

Πανεπιστήμιο Ιωαννίνων

Δ.Π.Μ.Σ “Ιατρική Χημεία”



Μεταπτυχιακή Εργασία

**Αναγνώριση Φαρμακευτικών Στόχων με Τεχνικές *In Silico* για Νέα
Θεραπευτικά Μέσα Έναντι των Αυτοφλεγμονωδών Νοσημάτων**

Γεωργία Κουνναπή, Α.Μ.: 304

Επιβλέπων Καθηγητής: Κωνσταντίνος Παπαλουκάς

Ιωάννινα, Ιανουάριος 2026

Συντομογραφίες

SAIDs	Συστημικά Αυτοφλεγμονώδη Νοσήματα
HPF	Κληρονομικός Περιοδικός Πυρετός
FMF	Οικογενής Μεσογειακός Πυρετός
TRAPS	Περιοδικό Εμπύρετο Σύνδρομο Σχετιζόμενο με τον Υποδοχέα TNF
HIDS	Σύνδρομο Υπερέκκρισης IgD με Περιοδικό Πυρετό
CAPS	Σχετιζόμενο με την Κρυοπυρίνη Περιοδικό Σύνδρομο
FCAS	Οικογενές Αυτοφλεγμονώδες Σύνδρομο εκ Ψύχους
CINCA	Χρόνιο Βρεφικό Νευρολογικό-Δερματικό Αρθρικό
NOMID	Πολυσυστηματική Φλεγμονώδης Νόσος Νεογνικής Εμφάνισης
IL	Ιντερλευκίνη
NLRP3	Τομέας Πυρίνης της Οικογένειας NLR που περιέχει 3
DAMPs	Μοριακά Μοτίβα Σχετιζόμενα με Βλάβη
PAMPs	Μοριακά Μοτίβα Σχετιζόμενα με το Παθογόνο
PRRs	Υποδοχείς Αναγνώρισης Προτύπων
TLRs	Υποδοχείς Αναγνώρισης Προτύπων Toll
NLRs	Υποδοχείς Τύπου NOD
CLRs	Υποδοχείς Λεκτίνης Τύπου C
RLRs	Υποδοχείς Τύπου RIG-I
PRAAS	Αυτοφλεγμονώδη Σύνδρομα Σχετιζόμενα με το Πρωτεάσωμα
TNFR1	Υποδοχέας Παράγοντα Νέκρωσης Όγκου

ER	Ενδοπλασματικό Δίκτυο
ROS	Αντιδραστικά Είδη Οξυγόνου
MAPK	Ενεργοποιημένη από Μιτογόνο Πρωτεϊνική Κινάση
CRD	Περιοχές Πλούσιες σε Κυστεΐνη
DD	Περιοχή Ενδοκυτταρικού Θανάτου
PLAD	Περιοχή Σύνδεσης προ-Συνδέτη
NF-κB	Πυρηνικός Παράγοντας κB
JNK	C-Jun N-τερματικής κινάσης
NALP3	Περιοχές Πλούσιες σε Λευκίνη και Πυρίνης που Περιέχουν Πρωτεΐνη 3
NLR	Τομέας Δέσμευσης Νουκλεοτιδίων
CRP	C-αντιδρώσας Πρωτεΐνης
SAA	Αμυλοειδές Α Ορού
AA	Αμυλοειδές Α
IgD	Ανοσοσφαιρίνη D
CNN	Συνελικτικό Νευρωνικό Δίκτυο
MSE	Μέσο Τετραγωνικό Σφάλμα
RMSE	Ρίζα Μέσου Τετραγωνικού Σφάλματος
MVA	Μεβαλονική Οξυουρία
MVK	Μεβαλονική Κινάση
MKD	Ανεπάρκεια Μεβαλονικής Κινάσης
PBMCs	Μονοπύρρηνα Κύτταρα Περιφερικού Αίματος
IPP	Ισοπεντενυλοπυροφωσφορικό Άλας
ΔGbind	Ελεύθερη Ενέργεια Δέσμευσης

Περιεχόμενα

Abstract.....	1
Περίληψη	2
1. Εισαγωγή	3
1.1 Αυτοφλεγμονώδη Νοσήματα.....	3
1.1.1 Παθογένεια των αυτοφλεγμονωδών νοσημάτων	4
1.2 Περιοδικό Εμπύρετο Σύνδρομο Σχετιζόμενο με τον Υποδοχέα TNF	6
1.2.1 Γενετική των TRAPS.....	7
1.2.2 Θεραπεία των TRAPS	9
1.3 Άλλα Αυτοφλεγμονώδη Νοσήματα.....	10
1.4 Βιοπληροφορική Προσέγγιση	14
1.4.1 Μοριακή πρόσδεση	15
2. Μέθοδοι και Υλικά	19
2.1 Υπολογιστική Προσέγγιση	19
2.1.1 Βάσεις δεδομένων και σύνολα δεδομένων δομής	20
2.1.2 Αξιολόγηση μοριακής πρόσδεσης.....	22
2.1.3 Εντοπισμός περιοχών πρόσδεσης στην πρωτεΐνη	23
2.1.4 Ομαδοποίηση ενώσεων	24
2.2 Αλγόριθμοι Μοριακής Πρόσδεσης	24
2.2.1 Μοριακή πρόσδεση χρησιμοποιώντας το AutoDock Vina.....	26
2.2.2 Μοριακή πρόσδεση χρησιμοποιώντας το AutoDock4.2.....	28
2.2.3 Ανάπτυξη μοντέλου Συνελκτικού Νευρωνικού Δικτύου για μοριακή πρόσδεση	31
3. Αποτελέσματα	35
3.1. Απόδοση και Αξιολόγηση του Συνελκτικού Νευρωνικού Δικτύου	35
3.2 Αποτελέσματα Μοριακής Πρόσδεσης.....	36
3.3 Ανάλυση Θέσεων Πρόσδεσης των Υποψηφίων Ενώσεων και των Κοιλοτήτων της Πρωτεΐνης TNFR1	40
3.4 Σύγκριση με Βιβλιογραφικά Δεδομένα	44
4. Συζήτηση και Συμπεράσματα	53
4.1 Αξιολόγηση της Υπολογιστικής Προσέγγισης.....	54

4.2 Ολοκληρωμένη Ανάλυση Υπολογιστικών Αποτελεσμάτων Μοριακής Πρόσδεσης ...	57
4.3 Περιορισμοί της Παρούσας Μελέτης	64
4.4 Μελλοντική Εργασία.....	65
5. Βιβλιογραφία.....	67
6. Παράρτημα.....	71
I. Δομή Αρχείου SDF.....	71
II. Δομή Αρχείου PDBQT	72
III. Παράμετροι για το AutoDock Vina.....	73
IV. Αποτελέσματα AutoDock Vina.....	74
V. Αρχείο Παραμέτρων Πλέγματος (grf).....	75
VI. Αποτελέσματα AutoDock4.2	76
VII. Ενέργειες δέσμευσης στα AutoDock Vina, AutoDock4.2, και στο μοντέλο CNN	77
VIII. Κώδικας CNN	83
IX. Κώδικας Πρόβλεψης Ενεργειών Δέσμευσης των Μορίων	86

Abstract

Systemic Autoinflammatory Diseases (SAIDs) constitute a diverse group of disorders characterized by recurrent systemic inflammation without evidence of infection or autoimmunity. Since the first identification of the MEFV gene in Familial Mediterranean Fever (FMF) in 1997, the understanding of SAIDs has significantly expanded, revealing the central role of the innate immune system and its molecular regulators. Among these diseases, the Tumor Necrosis Factor Receptor Associated Periodic Syndrome (TRAPS) represents a prototypical monogenic autoinflammatory disorder caused by mutations in the TNFRSF1A gene. This gene encodes the Tumor Necrosis Factor Receptor 1 (TNFR1), a crucial mediator of the TNF- α signaling pathway. Structural or functional alterations in TNFR1 lead to dysregulated inflammation, contributing to the recurrent febrile episodes and chronic inflammatory symptoms characteristic of TRAPS. The aim of this study was to employ computational and cheminformatics methodologies to identify small molecules capable of modulating TNFR1 activity, with the ultimate goal of exploring potential therapeutic candidates for TRAPS. The crystal structure of TNFR1 (PDB ID: 1EXT) was selected as a reference model for the molecular docking experiments. A blind docking approach was performed to investigate potential binding sites across the protein surface, enabling the identification of regions suitable for small molecule interaction.

Two docking tools, AutoDock 4.2 and AutoDock Vina, were employed to ensure methodological consistency and to provide comparative insight into the predicted binding affinities (ΔG_{bind}) and interaction modes. A dataset of more than 7,000 compounds was extracted from the DrugBank database, comprising approved drugs, withdrawn, and investigational compounds with known pharmacological relevance. The docking results were analyzed and ranked based on predicted binding energies, hydrogen bonding patterns, and interaction residues. In addition, a Convolutional Neural Network (CNN) model was developed and trained to predict and classify binding affinities based on structural and electronic descriptors of the compounds. The CNN predictions were in strong agreement with the results from AutoDock Vina, confirming the validity of the machine learning model as a complementary tool for *in silico* screening. The use of the CNN significantly reduced the computational cost and processing time, demonstrating the efficiency of artificial intelligence-assisted approaches in drug discovery. Overall, the results of this study highlight the power of integrated *in silico* methods, combining molecular docking with machine learning techniques, in accelerating the early stages of drug discovery. The computational identification of small molecules with high affinity for TNFR1 provides valuable insight into the structural determinants of receptor modulation and opens new perspectives for the development of targeted therapies against TRAPS and related autoinflammatory disorders.

Περίληψη

Οι Συστηματικές Αυτοφλεγμονώδεις Νόσοι (Systemic Autoinflammatory Diseases, SAIDs) αποτελούν μια ετερογενή ομάδα διαταραχών που χαρακτηρίζονται από επαναλαμβανόμενα επεισόδια συστηματικής φλεγμονής, χωρίς ενδείξεις λοίμωξης ή αυτοανοσίας. Από την πρώτη ταυτοποίηση του γονιδίου MEFV στον Οικογενή Μεσογειακό Πυρετό (Familial Mediterranean Fever, FMF) το 1997, η κατανόηση των SAIDs έχει διευρυνθεί σημαντικά, αναδεικνύοντας τον κεντρικό ρόλο της έμφυτης ανοσίας και των μοριακών ρυθμιστών της. Μεταξύ αυτών των νοσημάτων, το Σύνδρομο Περιοδικού Πυρετού που σχετίζεται με τον υποδοχέα TNF (Tumor Necrosis Factor Receptor Associated Periodic Syndrome, TRAPS) αποτελεί χαρακτηριστικό μονογονιδιακό αυτοφλεγμονώδες νόσημα που προκαλείται από μεταλλάξεις στο γονίδιο TNFRSF1A. Το γονίδιο αυτό κωδικοποιεί τον υποδοχέα TNFR1, έναν κρίσιμο διαμεσολαβητή της σηματοδοτικής οδού του TNF-α. Δομικές ή λειτουργικές αλλοιώσεις του TNFR1 οδηγούν σε απορρύθμιση της φλεγμονής, συμβάλλοντας στα επαναλαμβανόμενα επεισόδια πυρετού και στα χρόνια φλεγμονώδη συμπτώματα που χαρακτηρίζουν το TRAPS.

Στόχος της παρούσας μελέτης ήταν η αξιοποίηση υπολογιστικών και χημειοπληροφορικών μεθοδολογιών για τον εντοπισμό μικρών μορίων ικανών να τροποποιήσουν τη λειτουργικότητα του TNFR1, με απώτερο σκοπό τον προσδιορισμό πιθανών θεραπευτικών υποψηφίων για την αντιμετώπιση του TRAPS. Ως μοντέλο αναφοράς στη μοριακή πρόσδεση επιλέχθηκε η κρυσταλλική δομή του TNFR1 (PDB ID: 1EXT). Εφαρμόστηκε προσέγγιση τυφλής μοριακής πρόσδεσης (blind docking) για τη διερεύνηση πιθανών θέσεων πρόσδεσης σε ολόκληρη την επιφάνεια της πρωτεΐνης, επιτρέποντας τον εντοπισμό περιοχών κατάλληλων για αλληλεπίδραση με μικρά μόρια. Για την ενίσχυση της μεθοδολογικής αξιοπιστίας χρησιμοποιήθηκαν δύο εργαλεία μοριακής πρόσδεσης, το AutoDock 4.2 και το AutoDock Vina, τα οποία παρείχαν συγκριτικά δεδομένα ως προς τις προβλεπόμενες ενέργειες πρόσδεσης (ΔG_{bind}) και τους τρόπους αλληλεπίδρασης. Το σύνολο δεδομένων αποτελούνταν από περισσότερες από 7,000 χημικές ενώσεις εγκεκριμένων φαρμάκων, υπό διερεύνηση ενώσεων και φαρμάκων που έχουν αποσυρθεί. Τα αποτελέσματα ταξινομήθηκαν με βάση τις ενέργειες πρόσδεσης και τα αμινοξικά κατάλοιπα που συμμετέχουν στις αλληλεπιδράσεις. Παράλληλα, αναπτύχθηκε και εκπαιδεύτηκε ένα Συνελικτικό Νευρωνικό Δίκτυο (Convolutional Neural Network, CNN) για την πρόβλεψη και κατηγοριοποίηση των συγγενειών πρόσδεσης. Οι προβλέψεις του CNN παρουσίασαν υψηλή συμφωνία με τα αποτελέσματα του AutoDock Vina, επιβεβαιώνοντας την εγκυρότητα του μοντέλου ως συμπληρωματικού εργαλείου στην *in silico* αξιολόγηση μορίων. Επιπλέον, η χρήση του CNN επέφερε σημαντική μείωση του υπολογιστικού χρόνου, αναδεικνύοντας την αποδοτικότητα των προσεγγίσεων τεχνητής νοημοσύνης στη διαδικασία ανακάλυψης φαρμάκων. Η υπολογιστική αναγνώριση μικρών μορίων με υψηλή συγγένεια για τον TNFR1 προσφέρει πολύτιμες γνώσεις για τους δομικούς παράγοντες που καθορίζουν τη ρύθμιση του υποδοχέα και ανοίγει νέες προοπτικές για την ανάπτυξη στοχευμένων θεραπευτικών προσεγγίσεων έναντι του TRAPS και άλλων συγγενών αυτοφλεγμονωδών νοσημάτων.

1. Εισαγωγή

1.1 Αυτοφλεγμονώδη Νοσήματα

Η έννοια της αυτοφλεγμονής έχει χρησιμοποιηθεί εκτενώς για να περιγράψει μια ομάδα συνδρόμων που χαρακτηρίζονται από υποτροπιάζοντες πυρετούς και διάφορες μορφές συστηματικής φλεγμονής, χωρίς τυπικές εκδηλώσεις αυτοάνοσων ασθενειών. Αν και μοιράζονται ορισμένα κοινά χαρακτηριστικά με αυτοάνοσες ασθένειες, οι συστηματικές αυτοφλεγμονώδεις ασθένειες (Systemic Autoinflammatory Diseases, SAIDs) εμφανίζουν μοναδικά χαρακτηριστικά, συμπεριλαμβανομένης της περιοδικότητας καθώς και της έλλειψη ενδείξεων εμπλοκής της προσαρμοστικής ανοσίας, όπως η συσχέτιση με απλότυπους HLA, αυτοαντισώματα υψηλού τίτλου (high-titer autoantibodies) ή αντιγόνο-ειδικά T κύτταρα.^[1] Η αυτοφλεγμονή επιβεβαιώθηκε από τους McDermott και τους συνεργάτες του που πρότειναν αρχικά το περιοδικό σύνδρομο που σχετίζεται με τον υποδοχέα του παράγοντα νέκρωσης όγκων (Tumor Necrosis Factor Receptor-Associated Periodic Syndrome, TRAPS) ως το πρωτότυπο μιας οικογένειας κυρίαρχα κληρονομικών αυτοφλεγμονώδων συνδρόμων που μοιράζονται την εξασθενημένη κάθαρση του υποδοχέα κυτοκίνης ως τον κύριο μηχανισμό ασθένειας.^[2] Τα επόμενα χρόνια, πολλές μελέτες συμφώνησαν για να αποδείξουν ότι όχι μόνο η δυσλειτουργία των υποδοχέων αλλά και η κυτοκίνη μπορεί να προκαλέσει τον αυτοφλεγμονώδη φαινότυπο πολλών διαφορετικών γενετικών συνδρόμων. Η κυτοκίνη που εμπλέκεται στην πλειονότητα των συνδρόμων αυτών αποδείχθηκε ότι είναι η IL-1β η οποία είναι μια ισχυρή προφλεγμονώδης κυτοκίνη που προκαλεί συμπτώματα όπως πυρετό, ανορεξία και αυξημένα επίπεδα δεικτών φλεγμονής στον ορό. Σε περιπτώσεις όπου η δραστηριότητα της IL-1β είναι πολύ υψηλή, εμφανίζεται βλάβη στους ιστούς με αποτέλεσμα την καταστροφή των αρθρώσεων^[1,2].

Έκτοτε, η αξιοσημείωτη πρόοδος στη γονιδιωματική αλλά και ο αυξανόμενος αριθμός κλινικών αναφορών οδήγησαν στην αναγνώριση των αυτοφλεγμονωδών νοσημάτων ως ξεχωριστή οντότητα. Η πλειονότητα των SAIDs είναι κληρονομικές ασθένειες με το αιτιολογικό γονίδιο να έχει απομονωθεί. Αρκετά νέα μονογονιδιακά SAIDs ανακαλύφθηκαν την τελευταία δεκαετία με τη πιο γνωστή υποομάδα που περιγράφηκε για πρώτη φορά να είναι η Κληρονομικοί Περιοδικοί Πυρετοί (HPF), όπου συμπεριλαμβάνει τέσσερις ιστορικές νόσους που εξαρτώνται από την ιντερλευκίνη (IL)-1. Συγκεκριμένα, ο Οικογενής Μεσογειακός Πυρετός (Familial Mediterranean Fever, FMF) που είναι το πιο συχνό SAID, οι Κρυοπυρινοπάθειες που περιλαμβάνουν ένα φάσμα ασθενειών οι οποίες σχετίζονται με μεταλλάξεις του γονιδίου NLRP3, η Ανεπάρκεια Μεβαλονικής Κινάσης (Mevalonate Kinase Deficiency, MKD) και το σύνδρομο TRAPS σχετίζεται με τον υποδοχέα τύπου 1A. Ενδιαφέρον αποτελεί το γεγονός ότι ορισμένες μονογονιδιακές ασθένειες έχουν διαφορετική κλινική έκφραση ανάλογα με τη θέση της μετάλλαξης στο γονίδιο. Ως αποτέλεσμα, το κλινικό φάσμα των ασθενειών που σχετίζονται με μεταλλάξεις στα ακόλουθα γονίδια MEFV, NLRC4, NOD2, PSTPIP1 να έχει επεκταθεί. Επίσης σημαντικό αποτελεί η ανακάλυψη ασθενών που εμφανίζουν σωματικές μεταλλάξεις των γονιδίων

που σχετίζονται με τα SAIDs όμως μεταξύ αυτών των ασθενών μόνο ένα κλάσμα των κυττάρων είναι μεταλλαγμένο [2].

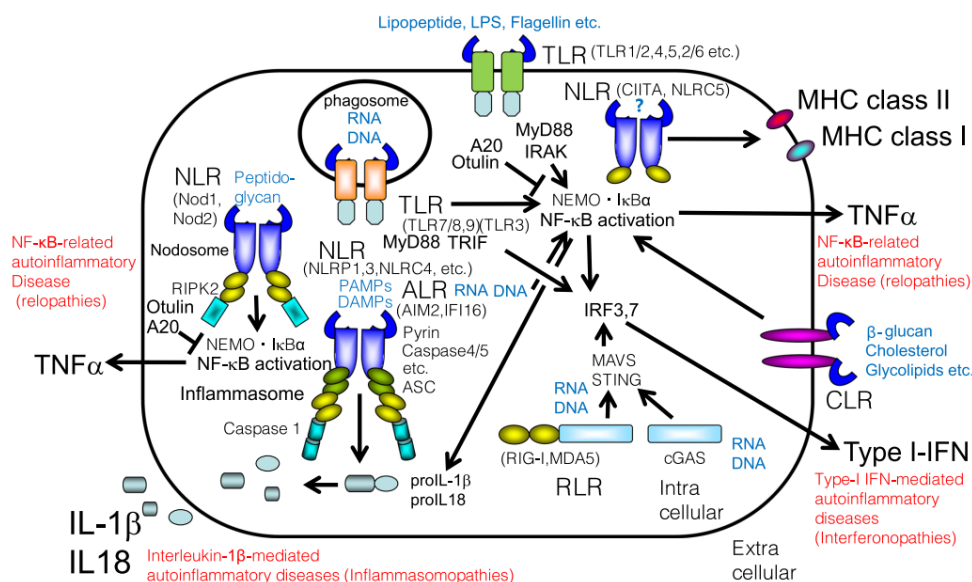
Αναντίρρητα, η βελτιωμένη κατανόηση της έμφυτης ανοσίας και των διαταραχών της αυτοφλεγμονής έχει συμβάλει σε ένα αυξανόμενο οπλοστάσιο στοχευμένων θεραπειών, όχι μόνο βελτιώνοντας δραστικά τη ζωή των ασθενών αλλά και επιβεβαιώνοντας υποθέσεις για τον περίπλοκο ρόλο του έμφυτου ανοσοποιητικού συστήματος στην αυτοφλεγμονή. Τα αυτοάνοσα νοσήματα ανταποκρίνονται σε βιολογικούς παράγοντες που στοχεύουν τις λειτουργίες των T- και B-κυττάρων, συμπεριλαμβανομένων των anti-TNFα, anti-IL-6 (αντι-ιντερλευκίνη) υποδοχέων και anti-IL-12/IL-23 αντισωμάτων. Αυτά τα θεραπευτικά, ωστόσο, δεν έχουν επιδράσεις στη θεραπεία αυτοφλεγμονωδών νόσων, αντίθετα εμφανίζουν δραματικές κλινικές αποκρίσεις στους αναστολείς της IL-1 [3]. Στα SAIDs, ο ένοχος είναι το μονοκύτταρο-μακροφάγο και όχι το T-κύτταρο και το πρόβλημα στις περισσότερες περιπτώσεις είναι μια απορρύθμιση της IL-1β, η οποία επάγει αρκετά άλλα προφλεγμονώδη γονίδια. Η φλεγμονή που προκαλείται από την κυτοκίνη πυροδοτεί επίσης την έκφραση γονιδίων που κωδικοποιούν αντιφλεγμονώδεις πρωτεΐνες που καταστέλλουν τη φλεγμονή. Ενώ σε ορισμένες ασθένειες η σχέση μεταξύ της γονιδιακής μετάλλαξης και του φλεγμονώδους φαινοτύπου που προκαλείται από την IL-1β είναι προφανής σε άλλες δεν είναι. Για αυτό τον λόγο, ιδιαίτερα σημαντικός είναι ο ανταγωνιστής του υποδοχέα IL-1 (IL-1Ra) που αναστέλλει ειδικά τη δραστηριότητα της IL-1. Η IL-1Ra είναι δομικά παρόμοια με την IL-1β αλλά στερείται βιολογικής δράσης. Συνδέεται στενά με τον υποδοχέα IL-1 εμποδίζοντας έτσι την πρόσβαση της IL-1. Τόσο η IL-1, όσο και η IL-1Ra παράγονται σε ασθενείς με λοιμώξεις, τραύματα ή άλλες φλεγμονώδεις καταστάσεις και ανταγωνίζονται για την κατάληψη του υποδοχέα. Ως εκ τούτου, το αποτέλεσμα μιας φλεγμονώδους διαδικασίας είναι πιθανό να επηρεαστεί από τις σχετικές ποσότητες IL-1β και IL-1Ra. Η ισορροπία μεταξύ IL-1β και IL-1Ra μεταβάλλεται σε αυτοφλεγμονώδεις νόσους, με κυρίαρχη τη δραστηριότητας της IL-1 έναντι της αναστολής της, εξηγώντας έτσι την αποτελεσματική απόκριση της θεραπείας αποκλεισμού της IL-1 σε πολλές από αυτές τις διαταραχές [4].

1.1.1 Παθογένεια των αυτοφλεγμονωδών νοσημάτων

Το ανθρώπινο σώμα έχει εξελίξει διάφορες προσαρμοστικές αποκρίσεις που προστατεύουν από τη βλάβη των κυττάρων και των ιστών που προκαλούνται από φυσικούς, χημικούς και βιολογικούς παράγοντες. Τέτοιοι παράγοντες περιλαμβάνουν μόρια που απελευθερώνονται από κατεστραμμένους ιστούς, μεταβολίτες και μολύνσεις (βακτήρια, ιούς, παράσιτα). Η έμφυτη ανοσολογική απόκριση έχει εξελιχθεί για να προστατεύει τους οργανισμούς από αυτούς τους εξωγενείς και ενδογενείς κινδύνους που απειλούν, βλάπτουν ή καταστρέφουν τον οργανισμό. Ο στόχος της ανοσολογικής απόκρισης είναι να εξαλείψει ή να απομονώσει την πηγή της διαταραχής και να αποκαταστήσει τη λειτουργία των ιστών και την ομοιόσταση. Το έμφυτο ανοσοποιητικό σύστημα χρησιμοποιεί περιορισμένο αριθμό υποδοχέων οι οποίοι αναγνωρίζουν τις εξωγενείς και ενδογενείς μοριακές μονάδες. Όταν η οδός ενεργοποιείται οδηγεί στην παραγωγή προφλεγμονωδών κυτοκινών και χημειοκινών για τη στρατολόγηση των ανοσοκυττάρων στους

ιστούς και στον συντονισμό της άμυνας των ιστών. Για την αποφυγή ακατάλληλων και επιζήμιων αντιδράσεων, υπάρχουν πολλοί ρυθμιστικοί μηχανισμοί. Οι αυτοφλεγμονώδεις ασθένειες προκαλούνται κυρίως από ανωμαλίες που μπορεί να εμφανιστούν σε κάθε στάδιο της ανοσολογικής απόκρισης και της ρύθμισής της. Ορισμένα SAIDs οφείλονται σε υπερβολική ανίχνευση εξωκυτταρικών και ενδοκυτταρικών σημάτων κινδύνου [5].

Αρχικά, η φλεγμονή, μια προσαρμοστική απόκριση στον κυτταρικό τραυματισμό, δημιουργεί μοριακά μοτίβα που σχετίζονται με βλάβη (DAMPs) και μοριακά μοτίβα που σχετίζονται με παθογόνα (PAMPs), τα οποία στη συνέχεια αναγνωρίζονται από υποδοχείς αναγνώρισης προτύπων (PRRs) που εκφράζονται κυρίως από τα έμφυτα ανοσοκύτταρα. Τα PRR περιλαμβάνουν υποδοχείς τύπου Toll (TLRs), υποδοχείς τύπου Nod (NLRs), υποδοχείς λεκτίνης τύπου C (CLRs) και υποδοχείς τύπου RIG-I (RLRs) που αναγνωρίζουν τα DAMPs και PAMPs για την έναρξη ανοσολογικών αποκρίσεων. Αυτοί οι υποδοχείς ονομάζονται επίσης έμφυτοι ανοσοϋποδοχείς (Εικόνα 1). Τα SAIDs συνδέονται στενά με τη απορρύθμιση αυτών των αλληλεπιδράσεων που περιέχουν PRR, οι οποίες περιλαμβάνουν φλεγμονοσώματα, σιναλοσώματα που ενεργοποιούν τον πυρηνικό παράγοντα (NF)-κΒ, και ιντερφερόνη τύπου Ι. Η δυσλειτουργία αυτών των αλληλεπιδράσεων έχει ως αποτέλεσμα φλεγμονώδεις παθήσεις, ιντερφερονοπάθειες και αυτοφλεγμονώδη σύνδρομα που σχετίζονται με το πρωτεάσωμα (PRAAS). Αυτό εξηγεί την παθογένεση των αυτοφλεγμονωδών νόσων που περιλαμβάνουν επαναλαμβανόμενες φλεγμονώδεις εξάρσεις απουσία αυτοαντισωμάτων ή ειδικών για αντιγόνο Τ λεμφοκυττάρων. Η γνώση των μοριακών μηχανισμών που βρίσκεται κάτω από τις λειτουργίες αυτών των έμφυτων υποδοχέων του ανοσοποιητικού είναι χρήσιμη για τη θεραπεία και τη διαχείριση των ατόμων με SAIDs [6].



Εικόνα 1: Υποδοχείς αναγνώρισης προτύπων σε έμφυτα κύτταρα του ανοσοποιητικού. Οι κόκκινοι χαρακτήρες υποδεικνύουν DAMP και PAMP. Οι κόκκινοι χαρακτήρες υποδεικνύουν ταξινομημένες αυτοφλεγμονώδεις ασθένειες [6].

1.2 Περιοδικό Εμπύρετο Σύνδρομο Σχετιζόμενο με τον Υποδοχέα TNF

Έχει περάσει σχεδόν μια δεκαετία από τότε που η έννοια της αυτοφλεγμονής προωθήθηκε για να περιγράψει καταστάσεις που χαρακτηρίζονται από φαινομενικά απρόκλητη φλεγμονή, χωρίς συσχέτιση με αυτοαντισώματα υψηλού τίτλου ή τα ειδικά για αντιγόνο T-κύτταρα που βρίσκονται στα γνωστά αυτοάνοσα νοσήματα. Αυτή η ιδέα εισήχθη μετά την αναγνώριση μεταλλάξεων στον υποδοχέα του παράγοντα νέκρωσης όγκου p55, TNFR1, σε ασθενείς με κυρίαρχο κληρονομικό σύνδρομο συστημικής φλεγμονής TRAPS. Οι συστημικές αυτοφλεγμονώδεις νόσοι προκαλούνται από ανώμαλη ενεργοποίηση του έμφυτου ανοσοποιητικού συστήματος και ορισμένα έχουν ως αποτέλεσμα επαναλαμβανόμενα επεισόδια πυρετού, εντοπισμένης οροίτιδας και δερματικών εξανθημάτων. Δύο κληρονομικοί υποτροπιάζοντες πυρετοί, ο Οικογενής Μεσογειακός Πυρετός και το Περιοδικό Σύνδρομο TRAPS ήταν πρωτότυπα για αυτή τη διαγνωστική κατηγορία. Τελευταία αυτή η έννοια επεκτάθηκε για να συμπεριλάβει μια σειρά από μεντελικές διαταραχές, συμπεριλαμβανομένων άλλων κληρονομικών επαναλαμβανόμενων πυρετών, καθώς και πιο κοινές ασθένειες με σύνθετους τρόπους κληρονομικότητας, όπως η νόσος Behcet και η νόσος του Crohn [7].

Συγκεκριμένα, το σύνδρομο TRAPS με τους υποδοχείς του παράγοντα νέκρωσης όγκων (Tumor Necrosis Factor, TNF) είναι μια αυτοσωμική επικρατούσα νόσος που οφείλεται σε μετάλλαξη του μέλους της υπεροικογένειας των υποδοχέων TNF (TNFRSF1A). Είναι μια πρόσφατα αναγνωρισμένη αυτοφλεγμονώδης νόσος όπου περιγράφηκε σε μια μεγάλη σκωτσέζικη/ιρλανδική οικογένεια και για αυτό πήρε και την ονομασία Οικογενής Ιρλανδικός Πυρετός. Ωστόσο, το σύνδρομο μπορεί να επηρεάσει όλες τις εθνότητες^[6]. Αποτελεί μια διαταραχή που στην πραγματικότητα περιλαμβάνει αρκετές παλαιότερες διαγνώσεις που παγιώθηκαν με πρόσφατη την αναγνώριση ότι όλες προκαλούνται από μεταλλάξεις στην εξωκυτταρική περιοχή του υποδοχέα TNF 55 kDa (TNFRSF1A, επίσης γνωστή ως TNFR1, CD120a και p55). Το TRAPS είναι η πρώτη και, μέχρι σήμερα, μοναδική ασθένεια που είναι γνωστό ότι προκαλείται από δομικές μεταλλάξεις σε έναν υποδοχέα που έχει αναγνωριστεί ότι παίζει σημαντικό ρόλο στη φλεγμονή και την ανοσία [8].

Τα κλινικά χαρακτηριστικά του συνδρόμου περιλαμβάνουν πυρετό, αρθραλγία, μυαλγία, μεταναστευτικό εξάνθημα, κοιλιακό άλγος, πλευρίτιδα, επιπεφυκίτιδα, περικογχικό οίδημα και στοματικά έλκη. Τυπικά το TRAPS χαρακτηρίζεται από παρατεταμένες προσβολές που διαρκούν έως και 3-4 εβδομάδες, ωστόσο, μερικές φορές τα φλεγμονώδη επεισόδια μπορεί να είναι μεγαλύτερα ή μπορεί να εξαφανιστούν σε λιγότερο από μία εβδομάδα, ανάλογα με τον γονότυπο. Ιδιαιτερότητες του TRAPS είναι το περικογχικό οίδημα και το εξάνθημα ψευδοκυτταρίτιδας. Η μέση ηλικία έναρξης είναι περίπου τα τέσσερα έτη με μέση διαγνωστική καθυστέρηση μία δεκαετία, με το ένα πέμπτο των ασθενών έχουν εμφανίσει τα πρώτα συμπτώματα κατά την ενήλικη ζωή. Οι δερματικές εκδηλώσεις περιλαμβάνουν μεταναστευτικό ερυθρηματώδες εξάνθημα ή μπαλώματα της ωχράς κηλίδας, εκχυμώσεις, οίδηματώδεις δερματικές πλάκες, σερπιγγώδεις ή δακτυλιοειδείς βλάβες και περικογχικό οίδημα. Η διάγνωση απαιτεί την αναγνώριση μιας

μετάλλαξης στο TNFRSF1A, επομένως, για ασθενείς με κλινικά συμπτώματα που οδηγούν στην υποψία TRAPS, οι γενετικές εξετάσεις είναι απαραίτητες [9].

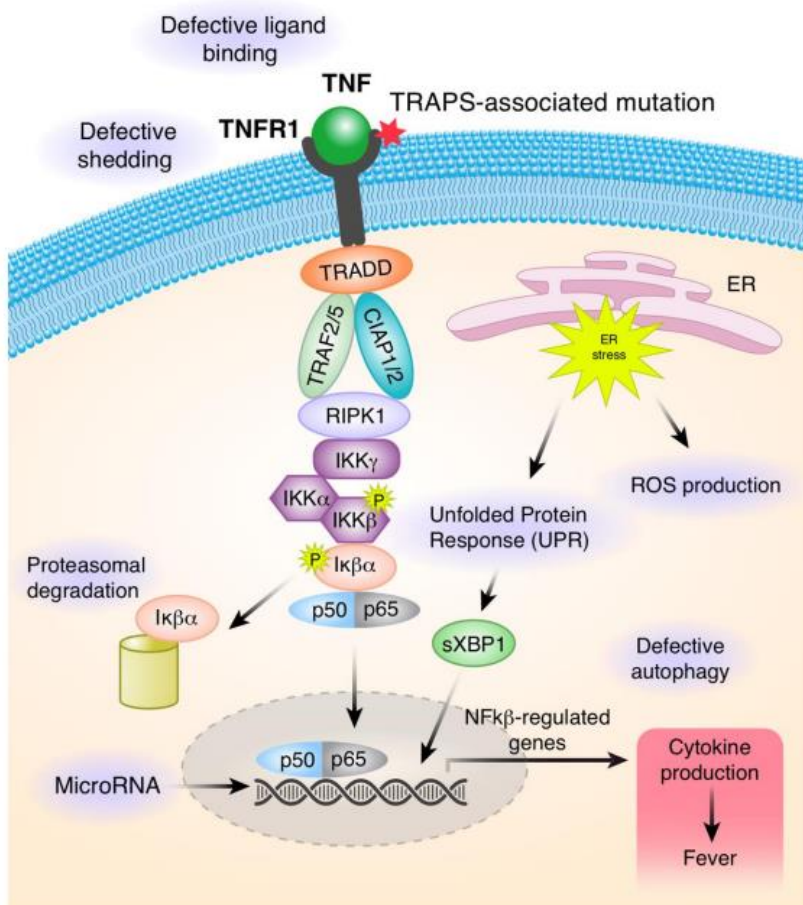
1.2.1 Γενετική των TRAPS

Το TRAPS (OMIM 142680) είναι μια αυτοσωμική κυρίαρχη νόσος που προκαλείται από συχνά λανθασμένες μεταλλάξεις στο γονίδιο TNFRSF1A, το οποίο αποτελείται από 10 εξόνια που κωδικοποιούν τον υποδοχέα p55 1A του TNF (TNFR1A) με την συντριπτική πλειονότητα των μεταλλάξεων να βρίσκεται στα εξόνια 3, 2, 4, και 6, και μπορούν να διακριθούν ως υψηλής ή χαμηλής διείσδυσης μεταλλάξεις. Οι πρώτες εντοπίζονται σε πλούσιες σε κυστεΐνη αμινοτελικές περιοχές, θεμελιώδεις για τη συναρμολόγηση της τρισδιάστατης δομής του υποδοχέα και χαρακτηρίζονται από την πρόωμη έναρξη της νόσου καθώς και με πιο σοβαρές κλινικές εκδηλώσεις. Οι μεταλλάξεις χαμηλής διείσδυσης, όπως οι R92Q και P46L, τείνουν να σχετίζονται με την εμφάνιση της νόσου στην ενήλικη ζωή και με λιγότερο έντονα ή άτυπα κλινικά χαρακτηριστικά. Αν και η βιολογική αλλοίωση περιλαμβάνει τον υποδοχέα TNF, η παθογένεση του TRAPS φαίνεται επίσης να σχετίζεται με απορρύθμιση στην έκκριση IL-1 και IL-6, καθώς και με οξειδωτική βλάβη που σχετίζεται με τη μιτοχονδριακή παραγωγή ελεύθερων ριζών [9,10].

Αρχικά, ο παράγοντας νέκρωσης όγκου (TNF) είναι μια πλειοτροπική κυτοκίνη που μεσολαβεί σε ένα ευρύ φάσμα κυτταρικών δραστηριοτήτων κυρίως μέσω της αλληλεπίδρασής του με τον υποδοχέα TNF-1 (TNFR1, TNFRSF1A, CD120a, p55/p60). Η ενεργοποίηση του καταρράκτη σηματοδότησης TNF οδηγεί σε μια σειρά αποκρίσεων, συμπεριλαμβανομένης της απόπτωσης, της φλεγμονής και της ρύθμισης της ανοσολογικής απόκρισης. Το αιτιολογικό γονιδιακό προϊόν του συνδρόμου, TNFR1/p55/CD120, είναι το μέλος 1A της υπεροικογένειας των υποδοχέων TNF που περιλαμβάνει 29 πρωτεΐνες στον άνθρωπο, οι οποίες αλληλεπιδρούν με μία ή περισσότερες κυτοκίνες της οικογένειας TNF [10].

Αυτή η μονογονική κυρίαρχη ασθένεια προκαλείται από μια εσφαλμένη μετάλλαξη του TNFRSF1A στο χρωμόσωμα 12 που κωδικοποιεί τον υποδοχέα 55-KF του TNF-α. Μέχρι στιγμής, έχουν αναφερθεί 180 παραλλαγές του γονιδίου TNFRSF1A. Η πλειονότητα των παθογόνων μεταλλάξεων δρα μεταβάλλοντας τους δισουλφιδικούς δεσμούς κυστεΐνης-κυστεΐνης στην εξωκυτταρική περιοχή (π.χ. T50 M και C33Y) του TNFRSF1A και έτσι πιστεύεται ότι είναι σημαντικοί για την παθογένεση της νόσου. Περισσότεροι από το ένα τρίτο των ασθενών φέρουν μεταλλάξεις R92Q και P46L, αν και αυτές οι μεταλλάξεις χαμηλής διείσδυσης έχουν ασαφή ρόλο και συχνά οδηγούν σε ηπιότερη νόσο. Παρά τις πολλές εικασίες σχετικά με την ελαττωματική αποβολή υποδοχέα και την απόπτωση, το πιο αναγνωρισμένο αποτέλεσμα της μετάλλαξης TNFRSF1 είναι η αλλοιωμένη διακίνηση του υποδοχέα TNF τύπου I (TNFR1) που συσσωρεύεται στο ενδοπλασματικό δίκτυο. Στο TRAPS, η κακή αναδίπλωση του μεταλλαγμένου TNFRSF1A οδηγεί σε συσσώρευση της πρωτεΐνης στο ενδοπλασματικό δίκτυο (ER), η οποία προκαλεί στρες ER και αυξημένη παραγωγή μιτοχονδριακών αντιδραστικών ειδών οξυγόνου (ROS) που ενισχύουν το μονοπάτι NF-kB και την ενεργοποιημένη από μιτογόνο πρωτεϊνική κινάση (MAPK).

Αυτό με τη σειρά του ενεργοποιεί τα φλεγμονώδη. Η υπερφόρτωση του TNFR1 σε ασθενείς με TRAPS ενισχύει την παραγωγή ROS και μπορεί έμμεσα να οδηγήσει σε ενεργοποίηση του φλεγμονώδους NLRP3. Μέσω της αυτοφαγίας οι ξεδιπλωμένες και συσσωρευμένες πρωτεΐνες απομακρύνονται από τα κύτταρα (Εικόνα 2) [8,11].



Εικόνα 2: Σύνοψη των προτεινόμενων παθογόνων μηχανισμών σε TRAPS. Η σύνδεση του TNFα στον TNFR1 οδηγεί στη συναρμολόγηση της οδού σηματοδότησης που τελικά ρυθμίζει προς τα πάνω τη γονιδιακή έκφραση πολλών προφλεγμονωδών κυτοκινών. Υπάρχουν πολλοί μηχανισμοί που συμβάλλουν στην παθογένεση του TRAPS [12].

Πιο συγκεκριμένα, ο TNFR1 εκφράζεται στην κυτταρική επιφάνεια στους περισσότερους ιστούς. Είναι ένα μεγάλο πολυπεπίτιδιο με μια εξωκυτταρική περιοχή, που αποτελείται από 4 πλούσιες σε κυστεΐνη περιοχές (CRD), μια διαμεμβρανική περιοχή και μια περιοχή ενδοκυτταρικού θανάτου (DD). Κάθε CRD περιλαμβάνει 3 ζεύγη κατάλοιπων κυστεΐνης που σχηματίζουν ενδομοριακούς δισουλφιδικούς δεσμούς σημαντικούς για τη διατήρηση της τρισδιάστατης διαμόρφωσης του εξωκυτταρικού τμήματος του υποδοχέα. Έτσι, μια υποκατάσταση σε ένα κατάλοιπο κυστεΐνης έχει τη δυνατότητα να διαταράξει τον σχετικό δισουλφιδικό δεσμό και να οδηγήσει σε σημαντική δομική διαταραχή. Η πρώτη CRD, που μερικές φορές ονομάζεται περιοχή σύνδεσης προ-συνδέτη

(PLAD), επιτρέπει αλληλεπιδράσεις ανεξάρτητες από TNF των μορίων TNFR1, ρυθμίζοντας την κυτταρική απόκριση στον TNF. Η δεύτερη CRD περιέχει τις θέσεις δέσμευσης προσδέτη για TNF. Οι περισσότερες μεταλλάξεις TNFR1 που σχετίζονται με την νόσο είναι λανθασμένες αλλαγές νουκλεοτιδίων που εντοπίζονται στους δύο πρώτους πλούσιους σε κυστεΐνη τομείς (CRD) της εξωκυτταρικής περιοχής του TNFR1. Η δέσμευση τριμερικών συμπλεγμάτων TNF προκαλεί τριμερισμό των μορίων TNFR1 και επιτρέπει στο DD να στρατολογεί μόρια σηματοδότησης. Ανάλογα με τον τύπο του κυττάρου, οι πρωτεΐνες προσαρμογής που εμπλέκονται από το DD μπορεί να οδηγήσουν σε ενεργοποίηση προ-αποπτωτικών, πυρηνικών οδών κB (NF-κB) ή C-Jun N-τερματικής κινάσης (JNK). Μια σειρά πολύπλοκων μηχανισμών ανάδρασης μπορεί να ενισχύσει ή να μειώσει αυτά τα σήματα. Ένας τρόπος περιορισμού της σηματοδότησης του TNFR1 είναι η πρωτεολυτική διάσπαση της εξωκυτταρικής περιοχής του από μεταλλοπρωτεϊνάσες, απελευθερώνοντας ένα διαλυτό θραύσμα του TNFR1 (sTNFR1), το οποίο δρα ως υποδοχέας δόλωμα και εξουδετερώνει τον TNF ορού ^[12].

