

Παπαϊωάννου Ι. Χαρίκλεια
ΑΕΜ : 113

Διπλωματική Εργασία

*“Μελέτες Μοριακής Δυναμικής μιας
Μορφής της ROP με Επανασχεδιασμένο
Υδρόφοβο Πυρήνα, Ala₂Ile₂-6”*

Επιβλέπων Καθηγητής : Νικόλαος Μ. Γλυκός









*Δημοκρίτειο Πανεπιστήμιο Θράκης
Τμήμα Μοριακής Βιολογίας & Γενετικής
Σεπτέμβριος 2007*



Θα ήθελα να ευχαριστήσω τον καθηγητή μου, κ. Νικόλαο Μ. Γλυκό, για την πολύτιμη βοήθειά του. Η κατανόηση που έδειξε, η επιμονή, αλλά κυρίως η υπομονή του, με δίδαξαν πολύ περισσότερα απ' όσα θα μπορούσαν να χωρέσουν στις σελίδες αυτής της διπλωματικής εργασίας.










Επίσης θα ήθελα να ευχαριστήσω τους συναδέλφους μου, κ. Γεωργούλια Παναγιώτα, κ. Δαρμανή Σπυρίδωνα, κ. Πάλιτοκάρη Δημήτριο και κ. Πουγούνια Αικατερίνη, για το ευχάριστο περιβάλλον εργασίας που βοήθησαν να δημιουργήσουμε, κατά τη διάρκεια της παραμονής μας στο εργαστήριο.

Στους γονείς μου Γιάννη και Ασημούλα, καθώς και στην αδερφή μου Βιβή. Η βοήθεια και η συμπαράστασή σας ήταν, είναι και θα είναι πολύ σημαντικές για μένα. Οι αξίες, οι συμβουλές, αλλά κυρίως η αγάπη που μου δώσατε, συνέβαλαν στο να γίνω ο άνθρωπος που είμαι σήμερα. Το λιγότερο που μπορώ να κάνω είναι να σας αφιερώσω τη διπλωματική μου εργασία. Σας ευχαριστώ για όλα!!!

Περιεχόμενα

Περίληψη	7
Κεφάλαιο 1	
Εισαγωγή	9
Κεφάλαιο 2	
Προσομοιώσεις Μοριακής Δυναμικής	20
 2.1 Προσομοιώσεις μοριακής δυναμικής της Ala_2Ile_2-6	25
Κεφάλαιο 3	
Προετοιμασία Συστημάτων	27
 3.1 Εισαγωγή	28
 3.2 Γενική προετοιμασία	35
 3.3 Προετοιμασία διμερούς της Ala_2Ile_2-6 με τοπολογία <i>syn</i>	37
 3.4 Προετοιμασία διμερούς της Ala_2Ile_2-6 με τοπολογία <i>anti</i>	44
 3.5 Προετοιμασία <i>native</i> μονομερούς της Ala_2Ile_2-6	47
 3.6 Προετοιμασία <i>artificial</i> μονομερούς της Ala_2Ile_2-6	49
 3.7 Προετοιμασία μονομερούς της Ala_2Ile_2-6 με αναδίπλωση της $A31P$	57

Κεφάλαιο 4	
Ανάλυση Προσομοιώσεων	60
 4.1 Ανάλυση προσομοίωσης διμερούς της Ala_2Ile_2-6 με τοπολογία <i>syn</i>	65
4.1.1 $RMSD$ (τυπική απόκλιση) από την αρχική δομή	68
4.1.1.1 Τι θα συμβεί αν “κοπούν οι ουρές;”	72
4.1.1.2 Διαφορές ανάμεσα στην Ala_2Ile_2-6 με τοπολογία <i>syn</i> με τα υπερκινητικά άκρα και την Ala_2Ile_2-6 με τοπολογία <i>syn</i> χωρίς αυτά	77
4.1.2 Η μέση δομή και οι αντίστοιχες διακυμάνσεις	79
4.1.2.1 $RMSD$ από τη μέση δομή	80
4.1.3 $RMSF$	81
4.1.4 Εξέλιξη του R_G , <i>Radius of Gyration</i>	83
4.1.5 Χάρτης μέσης απόστασης μεταξύ των $C\alpha$ ατόμων και το αντίστοιχο $RMSD$	84
4.1.6 <i>Frame-to-Frame RMSD</i>	85
4.1.7 <i>Variance-Covariance</i> και <i>Cross-Correlation</i>	85
4.1.8 Χαρακτηριστικές τιμές (<i>eigenvalues</i>) και χαρακτηριστικά ανύσματα (<i>eigenvectors</i>)	87
4.1.9 Ομαδοποίηση των δομών της προσομοίωσης (<i>cluster analysis</i>)	90
4.1.10 Σύγκριση μεταξύ μέσης δομής των 2 <i>clusters</i> και της αρχικής δομής	93
4.1.11 Ανάλυση δευτεροταγούς δομής	95
 4.2 Ανάλυση προσομοίωσης διμερούς της Ala_2Ile_2-6 με τοπολογία <i>anti</i> (<i>Rop-like</i>)	96
4.2.1 $RMSD$ (τυπική απόκλιση) από την αρχική δομή	97
4.2.2 Η μέση δομή και οι αντίστοιχες διακυμάνσεις	99
4.2.2.1 $RMSD$ από τη μέση δομή	99
4.2.3 $RMSF$	100
4.2.4 Εξέλιξη του R_G , <i>Radius of Gyration</i>	101
4.2.5 Χάρτης μέσης απόστασης μεταξύ των $C\alpha$ ατόμων και το αντίστοιχο $RMSD$	102
4.2.6 <i>Frame-to-Frame RMSD</i>	103

4.2.7	<i>Variance-Covariance και Cross-Correlation</i>	104
4.2.8	Χαρακτηριστικές τιμές (<i>eigenvalues</i>) και χαρακτηριστικά ανύσματα (<i>eigenvectors</i>)	105
4.2.9	Ομαδοποίηση των δομών της προσομοίωσης (<i>cluster analysis</i>)	108
4.2.10	Σύγκριση μεταξύ μέσης δομής των 2 <i>clusters</i> και της αρχικής δομής	111
4.2.11	Ανάλυση δευτεροταγούς δομής	113
	4.3 Σύγκριση αποτελεσμάτων ανάλυσης μεταξύ <i>Ala₂Ile₂₋₆</i> με <i>syn</i> τοπολογία και <i>Ala₂Ile₂₋₆</i> με <i>anti</i> (<i>Rop-like</i>) τοπολογία	114
	4.4 Ανάλυση προσομοίωσης <i>native</i> μονομερούς	121
	4.5 Ανάλυση προσομοίωσης <i>artificial</i> μονομερούς	124
	4.6 Ανάλυση προσομοίωσης <i>opened</i> (<i>A31P</i>) μονομερούς	125
	Επίλογος	131
	Βιβλιογραφία	132
	Παραρτήματα	141
	3-1 : <i>Heating up Script</i>	142
	3-2 : <i>Equilibration Script</i>	146
	4-1 : <i>Calculation of RMSD from starting structure (Xplor Script)</i>	148
	4-2 : <i>Calculation of RMSD from average structure (Xplor Script)</i>	150
	4-3 : <i>Calculation of Radius of Gyration R_G (Xplor Script)</i>	153

ΠΕΡΙΛΗΨΗ

Το πρόβλημα της αναδίπλωσης των πρωτεϊνών απασχολεί τους επιστήμονες εδώ και χρόνια και αποτελεί ένα σημαντικό πεδίο έρευνας. Μέχρι σήμερα έχουν αναπτυχθεί πολλές μέθοδοι για να μπορέσουμε να προσεγγίσουμε το πρόβλημα αυτό. Σαν μοντέλο στην έρευνα αυτή επιλέχθηκε η Rop, μια μικρή ομοδιμερής πρωτεΐνη που σχηματίζει ένα 4-α-ελικοειδές δεμάτιο. Εμπλέκεται στον έλεγχο των αντιγράφων του πλασμιδίου ColE1 της *Escherichia coli*, καθώς προσδέεται στο σύμπλοκο των μορίων RNAI-RNAII. Σαν σύστημα για τις μελέτες που μας αφορούν παρουσιάζει πολλά πλεονεκτήματα δομικά και λειτουργικά. Αξιοσημείωτο είναι το γεγονός ότι μπορεί να χρησιμοποιηθεί για τη δημιουργία παραλλαγών της, μια από τις οποίες είναι και η Al₂He₂-6 (Willis *et al.*, 2000), η παραλλαγή την οποία σκοπεύουμε να μελετήσουμε. Πρόκειται για μια παραλλαγή όπου ο υδρόφοβος πυρήνας έχει επανασχεδιαστεί ώστε τα έξι μεσαία στρώματά (layers) του να αποτελούνται μόνο από κατάλοιπα αλανίνης και ισολευκίνης. Η πρωτεΐνη αυτή έχει μελετηθεί διεξοδικά όσο αφορά τις κινητικές και θερμοδυναμικές ιδιότητές της (Willis *et al.*, 2000). Παρουσιάζει όμως μια ιδιαιτερότητα όσο αφορά τη δομή της. Παρατηρείται μια στροφή του ενός μονομερούς κατά 180° γύρω από την κοινή επιφάνεια των μονομερών. Η διαφορετική αυτή διαμόρφωση και η απώλεια της λειτουργικότητας της φυσικής πρωτεΐνης είναι κάποια από τα ερεθίσματα που μας ωθούν στο να ασχοληθούμε περαιτέρω με τη μελέτη της αναδίπλωσης της πρωτεΐνης μέσω προσομοιώσεων μοριακής δυναμικής. Οι προσομοιώσεις μοριακής δυναμικής αποτελούν ένα ισχυρό εργαλείο, που μπορεί να δώσει πολλές πληροφορίες σε ποικίλα επίπεδα ανάλυσης. Η προσομοίωση έχει αποδειχθεί χρήσιμη στην κατανόηση των μηχανισμών της πρωτεϊνικής αναδίπλωσης (Duan *et al.*, 1998). Για το λόγο αυτό χρησιμοποιούνται εδώ και χρόνια και θεωρούνται αρκετά αξιόπιστες (Karplus, 2003; Karplus *et al.*, 2002).

Εξαιτίας της διαφοράς στη δομή και τις ιδιότητες μεταξύ της Al₂He₂-6 και της φυσικής Rop, προτάθηκε μια υπόθεση σχετικά με το ποια δομή είναι τελικά πιο σταθερή (Willis *et al.*, 2000). Πραγματοποιήσαμε λοιπόν προσομοιώσεις για τα συστήματα δυο διμερών με διαφορετικές τοπολογίες (ένα με την τοπολογία της Al₂He₂-6, syn και ένα με την αλληλουχία της Al₂He₂-6 και την τοπολογία της φυσικής Rop, anti). Από τα αποτελέσματα δεν προκύπτει κάποιο στοιχείο που να εξηγεί το λόγο για τον οποίο η Al₂He₂-6 προτιμά τη διαμόρφωση syn. Παρόλα αυτά στόχος μας είναι να μπορέσουμε να παρακολουθήσουμε ένα γεγονός αναδίπλωσης. Έτσι χρησιμοποιούμε απομονωμένα μονομερή, με διαφορετικές αρχικές διαμορφώσεις, με σκοπό την πραγματοποίηση υπολογιστικών διμοριακών συγκρούσεων (bimolecular collisions). Εντούτοις, όντας χρονικά περιορισμένοι, στην εργασία αυτή θα ασχοληθούμε με τη μελέτη ορισμένων συστημάτων διμερών και μονομερών μορίων.

ABSTRACT

The protein folding problem has concerned scientists for many years and represents a very important field of research. Several methods have been developed in order to approach this problem. As a model in this research we have chosen Rop, a small homodimeric protein that forms a 4- α -helix bundle. Rop is involved in the control process of the copy number of *Escherichia coli*'s ColE1 plasmid, as it binds the RNAI-RNAPII complex. This protein shows a lot of advantages structural and functional. It has been used as a model system for designing several variants, such as Ala₂Ile₂-6 (Willis *et al.*, 2000), the variant we are about to study. Its hydrophobic core is redesigned so that the middle six layers consist only of alanine and isoleucine residues. This protein has been fully studied as far as it concerns its kinetic and thermodynamic properties (Willis *et al.*, 2000). Ala₂Ile₂-6 has a very interesting feature as far as it concerns its structure. A structural flip of one of the two monomers about 180° around the common surface of the two monomers is observed. This new conformation and the loss of the wild type protein's function are a few of the things that motivate us to study further the protein's folding process through molecular dynamics simulations. Molecular dynamics simulations have been a powerful tool that can provide rich information at various levels of resolution. Simulation has been instrumental in understanding protein folding mechanisms (Duan *et al.*, 1998). These are a few reasons why they are being used for years and they are considered very reliable (Karplus, 2003; Karplus *et al.*, 2002).

Due to the difference in the structure and the properties between Ala₂Ile₂-6 and wild type Rop, a suggestion has been made concerning their stabilities (Willis *et al.*, 2000). We have performed MD simulations for the systems of two dimers with different topologies (one with the syn topology of Ala₂Ile₂-6 and one with the anti topology of the wild type Rop). The results cannot explain why Ala₂Ile₂-6 prefers the syn conformation. Nevertheless our intention is to observe, if possible, a folding event, through bimolecular collisions of monomers, starting from different conformations. However, due to the lack of time, we will perform simulations of two dimers and a few monomers with different topologies.



Κεφάλαιο 1

Εισαγωγή

Η ικανότητα μιας πρωτεΐνης να αναδιπλώνεται αξιόπιστα σε μια προκαθορισμένη διαμόρφωση, παρά τον σχεδόν απεριόριστο αριθμό διαμορφώσεων, δεν έχει γίνει ακόμα κατανοητή. Είναι γνωστό, εδώ και πολύ καιρό, πως η δομή μιας πρωτεΐνης καθορίζεται αποκλειστικά από την αμινοξική αλληλουχία (Anfinsen, 1973) και η δομή μιας πρωτεΐνης καθορίζει τη λειτουργία της. Κατ' επέκταση, η λειτουργία μιας πρωτεΐνης εξαρτάται αποκλειστικά από την ικανότητα της πρωτεΐνης να αναδιπλώνεται γρήγορα και με αξιοπιστία στην φυσική της δομή.

Σε πολλές πρωτεΐνες παρατηρείται το φαινόμενο της αυθόρμητης αναδίπλωσης στη φυσική τους δομή, όταν βρεθούν μέσα σε υδατικό διάλυμα. Η διαδικασία αυτή της αναδίπλωσης πρέπει να ικανοποιεί δυο συνθήκες, μια θερμοδυναμική και μια κινητική. Από θερμοδυναμικής άποψης, η πρωτεΐνη αποκτά μια μοναδική, σταθερή και αναδιπλωμένη διαμόρφωση. Όσο για την κινητική αξίωση, η πρωτεΐνη πρέπει να αναδιπλώνεται στη φυσική της δομή σε προκαθορισμένη χρονική κλίμακα. Υπάρχει μια υπόθεση σύμφωνα με την οποία για μια πρωτεΐνη με 100 αμινοξέα, μια καθαρά τυχαία έρευνα διαμόρφωσης θα απαιτούσε ένα χρονικό διάστημα από κάποια δευτερόλεπτα έως κάποια χρόνια (Dobson *et al.*, 1998). Παρόλα αυτά οι πρωτεΐνες είναι ικανές να αναδιπλώνονται σε ένα χρονικό διάστημα της τάξεως των χιλιοστών του δευτερολέπτου (ms). Έχουν προταθεί πολλά θεωρητικά μοντέλα αναδίπλωσης, με σκοπό να εξηγήσουν τον τρόπο με τον οποίο πραγματοποιείται με τέτοιες ταχύτητες η αξιόπιστη αναδίπλωση των πρωτεϊνών (γνωστό και ως πρόβλημα της αναδίπλωσης των πρωτεϊνών, protein folding problem) (Wetlaufer, 1973; Kim *et al.*, 1990; Bashford *et al.*, 1988; Matheson *et al.*, 1978; Harrison *et al.*, 1985). Το ερώτημα αυτό είναι γνωστό ως παράδοξο του Levinthal και έχει επικρατήσει στις συζητήσεις για την αναδίπλωση πάνω από 30 χρόνια (Levinthal, 1968).

Από τη στιγμή που υπάρχει ένας μεγάλος αριθμός προσιτών διαμορφώσεων για μια μη αναδιπλωμένη πολυπεπτιδική αλυσίδα, δεν είναι πρακτικά δυνατό για μια πρωτεΐνη να περάσει από όλες αυτές τις διαμορφώσεις ώστε να λάβει τελικά την κατάλληλη δομή. Το γεγονός αυτό υποδηλώνει πως ελέγχεται μόνο ένα μικρό μέρος του χώρου διαμόρφωσης των πρωτεϊνών κατά τη διάρκεια της διαδικασίας αναδίπλωσης και αυτό με τη σειρά του υποδηλώνει την ύπαρξη των κινητικών μονοπατιών αναδίπλωσης όπως ισχυρίστηκε πρώτος και ο Levinthal το 1968 (Levinthal, 1968). Σε αυτά τα μονοπάτια λοιπόν, τα πρωτεϊνικά μόρια θα περνούσαν από καλά χαρακτηρισμένα και μερικώς δομημένα ενδιάμεσα μόρια.

Ένα αρχικό μοντέλο (Wetlaufer, 1973), υποστήριζε την εμφάνιση μέσω τυχαίων ταλαντεύσεων μιας αρκετά μικρής δομής “πυρήνα”, η οποία δε θα ήταν σταθερή, θα λειτουργούσε όμως σαν εκμαγείο πάνω στο οποίο το υπόλοιπο της πολυπεπτιδικής αλυσίδας θα μπορούσε να αναδιπλωθεί ταχύτατα. Ο ρυθμός της

αναδίπλωσης στο μοντέλο αυτό καθορίζεται από αρχικό βήμα του σχηματισμού του “πυρήνα”. Δυο άλλα μοντέλα, τα οποία είναι αρκετά όμοια, είναι το μοντέλο των δομών (framework model) (Kim *et al.*, 1990; Baldwin, 1989) και το μοντέλο διάχυσης - σύγκρουσης - συγκόλλησης (diffusion - collision -adhesion model) (Karplus *et al.*, 1979; Bashford *et al.*, 1988). Σύμφωνα με το πρώτο μοντέλο, ο σχηματισμός της αναδιπλωμένης δομής είναι αποτέλεσμα του πακεταρίσματος προϋπαρχόντων ανεξάρτητων στοιχείων δευτεροταγούς δομής. Εδώ ο ρυθμός καθορίζεται από το συνδυασμό και το πακετάρισμα των στοιχείων δευτεροταγούς δομής. Στο δεύτερο ενέχεται η υπόθεση ότι υπάρχουν ανεξάρτητα στοιχεία της φυσικής δευτεροταγούς δομής, ή “μικρο-επικράτειες”, τα οποία διαχέονται μαζί και συγκρούονται, ώστε να σχηματιστεί η τριτοταγής δομή. Αρχικά αυτές οι “μικρο-επικράτειες” είναι ασταθείς, παρόλα αυτά όμως όταν βρεθούν κοντά σταθεροποιεί η μια την άλλη μέσω των μεταξύ τους αλληλεπιδράσεων. Ο ρυθμός σχηματισμού ενός σταθερότερου συμπλόκου δίνεται από το γινόμενο του ρυθμού της ταυτόχρονης διάχυσης επί την πιθανότητα όλες οι “μικρο-επικράτειες” να έχουν την κατάλληλη διαμόρφωση (Northey, 2001). Ένας ακόμη μηχανισμός είναι το μοντέλο της υδρόφοβης “κατάρρευσης” (hydrophobic collapse model), σύμφωνα με το οποίο η αναδίπλωση ξεκινά με μια γρήγορη υδρόφοβη “κατάρρευση” (Matheson *et al.*, 1978; Wuthcich *et al.*, 1978; Dill, 1985, 1990). Η κινητήρια δύναμη της αναδίπλωσης απεικονίζεται με την έξοδο του νερού από έναν ταχύτατα σχηματιζόμενο υδρόφοβο πυρήνα, μέσα στον οποίο δημιουργούνται διαδοχικά η δευτεροταγής και η τριτοταγής δομή. Σε αντίθεση με τα πιο πάνω μοντέλα, εμφανίζεται το μοντέλο συναρμολόγησης (jigsaw puzzle model), σύμφωνα με το οποίο δεν υπάρχει κάποιο προτιμητέο σημείο εκκίνησης της αναδίπλωσης και κάθε προσπάθεια αναδίπλωσης ακολουθεί διαφορετικό μονοπάτι (Harrison *et al.*, 1985). Το μοντέλο αυτό είναι παρόμοιο με τη “νέα όψη” της πρωτεϊνικής αναδίπλωσης, σύμφωνα με την οποία οι πρωτεΐνες αναδιπλώνονται μέσω ενός συνόλου μονοπατιών και τα μεμονωμένα ενδιάμεσα μόρια δεν υπάρχουν (Bryngelson *et al.*, 1995).

Πιο πάνω αναφέραμε αλληλεπιδράσεις οι οποίες σταθεροποιούν τη δομή της πρωτεΐνης. Έχει βρεθεί πως η σταθερότητα της φυσικής κατάστασης μιας πρωτεΐνης, όσο αφορά τη μη αναδιπλωμένη μορφή της, είναι οριακή. Η ελεύθερη ενέργεια του Gibb (ΔG_u), σε σχέση με την πρωτεϊνική αναδίπλωση είναι γενικά μεταξύ -5 και -15 kcal/mol (Seckler *et al.*, 1992). Για να προσδιοριστεί όμως ποιες αλληλεπιδράσεις είναι σημαντικές στον καθορισμό της σταθερότητας και διαμόρφωσης των πρωτεϊνικών μορίων, αναπτύχθηκαν οι μέθοδοι της πρωτεϊνικής μηχανικής.

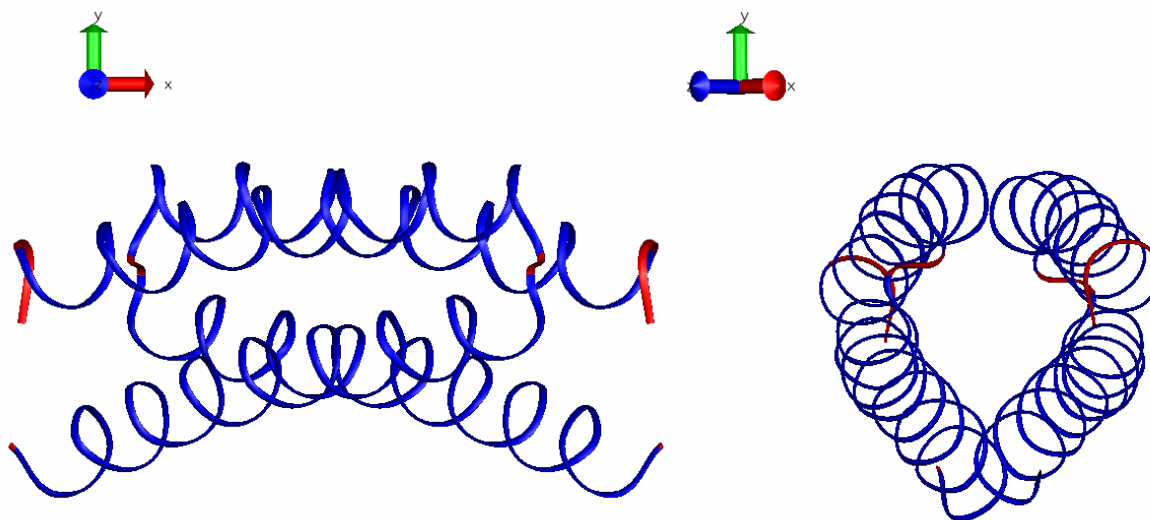
Παρόλα αυτά, τα τελευταία χρόνια έχει αναπτυχθεί ένας μεγάλος όγκος πειραματικών δεδομένων, που υποστηρίζει πως ο ρυθμός και οι μηχανισμοί της πρωτεϊνικής αναδίπλωσης είναι συντηρημένοι και καθορίζονται κατά ένα μεγάλο

μέρος από την τοπολογία της φυσικής δομής, παρά από τις ειδικές αλληλεπιδράσεις, οι οποίες σταθεροποιούν τη δομή μιας πρωτεΐνης.

Για να μπορέσουμε να προσεγγίσουμε κατά κάποιο τρόπο την αναδίπλωση των πρωτεϊνών, χρησιμοποιήσαμε σαν σύστημα αναφοράς την πρωτεΐνη Rop. Πρόκειται για μια πρωτεΐνη, η πολυπεπτιδική αλυσίδα της οποίας είναι από τις καταλληλότερες για τέτοιου είδους μελέτες, εξαιτίας των πλεονεκτημάτων που τη χαρακτηρίζουν. Κωδικοποιείται από το γονίδιο *rop / rom*, τμήμα του πλασμιδίου ColE1, της *E. coli*. Είναι μια μικρή, όξινη κατά βάση, πρωτεΐνη, η πολυπεπτιδική αλυσίδα της οποίας αποτελείται από 63 αμινοξικά κατάλοιπα (Polisky, 1988). Είναι το απλούστερο και πιο συμμετρικό ίσως μέλος μιας οικογένειας πρωτεϊνών, χαρακτηριστικό των οποίων είναι ένα δεμάτιο που σχηματίζεται από τέσσερις αντιπαράλληλες έλικες (Castagnoli *et al.*, 1994). Η δομή της είναι γνωστή στην κρυσταλλική μορφή (κρυσταλλογραφία ακτίνων-X σε ανάλυση 1,7 Å) (Banner *et al.*, 1987), καθώς και στο διάλυμα (2D NMR ανάλυση) (Eberle *et al.*, 1991). Ήταν η πρώτη δομή σπειρωμένου σπειράματος που επιλύθηκε σε τόσο υψηλή ανάλυση. Παρόλα αυτά όμως, είναι αξιοσημείωτο το γεγονός ότι τα τελευταία 7 αμινοξέα που βρίσκονται στο C-τελικό άκρο, δεν μπορούν να εντοπιστούν στην κρυσταλλική δομή, πιθανώς λόγω της υψηλής κινητικότητας αυτών των καταλοίπων (Steif *et al.*, 1993).

Η λειτουργία της Rop έγκειται στον έλεγχο των αντιγράφων του πλασμιδίου ColE1. Πλασμίδια με έλλειμμα μέρους ή και ολόκληρου του γονιδίου *rop* χάνουν την ενεργότητα γονιδίου καταστολής (Moser *et al.*, 1984). Σύμφωνα με πειράματα φθορισμού (χρησιμοποιώντας 2-aminopurine-2'-O-methyl ribose, 2-AP) για τη μέτρηση της ικανότητας συγκράτησης δύο μεταγράφων RNA του πλασμιδίου ColE1, του RNA I (κινητικά ελεγχόμενος καταστολέας της αντιγραφής) και του συμπληρωματικού του RNA II (εκκινητής της αντιγραφής του πλασμιδίου), προκύπτει ότι η Rop λειτουργεί ως ενισχυτής της σταθερότητας του ενδιάμεσου συμπλόκου RNA I – RNA II, προσδεδεμένη σ' αυτό (Castagnoli *et al.*, 1989; Rist *et al.*, 2001). Πειράματα μεταλλάξεων έδειξαν πως τα κατάλοιπα εκείνα που φαίνεται να προσδίδουν στη Rop τη δυνατότητα να αναγνωρίζει το RNA και να προσδέεται σ' αυτό, βρίσκονται στις θέσεις 10, 14, 18 και 25 στην πρώτη έλικα (1/1') καθενός από τα δύο μονομερή. Ένα επιπλέον κατάλοιπο στη θέση 3 φαίνεται να συμμετέχει στην πρόσδεση, χωρίς όμως να διαδραματίζει σημαντικό ρόλο (Predki *et al.*, 1995). Πλεονέκτημα στη χρήση της Rop σαν σύστημα αναφοράς, είναι η δυνατότητα δημιουργίας παραλλαγών της πρωτεΐνης, κυρίως μέσω σημειακών μεταλλάξεων, και η εκτεταμένη μελέτη τους (Magliery *et al.*, 2004), καθώς και η δυνατότητα μελέτης της πρόγνωσης περιορισμών που βασίζονται στην τοπολογία του πρωτεϊνικού σκελετού (Presnell *et al.*, 1989), κυρίως λόγω της δομής της, δηλαδή της δομής του 4-α-ελικοειδούς δεματίου.

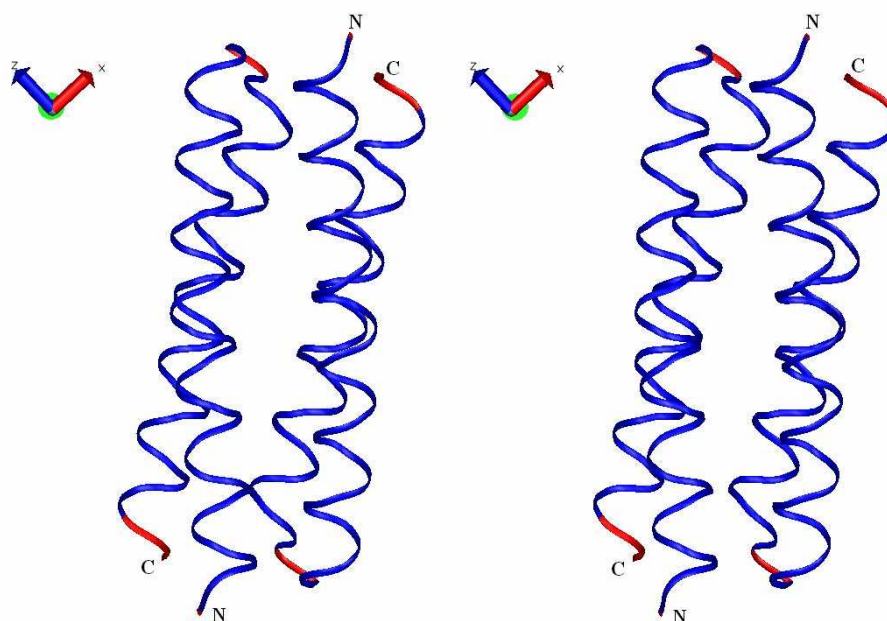
Το 4-α-ελικοειδές δεμάτιο αποτελεί ένα κοινό δομικό μοτίβο ανάμεσα στις φυσικές πρωτεΐνες, το οποίο δομικά μπορεί να παρατηρηθεί είτε ανεξάρτητα, είτε σαν συστατικό μεγαλύτερων πρωτεϊνών (Harris *et al.*, 1994). Για παράδειγμα η δομή αυτή χαρακτηρίζει πρωτεΐνες όπως, η ανθρώπινη αυξητική ορμόνη, η απολιποπρωτεΐνη και οι ιντερλευκίνες. Λειτουργικά μπορεί να συμμετέχει σε διάφορες διαδικασίες, από την αποθήκευση μετάλλων μέχρι και την κυτταρική διέγερση (Kamterkat *et al.*, 1995). Δυο μορφές των 4-α-ελικοειδών δεματίων είναι γνωστές. Στην πρώτη, το εσωτερικό των δεματίων λειτουργεί σαν υποδοχέας μιας προσθετικής ομάδας και σε αυτή την περίπτωση το δεμάτιο συνήθως αποκλίνει από το ένα άκρο. Στη δεύτερη δεν μετέχει κάποια προσθετική ομάδα και οι 4 α έλικες αποκλίνουν συμμετρικά από το κεντρικό σημείο στο οποίο βρίσκονται πιο στενά πακεταρισμένες (Cohen *et al.*, 1990). Το δεμάτιο αντιπροσωπεύεται από την αντιπαράλληλη διευθέτηση τεσσάρων α ελίκων, οι οποίες διασταυρώνονται με γωνία περίπου 20° (Presnell *et al.*, 1989) (**Εικόνα 1-1**). Η διευθέτηση των ελίκων κατά αυτή τη φορά οφείλεται σε μη ομοιοπολικές αλληλεπιδράσεις που οδηγούν στο ισχυρό πακετάρισμά τους, ενώ οι ηλεκτροστατικές αλληλεπιδράσεις είναι εκείνες που παίζουν σημαντικό ρόλο στην προτίμηση της αντιπαράλληλης διευθέτησης των γειτονικών ελίκων στο δεμάτιο (Chou *et al.*, 1988).



Εικόνα 1-1: Αριστερά: Πλαϊνή όψη ενός 4-α-ελικοειδούς δεματίου. Με μπλε χρώμα απεικονίζονται οι περιοχές των ελίκων και με κόκκινο οι περιοχές της θηλιάς, των 2 πρώτων N-τελικών και των 2 τελευταίων C-τελικών καταλοίπων κάθε μονομερούς, που δε συμμετέχουν στο σχηματισμό α έλικας. **Δεξιά:** Κάτοψη του 4-α-ελικοειδούς δεματίου. ⁽¹⁾

Στη φυσική Rop, δυο φουρκέτες της μορφής έλικα - θηλιά - έλικα σε μια διευθέτηση με φορά κεφαλή προς ουρά, σχηματίζουν ένα αριστερόστροφο, αντιπαράλληλο 4-α-ελικοειδές δεμάτιο με εσωτερικές γωνίες περίπου 20° και μια

δεξιόστροφη υπερελικοειδή τοπολογία (**Εικόνα 1-2**) (Predki *et al.*, 1995; Kresse *et al.*, 2001).



Εικόνα 1-2: Σχηματικό στερεοδιάγραμμα της φυσικής Rop. Τα δυο μονομερή απεικονίζονται με μπλε χρώμα, ενώ επισημαίνονται με κόκκινο χρώμα η περιοχή της στροφής, τα 2 N-τελικά και 2 C-τελικά κατάλοιπα κάθε μονομερούς, που δε συμμετέχουν στο σχηματισμό α έλικας.

Το ισχυρό πακετάρισμα στο εσωτερικό των πρωτεϊνών οδήγησε στην υπόθεση ότι συμπληρωματικά σχήματα και μεγέθη καταλοίπων του πυρήνα καθορίζουν την αναδίπλωση της πρωτεΐνης με έναν τρόπο ανάλογο με τον τρόπο που τα κομμάτια ενός παζλ καθορίζουν την τελική μορφή του (Plaxco *et al.*, 1998). Στα 4-α-ελικοειδή δεμάτια το πακετάρισμα του υδρόφοβου πυρήνα φαίνεται να είναι κρίσιμο για τη σταθεροποίηση της πρωτεϊνικής δομής. Οι σφαιρικές πρωτεΐνες χαρακτηρίζονται από καλά πακεταρισμένους υδρόφοβους πυρήνες, οι οποίοι διαδραματίζουν σημαντικό ρόλο τόσο στην αναδίπλωση όσο και στη σταθερότητα των πρωτεϊνών (Dill, 1990). Η κατανόηση της συμβολής μεμονωμένων υδρόφοβων καταλοίπων στην ελεύθερη ενέργεια των πρωτεϊνών, είναι ένα σημαντικό πρόβλημα, που σε αρκετές περιπτώσεις προσεγγίζεται με ανάλυση ενεργειακών και δομικών αλλαγών που εισάγονται με μεταλλάξεις (Vlassi *et al.*, 1999). Τέτοιου είδους μελέτες πραγματοποιήθηκαν χρησιμοποιώντας τη φυσική Rop (Steif *et al.*, 1995), καθώς και παραλλαγές της (Vlassi *et al.*, 1994). Οι μελέτες αυτές έδειξαν ότι η αναδίπλωση και η σταθερότητα καθορίζονται κατά ένα μεγάλο βαθμό από ένα συμμετρικό μοτίβο υδρόφοβων αλληλεπιδράσεων του πυρήνα. Το μοτίβο αυτό περιλαμβάνει μια επανάληψη επτά καταλοίπων, η οποία τοποθετεί συγκεκριμένα κατάλοιπα σε συγκεκριμένες θέσεις (Cohen *et al.*, 1986). Σχετικές στατιστικές αναλύσεις έδειξαν

ότι στα 4-α-ελικοειδή δεμάτια, οι θέσεις του υδρόφοβου πυρήνα έχουν μια προδιάθεση να προτιμούν κατάλοιπα όπως λευκίνη ή αλανίνη (Paliakasis *et al.*, 1992). Πειράματα υποκατάστασης των καταλοίπων του πυρήνα έδειξαν ότι δημιουργούνται πρωτεΐνες σταθερές, διμερείς που μοιάζουν με τη φυσική πρωτεΐνη, όμως εμφανίζουν αυξημένα επίπεδα αποδιάταξης και αναδίπλωσης (Munson *et al.*, 1996). Η Ror χρησιμοποιήθηκε για την εύρεση κρυμμένων ομοιοτήτων με άλλες πρωτεΐνες, μέσω μίας προσέγγισης η οποία δημιουργεί μία σειρά από λειτουργικά μεταλλάγματα (μέσω συνδυαστικής μεταλλαξιγένεσης) και στη συνέχεια χρησιμοποιεί αυτή τη συλλογή μεταλλαγμένων αλληλουχιών σε μία έρευνα των βάσεων δεδομένων για πρωτεϊνικές δομές. Αυτή η προσέγγιση (mutational envelope scanning) μπορεί να είναι αρκετά ισχυρή ώστε να ανιχνεύσει περιοχές περιορισμένης έκτασης, που είναι συντηρημένες λειτουργικά, ανεξάρτητα από την εξελικτική σχέση που έχουν οι πρωτεΐνες (Christ *et al.*, 2003).

Μια συχνή ερώτηση σχετικά με την αναδίπλωση των πρωτεϊνών είναι κατά πόσο παίζουν κάποιο ρόλο οι στροφές και οι θηλιές. Είναι απλά και μόνο σύνδεσμοι μεταξύ των στοιχείων της δευτεροταγούς δομής ή συμμετέχουν με πιο δραστικό τρόπο στον καθορισμό του μονοπατιού αναδίπλωσης και στον καθορισμό της τελικής δομής (Nagi *et al.*, 1999); Θεωρείται πως ο σχηματισμός θηλιάς ή στροφής πυροδοτεί την πρωτεϊνική αναδίπλωση (Linhananta *et al.*, 2002; Wright *et al.*, 1988) και φαίνεται να προσδίδει σταθερότητα στη δομή (Kresse *et al.*, 2001). Επίσης μελέτες έδειξαν ότι οι αλληλεπιδράσεις θηλιάς - έλικας είναι πιο ισχυρές από εκείνες μεταξύ των ελίκων (Chou *et al.*, 1992). Συγκεκριμένα μελέτες της Regan και των συνεργατών της (Nagi *et al.*, 1997) έδειξαν πως το μήκος της θηλιάς απλών θηλιών πολυγλυκίνης είναι καθοριστικός παράγοντας για τη σταθερότητα του δεματίου. Ανάλογες θεωρητικές προσεγγίσεις είχαν ήδη πραγματοποιηθεί σχετικά με το πόσο καθοριστικό ρόλο παίζει το μήκος της θηλιάς στη σταθερότητα της δομής των πρωτεϊνών (Thomas, 1990). Καθώς το μήκος της θηλιάς μεγαλώνει, αυξάνει επίσης και ο ρυθμός αποδιάταξης, ενώ αντίθετα ο ρυθμός επαναδιάταξης ελαττώνεται (Nagi *et al.*, 1997).

Η περιοχή της στροφής στη Ror προσελκύει το ενδιαφέρον, όχι μόνο εξαιτίας της διαμάχης για το ρόλο της θηλιάς στην αναδίπλωση και τη σταθερότητα των δεματίων και των πρωτεϊνών γενικότερα. Ιδιαίτερα τα κατάλοιπα 30, 31 και 32 έχουν αποτελέσει αντικείμενο για πολυάριθμες μελέτες μεταλλάξεων (Castagnoli *et al.*, 1994). Σημαντικό παράδειγμα σχετικά με τη σταθερότητα της δομής και το τι ρόλο παίζει η στροφή, αποτελεί η A31P, μια παραλλαγή της φυσικής Ror. Η μεταλλαγμένη αυτή πρωτεΐνη φέρει αντί για αλανίνη στη θέση 31, το κατάλοιπο της προλίνης (Glykos *et al.*, 1999). Η αλλαγή αυτή οδηγεί όχι μόνο σε διαφορετική διαμόρφωση της δομής του δεματίου, αλλά παράλληλα και σε αλλαγές στις ιδιότητες της επιφάνειας του μορίου και στο πακετάρισμα του υδρόφοβου πυρήνα (Glykos *et*

al., 2004). Πιο συγκεκριμένα εμφανίζεται μείωση κατά 7% στην περιεκτικότητα του μορίου σε α έλικα, καθώς και πολύ σημαντική μείωση στη σταθερότητα του μορίου σε σχέση με τη φυσική Rop, που πιθανόν να οφείλεται στη διαφορετική τοπολογία της A31P και την πιθανή απώλεια κάποιων υδρογονοδεσμών (Peters *et al.*, 1997). Επίσης, απομάκρυνση των πέντε αμινοξικών καταλοίπων της στροφής, οδηγεί στη δημιουργία της μεταλλαγμένης πρωτεΐνης RM6, η οποία παρουσιάζει δραματικές αλλαγές τόσο στη δομή όσο και στη σταθερότητα, σε σύγκριση με τη φυσική πρωτεΐνη. Το μοτίβο από ομοδιμερές 4-α-ελικοειδές μετατρέπεται σε ομοτετραμερές 4-α-ελικοειδές δεμάτιο και η θερμοκρασία τήξης, από 71° C στη φυσική πρωτεΐνη, αυξάνεται στους 101° C (Lassalle *et al.*, 1998).

Ο ρόλος της στροφής στον καθορισμό της δομής και της σταθερότητας των πρωτεϊνών δεν ήταν εύκολο να κατανοηθεί. Για το λόγο αυτό πραγματοποιήθηκε μια σειρά υποκαταστάσεων στην πρωτεΐνη Rop και πιο συγκεκριμένα υποκαταστάθηκε η Asp 30, το πρώτο κατάλοιπο της στροφής, με καθένα από τα υπόλοιπα 19 κατάλοιπα. Στην πλειοψηφία τους οι παραλλαγές της φυσικής Rop δεν απέκλιναν πολύ από τη φυσική δομή. Παρόλα αυτά όμως, δώδεκα από αυτές τις 30 παραλλαγές, αποτελούν πρωτεΐνες σταθερότερες από τη φυσική πρωτεΐνη. Συνεπώς φαίνεται πως υπάρχουν διακριτές θερμοδυναμικές προτιμήσεις για την ταυτότητα του αμινοξέος στη θέση 30 της πρωτεϊνικής αλληλουχίας (Predki *et al.*, 1996). Συνοπτικά μπορούμε να παρατηρήσουμε πως δεν είναι τόσο το είδος των καταλοίπων που παίζει ρόλο στη σημαντικότητα της στροφής, όσο το μήκος της.

Η έρευνα για την εύρεση των παραγόντων που συμμετέχουν στον καθορισμό της αναδίπλωσης ήταν πολύ σημαντική. Αυτό οδήγησε και στην πρόκληση που είναι γνωστή ως “Paracelsus Challenge” (Rose *et al.*, 1994). Επρόκειτο για τη μετατροπή της αναδίπλωσης μιας πρωτεΐνης σε μια διαφορετική, η οποία όμως θα διατηρούσε τουλάχιστον 50% ταυτότητα αλληλουχίας με την αρχική. Η πρωτεΐνη αυτή σχεδιάστηκε τελικά από τις Dalal και Regan έχοντας σαν μοντέλο τη Rop και ονομάστηκε Janus. Η διαμόρφωση της Janus είναι ελικοειδής και εμφανίζει 50% ομοιότητα στην αλληλουχία με την επικράτεια B1, η δομή της οποίας είναι κυρίως β φύλλο (Dalal *et al.*, 2000).

Μια παραλλαγή της Rop που παρουσιάζει ιδιαίτερο ενδιαφέρον, είναι η Ala₂Ile₂-6. Η πρωτεΐνη αυτή έχει έναν εκτεταμένα αναδιαμορφωμένο υδρόφοβο πυρήνα. Πιο συγκεκριμένα τα έξι από τα οκτώ στρώματα υδρόφοβων καταλοίπων, που αποτελούν τον υδρόφοβο πυρήνα, αποτελούνται από κατάλοιπα αλανίνης και ισολευκίνης (Πίνακας 1-1). Η δομή της Ala₂Ile₂-6 (Willis *et al.*, 2000) αποκαλύπτει μια τελείως νέα αναδίπλωση που δημιουργείται από μια στροφή του ενός από τα δυο μονομερή γύρω από την κοινή επιφάνεια του διμερούς. Το προφίλ της ελεύθερης ενέργειας της Ala₂Ile₂-6 είναι επίσης πολύ διαφορετικό από εκείνο της φυσικής

πρωτεΐνης. Η Ala₂Ile₂-6 έχει υψηλότερο σημείο τήξης από τη Rop, όμως υφίσταται μια ελαφρώς μικρότερη αλλαγή ελεύθερης ενέργειας κατά την αποδιάταξη.

Θέση καταλοίπου στον υδρόφοβο πυρήνα	Φυσική Rop	Ala ₂ Ile ₂ -6
5	Glu	Glu
8	Ala	Ile
12	Ala	Ala
15	Ile	Ile
19	Thr	Ala
22	Leu	Ile
26	Leu	Ala
29	Leu	Leu
31	Ala	Ala
34	Gln	Ile
38	Cys	Ala
41	Leu	Ile
45	Ala	Ala
48	Leu	Ile
52	Cys	Ala
56	Phe	Phe

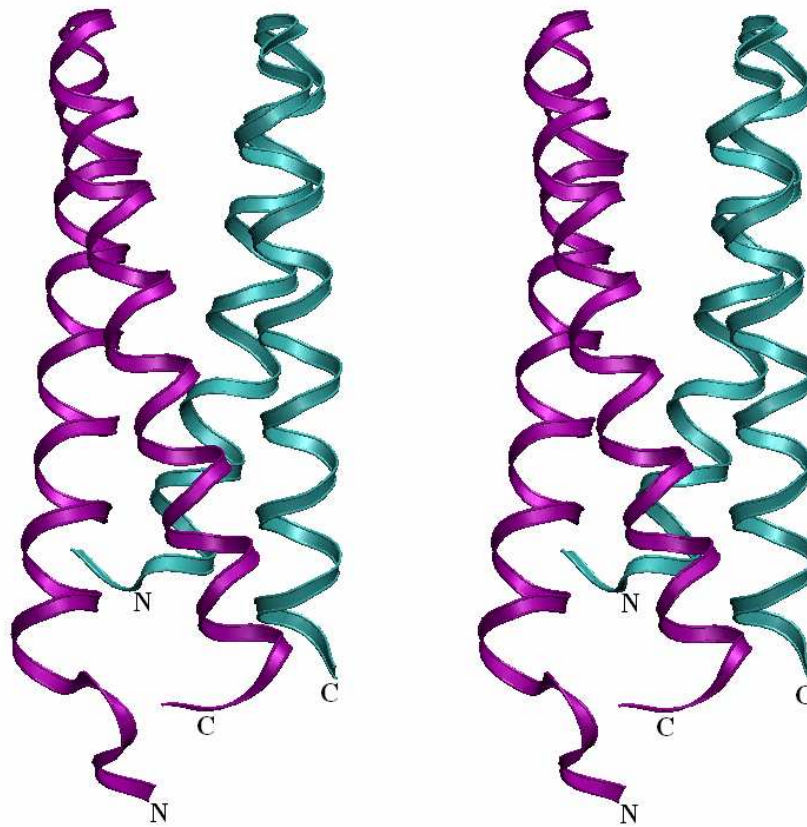
Πίνακας 1-1: Στην πρώτη στήλη φαίνεται η θέση του κάθε καταλοίπου στον υδρόφοβο πυρήνα. Στη δεύτερη στήλη είναι τα κατάλοιπα της φυσικής Rop που αντιστοιχούν σε αυτές τις θέσεις, ενώ στην τρίτη στήλη τα κατάλοιπα της Ala₂Ile₂-6 που αντιστοιχούν. Με μπλε και κόκκινο χρώμα φαίνονται τα κατάλοιπα που έχουν υποκατασταθεί.

Από διάφορες μελέτες πάνω στην Ala₂Ile₂-6 σε σύγκριση πάντα με τη φυσική Rop, προέκυψε ο **Πίνακας 1-2**. Η δομή της Ala₂Ile₂-6, που περιλαμβάνει ένα διμερές στην ασύμμετρη μονάδα, προσδιορίστηκε με κρυσταλλογραφία ακτίνων-X σε αναλυση 1,9Å (Willis *et al.*, 2000).

	Φυσική Rop	Ala₂Ile₂-6
Τάξη	I	III
Ενεργότητα in vitro	Y	N
Ενεργότητα in vivo	Y	N
T_m, °C	64	83
ΔG°, kcal/mol	-7,7	-5,1
Δομή (Μέθοδος)	anti (ακτίνες-X, NMR)	syn (ακτίνες-X)

Πίνακας 1-2: Οι παραλλαγές της Rop έχουν καταταχθεί σε πέντε τάξεις, ανάλογα με τη θερμοδυναμική της αναδίπλωσής τους, την ικανότητα πρόσδεσης του συμπλόκου RNAI-RNAII, την τοπολογία του διμερούς, και την κινητική της αναδίπλωσης. Τα αποτελέσματα βασίζονται στην in vitro ενεργότητα. Συνεπώς οι παραλλαγές με αντίτοπολογία ανήκουν στην I τάξη, όπως η φυσική Rop, ενώ οι παραλλαγές με syn τοπολογία, όπως η Ala₂Ile₂-6, ανήκουν στην III τάξη, στην οποία ανήκουν μόρια με υψηλά ποσοστά α έλικας. Y: πρωτεΐνη Rop που προσδένει RNA, N: καμία ενεργότητα, T_m: θερμοκρασία τήξης, ΔG°: κατά προσέγγιση ελεύθερη ενέργεια του Gibbs, χωρίς αποδιατακτικό. (Τα αποτελέσματα αναπαράγονται άνευ αδείας από Levy et al., 2005).

Αξιοσημείωτο είναι το γεγονός ότι η δομή της Ala₂Ile₂-6 αποκαλύπτει μια τελείως διαφορετική τοπολογία. Στροφή κατά 180° του ενός μονομερούς αλλάζει την αναδίπλωση της πρωτεΐνης, τοποθετώντας και τις δυο στροφές της Ala₂Ile₂-6 στην ίδια μεριά του 4-α-ελικοειδούς δεματίου και τα δυο N και C τελικά άκρα στην άλλη μεριά. Αυτή η δομική αλλαγή εξηγεί επίσης και την απώλεια της ικανότητας της πρωτεΐνης να προσδένει RNA (**Εικόνα 1-3**). Επειδή η δομή της Ala₂Ile₂-6 έχει ήδη επιλυθεί θα την χρησιμοποιήσουμε ως έχει, από τη βάση δεδομένων της PDB (Protein Data Bank) για να πραγματοποιήσουμε τις προσομοιώσεις που μας ενδιαφέρουν. Πιο συγκεκριμένα, και λόγω περιορισμού χρόνου, θα μελετήσουμε συστήματα διμερών και μονομερών που προκύπτουν από την Ala₂Ile₂-6. Μέσω των αποτελεσμάτων που θα προκύψουν από την πραγματοποίηση των προσομοιώσεων αυτών και της μετέπειτα ανάλυσής τους, θα προσπαθήσουμε να κατανοήσουμε το λόγο για τον οποίο η Ala₂Ile₂-6 προτιμά αυτή τη διαμόρφωση και όχι κάποια άλλη, όπως για παράδειγμα εκείνη της φυσικής πρωτεΐνης. Αυτός είναι και ο στόχος της συγκεκριμένης εργασίας.



Εικόνα 1-3: Σχηματικό στερεοδιάγραμμα της Ala_2Ile_2-6 . Είναι εμφανής η αλλαγή της τοπολογίας σε σχέση με τη φυσική *Rop*. Οι δυο στροφές βρίσκονται στη μια πλευρά, ενώ τα δυο *N* και *C* τελικά άκρα στην άλλη.

⁽¹⁾ Όλες οι **Εικόνες** έχουν προετοιμαστεί με τα προγράμματα Rasmol (Bernstein, 1999), VMD (Humphrey *et al.*, 1996) και Raster3D (Merritt *et al.*, 1997).



Κεφάλαιο 2

*Προσομοιώσεις
Μοριακής
Δυναμικής*

Η κατανόηση της αναδίπλωσης των πρωτεϊνών είναι ένας από τους πιο επιθυμητούς στόχους της σύγχρονης μοριακής βιολογίας. Η ικανότητα πρόβλεψης του μηχανισμού αναδίπλωσης και της τελικής δομής μιας πρωτεΐνης, βασιζόμενοι στην αλληλουχία, θα μπορούσε να επηρεάσει δραματικά διάφορα πεδία, από τη βιοχημεία και τη μοριακή ιατρική μέχρι τη νανοτεχνολογία. Ύστερα από δεκαετίες ανεξάρτητων πειραματικών και θεωρητικών προσεγγίσεων, το πεδίο της πρωτεϊνικής αναδίπλωσης εισέρχεται αργά αλλά αναμφισβήτητα σε μια πιο ώριμη περίοδο, όπου πείραμα και θεωρία συγκλίνουν. Οι πειραματικές τεχνικές έχουν πλέον αναπτυχθεί αρκετά ώστε να χρησιμοποιηθούν για την έρευνα της αναδίπλωσης μικρών, ταχέως αναδιπλούμενων πρωτεϊνών και πρωτεϊνικών στοιχείων, ενώ η υπολογιστική δύναμη και οι αλγόριθμοι έχουν φτάσει σε ένα σημείο όπου η προσομοίωση των παραπάνω γεγονότων να είναι πλέον εφικτή (Zagrovic *et al.*, 2001).

Για να είναι λειτουργικές οι πρωτεΐνες πρέπει να αναδιπλωθούν σε μια συγκεκριμένη τρισδιάστατη δομή (φυσική κατάσταση) (Caflish, 2004). Η πρωτεϊνική αναδίπλωση είναι μια περίπλοκη διαδικασία, η οποία περιλαμβάνει μη ομοιοπολικές αλληλεπιδράσεις σε ολόκληρο το μόριο, πολλούς βαθμούς ελευθερίας και μια ισορροπία ανάμεσα στην συμβολή της ενθαλπίας και της εντροπίας στην ελεύθερη ενέργεια (Dill *et al.*, 1997). Η πιο σταθερή αναδίπλωση μιας πρωτεΐνης εξαρτάται από την αμινοξική σύνθεση, από το περιβάλλον διάλυμα (σύνθεση, τιμή pH), και τη θερμοδυναμική κατάσταση (θερμοκρασία, πίεση). Αλληλεπιδράσεις σε ατομικό επίπεδο καθορίζουν την ισορροπία μεταξύ των αναδιπλωμένων και αποδιαταγμένων παραγόντων. Οι υπολογιστικές προσομοιώσεις μπορούν να χρησιμοποιηθούν για να υπολογιστεί η ισορροπία αναδιπλωμένης / αποδιαταγμένης δομής, ενώ οι προσομοιώσεις μοριακής δυναμικής (MD) μπορούν να αποδώσουν την κινητική της διαδικασίας αναδίπλωσης (van Gunsteren *et al.*, 2001).

Οι έρευνες σε αυτή τη βάση ξεκίνησαν όταν οι Bruce και Karplus αποφάσισαν να αναπτύξουν ένα πρόγραμμα το οποίο δεδομένων κάποιων πληροφοριών, όπως η αμινοξική αλληλουχία και ένα σετ συντεταγμένων, θα μπορούσε να υπολογίσει την ενέργεια του συστήματος και τα παράγωγά του σαν συνάρτηση των ατομικών θέσεων. Το πρόγραμμα αυτό ονομάστηκε Pre-CHARMM (Karplus, 2003). Δεδομένου ενός προγράμματος που θα μπορούσε να υπολογίσει τις δυνάμεις στα άτομα μιας πρωτεΐνης για ελαχιστοποίηση της ενέργειας, το επόμενο βήμα θα ήταν να χρησιμοποιηθούν αυτές οι δυνάμεις στην εξίσωση του Νεύτωνα ώστε να υπολογιστούν τα δυναμικά. Οι προσομοιώσεις μπορούν να δώσουν πολλές λεπτομέρειες που αφορούν μεμονωμένες κινήσεις σωματιδίων σε συνάρτηση του χρόνου. Επομένως μπορούν να χρησιμεύσουν ώστε να απευθύνουν ερωτήματα για τις ιδιότητες ενός συστήματος, μερικές φορές πιο εύκολα από πειράματα στο πραγματικό σύστημα.

Υπάρχουν τρεις κατηγορίες εφαρμογών των μεθόδων προσομοίωσης σε μακρομοριακό, καθώς και σε άλλα επίπεδα. Η πρώτη χρησιμοποιεί τις

προσομοιώσεις ως ένα μέσο ελέγχου του χώρου που διαμορφώνεται (sampling configuration space). Χρησιμοποιείται κατά βάση για τον καθορισμό ή τη βελτιστοποίηση δομών. Η δεύτερη χρησιμεύει για την περιγραφή του συστήματος σε κατάσταση ισορροπίας συμπεριλαμβανομένων των δομικών και κινητικών του ιδιοτήτων, καθώς και τις τιμές των θερμοδυναμικών παραμέτρων. Εδώ πρέπει να σημειώσουμε ότι είναι απαραίτητο να υπολογιστεί επαρκώς ο χώρος που διαμορφώνεται και κάθε σημείο να σταθμίζεται με τον κατάλληλο παράγοντα Boltzmann. Στην τρίτη εφαρμογή, εξετάζεται η πραγματική δυναμική του συστήματος, όπου δεν απαιτείται απλά ο υπολογισμός του χώρου (sampling configuration space) και η στάθμιση με τον κατάλληλο παράγοντα Boltzmann, αλλά θα πρέπει να γίνει με τρόπο που να αναπαριστά σωστά την εξέλιξη του συστήματος συναρτήσει του χρόνου. Για τις πρώτες δύο εφαρμογές μπορούν να χρησιμοποιηθούν οι προσομοιώσεις Monte Carlo αλλά και οι προσομοιώσεις μοριακής δυναμικής. Στην τρίτη περίπτωση που το ενδιαφέρον επικεντρώνεται στις κινήσεις και την εξέλιξή τους στο χρόνο, μόνο οι προσομοιώσεις μοριακής δυναμικής μας παρέχουν τις απαραίτητες πληροφορίες (Karplus *et al.*, 2002).

Οι ραγδαίες εξελίξεις της υπολογιστικής δύναμης αυξάνουν τις πιθανότητες χρήσης τεχνικών προσομοιώσεων μοριακής δυναμικής, ώστε να δώσουν σημαντικές πληροφορίες για την πρωτεϊνική αναδίπλωση. Οι προσομοιώσεις μοριακής δυναμικής χρησιμοποιούνται για την παραγωγή μοντέλων ελεύθερης ενέργειας αναδίπλωσης καθώς και για αποδιαταγμένες ή μερικώς αναδιπλωμένες μορφές μιας πρωτεΐνης (Smith *et al.*, 2002). Οι προσομοιώσεις μοριακής δυναμικής αντιπροσωπεύουν ένα ισχυρό εργαλείο για την ανάλυση των μηχανισμών που καθορίζουν τη μοναδική τρισδιάστατη δομή που θα υιοθετήσει μια πρωτεΐνη, εξασφαλίζοντας τις ατομικές λεπτομέρειες της διαδικασίας της αναδίπλωσης. Εξασφαλίζουν δηλαδή ένα ισχυρό μέσο για την έρευνα των βιολογικών συστημάτων σε ατομικό επίπεδο. Ένας ερευνητής μπορεί έτσι να εξετάσει το φυσικό σύστημα που τον ενδιαφέρει, αλλά ταυτόχρονα και να πραγματοποιήσει προσομοιώσεις με υπερβολικές ή τεχνητές συνθήκες, για να εξετάσει περαιτέρω τους προτεινόμενους μηχανισμούς (Grayson *et al.*, 2003).

Ο κυριότερος περιορισμός στη χρήση των προσομοιώσεων είναι ο χρόνος. Οι προσομοιώσεις πραγματοποιούνται σε μια χρονική διάρκεια της τάξεως των νανοδευτερολέπτων, ενώ στην πραγματικότητα οι πρωτεΐνες αναδιπλώνονται σε χιλιοστά του δευτερολέπτου ή ακόμα και σε δευτερόλεπτα. Για να παρακαμφθεί το πρόβλημα αυτό, προτιμώνται προσομοιώσεις αποδιάταξης για τη μελέτη των αλλαγών στη διαμόρφωση των πρωτεϊνών, επειδή πραγματοποιούνται υπό ακραίες συνθήκες που επιταχύνουν το φαινόμενο της αποδιάταξης, ώστε να προσεγγίζει τη χρονική κλίμακα των προσομοιώσεων μοριακής δυναμικής (Gruija *et al.*). Η ομοιοτική επικράτεια Engrailed της *Drosophila melanogaster* έχει τη μεγαλύτερη σταθερά ταχύτητας αναδίπλωσης και αποδιάταξης που έχει παρατηρηθεί μέχρι

σήμερα. Ο χρόνος ημιζωής της αποδιάταξης υπολογίζεται περίπου σε 7.5 ns στους 100°C, γεγονός που επιτρέπει την δοκιμασία, σε πραγματικό χρόνο, προσομοιώσεων αποδιάταξης του συστήματος αυτού, σε ρεαλιστική θερμοκρασία (Mayor *et al.*, 2000).

Εξαιτίας του περιορισμού που αναφέρθηκε προηγουμένως, αναπτύχθηκαν προσομοιώσεις μοριακής δυναμικής όπου ένα συγκεκριμένο γεγονός προκαλείται μέσω της εφαρμογής έμμεσων δυναμικών ή εξωγενών δυνάμεων οι οποίες οδηγούν το σύστημα σε μια προκαθορισμένη διαμόρφωση. Για το λόγο αυτό αναπτύχθηκαν δυο μέθοδοι, η δειγματοληψία κάλυψης (umbrella sampling) και η καθοδηγούμενη προσομοίωση μοριακής δυναμικής (steered molecular dynamics simulation), οι οποίες εφαρμόστηκαν σε διάφορα μοριακά συστήματα. Στην πράξη αυτές οι μέθοδοι απαιτούν προκαθορισμένες συντεταγμένες που συνδέουν τις αρχικές και τελικές διαμορφώσεις των συστημάτων. Μια πιο νέα μέθοδος προσομοίωσης είναι οι διαδραστικές προσομοιώσεις μοριακής δυναμικής (IMD, Interactive Molecular Dynamics) που χρησιμοποιείται για την εξέταση δυναμικών διαδικασιών (Grayson *et al.*, 2003).

Οι υπολογιστικές προσομοιώσεις λειτουργούν σαν μια γέφυρα μεταξύ μικροσκοπικού μήκους και κλίμακας χρόνου και του μακροσκοπικού κόσμου του εργαστηρίου, δηλαδή ξεκινάμε με μια πρόβλεψη για τις αλληλεπιδράσεις μεταξύ των μορίων και καταλήγουμε με “ακριβείς” προβλέψεις για τις πρωτεΐνες (Allen, 2004). Κατά πόσο οι προσομοιώσεις μοριακής δυναμικής είναι αξιόπιστες είναι ένα καίριο ερώτημα. Η απάντηση δίνεται μέσα από σύγκριση αποτελεσμάτων μεταξύ των αποτελεσμάτων των προσομοιώσεων και άλλων πειραματικών δεδομένων. Ένα χαρακτηριστικό παράδειγμα αποτελεί η υψηλή πιστότητα των αποτελεσμάτων από την ανάλυση time-resolved x-ray δομών και MD προσομοιώσεων στο μόριο της μυοσφαιρίνης (Hummer *et al.*, 2004).

