



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΤΟΜΕΑΣ ΙV : ΣΥΝΘΕΣΗΣ ΚΑΙ ΑΝΑΠΤΥΞΗΣ  
ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΔΙΚΑΣΙΩΝ

## Διπλωματική Εργασία

Ανάλυση κοιλοτήτων πρόσδεσης της ογκοπρωτεΐνης  
KRAS-4B με Μοριακές Δυναμικές προσομοιώσεις

---

Μαρία-Ελευθερία Κουστένη

Επιβλέπων

Ευάγγελος Τόπακας, Αναπληρωτής Καθηγητής, Σχολή  
Χημικών Μηχανικών, ΕΜΠ

**ΑΘΗΝΑ 2021**

# Περιεχόμενα

Ευρετήριο εικόνων.....	iii
Ευρετήριο πινάκων .....	viii
Ευρετήριο διαγραμμάτων .....	ix
Ευχαριστίες.....	xi
Περίληψη.....	xiii
Abstract .....	xv
Ορολογία και συντομογραφίες .....	xvii
1 ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ.....	1
1.1 Εισαγωγή.....	1
1.1.1 Δομή και καταλυτική λειτουργία των πρωτεϊνών RAS .....	2
1.1.2 Η πρωτεΐνη Raf και η σημασία της στον διμερισμό της KRAS-4B .....	5
1.1.3 Ο ρόλος της KRAS-4B στον καρκίνο και η μετάλλαξη G12D.....	6
1.2 Ορθολογικός σχεδιασμός φαρμάκων.....	8
1.2.1 Σχεδιασμός φαρμάκων με βάση τη δομή της πρωτεΐνης .....	9
1.2.2 Κοιλότητες πρόσδεσης σε μία πρωτεΐνη .....	11
1.3 Στρατηγικές στόχευσης της KRAS-4B.....	14
1.3.1 Γνωστές θέσεις πρόσδεσης και αναστολείς της KRAS .....	16
1.4 Σκοπός της εργασίας.....	20
2 ΜΕΘΟΔΟΛΟΓΙΑ-ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ.....	21
2.1 Εμπειρική συνάρτηση δυναμικού-πεδίο δυνάμεων .....	21
2.2 Προσομοιώσεις Μοριακής Δυναμικής .....	27
2.2.1 Βασικές αρχές προσομοιώσεων Μοριακής Δυναμικής.....	29
2.2.2 Ροή εργασίας προσομοιώσεων Μοριακής Δυναμικής .....	30

2.3	Ομαδοποίηση πρωτεϊνικών δομών από προσομοίωση .....	36
2.4	Αλγόριθμοι προσδιορισμού κοιλοτήτων πρόσδεσης.....	39
2.5	Υπολογισμοί όγκου κοιλοτήτων πρόσδεσης.....	45
3	ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ .....	47
3.1	Σύστημα διμερούς φυσικού τύπου KRAS-4B συνδεδεμένο με τις πρωτεΐνες Raf στην κυτταρική μεμβράνη .....	47
3.1.1	Εύρεση κυρίαρχων δομών-διαμορφώσεων μέσω του αλγορίθμου ομαδοποίησης gmx cluster του λογισμικού GROMACS.....	48
3.1.2	Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου SiteMap .....	52
3.1.3	Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου FTMap και σύγκριση των δύο αλγορίθμων .....	55
3.1.4	Σύγκριση αποτελεσμάτων κοιλοτήτων πρόσδεσης με τη βιβλιογραφία .....	57
3.1.5	Προσδιορισμός χρονικής μεταβολής των κοιλοτήτων πρόσδεσης .....	60
3.2	Σύστημα διμερούς της μεταλλαγμένης (G12D) KRAS-4B συνδεδεμένο με τις πρωτεΐνες Raf στην κυτταρική μεμβράνη .....	72
3.2.1	Εύρεση κυρίαρχων δομών-διαμορφώσεων μέσω του αλγορίθμου ομαδοποίησης gmx cluster του λογισμικού GROMACS.....	72
3.2.2	Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου SiteMap .....	74
3.2.3	Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου FTMap και σύγκριση των δύο αλγορίθμων .....	78
3.2.4	Σύγκριση αποτελεσμάτων κοιλοτήτων πρόσδεσης με τη βιβλιογραφία .....	79
3.2.5	Προσδιορισμός χρονικής μεταβολής των θέσεων πρόσδεσης .....	81
3.2.6	Σύγκριση αποτελεσμάτων με το σύστημα φυσικού τύπου .....	82
4	ΣΥΜΠΕΡΑΣΜΑΤΑ .....	85
5	ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ .....	87
6	ΒΙΒΛΙΟΓΡΑΦΙΑ .....	88

# Ευρετήριο εικόνων

Εικόνα 1-1: Οι μεταλλάξεις των πρωτεϊνών RAS στον καρκίνο.....	1
Εικόνα 1-2: Ο κύκλος της υδρόλυσης, μεσολαβούμενος από τις πρωτεΐνες GAP, GEF [5] .....	2
Εικόνα 1-3: Χημική δομή της ένωσης GDP (αριστερά) και της ένωσης GTP (δεξιά) .....	3
Εικόνα 1-4: Δομή της KRAS-4B. lobe 1: effector lobe και lobe 2: allosteric lobe [7] .....	4
Εικόνα 1-5: Δομικές περιοχές της KRAS-4B (A,B) με τις αντίστοιχες περιοχές αμινοξικών ακολουθιών (C) [9].....	4
Εικόνα 1-6: Μηχανισμοί αλληλεπίδρασης RAS-Raf για την ενεργοποίηση του μονοπατιού MEK/Erk [11].....	5
Εικόνα 1-7: Ενεργοποίηση της κινάσης Raf από την πρωτεΐνη RAS [12].....	6
Εικόνα 1-8: Συχνότητα εμφάνισης των μεταλλάξεων στα αμινοξέα 12, 13 και 61 στα διάφορα ισόμορφα RAS [2] .....	7
Εικόνα 1-9: Οι μεταλλάξεις στα αμινοξέα 12,13 και 61 εμποδίζουν την αντίδραση υδρόλυσης του GTP από την KRAS-4B, οδηγώντας σε μία μόνιμα ενεργή μορφή της πρωτεΐνης [14] .....	7
Εικόνα 1-10: Βασικά στάδια ανακάλυψης φαρμάκων [15] .....	8
Εικόνα 1-11: Διαδικασία σχεδιασμού φαρμάκων με βάση τη δομή της πρωτεΐνης [16] .....	10
Εικόνα 1-12: Σύμπλοκο της πρωτεΐνης KRAS με το ενδογενές πρόσδεμα GTP. Τα αμινοξέα του ενεργού κέντρου φαίνονται με μπλε χρώμα, το μόριο GTP με κόκκινο χρώμα και οι αλληλεπιδράσεις μεταξύ πρωτεΐνης-πρόσδεματος με πράσινες διακεκομμένες γραμμές. Η πρωτεΐνη απεικονίζεται με τη μορφή γκρι κορδελών (ribbons), ενώ φαίνεται η επιφάνειά της με διαφανές γκρι χρώμα (PDB ID: 5VQ2) [17] .....	12
Εικόνα 1-13: Μηχανισμοί δράσης των διαφόρων τύπων αλλοστερικών μετατροπέων [18].....	13
Εικόνα 1-14: Η χημική δομή της ένωσης sotorasib, του πρώτου εγκεκριμένου φαρμάκου κατά της G12C KRAS .....	15
Εικόνα 1-15: Κρυσταλλική δομή του συμπλόκου της KRAS με το πρόσδεμα GDP και τον αναστολέα AMG-510 (PDB ID 6OIM) [26]. Αμινοξέα της κοιλότητας πρόσδεσης: 9-11, 12 (ομοιοπολικά), 16, 34, 58-63, 68, 69, 72, 92, 95, 96, 99, 100, 103.....	17
Εικόνα 1-16: Σύμπλοκο της πρωτεΐνης G12D KRAS με τον αναστολέα BI-2852 (PDB ID: 6GJ8) [31]. Ο αναστολέας φαίνεται με πράσινο χρώμα και η κοιλότητα φαίνεται ως χάρτης επιφάνειας (γκρι δίχτυ). Η πρωτεΐνη απεικονίζεται με τη μορφή κορδελών (ribbons). Αμινοξέα κοιλότητας: 5-7, 37-39, 54-56,67,70,71,74,75.....	18

Εικόνα 1-17: A,B: Χημικές δομές των ενώσεων PPIN-1, PPIN2 αντίστοιχα C,D,E,F: Κρυσταλλικές δομές της πρωτεΐνης Q61H KRAS-4B με τις ενώσεις PPIN-1 (C,E) και PPIN-2 (D,F). Αμινοξέα της κοιλότητας πρόσδεσης: 5-7, 39-41, 54-56, 71, 74 [32].....	19
Εικόνα 2-1: Το εργαλείο Protein Preparation Wizard.....	32
Εικόνα 2-2: Χάρτης επιφάνειας κοιλότητας πρόσδεσης που προσδιορίστηκε στην πρωτεΐνη KRAS-4B φυσικού τύπου, αναπαράσταση ως κόκκινο δίκτυο. Η πρωτεΐνη αναπαρίσταται με τη μορφή κορδελών (ribbons). Ακόμα, τα "site points" διακρίνονται ως λευκές σφαίρες. ....	41
Εικόνα 2-3: Μορφή αποτελεσμάτων SiteScore.....	42
Εικόνα 2-4: Το περιβάλλον εργασίας του αλγορίθμου SiteMap.....	42
Εικόνα 2-5: Τα 16 μικρά οργανικά μόρια που χρησιμοποιούνται από τον αλγόριθμο FTMap [40] .....	44
Εικόνα 2-6: Απεικόνιση αποτελεσμάτων στο περιβάλλον εργασίας του αλγορίθμου FTMap για την ρενίνη (PDB ID: 2REN) [41]. Η πρωτεΐνη απεικονίζεται με πράσινο με τη μορφή κορδελών (ribbons) και τα μόρια-προσδέματα με τη μορφή δεσμών σε διάφορα χρώματα [40].....	44
Εικόνα 2-7: Το αρχείο εισόδου που χρησιμοποιήθηκε για τις κοιλότητες πρόσδεσης του μεταλλαγμένου διμερούς KRAS-4B .....	45
Εικόνα 2-8: Το περιβάλλον εργασίας του εργαλείου Erock [42] .....	46
Εικόνα 2-9: Οπτικοποίηση των κοιλοτήτων πρόσδεσης στο λογισμικό VMD [42] .....	46
Εικόνα 3-1: Το σύστημα διμερούς φυσικού τύπου KRAS-4B. Το διμερές KRAS-4B και οι πρωτεΐνες Raf απεικονίζονται με τη μορφή κορδελών (ribbons) σε γαλάζιο και φούξια χρώμα αντίστοιχα. Η μεμβράνη απεικονίζεται με τη μορφή επιφάνειας σε ροζ χρώμα. ....	47
Εικόνα 3-2: Αρχείο script που χρησιμοποιήθηκε για την διαίρεση των αρχείων τροχιάς των συστημάτων φυσικού τύπου διμερούς KRAS-4B .....	49
Εικόνα 3-3: Αποτελέσματα κοιλοτήτων πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 1 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίκτυο). ....	52
Εικόνα 3-4: Αποτελέσματα κοιλοτήτων πρόσδεσης για τις κυρίαρχες δομές των υπο-τροχιών του αντιγράφου 1 του συστήματος φυσικού KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίκτυο). ....	53
Εικόνα 3-5: Αποτελέσματα κοιλοτήτων πρόσδεσης για τις κυρίαρχες δομές των αντιγράφων 2-6 του συστήματος φυσικού τύπου KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή	

κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ). α) αντίγραφο 2.	
β) αντίγραφο 3. γ) αντίγραφο. 4 δ) αντίγραφο 5. ε) αντίγραφο 6. ....	54
Εικόνα 3-6: Σύγκριση των αποτελεσμάτων που προέκυψαν μέσω των αλγορίθμων SiteMap (αριστερά) και FTMap (δεξιά) για τα 6 αντίγραφα του συστήματος φυσικού τύπου. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ). ....	57
Εικόνα 3-7: Διμερές KRAS-4B φυσικού τύπου συνδεδεμένο με το ενδογενές πρόσδεμα GTP. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι ορθοστερικές κοιλότητες πρόσδεσης με τη μορφή χάρτη επιφάνειας (δίχτυ) γκρι χρώματος (PDB ID: 5VQ2) [17].....	57
Εικόνα 3-8: Ορθοστερικές κοιλότητες πρόσδεσης του διμερούς, όπως εντοπίστηκαν στις δομές των 6 αντιγράφων του συστήματος φυσικού τύπου, μέσω του αλγορίθμου SiteMap. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ) σε κόκκινο και μπλε χρώμα. ....	58
Εικόνα 3-9: Σύγκριση αποτελεσμάτων του αντιγράφου 1 του συστήματος φυσικού τύπου (μαύρο χρώμα) με την Q61H SI/II κοιλότητα πρόσδεσης, όπως προσδιορίστηκε από τους Cruz-Migoni et al[32] (μπλε χρώμα). Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ). ....	59
Εικόνα 3-10: Κοιλότητες πρόσδεσης στην κυρίαρχη δομή του αντιγράφου 1 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ). ....	60
Εικόνα 3-11: Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 2 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ). ....	63
Εικόνα 3-12: Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 3 του συστήματος φυσικού τύπου KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ). ....	65
Εικόνα 3-13: Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 4 του συστήματος φυσικού τύπου KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή	

SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	66
Εικόνα 3-14: Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 5 του συστήματος φυσικού τύπου KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	68
Εικόνα 3-15: Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 6 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	70
Εικόνα 3-16: Το σύστημα διμερούς μεταλλαγμένης πρωτεΐνης (G12D) KRAS-4B. Το διμερές KRAS-4B και οι πρωτεΐνες Raf απεικονίζονται με τη μορφή κορδελών (ribbons) σε γαλάζιο και μπλε χρώμα αντίστοιχα. Η μεμβράνη απεικονίζεται με τη μορφή επιφάνειας σε γκρι χρώμα. ..	72
Εικόνα 3-17: Αρχείο script που χρησιμοποιήθηκε για την διαίρεση των αρχείων τροχιάς των συστημάτων φυσικού τύπου διμερούς KRAS-4B .....	73
Εικόνα 3-18: Αποτελέσματα κοιλοτήτων πρόσδεσης για την κυρίαρχη δομή του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	74
Εικόνα 3-19: Αποτελέσματα κοιλοτήτων πρόσδεσης για τις κυρίαρχες δομές των υπο-τροχιών του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	77
Εικόνα 3-20: Σύγκριση αποτελεσμάτων κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου SiteMap (αριστερά) και μέσω του αλγορίθμου FTMap (δεξιά) για την κυρίαρχη δομή του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	78
Εικόνα 3-21: Ορθοστερικές κοιλότητες πρόσδεσης προσδιορισμένες στο μεταλλαγμένο (G12D) διμερές KRAS-4B, με χρήση του αλγορίθμου SiteMap. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ) κόκκινου και μπλε χρώματος. ....	79
Εικόνα 3-22: Σύγκριση της κοιλότητας G12D SI/II [29] [30] [31] (πράσινο χρώμα) με γειτονική κοιλότητα στο διμερές G12D KRAS-4B, προσδιορισμένη μέσω του SiteMap (ροζ χρώμα). Το	

διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι δύο κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	79
Εικόνα 3-23: Σύγκριση της βιβλιογραφικά προσδιορισμένης κοιλότητας Q61H SI/II [32](μπλε χρώμα) με γειτονική κοιλότητα στο διμερές G12D KRAS-4B (ροζ χρώμα), προσδιορισμένη μέσω του SiteMap. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι δύο κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	80
Εικόνα 3-24: Σύγκριση της βιβλιογραφικά προσδιορισμένης κοιλότητας G12C SII [26] [27] [28] (γκρι χρώμα) με γειτονική κοιλότητα στο διμερές G12D KRAS-4B (ροζ χρώμα), προσδιορισμένη μέσω του SiteMap. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι δύο κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).....	80
Εικόνα 3-25: Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες πρόσδεσης με τη μορφή χάρτη επιφάνειας (δίχτυ). ....	81
Εικόνα 3-26: Σύγκριση των αποτελεσμάτων στα συστήματα φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες πρόσδεσης με τη μορφή χάρτη επιφάνειας (πορτοκαλί δίχτυ). ....	84



## Ευρετήριο πινάκων

Πίνακας 3-1: Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 1 .....	62
Πίνακας 3-2: Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 2 .....	64
Πίνακας 3-3: Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 3 .....	66
Πίνακας 3-4: Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 4 .....	67
Πίνακας 3-5: Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 5 .....	69
Πίνακας 3-6: Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 6 .....	71
Πίνακας 3-7: Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο μεταλλαγμένο διμερές KRAS-4B .....	82

## Ευρετήριο διαγραμμάτων

Διάγραμμα 3-1: Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 1 του συστήματος φυσικού τύπου KRAS-4B.....	51
Διάγραμμα 3-2: Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 1 του συστήματος φυσικού τύπου) .....	61
Διάγραμμα 3-3: Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 2 του συστήματος φυσικού τύπου) .....	63
Διάγραμμα 3-4: Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 3 του συστήματος φυσικού τύπου) .....	65
Διάγραμμα 3-5: Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 4 του συστήματος φυσικού τύπου) .....	67
Διάγραμμα 3-6: Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 5 του συστήματος φυσικού τύπου) .....	68
Διάγραμμα 3-7: Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 6 του συστήματος φυσικού τύπου) .....	70
Διάγραμμα 3-8: Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το σύστημα διμερούς της μεταλλαγμένης (G12D) KRAS-4B .....	73
Διάγραμμα 3-9: Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (μεταλλαγμένη G12D KRAS-4B) .....	81
Διάγραμμα 6-1: Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 2 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 5 διατηρείται σε όλο το μήκος της τροχιάς. ....	94
Διάγραμμα 6-2: Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 3 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 7 διατηρείται στο μεγαλύτερο μέρος της τροχιάς.....	94
Διάγραμμα 6-3: Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 4 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 4 διατηρείται στο μεγαλύτερο μέρος της τροχιάς.....	95
Διάγραμμα 6-4: Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 5 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 3 διατηρείται στο μεγαλύτερο μέρος της τροχιάς.....	95

Διάγραμμα 6-5: Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 6 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 3 διατηρείται σε όλο το μήκος της τροχιάς. ....96

## Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Ίδρυμα Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών υπό την επίβλεψη της Ερευνήτριας Β' Δρ. Ζωής Κούρνια και σε συνεργασία με τη φαρμακευτική εταιρία Novartis κατά το ακαδημαϊκό έτος 2020-2021. Λόγω των ιδιαίτερων συνθηκών που δημιούργησε η πανδημία του COVID-19, το μεγαλύτερο κομμάτι της πραγματοποιήθηκε εξ' αποστάσεως.

Αρχικά, θα ήθελα να ευχαριστήσω τον Αναπληρωτή Καθηγητή Δρ. Ευάγγελο Τόπακα για την ευκαιρία που μου έδωσε να συνεργαστώ με το Ίδρυμα Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών και τη Δρ. Ζωή Κούρνια, γνωρίζοντας έτσι τον τόσο ενδιαφέροντα τομέα της Υπολογιστικής Χημείας. Τον ευχαριστώ επίσης για όλες τις γνώσεις που μου προσέφερε καθ' όλη τη διάρκεια των σπουδών μου. Θα ήθελα να ευχαριστήσω τη Δρ. Ζωή Κούρνια για την εμπιστοσύνη που μου έδειξε και για τη συνεχή επιστημονική καθοδήγηση κατά την εκπαίδευσή μου στο εργαστήριό της και κατά της εκπόνηση της διπλωματικής μου εργασίας. Την ευχαριστώ επίσης και για την ηθική παρότρυνση που με βοήθησε να συνεχίσω και να ολοκληρώσω την παρούσα εργασία. Θα ήθελα να ευχαριστήσω ακόμα τη φαρμακευτική εταιρία Novartis και συγκεκριμένα τον Δρ. John Manchester και τον Δρ. Camilo Velez-Vega για τη δυνατότητα να εκπονήσω τη διπλωματική μου εργασία σε ερευνητικό σχέδιο της εταιρίας.

Προχωρώντας, θα ήθελα να ευχαριστήσω τη Δρ. Σοφία Κυριακίδη και τον Δρ. Μιχάλη Παπαδουράκη, για τη βοήθεια και την υποστήριξη που μου παρείχαν αλλά και για τις συζητήσεις και τις συμβουλές που με βοήθησαν να ξεπεράσω κάθε δυσκολία που προέκυψε κατά την εξέλιξη της διπλωματικής μου εργασίας. Επιπλέον, θα ήθελα να ευχαριστήσω τον Δρ. Ιωάννη Ανδρεαδέλη για την παροχή δεδομένων και αποτελεσμάτων, τα οποία χρησιμοποιήθηκαν για τους σκοπούς της παρούσας εργασίας, καθώς και για τις πολύτιμες συμβουλές του. Ευχαριστώ επίσης τον υποψήφιο Διδάκτορα Αλέξη Χατζηγούλα για τη βοήθειά του σε τεχνικά ζητήματα που προέκυψαν κατά την εκπαίδευσή μου στο εργαστήριο και κατά την εκπόνηση της παρούσας εργασίας. Ομοίως, θα ήθελα να ευχαριστήσω και τα υπόλοιπα μέλη του εργαστηρίου, την υποψήφια διδάκτορα Ζωή Διαμάντη, την Ισιδώρα Διακογιαννάκη και τον Θεμιστοκλή Αϊβαλιώτη για τη στήριξή τους και τις συμβουλές που με βοήθησαν να βελτιωθώ ερευνητικά. Φυσικά, δεν θα μπορούσα να παραλείψω την Ελευθερία Κελεφιώτη-

Στρατηδάκη, για την ψυχολογική υποστήριξη και τη διαρκή της ενθάρρυνση, κατά την εκπόνηση των διπλωματικών μας εργασιών.

Με την παρούσα εργασία, ολοκληρώνεται ο κύκλος των προπτυχιακών μου σπουδών. Θα ήθελα λοιπόν να ευχαριστήσω όλους όσους ήταν μαζί μου σε αυτά τα 5 και πλέον χρόνια. Ευχαριστώ τις φίλες μου και συναδέλφους, για όλες τις όμορφες αναμνήσεις που δημιουργήσαμε μαζί στα χρόνια αυτά, καθώς επίσης και την οικογένεια του Chemecop, μέσα από το οποίο μπόρεσα να προσφέρω περισσότερα στην κοινότητα των Χημικών Μηχανικών. Τέλος, θα ήθελα να εκφράσω την αμέριστη ευγνωμοσύνη μου στην οικογένειά μου, τον πατέρα μου Νίκο, τη μητέρα μου Αντωνία και τον αδερφό μου Κωνσταντίνο, οι οποίοι ήταν στο πλευρό μου καθ'όλη τη διάρκεια των φοιτητικών μου χρόνων.

## Περίληψη

Οι πρωτεΐνες RAS είναι υδρολάσες της ένωσης GTP, οι οποίες ενεργοποιούν μονοπάτια σηματοδότησης, απαραίτητα για τον κυτταρικό πολλαπλασιασμό. Συγκεκριμένα, ο διμερισμός της KRAS παρουσία της πρωτεΐνης Raf στην κυτταρική μεμβράνη ενεργοποιεί μονοπάτια σηματοδότησης που σχετίζονται με την κυτταρική ανάπτυξη και τον κυτταρικό πολλαπλασιασμό, όπως το μονοπάτι Ras-Raf-MEK-Erk. Οι μεταλλάξεις στις πρωτεΐνες RAS που παρεμποδίζουν την υδρόλυση του GTP έχουν ως αποτέλεσμα τη σταθεροποίηση της ενεργής τους μορφής, η οποία είναι συνδεδεμένη με την ένωση GTP, και ως εκ τούτου την υπερ-ενεργοποίηση της σηματοδότησης που μπορεί να οδηγήσει σε ογκογένεση.

Η πρωτεΐνη KRAS-4B αποτελεί το πιο συχνά μεταλλαγμένο ισόμορφο της οικογένειας πρωτεϊνών RAS, συνιστώντας το 86% όλων των RAS μεταλλάξεων. Οι μεταλλάξεις στη γλυκίνη 12 της KRAS-4B είναι οι πιο συχνές και η μετάλλαξη της γλυκίνης 12 σε ασπαρτικό οξύ (G12D) είναι η πιο συχνά εμφανιζόμενη και για το λόγο αυτό επιλέχθηκε προς μελέτη στην παρούσα εργασία.

Η φαρμακευτική στόχευση της μεταλλαγμένης KRAS-4B μέσω ανταγωνισμού της πρόσδεσης των ενδογενών προσδεμάτων GTP, GDP στις ορθοστερικές κοιλοότητες του διμερούς δεν είναι εφικτή, καθώς τα υποστρώματα αυτά εμφανίζουν υψηλή χημική συγγένεια με την KRAS-4B. Για το λόγο αυτό, η έρευνα κατευθύνεται στην αναζήτηση αλλοστερικών κοιλοτήτων στην επιφάνεια της KRAS-4B. Σκοπός της παρούσας εργασίας είναι η εύρεση αλλοστερικών κοιλοτήτων στη διεπιφάνεια του διμερούς KRAS-4B, στις οποίες θα προσδεθούν μικρά μόρια με στόχο τη παρεμπόδιση του σχηματισμού του και ως εκ τούτου της ογκογένεσης. Η ανάλυση αυτή δεν υπάρχει μέχρι τώρα στη βιβλιογραφία και μπορεί να αποκαλύψει νέες προοπτικές για τη στόχευση του διμερούς KRAS-4B.

Ο εντοπισμός των κοιλοτήτων πρόσδεσης πραγματοποιείται σε κυρίαρχες δομές του διμερούς KRAS-4B, οι οποίες προκύπτουν από ομαδοποίηση πρωτεϊνικών δομών που προέρχονται από προσομοιώσεις Μοριακής Δυναμικής σε συστήματα φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B, σε μοντέλο κυτταρικής μεμβράνης. Οι κοιλοότητες που προκύπτουν εξετάζονται ως προς τη χρονική τους μεταβολή, προκειμένου να διευκρινιστεί αν διατηρούνται σταθερές κατά τη διάρκεια της προσομοίωσης.

Από την ανάλυση, προκύπτει αρχικά ότι οι αλγόριθμοι που χρησιμοποιήθηκαν εντοπίζουν τις ίδιες κοιλότητες πρόσδεσης στις δομές διμερών KRAS-4B, κάποιες από τις οποίες έχουν προηγουμένως προσδιοριστεί βιβλιογραφικά. Μέσω σύγκρισης των αποτελεσμάτων στα δύο συστήματα φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B, μια σταθερή κοιλότητα πρόσδεσης με κατάλληλη μέση τιμή όγκου εντοπίστηκε σε όλες τις δομές των διμερών. Ωστόσο, η εκλεκτική στόχευση της κοιλότητας στο μεταλλαγμένο (G12D) διμερές KRAS-4B, χρειάζεται να διερευνηθεί περαιτέρω. Στην κατεύθυνση αυτή, η περαιτέρω έρευνα των αμινοξέων και του δικτύου μορίων νερού των κοιλοτήτων πρόσδεσης, καθώς και η μελέτη ασταθών διαμορφώσεων του διμερούς KRAS-4B μπορούν να αποκαλύψουν νέες προοπτικές.

# Abstract

RAS proteins are small hydrolases of GTP, which mediate signaling pathways necessary for cell proliferation. In particular, KRAS dimerization in the cell membrane in the presence of Raf protein activates signaling pathways associated with cell growth and cell proliferation, such as the Ras-Raf-MEK-Erk pathway. Mutations in RAS proteins, whose effect is GTP hydrolysis inhibition, result in stabilization of RAS active form, which is bound to their endogenous ligand GTP, and hence cause overactivation of signaling which may lead to oncogenesis.

KRAS-4B protein is the most frequently mutated isoform of the RAS family of proteins, accounting for 86% of all RAS mutations. Mutations in glycine 12 of KRAS-4B are the most common and mutation of glycine 12 to aspartic acid (G12D) is the most common, which is why it is studied in this thesis.

Pharmacological targeting of the mutant KRAS-4B through competition for the binding of the endogenous ligands GTP, GDP to the orthosteric cavities of the dimer is not possible, as they show high chemical affinity for KRAS-4B. For this reason, drug design efforts are directed to the search for allosteric binding pockets on the surface of KRAS-4B. The aim of the present work is to discover allosteric binding pockets located on the dimerization interface of KRAS-4B dimer, in which small inhibitors can bind in order to disrupt dimerization, thus blocking oncogenesis. So far, this analysis cannot be found in the literature and may reveal new perspectives for the targeting of KRAS-4B dimer.

Identification of the binding pockets is performed in dominant structures of KRAS-4B dimer, emerged from clustering of dimer structures derived from Molecular Dynamics simulations in wild-type and mutant (G12D) KRAS-4B systems, in a cell membrane model. Resulting binding pockets are examined in terms of their temporal variation, in order to determine whether their volume is maintained throughout simulation time.

By comparing the algorithms used, it is concluded that they both identify the same binding pockets in KRAS-4B dimer structures and detect binding pockets that have been previously reported in the literature. Also, through comparison of the pockets identified in wild-type and mutant (G12D) systems, a stable binding pocket with appropriate average value of volume is discovered in all dominant dimer structures. However, the possibility of selectively targeting the



mutant (G12D) dimer has to be further investigated. In this regard, further research into the amino acids and the water molecules network of the binding pockets, as well as the study of unstable configurations of KRAS-4B dimer may reveal new perspectives.

# Ορολογία και συντομογραφίες

**ADME:** Absorption distribution metabolism and excretion, απορρόφηση κατανομή μεταβολισμός και απέκκριση

**Allostery:** αλλοστερισμός

**Allosteric modulators:** αλλοστερικοί μετατροπείς

**C:** Cysteine, κυστεΐνη

**D:** Aspartic acid, ασπαρτικό οξύ

**Docking:** προσαρμογή του προσδέματος στο ενεργό κέντρο

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

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

**Drug-likeness:** φαρμακο-ομοιότητα

**Effectors:** πρωτεΐνες σηματοδότησης

**Farnesyl-transferase:** φαρνεσυλο-τρανσφεράση

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

**Fragment screening:** έλεγχοι δέσμευσης χημικών θραυσμάτων

**G:** Glycine, γλυκίνη

**G domain:** catalytic domain, καταλυτικός τομέας

**G12C:** Μετάλλαξη της γλυκίνης 12 σε κυστεΐνη

**G12D:** Μετάλλαξη της γλυκίνης 12 σε ασπαρτικό οξύ

**G12V:** Μετάλλαξη της γλυκίνης 12 σε βαλίνη

**GAP:** GTPase activating-accerating protein, πρωτεΐνες ενεργοποίησης/επιτάχυνσης της GTP-άσης

**GDP:** διφωσφορική γουανοσίνη

**GEF:** Guanine Nucleotide Exchange Factor, Παράγοντες Ανταλλαγής Νουκλεοτιδίων Γουανίνης

**GTP:** τριφωσφορική γουανοσίνη

**H:** Histidine, ιστιδίνη

**Hit:** οδηγός-δομή

**HRAS:** Harvey Rat Sarcoma Virus

**HTS:** High Throughput Screening, ταχύς βιολογικός έλεγχος/διαλογή υψηλής απόδοσης

**HVR:** Hypervariable C-terminal Region, υπερμεταβλητή καρβοξυ-τελική περιοχή

**Intensive electrophile screening:** ηλεκτρόφιλος έλεγχος δέσμευσης μορίων

**KRAS:** Kirsten rat sarcoma virus

**Lead-compound:** ένωση-οδηγός

**Ligand:** πρόσδεμα

**NMR:** Nuclear Magnetic Resonance, πυρηνικός μαγνητικός συντονισμός

**NMR-based Saturation Transfer Difference:** NMR διαφοράς μεταφοράς κορεσμού

**NRAS:** Neuroblastoma Rat Sarcoma Virus

**NSCLC:** Non-small-cell lung carcinoma, μη μικροκυτταρικός καρκίνος του πνεύμονα (ΜΜΚΠ)

**PDB:** Protein Data Bank, βάση δεδομένων πρωτεϊνικών δομών

**PIP2:** Phosphatidylinositol 4,5-bisphosphate

**Q:** Glutamine, γλουταμίνη

**Raf:** Rapidly Accelerated Fibrosarcoma, ταχέως επιταχυνόμενο ινοσάρκωμα

**RAS:** Rat Sarcoma virus, ιός σαρκώματος αρουραίου

**RMSD:** Root Mean Square Deviation, τετραγωνική ρίζα της μέσης τιμής των τετραγώνων της απόκλισης

**SBDD:** Structure-based Drug Design, σχεδιασμός φαρμάκων με βάση τη δομή της πρωτεΐνης-στόχου

**Scoring functions:** συναρτήσεις βαθμολόγησης

**SOS:** Son of Sevenless

**Surface Plasmon Resonance (SPR):** συντονισμός επιφανειακών πλάσμονίων

**V:** Valine, βαλίνη

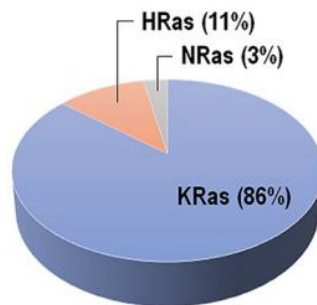
**WT:** Wild-type, φυσικού τύπου

**X-ray crystallography:** κρυσταλλογραφία ακτίνων Χ

# 1 ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ

## 1.1 Εισαγωγή

Η οικογένεια των πρωτεϊνών RAS αποτελείται από τις πρωτεΐνες HRAS, NRAS και τις δύο παραλλαγές της KRAS, KRAS-4A και KRAS-4B. Οι πρωτεΐνες RAS είναι μικρές GTP-άσες, οι οποίες ενεργοποιούν μονοπάτια σηματοδότησης, απαραίτητα για τον κυτταρικό πολλαπλασιασμό[1]. Περίπου το ένα τρίτο όλων των καρκίνων που εμφανίζονται σε ανθρώπους υποθάλπουν μια ογκογενή μετάλλαξη σε τουλάχιστον ένα γονίδιο που κωδικοποιεί πρωτεΐνες RAS, καθιστώντας τις ως τις πιο συχνά μεταλλαγμένες ογκοπρωτεΐνες[2]. Η πρωτεΐνη KRAS αποτελεί το πιο συχνά μεταλλαγμένο ισόμορφο στους καρκίνους που εμφανίζονται στον άνθρωπο, συνιστώντας το 86% όλων των RAS μεταλλάξεων, ενώ σε κάποια αδenoκαρκινώματα του παγκρέατος είναι μεταλλαγμένη σχεδόν στο 90% των όγκων. Παρόλο που οι πρωτεΐνες RAS αποτελούν μια εξαιρετικά διατηρημένη οικογένεια πρωτεϊνών, η πρωτεΐνη KRAS χαρακτηρίζεται από μοναδικές ογκογενείς ιδιότητες, υψηλή γονιδιακή έκφραση και χαμηλό ρυθμό επιδιόρθωσης του DNA, σε σχέση με τις HRAS και NRAS.

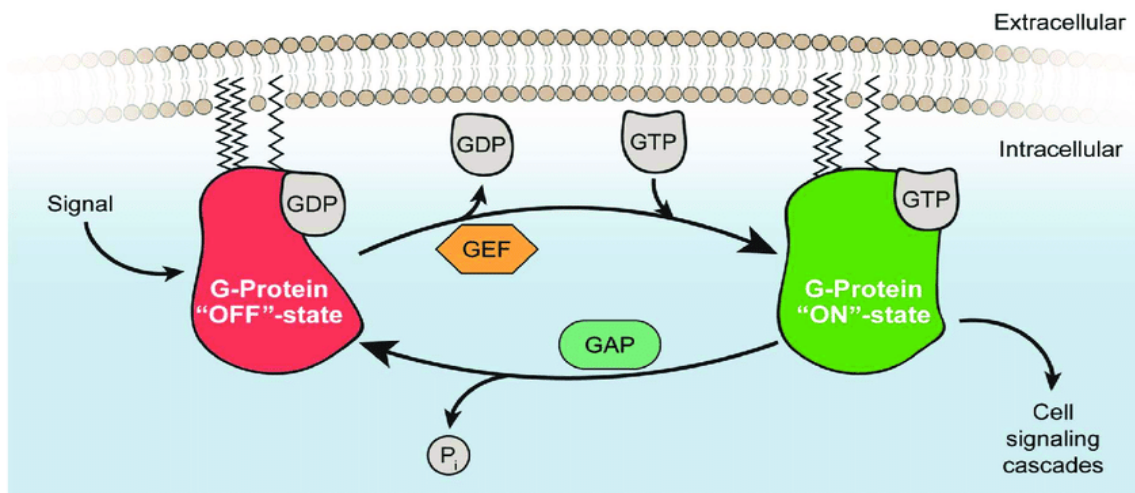


Ras mutation in cancer

**Εικόνα 1-1:** Οι μεταλλάξεις των πρωτεϊνών RAS στον καρκίνο

### 1.1.1 Δομή και καταλυτική λειτουργία των πρωτεϊνών RAS

Η κύρια λειτουργία των πρωτεϊνών RAS είναι η υδρόλυση της ένωσης GTP προς την ένωση GDP. Η ενεργή μορφή των πρωτεϊνών RAS είναι συνδεδεμένη με την ένωση GTP, ενώ η ανενεργή με την ένωση GDP. Κατ'αυτόν τον τρόπο, οι πρωτεΐνες RAS λειτουργούν ως δυαδικοί διακόπτες που εναλλάσσονται μεταξύ της ενεργούς και της ανενεργούς τους διαμόρφωσης. Ο κύκλος της υδρόλυσης ρυθμίζεται από τους Παράγοντες Ανταλλαγής Νουκλεοτιδίων Γουανίνης (Guanine Nucleotide Exchange Factors ή GEFs), πρωτεΐνες που εμφανίζονται ως απάντηση σε διάφορους υποδοχείς. Οι GEFs προάγουν την απελευθέρωση της ένωσης GDP υπέρ της ένωσης GTP, η οποία εμφανίζει δεκαπλάσια επίπεδα κυτταρικής συγκέντρωσης[3]. Κατ'αντιστοιχία, οι πρωτεΐνες ενεργοποίησης/επιτάχυνσης της GTP-άσης (GTPase activating/accelerating proteins ή GAPs) επιταχύνουν την υδρολυτική δραστηριότητα των πρωτεϊνών RAS, η οποία είναι εγγενώς χαμηλή, έως και τάξεις μεγέθους, οδηγώντας τις προς την ανενεργή τους δομή που προκύπτει μετά την υδρόλυση και συνδέεται με την ένωση GDP[4]. Οι ογκογενείς μεταλλάξεις στις πρωτεΐνες RAS παρεμποδίζουν την υδρόλυση που μεσολαβείται από τις GAPs, έχοντας ως αποτέλεσμα τη σταθεροποίηση της ενεργής μορφής που είναι συνδεδεμένη με την ένωση GTP και ακολούθως προκαλώντας υπερ-ενεργοποίηση της σηματοδότησης (upstream signaling).