1.2.2 Θεραπεία των TRAPS

Η θεραπεία του TRAPS βασίζεται κυρίως στον έλεγχο της κλινικής και υποκλινικής φλεγμονής προκειμένου να εξασφαλιστεί καλή ποιότητα ζωής και να αποφευχθεί η αντιδραστική αμυλοείδωση (AA). Από θεραπευτική άποψη, τα κορτικοστεροειδή μπορεί να αποδειχθούν χρήσιμα κατά τις οξείες φάσεις, αν και δεν μειώνουν τη συχνότητα των προσβολών και ούτε προλαμβάνουν την αμυλοείδωση. Η χορήγηση τους για μεγάλες χρονικές περιόδους μπορεί να προκαλέσουν σοβαρές συστηματικές παρενέργειες. Η κολχικίνη και οι ανοσοτροποποιητικοί ή ανοσοκατασταλτικοί παράγοντες έχουν επίσης αποδειχθεί ότι έχουν πολύ μικρή αποτελεσματικότητα στο TRAPS ^[13].

Λόγω του γενετικού ελαττώματος στην προέλευση της παθολογίας, ήταν σαφές ότι η χρήση αντι-TNF παραγόντων θα μπορούσε να έχει σημαντική επίδραση σε αυτούς τους ασθενείς. Στην πραγματικότητα, η Etanercept έχει αποδειχθεί χρήσιμη στη μείωση της έντασης και της διάρκειας των οξέων προσβολών, αν και σε ορισμένες περιπτώσεις χάνει σταδιακά την αποτελεσματικότητά της. Αντίθετα, το Infliximab και το Adalimumab, παραδόξως, μπορούν να προκαλέσουν τυπικές οξείες φλεγμονώδεις προσβολές της νόσου. Όπως και σε άλλα μονογονιδιακά, στα SAIDs η απόφραξη της IL-1 έχει αποδειχθεί ότι είναι αποτελεσματική για την πρόληψη επιθέσεων TRAPS και AA. Η πιο αποτελεσματική θεραπεία για το TRAPS είναι τα φάρμακα αναστολέα IL-1 συγκεκριμένα τα Anakira και Canakinumab. Ο αναστολέας IL-1, το Canakinumab (Ilaris), ένα ανθρώπινο μονοκλωνικό αντίσωμα κατά της IL-1β, το οποίο είναι εγκεκριμένο από την FDA για τη θεραπεία πολλαπλών αυτοφλεγμονωδών νόσων (συμπεριλαμβανομένων: CAPS, σύνδρομο MKD/Υπερανοσοφαιρίνης D και στο ανθεκτικό στην κολχικίνη FMF), εγκρίθηκε από τον FDA για το TRAPS το 2013 ^[14].

Καθώς επίσης, ο ανταγωνιστής του υποδοχέα IL-6 Tocilizumab σε ασθενείς ανθεκτικούς στο Etanercept και στην Anakira φαίνεται να έχει καλά αποτελέσματα, υποδηλώνοντας έτσι τον

πιθανό ρόλο της IL-6 στην παθογένεση του TRAPS. Δεδομένα από πολλαπλές κλινικές μελέτες επιβεβαιώνουν την αποτελεσματικότητα και την ασφάλεια των βιολογικών φαρμάκων στον έλεγχο των συμπτωμάτων του TRAPS και ένα μακροπρόθεσμο όφελος για αυτούς τους ασθενείς. Οι ανεπιθύμητες ενέργειες είναι ήπιες και συνίστανται σε ήπιες λοιμώξεις. Συνολικά, βιολογικές στρατηγικές που μπλοκάρουν συγκεκριμένους ανοσομεσολαβητές, όπως ο TNF-α ή η IL-1β, είναι αποτελεσματικές στην καταστολή των συμπτωμάτων TRAPS, στην πρόληψη της αντιδραστικής αμυλοείδωσης και στην αναστολή της εξέλιξης σε βλάβες οργάνων ^[14].

1.3 Άλλα Αυτοφλεγμονώδη Νοσήματα

Εκτός από το Περιοδικό Εμπύρετο Σύνδρομο Σχετιζόμενο με τον Υποδοχέα TNF σύνδρομο, δύο άλλες αυτοφλεγμονώδεις νόσοι είναι το περιοδικό σύνδρομο που σχετίζεται με την κρυοπυρίνη (Cryopyrin-Associated Periodic Syndromes, CAPS) και το σύνδρομο που σχετίζεται με την ανεπάρκεια της μεβαλονικής κινάσης (Mevalonate Kinase Deficiency, MKD) επίσης γνωστό ως σύνδρομο υπερανοσοφαιριναιμίας D.

Το περιοδικό σύνδρομο CAPS, μεταδίδεται από αυτοσωματική επικρατούσα κληρονομικότητα που προκαλείται από μεταλλάξεις στο γονίδιο NLRP3 (ονομάζεται επίσης CIAS1 ή PYPAF) το οποίο κωδικοποιεί την κρυοπυρίνη, μια κρίσιμη πρωτεΐνη φλεγμονώδους που ενεργοποιεί άμεσα την IL-1β. Το σύνδρομο αυτό περιλαμβάνει μια σειρά τριών φαινοτύπων, με το λιγότερο σοβαρό να είναι το Οικογενές Κρυοεπαγόμενο Αυτοφλεγμονώδες Σύνδρομο (Familial Cold Autoinflammatory Syndrome, FCAS, OMIM 120100), ακολουθεί το σύνδρομο Muckle-Wells (Muckle-Wells Syndrome, MWS, OMIM 191900) και τέλος το χρόνιο βρεφικό νευρολογικό-δερματικό αρθρικό (Chronic Infantile Neurological Cutaneous and Articular, CINCA, OMIM 607115), γνωστό και ως Πολυσυστηματική Φλεγμονώδης Νόσος Νεογνικής Εμφάνισης (Neonatal-Onset Multisystem Inflammatory Disease, NOMID) το οποίο παρουσιάζει μια σαφώς πιο σοβαρή συνολική κλινική εικόνα ^[15]. Ιστορικά, τα FCAS, MWS και NOMID/CINCA έχουν περιγραφεί ως τρεις διαφορετικές ασθένειες. Ωστόσο, και στους τρεις φαινότυπους, έχουν εντοπιστεί μεταλλάξεις στο γονίδιο NLRP3 και αυξημένη απέκκριση IL-1β ^[16]. Σήμερα, αναγνωρίζεται ότι αντιπροσωπεύουν μια ομάδα ασθενειών που ονομάζονται CAPS και θεωρείται ως ένα συνεχές φάσμα ασθενειών. Αυτές οι μεταλλάξεις προκαλούν ανισορροπία στην παραγωγή IL-1β, οδηγώντας σε κρίσεις πυρετού που σχετίζονται με άλλα πολλαπλά φλεγμονώδη συμπτώματα.

Η επικράτηση της CAPS εκτιμάται ότι είναι 1–10 περιπτώσεις ανά εκατομμύριο. Οι Καυκάσιοι φαίνεται να επηρεάζονται περισσότερο από άλλες φυλές και δεν υπάρχει υπεροχή ανδρών/γυναικών. Στις περισσότερες περιπτώσεις, το CAPS γίνεται εμφανές κατά την πρώιμη παιδική ηλικία. Η διάγνωση συχνά καθυστερεί και απαιτεί ενδελεχή ανασκόπηση των κλινικών συμπτωμάτων. Οι περισσότεροι ασθενείς με CAPS αναφέρουν μυαλγία, αρθραλγία, πονοκέφαλο και κόπωση, αν και αυτά τα συμπτώματα είναι συχνά δύσκολο να ποσοτικοποιηθούν αντικειμενικά. Η επιπεφυκίτιδα και η κερατίτιδα, αν και είναι λιγότερο διαδεδομένες από το

εξάνθημα, μπορούν να παρατηρηθούν σε ασθενείς και από τους τρεις υποτύπους. Η αξιοσημείωτη πρόοδος στην κατανόησή για τη γενετική και τη μοριακή οδό που είναι υπεύθυνη για τον κλινικό φαινότυπο του CAPS έχει οδηγήσει στην ανάπτυξη αποτελεσματικών θεραπειών. Έχει επίσης καταστεί σαφές ότι το φλεγμονώδες NLRP3 παίζει κρίσιμο ρόλο στην έμφυτη άμυνα του ανοσοποιητικού και ως εκ τούτου έχει ευρύτερες επιπτώσεις για άλλες καταστάσεις φλεγμονώδους νόσου ^[17].

Οι τρεις φαινότυποι του CAPS προκαλούνται από κυρίαρχα κληρονομικές ή de novo μεταλλάξεις κέρδους λειτουργίας στο γονίδιο NLRP3 (επίσης γνωστό ως CIAS1) που βρίσκεται στο χρωμόσωμα 1q44, με μεταβλητή διείσδυση. Το γονίδιο NLRP3 κωδικοποιεί την πρωτεΐνη NALP3 (κρυοπυρίνη), η οποία είναι μέλος της οικογένειας των υποδοχέων που μοιάζουν με τον τομέα ολιγομερισμού που δεσμεύουν τα νουκλεοτίδια (NOD) (NLRs) ^[15]. Η κρυοπυρίνη (πρωτεΐνη NALP3) παίζει κεντρικό ρόλο στην αναγνώριση των σημάτων κινδύνου και στον επακόλουθο ανοσολογικό καταρράκτη. Η αύξηση της λειτουργίας της κρυοπυρίνης περιλαμβάνει αυξημένη φλεγμονώδη δραστηριότητα, ένα από τα βασικά συστατικά του έμφυτου ανοσοποιητικού συστήματος (Εικόνα 3). Οι NLR συσσωματώνονται σε ένα σύμπλεγμα πολλαπλών πρωτεϊνών, το λεγόμενο φλεγμονώδες. Όλα τα NLR περιέχουν έναν τομέα NACHT που τους επιτρέπει να συγκεντρώνονται και να ολιγομερίζονται. Οι μεταλλάξεις στην περιοχή NACHT του NALP3 μπορεί να οδηγήσουν σε αυξημένη φλεγμονή. Για την παραγωγή της IL-1β απαιτούνται δύο σήματα. Η IL-1β είναι ισχυρή σε χαμηλές συγκεντρώσεις, ενώ ο φυσικός της αναστολέας, δηλαδή ο ανταγωνιστής του υποδοχέα IL-1 (IL-1RA), χρειάζεται πολύ υψηλές συγκεντρώσεις για να είναι αποτελεσματικός. Σε ασθενείς με CAPS, το φλεγμονώδες σώμα ενεργοποιείται ακόμη και απουσία χρόνιας λοίμωξης. Η IL-1 επάγει έναν φαύλο κύκλο παραγωγής IL-1β με αυτοκρινή τρόπο και οι συγκεντρώσεις IL-1Ra που επιτυγχάνονται στο CAPS δεν επαρκεί για την αναστολή της IL-1β. Ωστόσο, αυτός ο κύκλος παραγωγής IL-1β μπορεί να διακοπεί από το Anakinra, το οποίο μπλοκάρει τόσο την IL-1α όσο και την IL-1β. Αυτές οι διαφορετικές μεταλλάξεις εμφανίζουν ισχυρή συσχέτιση γονότυπου-φαινοτύπου, αν και μια συγκεκριμένη μετάλλαξη μπορεί να σχετίζεται με διαφορετικούς φαινότυπους μεταβλητής σοβαρότητας. Λίγοι ασθενείς με πειστικό προφίλ CAPS δεν έχουν δείξει μετάλλαξη. Ορισμένα από αυτά έχουν σωματικό μωσαϊκό, υποδηλώνοντας έναν ρόλο σωματικών μεταλλάξεων ^[17,18].

Χάρη στη σημαντική πρόοδο στην κατανόηση της γενετικής βάσης και των μηχανισμών της νόσου για αυτές τις διαταραχές, αναπτύσσονται επί του παρόντος νέες θεραπευτικές στρατηγικές που στοχεύουν στην οδό IL-1. Με βάση τους αιτιολογικούς μηχανισμούς που έχουν τις ρίζες τους στην υπερπαραγωγή της IL-1β, τα CAPS αντιμετωπίστηκαν με αντι-IL-1 παράγοντες. Το Kineret (Anakinra) της Biovitrum Pharmaceuticals, είναι ένα ανασυνδυασμένο, μη γλυκοζυλιωμένο ομόλογο του ανταγωνιστή του ανθρώπινου υποδοχέα IL-1 το οποίο αναστέλλει ανταγωνιστικά τη σύνδεση της IL-1α και της IL-1β στον υποδοχέα IL-1. Έχει χρησιμοποιηθεί εκτενώς σε ασθενείς με FCAS με εξαιρετικά κλινικά αποτελέσματα. Ωστόσο, δεν είναι επί του παρόντος εγκεκριμένο από τον FDA για τη θεραπεία του FCAS ή οποιασδήποτε από τις ασθένειες CAPS. Η ασφάλεια και η ανεκτικότητα του Rilonacept έχει αποδειχθεί σε μια ομάδα παιδιατρικών και ενηλίκων

ασθενών με CAPS, ενώ η Canakinumab έχει αποδειχθεί ασφαλής και αποτελεσματική τόσο στον έλεγχο των κλινικών/εργαστηριακών δεικτών της δραστηριότητας της νόσου όσο και στον έλεγχο των επιπλοκών που σχετίζονται με την αμυλοείδωση. Το Ilaris (Canakinumab) της Novartis Pharmaceuticals, ένα μονοκλωνικό αντίσωμα στην IL-1, εγκρίθηκε από τον FDA το 2009 ως θεραπεία για παιδιά και ενήλικες με CAPS, συμπεριλαμβανομένων των FCAS και MWS. Επίσης, το Arcalyst (Rilonacept) της Regeneron Pharmaceuticals, ένας αναστολέας της IL-1, εγκρίθηκε από τον FDA το 2008 για τη θεραπεία του CAPS, συμπεριλαμβανομένων των FCAS και MWS, σε ενήλικες και παιδιά 12 ετών και άνω [19,20].

Αναφορικά με το σύνδρομο που σχετίζεται με την ανεπάρκεια μεβαλονικής κινάσης, επίσης γνωστό ως σύνδρομο υπερανοσοσφαιριναιμίας D (OMIM 260920) είναι μια αυτοσωματική υπολειπόμενη νόσος που προκαλείται από μεταλλάξεις στο γονίδιο MVK, που κωδικοποιεί το ένζυμο μεβαλονική κινάση, που εμπλέκεται στην εξαρτώμενη από το ATP φωσφορύλωση του μεβαλονικού οξέος σε 5-φωσφομεβαλονικό. Περιλαμβάνει ένα φάσμα σπάνιων αυτοφλεγμονωδών διαταραχών που προκαλούνται από ένα υπολειπόμενο κληρονομικό ελάττωμα στην ισοπρενοειδή βιοσυνθετική οδό. Εμφανίζεται κλινικά ως σύνδρομο υπερανοσοσφαιριναιμίας D (Hyperimmunoglobulin-D Syndrome HIDS), που είναι ο λιγότερος σοβαρός φαινότυπος και πιο συχνή μορφή, και ως μεβαλονική οξουρία (Mevalonic Aciduria, MVA), που είναι ο πιο σοβαρός φαινότυπος και σπάνια μορφή. Το όνομα HIDS (υπερ-IgD σύνδρομο περιοδικού πυρετού) προήλθε από το γεγονός ότι σε πολλούς ασθενείς βρέθηκαν αυξημένα τα επίπεδα της ανοσοσφαιρίνης D (IgD) στον ορό. Οι μεταλλάξεις αυτές όχι μόνο παρεμβαίνουν στη βιοσύνθεση της χοληστερόλης και του ισοπρενίου αλλά επίσης οδηγούν σε επεισόδια υπερφλεγμονής με αυξημένη έκκριση IL-1β μέσω ενεργοποίησης της κασπάσης-1. Το MKD χαρακτηρίζεται από αυτοφλεγμονώδεις εξάρσεις με πυρετό, κοιλιακό άλγος, βλεννοειδείς και δερματικές αλλοιώσεις και αρθραλγίες. Επιπλέον, οι πιο σοβαρά προσβεβλημένοι ασθενείς με MVA έχουν κλασικά αναπτυξιακή καθυστέρηση, σωματικούς δυσμορφισμούς, ψυχοκινητική καθυστέρηση, ηπατοσπληνομεγαλία και οφθαλμικές ανωμαλίες. Καθώς το MKD παραμένει μυστηριώδες τόσο στην ακριβή παθογένειά του όσο και ως προς το πλήρες εύρος του φαινοτυπικού φάσματος, επιδιώκουμε να αναθεωρήσουμε τις πιο σύγχρονες προοπτικές αυτής της ασθένειας. Η διάγνωση βασίζεται στην ανάλυση μετάλλαξης του γονιδίου MVK. Στην πραγματικότητα, το φάσμα είναι πιθανώς ακόμη ευρύτερο, από εμβρυονικά θανατηφόρο έως φαινομενικά υγιές. Η παθογένεση είναι μόνο εν μέρει κατανοητή και δεν υπάρχει διαθέσιμη θεραπεία βασισμένη σε στοιχεία, με την αναστολή της IL-1 να είναι συνήθως αποτελεσματική. Οι σοβαρές περιπτώσεις απαιτούν αλλογενή μεταμόσχευση βλαστοκυττάρων ενώ απαιτούνται στοχευμένες θεραπείες [21,22].

Οι ανοσολογικές εκδηλώσεις της MKD ταιριάζουν σε μεγάλο βαθμό με το πρότυπο μιας αυτοφλεγμονώδους διαταραχής που έχει οριστεί από τους Masters et al ως φαινομενικά απρόκλητες κρίσεις συστηματικής φλεγμονής απουσία αυτοαντιδραστικών λεμφοκυττάρων ή αυτοαντισωμάτων υψηλού τίτλου^[22]. Οι αυτοφλεγμονώδεις διαταραχές στα περιοδικά σύνδρομα που σχετίζονται με την κρυοπυρίνη και το FMF, όπου παραλλαγές σε συγκεκριμένα φλεγμονώδη συστατικά οδηγούν άμεσα σε υπερπαραγωγή κυτοκινών συμπεριλαμβανομένης της IL-1β. Οι

ασθενείς με MKD παρουσιάζουν παρομοίως περιοδικά επεισόδια συστηματικής φλεγμονής, τα οποία επίσης χαρακτηρίζονται από υπερβολική έκκριση φλεγμονωδών μεσολαβητών. Κατά τη διάρκεια των επεισοδίων, οι ασθενείς με MKD έχουν αυξημένα επίπεδα πρωτεϊνών οξείας φάσης, συμπεριλαμβανομένης της C-αντιδρώσας πρωτεΐνης και κυτοκινών όπως ο TNF-α, η IL-6 και η ιντερφερόνη-γ. Επιπλέον, τα μονοκύτταρα περιφερικού αίματος (PBMCs) από ασθενείς με MKD παράγουν αυξημένες ποσότητες IL-1β, IL-6 και TNF-α τόσο σε ηρεμία όσο και μετά από διέγερση με ενεργοποιητικούς συνδέτες. Κρίσιμης σημασίας, η φλεγμονώδης υπερανταπόκριση στο MKD φαίνεται να οφείλεται στην έλλειψη ισοπρενοειδών προϊόντων και όχι στη συσσώρευση μεβαλονικού οξέος. Η κλινική παρατήρηση ότι η θεραπεία με στατίνες σε ασθενείς με MKD μειώνει τόσο τα επίπεδα του μεβαλονικού οξέος όσο και την υπολειμματική παραγωγή ισοπρενοειδών, στην πραγματικότητα μπορεί να προκαλέσει έξαρση της νόσου. Περαιτέρω έρευνα έχει τονίσει συγκεκριμένα την ανάγκη για γερανυλίωση αντί για τα άλλα προϊόντα της μεβαλονικής οδού, όπως η βιοσύνθεση χοληστερόλης, στη μεσολάβηση της υπερέκκρισης της IL-1β. Αυτά τα συσσωρευμένα στοιχεία οδήγησαν στην πρόταση ότι η MKD αντιπροσωπεύει μια μεταβολική αυτοφλεγμονώδη νόσο, όπου η ελαττωματική ροή μέσω της οδού ισοπρενοειδών ενεργοποιεί το έμφυτο ανοσοποιητικό σύστημα για υπερανταπόκριση [23,24,25].

Η θεραπεία στρέφεται προς τα συγκεκριμένα συμπτώματα που είναι εμφανή σε κάθε άτομο. Διάφορα φάρμακα, συμπεριλαμβανομένης της παρακεταμόλης (ακεταμινοφενόνης) έχουν χρησιμοποιηθεί για τη θεραπεία προσβεβλημένων ατόμων. Τα κορτικοστεροειδή, όπως η πρεδνιζόνη, έχουν χρησιμοποιηθεί για τη θεραπεία ορισμένων ατόμων. Μη στεροειδή αντιφλεγμονώδη και παυσίπονα (αναλγητικά) έχουν επίσης δοκιμαστεί. Αυτά τα φάρμακα έχουν οδηγήσει σε πλήρη ή μερική ύφεση κατά τη διάρκεια ενός επεισοδίου πυρετού, αλλά η επιτυχία είναι ασυνεπής και προσωρινή. Γενικά θεωρούνται τα καλύτερα για άτομα με ήπια συμπτώματα και με μόνο λίγα επεισόδια. Πολλά άλλα φάρμακα έχουν χρησιμοποιηθεί σε άτομα με ανεπάρκεια μεβαλονικής κινάσης. Το 2016, ο Οργανισμός Τροφίμων και Φαρμάκων των ΗΠΑ (FDA) ενέκρινε το Canakinumab (Ilaris) για τη θεραπεία ατόμων με ανεπάρκεια μεβαλονικής κινάσης. Είναι μια πλήρως εξανθρωπισμένη πρωτεΐνη που αναστέλλει την κυτοκίνη IL-1β, η οποία συμβάλλει στα επεισόδια πυρετού. Ωστόσο, απαιτείται περισσότερη έρευνα για τον προσδιορισμό της μακροπρόθεσμης ασφάλειας και αποτελεσματικότητας του Canakinumab για άτομα με MKD [26,27]. Μερικά προσβεβλημένα άτομα έχουν λάβει θεραπεία με Anakinra το οποίο είναι ένας ανταγωνιστής του υποδοχέα της IL-1β που αναστέλλει τη δραστηριότητα της ενώ μερικοί άλλοι ασθενείς έλαβαν θεραπεία με φάρμακα που εμποδίζουν τη δραστηριότητα του παράγοντα νέκρωσης όγκου, μιας εξειδικευμένης πρωτεΐνης (κυτοκίνης). Όπως είναι τα Etanercept και Adalimumab αλλά απαιτείται περισσότερη έρευνα για τον προσδιορισμό της μακροπρόθεσμης ασφάλειας και αποτελεσματικότητας αυτών των θεραπειών για άτομα με MKD. Το Tocilizumab είναι ένα φάρμακο που καταστέλλει το ανοσοποιητικό σύστημα αναστέλλοντας μια διαφορετική κυτοκίνη που ονομάζεται IL-6, έχει δοκιμαστεί ως θεραπεία για άτομα με ανεπάρκεια μεβαλονικής κινάσης και έχει δείξει όφελος στη μείωση της συχνότητας και της διάρκειας των επεισοδίων πυρετού [27].

1.4 Βιοπληροφορική Προσέγγιση

Για εκατοντάδες χρόνια, οι άνθρωποι ερευνούσαν για αποτελεσματικά φάρμακα. Η διαδικασία ανάπτυξης φαρμάκων είναι χρονοβόρα και μπορεί χονδρικά να χωριστεί σε δύο φάσεις. Κατά την πρώτη φάση, γίνεται αναζήτηση του στόχου και ανακάλυψη της βιολογικής προέλευσης μιας συγκεκριμένης ασθένειας. Ο στόχος του φαρμάκου που θα αναπτυχθεί μπορεί να είναι μια πρωτεΐνη, γονίδιο, DNA ή RNA. Στη δεύτερη φάση επιλέγεται ή αναπτύσσεται το μόριο, φάρμακο. Με την ανάπτυξη των πειραματικών διαδικασιών στην αναλυτική και συνθετική χημεία, το αποτέλεσμα της θεραπείας άρχισε να αποδίδεται σε συγκεκριμένα μόρια τα οποία μπορούσαν να τροποποιηθούν και να τελειοποιηθούν για να ενισχύσουν τις δράσεις τους αλλά και να περιορίσουν τις ανεπιθύμητες παρενέργειες. Ωστόσο, στις περισσότερες περιπτώσεις, ο στόχος (π.χ. πρωτεΐνη) δεν ήταν γνωστή. Ακόμη και σήμερα, δεν γνωρίζουμε τις πρωτεΐνες-στόχους για πολλά φάρμακα στην αγορά. Η υπολογιστική πλευρά του πεδίου αντικατοπτρίζει την εξέλιξη στην πειραματική πλευρά. Μπορούν να διακριθούν τρία υποπεδία, ο έλεγχος φαρμάκων, η πρόσδεση και η βιοπληροφορική. Τα δύο πρώτα ανήκουν στο κομμάτι της υπολογιστικής χημείας όπου ο έλεγχος των φαρμάκων στοχεύει στην υπολογιστική σύγκριση μεταξύ φαρμάκων ή μορίων που μοιάζουν με άλλα φάρμακα χωρίς να απαιτείται γνώση της δομής της πρωτεΐνης-στόχου. Ενώ κατά την διαδικασία πρόσδεσης του φαρμάκου, η δομή της πρωτεΐνης πρέπει να είναι διαθέσιμη και έτσι γίνεται δυνατή η ανάπτυξη ή η κατάλληλη επιλογή του φαρμάκου από μια μεγάλη βάση ενώσεων [28].

Αντίθετα, η βιοπληροφορική ανάλυση μπορεί όχι μόνο να επιταχύνει την ταυτοποίηση του στόχου του φαρμάκου, τον έλεγχο και τη βελτίωση του υποψηφίου φαρμάκου, αλλά διευκολύνει και την αναγνώριση των παρενεργειών και την αντοχή στα φάρμακα. Δεδομένα υψηλής απόδοσης, όπως γονιδιωμικά, επιγενετικά, αρχιτεκτονική γονιδιώματος, μεταγραφικά, πρωτεομικά και ριβοσωμικά έχουν όλα συνεισφέρει σημαντικά στην ανακάλυψη φαρμάκων με βάση τον μηχανισμό και τον επαναπροσδιορισμό τους. Η συσσώρευση δομών πρωτεϊνών και RNA, καθώς και η ανάπτυξη μοντελοποίησης ομολογίας και προσομοίωσης πρωτεϊνικής δομής, σε συνδυασμό με μεγάλες βάσεις δεδομένων μικρών μορίων και μεταβολιτών, άνοιξαν το δρόμο για πιο ρεαλιστικά πειράματα σύνδεσης φαρμάκου-στόχου και πιο ενημερωτική εικόνα διαλογής. Μία από τις κύριες ωθήσεις των τρεχουσών προσεγγίσεων βιοπληροφορικής είναι η πρόβλεψη και η αναγνώριση βιολογικά ενεργών υποψηφίων και η αποθήκευση σχετικών πληροφοριών. Τα φάρμακα συνήθως αναπτύσσονται μόνο όταν έχει εντοπιστεί και μελετηθεί ο συγκεκριμένος στόχος για τις δράσεις αυτών των φαρμάκων. Ο αριθμός των πιθανών στόχων για τη διαδικασία ανακάλυψης φαρμάκων αυξάνεται εκθετικά. Επομένως, η βιοπληροφορική επιτρέπει την αναγνώριση και την ανάλυση ολόένα και περισσότερων βιολογικών στόχων φαρμάκων. Έτσι αναμένεται να αυξήσει σημαντικά το αριθμό των πιθανών φαρμάκων στις βάσεις δεδομένων των φαρμακευτικών εταιρειών [29].

Η Βιοπληροφορική παρέχει επίσης στρατηγικές και αλγόριθμους για την πρόβλεψη νέων στόχων φαρμάκων και την αποθήκευση και διαχείριση των διαθέσιμων πληροφοριών στόχων φαρμάκων. Μετά την ανακάλυψη των «δυσνητικών» στόχων φαρμάκων, υπάρχει μια ανεκτίμητη ανάγκη να δημιουργηθεί μια ισχυρή συσχέτιση μεταξύ ενός υποτιθέμενου στόχου και της ασθένειας ενδιαφέροντος. Η δημιουργία μιας τέτοιας βασικής ένωσης δικαιολογεί τη διαδικασία ανάπτυξης φαρμάκων. Αυτή η διαδικασία, γνωστή ως επικύρωση στόχου, είναι ένας τομέας όπου η βιοπληροφορική παίζει σημαντικό ρόλο. Η επικύρωση του στόχου του φαρμάκου βοηθά στον μετριασμό της πιθανότητας αποτυχίας στις φάσεις κλινικών δοκιμών και έγκρισης [29].

Ωστόσο, στη διαδικασία ανάπτυξης φαρμάκων, μεγάλος αριθμός μορίων αποτυγχάνει σε μεταγενέστερο στάδιο λόγω ανησυχιών για την ασφάλεια ή την αποτελεσματικότητα που γενικά διερευνώνται με τη διεξαγωγή διαφόρων πειραμάτων σε ζώα, αλλά δυστυχώς, τα πειράματα σε ζώα συνδέονται με πολλούς περιορισμούς. Έτσι, υπάρχει ένα αυξανόμενο ενδιαφέρον για την ανάπτυξη εναλλακτικών προσεγγίσεων που απαιτούν λιγότερο αριθμό ζώων. Το docking είναι μια εναλλακτική προσέγγιση για τον έλεγχο των ενώσεων πριν από την πραγματική δοκιμή σε ζώα. Αποτελεί την καλύτερη επιλογή για την πρόβλεψη των ενεργειακά ευνοϊκών διαμορφώσεων σύνδεσης των προσδεμάτων στην κοιλότητα της ενεργής θέσης συγκεκριμένου υποδοχέα και έχει κερδίσει το ενδιαφέρον στην επιστημονική κοινότητα για εξοικονόμηση χρόνου και χρημάτων που εμπλέκονται στη διαδικασία ανάπτυξης φαρμάκων. Η χρήση τέτοιων τεχνικών αυξάνεται μέρα με τη μέρα μεταξύ των ερευνητών λόγω της διαθεσιμότητας φιλικών προς τον χρήστη λογισμικού. Λόγω αυτού, ο αριθμός των εγγράφων που σχετίζονται με τη σύνδεση έχει αυξηθεί δραματικά τις τελευταίες δεκαετίες. Επί του παρόντος, διατίθενται διάφορα λογισμικά σύνδεσης για την πρόβλεψη της θέσης δέσμευσης προσδέτη στον συγκεκριμένο υποδοχέα του, αλλά υπάρχει έλλειψη γνώσης και εμπιστοσύνης μεταξύ των επιστημόνων σχετικά με την ακρίβεια αυτών των λογισμικών. Υπάρχουν πολλές ανησυχίες που πρέπει να ληφθούν υπόψη, κατά τη διεξαγωγή οποιασδήποτε μελέτης σύνδεσης, όπως η επιλογή των σωστών σημείων δέσμευσης της πρωτεΐνης στόχου, η επιλογή της σωστής στάσης πρόσδεσης, η έλλειψη σαφήνειας σχετικά με το εάν η ένωση είναι ανταγωνιστής ή αγωνιστής, δύσκολη λάβετε τους καλύτερους αλγόριθμους αντιστοίχισης και τα σχήματα βαθμολόγησης. Πολλές ενώσεις παρουσιάζουν υψηλή βαθμολογία υπόβαθρου, αλλά δυστυχώς αποτυγχάνουν στις προκλινικές μελέτες. Αυτά τα ζητήματα θα πρέπει να είναι μια ευκαιρία γιατί για προσοχή και ανησυχία πριν πραγματοποιήσετε τη σύνδεση [30].

1.4.1 Μοριακή πρόσδεση

Ο τομέας της μοριακής πρόσδεσης έχει αναδυθεί κατά τη διάρκεια των τελευταίων τριών δεκαετιών καθοδηγούμενος από τις ανάγκες της δομικής μοριακής βιολογίας και της ανακάλυψης φαρμάκων με βάση τη δομή. Έχει διευκολυνθεί πολύ από τη δραματική αύξηση της διαθεσιμότητας και της ισχύος των υπολογιστών καθώς και την αυξανόμενη ευκολία πρόσβασης σε βάσεις δεδομένων μικρών μορίων και πρωτεϊνών. Η μελέτη πρόσδεσης είναι η εφαρμογή μαθηματικών, βιολογικών μοντέλων και μοντέλων που βασίζονται σε υπολογιστή για την πρόβλεψη της συγγένειας δέσμευσης μικρών μορίων προς συγκεκριμένο υποδοχέα. Αυτή η

έρευνα ενσωματώνει τις εξελίξεις στη μοριακή βιολογία, τη βιοτεχνολογία, τη βιοπληροφορική, τα μαθηματικά, τη χημεία με τη μοντελοποίηση και την επιστήμη των υπολογιστών, προκειμένου να αυξηθεί η προγνωστική ισχύς των λογισμικών σύνδεσης. Όπως όλοι γνωρίζουμε, η διαδικασία ανάπτυξης φαρμάκων είναι μακρά και δαπανηρή. Επιπλέον, η αυξανόμενη συχνότητα εμφάνισης της νόσου και η αντίσταση των μικροβίων στα εγκεκριμένα φάρμακα απαιτεί επιτάχυνση της διαδικασίας ανάπτυξης φαρμάκων. Για την εξοικονόμηση χρόνου και χρημάτων από τη διαδικασία ανακάλυψης και ανάπτυξης φαρμάκων, οι μελέτες σύνδεσης χρησιμοποιούνται ευρέως για την πρόβλεψη της αλληλεπίδρασης σχεδιασμένων ανάλογων προς συγκεκριμένο υποδοχέα [31,32].

Ο στόχος του αυτοματοποιημένου λογισμικού μοριακής πρόσδεσης είναι να κατανοήσει και να προβλέψει τη μοριακή αναγνώριση, τόσο δομικά, βρίσκοντας πιθανούς τρόπους δέσμευσης, όσο και ενεργειακά, προβλέποντας τη συγγένεια δέσμευσης. Η μοριακή πρόσδεση πραγματοποιείται συνήθως μεταξύ ενός μικρού μορίου και ενός μακρομορίου στόχου. Αυτό αναφέρεται συχνά ως σύνδεση προσδέτη-πρωτεΐνης, αλλά υπάρχει αυξανόμενο ενδιαφέρον για σύνδεση πρωτεΐνης-πρωτεΐνης. Μια ποικιλία προσεγγίσεων σύνδεσης είναι επί του παρόντος διαθέσιμες για την πρόβλεψη της αλληλεπίδρασης των μορίων έναντι συγκεκριμένων υποδοχέων, αλλά υπάρχει έλλειψη γνώσης και εμπιστοσύνης μεταξύ των ερευνητών. Δυστυχώς, μια μέθοδος που είναι κατάλληλη για ορισμένους τύπους πρόβλεψης μπορεί να μην λειτουργεί σωστά (ή να μην λειτουργεί καθόλου) για άλλους. Πολλές ενώσεις αποτυγχάνουν στην προσομοίωση μοριακής δυναμικής παρά το γεγονός ότι διαθέτουν υψηλή βαθμολογία εγγράφων. Επίσης, υπάρχει έλλειψη σαφήνειας σχετικά με το εάν η ένωση είναι αγωνιστής ή ανταγωνιστής. Εάν χρησιμοποιηθούν σωστά, τα εργαλεία σύνδεσης μπορούν να είναι πολύ αποτελεσματικά στην αξιολόγηση της ασφάλειας και της αποτελεσματικότητας των μορίων. Επομένως, για να διασφαλιστεί η ακριβής και αποτελεσματική εφαρμογή των μοντέλων σύνδεσης, είναι απαραίτητο να κατανοήσουμε τα δυνατά σημεία, τους περιορισμούς, το εύρος εφαρμογής και την ερμηνεία της μεθόδου [32].

Επομένως, η μοριακή πρόσδεση (Molecular Docking) είναι ένα είδος βιοπληροφορικής μοντελοποίησης που περιλαμβάνει την αλληλεπίδραση δύο ή περισσότερων μορίων για να δώσει ένα σταθερό προϊόν προσθήκης. Ανάλογα με τις ιδιότητες σύνδεσης του προσδέτη και του στόχου, προβλέπεται η τρισδιάστατη δομή οποιουδήποτε συμπλόκου. Η μοριακή πρόσδεση δημιουργεί διαφορετικές πιθανές δομές προσθήκης που ταξινομούνται και ομαδοποιούνται χρησιμοποιώντας μια λειτουργία βαθμολόγησης στο λογισμικό όπου προβλέπεται τη συνολική ενέργεια του συστήματος. Παρά όλες τις πιθανές προσεγγίσεις, η χημεία των προσδεμάτων (αυτομερισμός και ιονισμός), η ευελιξία του υποδοχέα (μονή διαμόρφωση του άκαμπτου υποδοχέα) και η λειτουργία βαθμολόγησης (διαφοροποίηση της πραγματικής κατάστασης σύνδεσης) εξακολουθούν να αποτελούν πρόκληση. Ενδιαφέρον παρουσιάζει η μεγάλη ποικιλία χρήσεων και εφαρμογών στην ανακάλυψη φαρμάκων, συμπεριλαμβανομένων μελετών δομής-δραστηριότητας, βελτιστοποίησης ηλεκτροδίων, εύρεση πιθανών απαγωγών μέσω εικονικής διαλογής, παροχή δεσμευτικών υποθέσεων για τη διευκόλυνση προβλέψεων για μελέτες μεταλλαξιογένεσης, υποβοήθηση

κρυσταλλογραφίας ακτίνων X στην τοποθέτηση υποστρωμάτων και αναστολείς της πυκνότητας ηλεκτρονίων, μελέτες χημικών μηχανισμών και συνδυαστικό σχεδιασμό βιβλιοθήκης [33].

Από την πρώτη εμφάνισή του στα μέσα της δεκαετίας του 1970, το docking έχει αποδειχθεί σημαντικό εργαλείο για την κατανόηση του τρόπου με τον οποίο οι χημικές ενώσεις αλληλεπιδρούν με τους μοριακούς στόχους τους και για την ανακάλυψη και ανάπτυξη φαρμάκων. Στην πραγματικότητα, ο αριθμός των μελετών που αναφέρουν την χρήση μοριακής πρόσδεσης για τον εντοπισμό δομικών καθοριστικών παραγόντων που είναι απαραίτητοι για την αποτελεσματική δέσμευση συνδέτη-υποδοχέα και την ανάπτυξη πιο ακριβών μεθόδων σύνδεσης, έχει αυξηθεί σημαντικά από την πρώτη εμφάνιση. Από τότε, η μοριακή πρόσδεση υπέστη δραματικές βελτιώσεις, όπως για παράδειγμα χρησιμοποιώντας ευέλικτους αλγόριθμους στους υπολογισμούς. Επιπλέον, άρχισε να χρησιμοποιείται για το σχεδιασμό και τη βελτιστοποίηση ενώσεων με θεραπευτικό ενδιαφέρον [32].

Η μοριακή πρόσδεση αποτελεί χρήσιμο εργαλείο για την κατανόηση των βιομοριακών αλληλεπιδράσεων του φαρμάκου για τον ορθολογικό σχεδιασμό και την ανακάλυψη του, καθώς και στη μηχανιστική μελέτη με την τοποθέτηση ενός μορίου (προσδέτη) στην προτιμώμενη θέση δέσμευσης της ειδικής περιοχής στόχου του DNA/πρωτεΐνης (υποδοχέας) κυρίως με μη ομοιοπολικό τρόπο για να σχηματιστεί ένα σταθερό σύμπλεγμα πιθανής αποτελεσματικότητας και μεγαλύτερης εξειδίκευσης. Οι πληροφορίες που λαμβάνονται από την σύνδεση μπορούν να χρησιμοποιηθούν για να υποδείξουν την ενέργεια δέσμευσης, την ελεύθερη ενέργεια και τη σταθερότητα των συμπλεγμάτων. Προς το παρόν, η τεχνική του docking χρησιμοποιείται για να προβλέψει εκ των προτέρων τις δοκιμαστικές παραμέτρους δέσμευσης του συμπλόκου προσδέτη-υποδοχέα. Η σύνδεση αυτή μπορεί να βοηθήσει στην πρόβλεψη της συγγένειας και των ιδιοτήτων δέσμευσης των φαρμάκων προς έναν συγκεκριμένο στόχο. Γενικά, το καλύτερο μόριο συγγένειας δέσμευσης θεωρείται ότι είναι πιο ισχυρό. Με βάση αυτό, ο αριθμός των λογισμικών σύνδεσης (όπως AutoDock, Mol Dock, Maestro κλπ.) είναι διαθέσιμος για ερευνητές που είναι φιλικό προς το χρήστη και είναι εύκολοι στη χρήση [33,34].

Ο κύριος στόχος της μοριακής πρόσδεσης είναι η επίτευξη του συμπλόκου προσδέτη-υποδοχέα με βελτιστοποιημένη διαμόρφωση και με σκοπό χαμηλή ελεύθερη ενέργεια δέσμευσης. Η καθαρή προβλεπόμενη ελεύθερη ενέργεια δέσμευσης (ΔG_{bind}) αποκαλύπτεται ως προς διάφορες παραμέτρους όπως ο δεσμός υδρογόνου (ΔG_{hbond}), ηλεκτροστατική ενέργεια (ΔG_{elec}), ελεύθερης στρέψης (ΔG_{tor}), διασπορά και απώθηση (ΔG_{vdw}), αποδιάλυση (ΔG_{desolv}), συνολική εσωτερική ενέργεια (ΔG_{total}) και την ενέργεια του αδέσμευτου συστήματος (ΔG_{unb}). Επομένως, η καλή κατανόηση της γενικής ηθικής που διέπει την προβλεπόμενη ελεύθερη ενέργεια δέσμευσης (ΔG_{bind}) παρέχει πρόσθετες ενδείξεις σχετικά με τη φύση των διαφόρων ειδών αλληλεπιδράσεων που οδηγούν στη μοριακή πρόσδεση [32].

Προκειμένου να γίνει βέλτιστη χρήση περιορισμένων υπολογιστικών πόρων, αναπτύχθηκαν διάφορες συναρτήσεις βαθμολόγησης για τη μοριακή πρόσδεση ώστε να προβλεφθεί γρήγορα πόσο ισχυρή είναι η δέσμευση μεταξύ μιας χημικής ένωσης και μιας πρωτεΐνης-μακρομορίου στόχου. Αν και οι συναρτήσεις συνήθως βαθμονομούνταν με πειραματικά αποτελέσματα, οι προβλεπόμενες τιμές δεν είναι η πραγματική αλλαγή δεσμευτικής ελεύθερης ενέργειας. Με άλλα λόγια, δεν παρατηρήθηκε ισχυρή σχέση μεταξύ της πειραματικά προσδιορισμένης βιολογικής δραστηριότητας και της προβλεπόμενης ισχύος δέσμευσης από διάφορες συναρτήσεις βαθμολόγησης. Για να ξεπεραστεί το πρόβλημα, αναφέρθηκαν αρκετές ακριβείς προσεγγίσεις για τον υπολογισμό της ελεύθερης ενέργειας δέσμευσης και της συγγένειας δέσμευσης, για παράδειγμα, διαταραχή ελεύθερης ενέργειας, θερμοδυναμική ολοκλήρωση και λιγότερο αυστηρές μέθοδοι τελικού σημείου όπως η ενέργεια γραμμικής αλληλεπίδρασης, η μοριακή μηχανική-Poisson-Boltzmann και η λ-δυναμική. Ωστόσο, λόγω της σημαντικής ζήτησης για υπολογιστικούς πόρους, κανένας από αυτούς δεν μπορούσε να εφαρμοστεί σε σύνδεση υψηλής απόδοσης λόγω του μεγάλου αριθμού συνδεδεμένων διαμορφώσεων κάθε ένωσης σε μια βάση χημικών δεδομένων [28].

