

## Συγκριτικός σχεδιασμός δομής πρωτεΐνης (Homology Modeling)

(Σχεδιασμός δομής πρωτεΐνης με βάση την ομοιότητα της ακολουθίας της με πρωτεΐνες γνωστής δομής)

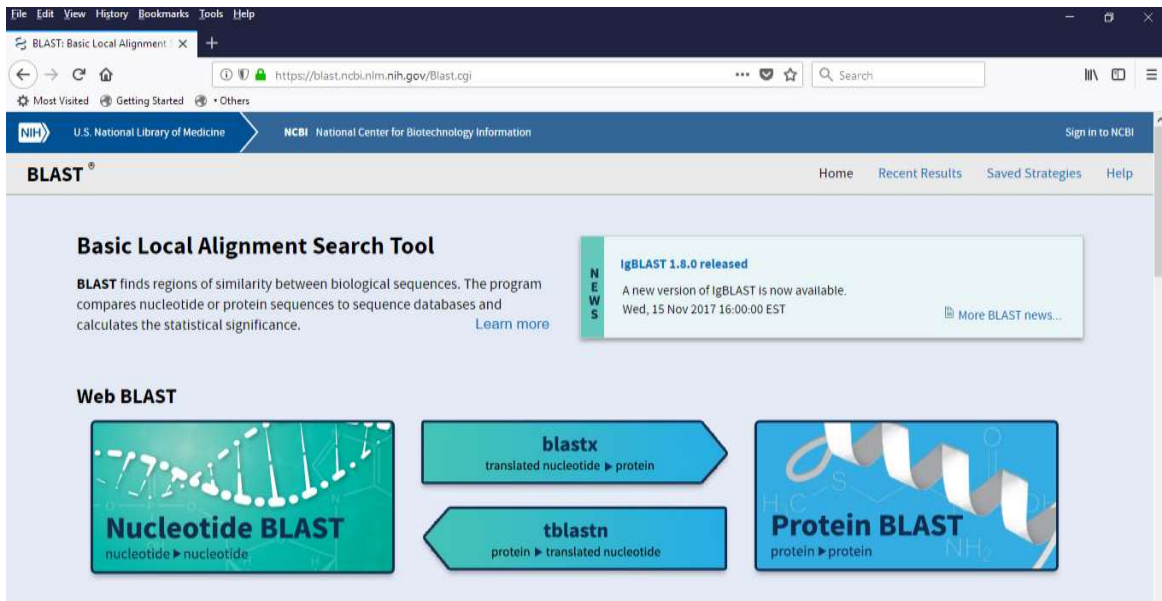
Ακολουθώντας τις οδηγίες που ακολουθούν σχεδιάστε με βάση κατάλληλο πρότυπο από την PDB την 3D δομή πρωτεΐνης με την ακολουθία (**unknown\_fasta.txt**):

```
PQFHLWKRPNVNTAHIIEGQPVEVLLDGTGAEDSIVTAIEIGVHYTPKIVGGIGGFINTKEYK  
NVEVEVLGKRIKGTIMTGNTPMNI FGRNLLTALGMSLNF
```

Προβαίνουμε στα εξής βήματα:

- 1) Εύρεση καταλλήλου προτύπου (template):
  - Αποστολή της ακολουθίας στόχου για Blast έναντι της PDB.

<http://www.ncbi.nlm.nih.gov/BLAST/>



- Πατήστε το πλήκτρο "Protein BLAST"

Protein BLAST: search protein

BLASTP programs search protein databases using a protein query. [more...](#) [Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)  Clear  Query subrange  From  To

Or, upload file  No file selected.

Job Title

Align two or more sequences

Choose Search Set

Database

Organism   Exclude

Exclude  Models (XM/XP)  Uncultured/environmental sample sequences

Entrez Query  [YouTube](#) [Create custom database](#)

- Αναζήτηση στην Protein Data Bank (Επικολλήστε την ακολουθία του αρχείου "unknown\_fasta.txt".)
- Πατήστε "Algorithmic parameters" και συμπληρώστε:

General Parameters

Max target sequences  Select the maximum number of aligned sequences to display

Short queries  Automatically adjust parameters for short input sequences

Expect threshold

Word size

Max matches in a query range

Scoring Parameters

Matrix

Gap Costs

Compositional adjustments

Filters and Masking

Filter  Low complexity regions

Mask  Mask for lookup table only  Mask lower case letters

**BLAST** Search database Protein Data Bank proteins(pdb) using Blastp (protein-protein BLAST)  Show results in a new window

Στο τέλος πατήστε το πλήκτρο "BLAST".

