



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΣΥΝΘΕΣΗΣ ΚΑΙ ΑΝΑΠΤΥΞΗΣ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΔΙΚΑΣΙΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

Αδροτοποιημένες Μοριακές Δυναμικές
προσομοιώσεις του διμερισμού
της ογκοπρωτεΐνης KRAS-4B

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
Ελευθερία Κελεφιώτη Στρατηδάκη

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ
Τόπακας Ευάγγελος

Αθήνα, Σεπτέμβριος 2021

Πίνακας Περιεχομένων

Ευχαριστίες.....	iii
Περίληψη	v
Abstract.....	vi
Ορολογία και συντομογραφίες.....	vii
1. Θεωρητικό μέρος	1
1.1. Εισαγωγή.....	1
1.1.1. Οικογένεια πρωτεϊνών RAS	1
1.1.2. Ο ρόλος της πρωτεΐνης KRAS-4B στο κύτταρο.....	8
1.1.3. Διεπιφάνειες διμερισμού της πρωτεΐνης KRAS-4B	11
1.2. Στόχος της εργασίας	14
2. Μέθοδοι	15
2.1. Μοριακή Μοντελοποίηση	15
2.2. Εμπειρική συνάρτηση δυναμικής ενέργειας - πεδίο δυνάμεων	15
2.3. Προσομοιώσεις Μοριακής Δυναμικής	20
2.3.1. Προετοιμασία συστήματος	23
2.3.2. Ελαχιστοποίηση ενέργειας.....	23
2.3.3. Θέρμανση συστήματος	25
2.3.4. Εξισορρόπηση συστήματος	25
2.3.5. Παραγωγή τροχιάς και ανάλυση προσομοίωσης.....	26
2.4. Αδροποιημένα μοντέλα Μοριακής Δυναμικής προσομοίωσης.....	27
2.4.1. Πεδίο δυνάμεων MARTINI 3	28
2.4.2. Ροή εργασίας martiniate	31
2.5. Ομαδοποίηση της τροχιάς.....	32
3. Αποτελέσματα και συζήτηση αποτελεσμάτων	34
3.1. Μετατροπή ατομιστικού μοντέλου διμερούς KRAS-4B σε αδροποιημένο μοντέλο. 34	
3.1.1. Δημιουργία αδροποιημένης δομής και αρχείου topology	34
3.1.2. Ενυδάτωση συστήματος και προσθήκη ιόντων	35
3.2. Προσομοιώσεις Μοριακής Δυναμικής στο σύστημα διμερούς KRAS-4B σε διάλυμα 35	
3.2.1. Ελαχιστοποίηση της ενέργειας στο σύστημα διμερούς KRAS-4B.....	36
3.2.2. Εξισορρόπηση του συστήματος.....	38
3.2.3. Παραγωγή τροχιάς και ανάλυση ιδιοτήτων.....	43
3.2.3.1. Υπολογισμός θερμοδυναμικών ιδιοτήτων	44
3.2.3.2. Ρίζα μέσης τετραγωνικής απόκλισης σε συνάρτηση με το χρόνο	46
3.2.3.3. Απόσταση κέντρων μαζών των δύο μονομερών KRAS-4B σε διάλυμα.....	47
3.3. Ομαδοποίηση των τροχιών για την ανάλυση των διαπιφανειών διμερισμού.....	49
4. Συμπεράσματα και συζήτηση.....	53
5. Μελλοντικές προοπτικές	55
Βιβλιογραφία.....	56

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Ίδρυμα Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών σε συνεργασία με την Ερευνήτρια Β' Δρ. Ζωή Κούρνια, κατά το ακαδημαϊκό έτος 2020-2021 και σε μεγάλο βαθμό εξ αποστάσεως λόγω των συνθηκών που επικρατούσαν κατά τη διάρκεια της πανδημίας του COVID-19.

Θα ήθελα πρωτίστως να ευχαριστήσω τον Αναπληρωτή Καθηγητή Δρ. Ευάγγελο Τόπακα για την ευκαιρία που μου έδωσε, προτείνοντάς μου να πραγματοποιήσω την εργασία στο Ίδρυμα Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών σε συνεργασία με την Δρ. Ζωή Κούρνια. Τον ευχαριστώ επίσης για τις πολύτιμες συμβουλές και την βοήθειά του καθόλη τη διάρκεια των σπουδών μου στη Σχολή Χημικών Μηχανικών.

Επίσης θα ήθελα να ευχαριστήσω τη Δρ. Ζωή Κούρνια για την εμπιστοσύνη που μου έδειξε και τη δυνατότητα που μου έδωσε να επεκτείνω τους ορίζοντές μου, γνωρίζοντας τον τομέα της υπολογιστικής χημείας, μέσα από την εκπαίδευση στο εργαστήριό της και την εκπόνηση της παρούσας διπλωματικής εργασίας. Ακόμα θα ήθελα να ευχαριστήσω τον Δρ. Camilo Velez-Vega και τον Δρ. John Manchester από την φαρμακευτική εταιρία Novartis Institutes of Biomedical Research (Boston), με τους οποίους συνεργαστήκαμε για το έργο στο οποίο εντάσσεται και η παρούσα διπλωματική εργασία.

Αισθάνομαι την ανάγκη να ευχαριστήσω τον Δρ. Ιωάννη Ανδρεαδέλη και τη Δρ. Σοφία Κυριακίδη για την υποστήριξη που μου παρείχαν στα προβλήματα και για τις συζητήσεις γύρω από τα ερωτήματα που προέκυπταν κατά την εξέλιξη της διπλωματικής μου εργασίας. Επιπλέον θα ήθελα να ευχαριστήσω τον Δρ. Μιχάλη Παπαδουράκη και τον υποψήφιο Διδάκτορα Αλέξη Χατζηγούλα για την βοήθειά τους σε τεχνικά ζητήματα κατά την εκπαίδευσή μου και την ενασχόλησή μου με την παρούσα εργασία. Ομοίως θα ήθελα να ευχαριστήσω τα υπόλοιπα μέλη του εργαστηρίου, την υποψήφια Διδάκτορα Ζωή Διαμαντή, την Ισιδώρα Διακογιαννάκη και τον Θεμιστοκλή Αϊβαλιώτη για τη στήριξη τους αυτούς τους μήνες. Δεν θα μπορούσα όμως να μην ευχαριστήσω την συμφοιτήτριά μου Μελίνα Κουστένη, για τις ατελείωτες ώρες συζητήσεων και κυρίως ψυχολογική υποστήριξή της κατά τη διάρκεια της εκπόνησης των διπλωματικών μας εργασιών.

Η ολοκλήρωση της παρούσας διπλωματικής, η οποία γράφτηκε υπό τους ήχους μεταμεσονύκτιων House mix, EDM DJ sets και Lofi beats, σηματοδοτεί το τέλος των προπτυχιακών μου σπουδών. Θα ήθελα λοιπόν να εκφράσω την ευγνωμοσύνη μου στα άτομα που ήταν μαζί μου τα τελευταία πέντε και πλέον χρόνια, στην Τζωρτζίνα, στη Μάγδα, στη Στελίνα και στη Μαρία, στα άτομα που γνώρισα μέσα από το Chemecon και φυσικά στην οικογένεια του iGEM Athens. Τέλος, δεν θα μπορούσα να έχω καταφέρει τίποτα απ'

όλα αυτά χωρίς την αμέριστη στήριξη, τη συμπαράσταση και την ενθάρρυνση της οικογένειάς μου, της μητέρας μου, Ειρήνης, του πατέρα μου, Δημήτρη και του αδερφού μου, Παναγιώτη, τους οποίους δεν θα μπορούσα να ευχαριστήσω αρκετά.

Περίληψη

Η πρωτεΐνη KRAS-4B αποτελεί μέλος της οικογένειας των πρωτεϊνών RAS, οι οποίες είναι μικρές ενδοκυτταρικές GTPασες που συμμετέχουν στη σηματοδότηση για την ανάπτυξη και αναπαραγωγή του κυττάρου. Η KRAS-4B βρίσκεται στην ενεργή της μορφή όταν είναι συνδεδεμένη με το νουκλεοτίδιο GTP και στη ανενεργή όταν το GTP υδρολύεται σε GDP. Η αντικατάσταση του αμινοξέος της γλυκίνης στη θέση 12 με ασπαρτικό οξύ (G12D) παρεμποδίζει την υδρόλυση του GTP σε GDP και επομένως η KRAS-4B παραμένει μονίμως στην ενεργή μορφή. Αυτό το φαινόμενο προκαλεί ανεξέλεγκτη ανάπτυξη στο κύτταρο, οδηγώντας έτσι στην ογκογένεση.

Η KRAS-4B ενεργοποιεί τα σηματοδοτικά μονοπάτια είτε ως διμερές είτε ως μονομερές. Παρόλο που το διμερές KRAS-4B μελετάται για περισσότερες από τέσσερις δεκαετίες, δεν είναι γνωστή η δομή του ούτε για την πρωτεΐνη φυσικού τύπου ούτε για τη μεταλλαγμένη πρωτεΐνη G12D. Η μελέτη του διμερισμού της KRAS-4B και η σύγκριση των δομών διμερούς του φυσικού τύπου σε σχέση με το μεταλλαγμένο G12D, μπορεί δυνητικά να συμβάλει στον σχεδιασμό φαρμάκων, τα οποία θα στοχεύουν επιλεκτικά στην αναστολή του διμερισμού της μεταλλαγμένης KRAS-4B.

Στόχος της παρούσας διπλωματικής εργασίας είναι η διερεύνηση του διμερισμού της KRAS-4B σε νερό με Μοριακές Δυναμικές προσομοιώσεις με αδροτοποιημένα μοντέλα. Πιο συγκεκριμένα, μοντελοποιείται η KRAS-4B φυσικού τύπου και η μεταλλαγμένη G12D. Για την μελέτη των συστημάτων αυτών παράγεται τροχιά 1 μ s, η οποία ομαδοποιείται σε κυρίαρχες διαμορφώσεις για να αναλυθούν οι διεπιφάνειες διμερισμού της KRAS-4B. Από την ανάλυση των ιδιοτήτων των δύο συστημάτων συμπεραίνεται ότι η σχετική θέση των μονομερών δεν είναι σταθερή και επομένως παρατηρούνται αλλαγές των διεπιφανειών διμερισμού. Οι νέες διεπιφάνειες διμερισμού της KRAS-4B που ταυτοποιούνται συμφωνούν με τις διεπιφάνειες που περιγράφονται στην βιβλιογραφία.

Abstract

KRAS-4B protein is a member of RAS family proteins, which are small GTPases involved in intracellular signalling that leads to cell growth and proliferation. KRAS-4B is in its active form when bound to the GTP nucleotide and inactive when GTP is hydrolyzed to GDP. Replacement of the amino acid glycine at position 12 with an aspartic acid (G12D) inhibits the hydrolysis of GTP to GDP and therefore KRAS-4B remains permanently active. This phenomenon causes uncontrollable cellular growth, leading to oncogenesis.

KRAS-4B activates signalling pathways either in its dimer or monomer form. However, although the KRAS-4B dimer has been studied for more than four decades, its dimer structure is not known for the wild type (WT) or the mutant G12D protein. Studying the dimerization of KRAS-4B and comparing the dimer structures of the WT and the mutant G12D proteins could potentially contribute to the design of drugs that selectively inhibit the dimerization of the mutant KRAS-4B.

The aim of this diploma thesis is to investigate the dimerization of KRAS-4B in water using coarse-grained Molecular Dynamics simulations. More specifically, the WT as well as the G12D mutant systems are modelled. To study these two systems, a 1 μ s trajectory is produced, which is then clustered into representative structures that are analyzed to identify KRAS-4B dimerization interfaces. From the analysis of the properties in both systems it is concluded that the relative position of the two monomers is not constant throughout the trajectory and therefore multiple dimerization interfaces are observed. The dimerization interfaces that are identified are consistent with the interfaces described in the literature.

Ορολογία και συντομογραφίες

A: Alanine, αλανίνη

Allosteric lobe: αλλοστερικό λοβό

C: Cysteine, κυστίνη

CADD: Computer-aided drug design, σχεδιασμός φαρμάκων με τη βοήθεια υπολογιστή

Canonical ensemble: Κανονικό σύνολο

COM distance: Center Of Mass distance, απόσταση των κέντρων μάζας

D: Aspartic Acid, ασπαρτικό οξύ

DOPC: 1,2-Dioleoyl-sn-glycero-3-phosphocholine

DOPS: 1,2-Dioleoyl-sn-glycero-3-phospho-L-serine

Effector: σηματοδοτική πρωτεΐνη

Effector lobe: λοβό τελεστή

EGF: Epidermal Growth Factor, επιδερμικός αυξητικός παράγοντας

EGFR: EGF Receptor, υποδοχέας επιδερμικού αυξητικού παράγοντα

Electron microscopy: ηλεκτρονική μικροσκοπία

Ensemble: στατιστικό σύνολο

FDA: Food & Drug Administration, Αμερικάνικος Οργανισμός Τροφίμων και Φαρμάκων

G: Glycine, ηλικίνη

G domain: καταλυτική περιοχή

G12D: μετάλλαξη της γλυκίνης 12 σε ασπαρτικό οξύ

GAP: GTPase-activating protein, πρωτεΐνη ενεργοποίησης της GTPασης

GDP: Guanosine diphosphate, διφωσφορική γουανοσίνη

GEF: guanine nucleotide exchange factor, παράγοντας ανταλλαγής νουκλεοτιδίων γουανίνης

GTP: Guanosine triphosphate, τριφωσφορική γουανοσίνη

HVR: hypervariable C-terminal region, υπερμεταβλητή καρβοξυτελική περιοχή

INSANE: INSert membraNE, εισαγωγή μεμβράνης

Isothermal–isobaric ensemble: Ισόθερμο – ισοβαρές σύνολο

LB-CADD: Ligand-Based CADD, σχεδιασμός φαρμάκων με τη βοήθεια υπολογιστή με τη βοήθεια υποκαταστάτη

MetaD: MetaDynamics, μεταδυναμική

Microcanonical ensemble: Μικρο-κανονικό σύνολο

NMR: Nuclear Magnetic Resonance, πυρηνικός μαγνητικός συντονισμός

NSCLC: Non-small-cell lung carcinoma, μη μικροκυτταρικός καρκίνος του πνεύμονα

PDB: Protein Data Bank, βάση δεδομένων πρωτεϊνών

Periodic boundary conditions: περιοδικές οριακές συνθήκες
PI3K: Phosphoinositide 3-kinase, 3-κινάση της φωσφοϊνοσιτίδη
PIP2: Phosphatidylinositol 4,5-bisphosphate
PR: Position Restrained, συγκρατείται στη θέση
PRE NMR: paramagnetic relaxation enhancement, NMR παραμαγνητικής ενίσχυσης χαλάρωσης
PT: Parallel Tempering, παράλληλη θερμοκρασία
Q: Glutamine, γλουταμίνη
Raf: Rapidly Accelerated Fibrosarcoma, ταχέως επιταχυνόμενο ινοσάρκωμα
RAS: RAt Sarcoma virus, ιός σαρκώματος αρουραίου
RMS: Root Mean Square, μέσο τετραγώνου ρίζας
RMSD: Root Mean Square Deviation, ρίζα της μέσης τετραγωνικής απόκλισης
RTK: Receptor Tyrosine Kinase, κινάση τυροσίνης υποδοχέα
V: Valine, βαλίνη
WT: Wild Type, φυσικός τύπος
WTE: Well-Tempered Ensemble, κανονικό σύνολο
X-ray crystallography: κρυσταλλογραφία ακτίνων Χ

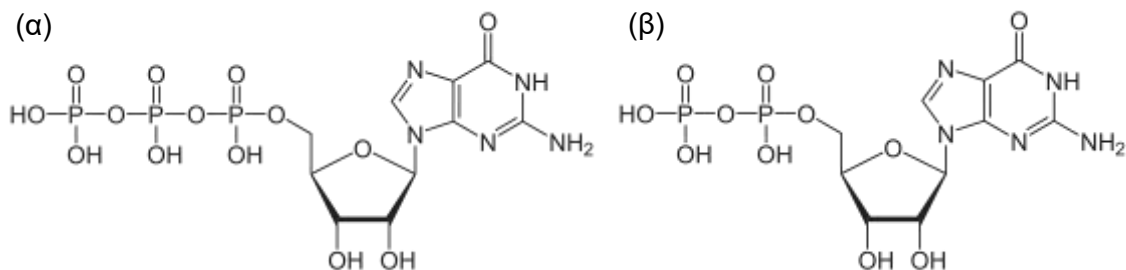
1. Θεωρητικό μέρος

1.1. Εισαγωγή

1.1.1. Οικογένεια πρωτεϊνών RAS

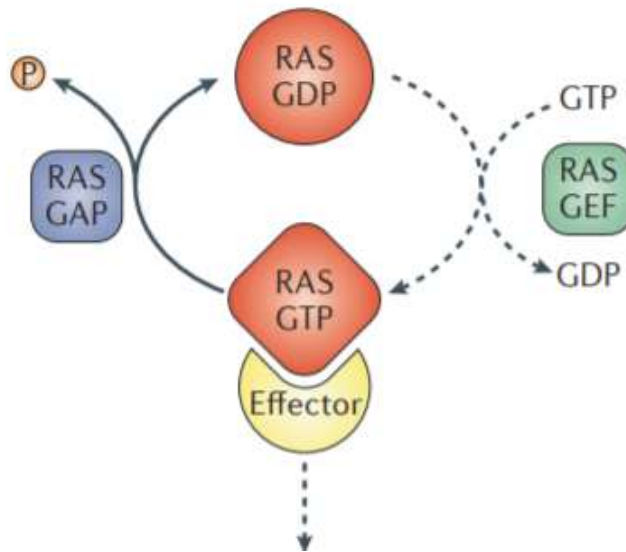
Η οικογένεια πρωτεϊνών RAS (RAt Sarcoma virus) ανήκει στην κλάση των πρωτεϊνών που ονομάζονται μικρές GTPασες (EC 3.6.5.2) και αποτελείται από τέσσερα μέλη: την KRAS-4A, την KRAS-4B, την HRAS και την NRAS. Οι πρωτεΐνες RAS χαρακτηρίζονται ως ογκοπρωτεΐνες και σχετίζονται με την εμφάνιση πολλών καρκίνων. Πιο συγκεκριμένα οι ογκογόνες μεταλλάξεις των πρωτεϊνών RAS εντοπίζονται σε περίπου 30% των καρκίνων του ανθρώπου, ενώ μεταλλάξεις βρίσκονται συχνά σε καρκίνους του πνεύμονα, του παχέος εντέρου και του παγκρέατος.

Ως προς τη λειτουργία τους, οι μικρές GTPασες εναλλάσσονται μεταξύ της ενεργής τους μορφής, όπου η πρωτεΐνη βρίσκεται δεσμευμένη με το μόριο της τριφωσφορικής γουανοσίνης (GTP) και της ανενεργής μορφής, δεσμευμένης με το μόριο της διφωσφορικής γουανοσίνης (GDP). Τα δύο αυτά μόρια διαφέρουν κατά μία φωσφορική ομάδα και απεικονίζονται στο Σχήμα 1.1 (α) και (β) αντίστοιχα. Η διεργασία της μετατροπή μιας ενεργής GTPασης σε ανενεργή γίνεται μέσω της υδρόλυσης του GTP σε GDP [1] [2].



Σχήμα 1.1: (α) Το μόριο της τριφωσφορικής γουανοσίνης (GTP) και (β) το μόριο της διφωσφορικής γουανοσίνης (GDP) [3].

Η εγγενής υδρόλυση του GTP από τις πρωτεΐνες RAS είναι αργή, αλλά επιταχύνεται σημαντικά από τις πρωτεΐνες ενεργοποίησης της GTPase (GTPase-activating proteins ή GAP), οι οποίες χρησιμεύουν ως αρνητικοί ρυθμιστές της δραστηριότητας των RAS επιταχύνοντας την υδρόλυση του GTP, οδηγώντας στην απενεργοποίηση της σηματοδότησης RAS. Η ενεργοποίηση των RAS ρυθμίζεται από παράγοντες ανταλλαγής νουκλεοτιδίων γουανίνης (guanine nucleotide exchange factors ή GEF), οι οποίοι λειτουργούν ως θετικοί ρυθμιστές της δραστηριότητας των RAS, μέσω της καταλύσεως της απελευθέρωσης του GDP και της δέσμευσης του πιο άφθονου GTP [4]. Η σχηματική αναπαράσταση της υδρόλυσης των πρωτεϊνών RAS απεικονίζεται στο Σχήμα 1.2 [5].

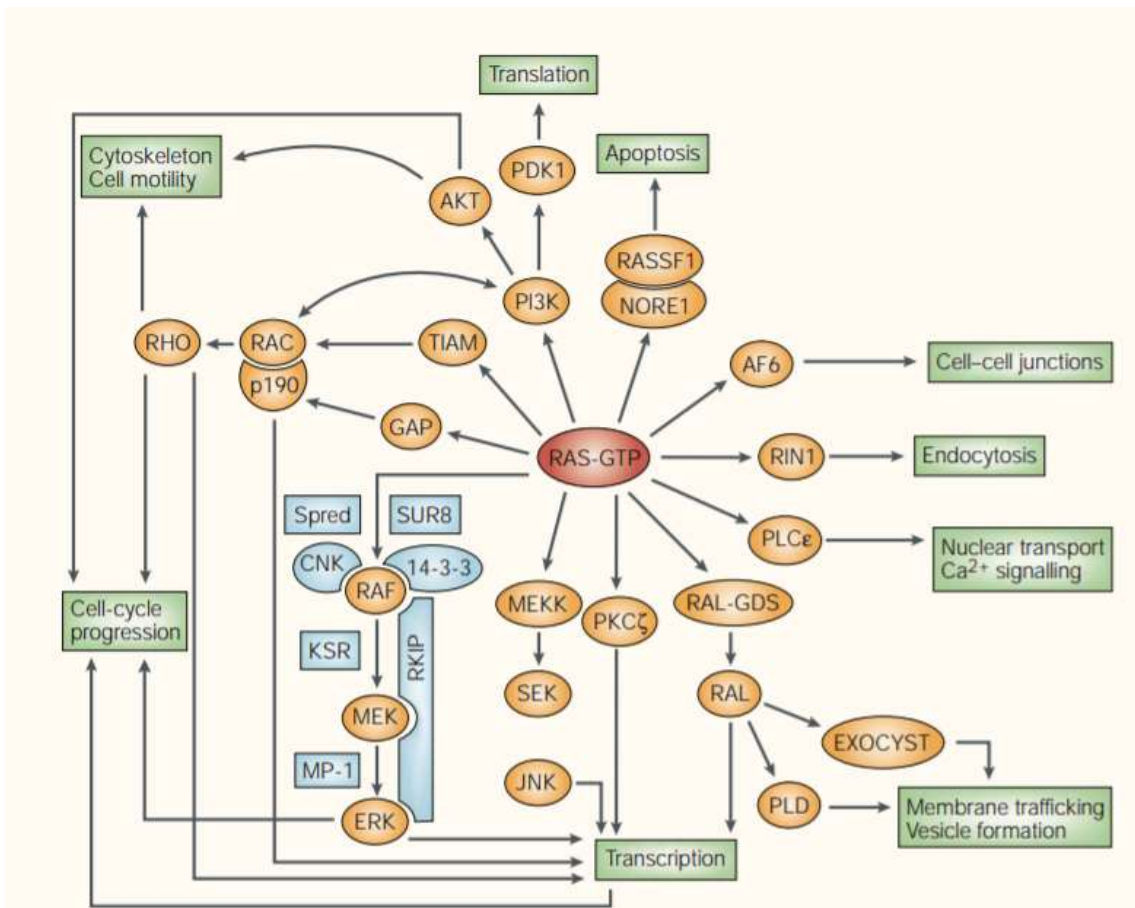


Σχήμα 1.2: Οι ενεργοποιημένες πρωτεΐνες GEF προωθούν τον παροδικό σχηματισμό RAS-GTP, ενώ οι πρωτεΐνες GAP επιστρέφουν την πρωτεΐνη στην ανενεργή δεσμευμένη κατάσταση του GDP [5].

Τα γονίδια που κωδικοποιούν τις πρωτεΐνες RAS στον άνθρωπο ανακαλύφθηκαν μόλις το 1981 [6], ενώ το 1984 ανακαλύφθηκε ότι οι πρωτεΐνες RAS είναι διαμεσολαβητές μεταφοράς σήματος μέσω της μεμβράνης πλάσματος.

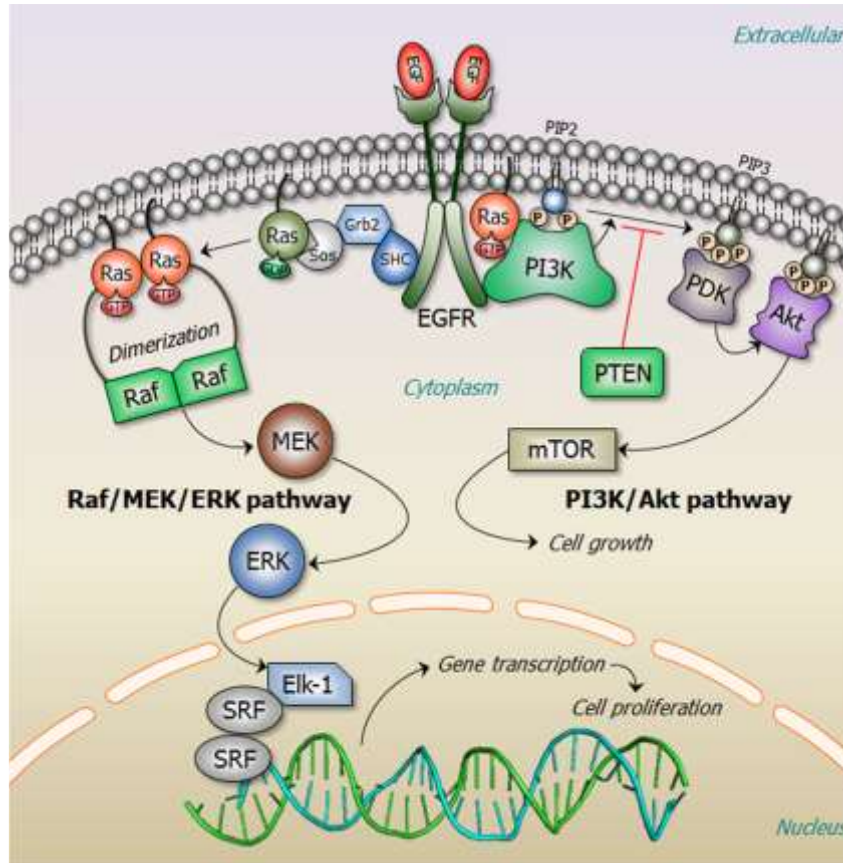
Ένας από τους ρόλους των πρωτεϊνών RAS αποτελεί μέρος του μεταβολικού μονοπατιού RAS-Raf-MEK-ERK, το οποίο περιγράφεται αναλυτικά παρακάτω, καθιστώντας τις πρωτεΐνες RAS υπεύθυνες για τη ρύθμιση της κυτταρικής ανάπτυξης, του πολλαπλασιασμού της διαφοροποίησης και της απόπτωσης. Οι βασικοί μηχανισμοί της σηματοδότησης RAS έχουν ερευνηθεί εκτενώς τις τελευταίες τέσσερις δεκαετίες. Ως αποτέλεσμα, υπάρχει πλέον μια λεπτομερής βιοχημική κατανόηση του τρόπου με τον οποίο η GTPαση RAS συνδέεται με GTP και πώς, σε αυτήν την ενεργοποιημένη κατάσταση, η RAS συνδέεται και ενεργοποιεί διάφορες σηματοδοτικές πρωτεΐνες [1].

Φυσιολογικά, οι πρωτεΐνες παράγονται στο κυτταρόπλασμα και βρίσκονται στην εσωτερική πλευρά της πλασματικής μεμβράνης, όπου έχουν την δυνατότητα να συμβάλλουν στη διάδοση σημάτων μέσω της σύνδεσης τους με άλλα σηματοδοτικά μόρια, όπως οι PI3K και Raf (Rapidly Accelerated Fibrosarcoma) [7] [8]. Στα μέσα της δεκαετίας του 1990 ο κατάλογος των σηματοδοτικών μορίων των πρωτεϊνών RAS άρχισε να αυξάνεται ραγδαία. Ενδεικτικά, στο Σχήμα 1.3 παρουσιάζονται συνοπτικά τα γνωστά σηματοδοτικά μόρια των πρωτεϊνών RAS και οι αντίστοιχες βιολογικές αποκρίσεις τους [8].



Σχήμα 1.3: Επισκόπηση των γνωστών σηματοδοτικών μορίων RAS και των αντίστοιχων βιολογικών τους αποκρίσεων. Το ενεργό RAS-GTP προκαλεί μια μεγάλη ποικιλία κυτταρικών διεργασιών, όπως μεταγραφή, μετάφραση, πρόοδο κυτταρικού κύκλου, απόπτωση ή κυτταρική επιβίωση, μέσω άμεσης αλληλεπίδρασης με διάφορες σηματοδοτικές πρωτεΐνες [8].

Πιο αναλυτικά ο μηχανισμός με τον οποίο επιτελείται η λειτουργία ενδιαφέροντος για την παρούσα εργασία της πρωτεΐνης RAS, παρουσιάζεται συνοπτικά στο Σχήμα 1.4 και περιγράφεται παρακάτω. Ο επιδερμικός αυξητικός παράγοντας (Epidermal Growth Factor ή EGF) που συνδέεται με την εξωκυτταρική περιοχή του υποδοχέα επιδερμικού αυξητικού παράγοντα (EGF Receptor ή EGFR), που είναι μία κινάση τυροσίνης υποδοχέα (RTK), έχει ως αποτέλεσμα τον διμερισμό και την ενεργοποίηση του EGFR. Ο ενεργοποιημένος EGFR προσελκύει πρωτεΐνες GEF, όπως η SOS μέσω των πρωτεϊνών προσαρμογής SHC και Grb2. Η SOS ανταλλάσσει το GDP με GTP και ενεργοποιεί την RAS. Το ενεργοποιημένο πλέον σύμπλεγμα RAS-GTP διμερίζεται και στη συνέχεια συνδέεται με τη σηματοδοτική Raf κινάση, οδηγώντας έτσι με τη σειρά του σε διμερισμό της σηματοδοτικής πρωτεΐνης Raf και ενεργοποίησή της [2] [9] [10]. Στο Σχήμα 1.5 παρουσιάζεται η πρωτεΐνη KRAS-4B με τρικουάζ, σε σύμπλοκο με ένα μόριο GTP, συνδεδεμένη με το μόριο Raf με πορτοκαλί και οι μεταξύ τους τις αλληλεπιδράσεις με πράσινο.