Η συνάρτηση βαθμολόγησης συνήθως αποτελείται από διαφορετικούς όρους για να περιγράψει διαφορετικές διαμοριακές αλληλεπιδράσεις και τη συμβολή τους στη συνολική δύναμη δέσμευσης. Ωστόσο, ορισμένοι όροι διαμοριακής αλληλεπίδρασης δύσκολα μπορούν να προβλεφθούν με ακρίβεια, όπως το φαινόμενο της διαλυτοποίησης και η μεταβολή της εντροπίας. Επιπλέον, ορισμένες διαμοριακές αλληλεπιδράσεις περιλαμβάνονται στην προκαταρκτική μελέτη, επομένως δεν περιλαμβάνονται στις συναρτήσεις βαθμολόγησης. Για παράδειγμα, ο δεσμός αλογόνου επαληθεύεται ότι συνεισφέρει στη συγγένεια δέσμευσης πρωτεΐνης-προσδέματος αλλά δεν αντιμετωπίζεται καλά από τις περισσότερες λειτουργίες βαθμολόγησης και ορισμένες ειδικές αλληλεπιδράσεις, όπως η αλληλεπίδραση γουανιδίνης-αργινίνης, δεν λαμβάνονται υπόψη σε καμία λειτουργίες. Επιπλέον, διαφορετικές διαμοριακές αλληλεπιδράσεις δεν είναι πάντα πρόσθετες στη συνολική ισχύ δέσμευσης. Για παράδειγμα, ο δεσμός αλογόνου θα μπορούσε να ενισχυθεί ή να εξασθενήσει τόσο από τη αλληλεπίδραση π-π όσο και από αλληλεπιδράσεις ανιόντος-π/Ip-π σε διαφορετικές περιπτώσεις και να επηρεαστεί από τη γέφυρα νερού. Εκτός από αυτό, μπορεί να υπάρχουν κάποιες άγνωστες διαμοριακές αλληλεπιδράσεις που όντως συμμετέχουν στη δέσμευση πρωτεΐνης-προσδέματος. Η επιτυχία μετριέται συχνά με όρους ρίζας μέσης τετραγωνικής απόκλισης (Root Mean Square Deviation, RMSD) των καρτεσιανών συντεταγμένων των ατόμων του προσδέτη στις συνδεδεμένες και κρυσταλλογραφικές διαμορφώσεις. Μία σύνδεση θεωρείται γενικά επιτυχής εάν είναι μικρότερη από το κάπως αυθαίρετο όριο των 2 Å και υπάρχουν εναλλακτικά μέτρα επιτυχίας, όπως εάν ανακτώνται οι σωστές αλληλεπιδράσεις συνδέτη-στόχου [32,33,34].

2. Μέθοδοι και Υλικά

Η παρούσα διπλωματική εργασία είχε ως στόχο την αξιοποίηση διαδικτυακών εργαλείων χημειοπληροφορικής και βιοπληροφορικής για τη διερεύνηση της μοριακής πρόσδεσης ενός μεγάλου αριθμού φαρμάκων, τα οποία είναι διαθέσιμα σε ανοικτές φαρμακευτικές βάσεις δεδομένων, με πρωτεΐνες που σχετίζονται με αυτοφλεγμονώδεις ασθένειες. Ειδικότερα, η μελέτη επικεντρώθηκε στην πρωτεΐνη που προκαλεί το σύνδρομο TRAPS, και συγκεκριμένα στην εξωκυτταρική περιοχή του υποδοχέα επίσης γνωστή ως TNFR1 (PDB ID: 1EXT) από την τράπεζα δεδομένων πρωτεϊνικών δομών (Protein Data Bank, PDB). Η προσέγγιση βασίζεται στη φιλοσοφία της μοριακής πρόσδεσης, σύμφωνα με την οποία υποψήφια μόρια αξιολογούνται με βάση την ικανότητά τους να σχηματίζουν ενεργειακά ευνοϊκές και σταθερές αλληλεπιδράσεις με το υπό μελέτη μακρομόριο-στόχο. Εφαρμόστηκαν υπολογιστικές μέθοδοι μοριακής πρόσδεσης (docking) και χρησιμοποιήθηκαν διάφορα εργαλεία πρόβλεψης της αλληλεπίδρασης του συμπλόκου φαρμάκου-πρωτεΐνης.

Παράλληλα, αναπτύχθηκε και εφαρμόστηκε ένα μοντέλο Συνελικτικού Νευρωνικού Δικτύου (Convolutional Neural Network, CNN) με στόχο την ταξινόμηση και πρόβλεψη μορίων με βάση τα χαρακτηριστικά δόμησης τους και τις πιθανές βιολογικές τους επιδράσεις. Η προσέγγιση ποικίλει ώστε να περιλαμβάνει εγκεκριμένα και επαναχρησιμοποιημένα φάρμακα ή υποψήφια υπό έρευνα φάρμακα ώστε να αναγνωριστούν οι αλληλεπιδράσεις τους με συγκεκριμένους στόχους που εμπλέκονται σε αυτοφλεγμονώδεις ασθένειες. Στη συνέχεια, τίθεται το ερώτημα κατά πόσο μπορούν να ανιχνευθούν ισχυρές αλληλεπιδράσεις πρόσδεσης μεταξύ αυτών των φαρμάκων και της πρωτεΐνης, ακόμη και όταν τα φάρμακα αυτά δεν έχουν μέχρι στιγμής συσχετιστεί με αυτοφλεγμονώδεις ασθένειες. Εφόσον προκύψουν θετικά αποτελέσματα, αυτά τα φάρμακα θα μπορούν να αξιολογηθούν ως νέες θεραπευτικές υποψήφιες ενώσεις ενάντια στις εν λόγω ασθένειες.

2.1 Υπολογιστική Προσέγγιση

Ο συνδυασμός της χημείας, βιολογίας και των υπολογιστικών μεθόδων παρέχει σημαντικά οφέλη σε διαφορετικούς επιστημονικούς τομείς με τη συνεργιστική τους σχέση, ιδιαίτερα στους τομείς της βιοχημείας, της μοριακής βιολογίας, της ιατρικής χημείας και της βιοπληροφορικής. Σε αυτή την εργασία, έγινε χρήση υπολογιστικών εργαλείων, για να καταστεί δυνατή η μελέτη βιολογικών συστημάτων σε ατομικό επίπεδο, η παροχή σχετικών πληροφοριών και, ως εκ τούτου, η ενίσχυση της διαδικασίας ορθολογικού σχεδιασμού φαρμάκων. Ξεκινώντας από τη συλλογή σύνθετων δομών από βάσεις δεδομένων μέχρι την αξιολόγηση της μοριακής πρόσδεσης, τα εργαλεία πληροφορικής αποδείχθηκαν ιδιαίτερα χρήσιμα. Η δημιουργία της βάσης δεδομένων είναι το πρώτο στάδιο ανάπτυξης φαρμάκων. Είναι ένα διερευνητικό βήμα που στοχεύει στην αποκάλυψη υποψήφιων φαρμάκων, γονιδιακών στόχων και αιτιολογικών παραγόντων μιας δεδομένης ασθένειας ή ενός δεδομένου χημικού παράγοντα.

Η υπολογιστική πρόσδεση (computational docking) χρησιμοποιείται για την πρόβλεψη των τρόπων σύνδεσης δύο ή περισσότερων μορίων και βασίζεται σε δύο μεθόδους, στο πεδίο δύναμης για την εκτίμηση της ελεύθερης ενέργειας δέσμευσης του συμπλόκου που τυπικά υπολογίζεται με βάση μια συγκεκριμένη δεσμευμένη διαμόρφωση, και στη μέθοδο αναζήτησης για την εξερεύνηση του διαμορφωτικού χώρου που είναι διαθέσιμος στο μόριο και στο στόχο. Συχνά, πολλές προσεγγίσεις πρέπει να ενσωματωθούν στη μέθοδο, τόσο στο πεδίο δύναμης όσο και στην αναζήτηση διαμόρφωσης, για να επιτραπεί η σύνδεση με μια λογική υπολογιστική προσπάθεια. Αυτά μπορεί να περιλαμβάνουν χρήση απλοποιημένων πεδίων δύναμης, περιορισμό του χώρου αναζήτησης ή περιορισμούς στη διαμορφωτική ευελιξία του προσδέτη ή και του στόχου. Η μοριακή πρόσδεση μπορεί να χρησιμοποιηθεί ως σύνθετο φίλτρο για να τονίσει μόνο τα πιο ενδιαφέροντα υποψήφια φάρμακα αλλά και για την ανίχνευση πιθανών παρενεργειών φαρμάκων ή μοριακών τοξικοτήτων. Η ανίχνευση του βέλτιστα συνδεδεμένου προσδέτη μεταξύ μιας βάσης δεδομένων είναι σημαντική. Ωστόσο, μπορεί να αποτελέσει τη βάση της έρευνας για την εύρεση νέων στόχων ή για τον χαρακτηρισμό πιθανών παρενεργειών, όπως η τοξικότητα, όπου ο ρόλος της πρωτεΐνης και του προσδέτη αλλάζει.

2.1.1 Βάσεις δεδομένων και σύνολα δεδομένων δομής

Η συλλογή σύνθετων δομών είναι ένα ουσιαστικό βήμα για τη διαδικασία αξιολόγησης της μοριακής πρόσδεσης που στηρίχτηκε κυρίως σε βάσεις δεδομένων που είναι διαθέσιμες στο διαδίκτυο, όπου τα σχετικά δεδομένα είναι εύκολο να ληφθούν και να διαβαστούν αυτόματα. Αυτό το κριτήριο επιτρέπει την εύκολη ενσωμάτωση δεδομένων σε μια υπολογιστική μέθοδο. Απαιτήθηκε επίσης οι βάσεις αυτές να περιέχουν υψηλής ποιότητας δεδομένα και να παρέχουν βαθμολογία εμπιστοσύνης σε μεγάλη κλίμακα ώστε να αποφεύγεται η ανάγκη συλλογής δεδομένων από διαφορετικές πηγές, οι οποίες μπορεί να υποβάλλονται σε προεπεξεργασία σε διαφορετικές, μη συγκρίσιμες μορφές. Εξάλλου, όποτε ήταν δυνατόν, αναζητήθηκαν βάσεις δεδομένων που επωφελήθηκαν από πρόσφατες ενημερώσεις (λιγότερο από ένα έτος), προκειμένου να διασφαλίσουμε ότι εξακολουθούν να ενημερώνονται και να είναι σχετικές. Για να δημιουργήσουμε μια βιβλιοθήκη φαρμάκων (εγκεκριμένα, επανασχεδιασμένα και σε συνεχείς κλινικές δοκιμές) για τις αυτοφλεγμονώδεις ασθένειες για την παρούσα διπλωματική εργασία, εστίασαμε σε τρεις βάσεις δεδομένων που είναι διαθέσιμες στο διαδίκτυο και τα δεδομένα τους μπορούν να ληφθούν εύκολα, συγκεκριμένα στις Malacards, OpenTargets και DrugBank (βλ. Πίνακα 1).^[35-37] Ακολούθως για την επιλογή της δομής της πρωτεΐνης χρησιμοποιήθηκε η βάση RCSB PDB.^[38]

Οι βάσεις Malacards και OpenTargets συλλέγουν δεδομένα από διάφορες πηγές και παρέχουν πληροφορίες για το προφίλ φαρμάκων-ασθένειας. Η DrugBank αποτελεί μία εκτενή φαρμακευτική βάση δεδομένων, η οποία περιλαμβάνει όχι μόνο εγκεκριμένα φάρμακα, αλλά και αποσυρόμενες ή επανασχεδιασμένες φαρμακευτικές ενώσεις. Ιδιαίτερης σημασίας είναι το γεγονός ότι η DrugBank περιέχει μεγάλο αριθμό εγκεκριμένων φαρμάκων για άλλες παθήσεις, γεγονός που καθιστά δυνατή τη διερεύνηση στρατηγικών επανατοποθέτησης φαρμάκων (drug

repurposing) για νέους θεραπευτικούς στόχους.. Επιτρέπει επίσης τη λήψη όλων των δομών φαρμάκων που περιλαμβάνονται στη βάση δεδομένων σε μορφή SDF. Θα πρέπει να σημειωθεί ότι τα περισσότερα σύνολα δεδομένων δομής παρέχουν τα δεδομένα με τη μορφή μεμονωμένων αρχείων mol για ξεχωριστά μόρια ή συνδυασμούς αρχείων mol για τα δεδομένα δομής. Τα σύνθετα δεδομένα μπορούν να εκφραστούν είτε ως μονοδιάστατες, δισδιάστατες ή τρισδιάστατες δομές (1D, 2D, 3D, αντίστοιχα). Οι δομές 1D προέρχονται από τον μοριακό τύπο. Οι 2D και 3D δομές παρουσιάζουν δομικές πληροφορίες με τις 2D να υπολογίζονται από τη δισδιάστατη δομή όπως ο αριθμός των δακτυλίων βενζολίου και των δοτών δεσμών υδρογόνου. Επιπλέον, οι τρισδιάστατοι δομές βασίζονται στην τρισδιάστατη αναπαράσταση ενός μορίου, για παράδειγμα την επιφάνεια που είναι προσβάσιμη σε διαλύτη η οποία έχει θετικό μερικό φορτίο. Πλέον οι ερευνητές εργάζονται με δισδιάστατες μοριακές δομές αντί για 3D, επειδή είναι λιγότερο περίπλοκες. Ενώ η τρισδιάστατη ομοιότητα δομής καταγράφει όλες τις πτυχές της μοριακής δομής, όπως το μέγεθος και το σχήμα, το 2D δεν το κάνει, κάτι που θα μπορούσε να είναι περιοριστικό σε ορισμένες περιπτώσεις. Ένα αντιπροσωπευτικό παράδειγμα δομής SDF παρουσιάζεται στο Παράρτημα Ι.^[35-37]

Για την μοριακή πρόσδεση χρησιμοποιείται ακόμα ένα είδος αρχείων, το αρχείο PDBQT το οποίο είναι μια τροποποιημένη μορφή του PDB, η οποία χρησιμοποιείται κυρίως από το λογισμικό AutoDock 4.2 και AutoDock Vina για προσομοιώσεις πρόσδεσης μορίων. Περιλαμβάνει τις τρισδιάστατες συντεταγμένες των ατόμων, τους χημικούς δεσμούς, καθώς και πρόσθετες πληροφορίες σχετικά με τις περιστροφικές ελευθερίες του μορίου, που καθορίζονται μέσω γωνιών στρέψης. Η μορφή αυτή που φαίνεται στο Παράρτημα ΙΙ είναι απαραίτητη για τον υπολογιστικό προσδιορισμό της αλληλεπίδρασης μεταξύ βιολογικών μακρομορίων, όπως οι πρωτεΐνες, και μικρών μορίων, επιτρέποντας έτσι την αποτελεσματική μελέτη πιθανών φαρμακολογικών στόχων και την ανακάλυψη νέων θεραπευτικών ενώσεων.

Επιπλέον, η γνώση της τρισδιάστατης δομής ενός βιολογικού μακρομορίου είναι απαραίτητη για την κατανόηση του ρόλου που αποτελεί στην υγεία και τις ασθένειες του ανθρώπου και των ζώων, τη λειτουργία του στα φυτά και την παραγωγή τροφίμων και ενέργειας και τη σημασία του σε άλλα θέματα που σχετίζονται με την παγκόσμια ευημερία και βιωσιμότητα. Το RCSB PDB (RCSB.org) αποτελεί το κέντρο δεδομένων των Ηνωμένων Πολιτειών Αμερικής για το αρχείο της παγκόσμιας τράπεζας πρωτεϊνών (Protein Data Bank, PDB) με δεδομένα τρισδιάστατης δομής για μεγάλα βιολογικά μόρια (πρωτεΐνες, DNA και RNA) απαραίτητα για την έρευνα και την εκπαίδευση στη θεμελιώδη βιολογία, υγεία, ενέργεια και βιοτεχνολογία. Μέσω μιας διαδικτυακής πύλης πληροφοριών και αρχείου δεδομένων με δυνατότητα λήψης, το PDB παρέχει πρόσβαση σε δεδομένα τρισδιάστατης δομής για τα μόρια της ζωής, που βρίσκονται σε όλους τους οργανισμούς στον πλανήτη. Ο τεράστιος πλούτος δεδομένων τρισδιάστατης δομής που αποθηκεύονται στο PDB έχει υποστηρίξει σημαντικές προόδους στην κατανόηση της αρχιτεκτονικής πρωτεϊνών, με αποκορύφωμα τις πρόσφατες ανακαλύψεις στην πρόβλεψη της δομής των πρωτεϊνών που επιταχύνθηκαν από προσεγγίσεις τεχνητής νοημοσύνης και μεθόδους βαθιάς ή μηχανικής μάθησης.^[38]

Πίνακας 1. Βάσεις δεδομένων από τις οποίες οι δομές των φαρμάκων (εγκεκριμένων, σε κλινικές δοκιμές και που έχουν αποσυρθεί) και της πρωτεΐνης έχουν εξαχθεί.

Βάση Δεδομένων	Περιγραφή	Δεδομένα
Malacards ^[35]	Ανθρώπινες ασθένειες και οι σχολιασμοί τους, κορυφαία γονίδια, μεταλλάξεις, σημασία, φαινότυποι, κλινικές δοκιμές. Η κάρτα ασθένειας περιέχει ενότητες: Ψευδώνυμα & Ταξινομήσεις. Ανατομικό πλαίσιο, Φάρμακα & Θεραπευτικά.	Χημικά, γονιδιωματικά δεδομένα
DrugBank ^[36]	Σχετιζόμενα φάρμακα, περιλαμβάνει εγκεκριμένα, αποσυρθέντα φάρμακα, καθώς και εργαλεία χημικών ενώσεων και αναφέρει τις πιθανές ενδείξεις τους.	Χημικά, Φαρμακευτικά Δεδομένα
OpenTargets ^[37]	Συσχετίσεις ασθενειών-γονιδίων; διαφορετικά είδη στο CTD, χειροκίνητες αλληλεπιδράσεις αντί υπολογιστικές. Δεδομένα από διάφορες επιμελημένες πηγές. Συντελεστές για συσχέτιση ασθένειας-γονιδίου ποσοτικοποιώντας το αντίστοιχο επίπεδο απόδειξης.	Γονιδιωματικά δεδομένα
Protein Data Bank (PDB) ^[38]	Τρισδιάστατα δομικά δεδομένα μεγάλων βιολογικών μορίων όπως οι πρωτεΐνες και τα νουκλεϊνικά οξέα	Δομή πρωτεΐνης, Κρυσταλλογραφία Ακτινών-X, Προσδιορισμός Δομής NMR

2.1.2 Αξιολόγηση μοριακής πρόσδεσης

Μετά την συλλογή όλων των δομών φαρμάκων δημιουργήθηκε μια βιβλιοθήκη δεδομένων από φάρμακα τα οποία είναι σε κλινικές δοκιμές, εγκεκριμένα και που έχουν αποσυρθεί. Ο στόχος ήταν ο εντοπισμός ενώσεων οι οποίες θα προσδένονται ισχυρά στην πρωτεΐνη στόχο. Η αρχή ότι ο υποδοχέας με τον προσδέτη προσανατολίζονται με τέτοιο τρόπο ώστε να εξασκούν ευνοϊκότερες στερεοηλεκτρονικές αλληλεπιδράσεις επάγοντας ένα συγκεκριμένο σήμα που μεταφράζεται σε βιολογικό αποτέλεσμα είναι μια υποκείμενη υπόθεση εικονικής διαλογής με βάση την δομή. Τα προγράμματα docking αξιολογούν και ταξινομούν τις πιθανές θέσεις πρόσδεσης με την βοήθεια συναρτήσεων βαθμολόγησης (scoring functions). Στην πράξη, η προτεινόμενη από τα προγράμματα docking ιεράρχηση των λύσεων του προβλήματος δεν εγγυάται ότι η ορθή λύση θα είναι η πρώτη ή μεταξύ των πρώτων. Πρόσθετες βιοχημικές πληροφορίες πρέπει να αξιοποιηθούν προκειμένου να επιλεγεί η σωστή. Δυο προγράμματα μοριακής πρόσδεσης και ένας αλγόριθμος ο οποίος αναπτύχθηκε χρησιμοποιήθηκαν ώστε να βρεθούν οι ενώσεις με την καλύτερη πρόσδεση

στην πρωτεΐνη. Για τον έλεγχο της μοριακής πρόσδεσης, ο αλγόριθμος που αναπτύχθηκε περιλαμβάνει εντολές που επικεντρώνονται στα δακτυλικά αποτυπώματα, στην ανάπτυξη ενός μοντέλου CNN χρησιμοποιώντας το πακέτο Keras και στις μετρήσεις αξιολόγησης. Οι τιμές που προέκυψαν από τη χρήση των δύο προγραμμάτων καθώς και από τον αλγόριθμο υποδεικνύουν αν η ένωση προσδένεται ισχυρά ή όχι με την πρωτεΐνη.

2.1.3 Εντοπισμός περιοχών πρόσδεσης στην πρωτεΐνη

Για τον εντοπισμό των πιθανών περιοχών πρόσδεσης μικρών μορίων στην πρωτεΐνη χρησιμοποιήθηκαν δύο διαδικτυακά διαθέσιμα εργαλεία, το CASTp και το DoGSiteScorer. Τα δύο αυτά συστήματα προσφέρουν συμπληρωματικές πληροφορίες, καθώς το CASTp χαρτογραφεί γεωμετρικά τις κοιλότητες της πρωτεΐνης, ενώ το DoGSiteScorer αναλύει τοπικές και συνολικές ιδιότητες των θυλάκων και αξιολογεί τη φαρμακολογική αξιοποιησιμότητα (druggability score) τους. Ο συνδυασμός των δύο εργαλείων επέτρεψε μια πιο ακριβή και αξιόπιστη πρόβλεψη των υποψήφιων ενεργών και αλλοστερικών περιοχών πρόσδεσης.

Συγκεκριμένα, το CASTp είναι ένα διαδικτυακό εργαλείο το οποίο εντοπίζει και μετρά θύλακες και κενά σε τρισδιάστατες δομές πρωτεϊνών. Ορίζεται ως ένας υπολογιζόμενος Άτλας της επιφανειακής τοπογραφίας των πρωτεϊνών με δομική και τοπογραφική χαρτογράφηση λειτουργικών κατάλοιπων. Χρησιμοποιεί μια ημι-καθολική μέθοδο ευθυγράμμισης αλληλουχίας ανά ζεύγη για τη χαρτογράφηση της αλληλουχίας μεταξύ καταχωρήσεων στο Swiss-Prot, στο OMIM και καταχωρήσεων στο PDB. Το CASTp μπορεί να χρησιμοποιηθεί για τη μελέτη χαρακτηριστικών επιφάνειας, λειτουργικών περιοχών και συγκεκριμένων ρόλων βασικών υπολειμμάτων πρωτεϊνών. Το CASTp βασίζεται σε πρόσφατα θεωρητικά και αλγοριθμικά αποτελέσματα της υπολογιστικής γεωμετρίας. Έχει πολλά πλεονεκτήματα όπως να αναγνωρίζει αναλυτικά θύλακες και κοιλότητες, να ορίζει με ακρίβεια το όριο μεταξύ του διαλύτη και του θύλακα και όλες οι υπολογισμένες παράμετροι είναι περιστροφικά αμετάβλητες, δεν περιλαμβάνουν διακριτοποίηση και δεν χρησιμοποιούν σημεία επιφάνειας κουκκίδων ή πλέγματος ^[51].

Το DoGSiteScorer, εργαλείο της πλατφόρμας ProteinsPlus, είναι ένας άλλος αυτόματος αλγόριθμος πρόβλεψης περιοχών πρόσδεσης σε πρωτεΐνες, ο οποίος χρησιμοποιείται ευρέως για την αξιολόγηση πιθανών φαρμακολογικά αξιοποιήσιμων θυλάκων ^[52]. Η μέθοδος βασίζεται αποκλειστικά στις ατομικές συντεταγμένες της πρωτεΐνης και χρησιμοποιεί ένα φίλτρο Διαφοράς Γκαουσιανού για τον εντοπισμό συνεχών κοιλοτήτων στην επιφάνειά της. Εκτός από τις συνολικές ιδιότητες του θύλακα, το DoGSiteScorer λαμβάνει υπόψη και τις τοπικές ομοιότητες που παρουσιάζονται ανάμεσα σε γειτονικές περιοχές, εντοπίζοντας έτσι τόσο κύριους όσο και υπό-θυλάκους. Για κάθε προβλεπόμενο θύλακα, το εργαλείο υπολογίζει μια σειρά από περιγραφείς (descriptors) που χαρακτηρίζουν το μέγεθος, το σχήμα και τα χημικά χαρακτηριστικά της περιοχής, όπως η υδροφοβικότητα. Με βάση αυτούς τους περιγραφείς υπολογίζεται μια απλή βαθμολογία φαρμακευσιμότητας (druggability score), η οποία προκύπτει από το γραμμικό

συνδυασμό βασικών παραμέτρων όπως ο όγκος, η υδροφοβικότητα και ο βαθμός «περίφραξης» του θύλακα. Επιπλέον, ένα υποσύνολο κρίσιμων περιγραφέων αξιοποιείται μέσω ενός μοντέλου για την παραγωγή μιας πιο εκλεπτυσμένης βαθμολογίας φαρμακολογικής στόχευσης, η οποία λαμβάνει τιμές από 0 έως 1. Όσο υψηλότερη είναι η τιμή αυτή, τόσο μεγαλύτερη θεωρείται η πιθανότητα ο θύλακας να μπορεί να στοχευθεί επιτυχώς από μικρά μόρια.

2.1.4 Ομαδοποίηση ενώσεων

Μετά την ολοκλήρωση της διαδικασίας μοριακής πρόσδεσης, ακολούθησε ομαδοποίηση (clustering) των ενώσεων. Η ομαδοποίηση επιτρέπει την κατάταξη των ενώσεων σε συστάδες με βάση τη μεταξύ τους ομοιότητα, συμβάλλοντας στην αποτελεσματικότερη επιλογή υποψηφίων μορίων για περαιτέρω αξιολόγηση. Με τον τρόπο αυτό, καθίσταται δυνατή η εστίαση σε αντιπροσωπευτικές ενώσεις από συστάδες που περιλαμβάνουν μόρια με γνωστή ή υψηλή βιολογική δραστηριότητα. Οι αλγόριθμοι ομαδοποίησης διακρίνονται γενικά σε δύο κύριες κατηγορίες: τους μη ιεραρχικούς και τους ιεραρχικούς. Στην παρούσα μελέτη εφαρμόστηκε ιεραρχική ομαδοποίηση, όπου κάθε ένωση αρχικά θεωρείται ως ανεξάρτητο σύμπλεγμα, και στη συνέχεια τα συμπλέγματα συγχωνεύονται σταδιακά με βάση κριτήρια ομοιότητας. Ως μέθοδος σύνδεσης χρησιμοποιήθηκε ο αλγόριθμος κατά Ward, με όριο ομαδοποίησης ίσο με 0.8. Η μέθοδος Ward βασίζεται στη βελτιστοποίηση της ενδοσυσταδικής διακύμανσης, επιλέγοντας σε κάθε βήμα τη συγχώνευση των δύο συστάδων που οδηγεί στη μικρότερη δυνατή αύξηση του συνολικού αθροίσματος των τετραγώνων των αποκλίσεων. Με τον τρόπο αυτό επιτυγχάνεται η δημιουργία συμπαγών και ομοιογενών συστάδων, με μειωμένη διακύμανση εντός συστάδας και αυξημένη διακριτότητα μεταξύ διαφορετικών συστάδων. Η ιεραρχική ομαδοποίηση κατά Ward έχει χρησιμοποιηθεί ευρέως στην ανάλυση και ταξινόμηση χημικών ενώσεων, παρουσιάζοντας ιδιαίτερα ικανοποιητικά αποτελέσματα στον διαχωρισμό ενεργών και ανενεργών μορίων. Σε πολλές μελέτες έχει αποδειχθεί ότι υπερέρχει ή εμφανίζει σταθερότερη απόδοση σε σύγκριση με εναλλακτικούς αλγορίθμους ομαδοποίησης, γεγονός που την καθιστά μέθοδο αναφοράς για εφαρμογές χημειοπληροφορικής και εικονικού φαρμακοσχεδιασμού.^[39]

2.2 Αλγόριθμοι Μοριακής Πρόσδεσης

Χρησιμοποιήθηκαν διαδικτυακά εργαλεία χημειοπληροφορικής, προκειμένου να πραγματοποιηθεί έλεγχος ενός τεράστιου αριθμού δομών φαρμάκων που είναι εμπορικά διαθέσιμα, σε κλινικές δοκιμές και που έχουν αποσυρθεί για τη μοριακή τους δέσμευση με την αντίστοιχη πρωτεΐνη των αυτοφλεγμονώδων ασθενειών. Η μοριακή πρόσδεση χρησιμοποιεί τις δομές δύο μορίων, του προσδέτη και του στόχου, για να προβλέψει τον προτιμώμενο προσανατολισμό του πρώτου σε σχέση με το δεύτερο όταν συνδέονται μεταξύ τους για να σχηματίσουν ένα σταθερό σύμπλεγμα. Συνήθως, ο προσδέτης είναι το μικρότερο μόριο. Στην ανακάλυψη φαρμάκων, το μόριο προσδέτης είναι ένα ενεργό συστατικό και ο στόχος είναι ένα βιολογικό μακρομόριο (π.χ. μια πρωτεΐνη ή DNA). Εδώ εστιάζουμε στη μοριακή σύνδεση πρωτεΐνης-μικρού μορίου (κοινώς αποκαλούμενη πρωτεΐνη-προσδέτη) επειδή καλύπτει μια

σημαντική επιλογή υφιστάμενων μεθόδων σύνδεσης. Στην παρούσα διπλωματική εργασία χρησιμοποιήθηκαν το AutoDock Vina και το AutoDock4.2 ως προγράμματα μοριακής πρόσδεσης καθώς επίσης και ένας αλγόριθμος εκμάθησης CNN ο οποίος έχει αναπτυχθεί χρησιμοποιώντας την γλώσσα προγραμματισμού Python (βλ. Πίνακα 2). Τα τελευταία χρόνια, μέθοδοι μηχανικής εκμάθησης (Machine Learning, ML), όπως η βαθιά εκμάθηση (Deep Learning, DL) έχουν εφαρμοστεί για τη βελτιστοποίηση της διαδικασίας σύνδεσης. Η μηχανική εκμάθηση στη μοριακή σύνδεση μορίου-πρωτεΐνης μπορεί να φέρει νέες στρατηγικές για τη βαθμολόγηση ενός συμπλέγματος. Η χρήση της στη μοριακή πρόσδεση έχει εξελιχθεί ταχέως και την προηγούμενη δεκαετία εμφανίστηκαν πολυάριθμες μέθοδοι, που όλες έφεραν σημαντική βελτίωση.

Πίνακας 2. Προγράμματα μοριακής πρόσδεσης, ομαδοποίησης και εντοπισμού περιοχών πρόσδεσης της πρωτεΐνης που χρησιμοποιήθηκαν στην εργασία.

Αλγόριθμοι Μοριακής Πρόσδεσης	Περιγραφή
AutoDock Vina ^[39,40]	Πρόγραμμα ανοιχτού κώδικα για μοριακή πρόσδεση.
AutoDock4.2 ^[41]	Σουίτα αυτοματοποιημένων εργαλείων μοριακής πρόσδεσης. Προβλέψεις μικρών μορίων, όπως υποστρώματα ή υποψήφια φάρμακα, για σύνδεση με έναν υποδοχέα γνωστής τρισδιάστατης δομής.
CNN Network ^[42-45]	Μοντέλο αρχιτεκτονικής βαθιάς μάθησης που χρησιμοποιείται ευρέως στην επεξεργασία εικόνων, αλλά και σε άλλες εφαρμογές πρόβλεψης και ταξινόμησης.
ChemBioserver ^[46]	Δωρεάν, βασισμένο στο διαδίκτυο, παρέχει προηγμένο φιλτράρισμα, ομαδοποίηση και δικτύωση χημικών ενώσεων για ανακάλυψη και επαναχρησιμοποίηση φαρμάκων.
CASTp ^[51]	Διαδικτυακό εργαλείο το οποίο εντοπίζει και μετρά θύλακες και κενά σε τρισδιάστατες δομές πρωτεϊνών. Ορίζεται ως υπολογιζόμενος Άτλας της επιφανειακής τοπογραφίας των πρωτεϊνών με δομική και τοπογραφική χαρτογράφηση λειτουργικών κατάλοιπων.
DoGSiteScorer ^[52]	Πλήρως αυτόματος αλγόριθμος για την πρόβλεψη θέσεων πρόσδεσης στη πρωτεΐνη και την αξιολόγηση πιθανών φαρμακολογικά αξιοποιήσιμων θυλάκων.

2.2.1 Μοριακή πρόσδεση χρησιμοποιώντας το AutoDock Vina

Το AutoDock Vina είναι ένας από τους κινητήρες σύνδεσης στο AutoDock Suite, μαζί με το AutoDock4 (AD4). Το καθιστά ένα από τους ταχύτερους και πιο ευρέως χρησιμοποιούμενους κινητήρες σύνδεσης ανοιχτού κώδικα. Βασίζεται σε μια απλή συνάρτηση βαθμολόγησης και ταχεία διαμορφωτική αναζήτηση βελτιστοποίησης. Η γενική λειτουργική μορφή του τμήματος που εξαρτάται από τη διαμόρφωση της συνάρτησης βαθμολόγησης του έχει σχεδιαστεί για να λειτουργεί όπως φαίνεται στη συνάρτηση παρακάτω. Το άθροισμα είναι όλα τα ζεύγη ατόμων που μπορούν να κινούνται μεταξύ τους, συνήθως εξαιρώντας 1,4 αλληλεπιδράσεις, δηλαδή άτομα που χωρίζονται από τρεις διαδοχικούς ομοιοπολικούς δεσμούς.^[41]

$$c = \sum_{i < j} f_{titj}(r_{ij}) \quad (1)$$

Η προβλεπόμενη ελεύθερη ενέργεια δέσμευσης υπολογίζεται από το διαμοριακό τμήμα της διαμόρφωσης με τη χαμηλότερη βαθμολογία, που ορίζεται ως 1, όπου η συνάρτηση g μπορεί να είναι μια αυθαίρετη αυστηρά αυξανόμενη ομαλή πιθανώς μη γραμμική συνάρτηση.^[41]

$$s_l = g(c_l - c_{intra}) = g(c_{interl}) \quad (2)$$

Για την μοριακή πρόσδεση με την χρήση του AutoDock Vina, η πρωτεΐνη 1EXT λήφθηκε από τη βάση δεδομένων PDB (Protein Data Bank) σε μορφή PDB και προετοιμάστηκε χρησιμοποιώντας το AutoDockTools. Αρχικά, αφαιρέθηκαν τα μόρια νερού και τα ετεροάτομα που θα μπορούσαν να επηρεάσουν τα αποτελέσματα και προστέθηκαν τα πολικά υδρογόνα καθώς εφαρμόστηκαν και τα φορτία Kollman. Τα φορτία Kollman βελτιστοποιούν την ακρίβεια των υπολογισμών κατά τη διαδικασία πρόσδεσης. Τα φορτία αυτά αντιπροσωπεύουν τη μερική κατανομή φορτίου στα άτομα της πρωτεΐνης, επιτρέποντας πιο ρεαλιστική μοντελοποίηση των ηλεκτροστατικών αλληλεπιδράσεων με τα μόρια. Επιπλέον, είναι ειδικά σχεδιασμένα για χρήση σε προσομοιώσεις με το AutoDock, διασφαλίζοντας συμβατότητα με το πεδίο δυνάμεών του. Με την εφαρμογή των φορτίων Kollman, βελτιώνεται η εκτίμηση της ενεργειακής σταθερότητας των πιθανών διαμορφώσεων πρόσδεσης, συμβάλλοντας στην αξιόπιστη αναγνώριση των πιο ευνοϊκών αλληλεπιδράσεων μεταξύ της πρωτεΐνης και των υποψήφιων αναστολέων. Στα άτομα της πρωτεΐνης αποδόθηκε ο τύπος AD4, προκειμένου να διασφαλιστεί η σωστή περιγραφή των ατομικών ιδιοτήτων και η συμβατότητα με το πεδίο δυνάμεων του AutoDock4. Ο τύπος αυτός καθορίζει κρίσιμες παραμέτρους, όπως η ακτίνα Van der Waals, η ενέργεια διάχυσης (dispersion energy) και οι ηλεκτροστατικές ιδιότητες, που είναι απαραίτητες για την ακριβή προσομοίωση των αλληλεπιδράσεων μεταξύ της πρωτεΐνης και των μορίων πρόσδεσης. Με αυτόν τον τρόπο, επιτράπηκε η σωστή εκτίμηση των ενεργειακών όρων που εμπλέκονται στη διαδικασία

πρόσδεσης, εξασφαλίζοντας μεγαλύτερη ακρίβεια στην πρόβλεψη των σταθερότερων συμπλόκων πρωτεΐνης-μορίου. Τέλος, το αρχείο αποθηκεύτηκε σε μορφή PDBQT. Είναι απαραίτητο τόσο τα μόρια των ενώσεων όσο και το αρχείο συντεταγμένων του υποδοχέα να είναι στη μορφή αυτή.^[38] Οι ενώσεις που χρησιμοποιήθηκαν για τη μοριακή πρόσδεση λήφθηκαν σε μορφή SDF και, μέσω του προγράμματος OpenBabel, μετατράπηκαν σε PDBQT αρχεία. Χρησιμοποιήθηκαν 7,000 ενώσεις για την μοριακή πρόσδεση στην πρωτεΐνη 1EXT και επιλέχθηκαν οι εκατό καλύτερες ενώσεις που είχαν την μικρότερη ενέργεια δέσμευσης και άρα την καλύτερη μοριακή πρόσδεση, μικρότερη από -8 kcal.

Το AutoDock Vina δεν απαιτεί αρχείο παραμέτρων πλέγματος και σύνδεσης, αλλά ένα αρχείο διαμόρφωσης σε μορφή κειμένου. Για τη ρύθμιση των παραμέτρων, το πλέγμα ορίστηκε χρησιμοποιώντας το AutoDockTools και συγκεκριμένα το AutoGrid, με συντεταγμένες που καλύπτουν ολόκληρη την πρωτεΐνη για τη διεξαγωγή τυφλής πρόσδεσης (blind docking) δηλαδή αντί να προσδιοριστεί εκ των προτέρων το σημείο της αλληλεπίδρασης, η διαδικασία επιτρέπει την εξερεύνηση ολόκληρης της επιφάνειας της πρωτεΐνης για την αναγνώριση του βέλτιστου σημείου δέσμευσης. Αυτός ο τύπος προσομοίωσης είναι χρήσιμος όταν δεν είναι γνωστό το σημείο αλληλεπίδρασης. Οι συντεταγμένες, καθώς και το όνομα της πρωτεΐνης, αποθηκεύτηκαν σε αρχείο κειμένου (Notepad), Παράρτημα III.

Το πλέγμα που χρησιμοποιήθηκε στο Vina ορίστηκε από το κέντρο του και τον αριθμό των σημείων σε κάθε διάσταση. Η απόσταση του πλέγματος είναι πάντα 1Å, και ο ρόλος του είναι αποκλειστικά να καθορίσει το μέγεθος του χώρου αναζήτησης, χωρίς να επηρεάζει την ακρίβεια του υπολογισμού. Η μοριακή πρόσδεση μεταξύ του υποδοχέα και των ενώσεων πραγματοποιήθηκε μέσω ενός αρχείου παρτίδας των Windows (Windows batch file) που είναι ένα αρχείο κειμένου που περιέχει μια σειρά από εντολές που εκτελούνται από τη γραμμή εντολών των Windows, Παράρτημα III. Η προεπιλεγμένη ρύθμιση του δεν είναι από μόνη της αρκετή για την ακριβή αξιολόγηση του προφίλ αλληλεπίδρασης και των ενεργειών δέσμευσης για αυτό για τον έλεγχο της υπολογιστικής ισχύος που χρησιμοποιείται κατά τη διάρκεια της ανάλυσης σύνδεσης, το Vina διαθέτει μια ρυθμιζόμενη παράμετρο που ονομάζεται εξαντλητικότητα (exhaustiveness), με προεπιλεγμένη τιμή 8.

Κατά την εκτέλεση του Vina εξήχθη ένα αρχείο συντεταγμένων της μορφής PDBQT, το οποίο περιέχει πληροφορίες σύνδεσης που προβάλλονται στο παράθυρο του τερματικού όπως φαίνεται στο Παράρτημα IV. Η προβλεπόμενη ελεύθερη ενέργεια δέσμευσης εκφράστηκε σε kcal/mol, ενώ υπολογίστηκαν και οι τιμές RMSD, οι οποίες συγκρίθηκαν με την καλύτερη προβλεπόμενη διαμόρφωση. Η τιμή RMSD είναι ένα μέτρο της διαφοράς ή της απόκλισης μεταξύ δύο συνόλων δεδομένων, και χρησιμοποιείται ευρέως στη βιοπληροφορική και τη μοριακή μοντελοποίηση για την αξιολόγηση της ομοιότητας ή της παραμόρφωσης στη θέση των ατόμων ή μορίων σε δύο διαφορετικές διαμορφώσεις. Στη μοριακή μοντελοποίηση και τη μοριακή δυναμική, η RMSD χρησιμοποιείται για τη σύγκριση των δομών δύο μορίων. Το Vina παρέχει δύο διαφορετικές μεθόδους μέτρησης της RMSD, οι οποίες χρησιμοποιήθηκαν για την αξιολόγηση των αποκλίσεων

μεταξύ των διαμορφώσεων πρόσδεσης. Η πρώτη μέθοδος, rmsd/lb (lower bound RMSD), υπολογίζει το κάτω όριο της απόκλισης, δηλαδή τη μικρότερη δυνατή απόσταση μεταξύ των ατόμων της εξεταζόμενης διαμόρφωσης και της αναφοράς και η δεύτερη μέθοδος, rmsd/ub (upper bound RMSD), αντιστοιχίζει κάθε άτομο μιας διαμόρφωσης με τον εαυτό του, παρέχοντας έτσι το άνω φράγμα της απόκλισης. Η δεύτερη μέτρηση είναι συνήθως μεγαλύτερη, καθώς λαμβάνει υπόψη όλες τις πιθανές αντιστοιχίσεις ατόμων.^[40,41]

Έτσι, ένας εξαιρετικά συμμετρικός άκαμπτος προσδέτης μπορούσε να περιστραφεί σε σχέση με μια διαμόρφωση αναφοράς ώστε η νέα διαμόρφωση να είναι ακριβώς ισοδύναμη με τη διαμόρφωση αναφοράς. Ωστόσο, η εσωτερική αρίθμηση των ατόμων δεν αλλάζει κατά τη διάρκεια της σύνδεσης, επομένως ο αλγόριθμος rmsd/ub, ο οποίος ταιριάζει άτομα με ίδια σήμανση (και όχι παρόμοια ή ισοδύναμα άτομα) αποδίδει σημαντικό rmsd, ενώ ο αλγόριθμος rmsd/lb αποδίδει ένα πιο ρεαλιστικό rmsd του μηδενός. Για δύο δομές, a και b, ενός ίδιου μορίου, ορίζεται το RMSD όπως φαίνεται παρακάτω και το άθροισμα είναι από όλα τα N βαρέα άτομα στη δομή a, το ελάχιστο είναι από όλα τα άτομα στη δομή b με τον ίδιο τύπο στοιχείου με το άτομο i στη δομή a.^[38]

$$RMSD_{ab} = \max(RMSD'_{ab}, RMSD'_{ba}) \quad (3)$$

$$RMSD'_{ab} = \sqrt{\frac{1}{N} \sum_i \min_j r_{ij}^2} \quad (4)$$

2.2.2 Μοριακή πρόσδεση χρησιμοποιώντας το AutoDock4.2

Για τη διερεύνηση της συγγένειας πρόσδεσης με το AutoDock4.2, χρησιμοποιήθηκαν τα αποτελέσματα του AutoDock Vina και συγκεκριμένα οι 100 ενώσεις με τη χαμηλότερη ελεύθερη ενέργεια δέσμευσης, καθώς θεωρούνται οι πιο υποσχόμενοι υποψήφιοι αναστολείς. Αυτές οι ενώσεις χρησιμοποιήθηκαν για επαναξιολόγηση της πρόσδεσής τους με το AutoDock4.2, αφού θεωρείται πιο έγκυρη και ακριβής μέθοδος μοριακής πρόσδεσης αλλά και να γίνει σύγκριση των αποτελεσμάτων των δύο μεθόδων. Η διαδικασία προετοιμασίας των μορίων πραγματοποιήθηκε μέσω του AutoDockTools. Η προετοιμασία της πρωτεΐνης ήταν η ίδια διαδικασία η οποία πραγματοποιήθηκε και στο AutoDock Vina δηλαδή στη πρωτεΐνη προστέθηκαν τα πολικά υδρογόνα και φορτία Kollman, ενώ στις ενώσεις προστέθηκαν φορτία Gasteiger. Όλα τα μόρια αποθηκεύτηκαν σε PDBQT μορφή, η οποία απαιτείται για τη διεξαγωγή των προσομοιώσεων πρόσδεσης. Στο AutoDock4.2, δημιουργήθηκε ένα πλέγμα (grid) που κάλυπτε όλη την περιοχή της πρωτεΐνης, χρησιμοποιώντας το AutoGrid4.2 για τον υπολογισμό των χαρτών αλληλεπίδρασης των διαφόρων ατόμων του μορίου με τον υποδοχέα (Παράρτημα V).

