

Δομική βιολογία

Διάλεξη 6η :

Ταξινόμηση πρωτεϊνικών δομών.
Τάξεις πρωτεϊνικών δομών. Η τάξη α :
Μία εικονογραφημένη περιήγηση στις α-ελικοειδείς δομές.

Εισαγωγή

Εξέταση των πρωτεϊνικών δομών σε επίπεδο τοπολογίας ή μοτίβου αναδίπλωσης οδήγησε στη γενίκευση ότι η ποικιλιότητα σε επίπεδο δομών είναι μικρότερη απ'ότι σε επίπεδο αλληλουχιών. Η ερμηνεία της παρατήρησης βασίζεται (1) στην πίεση της φυσικής επιλογής (η οποία δρα σε επίπεδο δομών και όχι αλληλουχιών), και, (2) στην αρχή της φειδωλότητας [το ίδιο domain (autonomously folding unit) μπορεί να γίνει βάση για για την υλοποίηση πολλών και διαφορετικών λειτουργιών]. Έτσι ξεκίνησε η προσπάθεια ταξινόμησης των πρωτεϊνικών δομών σε ομάδες και υπο-ομάδες.

SCOP

Οι δύο πλέον γνωστές ταξινομικές απόπειρες είναι η SCOP (Structural Classification Of Proteins) και η CATH (Class-Architecture-Homologous superfamily). Οι δύο ταξινομήσεις αυτές (τουλάχιστον στα υψηλά επίπεδα της ιεράρχησης) μοιάζουν μεταξύ τους. Για την περίπτωση της SCOP, στο χαμηλότερο επίπεδο βρίσκονται domains όπως αυτά παρατηρούνται στις πειραματικά προσδιορισμένες δομές. Συνεπώς η SCOP δεν ταξινομεί ολόκληρες δομές πάρα μόνο domains. Στο επόμενο επίπεδο, όσα από αυτά τα domains δείχνουν ομοιότητα σε επίπεδο δομής, λειτουργίας και αλληλουχίας (και συνεπώς είναι μάλλον ομόλογα) συγκροτούν το επόμενο επίπεδο, αυτό των οικογενειών.

SCOP

Όσες οικογένειες έχουν παρόμοιες δομές και λειτουργίες, αλλά διαφέρουν σημαντικά σε επίπεδο αλληλουχιών οργανώνονται σε υπερ-οικογένειες. Η ύπαρξη ομολογίας μεταξύ των μελών μίας υπερ-οικογένειας θεωρείται απλά πιθανή. Όσες από τις υπεροικογένειες έχουν αντίστοιχα μοτίβα αναδίπλωσης τουλάχιστον για ένα σημαντικό τους τμήμα, ομαδοποιούνται σε folds (μοτίβα αναδίπλωσης). Τα μοτίβα αναδίπλωσης δηλώνουν δομική ομοιότητα χωρίς αναφορά σε εξελικτικές σχέσεις. Η δομική ομοιότητα σημαίνει ότι έχουν τα στοιχεία δευτεροταγούς δομής (του πυρήνα της δομής) είναι σε αντίστοιχες θέσεις και με αντίστοιχες τοπολογικές μεταξύ τους συνδέσεις.

SCOP

Στην κορυφή της ιεραρχίας βρίσκονται οι πρωτεϊνικές τάξεις. Για την παρούσα έκδοση της SCOP (1.63) οι τάξεις αυτές είναι επτά :

- α (εξ ολοκλήρου ή κυρίως α)
- β (εξ ολοκλήρου ή κυρίως β)
- α/β (κυρίως παράλληλη β -πτυχωτή μέσω β - α - β μονάδων)
 - $\alpha+\beta$ (κυρίως αντι-παράλληλη β , με διακριτά τμήματα α και β)
 - Multi-domain (μοτίβα αναδίπλωσης αποτελούμενα από domains τα οποία ανήκουν σε διαφορετικές τάξεις).
- Μεμβρανικές
- Μικρές πρωτεΐνες

SCOP

Στατιστικά

18946 καταχωρήσεις στην PDB, ομαδοποιημένες σε 49497 domains, τα οποία ανήκουν σε :

Τάξη	Αναδιπλώσεις	Υπεροικογένειες	Οικογένειες
α	171	286	457
β	119	234	418
a/b	117	192	501
a+b	224	330	532
MD	39	39	50
Membr	34	64	71
Small	61	87	135
Total	765	1232	2164

Αυτόματες ταξινομήσεις

FSSP

Αυτές στηρίζονται σε μεθόδους στοίχισης δομών (στις τρεις διαστάσεις) για να ομαδοποιήσουν τις γνωστές δομές. Η πλέον γνωστή είναι η FSSP (Fold classification based on Structure-Structure alignment of Proteins !), η οποία δημιουργείται ως εξής :

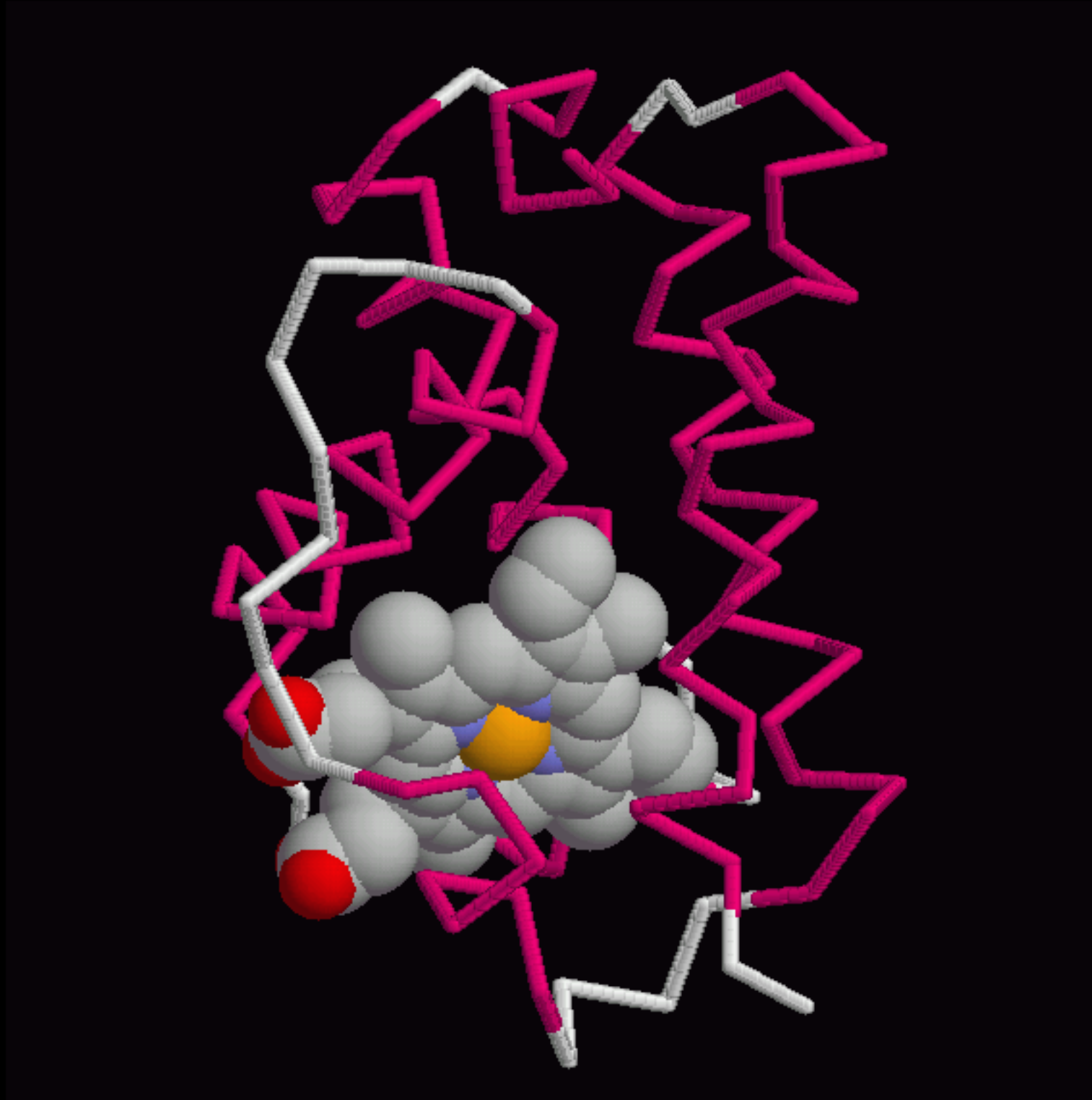
Για όλες τις πρωτεϊνικές αλυσίδες με περισσότερα από 30 κατάλοιπα και περισσότερο από 25% αμινοξική ταυτότητα (σε επίπεδο αλληλουχίας) επιλέγεται μία αντιπροσωπευτική δομή. Η δομή αυτή χρησιμοποιείται σε μια έρευνα της PDB για ταυτοποίηση όμοιων (στο χώρο) δομών (δηλ. με μικρό rmsd, αλγόριθμος DALI), οι οποίες και ομαδοποιούνται.

Τάξη α

Με 171 μοτίβα αναδίπλωσης στην SCOP, η τάξη αυτή είναι η δεύτερη πολυπληθέστερη μετά την α+β. Στην τάξη αυτή ανήκουν μερικές από τις πλέον καλά χαρακτηρισμένες (και όχι μόνο δομικά) οικογένειες, π.χ. σφαιρίνες και κυτοχρώματα c. Μερικά χαρακτηριστικά μοτίβα αναδίπλωσης α-ελικοειδών δομών είναι :

