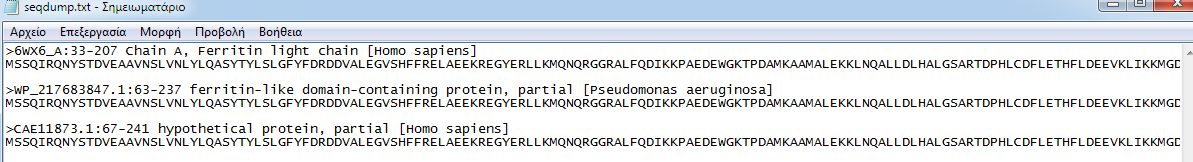
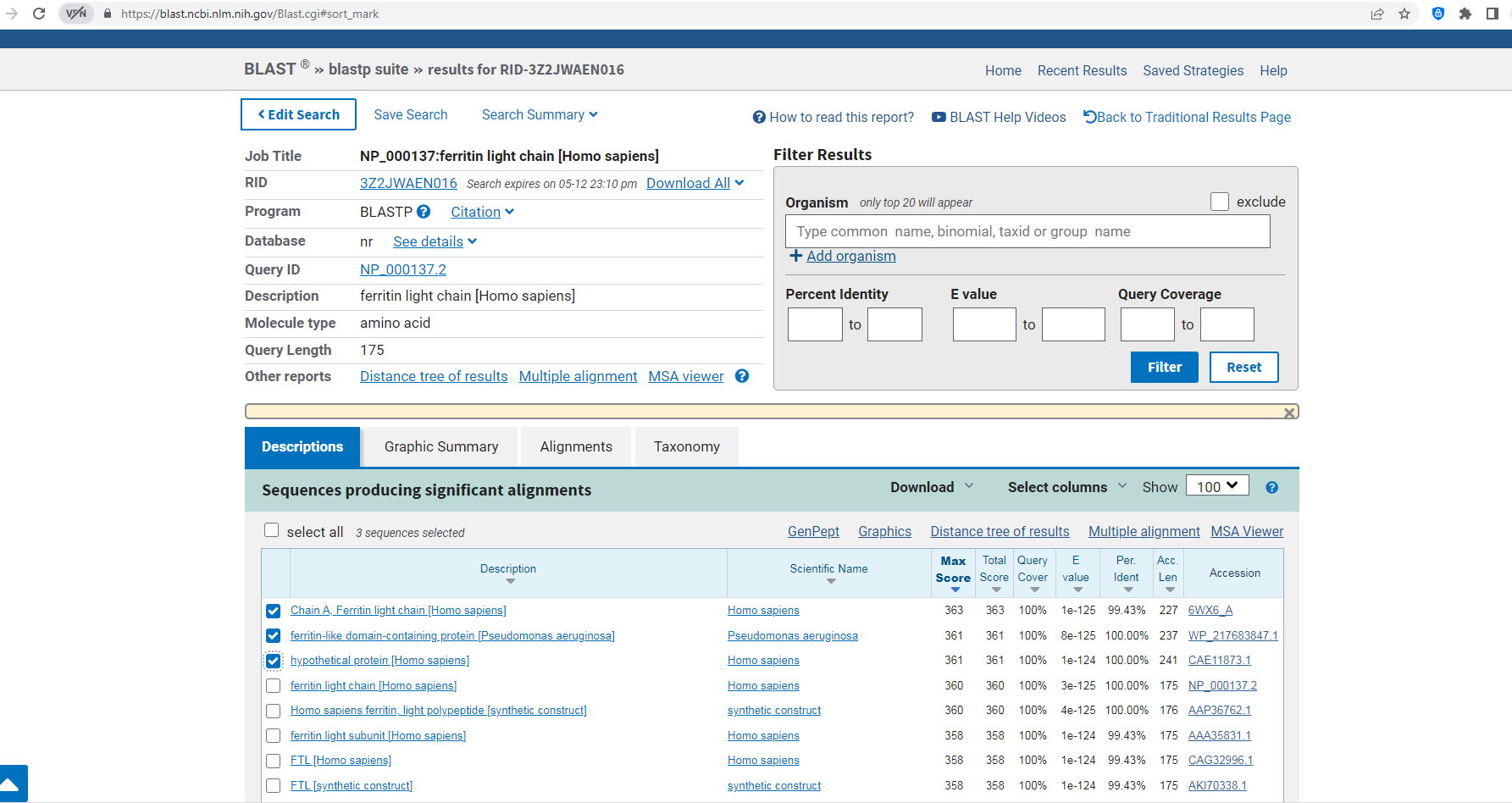
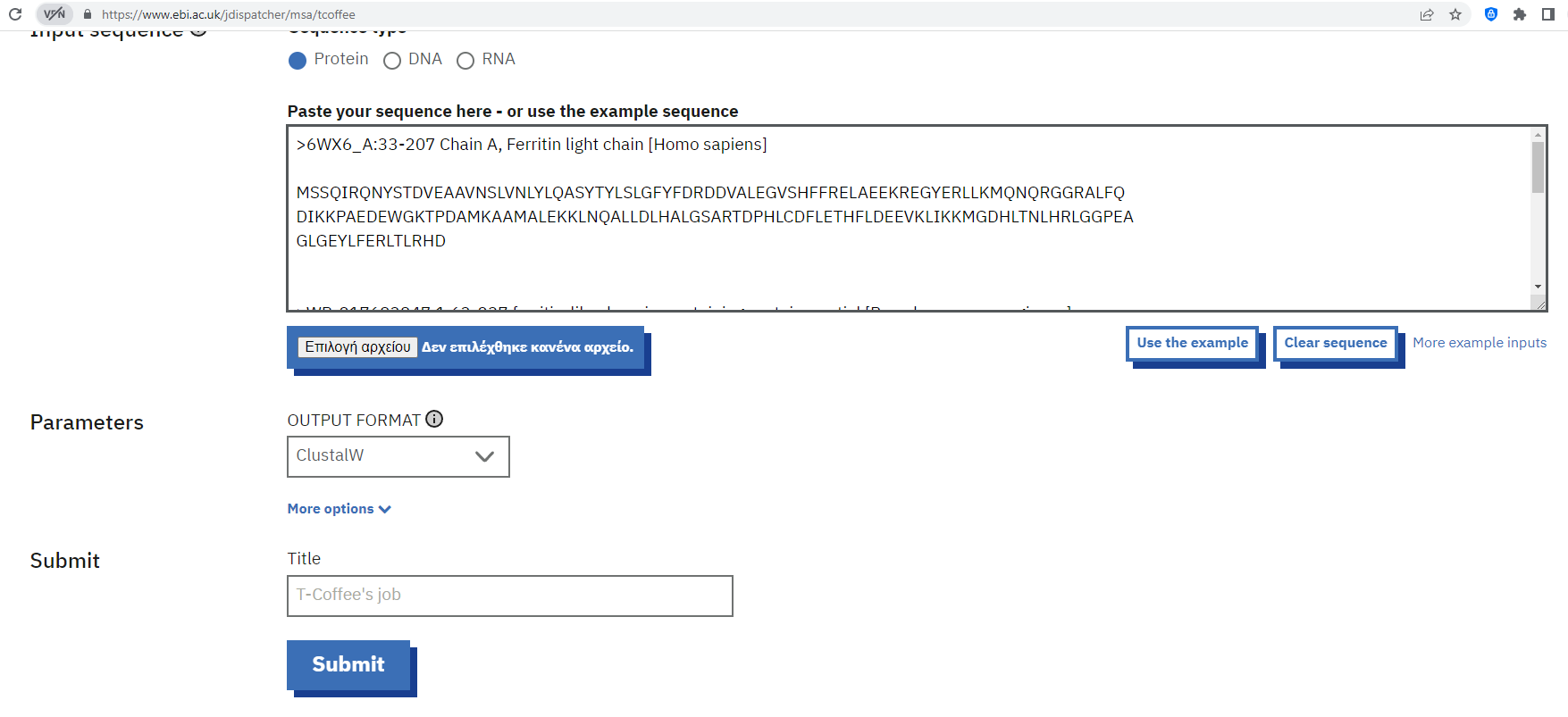
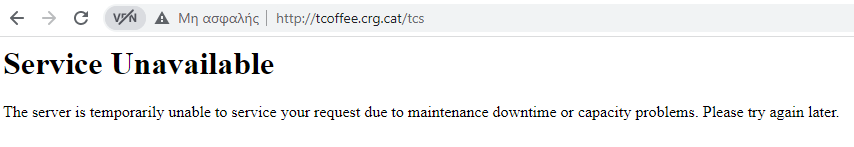
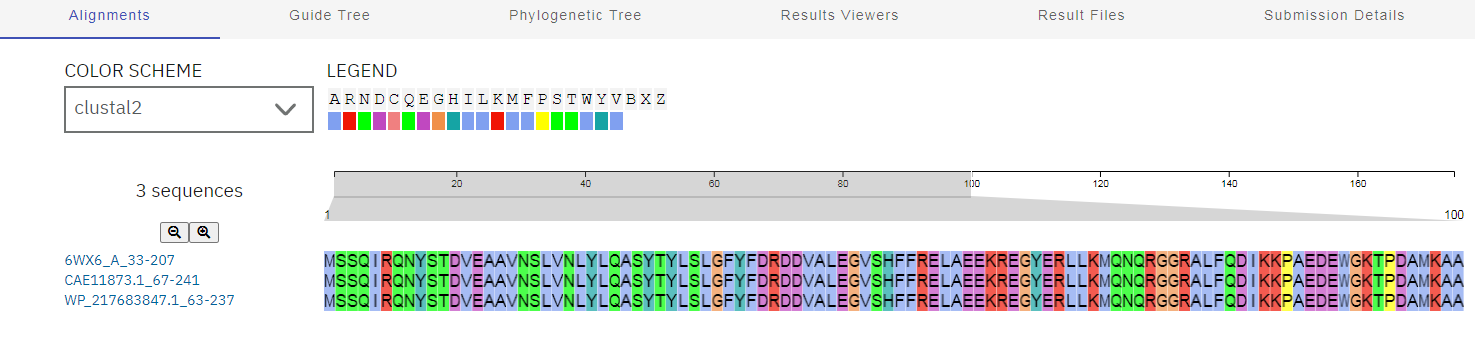
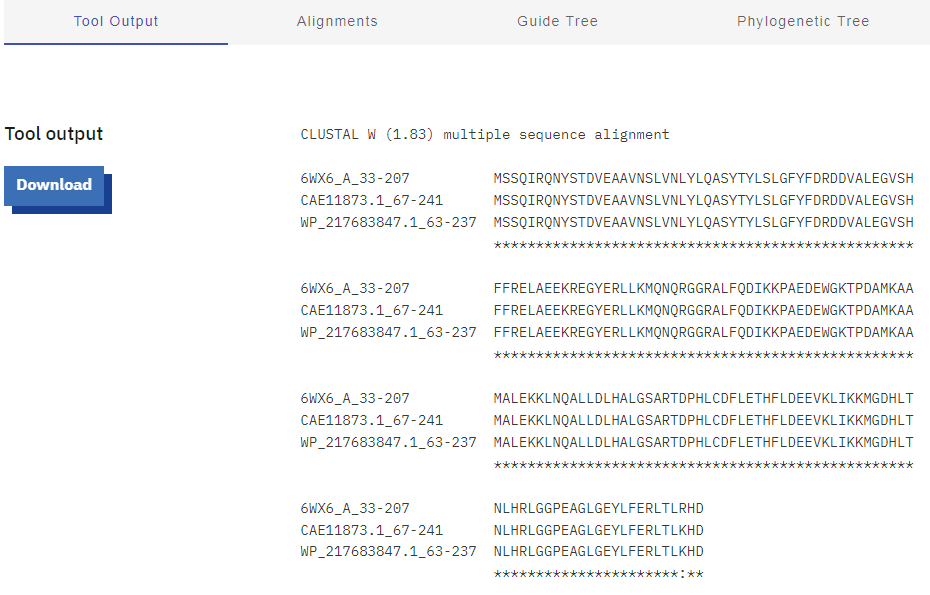
Παρακάτω πραγματοποιούμε αυτή την αναζήτηση NP\_000137 (διαλέγοντας BLASTP protein-to-protein) και διαλέγουμε να ταξινομήσουμε τις πρωτεϊνες από τις πιο όμοιες στις λιγότερο. Έπειτα επιλέγουμε τις 3 πρώτες.



