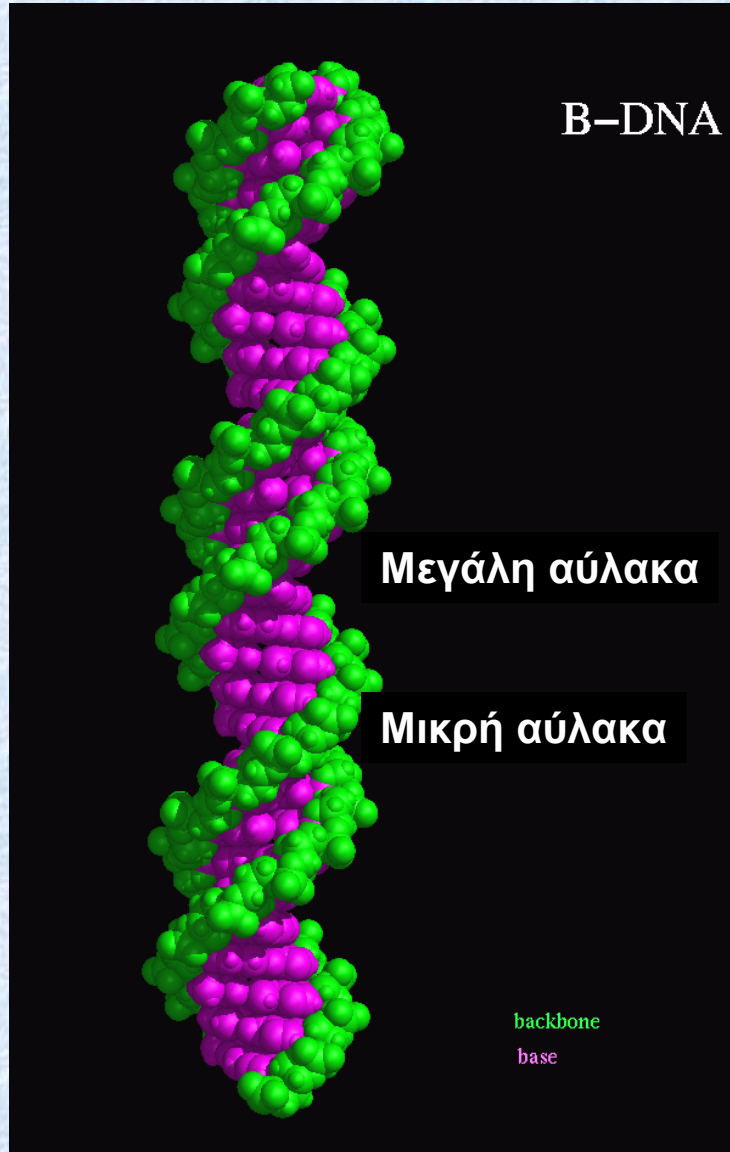
Τρία περιοριστικά ένζυμα αναγνωρίζουν την **μεγάλη αύλακα** χρησιμοποιώντας τον «κώδικα» που φαίνεται παραπάνω.

Παρατηρήστε ότι η ειδική αναγνώριση (sequence-specific recognition) επιτυγχάνεται με μικρό αριθμό βάσεων

Τυπική αναγνώριση DNA από πρωτεΐνες

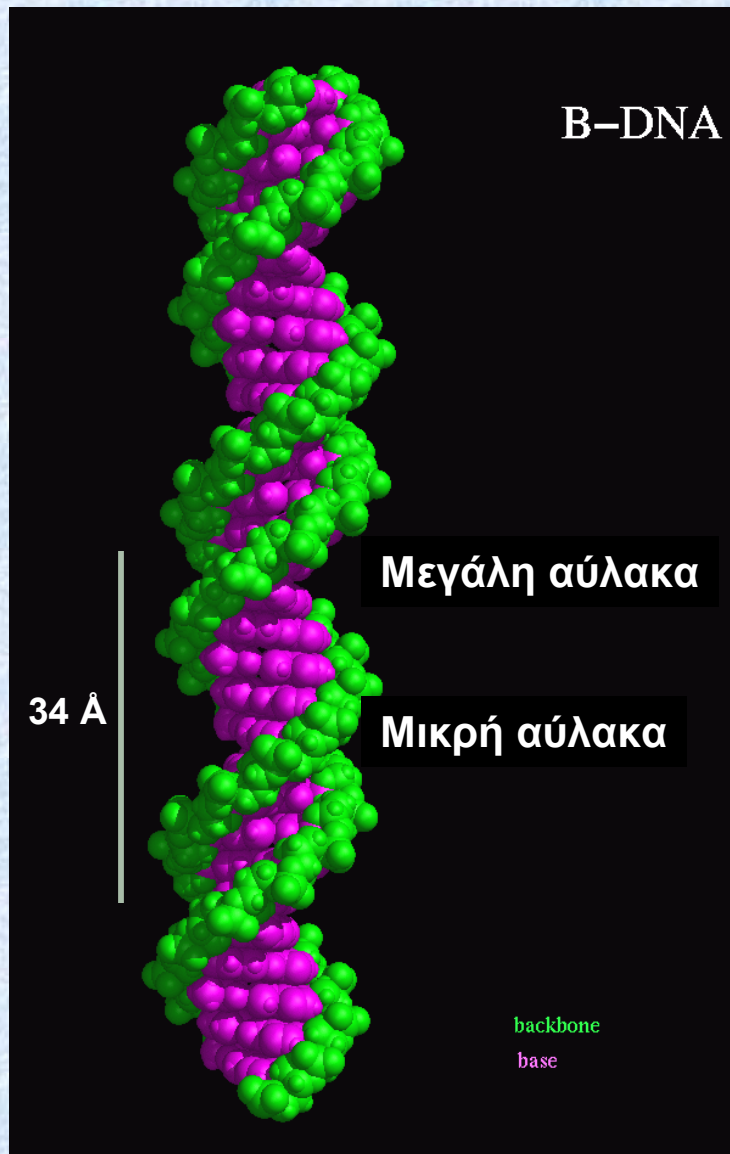


Η δομή του B-DNA



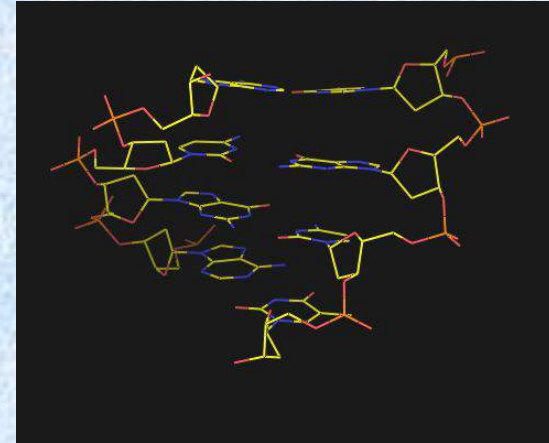
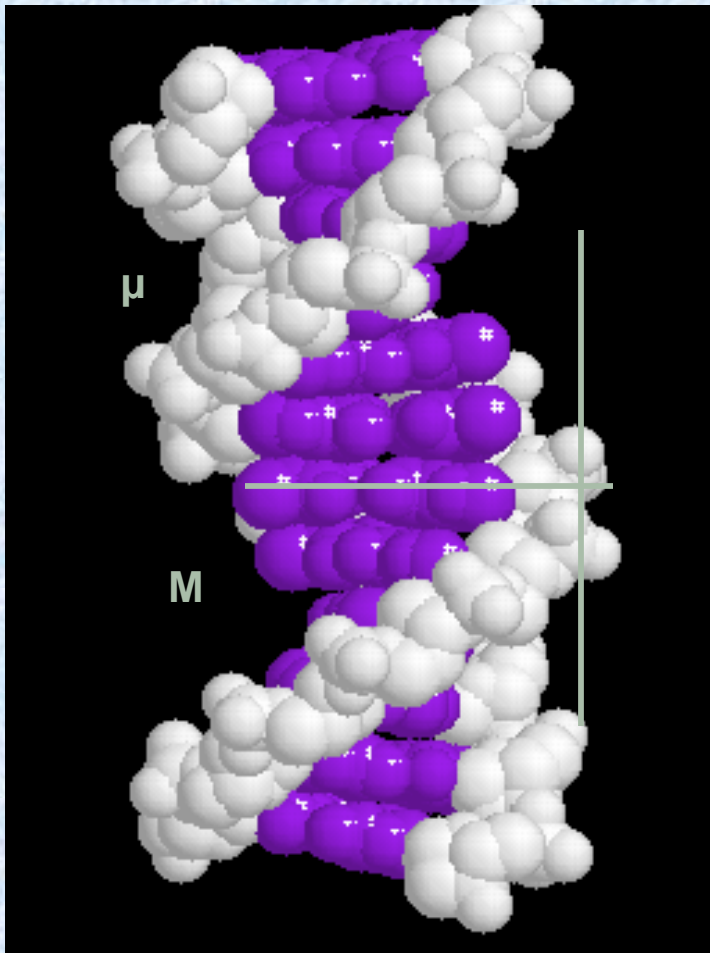
- Το B-DNA είναι η κλασική δεξιόστροφη διπλή έλικα Watson-Crick
- Η διάμετρος του είναι $\sim 20 \text{ \AA}$
- Η μεγάλη αύλακα έχει διάμετρο $\sim 12 \text{ \AA}$ και βάθος 8.5 \AA
- Η μικρή αύλακα έχει διάμετρο $\sim 6 \text{ \AA}$ και βάθος 7.5 \AA
- *In vitro* απαντάται σε πλήρως ενυδατωμένες ίνες DNA και είναι η κυρίως δομή *in vivo*