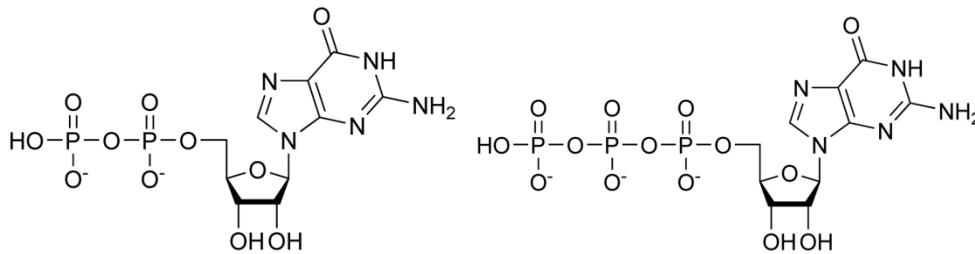


**Εικόνα 1-2:** Ο κύκλος της υδρόλυσης, μεσολαβούμενος από τις πρωτεΐνες GAP, GEF [5]

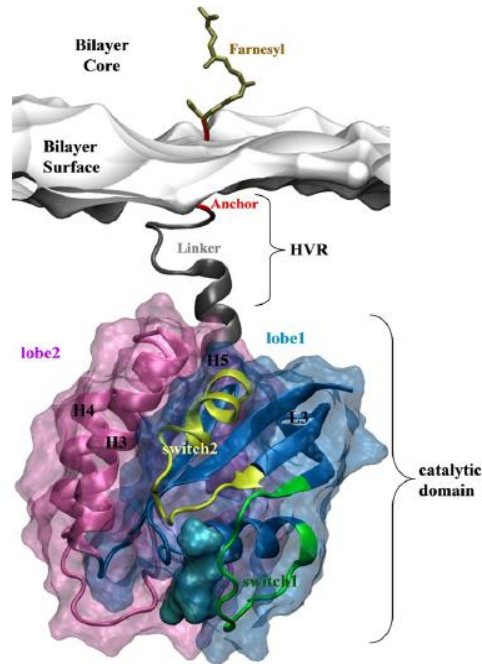
Οι πρωτεΐνες RAS αποτελούνται από τον καταλυτικό τομέα (αμινοξέα 1 έως 166) και μία υπερμεταβλητή καρβοξυ-τελική περιοχή ή αλλιώς HVR (hypervariable region) που βρίσκεται στο καρβοξυ-τελικό άκρο της πρωτεΐνης (αμινοξέα 167 έως 188/189). Η περιοχή HVR είναι εξαιρετικά ευέλικτη και διαφοροποιείται αρκετά μεταξύ των πρωτεϊνών RAS. Ο καταλυτικός τομέας, γνωστός και με την ονομασία G-domain, αποτελείται από τις εξής δύο δομικά διακριτές περιοχές:

- Περιοχή “effector lobe” (αμινοξέα 1 έως 86): στην περιοχή αυτή των πρωτεϊνών RAS, συνδέονται πρωτεΐνες που αφορούν τη διαδικασία της σηματοδότησης (effectors)
- Περιοχή “allosteric lobe” (αμινοξέα 87-166): η περιοχή αυτή εμπλέκεται σε αλληλεπιδράσεις των πρωτεϊνών RAS με την μεμβράνη

Ακόμα, ο καταλυτικός τομέας περιλαμβάνει τις περιοχές Switch I (SI) και Switch II (SII), οι οποίες συνδέονται μέσω του φωσφορικού που περιέχει το πρόσδεμα GTP και υφίστανται σημαντικές αλλαγές στη διαμόρφωσή τους κατά την υδρόλυση του GTP προς το GDP. Η απελευθέρωση του φωσφορικού που περιέχει η ένωση GTP, μετά την υδρόλυσή της, οδηγεί στην ανενεργή δομή των πρωτεϊνών RAS, η οποία είναι συνδεδεμένη με την ένωση GDP[6]. Η ενεργή μορφή της πρωτεΐνης KRAS-4B μπορεί να σταθεροποιηθεί στην κυτταρική μεμβράνη μετά την μετα-μεταφραστική φαρνεσυλίωση στην κυστεΐνη του καρβοξυ-τελικού της άκρου, μία από τις πιο συχνές μετα-μεταφραστικές τροποποιήσεις της KRAS-4B.

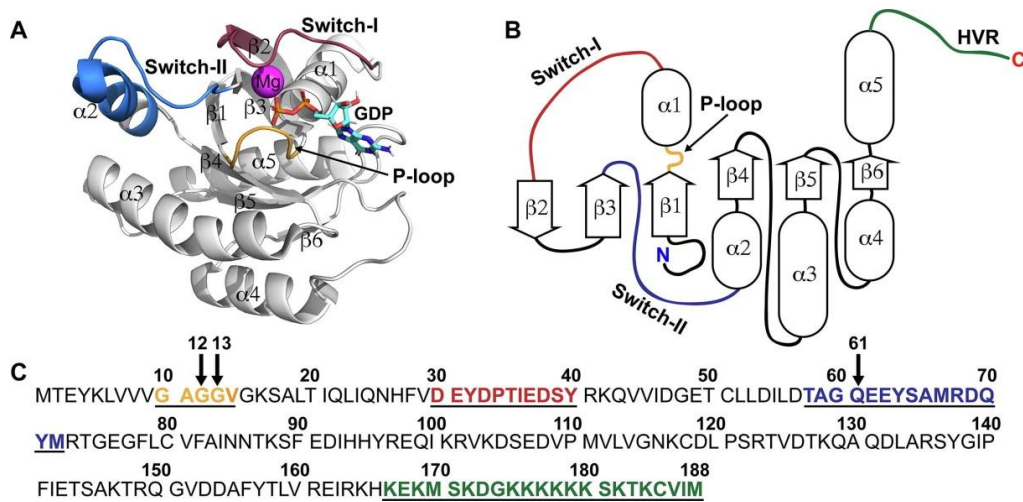


**Εικόνα 1-3:** Χημική δομή της ένωσης GDP (αριστερά) και της ένωσης GTP (δεξιά)



**Εικόνα 1-4:** Δομή της KRAS-4B. lobe 1: effector lobe και lobe 2: allosteric lobe [7]

Στην φυσικού τύπου πρωτεΐνη KRAS-4B που είναι συνδεδεμένη με την ένωση GDP, η περιοχή HVR αλληλεπιδρά στενά με τον καταλυτικό τομέα, περιορίζοντας την πρωτεΐνη σε μια εγγενώς αυτοανασταλμένη κατάσταση. Αντίθετα, όταν η πρωτεΐνη είναι συνδεδεμένη με την ένωση GTP, οι αλληλεπιδράσεις είναι πιο χαλαρές, οδηγώντας στην απελευθέρωση της περιοχής HVR και στην επακόλουθη σύνδεσή της με την κυτταρική μεμβράνη[8].

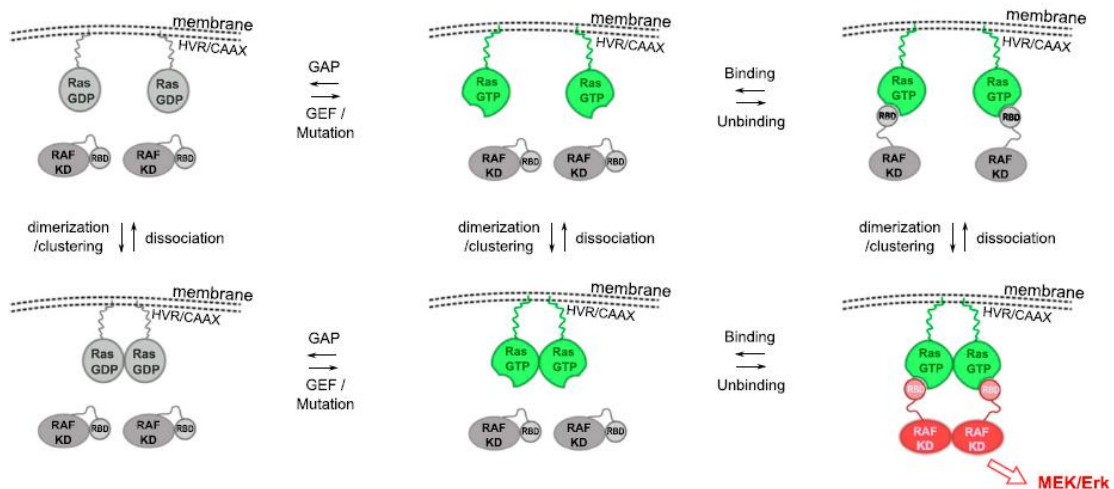


**Εικόνα 1-5:** Δομικές περιοχές της KRAS-4B (A,B) με τις αντίστοιχες περιοχές αμινοξικών ακολουθιών (C) [9]

### 1.1.2 Η πρωτεΐνη Raf και η σημασία της στον διμερισμό της KRAS-4B

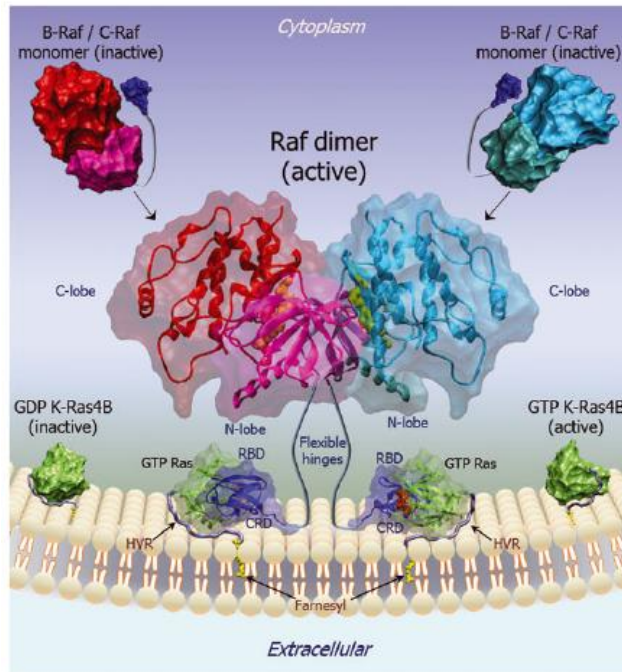
Η αλληλεπίδραση της KRAS-4B με την πρωτεΐνη Raf ενεργοποιεί μονοπάτια σηματοδότησης που σχετίζονται με την κυτταρική ανάπτυξη και τον κυτταρικό πολλαπλασιασμό, όπως το μονοπάτι Ras-Raf-MEK-Erk. Ο διμερισμός της KRAS στη μεμβράνη παίζει σημαντικό ρόλο για τη σύνδεση με τη Raf και την επακόλουθη ρύθμιση του μονοπατιού MEK-Erk. Όπως έχει προταθεί βιβλιογραφικά, η πρωτεΐνη Raf χρησιμοποιεί τη RAS ως ικρίωμα για να αγκυροβολήσει και να δημιουργήσει ομοδιμερή, διαδικασία αναγκαία για την ενεργοποίησή της[10].

Η διαδικασία έχει ως εξής. Μετά τη σταθεροποίηση της KRAS-4B στην κυτταρική μεμβράνη, ακολουθεί η σύνδεση πληθώρας πρωτεϊνών σηματοδότησης (effectors) στην περιοχή “effector lobe”, μία από τις οποίες είναι και η κινάση Raf (Rapidly Accelerated Fibrosarcoma). Συγκεκριμένα, η KRAS-4B προσδέεται στη μεμβράνη μέσω της περιοχής της HVR. Όταν η KRAS-4B βρίσκεται συνδεδεμένη με την ένωση GDP (ανενεργή μορφή), η Raf δε μπορεί να συνδεθεί σε αυτή και ως εκ τούτου υπάρχει σε μια ανενεργή μορφή στην κυττοσόλη, αλληλεπιδρώντας με τα λιπίδια της μεμβράνης. Συνδεδεμένη με την ένωση GTP, η KRAS-4B μπορεί να συνδεθεί με τη Raf ή να σχηματίσει ομοδιμερή. Όταν τα δύο ετεροδιμερή KRAS-Raf διμεριστούν ή όταν στο διμερές KRAS-GTP προσδεθούν δύο μόρια Raf, το ομοδιμερές Raf-Raf ενεργοποιεί την κινάση Raf και επακόλουθα το μονοπάτι σηματοδότησης MEK/Erk (MAPK) [11].



**Εικόνα 1-6:** Μηχανισμοί αλληλεπίδρασης RAS-Raf για την ενεργοποίηση του μονοπατιού MEK/Erk [11]

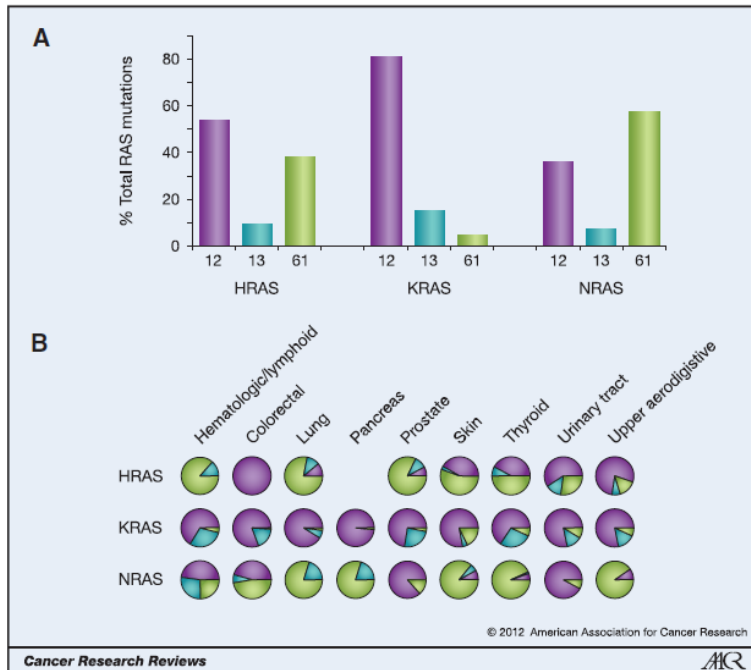




**Εικόνα 1-7:** Ενεργοποίηση της κινάσης Raf από την πρωτεΐνη RAS [12]

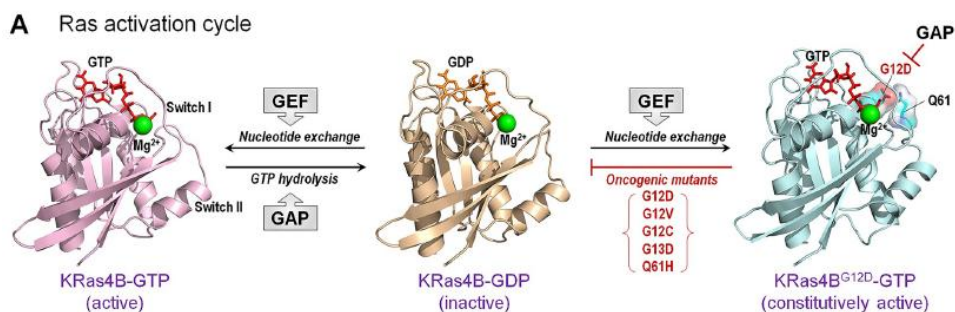
### 1.1.3 Ο ρόλος της KRAS-4B στον καρκίνο και η μετάλλαξη G12D

Παρά τη μεγάλη ομολογία μεταξύ των καταλυτικών περιοχών των ισόμορφων RAS (NRAS, HRAS, KRAS) και το γεγονός ότι οι μεταλλάξεις προκύπτουν σε αυτές τις περιοχές, η συχνότητα και η κατανομή τους στα ισόμορφα RAS ποικίλλει. Το ισόμορφο KRAS είναι το πιο συχνά μεταλλαγμένο (86%) σε σχέση με τα υπόλοιπα, ενώ ακολουθεί το ισόμορφο NRAS (11%) και τέλος το ισόμορφο HRAS (3%). Το 98% όλων των μεταλλάξεων βρίσκονται στο ενεργό κέντρο των RAS και συγκεκριμένα στα αμινοξέα γλυκίνη 12 (Gly12), γλυκίνη 13 (Gly13) και γλουταμίνη 61, παρεμποδίζοντας την υδρόλυση της ένωσης GTP[2][13]. Εξ αυτών, οι μεταλλάξεις στη γλυκίνη 12 της KRAS είναι οι πιο συχνές (89%). Ακολουθούν οι μεταλλάξεις στη γλυκίνη 13 (9%) και στη γλουταμίνη 61 (1%). Η μετάλλαξη της γλυκίνης 12 σε ασπαρτικό οξύ (G12D) είναι η πιο συχνά εμφανιζόμενη, για το λόγο αυτό έχουμε επιλέξει να τη μελετήσουμε στα πλαίσια αυτής της εργασίας[2].



**Εικόνα 1-8:** Συχνότητα εμφάνισης των μεταλλάξεων στα αμινοξέα 12, 13 και 61 στα διάφορα ισόμορφα RAS [2]

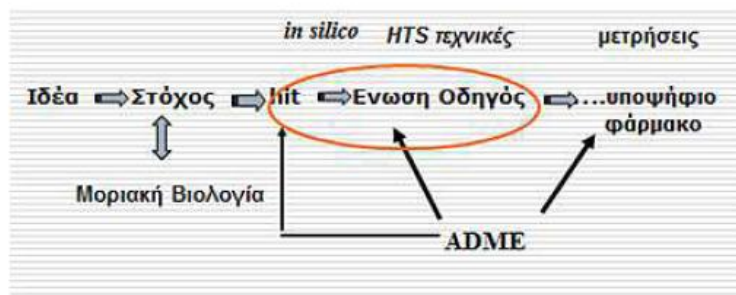
Οι μεταλλάξεις στα αμινοξέα γλυκίνη 12 και γλυκίνη 13 ορίζονται ως ενεργοποιητικές, καθώς δρουν παρεμποδίζοντας την αλληλεπίδραση της KRAS με τις πρωτεΐνες GAP, ενώ δεν έχουν επίδραση στην αλληλεπίδραση με τις πρωτεΐνες GEF και τις πρωτεΐνες σηματοδότησης. Συγκεκριμένα, η αντικατάσταση της γλυκίνης στη θέση 12 με οποιοδήποτε άλλο αμινοξύ εκτός της προλίνης προκαλεί παρεμβολές στη διαδικασία πρόσδεσης των πρωτεϊνών GAP στην KRAS, με τη μορφή στερεοχημικής παρεμπόδισης που αφορά στο αμινοξύ αργινίνη των πρωτεϊνών GAP. Κατ'αυτόν τον τρόπο, οι μεταλλάξεις οδηγούν στη συσσώρευση του συμπλόκου KRAS-GTP και ως εκ τούτου στην υπερ-ενεργοποίηση των μετέπειτα οδών σηματοδότησης.



**Εικόνα 1-9:** Οι μεταλλάξεις στα αμινοξέα 12,13 και 61 εμποδίζουν την αντίδραση υδρόλυσης του GTP από την KRAS-4B, οδηγώντας σε μία μόνιμα ενεργή μορφή της πρωτεΐνης [14]

## 1.2 Ορθολογικός σχεδιασμός φαρμάκων

Η διαδικασία για την ανακάλυψη νέων φαρμακομορίων έχει ως εξής. Αρχικά, επιλέγεται ο μοριακός στόχος για την αντιμετώπιση μιας συγκεκριμένης παθολογικής κατάστασης. Ακολούθως, η έρευνα προσανατολίζεται προς την ανακάλυψη μιας αρχικής βιοδραστικής οδηγού-δομής (hit) ή μιας ένωσης-οδηγού (lead compound), ανάλογα με τη στρατηγική που ακολουθείται. Η αρχική βιοδραστική οδηγός-δομή (hit) προκύπτει μέσω ταχύ βιολογικού ελέγχου/διαλογής υψηλής απόδοσης (HTS), διαδικασίας απαραίτητης για τον έλεγχο της ύπαρξης δραστηριότητας μιας ένωσης σε καθορισμένη συγκέντρωση. Οι ενώσεις-οδηγοί (lead compounds) φέρουν τη φαρμακοφόρο δομή, δηλαδή το σύνολο των ηλεκτρονικών και στερεικών χαρακτηριστικών που απαιτούνται για την επίτευξη των βέλτιστων αλληλεπιδράσεων με ένα συγκεκριμένο βιολογικό στόχο και την πρόκληση βιολογικού αποτελέσματος. Το επόμενο βήμα αφορά τη βελτιστοποίηση της ένωσης-οδηγού, μέσω σχεδιασμού, σύνθεσης και βιολογικής αποτίμησης μιας σειράς συγγενών ενώσεων. Στο σημείο αυτό, λαμβάνονται υπόψη οι φυσικοχημικές ιδιότητες και οι ιδιότητες ADME μέσω υπολογιστικών τεχνικών (in silico) ή με ταχείς βιολογικούς ελέγχους/διαλογές υψηλής απόδοσης (HTS). Για την τελική επιλογή των ενώσεων είναι απαραίτητοι ακριβείς πειραματικοί προσδιορισμοί.



Εικόνα 1-10: Βασικά στάδια ανακάλυψης φαρμάκων [15]

Ο ορθολογικός σχεδιασμός φαρμάκων βασίζεται στη γνώση της παθολογίας της ασθένειας, στον προσδιορισμό των μικρομορίων και μακρομορίων που εμπλέκονται και στις λειτουργίες που αυτά επιτελούν, για την ανακάλυψη νέων ενώσεων-οδηγών. Συγκεκριμένα, η λεπτομερής γνώση μιας παθοφυσιολογικής κατάστασης κατευθύνει την έρευνα ως προς τις δυνατότητες παρέμβασης προκειμένου να αποκατασταθεί η φυσιολογική λειτουργία. Στην περίπτωση

υπερέκφρασης μιας πρωτεΐνης, όπως συμβαίνει με την KRAS, στόχος είναι η εύρεση μορίων που θα παρεμποδίσουν το βιολογικό αποτέλεσμα της πρωτεΐνης-στόχου.

Οι δομές των συμπλόκων των πρωτεϊνών-στόχων με τα ενδογενή (ή μη) προσδέματα μπορούν να αποτελέσουν την αφετηρία για την ανακάλυψη νέων ενώσεων-οδηγών. Οι δομές αυτές μπορούν να προσδιοριστούν με τη χρήση πειραματικών τεχνικών, όπως η κρυσταλλογραφία ακτίνων-Χ και η φασματοσκοπία NMR. Όπως είναι φυσικό, τα ενδογενή προσδέματα, όπως είναι για την KRAS οι ενώσεις GTP και GDP, εμφανίζουν τη βέλτιστη δομή για τη σύνδεση με την πρωτεΐνη-υποδοχέα. Ωστόσο, δε μπορούν να χορηγηθούν ως φάρμακα, λόγω έλλειψης μεταβολικής σταθερότητας, χαμηλής διαθεσιμότητας, καθώς και της πιθανότητας εκδήλωσης δευτερευουσών δράσεων, στην περίπτωση που αλληλεπιδρούν και με άλλους υποδοχείς πλην της πρωτεΐνης-στόχου. Κατά τον ορθολογικό σχεδιασμό νέων μορίων-φαρμάκων, τα δομικά χαρακτηριστικά των ενδογενών προσδεμάτων που ευνοούν τη σύνδεση με την πρωτεΐνη-υποδοχέα διατηρούνται, ενώ παράλληλα ανεπιθύμητα χαρακτηριστικά ελαττώνονται.

### 1.2.1 Σχεδιασμός φαρμάκων με βάση τη δομή της πρωτεΐνης

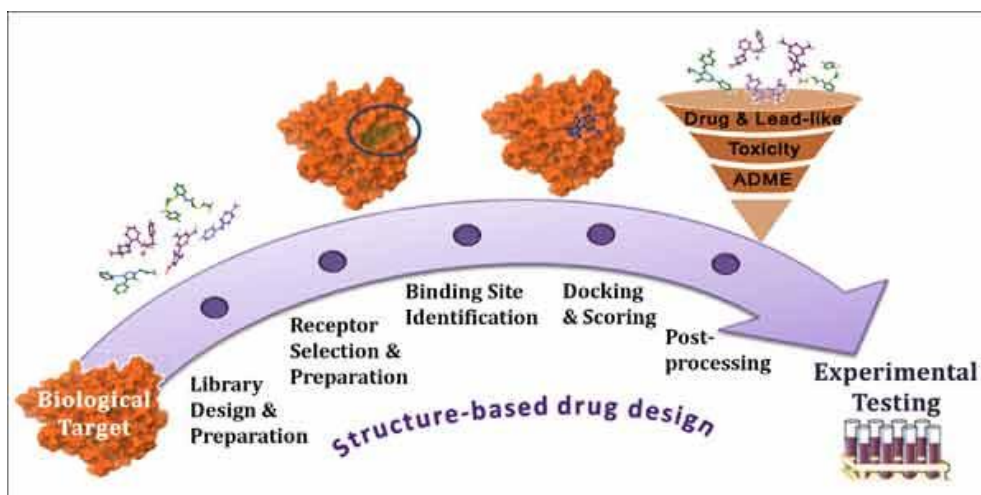
Ο σχεδιασμός φαρμάκων με βάση τη δομή της πρωτεΐνης-στόχου (structure-based drug design, SBDD) βασίζεται στη γνώση και κατανόηση του τρόπου με τον οποίο μικρά μόρια αναγνωρίζονται και αλληλεπιδρούν με τις πρωτεΐνες. Η τεχνική που χρησιμοποιείται είναι αυτή της προσαρμογής του προσδέματος στο ενεργό κέντρο (docking). Απαραίτητη είναι η επιλογή των κατάλληλων προσδεμάτων, συνήθως μέσα από διαθέσιμες βιβλιοθήκες ενώσεων. Ακόμα, είναι αναγκαία η γνώση της δομής της πρωτεΐνης-στόχου, κάτι το οποίο επιτυγχάνεται μέσω πειραματικών τεχνικών, όπως ήδη έχει αναφερθεί. Υπάρχουν διαθέσιμες πάνω από 100.000 τρισδιάστατες δομές πρωτεϊνών, καταχωρημένες στην τράπεζα δεδομένων Protein Data Bank (PDB).

Στην πλειοψηφία των περιπτώσεων, η δομή του υποδοχέα θεωρείται στατική κατά τη διαδικασία προσαρμογής (docking). Η δομή των προσδεμάτων μπορεί να θεωρηθεί άκαμπτη, οπότε η γεωμετρία τους βελτιστοποιείται με μοριακή μοντελοποίηση και προκύπτει μια διαμόρφωση που αντιστοιχεί στην ελάχιστη ενέργεια, η οποία χρησιμοποιείται σε όλη τη

διαδικασία docking. Αν η δομή των προσδεμάτων θεωρηθεί εύκαμπτη, υπάρχουν δύο επιλογές:

- Κατασκευή ομάδας διαμορφώσεων χαμηλής ενέργειας, οι οποίες χρησιμοποιούνται στη συνέχεια όλες ως άκαμπτες.
- Βελτιστοποίηση της διαμόρφωσης εντός του ενεργού κέντρου του υποδοχέα. Στην περίπτωση αυτή, οι διαδικασίες της βελτιστοποίησης και της προσαρμογής της διαμόρφωσης του προσδέματος πραγματοποιούνται ταυτόχρονα.

Κατά τη διαδικασία προσαρμογής των προσδεμάτων στο ενεργό κέντρο της πρωτεΐνης-υποδοχέα, επιτυγχάνεται η καλύτερη δυνατή αλληλεπίδραση μεταξύ των δύο. Τέλος, η προσαρμογή αυτή, καθώς και η ισχύς των αλληλεπιδράσεων μεταξύ προσδέματος και υποδοχέα αξιολογούνται μέσω εμπειρικών συναρτήσεων βαθμολόγησης (scoring functions), οι οποίες υπολογίζουν την ελεύθερη ενέργεια πρόσδεσης. Οι ενώσεις που προκύπτουν υποβάλλονται σε μια τελική αξιολόγηση των φυσικοχημικών τους ιδιοτήτων, της φαρμακομοιότητάς τους (drug-likeness), της μεταβολικής τους σταθερότητας, της τοξικότητας και των ιδιοτήτων ADME. Από τη διαδικασία αυτή προκύπτει μικρός αριθμός ενώσεων που προχωρούν στη φάση των πειραματικών δοκιμών.



**Εικόνα 1-11:** Διαδικασία σχεδιασμού φαρμάκων με βάση τη δομή της πρωτεΐνης [16]

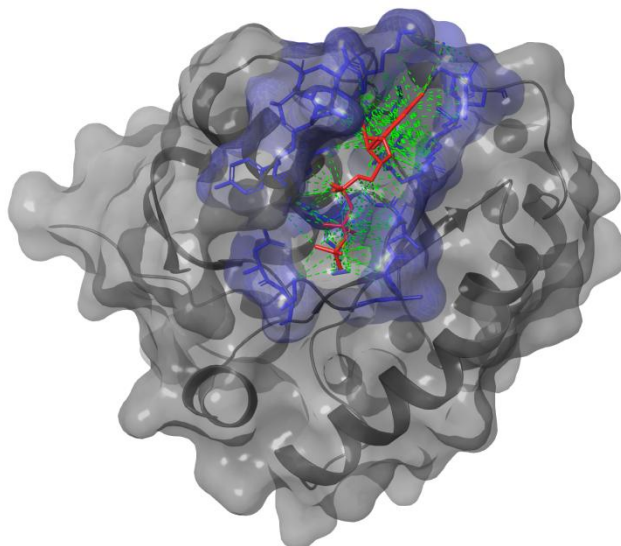
## 1.2.2 Κοιλότητες πρόσδεσης σε μία πρωτεΐνη

Οι κοιλότητες πρόσδεσης μια πρωτεΐνης είναι μικρές περιοχές στην τριτοταγή δομή της πρωτεΐνης, στις οποίες μικρά μόρια προσδένονται εκλεκτικά. Στην πλειοψηφία των περιπτώσεων, οι κοιλότητες είναι κοίλες στο σχήμα. Τα μόρια που προσδέονται στις κοιλότητες πρόσδεσης μιας πρωτεΐνης συνήθως ονομάζονται προσδέματα (ligands). Παραδείγματα προσδεμάτων είναι οι νευροδιαβιβαστές, τοξίνες, πεπτίδια ή και ορμόνες. Μόνο ένας μικρός αριθμός αμινοξέων της πρωτεΐνης συμμετέχουν στη σύνδεση με τα προσδέματα, τα υπόλοιπα λειτουργούν σαν πλαίσιο της κοιλότητας, παρέχοντάς της κατάλληλη διαμόρφωση και προσανατολισμό. Η πρόσδεση συνήθως συνοδεύεται από αλλαγές στη διαμόρφωση της πρωτεΐνης, οι οποίες μεταβάλλουν τη λειτουργία της.

Κατά τη στόχευση μιας πρωτεΐνης με μικρά μόρια (προσδέματα), δημιουργείται μεταξύ τους ένα μη ομοιοπολικό σύμπλοκο, το οποίο προκαλεί μια σειρά γεγονότων που καταλήγουν στο βιολογικό αποτέλεσμα. Οι αγωνιστές είναι προσδέματα, των οποίων το βιολογικό αποτέλεσμα αντιστοιχεί στο 100% του αποτελέσματος που προκαλεί το φυσικό πρόσδεμα. Σε περίπτωση που αυτό δε συμβαίνει, το πρόσδεμα καλείται μερικός αγωνιστής. Ακόμα, αν το πρόσδεμα παρεμποδίζει την εμφάνιση του βιολογικού αποτελέσματος τότε καλείται ανταγωνιστής [15].

## Ορθοστερικές κοιλότητες πρόσδεσης

Οι ορθοστερικές κοιλότητες πρόσδεσης είναι οι κοιλότητες πρόσδεσης της πρωτεΐνης, στις οποίες δεσμεύονται τα ενδογενή προσδέματα, προκειμένου να παραγάγουν το βιολογικό τους αποτέλεσμα. Στην περίπτωση των ενζύμων, η ορθοστερική θέση πρόσδεσης ονομάζεται ενεργό κέντρο, ενώ τα προσδέματα αναφέρονται ως υποστρώματα του ενζύμου. Η βασική λειτουργία των ενζύμων είναι η επαγωγή της πραγματοποίησης μια χημικής αντίδρασης στο ενεργό τους κέντρο, μέσω ελάττωσης του ενεργειακού φράγματος που είναι απαραίτητο για την αντίδραση. Για παράδειγμα, η πρωτεΐνη KRAS είναι μια υδρολάση, στις οποίες το ενεργό κέντρο πραγματοποιείται η υδρόλυση της ένωσης GTP προς την ένωση GDP.



**Εικόνα 1-12:** Σύμπλοκο της πρωτεΐνης KRAS με το ενδογενές πρόσδεμα GTP. Τα αμινοξέα του ενεργού κέντρου φαίνονται με μπλε χρώμα, το μόριο GTP με κόκκινο χρώμα και οι αλληλεπιδράσεις μεταξύ πρωτεΐνης-πρόσδεματος με πράσινες διακεκομμένες γραμμές. Η πρωτεΐνη απεικονίζεται με τη μορφή γκρι κορδελών (ribbons), ενώ φαίνεται η επιφάνειά της με διαφανές γκρι χρώμα (PDB ID: 5VQ2) [17]

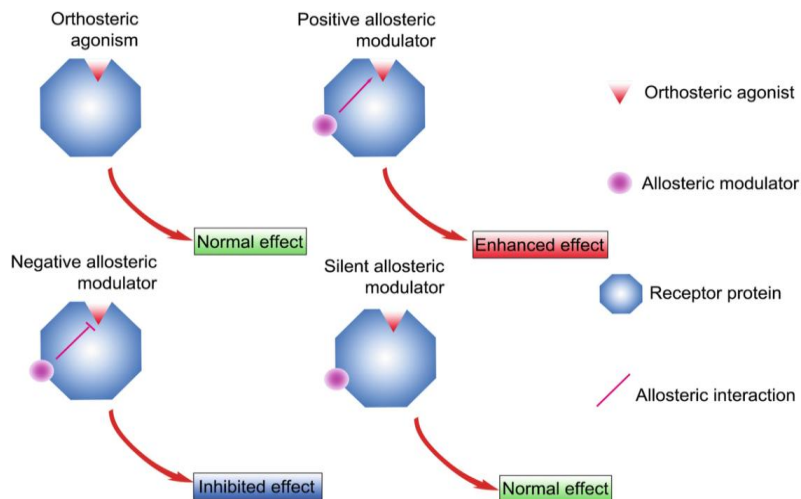
Τα φάρμακα που στοχεύουν τις ορθοστερικές κοιλότητες πρόσδεσης δρουν παρεμποδίζοντας την πρόσδεση των ενδογενών προσδεμάτων. Ενώσεις που έχουν υψηλή χημική συγγένεια με την επιφάνεια των ορθοστερικών θέσεων πρόσδεσης μπορούν να χρησιμοποιηθούν ως ορθοστερικά φάρμακα, περίπτωση η οποία είναι και η πιο συχνή. Μόρια που προσδέονται στις ορθοστερικές κοιλότητες πρόσδεσης, ανταγωνιζόμενα την πρόσδεση των ενδογενών προσδεμάτων, ονομάζονται ορθοστερικοί αγωνιστές.

## Αλλοστερικές κοιλότητες πρόσδεσης

Ως αλλοστερισμός (allostery) ορίζεται το φαινόμενο, κατά το οποίο η πρόσδεση σε μια κοιλότητα πρόσδεσης μιας πρωτεΐνης επηρεάζει τη δραστηριότητα πρόσδεσης σε μια διαφορετική κοιλότητα, επιτρέποντας τη ρύθμιση της αντίστοιχης λειτουργίας της πρωτεΐνης. Κατ'αυτόν τον τρόπο, μας δίνεται η δυνατότητα να ρυθμίσουμε τη λειτουργία μιας πρωτεΐνης, στοχεύοντας κοιλότητες διαφορετικές από τις ορθοστερικές κοιλότητες πρόσδεσης (αλλοστερικές κοιλότητες), διαδικασία γνωστή ως αλλοστερική ρύθμιση. Κατά τη δέσμευση

προσδεμάτων στις αλλοστερικές κοιλότητες πρόσδεσης, προκαλείται αλλαγή στην τρισδιάστατη δομή της πρωτεΐνης, γεγονός που συχνά επηρεάζει τις ορθοστερικές θέσεις πρόσδεσης. Οι ενώσεις που προσδένονται στις αλλοστερικές κοιλότητες πρόσδεσης ονομάζονται αλλοστερικοί μετατροπείς (allosteric modulators).

Οι αλλοστερικοί μετατροπείς μπορεί να είναι θετικοί, αρνητικοί ή ουδέτεροι. Οι θετικοί αλλοστερικοί μετατροπείς προσδένονται σε αλλοστερικές κοιλότητες και αυξάνουν τη συγγένεια των αγωνιστών με την πρωτεΐνη, ενισχύοντας τη βιολογική της δράση. Αντίστοιχα, οι αρνητικοί αλλοστερικοί μετατροπείς ελαττώνουν τη συγγένεια των αγωνιστών με την πρωτεΐνη, μειώνοντας ή παρεμποδίζοντας τη βιολογική δράση της πρωτεΐνης. Τέλος, υπάρχει και η περίπτωση η πρόσδεση των αλλοστερικών μετατροπέων να μην επηρεάζει την αγωνιστική δραστηριότητα άρα και τη βιολογική δράση της πρωτεΐνης, οπότε καλούνται ουδέτεροι. Στην παρακάτω εικόνα, φαίνεται ο τρόπος δράσης των αλλοστερικών μετατροπέων.



**Εικόνα 1-13:** Μηχανισμοί δράσης των διαφόρων τύπων αλλοστερικών μετατροπέων [18]



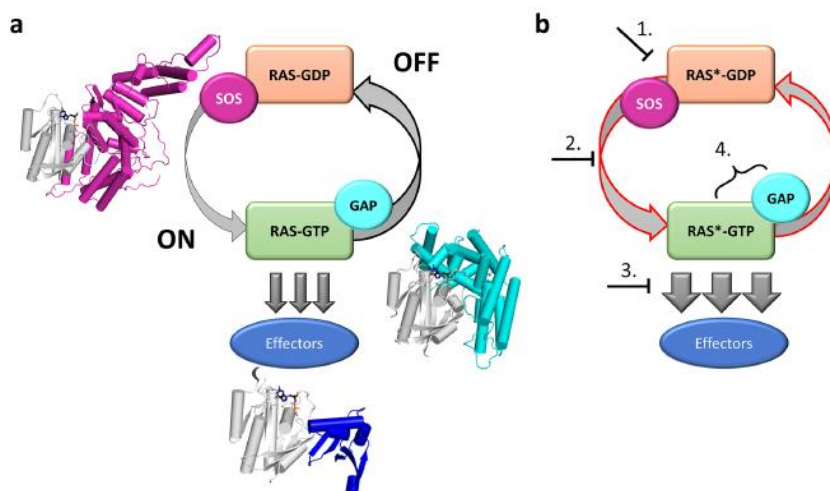
### 1.3 Στρατηγικές στόχευσης της KRAS-4B

Οι προσπάθειες για τη φαρμακευτική στόχευση των μεταλλαγμένων μορφών RAS που εμφανίζονται στους ανθρώπινους καρκίνους υπήρξαν μέχρι στιγμής ανεπιτυχείς. Καθώς η βιολογική δραστηριότητα των RAS απαιτεί την πρόσδεσή τους στη μεμβράνη, οι αρχικές προσπάθειες στόχευσης εστιάστηκαν στη δημιουργία αναστολέων της δράσης του ενζύμου φαρνεσυλο-τρανσφεράση (farnesyl transferase). Οι ενώσεις αυτές ελάττωσαν με επιτυχία την ανάπτυξη των όγκων που δημιουργούνται από το ισόμορφο HRAS[19], ωστόσο δεν είχαν επίδραση στους όγκους όπου κυριαρχούν μεταλλάξεις των ισόμορφων NRAS, KRAS[20]. Άλλες στρατηγικές, όπως η επαναφορά της υδρόλυσης της ένωσης GTP υπήρξαν επίσης αναποτελεσματικές[21]. Λόγω αυτών των δυσκολιών, το ενδιαφέρον στράφηκε προς τη στόχευση των σηματοδοτικών πρωτεϊνών που συνδέονται με την KRAS, όπως η Raf που μεσολαβεί στο μονοπάτι Ras-Raf-MEK-Erk. Παρότι η στρατηγική αυτή αποδείχθηκε κλινικά αποτελεσματική σε πλήθος καρκίνων[22], η αντίσταση τελικά δε μπόρεσε να αποφευχθεί, συχνά με τη μορφή εναλλακτικών σηματοδοτικών οδών. Συνεπώς, η έρευνα για νέους τρόπους στόχευσης των πρωτεϊνών RAS καθίσταται αναγκαία.

