

# Drug Design - Εισαγωγή στην Βιοπληροφορική

Χριστίνα Κρατημένου (1067495), Ιωάννης Λουδάρος (1067400)

Ιούνιος 2022

## Περίληψη

Ο σχεδιασμός φαρμάκων, και ειδικότερα ο υπολογιστικά υποβοηθούμενος, συνίσταται από ένα ευρύ φάσμα θεωρητικών και υπολογιστικών προσεγγίσεων που είναι μέρος της σύγχρονης ανακάλυψης φαρμάκων. Οι υπολογιστικές μέθοδοι σχεδιασμού φαρμάκων, ή αλλιώς CAAD (Computer-Aided Drug Design), έχουν συνεισφέρει σημαντικά στην ανάπτυξη φαρμάκων που βρίσκονται σε κλινικές δοκιμές ή και κλινική χρήση. Τέτοιες μέθοδοι έχουν εμφανιστεί και εφαρμοστεί παράλληλα με πειραματικές προσεγγίσεις που χρησιμοποιούνται στο drug design. Διεργασίες που έχουν φανεί βοηθητικές ή και απαραίτητες στον σχεδιασμό φαρμάκων είναι η ευέλικτη στοίχιση μικρών μορίων. Αποτελεί σημαντικό στοιχείο τόσο των ligand-based μεθόδων σχεδίασης, όσο και των structure-based. Παρόλη την σημασία της διαδικασίας αυτής, η ύπαρξη πακέτων λογισμικού υψηλής ποιότητας για την στοίχιση μορίων είναι περιορισμένη. Ωστόσο, το property-based πρόγραμμα BCL::MolAlign προσφέρει ευελιξία συνδετών (ligands) μέσω ενός συνδυασμού προσχεδιασμένων conformers και επιτόπου περιστροφή δεσμού. Χρησιμοποιεί τον αλγόριθμο Monte Carlo Metropolis Sampling (MCM) για τον υπολογισμό του κοινού δομικού χώρου των δύο μορίων που ευθυγραμμίζονται, γεγονός που το κάνει αποδοτικότερο σε σχέση με άλλα πακέτα που κυκλοφορούν σε αυτό τον τομέα. Άλλη μία διαδικασία που εμφανίζεται στο drug design και τα υπολογιστικά πειράματα ανακάλυψης φαρμάκων είναι η μοριακή δυναμική (Molecular Dynamics - MD), η οποία έχει σκοπό την δομική ανάλυση μορίων και ατόμων. Η σουίτα λογισμικού Thetis αναπτύχθηκε για την διευκόλυνση της παρακολούθησης των δομικών αλλαγών κατά την διάρκεια εξομοιώσεων μοριακής δυναμικής και, σε αντίθεση με άλλα πακέτα λογισμικού, η Thetis μπορεί να χρησιμοποιηθεί σε μία σειρά δομών MD. Πακέτα λογισμικών που εξειδικεύονται στην μοντελοποίηση φαρμακοφόρων, τόσο σε επίπεδο αρχάριου, όσο και σε επίπεδο επαγγελματία είναι τα Drugster και DrugOn (Pro). Οι λειτουργίες τους και τα βήματα σχεδίασης είναι παρόμοια, ωστόσο οι μικρές διαφορές που παρατηρούνται είναι αρκετές να διαχωρίσουν δραστηρικά το ένα από το άλλο.

# Εισαγωγή

Σύμφωνα με τον πιο πρόσφατο ορισμό της IUPAC (International Union of Pure and Applied Chemistry) ένα μοντέλο φαρμακοφόρου αποτελεί ένα σύνολο από στερεοχημικά και ηλεκτρονικά χαρακτηριστικά, το οποίο είναι απαραίτητο για την εξασφάλιση των βέλτιστων υπερμοριακών αλληλεπιδράσεων με έναν ορισμένο βιολογικό στόχο, και την πυροδότηση ή παρεμπόδιση της βιολογικής του αντίδρασης.

Οι πλήρως αυτοματοποιημένες μέθοδοι σχεδιασμού μοντέλου φαρμακοφόρου, μπορούν να διευκολύνουν της διαδικασίας της σύγχρονης *in silico* ανακάλυψης φαρμάκων. Οι υπολογιστές κερδίζουν αξιοπιστία στον τομέα της υπολογιστικής βιολογίας και σχεδιασμού φαρμάκων, όσο εγκαθιδρύονται νέοι και αποδοτικότεροι αλγόριθμοι και pipelines. Με τον υπολογιστικά υποβοηθούμενο σχεδιασμό φαρμάκων να είναι ένα βασικό κομμάτι της ανακάλυψης φαρμάκων και της διαδικασίας βελτιστοποίησης, τα μοντέλα φαρμακοφόρων έχουν γίνει στοιχείο-κλειδιά για την καλύτερη κατανόηση των αλληλεπιδράσεων μεταξύ υποδοχέων και συνδετών. Συγκεκριμένα, έχουν συνεισφέρει στην εξέλιξη της διαδικασίας του σχεδιασμού φαρμάκων δίνοντας έμφαση στην τρισδιάστατη διάταξη χημικών χαρακτηριστικών.

Οι μέθοδοι της 3D μοντελοποίησης, λαμβάνουν υπόψιν τις τρισδιάστατες δομές και συνδέσεις υποδοχέων και αναστολέων, με σκοπό να αναγνωρίζουν περιοχές που είναι ευνοϊκές ή μη, για κάποια συγκεκριμένη αλληλεπίδραση υποδοχέων και αναστολέων. Τα μοντέλα αυτά, συμβάλλουν στην ανακάλυψη φαρμάκων προσφέροντας πλήθος πλεονεκτημάτων, ενώ χαρακτηρίζονται από υπολογιστική αποδοτικότητα λόγω της απλότητάς τους. Παράλληλα, είναι εμπεριστατωμένα και επεξεργάσιμα, έτσι ώστε η πληροφορία να είναι εύκολα προσβάσιμη. Τα μοντέλα φαρμακοφόρων έχουν χρησιμοποιηθεί ήδη σε μία πληθώρα μελετών, ώστε να γίνει πλήρης εκμετάλλευση των οφελών τους. Η εκμετάλλευση αυτή, οδήγησε στην ανάπτυξη αυτοματοποιημένων εργαλείων και εφαρμογών μέσα στα τελευταία 20 χρόνια.

Ωστόσο, η μοντελοποίηση δεν έχει φτάσει ακόμα την πλήρη δυνατότητα της, καθώς περιορίζεται από διάφορα εμπόδια που σχετίζονται με την συνεχή ζήτηση για μείωση του υψηλού κόστους που έχει ο σχεδιασμός και η ανακάλυψη φαρμάκων.

Για τον υπολογιστικά υποβοηθούμενο σχεδιασμό φαρμάκων χρησιμοποιούνται ορισμένες διεργασίες, οι οποίες αποτελούν και βασικά βήματα κατά την σχεδίαση. Η ευέλικτη στοίχιση μικρών μορίων, είναι η διαδικασία οργάνωσης τρισδιάστατων μοριακών δομών στον χώρο, σύμφωνα με τις ομοιότητές τους. Είναι ένα απαραίτητο βήμα στις στρατηγικές ανακάλυψης φαρμάκων, οι οποίες χρησιμοποιούν την 3D δομική πληροφορία για να εκτιμήσουν πιθανούς συνδέτες. Η στοίχιση των συνδετών είναι απαραίτητη και αποτελεί μία από τις σημαντικότερες προκλήσεις στην μοντελοποίηση φαρμακοφόρων, είτε πρόκειται για *ligand-based* μεθόδους, είτε για *structure-based* μεθόδους. Η αναγκαιότητα της διεργασίας αυτής συναντάται στο γεγονός ότι, η πληροφορία σύνδεσης για το *ligand* δεν μπορεί να καθοριστεί αξιόπιστα αν ο συνδέτης μπορεί να έχει πολλαπλές δομές με συγκρίσιμη ενέργεια.