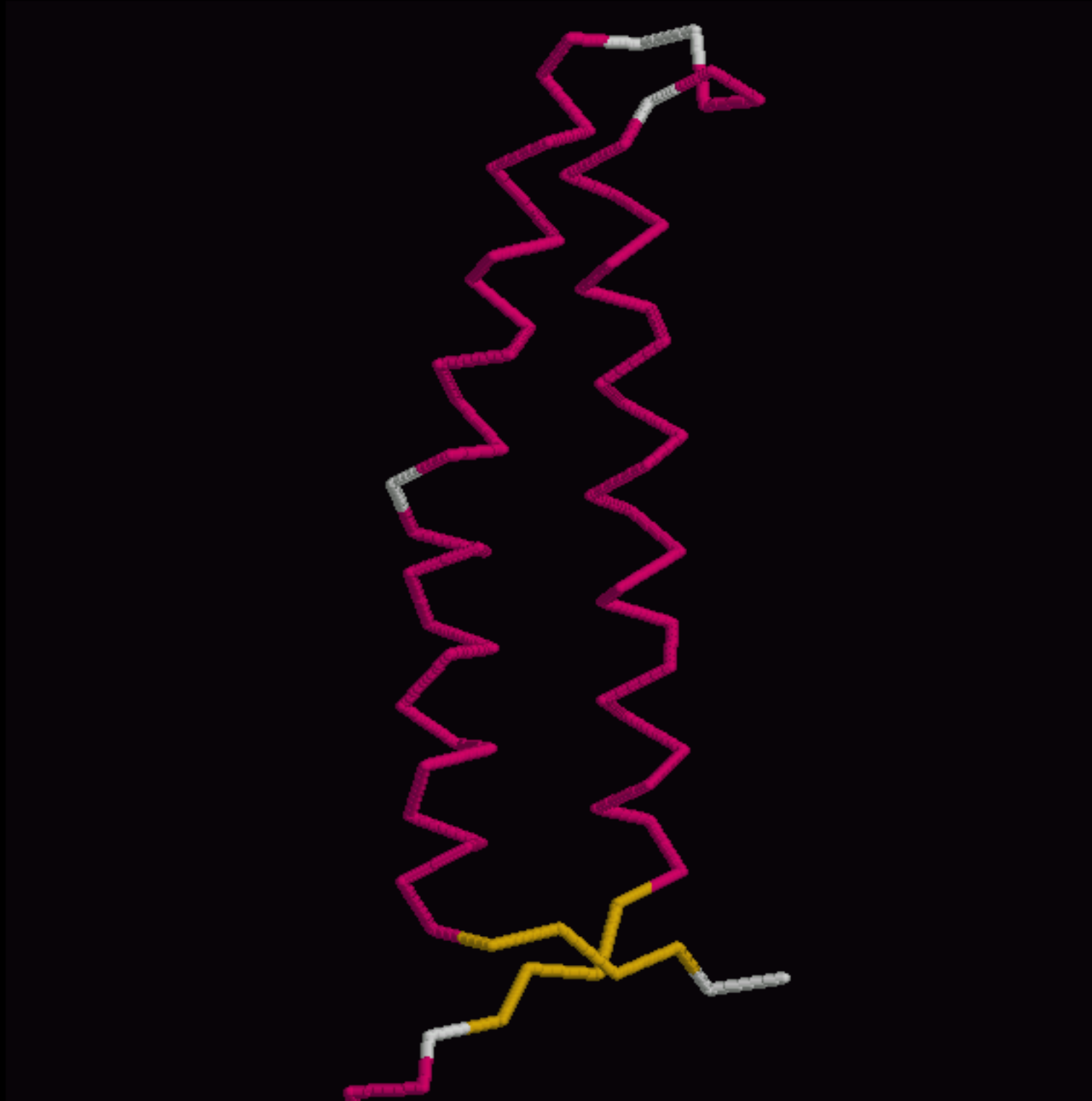
Σφαιρίνες

6 έλικες σε διπλωμένο φύλλο, ανοικτό



α-φουρκέτα

2 έλικες, αντιπαράλληλες

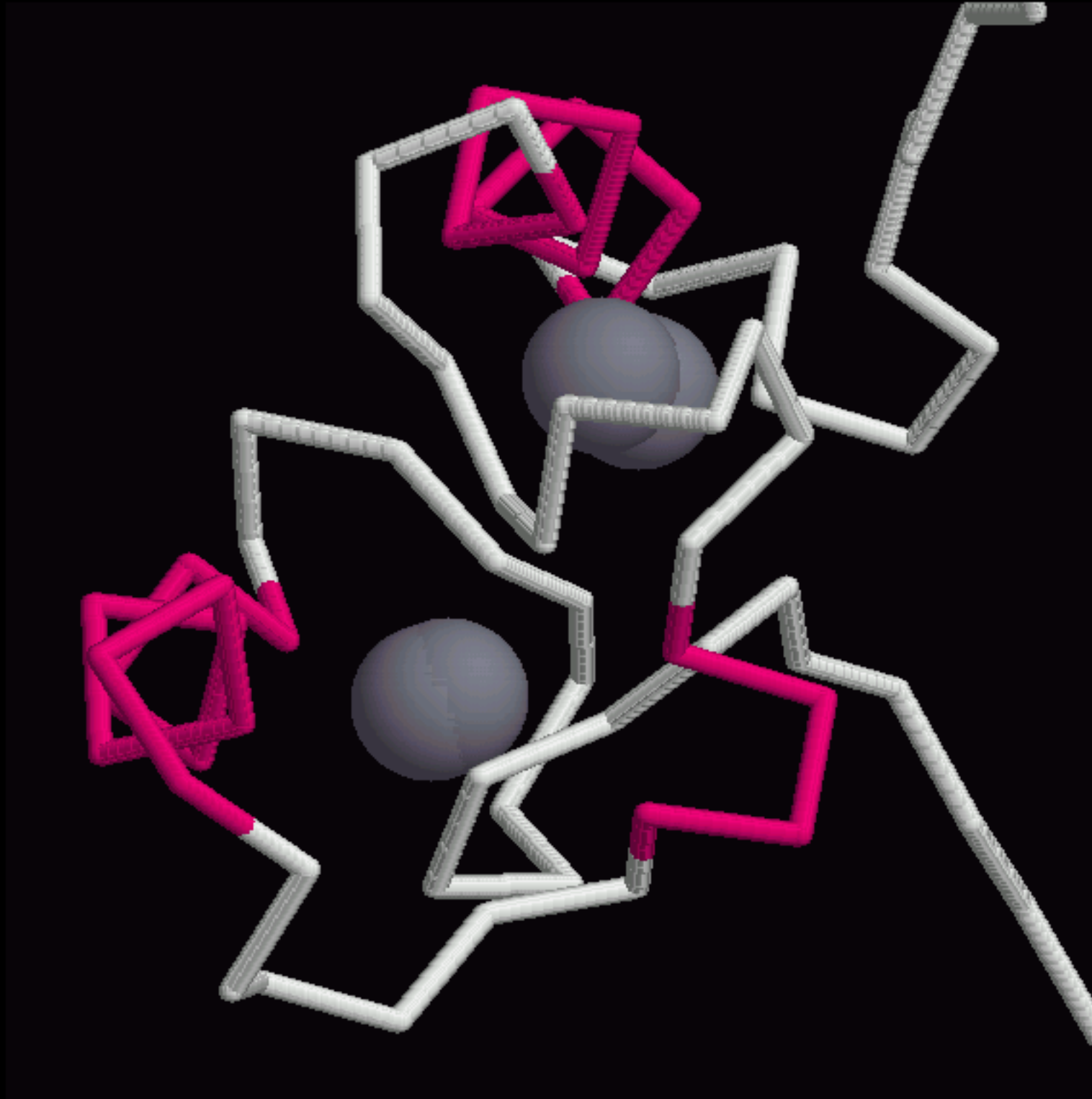


S10_1

S10_2

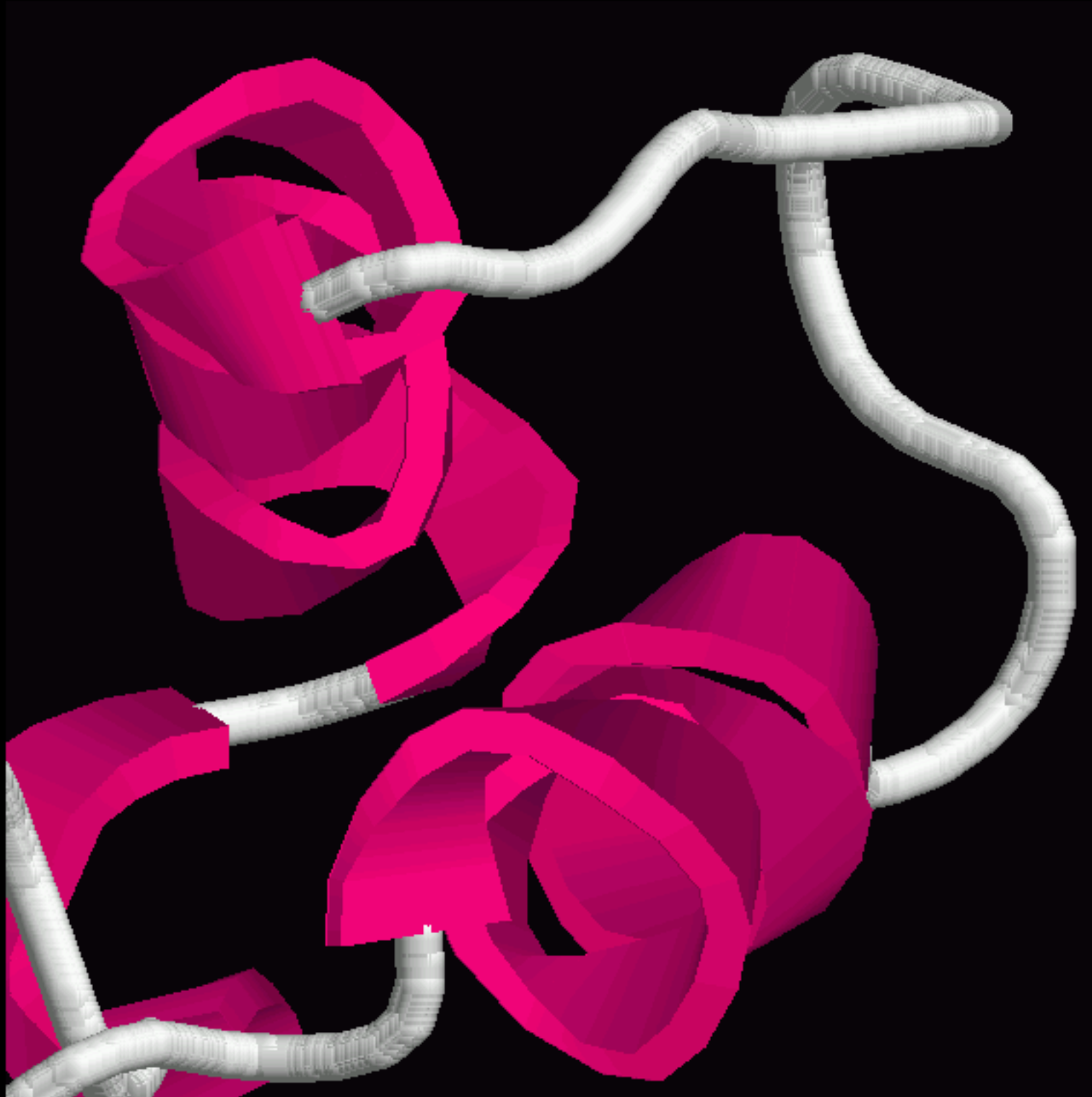
Θηλιά-έλικα

Θηλιά-έλικα για πρόσδεση Ca^{++} (όχι EF-hand)



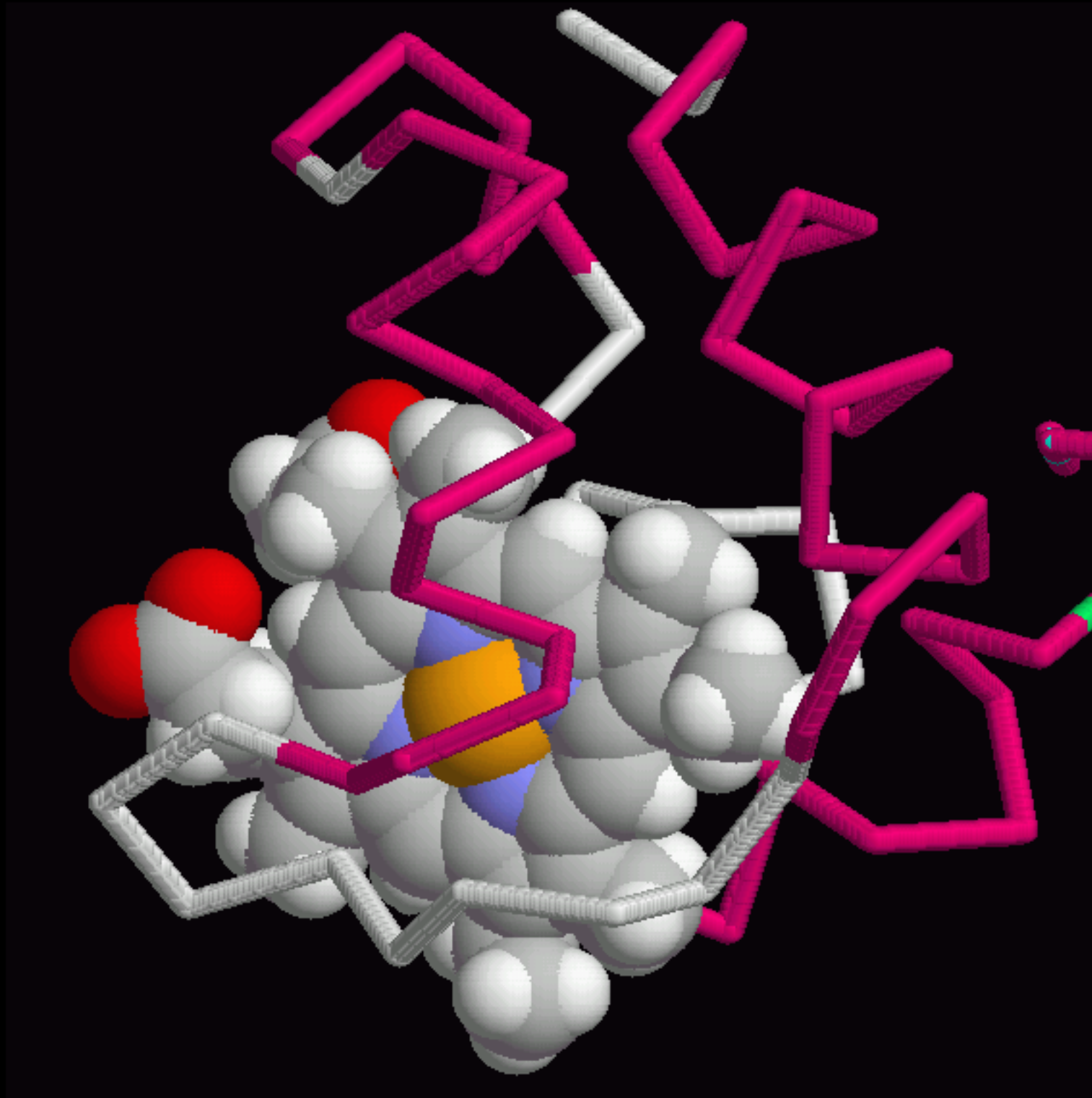
Έλικά-θηλιά-έλικά

Παράλληλες έλικες μέσω μακριάς θηλιάς



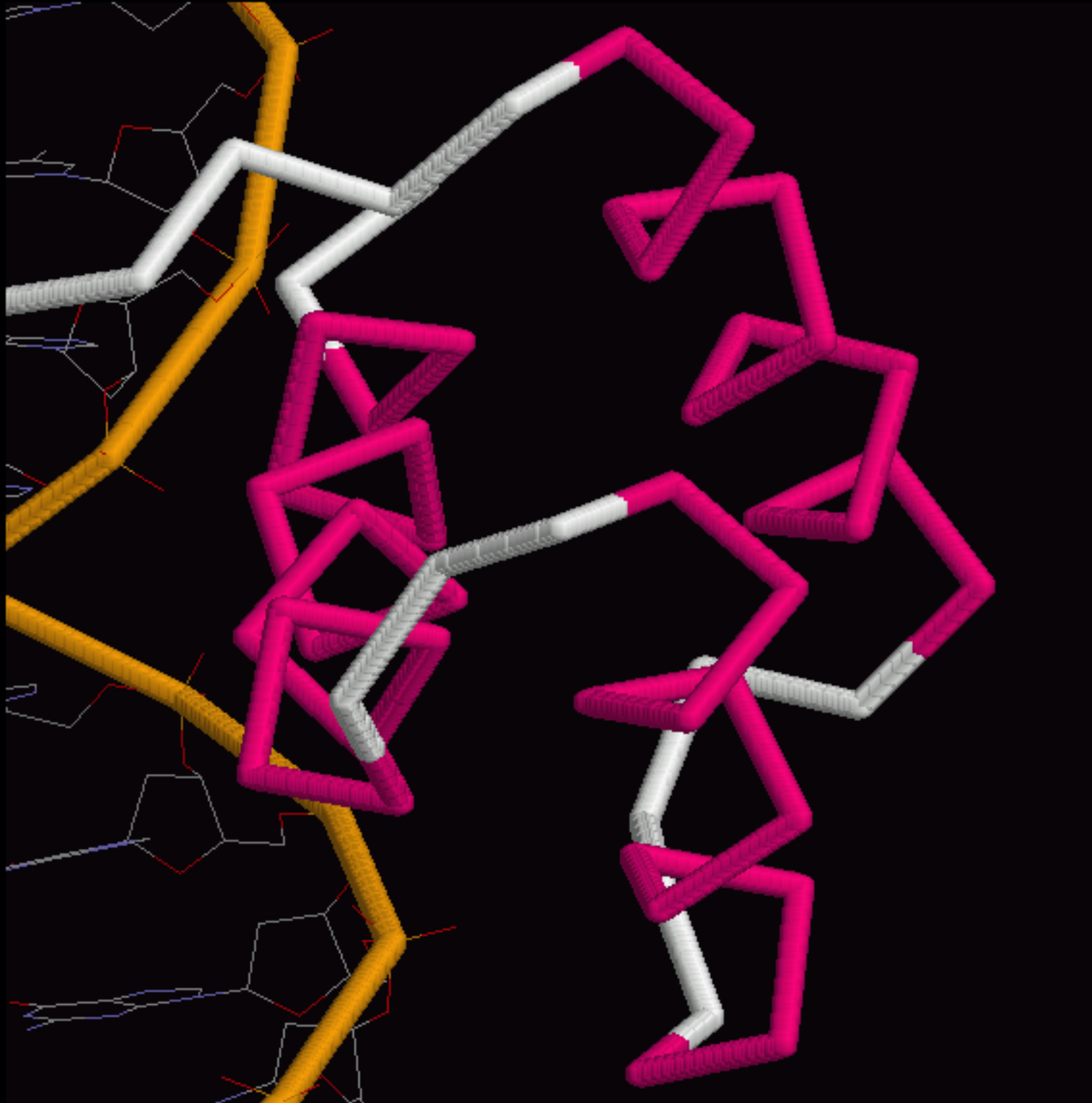
Κυτόχρωμα c

3 έλικες σε διπλωμένο φύλλο, ανοικτό



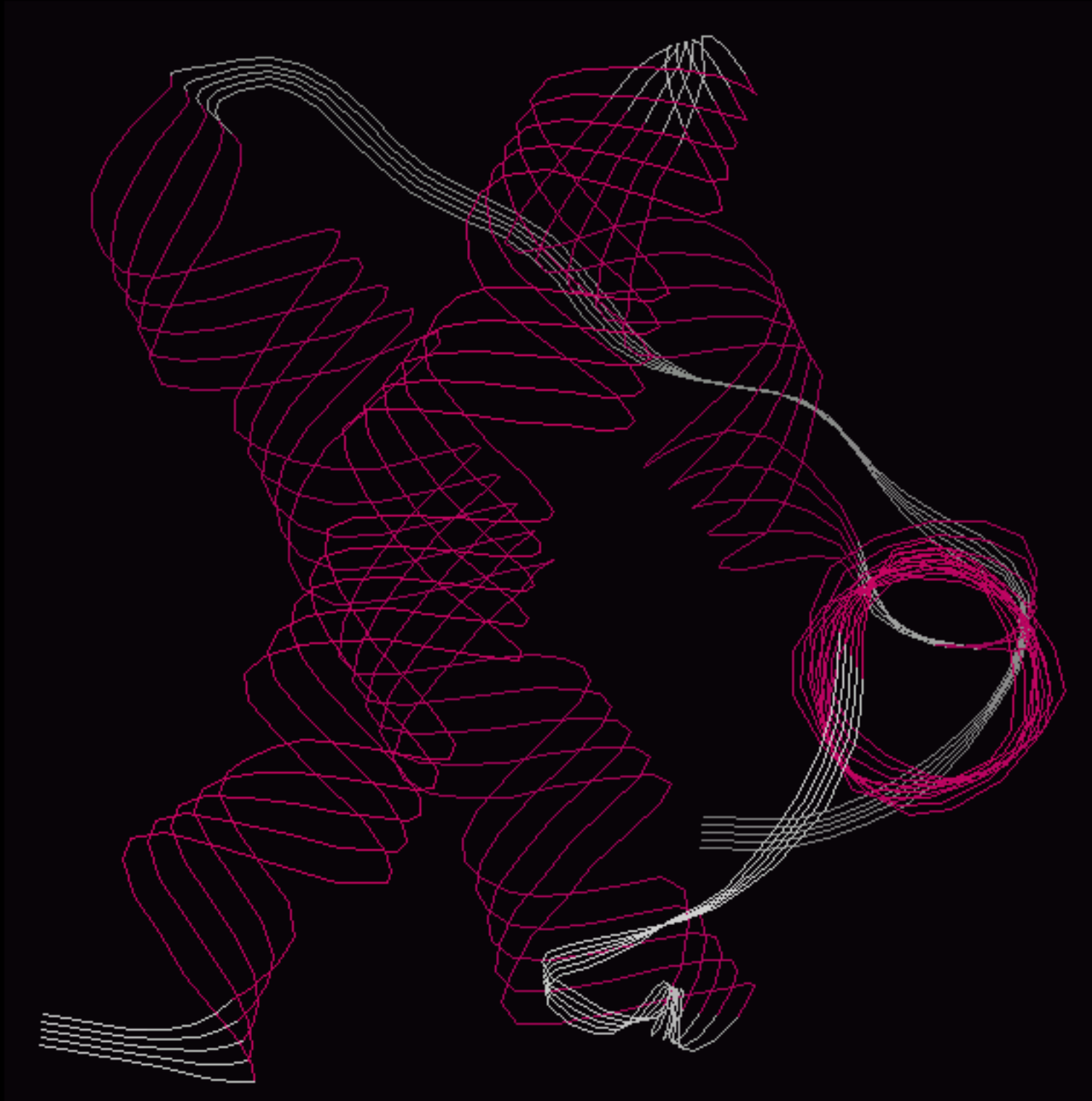
3-εΛΙΚΟΕΙΔΕΣ ΔΕΜΑΤΙΟ

3 έλικες, δεμάτιο (Engrailed homeodomain)