Η δομή του B-DNA



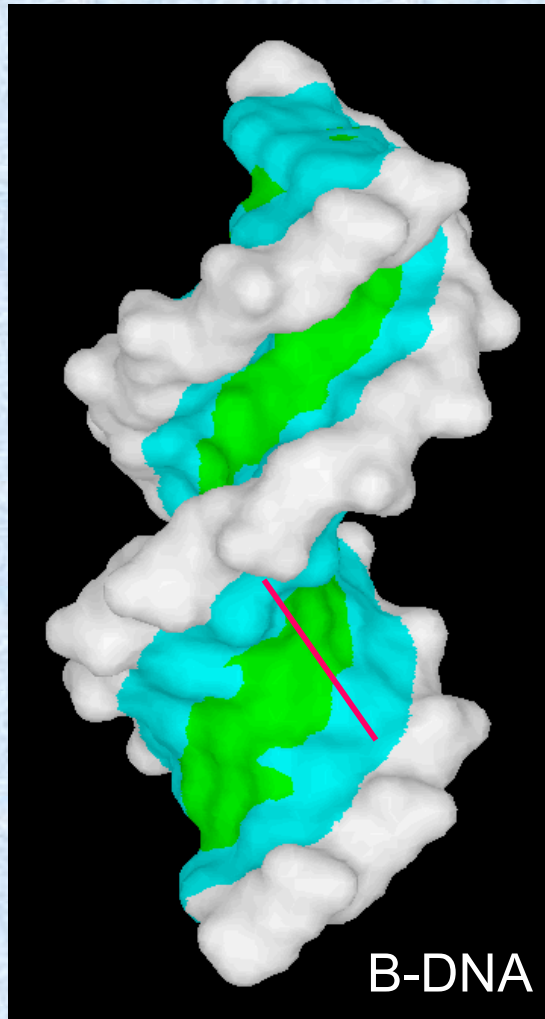
- Δεξιόστροφη έλικά
- 10 κατάλοιπα / στροφή
- Βήμα 34 Å
- Ανύψωση 3.4 Å / κατάλοιπο (ζεύγος)
- Περιστροφή 36° / κατάλοιπο (ζεύγος)

Η δομή του B-DNA



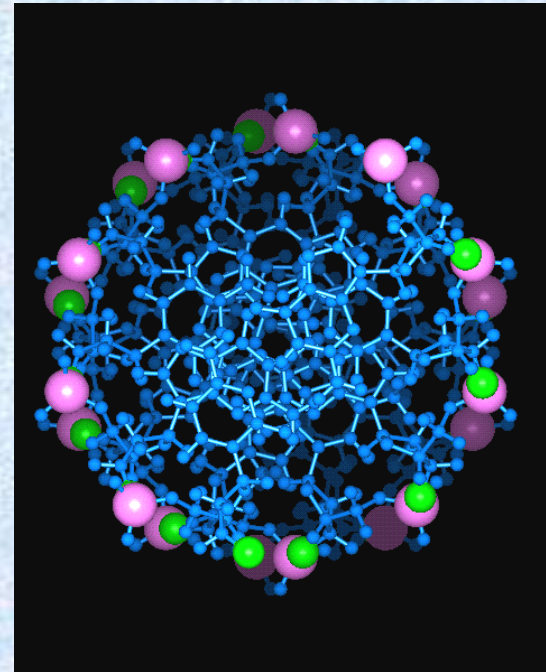
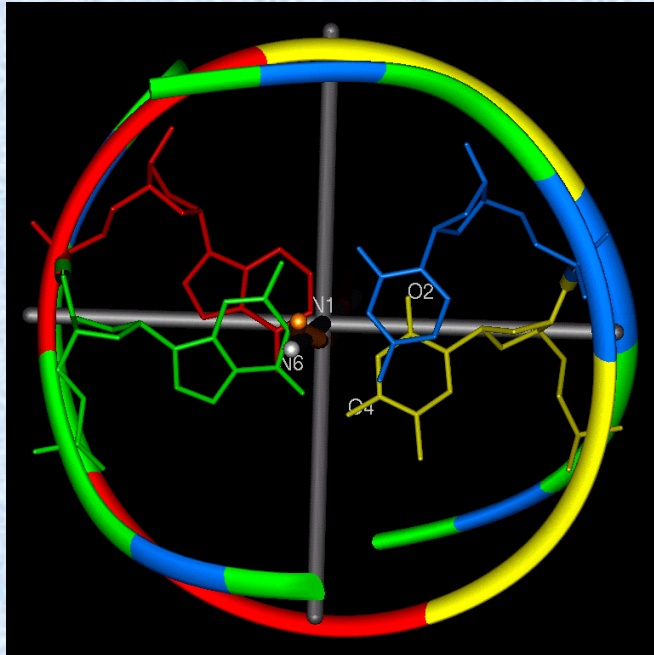
Οι βάσεις είναι σχεδόν κάθετες στον άξονα της έλικας (tilt 6°)

Η δομή του B-DNA



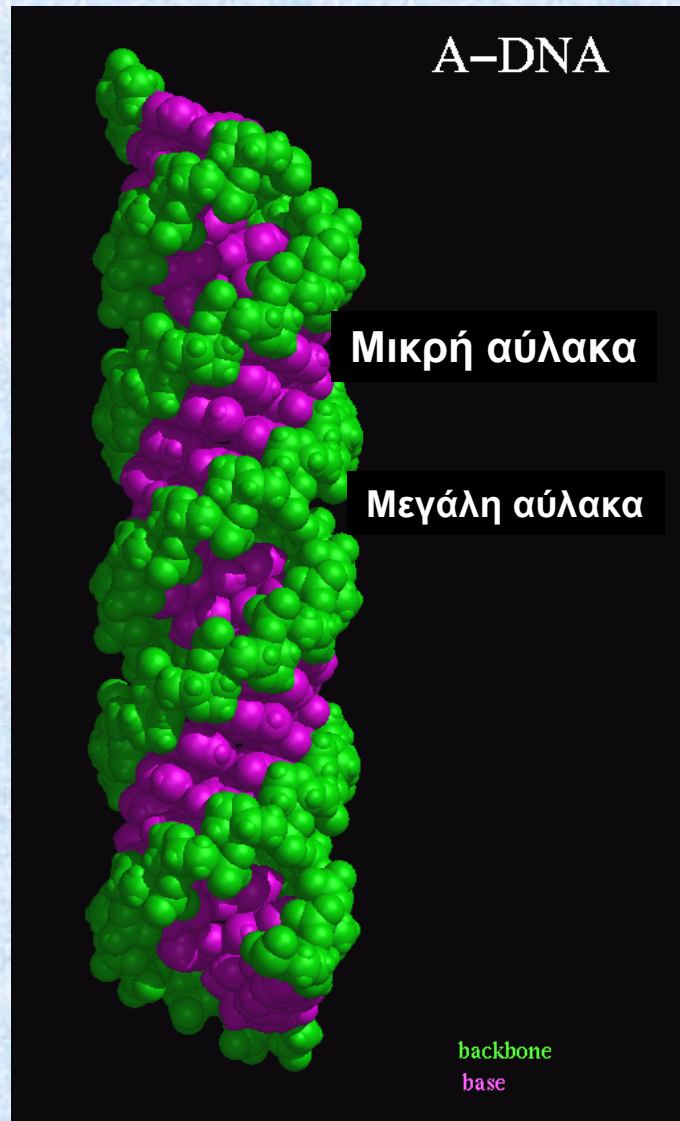
Μία α -έλικα και δύο
αντιπαράλληλοι β -κλώνοι
«ταιριάζουν» στις διαστάσεις της
μεγάλης αύλακας του B-DNA

Η δομή του B-DNA



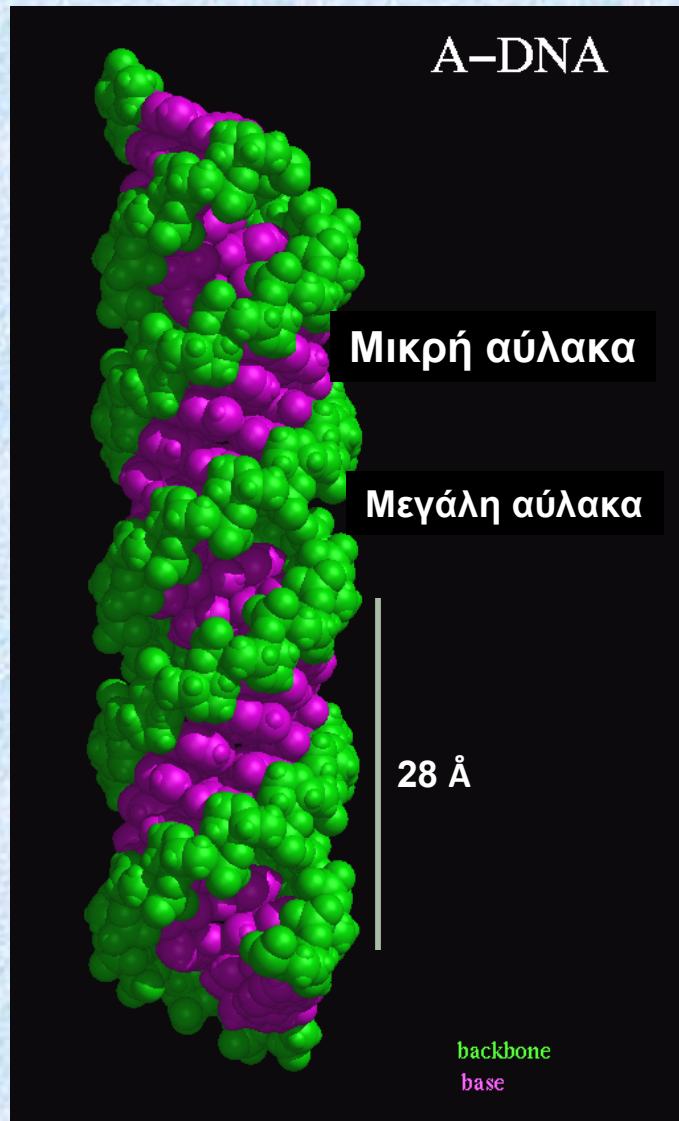
Ο άξονας της έλικας του B-DNA περνά σχεδόν από το κέντρο του κάθε ζεύγους βάσεων