Το t-coffee εδώ και κάποιες μέρες δεν είναι διαθέσιμο στην επίσημη σελίδα του. Παρόλα αυτά από το documentation του στην σελίδα [www.ebi.ac.uk](http://www.ebi.ac.uk), είναι αρκετά καλό εργαλείο για μικρές ευθυγραμμίσεις, άρα θα το χρησιμοποιήσουμε από την σελίδα του EBI.

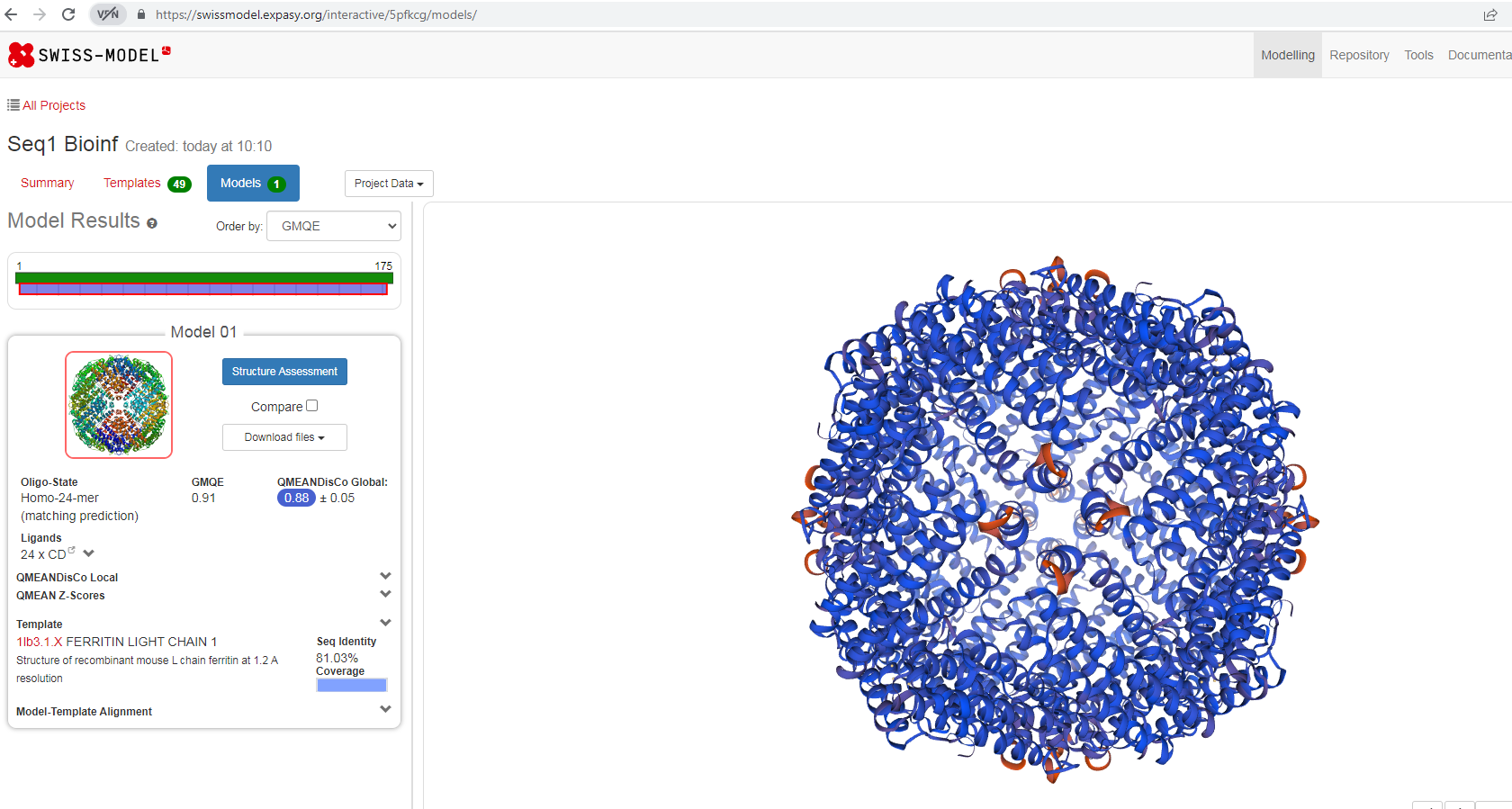


