

Εργαστηριακές Ασκήσεις Βιοπληροφορικής

Άσκηση 10η

“Μια μικρή άσκηση Βιοπληροφορικής”

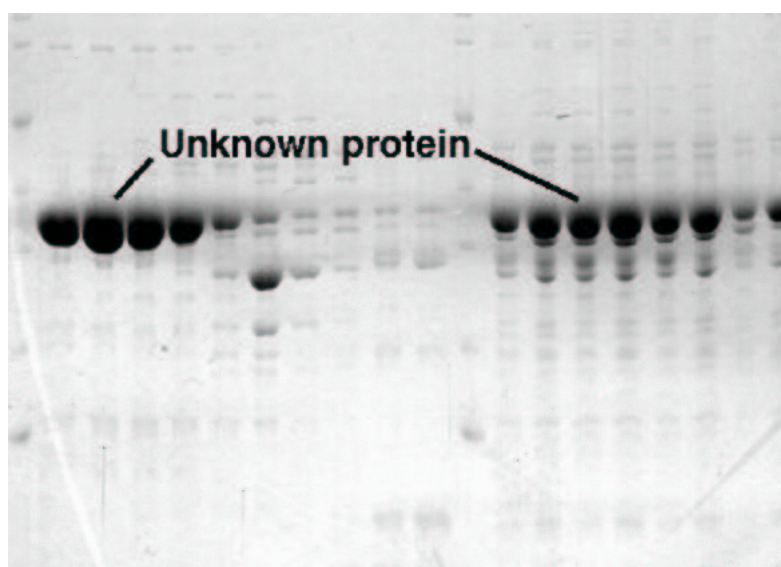
Όνοματεπώνυμο :

Αριθμ. Μητρώου :

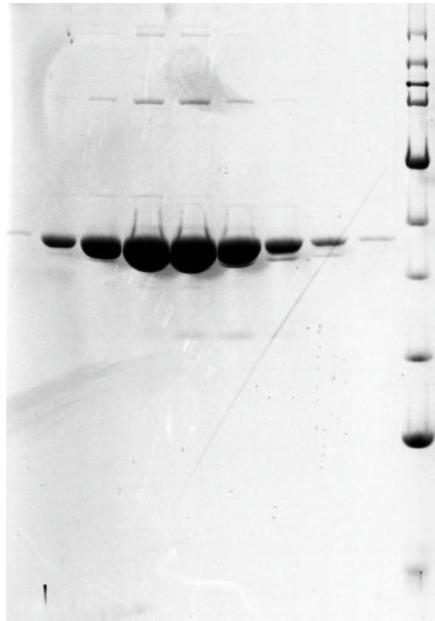
Ύλη ασκήσεως :

- Έρευνα και χρήση (μέσω του διαδικτύου) των βάσεων δεδομένων και των εφαρμογών τους με σκοπό την ταυτοποίηση και χαρακτηρισμό μιας πρωτεΐνης με βάση ελλιπή και αποσπασματικά στοιχεία.
-

Κατά την διάρκεια μιας απόπειρας απομόνωσης της πρωτεΐνης BseCI —μιας DNA μεθυλάσης από τον οργανισμό *Bacillus subtilis*— η οποία είχε υπερεκφραστεί σε κύτταρα *Escherichia coli*, παρατηρήθηκε ότι το φαινομενικό μοριακό βάρος της υπερεκφρασμένης πρωτεΐνης (όπως κρίθηκε από ένα πείραμα ηλεκτροφόρησης σε πήκτωμα ακρυλαμίδης (SDS-PAGE), δείτε εικόνα που ακολουθεί), ήταν περίπου 50 KDa, σημαντικά μικρότερο από το αναμενόμενο μοριακό βάρος των 75 KDa.