Η δομή του A-DNA



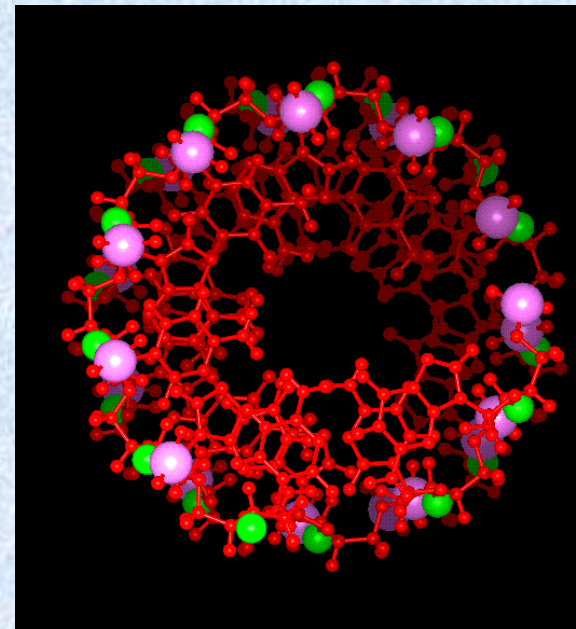
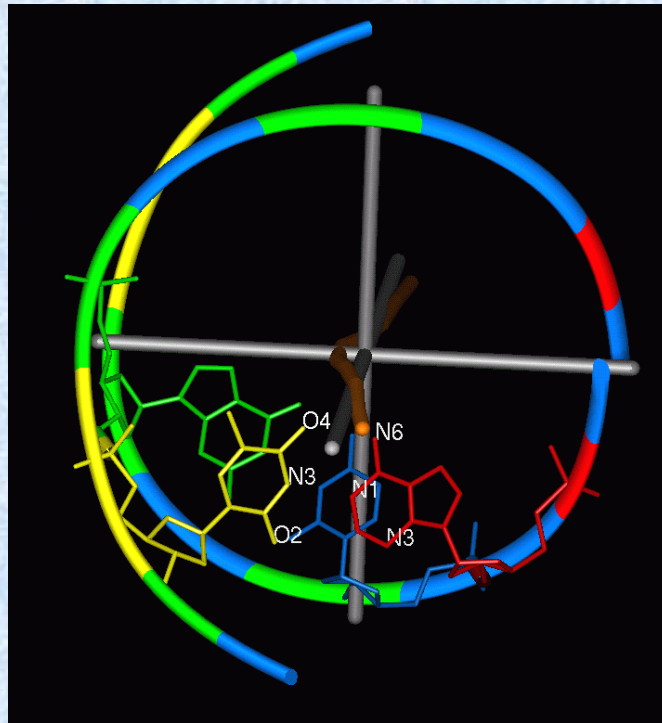
- Το A-DNA είναι δεξιόστροφο
- Η έλικα είναι πλατύτερη και κοντύτερη σε σχέση με το B-DNA
- Η μεγάλη αύλακα είναι πολύ βαθιά ενώ η μικρή αύλακα είναι πολύ ρηχή
- Είναι μια δομή DNA που παρατηρείται *in vitro* όταν η υγρασία είναι ~75 %

Η δομή του A-DNA



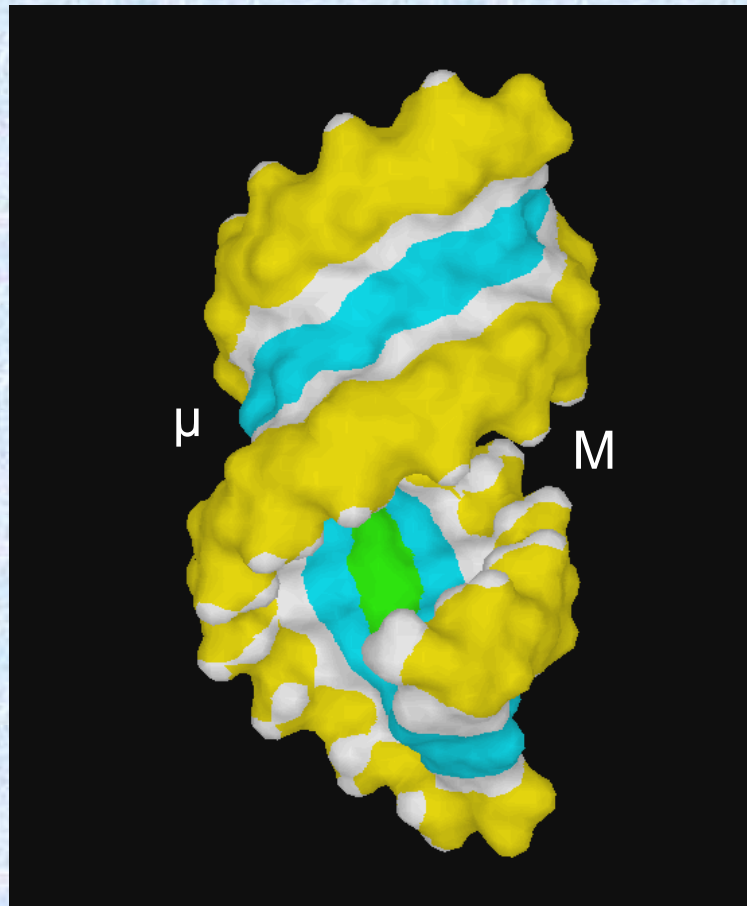
- Δεξιόστροφη έλικά
- 11 κατάλοιπα / στροφή
- Βήμα 28 Å
- Ανύψωση 2.55 Å / κατάλοιπο
- Περιστροφή 33° / κατάλοιπο

Η δομή του A-DNA



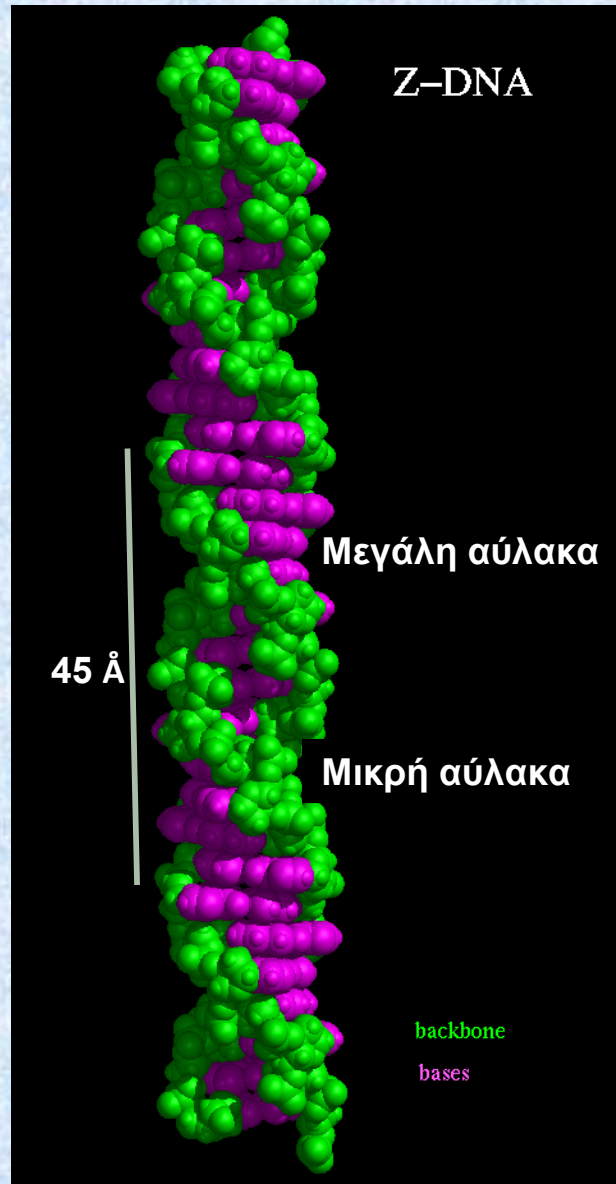
Ο άξονας της έλικας του A-DNA είναι μετατοπισμένος πρὸς τη μεγάλη αύλακα