Άλλη μία μέθοδος που χρησιμοποιείται στα υπολογιστικά πειράματα σχεδιασμού και ανακάλυψης είναι η μοριακή δυναμική, η οποία χρησιμοποιείται για την προσομοίωση υπολογιστή, με σκοπό την ανάλυση των φυσικών κινήσεων ατόμων και μορίων. Επιτρέπει την μελέτη των πολύπλοκων δυναμικών διεργασιών που λαμβάνουν μέρος σε ένα βιολογικό σύστημα, και υπολογίζει την εξαρτώμενη από τον χρόνο συμπεριφορά μοριακών συστημάτων. Οι εξομοιώσεις μοριακής δυναμικής μπορούν να χρησιμοποιηθούν σε μία πληθώρα βιολογικών διαδικασιών, όπως είναι η αναδίπλωση πρωτεϊνών, η μοριακή αναγνώριση σε πρωτεΐνες, ενώ έχουν εφαρμογή στα υπολογιστικά πειράματα φαρμάκων. Γνωρίζοντας πως οι τροχιές των μοριακών δυναμικών είναι εξαιρετικά πολύπλοκες συναρτήσεις των πρωτεϊνών και του περιβάλλοντος, η σύγκριση διαφορετικών τροχιών -ακόμα και κάτω από τις ίδιες συνθήκες- δεν είναι σαφής. Έχουν προταθεί πολλές μέθοδοι για την σύγκριση αυτή, με τις απλότερες να βασίζονται στην γεωμετρία και με τις πιο σύνθετες να βασίζονται στις διακυμάνσεις του χρόνου των διάφορων ιδιοτήτων του συστήματος κατά την εξομοίωση.

# Pharmacophore Modeling

Κατά την διάρκεια των τελευταίων χρόνων, η μοντελοποίηση φαρμακοφόρων έχει γίνει αναπόσπαστο κομμάτι της υπολογιστικά υποβοηθούμενης σχεδίασης και ανακάλυψης φαρμάκων. Τα συνεχώς αυξανόμενα οφέλη της μοντελοποίησης οδήγησαν στην ανάπτυξη ολοένα και περισσότερων υπολογιστικών εργαλείων και πακέτων λογισμικού, με σκοπό την καλύτερη προσέγγιση του ζητήματος.

Το Drugster είναι ένα pipeline, πλήρως διαδραστικό, σχεδιασμένο με σκοπό γνώμονα την εμπειρία του χρήστη. Συντελείται από ένα φιλικό προς τον χρήστη περιβάλλον, που του επιτρέπει να εκτελέσει, καθοδηγούμενος σε κάθε βήμα, την διαδικασία του σχεδιασμού φαρμάκων μέσα από έναν αποδοτικό συνδυασμό των εργαλείων PDB2PQR, Ligbuilder, Gromacs, Dock suites.

Λίγα λόγια για το καθένα:

- *PDB2PQR*: Το εργαλείο αυτό προετοιμάζει τις δομές για επιπλέον υπολογισμούς, επανακατασκευάζοντας άτομα που λείπουν, προσθέτοντας υδρογόνα, αναθέτοντας ατομικά φορτία και ακτίνες από ορισμένα δυναμικά πεδία και τέλος, δημιουργώντας αρχεία PQR.
- *Ligbuilder*: Χρησιμοποιείται για de novo drug design, δηλαδή τον σχεδιασμό νέων χημικών οντοτήτων που ανταποκρίνονται σε ένα σύνολο περιορισμών. Το Ligbuilder χρησιμοποιεί γενετικό αλγόριθμο για να κατασκευάσει συνδέτες (ligands) επαναληπτικά, με την χρήση βιβλιοθήκης οργανικών θραυσμάτων. Παράλληλα, μέσω μίας ποικιλίας μεθόδων, χειρίζεται και μεταβάλλει την μοριακή δομή.
- *Gromacs*: Είναι ένα από τα πιο διαδεδομένα εργαλεία στην χημεία και χρησιμοποιείται κυρίως για δυναμικές εξομοιώσεις μορίων. Παρέχει μία πληθώρα τύπων εργαλείων για υπολογισμό, προετοιμασία και ανάλυση.
- *Dock*: Χρησιμοποιείται για τα απαραίτητα docking simulations, με τα οποία γίνεται πρόβλεψη για τον ιδανικό προσανατολισμό του ligand στην ενεργή κατάσταση του υποδοχέα, όταν συνδέονται για να σχηματίσουν ένα σταθερό σύμπλεγμα. Παράλληλα, περιέχει λειτουργίες βαθμολόγησης και κατάταξης των ενώσεων.

Η πλατφόρμα Drugster είναι Perl-Tcl-Tk based και συνδυάζει τους παραπάνω αλγορίθμους προσφέροντας ένα κατανοητό και εύκολα διαχειρίσιμο περιβάλλον χρήσης. Σχεδιάστηκε για να

αυτοματοποιήσει την πλήρη εργασία της ρύθμισης του σχεδιασμού και των πειραμάτων βελτιστοποίησης. Το κύριο παράθυρο της εφαρμογής αποτελείται από διεπαφή με μενού, καθώς και ευκολονόητο, βήμα προς βήμα layout. Ο χρήστης έχει στην διάθεσή του ένα παράθυρο διεργασιών για να παρακολουθεί τους ενεργούς υπολογισμούς σε πραγματικό χρόνο και ένα παράθυρο που λειτουργεί ως γραμμή εντολών. Η διαδικασία χρήσης της σουίτας για σχεδιασμό φαρμάκων και πειράματα δομής, αποτελείται από πέντε βήματα.

Το πρώτο βήμα που ακολουθείται, είναι κάτι που απουσιάζει από αρκετά μεγάλα πακέτα λογισμικού και επιτρώνεται στην διόρθωση των κοινών ζητημάτων που παρουσιάζονται στα αρχεία PDB (Program Database). Κατά την διαδικασία αυτή, γίνεται πρόσθεση των υδρογόνων που λείπουν, υπολογίζονται οι μερικές αλλαγές, μπορεί να γίνει αφαίρεση των ετεροατόμων κ.α. Στη συνέχεια, με την χρήση της ευέλικτης σουίτας Gromacs, ο υποδοχέας ελαχιστοποιείται ενεργειακά πριν το de novo πείραμα σχεδιασμού φάρμακου. Έτσι, διορθώνεται ένα από τα μεγαλύτερα μειονεκτήματα των structure-based αλγορίθμων: η έλλειψη διαμορφωτικής βελτιστοποίησης του υποδοχέα. Αυτό το αυτοματοποιημένο βήμα, αντιμετωπίζει πολλές ασυνέπειες που μπορεί να προκύπτουν από την αφαίρεση ετεροατόμων, χωρίς να γίνεται επαναφορά στην 'relaxed' διαμόρφωση του PDB αρχείου του υποδοχέα, το οποίο θα χρησιμοποιηθεί για την γενία των νέων ligands.