Οι προσομοιώσεις μοριακής δυναμικής χρησιμοποιούνται επίσης για μελέτες σχετικά με τη δομή και την κίνηση και των DNA ολιγονουκλεοτιδίων στο διάλυμα. Πρόσφατες βελτιώσεις στα δυναμικά πεδία των νουκλεϊκών οξέων και η ανάπτυξη παραγόντων “δεύτερης γενιάς” ειδικά για προσομοιώσεις που περιλαμβάνουν το διαλύτη, έχουν προάγει την ικανότητα των προσομοιώσεων να προβλέψουν με ακρίβεια τη δομή του DNA. Για να διαπιστωθεί η ικανότητα των προσομοιώσεων στην πρόβλεψη της δομής των νουκλεϊκών οξέων, εφαρμόζεται ο πυρηνικός μαγνητικός συντονισμός NMR (Nuclear Magnetic Resonance). Το NMR, εκτός από τη συμμετοχή στον προσδιορισμό δομών, χρησιμεύει στη μελέτη της δυναμικής και θερμοδυναμικής των μακρομορίων. Τα αποτελέσματα που προκύπτουν τόσο από τον NMR όσο και από τις προσομοιώσεις δεν απέχουν πολύ μεταξύ τους, γεγονός που ενισχύει τη χρήση των προσομοιώσεων (Arthanari *et al.* 2003).

Είναι φανερό ότι οι MD προσομοιώσεις είναι αρκετά αξιόπιστες για τις μελέτες της κίνησης και αναδίπλωσης διαφόρων μορίων, όμως πόσο εύκολο είναι να πραγματοποιηθούν τέτοιου είδους προσομοιώσεις; Πόσοι υπολογιστές χρειάζονται

και πόσα δίκτυα πρέπει να δημιουργηθούν ώστε να καλυφθεί ο όγκος των δεδομένων που απαιτείται για να είναι επαρκή τα αποτελέσματα; Οι υπολογιστές που έχουμε στη διάθεσή μας σήμερα δεν είναι αρκετά γρήγοροι ώστε να ερευνηθούν όλες οι δυνατές διαμορφώσεις (Hess, 2002). Σε αυτή τη βάση, εργάστηκε ο Izaguirre και η ομάδα του και πρότεινε μια νέα μέθοδο που ονομάζεται Verlet-I/r-RESPA. Πρόκειται για μια μέθοδο όπου το αργό μέρος του δυναμικού ενέργειας τροποποιείται και έχουμε τη δυνατότητα μεγαλύτερων βημάτων ώστε να πραγματοποιηθούν προσομοιώσεις μεγαλύτερης διάρκειας (Izaguirre *et al.*, 1999). Η μεγαλύτερης διάρκειας γνωστή προσομοίωση αφορά την αναδίπλωση της επικράτειας κεφαλής της βιλίνης, μιας πρωτεΐνης 36 καταλοίπων και είναι της τάξης του 1μs (Duan *et al.*, 1998).

Οι MD προσομοιώσεις χρησιμοποιήθηκαν σε μια αρκετά μεγάλη ποικιλία συστημάτων. Χαρακτηριστικό παράδειγμα αποτελεί η μελέτη της διαδικασίας ανταλλαγής νουκλεοτιδίου της υπομονάδας G_{α} της μεταγωγής G_{α} . Τα αποτελέσματα αυτού του πειράματος αποκαλύπτουν έναν αριθμό δομικών και δυναμικών αλλαγών που εξαρτώνται από το νουκλεοτίδιο που δεν παρατηρήθηκαν στις διάφορες κρυσταλλικές δομές της G_{α} (Ceruso *et al.*, 2004). Ένα ακόμα παράδειγμα είναι η διερεύνηση της φύσης της παραμόρφωσης του DNA που προκαλείται από ένα διμερές κυκλοβουτανίου πυριμιδίνης (cyclobutane pyrimidine dimer) (Yamaguchi *et al.*, 1998). Οι προσομοιώσεις χρησιμοποιούνται επίσης στη διερεύνηση της δυναμικής και των διαμοριακών αλληλεπιδράσεων πρωτεϊνών όπως ο trans ενεργοποιητής του HIV-1, η Tat. Από τέτοιες μελέτες ταυτοποιήθηκαν διαμοριακές αλληλεπιδράσεις σημαντικές για τη σταθερότητα καθώς και σημαντικές ηλεκτροστατικές αλληλεπιδράσεις της N-τελικής περιοχής με τη βασική επικράτεια της πρωτεΐνης, σύμφωνα με την στοχευμένη μεταλλαξιγένεση που εφαρμόστηκε (Pantano *et al.*, 2004). Η χρήση των MD προσομοιώσεων βοήθησε στην αποκάλυψη μιας γέφυρας άλατος η οποία δημιουργεί μια κινητική παγίδα (kinetic trap) στην αποδιάταξη της truncated staphylococcal nuclease (SNase). Η γέφυρα αυτή σχηματίζεται μεταξύ της αργινίνης 105 και του γλουταμινικού 135 (Gruija *et al.*, 2003). Ιδιαίτερο ενδιαφέρον παρουσιάζουν οι αλληλεπιδράσεις μεταξύ πρωτεϊνών, οι οποίες διαδραματίζουν σημαντικούς ρόλους στη ρύθμιση ποικίλων βιολογικών διεργασιών. Πρόσφατα, έχει χαρακτηριστεί μία επικράτεια, η PB1 (Phox και Bem1), η οποία έχει συντηρηθεί κατά την εξέλιξη και εμφανίζεται σε διάφορες πρωτεΐνες που συμμετέχουν στη μεταγωγή σήματος, στην πολικότητα του κυττάρου και στην επιβίωση. Η επικράτεια αυτή και η συμπεριφορά της αποτέλεσε αντικείμενο μελέτης προσομοιώσεων (van Drogen-Petit *et al.*, 2004).

2.1 Προσομοιώσεις μοριακής δυναμικής της $\text{Ala}_2\text{Ile}_2\text{-6}$

Οι προσομοιώσεις της $\text{Ala}_2\text{Ile}_2\text{-6}$ πραγματοποιήθηκαν σε μια συστοιχία υπολογιστών τύπου Beowulf, με τη χρήση του δυναμικού πεδίου charmm27 το οποίο μας δίνει τις πληροφορίες για την τοπολογία και άλλες παραμέτρους της πρωτεΐνης (MacKerell *et al.*, 1998). Το δυναμικό πεδίο είναι μία μαθηματική έκφραση της δυναμικής ενέργειας των ατόμων του συστήματος. Υπάρχουν διάφορα δυναμικά πεδία για πρωτεΐνες, όπως τα AMBER, CHARMM, OPLS-AA, τα οποία φαίνεται να δίνουν παρόμοια αποτελέσματα (Gnanakaran *et al.*, 2005; Price *et al.*, 2002; Shirts *et al.*, 2003). Το αρχείο με τις παραμέτρους περιέχει πληροφορίες όπως η ισχύς και το μήκος ισορροπίας των ομοιοπολικών δεσμών. Το αρχείο με την τοπολογία περιέχει πληροφορία σχετικά με τους ατομικούς τύπους, τα φορτία και τον τρόπο σύνδεσης των ατόμων στο μόριο. Στο αρχείο pdb που παίρνουμε από τη βάση δεδομένων της PDB περιέχεται η πληροφορία για τις συντεταγμένες των ατόμων. Ακολουθεί η ευθυγράμμιση των αξόνων της αδράνειας του μορίου με το ορθοκανονικό σύστημα αξόνων και η τοποθέτηση του κέντρου βάρους του στην αρχή των αξόνων χρησιμοποιώντας το πρόγραμμα moleman (Kleywegt, 1995). Η προσομοίωση πραγματοποιείται σε όξινο pH και όχι σε ουδέτερο, κατά συνέπεια οι ιστιδίνες θα πρέπει να βρίσκονται σε πρωτονιωμένη μορφή, δηλαδή τα δύο άτομα αζώτου πρέπει να είναι σε πρωτονιωμένη μορφή (αντί για HIS χρησιμοποιούμε HSP). Εναλλακτικά, μπορεί να είναι πρωτονιωμένο μόνο το δ άζωτο (οπότε αναφέρεται σαν HSD), ή μόνο το ϵ (οπότε αναφέρεται σαν HSE). Τέλος, χρησιμοποιείται ένα αρχείο psf (protein structure file) στο οποίο περιέχεται η πληροφορία σχετικά με τη δομή της πρωτεΐνης και το αρχείο pdb που αναφέραμε προηγουμένως. Το psf δημιουργείται με τη βοήθεια του psfgen, το οποίο εμπεριέχεται στο NAMD (Kale *et al.*, 1999). Έτσι, έχουμε όλες τις απαραίτητες πληροφορίες για την πρωτεΐνη.

Ακολουθώς ασχολούμαστε με την προετοιμασία του συστήματος την οποία περιγράφουμε αναλυτικά στο 3^ο κεφάλαιο. Η πρωτεΐνη στο σημείο αυτό τοποθετείται σε ένα κουτί με νερά και ιόντα ώστε να δημιουργηθεί ένα κατά προσέγγιση φυσιολογικό υδατικό περιβάλλον. Τα μόρια νερού που χρησιμοποιούνται είναι τύπου TIP3 (Shirts *et al.*, 2003). Η παρουσία των ιόντων χρησιμεύει στην “εξουδετέρωση” του φορτίου και έτσι συμβάλλει στη σταθερότητα του συστήματος. Τυπικά τα ιόντα πρέπει να τοποθετηθούν σε περιοχές όπου εμφανίζονται τοπικά ελάχιστα του δυναμικού, αφού κατά τη διάρκεια της προσομοίωσης θα μετατοπιστούν προς αυτές τις περιοχές. Το φορτίο του συστήματος περιέχεται στο αρχείο psf. Στις προσομοιώσεις μοριακής δυναμικής με periodic boundary conditions η ενέργεια λόγω ηλεκτροστατικών αλληλεπιδράσεων υπολογίζεται με τη μέθοδο particle-mesh Ewald (PME), η οποία απαιτεί την ύπαρξη ουδέτερου συστήματος (Darden *et al.*,

1993). Εάν το μόριο φέρει σχετικά μικρό ηλεκτρικό φορτίο, ο πιο εύκολος τρόπος να τοποθετηθούν τα ιόντα στο σύστημα είναι αυτόματα με τη βοήθεια του VMD (Humphrey *et al.* 1996). Έτσι υπολογίζεται ο αριθμός των ιόντων νατρίου και χλωρίου που απαιτούνται και τοποθετούνται τυχαία στο διάλυμα, αλλά παράλληλα διατηρούνται οι ελάχιστες αποστάσεις ανάμεσα στα ιόντα και το μόριο, αλλά και μεταξύ των ιόντων. Το μέγεθος του κουτιού που χρησιμοποιείται είναι αρκετά μεγάλο ώστε να αποφευχθεί η σύγκρουση με κάποιο από τα γειτονικά είδωλα, αφού οι προσομοιώσεις πραγματοποιούνται σε periodic boundary conditions. Όταν έχουμε periodic boundary conditions, τοποθετούμε γύρω από το σύστημα που μελετάμε (primary cell) (και στις τρεις διαστάσεις) πανομοιότυπες μονάδες του εαυτού του ή “αντίγραφα” (replicas). Τα “αντίγραφα” αυτά λέγονται “image cells” (Haile, 1997). Κατ’ αυτό τον τρόπο τα περιβάλλοντα συστήματα αλληλεπιδρούν με τα άτομα του πραγματικού συστήματος, με αποτέλεσμα τη δημιουργία ενός περιβάλλοντος όπου το σύστημα δε βλέπει κενό. Οι συνθήκες αυτές ελαχιστοποιούν σε μεγάλο βαθμό την επιφανειακή αλληλεπίδραση των μορίων νερού. Έτσι το σύστημα διαφέρει σχετικά από ένα περιβάλλον που μοιάζει περισσότερο με την κρυσταλλική κατάσταση.

Μετά την προετοιμασία της προσομοίωσης έπεται η ελαχιστοποίηση της ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Τα βήματα που ακολουθούνται περιγράφονται αναλυτικά στο 3^ο κεφάλαιο. Κατά την παραγωγική φάση επιλύονται οι εξισώσεις του Νεύτωνα για κάθε άτομο του συστήματος και καταγράφεται η τροχιά που ακολουθεί. Οι συνθήκες που επικρατούν είναι ισοβαρής/ισόθερμη (NpT), δηλαδή σταθερός αριθμός ατόμων (N), σταθερή πίεση (P) και σταθερή θερμοκρασία (T). Για την διατήρηση της πίεσης και θερμοκρασίας χρησιμοποιούνται τα Nosé-Hoover Langevin Dynamics και η μέθοδος ελέγχου Langevin piston barostat. Για τις εξισώσεις της ταχύτητας επιλύεται ο αλγόριθμος Verlet. Ο αλγόριθμος Verlet είναι απευθείας χρονικά αναστρέψιμος (exactly time reversible), συμπλεκτικός (symplectic), έχει χαμηλή τάξη μεγέθους στο χρόνο (low order in time) οπότε επιτρέπει μεγαλύτερα βήματα (timesteps), απαιτεί μόνο έναν υπολογισμό δύναμης σε κάθε βήμα και είναι εύκολος ο προγραμματισμός (Allen, 2004). Εφόσον η λύση των εξισώσεων γίνεται κατά προσέγγιση σε διακριτά χρονικά βήματα, οι περιορισμοί (constraints) δεν θα τηρούνται σε όλη τη διάρκεια της προσομοίωσης. Αυτό το εμπόδιο μπορεί να παρακαμφθεί με τον προσδιορισμό των δυνάμεων περιορισμού ώστε οι περιορισμοί να ικανοποιούνται ακριβώς στο τέλος κάθε χρονικού βήματος (timestep). Αυτό επιτυγχάνεται μέσω του αλγόριθμου SHAKE (Ryckaert *et al.*, 1977).



Κεφάλαιο 3

*Προετοιμασία
Συστημάτων*

Στο κεφάλαιο αυτό θα γίνει μια πλήρης περιγραφή της προετοιμασίας κάθε συστήματος που χρησιμοποιήθηκε για τις προσομοιώσεις, χωριστά. Η προετοιμασία περιλαμβάνει την τροποποίηση του μορίου της πρωτεΐνης που μελετάμε, την κατασκευή του πλαισίου με τα νερά, την τοποθέτηση των ιόντων και άλλα. Τα μόρια που χρησιμοποιήσαμε είναι τα ακόλουθα:

- Διμερές της Ala_2Ile_2-6 με τοπολογία Ala_2Ile_2-6 και αλληλουχία Ala_2Ile_2-6 (Syn τοπολογία)
- Διμερές της Ala_2Ile_2-6 με τοπολογία της φυσικής Rop και αλληλουχία Ala_2Ile_2-6 (Anti τοπολογία)
- Φυσικό (native) μονομερές της Ala_2Ile_2-6
- Τεχνητό (artificial) μονομερές της Ala_2Ile_2-6
- Μονομερές της Ala_2Ile_2-6 με τη διαμόρφωση ενός A31P μονομερούς και αλληλουχία Ala_2Ile_2-6 (7 προσομοιώσεις)

Στον **Πίνακα 3-1** βλέπουμε συνοπτικά τι περιλαμβάνει καθένα από τα συστήματα.

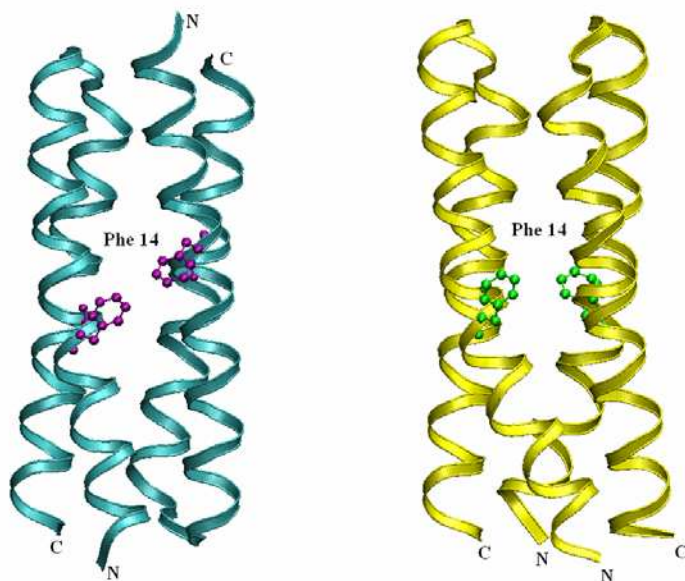
3.1 Εισαγωγή

Η διμερής πρωτεΐνη Rop, όπως είδαμε στο Κεφάλαιο 1, είναι ένα 4-α-ελικοειδές δεμάτιο με έναν καλοσχηματισμένο υδρόφοβο πυρήνα. Η Ala_2Ile_2-6 είναι μια παραλλαγή της Rop με έναν εκτεταμένα ανασχεδιασμένο υδρόφοβο πυρήνα. Αξιοσημείωτο είναι το γεγονός ότι η δομή της Ala_2Ile_2-6 αποκαλύπτει μια τελείως νέα μορφή αναδίπλωσης (**Εικόνα 3-1**), η οποία σχηματίζεται μέσω μιας δομικής περιστροφής των δύο μονομερών γύρω από την κοινή επιφάνεια του διμερούς (Willis *et al.*, 2000). Η δομή αυτή της Ala_2Ile_2-6 δείχνει τη σημαντικότητα του στενού πακεταρίσματος των καταλοίπων του πυρήνα και την υιοθέτηση των ευνοϊκών τιμών για τις πλευρικές αλυσίδες των στροφομερών του πυρήνα, στον καθορισμό των αλληλεπιδράσεων μεταξύ των ελίκων στο τέσσερα-ελικοειδές δεμάτιο.

Πρέπει επίσης να σημειώσουμε ότι η Ala_2Ile_2-6 έχει υψηλότερο σημείο τήξης από τη φυσική Rop, υφίσταται μια ελάχιστη μείωση στην ελεύθερη ενέργεια κατά το ξετύλιγμα και, αντίθετα με τη φυσική Rop, ή άλλες παραλλαγές της, δεν προσδένει RNA. Πιο αναλυτικά, η Ala_2Ile_2-6 προκύπτει από την τοποθέτηση καταλοίπων αλανίνης στις θέσεις “a” και καταλοίπων ισολευκίνης στις θέσεις “d” στα έξι μεσαία στρώματα (layers) του πυρήνα της φυσικής Rop (**Πίνακας 3-2, Εικόνα 3-2**).

Όνομα μορίου	Αριθμός καταλοίπων πρωτεΐνης	Αριθμός ατόμων πρωτεΐνης	Αριθμός ατόμων H ₂ O (TIP3)	Αριθμός Na ⁺	Αριθμός Cl ⁻	Διαστάσεις κυψελίδας (Å ³)	Διάρκεια προσομοίωσης σε ns
Ala₂Pe₂-6 με syn τοπολογία	A: 60	A: 929	24669	7	1	80.7 x 58.9 x 58.9	10
	B: 56	B: 891					
Ala₂Pe₂-6 με αντί τοπολογία (Rop-like)	A: 56	A: 901	23571	6	2	78.59 x 60.1 x 56.85	10
	B: 56	B: 901					
Native Ala₂Pe₂-6	A: 60	A: 929	18423	5	1	80.4 x 51.3 x 49.9	20
Artificial Ala₂Pe₂-6	A: 60	A: 929	18804	5	1	79.6 x 50.6 x 51.9	8
Ala₂Pe₂-6 με διαμόρφωση μονομερούς A31P (x 7)	A: 56	A: 901	20850	4	2	78.2 x 60.1 x 49.2	2
	A: 56	A: 901	20850	4	2	78.2 x 60.1 x 49.2	2
	A: 56	A: 901	20850	4	2	78.2 x 60.1 x 49.2	2
	A: 56	A: 901	20850	4	2	78.2 x 60.1 x 49.2	2
	A: 56	A: 901	20850	4	2	78.2 x 60.1 x 49.2	2
	A: 56	A: 901	20850	4	2	78.2 x 60.1 x 49.2	1.4
	A: 56	A: 901	20850	4	2	78.2 x 60.1 x 49.2	10

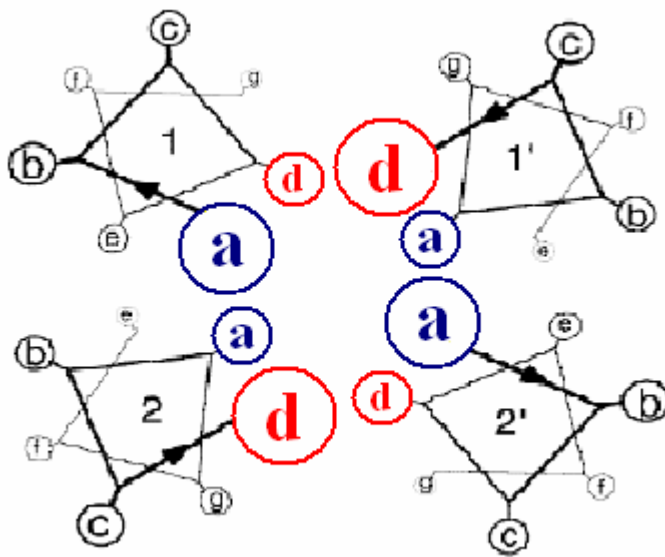
Πίνακας 3-1: Κάθε γραμμή αντιστοιχεί σε καθεμία από τις προσομοιώσεις που πραγματοποιήθηκαν. Δίδονται ο αριθμός των ατόμων του μορίου της πρωτεΐνης, των μορίων του νερού, τον αριθμό των ιόντων, οι διαστάσεις της κάθε κυψελίδας και το χρονικό διάστημα που διήρκεσε η κάθε προσομοίωση. Στη δεύτερη στήλη με A και B σημαίνονται οι αλυσίδες του μορίου της πρωτεΐνης. Με κyanό χρώμα σκιάζονται τα δεδομένα που αντιστοιχούν στα διμερή, ενώ με κίτρινο χρώμα εκείνα που αντιστοιχούν στα μονομερή.



Εικόνα 3-1: Αριστερά: Άποψη του 4α-ελικοειδούς δεματίου της Ror (με κυανό χρώμα). Με μοβ χρώμα μπορούμε να δούμε το κατάλοιπο Phe της θέσης 14. Δεξιά: Άποψη της δομής της Ala₂Ile₂-6 (με κίτρινο χρώμα). Με πράσινο χρώμα παρατηρούμε τη Phe 14. Τα κατάλοιπα Phe χρησιμοποιούνται για να δείξουμε την αλλαγή στην αναδίπλωση, με την περιστροφή των μονομερών γύρω από την κοινή επιφάνεια του διμερούς.

Πρωτεΐνη	Αριθμός καταλοίπων	Αλληλουχία	Μοριακό βάρος	Ανάλυση ακτίνων-X
Φυσική Ror (PDBID: 1ROP)	63	MET1 THR2 LYS3 GLN4 GLU5 LYS6 THR7 ALA8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 THR19 LEU20 THR21 LEU22 LEU23 GLU24 LYS25 LEU26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 GLN34 ALA35 ASP36 ILE37 CYS38 GLU39 SER40 LEU41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 LEU48 TYR49 ARG50 SER51 CYS52 LEU53 ALA54 ARG55 PHE56 GLY57 ASP58 ASP59 GLY60 GLU61 ASN62 LEU63	7210	1.70 Å
Ala ₂ Ile ₂ -6 (PDBID: 1F4N)	63	GLY1 THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57 ASP58 ASP59 GLY60 GLU61 ASN62 LEU63	7027	1.90 Å

Πίνακας 3-2: Στην πρώτη στήλη φαίνονται τα ονόματα και οι κωδικοί των δύο δομών που πήραμε από τη βάση δεδομένων της PDB (Protein Data Bank). Στη δεύτερη στήλη φαίνεται ο αριθμός των καταλοίπων ανά αλυσίδα. Η τρίτη στήλη περιλαμβάνει την αλληλουχία μιας αλυσίδας από κάθε μόριο. Με μπλε σκίαση μπορούμε να διακρίνουμε τα κατάλοιπα στις θέσεις “a”, ενώ με κόκκινη τα κατάλοιπα στις θέσεις “d” των έξι κεντρικών layers του υδρόφοβου πυρήνα. Τα κατάλοιπα των “a” θέσεων της φυσικής Ror έχουν αντικατασταθεί με ισολευκίνη, ενώ των “d” θέσεων με αλανίνη. Στην τέταρτη και πέμπτη στήλη φαίνονται το μοριακό βάρος ανά αλυσίδα του μορίου και η διακριτικότητα των κρυσταλλογραφικά προσδιορισμένων δομών, αντίστοιχα.

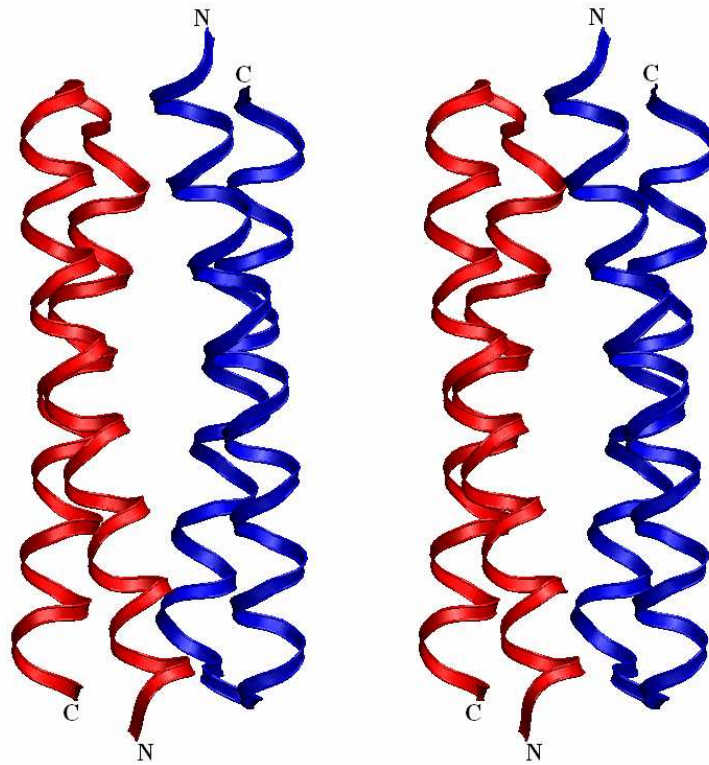


Εικόνα 3-2.α: Άποψη του πυρήνα της Rop. Τα κατάλοιπα στις θέσεις “a” (με μπλε) και “d” (με κόκκινο) σχηματίζουν τον υδρόφοβο πυρήνα και αντιστοιχούν στα κατάλοιπα που αντικαθίστανται στην ανασχεδιασμένη πρωτεΐνη Ala₂Ile₂-6. Φαίνονται δυο επίπεδα του πυρήνα. Με 1 και 2 ορίζονται οι έλικες του A μονομερούς και με 1’ και 2’ οι έλικες του B μονομερούς. Τα βέλη δείχνουν την κατεύθυνση της πολυπεπτιδικής αλυσίδας από το N-τελικό προς το C-τελικό άκρο (αναπαράγεται άνευ αδείας από Willis et al., 2000).

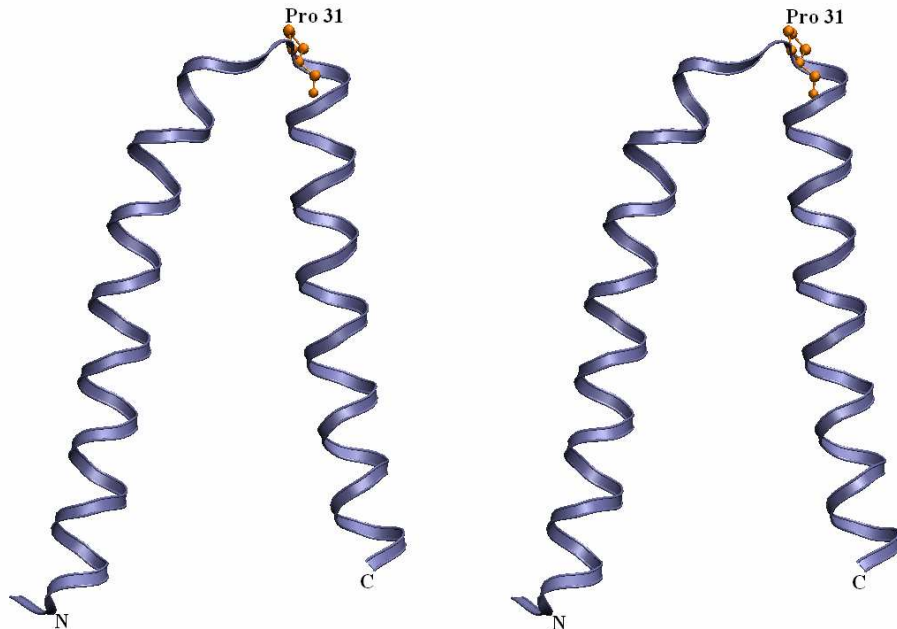
		1		2		3
	1	0		0		0
Rop	MTKQ BE KT A LN MARFI RSQ TLTLL EK LNE LD					
Ala Ile -6	GTKQ E KT I LN MARFI RSQ ALT I LEK A NE LD					
	3	4		5		6
	1	0		0		0
Rop	A DE Q AD IC ES L HD H ADE LYR S CLAR FGDDG EN L					
Ala Ile -6	A DE I AD IA ES I HD H ADE I YRS AL AR FG DDG EN L					

Εικόνα 3-2.β: Στοίχιση των αλληλουχιών της φυσικής Rop και της Ala₂Ile₂-6 με τον αριθμό των καταλοίπων τοποθετημένο από πάνω ανά δέκα κατάλοιπα. Τα κατάλοιπα στις θέσεις “a” και “d” χρωματίζονται ανάλογα με το διάγραμμα. Για την δημιουργία της Ala₂Ile₂-6, τα κατάλοιπα στις θέσεις “a” και “d” της Rop άλλαξαν σε αλανίνη και ισολευκίνη αντίστοιχα. Το τελευταίο στρώμα σε κάθε άκρο του τέσσερα ελικοειδούς δεματίου, που περιλαμβάνει τα κατάλοιπα 5, 29, 31 και 56, δεν άλλαξε. Το κατάλοιπο 56, στη θέση “e” λειτουργεί σαν “d” κατάλοιπο πακετάροντας την πλευρική του αλυσίδα στην κατάλληλη θέση στον πυρήνα.

Η δομή της Ala₂Ile₂-6, που περιλαμβάνει μόνο ένα διμερές στην ασύμμετρη μονάδα, αποκαλύπτει μια δραματική αλλαγή στην τοπολογία απ’ ότι εκείνη της φυσικής Rop, ή της παραλλαγής της Rop που προέκυψε από μια σημειακή μετάλλαξη (A31P: αλλαγή της αλανίνης στη θέση 31 σε προλίνη) και απέκτησε τελείως διαφορετική αναδίπλωση (Εικόνες 3-3, 3-4).

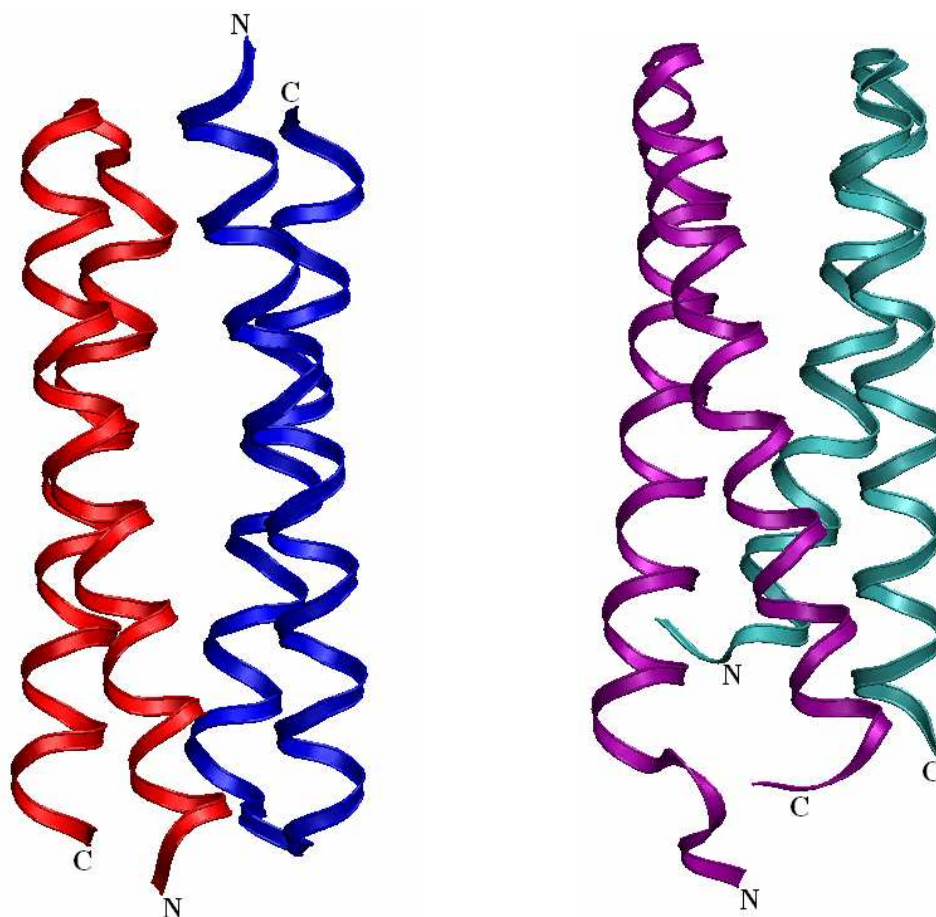


Εικόνα 3-3: Σχηματικό στερεοδιάγραμμα της φυσικής Rop (1ROP.pdb). Με κόκκινο χρώμα βλέπουμε την αλυσίδα A, ενώ με μπλε την αλυσίδα B. Σημειώνονται επίσης και τα N-τελικά και C-τελικά άκρα.



Εικόνα 3-4: Σχηματικό στερεοδιάγραμμα του μονομερούς της A31P που προέκυψε με χρήση των συντεταγμένων της δομής 1B6Q που πήραμε από τη βάση δεδομένων της PDB (Protein Data Bank). Με πορτοκαλί χρώμα βλέπουμε την προλίνη στη θέση 31 (σε CPK). Σημειώνονται επίσης το N-τελικό και το C-τελικό άκρο.

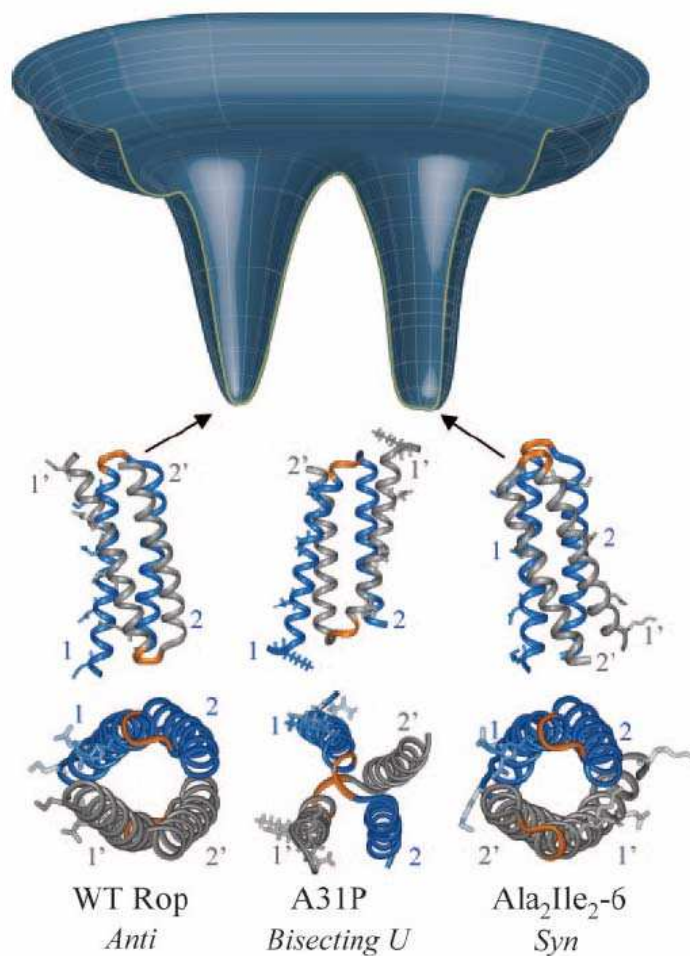
Η Ala_2Ple_2-6 αποτελείται από δυο μονομερή με τη μορφή έλικα-στροφής-έλικα, που σχηματίζουν ένα τέσσερα-ελικοειδές δεμάτιο. Η κοινή επιφάνεια του διμερούς, παρόλα αυτά, μετασχηματίζεται από μια περιστροφή κατά 180° του ενός μονομερούς γύρω από έναν άξονα κάθετο προς την κοινή επιφάνεια. Αυτή η αναπροσαρμογή μεταβάλλει της αναδίπλωση της πρωτεΐνης, τοποθετώντας τις δυο στροφές της Ala_2Ple_2-6 στη μια πλευρά του τέσσερα-ελικοειδούς δεματίου και τα N και C-τελικά άκρα στην άλλη πλευρά (**Εικόνα 3-5**). Αυτή η δομική αναστροφή εξηγεί επίσης και την απώλεια της δραστηριότητας πρόσδεσης RNA.



Εικόνα 3-5: Σε αντιπαράθεση οι δομές της φυσικής Rop (αριστερά, με κόκκινο η αλυσίδα A, με μπλε η αλυσίδα B) και στις Ala_2Ple_2-6 (δεξιά, με μοβ η αλυσίδα A και με κριανό η αλυσίδα B).

Είναι γεγονός ότι οι παραλλαγές τις Rop και οι διαφορετικές διαμορφώσεις και ιδιότητες που έχει η καθεμία, έχουν γίνει αντικείμενο αρκετών (πολλές φορές αντικρουόμενων) μελετών. Σε μια πρόσφατη έρευνα έγινε η υπόθεση (Levy *et al.*, 2005) ότι το αίνιγμα της Rop μπορεί να διαλευκανθεί με την παραδοχή της ύπαρξης ενός υποθετικού διπλού ενεργειακού πεδίου για την πρωτεΐνη Rop. Οι παραλλαγές της Rop, οι οποίες εμφανίζουν διαφορετικές κινητικές αναδίπλωσης, υιοθετούν τοπολογικά διαφορετικές δομές με ελάχιστη ενέργεια ή τουλάχιστον έχουν τοπολογίες που λειτουργούν ως παγίδες. Εξαιτίας των συμμετρικών αλληλουχιών

αυτών των μορίων, οι anti και syn τοπολογίες εμφανίζουν παρόμοιο πακετάρισμα του πυρήνα. Διαφέρουν κυρίως στη σχετική θέση της στροφής. Η τοπολογία U της A31P δε φαίνεται να ανταγωνίζεται ιδιαίτερα τις τοπολογίες anti και syn εξαιτίας του ελάχιστου πακεταρίσματος και των ιδιοτήτων που έχει ως διάχυτο σφαιρίδιο (molten globule). Παρόλα αυτά μπορεί να λειτουργήσει σαν ένα υψηλής ενέργειας ενδιάμεσο στην αναδίπλωση άλλων τοπολογιών. Έτσι υπέθεσαν ότι για τις ανασχεδιασμένες αλληλουχίες της Rop, κυρίως πακετάροντας εκ νέου τα άκρα του πυρήνα, μπορεί να συμβεί μια μετάβαση της δομής της Rop από την anti τοπολογία στη syn (Εικόνα 3-6).



Εικόνα 3-6: Σχηματικό διάγραμμα του διπλού ενεργειακού πεδίου του διμερούς της Rop και στις τρεις τοπολογίες που αντιστοιχούν στη φυσική (WT) Rop, και τις παραλλαγές A31P και Ala₂Ile₂-6. Στις δομές αυτές το ένα μονομερές χρωματίζεται γκρι και το άλλο μπλε. Η θηλιά ανάμεσα στις δυο έλικες χρωματίζεται πορτοκαλί(αναπαράγεται άνευ αδείας από Levy et al., 2005) .

3.2 Γενική προετοιμασία

Με βάση όλα τα προηγούμενα αποφασίσαμε να μελετήσουμε την πρωτεΐνη Ala_2Pe_2-6 πραγματοποιώντας μια σειρά από προσομοιώσεις. Οι προσομοιώσεις αυτές γίνονται με το πρόγραμμα NAMD, το οποίο όπως είδαμε και στο κεφάλαιο 2 χρησιμοποιεί το CHARMM27 δυναμικό πεδίο και περιοδικές οριακές συνθήκες (Periodic Boundary Conditions, PBC) σε τρεις διαστάσεις. Επίσης περιλαμβάνει τον αλγόριθμο Particle Mesh Ewald (PME). Πρόκειται για μια πλήρη αποτελεσματική ηλεκτροστατική μέθοδο που χρησιμοποιείται σε συνδυασμό με τις PBC. Για την πραγματοποίηση αυτών των προσομοιώσεων είναι απαραίτητη η προετοιμασία ενός συστήματος, το οποίο θα περιλαμβάνει την πρωτεΐνη μέσα σε ένα πλαίσιο γεμάτο μόρια νερού, στην προσπάθεια να αποδώσουμε όσο καλύτερα μπορούμε τις συνθήκες που επικρατούν *in vivo*.

Το νερό είναι ο πιο σημαντικός διαλύτης στη φύση. Οι ασυνήθιστες ιδιότητές του προκύπτουν τόσο από το μικρό μέγεθος του μορίου, όσο και από την ικανότητα των μορίων του νερού να σχηματίζουν στο χώρο ένα δίκτυο δεσμών υδρογόνου. Η αναπαραγωγή των ιδιοτήτων του νερού στις προσομοιώσεις είναι δύσκολη – μέχρι τώρα έχουν χρησιμοποιηθεί πολλά διαφορετικά μοντέλα για τη δυναμική ενέργεια των αλληλεπιδράσεων. Τα μοντέλα αυτά μας επιτρέπουν να αναπαράγουμε αρκετές από τις φυσικές ιδιότητες του νερού όπως για παράδειγμα την πυκνότητα, τη διηλεκτρική σταθερά, τη συμπίεση, το συντελεστή διάχυσης, καθώς και πολλές άλλες, συμπεριλαμβανομένης της θερμοκρασίας ή της εξάρτησης αυτών των ποσοτήτων από την πίεση. Τα μοντέλα αυτά χρησιμοποιήθηκαν επίσης για τη μελέτη πιο περίπλοκων διαδικασιών, όπως η υδροφοβική ενυδάτωση ή η επιρροή του νερού στη διαμόρφωση των βιομορίων στο διάλυμα. Σε τέτοιου είδους συστήματα, οι αλληλεπιδράσεις ανάμεσα στο διαλύτη και τη διαλυτή ουσία διαδραματίζουν πολύ σημαντικό ρόλο. Η τελική μορφή ενός μορίου μέσα στο διάλυμα είναι το αποτέλεσμα της ισορροπίας μεταξύ δυο διεργασιών: των αλληλεπιδράσεων μεταξύ διαλύτη – διαλυτής ουσίας (συμπεριλαμβανομένων και των υδρόφοβων αλληλεπιδράσεων) και της απώλειας της δομικής εντροπίας της πολυπεπτιδικής αλυσίδας. Συμπεραίνουμε λοιπόν ότι, αφού τα αποτελέσματα των προσομοιώσεων εξαρτώνται από το μοντέλο που θα χρησιμοποιηθεί, έτσι και η δομή των μορίων του νερού (σε υγρή μορφή) που χρησιμοποιούνται, εξαρτάται από τον τύπο του νερού που θα χρησιμοποιήσουμε.

Υπάρχουν διάφοροι τύποι νερού, όπως για παράδειγμα τα SPC, SPCE, TIP4P, TIP5P και TIP3. Κάθε τύπος έχει διαφορετικές ιδιότητες και χρησιμοποιείται από διαφορετικά προγράμματα ανάλογα με τις απαιτήσεις του καθενός και ανάλογα με τη θερμοκρασία που εφαρμόζουμε στο σύστημα που μελετάμε. Παρόλο που απ' όλους τους τύπους νερού που γνωρίζουμε, τα SPC και TIP4P αναπαράγουν τις δομικές του νερού με τον καλύτερο τρόπο (Zielkiewicz, J., 2005), στην δική μας περίπτωση θα

χρησιμοποιήσουμε νερά τύπου TIP3, μιας και αυτός είναι ο τύπος νερού με βάση τον οποίο παραμετροποιήθηκε το CHARMM27. Το CHARMM χρησιμοποιεί ένα επιπλέον δεσμό H-H, οπότε πρέπει να αποφύγουμε τη δημιουργία γωνιών και διεδρών για τα τμήματα που περιλαμβάνουν μόρια νερού. Έτσι ο τύπος νερού που χρησιμοποιεί είναι τα TIP3. Επιπλέον τα TIP3 σχεδιάστηκαν έτσι ώστε να αναπαράγουν τις ιδιότητες του νερού στην “υγρή” του μορφή και ιδιαίτερα στη θερμοκρασία που εφαρμόζουμε στο σύστημα.

Το πλαίσιο με τα νερά δημιουργείται από το πρόγραμμα VMD με την εντολή `solvate`. Η ολοκληρωμένη εντολή έχει ως εξής:

```
solvate psfgen.psf psfgen.pdb -b 1.80 -t 15.0
```

Τα `psfgen.psf` και `psfgen.pdb` είναι τα αρχεία που θα χρησιμοποιήσει το πρόγραμμα. Η επιλογή `-b` αφαιρεί όσα μόρια νερού βρίσκονται σε απόσταση μικρότερη των 1.80 Å από το μόριο της πρωτεΐνης. Η επιλογή `-t` δημιουργεί τις διαστάσεις του πλαισίου με τα νερά έτσι ώστε να υπάρχει ένα στρώμα νερού της τάξης των 15.0 Å από το άτομο με τη μεγαλύτερη συντεταγμένη μέχρι το άκρο του πλαισίου προς οποιαδήποτε κατεύθυνση. Προκύπτει λοιπόν, μια ελάχιστη απόσταση της τάξης των 30 Å μεταξύ διαδοχικών ειδώλων.

Ο αλγόριθμος PME που χρησιμοποιείται στις προσομοιώσεις απαιτεί το συνολικό φορτίο του συστήματος να ισούται με το μηδέν. Η εφαρμογή `autoionize` του προγράμματος VMD παρέχει μια γρήγορη και εύκολη μέθοδο για να γίνει το συνολικό φορτίο του συστήματος ίσο με το μηδέν, προσθέτοντας ιόντα νατρίου και χλωρίου στο διάλυμα. Η εντολή που χρησιμοποιεί το πρόγραμμα έχει ως εξής:

```
autoionize -psf hydrated.psf -pdb hydrated.pdb -is 0.50
```

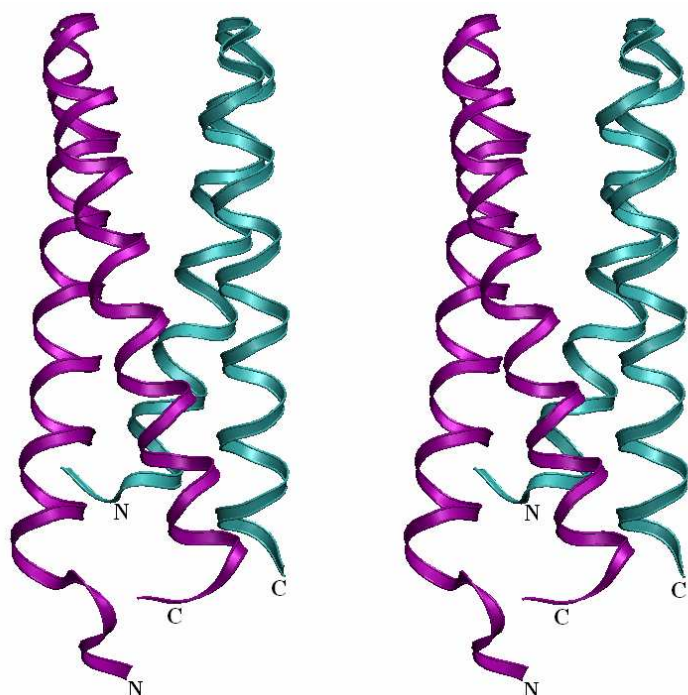
Με τον τρόπο αυτό υπολογίζεται ο αριθμός των ιόντων νατρίου και χλωρίου ώστε το συνολικό φορτίο να ισούται με το μηδέν. Τα αρχεία `hydrated.psf` και `hydrated.pdb` είναι αυτά που θα χρησιμοποιήσει το πρόγραμμα. Η επιλογή `-is` καθιστά τη μέση ιοντική ισχύ του διαλύματος ίση με 50mM.

Η τοποθέτηση των ιόντων γίνεται κατά τυχαίο τρόπο, παρόλα αυτά όμως, υπάρχει ένα ελάχιστο στις αποστάσεις τόσο ανάμεσα στα ιόντα και το μόριο, όσο και ανάμεσα στα ίδια τα ιόντα. Θα μπορούσαμε να μην τοποθετήσουμε τα ιόντα τυχαία, αλλά να βρούμε τα σημεία στα οποία το ηλεκτροστατικό δυναμικό είναι ελάχιστο και να τα τοποθετήσουμε μόνοι μας σε κατάλληλη απόσταση. Βέβαια αυτό δε μας εγγυάται το επιθυμητό αποτέλεσμα και θα ήταν αρκετά περίπλοκο και χρονοβόρο, οπότε προτιμούμε την αυτόματη τοποθέτηση των ιόντων.

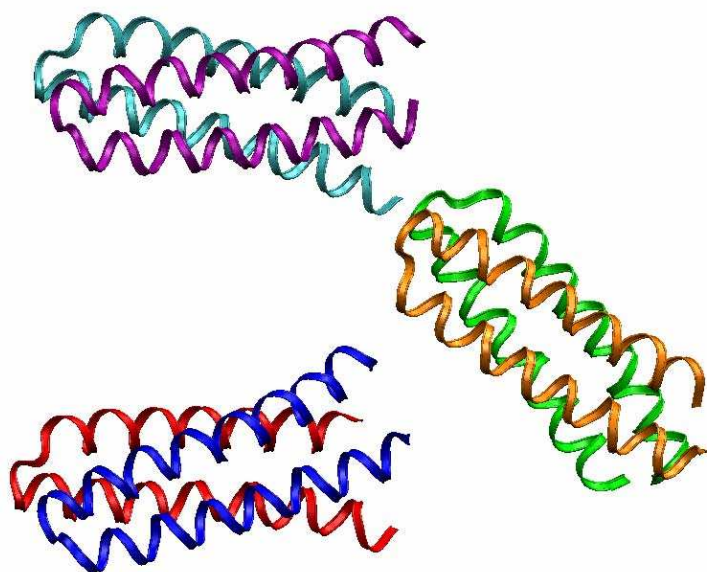
Παρακάτω θα δούμε αναλυτικά τη διαδικασία που ακολουθήθηκε για την προετοιμασία κάθε συστήματος.

3.3 Προετοιμασία διμερούς της Ala_2Pe_2-6 με τοπολογία *syn*

Το πρώτο βήμα ήταν να βρούμε τις συντεταγμένες του μορίου της πρωτεΐνης. Για το λόγο αυτό ανατρέξαμε στη βάση δεδομένων της Protein Data Bank (PDB), όπου βρήκαμε δύο δομές που αντιστοιχούσαν στην Ala_2Pe_2-6 . Πρόκειται για τις 1F4N (Εικόνα 3-7) και 1F4M (Εικόνα 3-8). Η πρώτη περιλαμβάνει ένα τέσσερα-ελικοειδές δεμάτιο σε διακριτικότητα 1.9 Å. Η δεύτερη περιλαμβάνει τρία τέσσερα-ελικοειδή δεμάτια σε διακριτικότητα 2.25 Å.



Εικόνα 3-7: Σχηματικό στερεοδιάγραμμα του 4α-ελικοειδούς δεματίου της δομής 1F4N σε διακριτικότητα 1.9 Å. Με μοβ φαίνεται η αλυσίδα A και με κωινό η αλυσίδα B (σημειώνονται και τα N-τελικά και C-τελικά άκρα).



Εικόνα 3-8: Τρία τέσσερα-ελικοειδή δεμάτια της δομής 1F4M σε διακριτικότητα 2.25 Å. Με μοβ φαίνεται η αλυσίδα A, με κωινό η αλυσίδα B, με κόκκινο η αλυσίδα C, με μπλε η αλυσίδα D, με πορτοκαλί η αλυσίδα E και με πράσινο η αλυσίδα F.

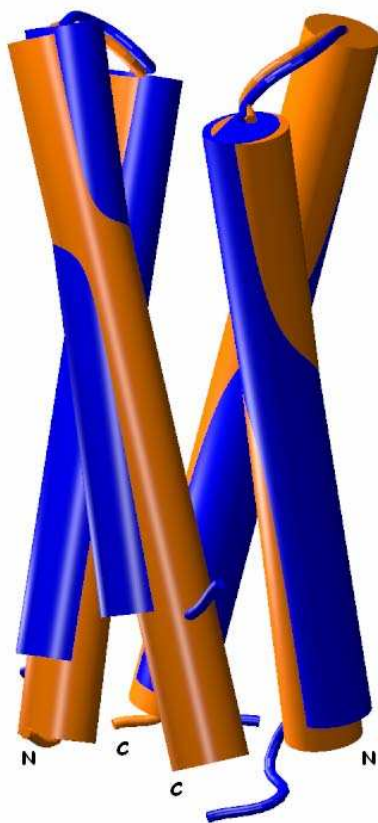
Αξιοσημείωτο είναι το γεγονός ότι κάθε δομή περιείχε διαφορετικό αριθμό καταλοίπων ανά δεμάτιο (Πίνακας 3-3).

Δομή	Αλυσίδες	Κατάλοιπα
1F4N	A	GLY1 THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57 ASP58 ASP59 GLY60
	B	GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55
1F4M	A	GLY1 THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56
	B	THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57
	C	THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55
	D	GLY1 THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57
	E	GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57
	F	GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57

Πίνακας 3-3: Βλέπουμε τις αλληλουχίες όλων των αλυσίδων των δύο δομών, της 1F4N (διακριτικότητα 1.9 Å) και της 1F4M (διακριτικότητα 2.25 Å). Με κίτρινο χρώμα σκιάζονται τα κατάλοιπα που δεν εμφανίζονται σε όλες τις αλυσίδες.

Το γεγονός αυτό μας οδήγησε στο εξής ερώτημα: γιατί σε υψηλότερη διακριτικότητα δεν εμφανίζονται στην κρυσταλλική δομή όλα τα κατάλοιπα της πρωτεΐνης; Έτσι αποφασίσαμε να δημιουργήσουμε μια καινούρια “δομή” η οποία θα συνδύαζε την υψηλότερη διακριτικότητα και το μεγαλύτερο αριθμό καταλοίπων. Για το λόγο αυτό, αρχικά, χρησιμοποιώντας το πρόγραμμα VMD, κάναμε υπέρθεση όλων των τέσσερα-ελικοειδών δεματίων, ανά δυο, για να ελέγξουμε κατά πόσο είναι όμοια (από άποψη πλευρικών αλυσίδων των καταλοίπων αφού τα κατάλοιπα είναι ίδια σε όλα τα δεμάτια). Παρατηρήσαμε λοιπόν, ότι δεν υπήρχε κάποια διαφορά στη διαμόρφωση των πλευρικών αλυσίδων των καταλοίπων. Έτσι προχωρήσαμε στη δημιουργία του μορίου.

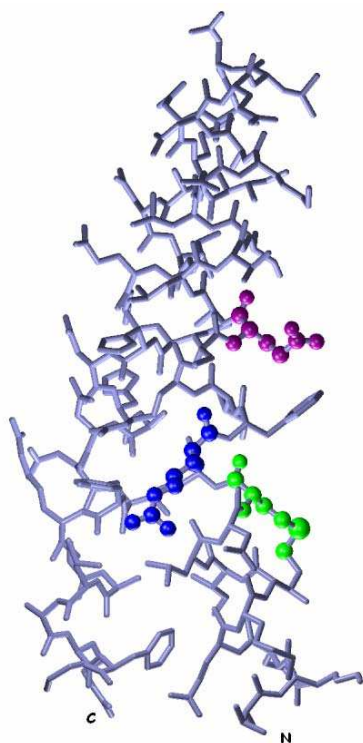
Χρησιμοποιήσαμε το δεμάτιο AB από τη δομή 1F4N και το δεμάτιο AB από τη δομή 1F4M. Με τη βοήθεια και πάλι του προγράμματος VMD κάναμε υπέρθεση των δυο μορίων (Εικόνα 3-9) και στη συνέχεια σώσαμε ένα αρχείο pdb το οποίο περιείχε τις συντεταγμένες και των δυο μορίων.



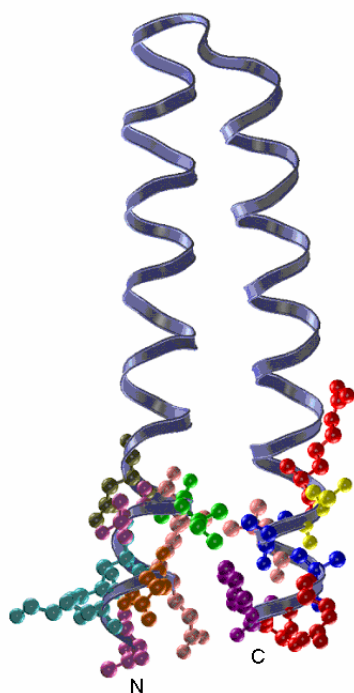
Εικόνα 3-9: Σχηματικό διάγραμμα των διμερών 1F4N (με μπλε χρώμα) και 1F4M (με πορτοκαλί χρώμα) σε υπέρθεση. Είναι εμφανής η διαφορά του αριθμού των καταλοίπων στα άκρα των αλυσίδων.

Ύστερα από επεξεργασία του pdb αρχείου και αφού κρατήσαμε τις συντεταγμένες μόνο των B αλυσίδων, καταφέραμε να προσθέσουμε στην αλυσίδα B της δομής 1F4N τα κατάλοιπα που έλειπαν, χρησιμοποιώντας τα αντίστοιχα από την αλυσίδα B της δομής 1F4M. Πιο συγκεκριμένα, προσθέσαμε τα κατάλοιπα 2-10 στο N-τελικό άκρο και 50-57 στο C-τελικό άκρο, παρόλο που από τη δομή έλειπαν μόνο τα κατάλοιπα 2-4 και 56-57. Αυτό έγινε για να “σχηματιστεί” και το πρώτο επίπεδο του υδρόφοβου πυρήνα, ένα από τα κατάλοιπα του οποίου (το κατάλοιπο της θέσης 56) δεν υπήρχε στο μονομερές B της 1F4N δομής, διατηρώντας τις απαραίτητες

αλληλεπιδράσεις με τα γειτονικά άτομα. Έπειτα, χρησιμοποιώντας το πρόγραμμα X-fit, αλλάξαμε τα στροφομερή τριών καταλοίπων (Met11, Arg13 και Gln18) (Εικόνα 3-10) ώστε οι πλευρικές αλυσίδες τους να μην παρεμποδίζουν τα υπόλοιπα κατάλοιπα, οπότε προέκυψε η ολοκληρωμένη αλυσίδα Β (Εικόνα 3-11). Τέλος, προσθέσαμε και τις συντεταγμένες της αλυσίδας Α από τη δομή 1F4N ώστε να σχηματιστεί το διμερές, οπότε και προκύπτει ο ακόλουθος πίνακας (Πίνακας 3-4). (Το νέο διμερές θα ονομάζεται Ala₂Pe₂-6).



Εικόνα 3-10: Σχηματικό διάγραμμα της αλυσίδας Β της Ala₂Pe₂-6. Με CPK μπορούμε να δούμε τα κατάλοιπα των οποίων αλλάξαμε τα στροφομερή. Με πράσινο φαίνεται η Met11, με μπλε η Arg13 και με μοβ η Gln18.

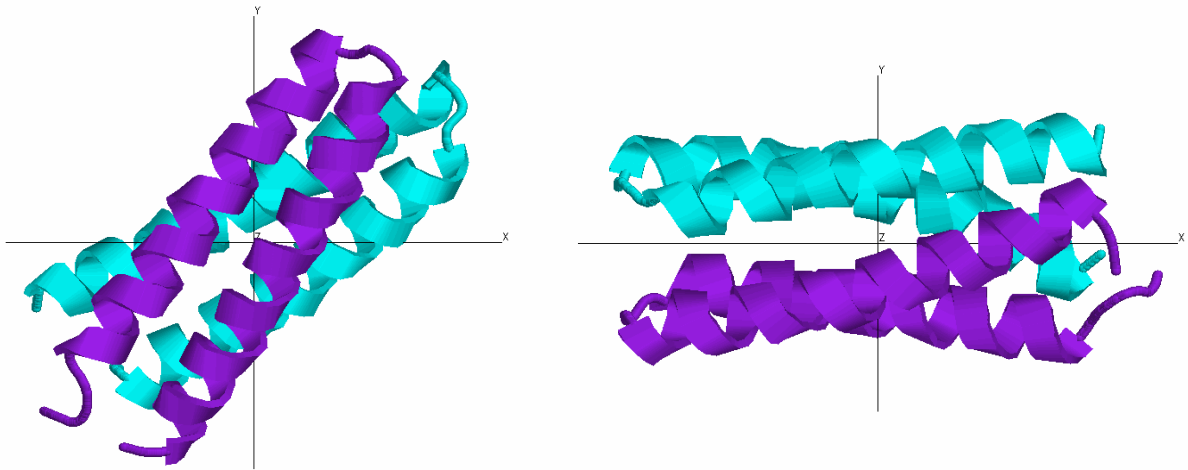


Εικόνα 3-11: Σχηματικό διάγραμμα της αλυσίδας Β της Ala₂Pe₂-6. Με CPK μπορούμε να δούμε τα κατάλοιπα που προσθέσαμε στο N-τελικό και C-τελικό άκρο της αλυσίδας.