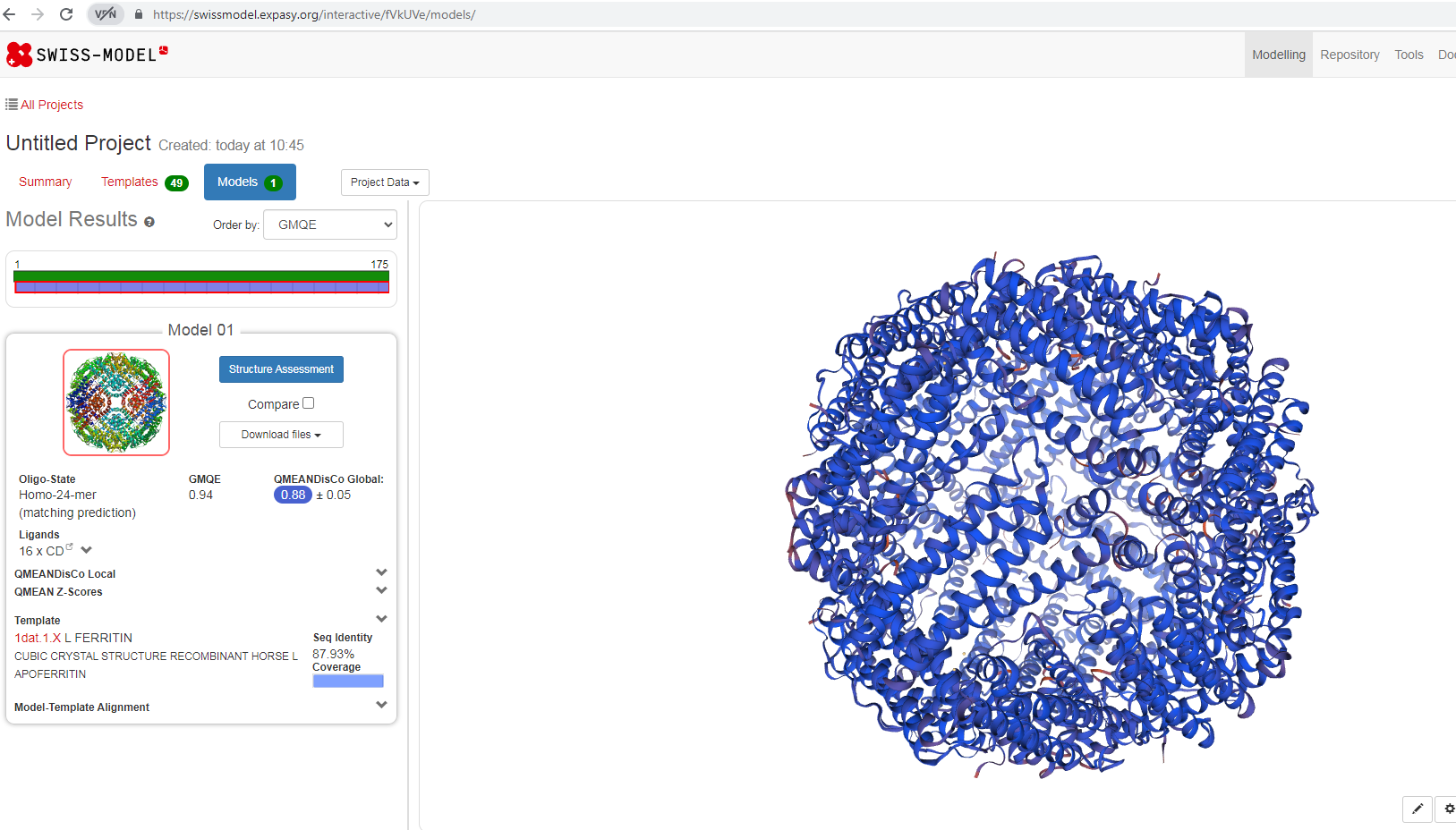
Αποτελέσματα

Ερώτημα (III)

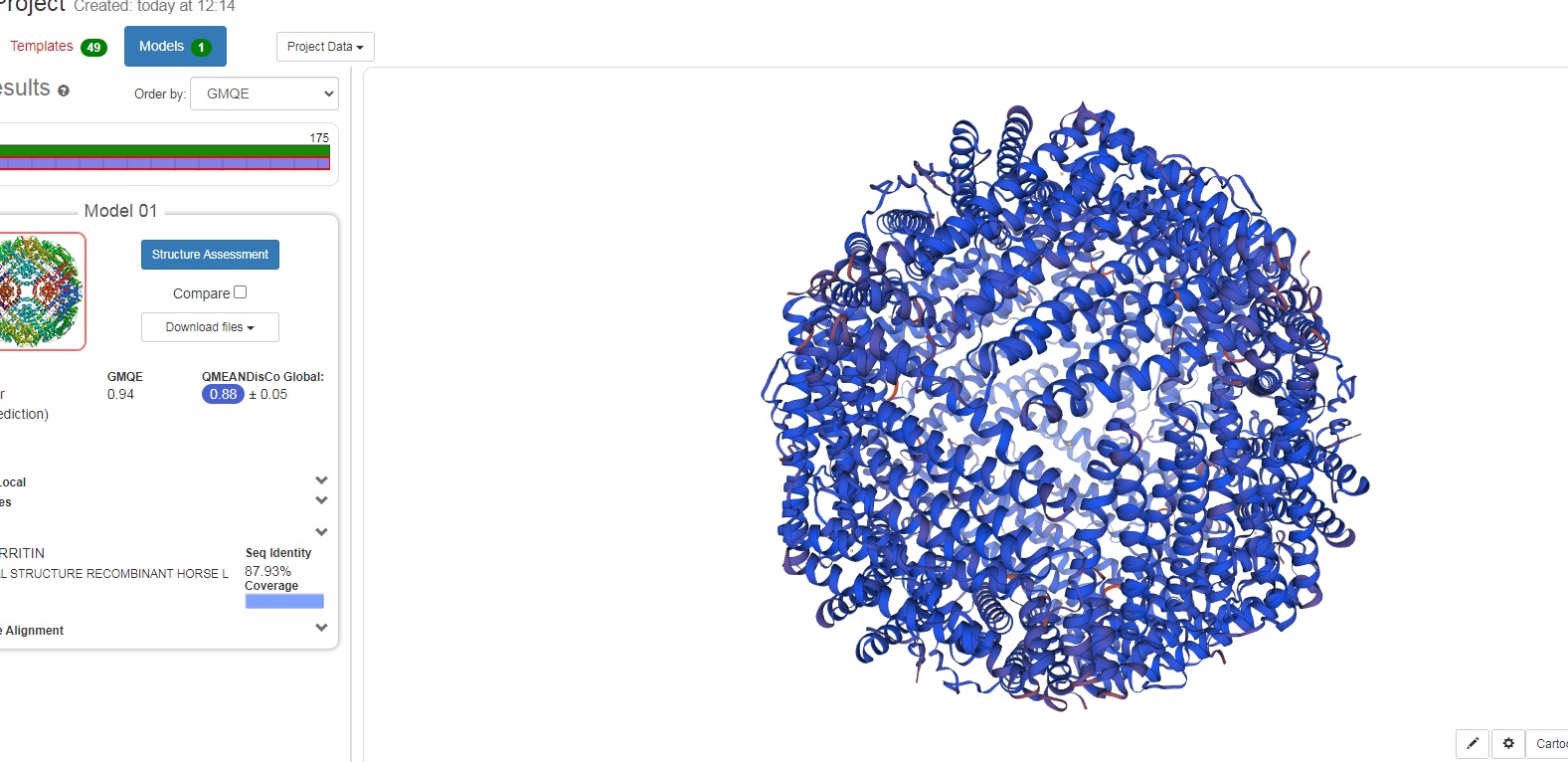
https://swissmodel.expasy.org

>6WX6\_A:33-207 Chain A, Ferritin light chain [Homo sapiens]

>WP\_217683847.1:63-237 ferritin-like domain-containing protein, partial [Pseudomonas aeruginosa]