Η δομή του A-DNA



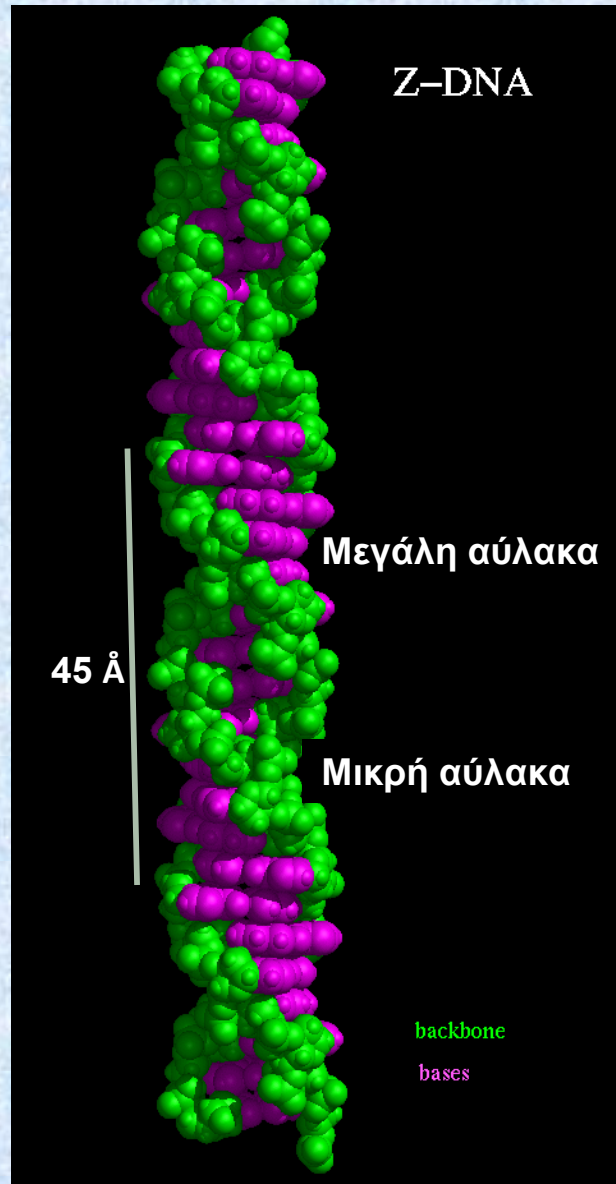
Μοντέλο A-DNA

Η δομή του Z-DNA



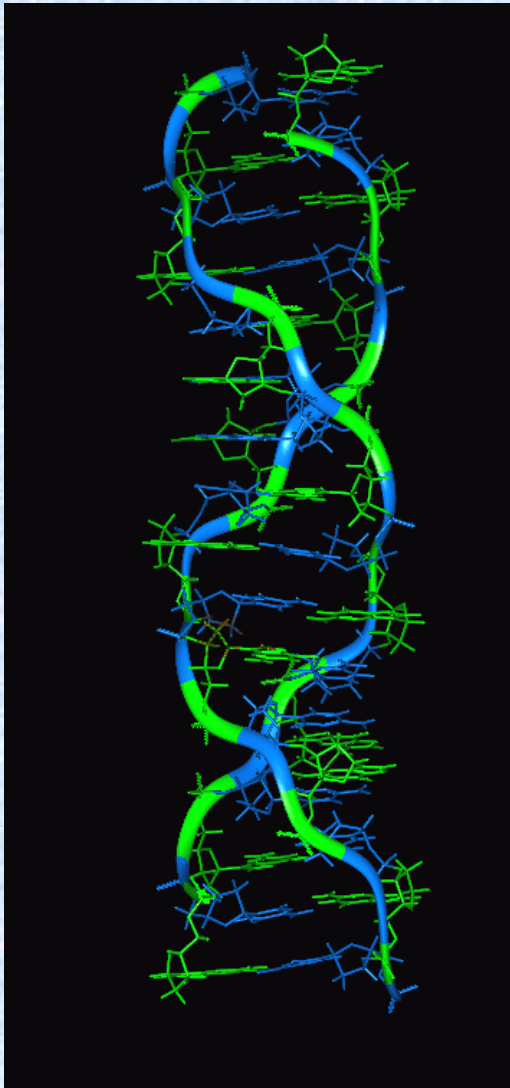
- Αριστερόστροφη έλικά
- Η έλικά είναι λεπτότερη και μακρύτερη σε σχέση με το B-DNA
- Τα φωσφορικά του σκελετού σχηματίζουν μία τεθλασμένη γραμμή ζιγκ-ζαγκ
- Η δομή αυτή ευνοείται από υψηλή αλατότητα και από επαναλαμβανόμενες ακολουθίες CG ή πυριμιδινών πουρινών
- Η μικρή αύλακα είναι πολύ βαθιά και στενή ενώ σχεδόν δεν υπάρχει μεγάλη αύλακα
- Μπορεί να βρίσκεται και *in vivo* π.χ. σε περιοχές μεθυλιωμένου DNA (C*)

Η δομή του Z-DNA



- Η επαναλαμβανόμενη μονάδα είναι ουσιαστικά ένα διμερές
- 12 κατάλοιπα / στροφή (6 διμερή)
- Βήμα 45 Å
- Ανύψωση 3.7 Å / κατάλοιπο

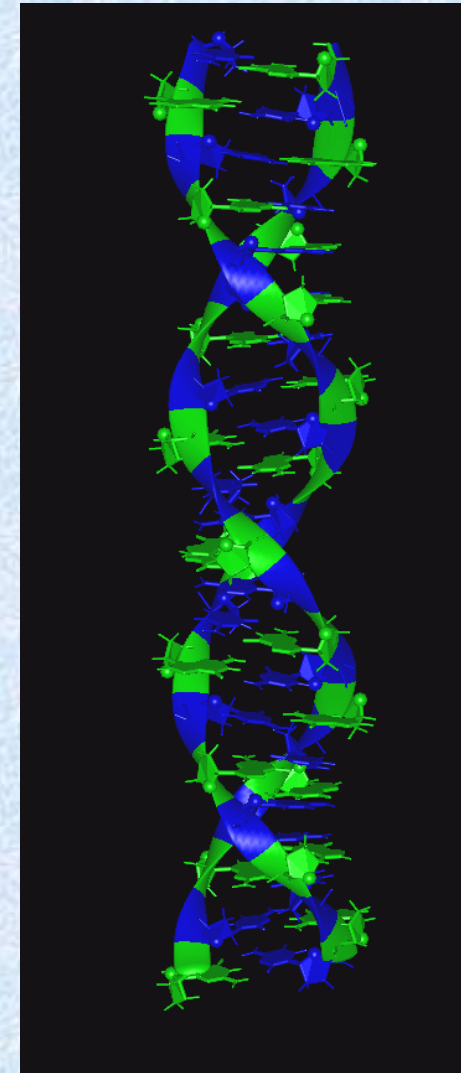
Η δομή του Z-DNA



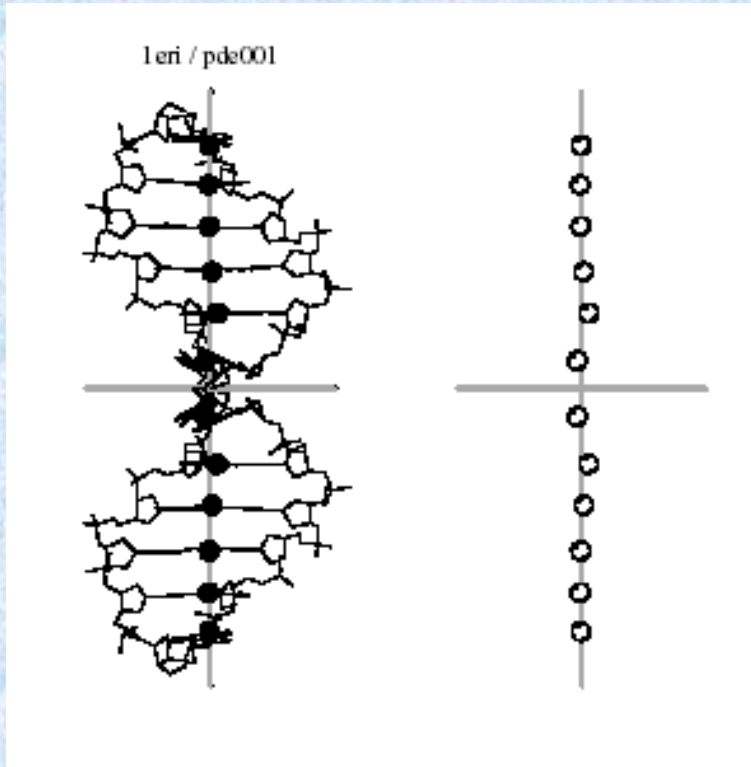
Το πακετάρισμα των βάσεων CG και GC δεν είναι ισοδύναμο στο Z-DNA.

Το πακετάρισμα GC είναι πολύ καλύτερο.