Δομή	Αλυσίδες	Κατάλοιπα
Ala₂Ile₂-6	A	GLY1 THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57 ASP58 ASP59 GLY60
	B	THR2 LYS3 GLN4 GLU5 LYS6 THR7 ILE8 LEU9 ASN10 MET11 ALA12 ARG13 PHE14 ILE15 ARG16 SER17 GLN18 ALA19 LEU20 THR21 ILE22 LEU23 GLU24 LYS25 ALA26 ASN27 GLU28 LEU29 ASP30 ALA31 ASP32 GLU33 ILE34 ALA35 ASP36 ILE37 ALA38 GLU39 SER40 ILE41 HIS42 ASP43 HIS44 ALA45 ASP46 GLU47 ILE48 TYR49 ARG50 SER51 ALA52 LEU53 ALA54 ARG55 PHE56 GLY57

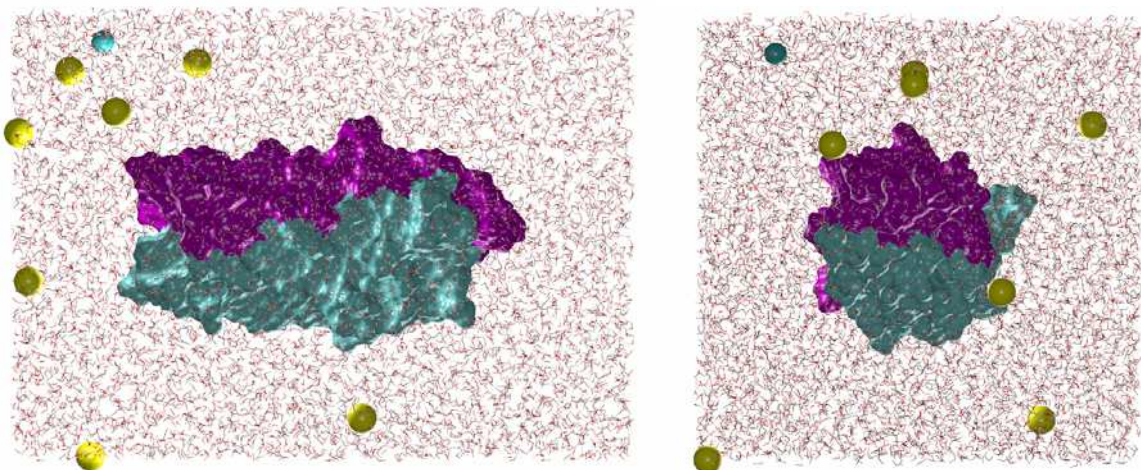
Πίνακας 3-4: Βλέπουμε τις αλληλουχίες των αλυσίδων A και B του μορίου Ala₂Ile₂-6 που κατασκευάσαμε. Με ροζ χρώμα σκιάζονται τα κατάλοιπα εκείνα που προσθέσαμε στην αλυσίδα B.

Έχοντας το μόριο που χρειαζόμαστε, πρέπει να προετοιμάσουμε το σύστημα που απαιτείται για την προσομοίωση. Αρχικά χρειαζόμαστε τα αρχεία που περιέχουν τις παραμέτρους και τις τοπολογίες του δυναμικού πεδίου CHARMM27. Από το pdb αρχείο που διαθέτουμε με κατάλληλη επεξεργασία, κρατάμε μόνο τα στοιχεία εκείνα που αντιστοιχούν στα διακριτικά ATOM και TER, καθώς και το τελικό END. Στο σημείο αυτό πρέπει να προσέξουμε οι αλυσίδες μας να έχουν διαφορετικά διακριτικά, δηλαδή A και B. Έπειτα ευθυγραμμίζουμε τους άξονες της αδράνειας του μορίου με το ορθογώνιο πλαίσιο (**Εικόνα 3-12**). Αυτό επιτυγχάνεται μέσω του προγράμματος moleman, το οποίο διαβάει τις συντεταγμένες του μορίου μας και γράφει ένα νέο αρχείο pdb το aligned.pdb. Στη συνέχεια κάνουμε δυο αντίγραφα του αρχείου aligned.pdb και τα ονομάζουμε A.pdb και B.pdb. Στο πρώτο διατηρούμε τις συντεταγμένες της αλυσίδας A και στο δεύτερο της αλυσίδας B με ένα END στο τέλος κάθε αρχείου. Ακολούθως στα δυο αυτά αρχεία βρίσκουμε και αντικαθιστούμε τα κατάλοιπα HIS με HSP. Πρόκειται για πλήρως πρωτονιωμένα κατάλοιπα ιστιδίνης, τα οποία ανταποκρίνονται καλύτερα στην πραγματοποίηση μιας προσομοίωσης σε pH < 6. (Για μια προσομοίωση σε ουδέτερο pH οι HIS θα έπρεπε να αντικαθίστανται είτε με HSE είτε με HSD ανάλογα με το αν η ιστιδίνη πρωτονιώνεται στο δ ή στο ε άτομο. Αυτό μπορεί να καθοριστεί από το μοτίβο του υδρογονοδεσμού της ιστιδίνης (στο περιεχόμενο της δομής).



Εικόνα 3-12: *Αριστερά:* Η θέση του μορίου σύμφωνα με το ορθοκανονικό σύστημα αξόνων πριν την ευθυγράμμιση των αξόνων της αδράνειας του μορίου. *Δεξιά:* Η θέση του μορίου μετά την ευθυγράμμιση των αξόνων. Με μοβ χρώμα βλέπουμε την αλυσίδα A και με κvano την αλυσίδα B.

Ύστερα, με χρήση ενός άλλου προγράμματος, του psfgen, το οποίο μετά από επεξεργασία της τοπολογίας του δυναμικού πεδίου και των αρχείων A.pdb και B.pdb, μας δίνει τρία νέα αρχεία, το psfgen.pdb, το psfgen.psf (το οποίο είναι απαραίτητο για το πρόγραμμα NAMD που θα χρησιμοποιήσουμε για να κάνουμε την προσομοίωση) και το psfgen.log.



Εικόνα 3-13 : *Αριστερά:* Πλαϊνή όψη του πλαισίου που περιλαμβάνει το μόριο της πρωτεΐνης (με μοβ η αλυσίδα A και με κvano η αλυσίδα B), τα ιόντα (με γαλάζιο τα ιόντα Cl^- και με κίτρινο τα ιόντα Na^+ , σε VDW, με κλίμακα σφαίρας 0.9) και τα μόρια του νερού. *Δεξιά:* Κάτοψη του ίδιου πλαισίου.

Μετά δημιουργούμε ένα script για το πρόγραμμα VMD, το `make_all.vmd`, το οποίο θα μας βοηθήσει να κατασκευάσουμε το πλαίσιο με τα νερά και να τοποθετήσουμε αυτόματα τα ιόντα, ώστε να εξουδετερωθεί το συνολικό φορτίο. Έτσι καταλήγουμε με διάφορα αρχεία, όπως τα `hydrated.psf`, `hydrated.pdb`, `restrain_ca.pdb`, `fix_backbone.pdb`, `ionized.psf` και `ionized.pdb`. Τα δυο τελευταία αντιστοιχούν στο μόριο της πρωτεΐνης μαζί με τα νερά και τα ιόντα. Συγκεκριμένα πρόκειται για ένα ορθογώνιο πλαίσιο με προ-ισορροπημένα μόρια νερού τύπου TIP3, διαστάσεων 80.7 x 58.9 x 58.9 Å³ (**Εικόνα 3-13**). Όλα τα μόρια νερού που βρίσκονταν σε απόσταση μικρότερη από 1.8 Å απομακρύνθηκαν. Το τελικό σύστημα περιλαμβάνει 1820 άτομα πρωτεΐνης, 24667 άτομα νερού, 7 ιόντα Na⁺ (SOD) και 1 ιόν Cl⁻ (CLA).

Το τελευταίο βήμα πριν τη χρήση του προγράμματος NAMD είναι να καθορίσουμε τα όρια του συστήματος (κατά μήκος του ορθογώνιου πλαισίου). Αυτό μπορεί να γίνει γρήγορα με το πρόγραμμα `pdbset` στη μορφή CCP4. Τις συντεταγμένες των ορίων του ορθογώνιου πλαισίου μπορούμε να τις βρούμε στο αρχείο `pdbset.log` που προκύπτει. Για την Ala₂Ile₂-6 αυτές έχουν ως εξής:

	Minimum	Maximum	Centre	Range
On X :	-39.84	40.89	0.53	80.74
On Y :	-29.18	29.69	0.26	58.87
On Z :	-27.91	30.95	1.52	58.85

Όλοι οι υπόλοιποι υπολογισμοί έγιναν με το πρόγραμμα NAMD χρησιμοποιώντας το δυναμικό πεδίο CHARMM27 ως ακολούθως. Αρχικά έγινε ελαχιστοποίηση της ενέργειας του συστήματος για 2000 βήματα (**Παράρτημα 3-1**) και στη συνέχεια εξισορρόπηση της ενέργειας για 0.2 ns υπό συνθήκες NpT, με μια βαθμίδωση ($\Delta = 20$ K) στην αύξηση της θερμοκρασίας από 10 έως 320K για ένα διάστημα 20 ps. Η πίεση διατηρήθηκε σταθερή 1 atm (1.01325 bar). Στο τέλος της εξισορρόπησης της ενέργειας οι διαστάσεις του πλαισίου με τα νερά ήταν 79.3 x 57.8 x 57.8 Å³. Μια NpT προσομοίωση μοριακής δυναμικής πραγματοποιήθηκε για 10 ns (**Παράρτημα 3-2**) χρησιμοποιώντας Nosé-Hoover Langevin dynamics και μεθόδους Langevin piston barostat control (όπως προβλέπεται από το πρόγραμμα NAMD) για τη διατήρηση της πίεσης και της θερμοκρασίας στη 1 atm και στους 320 K αντίστοιχα. Το βήμα ήταν 2 fs, εφαρμόστηκαν Periodic Boundary Conditions, οι πλήρεις ηλεκτροστατικές αλληλεπιδράσεις υπολογίζονταν κάθε τέσσερα βήματα με τη μέθοδο PME (Particle Mesh Ewald), εφαρμόστηκε ένα cutoff για τις van der Waals αλληλεπιδράσεις σε απόσταση 12 Å, και το Shake χρησιμοποιήθηκε για τον περιορισμό όλων των δεσμών που περιελάμβαναν τα άτομα υδρογόνου. Οι τροχιές (trajectories) διαμορφώθηκαν σώζοντας τις ατομικές συντεταγμένες ολόκληρου του συστήματος κάθε 0.4 ps.

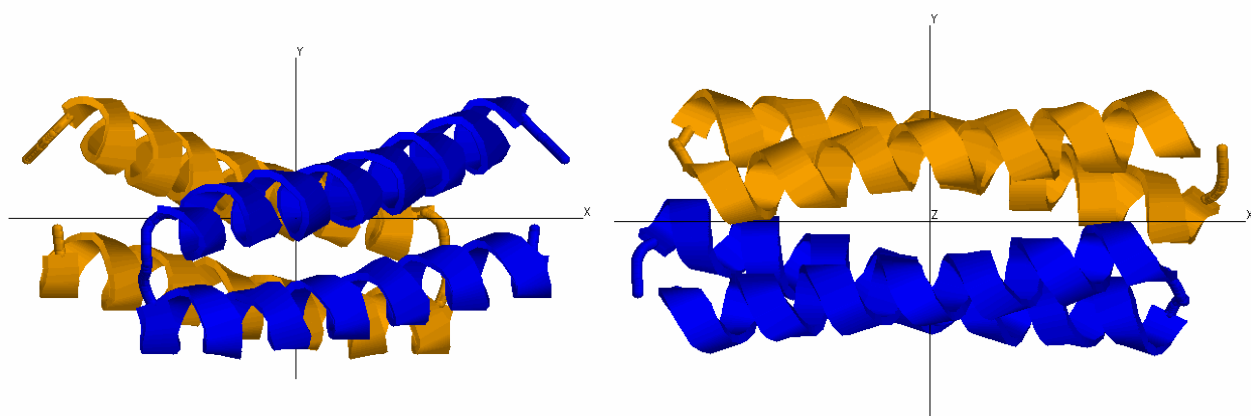
3.4 Προετοιμασία διμερούς της Ala_2Ile_2-6 με τοπολογία anti

Το γεγονός ότι η αναδίπλωση της Ala_2Ile_2-6 είναι τελείως διαφορετική από εκείνη της φυσικής Rop, εξαιτίας της περιστροφής του ενός μονομερούς κατά 180° γύρω από την κοινή επιφάνεια του διμερούς, μας οδήγησε στο εξής ερώτημα: τι θα συνέβαινε στο μόριο εάν διατηρούσε την τοπολογία της φυσικής Rop και την αλληλουχία της Ala_2Ile_2-6 ; Έτσι αποφασίσαμε να κατασκευάσουμε ένα μόριο με την τοπολογία της φυσικής Rop και την αλληλουχία της Ala_2Ile_2-6 , ώστε να μπορέσουμε να μελετήσουμε τους λόγους για τους οποίους η δομή του μορίου δε μοιάζει με εκείνη της φυσικής Rop (Rop-like).

Αυτό μπορεί να γίνει με δυο τρόπους. Ένας τρόπος είναι να πάρουμε το διμερές της Ala_2Ile_2-6 με τη syn τοπολογία και να κάνουμε μια αναστροφή του ενός μονομερούς κατά 180° γύρω από την κοινή επιφάνεια του διμερούς, με άλλα λόγια να αντιστρέψουμε την αρχική αναστροφή. Ένας άλλος τρόπος είναι να πάρουμε τη δομή της φυσικής Rop και χρησιμοποιώντας κάποιο πρόγραμμα να αλλάξουμε τα αμινοξέα που μας ενδιαφέρουν, ώστε να πάρουμε τελικά την αλληλουχία της Ala_2Ile_2-6 . Αποφασίσαμε να ακολουθήσουμε το δεύτερο τρόπο για λόγους ευκολίας.

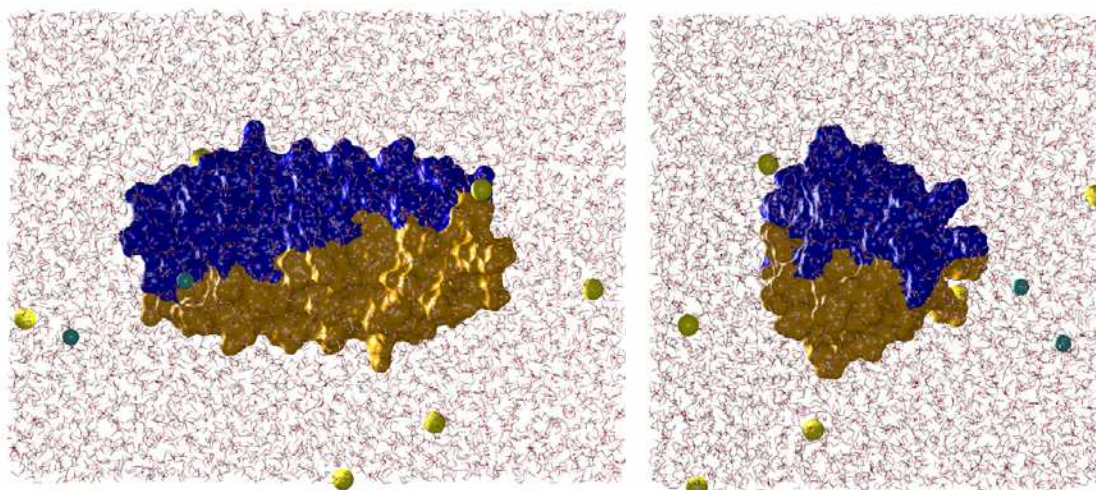
Ανατρέχοντας λοιπόν, στη βάση δεδομένων της Protein Data Bank (PDB), βρήκαμε τις συντεταγμένες της φυσικής Rop (1rop.pdb). Αυτές τις συντεταγμένες θα χρησιμοποιήσουμε σαν εκμαγείο για να κάνουμε τις τροποποιήσεις που είναι απαραίτητες ώστε να πάρουμε την Ala_2Ile_2-6 με τη νέα τοπολογία. Χρησιμοποιώντας το πρόγραμμα X-fit, κάναμε σημειακές μεταλλάξεις στα αμινοξέα των θέσεων “a” και “d” των έξι κεντρικών στρωμάτων του διμερούς. Ύστερα ακολούθησε μια βελτίωση της γεωμετρίας του μορίου ώστε να μην παρεμποδίζει η μια πλευρική αλυσίδα την άλλη. Στη συνέχεια έπρεπε να επιλέξουμε τα κατάλληλα στροφομερή για τις πλευρικές αλυσίδες των καταλοίπων της ισολευκίνης που συμμετείχαν στα έξι κεντρικά στρώματα. Έτσι κάναμε υπέρθεση της δομής που μέχρι τώρα έχουμε φτιάξει, με τη δομή 1F4N της Ala_2Ile_2-6 (μπορούσαμε να χρησιμοποιήσουμε οποιοδήποτε διμερές από τις δομές 1F4N και 1F4M, αφού όπως είδαμε προηγουμένως δεν έχουν καμία διαφορά). Παρατηρήσαμε προσεκτικά τη διεύθυνση των πλευρικών αλυσίδων των καταλοίπων της ισολευκίνης στη δομή 1F4N και χρησιμοποιώντας και πάλι το πρόγραμμα X-fit, αλλάξαμε τα στροφομερή στη δομή που φτιάχνουμε, ώστε να μοιάζουν όσο το δυνατό περισσότερο με εκείνα της κανονικής δομής. Καταλήξαμε λοιπόν, με ένα διμερές το οποίο έχει την τοπολογία της φυσικής Rop (anti) και την αλληλουχία της Ala_2Ile_2-6 .

Σειρά έχει η προετοιμασία του συστήματος, η οποία σε γενικές γραμμές είναι ίδια με εκείνη που περιγράψαμε προηγουμένως για την Ala_2Ile_2-6 με τοπολογία syn. Ένα από τα σημεία εκείνα στα οποία παρουσιάζονται διαφορές αφορά την ευθυγράμμιση των αξόνων της αδράνειας (**Εικόνα 3-14**).



Εικόνα 3-14: *Αριστερά:* Η θέση του μορίου σύμφωνα με το ορθοκανονικό σύστημα αξόνων πριν την ευθυγράμμιση των αξόνων της αδράνειας του μορίου. *Δεξιά:* Η θέση του μορίου μετά την ευθυγράμμιση των αξόνων. Με πορτοκαλί χρώμα βλέπουμε την αλυσίδα A και με μπλε την αλυσίδα B.

Επίσης διαφορά παρουσιάζεται στις διαστάσεις του ορθογώνιου πλαισίου με τα προ-ισοροπημένα μόρια νερού τύπου TIP3, οι οποίες είναι $78.59 \times 60.1 \times 56.85 \text{ \AA}^3$ (**Εικόνα 3-15**). Και εδώ όλα τα μόρια νερού που βρίσκονταν σε απόσταση μικρότερη από 1.8 \AA απομακρύνθηκαν. Το τελικό σύστημα περιλαμβάνει 1802 άτομα πρωτεΐνης, 23571 άτομα νερού, 6 ιόντα Na^+ (SOD) και 2 ιόντα Cl^- (CLA).



Εικόνα 3-15: *Αριστερά:* Πλαϊνή όψη του πλαισίου που περιλαμβάνει το μόριο της πρωτεΐνης (με πορτοκαλί η αλυσίδα A και με μπλε η αλυσίδα B), τα ιόντα (με γαλάζιο τα ιόντα Cl^- και με κίτρινο τα ιόντα Na^+ , σε VDW με κλίμακα σφαίρας 2.9) και τα μόρια του νερού. *Δεξιά:* Κάτοψη του ίδιου πλαισίου.

Το τελευταίο βήμα πριν τη χρήση του προγράμματος NAMD είναι ο καθορισμός των ορίων του συστήματος (κατά μήκος του ορθογώνιου πλαισίου). Αυτό γίνεται εύκολα με το πρόγραμμα `pdbset`. Τις συντεταγμένες των ορίων του ορθογώνιου πλαισίου μπορούμε να τις βρούμε στο αρχείο `pdbset.log` που προκύπτει. Για την Ala_2Ile_2 -6 με τοπολογία anti αυτές έχουν ως εξής:

	Minimum	Maximum	Centre	Range
On X :	-39.29	39.30	0.00	78.59
On Y :	-30.04	30.06	0.01	60.10
On Z :	-26.03	30.82	2.39	56.85

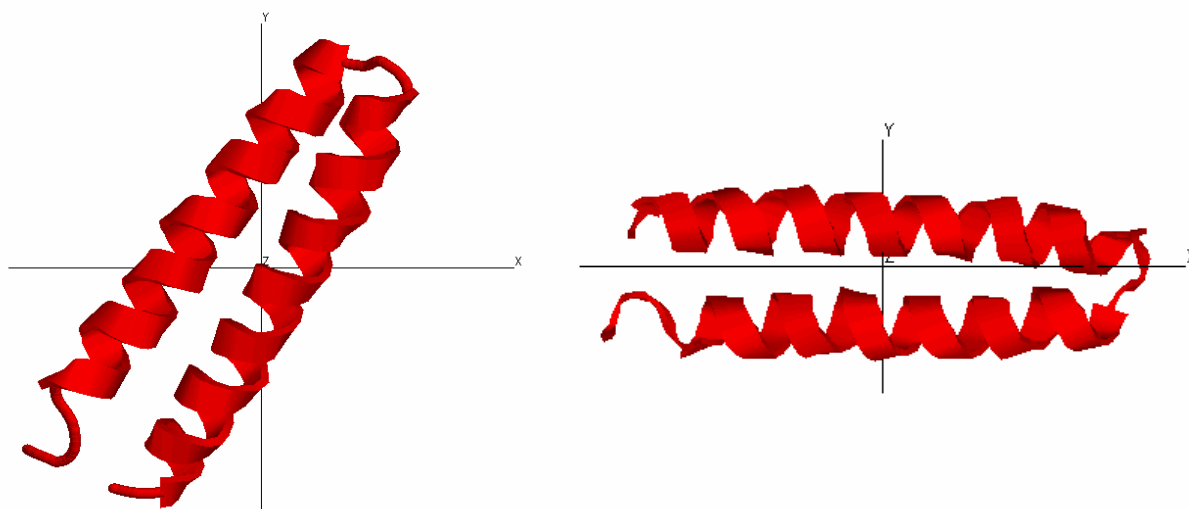
Όλοι οι υπόλοιποι υπολογισμοί έγιναν ομοίως με το πρόγραμμα NAMD χρησιμοποιώντας το δυναμικό πεδίο CHARMM27. Μετά την ελαχιστοποίηση και την εξισορρόπηση της ενέργειας οι διαστάσεις του πλαισίου με τα νερά ήταν $76.97 \times 58.68 \times 55.68 \text{ \AA}^3$. Μια NpT προσομοίωση μοριακής δυναμικής πραγματοποιήθηκε για 10 ns χρησιμοποιώντας Nosé-Hoover Langevin dynamics και μεθόδους Langevin piston barostat control (όπως προβλέπεται από το πρόγραμμα NAMD) για τη διατήρηση της πίεσης και της θερμοκρασίας στη 1 atm και στους 320 K αντίστοιχα. Το βήμα ήταν 2 fs, εφαρμόστηκαν Periodic Boundary Conditions, οι πλήρεις ηλεκτροστατικές αλληλεπιδράσεις υπολογίζονταν κάθε τέσσερα βήματα με τη μέθοδο PME (Particle Mesh Ewald), εφαρμόστηκε ένα cutoff για τις van der Waals αλληλεπιδράσεις σε απόσταση 12 Å, και το Shake χρησιμοποιήθηκε για τον περιορισμό όλων των δεσμών που περιελάμβαναν τα άτομα υδρογόνου. Οι τροχιές (trajectories) εξασφαλίστηκαν σώζοντας τις ατομικές συντεταγμένες ολόκληρου του συστήματος κάθε 0.4 ps.

3.5 Προετοιμασία native μονομερούς της $\text{Ala}_2\text{Pe}_2\text{-6}$

Μια από τις απορίες που μας δημιουργήθηκαν ήταν εάν το διμερές της $\text{Ala}_2\text{Pe}_2\text{-6}$ θα μπορούσε να προέλθει από σύγκρουση προσχηματισμένων μονομερών (δηλαδή μονομερών που ήδη είχαν αναδιπλωθεί ως έλικα-στροφή-έλικα). Έτσι προέκυψε το ερώτημα: είναι ένα προσχηματισμένο μονομερές σταθερό; Με άλλα λόγια θέλουμε να δούμε είναι εάν στο μονομερές μπορεί να δημιουργηθεί υδρόφοβος πυρήνας, ή εάν αυτό θα διαλυθεί μέσα στο πλαίσιο με τα νερά.

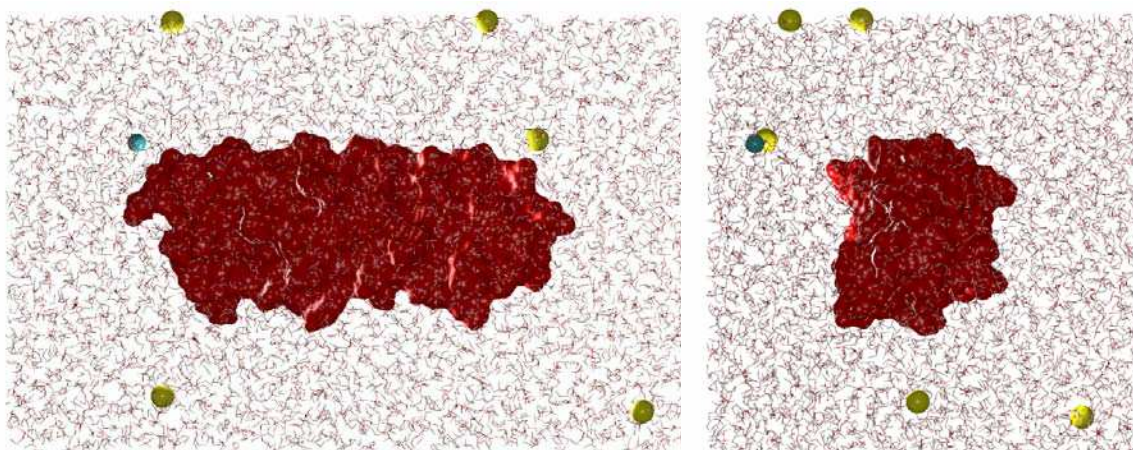
Έχοντας τις συντεταγμένες του μορίου 1F4N από τη βάση δεδομένων της PDB, με κατάλληλη επεξεργασία κρατήσαμε σε ένα αρχείο μόνο τις συντεταγμένες της αλυσίδας A. Οι συντεταγμένες του μονομερούς που προέκυψαν ήταν εκείνες που χρησιμοποιήσαμε για την προσομοίωση.

Ακολούθησε η προετοιμασία του συστήματος με τρόπο όμοιο με εκείνο των προηγούμενων προσομοιώσεων. Εδώ η ευθυγράμμιση των αξόνων της αδράνειας του μορίου με το ορθογώνιο πλαίσιο (**Εικόνα 3-16**) δίνει εικόνα διαφορετική από τις προηγούμενες, αφού έχουμε να κάνουμε πλέον με ένα μονομερές.



Εικόνα 3-16: *Αριστερά:* Η θέση του μορίου σύμφωνα με το ορθοκανονικό σύστημα αξόνων πριν την ευθυγράμμιση των αξόνων της αδράνειας του μορίου. *Δεξιά:* Η θέση του μορίου μετά την ευθυγράμμιση των αξόνων.

Οι διαστάσεις του πλαισίου με τα μόρια νερού τύπου TIP3 αντιστοιχούν σε $80.4 \times 51.3 \times 49.9 \text{ \AA}^3$ (**Εικόνα 3-17**). Όλα τα μόρια νερού που βρίσκονταν σε απόσταση μικρότερη από 1.8 \AA απομακρύνθηκαν. Το τελικό σύστημα περιλαμβάνει 929 άτομα πρωτεΐνης, 18,423 άτομα νερού, 5 ιόντα Na^+ (SOD) και 1 ιόν Cl^- (CLA).



Εικόνα 3-17: *Αριστερά:* Πλαϊνή όψη του πλαισίου που περιλαμβάνει το μόριο του μονομερούς, τα ιόντα (με γαλάζιο τα ιόντα Cl⁻ και με κίτρινο τα ιόντα Na⁺, σε VDW με κλίμακα σφαίρας 2.9) και τα μόρια του νερού. *Δεξιά:* Κάτοψη του ίδιου πλαισίου.

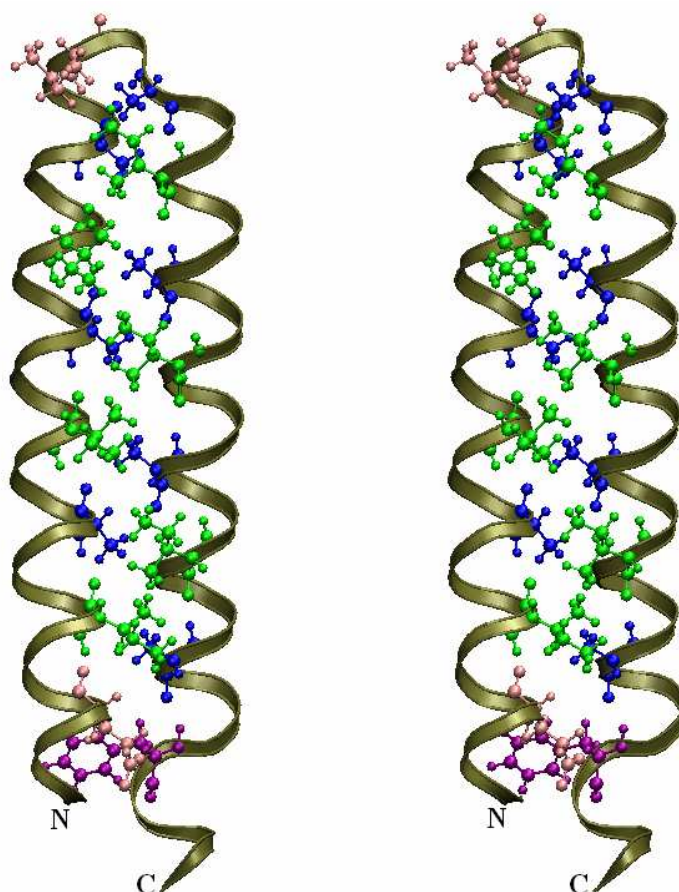
Ακολουθεί ο καθορισμός των ορίων του συστήματος (κατά μήκος του ορθογώνιου πλαισίου). Οι συντεταγμένες των ορίων του ορθογώνιου πλαισίου για το native μονομερές της Ala2Ile2-6 έχουν ως εξής:

	Minimum	Maximum	Centre	Range
On X :	-40.31	40.13	-0.09	80.44
On Y :	-26.57	24.73	-0.92	51.29
On Z :	-24.76	25.11	0.17	49.87

Έπειτα από την ελαχιστοποίηση και την εξισορρόπηση της ενέργειας οι διαστάσεις του πλαισίου με τα νερά ήταν 78.7 x 50.2 x 48.8 Å³. Τέλος πραγματοποιήσαμε μια NpT προσομοίωση μοριακής δυναμικής για 20 ns (ή αλλιώς 10.000.000 βήματα) αυτή τη φορά, με τις ατομικές συντεταγμένες του συστήματος να σώζονται κάθε 0.4 ps

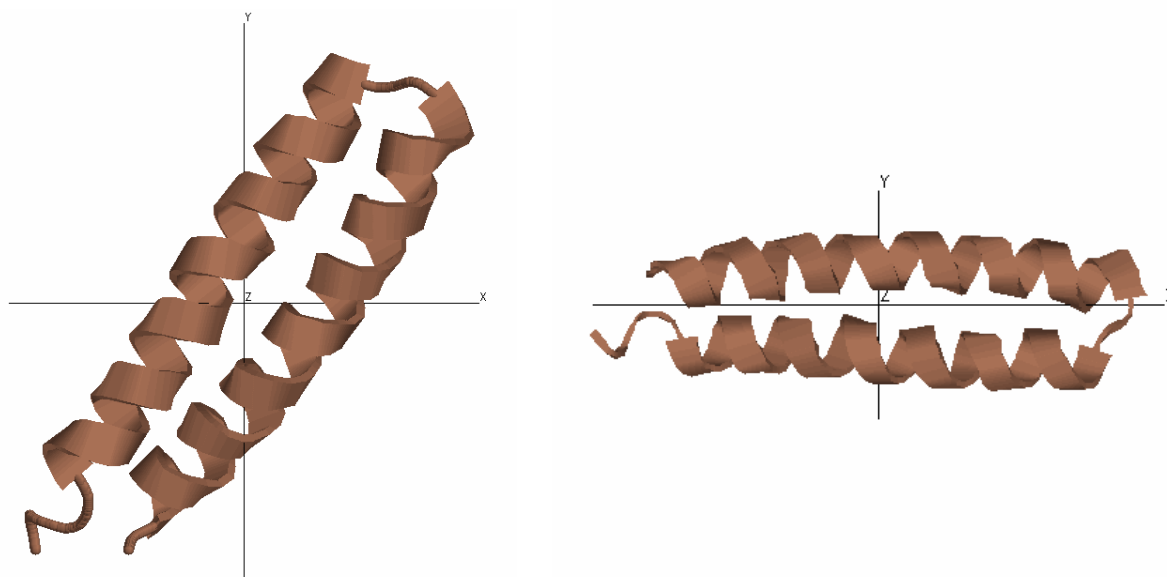
3.6 Προετοιμασία artificial μονομερούς της Ala₂Ile₂-6

Το γεγονός ότι υπάρχει πιθανότητα το μονομερές της Ala₂Ile₂-6 να μη μπορέσει από μόνο του να δημιουργήσει υδρόφοβο πυρήνα, μας ώθησε στο να σκεφτούμε ότι ίσως θα μπορούσαμε να στρέψουμε τις α έλικες προς τα “μέσα” ώστε να δημιουργήσουμε ένα “υπόβαθρο” για το σχηματισμό του πυρήνα. Οι πλευρικές αλυσίδες των α-ελίκων στο μονομερές είναι στραμμένες προς τα έξω, αφού η δομή του είναι ακριβώς όπως και στο διμερές. Αυτό που κάναμε ήταν, με τη χρήση του προγράμματος X-fit, να στρίψουμε τις α-έλικες προς τα μέσα, όσο το δυνατό περισσότερο, ώστε οι πλευρικές αλυσίδες των καταλοίπων που σχηματίζουν τον υδρόφοβο πυρήνα να βρίσκονται στο εσωτερικό ενός υποτιθέμενου πυρήνα (**Εικόνα 3-18**). Στη συνέχεια ελέγξαμε τις πλευρικές αλυσίδες ώστε να μην εμποδίζουν η μια την άλλη και κάναμε μια βελτίωση της γεωμετρίας του μορίου. Έτσι καταλήξαμε με το επιθυμητό μόριο, οπότε είμαστε έτοιμοι να προετοιμάσουμε το σύστημα για την προσομοίωση.



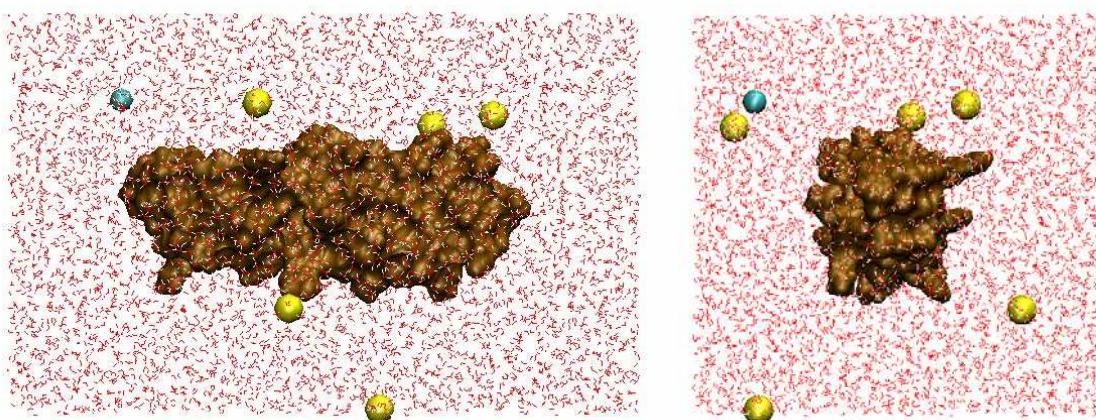
Εικόνα 3-18: Σχηματικό στερεοδιάγραμμα του artificial μονομερούς. Με καφέ χρώμα φαίνεται η αλυσίδα, με μπλε χρώμα μπορούμε να διακρίνουμε τα κατάλοιπα Ala, με πράσινο τα κατάλοιπα Ile, με ροζ τα κατάλοιπα Leu, ενώ με μοβ το κατάλοιπο Phe (σε CPK). Τα κατάλοιπα αυτά είναι εκείνα που σχηματίζουν τον υδρόφοβο πυρήνα. Μετά από κατάλληλη επεξεργασία τα κατάλοιπα βρίσκονται σχεδόν στο εσωτερικό ενός υποθετικού υδρόφοβου πυρήνα.

Η διαδικασία της προετοιμασίας δε διαφέρει ούτε εδώ από εκείνη που είδαμε στις προηγούμενες ενότητες. Στην παρακάτω εικόνα βλέπουμε τη θέση του μορίου πριν και μετά την ευθυγράμμιση των αξόνων της αδράνειας (**Εικόνα 3-17**). Η αλλαγή στη θέση είναι εμφανής.



Εικόνα 3-17: *Αριστερά:* Η θέση του μορίου σύμφωνα με το ορθοκανονικό σύστημα αξόνων πριν την ευθυγράμμιση των αξόνων της αδράνειας του μορίου. *Δεξιά:* Η θέση του μορίου μετά την ευθυγράμμιση των αξόνων.

Ακολουθεί ενυδάτωση του μονομερούς σε ορθογώνιο πλαίσιο με προ-ισοροπημένα μόρια νερού τύπου TIP3, διαστάσεων $79.6 \times 50.6 \times 51.9 \text{ \AA}^3$ (**Εικόνα 3-18**).



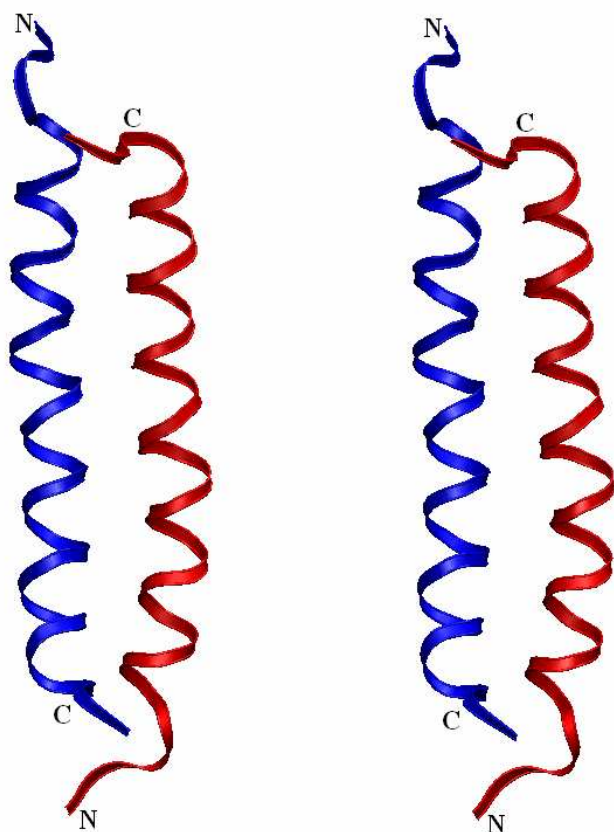
Εικόνα 3-18: *Αριστερά:* Πλαϊνή όψη του πλαισίου που περιλαμβάνει το μόριο του μονομερούς, τα ιόντα (με γαλάζιο τα ιόντα Cl^- και με κίτρινο τα ιόντα Na^+ , σε VDW με κλίμακα σφαίρας 2.9) και τα μόρια του νερού. *Δεξιά:* Κάτοψη του ίδιου πλαισίου.

Όλα τα μόρια νερού που βρίσκονταν σε απόσταση μικρότερη από 1.8 Å απομακρύνθηκαν. Το τελικό σύστημα περιλαμβάνει 929 άτομα πρωτεΐνης, 18804 άτομα νερού, 5 ιόντα Na⁺ (SOD) και 1 ιόν Cl⁻ (CLA). Ο καθορισμός των ορίων του ορθογώνιου πλαισίου για το artificial μονομερές Ala₂He₂-6 δίνει τα ακόλουθα αποτελέσματα:

	Minimum	Maximum	Centre	Range
On X :	-24.41	24.40	0.54	48.81
On Y :	-9.75	9.73	-0.11	19.47
On Z :	-10.23	10.15	1.01	20.38

Ακολουθεί η ελαχιστοποίηση και η εξισορρόπηση της ενέργειας. Στο τέλος και της φάσης εξισορρόπησης οι διαστάσεις του πλαισίου με τα νερά ήταν 78.0 x 49.6 x 50.8 Å³. Η προσομοίωση μοριακής δυναμικής διαρκεί 8 ns, με τις ατομικές συντεταγμένες του συστήματος να σώζονται κάθε 0.4 ps

Στο σημείο αυτό πρέπει να σημειώσουμε ότι χρησιμοποιήσαμε ακόμα ένα τρόπο για να γυρίσουμε τις πλευρικές αλυσίδες των καταλοίπων του πυρήνα προς τα μέσα. Σκεφτήκαμε να κάνουμε υπέρθεση του μονομερούς που διαθέτουμε με ένα εσπειραμένο σπείραμα, να βρούμε σε αυτό τα αντίστοιχα κατάλοιπα στις θέσεις “a” και “d” και με καλούπι αυτό το σπειρωμένο σπείραμα, να στρίψουμε τις πλευρικές αλυσίδες των καταλοίπων.



Εικόνα 3-19: Σχηματικό στερεοδιάγραμμα μιας δομής της πρωτεΐνης Prop, η οποία προκύπτει από την NMR δομή 1R48.pdb. Πρόκειται για ένα αντιπαράλληλο εσπειραμένο σπείραμα. Με μπλε χρώμα φαίνεται η αλυσίδα A και με κόκκινο η αλυσίδα B.

Έτσι ψάχνοντας στη βάση δεδομένων της PDB βρήκαμε την πρωτεΐνη Prop (1R48.pdb) της *E.coli*, από την οποία χρησιμοποιήσαμε την C-τελική κυτταροπλασματική επικράτεια. Πρόκειται για NMR δομή (51 δομές σε ένα αρχείο pdb). Από τις 51 αυτές δομές αποθηκεύσαμε τη μια σε ένα ξεχωριστό αρχείο pdb, το οποίο ονομάσαμε Prop.pdb (**Εικόνα 3-19**).

Έχοντας λοιπόν τα μόρια που χρειαζόμαστε, χρησιμοποιήσαμε το πρόγραμμα VMD για να κάνουμε υπέρθεση των δύο δομών (του σπειρωμένου σπειράματος και των δυο α-ελίκων του μονομερούς Ala2Ile2-6 χωρίς τα κατάλοιπα 28-32 της στροφής) και να βρούμε, μέσω της τυπικής απόκλισης (RMSD), ποιες έλικες ταιριάζουν καλύτερα και σε ποια σημεία (δηλαδή ποιος συνδυασμός έχει το μικρότερο RMSD). Δοκιμάσαμε όλους τους δυνατούς συνδυασμούς (**Πίνακες 3-5 έως 3-8**). Πιο συγκεκριμένα, με κίτρινο χρώμα στους πίνακες βλέπουμε οριζόντια τα τμήματα των α ελίκων της Ala2Ile2-6 και κάθετα της PROP, τα οποία χρησιμοποιήθηκαν για την υπέρθεση. Με γαλάζιο και ροζ βλέπουμε τους συνδυασμούς a-d και d-a αντίστοιχα, των θέσεων a και d του υδρόφοβου πυρήνα και με λευκό τις τιμές του RMSD που προέκυψαν από τους εκάστοτε συνδυασμούς.

RMSD	(Ala2-Ile2) ₆ κατάλοιπα 1-27		(Ala2-Ile2) ₆ κατάλοιπα 33-60	
PROP αλυσίδα A	a – d	5-26	d – a	34-56
	3-24	1.1484	6-28	0.9012
	10-31	1.1091	-	-
PROP αλυσίδα B	a – d	5-26	d – a	34-56
	3-24	1.1536	6-28	0.8997
	10-31	1.0869	-	-

Πίνακας 3-5: Υπολογισμός RMSD για 7 στρώματα του πυρήνα.

RMSD	(Ala2-Ile2) ₆ κατάλοιπα 1-27		(Ala2-Ile2) ₆ κατάλοιπα 33-60	
PROP αλυσίδα A	a – d	5-22	a – d	38-56
	3-20	0.9993	3-21	1.1671
	10-27	0.8307	10-28	0.7799
	d – a	8-26	d – a	34-52
	6-24	0.7762	6-24	0.7002
	13-31	0.9643	13-31	0.7657
PROP αλυσίδα B	a – d	5-22	a – d	38-56
	3-20	1.0128	3-21	1.1876
	10-27	0.7954	10-28	0.7667
	d – a	8-26	d – a	34-52
	6-24	0.7823	6-24	0.7086
	13-31	0.9459	13-31	0.7708

Πίνακας 3-6: Υπολογισμός RMSD για 6 στρώματα του πυρήνα.

RMSD	(Ala ₂ -Ile ₂) ₆ κατάλοιπα 1-27			(Ala ₂ -Ile ₂) ₆ κατάλοιπα 33-60		
PROP αλυσίδα A	a – d	5-19	12-26	a – d	38-52	-
	3-17	0.9680	0.8714	3-17	1.0376	-
	10-24	0.6329	0.5728	10-24	0.5615	-
	17-31	0.8128	0.7832	17-31	0.7745	-
	d – a	8-22	-	d – a	34-48	41-56
	6-20	0.6105	-	6-20	0.4998	0.8909
	13-27	0.7221	-	13-27	0.5240	0.7156
	a – d	5-19	12-26	a – d	38-52	-
PROP αλυσίδα B	3-17	0.9842	0.8775	3-17	1.0608	-
	10-24	0.6251	0.5673	10-24	0.5520	-
	17-31	0.8213	0.7768	17-31	0.7888	-
	d – a	8-22	-	d – a	34-48	41-56
	6-20	0.6297	-	6-20	0.5225	0.9115
	13-27	0.6838	-	13-27	0.5075	0.6996

Πίνακας 3-7: Υπολογισμός RMSD για 5 στρώματα του πυρήνα.

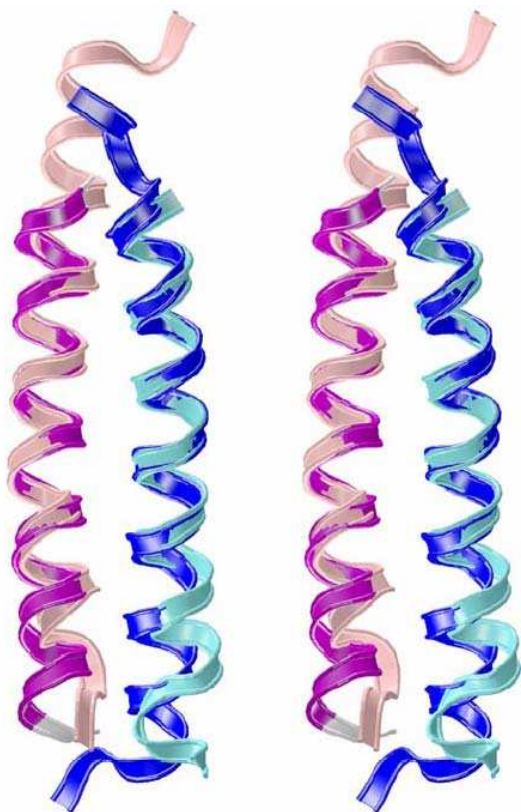
RMSD	(Ala ₂ -Ile ₂) ₆ κατάλοιπα 1-27			(Ala ₂ -Ile ₂) ₆ κατάλοιπα 33-60		
PROP αλυσίδα A	a – d	5-15	12-22	a – d	38-48	45-56
	3-13	0.9027	0.8930	3-14	0.8809	0.9836
	10-20	0.4865	0.3992	10-21	0.3825	0.6623
	17-27	0.4720	0.5531	17-28	0.3830	0.6515
	d – a	8-19	15-26	d – a	34-45	41-52
	6-17	0.5907	0.4526	6-17	0.4564	0.7653
	13-24	0.5052	0.5405	13-24	0.4336	0.4583
	20-31	0.7670	0.7286	20-31	0.6522	0.8017
PROP αλυσίδα B	a – d	5-15	12-22	a – d	38-48	45-56
	3-13	0.9073	0.8900	3-14	0.8868	0.9897
	10-20	0.4959	0.4069	10-21	0.3934	0.6634
	17-27	0.4491	0.5136	17-28	0.3689	0.6441
	d – a	8-19	15-26	d – a	34-45	41-52
	6-17	0.6138	0.4738	6-17	0.4800	0.7918
	13-24	0.4931	0.5360	13-24	0.4300	0.4386
	20-31	0.7795	0.7284	20-31	0.6535	0.8169

Πίνακας 3-8: Υπολογισμός RMSD για 4 στρώματα του πυρήνα.

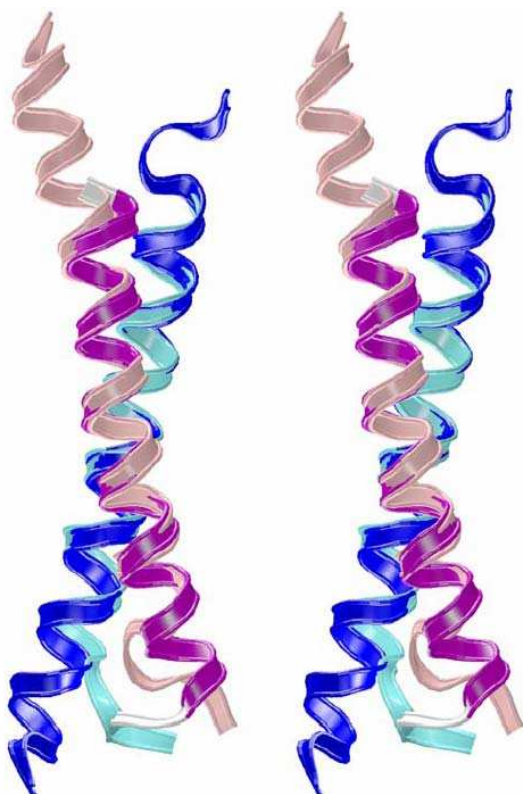
Αφού μελετήσαμε τα αποτελέσματα, καταλήξαμε ότι ταιριάζουν καλύτερα η αλυσίδα A της Prop με τα κατάλοιπα 33-60 του μονομερούς της Ala₂Ile₂-6 και η αλυσίδα B της Prop με τα κατάλοιπα 1-27 του μονομερούς της Ala₂Ile₂-6 (Εικόνες 3-20 έως 3-24). Πιο συγκεκριμένα τα κατάλοιπα 6-20 της αλυσίδας A της Prop με τα κατάλοιπα 34-48 του μονομερούς της Ala₂Ile₂-6 (με μοβ χρώμα στον Πίνακα 3-7)

και τα κατάλοιπα 6-20 της αλυσίδας B της Prop με τα κατάλοιπα 8-22 του μονομερούς της Ala2Ile2-6 (με πορτοκαλί χρώμα στον Πίνακα 3-7).

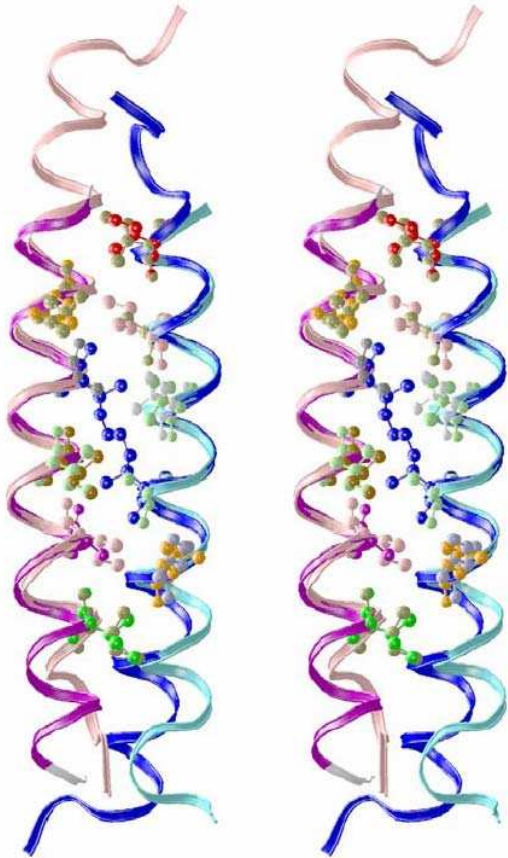
Παρόλο, βέβαια, που η όλη αυτή διαδικασία θα μπορούσε να μας δώσει ένα πιο σωστά δομημένο μονομερές, η διαδικασία ανακατασκευής της στροφής στο μονομερές Ala2Ile2-6 (δηλαδή η επανατοποθέτηση των καταλοίπων 28-32) ήταν αρκετά περίπλοκη και δύσκολη. Συνεπώς η προσπάθεια αυτή δεν κατέληξε κάπου.



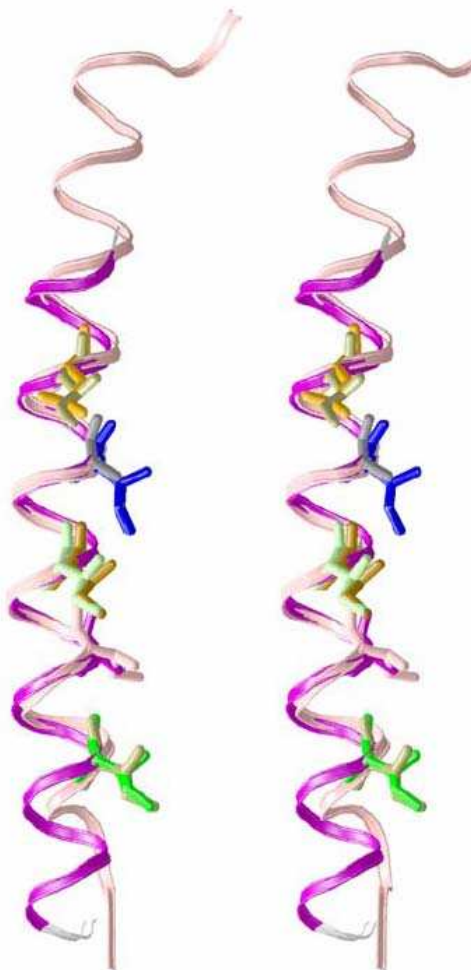
Εικόνα 3-20: Σχηματικό στερεοδιάγραμμα της δομής του μονομερούς της Ala2Ile2-6 σε υπέρθεση με την Prop. Βλέπουμε την αλυσίδα A της Prop με μπλε χρώμα, την αλυσίδα B της Prop με ροζ χρώμα, τα κατάλοιπα 1-27 της Ala2Ile2-6 με μοβ χρώμα και τα κατάλοιπα 33-60 της Ala2Ile2-6 με κίτρινο χρώμα.



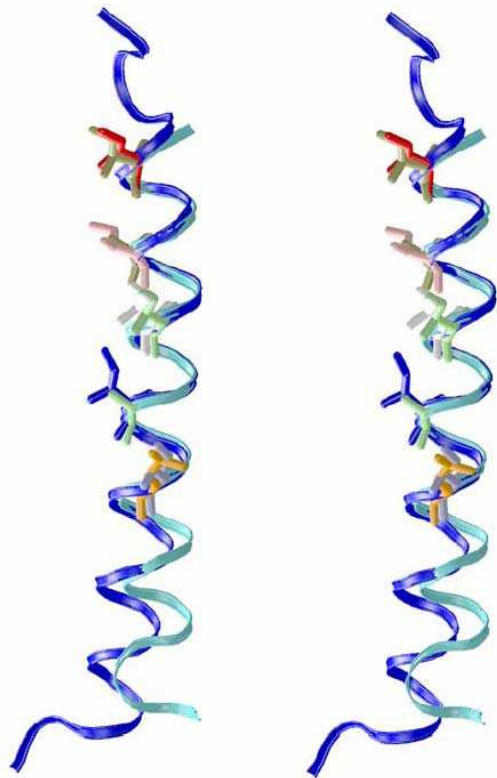
Εικόνα 3-21: Σχηματικό στερεοδιάγραμμα της δομής του μονομερούς της Ala2Ile2-6 σε υπέρθεση με την Prop, σε πλαϊνή όψη. Βλέπουμε την αλυσίδα A της Prop με μπλε χρώμα, την αλυσίδα B της Prop με ροζ χρώμα, τα κατάλοιπα 1-27 της Ala2Ile2-6 με μοβ χρώμα και τα κατάλοιπα 33-60 της Ala2Ile2-6 με κίτρινο χρώμα.



Εικόνα 3-22: Σχηματικό στερεοδιάγραμμα της δομής του μονομερούς της Ala2Ile2-6 σε υπέρθεση με την Pror, σε πλαϊνή όψη. Βλέπουμε την αλυσίδα A της Pror με μπλε χρώμα, την αλυσίδα B της Pror με ροζ χρώμα, τα κατάλοιπα 1-27 της Ala2Ile2-6 με μοβ χρώμα και τα κατάλοιπα 33-60 της Ala2Ile2-6 με κίτρινο χρώμα. Οι πλευρικές ομάδες των καταλοίπων που συμμετέχουν στον υδρόφοβο πυρήνα φαίνονται σε CPK (0.8).



Εικόνα 3-23: Σχηματικό στερεοδιάγραμμα της δομής του μονομερούς της Ala2Ile2-6 σε υπέρθεση με την Pror, σε πλαϊνή όψη. Βλέπουμε την αλυσίδα B της Pror με ροζ χρώμα και τα κατάλοιπα 1-27 της Ala2Ile2-6 με μοβ. Οι πλευρικές ομάδες των καταλοίπων που συμμετέχουν στον υδρόφοβο πυρήνα φαίνονται σε licorice.

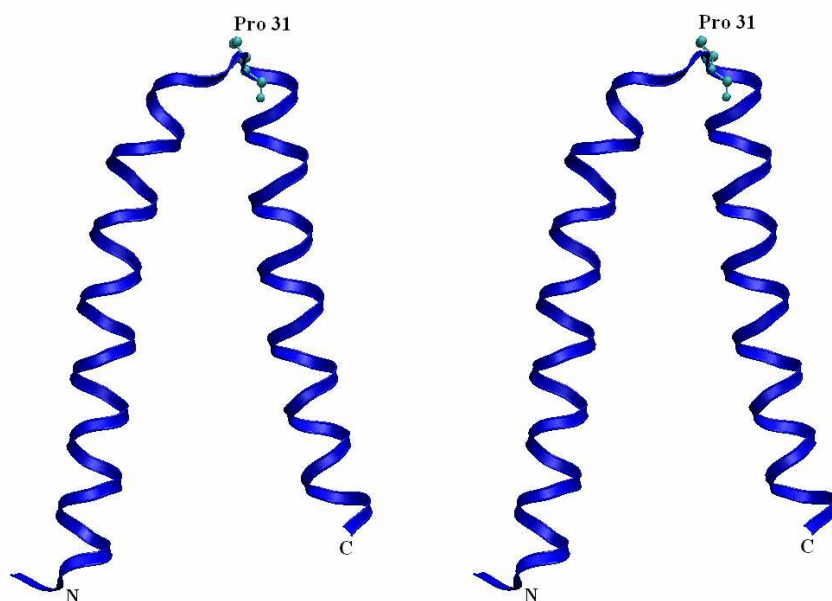


Εικόνα 3-24: Σχηματικό στερεοδιάγραμμα της δομής του μονομερούς της Ala2Ile2-6 σε υπέρθεση με την Prop, σε πλαϊνή όψη. Βλέπουμε την αλυσίδα A της Prop με μπλε χρώμα και τα κατάλοιπα 33-60 της Ala2Ile2-6 με κίτρινο χρώμα. Οι πλευρικές ομάδες των καταλοίπων που συμμετέχουν στον υδρόφοβο πυρήνα φαίνονται σε licorice.

3.7 Προετοιμασία μονομερούς της Ala₂Pe₂-6 με αναδίπλωση της A31P

Ύστερα από αυτές τις δοκιμές με τα μονομερή της Ala₂Pe₂-6, σκεφτήκαμε να κάνουμε ακόμα μια προσομοίωση με ένα μονομερές, το οποίο αυτή τη φορά θα είχε την αλληλουχία της Ala₂Pe₂-6, αλλά την αναδίπλωση της A31P. Επειδή, όπως γνωρίζουμε η δομή της A31P είναι ιδιαίτερη, εξαιτίας του άκαμπτου δακτυλίου της προλίνης στη θέση 31, θέλαμε να δούμε κατά πόσο το συγκεκριμένο μονομερές θα διαλυθεί στο διάλυμα.

Ανατρέξαμε για μια ακόμη φορά στη βάση δεδομένων PDB και βρήκαμε τη δομή της A31P που χρειαζόμαστε. Πρόκειται για την 1B6Q.pdb (**Εικόνα 3-25**). Χρησιμοποιώντας το πρόγραμμα X-fit κάναμε τις απαραίτητες αλλαγές στα κατάλοιπα που συμμετέχουν στο σχηματισμό του υδρόφοβου πυρήνα. Καταλήξαμε έτσι με την επιθυμητή δομή.

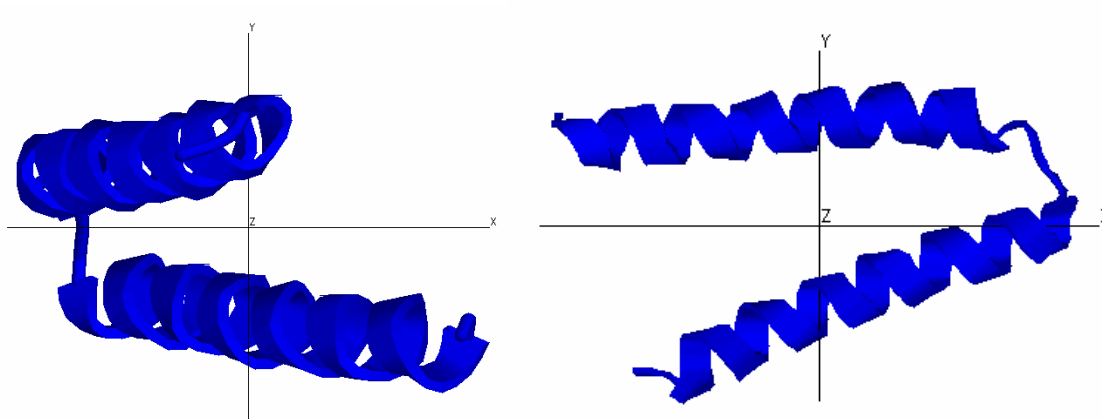


Εικόνα 3-25: Σχηματικό στερεοδιάγραμμα της 1B6Q με μπλε χρώμα. Στη θέση 31 φαίνεται το κατάλοιπο της προλίνης με χρώμα κυανό (CPK).

