



NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

**SCHOOL OF SCIENCE
DEPARTMENT OF INFORMATICS AND TELECOMMUNICATION**

**INTERDISCIPLINARY POSTGRADUATE PROGRAM
"INFORMATION TECHNOLOGIES IN MEDICINE AND BIOLOGY"**

MASTER THESIS

**Automating Free Energy Perturbation Calculations
For Drug Design**

Stamatia N. Zavitsanou

Supervisor: **Dr. Zoe Cournia**, Researcher - Assistant Professor Level,
Biomedical Research Foundation of the Academy of Athens
(BRFAA)

ATHENS

JULY 2018



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΔΙΑΤΜΗΜΑΤΙΚΟ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
"ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ"**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Αυτοματοποίηση υπολογισμού ελεύθερης ενέργειας
πρόσδεσης για τον σχεδιασμό φαρμάκων**

Σταματία Ν. Ζαβιτσάνου

**Επιβλέπουσα: Δρ. Ζωή Κούρνια, Ερευνήτρια Γ', Ίδρυμα Ιατροβιολογικών
Ερευνών Ακαδημίας Αθηνών (ΙΙΒΕΑΑ)**

**ΑΘΗΝΑ
ΙΟΥΛΙΟΣ 2018**

MASTER THESIS

Free energy perturbation calculations for drug design

Stamatia N. Zavitsanou
S.N.: PIV0153

SUPERVISOR: **Dr. Zoe Cournia**, Researcher - Assistant Professor Level,
Biomedical Research Foundation of the Academy of Athens
(BRFAA)

EXAMINATION COMMITTEE: **Dr. Zoe Cournia**, Researcher - Assistant Professor Level,
Biomedical Research Foundation of the Academy of Athens
(BRFAA)

Dr. Stavros Perantonis, Research Director, Institute of
Informatics and Telecommunications, National Center of
Scientific Research "Demokritos"

Dr. Ioannis Emiris, Professor Level, National and
Kapodistrian University of Athens (NKUA), Department of
Informatics and Telecommunications (DIT)

JULY 2018

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αυτοματοποίηση υπολογισμού ελεύθερης ενέργειας πρόσδεσης για τον σχεδιασμό
φαρμάκων

Σταματία Ν. Ζαβιτσάνου

A.M.: ΤΠΙΒ0153

ΕΠΙΒΛΕΠΟΥΣΑ: **Δρ. Ζωή Κούρνια**, Ερευνήτρια Γ', Ίδρυμα
Ιατροβιολογικών Ερευνών Ακαδημίας Αθηνών (ΙΙΒΕΑΑ)

**ΕΞΕΤΑΣΤΙΚΗ
ΕΠΙΤΡΟΠΗ:**

Δρ. Ζωή Κούρνια, Ερευνήτρια Γ', Ίδρυμα
Ιατροβιολογικών Ερευνών Ακαδημίας Αθηνών (ΙΙΒΕΑΑ)

Δρ. Σταύρος Περαντώνης, Διευθυντής Έρευνας, Ινστιτούτο
Πληροφορικής και Τηλεπικοινωνιών, ΕΚΕΦΕ "Δημόκριτος".

Δρ. Ιωάννης Εμίρης, Καθηγητής, Εθνικό και Καποδιστριακό
Πανεπιστήμιο Αθηνών (ΕΚΠΑ), Τμήμα Πληροφορικής και
Τηλεπικοινωνιών (ΤΠΛ)

ΙΟΥΛΙΟΣ 2018

ABSTRACT

The advent of technological advances in the field of computer-aided drug design has streamlined the drug design process, making it more cost- and time-efficient. Nowadays rational, structure-based drug design may quantify underlying molecular interactions involved in ligand-protein binding by utilizing the 3D structure of the therapeutic target in the process. One of the most important tasks in the lead optimization phase of the drug design process is to predict, among a series of lead candidates, which ones will bind more strongly to the therapeutic target. In this direction, relative binding free energy methodologies have been developed, which rely on physics-based molecular simulations and rigorous statistical mechanics to calculate the differences in the free energy of binding between a parent candidate drug and analogues. For example, Free Energy Perturbation (FEP) calculations coupled with Molecular Dynamics (MD) simulations calculate the free energy difference between an initial (reference) and a analog (target) molecule to an average of a function of their energy difference evaluated by sampling for the initial state.

Such methodologies have shown significant potential in the lead optimization process, however, they have been limited by technical challenges such as manual creation of large numbers of input files to setup/run/analyze free energy simulations. Automating free energy perturbation calculations would streamline the use of FEP calculations and would be a step forward to delivering high throughput calculations for accurate predictions of relative binding affinities before a compound is synthesized, and consequently save enormous experimental and human resources.

In this thesis, an algorithm that automates the set up procedure for relative binding free energy simulations is described. We have created FEPprepare, the first web-based server, which automates the set-up procedure for FEP calculations within the context of NAMD, one of the major MD engines. The user has to upload the structure files to the web-server. The algorithm, which is written in Python, performs atom renaming, and partial charge redistribution. Later it uses VMD, which is a molecular viewer program that can be used to help set up NAMD simulations and to help analyse and visualize NAMD output, to generate all needed files for the calculations. After the algorithm confirms compatibility with NAMD, it provides the user with everything needed to run a simulation

Computational methods in drug design have proven very useful as they render the lead optimization process both time and cost efficient. Automation of Free Energy Perturbation calculations to provide access to large-scale simulations for lead optimization has been presented in this thesis.