3-ελικοειδές δεμάτιο

3 έλικες, ίδια τοπολογία, διαφορετικό πακετάρισμα



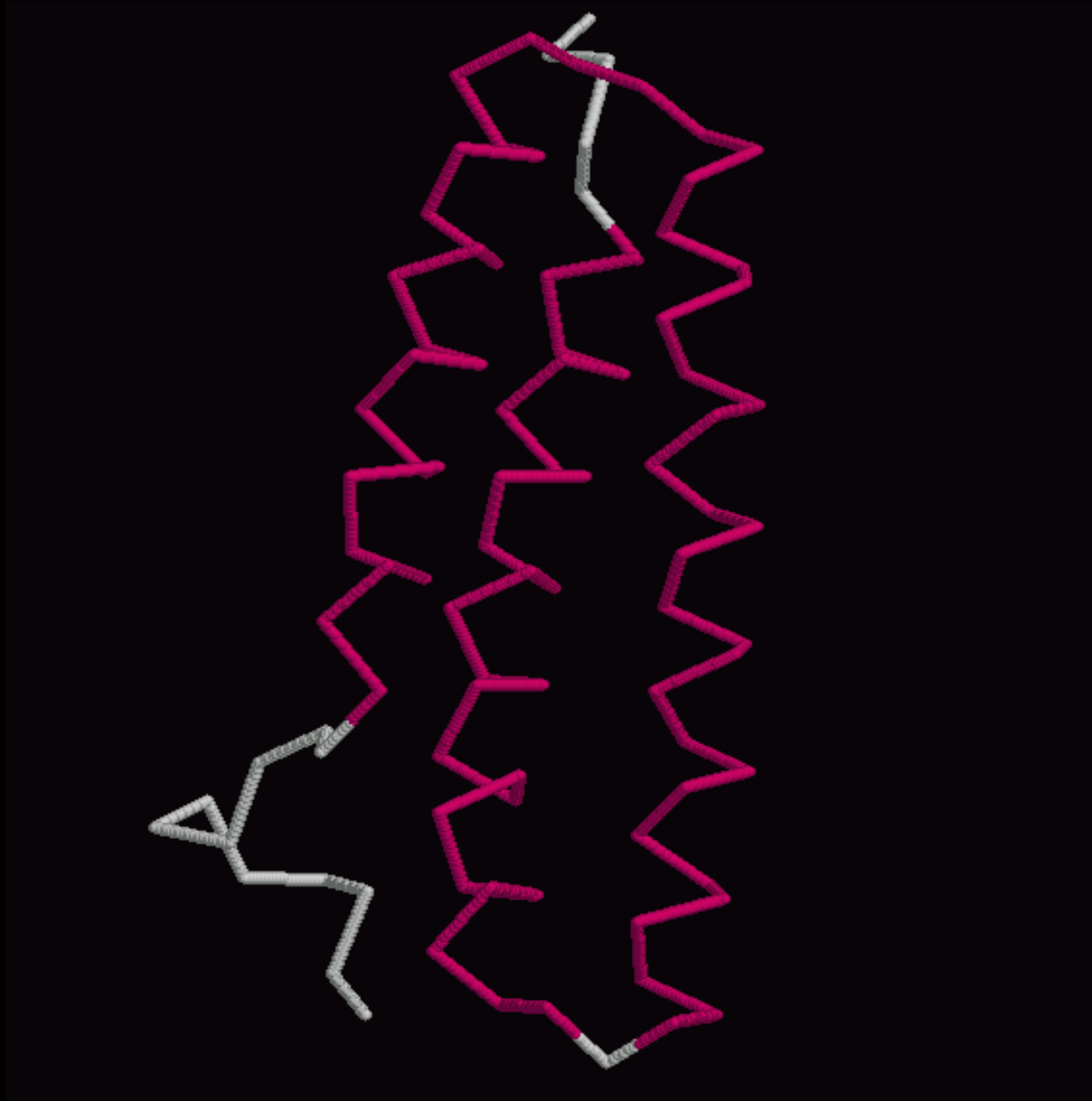
3-ελικοειδές δεμάτιο

3 έλικες, δεμάτιο (RuvA)



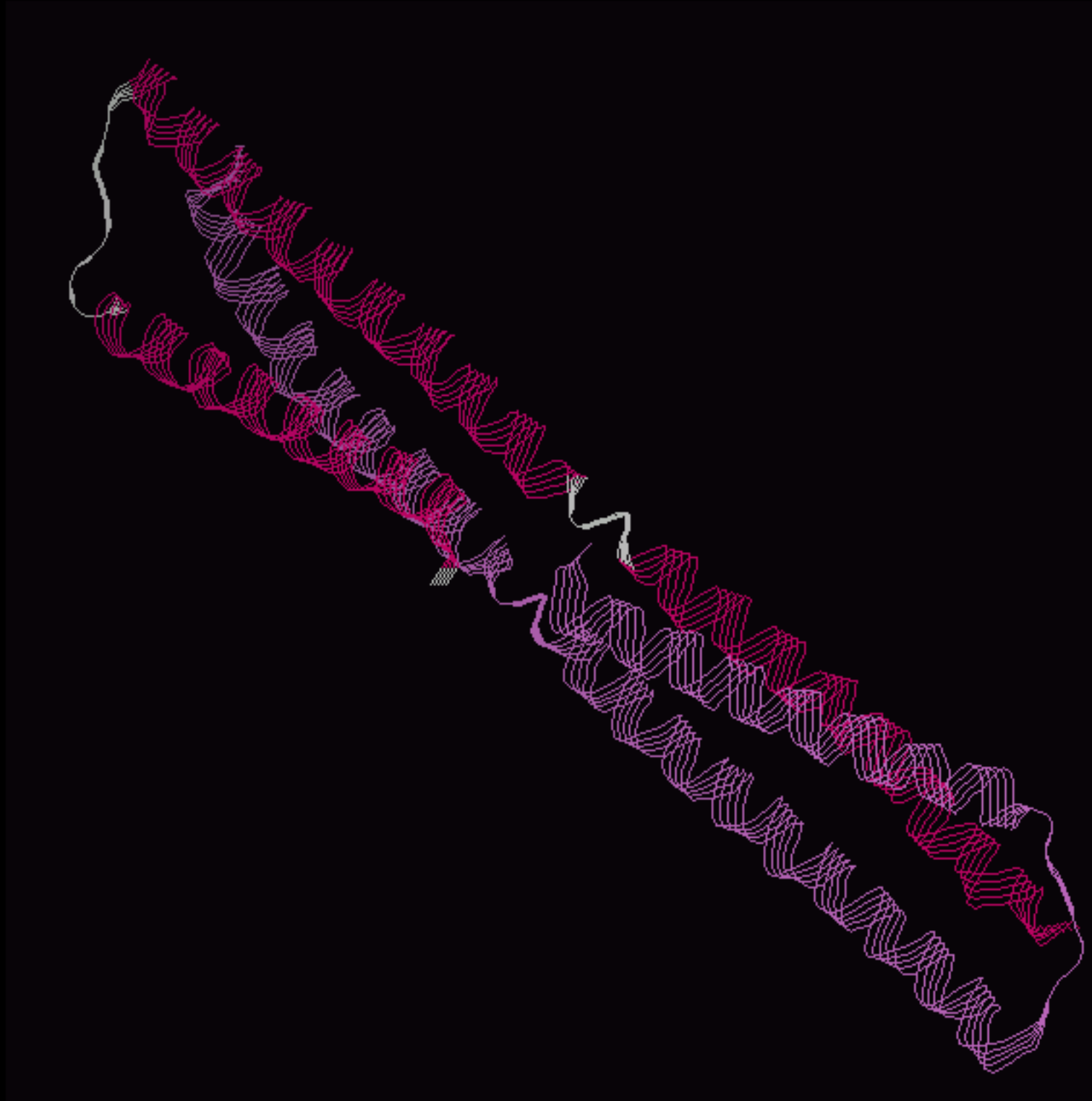
3-εΛΙΚΟΕΙΔΕΣ ΔΕΜΑΤΙΟ

3-έλικες, δεμάτιο (Chaperone protein HSCB)