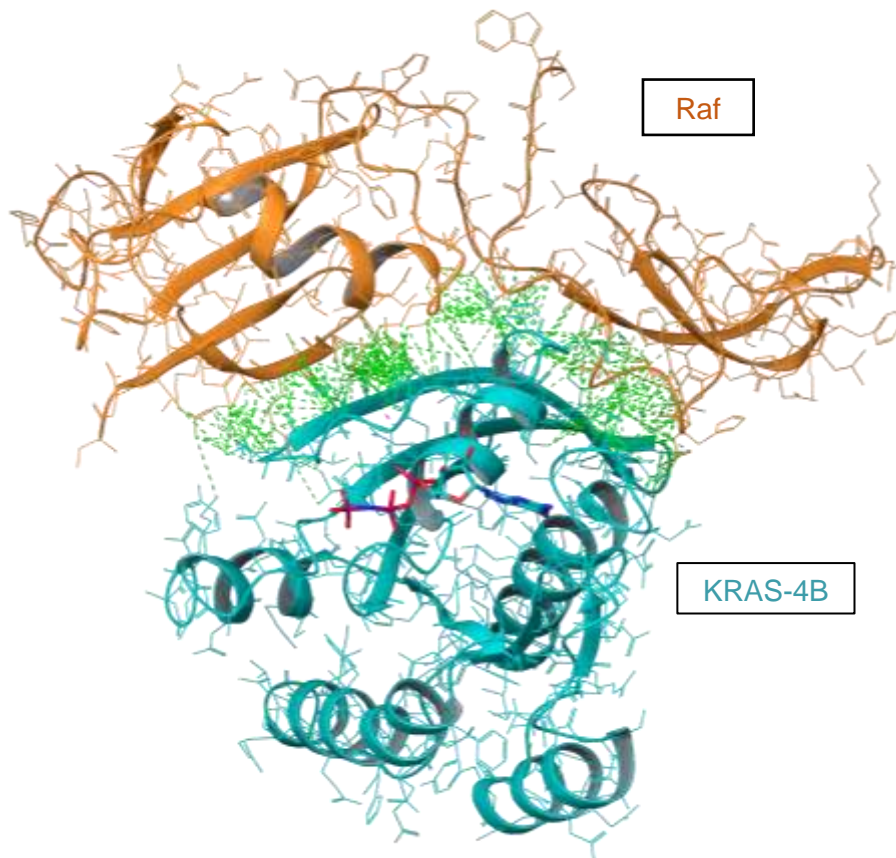
Σχήμα 1.4: Η σηματοδότηση του RAS στο μονοπάτι Raf-MEK-ERK [2].

Στο μονοπάτι RAS–Raf–MEK–ERK, το ενεργό διμερές Raf φωσφορυλιώνει και ενεργοποιεί την κινάση MEK1/2, η οποία με τη σειρά της φωσφορυλιώνει και ενεργοποιεί το ERK1/2. Το ενεργό ERK1/2 φωσφορυλιώνει και ενεργοποιεί το ELK-1, έναν μεταγραφικό παράγοντα. Το ELK-1 συνδέεται με τον συμπαράγοντα του, το διμερές SRF (Serum Response Factor), οδηγώντας στην ενεργοποίηση της μεταγραφής και στον πολλαπλασιασμό κυττάρων [2] [9] [10]. Με παρόμοιο τρόπο ενεργοποιείται και το μονοπάτι RAS–PI3K–Akt, το οποίο όμως δεν θα περιγραφεί, καθώς ξεφεύγει από το αντικείμενο της παρούσας εργασίας.

Οι μεταλλάξεις που συμβαίνουν στις πρωτεΐνες RAS βλάπτουν τη δραστηριότητα της ως GTPαση και ως εκ τούτου καθιστούν τις μεταλλαγμένες RAS μη ευαίσθητες στις πρωτεΐνες GAP. Επομένως οι μεταλλαγμένες πρωτεΐνες RAS βρίσκονται μόνιμως δεσμευμένες με GTP και επομένως ενεργές, επιμηκύνοντας έτσι την κατάντη σηματοδότηση που σχετίζεται με την ογκογόνο κυτταρική ανάπτυξη [2].

Ο εντοπισμός της μετάλλαξης που οδηγεί σε καρκινογένεση ήταν μία ακόμα σημαντική ανακάλυψη που έλαβε χώρα τη δεκαετία του 1980. Πρόκειται για μια μετάλλαξη σημείου, δηλαδή η αντικατάσταση ενός νουκλεοτιδίου στο γονίδιο που είναι υπεύθυνο για την

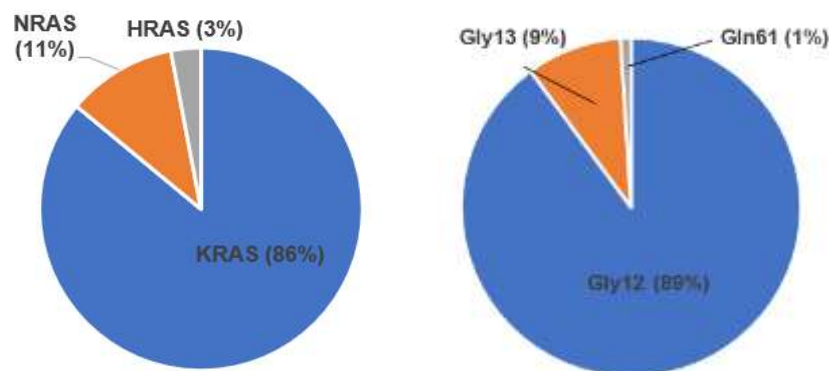
κωδικοποίηση της πρωτεΐνης και επομένως η αλλαγή του αμινοξέος που μεταφράζεται. Πιο συγκεκριμένα, μια μετάλλαξη σημείου στο κωδικόνιο 12 τόσο στα γονίδια που κωδικοποιούν τις πρωτεΐνες RAS είναι υπεύθυνη για την υποκατάσταση των αμινοξέων που βρίσκονται φυσιολογικά σε αυτή την θέση με άλλα τα οποία διαταράσσουν την δομή και επομένως την λειτουργία των πρωτεϊνών αυτών. Οι μεταλλάξεις που λαμβάνουν χώρα στο κωδικόνιο 12, μαζί με αυτές που εντοπίζονται στα κωδικόνια 13 και 61, είναι υπεύθυνες για το 97-99% των καρκίνων που συνδέονται με μεταλλάξεις των πρωτεϊνών RAS. Η ανακάλυψη αυτή ήταν σε μεγάλο βαθμό υπεύθυνη για την εντατικοποίηση της αναζήτησης της προέλευσης του καρκίνου σε μοριακό επίπεδο που ακολούθησε τα επόμενα χρόνια, αλλά και των σημαντικών κατακτήσεων που επιτεύχθηκαν στη συνέχεια για την καταπολέμησή του, όπως είναι η πλήρης αλληλούχιση του ανθρώπινου γονιδιώματος το 2001 [11].



Σχήμα 1.5: Σύμπλοκο της KRAS-4B με τη σηματοδοτική πρωτεΐνη Raf.

Ωστόσο, ανάμεσα στα μέλη της οικογένειας πρωτεϊνών RAS, δεν παρατηρούνται μεταλλάξεις με την ίδια συχνότητα. Μεταξύ των ισομορφών RAS, εντοπίζονται συχνότερα μεταλλάξεις στο γονίδιο που κωδικοποιεί την πρωτεΐνη KRAS (86%), λιγότερο συχνά σε αυτό που κωδικοποιεί την πρωτεΐνη HRAS (3%) και σπάνια στο γονίδιο που είναι υπεύθυνο για την κωδικοποίηση της πρωτεΐνης NRAS (11%), όπως φαίνεται και στο Σχήμα 1.6 (α) [2] [12]. Είναι επίσης αξιοσημείωτο ότι η συχνότητα μετάλλαξης σε κάθε θέση διαφέρει ευρέως

μεταξύ των ισομορφών, με τις μεταλλάξεις G12 να είναι οι πιο συχνές στο γονίδιο για την KRAS και το HRAS, ενώ οι μεταλλάξεις Q61 είναι οι πιο συχνές στο γονίδιο για την NRAS [12]. Ακόμα, στο Σχήμα 1.6 (β) φαίνεται ότι η συχνότερη καρκινογόνος μετάλλαξη στην KRAS-4B λαμβάνει χώρα στη γλυκίνη της θέσης 12 και είναι μάλιστα η αντικατάσταση της από το ασπαρτικό οξύ. Η μετάλλαξη αυτή συμβολίζεται G12D. Ως αποτέλεσμα, οι μεταλλάξεις της KRAS εμφανίζονται σε πολύ υψηλή συχνότητα στο αδενοκαρκίνωμα του παγκρέατος και στους όγκους των πνευμόνων και του παχέος εντέρου, οι μεταλλάξεις της NRAS ανιχνεύονται συνήθως σε αιματοποιητικούς όγκους και σε κακοήγη μελανώματα, ενώ οι μεταλλάξεις της HRAS είναι οι πιο συχνές σε όγκους της ουροδόχου κύστης και καρκίνωμα πλακωδών κυττάρων κεφαλής και λαιμού [2].

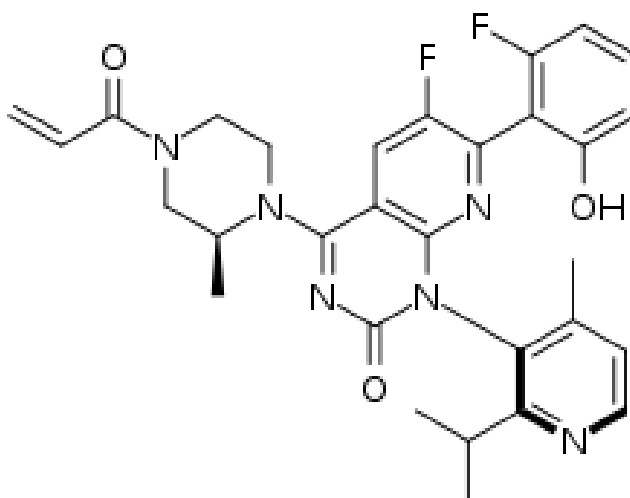


Σχήμα 1.6: (α) Η KRAS μεταλλάσσεται συχνότερα απ' όλες τις πρωτεΐνες RAS (β) στην KRAS-4B η μετάλλαξη της γλυκίνης στη θέση 12 είναι η συχνότερη ογκογενική μετάλλαξη [13].

Προκειμένου να αντιμετωπιστούν οι μεταλλάξεις αυτές που οδηγούν σε καρκίνο, γίνεται προσπάθεια από την ερευνητική κοινότητα και τη φαρμακοβιομηχανία για την ανακάλυψη μικρών μορίων - φαρμάκων που θα λειτουργήσουν σαν αναστολείς, με στόχο την απενεργοποίηση μόνο των μεταλλαγμένων μορφών των πρωτεϊνών αυτών. Ο σχεδιασμός φαρμάκων με τη βοήθεια υπολογιστών (Computer-Aided Drug Discovery ή CADD) χρησιμοποιεί υπολογιστικές προσεγγίσεις για την ανακάλυψη, ανάπτυξη και ανάλυση φαρμάκων και άλλων παρόμοιων βιολογικά ενεργών μορίων. Η ανακάλυψη φαρμάκων με τη βοήθεια υποκαταστάτη (Ligand-Based-CADD ή LB-CADD) περιλαμβάνει την ανάλυση υποκαταστατών που είναι γνωστό ότι αλληλεπιδρούν με έναν στόχο ενδιαφέροντος. Αυτές οι μέθοδοι χρησιμοποιούν ένα σύνολο δομών αναφοράς που συλλέγονται από ενώσεις, οι οποίες είναι γνωστό ότι αλληλεπιδρούν με τον στόχο που μας ενδιαφέρει και αναλύουν τις δισδιάστατες ή τριδιάστατες δομές τους. Ο βασικός στόχος αυτών των μεθόδων είναι η πρόβλεψη του είδους και της ισχύος της σύνδεσης ενός γνωστού υποκαταστάτη με τον στόχο ενδιαφέροντος.

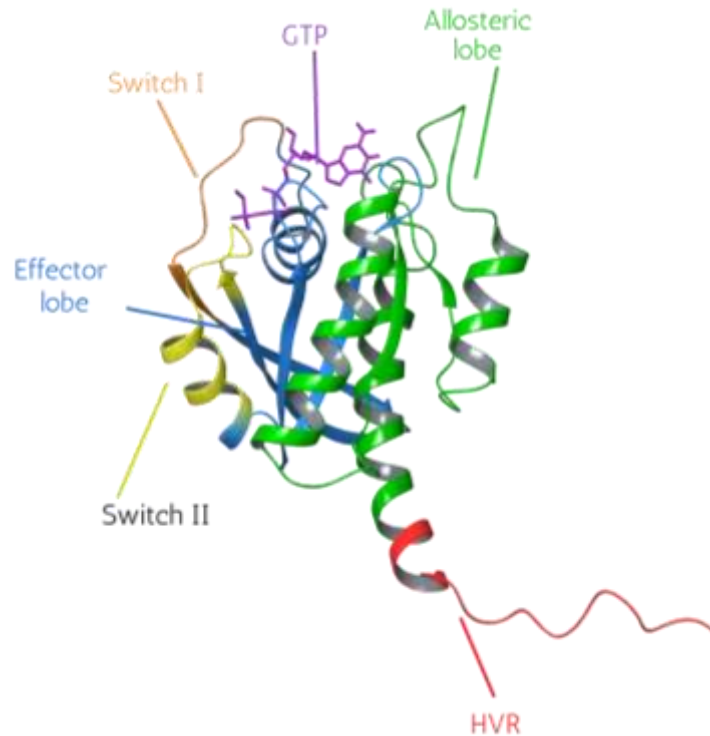
Στοιχεία από τομείς όπως η κβαντική χημεία αλλά και η θεωρία συναρτησιακής πυκνότητας συνδυάζονται και χρησιμοποιούνται συχνά για την παροχή βελτιστοποιημένων παραμέτρων για τους υπολογισμούς μοριακής μηχανικής, που έχουν στόχο την πρόβλεψη της δομής των μικρών μορίων – υποκαταστατών και για τη μοντελοποίηση των διαμορφωτικών αλλαγών που ενδέχεται να προκύψουν στο βιολογικό στόχο μετά τη σύνδεσή του με το μικρό μόριο που εξετάζεται. Οι μέθοδοι CADD μπορούν να αυξήσουν τις πιθανότητες αναγνώρισης ενώσεων με επιθυμητά χαρακτηριστικά, να αυξήσουν την ανάπτυξη της επιτυχίας και να αυξήσουν τις πιθανότητες απόκτησης μιας ένωσης πέρα από τα πολλά εμπόδια των προκλινικών δοκιμών [14].

Ωστόσο, στην περίπτωση των πρωτεϊνών RAS, η διεργασία αυτή έχει αποδειχθεί εξαιρετικά δύσκολη, όπως αποδεικνύεται από το γεγονός ότι παρόλο που το φαινόμενο των μεταλλάξεων στις πρωτεΐνες RAS και ο συσχετισμός τους με τις καρκινογενέσεις έχει παρατηρηθεί εδώ αρκετές δεκαετίες, το εγκεκριμένο μοναδικό φαρμακευτικό σκεύασμα για την αντιμετώπισή τους, είναι προϊόν πρόσφατων μελετών και πήρε την επίσημη έγκριση από τον Αμερικανικό Οργανισμό Τροφίμων και Φαρμάκων (Food & Drug Administration ή FDA) μόλις στις 28 Μαΐου του 2021 [15] [16]. Πρόκειται για το Sotorasib (Lumakras™, Amgen, Inc.), το οποίο απεικονίζεται στο Σχήμα 1.7, έναν αναστολέα της οικογένειας πρωτεϊνών RAS, για ενήλικες ασθενείς με την μετάλλαξη KRAS G12C σε τοπικά προχωρημένο ή μεταστατικό μη μικροκυτταρικό καρκίνο του πνεύμονα (Non-small-cell lung carcinoma ή NSCLC) [17]. Καθώς οι μεταλλάξεις της πρωτεΐνης KRAS θεωρούνταν για καιρό ανθεκτικές στη φαρμακευτική θεραπεία, η ελπιδοφόρα αυτή εξέλιξη αντιπροσωπεύει ένα σημαντικό βήμα προς ένα μέλλον, στο οποίο περισσότεροι ασθενείς θα έχουν εξατομικευμένη θεραπευτική προσέγγιση, σε ασθένειες όπως ο καρκίνος του πνεύμονα που είναι ο πιο κοινός τύπος καρκίνου και έχει τη μεγαλύτερη θνησιμότητα [18].



Σχήμα 1.7: Η χημική δομή του Sotorasib, του πρώτου εγκεκριμένου φαρμάκου κατά της KRAS G12C [19].

1.1.2. Ο ρόλος της πρωτεΐνης KRAS-4B στο κύτταρο

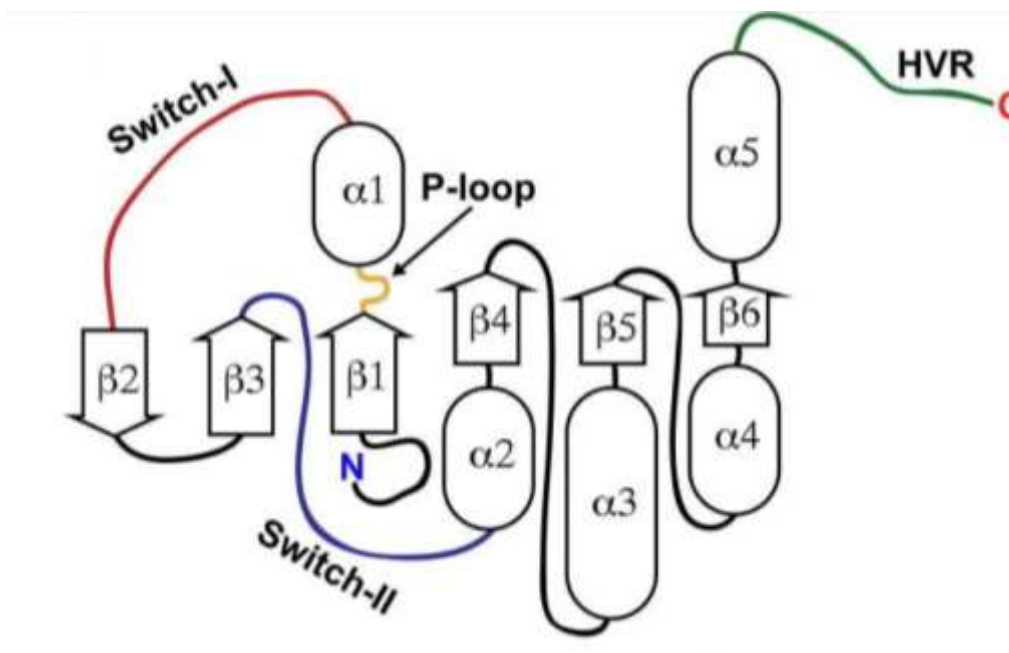


Σχήμα 1.8: Μονομερές της KRAS-4B (PDB ID: 6W4E), στο οποίο διακρίνονται οι διαφορετικές περιοχές της.

Σε όλα τα βιολογικά συστήματα η λειτουργία μιας πρωτεΐνης συνδέεται άρρηκτα με την δομή της. Υπό αυτό το πρίσμα, κρίνεται σημαντικό να περιγραφεί με λεπτομέρεια η δομή της πρωτεΐνης KRAS-4B εντός του κυττάρου. Η KRAS-4B αποτελείται από 188 αμινοξέα και χωρίζεται στην καταλυτική περιοχή (G domain), η οποία εκτείνεται από το αμινοξύ 1 έως το 166, και στην υπερμεταβλητή καρβοξυτελική περιοχή (hypervariable C-terminal region ή HVR), η οποία εκτείνεται από το αμινοξύ 167 έως το 188. Η καταλυτική περιοχή, αποτελείται από δύο δομικά διακριτές περιοχές: τον λοβό τελεστή (effector lobe, αμινοξέα 1 έως 86), ο οποίος προσλαμβάνει παράγοντες σηματοδότησης κατάντη και τον αλλοστερικό λοβό (allosteric lobe, αμινοξέα 87 έως 167), ο οποίος εμπλέκεται στις αλληλεπιδράσεις μεμβράνης [20]. Η καταλυτική περιοχή περιλαμβάνει επίσης τον βρόχο δέσμευσης φωσφορικών (P-loop, αμινοξέα 10 έως 17) και τις περιοχές Switch I (SI, αμινοξέα 25 έως 40) και Switch II (SII, αμινοξέα 60 έως 74), οι οποίες συνδέονται με το μόριο GTP και υφίστανται σημαντικές διαμορφωτικές αλλαγές κατά την υδρόλυση του GTP προς GDP, διαδικασία η οποία απενεργοποιεί την KRAS-4B. Η υπερμεταβλητή περιοχή του τερματικού C, HVR, είναι εξαιρετικά ευέλικτη και διαφέρει μεταξύ των μελών της οικογένειας RAS. Αυτές οι περιοχές

διακρίνονται χρωματισμένες στο Σχήμα 1.8, στο οποίο απεικονίζεται το ένα μονομερές από τη δομή με PDB ID 6W4E της Protein Data Bank [21].

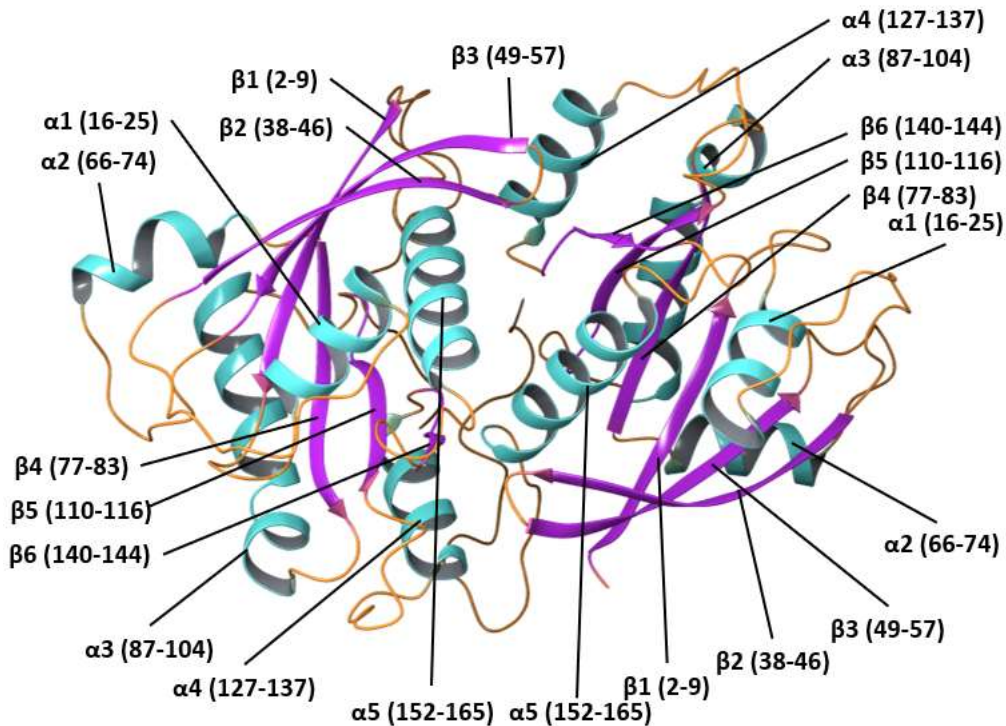
Όταν η πρωτεΐνη KRAS-4B φυσικού τύπου (WT KRAS-4B) συνδέεται με το GDP, το HVR αλληλεπιδρά στενά με τον καταλυτικό τομέα, περιορίζοντας την KRAS-4B σε εγγενώς αυτοανασταλμένη κατάσταση [22]. Αντίθετα, όταν η πρωτεΐνη KRAS-4B φυσικού τύπου συνδέεται με το GTP, παρατηρούνται χαλαρότερες αλληλεπιδράσεις HVR – καταλυτικής περιοχής, γεγονός που οδηγεί στην απελευθέρωση του HVR και την επακόλουθη συσχέτισή του με την κυτταρική μεμβράνη [23]. Σε όλη την έκταση της πρωτεΐνης παρατηρούνται πέντε α-έλικες και έξι β-πτυχωτά φύλλα, όπως φαίνεται στο Σχήμα 1.9 [2] [24].



Σχήμα 1.9: Δευτεροταγής δομή της πρωτεΐνης KRAS-4B, στην οποία οι α-έλικες απεικονίζονται ως κύλινδροι και τα β-πτυχωτά φύλλα ως βέλη [24].

Στο σχήμα 1.10 παρουσιάζεται επίσης το διμερές της πρωτεΐνης KRAS-4B, βασισμένο στην κρυσταλλική δομή με PDB ID 5VQ2 [25] της Protein Data Bank, στα οποία εμφανίζονται και οι δευτεροταγείς δομές στις τρεις διαστάσεις με τα αντίστοιχα αμινοξέα στα οποία εκτείνονται.

Η KRAS-4B είναι η πιο συχνά μεταλλαγμένη ογκοπρωτεΐνη που συναντάται σε πολλαπλούς ανθρώπινους καρκίνους [26]. Οι ασθενείς με ογκογονικές μεταλλάξεις KRAS-4B έχουν πολύ κακή ανταπόκριση στις τυπικές θεραπείες [27]. Η σημασία των ογκογόνων μεταλλάξεων στην KRAS-4B πηγάζει από τον κεντρικό ρόλο του στα δίκτυα σηματοδότησης που ελέγχουν την κυτταρική ανάπτυξη, πολλαπλασιασμό και διαφοροποίηση, όπως περιγράφηκε παραπάνω.



Σχήμα 1.10: Διμερές της KRAS-4B βασισμένο στη δομή με PDB ID 5VQ2 [25], στο οποίο φαίνονται οι α -έλικες με γαλάζιο χρώμα και τα β -πτυχωτά φύλλα με φούξια.

Ογκογόνες μεταλλάξεις στη KRAS-4B παρατηρούνται συχνά στις θέσεις αμινοξέων 12, 13 και 61 σε καρκινοπαθείς. Μεταξύ αυτών, η γλυκίνη 12 (G12) είναι το πιο συχνά μεταλλαγμένο αμινοξύ (89%) και συχνότερα μεταλλάσσεται σε ασπαρτικό οξύ (G12D, 36%) ακολουθούμενο από βαλίνη (G12V, 23%) και κυστεΐνη (G12C, 14%) [26]. Πιο συγκεκριμένα η δεύτερη γουανίνη στο κωδικόνιο που κωδικοποιεί την γλυκίνη (GGT) αντικαθιστάται από μία αδενίνη (GAT) μετατρέποντας το κωδικόνιο σε ασπαρτικό οξύ. Ως εκ τούτου, αυτή η μελέτη εστιάζει στην πιο συχνά παρατηρούμενη μετάλλαξη KRAS-4B σε ασθενείς με καρκίνο, την G12D. Το G12 βρίσκεται στη ενεργό κέντρο της πρωτεΐνης, και συγκεκριμένα στον βρόχο δέσμευσης φωσφορικών. Τα αμινοξέα στο ενεργό κέντρο συνδέονται με τις φωσφορικές ομάδες του GTP και είναι υπεύθυνα για τη λειτουργία της KRAS-4B ως GTPάση. Οι περιοχές μεταγωγής SI και SII είναι επιπρόσθετα υπεύθυνες για τον έλεγχο της πρόσδεσης στις σηματοδοτικές πρωτεΐνες. Ωστόσο, η μετάλλαξη της γλυκίνης στη θέση 12 προς ασπαρτικό οξύ (G12D) οδηγεί στην προβολή μιας πιο ογκώδους και αρνητικά φορτισμένης ομάδας στο ενεργό κέντρο, η οποία προκαλεί στερική παρεμπόδιση στην υδρόλυση του GTP. Με τον τρόπο αυτό, η μετάλλαξη επηρεάζει τη λειτουργία της KRAS-4B ως GTPάση και κλειδώνει την KRAS-4B στην ενεργή της κατάσταση, συνδεδεμένη με GTP [27].

Πέρα από την αλληλεπίδραση της κάθε πρωτεΐνης KRAS-4B με το μόριο GTP που έχει συνδεθεί, ένα σημαντικό δομικό χαρακτηριστικό το οποίο εικάζεται ότι επηρεάζεται από τις μεταλλάξεις που οδηγούν σε καρκινογένεση είναι οι διεπιφάνειες διμερισμού, δηλαδή τα δομικά τμήματα του κάθε μονομερούς που αλληλεπιδρά με το άλλο προς τη δημιουργία του διμερούς.

1.1.3. Διεπιφάνειες διμερισμού της πρωτεΐνης KRAS-4B

Όπως είναι γνωστό, η πειραματική επίλυση της δομής μιας πρωτεΐνης είναι μία πολύπλοκη διαδικασία. Κατά κύριο λόγο χρησιμοποιούνται τρεις μέθοδοι για την επίτευξή της: η κρυσταλλογραφία ακτίνων Χ (x-ray crystallography), η οποία είναι η μέθοδος που χρησιμοποιείται συχνότερα για την επίλυση πρωτεϊνών που κατατίθενται στην Protein Data Bank (88%), ο πυρηνικός μαγνητικός συντονισμός (Nuclear Magnetic Resonance ή NMR, 7.5%) και η ηλεκτρονική μικροσκοπία (electron microscopy, 4.3%) [28] [29]. Κάθε μέθοδος έχει πλεονεκτήματα και μειονεκτήματα και σε κάθε περίπτωση απαιτούνται πειραματικά δεδομένα για να δημιουργηθεί το τελικό ατομικό μοντέλο της πρωτεΐνης. Για κρυσταλλογραφία ακτίνων Χ, αυτό είναι το πρότυπο περίθλασης ακτίνων Χ. Για τη φασματοσκοπία NMR, είναι πληροφορίες σχετικά με την τοπική διαμόρφωση και την απόσταση μεταξύ ατόμων που βρίσκονται κοντά το ένα στο άλλο. Στην ηλεκτρονική μικροσκοπία, είναι μια εικόνα του συνολικού σχήματος του μορίου [30].