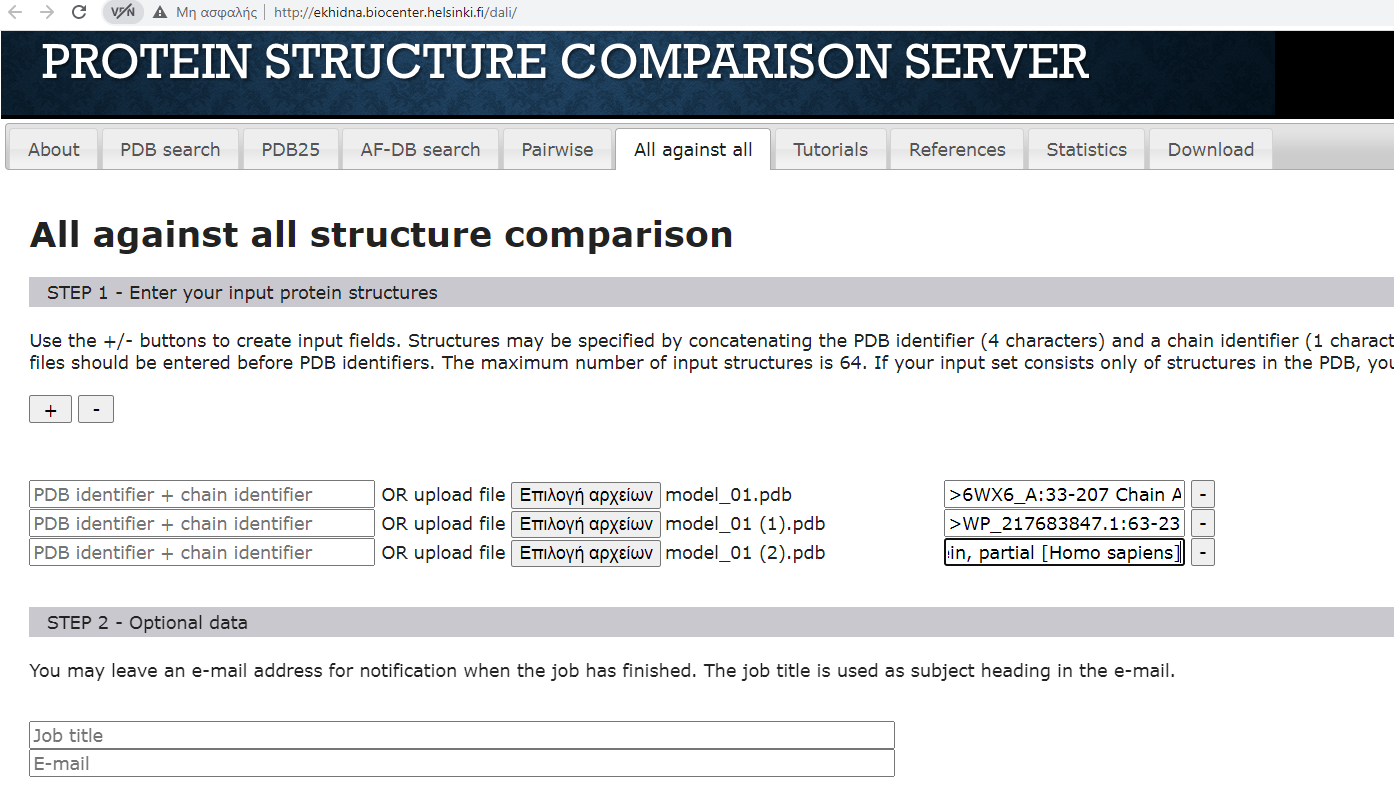


>CAE11873.1:67-241 hypothetical protein, partial [Homo sapiens]



Παρατηρήσαμε ότι η σελίδα χρειάζεται αρκετή ώρα για να ανταποκριθεί στους σχεδιασμούς (σχεδόν 2 ώρες για το κάθε ένα).

http://ekhidna.biocenter.helsinki.fi/dali/

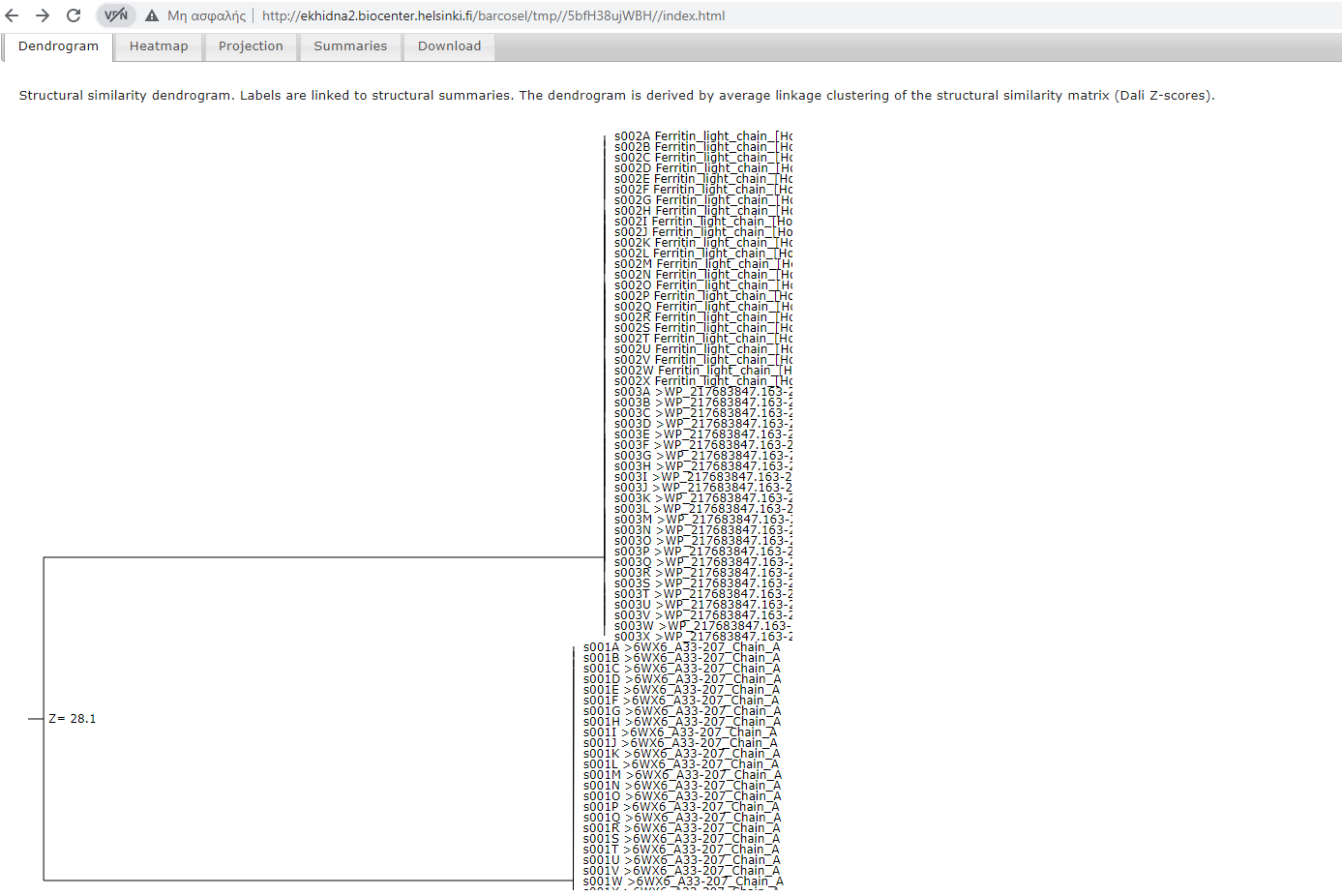
Παρακάτω βλέπουμε την ακολουθία που δώσαμε χωρισμένη σε επιμέρους κομματάκια διαφορετικών πρωτεϊνών.