Στο επόμενο βήμα, λαμβάνει μέρος ο πραγματικός de novo δομικά βασισμένος σχεδιασμός των νέων δομών ligands. Δίνεται στον χρήστη η δυνατότητα να παραμετροποιήσει πλήρως την διαδικασία με την συμβολή του Ligbuilder v1.2 και v2.0. Η διεργασία οργανώνεται σε τρεις, πλήρως προσαρμοζόμενες από τον χρήστη, φάσεις. Αρχικά, προετοιμάζεται ένας φαρμακοφόρος που, με την σειρά του προετοιμάζει και συνοψίζει τις τρισδιάστατες ιδιότητες του 'σκελετού', βασικές δομές που αργότερα θα παραχθούν και θα αναλυθούν. Έπειτα ο χρήστης έχει την επιλογή ανάμεσα στους growing ή τους linking αλγόριθμους του Ligbuilder. Ο συνδυασμός μοριακών θραυσμάτων αρχίζει αυτόματα όταν ο χρήστης έχει ολοκληρώσει την ρύθμιση των παραμέτρων. Η τρίτη φάση είναι η συνάρτηση αξιολόγησης ενώσεων, κατά την οποία τα ανώτερα μόρια επιλέγονται για το επόμενο βήμα. Όλα τα υποψήφια μόρια ligand που προετοιμάστηκαν στο προηγούμενο βήμα υποβάλλονται σε docking simulations με την χρήση του Dock, τα οποία ακολουθούνται από ενεργειακές ελαχιστοποιήσεις μέσα στον υποδοχέα. Έστω, τα μόρια ligand επαναβαθμολογούνται και επανακατατάσσονται, με την προσέγγιση της επαναβαθμολόγησης να βασίζεται στο *free energy perturbation* και στην ανάλυση της αλληλεπίδρασης μεταξύ ένωσης και υποδοχέα. Οι συναρτήσεις βαθμολόγησης που περιλαμβάνονται στο Dock είναι γρήγορες και ευέλικτες, προσφέροντας ένα σύνολο από εργαλεία για βαθμολόγηση και επανακατάταξη των υποψήφιων ενώσεων. Τέλος, οι τοπολογίες των ligands μπορούν είτε να ανατεθούν αυτόματα, είτε από τον χρήστη με την βοήθεια εξειδικευμένου λογισμικού.

Το τελευταίο βήμα του pipeline του Drugster είναι η αυτόματη εισαγωγή των υποψήφιων ενώσεων με το μεγαλύτερο σκορ στην μηχανή μοριακής δυναμικής του Dock. Οι επιλογές περιλαμβάνουν: έναν επιλογέα δυναμικού πεδίου, γεννήτρια για τα αρχεία τοπολογίας, την ρύθμιση ενός περιοδικού συστήματος, διαλύτωση, καθώς και την εξομοίωση μοριακής δυναμικής (MD). Τελικά, εισαγεται και μία ανάλυση τροχιών μετά την εξομοίωση MD, ώστε να βοηθήσει τον χρήστη να επισπεύσει την διαδικασία επιλογής υποψήφιων ενώσεων.

Άλλη μία πλατφόρμα που μπορεί να προσφέρει αποδοτική και φιλική στον χρήστη σχεδίαση και μοντελοποίηση φαρμακοφόρων είναι το DrugOn. Είναι ένα πλήρως διαδραστικό pipeline, σχεδιασμένο ώστε να εκμεταλλεύεται όλα τα οφέλη του σύγχρονου προγραμματισμού. Παρέχει δύο ξεχωριστά interfaces, ένα για τον αρχάριο χρήστη και ένα για τον έμπειρο, και συνδυάζει τις σουίτες PDB2PQR, Ligbuilder, Gromacs και PharmACOPhore. Όπως και το Drugster, το DrugOn αναπτύχθηκε σε Perl-Tcl-Tk και προσφέρει καθοδήγηση σε κάθε βήμα της διαδικα-

σίας. Με σκοπό να εξομαλύνει το task της τρισδιάστατης βελτιστοποίησης δομής, ενσωματώνει την πλατφόρμα Rymol. Έτσι, προσφέρεται στον χρήστη ένας 3D μοριακός viewer, ο οποίος αποτελεί σημείο ενδιαφέροντος στην σύγχρονη ανακάλυψη φαρμάκων και στον υπολογιστικά υποβοηθούμενο σχεδιασμό φαρμάκων. Με γνώμονα την δημιουργία ενός περιβάλλοντος επιστημονικά βάσιμο, που παραμένει εύκολο στην χρήση και στην κατανόηση, αναπτύχθηκαν τα ξεχωριστά interfaces, DrugOn και DrugOn Pro. Το DrugOn απευθύνεται στον αρχάριο χρήστη και παρέχει μία αυτοματοποιημένη πλατφόρμα, η οποία θα του επιτρέψει να μάθει εκτελώντας απλά πειράματα και να εξασκηθεί στο CADD, βάζοντας σε χρήση της ιδέες του, χωρίς την ανάγκη εμπειρίας. Από την άλλη, το DrugOn Pro έχει ένα πλήρως integrated interface με όλη την παραμετροποίηση που χρειάζεται ένας επαγγελματίας. Συγκεκριμένα, αντιμετωπίζει όλα τα συχνά προβλήματα που προκύπτουν με το formatting των PDB αρχείων, προχωρά σε δομική βελτιστοποίηση του υποδοχέα με την χρήση του Gromacs, χρησιμοποιεί την σουίτα Ligbuilder για την παραγωγή νέων μορίων ή την βελτίωση ήδη κατασκευασμένων ενώσεων και τέλος, το πρόγραμμα PharmACOPhore χρησιμοποιείται για την σύζευξη των ligands, καταλήγοντας στην κατασκευή ενός τρισδιάστατου μοντέλου φαρμακοφόρου.

Το κύριο παράθυρο του DrugOn συντελείται από χαρακτηριστικά που το κάνουν ευκολόχρηστο σε κάθε βήμα. Η διαδικασία για ένα πείραμα μοντελοποίησης φαρμακοφόρου έχει τέσσερα βήματα. Αρχικά, όμοια με το Drugster, ελέγχονται όλα τα PDB αρχεία, ώστε κάθε ζήτημα συμβατότητας να αντιμετωπιστεί και να υπολογιστούν οι βασικές χημικές πληροφορίες πριν την αρχή του πειράματος. Για την διευκόλυνση του άπειρου χρήστη υπάρχουν οι προεπιλογές για την απομάκρυνση των ετεροατόμων, το δυναμικό πεδίο και την εξουδετέρωση ή όχι μερικών στοιχείων της πρωτεΐνης. Έτσι, το μόνο που έχει να κάνει ο χρήστης είναι να επιλέξει το PDB αρχείο εισόδου, όπως και την τοποθεσία και το όνομα του PDB αρχείο εξόδου. Όλα τα παραπάνω επεξεργάζονται με τον αλγόριθμο PDB2PQR. Το δεύτερο αυτοματοποιημένο βήμα αφορά την βελτιστοποίηση της διαμόρφωσης του υποδοχέα, ένα θέμα που αποτελεί μεγάλο μειονέκτημα πολλών structure-based αλγόριθμους. Με την χρήση του Gromacs, ο υποδοχέας βελτιστοποιείται μέσω ενεργειακής ελαχιστοποίησης πριν το πείραμα. Σε αυτό το βήμα οι διαθέσιμες επιλογές του χρήστη είναι τα PDB αρχεία εισόδου και εξόδου.

Στη συνέχεια, γίνεται το πραγματικό structure-based drug design των νέων δομών συνδετών, όπου ο χρήστης μπορεί να παραμετροποιήσει πλήρως την διαδικασία μέσω του Ligbuilder v1.2. Τα υποψήφια φάρμακα που δεν συμμορφώνονται με τα κριτήρια του χρήστη αποκλείονται, με την χρήση ενός φίλτρου αποκοπής ομοιότητας, το οποίο είναι και αυτό ρυθμιζόμενο από τον χρήστη. Το τελικό στάδιο του DrugOn pipeline είναι η αυτόματη στοίχιση δομής των μορίων που παράχθηκαν προηγουμένως. Η similarity-based συνάρτηση βαθμολόγησης ρυθμισμένη για ligand-based pose prediction, συνδυάζεται με έναν υβριδικό ant colony optimization αλγόριθμο μέσω της σουίτας PharmACOPhore. Ως αποτέλεσμα, δημιουργείται ένα πλήρως παραμετροποιημένο configuration αρχείο για να εξυπηρετήσει τα πειράματα μοντελοποίησης φαρμακοφόρου.

