



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

**Δομική Ανάλυση και Υπολογιστική
Στόχευση του Υποδοχέα Ορμόνης
Απελευθέρωσης Αυξητικής Ορμόνης με
Δυνητικά Βιοδραστικά Μόρια**

ΤΣΙΑΝΑΚΑΣ ΚΩΝΣΤΑΝΤΙΝΟΣ
Αριθμός Μητρώου: 18388078

Επιβλέπων Καθηγητής
Μίνως Ματσούκας, Επίκουρος Καθηγητής

Αθήνα 30/10/2023

Η Τριμελής Εξεταστική Επιτροπή

Ο Επιβλέπων Καθηγητής

Μίνως Ματσούκας

Σπύρος Κωστόπουλος

Εμμανουήλ Αθανασιάδης

Επίκουρος Καθηγητής

Αναπληρωτής Καθηγητής

Επίκουρος Καθηγητής

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο υπογράφων **Τσιανάκας Κωνσταντίνος** του **Ευριπίδη**, με αριθμό μητρώου **18388078** φοιτητής του Τμήματος **Μηχανικών Βιοϊατρικής** της Σχολής **Μηχανικών** του Πανεπιστημίου Δυτικής Αττικής, δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του διπλώματός μου».

Ημερομηνία

30/10/2023

Ο Δηλών



ΠΕΡΙΛΗΨΗ

Η ορμόνη απελευθέρωσης αυξητικής ορμόνης (Growth Hormone Releasing Hormone, GHRH) είναι ένα νευροπεπτίδιο που αποτελείται από 44 αμινοξέα και παράγεται στον υποθάλαμο. Μετά την σύνθεση της απελευθερώνεται και οδηγείται προς τα σωματοτρόπα κύτταρα της υπόφυσης όπου προσδέεται με τον υποδοχέα ορμόνης απελευθέρωσης αυξητικής ορμόνης (Growth Hormone Releasing Hormone Receptor, GHRH-R) διεγείροντας την σύνθεση της αυξητικής ορμόνης (Growth Hormone, GH). Ο GHRH-R ανήκει στην Β οικογένεια των υποδοχέων συζευγμένων με G πρωτεΐνη (G-Proteins Coupled Receptors, GPCRs) και συγκεκριμένα στην Β1 υποκατηγορία των υποδοχέων σεκρετίνης. Σε περιπτώσεις σοβαρών διαταραχών συμπεριλαμβανομένου του καρκίνου του παγκρέατος, του μαστού, των ωοθηκών και του προστάτη, ο GHRH-R υπερεκφράζεται στα καρκινικά κύτταρα καθιστώντας τον έναν πολύ σημαντικό θεραπευτικό στόχο. Σκοπός της παρούσας εργασίας είναι η δομική ανάλυση του GHRH-R και η στόχευση του με μικρά μη πεπτιδικά μόρια. Αρχικά γίνεται η δομική ανάλυση του GHRH-R μέσω της μελέτης των αλληλεπιδράσεων που κάνει με το ενδογενές του πεπτίδιο. Έπειτα αναλύεται μια σειρά βιοδραστικών ενώσεων με την ένδειξη “MK”, που έχουν δοκιμαστεί πειραματικά και προέρχονται από δουλειά που έχει προηγηθεί. Στόχος είναι η μελέτη του μηχανισμού πρόσδεσης των ενώσεων στον υποδοχέα. Τέλος εφαρμόζεται η υπολογιστική τεχνική της εικονικής σάρωσης με σκοπό τον εντοπισμό δυναμικά βιοδραστικών μορίων με ανταγωνιστική δράση έναντι του GHRH-R.

Για την αξιολόγηση της σειράς των βιοδραστικών ενώσεων MK που έχουν δοκιμαστεί πειραματικά εφαρμόστηκε η υπολογιστική τεχνική της μοριακής πρόσδεσης. Για την επιτυχημένη εφαρμογή της μεθόδου απαιτείται μια σειρά από διαδικασίες όπως είναι η προετοιμασία της πρωτεΐνης-στόχου, των προσδετών καθώς και η ρύθμιση ενός συνόλου παραμέτρων. Η μοριακή πρόσδεση έλαβε μέρος στην ορθοστερική θέση της διαμεμβρανικής περιοχής του υποδοχέα τόσο για την ενεργή όσο και την ανενεργή του κατάσταση. Τα αποτελέσματα αξιολογήθηκαν με βάση το σκορ πρόσδεσης αλλά και μέσω των αλληλεπιδράσεων που σχηματίστηκαν μεταξύ των προσδετών και του GHRH-R. Ιδιαίτερη βαρύτητα δόθηκε στις ενώσεις που παρουσίασαν ενθαρρυντικά αποτελέσματα κατά την πειραματική τους δοκιμή. Συγκεκριμένα ξεχώρισε για την αγωνιστική της δράση η ένωση MK24 ενώ για την ανταγωνιστική της δράση η MK04 και οι παρόμοιες με αυτή MK04-6 και MK04-9. Τα υπολογιστικά αποτελέσματα που προέκυψαν ποικίλουν. Ειδικότερα το σκορ πρόσδεσης της ένωσης MK24 δεν αντικατοπτρίζει τα πειραματικά αποτελέσματα. Από την άλλη αλληλεπιδρά με συγκεκριμένα αμινοξέα του υποδοχέα που συμμετέχουν και κατά την διαδικασία σχηματισμού του συμπλόκου GHRHR – GHRH και φαίνεται να παίζουν σημαντικό ρόλο στην αγωνιστική δράση της ένωσης MK24. Η σειρά των ενώσεων με ανταγωνιστική δράση MK04, MK04-6 και MK04-9 σημείωσε καλό σκορ πρόσδεσης. Κατά την ανάλυση των αλληλεπιδράσεων τους, παρατηρήθηκε πως οι τρεις ενώσεις έχουν παρόμοιο τρόπο δέσμευσης όμως οι μικρές δομικές διαφορές τους συντελούν σε μοναδικές αλληλεπιδράσεις με τον υποδοχέα. Γενικότερα και οι τρεις ενώσεις αλληλεπιδρούν με συγκεκριμένα αμινοξέα του υποδοχέα που επιβεβαιώνουν εν μέρη τα πειραματικά αποτελέσματα και την ανταγωνιστική τους δράση.

Στην συνέχεια με σκοπό την ανακάλυψη δυναμικά βιοδραστικών μορίων με ανταγωνιστική δράση έναντι του GHRH-R εφαρμόστηκε η εικονική σάρωση. Χρησιμοποιήθηκαν δύο χημειοθήκες με 14.511 βιοδραστικές ενώσεις συνολικά οι

οποίες αντλήθηκαν ηλεκτρονικά και ακολούθησε η προ-επεξεργασία τους. Αφού πρώτα οι ενώσεις φιλτραρίστηκαν με βάση τον κανόνα των πέντε του Lipinski και προετοιμάστηκαν, ακολούθησε η εφαρμογή της μοριακής πρόσδεσης. Με βάση το σκορ πρόσδεσης ξεχώρισαν οι καλύτερες ενώσεις. Στις ενώσεις αυτές εφαρμόστηκε ο συντελεστής Tanimoto και ξεχώρισαν εκείνες με το χαμηλότερο ποσοστό ομοιότητας που κωδικοποιήθηκαν με τα μοριακά αποτυπώματα RDKit. Ακολούθησε η αξιολόγηση των ενώσεων από την οποία προέκυψαν οι τελικές ενώσεις με πιθανή ανταγωνιστική δράση έναντι του GHRH-R. Μάλιστα από αυτές ξεχώρισε μια ένωση που συνδυάζει καλό σκορ πρόσδεσης και αλληλεπιδράσεις με σημαντικά αμινοξέα του υποδοχέα.

Λέξεις Κλειδιά: GPCRs, GHRH-R, GHRH, Κρύο-Ηλεκτρονική Μικροσκοπία, Μοριακή Πρόσδεση, Μοριακοί Περιγραφείς, Μοριακά Αποτυπώματα, Συντελεστής Tanimoto, Χημειοθήκες, Εικονική Σάρωση

ABSTRACT

Growth Hormone Releasing Hormone (GHRH) is a neuropeptide consisting of 44 amino acids and produced in hypothalamus. After its synthesis, it is released and directed to the somatotrophic cells of the pituitary gland where it binds to the Growth Hormone Releasing Hormone Receptor (GHRH-R) stimulating the synthesis of Growth Hormone (GH). GHRH-R belongs to the B family of G-Proteins Coupled Receptors (GPCRs) and specifically in the B1 subclass of secretin receptors. In cases of severe disorders, including pancreatic, breast, ovarian and prostate cancer, GHRH-R is overexpressed in cancer cells, making it a significant therapeutic target. The aim of this study is the structural analysis of GHRH-R and its targeting with small non-peptide molecules. Initially, the structural analysis of GHRH-R is conducted by examining its interactions with its endogenous peptide. Subsequently, a series of bioactive compounds labeled as “MK”, which have been experimentally tested and derived from previous research, are analyzed. The purpose is to study the mechanism of binding of these compounds to the receptor. Finally, the computational technique of virtual screening is applied to identify potential bioactive molecules with competitive action against GHRH-R.

For the evaluation of the series of bioactive compounds MK that have been experimentally tested, the computational technique of molecular docking was applied. Successful application of the method requires a series of procedures such as the preparation of the protein target, the ligands and the adjustment of a set of parameters. Molecular docking took place in the orthosteric binding site of the transmembrane region of the receptor for both its active and inactive states. The results were assessed based on the docking score and the interactions formed between the ligands and GHRH-R. Special attention was given to compounds that exhibited encouraging results during their experimental testing. Specifically, the compound MK24 stood out for its agonistic action, while MK04 and its analogs MK04-6 and MK04-9, demonstrated competitive action. Computational results varied. Particularly the docking score of compound MK24 did not reflect the experimental outcomes. On the other hand, it interacted with specific amino acids of the receptor involved in the formation of the GHRH-R-GHRH complex, indicating a significant role in the agonistic action of MK24. The series of compounds with antagonistic action, MK04, MK04-6 and MK04-9 achieved a good docking score. In the analysis of their interactions, it was observed that all three compounds had a similar binding mode, yet their small structural differences led to unique interactions with the receptor. Overall, all three compounds interacted with specific amino acids of the receptor, partially confirming the experimental results and their competitive action.

Next, with the aim of discovering potential bioactive molecules with competitive action against GHRH-R, virtual screening was implemented. Two chemical libraries with a total of 14.511 bioactive compounds were utilized, electronically retrieved and subjected to pre-processing. After the compounds were filtered based on the Lipinski's rule of five and prepared, molecular docking was applied. The best compounds were distinguished based on the docking score. The Tanimoto coefficient was applied to these compounds, highlighting those with the lowest similarity percentage encoded with RDKit molecular fingerprints. The evaluation of the compounds followed, from which the final compounds with potential competitive action against GHRH-R

emerged. Notably, among these a compound stood out combining a good docking score and interactions with significant amino acids of the receptor.

Keywords: *GPCRs, GHRH-R, GHRH, Cryogenic-Electron Microscopy, Molecular Docking, Molecular Descriptors, Molecular Fingerprints, Tanimoto coefficient, Chemical Libraries, Virtual Screening*

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο εργαστήριο Υπολογιστικής Χημείας του τμήματος Μηχανικών Βιοϊατρικής του Πανεπιστημίου Δυτικής Αττικής κατά το ακαδημαϊκό έτος 2022-2023 υπό την επίβλεψη του καθηγητή Δρ. Μίνω Ματσούκα.

Θα ήθελα να εκφράσω τις ευχαριστίες μου σε όλους όσους συνέβαλαν στην εκπόνηση της διπλωματικής εργασίας. Πρωτίστως θα ήθελα να ευχαριστήσω θερμά τον επιβλέπων καθηγητή Δρ. Μίνω Ματσούκα για την καθοδήγηση του, τον πολύτιμο χρόνο που αφιέρωσε και τις ουσιαστικές συμβουλές του σε όλη την διάρκεια εκπόνησης της διπλωματικής εργασίας, καθώς και για την δυνατότητα που μου πρόσφερε να γνωρίσω τον τομέα της υπολογιστικής χημείας. Ακόμα θα ήθελα να τον ευχαριστήσω για την εμπιστοσύνη του απέναντι στο πρόσωπο μου σε επίπεδο πρακτικής άσκησης.

Στην συνέχεια θα ήθελα να ευχαριστήσω βαθύτατα τον Υποψήφιο Διδάκτορα Βασίλη Παναγιωτόπουλο για την υποστήριξη του καθ' όλη τη διάρκεια της διπλωματικής εργασίας. Με τις γνώσεις και την εμπειρία του συνέβαλε στο να ξεπεραστούν τα εμπόδια και να επιλυθούν τα διάφορα ζητήματα που παρουσιάστηκαν. Ήταν πάντα πρόθυμος και η βοήθειά του ήταν πολύτιμη για την ολοκλήρωση της εργασίας. Σε γενικότερο πλαίσιο θα ήθελα να τον ευχαριστήσω για την συνεργασία που είχαμε τόσο σε επίπεδο διπλωματικής όσο και πρακτικής άσκησης.

Η παρούσα διπλωματική εργασία αποτελεί επέκταση μιας δουλειάς που έχει προηγηθεί, συνεπώς θα ήθελα να ευχαριστήσω θερμά και όλους όσους συντέλεσαν στην ολοκλήρωση της. Επιπρόσθετα του Δρ. Μίνω Ματσούκα και του Υποψήφιου Διδάκτορα Βασίλη Παναγιωτόπουλου θα ήθελα να ευχαριστήσω από το εργαστήριο Φαρμακολογίας του Πανεπιστημίου της Πρετόρια την καθηγήτρια Dr. Claire Newton και την ομάδα της Tarryn Radomsky, Robin du Preez, Robert Millar και Ross Anderson για την υλοποίηση των πειραματικών διαδικασιών και την παροχή χρήσιμων αποτελεσμάτων. Επίσης θα ήθελα να ευχαριστήσω των καθηγητή Γεώργιο Σπυρούλια καθώς και τον ερευνητή Δρ. Μιχάλη Παπαδουράκη για την συνεισφορά τους στο έργο.

Τέλος θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους μου για όλη την υποστήριξή τους κατά την διάρκεια των σπουδών μου.

Περιεχόμενα

ΕΙΣΑΓΩΓΗ	15
1. Θεωρητικό υπόβαθρο	18
1.1 Κυτταρική Μembrάνη & Μembrανικές Πρωτεΐνες.....	18
1.2 Υποδοχείς συζευγμένοι με G πρωτεΐνη (GPCRs).....	18
1.2.1 Εισαγωγή στους GPCRs.....	18
1.2.1 Δομή των GPCR.....	19
1.2.2 Πορεία Μεταγωγής Σήματος.....	21
1.3 Μέθοδοι Προσδιορισμού Δομής.....	22
1.3.1 Κρυσταλλογραφία Ακτινών Χ.....	22
1.3.2 Κρύο-Ηλεκτρονική Μικροσκοπία (Cryogenic-Electron Microscopy - Cryo-EM).....	23
1.3.3 Φασματοσκοπία Πυρηνικού Μαγνητικού Συντονισμού (Nuclear Magnetic Resonance Spectroscopy - NMR).....	25
1.4 Υποδοχέας Ορμόνης Απελευθέρωσης Αυξητικής Ορμόνης.....	26
1.5 Ορμόνη Απελευθέρωσης Αυξητικής Ορμόνης.....	28
1.6 Μηχανισμός πρόσδεσης του πεπτιδίου (GHRH) στον υποδοχέα (GHRH-R) ..	29
1.7 Χημειοπληροφορική.....	31
1.7.1 Μοριακές Αναπαραστάσεις.....	32
1.7.2 Μοριακοί Περιγραφείς.....	37
1.7.3 Μοριακά Αποτυπώματα.....	38
1.7.4 Μοριακή Ομοιότητα - Αναζήτηση Ομοιότητας Tanimoto.....	41
1.8 Μοριακή Πρόσδεση.....	43
1.8.1 Μοριακή Πρόσδεση - Μεθοδολογίες.....	43
1.8.2 Διαμοριακές Αλληλεπιδράσεις.....	43
1.8.3 Αλγόριθμοι Διαμορφωτικής Αναζήτησης.....	44
1.8.4 Συναρτήσεις Βαθμολόγησης.....	46
1.9 Εικονική Σάρωση.....	47
1.9.1 Εισαγωγή στην Εικονική Σάρωση.....	47
1.9.2 Βιβλιοθήκες Χημικών Ενώσεων.....	48
1.9.3 Εικονική Σάρωση Βάση Δομής και Βάση Προσδέτη.....	49
2. Μέθοδοι	51
2.1 Πρωτεΐνη – Στόχος.....	51
2.2 Προσδέτες.....	51
2.3 Λογισμικά Χημειοπληροφορικής.....	52
2.4 Αξιολόγηση βιοδραστικών ενώσεων της σειράς MK.....	53
2.4.1 Προετοιμασία Προσδετών.....	53
2.4.2 Προετοιμασία Πρωτεΐνης – Επιλογή περιοχής πρόσδεσης.....	53
2.4.3 Εφαρμογή Μοριακής Πρόσδεσης.....	54
2.4.4 Αξιολόγηση Μοριακής Πρόσδεσης.....	54
2.5 Εικονική Σάρωση.....	55
2.5.1 Έλεγχος ομοιότητας.....	56

3. Αποτελέσματα	58
3.1 Δομική Ανάλυση GHRH-R	58
3.2 Υπολογιστική αξιολόγηση βιοδραστικών ενώσεων της σειράς MK	61
3.2.1 Αποτελέσματα Μοριακής Πρόσδεσης	61
3.2.2 Αλληλεπιδράσεις προσδετών με τον GHRH-R.....	64
3.4 Αποτελέσματα Εικονικής Σάρωσης	67
4. Συζήτηση και Συμπεράσματα	73
Βιβλιογραφία.....	76
Παράρτημα.....	81

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1.1: Δομή των GPCR.....	20
Εικόνα 1.2: Σύμπλοκο GHRHR-GHRH-G _s	27
Εικόνα 1.3: Δομή του GHRH-R.....	28
Εικόνα 1.4: Μηχανισμός έκκρισης της GH.....	29
Εικόνα 1.5: Διαδικασία δύο σταδίων πρόσδεσης του GHRH στον GHRH-R.....	30
Εικόνα 1.6: Σύγκριση δομής ενεργής και ανενεργής κατάστασης του GHRH-R.....	31
Εικόνα 1.7: Διαφορετικές μοριακές αναπαραστάσεις του ίδιου μορίου.....	33
Εικόνα 1.8: Οργάνωση Molfile αρχείου για την L-Αλανίνη.....	34
Εικόνα 1.9: Κωδικοποίηση οργανικών μορίων με SMILES.....	35
Εικόνα 1.10: Κωδικοποίηση με InChI και InChIKey.....	37
Εικόνα 1.11: Μοριακό αποτύπωμα βασισμένο σε δομικά κλειδιά υποομάδων.....	39
Εικόνα 1.12: Τοπολογικά μοριακά αποτυπώματα τύπου Daylight.....	40
Εικόνα 1.13: Κυκλικό μοριακό αποτύπωμα.....	41
Εικόνα 1.14: Διάγραμμα ροής CADD.....	50
Εικόνα 3.1: Σύνολο αλληλεπιδράσεων μεταξύ GHRH και διαμεμβρανικού τομέα του GHRH-R.....	59
Εικόνα 3.2: Αναλυτική αναπαράσταση αλληλεπιδράσεων μεταξύ GHRH και διαμεμβρανικού τομέα του GHRH-R.....	60
Εικόνα 3.3: Τα σημαντικά αμινοξέα του GHRH-R που ξεχώρισαν κατά την πρόσδεση του GHRH.....	61
Εικόνα 3.4: Αποτελέσματα πειραματικού ελέγχου αγωνιστικής δραστηριότητας των ενώσεων MK(01-50).....	62
Εικόνα 3.5: Αποτελέσματα μοριακής πρόσδεσης των ενώσεων MK(01-50) στην ενεργή κατάσταση του υποδοχέα.....	62
Εικόνα 3.6: Αποτελέσματα πειραματικού ελέγχου ανταγωνιστικής δραστηριότητας των ενώσεων MK(01-50).....	63
Εικόνα 3.7: Αποτελέσματα μοριακής πρόσδεσης των ενώσεων MK(01-50) στην ανενεργή κατάσταση του υποδοχέα.....	63
Εικόνα 3.8: Κώδικας για την μετατροπή των ενώσεων.....	63
Εικόνα 3.9: Κώδικας για την εφαρμογή της μοριακής πρόσδεσης.....	64
Εικόνα 3.10: Αλληλεπίδραση της ένωσης MK24 με την ενεργή κατάσταση του υποδοχέα.....	66

Εικόνα 3.11: Αλληλεπίδραση των ενώσεων MK04, MK04-6 και MK04-9 με την ανενεργή κατάσταση του υποδοχέα	66
Εικόνα 3.12: Κώδικας για την εφαρμογή της μοριακής πρόσδεσης για την υλοποίηση της εικονικής σάρωσης.....	67
Εικόνα 3.13: Κώδικας για την εξαγωγή μοριακών αποτυπωμάτων και την εφαρμογή της αναζήτησης ομοιότητας.....	68
Εικόνα 3.14: Κάτω τριγωνικός χάρτης συντελεστών Tanimoto.....	68
Εικόνα 3.15: Οι δέκα ενώσεις που προέκυψαν από την εικονική σάρωση.....	71
Εικόνα 3.16: Η μια ένωση που ξεχώρισε από την εικονική σάρωση και οι αλληλεπιδράσεις της με τον υποδοχέα.....	71

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1.1: Συντελεστές συσχέτισεων και αποστάσεων στην αναζήτηση ομοιότητας.....	42
Πίνακας 3.1: Δισδιάστατη δομή των ενώσεων MK24, MK04, MK04-6 και MK04-9.....	64
Πίνακας 3.2: Δισδιάστατη δομή των τελικών ενώσεων που προέκυψαν από την εικονική σάρωση.....	69
Πίνακας 6.1: Οι εκατό καλύτερες ενώσεις που προέκυψαν από την εικονική σάρωση.....	81

ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ

Συντομογραφίες	Ορισμοί
ADMET	Adsorption Distribution Metabolism Elimination and Toxicity
CADD	Computer Aided Drug Design
cAMP	Cyclic Adenosine Monophosphate
Cryo-EM	Cryogenic Electron Microscopy
Cryo-STEM	Cryogenic Scanning Transmission Electron Microscopy
Cryo-ET	Cryogenic Electron Tomography
CW	Continuous Ware
DED	Direct Electron Detectors
ECD	Extracellular Domain
ECFP	Extended Connectivity Fingerprints
ECL	Extracellular Loop
FCFP	Functional-Class Fingerprints
FDA	Food and Drug Administration
GA	Genetic Algorithms
GCGR	Glucagon Receptor
GPCR	G-Protein Coupled Receptor
GHIH	Growth Hormone Inhibiting Hormone
GHRH-R	Growth Hormone Releasing Hormone Receptor
GHRH	Growth Hormone Releasing Hormone
GH	Growth Hormone
GHS	Growth Hormone Secretagogue
GLP 1	Glucagon Like Peptide 1
HTS	High Throughput Screening
ICL	Intracellular Loop
InChI	International Chemical Identifier
IUPAC	International Union of Pure and Applied Chemistry
LBDD	Ligand Based Drug Design
LBVS	Ligand Based Virtual Screening
MACCS	Molecular Access System
Micro-ED	Micro Electron Diffraction
MC	Monte Carlo
NIST	National Institute of Standards and Technology
NMR	Nuclear Magnetic Resonance
SBDD	Structure Based Drug Design
SBVS	Structure Based Virtual Screening
SMILES	Simplified Molecular Input Line Entry System
SMARTS	SMILES Arbitrary Target Specification
SS	Somatostatin
SPA	Single Particle Analysis
SNR	Signal to Noise Ratio
PKA	Protein Kinase A
QSAR	Quality Structure Activity Relationship

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ
ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

TS	Tabu Search
Tc	Tanimoto coefficient
TMD	Transmembrane Domain
VS	Virtual Screening

ΕΙΣΑΓΩΓΗ

Η διαδικασία του σχεδιασμού και της ανακάλυψης ενός φαρμάκου αποτελεί μια πολύ δαπανηρή και χρονοβόρα διαδικασία. Σύμφωνα με πρόσφατες μελέτες εκτιμάται πως από την αρχή της έρευνας έως την προώθηση του φαρμάκου στην αγορά χρειάζονται 10-15 χρόνια και δαπανούνται κατά μέσο όρο 2.87 δισεκατομμύρια δολάρια. Τα τελευταία χρόνια όλο και περισσότερες φαρμακευτικές εταιρείες χρησιμοποιούν υπολογιστικές μεθόδους οι οποίες σημειώνουν μεγάλη επιτυχία στην ανακάλυψη υποψήφιων φαρμάκων μειώνοντας τόσο το κόστος όσο και τον χρόνο που απαιτείται για να ολοκληρωθεί η διαδικασία. Καθοριστικοί παράγοντες που εδραίωσαν την χρήση υπολογιστικών μεθόδων σε αυτόν τον τομέα είναι η ραγδαία τεχνολογική ανάπτυξη τόσο από άποψη υλικού όσο και λογισμικού, η βελτίωση αλγορίθμων, η εισαγωγή του cloud computing (παροχή υπηρεσιών υπολογιστικής ισχύος και αποθήκευσης δεδομένων στο διαδίκτυο) και η ανάπτυξη του κλάδου της δομικής βιολογίας με την βοήθεια τεχνικών όπως η κρύο-ηλεκτρονική μικροσκοπία για τον προσδιορισμό της τρισδιάστατης δομής πρωτεϊνών που μέχρι πρότινος ήταν άγνωστη. Χαρακτηριστικό παράδειγμα είναι οι πρωτεΐνες της οικογένειας των GPCR (G-Protein Coupled Receptors) που αποτελούν έναν από τους σημαντικότερους φαρμακευτικούς στόχους δεδομένου ότι το 1/3 των υπαρχόντων φαρμάκων στοχεύουν μια πρωτεΐνη αυτής της οικογένειας [1],[2].

Ο GHRH-R ανήκει στην Β οικογένεια των GPCR και συγκεκριμένα στην Β1 υποκατηγορία των υποδοχέων σεκρετίνης. Σε σοβαρές περιπτώσεις διαταραχών όπως είναι ο καρκίνος του παγκρέατος, του μαστού, των ωθηκών και του προστάτη, υπερεκφράζεται στα καρκινικά κύτταρα ο GHRH-R. Ως επόμενο οδηγεί σε αύξηση της παραγωγής των καρκινικών κυττάρων καθιστώντας τον έναν πολύ σημαντικό θεραπευτικό στόχο. Μέχρι και σήμερα δεν έχουν βρεθεί μικρά μη πεπτιδικά μόρια που να αναστέλλουν την δράση του GHRH-R. Η πρόσφατη επίλυση της τρισδιάστατης δομής του GHRH-R μαζί με το ενδογενές πεπτίδιο GHRH και την G_s πρωτεΐνη άνοιξε νέους ορίζοντες για την εφαρμογή υπολογιστικών τεχνικών με μεγαλύτερη ακρίβεια και αξιοπιστία με σκοπό τον εντοπισμό μικρών μη πεπτιδικών μορίων με ανταγωνιστική δράση έναντι του υποδοχέα [1],[2].

Πρωταρχικός σκοπός της παρούσας διπλωματικής εργασίας είναι η δομική ανάλυση του GHRH-R τόσο μακροσκοπικά μέσω της μελέτης του μηχανισμού του υποδοχέα και των διαμορφωτικών αλλαγών στις διάφορες καταστάσεις (ενεργή – ανενεργή) όσο και μικροσκοπικά αναλύοντας τις αλληλεπιδράσεις που σχηματίζονται κατά την πρόσδεση του πεπτιδίου στο ορθοστερικό κέντρο του υποδοχέα. Με αυτόν τον τρόπο θα εντοπιστούν εκείνα τα στοιχεία που ευνοούν την πρόσδεση και θα αποτελέσουν χρήσιμη πληροφορία για την πορεία της εργασίας. Στην συνέχεια της εργασίας λαμβάνει μέρος η αξιολόγηση μια ομάδας ενώσεων που έχουν δοκιμαστεί πειραματικά, η σύγκρισή τους με υπολογιστικά δεδομένα και η μελέτη του μηχανισμού πρόσδεσής τους. Τελειώνοντας, γίνεται εφαρμογή της υπολογιστικής τεχνικής της εικονικής σάρωσης με στόχο τον εντοπισμό δυναμικά βιοδραστικών ενώσεων με ανταγωνιστική δράση έναντι του GHRH-R.

Στο πρώτο κεφάλαιο καλύπτεται το θεωρητικό υπόβαθρο της παρούσας διπλωματικής εργασίας. Αναλύονται η θεωρία και οι βασικές έννοιες με σκοπό την καλύτερη κατανόηση του θέματος και την προετοιμασία του εδάφους για τα επόμενα κεφάλαια. Αρχικά γίνεται μια εισαγωγή στις διαμεμβρανικές πρωτεΐνες με εστίαση στους

υποδοχείς συζευγμένους με G πρωτεΐνες για τους οποίους αναλύεται η δομή και η λειτουργία τους καθώς και η σημαντικότητά τους ως θεραπευτικοί στόχοι. Έπειτα μελετώνται ως προς την αρχή λειτουργίας, την οργανολογία, τα πλεονεκτήματα και μειονεκτήματά τους, οι τρεις βασικές μέθοδοι προσδιορισμού της δομής μακρομορίων που είναι η κρυσταλλογραφία ακτινών X, η φασματοσκοπία πυρηνικού μαγνητικού συντονισμού και η κρύο-ηλεκτρονική μικροσκοπία. Στην συνέχεια γίνεται λόγος για τον GHRH-R που αποτελεί το επίκεντρο της εργασίας. Μελετάται η δομή του GHRH-R ο τρόπος ρύθμισης της παραγωγής της GH καθώς και η αλληλεπίδραση μεταξύ του GHRH-R και του ενδογενούς πεπτιδίου GHRH. Στην συνέχεια περιγράφονται κάποιες από τις βασικές έννοιες που συναντώνται στον τομέα της χημειοπληροφορικής όπως είναι οι τρόποι αναπαράστασης χημικών δομών, οι μοριακοί περιγραφείς, τα μοριακά αποτυπώματα και η μοριακή ομοιότητα. Τέλος γίνεται εισαγωγή στον κλάδο της υπολογιστικής χημείας και αναλύονται οι τεχνικές της μοριακής πρόσδεσης και της εικονικής σάρωσης που χρησιμοποιήθηκαν στην εργασία.

Στο δεύτερο κεφάλαιο παρουσιάστηκε η μεθοδολογία που ακολουθήθηκε για την διεκπεραίωση της εργασίας. Συγκεκριμένα παρουσιάζονται οι προσδέτες, η πρωτεΐνη-στόχος καθώς και οι βάσεις δεδομένων στις οποίες είναι αποθηκευμένα. Επίσης παρατίθενται και όλα τα λογισμικά χημειοπληροφορικής που χρησιμοποιήθηκαν για την προετοιμασία των πρωτεϊνών και των ενώσεων, για την εφαρμογή της μοριακής πρόσδεσης και για την ανάλυση των αποτελεσμάτων. Για την αξιολόγηση των βιοδραστικών ενώσεων της σειράς MK αναλύθηκε η προετοιμασία και προεπεξεργασία τόσο του υποδοχέα όσο και των ενώσεων, η διαδικασία μετατροπής τους στην κατάλληλη μορφή αρχείων που απαιτείται και ο τρόπος με τον οποίο εφαρμόστηκε η μοριακή πρόσδεση. Έπειτα αναλύθηκε η μεθοδολογία που ακολουθήθηκε για την εφαρμογή της εικονικής σάρωσης. Αναλυτικά, στην μεθοδολογία γίνεται λόγος και εδώ για την προετοιμασία των ενώσεων και του υποδοχέα, για το φιλτράρισμα των χημειοθηκών με συγκεκριμένες φυσικό-χημικές ιδιότητες, για τον τρόπο εφαρμογής της μοριακής πρόσδεσης και για τα κριτήρια διαλογής των τελικών ενώσεων.

Στο τρίτο κεφάλαιο παρατέθηκαν τα αποτελέσματα που προέκυψαν. Αναλύθηκαν σε ατομικό επίπεδο οι αλληλεπιδράσεις που συμβαίνουν κατά την δημιουργία του συμπλόκου GHRH-GHRH. Έπειτα καταγράφηκαν τα αποτελέσματα που προέκυψαν από την εφαρμογή των υπολογιστικών τεχνικών. Ειδικότερα παρουσιάστηκαν σε μορφή διαγράμματος τα σκορ πρόσδεσης των βιοδραστικών ενώσεων MK για τις δύο διαφορετικές καταστάσεις του υποδοχέα (ενεργή και ανενεργή κατάσταση) ενώ για τις ενώσεις που ξεχώρισαν για την ανταγωνιστική και αγωνιστική τους δράση κατά την πειραματική διαδικασία, παρουσιάστηκαν μέσω εικόνων οι αλληλεπιδράσεις τους με τον υποδοχέα. Για την εφαρμογή της εικονικής σάρωσης σε δύο βιβλιοθήκες χημικών ενώσεων, φαίνονται μέσω ενός κάτω τριγωνικού χάρτη οι συντελεστές Tanimoto για εκατό ενώσεις ενώ με την χρήση πίνακα καταγράφονται οι δέκα ενώσεις που ξεχώρισαν. Όπου χρειάστηκε, παρατέθηκαν μέρη κώδικα, όπως αυτά για την μετατροπή των ενώσεων, για την εφαρμογή του συντελεστή Tanimoto και για την εφαρμογή της μοριακής πρόσδεσης. Τέλος παρουσιάζονται σε μια εικόνα οι δέκα ενώσεις προσδεμένες στο ενεργό κέντρο του υποδοχέα καθώς επίσης και οι αλληλεπιδράσεις της μιας ένωσης που ξεχώρισε με τον υποδοχέα.

Στο τέταρτο και τελευταίο κεφάλαιο συζητήθηκαν τα αποτελέσματα, εξάχθηκαν συμπεράσματα και διατυπώθηκαν προτάσεις για τις μελλοντικές προοπτικές του

θέματος. Συγκεκριμένα τα υπολογιστικά αποτελέσματα για την αξιολόγηση της σειράς των βιοδραστικών ενώσεων ΜΚ, συγκρίθηκαν με τα πειραματικά αποτελέσματα προκειμένου να προσδιοριστεί η συσχέτισή τους. Σχολιάστηκε ο μηχανισμός πρόσδεσης των ενώσεων που ξεχώρισαν κατά την πειραματική διαδικασία, σημειώθηκαν εκείνα τα δομικά στοιχεία που ισχυροποίησαν την σύνδεση κάθε ένωσης ενώ παράλληλα κατά την πρόσδεση, αναζητήθηκαν αμινοξέα του υποδοχέα που έχει διαπιστωθεί ότι έχουν καταλυτικό ρόλο στην σύνδεση του GHRH. Τέλος εξάχθηκαν συμπεράσματα για τις τελικές ενώσεις που προέκυψαν καθώς και για την μια που ξεχώρισε από την εικονική σάρωση και σχολιάστηκε η συνεισφορά κάθε βήματος που εφαρμόστηκε, στο τελικό αποτέλεσμα.

1. Θεωρητικό υπόβαθρο

1.1 Κυτταρική Μεμβράνη & Μεμβρανικές Πρωτεΐνες

Οι κυτταρικές ή αλλιώς πλασματικές μεμβράνες αποτελούν απαραίτητο στοιχείο για την βιωσιμότητα ενός κυττάρου καθώς διαχωρίζουν το εσωτερικό από το εξωτερικό του περιβάλλον και επιτρέπουν σε ορισμένες μόνο ουσίες να εισέρχονται ή να εξέρχονται στο κύτταρο (επιλεκτική διαπερατότητα) [3]. Αποτελούνται από λιπίδια τα οποία έχουν μια υδρόφοβη υδρογονανθρακική ουρά και μια υδρόφιλη κεφαλή. Οι πολικές ομάδες της υδρόφιλης κεφαλής ενοούν την επαφή με το νερό. Από την άλλη, οι υδρογονανθρακικές ουρές αλληλεπιδρούν μεταξύ τους δημιουργώντας μια υδρόφοβη εσωτερική περιοχή μεταξύ των υδρόφιλων κεφαλών που έρχονται σε επαφή με το υδάτινο εξωτερικό περιβάλλον και το υδάτινο εσωτερικό περιβάλλον του κυττάρου σχηματίζοντας τελικά τις μεμβρανικές φωσφολιπιδικές διπλοστιβάδες. Η διαδικασία σχηματισμού τους γίνεται γρήγορα και αυθόρμητα μέσα από μια διεργασία αυτοσυγκρότησης. Καθοριστικός παράγοντας για την δημιουργία της κυτταρικής μεμβράνης είναι η ύπαρξη υδρόφοβων αλληλεπιδράσεων. Επίσης η στενή επαφή μεταξύ των υδροφοβικών ουρών οφείλεται στο γεγονός ότι ασκούνται μεταξύ τους δυνάμεις van der Waals ενώ μεταξύ πολικών κεφαλών και μορίων νερού (από εσωτερικό και εξωτερικό περιβάλλον του κυττάρου) ασκούνται ηλεκτροστατικές δυνάμεις.

Οι περισσότερες μεμβρανικές διεργασίες όπως είναι η μεταφορά πληροφοριών διαμέσου της μεμβράνης πραγματοποιούνται από πρωτεΐνες. Οι πρωτεΐνες της λιπιδικής διπλοστιβάδας χωρίζονται σε δύο βασικές κατηγορίες. Η πρώτη κατηγορία είναι οι ενσωματωμένες πρωτεΐνες, που είναι εμβαπτισμένες στην υδρόφοβη περιοχή και στην πλειονότητά τους διαπερνούν την κυτταρική μεμβράνη. Η δεύτερη κατηγορία είναι οι περιφερειακές πρωτεΐνες που είτε προσδένονται στις υδρόφιλες κεφαλές των λιπιδίων είτε στα τμήματα των ενσωματωμένων πρωτεϊνών που προεξέχουν της κυτταρικής μεμβράνης, με την πρόσδεση να λαμβάνει χώρα στην εξωκυττάρια ή στην κυτταροπλασματική πλευρά της μεμβράνης και επιτυγχάνεται μέσω δεσμών υδρογόνου και ηλεκτροστατικών αλληλεπιδράσεων. Ο πιο κοινός τρόπος με τον οποίο οι ενσωματωμένες πρωτεΐνες διαπερνούν την λιπιδική διπλοστιβάδα είναι με δομές α-έλικας. Χαρακτηριστικό παράδειγμα τέτοιου τύπου πρωτεϊνών είναι οι υποδοχείς συζευγμένοι με G πρωτεΐνη (G Protein Coupled Receptors – GPCRs) οι οποίοι θα αναλυθούν στο επόμενο κεφάλαιο [4].