Στις περισσότερες περιπτώσεις, αυτές οι πειραματικές πληροφορίες δεν επαρκούν για την δόμηση ενός ατομικού μοντέλου από την αρχή. Πρέπει να προστεθούν επιπλέον γνώσεις σχετικά με τη μοριακή δομή. Για παράδειγμα, συχνά γνωρίζουμε ήδη την αλληλουχία των αμινοξέων σε μια πρωτεΐνη και γνωρίζουμε την προτιμώμενη γεωμετρία των ατόμων σε μια τυπική πρωτεΐνη (για παράδειγμα, τα μήκη δεσμών και τις γωνίες δεσμών). Αυτές οι πληροφορίες επιτρέπουν την δημιουργία ενός μοντέλου που είναι συνεπές τόσο με τα πειραματικά δεδομένα όσο και με την αναμενόμενη σύνθεση και γεωμετρία του μορίου [31].

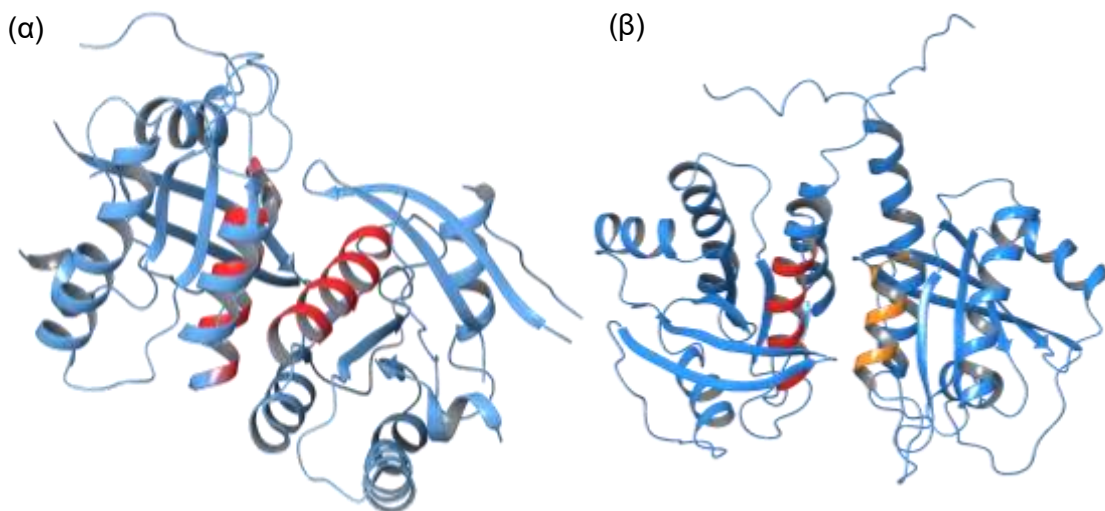
Ειδικά στη περίπτωση των πρωτεϊνών που συνδέονται με τη μεμβράνη, όπως είναι η KRAS-4B, είναι ακόμα πιο δύσκολο να επιλυθεί η δομή τους, καθώς δεν μπορεί να αναπαρασταθεί όλο το περιβάλλον της πρωτεΐνης και επομένως οι αλληλεπιδράσεις της με αυτό που ενδέχεται να επηρεάζουν τη δομή της. Επομένως αυτή τη στιγμή υπάρχουν μερικές χιλιάδες καταχωρήσεις για την KRAS-4B φυσικού τύπου και της μετάλλαξης στην Protein Data Bank, οι οποίες έχουν επιλυθεί με διάφορες πειραματικές ή υπολογιστικές μεθόδους, υπό διαφορετικές συνθήκες και σε διαφορετικά περιβάλλοντα. Ως αποτέλεσμα, μεταξύ των διαφορετικών αυτών δομών υπάρχουν ασυμφωνίες σχετικά με τη διεπιφάνεια διμερισμού

της πρωτεΐνης. Η πληροφορία αυτή είναι σημαντική, καθώς η ταυτοποίηση της τριδιάστατης διαμόρφωσης του διμερούς φυσικού τύπου σε σχέση με το μεταλλαγμένο μπορεί να οδηγήσει νέες ερευνητικές προσπάθειες για την ανάπτυξη φαρμάκων που στοχεύουν τη μεταλλαγμένη πρωτεΐνη KRAS-4B, με βάση τη δομή του στόχου. Παρακάτω περιγράφονται μερικές μόλις από τις διαθέσιμες δομές διμερούς KRAS-4B WT και KRAS-4B G12D και οι διεπιφάνειες διμερισμού τους.

Για το διμερές της KRAS-4B WT θα αναλυθούν οι δομές με PDB ID: 5VQ2, 6W4E της Protein Data Bank, οι οποίες παρουσιάζουν διαφορετικές διεπιφάνειες διμερισμού. Αντίστοιχα, για το διμερές της KRAS-4B G12D θα αναλυθούν οι δομές με PDB ID: 5US4, 5XCO, 6GJ5, 7ACA και 6GJ7, οι οποίες παρουσιάζουν διαφορετικές διεπιφάνειες.

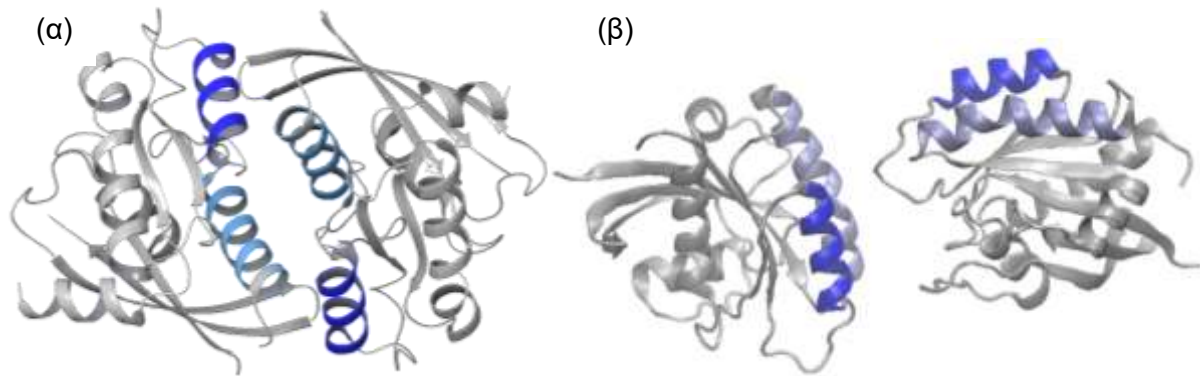
Η δομή με PDB ID 5VQ2, η οποία απεικονίζεται στο Σχήμα 1.11 (α), έχει χαρακτηριστεί με χρήση κρυσταλλογραφίας ακτίνων Χ και παρουσιάζει διεπιφάνεια διμερισμού μέσω της αλληλεπίδρασης των ελικών $\alpha 5$ του κάθε μονομερούς ($\alpha 5 - \alpha 5$), οι οποίες είναι χρωματισμένες με κόκκινο. Η δομή επιλύθηκε με την KRAS-4B συνδεδεμένη με το μόριο GTP, το οποίο όμως, όπως και όλα τα μικρά μόρια, δεν εμφανίζονται στην απεικόνιση με κυλίνδρους και κορδέλες που έχει χρησιμοποιηθεί [25].

Η δομή με PDB ID 6W4E στο Σχήμα 1.11 (β) είναι το διμερές της KRAS-4B WT συνδεδεμένο με το μόριο GTP σε ένα σύστημα νανοδίσκων, το οποίο δεν απεικονίζεται για ευκρίνεια. Το σύστημα επιλύθηκε με φασματοσκοπία NMR παραμαγνητικής ενίσχυσης χαλάρωσης (paramagnetic relaxation enhancement ή PRE NMR) και βρέθηκε ότι η διεπιφάνεια διμερισμού σχηματίζεται μέσω της αλληλεπίδρασης της έλικας $\alpha 4$ (κίτρινη) του ενός μονομερούς με την έλικα $\alpha 5$ (κόκκινη) του άλλου ($\alpha 4 - \alpha 5$) [21].



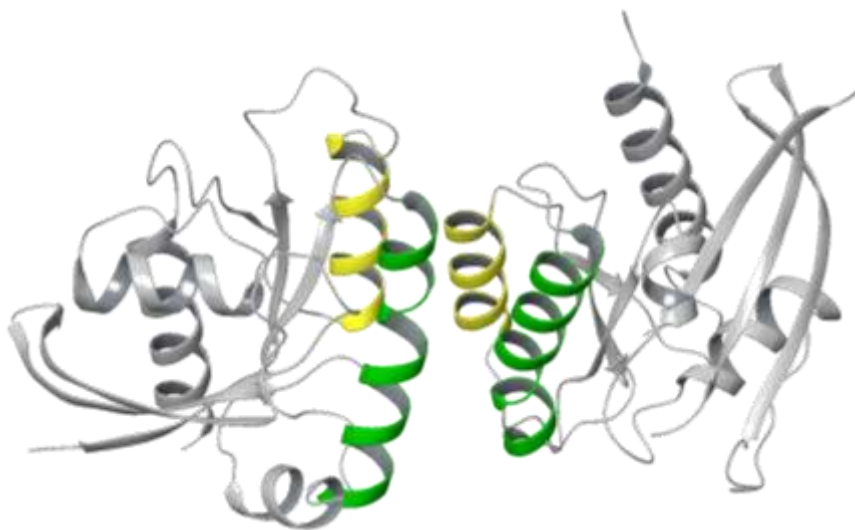
Σχήμα 1.11: Διμερή της KRAS-4B WT (α) βασισμένο στη δομή PDB ID 5VQ2 [25], με διεπιφάνεια διμερισμού $\alpha 5 - \alpha 5$, (β) βασισμένο στη δομή PDB ID 6W4E [21], με διεπιφάνεια διμερισμού $\alpha 4 - \alpha 5$.

Οι δομές με PDB ID 5US4 [32] και 5XCO [33], που απεικονίζονται στο Σχήμα 1.12 (α) και (β) επιλύθηκαν με την KRAS-4B G12D σε σύμπλοκο με GDP. Οι δομές προέκυψαν από κρυσταλλογραφία ακτίνων Χ και βρέθηκε ότι η διεπιφάνεια διμερισμού και για τις δύο δομές σχηματίζεται μέσω της αλληλεπίδρασης της έλικας α4 (γαλάζια) του ενός μονομερούς με την έλικα α5 (μπλε) του άλλου (α4 – α5).



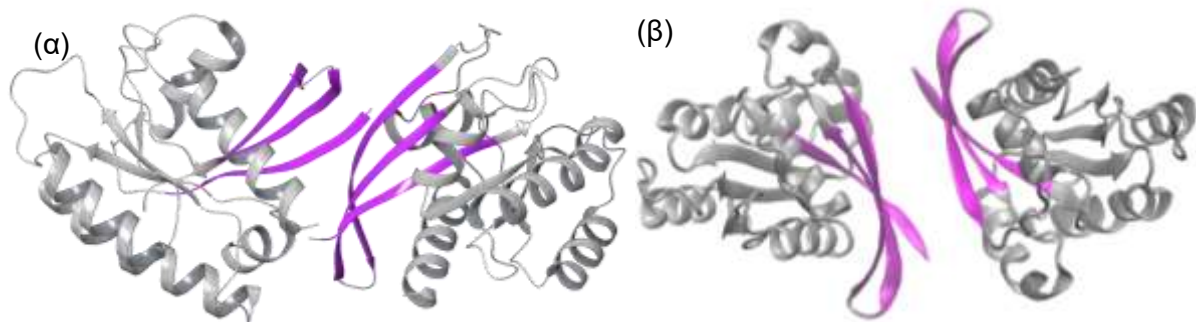
Σχήμα 1.12: Διμερή της KRAS-4B G12D με διεπιφάνεια διμερισμού α4 – α5 (α) βασισμένο στη δομή με PDB ID 5US4 [35] και (β) βασισμένο στη δομή με PDB ID 5XCO [36].

Η δομή με PDB ID 6GJ5, η οποία απεικονίζεται στο Σχήμα 1.13, έχει προκύψει από κρυσταλλογραφία ακτίνων Χ της KRAS-4B G12D και η διεπιφάνεια διμερισμού σχηματίζεται μέσω της μεταξύ τους αλληλεπίδρασης των ελικών α3 (γαλάζιες) του κάθε μονομερούς και μέσω της αλληλεπίδρασης των ελικών α4 (μπλε) μεταξύ τους (α3 – α3 και α4 – α4). Η δομή επιλύθηκε με την KRAS-4B συνδεδεμένη με ένα GTP ανάλογο [34].



Σχήμα 1.13: Διμερές της KRAS-4B G12D βασισμένο στη δομή με PDB ID 6GJ5, με διεπιφάνεια διμερισμού α3 – α3 και α4 – α4 [34].

Τέλος, τα διμερή που αναπαριστώνται στο Σχήμα 1.14 (α) και (β) είναι οι δομές με PDB ID 7ACA και 6GJ7, αντίστοιχα, με την KRAS-4B G12D να βρίσκεται συνδεδεμένη με ένα μόριο ανάλογο του GTP. Το σύστημα επιλύθηκε με κρυσταλλογραφία ακτίνων Χ και βρέθηκε ότι η διεπιφάνεια διμερισμού γίνεται μέσω της αλληλεπίδρασης των ροζ β-πτυχωτών φύλλων του κάθε μονομερούς ($\beta 1$, $\beta 2$, $\beta 3$) [34].



Σχήμα 1.14: Διμερές της KRAS-4B με διεπιφάνεια διμερισμού β-πτυχωτών φύλλων (α) βασισμένο στη δομή με PDB ID 7ACA και (β) βασισμένο στη δομή με PDB ID 6GJ7 [36].

1.2. Στόχος της εργασίας

Όπως φαίνεται και από την παραπάνω ανάλυση, η δομή του διμερούς KRAS-4B στη βιοενεργή της μορφή δεν έχει ταυτοποιηθεί. Ωστόσο ο εντοπισμός των διαφορών του διμερούς μεταξύ του φυσικού τύπου και της μετάλλαξης G12D είναι σημαντικός για την αντιμετώπιση της μετάλλαξης που οδηγεί σε καρκίνο. Ένας τρόπος αναστολής της μεταλλαγμένης πρωτεΐνης, θα ήταν να ανασταλεί ο σχηματισμός του διμερούς της πρωτεΐνης, η οποία όταν διμερίζεται, συμπλέκεται με άλλα σηματοδοτικά μόρια και δίνει σήμα στο κύτταρο για πολλαπλασιασμό. Επομένως, η στρατηγική για την θεραπεία του καρκίνου που οφείλεται στη μετάλλαξη G12D της KRAS-4B είναι η αναστολή του σχηματισμού του μεταλλαγμένου διμερούς.

Καθώς στη βιβλιογραφία δεν είναι ξεκάθαρα καθορισμένες οι διεπιφάνειες διμερισμού της KRAS-4B φυσικού τύπου και της μετάλλαξης G12D, σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάλυση των διεπιφανειών διμερισμού της ογκοπρωτεΐνης KRAS-4B με χρήση αδροποιημένων μοντέλων και προσομοιώσεις Μοριακής Δυναμικής στο νερό και η σύγκριση των διεπιφανειών διμερισμού με αυτών της βιβλιογραφίας. Η ανάλυση αυτή μπορεί στη συνέχεια να αξιοποιηθεί συνδυαστικά με πειραματικά δεδομένα για τον σχεδιασμό φαρμάκων κατά της μετάλλαξης G12D για τη θεραπεία του καρκίνου.

2. Μέθοδοι

2.1. Μοριακή Μοντελοποίηση

Η μοριακή μοντελοποίηση βασίζεται στην ανάπτυξη θεωρητικών και υπολογιστικών μεθόδων, για τη μοντελοποίηση και τη μελέτη της συμπεριφοράς και των ιδιοτήτων ατόμων και μορίων, από μικρά χημικά συστήματα έως μεγάλα βιολογικά μόρια και συστήματα υλικών. Στόχος των μοντελοποιήσεων αυτών είναι η μίμηση της συμπεριφοράς μορίων και συστημάτων μορίων. Οι μοριακές μοντελοποιήσεις βρίσκουν εφαρμογή σε τομείς όπως η υπολογιστική χημεία, ο σχεδιασμός φαρμάκων, η υπολογιστική βιολογία και η επιστήμη των υλικών. Η βασική μέθοδος για την ανάπτυξη μοριακών μοντελοποιήσεων είναι η προσομοίωση. Μέσω των μοριακών προσομοιώσεων, παρέχεται η δυνατότητα προσομοίωσης συνθηκών που είναι δύσκολες έως αδύνατες να πραγματοποιηθούν στο εργαστήριο, όπως μελέτη ενός συστήματος σε ακραίες θερμοκρασίες ή πιέσεις [35] [36].

2.2. Εμπειρική συνάρτηση δυναμικής ενέργειας - πεδίο δυνάμεων

Η θεωρητική μελέτη βιολογικών μορίων επιτρέπει τη μελέτη των σχέσεων μεταξύ δομής και λειτουργίας στο ατομικό επίπεδο. Σε μακροσκοπικό επίπεδο η κίνηση των σωματιδίων ενός συστήματος μπορεί να περιγραφεί πλήρως από την Κλασική Μηχανική, επιλύοντας τις εξισώσεις κίνησης του Νεύτωνα. Ωστόσο, οι νόμοι της Κλασικής Μηχανικής αδυνατούν να περιγράψουν την κίνηση τόσο μικρών σωματιδίων όπως τα ηλεκτρόνια. Για το σκοπό αυτό εφαρμόζονται οι εξισώσεις της Κβαντικής Μηχανικής, η οποία αναγνωρίζει το δυϊσμό των σωματιδίων αυτών θεωρώντας ότι ένα σωματίδιο αντί να ταξιδεύει κατά μήκος μιας καθορισμένης τροχιάς, διανέμεται διαμέσου του χώρου ως κύμα. Τα μοντέλα της Κβαντικής Μηχανικής πηγάζουν από την εξίσωση του Schrödinger που πρωτοεμφανίστηκε στα τέλη της δεκαετίας του 1920, η οποία αντιμετωπίζει τα μόρια ως συλλογές πυρήνων και ηλεκτρονίων, δίχως καμία αναφορά σε «χημικούς δεσμούς». Δεδομένου όμως ότι τα βιολογικά συστήματα περιλαμβάνουν πολλά άτομα, η κβαντομηχανική αναπαράσταση αυτών των ατόμων δεν είναι εφικτή. Ο πιο συνηθισμένος τρόπος για την αναπαράστασή τους είναι η χρήση εμπειρικών συναρτήσεων δυναμικής ενέργειας (force field), οι οποίες αποτελούν τη μαθηματική έκφραση του δυναμικού το οποίο ασκείται σε όλα τα άτομα ενός συστήματος. Οι συναρτήσεις αυτές ενώ είναι λιγότερο ακριβείς υπολογιστικά, περιλαμβάνουν παραδοχές που οδηγούν σε μία ρεαλιστική αναπαράσταση [35].

Πιο συγκεκριμένα, τα πεδία δυνάμεων, όπως ονομάζονται εναλλακτικά αυτές οι εμπειρικές συναρτήσεις δυναμικής ενέργειας, χρησιμοποιούνται για να περιγράψουν τις ενδομοριακές και διαμοριακές αλληλεπιδράσεις που λαμβάνουν χώρα μεταξύ των πυρήνων των ατόμων του συστήματος. Η περιγραφή των ιδιοτήτων του συστήματος είναι αρκετά ακριβής, ακόμη και όταν χρησιμοποιούνται απλές συναρτήσεις, για να περιγράψουν αυτές τις συνεισφορές στην ολική ενέργεια, όπως ο νόμος του Hooke για την αναπαράσταση των δεσμών σαν ελατήρια. Σημειώνεται επίσης ότι κάθε άτομο αναπαριστάται από ένα σωματίδιο, ενώ τα ηλεκτρόνια δεν αναπαριστώνται ξεχωριστά.

Τα πεδία δυνάμεων πραγματοποιούν έναν συμβιβασμό μεταξύ συμφωνίας με τα πειραματικά δεδομένα και υπολογιστικής ταχύτητας. Συχνά βαθμονομούνται με χρήση πειραματικών αποτελεσμάτων και με κβαντομηχανικούς υπολογισμούς απλών μοντέλων ενώσεων. Μεταξύ των συνηθέστερα χρησιμοποιούμενων συναρτήσεων δυναμικής ενέργειας είναι τα πεδία δύναμης AMBER, CHARMM, GROMOS και OPLS. Ένας από τους πιο σημαντικούς περιορισμούς των εμπειρικών πεδίων δυνάμεων είναι ότι δεν επιτρέπεται να πραγματοποιηθούν δραστικές αλλαγές στην ηλεκτρονική δομή. Δηλαδή δεν μπορούν να μοντελοποιηθούν γεγονότα όπως η δημιουργία ή το σπάσιμο δεσμών.

Πρακτικά οι εμπειρικές συναρτήσεις δυναμικής ενέργειας έχουν διαμορφωθεί έτσι ώστε να επιτρέπουν τον υπολογισμό της ενέργειας του συστήματος συναρτήσει των συντεταγμένων του, καθώς η ολική ενέργεια του συστήματος αναπαριστάται από την ενέργεια H (Εξίσωση 2.1).

$$H = T + V = \frac{1}{2} mv^2 + V(r) \quad (2.1)$$

Όπου:

T : η κινητική ενέργεια

V : η δυναμική ενέργεια

Η δυναμική ενέργεια του συστήματος $V(r)$ μπορεί να περιγραφεί ως το άθροισμα των ομοιοπολικών (bonded) και των μη ομοιοπολικών (non-bonded) αλληλεπιδράσεων (Εξίσωση 2.2).

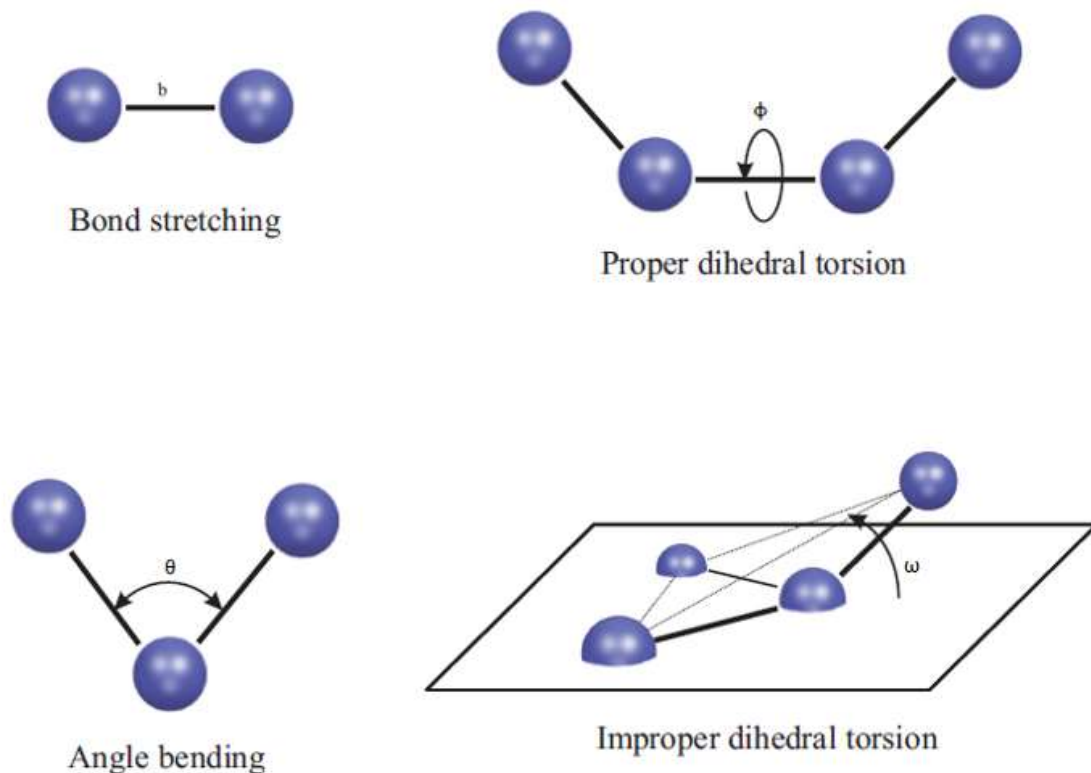
$$V(\mathbf{r}) = V_{bonded} + V_{non-bonded} \quad (2.2)$$

Στις ομοιοπολικές αλληλεπιδράσεις συμπεριλαμβάνονται οι όροι της τάσης των δεσμών (bond stretching), της κάμψης των γωνιών (angle bending), των δίεδρων γωνιών (dihedrals) και των ακατάλληλων δίεδρων γωνιών (improper dihedral) (Εξίσωση 2.3). Αυτοί οι όροι σχεδιάζονται σχηματικά στο Σχήμα 2.1.

$$V_{bonded} = V_{bonds} + V_{angles} + V_{dihedrals} + V_{improper\ dihedral} \quad (2.3)$$

Οι όροι των μη ομοιοπολικών αλληλεπιδράσεων $V_{non-bonded}$ περιγράφουν τις δυνάμεις van der Waals και τις ηλεκτροστατικές αλληλεπιδράσεις Coulomb μεταξύ των ατόμων (Εξίσωση 2.4). Οι διαφορετικοί όροι θα παρουσιαστούν παρακάτω με περισσότερες λεπτομέρειες.

$$V_{non-bonded} = V_{vdW} + V_{elec} \quad (2.4)$$



Σχήμα 2.1: Σχηματική αναπαράσταση των όρων των ομοιοπολικών αλληλεπιδράσεων.

Ο πρώτος ομοιοπολικός όρος αντιστοιχεί στη τάση των δεσμών και περιγράφει τις δυνάμεις που δρουν μεταξύ δύο ομοιοπολικά συνδεδεμένων ατόμων. Ο όρος μπορεί να περιγραφεί από ένα αρμονικό δυναμικό βασισμένο στο νόμο του Hooke (Εξίσωση 2.5).

$$V_{bonded} = k_b(b - b_0)^2 \quad (2.5)$$

Όπου:

b : είναι η απόσταση μεταξύ των δύο ατόμων.

b_0 : η μέση απόσταση μεταξύ των ατόμων στη θέση της ισορροπίας

k_b : η σταθερά της δύναμης που χαρακτηρίζει την αλληλεπίδραση

Ο δεύτερος ομοιοπολικός όρος αντιστοιχεί στη κάμψη των δεσμών και περιγράφει τη δόνηση μεταξύ τριών ομοιοπολικά συνδεδεμένων ατόμων. Ο όρος αυτός περιγράφεται επίσης χρησιμοποιώντας ένα αρμονικό δυναμικό βασιζόμενο στο νόμο του Hooke (Εξίσωση 2.6).

$$V_{angles} = k_{\theta}(\theta - \theta_0)^2 \quad (2.6)$$

Όπου:

θ : η γωνία μεταξύ των τριών ατόμων

θ_0 : η γωνία μεταξύ των τριών ατόμων στη θέση της ισορροπίας

k_{θ} : η σταθερά της δύναμης που χαρακτηρίζει τη δόνηση της γωνίας

Ο τρίτος ομοιοπολικός όρος αντιστοιχεί στα δίεδρα και περιγράφει το φράγμα περιστροφής μεταξύ τεσσάρων ομοιοπολικά συνδεδεμένων ατόμων. Τα δίεδρα (dihedrals) χαρακτηρίζονται από τη συνημιτονοειδή συνάρτηση δυναμικού (Εξίσωση 2.7).

$$V_{dihedrals} = k_{\varphi}(1 + \cos[n\varphi - \delta]), \quad n = 1,2,3,4,6 \quad (2.7)$$

Όπου:

φ : η γωνία μεταξύ των επιπέδων που σχηματίζονται από το πρώτο και τα τρία τελευταία από τα τέσσερα άτομα.

δ : καθορίζει τη γωνία με την ελάχιστη ενέργεια

k_{φ} : η σταθερά δυνάμεως

n : η περιοδικότητα

Ο τελευταίος ομοιοπολικός όρος αντιστοιχεί στα ακατάλληλα δίεδρα και διατηρεί τόσο την τετραεδρική γεωμετρία γύρω από ένα χειρόμορφο άτομο όσο και την επίπεδη διαμόρφωση γύρω από συγκεκριμένες ομάδες ατόμων, όπως πχ ο βενζολικός δακτύλιος και περιγράφεται από ένα αρμονικό δυναμικό (Εξίσωση 2.8).

$$V_{improper\ dihedral} = k_{\omega}(\omega - \omega_0)^2 \quad (2.8)$$

Όπου:

ω : η γωνία μεταξύ του επιπέδου που σχηματίζεται από το κεντρικό άτομο, τα δύο περιφερειακά άτομα και το επίπεδο που σχηματίζεται από τα περιφερειακά άτομα.