Το DrugOn Pro, από την άλλη, είναι μία εμπειριστατωμένη και σε βάθος προσέγγιση, στοχευμένη για ειδικούς και επαγγελματίες χρήστες. Τα βήματα που ακολουθούνται είναι τα ίδια τέσσερα με το DrugOn, με μερικές διαφορές στην παραμετροποίηση. Στο πρώτο στάδιο, αντίθετα με το DrugOn μπορεί να παραμετροποιήσει πλήρως το PDB αρχείο εισόδου, ξανά με την χρήση του αλγορίθμου PDB2PQR. Στο επόμενο βήμα, ο χρήστης έχει την επιλογή να χρησιμοποιήσει τις default παραμέτρους ή να τις προσαρμόσει πλήρως για το πείραμα. Μερικές από τις παραμέτρους παρουσιάζονται παρακάτω.

- Force Field

- Box Type
- Sol-Box Distance
- Solvate Protein in Water
- Water Model
- Neutralize System
- Remove Water/Ions from PDB File
- Show Resulting Energy Graph
- Parameter File

Ύστερα, για την παραμετροποίηση της ligand building διαδικασίας, ο χρήστης μπορεί να χρησιμοποιήσει, πέρα από το Ligbuilder v1.2, το Ligbuilder v2.0. Η πρώτη έκδοση του εργαλείου χρησιμοποιείται ακριβώς όπως στο DrugOn, ενώ στην δεύτερη έκδοση υπάρχει μεγαλύτερη αυτοματοποίηση σε ορισμένες διεργασίες. Όμοια με το Drugster το βήμα αυτό χωρίζεται σε τρεις διαφορετικές φάσεις, με τελευταία φάση την επιλογή των ανώτερων μορίων για το επόμενο και τελευταίο βήμα. Στο τελευταίο στάδιο, ο χρήστης εκτός από τις βασικές επιλογές που δίνονται και στην πιο απλοποιημένη έκδοση, έχει επίσης την δυνατότητα να δημιουργήσει ή να επεξεργαστεί το δικό του αρχείο configuration με τις παραμέτρους που χρειάζονται σε κάθε πείραμα.



# Molecular Dynamics

Η μοριακή δυναμική είναι μια μέθοδος προσομοίωσης υπολογιστή για την ανάλυση των φυσικών κινήσεων ατόμων και μορίων. Τα άτομα και τα μόρια αφήνονται να αλληλεπιδράσουν για ένα καθορισμένο χρονικό διάστημα, δίνοντας μια εικόνα της δυναμικής «εξέλιξης» του συστήματος. Οι εξομοιώσεις μοριακής δυναμικής (MD) δίνουν το μέσο για την διεξαγωγή μελετών βελτίωσης δομής, όπως και χρησιμοποιούνται ως εργαλείο κατά την εκτέλεση πειραμάτων σχεδιασμού φαρμάκων. Ενώ υπάρχουν πακέτα λογισμικού που επικεντρώνονται στην ανάλυση τροχιάς, δεν έχουν αναπτυχθεί προγράμματα που δίνουν τη δυνατότητα παρακολούθησης των αλλαγών στην διάταξη κατά την διάρκεια MD εξομοιώσεων.

Η σουίτα Thetis, έχει σχεδιαστεί με τον σκοπό να γίνεται monitoring στις δομικές αλλαγές μέσα σε μία MD εξομοίωση και προσφέρει εκτεταμένες ικανότητες για λεπτομερή conformational ανάλυση. Έχει αναπτυχθεί σε Basic και διαχειρίζεται με απλά pre-formated αρχεία κειμένου μέσα από ένα σύνολο από keywords. Οι παράμετροι που χρησιμοποιούνται από το πρόγραμμα μπορούν να τροποποιηθούν πλήρως για να συναντήσουν τις ανάγκες του εκάστοτε πειράματος και υπάρχει ευελιξία στον τύπο και την ποσότητα της εξόδου που δημιουργεί, ενώ μπορεί να διαβάσει δεδομένα συντεταγμένων πρωτεϊνικών δομών σε PDB και ENT format. Το πρόγραμμα μπορεί να λειτουργήσει σε δύο modes, είτε ανάγνωση, ανάλυση και οπτικοποίηση δομών κατά μόνας, είτε batch processing σε σύνολα δομών για μεγάλης κλίμακας ανάλυση. Επίσης, δύο ξεχωριστοί τρόποι λειτουργίας του προγράμματος αφορούν την αποθήκευση των αποτελεσμάτων που προκύπτουν από μεγάλους MD υπολογισμούς. Κατά τον πρώτο τρόπο, απαιτείται μεγάλη μνήμη, καθώς κάθε αρχείο coordinates προσπελαύνεται μία φορά και όλη η πληροφορία αποθηκεύεται μόνιμα στον υπολογιστή (Fast Implementation). Από την άλλη, στη δεύτερη επιλογή επαρκεί η μικρή μνήμη, αφού κάθε κομμάτι δεδομένων διαβάζεται μόνο όταν χρειάζεται και αμέσως μετά απομακρύνεται από την μνήμη του υπολογιστή.

Έχοντας αναγνωρίσει τους δεσμούς υδρογόνου, η Thetis προχωρά στην ανάλυση ξεχωριστών υπολειμμάτων σε έναν δευτερεύοντα τύπο δομής, ταιριάζοντας hydrogen-bonding μοτίβα σε αυτά που είναι χαρακτηριστικές ιδανικές δευτερεύουσες δομές. Οι τρεις κύριοι τύποι helical secondary structure που αναλύονται από το πρόγραμμα είναι οι  $\alpha$ ,  $\beta$ ,  $3_{10}$  έλικες. Ακολουθώντας τους ορισμούς Kabsh-Sander ή και Ramachandran, αυτά τα μοτίβα μπορούν να περιγραφούν από απλές λογικές συνθήκες που εκφράζονται σύμφωνα με το μητρώο συνδεσιμότητας hydrogen bond. Εφόσον το πρωτεϊνικό σύστημα έχει φορτωθεί και αναλυθεί, οι πιο αντιπροσωπευτικοί υπολογισμοί που μπορούν να γίνουν με το Thetis, παρουσιάζονται παρακάτω.

- Αναγνώριση των  $\alpha$ ,  $\beta$ ,  $3_{10}$  έλικων χρησιμοποιώντας ως κριτήριο το H bonding pattern της πολυπεπτιδικής αλυσίδας.
- Δημιουργία ενός αρχείου .txt, το οποίο περιέχει όλα τα  $\phi/\psi$ ,  $\omega$  conformational angles για κάθε αμινοξύ κάθε εισόδου MD.