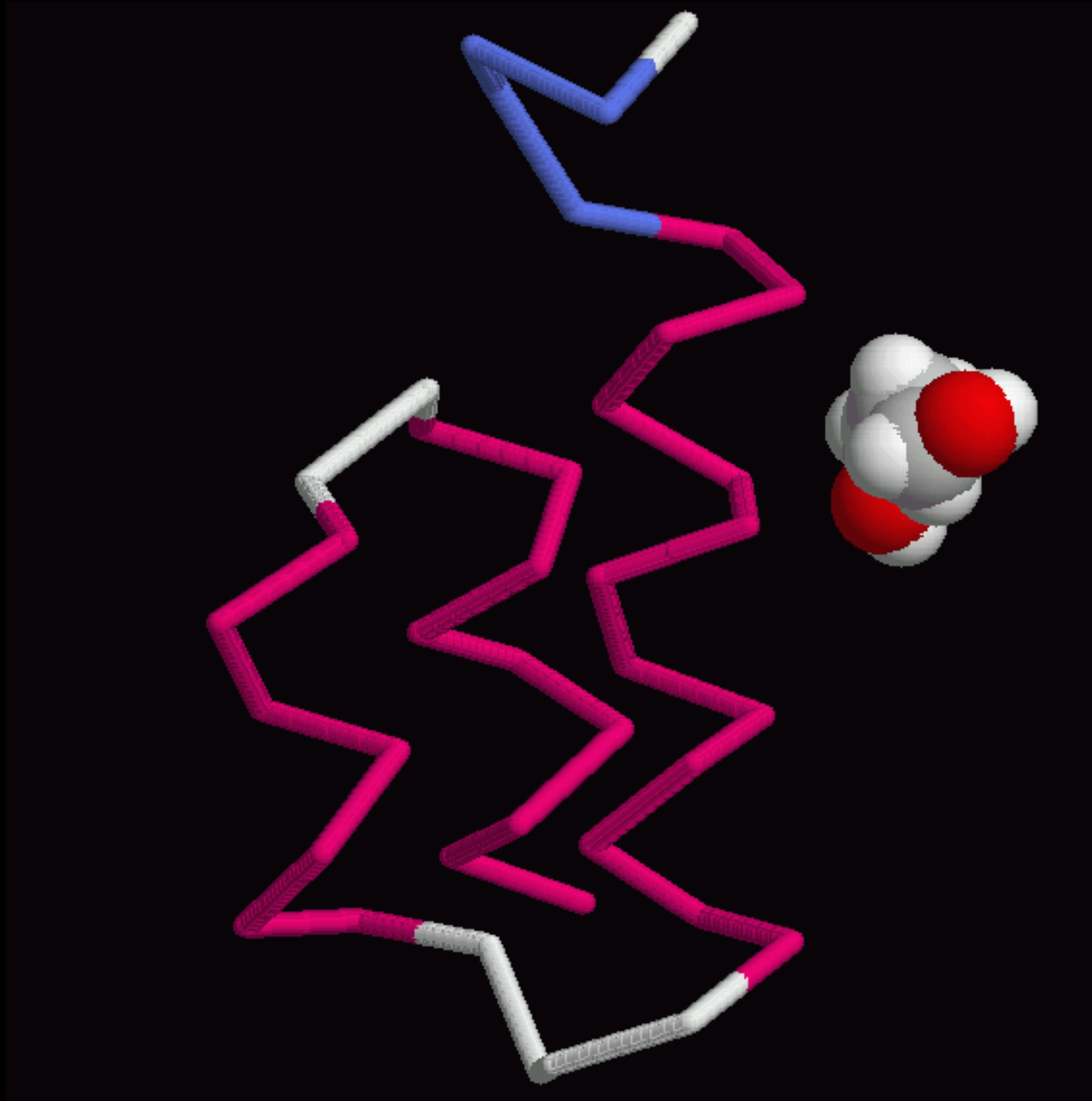
Spectrin

3-ελικοειδές δεμάτιο



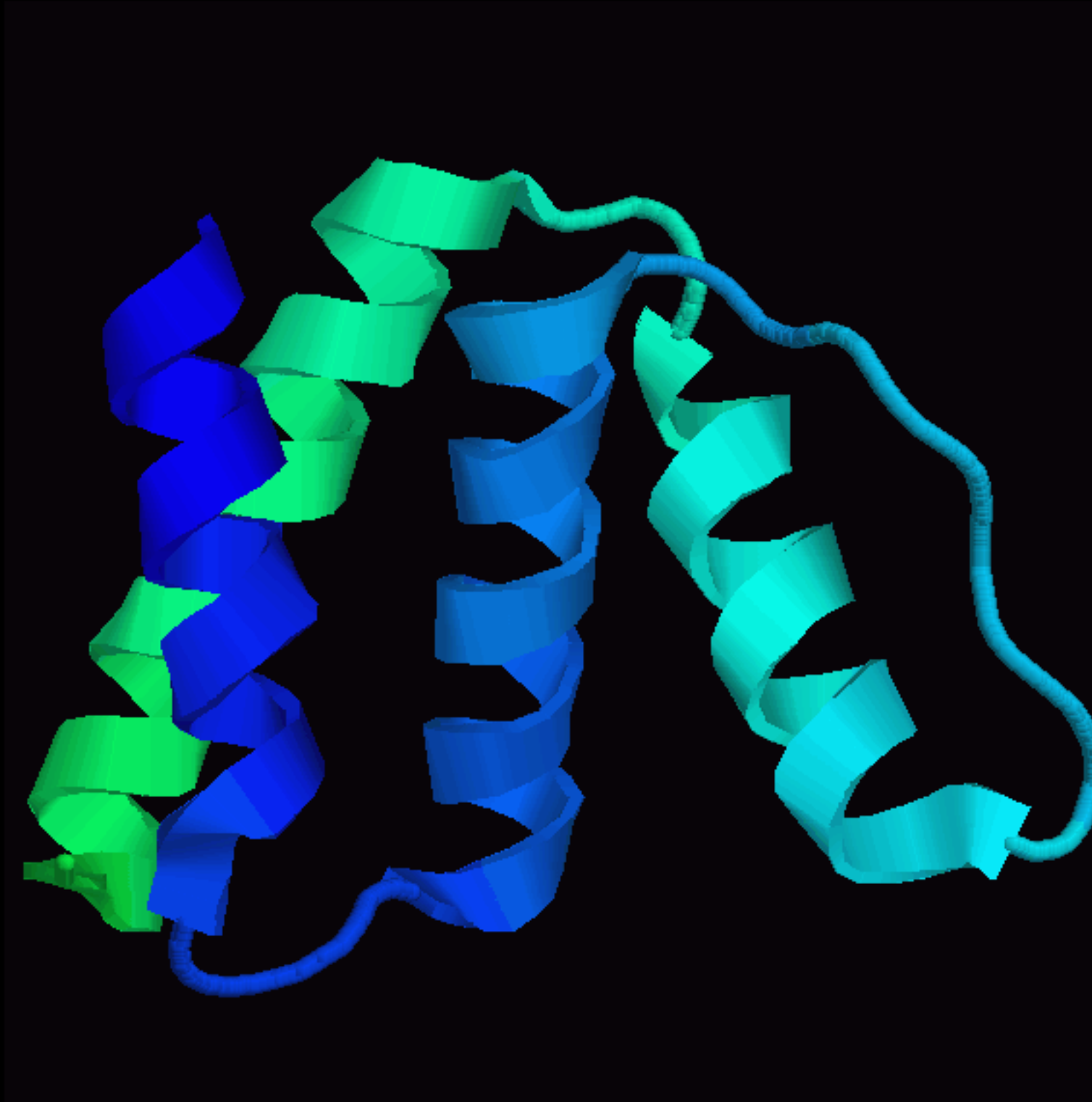
Φερομόνες πρωτοζώων

3-ελικοειδές δεμάτιο



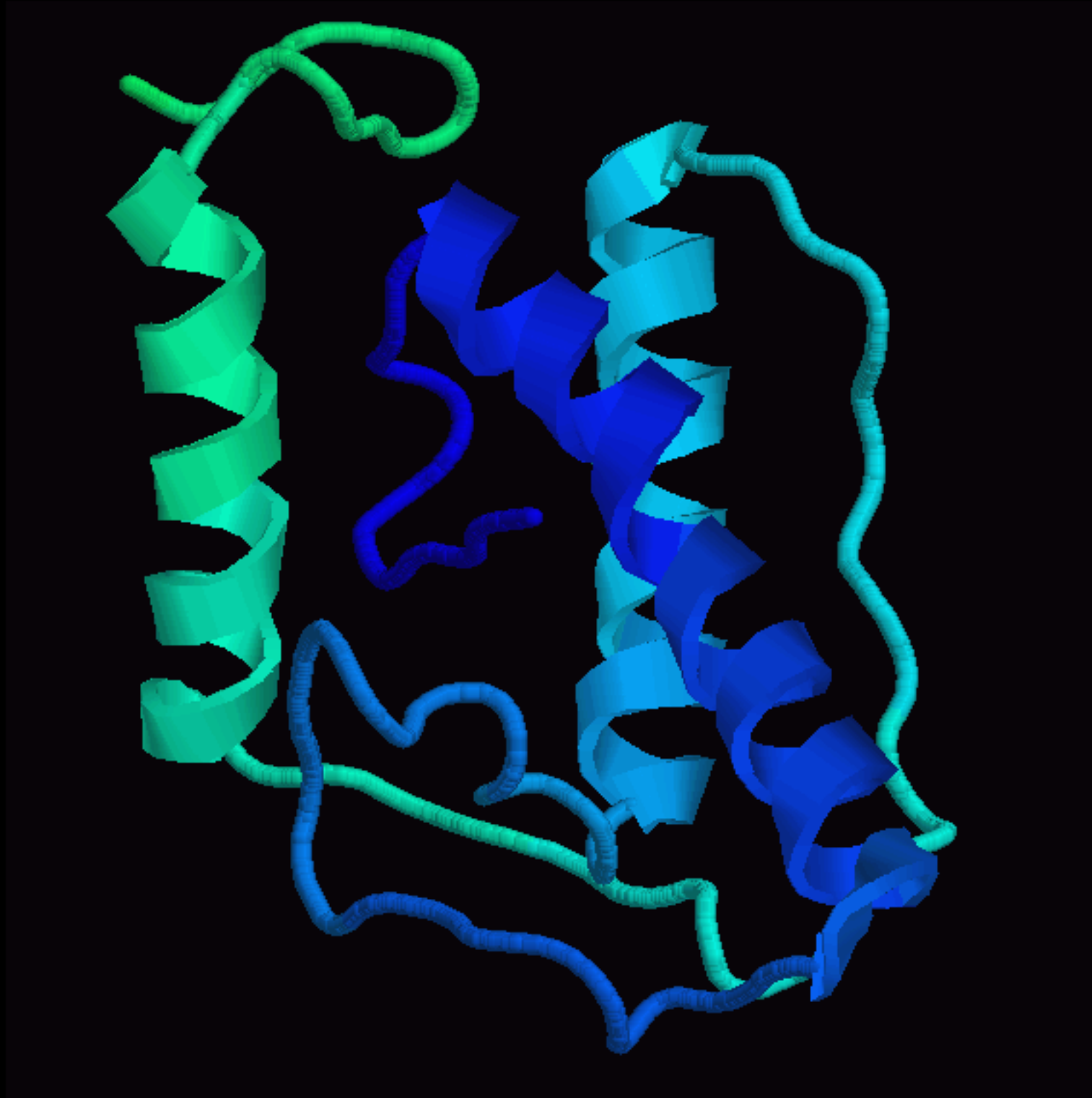
Ακετυλο-CoA προσδενόμενη

3-ελικοειδές δεμάτιο



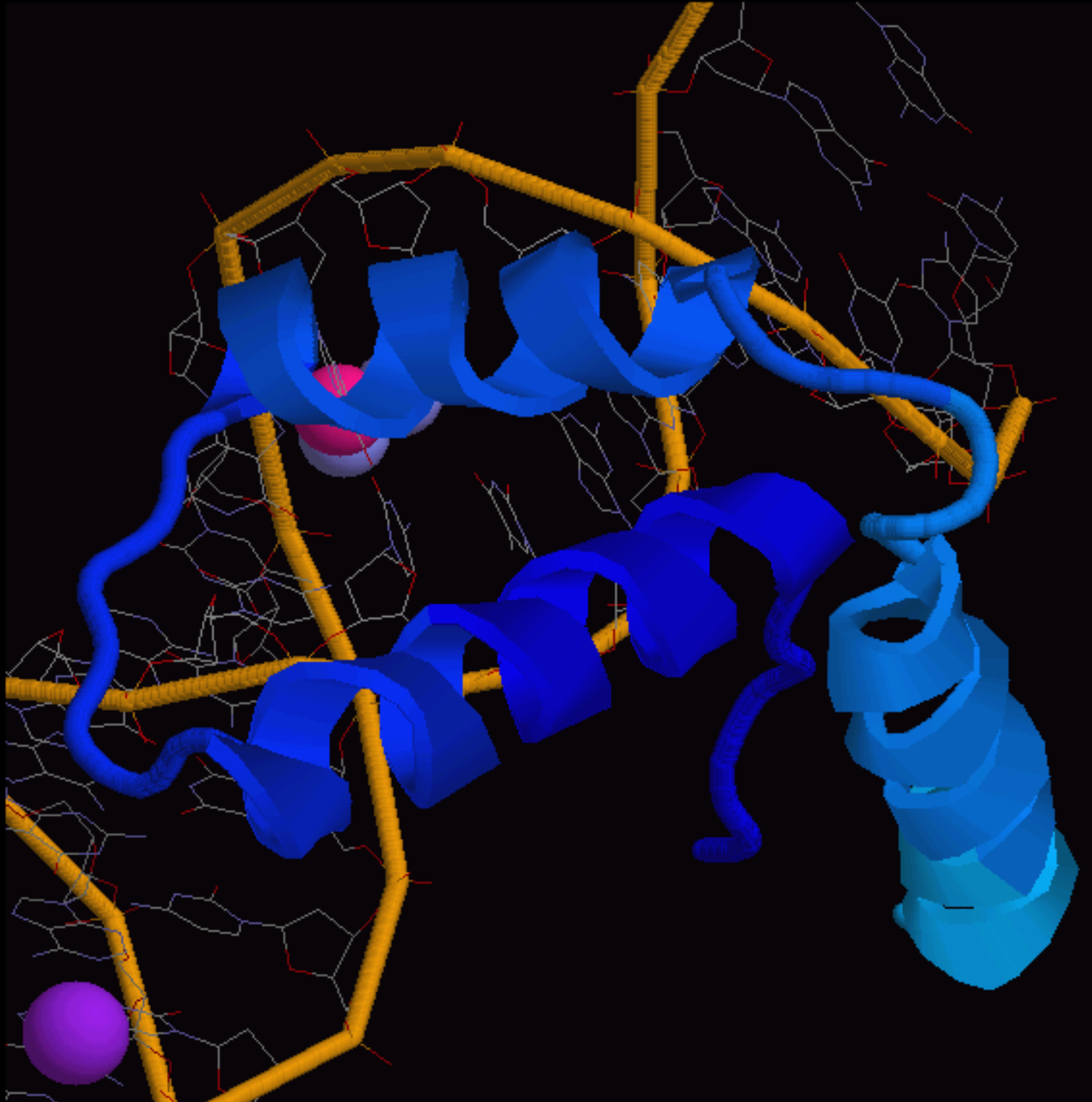
Τ4 ενδοουκλεάση V

3 έλικες, ακανόνιστη διάταξη



HMG box

3 έλικες, ακανόνιστη διάταξη



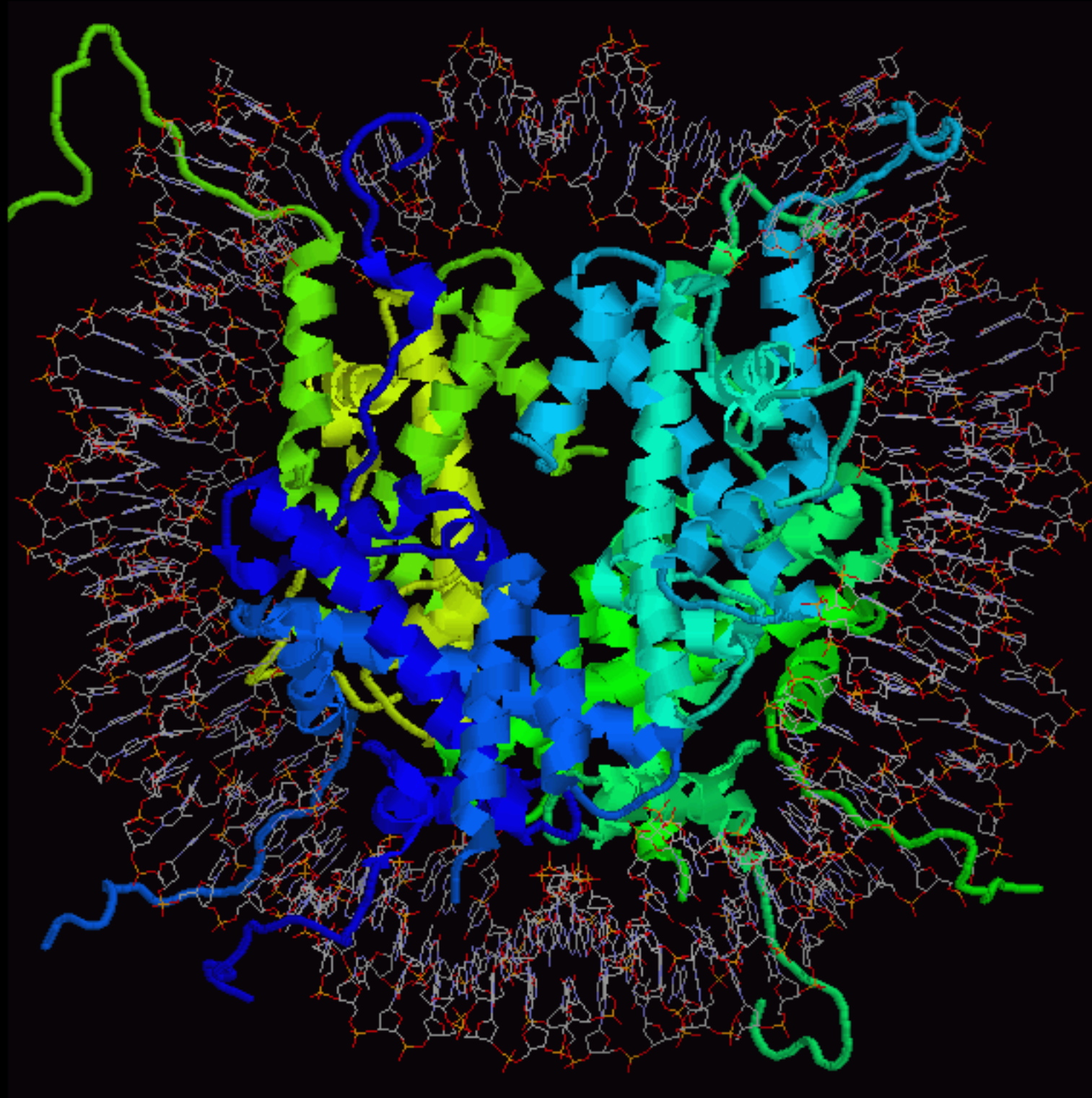
Histone fold

3 έλικες, περιστοιχιζόμενη κεντρική μεσαία έλικα



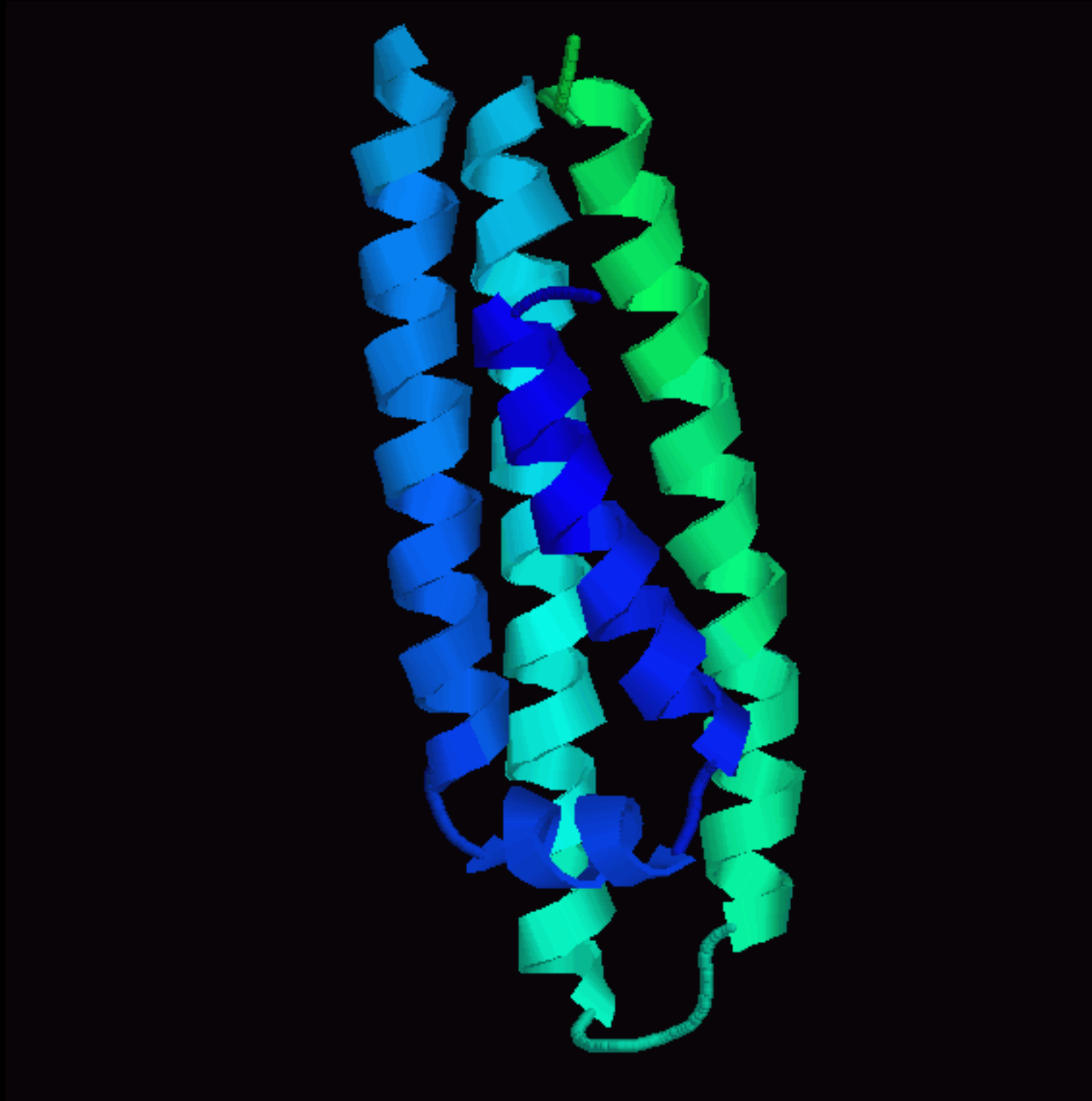