1.2 Υποδοχείς συζευγμένοι με G πρωτεΐνη (GPCRs)

1.2.1 Εισαγωγή στους GPCRs

Οι υποδοχείς που είναι συζευγμένοι με G πρωτεΐνη (G-Protein Coupled Receptors - GPCRs) είναι μια υπέρ-οικογένεια διαμεμβρανικών πρωτεϊνών και χωρίζονται σε έξι κατηγορίες A, B, C, D, E και F από τις οποίες οι τέσσερις (A, B, C και F) έχουν βρεθεί στον ανθρώπινο οργανισμό. Ο ρόλος τους βασίζεται στην μεταγωγή του σήματος από το εξωτερικό στο εσωτερικό του κυττάρου. Το ανθρώπινο γονιδίωμα κωδικοποιεί

περίπου 800 GPCRs με τις περισσότερες να ανήκουν στην κατηγορία A ή αλλιώς ροδοψίνη. Αναλυτικότερα η κατηγορία A περιέχει 719 GPCRs οι μισές εκ των οποίων είναι αισθητήριοι υποδοχείς και ενεργοποιούνται είτε από φερομόνες, μια χημική ουσία που σχετίζεται με την όσφρηση είτε από ροδοψίνες που σχετίζονται με την όραση. Οι άλλες μισές αποτελούν πολύ καλούς φαρμακευτικούς στόχους και ενεργοποιούνται από ορμόνες και νευροδιαβιβαστές. Παράδειγμα ενός υποδοχέα που ανήκει στην A κατηγορία των GPCR είναι ο υποδοχέας αδενοσίνης A_{2A} (Adenosine A_{2A} Receptor, ADORA2A) που αποτελεί στόχο στην ανοσο-ογκολογία και την νόσο του Parkinson. Η κατηγορία B αποτελείται από 48 GPCRs που διακρίνονται σε δύο υποκατηγορίες (secretin receptor, adhesion receptor). Η υποκατηγορία του υποδοχέα σεκρετίνης διαθέτει ένα μεγάλο εξωκυττάριο μέρος που κατά την διαδικασία της δημιουργίας συμπλόκου αλληλεπιδρά με το μεγαλύτερο κομμάτι του σηματοδοτικού μορίου ενώ μικρότερο είναι το μέρος του προσδέτη που έρχεται σε επαφή με το διαμεμβρανικό μέρος το υποδοχέα. Παράδειγμα υποδοχέα που ανήκει σε αυτή την κατηγορία των GPCR είναι ο υποδοχέας πεπτιδίου-1 τύπου γλυκαγόνης (Glucagon-Like Peptide-1 Receptor, GLP1R) ο οποίος αποτελεί στόχο για την θεραπεία του διαβήτη. Η κατηγορία C αποτελείται από 22 GPCRs και η F από 11 GPCRs με την δεύτερη να διακρίνεται στους frizzled υποδοχείς και στους smoothened υποδοχείς οι οποίοι ενεργοποιούνται από στεροειδή.

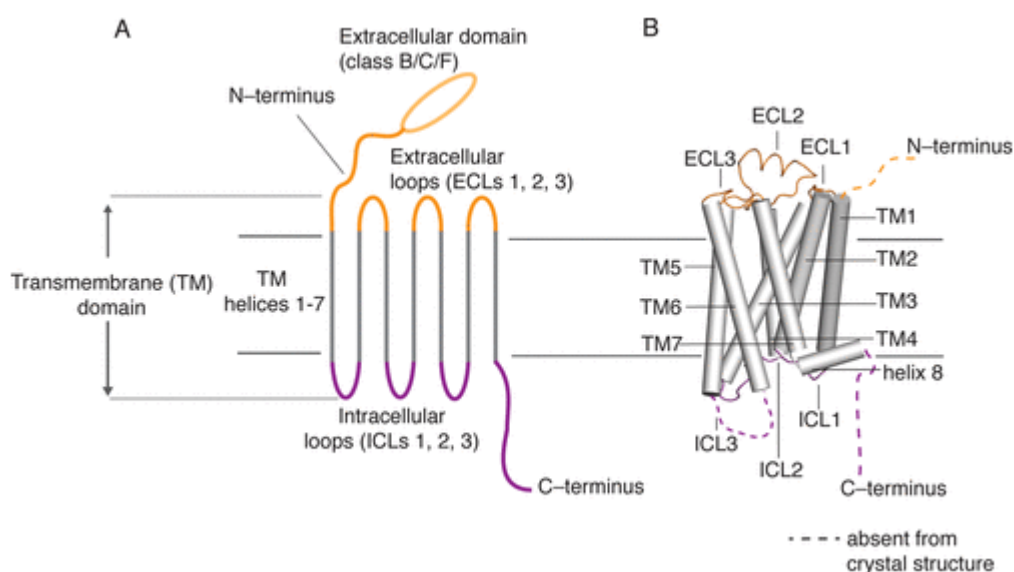
Οι GPCRs αποτελούν έναν από τους σημαντικότερους φαρμακευτικούς στόχους στο ανθρώπινο σώμα δεδομένου ότι περίπου το 1/3 (34%) των υπάρχοντων φαρμάκων που έχουν εγκριθεί από τον FDA (Food and Drug Administration) συνδέονται με μια GPCR πρωτεΐνη προκειμένου να επάγουν μια δραστηριότητα. Ένας υποδοχέας της οικογένειας των GPCR προσφέρει πολλά πλεονεκτήματα στον σχεδιασμό και στην ανακάλυψη νέων φαρμάκων. Για παράδειγμα, μέρος της ορθοστερικής θέσης δέσμευσης βρίσκεται εξωτερικά του κυττάρου. Στις περιπτώσεις που ο προσδέτης δεσμεύεται στην εξωκυττάρια περιοχή του υποδοχέα, αυτό διευκολύνει κατά πολύ τον σχεδιασμό και την ανάπτυξη φαρμάκων καθώς δεν χρειάζεται να ληφθεί υπόψη και να μελετηθεί όλο το μήκος της διαμεμβρανικής περιοχής του υποδοχέα. Επίσης η σύσταση σε αμινοξέα, το μέγεθος και το σχήμα της ορθοστερικής θέσης δέσμευσης την καθιστά επιθυμητή για την πρόσδεση ενώσεων με αγωνιστική ή ανταγωνιστική δράση. Επιπλέον οι GPCR αποτελούν πολύ καλούς φαρμακευτικούς στόχους καθώς είναι δυναμικές δομές που έχουν την ικανότητα να μεταβάλλουν την θέση των αελίκων τους και να αλλάζουν διαμορφώσεις κατά την δέσμευση είτε ενός αγωνιστή είτε ενός ανταγωνιστή μορίου.

Όλα τα παραπάνω σε συνδυασμό με το γεγονός πως μεγάλο μέρος των GPCR που εκφράζονται από τον άνθρωπο δεν αξιοποιούνται ακόμα ως φαρμακευτικοί στόχοι (συγκεκριμένα αξιοποιείται λιγότερο από το 25% των GPCR) προσδίδουν μεγάλες προοπτικές και εστιάζουν το ενδιαφέρον της επιστημονικής και ερευνητικής κοινότητας στην ανάλυση αυτών των δομών [5],[6],[7]. Εφόσον έγινε μια εισαγωγή στις GPCR και εξηγήθηκε η σημαντικότητα τους στην φαρμακευτική βιομηχανία, επόμενο είναι να μελετηθεί η δομή τους καθώς και η διαδικασία μεταγωγής του σήματος από το εξωτερικό προς το εσωτερικό του κυττάρου.

1.2.1 Δομή των GPCRs

Για αρκετά χρόνια η κρυσταλλογραφία ακτίνων X ήταν η βασική μέθοδος για τον προσδιορισμό της δομής των GPCR. Αυτό ανατράπηκε τα τελευταία χρόνια με την εισαγωγή της τεχνικής της κρύο-ηλεκτρονικής μικροσκοπίας (Cryogenic-Electron Microscopy, Cryo-EM) η οποία προσδιορίζει με πολύ καλή ανάλυση δομές πρωτεϊνών που μέχρι πρότινος ήταν άγνωστες. Από το 2017 και μετά επιλύθηκαν περισσότερες δομές GPCR στην ενεργή κατάσταση συζευγμένες με την ετεροτριμερή G πρωτεΐνη μέσω της μεθόδου Cryo-EM παρά με την κρυσταλλογραφία ακτίνων X. Αυτό οφείλεται κυρίως στην ικανότητα της Cryo-EM να προσδιορίζει την δομή μικρότερων πρωτεϊνών λόγω της ανάπτυξη των άμεσων ανιχνευτών ηλεκτρονίων καθώς και της εξέλιξης των αλγορίθμων επεξεργασίας εικόνας. Εκτιμάται πως η μέθοδος της Cryo-EM έχει ακόμα περιθώρια βελτίωσης και μελλοντικά θα συντελέσει στον χαρακτηρισμό πρωτεϊνών που η δομή τους δεν έχει προσδιοριστεί ακόμα συμβάλλοντας έτσι στην πρόοδο του σχεδιασμού φαρμάκων βάση δομής (Structure Based Drug Design) [5]. Περισσότερα για τις μεθόδους χαρακτηρισμού θα αναλυθούν στο κεφάλαιο 1.3. Παρακάτω αναλύονται τα βασικά δομικά μέρη μιας πρωτεΐνης της οικογένειας των GPCR.

Όλες οι πρωτεΐνες που ανήκουν στην υπέρ-οικογένεια των GPCR αποτελούνται από μια διαμεμβρανική περιοχή που είναι πολύ σημαντική για την διαδικασία μεταγωγής του σήματος από το εξωτερικό προς το εσωτερικό του κυττάρου. Συντίθεται από επτά α-έλικες (Transmembrane helices: TM1, TM2, TM3, TM4, TM5, TM6, TM7) οι οποίες είναι συνδεδεμένες με τρεις ενδοκυττάριους βρόγχους (Intracellular Loops: ICLs 1, 2, 3) και τρεις εξωκυττάριους βρόγχους (Extracellular Loops: ECLs 1, 2, 3). Πολλές κατηγορίες των GPCR όπως είναι η οικογένεια B, C και F αποτελούνται επιπρόσθετα από μια καρβοξυτελική ουρά (C-terminus) στην ενδοκυτταρική περιοχή και μια αμινοτελική ουρά (N-terminus) στην εξωκυτταρική περιοχή. Η εξωκυτταρική περιοχή συμβάλει στην αναγνώριση και δέσμευση των προσδετών στον υποδοχέα. Αντιθέτως η A κατηγορία των GPCR αποτελείται μόνο από την διαμεμβρανική περιοχή και συνεπώς οι προσδέτες δεσμεύονται κατευθείαν σε αυτή. Στην εικόνα 1.1 παρουσιάζεται η δομή ενός GPCR με εξωκυτταρική και ενδοκυτταρική περιοχή [8],[9].



Εικόνα 1.1: Δομή των GPCR (A) Με πορτοκαλί χρώμα φαίνεται η αμινοτελική ουρά και οι εξωκυττάριοι βρόγχοι, με γκρι χρώμα οι επτά διαμεμβρανικές α-έλικες και με μωβ χρώμα η καρβοξυτελική ουρά και οι εξωκυττάριοι βρόγχοι (B) Αναπαράσταση cartoon της (A) δομής [8].

1.2.2 Πορεία Μεταγωγής Σήματος

Για την μετάδοση της πληροφορίας στο εσωτερικό του κυττάρου ακολουθείτε μια σειρά βημάτων που είναι κοινά στην πλειοψηφία των περιπτώσεων. Η διαδικασία μεταγωγής σήματος που θα αναλυθεί στην συνέχεια αφορά υποδοχείς με επτά διαμεμβρανικές α-έλικες στις οποίες συγκαταλέγονται και οι GPCR. Η πορεία ξεκινάει με ένα ερέθισμα που απελευθερώνει το σηματοδοτικό μόριο το οποίο καλείται πρωτογενής αγγελιοφόρος ή πιο απλά προσδέτης.

Ακολουθεί ο σχηματισμός του συμπλόκου προσδέτη – υποδοχέα. Όπως αναφέρθηκε και στο κεφάλαιο 1.1, οι λιπιδικές διπλοστιβάδες έχουν χαμηλή διαπερατότητα σε μεγάλα μόρια καθώς και σε πολύ πολικά μόρια. Συνεπώς η πληροφορία που μεταφέρεται πρέπει να εισέλθει στο κύτταρο χωρίς τα ίδια τα μόρια και αυτό μπορεί να πραγματοποιηθεί μέσω ενσωματωμένων πρωτεϊνών. Στην εξωτερική πλευρά του υποδοχέα υπάρχει μια θέση δέσμευσης, η οποία αναγνωρίζει τον προσδέτη. Η δέσμευση του σηματοδοτικού μορίου στον υποδοχέα και η δημιουργία ενός συμπλόκου έχει ως αποτέλεσμα να αλλάξει η τριτοταγής και τεταρτοταγής δομή του υποδοχέα. Όπως προαναφέρθηκε, οι ενσωματωμένες πρωτεΐνες συχνά διαπερνούν την κυτταρική μεμβράνη ενώ στο ενδοκυτταρικό άκρο τους, κατά την διαδικασία ενεργοποίησης, είναι συζευγμένες με G πρωτεΐνες. Αλλαγή λοιπόν της στερεοδιάταξης του υποδοχέα οδηγεί σε αλλαγή της διαμόρφωσης στην κυτταροπλασματική περιοχή με αποτέλεσμα να ενεργοποιείται η G πρωτεΐνη η οποία με την σειρά της πυροδοτεί μια σειρά αλληλεπιδράσεων.

Όταν η G πρωτεΐνη είναι απενεργοποιημένη, είναι προσδεμένη στην μεμβράνη και υπάρχει ως ετεροτριμερές αποτελούμενη από τρεις διαφορετικές υπομονάδες, τις α, β και γ που συχνά αναφέρονται ως Gα, Gβ, Gγ. Στην υπομονάδα Gα είναι δεσμευμένη η διφωσφορική γουανοσίνη (GDP). Το σύμπλοκο προσδέτης – υποδοχέας που δημιουργείται, αλληλεπιδρά με το ετεροτριμερές με αποτέλεσμα να αποδεσμεύσει την GDP και να δώσει την δυνατότητα στην τριφωσφορική γουανοσίνη (GTP) να δεσμευτεί με την υπομονάδα Gα. Παράλληλα το ετεροτριμερές διασπάται στις μονάδες Gα και Gβγ. Η αποσύνθεση του ετεροτριμερούς αποτελεί το σήμα ότι έχει δημιουργηθεί το σύμπλοκο προσδέτη-υποδοχέα. Το σήμα αυτό δεν είναι αμελητέο αν σκεφτεί κανείς ότι για κάθε σηματοδοτικό μόριο που προσδένεται, εκατοντάδες ετεροτριμερείς G πρωτεΐνες διασπώνται. Παρόλα αυτά, η πληροφορία που παράγεται δεν επαρκεί ώστε να προκαλέσει την απόκριση του κυττάρου και σε αυτό το σημείο εισάγεται ο ρόλος του δεύτερου αγγελιοφόρου ο οποίος επιτελεί την διαδικασία της αναμετάδοσης και της ενίσχυσης του σήματος. Κάποιοι από τους σημαντικότερους δεύτερους αγγελιοφόρους είναι η κυκλική μονοφωσφορική αδενοσίνη (κυκλική AMP ή cAMP), η κυκλική μονοφωσφορική γουανοσίνη (κυκλική GMP ή cGMP), η 1,4,5 - τριφωσφορική ινοσιτόλη (IP₃), η διακυλογλυκερόλη (DAG) και το ιόν ασβεστίου.

Μετά από την ενεργοποίηση της πορείας μεταγωγής πληροφορίας και την ολοκλήρωση της διαδικασίας που αναλύθηκε παραπάνω, είναι σημαντικό η μεταγωγή του σήματος να λάβει τέλος. Αυτό επιτυγχάνεται αυθόρμητα μέσω της υδρόλυσης της GTPάσης που είναι δεσμευμένη στην υπομονάδα Gα της G πρωτεΐνης. Τα προϊόντα της υδρόλυσης

είναι το ανόργανο ορθοφωσφορικό (Pi) και το GDP. Πλέον η υπομονάδα Ga με δεσμευμένο το GDP συνδέεται και πάλι με το διμερές Gβγ και προκύπτει η ετεροτριμερής G πρωτεΐνη. Επίσης το σηματοδοτικό μόριο αποσπάται από τον υποδοχέα ο οποίος επιστρέφει στην απενεργοποιημένη κατάσταση. Η διαδικασία τερματισμού του σήματος είναι εξαιρετικά σημαντική καθώς επιτρέπει στα κύτταρα να δεχτούν νέα σήματα ενώ ταυτόχρονα αποφεύγεται η ανεξέλεγκτη κυτταρική αύξηση και ο καρκίνος [4],[10],[11].

1.3 Μέθοδοι Προσδιορισμού Δομής

Η κρυσταλλογραφία ακτινών X, η κρύο-ηλεκτρονική μικροσκοπία και η φασματοσκοπία πυρηνικού μαγνητικού συντονισμού αποτελούν τις τρεις βασικές μεθόδους που χρησιμοποιούνται ευρέως από την επιστημονική κοινότητα για τον προσδιορισμό της δομής ενός μακρομορίου. Παρακάτω θα αναλυθεί η λειτουργία και τα βασικά τους χαρακτηριστικά.

1.3.1 Κρυσταλλογραφία Ακτινών X

Ο ομοιοπολικός δεσμός είναι ο δεσμός του οποίου τα άτομα απέχουν την μικρότερη απόσταση στον χώρο από οποιαδήποτε άλλα άτομα και συγκεκριμένα έχει μήκος 0,12nm. Για να καταφέρει κανείς να απεικονίσει με αποτελεσματικότητα την δομή ενός μορίου θα πρέπει η μέθοδος που χρησιμοποιεί να είναι ικανή να διακρίνει ως ξεχωριστά δύο τέτοια άτομα (δηλαδή να έχει διακριτική ικανότητα μικρότερη ή ίση από 0,12nm). Μια από τις παραμέτρους που επηρεάζει την διακριτική ικανότητα της μεθόδου είναι το μήκος κύματος (λ) της προσπίπτουσας ακτινοβολίας και δίνεται από την σχέση 1. Σύμφωνα με αυτήν την σχέση, η τιμή του μήκους κύματος (λ) της ακτινοβολίας δεν πρέπει να ξεπερνάει τα 0,24nm. Η τιμή αυτή ταυτίζεται με την περιοχή του ηλεκτρομαγνητικού φάσματος στο οποίο ανήκουν οι ακτίνες X.

$$LR = \frac{\lambda}{2} \quad (1)$$

Το φαινόμενο που χρησιμοποιείται για τον προσδιορισμό της δομής των μακρομορίων είναι η περίθλαση καθώς είναι η μόνη που μπορεί να δώσει ακριβής και αξιόπιστες απεικονίσεις σε ατομικό επίπεδο. Πριν την εφαρμογή της μεθόδου απαιτείται η επεξεργασία του δείγματος ώστε να είναι σε μορφή κρυστάλλου.

Για την δημιουργία ενός κρυστάλλου πρέπει αρχικά να γίνει η μετατροπή του μακρομορίου από υδατικό διάλυμα σε στερεή μορφή που είναι βασικό χαρακτηριστικό ενός κρυστάλλου. Μια μέθοδος είναι να κατακρημνιστούν τα διαλυμένα μόρια του διαλύματος με ελεγχόμενο τρόπο. Μια άλλη συχνά χρησιμοποιούμενη τεχνική είναι να αφαιρεθεί μέρος του διαλύτη με αποτέλεσμα να αυξηθεί η συγκέντρωση του μακρομορίου στην υπέρκορη κατάσταση. Οι συνθήκες κρυστάλλωσης των μακρομορίων παίζουν καθοριστικό ρόλο για την επιτυχή ολοκλήρωση της διαδικασίας. Πολλές είναι οι παράμετροι που λαμβάνονται υπόψη κατά την δημιουργία ενός κρυστάλλου με σημαντικότερη αυτής της βιοχημικής και δομικής καθαρότητας του

δείγματος. Ανάλογα με την περίπτωση αλλάζουν, οι συνθήκες, οι διαλύτες και τα διαλύματα αλάτων που χρησιμοποιούνται για να επιτευχθεί βέλτιστη κρυστάλλωση.

Πρωταρχικό και απαραίτητο στάδιο για την δημιουργία ενός κρυστάλλου είναι η πυρήνωση δηλαδή ο σχηματισμός πυρήνων. Το στάδιο αυτό έχει χαμηλή πιθανότητα να συμβεί και εμφανίζεται μόνο όταν η συγκέντρωση του δείγματος ξεπεράσει κατά πολύ την εγγενή του διαλυτότητα S^0 δηλαδή φτάσει σε κατάσταση υπερκορεσμού. Δεύτερο στάδιο και με υψηλότερη πιθανότητα να συμβεί είναι η διάδοση του μεγέθους του κρυστάλλου. Για την διευκόλυνση των παραπάνω σταδίων χρησιμοποιούνται διάφορες μέθοδοι με κυριότερες αυτές της διάχυσης ατμών και της μικροδιαπίδυσης. Σκοπός και των δύο τεχνικών είναι η μείωση της διαλυτότητας των μακρομορίων με ελεγχόμενο τρόπο. Κοινά βήματα και στις δύο τεχνικές είναι η διάλυση του μορίου σε ρυθμιστικό διάλυμα και ο διαχωρισμός του από τον μεγαλύτερο όγκο του διαλύτη ο οποίος τοποθετείται σε μια δεξαμενή. Στην συνέχεια τα διαλύματα του μακρομορίου και της δεξαμενής τοποθετούνται σε ένα κλειστό σύστημα. Στην περίπτωση της διάχυσης ατμών γίνεται μεταφορά των διαλυτών λόγω διαφοράς τάσης ατμών που υπάρχει μεταξύ της δεξαμενής και του δείγματος ενώ στην περίπτωση της μικροδιαπίδυσης οι διαλύτες μεταφέρονται μέσω μιας ημιδιαπερατής μεμβράνης λόγω ωσμωτικών πιέσεων. Αφού προετοιμαστεί ο κρύσταλλος, μπορεί πλέον να ακτινοβοληθεί με ακτίνες X και μέσω του φαινομένου της περίθλασης να προσδιοριστεί η δομή του μακρομορίου.

Οι ανακλάσεις που προκύπτουν από την ακτινοβολήση του δείγματος εξάγουν μια εικόνα περίθλασης ή όπως αναφέρεται ένα περιθλασίγραμμα, η ανάλυση του οποίου οδηγεί στον προσδιορισμό της δομής του μακρομορίου. Σύμφωνα με τον νόμο του Bragg όπως περιγράφεται από την σχέση 2, η γωνία Bragg (θ) μεταβάλλεται αντιστρόφως ανάλογα από τις αποστάσεις των ανακλώντων επιπέδων (d). Αυτό πρακτικά σημαίνει ότι όσο μεγαλύτερη είναι η απόσταση μεταξύ των επιπέδων που αντανακλούν τις ακτίνες X τόσο μικρότερες θα είναι οι γωνίες περίθλασης που ανακλώνται από τον κρύσταλλο. Η πληροφορία αυτή μπορεί να αξιοποιηθεί για τον προσδιορισμό των μηκών της μοναδιαίας κυψελίδας του κρυστάλλου. Επιπρόσθετα ο προσανατολισμός που έχουν οι ανακλάσεις μετά την περίθλαση δίνει πληροφορία για τον προσανατολισμό των αξόνων της μοναδιαίας κυψελίδας. Τέλος πληροφορίες για τα μήκη και τις γωνίες των πλευρών της μοναδιαίας κυψελίδας λαμβάνονται μέσω της παρατήρησης της συμμετρίας του περιθλασιγράμματος. Όλα τα παραπάνω συνδυαστικά, προσδιορίζουν τελικά την δομή του μορίου προς ανάλυση [12].

$$\frac{2 \sin \theta}{\lambda} = \frac{n}{d} \quad (2)$$

1.3.2 Κρύο-Ηλεκτρονική Μικροσκοπία (Cryogenic-Electron Microscopy - Cryo-EM)

Η κρύο-ηλεκτρονική μικροσκοπία είναι ακόμα μια τεχνική που έχει εισέλθει δυναμικά στο πεδίο της δομικής βιολογίας και μπορεί να χρησιμοποιηθεί στον αποδοτικό χαρακτηρισμό της δομής μακρομορίων. Τα τελευταία χρόνια η πρόοδος της τεχνολογίας και συγκεκριμένα η ανάπτυξη των άμεσων ανιχνευτών ηλεκτρονίων (Direct Electron Detectors, DED) σε συνδυασμό με την εξέλιξη αλγορίθμων επεξεργασίας εικόνας έχουν εδραιώσει αυτήν την μεθοδολογία για χρήση στον

χαρακτηρισμό μακρομορίων. Μια από τις βασικές τεχνικές της Cryo-EM μεθόδου είναι η ανάλυση μεμονωμένων σωματιδίων (Single Particle Analysis, SPA) η αρχή λειτουργίας της οποίας βασίζεται στην λήψη πολλών προβολών διαφορετικού προσανατολισμού για πολλά αντίγραφα του μορίου προς χαρακτηρισμό ενώ το τελικό αποτέλεσμα προκύπτει από τον μέσο όρο αυτών των προβολών. Επειδή ακριβώς η δέσμη ηλεκτρονίων δεν περιορίζεται σε μια περιοχή ενός μορίου αλλά κατανέμεται σε διαφορετικές περιοχές πολλών αντιγράφων του ίδιου μορίου, έχει ως αποτέλεσμα να αυξάνεται η μέγιστη ανεκτή δόση ηλεκτρονίων που επιτρέπεται ώστε να μην υποστεί βλάβη το δείγμα. Επιπλέον οι πανομοιότυπες προβολές διαφορετικών αντιγράφων συμπληρώνουν η μια την άλλη αυξάνοντας σημαντικά το SNR (Signal to Noise Ratio) της εικόνας [13],[14]. Μια ακόμα βασική τεχνική είναι η κρύο-ηλεκτρονική τομογραφία (Cryogenic-Electron Tomography, Cryo-ET) στην οποία η πληροφορία που απαιτείται για την ανακατασκευή της τρισδιάστατης εικόνας του μορίου προς ανάλυση προέρχεται από κεκλιμένες προβολές προς μια περιοχή του μορίου. Για την υλοποίηση της μεθόδου δεν χρειάζονται πολλά αντίγραφα και χρησιμοποιείται συνήθως για τον χαρακτηρισμό ετερογενών δομών όπως είναι οι κυτταρικοί όγκοι. Παράλληλα αναπτύσσονται και νέες τεχνικές όπως αυτές της περίθλασης μικρό-ηλεκτρονίου (Micro-Electron Diffraction, Micro-ED) και η ηλεκτρονική μικροσκοπία μετάδοσης κρύο-σάρωσης (Cryogenic-Scanning Transmission Electron Microscopy, Cryo-STEM) [14].

Η διαδικασία που χρησιμοποιείται από την αρχή μέχρι και την εξαγωγή του αποτελέσματος μπορεί να διαφέρει ανάλογα την εφαρμογή αλλά ο κορμός της μεθοδολογίας παραμένει ο ίδιος. Πρωταρχικό βήμα είναι να απομονωθεί και να καθαριστεί το δείγμα. Μεγάλο πλεονέκτημα της μεθόδου σε αντίθεση με την κρυσταλλογραφία ακτίνων-X και την φασματοσκοπία πυρηνικού μαγνητικού συντονισμού είναι ότι δεν απαιτείται μεγάλη ποσότητα δείγματος (συνήθως μερικά ml σε συγκέντρωση 1 mg/ml) για την επιτυχή υλοποίηση της διαδικασίας. Επόμενο βήμα είναι να ελεγχθεί η ποιότητα του δείγματος και συγκεκριμένα η διαμορφωτική του ομοιομορφία κάτι που επιτυγχάνεται με την χρήση αρνητικής χρώσης στο δείγμα. Εάν η ποιότητα του δείγματος είναι καλή, η διαδικασία προχωράει με την συλλογή και ανάλυση των δεδομένων καθώς και με την τρισδιάστατη ανακατασκευή της εικόνας που θα φανερώσει την τελική δομή του μακρομορίου [15].

Η μέθοδος Cryo-EM σε πολλές περιπτώσεις φαίνεται πως έχει αντικαταστήσει τις παραδοσιακές τεχνικές όπως για παράδειγμα την κρυσταλλογραφία ακτίνων X. Βασικά πλεονεκτήματα της μεθόδου που την έχουν εδραιώσει στον τομέα της δομικής βιολογίας είναι το γεγονός πως κατά τον προσδιορισμό της δομής του, το δείγμα δεν χρειάζεται να είναι σε μορφή κρυστάλλου, μικρή ποσότητα δείγματος είναι αρκετή για την επιτυχή εφαρμογή της διαδικασίας, επιτρέπει την απεικόνιση πολλών διαφορετικών διαμορφωτικών καταστάσεων του δείγματος, είναι κατάλληλη και για μακρομόρια με μεγάλο μοριακό βάρος ενώ ταυτόχρονα η ακτινοβολία δεν καταστρέφει την δομή και την λειτουργία της δομής προς ανάλυση. Η ραγδαία ανάπτυξη των τελευταίων ετών της Cryo-EM μεθόδου τόσο σε επίπεδο λογισμικού όσο και υλικού κατέστησε εφικτή την απεικόνιση δομών σε πολύ υψηλή ανάλυση. Οι περισσότερες δομές που προσδιορίστηκαν με την μέθοδο Cryo-EM είναι σε ανάλυση 3.0-3.5 Angstrom ενώ οι μέθοδοι έχει καταφέρει να επιλύσει δομές σε πολύ υψηλή ανάλυση της τάξης των 1.2 Angstrom. Όλα τα παραπάνω έχουν ως αποτέλεσμα η μέθοδος της Cryo-EM να αποτελεί ένα βασικό εργαλείο στον τομέα της σχεδίασης φαρμάκων βάση δομής (Structure Based Drug Design, SBDD) [16],[17]. Η Cryo-EM μέθοδος είχε

τεράστια επίδραση στην διασάφηση της δομής μεμβρανικών πρωτεϊνών όπως είναι οι διάλυτοι ιόντων και οι GPCR συμβάλλοντας ουσιαστικά στην πρόοδο της φαρμακευτικής έρευνας καθώς διασαφηνίζει την δομή σημαντικών στόχων που μέχρι πρότινος ήταν δύσκολο να προσδιοριστούν. Χαρακτηριστικό παράδειγμα αποτελεί η πρωτεΐνη που αναλύεται στην παρούσα εργασία. Η δομή της GHRH-R σε σύζευξη με το πεπτίδιο GHRH και την Gs πρωτεΐνη έχει προσδιοριστεί μέσω της Cryo-EM μεθόδου [18].

1.3.3 Φασματοσκοπία Πυρηνικού Μαγνητικού Συντονισμού (Nuclear Magnetic Resonance Spectroscopy - NMR)

Η τελευταία μέθοδος που θα μελετηθεί σε αυτό το κεφάλαιο και χρησιμοποιείται για την διασάφηση της δομής μακρομορίων είναι η φασματοσκοπία πυρηνικού μαγνητικού συντονισμού (Nuclear Magnetic Resonance Spectroscopy - NMR). Παρακάτω θα αναλυθεί η αρχή λειτουργίας της μεθόδου ενώ θα γίνει και μια σύντομη αναφορά στην οργανολογία του συστήματος.

Η βασική αρχή της μεθόδου NMR βασίζεται στο γεγονός πως πολλοί πυρήνες έχουν μια αυτοστροφορμή ή αλλιώς spin και περιστρέφονται με τυχαίο προσανατολισμό. Όταν οι πυρήνες αυτοί εκτεθούν σε ένα εξωτερικό μαγνητικό πεδίο (H_0) τότε ανάλογα με τον κβαντικό αριθμό μαγνητικού spin l (όπου $l=0, 1/2, 1, 3/2, \dots$) που έχουν, υπάρχουν $2l + 1$ δυνατοί προσανατολισμοί μέσα στο μαγνητικό πεδίο. Συγκεκριμένα για τους πυρήνες του υδρογόνου (1H) και του άνθρακα (^{13}C) ισχύει ότι $l = 1/2$ συνεπώς μπορούν να πάρουν σε σχέση με το εξωτερικό πεδίο, δύο διαφορετικούς προσανατολισμούς. Ο ένας θα είναι παράλληλα με τον προσανατολισμό του H_0 και θα χαρακτηρίζεται από χαμηλότερη ενέργεια ενώ ο άλλος θα είναι αντιπαράλληλος του H_0 και θα χαρακτηρίζεται από υψηλότερη ενέργεια. Όταν οι πυρήνες του 1H ή του ^{13}C ακτινοβοληθούν με μαγνητική ακτινοβολία, παρατηρείται μετάβαση των πυρήνων από κατάσταση χαμηλότερης ενέργειας σε κατάσταση υψηλότερης ενέργειας με αντίστοιχη αναστροφή της αυτοστροφορμής του πυρήνα. Το παραπάνω φαινόμενο ονομάζεται συντονισμός και παρατηρείται σε πυρήνες με περιττό αριθμό πρωτονίων καθώς και σε πυρήνες με περιττό αριθμό νετρονίων. Για να επιτευχθεί συντονισμός απαιτείται κατάλληλη συχνότητα η οποία εξαρτάται τόσο από το ίδιο το είδος του πυρήνα όσο και από την ισχύ του H_0 . Σημαντική παράμετρο αποτελεί ο γυρομαγνητικός λόγος (γ), μια χαρακτηριστική σταθερά για κάθε πυρήνα που εκφράζεται από την σχέση 3:

$$\gamma = \frac{2\pi\nu}{H_0} \quad (3)$$

Από την παραπάνω σχέση, προκύπτει η βασική εξίσωση Larmor για την μέθοδο NMR στην οποία η συχνότητα συντονισμού (ν) δίνεται από την σχέση 4:

$$\nu = \frac{\gamma H_0}{2\pi} \quad (4)$$

Για τον προσδιορισμό της δομής των οργανικών ενώσεων εισάγεται μια νέα έννοια που ονομάζεται χημική μετατόπιση. Κάθε πυρήνας ενός μορίου έχει χαρακτηριστικό χημικό περιβάλλον γύρω του. Ειδικότερα έχει χαρακτηριστική ηλεκτρονιακή κατανομή η οποία τον προστατεύει από την επίδραση του εξωτερικού μαγνητικού

πεδίου H_0 δημιουργώντας ένα μαγνητικό πεδίο με προσανατολισμό αντίθετο από αυτό του H_0 . Από τα προηγούμενα συμπεραίνεται πως το πραγματικό εξωτερικό μαγνητικό πεδίο (H) που εφαρμόζεται είναι μικρότερο αν ληφθεί υπόψη και η προστασία (σ) που παρέχει στους πυρήνες η ηλεκτρονιακή πυκνότητα. Το πραγματικό εξωτερικό μαγνητικό πεδίο εκφράζεται από την σχέση 5:

$$H = H_0 - \sigma H_0 \quad (5)$$

Από τα παραπάνω προκύπτει πως η ηλεκτρονιακή πυκνότητα και η προστασία (σ) που προσδίδεται είναι διαφορετική για κάθε πυρήνα και συνεπώς η ένταση του πραγματικού εξωτερικού μαγνητικού πεδίου (H) είναι χαρακτηριστική για κάθε πυρήνα. Ως αποτέλεσμα πυρήνες με διαφορετικό χημικό περιβάλλον συντονίζονται σε διαφορετικές συχνότητες και δίνουν διαφορετικές κορυφές στο φάσμα NMR που προκύπτει. Η παραπάνω ιδιότητα κάνει εφικτό τον προσδιορισμό της δομής οργανικών ενώσεων και σε αυτή βασίζεται η μέθοδος NMR.

Παλαιότερα το όργανο NMR που χρησιμοποιούνταν ήταν το φασματόμετρο συνεχούς κύματος (Continuous Wave, CW) ενώ πλέον χρησιμοποιούνται κατά βάση τα παλμικά φασματόμετρα ή αλλιώς μετασχηματισμού Fourier (Fourier Transform-NMR, FT-NMR), στα οποία το δείγμα προς μελέτη ακτινοβολείται με περιοδικούς παλμούς ραδιοσυχνοτήτων και έπειτα εφαρμόζεται μετασχηματισμός Fourier. Τα κυριότερα τμήματα των δύο παραπάνω τύπων οργάνων NMR είναι κοινά και περιλαμβάνουν μια γεννήτρια ραδιοσυχνοτήτων που παράγει RF ακτινοβολία, έναν προ-ενισχυτή, έναν μαγνήτη ο οποίος αποτελεί το σημαντικότερο και ακριβότερο τμήμα του συστήματος, αισθητήρες, έναν ανιχνευτή ραδιοσυχνοτήτων, έναν ενισχυτή και τέλος το σύστημα καταγραφής του σήματος [19].

1.4 Υποδοχέας Ορμόνης Απελευθέρωσης Αυξητικής Ορμόνης

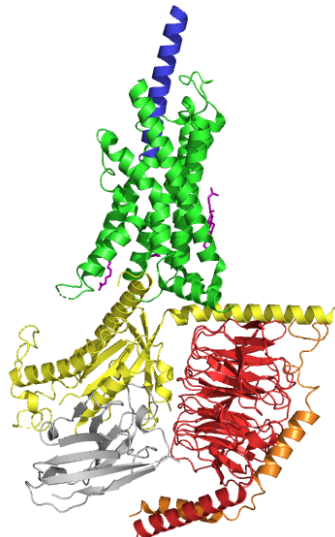
Η οικογένεια B των GPCR κατέχει σημαντικό ρόλο σε πολλές φυσιολογικές και παθοφυσιολογικές διαδικασίες. Η δομή των πρωτεϊνών που ανήκουν στην συγκεκριμένη οικογένεια αποτελείται από επτά διαμεμβρανικές α -έλικες, τρεις ενδοκυττάριους βρόγχους, τρεις εξωκυττάριους βρόγχους, μια αμινοτελική ουρά στην εξωκυττάρια περιοχή και μια καρβοξυτελική ουρά στην ενδοκυττάρια περιοχή του υποδοχέα, όπως ακριβώς αναλύθηκε και στην δομή των GPCR σε προηγούμενο κεφάλαιο. Η οικογένεια B των GPCR αποτελείται από δεκαπέντε υποδοχείς οι οποίοι είναι οι εξής:

- Calcitonin Receptor (CTR)
- Calcitonin Receptor-Like Receptor (CRLR)
- Corticotropin-Releasing Hormone Receptors type 1 (CRF1R)
- Corticotropin-Releasing Hormone Receptors type 2 (CRF2R)
- Glucagon Receptor (GCGR)
- Gastric Inhibitory Polypeptide Receptor (GIPR)
- Glucagon-Like Peptide 1 Receptor (GLP1R)
- Glucagon-Like Peptide 2 Receptor (GLP2R)
- Growth Hormone Releasing Hormone Receptor (GHRHR)

- Vasoactive Intestinal Peptide Receptors type 1 (VPAC1R)
- Vasoactive Intestinal Peptide Receptors type 2 (VPAC2R)
- Pituitary Adenylate Cyclase-Activating Protein Receptor (PAC1R)
- Parathyroid Hormone Receptors type 1 (PTH1R)
- Parathyroid Hormone Receptors type 2 (PTH2R)
- Secretin Receptor (SCTR)

Οι υποδοχείς της Β οικογένειας των GPCR αποτελούν σημαντικούς φαρμακευτικούς στόχους καθώς σε αυτούς προσδένονται μεγάλα πεπτίδια τα οποία εμπλέκονται σε πολλές φυσιολογικές και παθοφυσιολογικές διαδικασίες [20].