- Δομική ανάλυση του *a*-helical conformation κάθε υπολείμματος, σύμφωνα με τα κριτήρια του Ramachandran.
- Διαγράμματα Ramachandran μπορούν να σχεδιαστούν ειδικά για ένα καθορισμένο από τον χρήστη σύνολο αρχείων. Τα διαγράμματα αυτά, είναι ένας τρόπος οπτικοποίησης ενεργητικά επιτρεπόμενων περιοχών για τις διεδρικές γωνίες ραχοκοκαλιάς  $\psi$  έναντι  $\phi$  των υπολειμμάτων αμινοξέων στη δομή πρωτεϊνών.
- Μπορεί να παραχθεί η απόκλιση ρίζας-μέσου-τετραγώνου των ατομικών θέσεων (μέτρο της μέσης απόστασης μεταξύ των ατόμων των υπερτιθέμενων πρωτεϊνών) για τον άλφα άνθρακα, για ένα user-defined σύνολο υπολειμμάτων και χρονικό πλαίσιο. Η διαφορά μεταξύ δύο δομών της ίδιας πρωτεΐνης μπορεί να προσεγγιστεί, μετρώντας την σχετική απόσταση των αντίστοιχων ατόμων. Όταν αφαιρεθεί η συνεισφορά από την μετάφραση και την περιστροφή και οι απόλυτες τιμές είναι scaled σύμφωνα με τον συνολικό αριθμό ατόμων, τότε οι αποστάσεις αυτές γίνονται η απόκλιση ρίζας-μέσου-τετραγώνου των ατομικών θέσεων (RMSD).
- Αφού όλες οι RMS αποκλίσεις μεταξύ του δοσμένου residue-set έχουν υπολογιστεί, ένα μητρώο C-alpha RMSD κατασκευάζεται, έτσι ώστε να συγκριθούν οι διάφορες MD δομές μεταξύ τους.

Με χρήση των παραπάνω, ακόμη και οι πιο αμελητέες αλλαγές που μπορεί να προκύψουν σε ένα *a*-helical conformation κατά την διάρκεια μίας εξομώσεως μοριακής δυναμικής, μπορούν να εντοπιστούν. Η μέθοδος της εξομώσεως μοριακής δυναμικής εφαρμόζεται, όπως αναφέρθηκε προηγουμένως, στο τελευταίο βήμα του pipeline της σουίτας Drugster.

Όλες οι έξοδοι που παράγονται από τη Thetis είναι απλά αρχεία κειμένου, εκτός από τα αρχεία εικόνων.

# Small Molecule Flexible Alignment

Η ευέλικτη ευθυγράμμιση μικρών μορίων είναι ένα σημαντικό βήμα κατά την διαδικασία του computer-aided drug design και αφορά την οργάνωση τρισδιάστατων μοριακών δομών στον χώρο, σύμφωνα με τις ομοιότητές τους, με σκοπό την εύρεση των δομικών απαιτήσεων για μία βιολογική ενέργεια. Ο καθορισμός του πιο πιθανού binding-pose συνδέτη, συντελεί σημαντικό στοιχείο της ανακάλυψης φαρμάκων (ligand/structure-based). Ένα από τα σημαντικότερα ζητήματα στην μοντελοποίηση φαρμακοφόρων είναι η απόκτηση μίας ακριβούς και κατατοπιστικής μοριακής στοίχισης. Το αποτελεσματικό docking πρωτεΐνης και συνδέτη, συνήθως απαιτεί εκ των προτέρων γνώση για ένα προσεγγιστικό binding mode.

Ο αλγόριθμος BCL::MolAlign της σουίτας μοριακής μοντελοποίησης BIOLOGY AND CHEMISTRY LIBRARY, αφορά το απαιτητικό πρόβλημα του small molecule flexible alignment. Ένας επιτυχής αλγόριθμος ευθυγράμμισης πρέπει να παρέχει ικανοποιητική δειγματοληψία του conformational space κάθε μορίου, ικανοποιητική δειγματοληψία πιθανών στοίχισεων και βαθμολόγηση των aligned poses σύμφωνα με το ταίριασμά τους. Αλγόριθμοι και πακέτα λογισμικού που κυκλοφορούν για τον ίδιο σκοπό, φαίνεται να χρησιμοποιούν μεθόδους που καταλήγουν σε μη ρεαλιστικά ligand poses σε περιπτώσεις που υπάρχει δραστηκή διαφορά στα μεγέθη ή τα σχήματα των μορίων προς σύγκριση. Ο BCL::MolAlign, για την αντιμετώπιση του ζητήματος αυτού ενσωματώνει έναν συνδυασμό δειγματοληψίας της γωνίας του δεσμού και *rigid-body alignment* με ένα σύνολο μοριακών conformers.

Αρχικά, χρησιμοποιείται το BCL::Conf για την δημιουργία ενός σετ από conformers για το ένα ή και τα δύο μόρια. Το σχέδιο βαθμολόγησης που ακολουθεί το BCL::Conf, είναι ικανό να ανακτήσει περισσότερο native-like conformers σε σχέση με άλλα πρωτόκολλα. Στη συνέχεια, γίνεται περιορισμένη βελτίωση στον conformer-στόχο κατά την διάρκεια του *pose sampling*. Μία επιπλέον πρόκληση είναι η ανάπτυξη ενός εύρωστου αλγορίθμου αναζήτησης, ο οποίος θα πλοηγείται στο κοινό conformational space των μορίων (co-space) που πρόκειται να ευθυγραμμιστούν. Τα περισσότερα προγράμματα χρησιμοποιούν ντετερμινιστικό αλγόριθμο βασισμένο στην μέγιστη επικάλυψη του μοριακού όγκου. Μία τέτοια προσέγγιση, αν και γρήγορη, γίνεται λιγότερο αποτελεσματική όταν αυξάνεται ο αριθμός των *rotatable* δεσμών στο στοχευμένο μόριο. Αυτή η έλλειψη, αντιμετωπίζεται από το BCL::MolAlign με την χρήση δειγματοληψίας Monte Carlo Metropolis (multi-trajectory), ώστε να γίνει επικάλυψη των υποδομών του μορίου. Η μέθοδος αυτή επιτρέπει γρήγορη σύγκλιση στο (co-space) των μορίων, ενώ διατηρεί δυναμικό conformational sampling. Τελικά, χρειάζεται μία μετρική βαθμολόγησης που είναι ικανή να κατατάξει μοριακές υπερθέσεις, στον βαθμό όπου χημικά παρόμοιες ομάδες είναι καλύτερα superimposed. Αρκετοί αλγόριθμοι υλοποιούν έναν συντελεστή Tanimoto για να αξιολογήσουν

χημική ή και σχηματική ομοιότητα, μέθοδος η οποία δεν αποδίδει όταν τα μόρια που συγκρίνονται είναι ικανοποιητικά διαφορετικού μεγέθους. Άλλοι αλγόριθμοι χρησιμοποιούν Γκαουσιανή επικάλυψη, η οποία όταν πρόκειται για συνεχείς ιδιότητες χωλαίνει. Επίσης, οι Γκαουσιανές μέθοδοι τυπικά καθορίζουν μία κλίμακα μήκους για κάθε ιδιότητα, το οποίο είναι αυθαίρετο και ακατάλληλο για διαφορετικά επίπεδα flexibility. Μία εναλλακτική προσέγγιση είναι να δημιουργηθεί μία συνάρτηση σύγκρισης από weighted γραμμικούς συνδυασμούς χημικών ιδιοτήτων. Έτσι, η μέθοδος που χρησιμοποιείται από το BCL::MolAlign εφαρμόζει το παραπάνω και η συνάρτηση σύγκρισης που υλοποιείται χρησιμοποιώντας την απόσταση μεταξύ κοντινότερων γειτονικών ατόμων των μορίων προς στοίχιση. Η διαδικασία αναζήτησης του, βασίζεται κυρίως στον αλγόριθμο Monte Carlo Metropolis (MCM), κατά την οποία πραγματοποιείται τουλάχιστον ένα MC trajectory για κάθε στοίχιση, με την επιλογή της διευκρίνισης επιπρόσθετων ανεξάρτητων trajectories. Κάθε ένα θα περάσει από τρία επίπεδα βελτιστοποίησης, τα οποία σταδιακά ελέγχουν το σκορ κάθε conformer pair και αφαιρούν αυτά με την χαμηλότερη βαθμολογία. Κάθε βήμα του MCM βαθμολογείται και εάν το σκορ είναι το καλύτερο που έχει γίνει sample μέχρι εκείνο το βήμα, ή εάν έχει βελτιωθεί από το προηγούμενο αποδεκτό βήμα, τότε γίνεται αποδεκτό και αυτό.