Histone fold

3 έλικες, περιτοιχιζόμενη κεντρική μεσαία έλικα



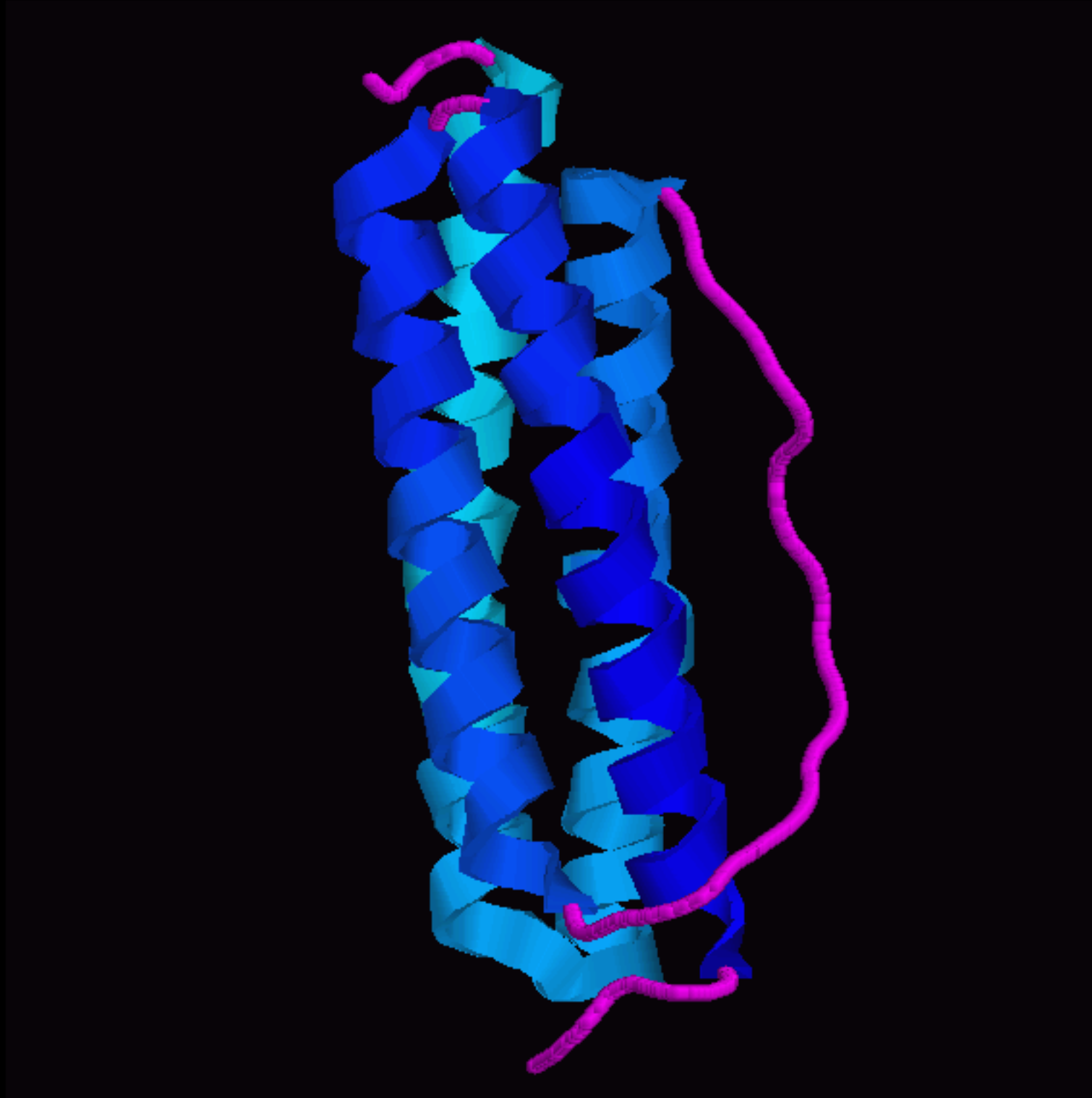
4-ελικοειδές δεμάτιο

4 έλικες, δεμάτιο



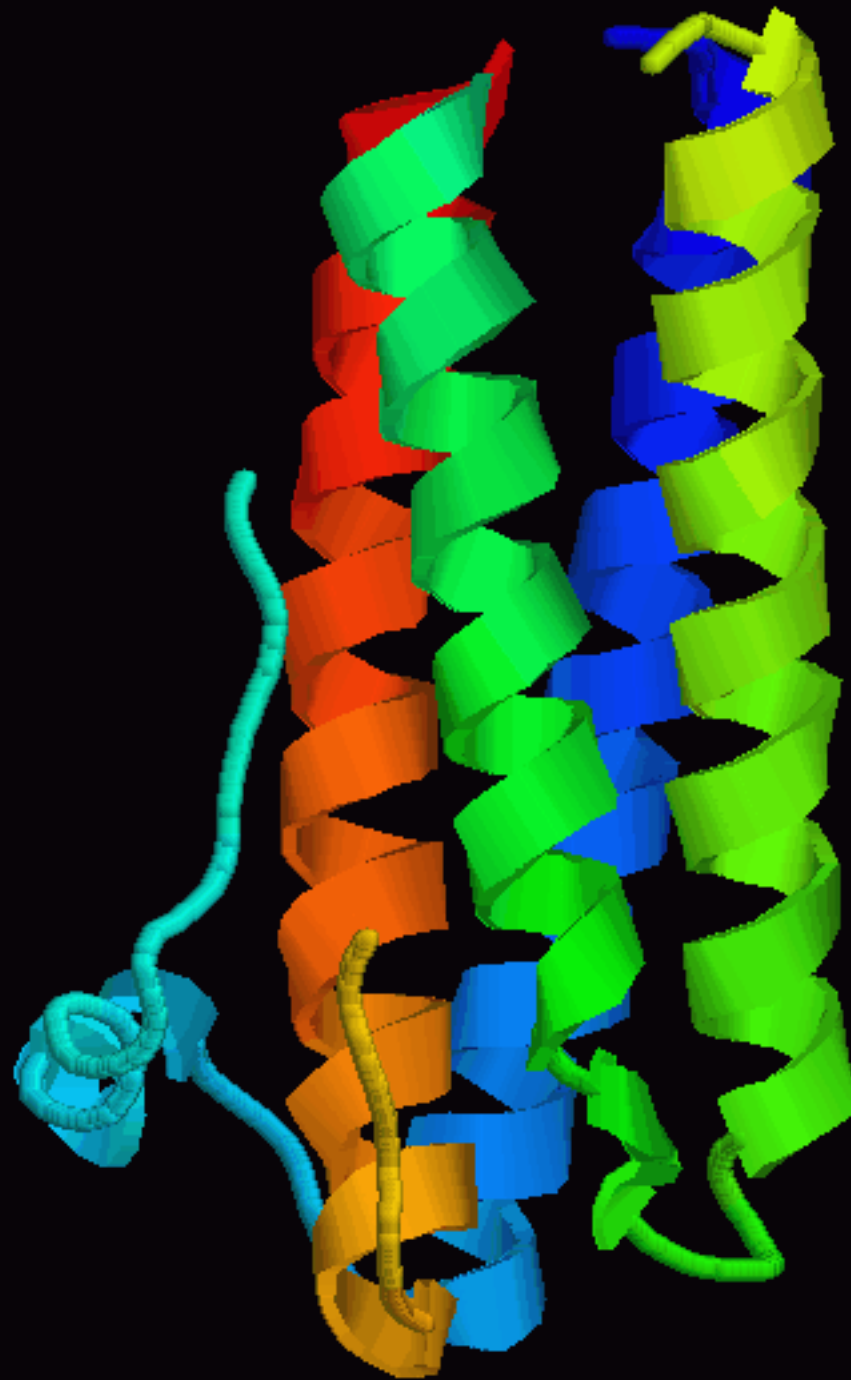
Ferritin-like

4 έλικες, δεμάτιο, 1 διασταύρωση



Cytokines (4-helical)

4 έλικες, δεμάτιο, 2 διασταυρώσεις



Acyl-carrier

4 έλικες, δεμάτιο, έλικα 3 κοντή (thiobutyric acid)



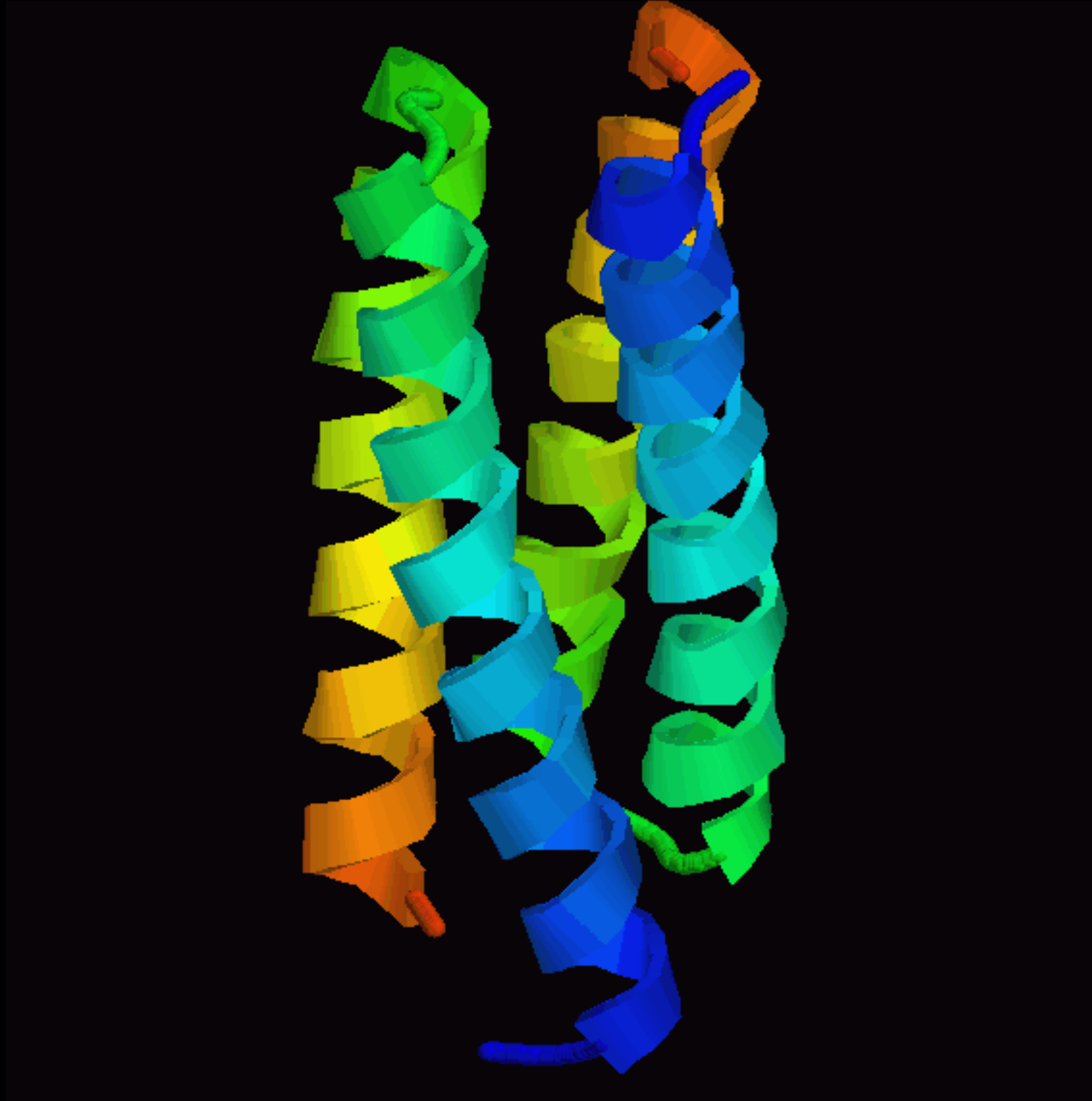
Bromodomain (GCN5)

4 έλικες, δεμάτιο, ακετυλιωμένο πεπτίδιο από H4



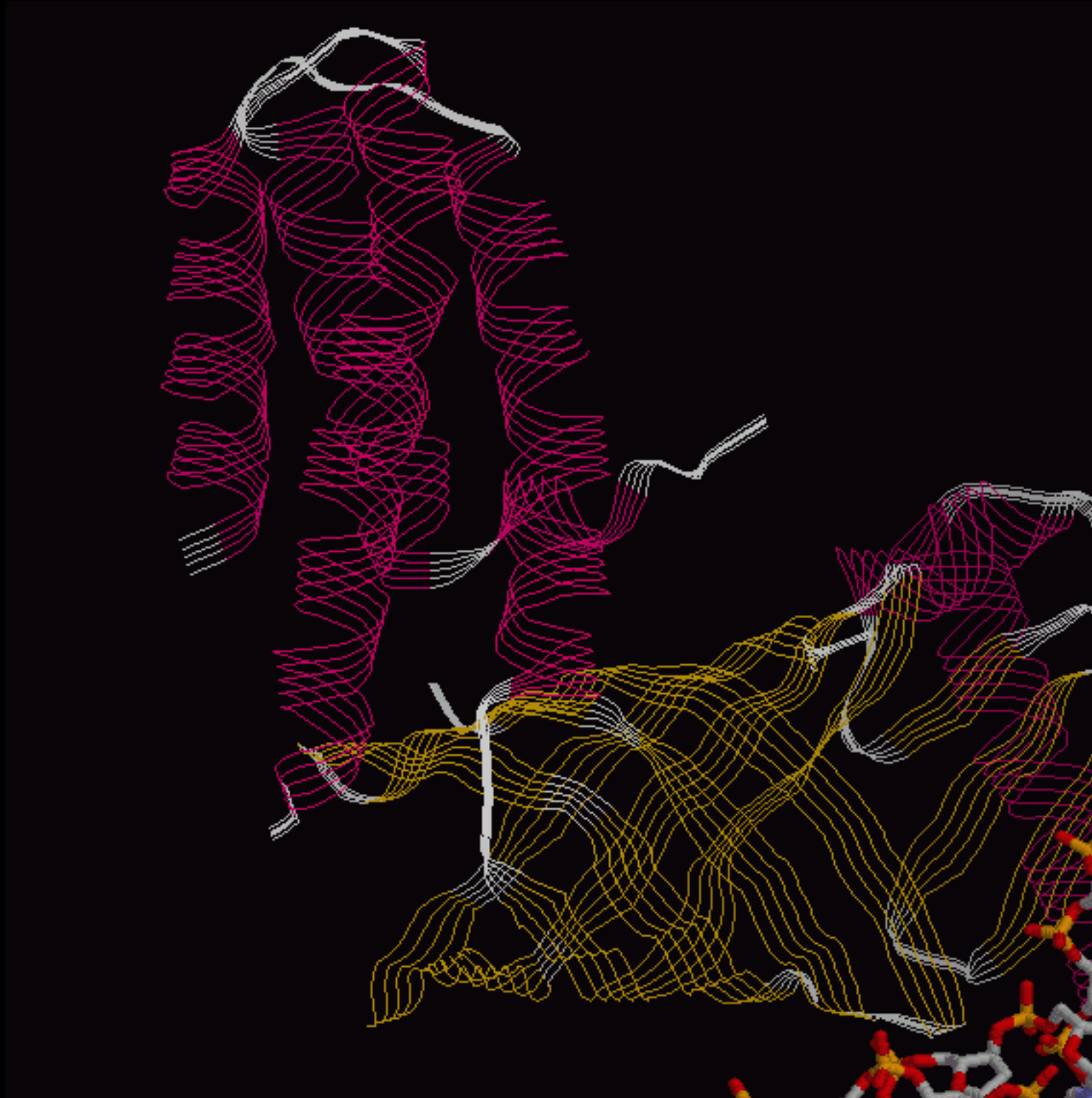
Rop (RNA-binding)