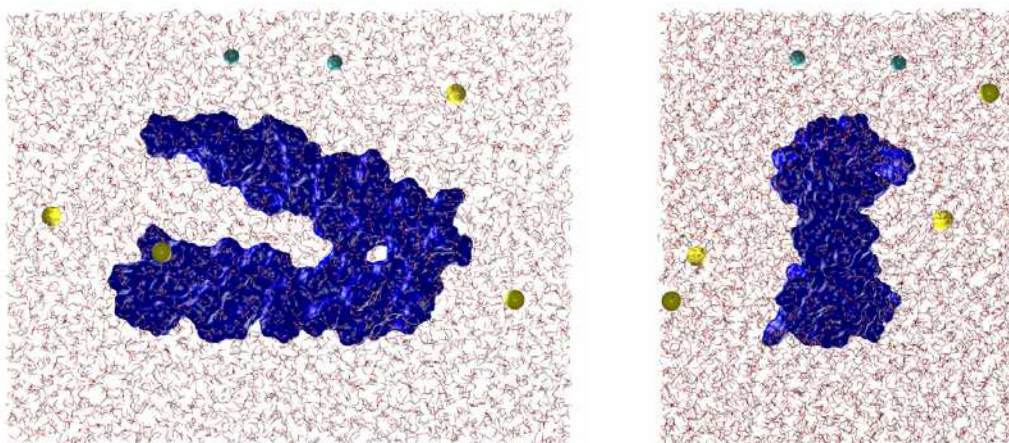
Είμαστε έτοιμοι για την προετοιμασία του συστήματος. Χάριν ευκολίας δε θα αναφέρουμε τη διαδικασία, αφού είναι ήδη γνωστή. Θα επισημάνουμε μόνο ορισμένα σημεία στα οποία υπάρχει διαφορά, όπως η ευθυγράμμιση των αξόνων της αδράνειας (**Εικόνα 3-26**).

Η ενυδάτωση του μονομερούς γίνεται σε ορθογώνιο πλαίσιο με με προ-ισορροπημένα μόρια νερού τύπου TIP3, διαστάσεων 78.2 x 60.1 x 49.2 Å³ (**Εικόνα 3-27**). Όλα τα μόρια νερού που βρίσκονταν σε απόσταση μικρότερη από 1.8 Å απομακρύνθηκαν. Το τελικό σύστημα περιλαμβάνει 901 άτομα πρωτεΐνης, 20850

άτομα νερού, 4 ιόντα Na^+ (SOD) και 2 ιόντα Cl^- (CLA). Χρησιμοποιώντας το πρόγραμμα Rasmol μελετήσαμε το σύστημα (ionized.pdb) για την περίπτωση που κάποια μόρια νερού ή ιόντα τοποθετήθηκαν τυχαία στο εσωτερικό του υδρόφοβου πυρήνα.



Εικόνα 3-26: *Αριστερά:* Η θέση του μορίου σύμφωνα με το ορθοκανονικό σύστημα αξόνων πριν την ευθυγράμμιση των αξόνων της αδράνειας του μορίου. *Δεξιά:* Η θέση του μορίου μετά την ευθυγράμμιση των αξόνων.



Εικόνα 3-27: *Αριστερά:* Πλαϊνή όψη του πλαισίου που περιλαμβάνει το μόριο του μονομερούς, τα ιόντα (με γαλάζιο τα ιόντα Cl^- και με κίτρινο τα ιόντα Na^+ , σε VDW με κλίμακα σφαίρας 2.9) και τα μόρια του νερού. *Δεξιά:* Κάτοψη του ίδιου πλαισίου.

Το τελευταίο βήμα πριν τη χρήση του προγράμματος NAMD είναι να καθορίσουμε τα όρια του συστήματος (κατά μήκος του ορθογώνιου πλαισίου). Οι συντεταγμένες των ορίων του ορθογώνιου πλαισίου για το μονομερές με την τοπολογία της A31P αυτές έχουν ως εξής:

	Minimum	Maximum	Centre	Range
On X :	-39.64	38.50	-0.57	78.15
On Y :	-28.24	31.87	1.82	60.11
On Z :	-25.31	23.88	-0.71	49.19

Ακολουθεί η ελαχιστοποίηση και η εξισορρόπηση της ενέργειας. Στο τέλος και της φάσης εξισορρόπησης οι διαστάσεις του πλαισίου με τα νερά ήταν $76.6 \times 58.9 \times 48.2 \text{ \AA}^3$. πραγματοποιήθηκαν επτά προσομοιώσεις μοριακής δυναμικής. Οι πέντε από αυτές είχαν διάρκεια 2 ns, μια 1.4 ns και άλλη μια 10 ns με τις ατομικές συντεταγμένες ολόκληρου του συστήματος να σώζονται κάθε 0.4 ps.



Κεφάλαιο 4

*Ανάλυση
Προσομοιώσεων*

Σ

το κεφάλαιο αυτό θα αναλύσουμε σταδιακά τα δεδομένα που εξήχθησαν από τις προσομοιώσεις των δυο διμερών, δηλαδή του διμερούς Ala_2Ile_2 -6 με syn τοπολογία και του Ala_2Ile_2 -6 με anti (Rop-like) τοπολογία, που αναφέραμε στο προηγούμενο κεφάλαιο. Επιπλέον θα δούμε και θα σχολιάσουμε συνοπτικά τα αποτελέσματα των προσομοιώσεων που πραγματοποιήσαμε για τα συστήματα των μονομερών (βλ. §3.5-§3.7). Στο στάδιο αυτό της ανάλυσης θα δούμε κατά πόσο οι δομές των δύο συστημάτων είναι σταθερές, αν δηλαδή παραμένει η δομή του 4-α ελικοειδούς δεματίου ή αν έχει αποδιαταχθεί στο διάλυμα. Επίσης θα μελετήσουμε διάφορες παραμέτρους όπως για παράδειγμα το RMSD (Root Mean Square Deviation, τυπική απόκλιση) τόσο από την αρχική, όσο και από τη μέση δομή, το RMSF (Root Mean Square Fluctuations, τυπικές διακυμάνσεις), το Radius of Gyration (την ακτίνα περιστροφής των μορίων) και άλλα.

Πριν ξεκινήσουμε όμως την ανάλυση, θα δούμε σε γενικές γραμμές πως ορίζεται η κάθε παράμετρος και για ποιο λόγο τη χρησιμοποιούμε. Ας ξεκινήσουμε με το RMSD από την αρχική δομή, το οποίο μας παρουσιάζει την απόκλιση των δομών που παράγονται μέσω της προσομοίωσης από τις αρχικές συντεταγμένες της δομής του μορίου. Πρόκειται για την πιο εύχρηστη παράμετρο όσο αφορά τον καθορισμό της σταθερότητας μιας προσομοίωσης μοριακής δυναμικής σε συνάρτηση με το χρόνο (Garemyr *et al.*, 1999). Το RMSD αποτελεί ένα αριθμητικό μέτρο της διαφοράς μεταξύ δύο δομών. Ορίζεται ως:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{\text{N}_{\text{atoms}}} [\chi_i - \chi_0]^2}{\text{N}_{\text{atoms}}}}$$

όπου, N_{atoms} είναι ο αριθμός των ατόμων των οποίων οι θέσεις συγκρίνονται, χ_i οι συντεταγμένες των ατόμων για την εκάστοτε δομή του μορίου κατά τη διάρκεια του trajectory και χ_0 οι συντεταγμένες των ατόμων της αρχικής δομής του μορίου (Παράρτημα 4-1).

Πέρα όμως από το RMSD από την αρχική δομή, θα μελετήσουμε και το RMSD από τη μέση δομή. Η μέση δομή του συστήματος δεν υφίσταται πραγματικά, δηλαδή δεν πρόκειται για μια δομή που λαμβάνει το μόριο κατά τη διάρκεια της προσομοίωσης, παρά για μια δομή που εμείς υπολογίζουμε, χρησιμοποιώντας σαν βάση το σύνολο των διαδοχικών διαμορφώσεων που λαμβάνει το μόριο και το οποίο είναι γνωστό ως τροχιά (trajectory). Πρέπει να σημειώσουμε εδώ, πως αναφερόμαστε σε οικογένεια δομών, επειδή οι ατομικές συντεταγμένες του συστήματος υπολογίζονται κάθε 0.4ps και κατά συνέπεια λαμβάνουμε υπόψη μια πληθώρα διαμορφώσεων από τις οποίες περνάει το μόριο. Ο υπολογισμός λοιπόν του RMSD από τη μέση δομή, θα μας βοηθήσει να δούμε εάν το μόριο τείνει να λάβει μια ή περισσότερες διαμορφώσεις. Ορίζεται ως:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{\text{N}_{\text{atoms}}} [\chi_i - \chi^{\text{aver}}]^2}{\text{N}_{\text{atoms}}}}$$

όπου, όπως είδαμε και προηγουμένως για το RMSD από την αρχική δομή, N_{atoms} είναι ο αριθμός των ατόμων των οποίων οι θέσεις συγκρίνονται, χ_i οι συντεταγμένες των ατόμων για την εκάστοτε δομή του μορίου κατά τη διάρκεια του trajectory, ενώ χ^{aver} οι συντεταγμένες των ατόμων της μέσης δομής του μορίου (**Παράρτημα 4-2**).

Για την αποτίμηση της συμφωνίας μεταξύ των ατομικών διακυμάνσεων και των πειραματικών B παραγόντων (B factors) υπολογίζουμε το RMSF (Root Mean Square Fluctuation) για τη μέση δομή. Ορίζεται ως:

$$\text{RMSF} = \sqrt{\frac{\sum \langle [r_i^t - r_i^{\text{aver}}]^2 \rangle_t}{\text{N}_{\text{atoms}}}}$$

Όπου r_i^t η θέση του ατόμου i στη χρονική στιγμή t , r_i^{aver} είναι η μέση θέση του ατόμου i κατά τη διάρκεια του trajectory και $r_i^t - r_i^{\text{aver}}$ η απόσταση μεταξύ της θέσης του ατόμου σε μια δεδομένη χρονική στιγμή και της μέσης θέσης. Οι διακυμάνσεις υπολογίστηκαν για όλα τα άτομα κάθε καταλοίπου εκτός από τα άτομα υδρογόνου (nonhydrogen atoms). Σε περίπτωση που οι διακυμάνσεις είναι ισοτροπικές (isotropic) υπάρχει μια ακριβής σύνδεση μεταξύ των τιμών του rms για κάθε ανεξάρτητο άτομο i και του αντίστοιχου πειραματικού παράγοντα θερμοκρασίας B_i , δηλαδή:

$$\text{RMSF}_i = \sqrt{\frac{3B_i}{8\pi^2}}$$

Οι παράγοντες B για όλα τα μη υδρογονικά άτομα, όλα τα άτομα του σκελετού και όλα τα άτομα των πλευρικών αλυσίδων αντίστοιχα υπολογίζονται ανά κατάλοιπο και μετατρέπονται σε rms διακυμάνσεις χρησιμοποιώντας τον παραπάνω τύπο (Garemyr *et al.*, 1999).

Ένας τρόπος για να ελέγξουμε τη σταθερότητα και τη διατήρηση της δομής είναι μέσω του radius of gyration (R_G) (Zagrovic *et al.*, 2005). Πρόκειται για μια παράμετρο η οποία είναι ενδεικτική του μεγέθους ενός σώματος οποιουδήποτε σχήματος και αντανακλά το κατά πόσο είναι το σώμα συμπαγές. Ισοδυναμεί με την υδροδυναμική ακτίνα μιας σφαίρας με παρόμοιες υδροδυναμικές ιδιότητες και ορίζεται ως:

$$R_G = \sqrt{\langle (\mathbf{r}_i^t - \mathbf{r}_i^{\text{aver}})^2 \rangle_t}$$

όπου r_i^t είναι η απόσταση του ατόμου i από το κέντρο βάρους του μορίου σε κάθε χρονική στιγμή t και r_i^{aver} η μέση απόσταση του ατόμου i από το κέντρο βάρους του μορίου καθόλη τη διάρκεια του trajectory. Έτσι μπορούμε να υπολογίσουμε τη μεταβολή του R_G του μορίου βάσει μιας μέσης τιμής (**Παράρτημα 4-3**).

Στη συνέχεια ακολουθεί μια σειρά αναλύσεων που έχουν να κάνουν με τις αποστάσεις μεταξύ των ατόμων. Πρόκειται για τη δημιουργία χαρτών αποστάσεων, μέσω του προγράμματος carma, οι οποίοι χάρτες περιέχουν πληροφορίες σχετικά με την κινητικότητα της πρωτεΐνης που μελετάμε κάθε φορά, για παράδειγμα σε ποιο τμήμα της παρατηρείται μεγαλύτερη κινητικότητα σε σχέση με άλλα κατά τη διάρκεια της προσομοίωσης. Τέτοιοι χάρτες είναι ο χάρτης αποστάσεων μεταξύ των Ca ατόμων (Ca-Ca distance map), όπου υπολογίζεται η μέση απόσταση κάθε Ca ατόμου από τα άλλα Ca άτομα κατά τη διάρκεια της προσομοίωσης και ο χάρτης των RMSD που αντιστοιχούν σε αυτές τις αποστάσεις. Ένας ακόμη είναι ο χάρτης που περιλαμβάνει τις RMS διακυμάνσεις του κάθε frame από τα άλλα frames (frame-to-frame rmsd), βάσει του οποίου μπορούμε να εξάγουμε συμπεράσματα για την κινητικότητα του μορίου, καθώς επίσης και να διακρίνουμε μια πρώτη ομαδοποίηση των δομών.

Έπειτα υπολογίζουμε τον πίνακα variance-covariance (Ichiye *et al.*, 1991; Mayer *et al.*, 2003), τον οποίο μετά από την κανονικοποίηση των τιμών χρησιμοποιούμε για τη δημιουργία του πίνακα cross-correlation για τα Ca άτομα του μορίου. Οι τιμές του πίνακα κυμαίνονται από -1 έως +1. Αυτός ο πίνακας μπορεί να μας δώσει πληροφορίες για το πώς συσχετίζονται οι διάφορες κινήσεις που πραγματοποιεί κάθε τμήμα του μορίου, καθώς και για τα δευτερογενή στοιχεία που κινούνται σαν σύνολα (Hünenberger *et al.*, 1995; Ichiye *et al.*, 1991; Lange *et al.*, 2005). Όταν λοιπόν δύο τμήματα κινούνται μαζί έχουμε θετική συσχέτιση και η τιμή θα είναι +1, αν κινούνται αντίθετα έχουμε αρνητική συσχέτιση και η τιμή θα είναι -1, ενώ αν δε σχετίζονται καθόλου η τιμή θα είναι 0 (Luo *et al.*, 2002).

Ακολουθεί η ανάλυση των κύριων συνιστωσών (principal component analysis), όπως ο υπολογισμός των χαρακτηριστικών τιμών (eigenvalues) και των χαρακτηριστικών ανυσμάτων (eigenvectors) (Amadei *et al.*, 1993; Cregut *et al.*, 1998). Ένα χαρακτηριστικό άνυσμα είναι ένα μη μηδενικό άνυσμα το οποίο ικανοποιεί την εξίσωση:

$$Ax = \lambda x$$

όπου A είναι ένας διδιάστατος πίνακας (π.χ. πίνακας variance-covariance), λ η χαρακτηριστική τιμή και x το χαρακτηριστικό άνυσμα που αντιστοιχεί στην τιμή αυτή. Ένα μόριο πραγματοποιεί μια σύνθετη κίνηση, η οποία μπορεί να αναλυθεί σε

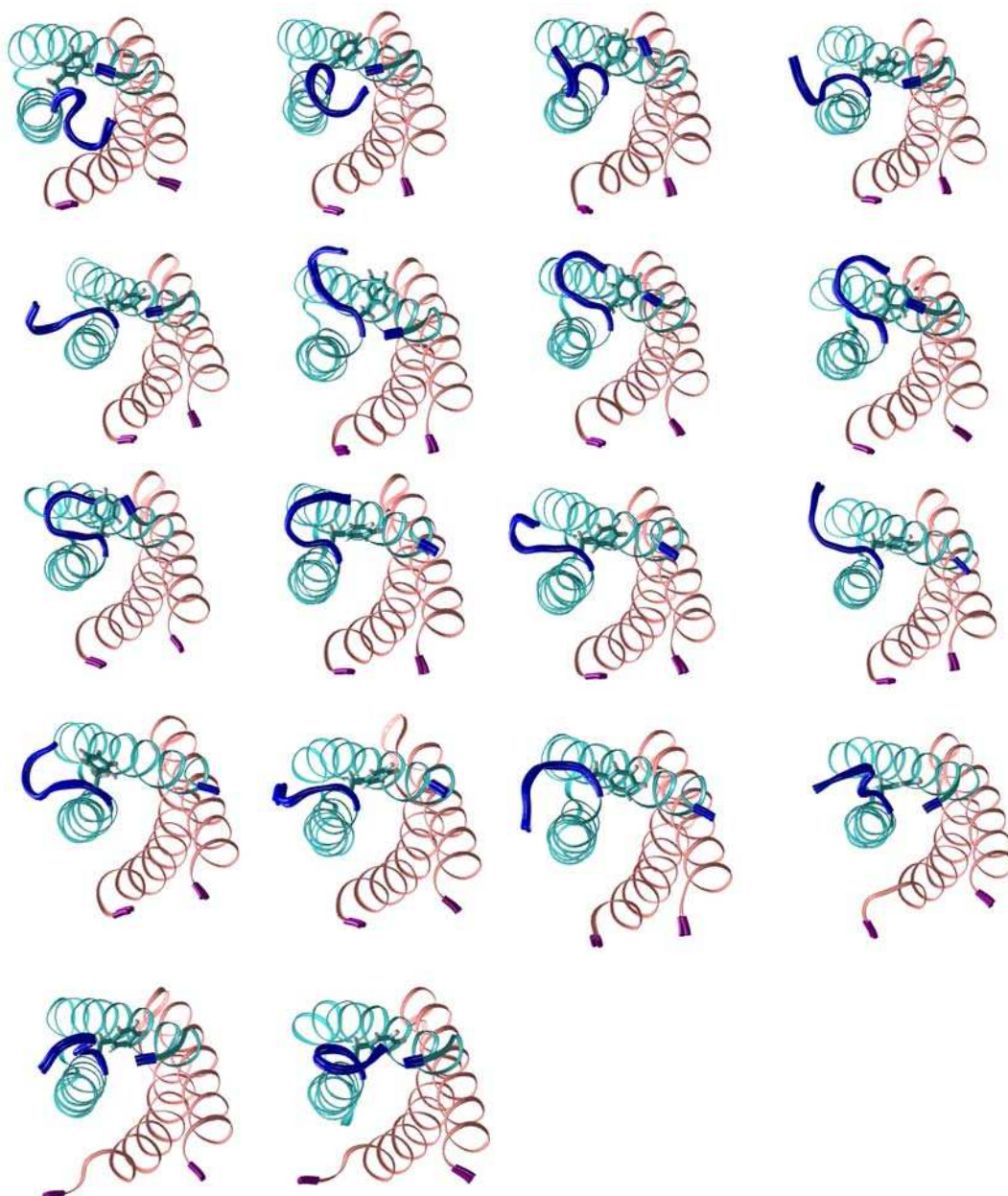
επιμέρους απλές κινήσεις, οι οποίες αντιπροσωπεύονται από κάθε ζεύγος χαρακτηριστικής τιμής - χαρακτηριστικού άνυσματος. Επίσης μπορούμε να διαπιστώσουμε εάν η διάρκεια της προσομοίωσης ήταν αρκετή ώστε να έχουμε ένα ικανοποιητικό δείγμα της κίνησης του μορίου. Αυτό γίνεται μέσω της προβολής των ατομικών διακυμάνσεων σε επιλεγμένα χαρακτηριστικά άνυσματα, όπως τα τρία πρώτα άνυσματα που αντιστοιχούν στις βασικές κινήσεις του μορίου. Και εδώ βέβαια παρατηρείται ο περιορισμός που αναφέρθηκε στο κεφάλαιο 2 σχετικά με τη χρονική διάρκεια των προσομοιώσεων, η οποία είναι της τάξης των nanosecond, σε σύγκριση με την κίνηση της πρωτεΐνης σε μακρομοριακό επίπεδο, που μπορεί να είναι από msec έως και κάποια δευτερόλεπτα. Παρόλο που δε μπορούμε να πούμε με σιγουριά ότι παρατηρούμε όλες τις δυνατές διαμορφώσεις ενός μορίου, κατά τη διάρκεια της αναδίπλωσης, το δείγμα μας (sufficient sampling) είναι αρκετά ικανοποιητικό. Τέλος, εφόσον παρατηρηθεί η ύπαρξη sufficient sampling για κάποιο χαρακτηριστικό άνυσμα, μπορούμε να δημιουργήσουμε ένα τεχνητό trajectory το οποίο θα αναπαριστά την κίνηση που αντιστοιχεί στο συγκεκριμένο άνυσμα και έτσι να μελετήσουμε μια συγκεκριμένη κίνηση που πραγματοποιεί το μόριο. Ακολουθώντας μελετάμε τη διακύμανση των eigenvectors ανά κατάλοιπο. Με αυτό τον τρόπο βλέπουμε ποια κατάλοιπα συμμετέχουν περισσότερο σε κάθε χαρακτηριστική κίνηση. Έπειτα με τον υπολογισμό της αθροιστικής διακύμανσης, έστω των πέντε πρώτων χαρακτηριστικών ανυσμάτων, βλέπουμε ποια κατάλοιπα εμφανίζουν τη μεγαλύτερη διακύμανση κατά την κίνησή τους.

Μια ακόμα ανάλυση που πραγματοποιήθηκε είναι η ομαδοποίηση των δομών (cluster analysis) που παρατηρούνται κατά τη διάρκεια της προσομοίωσης. Πρόκειται για έναν καλό τρόπο να δούμε εάν και κατά πόσο οι δομές συγκλίνουν μεταξύ τους. Στο σημείο αυτό είναι δυνατή η κατασκευή ενός δενδρογράμματος, ανάλογου με τα φυλογενετικά δένδρα, με τον πίνακα των RMSD να αποτελεί τον πίνακα των αποστάσεων. Ο αριθμός των ομάδων / clusters που θα προκύψει είναι ενδεικτικός των βασικών δομών από τις οποίες θα “περάσει” το μόριο. Εδώ μπορούμε ακόμα να υπολογίσουμε στατιστικά το ποσοστό των δομών που περιέχονται στο cluster, για να έχουμε μια εικόνα του πόσο μένει το μόριο σε μια συγκεκριμένη διαμόρφωση. Οι υπολογισμοί αυτοί γίνονται με το πρόγραμμα R (R Development Core Team, 2004).

Εκτός από τη δομή της Ala_2Ile_2-6 που είναι ήδη γνωστή και διαθέσιμη στην βάση δεδομένων της PDB, για την οποία έχει γίνει εκτενής μελέτη τόσο δομικά όσο και θερμοδυναμικά (Willis *et al.*, 2000), θα αναλύσουμε, όπως είδαμε και στο προηγούμενο κεφάλαιο και τα αποτελέσματα από την προσομοίωση της Ala_2Ile_2-6 με την τοπολογία της φυσικής Rop. Έτσι θα μπορέσουμε να δούμε κατά πόσο οι δομές αυτές είναι αρκετά σταθερές, καθώς και να προσπαθήσουμε να κατανοήσουμε γιατί προτιμάται η μια διαμόρφωση εις βάρος της άλλης κάτω από δεδομένες συνθήκες.

4.1 Ανάλυση προσομοίωσης διμερούς της Ala_2Ile_2-6 με τοπολογία *syn*

Η προσομοίωση έχει ολοκληρωθεί και αυτό που πρέπει να κάνουμε αρχικά είναι να μελετήσουμε την κίνηση του συστήματος.



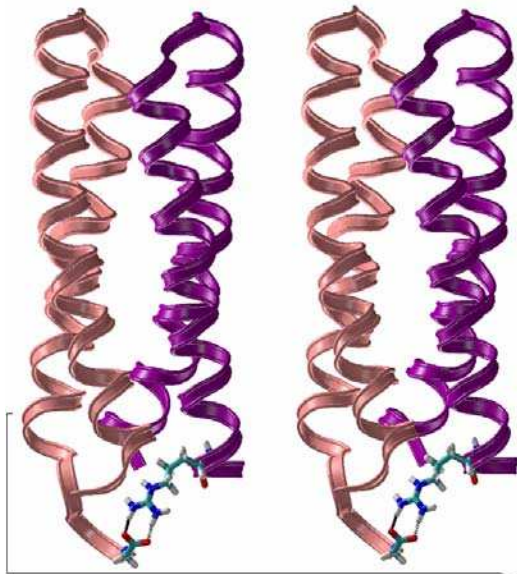
*Εικόνα 4-1: 18 στιγμιότυπα από την προσομοίωση της Ala_2Ile_2-6 με *syn* τοπολογία (1 στιγμιότυπο /1500 frames). Με κyanό χρώμα βλέπουμε τα κατάλοιπα 2-55 της έλικας A, με ροζ τα κατάλοιπα 3-56 της έλικας B, ενώ με μπλε και μοβ βλέπουμε τα κατάλοιπα 1 και 56-60 της A και 2 και 57 της B αλυσίδας αντίστοιχα. Παρατηρούμε επίσης και το κατάλοιπο Phe στη θέση 56 της αλυσίδας A το οποίο αποτελεί τμήμα της υπερκινητικής ουράς, οπότε δε συμμετέχει στο σχηματισμό του 8^{ου} και τελευταίου **layer** της Ala_2Ile_2-6 με *syn* τοπολογία.*

Με τη βοήθεια του προγράμματος VMD (Humphrey *et al.*, 1996) ελέγχουμε την τροχιά (trajectory) που ακολούθησε το σύστημα καθ' όλη τη διάρκεια της προσομοίωσης. Παρακολουθώντας το trajectory παρατηρούμε πως οι “ουρές” των δυο μονομερών, συγκεκριμένα τα κατάλοιπα 1 και 56-60 του μονομερούς Α, καθώς και τα 2 και 57 του μονομερούς Β, κινούνται ακανόνιστα σε σχέση με το υπόλοιπο σύστημα (**Εικόνα 4-1**).

Πιο συγκεκριμένα, φαίνεται πως οι “ουρές” αυτές κινούνται πολύ γρήγορα και η έλικα σε εκείνα τα σημεία έχει αρχίσει να αποδιατάσσεται. Το γεγονός αυτό έρχεται σε μερική συμφωνία με μια από τις κρυσταλλικές δομές της Ala_2Ile_2-6 που χρησιμοποιήσαμε για την κατασκευή του συστήματος, την 1F4M, από την Α έλικα της οποίας απουσιάζουν τα τέσσερα τελευταία κατάλοιπα. Αντιστοίχως παρατηρούμε την υπερκινητικότητα του πρώτου (2) και του τελευταίου (57) καταλοίπου από την αλυσίδα Β, σε σχέση με την αλυσίδα Β της κρυσταλλικής δομής 1F4N, από την οποία τα κατάλοιπα αυτά απουσιάζουν. Επίσης βλέπουμε πως το κατάλοιπο Phe στη θέση 56, λόγω της ύπαρξής του εντός της υπερκινητικής “ουράς”, αδυνατεί να συμμετάσχει στο σχηματισμό του 8^{ου} και τελευταίου στρώματος του υδρόφοβου πυρήνα του συστήματος. Στη συνέχεια θα μελετήσουμε κατά πόσο η υπερκινητικότητα αυτή των “ουρών” επηρεάζει τα δεδομένα της προσομοίωσης, καθώς και τη σταθερότητα του συστήματος. Με μια πρώτη ματιά το trajectory φαίνεται φυσιολογικό. Παρόλα αυτά παρατηρώντας το πιο προσεκτικά, είδαμε κάτι ασυνήθιστο. Επρόκειτο για το σχηματισμό μιας γέφυρας άλατος, κοντά στα C-τελικά άκρα των αλυσίδων Α και Β της Ala_2Ile_2-6 με syn τοπολογία, η οποία σχηματίστηκε μεταξύ της COO^- τελικής ομάδας του καταλοίπου Gly στη θέση 60 της αλυσίδας Α και του καταλοίπου Arg στη θέση 55 της αλυσίδας Β (**Εικόνα 4-2, 4-3**).

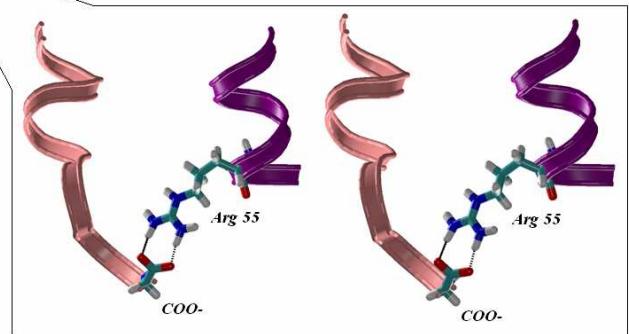
Αυτή η γέφυρα άλατος, υπό φυσιολογικές συνθήκες, δεν υπάρχει στη δομή. Ο σχηματισμός της βασίζεται στο γεγονός ότι η COO^- ομάδα της Gly 60 δεν υπάρχει φυσιολογικά στη δομή της αλυσίδας Α της Ala_2Ile_2-6 , η οποία αποτελείται από 63 κατάλοιπα, αλλά προστέθηκε κατά τη διάρκεια επεξεργασίας της δομής, όπως αναφέρθηκε στο προηγούμενο κεφάλαιο. Η γέφυρα άλατος φαίνεται να συγκρατεί την “ουρά”, που αποτελείται από τα κατάλοιπα 56-60 της αλυσίδας Α, κοντά στην αλυσίδα Β. Στη συνέχεια όμως, λόγω της υπερκινητικότητας της “ουράς”, η γέφυρα “σπάει”, οπότε δεν δημιουργεί πρόβλημα στην κίνηση του συστήματος.

Αφού ολοκληρώθηκε η μελέτη της κίνησης του συστήματος, σειρά έχει η ανάλυση των δεδομένων της προσομοίωσης. Για το λόγο αυτό θα χρησιμοποιήσουμε κυρίως δυο αρχεία, τα protein.dcd και protein.psf. Το αρχείο protein.psf περιλαμβάνει πληροφορίες για τα άτομα, τα κατάλοιπα, το διακριτικό των αλυσίδων (segment name) της πρωτεΐνης (τα νερά του συστήματος έχουν αφαιρεθεί), όμως όχι για τις συντεταγμένες του συστήματος, γι' αυτό χρησιμοποιείται σε συνδυασμό με το αρχείο protein.dcd.



Εικόνα 4-2: Σχηματικό στερεοδιάγραμμα της δομής και τοπολογίας του διμερούς της Ala_2Ple_2-6 . Με ροζ διακρίνουμε την αλυσίδα A, ενώ με μοβ την αλυσίδα B. Κοντά στα καρβοξυ-τελικά τους άκρα παρατηρούμε το σχηματισμό μιας γέφυρας άλατος.

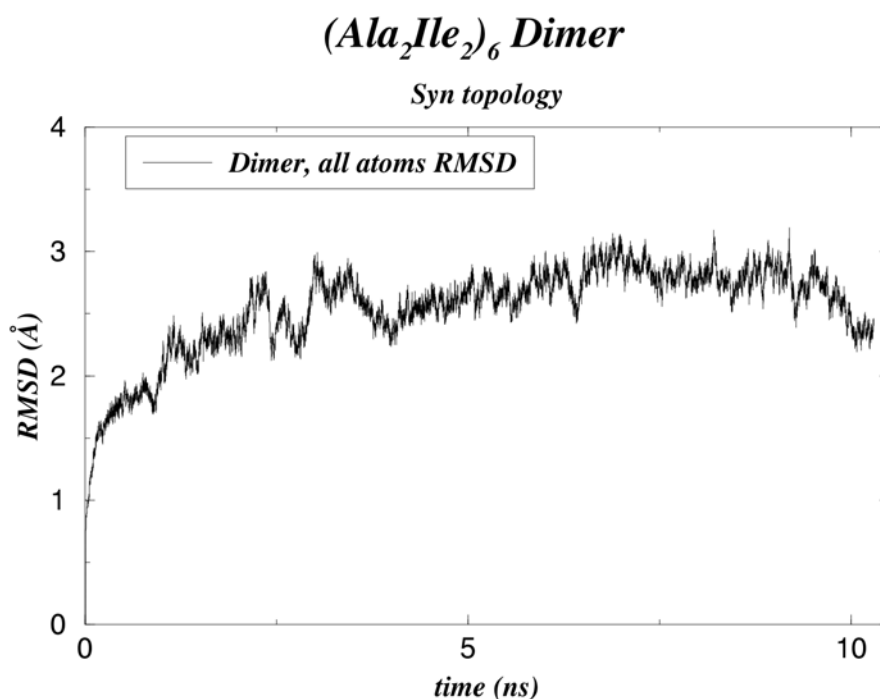
Εικόνα 4-3: Σχηματικό στερεοδιάγραμμα των καρβοξυ-τελικών άκρων των αλυσίδων A (με ροζ) και B (με μοβ) του διμερούς της Ala_2Ple_2-6 . Μπορούμε να διακρίνουμε το κατάλοιπο Arg στη θέση 55 της αλυσίδας B και την COO^- ομάδα στο άκρο της αλυσίδας A. Με διάστικτες γραμμές διακρίνουμε τη γέφυρα άλατος.



4.1.1 RMSD (τυπική απόκλιση) από την αρχική δομή

Στο σημείο αυτό θα μελετήσουμε την τυπική απόκλιση του συστήματος από την αρχική του δομή, κατά πόσο δηλαδή παρεκκλίνει η δομή από την αρχή του trajectory μέχρι το τέλος του, σε συνάρτηση με το χρόνο. Λόγω της ύπαρξης των υπερκινητικών “ουρών”, όπως είδαμε προηγουμένως, υποθέτουμε πως θα υπάρξει κάποια εμφανής αύξηση της τυπικής απόκλισης από την αρχική δομή.

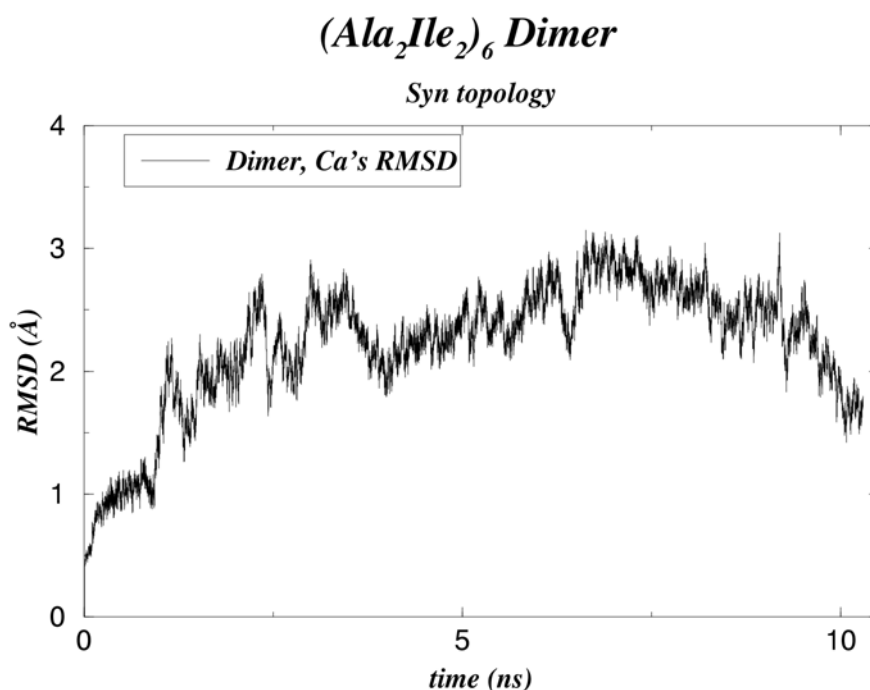
Χρησιμοποιώντας, λοιπόν, τα προγράμματα X-plor (Brünger, 1992) και Xmgr (Turner, P.J., 1991-1995), πήραμε την **Εικόνα 4-4**, στην οποία βλέπουμε μια καμπύλη η οποία απεικονίζει την τυπική απόκλιση (Å), συναρτήσει του χρόνου (ns). Η μέτρηση έχει γίνει με βάση όλα τα άτομα του συστήματος Ala_2Ile_2-6 (εξαιρούνται τα άτομα του νερού, τα οποία έχουν αφαιρεθεί, οπότε δεν θα συμπεριληφθούν σε καμία από τις μετρήσεις), και ως μέτρο σύγκρισης χρησιμοποιείται η αρχική δομή. Στην **Εικόνα 4-4** παρατηρούμε μια αρκετά μεγάλη αύξηση της τυπικής απόκλισης από 0,8Å μέχρι τα 3Å περίπου, για μια χρονική περίοδο της τάξεως των 9ns περίπου, από την αρχή του trajectory. Στα επόμενα 1.5ns φαίνεται πως ελαττώνεται αυτή η διαφορά της τυπικής απόκλισης, κατά 0.8Å περίπου (από 3Å πέφτει στα 2,4Å περίπου). Ενισχύεται συνεπώς, η αρχική μας υπόθεση πως η υπερκινητικότητα των “ουρών” επηρεάζει την τυπική απόκλιση του συστήματος.



Εικόνα 4-4: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του συστήματος. Μπορούμε να παρατηρήσουμε την εμφανή αύξηση της τυπικής απόκλισης όσο κατευθυνόμαστε προς το τέλος του trajectory.

Το αποτέλεσμα αυτό μας ώθησε να μελετήσουμε εκτενέστερα την τυπική απόκλιση του συστήματος, αυτή τη φορά όμως η μέτρηση έγινε βάσει των ατόμων Ca και όχι όλων των ατόμων του συστήματος. Στην **Εικόνα 4-5** βλέπουμε τα αποτελέσματα.

Πρέπει να σημειώσουμε πως, ενώ είναι εμφανής και σε αυτή την εικόνα η αύξηση της τυπικής απόκλισης από 0,5Å έως 3Å περίπου στα 9,5 πρώτα ns (με μεγαλύτερη τιμή τα 3,1Å στα 9ns) και η ελάττωση από 3Å σε 1,7Å στο τελευταίο 1ns της προσομοίωσης, δεν εξάγουμε κανένα διαφορετικό συμπέρασμα.



Εικόνα 4-5: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα Ca άτομα του διμερούς συστήματος. Μπορούμε να παρατηρήσουμε την αύξηση της τυπικής απόκλισης, καθώς και την σχετικά απότομη ελάττωσή της προς το τέλος του trajectory.

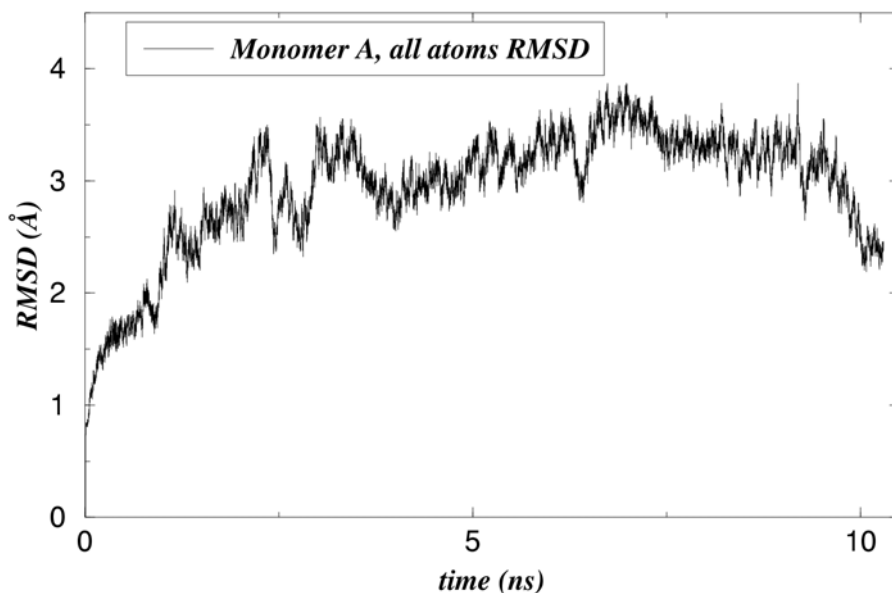
Συνεπώς, πρέπει να συνεχίσουμε να μελετάμε το σύστημα μέχρι να βρούμε τι είναι αυτό που επηρεάζει τόσο πολύ την τυπική απόκλιση. Αυτή τη φορά θα εξετάσουμε την τυπική απόκλιση για καθένα από τα μονομερή του συστήματος (αλυσίδες A και B) χωριστά.

Στην **Εικόνα 4-6** μπορούμε να δούμε την τυπική απόκλιση για όλα τα άτομα του μονομερούς A. Η καμπύλη είναι σε γενικές γραμμές παρόμοια με εκείνη του διμερούς. Η αυξομειώσεις της τυπικής απόκλισης είναι εμφανείς και μπορούμε να σημειώσουμε πως είναι δυνατό, ανάμεσα σε δυο στιγμιότυπα της προσομοίωσης, οι τιμές του RMSD να έχουν μεγάλη διαφορά. Συνεπώς σε καμία περίπτωση αυτό που παρατηρούμε στην **Εικόνα 4-6** δεν είναι μια φυσιολογική εικόνα της τυπικής απόκλισης του μονομερούς (όπως συνέβαινε άλλωστε και πιο πάνω με ολόκληρο το διμερές). Το γεγονός αυτό βέβαια, έρχεται σε συμφωνία με την υπόθεση των

υπερκινητικών “ουρών”, δεδομένου ότι η μεγαλύτερη “ουρά” αντιστοιχεί στο καρβοξυ-τελικό άκρο της αλυσίδας A, στην οποία ανήκουν τα κατάλοιπα 56-60.

$(Ala_2Ile_2)_6$ Dimer

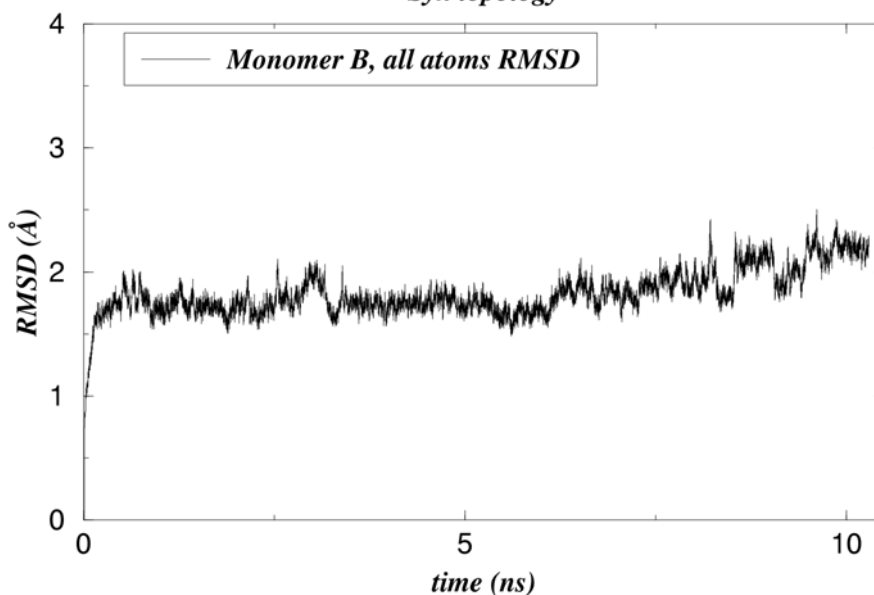
Syn topology



Εικόνα 4-6: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του μονομερούς A του συστήματος. Μπορούμε να διακρίνουμε αύξηση της τυπικής απόκλισης μέχρι το 9^ο ns περίπου της προσομοίωσης και ελάττωσή της προς το τέλος του trajectory.

$(Ala_2Ile_2)_6$ Dimer

Syn topology

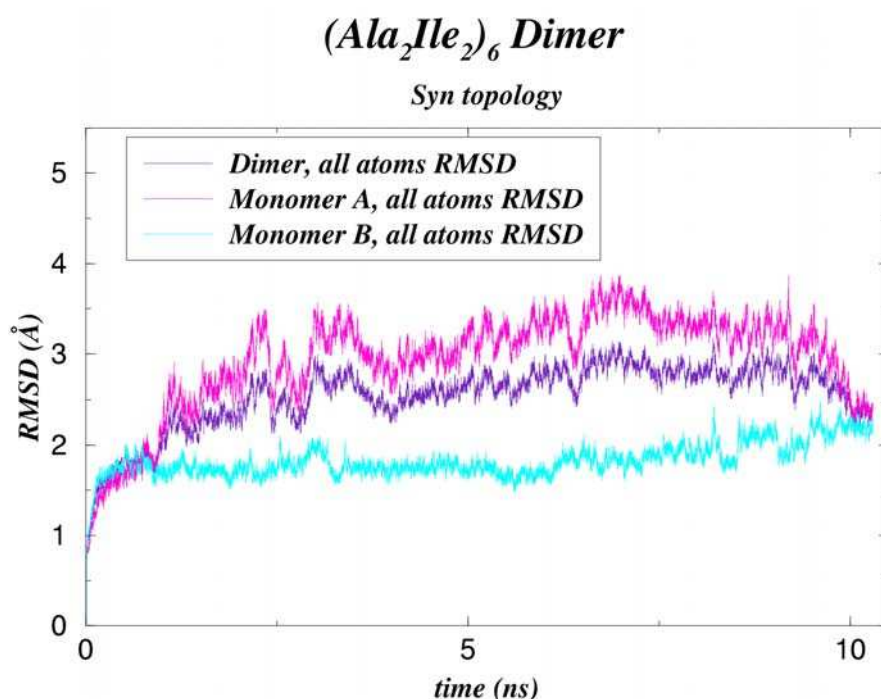


Εικόνα 4-7: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του μονομερούς B του συστήματος. Είναι φανερή η ομαλότερη διακύμανση της τυπικής απόκλισης τόσο σε σχέση με το μονομερές A, όσο και με όλο το διμερές.

Το ερώτημα τώρα είναι αν ισχύει κάτι ανάλογο και για την τυπική απόκλιση της αλυσίδας B. Στην **Εικόνα 4-7** βλέπουμε πως η τυπική απόκλιση για όλα τα άτομα

του μονομερούς B εμφανίζει μια πιο ομαλή διακύμανση, η οποία ξεκινάει από τα 0,8Å και στο τέλος της προσομοίωσης φτάνει στα 2,2Å. Η μικρότερη τιμή που παίρνει η τυπική απόκλιση είναι τα 1,5Å περίπου στα 5,4ns, ενώ η μεγαλύτερη τα 2,5Å στα 9,5ns του trajectory. Παρόλο που η διακύμανση της τυπικής απόκλισης για το μονομερές B φαίνεται πιο ομαλή τόσο σε σχέση με εκείνη του μονομερούς A, όσο και με του διμερούς, παρατηρούμε και εδώ τις απότομες αυξομειώσεις. Πρέπει όμως να λάβουμε υπ' όψιν ότι και στο μονομερές B έχουμε υπερκινητικά άκρα. Μπορεί να μη συμμετέχουν πολλά κατάλοιπα, παρόλα αυτά υπάρχει πιθανότητα να επηρεάζουν το σύστημα.

Πριν προχωρήσουμε παρακάτω, ας δούμε τα παραπάνω αποτελέσματα συγκεντρωτικά. Στην **Εικόνα 4-8** μπορούμε να δούμε την τυπική απόκλιση για όλα τα άτομα, τόσο για το διμερές, όσο και για τα μονομερή A και B χωριστά.

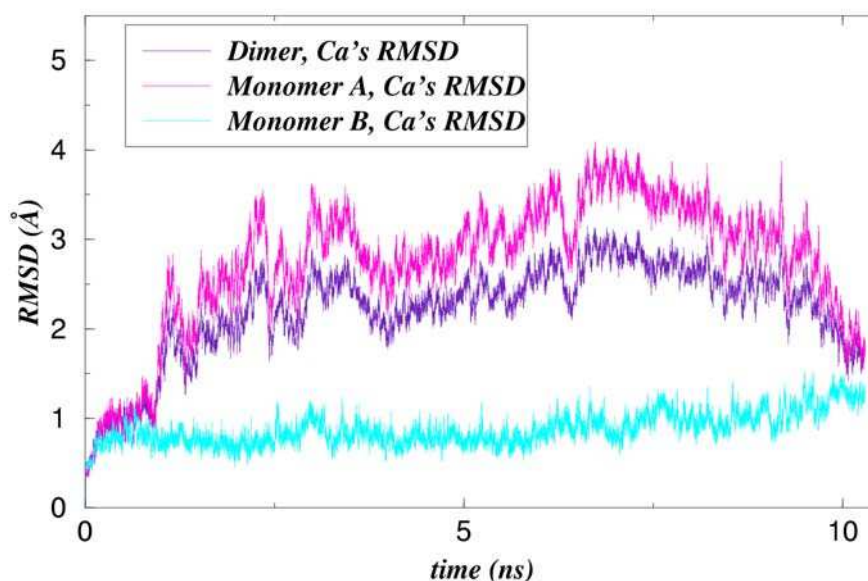


Εικόνα 4-8: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του συστήματος. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Παρατηρούμε πως το μονομερές A είναι αυτό που επηρεάζει περισσότερο την τυπική απόκλιση ολόκληρου του συστήματος.

Επίσης στην **Εικόνα 4-9** μπορούμε να δούμε την τυπική απόκλιση μόνο για τα Ca άτομα του διμερούς και των δυο μονομερών χωριστά. Στις δυο αυτές εικόνες (**4-8**, **4-9**) είναι εμφανές πως το μονομερές A επηρεάζει σε μεγάλο βαθμό την τυπική απόκλιση του διμερούς, ενώ η μικρότερη τυπική απόκλιση του μονομερούς B την αντισταθμίζει ελάχιστα.

(Ala₂He₂)₆ Dimer

Syn topology



Εικόνα 4-9: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα Ca, τόσο του διμερούς, όσο και των δυο μονομερών. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Παρόλο που το μονομερές A επηρεάζει σημαντικά την τυπική απόκλιση του διμερούς, βλέπουμε πως το μονομερές B αντισταθμίζει κατά ένα μέρος τη διακύμανση.

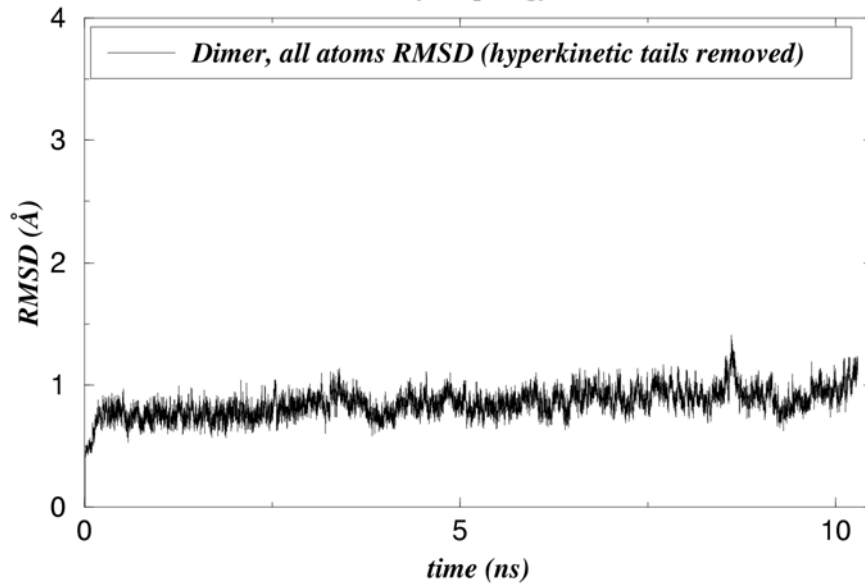
Όλα αυτά μάλλον στηρίζουν, παρά αποκλείουν την αρχική υπόθεσή μας ότι οι “ουρές” στα άκρα των μονομερών επηρεάζουν σημαντικά την τυπική απόκλιση του διμερούς. Στη συνέχεια θα προσπαθήσουμε να αποδείξουμε την υπόθεση αυτή και να δούμε τι τιμές θα πάρει η τυπική απόκλιση, αν θα παραμείνει ως έχει ή αν θα μεταβληθεί.

4.1.1.1 Τι θα συμβεί αν “κοπούν οι ουρές;”

Χρησιμοποιώντας και εδώ το πρόγραμμα X-plor, τροποποιώντας αυτή τη φορά το script ώστε να μην συμπεριληφθούν στη μέτρηση του RMSD τις “ουρές”, (δηλαδή τα κατάλοιπα 1 και 56-60 της αλυσίδας A και 2 και 57 της αλυσίδας B) καθώς και το Xmgr, και ακολουθώντας την ίδια σειρά στα γραφήματα, προκύπτει αρχικά η **Εικόνα 4-10**. Είναι εμφανές πως η τυπική απόκλιση έχει ελαττωθεί κατά ένα μεγάλο βαθμό. Η τιμή της κυμαίνεται από 0,6Å έως 1,5Å, όμως γενικότερα, κατά τη διάρκεια του trajectory παρατηρούμε μια διακύμανση γύρω από τα 0,9Å περίπου, που σημαίνει ότι αν αφαιρεθούν οι υπερκινητικές “ουρές” από τη μέτρηση του RMSD, η δομή παραμένει σε γενικές γραμμές ίδια με την αρχική. Η διαφορά με την αρχική μέτρηση είναι πολύ μεγάλη. Συμπεραίνουμε λοιπόν πως εξαιτίας των υπερκινητικών άκρων των μονομερών A και B προέκυπτε αυτή η ασυνήθιστη αύξηση στην τιμή της τυπικής απόκλισης. Ίσως τελικά αυτή να είναι και η αιτία της απουσίας των ακραίων καταλοίπων από την κρυσταλλική δομή.

$(Ala_2Ile_2)_6$ Dimer

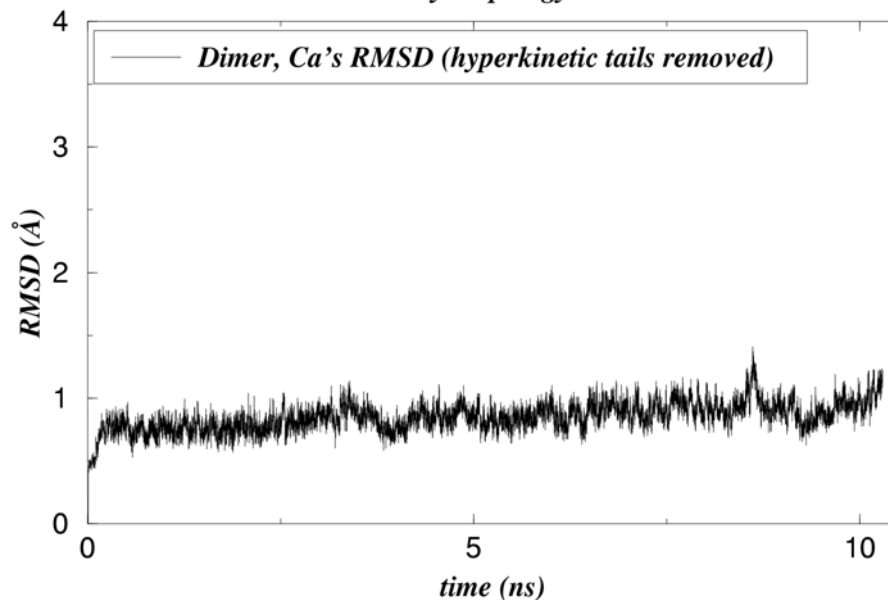
Syn topology



Εικόνα 4-10: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του διμερούς, εκτός από τα υπερκινητικά άκρα. Είναι φανερή η μεταβολή στην τιμή της τυπικής απόκλισης, η οποία σε αντίθεση με την αρχική τιμή, είναι κατά μέσο όρο 0,9Å.

$(Ala_2Ile_2)_6$ Dimer

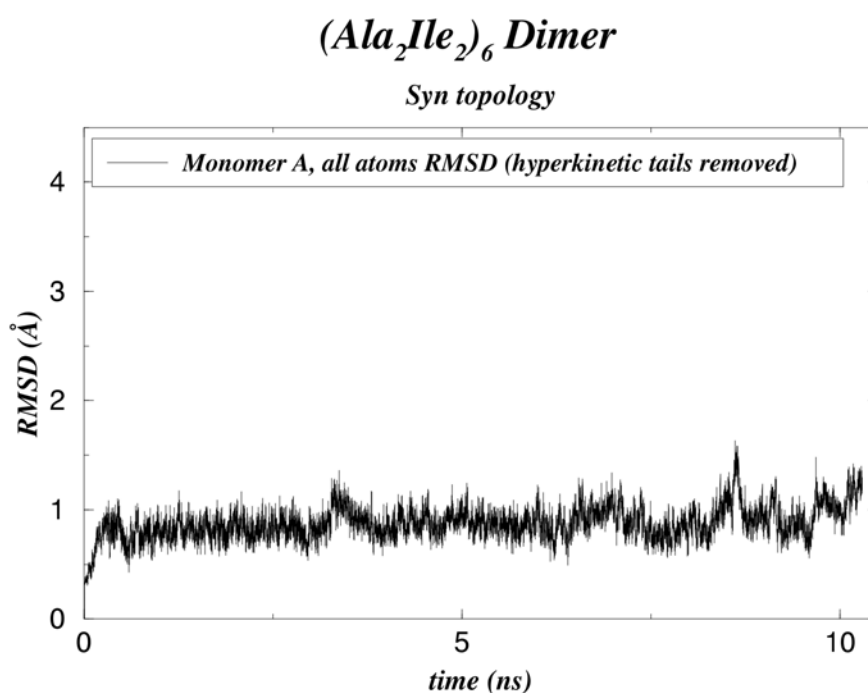
Syn topology



Εικόνα 4-11: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα Ca του διμερούς, εκτός από τα υπερκινητικά άκρα. Είναι φανερή η μεταβολή στην τιμή της τυπικής απόκλισης, η οποία σε αντίθεση με την αρχική τιμή, είναι κατά μέσο όρο 0,9Å.

Ας δούμε όμως τι γίνεται και με το RMSD λαμβάνοντας υπ' όψιν στην μέτρηση μόνο τα άτομα Ca του διμερούς (**Εικόνα 4-11**). Ενώ προηγουμένως υπήρχε μια διαφορά ανάμεσα στο RMSD του διμερούς όταν η μέτρηση έγινε με βάση όλα τα άτομα και τα άτομα Ca (στη δεύτερη περίπτωση η καμπύλη ήταν πιο κυρτή και οι αυξομειώσεις των τιμών πιο έντονες), στη συγκεκριμένη περίπτωση φαίνεται πως οι δυο καμπύλες δεν έχουν εμφανείς διαφορές. Το μέγιστο και στα δυο γραφήματα είναι 1,5Å στα 8,5ns, ενώ το ελάχιστο είναι 0,55Å στα 0,6ns.

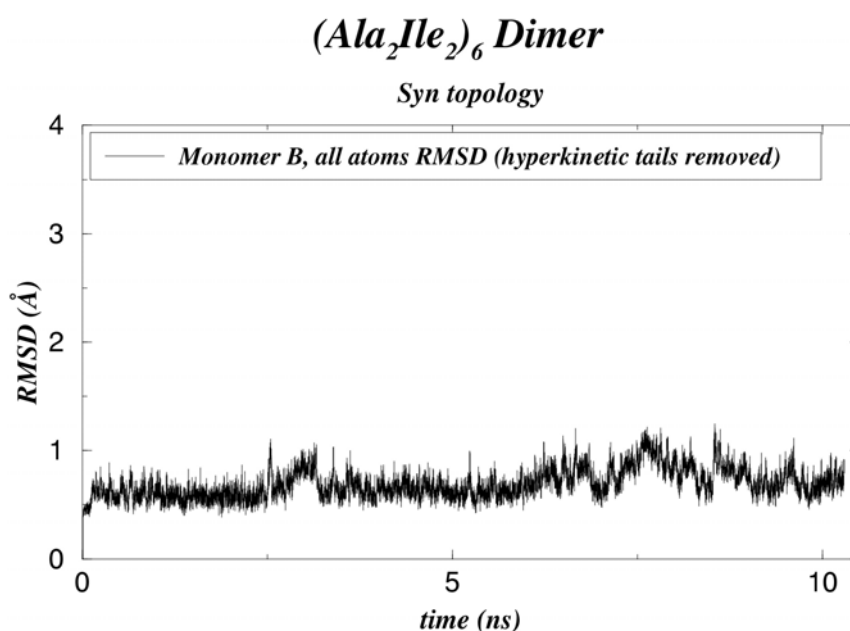
Παρόλο που παρατηρούμε ότι αγνοώντας τις υπερκινητικές “ουρές” το RMSD σταθεροποιείται σε μια φυσιολογική τιμή, θα συνεχίσουμε τις μετρήσεις για κάθε μονομερές χωριστά, όπως κάναμε και με ολόκληρη τη δομή, για να δούμε αν θα υπάρξουν και εκεί διαφορές με τα αρχικά γραφήματα. Στην **Εικόνα 4-12** είναι επίσης φανερή και εντυπωσιακή η “ομαλοποίηση” του RMSD του μονομερούς A. Η μέτρηση έγινε με βάση όλα τα άτομα του μονομερούς, εκτός από τα άτομα των καταλοίπων που συμμετέχουν στα υπερκινητικά άκρα, με μέγιστη τιμή είναι 1,6Å στα 8,5ns, ενώ η ελάχιστη 0,45Å στα 0,55ns. Η μέση τιμή κυμαίνεται και εδώ στα 0,9Å, γεγονός που σημαίνει ότι η δομή είναι αρκετά σταθερή.



Εικόνα 4-12: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του μονομερούς A, εκτός από τα υπερκινητικά άκρα. Είναι εντυπωσιακή η μεταβολή της τυπικής απόκλισης και πιο συγκεκριμένα η σταθεροποίησή της σε μια τιμή γύρω από τα 0,9Å.

Ας δούμε όμως τι συμβαίνει και με το μονομερές B. Στην **Εικόνα 4-13** μπορούμε να δούμε τι συμβαίνει με το RMSD αν στη μέτρηση ληφθούν υπ' όψιν τα άτομα όλων των καταλοίπων εκτός από τα Thr 2 και Gly 57, τα οποία κινούνται πολύ

σε σχέση με τα υπόλοιπα κατάλοιπα του μονομερούς, κατά τη διάρκεια της προσομοίωσης. Συγκρίνοντας την καμπύλη αυτή με την αρχική, παρατηρείται επίσης ελάττωση του RMSD κατά μέσο όρο 1,5Å περίπου. Μέχρι το 7^οns υπάρχει μια διακύμανση γύρω από τα 0,7Å, μετά όμως έχουμε συχνές αυξομειώσεις οι οποίες στο τέλος της προσομοίωσης οδηγούν σε ένα RMSD της τάξης των 0,8Å. Εδώ η μέγιστη τιμή είναι 1,25Å στα 8,1ns, ενώ η μικρότερη είναι 0,4Å στα 1,8ns της προσομοίωσης.



Εικόνα 4-13: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του μονομερούς B, εκτός από τα άτομα των καταλοίπων Thr 2 και Gly 57, τα οποία αποτελούν τις υπερκινητικές “ουρές” του συγκεκριμένου μονομερούς. Στο γράφημα αυτό, παρατηρείται ελάττωση του RMSD.