Στη διαδικασία αναζήτησης, ενσωματώνεται και η διαδικασία του conformational sampling μέσα από έναν συνδυασμό ανταλλαγής conformers και περιστροφή δεσμού. Κατά το conformer swapping, επιλέγεται και αντικαθιστάται ένας conformer με έναν από αυτούς που δημιουργήθηκαν στην αρχή της διαδικασίας. Το BondRotate (μέθοδος που χρησιμοποιεί το BCL::MolAlign κατά την δειγματοληψία) περιστρέφει μη κυκλικές, non-conjugated μονές ενώσεις ανάμεσα σε βαρέα άτομα (οποιοδήποτε άτομο εκτός του υδρογόνου) που σχηματίζουν διέδρες γωνίες με αντίστοιχα άτομα. Ο σκοπός του είναι να επιτρέψει την διόρθωση των αλλιώς στοιχισμένων ικανοποιητικά conformers, όταν υπάρχει μικρή πιθανότητα για την σωστή υποκατάσταση, λόγω του περιορισμένου conformational space. Παράλληλα, το BCL::MolAlign κάνει iterative δειγματοληψία ευθυγραμμίσεων μέσα από την υπέρθεση των ατόμων σε ένωση. Χρησιμοποιεί μία σειρά από εργαλεία για την πραγματοποίηση του βήματος αυτού (BondAlign, BondAlign2, MatchAtomNeighbors), τα οποία προσφέρουν αποτελέσματα υψηλής ποιότητας. Το σύστημα βαθμολόγησης χρησιμοποιεί την Ευκλείδεια απόσταση σε συνδυασμό με μία τιμή ιδιότητας ως πρόσθετη διάσταση, ενώ επεκτείνεται και στον υπολογισμό των weighted atom distances μεταξύ ζεύγων ατόμων. Τέλος, μέσα από ενδεδειγμένα πειράματα παρατηρήθηκε πως η πλατφόρμα BCL::MolAlign βελτιώνει την ανάκτηση των χρυσταλογραφικά καθορισμένων ligand binding poses, πως το σκορ ευθυγράμμισης μπορεί να προβλέψει την βιοδραστικότητα σε ένα δείγμα συνόλων από το Directory of Useful Decoys και πως όταν το starting pose είναι διαθέσιμο για ένα σύμπλεγμα πρωτεϊνών-ligands, είναι ικανό να αναγνωρίσει τα native-like binding poses.

## Σημείωση

Στο τέλος παρουσιάζονται μερικά σχήματα που περιγράφουν τις οθόνες των Drugster και DrugOn (Pro), το διάγραμμα ροής της λειτουργίας του Drugster και του BCL::MolAlign, ενώ δίνεται και ένα επεξηματικό σχήμα για τις τεχνικές δειγματοληψίας που χρησιμοποιεί το τελευταίο. Τα διαγράμματα ροής έχουν αναλυθεί στην αντίστοιχη ενότητα, όπως επίσης έχει γίνει και περιγραφή των οθονών. Οι εικόνες συμπεριλαμβάνονται για την αποτελεσματικότερη κατανόηση των χειμένων, καθώς πρόκειται για προγράμματα και αλγόριθμους με πολύπλοκη λειτουργία. Η βιβλιογραφία που έχει σημειωθεί παρακάτω περιέχει, εκτός των βασικών δοσμένων χειμένων, τα κείμενα που βοήθησαν στη διαδικασία συγγραφής της αναφοράς.

# Βιβλιογραφία

Dimitrios Vlachakis, Paraskevas Fakourelis, Vasileios Megalooikonomou, Christos Makris & Sophia Kossida, *DrugOn: a fully integrated pharmacophore modeling and structure optimization toolkit*. [Online Document], 2015. Available: PubMed Central [Accessed: March 2022]

Dimitrios Vlachakis, Dimosthenis Tsagrasoulis, Vasileios Megalooikonomou & Sophia Kossida, *Introducing Drugster: a comprehensive and fully integrated drug design, lead and structure optimization toolkit*. [Online Document], 2015. Available: Oxford Academic [Accessed: March 2022]

Dimitrios Vlachakis, Dimosthenis Tsagrasoulis, Vasileios Megalooikonomou & Sophia Kossida, *Introducing Drugster: a comprehensive and fully integrated drug design, lead and structure optimization toolkit*. [Online Document], 2012. Available: Oxford Academic [Accessed: March 2022]

Benjamin P. Brown, Jeffrey Mendenhall & Jens Meiler, *BCL::MolAlign: Three-Dimensional Small Molecule Alignment for Pharmacophore Mapping*. [Online Document], 2019. Available: PubMed Central [Accessed: March 2022]

Eleni Picasi, Athanasios Tartas, Vasileios Megalooikonomou & Dimitrios Vlachakis, *Introducing Thetis: a comprehensive suite for event detection in molecular dynamics*. [Online Document], 2018. Available: PubMed Central [Accessed: March 2022]

Fernando D. Prieto-Martínez, Edgar López-López, K. Eurídice Juárez-Mercado & José L. Medina-Franco, *Chapter 2 - Computational Drug Design Methods—Current and Future Perspectives*. [Online Document], 2019. Available: ScienceDirect [Accessed: May 2022]

R. K. Pathria & Paul D. Beale, *16 - Computer Simulations*. [Online Document], 2011. Available: ScienceDirect [Accessed: May 2022]

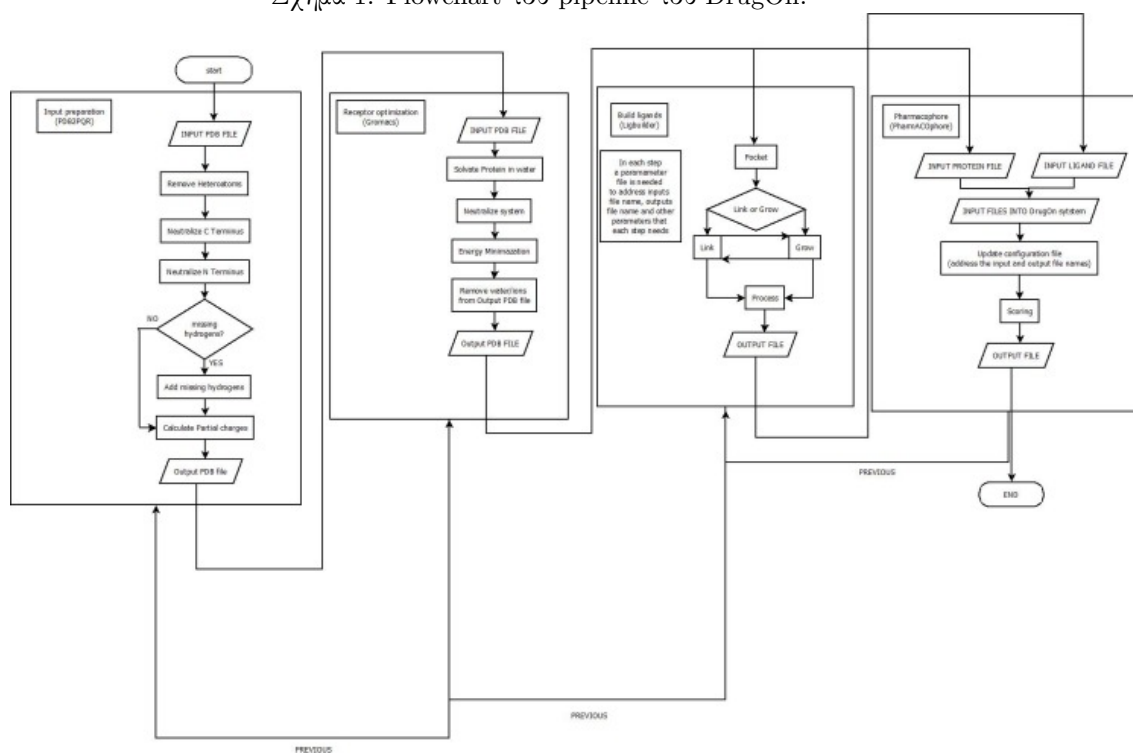
Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G & Baker NA, *PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations*. [Online Document], 2007. Available: PubMed [Accessed: June 2022]

