

Γεωργούλια Σ. Παπαγιώτα

Διπλωματική Εργασία

Προσομοιώσεις Μοριακής Δυναμικής

ενός τεχνητού α -ελικοειδούς

δεματίου: $[Ala_2Leu_2]_8$

Επιβλέπων Καθηγητής: Νικόλαος Μ. Τσικός

Δημοκρίτειο Πανεπιστήμιο Θράκης
Τμήμα Μοριακής Βιολογίας & Γενετικής

Ιούνιος 2006

ΠΕΡΙΕΧΟΜΕΝΑ

| | |
|--|----|
| Περίληψη | 4 |
| Κεφάλαιο 1 | |
| Εισαγωγή | 6 |
| Κεφάλαιο 2 | |
| Προσομοιώσεις Μοριακής Δυναμικής | 12 |
| α 2.1 Προσομοιώσεις μοριακής δυναμικής της [Ala ₂ Leu ₂] ₈ | 17 |
| Κεφάλαιο 3 | |
| Προετοιμασία Συστημάτων | 20 |
| α 3.1 Υποθέσεις που έχουν διατυπωθεί μέχρι σήμερα | 22 |
| α 3.2 Προετοιμασία συστήματος ομοδιμερούς [Ala ₂ Leu ₂] ₈ με native (Rop-like) τοπολογία | 25 |
| α 3.3 Προετοιμασία συστήματος ομοδιμερούς [Ala ₂ Leu ₂] ₈ με <i>syn</i> τοπολογία | 30 |
| α 3.4 Προετοιμασία συστήματος μονομερούς [Ala ₂ Leu ₂] ₈ με native (Rop-like) αναδίπλωση | 35 |
| α 3.5 Προετοιμασία συστήματος <i>artificial</i> μονομερούς [Ala ₂ Leu ₂] ₈ με native (Rop-like) αναδίπλωση | 37 |
| α 3.6 Προετοιμασία συστήματος μονομερούς [Ala ₂ Leu ₂] ₈ με A31P αναδίπλωση | 45 |
| α 3.7 Προετοιμασία συστήματος μονομερούς [Ala ₂ Leu ₂] ₈ με A31P αναδίπλωση και Pro στη θέση 31 | 48 |

| | |
|---|-----|
| Κεφάλαιο 4 | |
| Ανάλυση Προσομοιώσεων | 50 |
| α 4.1 Ανάλυση προσομοίωσης διμερούς $[Ala_2Leu_2]_8$ | 57 |
| με native (Rop-like) τοπολογία | |
| 4.1.1 Rmsd από την αρχική δομή | 57 |
| 4.1.2 Rmsd από τη μέση δομή | 59 |
| 4.1.3 Rmsf | 61 |
| 4.1.4 Εξέλιξη του Radius of Gyration (R_G) | 62 |
| 4.1.5 Χάρτης της μέσης απόστασης μεταξύ Ca ατόμων και αντίστοιχο rmsd | 63 |
| 4.1.6 Rmsd μεταξύ των δομών του τροχιακού | 64 |
| 4.1.7 Variance-Covariance και cross-correlation | 65 |
| 4.1.8 Eigenvalues και Eigenvectors | 66 |
| 4.1.9 Ομαδοποίηση δομών του trajectory | 75 |
| 4.1.10 Ανάλυση διεδρων γωνιών (χ -value) | 79 |
| 4.1.11 Ανάλυση δευτεροταγούς δομής | 90 |
| 4.1.12 Υπολογισμός της κατανομής των μορίων νερού | 91 |
| α 4.2 Ανάλυση προσομοίωσης διμερούς $[Ala_2Leu_2]_8$ | 94 |
| με <i>syn</i> τοπολογία | |
| 4.2.1 Rmsd από την αρχική δομή | 97 |
| 4.2.2 Rmsd από τη μέση δομή | 99 |
| 4.2.3 Rmsf | 101 |
| 4.2.4 Εξέλιξη του Radius of Gyration (R_G) | 102 |
| 4.2.5 Χάρτης της μέσης απόστασης μεταξύ Ca ατόμων και αντίστοιχο rmsd | 103 |
| 4.2.6 Rmsd μεταξύ των δομών του τροχιακού | 104 |
| 4.2.7 Variance-Covariance και cross-correlation | 106 |
| 4.2.8 Eigenvalues και Eigenvectors | 107 |
| 4.2.9 Ομαδοποίηση δομών του trajectory | 116 |
| 4.2.10 Ανάλυση διεδρων γωνιών (χ -value) | 120 |
| 4.2.11 Ανάλυση δευτεροταγούς δομής | 131 |
| α 4.3 Συγκριτική μελέτη ανάμεσα στην $[Ala_2Leu_2]_8$ | 133 |
| με native (Rop-like) και με <i>syn</i> τοπολογία | |
| α 4.5 Ανάλυση προσομοίωσης <i>native</i> μονομερούς | 137 |
| α 4.6 Ανάλυση προσομοίωσης <i>artificial</i> μονομερούς | 140 |
| α 4.7 Ανάλυση προσομοίωσης <i>opened</i> μονομερούς | 141 |
| α 4.8 Ανάλυση προσομοίωσης <i>opened P31</i> μονομερούς | 150 |
| Επίλογος | 153 |
| Βιβλιογραφία | 154 |
| Παράρτημα | 164 |

ΠΕΡΙΛΗΨΗ

Το πρόβλημα της αναδίπλωσης των πρωτεϊνών (protein folding) συνιστά ένα ιδιαίτερα ενδιαφέρον και ενεργό πεδίο έρευνας. Στα πλαίσια της μελέτης αυτής επιλέχθηκε ως σύστημα μοντέλο η Rop, μία μικρή ομοδιμερής πρωτεΐνη που σχηματίζει ένα 4-α-ελικοειδές δεμάτιο. Η Rop διαδραματίζει σημαντικό ρόλο στον έλεγχο των αντιγράφων του πλασμιδίου ColE1 της *Escherichia coli*, μέσω πρόσδεσης σε ένα σύμπλοκο δύο μικρών μορίων RNA. Παρουσιάζει πληθώρα πλεονεκτημάτων, μεταξύ των οποίων ότι έχει γίνει κινητική και θερμοδυναμική μελέτη και ότι υπάρχει μεγάλη ποικιλία μεταλλάξεων. Μία από τις πιο ενδιαφέρουσες παραλλαγές της, είναι η [Ala₂Leu₂]₈. Σε αυτήν την παραλλαγή ο εσωτερικός υδρόφοβος πυρήνας της πρωτεΐνης έχει επανασχεδιαστεί ώστε να αποτελείται αποκλειστικά από κατάλοιπα αλανίνης και λευκίνης. Η πρωτεΐνη αυτή παρουσιάζει μία αύξηση του χρόνου αναδίπλωσης μέχρι και τέσσερις τάξεις μεγέθους σε σχέση με τη φυσική πρωτεΐνη, δηλαδή πραγματοποιείται στην κλίμακα των ms. Το γεγονός αυτό την καθιστά καλό υποψήφιο για μελέτη γεγονότων αναδίπλωσης με προσομοιώσεις μοριακής δυναμικής. Οι προσομοιώσεις μοριακής δυναμικής συνιστούν χρήσιμα εργαλεία για την κατανόηση, σε φυσική βάση, της δομής και της λειτουργίας των βιολογικών μακρομορίων, καθώς δίνουν έμφαση στη δυναμική τους. Οι υπολογιστικές προσομοιώσεις έχουν χρησιμοποιηθεί εδώ και δεκαετίες από τους ερευνητές για να λύσουν θεμελιώδη προβλήματα της βιολογίας, όπως η αναδίπλωση των πρωτεϊνών. Εξαιτίας της μεγάλης πολυπλοκότητας του προβλήματος έχει αναπτυχθεί μία πληθώρα τεχνικών όπως, ο κατανεμημένος υπολογισμός (distributed computing), οι καθοδηγούμενες προσομοιώσεις (SMD) και οι διαδραστικές προσομοιώσεις (IMD).

Λόγω της μεγάλης διαφοράς που παρουσιάζει στην κινητική της, η [Ala₂Leu₂]₈ (καθώς και κάποιες άλλες παραλλαγές) σε σχέση με τη φυσική πρωτεΐνη, προτάθηκε μία υπόθεση αναφορικά με τη δομή που υιοθετούν τα μόρια αυτά. Έτσι πραγματοποιήθηκαν προσομοιώσεις για δύο συστήματα διμερών με διαφορετικές τοπολογίες, ένα με την τοπολογία της φυσικής πρωτεΐνης και ένα με την τοπολογία της [Ala₂Ile₂]₆. Ωστόσο από τη σύγκριση των αποτελεσμάτων δεν προκύπτει κάποια ισχυρή ένδειξη ότι μία από τις δύο τοπολογίες είναι περισσότερο σταθερή. Προκειμένου να παρακολουθήσουμε ένα πραγματικό γεγονός αναδίπλωσης, που είναι και ο μακροπρόθεσμος στόχος, χρησιμοποιούμε απομονωμένα μονομερή και ξεκινώντας και πάλι από διάφορες αναδιπλώσεις, κάνουμε υπολογιστικά συγκρούσεις (bimolecular collisions). Ωστόσο, στα πλαίσια της εργασίας αυτής, θα περιοριστούμε απλά στη μελέτη κάποιων συστημάτων διμερών και μονομερών μορίων.

ABSTRACT

Protein folding remains one of the most interesting and challenging areas of research. In this study we try to approach the mystery of how a protein finds the way to its tertiary structure through Rop, a small homodimeric protein that folds in a simple 4- α -helix bundle motif. This protein plays an essential role in the regulation of the copy number of plasmid ColE1 in *Escherichia coli*, by binding to an RNA complex. Rop has been extensively studied in thermodynamic and kinetic basis and there are available numerous mutants. One of the most interesting variants that has been described, is [Ala₂Leu₂]₈, which has a fully repacked hydrophobic core consisting exclusively of alanines and leucines. We focus in the particular variant due to a dramatic raise of over four orders of magnitude in the rate of folding (it takes place in the ms scale) and the lack of the characteristic two-phase kinetic of the wild-type protein. This assigns to a very good candidate for folding studies by molecular dynamics simulations. MD emphasizes to the motion and dynamics of a protein, thus provides useful information relative to its structure and function. MD simulations have been used for decades, to solve crucial problems, such as the protein folding. Due to the complication of such problems and the large quantity of data, there have been developed techniques such as distributed computing, steered MD and interactive MD.

Most of the variants of wild type Rop that have been described so far, are assumed to adopt the same tertiary structure as the native protein. However, due to the drastic change in the kinetics of [Ala₂Leu₂]₈ there has been proposed a theory, according to which there are more than one possible foldings for some variants. Thus, we prepared two sets of dimers, one with the native (Rop-like) topology and one with the topology of [Ala₂Ile₂]₆ referred as *syn*. The results of the dynamics show that the two alternative topologies are almost equally stable, with the native (Rop-like) topology slightly more stable. In order to watch an actual folding event, which is the long term goal of this study, we could perform bimolecular collisions of monomers, starting again from a number of different topologies. However, in this study (due to lack of time) we restricted to simple simulations of a number of monomers with different topologies. These data can then be used to perform different sets of bimolecular collisions.

Κεφάλαιο 1

Εισαγωγή

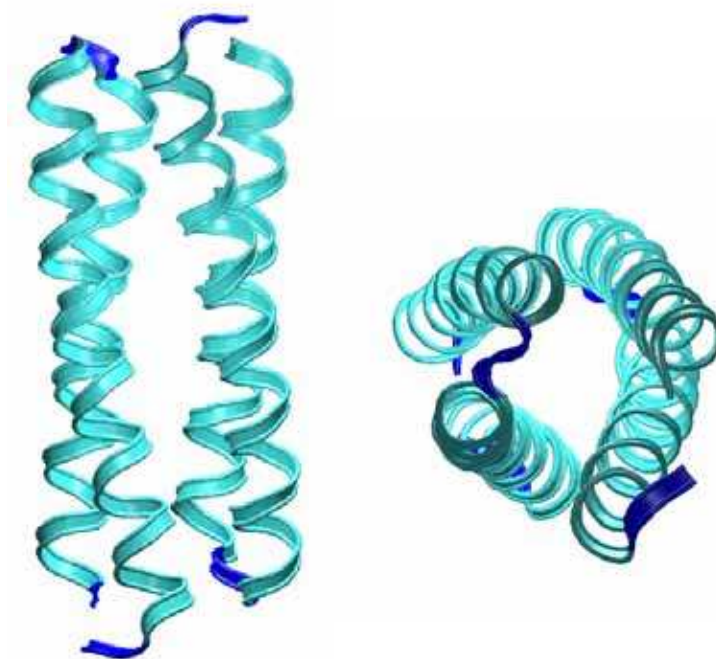
Η αναδίπλωση των πρωτεϊνών (protein folding) είναι μία σύνθετη διαδικασία που περιλαμβάνει μη ομοιοπολικές αλληλεπιδράσεις τόσο σε ολόκληρο το μόριο όσο και ανάμεσα στο μόριο και τον διαλύτη, αρκετούς βαθμούς ελευθερίας και ισορροπία στη συνεισφορά της ενθαλπίας και της εντροπίας στην ελεύθερη ενέργεια. Παλιότερα υπήρχε η πεποίθηση ότι η αναδίπλωση συντελείται μέσα από μία γραμμική αλληλουχία γεγονότων. Οι νεότερες απόψεις υποστηρίζουν ότι πρόκειται για παράλληλες μικροσκοπικές διαδικασίες που μοιάζουν με διάχυση και περιλαμβάνουν πολλαπλά μονοπάτια (parallel microscopic multi-pathway diffusion-like processes) (Caflisch, 2004). Αυτό όμως έρχεται σε αντίθεση με τα πειράματα του Anfinsen, ο οποίος, το 1961, απέδειξε ότι όλη η απαραίτητη πληροφορία για την αναδίπλωση του μορίου περιέχεται στην αμινοξική του αλληλουχία (Anfinsen, 1973). Αυτή η αντιπαράθεση είναι γνωστή και ως παράδοξο του Levinthal (Levinthal's paradox) (Levinthal, 1968). Οι σύγχρονες ιδέες δεν εμπεριέχουν αυτό το παράδοξο καθώς

ελαχιστοποιούν την ιδέα του μονοπατιού: η αναδίπλωση διοχετεύεται (funnel) προς μία σταθερή κατάσταση μέσα από πολλές οδούς διαμόρφωσης στο χώρο (Dill *et al.*, 1997). Έτσι, 45 χρόνια αργότερα, η έρευνα γύρω από τους μηχανισμούς που διέπουν την αναδίπλωση παραμένει ένα ενεργό και εξίσου άγνωστο (όσο και πριν 45 χρόνια) πεδίο. Αυτό που θα πρέπει να εξακριβωθεί είναι κατά πόσο η εξέλιξη έχει επιλέξει αλληλουχίες που αναδιπλώνονται μέσω ενός συγκεκριμένου μονοπατιού αντί να «σχεδιάσει» πρωτεΐνες που μπορούν να αναδιπλωθούν μέσω πολλών μονοπατιών. Ακόμα, γιατί δεν είναι περίεργο το γεγονός ότι υπάρχουν παράλληλα μονοπάτια, αλλά ότι μπορούμε να τα ανιχνεύσουμε και να τα παρακολουθήσουμε πειραματικά (Wright *et al.*, 2003).

Στα πλαίσια της μελέτης της αναδίπλωσης των πρωτεϊνών, επιλέξαμε σαν σύστημα μοντέλο την Rop. Η πρωτεΐνη αυτή έχει χρησιμοποιηθεί κατά κόρον, χάρη στα πλεονεκτήματα που εμφανίζει. Μπορεί να εκφραστεί εύκολα σε μεγάλες ποσότητες, είναι διαλυτή, είναι γνωστή η δομή της τόσο στην κρυσταλλική μορφή (Banner *et al.*, 1987) όσο και στο διάλυμα (Eberle *et al.*, 1991), ενώ έχουν ταυτοποιηθεί τα κατάλοιπα (10, 14, 18 και 25, στην πρώτη έλικα κάθε μονομερούς) που είναι απαραίτητα για τη λειτουργία της (Predki *et al.*, 1995). Η Rop διαδραματίζει σημαντικό ρόλο στον έλεγχο των αντιγράφων του πλασμιδίου ColE1 της *Escherichia coli*, μέσω πρόσδεσης σε ένα σύμπλοκο δύο μικρών μορίων RNA, του RNAI, που έχει ρόλο αναστολέα και είναι συμπληρωματικό με το RNAII που εξυπηρετεί σαν εκκινητής (primer) για την έναρξη της αντιγραφής του DNA από την DNA πολυμεράση I (Polisky, 1988). Ο μηχανισμός δημιουργίας του συμπλόκου έχει μελετηθεί με ένα φθορίζον ανιχνευτή (2-AP, 2-aminopurine-2'-*O*-methyl riboside) και φαίνεται να περιλαμβάνει δύο στάδια (Rist *et al.*, 2001). Η Rop υιοθετεί μία ιδιαίτερα απλή δομή, ένα 4-α-ελικοειδές δεμάτιο, γεγονός που επιτρέπει τη λογική (αλλά και μερικές φορές λανθασμένη) κατανόηση των επιπτώσεων των μεταλλάξεων. Επιπλέον, έχει γίνει εκτεταμένη γενετική και δομική ανάλυση (Castagnoli *et al.*, 1989). Μέχρι πρόσφατα, ένα από τα σημαντικά μειονεκτήματα του συστήματος της Rop ήταν η δυσκολία στην ανίχνευση παραλλαγών (variants) που εμφάνιζαν ενεργότητα και στη μέτρηση αυτής της ενεργότητας. Πλέον, έχει αναπτυχθεί ένας αρκετά αποτελεσματικός τρόπος ελέγχου ενεργότητας μεγάλων βιβλιοθηκών με παραλλαγές της Rop (Magliery *et al.*, 2004).

Το 4-ελικοειδές δεμάτιο είναι ένα αρκετά κοινό μοτίβο στις φυσικές πρωτεΐνες. Σε αυτό το μοτίβο τέσσερις α-έλικες πακετάρονται ισχυρά σε όλο τους το μήκος (**Εικόνα 1.1**). Οι ηλεκτροστατικές αλληλεπιδράσεις συμβάλλουν στην σταθεροποίηση της αντιπαράλληλης διεύθεσης των γειτονικών ελίκων ενώ οι μη ομοιοπολικές αλληλεπιδράσεις που οδηγούν στο ισχυρό πακετάρισμα, καθορίζουν την κατεύθυνση περιστροφής του δεματίου (Chou *et al.*, 1988). Από δομικής πλευράς ένα τέτοιο δεμάτιο μπορεί να απαντάται είτε ως μία αυτόνομη επικράτεια, είτε σαν

τμήμα μίας μεγαλύτερης πρωτεΐνης. Από λειτουργική σκοπιά μπορεί να εξυπηρετήσει διάφορους ρόλους, από αποθήκευση μετάλλων μέχρι ενεργοποίηση κυττάρων. Το μοτίβο του 4-ελικοειδούς δεματίου υπάρχει σε πολλές ιατρικά σημαντικές πρωτεΐνες, όπως η ανθρώπινη αυξητική ορμόνη, η απολιποπρωτεΐνη και οι ιντερλευκίνες. Χάρη στην απλότητα του μοτίβου αυτού σε σχέση με άλλα δομικά μοτίβα, έχει εξυπηρετήσει ως σύστημα μοντέλο τόσο για την κατανόηση της δομής φυσικών πρωτεϊνών όσο και για το σχεδιασμό νέων [(Kamtekar *et al.*, 1995), (Cohen *et al.*, 1990), (Harris *et al.*, 1994)]. Επιπλέον, το 4-ελικοειδές δεμάτιο έχει αποτελέσει ένα έξοχο πεδίο έρευνας στην πρόγνωση περιορισμών που προκύπτουν με βάση την τοπολογία του πρωτεϊνικού σκελετού (Presnell *et al.*, 1989).



Εικόνα 1.1 Αριστερά Άποψη ενός 4-α-ελικοειδούς δεματίου. Με μπλε ανοιχτό απεικονίζονται οι περιοχές α-έλικας (α-helix) και με μπλε σκούρο οι περιοχές θηλιάς (coils). Δεξιά Το ίδιο δεμάτιο μετά από περιστροφή 90° γύρω από έναν άξονα κάθετο στον επιμήκη άξονα του δεματίου.