4 έλικες, δεμάτιο, αντιπαράλληλο



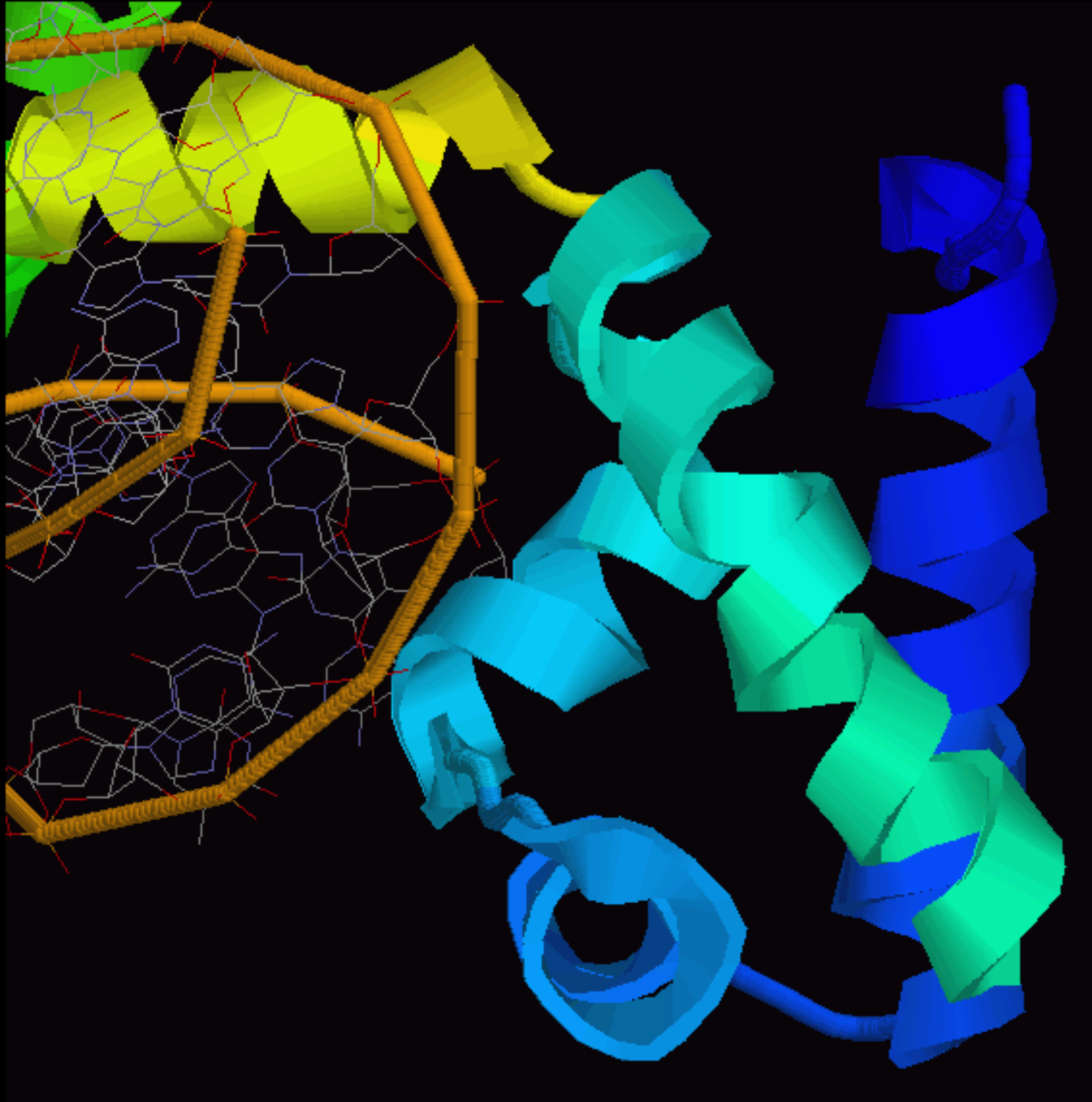
TFIIA, N-τελικό

4 έλικες, δεμάτιο



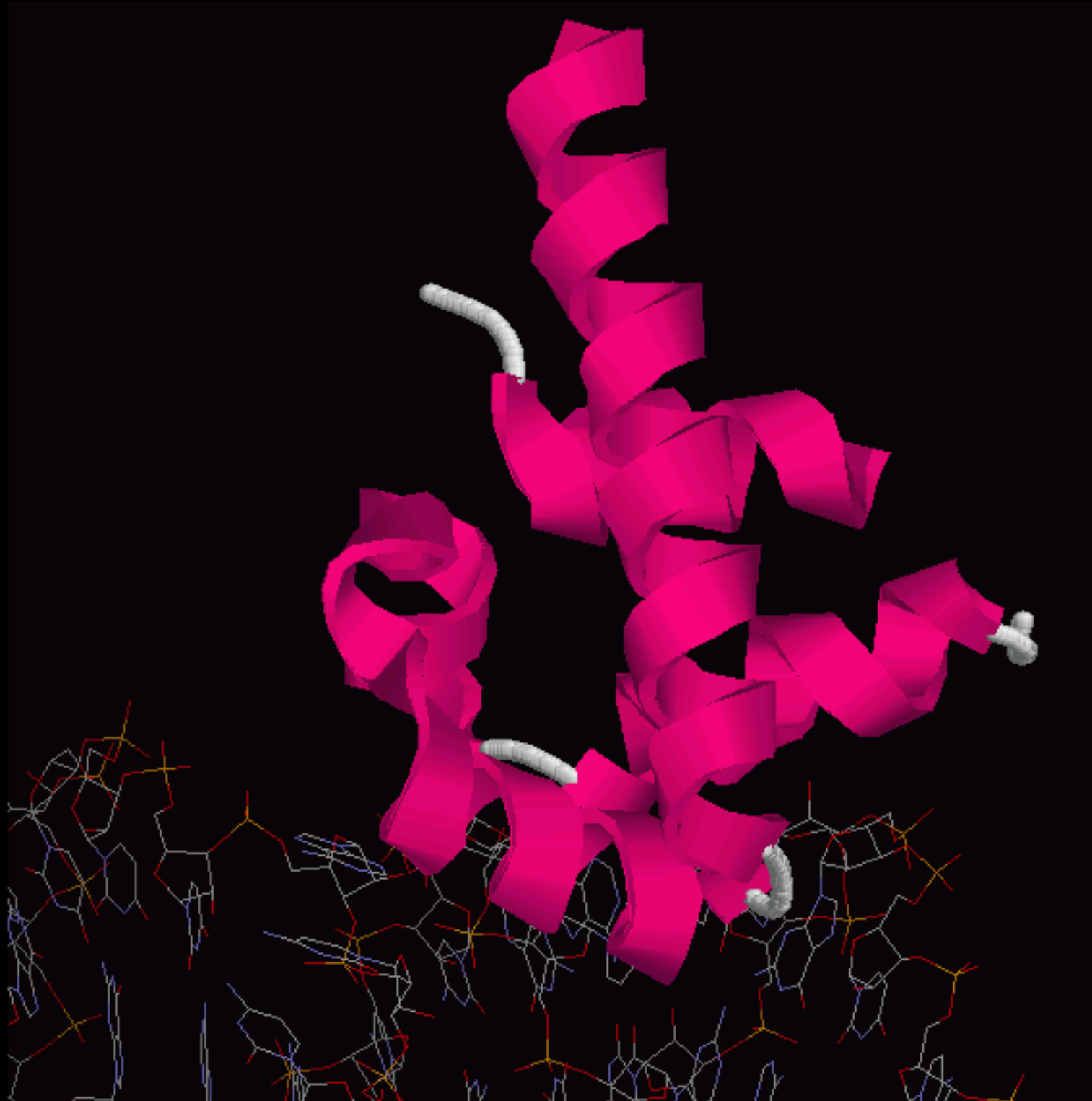
λ-repressor-like

4 έλικες, διπλωμένο φύλλο (TF OCT-1)



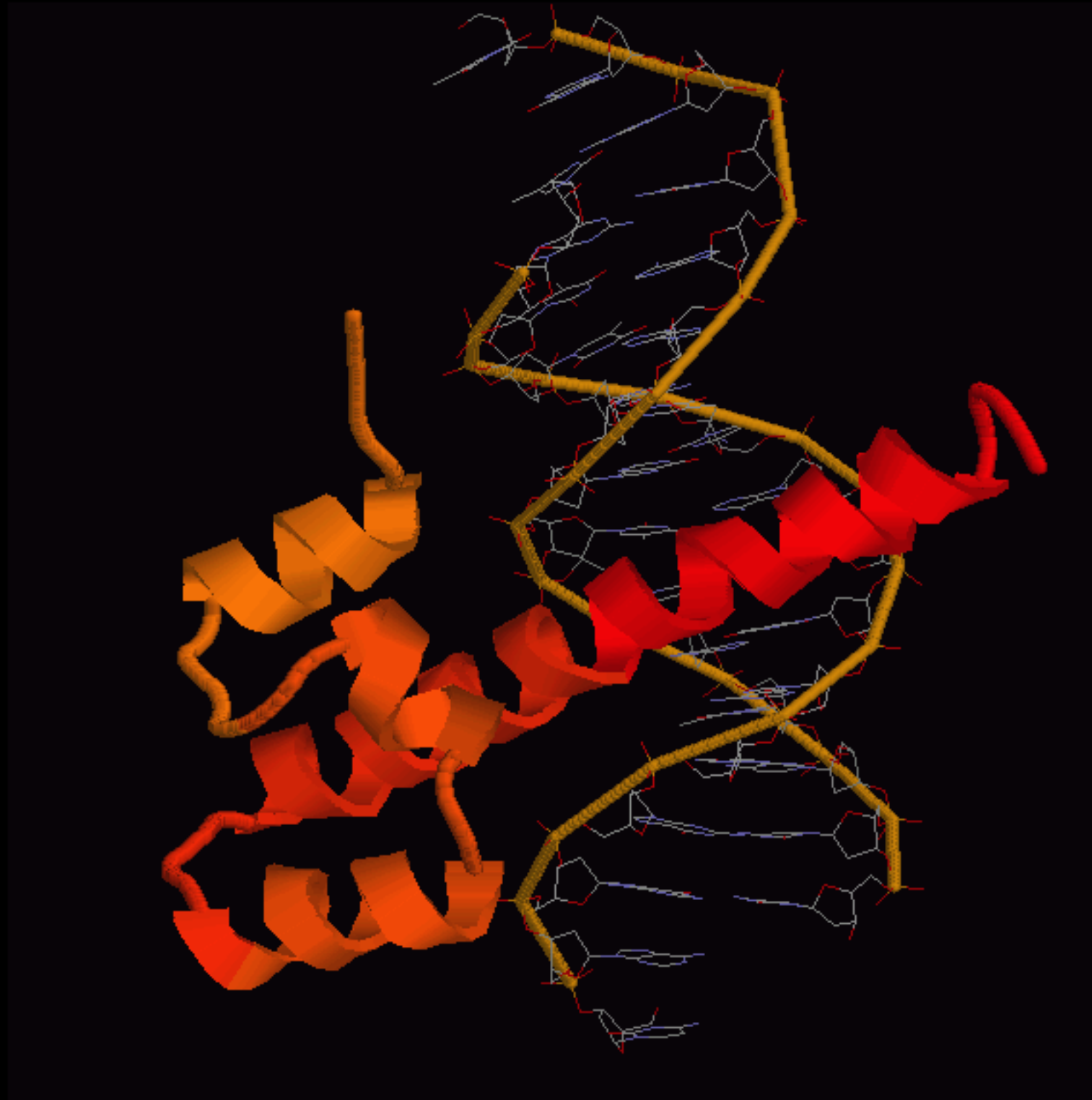
RNA-binding

4 έλικες, ορθογώνια διάταξη, signal peptide-binding



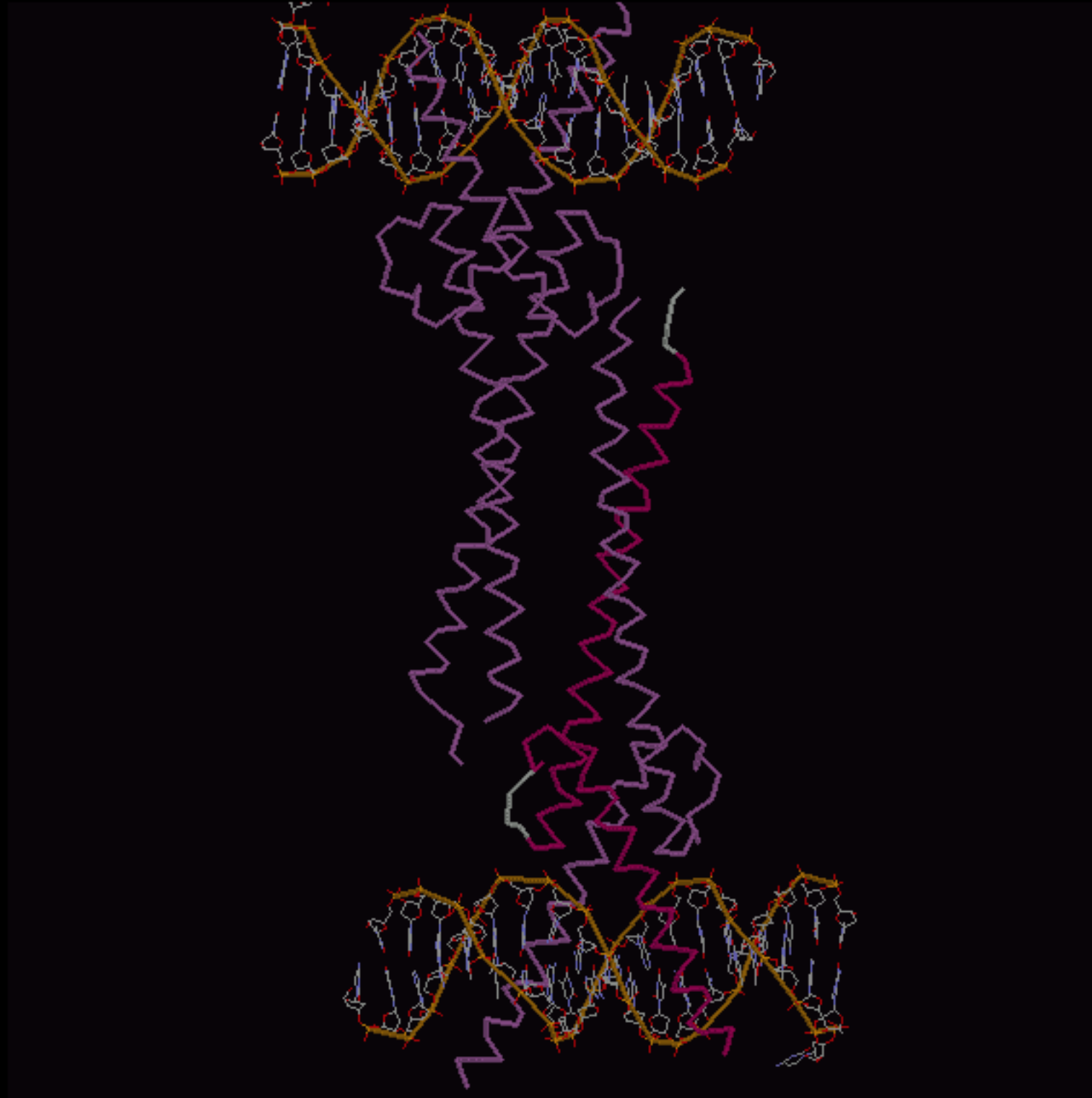
DNA-binding (ευκαρυωτικοί)