ω_0 : η γωνία μεταξύ των ατόμων στη θέση ισορροπίας

k_{ω} : η σταθερά της δύναμης που χαρακτηρίζει τη δόνηση της γωνίας

Ο πρώτος μη ομοιοπολικός όρος περιγράφει τις δυνάμεις van der Waals, οι οποίες δρουν σε δύο άτομα, i και j που είναι γειτονικά μεταξύ τους. Είναι απωστικές σε μικρή απόσταση μεταξύ των ατόμων, εμφανίζουν ένα βάθος δυναμικού στη βέλτιστη απόσταση αλληλεπίδρασης μεταξύ δύο ατόμων και στη συνέχεια είναι ελκτικές σε μεγαλύτερες αποστάσεις και τείνουν στο μηδέν σε μεγαλύτερες αποστάσεις. Η αλληλεπίδραση μπορεί να προσεγγιστεί από το δυναμικό Lennard-Jones, το οποίο απεικονίζεται στο Σχήμα 2.2 (Εξίσωση 2.9).

$$V_{vdW} = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (2.9)$$

Όπου:

r_{ij} : η απόσταση μεταξύ των δύο ατόμων

σ_{ij} : η ακτίνα van der Waals, η απόσταση για την οποία η δυναμική ενέργεια είναι μηδέν

ε_{ij} : το βάθος του πηγαδιού, υποδηλώνει την ισχύ της αλληλεπίδρασης

Ο δεύτερος μη ομοιοπολικός όρος είναι οι ηλεκτροστατικές αλληλεπιδράσεις μεταξύ δύο ατόμων i και j και περιγράφονται από τον νόμο του Coulomb (Εξίσωση 2.10).

$$V_{elec} = \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}} \quad (2.10)$$

Όπου:

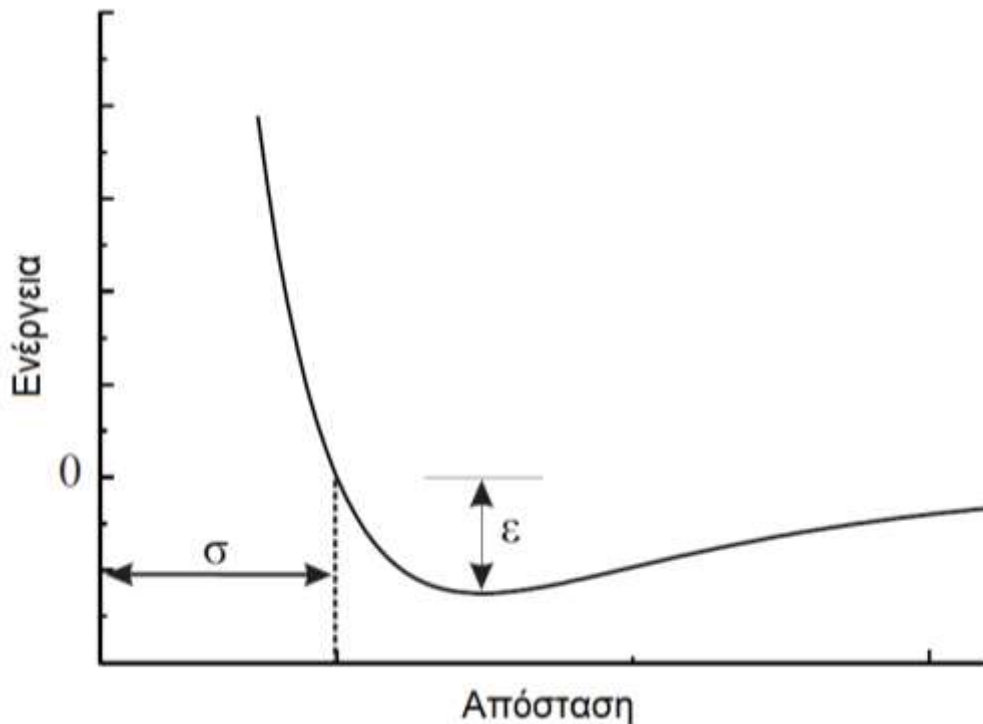
q_i και q_j : τα φορτία των δύο ατόμων

r_{ij} : η απόσταση μεταξύ των δύο ατόμων

ε_0 : είναι η διηλεκτρική σταθερά του κενού

Συνολικά, το πεδίο δυνάμεων περιγράφεται από την Εξίσωση 2.11.

$$\begin{aligned} V(\mathbf{r}) = & \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi (1 + \cos[n\phi - \delta]) \\ & + \sum_{impropers} k_\omega (\omega - \omega_0)^2 + \sum_{i,j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \\ & + \sum_{i,j} \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}} \quad (2.11) \end{aligned}$$



Σχήμα 2.2: Το δυναμικό Lennard-Jones. Η δυναμική ενέργειας (V) απεικονίζεται συναρτήσει της απόστασης των ατόμων (r). Στο διάγραμμα παρουσιάζεται και η ακτίνα van der Waals (σ) και το βάθος του πηγαδιού (ϵ) [37].

2.3. Προσομοιώσεις Μοριακής Δυναμικής

Ένα από τα κύρια εργαλεία για την προσομοίωση των βιομορίων στο χρόνο είναι η μέθοδος της Μοριακής Δυναμικής. Η συγκεκριμένη υπολογιστική μέθοδος υπολογίζει τη συμπεριφορά ενός μοριακού συστήματος σε συνάρτηση με τον χρόνο και χρησιμοποιείται για να περιγράψει ένα πολύπλοκο μοριακό σύστημα σε ατομικό επίπεδο. Στόχος της Μοριακής Δυναμικής είναι η κατανόηση και η πρόβλεψη μακροσκοπικών ιδιοτήτων, με βάση γνωστές ιδιότητες στην ατομική κλίμακα.

Κατά τη πραγματοποίηση προσομοιώσεων Μοριακής Δυναμικής χρησιμοποιείται μία εμπειρική συνάρτηση δυναμικής ενέργειας, προκειμένου να υπολογιστεί η ενέργεια ενός συστήματος, συνήθως με βάση τις θέσεις των πυρήνων των ατόμων. Προκειμένου να περιγραφεί το σύστημα με τις προσομοιώσεις Μοριακής Δυναμικής ο δεύτερος νόμος του Νεύτωνα (Εξίσωση 2.12) συνδυάζεται με την δυναμική ενέργεια (Εξίσωση 2.13) και προκύπτει η Εξίσωση 2.14, η οποία ολοκληρώνεται αριθμητικά και παράγει την εξέλιξη της θέσης των ατόμων του συστήματος στον χρόνο της προσομοίωσης, που ονομάζεται τροχιά (trajectory). Από την τροχιά αυτή προκύπτουν πληροφορίες σχετικά με το σύστημα σε

μικροσκοπικό επίπεδο, οι οποίες χρησιμοποιούνται για τον υπολογισμό των μακροσκοπικών φυσικών ιδιοτήτων του συστήματος.

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2} \quad (2.12)$$

$$\mathbf{F}_i = -\nabla V(\mathbf{r}_i) \quad (2.13)$$

$$-\nabla V(\mathbf{r}_i) = m_i \frac{d^2 \mathbf{r}_i}{dt^2} \quad (2.14)$$

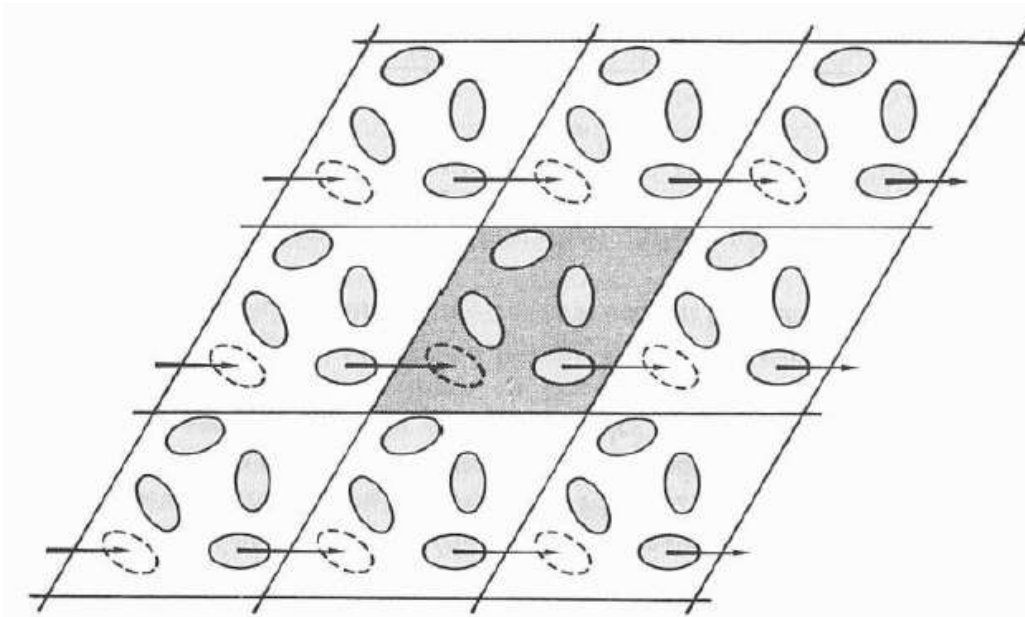
Στη συνέχεια, οι πληροφορίες που έχουν προκύψει στο μικροσκοπικό επίπεδο, όπως η θέση και η ταχύτητα των ατόμων, συνδέονται με τα μακροσκοπικά μεγέθη, όπως η πίεση και η ενέργεια, μέσω της στατιστικής μηχανικής. Αυτό επιτυγχάνεται μέσα από τη χρήση των στατιστικών συνόλων. Ένα στατιστικό σύνολο (ensemble) είναι μια συλλογή σημείων στο χώρο που ικανοποιεί τις συνθήκες μιας συγκεκριμένης θερμοδυναμικής κατάστασης. Τα στατιστικά μηχανικά σύνολα περιγράφουν μια συλλογή από πιθανά συστήματα τα οποία μπορεί να έχουν διαφορετικές μικροσκοπικές καταστάσεις αλλά παρόμοιες μακροσκοπικές καταστάσεις.

Η θερμοδυναμική κατάσταση ενός συστήματος καθορίζεται συνήθως από ένα μικρό σύνολο παραμέτρων, για παράδειγμα τη θερμοκρασία, την πίεση και τον αριθμό των σωματιδίων. Άλλες θερμοδυναμικές ιδιότητες μπορούν να προκύψουν από τις καταστατικές εξισώσεις και άλλες θεμελιώδεις θερμοδυναμικές εξισώσεις. Στη συνέχεια παρουσιάζεται μια σύνοψη των διαφορετικών στατιστικών συνόλων με διαφορετικά χαρακτηριστικά.

- Μικρο-κανονικό σύνολο (Microcanonical – NVE): οι τρεις σταθερές παράμετροι είναι ο αριθμός των σωματιδίων (N), ο όγκος (V) και η ενέργεια του συστήματος (E). Αυτό είναι το φυσικό σύνολο για προσομοιώσεις Μοριακής Δυναμικής, όπου οι Νευτώνειες εξισώσεις κίνησης μπορούν να εφαρμοστούν αμετάβλητες στο σύστημα. Η κατανομή αυτή περιγράφει ένα απομονωμένο σύστημα.
- Κανονικό σύνολο (Canonical – NVT): οι σταθερές παράμετροι του συστήματος είναι ο αριθμός σωματιδίων (N), ο όγκος (V) και η θερμοκρασία (T). Δεδομένου ότι η θερμοκρασία του συστήματος πρέπει να διατηρηθεί σταθερή για αυτό το στατιστικό σύνολο, πρέπει να εισαχθεί ένας θερμοστάτης, ώστε η κινητική και δυναμική ενέργεια που υπολογίζονται να είναι ακριβείς.
- Ισόθερμο – ισοβαρές σύνολο (Isothermal–isobaric – NpT): οι σταθερές παράμετροι του συστήματος είναι ο αριθμός σωματιδίων (N), η πίεση (p) και η θερμοκρασία (T). Στην περίπτωση αυτή, ο νόμος κίνησης του Νεύτωνα γίνεται ακόμα πιο πολύπλοκος

επειδή εκτός από το θερμοστάτη, το σύστημα συνδέεται και με βαροστάτη ούτως ώστε η κατανομή να προσομοιάζει καλύτερα τις κανονικές πειραματικές συνθήκες.

Επιπρόσθετα, για να αποφευχθούν τεχνητά οριακά φαινόμενα στις προσομοιώσεις Μοριακής Δυναμικής τα οποία μπορεί να έχουν ως αποτέλεσμα τη δημιουργία διεπιφανειών στα όρια του συστήματος με το περιβάλλον του, χρησιμοποιούνται περιοδικές οριακές συνθήκες (periodic boundary conditions). Με την προσέγγιση αυτή θεωρείται ουσιαστικά ότι το κουτί της προσομοίωσης είναι περιοδικά συνεχές. Έτσι, η μπροστά του επιφάνεια είναι απολύτως ίδια με την πίσω, η δεξιά του απολύτως ίδια με την αριστερή του και η άνω του επιφάνεια απολύτως ίδια με την κάτω (Σχήμα 2.4). Συνεπώς όταν ένα μόριο εγκαταλείπει το κεντρικό κουτί, τότε μια περιοδική του εικόνα εισέρχεται ακριβώς από την απέναντι πλευρά. Κατά τις μετακινήσεις των μορίων η αριθμητική πυκνότητα στο κεντρικό κουτί παραμένει σταθερή. Όπως φαίνεται και στο Σχήμα 2.3, το κεντρικό κουτί είναι το σκιασμένο με βάση το οποίο δημιουργούνται οκτώ κλώνοι του προς όλες τις κατευθύνσεις. Τα άτομα στα περιβάλλοντα εικονικά συστήματα αλληλεπιδρούν με τα άτομα στο πραγματικό σύστημα [36].



Σχήμα 2.3: Οι περιοδικές οριακές συνθήκες ενός συστήματος [36].

Για την πραγματοποίηση προσομοιώσεων Μοριακής Δυναμικής απαιτείται μία σειρά βημάτων, τα οποία περιγράφονται αναλυτικά στις επόμενες υποπαραγράφους. Πιο συγκεκριμένα αυτά είναι η προετοιμασία του συστήματος που προσομοιάζεται, η ελαχιστοποίηση της ενέργειας, η θέρμανση και η εξισορρόπηση του συστήματος και τέλος η παραγωγή της τροχιάς και η ανάλυση κάποιων ενδεικτικών μεγεθών κατά τη διάρκεια της προσομοίωσης.

Ένα από τα πιο δημοφιλή ελεύθερα πακέτα για την εκτέλεση προσομοιώσεων Μοριακής Δυναμικής είναι το GROMACS [38] [39], το οποίο και χρησιμοποιείται στη παρούσα εργασία, σε περιβάλλον Linux. Το GROMACS έχει σχεδιαστεί κυρίως για βιοχημικά μόρια όπως πρωτεΐνες, λιπίδια και νουκλεϊκά οξέα. Η αναλυτική καταγραφή των δυνατοτήτων του πακέτου, αλλά και η ενεργή κοινότητα που το χρησιμοποιεί αποτελεί και από τα κύρια πλεονεκτήματά του, που βοηθούν στην γρήγορη εξοικείωση με το πακέτο και την επίλυση των προβλημάτων που προκύπτουν. Η έκδοση 2020.6 είναι αυτή που χρησιμοποιήθηκε για τις προσομοιώσεις της παρούσας εργασίας [35] [36].

2.3.1. Προετοιμασία συστήματος

Προκειμένου να μοντελοποιηθεί ένα βιολογικό σύστημα, όπως το διμερές της πρωτεΐνης KRAS-4B σε διάλυμα νερού με τα κατάλληλα ιόντα, επιλέγεται μία αρχική δομή της πρωτεΐνης. Η δομή αυτή συνήθως προέρχεται από την βάση δεδομένων πρωτεϊνών [40]. Η συγκεκριμένη βάση δεδομένων έχει καθιερωθεί ως η κύρια πηγή για την ελεύθερη πρόσβαση σε δομές βιολογικών μορίων.

Οι δομές που βρίσκονται στη βάση δεδομένων πρωτεϊνών είναι αποτέλεσμα πειραματικών τεχνικών όπως η περίθλαση ακτίνων Χ, η ηλεκτρονική μικροσκοπία και ο πυρηνικός μαγνητικός συντονισμός (NMR). Η μέθοδος με την οποία έχει απεικονιστεί η κάθε δομή επηρεάζει τη τελική δομή του συστήματος, καθώς η κάθε μέθοδος απαιτεί διαφορετική προετοιμασία του συστήματος και αλληλεπίδρα με διαφορετικό τρόπο με αυτό. Επομένως, η επιλογή της δομής απαιτεί ιδιαίτερη προσοχή, καθώς επηρεάζει την μετέπειτα πορεία της προσομοίωσης.

2.3.2. Ελαχιστοποίηση ενέργειας

Η εμπειρική συνάρτηση δυναμικής ενέργειας ενός βιομοριακού συστήματος είναι ένα πολύ περίπλοκο και πολυδιάστατο τοπίο, το οποίο περιγράφει τις ενδομοριακές και διαμοριακές αλληλεπιδράσεις που λαμβάνουν χώρα μεταξύ των ατόμων του συστήματος, όπως αναφέρθηκε προηγουμένως. Όταν αναπαρασταθεί γραφικά ως συνάρτηση κάποιας συλλογικής μεταβλητής, όπως θα αναλυθεί στη συνέχεια, έχει ένα βαθύτερο σημείο, το ολικό ελάχιστο και έναν πολύ μεγάλο αριθμό τοπικών ελαχίστων. Ο στόχος της ελαχιστοποίησης της ενέργειας είναι να βρεθεί ένα τοπικό ελάχιστο. Η ενέργεια σε αυτό το τοπικό ελάχιστο μπορεί να είναι πολύ υψηλότερη από την ενέργεια του ολικού ελαχίστου. Από φυσικής

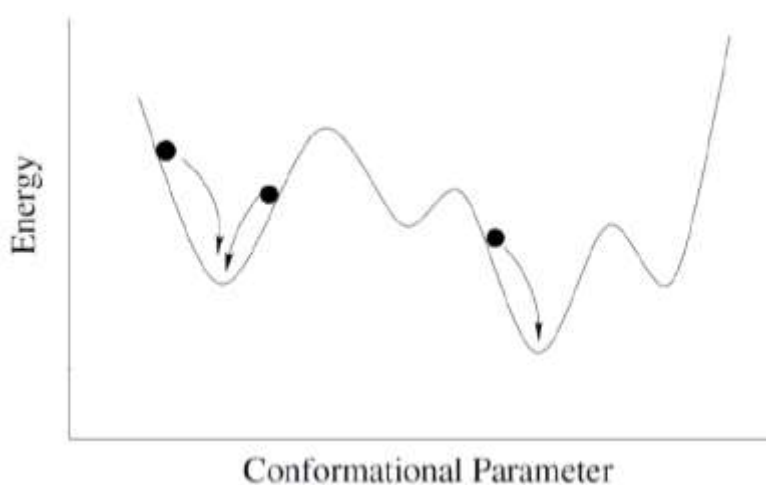
σκοπιάς, η ελαχιστοποίηση της ενέργειας αντιστοιχεί σε στιγμιαία «ψύξη» του συστήματος στους 0 K. Η πραγματοποίηση της ελαχιστοποίησης ενέργειας θα εγγυηθεί την κατάρρευση τυχόν δυσμενών αλληλεπιδράσεων van der Waals που ενδέχεται να υπάρχουν, οι οποίες διαφορετικά θα μπορούσαν να οδηγήσουν σε τοπική παραμόρφωση της δομής και να οδηγήσουν σε αστάθεια της προσομοίωσης.

Η εμπειρική συνάρτηση δυναμικής ενέργειας ενός συστήματος περιγράφεται από μία συνάρτηση, έστω f , η οποία εξαρτάται από τις μεταβλητές x_1, x_2, \dots, x_n και η μορφή της οποίας μπορεί να περιγραφεί από την Εξίσωση 2.11. Ως ελάχιστο της συνάρτησης f , ορίζεται το σημείο στο οποίο η πρώτη παράγωγος της συνάρτησης f ως προς κάθε μεταβλητή x_i είναι μηδέν (Εξίσωση 2.15) και όλες οι δεύτερες παράγωγοι έχουν θετικές τιμές (Εξίσωση 2.16).

$$\frac{\partial f}{\partial x_i} = 0 \quad (2.15)$$

$$\frac{\partial^2 f}{\partial x_i^2} > 0 \quad (2.16)$$

Οι συναρτήσεις δυναμικής ενέργειας των βιομοριακών συστημάτων αντιμετωπίζονται αριθμητικά, αλλάζοντας σταδιακά τις συντεταγμένες του συστήματος ώστε να δημιουργήσουν τις εγγύτερα προσεγγίσιμες διαμορφώσεις χαμηλότερης ενέργειας, μέχρις ότου βρεθεί η διαμόρφωση με την ελάχιστη ενέργεια τοπικά. Είναι σημαντικό να σημειωθεί ότι δεν υπάρχουν τρόποι να προσπεραστούν οι δομές που βρίσκονται ανάμεσα σε δύο τοπικά ελάχιστα προκειμένου να βρεθεί το ολικό ελάχιστο ενέργειας του συστήματος, χωρίς την παροχή κάποιας εξωτερικής ενέργειας στο σύστημα. Επομένως, όπως φαίνεται και στο Σχήμα 2.4, γίνεται εμφανές ότι είναι ιδιαίτερα σημαντική η επιλογή της αρχικής δομής που θα χρησιμοποιηθεί στην προσομοίωση.



Σχήμα 2.4: Παράδειγμα διαγράμματος ενέργειας συναρτήσει μιας διαμορφωτικής παραμέτρου, βάσει της οποίας χαρτογραφείται το υπό μελέτη σύστημα [36].

2.3.3. Θέρμανση συστήματος

Σε κάθε άτομο του συστήματος αντιστοιχίζεται μία αρχική ταχύτητα σε χαμηλή θερμοκρασία και οι εξισώσεις κίνησης του Νεύτωνα ολοκληρώνονται σε όλη την έκταση του συστήματος. Κατά τη θέρμανση του συστήματος, οι αρχικές ταχύτητες ανατίθενται στα άτομα σε χαμηλή θερμοκρασία, τυχαία, μέσω της κατανομής Maxwell-Boltzmann (Εξίσωση 2.17). Η προσομοίωση ξεκινά με περιοδική ανάθεση νέων ταχύτητας σε ελαφρώς υψηλότερη θερμοκρασία και αφήνοντας την προσομοίωση να συνεχιστεί. Αυτό το βήμα επαναλαμβάνεται έως ότου επιτευχθεί η επιθυμητή θερμοκρασία.

$$p(V_i) = \left(\frac{m_i}{2\pi k_B T} \right)^{\frac{1}{2}} e^{-\frac{1}{2} \frac{(m_i V_i^2)}{k_B T}} \quad (2.17)$$

Όπου:

$p(V_i)$: η πιθανότητα ενός ατόμου i να έχει ταχύτητα $V_i = (V_{ix}, V_{iy}, V_{iz})$

k_B : η σταθερά Boltzmann

m_i : η μάζα του ατόμου i

T : η απόλυτη θερμοκρασία στην οποία βρίσκεται το σύστημα

2.3.4. Εξισορρόπηση συστήματος

Μετά τη θέρμανση του συστήματος και αφού επιτευχθεί η επιθυμητή θερμοκρασία, η προσομοίωση συνεχίζεται με την φάση της εξισορρόπησης. Κατά τη διάρκεια αυτής της φάσης, ιδιότητες όπως η πίεση και η ενέργεια παρακολουθούνται σε συνάρτηση με τον χρόνο. Ο σκοπός της φάσης εξισορρόπησης είναι η πραγματοποίηση της προσομοίωσης έως ότου αυτές οι ιδιότητες να σταθεροποιηθούν. Εάν κατά τη διαδικασία αυτή υπάρχουν σημαντικές διακυμάνσεις στη θερμοκρασία, οι ταχύτητες αναπροσαρμόζονται έτσι ώστε η θερμοκρασία να επιστρέφει στην επιθυμητή τιμή και να επιτευχθεί θερμοδυναμική ισορροπία.

Κατά το στάδιο αυτό επιλέγεται υπό ποιες συνθήκες θα πραγματοποιηθεί η εξισορρόπηση. Δύο από τα συνηθέστερα θερμοδυναμικά σύνολα στα οποία πραγματοποιείται εξισορρόπηση είναι το Κανονικό (Canonical) σύνολο και το Ισόθερμο - Ισοβαρές σύνολο. Στο Κανονικό σύνολο διατηρείται σταθερός ο αριθμός των ατόμων του συστήματος (N), ο όγκος (V) και η θερμοκρασία του συστήματος (T). Επομένως για χάρη συντομίας το Κανονικό σύνολο συμβολίζεται ως NVT . Αντίστοιχα, στο Ισόθερμο - Ισοβαρές σύνολο διατηρείται σταθερός ο αριθμός των ατόμων του συστήματος (N), η πίεση (P) και η θερμοκρασία του συστήματος (T), ενώ συμβολίζεται με τα αρχικά NPT για συντομία.

2.3.5. Παραγωγή τροχιάς και ανάλυση προσομοίωσης

Το τελευταίο βήμα μιας προσομοίωσης Μοριακής Δυναμικής είναι η φάση της παραγωγής της τροχιάς, όπου το σύστημα προσομοιώνεται για το χρονικό διάστημα που απαιτείται. Το χρονικό διάστημα αυτό μπορεί να ανήκει τυπικά στο εύρος των μερικών εκατοντάδων ps έως ns ή και περισσότερο. Κατά τη διάρκεια αυτής της διαδικασίας οι συντεταγμένες του συστήματος σε διαφορετικούς χρόνους αποθηκεύονται με τη μορφή τροχιών. Οι τροχές αυτές χρησιμοποιούνται στη συνέχεια για να υπολογιστούν η μέση δυναμική και κινητική ενέργεια, οι διακυμάνσεις του μέσου τετραγώνου ρίζας (Root Mean Square ή RMS) μεταξύ δομών κ.λπ. Από τις προσομοιώσεις Μοριακής Δυναμικής, μπορούν επίσης να υπολογιστούν ιδιότητες συναρτήσεως του χρόνου, όπως οι συναρτήσεις συσχέτισης και αυτές με τη σειρά τους μπορούν να διασταυρωθούν με φασματοσκοπικές μετρήσεις. Δύο ενδιαφέρουσες ιδιότητες που μπορούν να υπολογιστούν από την τροχιά ενός συστήματος διμερισμού είναι η ρίζα μέσης τετραγωνικής απόκλισης σε συνάρτηση με το χρόνο και η απόσταση των κέντρων μαζών των μονομερών.

Η ρίζα της μέσης τετραγωνικής απόκλισης των ατομικών θέσεων του συστήματος (Root Mean Square Deviation ή RMSD) είναι το μέτρο της μέσης απόστασης μεταξύ των ατόμων δύο υπερτιθέμενων πρωτεϊνών. Στη περίπτωση μιας τροχιάς ενός διμερούς, το RMSD υπολογίζεται για υπερτιθέμενες δομές οι οποίες προκύπτουν σε κάθε βήμα της προσομοίωσης. Ο υπολογισμός του για ένα σύστημα n ατόμων στις τρεις διαστάσεις (x , y και z) γίνεται με βάση την Εξίσωση 2.18

$$\begin{aligned} RMSD(v, w) &= \sqrt{\frac{1}{n} \sum_{i=1}^n \|v_i - w_i\|^2} \\ &= \sqrt{\frac{1}{n} \sum_{i=1}^n ((v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2)} \quad (2.18) \end{aligned}$$

Όταν προσομοιάζεται ο διμερισμός ενός συστήματος, η τιμή του RMSD αναμένεται να παρουσιάζει σημαντικές διακυμάνσεις, καθώς τα δύο μονομερή με όλα τα άτομά τους θα έρχονται το ένα πιο κοντά στο άλλο.

Η απόσταση των κέντρων μάζας (center of mass distance ή COM distance) των δύο μονομερών είναι άλλη μια ενδεικτική τιμή για την προόδου του συστήματος κατά τη διάρκεια του διμερισμού. Όσο το σύστημα πλησιάζει στη κατάσταση ολοκλήρωσης του διμερισμού, η τιμή της μειώνεται προοδευτικά.

Τόσο η απεικόνιση της τροχιάς του τελικού συστήματος, όσο και οι υπολογισμοί των τιμών του RMSD και της απόστασης των κέντρων μάζας των δύο μονομερών γίνονται στο λογισμικό VMD. Το συγκεκριμένο λογισμικό έχει σχεδιαστεί για την μοντελοποίηση, οπτικοποίηση και ανάλυση βιολογικών συστημάτων όπως πρωτεΐνες, νουκλεϊνικά οξέα, συγκροτήματα λιπιδίων διπλής στιβάδας, κλπ. Μπορεί να χρησιμοποιηθεί για την προβολή γενικότερων μορίων, καθώς το VMD μπορεί να διαβάσει αρχεία δομών από την βάση δεδομένων πρωτεϊνών. Το VMD παρέχει μια πληθώρα μεθόδων για την απόδοση και το χρωματισμό ενός μορίου: απλά σημεία και γραμμές, σφαίρες και κύλινδροι, κυλίνδρους και κορδέλες, καρτούν και άλλες. Τα εργαλεία που περιλαμβάνονται στο VMD παρέχουν τη δυνατότητα απεικόνισης με κίνηση μιας τροχιάς από κάποια προσομοίωσης μοριακής δυναμικής, αλλά και την περαιτέρω ανάλυσή της.

Ένα χρήσιμο εργαλείο του VMD αποτελεί το RMSD trajectory tool, με τη χρήση του οποίου η τροχιά του συστήματος ευθυγραμμίζεται σε σχέση με ένα επιλεγμένο στιγμιότυπο και υπολογίζεται το RMSD κάθε στιγμιότυπου σε σχέση με το επιλεγμένο στιγμιότυπο. Στη παρούσα εργασία η ευθυγράμμιση και ο υπολογισμός του RMSD γίνεται σε σχέση με το πρώτο στιγμιότυπο της τροχιάς. Προκύπτουν έτσι οι τιμές του RMSD σε συνάρτηση με το χρόνο.

Αντίστοιχα, προκειμένου να υπολογιστεί η απόσταση των κέντρων μαζών δύο μονομερών του διμερούς της KRAS-4B σε συνάρτηση με τον χρόνο προσομοίωσης, αξιοποιείται το ελεύθερα διαθέσιμο script distance.tcl. Το συγκεκριμένο script αυτό καλείται στο Tk Console του VMD και δηλώνοντας τις δομές ενδιαφέροντος, εξάγεται το COM distance των δομών αυτών συναρτήσει του χρόνου.

2.4. Αδροτοποιημένα μοντέλα Μοριακής Δυναμικής προσομοίωσης

Οι προσομοιώσεις Μοριακής Δυναμικής αποτελούν πλέον ένα απαραίτητο εργαλείο για τις φυσικές επιστήμες, προσφέροντας μια χωροχρονική ανάλυση που μπορεί να προηγηθεί αλλά και να χρησιμοποιηθεί συμπληρωματικά με κάποια πειραματική τεχνική. Ωστόσο προκύπτει ένα σημαντικό εμπόδιο για τη χρήση της μοριακής δυναμικής, το οποίο είναι η περιορισμένη πρόσβαση σε μεγάλες κλίμακες χρόνου και μήκους, καθώς ένας μεγάλος αριθμός σημαντικών προβλημάτων της δομικής βιολογίας, αλλά και της χημείας και των υλικών λαμβάνουν χώρα σε χρονικό διάστημα ms έως s και αποστάσεων μερικών εκατοντάδων νανομέτρων έως μερικών μικρομέτρων. Αυτές οι κλίμακες χρόνου και μήκους εξακολουθούν να είναι ανέφικτες μέσω μεθόδων κλασική ή και κβαντικής μοριακής

δυναμικής, παρά τη σημαντική πρόοδο των υπολογιστικών συστημάτων και την ανάπτυξη ολοένα και πιο ισχυρού λογισμικού.