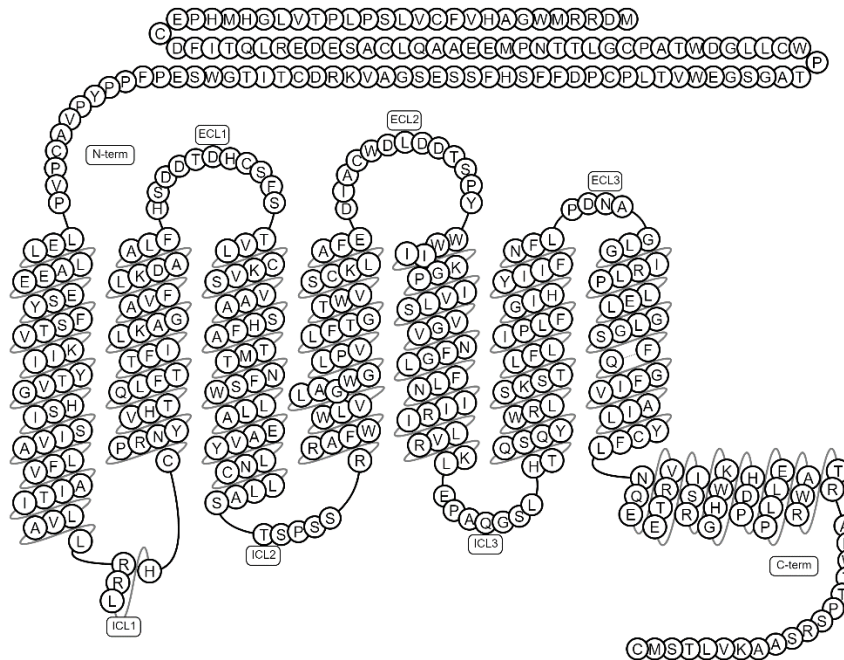
Παρακάτω γίνεται αναφορά για τον υποδοχέα της ορμόνης απελευθέρωσης αυξητικής ορμόνης (Growth Hormone Releasing Hormone Receptor - GHRHR) που αποτελεί την πρωτεΐνη - στόχο της εργασίας. Ο GHRH-R ανήκει στην Β οικογένεια των GPCRs και συγκεκριμένα στην υποκατηγορία Β1 των υποδοχέων σεκρετίνης (secretin receptor). Στην εικόνα 1.2 φαίνεται το σύμπλοκο GHRH-GHRHR-Gs που δημιουργείται κατά την ενεργοποίηση του υποδοχέα. Η εξωκυττάρια περιοχή του υποδοχέα αναγνωρίζει το πεπτίδιο και αλληλεπιδρά με αυτό. Ως αποτέλεσμα ο υποδοχέας υφίσταται διαμορφωτικές αλλαγές που δημιουργούν τις κατάλληλες συνθήκες για την πρόσδεση της Gs πρωτεΐνης στην ενδοκυτταρική περιοχή του. Το σύμπλοκο GHRH-GHRHR-Gs που δημιουργείται είναι ασταθές δυσκολεύοντας την διαδικασία επίλυσης της κρυσταλλικής του δομής. Για τον λόγο αυτό, κατά την διάρκεια του χαρακτηρισμού του συμπλόκου χρησιμοποιήθηκε το Nanobody 35 (Nb 35) το οποίο σταθεροποιεί το σύμπλοκο. Επιπλέον μαζί με τον υποδοχέα έχει συγκρυσταλλωθεί ένα μόριο νερού στην ορθοστερική θέση, δύο μόρια νερού στην κοιλότητα πρόσδεσης της Gs πρωτεΐνης καθώς και μερικά λιπίδια [21].



Εικόνα 1.2: Αναπαράσταση του συμπλόκου GHRHR-GHRH-Gs που δημιουργείται κατά την ενεργοποίηση του υποδοχέα. Πράσινο: GHRH-R, Μπλε: GHRH, Κίτρινο: Ga, Κόκκινο: Gβ, Πορτοκαλί: Gγ, Γκρι: Nanobody 35, Μωβ: Λιπίδια.

Στην εικόνα 1.3 φαίνεται η δομή του GHRH-R καθώς και τα αμινοξέα από τα οποία αποτελείται η πολυπεπτιδική αλυσίδα. Η δομή αυτή έρχεται σε συμφωνία με την δομή που περιεγράφηκε προηγουμένως για την δομή των GPCRs. Διακρίνονται οι επτά

διαμεμβρανικές α -έλικες (από αριστερά προς τα δεξιά TM1-TM7), οι τρεις εξωκυττάριοι βρόγχοι (ECL1-ECL3), οι τρεις ενδοκυττάριοι βρόγχοι (ICL1-ICL3), η εξωκυττάρια αμινοτελική ουρά (N-Terminus ή αλλιώς Extracellular Domain-ECD) και η ενδοκυττάρια καρβοξυτελική ουρά (C-Terminus). Το άνω διαμεμβρανικό κομμάτι του υποδοχέα μαζί με τους εξωκυττάριους βρόγχους καλείται J-Domain [21].

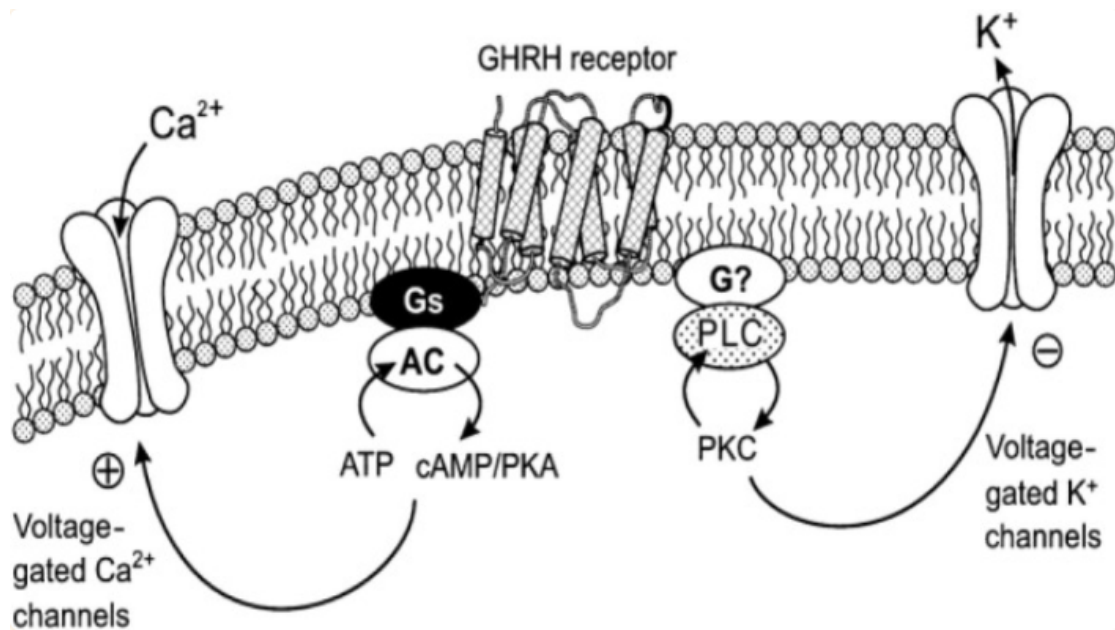


Εικόνα 1.3: Δομή του υποδοχέα ορμόνης απελευθέρωσης αυξητικής ορμόνης. Φαίνονται οι επτά διαμεμβρανικές α -έλικες που συνδέονται με τους ενδοκυττάριους και εξωκυττάριους βρόγχους καθώς και το αμινοτελικό και καρβοξυτελικό άκρο του υποδοχέα (εικόνα από GPCR database).

1.5 Ορμόνη Απελευθέρωσης Αυξητικής Ορμόνης

Η αυξητική ορμόνη (Growth Hormone - GH) ή αλλιώς σωματοτροπίνη είναι ένα πολυπεπτιδίο αποτελούμενο από 191 αμινοξέα με μοριακό βάρος περίπου 22 kDa, το οποίο εκκρίνεται από τα σωματοτρόπα κύτταρα της πρόσθιας υπόφυσης. Η GH κωδικοποιείται από το γονίδιο GH-1 που βρίσκεται στο 17^ο χρωμόσωμα του ανθρώπου [22],[23]. Η παραγωγή της GH ρυθμίζεται από πολλούς παράγοντες, βασικότεροι εκ των οποίων είναι οι εξής τρεις ορμόνες: Η ορμόνη απελευθέρωσης αυξητικής ορμόνης (Growth Hormone Releasing Hormone – GHRH) ή αλλιώς σωματοκρινίνη που παράγεται στον υποθάλαμο και μαζί με την αυξητική ορμόνη έκκρισης (Growth Hormone Secretagogue - GHS) η οποία παράγεται στον γαστρεντερικό σωλήνα, διεγείρουν την λειτουργία των σωματοτρόπων κυττάρων και ως επόμενο την έκκριση της GH. Αντιθέτως η ανασταλτική ορμόνη της αυξητικής ορμόνης (Growth Hormone Inhibiting Hormone – GHIH) ή αλλιώς σωματοστατίνη (Somatostatin - SS), αναστέλλει την έκκριση GHRH, την λειτουργία των σωματοτρόπων κυττάρων και κατά συνέπεια την παραγωγή της GH [22],[24].

Αναφορικά για την GHRH, είναι ένα πεπτιδίο που αποτελείται συνολικά από 44 αμινοξέα, από τα οποία τα πρώτα 29 έχουν βιολογική δράση και κωδικοποιείται από το γονίδιο που βρίσκεται στο 20^ο χρωμόσωμα του ανθρώπου [23]. Η σύνθεση της GHRH γίνεται στον υποθάλαμο και συγκεκριμένα στον τοξοειδή πυρήνα. Μετά την σύνθεση της απελευθερώνεται και οδηγείται μέσω των υποθαλαμικών-υποφυσικών αιμοφόρων αγγείων προς τα σωματοτρόπα κύτταρα της υπόφυσης. Εκεί προσδέεται με τον GHRH-R και ακολουθείται η διαδικασία όπως περιεγράφηκε στο κεφάλαιο της μεταγωγής σήματος. Η υπομονάδα G_s της G_s πρωτεΐνης ενεργοποιεί την αδενυλυλοκυκλάση, παράγεται cAMP, ενεργοποιείται η πρωτεϊνική κινάση A (Protein Kinase A – PKA) με αποτέλεσμα το άνοιγμα καναλιών Ca^{2+} που προκαλεί την έκκριση της GH [23],[25]. Η σχηματική αναπαράσταση της παραπάνω διαδικασίας φαίνεται στην εικόνα 1.4.

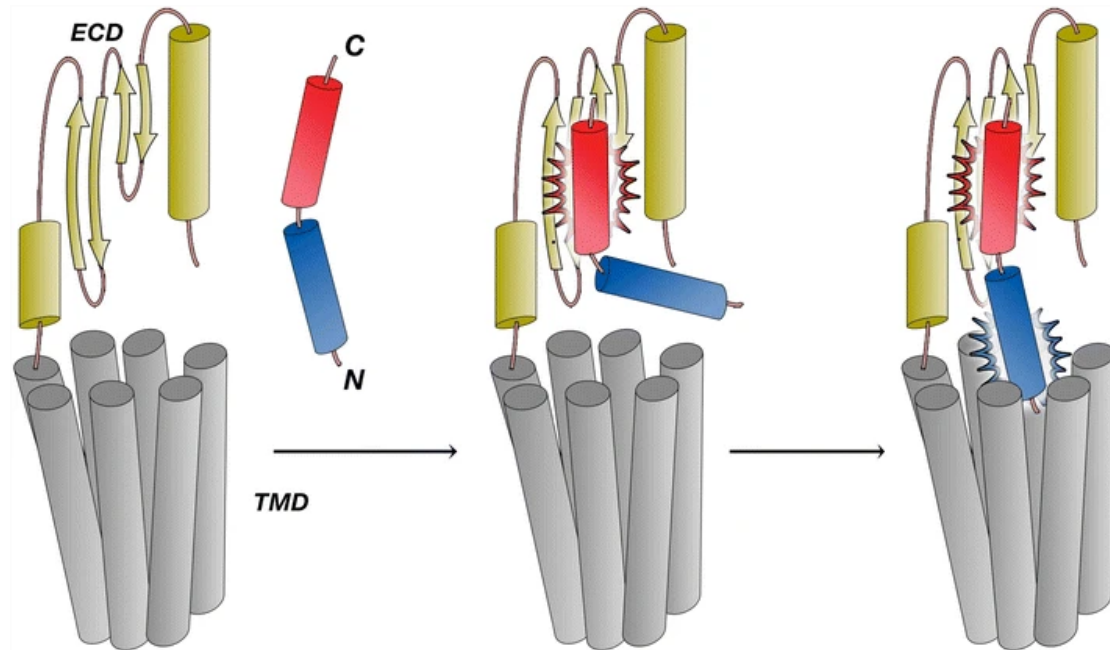


Εικόνα 1.4: Αναπαράσταση του μηχανισμού έκκρισης της GH μέσω της παραγωγής cAMP και του ανοίγματος των διαύλων Ca^{2+} [25].

1.6 Μηχανισμός πρόσδεσης του πεπτιδίου (GHRH) στον υποδοχέα (GHRH-R)

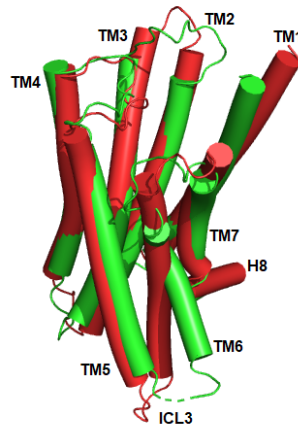
Τα πεπτιδία των GPCR της οικογένειας B χαρακτηρίζονται από μια καρβοξυτελική (C-terminus) περιοχή και μια αμινοτελική (N-terminus) περιοχή που συμβάλλουν ενεργά στον σχηματισμό του συμπλόκου GHRH-GHRHR, η κάθε μια με διαφορετικό τρόπο. Συγκεκριμένα η πρόσδεση του πεπτιδίου γίνεται με βάση ένα μοντέλο δύο σταδίων (εικόνα 1.5). Σύμφωνα με αυτό το μοντέλο, κατά το πρώτο στάδιο, η εξωκυττάρια περιοχή του υποδοχέα αλληλεπιδρά με την καρβοξυτελική περιοχή του πεπτιδίου. Σύμφωνα με διάφορες μελέτες που έχουν γίνει αποδεικνύεται ότι η καρβοξυτελική περιοχή του πεπτιδίου είναι υπεύθυνη για την σύνδεση του πεπτιδίου στον υποδοχέα. Ως αποτέλεσμα της παραπάνω αλληλεπίδρασης, η αμινοτελική περιοχή του πεπτιδίου τοποθετείται με τέτοιο τρόπο ώστε να ευνοείται η αλληλεπίδραση της με την διαμεμβρανική περιοχή και συγκεκριμένα με τον J-Domain του υποδοχέα που αποτελεί

και το δεύτερο στάδιο του μοντέλου. Η αμινοτελική περιοχή του πεπτιδίου είναι υπεύθυνη για την ενεργοποίηση του υποδοχέα. Η επιτυχία της δημιουργίας του συμπλόκου οφείλεται επιπρόσθετα και στην διαμόρφωση α -έλικας που παίρνει το πεπτίδιο στην πλασματική μεμβράνη και πριν την πρόσδεση του στον υποδοχέα [26].



Εικόνα 1.5: Αναπαράσταση cartoon της διαδικασίας δύο σταδίων πρόσδεσης του GHRH στον GHRHR. Με γκρι χρώμα φαίνεται το διαμεμβρανικό κομμάτι (J-Domain) του GHRHR ενώ για την καλύτερη απεικόνιση παραλείπονται οι ECLs. Με κίτρινο χρώμα φαίνεται το ECD και με βέλη αναπαρίστανται τα β -φύλλα. Με κόκκινο και μπλε χρώμα παρουσιάζεται το καρβοξυτελικό και το αμινοτελικό μέρος του πεπτιδίου αντίστοιχα [26].

Φυσικό επόμενο της πρόσδεσης του πεπτιδίου, είναι οι διαμορφωτικές αλλαγές που υφίσταται ο υποδοχέας κατά την ενεργοποίησή του. Μέσω της σύγκριση των δύο διαφορετικών καταστάσεων, της ανενεργής και της ενεργής κατάστασης του GHRHR παρατηρείται πως η TM6 στην ενεργή κατάσταση του υποδοχέα και συγκεκριμένα το ενδοκυτταρικό μισό της, κάνει μια περιστροφή προς τα έξω με σκοπό να φιλοξενήσει την G πρωτεΐνη. Επιπλέον της TM6, τόσο η TM5 όσο και η ICL3 υφίστανται μια αλλαγή στην διαμόρφωσή τους, γεγονός που δημιουργεί τις κατάλληλες συνθήκες για την σύνδεση της G πρωτεΐνης με τον GHRHR (εικόνα 1.6). Η πρωτεΐνη G και συγκεκριμένα η G_α υπομονάδα του ετεροτριμερούς, αλληλεπιδρά με τα μέρη TM2, TM3, TM6 και TM7 ενώ η G_β υπομονάδα του ετεροτριμερούς, αλληλεπιδρά με τα μέρη ICL1 και H8 μέσω ενός δικτύου ηλεκτροστατικών αλληλεπιδράσεων στο οποίο συμμετέχουν τα αμινοξέα $GLU386^{8.53b}$ και $ARG389^{8.56b}$ του τομέα H8, η $ARG156^{12.49b}$ της ICL1 και το $ASP312$ της υπομονάδας G_β [21],[27].



Εικόνα 1.6: Αναπαράσταση με κυλινδρικές έλικες των δύο διαφορετικών καταστάσεων του GHRH-R. Με κόκκινο χρώμα παρουσιάζεται η ανενεργή κατάσταση ενώ με πράσινο χρώμα η ενεργή κατάσταση του υποδοχέα. Στην ενεργή κατάσταση διακρίνονται οι περιοχές TM5, TM6 και ICL3 να έχουν αλλάξει διαμόρφωση ώστε να επιτραπεί η σύνδεση της G πρωτεΐνης.

1.7 Χημειοπληροφορική

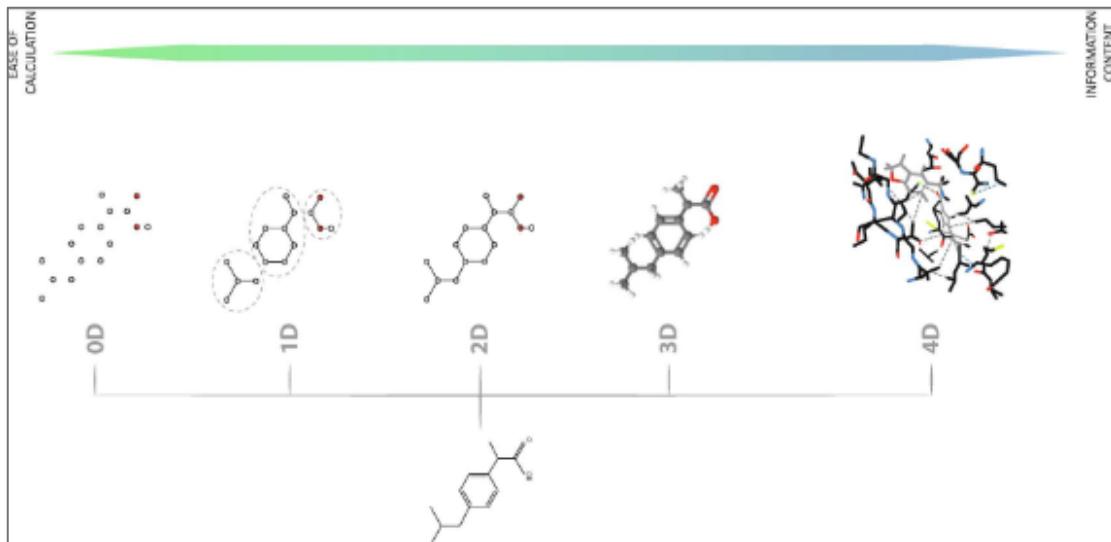
Ο όρος της χημειοπληροφορικής εισήχθη στην επιστημονική κοινότητα εδώ και πολλά χρόνια με πηγές να αναφέρουν ότι τα πρώτα βήματα έγιναν την εποχή 1960-1970 και αφορούσαν την πρόβλεψη δομών πρωτεϊνών, την ανάλυση των αλληλεπιδράσεων χημικών ενώσεων, την εφαρμογή Fourier για την επεξεργασία εικόνων που προέρχονται από κρυσταλλογραφία ακτινών X και την ανάλυση δεδομένων που προέρχονται από την μέθοδο της φασματοσκοπίας. Ένας από τους πιο γνωστούς ορισμούς που χρησιμοποιείται για τον κλάδο της χημειοπληροφορικής είναι η εφαρμογή της πληροφορικής στην επίλυση χημικών προβλημάτων. Εκτός από την πληροφορική και την χημεία συνδυάζει και τομείς όπως η βιολογία, η βιοχημεία, η φυσική, τα μαθηματικά και η στατιστική. Ο κλάδος αυτός αναπτύχθηκε με το πέρασμα των χρόνων λόγω της εξέλιξης της τεχνολογίας καθώς και της εκθετικά αυξανόμενης ποσότητας χημικής πληροφορίας η οποία θα μπορούσε να ελεγχθεί και να επεξεργαστεί πιο αποδοτικά με την χρήση υπολογιστών. Η χημειοπληροφορική εστιάζει στην αποθήκευση δεδομένων, στην ανάκτηση πληροφορίας από βάσεις δεδομένων που περιέχουν χημική πληροφορία, στον εντοπισμό αλλά και την πρόβλεψη φυσικό-χημικών ιδιοτήτων χημικών ενώσεων, στην αναζήτηση ομοιότητας ενώσεων με βάση αλγόριθμους, στην εφαρμογή προσομοιώσεων όπως η μοριακή πρόσδεση για την μελέτη αλληλεπιδράσεων καθώς και σε πολλές ακόμα εφαρμογές. Πλέον η χημειοπληροφορική χρησιμοποιείται σε διάφορα πεδία με τα πιο δημοφιλή να είναι η επιστήμη υλικών, η επιστήμη τροφίμων και η ανακάλυψη φαρμάκων. Αναφορικά με την ανακάλυψη φαρμάκων, η διαδικασία που απαιτείται για να βγει ένα φάρμακο στην αγορά είναι ο εντοπισμός της ασθένειας, του στόχου που σχετίζεται με αυτήν και της χημικής ένωσης που επιδρά στον στόχο, η διενέργεια προ-κλινικών και κλινικών δοκιμών και τέλος η έγκριση του φαρμάκου από τον FDA (Food and Drug Administration). Η εύρεση και βελτίωση της χημικής ένωσης αποτελεί συνήθως το πιο χρονοβόρο και δαπανηρό κομμάτι της διαδικασίας και επιλύεται σε μεγάλο βαθμό με την χρήση μεθόδων χημειοπληροφορικής και υπολογιστικής χημείας. Για τον εντοπισμό της ένωσης εφαρμόζεται σάρωση μεγάλων βιβλιοθηκών χημικών ενώσεων

για τον συγκεκριμένο υποδοχέα-στόχο, αξιολογείται η πρόσδεση και εφαρμόζονται αλγόριθμοι υπολογισμού ομοιότητας των ενώσεων ενώ η βελτιστοποίηση της ένωσης στοχεύει στην επιλογή των χαρακτηριστικών που θα επιτρέψουν την καλύτερη επιλεκτικότητα και πρόσδεση της ένωσης στον πρωτεϊνικό στόχο. Οι πιο γνωστές μέθοδοι που εφαρμόζονται στο κομμάτι του εντοπισμού και της βελτιστοποίησης της ένωσης είναι η Εικονική Σάρωση, η Εξόρυξη Δεδομένων, η ποσοτική σχέση δομής-δραστικότητας (Quality Structure – Activity Relationship, QSAR) και η ADMET (Adsorption Distribution Metabolism Elimination and Toxicity). Στα επόμενα κεφάλαια θα μελετηθούν στοιχεία που έχουν εφαρμογή στην χημειοπληροφορική όπως είναι οι τρόπος απεικόνισης των χημικών δομών, οι μοριακοί περιγραφείς, τα μοριακά αποτυπώματα και η μοριακή ομοιότητα [28],[29].

1.7.1 Μοριακές Αναπαράστασεις

Η ραγδαία αύξηση της πληροφορίας σχετικά με τις χημικές δομές μεταξύ άλλων οδήγησε και στην προσπάθεια αναπαράστασής τους με τέτοιο τρόπο ώστε να είναι περισσότερο αποδοτική η διαχείριση τους. Για την αποτελεσματικότερη διαχείριση των χημικών δεδομένων όπως είναι φυσικό απαιτείται η υπολογιστική επεξεργασία τους, μόνο που σε αυτήν την περίπτωση εκτός από κείμενο και αριθμούς εισάγονται και τα μόρια που αποτελούνται από άτομα, δεσμούς, έχουν τρισδιάστατη υπόσταση και φυσικοχημικές ιδιότητες που πρέπει να χαρακτηριστούν με ακρίβεια [28].

Η πιο απλή, μηδενικής διάστασης μοριακή αναπαράσταση είναι ο μοριακός τύπος της ένωσης και παρέχει πληροφορία σχετικά με τον αριθμό των διάφορων τύπων ατόμων από τα οποία αποτελείται ένα μόριο όμως δεν παρέχει καμία πληροφορία σχετικά με την δομή του μορίου και την συνδεσιμότητα των ατόμων που το απαρτίζουν. Η μονοδιάστατη μοριακή αναπαράσταση δεν προσδιορίζει την ολοκληρωμένη δομή του μορίου αλλά παρέχει μια λίστα με δομικά θραύσματα, λειτουργικές ομάδες και σημαντικούς υποκαταστάτες του μορίου. Η διςδιάστατη ή αλλιώς τοπολογική μοριακή αναπαράσταση παρέχει πληροφορίες σχετικά με την συνδεσιμότητα των ατόμων. Ανεβαίνοντας διάσταση αυξάνεται και η πολυπλοκότητα της αναπαράστασης όμως αυξάνεται και η παρεχόμενη πληροφορία που λαμβάνει ο χρήστης. Η τρισδιάστατη ή αλλιώς γεωμετρική μοριακή αναπαράσταση, επιπρόσθετα των πληροφοριών που παρέχονται στην διςδιάστατη αναπαράσταση, απεικονίζει και την διαμόρφωση του μορίου στον χώρο. Τέλος η μοριακή αναπαράσταση στις τέσσερις διαστάσεις ή αλλιώς στερεοηλεκτρονική αναπαράσταση απεικονίζει το μόριο σε σχέση με τις αλληλεπιδράσεις που κάνει με το γύρω περιβάλλον του [30]. Οι παραπάνω αναπαράστασεις φαίνονται και σχηματικά στην εικόνα 1.7.



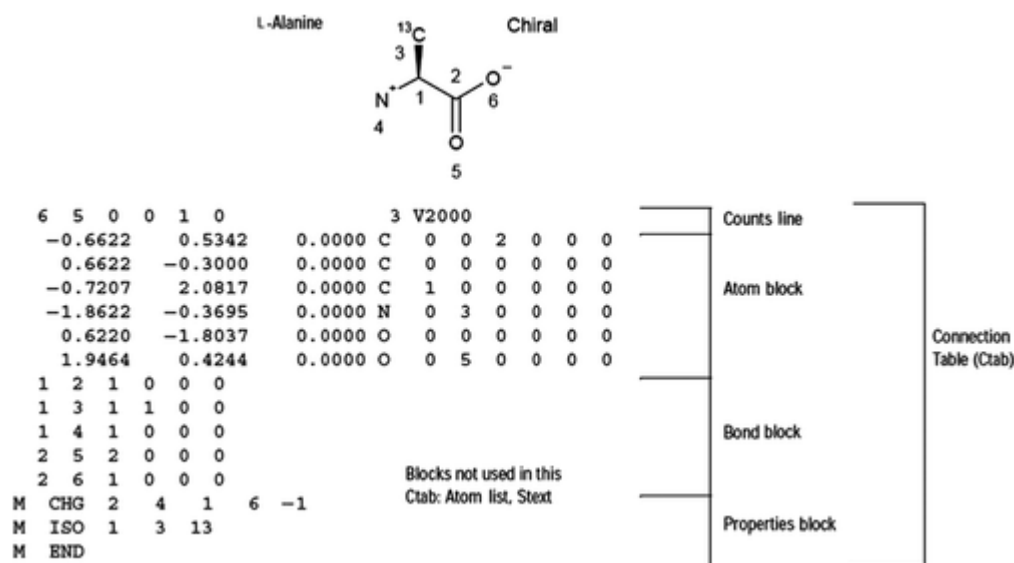
Εικόνα 1.7: Μοριακές αναπαραστάσεις (0D έως 4D) της ιβουπροφαίνης. Με την αύξηση της διάστασης αυξάνεται και η πολυπλοκότητα της απεικόνισης όμως εμπλουτίζεται η παρεχόμενη πληροφορία [31].

Ένας από τους πιο γνωστούς και συχνά χρησιμοποιούμενους τρόπους για την αναπαράσταση μοριακών δομών είναι αυτός μέσω γράφου, με τους κόμβους να αναπαρίστανται ως άτομα και τις ακμές ως δεσμοί που συνδέουν τα άτομα μεταξύ τους. Ένας γράφος (G) ορίζεται ως εξής (σχέση 6):

$$G=(V,E) \quad (6)$$

όπου V είναι το σύνολο των κόμβων και E το σύνολο των ακμών του γράφου. Στην πλειοψηφία τους, οι ακμές των μοριακών γράφων δεν έχουν κατεύθυνση δηλαδή μια ακμή συμβολίζει την ύπαρξη σχέσης μεταξύ δύο κόμβων χωρίς να τους διακρίνει σε αρχικό και τελικό κόμβο. Οι βασικές πληροφορίες που εξάγονται από έναν γράφο αφορούν την συνδεσιμότητα των ατόμων, τον τύπο των ατόμων που απαρτίζουν το μόριο και τον τύπο των δεσμών που συνδέουν τα άτομα μεταξύ τους. Η αναπαράσταση του γράφου από μια αφηρημένη μαθηματική έννοια σε ένα αντικείμενο που μπορεί να επεξεργαστεί εύκολα από τον υπολογιστή, βασίζεται συνήθως στην χρήση πινάκων. Ειδικότερα ένας πίνακας γειτνίασης (A) χρησιμοποιείται για να προσδιορίσει την συνδεσιμότητα των ατόμων. Εάν το στοιχείο του πίνακα $a_{i,j}$ ισούται με “ένα” ($a_{i,j}=1$) υποδηλώνει πως τα δύο άτομα v_i και v_j συνδέονται άμεσα με δεσμό ενώ αντίθετα αν $a_{i,j}$ ισούται με “μηδέν” ($a_{i,j}=0$) υποδηλώνει πως δεν υπάρχει σύνδεση μεταξύ των δύο κόμβων. Για τον χαρακτηρισμό του τύπου του ατόμου εισάγεται ένας νέος πίνακας (έστω X) ο οποίος περιέχει πληροφορίες σχετικά με ιδιότητες που αφορούν το άτομο όπως είναι ο τύπος του ατόμου (π.χ. C, N, O, S) και αν το άτομο είναι φορτισμένο (-1, 0, +1). Για τον χαρακτηρισμό του δεσμού και εδώ εισάγεται ένας νέος πίνακας (έστω E) ο οποίος παρέχει πληροφορία σχετικά με τον τύπο του δεσμού (απλός, διπλός, τριπλός, αρωματικός). Η παραπάνω πληροφορία είναι αποθηκευμένη σε αρχεία. Ένας από τους πιο γνωστούς τύπους αρχείων που χρησιμοποιείται ευρέως είναι τα SDF αρχεία τα οποία έχουν συγκεκριμένη δομή που βασίζεται στην Molfile μορφή. Ειδικότερα ο κορμός ενός τέτοιου αρχείου είναι ο πίνακας σύνδεσης (connection table) που αποτελείται από 6 μέρη τα οποία είναι η γραμμή μέτρησης (counts line), τα άτομα (atom block), οι δεσμοί (bond block), η λίστα με τα άτομα (atom list), το κείμενο (Stext) και οι ιδιότητες (properties block) όπως παρουσιάζεται στην εικόνα 1.8. Η γραμμή μέτρησης είναι η πρώτη γραμμή του αρχείου και παρέχει πληροφορία για τον αριθμό

των ατόμων, των δεσμών, την ύπαρξη ή όχι χειρομορφίας καθώς και την έκδοση του αρχείου που αναφέρεται ως V2000 (παλαιότερη έκδοση) ή V3000 (νέα έκδοση). Το μπλοκ των ατόμων πληροφορεί για τον τύπο των ατόμων, την στερεοχημεία, τα φορτία τους και πολλά ακόμα ενώ το μπλοκ των δεσμών για τον τύπο των δεσμών και την συνδεσιμότητα των ατόμων. Υπάρχουν διαθέσιμα πολλά λογισμικά που έχουν αναπτυχθεί τα τελευταία χρόνια και αποτελούν χρήσιμα “εργαλεία” στον κλάδο της χημειοπληροφορικής για την αναπαράσταση και επεξεργασία 2D και 3D γράφων. Μερικά από τα πιο γνωστά λογισμικά αναπαράστασης δομών είναι το PyMOL, το Jmol, το Chimera, το Avogadro, το MolView και το VMD [32].

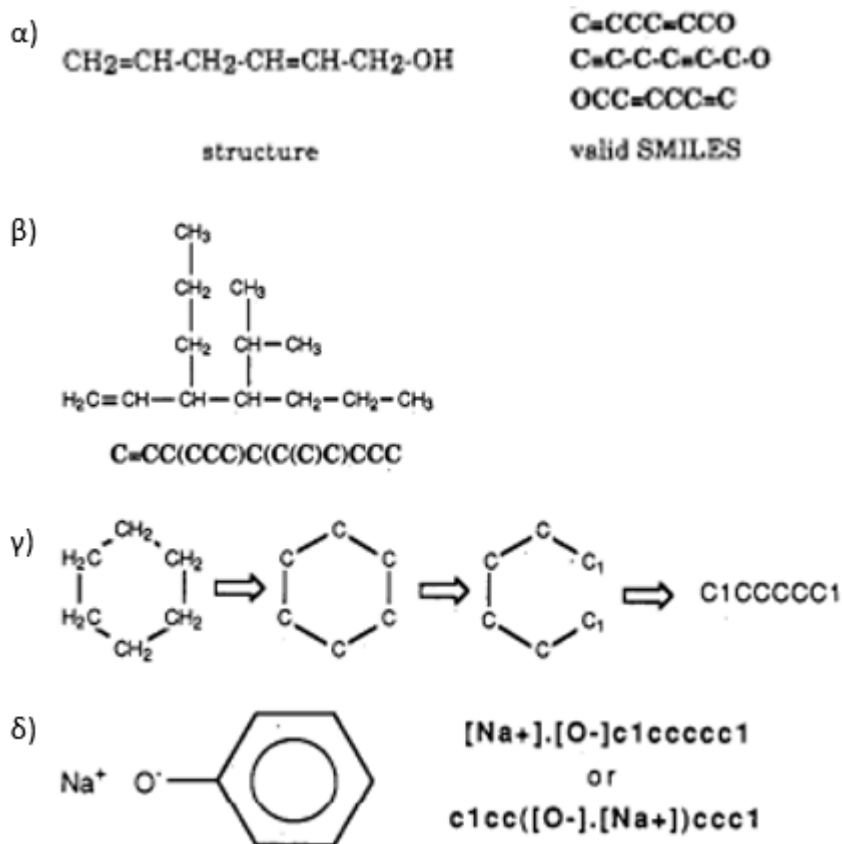


Εικόνα 1.8: Οργάνωση ενός Molfile αρχείου για την L-Αλανίνη. Φαίνονται τα μέρη που απαρτίζουν έναν πίνακα σύνδεσης (γραμμή μέτρησης, μπλοκ ατόμων, μπλοκ δεσμών, ιδιότητες) ενώ απουσιάζουν οι δομές για την λίστα ατόμων και την πληροφορία κειμένου. Το ‘M END’ σηματοδοτεί το τέλος του αρχείου [33].

Παρά τα πλεονεκτήματα που έχουν οι παραπάνω μοριακές αναπαραστάσεις, καταλαμβάνουν αρκετό χώρο στον δίσκο. Το γεγονός αυτό οδήγησε στην ανάπτυξη γραμμικών σημειογραφιών (linear notations) που αναπαριστούν τα μόρια με την χρήση συμβολοσειρών. Έτσι καταλαμβάνουν πολύ λιγότερο χώρο στον δίσκο ενώ ταυτόχρονα καθιστούν την διαχείριση τους από τον υπολογιστή ευκολότερη [32]. Κάποιες από τις πιο γνωστές αναπαραστάσεις αυτής της κατηγορίας είναι οι ακόλουθες: SMILES (Simplified Molecular Input Line Entry System), SMARTS (SMILES Arbitrary Target Specification), InChI (International Chemical Identifier) και InChIKeys.