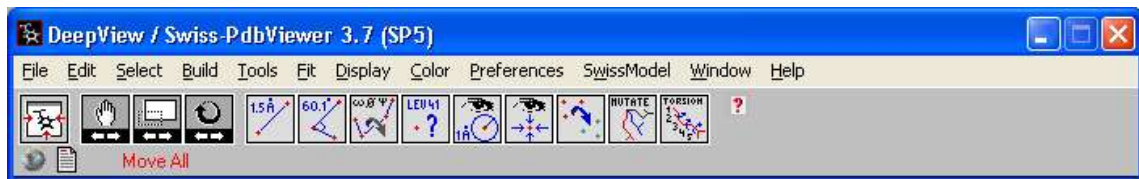
### Sequences producing significant alignments:

Select: All None Selected:0

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Chain A_SivProtease Crystallized With Peptide Product	191	191	100%	1e-64	94%	1YTI_A
<input type="checkbox"/> Chain A_STRUCTURE OF THE PROTEASE FROM SIMIAN IMMUNODEFICIENCY VIRUS. COMPLEX WITH AN IRREVERSIBLE NON-PEPTIDE INHIBIT	190	190	100%	2e-64	93%	2SAM_A
<input type="checkbox"/> Chain A_Alternative Native Flap Conformation Revealed By 2.3 Angstroms Resolution Structure Of Siv Proteinase	185	185	100%	1e-62	90%	1SIP_A
<input type="checkbox"/> Chain A_THREE-DIMENSIONAL STRUCTURE OF A SIV PROTEASE(SLASH)INHIBITOR COMPLEX. IMPLICATIONS FOR THE DESIGN OF HIV-1 AND HI	183	183	100%	1e-61	89%	1SIV_A
<input type="checkbox"/> Chain A_Crystal Structure Of Human Immunodeficiency Virus (Hiv) Type 2 Protease In Complex With A Reduced Amide Inhibitor And Comparison With I	171	171	100%	5e-57	81%	2MIP_A
<input type="checkbox"/> Chain A_THE CRYSTALLOGRAPHIC STRUCTURE OF THE PROTEASE FROM HUMAN IMMUNODEFICIENCY VIRUS TYPE 2 WITH TWO SYNTHETIC P	168	168	100%	6e-56	80%	1IVP_A
<input type="checkbox"/> Chain A_Hiv-1 (2:31-37, 47, 82) Protease Complexed With Inhibitor Sb203386	116	116	100%	4e-35	55%	1BDQ_A
<input type="checkbox"/> Chain A_Hiv-1 (2:31-37) Protease Complexed With Inhibitor Sb203386	115	115	100%	8e-35	54%	1BDL_A
<input type="checkbox"/> Chain A_Hiv-1 (2: 31, 33-37) Protease Complexed With Inhibitor Sb203386	115	115	100%	1e-34	53%	1BDR_A
<input type="checkbox"/> Chain A_Hiv-1 Pr Resistant Mutant + Lpv	114	114	100%	2e-34	52%	2RKF_A

- Αποδοχή της δομής με την καλύτερη βαθμολογία ως πρότυπο (κατά το δυνατόν χωρίς χάσματα). Αν η ομοιότητα δεν είναι ικανοποιητική δεν προχωρούμε στον σχεδιασμό.
- Λήψη του προτύπου αρχείου pdb από την βάση RCSB (<http://www.pdb.org/pdb/home/home.do>) με το κωδικό όνομα Accession (π.χ. 1YTI). Στην σημερινή άσκηση θα χρησιμοποιήσετε ως πρότυπο την δομή **1AZ5\_A\_prot.pdb**, την οποία θα βρείτε στον φάκελο ASK\_HOMOL\_MODEL. Η ακολουθία του προτύπου δίνεται στο αρχείο template\_fasta.txt.
- Χρησιμοποιήστε τον ιστότοπο [http://www.ch.embnet.org/software/LALIGN\\_form.html](http://www.ch.embnet.org/software/LALIGN_form.html) για να στοιχίσετε τις ακολουθίες unknown\_fasta.txt και template\_fasta.txt. Διαπιστώστε ότι από την ακολουθία του προτύπου λείπει το κομμάτι «GGIG», το οποίο είναι μια μικρή αναστροφή (loop).

- 2) Χρήση του προγράμματος **spdbv.exe** (από τον φάκελο spdbv) για την προσαρμογή της ακολουθίας στόχου στην δομή πρότυπο.



- Τρέξτε το πρόγραμμα SPDB και φορτώστε την ακολουθία «στόχο» unknown\_fasta.txt: SwissModel → Load Raw Sequence to Model.
- Ανοίξτε την δομή «πρότυπο» 1AZ5\_A\_prot.pdb: File → Open PDB file.

Η προσαρμογή της άγνωστης ακολουθίας στην δομή πρότυπο γίνεται με βάση την στοίχιση των δύο ακολουθιών. Λόγω της παρουσίας του χάσματος η προσαρμογή δεν θα περιλαμβάνει το κομμάτι GGIG, η δομή του οποίου θα σχεδιασθεί ανεξάρτητα.