Όλες οι εικόνες προετοιμάστηκαν με τα προγράμματα Rasmol (Bernstein, 1999), VMD (Humphrey *et al.*, 1996) και Raster3D (Merritt *et al.*, 1997).

Με βάση τα αποτελέσματα που έχουν προκύψει από τις πρώτες μελέτες της Rop, ο Sander πρότεινε την «υπόθεση του πυρήνα» (“core hypothesis”). Σύμφωνα με αυτήν, η ακριβής αλληλουχία στις περιοχές θηλιάς και στην επιφάνεια της πρωτεΐνης είναι δευτερεύουσας σημασίας, υπό τον όρο ότι η αλληλουχία θα είναι συμβατή με το σχηματισμό θηλιάς και την ενυδάτωση της επιφάνειας της πρωτεΐνης, αντίστοιχα (Sander, 1994). Έτσι γεννάται ένα ενδιαφέρον ζήτημα σχετικά με την αναδίπλωση των πρωτεϊνών. Πόσο σημαντικός είναι ο ρόλος που διαδραματίζουν στη διαδικασία της αναδίπλωσης οι περιοχές με δομές στροφής (turns) και θηλιάς (loop); Συνιστούν

απλώς συνδέσμους που παρεμβάλλονται ανάμεσα σε στοιχεία δευτεροταγούς δομής ή διαδραματίζουν περισσότερο ενεργούς ρόλους στο μονοπάτι που θα ακολουθηθεί (folding pathway) και στην τελική αναδιπλωμένη δομή; Μεταγενέστερες μελέτες της Ror και των πολυάριθμων παραλλαγών της φαίνεται να αντιτίθεται στην υπόθεση του Sander.

Έχει δειχθεί ότι το μήκος της θηλιάς έχει επιπτώσεις στην σταθερότητα του ενδιάμεσου μορίου κατά την αναδίπλωση (folding intermediate). Αυτή η αλλαγή στην κινητική της αναδίπλωσης και αποδιάταξης υποδεικνύει μία αλλαγή στην ενέργεια της μεταβατικής κατάστασης (Nagi *et al.*, 1999). Πιο συγκεκριμένα, με την αύξηση του μήκους της θηλιάς, μειώνεται προοδευτικά η σταθερότητα της Ror έναντι στη θερμική και χημική αποδιάταξη (Nagi *et al.*, 1997). Η σημαντικότητα του ρόλου των περιοχών θηλιάς στη διατήρηση της φυσικής δομής και στη σταθερότητα της διαμόρφωσης του μορίου υπογραμμίζεται μέσα από τις μελέτες της A31P, μίας μεταλλαγμένης πρωτεΐνης, στην οποία η αλανίνη στη θέση 31 στην περιοχή της θηλιάς έχει υποκατασταθεί με προλίνη (Glykos *et al.*, 1999). Αυτή η αλλαγή οδηγεί σε μείωση κατά 7% της περιεκτικότητας σε α-έλικα του μορίου και ακολούθως μία μείωση στη σταθερότητα, της τάξεως του 40% σε σχέση με τη φυσική πρωτεΐνη. Αυτή η σημαντική μείωση θεωρείται ότι οφείλεται στην απώλεια ενός τουλάχιστον υδρογονικού δεσμού και στην παραμόρφωση (deformation) της πρωτεϊνικής δομής (Peters *et al.*, 1997). Επίσης, εκτομή της περιοχής της θηλιάς, οδηγεί στη μεταλλαγμένη πρωτεΐνη RM6, η οποία επιδεικνύει δραματικές αλλαγές τόσο στη δομή όσο και στη σταθερότητα. Το μοτίβο μετατρέπεται από ομοδιμερές 4-α-ελικοειδές σε ομοτετραμερές 4-α-ελικοειδές δεμάτιο και η θερμοκρασία τήξης, T_m , αυξάνεται στους 101°C, από 71 °C που παρουσιάζει η φυσική πρωτεΐνη (Lassalle *et al.* 1998).

Σε μία προσπάθεια μελέτης του ρόλου των στροφών στη δομή και τη σταθερότητα της πρωτεΐνης, επιλέχθηκε η θέση 30 για μετάλλαξη και προς τα 19 άλλα αμινοξέα. Από αυτήν την προσπάθεια προέκυψε πως η θέση αυτή είναι ιδιαίτερα ανθεκτική στις μεταλλάξεις καθώς όλες οι παραλλαγές (variants) που προέκυψαν έχουν δομές παρόμοιες με αυτήν της φυσικής πρωτεΐνης. Ωστόσο, 12 από αυτές τις υποκαταστάσεις οδηγούν σε πιο σταθερές πρωτεΐνες, γεγονός που υποδεικνύει ότι υπάρχουν διακριτές θερμοδυναμικές προτιμήσεις όσον αφορά την ταυτότητα του αμινοξικού καταλοίπου στη θέση 30 της αλληλουχίας (Predki *et al.*, 1996). Η περιοχή της στροφής και ιδιαίτερα τα κατάλοιπα 30, 31 και 32 έχουν αποτελέσει στόχο πληθώρας μεταλλάξεων (Castagnoli *et al.*, 1994). Γενικότερα, φαίνεται η ταυτότητα των καταλοίπων στις περιοχές των στροφών να μην είναι καθοριστική για την αναδίπλωση σε ένα 4-ελικοειδές δεμάτιο. Το μήκος της στροφής, όμως μπορεί να είναι σημαντικό για τον καθορισμό της αναδίπλωσης (Kamtekar *et al.*, 1995). Γνωρίζουμε πολύ περισσότερα για τη συμβολή των στροφών

στη σταθερότητα της αναδιπλωμένης μορφής για στροφές που συνδέουν β-κλώνους σε δομές φουρκέτας (β-hairpin) (Lahr *et al.*, 2005).

Ενδεικτικό του πόσο σημαντική είναι η έρευνα για τους παράγοντες που καθορίζουν μία συγκεκριμένη αναδίπλωση (fold) και αποκλείουν τις εναλλακτικές είναι μία πρόκληση που τέθηκε από τους Creamer και Rose, γνωστή ως “Paracelsus Challenge” (Rose *et al.*, 1994). Η πρόκληση ήταν να γίνει μετατροπή μίας πρωτεϊνικής αναδίπλωσης σε κάποια άλλη διατηρώντας τουλάχιστον 50% ομοιότητα στην αλληλουχία, σε σχέση με την αρχική αναδίπλωση. Οι Dalal και Regan σχεδίασαν επιτυχώς μία τέτοια πρωτεΐνη, την οποία ονόμασαν Janus, έχοντας σαν μοντέλο την Rop. Η Janus υιοθετεί ελικοειδή διαμόρφωση και έχει 50% ομοιότητα στην αλληλουχία με την επικράτεια B1 (B1 domain), που κατά κύριο λόγο έχει δομή β-φύλλου (Dalal *et al.*, 2000).

Για την Rop διενεργήθηκαν λεπτομερείς θερμοδυναμικές και φασματοσκοπικές μελέτες για να προσδιοριστεί σε ποσοτική βάση η συνεισφορά των μη ομοιοπολικών αλληλεπιδράσεων στην σταθερότητα των 4-ελικοειδών δεματίων. Από αυτές τις μελέτες προέκυψε ένας μηχανισμός αποδιάταξης δύο σταδίων του τύπου: $N_2 \leftrightarrow 2N^* \leftrightarrow 2D$, όπου N_2 είναι το διμερές στη φυσική (Native) μορφή, N^* είναι τα μονομερή, τα οποία διατηρούν την ελικοειδή διαμόρφωση που έχουν και στο διμερές (όπως προκύπτει με βάση την mean residue ellipticity) και D είναι τα μονομερή στην αποδιαταγμένη κατάσταση (Denatured). Επίσης αποδείχτηκε ότι οι αλληλεπιδράσεις μεταξύ των υπομονάδων είναι μία σημαντική αιτία σταθεροποίησης της φυσικής δομής του 4-ελικοειδούς δεματίου της Rop (Steif *et al.*, 1993).

Χάρη στον απλουστευμένο υδρόφοβο πυρήνα της, η Rop χρησιμοποιήθηκε σε πειράματα σχεδιασμού ή επιλογής εξαιρετικά απλουστευμένων αμινοξικών αλληλουχιών που θα είναι σε θέση να σχηματίσουν σταθερές φυσικές (native ή native-like) πρωτεΐνες (Plaxco *et al.* 1998). Η Rop έχει χρησιμοποιηθεί και στην πρωτεϊνική μηχανική (protein engineering). Εισάγοντας συνδετικές θηλιές (connecting loops) μετατράπηκε σε ένα μονομερές 4-ελικοειδές δεμάτιο, με απώτερο σκοπό τη δημιουργία ενός μικρού και σταθερού μορίου για *in vivo* και *in vitro* παρουσίαση βιολογικά ενεργών πεπτιδίων (Kresse *et al.*, 2001). Τέλος, η Rop χρησιμοποιήθηκε για την εύρεση κρυμμένων ομοιοτήτων με άλλες πρωτεΐνες, μέσω μίας προσέγγισης η οποία δημιουργεί μία σειρά από λειτουργικά μεταλλάγματα (μέσω συνδυαστικής μεταλλαξιγένεσης, combinatorial mutagenesis) 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).

Μία από τις πιο εντυπωσιακές παραλλαγές της Rop είναι η $[Ala_2Leu_2]_8$, της οποίας ο εσωτερικός υδρόφοβος πυρήνας συνίσταται αποκλειστικά από κατάλοιπα αλανίνης και λευκίνης. Η πρωτεΐνη αυτή παρουσιάζει σταθερά αποδιάταξης πέντε τάξεις μεγέθους μεγαλύτερη από τη φυσική πρωτεΐνη. Για την πλήρη αποδιάταξη της Rop χρειάζονται περισσότερες από 24 ώρες, ενώ η $[Ala_2Leu_2]_8$ αποδιατάσσεται πλήρως σε λιγότερο από 30s. Η φυσική πρωτεΐνη παρουσιάζει διφασική κινητική, με την πρώτη φάση να διαρκεί πολύ λιγότερο από 1s και τη δεύτερη, αργή, φάση να ολοκληρώνεται σε αρκετά δευτερόλεπτα. Τόσο η αναδίπλωση όσο και η αποδιάταξη όμως της $[Ala_2Leu_2]_8$ λαμβάνουν χώρα στην κλίμακα των ms (Munson *et al.*, 1997). Χάρη στην γρήγορη κινητική της, η $[Ala_2Leu_2]_8$ προσφέρεται ιδιαίτερα για μελέτες αναδίπλωσης με προσομοιώσεις μοριακής δυναμικής. Επιπλέον, είναι διαθέσιμη πληθώρα θερμοδυναμικών στοιχείων (**Πίνακας 1.1**).

| | MRE (degree cm ² dmol ⁻¹) | T _m (°C) | ΔG° (kcal mol ⁻¹) | C _m (M) | m (kcal mol ⁻¹ M ⁻¹) | Binds RNA |
|------------------|---|------------------------|----------------------------------|-----------------------|--|--------------|
| $[Ala_2Leu_2]_8$ | 28,500 | 91 | -7.5 | 3.0 | 2.8 | yes |
| wild-type | 30,100 | 64 | -7.7 | 3.3 | 2.4 | yes |

Πίνακας 1.1 Σύνοψη της θερμοδυναμικής σταθερότητας και ενεργότητας της $[Ala_2Leu_2]_8$ και της φυσικής πρωτεΐνης, όπου:

MRE είναι η mean residue ellipticity όπως υπολογίστηκε στα 222nm και στους 25°C,

T_m είναι η θερμοκρασία τήξης,

ΔG° είναι η κατά προσέγγιση ελεύθερη ενέργεια του Gibbs απουσία αποδιατακτικού,

C_m είναι η συγκέντρωση του αποδιατακτικού όταν έχει αποδιαταχθεί η μισή πρωτεΐνη και

m είναι η κλίση της ευθείας στην γραφική παράσταση της ΔG με την συγκέντρωση του GuHCl (τα αποτελέσματα αναπαράγονται άνευ αδείας από Munson *et al.*, 1997).

Λόγω της μεγάλης διαφοράς που παρουσιάζει στην κινητική της, η $[Ala_2Leu_2]_8$ (καθώς και κάποιες άλλες παραλλαγές) σε σχέση με τη φυσική πρωτεΐνη, προτάθηκε από τους Levy *et al.* μία υπόθεση αναφορικά με τη δομή που υιοθετούν τα μόρια αυτά (για λεπτομέρειες βλ. Κεφάλαιο 3). Ξεκινώντας από διάφορες δομές (όπως της φυσικής πρωτεΐνης, της $[Ala_2Ile_2]_6$ και της A31P) μπορούμε να εισάγουμε υπολογιστικά τις μεταλλάξεις στα κατάλοιπα του πυρήνα και να παρακολουθήσουμε τη συμπεριφορά των μορίων αυτών. Εξαιτίας της δραματικής αύξησης του χρόνου αναδίπλωσης της συγκεκριμένης παραλλαγής, μακροπρόθεσμος στόχος είναι να παρακολουθήσουμε ένα πραγματικό γεγονός αναδίπλωσης. Ένας τρόπος για να γίνει αυτό, είναι χρησιμοποιώντας απομονωμένα μονομερή και ξεκινώντας και πάλι από διάφορες αναδιπλώσεις, να κάνουμε υπολογιστικά συγκρούσεις (bimolecular collisions). Στα πλαίσια της εργασίας αυτής, θα περιοριστούμε απλά στη μελέτη κάποιων συστημάτων διμερών και μονομερών μορίων.

Κεφάλαιο 2

Προσομοιώσεις Μοριακής Δυναμικής

Οι προσομοιώσεις μοριακής δυναμικής συνιστούν χρήσιμα εργαλεία για την κατανόηση, σε φυσική βάση, της δομής και της λειτουργίας των βιολογικών μακρομορίων. Η αρχική αντιμετώπιση των πρωτεϊνών ως άκαμπτες δομές έχει δώσει τη θέση της σε ένα πιο δυναμικό μοντέλο, στο οποίο οι εσωτερικές κινήσεις (internal motions) που οδηγούν σε αλλαγές στη διαμόρφωση (conformational changes) διαδραματίζουν καίριο ρόλο στη λειτουργία τους.

Η πρώτη προσομοίωση μακρομορίου βιολογικού ενδιαφέροντος δημοσιεύτηκε πριν από τριάντα χρόνια και αφορούσε την πρωτεΐνη BPTI. Το μόριο

αυτό έπαιξε το ρόλο του «υδρογόνου» για τις προσομοιώσεις, λόγω του μικρού του μεγέθους, της υψηλής του σταθερότητας και της σχετικά ακριβούς για την εποχή, κρυσταλλογραφικής του δομής (McCammon *et al.*, 1977). Η όλη ιδέα ξεκίνησε από τους Bruce και Karplus, που επιχείρησαν να αναπτύξουν ένα πρόγραμμα (που αργότερα ονόμασαν Pre-Charmm) το οποίο θα ήταν σε θέση να χρησιμοποιήσει μία δεδομένη αλληλουχία αμινοξέων και μία σειρά συντεταγμένων (όπως αυτές που προκύπτουν κρυσταλλογραφικά) για να υπολογίσει την ενέργεια του συστήματος και των παραγών του ως συνάρτηση των ατομικών θέσεων (Karplus, 2003).

Υπάρχουν τρεις κύριες κατηγορίες εφαρμογών των μεθόδων προσομοίωσης. Η πρώτη χρησιμοποιεί τις προσομοιώσεις ως ένα τρόπο για να μελετήσει διαμορφώσεις στο χώρο (sampling configuration space) και χρησιμοποιείται κυρίως στο προσδιορισμό ή στη βελτιστοποίηση δομών. Η δεύτερη εφαρμογή αφορά στην περιγραφή ενός συστήματος σε κατάσταση ισορροπίας (equilibrium) καθώς και των δομικών και κινητικών του ιδιοτήτων. Στην περίπτωση αυτή θα πρέπει να έχουμε επιτύχει sufficient sampling στο χρόνο της προσομοίωσης (για λεπτομέρειες βλ. Κεφάλαιο 4) και στάθμιση με τον κατάλληλο παράγοντα Boltzmann. Στην τελευταία εφαρμογή, εξετάζεται η ουσιαστική δυναμική του συστήματος, όπου όχι μόνο απαιτείται sufficient sampling και στάθμιση με τον κατάλληλο παράγοντα Boltzmann, αλλά θα πρέπει να γίνει και με τρόπο τέτοιο ώστε να αναπαρασταθεί σωστά η εξέλιξη του συστήματος στο χρόνο. Για τις πρώτες δύο εφαρμογές μπορούν να χρησιμοποιηθούν τόσο οι προσομοιώσεις Monte Carlo όσο και οι προσομοιώσεις μοριακής δυναμικής. Στην τρίτη περίπτωση που το ενδιαφέρον μας έγκειται στις κινήσεις και την εξέλιξή τους στο χρόνο, μόνο οι προσομοιώσεις μοριακής δυναμικής μας παρέχουν τις απαραίτητες πληροφορίες (Karplus *et al.*, 2002).

Ένας τομέας στον οποίο υπήρχε και υπάρχει συμβιωτική σχέση ανάμεσα στις προσομοιώσεις και τα πειράματα, είναι ο πυρηνικός μαγνητικός συντονισμός (NMR, Nuclear Magnetic Resonance). Το NMR, εκτός από την εφαρμογή του στον προσδιορισμό δομών, αποτελεί και ένα ισχυρό εργαλείο για τη μελέτη της δυναμικής και θερμοδυναμικής των μακρομορίων. Το NMR έχει χρησιμοποιηθεί για να διαπιστωθεί η ακρίβεια των μοντέλων της μοριακής δυναμικής (MD, Molecular Dynamics) για τα νουκλεϊκά οξέα. Τα αποτελέσματα που παρατηρήθηκαν μέσω NMR και υπολογίστηκαν μέσω MD παρουσιάζουν μεγάλη συμφωνία μεταξύ τους, γεγονός που συμβάλλει σημαντικά στην αξιοπιστία των προσομοιώσεων (Arthanari *et al.*, 2003). Η αξιοπιστία των προσομοιώσεων ενισχύεται και από την υψηλή πιστότητα των αποτελεσμάτων που προέκυψαν από συνδυασμένη ανάλυση time-resolved x-ray δομών και προσομοιώσεων MD στη μυοσφαιρίνη (Hummer *et al.*, 2004).

Οι υπολογιστικές προσομοιώσεις έχουν χρησιμοποιηθεί εδώ και δεκαετίες από τους ερευνητές για να λύσουν θεμελιώδη προβλήματα της βιολογίας. Πολλές

από αυτές τις μεγάλες προκλήσεις της υπολογιστικής βιολογίας, όπως η αναδίπλωση των πρωτεϊνών, παραμένουν άλυτες εξαιτίας της μεγάλης πολυπλοκότητας του προβλήματος. Για την προσομοίωση της αναδίπλωσης της ταχύτερα αναδιπλούμενης πρωτεΐνης απαιτούνται χρόνια υπολογιστικού χρόνου ακόμα και στους ταχύτερους διαθέσιμους υπολογιστές. Μία καινοτόμος μέθοδος που έρχεται να δώσει λύση στο πρόβλημα της ταχύτητας είναι ο κατανεμημένος υπολογισμός (distributed computing). Με τη μέθοδο αυτή έγινε μία προσπάθεια μελέτης της αναδίπλωσης της καρβοξυτελικής β-φουρκέτας της πρωτεΐνης G, σε ατομική λεπτομέρεια. Η προσομοίωση ήταν της τάξης των 38μs και περιλάμβανε οκτώ ολοκληρωμένα και ανεξάρτητα trajectories (Zagrovic *et al.*, 2001). Εκμεταλλευόμενοι τις εκατοντάδες χιλιάδες υπολογιστών σε ολόκληρο τον κόσμο, καθίσταται δυνατή η υπέρβαση των υπολογιστικών φραγμών. Ωστόσο το distributed computing θέτει νέες προκλήσεις όπως το πώς να κατανεμηθεί επικοινωνητικά ένας πολύπλοκος υπολογισμός σε ένα τόσο μεγάλο αριθμό υπολογιστών που συνδέονται μέσω αργού διαδικτύου. Αλλά ακόμη κι αν βρεθεί λύση σε αυτό, πως μπορούν να χρησιμοποιηθούν τα αποτελέσματα που θα προκύψουν από αυτές τις προσομοιώσεις (που θα είναι αρκετές εκατοντάδες gigabyte δεδομένων) ώστε να δοθούν απαντήσεις σε θεμελιώδη ερωτήματα; Η πρόκληση της ανάλυσης του τεράστιου όγκου των δεδομένων από τέτοιες προσομοιώσεις μεγάλης κλίμακας, θα παραμείνει για αρκετές δεκαετίες (Larson *et al.*, 2003). Οι Izaguirre *et al.* (1999) πρότειναν μία εναλλακτική μέθοδο, την Verlet-I/r-RESPA, η οποία βασίζεται στην τροποποίηση του «αργού» (slow) μέρους του δυναμικού ενέργειας και δίνει τη δυνατότητα μεγαλύτερων βημάτων (timesteps) και κατά συνέπεια επιτρέπει την πραγματοποίηση προσομοιώσεων μεγαλύτερης διάρκειας.