Yaxia Yuan, Jianfeng Pei & Luhua Lai, *LigBuilder 2: A Practical de Novo Drug Design Approach*. [Online Document], 2011. Available: Journal of Chemical Information and Modeling [Accessed: June 2022]

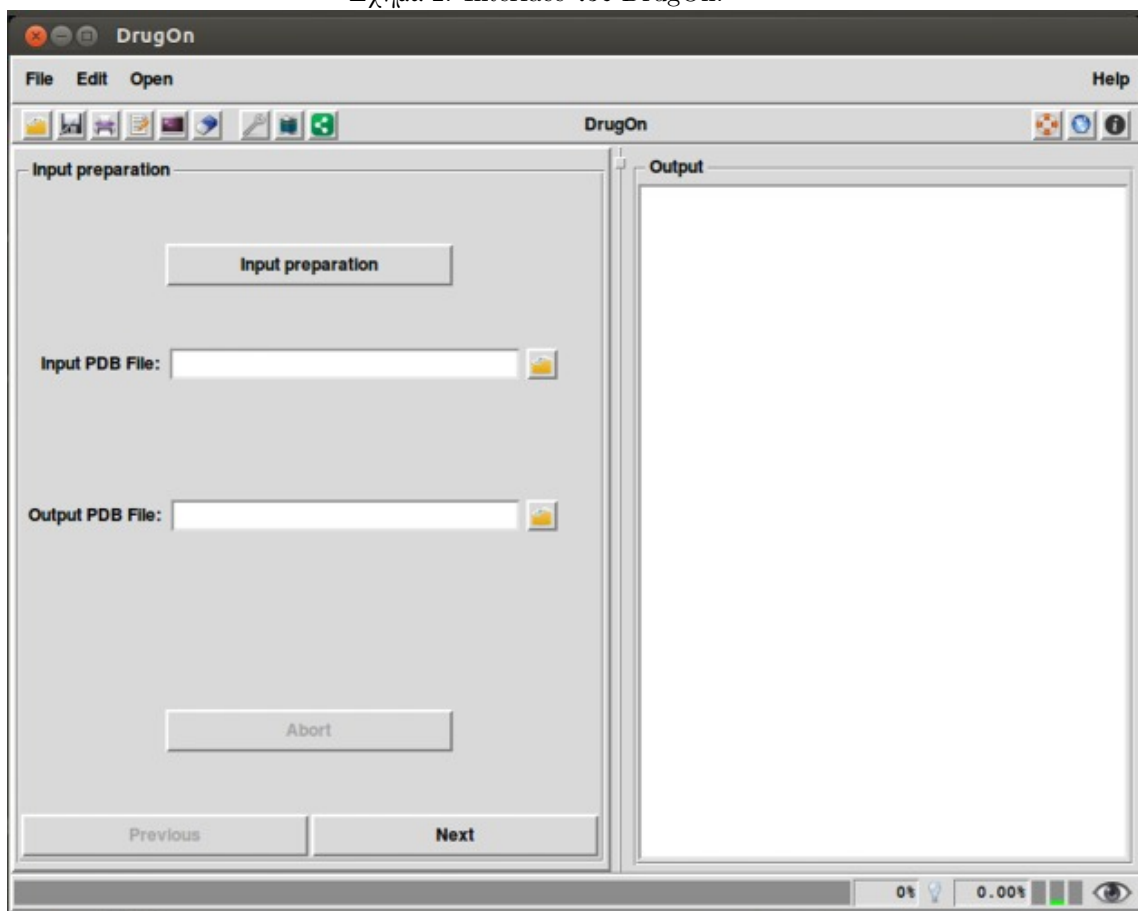
Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C. Smith, Berk Hess & Erik Lindahl, *GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers*. [Online Document], 2015. Available: ScienceDirect [Accessed: June 2022]

J. Polanski, *Comprehensive Chemometrics, Volume 4*. [Online Document], 2009. Available: ScienceDirect [Accessed: June 2022]

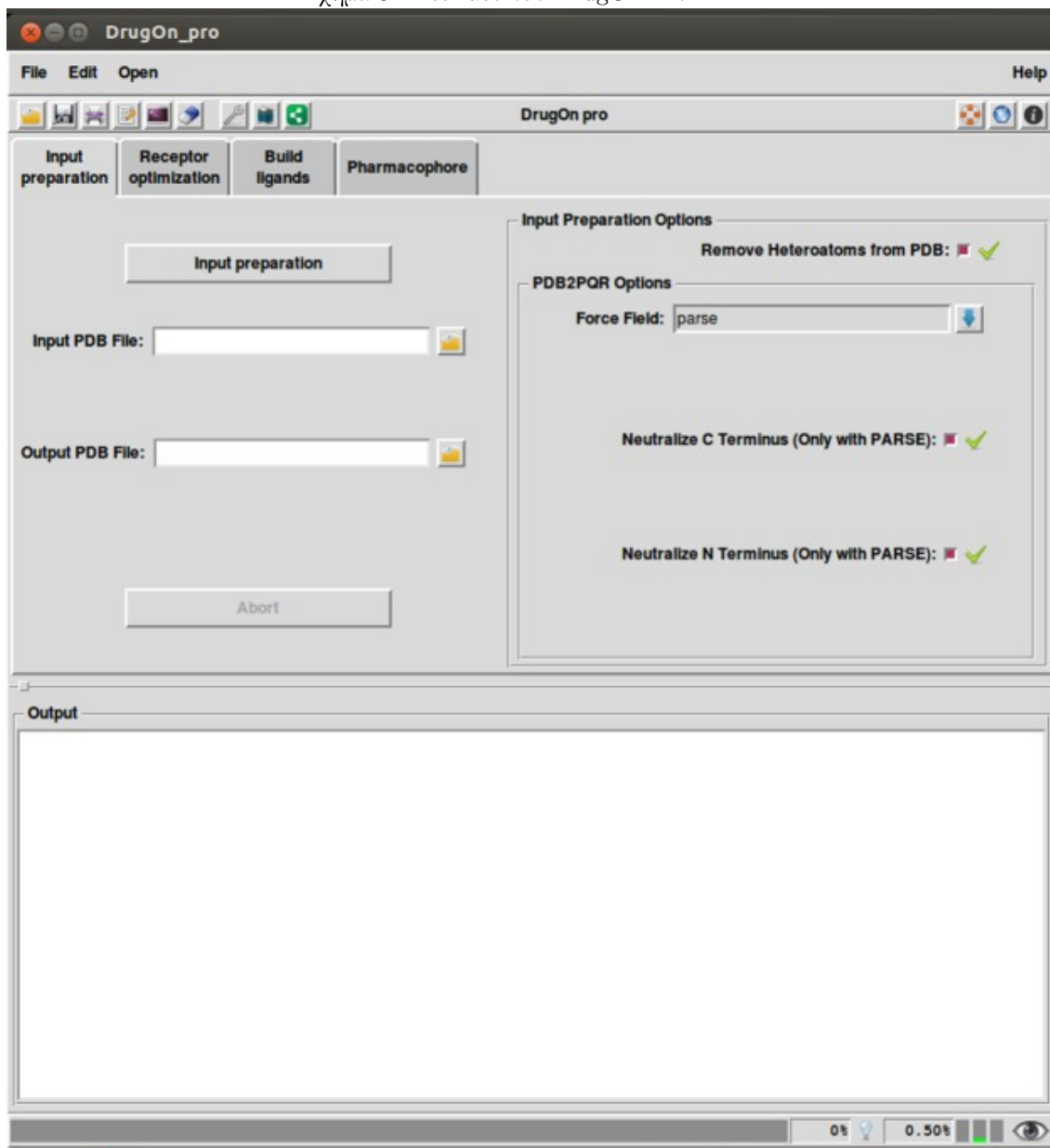
Σχήμα 1: Flowchart του pipeline του DrugOn.