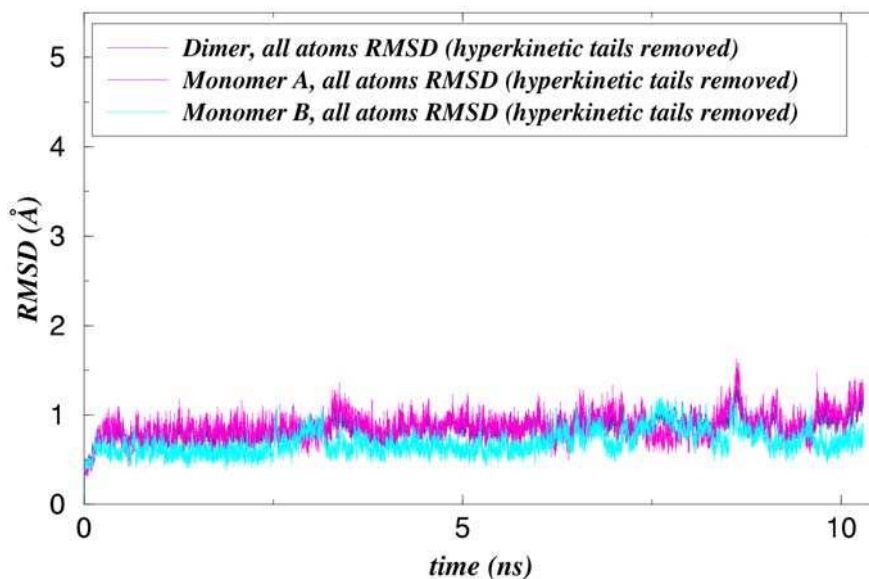
Στο σημείο αυτό θα δούμε τα αποτελέσματα συγκεντρωτικά, τόσο για το διμερές όσο και για τα δυο μονομερή χωριστά. Η μέτρηση έχει γίνει με βάση όλα τα άτομα, εκτός από εκείνα που απαρτίζουν τις “ουρές” (**Εικόνα 4-14**).

Στην **Εικόνα 4-15** βλέπουμε τις αντίστοιχες καμπύλες, μόνο που αυτή τη φορά η μέτρηση έγινε με βάση μόνο τα άτομα Cα, εξαιρουμένων και πάλι των ατόμων των “ουρών”.

Στις δυο παρακάτω εικόνες βλέπουμε πως η καμπύλη του RMSD του διμερούς και του μονομερούς A σχεδόν ταυτίζονται. Ενισχύεται λοιπόν η άποψη μας ότι τελικά να είναι το μονομερές A εκείνο που επηρεάζει περισσότερο την τυπική απόκλιση ολόκληρου του διμερούς.

$(Ala_2Ile_2)_6$ Dimer

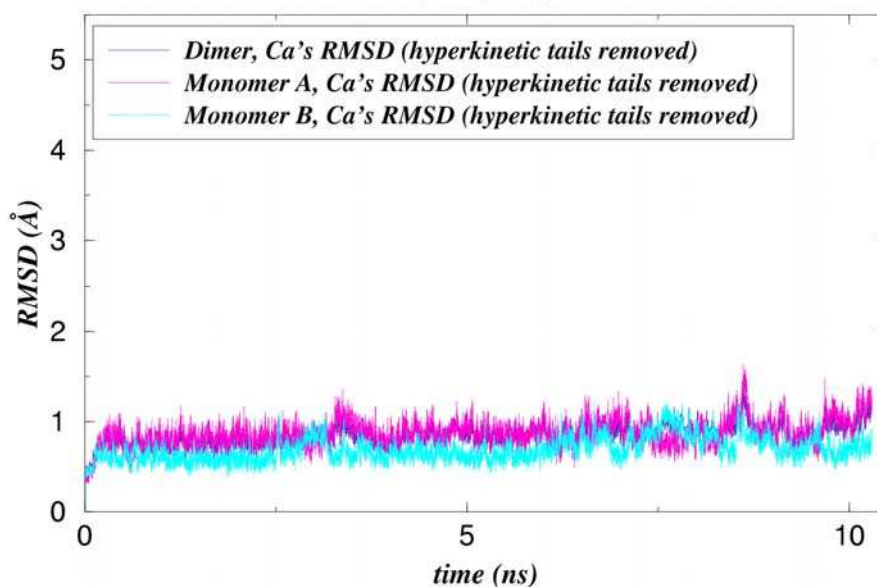
Syn topology



Εικόνα 4-14: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του συστήματος, εκτός από τα άτομα των “ουρών”. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Παρατηρούμε πως το RMSD του διμερούς σχεδόν συμπίπτει με εκείνο του μονομερούς A, ενώ αυτό που αντιστοιχεί στο μονομερές B ξεχωρίζει.

$(Ala_2Ile_2)_6$ Dimer

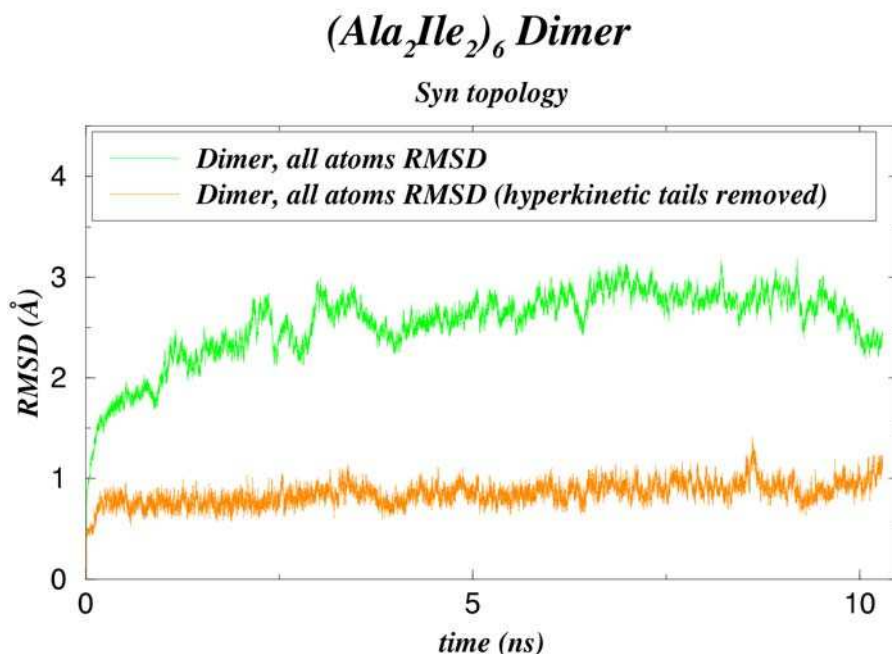
Syn topology



Εικόνα 4-15: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα Ca, τόσο του διμερούς, όσο και των δυο μονομερών με εξαίρεση τα άτομα των “ουρών”. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Η καμπύλη του διμερούς παρόλα αυτά σχεδόν συμπίπτει με την καμπύλη του μονομερούς A.

4.1.1.2 Διαφορές ανάμεσα στην Ala_2Ile_2-6 με τοπολογία syn με τα υπερκινητικά άκρα και την Ala_2Ile_2-6 με τοπολογία syn χωρίς αυτά

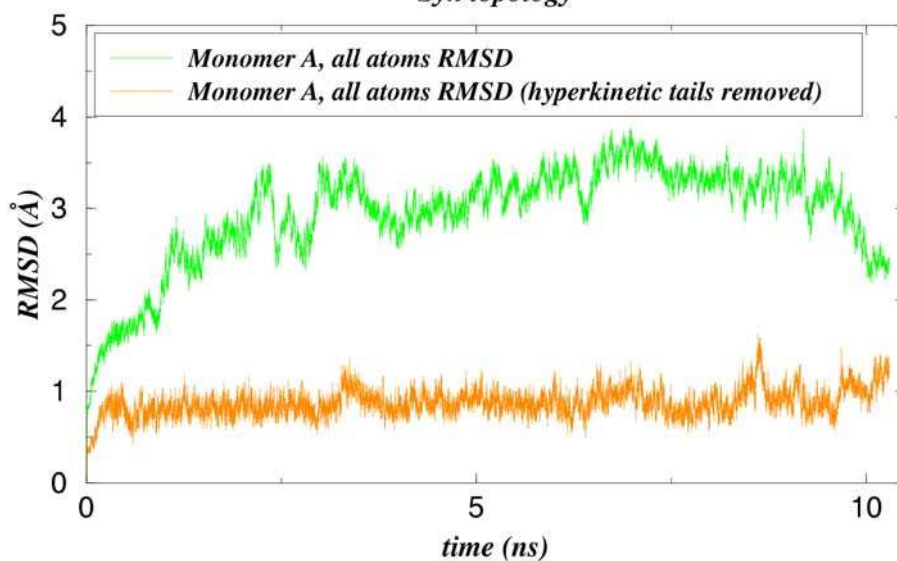
Ας δούμε όμως πως φαίνεται η διαφορά ανάμεσα στις καμπύλες πριν και μετά την “αφαίρεση” των υπερκινητικών άκρων. Στην **Εικόνα 4-16** βλέπουμε το RMSD του διμερούς. Η πράσινη καμπύλη αντιστοιχεί σε ολόκληρο το διμερές, ενώ η πορτοκαλί στο διμερές χωρίς τις “ουρές”. Στην **Εικόνα 4-17** βλέπουμε το RMSD του μονομερούς A. Η πράσινη καμπύλη αντιστοιχεί σε ολόκληρο το μονομερές, ενώ η πορτοκαλί στο μονομερές χωρίς τις “ουρές”. Και στα δυο γραφήματα παρατηρούμε αυτό που περιγράψαμε προηγουμένως, δηλαδή την ομαλοποίηση του RMSD. Τέλος στην **Εικόνα 4-18** βλέπουμε το RMSD του μονομερούς B. Και εδώ η πράσινη καμπύλη αντιστοιχεί σε ολόκληρο το μονομερές, ενώ η πορτοκαλί στο μονομερές χωρίς τις “ουρές”. Η διαφορά με τα δυο προηγούμενα γραφήματα (**Εικόνες 4-16, 4-17**) είναι ότι το RMSD του μονομερούς B μέχρι τα 7ns φαίνεται σαν να είχε μετατοπιστεί 1Å πιο πάνω. Μετά τα 7ns υπάρχει διαφορά ανάμεσα στις δυο καμπύλες με τη δεύτερη (πορτοκαλί) να μοιάζει πιο φυσιολογική.



Εικόνα 4-16: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Η καμπύλη με το πράσινο χρώμα αντιστοιχεί στο διμερές με τα υπερκινητικά άκρα, ενώ εκείνη με το πορτοκαλί χρώμα αντιστοιχεί στο διμερές χωρίς τα υπερκινητικά άκρα. Η μέτρηση έγινε με βάση όλα τα άτομα.

$(Ala_2Ile_2)_6$ Dimer

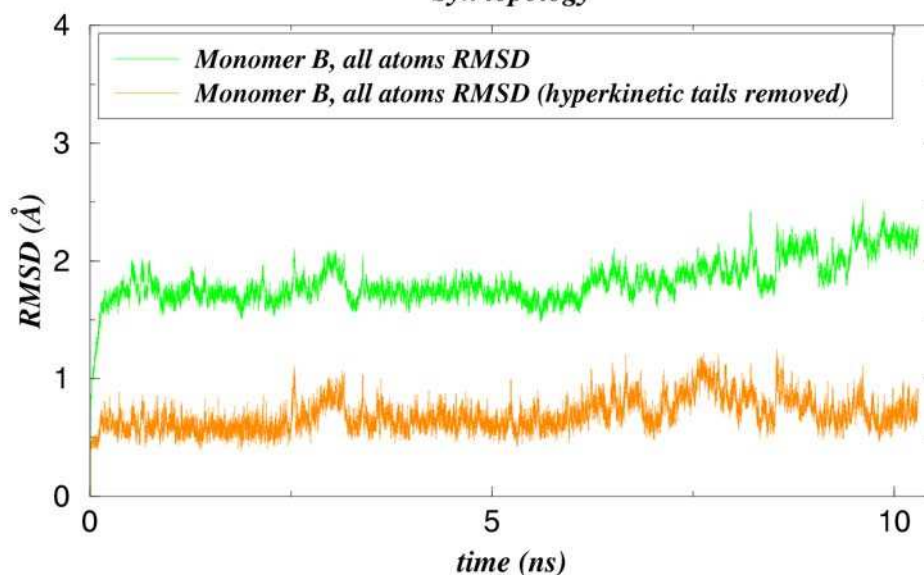
Syn topology



Εικόνα 4-17: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Η καμπύλη με το πράσινο χρώμα αντιστοιχεί στο μονομερές A με τα υπερκινητικά άκρα, ενώ εκείνη με το πορτοκαλί χρώμα αντιστοιχεί στο μονομερές A χωρίς τα υπερκινητικά άκρα. Η μέτρηση έγινε με βάση όλα τα άτομα.

$(Ala_2Ile_2)_6$ Dimer

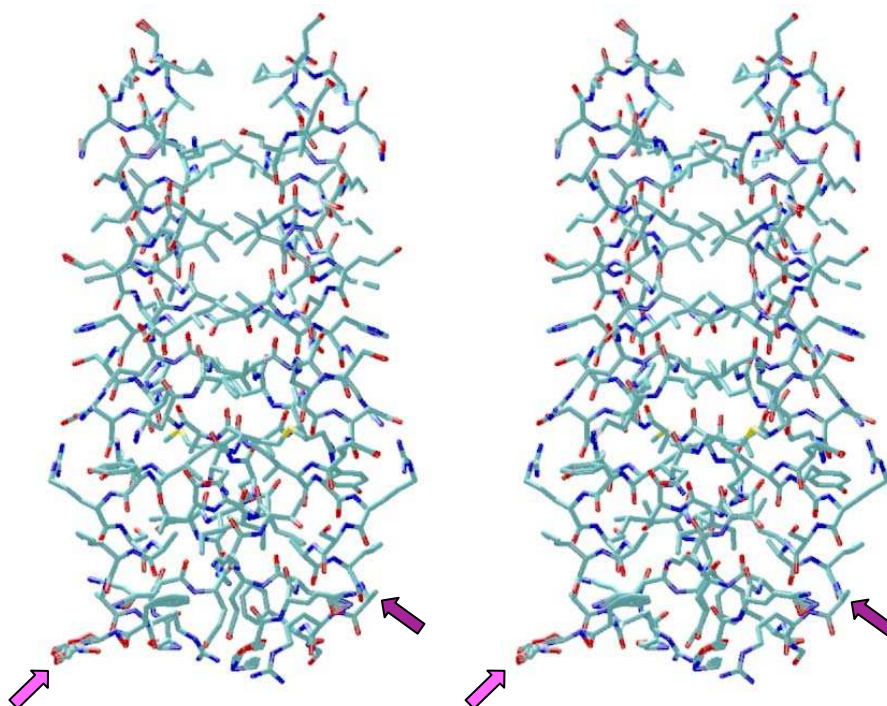
Syn topology



Εικόνα 4-18: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Η καμπύλη με το πράσινο χρώμα αντιστοιχεί στο μονομερές B με τα υπερκινητικά άκρα, ενώ εκείνη με το πορτοκαλί χρώμα αντιστοιχεί στο μονομερές B χωρίς τα υπερκινητικά άκρα. Η μέτρηση έγινε με βάση όλα τα άτομα.

4.1.2 Η μέση δομή και οι αντίστοιχες διακυμάνσεις

Ύστερα από επεξεργασία ολόκληρου του trajectory με τη βοήθεια του προγράμματος X-plor, καταφέραμε να ανακτήσουμε ένα αρχείο pdb στο οποίο εμπεριέχονται οι συντεταγμένες ενός διμερούς που αντιστοιχεί σε μια μέση δομή του αρχικού συστήματος. Ύστερα από μια επεξεργασία του αρχείου αυτού με το πρόγραμμα VMD, δημιουργήσαμε ένα στερεοδιάγραμμα της μέσης αυτής δομής (Εικόνα 4-19).

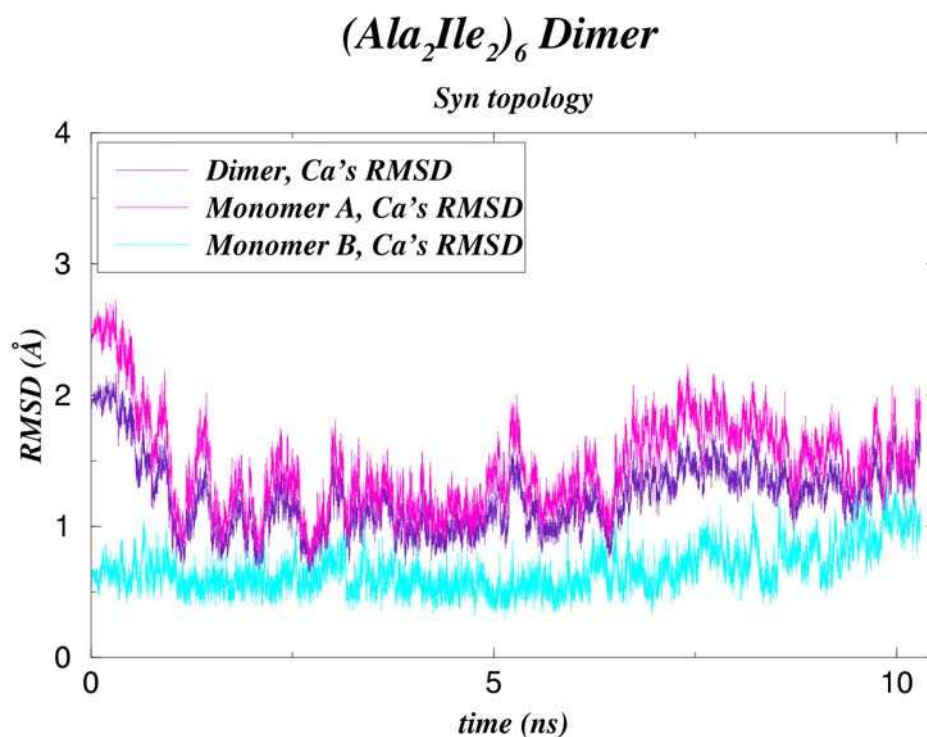


Εικόνα 4-19: Σχηματικό στερεοδιάγραμμα της μέσης δομής του συστήματος Ala₂Ile₂-6. Αριστερά με το ροζ βέλος βρίσκεται η αλυσίδα A και δεξιά με το μοβ βέλος η αλυσίδα B. Σε γενικές γραμμές δεν παρατηρούμε μεγάλες διαφορές από την αρχική δομή του συστήματος, με ελάχιστες εξαιρέσεις.

Στην εικόνα διακρίνουμε την αλυσίδα A αριστερά και τη B δεξιά, σε κάθε τμήμα του στερεοδιαγράμματος. Δεν υπάρχουν πολλές διαφορές όσο αφορά στη δομή του συστήματος σε σχέση με την αρχική. Εξαιρέση αποτελούν ορισμένα κατάλοιπα στα άκρα των αλυσίδων, όπως είναι τα κατάλοιπα Asp58, Asp59 της αλυσίδας A και Arg55 της αλυσίδας B, καθώς και στις στροφές, όπως για παράδειγμα τα κατάλοιπα Leu29 και Asp30 των αλυσίδων A και B, τα οποία φαίνεται πως δεν έχουν λάβει μια συγκεκριμένη διαμόρφωση στο χώρο κατά τη διάρκεια της προσομοίωσης.

4.1.2.1 RMSD από τη μέση δομή

Στο αρχείο pdb που δημιουργήσαμε, υπάρχει μια στήλη η οποία αντιστοιχεί στα RMSF των ατόμων. Από τη στήλη αυτή κρατήσαμε τα RMSF μόνο των ατόμων Ca για να υπολογίσουμε τη διακύμανση για καθένα από τα άτομα αυτά. Στην **Εικόνα 4-20** μπορούμε να δούμε το RMSD από τη μέση δομή για τα Ca άτομα τόσο του διμερούς, όσο και των μονομερών A και B.

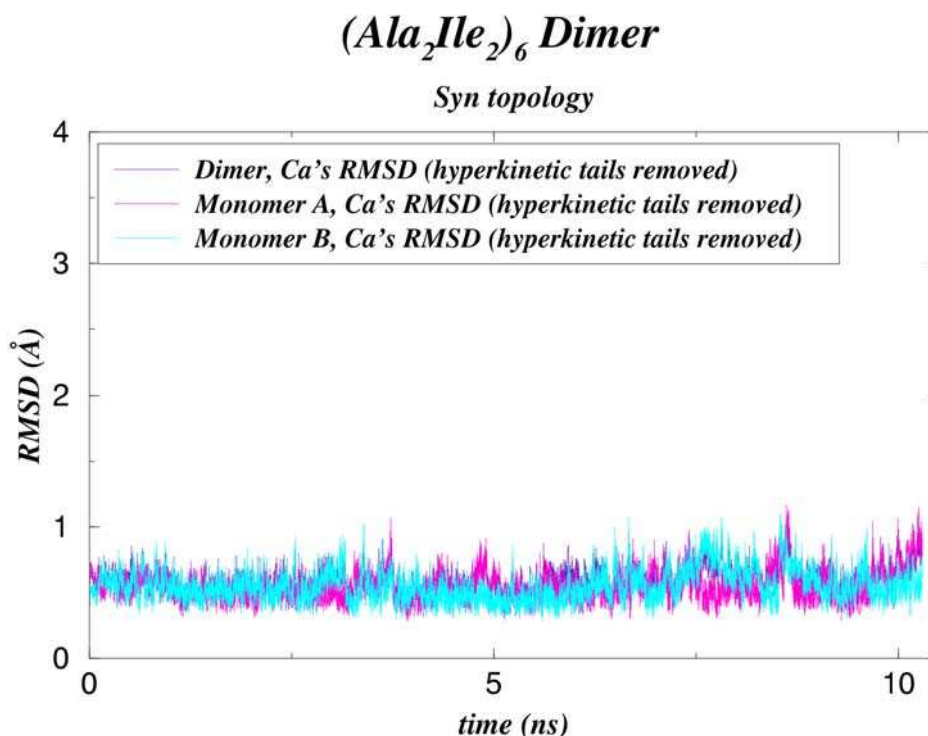


Εικόνα 4-20: Καμπύλη τυπικής απόκλισης RMSD (Å) από τη μέση δομή συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα Ca, τόσο του διμερούς, όσο και των δυο μονομερών. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Η καμπύλη του διμερούς παρόλα αυτά σχεδόν συμπίπτει με την καμπύλη του μονομερούς A.

Η τυπική απόκλιση για το διμερές και το μονομερές A φαίνεται πως και εδώ είναι σχεδόν ταυτόσημες. Παρατηρείται η ίδια διακύμανση με μετατόπιση περίπου 0,5Å. Για το διμερές η διακύμανση είναι από 0,7Å έως 2Å, με μέγιστο την αρχική τιμή των 2Å και ελάχιστο τα 0,7Å στα 2,6 ns περίπου του trajectory. Για το μονομερές A έχουμε μέγιστο 2,5Å στην αρχή του trajectory και ελάχιστο στα 0,6Å επίσης στα 2,6 ns του trajectory. Το μονομερές B και εδώ δεν επηρεάζεται όσο αφορά τη διακύμανση, η τιμή της οποίας κινείται γύρω από τα 0,6Å σε όλη τη διάρκεια του trajectory, με μια μικρή αύξηση προς το τέλος του και ένα τοπικό μέγιστο στα 1,2Å, στο δέκατο ns του trajectory.

Στην **Εικόνα 4-21** βλέπουμε το RMSD από τη μέση δομή για τα Ca άτομα του διμερούς και των μονομερών A και B χωρίς να υπολογίσουμε τα άτομα των

“ουρών”. Είναι ολοφάνερη η διαφορά στη διακύμανση. Πλέον η τιμή της διακύμανσης τόσο για το διμερές, όσο και για τα δυο μονομερή κυμαίνεται γύρω από τα 0,5Å σε όλη τη διάρκεια του trajectory.



Εικόνα 4-21: Καμπύλη τυπικής απόκλισης RMSD (Å) από τη μέση δομή συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα Ca, τόσο του διμερούς, όσο και των δυο μονομερών με εξαίρεση τα άτομα των “ουρών”. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούζια στο μονομερές A και η κωανή στο μονομερές B. Η καμπύλη του διμερούς παρόλα αυτά σχεδόν συμπίπτει με την καμπύλη του μονομερούς A.

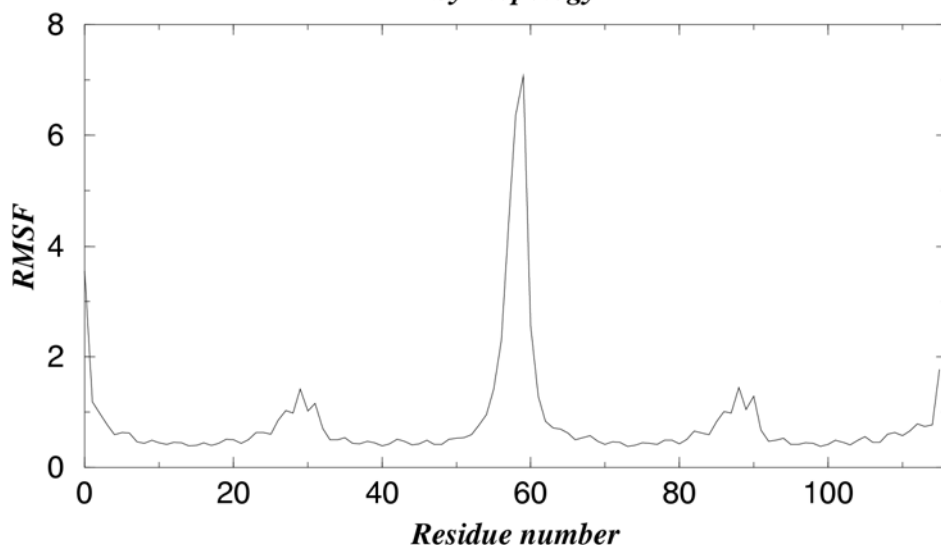
Με βάση τα έως τώρα δεδομένα, μπορούμε να πούμε με σχετική σιγουριά πως τα άτομα των υπερκινητικών άκρων μπορούν να παραληφθούν από τις μετρήσεις, αφού το μόνο που κάνουν είναι να αλλοιώνουν τα αποτελέσματα. Οπότε από εδώ και στο εξής δε θα λαμβάνονται υπ’ όψιν.

4.1.3 RMSF

Το RMSF είναι μια παράμετρος που αντικατοπτρίζει τη διακύμανση ενός ατόμου σε σχέση με τη θέση του στη μέση δομή. Στην **Εικόνα 4-22** μπορούμε να δούμε τα κατάλοιπα που παρουσιάζουν τη μεγαλύτερη κινητικότητα (↑διακύμανση → ↑κινητικότητα).

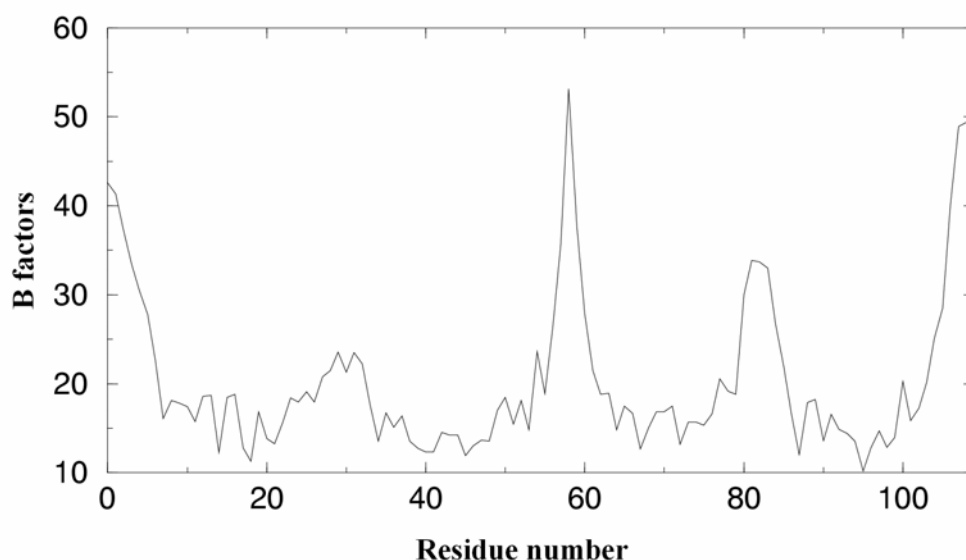
$(Ala_2Ile_2)_6$ Dimer

Syn topology



Εικόνα 4-22: RMSF των ατόμων του διμερούς σε σχέση με τη θέση τους στη μέση δομή. Το RMSF υπολογίζεται σε Å.

Τις υψηλότερες τιμές εμφανίζουν τα κατάλοιπα στην αρχή (1-2), στις στροφές (28-31) και στο τέλος (54-56) κάθε μονομερούς. Τα κατάλοιπα αυτά όπως είδαμε και στο κεφάλαιο 1 δε συμμετέχουν στο σχηματισμό έλικας, επομένως δικαιολογημένα εμφανίζουν το μεγαλύτερο βαθμό κινητικότητας σε σύγκριση με τα υπόλοιπα κατάλοιπα. Στην **Εικόνα 4-23** μπορούμε να δούμε τις τιμές των κρυσταλλογραφικών B παραγόντων (B factors) των Ca ατόμων της δομής 1F4N.



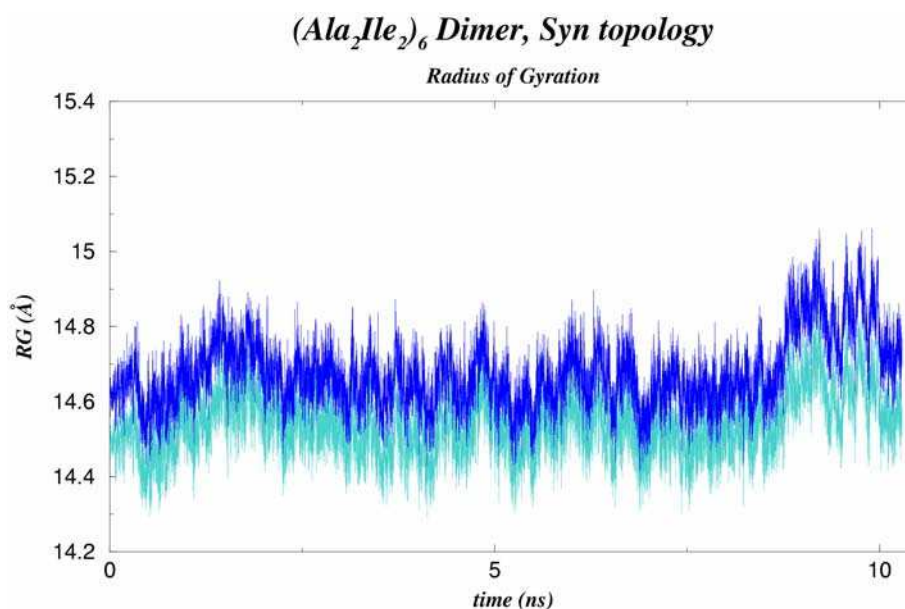
Εικόνα 4-23: B factor ανά Ca άτομο της δομής 1F4N.

Αν συγκρίνουμε τα κρυσταλλογραφικά Bs και τα RMSF της προηγούμενης εικόνας θα δούμε πως, παρόλο που οι κρυσταλλογραφικές τιμές είναι πολύ πιο αυξημένες, έχουμε μέγιστες τιμές στα ίδια κατάλοιπα, δηλαδή στα κατάλοιπα των

στροφών και στα κατάλοιπα που βρίσκονται στην αρχή και στο τέλος κάθε μονομερούς. Εδώ να σημειώσουμε πως στις δυο εικόνες δεν αντιστοιχεί ο ίδιος αριθμός καταλοίπων. Στη δομή 1F4N τα κατάλοιπα είναι λιγότερα, όπως έχουμε ήδη αναφέρει στο κεφάλαιο 3.

4.1.4 Εξέλιξη του R_G , Radius of Gyration

Στην εικόνα 4-24 μπορούμε να δούμε την εξέλιξη του R_G των μονομερών A και B της Ala_2Ile_2-6 . Η τιμή του R_G κυμαίνεται για το μονομερές A γύρω από τα 14,6Å και για το μονομερές B γύρω από τα 14,5Å και δε φαίνεται να παρουσιάζει σημαντικές διακυμάνσεις, αν εξαιρέσουμε τη μικρή αύξηση στα τελευταία 2 ns περίπου του trajectory. Συνεπώς παρατηρούμε πως το μόριο είναι αρκετά σταθερό, πράγμα που γνωρίζουμε ήδη από τα προηγούμενα αποτελέσματα.

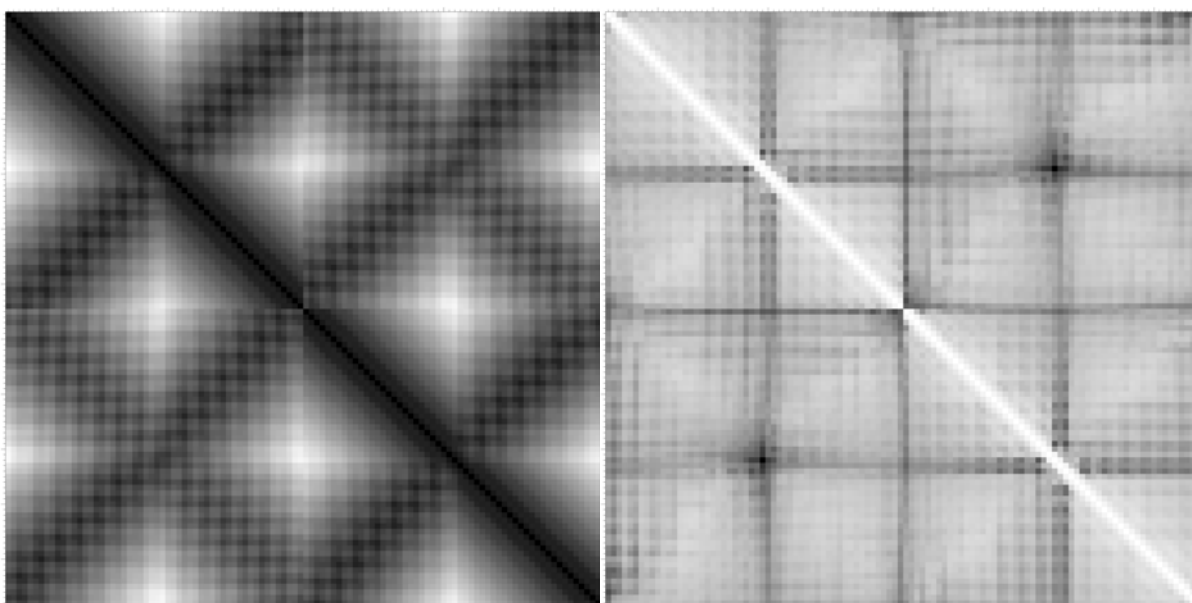


Εικόνα 4-24: R_G των μονομερών A (με μπλε χρώμα) και B (με κίτρινο χρώμα).

4.1.5 Χάρτης μέσης απόστασης μεταξύ των Ca ατόμων και το αντίστοιχο RMSD

Σύμφωνα με όσα αναφέρθηκαν στην αρχή του κεφαλαίου, είναι δυνατόν να φτιάξουμε έναν πίνακα που να περιλαμβάνει τις μέσες αποστάσεις μεταξύ των ατόμων Ca. Ένας τέτοιος πίνακας φαίνεται στην **Εικόνα 4-25**. Η οριζόντια γραμμή στην πάνω πλευρά και η κάθετη γραμμή στην αριστερή πλευρά αντιστοιχούν στα άτομα Ca των καταλοίπων των δύο μονομερών A και B. Βλέπουμε ότι ο πίνακας λόγω συμμετρίας χωρίζεται σε τετράγωνα. Το πρώτο τετράγωνο αντιστοιχεί στο πρώτο μονομερές και το δεύτερο στο δεύτερο μονομερές, ενώ τα κατάλοιπα της στροφής βρίσκονται στο κέντρο του τετραγώνου. Οι μικρότερες αποστάσεις μεταξύ των Ca ατόμων απεικονίζονται με μαύρο χρώμα και οι μεγαλύτερες με λευκό χρώμα. Η μέγιστη απόσταση υπολογίζεται στα 45,23Å και εμφανίζεται στα άτομα της στροφής του μονομερούς A σε σχέση με το μονομερές B.

Στην **Εικόνα 4-26** βλέπουμε μια γραφική αναπαράσταση των RMSD που αντιστοιχούν στην **Εικόνα 4-25**. Και εδώ ισχύουν όσα και προηγουμένως με τη μόνη διαφορά ότι έχουμε αντίστροφη των χρωμάτων. Η μεγαλύτερη τιμή του RMSD είναι τα 5,2Å. Μπορούμε και εδώ να δούμε ποια κατάλοιπα εμφανίζουν τη μεγαλύτερη κινητικότητα. Πρόκειται για τα κατάλοιπα της στροφής του μονομερούς A σε σχέση με τη στροφή του μονομερούς B, τα οποία δε σχηματίζουν έλικα. Τα κατάλοιπα των υπερκινητικών άκρων δεν υπολογίζονται όπως αναφέρθηκε προηγουμένως.

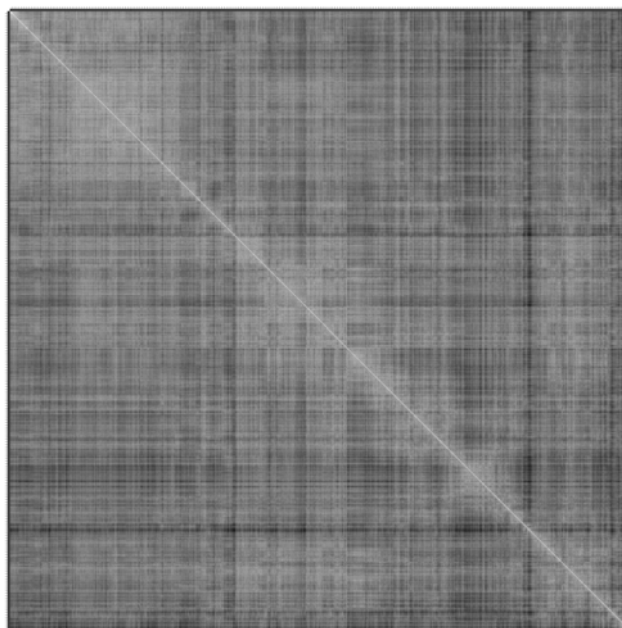


Εικόνα 4-25: Χάρτης αποστάσεων Ca-Ca

Εικόνα 4-26: RMSD που αντιστοιχεί στο χάρτη αποστάσεων Ca-Ca

4.1.6 Frame-to-Frame RMSD

Με το frame-to-frame RMSD έχουμε μια πιο ολοκληρωμένη άποψη για την κινητικότητα του μορίου από εκείνη που μας δίνει το RMSD από την αρχική δομή. Στην **Εικόνα 4-27** απεικονίζεται το frame-to-frame RMSD.

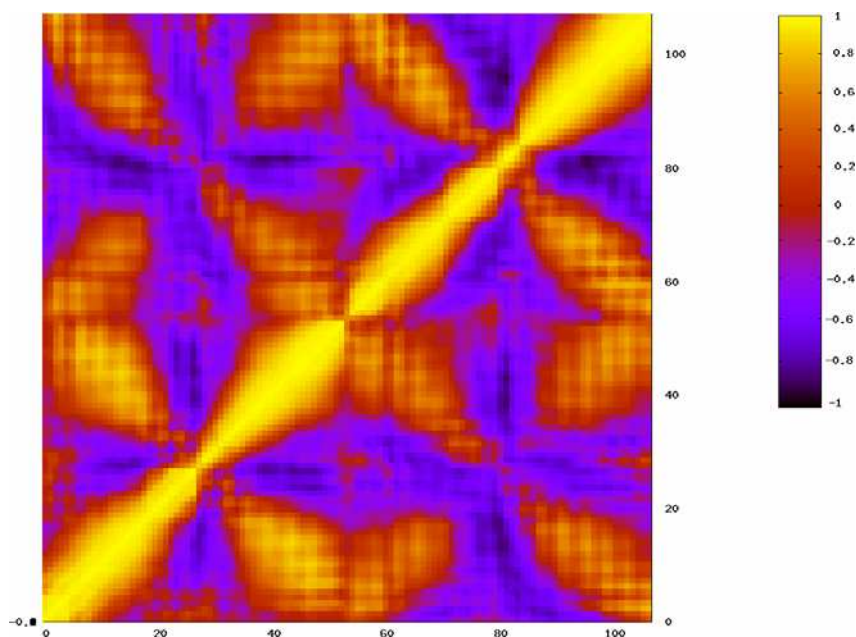


Εικόνα 4-27: Frame-to-Frame RMSD. Στον οριζόντιο και κάθετο άξονα αντιστοιχούν frames και ο υπολογισμός έγινε με βήμα 10.

Κάθε σημείο στον οριζόντιο και κάθετο άξονα αντιστοιχεί σε frames με βήμα 10. Μαζί με τη φάση του heating είναι συνολικά $[2573 \times 2573] = 6.620.329$. Η μέγιστη τιμή είναι $1,5\text{\AA}$ όπου το χρώμα είναι μαύρο, ενώ η ελάχιστη 0\AA όπου το χρώμα είναι άσπρο, πράγμα που φαίνεται και στην **Εικόνα 4-5**. Επιπλέον παρατηρείται και μια ομαδοποίηση των δομών (clusters) την οποία όμως θα περιγράψουμε αναλυτικότερα σε επόμενη ενότητα.

4.1.7 Variance-Covariance και Cross-Correlation

Όπως είδαμε προηγουμένως ο πίνακας των variance-covariance και cross-correlation μας δείχνει ποια τμήματα του μορίου κινούνται προς την ίδια ή προς αντίθετη κατεύθυνση. Ο πίνακας αυτός παρουσιάζεται στην **Εικόνα 4-28**.

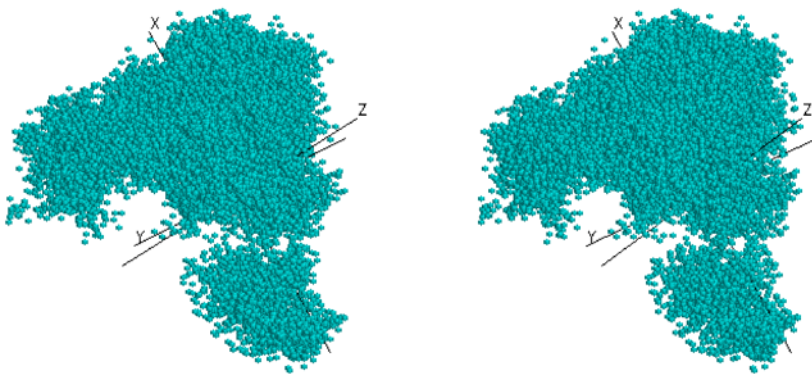
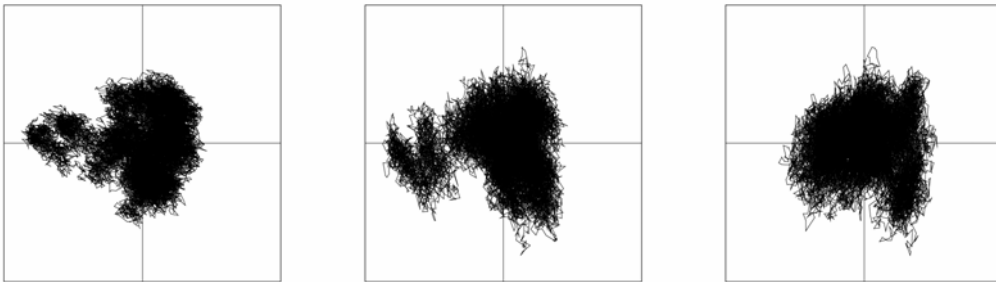
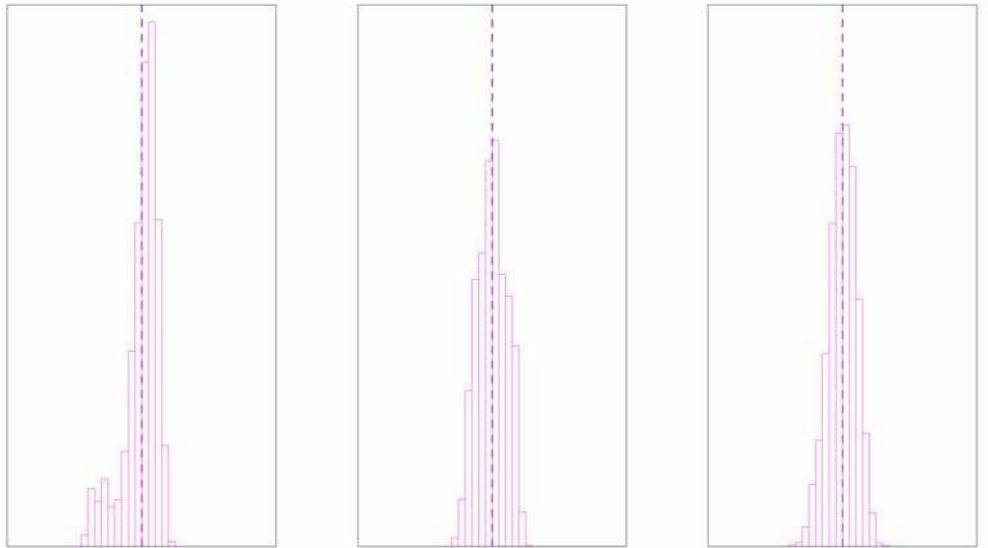


Εικόνα 4-28: Variance-covariance και cross-correlation

Στους δυο άξονες φαίνονται τα κατάλοιπα του Α μονομερούς (1-56) και του Β μονομερούς (57-112). Δεξιά δίπλα από τον πίνακα φαίνεται η χρωματική διαβάθμιση από μαύρο (-1, αρνητική συσχέτιση) σε κίτρινο (+1, θετική συσχέτιση), ενώ εκεί όπου δεν υπάρχει καθόλου συσχέτιση, δηλαδή όπου η τιμή είναι 0 αντιστοιχεί το κόκκινο χρώμα. Η περιοχή με το έντονο κίτρινο (η διαγώνιος) αντιστοιχεί στο ίδιο κατάλοιπο το οποίο προφανώς έχει πλήρη συσχέτιση με τον εαυτό του. Το ίδιο ισχύει και για τα πλαϊνά σε αυτό κατάλοιπα τα οποία είναι ομοιοπολικά συνδεδεμένα με αυτό. Άλλα κατάλοιπα που φαίνεται να συσχετίζεται η κίνησή τους βρίσκονται στα σημεία με το πορτοκαλί χρώμα, που είναι περίπου στα κατάλοιπα που βρίσκονται στις στροφές των μονομερών. Το ίδιο ισχύει και για τα κατάλοιπα της κάθε α έλικας (πορτοκαλί διαγώνιες κάθετες στην κίτρινη διαγώνιο). Επίσης τα κατάλοιπα που δημιουργούν τα επίπεδα ή διπλανά επίπεδα κινούνται μαζί λόγω των αλληλεπιδράσεων που αναπτύσσονται ώστε να σταθεροποιηθεί ο πυρήνας. Οι κινήσεις που δε σχετίζονται καθόλου μεταξύ τους εμφανίζονται μεταξύ των καταλοίπων που ανήκουν σε διαφορετικές έλικες και δεν κινούνται μαζί. Τα αντίστοιχα σημεία απεικονίζονται με μωβ και μαύρο χρώμα, όπως για παράδειγμα τα κατάλοιπα της στροφής του ενός μονομερούς σε σχέση με τα κατάλοιπα της στροφής του άλλου μονομερούς.

4.1.8 Χαρακτηριστικές τιμές (eigenvalues) και χαρακτηριστικά ανύσματα (eigenvectors)

Για να δούμε τις χαρακτηριστικές κινήσεις του μορίου, όπως αναφέρθηκε και στην αρχή του κεφαλαίου, πρέπει να υπολογίσουμε τις χαρακτηριστικές τιμές (eigenvalues) και τα χαρακτηριστικά ανύσματα (eigenvectors). Αρχικά όμως πρέπει να ελέγξουμε αν υπάρχει sufficient sampling, οπότε υπολογίζουμε τις διακυμάνσεις των τριών πρώτων eigenvectors που φαίνεται ότι θα αντιπροσωπεύουν την πλειοψηφία των βασικών κινήσεων του μορίου. Στην **Εικόνα 4-29** έχουμε τη γραφική απεικόνιση των 3 πρώτων eigenvectors. Εάν ο χρόνος της προσομοίωσης είναι αρκετός για sufficient sampling ορισμένων κινήσεων, το ιστόγραμμα θα αντιστοιχεί σε μονοδιάστατη (1D) κανονική κατανομή γύρω από το μηδέν, το επίπεδο που σχηματίζουν οι 3 πρώτοι eigenvectors, συνδυαζόμενοι ανά δυο, θα αντιστοιχεί σε δισδιάστατη (2D) κανονική κατανομή γύρω από τη θέση 0,0 και το επίπεδο που σχηματίζουν οι 3 πρώτοι eigenvectors θα αντιστοιχεί σε τρισδιάστατη (3D) κανονική κατανομή γύρω από τη θέση 0,0,0. Βλέπουμε πως έχει επιτευχθεί σχεδόν sufficient sampling για τον πρώτο eigenvector, όμως βλέπουμε την ύπαρξη ενός μεγάλου cluster και ίσως κάποιου μικρότερου. Κάτι αντίστοιχο συμβαίνει σε μικρότερο βαθμό και με τον δεύτερο eigenvector, ενώ στον τρίτο ο χρόνος της προσομοίωσης φαίνεται πως είναι αρκετός για να τον παρακολουθήσουμε. Στο κάτω μέρος της εικόνας, όπου έχουμε την τρισδιάστατη απεικόνιση της προβολής των 3 eigenvectors φαίνεται πως δημιουργούνται δυο clusters. Όμως η απεικόνιση αυτή δε μπορεί να μας δείξει τη συνεισφορά σε κάθε cluster. Χρειαζόμαστε λοιπόν μια απεικόνιση στην οποία θα επισημαίνεται η πληροφορία σχετικά με την πυκνότητα των σημείων. Κάτι ανάλογο δηλαδή με την **Εικόνα 4-30**, όπου η χρωματική διαβάθμιση από μαύρο (0) σε κίτρινο (90) ακολουθεί την αυξητική πορεία της πυκνότητας των σημείων. Συνεπώς φαίνονται δυο clusters τα οποία σχεδόν επικαλύπτονται ενώ ένα τρίτο μικρότερο δεν εμφανίζει σημαντικό αριθμό σημείων.

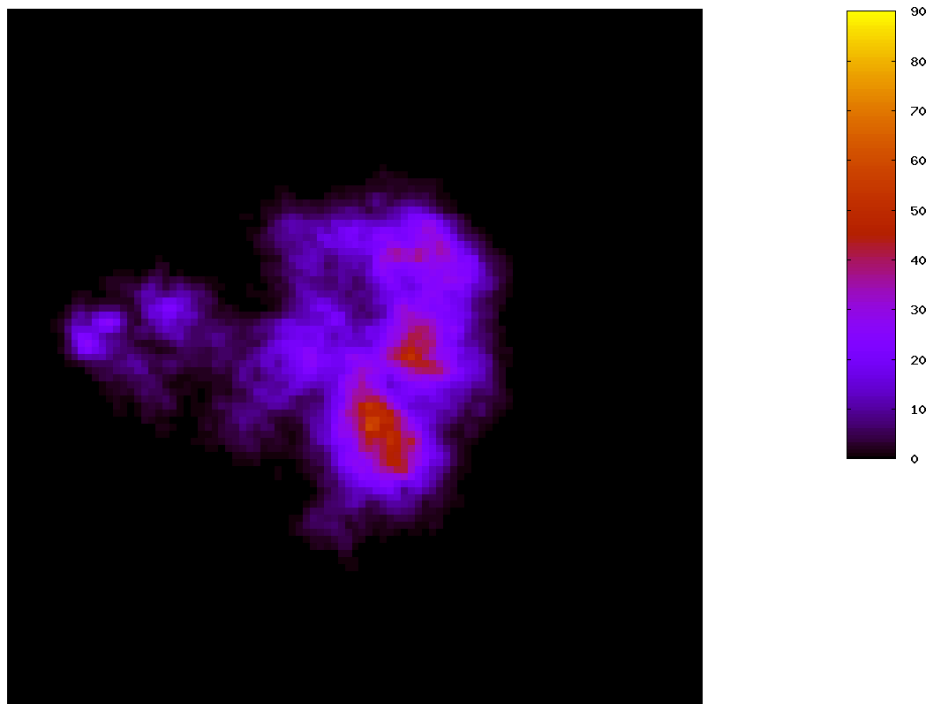


Εικόνα 4-29

Πάνω: Προβολή διακυμάνσεων των 3 πρώτων eigenvectors σε μια διάσταση

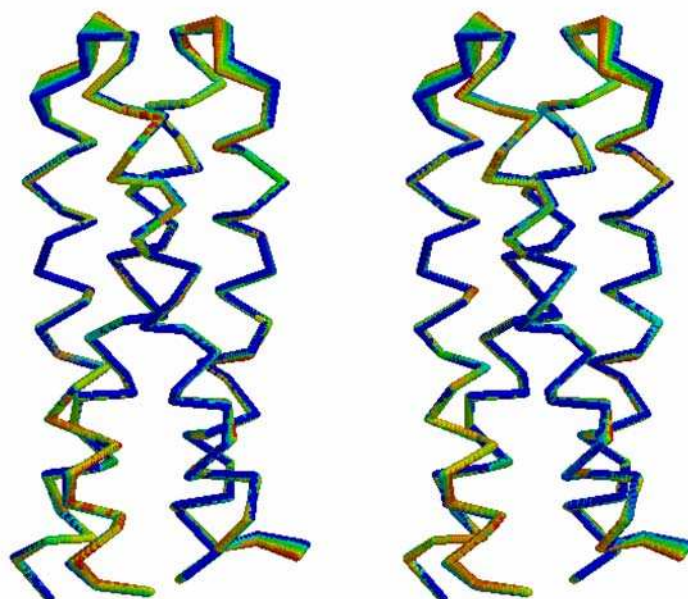
Στη μέση: Προβολή διακυμάνσεων των 3 πρώτων eigenvectors, συνδυασμένων ανά 2 (1-2, 1-3, 2-3 από αριστερά προς δεξιά) στις δυο διαστάσεις

Κάτω: Σχηματικό στερεοδιάγραμμα της προβολής των 3 πρώτων eigenvectors σε τρεις διαστάσεις.



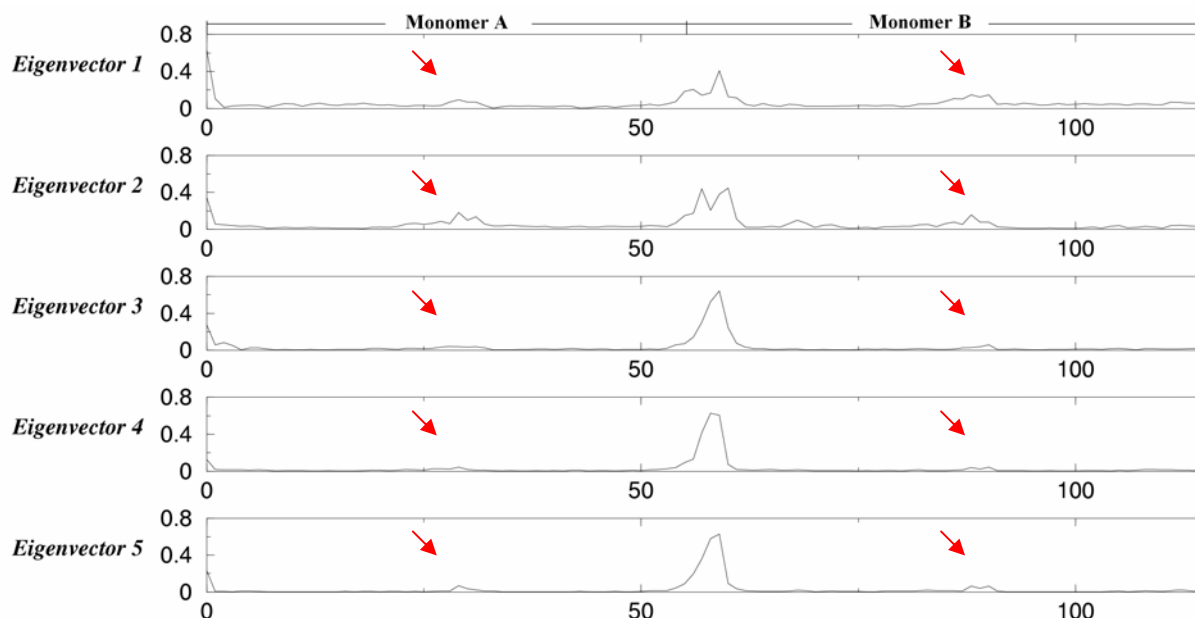
Εικόνα 4-30: Προβολή διακυμάνσεων ανά ζεύγη eigenvectors με βάση την πυκνότητα των σημείων.

Στην **Εικόνα 4-31** μπορούμε να δούμε ένα σχηματικό στερεοδιάγραμμα που αντιστοιχεί στην κίνηση του μορίου βάσει του πρώτου ζεύγους eigenvalue-eigenvector για ολόκληρη τη διάρκεια της προσομοίωσης. Η χρωματική διαβάθμιση από μπλε σε κόκκινο υποδεικνύει το εύρος της κίνησης.



Εικόνα 4-31: Σχηματικό στερεοδιάγραμμα της κίνησης του μορίου, βάσει του πρώτου ζεύγους eigenvalue-eigenvector σε όλη τη διάρκεια του trajectory.

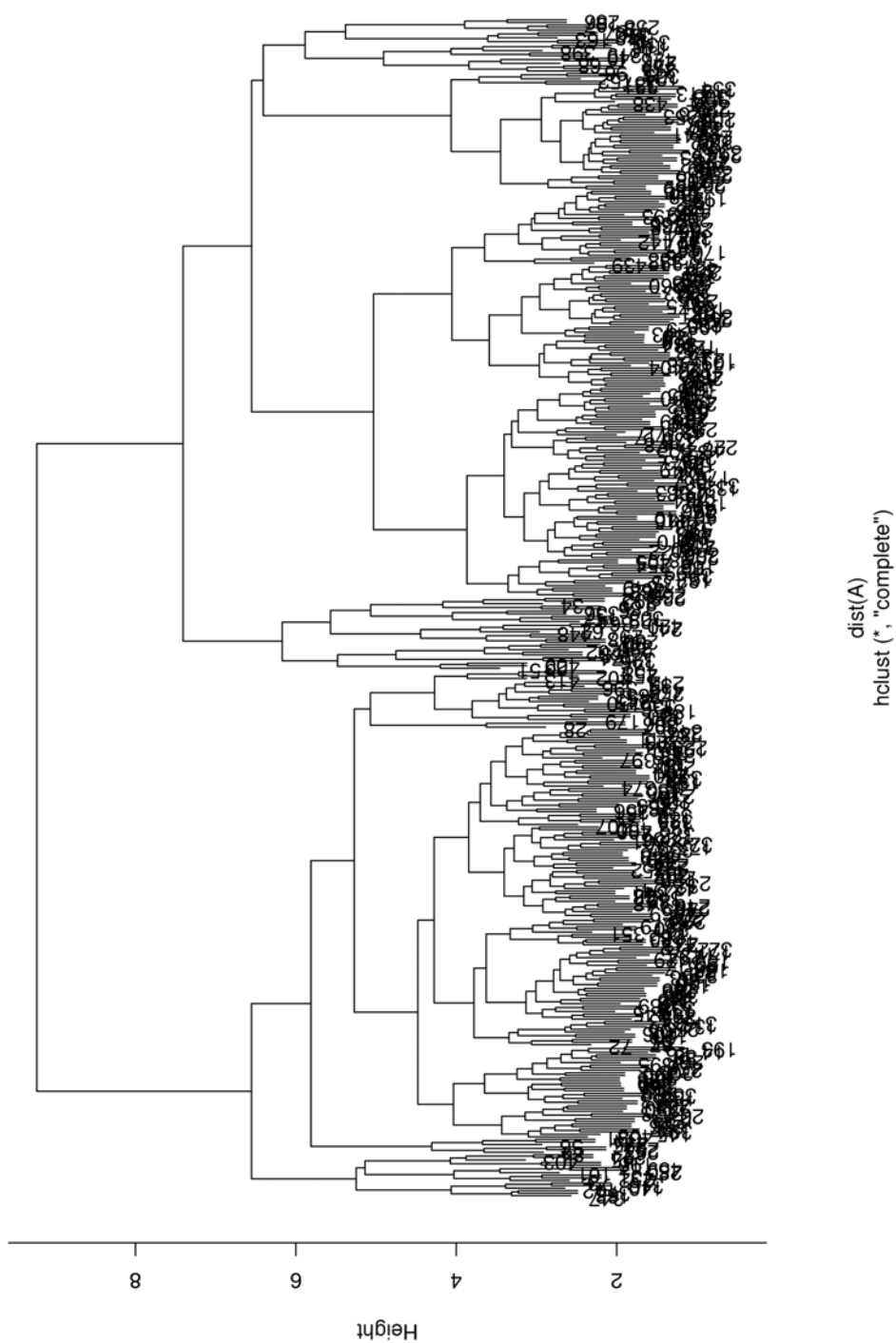
Για να ολοκληρώσουμε τη μελέτη των eigenvectors μένει να μελετήσουμε την κατανομή των διακυμάνσεων ανά κατάλοιπο. Έτσι μπορούμε να έχουμε μια ολοκληρωμένη άποψη για την ποσότητα των σχετιζόμενων διακυμάνσεων του χαρακτηριστικού ανύσματος μεταξύ των καταλοίπων. Στην **Εικόνα 4-32** φαίνεται η διακύμανση των πέντε πρώτων χαρακτηριστικών ανυσμάτων χωριστά για κάθε κατάλοιπο.



Εικόνα 4-32: Αναπαράσταση των διακυμάνσεων των πέντε πρώτων χαρακτηριστικών ανυσμάτων ανά κατάλοιπο. Με κόκκινα βέλη σημαίνονται οι στροφές κάθε μονομερούς.

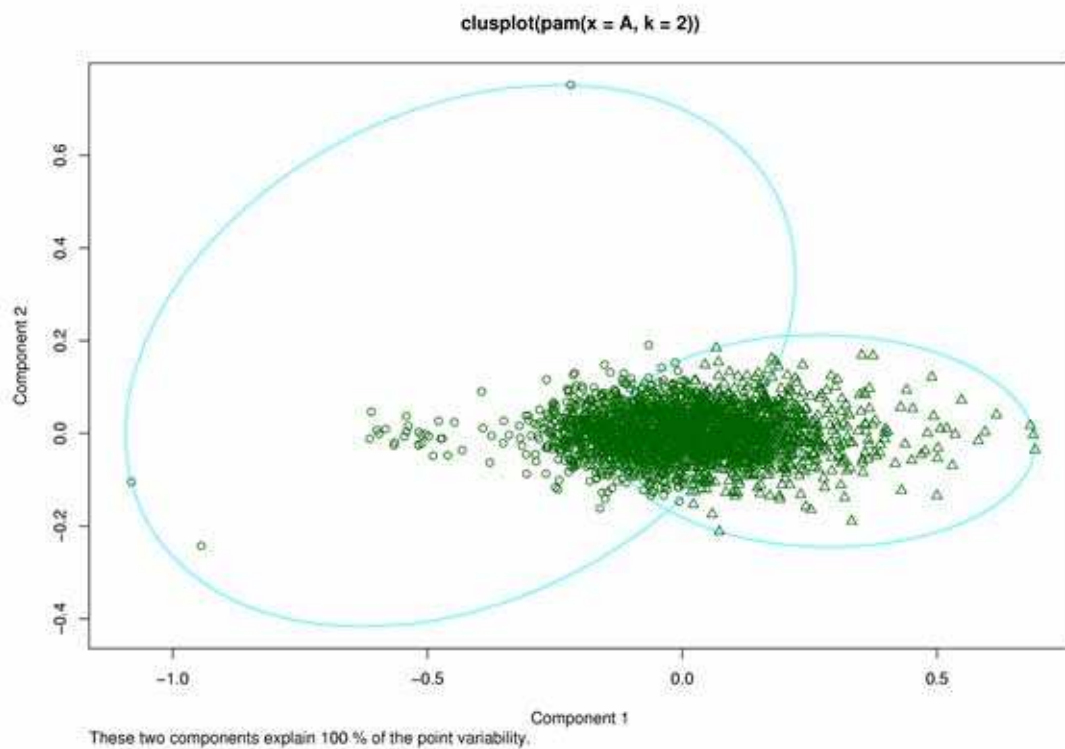
4.1.9 Ομαδοποίηση των δομών της προσομοίωσης (cluster analysis)

Όπως είδαμε και στην αρχή του κεφαλαίου, για να ομαδοποιήσουμε τις δομές του trajectory μπορούμε να φτιάξουμε ένα δενδρόγραμμα, όπως αυτό της **Εικόνας 4-33**, όπου θα λαμβάνεται υπ' όψιν το RMSD ανάμεσα στις διάφορες δομές. Έτσι μας δίνεται η δυνατότητα να προσδιορίσουμε τον αριθμό των κύριων clusters. Για να γίνει κάτι τέτοιο μπορούμε να χρησιμοποιήσουμε το κατάλληλο πρόγραμμα, που θα μας βοηθήσει να πάρουμε τις απαραίτητες πληροφορίες για τα βασικά clusters.