Οι μεταλλάξεις που επιδρούν στη λειτουργική πρόσδεση των πρωτεϊνών GAP, GEF και των πρωτεϊνών σηματοδότησης (effectors) διαταράσσουν τον μηχανισμό ρύθμισης και οδηγούν στην ογκογένεση. Τη σήμερον ημέρα, η έρευνα για τη στόχευση των πρωτεϊνών RAS επικεντρώνεται στις εξής στρατηγικές:

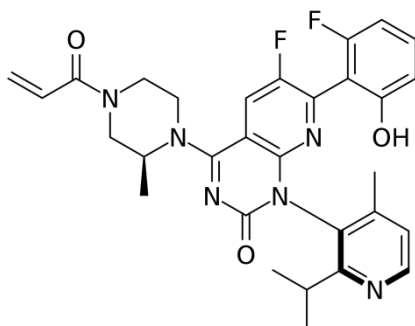
1. Διαταραχή του σχηματισμού του συμπλόκου της KRAS με τις πρωτεΐνες GEF, προκειμένου να ελαττωθεί η εναλλαγή του GDP προς GTP
2. Αύξηση της συγκέντρωσης του συμπλόκου KRAS-GDP σε βάρος του συμπλόκου KRAS-GTP (κυρίως με ομοιοπολικά προσδέματα)
3. Παρεμπόδιση της πρόσδεσης πρωτεϊνών που αφορούν τη διαδικασία σηματοδότησης (effectors), προκειμένου να εξασθενίσει η σηματοδότηση
4. Βελτίωση της πρόσδεσης των πρωτεϊνών GAP στις μεταλλαγμένες μορφές της KRAS προκειμένου να ελαττωθεί η συγκέντρωση του συμπλόκου KRAS-GTP

Οι μηχανισμοί αυτοί στόχευσης περιγράφονται στο παρακάτω σχήμα.



**Σχήμα 1-1:** Μηχανισμοί στόχευσης των πρωτεϊνών RAS [23]

Στις 28 Μαΐου 2021, εγκρίθηκε από τον Αμερικάνικο Οργανισμό Τροφίμων και Φαρμάκων (FDA) το φάρμακο sotorasib (Lumakras™, Amgen, Inc.), το οποίο απευθύνεται σε ενήλικες ασθενείς με τοπικά προχωρημένο ή μεταστατικό μικροκυτταρικό καρκίνο του πνεύμονα (ΜΜΚΠ), στους οποίους εντοπίζεται η μετάλλαξη G12C της KRAS[24]. Το sotorasib λειτουργεί σταθεροποιώντας την G12C KRAS στην ανενεργή της μορφή. Η έγκριση βασίστηκε σε κλινικές δοκιμές σε 124 ασθενείς με προχωρημένη ασθένεια ή με τουλάχιστον μία προηγούμενη συστηματική θεραπεία, όπως η χημειοθεραπεία. Το sotorasib προκάλεσε συρρίκνωση των όγκων στο 36% των ασθενών με ΜΜΚΠ, σε αντίθεση με τις συμβατικές θεραπείες που προκαλούν συρρίκνωση των όγκων σε λιγότερο από 20% των ασθενών, εξέλιξη αρκετά ελπιδοφόρα[25]. Η ανακάλυψη και έγκριση του sotorasib ανοίγει το δρόμο για την ανάπτυξη ενώσεων που θα στοχεύουν επιλεκτικά και τις μεταλλάξεις G12D, G12V της KRAS.



**Εικόνα 1-14:** Η χημική δομή της ένωσης sotorasib, του πρώτου εγκεκριμένου φαρμάκου κατά της G12C KRAS

### 1.3.1 Γνωστές θέσεις πρόσδεσης και αναστολείς της KRAS

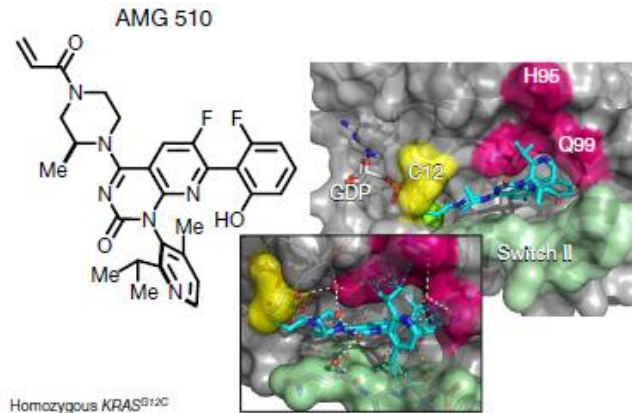
Τα ενδογενή προσδέματα GTP, GDP εμφανίζουν πολύ υψηλή συγγένεια με την ορθοστερική κοιλότητα των πρωτεϊνών RAS. Πρακτικά, η KRAS-4B, όπως και τα υπόλοιπα ισόμορφα RAS, βρίσκεται μόνιμα σε μορφή συμπλόκου με ένα από τα δύο αυτά προσδέματα. Για το λόγο αυτό, ο ανταγωνισμός της πρόσδεσης των ενώσεων GTP, GDP μέσω στόχευσης της ορθοστερικής κοιλότητας πρόσδεσης είναι αρκετά δύσκολος έως και αδύνατος.

Η πρωτεΐνη KRAS δε διαθέτει βαθιές κοιλότητες στην επιφάνειά της, η οποία ωστόσο είναι ευέλικτη και ικανή να σχηματίσει νέες κοιλότητες με την πρόσδεση μικρών μορίων. Δύο ευέλικτες κοιλότητες πρόσδεσης έχουν εντοπιστεί στις δύο πλευρές της α2 έλικας της KRAS, γνωστές ως κοιλότητες S-IIp και SI/II.

#### **Κοιλότητα S-IIp**

Η κοιλότητα αυτή εντοπίστηκε στην ανενεργή μορφή της G12C KRAS, συνδεδεμένη με το πρόσδεμα GTP. Στην ενεργή μορφή της G12C KRAS, η κοιλότητα δεν εντοπίζεται λόγω διαφορετικής διαμόρφωσης της περιοχής Switch II. Οι ενώσεις που έχουν αναπτυχθεί για τη στόχευση της κοιλότητας S-IIp αφορούν κυρίως μικρά μόρια που συνδέονται ομοιοπολικά με την κυστεΐνη της θέσης 12, παρεμποδίζοντας τη σύνδεση του προσδέματος GTP στην KRAS. Η σύνδεση αυτή προκαλεί επίσης μετακίνηση του αμινοξέος γλυκίνη 60, το οποίο είναι υπεύθυνο για τη σύνδεση της περιοχής Switch II με το πρόσδεμα GTP. Κατ'αυτόν τον τρόπο, αποτρέπεται ο σχηματισμός της ενεργής μορφής της KRAS, δηλαδή του συμπλόκου KRAS-GTP με τη σύνδεση ενός ανταγωνιστή που έχει και αλλοστερικό αντίκτυπο στη διαμόρφωση της περιοχής Switch II. Ένα τέτοιο παράδειγμα αποτελεί η ένωση AMG-510, η οποία ανακαλύφθηκε από τους Canon et al[26] μέσω του συνδυασμού της πειραματικής τεχνικής ηλεκτρόφιλου ελέγχου δέσμησης μορίων (intensive electrophile screening) και του σχεδιασμού φαρμάκων με βάση τη δομή του υποδοχέα (SBDD) και βρίσκεται στη φάση των κλινικών δοκιμών. Η ένωση αυτή σταθεροποιεί εκλεκτικά την ανενεργή μορφή της G12C KRAS, δηλαδή το σύμπλοκο KRAS-GDP, μέσω σχηματισμού δεσμού υδρογόνου με ένα από τα οξυγόνα του φωσφορικού που περιέχει το πρόσδεμα GDP. Ενώσεις με παρόμοια δράση στην κοιλότητα S-IIp που ανακαλύφθηκαν προηγουμένως είναι οι ARS-1620 από τους Janes et al[27] και ARS-853 από τους Patricelli et al[28].

Η εκλεκτική στόχευση της G12C KRAS με ομοιοπολικούς αναστολείς έθεσε το ερώτημα αν αυτή η στρατηγική θα ήταν βιώσιμη και για την G12D KRAS, καθώς η καρβοξυλική ομάδα του ασπαρτικού οξέος μπορεί να αντιδράσει με λειτουργικές ομάδες όπως η αζιριδίνη και η χλωροακεταμίδα. Παρόλα αυτά, αποδείχθηκε ότι στην περίπτωση της KRAS, ενώσεις που διαθέτουν ηλεκτρόφιλες ομάδες κατάλληλες για αντίδραση με το ασπαρτικό οξύ αδυνατούν να το αναγνωρίσουν επιτυχώς, πιθανώς λόγω του ακατάλληλου προσανατολισμού του.

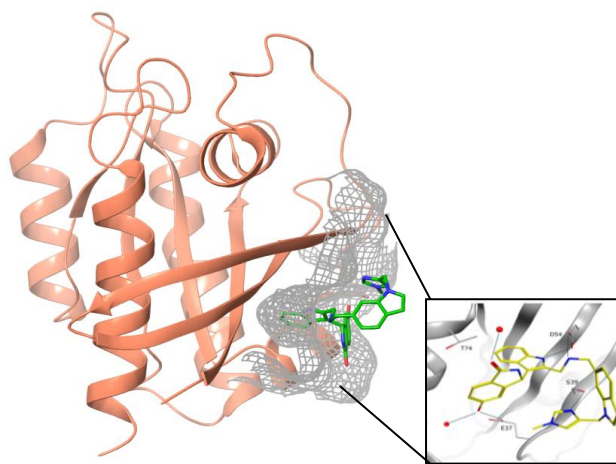


**Εικόνα 1-15:** Κρυσταλλική δομή του συμπλόκου της KRAS με το πρόσδεμα GDP και τον αναστολέα AMG-510 (PDB ID 6OIM) [26]. Αμινοξέα της κοιλότητας πρόσδεσης: 9-11, 12 (ομοιοπολικά), 16, 34, 58-63, 68, 69, 72, 92, 95, 96, 99, 100, 103

## Κοιλότητα SI/II

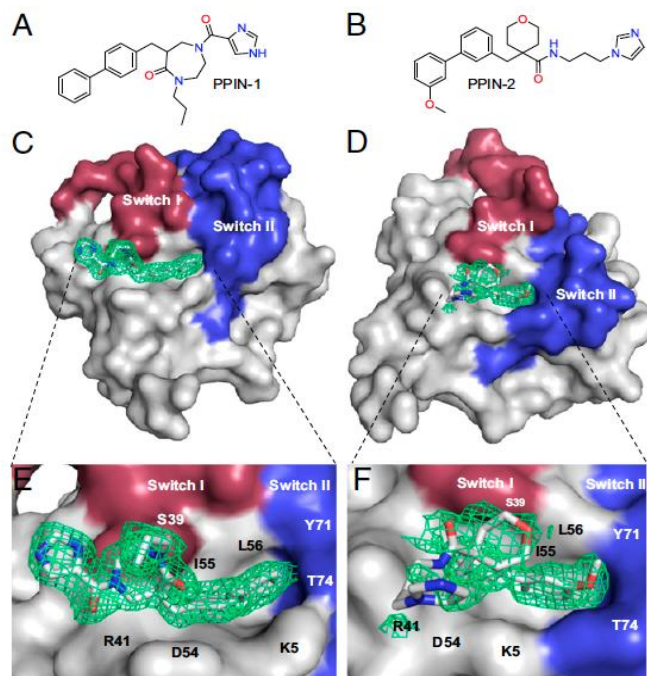
Όσο αφορά στην περιοχή αλληλεπίδρασης της πρωτεΐνης KRAS με τις σηματοδοτικές πρωτεΐνες, όπως η Raf, η κοιλότητα SI/II αποτελεί την πιο πιθανή τοποθεσία στόχευσης από μικρά μόρια. Η κοιλότητα βρίσκεται ανάμεσα στην α2 έλικα και τα β1-β2 πτυχωτά φύλλα και παρουσιάζει σημαντική ευελιξία κατά τη σύνδεση μορίων. Αναγνωρίστηκε αρχικά από τους Maurer et al[29] στη δομή της G12D KRAS-4B μέσω της πειραματικής μεθόδου NMR διαφοράς μεταφοράς κορεσμού (NMR-based Saturation Transfer Difference). Οι Maurer et al. ανακάλυψαν μικρές ενώσεις ινδολών και βενζαμιδίνης, οι οποίες παρεμποδίζουν το σχηματισμό του συμπλόκου της KRAS με την πρωτεΐνη SOS, μίας από τις πρωτεΐνες GEF, μέσω πρόσδεσης στην κοιλότητα SI/II. Η ίδια κοιλότητα αναγνωρίστηκε από τους Sun et al[30], οι οποίοι έδειξαν ότι η πρόσδεση μορίων στην κοιλότητα SI/II επάγει μια μικρή αλλαγή στη διαμόρφωση της α2 έλικας και της περιοχής Switch II, ενώ αποτρέπει το αμινοξύ τυροσίνη 71

από τον σχηματισμό δεσμού υδρογόνου με το αμινοξύ τυροσίνη 910 της πρωτεΐνης SOS. Αργότερα, οι Kessler et al[31] ανακάλυψαν τον αναστολέα BI-2852 πραγματοποιώντας ελέγχους δέσμευσης χημικών θραυσμάτων (fragment screening) στο σύμπλοκο της G12D KRAS με το πρόσδεμα GDP. Ο αναστολέας BI-2852 παρεμποδίζει την αλληλεπίδραση μεταξύ του συμπλόκου G12D KRAS-GDP και της πρωτεΐνης SOS, αλλά και τις αλληλεπιδράσεις μεταξύ του συμπλόκου G12D KRAS-GTP και της πρωτεΐνης SOS ή τις σηματοδοτικές πρωτεΐνες, όπως η πρωτεΐνη Raf. Αυτό έχει ως αποτέλεσμα την παρεμπόδιση της ανταλλαγής του GDP προς GTP, διαδικασία μεσολαβούμενη από τις πρωτεΐνες GEF, όπως είναι και η πρωτεΐνη SOS.



**Εικόνα 1-16:** Σύμπλοκο της πρωτεΐνης G12D KRAS με τον αναστολέα BI-2852 (PDB ID: 6GJ8) [31]. Ο αναστολέας φαίνεται με πράσινο χρώμα και η κοιλότητα φαίνεται ως χάρτης επιφάνειας (γκρι δίχτυ). Η πρωτεΐνη απεικονίζεται με τη μορφή κορδελών (ribbons). Αμινοξέα κοιλότητας: 5-7, 37-39, 54-56,67,70,71,74,75.

Η κοιλότητα SI/II έχει επίσης προσδιοριστεί από τους Cruz-Migoni et al[32] σε δομές μεταλλαγμένης KRAS-4B με τη μετάλλαξη Q61H (γλουταμίνη σε ιστιδίνη), μέσω ελέγχου βιβλιοθήκης μορίων με τη χρήση συντονισμού επιφανειακών πλασμονίων (Surface Plasmon Resonance). Συγκεκριμένα, διαπιστώθηκε η πρόσδεση των ενώσεων PPIN-1, PPIN-2 μέσω συγκρουσάλλωσης με την πρωτεΐνη Q61H KRAS-4B.



**Εικόνα 1-17:** A,B: Χημικές δομές των ενώσεων PPIN-1, PPIN2 αντίστοιχα C,D,E,F: Κρυσταλλικές δομές της πρωτεΐνης Q61H KRAS-4B με τις ενώσεις PPIN-1 (C,E) και PPIN-2 (D,F). Αμινοξέα της κοιλότητας πρόσδεσης: 5-7, 39-41, 54-56, 71, 74 [32]

## 1.4 Σκοπός της εργασίας

Όπως ήδη αναφέρθηκε, τα ενδογενή προσδέματα GDP και GTP εμφανίζουν υψηλή συγγένεια με την ορθοστερική κοιλότητα πρόσδεσης της KRAS-4B. Για το λόγο αυτό, είναι αρκετά δύσκολο μικρά μόρια να ανταγωνιστούν την πρόσδεσή των ενδογενών προσδεμάτων στην KRAS-4B, κατά την πρόσδεσή τους στην ορθοστερική κοιλότητα. Ένας εναλλακτικός μηχανισμός αναστολής της δράσης της KRAS-4B είναι μέσω πρόσδεσης μικρών μορίων σε αλλοστερικές κοιλοότητες. Το πρώτο βήμα για το σχεδιασμό αλλοστερικών αναστολέων είναι η αναγνώριση αλλοστερικών κοιλοτήτων στην επιφάνεια της πρωτεΐνης. Επιπλέον, γνωρίζουμε ότι η KRAS-4B διμερίζεται παρουσία της σηματοδοτικής πρωτεΐνης Raf, όπως αναφέρθηκε στην Παράγραφο 1.1.2, και ο διμερισμός ενεργοποιεί μονοπάτια σηματοδότησης που σχετίζονται με την κυτταρική ανάπτυξη και τον κυτταρικό πολλαπλασιασμό.

Σκοπός της παρούσας εργασίας είναι η εύρεση αλλοστερικών κοιλοτήτων και συγκεκριμένα στη διεπιφάνεια του διμερούς KRAS-4B, ώστε μελλοντικά να μπορούν να σχεδιαστούν αλλοστερικοί αναστολείς που θα διαταράσσουν το διμερισμό της KRAS-4B, με αποτέλεσμα την παρεμπόδιση της σηματοδότησης και ως εκ τούτου της ογκογένεσης. Οι κοιλοότητες αυτές θα εξεταστούν ως προς τη χρονική τους μεταβολή σε τροχιές Μοριακών Δυναμικών προσομοιώσεων, προκειμένου να διευκρινιστεί αν διατηρούνται κατά τη διάρκεια της προσομοίωσης.

Δεδομένου ότι η μετάλλαξη της γλυκίνης σε ασπαρτικό οξύ (G12D) είναι αυτή που εμφανίζεται πιο συχνά στους καρκίνους που σχετίζονται με το ισόμορφο KRAS-4B, στόχος της εργασίας είναι η εύρεση κοιλοτήτων που θα εντοπίζονται εκλεκτικά στο μεταλλαγμένο διμερές G12D KRAS-4B. Για την διερεύνηση της εκλεκτικότητας, αναγκαία είναι η σύγκριση των κοιλοτήτων που προκύπτουν στο διμερές φυσικού τύπου και στο μεταλλαγμένο (G12D) διμερές.

## 2 ΜΕΘΟΔΟΛΟΓΙΑ - ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ

Οι προσομοιώσεις Μοριακής Δυναμικής αποτελούν ένα από τα κύρια υπολογιστικά εργαλεία για τη μελέτη της εξέλιξης ενός μοριακού συστήματος στο χρόνο. Η τεχνική αυτή παράγει διαμορφώσεις του συστήματος σε διαφορετικές χρονικές στιγμές. Από την ομαδοποίηση των διαμορφώσεων αυτών, προκύπτουν αντιπροσωπευτικές ή αλλιώς κυρίαρχες δομές του συστήματος στις οποίες έπειτα μπορούν να εντοπιστούν κοιλοότητες πρόσδεσης μέσω των αλγορίθμων SiteMap και FTMap. Τέλος, η χρονική μεταβολή του όγκου των κοιλοτήτων μπορεί να υπολογιστεί μέσω του εργαλείου Eroch του λογισμικού VMD. Οι παραπάνω υπολογιστικές τεχνικές περιγράφονται αναλυτικά στο παρόν κεφάλαιο.

### 2.1 Εμπειρική συνάρτηση δυναμικού-πεδίο δυνάμεων

Η ολική ενέργεια του συστήματος υπολογίζεται ως το άθροισμα της κινητικής και της δυναμικής ενέργειας του συστήματος:

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

όπου:

- T: η κινητική ενέργεια του συστήματος
- V: η δυναμική ενέργεια του συστήματος

Για τον υπολογισμό της δυναμικής ενέργειας, συνήθως χρησιμοποιούνται εμπειρικές συναρτήσεις, οι οποίες έχουν μικρότερο υπολογιστικό κόστος, ωστόσο περιλαμβάνουν αριθμητικές προσεγγίσεις που οδηγούν σε περιορισμούς. Για παράδειγμα, οι δραστικές αλλαγές στην ηλεκτρονιακή δομή, όπως είναι η δημιουργία ή το σπάσιμο δεσμών, δε μπορούν να μοντελοποιηθούν. Οι σύγχρονες εμπειρικές συναρτήσεις ή αλλιώς πεδία δυνάμεων παρέχουν έναν ικανοποιητικό συμβιβασμό μεταξύ ακρίβειας και υπολογιστικής αποτελεσματικότητας, καθώς είναι συχνά βαθμονομημένες σύμφωνα με πειραματικά.

Το εμπειρικό δυναμικό που χρησιμοποιείται για τον υπολογισμό της ενέργειας ενός συστήματος βασίζεται μόνο στη θέση των πυρήνων αγνοώντας τα ηλεκτρόνια, τα οποία θεωρείται ότι κινούνται πολύ γρήγορα σε σχέση με τους πυρήνες και έτσι μπορεί να θεωρηθεί



ότι η κίνησή τους ακολουθεί αυτή των πυρήνων (προσέγγιση Born-Orppenheimer). Συνεπώς, οι προσομοιώσεις Μοριακής Δυναμικής στηρίζονται στην κλασική μηχανική για να μπορέσουν να περιγράψουν τη δυναμική του συστήματος.

Το άθροισμα των διαμοριακών αλληλεπιδράσεων που συνεισφέρουν στη δυναμική ενέργεια του συστήματος ορίζεται ως η δυναμική του ενέργεια. Ειδικότερα, οι επιμέρους ενδομοριακές αλληλεπιδράσεις της πρωτεΐνης και του προσδέματος αλλά και οι διαμοριακές αλληλεπιδράσεις μεταξύ πρωτεΐνης-προσδέματος χρειάζεται να αναλυθούν, προκειμένου να υπολογιστεί η δυναμική ενέργεια του συστήματος. Ως εκ τούτου, η δυναμική ενέργεια μπορεί να περιγραφεί ως ένα σύνολο εξισώσεων που περιγράφει τις βασικές αλληλεπιδράσεις δίνοντάς μας την ολική δυναμική ενέργεια των ατόμων ενός συστήματος και ονομάζεται πεδίο δυνάμεων.

Το πεδίο δυνάμεων μπορεί να χωριστεί σε δύο κατηγορίες αλληλεπιδράσεων, τις ομοιοπολικές και τις μη ομοιοπολικές αλληλεπιδράσεις (εξίσωση 2.2). Οι ομοιοπολικές αλληλεπιδράσεις περιλαμβάνουν ενδομοριακές δυνάμεις που προκαλούνται λόγω της τάσης δεσμών, της κάμψης και της στρέψης γωνιών (εξίσωση 2.3). Οι μη ομοιοπολικές αλληλεπιδράσεις περιλαμβάνουν τις διαμοριακές αλληλεπιδράσεις van der Waals και τις ηλεκτροστατικές αλληλεπιδράσεις Coulomb (εξίσωση 2.4).

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

$$E_{bonded} = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi (1 + \cos[n\phi - \delta]) + \sum_{impropers} k_\omega (\omega - \omega_0)^2 \quad (2.3)$$

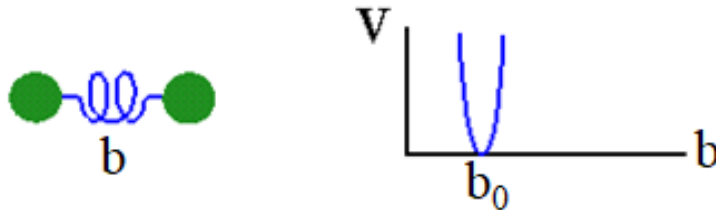
$$E_{non-bonded} = \sum_{i,j} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i,j} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \quad (2.4)$$

Οι όροι των εξισώσεων 2.3 και 2.4 αποτελούν το πεδίο δυνάμεων και περιγράφονται αναλυτικά παρακάτω.

### Τάση δεσμών (Bond stretching)

Ο όρος αυτός περιγράφει τη δόνηση ανάμεσα σε δύο άτομα συνδεδεμένα με ομοιοπολικό δεσμό και μπορεί να περιγραφεί από ένα αρμονικό δυναμικό βασιζόμενο στο νόμο του Hooke:

$$V_B = k_b(b - b_0)^2$$



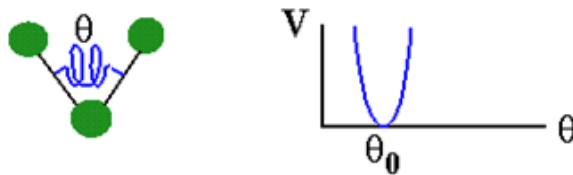
Όπου:

- **b**: η απόσταση μεταξύ των δύο ατόμων
- **b<sub>0</sub>**: η απόσταση μεταξύ των ατόμων στη θέση ισορροπίας
- **k<sub>b</sub>**: η σταθερά της δύναμης που χαρακτηρίζει τη δόνηση

### Κάμψη δεσμών (Angle bending)

Ο όρος της κάμψης δεσμών περιγράφει τη δόνηση μεταξύ 3 ομοιοπολικά συνδεδεμένων ατόμων. Ο όρος αυτός περιγράφεται επίσης χρησιμοποιώντας ένα αρμονικό δυναμικό βασιζόμενο στο νόμο του Hooke:

$$V_\theta = k_\theta(\theta - \theta_0)^2$$



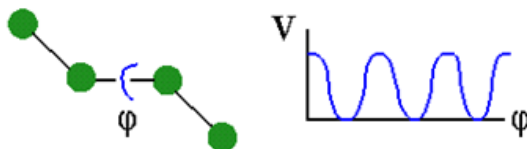
Όπου:

- **θ**: η γωνία μεταξύ τριών ατόμων
- **θ<sub>0</sub>**: η γωνία μεταξύ των ατόμων στη θέση ισορροπίας
- **k<sub>θ</sub>**: η σταθερά της δύναμης που χαρακτηρίζει τη δόνηση της γωνίας

## Παράγοντες στρέψης

Οι παράγοντες στρέψης περιγράφουν το φράγμα περιστροφής μεταξύ τεσσάρων ομοιοπολικά συνδεδεμένων ατόμων. Υπάρχουν δύο είδη παραγόντων στρέψης: των επιτρεπτών και μη επιτρεπτών διεδρων. Οι επιτρεπτές διεδρες (proper dihedrals) χαρακτηρίζονται από τη συνημιτονοειδή συνάρτηση δυναμικού:

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

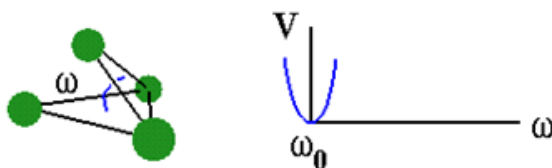


Όπου:

- $\varphi$ : η γωνία μεταξύ των επιπέδων που σχηματίζονται από το πρώτο και τα τρία τελευταία από τα τέσσερα άτομα.
- $\delta$ : καθορίζει τη γωνία με την ελάχιστη ενέργεια
- $k_{\varphi}$ : η σταθερά δυνάμεως
- $n$ : η περιοδικότητα

Ο όρος των μη επιτρεπτών διεδρων έχει σχεδιαστεί ώστε να 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. Περιγράφεται από το αρμονικό δυναμικό:

$$V_{\omega} = k_{\omega}(\omega - \omega_0)^2$$



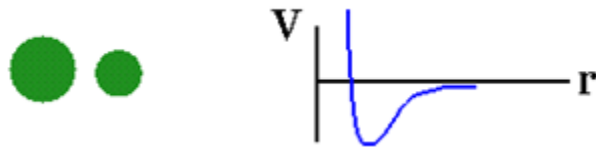
Όπου :

- $\omega$ : η γωνία μεταξύ του επιπέδου που σχηματίζεται από το κεντρικό άτομο, τα δύο περιφερειακά άτομα και το επίπεδο που σχηματίζεται από τα περιφερειακά άτομα.
- $\omega_0$ : η γωνία μεταξύ των ατόμων στη θέση ισορροπίας
- $k_\omega$ : η σταθερά της δύναμης που χαρακτηρίζει τη δόνηση της γωνίας

### Αλληλεπιδράσεις van der Waals

Οι δυνάμεις van der Waals (vdW) δρουν σε άτομα 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. Σε αποστάσεις που τείνουν στο άπειρο, οι δυνάμεις vdW τείνουν προς το μηδέν. Η αλληλεπίδραση μπορεί να προσεγγιστεί από το δυναμικό Lennard-Jones:

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



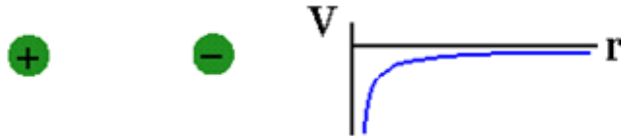
Όπου:

- $r_{ij}$ : η απόσταση μεταξύ των δύο ατόμων
- $\sigma_{ij}$ : προσδιορίζει τη διάμετρο του μορίου (ακτίνα Van der Waals) και είναι το σημείο όπου η δυναμική ενέργεια είναι μηδέν
- $\varepsilon_{ij}$ : υποδηλώνει την ισχύ της αλληλεπιδράσεως (το βάθος του δυναμικού)

## Ηλεκτροστατικές αλληλεπιδράσεις

Τέλος, οι ηλεκτροστατικές αλληλεπιδράσεις μεταξύ δύο ατόμων περιγράφονται από το νόμο του Coulomb:

$$V_{Elec} = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$

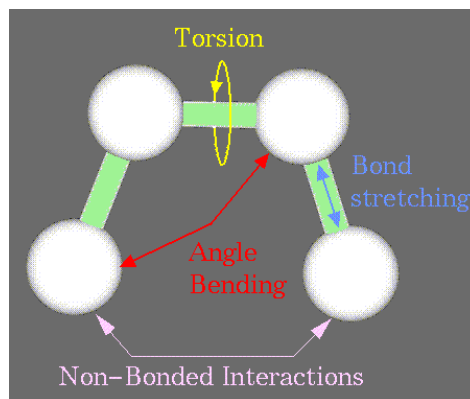


Όπου:

- $q_i$  και  $q_j$ : τα φορτία των δύο ατόμων
- $r_{ij}$ : η απόσταση μεταξύ των δύο ατόμων
- $\epsilon_0$ : είναι η διηλεκτρική σταθερά του κενού

Ο υπολογισμός της δυναμικής ενέργειας του συστήματος είναι απαραίτητος για τις προσομοιώσεις Μοριακής Δυναμικής.

Οι παραπάνω αλληλεπιδράσεις συνοψίζονται στο παρακάτω σχήμα.



**Σχήμα 2-1:** Αναπαράσταση των αλληλεπιδράσεων από τις οποίες προκύπτει η συνάρτηση δυναμικής ενέργειας της Μοριακής Δυναμικής [33]

## 2.2 Προσομοιώσεις Μοριακής Δυναμικής

Οι προσομοιώσεις Μοριακής Δυναμικής αποτελούν ένα από τα κύρια εργαλεία για τη θεωρητική μελέτη των βιομορίων. Μέσω της υπολογιστικής αυτής τεχνικής, εκτιμάται η συμπεριφορά ενός μοριακού συστήματος συναρτήσει του χρόνου. Συγκεκριμένα, μέσω των προσομοιώσεων Μοριακής Δυναμικής, μας δίνεται η δυνατότητα να κατανοήσουμε και να προβλέψουμε τις μακροσκοπικές θερμοδυναμικές ιδιότητες ενός περίπλοκου μοριακού συστήματος, βασισμένοι στις μικροσκοπικές του ιδιότητες, όπως οι θέσεις και οι ταχύτητες των ατόμων.

Γνωρίζοντας, λοιπόν, τις θέσεις και τις ταχύτητες όλων των ατόμων σε κάθε χρονική στιγμή, υπολογίζεται η ενέργεια του συστήματος  $N$  ατόμων, σύμφωνα με τις εξισώσεις του δεύτερου νόμου του Νεύτωνα:

$$m_i \frac{d^2 r_i}{dt^2} = F_i \quad i = 1, \dots, N. \quad (2.5)$$

Όπου  $m_i$  η μάζα του κάθε ατόμου του συστήματος,  $r_i$  η θέση του κάθε ατόμου του συστήματος,  $F_i$  η συνολική δύναμη που ασκείται στο κάθε άτομο  $i$  του συστήματος και  $t$  ο χρόνος. Οι δυνάμεις υπολογίζονται ως αρνητικές παράγωγοι της δυναμικής ενέργειας  $V(r_1, r_2, r_3, \dots, r_N)$ :

$$F_i = -\nabla_i V(r) \quad (2.6)$$

Όπου  $V(r)$  η συνολική δυναμική ενέργεια του συστήματος, ως συνάρτηση των ατομικών θέσεων. Εξισώνοντας τις δύο παραπάνω σχέσεις 2.5 και 2.6, λαμβάνουμε την εξίσωση:

$$\frac{dV(r)}{dr} = -m_i \frac{d^2 r_i}{dt^2} \quad (2.7)$$

Η παραπάνω διαφορική εξίσωση είναι δευτέρας τάξεως και επομένως δύσκολο να επιλυθεί μέσω κάποιας αναλυτικής μεθόδου, λόγω της πολυπλοκότητάς της. Γι αυτό το λόγο, χρησιμοποιούνται μέθοδοι αριθμητικής ολοκλήρωσης για την επίλυσή της.

Ο αλγόριθμος Verlet αποτελεί ένα παράδειγμα μεθόδου αριθμητικής ολοκλήρωσης. Ο συγκεκριμένος αλγόριθμος χρησιμοποιείται για τον υπολογισμό της τροχιάς των σωματιδίων σε ένα σύστημα, ολοκληρώνοντας τις εξισώσεις κίνησης του Νεύτωνα, με βάση τις παρακάτω εξισώσεις[34]:

$$\mathbf{p}_i(t_i) = \mathbf{p}_i(t_0) + \int_{t_0}^{t_i} \mathbf{a}_i(t) dt \quad (2.8)$$

$$\mathbf{q}_i(t_i) = \mathbf{q}_i(t_0) + \int_{t_0}^{t_i} \frac{\mathbf{q}_i(t)}{m} dt \quad (2.9)$$

$$\mathbf{a}_i(t) = - - \frac{1}{m_i} \frac{\partial V(\mathbf{r}_i, \dots, \mathbf{r}_M)}{\partial \mathbf{r}_i} \quad (2.10)$$

Οι εξισώσεις 2.5 και 2.6 επιλύονται ταυτόχρονα σε πολλαπλά μικρά διαστήματα ολοκλήρωσης (time steps). Το διάστημα ολοκλήρωσης υπαγορεύεται από τη μεγαλύτερη συχνότητα δόνησης μεταξύ δύο ατόμων στο σύστημα, η οποία στην περίπτωση μίας πρωτεΐνης αντιστοιχεί στη δόνηση του δεσμού C-H.

Μέσω της συνεχούς ολοκλήρωσης της διαφορικής εξίσωσης, είμαστε σε θέση να παρακολουθήσουμε τη χρονική εξέλιξη ενός συστήματος για ένα συγκεκριμένο χρονικό διάστημα, ελέγχοντας παράλληλα ότι η πίεση και η θερμοκρασία βρίσκονται στις επιθυμητές τιμές. Οι συντεταγμένες καταγράφονται σε ένα εξερχόμενο αρχείο σε συγκεκριμένα χρονικά διαστήματα και αναπαριστούν την πορεία του συστήματος. Μετά από κάποιες αρχικές αλλαγές, το σύστημα συνήθως εισέρχεται σε κατάσταση ισορροπίας. Από τον μέσο όρο της πορείας ενός συστήματος σε ισορροπία, μπορούν να εξαχθούν συμπεράσματα για τις μακροσκοπικές ιδιότητες μέσω ενός εξερχόμενου αρχείου.

### 2.2.1 Βασικές αρχές προσομοιώσεων Μοριακής Δυναμικής

Η Μοριακή Δυναμική αποτελεί τεχνική προσομοίωσης με τη χρήση υπολογιστών, κατά την οποία αξιοποιούνται οι εξισώσεις κίνησης της κλασικής μηχανικής για την περιγραφή της κίνησης ατόμων και μορίων. Απώτερος σκοπός η παρατήρηση της εξέλιξης ενός συστήματος στον χρόνο. Η μέθοδος αυτή κάνει χρήση των εξισώσεων κίνησης του Νεύτωνα (δεύτερος νόμος του Νεύτωνα), όπως ήδη έχει αναφερθεί παραπάνω. Για ένα μοριακό σύστημα  $N$  ατόμων, η δύναμη που δέχεται κάθε σωματίδιο  $i$  από το σύνολο των ατόμων σε κάθε χρονική στιγμή υπολογίζεται ως εξής:

$$F_i(\mathbf{r}, t) = -\nabla V_{total}(\mathbf{r}, t) = m_i \frac{d^2 \mathbf{r}_i}{dt^2} \quad (2.11)$$

Όπου  $V_{total}$  η ολική δυναμική ενέργεια του συστήματος, υπολογισμένη σύμφωνα με το πεδίο δυνάμεων, και  $\mathbf{r}_i$  η θέση κάθε ατόμου κάθε χρονική στιγμή. Η ενέργεια αυτή προσδιορίζει την ενέργεια αλληλεπίδρασης κάθε ατόμου με τα υπόλοιπα στοιχεία του συστήματος.