Ο περιορισμός που περιγράφεται προκύπτει επειδή το βήμα ολοκλήρωσης dt της Εξίσωσης 2.14 καθορίζεται πρακτικά από την πιο γρήγορη δόνηση του βιομοριακού συστήματος, δηλαδή από τη συχνότητα δόνησης του δεσμού C-H. Η δόνηση του δεσμού αυτού προσδιορίζεται στο 1 fs και επομένως για τα συστήματα στα οποία εφαρμόζεται Μοριακή Δυναμική ισχύει $dt = 1 fs = 10^{-15} s$. Ωστόσο, ένα πολύ μικρό βήμα, αυξάνει τον απαιτούμενο υπολογιστικό χρόνο.

Επομένως, μια προσέγγιση που έχει χρησιμοποιηθεί για να μπορέσουν να μελετηθούν αυτά τα πολύπλοκα προβλήματα είναι να μειωθεί η απαιτητικότητα της προσομοίωσης μειώνοντας τον αριθμό των ατόμων που μοντελοποιούνται και επομένως τους βαθμούς ελευθερίας του προσομοιωμένου συστήματος. Αυτή η διαδικασία μείωσης του αριθμού των ατόμων που μοντελοποιούνται σε ένα σύστημα γίνεται με την ομαδοποίηση των ατόμων μαζί και την αναπαράστασή τους ως μια ενιαίο (ή αδροποιημένο) σφαιρίδιο αλληλεπίδρασης. Η μέθοδος αυτή επιτρέπει την αύξηση του χρονικού διαστήματος dt μέχρι και μερικές δεκάδες fs, λόγω της παράλειψης των κινήσεων υψηλής συχνότητας, όπως οι δονήσεις των ατόμων υδρογόνου. Για το πεδίο δυνάμεων MARTINI 3, το οποίο αξιοποιήθηκε στην εργασία αυτή, χρησιμοποιείται χρονικό διάστημα $dt = 20 fs$. Άρα μπορούν να επιλεγθούν μεγαλύτερα βήματα για τις προσομοιώσεις, επιτρέποντας έτσι μεγαλύτερες προσομοιώσεις και μειώνοντας τον απαιτούμενο υπολογιστικό χρόνο [41].

Το μοντέλο MARTINI με τις διαφορετικές εκδόσεις του είναι ένα από τα πιο δημοφιλή αδροποιημένα μοντέλα στον τομέα της βιομοριακής προσομοίωσης, λόγω της πολύ καλής σύγκρισης που έχουν επειδείξει οι προσομοιώσεις με αυτό το μοντέλο σε σχέση με πειραματικά δεδομένα.

2.4.1. Πεδίο δυνάμεων MARTINI 3

Το MARTINI 3 είναι ένα ελεύθερα διαθέσιμο αδροποιημένο πεδίο δυνάμεων [42], το οποίο έχει αναπτυχθεί στο Πανεπιστήμιο του Groningen και είναι συμβατό με το λογισμικό GROMACS [39] [38], το οποίο χρησιμοποιείται στην παρούσα εργασία. Βασίζεται σε μία λογική αντιστοίχισης 4:1, δηλαδή, σε κατά μέσο όρο τέσσερα βαρέα άτομα (δηλ. όλα τα άτομα εκτός του υδρογόνου) και τα υδρογόνα που σχετίζονται με αυτά αντικαθίστανται και αναπαριστώνται ως ένα αδροποιημένο σφαιρίδιο. Τα σφαιρίδια αυτά έχουν παραμετροποιηθεί χρησιμοποιώντας τα θερμοδυναμικά δεδομένα των επιμέρους στοιχείων που προσομοιάζουν. Οι ασθενείς αλληλεπιδράσεις μεταξύ ουδέτερων σφαιριδίων

περιγράφονται αποκλειστικά από το δυναμικό Lennard-Jones, ενώ τα φορτισμένα σφαιρίδια περιλαμβάνουν επίσης τις αλληλεπιδράσεις Coulomb. Για τη διάκριση μεταξύ διαφορετικών επιπέδων πολικότητας των αδροτοποιημένων σφαιριδίων χρησιμοποιείται η ισχύς της αλληλεπίδρασης του δυναμικού Lennard-Jones. Οι ασθενείς αλληλεπιδράσεις βασίζονται σε πειραματικά δεδομένα ελεύθερων ενεργειών διαχωρισμού μεταξύ πολικών και μη πολικών φάσεων ενός μεγάλου αριθμού χημικών ενώσεων, ενώ οι ισχυρές αλληλεπιδράσεις προκύπτουν από ατομιστικές προσομοιώσεις αναφοράς.

Πιο συγκεκριμένα τα πεδία δυνάμεων MARTINI περιγράφονται από ασθενείς και ισχυρές αλληλεπιδράσεις [41]. Στις ασθενείς αλληλεπιδράσεις περιλαμβάνεται το δυναμικό Lennard-Jones U_{LJ} , όπως περιγράφεται στην Εξίσωση 2.19 και το δυναμικό Coulomb U_{el} , όπως περιγράφεται στην Εξίσωση 2.20.

$$U_{LJ}(r) = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right] \quad (2.19)$$

Όπου:

$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} - s_{ij}$: η μέση απόσταση μεταξύ των σωματιδίων στη θέση της ισορροπίας, όπου η δυναμική ενέργεια είναι μηδέν, εκφρασμένη ως το μέσο όρο της απόστασης του κάθε σωματιδίου από ένα σωματίδιο του ίδιου τύπου, μειωμένο κατά $s_{ij} = 0.01 \text{ nm}$

ε_{ij} : το βάθος του πηγαδιού, υποδηλώνει την ισχύ της αλληλεπίδρασης

$$U_{el}(r) = \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r} \quad (2.20)$$

Όπου:

q_i και q_j : τα φορτία των δύο ατόμων

ε_0 : είναι η διηλεκτρική σταθερά του κενού

ε_r : είναι η σχετική διηλεκτρική σταθερά για τη διαλογή των ατόμων

Οι ισχυρές αλληλεπιδράσεις μεταξύ χημικά συνδεδεμένων σωματιδίων περιλαμβάνουν την περιγραφή των δεσμών από ένα ασθενές αρμονικό δυναμικό V_{bond} (Εξίσωση 2.21), των γωνιών από μία συνάρτηση συνημιτόνου V_{angle} (Εξίσωση 2.22) και για πιο περίπλοκες γεωμετρίες και την αποφυγή στρεβλώσεων εκτός επιπέδου των ακατάλληλων διέδρων χρησιμοποιείται ένα αρμονικό δυναμικό V_{id} (Εξίσωση 2.23).

$$V_{bond}(R) = \frac{1}{2} K_{bond} (R - R_{bond})^2 \quad (2.21)$$

Όπου:

R : είναι η απόσταση μεταξύ των δύο ατόμων.

$R_{bond} = \sigma_{ij}$: η μέση απόσταση μεταξύ των σωματιδίων στη θέση της ισορροπίας, υπολογισμένη όπως στην Εξίσωση 2.19

K_{bond} : η σταθερά της δύναμης που χαρακτηρίζει την αλληλεπίδραση

$$V_{angle}(\theta) = \frac{1}{2}K_{angle}\{\cos(\theta) - \cos(\theta_0)\}^2 \quad (2.22)$$

Όπου:

θ : η γωνία μεταξύ των τριών ατόμων

θ_0 : η γωνία μεταξύ των τριών ατόμων στη θέση της ισορροπίας

K_{angle} : η σταθερά της δύναμης που χαρακτηρίζει τη δόνηση της γωνίας

$$V_{id}(\theta) = K_{id}(\theta - \theta_{id})^2 \quad (2.23)$$

Όπου:

θ : η γωνία μεταξύ του επιπέδου που σχηματίζεται από το κεντρικό άτομο, τα δύο περιφερειακά άτομα και το επίπεδο που σχηματίζεται από τα περιφερειακά άτομα.

θ_{id} : η γωνία μεταξύ των ατόμων στη θέση ισορροπίας

K_{id} : η σταθερά της δύναμης που χαρακτηρίζει τη δόνηση της γωνίας

Επομένως τα πεδία δυνάμεων MARTINI περιγράφονται από την Εξίσωση 2.24, με κάθε μια από τις παραμέτρους της συνάρτησης αυτής να παίρνει διαφορετικές τιμές αναλόγως τον τύπο των σφαιριδίων που χαρακτηρίζει [41].

$$\begin{aligned} V(\mathbf{r}) &= U_{LJ} + U_{el} + V_{bond} + V_{angle} + V_{id} \\ &= 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r} + \frac{1}{2}K_{bond}(R - R_{bond})^2 \\ &\quad + \frac{1}{2}K_{angle}\{\cos(\theta) - \cos(\theta_0)\}^2 + K_{id}(\theta - \theta_{id})^2 \quad (2.24) \end{aligned}$$

Το μοντέλο διαθέτει τέσσερις κύριες κατηγορίες αδροποιημένων σφαιριδίων, που συμβολίζονται με τα γράμματα C, N, P και Q που αντιπροσωπεύουν αντιστοίχως μη πολικές, ενδιάμεσα πολικές, πολικές και φορτισμένες ομάδες. Επιπλέον χρησιμοποιούνται δευτερεύοντες συμβολισμοί προκειμένου να επιτευχθεί μια επιπλέον διάκριση εντός κάθε κατηγορίας ως προς τον βαθμό πολικότητας ή τις δυνατότητες του δότη/δέκτη υδρογόνου.

Κατά βάση, όλες οι σφαίρες έχουν το ίδιο μέγεθος και συμβολίζονται με το γράμμα R από τη λέξη 'regular' ή κανονικός. Κατ' εξαίρεση εισάγονται μικρές σφαίρες, οι οποίες συμβολίζονται με το γράμμα S από τη λέξη small, σε μοντέλα ενώσεων που προσεγγίζουν έναν δακτύλιο. Σε αυτές τις περιπτώσεις, η μέθοδος χαρτογράφησης 4:1 είναι ανεπαρκής. Για την ορθή αναπαράσταση επιστοιβαγμάτων π-π και για τις αποστάσεις των δεσμών υδρογόνου μεταξύ νουκλεοτιδίων, χρησιμοποιούνται ακόμα μικρότερες σφαίρες, οι οποίες συμβολίζονται με το γράμμα T από τη λέξη 'tiny' ή μικροσκοπικός.

Τα πεδία δυνάμεων MARTINI έχουν χρησιμοποιηθεί σε ένα ευρύ φάσμα εφαρμογών σε διαφορετικά πεδία, όπως η δομική βιολογία, η βιοφυσική, η βιοϊατρική, η νανοτεχνολογία και ο σχεδιασμός υλικών. Πιο συγκεκριμένα, η έκδοση MARTINI 3 αντιμετωπίζει τα προβλήματα των έντονων αλληλεπιδράσεων μεταξύ ορισμένων μορίων που ήταν εμφανή στις προηγούμενες εκδόσεις, επιτρέποντας πιο ακριβείς προσομοιώσεις μοριακών συστημάτων [42].

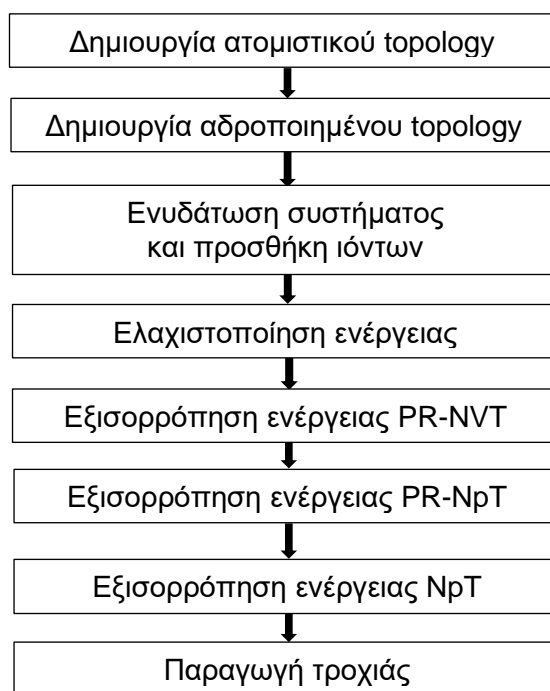
2.4.2. Ροή εργασίας martiniate

Προκειμένου να πραγματοποιηθούν σταδιακά στη σειρά οι διαδικασίες της Μοριακής Δυναμικής που περιγράφονται στην προηγούμενη παράγραφο, δηλαδή η προετοιμασία του συστήματος που προσομοιάζεται, η ελαχιστοποίηση της ενέργειας, η θέρμανση και η εξισορρόπηση του συστήματος και η παραγωγή της τροχιάς, χρησιμοποιείται η ροή εργασίας martiniate [43].

Πρόκειται για ένα shell script, στο οποίο περιέχονται οι εντολές που απαιτούνται για τα διαδοχικά βήματα της Μοριακής Δυναμικής. Αυτά με τη σειρά είναι η δημιουργία του ατομιστικού και του αδροτοποιημένου αρχείου topology από την αρχική δομή, η ενυδάτωση του συστήματος και η προσθήκη των απαραίτητων ιόντων, η ελαχιστοποίηση της ενέργειας του συστήματος, η εξισορρόπηση του συστήματος υπό σταθερό αριθμό ατόμων, όγκο και θερμοκρασία, ενώ παράλληλα εφαρμόζεται ένα αρμονικό δυναμικό ώστε τα άτομα να συγκρατούνται στις θέσεις που κατέλαβαν στο βήμα της ελαχιστοποίησης ενέργειας (Position Restrained NVT ή PR-NVT), η εξισορρόπηση του συστήματος υπό σταθερό αριθμό ατόμων, πίεση και θερμοκρασία και εφαρμόζοντας ένα αρμονικό δυναμικό, ώστε τα άτομα να συγκρατούνται στις θέσεις που κατέλαβαν στο προηγούμενο βήμα της εξισορρόπησης ενέργειας (Position Restrained NpT ή PR-NpT), η εξισορρόπηση του συστήματος υπό τις συνθήκες της τελικής προσομοίωσης υπό σταθερό αριθμό ατόμων, πίεση και θερμοκρασία αλλά χωρίς την συγκράτηση των ατόμων σε συγκεκριμένες θέσεις (Unrestrained NpT ή NpT),

και τέλος παραγωγή της τροχιάς του συστήματος. Τα βήματα αυτά συνοψίζονται στο σχήμα 2.5, υπό τη μορφή διαγράμματος ροής.

Σημειώνεται ότι για την ενυδάτωση του συστήματος και τη προσθήκη ιόντων γίνεται χρήση του εργαλείου INSANE (INSert membraNE), το οποίο είναι ένα python script που χρησιμοποιείται κυρίως για την δημιουργία μεμβρανών σε αδροποιημένα συστήματα.



Σχήμα 2.5: Διάγραμμα ροής της ροής εργασίας martinate.

2.5. Ομαδοποίηση της τροχιάς

Η γεωμετρική ομαδοποίηση πραγματοποιείται για τον εντοπισμό παρόμοιων διαμορφώσεων που αποτελούν στιγμιότυπα μιας τροχιάς Μοριακής Δυναμικής. Υπάρχουν διάφορες μέθοδοι ομαδοποίησης. Ένας από τους πιο συχνά χρησιμοποιούμενους αλγορίθμους ομαδοποίησης που εφαρμόζονται σε τροχίες Μοριακής Δυναμικής είναι το gromos και εκτελείται με την εντολή gmx cluster στο GROMACS. Η μέθοδος αυτή βασίζεται στην κοινή ακτίνα αποκοπής του RMSD μεταξύ των διαμορφώσεων της τροχιάς Μοριακής Δυναμικής.

Αρχικά επιλέγεται μια ακτίνα αποκοπής RMSD, με βάση την οποία θα γίνει η ομαδοποίηση σε συστάδες. Στη συνέχεια υπολογίζεται η τιμή του RMSD για κάθε ζεύγος διαμορφώσεων, σε όλη την έκταση της τροχιάς. Αφού ολοκληρωθεί αυτό το βήμα, μετράται ο αριθμός των γειτόνων της κάθε διαμόρφωσης που βρίσκονται εντός της ακτίνας αποκοπής και προσδιορίζεται η διαμόρφωση με τον μεγαλύτερο αριθμό γειτόνων, δηλαδή, άλλες

διαμορφώσεις, ως το κέντρο της συστάδας. Η διαμόρφωση αυτή, όπως και οι γειτονικές διαμορφώσεις που ομαδοποιήθηκαν σε μία συστάδα, αποκλείονται από την αρχική συλλογή στιγμιότυπων. Η διαδικασία αυτή επαναλαμβάνεται για τις υπόλοιπες διαμορφώσεις που έχουν απομείνει από την αρχική συλλογή, μέχρι να αντιστοιχιστούν όλα τα στιγμιότυπα σε μία συστάδα.

Όλες οι συστάδες είναι αμοιβαία αποκλειστικές, επομένως μια διαμόρφωση μπορεί να είναι μέλος μόνο ενός συμπλέγματος. Η διαμόρφωση με τον μεγαλύτερο αριθμό γειτόνων σε καθένα από τα συμπλέγματα χαρακτηρίζεται ως η αντιπροσωπευτική διαμόρφωση του συμπλέγματος, δηλαδή, το κέντρο της συστάδας. Οποιαδήποτε διαμόρφωση μείνει χωρίς γείτονες αντιπροσωπεύει τη δική της μονομελή συστάδα [38] [44].

3. Αποτελέσματα και συζήτηση αποτελεσμάτων

3.1. Μετατροπή ατομιστικού μοντέλου διμερούς KRAS-4B σε αδροποιημένο μοντέλο

Για τη παρούσα εργασία χρησιμοποιήθηκε το διμερές της KRAS-4B σε δύο μορφές: τον φυσικό τύπο (WT) και την μετάλλαξη G12D. Η αρχική δομή που χρησιμοποιήθηκε με PDB ID 5VQ2 είναι η κρυσταλλική δομή της ανθρώπινης KRAS-4B φυσικού τύπου σε σύμπλεγμα με το μόριο GTP και παρουσιάζει διεπιφάνεια διμερισμού α5–α5 [25]. Το αρχικό ατομιστικό σύστημα είχε προηγουμένως αναπτυχθεί στο πλαίσιο της ερευνητικής εργασίας του εργαστηρίου και η μετάλλαξη δημιουργήθηκε αντικαθιστώντας την γλυκίνη της θέσης 12 με ασπαρτικό οξύ και πραγματοποιώντας μία τοπική ελαχιστοποίηση ενέργειας. Η περιγραφή που ακολουθεί θα αναφέρεται στην φυσιολογική μορφή αλλά εφεξής θα εννοείται ότι τα ίδια ισχύουν και για την μετάλλαξη G12D.

Το συγκεκριμένο σύστημα αποτελείται από το διμερές της πρωτεΐνης KRAS-4B σε μεμβράνη, μέσα σε νερό με ιόντα. Για τις ανάγκες της εργασίας αυτής, αφαιρούνται από το σύστημα τα άτομα που αντιστοιχούν στα λιπίδια της μεμβράνης και το νερό με τα ιόντα, ώστε να προστεθούν αργότερα στις σωστές αναλογίες.

Τα παρακάτω βήματα που περιγράφονται πραγματοποιούνται από τη ροή εργασίας Martinate που αναφέρθηκε προηγουμένως.

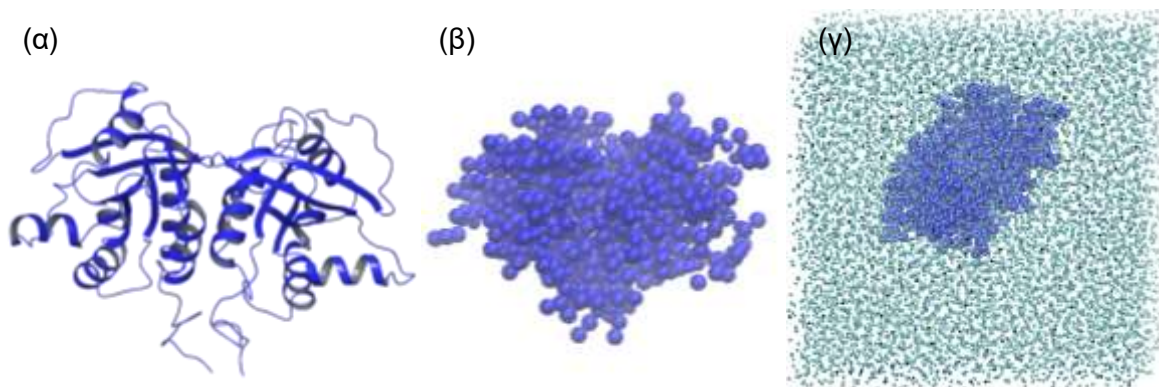
3.1.1. Δημιουργία αδροποιημένης δομής και αρχείου topology

Για τη δημιουργία της αδροποιημένης δομής χρησιμοποιείται η εντολή `pdb2gmx`, η οποία μετατρέπει το ατομιστικό αρχείο `.pdb`, που δίνεται ως `input` (Σχήμα 3.1 (α)), σε αρχείο της μορφής `.gmx`, το οποίο αναγνωρίζεται και αξιοποιείται από το GROMACS. Παράγει επίσης και ένα αρχείο `topology (.top)` το οποίο περιέχει μια πλήρη περιγραφή όλων των αλληλεπιδράσεων της πρωτεΐνης.

Για τη δημιουργία της αδροποιημένης δομής του διμερούς της KRAS-4B (Σχήμα 3.1 (β)), καλείται το `python script martinize`, στόχος του οποίου είναι μετατροπή μιας ατομιστικής δομής σε αδροποιημένη. Το `martinize` έχει αναπτυχθεί από τους δημιουργούς του MARTINI.

3.1.2. Ενυδάτωση συστήματος και προσθήκη ιόντων

Για την ενυδάτωση του συστήματος και την προσθήκη ιόντων χρησιμοποιείται το πρόγραμμα INSANE (INSert membrANE). Το πρόγραμμα INSANE έχει αναπτυχθεί για να εισάγει αποτελεσματικά μεμβράνες από διαφορετικά λιπίδια σε αδροποιημένα συστήματα. Στη παρούσα εργασία το πρόγραμμα αξιοποιείται για την δημιουργία ενός κυβικού κουτιού ακμής 5 nm, το οποίο ενυδατώνεται με νερό και προστίθενται ιόντα έως ότου να επιτευχθεί συγκέντρωση 0.15 M, η οποία αντιστοιχεί στην φυσιολογική συγκέντρωση ενός κυττάρου, όπως απεικονίζεται στο Σχήμα 3.1 (γ).



Σχήμα 3.1: Απεικόνιση (α) της ατομιστικής δομής του διμερούς της KRAS-4B, (β) της αδροποιημένης δομής του διμερούς της KRAS-4B και (γ) της αδροποιημένης δομής του διμερούς της KRAS-4B στο ενυδατωμένο σύστημα με τα ιόντα.

3.2. Προσομοιώσεις Μοριακής Δυναμικής στο σύστημα διμερούς KRAS-4B σε διάλυμα

Μετά την προετοιμασία και την ενυδάτωση του συστήματος απαιτείται μια σειρά βημάτων μέχρι να προκύψει η τελική τροχιά του συστήματος και να αναλυθούν οι επιθυμητές ιδιότητες, όπως αυτά που περιγράφηκαν στις Μεθόδους. Τα βήματα αυτά είναι επίσης ενσωματωμένα στη ροή εργασίας Martinate. Πρακτικά, για κάθε βήμα, αφού τεθούν οι επιθυμητές συνθήκες, πραγματοποιούνται δύο εντολές: το `gmx grompp` και το `gmx mdrun`. Η πρώτη εντολή ενημερώνει το αρχείο `topology` και η δεύτερη εκτελεί το βήμα της ελαχιστοποίησης ενέργειας, θερμανσης, εξισορρόπησης, μοριακής δυναμικής ή ό,τι απαιτείται στο συγκεκριμένο βήμα.

3.2.1. Ελαχιστοποίηση της ενέργειας στο σύστημα διμερούς KRAS-4B

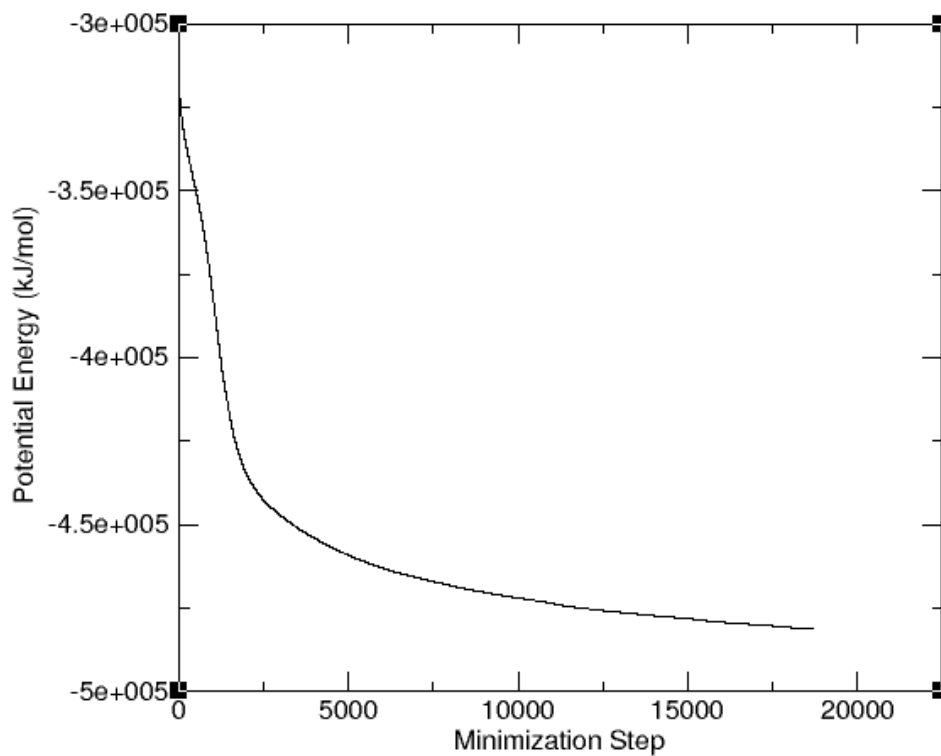
Μετά την ενυδάτωση του συστήματος απαιτείται να γίνει ελαχιστοποίηση της ενέργειας του συστήματος, προκειμένου να καταργηθούν τυχόν δυσμενείς αλληλεπιδράσεις van der Waals που μπορεί να έχουν σχηματιστεί. Εάν δεν πραγματοποιηθεί η ελαχιστοποίηση της ενέργειας, το διμερές ενδέχεται να παραμορφωθεί τοπικά και επομένως η προσομοίωση να καταστεί ασταθής.

Στον φυσικό τύπο επιτυγχάνεται ελαχιστοποίηση της ενέργειας σε περίπου 19 χιλιάδες βήματα, ενώ στην μετάλλαξη σε περίπου 21 χιλιάδες βήματα. Στο σχήμα 3.2 (α) και (β) απεικονίζονται τα διαγράμματα της δυναμικής ενέργειας σε συνάρτηση με τον χρόνο προσομοίωσης για τον φυσικό τύπο και τη μετάλλαξη G12D αντίστοιχα. Η τιμή της δυναμικής ενέργειας για στην οποία συγκλίνει το σύστημα του φυσικού τύπου είναι $-4.81 \times 10^5 \text{ kJ/mol}$, ενώ για το σύστημα της μετάλλαξης G12D είναι $-4.06 \times 10^5 \text{ kJ/mol}$.

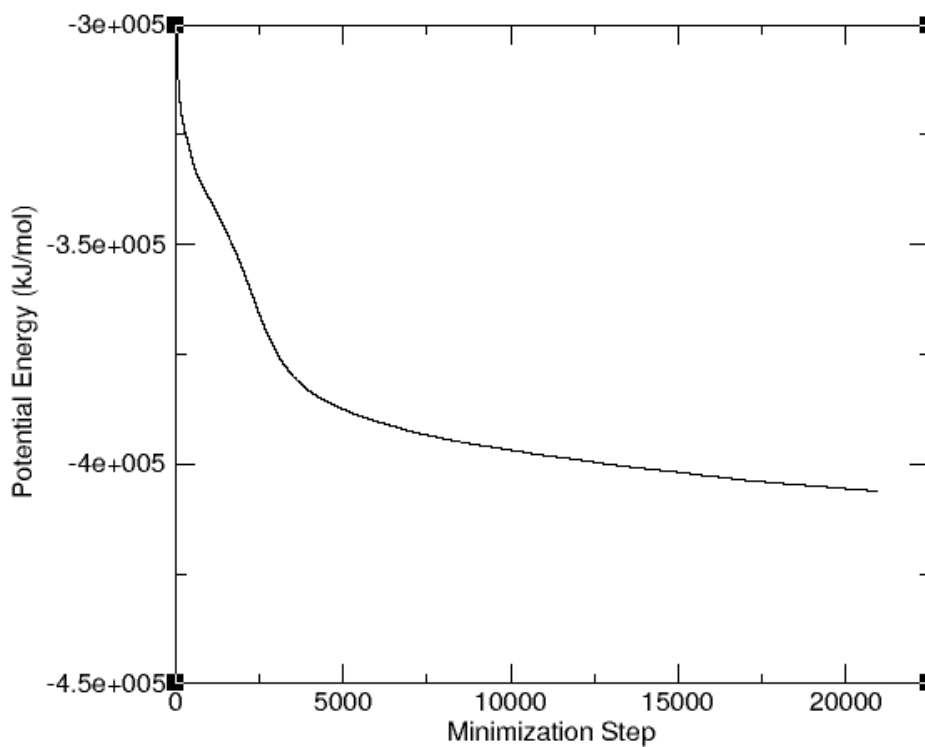
Παρατηρείται ότι η μεγαλύτερη βαθμίδα μείωσης της δυναμικής ενέργειας παρουσιάζεται στα πρώτα 5000 βήματα και για τα δύο συστήματα. Ο λόγος για τον οποίο συμβαίνει αυτό είναι επειδή η κρυσταλλική δομή έχει κάποιες ανεπιθύμητες αλληλεπιδράσεις υψηλής ενέργειας. Καθώς ο στόχος του βήματος αυτού είναι η ελαχιστοποίηση αυτών των αλληλεπιδράσεων και η εύρεση ενός τοπικού ελάχιστου στο ενεργειακό τοπίο του συστήματος, δηλαδή, το μέρος στο οποίο το μόριο χαλαρώνει, οι θέσεις των ατόμων μεταβάλλονται συστηματικά και υπολογίζεται η δυναμική ενέργεια. Είναι λοιπόν αναμενόμενο στα πρώτα βήματα κατά τα οποία μεταβάλλονται οι θέσεις των ατόμων να υπάρχει μεγάλη διαφορά δυναμικής ενέργεια, καθώς καταργούνται οι αλληλεπιδράσεις υψηλής ενέργειας.

