

Εργαστηριακές Ασκήσεις
Εισαγωγής στην Υπολογιστική Βιολογία.

Άσκηση 8η
«C, μέρος τέταρτο (τελευταίο και φαρμακερό)»

Όνοματεπώνυμο :

Αριθμ. Μητρώου :

Δημιουργήστε τον υποκατάλογο **practicals/8/**, κάντε τον τρέχοντα κατάλογο σας, και στη συνέχεια διαβάστε προσεκτικά ολόκληρη την εκφώνηση του προβλήματος πριν ασχοληθείτε με την λύση του.

Η βακτηριοροδοψίνη είναι μία διαμεμβρανική πρωτεΐνη η οποία εμπλέκεται στη μετατροπή φωτεινής ακτινοβολίας σε χημική στα αρχαιοβακτήρια. Η δομή της πρωτεΐνης αποτελείται από επτά διαμεμβρανικές α-έλικες, με κάθε έλικα να έχει μήκος περίπου 20-25 αμινοξικά κατάλοιπα. Τα αμινοξέα που ανήκουν στις έλικες είναι κυρίως υδρόφοβα (αφού είναι στο εσωτερικό της μεμβράνης). Θέλουμε με βάση την αλληλουχία της πρωτεΐνης και την υδροφοβικότητα των αμινοξέων να βρούμε ποιά αμινοξέα ανήκουν σε ποιά από τις επτά έλικες.

Η αλληλουχία της πρωτεΐνης (την έχετε έτοιμη στον κατάλογο **~/data/**) είναι :

```
MLELLPTAVE GVSQAQITGR PEWIWLALGT ALMGLGTYLF LVKGMGVSDP
DAKKFYAITT LVPAlAFTMY LSMLLGYGLT MVPFGGEQNP IYWARYADWL
FTTPLLLLDL ALLVDADQGT ILALVGADGI MIGTGLVGAL TKVYSYRFVW
WAISTAAMLY ILYVLFFGFT SKAESMRPEV ASTFKVLRNV TVVLWSAYPV
VWLIGSEGAG IVPLNIETLL FMVLDVSAKV GFGLILLRSR AIFGEAEAPE
PSAGDGAAT SD
```

Οι υδροφοβικότητες των αμινοξικών καταλοίπων (όσο πιο μεγάλη η τιμή, τόσο πιο υδρόφοβο το αντίστοιχο αμινοξύ) είναι :

Alanine	Ala A	0.616	Methionine	Met M	0.738
Cysteine	Cys C	0.680	Asparagine	Asn N	0.236
Aspartate	Asp D	0.028	Proline	Pro P	0.711
Glutamate	Glu E	0.043	Glutamine	Gln Q	0.251
Phenylalanine	Phe F	1.000	Arginine	Arg R	0.000
Glycine	Gly G	0.501	Serine	Ser S	0.359
Histidine	His H	0.165	Threonine	Thr T	0.450
Isoleucine	Ile I	0.943	Valine	Val V	0.825
Lysine	Lys K	0.283	Tryptophan	Trp W	0.878
Leucine	Leu L	0.943	Tyrosine	Tyr Y	0.880

Γράψτε ένα πρόγραμμα σε C το οποίο θα διαβάζει από την καθιερωμένη είσοδο μία πρωτεϊνική αλληλουχία και για κάθε αμινοξύ που διάβασε θα τυπώνει στην καθιερωμένη έξοδο έναν από τους παρακάτω χαρακτήρες ανάλογα με την υδροφοβικότητα του αντίστοιχου αμινοξέος :

- Κενό : τιμή υδροφοβικότητας μικρότερη ή ίση του 0.30.
- Τελεία : τιμή υδροφοβικότητας μεγαλύτερη του 0.30 και μικρότερη ή ίση του 0.50.
- Άνω-Κάτω τελεία : τιμή υδροφοβικότητας μεγαλύτερη του 0.50 και μικρότερη ή ίση του 0.75.
- Αστέρι : τιμή υδροφοβικότητας μεγαλύτερη του 0.75.

Ένα παράδειγμα χρήσης του προγράμματος που ζητείται είναι :

```
# ./a.out
MLELLPTAVEGVSQAQITGR
:* **:..:* :*. : *.:
```

(α) Χρησιμοποιήστε το πρόγραμμα σας για να υπολογίσετε την κατανομή υδροφοβικότητας στην βακτηριοροδοψίνη. Στη συνέχεια χρησιμοποιήστε το αποτέλεσμα από τον υπολογισμό σας για να βρείτε τις θέσεις των διαμεμβρανικών ελίκων που θεωρείτε ως τις πλέον πιθανές. Τι βρήκατε ; (τα αποτελέσματα σας πρέπει να έχουν τη μορφή "Έλικα A : 11-21, Έλικα B : 22-29, ...). Συγκρίνετε τα αποτελέσματα που εσείς βρήκατε με τα πειραματικά προσδιορισμένα όρια των ελίκων όπως δίδονται στην σελίδα <https://www.rcsb.org/sequence/1m0k> (Secondary structure). Πως τα πήγατε από άποψη σφαλμάτων ?

(β) Στο χώρο που ακολουθεί γράψτε το πρόγραμμα που δημιουργήσατε.