Το AutoDock απαιτεί προ-υπολογισμένους χάρτες πλέγματος (grid maps), έναν για κάθε τύπο ατόμου που υπάρχει στο μόριο που προσδένεται. Αυτό βοηθά τους υπολογισμούς πρόσδεσης να γίνουν πιο γρήγορα. Αυτοί οι χάρτες αποτελούνται από ένα τρισδιάστατο πλέγμα σημείων σε κανονική απόσταση, που περιβάλλουν εξολοκλήρου και επικεντρώνονται σε όλη τη περιοχή του μακρομορίου. Η τυπική απόσταση σημείων πλέγματος κυμαίνεται από 0.2Å έως 1.0Å και η προεπιλογή είναι 0.375Å (περίπου το ένα τέταρτο του μήκους ενός απλού δεσμού άνθρακα-άνθρακα). Κάθε σημείο εντός του χάρτη πλέγματος αποθηκεύει τη δυναμική ενέργεια ενός ατόμου «ανιχνευτή» ή μιας λειτουργικής ομάδας που οφείλεται σε όλα τα άτομα του μακρομορίου. Το αρχείο παραμέτρων πλέγματος έχει την επέκταση ".grpf". Μια πλήρης περιγραφή του αρχείου παραμέτρων πλέγματος περιλαμβάνεται στο Παράρτημα V.^[39,40]

Η αξιολόγηση της αλληλεπίδρασης προσδέτη-πρωτεΐνης πραγματοποιήθηκε με τη χρήση των χαρτών πλέγματος (grid maps), οι οποίοι δημιουργήθηκαν από το AutoGrid και περιγράφουν την ενεργειακή συμβολή κάθε τύπου ατόμου. Η διαδικασία πρόσδεσης έλαβε υπόψη πολλαπλές διαμορφώσεις του προσδέτη (poses) και βασίστηκε σε χάρτες πλέγματος για κάθε τύπο ατόμου στο σύστημα, υπολογισμένους μέσω του AutoGrid. Η προσομοίωση πραγματοποιήθηκε χρησιμοποιώντας ένα αρχείο PDBQT για τον υποδοχέα και το προσδέτη, καθώς και ένα αρχείο παραμέτρων πρόσδεσης (docking parameter file, .dpf), το οποίο δημιουργήθηκε μέσω του AutoDockTools. Το αρχείο αυτό καθορίζει τα απαραίτητα δεδομένα και ρυθμίσεις για την πρόσδεση, όπως οι χάρτες πλέγματος που χρησιμοποιούνται, τα αρχεία συντεταγμένων του προσδέτη, καθώς και τα κριτήρια αναζήτησης της βέλτιστης διαμόρφωσης.

Για την αναζήτηση της βέλτιστης διαμόρφωσης χρησιμοποιήθηκε ο Λαμαρκιάνιος Γενετικός Αλγόριθμος (Lamarckian Genetic Algorithm, LGA), ο οποίος συνδυάζει γενετικούς αλγορίθμους με τοπική αναζήτηση, επιτρέποντας τη δυναμική προσαρμογή του προσδέτη εντός της θέσης πρόσδεσης, διασφαλίζοντας έναν αποδοτικό και στοχευμένο τρόπο εξερεύνησης του διαμορφωτικού χώρου του προσδέτη.^[40] Το AutoDock 4.2 βασίζεται σε ένα ημι-εμπειρικό πεδίο ελεύθερης ενέργειας για την πρόβλεψη τόσο της γεωμετρίας του συμπλόκου όσο και της ελεύθερης ενέργειας πρόσδεσης ($\Delta G_{\text{binding}}$). Η αξιολόγηση της πρόσδεσης πραγματοποιείται σε δύο διακριτά στάδια. Στο πρώτο στάδιο υπολογίζονται οι ενδομοριακές ενέργειες, οι οποίες αντιστοιχούν στη μεταβολή της ενέργειας του προσδέτη και της πρωτεΐνης κατά τη μετάβαση από την αδέσμευτη κατάσταση στη δεσμευμένη διαμόρφωση. Στο δεύτερο στάδιο εκτιμώνται οι διαμοριακές ενέργειες, που προκύπτουν από τις αλληλεπιδράσεις μεταξύ προσδέτη και πρωτεΐνης στο τελικό σύμπλοκο πρόσδεσης (Εικόνα 3).^[40,41] Το πεδίο δύναμης του AutoDock4.2 περιλαμβάνει έξι όρους αλληλεπίδρασης κατά ζεύγη, καθώς και έναν πρόσθετο όρο που λαμβάνει υπόψη την απώλεια διαμορφωτικής εντροπίας του προσδέτη κατά τη διαδικασία της πρόσδεσης (ΔS_{conf}).^[47,48]

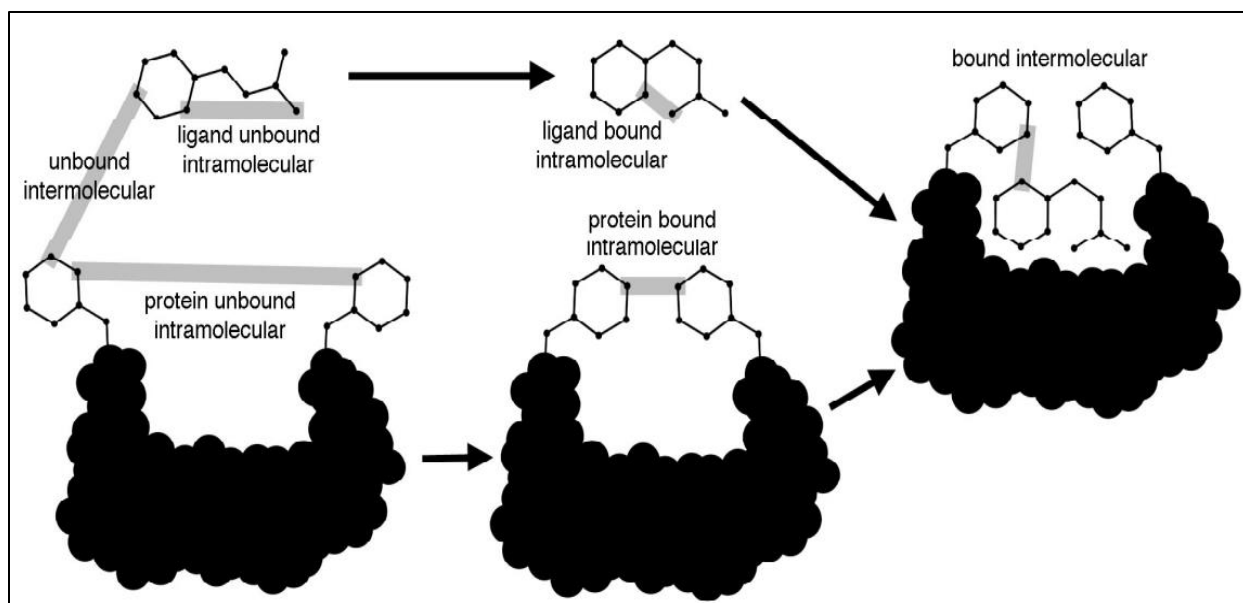
$$\Delta G = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + (V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S_{conf}) \quad (5)$$

Κάθε ενεργειακός όρος που υπολογίζεται μεταξύ ζευγών ατόμων περιλαμβάνει συνεισφορές από διασπορά/απώθηση, δεσμούς υδρογόνου, ηλεκτροστατικές αλληλεπιδράσεις και αποδιάλυση. Οι σταθερές στάθμισης W έχουν βελτιστοποιηθεί ώστε να βαθμονομήσουν την εμπειρική ελεύθερη ενέργεια δέσμευσης με βάση ένα σύνολο πειραματικά καθορισμένων σταθερών πρόσδεσης.

$$V = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^6} \right) + W_{elec} \sum_{i,j} \frac{q_i q_j}{e(r_{ij}) r_{ij}} + W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{-r_{ij}^2 / 2\sigma^2} \quad (6)$$

Ο πρώτος όρος αφορά τις αλληλεπιδράσεις διασποράς, απώθησης και περιγράφεται από ένα δυναμικό Lennard-Jones 6/12. Οι παράμετροι του όρου αυτού προέρχονται από το πεδίο δύναμης Amber. Ο δεύτερος όρος αφορά τους δεσμούς υδρογόνου και περιγράφεται από ένα κατευθυντικό δυναμικό 10/12. Όπου C και D είναι παράμετροι που καθορίζουν το βάθος του φρεατίου ενέργειας και έχουν τιμές που εξασφαλίζουν τη σωστή γεωμετρία των δεσμών υδρογόνου, ενώ η συνάρτηση $E(t)$ παρέχει κατευθυντικότητα ανάλογα με τη γωνία t από την ιδανική γεωμετρία του δεσμού υδρογόνου. Ο τρίτος όρος αφορά την ηλεκτροστατική αλληλεπίδραση και περιγράφεται από τη νόμο του Coulomb, όπου χρησιμοποιείται ένα ελεγμένο δυναμικό Coulomb. Ο τελευταίος όρος περιγράφει την ενέργεια αποδιάλυσης (desolvation energy), η οποία βασίζεται στον όγκο των ατόμων (V) που περιβάλλουν ένα δεδομένο άτομο και το προστατεύουν από το διαλύτη. Ο όρος αυτός σταθμίζεται από μια παράμετρο διαλυτοποίησης (S) και έναν εκθετικό όρο με συντελεστή απόστασης σ ίσο $3.5 \text{ \AA}$. Αυτοί οι ενεργειακοί όροι συνδυάζονται ώστε να υπολογιστεί η συνολική εμπειρική ελεύθερη ενέργεια πρόσδεσης ($\Delta G_{binding}$), η οποία καθορίζει τη σταθερότητα της αλληλεπίδρασης προσδέτη-υποδοχέα και αξιοποιείται για την πρόβλεψη της βέλτιστης διαμόρφωσης (pose) κατά τη μοριακή πρόσδεση.^[48]

Μετά την ολοκλήρωση της μοριακής πρόσδεσης, καταγράφηκαν οι συντεταγμένες κάθε διαμόρφωσης με τη βέλτιστη προβλεπόμενη ενέργεια στο αρχείο σύνδεσης, μαζί με πληροφορίες για τις ενεργειακές τιμές ομαδοποίησης και τις αλληλεπιδράσεις (Παράρτημα VI). Οι διαμορφώσεις αναλύθηκαν μέσω του αρχείου παραμέτρων σύνδεσης, χρησιμοποιώντας ιστόγραμμα και πίνακα RMSD για κάθε σύμπλεγμα, προκειμένου να παρουσιαστούν οι διαφορετικές διαμορφώσεις σύνδεσης και οι αντίστοιχες ελάχιστες ενεργειακές τους τιμές (Παράρτημα VII). Τα αποτελέσματα στην συνέχεια ομαδοποιήθηκαν για την αναγνώριση παρόμοιων διαμορφώσεων.

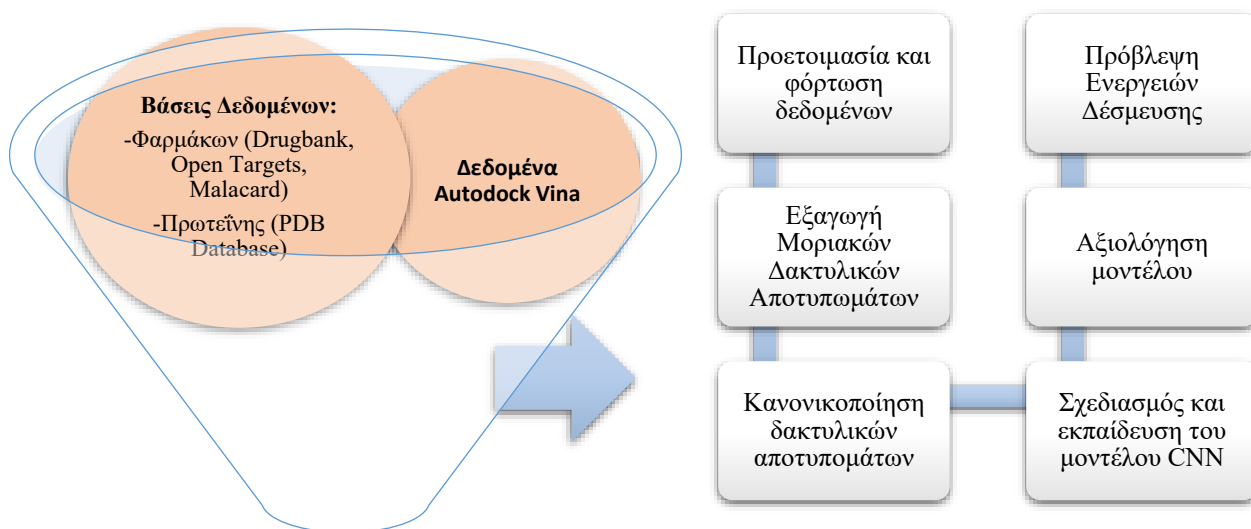


Εικόνα 3: Επισκόπηση της λειτουργίας βαθμολόγησης της ελεύθερης ενέργειας ^[41]

2.2.3 Ανάπτυξη μοντέλου Συνελικτικού Νευρωνικού Δικτύου για μοριακή πρόσδεση

Η προσέγγιση αυτή ακολουθεί μια φιλοσοφία ανάπτυξης ενός μοντέλου CNN μέσω της γλώσσας προγραμματισμού Python για την πρόβλεψη συγγενειών δέσμευσης με βάση τη δομή του προσδέτη προκειμένου να εντοπιστούν οι ενώσεις με την καλύτερη δέσμευση. Ο κώδικας είχε πρωταρχικό στόχο τη πρόβλεψη των συγγενειών δέσμευσης των μορίων σε μια πρωτεΐνη στόχο. Συγκεκριμένα, στοχεύει στην εκτέλεση της μοριακής σύνδεσης για να δημιουργήσει θέσεις δέσμευσης για προσδέτη εντός της θέσης δέσμευσης της πρωτεΐνης στόχου. Ακόμη, στοχεύει στην εκπαίδευση ενός μοντέλου CNN χρησιμοποιώντας δομές προσδέτη και αντίστοιχες συγγενείες δέσμευσης για να προβλέψει νέες συγγενείες δέσμευσης προσδέτη-πρωτεΐνης. Παρέχετε ένα πλαίσιο για την ανάλυση και την πρόβλεψη των αλληλεπιδράσεων προσδέτη-πρωτεΐνης.^[49,50]

Σκοπός είναι να βρεθούν οι ενώσεις με την ισχυρότερη δέσμευση με την πρωτεΐνη στόχο χρησιμοποιώντας το μοντέλο CNN και ακολούθως να γίνει σύγκριση των αποτελεσμάτων με τα άλλα υπολογιστικά μέσα ώστε να βρεθούν νέες ενώσεις που μπορούν να χρησιμοποιηθούν σαν φαρμακευτικοί στόχοι στην πρωτεΐνη της ασθένειας. Στην περίπτωση αυτή η προσέγγιση ποικίλει για να περιλαμβάνει εγκεκριμένα φάρμακα καθώς και αποσυρόμενα και επανασχεδιασμένα φάρμακα. Στη συνέχεια, τίθεται το ερώτημα εάν μπορούν να ανιχνευθούν υψηλές ομοιότητες μεταξύ αυτών των ήδη επικυρωμένων φαρμακευτικών δομών που δεν έχουν ακόμη συσχετιστεί με αυτοφλεγμονώδη ασθένειες. Ο τρόπος που αναπτύχθηκε και χρησιμοποιήθηκε ο κώδικας για την ανάλυση και την πρόβλεψη των συγγενειών δέσμευσης των μορίων και μιας πρωτεΐνης στόχου καθώς και τα στάδια αυτής της διαδικασίας φαίνονται στο σχήμα παρακάτω (Εικόνα 4).



Εικόνα 4: Διάγραμμα ροής που επισημαίνει τα ευδιάκριτα στάδια της προσέγγισης που ακολουθείται στην ανάπτυξη του μοντέλου CNN (βλ. Παράρτημα VIII).

Η διαδικασία που εφαρμόστηκε για την ανάπτυξη και αξιολόγηση του προτεινόμενου υπολογιστικού μοντέλου περιλαμβάνει μια ολοκληρωμένη ροή εργασίας, η οποία απεικονίζεται στο παραπάνω διάγραμμα. Πραγματοποιήθηκε η συλλογή δεδομένων από επιλεγμένες βάσεις δεδομένων φαρμάκων και βιολογικών στόχων, όπως οι DrugBank, Open Targets και Malacards, καθώς και από τη βάση Protein Data Bank (PDB) για τη δομή της πρωτεΐνης TNFR1. Τα δεδομένα αυτά ενσωματώθηκαν με τα αποτελέσματα της μοριακής πρόσδεσης μέσω του AutoDock Vina, προκειμένου να δημιουργηθεί ένα σύνολο εισόδου. Ακολούθησε η εξαγωγή των μοριακών δακτυλικών αποτυπωμάτων, δηλαδή η μετατροπή των χημικών δομών σε ψηφιακές αναπαραστάσεις κατάλληλες για χρήση σε μοντέλα μηχανικής μάθησης. Τα αποτυπώματα αυτά κανονικοποιήθηκαν ώστε να μειωθούν τυχόν αποκλίσεις και να ενισχυθεί η συνέπεια των χαρακτηριστικών. Με βάση τα προεπεξεργασμένα δεδομένα, δημιουργήθηκε και εκπαιδεύτηκε ένα μοντέλο CNN, το οποίο σχεδιάστηκε για να μαθαίνει τη συσχέτιση μεταξύ της μοριακής δομής και της ενέργειας δέσμευσης των μορίων. Μετά την εκπαίδευση, το μοντέλο αξιολογήθηκε με σκοπό να ελεγχθεί η ακρίβεια και η αξιοπιστία των προβλέψεών του σε σχέση με τα πειραματικά ή υπολογιστικά δεδομένα αναφοράς. Σκοπός η εφαρμογή του μοντέλου για την πρόβλεψη των ενεργειών δέσμευσης νέων υποψηφίων ενώσεων, συμβάλλοντας στην επιτάχυνση της διαδικασίας αναγνώρισης μορίων με υψηλή συγγένεια πρόσδεσης προς τον υποδοχέα TNFR1 και στην προοπτική σχεδιασμού νέων φαρμακολογικών παρεμβάσεων για ασθένειες που σχετίζονται με τον άξονα TNF–TNFR1, όπως οι TRAPS.

Συγκεκριμένα, για την ανάπτυξη του αλγορίθμου, οι ενώσεις φορτώθηκαν ως αρχεία της μορφής SDF (Structure Data File), και εξήχθησαν οι αναπαραστάσεις SMILES (Simplified Molecular Input Line Entry System). Τα SMILES είναι ένας γραμμικός τρόπος αναπαράστασης των χημικών μορίων χρησιμοποιώντας αλφαριθμητικά σύμβολα. Δηλαδή, μετατρέπει τη δομή ενός μορίου σε συμβολοσειρά χαρακτήρων, επιτρέποντας την εύκολη αποθήκευση και επεξεργασία της σε προγράμματα χημειοπληροφορικής. Ο κώδικας σχεδιάστηκε έτσι ώστε να διαχωρίζει τα δεδομένα σε τρία σύνολα, εκπαίδευσης, επικύρωσης και δοκιμής. Τα SMILES strings μετατράπηκαν σε μοριακά δακτυλικά αποτυπώματα Morgan, τα οποία αποτελούν έναν τύπο κυκλικών αποτυπωμάτων, και χρησιμοποιήθηκαν ως είσοδοι για το μοντέλο CNN αφότου μετατράπηκαν σε πίνακες NumPy και οργανώθηκαν στα αντίστοιχα σύνολα δεδομένων, συνοδευόμενα από τις τιμές-στόχους (συγγένειες δέσμευσης).^[42] Για τη βελτιστοποίηση της διαδικασίας εκπαίδευσης, τα δεδομένα κανονικοποιήθηκαν με τη Min-Max κλίμακα, εξασφαλίζοντας ότι όλα τα χαρακτηριστικά βρίσκονται σε παρόμοια κλίμακα, διευκολύνοντας τη σύγκλιση του νευρωνικού δικτύου (βλ. Παράρτημα VIII).

Η αρχιτεκτονική του μοντέλου CNN ορίστηκε και υλοποιήθηκε χρησιμοποιώντας το πακέτο Keras όπως φαίνεται στην Εικόνα 5.^[42-45] Το μοντέλο αποτελείται από συνελκτικά στρώματα (convolutional layers), στρώματα συγκέντρωσης (pooling layers) και πλήρως συνδεδεμένα στρώματα (fully connected layers). Το πρώτο συνελκτικό επίπεδο (Conv1D) έχει 32 φίλτρα, μέγεθος πυρήνα 3 και ενεργοποίηση ReLU ως συνάρτηση ενεργοποίησης, ώστε να ανιχνεύει τοπικά μοριακά μοτίβα στα μοριακά αποτυπώματα και να καταγράφει βασικά χαρακτηριστικά σε κάθε δακτυλικό αποτύπωμα. Για την εκπαίδευση του μοντέλου, εφαρμόστηκε η πρόωρη διακοπή (early stopping), διακόπτοντας τη διαδικασία όταν δεν υπήρχε περαιτέρω βελτίωση της απώλειας επικύρωσης, μειώνοντας έτσι τον κίνδυνο υπερπροσαρμογής (overfitting). Η υπομονή (patience) ορίστηκε σε 3 εποχές, ενώ το μοντέλο εκπαιδεύτηκε έως και 100 εποχές (epochs) με μέγεθος 32 (Εικόνα 5). Χρησιμοποιήθηκε ο βελτιστοποιητής Adam, γνωστός για τη διαχείριση του προσαρμοστικού ρυθμού μάθησης και την αποτελεσματικότητα σε πολύπλοκα μοντέλα (βλ. Παράρτημα VIII).

Η αξιολόγηση της απόδοσης του εκπαιδευμένου μοντέλου πραγματοποιήθηκε στο σύνολο δοκιμής, χρησιμοποιώντας μετρήσεις όπως το μέσο τετραγωνικό σφάλμα (Mean Squared Error, MSE), η ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Squared Error, RMSE), το μέσο απόλυτο σφάλμα (Mean Absolute Error, MAE) και ο συντελεστής προσδιορισμού (R^2 score). Ο συντελεστής R^2 (R-squared) υποδεικνύει την ικανότητα του μοντέλου να εξηγεί τη διακύμανση της μεταβλητής-στόχου. Τιμές κοντά στο 1 υποδηλώνουν καλύτερη προσαρμογή του μοντέλου στα δεδομένα. Το εκπαιδευμένο μοντέλο χρησιμοποιήθηκε για την πρόβλεψη της συγγένειας δέσμευσης διαφόρων μορίων (ligands) με συγκεκριμένες πρωτεΐνες. Επιπλέον, η ίδια διαδικασία μπορεί να εφαρμοστεί και σε άλλα σύνολα δεδομένων προσδετών ή πρωτεϊνών, επιτρέποντας τη γενίκευση των προβλέψεων σε διαφορετικούς συνδυασμούς προσδέτη-πρωτεΐνης. Αυτή η προσέγγιση παρέχει ένα βελτιστοποιημένο υπολογιστικό εργαλείο για την ανακάλυψη νέων

φαρμάκων, καθώς και για την αξιολόγηση της συγγένειας δέσμευσης νέων μορίων σε ποικίλες πρωτεΐνες-στόχους.^[49,50]

```
model = Sequential()
model.add(Conv1D(32, kernel_size=3, activation='relu',
input_shape=(X_train.shape[1], 1)))
model.add(MaxPooling1D(pool_size=2))
model.add(Flatten())
model.add(Dropout(0.2))
model.add(Dense(64, activation='relu', kernel_regularizer=regularizers.l2(0.01)))
model.add(Dense(1, activation='linear', kernel_regularizer=regularizers.l2(0.01)))
model.compile(loss='mean_squared_error', optimizer='adam')
history = model.fit(X_train.reshape(X_train.shape[0], X_train.shape[1], 1), y_train,
validation_data=(X_valid.reshape(X_valid.shape[0],
X_valid.shape[1], 1), y_valid), epochs=100, batch_size=32,
callbacks=[early_stopping])
mse = model.evaluate(X_test, y_test)
print("Mean Squared Error:", mse)
```

Εικόνα 5: Ορισμός και εκπαίδευση του 1D CNN με χρήση Keras/TensorFlow για την πρόβλεψη των ενεργειών πρόσδεσης.

Για την πρόβλεψη της συγγένειας δέσμευσης χρησιμοποιήθηκε το εκπαιδευμένο μοντέλο. Τα αρχεία SDF των υποψήφιων προσδετών διαβάστηκαν με τον SDMolSupplier, εξήχθησαν κανονικοποιημένες αναπαραστάσεις SMILES και στη συνέχεια παράχθηκαν δυαδικά δακτυλικά αποτυπώματα Morgan με ακτίνα 2 (Εικόνα 6). Τα αποτυπώματα αυτά κανονικοποιήθηκαν μέσω ενός εκπαιδευμένου scaler, ώστε να διασφαλιστεί η ομοιομορφία εισόδου, και τροφοδοτήθηκαν στο 1D CNN που είχε εκπαιδευτεί προηγουμένως σε δεδομένα docking. Το μοντέλο παρήγαγε εκτιμώμενες τιμές συγγένειας (kcal/mol) για κάθε μόριο, οι οποίες αποθηκεύτηκαν σε αρχείο Excel μαζί με τα αναγνωριστικά των ενώσεων και τα αντίστοιχα SMILES (Παράρτημα IX). Η διαδικασία αυτή επέτρεψε την ταχεία και αυτοματοποιημένη αξιολόγηση μεγάλου αριθμού προσδετών με ελάχιστη υπολογιστική επιβάρυνση, προσφέροντας ένα αξιόπιστο συμπληρωματικό εργαλείο σε σχέση με τις παραδοσιακές προσομοιώσεις docking.

```
new_ligands['Mol'] = new_ligands['SMILES'].apply(lambda x: Chem.MolFromSmiles(x))
new_ligands['Mol_Hs'] = new_ligands['Mol'].apply(lambda x: Chem.AddHs(x) if x is not None
else None)
new_ligands['MorganFingerprint'] = new_ligands['Mol_Hs'].apply(lambda x:
AllChem.GetMorganFingerprintAsBitVect(x, 2) if x is not None else None)
X_new = np.array(list(new_ligands['MorganFingerprint']), dtype=object)
X_new_filtered = [x for x in X_new if x is not None]
max_length = max(len(x) for x in X_new_filtered)
X_new_padded = [np.pad(x, (0, max_length - len(x)), mode='constant') for x in
X_new_filtered]
X_new_normalized = scaler.transform(np.vstack(X_new_padded))
predictions = model.predict(X_new_normalized.reshape(X_new_normalized.shape[0],
X_new_normalized.shape[1], 1))
new_ligands['Predicted Affinity'] = predictions
output_file_path = 'predicted_affinities.xlsx'
new_ligands.to_excel(output_file_path, index=False)
```

Εικόνα 6: Δημιουργία δακτυλικών αποτυπωμάτων Morgan, κανονικοποίηση και πρόβλεψη συγγένειας πρόσδεσης με το εκπαιδευμένο CNN.

3. Αποτελέσματα

3.1. Απόδοση και Αξιολόγηση του Συνελικτικού Νευρωνικού Δικτύου

Στη παρούσα διπλωματική εργασία πραγματοποιήθηκε ανάλυση της μοριακής πρόσδεσης περίπου 7,000 ενώσεων με τη χρήση του AutoDock Vina έναντι της πρωτεΐνης που σχετίζεται με το σύνδρομο TRAPS, και συγκεκριμένα του υποδοχέα TNFR1 (PDB ID: 1EXT). Τα αποτελέσματα αυτής της διαδικασίας αποτέλεσαν τη βάση για την ανάπτυξη του εκπαιδευμένου μοντέλου CNN για την αξιολόγηση της μοριακής πρόσδεσης μικρών μορίων έναντι της πρωτεΐνης που εφαρμόστηκε στην μελέτη. Για την υλοποίηση του CNN αξιοποιήθηκε η γλώσσα προγραμματισμού Python, ενώ η επεξεργασία και η προετοιμασία των μοριακών δεδομένων βασίστηκε στο πακέτο RDKit. Η αρχιτεκτονική και η εκπαίδευση του μοντέλου πραγματοποιήθηκαν μέσω του πακέτου TensorFlow/Keras, το οποίο παρέχει τα απαραίτητα εργαλεία για τον σχεδιασμό και τη βελτιστοποίηση του μοντέλου. Αφού ολοκληρώθηκε η δημιουργία του συνόλου δεδομένων και ο σχεδιασμός της αρχιτεκτονικής του μοντέλου, ακολούθησε η διαδικασία εκπαίδευσης, η οποία ανέδειξε τα πρώτα συμπεράσματα σχετικά με την απόδοση και τη σταθερότητα του εκπαιδευμένου μοντέλου CNN.

Κατά την εκπαίδευση του μοντέλου CNN, παρατηρήθηκε σταδιακή και σταθερή μείωση του σφάλματος εκπαίδευσης (Training Loss), από την αρχική τιμή στα 3.65 έως 0.59 στα τελικά στάδια της διαδικασίας (Εικόνα 7). Τα epochs, δηλαδή οι πλήρεις επαναλήψεις της εκπαίδευσης πάνω σε ολόκληρο το σύνολο δεδομένων, κατέδειξαν ότι το μοντέλο μάθαινε αποτελεσματικά τις συσχετίσεις μεταξύ των μοριακών χαρακτηριστικών και των αντίστοιχων ενεργειών δέσμευσης. Αντίστοιχη συμπεριφορά παρατηρήθηκε και στο σφάλμα επαλήθευσης (Validation Loss), το σφάλμα δηλαδή που υπολογίζεται σε σύνολο δεδομένων το οποίο δεν συμμετέχει στην εκπαίδευση, αλλά χρησιμοποιείται για την εκτίμηση της ικανότητας γενίκευσης του μοντέλου. Το σφάλμα επαλήθευσης εμφάνισε σημαντική μείωση κατά τα πρώιμα έως μεσαία στάδια της εκπαίδευσης, φθάνοντας στο ελάχιστο (περίπου 0.84), πριν αρχίσει να παρουσιάζει μια ήπια ανοδική τάση. Η συμπεριφορά αυτή αποτελεί χαρακτηριστική ένδειξη της έναρξης φαινομένου υπερπροσαρμογής (overfitting), κατά το οποίο το μοντέλο αρχίζει πλέον να προσαρμόζεται υπερβολικά στα δεδομένα εκπαίδευσης εις βάρος της γενίκευσής του.

Η τελική απόδοση του μοντέλου σε ένα ανεξάρτητο σύνολο δεδομένων για την τελική αξιολόγηση του συνόλου δοκιμής ανήλθε σε τιμή μέσου τετραγωνικού σφάλματος (Mean Squared Error, MSE) ίση με 0.9555, που αντιστοιχεί σε Ρίζα Μέσου Τετραγωνικού Σφάλματος (Root Mean Square Error, RMSE) περίπου 0.98 kcal/mol, τιμή η οποία θεωρείται ιδιαίτερα ικανοποιητική για προβλέψεις ενεργειών δέσμευσης. Το χαμηλό σφάλμα υποδεικνύει ότι το μοντέλο πέτυχε σχετικά υψηλή ακρίβεια γενίκευσης, καθώς οι προβλεπόμενες τιμές βρίσκονται σε καλή συμφωνία με εκείνες που παρήχθησαν από τα προγράμματα μοριακής πρόσδεσης AutoDock Vina και AutoDock4.2. Η καμπύλη εκπαίδευσης και επαλήθευσης επιβεβαιώνει την ορθή σύγκλιση του μοντέλου, με σαφή διαχωρισμό μεταξύ της φάσης μάθησης και του σημείου βέλτιστης γενίκευσης.

Το βέλτιστο σημείο επιτεύχθηκε στα epochs όπου το σφάλμα επαλήθευσης έφτασε το ελάχιστο, προτού αρχίσει η απόδοση να σταθεροποιείται. Αυτό υποδεικνύει ότι το εκπαιδευόμενο μοντέλο CNN κατάφερε να αποτυπώσει τις μη γραμμικές συσχετίσεις μεταξύ των μοριακών δακτυλικών αποτυπωμάτων και των τιμών ΔG χωρίς να υπερπροσαρμόζεται στα δεδομένα εκπαίδευσης.

```
Epoch 1/100
66/66 [=====] - 9s 103ms/step - loss: 3.6534 - val_loss: 2.4248
Epoch 2/100
66/66 [=====] - 6s 94ms/step - loss: 0.8765 - val_loss: 2.1582
Epoch 3/100
66/66 [=====] - 6s 95ms/step - loss: 0.6768 - val_loss: 1.7682
Epoch 4/100
66/66 [=====] - 6s 93ms/step - loss: 0.6141 - val_loss: 1.5353
Epoch 5/100
66/66 [=====] - 6s 94ms/step - loss: 0.5794 - val_loss: 1.3602
Epoch 6/100
66/66 [=====] - 6s 92ms/step - loss: 0.6046 - val_loss: 1.7660
Epoch 7/100
66/66 [=====] - 6s 95ms/step - loss: 0.6303 - val_loss: 0.8403
Epoch 8/100
66/66 [=====] - 6s 95ms/step - loss: 0.6286 - val_loss: 1.9430
Epoch 9/100
66/66 [=====] - 6s 97ms/step - loss: 0.5809 - val_loss: 1.2379
Epoch 10/100
66/66 [=====] - 6s 94ms/step - loss: 0.5888 - val_loss: 1.8531
14/14 [=====] - 0s 15ms/step - loss: 0.9555
Mean Squared Error: 0.9554837942123413
```

Εικόνα 7: Η απόδοση του εκπαιδευόμενου μοντέλου CNN, όπου χαρακτηρίστηκε από σταδιακή και σταθερή μείωση τόσο του σφάλματος εκπαίδευσης όσο και του σφάλματος επαλήθευσης.

Η συσχέτιση μεταξύ των προβλεπόμενων και των πραγματικών τιμών ενεργειών δέσμευσης παρουσίασε υψηλό βαθμό συμφωνίας, με τις αποκλίσεις να παραμένουν κυρίως εντός του 1 kcal/mol. Η γραμμική παλινδρόμηση των αποτελεσμάτων έδειξε συντελεστή συσχέτισης (R^2) κοντά στο 0.89, υποδεικνύοντας ισχυρή συσχέτιση μεταξύ των προβλέψεων του μοντέλου και των υπολογισμένων τιμών από το AutoDock Vina. Σε σύγκριση με τις μεθόδους μοριακής πρόσδεσης, το CNN εμφάνισε ταχύτερη και πιο αυτοματοποιημένη πρόβλεψη των ενεργειών δέσμευσης, επιτρέποντας την αξιολόγηση μεγάλου αριθμού ενώσεων σε μικρότερο χρόνο. Παρότι δεν αντικαθιστά πλήρως τις παραδοσιακές μεθόδους, η συνδυαστική χρήση CNN και AutoDock παρείχε πιο ολοκληρωμένα και συνεπή αποτελέσματα, ενισχύοντας την αξιοπιστία των προβλέψεων. Συνολικά, τα αποτελέσματα δείχνουν ότι το CNN μοντέλο μπορεί να αποτελέσει αποτελεσματικό εργαλείο πρόβλεψης συγγενειών πρόσδεσης, ικανό να υποστηρίξει διαδικασίες προκαταρκτικού ελέγχου (virtual screening) και επαναπροσδιορισμού φαρμάκων (drug repurposing) σε πρωτεΐνες που σχετίζονται με αυτοφλεγμονώδεις νόσους, όπως ο υποδοχέας TNFR1 στο σύνδρομο TRAPS.

3.2 Αποτελέσματα Μοριακής Πρόσδεσης

Οι εκατό ενώσεις με τις χαμηλότερες τιμές ελεύθερης ενέργειας πρόσδεσης που προέκυψαν από το AutoDock Vina, υποβλήθηκαν σε περαιτέρω διερεύνηση μέσω του AutoDock4.2 για πιο

λεπτομερή υπολογισμό της συγγένειας (Παράρτημα VII). Για την καλύτερη κατανόηση των αλληλεπιδράσεων και την αναγνώριση αντιπροσωπευτικών δομικών μοτίβων, οι τελικές ενώσεις ταξινομήθηκαν μέσω ομαδοποίησης, προκειμένου να αναδειχθούν αντιπροσωπευτικά δομικά μοτίβα και να αποκτηθεί μια πιο ολοκληρωμένη εικόνα των μορίων πρόσδεσης. Η διαδικασία ομαδοποίησης των εκατό ενώσεων με τη αποτελεσματικότερη πρόσδεση ανέδειξε επιμέρους ομάδες με κοινά δομικά χαρακτηριστικά. Μέσα από αυτήν την ανάλυση κατέστη δυνατή η επιλογή αντιπροσωπευτικών μορίων από κάθε ομάδα, γεγονός που παρείχε μια πιο ολοκληρωμένη εικόνα αναφορικά με τη δομή των μορίων. Η προσέγγιση αυτή συνέβαλε στην ανάδειξη χαρακτηριστικών δομικών μοτίβων (scaffolds) που θα μπορούσαν να αποτελέσουν τη βάση για περαιτέρω μελέτη και βελτιστοποίηση. Η συσχέτιση αυτή μπορεί να αναδείξει κοινά χαρακτηριστικά πρόσδεσης και να ενισχύσει την ανάγκη για συγκριτική μελέτη των συγκεκριμένων ενώσεων, ώστε να διερευνηθεί εάν παρουσιάζουν παρόμοιο μηχανισμό δράσης ή εάν μπορούν να αποτελέσουν εναλλακτικά σημεία για περαιτέρω βελτιστοποίηση.

Η σύγκριση μεταξύ των τιμών συγγένειας που προέκυψαν από το AutoDock Vina και τον κώδικα CNN κατέδειξε μικρές αποκλίσεις, με μέση διαφορά -0.62 kcal/mol και μέτρια θετική συσχέτιση ($r = 0.40$). Τα αποτελέσματα αυτά καταδεικνύουν ότι οι δύο προσεγγίσεις παρουσιάζουν παρόμοιες τάσεις στην εκτίμηση της συγγένειας πρόσδεσης για το σύνολο των εξεταζόμενων ενώσεων. Αντίθετα, τα αποτελέσματα του AutoDock4.2 εμφάνισαν μεγαλύτερες αποκλίσεις τόσο από το AutoDock Vina όσο και από το CNN, με μέση διαφορά περίπου -2 kcal/mol και σχεδόν μηδενική ή και αρνητική συσχέτιση (τιμές r περίπου -0.1 έως -0.2). Συνεπώς, παρότι οι τρεις μέθοδοι κατέγραψαν παρόμοιες τάσεις για ορισμένες ενώσεις, η συνολική εικόνα δείχνει ότι το AutoDock Vina και το CNN παρουσιάζουν μεγαλύτερη συμφωνία μεταξύ τους, ενώ το AutoDock4.2 αποκλίνει περισσότερο. Αυτό μπορεί να εξηγηθεί από το γεγονός ότι το CNN εκπαιδεύτηκε με δεδομένα του Vina, ενώ το AutoDock4.2 βασίζεται σε διαφορετικό ενεργειακό μοντέλο και αλγόριθμο, με συνέπεια οι εκτιμήσεις του να διαφοροποιούνται.

Με βάση τις τρεις μεθόδους μοριακής πρόσδεσης που εφαρμόστηκαν (AutoDock Vina, AutoDock 4.2 και CNN), πραγματοποιήθηκε συστηματική σύγκριση της συγγένειας πρόσδεσης των υποψηφίων μορίων στο σύνολο της επιφάνειας της πρωτεΐνης TNFR1. Η ανάλυση των ενεργειών δέσμευσης και των γεωμετριών πρόσδεσης επέτρεψε την κατάταξη των ενώσεων και την αναγνώριση χαρακτηριστικών προτύπων που αντανακλούν τη σταθερότητα και την ποιότητα των αλληλεπιδράσεων σε κάθε μέθοδο. Συγκεκριμένα, στην ανάλυση με το AutoDock Vina, η ένωση με την καλύτερη ενέργεια δέσμευσης ήταν η DB03219 με τιμή -10.6 kcal/mol, ακολουθεί η ένωση DB07072 με ενέργεια -8.6 kcal/mol, ενώ η DB01410 εμφάνισε τη λιγότερο αρνητική τιμή, -8.0 kcal/mol, και επομένως θεωρείται η ασθενέστερη από τις τρεις σε σχέση με τη συγγένεια πρόσδεσης (Παράρτημα VII). Κατά τη μοριακή πρόσδεση με το AutoDock4.2, η ένωση DB01199 κατέγραψε την πλέον ευνοϊκή ενέργεια δέσμευσης, -11.01 kcal/mol. Η ένωση DB04708 κατατάχθηκε στη μεσαία θέση με τιμή -6.89 kcal/mol, ενώ η DB02977 παρουσίασε θετική τιμή, $+1.06$ kcal/mol, γεγονός που καταδεικνύει ασθενή ή μη ευνοϊκή συγγένεια με την πρωτεΐνη. Κατά την αξιολόγηση με το μοντέλο CNN, η DB03219 παρουσίασε την πιο ευνοϊκή τιμή ενέργειας

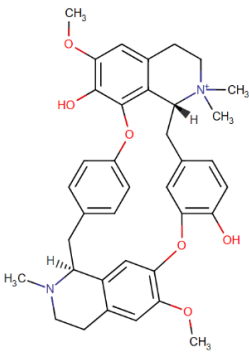
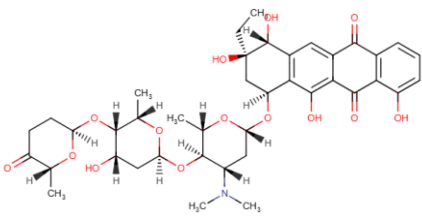
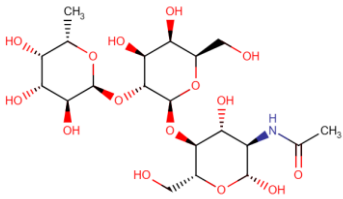
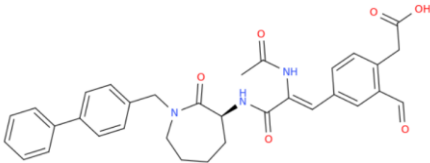
δέσμευσης, -10.3795 kcal/mol. Αντίθετα, η DB08300 εμφάνισε τη λιγότερο αρνητική τιμή (-5.4135 kcal/mol), ενώ η DB01339 κατατάχθηκε σε ενδιάμεση θέση με -8.2185 kcal/mol. Τα αναλυτικά δεδομένα παρουσιάζονται στο Πίνακα 3.

Πίνακας 3. Τιμές ελεύθερης ενέργειας πρόσδεσης (kcal/mol) των επιλεγμένων ενώσεων, όπως προέκυψαν από τις αναλύσεις με AutoDock Vina, AutoDock 4.2 και το μοντέλο CNN.