Ο λόγος για τον οποίο απαιτούνται περισσότερα βήματα για την ελαχιστοποίηση ενέργειας του συστήματος της μετάλλαξης G12D είναι επειδή εισήχθη το αμινοξύ του ασπαρτικού οξέος στη θέση 12, αντικαθιστώντας την γλυκίνη που υπήρχε εκεί αρχικά. Η αντικατάσταση αυτή δημιουργεί επιπλέον αλληλεπιδράσεις μεταξύ των αμινοξέων σε εκείνο το δομικό τμήμα της πρωτεΐνης.

(α) Potential Energy vs Step for KRAS-4B WT Energy Minimization



(β) Potential Energy vs Step for KRAS-4B G12D Energy Minimization



Σχήμα 3.2: Διαγράμματα δυναμικής ενέργειας συναρτήσει του χρόνου για το βήμα της ελαχιστοποίησης ενέργειας στο σύστημα του διμερούς της KRAS-4B (α) φυσικού τύπου και (β) μετάλλαξη G12D.

3.2.2. Εξισορρόπηση του συστήματος

Η εξισορρόπηση του συστήματος γίνεται σε τρία στάδια και οι τιμές των ιδιοτήτων του συστήματος αποθηκεύονται κάθε 8 βήματα:

1. Στο κανονικό θερμοδυναμικό σύνολο και εφαρμόζοντας ένα αρμονικό δυναμικό, ώστε τα άτομα να συγκρατούνται στις θέσεις τους από το βήμα της ελαχιστοποίησης ενέργειας (Position Restrained NVT ή PR-NVT). Το σύστημα μοντελοποιείται για 50000 βήματα με $dt = 2 fs$, δηλαδή συνολικά 100 ps.
2. Στο ισόθερμο - ισοβαρές θερμοδυναμικό σύνολο και εφαρμόζοντας ένα αρμονικό δυναμικό, ώστε τα άτομα να συγκρατούνται στις θέσεις τους από το προηγούμενο βήμα της εξισορρόπησης ενέργειας (Position Restrained NpT ή PR-NpT). Το σύστημα μοντελοποιείται για 50000 βήματα με $dt = 20 fs$, δηλαδή συνολικά 1000 ps.
3. Στο ισόθερμο - ισοβαρές θερμοδυναμικό σύνολο αλλά χωρίς την συγκράτηση των ατόμων σε συγκεκριμένες θέσεις (Unrestrained NpT ή NpT). Οι συνθήκες αυτές αποτελούν και τις συνθήκες στις οποίες θα πραγματοποιηθεί και η παραγωγή τροχιάς, επομένως το στάδιο αυτό αποτελεί την προετοιμασία της παραγωγής τροχιάς. Το σύστημα μοντελοποιείται για 50000 βήματα με $dt = 20 fs$, δηλαδή συνολικά 1000 ps.

Στο σχήμα 3.3 (α) και (β) απεικονίζονται τα διαγράμματα της δυναμικής ενέργειας σε συνάρτηση με τον χρόνο προσομοίωσης κατά την εξισορρόπηση ενέργειας στο κανονικό θερμοδυναμικό σύνολο και με σταθερές τις θέσεις των ατόμων στο σύστημα (στάδιο PR-NVT) για τον φυσικό τύπο και τη μετάλλαξη G12D αντίστοιχα. Στο κανονικό θερμοδυναμικό σύνολο διατηρείται σταθερός ο αριθμός ατόμων (N) του συστήματος, καθώς και ο όγκος του (V) και η θερμοκρασία του (T).

Και στα δύο συστήματα παρατηρείται μία αρχική αύξηση της δυναμικής ενέργειας του συστήματος, ενώ στη συνέχεια η δυναμική ενέργεια συγκλίνει με μικρές διακυμάνσεις. Η τιμή της δυναμικής ενέργειας για την οποία συγκλίνει το σύστημα του φυσικού τύπου είναι $-4.04 \times 10^5 kJ/mol$, ενώ για το σύστημα της μετάλλαξης G12D είναι $-3.41 \times 10^5 kJ/mol$.

Η αρχική αύξηση της δυναμικής ενέργειας οφείλεται στη σταθεροποίηση των θέσεων των ατόμων του διμερούς κατά τη διάρκεια της προσομοίωσης. Η σταθεροποίηση αυτή επιτυγχάνεται περιορίζοντας την κίνηση των ατόμων χωρίς όμως να τα ακινητοποιεί ολοκληρωτικά. Η σύγκλιση αυτή οδηγεί στο συμπέρασμα ότι η εξισορρόπηση υπό αυτές τις συγκεκριμένες συνθήκες ήταν επιτυχής και το κάθε ένα από τα δύο συστήματα μπορούν να περάσουν στο επόμενο στάδιο της εξισορρόπησης.

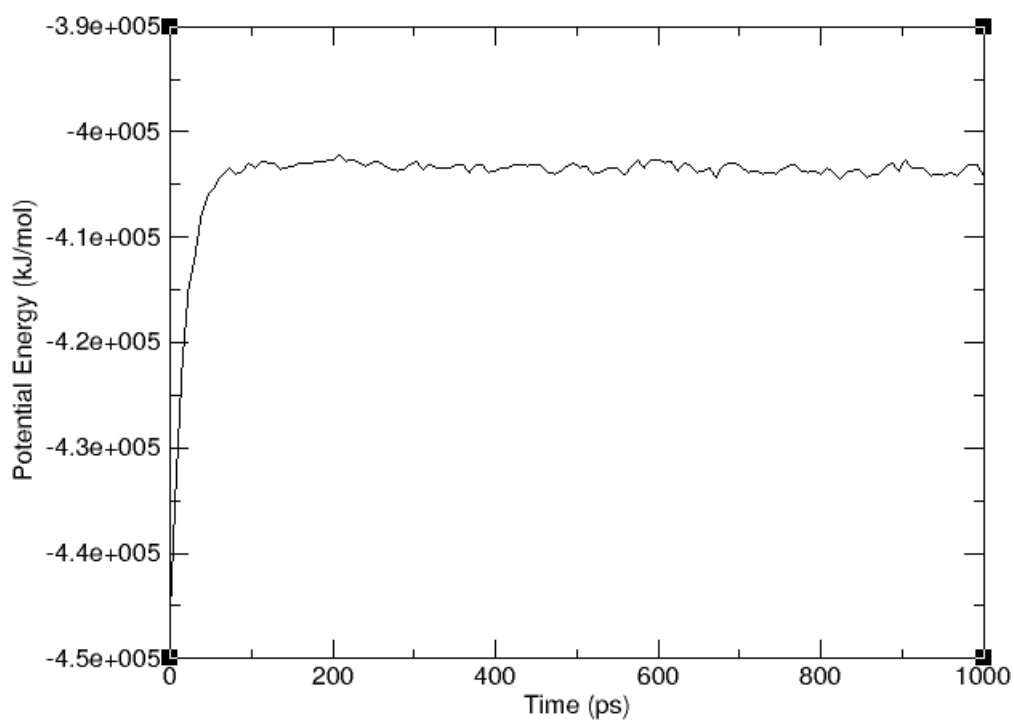
Στο σχήμα 3.4 (α) και (β) απεικονίζονται τα διαγράμματα της δυναμικής ενέργειας σε συνάρτηση με τον χρόνο προσομοίωσης κατά την εξισορρόπηση ενέργειας στο ισόθερμο – ισοβαρές θερμοδυναμικό σύνολο και με σταθερές τις θέσεις των ατόμων στο σύστημα (στάδιο PR-NpT) για τον φυσικό τύπο και τη μετάλλαξη G12D αντίστοιχα. Στο ισόθερμο – ισοβαρές θερμοδυναμικό σύνολο διατηρείται σταθερός ο αριθμός ατόμων (N) του συστήματος, καθώς και η πίεσή του (p) και η θερμοκρασία του (T).

Σε αμφότερα τα συστήματα παρατηρείται ότι η δυναμική ενέργεια παραμένει σταθερή καθ' όλη τη διάρκεια της προσομοίωσης, ενώ παρατηρούνται διακυμάνσεις της τάξεως του 0.25-0.5% γύρω από την τιμή στην οποία ισορροπεί το κάθε σύστημα. Το αποτέλεσμα αυτό είναι αναμενόμενο καθώς η δυναμική ενέργεια του συστήματος είχε ήδη σταθεροποιηθεί από το προηγούμενο στάδιο της εξισορρόπησης, διατηρώντας σταθερές τις θέσεις των ατόμων. Η τιμή της δυναμικής ενέργειας στην οποία έχει σταθεροποιηθεί το σύστημα του φυσικού τύπου είναι $-3.97 \times 10^5 \text{ kJ/mol}$, ενώ για το σύστημα της μετάλλαξης G12D είναι $-3.53 \times 10^5 \text{ kJ/mol}$. Αξίζει να σημειωθεί ότι σε σύγκριση με τις τιμές της δυναμικής ενέργειας στις οποίες συνέκλιναν τα συστήματα στο πρώτο στάδιο PR-NVT της εξισορρόπησης, η δυναμική ενέργεια ενέργεια του συστήματος του φυσικού τύπου αυξάνεται στο δεύτερο στάδιο PR-NpT, ενώ η αντίστοιχη για το σύστημα της μετάλλαξης μειώνεται.

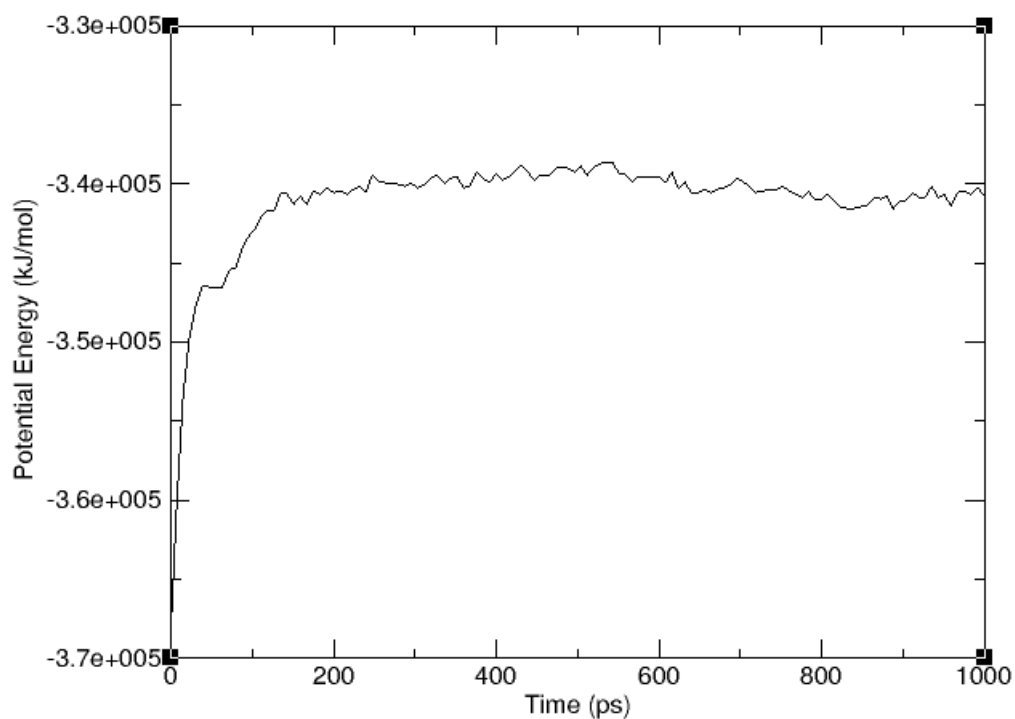
Στο σχήμα 3.5 (α) και (β) απεικονίζονται τα διαγράμματα της δυναμικής ενέργειας σε συνάρτηση με τον χρόνο προσομοίωσης κατά την εξισορρόπηση ενέργειας στο ισόθερμο – ισοβαρές θερμοδυναμικό σύνολο, χωρίς κάποια εφαρμογή δυνάμεων στα άτομα του συστήματος (στάδιο NpT) για τον φυσικό τύπο και τη μετάλλαξη G12D αντίστοιχα.

Όπως και πριν, παρατηρείται και στα δύο συστήματα ότι η δυναμική ενέργεια παραμένει σταθερή κατά τη διάρκεια της προσομοίωσης, με διακυμάνσεις της τάξεως του 0.25-0.5% γύρω από την τιμή στην οποία ισορροπεί το κάθε σύστημα. Το αποτέλεσμα αυτό είναι αναμενόμενο καθώς η δυναμική ενέργεια του συστήματος είχε ήδη σταθεροποιηθεί από το προηγούμενο στάδιο της εξισορρόπησης, διατηρώντας σταθερές τις θέσεις των ατόμων. Η τιμή της δυναμικής ενέργειας στην οποία έχει σταθεροποιηθεί το σύστημα του φυσικού τύπου είναι $-3.97 \times 10^5 \text{ kJ/mol}$, ενώ για το σύστημα της μετάλλαξης G12D είναι $-3.53 \times 10^5 \text{ kJ/mol}$. Σημειώνεται ότι οι τιμές αυτές είναι ίδιες με τις τιμές στις οποίες είχαν σταθεροποιηθεί τα συστήματα στο προηγούμενο στάδιο PR-NpT και επομένως επιβεβαιώνεται ότι έχει επιτευχθεί εξισορρόπηση του καθενός από τα δύο συστήματα.

(α) Potential Energy vs Time for KRAS-4B WT PR-NVT

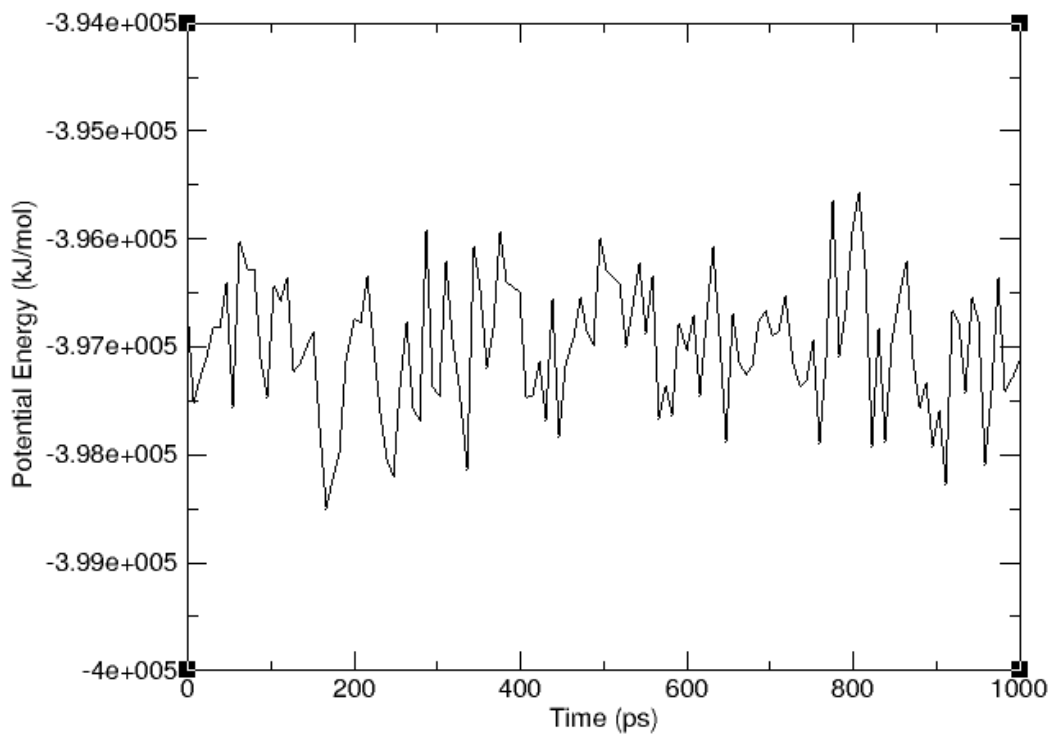


(β) Potential Energy vs Time for KRAS-4B G12D PR-NVT

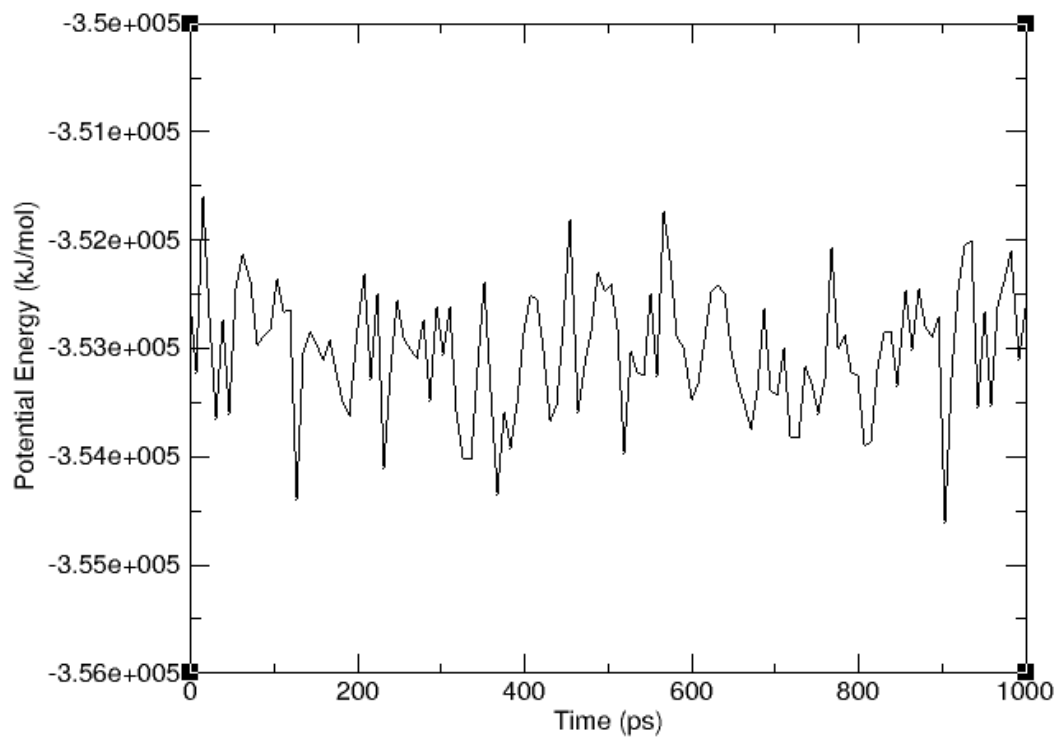


Σχήμα 3.3: Διαγράμματα δυναμικής ενέργειας συναρτήσει του χρόνου για το βήμα της εξισορρόπησης ενέργειας στο κανονικό (NVT) θερμοδυναμικό σύνολο με τα άτομα σε σταθερές θέσεις στο σύστημα του διμερούς της KRAS-4B (α) φυσικού τύπου και (β) μετάλλαξη G12D.

(α) Potential Energy vs Time for KRAS-4B WT PR-NpT

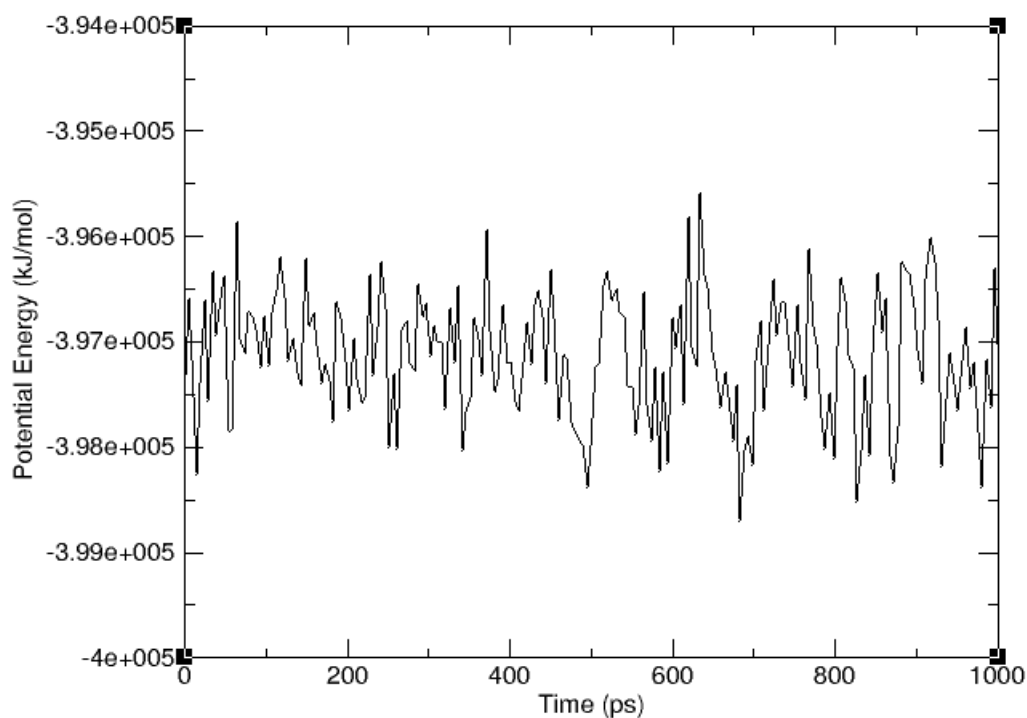


(β) Potential Energy vs Time for KRAS-4B G12D PR-NpT

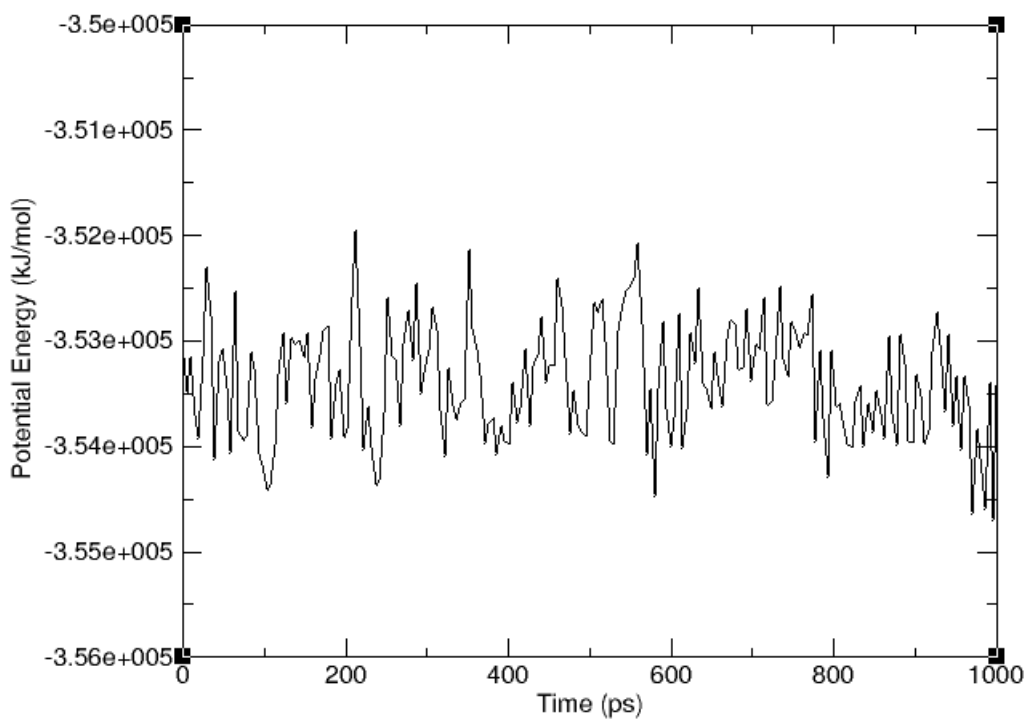


Σχήμα 3.4: Διαγράμματα δυναμικής ενέργειας συναρτήσει του χρόνου για το βήμα της εξισορρόπησης ενέργειας στο κανονικό (NpT) θερμοδυναμικό σύνολο με τα άτομα σε σταθερές θέσεις στο σύστημα του διμερούς της KRAS-4B (α) φυσικού τύπου και (β) μετάλλαξη G12D.

(α) Potential Energy vs Time for KRAS-4B WT NpT



(β) Potential Energy vs Time for KRAS-4B G12D NpT



Σχήμα 3.5: Διαγράμματα δυναμικής ενέργειας συναρτήσει του χρόνου για το βήμα της εξισορρόπησης ενέργειας στο κανονικό (NpT) θερμοδυναμικό σύνολο με ελεύθερα άτομα στο σύστημα του διμερούς της KRAS-4B (α) φυσικού τύπου και (β) μετάλλαξη G12D.

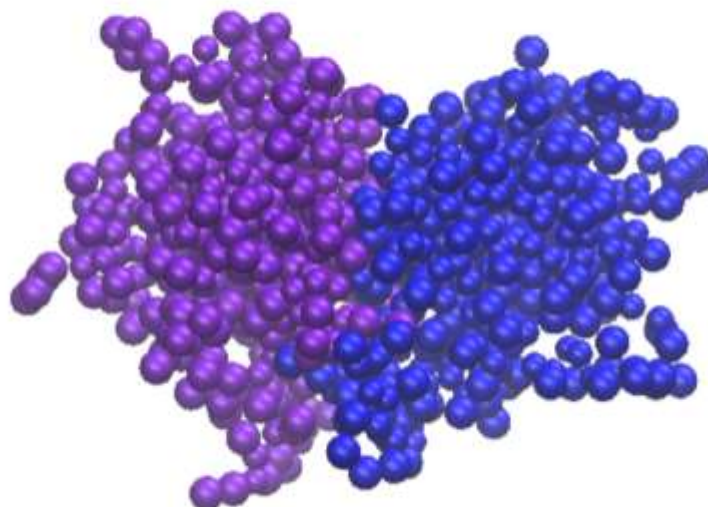
3.2.3. Παραγωγή τροχιάς και ανάλυση ιδιοτήτων

Το τελευταίο βήμα για την ολοκλήρωση της Μοριακής Δυναμικής είναι η παραγωγή της τροχιάς και αποτελεί τον λόγο για τον οποίο πραγματοποιούνται τα προηγούμενα βήματα, ώστε το σύστημα να φτάσει στη κατάλληλη κατάσταση για να μοντελοποιηθεί. Το συγκεκριμένο βήμα μοντελοποιείται για 5000000 βήματα και $dt = 20 fs$, δηλαδή συνολικός χρόνος προσομοίωσης 1 μs . Καθώς για το συγκεκριμένο βήμα εκτελούνται πολλά βήματα, οι τιμές των ιδιοτήτων για τα δύο συστήματα που επιλέγεται να αποθηκευτούν λαμβάνονται κάθε 1 ns, ώστε τόσο η τροχιά, όσο και τα δεδομένα για τις ενέργειες του κάθε συστήματος να είναι πιο εύκολα διαχειρίσιμες. Ο χρόνος προσομοίωσης (1 μs) επιτυγχάνεται σε διάστημα περίπου 24 ωρών σε σταθερό υπολογιστή. Σημειώνεται ότι για να επιτευχθεί ο ίδιος υπολογιστικός χρόνος για αυτό το σύστημα και με την ίδια υπολογιστική ισχύ, θα απαιτούνταν περισσότεροι από τρεις μήνες.

Από αυτό το βήμα προκύπτει ένα δυαδικό αρχείο .trr, το οποίο μετατρέπεται σε αρχείο .xtc με την εντολή του GROMACS `gmx trjconv`. Στη συνέχεια η τροχιά επεξεργάζεται, ώστε η εικόνα που θα απεικονιστεί να είναι συνεχής. Αντίστοιχα, χρησιμοποιώντας την ίδια εντολή και αξιοποιώντας τη δυνατότητα `-pbc mol -center`, το κέντρο μάζας της πρωτεΐνης κεντράρεται στο κουτί της προσομοίωσης. Χρησιμοποιώντας την δυνατότητα `-pbc nojump`, αποφεύγεται το σπάσιμο των πρωτεϊνών σε περίπτωση που κάποιο μέρος τους περάσει τα σύνορα του κουτιού και η αναπλήρωσή τους μέσω των περιοδικών οριακών συνθηκών. Τέλος, χρησιμοποιώντας τη δυνατότητα `-fit rot+trans` εξασφαλίζεται ότι η πρωτεΐνη δεν θα περιστρέφεται μέσα στο κουτί, ώστε να μην προσεγγιστούν τα σύνορα του κουτιού.

Το τελικό αρχείο .xtc που προκύπτει αντιστοιχεί στην τροχιά που εξήχθη από το τελικό βήμα της Μοριακής Δυναμικής. Επομένως αυτό το αρχείο εισάγεται στο πρόγραμμα VMD πάνω στο αρχείο .gro, στο οποίο απεικονίζεται το σύστημα. Ο συνδυασμός των δύο αρχείων .xtc και .gro μπορεί να αναπαραστήσει την πλήρη τροχιά του συστήματος στην εξέλιξη του χρόνου. Στο σχήμα 3.6 απεικονίζεται ένα στιγμιότυπο από την τροχιά του συστήματος, στο οποίο όμως είναι ορατό μονάχα το διμερές της KRAS-4B, με κάθε μονομερές να είναι χρωματισμένο διαφορετικά και τα μόρια ιόντα και νερού να αποκρύπτονται για απλούστευση.

Όπως και προηγουμένως, παρακολουθείται η δυναμική ενέργεια συναρτήσει του χρόνου της προσομοίωσης, προκειμένου να επιβεβαιωθεί ότι το σύστημα παραμένει εξισορροπημένο. Ωστόσο, οι ιδιότητες ενδιαφέροντος που παρακολουθούνται είναι η ρίζα μέσης τετραγωνικής απόκλισης σε συνάρτηση με το χρόνο και η απόσταση των κέντρων μαζών των δύο μονομερών KRAS-4B σε διάλυμα. Η μεταβολή αυτών των μεγεθών έχει βρεθεί ότι είναι πραγματικά αντιπροσωπευτική της εξέλιξης του διμερισμού ενός συστήματος πρωτεΐνης της οικογένειας RAS [45].