Αφότου υπολογιστούν όλες οι δυνάμεις, η εξίσωση 2.11 ολοκληρώνεται αριθμητικά με τη χρήση κατάλληλου αλγορίθμου, ώστε να παραχθούν νέες θέσεις και νέες ταχύτητες για κάθε σωματίδιο του συστήματος. Κατόπιν, η δυναμική ενέργεια υπολογίζεται με χρήση των νέων συντεταγμένων. Κατά τη διάρκεια μιας προσομοίωσης Μοριακής Δυναμικής σε σύστημα που βρίσκεται σε ισορροπία, τα βήματα είναι τα εξής:

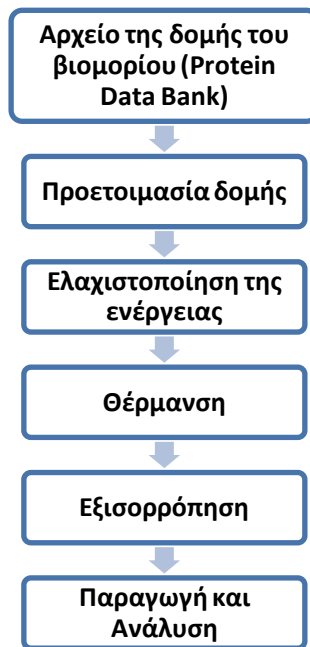
- 1) Επιλογή των αρχικών συνθηκών (συντεταγμένες) των ατόμων του μοριακού συστήματος
- 2) Επιλογή των αρχικών ταχυτήτων των ατόμων
- 3) Υπολογισμός της ορμής του κάθε ατόμου από την ταχύτητα και τη μάζα αυτού
- 4) Υπολογισμός των δυνάμεων που ασκούνται σε κάθε άτομο, ως αποτέλεσμα των αλληλεπιδράσεων του με άλλα άτομα
- 5) Υπολογισμός των νέων θέσεων των ατόμων στο χώρο μετά από χρόνο  $dt$ , ο οποίος αποτελεί και το βήμα προσομοίωσης
- 6) Υπολογισμός των νέων ταχυτήτων και επιταχύνσεων των ατόμων
- 7) Επανάληψη των βημάτων 3-6, έως ότου το σύστημα φθάσει σε ισορροπία



Στην υποενότητα 3.2.2 αναλύονται τα βήματα, τα οποία απαιτούνται για την οργάνωση και την πραγματοποίηση μιας προσομοίωσης Μοριακής Δυναμικής.

### 2.2.2 Ροή εργασίας προσομοιώσεων Μοριακής Δυναμικής

Το πρώτο βήμα για την πραγματοποίηση μιας προσομοίωσης Μοριακής Δυναμικής, σχετίζεται με την επιλογή της αρχικής κατάστασης του συστήματος, τη χρονική στιγμή  $t=0$ . Σε προσομοιώσεις βιομορίων, επιλέγεται η πειραματική δομή της πρωτεΐνης, η οποία έχει προκύψει έπειτα από κρυσταλλογραφία ακτίνων Χ, NMR ή από άλλη τεχνική δομική βιολογία και είναι καταχωρημένη στην τράπεζα δεδομένων Protein Data bank (<http://www.rcsb.org>). Αυτή είναι η αρχική δομή του συστήματος. Η επιλογή της αρχικής διαμόρφωσης είναι καίριας σημασίας, καθώς μπορεί να επηρεάσει την ποιότητα της προσομοίωσης. Οι προσομοιώσεις Μοριακής Δυναμικής περιέχουν 5 ακόμη βήματα, τα οποία αναφέρονται παρακάτω και απεικονίζονται σχηματικά στο Σχήμα 2-2 και περιγράφονται στις παρακάτω υποενότητες.



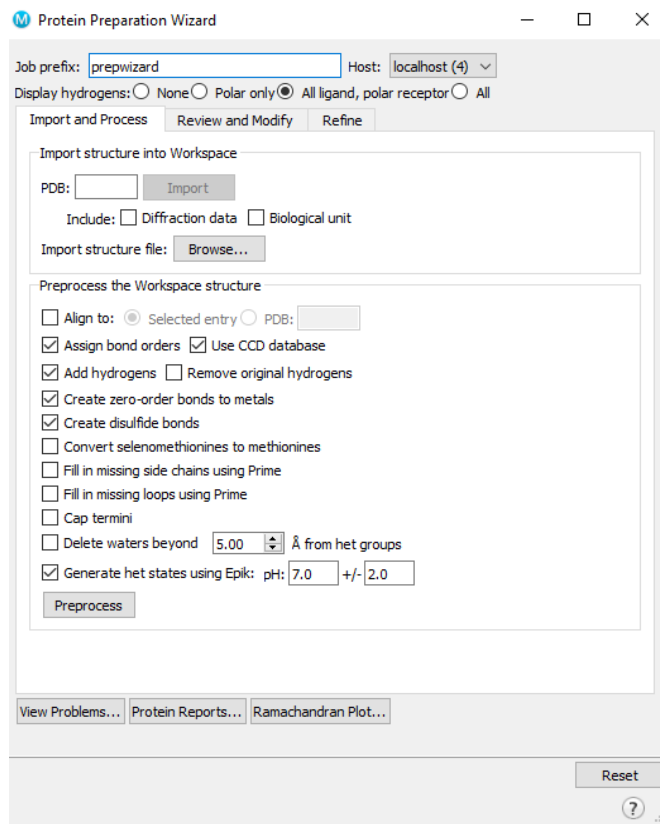
**Σχήμα 2-2:** Τα στάδια των προσομοιώσεων Μοριακής Δυναμικής

### 2.2.2.1 Προετοιμασία των πρωτεϊνικών δομών για προσομοίωση

Οι τρισδιάστατες δομές πρωτεϊνών που προέρχονται από μεθόδους δομικής βιολογίας χρειάζονται επεξεργασία πριν τη μοριακή μοντελοποίηση και τη χρήση υπολογιστικών τεχνικών, όπως η Μοριακή Δυναμική. Αυτό συμβαίνει διότι ένα τυπικό αρχείο δομής για παράδειγμα σε τύπο protein data bank «.pdb» προερχόμενο από κρυσταλλογραφία ακτίνων Χ περιέχει μόνο βαρέα άτομα, ενώ μπορεί να περιέχει και ένα συγκρυσταλλωμένο πρόσδεμα, μόρια νερού, μεταλλικά ιόντα και συμπαραγοντες. Λόγω της μειωμένης διακριτικής ικανότητας των πειραμάτων κρυσταλλογραφίας ακτίνων Χ, μπορεί επίσης να είναι δύσκολη η διάκριση μεταξύ αζώτου και οξυγόνου καθώς έχουν μόλις 1 ηλεκτρόνιο διαφορά. Κατά συνέπεια, χρειάζεται προσεκτικός έλεγχος στην τοποθέτηση των ομάδων αυτών, καθώς μπορεί να μην έχουν τοποθετηθεί σωστά κατά τη μοντελοποίηση της ηλεκτρονιακής πυκνότητας. Η επεξεργασία περιλαμβάνει προσθήκη ατόμων υδρογόνου, βελτιστοποίηση των δεσμών υδρογόνου, ελαχιστοποίηση της ενέργειας της πρωτεΐνης κ.ά.[35].

Το Protein Preparation Wizard είναι ένα εργαλείο του λογισμικού Schrödinger, το οποίο σχεδιάστηκε με σκοπό να βοηθήσει τους ερευνητές να επιβεβαιώνουν την ορθότητα της υπό εξέταση δομής, αλλά και να παρέχει μια ποιοτική δομή, ιδανική για την εφαρμογή πληθώρας εφαρμογών μοντελοποίησης[36]. Οι επιλογές που περιλαμβάνει για την κατάλληλη προετοιμασία πρωτεϊνικών δομών είναι οι παρακάτω:

- Ανάθεση κατάλληλων τάξεων δεσμών και τυπικών φορτίων
- Προσθήκη υδρογόνων (όπου υπάρχει έλλειψη)
- Δημιουργία δεσμών μηδενικής τάξης στα μέταλλα
- Δημιουργία δισουλφιδικών δεσμών
- Παραγωγή πρωτονιωμένων και ταυτομερικών μορφών (για παράδειγμα η ιστιδίνη έχει ταυτομερικές μορφές, ενώ το ασπαρτικό και το γλουταμικό οξύ έχουν πρωτονιωμένες μορφές)
- Βελτιστοποίηση του δικτύου δεσμών υδρογόνου (συνήθως μέσω επαναπροσανατολισμού των υδροξυλικών ομάδων της πρωτεΐνης)
- Αφαίρεση των μορίων νερού (όπου και αν κρίνεται απαραίτητο)



**Εικόνα 2-1:** Το εργαλείο Protein Preparation Wizard

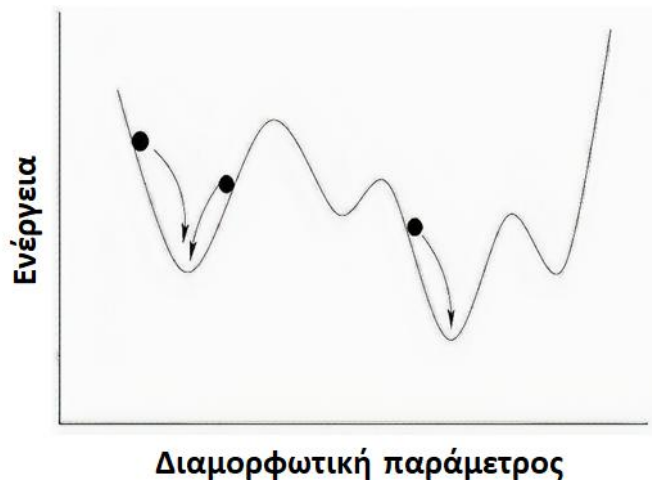
### 2.2.2.2 Ελαχιστοποίηση της ενέργειας

Η εξίσωση δυναμικής ενέργειας ενός συστήματος βιομορίων αποτελεί ένα αρκετά πολύπλοκο και αρκετά πολυδιάστατο πεδίο, ενώ εμφανίζει ένα ολικό ελάχιστο και μεγάλο αριθμό τοπικών ελαχίστων. Ο στόχος της ελαχιστοποίησης της ενέργειας είναι να εντοπίσει ένα τοπικό ελάχιστο, στο οποίο ωστόσο η ενέργεια μπορεί να είναι πολύ υψηλότερη απ'ό,τι στο ολικό ελάχιστο. Η ελαχιστοποίηση της ενέργειας του συστήματός εξασφαλίζει την απομάκρυνση των ανεπιθύμητων van der Waals αλληλεπιδράσεων που ενδέχεται να υπάρχουν, οι οποίες είναι πιθανόν σε αντίθετη περίπτωση να οδηγήσουν σε τοπική στρέβλωση της δομής, με αποτέλεσμα μια ασταθή προσομοίωση.

Για την αναγνώριση των διαμορφώσεων που αντιστοιχούν στη χαμηλότερη ενέργεια του συστήματος, χρησιμοποιούνται οι αλγόριθμοι ελαχιστοποίησης. Με δεδομένη μια συνάρτηση  $f$  των μεταβλητών  $x_1, x_2, \dots, x_n$ , ως ελάχιστο της  $f$  ορίζεται το σημείο στο οποίο η πρώτη παράγωγος της συνάρτησης ως προς κάθε μεταβλητή είναι μηδέν (εξίσωση 2. 12) ενώ παράλληλα οι δεύτερες παράγωγοι εμφανίζουν όλες θετικές τιμές (εξίσωση 2. 13):

$$\frac{\partial f}{\partial x_i} = 0; \quad (2.12)$$

$$\frac{\partial^2 f}{\partial x_i^2} > 0 \quad (2.13)$$



**Σχήμα 2-3:** Η τροχιά της δυναμικής ενέργειας σε συνάρτηση με μια παράμετρο. Οι μέθοδοι ελαχιστοποίησης κινούνται καθοδικά προς το εγγύτερο ελάχιστο. [37]

Στην περίπτωση των αναλυτικών συναρτήσεων, τα ελάχιστα μπορούν να εντοπιστούν με τη χρήση υπολογιστικών μεθόδων. Στην περίπτωση των συστημάτων βιομορίων αυτό δεν είναι εφικτό, καθώς η μορφή της δυναμικής ενέργειας δε μπορεί να εκφρασθεί μέσω μιας αναλυτικής εξίσωσης. Ως εκ τούτου, γι' αυτά τα συστήματα, τα ελάχιστα μπορούν να υπολογιστούν χρησιμοποιώντας αριθμητικές μεθόδους που αλλάζουν σταδιακά τις συντεταγμένες, ώστε να δημιουργήσουν διαμορφώσεις ελαττωμένης ενέργειας, μέχρις ότου βρεθεί η διαμόρφωση με τη χαμηλότερη ενέργεια. Οι πιο κοινοί αλγόριθμοι ελαχιστοποίησης χρησιμοποιούν παραγώγους της ενέργειας ως προς τις συντεταγμένες, με σκοπό να προβλέψουν τη θέση του κοντινότερου ενεργειακού ελαχίστου. Δύο παράγοντες είναι σημαντικοί σε έναν αλγόριθμο ελαχιστοποίησης:

- Η ταχύτητα της σύγκλισης (speed of convergence)
- Οι απαιτήσεις της μνήμης

Δεδομένου ότι δεν έχει βρεθεί η ιδανική μέθοδος που να πληροί όλα τα κριτήρια, εφαρμόζεται συνδυασμός διαφορετικών αλγορίθμων για την ελαχιστοποίηση της ενέργειας των βιολογικών μορίων. Οι περισσότεροι αλγόριθμοι ελαχιστοποίησης μπορούν να κινηθούν καθοδικά πάνω στην ενεργειακή τροχιά, εντοπίζοντας τα ελάχιστα που είναι πλησιέστερα στο σημείο έναρξης (βλ. Σχήμα 2-3). Για τον εντοπισμό περισσότερων ελάχιστων σημείων, απαιτούνται περισσότερα σημεία έναρξης.

### 2.2.2.3 Θέρμανση του συστήματος

Κατά τη φάση της θέρμανσης του συστήματος, οι αρχικές ταχύτητες των ατόμων προσδιορίζονται τυχαία μέσω της κατανομής Maxwell-Boltzmann:

$$p(\mathbf{V}_i) = \left( \frac{m_i}{2\pi k_B T} \right)^{\frac{3}{2}} e^{-\frac{m_i \mathbf{V}_i^2}{2 k_B T}} \quad (2.14)$$

Η εξίσωση 2.14 υπολογίζει την πιθανότητα ενός ατόμου  $i$  και μάζας  $m_i$  σε θερμοκρασία  $T$  να έχει ταχύτητα  $\mathbf{V}_i = (V_{ix}, V_{iy}, V_{iz})$ .

Ο υπολογισμός πραγματοποιείται αρχικά σε χαμηλή θερμοκρασία και η προσομοίωση ξεκινά με τον σταδιακό προσδιορισμό νέων ταχυτήτων σε ελαφρώς μεγαλύτερη θερμοκρασία. Η διαδικασία επαναλαμβάνεται μέχρι το σύστημα να φτάσει την επιθυμητή θερμοκρασία.

### 2.2.2.4 Εξισορρόπηση του συστήματος

Με την ολοκλήρωση του σταδίου της θέρμανσης, το σύστημα έχει αποκτήσει την επιθυμητή θερμοκρασία. Η προσομοίωση συνεχίζεται προκειμένου να ρυθμιστούν παράγοντες όπως η πίεση, η θερμοκρασία και η ενέργεια. Συγκεκριμένα, σκοπός της εξισορρόπησης είναι να συνεχιστεί η προσομοίωση έως ότου αυτοί οι παράγοντες παραμένουν σταθεροί σε συνάρτηση με τον χρόνο. Αν στη συγκεκριμένη φάση, υπάρξει απότομη αύξηση ή μείωση της θερμοκρασίας, οι ταχύτητες κλιμακώνονται με τέτοιο τρόπο ώστε η θερμοκρασία να επιστρέφει στην επιθυμητή τιμή και το σύστημα να φτάσει σε θερμοδυναμική ισορροπία.

#### 2.2.2.5 Παραγωγή τροχιάς

Η τελευταία φάση της προσομοίωσης είναι η αυτή της παραγωγής, κατά την οποία παράγεται η τροχιά του συστήματος (trajectory) στο σημείο θερμοδυναμικής ισορροπίας σε πραγματικό χρόνο - από μερικές χιλιάδες ps μέχρι μs. Οι συντεταγμένες του συστήματος αποθηκεύονται σε διαφορετικές χρονικές στιγμές, ενώ οι θερμοδυναμικές ιδιότητες μπορούν να αποδοθούν με τη μορφή γραφημάτων καταγραφής της πορείας του συστήματος. Τα γραφήματα παρέχουν απαραίτητες πληροφορίες για τον προσδιορισμό των ιδιοτήτων όπως η ενέργεια του συστήματος.

## 2.3 Ομαδοποίηση πρωτεϊνικών δομών από προσομοίωση

Η τροχιά που παράγεται στην τελευταία φάση της προσομοίωσης αποτελείται από διαμορφώσεις του συστήματος σε διαφορετικές χρονικές στιγμές, το πλήθος των οποίων εξαρτάται από το βήμα προσομοίωσης που χρησιμοποιείται και τον συνολικό χρόνο προσομοίωσης. Στην πραγματικότητα, κάθε διαμόρφωση του συστήματος αντιστοιχεί σε μια διαφορετική δομή-διάταξη στο χώρο, καθώς οι θέσεις των ατόμων συνεχώς μεταβάλλονται. Μέσω της ομαδοποίησης των δομών αυτών, με κριτήριο τον βαθμό ομοιότητάς τους, προκύπτουν σύνολα όμοιων δομών. Το σύνολο με τις περισσότερες δομές μπορεί να θεωρηθεί ως το επικρατέστερο και από αυτό μπορεί να εξαχθεί η αντιπροσωπευτική ή αλλιώς κυρίαρχη δομή για την τροχιά του συστήματος.

Μέτρο σύγκρισης του βαθμού ομοιότητας μεταξύ δομών αποτελεί η τετραγωνική ρίζα της μέσης τιμής των τετραγώνων της απόκλισης (Root Mean Square Deviation, RMSD), μέγεθος που υπολογίζει την απόκλιση μεταξύ των συντεταγμένων των ατόμων. Υπολογίζεται ως εξής:

$$RMSD = \sqrt{\frac{\sum_{i=1}^N (r_i - r_j)^2}{N}}$$

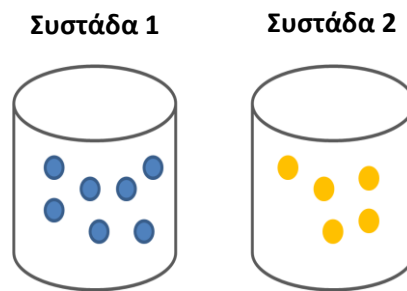
Όπου N ο αριθμός των ατόμων που συγκρίνονται σε κάθε μόριο και  $r_i, r_j$  οι θέσεις των ατόμων που συγκρίνονται. Συνήθως, το μέγεθος RMSD υπολογίζεται λαμβάνοντας υπόψιν μόνο τα άτομα της κυρίας αλυσίδας (κορμού) των πρωτεϊνών που συγκρίνονται. Όσο η τιμή του RMSD μεταξύ των δομών που συγκρίνονται πλησιάζει το 0, τόσο πλησιάζουμε στην πλήρη ταύτισή τους.

Ο αλγόριθμος που χρησιμοποιείται στην παρούσα εργασία για την ομαδοποίηση των δομών που προκύπτουν από προσομοιώσεις είναι ο `gmx cluster` του λογισμικού GROMACS[38]. Η μέθοδος που χρησιμοποιείται είναι αυτή της απλής σύνδεσης (simple linkage). Τα βήματα είναι τα παρακάτω:

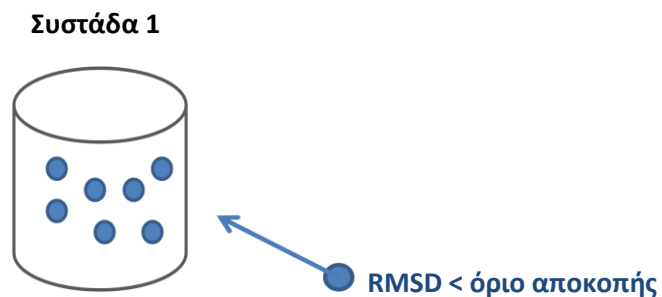
1. Υπολογισμός RMSD για κάθε διαμόρφωση της τροχιάς
2. Αν η τιμή του RMSD για μία διαμόρφωση είναι μικρότερη από το όριο αποκοπής, τότε η δομή προστίθεται στην συστάδα
3. Επανάληψη των δύο παραπάνω βημάτων για όλες τις διαμορφώσεις της τροχιάς

4. Αν η τιμή του RMSD για μία διαμόρφωση είναι μεγαλύτερη από το όριο αποκοπής, συγκριτικά με κάθε ήδη υπάρχουσα συστάδα, τότε μια νέα συστάδα δημιουργείται με τη διαμόρφωση αυτή

Για παράδειγμα, έστω ότι έχουμε τις δύο παρακάτω συστάδες, οι οποίες έχουν διαμορφωθεί μέσω ομαδοποίησης και αποτελούνται από όμοιες πρωτεϊνικές δομές.



Συγκρίνοντας μία νέα δομή με όλες τις δομές στην συστάδα 1, καταλήγουμε στο συμπέρασμα ότι η τιμή του RMSD μεταξύ της δομής αυτής και όλων των δομών της συστάδας 1 είναι κάτω από το όριο αποκοπής. Συνεπώς, η δομή αυτή μπορεί να προστεθεί στην συστάδα 1.





Αντίστοιχα, αν η τιμή του RMSD μεταξύ μιας δομής και όλων των δομών της συστάδας 1 και της συστάδας 2 (υπάρχουσες συστάδες) είναι μεγαλύτερη από το όριο αποκοπής, τότε δημιουργείται μία νέα συστάδα με τη δομή αυτή.



Αφού ολοκληρωθεί αυτή η διαδικασία για όλες τις διαμορφώσεις της τροχιάς, επιλέγεται η συστάδα που περιέχει τις περισσότερες δομές, ως η επικρατέστερη. Συγκρίνοντας όλες τις δομές της συστάδας αυτής μεταξύ τους, η δομή που εμφανίζει τη μικρότερη τιμή RMSD είναι και η αντιπροσωπευτική ή μέση δομή (representative or average structure) της συστάδας.

## 2.4 Αλγόριθμοι προσδιορισμού κοιλοτήτων πρόσδεσης

Η αναγνώριση και ο χαρακτηρισμός κοιλοτήτων πρόσδεσης είναι καίριας σημασίας για τη μετέπειτα διαδικασία του σχεδιασμού φαρμάκων. Οι αλγόριθμοι προσδιορισμού κοιλοτήτων πρόσδεσης είναι χρήσιμοι σε περιπτώσεις που είτε δεν υπάρχουν πληροφορίες για τις κοιλοότητες πρόσδεσης στην πρωτεΐνη-υποδοχέα ενδιαφέροντος είτε είναι αναγκαίος ο προσδιορισμός της δυνατότητας στόχευσης ήδη προσδιορισμένων κοιλοτήτων πρόσδεσης. Αξιοσημείωτο είναι ότι το 60% των μελετών που αφορούν στόχευση υποδοχέων με μικρά μόρια αποτυγχάνουν λόγω του ότι οι υποδοχείς τελικά χαρακτηρίζονται ως μη κατάλληλοι για στόχευση. Μόλις το 10% των πρωτεϊνών που κωδικοποιούνται από το ανθρώπινο γονιδίωμα έχουν χαρακτηριστεί ως κατάλληλες για στόχευση με μικρά μόρια. Επιπλέον, σε συγκεκριμένες περιπτώσεις, είναι επιθυμητή η εύρεση νέων κοιλοτήτων πρόσδεσης με σκοπό την παραγωγή ορισμένης βιολογικής απόκρισης. Οι κοιλοότητες αυτές μπορούν να προσδιοριστούν με τη χρήση υπολογιστικών μεθόδων, όταν δεν υπάρχουν πειραματικά δεδομένα ή δεδομένα από 3D δομές. Παρακάτω παρουσιάζονται υπολογιστικές μέθοδοι προσδιορισμού κοιλοτήτων πρόσδεσης, οι οποίες χρησιμοποιήθηκαν στα πλαίσια της παρούσας εργασίας.

### Ο αλγόριθμος SiteMap

Ο αλγόριθμος SiteMap του λογισμικού Schrodinger Suite έχει αναπτυχθεί με σκοπό την αναγνώριση κοιλοτήτων πρόσδεσης και την πρόβλεψη της δυνατότητας στόχευσής τους. Ο υπολογισμός αποτελείται από τρία στάδια.

Στο πρώτο στάδιο του υπολογισμού, σκοπός είναι ο εντοπισμός των κοιλοτήτων πρόσδεσης. Αρχικά, ένα πλέγμα ακτίνας 1 Å τοποθετείται γύρω από την πρωτεΐνη. Κάθε σημείο του πλέγματος ταξινομείται ως «εσωτερικό» ή «εξωτερικό» της πρωτεΐνης, συγκρίνοντας την απόστασή του από γειτονικά άτομα με την ακτίνα van der Waals (vdW) των ατόμων αυτών. Αν η αναλογία των τετραγώνων αυτών των αποστάσεων είναι μεγαλύτερη από την τιμή 2.5 (κατώφλι), τότε το σημείο θεωρείται «εξωτερικό».

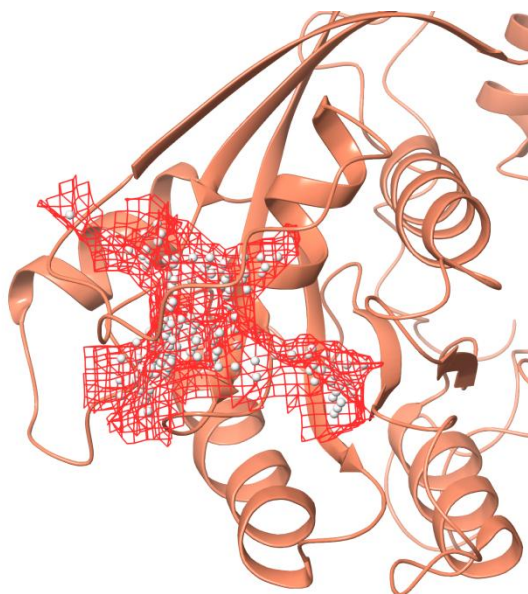
Στη συνέχεια, τα «εξωτερικά» σημεία εξετάζονται προς εξακρίβωση αυτών που έχουν καλές επαφές vdW με την πρωτεΐνη, καθώς επίσης και έναν ικανοποιητικό βαθμό εγκλεισμού ως προς αυτήν. Ο βαθμός εγκλεισμού υπολογίζεται δοκιμάζοντας 110 ομοιόμορφα

χωροθετημένες κατευθύνσεις και προσδιορίζοντας το κλάσμα των ακτινών εντός μιας συγκεκριμένης απόστασης (8 Å). Αν το κλάσμα είναι μικρότερο από ένα κατώφλι (0.5), το σημείο απορρίπτεται. Οι καλές επαφές vdW με την πρωτεΐνη προσδιορίζονται θέτοντας όριο στην ενέργεια vdW στα σημεία του πλέγματος (-1.1 kcal/mol). Το σημείο απορρίπτεται αν η ενέργεια αυτή είναι θετική ή πολύ μικρής τάξης μεγέθους. Τα σημεία που ικανοποιούν και τα δύο παραπάνω κριτήρια προστίθενται σε μία λίστα ως υποψήφια “site points”.

Κατόπιν, τα επιλεγμένα “site points” ομαδοποιούνται για να χαρακτηρίσουν τις διάφορες κοιλοότητες πρόσδεσης. Για να θεωρηθεί ένα “site point” υποψήφιο μέλος μίας ομάδας, τουλάχιστον τρία (3) υποψηφία “site points” πρέπει να βρίσκονται εντός μίας ορισμένης απόστασης από αυτό (συνήθως 1.76 Å). Τα “site points” που δεν ικανοποιούν αυτό το κριτήριο απορρίπτονται. Ακόμα, ομάδες που έχουν λιγότερα από τρία (3) site points απορρίπτονται.

Τελικά, οι ομάδες των “site points” συγχωνεύονται, όταν η κοντινότερη μεταξύ τους απόσταση είναι μικρότερη ή ίση από μία ορισμένη απόσταση (6.5 Å). Ο αλγόριθμος SiteMap συγχωνεύει ομάδες μόνο όταν τουλάχιστον μία από αυτές έχει το πολύ 100 “site points”. Οι ομάδες που τελικά σχηματίζονται συνιστούν τις κοιλοότητες πρόσδεσης. Από προεπιλογή, ο αλγόριθμος επιστρέφει πέντε (5) κοιλοότητες πρόσδεσης, ενώ υπάρχει δυνατότητα για εύρεση έως 10 κοιλοτήτων πρόσδεσης, ελάχιστου μεγέθους 15 “site points”. Ακόμα, υπάρχει και επιλογή για εντοπισμό «ρηχών» κοιλοτήτων πρόσδεσης στην επιφάνεια της πρωτεΐνης, αν αυτό κριθεί απαραίτητο.

Το δεύτερο στάδιο είναι αυτό της οπτικοποίησης. Στο στάδιο αυτό, παράγονται πλέγματα ηλεκτρικού πεδίου και vdW, υπολογίζοντας τις αντίστοιχες ηλεκτροστατικές και vdW αλληλεπιδράσεις μίας ένωσης (probe). Ακολούθως, τα πλέγματα χρησιμοποιούνται για να δημιουργηθούν τα δυναμικά που παράγουν τους χάρτες υδροφιλίας, υδροφοβίας, δοτών και δεκτών δεσμού υδρογόνου, επιφάνειας και μεταλλικών δεσμών, όταν είναι απαραίτητο. Στην παρούσα εργασία, χρησιμοποιούνται για την οπτικοποίηση των κοιλοτήτων οι χάρτες επιφάνειας, οι οποίοι παρουσιάζουν το μέγεθος και το σχήμα του χώρου των κοιλοτήτων που είναι διαθέσιμος για στόχευση από προσδέματα (ligands). Οι χάρτες αυτοί δημιουργούνται λαμβάνοντας υπόψιν τον απωστικό όρο του πλέγματος vdW σε μία θετική τιμή ενεργειακού κατωφλίου (+1 kcal/mol).



**Εικόνα 2-2:** Χάρτης επιφάνειας κοιλότητας πρόσδεσης που προσδιορίστηκε στην πρωτεΐνη KRAS-4B φυσικού τύπου, αναπαράσταση ως κόκκινο δίκτυο. Η πρωτεΐνη αναπαρίσταται με τη μορφή κορδελών (ribbons). Ακόμα, τα "site points" διακρίνονται ως λευκές σφαίρες.

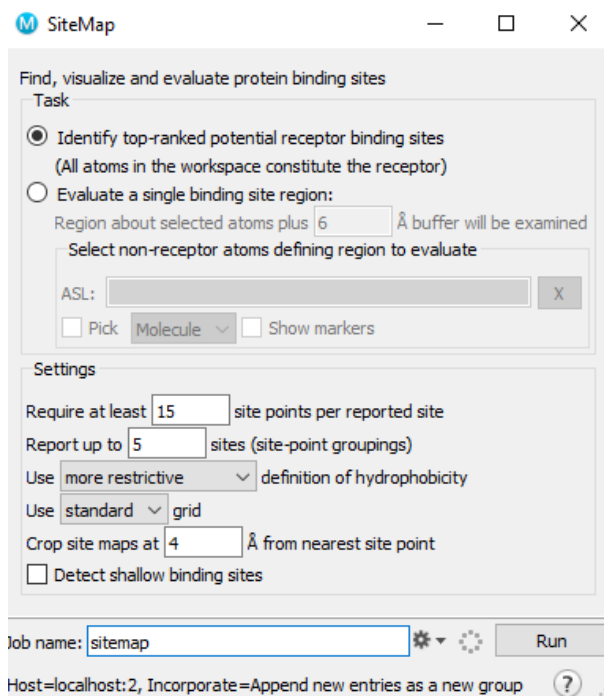
Το τρίτο και τελευταίο στάδιο είναι αυτό της αξιολόγησης των κοιλοτήτων πρόσδεσης, με βάση τα "site points" και τις ενεργειακές ιδιότητες που έχουν ήδη υπολογιστεί στα δύο προηγούμενα στάδια. Για τη σύγκριση και αξιολόγηση των κοιλοτήτων υπολογίζεται το μέγεθος SiteScore, ως ένα σταθμισμένο άθροισμα ιδιοτήτων[39]:

$$\text{SiteScore} = 0.0733n^{\frac{1}{2}} + 0.6688e - 0.20p$$

Όπου  $n$  ο αριθμός των site points (ανώτερο όριο 100),  $e$  ο βαθμός εγκλεισμού και  $p$  το υδροφιλικό σκορ, με μέγιστη τιμή το 1, ούτως ώστε να περιοριστεί η επίδραση της υδροφιλίας σε φορτισμένες ή υψηλά πολικές κοιλότητες πρόσδεσης. Οι πιο σημαντικοί όροι που επιδρούν στην τιμή του SiteScore είναι το μέγεθος της εκάστοτε κοιλότητας πρόσδεσης, μετρούμενο ως αριθμός site points ( $n$ ), καθώς και το πόσο «ανοιχτή» είναι η κοιλότητα ως προς τον διαλύτη, το οποίο υποδεικνύεται από τον βαθμό εγκλεισμού του ( $e$ ). Οι κοιλότητες πρόσδεσης για τις οποίες η τιμή SiteScore είναι κοντά στο ένα χαρακτηρίζονται ως οι περισσότερο υποσχόμενες. Συνήθως, χρησιμοποιείται η τιμή SiteScore 0.8 για τον διαχωρισμό μεταξύ των κοιλοτήτων πρόσδεσης στις οποίες μπορούν να συνδεθούν προσδέματα ( $\text{SiteScore} \geq 0.8$ ) και σε αυτές που δεν έχουν αυτή τη δυνατότητα ( $\text{SiteScore} \leq 0.8$ ).

Title	prepared	SiteScore
▼ job (6)		
site_1	<b>S</b>	1.050215
site_2	<b>S</b>	0.991565
site_3	<b>S</b>	0.964802
site_4	<b>S</b>	0.962980
site_5	<b>S</b>	0.939618
protein	<input checked="" type="checkbox"/>	

**Εικόνα 2-3:** Μορφή αποτελεσμάτων SiteScore



**Εικόνα 2-4:** Το περιβάλλον εργασίας του αλγορίθμου SiteMap

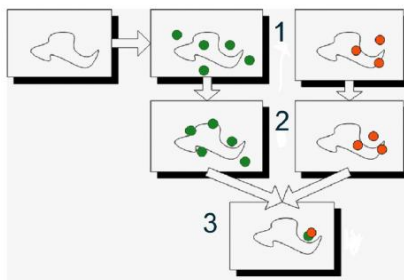
## Ο αλγόριθμος FTMap

Ο αλγόριθμος FTMap αναγνωρίζει κοιλάτητες πρόσδεσης ως “binding hot spots”, δηλαδή περιοχές της επιφάνειας των πρωτεϊνών με σημαντική συνεισφορά στην ελεύθερη ενέργεια πρόσδεσης ενώσεων-προσδεμάτων (ligand binding free energy). Ως εκ τούτου, τα “hot spots” αποτελούν τις ενεργειακά σημαντικές περιοχές των κοιλοτήτων πρόσδεσης. Για το σκοπό αυτό, μελετάται η πρόσδεση μικρών οργανικών μορίων (probes) στην πρωτεΐνη στόχο. Οι περιοχές στις οποίες προσδέονται ομάδες των μορίων αυτών χαρακτηρίζονται τελικά ως “hot spots”. Κύριο πλεονέκτημα των “hot spots” αποτελεί η μικρότερη ευαισθησία τους σε δομικές αλλαγές συγκριτικά με τις κοιλάτητες πρόσδεσης, ενώ μπορούν να αναγνωριστούν σχεδόν σε κάθε μορφή πρωτεϊνικής δομής, με ή χωρίς πρόσδεμα.

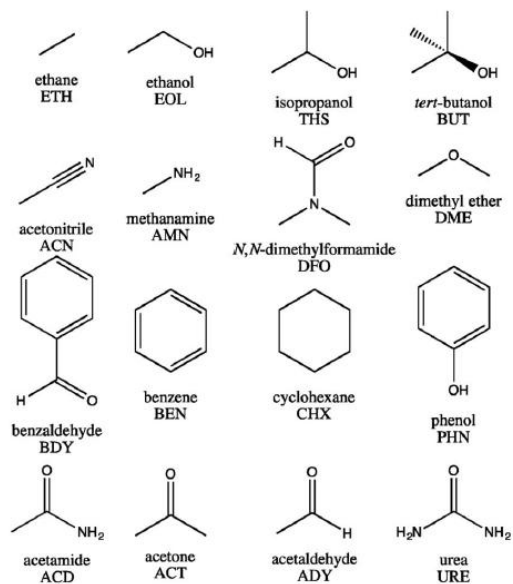
Ο βασικός αλγόριθμος που χρησιμοποιείται αποτελεί υπολογιστικό ανάλογο των πειραματικών τεχνικών κρυσταλλογραφίας ακτίνων Χ ή πυρηνικού μαγνητικού συντονισμού (NMR) και είναι διαθέσιμος διαδικτυακά (<http://ftmap.bu.edu>). Ως είσοδος απαιτείται μόνο η πρωτεϊνική δομή ενδιαφέροντος σε μορφή αρχείου «.pdb». Τα 3 βασικά βήματα που ακολουθούνται είναι τα εξής:

1. 16 μικρά οργανικά μόρια μεταβαλλόμενου μεγέθους, σχήματος και πολικότητας διανέμονται στη μακρομοριακή επιφάνεια της πρωτεΐνης
2. Έπειτα από ελαχιστοποίηση της ενέργειας, εντοπίζονται οι πιο ευνοϊκές διαμορφώσεις για κάθε ένα από τα μόρια και δημιουργούνται ομάδες μορίων ίδιου τύπου
3. Τέλος, οι ομάδες μορίων συγχωνεύονται με βάση τη μέση ελεύθερη ενέργεια και δημιουργούνται τα “hot spots”

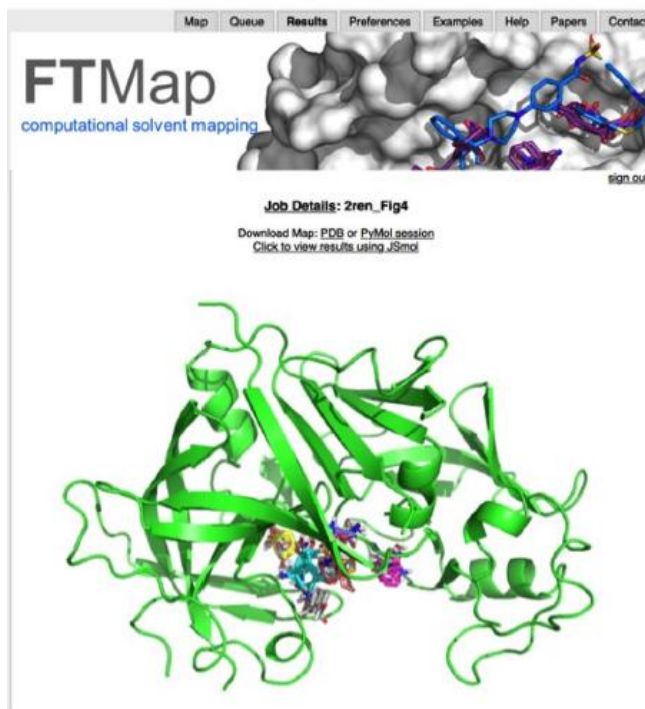
Η διαδικασία περιγράφεται και στο παρακάτω σχήμα.



