



NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

**SCHOOL OF SCIENCES
DEPARTMENT OF INFORMATICS AND TELECOMMUNICATIONS**

**GRADUATE PROGRAM
"INFORMATION TECHNOLOGIES IN MEDICINE AND BIOLOGY"**

MSc Thesis

**Computational evaluation of potential binding molecules to
protein targets with geometric criteria**

FYTROS MARIOS

Supervisor: **Ioannis Z. Emiris, Professor**

ATHENS

FEBRUARY 2014



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
"ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ"**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Υπολογιστική αξιολόγηση της εν δυνάμει πρόσδεσης μορίων σε
πρωτεϊνικούς στόχους με γεωμετρικά κριτήρια**

Μάριος Φύτρος

Επιβλέποντες: Ιωάννης Εμίρης, Καθηγητής

ΑΘΗΝΑ

ΦΕΒΡΟΥΑΡΙΟΣ 2014

MSc Thesis

Computational evaluation of potential binding molecules to protein targets with geometric criteria

Marios A. Fytros

R.N.: PIB 050

SUPERVISOR: **Ioannis Z. Emiris**, Professor

EXAMINING COMMITTEE: **Chrysin Evangelia** , Researcher
Theoharis Theoharis, Professor

February 2014

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Υπολογιστική αξιολόγηση της εν δυνάμει πρόσδεσης μορίων σε πρωτεϊνικούς στόχους με γεωμετρικά κριτήρια

Μάριος Α. Φύτρος

A.M.: ΠΙΒ050

ΕΠΙΒΛΕΠΟΝΤΕΣ: **Ιωάννης Εμίρης**, Καθηγητής

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: **Χρυσίνα Ευαγγελία**, Ερευνήτρια
Θεοχάρης Θεοχάρης, Καθηγητής

Φεβρουάριος 2014

ABSTRACT

The docking study of small molecules, mainly in organic compounds, to cavities or to binding sites of macromolecules (proteins), with computational methods (in silico) is a fundamental question of drug design and one of the main problems in contemporary structural bioinformatics. A subset of this problem is the docking of small molecule compounds (potential drug) where is determined by the geometrical and/or chemical complementarity with the binding site (e.g. the catalytic site of the enzyme) of the receptor.

Docking can be distinguished in the following stages: (1) Identifying the binding sites (enzyme active site or other cavities of interest) of the macromolecule. The receptor is composed of amino acids that make up the binding site. (2) Searching for ligands that match with geometric criteria with the binding sites. (3) Evaluating (scoring) the energy of formation for each complex to succeed in minimization.

In this thesis, we focused on (3), developing a Protein-Ligand docking algorithm with graphical methods in order to calculate the geometrical complementarity of the two molecules. This method was based on an algorithm [68] which matches separate parts or fragments belonging to an object, using their surface geometry, and calculating a complementarity score.

In addition, we compared our method to algorithms which detect binding sites and to others existing docking software methods. In this comparison, we select 20 PDB complexes from the PLD database.

SUBJECT AREA: Protein Docking

KEYWORDS: Graphics, Enhanced Simulated Annealing, Geometry, Shape Complementarity, Binding Sites, PDB

ΠΕΡΙΛΗΨΗ

Η μελέτη του τρόπου πρόσδεσης (docking) μικρών μορίων, συνήθως οργανικών ενώσεων, σε κοιλότητες ή κέντρα σύνδεσης μακρομορίων (πρωτεΐνες), με υπολογιστικές μεθόδους (in silico) αποτελεί θεμελιώδες ερώτημα στον σχεδιασμό φαρμάκων και ένα από τα κεντρικά σύγχρονα προβλήματα στην δομική βιοπληροφορική. Η πρόσδεση της μικρομοριακής ένωσης (εν δυνάμει φάρμακο) καθορίζεται, μεταξύ άλλων, από την γεωμετρική συμπληρωματικότητα αυτής με την περιοχή σύνδεσης (π.χ. καταλυτικό κέντρο του ενζύμου).

Το πρόβλημα της πρόσδεσης διακρίνεται στα εξής στάδια: (1) Ο εντοπισμός των κέντρων σύνδεσης (ενεργό κέντρο ενζύμου ή άλλες κοιλότητες που παρουσιάζουν ενδιαφέρον) του μακρομορίου, αν δεν είναι γνωστά. Ο υποδοχέας αποτελείται από αμινοξέα που συνθέτουν το κέντρο σύνδεσης. (2) Αναζητούνται προσδέτες που ταιριάζουν γεωμετρικά στα κέντρα πρόσδεσης. (3) Αξιολογείται (scoring) η ενέργεια σχηματισμού κάθε συμπλόκου, με σκοπό την ελαχιστοποίησή της.

Σε αυτή την εργασία, εστίασαμε στο (3), αναπτύσσοντας ένα Protein-Ligand docking αλγόριθμο με μεθόδους γραφικών, προκειμένου να υπολογιστεί η γεωμετρική συμπληρωματικότητα αυτών των δύο μορίων. Η μέθοδος αυτή βασίστηκε σε έναν υπάρχον αλγόριθμο [68], ο οποίος ταιριάζει ξεχωριστά τμήματα ή θραύσματα αντικειμένων που ανήκουν σε αυτό, χρησιμοποιώντας μόνο την γεωμετρική επιφάνεια αυτών, υπολογίζοντας ένα συμπληρωματικό score.

Επιπλέον, συγκρίναμε τη μέθοδο μας σε υπάρχοντα λογισμικά που ανιχνεύουν binding sites και σε λογισμικά Protein docking. Για την σύγκριση, χρησιμοποιηθήκανε 20 PDB complex από τη βάση δεδομένων PLD.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Πρόσδεση Πρωτεϊνικών Μορίων (Protein Docking)

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Πρόσδεση Πρωτεϊνικών Μορίων, Γραφικά, Ενισχυμένη προσομοίωση απόπτωσης (Enhanced Simulated Annealing), Συμπληρωματικότητα Μέσω Γεωμετρίας, Ενεργά Κέντρα Μορίων, PDB