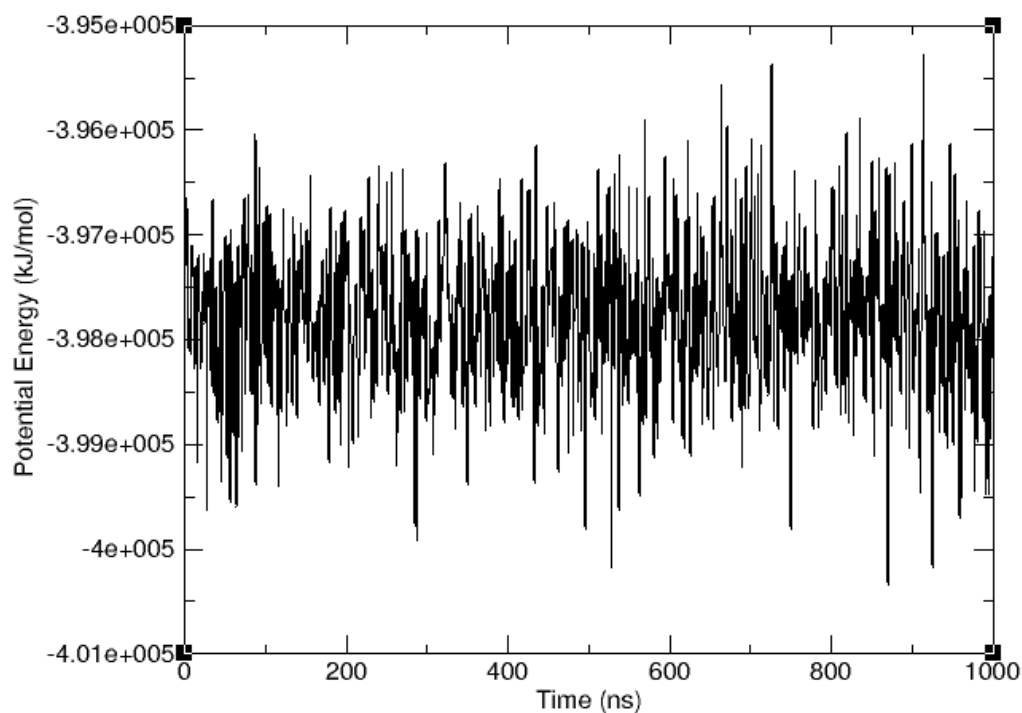


Σχήμα 3.6: Στιγμιότυπο από την τροχιά του συστήματος, όπου απεικονίζεται το διμερές της KRAS-4B με διαφορετικά χρωματισμένα μονομερή και αποκρύπτονται τα μόρια ιόντων και νερού.

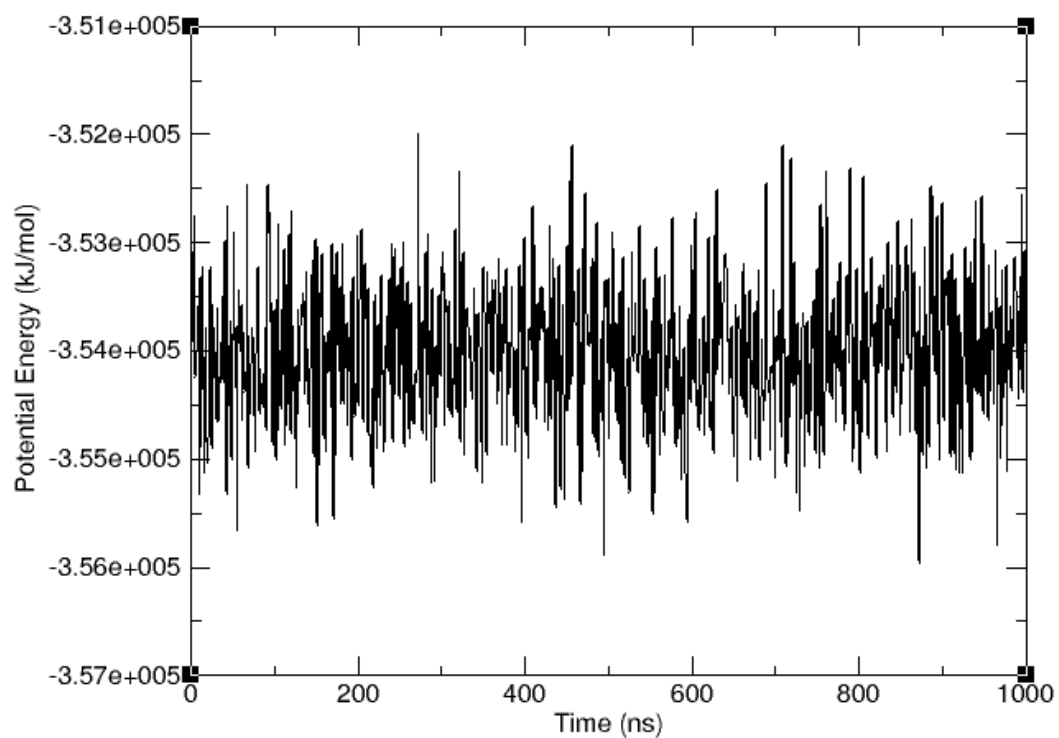
3.2.3.1. Υπολογισμός θερμοδυναμικών ιδιοτήτων

Στο σχήμα 3.7 (α) και (β) απεικονίζονται τα διαγράμματα της δυναμικής ενέργειας σε συνάρτηση με τον χρόνο προσομοίωσης κατά το στάδιο της παραγωγής τροχιάς για τον φυσικό τύπο και τη μετάλλαξη G12D αντίστοιχα. Παρατηρείται και στα δύο συστήματα ότι η δυναμική ενέργεια παραμένει σταθερή κατά τη διάρκεια της προσομοίωσης, όπως αναμένεται μετά την εξισορρόπηση ενέργειας. Η τιμή της δυναμικής ενέργειας στην οποία έχει σταθεροποιηθεί το σύστημα του φυσικού τύπου είναι $-3.98 \times 10^5 \text{ kJ/mol}$, ενώ για το σύστημα της μετάλλαξης G12D είναι $-3.54 \times 10^5 \text{ kJ/mol}$, τιμές αντίστοιχες με αυτές του βήματος της εξισορρόπησης.

(α) Potential Energy vs Time for KRAS-4B WT Production Run



(β) Potential Energy vs Time for KRAS-4B G12D Production Run



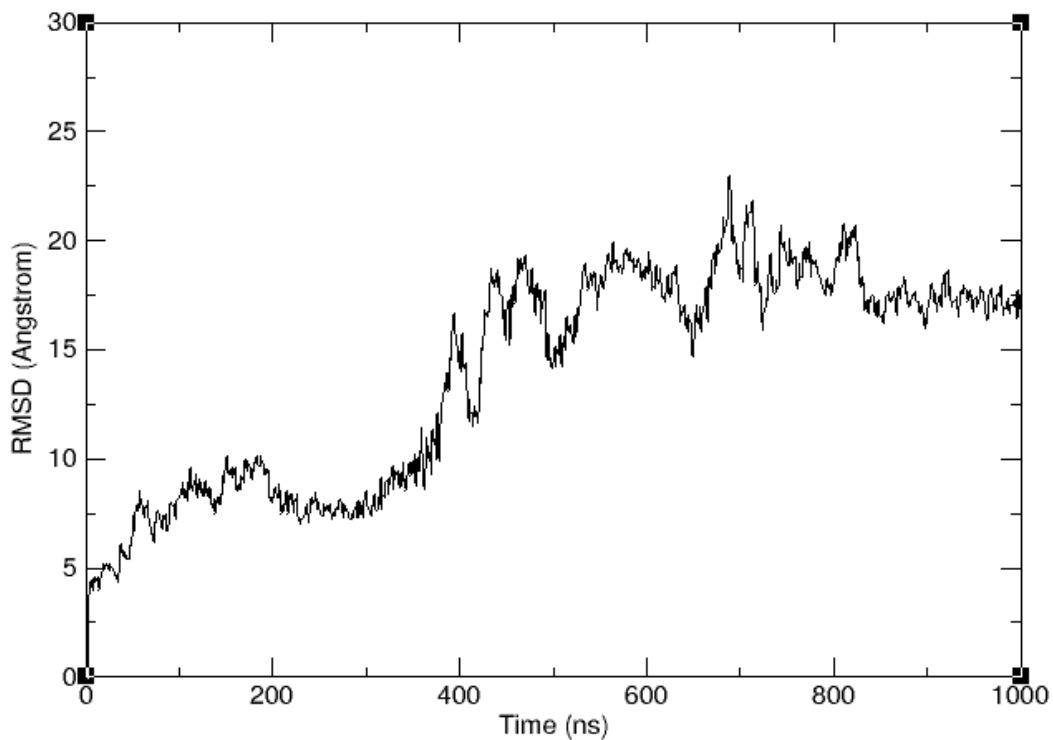
Σχήμα 3.7: Διαγράμματα δυναμικής ενέργειας συναρτήσει του χρόνου για το βήμα της παραγωγής τροχιάς στο σύστημα του διμερούς της KRAS-4B (α) φυσικού τύπου και (β) μετάλλαξη G12D.

3.2.3.2. Ρίζα μέσης τετραγωνικής απόκλισης σε συνάρτηση με το χρόνο

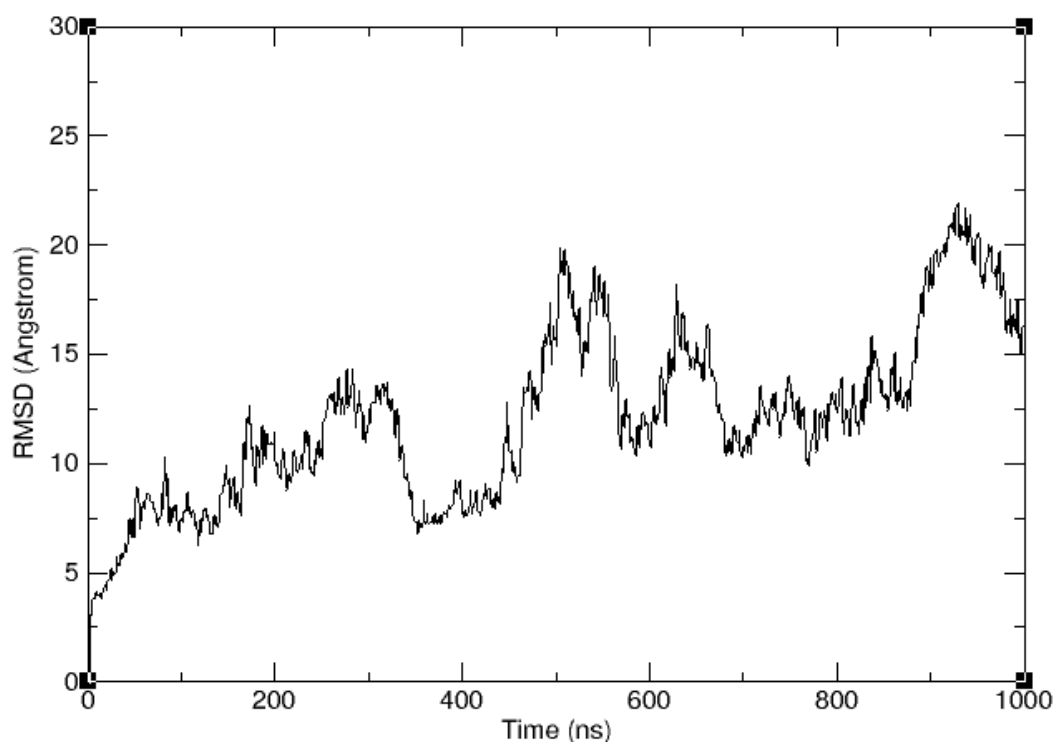
Μία ενδιαφέρουσα ιδιότητα προς παρακολούθηση για το σύστημα είναι η ρίζα της μέσης τετραγωνικής απόκλισης ή RMSD. Στο σχήμα 3.8 (α) και (β) απεικονίζονται τα διαγράμματα του RMSD σε συνάρτηση με τον χρόνο προσομοίωσης κατά το στάδιο της παραγωγής τροχιάς για τον φυσικό τύπο και τη μετάλλαξη G12D αντίστοιχα.

Παρατηρείται και για τα δύο συστήματα ότι οι τιμές του RMSD μεταβάλλονται καθόλη τη διάρκεια της προσομοίωσης. Η παρατήρηση αυτή αποτελεί ένδειξη ότι τα δύο μονομερή αλλάζουν θέση και προσανατολισμό στο χώρο αλληλεπιδρώνοντας μεταξύ τους, γεγονός που οδηγεί σε επιπλέον αλληλεπιδράσεις και επομένως αλλαγή των ατομικών θέσεων στο χώρο, με βάση τις οποίες πραγματοποιείται ο υπολογισμός του RMSD. Και στα δύο συστήματα, το RMSD παίρνει τιμές μεγαλύτερες των 5 Å, γεγονός που δεν μπορεί να δικαιολογηθεί αποκλειστικά από την κίνηση της πρωτεΐνης σε κάποιο τοπικό ελάχιστο. Η μέση τιμή του RMSD για το σύστημα του φυσικού τύπου είναι 14.0 Å, ενώ για το σύστημα της μετάλλαξης G12D είναι 12.1 Å.

(α) RMSD vs Time for KRAS-4B WT Production Run



(β) RMSD vs Time for KRAS-4B G12D Production Run



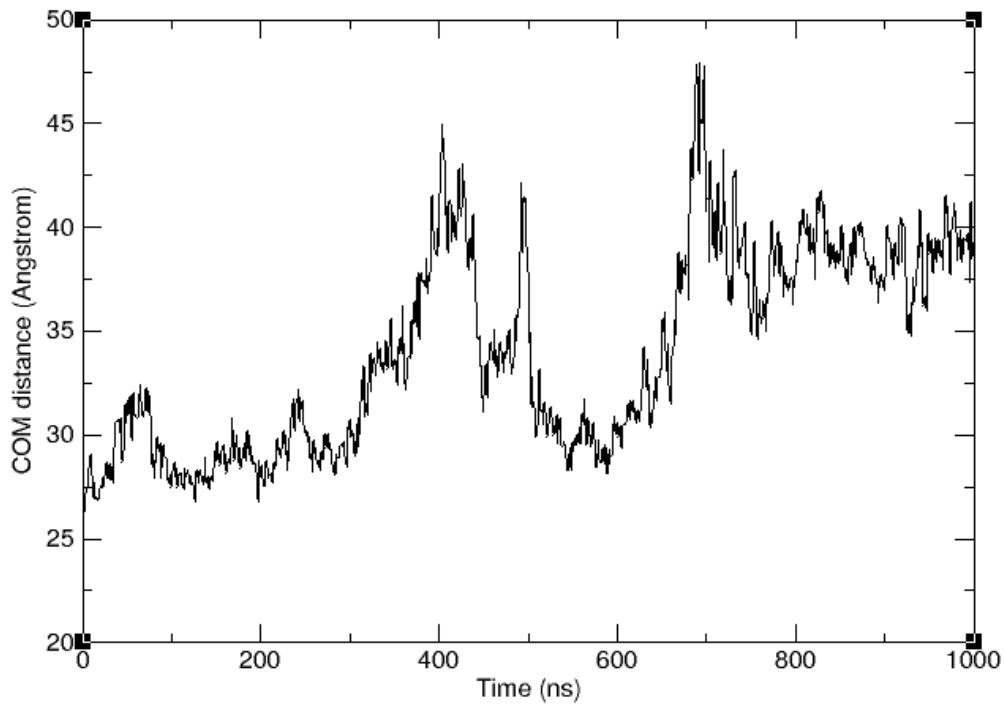
Σχήμα 3.8: Διαγράμματα RMSD συναρτήσει του χρόνου για το βήμα της παραγωγής τροχιάς στο σύστημα του διμερούς της KRAS-4B (α) φυσικού τύπου και (β) μετάλλαξη G12D.

3.2.3.3. Απόσταση κέντρων μαζών των δύο μονομερών KRAS-4B σε διάλυμα

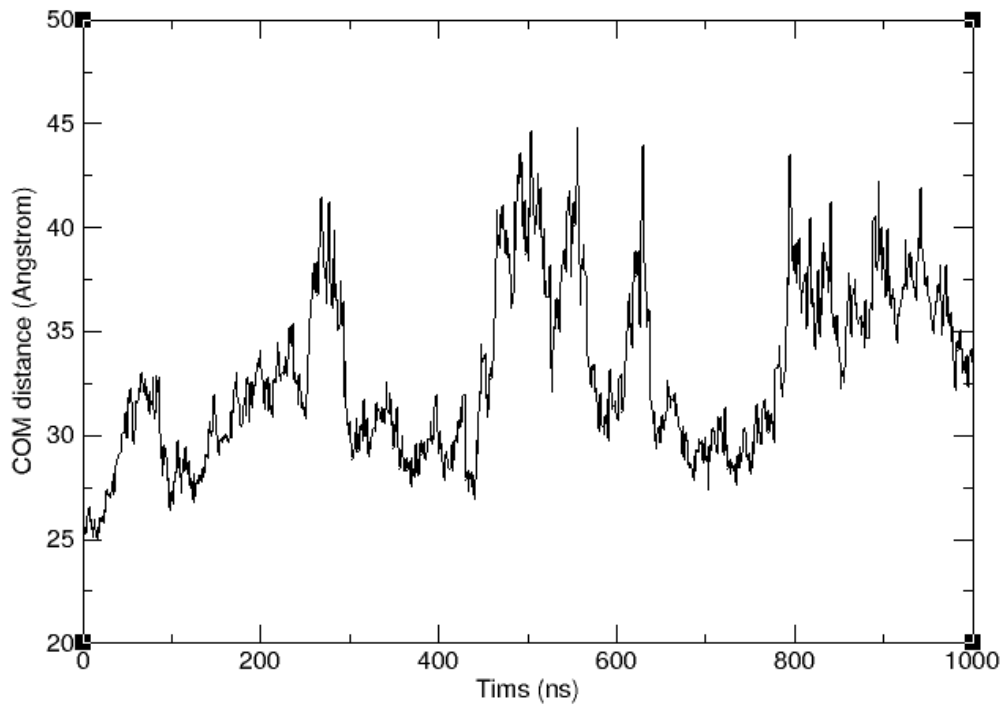
Στο σχήμα 3.9 (α) και (β) απεικονίζονται τα διαγράμματα της απόστασης των κέντρων μαζών των δύο μονομερών του διμερούς της KRAS-4B σε συνάρτηση με τον χρόνο προσομοίωσης κατά το στάδιο της παραγωγής τροχιάς για τον φυσικό τύπο και τη μετάλλαξη G12D αντίστοιχα.

Για τα δύο συστήματα παρατηρείται ότι η απόσταση των κέντρων μαζών των δύο μονομερών μεταβάλλεται σημαντικά με τη πρόοδο της προσομοίωσης, καθώς τα μονομερή αλλάζουν τον προσανατολισμό τους στο χώρο και τις μεταξύ τους θέσεις και άρα αποστάσεις. Οι διακυμάνσεις που παρατηρούνται μπορούν να ερμηνευτούν ως η απομάκρυνση και η επανασύνδεση των μονομερών. Η μέση τιμή της απόστασης των κέντρων μαζών των δύο μονομερών του διμερούς της KRAS-4B στο σύστημα του φυσικού τύπου είναι 34.1 Å, ενώ για το σύστημα της μετάλλαξης G12D είναι 33.0 Å.

(α) COM distance vs Time for KRAS-4B WT Production Run



(β) COM distance vs Time for KRAS-4B G12D Production Run



Σχήμα 3.9: Διαγράμματα απόστασης κέντρου μαζών των μονομερών συναρτήσει του χρόνου για το βήμα της παραγωγής τροχιάς στο σύστημα του διμερούς της KRAS-4B (α) φυσικού τύπου και (β) μετάλλαξη G12D.

3.3. Ομαδοποίηση των τροχιών για την ανάλυση των διαπιφανειών διμερισμού

Για την καλύτερη ανάλυση των παραπάνω αποτελεσμάτων, έγινε ομαδοποίηση των τροχιών βάσει του RMSD. Η κάθε τροχιά περιλαμβάνει 1000 στιγμιότυπα που αντιστοιχούν στις διαφορετικές διαμορφώσεις που ομαδοποιούνται. Οι ακτίνες αποκοπής που χρησιμοποιήθηκαν είναι 5 Å, 10 Å και 15 Å. Στον πίνακα 3.1 φαίνεται ο αριθμός των συστάδων που προκύπτουν για αυτές τις ακτίνες αποκοπής για το σύστημα φυσικού τύπου και τη μετάλλαξη G12D.

Πίνακας 3.1: Αριθμός συστάδων για διαφορετικές ακτίνες αποκοπής στα δύο συστήματα

Ακτίνα αποκοπής (Å)	WT KRAS-4B	G12D KRAS-4B
5	57	99
10	8	11
15	4	3

Η ομαδοποίηση που προκύπτει όταν η ακτίνα αποκοπής τίθεται 5 Å για τα συστήματα φυσικού τύπου και της μετάλλαξης δημιουργεί 57 και 99 συστάδες αντίστοιχα με τη πρώτη και μεγαλύτερη συστάδα να περιέχει μόλις 119 και 53 διαμορφώσεις αντίστοιχα, δηλαδή μεγάλο αριθμό συστάδων με μικρό πληθυσμό.

Αντίθετα, η ομαδοποίηση που προκύπτει όταν η ακτίνα αποκοπής τίθεται 15 Å για τα συστήματα φυσικού τύπου και της μετάλλαξης δημιουργεί μόλις 4 και 3 συστάδες αντίστοιχα με τη πρώτη και μεγαλύτερη συστάδα να περιέχει μόλις 868 και 986 διαμορφώσεις αντίστοιχα, δηλαδή μικρό αριθμό συστάδων με μεγάλο πληθυσμό.

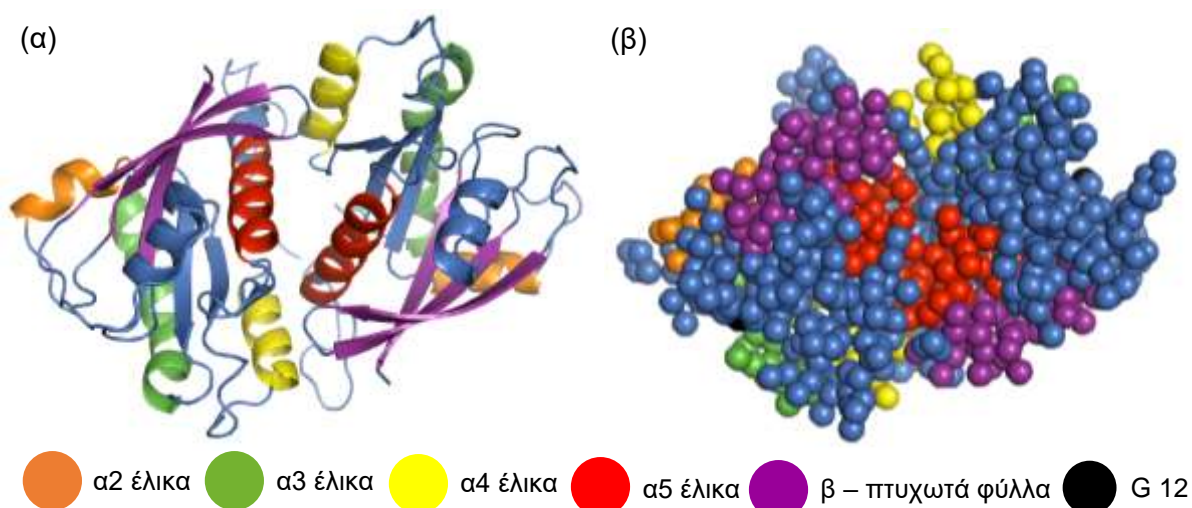
Επιλέγεται λοιπόν η ομαδοποίηση που προκύπτει από την χρήση ακτίνας αποκοπής 10 Å, η οποία παρέχει έναν καλό συμβιβασμό ανάμεσα στον αριθμό των συστάδων που δημιουργούνται και στον πληθυσμό τους. Στον πίνακα 3.2 παρουσιάζεται το πλήθος των στιγμιότυπων που έχουν ομαδοποιηθεί στις τρεις πρώτες και κύριες συστάδες των δύο συστημάτων. Επιλέχθηκαν οι τρεις πρώτες συστάδες επειδή μόνο αυτές είχαν πληθυσμό διαμορφώσεων περίπου ίσο ή μεγαλύτερο του 100.

Παρατηρείται ότι η πρώτη συστάδα του συστήματος της μετάλλαξης G12D έχει μεγαλύτερο πλήθος σε σχέση με το σύστημα φυσικού τύπου. Αυτό μπορεί ενδεχομένως να ερμηνευτεί σαν μεγαλύτερη διαμορφωτική σταθερότητα του συστήματος της μετάλλαξης.

Πίνακας 3.2: Πλήθος διαμορφώσεων στις αρχικές συστάδες με ακτίνα αποκοπής 10 Å στα δύο συστήματα

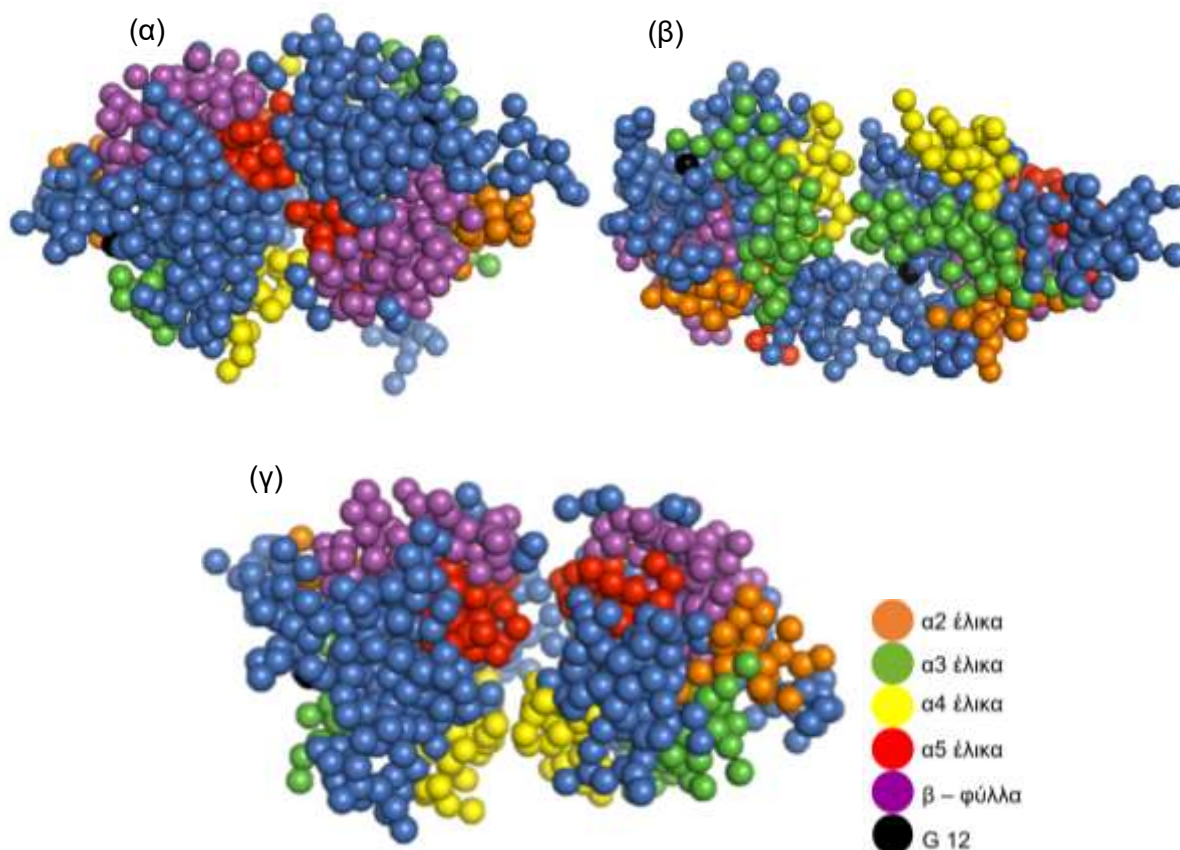
Αριθμός ομάδας	WT KRAS-4B	G12D KRAS-4B
1	402	640
2	280	189
3	269	95

Για την οπτικοποίηση των αντιπροσωπευτικών δομών των κύριων συστάδων, σημειώνονται με διαφορετικά χρώματα οι δευτεροταγείς δομές που εξετάζονται αν συμμετέχουν στη διεπιφάνεια διμερισμού. Σημειώνεται ότι η αρχική δομή με PDB ID 5nq2, η οποία χρησιμοποιείται για τις προσομοιώσεις εμφανίζει διεπιφάνεια διμερισμού $\alpha 5 - \alpha 5$. Στο σχήμα 3.10 παρουσιάζεται η αντιστίχιση των δευτεροταγών δομών που εξετάζονται αν συμμετέχουν στη διεπιφάνεια διμερισμού σε αναπαράσταση με κορδέλες (Σχήμα 3.10 (α)) και σε αδροτοποιημένη μορφή (Σχήμα 3.10 (β)).