**Σχήμα 2-4:** Η διαδικασία που ακολουθείται από τον αλγόριθμο FTMap [40]



**Εικόνα 2-5:** Τα 16 μικρά οργανικά μόρια που χρησιμοποιούνται από τον αλγόριθμο FTMap [40]



**Εικόνα 2-6:** Απεικόνιση αποτελεσμάτων στο περιβάλλον εργασίας του αλγορίθμου FTMap για την ρενίνη (PDB ID: 2REN) [41]. Η πρωτεΐνη απεικονίζεται με πράσινο με τη μορφή κορδελών (ribbons) και τα μόρια-προσδέματα με τη μορφή δεσμών σε διάφορα χρώματα [40]

## 2.5 Υπολογισμοί όγκου κοιλοτήτων πρόσδεσης

Μετά τον εντοπισμό των κοιλοτήτων πρόσδεσης, με βάση τους δύο παραπάνω αλγορίθμους, είναι αναγκαίος ο υπολογισμός του όγκου των κοιλοτήτων και ο προσδιορισμός της χρονικής τους μεταβολής. Για τον σκοπό αυτό, χρησιμοποιείται το εργαλείο «Erock» του λογισμικού VMD. Η διαδικασία είναι η εξής:

- Προσδιορίζονται τα αμινοξέα που σχηματίζουν τις κοιλότητες πρόσδεσης, σύμφωνα με τα αποτελέσματα του αλγορίθμου SiteMap
- Τα αμινοξέα εισάγονται στο λογισμικό VMD, όπου και υπολογίζεται το γεωμετρικό κέντρο της κάθε κοιλότητας πρόσδεσης
- Οι συντεταγμένες x,y,z των γεωμετρικών κέντρων που υπολογίστηκαν μεταφέρονται σε αρχεία εισόδου (configuration files), σύμφωνα με το οποία γίνεται ο υπολογισμός του όγκου στο Erock. Για κάθε κοιλότητα, δημιουργούνται σφαίρες, προκειμένου να προσομοιωθεί το μέγεθος και το σχήμα της. Εισάγοντας το αντίστοιχο αρχείο τροχιάς και επιλέγοντας τη δυναμική ανάλυση, το αρχείο εξόδου περιλαμβάνει τιμές του όγκου κοιλοτήτων στις αντίστοιχες τιμές χρόνου.

```
[DEFAULT]
grid_spacing = 0.5
contiguous = false
padding = 1.2
profile = true
contribution = residue

[MUT-1noeff-site1]
include_sphere = 75.2042236328125 62.85306167602539 68.21466064453125 9.5

[MUT-1noeff-site3]
include_sphere = 77.84920501708984 52.09016799926758 76.77365112304688 9.5

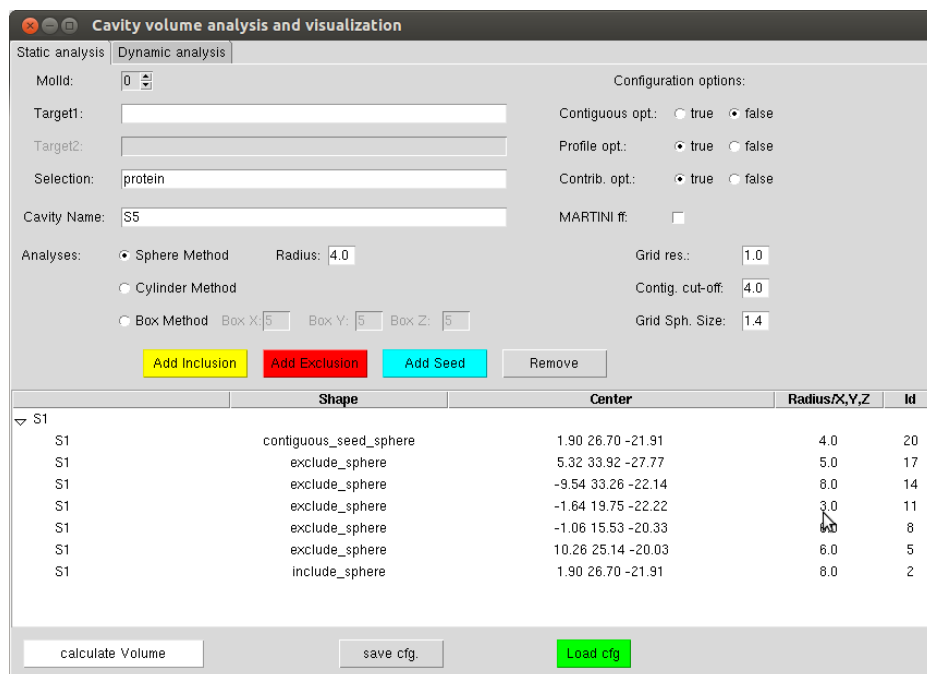
[MUT-1noeff-site2]
include_sphere = 55.5492057800293 56.33028030395508 90.24234008789063 9.5

[MUT-1noeff-site4]
include_sphere = 98.06147003173828 69.90445709228516 86.13387298583984 9.5

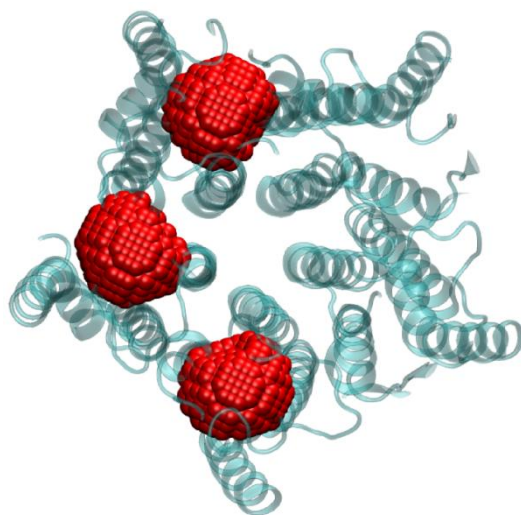
[MUT-1noeff-site5]
include_sphere = 101.47189331054688 62.903594970703125 70.08601379394531 9.5
```

**Εικόνα 2-7:** Το αρχείο εισόδου που χρησιμοποιήθηκε για τις κοιλότητες πρόσδεσης του μεταλλαγμένου διμερούς KRAS-4B





**Εικόνα 2-8:** Το περιβάλλον εργασίας του εργαλείου Erock [42]



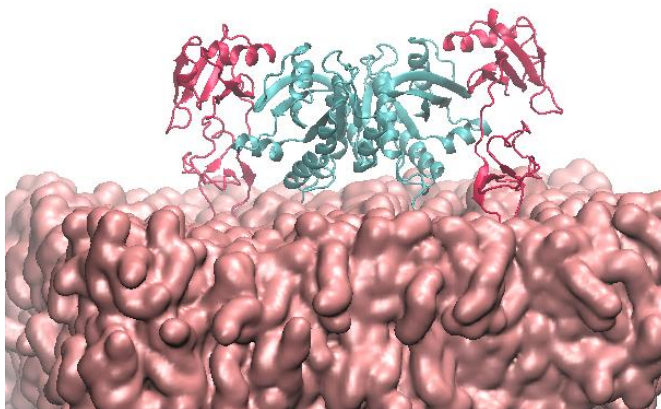
**Εικόνα 2-9:** Οπτικοποίηση των κοιλοτήτων πρόσδεσης στο λογισμικό VMD [42]

### 3 ΑΠΟΤΕΛΕΣΜΑΤΑ - ΣΥΖΗΤΗΣΗ

Στο παρόν κεφάλαιο, παρουσιάζονται τα αποτελέσματα ομαδοποίησης πρωτεϊνικών δομών και εύρεσης των κυρίαρχων δομών, προσδιορισμού κοιλοτήτων πρόσδεσης στις δομές αυτές και χρονικής μεταβολής του όγκου των κοιλοτήτων, για τα συστήματα διμερούς φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B. Οι κοιλοότητες που προκύπτουν συγκρίνονται με τη βιβλιογραφία. Ακόμα, γίνεται σύγκριση μεταξύ των κοιλοτήτων των δύο συστημάτων αλλά και μεταξύ των διαφόρων αντιγράφων του συστήματος φυσικού τύπου.

#### 3.1 Σύστημα διμερούς φυσικού τύπου KRAS-4B συνδεδεμένο με τις πρωτεΐνες Raf στην κυτταρική μεμβράνη

Οι κοιλοότητες πρόσδεσης προσδιορίστηκαν σε δύο συστήματα διμερούς της πρωτεΐνης KRAS-4B, συνδεδεμένης με τις πρωτεΐνες Raf σε μοντέλο κυτταρικής μεμβράνης. Το πρώτο σύστημα αποτελεί αυτό του διμερούς φυσικού τύπου (wild-type) KRAS-4B, το οποίο απεικονίζεται παρακάτω (Andreadelis et al, J Phys Chem B, 2021, in press).



**Εικόνα 3-1:** Το σύστημα διμερούς φυσικού τύπου KRAS-4B. Το διμερές KRAS-4B και οι πρωτεΐνες Raf απεικονίζονται με τη μορφή κορδελών (ribbons) σε γαλάζιο και φούξια χρώμα αντίστοιχα. Η μεμβράνη απεικονίζεται με τη μορφή επιφάνειας σε ροζ χρώμα.

Παραλήφθηκαν προσομοιώσεις Μοριακής Δυναμικής από τη βιβλιογραφία (Andreadelis et al, J Phys Chem B, 2021, in press) και έγινε ανάλυση για να βρεθούν οι δομές κυρίαρχων διαμορφώσεων.

### 3.1.1 Εύρεση κυρίαρχων δομών-διαμορφώσεων μέσω του αλγορίθμου ομαδοποίησης gmx cluster του λογισμικού GROMACS

Η διαδικασία της εύρεσης κυρίαρχων ή αλλιώς αντιπροσωπευτικών δομών πραγματοποιήθηκε για τα έξι (6) αντίγραφα του συστήματος φυσικού τύπου, με χρήση του λογισμικού GROMACS. Αρχικά, τα αρχεία τροχιάς του συστήματος μεταποιήθηκαν, διατηρώντας μία (1) διαμόρφωση του συστήματος ανά πενήντα (50), ελαττώνοντας έτσι τον συνολικό αριθμό των διαμορφώσεων, προκειμένου να διατηρηθεί ο επιθυμητός βαθμός λεπτομέρειας στο σύστημα. Για το σκοπό αυτό, χρησιμοποιήθηκε η παρακάτω εντολή[43]:

```
gmx trjconv -s structure.pdb -f trajectory_file.xtc -o output_wrapped_50.xtc -skip 50
```

Αρχεία εισόδου:

- trajectory\_file.xtc : τροχιά του συστήματος
- structure.pdb : συντεταγμένες ατόμων του συστήματος

Εισάγεται επίσης και ο αριθμός των διαμορφώσεων που παραλείπονται: -skip 50

Αρχεία εξόδου:

- output\_wrapped\_50 : αρχείο τροχιάς με τον ελαττωμένο αριθμό διαμορφώσεων

Έπειτα, το κάθε ένα από τα έξι (6) αρχεία τροχιάς που προκύπτουν διαιρείται σε μικρότερα αρχεία υπο-τροχιών μήκους 100 ns, χρησιμοποιώντας το παρακάτω αρχείο script με όνομα **cut\_traj.sh**:

```

#!/bin/bash
#####
#
#   cut trajectories into   #
#   100ns pieces         #
#
#####

gmx trjconv -f output_wrapped_50.xtc -b 000000 -e 100000 -o 0_100ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 100000 -e 200000 -o 100_200ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 200000 -e 300000 -o 200_300ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 300000 -e 400000 -o 300_400ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 400000 -e 500000 -o 400_500ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 500000 -e 600000 -o 500_600ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 600000 -e 700000 -o 600_700ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 700000 -e 800000 -o 700_800ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 800000 -e 900000 -o 800_900ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 900000 -e 1000000 -o 900_1000ns.xtc

```

**Εικόνα 3-2:** Αρχείο script που χρησιμοποιήθηκε για την διαίρεση των αρχείων τροχιάς των συστημάτων φυσικού τύπου διμερούς KRAS-4B

Για την κατασκευή του παραπάνω script, χρησιμοποιήθηκε η ίδια εντολή με παραπάνω (gmx trjconv), εισάγοντας το αρχείο τροχιάς του συστήματος και ορίζοντας τα διαστήματα που διαβάζονται και γράφονται κάθε φορά στο αρχείο εξόδου με χρήση των επιλογών:

- -b <χρόνος>: πρώτη διαμόρφωση της τροχιάς που διαβάζεται
- -e <χρόνος>: τελευταία διαμόρφωση της τροχιάς που διαβάζεται

Η εντολή που χρησιμοποιήθηκε για την εκτέλεση του script είναι η παρακάτω:

```
bash ../cut_traj.sh
```

Ως έξοδος, δημιουργούνται αρχεία τροχιάς με μήκος 100 ns, ο αριθμός των οποίων διαφέρει, ανάλογα με το αρχικό μήκος τροχιάς που δίνεται ως είσοδος.

Έχοντας ολοκληρώσει τη διαίρεση όλων των αντιγράφων του συστήματος σε τροχιές μήκους 100 ns, ακολουθεί το στάδιο της ομαδοποίησης πρωτεϊνικών δομών και της εύρεσης των κυρίαρχων δομών, με τον τρόπο που περιγράφηκε στην παράγραφο 2.3. Η διαδικασία αυτή πραγματοποιήθηκε τόσο στα αρχεία τροχιάς των διαφορετικών αντιγράφων του συστήματος, όσο και στις αντίστοιχες υπο-τροχιές τους, μήκους 100 ns. Η εντολή που χρησιμοποιήθηκε για την ομαδοποίηση των δομών είναι η παρακάτω:

```
gmx cluster -f trajectory_file.xtc -s structure.pdb -clid clust-id.xvg -cl clusters.pdb -g cluster.log
-cutoff 0.25
```

Αρχεία εισόδου:

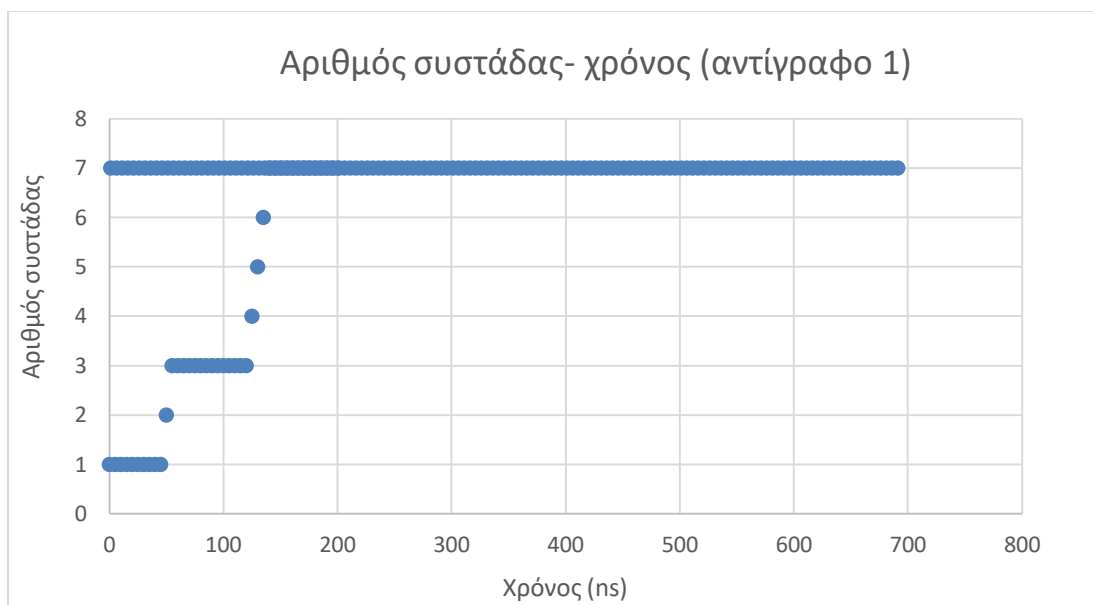
- trajectory\_file.xtc : τροχιά του συστήματος
- structure.pdb : συντεταγμένες ατόμων του συστήματος

Εισάγεται επίσης και η τιμή αποκοπής για το RMSD σε nm, για την οποία γίνονται δοκιμές βελτιστοποίησης (συνήθως 0.25 nm).

Αρχεία εξόδου:

- clust-id.xvg : αρχείο κειμένου που περιέχει δεδομένα αριθμού συστάδας - χρόνου, τα οποία χρησιμοποιούνται για την κατασκευή διαγραμμάτων
- clusters.pdb : αρχείο «.pdb» που περιέχει τις μέσες δομές όλων των ομάδων που προκύπτουν από την ομαδοποίηση
- cluster.log : αρχείο κειμένου που περιέχει τον αριθμό συστάδων που προέκυψαν και τις δομές της κάθε συστάδας

Προκειμένου να προσδιοριστεί η συστάδα πρωτεϊνικών δομών που διατηρείται στο σύνολο της εκάστοτε τροχιάς, συνεπώς περιέχει και τον μεγαλύτερο αριθμό διαμορφώσεων, κατασκευάστηκαν διαγράμματα αριθμού συστάδας - χρόνου. Για παράδειγμα, το διάγραμμα που αφορά τη συνολική τροχιά του αντιγράφου 1 του συστήματος φυσικού τύπου φαίνεται παρακάτω.

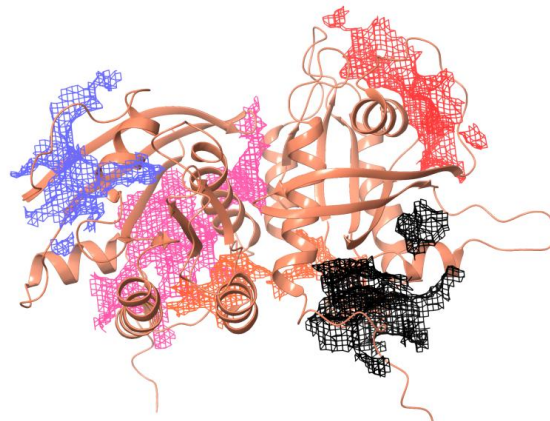


**Διάγραμμα 3-1:** Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 1 του συστήματος φυσικού τύπου KRAS-4B

Από το παραπάνω διάγραμμα, φαίνεται ότι η συστάδα 7 διατηρείται σε όλο το μήκος της τροχιάς του αντιγράφου 1. Συνεπώς, η μέση δομή που αντιστοιχεί σε αυτή την συστάδα επιλέχθηκε ως κυρίαρχη δομή του αντιγράφου 1. Ομοίως, επιλέχθηκαν οι κυρίαρχες δομές για τις υπο-τροχιές των 100 ns του αντιγράφου 1, όπως επίσης και για όλες τις υπόλοιπες τροχιές και υπο-τροχιές των αντιγράφων 2 έως 6. Τα αντίστοιχα διαγράμματα που αφορούν τις συνολικές τροχιές των αντιγράφων 2 έως 6 του συστήματος φυσικού τύπου φαίνονται στο παράρτημα.

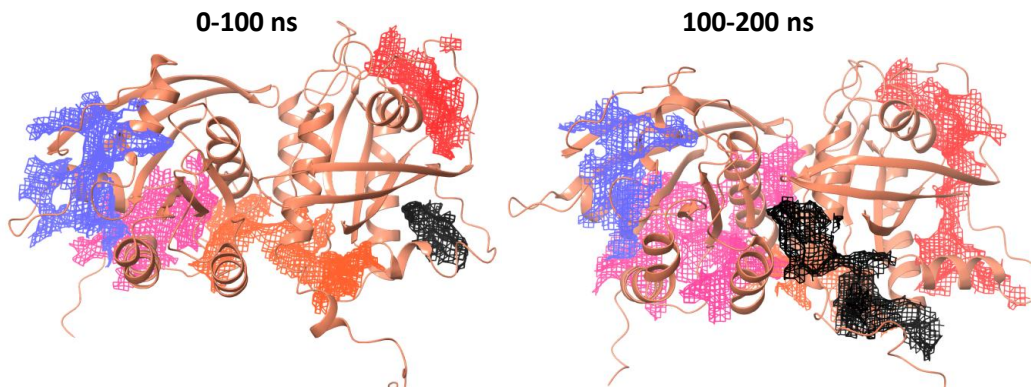
### 3.1.2 Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου SiteMap

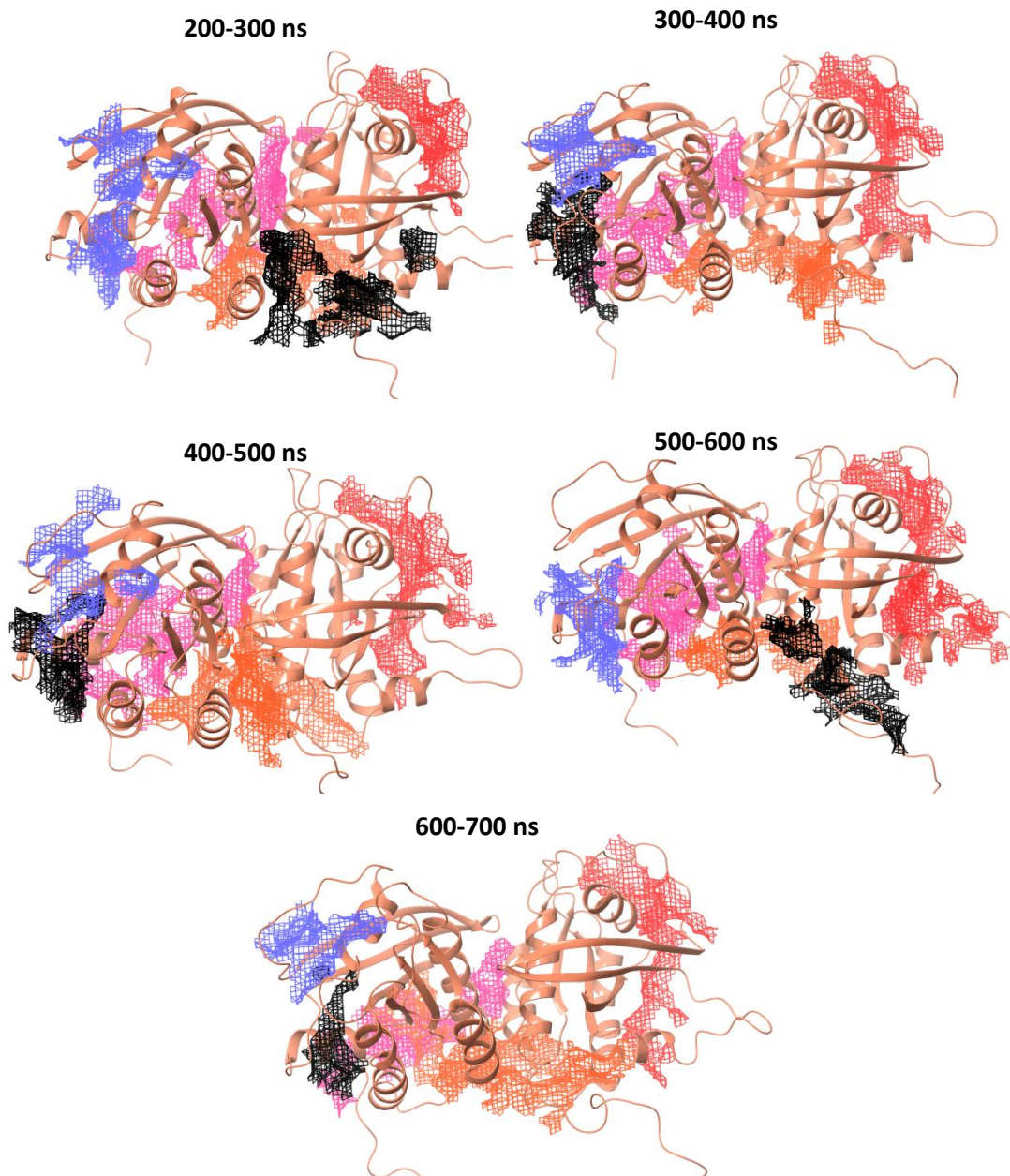
Ο προσδιορισμός των κοιλοτήτων πρόσδεσης πραγματοποιήθηκε στις κυρίαρχες δομές του διμερούς, όπως προέκυψαν από την διαδικασία της ομαδοποίησης. Σε κάθε πρωτεϊνική δομή, προσδιορίστηκαν πέντε (5) κοιλοότητες πρόσδεσης. Τα αποτελέσματα για το αντίγραφο 1 του συστήματος φυσικού τύπου φαίνονται παρακάτω.



**Εικόνα 3-3:** Αποτελέσματα κοιλοτήτων πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 1 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Παρακάτω, φαίνονται τα αποτελέσματα για τις κυρίαρχες δομές των υπο-τροχιών του αντιγράφου 1.



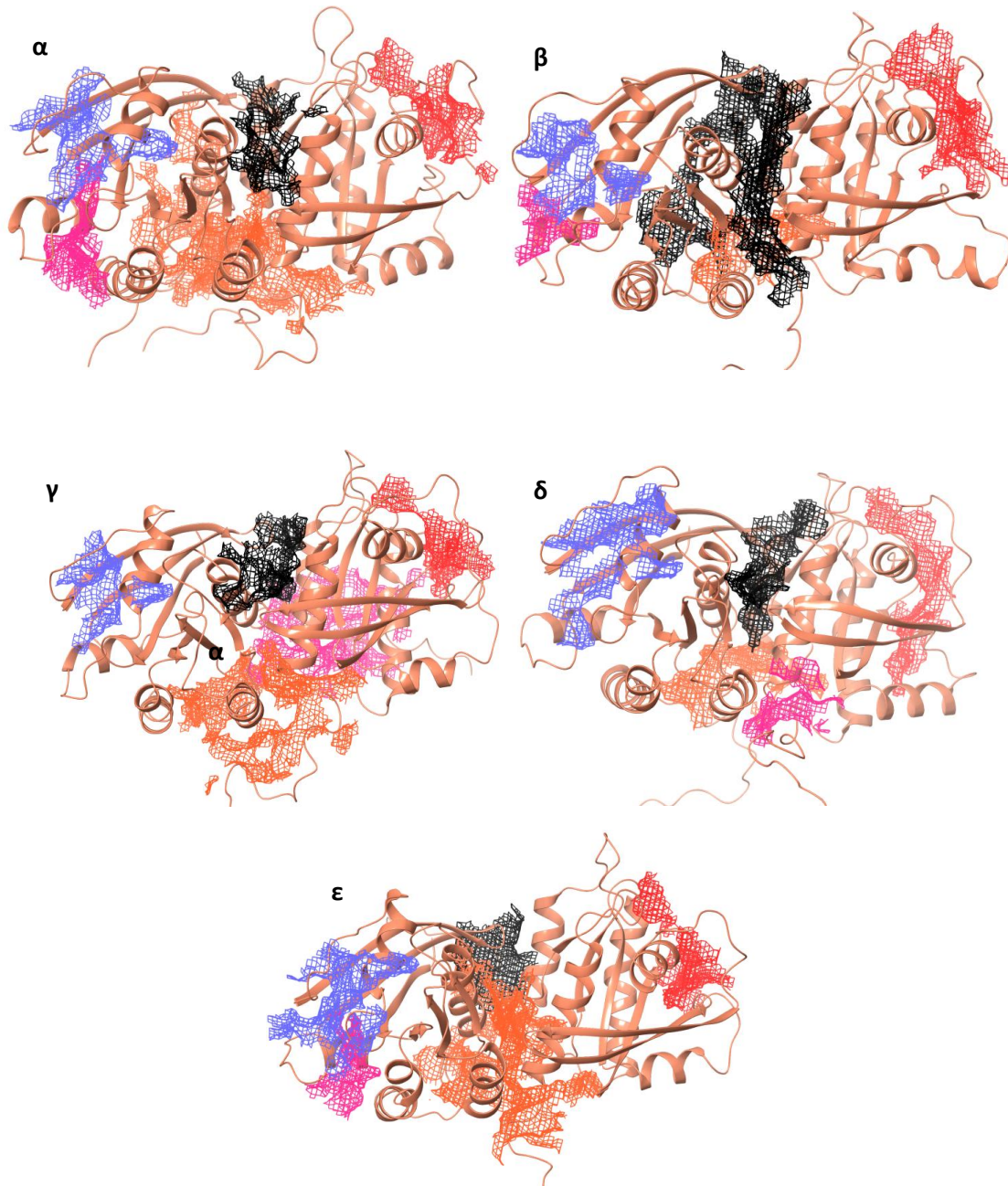


**Εικόνα 3-4:** Αποτελέσματα κοιλοτήτων πρόσδεσης για τις κυρίαρχες δομές των υπο-τροχιών του αντιγράφου 1 του συστήματος φυσικού KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Συγκρίνοντας τις παραπάνω δομές, φαίνεται ότι οι κοιλοότητες πρόσδεσης που προκύπτουν στην κυρίαρχη δομή ολόκληρης της τροχιάς είναι πανομοιότυπες με αυτές που προκύπτουν στις μικρότερες υπο-τροχιές των 100 ns. Η ίδια παρατήρηση ισχύει και για τα υπόλοιπα αντίγραφα του συστήματος φυσικού τύπου. Για το λόγο αυτό, οι κυρίαρχες δομές ολόκληρων των τροχιών είναι και αυτές που παρουσιάζονται παρακάτω και χρησιμοποιούνται για την ανάλυση των αποτελεσμάτων. Παρακάτω παρουσιάζονται τα αποτελέσματα κοιλοτήτων



πρόσδεσης στις κυρίαρχες δομές των υπολοίπων αντιγράφων του συστήματος φυσικού τύπου (αντίγραφο 2 έως 6).



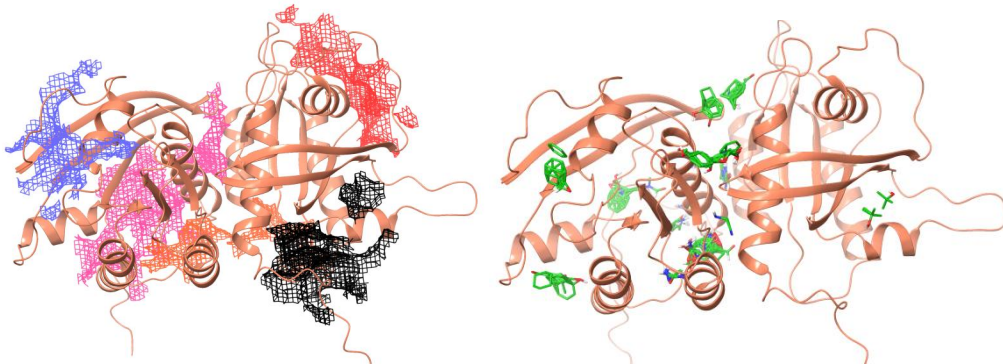
**Εικόνα 3-5:** Αποτελέσματα κοιλοτήτων πρόσδεσης για τις κυρίαρχες δομές των αντιγράφων 2-6 του συστήματος φυσικού τύπου KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ). α) αντίγραφο 2. β) αντίγραφο 3. γ) αντίγραφο 4. δ) αντίγραφο 5. ε) αντίγραφο 6.

Συγκρίνοντας τις κοιλοότητες που προκύπτουν στα διάφορα αντίγραφα του συστήματος φυσικού τύπου, παρατηρείται ότι οι διαφορές στην επιφάνεια της πρωτεΐνης που καλύπτουν είναι μικρές, όπως είναι άλλωστε αναμενόμενο, εφόσον πρόκειται για αντίγραφα του ίδιου συστήματος.

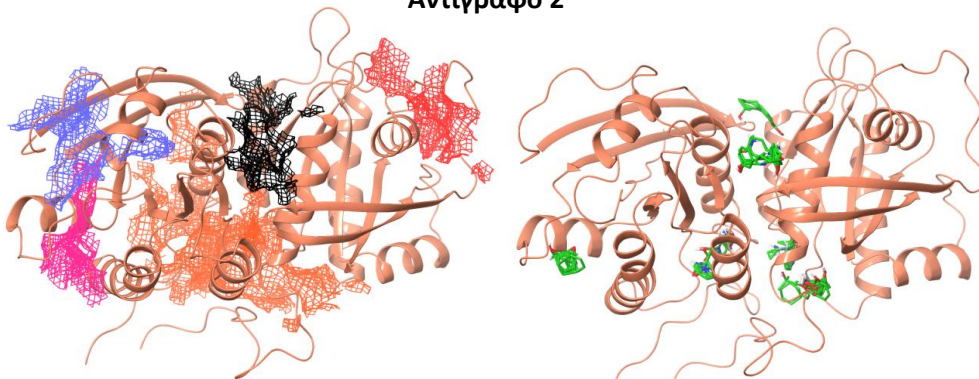
### 3.1.3 Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου FTMap και σύγκριση των δύο αλγορίθμων

Τα αποτελέσματα εντοπισμού κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου FTMap παρουσιάζονται παρακάτω, συγκριτικά με τα αποτελέσματα που προέκυψαν μέσω του αλγορίθμου SiteMap. Στα αποτελέσματα μέσω του αλγορίθμου FTMap, φαίνονται οι ομάδες των προσδεμάτων, οι οποίες και προσδιορίζουν τις πιθανές κοιλοότητες πρόσδεσης, όπως αναφέρθηκε στην παράγραφο 2.4.

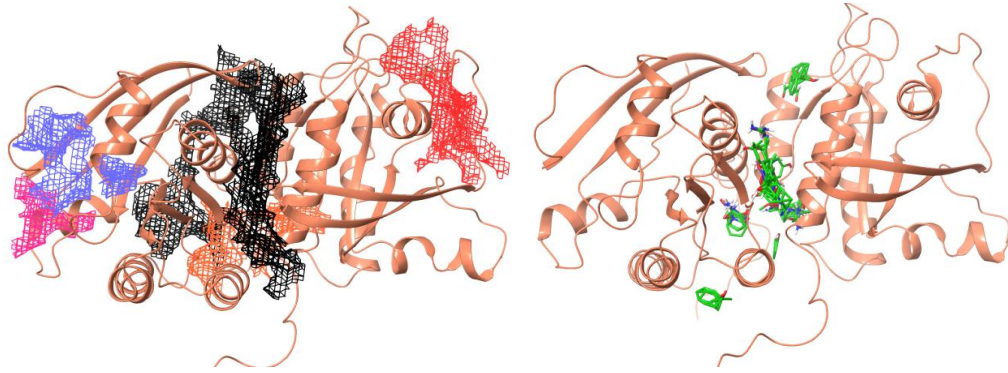
**Αντίγραφο 1**



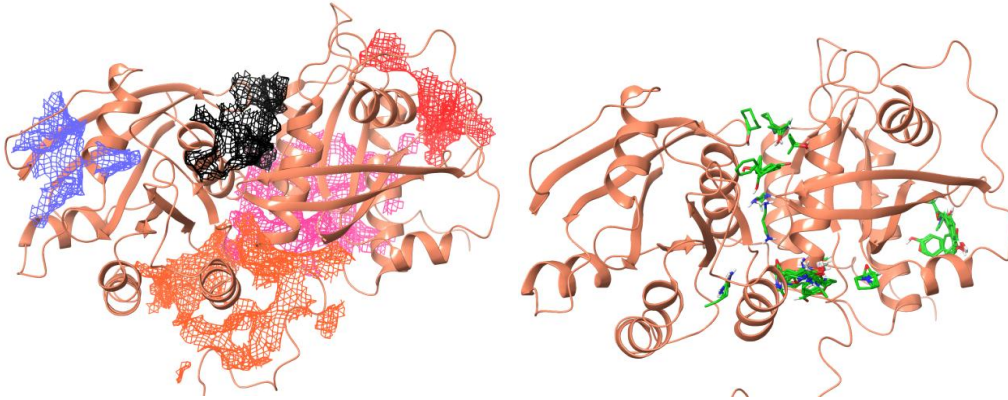
**Αντίγραφο 2**



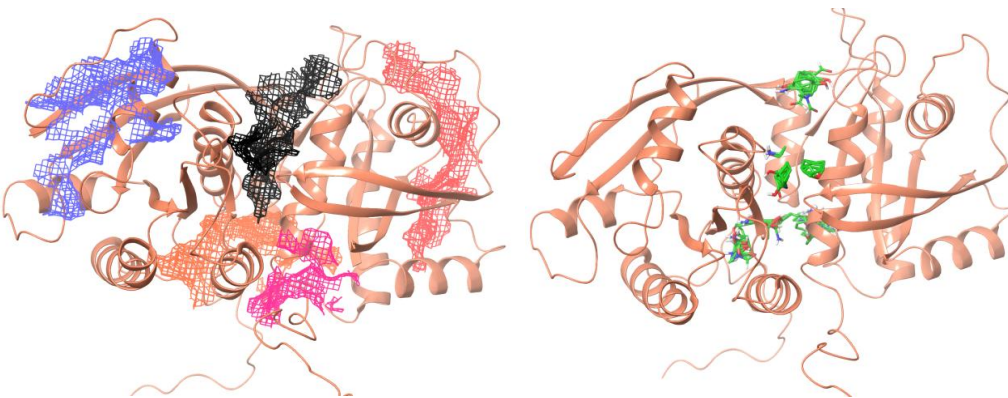
**Αντίγραφο 3**



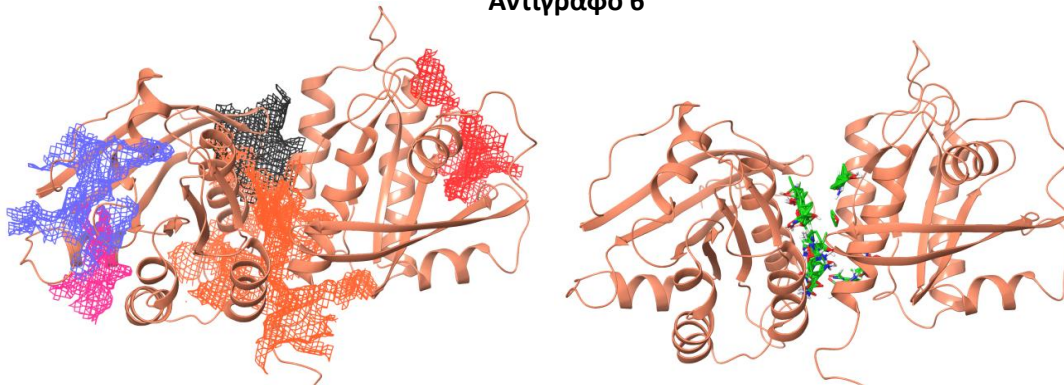
**Αντίγραφο 4**



**Αντίγραφο 5**



Αντίγραφο 6

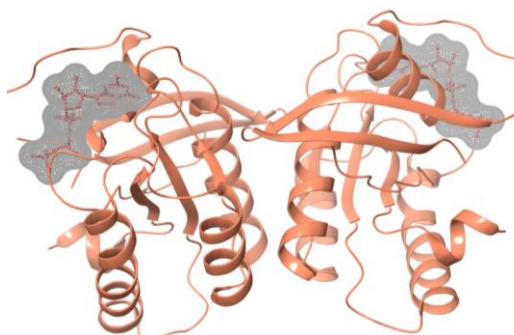


**Εικόνα 3-6:** Σύγκριση των αποτελεσμάτων που προέκυψαν μέσω των αλγορίθμων SiteMap (αριστερά) και FTMap (δεξιά) για τα 6 αντίγραφα του συστήματος φυσικού τύπου. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Οι ομοιότητες μεταξύ των κοιλοτήτων πρόσδεσης που προκύπτουν από τους δύο αλγορίθμους είναι εμφανείς, καθώς οι ομάδες προσδεμάτων του FTMap καταλαμβάνουν τις περιοχές τις πρωτεΐνης που προσδιορίζονται από το SiteMap ή έστω παραπλήσιες περιοχές. Συνεπώς, καταλήγουμε ότι οι δύο αλγόριθμοι SiteMap και FTMap εντοπίζουν τις ίδιες κοιλότητες πρόσδεσης στις διμερείς πρωτεϊνικές δομές KRAS-4B φυσικού τύπου.