Κωδικός DrugBank	Ενέργειες Δέσμευσης AutoDock Vina (kcal/mol)	Ενέργειες Δέσμευσης AutoDock4.2 (kcal/mol)	Ενέργειες Δέσμευσης CNN (kcal/mol)
DB03219	-10.6	-4.45	-10.3795
DB07072	-8.6	-9.24	-8.10286
DB01410	-8.0	-7.49	-8.44889
DB01199	-10.0	-11.01	-9.75693
DB04708	-8.5	-6.89	-8.694
DB02977	-8.7	+1.06	-8.67093
DB08300	-8.4	-7.08	-5.4135
DB01339	-8.8	-7.62	-8.2185

Στη συνολική αξιολόγηση των εκατό υποψηφίων μορίων, αναδείχθηκαν τρεις χαρακτηριστικές ενώσεις που αντιπροσωπεύουν τη καλύτερη, τη χειρότερη και τη μέτρια ένωση όσον αφορά την πρόσδεση τους στην πρωτεΐνη με τις τρεις μεθόδους (AutoDock Vina, μοντέλο CNN και AutoDock4.2). Ως καλύτερη κατατάχθηκε η ένωση DB01199, η οποία εμφάνισε εξαιρετικά αρνητικές τιμές συγγένειας και στις τρεις προσεγγίσεις (-10.0 kcal/mol στο Vina, -9.76 kcal/mol μέσω CNN και -11.01 kcal/mol στο AutoDock 4.2), με μέσο όρο περίπου -10.26 kcal/mol. Ο μέσος όρος των τιμών αυτών (-10.26 kcal/mol) χρησιμοποιήθηκε ως περιγραφικός δείκτης και υποδηλώνει ότι η συγκεκριμένη ένωση κατατάσσεται μεταξύ των ισχυρότερων υποψηφίων ανεξαρτήτως μεθόδου. Ως λιγότερη ευνοϊκή, αναγνωρίστηκε η DB04678, με αποτελέσματα που διαφοροποιήθηκαν σημαντικά μεταξύ των μεθόδων (-8.60 kcal/mol στο Vina, -7.49 kcal/mol στο CNN και 0.00 kcal/mol στο AutoDock 4.2), δίνοντας συνολικό μέσο όρο -5.36 kcal/mol που αντανακλά ασυνέπεια και υποδηλώνει αδύναμη ή μη σταθερή πρόσδεση. Στο μέσο της κατάταξης εντοπίζεται η ένωση DB03104, η οποία κατέγραψε τιμές συγγένειας -8.40 kcal/mol στο AutoDock Vina, -8.64 kcal/mol στο CNN και -6.42 kcal/mol στο AutoDock4.2. Αποδίδοντας μέσο όρο -7.82 kcal/mol, υποδηλώνοντας ότι η συγκεκριμένη ένωση αντιπροσωπεύει ένα ενδιάμεσο επίπεδο πρόσδεσης σε σχέση με τις υπόλοιπες (Πίνακας 4).

Πίνακας 4. Τιμές ελεύθερης ενέργειας πρόσδεσης (kcal/mol) για τις χαρακτηριστικές ενώσεις, όπως υπολογίστηκαν με το AutoDock Vina, το μοντέλο CNN και το AutoDock 4.2.

Κωδικός DrugBank	Δομή Μορίου	Ενέργειες Δέσμευσης AutoDock Vina (kcal/mol)	Ενέργειες Δέσμευσης CNN (kcal/mol)	Ενέργειες Δέσμευσης AutoDock 4.2 (kcal/mol)
DB01199		-10 kcal/mol	-9.75693 kcal/mol	-11.01 kcal/mol
DB03219		-10.6 kcal/mol	-10.3795 kcal/mol	-4.45 kcal/mol
DB04678		-8.60 kcal/mol	-7.49 kcal/mol	0.00 kcal/mol
DB03104		-8.40 kcal/mol	-8.64 kcal/mol	-6.42 kcal/mol

Η σύγκριση αυτή των αποτελεσμάτων ανέδειξε την DB01199 ως τον πλέον υποσχόμενο υποψήφιο, δεδομένης της ευνοϊκής συγγένειας πρόσδεσης που καταγράφηκε και στις τρεις μεθόδους. Η ένωση DB01199 μπορεί να αποτελέσει αντικείμενο περαιτέρω μελέτης μέσω

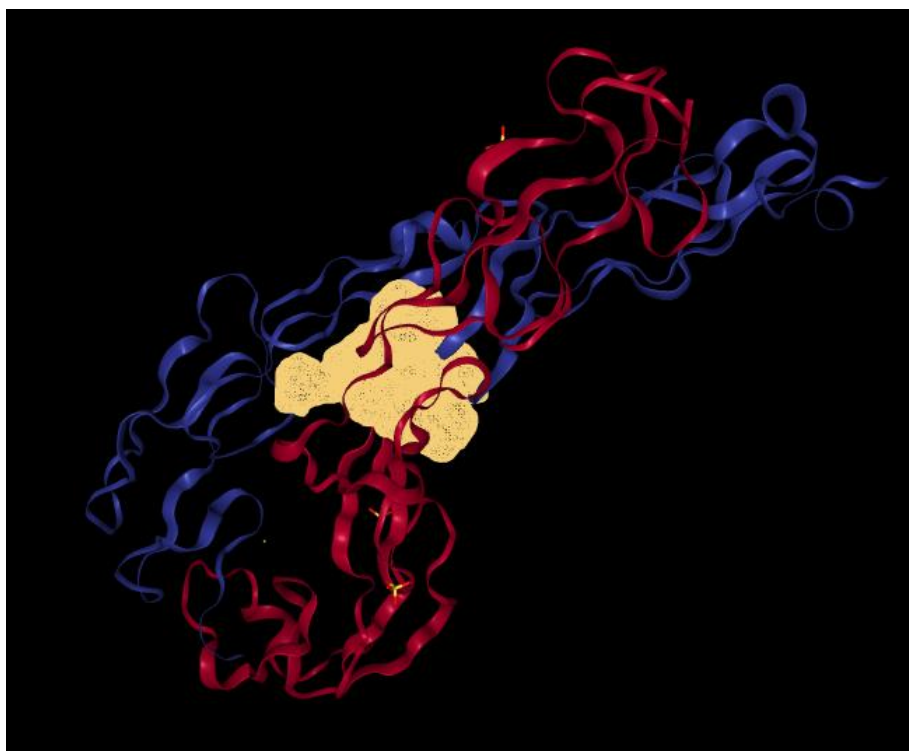
προσομοιώσεων μοριακής δυναμικής και πειραματικής επικύρωσης. Αντίθετα, η DB04678 παρουσίασε ασυνεπή και μη ευνοϊκά αποτελέσματα. Η ένωση DB03104 παρουσίασε ενδιάμεσες τιμές συγγένειας και μπορεί να αξιοποιηθεί ως σημείο αναφοράς για τη βελτιστοποίηση δομικών παραλλαγών καθώς παρουσιάζει περιθώριο βελτίωσης χωρίς να θεωρηθεί αποτυχημένος υποψήφιος. Επιπλέον, η ένωση DB03219 η οποία εμφάνισε τις καλύτερες τιμές πρόσδεσης στο AutoDock Vina (-10.6 kcal/mol) και στο μοντέλο CNN (-10.38 kcal/mol), στοιχείο που την καθιστά επίσης ενδιαφέρον υποψήφιο για περαιτέρω διερεύνηση με μοριακή δυναμική και πειραματική αξιολόγηση (βλ. Πίνακα 4). Ωστόσο, στο AutoDock4.2 η εκτίμηση συγγένειας της DB03219 ήταν αισθητά χαμηλότερη (-4.45 kcal/mol), γεγονός που υποδεικνύει ότι η αξιολόγηση της ένωσης απαιτεί προσοχή και περαιτέρω επικύρωση, ώστε να διευκρινιστεί αν οι υψηλές τιμές στα άλλα δύο μοντέλα αντανakλούν πραγματική βιολογική δραστικότητα ή περιορισμούς της μεθοδολογίας.

3.3 Ανάλυση Θέσεων Πρόσδεσης των Υποψηφίων Ενώσεων και των Κοιλοτήτων της Πρωτεΐνης TNFR1

Η διερεύνηση της τοπογραφίας του TNFR1 με τα εργαλεία CASTp και DoGSiteScorer αποτέλεσε κρίσιμο στάδιο για τον εντοπισμό των δομικών χαρακτηριστικών της πρωτεΐνης και την αναγνώριση πιθανών φαρμακολογικά αξιοποιήσιμων κοιλοτήτων της. Και οι δύο μέθοδοι ανέδειξαν με συνέπεια την παρουσία ενός μεγάλου, καλά διαμορφωμένου θυλάκου στην επιφάνεια του TNFR1, ο οποίος εντοπίζεται στη διεπαφή των περιοχών CRD2–CRD3 και αντιστοιχεί στο γνωστό σημείο πρόσδεσης του TNF- α , όπως έχει περιγραφεί στη διεθνή βιβλιογραφία. Ο συγκεκριμένος θύλακας παρουσιάζει σημαντικό όγκο και ευρεία πρόσβαση, χαρακτηριστικά που καθιστούν τη δομή του φαρμακολογικά αξιοποιήσιμη, επιτρέποντας τη σταθερή πρόσδεση μικρών μορίων. Η ανάλυση των αμινοξικών κατάλοιπων που διαμορφώνουν την κοιλοότητα αυτή ανέδειξε κατ' επανάληψη κατάλοιπα που αποτελούν θεμελιώδη στοιχεία της φυσιολογικής θέσης πρόσδεσης του TNF- α . Τα εργαλεία CASTp και DoGSiteScorer ταυτοποίησαν ως βασικά διαμορφωτικά στοιχεία του κύριου θυλάκου τα κατάλοιπα Ser74, Lys75, Arg77, Val91, Asn92, Leu93, Ser95, Glu109, Asn110, Leu111 και Phe124 της αλυσίδας A του TNFR1 (Εικόνα 6). Μεταξύ αυτών ξεχωρίζουν τα Ser74, Lys75 και Arg77 του CRD2, καθώς και τα Glu109, Asn110 και Leu111 του CRD3, τα οποία επιβεβαιώθηκαν και από τα δύο εργαλεία ως τα κεντρικά δομικά στοιχεία του ενεργού θυλάκου (Εικόνα 8). Ο συνδυασμός αυτών των καταλοίπων δημιουργεί έναν ευρύ και βιολογικά λειτουργικό χώρο πρόσδεσης που έχει ήδη αναγνωριστεί ως ο κύριος στόχος για φαρμακολογική παρεμβολή, τόσο σε μελέτες αναστολής της αλληλεπίδρασης TNF α –TNFR1 όσο και σε *in silico* αναλύσεις προηγούμενων ερευνών.

Η σύγκλιση των αποτελεσμάτων των CASTp και DoGSiteScorer, ως προς το μέγεθος, τον όγκο, τη γεωμετρία και τη σύσταση των καταλοίπων της ενεργής κοιλότητας, ενισχύει την αξιοπιστία του εντοπισμένου θύλακα και το καθιστά κατάλληλο σημείο αναφοράς για τη σύγκριση των θέσεων πρόσδεσης που προέκυψαν από τους υπολογισμούς της μοριακής πρόσδεσης. Αποτελεί τον κύριο λειτουργικό χώρο πρόσδεσης και τον σημαντικότερο στόχο για την *in silico* διερεύνηση

μικρών μορίων. Με βάση αυτήν τη θεμελιωμένη ταυτοποίηση, εφαρμόστηκε τυφλή μοριακή πρόσδεση (blind docking), με το πλέγμα να καλύπτει ολόκληρη την επιφάνεια της πρωτεΐνης, χωρίς προκαθορισμό συγκεκριμένης θέσης πρόσδεσης. Η επιλογή της τυφλής προσέγγισης επέτρεψε τη διερεύνηση όλων των πιθανών κοιλοτήτων, τόσο της κύριας αύλακας πρόσδεσης του TNF (περιοχή CRD2–CRD3) όσο και τυχόν δευτερευουσών ή αλλοστερικών θυλάκων που ενδέχεται να έχουν λειτουργική σημασία και δεν αποτυπώνονται πλήρως από γεωμετρικά εργαλεία ανάλυσης. Η στρατηγική αυτή κατέστησε δυνατή την ανακάλυψη εναλλακτικών θέσεων αλληλεπίδρασης, οι οποίες πιθανώς επηρεάζουν τη δομή ή τη ρύθμιση του TNFR1 και αποτελούν ενδιαφέροντες στόχους για την ανάπτυξη μικρομοριακών αναστολέων.



Εικόνα 8: Η πιθανή περιοχή πρόσδεσης μικρών μορίων στην πρωτεΐνη IEXT από την βάση PDB χρησιμοποιώντας τα προγράμματα CASTp και DoGSiteScorer.^[50,51]

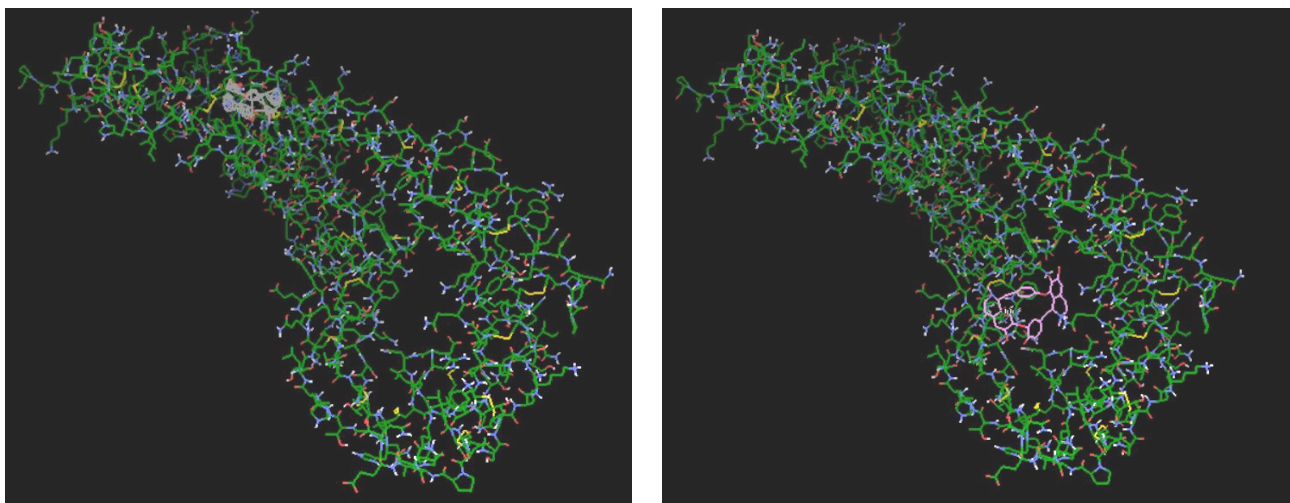
Η ταυτοποίηση του κύριου θυλάκου αποτέλεσε σημείο αναφοράς για την ερμηνεία των αποτελεσμάτων της μοριακής πρόσδεσης. Με δεδομένο ότι ο συγκεκριμένος θύλακας αντιστοιχεί στο σημείο αλληλεπίδρασης του TNF-α με τον TNFR1, αναμένονταν ότι οι ενώσεις με την υψηλότερη συγγένεια πρόσδεσης θα προσδένονταν εντός ή πλησίον της ίδιας κοιλότητας. Ωστόσο, τα αποτελέσματα της τυφλής μοριακής πρόσδεσης αποκάλυψαν ένα διαφορετικό πρότυπο. Οι ενώσεις που αναδείχθηκαν από τις τρεις μεθοδολογικές προσεγγίσεις (AutoDock Vina, AutoDock4.2 και CNN) δεν προσδέθηκαν στο κλασικό TNF-binding site, αλλά εντοπίστηκαν σε εναλλακτικές, γεωμετρικά διακριτές και λειτουργικά πιθανές αλλοστερικές κοιλότητες. Οι θέσεις αυτές βρίσκονται εκτός της αύλακας CRD2–CRD3 και σχηματίζονται από διαφορετικά αμινοξικά κατάλοιπα, υποδεικνύοντας την ύπαρξη πρόσθετων θυλάκων στην

επιφάνεια του TNFR1. Η απόκλιση αυτή μεταξύ προβλεπόμενης και πραγματικής θέσης πρόσδεσης αναδεικνύει τη δυναμική φύση της πρωτεΐνης και τη δυνατότητα μικρών μορίων να σταθεροποιούν διαμορφώσεις που δεν συνδέονται άμεσα με την κλασική ενεργό περιοχή. Στο πλαίσιο αυτό, η περαιτέρω ανάλυση των θέσεων πρόσδεσης των υποψηφίων ενώσεων καθίσταται κρίσιμη, καθώς αποκαλύπτει νέους πιθανούς μηχανισμούς ρύθμισης της λειτουργίας του TNFR1.

Η ανάλυση των θέσεων πρόσδεσης των ενώσεων που εμφάνισαν τις πιο ευνοϊκές ενεργειακές τιμές αποκάλυψε ένα σαφές μοτίβο απόκλισης από τον κλασικό θυλάκα. Χαρακτηριστικά, η ένωση DB01199, η οποία αναδείχθηκε ως η καλύτερη στις τρεις μεθοδολογικές προσεγγίσεις (AutoDock Vina, AutoDock4.2 και CNN), δεν προσδέθηκε στο γνωστό TNF-binding site. Σύμφωνα με το AutoDock4.2, η DB01199 αλληλεπιδρά στερεοχημικά με τα κατάλοιπα Gln130, Glu131, Arg104, Lys100, Cys98, Asn101 και Gln102, τα οποία βρίσκονται σε διακριτή περιοχή της εξωκυττάριας δομής του TNFR1, εκτός του Pocket 1 που αναδείχθηκε από CASTp και DoGSiteScorer. Η πρόσδεση σε αυτή την κοιλότητα, η οποία δεν ταυτίζεται με την κύρια ενεργή αύλακα της πρωτεΐνης, υποδεικνύει την πιθανή ύπαρξη λειτουργικά σημαντικού αλλοστερικού σημείου. Η υπόθεση αυτή ενισχύεται περαιτέρω από τα αποτελέσματα του AutoDock Vina, σύμφωνα με τα οποία η ίδια ένωση προσδένεται σε διαφορετικό σύνολο καταλοίπων, Gln133, Asn134, Tyr103, Phe115, His105, Gln113 και Trp107, υποδηλώνοντας ότι η DB01199 δύναται να αναγνωρίζει πολλαπλές γειτονικές κοιλότητες με παρόμοια ενέργεια δέσμευσης (Εικόνα 9). Η παρατηρούμενη μετατόπιση του σημείου πρόσδεσης μεταξύ των δύο εργαλείων αποτελεί ένδειξη ότι η συγκεκριμένη ένωση δεν δρα ως κλασικός ανταγωνιστής που καταλαμβάνει το TNF-binding site, αλλά πιθανώς ασκεί αλλοστερική ρύθμιση στη διαμόρφωση ή στη λειτουργική σταθερότητα του TNFR1.

Παρόμοια συμπεριφορά καταγράφηκε και για την ένωση DB03219, η οποία προσδένεται σε διαφορετική αλλά εξίσου διακριτή κοιλότητα, αλληλεπιδρώντας με τα κατάλοιπα His69, Tyr38, Arg53, Ser59, Glu56 και Cys70. Η θέση αυτή, επίσης απομακρυσμένη από το CRD2–CRD3, σχηματίζει έναν δευτερεύοντα θύλακα ο οποίος φαίνεται να έχει τη δυνατότητα να φιλοξενεί σταθερά μικρά μόρια, γεγονός που υποδηλώνει ότι η επιφάνεια του TNFR1 είναι περισσότερο δομικά εύκαμπτη και λειτουργικά πολυτροπική από όσο θεωρούνταν. Το κοινό χαρακτηριστικό των ενώσεων με τη βέλτιστη ενέργεια σύνδεσης είναι η σταθερή προτίμηση για μη κλασικές κοιλότητες, κάτι που προκύπτει τόσο από τη φύση της τυφλής μοριακής πρόσδεσης όσο και από τις διαφοροποιήσεις των αλγορίθμων βαθμολόγησης, οι οποίοι ενδέχεται να αναγνωρίζουν ενεργειακά ευνοϊκές αλλά μη παραδοσιακές θέσεις πρόσδεσης.

Συνολικά, τα δεδομένα αυτά υποδεικνύουν ότι οι ισχυρότεροι υποψήφιοι αναστολείς που αναγνωρίστηκαν στην παρούσα εργασία δεν στοχεύουν τη κλασική θέση του TNFα στο TNFR1, αλλά προσδένονται σε εναλλακτικούς θύλακες, πιθανώς με αλλοστερικό μηχανισμό δράσης. Η παρατήρηση αυτή ανοίγει νέες προοπτικές για την ανακάλυψη μη-ανταγωνιστικών τροποποιητών του TNFR1, οι οποίοι θα μπορούσαν να ρυθμίζουν τη λειτουργία του υποδοχέα χωρίς να παρεμβαίνουν άμεσα στη δέσμευση του TNF-α.



Εικόνα 9: Η πρόσδεση της ένωσης DB01199 στην πρωτεΐνη TNFR1 (PDB ID: 1EXT) σύμφωνα με τα αποτελέσματα του AutoDock4.2 (αριστερά) και AutoDock Vina (δεξιά)

Περαιτέρω ενίσχυση της πρόσδεσης στον TNFR1 προέκυψε από την ανάλυση της DB01336, η οποία, σύμφωνα με τα αποτελέσματα της ομαδοποίησης, παρουσιάζει δομική συγγένεια με τη DB01336. Παρότι η διαμόρφωση πρόσδεσης της DB01336 δεν ταυτίζεται με εκείνη της DB01199, εμφανίζει μια σταθερή προτίμηση προς μια γειτονική αλλοστερική περιοχή της πρωτεΐνης, αλληλεπιδρώντας με τα κατάλοιπα Phe115, Gln113, His66, Leu67, Trp107, Lys32, Glu64 και Gln117. Ιδιαίτερο ενδιαφέρον παρουσιάζει το γεγονός ότι η DB01336 μοιράζεται κρίσιμα σημεία επαφής, κυρίως τα Phe115, Gln113 και Trp107, με την ένωση DB01199, η οποία, όπως προέκυψε από την ανάλυση, προσδένεται σε κοντινό αλλά διακριτό θύλακα της πρωτεΐνης. Η επικάλυψη αυτή υποδεικνύει ότι οι δύο ενώσεις ενδέχεται να αναγνωρίζουν συμπληρωματικές περιοχές μιας ευρύτερης αλλοστερικής κοιλότητας, η οποία εκτείνεται πέρα από το παραδοσιακά θεωρούμενο ενεργό σημείο CRD2–CRD3. Η σύγκλιση των αλληλεπιδράσεων μεταξύ της των DB01336, DB03219 και DB01199, παρά τις δομικές και χημικές τους διαφορές, καταδεικνύει ότι ο TNFR1 δεν περιορίζεται σε ένα μόνο λειτουργικό θύλακα πρόσδεσης· αντίθετα, φαίνεται να διαθέτει πολλαπλές αλλοστερικές εστίες που μπορούν να φιλοξενήσουν σταθερά μικρά μόρια. Το εύρημα αυτό αποτελεί σημαντική ένδειξη ότι ο TNFR1 μπορεί να ρυθμιστεί φαρμακολογικά μέσω μη κλασικών σημείων παρέμβασης, γεγονός που διευρύνει τις δυνατότητες αναζήτησης νέων θεραπευτικών μορίων και επιβεβαιώνει την πολυπλοκότητα της επιφανειακής τοπογραφίας της πρωτεΐνης.

Η συγκριτική ανάλυση των θέσεων πρόσδεσης των τριών μορίων με την υψηλότερη συγγένεια (DB01199, DB03219 και DB01336) αποκάλυψε ότι, παρά τις μεταξύ τους δομικές διαφορές, οι ενώσεις αυτές δεν συγκλίνουν προς τον κύριο λειτουργικό θύλακα CRD2–CRD3, αλλά εντοπίζονται σε τρεις διακριτές, χωρικά γειτονικές περιοχές στην επιφάνεια του TNFR1. Η DB01199 εμφάνισε πρόσδεση σε μια κοιλότητα που περιλαμβάνει τα κατάλοιπα Gln130, Glu131, Arg104 και Lys100, ενώ η DB03219 προτίμησε μια διαφορετική περιοχή, αλληλεπιδρώντας με

υπολείμματα όπως His69, Tyr38 και Arg53. Αντίστοιχα, η DB01227 προσδέθηκε σε τρίτη, εν μέρει επικαλυπτόμενη θέση, αλληλεπιδρώντας με τα Phe115, Gln113 και Trp107, ορισμένα εκ των οποίων εμφανίζονται επίσης στις διαμορφώσεις πρόσδεσης των άλλων δύο μορίων. Η χαρτογράφηση αυτών των σημείων έδειξε ότι οι τρεις ενώσεις κατανέμονται σε περιοχές που σχηματίζουν ένα συνεκτικό σύμπλεγμα γειτονικών θυλάκων, οι οποίοι εμφανίζονται επανειλημμένα ως ενεργειακά ευνοϊκοί στα διαφορετικά docking πρωτόκολλα. Η κατανομή αυτή υποδεικνύει, σε επίπεδο υπολογιστικών παρατηρήσεων, την παρουσία περισσότερων του ενός προσβάσιμων κοιλοτήτων στην επιφάνεια του TNFR1, οι οποίες μπορούν να φιλοξενήσουν σταθερές αλληλεπιδράσεις με μικρά μόρια.

3.4 Σύγκριση με Βιβλιογραφικά Δεδομένα

Μετά την παρουσίαση των αποτελεσμάτων πρόσδεσης των ενώσεων DB01199, DB03219 και DB01227 στον TNFR1, είναι χρήσιμο να εξεταστούν τα ευρήματα προηγούμενων μελετών που έχουν επίσης διερευνήσει μικρομοριακούς αναστολείς του άξονα TNF- α /TNFR1, προκειμένου να αξιολογηθεί εάν τα υπολογιστικά αποτελέσματα της παρούσας εργασίας ευθυγραμμίζονται με τις τάσεις που αναφέρονται στη διεθνή βιβλιογραφία. Η αξιολόγηση των αποτελεσμάτων αποκτά ιδιαίτερο ενδιαφέρον όταν συγκριθεί με δεδομένα που έχουν δημοσιευτεί στη διεθνή βιβλιογραφία. Για την περεταίρω έρευνα και σύγκριση των αποτελεσμάτων των μορίων που βρέθηκαν να στοχεύουν τον υποδοχέα TNFR1, αξίζει να σημειωθεί η αναζήτηση βιβλιογραφικών δεδομένων που στοχεύουν εκτός από τον υποδοχέα TNFR1, το TNF- α με σκοπό την αναζήτηση μικρών μορίων που αναστέλλουν τη δέσμευσή του στον TNFR1. Έχουν πραγματοποιηθεί εκτενείς υπολογιστικές αναλύσεις μοριακής πρόσδεσης στο TNF- α . Οι προσεγγίσεις αυτές στοχεύουν στην ίδια τη διεπαφή πρόσδεσης TNF- α /TNFR1, συμβάλλοντας στην αποδυνάμωση του παθολογικού σήματος που χαρακτηρίζει νοσήματα με δυσλειτουργία του TNFR1, όπως οι TRAPS. Πέρα από τον υποδοχέα TNFR1, ο TNF- α έχει αποτελέσει επίσης αντικείμενο μελέτης μέσω υπολογιστικών μεθόδων docking και εικονικό έλεγχο, στοχεύοντας στην αναγνώριση μικρών μορίων που δεσμεύονται στον TNF- α και εμποδίζουν τη δημιουργία του συμπλόκου TNF- α -TNFR1.

Μετά τον εικονικό έλεγχο μορίων, σύμφωνα με μια μελέτη των Saddala et al., με στόχο τόσο την κυτοκίνη TNF- α όσο και τον υποδοχέα της TNFR1 (PDB ID: 1EXT), καθώς και το σύμπλοκο TNF- α /TNFR1 αναδείχθηκαν ενώσεις με υψηλές τιμές συγγένειας δέσμευσης.^[53] Οι ενώσεις αυτές κυμάνθηκαν κυρίως στην περιοχή των -8.0 έως -11.0 kcal/mol, όρια που θεωρούνται ενδεικτικά ισχυρής και ευνοϊκής πρόσδεσης. Οι τιμές αυτές είναι απόλυτα συγκρίσιμες με τα δεδομένα που προέκυψαν στην παρούσα εργασία, όπου οι ενώσεις με την ισχυρότερη δέσμευση κατέγραψαν συγγένειες στο ίδιο εύρος. Οι ενώσεις που χρησιμοποιήθηκαν προέρχονται από τη βάση δεδομένων ZINC, οι οποίες αρχικά φιλτραρίστηκαν μέσω φαρμακοφόρων και στη συνέχεια υποβλήθηκαν σε μοριακή πρόσδεση στον TNF- α , στον υποδοχέα TNFR1 (PDB ID: 1EXT) και στο σύμπλοκο τους. Όπως φαίνεται στους πίνακες, Πίνακα 5 και Πίνακα 6, οι κορυφαίες ενώσεις από τη βάση ZINC κατέγραψαν τιμές πρόσδεσης στην περιοχή -8.0 έως -11.0 kcal/mol, τόσο για την κυτοκίνη TNF- α όσο και για τον υποδοχέα TNFR1.^[53] Ενδιαφέρον παρουσιάζει το γεγονός

ότι οι ίδιες ενώσεις εμφάνισαν παρόμοιες τιμές πρόσδεσης σε διαφορετικούς στόχους, γεγονός που υποδηλώνει ότι η αναστολή του συμπλόκου TNF- α /TNFR1 μπορεί να επιτευχθεί με περισσότερους από έναν μηχανισμούς. Οι ενώσεις μπορούν να αλληλεπιδράσουν άμεσα με τον υποδοχέα TNFR1, με την κυτοκίνη TNF- α , καθώς και με το σύμπλοκο τους και να παρουσιάσουν σημαντική συγγένεια πρόσδεσης. Αυτό υποδεικνύει ότι η φαρμακολογική αναστολή του συμπλόκου TNF- α /TNFR1 μπορεί να επιτευχθεί μέσω πολλών μηχανισμών, όπως άμεσης δέσμευσης του TNF- α , αναστολή του υποδοχέα, είτε μέσω παρεμπόδισης της αλληλεπίδρασης στον σχηματισμό του συμπλόκου. Η πολυτροπικότητα αυτή ενισχύει τη φαρμακολογική αξία των υποψήφιων μορίων, καθώς αυξάνει την πιθανότητα επίτευξης ανασταλτικού αποτελέσματος μέσω εναλλακτικών μοριακών διαδρομών.

Τα αμινοξικά κατάλοιπα που βρέθηκαν ότι σχετίζονται με το ενεργό κέντρο και συμμετέχουν στη διαδικασία πρόσδεσης για το TNF- α είναι τα Val91, Asn92, Leu93, Phe124 της αλυσίδας A και τα His15, Val17, Ala18, Pro20, Arg32, Ala33, Asn34, Ala35, Phe144, Glu146, Ser147, Gly148, Gln149 και Val150 της αλυσίδας B. Για το σύμπλοκο TNF- α /TNFR1 τα κατάλοιπα στο ενεργό κέντρο είναι τα His15, Val17, Ala18, Pro20, Arg32, Ala33, Asn34, Ala35, Phe144, Ala145, Glu146, Ser147, Gly148, Gln149, Val150 και Tyr151 της αλυσίδας A, τα Thr77, His78, Thr79, Ser81, Pro90, Val91, Asn92, Leu93, Ser95, Ile97, Phe124, Glu135, Ile136 και Asn137 της αλυσίδας B, καθώς και τα Phe60, Leu71, Ser72, Cys73, Ser74, Lys75, Arg77, Gln82, Cys96 και Leu111 της αλυσίδας R. Οι ενεργές αυτές θέσεις θεωρούνται κρίσιμες για τη λειτουργία και τη ρύθμιση του TNF- α /TNFR1, καθώς συμμετέχουν καθοριστικά στις διεπαφές πρόσδεσης που διέπουν τη σηματοδότηση του TNF.^[53] Η λεπτομερής χαρτογράφηση αυτών των περιοχών συνέβαλε στην κατανόηση των μοριακών μηχανισμών της δυσλειτουργίας του υποδοχέα στις TRAPS και αποτελεί σημείο αναφοράς για τον σχεδιασμό νέων μικρομοριακών αναστολέων που θα μπορούσαν να ρυθμίσουν επιλεκτικά την παθολογική ενεργοποίηση της οδού TNF α -TNFR1.

Σύμφωνα με τη μελέτη των Saddala et al., αρκετά μικρά μόρια εμφάνισαν σημαντική συγγένεια πρόσδεσης με την πρωτεΐνη TNF α αλληλεπιδρώντας με βασικά αμινοξέα της ενεργού περιοχής (Πίνακας 5). Μεταξύ αυτών, το μόριο ZINC09609430 παρουσίασε την ισχυρότερη αλληλεπίδραση σχηματίζοντας δεσμό υδρογόνου με το κατάλοιπο Gly121, γεγονός που αντικατοπτρίζεται στη χαμηλότερη ενέργεια δέσμευσης (-9.2 kcal/mol) υποδηλώνοντας εξαιρετική σταθερότητα πρόσδεσης ικανή να επηρεάσει τη διαμόρφωση της πρωτεΐνης. Το μόριο ZINC49467549 κατέγραψε επίσης υψηλό βαθμό σταθερότητας (-9.0 kcal/mol), με κύριες αλληλεπιδράσεις ηλεκτροστατικής φύσης και δεσμούς Van der Waals με τα κατάλοιπα Ile58, Leu120, Gly121 και Tyr151. Αντίστοιχα, το ZINC13113075 σχημάτισε δεσμό υδρογόνου με το κατάλοιπο Tyr151, γεγονός που υποστηρίζει την ενέργεια δέσμευσης των -8.8 kcal/mol και υπογραμμίζεται η σημασία του συγκεκριμένου κατάλοιπου ως κομβικού σημείου αλληλεπίδρασης. Το ZINC39907639 εμφάνισε παρόμοιο μοτίβο αλληλεπίδρασης με τα ίδια κρίσιμα αμινοξέα της ενεργού θέσης και ενέργεια δέσμευσης -8.5 kcal/mol. Το μόριο ZINC25251930 διακρίνεται για τον σχηματισμό τριών δεσμών υδρογόνου με τα κατάλοιπα Ile58, Leu120, Gly121 και Tyr151, παρουσιάζοντας επίσης ικανοποιητική σταθερότητα (-8.1 kcal/mol).

Αντιθέτως, ο αναστολέας 307, που χρησιμοποιείται ως σημείο αναφοράς, εμφάνισε αισθητά χαμηλότερη συγγένεια (-6.8 kcal/mol) παρότι αλληλεπιδρά με τα ίδια κρίσιμα αμινοξέα, ενισχύοντας την υπεροχή των νέων υποψηφίων μορίων έναντι του υπάρχοντος αναστολέα.^[53] Συνολικά, τα αποτελέσματα αυτά αναδεικνύουν τα κατάλοιπα Ile58, Leu120, Gly121 και Tyr151 ως κεντρικούς δομικούς στόχους για την παρεμπόδιση του σχηματισμού του τριμερούς συμπλόκου του TNF- α , και ενισχύουν την προοπτική σχεδιασμού πιο αποδοτικών αναστολέων με βάση τα μόρια που εντοπίστηκαν.

Παρόμοια ευρήματα καταγράφηκαν και για την πρωτεΐνη TNFR1, καθώς τα αποτελέσματα αποκάλυψαν μόρια με ιδιαίτερα υψηλή σταθερότητα πρόσδεσης και σημαντικό αριθμό δεσμών υδρογόνου που ενισχύουν την αλληλεπίδραση με τα κρίσιμα αμινοξέα της ενεργούς περιοχής της πρωτεΐνης (Πίνακας 6).^[53] Το μόριο ZINC02968981 παρουσίασε τη μεγαλύτερη συγγένεια, σχηματίζοντας πέντε δεσμούς υδρογόνου μέσω των ομάδων ON και OC με τα κατάλοιπα Lys75, Gln82, Arg104 και Tyr106, με ενέργεια δέσμευσης -10.1 kcal/mol, καθιστώντας το μόριο αυτό ως τον πιο υποσχόμενο αναστολέα της συγκεκριμένης ομάδας. Αντίστοιχα, το ZINC09544246 εμφάνισε επίσης εξαιρετική σταθερότητα (-9.8 kcal/mol), δημιουργώντας επτά δεσμούς υδρογόνου μέσω ομάδων OC και NH με τα κατάλοιπα Glu56, Ser57, Ser59, Cys70, Cys73 και Ser74, γεγονός που καταδεικνύει την εκτεταμένη συμμετοχή των καταλοίπων της περιοχής αυτής στις αλληλεπιδράσεις. Το μόριο ZINC58047088 σχημάτισε δύο δεσμούς υδρογόνου με τα κατάλοιπα Ser74 και Asn110, παρουσιάζοντας ενέργεια δέσμευσης -9.5 kcal/mol. Το ZINC72021182 κατέγραψε έναν δεσμό υδρογόνου με το αμινοξύ Arg104, με ενέργεια -9.3 kcal/mol, γεγονός που αποδεικνύει ότι η σταθερή πρόσδεση μπορεί να επιτευχθεί και με περιορισμένο αριθμό δεσμών υδρογόνου. Το ZINC08704414 εμφάνισε τέσσερις δεσμούς υδρογόνου με τα αμινοξικά κατάλοιπα Ser74, Lys75, Arg77 και Asn110, επιτυγχάνοντας ενέργεια δέσμευσης -9.1 kcal/mol, κάτι που υποδηλώνει ευνοϊκή αλληλεπίδραση με πολλαπλά κρίσιμα αμινοξέα της πρωτεΐνης.

Τέλος, μια δεύτερη μορφή του μορίου ZINC09544246 κατέγραψε ακόμα μεγαλύτερο αριθμό δεσμών υδρογόνου, συνολικά οκτώ, με τα κατάλοιπα Arg77, Cys96, Arg104, Tyr106 και Asn110. Παρόλα αυτά, η αντίστοιχη ενέργεια πρόσδεσης ήταν σημαντικά μικρότερη (-7.6 kcal/mol), γεγονός που υποδηλώνει ότι ο αριθμός των δεσμών υδρογόνου δεν είναι πάντα ο καθοριστικός παράγοντας για την ενέργεια πρόσδεσης, αλλά η γεωμετρική διάταξη και η φύση των επιμέρους αλληλεπιδράσεων διαδραματίζουν εξίσου σημαντικό ρόλο επηρεάζοντας καθοριστικά την αποτελεσματικότητα της σύνδεσης. Συνολικά, από τα αποτελέσματα προκύπτει το μόριο ZINC02968981 ως ο πλέον ισχυρός υποψήφιος αναστολέας για τον TNFR1. Ταυτόχρονα, τα υπόλοιπα μόρια προσφέρουν πολύτιμες ενδείξεις για τη δομή και τα κρίσιμα κατάλοιπα που συμμετέχουν στις αλληλεπιδράσεις. Συγκεκριμένα, τα κατάλοιπα Lys75, Arg104, Tyr106, Ser74 και Asn110, αναδεικνύονται βασικοί στόχοι για τον περαιτέρω σχεδιασμό νέων θεραπευτικών παραγόντων.^[53]

Πίνακας 5. Αποτελέσματα μοριακής πρόσδεσης για τις κορυφαίες ενώσεις με τον υποδοχέα TNF- α .^[53]

TNF- α			
Κωδικός ZINC	Όνομα Χημικής Ένωσης	Ενέργεια Δέσμευσης (ΔG) kcal/mol ⁻¹	Αλληλεπιδράσεις Υδρογονικών Δεσμών Πρωτεΐνης– Προσδέτη
ZINC09609430	<i>N</i> -[1-(2,5-dioxabicyclo[4.4.0]deca-7,9,11-trien-8-yl)ethyl]-3-[7-(4-fluorophenyl)-2,4,8-trimethyl-1,5	– 9.2	Gly ¹²¹ CH–HN
ZINC49467549	<i>N</i> -[2-[3-(4-methyl-2-phenyl-thiazol-5-yl)-6-oxo-pyridazin-1-yl]ethyl]-2-(1-naphthyloxy)acetamide	– 9.0	Ile ⁵⁸ , Leu ¹²⁰ , Gly ¹²¹ and Tyr ¹⁵¹
ZINC13113075	(2 <i>R</i>)-2-[[5-[(1 <i>S</i>)-1-(4-fluorophenoxy)ethyl]-1,3,4-oxadiazol-2-yl]sulfanyl]-1-(2-methyl-1 <i>H</i> -indol-3-yl)	– 8.8	Tyr ¹⁵¹ CO–OC
ZINC39907639	2-[(3,4-dimethoxyphenyl)methyl]-1-[2-(1-naphthyloxy)ethyl]benzimidazole	– 8.5	Ile ⁵⁸ , Leu ¹²⁰ , Gly ¹²¹ and Tyr ¹⁵¹
ZINC25251930	1-[[(<i>R</i>)-(3-methoxyphenyl)-(1-methylimidazol-2-yl)methyl]amino]-3-methyl-pyrido[1,2- <i>a</i>]benzimidazole-4	– 8.1	Ile ⁵⁸ CO–NC
			Leu ¹²⁰ CO–NH
			Gly ¹²¹ CN–OC
307 (Query)	6,7-dimethyl-3-[(methyl{2-[methyl({1-[3-(trifluoro methyl)phenyl]-1 <i>H</i> -indol-3-yl}methyl)amino]ethyl}amino)methyl]-4 <i>H</i> -chromen-4-one	– 6.8	Ile ⁵⁸ , Leu ¹²⁰ , Gly ¹²¹ and Tyr ¹⁵¹

Πίνακας 6. Αποτελέσματα μοριακής πρόσδεσης για τις κορυφαίες ενώσεις με τον υποδοχέα TNFR1.^[53]

TNFR1			
Κωδικό ZINC	Όνομα Χημικής Ένωσης	Ενέργεια Δέσμευσης (ΔG) kcal/mol ⁻¹	Αλληλεπιδράσεις Υδρογονικών Δεσμών Πρωτεΐνης-Προσδέτη
ZINC02968981	<i>N</i> -benzyl-2-[[2-(4-nitrophenyl)-[1, 2, 4]triazolo[1,5- <i>c</i>]quinazolin-5-yl]thio]acetamide	– 10.1	Lys ⁷⁵ CO–ON Gln ⁸² CO–ON Gln ⁸² CN–ON Arg ¹⁰⁴ CN–OC Tyr ¹⁰⁶ CO–OC
ZINC09544246	2-[[5-(1 <i>H</i> -indol-3-yl)-1,3,4-oxadiazol-2-yl]sulfanyl]- <i>N</i> -(3-sulfamoylphenyl)-acetamide	– 9.8	Glu ⁵⁶ CN–OC Glu ⁵⁶ CO–NH Ser ⁵⁷ CO–NH Ser ⁵⁹ CO–OC Cys ⁷⁰ CN–OC Cys ⁷³ CO–NH Ser ⁷⁴ CO–NH
ZINC58047088	<i>N</i> -[(4-fluorophenyl)methyl]-2-[(3-methyl-2-oxo-[1, 2, 4]triazino[2,3- <i>c</i>]quinazolin-6-yl)sulfanyl]acetami	– 9.5	Ser ⁷⁴ CO–NH Asn ¹¹⁰ CO–NH
ZINC72021182	2-(4-chlorobenzoyl)- <i>N</i> -[2-[(2-hydroxybenzoyl)amino]ethyl]benzamide	– 9.3	Arg ¹⁰⁴ CN–OH
ZINC08704414	2-(methylBLAHyl)sulfanyl- <i>N</i> -(<i>p</i> -tolylmethyl)propanamide	– 9.1	Ser ⁷⁴ CO–NH Lys ⁷⁵ CO–NH Arg ⁷⁷ CN–OC Asn ¹¹⁰ CO–NH
ZINC33832439 (query)	1-Hydroxy-3-methoxy-6-methyl-8-(((2 <i>S</i> ,3 <i>R</i> ,4 <i>S</i> ,5 <i>S</i> ,6 <i>R</i>)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2 <i>H</i> -pyran-2-yl)oxy)anthracene-9,10-dione	– 7.6	Arg ⁷⁷ CN–OC Cys ⁹⁶ CN–OH Cys ⁹⁶ CO–OH Arg ¹⁰⁴ CN–OH Arg ¹⁰⁴ CN–OH Arg ¹⁰⁴ CN–OH Tyr ¹⁰⁶ CO–OH Asn ¹¹⁰ CO–OH

Όσον αφορά τους αναστολείς του συμπλόκου TNF-α/TNFR1, τα δεδομένα της βιβλιογραφίας αναδεικνύουν ορισμένα μόρια με ιδιαίτερα υψηλή συγγένεια πρόσδεσης και πολλαπλές σταθερές αλληλεπιδράσεις με κρίσιμα αμινοξικά κατάλοιπα του συμπλόκου όπως φαίνεται στον Πίνακα 7.