Για την παρακολούθηση γεγονότων που λαμβάνουν χώρα σε μεγαλύτερη κλίμακα χρόνου από αυτήν που μπορούν να παρακολουθήσουν οι προσομοιώσεις μοριακής δυναμικής, όπως μεγάλες αλλαγές στη στερεοδιαμόρφωση των μακρομορίων, είσοδος διαλυτών μέσα από κανάλια, δέσμευση και αποδέσμευση προσδετών (ligands) από τις θέσεις πρόσδεσής τους, υπάρχουν εναλλακτικές τεχνικές. Σε αυτήν την κατηγορία ανήκουν οι καθοδηγούμενες προσομοιώσεις (SMD, steered molecular dynamics) στις οποίες με εφαρμογή εξωτερικής δύναμης καθοδηγείς το σύστημα προς μία συγκεκριμένη διαμόρφωση (Izrailev *et al.*, 1998) και οι διαδραστικές προσομοιώσεις (IMD, interactive molecular dynamics) στις οποίες χειρίζεσαι ενεργά τα συστήματα των βιομορίων με σκοπό τη διερεύνηση δυναμικών διεργασιών [(Rapaport, 1997), (Grayson *et al.*, 2003)].

Οι προσομοιώσεις μοριακής δυναμικής συνιστούν τη μόνη μέθοδο που μπορεί να ακολουθήσει, σε ατομικό επίπεδο, την πορεία μίας πρωτεΐνης μέσα από διάφορες συνθήκες, στο μονοπάτι που ακολουθεί και παρέχει τη δυνατότητα απευθείας παρατήρησης της ατομικής δομής των συστατικών του συνόλου των

δομών (ensemble) (DeMarco *et al.*, 2004). Οι πρωτεΐνες που επιδεικνύουν ταχεία κινητική αναδίπλωσης και αποδιάταξης αποτελούν εξαιρετους υποψήφιους για μελέτες μοριακής δυναμικής.

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 (unfolding half-life) υπολογίζεται κατά προσέγγιση σε 7.5ns στους 100°C. Το γεγονός αυτό επιτρέπει την δοκιμασία, σε πραγματικό χρόνο, προσομοιώσεων αποδιάταξης (real-time molecular dynamics unfolding simulations) του συστήματος αυτού, σε ρεαλιστική θερμοκρασία (Mayor *et al.*, 2000). Η μεγαλύτερης διάρκειας προσομοίωση που είναι γνωστή είναι της τάξης του 1μs και αφορά την αναδίπλωση μίας μικρής πρωτεΐνης 36 καταλοίπων, την επικράτεια κεφαλής της villin (villin headpiece subdomain) (Duan *et al.*, 1998).

Μία από τις μεγαλύτερες προκλήσεις των προσομοιώσεων μοριακής δυναμικής των πρωτεϊνών και των νουκλεϊκών οξέων είναι η εύρεση της σωστής (πειραματικά προσδιορισμένης) γεωμετρίας όταν η προσομοίωση ξεκινά από μία σημαντική απόσταση από αυτήν. Οι Cheatham και Kollman έδειξαν μία μετάβαση από την δομή A-DNA στη B-DNA, το οποίο σημαίνει μία απόκλιση στο rmsd της τάξης των 6Å, σε μόλις 0.5ns. Το γεγονός ότι ανεξάρτητες γεωμετρίες συγκλίνουν σε μία κοινή δομή, η οποία έρχεται σε συμφωνία με πειραματικά δεδομένα, χρησιμοποιώντας ένα μοντέλο μοριακής δυναμικής που κάνει χρήση state of the art αλγορίθμων, πρωτοκόλλων προσομοίωσης και αναπαράσταση ηλεκτροστατικών αλληλεπιδράσεων μεγάλης κλίμακας, είναι ιδιαίτερα ενθαρρυντικό (Cheatham *et al.*, 1996).

Οι προσομοιώσεις μοριακής δυναμικής έχουν χρησιμοποιηθεί σε μία μεγάλη ποικιλία εφαρμογών όπως, στη διερεύνηση της φύσης της παραμόρφωσης του DNA που προκαλείται από ένα διμερές κυκλοβουτανίου πυριμιδίνης (cyclobutane pyrimidine dimer) (Yamaguchi *et al.*, 1998) και στη μελέτη της διαδικασίας ανταλλαγής νουκλεοτιδίου της υπομονάδας Gα της μεταγωγάσης Gα_i (transducin), από όπου προέκυψε ένας μηχανισμός ανάλογος με αυτόν που έχει βρεθεί στις GTPάσες της οικογένειας Arf (Ceruso *et al.*, 2004). Ακόμα, εφαρμόστηκαν στη μελέτη των ακουαπορινών (aquaporins), διαμεμβρανικών καναλιών που απαντώνται στις κυτταρικές μεμβράνες. Η «παράδοξη» λειτουργία τους έγκειται στο ότι επιτρέπουν την εισχώρηση μορίων νερού και ταυτόχρονα αποκλείουν τα πρωτόνια, διαδικασία που είναι εξαιρετικής σημασίας για τη διατήρηση του ηλεκτροχημικού δυναμικού κατά μήκος της κυτταρικής μεμβράνης. Με προσομοιώσεις των 12ns προσδιορίστηκε χωρικά και χρονικά η πιθανότητα της κατανομής και του προσανατολισμού μίας σειράς επτά έως εννιά μορίων νερού μέσα στο κανάλι (Tajkhorshid *et al.*, 2002). Οι προσομοιώσεις επιτρέπουν την διερεύνηση της δυναμικής και των διαμοριακών αλληλεπιδράσεων σε πρωτεΐνες όπως ένας 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. Τέτοιες πληροφορίες συνεισφέρουν στην κατανόηση της λειτουργίας της πρωτεΐνης και ακολούθως στη στοχευμένη απενεργοποίησή της (Pantano *et al.*, 2003). Με τη βοήθεια των προσομοιώσεων επιβεβαιώθηκε και ο μηχανισμός που έχει προταθεί για τις αννεξίνες (annexins). Τα μέλη της οικογένειας αυτής σχηματίζουν τασοελεγχόμενους διαύλους καλίου στις λιπιδικές διπλοστοιβάδες. Θεωρείται ότι ο υδρόφιλος πόρος στο κέντρο του μορίου συνιστά το μονοπάτι αγωγής των ιόντων και ότι η πύλη του διαύλου ανοίγει μέσω μίας περιστροφικής κίνησης (hinge). Οι αναλύσεις των MD και ED (essential dynamics) όντως έδειξαν τέτοιου είδους (hinge-bending) κινήσεις (Cregut *et al.*, 1998).

Ένας άλλος τομέας με ιδιαίτερο ενδιαφέρον είναι οι αλληλεπιδράσεις μεταξύ πρωτεϊνών, οι οποίες διαδραματίζουν καίριους ρόλους στη ρύθμιση ποικίλων βιολογικών διεργασιών. Πρόσφατα, έχει χαρακτηριστεί μία επικράτεια, η PB1 (Phox και Bem1), η οποία έχει συντηρηθεί κατά την εξέλιξη και είναι παρούσα σε εξελικτικά απομακρυσμένες πρωτεΐνες που συμμετέχουν στη μεταγωγή σήματος, στην πολικότητα του κυττάρου και στην επιβίωση. Η επικράτεια αυτή και η συμπεριφορά της αποτέλεσε αντικείμενο μελέτης προσομοιώσεων (Drogen-Petit *et al.*, 2004).

Μία διαφορετική προσέγγιση του προβλήματος της αναδίπλωσης των πρωτεϊνών έγινε μελετώντας την αντίστροφη πορεία της αναδίπλωσης (reversible folding) ενός επταπεπτιδίου β-αμινοξέων. Με ανάλυση τέτοιου είδους προσομοιώσεων, με πεπτίδια διαφορετικής σύνθεσης, σε διάφορους διαλύτες και σε διαφορετικές θερμοκρασίες, προέκυψαν διαφορετικά στοιχεία δευτεροταγούς δομής (ή αναδιπλώσεις). Όμως η αποδιαταγμένη κατάσταση αυτών των πεπτιδίων αποτελείται από σημαντικά λιγότερες διαμορφώσεις από αυτές που θα περίμενε κανείς βάση του αριθμού των βαθμών ελευθερίας. Έτσι, ενώ θεωρητικά ένα δεκαπεπτίδιο μπορεί να δώσει περίπου 10^9 διαφορετικές διαμορφώσεις (conformers), ένα πεπτίδιο με 20 περιστρέψιμους (rotatable) δεσμούς στην αποδιαταγμένη κατάσταση παρουσιάζει μόλις $10^2 - 10^3$ conformers. Ο μικρότερος αριθμός των αποδιαταγμένων καταστάσεων εξηγεί γιατί τα πεπτίδια είναι σε θέση να αναδιπλώνονται στην κλίμακα των nanosecond (van Gunsteren *et al.*, 2001).

2.1 Προσομοιώσεις μοριακής δυναμικής της [Ala₂Leu₂]₈

Στην παρούσα εργασία οι προσομοιώσεις πραγματοποιήθηκαν σε μία συστοιχία υπολογιστών τύπου Beowulf, χρησιμοποιώντας το charmm27 force field που παρείχε τις πληροφορίες σχετικά με την τοπολογία και άλλες παραμέτρους της πρωτεΐνης (MacKerell *et al.*, 1998). Το force field (δυναμικό πεδίο) είναι μία μαθηματική έκφραση της δυναμικής ενέργειας των ατόμων του συστήματος. Υπάρχουν διάφορα force fields για πρωτεΐνες, όπως AMBER, CHARMM, OPLS-AA, τα οποία φαίνεται να δίνουν παρόμοια αποτελέσματα και είναι αρκετά αξιόπιστα [(Gnanakaran *et al.*, 2005), (Shirts *et al.*, 2003), (Price *et al.*, 2002)]. Το αρχείο με τις παραμέτρους περιέχει πληροφορίες όπως η ισχύς και τα μήκη ισορροπίας των ομοιοπολικών δεσμών (bond strengths, equilibrium lengths). Το αρχείο με την τοπολογία περιέχει πληροφορία αναφορικά με τους ατομικούς τύπους, τα φορτία και τον τρόπο σύνδεσης των ατόμων στο μόριο. Η πληροφορία για τις συντεταγμένες των ατόμων περιέχεται στο αρχείο της PDB. Το επόμενο βήμα είναι να ευθυγραμμίσουμε τους αδρανειακούς άξονες (axes of inertia) του μορίου με το ορθοκανονικό σύστημα αξόνων (orthogonal frame) και να θέσουμε το κέντρο βάρους του στην αρχή των αξόνων (με τη βοήθεια του προγράμματος moleman, Kleywegt, 1995). Επειδή η προσομοίωση πραγματοποιείται σε όξινο και όχι σε ουδέτερο pH, οι ιστιδίνες θα πρέπει να απαντούν σε πλήρως πρωτονιωμένη μορφή, δηλαδή και τα δύο άτομα αζώτου να είναι πρωτονιωμένα (οπότε αντί για HIS θα χρησιμοποιήσουμε HSP). Εναλλακτικά, μπορεί να είναι πρωτονιωμένο μόνο το δ άζωτο (οπότε αναφέρεται σαν HSD), ή μόνο το ε (οπότε αναφέρεται σαν HSE). Τέλος, χρειαζόμαστε ένα αρχείο psf (protein structure file) το οποίο αποθηκεύει τη δομική πληροφορία της πρωτεΐνης, όπως τους διάφορους τύπους των δεσμικών αλληλεπιδράσεων και ένα αρχείο pdb που περιέχει αρχικές συντεταγμένες των ατόμων (ακόμα και εκείνων που απουσιάζουν από το αρχικό αρχείο pdb). Το psf δημιουργείται με τη βοήθεια του psfgen, το οποίο περιλαμβάνεται στο NAMD (Kale *et al.*, 1999). Έτσι, έχουμε ολοκληρώσει με τις πληροφορίες που απαιτούνται για την πρωτεΐνη.

Στη συνέχεια ακολουθεί η προετοιμασία του συστήματος η οποία περιγράφεται αναλυτικά στο επόμενο κεφάλαιο. Στα πλαίσια αυτής της διαδικασίας η πρωτεΐνη ενυδατώνεται σε ένα κουτί με νερά και ιόντα με σκοπό την αναπαράσταση ενός κατά το δυνατόν φυσιολογικού υδατικού περιβάλλοντος.

Τα μόρια νερού τα οποία χρησιμοποιήσαμε στο σύστημα είναι τύπου TIP3. Έχει δείχθει πως τα SPC και TIP4 περιγράφουν καλύτερα τις δομικές ιδιότητες του νερού και αναπαράγουν με μεγαλύτερη ακρίβεια τις πειραματικά προσδιορισμένες τιμές, για αυτό και προτείνονται για προσομοιώσεις βιομορίων σε υδατικά διαλύματα

(Zielkiewicz, J., 2005). Ωστόσο, τα TIP3 αναπαριστούν καλύτερα την ενέργεια και πυκνότητα του υγρού νερού, φαίνεται να είναι πιο κοντά σε αυτό που συμβαίνει στην πραγματικότητα, ταιριάζουν περισσότερο με τις συνθήκες που εφαρμόζουμε στο σύστημα (NpT), δηλαδή σταθερή πίεση και θερμοκρασία (Jorgensen *et al.*, 1983) και επιπλέον η παραμετροποίηση τους έγινε μαζί με τις υπόλοιπες παραμέτρους του CHARMM και με αυτόν τον τρόπο εξασφαλίζουμε συμβατότητα.

Η παρουσία των ιόντων προασπίζει ηλεκτροστατικά τις περιοχές της πρωτεΐνης που φέρουν φορτία και έτσι συμβάλλουν στη σταθερότητα του συστήματος. Τυπικά τα ιόντα πρέπει να τοποθετηθούν σε περιοχές που έχουμε τοπικά ελάχιστα του δυναμικού, αφού κατά τη διάρκεια της προσομοίωσης θα «αναγκασθούν» να κινηθούν προς αυτές τις περιοχές. Το φορτίο του συστήματος περιέχεται στο αρχείο psf. Στις προσομοιώσεις μοριακής δυναμικής με συνθήκες περιοδικής οριοθέτησης 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, όπως συμβαίνει εδώ, ο πιο εύκολος τρόπος να τοποθετηθούν τα ιόντα στο σύστημα είναι αυτόματα με τη βοήθεια του VMD (Humphrey *et al.* 1996). Στην περίπτωση αυτή υπολογίζεται πόσα ιόντα νατρίου και χλωρίου απαιτούνται και τοποθετούνται τυχαία στο διάλυμα, αλλά τηρούνται ελάχιστες αποστάσεις τόσο ανάμεσα στα ιόντα και το μόριο, όσο και μεταξύ των ιόντων. Εάν όμως έχουμε να εξουδετερώσουμε μεγάλο φορτίο, όπως για παράδειγμα όταν η μελέτη μας αφορά ένα τμήμα δίκλωνου DNA, το οποίο φέρει δύο αρνητικά φορτία ανά ζεύγος βάσεων, η αυτόματη τοποθέτηση των ιόντων δεν θα έχει σαν αποτέλεσμα μία ομοιόμορφη κατανομή τους στο διάλυμα, αλλά τη δημιουργία ενός νέφους γύρω από το μόριο. Σε αυτήν την περίπτωση θα πρέπει να βρεθεί ένας εναλλακτικός τρόπος, ώστε να τοποθετηθούν με βάση το ηλεκτροστατικό δυναμικό του μορίου.

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 συνθήκες περιοδικής οριοθέτησης (periodic boundary conditions). Κατά τις συνθήκες περιοδικής οριοθέτησης τοποθετούνται πανομοιότυπες μονάδες γύρω από το προς μελέτη σύστημα και στις τρεις διαστάσεις. Τα περιβάλλοντα συστήματα αλληλεπιδρούν με τα άτομα του πραγματικού συστήματος και έτσι δημιουργείται ένα περιβάλλον στο οποίο το σύστημα δεν «βλέπει» κενό. Οι συνθήκες αυτές ελαττώνουν αποτελεσματικά την επιφανειακή αλληλεπίδραση των μορίων νερού, με αποτέλεσμα όμως το σύστημα να διαφέρει σε κάποιο βαθμό από ένα περιβάλλον το οποίο να είναι πλησιέστερα στην κρυσταλλική κατάσταση (Haile 1997).

Μετά την προετοιμασία του συστήματος ακολουθεί η ελαχιστοποίηση ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Τα βήματα που

ακολουθούνται περιγράφονται αναλυτικά στο επόμενο κεφάλαιο. Κατά την παραγωγική φάση (production phase) 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 ταχύτητας επιλύεται ο αλγόριθμος Verlet. Τα σημαντικά χαρακτηριστικά του αλγόριθμου Verlet είναι πως είναι απευθείας χρονικά αναστρέψιμος (exactly time reversible), συμπλεκτικός (symplectic), έχει χαμηλή τάξη μεγέθους στο χρόνο (low order in time) επιτρέποντας μεγαλύτερα βήματα (timesteps), απαιτεί μόνο έναν (αλλά χρονοβόρο) υπολογισμό δύναμης σε κάθε βήμα και είναι εύκολος ο προγραμματισμός (Allen, 2004). Από τη στιγμή που οι εξισώσεις κίνησης λύνονται κατά προσέγγιση σε διακριτά χρονικά βήματα, οι περιορισμοί (constraints) δεν θα τηρούνται όσο προχωρά η προσομοίωση. Μία λύση σε αυτό το πρόβλημα είναι να προσδιοριστούν οι δυνάμεις περιορισμού (constraint forces) με τρόπο τέτοιο 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

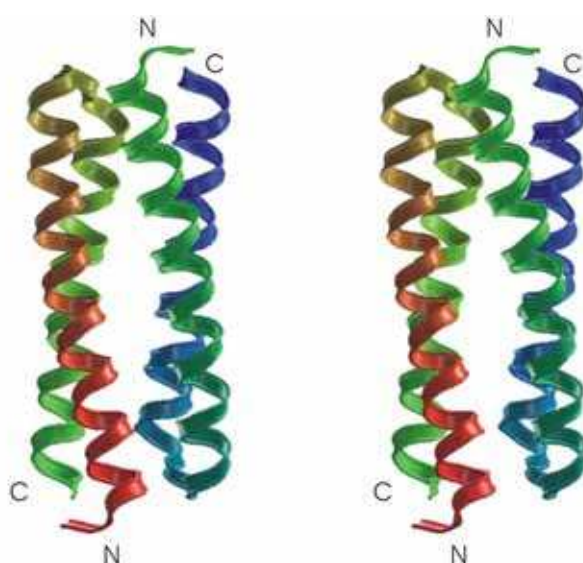
Στο κεφάλαιο αυτό αρχικά γίνεται μία ανασκόπηση των υποθέσεων που έχουν διατυπωθεί σχετικά με τη δομή που υιοθετούν οι διάφορες παραλλαγές (variants) της Rop σε σχέση με τη φυσική πρωτεΐνη. Ακολουθεί η περιγραφή και η προετοιμασία του κάθε συστήματος που επιλέχθηκε για μελέτη. Σε κάθε περίπτωση έχουμε σαν βάση την αλληλουχία της $[Ala_2Leu_2]_8$ και χρησιμοποιούμε διαφορετικές τοπολογίες, τόσο στη μορφή του ομοδιμερούς όσο και στη μορφή του μονομερούς. Στον **Πίνακα 3.1** παραθέτονται συγκριτικά όλα τα πειράματα που διεξήχθησαν και τα χαρακτηριστικά του κάθε συστήματος.

| | Όνομα μορίου | Διάρκεια προσομοίωσης | Διαστάσεις box | Συνολικός αριθμός ατόμων συστήματος | Αριθμός ατόμων πρωτεΐνης | Αριθμός ατόμων νερού | Αριθμός ιόντων Na ⁺ | Αριθμός ιόντων Cl ⁻ |
|-----------------------------|---|-----------------------|--------------------|-------------------------------------|--------------------------|----------------------|--------------------------------|--------------------------------|
| Διμερή | [Ala ₂ Leu ₂] ₈ με native (Rop-like) τοπολογία | 10ns | 78.5 x 60.1 x 56.8 | 25375 | 1790 | 23577 | 5 | 3 |
| | [Ala ₂ Leu ₂] ₈ με <i>syn</i> τοπολογία | 10ns | 79.0 x 58.9 x 58.6 | 25839 | 1801 | 24030 | 6 | 2 |
| Μονομερή | <i>Native</i> μονομερές | 20ns | 78.0 x 52.2 x 49.7 | 19032 | 895 | 18132 | 3 | 2 |
| | <i>Artificial</i> μονομερές | 8ns | 77.0 x 52.2 x 47.5 | 17922 | 895 | 17022 | 3 | 2 |
| | <i>Opened</i> μονομερές | 6 x 2ns | 78.1 x 60.2 x 49.1 | 21749 | 895 | 20847 | 4 | 3 |
| | | 1 x 10ns | | | | | | |
| <i>Opened P31</i> μονομερές | 10ns | 76.5 x 59.0 x 48.0 | 21735 | 899 | 20829 | 4 | 3 | |

Πίνακας 3.1 Συγκεντρωτικός πίνακας των προσομοιώσεων που πραγματοποιήθηκαν.