Τα SMILES αναπτύχθηκαν το 1988 από τον David Weininger του κολεγίου Pomona στην Καλιφόρνια και αργότερα υλοποιήθηκαν από το σύστημα χημικών πληροφοριών Daylight (Daylight Chemical Information Systems). Εκτοτε αποτελούν την πιο γνωστή και ευρέως χρησιμοποιούμενη γραμμική σημειογραφία [32]. Για να αναπαραστήσουν τα οργανικά μόρια χρησιμοποιούν συμβολοσειρές που αποτελούνται από χαρακτήρες του κώδικα ASCII [34]. Ειδικότερα τα άτομα συμβολίζονται με τα ατομικά τους σύμβολα (όπως ακριβώς και στον περιοδικό πίνακα) ενώ ο δεσμός που συνδέει τα

άτομα μεταξύ τους, ανάλογα με το αν είναι απλός, διπλός, τριπλός, τετραπλός ή αρωματικός αναπαρίσταται ως -, =, #, \$, και : αντίστοιχα (εικόνα 1.9α). Σε περίπτωση που κάποιο άτομο έχει φορτίο (+ ή -) τότε το άτομο μαζί με το αντίστοιχο φορτίο εισάγεται σε αγκύλες. Είναι σημαντικό να αναφερθεί πως ο απλός και ο αρωματικός δεσμός στην πλειοψηφία των περιπτώσεων παραλείπονται ενώ είναι υποχρεωτικός ο μη συμβολισμός των υδρογόνων του μορίου. Η ύπαρξη διακλαδώσεων συμβολίζεται μέσω παρενθέσεων (εικόνα 1.9β) ενώ για την γραμμική αναπαράσταση των κυκλικών δομών “σπάει” ένας δεσμός και σημειώνεται ο αρχικός και ο τελικός άνθρακας με τον αριθμό ‘1’ αμέσως μετά το αντίστοιχο ατομικό σύμβολο (εικόνα 1.9γ). Σε μη συνδεδεμένες δομές όπως είναι η αναπαράσταση ενός μορίου και ενός ιόντος, οι δύο δομές διαχωρίζονται με τελεία (εικόνα 1.9δ). Κλείνοντας με τους βασικούς κανόνες για την κωδικοποίηση με SMILES, η αναπαράσταση του αρωματικού δακτυλίου ενός μορίου επιτυγχάνεται με την χρήση πεζών συμβόλων για κάθε άτομο του δακτυλίου (εικόνα 1.9δ) [34],[35]. Τα SMILES αν και είναι πολύ αποτελεσματικά στο να αναπαριστούν δομές ως γραμμικές σημειογραφίες έχουν έναν βασικό περιορισμό: δεν είναι μοναδικά, γεγονός που σημαίνει πως διαφορετικά SMILES μπορούν να κωδικοποιούν το ίδιο μόριο. Σε περιπτώσεις όπου χρειάζεται να εξασφαλιστεί η μοναδικότητα μιας ένωσης και να αφαιρεθούν τυχόν διπλότυπα χρησιμοποιούνται τα κανονικά SMILES στα οποία η επιλογή του αρχικού κόμβου δεν είναι τυχαία και καθορίζεται με βάση το εκάστοτε πρόγραμμα που χρησιμοποιείται. Αυτή την αδυναμία των SMILES επιλύει με αποδοτικό τρόπο η κωδικοποίηση InChI που θα αναλυθεί στην συνέχεια [32],[36].



Εικόνα 1.9: Κωδικοποίηση οργανικών μορίων με SMILES. Παράδειγμα γραμμικής αναπαράστασης (α) ατόμων και δεσμών, (β) διακλαδώσεων, (γ) κυκλικών ενώσεων (δ) ασύνδετων δομών και αρωματικών δακτυλίων [35].

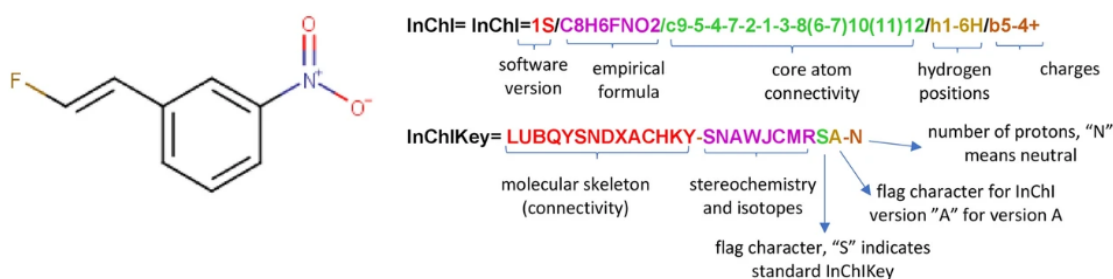
Η κωδικοποίηση που χρησιμοποιούν τα SMARTS αποτελεί προέκταση των SMILES. Συγκεκριμένα δίνουν την δυνατότητα χρήσης λογικών τελεστών και ειδικών συμβόλων που περιγράφουν τις ιδιότητες των ατόμων και των δεσμών, επιπρόσθετα όσων αναλύθηκαν παραπάνω για τα SMILES, οδηγώντας στην αύξηση της παρεχόμενης πληροφορίας για την δομή ενός μορίου. Τα SMARTS είναι ιδανικά για την αναζήτηση υποομάδας σε μια βάση δεδομένων, την αξιολόγηση της εκλεκτικότητας μορίων και τον προσδιορισμό ενεργών και ανενεργών ενώσεων [36].

Η κωδικοποίηση InChI αποτελεί μια γραμμική σημειογραφία που αναπτύχθηκε από την IUPAC (International Union of Pure and Applied Chemistry) το 2005, με την συνεισφορά των NIST (National Institute of Standards and Technology) και InChI Trust [34],[36]. Η κωδικοποίηση InChI στοχεύει στην εξαγωγή μοναδικής συμβολοσειράς για κάθε ένωση διευκολύνοντας με αυτόν τον τρόπο την σύνδεση διαφορετικών δεδομένων [33]. Αποτελεί ένα δωρεάν διαθέσιμο, ανοιχτού κώδικα αναγνωριστικό που βρίσκει εφαρμογή στην κωδικοποίηση των περισσότερων οργανικών και ανόργανων χημικών ενώσεων [34]. Στα θετικά της κωδικοποίησης InChI συγκαταλέγεται επίσης η επίλυση ασαφειών που δημιουργούνται από την εφαρμογή των SMILES και κυρίως αφορούν τα στερεογονικά κέντρα (άτομα συνδεδεμένα με τέσσερα διαφορετικά άτομα) και τις περιπτώσεις ταυτομέρειας. Από την άλλη τα επιπρόσθετα στοιχεία καθιστούν την μέθοδο πολυπλοκότερη συγκριτικά με προηγούμενα συστήματα γραμμικών σημειογραφιών και δυσκολότερα ερμηνεύσιμη από τον άνθρωπο [36]. Η δομή με την οποία αναπαρίστανται είναι ιεραρχική και χωρίζεται σε 6 διαφορετικά επίπεδα πληροφορίας που διαχωρίζονται μεταξύ τους με το σύμβολο “/” και περιγράφουν την χημική δομή της ένωσης. Η αρχή κάθε επιπέδου συμβολίζεται με συγκεκριμένο πρόθεμα ενώ κάποια από τα επίπεδα έχουν και δικά τους υπό-επίπεδα. Συγκεκριμένα αποτελείται από τα εξής [34]:

1. Κύριο επίπεδο (Main Layer): Παρέχει πληροφορία για τα άτομα ενός μορίου και την συνδεσιμότητα τους. Τα βασικά επίπεδα από τα οποία αποτελείται είναι ο μοριακός τύπος της ένωσης, η συνδεσιμότητα των ατόμων που συμβολίζεται με το πρόθεμα “/c” και τα υδρογόνα που συμβολίζονται με το πρόθεμα “/h”.
2. Επίπεδο φορτίου (Charge Layer): Σε αυτό το επίπεδο αναλύεται το φορτίο του ατόμου. Η περιγραφή γίνεται μέσω του φορτίου (πρόθεμα : “/q”) και της πρωτονίωσης και αποπρωτονίωσης (πρόθεμα “/p”).
3. Επίπεδο Στεροχημικής Διαμόρφωσης (Stereochemical Layer): Περιγράφει την στεροχημική διαμόρφωση της ένωσης και αποτελείται από τα επίπεδα διπλού δεσμού και τετραέδρου που συμβολίζονται με το πρόθεμα “/b” και “/t” αντίστοιχα.
4. Επίπεδο Ισοτοπικής Σύνθεσης (Isotopic Layer): Συμβολίζεται με το πρόθεμα “/i”
5. Επίπεδο Σταθερών Υδρογόνων (Fixed-H Layer): Συμβολίζεται με το πρόθεμα “/f”

6. Επανασυνδεδεμένο Επίπεδο (Reconnected Layer): Συμβολίζεται με το πρόθεμα “/r”

Η πολυπλοκότητα της κωδικοποίησης InChI δυσκολεύει την αναζήτηση χημικών δομών στο διαδίκτυο. Για τον σκοπό αυτό αναπτύχθηκε το InChIKey που αποτελεί μια συμπυκνωμένη αναπαράσταση του InChI. Το InChIKey διασφαλίζει την ακριβή και μοναδική περιγραφή των χημικών δομών και κατοχυρώνει την συμβατότητα των κωδικοποιήσεων μεταξύ διαφορετικών μηχανών αναζήτησης. Αποτελείται αυστηρά από 27 χαρακτήρες από τους οποίους οι πρώτοι 14 περιγράφουν την δομή του μορίου, την συνδεσιμότητα των ατόμων και τα υδρογόνα (ότι περιγράφεται από τα κύριο επίπεδο της κωδικοποίησης InChI). Έπειτα από μια παύλα ακολουθούν 10 ακόμα χαρακτήρες από τους οποίους οι πρώτοι 8 κωδικοποιούν πληροφορία που αφορά την στερεοχημεία και τα ισότοπα της ένωσης ενώ οι 2 επόμενοι χαρακτήρες προσδιορίζουν αν πρόκειται για ένα τυπικό InChIKey και δηλώνουν την έκδοση του. Η συμβολοσειρά του InChIKey ολοκληρώνεται με μια παύλα και έναν χαρακτήρα που δηλώνει την κατάσταση πρωτονίωσης ή αποπρωτονίωσης της ένωσης [33],[36]. Στην εικόνα 1.10 φαίνεται ένα παράδειγμα χρήσης των δύο κωδικοποιήσεων InChI και InChIKey για την περιγραφή μιας χημικής ένωσης.



Εικόνα 1.10: Κωδικοποίηση με InChI και InChIKey της ένωσης 1-[(E)-2-φθοροβενζύλιο]-3-νιτροβενζόλιο [36].

1.7.2 Μοριακοί Περιγραφείς

Οι μοριακοί περιγραφείς κατέχουν ένα βασικό ρόλο στο κομμάτι της έρευνας καθώς έχουν την ικανότητα να εξάγουν χρήσιμη πληροφορία από τα μόρια και να την κωδικοποιούν σε μορφή φιλική προς τους υπολογιστές. Διακρίνονται σε δύο κατηγορίες, σε αυτούς που προέκυψαν από πειραματικές διαδικασίες όπως είναι για παράδειγμα η τιμή logP, η διπολική ροπή, η διάθλαση του μορίου καθώς και άλλες φυσικοχημικές ιδιότητες και στους θεωρητικούς μοριακούς περιγραφείς που ταξινομούνται σε πέντε κλάσεις, τους 0D, 1D, 2D, 3D και 4D περιγραφείς [31]. Χαρακτηριστικοί περιγραφείς μηδενικής διάστασης είναι ο αριθμός των ατόμων, των δεσμών και το μοριακό βάρος, πληροφορίες που μπορούν πολύ εύκολα να εξαχθούν από τον μοριακό τύπο της ένωσης. Οι περιγραφείς που ανήκουν σε αυτή την κατηγορία δεν απαιτούν γνώση για την δομή των ενώσεων και την συνδεσιμότητα των ατόμων. Οι μονοδιάστατοι περιγραφείς βασίζονται σε πληροφορία που προέρχεται από δομικά θραύσματα, λειτουργικές ομάδες και σημαντικούς υποκαταστάτες του μορίου και χρησιμοποιούνται συνήθως σε εφαρμογές εικονικής σάρωσης, σχεδιασμού βιβλιοθηκών χημικών ενώσεων, υπολογισμού ομοιότητας καθώς και σε μοντελοποίηση φυσικοχημικών και βιολογικών ιδιοτήτων μορίων. Η πληροφορία

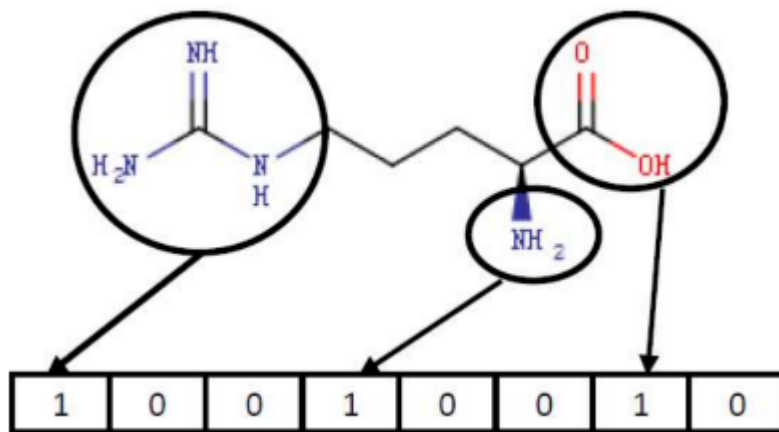
παρέχεται σε μορφή διανύσματος. Χαρακτηριστικό παράδειγμα μονοδιάστατου μοριακού περιγραφέα είναι τα μοριακά αποτυπώματα που κωδικοποιούν την παρουσία ή απουσία του δομικού θραύσματος μέσω μιας σειράς δυαδικών ψηφίων (bits) όπου το '0' αντιστοιχεί στην απουσία και το '1' στην παρουσία του αντίστοιχου θραύσματος. Τόσο οι 0D όσο και οι 1D μοριακοί περιγραφείς είναι εύκολα υπολογίσιμοι και παρέχουν μια βασική πληροφορία για το κάθε μόριο. Επιπλέον χαρακτηρίζονται από υψηλό εκφυλισμό (κυρίως οι 0D μοριακοί περιγραφείς) με πολλές ενώσεις, είτε ισομερή είτε όχι, να παρουσιάζουν ίδιες ή παρόμοιες τιμές με αποτέλεσμα πολλές φορές η χρήση τους να είναι συνδυαστική με πολυπλοκότερα μοντέλα που περιλαμβάνουν μοριακούς περιγραφείς μεγαλύτερης διάστασης όπως αυτοί που θα αναλυθούν στην συνέχεια [31],[37],[38]. Οι 2D μοριακοί περιγραφείς υπολογίζονται μέσω αλγορίθμων που χαρτογραφούν τοπολογικές αναπαραστάσεις (2D αναπαραστάσεις) και τις μετατρέπουν σε δεδομένα που επεξεργάζονται ευκολότερα σε έναν υπολογιστή. Για να επιτευχθεί αυτό εκμεταλλεύονται την θεωρία των γράφων κατά την οποία οι κόμβοι αντιπροσωπεύουν τα άτομα της ένωσης και οι ακμές τους δεσμούς της ένωσης. Η πληροφορία που παρέχεται από αυτούς τους μοριακούς περιγραφείς αφορά την συνδεσιμότητα των ατόμων το είδος και τις γωνίες των δεσμών, τις γωνίες στρέψης και τις τοπολογικές ή γεωμετρικές αποστάσεις και παρέχεται σε μορφή πινάκων. Η πληροφορία για την στερεοχημεία δηλαδή την χωρική διάταξη των ατόμων παρέχεται από τους 3D μοριακούς περιγραφείς. Ο πιο γνωστός αναπαριστά την ένωση ως μια φαρμακοφόρο δομή που παρέχει πληροφορία για τις υδρόφιλες και υδρόφοβες περιοχές του μορίου, για τους δότες και δέκτες δεσμών υδρογόνου καθώς και για πολλές ακόμα ιδιότητες του μορίου. Οι 4D μοριακοί περιγραφείς προέρχονται από πολλές διαφορετικές διαμορφώσεις του ίδιου μορίου και δίνουν πληροφορία σχετικά με την κατανομή των ηλεκτρονίων, με τις αλληλεπιδράσεις που κάνει το μόριο με τον περιβάλλοντα χώρο καθώς και με ενεργά κέντρα βιολογικών στόχων-υποδοχέων. Ο υπολογισμός των δύο τελευταίων μοριακών περιγραφέων είναι αρκετά πολύπλοκος όμως εξάγει χρήσιμη πληροφορία σχετικά με τις φυσικοχημικές και βιολογικές ιδιότητες του μορίου [31],[34],[39].

1.7.3 Μοριακά Αποτυπώματα

Τα μοριακά αποτυπώματα αρχικά κατασκευάστηκαν με σκοπό την αναζήτηση σε βιβλιοθήκες και βάσεις δεδομένων ενώ στην συνέχεια χρησιμοποιήθηκαν για αναζήτηση ομοιότητας και ομαδοποιήσεις χημικών ενώσεων [40]. Πλέον έχουν εφαρμογή σε διαδικασίες όπως η ανακάλυψη φαρμάκων και η εικονική σάρωση. Βασικός λόγος που η χρήση τους έχει εδραιωθεί παγκοσμίως είναι το γεγονός πως συνδυάζουν την ευκολία στην χρήση σε συνδυασμό με την ταχύτητα στον υπολογισμό τους. Υπάρχουν διάφορων ειδών μοριακά αποτυπώματα ανάλογα με τον τρόπο με τον οποίο η πληροφορία του μορίου μετατρέπεται σε δυαδική κατά βάση συμβολοσειρά, με το κάθε ένα να χαρακτηρίζει διαφορετικές πτυχές του ίδιου μορίου. Σε αυτό το κεφάλαιο θα αναλυθούν τα πιο δημοφιλή μοριακά αποτυπώματα. Οι πιο γνωστές και ευρέως χρησιμοποιούμενες κατηγορίες μοριακών αποτυπωμάτων είναι τα αποτυπώματα βασισμένα σε δομικά κλειδιά υποομάδων (substructure key-based fingerprints), τα τοπολογικά αποτυπώματα (topological fingerprints) και τα κυκλικά αποτυπώματα (circular fingerprints) [41],[42].

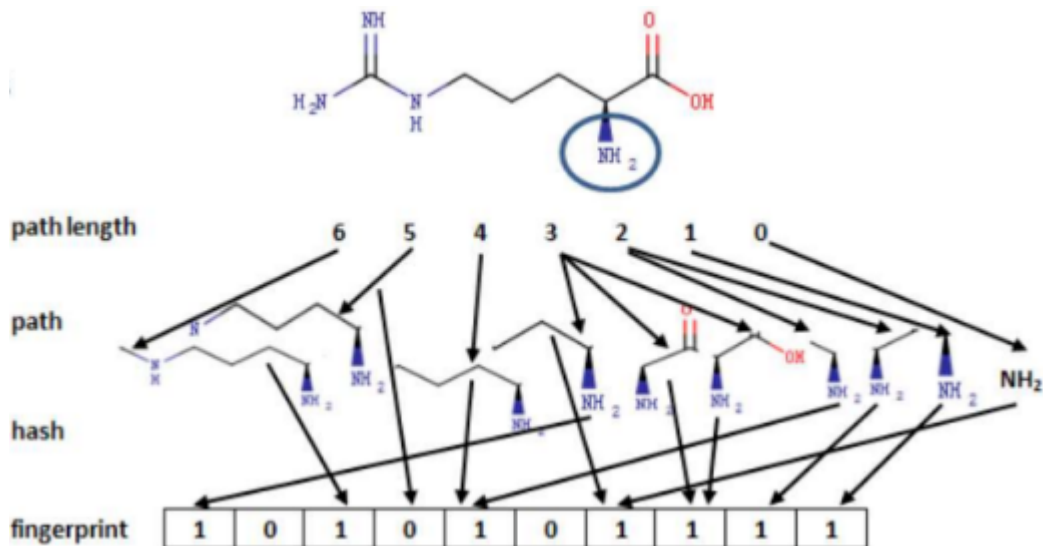
Τα αποτυπώματα που είναι βασισμένα σε δομικά κλειδιά υποομάδων μετατρέπουν την πληροφορία που προέρχεται από την αναπαράσταση του μορίου σε συμβολοσειρά με

βάση την ύπαρξη ή όχι συγκεκριμένων δομικών κλειδιών. Τα δομικά κλειδιά είναι αποθηκευμένα σε μια λίστα και η παρουσία τους στο μόριο συμβολίζεται με '1' ενώ η απουσία τους με '0'. Όπως είναι φυσικό, ο αριθμός των bits της συμβολοσειράς καθορίζεται από τον αριθμό των δομικών κλειδιών που υπάρχουν στην λίστα. Ένα από τα πιο χαρακτηριστικά παραδείγματα αυτής της κατηγορίας αποτυπώματων είναι τα MACCS (Molecular Access System) με δύο εκδοχές που διαφέρουν ως προς τον αριθμό των δομικών κλειδιών που περιέχουν. Η πρώτη αποτελείται από 960 δομικά κλειδιά (960 bits) ενώ η δεύτερη από 166 δομικά κλειδιά (166 bits) καθιστώντας την τελευταία την πιο συχνά χρησιμοποιούμενη τόσο λόγω των λιγότερων bit που καταλαμβάνει όσο και της δυνατότητάς που έχει να καλύψει τα πιο σημαντικά χαρακτηριστικά που απαιτούνται σε εφαρμογές εικονικής σάρωσης και ανακάλυψης φαρμάκων. Άλλα γνωστά μοριακά αποτυπώματα αυτής της κατηγορίας είναι τα PubChem αποτυπώματα με 881 δομικά κλειδιά και το BCI αποτυπώματα με μεταβαλλόμενο αριθμό δομικών κλειδιών (συνήθης τιμή 1052 bits) [41],[42]. Στην εικόνα 1.11 παρουσιάζεται ένα παράδειγμα αποτυπώματος βασισμένο σε δομικά κλειδιά υποομάδων στο οποίο από τα οκτώ δομικά κλειδιά που μελετήθηκαν, τα τρία ήταν παρόν στην συγκεκριμένη ένωση [43].



Εικόνα 1.11: Παράδειγμα εφαρμογής μοριακού αποτυπώματος βασισμένο σε δομικά κλειδιά υποομάδων. Από ένα σύνολο οκτώ δομικών κλειδιών, τα τρία που είναι κυκλωμένα και αντιστοιχισμένα με '1' βρέθηκαν στην συγκεκριμένη ένωση [43].

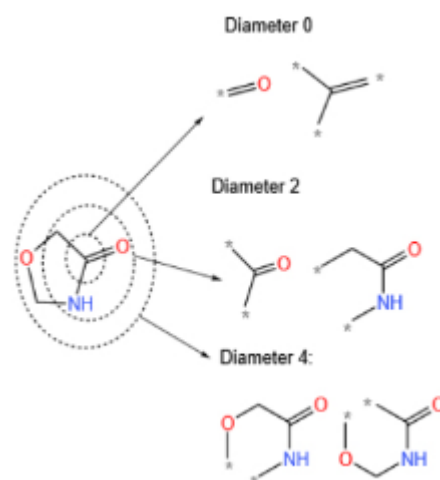
Τα τοπολογικά ή αλλιώς τα βασισμένα στην διαδρομή (path-based) μοριακά αποτυπώματα στηρίζονται στην ανάλυση των θραυσμάτων ενός μορίου. Για την κωδικοποίηση του αποτυπώματος ακολουθούν συνήθως ένα γραμμικό μονοπάτι έως έναν συγκεκριμένο αριθμό δεσμών που καθορίζεται από τον χρήστη και θέτουν την αντίστοιχη θέση του θραύσματος ως '1' εάν είναι παρόν στην ένωση και ως '0' αν είναι απών. Με την καταχώρηση κάθε διαδρομής δημιουργείται το τελικό μοριακό αποτύπωμα το οποίο χρησιμοποιείται σε εφαρμογές αναζήτησης θραυσμάτων και φιλτραρίσματος βιβλιοθηκών χημικών μορίων. Ένα από τα πιο γνωστά τοπολογικά μοριακά αποτυπώματα είναι το Daylight το οποίο αποτελείται από 2048 bits και έχει την ικανότητα να κωδικοποιήσει όλα τα πιθανά μονοπάτια οποιουδήποτε μορίου [41],[42],[43]. Στην εικόνα 1.12 παρουσιάζεται ένα παράδειγμα ενός Daylight μοριακού αποτυπώματος.



Εικόνα 1.12: Στην εικόνα φαίνεται ένα τοπολογικό αποτύπωμα τύπου Daylight που αποτελείται από 10 bits. Η ανάλυση ξεκινάει από την NH_2 που είναι κυκλωμένη και καταλήγει σε ένα μονοπάτι με 6 δεσμούς δημιουργώντας δέκα διαφορετικά θραύσματα τα οποία καταχωρούνται με αποτέλεσμα μια δυαδική συμβολοσειρά που αποτελεί το τελικό αποτύπωμα. Στις θέσεις 5, 7 και 8 έχουμε σύμπτωση θραυσμάτων στο ίδιο bit (bit collision) και οφείλεται πιθανώς στον μικρό αριθμό διαθέσιμων bits. Για να είναι ολοκληρωμένο το συγκεκριμένο τοπολογικό αποτύπωμα πρέπει να ακολουθηθεί η διαδικασία που έγινε για το NH_2 και για τα υπόλοιπα άτομα της ένωσης [43].

Όπως και τα τοπολογικά έτσι και τα κυκλικά μοριακά αποτυπώματα εξάγουν δομικά θραύσματα από τα οποία προκύπτει το τελικό μοριακό αποτύπωμα, όμως διαφέρουν στον τρόπο με τον οποίο κωδικοποιούν την πληροφορία. Τα κυκλικά μοριακά αποτυπώματα δεν χρησιμοποιούν γραμμικές διαδρομές αλλά μελετούν το περιβάλλον γύρω από ένα άτομο μέχρι μια καθορισμένη ακτίνα (εικόνα 1.13). Όπως και στα τοπολογικά αποτυπώματα έτσι και εδώ, η διαδικασία συνεχίζεται χρησιμοποιώντας διαδοχικά όλα τα άτομα της ένωσης ως σημείο εκκίνησης. Τα πιο κοινά μοριακά αποτυπώματα αυτής της κατηγορίας είναι τα ECFP (Extended-Connectivity Fingerprints) που παράγουν διαφορετικού μήκους δομικά θραύσματα ανάλογα με την ακτίνα του κύκλου που έχει καθοριστεί. Συνήθως τιμές για την διάμετρο του κύκλου είναι 4 και 6 Angstrom και τα μοριακά αποτυπώματα συμβολίζονται ως ECFP4 και ECFP6 αντίστοιχα. Πολλές φορές στον χαρακτηρισμό του αποτυπώματος αναφέρεται και ο αριθμός των bits από τα οποία αποτελείται ECFP4_{bits} (για παράδειγμα ECFP4₂₀₄₈) Παραλλαγή του ECFP αποτελεί το FCFP (Functional-Class Fingerprints) με την διαφορά τους να εντοπίζεται στο γεγονός πως στο FCFP λαμβάνεται υπόψη η λειτουργία κάθε ατόμου με αποτέλεσμα άτομα με ίδια ή παρόμοια λειτουργία να μην διακρίνονται ως διαφορετικά επιτρέποντας τα FCFP να χρησιμοποιηθούν ως φαρμακοφόρα αποτυπώματα. Η παρουσία ή απουσία των δομικών θραυσμάτων καταχωρείται ως '1' ή '0' αντίστοιχα με αποτέλεσμα την δημιουργία του τελικού μοριακού αποτυπώματος. Τα κυκλικά αποτυπώματα είναι κατάλληλα για αναζητήσεις ομοιότητας με βάση ολόκληρη την δομή της ένωσης, έχουν υψηλές αποδόσεις σε εφαρμογές εικονικής σάρωσης, σε μελέτες του μεταβολισμού και σε προβλέψεις των τιμών pKa. Ένας βασικός περιορισμός αυτής της κατηγορίας μοριακών αποτυπωμάτων εντοπίζεται στο γεγονός πως δεν είναι εφικτός ο προσδιορισμός της συνδεσιμότητας μεταξύ ατόμων που βρίσκονται σε διαφορετικά επίπεδα που έχουν καθοριστεί για συγκεκριμένες ακτίνες γύρω από το κεντρικό άτομο. Επιπλέον, συχνά δεν είναι

δυνατόν να εντοπισθούν όλοι οι δεσμοί που κάνει ένα άτομο αν και αυτός ο περιορισμός ξεπερνιέται μέσω των αλληλοεπικαλύψεων που εμφανίζονται κατά την διάρκεια της διαδικασίας [41],[42],[44].



Εικόνα 1.13: Στην εικόνα παρουσιάζεται ένα παράδειγμα κυκλικού αποτυπώματος με κεντρικό άτομο το άτομο άνθρακα και τρία επίπεδα με διαμέτρους 0, 2 και 4. Για διάμετρο 0 παράγονται δομικά θραύσματα με άτομα που είναι άμεσα συνδεδεμένα με το κεντρικό άτομο άνθρακα. Όσο μεγαλώνει η διάμετρος του κυκλικού αποτυπώματος εμφανίζονται μεγαλύτερα δομικά θραύσματα με έμμεσα συνδεδεμένα άτομα με το κεντρικό άτομο άνθρακα [45].

1.7.4 Μοριακή Ομοιότητα - Αναζήτηση Ομοιότητας Tanimoto

Η μοριακή ομοιότητα χρησιμοποιείται ευρέως στον τομέα της χημειοπληροφορικής για την σύγκριση και τον ποσοτικό προσδιορισμό της ομοιότητας δύο ή περισσότερων μορίων σε επίπεδο δομικό, λειτουργικό και φυσικό-χημικών ιδιοτήτων. Βρίσκει εφαρμογή στον σχεδιασμό και στην ανακάλυψη νέων φαρμάκων αλλά και στην βελτιστοποίηση ήδη υπαρχόντων. Για παράδειγμα η αναζήτηση ομοιότητας αποτελεί βασικό μέρος μιας διαδικασίας εικονικής σάρωσης και στοχεύει από μια χημειοθήκη χιλιάδων ενώσεων, τον εντοπισμό εκείνων που παρουσιάζουν υψηλό ποσοστό ομοιότητας με μια ένωση αναφοράς. Τρεις είναι οι βασικοί παράμετροι που επηρεάζουν τον υπολογισμό της μοριακής ομοιότητας. Ένας από αυτούς είναι ο τρόπος αναπαράστασης των μοριακών δομών, με συνηθέστερο, την χρήση διάφορων τύπων μοριακών περιγραφών όπως αυτοί που μελετήθηκαν σε προηγούμενο κεφάλαιο. Επιπλέον η προσθήκη βαρών είναι απαραίτητη στην περίπτωση που κάποια χαρακτηριστικά του μορίου έχουν προτεραιότητα έναντι άλλων. Τέλος η συνάρτηση ομοιότητας ή αλλιώς όπως συνηθίζεται να λέγεται ο συντελεστής ομοιότητας, χρησιμοποιείται για να υπολογίσει ποσοτικά την ομοιότητα μεταξύ μορίων [46]. Γνωστές συναρτήσεις ομοιότητας είναι ο συντελεστής συσχέτισης και ο συντελεστής απόστασης που εφαρμόζονται σε συνδυασμό με δυαδικές αναπαραστάσεις και η τιμή τους συνήθως κυμαίνεται μεταξύ “0” (καμία ταύτιση μεταξύ των συγκρινόμενων ενώσεων) και “1” (πανομοιότυπες ενώσεις) (Πίνακας 1.1) [47]. Ο πιο γνωστός συντελεστής απόστασης είναι η Ευκλείδεια απόσταση (Euclidean distance) ενώ ο πιο γνωστός συντελεστής συσχέτισης είναι ο συντελεστής Tanimoto (Tanimoto coefficient, T_c) ή αλλιώς συντελεστής Jaccard ο οποίος θα αναλυθεί παρακάτω [41].

Πίνακας 1.1: Παραδείγματα συντελεστών συσχετίσεων και αποστάσεων που χρησιμοποιούνται για την σύγκριση ομοιότητας μοριακών αποτυπωμάτων [41].

Measure	Expression	Range
Tanimoto/Jaccard coefficient	$\frac{c}{a+b-c}$	0 to 1
Euclidean distance	$\sqrt{a+b-2c}$	0 to N
City-block/Manhattan/Hamming distance	$a+b-2c$	0 to N
Dice coefficient	$\frac{2c}{a+b}$	0 to 1
Cosine similarity	$\frac{c}{\sqrt{ab}}$	0 to 1
Russell-RAO coefficient	$\frac{c}{m}$	0 to 1
Forbes coefficient	$\frac{cm}{ab}$	0 to 1
Soergel distance	$\frac{a+b-2c}{a+b-c}$	0 to 1

Ο συντελεστής Tanimoto αποτελεί την πιο γνωστή και συχνά χρησιμοποιούμενη μέθοδο για την σύγκριση μοριακών αποτυπωμάτων [41]. Παρόλο που δεν αποτελεί την καλύτερη διαδικασία για τον έλεγχο ομοιότητας μεταξύ μορίων, η ευκολία του στην εφαρμογή και η ταχύτητα του, τον έχουν ξεχωρίσει μεταξύ άλλων [46]. Ο συντελεστής Tanimoto υπολογίζεται διαιρώντας τον συνολικό αριθμό κοινών χαρακτηριστικών μεταξύ δύο ενώσεων με το σύνολο των χαρακτηριστικών που είναι παρόν στις δύο ενώσεις (σχέση 7). Η τιμή του κατωφλίου ομοιότητας πάνω από την οποία θεωρείται ότι δύο ενώσεις μοιράζονται όμοια δραστηριότητα, είναι $T_c=0.85$. Η τιμή αυτή δεν αποτελεί ένα γενικό και σταθερό κατώφλι αλλά εξαρτάται από την κάθε εφαρμογή και από περιορισμούς που θα αναλυθούν στην συνέχεια [41],[46],[48].

$$T_c(A, B) = \frac{c}{a + b - c} \quad (7)$$

Στην σχέση 7 παρουσιάζεται ο υπολογισμός του συντελεστή Tanimoto για δύο ενώσεις A και B. Με “a” συμβολίζεται ο αριθμός των χαρακτηριστικών που είναι παρόν στην ένωση A, με “b” ο αριθμός των χαρακτηριστικών που είναι παρόν στην ένωση B και με “c” ο αριθμός των κοινών χαρακτηριστικών στις ενώσεις A και B. Το “c” χρησιμοποιείται στον παρονομαστή για να αποφευχθεί η χρήση ενός χαρακτηριστικού δύο φορές [46]. Από τον παραπάνω τύπο αποδεικνύεται πως η τιμή του συντελεστή Tanimoto κυμαίνεται μεταξύ του “0” και του “1” ανεξάρτητα με το μήκος του μοριακού αποτυπώματος. Το γεγονός αυτό περιορίζει τον συγκεκριμένο συντελεστή ομοιότητας καθώς δεν είναι τόσο αντιπροσωπευτικός όσο μεγαλώνει το μήκος των μοριακών αποτυπωμάτων. Συμπεραίνεται λοιπόν πως το ποσοστό της ομοιότητας μεταξύ δύο ενώσεων εξαρτάται άμεσα από τον τύπο του μοριακού αποτυπώματος που χρησιμοποιείται καθιστώντας δύσκολη την εδραίωση ενός κατωφλίου που θα χρησιμοποιείται για να προσδιορίσει την ομοιότητα η ανομοιότητα δύο δομών [41],[48].

1.8 Μοριακή Πρόσδεση

1.8.1 Μοριακή Πρόσδεση - Μεθοδολογίες

Η μοριακή πρόσδεση αποτελεί μια από τις βασικές μεθόδους που χρησιμοποιείται κατά την διαδικασία του σχεδιασμού φαρμάκου με την υποστήριξη του υπολογιστή (Computer Aided Drug Design, CADD). Εφαρμόζεται για την μοντελοποίηση της αλληλεπίδρασης διάφορων συστημάτων όπως είναι το σύστημα πρωτεΐνης-πρωτεΐνης, πρωτεΐνης-προσδέτη και πρωτεΐνης-νουκλεϊκού οξέος. Στα πλαίσια της παρούσας διπλωματικής εργασίας, γίνεται αναφορά για το σύστημα πρωτεΐνης-προσδέτη. Ο όρος “προσδέτης” αναφέρεται σε μικρό φαρμακευτικό μόριο το οποίο ανάλογα με την δράση του επάγει ένα συγκεκριμένο βιολογικό αποτέλεσμα (αναστολή ή ενεργοποίηση) όταν συνδεθεί στον υποδοχέα. Η περιοχή στην οποία προσδέεται, ονομάζεται θήκη πρόσδεσης ή περιοχή πρόσδεσης ή ενεργό κέντρο του υποδοχέα. Η γνώση της περιοχής πρόσδεσης του υποδοχέα είναι ουσιαστική για την αποτελεσματικότητα της μεθόδου ενώ στις περιπτώσεις που δεν είναι γνωστή, είναι δυνατή η εύρεση της συγκρίνοντας την δομή του υποδοχέα-στόχου με άλλες πρωτεΐνες που έχουν παρόμοια λειτουργία ή που είναι συγκρυσταλλωμένες με κάποιον προσδέτη [49],[50].

Το μοντέλο που χρησιμοποιήθηκε αρχικά για να περιγράψει την μοριακή πρόσδεση είναι το μοντέλο κλειδιού-κλειδαριάς που προτάθηκε από τον Emil Fischer το 1890 και παρομοιάζει τον τρόπο με τον οποίο ένας προσδέτης ταιριάζει στο ενεργό κέντρο του υποδοχέα. Σύμφωνα με αυτό το μοντέλο, τόσο ο υποδοχέας όσο και ο προσδέτης αντιμετωπίζονται ως άκαμπτα σώματα αποτελούμενα από τρεις περιστροφικούς και τρεις μεταφορικούς βαθμούς ελευθερίας γεγονός που περιορίζει σε μεγάλο βαθμό τον χώρο αναζήτησης. Για να υπερβεί τον παραπάνω περιορισμό προτάθηκε το 1958 από τον Daniel Koshland το μοντέλο της επαγόμενης προσαρμογής. Με βάση αυτό το μοντέλο τόσο ο υποδοχέας όσο και ο προσδέτης αντιμετωπίζονται ως εύκαμπτα σώματα που προσαρμόζονται αμοιβαία μεταξύ τους αποσκοπώντας στην εύρεση του συμπλόκου με την ελάχιστη ενέργεια και ως επόμενο στον εντοπισμό της βέλτιστης πρόσδεσης. Η μέθοδος της επαγόμενης προσαρμογής προσφέρει μεγαλύτερη ακρίβεια στην προσομοίωση της αλληλεπίδρασης υποδοχέα-προσδέτη συγκριτικά με την άκαμπτη πρόσδεση όμως η ενσωμάτωση της ευελιξίας του υποδοχέα παρουσιάζει διάφορες δυσκολίες και απαιτεί πολλούς υπολογιστικούς πόρους γεγονός που έχει περιορίσει την χρήση της στην μοριακή πρόσδεση. Πολλοί μέθοδοι έχουν προταθεί για την ενσωμάτωση της ευελιξίας της πρωτεΐνης κατά την εφαρμογή της μοριακής πρόσδεσης όπως η μέθοδος του “soft docking”, η ενσωμάτωση της ευελιξίας της πλευρικής αλυσίδας με την χρήση βιβλιοθηκών ροταμερών και η μέθοδος πολλαπλών πρωτεϊνικών δομών. Παρά τις προσπάθειες που έχουν γίνει, η ενσωμάτωση της ευελιξίας της πρωτεΐνης και συγκεκριμένα της ραχοκοκαλιάς παρουσιάζει μέχρι και σήμερα μεγάλη πρόκληση κατά την υλοποίηση της μοριακής πρόσδεσης. Για τους παραπάνω λόγους επιλέγεται συχνά μια ισορροπία μεταξύ υπολογιστικού κόστους και ακρίβειας της μοντελοποίησης, ενσωματώνοντας μόνο την ευελιξία του προσδέτη και αντιμετωπίζοντας τον υποδοχέα ως άκαμπτο σώμα [50],[51].