4 έλικες, SKN-1 transcription factor



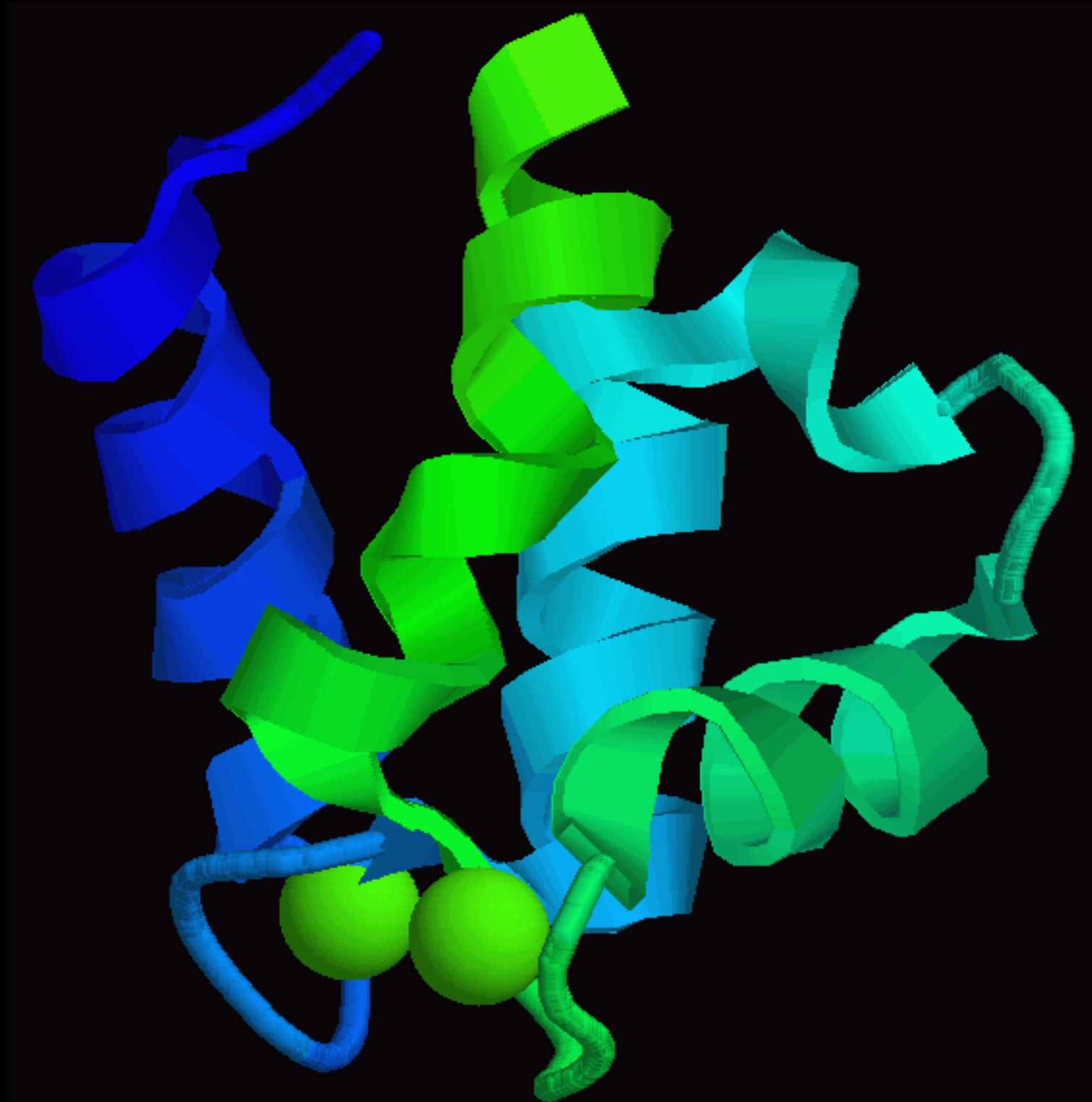
Helix-loop-helix (MYC-MAX)

4 έλικες, δεμάτιο, 2 διασταυρώσεις



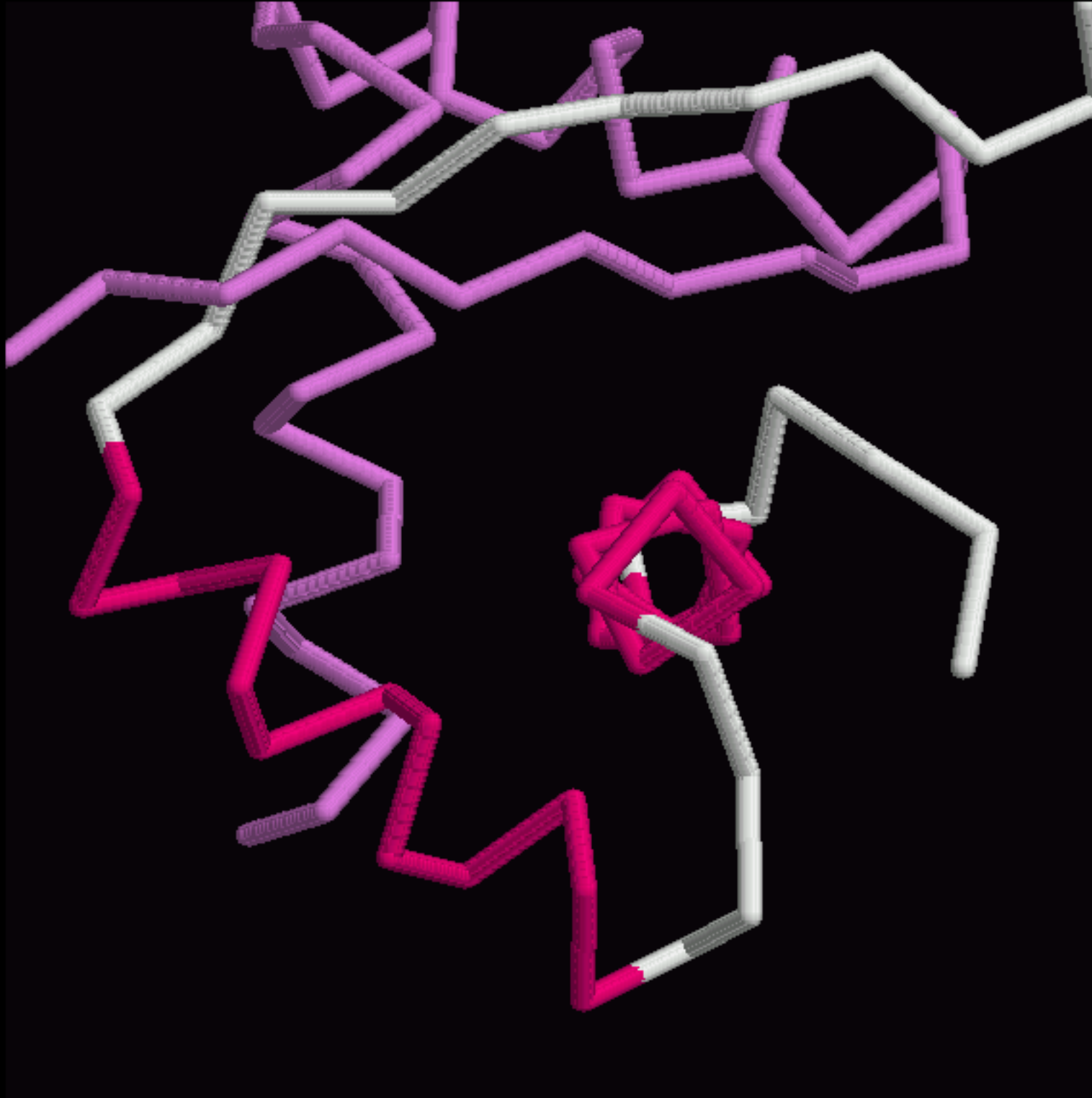
EF-hand

4 έλικες, δυο α-φουρκέτες, 2 Ca^{++} , Ca-binding



Met-repressor-like

4 έλικες, δυο α-φουρκέτες



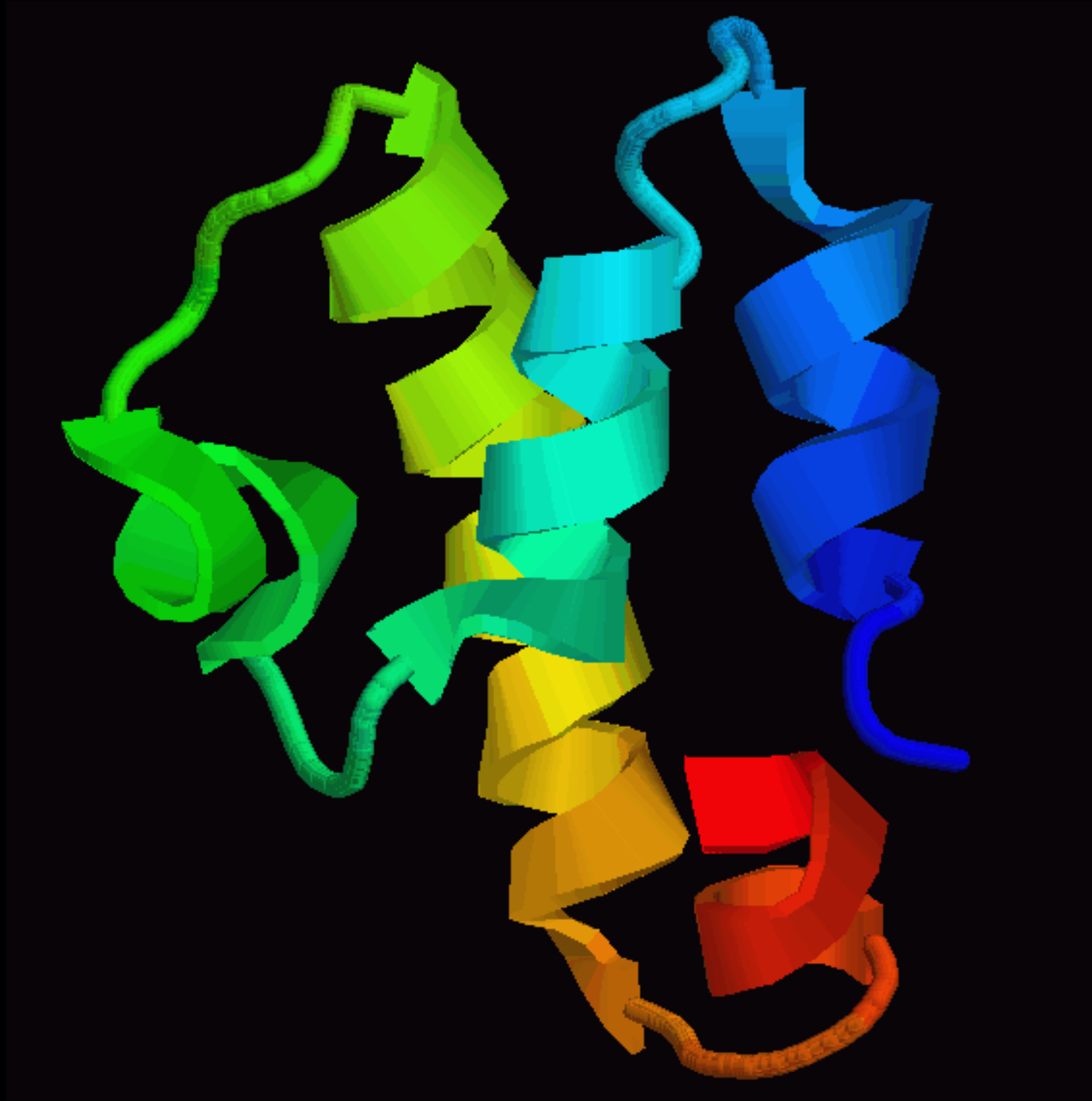
RNA-pol, υπομονάδα ω

3 έλικες, ακανόνιστη διάταξη



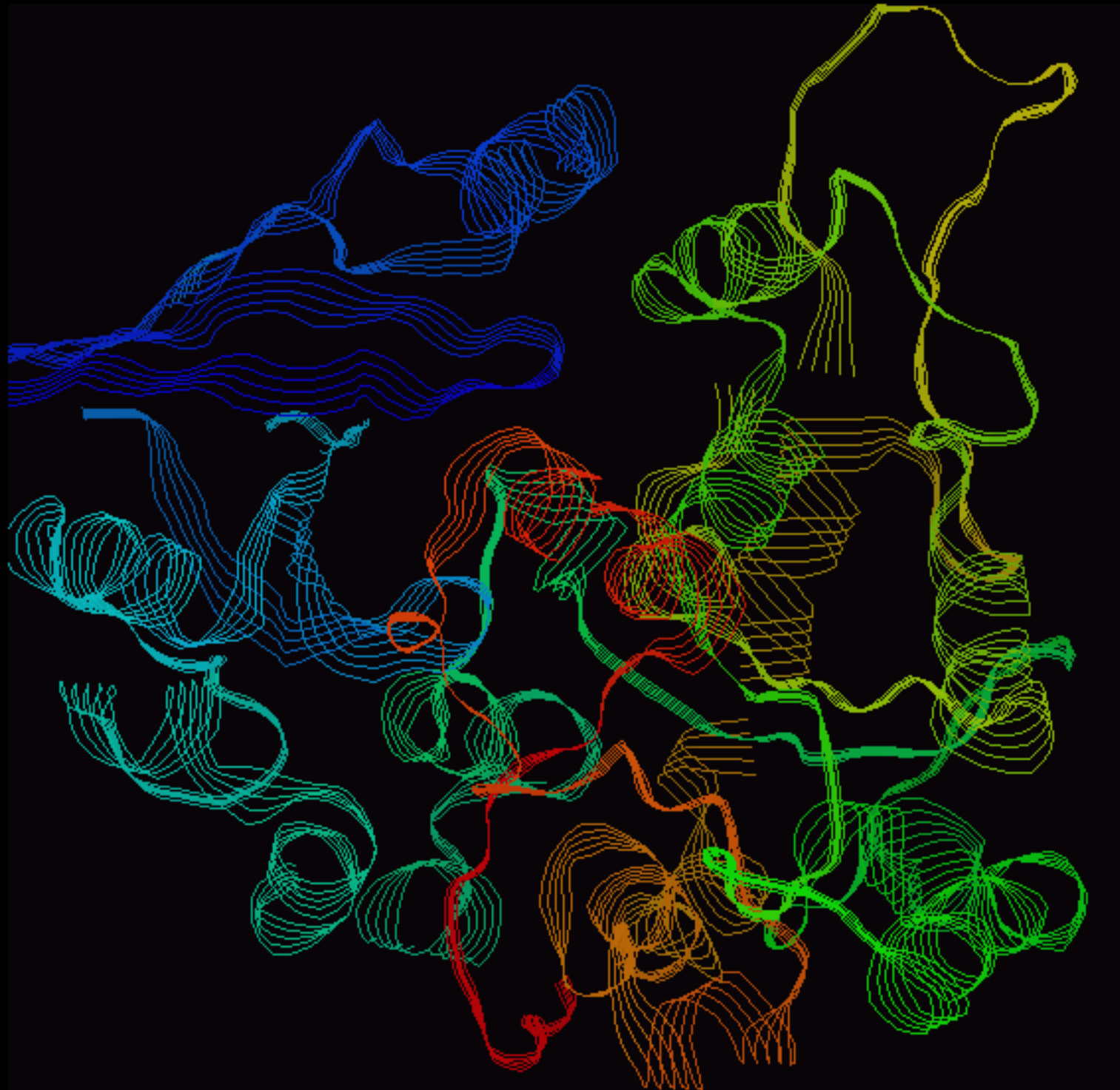
Συμπλήρωμα, αναφυλοτοξίνες

4 έλικες, ακανόνιστη διάταξη, δισουλφιδικές γέφυρες



BEACH domain

Πολυελικοειδής, ακανόνιστη διάταξη (vesicle trafficking)