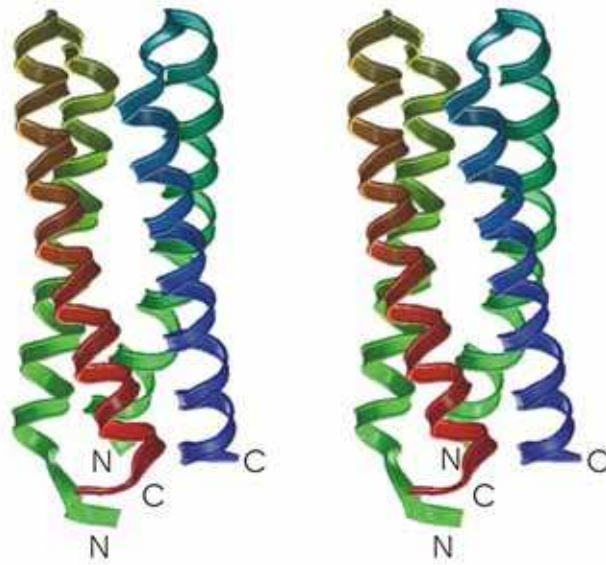
3.1 Υποθέσεις που έχουν διατυπωθεί μέχρι σήμερα

Στην προσπάθειά μας να μελετήσουμε το μηχανισμό αναδίπλωσης (folding) του τεχνητού 4-α-ελικοειδούς δεματίου [Ala₂-Leu₂]₈, που συνιστά μία από τις πολλές παραλλαγές (variants) της Rop που έχουν περιγραφεί λεπτομερώς μέχρι σήμερα, εστιαστήκαμε αρχικά στο ομοδιμερές που συνιστά τη λειτουργική μονάδα της πρωτεΐνης. Για τη συγκεκριμένη μεταλλαγμένη πρωτεΐνη δεν υπάρχει ακόμα γνωστή δομή. Κατά συνέπεια οι μεταλλάξεις έγιναν υπολογιστικά, έχοντας σαν αρχική δομή αυτήν της φυσικής πρωτεΐνης (Banner *et al.*, 1987) (Εικόνα 3.1).



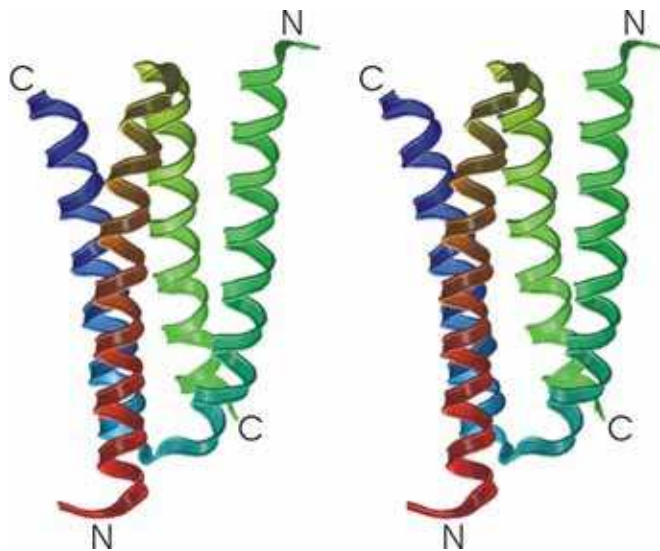
Εικόνα 3.1 Stereo άποψη της native (Rop-like) τοπολογίας.

Οι Levy *et al.* το 2004 παρουσίασαν στο περιοδικό PNAS μία θεωρητική πρόγνωση σχετικά με την τοπολογία που υιοθετούν οι διάφορες παραλλαγές της Rop. Η υπόθεση τους αυτή βασίζεται στο γεγονός ότι στις διάφορες παραλλαγές της Rop που έχουν περιγραφεί μέχρι σήμερα, έχουν βρεθεί, εκτός από την τοπολογία της φυσικής πρωτεΐνης, άλλες δύο τοπολογίες. Η μία αφορά τη μεταλλαγμένη πρωτεΐνη [Ala₂-Ile₂]₆, η οποία υιοθετεί μία *syn* τοπολογία, δηλαδή το ένα μονομερές έχει περιστραφεί 180° σε σχέση με το άλλο μονομερές (Willis *et al.*, 2000) (Εικόνα 3.2). Η άλλη τοπολογία έχει βρεθεί στην A31P, μία πρωτεΐνη που έχει προκύψει μετά από σημειακή μετάλλαξη της Ala 31 σε Pro, η οποία οδήγησε σε μετατροπή του κανονικού αριστερόστροφου δεματίου, όπου όλες οι έλικες είναι αντιπαράλληλες, σε ένα δεξιόστροφο μικτό δεμάτιο με παράλληλες και αντιπαράλληλες έλικες (Glykos *et al.*, 1999) (Εικόνα 3.3).



Εικόνα 3.2 Stereo άποψη της *syn* τοπολογίας.

Τα περισσότερα μεταλλάγματα της Rop θεωρήθηκε, βάσει της *in vitro* δραστηρότητας τους, ότι υιοθετούν την τοπολογία της native Rop (Munson *et al.*, 1997). Ωστόσο, λόγω έλλειψης της *in vivo* δραστηρότητας και επιπλέον εξαιτίας της αύξησης μέχρι και τέσσερις τάξεις μεγέθους, του χρόνου αναδίπλωσης και αποδιάταξης (folding/unfolding) και της έλλειψης της διαφασικής κινητικής κάποιων από αυτά, αμφισβητήθηκε η παραπάνω θεώρηση.



Εικόνα 3.3 Stereo άποψη της A31P τοπολογίας.

Η υπόθεση, λοιπόν, των Levy *et al.*, υποστηρίζει ότι τα διάφορα μεταλλάγματα της Rop απαντούν σε μίγμα των δύο τοπολογιών [της native (Rop-like) τοπολογίας και της *syn*]. Η τοπολογία A31P θεωρήθηκε ως παγιδευμένο μόριο (molten globule) ενώ, επιπλέον, το πακετάρισμα των υδρόφοβων ομάδων δεν είναι τόσο ισχυρό ώστε να

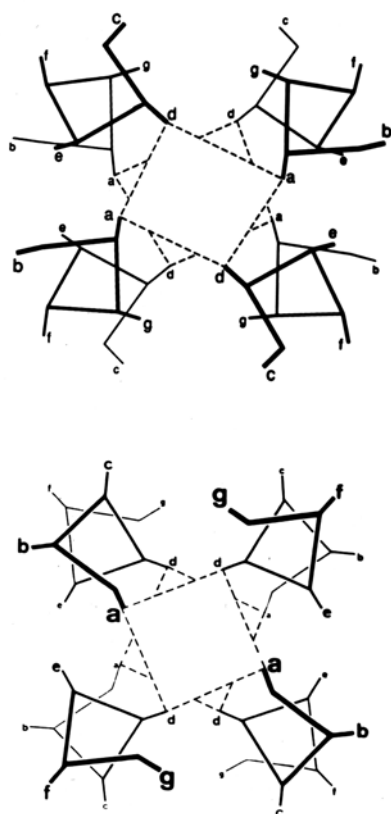
θεωρηθεί σαν μία ακόμη ανταγωνιστική τοπολογία (Glykos *et al.*, 2004). Μάλιστα, η υπόθεση ήταν ότι τα μεταλλάγματα εκείνα που επιδεικνύουν γρήγορη αναδίπλωση μάλλον θα δείχνουν κάποια προτίμηση προς την *syn* τοπολογία. Ο λόγος για αυτήν την προτίμηση αποδίδεται στο χαμηλότερο ενεργειακό φράγμα που θα πρέπει να ξεπεράσει ένα μόριο για να αποκτήσει την *syn* τοπολογία. Αυτό ισχύει γιατί στην περίπτωση της native (Rop-like) τοπολογίας τα τελικά κατάλοιπα του ενός μονομερούς αλληλεπιδρούν με τα κατάλοιπα της στροφής του άλλου μονομερούς.

Προκειμένου να εξετασθεί η παραπάνω υπόθεση, σχεδιάσαμε δύο ομοδιμερή με τη μεταλλαγμένη αλληλουχία, ένα με τη native (Rop-like) τοπολογία και ένα με την *syn* τοπολογία. Τα κύρια ερωτήματα τα οποία είχαμε σαν στόχο να διερευνήσουμε ήταν κατά πόσο οι δομές αυτές είναι σταθερές, αν είναι πιο σταθερές από την φυσική πρωτεΐνη, ή αν κάποια από τις δύο τοπολογίες επιδεικνύει ενδεχομένως, μεγαλύτερη σταθερότητα. Ακόμα, θα θέλαμε να εξετάσουμε αν οι πρωτεΐνες αυτές διαθέτουν ιδιαίτερα ευκίνητα τμήματα ή αν κάνουν κάποια χαρακτηριστική κίνηση μέσα στο διάλυμα. Καθώς το μονομερές συνιστά τη δομική μονάδα, μας ενδιαφέρει να εξετάσουμε πως συμπεριφέρεται το μονομερές στο δεμάτιο. Σύγκριση των αποτελεσμάτων από τις δύο τοπολογίες θα μας επιτρέψει να εξετάσουμε την ορθότητα της υπόθεσης ότι, αυτής της τάξης οι μεταλλαγμένες πρωτεΐνες, απαντούν σε μίγμα των δύο τοπολογιών και ακόμα ότι τα μόρια που αναδιπλώνονται γρήγορα τείνουν να υιοθετήσουν την *syn* τοπολογία.

Το επόμενο βήμα στη μελέτη μας ήταν να μελετήσουμε τα μονομερή στο διάλυμα. Σε πρώτη προσέγγιση θέλαμε να εξετάσουμε εάν τα μονομερή είναι σε θέση να διατηρήσουν τη δομή τους στο διάλυμα ή διαλύονται, δεδομένου ότι οι υδρόφοβες πλευρικές ομάδες που συγκροτούν τον πυρήνα του διμερούς θα είναι πλέον εκτεθειμένες προς τον διαλύτη, ή ακόμα εάν το μονομερές μπορεί να σχηματίσει ένα coiled coil από μόνο του. Μετά από μελέτη μεμονωμένων μονομερών είναι εφικτή η πραγματοποίηση, υπολογιστικά, συγκρούσεων μεταξύ μονομερών με σκοπό να παρατηρηθεί ένα πραγματικό γεγονός αναδίπλωσης, δεδομένου ότι η αναδίπλωση αλλά και η αποδιάταξη της συγκεκριμένης πρωτεΐνης πραγματοποιείται στην κλίμακα των msec.

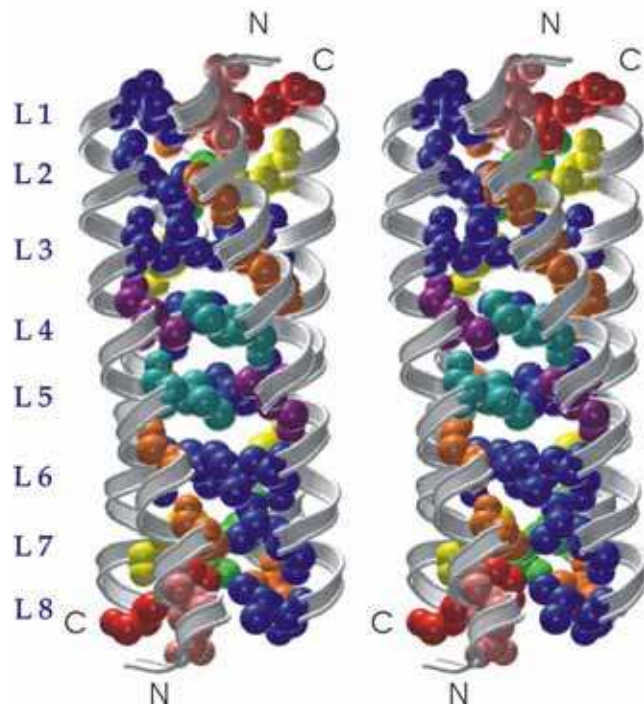
3.2 Προετοιμασία συστήματος ομοδιμερούς $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία

Η αρχική δομή που χρησιμοποιήθηκε ήταν, όπως προαναφέρθηκε, της φυσικής Rop (PDB entry 1rop στην PDB) σε διακριτικότητα 1.7 Å (Banner *et al.*, 1987). Η Rop είναι ένα ομοδιμερές που οργανώνεται σε ένα 4-α-ελικοειδές δεμάτιο (4- α -helix bundle). Η κάθε μονομερής υπομονάδα συνίσταται από 63 κατάλοιπα και σχηματίζει μία δομή τύπου έλικα-θηλιά-έλικα (helix-loop-helix).



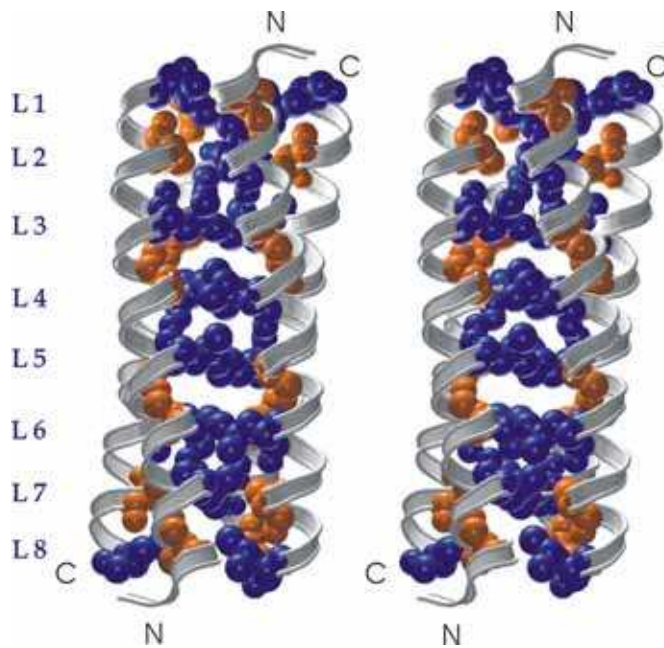
Εικόνα 3.4 Σχηματική απεικόνιση δύο επιπέδων (layers), όπου υποδεικνύονται οι θέσεις των καταλοίπων των ελίκων, σύμφωνα με την ονοματολογία του μοτίβου της επανάληψης επτάδας. Με διακεκομμένες γραμμές απεικονίζονται τα δύο επίπεδα που το καθένα αποτελείται από 2 κατάλοιπα τύπου a και 2 τύπου d (αναπαράγεται άνευ αδείας από τους Paliakasis *et al.*, 1992).

Ο εσωτερικός υδρόφοβος πυρήνας της Rop αποτελείται από 8 επίπεδα (layers) κάθετα προς τον επιμήκη άξονα του δεματίου. Τα επίπεδα αυτά τα σχηματίζουν πλευρικές αλυσίδες υδρόφοβων καταλοίπων και κάθε επίπεδο αποτελείται από τέσσερις τέτοιες πλευρικές ομάδες. Σύμφωνα με την ονοματολογία που εΐθισται να χρησιμοποιείται για συσχετιζόμενες έλικες, τα κατάλοιπα που συμμετέχουν στον υδρόφοβο πυρήνα μπορούν να προσδιοριστούν μέσω ενός μοτίβου επανάληψης επτάδας (heptad repeat nomenclature) (Munson *et al.*, 1994) (**Εικόνα 3.4**). Σύμφωνα λοιπόν με την παραπάνω ονοματολογία, ένα επίπεδο σχηματίζεται από 2 κατάλοιπα τύπου a και 2 κατάλοιπα τύπου d. Για την φυσική πρωτεΐνη τα κατάλοιπα αυτά υποδεικνύονται στην **Εικόνα 3.5**. Οι θέσεις a και d καταλαμβάνονται από υδρόφοβα κατάλοιπα τα οποία πακετάρονται μεταξύ τους ώστε να σχηματιστεί ο υδρόφοβος πυρήνας.



Εικόνα 3.5 Stereo άποψη της φυσικής Rop. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals (0.8 της πραγματικής ακτίνας). Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα και της Leu με μπλε. Το Glu με ροζ, η Phe με κόκκινο, η Cys με κίτρινο, η Gln με πράσινο, η Ile με γαλάζιο και η Thr με μωβ. Αριστερά αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Εκ πρώτης όψης φαίνεται πως τα κατάλοιπα αλανίνης και λευκίνης έχουν προτιμηθεί εξελικτικά σε μεγαλύτερο βαθμό σε σχέση με άλλα κατάλοιπα. Φαίνεται πως τα κατάλοιπα λευκίνης συμβάλλουν γενικότερα στη σταθεροποίηση ελίκων με κατάλοιπα αλανίνης (Luo *et al.*, 2002). Επιπλέον, έχει δειχθεί μία σαφής κυριαρχία καταλοίπων αλανίνης και λευκίνης σε εσωτερικές θέσεις (a και d) στα 4-α-ελικοειδή δεμάτια. Πιο συγκεκριμένα, η αλανίνη συναντάται στο 35% των θέσεων και η λευκίνη στο 13%, ενώ τα υπόλοιπα υδρόφοβα κατάλοιπα συναντώνται σε ποσοστό λιγότερο από 10% (Fadoulouglou *et al.*, 2001). Μέχρι σήμερα έχουν περιγραφεί πολυάριθμα μεταλλάγματα της Rop [(Ceruso *et al.*, 1999), (Munson *et al.*, 1994), (Munson *et al.*, 1996), (Munson *et al.*, 1997), (Predki *et al.*, 1995), (Steif *et al.*, 1995), (Vlassi *et al.*, 1999), (Willis *et al.*, 2000)]. Τα περισσότερα αφορούν μεταλλάξεις στον υδρόφοβο πυρήνα, δηλαδή αλλαγή των καταλοίπων που καταλαμβάνουν τις θέσεις a και d. Έτσι στην [Ala₂-Leu₂]₈, οι μεταλλάξεις είναι σχεδιασμένες κατά τρόπο ώστε η Ala να καταλαμβάνει τις a θέσεις και η Leu τις d θέσεις (**Εικόνα 3.6**). Οι μεταλλάξεις έγιναν υπολογιστικά με τη βοήθεια του προγράμματος xfit (McRee, 1992). Στην περίπτωση της λευκίνης η επιλογή του εκάστοτε στροφομερούς (rotamer) έγινε κατά τρόπο τέτοιο ώστε να αποφευχθούν οι στερικές παρεμποδίσεις με γειτονικά κατάλοιπα. Εάν δύο (ή περισσότερα) στροφομερή ήταν συμβατά, η επιλογή μεταξύ αυτών ήταν αυθαίρετη.