Υποθέτοντας ότι η πρωτεΐνη αυτή (των ~50 KDa) αντιστοιχεί σε ένα πρωτεολυτικό κομμάτι της πλήρους πρωτεΐνης (της BseCI δηλαδή), αποφασίστηκε να συνεχιστεί ο πρωτεϊνικός καθαρισμός με στόχο την απομόνωση αυτού του τμήματος (και με απώτερο σκοπό τον προσδιορισμό της τριτοταγούς του δομής μέσω κρυσταλλογραφίας ακτίνων X). Μετά την επιτυχή απομόνωση της άγνωστης πρωτεΐνης (με τεχνικές χρωματογραφίας), πραγματοποιήθηκε έλεγχος της κατάστασης ολιγομερισμού της μέσω ενός πειράματος χρωματογραφίας με μια βαθμονομημένη στήλη (χρωματογραφίας) τύπου gel filtration (δείτε εικόνα που ακολουθεί).



Το πείραμα αυτό έδειξε ότι η πλέον πιθανή κατάσταση ολιγομερισμού αυτής της πρωτεΐνης αντιστοιχεί σε ένα **ομοτετραμερές**.

Προκειμένου να ελεγχθεί η υπόθεση ότι αυτή η πρωτεΐνη είναι ένα πρωτεολυτικό τμήμα της BseCI (που ήταν το αρχικό αντικείμενο της μελέτης), ένα δείγμα της πρωτεΐνης υποβλήθηκε σε ένα πείραμα φασματοσκοπίας μαζών, ενώ ένα άλλο δείγμα χρησιμοποιήθηκε για την εύρεση της αλληλουχίας των πέντε πρώτων (N-τελικών) αμινοξέων (N-terminal sequencing).

Το πείραμα φασματοσκοπίας μαζών έδειξε ότι το μοριακό βάρος της κυρίαρχης μονομερούς μορφής αυτής της πρωτεΐνης είναι **52747 Da**, ενώ η N-τελική αλληλουχία προσδιορίστηκε ως **MENFK** (δηλαδή Μεθειονίνη - Γλουταμικό οξύ - Ασπαραγίνη - Φαινυλαλανίνη - Λυσίνη).

Με βάση αυτά τα δεδομένα και χρησιμοποιώντας τα εργαλεία και τις βάσεις δεδομένων που προσφέρονται μέσω του διαδικτύου (ξεκινώντας για παράδειγμα από τη διεύθυνση <http://www.expasy.ch/>), απαντήστε τις κάτωθι ερωτήσεις :

- Ποιά είναι η άγνωστη πρωτεΐνη, ποιά η λειτουργία της, ποιό το ακριβές μοριακό της βάρος και ποιό το κωδικό της όνομα στη βάση δεδομένων Swiss-Prot.
- Χρησιμοποιήστε ένα πρόγραμμα πρόβλεψης δευτεροταγούς δομής (για παράδειγμα το GOR-IV) προκειμένου να προσδιορίσετε την αναμενόμενη περιεκτικότητα της τριτοταγούς δομής αυτής της πρωτεΐνης σε α-έλικια, β-πτυχωτή και τυχαίο σπείραμα (random coil). Τι ποσοστά βρήκατε ;
- Χρησιμοποιήστε το πρόγραμμα BlastP για να βρείτε ομόλογες πρωτεϊνικές αλληλουχίες. Βρείτε για ποιά από αυτές τις ομόλογες πρωτεΐνες είναι γνωστή η τριτοταγής της δομή. Ποιός είναι ο κωδικός αυτής της ομόλογης πρωτεΐνης (με γνωστή τριτοταγή δομή) στη βάση δεδομένων Swiss-Prot ; Τι ποσοστό των αμινοξέων των δύο αλληλουχιών (αυτής που μελετάτε και αυτής με τη γνωστή τριτοταγή δομή) είναι πανομοιότυπα στις δυο αλληλουχίες (sequence identity) ;
- Βρείτε και δείξτε στους διδάσκοντες μια σχηματική αναπαράσταση της τριτοταγούς δομής αυτής της ομόλογης πρωτεΐνης (από τη βάση δεδομένων της Protein Data Bank, PDB).