### 3.1.4 Σύγκριση αποτελεσμάτων κοιλοτήτων πρόσδεσης με τη βιβλιογραφία

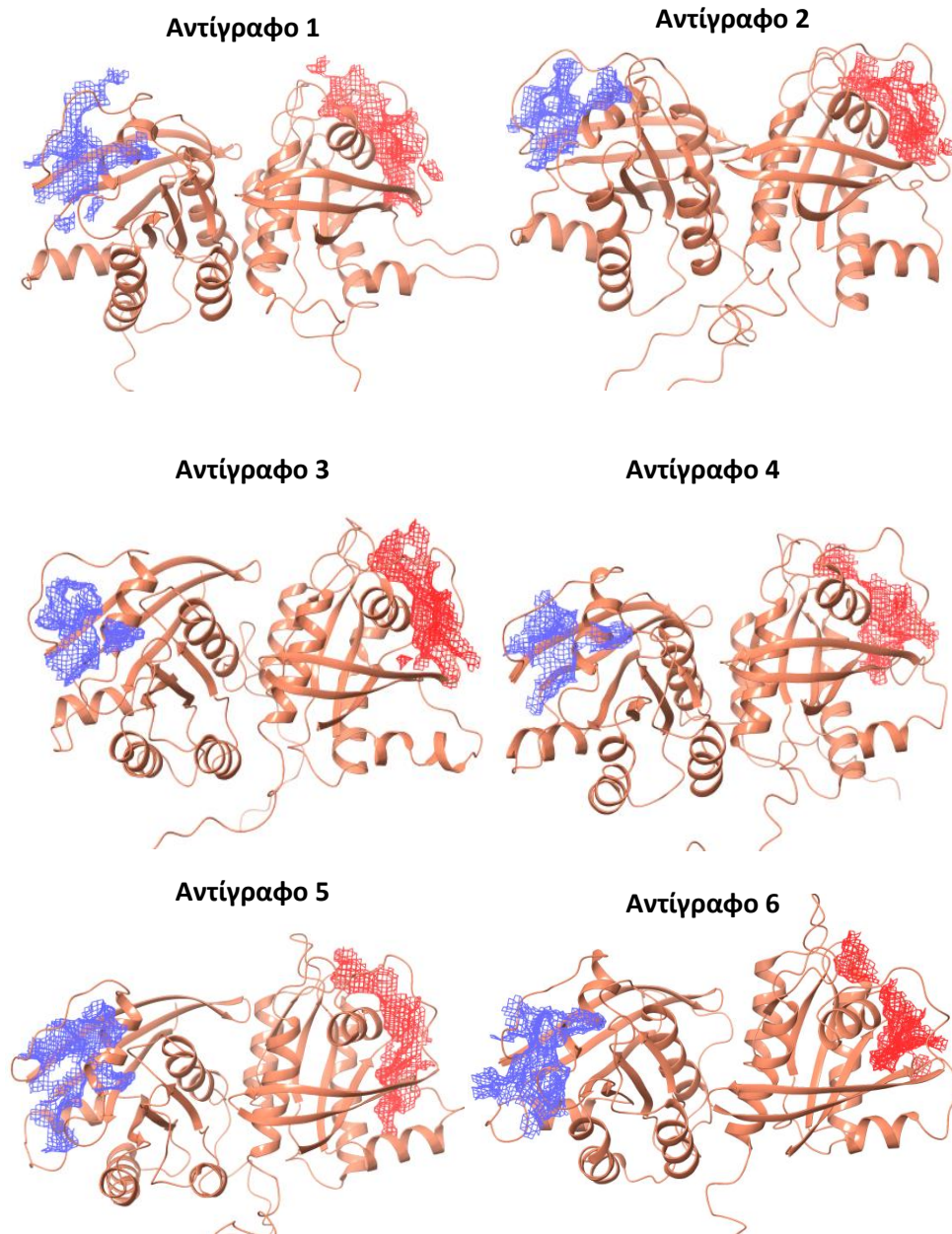
Παρακάτω, φαίνονται οι ορθοστερικές κοιλότητες πρόσδεσης της πρωτεΐνης KRAS-4B , όπως περιγράφηκαν στην παράγραφο 1.3.1.



**Εικόνα 3-7:** Διμερές KRAS-4B φυσικού τύπου συνδεδεμένο με το ενδογενές πρόσδεμα GTP. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι ορθοστερικές κοιλότητες πρόσδεσης με τη μορφή χάρτη επιφάνειας (δίχτυ) γκρι χρώματος (PDB ID: 5VQ2)

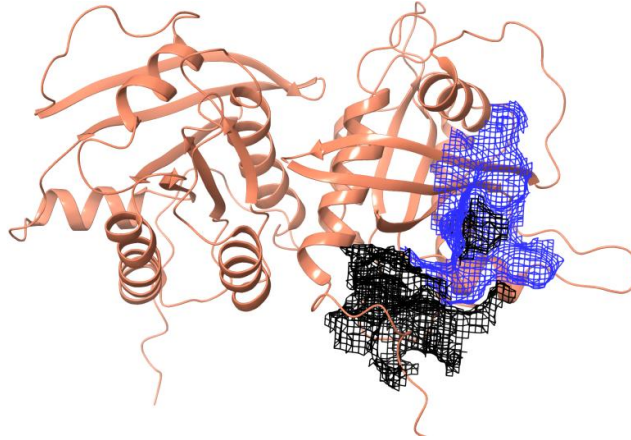
[17]

Τα αμινοξέα του ενεργού κέντρου των κοιλοτήτων είναι τα εξής: 10-16 (p-loop), 17, 18, 28, 29, 30-34 (switch I), 57-58 (switch II), 116-120, 145-147. Συγκρίνοντας με τα αποτελέσματα στις δομές από τα έξι (6) αντίγραφα του συστήματος φυσικού τύπου, παρατηρείται ότι οι ορθοστερικές κοιλοότητες έχουν προσδιοριστεί σε όλες τις δομές, μέσω του αλγορίθμου SiteMap, όπως φαίνεται και στην παρακάτω εικόνα.



**Εικόνα 3-8:** Ορθοστερικές κοιλοότητες πρόσδεσης του διμερούς, όπως εντοπίστηκαν στις δομές των 6 αντιγράφων του συστήματος φυσικού τύπου, μέσω του αλγορίθμου SiteMap. Το διμερές KRAS-4B αναπαρίστανται με τη μορφή κορδελών (ribbons) και οι κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ) σε κόκκινο και μπλε χρώμα.

Επιπλέον, στην παρακάτω εικόνα φαίνεται η σύγκριση της Q61H SI/II κοιλότητας πρόσδεσης, όπως προσδιορίστηκε από τους Cruz-Migoni et al[32], με τα αποτελέσματα του αντιγράφου 1, όπως έχουν προκύψει από τον αλγόριθμο SiteMap.



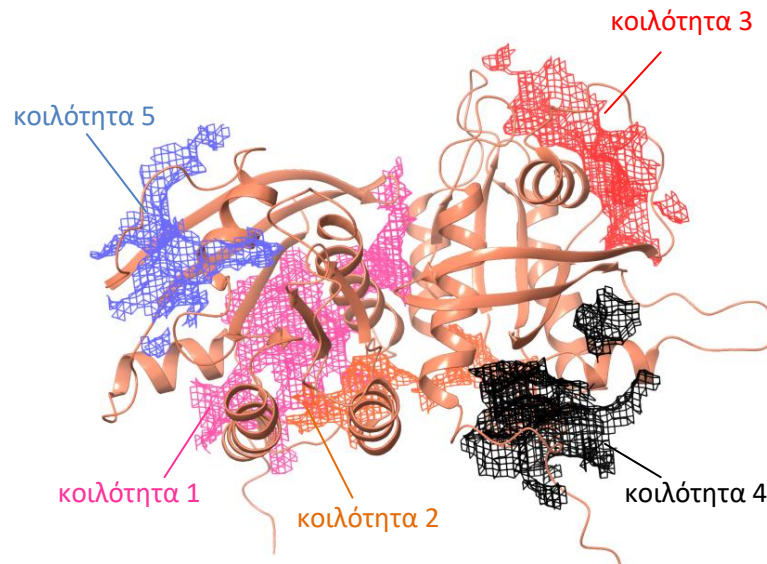
**Εικόνα 3-9:** Σύγκριση αποτελεσμάτων του αντιγράφου 1 του συστήματος φυσικού τύπου (μαύρο χρώμα) με την Q61H SI/II κοιλότητα πρόσδεσης, όπως προσδιορίστηκε από τους Cruz-Migoni et al[32] (μπλε χρώμα). Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Όπως είναι φανερό, ένα μέρος της κοιλότητας που προσδιορίστηκε μέσω SiteMap αλληλοεπικαλύπτεται με ένα μέρος της κοιλότητας SI/II. Παρότι αρκετά παρόμοιες, οι διαφορές μεταξύ των δύο κοιλοτήτων οφείλονται στις διαφορετικές πρωτεϊνικές δομές, στις οποίες έχουν προσδιορισθεί. Συγκεκριμένα, η δομή που χρησιμοποιήθηκε από τους Cruz-Migoni et al είναι αυτή της μεταλλαγμένης KRAS Q61H, σε αντίθεση με την KRAS φυσικού τύπου που χρησιμοποιήθηκε για τους υπολογισμούς μέσω του SiteMap.

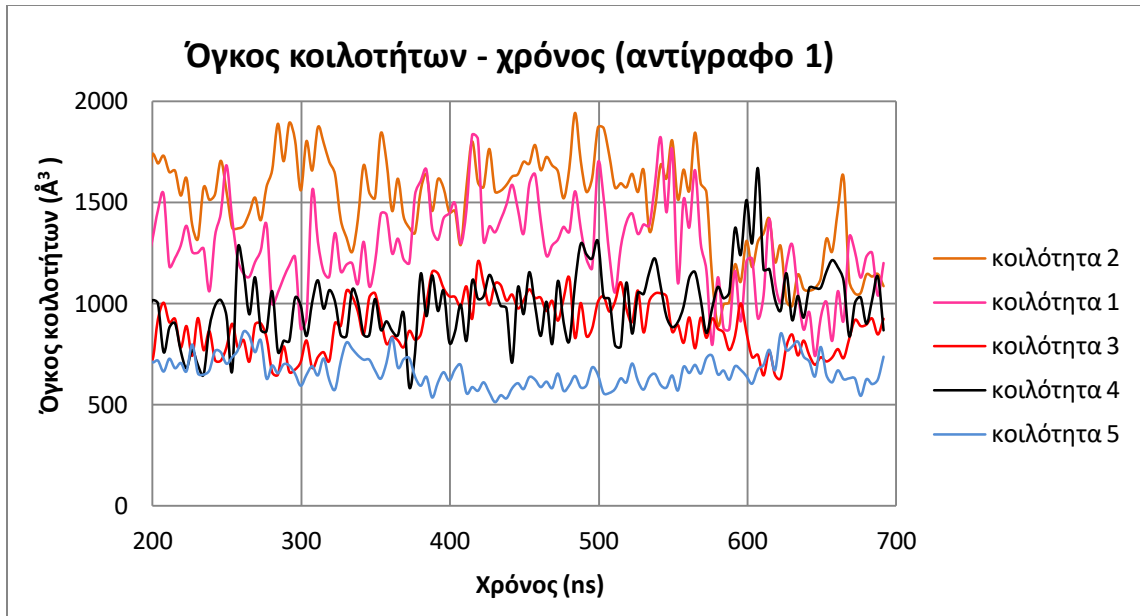
### 3.1.5 Προσδιορισμός χρονικής μεταβολής των κοιλοτήτων πρόσδεσης

Τα αποτελέσματα χρονικής μεταβολής του όγκου των κοιλοτήτων πρόσδεσης για το έξι (6) αντίγραφα του συστήματος φυσικού τύπου παρουσιάζονται παρακάτω. Οι κοιλότητες πρόσδεσης έχουν αριθμηθεί, με τρόπο ώστε η κοιλότητα 1 να αντιστοιχεί στη μεγαλύτερη τιμή SiteScore και η κοιλότητα 5 στη μικρότερη τιμή SiteScore.

#### Αντίγραφο 1



**Εικόνα 3-10:** Κοιλότητες πρόσδεσης στην κυρίαρχη δομή του αντιγράφου 1 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίστανται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).



**Διάγραμμα 3-2:** Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 1 του συστήματος φυσικού τύπου)

Προκειμένου να προσδιοριστεί η διακύμανση του όγκου, υπολογίζεται ο συντελεστής μεταβλητότητας (coefficient of variation) για το σύνολο δεδομένων του όγκου κοιλοτήτων ως εξής[44]:

$$CV(\%) = \left( \frac{\sigma}{\mu} \right) \times 100$$

Όπου  $\sigma$  η τυπική απόκλιση και  $\mu$  η μέση τιμή του όγκου σε κάθε κοιλότητα. Τα δύο αυτά μεγέθη υπολογίστηκαν κάνοντας χρήση των αντίστοιχων συναρτήσεων του EXCEL, STDEV και AVERAGE.

Στην ουσία, ο συντελεστής μεταβλητότητας εκφράζει την τυπική απόκλιση των τιμών ενός μεγέθους ως ποσοστό της μέσης τιμής του, αποτελώντας έτσι μέτρο σχετικής μεταβλητότητας. Γι αυτό το λόγο, μπορεί να χρησιμοποιηθεί ως μέτρο σύγκρισης της μεταβλητότητας δύο ή περισσότερων μεγεθών που έχουν διαφορετικές μέσες τιμές, όπως είναι στην περίπτωσή μας οι όγκοι των διαφορετικών κοιλοτήτων. Σημειώνεται ότι οι συντελεστές μεταβλητότητας υπολογίζονται σε κάθε περίπτωση για το χρονικό διάστημα της τροχιάς, στο οποίο η διακύμανση όγκου φαίνεται να έχει σταθεροποιηθεί, αφήνοντας έξω ένα μικρό κομμάτι της τροχιάς, συνήθως στην αρχή, το οποίο είναι ασταθές. Το χρονικό διάστημα αυτό φαίνεται στα



αντίστοιχα διαγράμματα όγκου κοιλοτήτων-χρόνου. Οι τιμές της τυπικής απόκλισης και του συντελεστή μεταβλητότητας για τις κοιλοότητες πρόσδεσης του αντιγράφου 1 παρουσιάζονται στον παρακάτω πίνακα. Η τυπική απόκλιση παρουσιάζεται ως εξής:

$$\text{Μέση τιμή} \pm \text{τυπική απόκλιση} (\text{\AA})$$

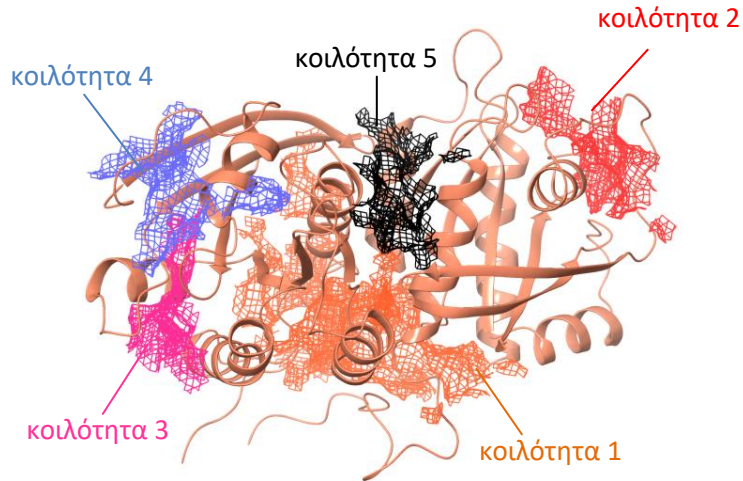
**Σημείωση:** Η μέση τιμή όγκου των κοιλοτήτων πρόσδεσης που θεωρείται ιδανική, ώστε να μπορέσουν να στοχευθούν από μικρά μόρια, είναι  $900 \text{ \AA}^3$ , ενώ κάτω από  $300 \text{ \AA}^3$  και πάνω από  $1800 \text{ \AA}^3$  θεωρούνται ακατάλληλες[45]. Ως εκ τούτου, οι μέσες τιμές των όγκων κοιλοτήτων συγκρίνονται με την το διάστημα αυτό.

**Πίνακας 3-1:** Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 1

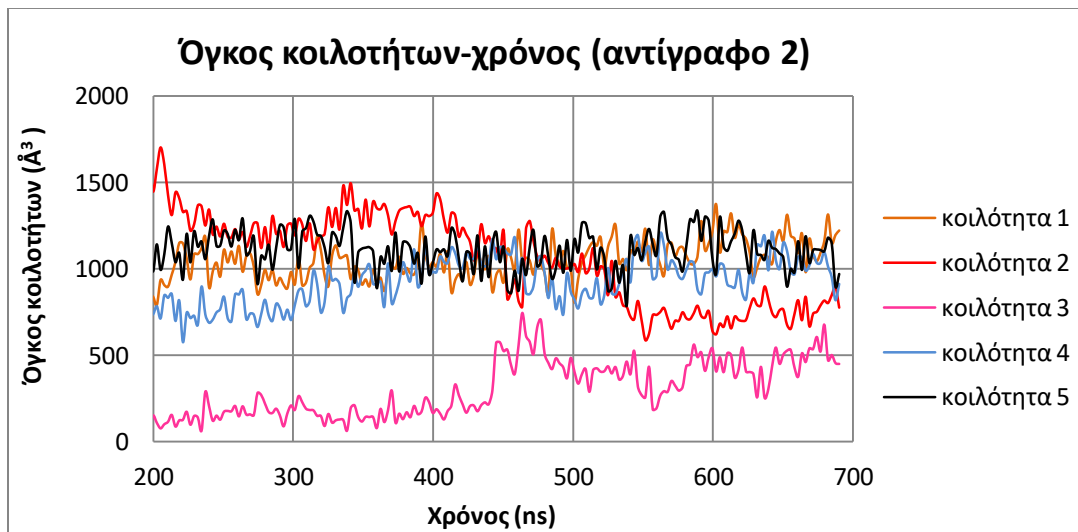
	Κοιλότητα 1	Κοιλότητα 2	Κοιλότητα 3	Κοιλότητα 4	Κοιλότητα 5
<b>Τυπική απόκλιση (Å)</b>	1279.9 ± 224.6	1498.5 ± 239.9	889.0 ± 132.4	995.6 ± 170.7	664.0 ± 74.5
<b>Συντελεστής μεταβλητότητας (Å)</b>	17.5%	16.0%	14.9%	17.1%	11.2%
<b>Κατάλληλη τιμή όγκου</b>	ναι	ναι	ναι	ναι	ναι

Γενικά, δεν υπάρχει κάποια συγκεκριμένη τιμή συντελεστή μεταβλητότητας, η οποία να θεωρείται «καλή». Όπως φαίνεται και από τον τρόπο υπολογισμού του, χαμηλότερες τιμές είναι επιθυμητές, καθώς αντιστοιχούν σε μικρότερη διακύμανση-μεγαλύτερη ανομοιογένεια του όγκου. Σύμφωνα με τον Πίνακα 3-4, η μικρότερη τιμή συντελεστή μεταβλητότητας αντιστοιχεί στην κοιλοότητα 5 και η αμέσως μεγαλύτερη στην κοιλοότητα 3, υποδεικνύοντάς τες ως αυτές με τη μικρότερη διακύμανση όγκου. Η παρατήρηση αυτή είναι αναμενόμενη, καθώς οι δύο αυτές κοιλοότητες αντιστοιχούν στις ορθοστερικές θέσεις πρόσδεσης της KRAS-4B, οι οποίες είναι εκ φύσεως σταθερές. Θεωρώντας ως άνω όριο την τιμή 20%, οι τιμές του συντελεστή μεταβλητότητας είναι και για τις πέντε (5) κοιλοότητες πρόσδεσης του αντιγράφου 1 αρκετά κοντινές και μικρότερες από το άνω όριο, επισημαίνοντας τη μικρή διακύμανση του όγκου σε όλες τις κοιλοότητες. Ως εκ τούτου, συμπεραίνεται η διατήρηση των παραπάνω κοιλοτήτων στον χρόνο.

## Αντίγραφο 2



**Εικόνα 3-11:** Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 2 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).



**Διάγραμμα 3-3:** Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 2 του συστήματος φυσικού τύπου)

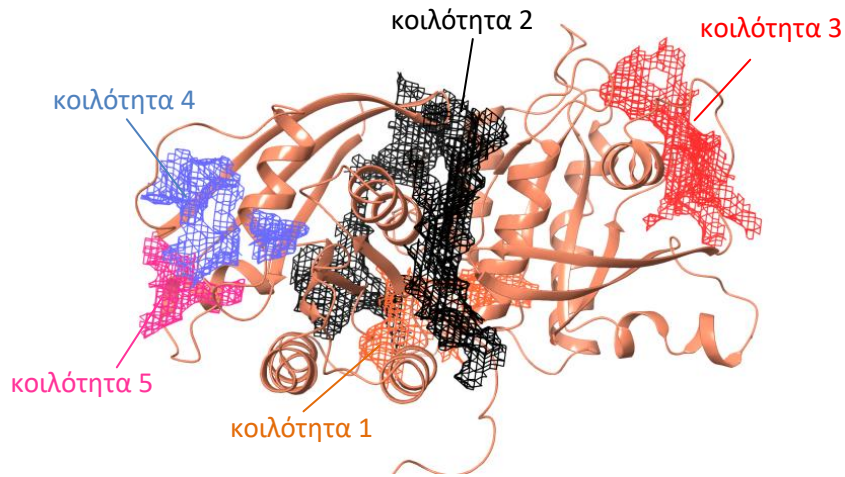
Οι τιμές της τυπικής απόκλισης και του συντελεστή μεταβλητότητας για τις κοιλότητες πρόσδεσης του αντιγράφου 2 παρουσιάζονται στον παρακάτω Πίνακα.

**Πίνακας 3-2:** Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 2

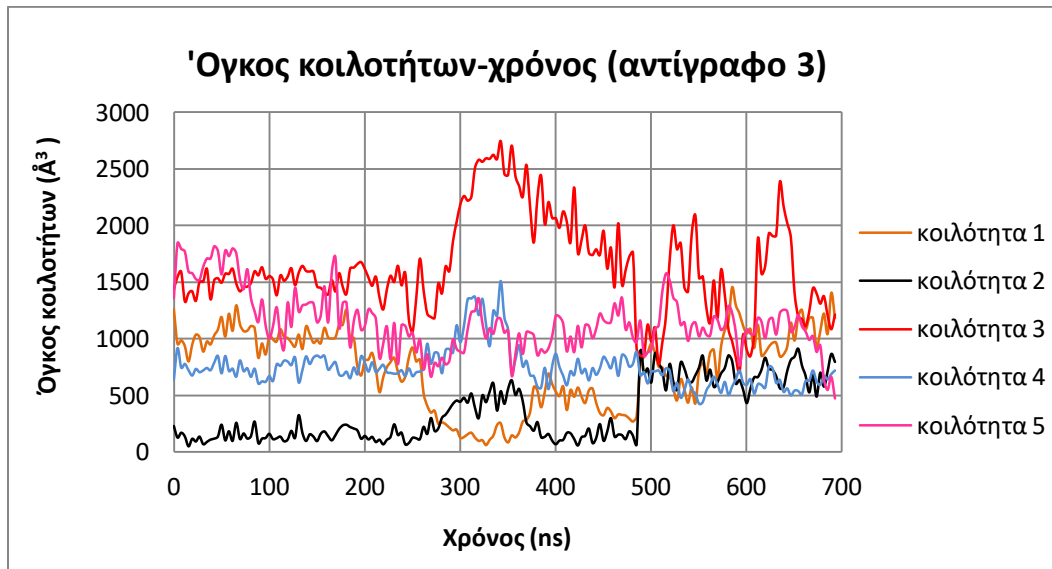
	Κοιλότητα 1	Κοιλότητα 2	Κοιλότητα 3	Κοιλότητα 4	Κοιλότητα 5
<b>Τυπική απόκλιση (Å)</b>	1054.3 ± 108.2	1057.1 ± 260.2	308.9 ± 161.9	930.9 ± 135.3	1109.7 ± 111.0
<b>Συντελεστής μεταβλητότητας (Å)</b>	10.3%	24.6%	52.4%	14.5%	10.0%
<b>Κατάλληλη τιμή όγκου</b>	ναι	ναι	όχι	ναι	ναι

Σύμφωνα με τον Πίνακα 3-5, η μικρότερη τιμή συντελεστή μεταβλητότητας αντιστοιχεί στην κοιλότητα 5, ενώ με ελάχιστη διαφορά ακολουθεί η κοιλότητα 1. Σε αυτή την περίπτωση, ενδιαφέρον προκαλεί ότι αυτές οι κοιλοότητες βρίσκονται στην επιφάνεια διμερισμού της πρωτεΐνης KRAS-4B. Ακόμα, η κοιλότητα 1 είναι αυτή με το μεγαλύτερο SiteScore, γεγονός που την καθιστά υποσχόμενη κοιλότητα για τη στόχευση της επιφάνειας διμερισμού. Οι κοιλοότητες 2 και 4 που αντιστοιχούν στις ορθοστερικές θέσεις πρόσδεσης έχουν και εδώ μικρή διακύμανση όγκου, ωστόσο η τιμή για την κοιλότητα 2 ξεπερνά για λίγο το άνω όριο. Η κοιλότητα 3 παρουσιάζει μεγάλη τιμή συντελεστή μεταβλητότητας όγκου που ξεπερνά επίσης το άνω όριο, συνεπώς ο όγκος της μεταβάλλεται με το χρόνο. Τελικά, οι κοιλοότητες 1,4 και 5 διατηρούν τον όγκο τους.

## Αντίγραφο 3



**Εικόνα 3-12:** Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 3 του συστήματος φυσικού τύπου KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).



**Διάγραμμα 3-4:** Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 3 του συστήματος φυσικού τύπου)

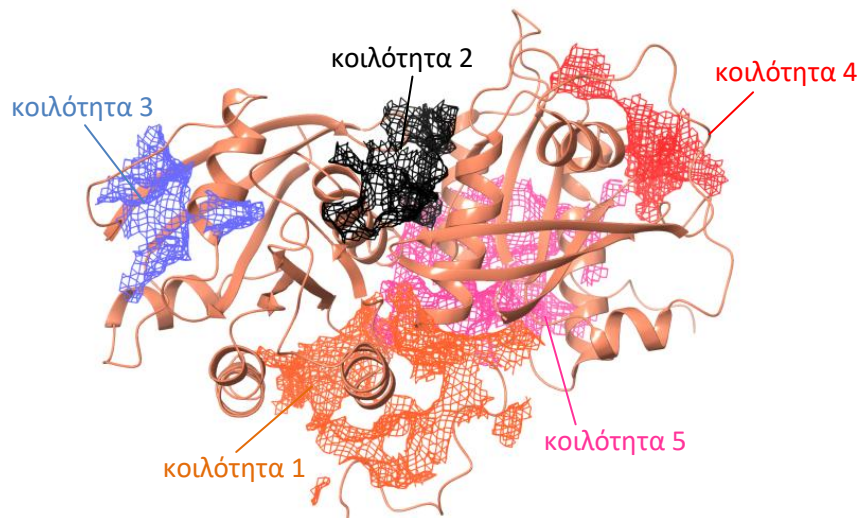
Οι τιμές της τυπικής απόκλισης και του συντελεστή μεταβλητότητας για τις κοιλότητες πρόσδεσης του αντιγράφου 3 παρουσιάζονται στον παρακάτω πίνακα.

**Πίνακας 3-3:** Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 3

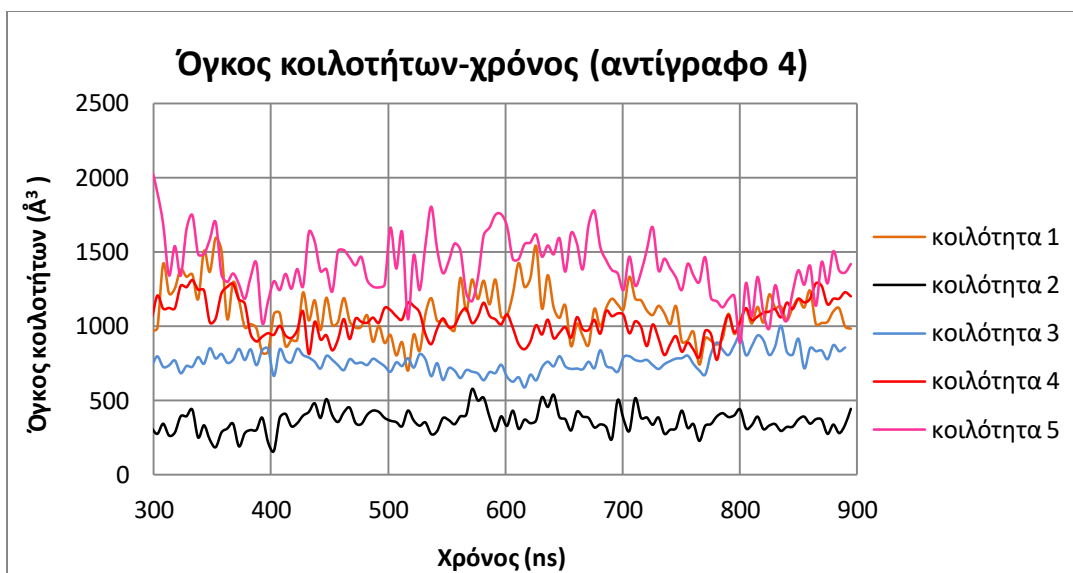
	Κοιλότητα 1	Κοιλότητα 2	Κοιλότητα 3	Κοιλότητα 4	Κοιλότητα 5
<b>Τυπική απόκλιση (Å)</b>	741.6 ± 344.8	353.0 ± 258.9	1632.6 ± 429.0	743.5 ± 177.9	1150.1 ± 265.6
<b>Συντελεστής μεταβλητότητας (Å)</b>	46.5%	73.3%	26.3%	23.9%	23.1%
<b>Κατάλληλη τιμή όγκου</b>	ναι	ναι	ναι	ναι	ναι

Σύμφωνα με τον Πίνακα 3-6, όλες οι τιμές μεταβλητότητας ξεπερνούν το όριο του 20%, οπότε οι κοιλοότητες που αντιστοιχούν στο αντίγραφο 3 του συστήματος φυσικού τύπου δεν διατηρούν τον όγκο τους. Η διαπίστωση αυτή μπορεί να δικαιολογηθεί από το γεγονός ότι σε αυτό το αντίγραφο οι πρωτεΐνες Raf είναι ασταθείς και έχουν αποσπαστεί από την κυτταρική μεμβράνη, δημιουργώντας έτσι ένα ασταθές σύστημα.

## Αντίγραφο 4



**Εικόνα 3-13:** Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 4 του συστήματος φυσικού τύπου KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίκτυ).



**Διάγραμμα 3-5:** Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 4 του συστήματος φυσικού τύπου)

Οι τιμές της τυπικής απόκλισης και του συντελεστή μεταβλητότητας για τις κοιλοότητες πρόσδεσης του αντιγράφου 4 παρουσιάζονται στον παρακάτω πίνακα.

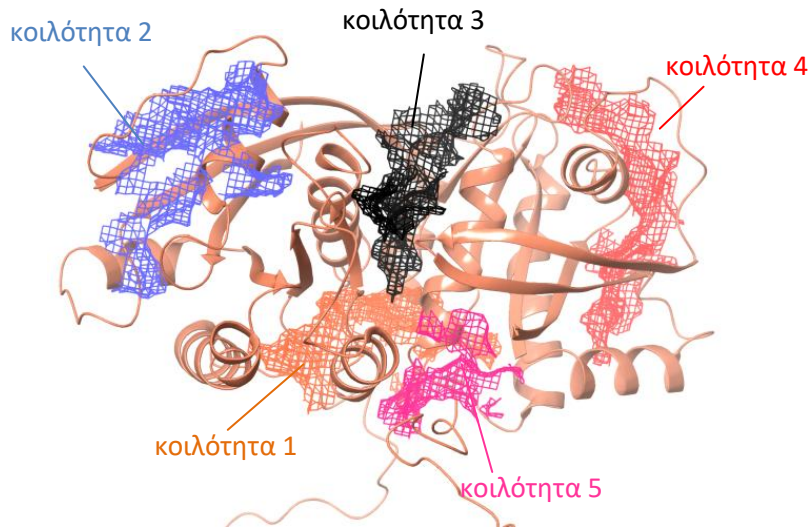
**Πίνακας 3-4:** Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 4

	Κοιλότητα 1	Κοιλότητα 2	Κοιλότητα 3	Κοιλότητα 4	Κοιλότητα 5
<b>Τυπική απόκλιση (Å)</b>	1092.4 ± 169.0	360.4 ± 72.8	766.8 ± 71.8	1038.5 ± 120.6	1399.3 ± 189.1
<b>Συντελεστής μεταβλητότητας (Å)</b>	15.5%	20.2%	9.4%	11.6%	13.5%
<b>Κατάλληλη τιμή όγκου</b>	ναι	ναι	ναι	ναι	ναι

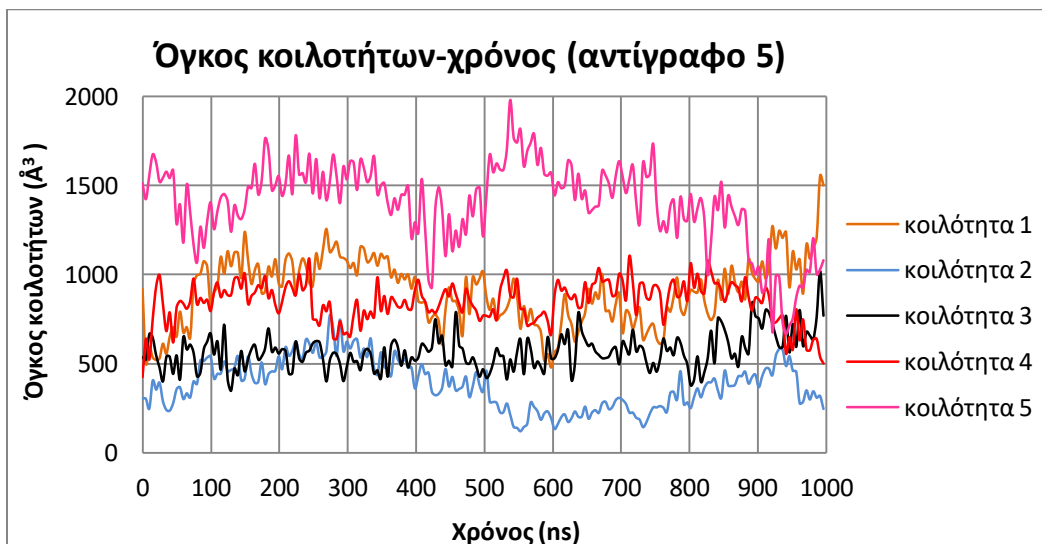
Όπως φαίνεται στον Πίνακα 3-7, οι μικρότερες τιμές του συντελεστή μεταβλητότητας αντιστοιχούν και πάλι στις ορθοστερικές κοιλοότητες πρόσδεσης 3 και 4. Ακολουθούν με μικρή διαφορά οι κοιλοότητες 1 και 5, μέρος των οποίων βρίσκεται στη διεπιφάνεια του διμερούς KRAS-4B. Για όλες τις παραπάνω κοιλοότητες, ο συντελεστής μεταβλητότητας όγκου βρίσκεται κάτω από το άνω όριο που έχει τεθεί, συνεπώς ο όγκος τους διατηρείται. Όσο αφορά στην κοιλότητα 2, ο συντελεστής μεταβλητότητας όγκου που της αντιστοιχεί ξεφεύγει ελάχιστα πάνω από το 20%, οπότε και πάλι θα μπορούσαμε να θεωρήσουμε ότι ο όγκος της διατηρείται. Κατά συνέπεια, όλες οι κοιλοότητες που αντιστοιχούν στο αντίγραφο 4 διατηρούν τον όγκο τους.

Η μέση τιμή του όγκου βρίσκεται για όλες εκτός της κοιλότητας 2 εντός του επιθυμητού διαστήματος.

## Αντίγραφο 5



**Εικόνα 3-14:** Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 5 του συστήματος φυσικού τύπου KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).



**Διάγραμμα 3-6:** Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 5 του συστήματος φυσικού τύπου)

Οι τιμές της τυπικής απόκλισης και του συντελεστή μεταβλητότητας για τις κοιλότητες πρόσδεσης του αντιγράφου 5 παρουσιάζονται στον παρακάτω πίνακα.