Πίνακας 7. Αποτελέσματα μοριακής πρόσδεσης για τις κορυφαίες ενώσεις με το σύμπλοκο TNF- α /TNFR1.^[53]

Σύμπλοκο TNF- α /TNFR1			
Κωδικό ZINC	Όνομα Χημικής Ένωσης	Ενέργεια Δέσμευσης (ΔG) kcal/mol ⁻¹	Αλληλεπιδράσεις Υδρογονικών Δεσμών Πρωτεΐνης-Προσδέτη
ZINC05462670	Chaetochromin B	- 10.0	Ser ⁷⁴ CO-OH Thr ⁹⁴ CO-OH Glu ¹⁰⁹ CO-OC Asn ¹¹⁰ CO-OH
ZINC35681945	<i>N</i> 2-[(2 <i>R</i>)-2-(dimethylamino)-2-(2-methoxyphenyl)ethyl]- <i>N</i> 4-(4-ethoxyphenyl)-5-nitro-pyrimidine-2,4,6-tr	- 9.7	Lys ⁷⁵ CO-NH Gln ⁸² CO-ON Gln ⁸² CN-ON Asp ⁹³ CO-NH Arg ¹⁰⁴ CN-OC Asn ¹¹⁰ CO-NH
ZINC23553920	6-[[4-allyl-5-(2-methoxyphenyl)-1,2,4-triazol-3-yl]sulfanylmethyl]- <i>N'</i> -(2-ethylphenyl)-1,3,5-triazine	- 9.5	Ser ⁷⁴ CO-OC Thr ⁹⁴ CO-NH Thr ⁹⁴ CO-NH
ZINC05328058	DNC011371	- 8.9	Ser ⁷⁴ CO-OH Gln ⁸² CO-OC Pro ⁹⁰ CO-OH Asn ⁹² CN-OH Cys ⁹⁶ CN-OH
ZINC17206695	(2 <i>Z</i>)-2-(2,5-dimethoxyphenyl)imino-6-[(3 <i>R</i> ,5 <i>S</i>)-3,5-dimethyl-1-piperidyl]-5-nitro-1 <i>H</i> -pyrimidin-4-amine	- 8.5	Ser ⁷⁴ CO-NH Ser ⁷⁴ CO-NH Asn ⁹³ CO-NH Asn ¹¹⁰ CN-ON Asn ¹¹⁰ CN-NO Ser ¹⁴⁷ OC-OC
ZINC08214556 (query)	Erythrosine sodium (USP)	- 7.2	Thr ⁹⁴ CO-OC Asn ¹¹⁰ CN-OC

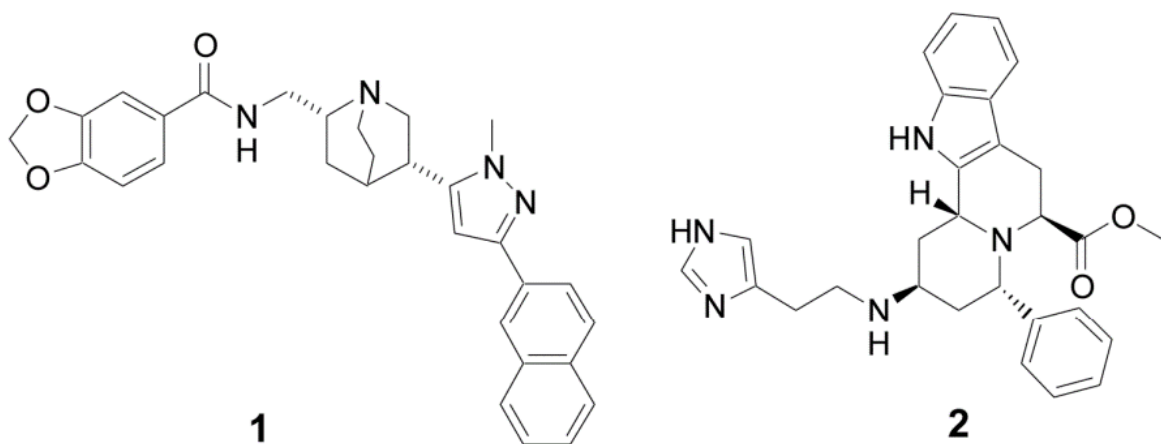
Το μόριο ZINC05462670 σχημάτισε τέσσερις δεσμούς υδρογόνου με τα κατάλοιπα Ser74, Thr94, Glu109 και Asn110, καταγράφοντας εξαιρετικά χαμηλή και επομένως ευνοϊκή ενέργεια πρόσδεσης, -10.0 kcal/mol, που το κατατάσσει στους πιο ισχυρούς υποψήφιους αναστολείς. Το ZINC35681945 παρουσίασε επίσης εξαιρετική σταθερότητα (-9.7 kcal/mol), σχηματίζοντας έξι δεσμούς υδρογόνου με τα κατάλοιπα Lys75, Gln82, Asp93, Arg104 και Asn110, γεγονός που δείχνει έναν μηχανισμό πρόσδεσης πολλαπλών επαφών. Το μόριο ZINC23553920 (-9.5 kcal/mol) σχημάτισε τρεις δεσμούς υδρογόνου με τα κατάλοιπα Ser74 και Thr94, ενώ το ZINC05328058 (-8.9 kcal/mol) δημιούργησε πέντε δεσμούς υδρογόνου με τα Ser74, Gln82, Pro90, Asn92 και Cys96, ενισχύοντας την αντίληψη ότι αυτά τα σημεία αποτελούν βασικά δομικά στοιχεία για την πρόσδεση. Επίσης, το μόριο ZINC17206695 εμφάνισε έξι δεσμούς υδρογόνου με τα κατάλοιπα Ser74, Asn93, Asn110 και Ser147, παρουσιάζοντας σταθερότητα -8.5 kcal/mol, υπογραμμίζοντας τη σημασία των κατάλοιπων Asn και Ser στην πρόσδεση.

Το μόριο ZINC08214556 το οποίο εμφάνισε μικρότερη σταθερότητα πρόσδεσης (-7.2 kcal/mol), παρά το γεγονός ότι σχημάτισε έξι δεσμούς υδρογόνου με τα κατάλοιπα Thr94 και Asn110, υποδηλώνει ότι η απλή ύπαρξη πολλών δεσμών υδρογόνου δεν εγγυάται υψηλή ενέργεια πρόσδεσης, καθώς η χωρική διάταξη και η φύση των επιμέρους αλληλεπιδράσεων διαδραματίζουν εξίσου κρίσιμο ρόλο. Συνολικά, τα δεδομένα αυτά υποδεικνύουν ότι τα κατάλοιπα Ser74, Thr94, Glu109, Asn110, Arg104 και Ser147 αποτελούν λειτουργικά κρίσιμα σημεία-κλειδιά για τη σταθεροποίηση του συμπλόκου TNF α /TNFR1 και κατά συνέπεια, για την ενεργοποίηση της σηματοδοτικής οδού του TNF α , η οποία σχετίζεται άμεσα με τις φλεγμονώδεις αποκρίσεις. Η κατανόηση αυτών των μοριακών αλληλεπιδράσεων είναι ιδιαίτερα σημαντική, καθώς η στοχευμένη αναστολή τους θα μπορούσε να συμβάλει στον περιορισμό της παθολογικής φλεγμονής που χαρακτηρίζει νοσήματα όπως οι TRAPS. Όπως φάνηκε, η αλληλεπίδραση TNF α -TNFR1 αποτελεί καθοριστικό βήμα για την ενεργοποίηση της σηματοδοτικής οδού του TNF και, συνεπώς, έναν από τους σημαντικότερους μοριακούς στόχους για την ανάπτυξη θεραπευτικών παρεμβάσεων σε αυτοφλεγμονώδεις παθήσεις. Στο πλαίσιο αυτό, ιδιαίτερο ενδιαφέρον παρουσιάζουν μελέτες που διερευνούν μικρομοριακούς αναστολείς ικανούς να παρεμποδίσουν τη δέσμευση του TNF α στον TNFR1.

Ακόμα ένα χαρακτηριστικό παράδειγμα αποτελεί η εργασία των Chan et al., στην οποία εφαρμόστηκε μια ολοκληρωμένη προσέγγιση δομικά καθοδηγούμενου σχεδιασμού φαρμάκων (structure-based drug design), συνδυάζοντας εικονικό έλεγχο μεγάλης κλίμακας (large-scale virtual screening) και μοριακή πρόσδεση.^[54] Η μελέτη αξιολόγησε περισσότερες από 20,000 ενώσεις φυσικών προϊόντων και παραγώγων τους, με στόχο τον εντοπισμό μικρών μορίων που θα μπορούσαν να αναστείλουν την πρόσδεση του TNF- α στον TNFR1 μέσω παρεμβολής στη διεπαφή του συμπλόκου. Η προσέγγιση αυτή ανέδειξε πολλαπλές ενώσεις με ιδιαίτερα ευνοϊκά ενεργειακά προφίλ, οι οποίες παρουσιάζουν σημαντικό ενδιαφέρον ως πιθανοί φαρμακολογικοί αναστολείς και μπορούν να χρησιμεύσουν ως δομικοί μοριακοί σκελετοί (scaffolds) για περαιτέρω φαρμακοχημική βελτιστοποίηση.

Η μελέτη των Chan et al. παρείχε ένα εκτεταμένο παράδειγμα εφαρμογής εικονικού ελέγχου μεγάλης κλίμακας με στόχο τον εντοπισμό φυσικών προϊόντων ικανών να παρεμποδίσουν τη δέσμευση του TNF-α στον TNFR1.^[54] Κατά τον έλεγχο περισσότερων από 20,000 ενώσεων, αρκετές υποψήφιες βρέθηκαν να παρουσιάζουν υψηλές ενέργειες δέσμευσης (−7 έως −10.2 kcal/mol) με τις πιο υποσχόμενες να συγκεντρώνουν ενέργειες δέσμευσης της τάξης των −9.5 έως −10.2 kcal/mol, υποδηλώνοντας σημαντική σταθερότητα αλληλεπίδρασης με κρίσιμα αμινοξικά κατάλοιπα της περιοχής πρόσδεσης του TNF-α.

Οι ενώσεις με την υψηλότερη κατάταξη εμφάνισαν ισχυρούς δεσμούς υδρογόνου σε συνδυασμό με υδρόφοβες αλληλεπιδράσεις με κατάλοιπα-κλειδιά της διεπαφής TNF-α/TNFR1, γεγονός που ενίσχυσε την υπολογιστική εκτίμηση της ανασταλτικής τους δράσης. Από την εικονική διαλογή επιλέχθηκαν 16 μόρια για πειραματική αξιολόγηση, εκ των οποίων δύο, μία πυραζόλη συνδεδεμένη με κινουκλιδίνη και μία ινδολ[2,3-α]κινολιζιδίνη, παρουσίασαν αξιοσημείωτη ικανότητα αναστολής της πρόσδεσης του TNF-α στον TNFR1 σε δοκιμασίες ELISA. Οι ενώσεις αυτές ήταν χημικά ετερογενείς, γεγονός που υποδηλώνει ότι διαφορετικές χημικές οικογένειες μπορούν να λειτουργήσουν ως μοριακοί σκελετοί για περαιτέρω φαρμακοχημική ανάπτυξη (Εικόνα 10).



Εικόνα 10: Οι δύο ενώσεις που προέκυψαν από την μελέτη που παρουσίασαν αξιοσημείωτη ικανότητα αναστολής της πρόσδεσης του TNF-α στον TNFR1 σε δοκιμασίες ELISA.^[54]

Η μελέτη των Chan et al. αποτελεί χαρακτηριστικό παράδειγμα, καθώς ανέδειξε φυσικές ενώσεις με συγκρίσιμες ενεργειακές τιμές (−7 έως −10 kcal/mol), επιβεβαιώνοντας ότι το σύμπλοκο TNF-α/TNFR1 μπορεί πράγματι να στοχευθεί αποτελεσματικά μέσω μικρών μορίων.^[54] Παράλληλα, έδειξε ότι διαφορετικές χημικές κατηγορίες δομών είναι ικανές να παρεμβληθούν στη διεπαφή του συμπλόκου, προσφέροντας πολυμορφία στη φαρμακοχημική βελτιστοποίηση. Τα δεδομένα

αυτά επιβεβαιώνουν ότι η χρήση της κρυσταλλικής δομής 1EXT ως υποδοχέα-στόχου αποτελεί έγκυρη και καθιερωμένη στρατηγική στη μελέτη αναστολέων του TNFR1, ενώ η ενσωμάτωση των βιβλιογραφικών ευρημάτων στο πλαίσιο της παρούσας εργασίας ενισχύει περαιτέρω την αξιολόγηση και τη συγκριτική ερμηνεία των αποτελεσμάτων.

Σύμφωνα με τα δεδομένα της παρούσας μελέτης, η ένωση DB01199 (Πίνακας 4) κατέγραψε τις πιο αρνητικές τιμές συγγένειας πρόσδεσης προς τον TNFR1 (PDB ID: 1EXT), με ενέργεια περίπου -10.0 kcal/mol, ενώ οι υπόλοιπες υψηλόβαθμες ενώσεις εμφάνισαν τιμές γύρω από -8.0 kcal/mol. Παρόμοια ενεργειακά προφίλ έχουν αναφερθεί και στη βιβλιογραφία για αναστολείς που στοχεύουν είτε τον TNF- α είτε το σύμπλοκο TNF- α /TNFR1, γεγονός που υποδεικνύει ότι οι προβλεπόμενες συγγένειες της τάξης των -8.0 έως -11.0 kcal/mol βρίσκονται εντός του αναμενόμενου και φαρμακολογικά ρεαλιστικού εύρους.^[54] Η σύγκλιση αυτή με τις δημοσιευμένες μελέτες ενισχύει την αξιοπιστία των υπολογιστικών αποτελεσμάτων της παρούσας εργασίας. Συνολικά, τα αποτελέσματα της εργασίας ανέδειξαν ότι ο TNFR1 διαθέτει πολλαπλές δομικές κοιλότητες με φαρμακολογικό ενδιαφέρον, πέραν του κλασικού TNF-binding site. Τα κορυφαία μόρια που προέκυψαν από το συνδυασμό AutoDock Vina, AutoDock4.2 και CNN, ιδίως οι ενώσεις DB01199, DB03219 και DB01227, εμφάνισαν σαφή προτίμηση προς εναλλακτικές θέσεις πρόσδεσης, οι οποίες συγκλίνουν χωρικά αλλά δεν επικαλύπτονται με την πρωτογενή θέση δέσμευσης του TNF- α . Η παρατήρηση αυτή υποστηρίζει την ύπαρξη πιθανών αλλοστερικών σημείων, ικανών να επηρεάσουν τη δομή ή τη λειτουργική κατάσταση του υποδοχέα.

Η σύγκλιση των αποτελεσμάτων μεταξύ των τριών μεθοδολογιών, αλλά και η συμφωνία τους με βιβλιογραφικά δεδομένα από μελέτες στο TNF- α , στον TNFR1 και στο σύμπλοκο TNF- α /TNFR1, υπογραμμίζουν τη συνέπεια και αξιοπιστία των υπολογιστικών προβλέψεων. Επιπλέον, η ταύτιση των ενεργειακών εύρων πρόσδεσης με αυτά της διεθνούς βιβλιογραφίας ενισχύει την εγκυρότητα των ευρημάτων και θέτει μια γερά θεμελιωμένη βάση για τη μετάβαση στη φάση πειραματικής επαλήθευσης. Τα δεδομένα αυτά είναι ιδιαίτερος σημαντικός στο πλαίσιο των TRAPS, δεδομένου ότι οι μεταλλάξεις στο TNFRSF1A, το γονίδιο που κωδικοποιεί τον TNFR1, οδηγούν σε παθολογική ενεργοποίηση της σηματοδοτικής οδού του TNF α . Οι μεταλλάξεις αυτές επηρεάζουν κρίσιμες δομικές περιοχές του TNFR1. Η υπερενεργοποίηση αυτή οφείλεται σε δυσλειτουργία του υποδοχέα, με αποτέλεσμα την υπερβολική παραγωγή φλεγμονωδών σημάτων. Συνεπώς, η στοχευμένη αναστολή του TNF α -TNFR1 άξονα μέσω μικρών μορίων μπορεί να αποτελέσει μια εναλλακτική στρατηγική παρέμβασης, ικανή να αντισταθμίσει τις δυσλειτουργίες που προκαλούν οι μεταλλάξεις.

4. Συζήτηση και Συμπεράσματα

Οι SAIDs αποτελούν μια ετερογενή ομάδα νοσημάτων που χαρακτηρίζονται από επαναλαμβανόμενα επεισόδια συστηματικής φλεγμονής και δυσλειτουργία της έμφυτης ανοσίας. Από την αναγνώριση του πρώτου σχετιζόμενου γονιδίου το 1997, η κατανόηση των SAIDs έχει επεκταθεί σημαντικά, αποκαλύπτοντας πλήθος μονογονιδιακών και πολυγονιδιακών συνδρόμων με κοινό παθοφυσιολογικό υπόβαθρο. Παρά τη διαφοροποίηση της κλινικής εικόνας μεταξύ των διαφόρων μορφών, οι περισσότερες αυτοφλεγμονώδεις διαταραχές συνδέονται με ανώμαλη ενεργοποίηση ή απορρύθμιση βασικών φλεγμονωδών μονοπατιών, όπως αυτών που περιλαμβάνουν τις ιντερλευκίνες IL-1 και IL-6, καθώς και τον παράγοντα νέκρωσης όγκων TNF-α. Στο πλαίσιο αυτό, η διερεύνηση των μοριακών μηχανισμών που διέπουν την αλληλεπίδραση TNF-α/TNFR1 αποκτά ιδιαίτερη σημασία, καθώς η απορρύθμισή της συνδέεται άμεσα με παθήσεις όπως το σύνδρομο TRAPS. Η παρούσα μελέτη εντάσσεται σε αυτό το επιστημονικό πεδίο, εξετάζοντας υπολογιστικά τη δυνατότητα μικρών μορίων να τροποποιήσουν ή να αναστείλουν τη λειτουργία του TNFR1.

Στο πλαίσιο των αυτοφλεγμονωδών νοσημάτων, το περιοδικό σύνδρομο που σχετίζεται με τον υποδοχέα του παράγοντα νέκρωσης όγκων, TRAPS, αποτελεί μία από τις πλέον χαρακτηριστικές μονογονιδιακές διαταραχές, όπου η δυσλειτουργία του άξονα TNF-TNFR1 διαδραματίζει κεντρικό ρόλο. Το TRAPS προκαλείται από μεταλλάξεις στο γονίδιο TNFRSF1A, το οποίο κωδικοποιεί τον υποδοχέα TNFR1, οδηγώντας σε λανθασμένη αναδίπλωση και παγίδευσή του στο ενδοπλασματικό δίκτυο, με ταυτόχρονη ενεργοποίηση παθολογικών φλεγμονωδών μονοπατιών. Οι μοριακές αυτές διαταραχές μεταφράζονται κλινικά σε επαναλαμβανόμενα επεισόδια πυρετού, άλγους και συστηματικής φλεγμονής, καθιστώντας το TRAPS ένα υπόδειγμα νόσου όπου η απορρύθμιση της έμφυτης ανοσίας συνδέεται άμεσα με τη δομική και λειτουργική αλλοίωση ενός υποδοχέα-κλειδιού. Η μελέτη του TNFR1 είναι, επομένως, ιδιαίτερα κρίσιμη για την κατανόηση των παθοφυσιολογικών μηχανισμών που διέπουν το σύνδρομο, αλλά και για τον εντοπισμό νέων θεραπευτικών στόχων. Ως βασικός μεσολαβητής της σηματοδότησης του TNF-α, ο TNFR1 ρυθμίζει ένα ευρύ φάσμα φλεγμονωδών και κυτταρικών διεργασιών· συνεπώς, η δομική του ανάλυση και η διερεύνηση μικρών μορίων που μπορούν να επηρεάσουν τη δραστηριότητά του αποτελούν σημαντικό βήμα προς την ανάπτυξη στοχευμένων φαρμακολογικών παρεμβάσεων για το TRAPS και συγγενείς αυτοφλεγμονώδεις παθήσεις.

Τα υπολογιστικά και χημειοπληροφορικά εργαλεία αποτελούν αναπόσπαστο τμήμα της σύγχρονης φαρμακευτικής έρευνας, καθώς προσφέρουν τη δυνατότητα συστηματικής διερεύνησης της μοριακής αλληλεπίδρασης μεταξύ μικρών μορίων και πρωτεϊνικών στόχων, όπως ο TNFR1. Μέσα από διαδικτυακά διαθέσιμες πλατφόρμες και εξειδικευμένα λογισμικά, είναι δυνατή η ανάλυση εκτεταμένων βιβλιοθηκών χημικών ενώσεων και η ανάδειξη «προνομιούχων» μοριακών δομών με ιδιαίτερα ευνοϊκά χαρακτηριστικά πρόσδεσης και πιθανή βιοδραστηριότητα. Η εφαρμογή τεχνικών όπως η εικονική διαλογή (virtual screening), η μοριακή πρόσδεση, η

μοντελοποίηση τρισδιάστατης δομής και οι προσομοιώσεις μοριακής δυναμικής επιτρέπουν τον *in silico* εντοπισμό και την αξιολόγηση υποψηφίων φαρμακολογικών παραγόντων. Με τον τρόπο αυτό, καθίσταται εφικτή η στοχευμένη επιλογή ενώσεων με υψηλή συγγένεια προς συγκεκριμένους βιολογικούς στόχους, μειώνοντας σημαντικά τον χρόνο και το κόστος των πρώιμων σταδίων της φαρμακευτικής ανάπτυξης και προσφέροντας πολύτιμα εργαλεία για την επιτάχυνση της ανακάλυψης νέων θεραπευτικών προσεγγίσεων.

4.1 Αξιολόγηση της Υπολογιστικής Προσέγγισης

Η παρούσα εργασία αξιοποίησε υπολογιστικά και χημειοπληροφορικά εργαλεία με στόχο την αναζήτηση μικρών μορίων ικανών να τροποποιήσουν τη λειτουργία του υποδοχέα TNFR1, ενός καίριου ρυθμιστή της φλεγμονώδους σηματοδότησης του TNF- α και βασικού παράγοντα στην παθογένεση του συνδρόμου TRAPS. Η μεθοδολογική προσέγγιση στηρίχθηκε στις αρχές του δομικά καθοδηγούμενου σχεδιασμού φαρμάκων (structure-based drug design), ενσωματώνοντας τεχνικές μοριακής πρόσδεσης για την πρόβλεψη και αξιολόγηση πιθανών σημείων αλληλεπίδρασης μεταξύ μικρών μορίων και του TNFR1. Μέσω της εφαρμογής τυφλής μοριακής πρόσδεσης (blind docking) εξετάστηκε ολόκληρη η επιφάνεια του υποδοχέα, επιτρέποντας τον εντοπισμό τόσο της κύριας περιοχής πρόσδεσης όσο και πιθανών δευτερευόντων ή αλλοστερικών θυλάκων, οι οποίοι δύνανται να αποτελέσουν νέους φαρμακολογικούς στόχους.

Ειδικότερα, επιλέχθηκε για τη μελέτη, η κρυσταλλική δομή του TNFR1 (PDB ID: 1EXT), με στόχο τη διερεύνηση πιθανών περιοχών πρόσδεσης μικρών μορίων και την ανάλυση των αμινοξικών κατάλοιπων που συμμετέχουν στη σταθεροποίηση των συμπλόκων. Για την προσομοίωση των αλληλεπιδράσεων, πραγματοποιήθηκαν blind docking υπολογισμοί με τη χρήση των λογισμικών AutoDock4.2 και AutoDock Vina, προκειμένου να επιτευχθεί πιο ολοκληρωμένη προσέγγιση και σύγκριση των αποτελεσμάτων. Από τα δεδομένα docking εξήχθησαν οι ενέργειες πρόσδεσης και οι γεωμετρίες των συμπλόκων, ενώ επιλέχθηκαν τα πιο σταθερά σύμπλοκα για περαιτέρω ανάλυση. Παράλληλα, κατά την ερμηνεία των αποτελεσμάτων, λήφθηκαν υπόψη δημοσιευμένα δεδομένα από τη διεθνή βιβλιογραφία, τα οποία αφορούσαν αντίστοιχες *in silico* μελέτες στο σύμπλοκο TNF- α /TNFR1. Μελέτες όπως αυτές των Chan et al., Saddala et al., προσέφεραν σημαντικά συγκριτικά στοιχεία σχετικά με τις ενέργειες πρόσδεσης, τα κρίσιμα αμινοξικά κατάλοιπα και τα δομικά χαρακτηριστικά των συμπλόκων, διευκολύνοντας τη συσχέτιση των αποτελεσμάτων της παρούσας εργασίας με ήδη τεκμηριωμένα δεδομένα.

Για την υλοποίηση αυτής της διαδικασίας εξετάστηκαν οι δομές εγκεκριμένων κλινικών φαρμάκων, φαρμάκων που έχουν αποσυρθεί, επαναχρησιμοποιηθεί και υποψηφίων προς δοκιμή χημικών ενώσεων, με επίκεντρο τον υποδοχέα TNFR1 (PDB ID: 1EXT). Η ανάλυση βασίστηκε σε πάνω από 7,000 χημικές δομές που αντλήθηκαν από εξειδικευμένες βάσεις δεδομένων όπως η DrugBank, Malacards, OpenTargets επιτρέποντας μια ευρεία και συστηματική διερεύνηση της χημικής ποικιλότητας. Ιδιαίτερης σημασίας αποτελεί το γεγονός ότι η χρήση της DrugBank περιλάμβανε εγκεκριμένα φάρμακα για άλλες παθήσεις, γεγονός που καθιστά την παρούσα μελέτη

άμεσα συναφή με στρατηγικές επανατοποθέτησης φαρμάκων (drug repurposing). Η προσέγγιση αυτή ενισχύει τη μεταφραστική αξία των αποτελεσμάτων, καθώς τα φάρμακα αυτά διαθέτουν ήδη γνωστό προφίλ ασφάλειας και φαρμακοκινητικών ιδιοτήτων. Μέσω υπολογιστικής μοριακής πρόσδεσης και εφαρμογής ειδικά σχεδιασμένων αλγορίθμων βαθμολόγησης, πραγματοποιήθηκε συγκριτική αξιολόγηση των ενώσεων ως προς τη συγγένεια πρόσδεσης, τις δομικές τους ιδιότητες και την επικάλυψη με γνωστές ή βιβλιογραφικά επικυρωμένες δραστικές χημικές ομάδες. Η προσέγγιση αυτή επέτρεψε την αναγνώριση υποσχόμενων μορίων, τα οποία εμφανίζουν ευνοϊκά ενεργειακά και δομικά χαρακτηριστικά και θα μπορούσαν να αποτελέσουν τη βάση για μελλοντική φαρμακοχημική βελτιστοποίηση.

Η επεξεργασία των δεδομένων μοριακής πρόσδεσης που προέκυψαν από τα δύο λογισμικά (AutoDock Vina και AutoDock 4.2), καθώς και από το εκπαιδευμένο μοντέλο CNN, επέτρεψε τη συστηματική αξιολόγηση των πιθανών αλληλεπιδράσεων των υποψηφίων ενώσεων με τον υποδοχέα TNFR1. Η συγκριτική ανάλυση των τριών μεθοδολογικών προσεγγίσεων ανέδειξε διακριτές κατηγορίες μορίων, οι οποίες χαρακτηρίζονται από υψηλή, ενδιάμεση και χαμηλή συγγένεια πρόσδεσης, αποτυπώνοντας έτσι το πλήρες φάσμα της ενεργειακής σταθερότητας των σχηματιζόμενων συμπλόκων. Τα δεδομένα αυτά συνέβαλαν στην αναγνώριση ιδιαίτερα υποσχόμενων υποψηφίων με ευνοϊκό ενεργειακό προφίλ και σταθερή συμπεριφορά μεταξύ των διαφορετικών εργαλείων, καθώς και ενώσεων με οριακή ή μη ευνοϊκή πρόσδεση, οι οποίες λειτουργούν ως χρήσιμα συγκριτικά σημεία για την κατανόηση των δομικών παραγόντων που καθορίζουν την αποτελεσματική δέσμευση στον TNFR1. Η σύγκλιση των προβλέψεων μεταξύ AutoDock Vina και CNN ενισχύει την αξιοπιστία της συνολικής προσέγγισης, ενώ η συνδυαστική αξιοποίηση συμβατικών και υπολογιστικών τεχνικών μηχανικής μάθησης προσφέρει ένα ισχυρό εργαλείο για την ανίχνευση χημικών μορίων με πιθανό φαρμακολογικό ενδιαφέρον έναντι του άξονα TNF- α /TNFR1.

Η ανάπτυξη και εφαρμογή του κώδικα CNN πρόσφερε ένα επιπλέον επίπεδο αξιολόγησης των αποτελεσμάτων μοριακής πρόσδεσης καθώς οι προβλέψεις του ήταν συνεπείς με τα αποτελέσματα του AutoDock Vina, υποδεικνύοντας ότι η εκπαίδευση του μοντέλου ήταν επιτυχής και μπορεί να αξιοποιηθεί ως συμπληρωματικό εργαλείο για την ταχεία εκτίμηση συγγενειών πρόσδεσης. Το CNN απέδειξε ότι μπορεί να λειτουργήσει ως συμπληρωματικό εργαλείο για την ταχεία εκτίμηση των συγγενειών πρόσδεσης, μειώνοντας σημαντικά το χρόνο και την υπολογιστική επιβάρυνση σε σχέση με τις παραδοσιακές μεθόδους docking. Ωστόσο, οι προβλέψεις του CNN παραμένουν εξαρτώμενες άμεσα από την ποιότητα και ποικιλία των δεδομένων εκπαίδευσης και συνεπώς δεν μπορεί να υποκαταστήσει πλήρως τα κλασικά υπολογιστικά εργαλεία. Για το λόγο αυτό, η συνδυαστική αξιοποίηση διαφορετικών μεθόδων, όπως το AutoDock4.2, το AutoDock Vina και μοντέλα μηχανικής μάθησης, κρίνεται απαραίτητη προκειμένου να εξασφαλιστεί μεγαλύτερη αξιοπιστία. Η συνέπεια των αποτελεσμάτων και η συγκλίνουσα κατάταξη των ενώσεων ανάμεσα στα διαφορετικά προγράμματα τεκμηριώνουν την εγκυρότητα της υπολογιστικής προσέγγισης.

Η αξιοποίηση των αποτελεσμάτων της μοριακής πρόσδεσης δεν περιορίστηκε μόνο στη σύγκριση των ενεργειών δέσμευσης, αλλά αποτέλεσε και τη βάση για την ανάπτυξη ενός μοντέλου CNN, το οποίο σχεδιάστηκε με στόχο την πρόβλεψη και ταξινόμηση των ενεργειών πρόσδεσης νέων υποψηφίων μορίων. Η ενσωμάτωση του CNN στην ερευνητική διαδικασία επέτρεψε την αυτοματοποιημένη αναγνώριση μορίων με υψηλή συγγένεια πρόσδεσης, βελτιστοποιώντας σημαντικά τη διαδικασία *in silico* ανίχνευσης πιθανών αναστολέων του TNFR1. Η συνδυαστική χρήση docking και μηχανικής μάθησης ενίσχυσε περαιτέρω την αξιοπιστία της αξιολόγησης και συνέβαλε στον εντοπισμό ενώσεων που θα μπορούσαν να επηρεάσουν τη διαμόρφωση ή τη λειτουργικότητα του υποδοχέα, δημιουργώντας νέες προοπτικές για την ανάπτυξη καινοτόμων θεραπευτικών παρεμβάσεων σε αυτοφλεγμονώδεις παθήσεις, και ιδιαίτερα στο σύνδρομο TRAPS.

Η ανάπτυξη και εφαρμογή του μοντέλου CNN αποτέλεσε έναν από τους πιο καινοτόμους άξονες της παρούσας μελέτης, ενισχύοντας ουσιαστικά τη διαδικασία πρόβλεψης των ενεργειών πρόσδεσης και προσφέροντας μία σύγχρονη κατεύθυνση στην *in silico* ανακάλυψη υποψήφιων φαρμακολογικών μορίων. Το μοντέλο εκπαιδεύτηκε αποκλειστικά με δεδομένα που παράχθηκαν από τις προσομοιώσεις μοριακής πρόσδεσης AutoDock Vina, επιτρέποντας στο CNN να μάθει τις σχέσεις μεταξύ μοριακών χαρακτηριστικών, όπως αυτά κωδικοποιήθηκαν μέσω RDKit, και των αντίστοιχων ενεργειών δέσμευσης (ΔG). Κατά την εκπαίδευση, παρατηρήθηκε σταθερή και προοδευτική μείωση του σφάλματος εκπαίδευσης από την αρχική τιμή των 3.65 έως τα 0.59 στα τελευταία epochs, γεγονός που υποδηλώνει αποτελεσματική εκμάθηση των υποκείμενων μοτίβων που καθορίζουν την αλληλεπίδραση ενός μικρού μορίου με την πρωτεΐνη TNFR1. Παράλληλα, το σφάλμα επαλήθευσης εμφάνισε σημαντική πτώση στα πρώιμα, μεσαία στάδια εκπαίδευσης, φθάνοντας στο ελάχιστο (περίπου 0.84), προτού αρχίσει να αυξάνεται ελαφρά, σηματοδοτώντας την έναρξη υπερπροσαρμογής (overfitting). Η συμπεριφορά αυτή, η οποία απεικονίστηκε καθαρά στις καμπύλες εκπαίδευσης-επαλήθευσης, επιβεβαιώνει ότι το μοντέλο πέτυχε ένα σημείο βέλτιστης γενίκευσης, όπου η προβλεπτική του ικανότητα μεγιστοποιείται χωρίς να υπερπροσαρμόζεται στα δεδομένα. Η αξιολόγηση του CNN σε ανεξάρτητο σύνολο δοκιμής (test set) επιβεβαίωσε την ισχυρή προβλεπτική του απόδοση, με την MSE να είναι στα 0.9555 και αντίστοιχο RMSE στα 0.98 kcal/mol, τιμές που θεωρούνται ιδιαίτερες ικανοποιητικές για μοντέλα που στοχεύουν στην αναπαραγωγή των αποτελεσμάτων μοριακής πρόσδεσης, όπου ακόμη και αποκλίσεις 1 kcal/mol είναι αποδεκτές. Επιπλέον, ο συντελεστής συσχέτισης R^2 στα 0.89 κατέδειξε ότι το CNN μπορεί να προβλέπει με συνέπεια τις τάσεις πρόσδεσης και να αναγνωρίζει τις ενώσεις με ισχυρή, μέτρια ή ασθενή συγγένεια προς την πρωτεΐνη. Το γεγονός ότι οι προβλέψεις του μοντέλου συμφωνούσαν σε μεγάλο βαθμό με τα αποτελέσματα των AutoDock Vina, υπογραμμίζει την αξιοπιστία του CNN ως συμπληρωματικού εργαλείου.

Ένα ακόμη κρίσιμο πλεονέκτημα του μοντέλου είναι η ταχύτητα και ο βαθμός αυτοματοποίησης. Σε αντίθεση με τις κλασικές μεθόδους docking που απαιτούν σημαντικό υπολογιστικό χρόνο, το CNN επιτρέπει την άμεση πρόβλεψη ενεργειών δέσμευσης για νέες ενώσεις, κάνοντας εφικτή την αξιολόγηση χιλιάδων μορίων σε ελάχιστο χρόνο. Έτσι, η προσέγγιση αυτή μπορεί να λειτουργήσει

ως πολύτιμο εργαλείο για την διαλογή φαρμακολογικών βιβλιοθηκών, επιταχύνοντας τον εντοπισμό ενώσεων που αξίζει να εξεταστούν περαιτέρω με πιο απαιτητικές υπολογιστικές ή πειραματικές μεθόδους. Συνολικά, η ενσωμάτωση του CNN στην παρούσα εργασία αποδεικνύει τη δυνατότητα των μεθόδων τεχνητής νοημοσύνης να ενισχύσουν και σε πολλές περιπτώσεις να επιταχύνουν τις διαδικασίες ανακάλυψης φαρμάκων. Το μοντέλο που αναπτύχθηκε μπορεί να αποτελέσει τη βάση για μελλοντικές επεκτάσεις, όπως η ενσωμάτωση τρισδιάστατων περιγραφών, η χρήση γραφημάτων μορίων (graph neural networks) ή η εκπαίδευση σε μεγαλύτερα σύνολα δεδομένων docking. Επιπλέον, το CNN δείχνει να έχει τη δυνατότητα να εφαρμοστεί και σε άλλους φλεγμονώδεις στόχους πέραν του TNFR1, όπως το NLRP3 ή το IL-1R, διευρύνοντας τον ρόλο του στην έρευνα αυτοφλεγμονωδών νοσημάτων.

Με βάση την αξιολόγηση του CNN και τη συμφωνία του με τα αποτελέσματα των AutoDock Vina και AutoDock4.2, κρίθηκε αναγκαίο να εξεταστούν συγκριτικά οι τρεις μέθοδοι, προκειμένου να εντοπιστούν τα μόρια με τη μεγαλύτερη και τη μικρότερη συγγένεια πρόσδεσης στον TNFR1. Από τη συνολική αξιολόγηση των 100 ενώσεων προέκυψαν τρεις χαρακτηριστικοί εκπρόσωποι διαφορετικών επιπέδων πρόσδεσης. Η ένωση DB01199 κατέγραψε τις πιο αρνητικές τιμές συγγένειας και στις τρεις μεθοδολογικές προσεγγίσεις (-10.0 kcal/mol στο Vina, -9.76 kcal/mol μέσω CNN και -11.01 kcal/mol στο AutoDock4.2), επιβεβαιώνοντας τη σταθερή και ισχυρή αλληλεπίδρασή της με την πρωτεΐνη. Αντίθετα, η ένωση DB04678 εμφάνισε τιμές που διαφοροποιούνταν σημαντικά μεταξύ των μεθόδων και συνολικό μέσο όρο -5.36 kcal/mol, γεγονός που υποδεικνύει ασταθή ή ασθενή πρόσδεση. Ως ένωση μέσης δραστηριότητας αναγνωρίστηκε η DB03104, παρουσιάζοντας ενδιάμεσες τιμές συγγένειας στις τρεις μεθόδους, με μέσο όρο -7.80 kcal/mol, αντιπροσωπεύοντας χαρακτηριστικά το μέσο επίπεδο δέσμευσης.

4.2 Ολοκληρωμένη Ανάλυση Υπολογιστικών Αποτελεσμάτων Μοριακής Πρόσδεσης

Οι ενώσεις επέτρεψαν την περαιτέρω εμβάθυνση στη συγκριτική μελέτη των τριών υπολογιστικών προσεγγίσεων και την εξαγωγή χρήσιμων συμπερασμάτων ως προς τα χαρακτηριστικά που καθορίζουν την πρόσδεση στον TNFR1. Η ανάλυση των δεδομένων πρόσδεσης που προέκυψαν από τις τρεις υπολογιστικές μεθόδους (AutoDock Vina, AutoDock 4.2 και CNN) αποκάλυψε σαφή πρότυπα ως προς τη συμπεριφορά των εξεταζόμενων ενώσεων, παρέχοντας μια ολοκληρωμένη εικόνα των πιθανών αλληλεπιδράσεων με τον TNFR1. Η εφαρμογή blind docking αποδείχθηκε καθοριστική, καθώς επέτρεψε την εξερεύνηση ολόκληρης της επιφάνειας του υποδοχέα, και ανέδειξε τόσο το κλασικό TNF-binding site (CRD2–CRD3) όσο και δευτερεύουσες, δυνητικά αλλοστερικές θέσεις πρόσδεσης. Παράλληλα, η σύγκλιση μεταξύ AutoDock Vina και του μοντέλου CNN ήταν αξιοσημείωτη, καθώς οι δύο μέθοδοι εντόπισαν την ένωση DB03219 ως από τους ισχυρότερους υποψήφιους αναστολείς, υποδηλώνοντας ότι κοινά δομικά χαρακτηριστικά της ενώσεως αναγνωρίζονται από αλγορίθμους διαφορετικής φιλοσοφίας. Το AutoDock 4.2, από την άλλη, ανέδειξε ως κορυφαία την DB01199, η οποία, αξιοσημείωτα, διατήρησε ισχυρό ενεργειακό προφίλ και στις τρεις μεθοδολογίες, καθιστώντας την τον πιο συνεπή και σταθερό υποψήφιο. Η παρουσία ενώσεων με χαμηλή ή ασυνεπή συγγένεια πρόσδεσης,

όπως η DB04678, ανέδειξε τις μεθοδολογικές διαφορές των αλγορίθμων και την ανάγκη χρήσης συμπληρωματικών εργαλείων για την αξιόπιστη επιλογή υποψηφίων.

Η σύγκριση αυτή δεν έδωσε μόνο κατάταξη μορίων, αλλά προσέφερε και δομικές ενδείξεις για τους παράγοντες που καθορίζουν την αποτελεσματική πρόσδεση. Η σταθερά υψηλή συγγένεια πρόσδεσης που παρουσίασαν ορισμένες ενώσεις και στις τρεις μεθόδους αποτελεί ένδειξη ότι διαθέτουν ευνοϊκά δομικά χαρακτηριστικά για την αναγνώριση του TNFR1, γεγονός που δικαιολογεί τη στοχοποίηση τους σε επόμενα στάδια βιολογικής και φαρμακολογικής αξιολόγησης. Υποδηλώνει την ύπαρξη μοριακών χαρακτηριστικών (όπως κατάλληλη χωρική γεωμετρία, επιφάνεια αλληλεπίδρασης και ηλεκτροστατικές ιδιότητες) που ευνοούν τη δέσμευση στον TNFR1. Επιπλέον, η ανίχνευση ενώσεων με ενδιάμεσο ενεργειακό προφίλ, όπως η DB03104, δείχνει ότι υπάρχουν μόρια με σημαντικό περιθώριο βελτίωσης μέσω φαρμακοχημικών τροποποιήσεων. Ταυτόχρονα, η διαφοροποίηση των αποτελεσμάτων για ορισμένες ενώσεις ανάμεσα στις τρεις μεθόδους παρέχει πολύτιμη πληροφορία για την ακρίβεια και τις αδυναμίες κάθε τεχνικής και δημιουργεί ένα πλαίσιο για την ορθολογική επιλογή των πιο υποσχόμενων υποψηφίων. Η συγκριτική ανάλυση μεταξύ των εργαλείων AutoDock Vina, AutoDock4.2 και του κώδικα CNN, ανέδειξε τη συνέπεια και την αξιοπιστία των τριών μεθόδων στην εκτίμηση των ενεργειών δέσμευσης (ΔG_{bind}), καθώς και την ικανότητά τους να διακρίνουν μοριακές ενώσεις με διαφορετική συγγένεια προς τον TNFR1. Παράλληλα, η σύγκριση μεταξύ AutoDock Vina και AutoDock4.2 ανέδειξε μικρές αποκλίσεις στις προβλεπόμενες ενέργειες πρόσδεσης, γεγονός που αποδίδεται σε διαφοροποιήσεις στους αλγορίθμους υπολογισμού και στις παραμέτρους αναζήτησης των εργαλείων, χωρίς ωστόσο να επηρεάζεται η συνολική ταξινόμηση των ενώσεων.