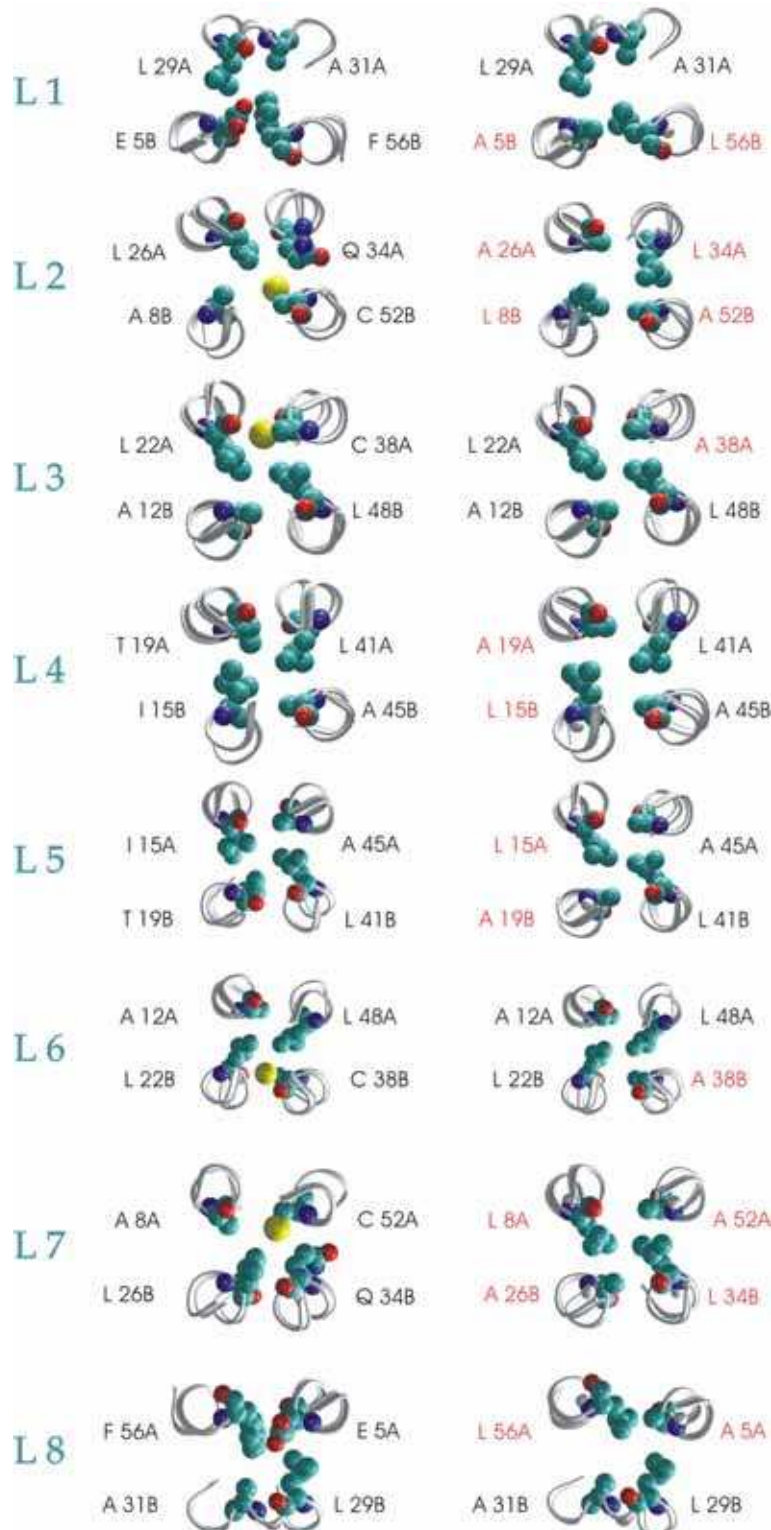
Εικόνα 3.6 Stereo άποψη της $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals (0.8 της πραγματικής ακτίνας). Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα και της Leu με μπλε. Αριστερά αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Στην **Εικόνα 3.7** γίνεται συγκριτική αντιπαράθεση των καταλοίπων που συνιστούν τα 8 επίπεδα στη φυσική πρωτεΐνη και στην $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία και υποδεικνύονται οι μεταλλάξεις που πραγματοποιήθηκαν.

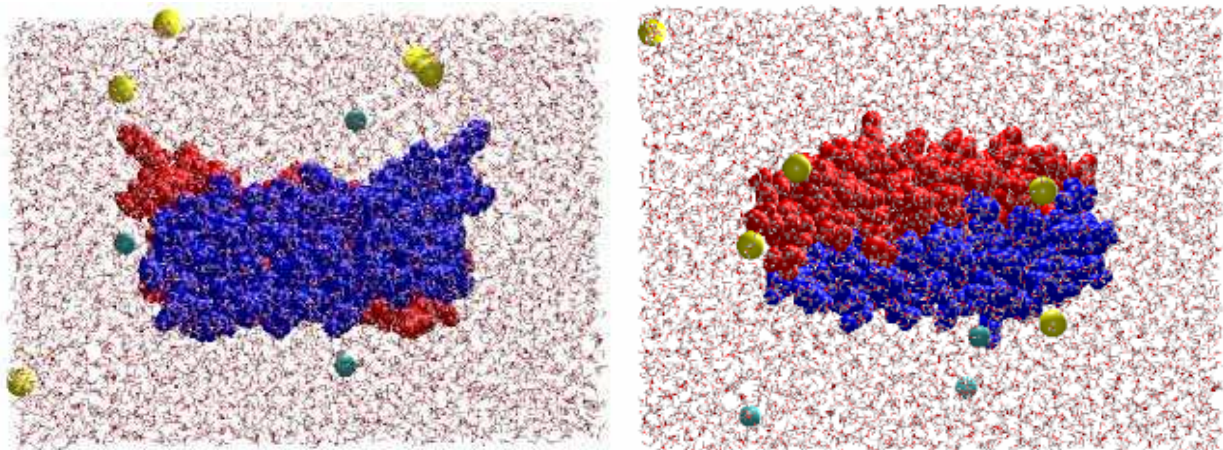
Από τη στιγμή που έχουμε τις συντεταγμένες των ατόμων της δομής της πρωτεΐνης, το επόμενο βήμα, όπως αναφέρθηκε στο *Κεφάλαιο 2*, είναι η ενυδάτωσή της με νερά και ιόντα. Έτσι, η πρωτεΐνη ενυδατώνεται (solvate) σε ένα ορθογώνιο κουτί διαστάσεων $78.5 \times 60.1 \times 56.8 \text{ \AA}^3$ με εκ των προτέρων εξισορροπημένα μόρια νερού (**Εικόνα 3.8**). Οι διαστάσεις του κουτιού είναι τέτοιες ώστε να υπάρχει ένα στρώμα νερού 15 \AA προς κάθε κατεύθυνση σε σχέση με τα πρωτεϊνικά μόρια. Όλα τα νερά που βρίσκονταν σε απόσταση 1.8 \AA από την πρωτεΐνη αφαιρέθηκαν. Το τελικό σύστημα αποτελείται από συνολικά 25.375 άτομα, από τα οποία τα 1790 είναι άτομα της πρωτεΐνης, 23.577 άτομα νερού, 5 ιόντα νατρίου και 3 ιόντα χλωρίου. Η ισοδύναμη συγκέντρωση χλωριούχου νατρίου (NaCl) ενός διαλύματος με τις ίδιες αναλογίες μορίων νερού και ιόντων είναι 50mM. Το σύστημα αυτό που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα ετοιμάζεται με τη βοήθεια του VMD (Humphrey *et al.* 1996).

Φυσική πρωτεΐνη

[Ala₂-Leu₂]₈



Εικόνα 3.7 Συγκριτική απεικόνιση των 8 επιπέδων στη φυσική πρωτεΐνη και στην [Ala₂-Leu₂]₈ αντίστοιχα. Τμήμα του σκελετού του μορίου φαίνεται με λευκή κορδέλα (ribbons) ενώ τα κατάλοιπα που συνιστούν το κάθε επίπεδο αναπαρίστανται με βάση την ακτίνα Van der Waals (0.7 της πραγματικής ακτίνας). Κάθε κατάλοιπο ονομάζεται με βάση τον κανόνα του ενός γράμματος, τη θέση του στην αλυσίδα και το μονομερές στο οποίο ανήκει. Τα κατάλοιπα τα οποία έχουν αλλαχθεί υποδεικνύονται με κόκκινο χρώμα. *Αριστερά* αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.



Εικόνα 3.8 *Αριστερά* Σχηματική αναπαράσταση του προς προσομοίωση συστήματος που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα. Διαφαίνεται η κυρτότητα του μορίου. Η πρωτεΐνη και τα ιόντα αναπαρίστανται με βάση την ακτίνα Van der Waals, ενώ τα νερά απεικονίζονται με wireframe . Με κόκκινο και μπλε χρώμα απεικονίζονται οι δύο υπομονάδες του ομοδιμερούς, με κίτρινο τα ιόντα νατρίου και με γαλάζιο τα ιόντα χλωρίου.
Δεξιά Το ίδιο σύστημα μετά από περιστροφή 90° γύρω από έναν άξονα παράλληλο στον επιμήκη άξονα του μορίου.

Μετά την προετοιμασία του συστήματος ακολουθεί η ελαχιστοποίηση της ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Η ελαχιστοποίηση ενέργειας, αρχικά εφαρμόζεται στα άτομα που δεν ανήκουν στο σκελετό του πρωτεϊνικού μορίου για 2.000 βήματα και στη συνέχεια σε όλα τα άτομα για ακόμα 2.000 βήματα, ώστε η πρωτεΐνη να φτάσει σε ένα σημείο όπου η γεωμετρία και τα μήκη δεσμών να είναι σωστά και να μη δημιουργείται ενεργειακό πρόβλημα. Ακολουθεί μία φάση κατά την οποία αυξάνεται σταδιακά η θερμοκρασία ανά 20° για 2.000 βήματα (heating). Στη συνέχεια γίνεται εξισορρόπηση του συστήματος πρώτα κρατώντας σταθερά τα Ca άτομα για 10.000 βήματα και μετά χωρίς κανένα περιορισμό για 100.000 βήματα. Ένα βήμα αντιστοιχεί σε 2fs (1fs = 10^{-15} s). Η διαδικασία αυτή γίνεται μέχρι η θερμοκρασία να φτάσει τους 320° K. Η επιλογή της θερμοκρασίας αυτής έγινε γιατί είναι αρκετά υψηλή (σε σχέση για παράδειγμα με τους 280° K) ώστε να επιταχυνθούν οι διαδικασίες που θέλουμε να παρακολουθήσουμε αλλά όχι τόσο ώστε να οδηγήσει σε αποδιάταξη (όπως στους 360° K). Ακολουθεί η παραγωγική φάση (production phase) για 5.000.000 βήματα, δηλαδή για 10ns με τις ατομικές συντεταγμένες να σώζονται κάθε 200 βήματα, δηλαδή κάθε 0.4ps. Ενδεικτικά αναφέρονται στο *Παράρτημα* τα script που χρησιμοποιήθηκαν για το NAMD, όπου αναφέρονται και λεπτομερώς όλες οι παράμετροι και οι τιμές που τους δόθηκαν.

3.3 Προετοιμασία συστήματος ομοδιμερούς $[Ala_2Leu_2]_8$ με *syn* τοπολογία

Για τη συγκεκριμένη μετάλλαξη της Rop, την $[Ala_2-Ile_2]_6$, που υιοθετεί την *syn* τοπολογία έχουν κατατεθεί δύο δομές σε διαφορετική διακριτικότητα. Η μία δομή (PDB entry 1F4N στην PDB) έχει διακριτικότητα 1.90 Å και έχουν κατατεθεί οι ατομικές συντεταγμένες για τα κατάλοιπα 1-60 για το ένα μονομερές (A) και οι συντεταγμένες για τα κατάλοιπα 5-55 για το άλλο μονομερές (B). Η άλλη δομή (PDB entry 1F4M στην PDB) έχει διακριτικότητα 2.25 Å, υπάρχουν 3 δεμάτια (bundles) στην ασύμμετρη μονάδα του κρυστάλλου και για το καθένα έχουν προσδιορισθεί ατομικές συντεταγμένες για κάθε μονομερές, όπως αναφέρεται στον **Πίνακα 3.2**.

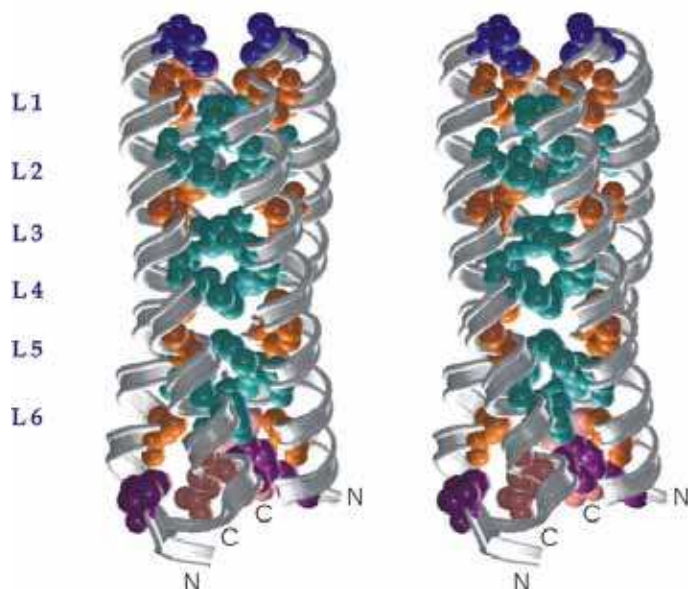
| | 1F4N | 1F4M | | | $[Ala_2Ile_2]_6$ |
|---|------|------|------|------|------------------|
| | | 1 | 2 | 3 | |
| A | 1-60 | 1-56 | 2-55 | 5-57 | 1-60 |
| B | 5-55 | 2-57 | 1-57 | 4-57 | 2-57 |

Πίνακα 3.2 Αριθμός καταλοίπων για τα οποία έχουν προσδιορισθεί κρυσταλλογραφικά οι ατομικές συντεταγμένες για τις δύο δομές διαφορετικής διακριτικότητας. Η στήλη που σημειώνεται ως $[Ala_2Ile_2]_6$, αντιστοιχεί στο μόριο το οποίο προέκυψε μετά από την υπέρθεση και την προσθήκη των καταλοίπων στο μονομερές B.

Από τη δομή χαμηλής διακριτικότητας επιλέχθηκε το δεμάτιο με το μεγαλύτερο αριθμό προσδιορισμένων καταλοίπων (δηλαδή το 1) και χρησιμοποιήθηκε για υπέρθεση πάνω στη δομή της υψηλής διακριτικότητας (1F4N). Η υπέρθεση έγινε κατά τρόπο τέτοιο ώστε να ελαχιστοποιηθεί το rmsd (root mean square deviation). Το αποτέλεσμα ήταν να προκύψει (υπολογιστικά) μία δομή, η $[Ala_2Ile_2]_6$, στην οποία το μονομερές A είναι αυτό της δομής υψηλής διακριτικότητας και συνίσταται από 60 κατάλοιπα, ενώ στο μονομερές B έχουν προστεθεί στη δομή της υψηλής διακριτικότητας τα κατάλοιπα 2-10 και 50 -57 από το δεμάτιο 1 της δομής χαμηλής διακριτικότητας [τα επιπλέον κατάλοιπα, 5-10 και 50-55, προστέθηκαν για λόγους σταθερότητας (για περισσότερες λεπτομέρειες βλ. Παπαϊωάννου, Χ.Ι. (2006). Μελέτες μοριακής δυναμικής μίας μορφής της Rop με επανασχεδιασμένο υδρόφοβο πυρήνα. Διπλωματική Εργασία Δημοκρίτειου Πανεπιστημίου Θράκης, Τμήματος Μοριακής Βιολογίας & Γενετικής)].

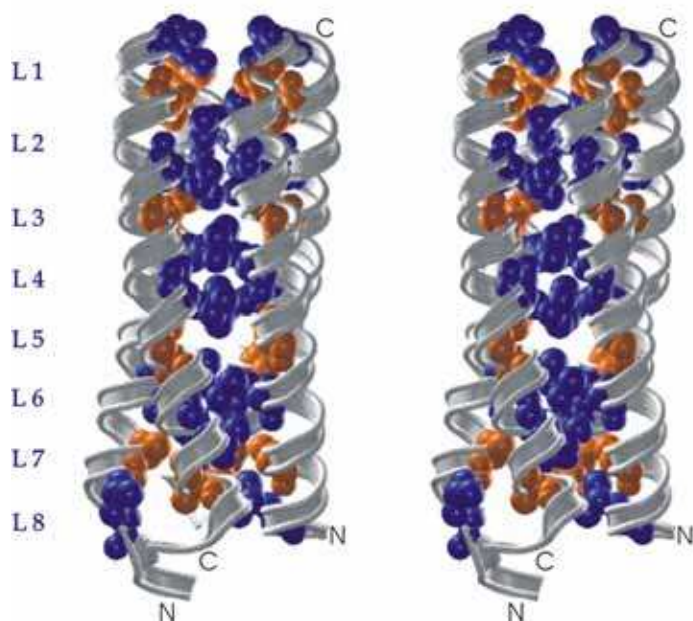
Ο υδρόφοβος πυρήνας αυτής της μεταλλαγμένης πρωτεΐνης συνίσταται από κατάλοιπα αλανίνης στις a θέσεις και κατάλοιπα ισολευκίνης στις d θέσεις με εξαίρεση τα επίπεδα 1 και 8 στα οποία τα κατάλοιπα των θέσεων αυτών δεν έχουν

μεταλλαχθεί. Στην **Εικόνα 3.9** βλέπουμε μία άποψη της πρωτεΐνης με τα κατάλοιπα που σχηματίζουν τον υδρόφοβο πυρήνα.



Εικόνα 3.9 Stereo άποψη της $[Ala_2Ile_2]_6$. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals (0.8 της πραγματικής ακτίνας). Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα, της Leu με μπλε, η Ile με γαλάζιο, το Glu με ροζ και η Phe με μωβ. *Αριστερά* αναφέρονται ενδεικτικά με L1 έως L6 τα έξι επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Ετοιμάσαμε λοιπόν υπολογιστικά την $[Ala_2Leu_2]_8$, χρησιμοποιώντας σαν αρχική δομή αυτήν της $[Ala_2Ile_2]_6$ που περιγράφηκε παραπάνω, δηλαδή προέκυψε ένα μόριο με την τοπολογία *syn* και κατάλοιπα αλανίνης και λευκίνης στις θέσεις a και d, αντίστοιχα, και στα οκτώ επίπεδα. Μία άποψη της πρωτεΐνης αυτής και των καταλοίπων που συνιστούν τα επίπεδα που συγκροτούν τον υδρόφοβο πυρήνα φαίνεται στην **Εικόνα 3.10**. Και σε αυτήν την περίπτωση οι μεταλλάξεις έγιναν υπολογιστικά με τη βοήθεια του προγράμματος xfit ενώ η επιλογή των στροφομερών (rotamers) της λευκίνης έγινε κατά τρόπο τέτοιο ώστε να αποφευχθούν οι στερικές παρεμποδίσεις με γειτονικά κατάλοιπα. Θα πρέπει εδώ να σημειωθεί πως στην κρυσταλλογραφικά προσδιορισμένη δομή της $[Ala_2Ile_2]_6$, σε όλα τα κατάλοιπα ισολευκίνης που καταλαμβάνουν τις θέσεις d μέσα στον υδρόφοβο πυρήνα το rotamer είναι το ίδιο, τόσο στη δομή της υψηλής διακριτικότητας, όσο και στη δομή χαμηλής διακριτικότητας (και στα τρία δεμάτια). Στην $[Ala_2Leu_2]_8$ όμως, τόσο η *syn* τοπολογία όσο και η native (Rop-like) τοπολογία, που περιγράφηκε στην Ενότητα 2.1, ετοιμάστηκαν όπως προαναφέρθηκε υπολογιστικά και για κάθε κατάλοιπο λευκίνης χρησιμοποιήθηκε διαφορετικό rotamer. Για το λόγο αυτό τα rotamers (καθώς και το κατά πόσο παραμένουν τα ίδια ή αλλάζουν) θα μελετηθούν ενδελεχώς στο Κεφάλαιο 4 (Ενότητα 4.1.10 και 4.2.10).