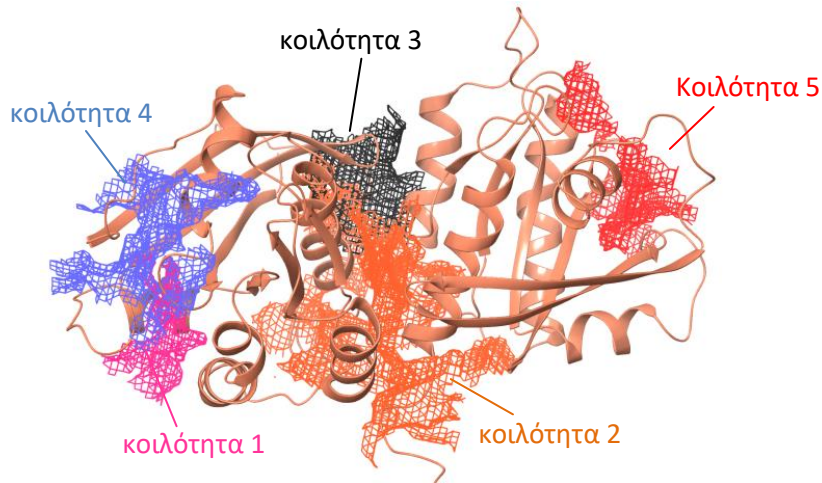
**Πίνακας 3-5:** Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 5

	Κοιλότητα 1	Κοιλότητα 2	Κοιλότητα 3	Κοιλότητα 4	Κοιλότητα 5
<b>Τυπική απόκλιση (Å)</b>	906.5 ± 188.0	382.2 ± 136.5	567.5 ± 100.2	844.5 ± 116.8	1398.9 ± 232.8
<b>Συντελεστής μεταβλητότητας (Å)</b>	20.7%	35.7%	17.7%	13.8%	16.6%
<b>Κατάλληλη τιμή όγκου</b>	ναι	ναι	ναι	ναι	ναι

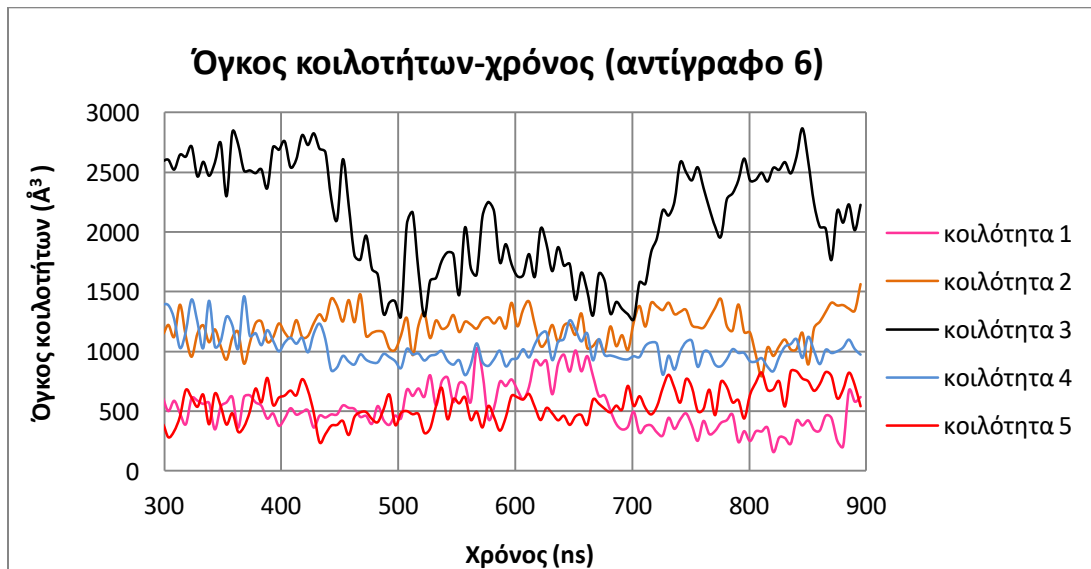
Όπως φαίνεται στον Πίνακα 3-8, η κοιλότητα 4 που αντιστοιχεί στην ορθοστερική θέση πρόσδεσης του ενός μονομερούς παρουσιάζει τη χαμηλότερη τιμή συντελεστή μεταβλητότητας όγκου. Ακολουθούν οι κοιλότητες 5 και 3 με χαμηλές τιμές συντελεστή μεταβλητότητας, κάτω από το άνω όριο. Οι κοιλότητες 3,4 και 5 διατηρούν τον όγκο τους στο χρόνο. Η κοιλότητα 1 παρουσιάζει συντελεστή μεταβλητότητας που ξεπερνά κατάτι το όριο, συνεπώς μπορεί να θεωρηθεί ότι και αυτή διατηρεί τον όγκο της. Ακόμα, οι μέσες τιμές του όγκου των κοιλοτήτων 1,3,4 και 4 βρίσκονται εντός του επιθυμητού διαστήματος. Η κοιλότητα 2 φαίνεται να έχει αρκετά χαμηλή μέση τιμή όγκου, εκτός του επιθυμητού διαστήματος, όπως επίσης και συντελεστή μεταβλητότητας όγκου που ξεφεύγει από το άνω όριο, υποδεικνύοντας την ως ασταθή.



## Αντίγραφο 6



**Εικόνα 3-15:** Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του αντιγράφου 6 του συστήματος φυσικού τύπου διμερούς KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίστανται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).



**Διάγραμμα 3-7:** Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (αντίγραφο 6 του συστήματος φυσικού τύπου)

Οι τιμές της τυπικής απόκλισης και του συντελεστή μεταβλητότητας για τις κοιλότητες πρόσδεσης του αντιγράφου 6 παρουσιάζονται στον παρακάτω πίνακα.

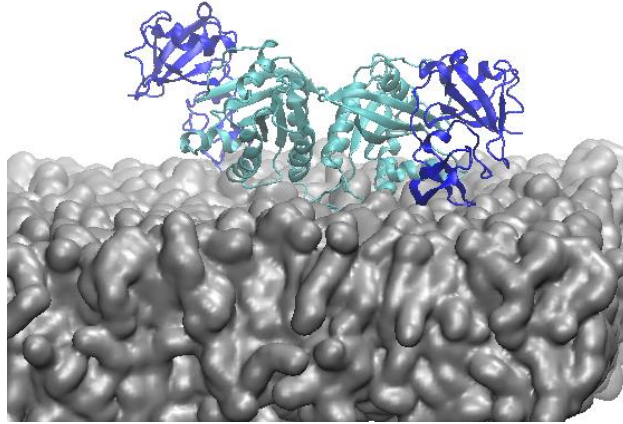
**Πίνακας 3-6:** Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο αντίγραφο 6

	Κοιλότητα 1	Κοιλότητα 2	Κοιλότητα 3	Κοιλότητα 4	Κοιλότητα 5
<b>Τυπική απόκλιση (Å)</b>	520.7 ± 187.3	1201.5 ± 140.7	2111.7 ± 453.8	1019.3 ± 127.8	557.8 ± 139.8
<b>Συντελεστής μεταβλητότητας (Å)</b>	36.0%	11.7%	21.5%	12.5%	25.1%
<b>Κατάλληλη τιμή όγκου</b>	ναι	ναι	όχι	ναι	ναι

Όπως φαίνεται στον Πίνακα 3-9, οι κοιλοότητες 2 και 4 έχουν συντελεστή μεταβλητότητας χαμηλότερο από το άνω όριο, σε αντίθεση με τις κοιλοότητες 1,3 και 5. Όμοια με το αντίγραφο 3, οι πρωτεΐνες Raf είναι ασταθείς και έχουν αποσπαστεί από την κυτταρική μεμβράνη, δημιουργώντας ένα ασταθές σύστημα.

## 3.2 Σύστημα διμερούς της μεταλλαγμένης (G12D) KRAS-4B συνδεδεμένο με τις πρωτεΐνες Raf στην κυτταρική μεμβράνη

Το σύστημα διμερούς της μεταλλαγμένης (G12D) KRAS-4B απεικονίζεται παρακάτω.



**Εικόνα 3-16:** Το σύστημα διμερούς μεταλλαγμένης πρωτεΐνης (G12D) KRAS-4B. Το διμερές KRAS-4B και οι πρωτεΐνες Raf απεικονίζονται με τη μορφή κορδελών (ribbons) σε γαλάζιο και μπλε χρώμα αντίστοιχα. Η μεμβράνη απεικονίζεται με τη μορφή επιφάνειας σε γκρι χρώμα.

Παραλήφθηκαν προσομοιώσεις Μοριακής Δυναμικής από τη βιβλιογραφία (Andreadelis et al, J Phys Chem B, 2021, in press) και έγινε ανάλυση για να βρεθούν οι δομές κυρίαρχων διαμορφώσεων.

### 3.2.1 Εύρεση κυρίαρχων δομών-διαμορφώσεων μέσω του αλγορίθμου ομαδοποίησης gmx cluster του λογισμικού GROMACS

Η διαδικασία που ακολουθείται για την εύρεση κυρίαρχων δομών είναι η ίδια με αυτή που εφαρμόστηκε στα συστήματα διμερούς φυσικού τύπου (Παράγραφος 2.3). Μετά τη διαδικασία ελάττωσης των διαμορφώσεων, το αρχείο τροχιάς που προκύπτει διαιρείται σε μικρότερες αρχεία υπο-τροχιών μήκους 100 ns, χρησιμοποιώντας το παρακάτω αρχείο script.

```

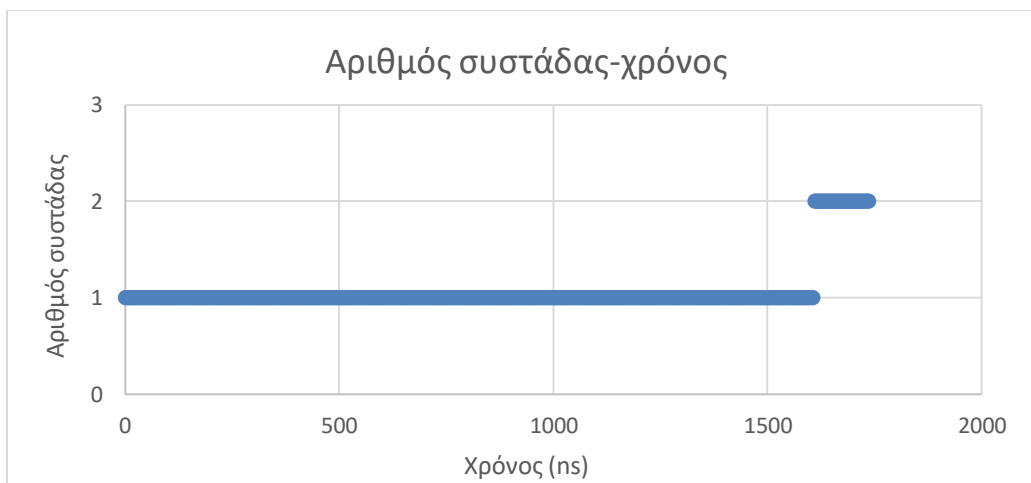
#!/bin/bash
#####
#
#   cut trajectories into   #
#   100ns pieces         #
#
#####

gmx trjconv -f output_wrapped_50.xtc -b 000000 -e 100000 -o 0_100ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 100000 -e 200000 -o 100_200ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 200000 -e 300000 -o 200_300ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 300000 -e 400000 -o 300_400ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 400000 -e 500000 -o 400_500ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 500000 -e 600000 -o 500_600ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 600000 -e 700000 -o 600_700ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 700000 -e 800000 -o 700_800ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 800000 -e 900000 -o 800_900ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 900000 -e 1000000 -o 900_1000ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1000000 -e 1100000 -o 1000_1100ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1100000 -e 1200000 -o 1100_1200ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1200000 -e 1300000 -o 1200_1300ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1300000 -e 1400000 -o 1300_1400ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1400000 -e 1500000 -o 1400_1500ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1500000 -e 1600000 -o 1500_1600ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1600000 -e 1700000 -o 1600_1700ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1700000 -e 1800000 -o 1700_1800ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1800000 -e 1900000 -o 1800_1900ns.xtc
gmx trjconv -f output_wrapped_50.xtc -b 1900000 -e 2000000 -o 1900_2000ns.xtc

```

**Εικόνα 3-17:** Αρχείο script που χρησιμοποιήθηκε για την διαίρεση των αρχείων τροχιών των συστημάτων φυσικού τύπου διμερούς KRAS-4B

Έχοντας ολοκληρώσει τη διαίρεση όλων των αντιγράφων του συστήματος σε τροχιές μήκους 100 ns, ακολουθεί το στάδιο της ομαδοποίησης πρωτεϊνικών δομών και της εύρεσης των κυρίαρχων δομών. Προκειμένου να προσδιοριστεί η συστάδα πρωτεϊνικών δομών που διατηρείται στο σύνολο της τροχιάς, συνεπώς περιέχει και τον μεγαλύτερο αριθμό διαμορφώσεων, κατασκευάστηκε το διαγράμματα αριθμού συστάδας-χρόνου που φαίνεται παρακάτω.

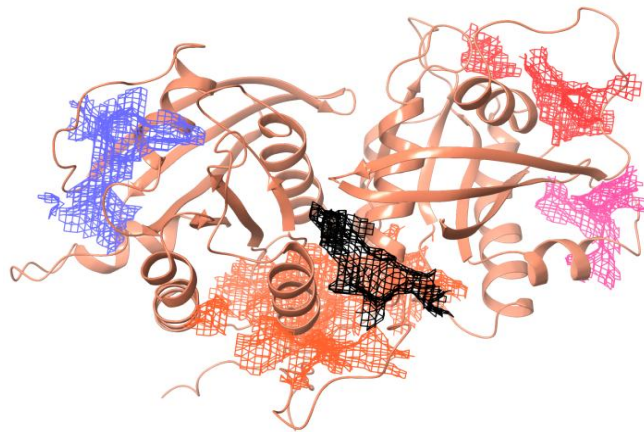


**Διάγραμμα 3-8:** Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το σύστημα διμερούς της μεταλλαγμένης (G12D) KRAS-4B

Από το παραπάνω διάγραμμα, φαίνεται ότι η συστάδα 1 διατηρείται σε όλο το μήκος της τροχιάς του αντιγράφου 1. Συνεπώς, η μέση δομή που αντιστοιχεί σε αυτή την συστάδα επιλέχθηκε ως κυρίαρχη δομή για την τροχιά του μεταλλαγμένου (G12D) KRAS-4B. Ομοίως, επιλέχθηκαν οι κυρίαρχες δομές για τις υπο-τροχιές των 100 ns.

### 3.2.2 Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου SiteMap

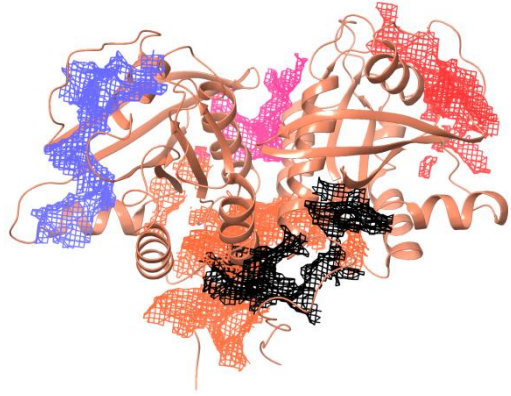
Ο προσδιορισμός των κοιλοτήτων πρόσδεσης πραγματοποιήθηκε στις κυρίαρχες δομές του διμερούς, όπως προέκυψαν από την διαδικασία της ομαδοποίησης. Σε κάθε πρωτεϊνική δομή, προσδιορίστηκαν πέντε (5) κοιλότητες πρόσδεσης. Τα αποτελέσματα για την κυρίαρχη δομή ολόκληρης της τροχιάς φαίνονται παρακάτω.



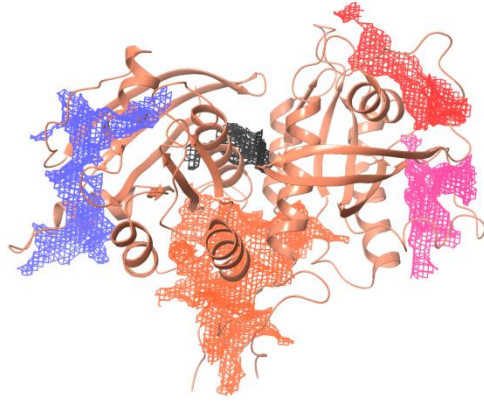
**Εικόνα 3-18:** Αποτελέσματα κοιλοτήτων πρόσδεσης για την κυρίαρχη δομή του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Παρακάτω, φαίνονται τα αποτελέσματα για τις κυρίαρχες δομές των υπο-τροχιών του συστήματος μεταλλαγμένης (G12D) KRAS-4B.

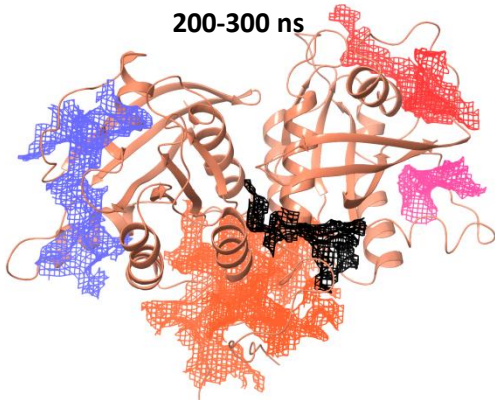
0-100 ns



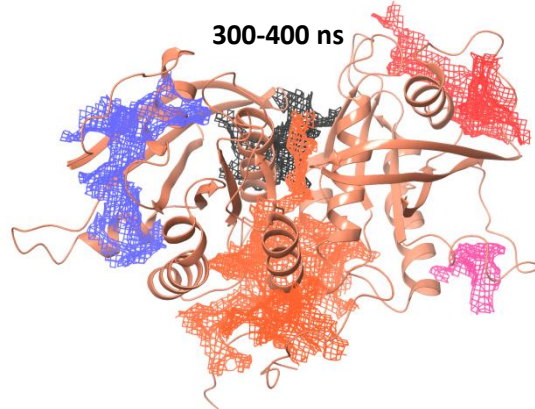
100-200 ns



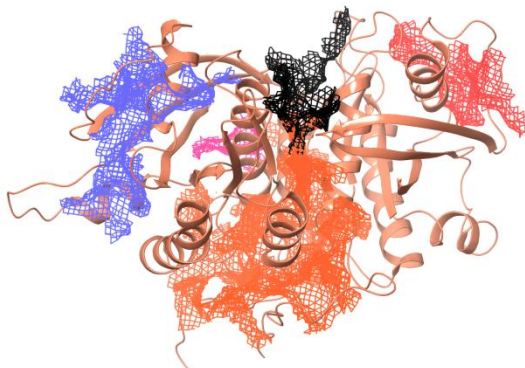
200-300 ns



300-400 ns



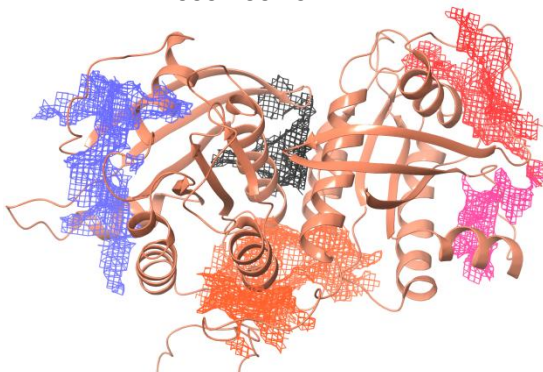
400-500 ns



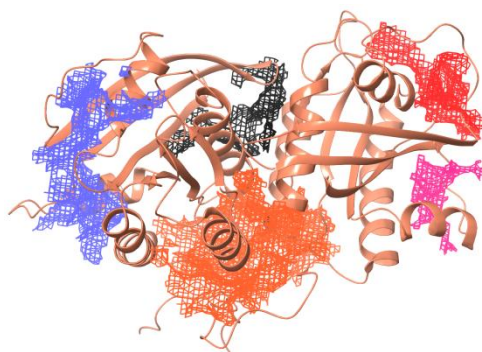
500-600 ns



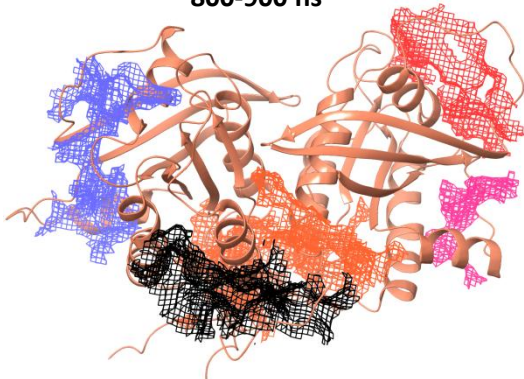
600-700 ns



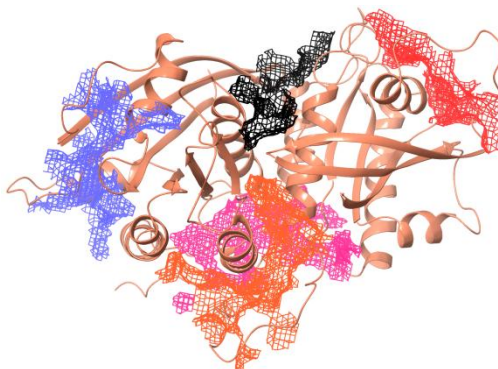
700-800 ns



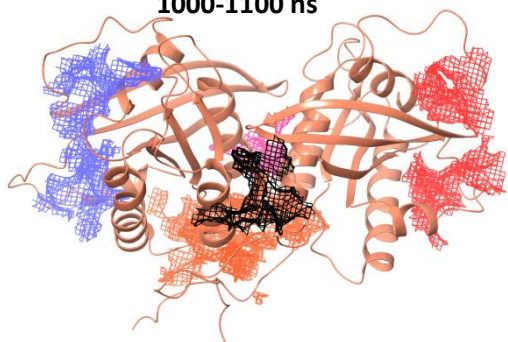
800-900 ns



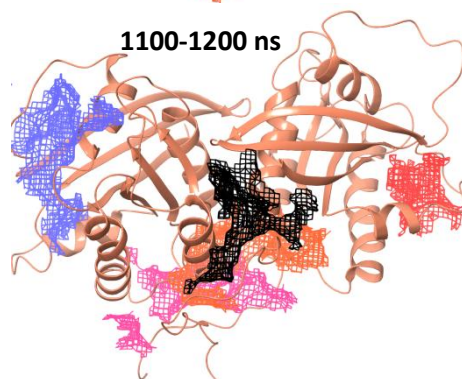
900-1000 ns



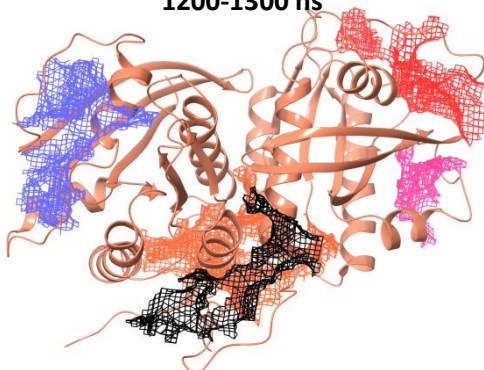
1000-1100 ns



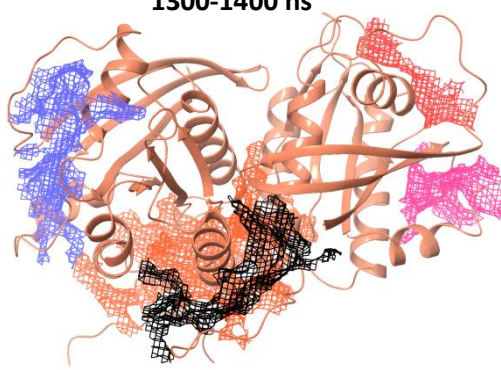
1100-1200 ns



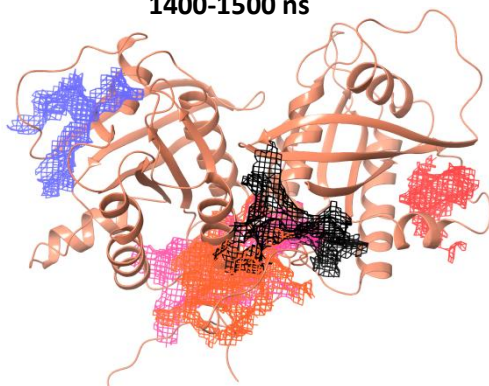
1200-1300 ns



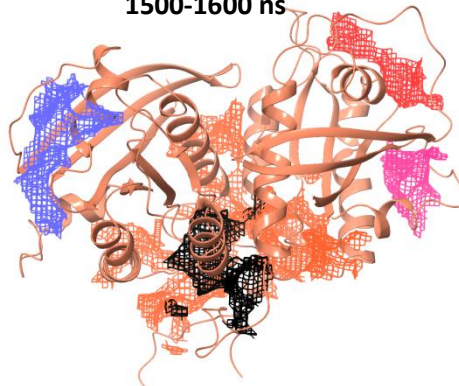
1300-1400 ns

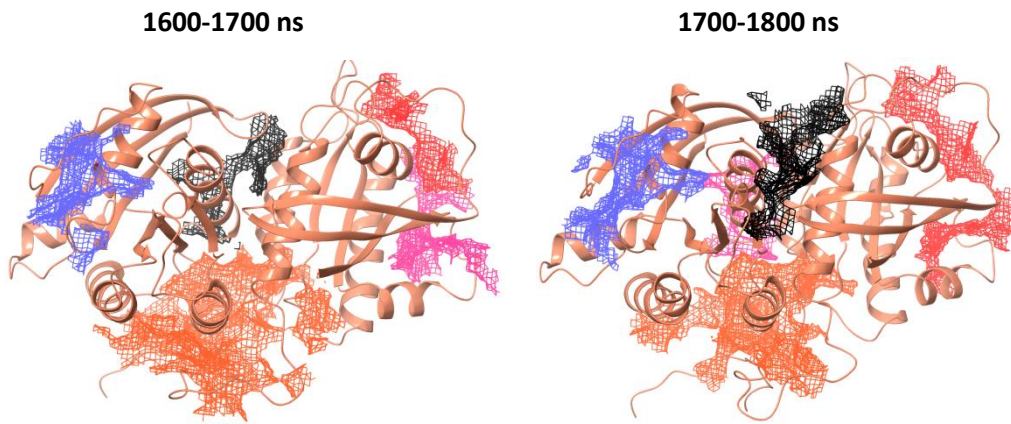


1400-1500 ns



1500-1600 ns





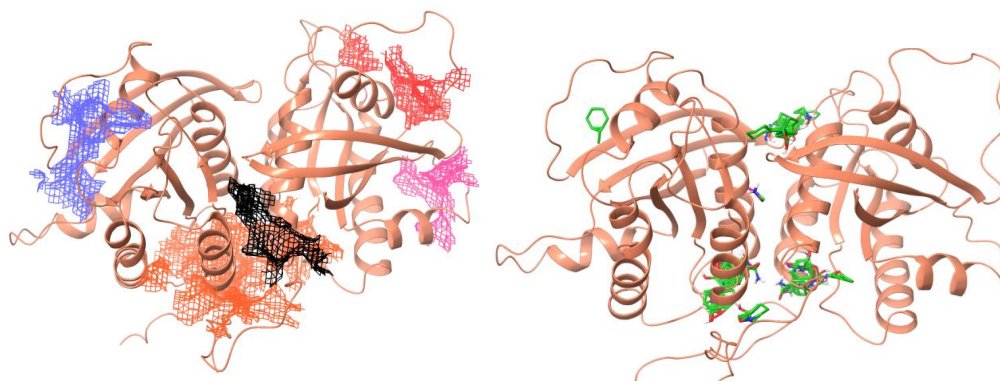
**Εικόνα 3-19:** Αποτελέσματα κοιλοτήτων πρόσδεσης για τις κυρίαρχες δομές των υπο-τροχιών του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Συγκρίνοντας τις παραπάνω δομές, φαίνεται ότι οι κοιλοότητες πρόσδεσης που προκύπτουν στην κυρίαρχη δομή ολόκληρης της τροχιάς είναι πανομοιότυπες με αυτές που προκύπτουν στις μικρότερες υπο-τροχιές των 100 ns. Για το λόγο αυτό, η ανάλυση των αποτελεσμάτων γίνεται με βάση την κυρίαρχη δομή ολόκληρης της τροχιάς.



### 3.2.3 Προσδιορισμός κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου FTMap και σύγκριση των δύο αλγορίθμων

Τα αποτελέσματα εντοπισμού των κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου FTMap παρουσιάζονται παρακάτω, συγκριτικά με τα αποτελέσματα που προέκυψαν μέσω του αλγορίθμου SiteMap. Στα αποτελέσματα που προέκυψαν μέσω του αλγορίθμου FTMap, φαίνονται οι ομάδες των προσδεμάτων, οι οποίες και προσδιορίζουν τις πιθανές κοιλοότητες πρόσδεσης, όπως αναφέρθηκε στην παράγραφο 2.4.

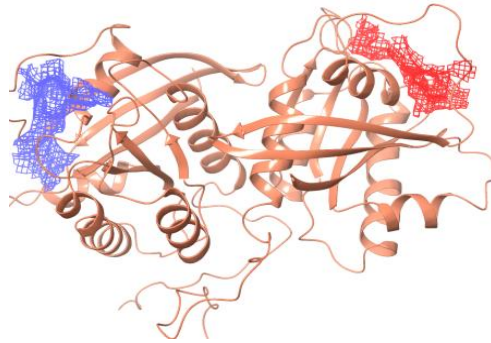


**Εικόνα 3-20:** Σύγκριση αποτελεσμάτων κοιλοτήτων πρόσδεσης μέσω του αλγορίθμου SiteMap (αριστερά) και μέσω του αλγορίθμου FTMap (δεξιά) για την κυρίαρχη δομή του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Οι ομοιότητες μεταξύ των κοιλοτήτων πρόσδεσης που προκύπτουν από τους δύο αλγορίθμους είναι εμφανείς, καθώς οι ομάδες προσδεμάτων του FTMap καταλαμβάνουν τις περιοχές τις πρωτεΐνης που προσδιορίζονται από το SiteMap ή έστω παραπλήσιες περιοχές. Συνεπώς, καταλήγουμε ότι οι δύο αλγόριθμοι SiteMap και FTMap εντοπίζουν τις ίδιες κοιλοότητες πρόσδεσης στις διμερείς πρωτεϊνικές δομές μεταλλαγμένης (G12D) KRAS-4B.

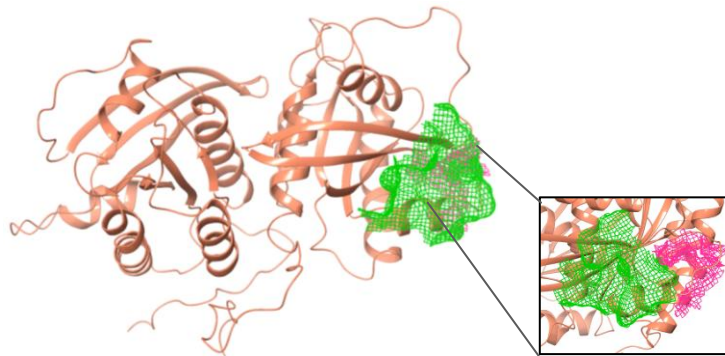
### 3.2.4 Σύγκριση αποτελεσμάτων κοιλοτήτων πρόσδεσης με τη βιβλιογραφία

Αρχικά, οι ορθοστερικές θέσεις πρόσδεσης, οι οποίες εντοπίστηκαν στο διμερές φυσικού τύπου KRAS-4B μέσω του SiteMap, όπως αναφέρεται στην Παράγραφο 3.1.4, εντοπίζονται και στο μεταλλαγμένο διμερές, όπως φαίνεται παρακάτω.



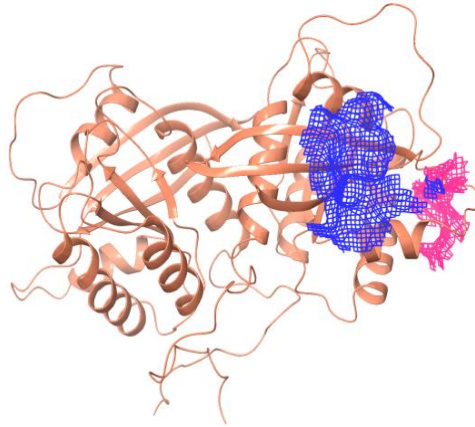
**Εικόνα 3-21:** Ορθοστερικές κοιλοότητες πρόσδεσης προσδιορισμένες στο μεταλλαγμένο (G12D) διμερές KRAS-4B, με χρήση του αλγορίθμου SiteMap. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ) κόκκινου και μπλε χρώματος.

Ακόμα, κοιλοότητα γειτονική με την κοιλοότητα πρόσδεσης G12D SI/II, όπως προσδιορίστηκε από τους Kessler et al[31] αλλά και από τους Sun et al[30] και Maurer et al[29] σε δομές G12D μεταλλαγμένης KRAS-4B, εντοπίζεται στο παραπάνω διμερές μέσω του αλγορίθμου SiteMap.



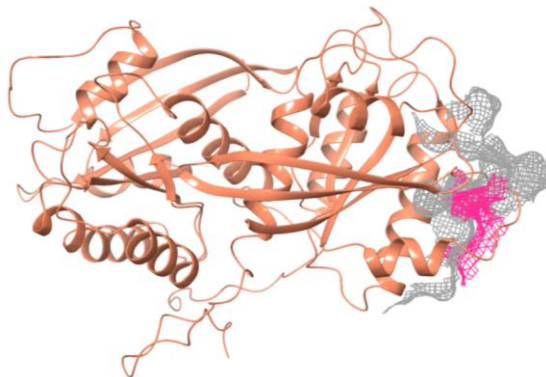
**Εικόνα 3-22:** Σύγκριση της κοιλοότητας G12D SI/II [29] [30] [31] (πράσινο χρώμα) με γειτονική κοιλοότητα στο διμερές G12D KRAS-4B, προσδιορισμένη μέσω του SiteMap (ροζ χρώμα). Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι δύο κοιλοότητες με τη μορφή χάρτη επιφάνειας (δίχτυ).

Η κοιλότητα SI/II έχει προσδιοριστεί και σε διμερές με διαφορετική μετάλλαξη, το Q61H KRAS-4B, από τους Cruz-Migoni et al[32]. Η σύγκριση οδηγεί στο συμπέρασμα ότι και πάλι οι κοιλότητες είναι γειτονικές αλλά διαφορετικές, καθώς χρησιμοποιήθηκαν αρχικές δομές διμερούς KRAS-4B με διαφορετικές μεταλλάξεις.



**Εικόνα 3-23:** Σύγκριση της βιβλιογραφικά προσδιορισμένης κοιλότητας Q61H SI/II [32](μπλε χρώμα) με γειτονική κοιλότητα στο διμερές G12D KRAS-4B (ροζ χρώμα), προσδιορισμένη μέσω του SiteMap. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι δύο κοιλότητες με τη μορφή χάρτη επιφάνειας (δίκτυ).

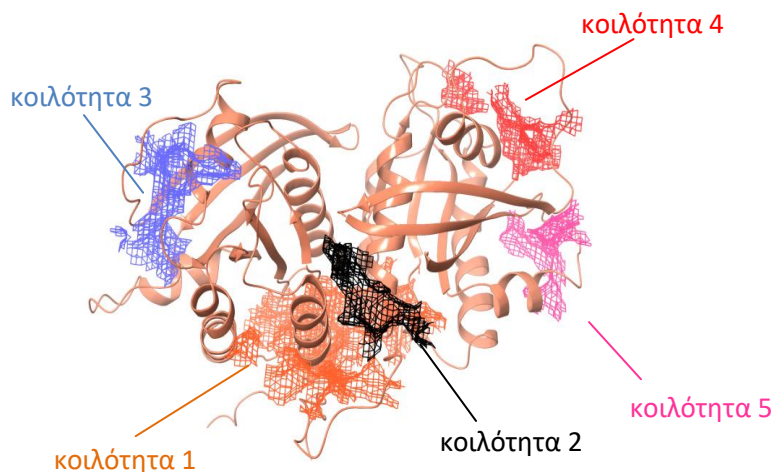
Τέλος, οι κοιλότητες πρόσδεσης του G12D μεταλλαγμένου διμερούς KRAS-4B συγκρίνονται με την G12C S-IIp κοιλότητα πρόσδεσης, όπως προσδιορίστηκε από τους Canon et al[26], Janes et al[27], Patricelli et al[28]. Όπως διακρίνεται και παρακάτω, σε αυτή την περίπτωση οι δύο κοιλότητες αλληλοεπικαλύπτονται.



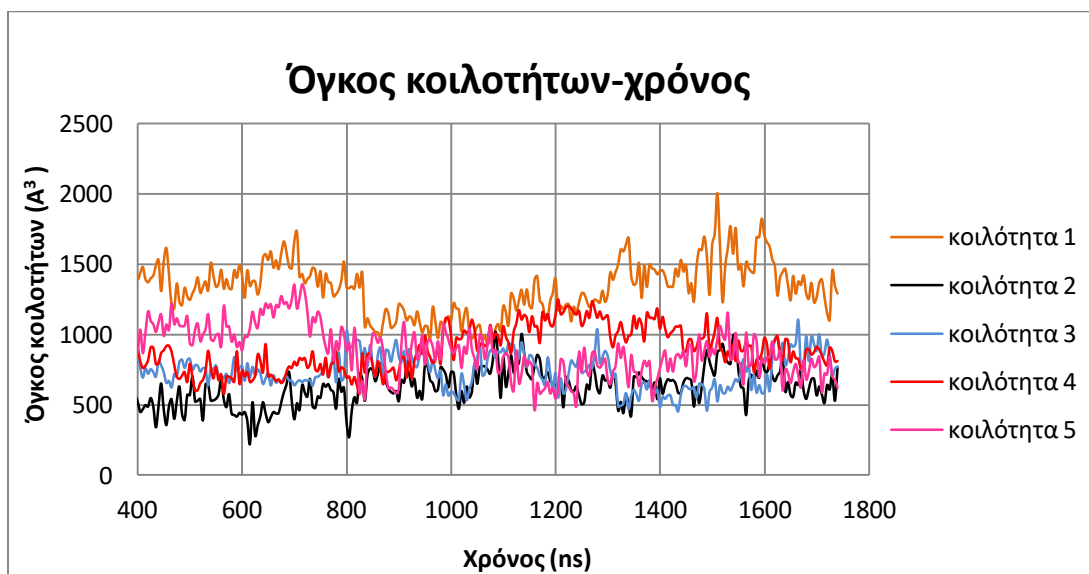
**Εικόνα 3-24:** Σύγκριση της βιβλιογραφικά προσδιορισμένης κοιλότητας G12C SII [26] [27] [28] (γκρι χρώμα) με γειτονική κοιλότητα στο διμερές G12D KRAS-4B (ροζ χρώμα), προσδιορισμένη μέσω του SiteMap. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι δύο κοιλότητες με τη μορφή χάρτη επιφάνειας (δίκτυ).

### 3.2.5 Προσδιορισμός χρονικής μεταβολής των θέσεων πρόσδεσης

Τα αποτελέσματα χρονικής μεταβολής του όγκου των κοιλοτήτων πρόσδεσης για το σύστημα μεταλλαγμένης (G12D) KRAS-4B παρουσιάζονται παρακάτω. Οι κοιλότητες πρόσδεσης έχουν αριθμηθεί, με τρόπο ώστε η κοιλότητα 1 να αντιστοιχεί στη μεγαλύτερη τιμή SiteScore και η κοιλότητα 5 στη μικρότερη τιμή SiteScore.



**Εικόνα 3-25:** Κοιλότητες πρόσδεσης για την κυρίαρχη δομή του συστήματος μεταλλαγμένης (G12D) KRAS-4B. Κοιλότητα 1: μεγαλύτερη τιμή SiteScore, κοιλότητα 5: μικρότερη τιμή SiteScore. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες πρόσδεσης με τη μορφή χάρτη επιφάνειας (δίχτυ).



**Διάγραμμα 3-9:** Όγκος κοιλοτήτων πρόσδεσης σε συνάρτηση με τον χρόνο (μεταλλαγμένη G12D KRAS-4B)

Οι τιμές της τυπικής απόκλισης και του συντελεστή μεταβλητότητας για τις κοιλότητες πρόσδεσης που προσδιορίστηκαν στο μεταλλαγμένο διμερές KRAS-4B παρουσιάζονται στον παρακάτω πίνακα.

