

Εργαστηριακές Ασκήσεις Βιοπληροφορικής

Άσκηση 8η “C, μέρος τρίτο”

Όνοματεπώνυμο :

Αριθμ. Μητρώου :

Ύλη ασκήσεως :

- Ανάλυση του προγράμματος υπολογισμού μοριακών βαρών πρωτεϊνών.
- Τροποποίηση του προγράμματος υπολογισμού μοριακών βαρών ώστε να υπολογίζεται το άθροισμα των δεικτών υδροφοβικότητας των αμινοξέων μιας πρωτεΐνης.
- Τροποποίηση του προηγούμενου προγράμματος ώστε να παρουσιάζει τους δείκτες υδροφοβικότητας ανά αμινοξύ (και ανά γραμμή εξόδου).
- Περαιτέρω τροποποίηση του προγράμματος υδροφοβικότητας ώστε να παρουσιάζει στην έξοδο μια γραφική απεικόνιση της κατανομής υδροφοβικότητας για ολόκληρη την πρωτεΐνη.
- Χρήση του τελευταίου αυτού προγράμματος για τον προσδιορισμό του αριθμού και της θέσεως υδρόφοβων τμημάτων (διαμεμβρανικών α-ελίκων) στην Βακτηριοροδοψίνη.

Δημιουργήστε τον υποκατάλογο `practicals/8/` και κάντε τον τον τρέχοντα κατάλογο σας.

Αντιγράψτε από τον κατάλογο `/c0d4/people/biologic/nglykos/pract/` τα αρχεία με τα ονόματα `mw.c`, `bacteriorhodopsin` και `hydro.dat` στον τρέχοντα κατάλογο σας (δηλ. στον κατάλογο `practicals/8/`).

Χρησιμοποιήστε τον κειμενογράφο αρεσκείας σας για να εξετάσετε το περιεχόμενο του αρχείου `mw.c` και προσπαθήστε να κατανοήσετε τον τρόπο λειτουργίας αυτού του προγράμματος [το οποίο υλοποιεί μια μέθοδο υπολογισμού του μοριακού βάρους πρωτεϊνών με βάση την πρωτοταγή τους αλληλουχία, για περισσότερες πληροφορίες δείτε (μέσω netscape) τις διαφάνειες που αρχίζουν από τη σελίδα http://aspera.admin.uoi.gr/html/c_03/sld043.htm].

Μεταγλωττίστε το πρόγραμμα που περιέχεται στο `mw.c` και μετονομάστε το προκύπτον εκτελέσιμο αρχείο σε `MWcalc` (από `a.out` που ήταν, βέβαια, το αρχικό του όνομα).

Χρησιμοποιήστε το πρόγραμμα `MWcalc` για να υπολογίσετε το μοριακό βάρος της πρωτεΐνης βακτηριοροδοψίνης της οποίας η αλληλουχία περιέχεται στο αρχείο `bacteriorhodopsin`. Ποιά εντολή χρησιμοποιήσατε και ποιο το αποτέλεσμα του υπολογισμού ;

Το μοριακό βάρος αυτής της πρωτεΐνης είναι 28256 Daltons. Σε τι νομίζετε ότι οφείλεται η διαφορά των 18 Daltons από την τιμή που υπολογίσατε με το πρόγραμμα `MWcalc` ; Πως θα πρέπει να τροποποιηθεί το `mw.c` ώστε να διορθωθεί αυτή η απόκλιση ;

Αντιγράψτε τα περιεχόμενα του αρχείου `mw.c` σε ένα αρχείο με το όνομα `hydro1.c` (στον τρέχοντα κατάλογο σας).

Χρησιμοποιήστε την εντολή `more` ή `cat` για να εξετάσετε το περιεχόμενο του αρχείου `hydro.dat` το οποίο περιέχει (στην τελευταία του στήλη) μια αριθμητική προσέγγιση (σε κανονικοποιημένη λογαριθμική κλίμακα) των συντελεστών (ή δεικτών) υδροφοβικότητας των 20 L-αμινοξέων (όπου όσο μεγαλύτερη η τιμή, τόσο πιο υδρόφοβο είναι το αμινοξύ, και όσο μικρότερη, τόσο πιο υδρόφιλο).