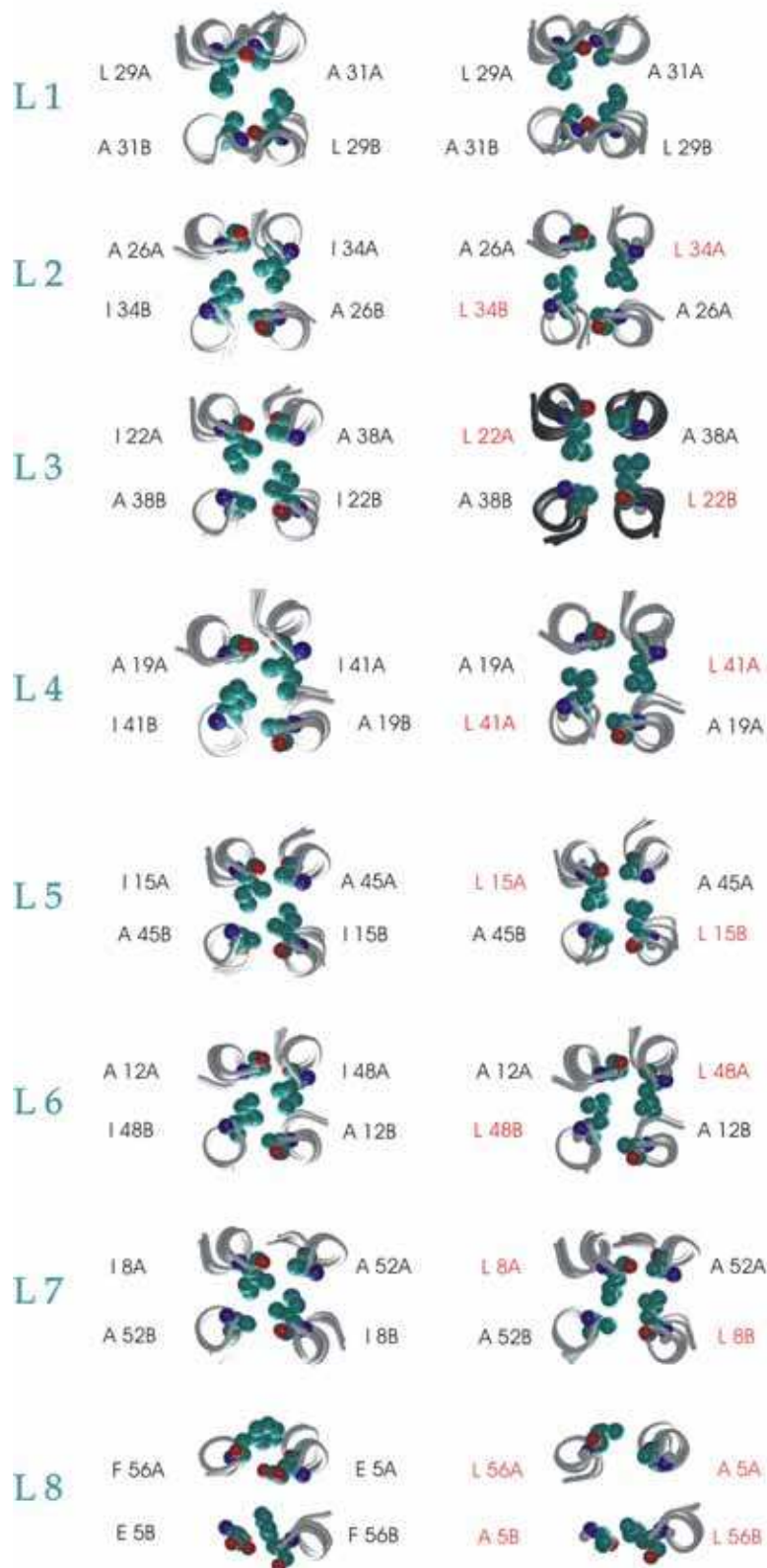
Εικόνα 3.10 Stereo άποψη της $[Ala_2Leu_2]_8$ με τοπολογία *syn*. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals (0.8 της πραγματικής ακτίνας). Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα και της Leu με μπλε. Αριστερά αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Στην **Εικόνα 3.11** γίνεται συγκριτική αντιπαράθεση των καταλοίπων που συνιστούν τα 8 επίπεδα στην $[Ala_2Ile_2]_6$ και στην $[Ala_2Leu_2]_8$ με τοπολογία *syn* και υποδεικνύονται οι μεταλλάξεις που πραγματοποιήθηκαν.

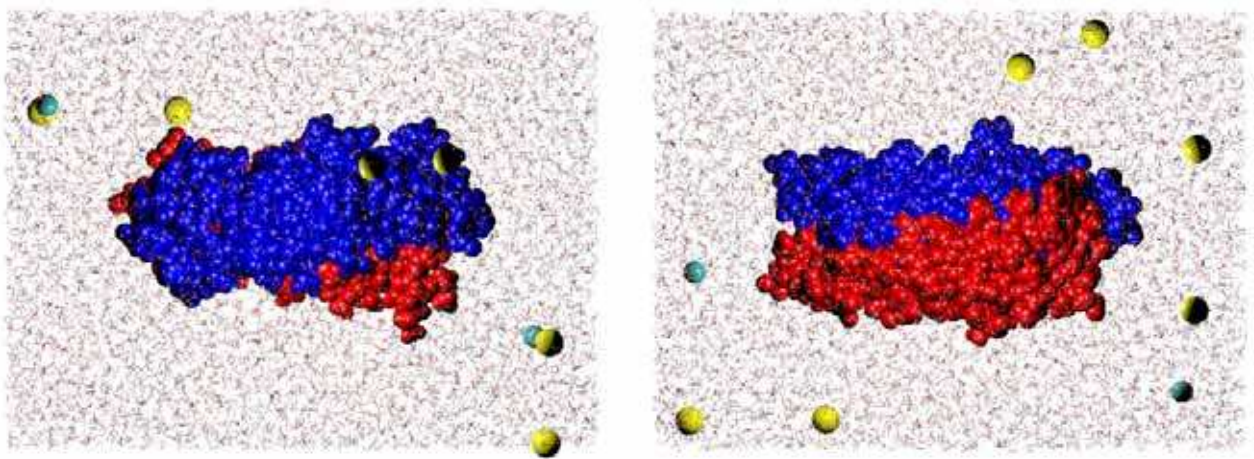
Ακολούθως η πρωτεΐνη ενυδατώνεται σε ένα ορθογώνιο κουτί διαστάσεων $79.0 \times 58.9 \times 58.6 \text{ \AA}^3$ με εκ των προτέρων εξισορροπημένα μόρια νερού (**Εικόνα 3.12**). Οι διαστάσεις του κουτιού είναι τέτοιες ώστε να υπάρχει ένα στρώμα νερού 15 \AA προς κάθε κατεύθυνση σε σχέση με τα πρωτεϊνικά μόρια. Όλα τα νερά που βρίσκονταν σε απόσταση 1.8 \AA από την πρωτεΐνη αφαιρέθηκαν. Το τελικό σύστημα αποτελείται από συνολικά 25.839 άτομα, από τα οποία τα 1801 είναι άτομα της πρωτεΐνης, 24.030 άτομα νερού, 6 ιόντα νατρίου και 2 ιόντα χλωρίου. Η ισοδύναμη συγκέντρωση χλωριούχου νατρίου (NaCl) ενός διαλύματος με τις ίδιες αναλογίες μορίων νερού και ιόντων είναι 50mM. Το σύστημα αυτό που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα ετοιμάζεται με τη βοήθεια του VMD.

Φυσική πρωτεΐνη

[Ala₂-Leu₂]₈



Εικόνα 3.11 Συγκριτική απεικόνιση των 8 επιπέδων στην [Ala₂Ile₂]₆ και στην [Ala₂-Leu₂]₈ με τοπολογία *syn*, αντίστοιχα. Τμήμα του σκελετού του μορίου φαίνεται με λευκή κορδέλα (ribbons) ενώ τα κατάλοιπα που συνιστούν το κάθε επίπεδο αναπαρίστανται με βάση την ακτίνα Van der Waals (0.7 της πραγματικής ακτίνας). Κάθε κατάλοιπο ονομάζεται με βάση τον κανόνα του ενός γράμματος, τη θέση του στην αλυσίδα και το μονομερές στο οποίο ανήκει. Τα κατάλοιπα τα οποία έχουν αλλαχθεί υποδεικνύονται με κόκκινο χρώμα. *Αριστερά* αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

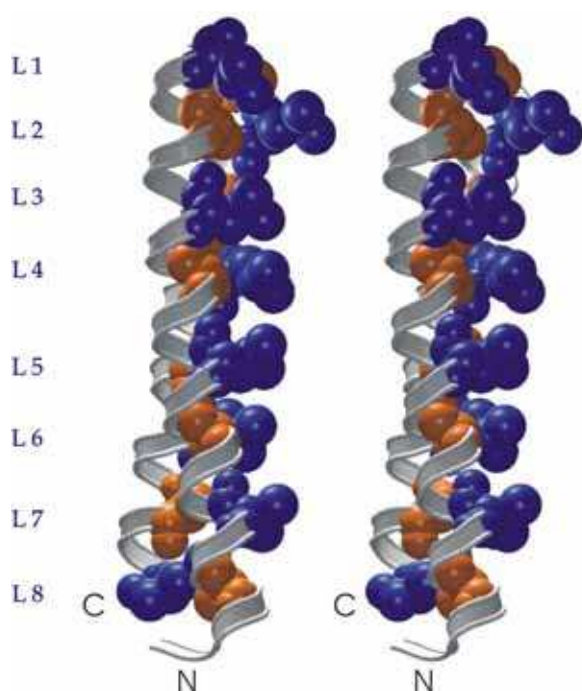


Εικόνα 3.12 Αριστερά Σχηματική αναπαράσταση του προς προσομοίωση συστήματος που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα. Η πρωτεΐνη και τα ιόντα αναπαρίστανται με βάση την ακτίνα Van der Waals, ενώ τα νερά απεικονίζονται με wireframe. Με κόκκινο και μπλε χρώμα απεικονίζονται οι δύο υπομονάδες του ομοδιμερούς, με κίτρινο τα ιόντα νατρίου και με γαλάζιο τα ιόντα χλωρίου.
Δεξιά Το ίδιο σύστημα μετά από περιστροφή 90° γύρω από έναν άξονα παράλληλο στον επιμήκη άξονα του μορίου.

Μετά την προετοιμασία του συστήματος ακολουθεί η ελαχιστοποίηση της ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Η φάση της ελαχιστοποίησης της ενέργειας εφαρμόζεται με τον ίδιο τρόπο που περιγράφηκε στην Ενότητα 3.2 για το σύστημα του ομοδιμερούς $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία και η φάση του heating εφαρμόζεται επίσης για συνολικά 146.000 βήματα, δηλαδή 292.000fs (0.292ns). Η παραγωγική φάση (production phase) εφαρμόζεται για 5.000.000 βήματα, δηλαδή για 10ns με τις ατομικές συντεταγμένες να σώζονται κάθε 200 βήματα, δηλαδή κάθε 0.4ps. Ενδεικτικά αναφέρονται στο *Παράρτημα* τα script που χρησιμοποιήθηκαν για το Namd, όπου αναφέρονται και λεπτομερώς όλες οι παράμετροι και οι τιμές που τους δόθηκαν.

3.4 Προετοιμασία συστήματος μονομερούς $[Ala_2Leu_2]_8$ με native (Rop-like) αναδίπλωση

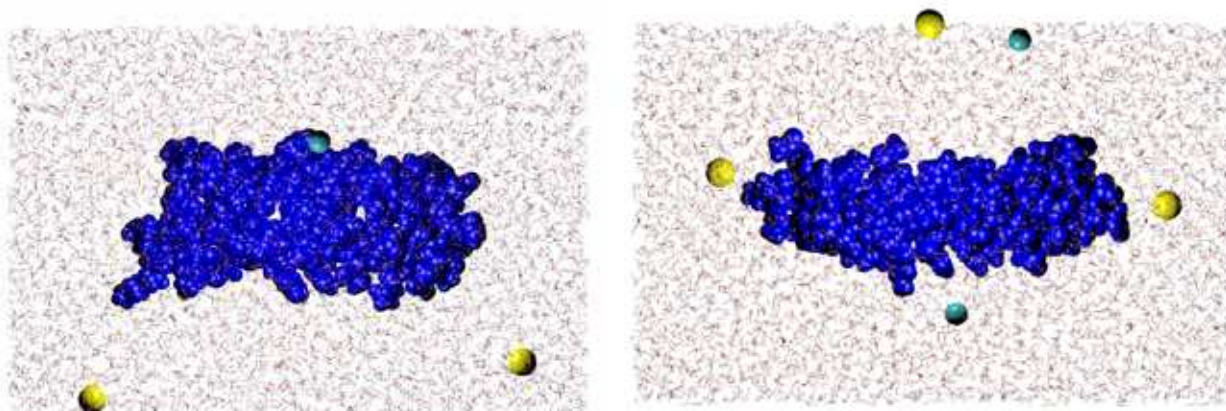
Όπως αναφέραμε στην αρχή του Κεφαλαίου, μετά από τη μελέτη του ομοδιμερούς που συνιστά τη λειτουργική μονάδα, το ενδιαφέρον στράφηκε στη μονομερή υπομονάδα. Αρχικά θέλαμε να εξετάσουμε τη συμπεριφορά του μονομερούς στη native μορφή του. Όταν αναφερόμαστε στον όρο native μονομερές, εννοούμε το μονομερές στο 4- α -ελικοειδές δεμάτιο, στη μορφή ακριβώς που το συναντάμε, δηλαδή με τις πλευρικές ομάδες των καταλοίπων στα οκτώ επίπεδα, που συνιστούν τον υδρόφοβο πυρήνα, στραμμένες αυτή τη φορά προς το διαλύτη (αφού στο δεμάτιο είναι στραμμένες προς τις αντίστοιχες του άλλου μονομερούς του δεματίου που βρίσκεται ακριβώς απέναντι) (Εικόνα 3.13). Η δομή προέκυψε υπολογιστικά διατηρώντας μόνο τις ατομικές συντεταγμένες των καταλοίπων (1-56) που ανήκουν στην αλυσίδα A (δηλαδή του ενός μονομερούς).



Εικόνα 3.13 Stereo άποψη του native μονομερούς. Οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τον υδρόφοβο πυρήνα είναι στραμμένες προς τα έξω. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals. Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα και της Leu με μπλε. Αριστερά αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Ακολούθως το native μονομερές ενυδατώνεται (solvate) σε ένα ορθογώνιο κουτί διαστάσεων $78.0 \times 52.2 \times 49.7 \text{ \AA}^3$ με εκ των προτέρων εξισορροπημένα μόρια νερού (Εικόνα 3.14). Οι διαστάσεις του κουτιού είναι τέτοιες ώστε να υπάρχει ένα στρώμα νερού 15 \AA προς κάθε κατεύθυνση σε σχέση με τα πρωτεϊνικά μόρια. Όλα τα νερά που βρίσκονταν σε απόσταση 1.8 \AA από την πρωτεΐνη αφαιρέθηκαν. Το τελικό σύστημα αποτελείται από συνολικά 19.032 άτομα, από τα οποία τα 895 είναι άτομα της πρωτεΐνης, 18.132 άτομα νερού, 3 ιόντα νατρίου και 2 ιόντα χλωρίου. Η

ισοδύναμη συγκέντρωση χλωριούχου νατρίου (NaCl) ενός διαλύματος με τις ίδιες αναλογίες μορίων νερού και ιόντων είναι 50mM. Το σύστημα αυτό που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα ετοιμάζεται με τη βοήθεια του VMD.



Εικόνα 3.14 *Αριστερά* Σχηματική αναπαράσταση του προς προσομοίωση συστήματος που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα. Η πρωτεΐνη και τα ιόντα αναπαρίστανται με βάση την ακτίνα Van der Waals, ενώ τα νερά απεικονίζονται με wireframe. Με μπλε χρώμα απεικονίζεται το μονομερές που χρησιμοποιήθηκε, με κίτρινο τα ιόντα νατρίου και με γαλάζιο τα ιόντα χλωρίου. *Δεξιά* Το ίδιο σύστημα μετά από περιστροφή 90° γύρω από έναν άξονα παράλληλο στον επιμήκη άξονα του μορίου.

Μετά την προετοιμασία του συστήματος ακολουθεί η ελαχιστοποίηση της ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Η φάση της ελαχιστοποίησης της ενέργειας εφαρμόζεται με τον ίδιο τρόπο που περιγράφηκε στην Ενότητα 3.2 για το σύστημα του ομοδιμερούς $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία και η φάση του heating εφαρμόζεται επίσης για συνολικά 146.000 βήματα, δηλαδή 292.000fs (0.292ns). Η παραγωγική φάση (production phase) εφαρμόζεται για 10.000.000 βήματα, δηλαδή για 20ns με τις ατομικές συντεταγμένες να σώζονται κάθε 200 βήματα, δηλαδή κάθε 0.4ps. Ενδεικτικά αναφέρονται στο *Παράρτημα* τα script που χρησιμοποιήθηκαν για το Namd, όπου αναφέρονται και λεπτομερώς όλες οι παράμετροι και οι τιμές που τους δόθηκαν.

3.5 Προετοιμασία συστήματος *artificial* μονομερούς $[Ala_2Leu_2]_8$ με *native* (Rop-like) αναδίπλωση

Δεδομένου ότι το μονομερές σχηματίζει μία δομή τύπου έλικα-στροφή-έλικα, προκύπτει λογικά το ερώτημα κατά πόσο το μόριο αυτό μπορεί να σχηματίσει ένα ελικοειδές σπείραμα (coiled-coil) αν βρεθεί μόνο του στο διάλυμα, χωρίς κάποιο άλλο μονομερές για να αλληλεπιδράσει. Για το σκοπό αυτό ανατρέξαμε στις βάσεις δεδομένων, PDB (Protein Data Bank) (Berman *et al.*, 2000), SCOP (Structural Classification Of Proteins) και CATH (Class, Architecture, Topology, Homologous superfamily) για να βρούμε μία φυσική πρωτεΐνη που να σχηματίζει ένα ελικοειδές σπείραμα αντιπαράλληλων ελίκων. Έχοντας στα χέρια μας ένα μόριο που υιοθετεί αυτή τη δομή μπορούμε να εξετάσουμε τη διαμόρφωση τόσο των ελίκων όσο και των πλευρικών ομάδων των καταλοίπων που συνιστούν τον υδρόφοβο πυρήνα. Έτσι θα έχουμε ένα φυσικό παράδειγμα για το πως θα μπορούσε να είναι η διαμόρφωση του μονομερούς προκειμένου να σχηματίσει μία δομή ελικοειδούς σπειράματος που να είναι ενεργειακά αποδεκτή.



Εικόνα 3.15 Stereo άποψη της κυτταροπλασματικής επικράτειας της ProP. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals (0.8 της πραγματικής ακτίνας). Τα κατάλοιπα Leu απεικονίζονται με μπλε χρώμα και της Pe με γαλάζιο. Η Arg με ανοιχτό πράσινο, η Gly με ανοιχτό μπλε και η His με καφέ. Αριστερά αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Το καλύτερο αποτέλεσμα από αυτήν την έρευνα των βάσεων δεδομένων είναι η πρωτεΐνη ProP. Πρόκειται για ένα συµμεταφορέα πρωτονίων του βακτηρίου *Escherichia coli*, ο οποίος αποκρίνεται σε οσµωτικές µεταβολές. Το καρβοξυτελικό

της άκρο (κατάλοιπα 468-497) θεωρείται ότι είναι σημαντικό για τη δραστηριότητά της και συνιστά την κυτταροπλασματική επικράτεια (cytoplasmic domain) (**Εικόνα 3.15**). Η δομή της (PDB entry 1R48) έχει προσδιορισθεί μέσω φασματοσκοπίας NMR (Nuclear Magnetic Resonance, πυρηνικός μαγνητικός συντονισμός) και όπως αναφέρθηκε, πρόκειται για ένα ομοδιμερές αντιπαράλληλο α-ελικοειδές σπείραμα (Zoetewey *et al.*, 2003). Η αμινοξική αλληλουχία της πρωτεΐνης αυτής παρουσιάζεται στον **Πίνακα 3.3**.

| | | | | | | | | |
|---------|------|------|------|------|------|------|------|------|
| Helix A | C 1 | G 2 | G 3 | D 4 | N 5 | I 6 | E 7 | Q 8 |
| | K 9 | I 10 | D 11 | D 12 | I 13 | D 14 | H 15 | E 16 |
| | I 17 | A 18 | D 19 | L 20 | Q 21 | A 22 | K 23 | R 24 |
| | T 25 | R 26 | L 27 | V 28 | Q 29 | Q 30 | H 31 | P 32 |
| | R 33 | | | | | | | |
| Helix B | C 1 | G 2 | G 3 | D 4 | N 5 | I 6 | E 7 | Q 8 |
| | K 9 | I 10 | D 11 | D 12 | I 13 | D 14 | H 15 | E 16 |
| | I 17 | A 18 | D 19 | L 20 | Q 21 | A 22 | K 23 | R 24 |
| | T 25 | R 26 | L 27 | V 28 | Q 29 | Q 30 | H 31 | P 32 |
| | R 33 | | | | | | | |

Πίνακας 3.3 Αλληλουχία των αμινοξέων της κυτταροπλασματικής επικράτειας της ProP, που σχηματίζει ένα ομοδιμερές ελικοειδές σπείραμα. Τα αμινοξέα ονομάζονται με βάση τον κανόνα του ενός γράμματος. Για λόγους ευκολίας, τα κατάλοιπα που καταλαμβάνουν τις a και d θέσεις υποδεικνύονται με πορτοκαλί και μπλε χρώμα, αντίστοιχα. Τα κατάλοιπα της δεύτερης έλικας παρατίθενται για λόγους συσχέτισης με τους πίνακες που ακολουθούν.