Εικόνα 4-33: Δενδρόγραμμα των δομών από τις οποίες περνά κατά τη διάρκεια του trajectory η Ala_2Ile_2-6 , βάσει του RMSD.

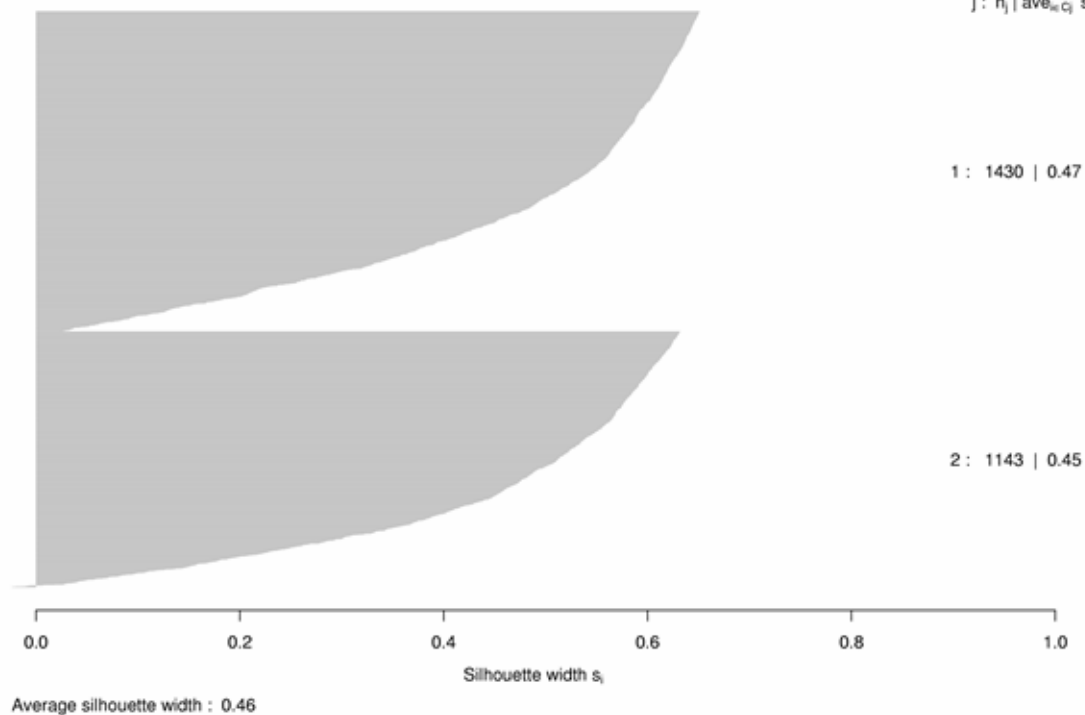
Στην **Εικόνα 4-34** φαίνονται δυο σημαντικές ομάδες δομών οι οποίες επικαλύπτονται, όπως είδαμε και προηγουμένως (Ενότητα 4.1.8). Χρησιμοποιήσαμε τον αλγόριθμο ram για να κάνουμε τους υπολογισμούς.



Silhouette plot of pam(x = A, k = 2)

n = 2573

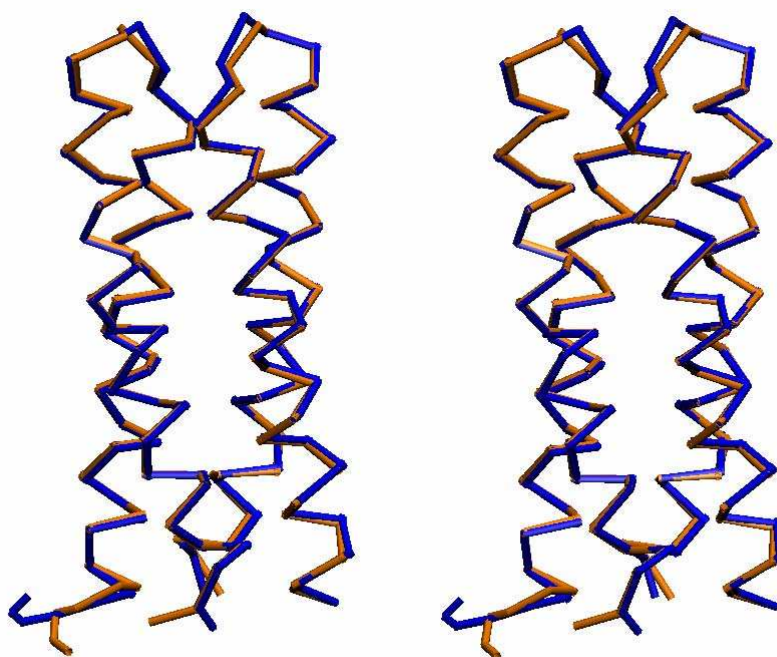
2 clusters C_j
j : n_j | $ave_{i \in C_j} s_i$



Εικόνα 4-34: Πάνω: Προσδιορισμός των clusters **Κάτω:** Κατανομή των δομών στα δυο clusters.

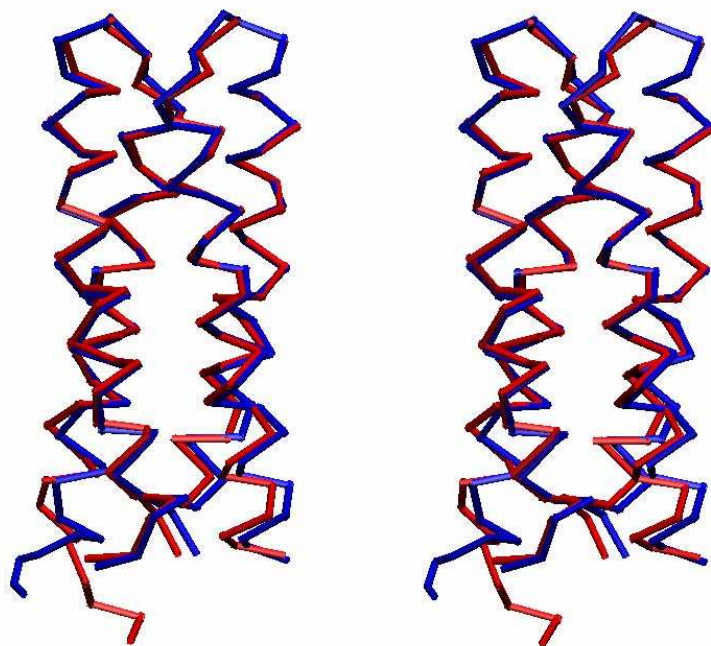
4.1.10 Σύγκριση μεταξύ μέσης δομής των 2 clusters και της αρχικής δομής

Εφόσον προσδιορίστηκε ο αριθμός των clusters ας δούμε αν υπάρχουν διαφορές στη μέση δομή καθενός εκ των δυο αυτών clusters (**Εικόνα 4-35**), εκ των οποίων το πρώτο αντιστοιχεί στα frames 1-12590 (49% του συνόλου των frames) και το δεύτερο στα frames 12960-22540 (50% του συνόλου των frames), καθώς και ανάμεσα στη μέση δομή καθενός cluster και στην αρχική δομή της Ala_2Ile_2-6 με τοπολογία syn (**Εικόνες 4-36 και 4-37**).

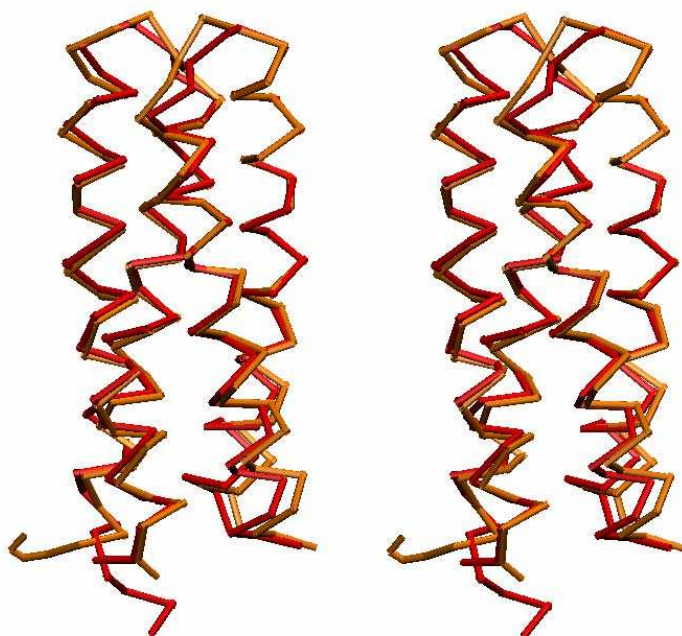


Εικόνα 4-35: Σχηματικό στερεοδιάγραμμα της μέσης δομής του πρώτου (με μπλε χρώμα) και του δεύτερου (με πορτοκαλί χρώμα) cluster σε υπέρθεση.

Στην **Εικόνα 4-35** βλέπουμε πως δεν υπάρχει κάποια σημαντική διαφορά ανάμεσα στις δυο μέσες δομές, εκτός από μια απόκλιση στο πρώτο N-τελικό και τα τρία τελευταία C-τελικά κατάλοιπα της Α έλικας. Στην **Εικόνα 4-36** βλέπουμε πως η μεγαλύτερη απόκλιση μεταξύ της μέσης δομής του πρώτου cluster και της αρχικής δομής της Ala_2Ile_2-6 με τοπολογία syn βρίσκεται στα κατάλοιπα 53-60 της Α έλικας. Τέλος στην **Εικόνα 4-37** παρατηρούμε επίσης τη μεγαλύτερη απόκλιση στην Α έλικα της μέσης δομής του δεύτερου cluster και της αρχικής δομής του μορίου, αυτή τη φορά μεταξύ των καταλοίπων 56-60.



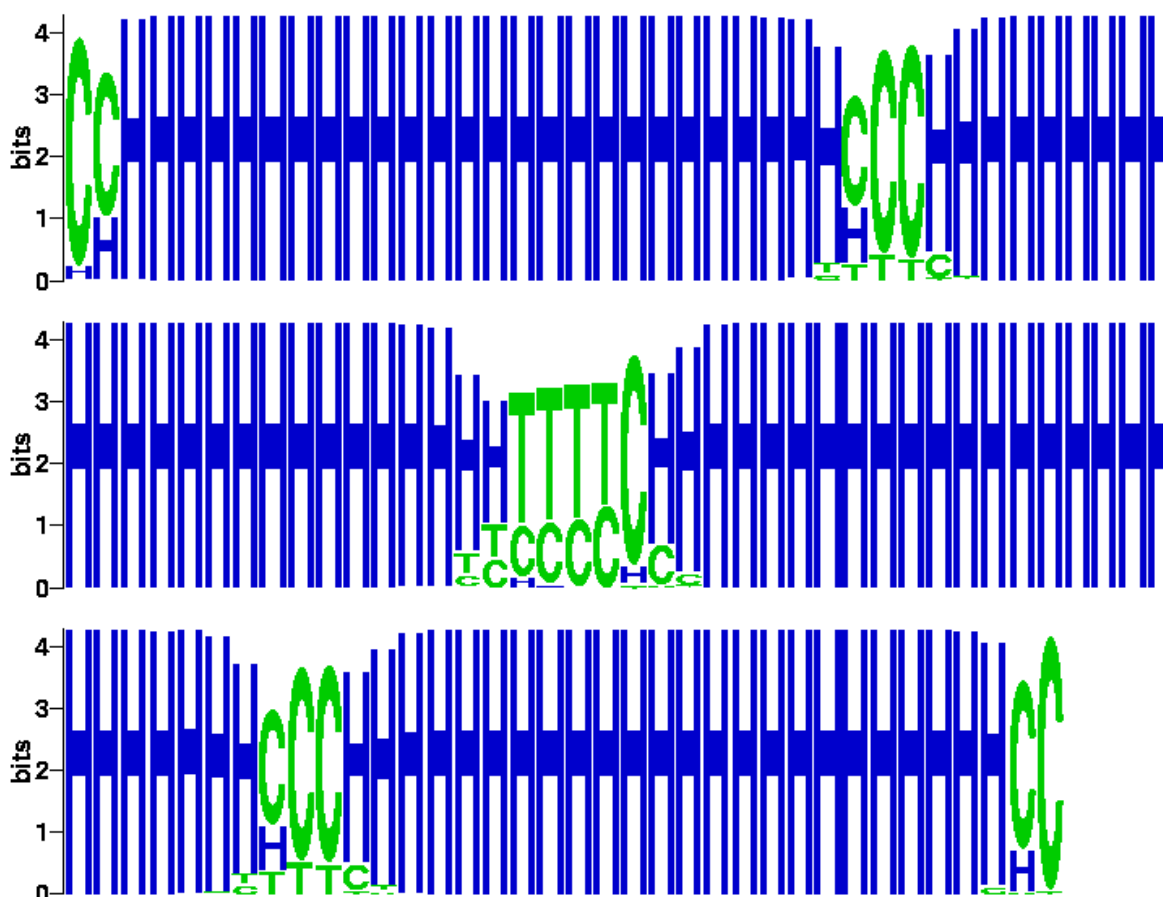
Εικόνα 4-36: Σχηματικό στερεοδιάγραμμα της μέσης δομής του πρώτου (με μπλε χρώμα) cluster και της αρχικής δομής του μορίου (με κόκκινο χρώμα) σε υπέρθεση.



Εικόνα 4-37: Σχηματικό στερεοδιάγραμμα της μέσης δομής του δεύτερου (με πορτοκαλί χρώμα) cluster και της αρχικής δομής του μορίου (με κόκκινο χρώμα) σε υπέρθεση.

4.1.11 Ανάλυση δευτεροταγούς δομής

Στο σημείο αυτό θα μελετήσουμε κατά πόσο διατηρούνται τα στοιχεία δευτεροταγούς δομής, δηλαδή οι α έλικες, κατά τη διάρκεια του trajectory. Για το σκοπό αυτό χρησιμοποιούμε ένα logo (Frishman *et al.*, 1995), που φαίνεται στην **Εικόνα 4-38**.

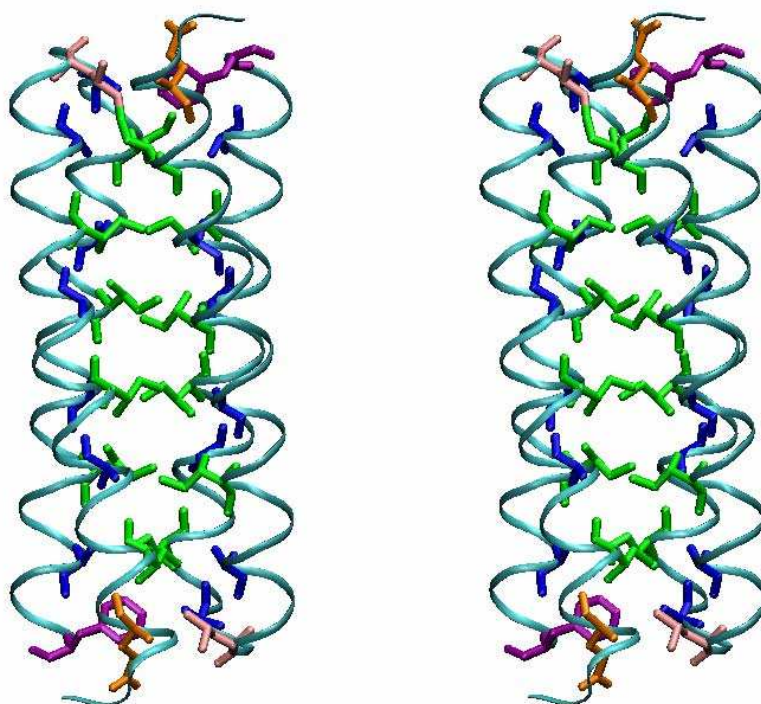


Εικόνα 4-38: Ανάλυση της δευτεροταγούς δομής κατά μήκος της αλληλουχίας

Με Η παριστάνονται οι α έλικες (Helix) και με C οι θηλιές (Coils). Σε κάθε σειρά βρίσκονται 40 θέσεις αλληλουχίας. Οι α έλικες διατηρούνται καθ' όλη τη διάρκεια του trajectory και είναι αρκετά σταθερές. Οι περιοχές που δε σχηματίζουν σταθερή δομή, αλλά θηλιά, αντιστοιχούν στα κατάλοιπα 1-2 του N-τελικού άκρου, 29-31 της στροφής και 57-59 του C-τελικού άκρου του μονομερούς A, καθώς και στα κατάλοιπα 1-2 του N-τελικού άκρου, 29-31 της στροφής και 56-57 του C-τελικού άκρου του μονομερούς B.

4.2 Ανάλυση προσομοίωσης διμερούς της Ala_2Ple_2-6 με τοπολογία anti (Rop-like)

Όπως είδαμε και στην ενότητα 3.2 για να δημιουργήσουμε το διμερές Ala_2Ple_2-6 με τοπολογία όμοια με εκείνη της φυσικής Rop, χρησιμοποιήσαμε αρχικά τις συντεταγμένες της φυσικής Rop (1rop.pdb), τις οποίες βρήκαμε στη βάση δεδομένων της PDB. Στη συνέχεια με τη βοήθεια του προγράμματος X-fit τροποποιήσαμε την αλληλουχία ώστε να πάρουμε τελικά την Ala_2Ple_2-6 με τοπολογία anti και βελτιστοποιήσαμε τη γεωμετρία του μορίου ώστε να μην παρεμποδίζει η μια πλευρική αλυσίδα την άλλη στα στρώματα του υδρόφοβου πυρήνα. Εδώ ξεκινάμε με όλα τα στρώματα να συμμετέχουν στην οργάνωση του υδρόφοβου πυρήνα, εφόσον χρησιμοποιήσαμε σαν εκμαγείο την τοπολογία της φυσικής Rop. Προέκυψε λοιπόν η δομή που θα χρησιμοποιήσουμε για την ανάλυση (**Εικόνα 4-39**).

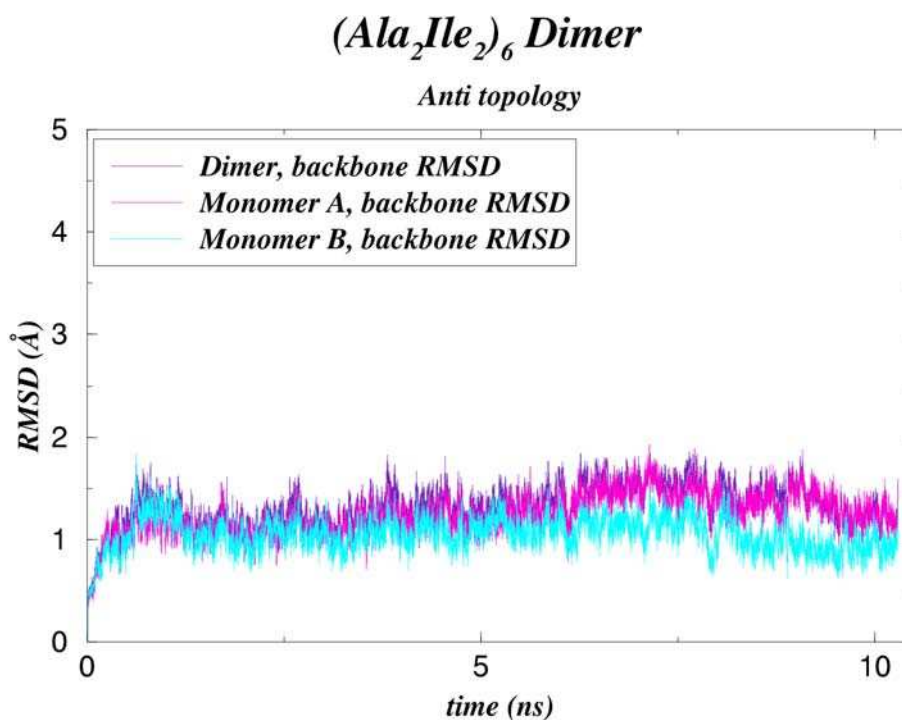


Εικόνα 4-39: Σχηματικό στερεοδιάγραμμα της Ala_2Ple_2-6 με τοπολογία της φυσικής Rop. Με κίτρινο χρώμα βλέπουμε το σκελετό της δομής, ενώ με μπλε χρώμα τα κατάλοιπα Ala, με πράσινο τα κατάλοιπα Ple, με μωβ τα κατάλοιπα Phe, με πορτοκαλί τα κατάλοιπα Gln και με κόκκινο τα κατάλοιπα Leu, που αποτελούν τα οκτώ στρώματα (layers) του σχηματιζόμενου υδρόφοβου πυρήνα.

4.2.1 RMSD (τυπική απόκλιση) από την αρχική δομή

Στο σημείο αυτό θα μελετήσουμε την τυπική απόκλιση του συστήματος από την αρχική του δομή, κατά πόσο δηλαδή παρεκκλίνει η δομή από την αρχή του trajectory μέχρι το τέλος του, σε συνάρτηση με το χρόνο.

Χρησιμοποιώντας, λοιπόν, τα προγράμματα X-plor (Brunger, 1992) και Xmgr, πήραμε την **Εικόνα 4-40**, στην οποία βλέπουμε μια καμπύλη η οποία απεικονίζει την τυπική απόκλιση (Å), συναρτήσει του χρόνου (ns). Η μέτρηση έχει γίνει με βάση τα άτομα του σκελετού backbone) του συστήματος Ala₂Ile₂-6 και ως μέτρο σύγκρισης χρησιμοποιείται η αρχική δομή. Στην **Εικόνα 4-40** παρατηρούμε ότι η τυπική απόκλιση κυμαίνεται γύρω από μια τιμή περίπου 1,3Å σε όλη τη διάρκεια του trajectory.



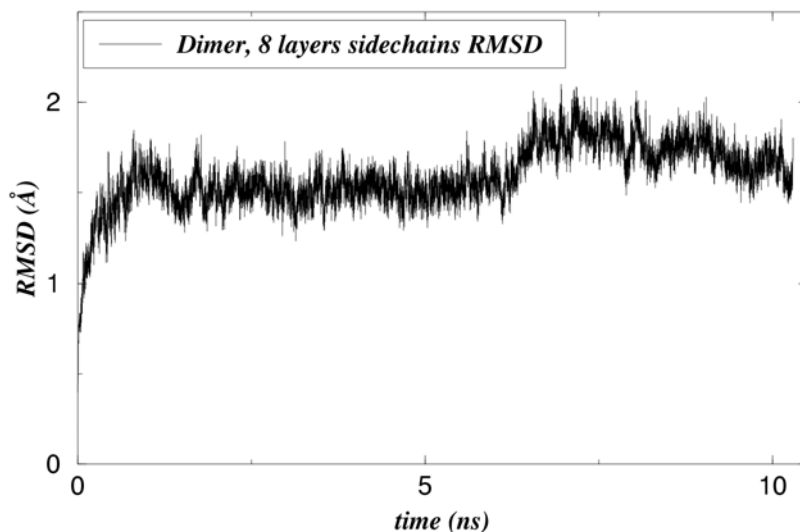
Εικόνα 4-40: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται όλα τα άτομα του σκελετού του συστήματος. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B.

Αν παρατηρήσουμε όμως τη διακύμανση από την αρχική δομή των πλευρικών ομάδων των καταλοίπων που σχηματίζουν τα στρώματα του υδρόφοβου πυρήνα (**Εικόνες 4-41, 4-42**), θα δούμε πως υπάρχει μια αύξηση της διακύμανσης από το έκτο nsec της προσομοίωσης μέχρι το τέλος της. Αρχικά η διακύμανση κυμαίνεται γύρω από τα 1,5Å και στη συνέχεια γύρω από τα 1,8Å. Η αύξηση δεν είναι μεγάλη και μπορεί να εξηγηθεί από την άποψη ότι οι πλευρικές ομάδες έχουν τοποθετηθεί από εμάς κατά προσέγγιση και είναι λογικό να παρουσιάζουν μια σχετική κινητικότητα. Αξιοσημείωτο είναι και εδώ το γεγονός ότι τη διακύμανση του

διμερούς επηρεάζει η διακύμανση του μονομερούς A, ενώ η διακύμανση του μονομερούς B είναι αρκετά σταθερή.

(Ala₂Ile₂)₆ Dimer

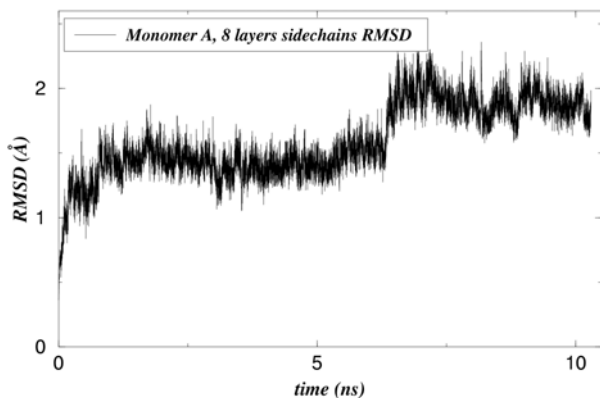
Anti topology



Εικόνα 4-41: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται τα άτομα των πλευρικών ομάδων των στρωμάτων που δημιουργούν τον υδρόφοβο πυρήνα στο διμερές.

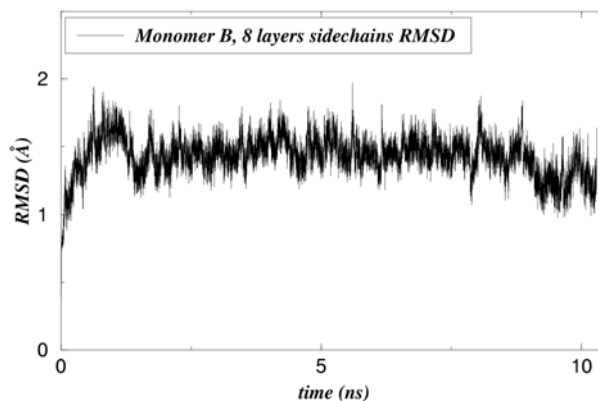
(Ala₂Ile₂)₆ Dimer

Anti topology



(Ala₂Ile₂)₆ Dimer

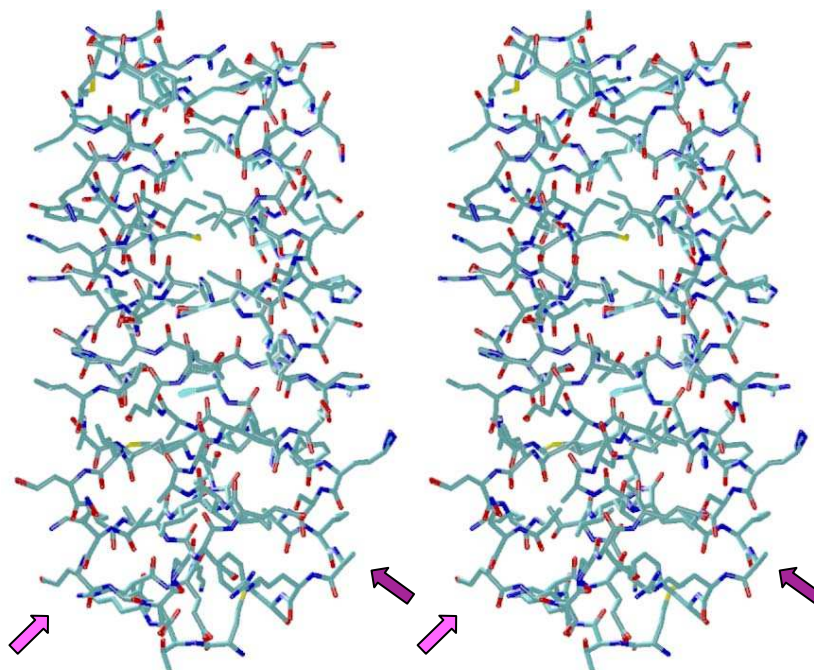
Anti topology



Εικόνα 4-42: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται τα άτομα των πλευρικών ομάδων των στρωμάτων που δημιουργούν τον υδρόφοβο πυρήνα στο μονομερές A αριστερά και στο μονομερές B δεξιά.

4.2.2 Η μέση δομή και οι αντίστοιχες διακυμάνσεις

Ύστερα από επεξεργασία ολόκληρου του trajectory με τη βοήθεια του προγράμματος X-plor, καταφέραμε να ανακτήσουμε ένα αρχείο pdb στο οποίο περιέχονται οι συντεταγμένες ενός διμερούς που αντιστοιχεί σε μια μέση δομή του αρχικού συστήματος. Ύστερα από μια επεξεργασία του αρχείου αυτού με το πρόγραμμα VMD, δημιουργήσαμε ένα στερεοδιάγραμμα της μέσης αυτής δομής (**Εικόνα 4-43**).



Εικόνα 4-43: Σχηματικό στερεοδιάγραμμα της μέσης δομής του συστήματος Ala₂Ile₂-6 με τοπολογία anti. Αριστερά με το ροζ βέλος βρίσκεται η αλυσίδα A και δεξιά με το μοβ βέλος η αλυσίδα B.

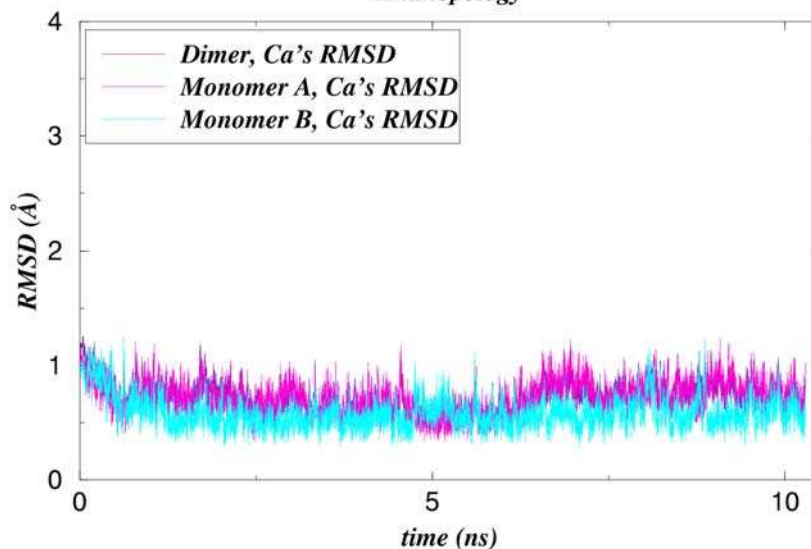
Στην εικόνα διακρίνουμε την αλυσίδα A αριστερά και τη B δεξιά, σε κάθε τμήμα του στερεοδιαγράμματος. Δεν υπάρχουν πολλές διαφορές όσο αφορά στη δομή του συστήματος σε σχέση με την αρχική.

4.2.2.1 RMSD από τη μέση δομή

Στο αρχείο pdb που δημιουργήσαμε, υπάρχει μια στήλη η οποία αντιστοιχεί στα RMSF των ατόμων. Από τη στήλη αυτή κρατήσαμε τα RMSF μόνο των ατόμων Ca για να υπολογίσουμε τη διακύμανση για καθένα από τα άτομα αυτά. Στην **Εικόνα 4-44** μπορούμε να δούμε το RMSD από τη μέση δομή για τα Ca άτομα τόσο του διμερούς, όσο και των μονομερών A και B.

(Ala₂Ile₂)₆ Dimer

Anti topology



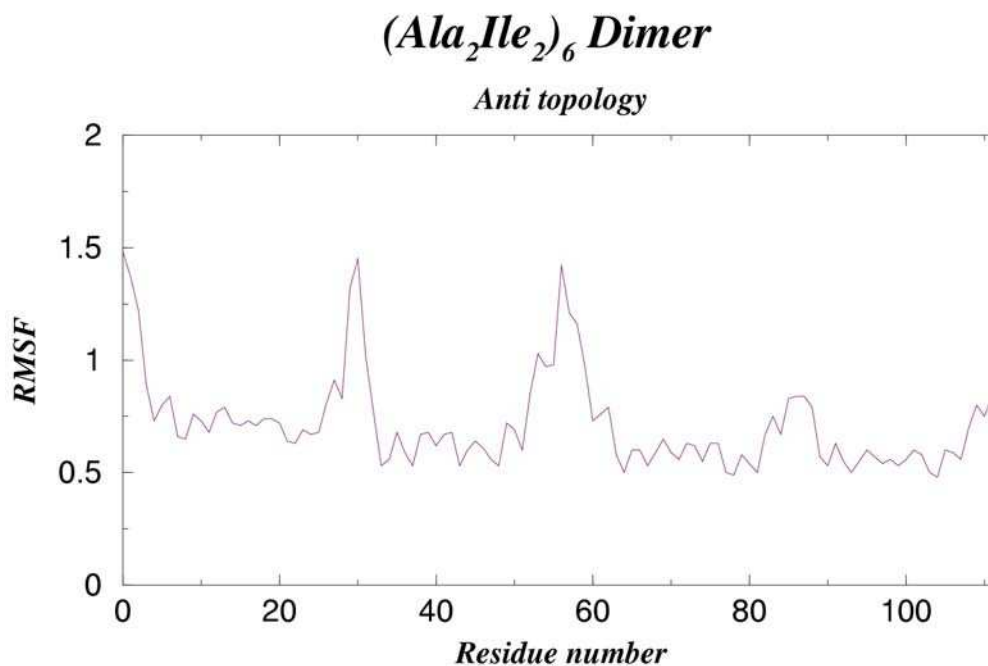
Εικόνα 4-44: Καμπύλη τυπικής απόκλισης RMSD (Å) από τη μέση δομή συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα Ca, τόσο του διμερούς, όσο και των δυο μονομερών. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κωανή στο μονομερές B. Η καμπύλη του διμερούς σχεδόν συμπίπτει με την καμπύλη του μονομερούς A.

Η τυπική απόκλιση για το διμερές και το μονομερές A είναι σχεδόν ταυτόσημες από 0,5Å έως 1,1Å, με μέγιστο την αρχική τιμή των 1,1Å και ελάχιστο τα 0,5Å στα 2,5 ns περίπου του trajectory. Το μονομερές B και εδώ δεν επηρεάζεται όσο αφορά τη διακύμανση, η τιμή της οποίας κινείται γύρω από τα 0,5Å σε όλη τη διάρκεια του trajectory, με ένα τοπικό μέγιστο στα 1,2Å στην αρχή του trajectory.

4.2.3 RMSF

Το RMSF είναι μια παράμετρος που αντικατοπτρίζει τη διακύμανση ενός ατόμου σε σχέση με τη θέση του στη μέση δομή. Στην **Εικόνα 4-45** μπορούμε να δούμε τα κατάλοιπα που παρουσιάζουν τη μεγαλύτερη κινητικότητα (↑διακύμανση → ↑κινητικότητα).

Εικόνα 4-45: RMSF των ατόμων του διμερούς σε σχέση με τη θέση τους στη μέση δομή. Το RMSF υπολογίζεται σε Å.

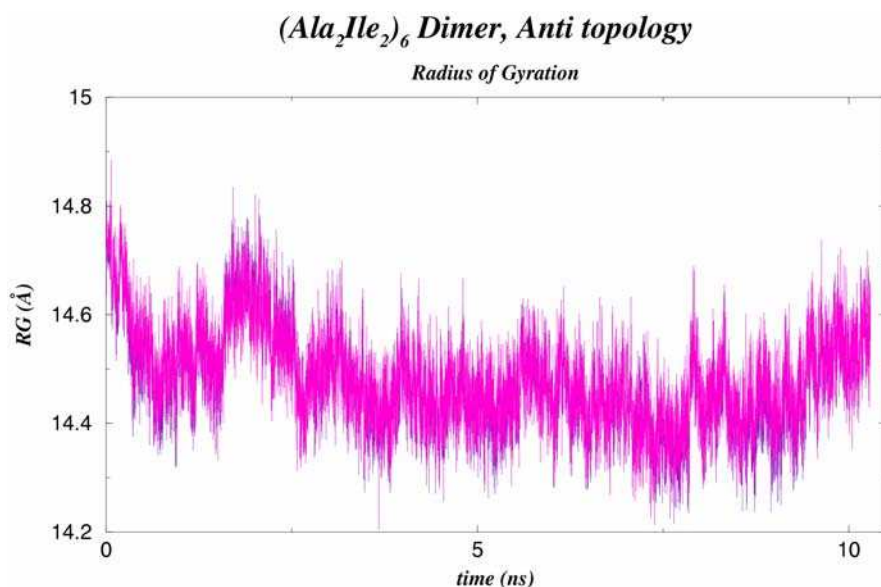


θέση τους στη μέση δομή. Το RMSF υπολογίζεται σε Å.

Τις υψηλότερες τιμές εμφανίζουν τα κατάλοιπα στην αρχή (1-2), στις στροφές (28-31) και στο τέλος (54-56) κάθε μονομερούς. Τα κατάλοιπα αυτά όπως είδαμε και στο κεφάλαιο 1 δε συμμετέχουν στο σχηματισμό έλικας και αποτελούν μέρος θηλιάς, επομένως δικαιολογημένα εμφανίζουν το μεγαλύτερο βαθμό κινητικότητας σε σύγκριση με τα υπόλοιπα κατάλοιπα. Παρόλα αυτά πρέπει να επισημάνουμε ότι η κινητικότητα είναι σχετικά αυξημένη σε σύγκριση με την κινητικότητα που παρατηρείται στη δομή της Ala₂Ile₂-6 με τοπολογία syn.

4.2.4 Εξέλιξη του R_G, Radius of Gyration

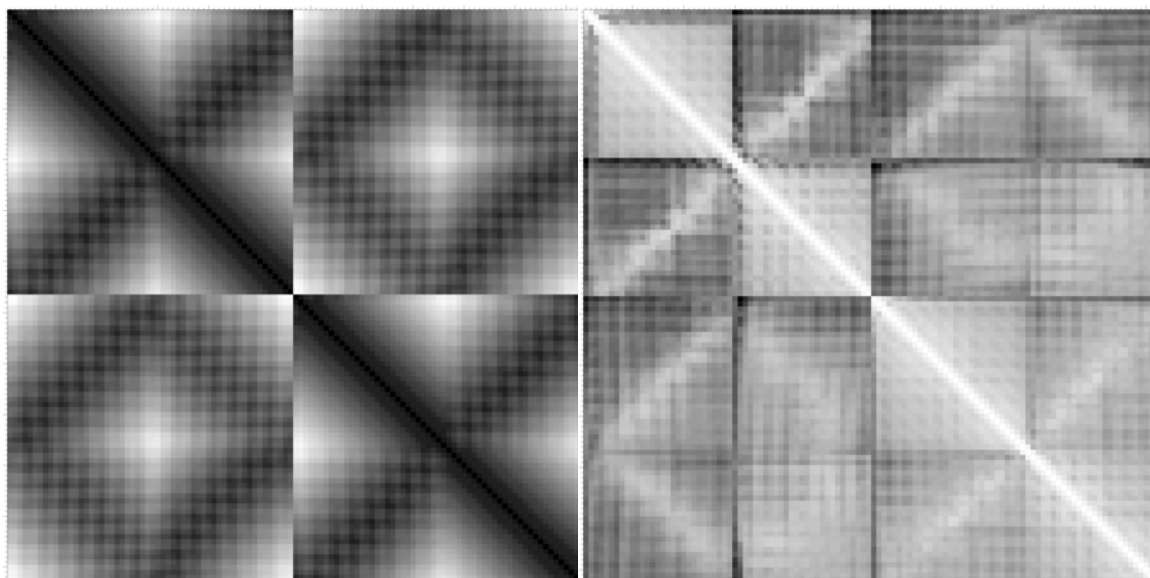
Στην εικόνα 4-46 μπορούμε να δούμε την εξέλιξη του R_G των μονομερών A και B της Ala₂Ile₂-6 με τοπολογία anti. Η τιμή του R_G κυμαίνεται για το μονομερές A γύρω από τα 14,6Å και για το μονομερές B γύρω από τα 14,5Å και δε φαίνεται να παρουσιάζει σημαντικές διακυμάνσεις, αν εξαιρέσουμε τη μεγαλύτερη τιμή που αγγίζει τα 14,9Å στην αρχή του trajectory και τη μικρότερη που αγγίζει τα 14,2Å περίπου στα 3,2ns του trajectory. Συνεπώς παρατηρούμε πως το μόριο είναι αρκετά σταθερό.



Εικόνα 4-46: R_G των μονομερών A (με μωβ χρώμα) και B (με φούζια χρώμα). Παρατηρούμε πως οι τιμές της διακύμανσης σχεδόν συμπίπτουν για τα δυο μονομερή.

4.2.5 Χάρτης μέσης απόστασης μεταξύ των Ca ατόμων και το αντίστοιχο RMSD

Σύμφωνα με όσα αναφέρθηκαν στην αρχή του κεφαλαίου, είναι δυνατόν να φτιάξουμε έναν πίνακα που να περιλαμβάνει τις μέσες αποστάσεις μεταξύ των ατόμων Ca. Ένας τέτοιος πίνακας φαίνεται στην **Εικόνα 4-47**.



Εικόνα 4-47: Χάρτης αποστάσεων Ca-Ca

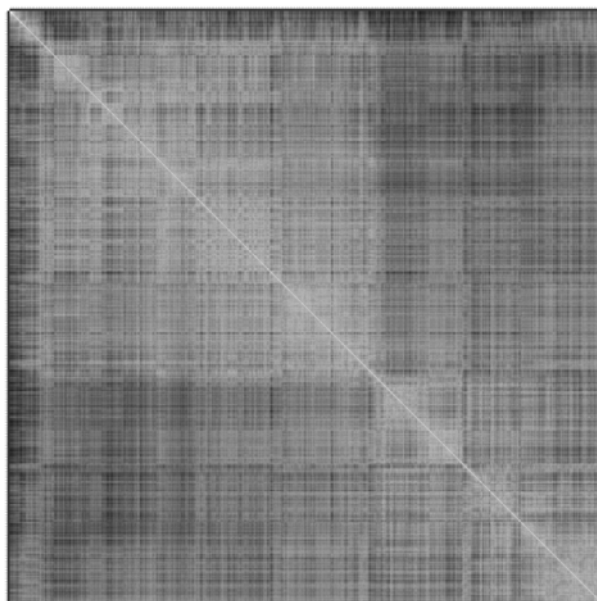
Εικόνα 4-48: RMSD που αντιστοιχεί στο χάρτη αποστάσεων Ca-Ca

Η οριζόντια γραμμή στην πάνω πλευρά και η κάθετη γραμμή στην αριστερή πλευρά αντιστοιχούν στα άτομα Ca των καταλοίπων των δύο μονομερών A και B. Βλέπουμε ότι ο πίνακας λόγω συμμετρίας χωρίζεται σε τετράγωνα. Το πρώτο τετράγωνο αντιστοιχεί στο πρώτο μονομερές και το δεύτερο στο δεύτερο μονομερές, ενώ τα κατάλοιπα της στροφής βρίσκονται στο κέντρο του τετραγώνου. Οι μικρότερες αποστάσεις μεταξύ των Ca ατόμων απεικονίζονται με μαύρο χρώμα και οι μεγαλύτερες με λευκό χρώμα. Η μέγιστη απόσταση εμφανίζεται στα άτομα της στροφής του μονομερούς A σε σχέση με το μονομερές B.

Στην **Εικόνα 4-48** βλέπουμε μια γραφική αναπαράσταση των RMSD που αντιστοιχούν στην **Εικόνα 4-47**. Και εδώ ισχύουν όσα και προηγουμένως με τη μόνη διαφορά ότι έχουμε αντίστροφη των χρωμάτων. Μπορούμε και εδώ να δούμε ποια κατάλοιπα εμφανίζουν τη μεγαλύτερη κινητικότητα. Πρόκειται για τα κατάλοιπα της στροφής του μονομερούς A σε σχέση με τη στροφή του μονομερούς B, τα οποία δε σχηματίζουν έλικα. Τα κατάλοιπα των υπερκινητικών άκρων δεν υπολογίζονται όπως αναφέρθηκε προηγουμένως.

4.2.6 Frame-to-Frame RMSD

Με το frame-to-frame RMSD έχουμε μια πιο ολοκληρωμένη άποψη για την κινητικότητα του μορίου από εκείνη που μας δίνει το RMSD από την αρχική δομή. Στην **Εικόνα 4-49** απεικονίζεται το frame-to-frame RMSD.



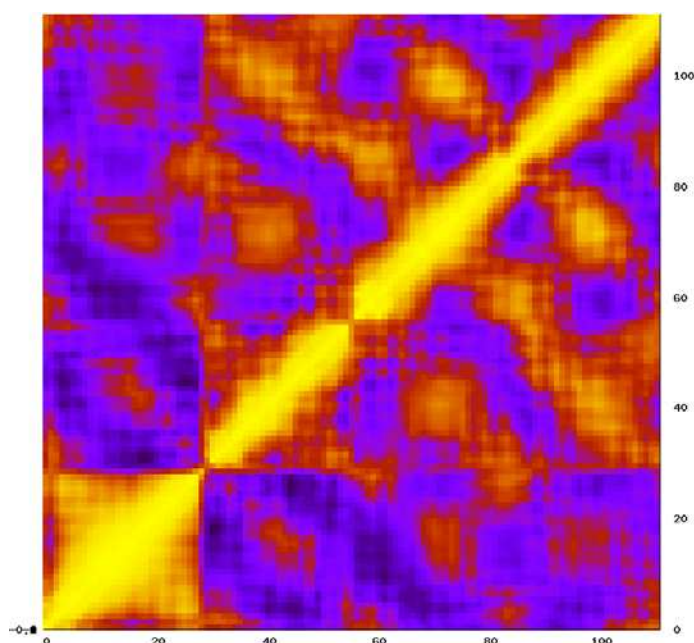
Εικόνα 4-49: *Frame-to-Frame RMSD*. Στον οριζόντιο και κάθετο άξονα αντιστοιχούν frames και ο υπολογισμός έγινε με βήμα 10.

Κάθε σημείο στον οριζόντιο και κάθετο άξονα αντιστοιχεί σε frames με βήμα 10. Μαζί με τη φάση του heating είναι συνολικά $[2573 \times 2573] = 6.620.329$. Στον οριζόντιο και κάθετο άξονα αντιστοιχούν τα frames τα οποία μαζί με τη φάση του

heating είναι συνολικά 6.620.329. Η μέγιστη τιμή είναι 1,9Å όπου το χρώμα είναι μαύρο, ενώ η ελάχιστη 0Å όπου το χρώμα είναι άσπρο. Επιπλέον παρατηρείται και μια ομαδοποίηση των δομών (clusters) την οποία όμως θα περιγράψουμε αναλυτικότερα σε επόμενη ενότητα.

4.2.7 Variance-Covariance και Cross-Correlation

Όπως είδαμε προηγουμένως ο πίνακας των variance-covariance και cross-correlation μας δείχνει ποια τμήματα του μορίου κινούνται προς την ίδια ή προς αντίθετη κατεύθυνση. Ο πίνακας αυτός παρουσιάζεται στην **Εικόνα 4-50**.



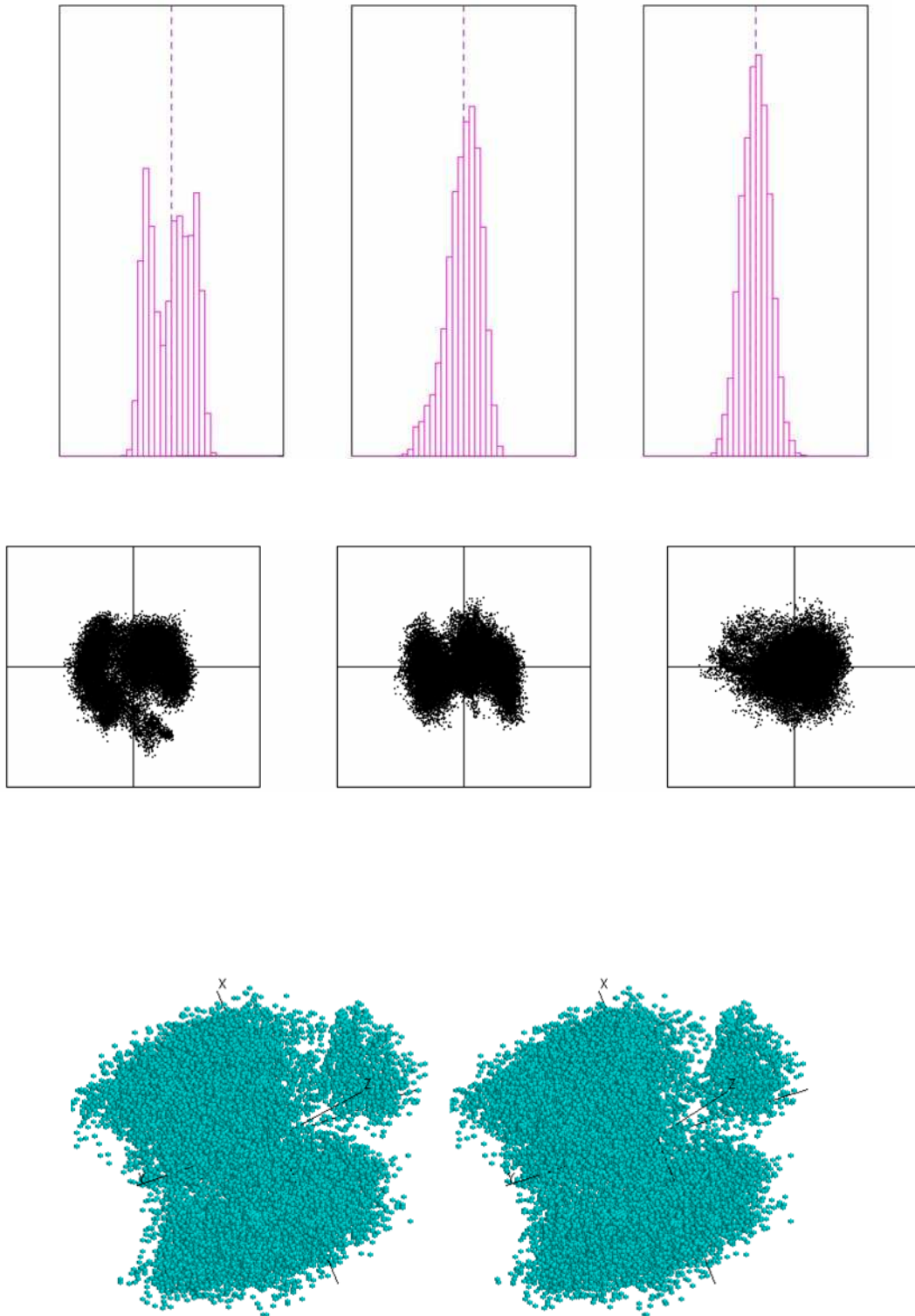
Εικόνα 4-50: Variance-covariance και cross-correlation

Στους δυο άξονες φαίνονται τα κατάλοιπα του Α μονομερούς (1-56) και του Β μονομερούς (57-112). Δεξιά δίπλα από τον πίνακα φαίνεται η χρωματική διαβάθμιση από μαύρο (-1, αρνητική συσχέτιση) σε κίτρινο (+1, θετική συσχέτιση), ενώ εκεί όπου δεν υπάρχει καθόλου συσχέτιση, δηλαδή όπου η τιμή είναι 0 αντιστοιχεί το κόκκινο χρώμα. Η περιοχή με το έντονο κίτρινο (η διαγώνιος) αντιστοιχεί στο ίδιο κατάλοιπο το οποίο προφανώς έχει πλήρη συσχέτιση με τον εαυτό του. Το ίδιο ισχύει και για τα πλαϊνά σε αυτό κατάλοιπα τα οποία είναι ομοιοπολικά συνδεδεμένα με αυτό. Άλλα κατάλοιπα που φαίνεται να συσχετίζεται η κίνησή τους βρίσκονται στα σημεία με το πορτοκαλί χρώμα, που είναι περίπου στα κατάλοιπα που βρίσκονται στις στροφές των μονομερών. Το ίδιο ισχύει και για τα κατάλοιπα της κάθε α έλικας (πορτοκαλί διαγώνιες κάθετες στην κίτρινη διαγώνιο).

Επίσης τα κατάλοιπα που δημιουργούν τα επίπεδα ή διπλανά επίπεδα κινούνται μαζί λόγω των αλληλεπιδράσεων που αναπτύσσονται ώστε να σταθεροποιηθεί ο πυρήνας. Οι κινήσεις που δε σχετίζονται καθόλου μεταξύ τους εμφανίζονται μεταξύ των καταλοίπων που ανήκουν σε διαφορετικές έλικες και δεν κινούνται μαζί. Τα αντίστοιχα σημεία απεικονίζονται με μωβ και μαύρο χρώμα. Είναι αξιοσημείωτο το γεγονός ότι τα κατάλοιπα της πρώτης έλικας του μονομερούς A κινούνται σε απόλυτη συσχέτιση μεταξύ τους. Είναι εμφανής επίσης και η δημιουργία clusters.

4.2.8 Χαρακτηριστικές τιμές (eigenvalues) και χαρακτηριστικά ανύσματα (eigenvectors)

Για να δούμε τις χαρακτηριστικές κινήσεις του μορίου, όπως αναφέρθηκε και στην αρχή του κεφαλαίου, πρέπει να υπολογίσουμε τις χαρακτηριστικές τιμές (eigenvalues) και τα χαρακτηριστικά ανύσματα (eigenvectors). Αρχικά όμως πρέπει να ελέγξουμε αν υπάρχει sufficient sampling, οπότε υπολογίζουμε τις διακυμάνσεις των τριών πρώτων eigenvectors που φαίνεται ότι θα αντιπροσωπεύουν την πλειοψηφία των βασικών κινήσεων του μορίου. Στην **Εικόνα 4-51** έχουμε τη γραφική απεικόνιση των 3 πρώτων eigenvectors. Εάν ο χρόνος της προσομοίωσης είναι αρκετός για sufficient sampling ορισμένων κινήσεων, το ιστόγραμμα θα αντιστοιχεί σε μονοδιάστατη (1D) κανονική κατανομή γύρω από το μηδέν, το επίπεδο που σχηματίζουν οι 3 πρώτοι eigenvectors, συνδυαζόμενοι ανά δυο, θα αντιστοιχεί σε δισδιάστατη (2D) κανονική κατανομή γύρω από τη θέση 0,0 και το επίπεδο που σχηματίζουν οι 3 πρώτοι eigenvectors θα αντιστοιχεί σε τρισδιάστατη (3D) κανονική κατανομή γύρω από τη θέση 0,0,0. Βλέπουμε πως έχει επιτευχθεί σχεδόν sufficient sampling για τον πρώτο eigenvector, όμως βλέπουμε την ύπαρξη ενός μεγάλου cluster και ίσως κάποιου μικρότερου. Κάτι αντίστοιχο συμβαίνει σε μικρότερο βαθμό και με τον δεύτερο eigenvector, ενώ στον τρίτο ο χρόνος της προσομοίωσης φαίνεται πως είναι αρκετός για να τον παρακολουθήσουμε. Στο κάτω μέρος της εικόνας, όπου έχουμε την τρισδιάστατη απεικόνιση της προβολής των 3 eigenvectors φαίνεται πως δημιουργούνται δυο clusters. Όμως η απεικόνιση αυτή δε μπορεί να μας δείξει τη συνεισφορά σε κάθε cluster. Χρειαζόμαστε λοιπόν μια απεικόνιση στην οποία θα επισημαίνεται η πληροφορία σχετικά με την πυκνότητα των σημείων. Κάτι ανάλογο δηλαδή με την **Εικόνα 4-52**, όπου η χρωματική διαβάθμιση από μαύρο (0) σε κίτρινο (30) ακολουθεί την αυξητική πορεία της πυκνότητας των σημείων. Συνεπώς φαίνονται δύο βασικά clusters με σημαντικό αριθμό σημείων.

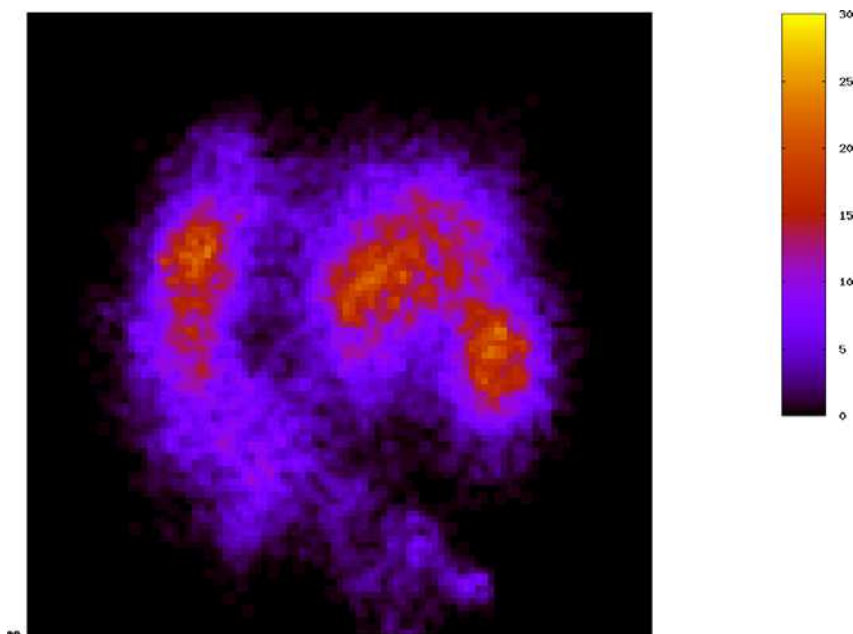


Εικόνα 4-51

Πάνω: Προβολή διακυμάνσεων των 3 πρώτων *eigenvectors* σε μια διάσταση

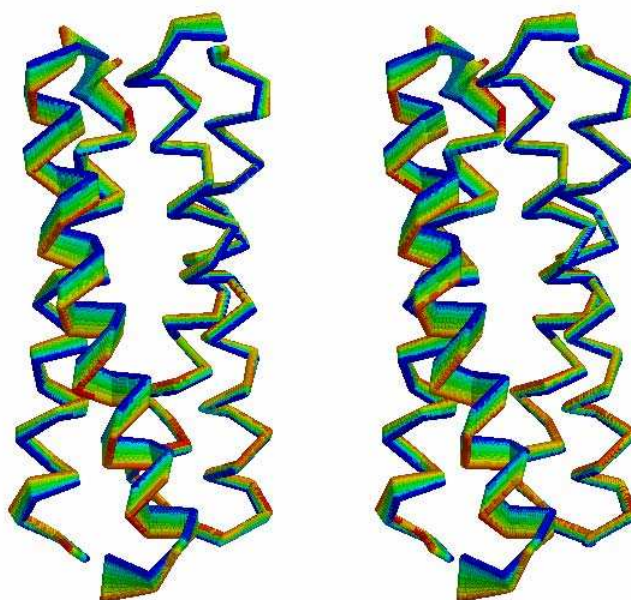
Στη μέση: Προβολή διακυμάνσεων των 3 πρώτων *eigenvectors*, συνδυασμένων ανά 2 (1-2, 1-3, 2-3 από αριστερά προς δεξιά) στις δυο διαστάσεις

Κάτω: Σχηματικό στερεοδιάγραμμα της προβολής των 3 πρώτων *eigenvectors* σε τρεις διαστάσεις.



Εικόνα 4-52: Προβολή διακυμάνσεων ανά ζεύγη eigenvectors με βάση την πυκνότητα των σημείων.

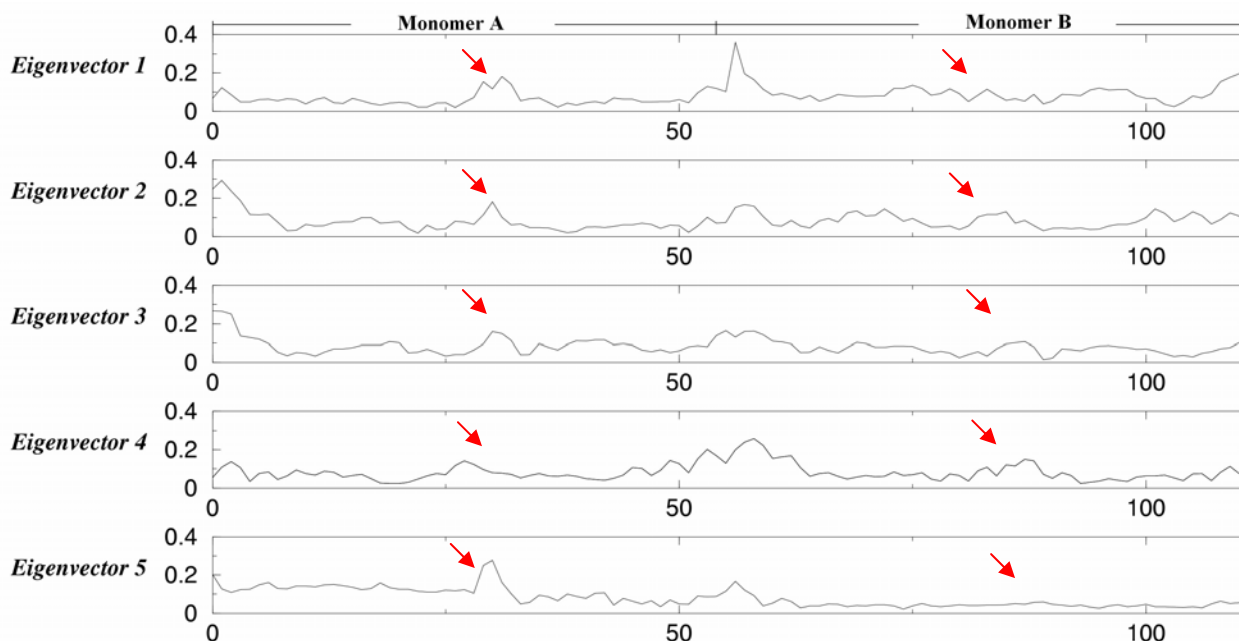
Στην **Εικόνα 4-53** μπορούμε να δούμε ένα σχηματικό στερεοδιάγραμμα που αντιστοιχεί στην κίνηση του μορίου βάσει του πρώτου ζεύγους eigenvalue-eigenvector για ολόκληρη τη διάρκεια της προσομοίωσης. Η χρωματική διαβάθμιση από μπλε σε κόκκινο υποδεικνύει το εύρος της κίνησης.