SUBJECT AREA: Computational Chemistry

KEYWORDS: free energy, drug design, NAMD, dual topology file, ligands, web development

ΠΕΡΙΛΗΨΗ

Η διαδικασία σχεδιασμού φαρμάκων έχει βελτιστοποιηθεί με τη βοήθεια των ηλεκτρονικών υπολογιστών, έχοντας γίνει πιο αποδοτική από πλευράς κόστους και χρόνου. Σήμερα, χρησιμοποιώντας τη 3D δομή του θεραπευτικού στόχου ο ορθολογιστικός σχεδιασμός φαρμάκων μπορεί να ποσοτικοποιήσει τις μοριακές αλληλεπιδράσεις που εμπλέκονται στη δέσμευση συνδέτη-πρωτεΐνης. Ένα από τα πιο σημαντικά καθήκοντα στον σχεδιασμό φαρμάκων είναι να προβλέψουμε μεταξύ μιας σειράς υποψήφιων ποιά από αυτά θα δεσμευτούν καλύτερα στον θεραπευτικό στόχο. Σε αυτή την κατεύθυνση έχουν αναπτυχθεί μεθοδολογίες σχετικής δέσμευσης της ελεύθερης ενέργειας, οι οποίες βασίζονται σε μοριακές προσομοιώσεις, στη φυσική και στην αυστηρή στατιστική μηχανική για τον υπολογισμό των διαφορών στην ελεύθερη ενέργεια σύνδεσης μεταξύ ενός γονικού υποψήφιου φαρμάκου και αναλόγων του. Για παράδειγμα, οι υπολογισμοί της Ελεύθερης Ενεργειακής Διαταραχής (FEP) σε συνδυασμό με τις προσομοιώσεις Μοριακής Δυναμικής (MD) υπολογίζουν την ελεύθερη διαφορά ενέργειας μεταξύ ενός αρχικού και ενός τελικού μορίου.

Τέτοιες μεθοδολογίες έχουν δείξει σημαντικές δυνατότητες στη διαδικασία βελτιστοποίησης οδηγών μορίων, ωστόσο έχουν περιοριστεί από τεχνικές προκλήσεις όπως η χειροκίνητη δημιουργία μεγάλου αριθμού αρχείων εισόδου για την εγκατάσταση / εκτέλεση / ανάλυση ελεύθερων προσομοιώσεων ενέργειας. Η αυτοματοποίηση των υπολογισμών της διατάραξης της ελεύθερης ενέργειας θα απλοποιήσει τη χρήση των υπολογισμών FEP και είναι ένα βήμα προς τα εμπρός για την παροχή υπολογισμών υψηλής απόδοσης για ακριβείς προβλέψεις πριν από την σύνθεση μιας ένωσης και συνεπώς θα εξοικονομήσει τεράστιο χρόνο και κόστος.

Σε αυτή τη διατριβή περιγράφεται ένας αλγόριθμος που αυτοματοποιεί τη διαδικασία εγκατάστασης για σχετικές δεσμευτικές προσομοιώσεις ελεύθερης ενέργειας. Δημιουργήσαμε το FEPprepare, τον πρώτο ιστότοπο, που αυτοματοποιεί τη διαδικασία του στησίματος για υπολογισμούς FEP στο πλαίσιο του NAMD, ενός από τους σημαντικότερους μηχανισμούς MD. Ο χρήστης ανεβάζει τα αρχεία δομής πρωτεΐνης και των προσδεμάτων, ο αλγόριθμος ο οποίος είναι γραμμένος σε Python μετονομάζει τα άτομα, και αναδιανείμει τα φορτία των ατόμων. Αργότερα χρησιμοποιείται το VMD, το οποίο είναι ένα πρόγραμμα μοριακής προβολής που μπορεί να χρησιμοποιηθεί για να βοηθήσει στη δημιουργία προσομοίωσης του NAMD και να βοηθήσει στην ανάλυση και απεικόνιση των δεδομένων που παράγει το NAMD, για να παράξει τα αρχεία που χρειαζόμαστε, και να κάνει το Solvation και τον ιονισμό. Αφού ο αλγόριθμος επιβεβαιώσει ότι όλα τα αρχεία είναι συμβατά με το NAND τα παρέχει στον χρήστη .

Οι υπολογιστικές μέθοδοι στον σχεδιασμό φαρμάκων έχουν αποδειχθεί πολύ χρήσιμες καθώς κάνουν την διαδικασία της βελτιστοποίησης πολύ πιο γρήγορη και φθηνή. Στόχος μας σε αυτή την διπλωματική εργασία είναι να αυτοματοποιήσουμε τους υπολογισμούς ελεύθερης ενέργειας πρόσδεσης για να παρέχουμε πρόσβαση σε μεγάλη κλίμακα προσομοιώσεων.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Υπολογιστική Χημεία

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: ελεύθερη ενέργεια, σχεδιασμός φαρμάκων, NAMD, αρχείο διπλής τοπολογίας, ανάπτυξη διαδικτυακού εργαλείου