Το αποτέλεσμα είναι η επαναλαμβανόμενη μονάδα να είναι ένα δινουκλεοτίδιο

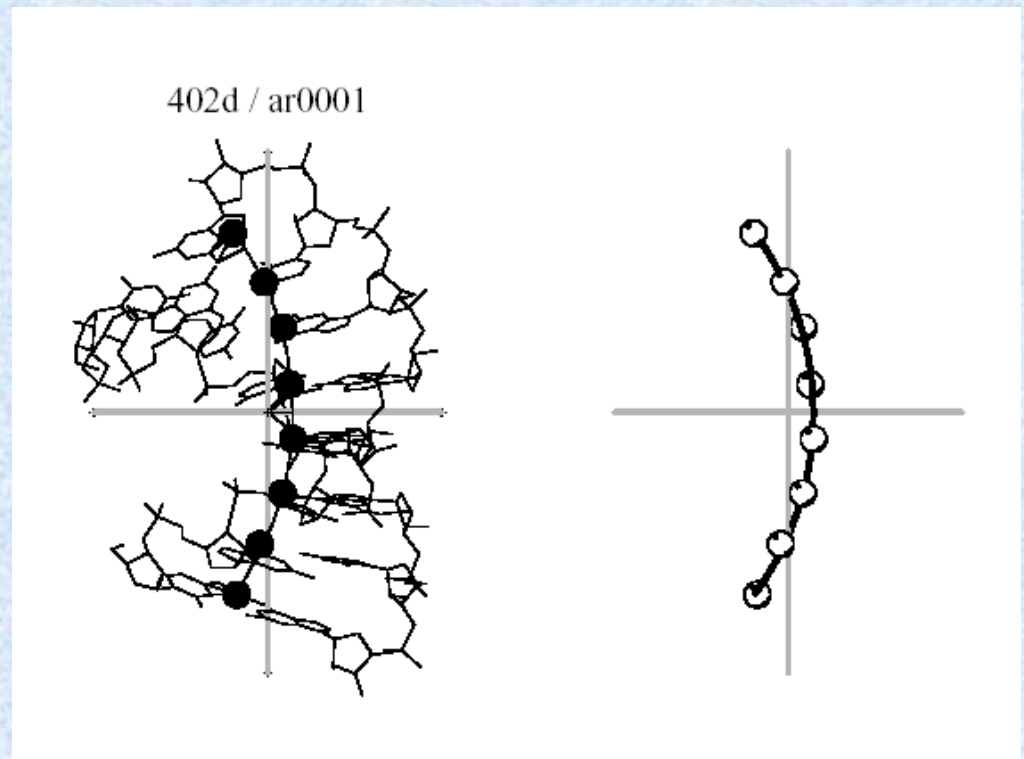


Το DNA είναι εύκαμπτο μόριο



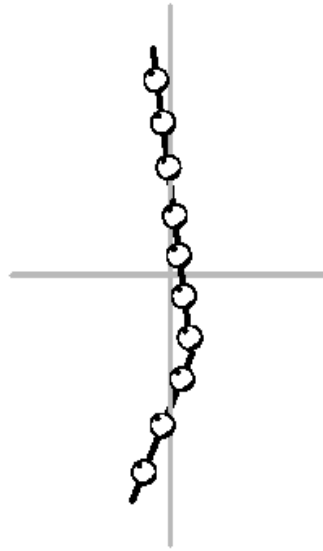
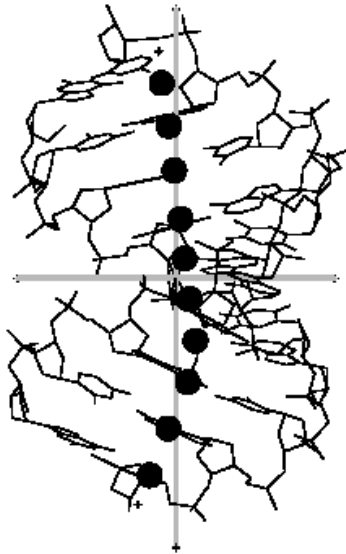
Μη παραμορφωμένο B-DNA

Τοξοειδής παραμόρφωση



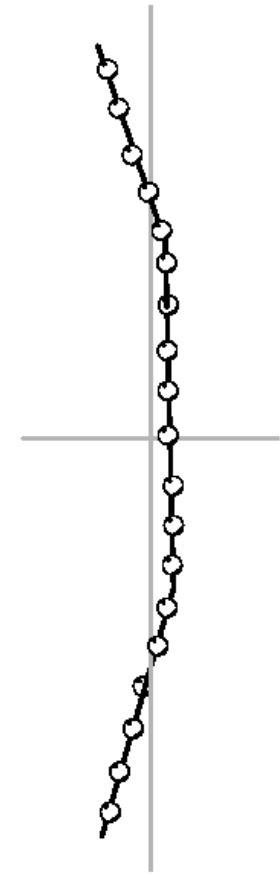
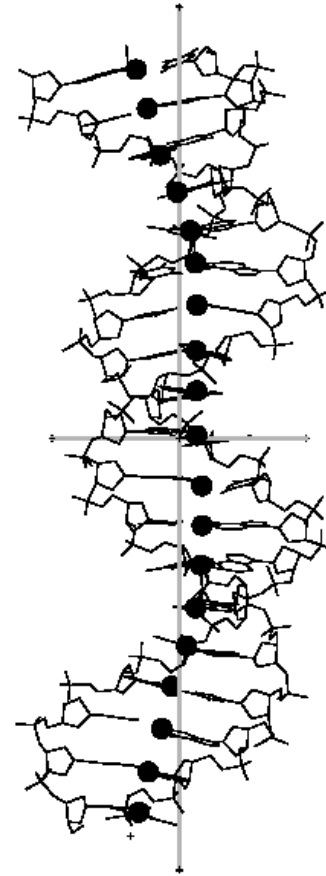
Το DNA είναι εύκαμπτο μόριο

1d88 / ahj044



Απλή στρέβλωση

3cro / pdr001



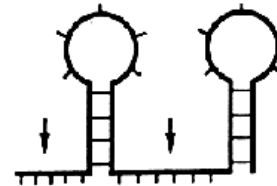
Σύνθετη στρέβλωση

Το RNA είναι πολύ πιο εύκαμπτο

a. DUPLEXES



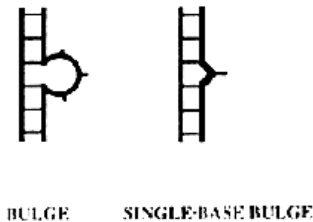
b. SINGLE STRANDED REGIONS



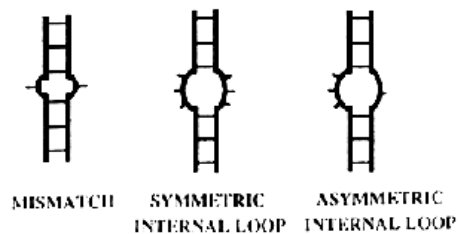
c. HAIRPINS



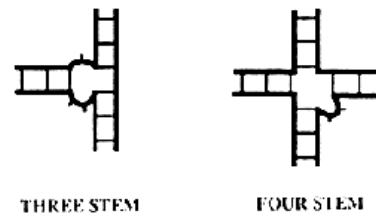
d. BULGES



e. INTERNAL LOOPS

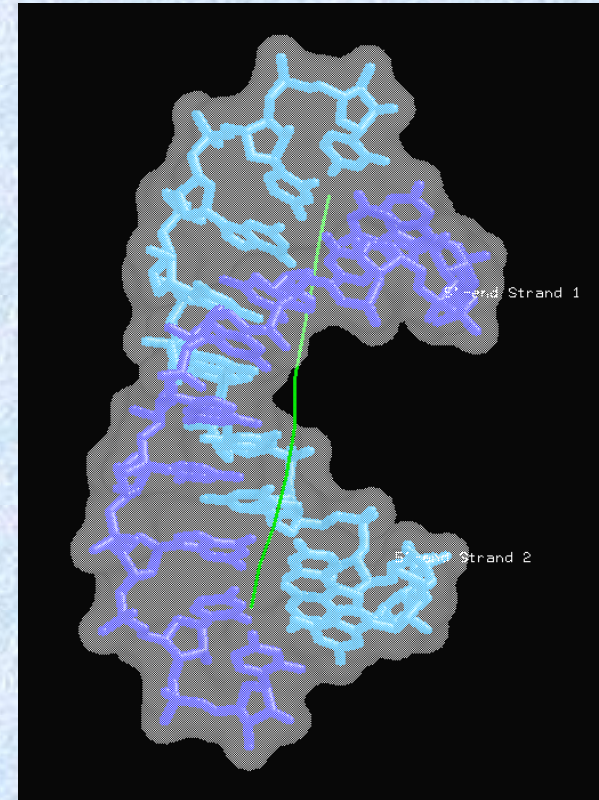
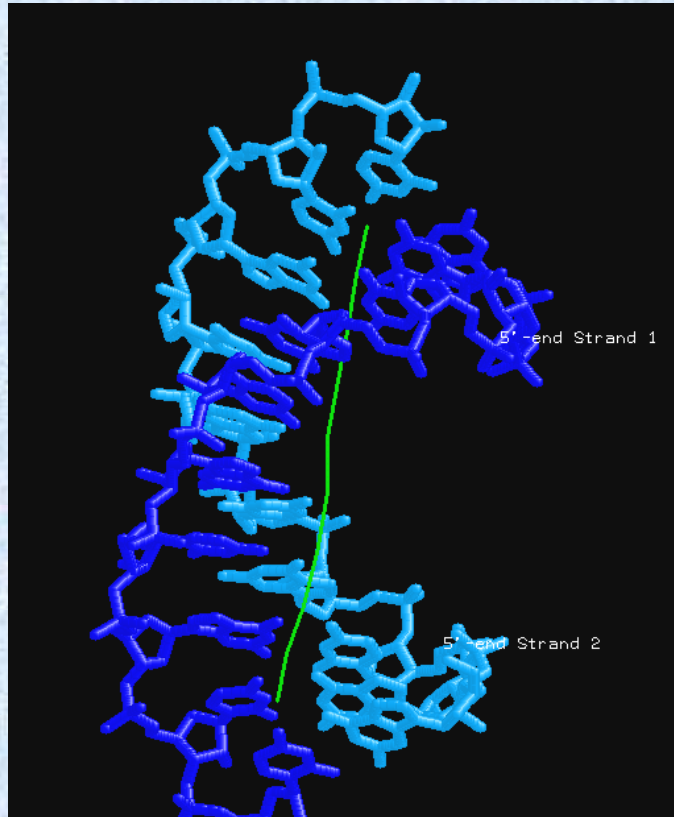
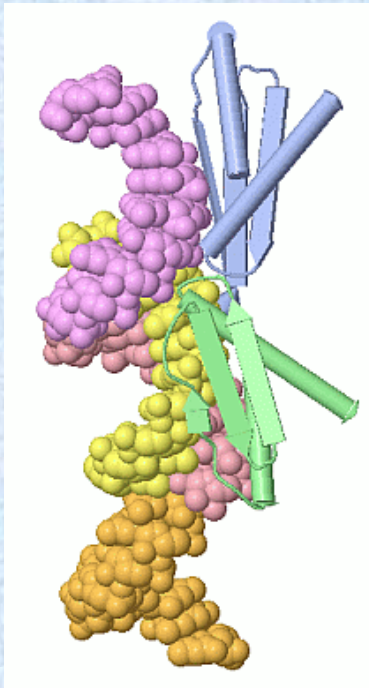