1.8.2 Διαμοριακές Αλληλεπιδράσεις

Η μελέτη των διαμοριακών αλληλεπιδράσεων του φαρμακευτικού μορίου με τον υποδοχέα είναι ουσιαστική καθώς σκοπεύει στην θεωρητική μελέτη, στην βελτιωμένη φαρμακευτική δράση καθώς και στον σχεδιασμό νέων φαρμακευτικών προϊόντων. Η αλληλεπίδραση φαρμάκου-υποδοχέα μπορεί να οριστεί από την ελεύθερη ενέργεια Gibbs (ΔG) που ισούται με το άθροισμα των μεταβολών των ελεύθερων ενεργειών των ηλεκτροστατικών (Electrostatic Interactions, EI), επαγωγικών (Inductive Interactions, II), μη πολικών (Non Polar Interactions, NPI) και υδρόφοβων αλληλεπιδράσεων (Hydrophobic Interactions, HI) αφαιρώντας τις απώλειες ενέργειας (ΔG^*) που οφείλονται σε εντροπικούς ή ενθαλπικούς παράγοντες κατά την πρόσδεση. (σχέση 8)

$$\Delta G_{\text{total}} = \Delta G_{\text{EI}} + \Delta G_{\text{II}} + \Delta G_{\text{NPI}} + \Delta G_{\text{HI}} - \Delta G^* \quad (8)$$

Στην συνέχεια αναλύονται ορισμένες από τις αλληλεπιδράσεις που αναφέρθηκαν παραπάνω. Οι ηλεκτροστατικές αλληλεπιδράσεις δημιουργούνται λόγω της έλξης των πυρήνων και των ηλεκτρονίων μεταξύ των δύο μορίων. Η ελκτική αυτή δύναμη έχει ως αποτέλεσμα την δημιουργία τριών τύπων αλληλεπιδράσεων: Αλληλεπίδραση φορτίου-φορτίου, φορτίου-διπόλου και διπόλου-διπόλου. Οι δύο τελευταίες είναι αρκετά διαδεδομένες καθώς εμφανίζονται σε κάθε μόριο του οποίου τα άτομα έχουν διαφορετική τιμή ηλεκτραρνητικότητας όμως είναι ασθενής αλληλεπιδράσεις. Αντιθέτως, οι ηλεκτροστατικές αλληλεπιδράσεις φορτίου-φορτίου είναι πολύ ισχυρές καθώς τα κατιοντικά και ανιοντικά τμήματα του φαρμακευτικού μορίου σε φυσιολογικό pH αλληλεπιδρούν με τις αντίθετα φορτισμένες ομάδες της πρωτεΐνης. Επίσης οι δεσμοί υδρογόνου είναι πολύ σημαντικοί κατά την πρόσδεση καθώς σταθεροποιούν το σύμπλοκο φαρμακευτικού μορίου-υποδοχέα. Ο δεσμός υδρογόνου αποτελείται από ένα υδρογόνο που είναι συνδεδεμένο με ένα ισχυρό ηλεκτραρνητικό στοιχείο όπως το οξυγόνο, το άζωτο ή το φθόριο το οποίο δρα ως δότης δεσμού υδρογόνου. Το ηλεκτραρνητικό άτομο εκτρέπει το ηλεκτρονικό νέφος που υπάρχει γύρω από τον πυρήνα του υδρογόνου αφήνοντας το με μερικό θετικό φορτίο το οποίο προσελκύει ένα ασύζευκτο ζεύγος ηλεκτρονίων από ένα άλλο άτομο που λειτουργεί ως δέκτης δεσμού υδρογόνου, σχηματίζοντας τελικά τον δεσμό υδρογόνου. Στα βιολογικά συστήματα οι δεσμοί υδρογόνου που σχετίζονται με την διατήρηση της τριτοταγούς δομής των πρωτεϊνών και με την πρόσδεση ενώσεων στον υποδοχέα, συνήθως αφορούν άτομα οξυγόνου των καρβοξυλομάδων, των καρβονυλομάδων και των υδροξυλομάδων καθώς και άτομα αζώτου των αμινομάδων. Οι υδρόφοβες αλληλεπιδράσεις είναι οι δυνάμεις που διατηρούν την συνοχή μη πολικών μορίων σε διάλυμα ή πολικό διαλύτη. Κατά την σύνδεση δύο άπολων μορίων στο νερό, ελαττώνεται η επιφάνεια που έρχεται σε επαφή με το διαλύτη με αποτέλεσμα να παρέχεται ευνοϊκή εντροπία για τη σύνδεση τους. Για τον ολοκληρωμένο θεωρητικό χαρακτηρισμό όλων των αλληλεπιδράσεων απαιτείται η κβαντομηχανική προσέγγιση λαμβάνοντας υπόψη τον υπολογισμό της χρονικά εξαρτώμενης εξίσωσης Schrödinger για τον υποδοχέα, τον προσδέτη και τον διαλύτη [49].

1.8.3 Αλγόριθμοι Διαμορφωτικής Αναζήτησης

Για την υλοποίηση της μοριακής πρόσδεσης απαιτούνται δύο διακριτά αλλά στενά συνδεδεμένα βήματα. Το πρώτο αφορά την διαμορφωτική αναζήτηση του προσδέτη στο ενεργό κέντρο του υποδοχέα ενώ το δεύτερο σχετίζεται με την αξιολόγηση της πρόσδεσης μέσω μιας συνάρτησης βαθμολόγησης. Στο πρώτο στάδιο γίνεται εφαρμογή αλγορίθμων διαμορφωτικής αναζήτησης που τροποποιούν διάφορες δομικές ιδιότητες

(δίεδρη γωνία, μεταφορικοί και περιστροφικοί βαθμοί ελευθερίας) του φαρμακευτικού μορίου εφαρμόζοντας μεθόδους στοχαστικής ή συστηματικής διαμορφωτικής αναζήτησης [52].

Οι συστηματικές μέθοδοι εξερευνούν τον διαμορφωτικό χώρο κάνοντας συνεχή τροποποιήσεις στις δομικές παραμέτρους του προσδέτη και συγκλίνουν στην ελάχιστη ενεργειακή διαμόρφωση. Είναι πολύ αποτελεσματικές για ενώσεις με περιορισμένη ευελιξία όμως για ενώσεις με εκτεταμένη ευελιξία και πολλούς βαθμούς ελευθερίας, παρουσιάζουν δυσκολίες καθώς αυξάνεται εκθετικά ο χώρος αναζήτησης οδηγώντας σε ένα φαινόμενο που ονομάζεται “συνδυαστική έκρηξη”. Επίσης είναι πιθανό να παγιδευτούν σε τοπικό ελάχιστο αδυνατώντας να βρουν την καλύτερη λύση. Οι στοχαστικές μέθοδοι εξερευνούν τον διαμορφωτικό χώρο εφαρμόζοντας τυχαίες αλλαγές στις δομικές παραμέτρους ενός μορίου ή μιας ομάδας μορίων. Η προσέγγιση αυτή είναι αποδοτικότερη για ευέλικτα μόρια με πολλούς βαθμούς ελευθερίας ενώ επίσης αυξάνεται η πιθανότητα εύρεσης του ολικού ελάχιστου δηλαδή της βέλτιστης λύσης. Τρία τυπικά παραδείγματα στοχαστικών μεθόδων για διαμορφωτική αναζήτηση είναι η μέθοδος Monte Carlo (MC), η μέθοδος των γενετικών αλγορίθμων (Genetic Algorithms, GA) και η μέθοδος Tabu αναζητήσεων (Tabu Search, TS) οι οποίες και θα αναλυθούν παρακάτω [52],[53].

Η μέθοδος Monte Carlo πραγματοποιεί τυχαίες αλλαγές σε όλες τις δομικές παραμέτρους του φαρμακευτικού μορίου και σε κάθε πόζα που παράγεται, εφαρμόζεται ελαχιστοποίηση ενέργειας. Οι παραγόμενες πόζες γίνονται δεκτές ή απορρίπτονται με βάση τον παράγοντα του Boltzmann ο οποίος λαμβάνει υπόψη την απόλυτη θερμοκρασία καθώς και τις ενέργειες των λύσεων πριν και μετά από κάθε διαμορφωτική αλλαγή που συμβαίνει. Στα θετικά της μεθόδου συγκαταλέγεται το γεγονός πως οι συναρτήσεις υπολογισμού των ενεργειών είναι πιο απλές συγκριτικά με άλλες μεθόδους καθώς και το ότι είναι αποδοτική στο να ξεπερνάει ενεργειακά φράγματα εξερευνώντας έτσι καλύτερα τον διαμορφωτικό χώρο αναζήτησης [54],[55]. Οι γενετικοί αλγόριθμοι είναι επίσης μια πολύ γνωστή στοχαστική μέθοδος που στηρίζεται στην θεωρία της εξέλιξης και της φυσικής επιλογής του Δαρβίνου: “Η επικράτηση του ισχυρότερου”. Όλες οι δομικές παράμετροι κωδικοποιούνται σε ένα διάλυμα το οποίο καλείται χρωμόσωμα και αναπαριστά μια πόζα του προσδέτη δηλαδή μια πιθανή λύση. Οι δομικές παράμετροι από τις οποίες αποτελείται το χρωμόσωμα ονομάζονται γονίδια. Ο τυχαίος αλγόριθμος αναζήτησης, βασιζόμενος στο αρχικό χρωμόσωμα, παράγει έναν πληθυσμό χρωμοσωμάτων που καλύπτει ένα μεγάλο εύρος του διαθέσιμου χώρου αναζήτησης. Τα χρωμοσώματα αυτά στην συνέχεια αξιολογούνται μέσω μιας συνάρτησης αξιολόγησης και επιλέγονται εκείνα με τις χαμηλότερες τιμές ενέργειας ώστε να αποτελέσουν τα πρότυπα για την δημιουργία του νέου πληθυσμού χρωμοσωμάτων. Ο νέος πληθυσμός χρωμοσωμάτων στους γενετικούς αλγόριθμους αναφέρεται ως “απόγονοι”. Για την δημιουργία νέων υποψήφιων λύσεων εφαρμόζονται οι γενετικοί τελεστές της μετάλλαξης και της διασταύρωσης. Ο τελεστής της μετάλλαξης εφαρμόζεται τυχαία σε ένα από τα γονίδια του χρωμοσώματος εξασφαλίζοντας την ποικιλομορφία του πληθυσμού και συνεπώς την καλύτερη διερεύνηση του χώρου αναζήτησης. Ο τελεστής της διασταύρωσης εφαρμόζεται μεταξύ δύο χρωμοσωμάτων όπου ανταλλάσσουν αμοιβαία ένα τμήμα των γονιδίων τους. Τα χρωμοσώματα που παράγονται κάθε φορά αξιολογούνται από μια συνάρτηση αξιολόγησης και αυτός ο κύκλος συνεχίζεται για συγκεκριμένο αριθμό επαναλήψεων όπου και εξάγεται το τελικό χρωμόσωμα που αποτελεί το ολικό ελάχιστο δηλαδή την βέλτιστη λύση. Ο γενετικός αλγόριθμος καταφέρνει να συγκλίνει προς την βέλτιστη

λύση καθώς από τον έναν πληθυσμό προς τον επόμενο μεταφέρονται μόνο εκείνα τα χρωμοσώματα με την χαμηλότερη ενέργεια που αποτελούνται από ευνοϊκά χαρακτηριστικά. Με αυτόν τον τρόπο όχι μόνο ο αλγόριθμος συγκλίνει προς το ολικό ελάχιστο αλλά μειώνει επίσης τον χρόνο και το υπολογιστικό κόστος που απαιτείται για την ολοκλήρωση της διαμορφωτικής αναζήτησης [50],[52],[56]. Τέλος η αναζήτηση Tabu αποτελεί επίσης μια στοχαστική μέθοδο που πραγματοποιεί τυχαίες αλλαγές στις δομικές ιδιότητες των προσδετών και όπως και οι προηγούμενες μέθοδοι κρατάει εκείνες με την χαμηλότερη ενέργεια. Το διαφορετικό που προσφέρει αυτός ο αλγόριθμος είναι η λίστα Tabu που περιέχει όλες τις λύσεις χαμηλής ενέργειας που έχουν παραχθεί στο παρελθόν. Μια νέα πόζα-υποψήφια λύση αξιολογείται μόνο εάν δεν είναι όμοια με τις πόζες που υπάρχουν στην λίστα, αποτρέποντας με αυτόν τον τρόπο τον αλγόριθμο να εξετάσει περιοχές του χώρου που έχουν ερευνηθεί, οδηγώντας τον έτσι σε νέες μη επισκέψιμες περιοχές [54].

1.8.4 Συναρτήσεις Βαθμολόγησης

Ο αλγόριθμος διαμορφωτικής αναζήτησης διερευνά ένα μεγάλο πλήθος διαμορφωτικών καταστάσεων του προσδέτη οι οποίες στην συνέχεια πρέπει να αξιολογηθούν από μια συνάρτηση βαθμολόγησης. Η συνάρτηση βαθμολόγησης εκτιμά την ενέργεια πρόσδεσης του φαρμακευτικού μορίου στον υποδοχέα και ξεχωρίζει εκείνες τις διαμορφώσεις με τις χαμηλότερες τιμές ενέργειας. Η εκτίμηση της ενέργειας βασίζεται στην σταθερά πρόσδεσης (K_d) και στον υπολογισμό της ελεύθερης ενέργειας Gibbs (ΔG) η οποία λαμβάνει υπόψη διάφορες φυσικοχημικές ιδιότητες που αφορούν την δημιουργία του συμπλόκου όπως είναι οι διαμοριακές αλληλεπιδράσεις και οι εντροπικοί παράγοντες. Όσες περισσότερες είναι οι παράμετροι που χρησιμοποιούνται για την πρόβλεψη της ενέργειας πρόσδεσης τόσο πιο ακριβής είναι ο υπολογισμός εις βάρος όμως του χρόνου και των υπολογιστικών πόρων που απαιτούνται. Οι συναρτήσεις βαθμολόγησης διακρίνονται στις εξής τρεις κατηγορίες: συνάρτηση βαθμολόγησης βασισμένη στο δυναμικό πεδίο, εμπειρική και βασισμένη στην γνώση [52],[54].

Η συνάρτηση βαθμολόγησης που βασίζεται στο δυναμικό πεδίο αποτελείται από το άθροισμα της εσωτερικής ενέργειας του προσδέτη και της ενέργειας αλληλεπίδρασης που αναπτύσσεται κατά την δημιουργία του συμπλόκου και αφορά ηλεκτροστατικές και van der Waals αλληλεπιδράσεις. Οι van der Waals αλληλεπιδράσεις υπολογίζονται από το μαθηματικό μοντέλο Lennard-Jones ενώ οι ηλεκτροστατικές αλληλεπιδράσεις μέσω της μεθόδου του Coulomb. Λόγω της δυσκολίας μοντελοποίησης του περιβάλλοντος της πρωτεΐνης, η μέθοδος του Coulomb συνδυάζεται με μια διηλεκτρική συνάρτηση εξαρτώμενη από την απόσταση η οποία μεταβάλλει την συνεισφορά των αλληλεπιδράσεων φορτίου-φορτίου ανάλογα με την απόσταση των φορτισμένων ατόμων. Η συνάρτηση βαθμολόγησης βασισμένη στο δυναμικό πεδίο ακολουθεί την μέθοδο “ab initio” που σημαίνει “από την αρχή” χωρίς δηλαδή να βασίζεται σε πειραματικά δεδομένα και εφαρμόζει εξισώσεις της κλασικής μηχανικής για τον υπολογισμό κάθε όρου της εξίσωσης. Η βασικοί περιορισμοί της συνάρτησης σχετίζονται με την αδυναμία υπολογισμού των εντροπικών παραγόντων, των ενεργειών αποδιαλύτωσης και στην ανακρίβεια υπολογισμού των όρων που σχετίζονται με την πρόσδεση και βρίσκονται σε μεγάλη απόσταση από το σημείο πρόσδεσης [50],[52],[54],[55].

Η εμπειρική συνάρτηση βαθμολόγησης αποτελεί μια ακόμα μέθοδο αξιολόγησης της πρόσδεσης. Η συνάρτηση περιλαμβάνει μη σχετιζόμενους, σταθμισμένους όρους που εμπλέκονται στην πρόσδεση του μορίου στον υποδοχέα όπως είναι οι ηλεκτροστατικές αλληλεπιδράσεις, οι αλληλεπιδράσεις van der Waals, οι δεσμοί υδρογόνου, οι ενέργειες αποδιαλύτωσης και επιδιαλύτωσης και οι εντροπικοί παράγοντες. Οι συντελεστές που χρησιμοποιούνται για να σταθμιστεί κάθε μεταβλητή της συνάρτησης εξάγονται μέσω της εφαρμογής της ανάλυσης παλινδρόμησης σε ένα σύνολο συμπλόκων πρωτεΐνης-προσδέτη με γνωστή συγγένεια δέσμευσης που χρησιμοποιείται ως σετ εκπαίδευσης. Προκύπτει τελικά μια συνάρτηση υπολογισμού της ελεύθερης ενέργειας (σχέση 9) με σταθμισμένες τις μεταβλητές της. Η ακρίβεια του υπολογισμού εξαρτάται άμεσα από τα γνωστά δεδομένα που χρησιμοποιήθηκαν για τον σχεδιασμό του μοντέλου [52],[54].

$$\Delta G_{\text{binding}} = c_1 \Delta G_{\text{vdW}} + c_2 \Delta G_{\text{Hbond}} + c_3 \Delta G_{\text{entropy}} + \dots \quad (9)$$

Η συνάρτηση βαθμολόγησης που βασίζεται στην γνώση αντλεί πληροφορίες από την στατιστική ανάλυση συμπλόκων πρωτεΐνης-προσδέτη που έχουν δοκιμαστεί πειραματικά. Συγκεκριμένα εξάγει ενεργειακά δυναμικά που αντιστοιχούν σε διάφορα ζεύγη ατόμων μεταξύ πρωτεΐνης και προσδέτη. Οι αλληλεπιδράσεις που εντοπίζονται ταξινομούνται ανάλογα με την συχνότητα εμφάνισής τους. Έτσι μια αλληλεπίδραση με υψηλή συχνότητα εμφάνισης κατά την στατιστική ανάλυση των πειραματικών δεδομένων καταχωρείται με μεγάλο συντελεστή βαρύτητας υποδηλώνοντας την σημαντικότητα της στην δημιουργία ενός συμπλόκου. Το τελικό αποτέλεσμα της συνάρτησης βαθμολόγησης βάση γνώσης αποτελείται από το άθροισμα όλων των σταθμισμένων ενεργειών. Η μέθοδος αυτή χαρακτηρίζεται από ισορροπία μεταξύ ακρίβειας και χρόνου που απαιτείται για την ολοκλήρωση της και είναι κατάλληλη για εφαρμογές εικονικής σάρωσης μεγάλων χημειοθηκών όμως εξαρτάται σε πολύ μεγάλο βαθμό από το πλήθος και την ποικιλομορφία των διαθέσιμων πειραματικών δεδομένων [52],[54].

Όλες οι παραπάνω μέθοδοι αξιολόγησης της πρόσδεσης παρουσιάζουν θετικά και αρνητικά στοιχεία. Δεν υπάρχει μια καθολική συνάρτηση βαθμολόγησης που να λειτουργεί αποδοτικά σε κάθε πρόβλημα. Προτείνεται η συνάρτηση βαθμολόγησης να επιλέγεται με βάση την κάθε περίπτωση ενώ μια αποδοτική προσέγγιση είναι η επιλογή εκείνης της συνάρτησης που έχει δοκιμαστεί για την συγκεκριμένη οικογένεια πρωτεΐνης και την κατηγορία του μορίου που μελετάται. Τέλος μερικές φορές επιλέγεται ο συνδυασμός διαφορετικών συναρτήσεων βαθμολόγησης με σκοπό την βελτίωση τόσο της αξιοπιστίας όσο και της ακρίβειας της αξιολόγησης της πρόσδεσης [52],[54].

1.9 Εικονική Σάρωση

1.9.1 Εισαγωγή στην Εικονική Σάρωση

Η ανακάλυψη και ο σχεδιασμός ενός νέου φαρμάκου αποτελεί μια πολύ χρονοβόρα και δαπανηρή διαδικασία. Από το πρώτο στάδιο της ανακάλυψης ενός φαρμάκου έως την προώθηση του στην αγορά χρειάζονται περίπου 10 με 15 χρόνια ή και περισσότερα ενώ σύμφωνα με μελέτη που έγινε το 2014, απαιτούνται κατά μέσο όρο 2.87 δισεκατομμύρια δολάρια. Το πιο απαιτητικό κομμάτι κατά την διαδικασία ανάπτυξης

ενός φαρμάκου είναι ο εντοπισμός του φαρμάκου-οδηγού μεταξύ εκατομμυρίων άλλων, κάτι που πειραματικά επιτυγχάνεται μέσω της σάρωσης υψηλής ανάλυσης (High-throughput Screening, HTS). Τα τελευταία χρόνια λόγω της ραγδαίας εξέλιξης της τεχνολογίας, της βελτιστοποίησης αλγορίθμων, της απόκτησης επιπρόσθετων γνώσεων σχετικά με τις φυσικοχημικές ιδιότητες των ενώσεων και τις δομικές και λειτουργικές ιδιότητες των πρωτεϊνών-στόχων έχει αναπτυχθεί ένα αποτελεσματικό υπολογιστικό εργαλείο για την ανακάλυψη φαρμάκων, η εικονική σάρωση (Virtual screening, VS). Η VS είναι μια υπολογιστική τεχνική που δρα συνδυαστικά με την HTS μειώνοντας τόσο το κόστος όσο και τον χρόνο που απαιτείται για την ολοκλήρωση της διαδικασίας [57],[58],[59]. Η μέθοδος VS δίνει την δυνατότητα σάρωσης μεγάλων βάσεων δεδομένων χημικών ενώσεων σε γνωστές κρυσταλλικές δομές ή ομόλογα μοντέλα, ώστε να περιορίσει τις ενώσεις που θα προχωρήσουν σε σύνθεση και πειραματική δοκιμή προκειμένου να αξιολογηθούν. Μια ένωση που θα ξεχωρίσει από την VS είναι εφικτό να δοκιμαστεί πειραματικά και να προχωρήσει σε διαδικασία βελτιστοποίησης ώστε να προσαρμοστεί η βιολογική της δραστηριότητα ή το φαρμακοκινητικό της προφίλ. Επιπλέον σε αντίθεση με την HTS που υστερεί στην ικανότητα της να προσπεράσει το στάδιο ADME, η VS παρουσιάζει πολύ καλή αποδοτικότητα στην πρόβλεψη της ADME και της τοξικότητας χημικών ενώσεων, με την βοήθεια σχεσιακών βάσεων δεδομένων που έχουν δημιουργηθεί για αυτόν τον σκοπό [58].

1.9.2 Βιβλιοθήκες Χημικών Ενώσεων

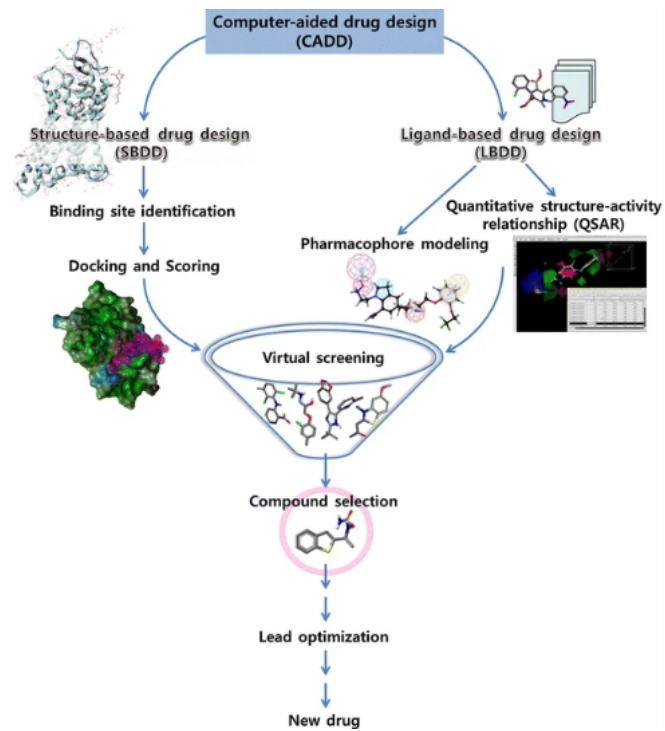
Η επιλογή της κατάλληλης βιβλιοθήκης χημικών ενώσεων είναι πολύ σημαντική κατά την διαδικασία εφαρμογής της εικονικής σάρωσης. Υπάρχουν πολλές διαθέσιμες βάσεις δεδομένων, είτε δωρεάν είτε εμπορικές, που διαφέρουν στο μέγεθος τους καθώς και στο είδος των ενώσεων που περιέχουν. Οι βασικοί τύποι που μπορούν να κατηγοριοποιηθούν οι χημικές βιβλιοθήκες είναι οι εξής: (α) βάσεις δεδομένων χημικών ενώσεων βασισμένες σε αναφορές (chemical substance reference databases) οι οποίες αποτελούν πολύ μεγάλα, δωρεάν αποθετήρια με ενώσεις που έχουν συλλεχθεί από διάφορες πηγές, (β) βάσεις δεδομένων βιολογικά ενεργών ενώσεων (databases of biologically active compounds) που περιέχουν ενώσεις με γνωστή βιολογική δράση, (γ) “fine chemical databases” στις οποίες περιέχονται ενώσεις που δεν έχουν γνωστή βιολογική δραστηριότητα όμως μπορούν να χρησιμοποιηθούν ως αρχικές δομές που θα βελτιστοποιηθούν στην συνέχεια και (δ) βάσεις δεδομένων σάρωσης (screening databases) που όπως αναφέρεται και από το όνομα αποτελούν κατάλληλα αποθετήρια για σάρωση. Οι πιο συχνά χρησιμοποιούμενες, ανοιχτής πρόσβασης, βάσεις δεδομένων είναι η ZINC, η PubChem, η ChEMBL και η DrugBank [59]. Μετά από την επιλογή της χημειοθήκης από όπου θα προμηθευτούν οι ενώσεις, ακολουθεί το φιλτράρισμα της για συγκεκριμένες φυσικοχημικές ιδιότητες που συμβάλλουν στην βελτίωση της φαρμακοκινητικής του φαρμάκου στον ανθρώπινο οργανισμό συμπεριλαμβανομένης της απορροφητικότητας, της κατανομής, του μεταβολισμού και της απέκκρισης του. Μια από τις πιο γνωστές μεθόδους για το φιλτράρισμα βιβλιοθηκών χημικών ενώσεων είναι ο κανόνας των πέντε του Lipinski. Πήρε το όνομα του από τα κατώφλια των παραμέτρων στα οποία αναφέρεται και έχουν τιμή κοντά στο πέντε ή πολλαπλάσια με το πέντε. Σύμφωνα με τον κανόνα του Lipinski υπάρχουν τέσσερις παράμετροι οι οποίες σχετίζονται με την διαλυτότητα και την διαπερατότητα των ενώσεων. Συγκεκριμένα οι ενώσεις παρουσιάζουν καλύτερη φαρμακοκινητική στο ανθρώπινο σώμα όταν: (α) Το μοριακό βάρος δεν υπερβαίνει τα 500 Dalton, (β) ο συντελεστής

κατανομής οκτανόλης/νερού (logP) είναι μικρότερος του 5, (γ) οι δότες δεσμού υδρογόνου δεν υπερβαίνουν τους 5 και (δ) οι δέκτες δεσμού υδρογόνου δεν υπερβαίνουν τους 10 [59],[60].

1.9.3 Εικονική Σάρωση Βάση Δομής και Βάση Προσδέτη

Η εικονική σάρωση διακρίνεται σε δύο κατηγορίες, στην εικονική σάρωση βάση δομής (Structure Based Virtual Screening, SBVS) και στην εικονική σάρωση βάση προσδέτη (Ligand Based Virtual Screening, LBVS) [61]. Στην εικόνα 1.14 παρουσιάζεται ο διαχωρισμός των δύο διαφορετικών μεθόδων. Στην περίπτωση της SBVS είναι γνωστή τόσο η δομή των ενώσεων όσο και η τρισδιάστατη δομή της πρωτεΐνης-στόχου επιτρέποντας την μελέτη των αλληλεπιδράσεων που γίνονται μεταξύ τους. Μέσω της ανάλυσης της δομής του υποδοχέα καθώς και της ανάλυσης των μεταλλάξεων στο ενεργό κέντρο, δίνεται η δυνατότητα να ξεχωρίσουν τα σημαντικά αμινοξέα που συμμετέχουν στην πρόσδεση των μορίων, πληροφορία πολύ χρήσιμη κατά την διαδικασία σχεδιασμού και ανακάλυψης φαρμάκων. Στην SBVS χρησιμοποιείται συνήθως η μέθοδος της μοριακής πρόσδεσης, με τις ενώσεις να τοποθετούνται στην περιοχή δέσμευσης της πρωτεΐνης και να εξάγεται ένα σκορ ανάλογα με την συγγένεια πρόσδεσης μεταξύ προσδέτη και υποδοχέα όπως αναλύθηκε και σε προηγούμενο κεφάλαιο. Επίσης μια ακόμα μέθοδος που χρησιμοποιείται τόσο στην SBVS όσο και στην LBVS είναι τα φαρμακοφόρα μοντέλα. Βασίζονται στην χαρτογράφηση ηλεκτρονιακών και στερεικών χαρακτηριστικών που αφορούν την αλληλεπίδραση μεταξύ προσδέτη και υποδοχέα τα οποία στην συνέχεια χρησιμοποιούνται για την αναζήτηση σε μια χημειοθήκη με σκοπό τον εντοπισμό εκείνων των ενώσεων με τα συγκεκριμένα χαρακτηριστικά. Η προσέγγιση της LBVS επιλέγεται όταν δεν είναι γνωστή η κρυσταλλική δομή του υποδοχέα ή η δημιουργία ενός ομόλογου μοντέλου παρουσιάζει δυσκολίες. Στην LBVS δεν απαιτείται η γνώση της τρισδιάστατης δομής του υποδοχέα αλλά η γνώση των χημικών ιδιοτήτων δοκιμασμένων ενώσεων με γνωστή βιολογική δράση. Η λογική της βασίζεται στο γεγονός πως ενώσεις με παρόμοια δομική ομοιότητα με τις δοκιμασμένες ενώσεις θα εμφανίσουν παρόμοια βιολογική δραστηριότητα. Κατά την εφαρμογή της, εκμεταλλεύεται και συσχετίζει την πληροφορία που εξάγεται τόσο από τις βιολογικά ενεργές όσο και ανενεργές ενώσεις καθιστώντας την κατάλληλη τεχνική για την πρόβλεψη της βιολογικής δράσης άγνωστων ενώσεων. Εκτός από τα πλεονεκτήματα που αναλύθηκαν παραπάνω, η LBVS παρουσιάζει και μειονεκτήματα με βασικό την αδυναμία πρόβλεψης της σωστής διαμόρφωσης της ένωσης λόγω της έλλειψης πληροφορίας σχετικά με την δομή του υποδοχέα [59],[62],[63]. Οι μέθοδοι που χρησιμοποιούνται συνήθως κατά την εφαρμογή της LBVS επιπρόσθετα των φαρμακοφόρων μοντέλων είναι η ποσοτική σχέση δομής-δραστικότητας (Quantitative Structure-Activity Relationship, QSAR) και η αναζήτηση ομοιότητας. Σε πολλές περιπτώσεις οι ερευνητές επιλέγουν την προσέγγιση της υβριδικής εικονικής σάρωσης όπου χρησιμοποιούν συνδυαστικά και τις δύο τεχνικές (SBVS και LBVS). Η εικονική σάρωση χρησιμοποιείται πλέον ευρέως στον τομέα του σχεδιασμού και της ανακάλυψης φαρμάκων και έχει συμβάλει ουσιαστικά στην ανάπτυξη μικρών θεραπευτικών μορίων πολλά εκ των οποίων εξελίχθηκαν σε εμπορικά διαθέσιμα φάρμακα [58].

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ



Εικόνα 1.14: Απλοποιημένο διάγραμμα ροής CADD. Περιλαμβάνει την επιλογή μεταξύ των δύο διαφορετικών κατευθύνσεων SBDD-LBDD, την εφαρμογή της εικονικής σάρωσης, την επιλογή ενώσεων, την βελτιστοποίηση τους και την πρόταση του υποψήφιου φαρμάκου για πειραματική αξιολόγηση [64].

2. Μέθοδοι

2.1 Πρωτεΐνη – Στόχος

Η πρωτεΐνη-στόχος της παρούσας εργασίας είναι ο υποδοχέας ορμόνης απελευθέρωσης αυξητικής ορμόνης (GHRH-R). Για τις ανάγκες και τους στόχους της εργασίας υπήρχε η απαίτηση ο υποδοχέας να βρίσκεται τόσο στην ενεργή όσο και στην ανενεργή του κατάσταση. Η δομή της ενεργής κατάστασης του υποδοχέα έχει επιλυθεί πειραματικά με την χρήση της κρύο-ηλεκτρονικής μικροσκοπίας (Cryo-EM) και διατίθεται στην βάση δεδομένων Protein Data Bank (PDB) με τον τετραψήφιο κωδικό “7CZ5”. Ειδικότερα μαζί με τον υποδοχέα GHRH-R έχει συγκρυσταλλωθεί επίσης το πεπτίδιο GHRH, η ετεροτριμερής πρωτεΐνη G_s και το Nanobody που χρησιμοποιήθηκε για την σταθεροποίηση του συμπλόκου κατά την διαδικασία χαρακτηρισμού του. Η ενεργή κατάσταση του υποδοχέα αντλήθηκε ηλεκτρονικά από την Protein Data Bank σε “pdb” μορφή. Αντιθέτως η ανενεργή κατάσταση του υποδοχέα μέχρι και σήμερα δεν έχει επιλυθεί πειραματικά. Για τον λόγο αυτό χρησιμοποιήθηκε το μοντέλο που έχει δημιουργηθεί με σχετικά καλή ακρίβεια από το λογισμικό πρόβλεψης δομής πρωτεϊνών AlphaFold. Το μοντέλο της AlphaFold διατίθεται στην GPCR Data Base (GPCRdb) όπου αντλήθηκε ηλεκτρονικά σε “pdb” μορφή.

2.2 Προσδέτες

Πριν γίνει αναφορά στους προσδέτες που χρησιμοποιήθηκαν, είναι σημαντικό να γίνει μια επισκόπηση της διαδικασίας που έχει ήδη πραγματοποιηθεί. Η παρούσα εργασία είναι επέκταση της μελέτης που έχει προηγηθεί για τον εντοπισμό μη πεπτιδικών μικρών μορίων που στοχεύουν τον GHRH-R και έχει υλοποιηθεί μέσω της συνεργασίας του κέντρου Νευροενδοκρινολογίας του πανεπιστημίου της Πρετόρια με τον Δρ. Μίνω Μαρτσούκα, υπολογιστικό χημικό του τμήματος Μηχανικών Βιοϊατρικής του Πανεπιστημίου Δυτικής Αττικής. Η μελέτη αυτή αποτελείται από τον συνδυασμό πειραματικών (in vitro) και υπολογιστικών (in silico) δοκιμασιών που στόχο έχουν την εύρεση πιθανών αγωνιστών και ανταγωνιστών έναντι του GHRH-R. Η μελέτη απαιτούσε την γνώση της δομής ολόκληρου του υποδοχέα δηλαδή το εξωκυττάριο κομμάτι (ECD) μαζί με το διαμεμβρανικό (TM) κομμάτι. Ακόμα και σήμερα ολόκληρη η δομή του GHRH-R δεν έχει κρυσταλλογραφηθεί. Μάλιστα κατά την περίοδο της μελέτης δεν είχε ακόμα επιλυθεί πειραματικά το διαμεμβρανικό κομμάτι του υποδοχέα. Για τον σκοπό αυτό ο Δρ. Μίνω Μαρτσούκας έφτιαξε ένα ομόλογο μοντέλο του GHRH-R στο οποίο είναι προσδεμένο το πεπτίδιο GHRH. Το ECD κομμάτι του μοντέλου βασίστηκε στην κρυσταλλική δομή του συμπλόκου του πεπτιδίου 1 τύπου γλυκαγόνης (Glucagon Like Peptide 1 – GLP 1) με τον υποδοχέα του GLP-1R ενώ το TM κομμάτι βασίστηκε στην κρυσταλλική δομή του υποδοχέα γλυκαγόνης (Glucagon Receptor - GCGR). Χρησιμοποιώντας το παραπάνω μοντέλο υλοποιήθηκε μοριακή πρόσδεση βασισμένη στην δομή του υποδοχέα, 5.5 εκατομμυρίων άκαμπτων χημικών ενώσεων από τις οποίες ξεχώρισαν λόγω του καλού σκορ πρόσδεσης οι 2000 ενώσεις. Έπειτα από την ενδελεχή ανάλυση και αξιολόγηση τους, επιλέχθηκαν οι 50 (MK01-MK50) από αυτές για περαιτέρω μελέτη. Συγκεκριμένα στο κέντρο Νευροενδοκρινολογίας ακολούθησε μια σειρά από πειραματικές διαδικασίες από τις οποίες διαπιστώθηκε ότι η ένωση MK24 έχει αγωνιστική δράση ενώ η ένωση MK04

έχει ανταγωνιστική δράση. Με βάση τα παραπάνω πειραματικά αποτελέσματα ο Δρ. Μίνως Ματσούκας εφάρμοσε εικονική σάρωση σε μια βιβλιοθήκη 7 εκατομμυρίων χημικών ενώσεων από την Molport με σκοπό τον εντοπισμό παρόμοιων δομικά ενώσεων με τις MK04 και MK24. Εκτός από την αναζήτηση δομικής ομοιότητας με τις αρχικές ενώσεις η οποία υλοποιήθηκε με χρήση του αλγορίθμου Tanimoto, λήφθηκαν υπόψη και χαρακτηριστικά που βελτιστοποιούν την πρόσδεση των ενώσεων με τον υποδοχέα. Προέκυψαν 22 νέες ενώσεις με τις 9 από αυτές να έχουν βασιστεί στην δομή του MK04 και 13 στην δομή του MK24. Ακολούθησε και πειραματική αξιολόγηση από όπου και προέκυψε ότι οι ενώσεις MK04-6 και MK04-9 έχουν ανταγωνιστική δράση [65],[66],[67].

Σε συνέχεια της παραπάνω μελέτης, οι 50 ενώσεις MK(01-50) που ξεχώρισαν στην πρώτη φάση καθώς και οι 9 ενώσεις MK04(2-10) που επιλέχθηκαν λόγω της ικανότητας της ένωσης MK04 να δρα ως ανταγωνιστής, αντλήθηκαν σε ηλεκτρονική μορφή για περαιτέρω ανάλυση. Ειδικότερα δύο από τους προσδέτες (MK16 και MK28) που θα χρησιμοποιηθούν στην διαδικασία της μοριακής πρόσδεσης αντλήθηκαν ηλεκτρονικά από την Zinc20, μια δωρεάν βάση δεδομένων εμπορικά διαθέσιμων ενώσεων με πάνω από 230 εκατομμύρια ενώσεις που παρέχει τις απαραίτητες πληροφορίες για κάθε μια από αυτές. Οι υπόλοιποι 57 προσδέτες αντλήθηκαν ηλεκτρονικά από την Molport, μια ευρέως γνωστή και χρησιμοποιούμενη διαδικτυακή πλατφόρμα όπου επιστήμονες μπορούν να αναζητήσουν και να προμηθευτούν τις απαραίτητες χημικές ενώσεις για τις ανάγκες τους.