**Πίνακας 3-7:** Τιμές τυπικής απόκλισης και συντελεστή μεταβλητότητας για τα δεδομένα όγκου των κοιλοτήτων πρόσδεσης που αντιστοιχούν στο μεταλλαγμένο διμερές KRAS-4B

	Κοιλότητα 1	Κοιλότητα 2	Κοιλότητα 3	Κοιλότητα 4	Κοιλότητα 5
<b>Τυπική απόκλιση (Å)</b>	1329.7 ± 188.3	624.0 ± 132.9	738.5 ± 124.5	893.9 ± 155.4	882.6 ± 177.5
<b>Συντελεστής μεταβλητότητας (Å)</b>	14.2%	21.3%	16.9%	17.4%	20.1%
<b>Κατάλληλη τιμή όγκου</b>	ναι	ναι	ναι	ναι	ναι

Σύμφωνα με τον Πίνακα 3-12, η μικρότερη τιμή συντελεστή μεταβλητότητας αντιστοιχεί στην κοιλότητα 1, η οποία αντιστοιχεί και στη μεγαλύτερη τιμή SiteScore. Ενδιαφέρον προκαλεί ότι η συγκεκριμένη κοιλότητα βρίσκεται στη διεπιφάνεια του διμερούς G12D KRAS-4B, καθώς και ότι η μέση τιμή του όγκου της βρίσκεται εντός του επιθυμητού διαστήματος. Ως εκ τούτου, η κοιλότητα αυτή που διατηρεί τον όγκο της κρίνεται ως υποσχόμενη. Με μικρή διαφορά, ακολουθούν οι κοιλότητες 3 και 4, οι οποίες αντιστοιχούν στις ορθοστερικές θέσεις πρόσδεσης. Οι κοιλότητες 2 και 5 ξεπερνούν ελάχιστα το όριο του συντελεστή μεταβλητότητας (20%), οπότε μπορεί να θεωρηθεί ότι διατηρούνται. Συνολικά, όλες οι κοιλότητες του μεταλλαγμένου διμερούς G12D KRAS-4B διατηρούνται στο χρόνο.

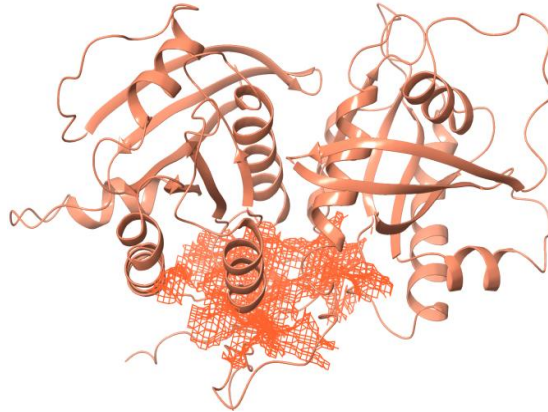
### 3.2.6 Σύγκριση αποτελεσμάτων με το σύστημα φυσικού τύπου

Συγκρίνοντας τις κοιλότητες που προκύπτουν στις δομές διμερών φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B, διαπιστώνουμε ότι σε όλες τις δομές έχει προσδιοριστεί μια κοιλότητα πρόσδεσης στη διεπιφάνεια του διμερούς με τα παρακάτω χαρακτηριστικά:

- Μικρή διακύμανση όγκου
- Κατάλληλη μέση τιμή όγκου (300-1800 Å)
- Κατάλληλη τιμή SiteScore (κοντά στην τιμή 1)

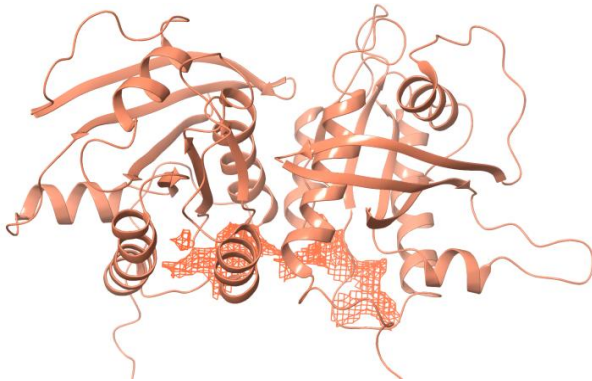
Η κοιλότητα φαίνεται παρακάτω, για τις διάφορες δομές του φυσικού τύπου και μεταλλαγμένης KRAS-4B.

**G12D KRAS-4B**



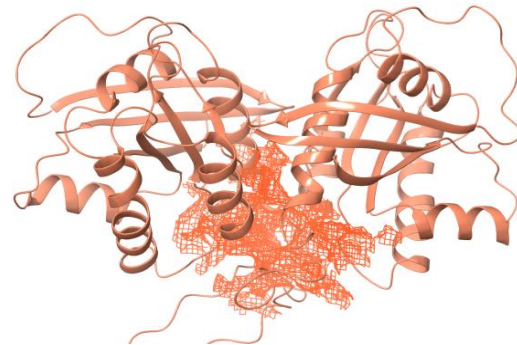
$V (\text{\AA}^3) = 1329.7 \pm 188.3$   
SiteScore = 1.012

**Φυσικού τύπου KRAS-4B  
αντίγραφο 1**



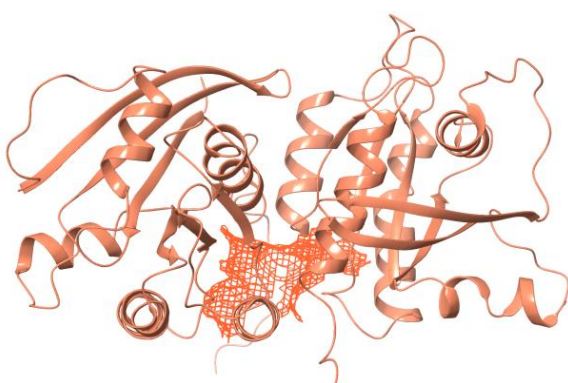
$V (\text{\AA}^3) = 1498.5 \pm 239.9$   
SiteScore = 0.999

**Φυσικού τύπου KRAS-4B  
αντίγραφο 2**



$V (\text{\AA}^3) = 1054.3 \pm 108.2$   
SiteScore = 1.050

**Φυσικού τύπου KRAS-4B  
αντίγραφο 3**



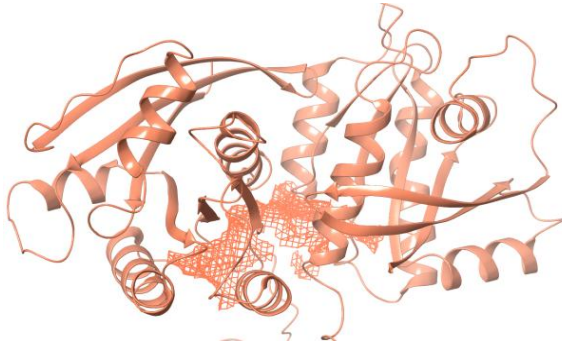
$V (\text{\AA}^3) = 741.6 \pm 344.8$   
SiteScore = 1.045

**Φυσικού τύπου KRAS-4B  
αντίγραφο 4**



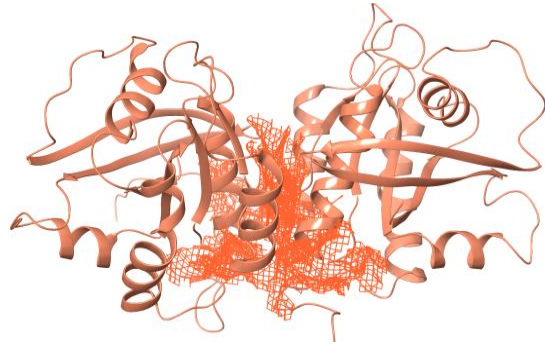
$V (\text{\AA}^3) = 1092.4 \pm 169.0$   
SiteScore = 1.021

Φυσικού τύπου KRAS-4B  
αντίγραφο 5



$V (\text{\AA}^3) = 906.5 \pm 188.0$   
SiteScore = 1.020

Φυσικού τύπου KRAS-4B  
αντίγραφο 6



$V (\text{\AA}^3) = 1201.5 \pm 140.7$   
SiteScore = 0.997

**Εικόνα 3-26:** Σύγκριση των αποτελεσμάτων στα συστήματα φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B. Το διμερές KRAS-4B αναπαρίσταται με τη μορφή κορδελών (ribbons) και οι κοιλότητες πρόσδεσης με τη μορφή χάρτη επιφάνειας (πορτοκαλί δίχτυ).

## 4 ΣΥΜΠΕΡΑΣΜΑΤΑ

Η πρωτεΐνη KRAS-4B αποτελεί στόχο φαρμακευτικών ερευνών, λόγω της μεγάλης συχνότητας των μεταλλάξεων που εμφανίζει και οδηγούν σε ογκογένεση. Η υψηλή συγγένεια της KRAS-4B με τα ενδογενή της προσδέματα GTP και GDP καθιστά δύσκολη τη στόχευση των ορθοστερικών της κοιλοτήτων και κατευθύνει την έρευνα προς την εύρεση αλλοστερικών θέσεων πρόσδεσης στην επιφάνεια της πρωτεΐνης. Η δέσμευση προσδεμάτων στις αλλοστερικές αυτές κοιλοότητες προκαλεί διαμορφωτικές αλλαγές στην επιφάνεια της πρωτεΐνης, επηρεάζοντας έτσι και την πρόσδεση των ενώσεων GTP, GDP. Έως τώρα, οι στρατηγικές στόχευσης έχουν επικεντρωθεί στην διατάραξη των συμπλόκων της KRAS-4B με τις πρωτεΐνες GEFs και τις πρωτεΐνες σηματοδότησης, όπως η Raf, στη βελτίωση του σχηματισμού του συμπλόκου μεταλλαγμένης KRAS-4B τις πρωτεΐνες GAP, καθώς και στην παρεμπόδιση της πρόσδεσης της KRAS-4B στην κυτταρική μεμβράνη, χωρίς ωστόσο ιδιαίτερη επιτυχία.

Σύμφωνα με τη βιβλιογραφία, ο διμερισμός της KRAS-4B παρουσία της πρωτεΐνης Raf στην κυτταρική μεμβράνη ενεργοποιεί μονοπάτια σηματοδότησης που σχετίζονται με την κυτταρική ανάπτυξη και τον κυτταρικό πολλαπλασιασμό, όπως είναι το μονοπάτι Ras-Raf-MEK-Erk. Ως εκ τούτου, μία νέα στρατηγική μπορεί να αποτελέσει η πρόσδεση μικρών μορίων σε αλλοστερικές κοιλοότητες που εντοπίζονται στη διεπιφάνεια του διμερούς KRAS-4B, με σκοπό τη διατάραξη του διμερισμού και ως εκ τούτου την παρεμπόδιση της σηματοδότησης και της ογκογένεσης. Δεδομένου ότι η μετάλλαξη της γλυκίνης σε ασπαρτικό οξύ (G12D) είναι αυτή που εμφανίζεται πιο συχνά στους καρκίνους που σχετίζονται με το ισόμορφο KRAS-4B, η παρούσα εργασία επικεντρώνεται στον προσδιορισμό αλλοστερικών κοιλοτήτων στη διεπιφάνεια του μεταλλαγμένου (G12D) διμερούς KRAS-4B.

Οι προσομοιώσεις Μοριακής Δυναμικής αποτελούν ένα από τα κύρια εργαλεία για τη μελέτη της συμπεριφοράς ενός μοριακού συστήματος συναρτήσει του χρόνου. Μέσω της αναγνώρισης κοιλοτήτων πρόσδεσης σε δομές που προκύπτουν από προσομοιώσεις Μοριακής Δυναμικής, μας δίνεται η δυνατότητα προσδιορισμού της χρονικής μεταβολής των κοιλοτήτων. Στην παρούσα εργασία, πραγματοποιείται αναγνώριση κοιλοτήτων πρόσδεσης σε αντιπροσωπευτικές ή αλλιώς κυρίαρχες δομές διμερούς KRAS-4B που προκύπτουν από την ομαδοποίηση πρωτεϊνικών δομών προερχόμενων από προσομοιώσεις των συστημάτων φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B, σε μοντέλο κυτταρικής μεμβράνης.



Ο προσδιορισμός των κοιλοτήτων πρόσδεσης πραγματοποιείται στις κυρίαρχες δομές του διμερούς KRAS-4B, με τη χρήση των αλγορίθμων SiteMap και FTMap. Μέσω σύγκρισης των αποτελεσμάτων, προκύπτει ότι οι δύο αλγόριθμοι αναγνωρίζουν τις ίδιες κοιλοότητες πρόσδεσης στην επιφάνεια του διμερούς KRAS-4B. Ακόμα, ο αλγόριθμος SiteMap εντοπίζει τις κοιλοότητες που έχουν προηγουμένως προσδιοριστεί βιβλιογραφικά, γεγονός που τον καθιστά έγκυρη μέθοδο για την αναγνώριση κοιλοτήτων. Για παράδειγμα, οι ορθοστερικές κοιλοότητες πρόσδεσης του διμερούς KRAS-4B εντοπίζονται σε όλες τις κυρίαρχες δομές του διμερούς, μέσω του SiteMap. Επιπλέον, συγκρίνοντας τις κοιλοότητες που προκύπτουν στις κυρίαρχες δομές των έξι (6) αντιγράφων φυσικού τύπου, παρατηρείται ότι αυτές διαφοροποιούνται ελάχιστα μεταξύ των αντιγράφων.

Οι κοιλοότητες εξετάζονται επίσης και ως προς τη χρονική τους μεταβολή, προκειμένου να διευκρινιστεί αν διατηρούνται καθ'όλη την κίνηση της KRAS-4B στο χρόνο. Για το σκοπό αυτό, υπολογίζεται η διακύμανση του όγκου τους, σύμφωνα με τις τροχιές της κίνησης του διμερούς KRAS-4B στον χρόνο. Από την ανάλυση αυτή συμπεραίνεται ότι σε γενικές γραμμές οι κοιλοότητες διατηρούνται σε όλα τα συστήματα φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B, εκτός των αντιγράφων τρία (3) και εξί (6) του συστήματος φυσικού τύπου, στα οποία οι πρωτεΐνες Raf έχουν αποσπαστεί από τη μεμβράνη οδηγώντας σε αστάθειες. Επίσης, μια σταθερή κοιλότητα πρόσδεσης εντοπίζεται στη διεπιφάνεια του διμερούς σε όλες τις κυρίαρχες δομές των συστημάτων φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B. Η κοιλότητα αυτή εμφανίζει κατάλληλη μέση τιμή όγκου και κατάλληλες τιμές SiteScore.

Καθώς σκοπός είναι η εκλεκτική στόχευση του διμερισμού της μεταλλαγμένης (G12D) KRAS-4B, αναγκαία είναι η διερεύνηση της εκλεκτικότητας μεταξύ των συστημάτων φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B. Ωστόσο, στα πλαίσια αυτής της εργασίας, δεν μπορούμε να αποφανθούμε σχετικά με την ύπαρξη ή όχι εκλεκτικότητας, καθώς περαιτέρω έρευνα των χαρακτηριστικών των κοιλοτήτων είναι αναγκαία.

## 5 ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ

Αρχικά, χρειάζεται να διερευνηθεί περαιτέρω η εκλεκτικότητα στόχευσης του μεταλλαγμένου (G12D) διμερούς. Για το σκοπό αυτό, κρίνεται σκόπιμη η μελέτη των προσανατολισμών των αμινοξέων που συνιστούν την κοιλότητα πρόσδεσης που εντοπίστηκε στη διεπιφάνεια των διμερών φυσικού τύπου και μεταλλαγμένης (G12D) KRAS-4B. Είναι επίσης απαραίτητη η μελέτη της χρονικής μεταβολής του προσανατολισμού των αμινοξέων, μέσω της κίνησής τους χρόνο, με βάση της τροχιές των συστημάτων.

Τα συστήματα φυσικού τύπου και μεταλλαγμένου (G12D) διμερούς KRAS-4B βρίσκονται υπό διάλυση σε υδάτινο περιβάλλον. Μελετώντας τις θέσεις των μορίων νερού, μπορεί να εξεταστεί κατά πόσο υπάρχουν στο χώρο των κοιλοτήτων. Στην περίπτωση αυτή, οι θέσεις των νερών στις κοιλοότητες μπορούν να καθοδηγήσουν την έρευνα ως προς τις ακριβείς τοποθεσίες στις οποίες μπορούν να προσδεθούν μικρά μόρια, σύμφωνα με το δίκτυο δεσμών υδρογόνου που εντοπίζεται.

Ακολούθως, γνωρίζοντας την επιφάνεια πρόσδεσης στην πρωτεΐνη G12D KRAS-4B, η έρευνα θα μπορούσε να προχωρήσει στο σχεδιασμό φαρμάκων με βάση τη δομή της πρωτεΐνης-στόχου (SBDD). Με τη χρήση κατάλληλης βιβλιοθήκης μορίων και μέσω της τεχνικής προσαρμογής του προσδέματος στο ενεργό κέντρο (docking), μπορούν να βρεθούν μόρια που αλληλεπιδρούν ευνοϊκά με τα αμινοξέα της κοιλότητας πρόσδεσης που θα επιλεγεί για στόχευση. Από τη διαδικασία αυτή, μπορεί να προκύψει μικρός αριθμός ενώσεων που θα προχωρήσουν σε πειραματικές δοκιμές.

Τέλος, για την ανάλυση των αποτελεσμάτων χρησιμοποιήθηκε σε κάθε περίπτωση το χρονικό διάστημα που παρουσίασε σταθερή διακύμανση όγκου, αφήνοντας έξω ένα μικρό κομμάτι στην αρχή της τροχιάς το οποίο είναι ασταθές. Παρότι η θεώρηση αυτή είναι θεμιτή για τον ορθό υπολογισμό της διακύμανσης του όγκου, θα πρέπει να διερευνηθούν εξίσου και τα διαστήματα στα οποία υπάρχουν αστάθειες, καθώς υπάρχει πιθανότητα σχηματισμού κάποιας ασταθούς κοιλότητας που μπορεί να ενδιαφέρει την έρευνα. Για παράδειγμα, αν στο ασταθές διάστημα 0-400 ns της τροχιάς του συστήματος μεταλλαγμένου (G12D) διμερούς εμφανίζεται κάποια κοιλότητα στην διεπιφάνεια, η οποία δεν εμφανίζεται στην υπόλοιπη τροχιά ή στην

τροχιά του συστήματος φυσικού τύπου, η κοιλότητα αυτή αξίζει να εξεταστεί περαιτέρω, καθώς μπορεί με αυτόν τον τρόπο να επιτευχθεί η ζητούμενη εκλεκτικότητα.

## 6 ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Simanshu, D. K., Nissley, D. V., & McCormick, F. (2017). RAS proteins and their regulators in human disease. *Cell*, *170*(1), 17-33.

[2] Prior, I. A., Lewis, P. D., & Mattos, C. (2012). A comprehensive survey of Ras mutations in cancer. *Cancer research*, *72*(10), 2457-2467.

[3] Bos, J. L., Rehmann, H., & Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. *Cell*, *129*(5), 865-877.

[4] Scheffzek, K., Lautwein, A., Kabsch, W., Ahmadian, M. R., & Wittinghofer, A. (1996). Crystal structure of the GTPase-activating domain of human p120GAP and implications for the interaction with Ras. *Nature*, *384*(6609), 591-596.

[5] Schöneborn, H., Raudzus, F., Copepy, M., Neumann, S., & Heumann, R. (2018). Perspectives of RAS and RHEB GTPase signaling pathways in regenerating brain neurons. *International journal of molecular sciences*, *19*(12), 4052.

[6] Ostrem, J. M., & Shokat, K. M. (2016). Direct small-molecule inhibitors of KRAS: from structural insights to mechanism-based design. *Nature reviews Drug discovery*, *15*(11), 771-785.

[7] Prakash, P., & Gorfe, A. A. (2013). Lessons from computer simulations of Ras proteins in solution and in membrane. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1830*(11), 5211-5218.

[8] Lu, S., Banerjee, A., Jang, H., Zhang, J., Gaponenko, V., & Nussinov, R. (2015). GTP binding and oncogenic mutations may attenuate hypervariable region (HVR)-catalytic domain interactions in small GTPase K-Ras4B, exposing the effector binding site. *Journal of Biological Chemistry*, *290*(48), 28887-28900.

- [9] Pantsar, T. (2020). The current understanding of KRAS protein structure and dynamics. *Computational and structural biotechnology journal*, 18, 189-198.
- [10] Muratcioglu, S., Chavan, T. S., Freed, B. C., Jang, H., Khavrutskii, L., Freed, R. N., ... & Nussinov, R. (2015). GTP-dependent K-Ras dimerization. *Structure*, 23(7), 1325-1335.
- [11] Nan, X., Tamgüney, T. M., Collisson, E. A., Lin, L. J., Pitt, C., Galeas, J., ... & Chu, S. (2015). Ras-GTP dimers activate the mitogen-activated protein kinase (MAPK) pathway. *Proceedings of the National Academy of Sciences*, 112(26), 7996-8001.
- [12] Nussinov, R., Jang, H., & Tsai, C. J. (2015). Oligomerization and nanocluster organization render specificity. *Biological Reviews*, 90(2), 587-598.
- [13] Bryant, K. L., Mancias, J. D., Kimmelman, A. C., & Der, C. J. (2014). KRAS: feeding pancreatic cancer proliferation. *Trends in biochemical sciences*, 39(2), 91-100.
- [14] Nussinov, R., Jang, H., Tsai, C. J., & Cheng, F. (2019). Correction: Review: Precision medicine and driver mutations: Computational methods, functional assays and conformational principles for interpreting cancer drivers. *PLoS computational biology*, 15(6), e1007114.
- [15] Δημόπουλος, Β., & Τσαντίλη-Κακουλίδου, Α. (2015). Βασικές αρχές σχεδιασμού και ανάπτυξης φαρμάκων.
- [16] Lionta, E., Spyrou, G., K Vassilatis, D., & Cournia, Z. (2014). Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current topics in medicinal chemistry*, 14(16), 1923-1938.
- [17] Xu, S., Long, B. N., Boris, G. H., Chen, A., Ni, S., & Kennedy, M. A. (2017). Structural insight into the rearrangement of the switch I region in GTP-bound G12A K-Ras. *Acta Crystallographica Section D: Structural Biology*, 73(12), 970-984.
- [18] Zhang, J., & Nussinov, R. (Eds.). (2019). *Protein allostery in drug discovery*. Singapore: Springer.
- [19] Schaber, M. D., O'hara, M. B., Garsky, V. M., Mosser, S. C., Bergstrom, J. D., Moores, S. L., ... & Gibbs, J. B. (1990). Polyisoprenylation of Ras in vitro by a farnesyl-protein transferase. *Journal of Biological Chemistry*, 265(25), 14701-14704.

- [20] Fiordalisi, J. J., Johnson, R. L., Weinbaum, C. A., Sakabe, K., Chen, Z., Casey, P. J., & Cox, A. D. (2003). High affinity for farnesyltransferase and alternative prenylation contribute individually to K-Ras4B resistance to farnesyltransferase inhibitors. *Journal of Biological Chemistry*, 278(43), 41718-41727.
- [21] Guan, K. L., & Stewart, S. (2000). The dominant negative Ras mutant, N17Ras, can inhibit signaling independently of blocking Ras activation. *Journal of Biological Chemistry*, 275(12), 8854-8862.
- [22] Chapman, P. B., Hauschild, A., Robert, C., Haanen, J. B., Ascierto, P., Larkin, J., ... & McArthur, G. A. (2011). Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *New England Journal of Medicine*, 364(26), 2507-2516.
- [23] Nyíri, K., Koppány, G., & Vértessy, B. G. (2020). Structure-based inhibitor design of mutant RAS proteins—a paradigm shift. *Cancer and Metastasis Reviews*, 39(4), 1091-1105.
- [24] FDA grants accelerated approval to sotorasib for KRAS G12C mutated NSC. U.S. Food and Drug Administration. (2021). Retrieved from <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-sotorasib-kras-g12c-mutated-nscl>.
- [25] Sotorasib is First KRAS Inhibitor Approved by FDA. National Cancer Institute. (2021). Retrieved from <https://www.cancer.gov/news-events/cancer-currents-blog/2021/fda-sotorasib-lung-cancer-kras>.
- [26] Canon, J., Rex, K., Saiki, A. Y., Mohr, C., Cooke, K., Bagal, D., ... & Lipford, J. R. (2019). The clinical KRAS (G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature*, 575(7781), 217-223.
- [27] Janes, M. R., Zhang, J., Li, L. S., Hansen, R., Peters, U., Guo, X., ... & Liu, Y. (2018). Targeting KRAS mutant cancers with a covalent G12C-specific inhibitor. *Cell*, 172(3), 578-589.
- [28] Patricelli, M. P., Janes, M. R., Li, L. S., Hansen, R., Peters, U., Kessler, L. V., ... & Liu, Y. (2016). Selective inhibition of oncogenic KRAS output with small molecules targeting the inactive state. *Cancer discovery*, 6(3), 316-329.
- [29] Maurer, T., Garrenton, L. S., Oh, A., Pitts, K., Anderson, D. J., Skelton, N. J., ... & Fang, G. (2012). Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated

nucleotide exchange activity. *Proceedings of the National Academy of Sciences*, 109(14), 5299-5304.

[30] Sun, Q., Burke, J. P., Phan, J., Burns, M. C., Olejniczak, E. T., Waterson, A. G., ... & Fesik, S. W. (2012). Discovery of small molecules that bind to K-Ras and inhibit Sos-mediated activation. *Angewandte Chemie International Edition*, 51(25), 6140-6143.

[31] Kessler, D., Gmachl, M., Mantoulidis, A., Martin, L. J., Zoephel, A., Mayer, M., ... & McConnell, D. B. (2019). Drugging an undruggable pocket on KRAS. *Proceedings of the National Academy of Sciences*, 116(32), 15823-15829.

[32] Cruz-Migoni, A., Canning, P., Quevedo, C. E., Bataille, C. J., Bery, N., Miller, A., ... & Rabbitts, T. H. (2019). 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*, 116(7), 2545-2550.

[33] Ζουμπουλάκης, Π., & Μαυρομούστακος, Θ. (2008). *Μοριακή Μοντελοποίηση: Εφαρμογές στην Οργανική και Φαρμακευτική Χημεία*. Ιατρικές Εκδόσεις Γιάννης Β. Παρισιάνος.

[34] Verlet, L. (1967). Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical review*, 159(1), 98.

[35] Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design*, 27(3), 221-234.

[36] Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2021; Impact, Schrödinger, LLC, New York, NY; Prime, Schrödinger, LLC, New York, NY, 2021

[37] Cournia, Z. (2006). *Structure and Dynamics of Biomembranes containing Cholesterol and other Biologically-Important Sterols: a computational perspective* (Doctoral dissertation).

[38] *gmx cluster — GROMACS 2021.3 documentation*. Manual.gromacs.org. (2021). Retrieved from <https://manual.gromacs.org/documentation/current/onlinehelp/gmx-cluster.html>.

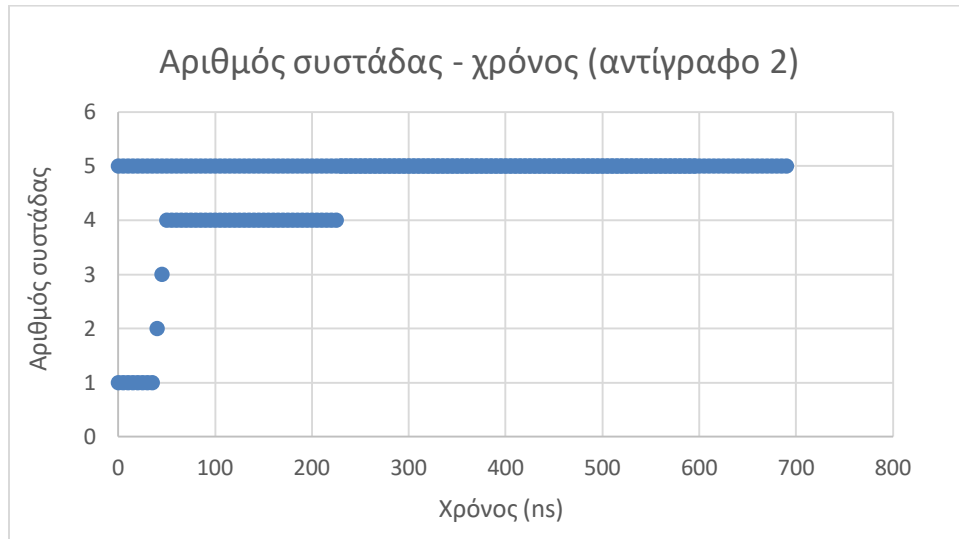
[39] Halgren, T. A. (2009). Identifying and characterizing binding sites and assessing druggability. *Journal of chemical information and modeling*, 49(2), 377-389.

- [40] Kozakov, D., Grove, L. E., Hall, D. R., Bohnuud, T., Mottarella, S. E., Luo, L., ... & Vajda, S. (2015). The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nature protocols*, *10*(5), 733-755.
- [41] Sielecki, A. R., Hayakawa, K., Fujinaga, M., Murphy, M. E., Fraser, M., Muir, A. K., ... & James, M. N. (1989). Structure of recombinant human renin, a target for cardiovascular-active drugs, at 2.5 Å resolution. *Science*, *243*(4896), 1346-1351.
- [42] *Epock's documentation*. Epock.bitbucket.io. (2021). Retrieved from <https://epock.bitbucket.io/docs/index.html>.
- [43] *gmx trjconv — GROMACS 5.1 documentation*. Manual.gromacs.org. (2015). Retrieved from <https://manual.gromacs.org/documentation/5.1/onlinehelp/gmx-trjconv.html>.
- [44] Everitt, B. S., & Skrondal, A. (2010). *The Cambridge dictionary of statistics*.
- [45] Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., & Rarey, M. (2012). Combining global and local measures for structure-based druggability predictions. *Journal of chemical information and modeling*, *52*(2), 360-372.
- [46] Halgren, T. (2007). New method for fast and accurate binding-site identification and analysis. *Chemical biology & drug design*, *69*(2), 146-148.
- [47] Laurent, B., Chavent, M., Cragolini, T., Dahl, A. C. E., Pasquali, S., Derreumaux, P., ... & Baaden, M. (2015). Epock: rapid analysis of protein pocket dynamics. *Bioinformatics*, *31*(9), 1478-1480.
- [48] Mazhab-Jafari, M. T., Marshall, C. B., Smith, M. J., Gasmi-Seabrook, G. M., Stathopoulos, P. B., Inagaki, F., ... & Ikura, M. (2015). Oncogenic and RASopathy-associated K-RAS mutations relieve membrane-dependent occlusion of the effector-binding site. *Proceedings of the National Academy of Sciences*, *112*(21), 6625-6630.
- [49] Fetics, S. K., Guterres, H., Kearney, B. M., Buhrman, G., Ma, B., Nussinov, R., & Mattos, C. (2015). Allosteric effects of the oncogenic RasQ61L mutant on Raf-RBD. *Structure*, *23*(3), 505-516.

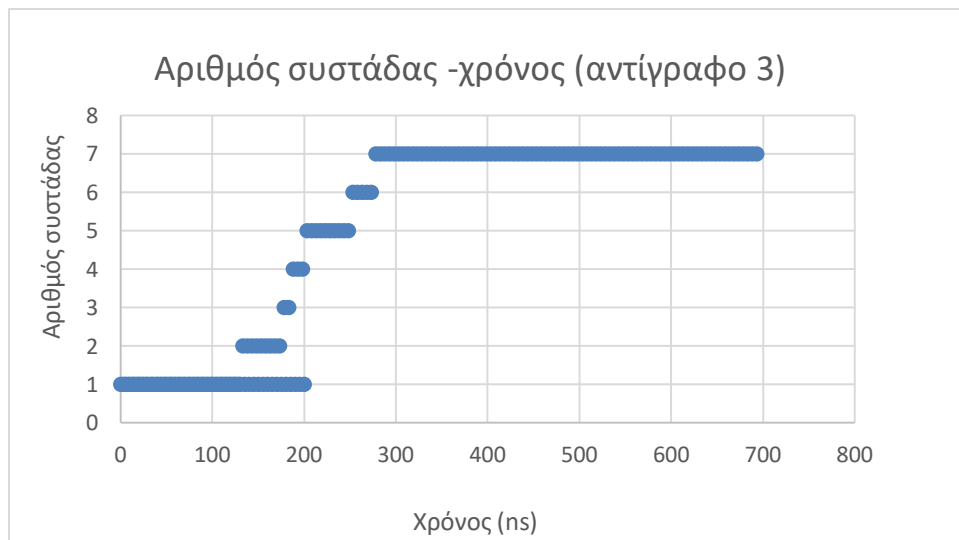
[50] Mott, H. R., Carpenter, J. W., Zhong, S., Ghosh, S., Bell, R. M., & Campbell, S. L. (1996). The solution structure of the Raf-1 cysteine-rich domain: a novel ras and phospholipid binding site. *Proceedings of the National Academy of Sciences*, 93(16), 8312-8317.



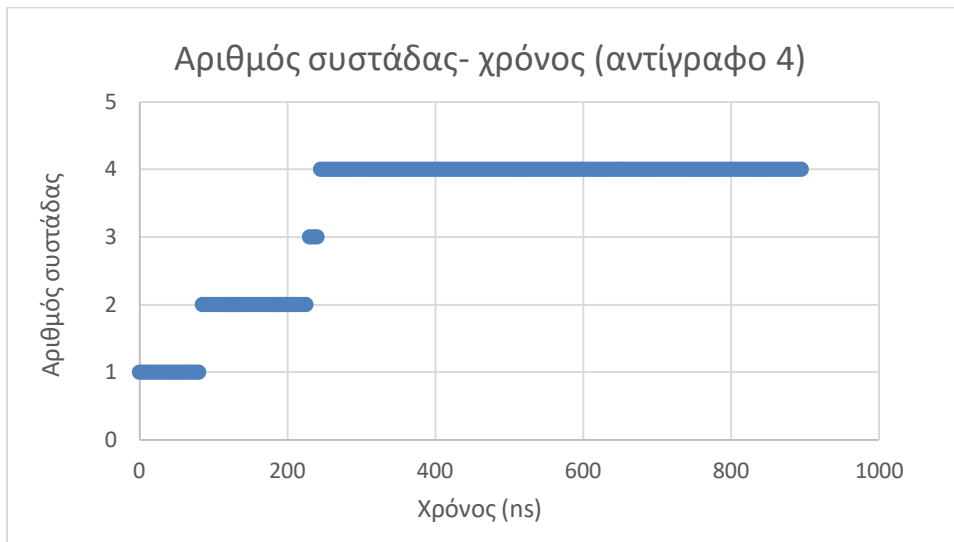
## ΠΑΡΑΡΤΗΜΑ



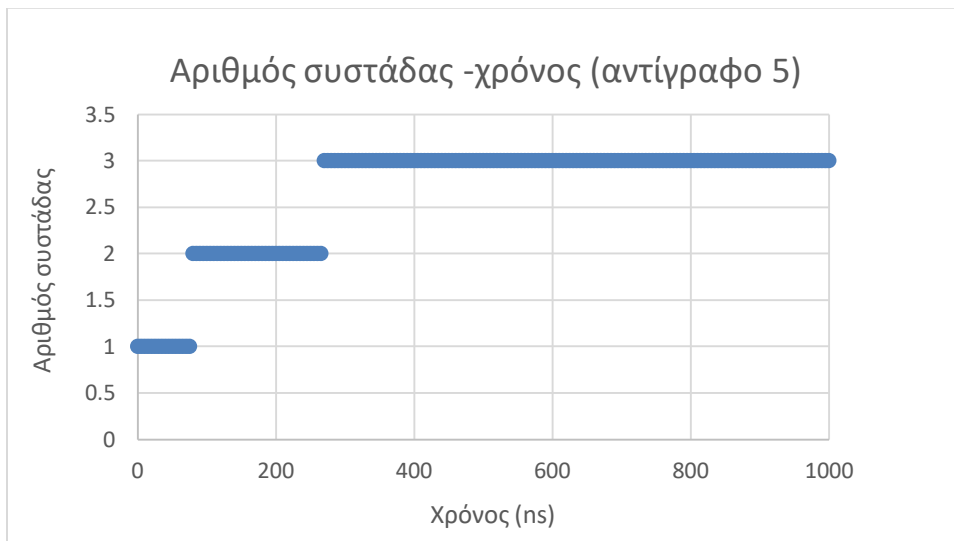
**Διάγραμμα 6-1:** Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 2 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 5 διατηρείται σε όλο το μήκος της τροχιάς.



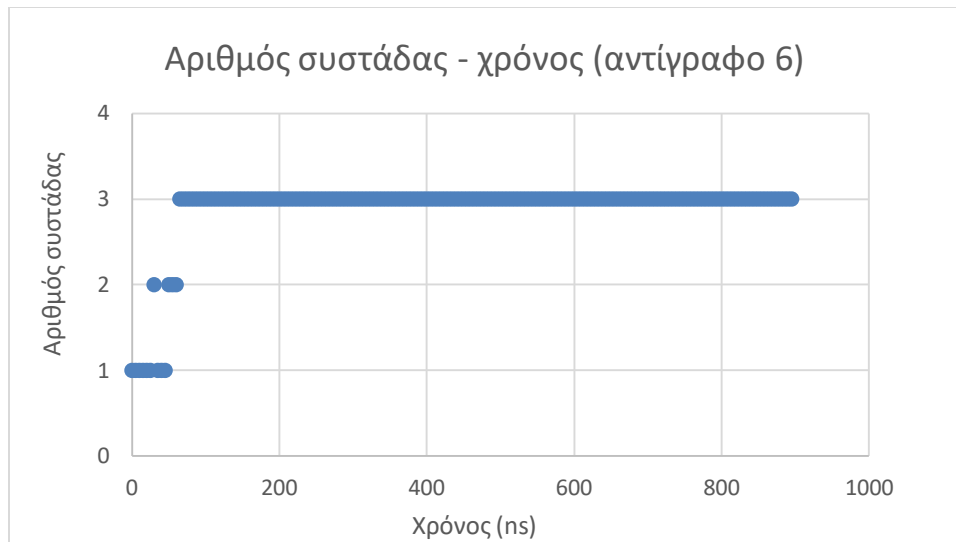
**Διάγραμμα 6-2:** Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 3 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 7 διατηρείται στο μεγαλύτερο μέρος της τροχιάς.



**Διάγραμμα 6-3:** Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 4 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 4 διατηρείται στο μεγαλύτερο μέρος της τροχιάς.



**Διάγραμμα 6-4:** Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 5 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 3 διατηρείται στο μεγαλύτερο μέρος της τροχιάς.



**Διάγραμμα 6-5:** Αριθμός συστάδας πρωτεϊνικών δομών σε συνάρτηση με τον χρόνο για το αντίγραφο 6 του συστήματος φυσικού τύπου KRAS-4B. Η συστάδα 3 διατηρείται σε όλο το μήκος της τροχιάς.