f. JUNCTIONS



Η δομή του RNA σε ένα σύμπλοκο

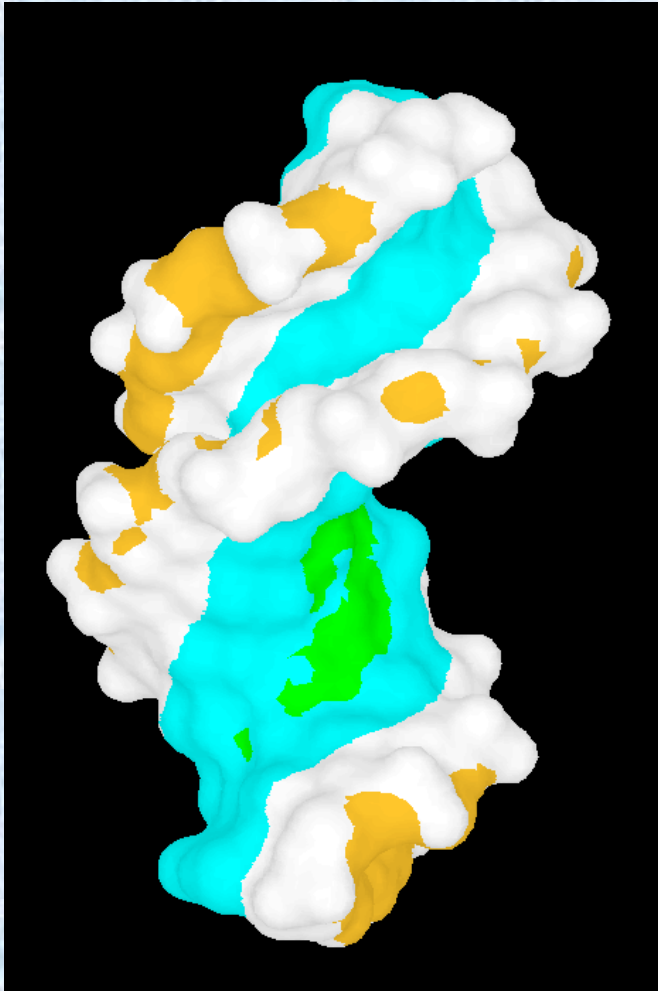
Xenopus laevis RNA-binding protein-A σε σύμπλοκο με dsRNA



Το RNA συνήθως λαμβάνει την A διαμόρφωση, δηλ. αντίστοιχη διαμόρφωση με το A-DNA

Η δομή ενός RNA/DNA υβριδίου

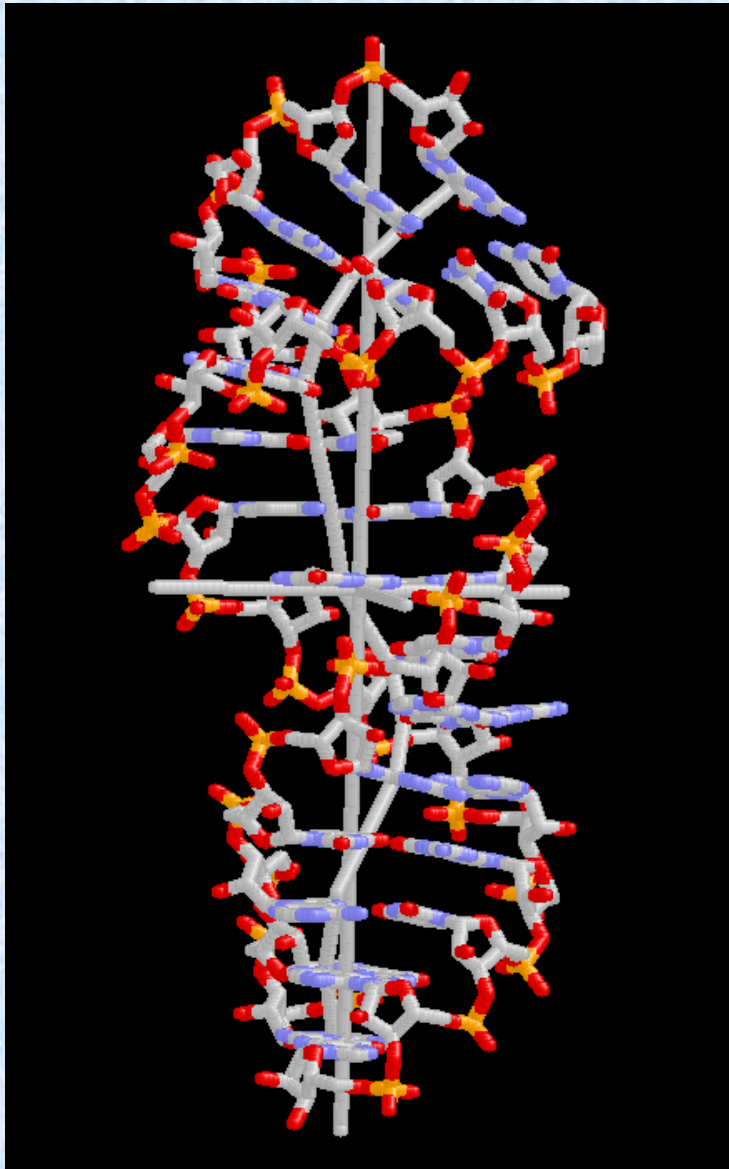
DNA/RNA (σε σύμπλοκο με Τορο Ι)



Και εδώ η δομή
είναι A

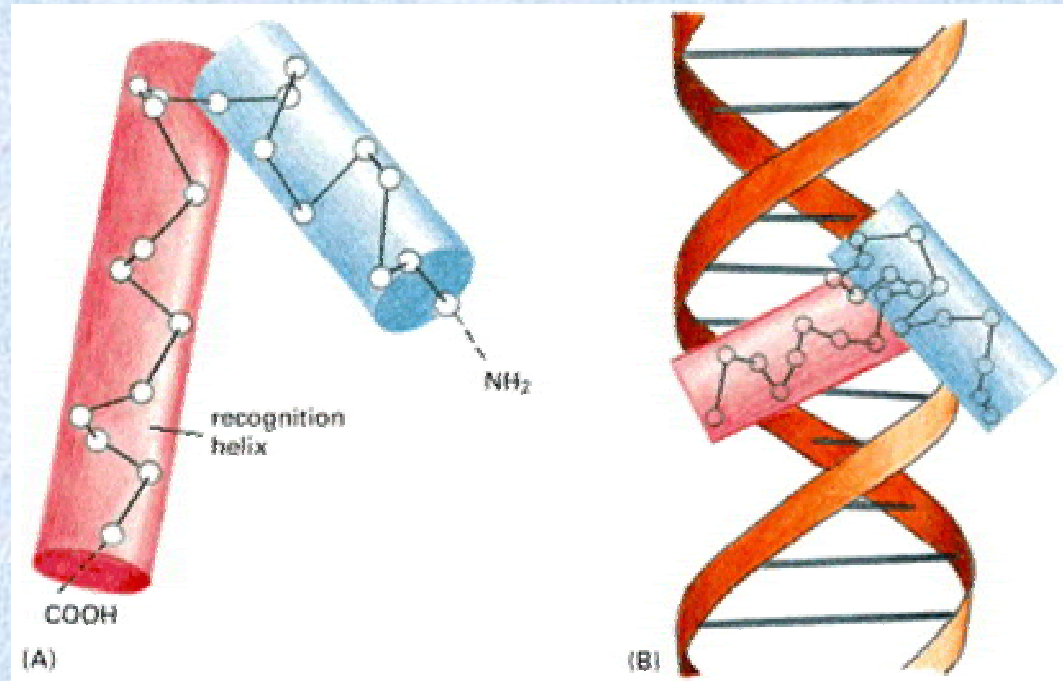
5' G A A A A U U U U T A 3'