Τροποποιήστε το πρόγραμμα `hydro1.c` με τέτοιο τρόπο ώστε αντί για το άθροισμα των μοριακών βαρών των αμινοξέων, να υπολογίζει το άθροισμα των συντελεστών υδροφοβικότητας. Μεταγλωττίστε το `hydro1.c` και μετονομάστε το προκύπτον εκτελέσιμο αρχείο σε `hydro1` (από `a.out` που ήταν, βέβαια, το αρχικό του όνομα).

Χρησιμοποιήστε το πρόγραμμα `hydro1` για να υπολογίσετε το άθροισμα των συντελεστών υδροφοβικότητας της βακτηριοροδοψίνης. Ποιά είναι η τιμή του αθροίσματος ;

Αντιγράψτε τα περιεχόμενα του αρχείου `hydro1.c` σε ένα αρχείο με το όνομα `hydro2.c` (στον τρέχοντα κατάλογο σας).

Τροποποιήστε το πρόγραμμα `hydro2.c` με τέτοιο τρόπο ώστε αντί να υπολογίζει το άθροισμα των συντελεστών υδροφοβικότητας να τυπώνει (στην καθιερωμένη έξοδο) τα σύμβολα των αμινοξέων ακολουθούμενων από τις τιμές των συντελεστών υδροφοβικότητας τους (ένα αμινοξύ ανά γραμμή εξόδου), όπως φαίνεται παρακάτω :

```
M 0.738
L 0.943
E 0.043
L 0.943
L 0.943
P 0.711
.....
T 0.450
S 0.359
D 0.028
```

Μεταγλωττίστε το τροποποιημένο `hydro2.c`, μετονομάστε το προκύπτον εκτελέσιμο αρχείο σε `hydro2` (από `a.out` που ήταν, βέβαια, το αρχικό του όνομα), και επιβεβαιώστε ότι οι τροποποιήσεις σας είχαν το επιθυμητό αποτέλεσμα (χρησιμοποιώντας την αλληλουχία της βακτηριοροδοψίνης ως πηγή δεδομένων).

Αντιγράψτε τα περιεχόμενα του αρχείου `hydro2.c` σε ένα αρχείο με το όνομα `hydro3.c` (στον τρέχοντα κατάλογο σας).

Τροποποιήστε το πρόγραμμα `hydro3.c` με τέτοιο τρόπο ώστε η έξοδος του για την ακολουθία της βακτηριοροδοψίνης να είναι ανάλογη με αυτό που φαίνεται παρακάτω :

```
* **:.:* :*. : *.: : ****:*.:.:*.:.***** :*:*. : : **:*..**:.:*.:.**:.**:*:*.
*:*:. :***: * : ***..***** *:*** : :.*****:. :*:*.:.**:.* .**.* ******..:.*
*****:*. . : . : *:..* ** *.*****:*.*****:. :*:*. * * .***:** *.: *:*:*****
**: : : : :.:. .:..
```

όπου, το νόημα των συμβόλων είναι :

- 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.

Σημειώστε ότι το σημείο αλλαγής γραμμών στην έξοδο από το πρόγραμμα σας θα εξαρτάται από το μέγεθος του παραθύρου σας (δηλαδή το πρόγραμμα σας θα αλλάζει γραμμή αφού θα έχει γεμίσει την προηγούμενη γραμμή).

Μεταγλωττίστε το `hydro3.c`, μετονομάστε το προκύπτον εκτελέσιμο αρχείο σε `hydro3` (από `a.out` που ήταν, βέβαια, το αρχικό του όνομα), και επιβεβαιώστε ότι οι τροποποιήσεις σας είχαν το επιθυμητό αποτέλεσμα (χρησιμοποιώντας την αλληλουχία της βακτηριοροδοψίνης ως πηγή δεδομένων).

Αντιγράψτε τα περιεχόμενα του αρχείου `hydro3.c` σε ένα αρχείο με το όνομα `Hydro.c` (στον τρέχοντα κατάλογο σας).

Τροποποιήστε το πρόγραμμα `Hydro.c` με τέτοιο τρόπο ώστε η έξοδος του για την ακολουθία της βακτηριοροδοψίνης να είναι πανομοιότυπη με αυτό που φαίνεται στην επόμενη σελίδα :