6WX6\_A:33-207

WP\_217683847.1:63-237

CAE11873.1:67-241

Σύνολο 522 κομματάκια πρωτεϊνών.

Ζ=28.1

Η ομοιότητα μετράται με βάση τα Z-scores του Dali. Οι ‘σημαντικές ομοιότητες’ έχουν Z-score πάνω από 2 και συνήθως αντιστοιχούν σε παρόμοια folds. Οι ‘δυνατές αντιστοιχίσεις’ έχουν ταυτότητα ακολουθίας πάνω από 20% ή ένα Z-score πάνω από ένα όριο που εξαρτάται από το μέγεθος της πρωτεΐνης ερωτήματος. Το όριο Z-score ορίστηκε εμπειρικά σε n/10 - 4, όπου n είναι το πλήθος των αμινοξέων στη δομή της πρωτεΐνης ερωτήματος. Επιπλέον, απαιτούμε ότι η πλήρης δομή καλύπτεται από δομικές αντιστοιχίσεις· ένα τμήμα της δομής της πρωτεΐνης ερωτήματος μεγαλύτερο από 80 αμινοξέα χωρίς καμία δομική αντιστοίχιση αποκλείει πάντα μια δυνατή αντιστοίχιση.

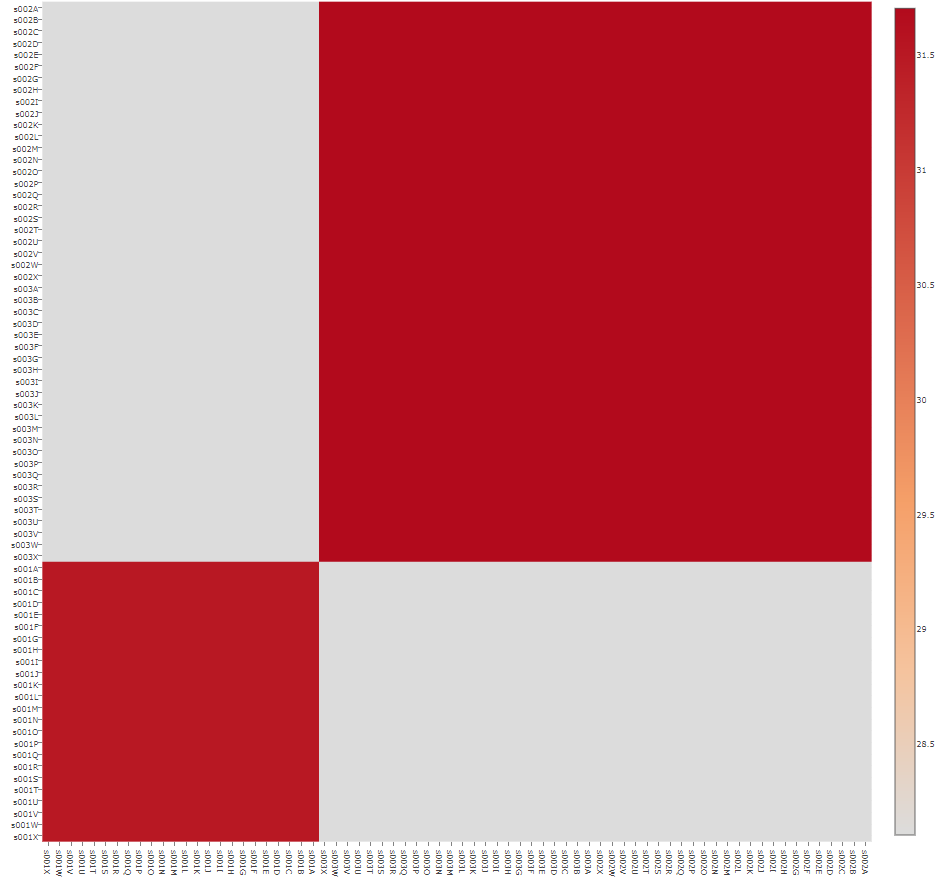
Το Ζ =28.1 δείχνει μεγάλη ομοιότητα όπως και περιμέναμε.

Πηγή : Searching protein structure databases with DaliLite v.3 L. Holm1,2,∗, S. Kääriäinen2, P. Rosenström2 and A. Schenkel2