Σχήμα 2: Interface του DrugOn.

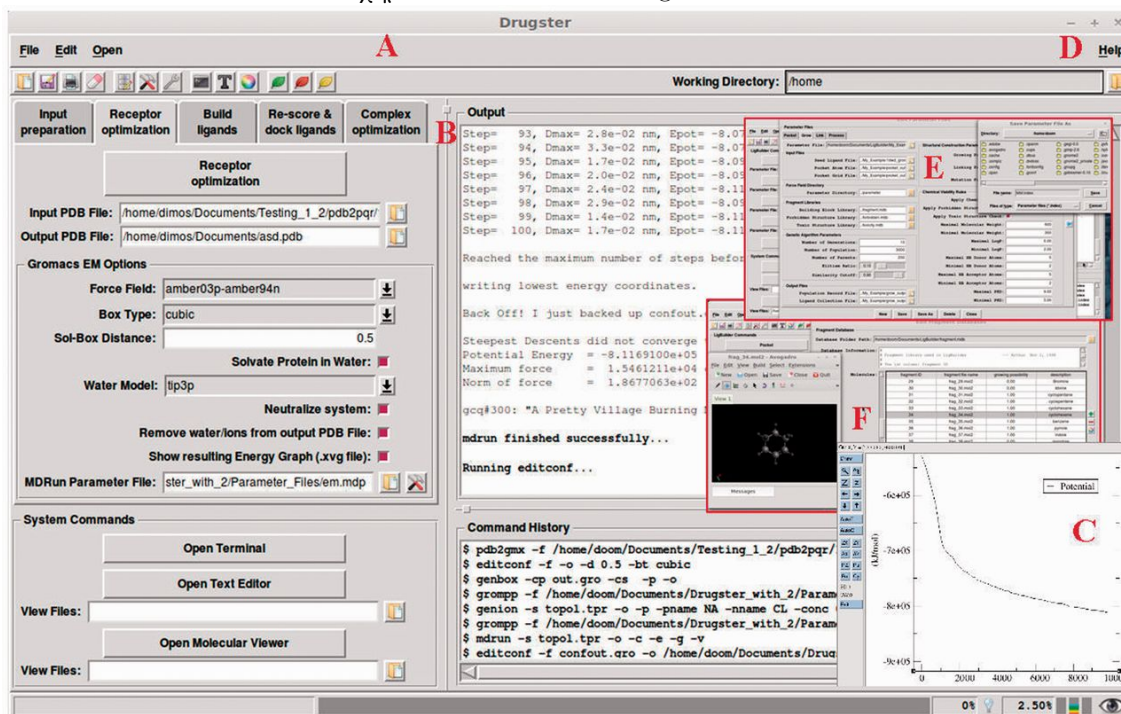


Σχήμα 3: Interface του DrugOn Pro.

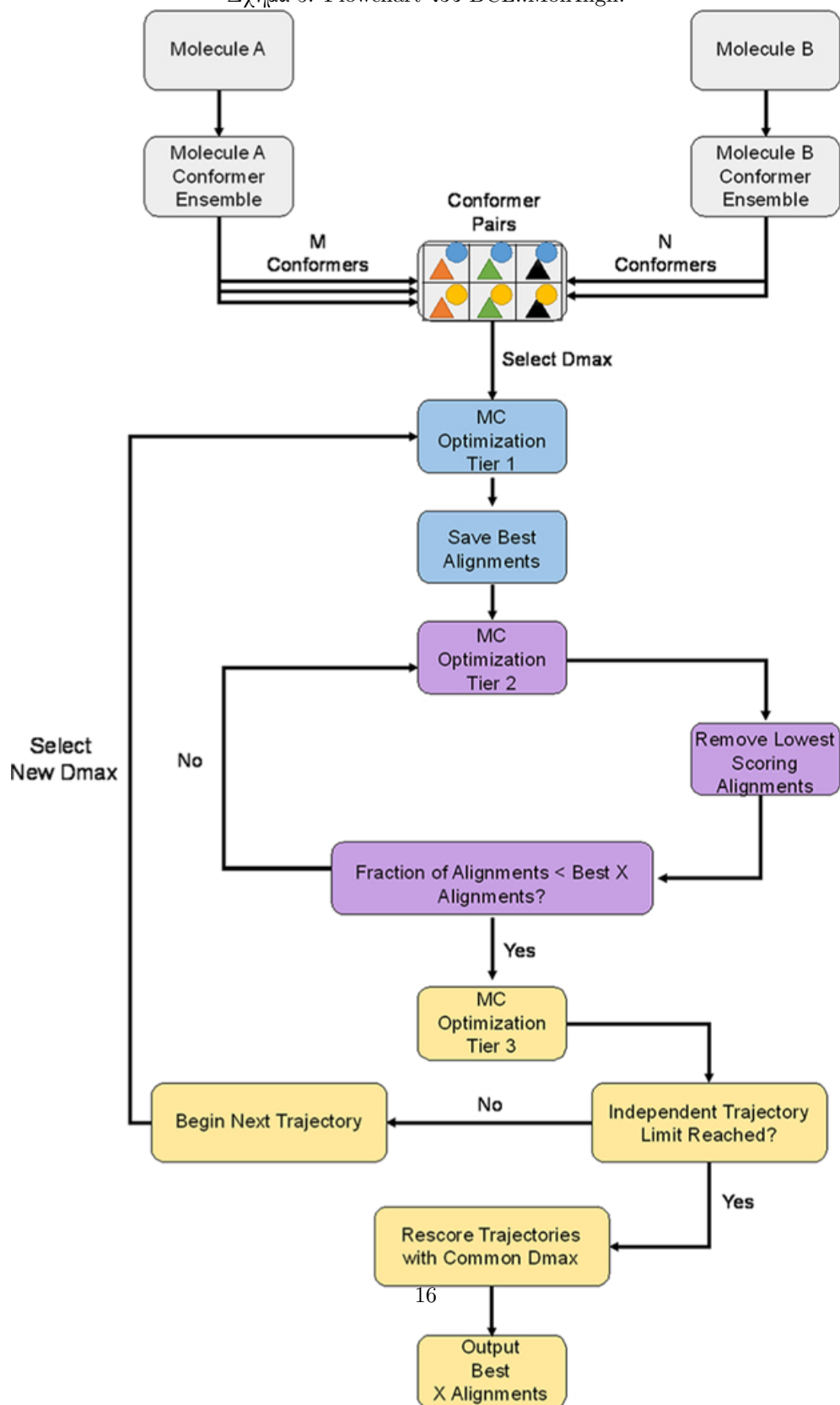




Σχήμα 4: Interface του Drugster.



Σχήμα 5: Flowchart του BCL::MolAlign.



Σχήμα 6: Στρατηγικές sampling του BCL::MolAlign.