Εικόνα 4-53: Σχηματικό στερεοδιάγραμμα της κίνησης του μορίου, βάσει του πρώτου ζεύγους eigenvalue-eigenvector σε όλη τη διάρκεια του trajectory.

Για να ολοκληρώσουμε τη μελέτη των eigenvectors μένει να μελετήσουμε την κατανομή των διακυμάνσεων ανά κατάλοιπο. Έτσι μπορούμε να έχουμε μια

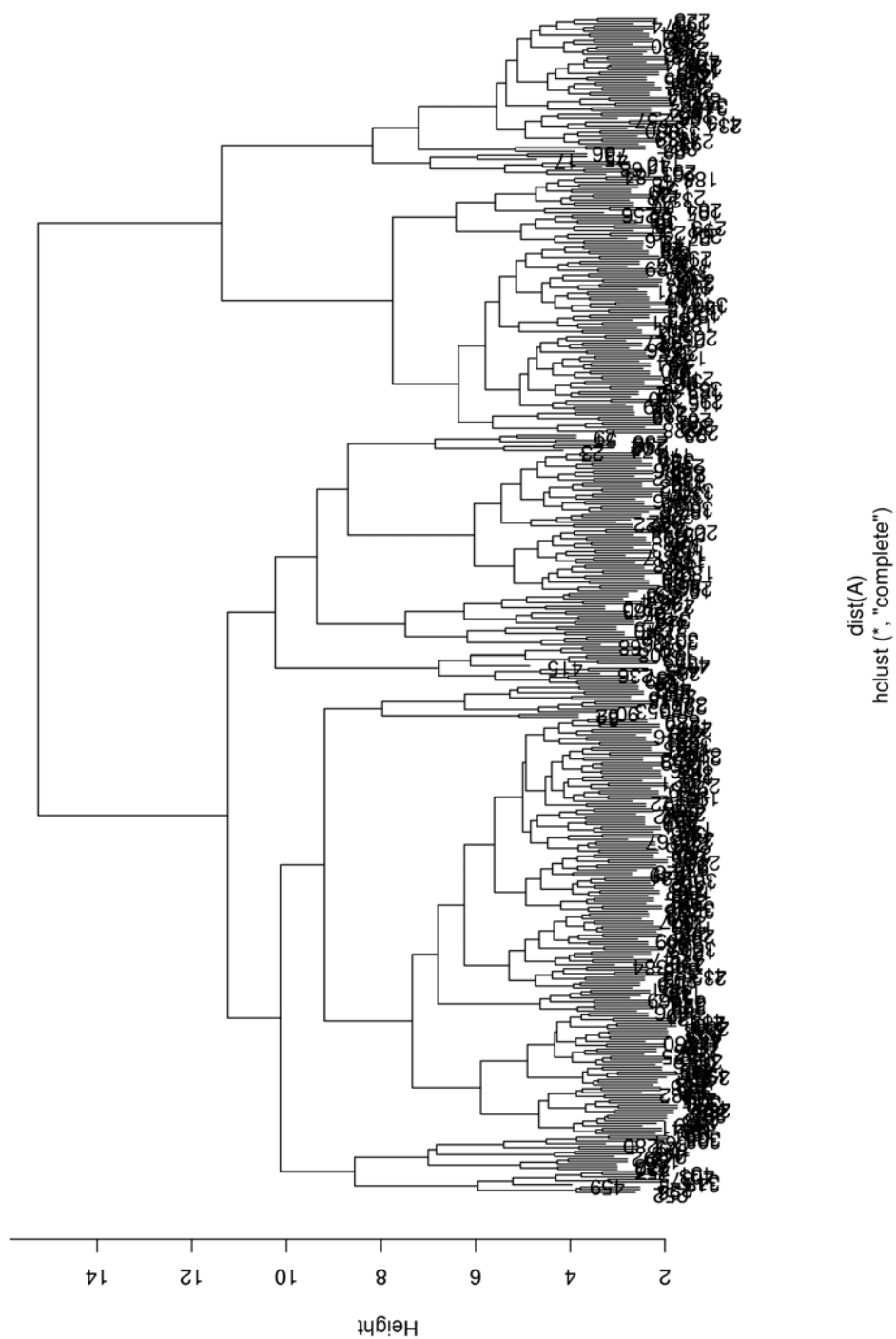
ολοκληρωμένη άποψη για την ποσότητα των σχετιζόμενων διακυμάνσεων του χαρακτηριστικού ανύσματος μεταξύ των καταλοίπων. Στην **Εικόνα 4-54** φαίνεται η διακύμανση των πέντε πρώτων χαρακτηριστικών ανυσμάτων χωριστά για κάθε κατάλοιπο.



Εικόνα 4-54: Αναπαράσταση των διακυμάνσεων των πέντε πρώτων χαρακτηριστικών ανυσμάτων ανά κατάλοιπο. Με κόκκινα βέλη σημαίνονται οι στροφές κάθε μονομερούς.

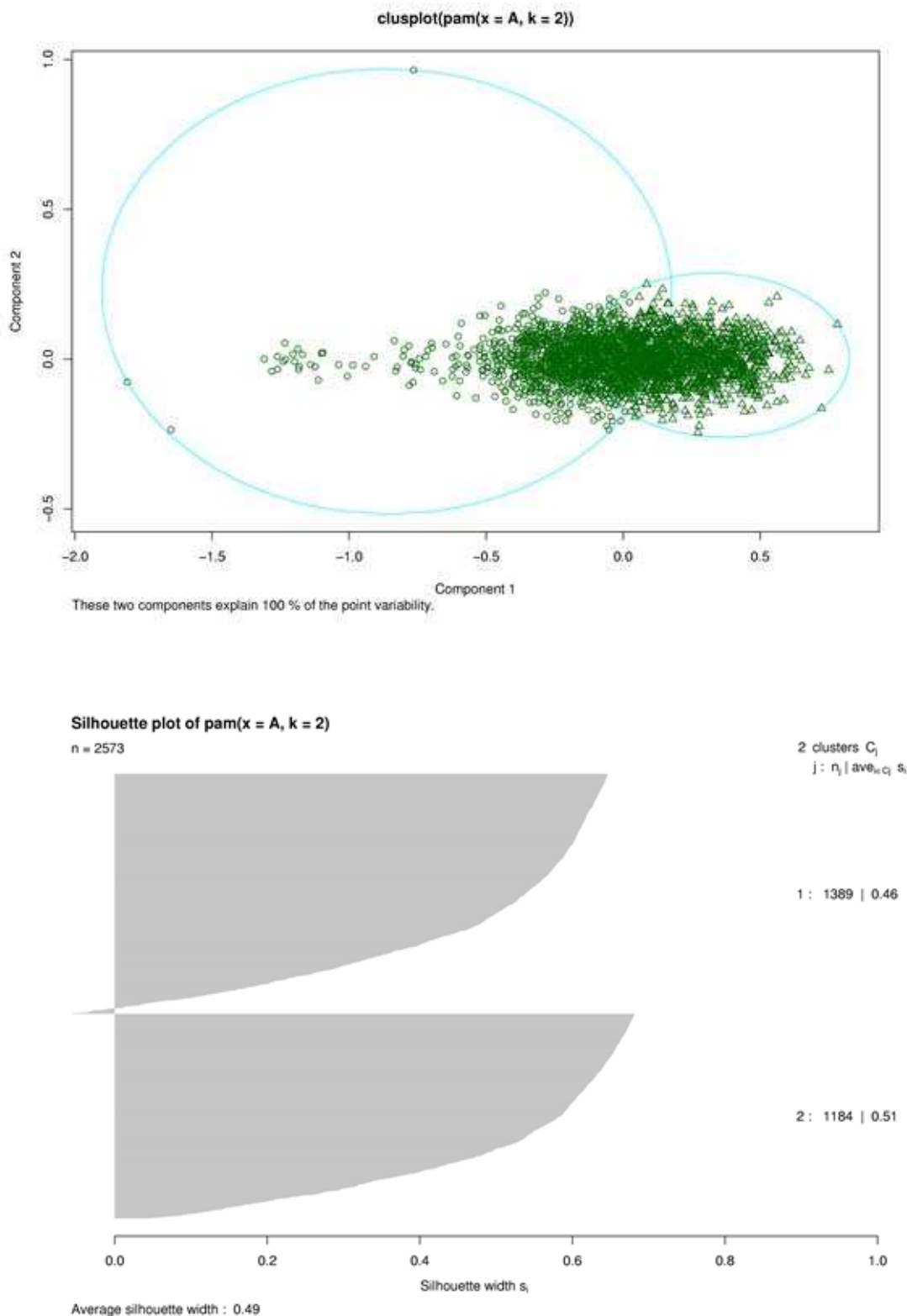
4.2.9 Ομαδοποίηση των δομών της προσομοίωσης (cluster analysis)

Όπως είδαμε και στην αρχή του κεφαλαίου, για να ομαδοποιήσουμε τις δομές του trajectory μπορούμε να φτιάξουμε ένα δενδρόγραμμα, όπως αυτό της **Εικόνας 4-55**, όπου θα λαμβάνεται υπ' όψιν το RMSD ανάμεσα στις διάφορες δομές. Έτσι μας δίνεται η δυνατότητα να προσδιορίσουμε τον αριθμό των κύριων clusters. Για να γίνει κάτι τέτοιο μπορούμε να χρησιμοποιήσουμε το κατάλληλο πρόγραμμα, που θα μας βοηθήσει να πάρουμε τις απαραίτητες πληροφορίες για τα βασικά clusters.



Εικόνα 4-55: Δενδρόγραμμα των δομών από τις οποίες περνά κατά τη διάρκεια του trajectory η Ala_2Ile_2-6 με τοπολογία anti, βάσει του RMSD.

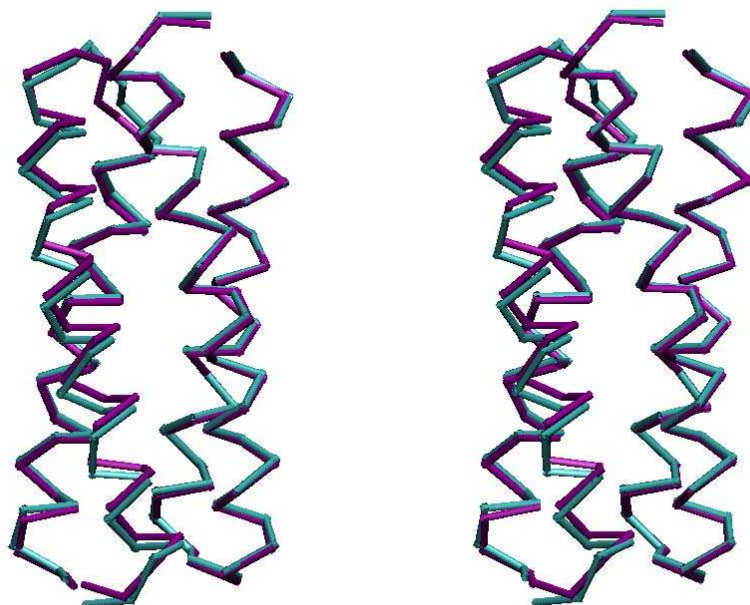
Στην **Εικόνα 4-56** φαίνονται δυο σημαντικές ομάδες δομών οι οποίες επικαλύπτονται, όπως είδαμε και προηγουμένως (Ενότητα 4.1.8). Χρησιμοποιήσαμε τον αλγόριθμο pam για να κάνουμε τους υπολογισμούς.



Εικόνα 4-56: Πάνω: Προσδιορισμός των clusters **Κάτω:** Κατανομή των δομών στα δυο clusters.

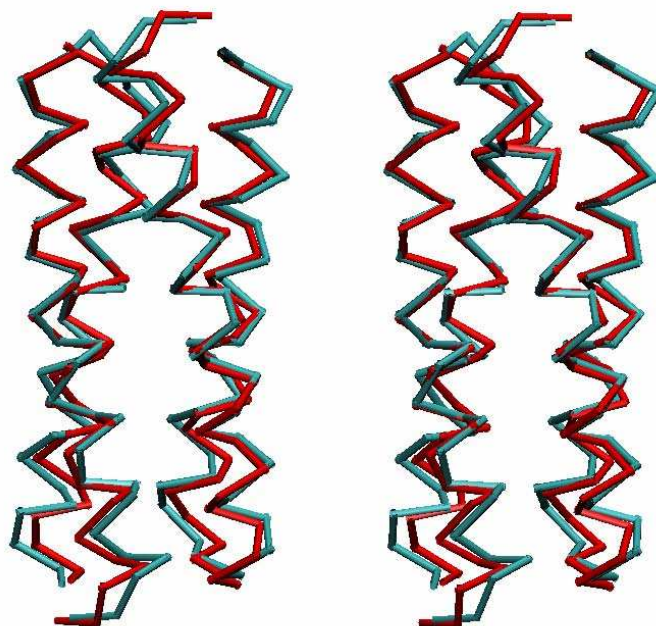
4.2.10 Σύγκριση μεταξύ μέσης δομής των 2 clusters και της αρχικής δομής

Εφόσον προσδιορίστηκε ο αριθμός των clusters ας δούμε αν υπάρχουν διαφορές στη μέση δομή καθενός εκ των δυο αυτών clusters (**Εικόνα 4-57**), εκ των οποίων το πρώτο αντιστοιχεί στα frames 1319-15833 (56% του συνόλου των frames) και το δεύτερο στα frames 16163-23750 (29% του συνόλου των frames), καθώς και ανάμεσα στη μέση δομή καθενός cluster και στην αρχική δομή της Ala_2Ile_2-6 με τοπολογία anti (Rop-like) (**Εικόνες 4-58 και 4-59**).

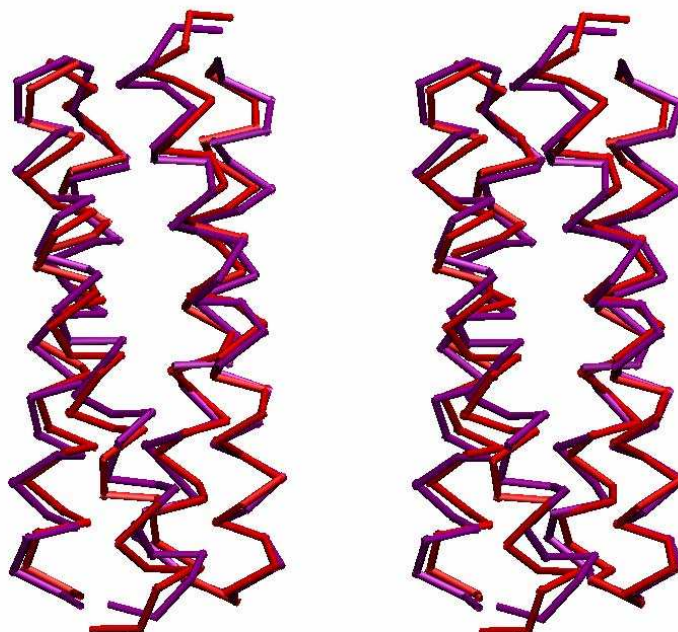


***Εικόνα 4-57:** Σχηματικό στερεοδιάγραμμα της μέσης δομής του πρώτου (με κυανό χρώμα) και του δεύτερου (με μοβ χρώμα) cluster σε υπέρθεση.*

Στην **Εικόνα 4-57** βλέπουμε πως οι δομές σχεδόν συμπίπτουν, δεν υπάρχει κάποια σημαντική διαφορά ανάμεσά τους. Τέλος παρατηρούμε μόνο μια μικρή απόκλιση μεταξύ των πέντε πρώτων N-τελικών καταλοίπων της Α έλικας, τόσο μεταξύ της αρχικής δομής και της μέσης δομής του πρώτου cluster (**Εικόνα 4-58**), όσο και μεταξύ της αρχικής δομής και της μέσης δομής του δεύτερου cluster (**Εικόνα 4-59**).



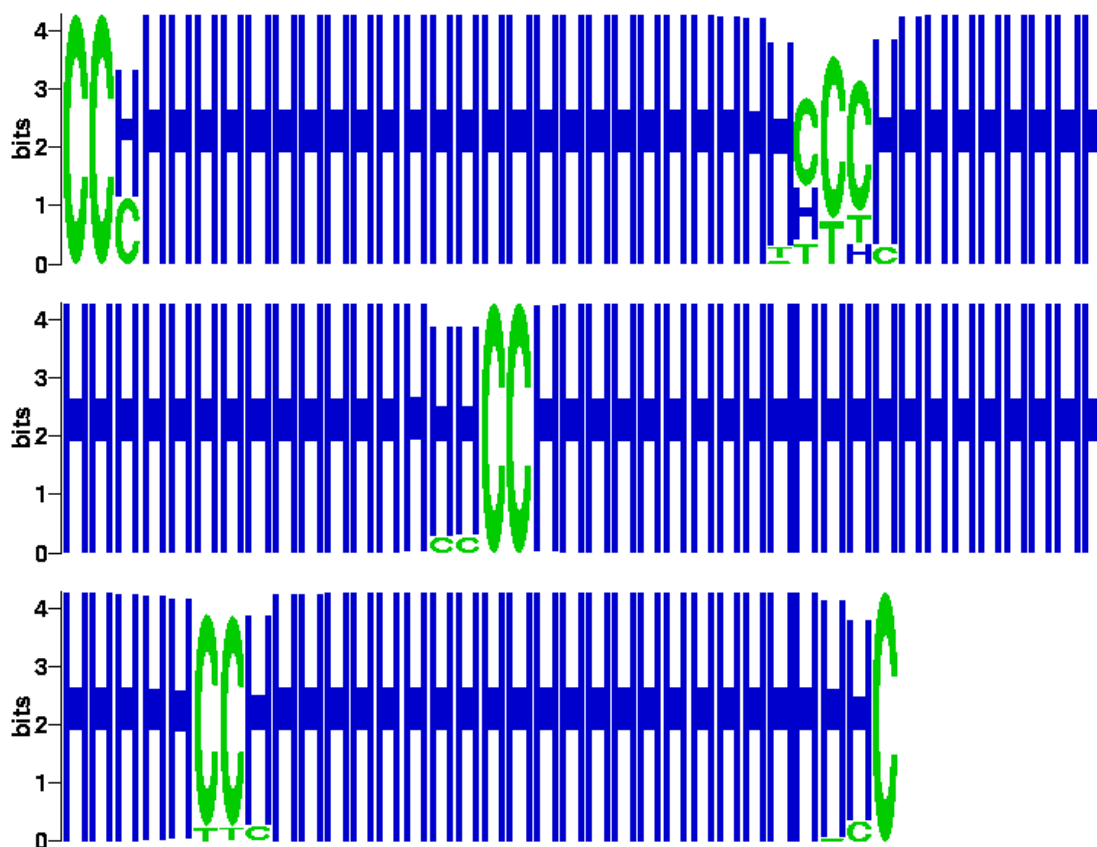
Εικόνα 4-58: Σχηματικό στερεοδιάγραμμα της μέσης δομής του πρώτου (με κυανό χρώμα) cluster και της αρχικής δομής του μορίου (με κόκκινο χρώμα) σε υπέρθεση.



Εικόνα 4-59: Σχηματικό στερεοδιάγραμμα της μέσης δομής του δεύτερου (με μοβ χρώμα) cluster και της αρχικής δομής του μορίου (με κόκκινο χρώμα) σε υπέρθεση.

4.2.11 Ανάλυση δευτεροταγούς δομής

Στο σημείο αυτό θα μελετήσουμε κατά πόσο διατηρούνται τα στοιχεία δευτεροταγούς δομής, δηλαδή οι α έλικες, κατά τη διάρκεια του trajectory. Για το σκοπό αυτό χρησιμοποιούμε ένα logo (Frishman *et al.*, 1995), που φαίνεται στην **Εικόνα 4-60**.



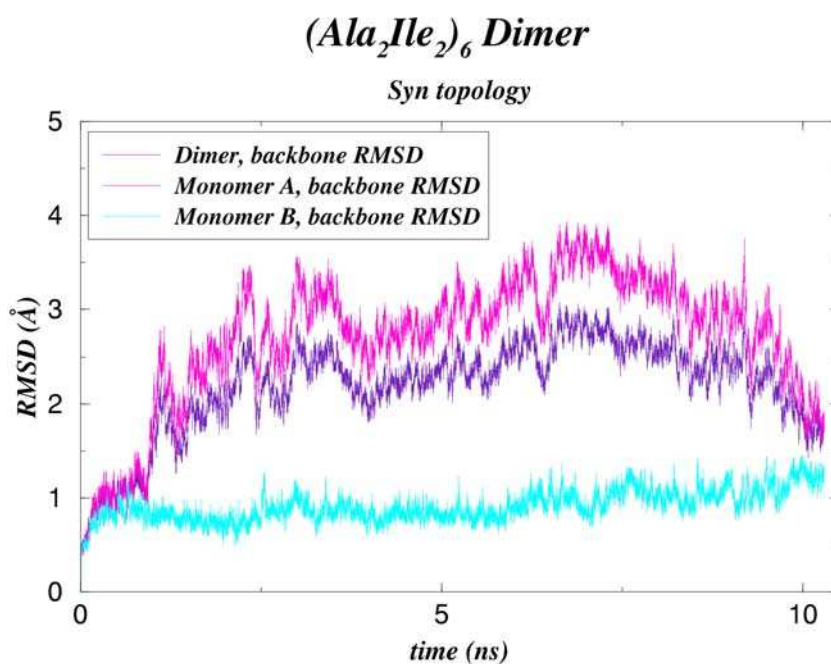
Εικόνα 4-60: Ανάλυση της δευτεροταγούς δομής κατά μήκος της αλληλουχίας

Με Η παριστάνονται οι α έλικες (Helix) και με C οι θηλιές (Coils). Σε κάθε σειρά βρίσκονται 40 θέσεις αλληλουχίας. Οι α έλικες διατηρούνται καθ' όλη τη διάρκεια του trajectory και είναι αρκετά σταθερές. Οι περιοχές που δε σχηματίζουν σταθερή δομή, αλλά θηλιά, αντιστοιχούν στα κατάλοιπα 1-3 του N-τελικού άκρου, 30-32 της στροφής και 55-56 του C-τελικού άκρου του μονομερούς A, καθώς και στα κατάλοιπα 1-2 του N-τελικού άκρου, 30-31 της στροφής και 55-56 του C-τελικού άκρου του μονομερούς B.

4.3 Σύγκριση αποτελεσμάτων ανάλυσης μεταξύ Ala_2Ile_2-6 με *syn* τοπολογία και Ala_2Ile_2-6 με *anti* (Ror-like) τοπολογία

Στις ενότητες 4.1 και 4.2 αναλύσαμε βήμα προς βήμα τα αποτελέσματα των προσομοιώσεων των δυο διμερών, Ala_2Ile_2-6 με τοπολογία *syn* και Ala_2Ile_2-6 με τοπολογία *anti*. Απώτερος σκοπός της μελέτης αυτής ήταν να προσπαθήσουμε να κατανοήσουμε, στο βαθμό που είναι εφικτό, το λόγο για τον οποίο αναδιπλώνεται κατά αυτό τον τρόπο η Ala_2Ile_2-6 . Η όλη προσπάθεια βασίστηκε στην αντιπαραβολή και σύγκριση των αποτελεσμάτων εκείνων που μπορούν να μας δώσουν μια πιο πλήρη εικόνα όσο αφορά την αναδίπλωση των δυο διμερών και τις διαφορές που πιθανώς προκύπτουν, όσο αφορά τη σταθερότητα των δυο συστημάτων.

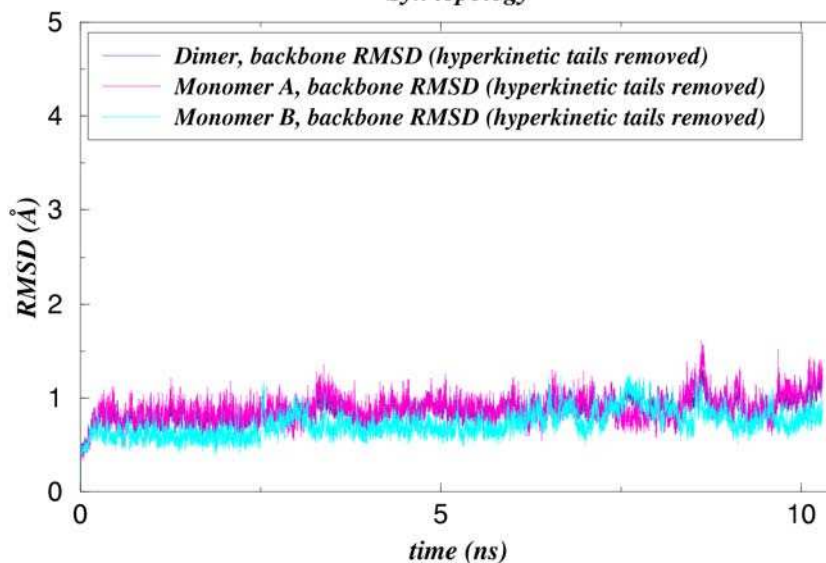
Αρχικά ας δούμε τι συμπεράσματα μπορούμε να εξάγουμε από τη μελέτη της τυπικής απόκλισης (RMSD) από την αρχική δομή, στη μέτρηση της οποίας λαμβάνονται υπ' όψιν τα άτομα του σκελετού (backbone) καθενός μορίου (**Εικόνες 4-61, 4-62, 4-63**).



Εικόνα 4-61: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα του σκελετού, τόσο του διμερούς, όσο και των δυο μονομερών. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Παρόλο που το μονομερές A επηρεάζει σημαντικά την τυπική απόκλιση του διμερούς, βλέπουμε πως το μονομερές B αντισταθμίζει κατά ένα μέρος τη διακύμανση.

$(Ala_2Ile_2)_6$ Dimer

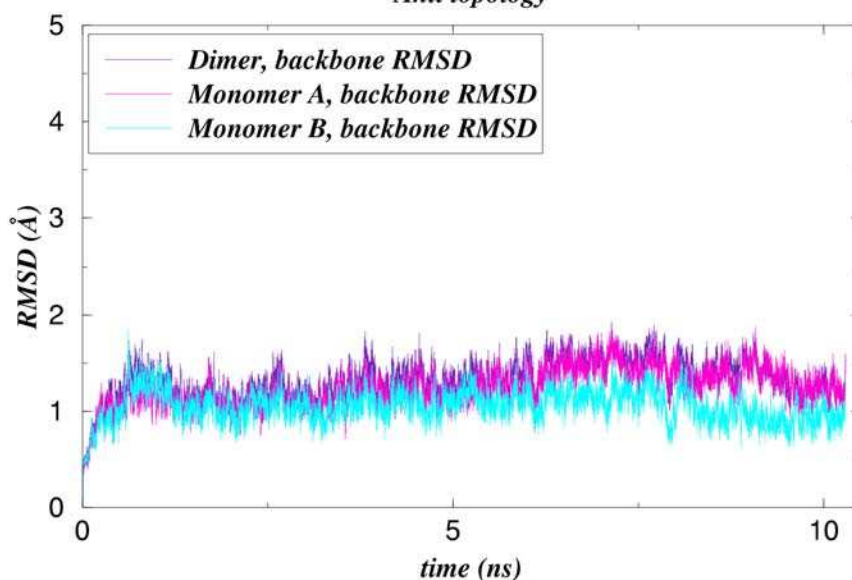
Syn topology



Εικόνα 4-62: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα του σκελετού, τόσο του διμερούς, όσο και των δυο μονομερών, με εξαίρεση τα υπερκινητικά άκρα. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Παρόλο που το μονομερές A επηρεάζει σημαντικά την τυπική απόκλιση του διμερούς, βλέπουμε πως το μονομερές B αντισταθμίζει κατά ένα μέρος τη διακύμανση.

$(Ala_2Ile_2)_6$ Dimer

Anti topology



Εικόνα 4-63: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα του σκελετού, τόσο του διμερούς, όσο και των δυο μονομερών. Η μοβ καμπύλη αντιστοιχεί στο διμερές, η φούξια στο μονομερές A και η κυανή στο μονομερές B. Παρόλο που το μονομερές A επηρεάζει σημαντικά την τυπική απόκλιση του διμερούς, βλέπουμε πως το μονομερές B αντισταθμίζει κατά ένα μέρος τη διακύμανση.

Όπως αναφέραμε και στην ενότητα 4.1 και όπως μπορούμε να παρατηρήσουμε και στην **Εικόνα 4-61**, τα άκρα των δυο μονομερών, A και B, που συγκροτούν το διμερές $\text{Ala}_2\text{Ile}_2-6$ με τοπολογία syn, κινούνται ακανόνιστα με αποτέλεσμα να επηρεάζουν σε μεγάλο βαθμό τη σταθερότητα του συστήματος. Φανερό είναι το γεγονός ότι το μονομερές A είναι εκείνο που επηρεάζει πιο πολύ τη διακύμανση του διμερούς. Οι μετρήσεις της διακύμανσης φαίνονται στον **Πίνακα 4-1**.

Σύστημα υπό μελέτη/ Τιμή RMSD (Å)	$\text{Ala}_2\text{Ile}_2-6$ syn			$\text{Ala}_2\text{Ile}_2-6$ syn *			$\text{Ala}_2\text{Ile}_2-6$ anti		
	A	B	Διμερές	A	B	Διμερές	A	B	Διμερές
Αρχική	0,5	0,5	0,5	0,4	0,4	0,4	0,4	0,4	0,4
Τελική	1,8	1,3	1,7	1,3	0,8	1,2	1,3	1,0	1,3
Μέγιστη	3,9	1,5	3,0	1,6	1,3	1,5	2,0	1,8	1,7
Ελάχιστη**	1,0	0,5	1,0	0,6	0,4	0,7	0,6	0,6	0,8

* Στις μετρήσεις δεν περιλαμβάνονται τα άτομα του σκελετού των υπερκινητικών άκρων.

** Η ελάχιστη τιμή από τη φάση heating και μετά.

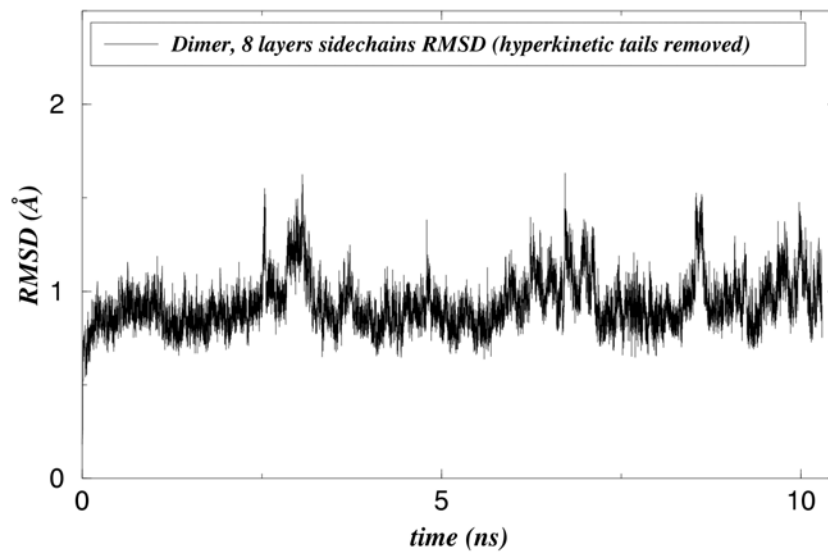
Πίνακας 4-1: Μετρήσεις τυπικής απόκλισης από την αρχική δομή για τα συστήματα που μελετάμε. Οι τιμές είναι κατά προσέγγιση δεκάτου.

Βλέπουμε λοιπόν πως η διακύμανση του διμερούς $\text{Ala}_2\text{Ile}_2-6$ με τοπολογία anti (0,8-1,7 Å) είναι ελαφρώς μεγαλύτερη από εκείνη του διμερούς $\text{Ala}_2\text{Ile}_2-6$ με τοπολογία syn (0,7-1,5 Å) χωρίς τα υπερκινητικά άκρα, ενώ η διακύμανση του διμερούς $\text{Ala}_2\text{Ile}_2-6$ με τοπολογία syn ξεπερνά κατά πολύ τις άλλες δυο (1,0-3,0 Å). Όμως όπως είπαμε και προηγουμένως τα αποτελέσματα για το σύστημα του διμερούς $\text{Ala}_2\text{Ile}_2-6$ με τοπολογία syn δε θα περιλαμβάνουν τα υπερκινητικά άκρα, συνεπώς η σύγκριση των δυο συστημάτων θα γίνει μεταξύ των $\text{Ala}_2\text{Ile}_2-6$ με τοπολογία anti και $\text{Ala}_2\text{Ile}_2-6$ με τοπολογία syn χωρίς τα υπερκινητικά άκρα. Γενικά όμως και οι δυο διακυμάνσεις που μας αφορούν είναι αρκετά σταθεροποιημένες, η πρώτη (τοπολογία syn) γύρω από τα 0,9Å και η δεύτερη (τοπολογία anti) γύρω από τα 1,2Å.

Ας δούμε στο σημείο αυτό τι γίνεται με τους πυρήνες των δυο συστημάτων, όσο αφορά την τυπική απόκλιση από την αρχική δομή (**Εικόνες 4-64, 4-65**). Στις δυο εικόνες βλέπουμε τη μεταβολή της τυπικής απόκλισης κατά τη διάρκεια της προσομοίωσης για τα άτομα των πλευρικών ομάδων των καταλοίπων που συμμετέχουν στο σχηματισμό των 8 στρωμάτων (layers) που αποτελούν τον υδρόφοβο πυρήνα καθενός από τα δυο συστήματα.

(Ala₂Ile₂)₆ Dimer

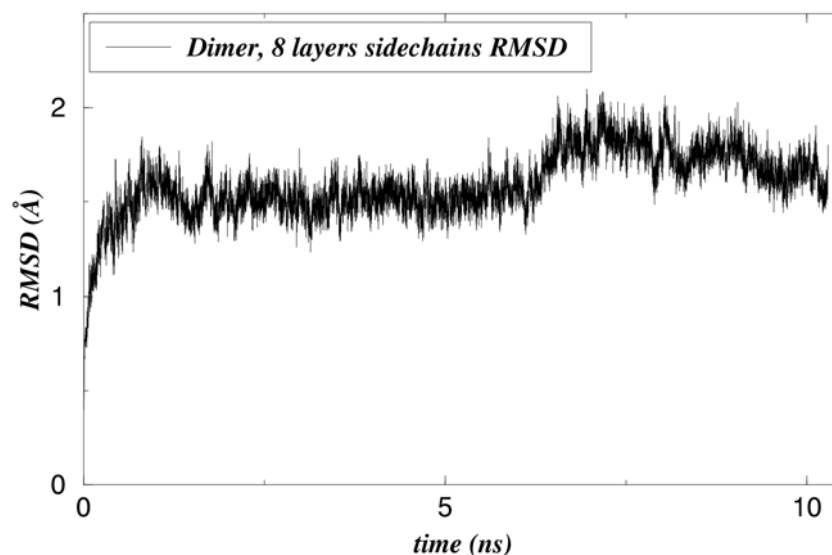
Syn topology



Εικόνα 4-64: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα των πλευρικών αλυσίδων των καταλοίπων του διμερούς που συμμετέχουν στη δημιουργία του υδρόφοβου πυρήνα. Εξαιρούνται τα άτομα των υπερκινητικών άκρων.

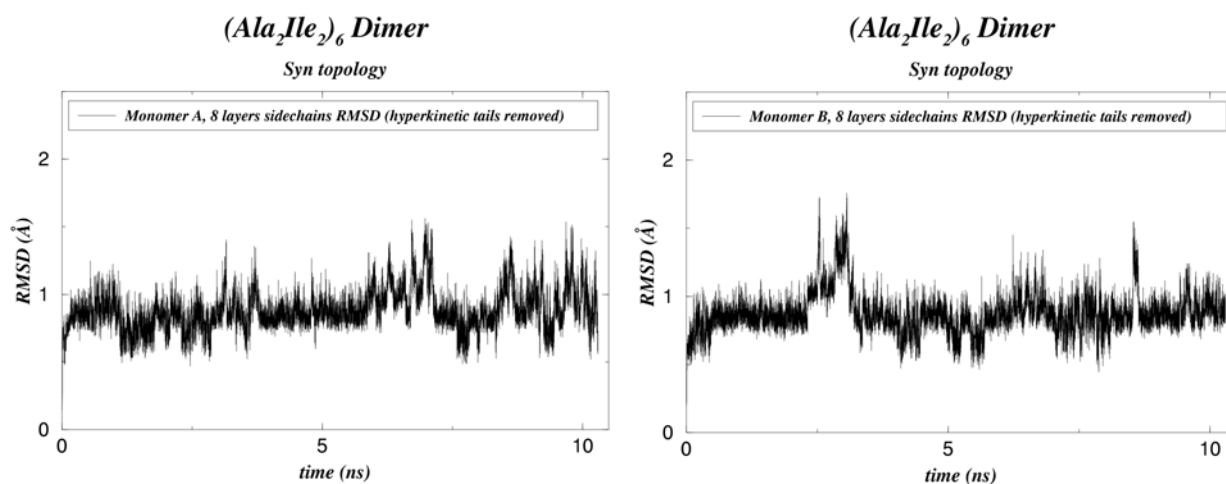
(Ala₂Ile₂)₆ Dimer

Anti topology

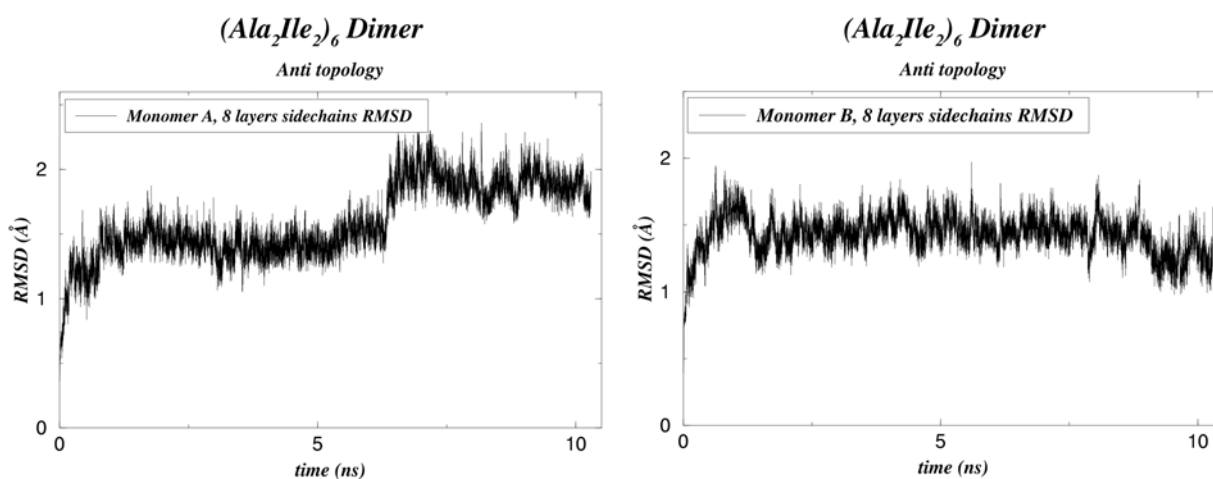


Εικόνα 4-65: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα των πλευρικών αλυσίδων των καταλοίπων του διμερούς που συμμετέχουν στη δημιουργία του υδρόφοβου πυρήνα.

Για το διμερές Ala_2Ile_2 -6 syn βλέπουμε πως οι τιμές της τυπικής απόκλισης είναι μεταξύ $0,7\text{\AA}$ και $1,6\text{\AA}$. Παρατηρούμε πως δεν σταθεροποιείται ιδιαίτερα, έχουμε αρκετά μέγιστα κυρίως αλλά και ελάχιστα, παρόλα αυτά όμως έχουμε μια γενικότερη διακύμανση γύρω από τα $1,0\text{\AA}$. Για το διμερές Ala_2Ile_2 -6 anti οι τιμές είναι μεταξύ $1,2\text{\AA}$ και $2,1\text{\AA}$. Η σταθεροποίηση είναι ευκολότερη στο σύστημα αυτό απ' ότι στο προηγούμενο, όμως η τιμή του RMSD είναι γενικότερα υψηλότερη, γύρω από τα $1,7\text{\AA}$ κατά τη διάρκεια του trajectory.



Εικόνα 4-66: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα των πλευρικών αλυσίδων των καταλοίπων των μονομερών A και B που συμμετέχουν στη δημιουργία του υδρόφοβου πυρήνα. Εξαιρούνται τα άτομα των υπερκινητικών άκρων.



Εικόνα 4-67: Καμπύλη τυπικής απόκλισης RMSD (Å) συναρτήσει του χρόνου (ns). Στη μέτρηση συμπεριλαμβάνονται μόνο τα άτομα των πλευρικών αλυσίδων των καταλοίπων των μονομερών A και B που συμμετέχουν στη δημιουργία του υδρόφοβου πυρήνα.

Στις **Εικόνες 4-66** και **4-67** μπορούμε να διακρίνουμε τις τυπικές αποκλίσεις για τα μονομερή A και B χωριστά. Οι τιμές δεν αποκλίνουν ιδιαίτερα από εκείνες των διμερών και των μονομερών μεταξύ τους. Αξιοσημείωτο είναι το γεγονός ότι στην τυπική απόκλιση του μονομερούς A της Ala_2Ile_2-6 anti είναι πιο φανερή η ύπαρξη δυο clusters.

Στη συνέχεια έχουμε τη σύγκριση των αποτελεσμάτων της τυπικής απόκλισης από τη μέση δομή, η οποία αποτελεί ένα μέτρο σταθερότητας του συστήματος, καθώς υποδεικνύει κατά πόσο διαφέρει η δομή του συστήματος από τη μέση δομή. Όπως φαίνεται στις **Εικόνες 4-21** και **4-44**, η διακύμανση της τυπικής απόκλισης για το διμερές της Ala_2Ile_2-6 syn, είναι γύρω από τα $0,5\text{\AA}$ σε όλη τη διάρκεια του trajectory, ενώ για το διμερές της Ala_2Ile_2-6 anti είναι γύρω από τα $0,7\text{\AA}$ καθ' όλη τη διάρκεια του trajectory εκτός από την εκκίνηση της προσομοίωσης που είναι στα 1\AA , σε αντίθεση με το άλλο διμερές που είναι και εκεί στα $0,5\text{\AA}$. Οι αντίστοιχες τιμές είναι ελάχιστα πιο υψηλές για τα μονομερή A και ελάχιστα πιο χαμηλές για τα μονομερή B.

Εάν μελετήσουμε το RMSF και για τα δύο συστήματα (**Εικόνες 4-22** και **4-45**), παρατηρούμε πως έχουμε τις μέγιστες τιμές στις ίδιες περιοχές, δηλαδή κοντά στα άκρα των μονομερών και κοντά στις στροφές, με τη διαφορά ότι στην Ala_2Ile_2-6 anti οι τιμές είναι οι τιμές του RMSF είναι χαμηλότερες στις περιοχές που αναφέραμε προηγουμένως (και αυτό λόγω των υπερκινητικών άκρων της Ala_2Ile_2-6 syn) και υψηλότερες σε όλες τις υπόλοιπες περιοχές.

Ας δούμε τι συμβαίνει με την R_G . Στις **Εικόνες 4-24** και **4-46** δε φαίνεται κάποια σημαντική διαφορά. Οι διακυμάνσεις είναι για τα δυο A μονομερή γύρω από τα $14,6\text{\AA}$ και για τα δυο B μονομερή γύρω από τα $14,5\text{\AA}$. Η μόνη σημαντική διαφορά φαίνεται στην αρχή του trajectory, όπου στο σύστημα Ala_2Ile_2-6 syn η ακτίνα περιστροφής είναι ομαλή, ενώ στο σύστημα Ala_2Ile_2-6 anti έχουμε μια απότομη ελάττωση από τα $14,75\text{\AA}$. Τα αποτελέσματα αυτά όμως δε μας βοηθούν να εξάγουμε κάποιο συμπέρασμα για το ποιο σύστημα είναι πιο σταθερό.

Στις **Εικόνες 4-32** και **4-54** βλέπουμε τη διακύμανση των πέντε πρώτων eigenvectors ανά κατάλοιπο και για τα δυο συστήματα. Στην τοπολογία anti οι τιμές είναι μικρότερες στα κατάλοιπα των στροφών και των άκρων, αλλά σχετικά μεγαλύτερες και με περισσότερες διακυμάνσεις στις υπόλοιπες περιοχές, σε σύγκριση με την τοπολογία syn.

Ένας επιπλέον τρόπος που θα μπορούσε ίσως να δώσει κάποιο αποτέλεσμα πιο ξεκάθαρο, είναι ο υπολογισμός της έκτασης της προσβάσιμης επιφάνειας του μορίου (ASA, Accessible Surface Area) τόσο στα διμερή, όσο και στα μονομερή. Η διαφορά της ASA του διμερούς από το άθροισμα της ASA των δυο μονομερών θα πρέπει να αντικατοπτρίζει την επιφάνεια που θάβεται όταν δημιουργείται το σύμπλοκο. Κατά συνέπεια όσο περισσότερη επιφάνεια θάβεται τόσο πιο σταθερό είναι το σύμπλοκο του διμερούς.

Για τον υπολογισμό αυτό χρησιμοποιήσαμε σαν βάση τη μέση δομή για κάθε σύστημα (**Πίνακας 4-2**). Χρησιμοποιήθηκε το πρόγραμμα areaimol (από το σύνολο των προγραμμάτων CCP4) (Kabsch *et al.*, 1994). Η ASA για το διμερές της Ala_2Ile_2-6 με τοπολογία syn είναι 6399\AA^2 , ενώ για το μονομερές A είναι 4517\AA^2 και για το μονομερές B είναι 4453\AA^2 . Η ASA για το διμερές της Ala_2Ile_2-6 με τοπολογία anti (Rop-like) είναι 6336\AA^2 , ενώ για το μονομερές A είναι 4529\AA^2 και για το μονομερές B είναι 4554\AA^2 . Συνεπώς η επιφάνεια που θάβεται στην Ala_2Ile_2-6 με τοπολογία syn αντιστοιχεί σε $[4517 + 4453] - 6399 = 2571\text{\AA}^2$, ενώ στην Ala_2Ile_2-6 με τοπολογία anti (Rop-like) είναι $[4529 + 4554] - 6336 = 2747\text{\AA}^2$. παρατηρούμε λοιπόν πως η επιφάνεια που θάβεται στα δυο συστήματα δεν έχει κάποια σημαντική διαφορά, με τη δεύτερη να είναι λίγο μεγαλύτερη. Η Ala_2Ile_2-6 φαίνεται πως είναι αρκετά σταθερή, ειδικά από τη στιγμή που υπάρχουν πιο πολλά υδρόφοβα κατάλοιπα στον πυρήνα απ' ότι στη φυσική πρωτεΐνη. Ας μην ξεχνάμε πως η επιφάνεια της φυσικής Rop που θάβεται στον πυρήνα αντιστοιχεί μόλις σε 1345\AA^2 .

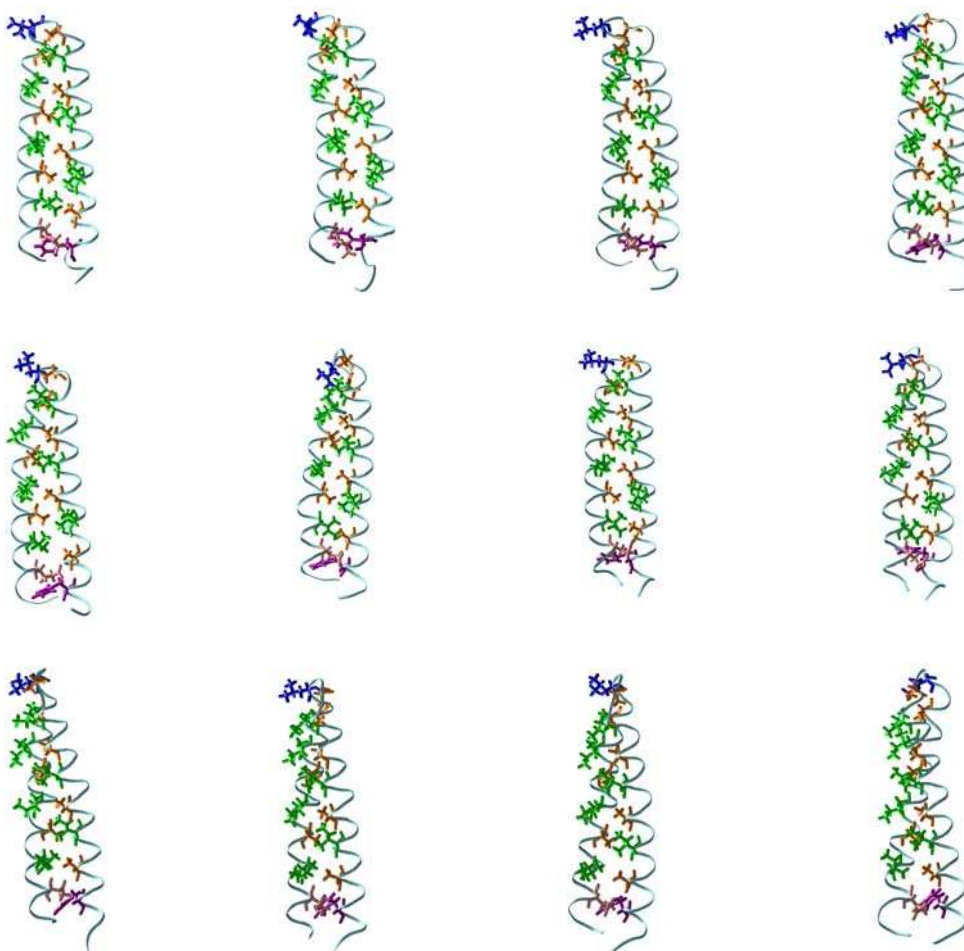
		Ala_2Ile_2-6 με τοπολογία syn	Ala_2Ile_2-6 με τοπολογία anti (Rop-like)
ASA (\AA^2)	A μονομερές	4517	4529
	B μονομερές	4435	4554
	Διμερές	6399	6336
BURIED SURFACE (\AA^2)		2571	2747

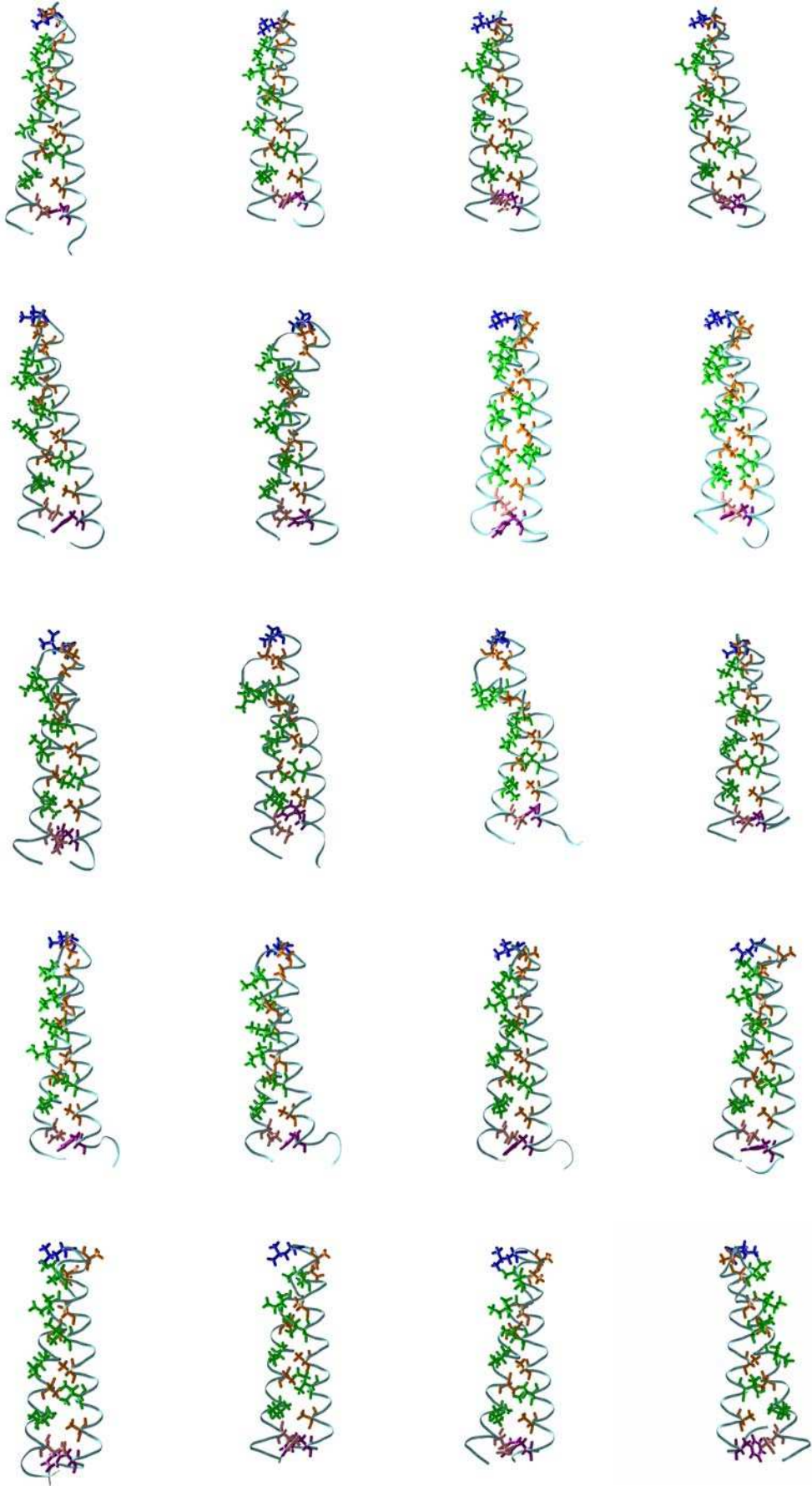
Πίνακας 4-2: Υπολογισμός της ASA για τα διμερή και μονομερή των δυο συστημάτων που μελετάμε, καθώς και υπολογισμός της επιφάνειας που θάβεται στον πυρήνα (buried surface).

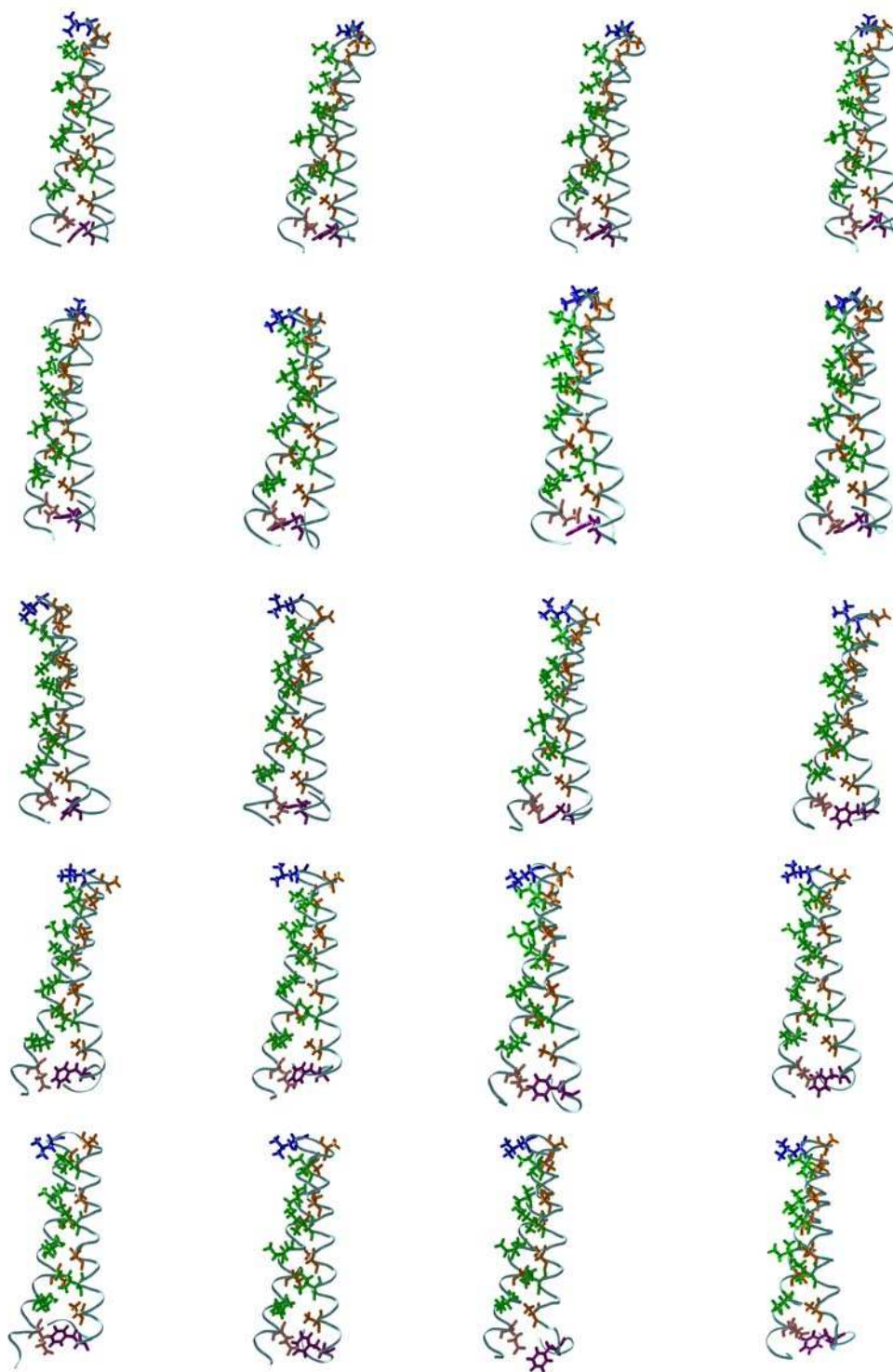
Συνοψίζοντας μπορούμε να πούμε ότι δεν υπάρχει κάποιο δεδομένο που να μας ωθεί στο συμπέρασμα ότι κάποια από τις δυο δομές είναι σταθερότερη. Δε μπορούμε δηλαδή να πούμε με βεβαιότητα ποια διαμόρφωση είναι καταλληλότερη για την παραλλαγή Ala_2Ile_2-6 . Τελικά ίσως κατά τη διάρκεια της αναδίπλωσης η Ala_2Ile_2-6 να παίρνει και τις δυο διαμορφώσεις, όπως διατύπωσε ο Levy και οι συνεργάτες του (Levy *et al.*, 2005). Ίσως δηλαδή να μην υιοθετείται μια μόνο διαμόρφωση, αλλά ένα σύνολο διαμορφώσεων. Κάτι τέτοιο όμως, για να ειπωθεί με σιγουριά χρειάζεται περαιτέρω και εκτενέστερη μελέτη.

4.4 Ανάλυση προσομοίωσης native μονομερούς

Επειδή ο διαθέσιμος χρόνος ήταν λίγος σε σχέση με τον όγκο των δεδομένων, ήμαστε σε θέση να αναλύσουμε διεξοδικά μόνο τα συστήματα των δυο διμερών. Όσο αφορά στις προσομοιώσεις των μονομερών, η ανάλυση αναγκαστικά θα περιοριστεί στην παρουσίαση επιλεγμένων στιγμιότυπων από τα trajectories που αποτελούν αντιπροσωπευτικό δείγμα της κίνησης και της δομής κάθε μονομερούς κατά τη διάρκεια της προσομοίωσης. Το native μονομερές δεν εμφανίζει σημαντική απόκλιση από την αρχική δομή (**Εικόνα 4-68**). Οι πλευρικές ομάδες των καταλοίπων που δημιουργούν τα οκτώ στρώματα του υδρόφοβου πυρήνα της πρωτεΐνης είναι αρκετά σταθερές, με εξαίρεση τα κατάλοιπα στη στροφή του μονομερούς και τα άκρα του μονομερούς, όπου σχηματίζεται το πρώτο στρώμα του πυρήνα. Το γεγονός αυτό δε μας κάνει εντύπωση, αφού γνωρίζουμε ήδη πως τα άκρα είναι ιδιαίτερα ευκίνητα. Μέχρι το τέλος της προσομοίωσης, οι πλευρικές ομάδες παραμένουν στραμμένες προς το άλλο μονομερές, όπως ήταν δηλαδή στο ομοδιμερές. Συνεπώς το native μονομερές δεν είναι αρκετό να αποτελέσει μια σταθερή δομή από μόνο του.



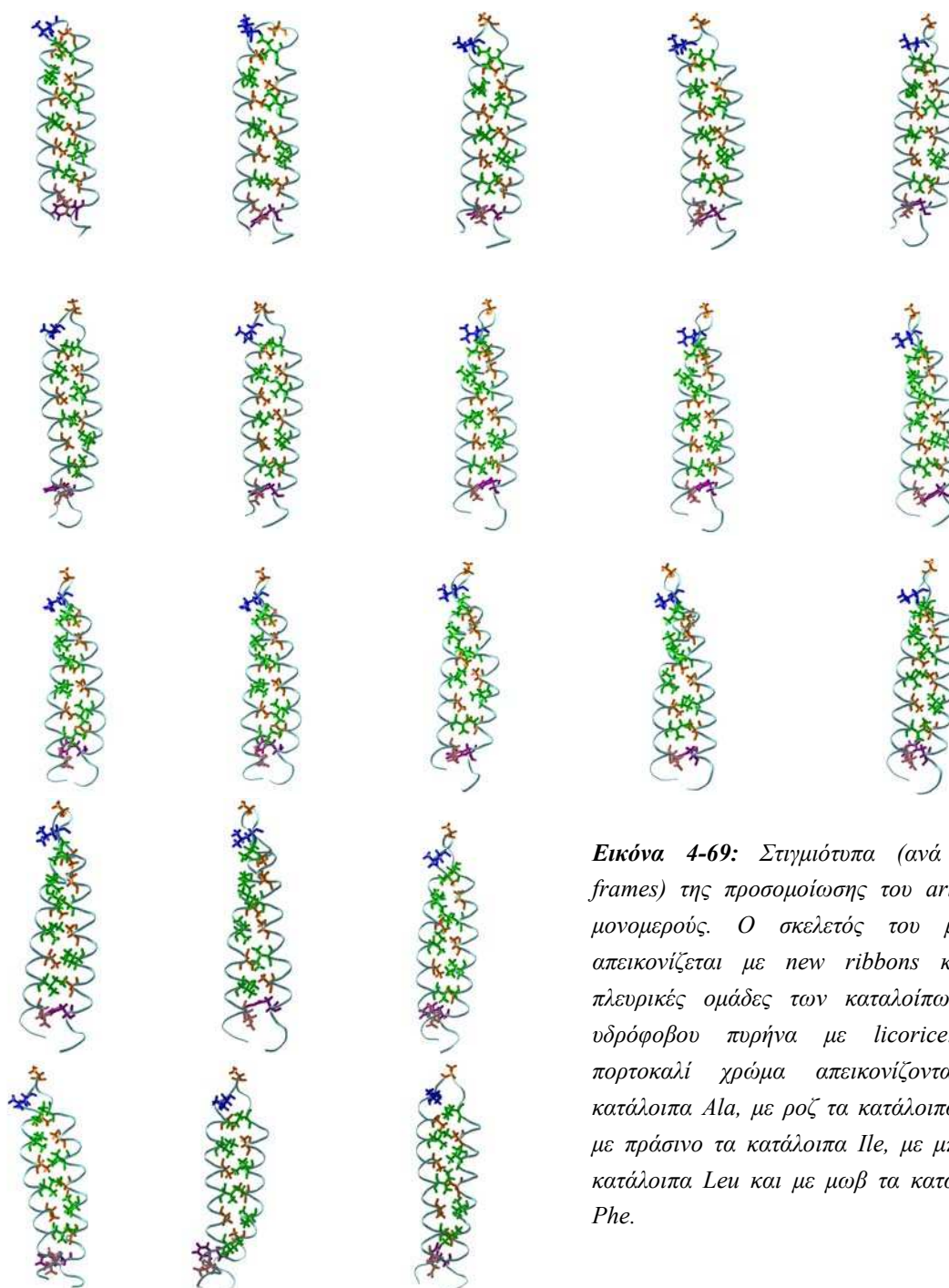




Εικόνα 4-68: Στιγμιότυπα (ανά 1000 frames) της προσομοίωσης του native μονομερούς. Ο σκελετός του μορίου απεικονίζεται με new ribbons και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με licorice. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα Ala, με ροζ τα κατάλοιπα Glu, με πράσινο τα κατάλοιπα Pro, με μπλε τα κατάλοιπα Leu και με μωβ τα κατάλοιπα Phe.