Η ανάλυση της πρωτεΐνης TNFR1 (PDB ID: 1EXT) με τα εργαλεία CASTp και DoGSiteScorer επέτρεψε την αναγνώριση των κατάλληλων θυλάκων για την πρόσδεση μικρών μορίων, συμβάλλοντας ουσιαστικά στην κατανόηση των δομικών ιδιοτήτων του υποδοχέα και στη διερεύνηση της δυνατότητας φαρμακολογικής του τροποποίησης. Δεδομένου ότι η κρυσταλλική δομή 1EXT περιλαμβάνει αποκλειστικά τον εξωκυττάριο τομέα του TNFR1, χωρίς τον φυσικό του προσδέτη TNF- α , η ακριβής θέση πρόσδεσης του TNF α δεν αποτυπώνεται πλήρως στην παρούσα δομή. Ωστόσο, η ανάλυση των θυλάκων ανέδειξε σαφώς ορισμένες περιοχές με υψηλή φαρμακολογικά αξιοποίηση, ιδιαίτερα στην ευρύτερη περιοχή των δομικών τομέων CRD2–CRD3, που εμπλέκονται στη φυσιολογική λειτουργία του υποδοχέα.

Η ανάλυση της κρυσταλλικής δομής του TNFR1 (PDB ID: 1EXT) με τα δύο εργαλεία (CASTp, DoGSiteScorer) ανέδειξε έναν καλά διαμορφωμένο και φαρμακολογικά αξιοποιήσιμο θύλακα στην εξωκυττάρια περιοχή του υποδοχέα. Ο θύλακας αυτός εντοπίζεται στην περιοχή που εκτείνεται ανάμεσα στους δομικούς τομείς CRD2–CRD3 και σχηματίζεται από κρίσιμα αμινοξικά κατάλοιπα της αλυσίδας A, όπως τα Lys75, Ser74, Arg77, Glu109, Asn110 και Leu111. Τα συγκεκριμένα κατάλοιπα έχουν αναφερθεί και σε προηγούμενες υπολογιστικές μελέτες ως σημαντικά για τη λειτουργική διαμόρφωση του TNFR1, επιβεβαιώνοντας την αξιοπιστία των υπολογιστικών προβλέψεων. Επιπλέον, στην ίδια περιοχή εντοπίστηκαν τα κατάλοιπα Val91,

Leu93, Ser95 και Phe124, τα οποία συμβάλλουν στη διαμόρφωση της κοιλότητας και ενισχύουν τη σταθερή χωροταξική της οργάνωση. Η επιτυχής χαρτογράφηση του συγκεκριμένου θύλακα επέτρεψε τη συστηματική σύγκριση ανάμεσα στη φυσιολογική θέση πρόσδεσης του TNF-α και στις θέσεις που εντοπίστηκαν κατά τη διαδικασία της τυφλής μοριακής πρόσδεσης (blind docking) των υποψήφιων ενώσεων της παρούσας μελέτης. Η σύγκριση αυτή είχε ιδιαίτερη σημασία, καθώς κατέδειξε ότι αρκετά από τα υποψήφια μόρια υψηλής συγγένειας δεν προσδένονται στον κλασικό ενεργό θύλακα, αλλά σε διαφορετικές, μέχρι σήμερα ανεξερεύνητες περιοχές του TNFR1.

Σε αντίθεση με τη διεπαφή TNF-α/TNFR1, η οποία απαιτεί μοντελοποιημένες δομές του συμπλόκου για να αποτυπωθεί πλήρως, οι φαρμακολογικά αξιοποιήσιμες κοιλότητες που αποκαλύφθηκαν στη δομή 1EXT αποτελούν εγγενείς, σαφώς οριοθετημένες δομικές κοιλότητες του υποδοχέα, κατάλληλες για πρόσδεση μικρών οργανικών μορίων. Η γεωμετρία τους, με ικανοποιητικό βάθος, όγκο και ποικιλία χημικών χαρακτηριστικών, υποδεικνύει ότι μπορούν να φιλοξενήσουν ενώσεις με διαφορετικά προφίλ υδροφοβικότητας ή πολικότητας, καθιστώντας τους ελκυστικούς στόχους για σχεδιασμό αναστολέων. Η ύπαρξη τέτοιων θύλακων είναι ιδιαίτερα σημαντική για νοσήματα όπως το TRAPS, όπου η δυσλειτουργία του TNFR1 αποτελεί κεντρικό παθοφυσιολογικό μηχανισμό. Η δυνατότητα εντοπισμού μικρών μορίων που μπορούν να προσδεθούν σε αυτές τις θέσεις παρέχει νέες προοπτικές για την ανάπτυξη αλλοστερικών ή ανταγωνιστικών αναστολέων του TNFR1, ικανών να τροποποιήσουν επιλεκτικά τη δραστηριότητα του υποδοχέα και να επηρεάσουν τη φλεγμονώδη σηματοδότηση. Συνολικά, ο εντοπισμός του συγκεκριμένου θύλακα από δύο ανεξάρτητα εργαλεία (CASTp και DoGSiteScorer) επιβεβαίωσε τη δομική του αξιοπιστία και θέτει ένα ισχυρό υπόβαθρο για την περαιτέρω αξιολόγηση των ενώσεων που προσδέθηκαν στη συγκεκριμένη περιοχή στα πλαίσια της παρούσας μελέτης.

Η ανάλυση των θέσεων πρόσδεσης των κορυφαίων ενώσεων αποκάλυψε ότι καμία από αυτές δεν προσδέθηκε στη κλασική θέση πρόσδεσης του TNFR1, όπως αυτό προσδιορίστηκε από τα CASTp και DoGSiteScorer, αλλά σε εναλλακτικές κοιλότητες της επιφάνειας του TNFR1. Στην ένωση DB01199, η οποία εμφάνισε τη σταθερότερη συνολικά συγγένεια πρόσδεσης και στις τρεις μεθόδους (AutoDock Vina, AutoDock4.2 και CNN), η θέση πρόσδεσης δεν εντοπίστηκε στο ενεργό σημείο πρόσδεσης του TNF-α στο TNFR1. Τα αποτελέσματα του AutoDock4.2 έδειξαν ισχυρές αλληλεπιδράσεις με τα κατάλοιπα Gln130, Glu131, Arg104, Lys100, Cys98, Asn101 και Gln102, ενώ το AutoDock Vina εντόπισε διαφορετικό, αλλά γειτονικό θύλακα, εμπλέκοντας τα Gln133, Asn134, Tyr103, Phe115, His105, Gln113 και Trp107. Η διαφοροποίηση αυτή των σημείων πρόσδεσης δεν αποτελεί ασυνέπεια, αλλά ισχυρή ένδειξη ότι η DB01199 μπορεί να δρα ως αλλοστερικός τροποποιητής, αναγνωρίζοντας δευτερεύοντες θύλακες που ενδέχεται να επηρεάζουν τη διαμόρφωση ή τη λειτουργική κατάσταση του TNFR1 χωρίς να καταλαμβάνει την πρωτογενή θέση δέσμευσης του TNF-α. Παρόμοιες παρατηρήσεις προέκυψαν και για την ένωση DB03219, η οποία προσδένεται στα κατάλοιπα His69, Tyr38, Arg53, Ser59, Glu56 και Cys70, σε περιοχή διαφορετική από το κλασικό ενεργό κέντρο αλλά χωρικά γειτονική προς τον θύλακα που αναγνωρίζει η DB01199. Η χωρική αυτή εγγύτητα υποδηλώνει την ύπαρξη «γειτονικών» ή επικαλυπτόμενων αλλοστερικών περιοχών οι οποίες μπορεί να ασκούν ρυθμιστική επίδραση στον

υποδοχέα. Η ένωση DB01227, που εντοπίστηκε μέσω ομαδοποίησης ως δομικά συγγενής προς την DB01199, εμφάνισε επίσης πρόσδεση σε εναλλακτικό θύλακα, αλληλεπιδρώντας με τα κατάλοιπα Phe115, Gln113, His66, Leu67, Trp107, Lys32, Glu64 και Gln117. Η σύγκλιση αυτή των αποτελεσμάτων, η αναγνώριση διαφορετικών αλλά γειτονικών σημείων πρόσδεσης από τρεις ανεξάρτητες ενώσεις, ενισχύει σημαντικά την υπόθεση ότι ο TNFR1 διαθέτει πολλαπλές αλλοστερικές θέσεις, οι οποίες ενδέχεται να ρυθμίζουν την ενεργοποίησή του και τη δέσμευση του TNF-α.

Η μετατόπιση της πρόσδεσης εκτός του κύριου θυλάκου μπορεί να αποδοθεί, αφενός, στη χρήση blind docking, διαδικασίας που εξερευνά ολόκληρη την επιφάνεια της πρωτεΐνης χωρίς τον περιορισμό στην κλασική περιοχή και, αφετέρου, στις διαφορές στη βαθμολόγηση μεταξύ των διαθέσιμων θέσεων πρόσδεσης, οι οποίες ενδέχεται να ευνοούν εναλλακτικές κοιλότητες με ευνοϊκότερες εκτιμήσεις ελεύθερης ενέργειας πρόσδεσης. Ωστόσο, όπως αναφέρεται και στη σχετική βιβλιογραφία, η πιο αρνητική ενεργειακή τιμή δεν ταυτίζεται απαραίτητα με την επιθυμητή βιολογική περιοχή πρόσδεσης, γεγονός που καθιστά απαραίτητη τη συμπληρωματική αξιολόγηση μέσω μοριακής δυναμικής και *in vitro* βιοχημικών μελετών. Το γεγονός ότι οι ενώσεις με την ισχυρότερη συγγένεια πρόσδεσης εντοπίστηκαν σε μη συμβατικές θέσεις ενισχύει την υπόθεση ύπαρξης νέων αλλοστερικών κοιλοτήτων στον TNFR1. Αυτές οι κοιλότητες θα μπορούσαν να λειτουργούν ως ρυθμιστικά σημεία, ικανά να τροποποιήσουν τη λειτουργικότητα του υποδοχέα με μη ανταγωνιστικό τρόπο, επηρεάζοντας έμμεσα τη διαμόρφωση ή τη σταθερότητα του συμπλόκου TNF–TNFR1. Η διαφοροποίηση των θέσεων πρόσδεσης μεταξύ των μορίων υποδηλώνει ότι ο υποδοχέας διαθέτει περισσότερους από έναν λειτουργικά αξιοποιήσιμους θύλακες και ότι οι διαφορετικές βαθμολογήσεις των εργαλείων μοριακής πρόσδεσης μπορεί να κατευθύνουν την αναζήτηση προς ενεργειακά σταθερές αλλά μη κλασικές κοιλότητες. Συνολικά, η παρατήρηση ότι τα ισχυρότερα υποψήφια μόρια προσδένονται εκτός της γνωστής λειτουργικής περιοχής υποστηρίζει την ιδέα ότι η αλλοστερική ρύθμιση του TNFR1 αποτελεί έναν πολλά υποσχόμενο και σχετικά ανεξερευνήτο άξονα για τη μελλοντική ανάπτυξη εξειδικευμένων θεραπευτικών παρεμβάσεων στα σύνδρομα TRAPS.

Πέρα από τα αποτελέσματα της παρούσας εργασίας, τα οποία ανέδειξαν τη σημασία των εναλλακτικών θυλάκων στην επιφάνεια του TNFR1, κρίθηκε απαραίτητο να ληφθούν υπόψη και δεδομένα προηγούμενων *in silico* μελετών που έχουν διερευνήσει τον άξονα TNFα–TNFR1. Μεταξύ αυτών, ιδιαίτερα χρήσιμα ήταν τα αποτελέσματα της δημοσιευμένης εργασίας που χρησιμοποιεί τη βάση δεδομένων ZINC για την αναζήτηση μικρομοριακών αναστολέων. Η μελέτη αυτή επιβεβαιώνει ότι η αλληλεπίδραση TNFα–TNFR1 αποτελεί κομβικό σημείο ρύθμισης της φλεγμονώδους απόκρισης, καθώς η πρόσδεση του TNFα στον TNFR1 πυροδοτεί σηματοδοτικά μονοπάτια που οδηγούν σε παραγωγή κυτταροκινών, ενεργοποίηση του NF-κΒ και εγκαθίδρυση συστηματικής φλεγμονής. Στα TRAPS, όπου ο TNFR1 είναι δομικά ή λειτουργικά δυσλειτουργικός, η παθολογική αυτή ενεργοποίηση γίνεται ακόμη εντονότερη, καθιστώντας τη στοχευμένη παρεμπόδιση του συμπλόκου μια πολλά υποσχόμενη θεραπευτική στρατηγική.

Τα αποτελέσματα της έρευνας των Saddala et al ενίσχυσαν τα συμπεράσματα, δείχνοντας ότι μικρά μόρια μπορούν πράγματι να προσδεθούν σε κρίσιμες θέσεις τόσο του TNFα όσο και του TNFR1, με ενέργειες δέσμευσης που κυμαίνονται σε εξαιρετικά χαμηλές τιμές (-10.1 έως -8.1 kcal/mol). Στον TNFα, αναδείχθηκαν κατάλοιπα όπως τα Ile58, Leu120, Gly121 και Tyr151, τα οποία φαίνεται να διαδραματίζουν καθοριστικό ρόλο στη σταθεροποίηση του τριμερούς συμπλόκου της κυτταροκίνης, μια διαδικασία που αποτελεί το πρώτο βήμα για την ενεργοποίηση της φλεγμονώδους οδού. Στον TNFR1, αντίστοιχα, οι ενώσεις με την υψηλότερη συγγένεια σχημάτιζαν πολλαπλούς δεσμούς υδρογόνου με τα Lys75, Ser74, Arg104, Tyr106 και Asn110, καταδεικνύοντας ότι ο συγκεκριμένος υποδοχέας διαθέτει όχι μόνο ένα κύριο σημείο πρόσδεσης, αλλά και περισσότερες λειτουργικά αξιοποιήσιμες κοιλότητες. Αντίστοιχα, η χαρτογράφηση των ενεργών θέσεων του συμπλόκου TNFα–TNFR1 αποκάλυψε ένα εκτεταμένο σύνολο αμινοξικών υπολειμμάτων (His15, Val17, Ala18, Arg32, Phe144, Glu146, Ser147, Gly148, Gln149, Val150, Tyr151), τα οποία συνιστούν τη λειτουργική διεπαφή που επιτρέπει την ενεργοποίηση της σηματοδοτικής οδού. Ορισμένα μόρια της βάσης ZINC εμφάνισαν ιδιαίτερα υψηλή σταθερότητα πρόσδεσης μέσω πολλαπλών αλληλεπιδράσεων με κρίσιμα υπολείμματα όπως Ser74, Thr94, Glu109, Asn110, Arg104 και Ser147. Οι αλληλεπιδράσεις αυτές αποτελούν δομικά σημεία-κλειδιά για τη συγκρότηση και διατήρηση της διεπαφής του συμπλόκου, ενισχύοντας την ιδέα ότι η στοχευμένη αναστολή τους θα μπορούσε να περιορίσει αποτελεσματικά την παθολογική ενεργοποίηση του TNFR1 σε φλεγμονώδεις νόσους όπως οι TRAPS. Η σύγκλιση των αποτελεσμάτων της παρούσας εργασίας με εκείνα της βιβλιογραφίας είναι ιδιαίτερα σημαντική. Παρότι οι υποψήφιες ενώσεις που εντοπίστηκαν στο blind docking προσδέθηκαν κυρίως σε εναλλακτικές, μη κλασικές κοιλότητες, υποδηλώνοντας αλλοστερική ρύθμιση, η βιβλιογραφία δείχνει ότι τόσο ο TNFα όσο και ο TNFR1 διαθέτουν πολλαπλές περιοχές δυνητικής φαρμακευτικής στόχευσης. Αυτό υπογραμμίζει το γεγονός ότι η παρεμβολή στο σύμπλοκο TNFα–TNFR1 μπορεί να επιτευχθεί με ποικίλους μηχανισμούς: είτε μέσω άμεσης αναστολής της κυτταροκίνης, είτε μέσω παρεμπόδισης του υποδοχέα, είτε μέσω αλλοστερικής αποσταθεροποίησης της διεπαφής τους.

Στο ίδιο πλαίσιο, η μελέτη των Chan et al. πρόσφερε πρόσθετα στοιχεία, αποτέλεσε μία από τις πρώτες ολοκληρωμένες εφαρμογές δομικά καθοδηγούμενου σχεδιασμού φαρμάκων στον άξονα TNFα–TNFR1. Η έρευνα αυτή απέδειξε ότι η αναστολή της αλληλεπίδρασης του TNFα με τον TNFR1 μπορεί πράγματι να επιτευχθεί μέσω μικρών μορίων φυσικής προέλευσης, αναδεικνύοντας τη διεπαφή TNFα–TNFR1 ως φαρμακολογικά αξιοποιήσιμο στόχο. Με τον εικονικό έλεγχο άνω των 20.000 φυσικών προϊόντων και τη μοριακή πρόσδεση βασισμένη στη δομή του συμπλόκου TNFα–TNFR1, η μελέτη εντόπισε ενώσεις με ενέργειες δέσμευσης στην περιοχή των -7 έως -10 kcal/mol και σταθερές αλληλεπιδράσεις με κρίσιμα κατάλοιπα του TNFα. Η πειραματική επιβεβαίωση (ELISA) δύο κορυφαίων ενώσεων, της πυραζόλης συνδεδεμένης με κινουκλιδίνη και της ινδολ[2,3-α]κινολιζιδίνης, απέδειξε ότι συγκεκριμένα μικρά μόρια μπορούν να αναστείλουν αποτελεσματικά τη δέσμευση του TNFα στον TNFR1, μειώνοντας τη σταθερότητα του συμπλόκου. Τα δεδομένα αυτά ενισχύουν το συμπέρασμα της παρούσας εργασίας ότι, πέρα από την κλασική ενεργό θέση του TNFR1, υπάρχουν χημικές κατηγορίες

μορίων ικανές να επηρεάσουν λειτουργικά τον άξονα TNFα–TNFR1, είτε μέσω άμεσης παρέμβασης στη διεπαφή, είτε μέσω αλλοστερικής ρύθμισης. Επιπλέον, η μελέτη Chan et al. αναδεικνύει ότι διαφορετικές χημικές οικογένειες μπορούν να λειτουργήσουν ως scaffolds για την ανάπτυξη νέων αναστολέων, γεγονός που συνδέεται άμεσα με την παρατήρηση της παρούσας εργασίας ότι ο TNFR1 διαθέτει πολλαπλές, δομικά διακριτές θέσεις πρόσδεσης, οι οποίες δύνανται να αξιοποιηθούν θεραπευτικά.

Ένα από τα σημαντικότερα συμπεράσματα της παρούσας μελέτης αφορά τη συμπληρωματικότητα και συνέργεια των υπολογιστικών προσεγγίσεων που εφαρμόστηκαν. Η ταυτόχρονη χρήση διαφορετικών μεθόδων μοριακής πρόσδεσης (AutoDock Vina, AutoDock 4.2) σε συνδυασμό με την ανάπτυξη ενός μοντέλου βαθιάς μάθησης (CNN) επέτρεψε την πολυδιάστατη αξιολόγηση των υποψήφιων ενώσεων, αναδεικνύοντας τόσο τις σταθερές τάσεις όσο και τις αποκλίσεις μεταξύ των εργαλείων. Η παρατήρηση ότι ορισμένα μόρια, όπως η DB01199, εμφανίζονται μεταξύ των κορυφαίων ενώσεων σε όλες τις μεθόδους, υποδηλώνει υψηλή πιθανότητα πραγματικής βιολογικής σημασίας και ενισχύει την αξιοπιστία των προβλέψεων. Παράλληλα, η ανίχνευση αλλοστερικών περιοχών πρόσδεσης που δεν έχουν περιγραφεί εκτενώς στη βιβλιογραφία, προτείνει νέα μονοπάτια φαρμακολογικής παρέμβασης, ιδιαίτερα σημαντικά για παθήσεις όπως οι TRAPS όπου οι κλασικοί μηχανισμοί απευθείας ανταγωνισμού της TNFα–TNFR1 αλληλεπίδρασης έχουν περιορισμένη αποτελεσματικότητα. Συνολικά, η μελέτη καταδεικνύει πως η ενσωμάτωση μεθοδολογιών docking με σύγχρονα μοντέλα μηχανικής μάθησης μπορεί να προσφέρει βαθύτερη κατανόηση των δομικών δυναμικών του TNFR1, συμβάλλοντας τόσο στη χαρτογράφηση λειτουργικών θυλάκων όσο και στον εντοπισμό νέων υποσχόμενων μικρομοριακών αναστολέων που ενδέχεται να αξιοποιηθούν σε μελλοντικές θεραπευτικές στρατηγικές.

Συνοψίζοντας, η παρούσα μελέτη ανέδειξε ότι η υπολογιστική διερεύνηση του TNFR1, ενός βασικού ρυθμιστή της φλεγμονώδους σηματοδότησης και κεντρικού παράγοντα παθογένεσης στο σύνδρομο TRAPS, μπορεί να προσφέρει πολύτιμες πληροφορίες για την αναγνώριση νέων υποψήφιων μικρομοριακών αναστολέων. Μέσω μιας ολοκληρωμένης προσέγγισης που συνδύασε blind molecular docking, συγκριτική αξιολόγηση τριών διαφορετικών μεθοδολογιών (AutoDock Vina, AutoDock4.2 και CNN), ανάλυση ομαδοποίησης και επιβεβαίωση των πιθανών θυλάκων πρόσδεσης με CASTp και DoGSiteScorer, κατέστη δυνατή η χαρτογράφηση λειτουργικά αξιοποιήσιμων θέσεων και η ταυτοποίηση μορίων με υψηλή εκτιμώμενη συγγένεια πρόσδεσης στον TNFR1.

Τα αποτελέσματα ανέδειξαν ότι οι περισσότερες ενώσεις με ευνοϊκό ενεργειακό προφίλ δεν προσδένονται στο κλασικό TNF-binding site (CRD2–CRD3), αλλά εντοπίζονται σε δευτερεύουσες ή γειτονικές κοιλότητες της πρωτεΐνης. Οι ισχυρότεροι υποψήφιοι, όπως οι DB01199 και DB03219, εμφάνισαν συνεχή και σταθερή συμπεριφορά πρόσδεσης, αλλά σε εναλλακτικές περιοχές που δεν ταυτίζονται με τον κύριο θύλακα. Το εύρημα αυτό υποστηρίζει την ύπαρξη πολλαπλών αλλοστερικών θυλάκων στον TNFR1 και υποδηλώνει ότι η

φαρμακολογική ρύθμιση του υποδοχέα μπορεί να επιτευχθεί και με μη ανταγωνιστικούς μηχανισμούς, οι οποίοι δεν απαιτούν άμεση αναστολή της φυσικής διεπαφής TNF α –TNFR1. Επιπλέον, η ομαδοποίηση (clustering) ανέδειξε ότι μοριακά παρόμοιες ενώσεις, όπως η DB01199 και η DB01227, τείνουν να αναγνωρίζουν γειτονικές ή εν μέρει επικαλυπτόμενες θέσεις, ενισχύοντας την υπόθεση ότι ο TNFR1 διαθέτει δομικά συγγενείς αλλά λειτουργικά διακριτούς αλλοστερικούς θύλακες.

Η ανάπτυξη του μοντέλου CNN αποτέλεσε σημαντική συνιστώσα της μελέτης, καθώς επέτρεψε την αυτοματοποιημένη πρόβλεψη ενεργειών δέσμευσης με υψηλό βαθμό ακρίβειας (RMSE περίπου 0.98 kcal/mol) και ισχυρή συσχέτιση με τα αποτελέσματα του AutoDock Vina (R^2 περίπου 0.89). Το CNN μοντέλο επέδειξε ικανότητα γενίκευσης, αναγνωρίζοντας με συνέπεια τις ενώσεις υψηλής συγγένειας, και επιβεβαίωσε τα ενεργειακά πρότυπα που προέκυψαν από τα κλασικά docking εργαλεία. Η σύγκλιση μεταξύ των δύο προσεγγίσεων υπογραμμίζει την αυξανόμενη αξία των τεχνητών νευρωνικών δικτύων σε εφαρμογές virtual screening και δείχνει ότι τα μοντέλα μηχανικής μάθησης μπορούν να λειτουργήσουν ως αξιόπιστοι επιταχυντές στις πρώιμες φάσεις ανακάλυψης φαρμάκων.

Η συνολική ερμηνεία των δεδομένων τοποθετεί τα ευρήματα της παρούσας εργασίας σε συμφωνία με τη διαθέσιμη βιβλιογραφία. Η μελέτη των Saddala et al. είχε δείξει ότι τόσο ο TNF α όσο και ο TNFR1 διαθέτουν πολλαπλά σημεία πρόσδεσης, με κρίσιμα κατάλοιπα όπως τα Ser74, Lys75, Thr94, Arg104, Glu109 και Asn110 να διαδραματίζουν καθοριστικό ρόλο στη σταθεροποίηση των συμπλόκων. Παρομοίως, η παρούσα εργασία κατέγραψε αλληλεπιδράσεις σε αυτές και σε γειτονικές θέσεις, ενισχύοντας τα συμπεράσματα σχετικά με την ύπαρξη περισσότερων του ενός φαρμακολογικά αξιοποιήσιμων θυλάκων. Επιπλέον, η μελέτη των Chan et al., η οποία αναζήτησε φυσικά προϊόντα ως αναστολείς της διεπαφής TNF α –TNFR1 μέσω large-scale virtual screening και πειραματικής επιβεβαίωσης, ανέδειξε ότι μικρά μόρια με κατάλληλα δομικά χαρακτηριστικά μπορούν να παρεμποδίσουν αποτελεσματικά τον σχηματισμό του συμπλόκου, οδηγώντας σε μετρήσιμη αναστολή της φλεγμονώδους σηματοδότησης.

Η μελέτη πιθανών αναστολέων τόσο για το TNF α όσο και για τον TNFR1 έχει άμεσες θεραπευτικές προεκτάσεις. Η ταυτοποίηση μικρών μορίων με ισχυρή συγγένεια πρόσδεσης και ευνοϊκή ενεργειακή σταθερότητα δείχνει ότι υπάρχουν εναλλακτικοί στόχοι για φαρμακολογική παρέμβαση πέρα από τους ήδη υπάρχοντες βιολογικούς αναστολείς του TNF α , οι οποίοι παρουσιάζουν περιορισμούς όπως ανεπιθύμητες ενέργειες, υψηλό κόστος ή περιορισμένη αποτελεσματικότητα σε ορισμένους ασθενείς. Οι ενώσεις που αναφέρθηκαν στη βιβλιογραφία θα μπορούσαν, μετά από περαιτέρω *in vitro* και *in vivo* έλεγχο, να αποτελέσουν τη βάση για την ανάπτυξη νέων στοχευμένων θεραπειών που θα αντιμετωπίζουν πιο αποτελεσματικά τη χρόνια φλεγμονή που χαρακτηρίζει τις TRAPS.

Η συνέργεια ανάμεσα στα αποτελέσματα της παρούσας εργασίας και των προαναφερθεισών μελετών καταδεικνύει ότι η στοχευμένη ρύθμιση του TNFR1 μπορεί να επιτευχθεί τόσο μέσω

αναστολής της φυσικής διεπαφής με τον TNFα όσο και μέσω αλλοστερικών μηχανισμών που επηρεάζουν τη διαμόρφωση και σταθερότητα του υποδοχέα. Η ταυτοποίηση εναλλακτικών θυλάκων και η επιβεβαίωση της φαρμακολογικής τους δυννητικότητας ενισχύει την αντίληψη ότι ο TNFR1 αποτελεί έναν πολυδιάστατο στόχο, με εφαρμογές όχι μόνο για τα TRAPS αλλά και για ένα ευρύτερο φάσμα αυτοφλεγμονωδών διαταραχών που χαρακτηρίζονται από δυσλειτουργία της οδού TNFα–TNFR1. Η μελέτη αυτή παρέχει μια ολοκληρωμένη τεκμηρίωση της δυνατότητας χρήσης υπολογιστικών μεθόδων υψηλής απόδοσης και τεχνητής νοημοσύνης στην αναζήτηση νέων θεραπευτικών προσεγγίσεων και θέτει τις βάσεις για περαιτέρω διερεύνηση με προσομοιώσεις μοριακής δυναμικής και πειραματικές επιβεβαιώσεις. Η σύγκλιση των βιβλιογραφικών δεδομένων με τα αποτελέσματα της παρούσας εργασίας υποδηλώνει ότι τόσο η άμεση στόχευση της διεπαφής TNFα–TNFR1 όσο και η αναστολή μέσω αλλοστερικών μηχανισμών αποτελούν βιώσιμες και συμπληρωματικές στρατηγικές για τον σχεδιασμό νέων αναστολέων. Συνολικά, τα ευρήματα καταδεικνύουν ότι ο TNFR1 διαθέτει μεγαλύτερη φαρμακολογική πλαστικότητα απ' ό,τι είχε αρχικά θεωρηθεί, ενώ η ταυτοποίηση πολλαπλών περιοχών υψηλής συγγένειας πρόσδεσης επεκτείνει το φάσμα των δυνητικών θεραπευτικών παρεμβάσεων όχι μόνο για τα TRAPS αλλά και για άλλες αυτοφλεγμονώδεις διαταραχές που σχετίζονται με τη δυσλειτουργία του άξονα TNFα–TNFR1.

4.3 Περιορισμοί της Παρούσας Μελέτης

Παρά τα ενθαρρυντικά αποτελέσματα που προέκυψαν από την παρούσα εργασία, είναι σημαντικό να αναγνωριστούν ορισμένοι περιορισμοί, οι οποίοι σχετίζονται τόσο με τη φύση των υπολογιστικών μεθόδων που εφαρμόστηκαν όσο και με το εύρος των διαθέσιμων δεδομένων. Η αναγνώριση αυτών των περιορισμών είναι απαραίτητη για την ορθή ερμηνεία των αποτελεσμάτων και για τον σχεδιασμό μελλοντικών ερευνητικών προσεγγίσεων.

Ένας βασικός περιορισμός της μελέτης αφορά τη χρήση μεθόδων μοριακής πρόσδεσης, οι οποίες βασίζονται σε στατικές δομές πρωτεϊνών. Στην παρούσα εργασία χρησιμοποιήθηκε η κρυσταλλική δομή του TNFR1 (PDB ID: 1EXT), η οποία παρέχει πολύτιμες πληροφορίες για τη δομή του υποδοχέα, ωστόσο δεν αποτυπώνει πλήρως τη δυναμική φύση της πρωτεΐνης σε φυσιολογικές συνθήκες. Η απουσία πρωτεϊνικής ευελιξίας ενδέχεται να επηρεάζει την ακρίβεια της πρόβλεψης των θέσεων πρόσδεσης και της πραγματικής σταθερότητας των συμπλόκων, ιδιαίτερα στην περίπτωση αλλοστερικών θυλάκων που μπορεί να εμφανίζονται ή να μεταβάλλονται δυναμικά. Επιπλέον, η εφαρμογή τυφλής μοριακής πρόσδεσης, παρότι επιτρέπει τη διερεύνηση ολόκληρης της επιφάνειας του υποδοχέα και την ανακάλυψη εναλλακτικών κοιλοτήτων, εισάγει εγγενή αβεβαιότητα ως προς τη βιολογική συνάφεια των ενεργειακά ευνοϊκότερων διαμορφώσεων. Όπως αναδεικνύεται και στη βιβλιογραφία, η διαμόρφωση με τη χαμηλότερη υπολογιστική ενέργεια δεν αντιστοιχεί πάντοτε στο λειτουργικά ή φαρμακολογικά πιο σημαντικό σημείο πρόσδεσης, γεγονός που απαιτεί προσεκτική ερμηνεία των αποτελεσμάτων.

Ένας ακόμη περιορισμός σχετίζεται με τα συστήματα βαθμολόγησης (scoring functions) των προγραμμάτων AutoDock Vina και AutoDock4.2. Παρότι τα εργαλεία αυτά είναι ευρέως αποδεκτά και χρησιμοποιούνται εκτενώς στη φαρμακολογική έρευνα, οι υπολογιστικές ενέργειες πρόσδεσης αποτελούν προσεγγίσεις και όχι πραγματικές θερμοδυναμικές τιμές. Οι διαφορές που παρατηρήθηκαν στις θέσεις πρόσδεσης και στις τιμές συγγένειας μεταξύ των μεθόδων αντανakλούν τους περιορισμούς των επιμέρους αλγορίθμων και υπογραμμίζουν την ανάγκη συνδυαστικής αξιολόγησης, όπως εφαρμόστηκε στην παρούσα εργασία. Όσον αφορά το μοντέλο CNN, αν και παρουσίασε ικανοποιητική απόδοση και καλή γενίκευση, η αποτελεσματικότητά του εξαρτάται άμεσα από την ποιότητα και το εύρος των δεδομένων εκπαίδευσης. Το μοντέλο εκπαιδεύτηκε σε σύνολο δεδομένων που βασίζεται σε υπολογιστικά παραγόμενες τιμές ενέργειας πρόσδεσης και όχι σε πειραματικά επικυρωμένα δεδομένα, γεγονός που ενδέχεται να περιορίζει την ακρίβεια των προβλέψεων σε πραγματικές βιολογικές συνθήκες.

Τέλος, σημαντικός περιορισμός της παρούσας μελέτης είναι η απουσία πειραματικής επικύρωσης των υπολογιστικών ευρημάτων. Παρότι τα αποτελέσματα παρουσιάζουν ισχυρή συμφωνία με δεδομένα της διεθνούς βιβλιογραφίας, η βιολογική δραστηριότητα των υποψήφιων μορίων δεν επιβεβαιώθηκε με *in vitro* ή *in vivo* πειράματα. Συνεπώς, τα συμπεράσματα της εργασίας θα πρέπει να ερμηνεύονται ως υπολογιστικές ενδείξεις υψηλού ενδιαφέροντος και όχι ως τελικές αποδείξεις φαρμακολογικής αποτελεσματικότητας. Συνολικά, οι παραπάνω περιορισμοί δεν αναιρούν τη σημασία των αποτελεσμάτων της παρούσας διπλωματικής εργασίας, αλλά αναδεικνύουν τα όρια των υπολογιστικών προσεγγίσεων και τονίζουν την ανάγκη συνδυασμού τους με δυναμικές προσομοιώσεις και πειραματικές μελέτες, στο πλαίσιο μιας ολοκληρωμένης στρατηγικής ανακάλυψης φαρμάκων.

4.4 Μελλοντική Εργασία

Τα αποτελέσματα της παρούσας μελέτης υποδεικνύουν ότι ο άξονας TNF- α /TNFR1 αποτελεί έναν ιδιαίτερα υποσχόμενο στόχο για φαρμακολογική παρέμβαση μέσω μικρών μορίων, αναδεικνύοντας σημαντικές κατευθύνσεις για μελλοντική έρευνα. Η αναγνώριση εναλλακτικών κοιλοτήτων πρόσδεσης στην επιφάνεια του TNFR1, πέραν του κλασικού TNF-binding site (CRD2–CRD3), υποδηλώνει ότι ο υποδοχέας διαθέτει μεγαλύτερη δομική και λειτουργική πολυπλοκότητα από ό,τι θεωρούνταν μέχρι σήμερα. Το γεγονός ότι οι ισχυρότερες ενώσεις της παρούσας εργασίας προσδέθηκαν σε μη συμβατικές θέσεις ενισχύει την υπόθεση ύπαρξης αλλοστερικών κέντρων ικανών να μεταβάλλουν τη ρυθμιστική συμπεριφορά του υποδοχέα, προσφέροντας νέες δυνατότητες για την ανάπτυξη μη-ανταγωνιστικών τροποποιητών της σηματοδότησης. Ένα κρίσιμο επόμενο βήμα αποτελεί η διερεύνηση της σταθερότητας και της λειτουργικής συνέπειας των εντοπισμένων συμπλόκων μέσω προσομοιώσεων μοριακής δυναμικής (Molecular Dynamics, MD). Οι MD προσομοιώσεις μπορούν να επιβεβαιώσουν κατά πόσο οι θέσεις πρόσδεσης που προέκυψαν από το μοριακή πρόσδεση αποτελούν βιώσιμες διαμορφώσεις σε χρονικά εξελισσόμενο περιβάλλον, καθώς και αν η πρόσδεση οδηγεί σε αλλοστερικές μεταβολές που επηρεάζουν την ενεργοποίηση του TNFR1 ή τη δέσμευση του TNF-

α. Αντίστοιχα, η χρήση προσεγγίσεων πολλαπλών διαμορφώσεων πρόσδεσης θα επιτρέψει την εξερεύνηση διαφορετικών διαμορφώσεων του TNFR1, αποκαλύπτοντας ενδεχομένως κρυφές, δυναμικά επαγόμενες κοιλότητες με φαρμακολογικό ενδιαφέρον. Στο πεδίο της φαρμακοχημικής βελτιστοποίησης, τα υποψήφια μόρια που αναδείχθηκαν από τη μελέτη απαιτούν περαιτέρω ανάλυση μέσω προσδιορισμού του προφίλ ADMET, μοντελοποίησης QSAR και βελτιστοποίησης του φαρμακοφόρου μοντέλου. Η διερεύνηση των φυσικοχημικών και φαρμακοκινητικών χαρακτηριστικών τους θα επιτρέψει την επιλογή των πλέον κατάλληλων δομικών σκελετών για περαιτέρω ανάπτυξη, ενώ ο σχεδιασμός βιβλιοθηκών δομικών παραλλαγών (scaffold hopping) και η εφαρμογή υψηλής απόδοσης μοριακής πρόσδεσης (high-throughput docking) μπορούν να οδηγήσουν στην ανακάλυψη δομικά συγγενών ενώσεων με αυξημένη εκλεκτικότητα ή βελτιωμένη βιοδραστικότητα.

Κομβικής σημασίας για τη συνέχεια της έρευνας αποτελεί η πειραματική επικύρωση των υπολογιστικών αποτελεσμάτων. Η ένωση που κατέγραψε τη σταθερότερη πρόσδεση και στις τρεις μεθοδολογίες (AutoDock Vina, CNN, AutoDock4.2) πρέπει να διερευνηθεί μέσω *in vitro* και *in vivo* προσεγγίσεων. Επιπλέον, η εφαρμογή κατευθυνόμενης κατά τόπους μεταλλαξιογένεσης (site-directed mutagenesis) στα κρίσιμα αμινοξικά κατάλοιπα του TNFR1 αποτελεί ένα ιδιαίτερα σημαντικό βήμα για την πειραματική επιβεβαίωση των υπολογιστικών ευρημάτων. Θα συμβάλει στην επαλήθευση των προβλεπόμενων θέσεων πρόσδεσης και στην κατανόηση πιθανών αλλοστερικών μηχανισμών. Η στοχευμένη τροποποίηση των καταλοίπων Lys75, Ser74, Arg77, Glu109, Asn110, Leu111 και Phe112, τα οποία συμμετέχουν στη διαμόρφωση της κύριας περιοχής πρόσδεσης ή εναλλακτικών αλλοστερικών θυλάκων, θα μπορούσε να αποσαφηνίσει τον ακριβή ρόλο τους στη δέσμευση των υποψήφιων μορίων. Μέσω της σύγκρισης των ενεργειακών παραμέτρων και της σταθερότητας πρόσδεσης πριν και μετά τις μεταλλάξεις, είναι δυνατόν να διαπιστωθεί αν οι θέσεις αυτές αποτελούν δομικά ή λειτουργικά σημεία-κλειδιά για τη ρύθμιση της ενεργότητας του υποδοχέα.

Το υπολογιστικό πλαίσιο που αναπτύχθηκε στη συγκεκριμένη εργασία προσφέρει σημαντικές δυνατότητες επέκτασης. Ο συνδυασμός τυφλής μοριακής πρόσδεσης, βαθμολόγησης AutoDock 4.2 και πρόβλεψης μέσω του CNN μπορεί να εφαρμοστεί για την εξέταση και άλλων πρωτεϊνών που σχετίζονται με αυτοφλεγμονώδεις νόσους, όπως οι υποδοχείς IL-1R και IL-6R, καθώς και το σύμπλοκο NLRP3. Η δυνατότητα χρήσης του ήδη εκπαιδευμένου μοντέλου CNN σε νέα χημικά σύνολα ή σε εναλλακτικούς στόχους καθιστά τη μεθοδολογία επεκτάσιμη και ευέλικτη, προσφέροντας ένα ενιαίο πλαίσιο για εικονικό σχεδιασμό φαρμάκων μεγάλης κλίμακας. Τέλος, η διεύρυνση της προσέγγισης αυτής σε άλλα αυτοφλεγμονώδη σύνδρομα, όπως το FMF (MEFV), το HIDS/MKD (MVK) και οι παθήσεις που σχετίζονται με το NLRP3, θα μπορούσε να οδηγήσει στη δημιουργία νέων θεραπευτικών στρατηγικών για παθήσεις που σήμερα αντιμετωπίζονται κυρίως συμπτωματικά. Η συστηματική ενσωμάτωση μοριακής πρόσδεσης, μηχανικής μάθησης και προηγμένων χημειοπληροφορικών τεχνικών έχει τη δυνατότητα να επιταχύνει σημαντικά τις πρώιμες φάσεις ανακάλυψης φαρμάκων, συμβάλλοντας ουσιαστικά στην κατανόηση και αντιμετώπιση σπάνιων αλλά κλινικά σημαντικών αυτοφλεγμονωδών διαταραχών.

5. Βιβλιογραφία

- 1) Havnaer A, Han G. Autoinflammatory Disorders: A Review and Update on Pathogenesis and Treatment. *Am J Clin Dermatol*. 2019 Aug 17;20(4).
- 2) McDermott MF, Aksentijevich I. The autoinflammatory syndromes. *Curr Opin Allergy Clin Immunol*. 2002 Dec;2(6).
- 3) Arend WP, Malyak M, Guthridge CJ, Gabay C. Interleukin-1 Receptor Antagonist: Role in Biology. *Annu Rev Immunol*. 1998 Apr;16(1).
- 4) Rubartelli A. Autoinflammatory diseases. *Immunol Lett*. 2014 Oct;161(2).
- 5) Georgin-Lavialle S, Fayand A, Rodrigues F, Bachmeyer C, Savey L, Grateau G. Autoinflammatory diseases: State of the art. *Presse Med*. 2019 Feb;48(1).
- 6) Masumoto J, Zhou W, Morikawa S, Hosokawa S, Taguchi H, Yamamoto T, et al. Molecular biology of autoinflammatory diseases. *Inflamm Regen*. 2021 Dec 11;41(1).
- 7) Ryan JG, Aksentijevich I. Tumor necrosis factor receptor-associated periodic syndrome: Toward a molecular understanding of the systemic autoinflammatory diseases. *Arthritis Rheum*. 2009 Jan 30;60(1).
- 8) Galon J, Aksentijevich I, McDermott MF, O'Shea JJ, Kastner DL. TNFRSF1A mutations and autoinflammatory syndromes. *Curr Opin Immunol*. 2000 Aug;12(4).
- 9) Caso F, Rigante D, Vitale A, Lucherini OM, Costa L, Atteno M, et al. Monogenic Autoinflammatory Syndromes: State of the Art on Genetic, Clinical, and Therapeutic Issues. *Int J Rheumatol*. 2013;2013.
- 10) Simon A, Park H, Maddipati R, Lobito AA, Bulua AC, Jackson AJ, et al. Concerted action of wild-type and mutant TNF receptors enhances inflammation in TNF receptor 1-associated periodic fever syndrome. *Proceedings of the National Academy of Sciences*. 2010 May 25;107(21).
- 11) Cantarini L, Lucherini OM, Cimaz R, Baldari CT, Bellisai F, Paccani SR, et al. Idiopathic Recurrent Pericarditis Refractory to Colchicine Treatment Can Reveal Tumor Necrosis Factor Receptor-Associated Periodic Syndrome. *Int J Immunopathol Pharmacol*. 2009 Oct 1;22(4).
- 12) Cudrici C, Deutch N, Aksentijevich I. Revisiting TNF Receptor-Associated Periodic Syndrome (TRAPS): Current Perspectives. *Int J Mol Sci*. 2020 May 5;21(9).
- 13) Hull KM, Drewe E, Aksentijevich I, Singh HK, Wong K, McDermott EM, et al. The TNF Receptor-Associated Periodic Syndrome (TRAPS). *Medicine*. 2002 Sep;81(5).
- 14) Marino A, Tirelli F, Giani T, Cimaz R. Periodic fever syndromes and the autoinflammatory diseases (AIDs). *J Transl Autoimmun*. 2020;3.
- 15) Kuemmerle-Deschner JB. CAPS- pathogenesis, presentation and treatment of an autoinflammatory disease. *Semin Immunopathol*. 2015 Jul 12;37(4).
- 16) Neven B, Prieur AM, dit Maire PQ. Cryopyrinopathies: update on pathogenesis and treatment. *Nat Clin Pract Rheumatol*. 2008 Sep 29;4(9).