Ακολουθώς, συγκρίναμε τις διαμορφώσεις των καταλοίπων που καταλαμβάνουν τις a και d θέσεις, στο native ελικοειδές σπείραμα και το native μονομερές της $[Ala_2Leu_2]_8$. Ετοιμάσαμε ένα εκτελέσιμο αρχείο (script) για το πρόγραμμα VMD, το οποίο κάνει υπέρθεση των δύο δομών και υπολογίζει το rmsd. Η υπέρθεση γίνεται στοιχίζοντας μόνο τα άτομα του σκελετού των μορίων (backbone) και διαφορετικό κάθε φορά αριθμό από κατάλοιπα στις a και d θέσεις, με σκοπό να βρούμε τη καλύτερη δυνατή στοίχιση (με τη μικρότερη τιμή rmsd). Σε κάθε περίπτωση όμως, στοιχίζουμε τα κατάλοιπα στις a θέσεις της μίας δομής με τα αντίστοιχα κατάλοιπα των a θέσεων της άλλης δομής και τα κατάλοιπα στις d θέσεις της μίας δομής με τα αντίστοιχα κατάλοιπα των d θέσεων της άλλης δομής. Επειδή η κάθε δομή διαθέτει δύο έλικες έπρεπε να γίνει υπέρθεση της κάθε έλικας και προς τις άλλες τρεις. Ωστόσο, στο native μονομερές οι δύο έλικες συνδέονται μέσω μίας περιοχής θηλιάς. Κατά συνέπεια τα κατάλοιπα της στροφής αφαιρέθηκαν (υπολογιστικά) ώστε να προκύψουν δύο απομονωμένες έλικες (κατάλοιπα 1-27 και

33-56). Να σημειωθεί ότι με τον τρόπο αυτό χάνεται το όγδοο επίπεδο (το οποίο σχημάτιζαν τα κατάλοιπα της στροφής). Στον **Πίνακα 3.4** παρουσιάζεται, για λόγους σύγκρισης, η αμινοξική αλληλουχία του native μονομερούς μετά από τις αλλαγές που προαναφέρθηκαν.

| | | | | | | | | |
|---------|------|------|------|------|------|------|------|------|
| Helix A | M 1 | T 2 | K 3 | Q 4 | A 5 | K 6 | T 7 | L 8 |
| | L 9 | N 10 | M 11 | A 12 | R 13 | F 14 | L 15 | R 16 |
| | S 17 | Q 18 | A 19 | L 20 | T 21 | L 22 | L 23 | E 24 |
| | K 25 | A 26 | N 27 | | | | | |
| Helix B | E 33 | L 34 | A 35 | D 36 | I 37 | A 38 | E 39 | S 40 |
| | L 41 | H 42 | D 43 | H 44 | A 45 | D 46 | E 47 | L 48 |
| | Y 49 | R 50 | S 51 | A 52 | L 53 | A 54 | R 55 | L 56 |

Πίνακας 3.4 Αλληλουχία των αμινοξέων του native μονομερούς μετά την αφαίρεση των καταλοίπων της στροφής. Τα αμινοξέα ονομάζονται με βάση τον κανόνα του ενός γράμματος. Και σε αυτήν την περίπτωση τα κατάλοιπα που καταλαμβάνουν τις a και d θέσεις υποδεικνύονται με πορτοκαλί και μπλε χρώμα, αντίστοιχα.

Η υπέρθεση έγινε στοιχίζοντας αρχικά επτά επίπεδα και ακολούθως, έξι, πέντε και τέσσερα. Η έρευνα δεν επεκτάθηκε σε λιγότερα επίπεδα, καθώς οι τιμές των rmsd που θα προέκυπταν θα ήταν σχετικά μικρές αλλά δεν θα περιείχαν αξιοποιήσιμη πληροφορία. Τα αποτελέσματα παρατίθενται στους **Πίνακες 3.5, 3.6, 3.7** και **3.8** που ακολουθούν.

| 7 | | Native μονομερές της [Ala ₂ Leu ₂] ₈ | | | |
|----------------------------------|----------------|--|--------|-----------------|-------|
| native ελικοειδής σπείραμα | | A helix (1-27) | | B helix (33-56) | |
| | | | 5-26 | | 34-56 |
| | A helix (1-33) | 3-24 | 1.5774 | 6-27 | - * |
| | | 10-31 | 1.3030 | | |
| | B helix (1-33) | 3-24 | 1.5887 | 6-27 | - * |
| | 10-31 | 1.2904 | | | |

Πίνακας 3.5 Αποτελέσματα στοίχισης επτά επιπέδων. Ο χρωματικός κώδικας (πορτοκαλί για τα κατάλοιπα στις a θέσεις και μπλε για τα κατάλοιπα στις d θέσεις) διατηρείται.

| 6 | | Native μονομερές της [Ala ₂ Leu ₂] ₈ | | | | | | |
|----------------------------------|----------------------|--|--------|------|-----------------|-------|--------|-------|
| native ελικοειδές σπείραμα | | A helix (1-27) | | | B helix (33-56) | | | |
| | A helix (1-33) | | | 5-22 | 8-26 | | 34-52 | 38-56 |
| | | 3-20 | 1.3743 | | | 6-24 | 0.9278 | |
| | | 10-27 | 1.0876 | | | 13-31 | 1.0684 | |
| | | 6-24 | | | 1.2110 | 3-20 | | -* |
| | | 13-31 | | | 1.1036 | 10-27 | | -* |
| | B helix (1-33) | 3-20 | 1.3889 | | | 6-24 | 0.9322 | |
| | | 10-27 | 1.0687 | | | 13-31 | 1.0571 | |
| | | 6-24 | | | 1.2222 | 3-20 | | -* |
| | | 13-31 | | | 1.0892 | 10-27 | | -* |

Πίνακας 3.6 Αποτελέσματα στοίχισης έξι επιπέδων. Ο χρωματικός κώδικας (πορτοκαλί για τα κατάλοιπα στις a θέσεις και μπλε για τα κατάλοιπα στις d θέσεις) διατηρείται.

| 5 | | Native μονομερές της [Ala ₂ Leu ₂] ₈ | | | | | | | | |
|----------------------------------|----------------------|--|-------|-------|-----------------|-------|-------|-------|-------|-------|
| native ελικοειδές σπείραμα | | A helix (1-27) | | | B helix (33-56) | | | | | |
| | A helix (1-33) | | | 5-19 | 8-22 | 12-26 | | 34-48 | 38-52 | 41-56 |
| | | 3-17 | 1.306 | | | 1.261 | 6-20 | 0.759 | | -* |
| | | 10-24 | 1.043 | | | 1.053 | 13-27 | 0.966 | | -* |
| | | 17-31 | 1.043 | | | 1.037 | 3-17 | | 1.245 | |
| | | 6-20 | | | 0.999 | | 10-24 | | 0.956 | |
| | 13-27 | | | 0.872 | | 17-31 | | 1.013 | | |
| | B helix (1-33) | 3-17 | 1.318 | | | 1.264 | 6-20 | 0.769 | | -* |
| | | 10-24 | 1.042 | | | 1.053 | 13-27 | 0.946 | | -* |
| | | 17-31 | 1.051 | | | 1.033 | 3-17 | | 1.252 | |
| 6-20 | | | | 1.016 | | 10-24 | | 0.955 | | |
| 13-27 | | | | 0.845 | | 17-31 | | 1.011 | | |

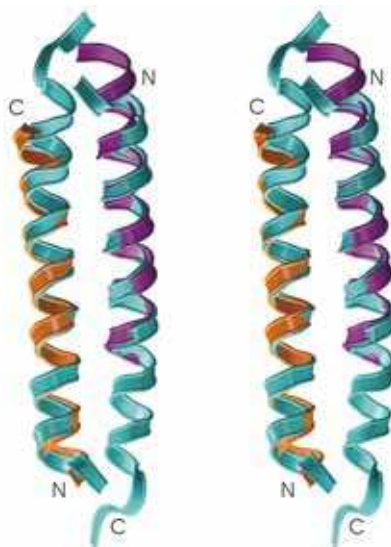
Πίνακας 3.7 Αποτελέσματα στοίχισης πέντε επιπέδων. Ο χρωματικός κώδικας (πορτοκαλί για τα κατάλοιπα στις a θέσεις και μπλε για τα κατάλοιπα στις d θέσεις) διατηρείται. Με ροζ χρώμα υποδεικνύονται οι τιμές rmsd με βάση τις οποίες έγινε η επιλογή των ατόμων που θα στοιχισθούν κατά την υπέρθεση.

* Στην [Ala₂Leu₂]₈ (όπως και στην φυσική Rop), η Leu56 καταλαμβάνει τη d θέση αντί της Arg55. Έτσι η στοίχιση δεν μπορούσε να πραγματοποιηθεί επειδή ο αριθμός ατόμων είναι διαφορετικός.

| 4 | | Native μονομερές της [Ala ₂ Leu ₂] ₈ | | | | | | | | | |
|----------------------------------|-------------------|--|-------|-------|-------|-----------------|-------|-------|-------|-------|-------|
| native ελικοειδές σπείραμα | | A helix (1-27) | | | | B helix (33-56) | | | | | |
| | | | 5-15 | 8-19 | 12-22 | 15-26 | | 34-45 | 38-48 | 41-52 | 45-56 |
| | A helix (1-33) | 3-13 | 1.087 | | 1.162 | | 6-17 | 0.810 | | 0.905 | |
| | | 10-20 | 0.863 | | 0.865 | | 13-24 | 0.892 | | 0.778 | |
| | | 17-27 | 0.805 | | 0.866 | | 20-31 | 0.972 | | 0.823 | |
| | | 6-17 | | 0.964 | | 1.096 | 3-13 | | 1.127 | | -* |
| | | 13-24 | | 0.803 | | 1.098 | 10-20 | | 0.798 | | -* |
| | B helix (1-33) | 20-31 | | 0.867 | | 1.054 | 17-27 | | 0.886 | | -* |
| | | 3-13 | 1.088 | | 1.158 | | 6-17 | 0.820 | | 0.921 | |
| | | 10-20 | 0.869 | | 0.869 | | 13-24 | 0.891 | | 0.776 | |
| | 17-27 | 0.797 | | 0.846 | | 20-31 | 0.976 | | 0.823 | | |
| | 6-17 | | 0.981 | | 1.104 | 3-13 | | 1.122 | | -* | |
| | 13-24 | | 0.798 | | 1.099 | 10-20 | | 0.802 | | -* | |
| | 20-31 | | 0.876 | | 1.053 | 17-27 | | 0.871 | | -* | |

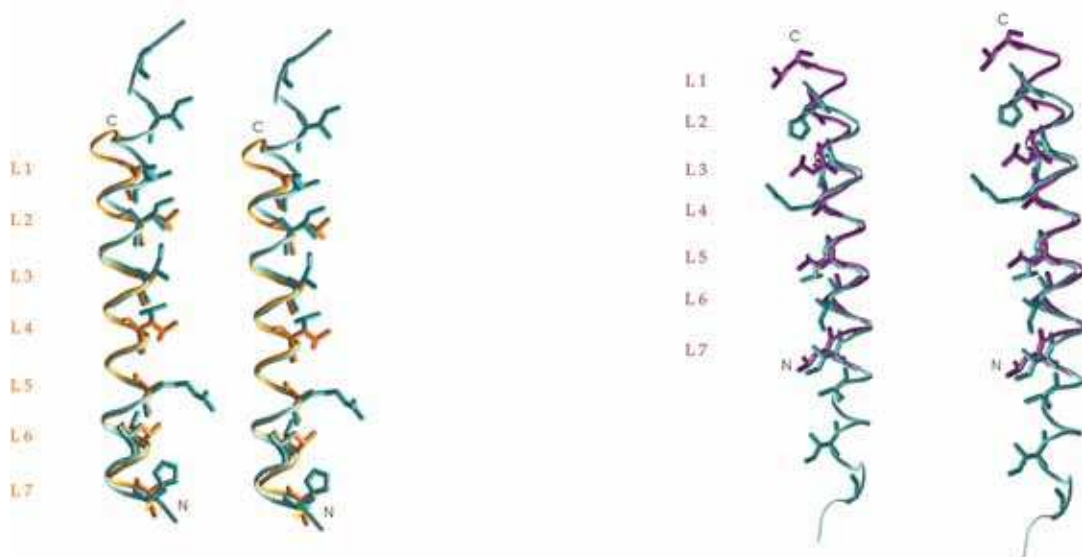
Πίνακας 3.8 Αποτελέσματα στοίχισης τεσσάρων επιπέδων. Ο χρωματικός κώδικας (πορτοκαλί για τα κατάλοιπα στις a θέσεις και μπλε για τα κατάλοιπα στις d θέσεις) διατηρείται.

Η προσπάθειά μας ήταν να συνδυάσουμε, κατά το δυνατόν, μεγαλύτερο αριθμό επιπέδων που στοιχίζονται (και άρα μεγαλύτερο τμήμα της πρωτεΐνης) με μικρές τιμές rmsd (τιμές μικρότερες της μονάδας θεωρούνται ικανοποιητικές). Έτσι επιλέξαμε για τη στοίχιση, τα κατάλοιπα 8-22 της A έλικας του native μονομερούς της $[Ala_2Leu_2]_8$ με τα κατάλοιπα 13-27 της A έλικας του native ελικοειδούς σπειράματος, με τιμή rmsd 0.8727 και τα κατάλοιπα 34-48 της B έλικας του native μονομερούς της $[Ala_2Leu_2]_8$ με τα κατάλοιπα 13-27 της B έλικας του native ελικοειδούς σπειράματος, με τιμή rmsd 0.9463 (**Εικόνα 3.16**).

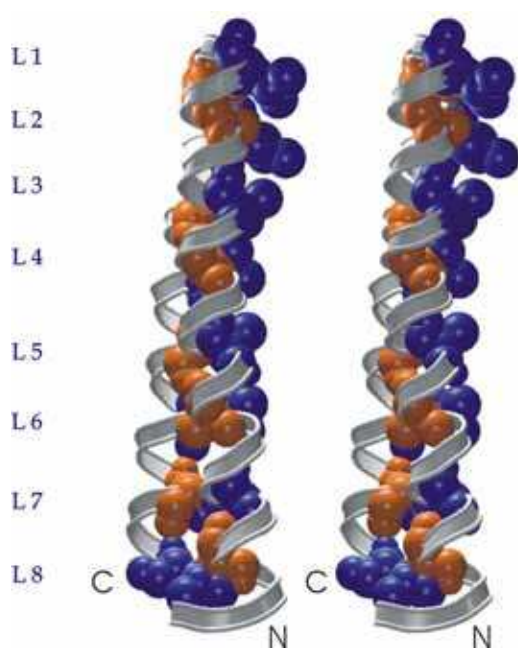


Εικόνα 3.16 Πάνω Stereo άποψη της υπέρθεσης των δύο δομών. Ο σκελετός (backbone) του native ελικοειδούς σπειράματος φαίνεται με γαλάζια κορδέλα (ribbons). Ο σκελετός της A έλικας του native μονομερούς της $[Ala_2Leu_2]_8$ φαίνεται με πορτοκαλί κορδέλα, ενώ της B έλικας με μωβ.

Κάτω Stereo άποψη της υπέρθεσης των δύο δομών αριστερά μόνο για τις A έλικες και δεξιά μόνο για τις B έλικες. Οι πλευρικές αλυσίδες που συνιστούν τα επίπεδα (αναφέρονται στο αριστερό τμήμα με L1 έως L7) απεικονίζονται με όλα τους τα άτομα.



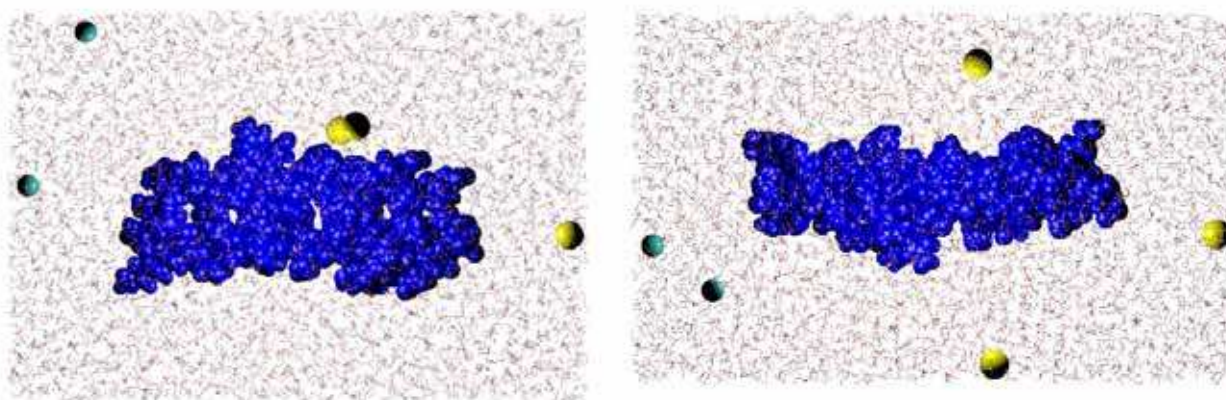
Το επόμενο βήμα ήταν να στρέψουμε υπολογιστικά τις πλευρικές ομάδες του native μονομερούς ώστε να μην είναι πλέον εκτεθειμένες στο διάλυμα. Προκειμένου να λάβει τη διαμόρφωση ενός native ελικοειδούς σπειράματος που περιγράφηκε παραπάνω, προσπαθήσαμε να περιστρέψουμε τις έλικες κατά τρόπο τέτοιο ώστε οι πλευρικές ομάδες να πλησιάσουν μεταξύ τους και στρέψαμε τα κατάλοιπα της στροφής ώστε να μεταβληθούν οι phi και psi γωνίες τους (**Εικόνα 3.17**). Στη συνέχεια έγινε μία γεωμετρική βελτίωση (geometrical refinement) για να διορθωθούν οι γωνίες και τα μήκη δεσμών. Η παραπάνω διαδικασία έγινε με τη βοήθεια του προγράμματος xfit.



Εικόνα 3.17 Stereo άποψη του artificial μονομερούς. Οι πλευρικές ομάδες των καταλοίπων της μίας έλικας που συγκροτούν τον υδρόφοβο πυρήνα έχουν στραφεί (υπολογιστικά) προς τις αντίστοιχες της άλλης έλικας. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (gibbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals. Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα και της Leu με μπλε. Αριστερά αναφέρονται ενδεικτικά με L1 έως L8 τα οκτώ επίπεδα που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Ακολούθως το artificial μονομερές ενυδατώνεται σε ένα ορθογώνιο κουτί διαστάσεων $77.0 \times 52.2 \times 47.5 \text{ \AA}^3$ με εκ των προτέρων εξισορροπημένα μόρια νερού (**Εικόνα 3.18**). Οι διαστάσεις του κουτιού είναι τέτοιες ώστε να υπάρχει ένα στρώμα νερού 15 \AA προς κάθε κατεύθυνση σε σχέση με τα πρωτεϊνικά μόρια. Όλα τα νερά που βρίσκονταν σε απόσταση 1.8 \AA από την πρωτεΐνη αφαιρέθηκαν. Το τελικό σύστημα αποτελείται από συνολικά 17.922 άτομα, από τα οποία τα 895 είναι άτομα της πρωτεΐνης, 17.022 άτομα νερού, 3 ιόντα νατρίου και 2 ιόντα χλωρίου. Η ισοδύναμη συγκέντρωση χλωριούχου νατρίου (NaCl) ενός διαλύματος με τις ίδιες αναλογίες μορίων νερού και ιόντων είναι 50mM. Το σύστημα αυτό που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα ετοιμάζεται με τη βοήθεια του VMD (Humphrey *et al.* 1996).