Ριβοσωμικό RNA

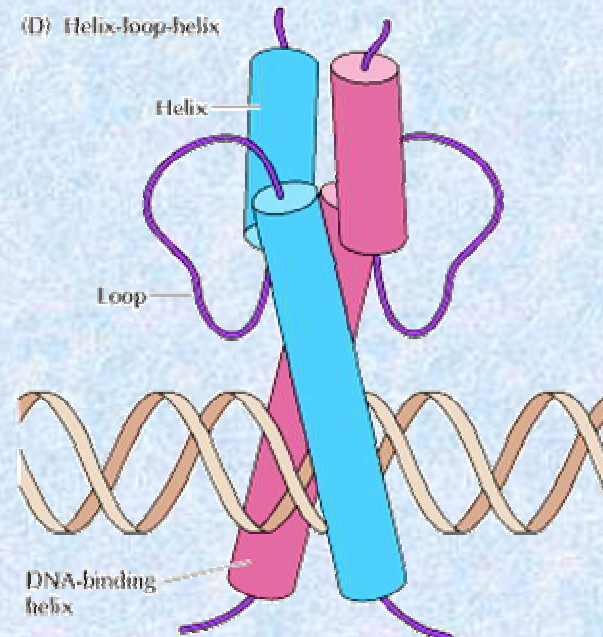
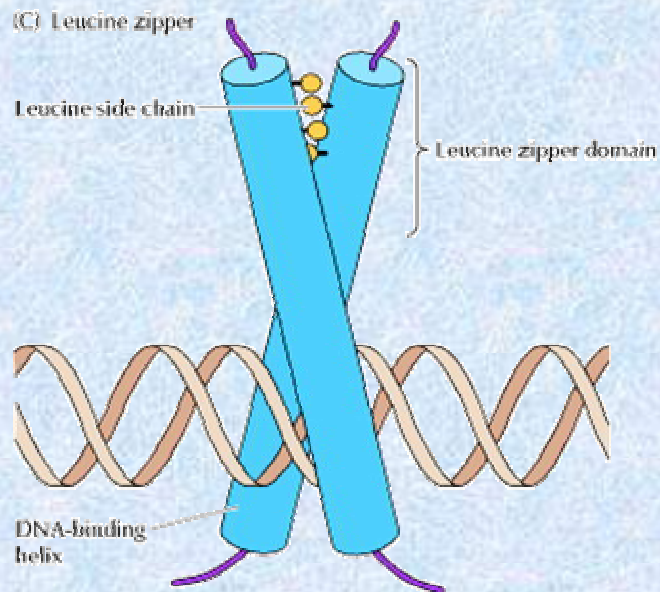
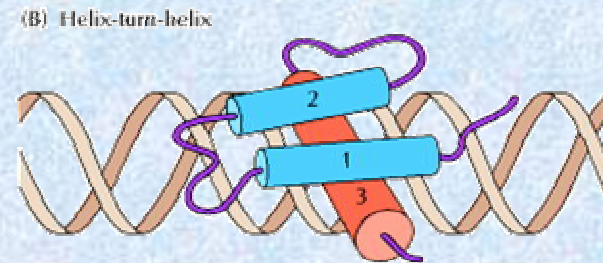
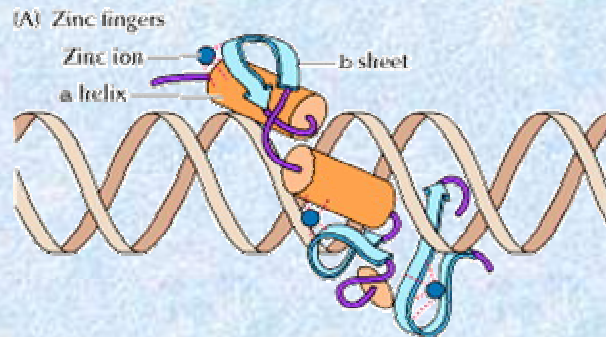


Κομμάτι 5S ριβοσωμικού RNA
σε σύμπλοκο με την πρωτεΐνη
L25 (δεν φαίνεται)

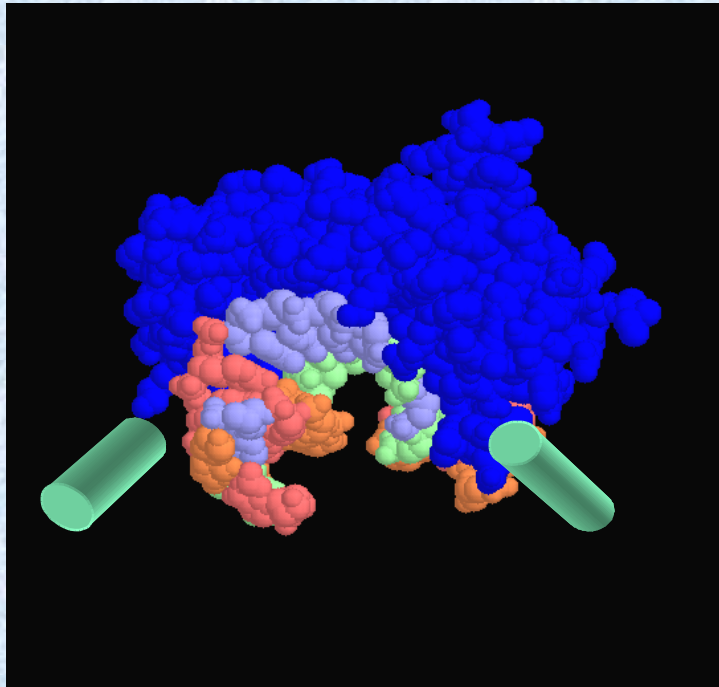
Μοτίβο έλικά-στροφή-έλικά



Διάφορα μοτίβα πρόσδεσης DNA

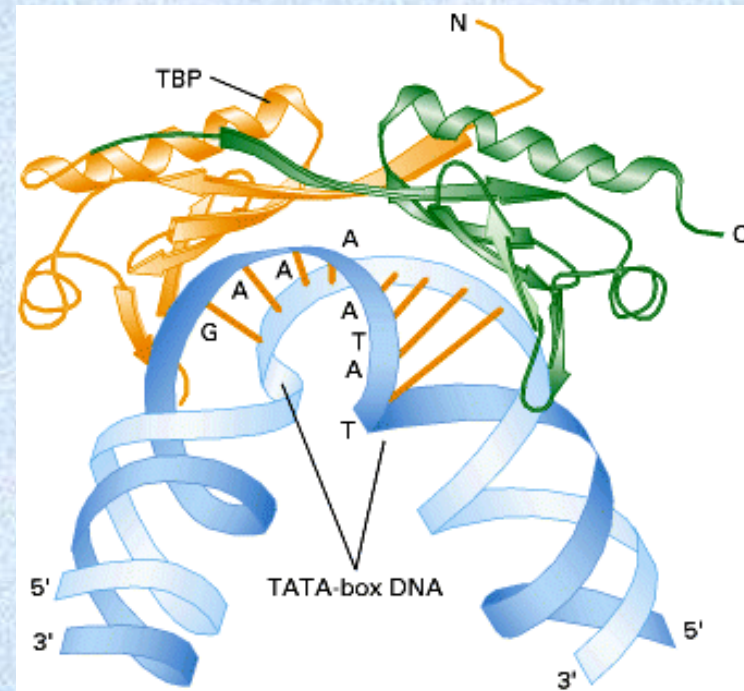


Δέσμευση της TBP με το TATA-box



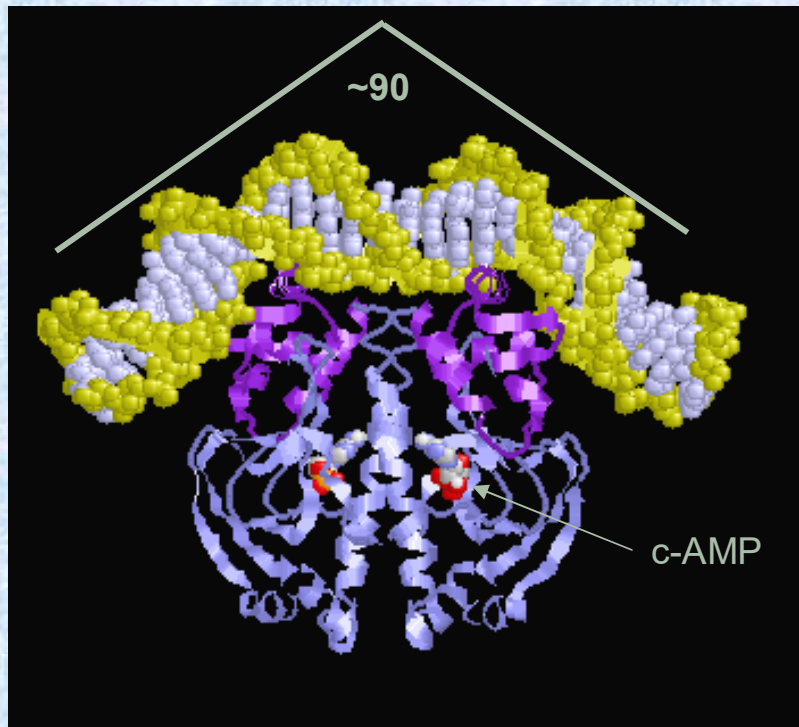
Η TBP δεσμεύεται στη μικρή αύλακα την οποία ανοίγει διάπλστα (induced fit)

