

## Ανάπτυξη λογισμικού για συνδυαστική κατάταξη (consensus scoring) αποτελεσμάτων μοριακής μοντελοποίησης και υπολογισμών πρόσδεσης (molecular docking) μικρών μορίων σε φαρμακευτικούς στόχους.

Διπλωματική Εργασία Βιοπληροφορικής  
Επιβλέπων: Καθηγητής Ι. Εμίρης, Τμήμα Πληροφορικής & Τηλεπικοινωνιών  
Συνεπιβλέπων: Καθηγητής Ε. Μικρός, Τμήμα Φαρμακευτικής

Ο υπολογισμός πρόσδεσης πρωτεΐνης-προσδέματος (protein-ligand docking) είναι μια τεχνική μοριακής μοντελοποίησης με σκοπό την εύρεση της θέσης και της διαμόρφωσης ενός μικρού μορίου (ligand) σε σύμπλοκο με έναν πρωτεϊνικό υποδοχέα ή ένζυμο. Τέτοιοι υπολογισμοί είναι ευρέως διαδεδομένοι στην ανακάλυψη νέων φαρμάκων, καθώς χρησιμοποιούνται για την εικονική σάρωση μεγάλων ψηφιακών βιβλιοθηκών μικρών μορίων (virtual screening) με σκοπό την επιλογή νέων ενώσεων με πιθανο φαρμακολογικό ενδιαφέρον .

Κάθε λογισμικό υπολογισμών πρόσδεσης, χρησιμοποιεί μια συνάρτηση κατάταξης (scoring function) για να ποσοτικοποιήσει την ικανότητα πρόσδεσης των διαμορφώσεων και προσανατολισμών που λαμβάνουν τα μικρά μόρια στην κοιλότητα πρόσδεσης του υποδοχέα. Κάθε scoring function έχει διαφορετική ακρίβεια για το κάθε πρωτεϊνικό υποδοχέα, το οποίο εξηγεί και την μεγάλη διαθεσιμότητα τόσο των προγραμμάτων για docking όσο και των scoring functions για επανακατάταξης (rescoring)<sup>1-3</sup> των προσανατολισμών πρόσδεσης (binding poses). Σύμφωνα τόσο με την βιβλιογραφία, όσο και με την δική μας εμπειρία, ο συνδυασμός λίγων scoring functions που αποδίδουν καλά για το υποδοχέα στόχο αυξάνει την ακρίβεια και την ευαισθησία του virtual screening<sup>4</sup>.

**Στόχος.** Η επέκταση και βελτίωση αναπτυσσόμενου προγράμματος σε γλώσσα Python για την αξιολόγηση αποτελεσμάτων που προκύπτουν από λογισμικά protein-ligand docking. Ανάλογα με το επίπεδο και τον χρόνο που διαθέτει, ο ενδιαφερόμενος θα έχει τη δυνατότητα να ασχοληθεί με ορισμένα από τις παρακάτω αντικείμενα: (1) ανάπτυξη κώδικα για την χρήση και σύγκριση πολλών διαφορετικών scoring functions· (2) κατασκευή μαθηματικών μοντέλων για (α) την περιγραφή του βέλτιστου συνδυασμού scoring functions (consensus scoring scheme) για τον υπό-μελέτη υποδοχέα, (β) βαθμονόμηση (calibration) μιας scoring function και βελτίωση της απόδοσής της για τον υπό-μελέτη υποδοχέα· (3) χρήση διαφόρων στατιστικών δεικτών για την αξιολόγηση της κάθε scoring function· (4) χρήση παράλληλου προγραμματισμού για την βελτιστοποίηση της ταχύτητας του προγράμματος· (5) ανάπτυξη κώδικα για παρουσίαση των αποτελεσμάτων σε κατανοητή μορφή καθώς και την προβολή τους σε υπάρχοντα προγράμματα μοριακής απεικόνισης. Τέλος ο υποψήφιος αναμένεται να αξιολογήσει την απόδοση των μεθόδων που θα υλοποιηθούν σε κώδικα Python πάνω σε πραγματικά δεδομένα.

Η παρούσα εργασία, ανάλογα με το αποτέλεσμα της δουλειάς, αναμένεται να οδηγήσει τουλάχιστον σε μία δημοσίευση σε διεθνές περιοδικό.

**Γνώσεις.** Απαραίτητη η γνώση αντικειμενοστραφούς προγραμματισμού, κατά προτίμηση σε Python. Επιθυμητή η γνώση αλγοριθμικών τεχνικών, κυρίως εφαρμοσμένης παλινδρόμησης και βελτιστοποίησης συναρτήσεων, καθώς και πολυπαραγοντικής στατιστικής. Επίσης επιθυμητή (αλλά όχι απαραίτητη) είναι η εμπειρία με δομές πρωτεϊνών και υπολογισμών πρόσδεσης (protein-ligand docking).

### **Βιβλιογραφικές Αναφορές:**

1. Durrant, J. D. & McCammon, J. A. NNScore: a neural-network-based scoring function for the characterization of protein-ligand complexes. *J. Chem. Inf. Model.* **50**, 1865–1871 (2010).
2. Durrant, J. D. & McCammon, J. A. NNScore 2.0: a neural-network receptor-ligand scoring function. *J. Chem. Inf. Model.* **51**, 2897–2903 (2011).
3. Neudert, G. & Klebe, G. DSX: a knowledge-based scoring function for the assessment of protein-ligand complexes. *J. Chem. Inf. Model.* **51**, 2731–2745 (2011).
4. Zhong, S., Zhang, Y. & Xiu, Z. Rescoring ligand docking poses. *Curr. Opin. Drug Discov. Devel.* **13**, 326–334 (2010).