- 17) Yu JR, Leslie KS. Cryopyrin-Associated Periodic Syndrome: An Update on Diagnosis and Treatment Response. *Curr Allergy Asthma Rep.* 2011 Feb 23;11(1).
- 18) Booshehri LM, Hoffman HM. CAPS and NLRP3. *J Clin Immunol.* 2019 Apr 10;39(3).
- 19) Tran TA. Muckle-Wells syndrome: clinical perspectives. *Open Access Rheumatol.* 2017 Jul 11; 9.
- 20) Paccaud Y, Berthet G, Von Scheven-Gête A, Vaudaux B, Mivelaz Y, Hofer M, et al. Neonatal treatment of CINCA syndrome. *Pediatric Rheumatology.* 2014 Dec 15;12(1).
- 21) Lengvári L, Takács K, Lengyel A, Pálkás A, Wouters CH, Koné-Paut I, et al. Mevalonate kinase deficiency: an updated clinical overview and revision of the SHARE recommendations. *Front Immunol.* 2024 Nov 12;15.
- 22) Barron K, Athreya B, Kastner D. Periodic Fever Syndromes and other inherited autoinflammatory diseases. In: *Textbook of Pediatric Rheumatology.* Elsevier; 2011.
- 23) Haas D, Hoffmann GF. Mevalonate kinase deficiencies: from mevalonic aciduria to hyperimmunoglobulinemia D syndrome. *Orphanet J Rare Dis.* 2006 Dec 26;1(1).
- 24) Schulert G, Favier L. Mevalonate kinase deficiency: current perspectives. *Appl Clin Genet.* 2016 Jul;Volume 9.
- 25) van der Burgh R, ter Haar NM, Boes ML, Frenkel J. Mevalonate kinase deficiency, a metabolic autoinflammatory disease. *Clinical Immunology.* 2013 Jun;147(3).
- 26) Zhang S. Natural history of mevalonate kinase deficiency: a literature review. *Pediatric Rheumatology.* 2016 Dec 4;14(1).
- 27) Rafiq NK, Lachmann H, Joensen F, Herlin T, Brogan PA. Tocilizumab for the Treatment of Mevalonate Kinase Deficiency. *Case Rep Pediatr.* 2018 Jan 26;2018(1).
- 28) Blundell TL, Sibanda BL, Montalvão RW, Brewerton S, Chelliah V, Worth CL, et al. Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 2006 Mar 29;361(1467).
- 29) Katara P. Role of bioinformatics and pharmacogenomics in drug discovery and development process. *Network Modeling Analysis in Health Informatics and Bioinformatics.* 2013 Dec 10;2(4).
- 30) Gupta M, Sharma R, Kumar A. Docking techniques in pharmacology: How much promising? *Comput Biol Chem.* 2018 Oct;76.
- 31) Morris GM, Lim-Wilby M. Molecular Docking. In 2008.
- 32) Dhaliwal B, Chen Y. Computational Resources for Protein Modelling and Drug Discovery Applications. *Infect Disord Drug Targets.* 2009 Oct 1;9(5).
- 33) Wang G, Zhu W. Molecular Docking for Drug Discovery and Development: A Widely Used Approach but Far from Perfect. *Future Med Chem.* 2016 Sep 31;8(14).
- 34) Espe S. MalaCards: The Human Disease Database. *Journal of the Medical Library Association.* 2018 Jan 12;106(1). (Available at: <http://www.malacards.org>)

- 35) Koscielny G, An P, Carvalho-Silva D, Cham JA, Fumis L, Gasparyan R, et al. Open Targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res.* 2017 Jan 4;45(D1). (Available at: <http://www.opentargets.org>)
- 36) Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* 2008 Jan 1;36. (Available at: <http://go.drugbank.com>)
- 37) Berman HM, Burley SK. Protein Data Bank (PDB): Fifty-three years young and having a transformative impact on science and society. *Q Rev Biophys.* 2025 Feb 20;58. (Available at: <https://www.rcsb.org/>)
- 38) Böcker A, Derksen S, Schmidt E, Teckentrup A, Schneider G. A Hierarchical Clustering Approach for Large Compound Libraries. *J Chem Inf Model.* 2005 Jul 1;45(4).
- 39) Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J Chem Inf Model.* 2021 Aug 23;61(8).
- 40) Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010 Jan 30;31(2).
- 41) Hou X, Du J, Zhang J, Du L, Fang H, Li M. How to Improve Docking Accuracy of AutoDock4.2: A Case Study Using Different Electrostatic Potentials. *J Chem Inf Model.* 2013 Jan 28;53(1).
- 42) Goodsell DS, Morris GM, Olson AJ. Automated docking of flexible ligands: Applications of autodock. *Journal of Molecular Recognition.* 1996 Jan;9(1).
- 43) Ragoza M, Hochuli J, Idrobo E, Sunseri J, Koes DR. Protein–Ligand Scoring with Convolutional Neural Networks. *J Chem Inf Model.* 2017 Apr 24;57(4).
- 44) Kiranyaz S, Avci O, Abdeljaber O, Ince T, Gabbouj M, Inman DJ. 1D convolutional neural networks and applications: A survey. *Mech Syst Signal Process.* 2021 Apr;151.
- 45) Vittorio S, Lunghini F, Morerio P, Gadioli D, Orlandini S, Silva P, et al. Addressing docking pose selection with structure-based deep learning: Recent advances, challenges and opportunities. *Comput Struct Biotechnol J.* 2024 Dec;23.
- 46) Athanasiadis E, Cournia Z, Spyrou G. ChemBioServer: a web-based pipeline for filtering, clustering and visualization of chemical compounds used in drug discovery. *Bioinformatics.* 2012 Nov 15;28(22). (Available at: <https://chembioserver.vi-seem.eu/>).
- 47) Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 2009 Dec 27;30(16).
- 48) Cosconati S, Forli S, Perryman AL, Harris R, Goodsell DS, Olson AJ. Virtual screening with AutoDock: theory and practice. *Expert Opin Drug Discov.* 2010 Jun 23;5(6).
- 49) Forli S, Olson AJ. A Force Field with Discrete Displaceable Waters and Desolvation Entropy for Hydrated Ligand Docking. *J Med Chem.* 2012 Jan 26;55(2).

- 50) Wang Y, Jiao Q, Wang J, Cai X, Zhao W, Cui X. Prediction of protein-ligand binding affinity with deep learning. *Comput Struct Biotechnol J*. 2023;21.
- 51) Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res*. 2018 Jul 2;46(W1). (Available at: <http://cast.engr.uic.edu>).
- 52) Volkamer A, Kuhn D, Grombacher T, Rippmann F, Rarey M. Combining Global and Local Measures for Structure-Based Druggability Predictions. *J Chem Inf Model*. 2012 Feb 27;52(2). (Available at: <https://proteins.plus/help/dogsite>).
- 53) Saddala MS, Huang H. Identification of novel inhibitors for TNF α , TNFR1 and TNF α -TNFR1 complex using pharmacophore-based approaches. *J Transl Med*. 2019 Dec 2;17(1).
- 54) Chan DS, Lee H, Yang F, Che C, Wong CCL, Abagyan R, Leung C, Ma D. Structure-Based Discovery of Natural-Product-like TNF- α Inhibitors. *Angewandte Chemie International Edition*. 2010 Apr 6;49(16).

6. Παράρτημα

I. Δομή Αρχείου SDF

Η βασική δομή ενός αρχείου τύπου SDF (Structure Data File), το οποίο περιλαμβάνει πληροφορίες για τη γεωμετρία των ατόμων, τους δεσμούς, καθώς και επιπλέον ιδιότητες που σχετίζονται με τη χημική ταυτότητα κάθε ένωσης.

```
Mrv0541 02231214272D

  6  5  0  0  0  0          999 V2000
    2.0928   0.2063   0.0000  O   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
    3.5217   1.0312   0.0000  O   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
    2.8072  -1.0312   0.0000  O   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
    3.5217   0.2063   0.0000  C   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
    4.2361  -0.2063   0.0000  C   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
    2.8072  -0.2063   0.0000  C   0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
  1  6  1  0  0  0  0
  2  4  2  0  0  0  0
  3  6  2  0  0  0  0
  4  5  1  0  0  0  0
  4  6  1  0  0  0  0
M  END
> <DATABASE_ID>
DB00119

> <DATABASE_NAME>
drugbank

> <SMILES>
CC(=O)C(O)=O

> <INCHI_IDENTIFIER>
InChI=1S/C3H4O3/c1-2(4)3(5)6/h1H3,(H,5,6)

> <INCHI_KEY>
LCTONWCANYUPML-UHFFFAOYSA-N

> <FORMULA>
C3H4O3

> <MOLECULAR_WEIGHT>
88.0621

> <EXACT_MASS>
88.016043994
```

II. Δομή Αρχείου PDBQT

Η βασική δομή ενός αρχείου PDBQT. Περιλαμβάνει πληροφορίες για τις συντεταγμένες, τα ατομικά φορτία και τις περιστρεφόμενες δεσμικές γωνίες των μορίων.

```
REMARK Name = 114
REMARK 4 active torsions:
REMARK status: ('A' for Active; 'I' for Inactive)
REMARK 1 A between atoms: P_1 and O_2
REMARK 2 A between atoms: O_2 and C_11
REMARK 3 A between atoms: C_9 and C_11
REMARK 4 A between atoms: C_10 and C_16
REMARK
REMARK      x      y      z      vdW Elec      q      Type
REMARK
ROOT
ATOM      1 O LIG A 1      2.093  0.206  0.000  0.00  0.00      -0.285 OA
ATOM      2 N LIG A 1      3.522 -1.444  0.000  0.00  0.00      -0.245 NA
ATOM      3 C LIG A 1      4.237 -0.206  0.000  0.00  0.00      +0.033 A
ATOM      4 C LIG A 1      3.522  0.206  0.000  0.00  0.00      +0.082 A
ATOM      5 C LIG A 1      2.808 -0.206  0.000  0.00  0.00      +0.212 A
ATOM      6 C LIG A 1      2.808 -1.031  0.000  0.00  0.00      +0.098 A
ATOM      7 C LIG A 1      4.237 -1.031  0.000  0.00  0.00      +0.097 A
ATOM      8 C LIG A 1      2.093 -1.444  0.000  0.00  0.00      +0.056 C
ENDROOT
BRANCH 4 10
ATOM      9 O LIG A 1      2.808  1.444  0.000  0.00  0.00      -0.283 OA
ATOM     10 C LIG A 1      3.522  1.031  0.000  0.00  0.00      +0.237 C
ENDBRANCH 4 10
BRANCH 3 11
ATOM     11 C LIG A 1      4.951  0.206  0.000  0.00  0.00      +0.182 C
BRANCH 11 12
ATOM     12 O LIG A 1      5.665 -0.206  0.000  0.00  0.00      -0.265 OA
BRANCH 12 13
ATOM     13 P LIG A 1      6.380  0.206  0.000  0.00  0.00      +0.533 P
ATOM     14 O LIG A 1      6.793 -0.508  0.000  0.00  0.00      -0.115 OA
ATOM     15 O LIG A 1      7.094  0.619  0.000  0.00  0.00      -0.115 OA
ATOM     16 O LIG A 1      5.967  0.921  0.000  0.00  0.00      -0.222 OA
ENDBRANCH 12 13
ENDBRANCH 11 12
ENDBRANCH 3 11
TORSDOF 4
```

III. Παράμετροι για το AutoDock Vina

Το αρχείο ρυθμίσεων (configuration file) που χρησιμοποιήθηκε για την εκτέλεση του προγράμματος AutoDock Vina με τις παραμέτρους της διαδικασίας μοριακής πρόσδεσης. Η εκτέλεση πραγματοποιήθηκε μέσω αρχείου τύπου .bat, το οποίο περιείχε την εντολή κλήσης του Vina με τις αντίστοιχες παραμέτρους εισόδου και εξόδου.

```
receptor = 1ext_w_o_heter.pdbqt  
center_x = -1.433  
center_y = 26.755  
center_z = -6.968  
size_x = 126  
size_y = 126  
size_z = 126  
num_modes = 10  
energy_range = 4
```

```
@echo off  
  
for %%a in ("C:\Users\Georgia\Desktop\Autoinflam\babel5\*.pdbqt") do (  
    "C:\Users\Georgia\Desktop\AutoDock\Vina\vina.exe" --ligand "%%a" --config conf.txt --log "%%a".txt  
)  
pause
```

IV. Αποτελέσματα AutoDock Vina

Ενδεικτικά η έξοδος των αποτελεσμάτων του προγράμματος AutoDock Vina, όπως προέκυψε μετά την ολοκλήρωση της διαδικασίας μοριακής πρόσδεσης. Περιλαμβάνονται οι διαφορετικοί τρόποι πρόσδεσης (modes), οι αντίστοιχες ενέργειες δέσμευσης (affinity, kcal/mol), καθώς και οι τιμές RMSD.

```
WARNING: The search space volume > 27000 Angstrom^3 (See FAQ)
Output will be C:\Users\Georgia\Desktop\Autoinflam\babel\structures1004_out.pdbqt
Detected 4 CPUs
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: -1192096100
Performing search ... done.
Refining results ... done.
```

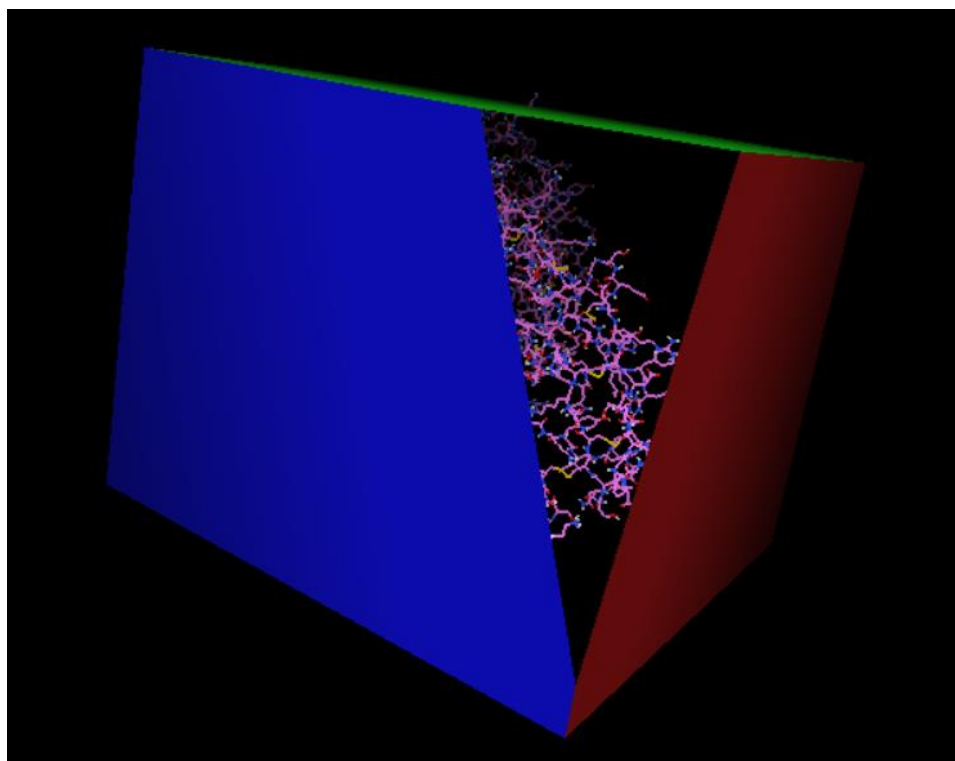
mode	affinity (kcal/mol)	dist from best mode rmsd l.b. rmsd u.b.
1	-6.0	0.000 0.000
2	-5.7	19.793 21.076
3	-5.7	9.627 10.447
4	-5.4	17.234 18.041
5	-5.4	14.302 17.457
6	-5.4	56.783 57.878
7	-5.4	9.856 11.941
8	-5.1	18.460 20.813
9	-5.1	17.297 18.408
10	-5.1	11.981 13.790

```
Writing output ... done.
```

V. Αρχείο Παραμέτρων Πλέγματος (gpf)

Το αρχείο παραμέτρων πλέγματος (grid parameter file, gpf) που χρησιμοποιήθηκε στο πρόγραμμα AutoDock 4.2 για τον καθορισμό της περιοχής αναζήτησης πρόσδεσης γύρω από την πρωτεΐνη TNFR1. Η συνοδευτική εικόνα απεικονίζει το τρισδιάστατο πλέγμα πρόσδεσης (grid box), όπως ορίστηκε γύρω από την επιφάνεια της πρωτεΐνης κατά τη φάση προετοιμασίας του docking.

```
npts 94 76 86          # num.grid points in xyz
gridfld 1ext.maps.fld   # grid_data_file
spacing 0.883333333333  # spacing(A)
receptor_types A C HD Mg N OA S SA # receptor atom types
ligand_types A C NA OA N S # ligand atom types
receptor 1ext.pdbqt      # macromolecule
gridcenter -0.205 28.667 -4.696    # xyz-coordinates or auto
smooth 0.5              # store minimum energy w/in rad(A)
map 1ext.A.map          # atom-specific affinity map
map 1ext.C.map          # atom-specific affinity map
map 1ext.NA.map         # atom-specific affinity map
map 1ext.OA.map         # atom-specific affinity map
map 1ext.N.map          # atom-specific affinity map
map 1ext.S.map          # atom-specific affinity map
elecmap 1ext.e.map      # electrostatic potential map
dsolvmap 1ext.d.map     # desolvation potential map
dielectric -0.1465      # <0, AD4 distance-dep.diel;>0, constant
```



VI. Αποτελέσματα AutoDock4.2

Αποσπάσματα από το αρχείο αποτελεσμάτων του προγράμματος AutoDock4.2, τα οποία περιλαμβάνουν την ανάλυση συσσωμάτωσης (cluster analysis) και τον πίνακα RMSD των παραγόμενων διαμορφώσεων.

CLUSTER ANALYSIS OF CONFORMATIONS												
Number of conformations = 10												
RMSD cluster analysis will be performed using the ligand atoms only (35 / 35 total atoms).												
Outputting structurally similar clusters, ranked in order of increasing energy.												
Number of distinct conformational clusters found = 8, out of 10 runs, Using an rmsd-tolerance of 2.0 Å												
CLUSTERING HISTOGRAM												
Clus- ter Rank	Lowest Binding Energy	Run	Mean Binding Energy	Num in Clus	Histogram							
					5	10	15	20	25	30	35	
1	-7.95	7	-7.68	3	###							
2	-7.45	2	-7.45	1	#							
3	-7.41	4	-7.41	1	#							
4	-6.80	5	-6.80	1	#							
5	-6.52	3	-6.52	1	#							
6	-5.83	10	-5.83	1	#							
7	-5.64	1	-5.64	1	#							
8	-5.59	9	-5.59	1	#							

Number of multi-member conformational clusters found = 1, out of 10 runs.						
RMSD TABLE						
Rank	Sub- Rank	Run	Binding Energy	Cluster RMSD	Reference RMSD	Grep Pattern
1	1	7	-7.95	0.00	16.13	RANKING
1	2	8	-7.70	0.92	15.95	RANKING
1	3	6	-7.40	0.88	16.16	RANKING
2	1	2	-7.45	0.00	16.86	RANKING
3	1	4	-7.41	0.00	9.08	RANKING
4	1	5	-6.80	0.00	27.83	RANKING
5	1	3	-6.52	0.00	64.43	RANKING
6	1	10	-5.83	0.00	32.90	RANKING
7	1	1	-5.64	0.00	17.80	RANKING
8	1	9	-5.59	0.00	63.45	RANKING

VII. Ενέργειες δέσμευσης στα AutoDock Vina, AutoDock4.2, και στο μοντέλο CNN

Συγκεντρωτικός πίνακα των 100 υποψήφιων ενώσεων, με τις αντίστοιχες ενέργειες δέσμευσης (ΔG) που προέκυψαν από την ανάλυση στα προγράμματα AutoDock Vina και AutoDock4.2, καθώς και από το μοντέλο CNN.

Κωδικός Drugbank	Όνομα Χημικής Ένωσης	Ενέργειες Δέσμευσης AutoDock Vina (kcal/mol)	Ενέργειες Δέσμευσης CNN Μοντέλο (kcal/mol)	Ενέργειες Δέσμευσης AutoDock4.2 (kcal/mol)
DB01177	Idarubicin	-9	-8.4825	-6.08
DB01199	Tubocurarine	-10	-9.75693	-11.01
DB01200	Bromocriptine	-9.3	-9.8129	-7.14
DB01211	Clarithromycin	-8.2	-8.09795	-5.54
DB01251	Gliquidone	-8.6	-9.16535	-8.23
DB01336	Metocurine	-8.3	-8.27441	-6.93
DB01338	Pipecuronium	-8.2	-7.49985	-5.7
DB01339	Vecuronium	-8.8	-8.21848	-7.62
DB01410	Ciclesonide	-8	-8.44889	-7.49
DB01411	Pranlukast	-8.4	-8.35004	-6.94
DB01697	beta-cellobiose	-8.9	-6.81886	-5.97
DB01720	(2Z)-2-(Benzoylamino)-3-[4-(2-bromophenoxy)phenyl]acrylic acid	-8.2	-7.22365	-5.33
DB01732	(4R,5S,6S,7R)-1,3-dibenzyl-4,7-bis(phenoxyethyl)-5,6-dihydroxy-1,3 diazepan-2-one	-8.1	-6.68528	-3.74
DB01873	Epothilone D	-8.8	-8.33063	-6.92
DB01933	7-Hydroxystaurosporine	-8.7	-9.66351	-7.03
DB02114	Daidzin	-9.1	-8.26831	-5.64
DB02118	(6S)-4-Benzyl-6-(1-methyl-2,2-dioxido-1,3-dihydro-2,1-benzothiazol-5-yl)-3-morpholinone	-8.6	-7.59864	-8.06
DB02152	K-252a	-8.9	-9.41252	-6.95
DB02285	Protoporphyrin	-8.3	-8.13436	-7.5
DB02322	Heparin Disaccharide I-S	-8	-7.87137	-7.79
DB02519	Indirubin-5-sulphonate	-8.4	-8.15508	-7.91
DB02621	Latrunculin A	-8.4	-6.97702	-7.21
DB02639	7-(alpha-D-Glucopyranosyloxy)-4-methyl-2H-1-benzopyran-2-one	-8.6	-7.17545	-5.3
DB02703	Fusidic acid	-8	-8.02462	-6.96
DB02704	Diol-Based HIV-1 protease inhibitor 4	-10.3	-7.52113	-5.34
DB02705	6-[N-(1-Isopropyl-1,2,3,4-Tetrahydro-7-Isoquinoliny)Carbamyl]-2-Naphthalenecarboxamidine	-8.5	-7.75972	-6.95
DB02717	Cellotetraose	-9.1	-6.81886	-5.40

DB02759	4-methyl-umbelliferyl-N-acetyl-chitobiose	-9.4	-7.69229	-6.33
DB02859	Soraphen A(1 α)	-8.2	-8.14786	-6.41
DB02977	PNU177836	-8.7	-8.67093	0.54, +1.06
DB03005	3,8-Diamino-6-Phenyl-5-[6-[1-[2-[(1,2,3,4-Tetrahydro-9-Acridinyl)Amino]Ethyl]-1h-1,2,3-Triazol-5-Yl]Hexyl]-Phenanthridinium	-8.6	-8.89336	-5.79
DB03038	LY341770	-8.8	-9.01204	-6.64
DB03067	4-{2,4-Bis[(3-Nitrobenzoyl)Amino]Phenoxy}Phthalic Acid	-8.6	-8.10379	-7.75
DB03104	2-[4-[(Z)-2-Acetamido-3-oxo-3-[[[(3S)-2-oxo-1-[(4-phenylphenyl)methyl]azepan-3-yl]amino]prop-1-enyl]-2-formylphenyl]acetic acid	-8.4	-8.63903	-6.42
DB03159	2-{5-[amino(iminio)methyl]-1H-indol-2-yl}-6-phenylbenzenolate	-9	-8.57386	-7.71
DB03219	11-Deoxy-Beta-Rhodomyacin	-10.6	-10.3795	-4.45
DB03268	4-[2-Acetylamino-2-(1-biphenyl-4-ylmethyl-2-oxo-azepan-3-ylcarbamoyl)-ethyl]-2-formylbenzoic acid	-8.4	-8.24448	-7.43
DB03344	Inositol-(1,3,4,5,6)-Pentakisphosphate	-8.4	-6.02301	-4.27
DB03480	Brequinar Analog	-8.1	-7.60176	-8.21
DB03495	4,6-Dideoxy-4-{[4,5,6-Trihydroxy-3-(Hydroxymethyl)Cyclohex-2-En-1-Yl]Amino}-Alpha-D-Lyxohexopyranosyl-(1->4)-Alpha-D-Threo-Hexopyranosyl-(1->6)-Alpha-L-Threo-Hexopyranose	-8.7	-8.72975	-5.81
DB03621	(1R,9S,12S,13R,14S,17R,18E,20S,21R,23S,24R,25S,27R)-17-ethyl-1,14,20-trihydroxy-23,25-dimethoxy-12-[(1E)-1-[(1R,3R,4R)-3-methoxy-4-[(1-methylindol-5-yl)oxy]cyclohexyl]prop-1-en-2-yl]-13,19,21,27-tetramethyl-11,28-dioxa-4-azatricyclo[22.3.1.0 ^{4,9}]octacos-18-ene-2,3,10,16-tetrone, L-709,587	-9.6	-9.82245	-2.96
DB03671	Digoxigenin	-8.5	-8.42717	-5.79
DB03823	Epigallocatechin	-8.3	-7.00106	-5.81
DB03878	N-[4-Methyl-3-[[4-(3-Pyridinyl)-2-Pyrimidinyl]Amino]Phenyl]-3-Pyridinecarboxamide	-8.2	-8.4334	-7.63
DB03971	Acarbose Derived Hexasaccharide	-9.7	-9.29756	0.77

DB03990	4'-nitrophenyl-3i-thiolamaritrioside	-8.1	-7.85478	-3.13
DB04011	2'-(4-Dimethylaminophenyl)-5-(4-Methyl-1-Piperazinyl)-2,5'-Bi-Benzimidazole	-8	-7.39192	-7.96
DB04030	1,3,4,9-Tetrahydro-2-(Hydroxybenzoyl)-9-[(4-Hydroxyphenyl)methyl]-6-Methoxy-2h-Pyrido[3,4-B]Indole	-8.2	-8.04448	-6.91
DB04118	N-Coeleneterazine	-8.9	-9.35007	-7.07
DB04131	10-(4-Dimethylamino-5-Hydroxy-6-Methyl-Tetrahydro-Pyran-2-Yloxy)-8-Ethyl-1,8,11-Trihydroxy-7,8,9,10-Tetrahydro-Naphthacene-5,12-Dione	-9.4	-9.57643	-9.56
DB04154	N-Methyl-N-[3-(6-Phenyl[1,2,4]Triazolo[4,3-B]Pyridazin-3-Yl)Phenyl]Acetamide	-8	-7.92805	-6.83
DB04226	Dihydro-Acarbose	-9.2	-7.63251	-6.95
DB04255	[5-(2-hydroxy-indan-1-ylcarbamoyl)-3,4-dihydroxy-2,5-[dibenzyl-oxy]-pentanoyl]-valinyl-amido-methane	-9.3	-8.67991	-2.91
DB04268	Methyl-umbelliferol-N-acetyl-chitotriose	-9.1	-7.69229	-5.45
DB04405	(2S)-2-[2-(1,3-dioxoisindol-2-yl)ethyl]-4-[4-(4-ethoxyphenyl)phenyl]-4-oxobutanoic acid	-9.2	-8.80922	-6.38
DB04439	Modified Acarbose Pentasaccharide	-9	-9.00008	-2.14
DB04453	4-O-(4,6-Dideoxy-4-{[4-[(4-O-Hexopyranosylhexopyranosyl)Oxy]-5,6-Dihydroxy-3-(Hydroxymethyl)Cyclohex-2-En-1-Yl]Amino}Hexopyranosyl)Hexopyranose	-10.2	-9.09486	-0.25
DB04678	H TYPE II TRISACCHARIDE	-8.6	-7.49038	-0.00
DB04703	(S)-(-)-hesperidin	-9.4	-8.72058	-2.37
DB04708	(3S)-oxolan-3-yl N-[(2S,3S)-4-[(2S)-2-benzyl-4-[(1R,3R)-3-(carbamoylmethyl)-2,3-dihydro-1H-inden-1-yl]-3-oxo-1H-pyrrol-2-yl]-3-hydroxy-1-phenylbutan-2-yl]carbamate	-8.5	-8.694	-6.89
DB04739	4-[(4-methylpiperazin-1-yl)methyl]-N-(3-{[4-(pyridin-3-yl)pyrimidin-2-yl]amino}phenyl)benzamide	-8.7	-9.06823	-6.45

DB04805	virginiamycin S1	-9.4	-9.99962	-6.98
DB04845	Ixabepilone	-9.1	-9.15599	-8.01
DB05109	Yondelis	-9.5	-10.3556	-7.71
DB05129	Elsamitrucin	-9.2	-9.31168	-5.16
DB05423	ORG-34517	-8.6	-8.61162	-7.25
DB05450	Smilagenin	-8.6	-8.24446	-5.75
DB05983	Bardoxolone methyl	-8.6	-8.46747	-2.66
DB06011	AMG-222	-8.8	-8.58202	-9.46
DB06409	Morphine glucuronide	-8.4	-8.55714	-6.12
DB06419	Cethromycin	-9.3	-10.0116	-6.52
DB06532	Pafuramidine	-8.6	-6.7227	-4.89
DB06569	P57 (glycoside)	-8.6	-8.41048	-4.29
DB06595	Midostaurin	-9.9	-10.1063	-7.06
DB06743	Ginkgolide A	-9.2	-7.8322	-6.07
DB06922	2-[(1R)-1-carboxy-2-naphthalen-1-ylethyl]-1,3-dioxo-2,3-dihydro-1H-isoindole-5-carboxylic acid	-8.8	-8.30631	-5.26
DB07021	(7R,8R)-8-(2,4,5-trifluorophenyl)-6,7,8,9-tetrahydroimidazo[1,2-a:4,5-c']dipyridin-7-amine	-9	-7.75925	-8.44
DB07072	(1S,2R,5S)-5-[3-(trifluoromethyl)-5H,6H,8H-[1,2,4]triazolo[4,3-a]pyrazin-7-yl]-2-(2,4,5-trifluorophenyl)cyclohexan-1-amine	-8.6	-8.10286	-9.24
DB07141	5-[(3R)-3-(5-methoxy-4'-methylbiphenyl-3-yl)but-1-yn-1-yl]-6-methylpyrimidine-2,4-diamine	-8.7	-8.83452	-8.37
DB07189	(1S,3R,6S)-4-oxo-6-{4-[(2-phenylquinolin-4-yl)methoxy]phenyl}-5-azaspiro[2.4]heptane-1-carboxylic acid	-8.9	-8.53358	-7.82
DB07213	(5-{3-[5-(piperidin-1-ylmethyl)-1H-indol-2-yl]-1H-indazol-6-yl}-2H-1,2,3-triazol-4-yl)methanol	-8.6	-8.92146	-8.39
DB07233	N-{[4-(but-2-yn-1-yloxy)phenyl]sulfonyl}-5-methyl-D-tryptophan	-8.5	-7.07316	-8.32
DB07354	2,3-DIMETHOXY-12H-[1,3]DIOXOLO[5,6]INDENO[1,2-C]ISOQUINOLIN-6-IUM	-8.6	-7.77355	-6.16
DB07404	1-hydroxy-2-{3'-phenyl-[1,1'-biphenyl]-3-yl}-1-phosphonoethylphosphonic acid	-8.3	-8.21567	-8.74
DB07410	[2-(3-DIBENZOFURAN-4-YL-PHENYL)-1-HYDROXY-1-PHOSPHONO-ETHYL]-PHOSPHONIC ACID	-8.2	-9.01245	-4.63

DB08006	N-(anthracen-2-yl)-5-methyl-[1,2,4]triazolo[1,5-a]pyrimidin-7-amine	-9.7	-6.54121	-8.54
DB07837	[4-(5-naphthalen-2-yl-1H-pyrrolo[2,3-b]pyridin-3-yl)phenyl]acetic acid	-9.4	-6.79749	-8.98
DB08067	4-[(2-{4-[(cyclopropylcarbamoyl)amino]-1H-pyrazol-3-yl}-3H-1,3-benzodiazol-5-yl)methyl]morpholin-4-ium	-9.4	-7.16249	-7.40
DB08572	4-{[4-amino-6-(cyclohexylmethoxy)-5-nitrosopyrimidin-2-yl]amino}benzamide	-9.2	-7.07091	-6.87
DB08674	(20S)-19,20,21,22-TETRAHYDRO-19-OXO-5H-18,20-ETHANO-12,14-ETHENO-6,10-METHENO-18H-BENZ[D]IMIDAZO[4,3-K][1,6,9,12]OXATRIAZA-CYCLOOCTADECOSINE-9-CARBONITRILE	-9.2	-8.04277	-9.99
DB07978	2-({2,3,5,6-tetrafluoro-4-[3-(trifluoromethoxy)phenyl]phenyl} carbamoyl)cyclopenta-1,3-diene-1-carboxylic acid	-9.1	-7.28522	-6.03
DB07861	(2R)-N-hydroxy-3-(naphthalen-2-yl)-2-(naphthalene-2-sulfonamido)propanamide	-8.7	-6.2479	-8.06
DB08676	(5S,18R)-31-oxo-20-oxa-2,6,11,13-tetraazahexacyclo[19.6.2.1^{2,5}.1^{15,19}.0^{9,13}.0^{24,28}]hentr iaconta-1(27),9,11,15,21(29),22,24(28),25-octaene-18-carbonitrile	-8.7	-7.13023	-9.93
DB08197	(5E,7S)-2-amino-7-(4-fluoro-2-pyridin-3-ylphenyl)-4-methyl-7,8-dihydroquinazolin-5(6H)-one oxime	-8.6	-7.57002	-6.78
DB08683	REL-(9R,12S)-9,10,11,12-TETRAHYDRO-9,12-EPOXY-1H-DIINDOLO[1,2,3-FG:3',2',1'-KL]PYRROLO[3,4-I][1,6]BENZODIAZOCINE-1,3(2H)-DIONE	-8.6	-6.81477	-7.77
DB08300	1-methyl-3-(naphthalen-2-yl)pyrazolo[3,4-d]pyrimidin-4-amine	-8.4	-5.41352	-7.08

DB08448	(4aS)-5-[(2,4-diaminopteridin-6-yl)methyl]-4a,5-dihydro-2H-dibenzo[b,f]azepin-8-ol	-8.4	-7.01459	-9.56
DB07834	N-(cyclopropylmethyl)-4-[2-methyl-5-(5-methyl-1,3,4-oxadiazol-2-yl)phenyl]benzamide	-8.3	-7.09844	-6.78
DB07855	(S)-phenyl[4-(9H-purin-6-yl)phenyl]methanamine	-8.3	-6.22405	-6.52

VIII. Κώδικας CNN

Ο κώδικας που αναπτύχθηκε για την εκπαίδευση και αξιολόγηση του CNN για την πρόβλεψη των ενεργειών δέσμησης (binding affinities) των μορίων προς την πρωτεΐνη TNFR1 (PDB ID: 1EXT).

```
from sklearn.decomposition import PCA
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import silhouette_samples, silhouette_score
from sklearn.cluster import KMeans
from sklearn import datasets, decomposition
from sklearn.manifold import TSNE
from rdkit import Chem, DataStructs
from rdkit.Chem import AllChem, MACCSkeys, Descriptors, Descriptors3D, Draw,
rdMolDescriptors, Draw, PandasTools
from rdkit.Chem.Draw import IPythonConsole
from sklearn.cluster import KMeans
from rdkit import Chem
from rdkit.Chem import AllChem
from rdkit import DataStructs
from rdkit import Chem
from rdkit.Chem import AllChem
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Conv1D, MaxPooling1D, Flatten
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Conv1D, MaxPooling1D, Flatten, Dropout
from tensorflow.keras import regularizers
from tensorflow.keras.callbacks import EarlyStopping
import matplotlib.pyplot as plt
from sklearn.preprocessing import MinMaxScaler

data['SMILES'] = ''

folder_path = 'C:/Users/Georgia/Desktop/Autoinflam/sdf/1-6000/'

for index, row in data.iterrows():
    ligand_name = row['file_name']
    sdf_file_path = folder_path + ligand_name + '.sdf'

    try:
        suppl = Chem.SDMolSupplier(sdf_file_path)
        molecule = next(suppl)

        if molecule:
            smiles = Chem.MolToSmiles(molecule)
            data.at[index, 'SMILES'] = smiles
        else:
            data.at[index, 'SMILES'] = 'Not found'

    except Exception as e:
        data.at[index, 'SMILES'] = 'Error: ' + str(e)
data_f = data[data['SMILES'].str.contains('Error:') == False]
```

```

train_data = data_f.sample(frac=0.7, random_state=42)
remaining_data = data_f.drop(train_data.index)
valid_data = remaining_data.sample(frac=0.5, random_state=42)
test_data = remaining_data.drop(valid_data.index)
print("Train samples:", len(train_data))
print("Validation samples:", len(valid_data))
print("Test samples:", len(test_data))

train_data['Mol'] = train_data['SMILES'].apply(lambda x: Chem.MolFromSmiles(x))
train_data['Mol_Hs'] = train_data['Mol'].apply(lambda x: Chem.AddHs(x) if x is not None else None)
train_data['MorganFingerprint'] = train_data['Mol_Hs'].apply(lambda x:
AllChem.GetMorganFingerprintAsBitVect(x, 2) if x is not None else None)
valid_data['MorganFingerprint'] = valid_data['SMILES'].apply(lambda x:
AllChem.GetMorganFingerprintAsBitVect(Chem.MolFromSmiles(x), 2))
test_data['MorganFingerprint'] = test_data['SMILES'].apply(lambda x:
AllChem.GetMorganFingerprintAsBitVect(Chem.MolFromSmiles(x), 2))

X_train = np.array(list(train_data['MorganFingerprint']), dtype=object)
X_valid = np.array(list(valid_data['MorganFingerprint']))
X_test = np.array(list(test_data['MorganFingerprint']))

X_train_filtered = [x for x in X_train if x is not None]
X_valid_filtered = [x for x in X_valid if x is not None]
X_test_filtered = [x for x in X_test if x is not None]

max_length = max(len(x) for x in X_train_filtered)

X_train_padded = [np.pad(x, (0, max_length - len(x)), mode='constant') for x in
X_train_filtered]
X_valid_padded = [np.pad(x, (0, max_length - len(x)), mode='constant') for x in
X_valid_filtered]
X_test_padded = [np.pad(x, (0, max_length - len(x)), mode='constant') for x in
X_test_filtered]

y_train = train_data['affinity'].values
y_valid = valid_data['affinity'].values
y_test = test_data['affinity'].values

X_train = np.array(X_train_padded, dtype=np.float32)
X_valid = np.array(X_valid_padded, dtype=np.float32)
X_test = np.array(X_test_padded, dtype=np.float32)

y_train = y_train.astype(np.float32)
y_valid = y_valid.astype(np.float32)
y_test = y_test.astype(np.float32)

X_train = X_train.reshape(X_train.shape[0], X_train.shape[1], 1)
X_valid = X_valid.reshape(X_valid.shape[0], X_valid.shape[1], 1)
X_test = X_test.reshape(X_test.shape[0], X_test.shape[1], 1)

scaler = MinMaxScaler()
X_train = scaler.fit_transform(X_train.reshape(X_train.shape[0], X_train.shape[1]))
X_valid = scaler.transform(X_valid.reshape(X_valid.shape[0], X_valid.shape[1]))

```

```

X_test = scaler.transform(X_test.reshape(X_test.shape[0], X_test.shape[1]))

X_train = X_train.reshape(X_train.shape[0], X_train.shape[1], 1)
X_valid = X_valid.reshape(X_valid.shape[0], X_valid.shape[1], 1)
X_test = X_test.reshape(X_test.shape[0], X_test.shape[1], 1)

early_stopping = EarlyStopping(monitor='val_loss', patience=3, restore_best_weights=True)

model = Sequential()
model.add(Conv1D(32, kernel_size=3, activation='relu', input_shape=(X_train.shape[1], 1)))
model.add(MaxPooling1D(pool_size=2))
model.add(Flatten())
model.add(Dropout(0.2))

model.add(Dense(64, activation='relu', kernel_regularizer=regularizers.l2(0.01)))
model.add(Dense(1, activation='linear', kernel_regularizer=regularizers.l2(0.01)))

model.compile(loss='mean_squared_error', optimizer='adam')

history = model.fit(X_train.reshape(X_train.shape[0], X_train.shape[1], 1), y_train,
                    validation_data=(X_valid.reshape(X_valid.shape[0], X_valid.shape[1], 1),
y_valid), epochs=100, batch_size=32, callbacks=[early_stopping])

mse = model.evaluate(X_test, y_test)
print("Mean Squared Error:", mse)

from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
y_pred = model.predict(X_test.reshape(X_test.shape[0], X_test.shape[1], 1))

mse = mean_squared_error(y_test, y_pred)
rmse = np.sqrt(mse)
mae = mean_absolute_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print("Mean Squared Error:", mse)
print("Root Mean Squared Error:", rmse)
print("Mean Absolute Error:", mae)
print("R-squared (R2) Score:", r2)

```

IX. Κώδικας Πρόβλεψης Ενεργειών Δέσμευσης των Μορίων

Ο κώδικας που χρησιμοποιήθηκε για την πρόβλεψη των ενεργειών δέσμευσης των μορίων προς την πρωτεΐνη TNFR1 (PDB ID: 1EXT).

```
import pandas as pd
from rdkit import Chem
from rdkit.Chem import PandasTools

import pandas as pd
new_ligands = pd.read_excel('top100-babel1_4.xlsx')

new_ligands['SMILES'] = ''

folder_path = 'C:/Users/Georgia/Desktop/Autoinflam/sdf/1-6000/'

for index, row in new_ligands.iterrows():
    ligand_name = row['file_name']
    sdf_file_path = folder_path + ligand_name + '.sdf'

    try:
        suppl = Chem.SDMolSupplier(sdf_file_path)
        molecule = next(suppl)

        if molecule:
            smiles = Chem.MolToSmiles(molecule)
            new_ligands.at[index, 'SMILES'] = smiles
        else:
            new_ligands.at[index, 'SMILES'] = 'Not found'

    except Exception as e:
        new_ligands.at[index, 'SMILES'] = 'Error: ' + str(e)
print(new_ligands)

new_ligands['Mol'] = new_ligands['SMILES'].apply(lambda x: Chem.MolFromSmiles(x))
new_ligands['Mol_Hs'] = new_ligands['Mol'].apply(lambda x: Chem.AddHs(x) if x is not None else None)
new_ligands['MorganFingerprint'] = new_ligands['Mol_Hs'].apply(lambda x:
AllChem.GetMorganFingerprintAsBitVect(x, 2) if x is not None else None)

X_new = np.array(list(new_ligands['MorganFingerprint']), dtype=object)
X_new_filtered = [x for x in X_new if x is not None]
max_length = max(len(x) for x in X_new_filtered)
X_new_padded = [np.pad(x, (0, max_length - len(x)), mode='constant') for x in X_new_filtered]
X_new_normalized = scaler.transform(np.vstack(X_new_padded))

predictions = model.predict(X_new_normalized.reshape(X_new_normalized.shape[0],
X_new_normalized.shape[1], 1))
new_ligands['Predicted Affinity'] = predictions
output_file_path = 'predicted_affinities.xlsx'
new_ligands.to_excel(output_file_path, index=False)
print(new_ligands[['file_name', 'SMILES', 'Predicted Affinity']])
```