Τέλος εφαρμόστηκε η υπολογιστική τεχνική της εικονικής σάρωσης με στόχο την ανακάλυψη δυναμικά βιοδραστικών μορίων που έχουν ανασταλτική δράση και μειώνουν την διέγερση του GHRH-R από το ενδογενές του πεπτίδιο GHRH. Για τον σκοπό αυτό αντλήθηκαν ηλεκτρονικά από την Selleckchem 14.511 βιοδραστικές ενώσεις χωρισμένες σε δύο βιβλιοθήκες, την Bioactive Compound Library I με 9.202 ενώσεις και την Bioactive Compound Library II με 5.309 ενώσεις. Ο λόγος που επιλέχθηκαν οι συγκεκριμένες βιβλιοθήκες της Selleckchem είναι διότι αποτελούνται από ενώσεις που έχουν δραστηριότητα έναντι άλλων πρωτεϊνών. Συνεπώς είναι πιθανό να παρουσιάσουν καλά χαρακτηριστικά και για τον GHRH-R.

2.3 Λογισμικά Χημειοπληροφορικής

Για την διεκπεραίωση των διαδικασιών της εργασίας χρησιμοποιήθηκε μια σειρά από λογισμικά χημειοπληροφορικής η χρήση των οποίων παρουσιάζεται παρακάτω. Το PyMOL είναι το λογισμικό που επιλέχθηκε για την αναπαράσταση των χημικών δομών και την προετοιμασία των εικόνων. Η επεξεργασία των χημικών ενώσεων που προέκυψαν από την πειραματική διαδικασία καθώς και οι ενώσεις που αντλήθηκαν από την Selleckchem υλοποιήθηκε μέσω του Open Babel. Η προετοιμασία της πρωτεΐνης στόχου GHRH-R τόσο στην ενεργή όσο και στην ανενεργή κατάσταση έγινε αρχικά στο PyMOL και ολοκληρώθηκε στο AutoDockTools. Στο Data Warrior έγινε το φιλτράρισμα των ενώσεων με συγκεκριμένους μοριακούς περιγραφείς ενώ μέσω του AutoDock Vina πραγματοποιήθηκαν οι υπολογισμοί της μοριακής πρόσδεσης. Η επεξεργασία των προσδετών καθώς και η εφαρμογή της μοριακής πρόσδεσης έγινε σε περιβάλλον Linux μέσω της γλώσσας εντολών bash ενώ όπου χρειάστηκε διαβάστηκαν και εκτελέστηκαν εντολές από αρχεία script shell. Η αυτοματοποίηση και ανάλυση των

αποτελεσμάτων της εικονικής σάρωσης έγινε σε γλώσσα προγραμματισμού python 3.10.6 σε περιβάλλον Spyder καθώς και στο Colab της Google.

2.4 Αξιολόγηση βιοδραστικών ενώσεων της σειράς MK

Με σκοπό την αξιολόγηση και την μελέτη του μηχανισμού δράσης των βιοδραστικών ενώσεων της σειράς MK εφαρμόστηκε η υπολογιστική τεχνική της μοριακής πρόσδεσης. Μελετήθηκε τόσο ποσοτικά μέσω του σκορ πρόσδεσης όσο και ποιοτικά μέσω της ανάλυσης των αλληλεπιδράσεων, η δέσμευση των χημικών ενώσεων με τον υποδοχέα-στόχο. Για την υλοποίηση της μοριακής πρόσδεσης είναι σημαντικό να προετοιμαστούν τόσο οι προσδέτες όσο και η πρωτεΐνη καθώς και να ρυθμιστεί ένα σύνολο παραμέτρων που έχουν καταλυτική σημασία στην επιτυχημένη εφαρμογή της διαδικασίας.

2.4.1 Προετοιμασία Προσδετών

Για την προετοιμασία των προσδετών χρησιμοποιήθηκε το λογισμικό Open Babel. Το Open Babel είναι ένα ελεύθερο λογισμικό και αποτελεί ένα βασικό εργαλείο στον τομέα της χημειοπληροφορικής και της υπολογιστικής χημείας. Χρησιμοποιείται για την αναζήτηση, την μετατροπή, την ανάλυση και την αποθήκευση χημικών δεδομένων. Οι 57 ενώσεις από την Molport αντλήθηκαν ηλεκτρονικά σε δισδιάστατη μορφή (2D sdf). Αρχικά έγινε η μετατροπή τους σε τρισδιάστατη μορφή (3D sdf), έπειτα αφαιρέθηκαν τα άλατα και ρυθμίστηκε το pH=7 ώστε να προσομοιώνει τις συνθήκες του ανθρώπινου οργανισμού. Τέλος οι προσδέτες μετατράπηκαν σε “pdbqt” μορφή καθώς είναι αυτή που απαιτείται για να γίνει η μοριακή πρόσδεση. Οι 2 εναπομείναντες ενώσεις αντλήθηκαν ηλεκτρονικά από την Zinc20 σε μορφή Smiles. Μετατράπηκαν σε δισδιάστατη μορφή (2D sdf) και από εκεί και πέρα η διαδικασία είναι πανομοιότυπη με αυτή που ακολουθήθηκε για τις υπόλοιπες ενώσεις.

2.4.2 Προετοιμασία Πρωτεΐνης – Επιλογή περιοχής πρόσδεσης

Ακολουθεί η προετοιμασία της ενεργής κατάστασης του υποδοχέα GHRH-R. Μέσω του PyMOL απομακρύνθηκαν όλα τα μέρη της δομής εκτός από τον GHRH-R. Συγκεκριμένα απομακρύνθηκαν το πεπτιδίο GHRH, η ετεροτριμερής πρωτεΐνη G_s, το Nanobody που χρησιμοποιήθηκε για την σταθεροποίηση του συμπλόκου κατά την διαδικασία χαρακτηρισμού του καθώς και συγκρυσταλλωμένα λιπίδια. Η προετοιμασία συνεχίστηκε στο AutoDockTools όπου αφαιρέθηκαν τα νερά για να μην επηρεάσουν τις αλληλεπιδράσεις κατά την διάρκεια των υπολογισμών, προστέθηκαν πολικά υδρογόνα ώστε να μελετηθούν οι πολικές αλληλεπιδράσεις και τέλος υπολογίστηκαν τα μερικά φορτία των ατόμων με την μέθοδο Kollman καθώς η φύση των αλληλεπιδράσεων είναι ηλεκτροστατική. Στην συνέχεια έγινε η προετοιμασία του μοντέλου του υποδοχέα GHRH-R που βρίσκεται στην ανενεργή κατάσταση. Η δομή του υποδοχέα περιλαμβάνει εκτός από το διαμεμβρανικό και το εξωκυττάριο μέρος. Για την προετοιμασία και εδώ αφαιρέθηκαν τα νερά, προστέθηκαν πολικά υδρογόνα και υπολογίστηκαν τα μερικά φορτία των ατόμων με την μέθοδο Kollman. Τόσο στην

ενεργή όσο και στην ανενεργή κατάσταση η πρωτεΐνη αποθηκεύτηκε σε “pdbqt” μορφή καθώς είναι αυτή που απαιτείται για την διαδικασία της μοριακής πρόσδεσης.

Οι δύο καταστάσεις του υποδοχέα, η ενεργή και η ανενεργή, έχουν καταχωρηθεί με διαφορετικές συντεταγμένες στις αντίστοιχες βάσεις δεδομένων. Καθώς η περιοχή που θα γίνει η μοριακή πρόσδεση θα είναι κοινή, πρέπει και οι δύο καταστάσεις του υποδοχέα να έχουν τις ίδιες συντεταγμένες. Αυτό επιτυγχάνεται μέσω της υπέρθεσης της ανενεργής κατάστασης του GHRH-R στην ενεργή, με την εντολή “align” του PyMOL. Μετά την προετοιμασία τους, οι πρωτεΐνες είναι έτοιμες για το εφαρμογή της μοριακής πρόσδεσης.

2.4.3 Εφαρμογή Μοριακής Πρόσδεσης

Η μοριακή πρόσδεση των 59 ενώσεων έγινε στην ορθοστερική θέση της διαμεμβρανικής περιοχής του υποδοχέα, δηλαδή στην θέση που δεσμεύεται το ενδογενές πεπτίδιο. Για την εφαρμογή της μοριακής πρόσδεσης ορίστηκε το grid box δηλαδή η περιοχή που θα γίνει η αναζήτηση της καταλληλότερης διαμόρφωσης του μορίου στον υποδοχέα. Χρησιμοποιήθηκε grid box με “spacing=1” (δηλαδή αποστάσεις σε Angstrom του grid box) και διαστάσεις 20×20×25 Angstrom ώστε να περιλαμβάνει όλο το ορθοστερικό κέντρο που είναι το σημείο ενδιαφέροντος. Κατά την διάρκεια των υπολογισμών ο υποδοχέας ήταν σταθερός ενώ οι ενώσεις ήταν ελεύθερες να πάρουν διάφορες διαμορφώσεις στο ενεργό κέντρο. Για τις διάφορες παραμέτρους (cpu, exhaustiveness, max_evals, min_rmsd, energy_range, num_modes, verbosity) χρησιμοποιήθηκαν οι προκαθορισμένες τιμές. Για παράδειγμα η εξαντλητικότητα, που χαρακτηρίζει το πόσο ενδεδειγμένη θα είναι η αναζήτηση των διαμορφώσεων του μορίου ορίστηκε ίση με “exhaustiveness=8”. Όσο μεγαλύτερη είναι αυτή η τιμή τόσο περισσότερες διαμορφώσεις του μορίου αξιολογούνται αυξάνοντας τις πιθανότητες να βρεθεί η βέλτιστη λύση εις βάρος όμως του υπολογιστικού κόστους. Οι πόζες δηλαδή οι διαφορετικές διαμορφώσεις του μορίου που θα παραχθούν ορίστηκαν ίσες με “num_modes=9”.

2.4.4 Αξιολόγηση Μοριακής Πρόσδεσης

Ως αποτέλεσμα της μοριακής πρόσδεσης εξάγεται για κάθε ένωση ένας πίνακας όπου στην πρώτη στήλη έχει τον αριθμό κάθε πόζας, στην δεύτερη στήλη το σκορ πρόσδεσης που εκφράζεται ως “affinity” με μονάδες μέτρησης τα kcal/mol, στην τρίτη και τέταρτη στήλη αντίστοιχα έχει το κάτω και το πάνω όριο της απόστασης από το RMSD (Root Mean Square Deviation). Η αξιολόγηση της μοριακής πρόσδεσης γίνεται μέσω του σκορ πρόσδεσης το οποίο παίρνει αρνητικές τιμές και αποτελεί μια ένδειξη, για το πόσο καλή είναι η δέσμευση του μορίου στον στόχο. Όσο πιο αρνητική είναι αυτή η τιμή τόσο πιο σταθερή και ισχυρή είναι η δέσμευση της ένωσης στον υποδοχέα. Τα σκορ πρόσδεσης για κάθε ένωση συγκρίθηκαν με τα πειραματικά αποτελέσματα.

Εκτός από την ποσοτική ανάλυση μέσω του σκορ πρόσδεσης, οι ενώσεις που ξεχώρισαν πειραματικά αναλύθηκαν περαιτέρω με σκοπό τον προσδιορισμό του μηχανισμού πρόσδεσής τους στον υποδοχέα. Συγκεκριμένα μελετήθηκαν οι αλληλεπιδράσεις του υποδοχέα με τις ενώσεις MK24, MK04, MK04-6 και MK04-9. Για το μόριο με αγωνιστική δράση (MK24) η ανάλυση έγινε χρησιμοποιώντας την

ενεργή κατάσταση του υποδοχέα ενώ για τις ενώσεις με ανταγωνιστική δράση (MK04, MK04-6 και MK04-9), η ανάλυση έγινε χρησιμοποιώντας την ανενεργή κατάσταση του υποδοχέα.

2.5 Εικονική Σάρωση

Στην συνέχεια με σκοπό την ανακάλυψη δυνητικά βιοδραστικών ενώσεων με ανταγωνιστική δράση έναντι του GHRH-R, έγινε εφαρμογή της υπολογιστικής τεχνικής της εικονικής σάρωσης. Χρησιμοποιήθηκαν δύο βιβλιοθήκες αποτελούμενες συνολικά από 14.511 βιοδραστικές ενώσεις (Bioactive Compound Library 1 & Bioactive Compound Library 2). Οι ενώσεις αντλήθηκαν ηλεκτρονικά σε δισδιάστατη μορφή από την Selleckchem και ακολούθησε η προ-επεξεργασία τους ώστε να προετοιμαστούν για την μοριακή πρόσδεση που θα ακολουθήσει στην συνέχεια.

Αρχικά η προ-επεξεργασία και των δύο βιβλιοθηκών έγινε στο DataWarrior όπου οι ενώσεις φιλτραρίστηκαν για συγκεκριμένους μοριακούς περιγραφείς σύμφωνα με τον κανόνα των πέντε του Lipinski. Με βάση τον παραπάνω κανόνα καθορίζεται πως οι ενώσεις δεν θα έχουν περισσότερους από πέντε δότες δεσμών υδρογόνου, δεν θα έχουν περισσότερους από δέκα δέκτες δεσμών υδρογόνου, το μοριακό βάρος δεν θα ξεπερνάει τα 500 Dalton και ο συντελεστής κατανομής οκτανόλης/νερού (logP) δεν θα υπερβαίνει το 5. Η βιβλιοθήκη I έχει συνολικά 9202 ενώσεις. Κατά την εισαγωγή της στο DataWarrior, 8 ενώσεις αφαιρέθηκαν αυτομάτως καθώς δεν ήταν δυνατή η παραγωγή της δομής τους λόγω σφαλμάτων που υπήρχαν στα αρχεία που λήφθηκαν από την βάση Selleckchem οπότε μείνανε 9194 ενώσεις. Έπειτα αφαιρέθηκαν 1.682 ενώσεις διότι ξεπερνούσαν το μοριακό βάρος των 500 Dalton, 349 ενώσεις καθώς η τιμή logP υπέρβαινε το όριο, 162 ενώσεις επειδή είχαν πάνω από δέκα δέκτες δεσμού υδρογόνου και 74 ενώσεις που είχαν πάνω από πέντε δότες δεσμού υδρογόνου απομένοντας συνολικά 6927 ενώσεις. Τέλος αφαιρέθηκαν 21 ενώσεις που βρέθηκε ότι τα sdf αρχεία τους ήταν κενά και 6 ενώσεις που δεν κατάφεραν από το πρόγραμμα να παραχθούν σωστά οι τρισδιάστατες συντεταγμένες τους, καταλήγοντας τελικά με 6900 ενώσεις που θα χρησιμοποιηθούν για την διαδικασία της μοριακής πρόσδεσης. Η βιβλιοθήκη II έχει συνολικά 5309 ενώσεις. Για την βιβλιοθήκη II δεν υπήρχαν σφάλματα στα αρχεία. Μετά την εφαρμογή της παραμέτρου για την μοριακή μάζα, τους δέκτες και τους δότες δεσμού υδρογόνου δεν άλλαξε ο αριθμός των ενώσεων καθώς όλες οι ενώσεις έχουν τιμές στα όρια που ορίζει ο κανόνας των πέντε του Lipinski. Μετά την εφαρμογή της παραμέτρου logP μείνανε 5243 ενώσεις για να χρησιμοποιηθούν στην την διαδικασία της μοριακής πρόσδεσης.

Οι διαδικασίες που ακολουθούν παρακάτω και αφορούν το κεφάλαιο της εικονικής σάρωσης έγιναν στον διακομιστή του εργαστηρίου Υπολογιστικής Χημείας του Τμήματος Μηχανικών Βιοϊατρικής με επεξεργαστή 64 πυρήνων, μνήμη RAM 256GB και 2 κάρτες γραφικών. Μετά την εφαρμογή του κανόνα του Lipinski ακολούθησε περαιτέρω επεξεργασία στην οποία η κάθε μια βιβλιοθήκη ξεχωριστά υποβλήθηκε σε αφαίρεση αλάτων μέσω της εντολής “-r” . Έπειτα έγινε παραγωγή των τρισδιάστατων συντεταγμένων κάθε ένωσης με την εντολή --gen 3d, οπότε και παράχθηκαν δύο βιβλιοθήκες σε μορφή 3D SDF. Στο συγκεκριμένο βήμα, για κάποιες από τις ενώσεις της βιβλιοθήκης I το Open Babel δεν κατάφερε να προβλέψει την στερεοϊσομέρεια τους με αποτέλεσμα ο ‘z’ άξονας να είναι κενός. Το πρόβλημα επιλύθηκε με την χρήση

του RDKit, ένα εργαλείο χημειοπληροφορικής ανοιχτού κώδικα που χρησιμοποιείται για την επεξεργασία χημικών ενώσεων [68]. Συγκεκριμένα έγινε χρήση της βιβλιοθήκης Chem.SDMolSupplier ώστε να διαβαστούν ομάδες μορίων, δηλαδή όλες οι ενώσεις της βιβλιοθήκης I και να μετατραπούν από μορφή “sdf” σε μορφή “mol”. Έπειτα χρησιμοποιήθηκαν οι συναρτήσεις GetConformer και GetPositions ώστε να καθοριστεί, ποια από τα αρχεία δεν είχαν συντεταγμένες για τον ‘z’ άξονα. Μετά από αυτή την επεξεργασία η βιβλιοθήκη I έχει συνολικά 6.642 ενώσεις. Επόμενο βήμα ήταν η ρύθμιση του pH στο 7 ώστε να προσομοιάζονται οι συνθήκες στον ανθρώπινο οργανισμό. Στην συνέχεια και οι δύο βιβλιοθήκες επεξεργάστηκαν έτσι ώστε να παραχθούν πολλαπλά αρχεία μέσω της εντολής “-m”. Τέλος, κατά την μοριακή πρόσδεση η μορφή των αρχείων που απαιτείται είναι “pdbqt”, συνεπώς όλα τα αρχεία που παράχθηκαν μετατράπηκαν από “sdf” σε “pdbqt”. Αφού προετοιμάστηκαν και οι δύο βιβλιοθήκες και συγχωνεύτηκαν σε μια βιβλιοθήκη, επόμενο βήμα ήταν η πραγματοποίηση του υπολογισμού της μοριακής πρόσδεσης.

Η μοριακή πρόσδεση έγινε και εδώ στην ορθοστερική θέση της διαμεμβρανικής περιοχής του υποδοχέα. Για το grid box χρησιμοποιήθηκαν οι διαστάσεις 20×20×25 Angstrom και το spacing τέθηκε ίσο με 1. Κατά την διάρκεια των υπολογισμών ο υποδοχέας ήταν σταθερός και οι ενώσεις ελεύθερες να πάρουν διάφορες διαμορφώσεις στο ενεργό κέντρο. Επιπλέον επειδή οι ενώσεις ήταν αρκετές, για μεγαλύτερη ταχύτητα των υπολογισμών αξιοποιήθηκαν 8 πυρήνες “cpu=8” ενώ παράχθηκε μία πόζα “num_modes=1” για κάθε ένωση καθώς έτσι συνηθίζεται σε διαδικασίες εικονικής σάρωσης. Για τις υπόλοιπες παραμέτρους (exhaustiveness, max_evals, min_rmsd, energy_range, verbosity) χρησιμοποιήθηκαν οι προκαθορισμένες τιμές.

Με το τέλος της μοριακής πρόσδεσης 24 ενώσεις από την βιβλιοθήκη I δεν κατάφεραν να υπολογιστούν με αποτέλεσμα να έχουμε τελικά 11.861 pdbqt αρχεία αντί για 11.885. Για κάθε μια από αυτές τις ενώσεις προέκυψε ένα αρχείο κειμένου που περιείχε το αποτελέσματα της μοριακής πρόσδεσης. Τα σκορ της πρώτης πόζας για κάθε μια από τις ενώσεις ταξινομήθηκαν ώστε να βρεθούν οι εκατό ενώσεις με το καλύτερο σκορ και να αξιολογηθούν στην συνέχεια.

2.5.1 Έλεγχος ομοιότητας

Στην παρούσα εργασία η εικονική σάρωση γίνεται με σκοπό την ανακάλυψη δυναμικά βιοδραστικών ενώσεων που στοχεύουν στην αναστολή της δράσης του υποδοχέα GHRH-R. Για τον σκοπό αυτό απαιτείται να απομονωθούν οι ενώσεις που έχουν την λιγότερη ομοιότητα μεταξύ τους ώστε να επιτευχθεί χημική ποικιλομορφία. Για τον έλεγχο ομοιότητας και την αξιολόγηση των ενώσεων θα χρησιμοποιηθεί ο συντελεστής Tanimoto. Τα sdf αρχεία (που είχαν προετοιμαστεί κατά το βήμα της μετατροπής σε pH=7) των εκατό καλύτερων ενώσεων χρησιμοποιήθηκαν ώστε να εφαρμοστεί ο έλεγχος ομοιότητας Tanimoto και να επιλεγθούν οι είκοσι εκείνες ενώσεις με το χαμηλότερο ποσοστό ομοιότητας.

Χρησιμοποιήθηκε η βιβλιοθήκη Chem.SDMolSupplier ώστε να γίνει η μετατροπή των sdf αρχείων σε mol αρχεία. Στην συνέχεια για τα mol αρχεία έγινε εξαγωγή των μοριακών αποτυπωμάτων MACCS μέσω της συνάρτησης GenMACCSKeys της βιβλιοθήκης MACCSkeys. Όπως αναλύθηκε και στο θεωρητικό υπόβαθρο τα μοριακά αποτυπώματα MACCS είναι βασισμένα σε δομικά κλειδιά υποομάδων και

αποτελούνται από 166 δομικά κλειδιά. Τέλος με την χρήση της συνάρτησης TanimotoSimilarity από την βιβλιοθήκη DataStructs υπολογίστηκαν οι ομοιότητες μεταξύ όλων των συνδυασμών των ζευγών των ενώσεων. Για να εξασφαλιστεί ότι όλες οι επιλεγόμενες ενώσεις θα διαφέρουν μεταξύ τους όσο γίνεται περισσότερο ακολουθήθηκε η εξής διαδικασία: υπολογίστηκε ο μέσος όρος της ομοιότητας κάθε ένωσης με τις υπόλοιπες και η τιμή που εξάχθηκε, χρησιμοποιήθηκε για την ταξινόμηση των ενώσεων και την επιλογή εκείνων με τον χαμηλότερο μέσο όρο. Η ίδια διαδικασία ακολουθήθηκε και για τα μοριακά αποτυπώματα RDKit που αντλήθηκαν μέσω της συνάρτησης GetRDKitFPGenerator της βιβλιοθήκης AllChem. Τα μοριακά αποτυπώματα RDKit ανήκουν στην κατηγορία των τοπολογικών αποτυπωμάτων και βασίζονται στα αποτυπώματα Daylight. Τα αποτελέσματα των δύο μοριακών αποτυπωμάτων συγκρίθηκαν μεταξύ τους και επιλέχθηκε εκείνο με τα χαμηλότερα ποσοστά ομοιότητας μεταξύ των ενώσεων.

Οι είκοσι ενώσεις αξιολογήθηκαν αρχικά με βάση τις αλληλεπιδράσεις τους με τον υποδοχέα και την ύπαρξη ή όχι των σημαντικών αμινοξέων που παρατηρούνται κατά την πρόσδεση του πεπτιδίου GHRH στον GHRH-R. Σε όσες ενώσεις η ανάλυση των αλληλεπιδράσεων δεν ήταν αρκετή για να δώσει ξεκάθαρο αποτέλεσμα, η επιλογή έγινε με βάση το σκορ που προέκυψε από την μοριακή πρόσδεση. Ως αποτέλεσμα έμειναν τελικά δέκα ενώσεις με πιθανή ανταγωνιστική δράση έναντι του GHRH-R για μετέπειτα επεξεργασία.

3. Αποτελέσματα

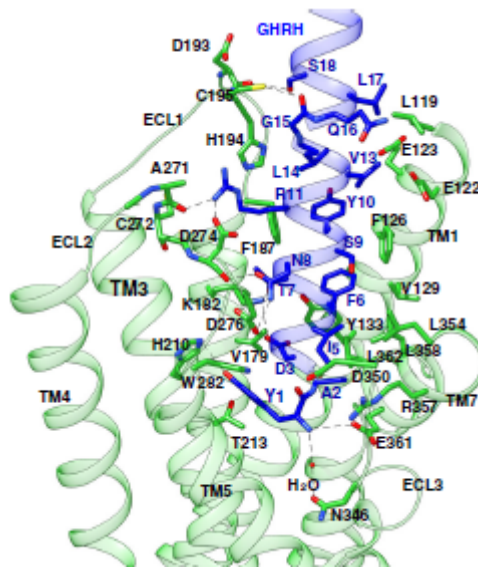
3.1 Δομική Ανάλυση GHRH-R

Η ενεργοποίηση του υποδοχέα εξαρτάται σε πολύ μεγάλο βαθμό από συγκεκριμένες αλληλεπιδράσεις που κάνει με το ενδογενές του πεπτιδίο. Είναι λοιπόν πολύ σημαντικό για την συνέχεια της εργασίας να γίνει ανάλυση των αλληλεπιδράσεων μεταξύ του πεπτιδίου GHRH και της πρωτεΐνης GHRH-R με σκοπό να εντοπιστούν εκείνα τα αμινοξέα του υποδοχέα που ισχυροποιούν την δέσμευση του πεπτιδίου. Μέχρι σήμερα έχει δημοσιευθεί η κρυσταλλική δομή της διαμεμβρανικής περιοχής του GHRH-R σε σύζευξη με το πεπτιδίο GHRH και την Gs πρωτεΐνη ενώ για την εξωκυττάρια περιοχή του υποδοχέα έχει δημοσιευθεί η κρυσταλλική δομή ξεχωριστά από την υπόλοιπη πρωτεΐνη. Η μελέτη των αλληλεπιδράσεων έγινε με την χρήση της διαμεμβρανικής περιοχής του GHRH-R χρησιμοποιώντας την κρυσταλλική δομή που έχει επιτευχθεί μέσω της μεθόδου κρύο-ηλεκτρονικής μικροσκοπίας και αντλήθηκε ηλεκτρονικά από την Protein Data Bank.

Μετά από δομική ανάλυση που έγινε παρατηρήθηκε ότι η αμινοτελική περιοχή του πεπτιδίου που εισέρχεται στο διαμεμβρανικό κομμάτι του υποδοχέα κάνει πλήθος πολικών και υδρόφοβων αλληλεπιδράσεων (εικόνα 3.1). Η Tyr^{1P} φαίνεται πως εμπλέκεται σε αρκετές αλληλεπιδράσεις και έχει σημαντικό ρόλο στην διαμόρφωση του υποδοχέα. Αρχικά η υδροξυλομάδα της πλευρικής αλυσίδας της Tyr^{1P} κάνει έναν δεσμό υδρογόνου με την HIS210^{3.37b} και Van Der Waals αλληλεπιδράσεις με τα αμινοξέα THR213^{3.40b} και TRP282^{5.36b}. Επιπλέον η αμινομάδα της κύριας αλυσίδας της κάνει δεσμό υδρογόνου με την ARG357^{7.38b} και παράλληλα συνδέεται με ένα μόριο νερού το οποίο είναι συνδεδεμένο με το αμινοξύ ASN346^{6.57b}, που βρίσκεται στο σημείο σύνδεσης της διαμεμβρανικής έλικας 6 (TM6) με τον εξωκυττάριο βρόγχο 3 (ECL3) και πιθανώς να σταθεροποιεί τον δεύτερο στην ενεργή του κατάσταση. Η Ala^{2P} μαζί με την Ile^{5P} εμπλέκονται σε ένα δίκτυο υδρόφοβων αλληλεπιδράσεων με τα αμινοξέα LEU362^{7.43b}, LEU358^{7.39b}, LEU354^{7.35b} και ASP350^{ECL3} ενώ η Ala^{4P} φαίνεται πως δεν συμμετέχει στην πρόσδεση του GHRH. Αντιθέτως το Asp^{3P} σχηματίζει πολύ ισχυρές αλληλεπιδράσεις με τον υποδοχέα. Ειδικότερα κάνει μια γέφυρα άλατος με την LYS182^{2.76b} ενώ οι πολικές αλληλεπιδράσεις που κάνει με την TYR133^{1.43b} και το ASP183^{2.68b} ισχυροποιούν την πρόσδεση και θέτουν το Asp^{3P} και την LYS182^{2.76b} δύο από τα σημαντικότερα αμινοξέα στην δημιουργία του συμπλόκου GHRH-GHRHR. Επιπλέον αυτή η περιοχή σταθεροποιείται ακόμα περισσότερο μέσω των πολικών αλληλεπιδράσεων στις οποίες συμμετέχει η Thr^{7P}. Σε ένα δίκτυο υδρόφοβων αλληλεπιδράσεων εμπλέκεται η Phe^{6P} σχηματίζοντας υδρόφοβες αρωματικές-αρωματικές αλληλεπιδράσεις με την PHE126^{1.36b} και την TYR133^{1.43b} καθώς και αλειφατικές-αρωματικές αλληλεπιδράσεις με την LYS130^{1.40b}, την VAL129^{1.39b} και την LEU362^{7.43b}. Στην πρόσδεση και σταθεροποίηση του πεπτιδίου συνεισφέρουν ενεργά και οι περιοχές ECL1 και ECL2. Ειδικότερα το ASP274^{ECL2} και το ASP276^{ECL2} μέσω δεσμού υδρογόνου με το Asn^{8P}, το ASP193^{ECL1} επίσης με δεσμό υδρογόνου με την Ser^{18P}, και η PHE187^{ECL1} μέσω υδρόφοβης αρωματικής-αρωματικής αλληλεπίδρασης με την Tyr^{10P}. Επίσης η Tyr^{10P} σχηματίζει και ένα δίκτυο υδρόφοβων και πολικών αλληλεπιδράσεων με αμινοξέα της διαμεμβρανικής έλικας 1. Τέλος η Arg^{11P} σχηματίζει γέφυρα άλατος με το ASP274^{ECL2} και δεσμούς υδρογόνου με τις κύριες αλυσίδες των ALA271^{ECL2} και HIS194^{ECL1}. Τα υπόλοιπα αμινοξέα φαίνεται πως

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ
ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

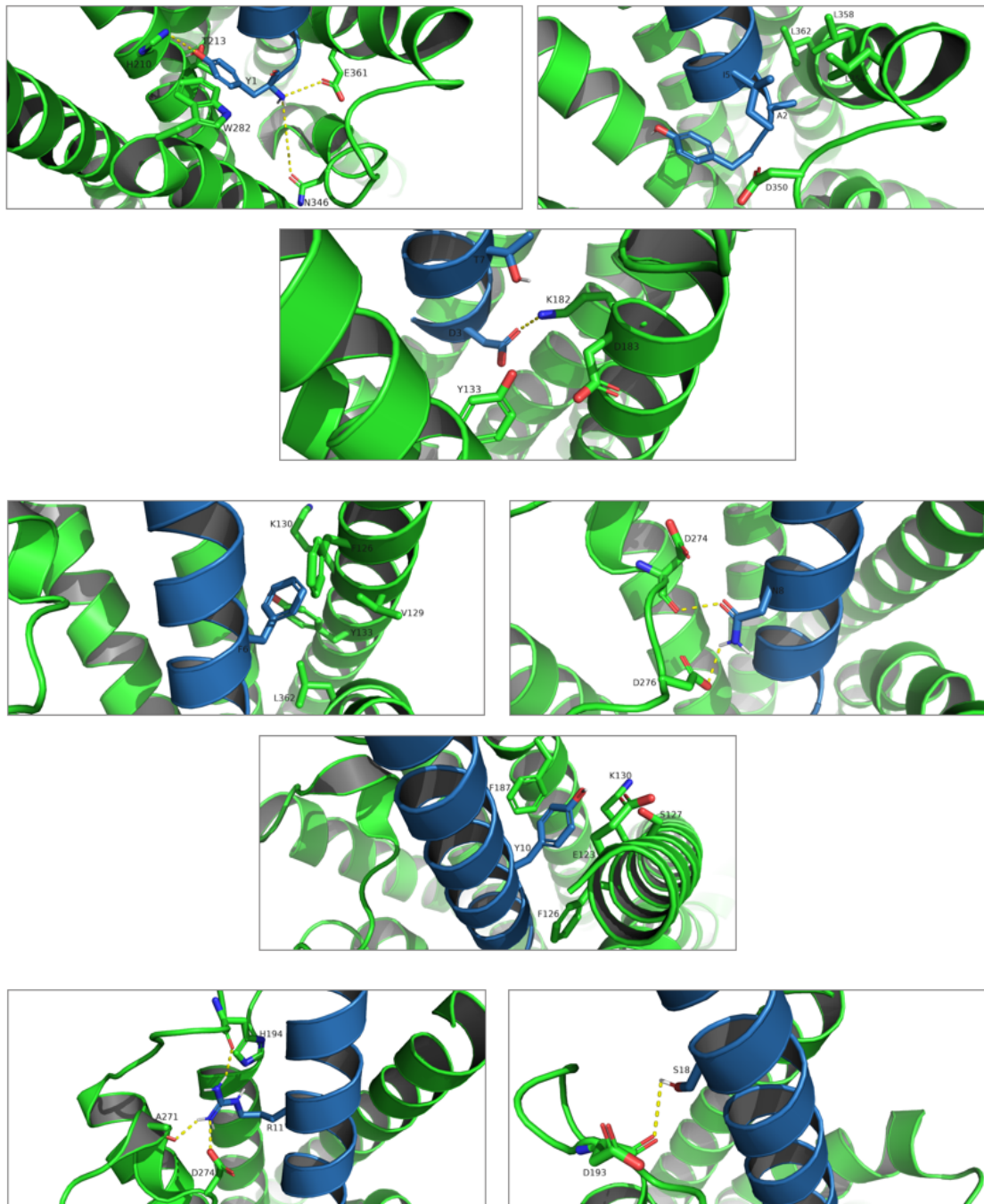
δεν συμμετέχουν ουσιαστικά στην πρόσδεση του πεπτιδίου ενώ από την ARG^{20P} και μετά ξεκινάνε οι αλληλεπιδράσεις με το ECD κομμάτι του υποδοχέα που δεν αποτελούν στόχο της παρούσας εργασίας [21].



Εικόνα 3.1: Σύνολο αλληλεπιδράσεων μεταξύ GHRH και διαμεμβρανικού μέρους του GHRH-R. Με πράσινο χρώμα φαίνεται ο υποδοχέας και με αναπαράσταση sticks τα αμινοξέα που συμμετέχουν στην πρόσδεση του πεπτιδίου. Με μπλε χρώμα φαίνεται το πεπτίδιο GHRH. Εικόνα από πηγή [21].

Οι αλληλεπιδράσεις του GHRH με τον GHRH-R που μελετήθηκαν παραπάνω, παρουσιάζονται αναλυτικά στην εικόνα 3.2 που ακολουθεί.

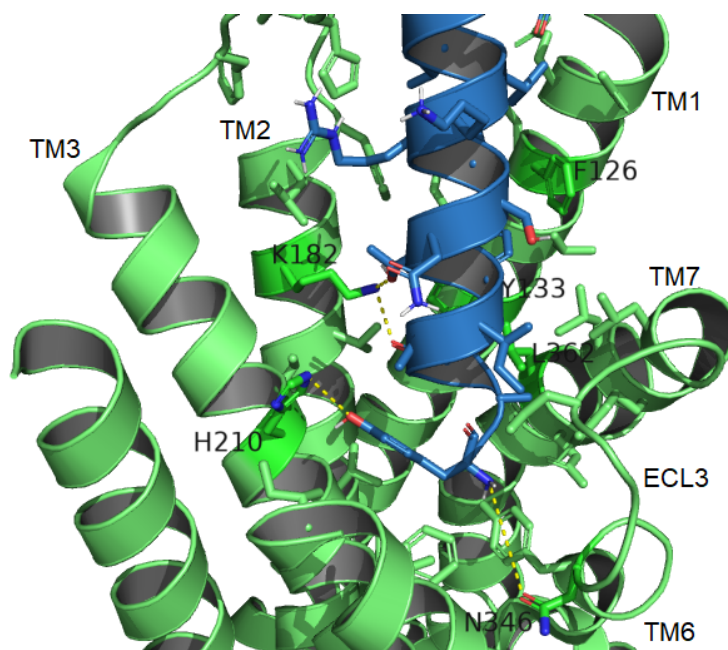
ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ



Εικόνα 3.2: Αλληλεπιδράσεις του πεπτιδίου GHRH (μπλε χρώμα) με τον υποδοχέα GHRH-R (πράσινο χρώμα). Από αριστερά προς τα δεξιά και από πάνω προς τα κάτω φαίνονται οι αλληλεπιδράσεις για τα εξής αμινοξέα του GHRH: Tyr^{1P}, Ala^{2P} & Ile^{5P}, Asp^{3P} & Thr^{7P}, Phe^{6P}, Asn^{8P}, Tyr^{10P}, Arg^{11P}, Ser^{18P}.

Με βάση την παραπάνω δομική ανάλυση παρατηρείται πως τα σημαντικά αμινοξέα του υποδοχέα που συμμετέχουν στην πρόσδεση του πεπτιδίου είναι η TYR133^{1.43b} που εμπλέκεται σε ένα δίκτυο πολικών και υδρόφοβων αλληλεπιδράσεων, η PHE126^{1.36b} και η LEU362^{7.43b} που κάνουν υδρόφοβες αλληλεπιδράσεις, η LYS182^{2.76b} που κάνει γέφυρα άλατος και δεσμούς υδρογόνου καθώς και η HIS210^{3.37b} και ASN346^{6.57b} που μέσω πολικών αλληλεπιδράσεων ισχυροποιούν την σύνδεση του αμινοτελικού άκρου του πεπτιδίου στον υποδοχέα. Επιπλέον το γεγονός ότι η ASN346^{6.57b} βρίσκεται στο σημείο σύνδεσης της διαμεμβρανικής έλικας 6 και του εξωκυττάριου βρόγχου 3 και

πιθανώς σταθεροποιεί τον βρόγχο στην ενεργή του κατάσταση, καθιστά το συγκεκριμένο αμινοξύ σημαντικό στόχο (εικόνα 3.3).



Εικόνα 3.3: Τα αμινοξέα που ξεχώρισαν για την συμμετοχή τους στην ισχυροποίηση της δέσμευσης του πεπτιδίου. Με πράσινο χρώμα φαίνεται ο GHRH-R ενώ με μπλε χρώμα η αμινοτελική περιοχή του GHRH.