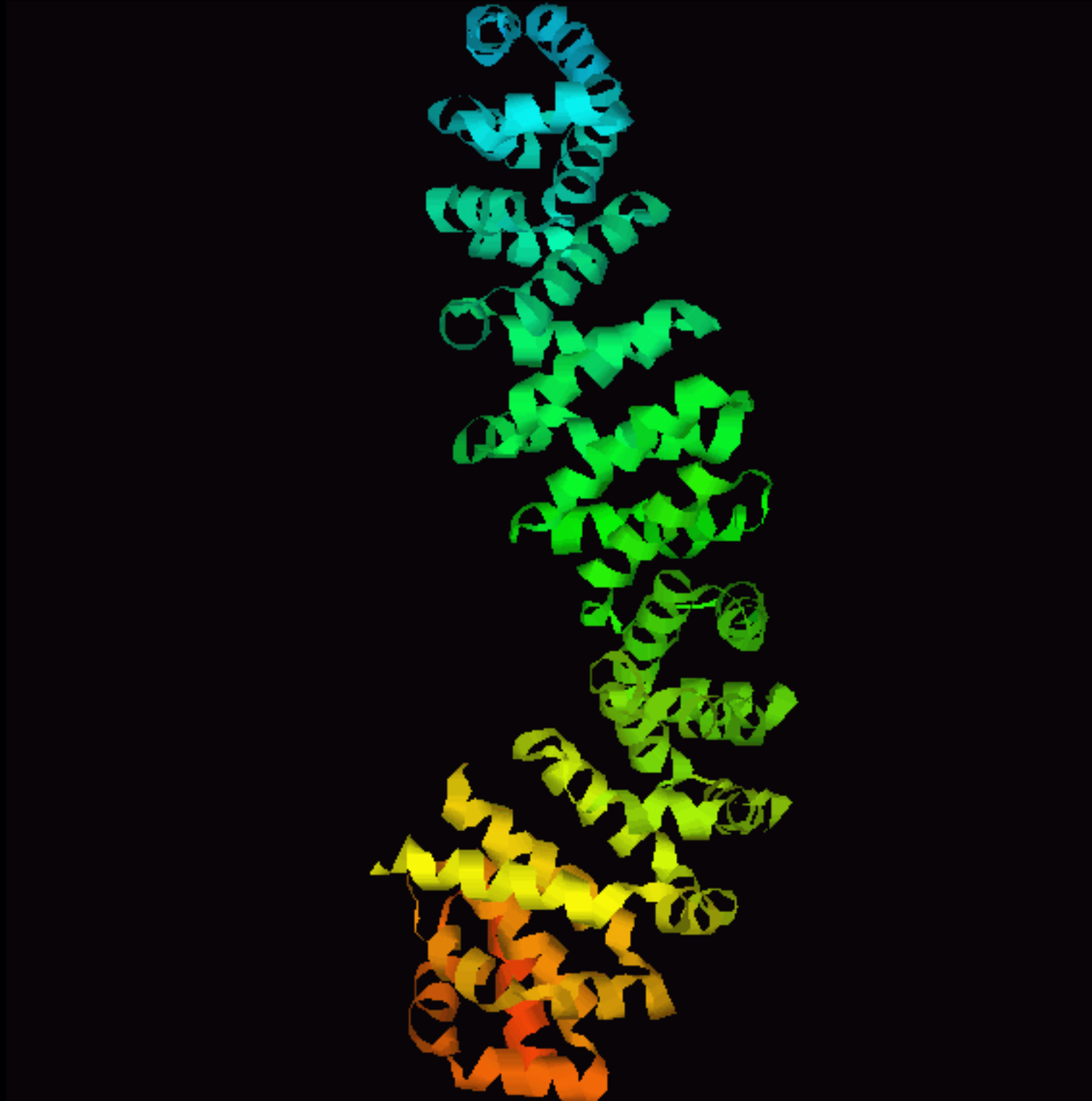
DEATH domain

6 έλικες, ψευδής συμμετρία (neurotrophin receptor)



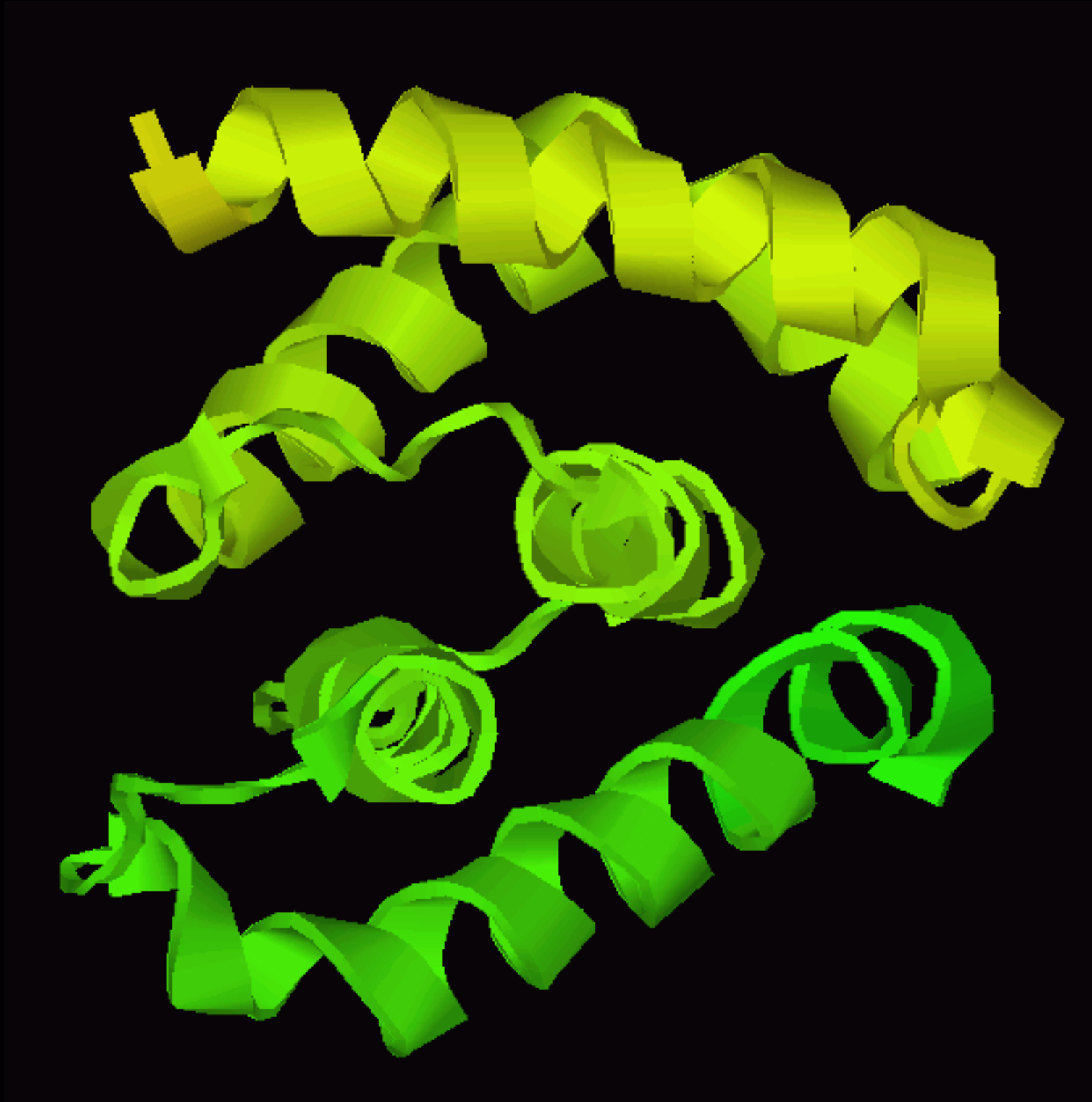
α - α -υπερέλικά

Δεξιόστροφη υπερέλικά (cell adhesion)

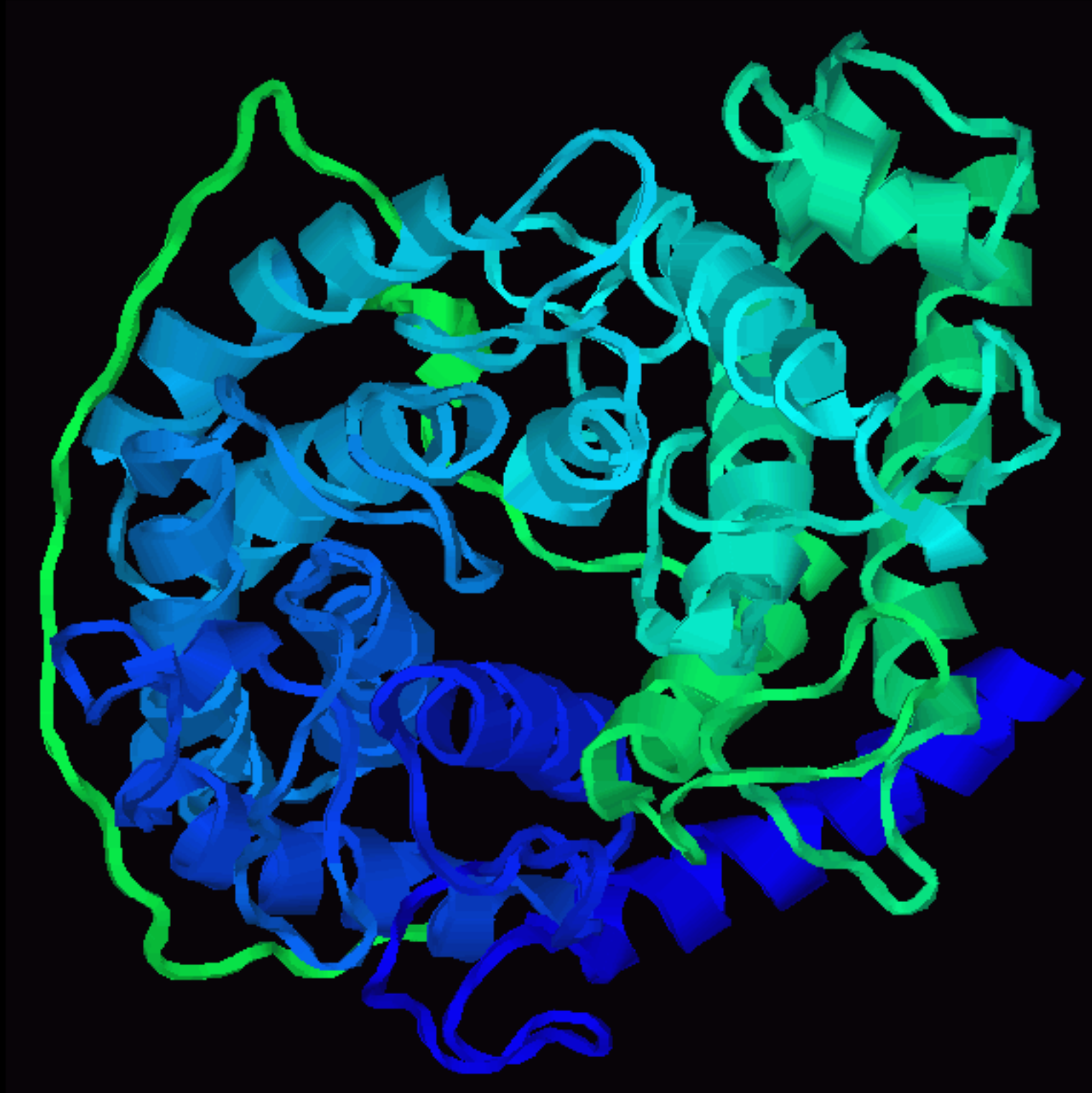


α-α-υπερέλικά

Αριστερόστροφη υπερέλικά

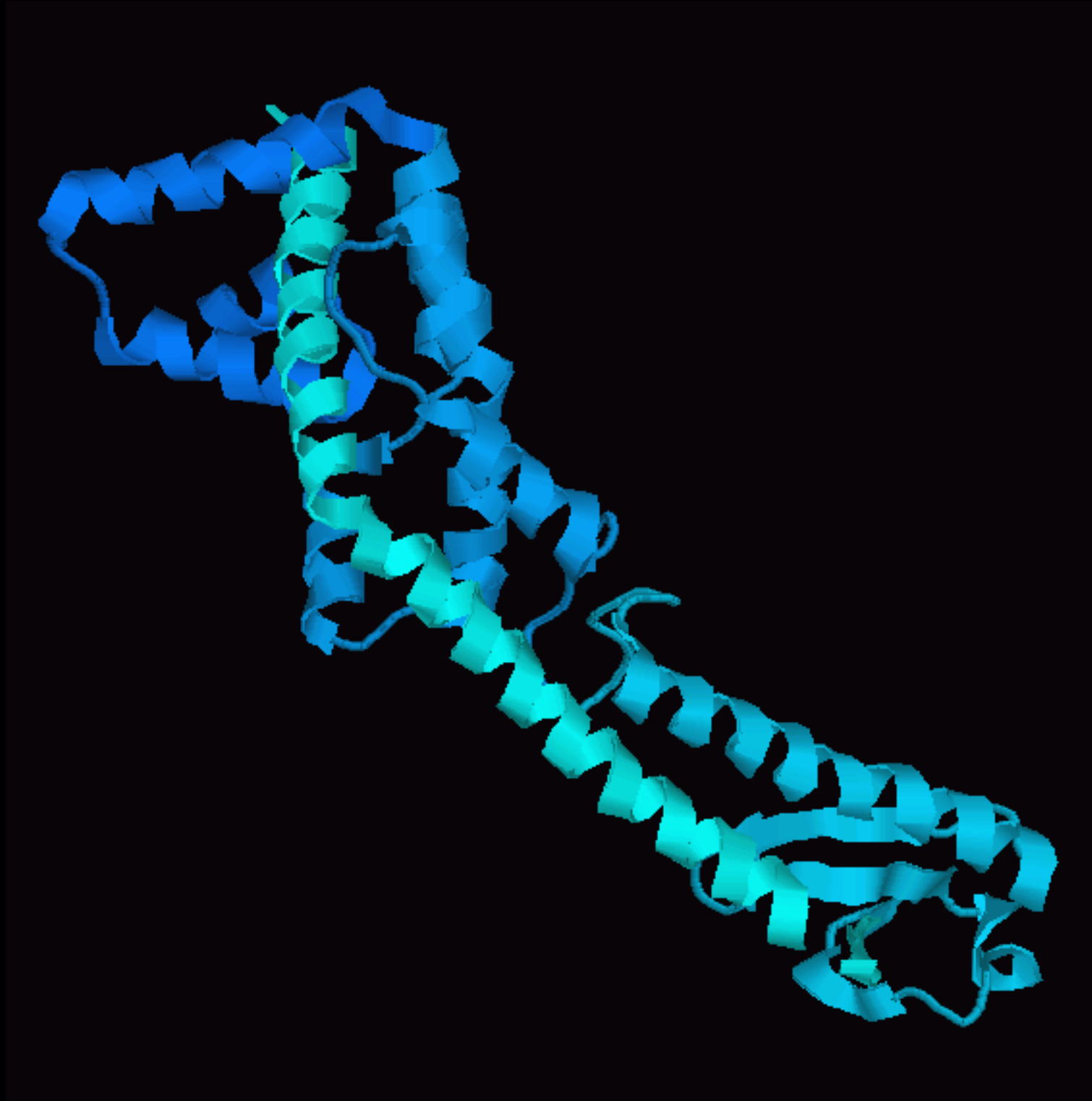


α-α-δίσκος



DNA-repair, MutS

Domain III : Πολυελικοειδές



Ηπατίτιδα Β : καψίδιο

5 έλικες, διμερισμός μέσω δύο α -φουρκετών