Ο άξονας του DNA σχηματίζει γωνία $\sim 100^\circ$



Σύμπλοκα DNA/πρωτεϊνών

CAP-DNA



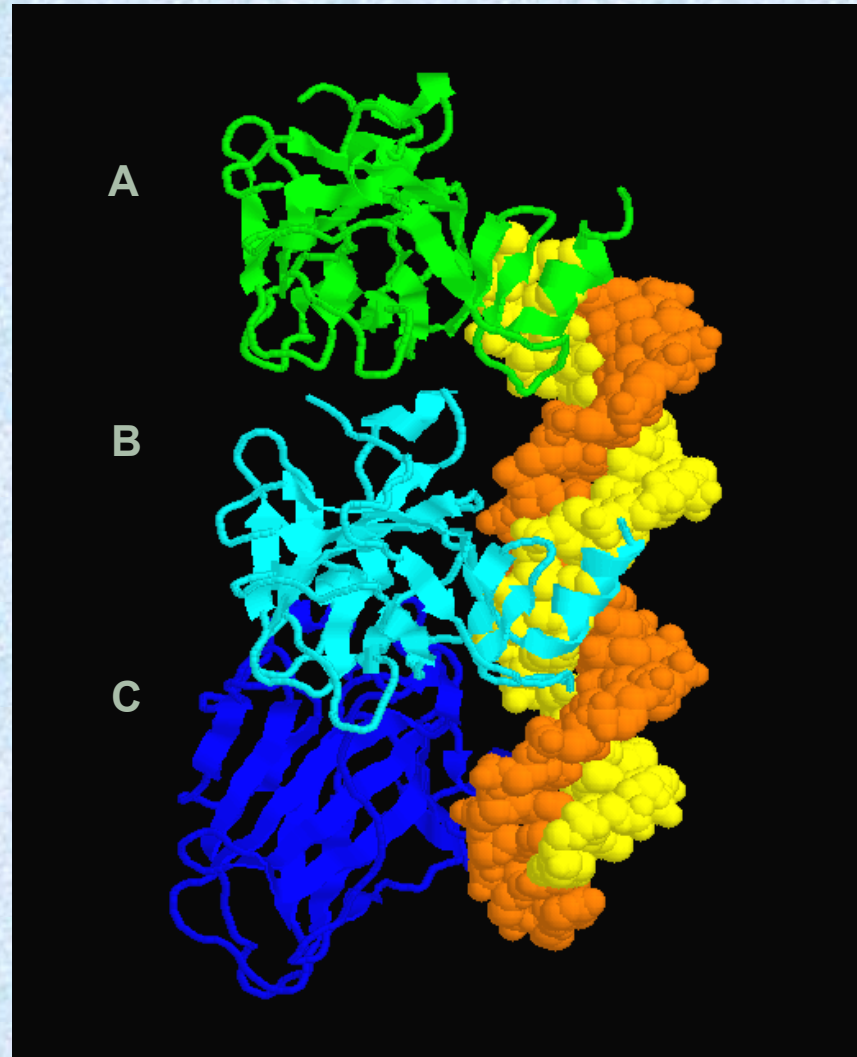
Είναι διμερές

Αναγνωρίζει την ακολουθία
TGTGA***TCACA**

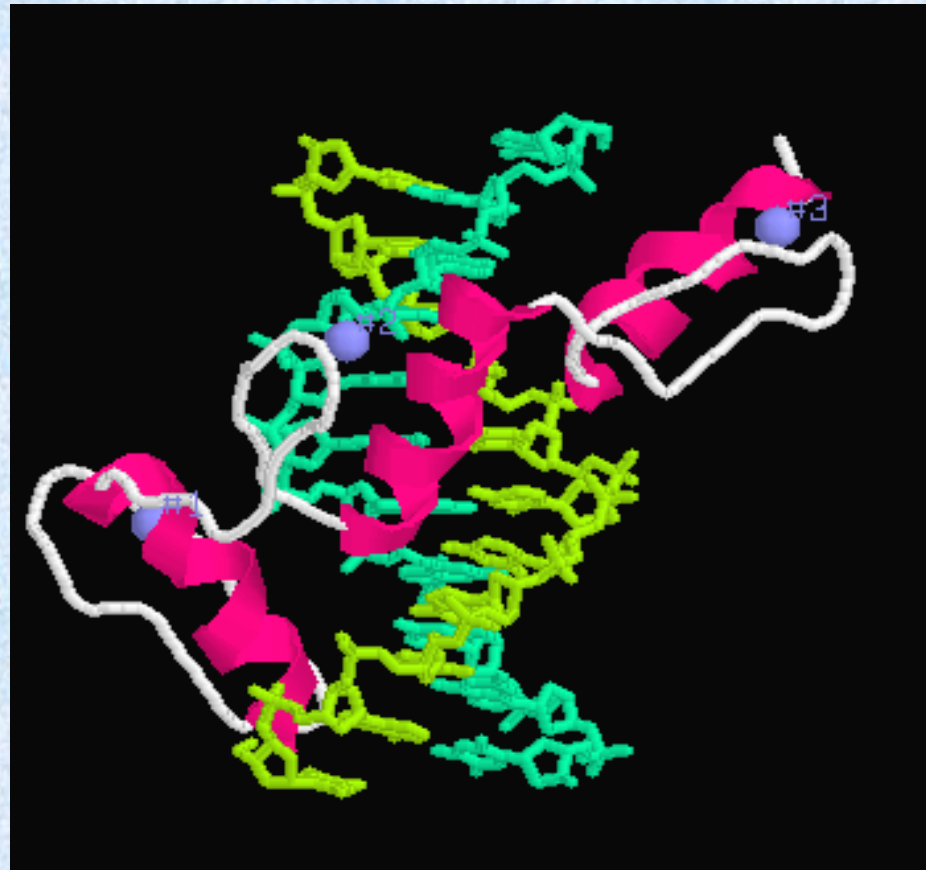
Δεσμεύεται στη μεγάλη αύλακα
και κάμπτεται το DNA κατά ~90

Αλληλεπιδρά με το DNA μόνο
όταν δεσμεύσει
c-AMP

Σύμπλοκο p53 με DNA

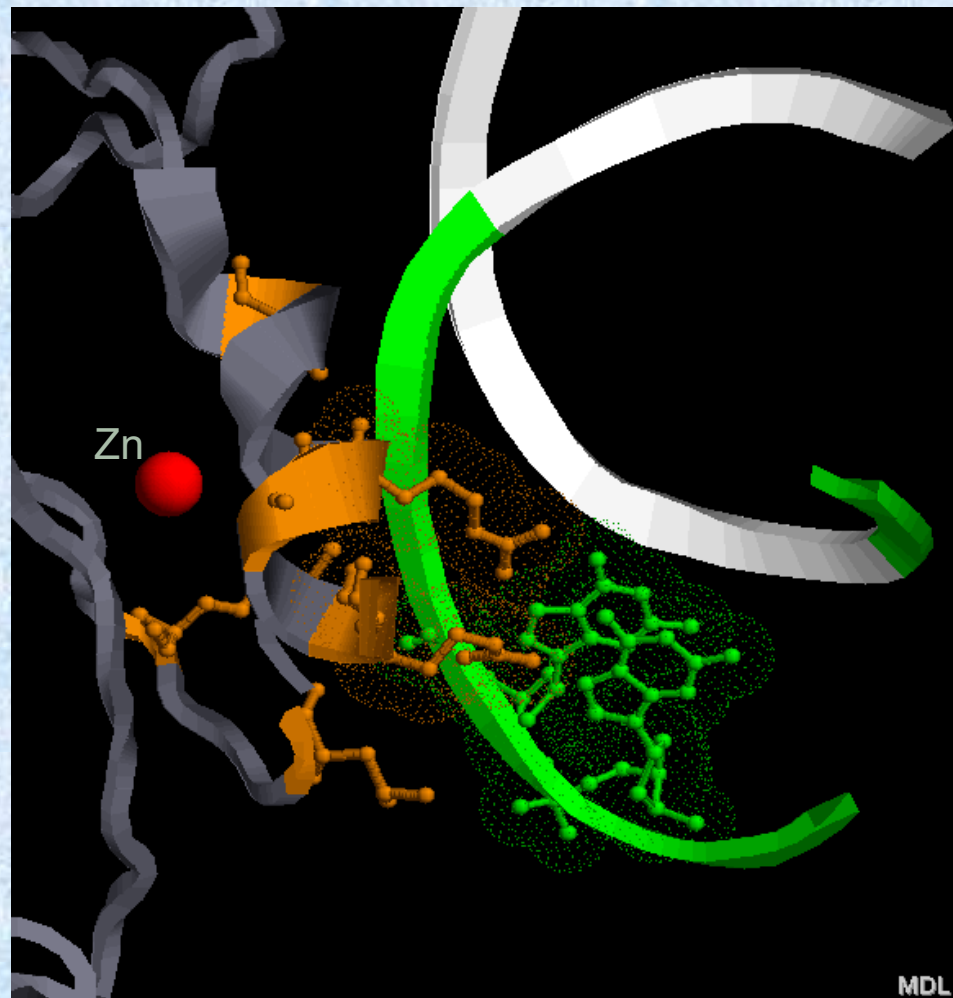


Σύμπλοκο Zn-finger με DNA



Zif 268

Σύμπλοκο Zn-finger με DNA



Υποδοχέας γλυκοκορτικοϊδών