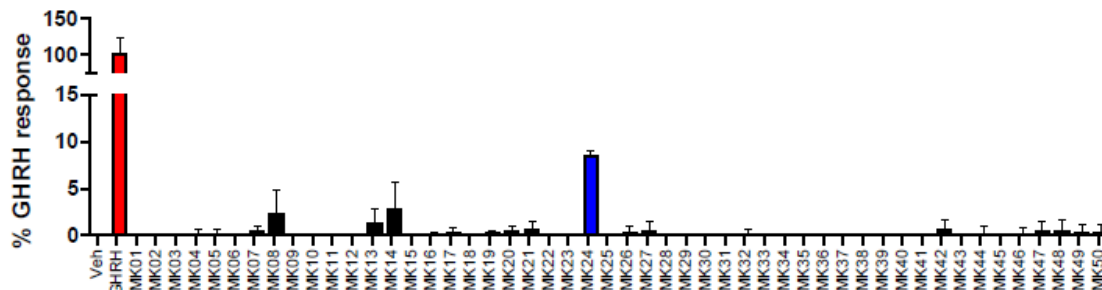
3.2 Υπολογιστική αξιολόγηση βιοδραστικών ενώσεων της σειράς MK

3.2.1 Αποτελέσματα Μοριακής Πρόσδεσης

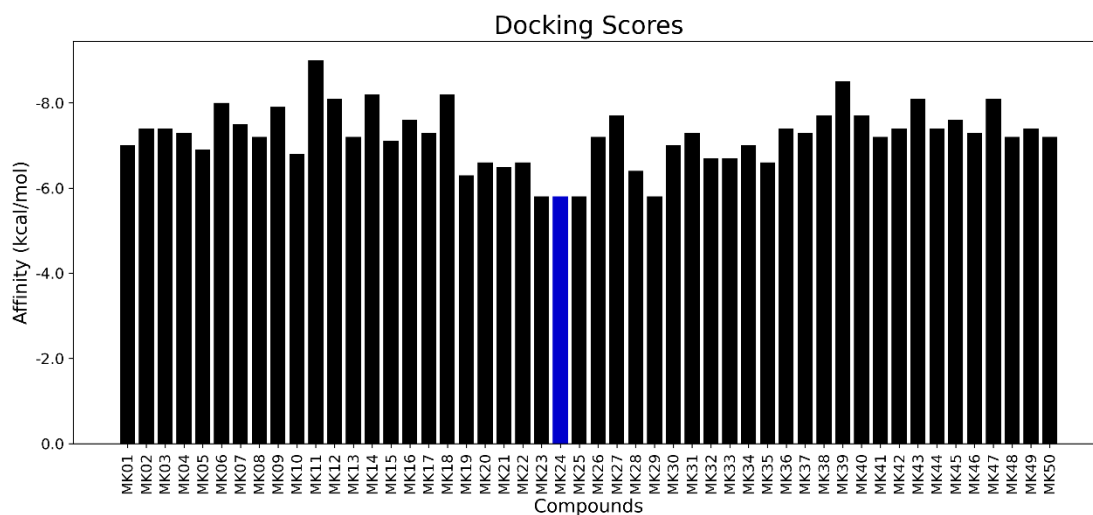
Όπως έχει αναφερθεί και σε προηγούμενο κεφάλαιο η παρούσα εργασία αποτελεί επέκταση μελέτης που έχει προηγηθεί μέσω της συνεργασίας του κέντρου Νευροενδοκρινολογίας του πανεπιστημίου της Πρετόρια με τον Δρ. Μίνω Ματσούκα και οδήγησε στον εντοπισμό μιας σειράς βιοδραστικών ενώσεων που δοκιμάστηκαν πειραματικά. Τα αποτελέσματα που ακολουθούν αφορούν την εφαρμογή της μοριακής πρόσδεσης για την δοκιμή αυτών των ενώσεων και υπολογιστικά. Η μοριακή πρόσδεση έλαβε μέρος στην ορθοστερική θέση της διαμεμβρανικής περιοχής του GHRH-R, τόσο για την ενεργή όσο και για την ανενεργή του κατάσταση. Για του υπολογισμούς χρησιμοποιήθηκε το λογισμικό AutoDock Vina.

Στην συνέχεια παρουσιάζονται τα αποτελέσματα για την μοριακή πρόσδεση που έγινε στην ενεργή κατάσταση του υποδοχέα με σκοπό να ελεγχθεί η αγωνιστική δράση των ενώσεων. Στο διάγραμμα της εικόνας 3.4 φαίνονται τα πειραματικά αποτελέσματα από την μελέτη που έχει προηγηθεί. Στον οριζόντιο άξονα είναι οι 50 ενώσεις MK(01-50) ενώ στον κάθετο άξονα είναι το παραγόμενο cAMP. Στο διάγραμμα της εικόνας 3.5 παρουσιάζονται τα υπολογιστικά αποτελέσματα της μοριακής πρόσδεσης όπου στον οριζόντιο άξονα είναι και πάλι οι 50 ενώσεις ενώ στον κάθετο άξονα είναι το σκορ πρόσδεσης με μονάδες μέτρησης τα kcal/mol. Η ένωση που ξεχώρισε από μοριακή πρόσδεση είναι η MK11 με σκορ πρόσδεσης “Affinity= -9 kcal/mol”. Αντιθέτως η

ένωση MK24 που πειραματικά παρουσίασε καλή αγωνιστική δράση και οδήγησε στην παραγωγή cAMP, δεν είχε την ίδια δράση και υπολογιστικά καθώς όπως φαίνεται και από το διάγραμμα της εικόνας 3.5 είχε ένα από τα χαμηλότερα σκορ πρόσδεσης (Affinity= -5.8 kcal/mol).



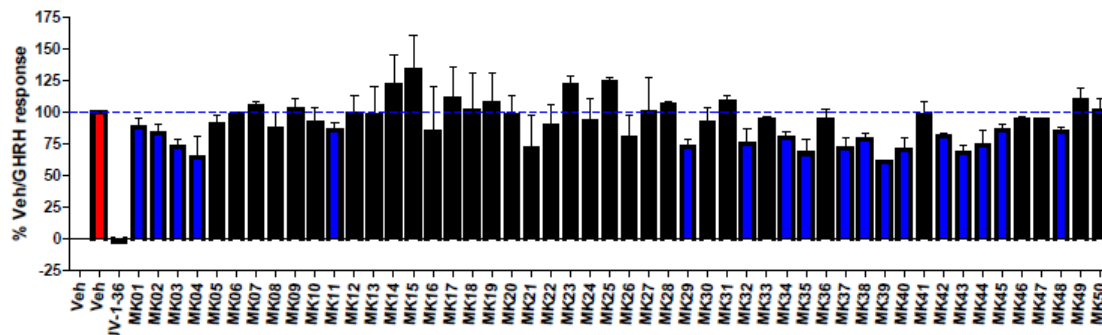
Εικόνα 3.4: Πειραματικά αποτελέσματα από δουλειά που έχει προηγηθεί για την μελέτη της αγωνιστικής δράσεις των ενώσεων MK(01-50). Στο διάγραμμα με μπλε χρώμα ξεχωρίζει για την αγωνιστική της δράση η ένωση MK24.



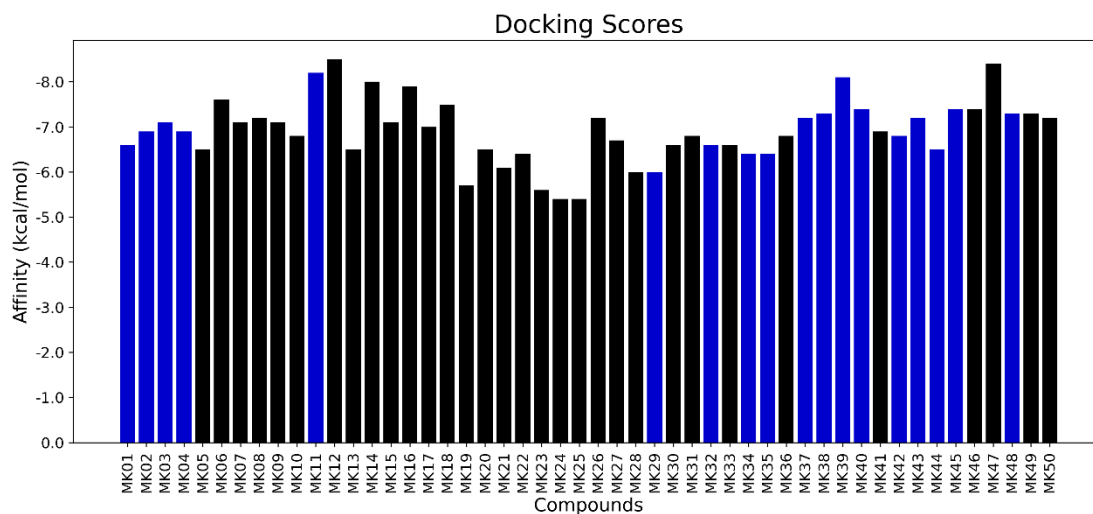
Εικόνα 3.5: Αποτελέσματα μοριακής πρόσδεσης για την ενεργή κατάσταση του υποδοχέα. Με μπλε χρώμα φαίνεται η ένωση MK24 που πειραματικά διαπιστώθηκε ότι μπορεί να προκαλέσει διέγερση του υποδοχέα.

Ακολουθως παρουσιάζονται τα αποτελέσματα για την μοριακή πρόσδεση που έγινε στην ανενεργή κατάσταση του υποδοχέα με σκοπό να ελεγχθεί η ανταγωνιστική δράση των ενώσεων. Στο διάγραμμα της εικόνας 3.6 με τα πειραματικά αποτελέσματα από την μελέτη που έχει προηγηθεί, φαίνονται με μπλε χρώμα οι 18 εκ των 50 ενώσεων που παρουσίασαν ανταγωνιστική δράση λόγω της ικανότητάς τους να μειώνουν το παραγόμενο cAMP. Συγκρίνοντας αυτές τις ενώσεις με τα υπολογιστικά αποτελέσματα που φαίνονται στο διάγραμμα της εικόνας 3.7, παρατηρείται πως τα πειραματικά αποτελέσματα επαληθεύονται για τις ενώσεις MK02, MK03, MK04, MK11, MK37, MK38, MK39, MK40, MK43, MK45 και MK48 των οποίων το σκορ πρόσδεσης ξεπερνούσε ένα κατώφλι που ορίστηκε στα -7 kcal/mol. Το συγκεκριμένο κατώφλι ορίστηκε με βάση τον μέσο των σκορ πρόσδεσης όλων των ενώσεων που έλαβαν μέρος στην μοριακή πρόσδεση που έγινε στην ανενεργή κατάσταση του υποδοχέα. Η ένωση που σημείωσε το καλύτερο σκορ πρόσδεσης είναι η MK12 με “Affinity= -8.5 kcal/mol”.

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ



Εικόνα 3.6: Πειραματικά αποτελέσματα από δουλειά που έχει προηγηθεί για την μελέτη της ανταγωνιστικής δράσης των ενώσεων MK(01-50). Στο διάγραμμα ξεχωρίζουν ως ανταγωνιστές έναντι του ενδογενούς πεπτιδίου GHRH οι ενώσεις που είναι επισημασμένες με μπλε χρώμα. Μεταξύ αυτών βρίσκεται και η ένωση MK04.



Εικόνα 3.7: Αποτελέσματα μοριακή πρόσδεσης για την ανενεργή κατάσταση του υποδοχέα. Με μπλε χρώμα φαίνονται οι ενώσεις που πειραματικά διαπιστώθηκε πως έχουν ανταγωνιστική δράση.

Επιπλέον παρατίθενται και μέρη από τους κώδικες που χρησιμοποιήθηκαν για την προετοιμασία των ενώσεων (εικόνα 3.8) και την εφαρμογή της μοριακής πρόσδεσης. (εικόνα 3.9).

```
for sdf2d in "$sdf2d_dir"/*.sdf; do
    #extract file name
    filename=$(basename $sdf2d .sdf)

    #convert from 2d sdf to 3d sdf
    obabel -isdf $sdf2d -osdf -O $sdf3d_dir/$filename.sdf --gen3D

    #convert from 3d sdf to 3d sdf -r
    obabel -isdf $sdf3d_dir/$filename.sdf -osdf -O $rsalt_dir/$filename.sdf -r

    #convert from 3d sdf -r to 3d sdf at ph=7
    obabel -isdf $rsalt_dir/$filename.sdf -osdf -O $ph_dir/$filename.sdf -p 7.0

    #convert from 3d sdf at ph=7 to pdbqt
    obabel -isdf $ph_dir/$filename.sdf -opdbqt -O $pdbqt_dir/$filename.pdbqt
```

Εικόνα 3.8: Διαδικασία μετατροπής των ενώσεων από 2d sdf σε pdbqt που είναι η απαιτούμενη μορφή αρχείου για την εφαρμογή της μοριακής πρόσδεσης. Έγινε χρήση της εντολής obabel.

```
mkdir $docking_directory/Results

for R in Receptor_GHRHr_inactive/*.pdbqt; do

    r=`basename $R .pdbqt` #extract receptor name
    mkdir $docking_directory/Results/$r

    for L in pdbqts/*.pdbqt;do
        l=`basename $L .pdbqt` #extract ligand name

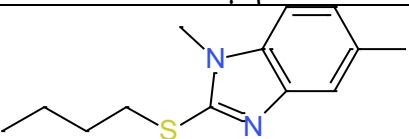
        #run docking
        /opt/vina --receptor Receptor_GHRHr_inactive/${r}.pdbqt
        --ligand pdbqts/${l}.pdbqt --config Config.txt
        --out $docking_directory/Results/${r}/${l}.pdbqt
        --log Log/${l}.log --cpu=4 & nwait NR_CPUS
    done
done
```

Εικόνα 3.9: Μέρος του κώδικα που χρησιμοποιήθηκε για την εφαρμογή της μοριακής πρόσδεσης των 59 ενώσεων στο ορθοστερικό κέντρο της ανενεργής/ενεργής κατάστασης του GHRH-R με την χρήση της εντολής /opt/vina.

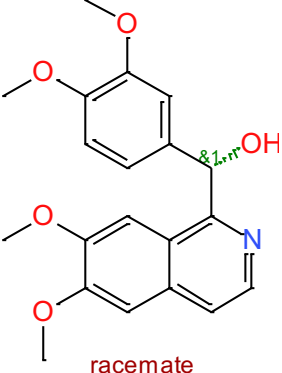
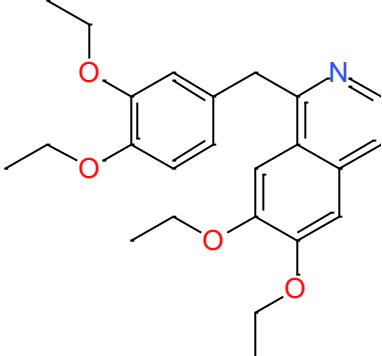
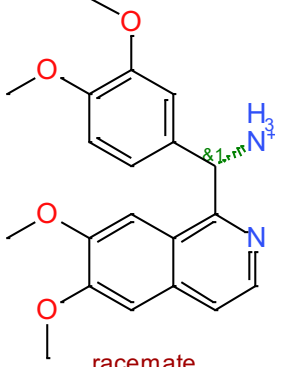
3.2.2 Αλληλεπιδράσεις προσδετών με τον GHRH-R

Η ανάλυση του σκορ πρόσδεσης αποτελεί μια ένδειξη της δέσμευσης των ενώσεων στον υποδοχέα όμως για την μεγαλύτερη αξιοπιστία των αποτελεσμάτων χρειάζεται επίσης η οπτική παρατήρηση και η μελέτη των αλληλεπιδράσεων. Οι ενώσεις που ξεχώρισαν πειραματικά αναλύονται περαιτέρω σε ατομικό επίπεδο μέσω της μελέτης των αλληλεπιδράσεων που κάνουν με τον υποδοχέα προκειμένου να γίνει κατανοητός ο μηχανισμός πρόσδεσής τους και ο λόγος για τον οποίο η κάθε ένωση είχε καλά in vitro αποτελέσματα. Για την ένωση MK24 που παρουσίασε αγωνιστική δράση, αναλύεται η αλληλεπίδραση της με την ενεργή κατάσταση του υποδοχέα ενώ για την ένωση MK04 και τις παρεμφερείς με αυτήν MK04-6 και MK04-9 που παρουσίασαν ανταγωνιστική δράση μελετάται η αλληλεπίδραση τους με την ανενεργή κατάσταση του υποδοχέα. Προκειμένου να διευκολυνθεί η μελέτη των αλληλεπιδράσεων, εμφανίστηκαν στο PyMOL οι πλευρικές αλυσίδες των αμινοξέων του υποδοχέα που απέχουν απόσταση πέντε Angstrom από την κάθε ένωση. Στον πίνακα 3.1 που ακολουθεί φαίνεται η δισδιάστατη δομή των ενώσεων.

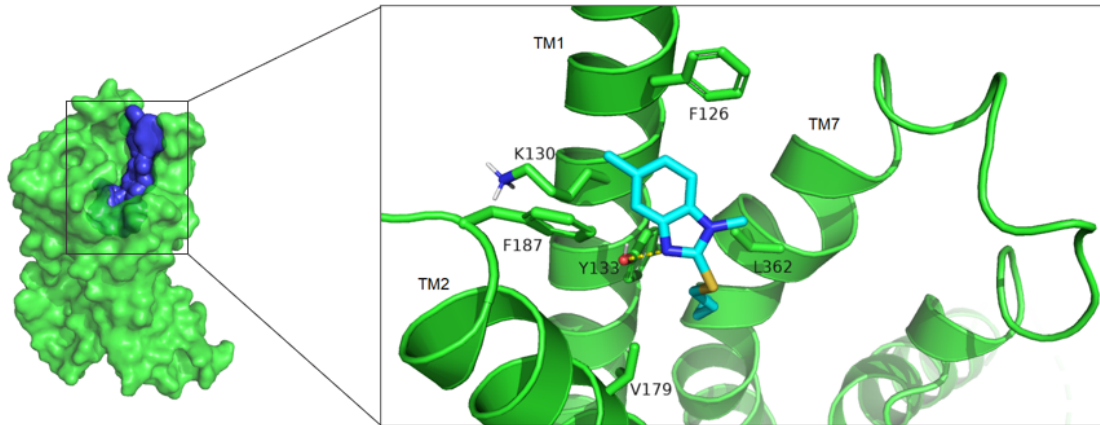
Πίνακας 3.1: Δισδιάστατη δομή των ενώσεων MK24, MK04, MK04-6 και MK04-9

Ένωση	2D Δομή
MK24	

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

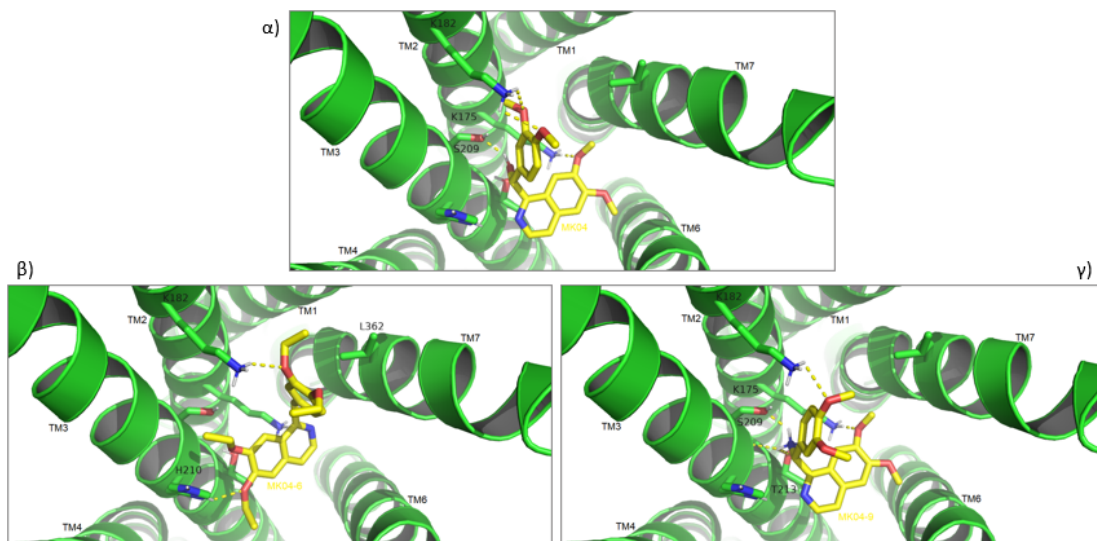
MK04	
MK04-6	
MK04-9	

Οι αλληλεπιδράσεις του υποδοχέα GHRH-R που βρίσκεται στην ενεργή κατάσταση με την ένωση MK24 παρουσιάζονται στην εικόνα 3.10. Η ένωση MK24 προσδένεται στον υποδοχέα μέσω έξι υδρόφοβων αλληλεπιδράσεων. Ειδικότερα δύο υδρόφοβες αρωματικές αλληλεπιδράσεις σχηματίζονται μεταξύ του βενζολίου της ένωσης και των αρωματικών δακτυλίων της PHE187^{ECL1} και της PHE126^{1.36b}, ενώ το ίδιο βενζόλιο κάνει μια υδρόφοβη αρωματική-αλειφατική αλληλεπίδραση με την αλειφατική περιοχή της πλευρικής αλυσίδας της LYS130^{1.40b}. Με την σειρά της η βουτυλομάδα της ένωσης σχηματίζει μια υδρόφοβη αρωματική-αλειφατική αλληλεπίδραση με τον αρωματικό δακτύλιο της TYR133^{1.43b} και μια υδρόφοβη αλειφατική-αλειφατική αλληλεπίδραση με την πλευρική αλυσίδα της LEU362^{7.43b} και της VAL179^{2.64b}. Τέλος το άζωτο του δακτυλίου της ένωσης κάνει δεσμό υδρογόνου με την υδροξυλομάδα της πλευρικής αλυσίδας της TYR133^{1.43b}.



Εικόνα 3.10: Αλληλεπίδραση της ένωσης MK24 με την ενεργή κατάσταση του GHRH-R. Στα αριστερά με μπλε χρώμα φαίνεται η περιοχή στην οποία πραγματοποιήθηκε η μοριακή πρόσδεση. Στην μεγεθυμένη εικόνα με πράσινο χρώμα αναπαρίσταται ο GHRH-R ενώ με κυανό χρώμα η ένωση MK24. Με διακεκομμένη γραμμή φαίνεται ο δεσμός υδρογόνου.

Παρακάτω φαίνονται οι αλληλεπιδράσεις του υποδοχέα GHRH-R που βρίσκεται στην ανενεργή κατάσταση, με τις ενώσεις MK04, MK04-6 και MK04-9 (εικόνα 3.11). Η ένωση MK04 μέσω των μεθόξυ οξυγόνων σχηματίζει δεσμούς υδρογόνου με την LYS182^{2.67b} και την LYS175^{2.60b} ενώ η υδροξυλομάδα της ένωσης κάνει δεσμό υδρογόνου με την SER209^{3.36b} (εικόνα 3.11α). Η ένωση MK04-6 προσδέεται στον υποδοχέα μέσω δύο δεσμών υδρογόνου και μιας υδρόφοβης αλληλεπίδρασης. Ειδικότερα μέσω των αιθόξυ οξυγόνων σχηματίζει δεσμούς υδρογόνου με την LYS182^{2.67b} και την HIS210^{3.37b} ενώ η υδρόφοβη αλειφατική-αλειφατική αλληλεπίδραση σχηματίζεται μεταξύ της αιθυλομάδας της ένωσης και της πλευρικής αλυσίδας της LEU362^{7.43b} (εικόνα 3.11β). Τέλος η ένωση MK04-9 συνδέεται στον υποδοχέα μέσω πέντε δεσμών υδρογόνου. Στους δύο από αυτούς συμμετέχουν τα μεθόξυ οξυγόνα της ένωσης με τα αμινοξέα LYS182^{2.67b} και LYS175^{2.60b}. Στους άλλους τρεις συμμετέχει η αμινομάδα της ένωσης η οποία σχηματίζει δεσμούς υδρογόνου με την πλευρική και την κύρια αλυσίδα της SER209^{3.36b} καθώς και με την πλευρική αλυσίδα της THR213^{3.40} (εικόνα 3.11γ).



Εικόνα 3.11: Αλληλεπίδραση των ενώσεων (κίτρινα sticks) με τον υποδοχέα (πράσινο cartoon). Οι πλευρικές αλυσίδες των σημαντικών αμινοξέων του υποδοχέα που συμμετέχουν στην πρόσδεση των ενώσεων φαίνονται με πράσινα sticks. Με διακεκομμένες γραμμές φαίνονται οι δεσμοί υδρογόνου. Στην εικόνα (3.11α) παρουσιάζεται η ένωση MK04 στην (3.11β) η MK04-6 και στην (3.11γ) η MK04-9.

3.4 Αποτελέσματα Εικονικής Σάρωσης

Μετά την εφαρμογή της μοριακής πρόσδεσης επιλέχθηκαν από τις δύο βιβλιοθήκες, με βάση το σκορ πρόσδεσης, οι εκατό καλύτερες ενώσεις. Οι δομές των ενώσεων αυτών βρίσκονται στο παράρτημα στον πίνακα 6.1. Σε αυτές εφαρμόστηκε έλεγχος ομοιότητας με την χρήση του συντελεστή Tanimoto για δύο διαφορετικά μοριακά αποτυπώματα (MACCS και RDKit). Οι εκατό ενώσεις συγκρίθηκαν μεταξύ τους και ξεχώρισαν οι είκοσι ενώσεις με το χαμηλότερο ποσοστό δομικής ομοιότητας και κωδικοποιήθηκαν με τα μοριακά αποτυπώματα RDKit. Οι ενώσεις αυτές είναι επισημασμένες με έντονο χρώμα στον πίνακα 6.1 του παραρτήματος. Τα αποτελέσματα εφαρμογής του συντελεστή Tanimoto με την χρήση της κωδικοποίησης των μοριακών αποτυπωμάτων RDKit για τις εκατό ενώσεις παρουσιάζονται στον τριγωνικό χάρτη συντελεστών ομοιότητας στην εικόνα 3.14. Το εύρος των τιμών των συντελεστών Tanimoto είναι από 0 έως 1. Με σκούρο μπλε χρώμα φαίνονται οι ενώσεις με υψηλό ποσοστό ομοιότητας ενώ με σκούρο κόκκινο χρώμα οι ενώσεις με χαμηλό ποσοστό ομοιότητας. Επιπρόσθετα στις εικόνες 3.12 και 3.13 παρατίθενται οι κώδικες που χρησιμοποιήθηκαν για τη εφαρμογή της μοριακής πρόσδεσης και του συντελεστή Tanimoto για αναζήτηση ομοιότητας αντίστοιχα.

```
mkdir $docking_directory/Results
```

```
for R in Receptor_GHRHr_inactive/*.pdbqt; do
```

```
    r=`basename $R .pdbqt` #extract receptor name  
    mkdir $docking_directory/Results/$r
```

```
    for L in pdbqts_L1_L2/*.pdbqt;do
```

```
        l=`basename $L .pdbqt` #extract ligand name
```

```
        #run docking
```

```
        /opt/vina --receptor Receptor_GHRHr_inactive/${r}.pdbqt  
--ligand pdbqts_L1_L2/${l}.pdbqt --config Config.txt  
--out $docking_directory/Results/${r}/${l}.pdbqt  
--log Log/${l}.log --cpu=4 --num_modes 1 & nrmwait NR_CPUS
```

Εικόνα 3.12: Μέρος του κώδικα που χρησιμοποιήθηκε για την εφαρμογή της μοριακής πρόσδεσης των 11.885 ενώσεων στο ορθοστερικό κέντρο της ανενεργής κατάστασης του GHRH-R με την χρήση της εντολής /opt/vina.

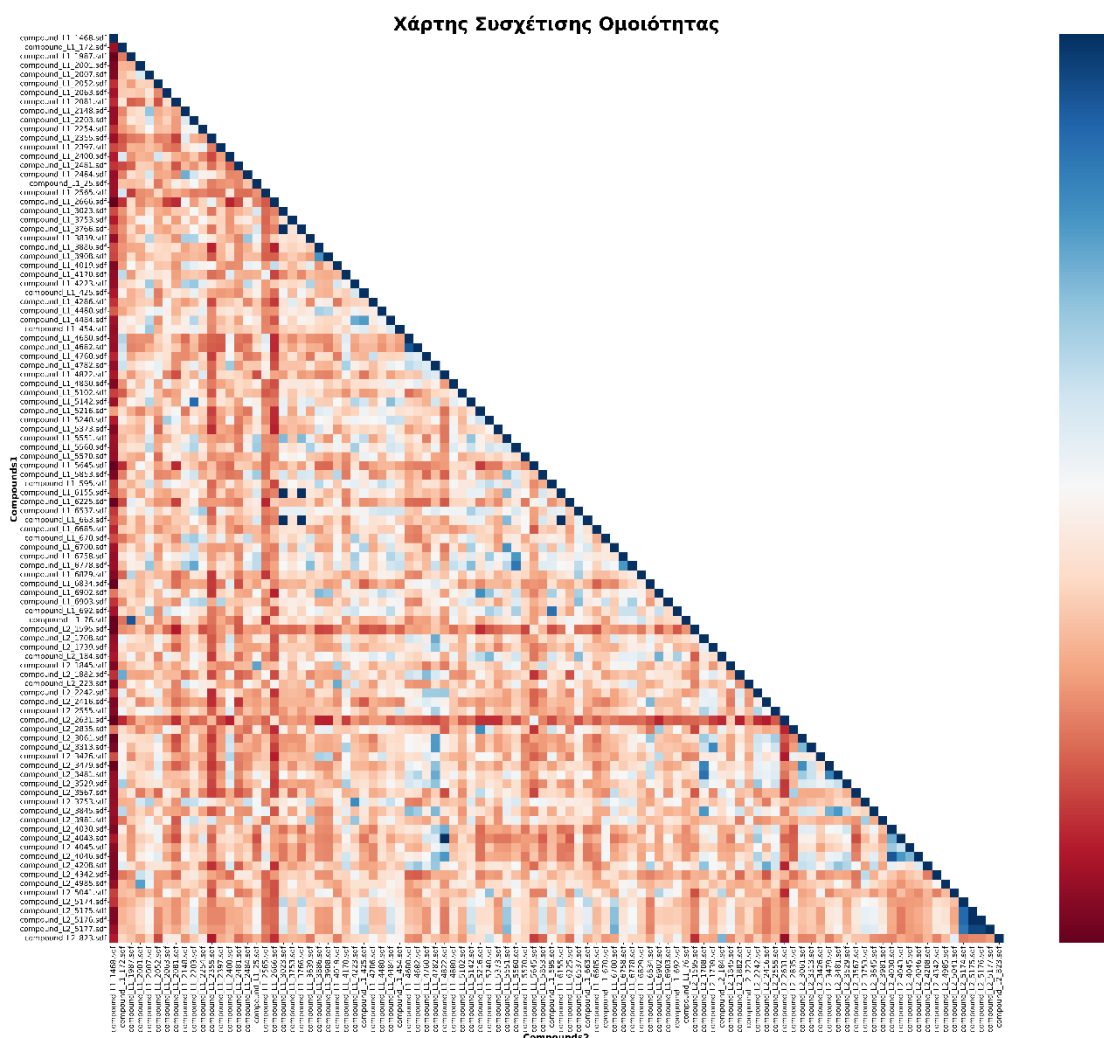
ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

```
# Calculate MACCS for each compound
#fps = [MACCSkeys.GenMACCSKeys(mol) for mol in mols]

# Calculate rdkit fingerprint for each compound
fpngen = AllChem.GetRDKitFPGenerator(maxPath=2, fpSize=1024)
fps = [fpngen.GetFingerprint(mol) for mol in mols]

# Calculate tanimoto similarity for each pair
num_mols = len(mols)
num_mols
similarity_data=[]
for i in range(num_mols):
    for j in range(i, num_mols):
        similarity = DataStructs.TanimotoSimilarity(fps[i], fps[j])
        similarity_data.append([file_names[i], file_names[j], similarity])
    if i != j:
        similarity_data.append([file_names[j], file_names[i], similarity])
```

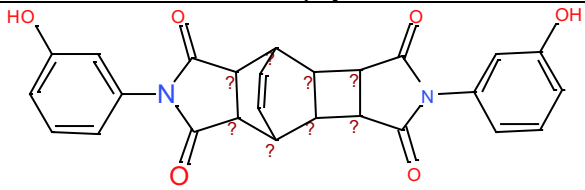
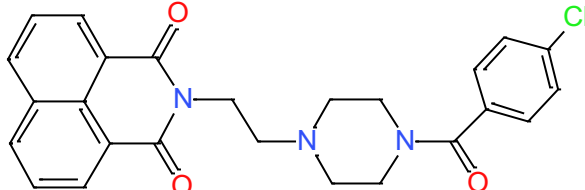
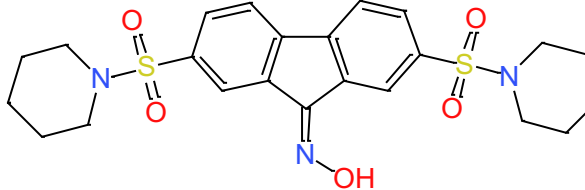
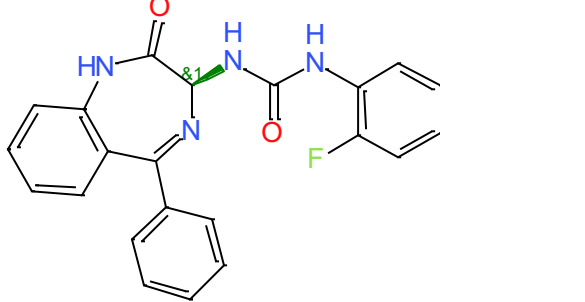
Εικόνα 3.13: Μέρος του κώδικα που χρησιμοποιήθηκε για την εξαγωγή των MACCS και RDKit μοριακών αποτυπωμάτων μέσω των συναρτήσεων GenMACCSKeys και GetRDKitFPGenerator αντίστοιχα. Υπολογισμός ομοιοτήτων μεταξύ όλων των ζευγών των ενώσεων μέσω της συνάρτησης TanimotoSimilarity της βιβλιοθήκης DataStructs.



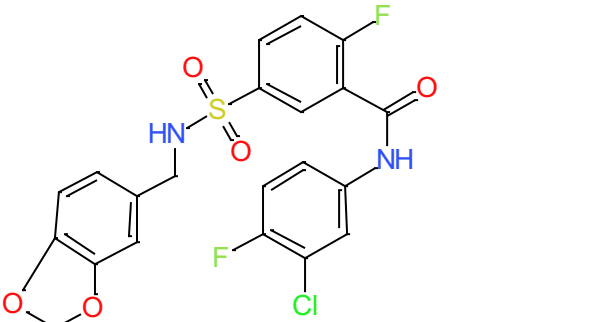
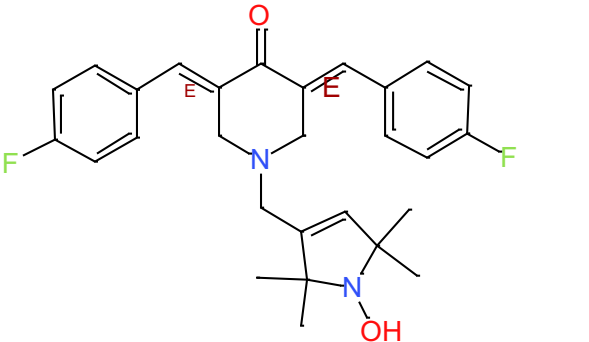
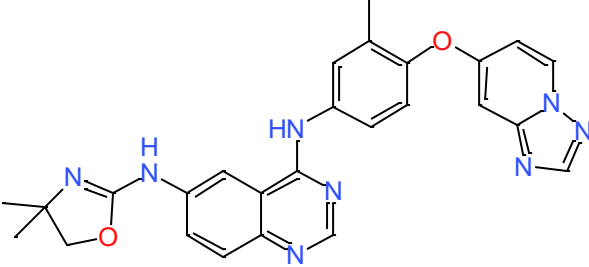
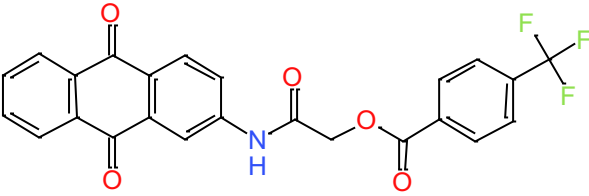
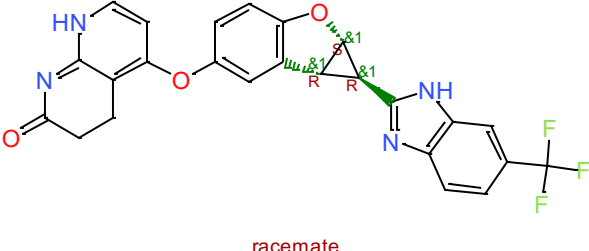
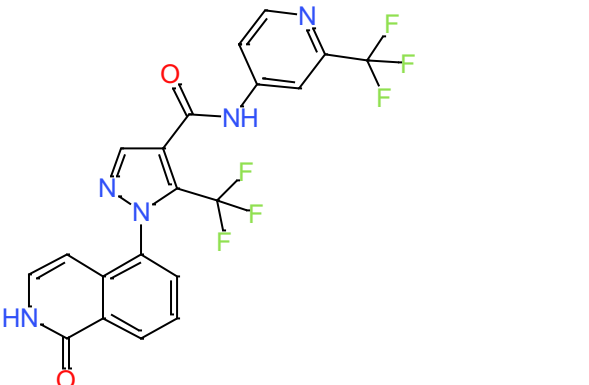
Εικόνα 3.14: Κάτω τριγωνικός χάρτης (triangular heatmap) συντελεστών Tanimoto για τις εκατό καλύτερες ενώσεις που προέκυψαν από την μοριακή πρόσδεση και κωδικοποιούνται από τα μοριακά αποτυπώματα RDKit. Με σκούρο μπλε χρώμα φαίνονται οι ενώσεις με υψηλό ποσοστό ομοιότητας ενώ με σκούρο κόκκινο χρώμα οι ενώσεις με χαμηλό ποσοστό ομοιότητας.

Σε συνέχεια της ανάλυσης και της αξιολόγησης των ενώσεων, επιλέχθηκαν εκείνες με τις καλύτερες αλληλεπιδράσεις με σκοπό τον εντοπισμό των ισχυρότερων, με βάση την συμμετοχή των σημαντικών αμινοξέων όπως αναφέρονται στο κεφάλαιο της δομικής ανάλυσης του GHRH-R. Όπου αυτό δεν ήταν αρκετό η επιλογή έγινε με βάση το καλύτερο σκορ πρόσδεσης. Στον πίνακα 3.2 φαίνονται οι δέκα καλύτερες ενώσεις που προέκυψαν από την εικονική σάρωση. Στην πρώτη στήλη βρίσκεται το αναγνωριστικό κάθε ένωσης (Mol ID) και σε παρένθεση ο αύξων αριθμός όπως προέκυψε μετά το φιλτράρισμα Lipinski. Στην δεύτερη στήλη παρουσιάζεται η δισδιάστατη δομή κάθε ένωσης και στην τρίτη στήλη το όνομα.