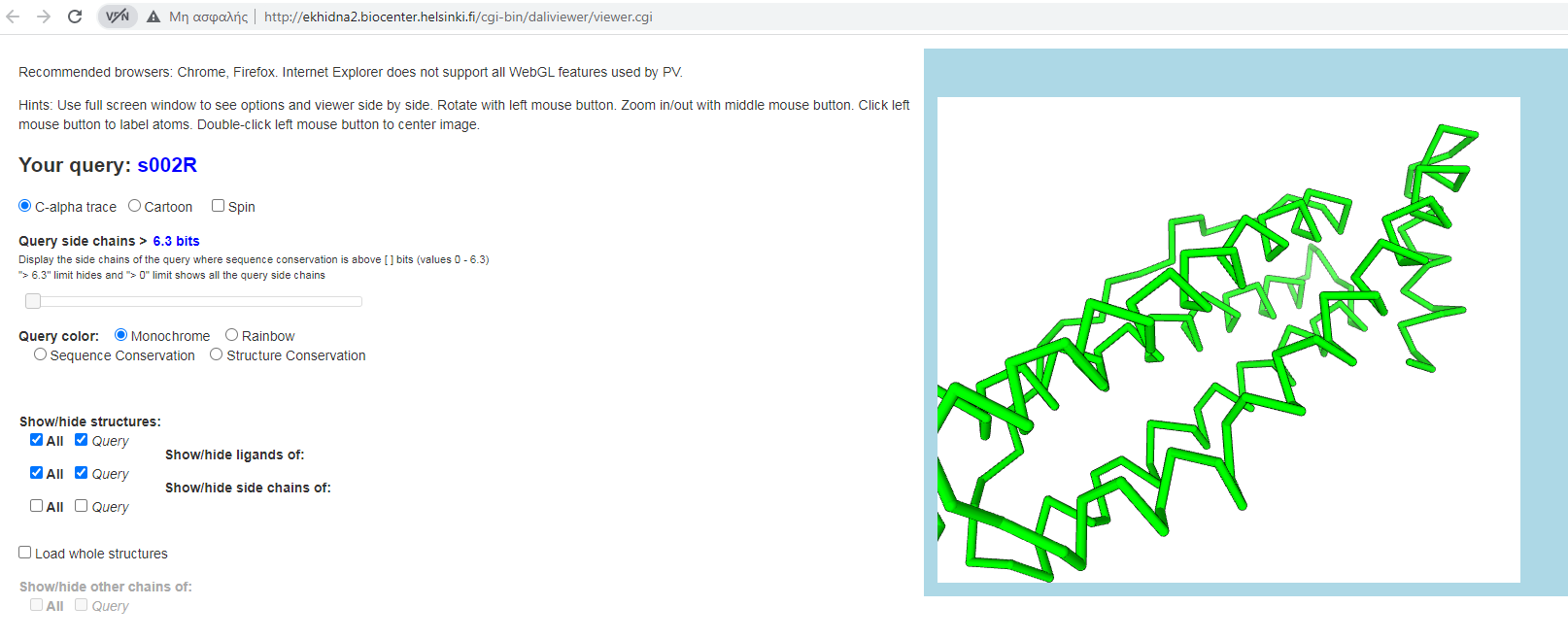
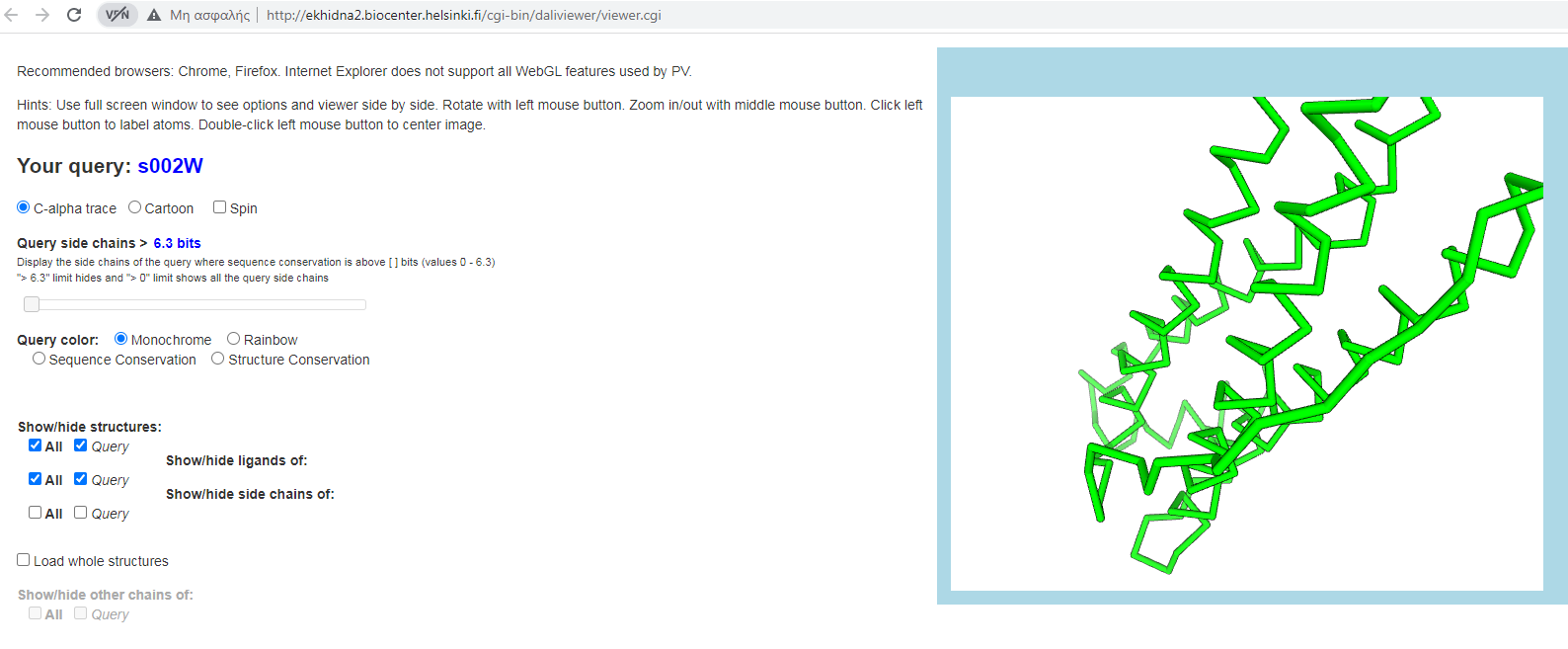
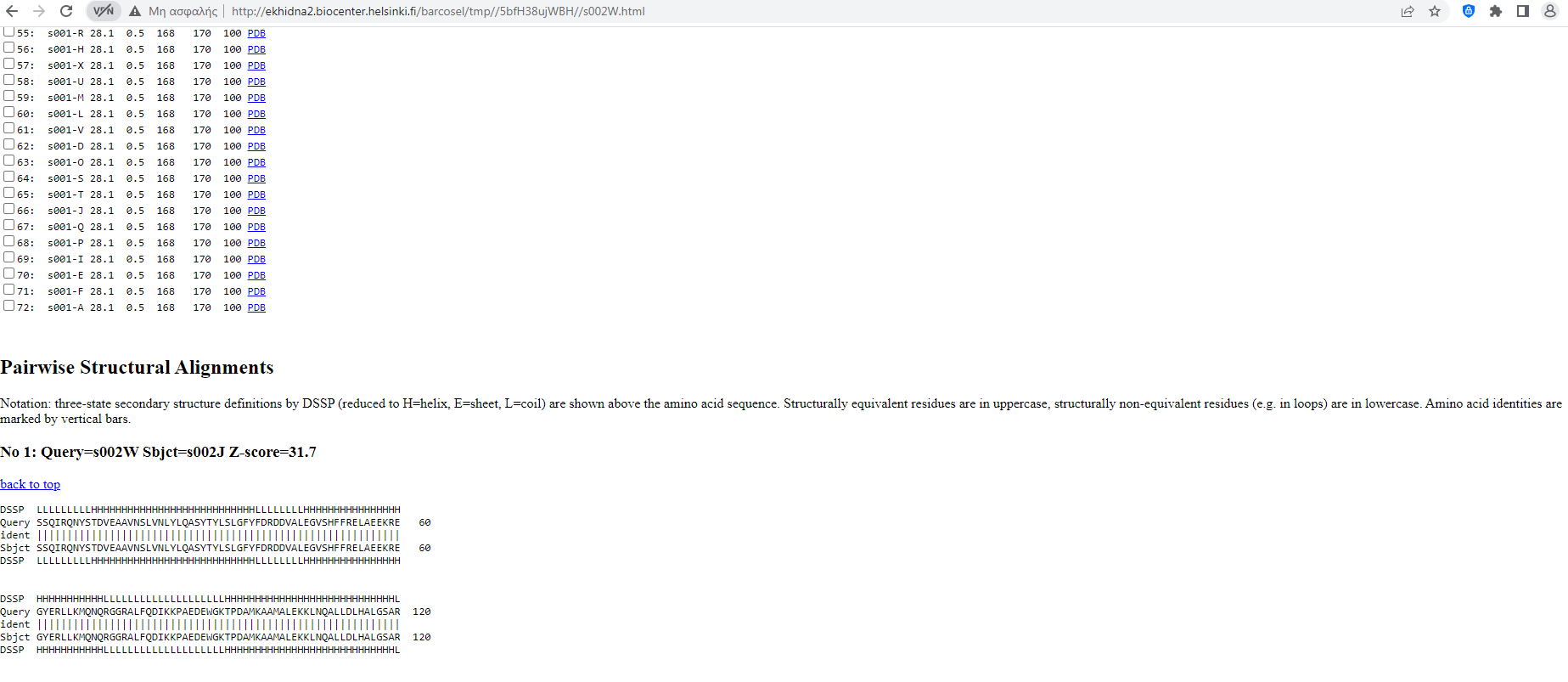
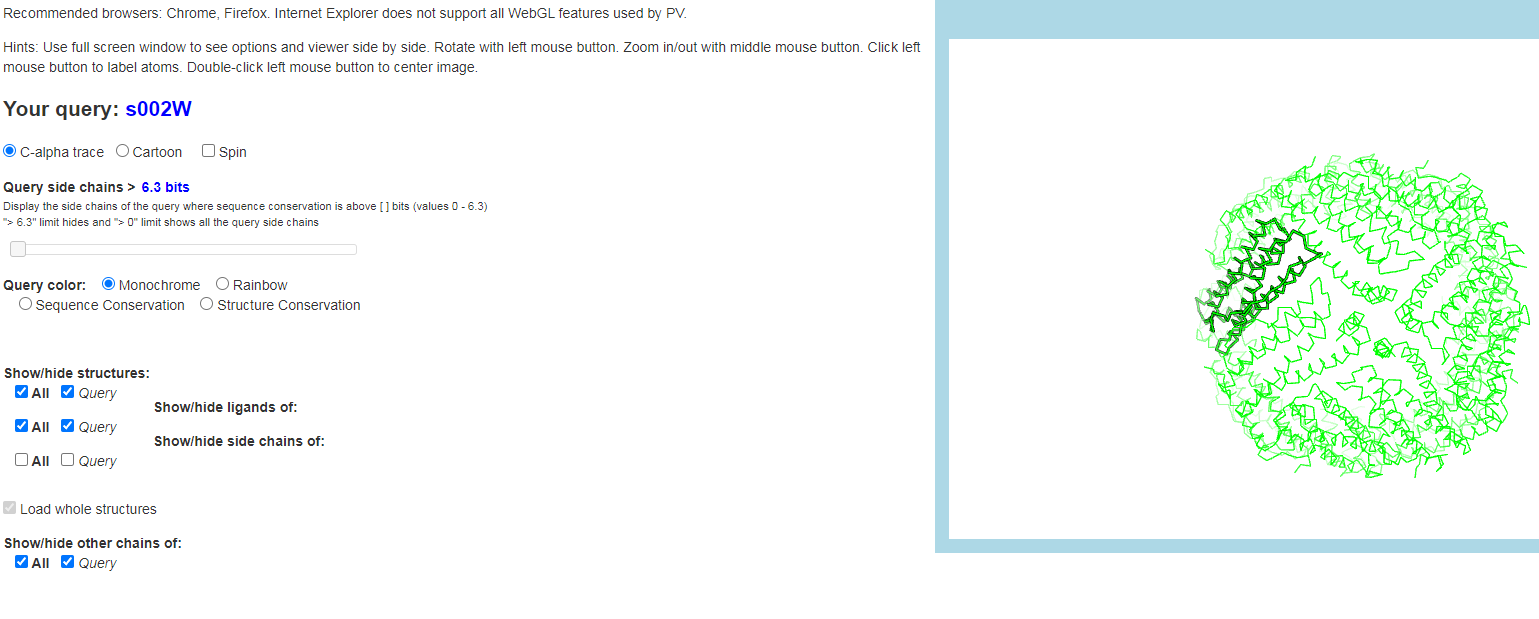
Όπως βλέπουμε το ιστόγραμμα μπορούμε να καταλάβουμε κάποια επιπλέον πράγματα. Η ferritin-like domain-containing protein, partial [Pseudomonas aeruginosa] και η hypothetical protein, partial [Homo sapiens] είναι λίγο πιο κοντά σχηματικά μεταξύ τους παρά με την 6WX6\_A:33-207 Chain A, Ferritin light chain [Homo sapiens]. Αυτή την παρατήρηση τυχαίναι να μπορούμε να την αντιληφθούμε και από τις εικόνες των 3D δομών παραπάνω. Αυτή είναι μια πληροφορία που η σειρά αμινοξέων των πρωτεϊνών δεν μας τη δίνει (χωρίς να ξέρουμε τα δομικά στοιχεία).

Πηγή: https://www.displayr.com/what-is-dendrogram/

Παρακάτω βλέπουμε ένα διαφορετικό διάγραμμα ομοιότητας για τις επιμέρους πρωτεϊνες ξεχωριστά. Το έντονο κόκκινο υποδεικνύει μεγάλη ομοιότητα.



Τέλος αν πατήσουμε πάνω σε κάποιο επιμέρους κομματάκι πρωτεϊνης βλέπουμε αναλυτικά κάποια πράγματα για αυτήν, όπως και το επιμέρους σχήμα της.

Μπορούμε να δούμε τα κομματάκια πρωτεϊνών και σε σύγκριση με ολόκληρη τη δομή. Τέλος η προβολή στο συγκεκριμένο προβλημα δείχνει μόνο 2 επιμέρους τμήματα πρωτεϊνών. Μάλλον επειδή τα υπόλοιπα είναι πολύ απομακρυσμένα και δεν υπάρχει κάποια επιλογή στοίχισης.