4.5 Ανάλυση προσομοίωσης artificial μονομερούς

Παρόμοια αποτελέσματα παρατηρούμε και στο artificial μονομερές. Και εδώ δεν παρουσιάζεται καμία σημαντική απόκλιση από την αρχική δομή, η κινητικότητα παρατηρείται στα κατάλοιπα της στροφής και του πρώτου επιπέδου του υδρόφοβου πυρήνα, ενώ οι πλευρικές ομάδες τείνουν να παραμείνουν στην αρχική κατεύθυνση (Εικόνα 4-69).



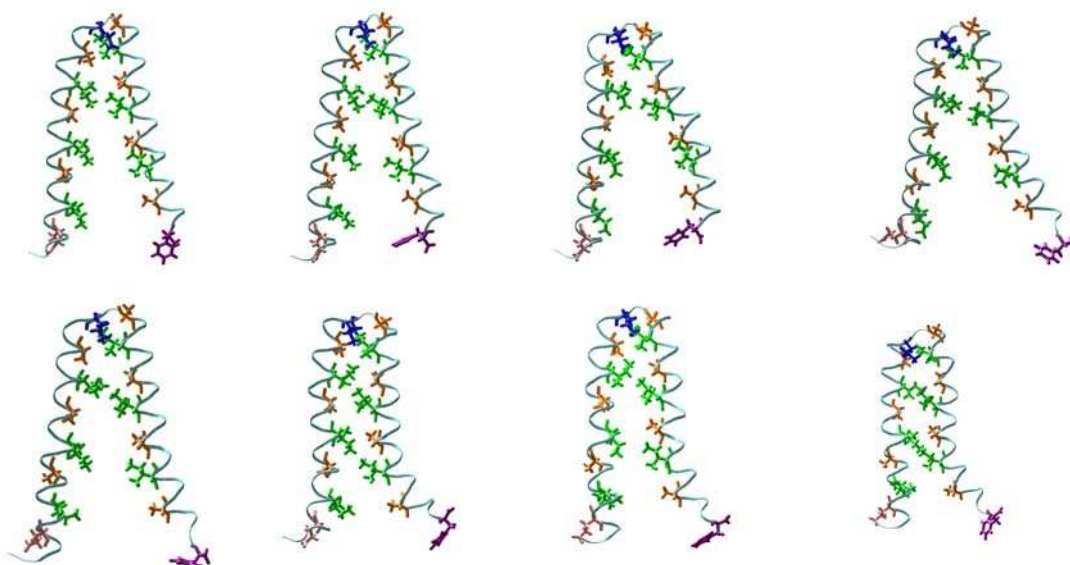
Εικόνα 4-69: Στιγμιότυπα (ανά 1000 frames) της προσομοίωσης του artificial μονομερούς. Ο σκελετός του μορίου απεικονίζεται με new ribbons και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με licorice. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα Ala, με ροζ τα κατάλοιπα Glu, με πράσινο τα κατάλοιπα Ile, με μπλε τα κατάλοιπα Leu και με μωβ τα κατάλοιπα Phe.

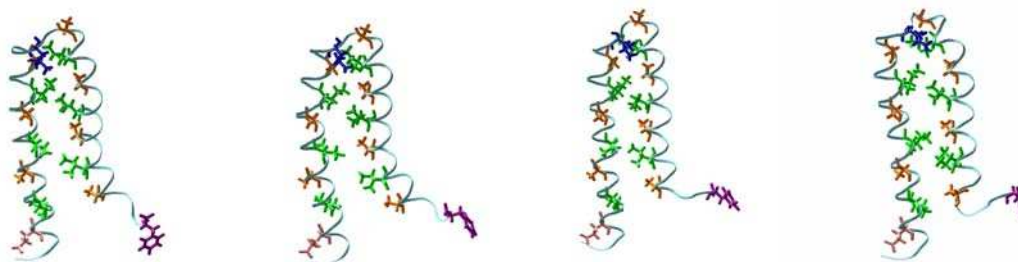
4.6 Ανάλυση προσομοίωσης opened (A31P) μονομερούς

Η τοπολογία A31P υιοθετεί μια ανοιχτή διαμόρφωση επιτρέποντας στις πλευρικές ομάδες να κινούνται ελεύθερα. Οι έξι από τις επτά προσομοιώσεις που πραγματοποιήθηκαν, είχαν μικρή διάρκεια (οι πέντε ήταν διάρκειας 2 ns και η μια διάρκειας 1,4 ns). Πραγματοποιώντας πολλές και μικρής διάρκειας προσομοιώσεις αποσκοπούμε στο να προσδιορίσουμε εάν το σύστημα ακολουθεί κάθε φορά το ίδιο trajectory ή διαφορετικό. Είναι αξιοσημείωτο το γεγονός ότι οι πλευρικές ομάδες των καταλοίπων που συμμετέχουν στο σχηματισμό του υδρόφοβου πυρήνα κινούνται γρήγορα, σχηματίζοντας αρκετά γρήγορα κατά τη διάρκεια του trajectory τα μεσαία επίπεδα του πυρήνα. Ταυτόχρονα παρατηρείται αποδιάταξη στα αρχικά επίπεδα του πυρήνα και στην περιοχή της στροφής, λόγω μεγάλης κινητικότητας των καταλοίπων (**Εικόνες 4-70 έως 4-75**).

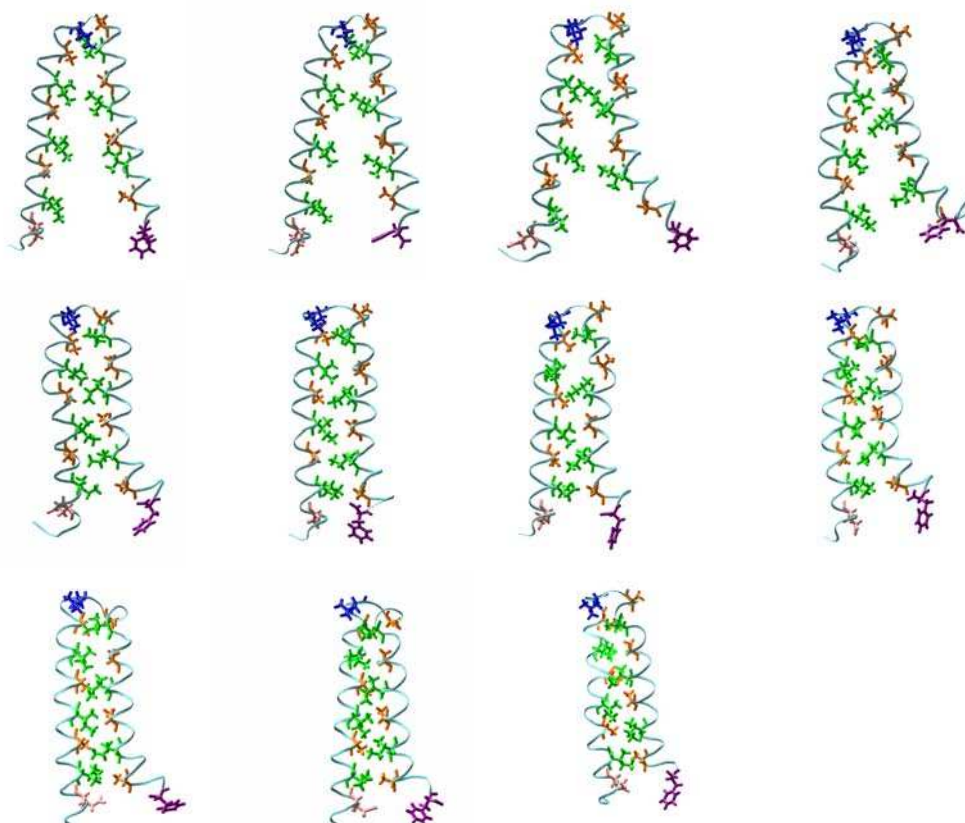
Στην πρώτη προσομοίωση τα έξι από τα οκτώ επίπεδα κλείνουν περίπου στα 2500 frames, στη δεύτερη γύρω στα 2000 frames, στην τρίτη στα 3500 frames, στην τέταρτη στα 4500 frames, στην πέμπτη στα 1500 frames και στην τελευταία προσομοίωση δεν κλείνουν καθόλου.

Όλα αυτά τα γεγονότα μπορούμε να τα παρακολουθήσουμε πιο εύκολα στην τελευταία προσομοίωση του opened μονομερούς που διαρκεί 10ns (**Εικόνα 4-76**). Τα έξι επίπεδα του υδρόφοβου πυρήνα κλείνουν γύρω στα 2000 frames, παρόλα αυτά όμως στα επόμενα 3000 με 4000 frames η δομή αρχίζει να αποδιατάσσεται. Η διαδικασία αυτή επαναλαμβάνεται μέχρι το τέλος της προσομοίωσης χωρίς όμως να κλείνουν όλα τα επίπεδα του υδρόφοβου πυρήνα και χωρίς να αποδιατάσσεται τελείως η δομή, με αποτέλεσμα η δομή να μην είναι καθόλου σταθερή.

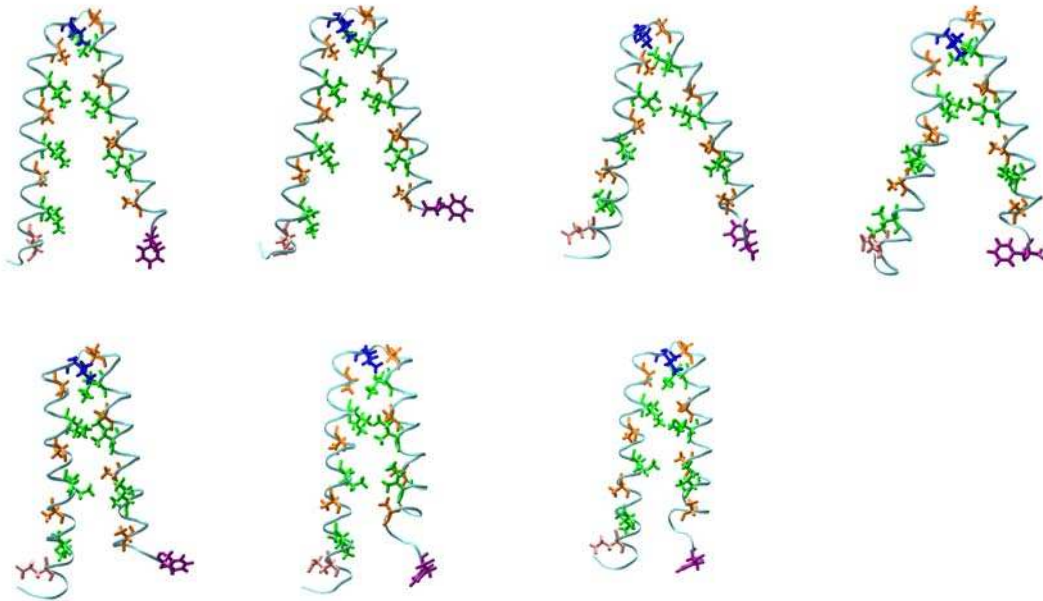




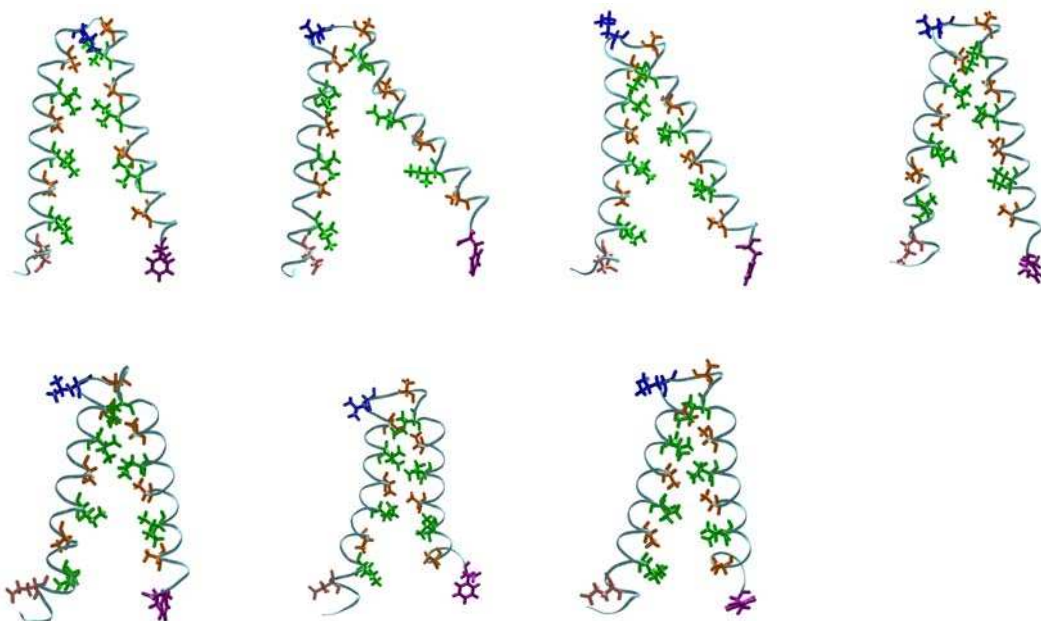
Εικόνα 4-70: Στιγμιότυπα (ανά 500 frames) της προσομοίωσης του *opened* μονομερούς. Ο σκελετός του μορίου απεικονίζεται με *new ribbons* και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με *licorice*. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα Ala, με ροζ τα κατάλοιπα Glu, με πράσινο τα κατάλοιπα Pro, με μπλε τα κατάλοιπα Leu και με μωβ τα κατάλοιπα Phe.



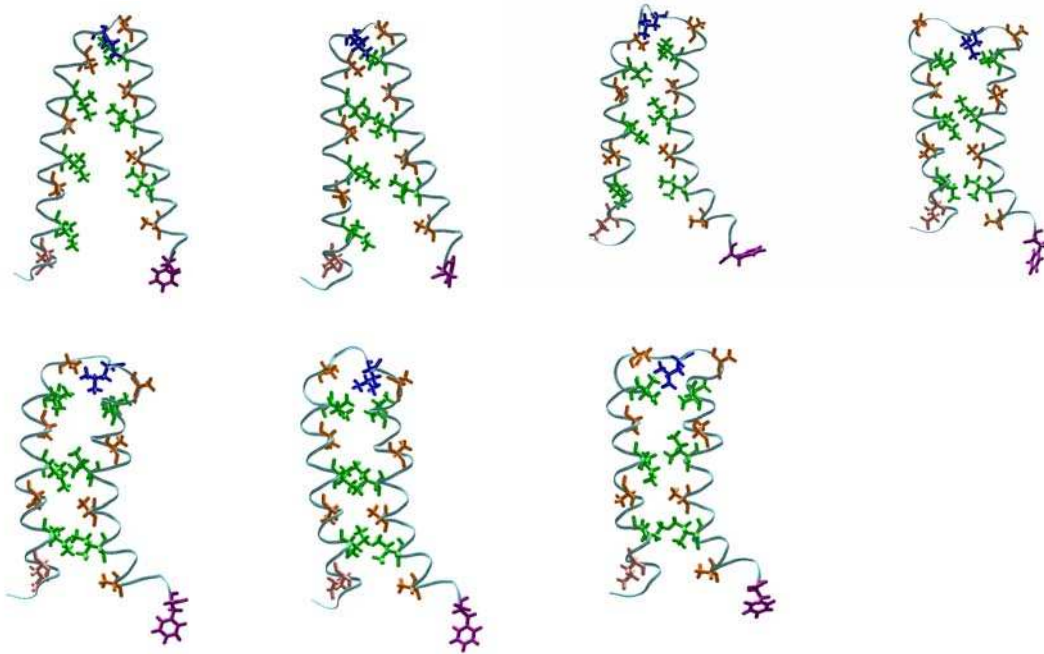
Εικόνα 4-71: Στιγμιότυπα (ανά 500 frames) της προσομοίωσης του *opened* μονομερούς. Ο σκελετός του μορίου απεικονίζεται με *new ribbons* και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με *licorice*. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα Ala, με ροζ τα κατάλοιπα Glu, με πράσινο τα κατάλοιπα Pro, με μπλε τα κατάλοιπα Leu και με μωβ τα κατάλοιπα Phe.



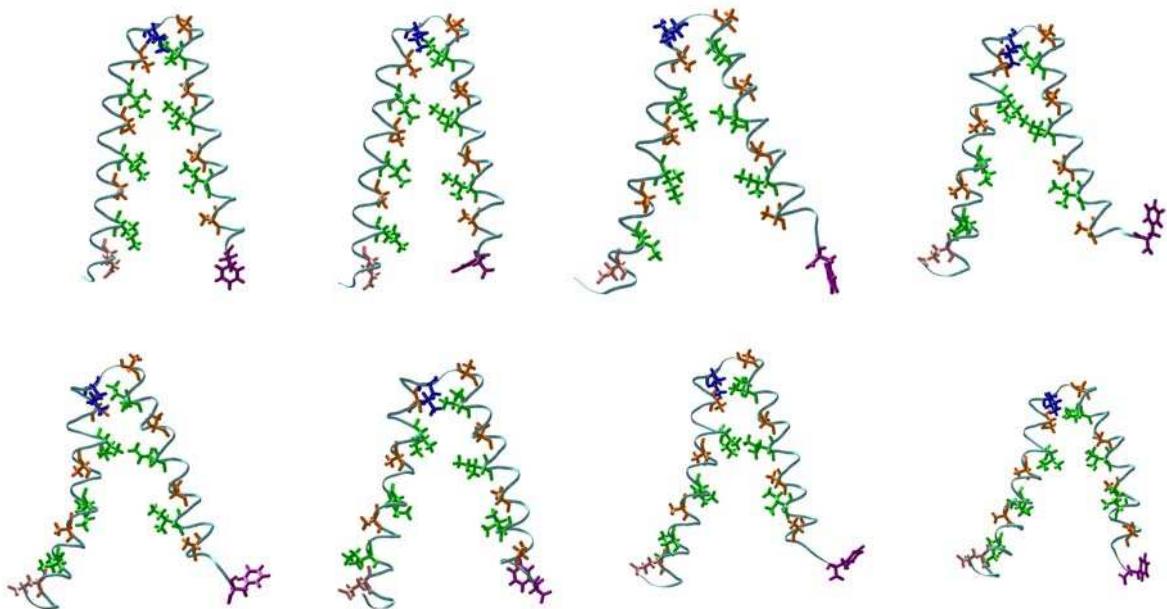
Εικόνα 4-72: Στιγμιότυπα (ανά 1000 frames) της προσομοίωσης του *opened* μονομερούς. Ο σκελετός του μορίου απεικονίζεται με *new ribbons* και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με *licorice*. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα *Ala*, με ροζ τα κατάλοιπα *Glu*, με πράσινο τα κατάλοιπα *Pro*, με μπλε τα κατάλοιπα *Leu* και με μωβ τα κατάλοιπα *Phe*.



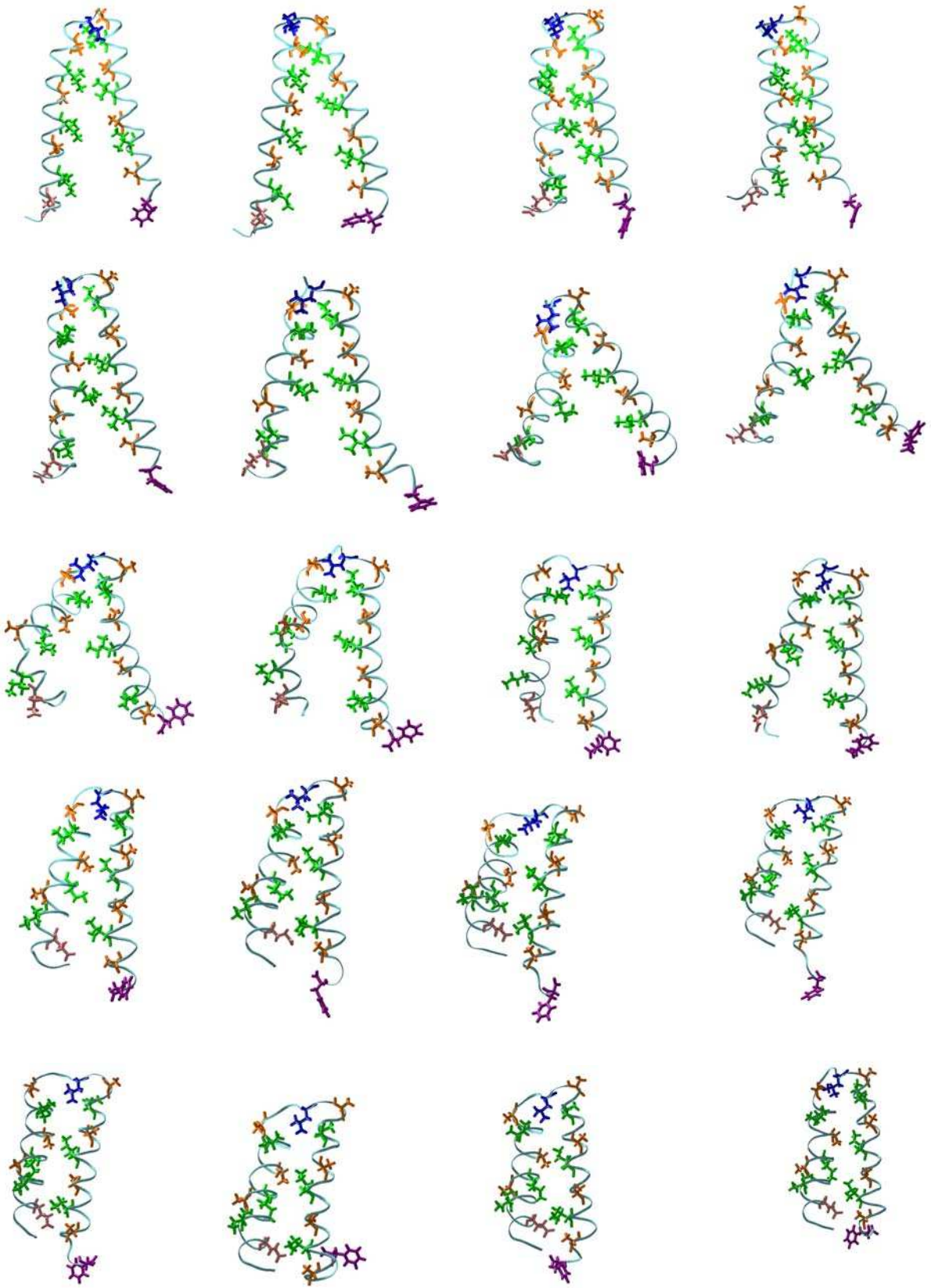
Εικόνα 4-73: Στιγμιότυπα (ανά 1000 frames) της προσομοίωσης του *opened* μονομερούς. Ο σκελετός του μορίου απεικονίζεται με *new ribbons* και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με *licorice*. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα *Ala*, με ροζ τα κατάλοιπα *Glu*, με πράσινο τα κατάλοιπα *Pro*, με μπλε τα κατάλοιπα *Leu* και με μωβ τα κατάλοιπα *Phe*.

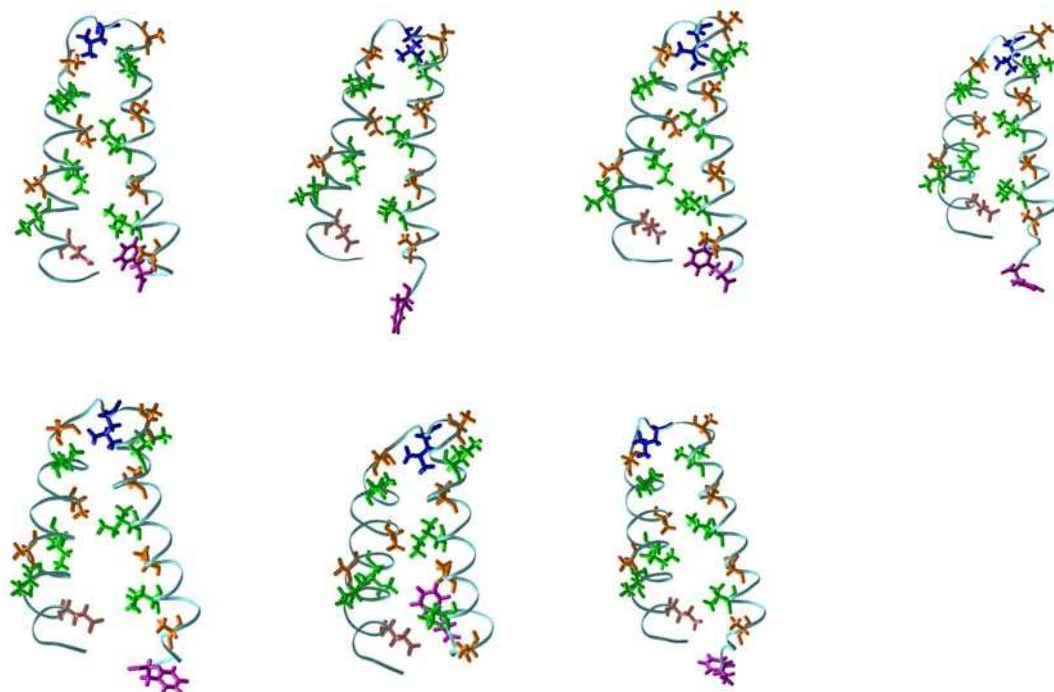


Εικόνα 4-74: Στιγμιότυπα (ανά 1000 frames) της προσομοίωσης του opened μονομερούς. Ο σκελετός του μορίου απεικονίζεται με new ribbons και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με licorice. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα Ala, με ροζ τα κατάλοιπα Glu, με πράσινο τα κατάλοιπα Ile, με μπλε τα κατάλοιπα Leu και με μωβ τα κατάλοιπα Phe.



Εικόνα 4-75: Στιγμιότυπα (ανά 500 frames) της προσομοίωσης του opened μονομερούς. Ο σκελετός του μορίου απεικονίζεται με new ribbons και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με licorice. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα Ala, με ροζ τα κατάλοιπα Glu, με πράσινο τα κατάλοιπα Ile, με μπλε τα κατάλοιπα Leu και με μωβ τα κατάλοιπα Phe.





Εικόνα 4-76: Στιγμιότυπα (ανά 1000 frames) της προσομοίωσης των 10 ns του opened μονομερούς. Ο σκελετός του μορίου απεικονίζεται με new ribbons και οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα με licorice. Με πορτοκαλί χρώμα απεικονίζονται τα κατάλοιπα Ala, με ροζ τα κατάλοιπα Glu, με πράσινο τα κατάλοιπα Ile, με μπλε τα κατάλοιπα Leu και με μωβ τα κατάλοιπα Phe.

ΕΠΙΛΟΓΟΣ

Η Rop είναι μια μικρή ομοδιμερής πρωτεΐνη, η οποία υιοθετεί το μοτίβο του 4-α-ελικοειδούς δεματίου. Το μικρό της μέγεθος δομικά δεν έχει σε τίποτα να κάνει με το τεράστιο μέγεθος ερευνών που βασίζονται σε αυτή. Εδώ και πολλά χρόνια η αναδίπλωσή της, καθώς και η αναδίπλωση των παραλλαγών της αποτελεί αντικείμενο μελέτης. Υπάρχουν παραλλαγές της Rop των οποίων η διαμόρφωση στο χώρο δεν έχει καμία σχέση με εκείνη της φυσικής πρωτεΐνης, όπως για παράδειγμα η Ala_2Ile_2-6 η οποία προέκυψε από τον επανασχεδιασμό του υδρόφοβου πυρήνα. Ξεκίνησε λοιπόν μια διαδικασία για να μελετήσουμε ποια δομή είναι πιο σταθερή, εκείνη της Ala_2Ile_2-6 με τοπολογία *syn* ή εκείνη της Ala_2Ile_2-6 με τοπολογία που μοιάζει με τη φυσική Rop. Από τα αποτελέσματα των προσομοιώσεων βλέπουμε πως δεν υπάρχει κάποιο στοιχείο που να μας ωθεί στο συμπέρασμα ότι κάποια διαμόρφωση είναι σταθερότερη από την άλλη. Επομένως δε μπορούμε να αποκλείσουμε κάποια διαμόρφωση. Πρέπει όμως να σημειώσουμε ότι εάν δε λάβουμε υπ' όψιν τα υπερκινητικά άκρα της Ala_2Ile_2-6 στη *syn* τοπολογία, οι τιμές που προκύπτουν είναι σχετικά πιο χαμηλές. Έτσι είναι επόμενο να υπάρχει μια ελαφρά προτίμηση προς την *syn* τοπολογία όσο αφορά την Ala_2Ile_2-6 . Ίσως τελικά η κάθε πρωτεΐνη να περνά από διάφορες διαμορφώσεις μέχρι να καταλήξει στην πιο σταθερή. Στη βάση αυτή πραγματοποιήθηκαν οι προσομοιώσεις μεμονωμένων μονομερών (native, artificial, opened), με διαφορετικές αρχικές διαμορφώσεις. Ίσως η ανάλυση των δεδομένων αυτών των προσομοιώσεων (η οποία λόγω στενότητας χρόνου δεν πραγματοποιήθηκε), να είναι ικανή να δώσει κάποιες απαντήσεις.

Βιβλιογραφία

1. Allen, M.P. (2004). Introduction to molecular dynamics simulation. NIC Series 23, 1-28.
2. Amadei, A., Linssen, A. B.M. & Berendsen, H. J.C. (1993). Essential dynamics of proteins. *Proteins* 17, 412-425.
3. Anfisen, C.B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223-230.
4. Arthanari, H., McConnell, K. J., Beger, R., Young, M. A., Beveridge, D.L. & Bolton, P. H. (2003). Assessment of the molecular dynamics structure of DNA in solution based on calculated and observed NMR NOESY volumes and dihedral angles from scalar coupling constants. *Biopolymers* 68, 3-15.
5. Baldwin, R.L. (1989). How does protein folding get started? *Trends Biochem. Sci.* 14, 291.
6. Banner, D. W., Kokkinidis, M. & Tsernoglou, D. (1987). Structure of the ColE1 Rop protein at 1.7 Å resolution. *J. Mol. Biol.* 196, 657-675.
7. Bashford, D., Cohen, FE., Karplus, M., Kuntz, LD., & Weaver, DL (1988). Diffusion-collision model for the folding kinetics of myoglobin. *Proteins: Smtct. Funct. Gen.* 4,211.
8. Bernstein, H. J. (1999). Rasmol 2.7.1. Molecular Graphics Visualisation Tool.
9. Brünger, A. T. (1992). X-PLOR Version 3.1. A system for X-ray crystallography and NMR. Connecticut: Yale University Press.
10. Bryngelson, J.D., Onuchic, JN., Socci, N.D. & Wolynes, P.G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins, Struct. Funct. Gen.* 21, 167-195.
11. Caflisch, A. (2004). Protein folding: simple models for a complex process. *Structure* 12, 1750-1752.

12. Castagnoli, L., Scarpa, M., Kokkinidis, M., Banner, D. W., Tsernoglou, D. & Cesareni, G. (1989). Genetic and structural analysis of the ColE1 Rop (Rom) protein. *EMBO J.* 8, 621-629.
13. Castagnoli, L., Vetriani, C. & Cesareni, G. (1994). Linking an easily detectable phenotype to the folding of a common structural motif. Selection of rare turn mutations that prevents the folding of Rop. *J. Mol. Biol.* 237, 378-387.
14. Ceruso, M. A., Periole, X. & Weinstein, H. (2004). Molecular dynamics simulations of transducin: interdomain and front to back communication in activation and nucleotide exchange. *J. Mol. Biol.* 338, 469-481.
15. Chou, K-C, Maggiora, G. M., Nemethy, G. & Scheraga, H. A. (1988). Energetics of the structure of the four- α -helix bundle in proteins. *PNAS* 85, 4295-4299.
16. Chou, K-C, Maggiora, G. M. & Scheraga, H.A. (1992). Role of loop-helix interactions in stabilizing four-helix bundle proteins. *Proc. Natl Acad. Sci. USA* 89, 7315-7319.
17. Christ, D. & Winter, G. (2003). Identification of functional similarities between proteins using directed evolution. *PNAS* 100, 13202-13206.
18. Cohen, C. & Parry, D.A.D. (1986). α -Helical coiled coils a widespread motif in proteins. *Trends Biochem. Sci.* 11, 245-248.
19. Cohen, C. & Parry, D.A.D. (1990). α -helical coiled coils and bundles: how to design an α -helical protein. *Proteins* 7, 1-15.
20. Cregut, D., Drin, G., Liautard, J.-P. & Chiche, L. (1998). Hinge-bending motions in annexins: molecular dynamics and essential dynamics of apo-annexin V and of calcium bound annexin V and I. *Protein Eng.* 11, 891-900.
21. Dalal, S. & Regan, L. (2000). Understanding the sequence determinants of conformational switching using protein design. *Protein Sci.* 9, 1651-1659.
22. Darden, T., York, D. & Pedersen, L. (1993). An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089-10092.
23. Dill, K.A. (1990). Dominant forces in protein folding. *Biochemistry* 29, 7133-7155.
24. Dill, K.A. (1985). Theory for the folding and stability of globular proteins. *Biochemistry* 24, 1501.

25. Dill, K.A. & Chan, H.S. (1997). From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4, 10-19.
26. Dobson, C., Sali, A. & Karplus, M. (1998). Protein folding: A perspective from theory and experiment. *Angew. Chem. Int. Ed.* 37, 7, 868-893.
27. Duan, Y., & Kollman, P. A. (1998). Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 282, 740-744.
28. Eberle, W., Pastore, A., Sander, C. & Rosch, P. (1991). The structure of the ColE1 rop in solution. *J. Biomol. NMR* 1, 71-82.
29. Frishman, D & Argos, P. (1995). Knowledge-based secondary structure assignment. *Proteins* 23, 566-579.
30. Garemyr, R. & Elofsson A. (1999). Study of the Electrostatics Treatment in Molecular Dynamics Simulations. *PROTEINS: Structure, Function and Genetics* 37, 417-428.
31. Glykos, N. M., Cesareni, G. & Kokkinidis, M. (1999). Protein plasticity to the extreme: changing the topology of a 4- α -helical bundle with a single amino acid substitution. *Structure* 7, 597-603.
32. Glykos, N. M. & Kokkinidis, M. (2004). Structural polymorphism of a marginally stable 4- α -helical bundle. Images of a trapped molten globule? *Proteins* 56, 420-425.
33. Gnanakaran, S. & Garcia, A. E. (2005). Helix-coil transition of alanine peptides in water: force field dependence of the folded and unfolded structures. *Proteins* 59, 773- 782.
34. Grayson, P., Tajkhorshid, E. & Schulten, K. (2003). Mechanisms of selectivity in channels and enzymes studied with interactive molecular dynamics. *Biophys. J.* 85, 36- 48.
35. Gruia, A.D., Fischer, S. & Smith, J.C. *Computer Simulation of Protein Unfolding.*
36. Gruia, A.D., Fischer, S. & Smith, J.C. (2003). Molecular Dynamics Simulation Reveals a Surface Salt Bridge Forming a Kinetic Trap in Unfolding of Truncated Staphylococcal Nuclease. *Proteins* 50, 507-515.
37. Haile, J. M. (1997). *Molecular dynamics simulation.* John Wiley & Sons.
38. Harris, N. L., Presnell, S. R. & Cohen, F. E. (1994). Four helix bundle diversity in globular proteins. *J. Mol. Biol.* 236, 1356-1368.

39. Harrison, S.C. & Durbin, R. (1985). Is there a single pathway for the folding of a polypeptide chain? *Proc. Natl. Acad. Sci. USA* 82, 4028-4030.
40. Hess, B. (2002). Convergence of sampling in protein simulations. *Phys. Rev. E* 65, 031910, 107-128.
41. Hummer, G., Schotte, F. & Anfinrud, P. A. (2004). Unveiling functional protein motions with picosecond x-ray crystallography and molecular dynamics simulations. *PNAS* 101, 15330-15334.
42. Humphrey, W., Dalke, A. & Schulten, K. (1996). "VMD — Visual Molecular Dynamics". *J. Mol. Graphics* 14.1, 33-38.
43. Hünenberger, P. H., Mark, A. E. & van Gunsteren, W. F. (1995). Fluctuation and Cross-correlation Analysis of Protein Motions Observed in Nanosecond Molecular Dynamics Simulations. *J. Mol. Biol.* 252, 492-503.
44. Ichiye, T. & Karplus, M. (1991). Collective Motions in Proteins: a Covariance Analysis of Atomic Fluctuations in Molecular Dynamics and Normal Mode Simulations. *Proteins* 11, 205-217.
45. Izaguirre, J. A., Reich, S. & Skeel, R. D (1999). Longer time steps for molecular dynamics. *J. Chem. Phys.* 110, 9853-9864.
46. Kale, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. & Schulten, K. (1999). NAMD2: Greater scalability for parallel molecular dynamics. *J. Comp. Phys.* 151, 283-312.
47. Kamtekar, S. & Hecht, M. H. (1995). The four-helix bundle: what determines a fold? *FASEB* 9, 1013-1022.
48. Karplus, M. (2003). Molecular Dynamics of biological macromolecules: a brief history and perspective. *Biopolymers* 68, 350-358.
49. Karplus, M. & McCammon, J. A. (2002). Molecular dynamics simulation of biomolecules. *Nat. Struct. Biol.* 9, 646-652.
50. Karplus, M. and Weaver, D.L. (1979). Diffusion-collision model for protein folding. *Biopolymers* 18, 1421.
51. Kim, P.S. & Baldwin, R.L. (1990). Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* 59,631.
52. Kleywegt, G.J. (1995). [O/X-PLOR DICTIONARIES] Dictionaries for Heteros, *ESF/CCP4 Newsletter* 31, 45-50.

53. Kresse, H. P., Czubayko, M. Nyakatura, G., Vriend, G., Sander, C. & Bloecker, H (2001). Four-helix bundle topology re-engineered: monomeric Rop protein variants with different loop arrangements. *Protein Eng.* 14, 897-901.
54. Lange, O. F., Grubmuller, H. & de Groot, B. L. (2005). Molecular dynamics simulations of protein G challenge NMR-derived correlated backbone motions. *Angew. Chem. Int. Ed.* 44, 3394-3399.
55. Lassalle, M. W., Hinz, H-J. Wenzel, H., Vlassi, M., Kokkinidis, M. & Cesareni, G. (1998). Dimer-to tetramer transformation: loop excision dramatically alters structure and stability of the ROP four α -helix bundle protein. *J. Mol. Biol.* 279, 987-1000.
56. Lee, B. & Richards, F. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55, 397-400.
57. Levinthal, C. (1968). Are there pathways for protein folding? *J. Chim. Phys.* 65, 44-45.
58. Levy, Y., Cho, S. S., Shen, T., Onuchic, J. N. & Wolynes, P. G. (2005). Symmetry and frustration in protein energy landscapes: A near degeneracy resolves the Rop dimerfolding mystery. *PNAS* 102, 2373-2378.
59. Linhananta, A., Zhou, H. & Zhou, Y. (2002). The dual role of a loop with low loop contact distance in folding and domain swapping *Protein Science* 11, 1695–1701.
60. Luo, B. & Baldwin, R. L. (2002). Origin of the different strengths of the (i, i+4) and (i, i+3) leucine pair interactions in helices. *Biophys. Chem.* 96, 103-108.
61. Luo, J. & Bruice, T. C. (2002). Ten-nanosecond molecular dynamics simulation of the motions of the horse liver alcohol dehydrogenase·PhCH₂O— complex. *PNAS* 99, 16597- 16600.
62. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fisher, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, C., Stote, R., Straub, J., Watanabe, M., Wiorcikiewicz-Kuczera, J., Yin, D. & Karplus, M. (1998). All-atom empirical potential for molecular modelling and dynamic studies of proteins. *J. Phys. Chem. Series B* 102, 3586-3616.

63. Magliery, T. J. & Regan, L. (2004). A cell-based screen for function of the four-helix bundle protein Rop: a new tool for combinatorial experiments in biophysics. *Protein Engineering, Design & Selection* 17, 77-83.
64. Matheson, R.R. and Scheraga, H.A. (1978). A method for predicting nucleation sites for protein folding based on hydrophobic contacts. *Macromolecules* 11, 819.
65. Mayer, K. L., Earley, M. R., Gupta, S., Pichumani, K., Regan, L. & Stone, M. J. (2003). Covariation of backbone motion throughout a small protein domain. *Nat. Struct. Biol.* 10, 962-965.
66. Mayor, U., Johnson, C. M., Daggett, V. & Fersht, A. R. (2000). Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *PNAS* 97, 13518-13522.
67. Merritt, E. A. & Bacon, D. J. (1997). Raster3D – photorealistic molecular graphics. *Methods Enzymol.* 277, 505-524.
68. Moser, D.R., Ma, D., Moser, C. & Campbell J.L. (1984). cis-acting mutations that affect rop protein control of plasmid copy number. *Proc. Natl. Acad. Sci. USA* 81, 4465-4469.
69. Munson, M., Balasubramanian, S., Fleming, K. G., Nagi, A. D., O'Brien, R., Sturtevant, J. M. & Regan, L. (1996). What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Prot. Sci.* 5, 1584-1593.
70. Nagi, A. D., Anderson, K. S. & Regan, L. (1999). Using loop length variants to dissect the folding pathway of a four-helix-bundle protein. *J. Mol. Biol.* 286, 257-265.
71. Nagi, A. D. & Regan, L. (1997). An inverse correlation between loop length and stability in a four-helix-bundle protein. *Fold. Des.* 2, 67-75.
72. Northey, J.G.B. (2001). Protein folding determinants and transition state analysis of the Fyn SH3 domain.
73. Paliakasis, C.D. & Kokkinidis, M. (1992). Relationships between sequence and structure for the four-alpha-helix bundle tertiary motif in proteins. *Protein Eng.* 5, 739-748.
74. Pantano, S., Tyagi, M., Giacca, M. & Carloni, P. (2004). Molecular dynamics simulation on HIV-1 Tat. *Eur. Biophys. J.* 33, 344-351.

75. Peters, K., Hinz, H-J. & Cesareni, G. (1997). Introduction of a proline residue into position 31 of the loop of the dimeric 4-alpha-helical protein ROP causes a drastic destabilization. *Biol. Chem.* 378, 1141-1152.
76. Plaxco, K.W., Riddle, D.S., Grantcharova, V. & Baker, D. (1998). Simplified proteins: minimalist solutions to the 'protein folding problem'. *Curr. Opin. Struct. Biol.* 8, 80-85.
77. Predki, P. F., Agrawal, V., Brünger, A. T. & Regan, L. (1996). Amino-acid substitutions in a surface turn modulate protein stability. *Nat. Struct. Biol.* 3, 54-58.
78. Predki, P. F. Nayak, L. M., Gottlieb, M. B. C & Regan, L. (1995). Dissecting RNAprotein interactions: RNA-RNA recognition by Rop. *Cell* 80, 41-50.
79. Predki, P. F. & Regan, L. (1995). Redesigning the topology of a four-helix-bundle protein: monomeric Rop. *Biochemistry* 34, 9834-9839.
80. Presnell, S. R. & Cohen, F. E. (1989). Topological distribution of four- α -helix bundles. *PNAS* 86, 6592-6596.
81. Price, D. J. & Brooks III, C. L. (2002). Modern protein force fields behave comparably in molecular dynamics simulations. *J. Comput Chem.* 23, 1045-1057.
82. Polisky, B. (1988). ColE1 replication control circuitry: sense from antisense. *Cell* 55, 929-932.
83. R Development Core Team (2004). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051- 07-0. URL <http://www.R-project.org>.
84. Rist, M. J. & Marino, J. P. (2001). Association of an RNA kissing complex analysed using 2-aminopurine fluorescence. *Nucleic Acids Res.* 29, 2401-2408.
85. Rose, G.D. & Creamer, T.P. (1994). Protein folding: Predicting predicting. *Proteins* 19, 502-513.
86. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of nalkanes. *J. Comput. Phys.* 23, 327-341.
87. Seckler, R. & Jaenicke, R. (1992). Protein folding and protein refolding. *FASEB J.* 6, 2545-2552.

88. Shirts, M. R., Pitera, J. W., Swope, W. C. & Pande, V. S. (2003). Extremely precise free energy calculations of amino acid side chain analogs : comparison of common molecular mechanics force fields for proteins. *J. Chem, Phys.* 119, 5740-5761.
89. Smith, L. J., Daura, X. & van Gunsteren, W. F. (2002). Assessing equilibration and convergence in biomolecular simulations. *Proteins* 48, 487-496.
90. Steif, C., Hinz, H-J & Cesareni, G. (1995). Effects of cavity-creating mutations on conformational stability and structure of the dimeric 4- α -helical protein ROP: thermal unfolding studies. *Proteins* 23, 83-96.
91. Steif, C., Weber, P., Hinz, H-J, Flossdorf, J., Cesareni, G. & Kokkinidis, M.(1993). Subunit interactions provide a significant contribution to the stability of the dimeric four- α -helical-bundle protein Rop. *Biochemistry* 32, 3867-3876.
92. Thomas, D. (1990). *J. Mol. Biol.* 216, 459-465.
93. Turner, P.J. (1991-1995) & ACE/gr Development Team (1996-1998), Xmgr v4.1.2.
94. van Drogen-Petit, A., Zwahlen, C., Peter, M. & Bonvin, A. M.J.J. (2004). Insight into molecular interactions between two PB1 domains. *J. Mol. Biol* 336, 1195-1210.
95. van Gunsteren, W. F., Burgi, R., Peter, C. & Daura, X. (2001). The key to solving the protein-folding problem lies in an accurate description of the denatured state. *Angew. Chem. Int. Ed.* 40, 352-355.
96. Vlassi, M., Cesareni, G. & Kokkinidis, M.(1999). A correlation between the loss of hydrophobic core packing interactions and protein stability. *J. Mol. Biol.* 285, 817-827.
97. Vlassi, M., Steif, Ch., Weber, P., Tsernoglou, D., Wilson, K. S., Hinz, H.-J. & Kokkinidis, M. (1994). Restored heptad pattern continuity does not alter the folding a four- α -helix bundle. *Nature Struct. Biol.* 1, 706-714.
98. Wetlaufer, D.B. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Nad. Acad. Sci. USA* 70, 697.
99. Willis, M. A., Bishop, B., Regan, L. & Brünger, A. T. (2000). Dramatic structural and thermodynamic consequences of repacking a protein's hydrophobic core. *Structure* 8, 1319-1328.

100. Wright, P.E., Dyson, H.J. & Lerner, R.A. (1988). Conformation of peptide fragments of proteins in aqueous solution: implications for initiation of protein folding. *Biochemistry* 27, 7167-7175.
101. Wuthrich, K. and Wagner, G. (1978). Internal motion in globular proteins. *Trends Biochem. Sci.* 3, 227.
102. Yamaguchi, H., Aalten, D.M.F., Pinak, M., Furukawa, A. & Osman, R. (1998). Essential dynamics of DNA containing a cis-syn cyclobutane thymine dimer lesion. *Nucleic Acid Res.* 26, 1939-1946.
103. Zagrovic, B., Jayachandran, Millett I. S., Doniach, S. & Pande V. S. (2005). How large is an α -helix? Studies of the radii of gyration of helical peptides by small-angle X-ray scattering and molecular dynamics. *J. Mol. Biol.* 353, 232-241.
104. Zagrovic, B., Sorin, E. J. & Pande, V. (2001). β -hairpin folding simulations in atomistic detail using an implicit solvent model. *J. Mol. Biol.* 313, 151-169.
105. Zielkiewicz, J. (2005). Structural properties of water: comparison of the SPC, SPCE, TIP4P and TIP5P models of water. *J. Chem. Phys.* 123, 104501.
106. <http://origin.imbb.forth.gr/~glykos/carma>, locally written software.



Παράρτηματα

3-1 : Heating up Script

```
#
# Input files
#
structure      ionized.psf
coordinates    ionized.pdb
parameters     par_all27_prot_na.inp
paraTypeCharmm on

#
# Output files & writing frequency for DCD
# and restart files
#
outputname     output/heat_out
binaryoutput   off
restartname    output/restart
restartfreq    1000
binaryrestart   yes
dcdFile        output/heat_out.dcd
dcdFreq        200

#
# Frequencies for logs and the xst file
#
outputEnergies 20
outputTiming   200
xstFreq        200

#
# Timestep & friends
#
timestep        2.0
stepsPerCycle   8
nonBondedFreq  2
fullElectFrequency 4

#
# Simulation space partitioning
#
switching       on
switchDist      10
cutoff          12
pairlistdist    13.5
```

```

#
# Basic dynamics
#
temperature      0
COMmotion        no
dielectric        1.0
exclude          scaled1-4
1-4scaling       1.0
rigidbonds       all

#
# Particle Mesh Ewald parameters.
#
Pme              on
PmeGridsizeX     80      # <===== CHANGE ME
PmeGridsizeY     64      # <===== CHANGE ME
PmeGridsizeZ     60      # <===== CHANGE ME

#
# Periodic boundary things
#
wrapWater        on
wrapNearest      on

cellBasisVector1 78.60 00.00 00.00 # <===== CHANGE ME
cellBasisVector2 00.00 60.10 00.00 # <===== CHANGE ME
cellBasisVector3 00.00 00.00 56.83 # <===== CHANGE ME
cellOrigin       0.00 0.00 2.44 # <===== CHANGE ME

#
# Fixed atoms for initial heating-up steps
#
fixedAtoms        on
fixedAtomsForces  on
fixedAtomsFile    fix_backbone.pdb
fixedAtomsCol     B

#
# Restrained atoms for initial heating-up steps
#
constraints       on
consRef           restrain_ca.pdb
consKFile         restrain_ca.pdb
consKCol          B

```

```

#
# Langevin dynamics parameters
#
langevin      on
langevinDamping  10
langevinTemp   320      # <===== Check me
langevinHydrogen  on

langevinPiston  on
langevinPistonTarget  1.01325
langevinPistonPeriod  200
langevinPistonDecay  100
langevinPistonTemp   320      # <===== Check me

useGroupPressure  yes

#####
# The actual minimisation and heating-up
# protocol follows. The number of steps
# shown below are too small for a real run
#####

#
# run one step to get into scripting mode
#
minimize      0

#
# turn off pressure control until later
#
langevinPiston  off

#
# minimize nonbackbone atoms
#
minimize      2000      ;# <===== CHANGE ME
output        output/min_fix

#
# min all atoms
#
fixedAtoms    off
minimize      2000      ;# <===== CHANGE ME
output        output/min_all

```




```
#
# heat with CAs restrained
#
set temp 20;
while { $temp < 321 } {           ;# <===== Check me
  langevinTemp      $temp
  run                2000         ;# <===== CHANGE ME
  output            output/heat_ca
  set temp [expr $temp + 20]
}

#
# equilibrate volume with CAs restrained
#
langevinPiston      on
run                 10000        ;# <===== CHANGE ME
output              output/equil_ca

#
# equilibrate volume without restraints
#
constraintScaling    0
run                  100000      ;# <===== CHANGE ME
```

3-2 : *Equilibration Script*

```
#
# Input files
#
structure          ionized.psf
coordinates        heat_out.coor
velocities         heat_out.vel
extendedSystem    heat_out.xsc
parameters         par_all27_prot_na.inp
paraTypeCharmm    on

#
# Output files & writing frequency for DCD
# and restart files
#
outputname        output/equi_out
binaryoutput      off
restartname       output/restart
restartfreq       1000
binaryrestart     yes
dcdFile          output/equi_out.dcd
dcdFreq          200

#
# Frequencies for logs and the xst file
#
outputEnergies    20
outputTiming      200
xstFreq           200

#
# Timestep & friends
#
timestep          2.0
stepsPerCycle     8
nonBondedFreq    2
fullElectFrequency 4

#
# Simulation space partitioning
#
switching         on
switchDist        10
cutoff            12
pairlistdist      13.5
```

```

#
# Basic dynamics
#
COMmotion      no
dielectric      1.0
exclude        scaled1-4
1-4scaling     1.0
rigidbonds     all

#
# Particle Mesh Ewald parameters.
#
Pme            on
PmeGridsizeX   80      # <===== CHANGE ME
PmeGridsizeY   64      # <===== CHANGE ME
PmeGridsizeZ   60      # <===== CHANGE ME

#
# Periodic boundary things
#
wrapWater      on
wrapNearest    on
wrapAll        on

#
# Langevin dynamics parameters
#
langevin       on
langevinDamping 1
langevinTemp   320     # <===== Check me
langevinHydrogen on

langevinPiston    on
langevinPistonTarget 1.01325
langevinPistonPeriod 200
langevinPistonDecay 100
langevinPistonTemp 320     # <===== Check me

useGroupPressure yes

firsttimestep   146000     # <===== CHANGE ME
run            5000000     ;# <===== CHANGE ME

```

4-1 : Calculation of *RMSD* from starting structure (*Xplor Script*)

```
structure @../..../..../backup/dimer/ala2_ile2_syn/protein.psf end
```

```
coordinates disposition=comp @psfgen.pdb
```

```
evaluate ($first = 0)
```

```
evaluate ($step = 1)
```

```
evaluate ($current = $first)
```

```
read trajectory
```

```
ascii=false
```

```
input=../..../..../backup/dimer/ala2_ile2_syn/protein.dcd
```

```
begin=$first
```

```
skip=$step
```

```
stop=2000000000
```

```
end
```

```
set display=all.rmsd end
```

```
while ($status # "COMPLETE") loop traj
```

```
coor fit selection=(((name CA or name C or name O or name HN or \  
name HA or name N) and not hydrogen) and not (\  
(resid 56:60 and segid="A ") or\  
(resid 1 and segid="A ") or\  
(resid 2 and segid="B ") or\  
(resid 57 and segid="B ") )) end
```

```
coor rms selection= (((name CA or name C or name O or name HN or \  
name HA or name N) and not hydrogen) and not (\  
(resid 56:60 and segid="A ") or\  
(resid 1 and segid="A ") or\  
(resid 2 and segid="B ") or\  
(resid 57 and segid="B ") )) end
```

```
evaluate ($rms = $RESULT)
```

```
display $current $rms
```

```
evaluate ($current = $current + $step)
```

```
read trajectory next end
```

```
end loop traj
```

```
stop
```

4-2 : Calculation of $RMSD$ from average structure (Xplor Script)

```
structure @../..../..../backup/dimer/ala2_ile2_syn/protein.psf end

coordinates disposition=comp @aver_and_rms.pdb

evaluate ($first = 0)
evaluate ($step = 1)

evaluate ($current = $first)

read trajectory

ascii=false
input=../remove-global-rotation-translation/dcd/ALL_fitted.dcd
begin=$first
skip=$step
stop=20000000
end

set display=aver_vs_MD_notails1.rmsd end

while ($status # "COMPLETE") loop traj

  coor fit selection=( (not hydrogen) and not (\
    (resid 1 and segid "A ") or \
    (resid 56:60 and segid "A ") or \
    (resid 2 and segid "B ") or \
    (resid 57 and segid "B ") ) ) end

  coor rms selection=( (not hydrogen) and not (\
    (resid 1 and segid "A ") or \
    (resid 56:60 and segid "A ") or \
    (resid 2 and segid "B ") or \
    (resid 57 and segid "B ") ) ) end

  evaluate ($rms_ALL_ALL = $RESULT)

  coor fit selection=( (segid="A " and not hydrogen) and not (\
    (resid 1 and segid "A ") or \
    (resid 56:60 and segid "A ") ) ) end
  coor rms selection=( (segid="A " and not hydrogen) and not (\
    (resid 1 and segid "A ") or \
    (resid 56:60 and segid "A ") ) ) end
```

```
evaluate ($rms_ALL_A = $RESULT)
```

```
coor fit selection=( (segid="B " and not hydrogen) and not (\
  (resid 2 and segid "B ") or \
  (resid 57 and segid "B ") ) ) end
```

```
coor rms selection=( (segid="B " and not hydrogen) and not (\
  (resid 2 and segid "B ") or \
  (resid 57 and segid "B ") ) ) end
```

```
evaluate ($rms_ALL_B = $RESULT)
```

```
coor fit selection=( (name ca) and not (\
  (resid 1 and segid "A ") or \
  (resid 56:60 and segid "A ") or \
  (resid 2 and segid "B ") or \
  (resid 57 and segid "B ") ) ) end
```

```
coor rms selection=( (name ca) and not (\
  (resid 1 and segid "A ") or \
  (resid 56:60 and segid "A ") or \
  (resid 2 and segid "B ") or \
  (resid 57 and segid "B ") ) ) end
```

```
evaluate ($rms_CA_ALL = $RESULT)
```

```
coor fit selection=( (name ca and segid="A ") and not(\
  (resid 1 and segid "A ") or \
  (resid 56:60 and segid "A ") ) ) end
```

```
coor rms selection=( (name ca and segid="A ") and not(\
  (resid 1 and segid "A ") or \
  (resid 56:60 and segid "A ") ) ) end
```

```
evaluate ($rms_CA_A = $RESULT)
```

```
coor fit selection=((name ca and segid="B ") and not (\
  (resid 2 and segid "B ") or \
  (resid 57 and segid "B ") ) ) end
```

```
coor rms selection=((name ca and segid="B ") and not (\
  (resid 2 and segid "B ") or \
  (resid 57 and segid "B ") ) ) end
```

```
evaluate ($rms_CA_B = $RESULT)
```

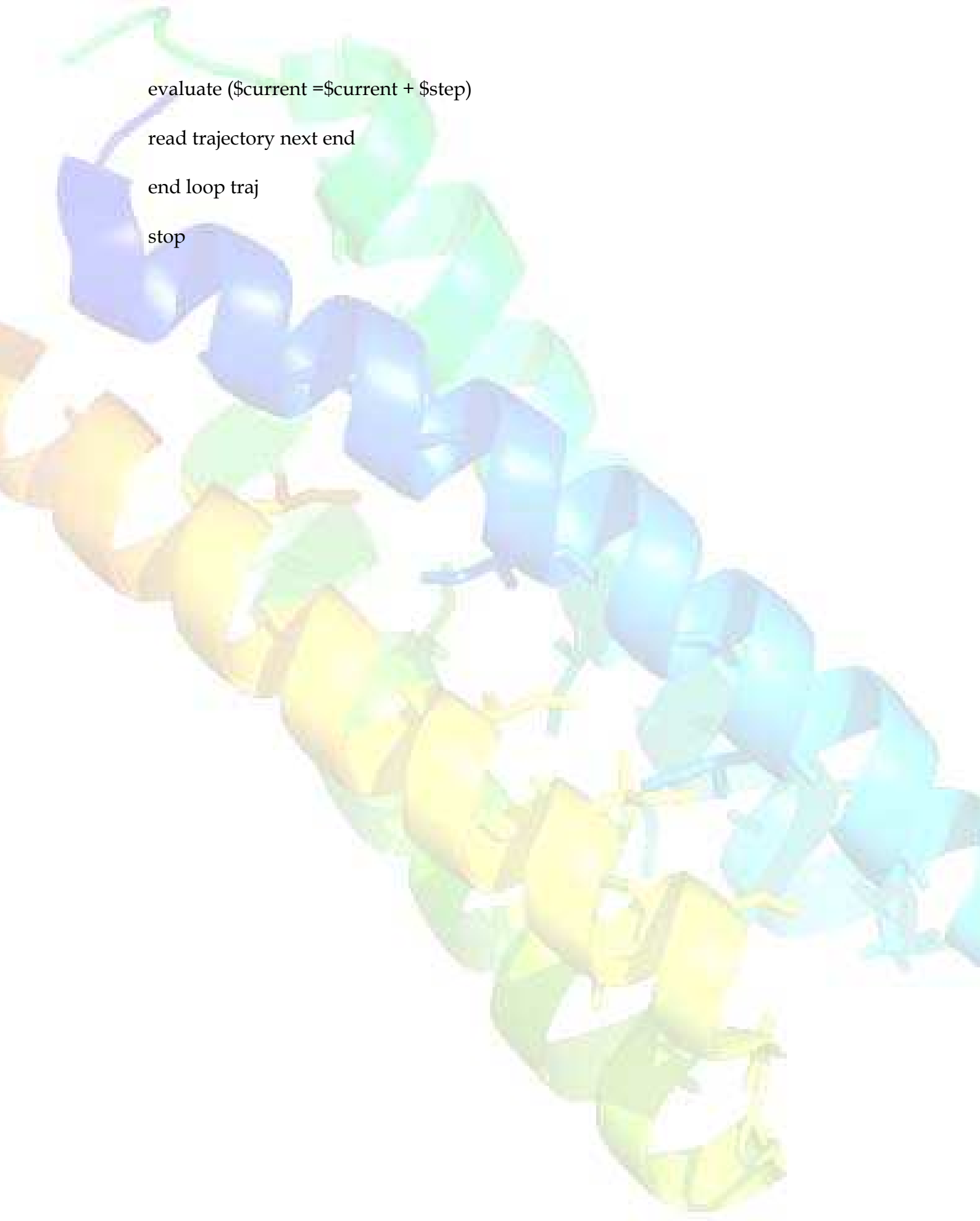
```
display $current $rms_CA_ALL $rms_CA_A $rms_CA_B $rms_ALL_ALL
$rms_ALL_A $rms_ALL_B
```

```
evaluate ($current =$current + $step)
```

```
read trajectory next end
```

```
end loop traj
```

```
stop
```



4-3 : Calculation of Radius of Gyration R_G (Xplor Script)

```
structure @../..../..../backup/dimer/ala2_ile2_syn/protein.psf end

evaluate ($first = 0)
evaluate ($step = 1)

evaluate ($current = $first)

read trajectory

asci=false
input=../remove-global-rotation-translation/dcd/ALL_fitted.dcd
begin=$first
skip=$step
stop=20000000
end

set display=gyration.list end

while ($status # "COMPLETE") loop traj

  coor rgyration end
  evaluate ($rg1 = $RG)

  coor rgyration mass=true end

  display $current $rg1 $RG

  evaluate ($current = $current + $step)

  read trajectory next end

end loop traj

stop
```