Σχήμα 3.10: Αντιστοίχιση των δευτεροταγών δομών υπό μελέτη (α) σε ατομιστική αναπαράσταση και (β) σε αδροτοποιημένη αναπαράσταση

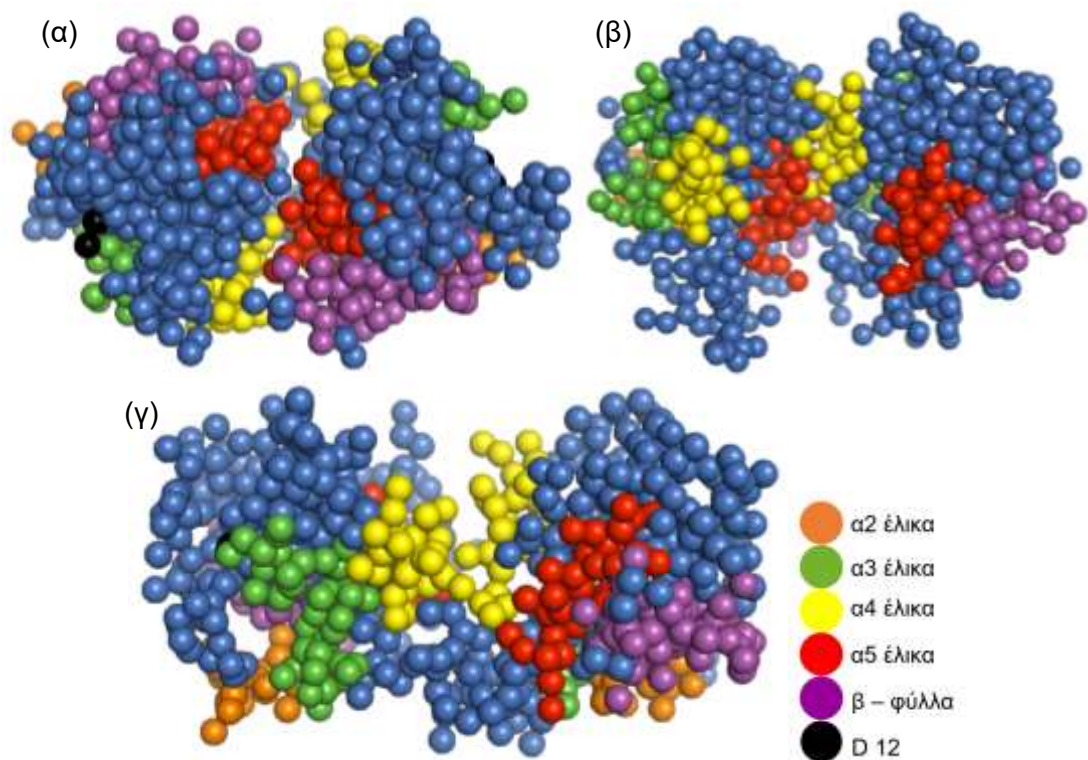
Στο Σχήμα 3.11 απεικονίζονται οι αντιπροσωπευτικές δομές των τριών κύριων συστάδων του διμερούς φυσικού τύπου. Η πρώτη συστάδα (Σχήμα 3.11 (α)) περιλαμβάνει τις δομές με διεπιφάνεια διμερισμού $\alpha 4 - \alpha 5$. Η δεύτερη συστάδα (Σχήμα 3.11 (β)) αντιστοιχεί στις δομές με διεπιφάνεια διμερισμού που οφείλεται στην αλληλεπίδραση των ελικών $\alpha 3 - \alpha 4$. Η τρίτη συστάδα (Σχήμα 3.11 (γ)) περιλαμβάνει τις δομές στις οποίες οι αλληλεπιδράσεις μεταξύ των ελικών $\alpha 4 - \alpha 4$ και $\alpha 5 - \alpha 5$ σχηματίζουν τη διεπιφάνεια διμερισμού.



Σχήμα 3.11: Αντιπροσωπευτικές δομές των τριών πρώτων συστάδων της KRAS-4B φυσικού τύπου. Παρουσιάζουν διεπιφάνειες διμερισμού (α) ελικών α4 – α5, (β) ελικών α3 – α4 και (γ) ελικών α4 – α4 και α5 – α5

Αντίστοιχα, το Σχήμα 3.12 απεικονίζει τις αντιπροσωπευτικές δομές των τριών κύριων συστάδων του μεταλλαγμένου διμερούς D12G. Και πάλι στην αντιπροσωπευτική διαμόρφωση της πρώτης συστάδας (Σχήμα 3.12 (α)) παρατηρείται η διεπιφάνεια διμερισμού α4 – α5, ενώ στην αντιπροσωπευτική διαμόρφωση της δεύτερης (Σχήμα 3.12 (β)) φαίνεται ότι τα μονομερή αλληλεπιδρούν μεταξύ τους μέσω μιας διαφορετικής διαμόρφωσης των ελικών α4 – α5. Τέλος η τρίτη συστάδα (Σχήμα 3.12 (γ)) περιλαμβάνει τις δομές στις οποίες οι αλληλεπιδράσεις μεταξύ των ελικών α4 – α4 σχηματίζουν τη διεπιφάνεια διμερισμού.

Όλες αυτές οι διεπιφάνειες διμερισμού που παρατηρούνται τόσο για το σύστημα φυσικού τύπου, όσο και για το σύστημα της μετάλλαξης G12D, συμφωνούν με τις διεπιφάνειες που αναφέρονται στην βιβλιογραφία. Ωστόσο είναι πιθανόν να υπάρχουν και άλλες διεπιφάνειες διμερισμού, οι οποίες να μην φαίνονται σε αυτό το χρονικό εύρος της προσομοίωσης και άρα να απαιτείται επιπλέον υπολογιστικός χρόνος για να παρατηρηθούν.



Σχήμα 3.12: Αντιπροσωπευτικές δομές των τριών πρώτων συστάδων της KRAS-4B μετάλλαξης G12D. Παρουσιάζουν διεπιφάνειες διμερισμού (α) ελικών α4 – α5, (β) ελικών α4 – α5 και (γ) ελικών α4 – α4

4. Συμπεράσματα και συζήτηση

Η πρωτεΐνη KRAS-4B είναι μια GTPάση που συμμετέχει στην ενδοκυτταρική σηματοδότηση και είναι ενεργή όταν είναι συνδεδεμένη με το μόριο GTP. Η KRAS-4B διμερίζεται παρουσία της πρωτεΐνης Raf και η Raf χρησιμοποιεί την KRAS-4B σαν ικρίωμα για να διμεριστεί η ίδια και να σηματοδοτήσει το κύτταρο για πολλαπλασιασμό. Από άποψη δομής, στην ενεργή διμερισμένη KRAS-4B φυσικού τύπου, οι περιοχές διακοπών P-loop, SI και SII αλληλεπιδρούν με τις φωσφορικές ομάδες του μορίου GTP και είναι υπεύθυνες για τη λειτουργία της ως GTPάση. Η KRAS-4B απενεργοποιείται όταν υδρολυθεί το GTP σε GDP. Ωστόσο, όταν μεταλλάσσεται το αμινοξύ γλυκίνη στη θέση 12 με ασπαρτικό οξύ (G12D) στην περιοχή P-loop, η υδρόλυση του μορίου GTP σε GDP αναστέλεται και η KRAS-4B παγώνει στην ενεργή κατάσταση, προκαλώντας ανεξέλεγκτη κυτταρική ανάπτυξη, διακοπή δημιουργίας αποπτωτικών σημάτων και τελικά καρκίνο.

Παρά την εκτεταμένη βιβλιογραφία σχετικά με τις επιδράσεις της μετάλλαξης G12D στην KRAS-4B, οι επιδράσεις της μετάλλαξης στην τοπική διαμόρφωση και δομή της ενεργής και ανενεργής πρωτεΐνης παραμένουν άγνωστες. Ταυτόχρονα, δεδομένου ότι η λειτουργία της πρωτεΐνης σχετίζεται εγγενώς με τη δυναμική της, η γνώση των δομικών διαφορών μεταξύ του φυσικού τύπου και της μετάλλαξης G12D μπορεί να χρησιμοποιηθεί για την αποτελεσματική στόχευση της μεταλλαγμένης KRAS-4B. Αυτό μπορεί να επιτευχθεί ακολουθώντας μια ιδιαίτερα υποσχόμενη και ενδιαφέρουσα κατεύθυνση για την θεραπεία αυτού του τύπου των καρκίνων, όπου μικρά μόρια σχεδιάζονται ώστε να παρεμποδίζουν τον διμερισμό της μεταλλαγμένης μορφής με σκοπό να αποτρέψουν την ανεξέλεγκτη και συνεχόμενη αναπαραγωγή των κυττάρων. Στόχος της παρούσας εργασίας ήταν η διερεύνηση του διμερισμού του φυσικού και μεταλλαγμένου τύπου KRAS-4B με προσομοιώσεις Μοριακής Δυναμικής με αδροποιημένα μοντέλα.

Οι μοριακές προσομοιώσεις και συγκεκριμένα η μέθοδος της Μοριακής Δυναμικής αποτελεί έναν αποτελεσματικό τρόπο να ερευνηθούν προβλήματα όπως αυτό του διμερισμού της ογκοπρωτεΐνης KRAS-4B. Ωστόσο, καθώς το σύστημα είναι μεγάλο και η κλίμακα χρόνου στην οποία εξελίσσεται ο διμερισμός είναι μεγαλύτερη από αυτή στην οποία μπορεί να μελετηθεί ένα φαινόμενο με ατομιστικά μοντέλα Μοριακής Δυναμικής, γίνεται χρήση των αδροποιημένων μοντέλων Μοριακής Δυναμικής, στα οποία ομάδες ατόμων αναπαριστώνται από σφαίρες με τις αντίστοιχες φυσικοχημικές ιδιότητες.

Στη παρούσα εργασία μοντελοποιούνται δύο συστήματα σε υδατικό περιβάλλον και σε αδροποιημένη μορφή, αυτό του φυσικού τύπου της πρωτεΐνης με PDB ID 5VQ2 [25] και αυτό της μετάλλαξης G12D της ίδιας δομής. Για τα στάδια της ελαχιστοποίησης ενέργειας,

της εξισορρόπησης ενέργειας στο κανονικό θερμοδυναμικό σύνολο και με σταθερές τις θέσεις των ατόμων στο σύστημα (στάδιο PR-NVT), της εξισορρόπησης ενέργειας στο ισόθερμο – ισοβαρές θερμοδυναμικό σύνολο και με σταθερές τις θέσεις των ατόμων στο σύστημα (στάδιο PR-NpT) και της εξισορρόπησης ενέργειας στο ισόθερμο – ισοβαρές θερμοδυναμικό σύνολο, χωρίς κάποια εφαρμογή δυνάμεων στα άτομα του συστήματος (στάδιο NpT), η δυναμική ενέργεια των δύο συστημάτων παρακολουθείται και αναλύεται. Προκύπτει ότι σε κάθε στάδιο η προσομοίωση συγκλίνει και εξελίσσεται με τρόπο που επιτρέπει την περαιτέρω εξέλιξη της μοντελοποίησης. Στο τελικό βήμα της παραγωγής τροχιάς, η οποία μοντελοποιείται για συνολικό χρόνο προσομοίωσης 1 μ s, παρακολουθείται και αναλύεται η δυναμική ενέργεια του συστήματος. Προκύπτει ότι κατά τη διάρκεια της παραγωγής τροχιάς το σύστημα είναι σταθερό. Επιπλέον παρακολουθούνται τα μεγέθη του RMSD και του COM distance, τα οποία είναι αντιπροσωπευτικά της εξέλιξης του διμερισμού της πρωτεΐνης KRAS-4B. Η εξέλιξη των μεγεθών αυτών στον χρόνο μπορούν να ερμηνευτούν σαν αλλαγή της διεπιφάνειας διμερισμού, σε καθένα από τα δύο συστήματα. Επιπλέον, οι αναπαραστάσεις της εξέλιξης των δύο τροχιών στον χρόνο που παρήχθησαν από τις προσομοιώσεις, επιβεβαιώνουν ότι τα συστήματα αλλάζουν τον σχετικό τους προσανατολισμό και την διεπιφάνειά τους. Για την ποσοτικοποίηση των παρατηρήσεων αυτών πραγματοποιείται ομαδοποίηση με βάση το RMSD και επιλέγεται η ομαδοποίηση με ακτίνα αποκοπής 10 Å. Οι αντιπροσωπευτικές διαμορφώσεις των τριών κύριων συστάδων οπτικοποιούνται και παρατηρούνται διεπιφάνειες διμερισμού που έχουν προηγουμένως αναφερθεί στην βιβλιογραφία.

Επομένως συμπεραίνεται ότι κατά τη διάρκεια της συγκεκριμένης προσομοίωσης, στο σύστημα φυσικού τύπου παρατηρείται ο σχηματισμός διεπιφανειών διμερισμού $\alpha_4 - \alpha_5$, $\alpha_3 - \alpha_4$ και $\alpha_4 - \alpha_4$ με $\alpha_5 - \alpha_5$. Αντίστοιχα για το σύστημα της μετάλλαξης G12D παρατηρείται ο σχηματισμός των διεπιφανειών διμερισμού $\alpha_4 - \alpha_5$ και $\alpha_4 - \alpha_4$. Οι διεπιφάνειες διμερισμού που παρατηρούνται στην προσομοίωση που έγινε συμφωνούν με δομές οι οποίες είναι γνωστές από τη βιβλιογραφία. Ωστόσο δεν βρέθηκαν όλες οι γνωστές διεπιφάνειες διμερισμού και αυτό οφείλεται στον περιορισμένο χρόνο προσομοίωσης (1 μ s) και την μειωμένη δειγματοληψία.

5. Μελλοντικές προοπτικές

Από τις προσομοιώσεις που παρουσιάζονται στη συγκεκριμένη εργασία δεν μπορεί να χαρτογραφηθεί το πλήρες τοπίο δυναμικής ενέργειας του διμερισμού της πρωτεΐνης KRAS-4B, λόγω του περιορισμένου χρόνου προσομοίωσης αλλά και των υψηλών φραγμάτων ενέργειας μεταξύ των διαφορετικών διαμορφώσεων του διμερούς. Σε επόμενα βήματα αυτού του έργου πρόκειται να μοντελοποιηθεί το φαινόμενο για μεγαλύτερο χρονικό διάστημα (πχ 20 μ s), για να μελετηθεί εάν υπάρχουν επιπλέον διαμορφώσεις. Προκειμένου να ξεπεραστούν τα ενεργειακά φράγματα μεταξύ των τοπικών ελαχίστων, τα επόμενα βήματα αυτής της εργασίας περιλαμβάνουν την πραγματοποίηση των προσομοιώσεων Μεταδυναμικής παράλληλης θερμοκρασίας στο κανονικό σύνολο (PT-MetaD-WTE), προκειμένου να ξεπεραστούν τα υψηλά φράγματα ενέργειας και να χαρτογραφηθεί το πλήρες τοπίο δυναμικής ενέργειας του διμερισμού της πρωτεΐνης KRAS-4B. Τα μεγέθη του RMSD και της απόστασης των κεντρών μάζας των μονομερών πρόκειται να χρησιμοποιηθούν ως συλλογικές μεταβλητές για την καθοδήγηση αλλά και την παρακολούθηση του κάθε συστήματος.

Μετά την ολοκλήρωση των προσομοιώσεων PT-MetaD-WTE, είναι χρήσιμο να πραγματοποιηθούν οι ίδιες προσομοιώσεις για τα συστήματα που μελετήθηκαν, προσθέτοντας σηματοδοτικές πρωτεΐνες και τα μόρια GTP, ώστε να επιτευχθεί μια απεικόνιση των συστημάτων πιο κοντά στην πραγματικότητα. Επιπλέον, οι προσομοιώσεις μπορούν μοντελοποιηθούν σε συστήματα παρουσία της μεμβράνης, ώστε να γίνει σύγκριση με τις προσομοιώσεις που πραγματοποιήθηκαν σε υδατικό περιβάλλον, όπως αυτό που περιγράφεται στη παρούσα εργασία. Η εισαγωγή της μεμβράνης θα οδηγήσει στην πιο ρεαλιστική περιγραφή του συστήματος, καθώς το διμερές της KRAS-4B αλληλεπιδρά με τη μεμβράνη κατά τη διεργασία του διμερισμού.

Ακόμα, οι προσομοιώσεις που αναπτύχθηκαν για την παρούσα εργασία μπορούν να πραγματοποιηθούν για περισσότερες μεταλλάξεις, τόσο στο κωδικόνιο 12, όπως οι G12V, G12C και G12A, όσο και στα κωδικόνια 13 και 61. Τα αποτελέσματα από μια μελέτη σαν αυτή, μπορούν να δώσουν μια σφαιρική εικόνα και να συντελέσουν στον σχεδιασμό φαρμάκων και της εξατομικευμένης θεραπείας διαφόρων ειδών καρκίνου, οι οποίοι οφείλονται στην εκάστοτε μετάλλαξη. Δηλαδή, θα πραγματοποιείται ανάγνωση του γονιδιώματος του ασθενή και ανάλογα με τον τύπο μετάλλαξης του γονιδίου που κωδικοποιεί την πρωτεΐνη KRAS-4B, θα χορηγείται στον εκάστοτε ασθενή η κατάλληλη φαρμακευτική αγωγή, η οποία θα έχει σχεδιαστεί με βάση τα ειδικά χαρακτηριστικά της μετάλλαξης αυτής, μειώνοντας έτσι τις ανεπιθύμητες ενέργειες. Τέλος η ίδια διαδικασία μπορεί να επαναληφθεί για τα άλλα μέλη της οικογένειας των πρωτεϊνών RAS.

Βιβλιογραφία

- [1] J. F. Hancock, «Ras proteins: different signals from different locations,» *Nature Reviews Molecular Cell Biology*, τόμ. 4, αρ. 5, pp. 373-385, 2003.
- [2] S. Lu, H. Jang, S. Muratcioglu, A. Gursoy, O. Keskin, R. Nussinov και J. Zhang, «Ras Conformational Ensembles, Allostery, and Signaling,» *Chemical Reviews*, τόμ. 116, pp. 6607-6665, 2016.
- [3] «Wikimedia Commons,» [Ηλεκτρονικό]. Available: https://commons.wikimedia.org/wiki/File:GTP_chemical_structure.svg. [Πρόσβαση 23 October 2021].
- [4] S. Lu, H. Jang, S. Gu, J. Zhang και R. Nussinov, «Drugging Ras GTPase: a comprehensive mechanistic and signaling structural view,» *Chemical Society Reviews*, τόμ. 45, αρ. 18, p. 4929–4952, 2016.
- [5] A. D. Cox, S. W. Fesik, A. C. Kimmelman, J. Luo και C. J. Der, «Drugging the undruggable RAS: Mission Possible?,» *Nature Reviews Drug Discovery*, τόμ. 13, αρ. 11, p. 828–851, 2014.
- [6] E. H. Chang, M. A. Gonda, R. W. Ellis, E. M. Scolnick και D. R. Lowy, «Human genome contains four genes homologous to transforming genes of Harvey and Kirsten murine sarcoma viruses,» *Proceedings of the National Academy of Sciences*, τόμ. 79, αρ. 16, pp. 4848-4852, 1982.
- [7] D. S. Goodsell, «The Molecular Perspective: The ras Oncogene,» *The Oncologist*, τόμ. 4, αρ. 3, pp. 263-264, 1999.
- [8] M. Malumbres και M. Barbacid, «RAS oncogenes: the first 30 years,» *Nature Reviews Cancer*, τόμ. 3, αρ. 6, pp. 459-465, 2003.
- [9] C. Ambrogio, J. Kohler, Z.-W. Zhou, H. Wang, R. Paranal, J. Li, M. Capelletti, C. Caffarra, S. Li, Q. Lv, S. Gondi, J. C. Hunter, J. Lu, R. Chiarle, D. Santamaria, K. D. Westover και P. A. Janne, «KRAS Dimerization Impacts MEK Inhibitor Sensitivity and Oncogenic Activity of Mutant KRAS,» *Cell*, τόμ. 172, p. 857–868, 2018.
- [10] J. Yuan, X. Dong, J. Yap και J. Hu, «The MAPK and AMPK signalings: interplay and implication in targeted cancer therapy,» *Journal of Hematology & Oncology*, τόμ. 13, αρ.

113, 2020.

- [11] A. D. Cox και C. J. Der, «Ras history,» *Small GTPases*, τόμ. 1, αρ. 1, p. 2–27, 2010.
- [12] R. A. Weinberg και A. E. Karnoub, «Ras oncogenes: split personalities,» *Nature Reviews Molecular Cell Biology*, τόμ. 9, pp. 517-531, 2008.
- [13] R. Nussinov, H. Jang, C.-J. Tsai και F. Cheng, «Review: Precision medicine and driver mutations: Computational methods, functional assays and conformational principles for interpreting cancer drivers,» *PLOS Computational Biology*, τόμ. 15, αρ. 3, 2019.
- [14] A. Kumar και A. Jha, «Chapter 7 - Drug Development Strategies,» σε *Anticandidal Agents*, 2017, pp. 63-71.
- [15] «U.S. Food & Drug Administration,» 28 05 2021. [Ηλεκτρονικό]. Available: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-sotorasib-kras-g12c-mutated-nsclc>. [Πρόσβαση 24 09 2021].
- [16] N. Staff, «National Cancer Institute,» 25 06 2021. [Ηλεκτρονικό]. Available: <https://www.cancer.gov/news-events/cancer-currents-blog/2021/fda-sotorasib-lung-cancer-kras>. [Πρόσβαση 24 09 2021].
- [17] F. Skoulidis, B. T. Li, G. K. Dy, T. J. Price, G. S. Falchook, J. Wolf, A. Italiano, M. Schuler, H. Borghaei, F. Barlesi, T. Kato, A. Curioni-Fontecedro, A. Sacher, A. Spira, S. S. Ramalingam, T. Takahashi, B. Besse, A. Anderson, A. Ang, Q. Tran, O. Mather, H. Henary, G. Ngarmchamnanrith, G. Friberg, V. Velcheti και R. Govindan, «Sotorasib for Lung Cancers with KRAS p.G12C Mutation,» *The New England Journal of Medicine*, pp. 2371-2381, 2021.
- [18] «U.S. Food & Drug Administration,» 28 05 2021. [Ηλεκτρονικό]. Available: <https://www.fda.gov/news-events/press-announcements/fda-approves-first-targeted-therapy-lung-cancer-mutation-previously-considered-resistant-drug>. [Πρόσβαση 24 09 2021].
- [19] «National Center for Biotechnology Information. PubChem Compound Database,» U.S. National Library of Medicine, [Ηλεκτρονικό]. Available: <https://pubchem.ncbi.nlm.nih.gov/compound/Sotorasib>. [Πρόσβαση 23 October 2021].
- [20] J. A. Parker και C. Mattos, «The Ras–Membrane Interface: Isoform-Specific Differences in the Catalytic Domain,» *Molecular Cancer Research*, τόμ. 13, αρ. 4, pp. 595-603,

2015.

- [21] K.-Y. Lee, Z. Fang, M. Enomoto, G. Gasmi-Seabrook, L. Zheng, S. Koide, M. Ikura και C. B. Marshall, «Two Distinct Structures of Membrane-Associated Homodimers of GTP- and GDP-Bound KRAS4B Revealed by Paramagnetic Relaxation Enhancement,» *Angewandte Chemie*, τόμ. 59, αρ. 27, pp. 11037-11045, 2020.
- [22] M. V. Milburn, L. Tong, A. M. DeVos, A. Brunger, Z. Yamaizumi, S. Nishimura και S. H. Kim, «Molecular Switch for Signal Transduction: Structural Differences between Active and Inactive Forms of Protooncogenic Ras,» *Science*, τόμ. 247, pp. 939-945, 1990.
- [23] S. Lu, A. Banerjee, H. Jang, J. Zhang, V. Gaponenko και R. Nussinov, «Gtp Binding and Oncogenic,» *Journal of Biological Chemistry*, τόμ. 290, pp. 28887-28900, 2015.
- [24] T. Pantsar, «The current understanding of KRAS protein structure and dynamics,» *Computational and structural biotechnology journal*, τόμ. 18, pp. 189-198, 2019.
- [25] S. Xu, B. N. Long, G. H. Boris, A. Chen, S. Ni και M. A. Kennedy, «Structural insight into the rearrangement of the switch I region in GTP-bound G12A K-Ras,» *Acta Crystallographica Section D Structural Biology*, τόμ. 73, αρ. 12, p. 970–984, 2017.
- [26] I. A. Prior, P. D. Lewis και C. Mattos, «A Comprehensive Survey of Ras Mutations in Cancer,» *Cancer Research*, τόμ. 72, αρ. 10, p. 2457–2467, 2012.
- [27] S. Vatansever, B. Erman και Z. H. Gümüş, «Oncogenic G12D mutation alters local conformations and dynamics of K-Ras,» *Scientific Reports*, τόμ. 9, αρ. 11730, 2019.
- [28] J. P. Glusker, «X-ray crystallography of proteins,» *Methods of Biochemical Analysis*, τόμ. 37, αρ. 1, pp. 1-72, 1994.
- [29] Y. Cheng, «Single-Particle Cryo-EM at Crystallographic Resolution,» *Cell*, τόμ. 161, αρ. 3, pp. 450-457, 2015.
- [30] R. Pearce και Y. Zhang, «Toward the solution of the protein structure prediction problem,» *Journal of Biological Chemistry*, τόμ. 297, αρ. 1, 2021.
- [31] «Protein Data Bank,» [Ηλεκτρονικό]. Available: <https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/methods-for-determining-structure>. [Πρόσβαση 15 09 2021].
- [32] M. E. Welsch, A. Kaplan, J. M. Chambers, M. E. Stokes, P. H. Bos, A. Zask, Y. Zhang, M. Sanchez-Martin, M. A. Badgley, C. S. Huang, T. H. Tran, H. Akkiraju, L. M. Brown, R.

- Nandakumar, S. Cremers, W. Seok Yang, L. Tong, K. P. Olive, A. Ferrando και B. R. Stockwell, «Multivalent Small-Molecule Pan-RAS Inhibitors,» *Cell*, τόμ. 168, αρ. 5, pp. 878-889, 2017.
- [33] S. Sogabe, Y. Kamada, M. Miwa, A. Niida, T. Sameshima, M. Kamaura, K. Yonemori, S. Sasaki, J.-i. Sakamoto και K. Sakamoto, «Crystal Structure of a Human K-Ras G12D Mutant in Complex with GDP and the Cyclic Inhibitory Peptide KRpep-2d,» *ACS Medicinal Chemistry Letters*, τόμ. 8, αρ. 7, pp. 732-736, 2017.
- [34] D. Kessler, M. Gmachl, A. Mantoulidis, L. J. Martin, A. Zoepfel, M. Mayer, A. Gollner, D. Covini, S. Fischer, T. Gerstberger, T. Gmaschitz, C. Goodwin, P. Greb, D. Häring, W. Hela, J. Hoffmann, J. Karolyi-Oezguer, P. Knesl, S. Kornigg, M. Koegl, R. Kousek, L. Lamarre, F. Moser, S. Munico-Martinez, C. Peinsipp, J. Phan, J. Rinenthal, J. Sai, C. Salamon, Y. Scherbantin, K. Schipany, R. Schnitzer, A. Schrenk, B. Sharps, G. Siszler, Q. Sun, A. Waterson, B. Wolkerstorfer, M. Zeeb, M. Pearson, S. W. Fesik και D. B. McConnell, «Drugging an undruggable pocket on KRAS,» *Proceedings of the National Academy of Sciences*, τόμ. 116, αρ. 32, pp. 15823-15829, 2019.
- [35] D. Frenkel και B. Smit, *Understanding Molecular Simulations From Algorithms to Applications*, Orlando, Florida: Academic Press, 2002.
- [36] Z. Cournia, *Structure and Dynamics of Biomembranes containing Cholesterol and other Biologically-Important Sterols. A computational perspective.*, Heidelberg, 2006.
- [37] P. Atkins και J. De Paula, *Φυσικοχημεία, Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης*, 2015.
- [38] E. Lindahl, M. J. Abraham, B. Hess και van der Spoel, «GROMACS 2020.6 Manual (2020.6),» 2021. [Ηλεκτρονικό]. Available: <https://doi.org/10.5281/zenodo.4576060>. [Πρόσβαση 22 9 2021].
- [39] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess και E. Lindahl, «GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers,» *SoftwareX*, Τόμ. 1 από 2, pp. 19-25, 2015.
- [40] «Protein Data Bank,» [Ηλεκτρονικό]. Available: <http://www.rcsb.org>. [Πρόσβαση 20 9 2021].
- [41] S. J. Marrink, H. J. Risselada και S. Yef, «The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations,» *The Journal of Physical Chemistry B*, τόμ. 111,

αρ. 27, p. 7812–7824, 2007.

- [42] P. C. T. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grünewald, I. Patmanidis, H. Abdizadeh, B. M. H. Bruininks, T. A. Wassenaar, P. C. Kroon, J. Melcr, V. Nieto, V. Corradi, H. M. Khan, J. Domański, M. Javanainen, H. Martinez-Seara, N. Reuter, R. B. Best, I. Vattulainen, L. Monticelli, X. Periole, D. P. Tieleman, A. H. de Vries και S. J. Marrink, «Martini 3: a general purpose force field for coarse-grained molecular dynamics,» *Nature Methods*, τόμ. 18, p. 382–388, 2021.
- [43] L. Marrink, «github.com/marrink-lab/gromit,» 23 2020. [Ηλεκτρονικό]. Available: <https://github.com/marrink-lab/gromit/blob/master/martinate.sh>. [Πρόσβαση 22 9 2021].
- [44] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren και A. E. Mark, «Peptide Folding: When Simulation Meets Experiment,» *Angewandte Chemie International Edition*, τόμ. 38, αρ. 1-2, pp. 236-240, 1999.
- [45] C. Lamprakis, I. Andreadelis, J. Manchester, C. Velez-Vega, J. S. Duca και Z. Cournia, «Evaluating the Efficiency of the Martini Force Field to Study Protein Dimerization in Aqueous and Membrane Environments,» *Journal of Chemical Theory and Computation*, τόμ. 17, αρ. 5, pp. 3088-3102, 2021.
- [46] T. Maurer, L. S. Garrenton, A. Oh, K. Pitts, D. J. Anderson, N. J. Skelton, B. P. Fauber, B. Pan, S. Malek, D. Stokoe, M. J. C. Ludlam, K. K. Bowman, J. Wu, A. M. Giannetti, M. A. Starovasnik, I. Mellman, P. K. Jackson, J. Rudolph, W. Wang και G. Fang, «Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity,» *Proceedings of the National Academy of Sciences*, τόμ. 109, αρ. 14, pp. 5299-5304, 2012.
- [47] S. Muratcioglu, T. S. Chavan, B. C. Freed, H. Jang, L. Khavrutskii, R. N. Freed, M. A. Dyba, K. Stefanisko, S. G. Tarasov, A. Gursoy, O. Keskin, N. I. Tarasova, V. Gaponenko και R. Nussinov, «GTP-Dependent K-Ras Dimerization,» *Structure*, τόμ. 23, αρ. 7, pp. 1325-1335, 2015.
- [48] A. Cruz-Migoni, P. Canning, C. E. Quevedo, C. J. R. Bataille, N. Bery, A. Miller, A. J. Russell, S. E. V. Phillips, S. B. Carr και T. H. Rabbitts, «Structure-based development of new RAS-effector inhibitors from a combination of active and inactive RAS-binding compounds,» *Proceedings of the National Academy of Sciences*, τόμ. 116, αρ. 7, pp. 2545-2550, 2019.

[49] M. E. Welsch , A. Kaplan , J. M. Chambers , M. E. Stokes , P. H. Bos , A. Zask , Y. Zhang , M. Sanchez-Martin , M. A. Badgley , C. S. Huang , T. H. Tran , H. Akkiraju , L. M. Brown , R. Nandakumar , S. Cremers , W. S. Yang , L. Tong , K. P. Olive , A. Ferrando και B. R. Stockwell, «Multivalent Small-Molecule Pan-RAS Inhibitors,» *Cell*, τόμ. 168, αρ. 5, pp. 878-889.e29, 2017.