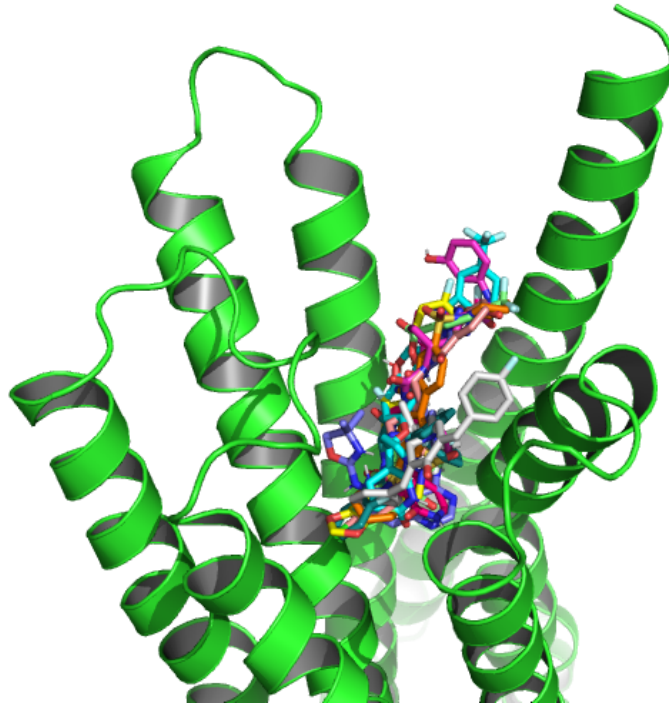
Πίνακας 3.2: Δισδιάστατη δομή τελικών ενώσεων που προέκυψαν από την εικονική σάρωση.

Mol ID (AA)	Δομή	Όνομα
835(823)	 unknown chirality	WAY-300758
3610(3567)		WAY-278599
3498 (2565)		CA3(CIL56)
6932(5373)	 racemate	RSV604

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ
ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

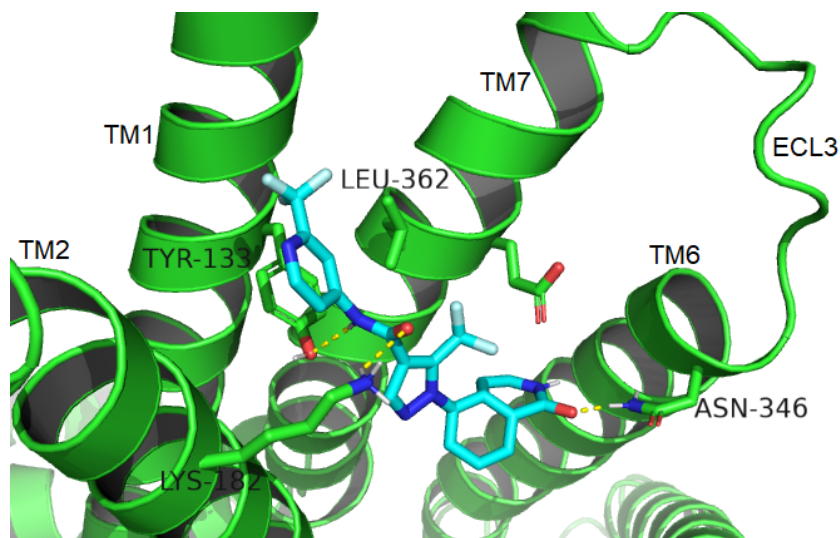
2446(2416)		WAY-335217
2812(2081)		HO-3867
3270(2397)		Tucatinib
5343(3908)		WAY-381244
8782(6685)		Lifirafenib (BGB-283)
8150(6225)		JNJ-67856633

Στην εικόνα 3.15 παρουσιάζονται και οι δέκα ενώσεις στο ορθοστερικό κέντρο της διαμεμβρανικής περιοχής του υποδοχέα.



Εικόνα 3.15: Οι δέκα ενώσεις που προέκυψαν από την μοριακή πρόσδεση τοποθετημένες στο ορθοστερικό κέντρο της διαμεμβρανικής περιοχής του υποδοχέα.

Από τις δέκα ενώσεις ξεχώρισε η ένωση “6225” τόσο λόγω του υψηλού σκορ πρόσδεσης (-9 kcal/mol) όσο και λόγω των αλληλεπιδράσεων που κάνει με σημαντικά αμινοξέα του υποδοχέα. Συγκεκριμένα όπως φαίνεται και στην εικόνα 3.16 κάνει δύο υδρόφοβες αλληλεπιδράσεις με την TYR133^{1.43b} και την LEU362^{7.43b}, δύο δεσμούς υδρογόνου με την TYR133^{1.43b} και την LYS182^{2.76} ενώ καταφέρνει να κάνει και έναν τρίτο δεσμό υδρογόνου με την ASN346^{6.57b}.



Εικόνα 3.16: Οι αλληλεπιδράσεις του GHRH-R με την ένωση που ξεχώρισε από την εικονική σάρωση για την δυναμικά ανταγωνιστική της δράση. Με διακεκομμένες γραμμές φαίνονται οι δεσμοί υδρογόνου.

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ
ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

4. Συζήτηση και Συμπεράσματα

Η εργασία ολοκληρώνεται με τα συμπεράσματα που προέκυψαν από την ανάλυση των αποτελεσμάτων. Η αξιολόγηση των βιοδραστικών ενώσεων MK24, MK04, MK04-6 και MK04-9 γίνεται τόσο με βάση το σκορ πρόσδεσης όσο και μέσω των αλληλεπιδράσεων που κάνουν με τον υποδοχέα προκειμένου να μελετηθεί ο μηχανισμός πρόσδεσής τους και να εντοπιστούν τα αμινοξέα του υποδοχέα που συμμετέχουν στην δέσμευση.

Συγκεκριμένα το σκορ πρόσδεσης (-5.8 kcal/mol) για την ένωση MK24 δεν αντικατοπτρίζει τα πειραματικά αποτελέσματα. Από την άλλη όμως, η ένωση αλληλεπιδρά με συγκεκριμένα αμινοξέα όπως είναι η LEU362^{7.43b}, η TYR133^{1.43b}, η PHE126^{1.36b} και η PHE187^{ECL1} που φαίνεται να παίζουν σημαντικό ρόλο στην αγωνιστική δράση της ένωσης, καθώς η δράση των αμινοξέων αυτών έχει παρατηρηθεί και κατά την πρόσδεση του ενδογενούς πεπτιδίου στον υποδοχέα. Η ένωση δεσμεύεται στον υποδοχέα μέσω πέντε υδρόφοβων αλληλεπιδράσεων και ενός δεσμού υδρογόνου. Στις υδρόφοβες αλληλεπιδράσεις συμμετέχει ο αρωματικός δακτύλιος και η βουτυλομάδα της ένωσης ενώ στον δεσμό υδρογόνου συμμετέχει το άζωτο του δακτυλίου.

Αναφορικά με την ένωση MK04 καθώς και τις 9 παρόμοιες με αυτήν MK04(2-10), έχουν μέσο όρο σκορ πρόσδεσης -7 kcal/mol, ένας δείκτης που επαληθεύει εν μέρη την ανταγωνιστική δράση των πειραματικών δεδομένων. Συγκεκριμένα οι ενώσεις ενδιαφέροντος MK04, MK04-6 και MK04-9 έχουν σκορ πρόσδεσης “-7 kcal/mol”, “-6.5 kcal/mol” και “-7 kcal/mol” αντίστοιχα. Πριν την παράθεση των συμπερασμάτων που προέκυψαν από τις αλληλεπιδράσεις, γίνεται ανάλυση των δομικών διαφορών που έχουν οι τρεις ενώσεις. Αντίθετα με την MK04-6, οι ενώσεις MK04 και MK04-9 έχουν ένα επιπρόσθετο δομικό στοιχείο. Συγκεκριμένα η ένωση MK04 έχει μια παραπάνω υδροξυλομάδα ενώ η ένωση MK04-9 μια παραπάνω αμινομάδα. Από την άλλη η MK04-6 έχει στα άκρα της αιθυλομάδες αντί για μεθυλομάδες που έχουν οι άλλες δύο ενώσεις. Οι μικρές αυτές δομικές διαφορές έχουν ως αποτέλεσμα διαφορετική βιοδραστικότητα στα πειραματικά δεδομένα. Μελετώντας τις αλληλεπιδράσεις που κάνουν με τον υποδοχέα θα παρατηρηθούν τα κοινά σημεία πρόσδεσης αλλά και πως αυτές οι δομικές διαφορές τροποποιούν την δέσμευση κάθε ένωσης. Προχωρώντας με τα συμπεράσματα που προέκυψαν από την μελέτη του μηχανισμού δράσης των ενώσεων, παρατηρήθηκε πως και οι τρεις έχουν ένα κοινό σημείο πρόσδεσης με τον υποδοχέα. Συγκεκριμένα οι ενώσεις MK04 και MK04-9 μέσω του μεθόξυ οξυγόνου και η ένωση MK04-6 μέσω του αιθόξυ οξυγόνου, σχηματίζουν δεσμό υδρογόνου με την πλευρική αλυσίδα της LYS182^{2.67b} ενώ οι ενώσεις MK04 και MK04-9 έχουν ένα επιπλέον κοινό σημείο πρόσδεσης μέσω του δεσμού υδρογόνου με την LYS175^{2.60b} (εικόνα 3.11α, 3.11β, 3.11γ). Η ένωση MK04-6 σχηματίζει έναν δεσμό υδρογόνου μέσω του αιθόξυ οξυγόνου και της πλευρικής αλυσίδας της HIS210^{3.37b} ενώ επίσης οι αιθυλομάδες της ένωσης δίνουν επιπλέον σημεία πρόσδεσης για υδρόφοβες αλληλεπιδράσεις όπως για παράδειγμα με την πλευρική αλυσίδα της LEU362^{7.43b}. Το επιπρόσθετο δομικό στοιχείο των ενώσεων MK04 και MK04-9 προσφέρει επιπλέον σημεία για πολικές αλληλεπιδράσεις. Ειδικότερα για την ένωση MK04 η επιπρόσθετη υδροξυλομάδα προσφέρει ένα επιπλέον σημείο πρόσδεσης μέσω ενός δεσμού υδρογόνου που κάνει με την πλευρική αλυσίδα της SER209^{3.36b}. Από την άλλη η αμινομάδα της ένωσης MK04-9 προσφέρει ένα ακόμα μεγαλύτερο δίκτυο

αλληλεπιδράσεων κάνοντας δεσμό υδρογόνου τόσο με την πλευρική όσο και με την κύρια αλυσίδα της SER209^{3,36b} καθώς και με την πλευρική αλυσίδα της THR213^{3,40}.

Τα επιπρόσθετα δομικά στοιχεία των ενώσεων MK04 και MK04-9 φαίνεται πως ισχυροποιούν την πρόσδεση τους έναντι της ένωσης MK04-6 κάτι που δικαιολογείται από την διαφορά στο σκορ πρόσδεσης (-7 kcal/mol έναντι -6.5 kcal/mol) αν και η διαφορά 0.5 kcal/mol είναι μικρή. Το γενικότερο συμπέρασμα που προκύπτει είναι πως και οι τρεις ενώσεις αλληλεπιδρούν με σημαντικά αμινοξέα του υποδοχέα γεγονός που επιβεβαιώνει εν μέρη τα πειραματικά αποτελέσματα για την ανταγωνιστική δράση των ενώσεων.

Ακολουθούν τα συμπεράσματα που προέκυψαν από την εφαρμογή της υπολογιστικής τεχνικής της εικονικής σάρωσης σε δύο βιβλιοθήκες 14.511 συνολικά βιοδραστικών ενώσεων. Στόχος της εικονικής σάρωσης ήταν η εύρεση μικρών μη πεπτιδικών χημικών ενώσεων με πιθανή ανταγωνιστική δράση έναντι του GHRH-R. Μετά το φιλτράρισμα των ενώσεων με βάση τον κανόνα των πέντε του Lipinski με σκοπό την αφαίρεση ενώσεων των οποίων τα χαρακτηριστικά δεν ευνοούν στην φαρμακοκινητική των φαρμάκων, εφαρμόστηκε η μοριακή πρόσδεση και επιλέχθηκαν οι εκατό ενώσεις με το υψηλότερο σκορ πρόσδεσης. Η αξιολόγηση των ενώσεων συνέχισε με την εφαρμογή ελέγχου ομοιότητας με τον συντελεστή Tanimoto με σκοπό τον εντοπισμό εκείνων των είκοσι ενώσεων που διαφέρουν, όσο το δυνατόν περισσότερο, δομικά μεταξύ τους. Ο λόγος που επιλέχθηκαν ενώσεις με συντελεστή ομοιότητας κοντά στο μηδέν είναι διότι η παρούσα εργασία σκοπεύει στην αξιολόγηση των ευρημάτων και πειραματικά και έχει ως στόχο να εξάγει από τις ενώσεις όσο το δυνατόν περισσότερη πληροφορία. Η προσέγγιση που επιλέχθηκε για την εκπλήρωση αυτού του σκοπού είναι οι ενώσεις να χαρακτηρίζονται από δομική ποικιλομορφία. Έπειτα από αξιολόγηση των είκοσι ενώσεων τόσο με βάση το σκορ πρόσδεσης όσο και με βάση τις αλληλεπιδράσεις, προέκυψαν οι τελικές δέκα ενώσεις με πιθανή ανταγωνιστική δράση.

Στην εικόνα 3.15 φαίνονται οι δέκα ενώσεις που προέκυψαν οι οποίες παρουσιάζουν ποικιλομορφία στις δομές τους, έχουν διαφορετικό τρόπο δέσμευσης και συμμετέχουν διαφορετικά αμινοξέα του υποδοχέα στην πρόσδεση τους. Ειδικότερα οι ενώσεις L2-823, L1-2565, L1-5373 επιλέχθηκαν τόσο για το σκορ πρόσδεσης (-9.5 kcal/mol, -9.1 kcal/mol και -9.1 kcal/mol αντίστοιχα) όσες και για τις ισχυρές αλληλεπιδράσεις που σχηματίζουν με καταλυτικά αμινοξέα του υποδοχέα. Οι ενώσεις L1-6685, L1-2081, L1-2397, L1-3908, L1-6225 επιλέχθηκαν κυρίως για τις αλληλεπιδράσεις που κάνουν με τον υποδοχέα. Ειδικότερα οι δύο πρώτες σχηματίζουν γέφυρα άλατος με το αρνητικά φορτισμένο GLU361^{7,42b}, που μαζί με τις υπόλοιπες αλληλεπιδράσεις συνιστούν μια αρκετά ισχυρή σύνδεση των ενώσεων στον υποδοχέα. Οι επόμενες δύο επιλέχθηκαν καθώς σχηματίζουν ένα δίκτυο ηλεκτροστατικών και υδρόφοβων αλληλεπιδράσεων με σημαντικά αμινοξέα του υποδοχέα. Η τελευταία (L1-6225) επιλέχθηκε τόσο για τις υδρόφοβες αλληλεπιδράσεις που κάνει με τα αμινοξέα LEU362^{7,43b} και TYR133^{1,43b} όσο και για τους δεσμούς υδρογόνου που κάνει με τις πλευρικές αλυσίδες των LYS182^{2,76}, TYR133^{1,43b} και ASN346^{6,57b}. Τέλος οι ενώσεις L2-2416 και L2-3567 επιλέχθηκαν κατά βάση λόγω του καλού σκορ πρόσδεσης που είχαν και δευτερευόντως για τις αλληλεπιδράσεις που κάνανε με τον υποδοχέα.

Η ένωση που ξεχώρισε κατά την εφαρμογή της υπολογιστικής τεχνικής της εικονικής σάρωσης ήταν η ένωση L1-6225 καθώς είχε υψηλό σκορ πρόσδεσης (-9 kcal/mol) αλλά

το βασικότερο ήταν ότι έκανε αλληλεπιδράσεις με σημαντικά αμινοξέα του υποδοχέα. Ειδικότερα όλα τα αμινοξέα με τα οποία αλληλεπιδράει (LEU362^{7.43b}, TYR133^{1.43b}, LYS182^{2.76} και ASN346^{6.57b}) παρατηρήθηκαν κατά την δομική ανάλυση του GHRH-R και την πρόσδεση του ενδογενούς πεπτιδίου. Επίσης το χαρακτηριστικό που ξεχώρισε αυτήν την ένωση είναι ο δεσμός υδρογόνου που σχηματίζει με το ASN346^{6.57b}, ένα αμινοξύ που βρίσκεται στο σημείο σύνδεσης της διαμεμβρανικής έλικας 6 και του εξωκυττάριου βρόγχου 3 και όπως αναφέρεται από την βιβλιογραφία πιθανώς σταθεροποιεί τον βρόγχο στην ενεργή του κατάσταση, συνιστώντας το συγκεκριμένο αμινοξύ έναν σημαντικό στόχο.

Μελλοντικοί στόχοι για την αξιολόγηση των βιοδραστικών ενώσεων είναι η εφαρμογή της υπολογιστικής τεχνικής της μοριακής δυναμικής για τις ενώσεις MK04, MK04-6 και MK04-9 με σκοπό να ελεγχθεί η σταθερότητά τους συναρτήσει του χρόνου. Επίσης μέσω της μοριακής δυναμικής θα ελεγχθεί και ο υποδοχέας GHRH-R εάν κατά την διάρκεια των αλληλεπιδράσεων παραμένει στην ενεργή η ανενεργή του κατάσταση. Τέλος για τις δέκα ενώσεις που προέκυψαν από την μοριακή πρόσδεση επόμενος στόχος είναι η πειραματική τους αξιολόγηση.

Βιβλιογραφία

- [1] Frye L., Bhat S., Akinsanya K. & Abel R., (2021). “From computer-aided drug discovery to computer driven drug discovery”, *Drug Discovery Today: Technologies*, **39**, pp 111-117. <https://doi.org/10.1016/j.ddtec.2021.08.001>
- [2] Kapetanovic I., (2008). “Computer-aided drug discovery and development (CADD): In silico-chemico-biological approach”, *Chemico-Biological Interactions*, **171(2)**, pp 165-176. <https://doi.org/10.1016/j.cbi.2006.12.006>
- [3] Campbell N. A., Reece J. B., Urry L., et al., (2013). “ΒΙΟΛΟΓΙΑ: Η χημεία της ζωής – το κύτταρο – γενετική”, *Πανεπιστημιακές Εκδόσεις Κρήτης*, ISBN 978-960-524-305-0, pp 153-159.
- [4] Tymoczko J. L., Berg J. M., Stryer L., (2015). “ΒΙΟΧΗΜΕΙΑ ΒΑΣΙΚΕΣ ΑΡΧΕΣ”, *Εκδόσεις Π.Χ. Πασχαλίδης*, ISBN 978-1-4641-2613-0, pp 261-294.
- [5] Congreve M., de Graaf C., Swain N. A. & Tate C. G., (2020). “Impact of GPCR Structures on Drug Discovery”, *Cell*, **181(1)**, pp 81-91. <https://doi.org/10.1016/j.cell.2020.03.003>
- [6] Pándy-Szekeres G., Munk C., Tsonkov T. M., et al., (2018) “GPCRdb in 2018: adding GPCR structure models and ligands”, *Nucleic Acids Research*, **46(D1)**, pp D440-D446. <https://doi.org/10.1093/nar/gkx1109>
- [7] Shonberg J., Kling R. C., Gmeiner P. & Löber S., (2015). “GPCR crystal structures: medicinal chemistry in the pocket”, *Bioorganic & Medicinal Chemistry*, **23(14)**, pp 3880-3906. <https://doi.org/10.1016/j.bmc.2014.12.034>
- [8] Latorraca N. R., Venkatakrishnan A.J. & Dror R. O., (2017). “GPCR Dynamics: Structures in Motion”, *Chemical Reviews*, **117(1)**, pp 139-155. <https://doi.org/10.1021/acs.chemrev.6b00177>
- [9] Kobilka B. K., (2007). “G protein coupled receptor structure and activation”, *Biochimica et Biophysica Acta (BBA)-Biomembranes*, **1768(4)**, pp 794-807. <https://doi.org/10.1016/j.bbamem.2006.10.021>
- [10] Tuteja N., (2009). “Signaling through G protein coupled receptors”, *Plant Signaling & Behavior*, **4(10)**, pp 942-947. <https://doi.org/10.4161/psb.4.10.9530>
- [11] Hilger D., Masureel M., & Kobilka B. K., (2018). “Structure and dynamics of GPCR signaling complexes”, *Nature Structural & Molecular Biology*, **25(1)**, pp 4-12. <https://doi.org/10.1038%2Fs41594-017-0011-7>
- [12] Kensal E. van Holde, W. Curtis Johnson, P. Shing Ho, (2006). “Βιοφυσική, αρχές φυσικής βιοχημείας”, *Εκδόσεις Εμβryo*, ISBN 978-960-8002-55-5, pp 331-367.
- [13] Murata K. & Wolf M., (2018). “Cryo-electron microscopy for structural analysis of dynamic biological macromolecules”, *Biochimica et Biophysica Acta (BBA)-General Subjects*, **1862(2)**, pp 324-334. <https://doi.org/10.1016/j.bbagen.2017.07.020>
- [14] Danev R., Yanagisawa H. & Kikkawa M., (2019). “Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions”, *Trends in Biochemical Sciences*, **44(10)**, pp 837-848. <https://doi.org/10.1016/j.tibs.2019.04.008>
- [15] Van Drie J.H. & Tong L., (2020). “Cryo-EM as a powerful tool for drug discovery”, *Bioorganic & Medicinal Chemistry Letters*, **30(22)**, 127524. <https://doi.org/10.1016/j.bmcl.2020.127524>
- [16] Benjin X. & Ling L., (2020). “Developments, applications and prospects of cryo-electron microscopy”, *Protein Science*, **29(4)**, pp 872-882. <https://doi.org/10.1002/pro.3805>

- [17] de Oliveira T. M., van Beek L., Shilliday F., et al., (2021). “Cryo-EM: The Resolution Revolution and Drug Discovery”, *SLAS Discovery: Advancing the Science of Drug Discovery*, **26(1)**, pp 17-31. <https://doi.org/10.1177/2472555220960401>
- [18] Robertson M. J., Meyerowitz J. G. & Skiniotis G., (2022). “Drug Discovery in the era of cryo-electron microscopy”, *Trends in Biochemical Sciences*, **47(2)**, pp 124-135. <https://doi.org/10.1016/j.tibs.2021.06.008>
- [19] Μαρία Οξενκιουν-Πετροπούλου, (2020). ‘Φασματομετρικές Μέθοδοι στην Σύγχρονη Ενόργανη Ανάλυση’, *Εκδόσεις Τσότρας*, ISBN 978-618-5309-88-6, pp 397-423.
- [20] Liapakis G., Matsoukas M. T., Karageorgos V., et al., (2017). “Family B G Protein-coupled Receptors and their Ligands: From Structure to Function”, *Current Medicinal Chemistry*, **24(31)**, pp 3323-3355. <https://doi.org/10.2174/0929867324666170303162416>
- [21] Zhou F., Zhang H., Cong Z., et al., (2020). “Structural basis for activation of the growth hormone-releasing hormone receptor”, *Nature Communications*, **11(1)**, 5205. <https://doi.org/10.1038/s41467-020-18945-0>
- [22] Lin-Su K. & Wajnrath M. P., (2002). “Growth Hormone Releasing Hormone (GHRH) and the GHRH Receptor”, *Reviews in Endocrine & Metabolic Disorders*, **3**, pp 313-323. <https://doi.org/10.1023/A:1020949507265>
- [23] Kato Y., Murakami Y., Sohmiya M. & Nishiki M., (2002). “Regulation of Human Growth Hormone Secretion and Its Disorders”, *Internal Medicine*, **41(1)**, pp 7-13. <https://doi.org/10.2169/internalmedicine.41.7>
- [24] Brinkman J. E., Tariq M. A., Leavitt L. & Sharma S., (2022). “Physiology, Growth Hormone”, *In StatPearls. StatPearls Publishing*.
- [25] Fridlyand L.E., Tamarina N. A., Schally A. V. & Philipson L. H., (2016). “Growth Hormone-Releasing Hormone in Diabetes”, *Frontiers in Endocrinology*, **7**, 129. <https://doi.org/10.3389/fendo.2016.00129>
- [26] Karageorgos V., Venihaki M., Sakellaris S., et al., (2018). “Current understanding of the structure and function of family B GPCRs to design novel drugs”, *Hormones*, **17**, pp 45-59. <https://doi.org/10.1007/s42000-018-0009-5>
- [27] Hilger D., Kumar K. K., Hu H., et al., (2020). “Structural insights into differences in G protein activation by family A and family B GPCRs”, *Science*, **369(6503)**, eaba3373. <https://doi.org/10.1126/science.aba3373>
- [28] Gasteiger J. & Engel T. (Eds.). (2003). “Cheminformatics: A Textbook”, *Wiley-VCH Verlag GmbH & Co. KGaA.*, ISBN: 3-527-30681-1
- [29] Begam B. F. & Kumar J. S., (2012). “A study on Cheminformatics and its Applications on Modern Drug Discovery”, *Procedia Engineering*, **38**, pp 1264-1275. <https://doi.org/10.1016/j.proeng.2012.06.156>
- [30] Todeschini R. & Consonni V., (2008). “Handbook of Molecular Descriptors”, *Wiley-VCH Verlag GmbH.*, ISBN 3-52-29913-0
- [31] Carracedo-Reboredo P., Liñares-Blanco J., Rodríguez-Fernández N., et al., (2021). “A review on machine learning approaches and trends in drug discovery”, *Computational and Structural Biotechnology Journal*, **19**, pp 4538-4558. <https://doi.org/10.1016/j.csbj.2021.08.011>
- [32] David L., Thakkar A., Mercado R. & Engkvist O., (2020). “Molecular representations in AI-driven drug discovery: a review and practical guide”, *Journal of Cheminformatics*, **12(1)**, pp 1-22. <https://doi.org/10.1186/s13321-020-00460-5>

- [33] Warr W. A., (2011). "Representation of chemical structures", *Wiley Interdisciplinary Reviews: Computational Molecular Science*, **1(4)**, p.557-579. <https://doi.org/10.1002/wcms.36>
- [34] Wigh D. S., Goodman J. M., Lapkin A. A., (2022). "A review of molecular representation in the age of machine learning", *Wiley Interdisciplinary Reviews: Computational Molecular Science*, **12(5)**, e1603. <https://doi.org/10.1002/wcms.1603>
- [35] Weininger D., (1988). "SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules", *Journal of Chemical Information and Computer Sciences*, **28(1)**, pp 31-36.
- [36] Saldivar-González F.I., Huerta-García C.S. & Medina-Franco J.L., (2020). "Chemoinformatics-based enumeration of chemical libraries: a tutorial", *Journal of Cheminformatics*, **12(1)**, 64. <https://doi.org/10.1186/s13321-020-00466-z>
- [37] Karthikeyan A. & Priyakumar U. D., (2022). "Artificial intelligence: machine learning for chemical sciences", *Journal of Chemical Sciences*, **134**, pp 1-20. <https://doi.org/10.1007/s12039-021-01995-2>
- [38] Consonni V., Todeschini R., (2010). "Molecular Descriptors", *Recent Advances in QSAR Studies: Methods and Applications*, pp 29-102. https://doi.org/10.1007/978-1-4020-9783-6_3
- [39] Polanski J. & Gasteiger J., (2016). "Computer Representation of Chemical Compounds", *Handbook of Computational Chemistry: Leszczynski J., Puzyn T., Eds*, pp 1-43
- [40] Hassan M., Brown R. D., Varma-O' Brien S., Rogers D., (2006). "Cheminformatics analysis and learning in a data pipelining environment", *Molecular Diversity*, **10(3)**, pp 283-299. <https://doi.org/10.1007/s11030-006-9041-5>
- [41] Cereto-Massagué A., Ojeda M. J., Valls C., et al., (2015). "Molecular fingerprint similarity search in virtual screening", *Methods*, **71**, pp 58-63. <https://doi.org/10.1016/j.ymeth.2014.08.005>
- [42] Παναγιωτόπουλος Β., (2021), "Μελέτες ταξινόμησης φυσικών προϊόντων με εργαλεία Μηχανικής Μάθησης και Χημειοπληροφορικής", Διπλωματική εργασία. <https://polynoe.lib.uniwa.gr/xmlui/handle/11400/1013>
- [43] Huang G., Li J. & Zhao C., (2018). "Computational Prediction and Analysis of Associations between Small Molecules and Binding-Associated S-Nitrosylation Sites", *Molecules*, **23(4)**, 954. <https://doi.org/10.3390/molecules23040954>
- [44] Glen R. C., Bender A., Arnby C. H., et al., (2006). "Circular fingerprints: Flexible molecular descriptors with applications from physical chemistry to ADME", *IDrugs*, **9(3)**, pp 199-204.
- [45] Yin Z., Song W., Li B., et al., (2023). "Neural networks prediction of the protein-ligand binding affinity with circular fingerprints", *Technology and Health Care*, **31**, pp 487-495.
- [46] Maggiora G., Vogt M., Stumpfe D. & Bajorath J., (2014). "Molecular Similarity in Medicinal Chemistry", *Journal of Medicinal Chemistry*, **57(8)**, pp 3186-3204. <https://doi.org/10.1021/jm401411z>
- [47] Salim N., Holliday J. & Willet P., (2003). "Combination of Fingerprint-Based Similarity Coefficients Using Data Fusion", *Journal of Chemical Information and Computer Sciences*, **43(2)**, pp 435-442. <https://doi.org/10.1021/ci025596j>
- [48] Willet P., (2006). "Similarity-based virtual screening 2D fingerprints", *Drug Discovery Today*, **11(23-24)**, pp. 1046-1053. <https://doi.org/10.1016/j.drudis.2006.10.005>
- [49] Θωμάς Μαυρομούστακος & Παναγιώτης Ζουμπουλάκης, (2008). "Μοριακή Μοντελοποίηση : Εφαρμογές στην Οργανική και Φαρμακευτική Χημεία", *Εκδόσεις Γιάννης Β. Παριανός*, ISBN 978-960-89486-5-5

- [50] Meng X. Y., Zhang H. X., Mezei M. & Cui M., (2011). “Molecular Docking: A powerful approach for structure-based drug discovery”, *Current Computer-Aided Drug Design*, **7(2)**, pp 146-157. <https://doi.org/10.2174/157340911795677602>
- [51] Ματσούκας Μ. Τ., (2023). “Φαρμακευτική Μηχανική”, Σημειώσεις Μαθήματος.
- [52] Ferreira L. G., Dos Santos R. N., Oliva G. & Andricopulo A. D., (2015). “Molecular Docking and Structure-Based Drug Design Strategies”, *Molecules*, **20(7)**, pp 13384-13421. <https://doi.org/10.3390/molecules200713384>
- [53] Agrafiotis D. K., Gibbs A. C., Zhu F., Izrailev S. & Martin E., (2007). “Conformational Sampling of Bioactive Molecules: A Comparative Study”, *Journal of Chemical Information and Modeling*, **47(3)**, pp 1067-1086. <https://doi.org/10.1021/ci6005454>
- [54] Guedes I. A., de Magalhães C.S. & Dardenne L. E., (2014). “Receptor-ligand molecular docking”, *Biophysical Reviews*, **6**, pp. 75-87. <https://doi.org/10.1007/s12551-013-0130-2>
- [55] Sousa S. F., Fernandes P. A. & Ramos M. J., (2006). “Protein-ligand docking: Current status and future challenges”, *Proteins: Structure, Function and Bioinformatics*, **65(1)**, pp 15-26. <https://doi.org/10.1002/prot.21082>
- [56] Ντούνης Α., (2023). “Εξελκτικός Υπολογισμός”, Σημειώσεις Μαθήματος.
- [57] Carpenter K. A. & Huang X., (2018). “Machine Learning-based Virtual Screening and Its Applications to Alzheimer’s Drug Discovery: A Review”, *Current Pharmaceutical Design*, **24(8)**, pp 3347-3358. <https://doi.org/10.2174/1381612824666180607124038>
- [58] Bhunia S. S., Saxena M. & Saxena A. K., (2021). “Ligand- and Structure-Based Virtual Screening in Drug Discovery”, *Biophysical and Computational Tools in Drug Discovery*, **37**, pp 281-339. https://doi.org/10.1007/7355_2021_130
- [59] Sabe V. T., Ntombela T., Jhamba L. A., et al., (2021). “Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review”, *European Journal of Medicinal Chemistry*, **224**, 113705. <https://doi.org/10.1016/j.ejmech.2021.113705>
- [60] Lipinski C. A., Lombardo F., Dominy B. W. & Feeney P. J., (2012). “Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings”, *Advanced Drug Delivery Reviews*, **64**, pp 4-17. <https://doi.org/10.1016/j.addr.2012.09.019>
- [61] Batool M., Ahmad B. & Choi S., (2019). “A Structure-Based Drug Discovery Paradigm”, *International Journal of Molecular Sciences*, **20(11)**, 2783. <https://doi.org/10.3390/ijms20112783>
- [62] Katsila T., Spyroulias G. A., Patrinos G. P. & Matsoukas M. T., (2016). “Computational approaches in target identification and drug discovery”, *Computational and Structural Biotechnology Journal*, **14**, pp 177-184. <https://doi.org/10.1016/j.csbj.2016.04.004>
- [63] Ballante F., Kooistra A. J., Kampen S., et al., (2021). “Structure-based virtual screening for ligands of G protein-coupled receptors: what can molecular docking do for you?”, *Pharmacological Reviews*, **73(4)**, pp 1698-1736. <https://doi.org/10.1124/pharmrev.120.000246>
- [64] Macalino S. J. Y., Gosu V., Hong S. & Choi S., (2015). “Role of computer-aided drug design in modern drug discovery”, *Archives of Pharmacol Research*, **38**, pp 1686-1701. <https://doi.org/10.1007/s12272-015-0640-5>
- [65] Tarryn Radomsky, (2021). “Identification and characterization of small-molecule growth hormone-releasing hormone receptor-interacting compounds”, Διπλωματική εργασία.

[66] Matsoukas M. T. & Spyroulias G. A., (2017). “Dynamic properties of the growth hormone releasing hormone receptor (GHRHR) and molecular determinants of GHRH binding”, *Molecular BioSystems*, **13(7)**, pp 1313-1322. <https://doi.org/10.1039/C7MB00130D>

[67] Matsoukas M. T., Radomsky T., Panagiotopoulos V., et al. (2023). “Identification of small-molecule antagonists targeting the Growth Hormone Releasing Hormone Receptor (GHRHR)”, *Authorea*. <https://doi.org/10.22541/au.169407025.54926709/v1>

[68] RDKit, <https://www.rdkit.org/docs/GettingStartedInPython.html>

Παράρτημα

Σύνδεσμος για το αποθετήριο στο GitHub που περιλαμβάνει τους κώδικες που χρησιμοποιήθηκαν στην εργασία: <https://github.com/TsianKon/GHRH-R>

Πίνακας 6.1: Οι εκατό καλύτερες ενώσεις που προέκυψαν από την εικονική σάρωση ταξινομημένες με βάση το σκορ πρόσδεσης.

Ενώσεις	Affinity (kcal/mol)
compound L1 6700	-9.9
compound L2 5176	-9.9
compound L1 595	-9.8
compound L1 4480	-9.7
compound L1 5216	-9.7
compound L2 1845	-9.7
compound L2 1882	-9.6
compound L1 2355	-9.5
compound L1 4223	-9.5
compound L1 6778	-9.5
compound L1 76	-9.5
compound L2 823	-9.5
compound L1 4019	-9.4
compound L1 4782	-9.4
compound L1 5102	-9.4
compound L1 6903	-9.4
compound L2 223	-9.4
compound L2 2631	-9.4
compound L2 4985	-9.4
compound L1 2063	-9.3
compound L1 2203	-9.3
compound L1 3023	-9.3
compound L1 5551	-9.3
compound L1 6537	-9.3
compound L1 6834	-9.3
compound L1 6902	-9.3
compound L1 692	-9.3
compound L2 184	-9.3
compound L2 3061	-9.3
compound L2 3753	-9.3
compound L2 4043	-9.3
compound L2 5174	-9.3
compound L1 2001	-9.2
compound L1 3766	-9.2
compound L1 454	-9.2
compound L1 4760	-9.2
compound L1 4880	-9.2
compound L1 5240	-9.2
compound L1 5570	-9.2

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ
ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

compound_L1_6155	-9.2
compound_L2_2835	-9.2
compound_L2_3313	-9.2
compound_L2_3426	-9.2
compound_L2_3479	-9.2
compound_L2_3481	-9.2
compound_L2_3567	-9.2
compound_L2_3981	-9.2
compound_L2_4030	-9.2
compound_L2_4046	-9.2
compound_L2_5177	-9.2
compound_L1_1468	-9.1
compound_L1_2400	-9.1
compound_L1_2565	-9.1
compound_L1_4682	-9.1
compound_L1_5373	-9.1
compound_L2_1739	-9.1
compound_L2_2416	-9.1
compound_L2_3845	-9.1
compound_L2_4208	-9.1
compound_L2_5175	-9.1
compound_L1_172	-9.0
compound_L1_2081	-9.0
compound_L1_2397	-9.0
compound_L1_2481	-9.0
compound_L1_2484	-9.0
compound_L1_25	-9.0
compound_L1_3839	-9.0
compound_L1_3908	-9.0
compound_L1_4170	-9.0
compound_L1_4286	-9.0
compound_L1_4484	-9.0
compound_L1_4680	-9.0
compound_L1_663	-9.0
compound_L1_6685	-9.0
compound_L1_670	-9.0
compound_L1_6758	-9.0
compound_L1_6829	-9.0
compound_L2_1595	-9.0
compound_L2_1708	-9.0
compound_L2_2242	-9.0
compound_L2_2555	-9.0
compound_L2_3529	-9.0
compound_L2_4045	-9.0
compound_L2_4342	-9.0
compound_L2_5041	-9.0
compound_L1_1987	-8.9
compound_L1_2007	-8.9

ΔΟΜΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΣΤΟΧΕΥΣΗ ΤΟΥ ΥΠΟΔΟΧΕΑ ΟΡΜΟΝΗΣ
ΑΠΕΛΕΥΘΕΡΩΣΗΣ ΑΥΞΗΤΙΚΗΣ ΟΡΜΟΝΗΣ ΜΕ ΔΥΝΗΤΙΚΑ ΒΙΟΔΡΑΣΤΙΚΑ ΜΟΡΙΑ

compound_L1_2052	-8.9
compound_L1_2148	-8.9
compound_L1_2254	-8.9
compound_L1_2666	-8.9
compound_L1_3753	-8.9
compound_L1_3886	-8.9
compound_L1_425	-8.9
compound_L1_4822	-8.9
compound_L1_5142	-8.9
compound_L1_5560	-8.9
compound_L1_5645	-8.9
compound_L1_5853	-8.9
compound_L1_6225	-8.9