- Window → Alignment



- Διαπιστώστε ότι μετά από την θέση 47 της ακολουθίας, λόγω της παρουσίας του χάσματος η στοίχιση δεν είναι σωστή. Διορθώστε την στοίχιση κάνοντας κλικ πάνω στην θέση 48 (G) της 1AZ5\_A\_prot.pdb και δημιουργώντας το κατάλληλο χάσμα με το πλήκτρο <SPACE>.
- Επιλέξτε στο παράθυρο Alignment με αριστερό κλικ και σύροντας (όπου χρειάζεται και με <Ctrl>) τα μέρη των ακολουθιών προτύπου και στόχου αποκλείοντας το κομμάτι GGIG.



- Προσαρμόστε την πρωτεΐνη «στόχο» στο πρότυπο για να δημιουργήσετε ένα μοντέλο:  
Fit → Realign Selection
- Μετατρέψτε σε ενεργό το επίπεδο (Layer) που περιέχει το μοντέλο σας με:  
Window → Layer infos  
και με κλικ στην πρωτεΐνη «στόχο» ώστε να κοκκινίσει.
- Σώστε το μοντέλο που μόλις κατασκευάσατε:  
File → Save → Save selected Residues ως model\_0.pdb.
- Ανοίξτε με έναν text editor το model\_0.pdb και διαπιστώστε ότι υπάρχει ασυνέχεια μεταξύ VAL47 και GLY52.

Με πατημένο δεξί κλικ μπορείτε να μετακινήσετε την δομή στο κέντρο της οθόνης. Κεντράρετε με κλικ στο  και μετά σε σημείο κάπου στο κέντρο της πρωτεΐνης. Μεγεθύνετε με  και περιστρέψτε με .

#### Σχεδιασμός της απουσίας αναστροφής GGIG για το model\_0.pdb:

- 1) Δημιουργήστε αρχείο fasta με την ακολουθία της απουσίας αναστροφής (π.χ. missing\_loop\_fasta.txt)
- 2) Φορτώστε την ακολουθία στο spdb: SwissModel → Load Raw Sequence to Model
- 3) Σώστε το προσωρινό μοντέλο της αναστροφής: Save → Layer (loop\_0.pdb)
- 4) Επαναρίθμηση ατόμων και αμινοξέων του loop\_0.pdb ώστε να συμπληρώνουν το χάσμα στο model\_0.pdb καθώς και δήλωση ονόματος αλυσίδας με το πρόγραμμα PDBrenumb.exe. Μετονομασία του προκύπτοντος αρχείου σε loop\_1.pdb.
- 5) Ανοίξτε με text editor το μοντέλο με το χάσμα (model\_0.pdb) και ενσωματώστε το κενό με την δομή του loop\_1.pdb. Σώστε ως model\_1.pdb.
- 6) Φορτώστε το model\_1.pdb στο spdbn για να συγκολλήσετε τον σκελετό.
  - α) Συγκόλληση αμινοτελικού άκρου της αναστροφής  
Build → Ligate Backbone → Pick atom (κλικ στο αμινοτελικό άκρο του loop)
  - β) Συγκόλληση καρβοξυτελικού άκρου της αναστροφής  
Build → Ligate Backbone → Pick atom (κλικ στο καρβοξυτελικό άκρο του loop)
 Σώστε το layer ως model\_2.pdb.
- 7) Build → Build Loop  
Pick 1st anchor: κλικ ένα αμινοξύ προ της αναστροφής.  
Pick 2nd anchor: κλικ ένα αμινοξύ μετά από την αναστροφή.  
Από την λίστα που προκύπτει διαλέξτε εκείνο που βολεύει.
- 8) Ελαχιστοποίηση ενέργειας για την αναστροφή.  
Από το Window → Control Panel επιλέξτε τα κατάλοιπα G48, G49, I50, G51.  
Tools → Energy Minimization  
Σώστε το τελικό μοντέλο ως model.pdb

#### Έλεγχος της ποιότητας του μοντέλου

- Με την βοήθεια του "Ramachandran Plot": Ένα καλό μοντέλο έχει > 90% των καταλοίπων του στα πλαίσια των στενά επιτρεπών περιοχών του Ramachandran plot και > 98% στις ευρύτερα επιτρεπές.

Select → All  
Window → Ramachandran Plot

Μετρήστε τα αμινοξέα έξω από τις κίτρινες περιοχές (αγνοήστε τις Gly) και βρείτε το ποσοστό των αμινοξέων του μοντέλου που εμπίπτουν σ' αυτές. Πώς κρίνετε το μοντέλο σας;

- Με την βοήθεια του ProQ: <http://www.sbc.su.se/~bjornw/ProQ/ProQ.cgi>  
Επικολλήστε τις συντεταγμένες του μοντέλου στο προβλεπόμενο πλαίσιο.