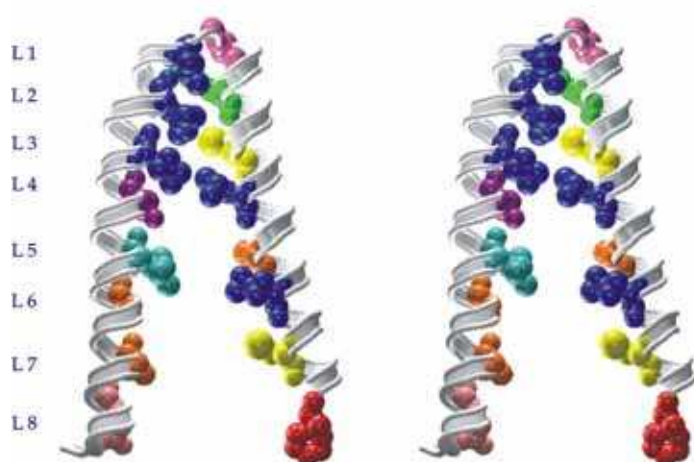
Μετά την προετοιμασία του συστήματος ακολουθεί η ελαχιστοποίηση της ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Η φάση της ελαχιστοποίησης της ενέργειας εφαρμόζεται με τον ίδιο τρόπο που περιγράφηκε στην Ενότητα 3.2 για το σύστημα του ομοδιμερούς $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία και η φάση του heating εφαρμόζεται επίσης για συνολικά 146.000 βήματα, δηλαδή 292.000fs (0.292ns). Η παραγωγική φάση (production phase) εφαρμόζεται για 4.000.000 βήματα, δηλαδή για 8ns με τις ατομικές συντεταγμένες να σώζονται κάθε 200 βήματα, δηλαδή κάθε 0.4ps. Ενδεικτικά αναφέρονται στο *Παράρτημα τα script* που χρησιμοποιήθηκαν για το Namd, όπου αναφέρονται και λεπτομερώς όλες οι παράμετροι και οι τιμές που τους δόθηκαν.



Εικόνα 3.18 *Αριστερά* Σχηματική αναπαράσταση του προς προσομοίωση συστήματος που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα. Η πρωτεΐνη και τα ιόντα αναπαρίστανται με βάση την ακτίνα Van der Waals, ενώ τα νερά απεικονίζονται με wireframe. Με μπλε χρώμα απεικονίζεται το μονομερές που χρησιμοποιήθηκε, με κίτρινο τα ιόντα νατρίου και με γαλάζιο τα ιόντα χλωρίου. *Δεξιά* Το ίδιο σύστημα μετά από περιστροφή 90° γύρω από έναν άξονα παράλληλο στον επιμήκη άξονα του μορίου.

3.6 Προετοιμασία συστήματος μονομερούς $[Ala_2Leu_2]_8$ με A31P αναδιπλώση

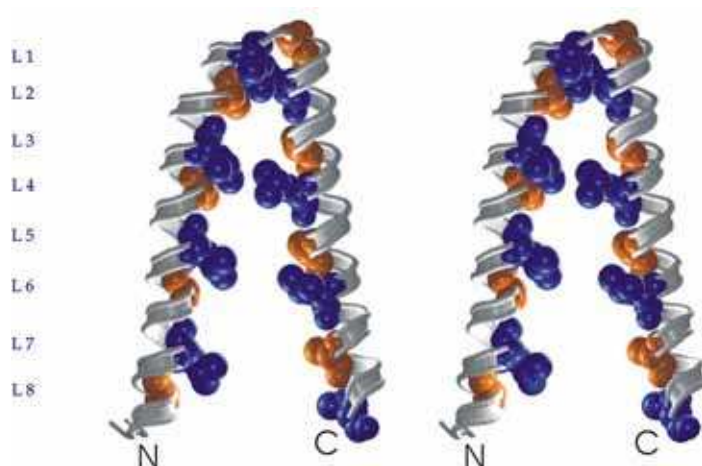
Στα πλαίσια της προσπάθειά μας να εξετάσουμε κατά πόσο το μονομερές μπορεί να σχηματίσει ένα ελικοειδές σπείραμα, στραφήκαμε στην τοπολογία της A31P (βλ. Εικόνα 2.3). Ο λόγος ήταν πως σε αυτήν την τοπολογία το μονομερές υιοθετεί μία αρκετά ανοιχτή διαμόρφωση (Εικόνα 3.19), δηλαδή οι έλικες είναι απομακρυσμένες και η διαμόρφωση της στροφής είναι διαφορετική (λόγω της προλίνης στη θέση 31). Έτσι υπάρχει διαθέσιμος χώρος ώστε οι πλευρικές ομάδες να κινηθούν και να στραφούν η μία προς την άλλη.



Εικόνα 3.19 Stereo άποψη του A31P μονομερούς. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals (0.8 της πραγματικής ακτίνας). Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα και της Leu με μπλε. Το Glu με ροζ, η Phe με κόκκινο, η Cys με κίτρινο, η Ile με γαλάζιο, η Thr με μωβ, η Gln με πράσινο και η Pro με ανοιχτό μωβ. *Αριστερά* αναφέρονται ενδεικτικά (για λόγους σύγκρισης) με L1 έως L8 τα οκτώ επίπεδα που θα μπορούσαν να συγκροτήσουν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Η αρχική δομή που χρησιμοποιήσαμε είναι της A31P (PDB entry 1B6Q) σε διακριτικότητα 1.8 Å. Σε πρώτη φάση απομονώσαμε τις ατομικές συντεταγμένες του μονομερούς A. Στη συνέχεια αλλάξαμε υπολογιστικά με τη βοήθεια του προγράμματος xfit τα κατάλοιπα στις θέσεις a και d έτσι ώστε να προκύψει ένα μόριο με την τοπολογία A31P και την αλληλουχία της $[Ala_2Leu_2]_8$. Στην **Εικόνα 3.20**

βλέπουμε μία άποψη του μονομερούς $[Ala_2Leu_2]_8$ με τοπολογία A31P. Για λόγους συντομίας το μονομερές αυτό θα αναφέρεται σαν *opened*.

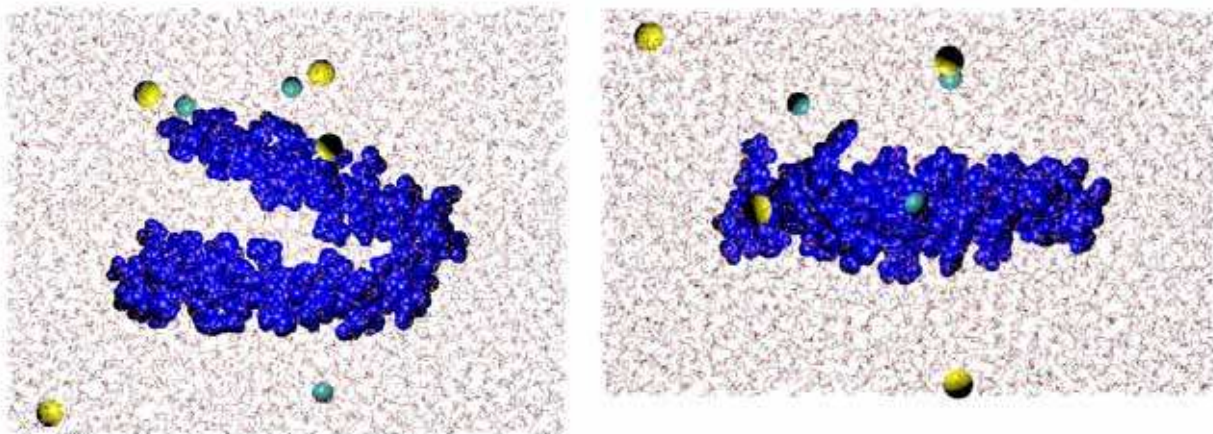


Εικόνα 3.20 Stereo άποψη του *opened* μονομερούς. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals. Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα και της Leu με μπλε. Αριστερά αναφέρονται ενδεικτικά (για λόγους σύγκρισης) με L1 έως L8 τα οκτώ επίπεδα που θα μπορούσαν να συγκροτήσουν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Ακολουθώς το *opened* μονομερές ενυδατώνεται σε ένα ορθογώνιο κουτί διαστάσεων $78.1 \times 60.2 \times 49.1 \text{ \AA}^3$ με εκ των προτέρων εξισορροπημένα μόρια νερού (Εικόνα 3.21). Οι διαστάσεις του κουτιού είναι τέτοιες ώστε να υπάρχει ένα στρώμα νερού 15 \AA προς κάθε κατεύθυνση σε σχέση με τα πρωτεϊνικά μόρια. Όλα τα νερά που βρίσκονταν σε απόσταση 1.8 \AA από την πρωτεΐνη αφαιρέθηκαν. Το τελικό σύστημα αποτελείται από συνολικά 21.749 άτομα, από τα οποία τα 895 είναι άτομα της πρωτεΐνης, 20.847 άτομα νερού, 4 ιόντα νατρίου και 3 ιόντα χλωρίου. Η ισοδύναμη συγκέντρωση χλωριούχου νατρίου (NaCl) ενός διαλύματος με τις ίδιες αναλογίες μορίων νερού και ιόντων είναι 50mM. Το σύστημα αυτό που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα ετοιμάζεται με τη βοήθεια του VMD.

Μετά την προετοιμασία του συστήματος ακολουθεί η ελαχιστοποίηση της ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Η φάση της ελαχιστοποίησης της ενέργειας εφαρμόζεται με τον ίδιο τρόπο που περιγράφηκε στην Ενότητα 3.2 για το σύστημα του ομοδιμερούς $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία και η φάση του heating εφαρμόζεται επίσης για συνολικά 146.000 βήματα, δηλαδή 292.000fs (0.292ns). Σε αυτήν την περίπτωση δεν πραγματοποιήσαμε μία προσομοίωση μεγάλης διάρκειας αλλά έξι μικρής διάρκειας. Σε πρώτη φάση έγιναν έξι παραγωγικές φάσεις (production phases) για 1.000.000 βήματα, δηλαδή για 2ns

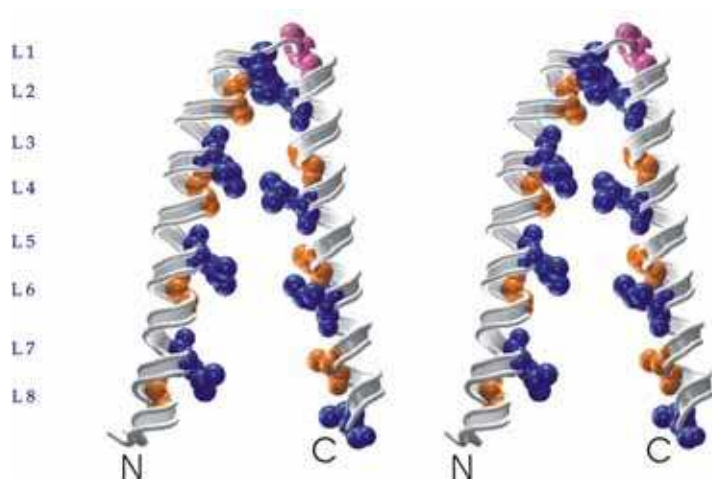
με τις ατομικές συντεταγμένες να σώζονται κάθε 200 βήματα, δηλαδή κάθε 0.4ps. Σε κάθε μία από τις παραγωγικές φάσεις, τα άτομα του συστήματος ακολουθούν διαφορετική τροχιά (trajectory) και με τον τρόπο αυτό μπορούμε να μελετήσουμε σε καλύτερο βαθμό τη συμπεριφορά των πλευρικών ομάδων και να δούμε αν το αποτέλεσμα που παρατηρούμε είναι επαναλήψιμο. Σε δεύτερη φάση πραγματοποιήσαμε μία προσομοίωση μεγαλύτερης διάρκειας για να μελετήσουμε την εν γένει συμπεριφορά του μονομερούς, δηλαδή αν οι πλευρικές ομάδες πλησιάζουν αρκετά ώστε να αλληλεπιδράσουν μεταξύ τους και να σχηματίσουν τον πυρήνα ενός ελικοειδούς σπειράματος. Επίσης, αν όντως αλληλεπιδράσουν, κατά πόσο παραμένουν σε αυτή τη διαμόρφωση και αν το αποτέλεσμα είναι ένα σταθεροποιημένο μόριο. Η διάρκεια της παραγωγικής φάσης (production phase) ήταν 5.000.000 βήματα, δηλαδή 10ns, με τις ατομικές συντεταγμένες να σώζονται κάθε 200 βήματα, δηλαδή κάθε 0.4ps. Ενδεικτικά αναφέρονται στο *Παράρτημα* τα script που χρησιμοποιήθηκαν για το NAMD, όπου αναφέρονται και λεπτομερώς όλες οι παράμετροι και οι τιμές που τους δόθηκαν.



Εικόνα 2.20 *Αριστερά* Σχηματική αναπαράσταση του προς προσομοίωση συστήματος που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα. Η πρωτεΐνη και τα ιόντα αναπαρίστανται με βάση την ακτίνα Van der Waals, ενώ τα νερά απεικονίζονται με wireframe. Με μπλε χρώμα απεικονίζεται το μονομερές που χρησιμοποιήθηκε, με κίτρινο τα ιόντα νατρίου και με γαλάζιο τα ιόντα χλωρίου. *Δεξιά* Το ίδιο σύστημα μετά από περιστροφή 90° γύρω από έναν άξονα παράλληλο στον επιμήκη άξονα του μορίου.

3.7 Προετοιμασία συστήματος μονομερούς [Ala₂Leu₂]₈ με A31P τοπολογία και Pro στη θέση 31

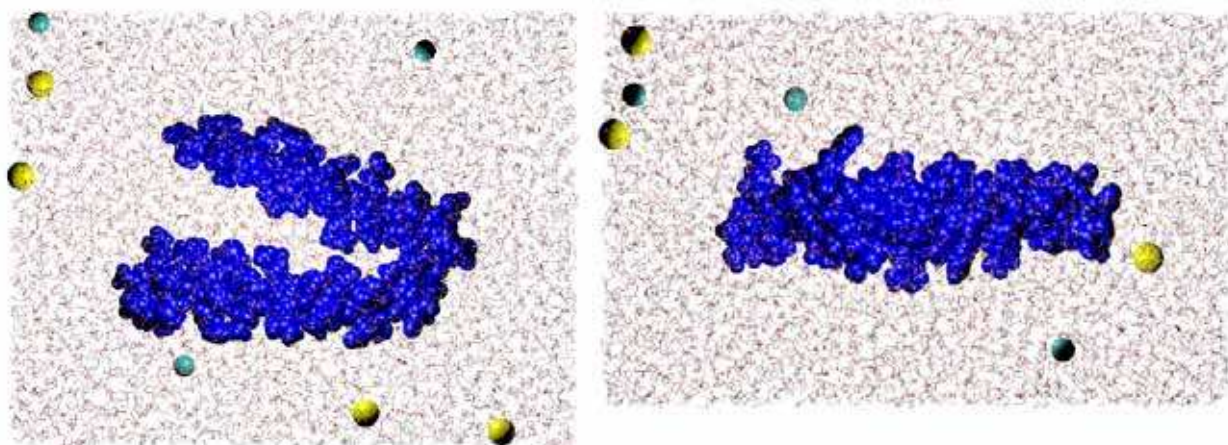
Όπως αναφέρθηκε στην αρχή του Κεφαλαίου, μία σημειακή μετάλλαξη στη θέση 31 της αλανίνης σε προλίνη οδηγεί σε δραματικές αλλαγές στην τοπολογία της πρωτεΐνης. Τόσο η διαμόρφωση των ελίκων όσο και η διαμόρφωση της στροφής είναι αρκετά διαφορετικές σε σχέση με τη φυσική πρωτεΐνη. Στην προηγούμενη ενότητα το μονομερές που ετοιμάσαμε είχε την τοπολογία της A31P, αλλά είχε αλανίνη στη θέση 31. Με τη διαδικασία που περιγράφηκε παραπάνω, ετοιμάσαμε ένα αντίστοιχο μονομερές, που αυτή τη φορά είχε προλίνη στη θέση 31, προκειμένου να φανεί η επίδραση της παρουσίας της προλίνης στη στροφή (Εικόνα 2.21). Για λόγους συντομίας το μονομερές αυτό θα αναφέρεται σαν *opened P31*. Το ερώτημα που τίθενται είναι εάν η παρουσία της προλίνης αλλάζει τη συμπεριφορά του μονομερούς κατά τη διάρκεια της προσομοίωσης. Ο λόγος για αυτήν την υπόθεση είναι πως έχει δειχθεί ότι μία υποθετική δομή, με δομή ίδια με αυτή της φυσικής πρωτεΐνης στην οποία έχει αλλαχθεί υπολογιστικά η αλανίνη στη θέση 31 σε προλίνη φαίνεται από τις προσομοιώσεις να είναι πιο σταθερή (Glykos *et al.*, 2004). Συνεπώς, θα πρέπει να διερευνηθεί εάν η παρουσία της προλίνης στη στροφή επιδρά στη διαμόρφωση του μονομερούς, αναγκάζοντας το μόριο να λάβει μία διαμόρφωση λιγότερο σταθερή σε σχέση με αυτήν της φυσικής πρωτεΐνης.



Εικόνα 2.21 Stereo άποψη του *opened P31* μονομερούς. Ο σκελετός του μορίου απεικονίζεται με λευκή κορδέλα (ribbons) και οι πλευρικές ομάδες των καταλοίπων που συγκροτούν τα 8 επίπεδα απεικονίζονται με βάση την ακτίνα Van der Waals. Τα κατάλοιπα Ala φαίνονται με πορτοκαλί χρώμα, της Leu με μπλε και η Pro με μωβ ανοιχτό. Αριστερά αναφέρονται ενδεικτικά (για λόγους σύγκρισης) με L1 έως L8 τα οκτώ επίπεδα που θα μπορούσαν να συγκροτήσουν τον υδρόφοβο πυρήνα της πρωτεΐνης.

Ακολουθώς το *opened P31* μονομερές ενυδατώνεται σε ένα ορθογώνιο κουτί διαστάσεων $76.5 \times 59.0 \times 48.0 \text{ \AA}^3$ με εκ των προτέρων εξισορροπημένα μόρια νερού (Εικόνα 2.22). Οι διαστάσεις του κουτιού είναι τέτοιες ώστε να υπάρχει ένα στρώμα νερού 15 \AA προς κάθε κατεύθυνση σε σχέση με τα πρωτεϊνικά μόρια. Όλα τα νερά που βρίσκονταν σε απόσταση 1.8 \AA από την πρωτεΐνη αφαιρέθηκαν. Το τελικό σύστημα αποτελείται από συνολικά 21.735 άτομα, από τα οποία τα 899 είναι άτομα της πρωτεΐνης, 20.829 άτομα νερού, 4 ιόντα νατρίου και 3 ιόντα χλωρίου. Η ισοδύναμη συγκέντρωση χλωριούχου νατρίου (NaCl) ενός διαλύματος με τις ίδιες αναλογίες μορίων νερού και ιόντων είναι 50mM. Το σύστημα αυτό που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα ετοιμάζεται με τη βοήθεια του VMD.

Μετά την προετοιμασία του συστήματος ακολουθεί η ελαχιστοποίηση της ενέργειας (energy minimization) και η εξισορρόπηση (equilibration). Η φάση της ελαχιστοποίησης της ενέργειας εφαρμόζεται με τον ίδιο τρόπο που περιγράφηκε στην Ενότητα 3.2 για το σύστημα του ομοδιμερούς $[\text{Ala}_2\text{Leu}_2]_8$ με native (Rop-like) τοπολογία και η φάση του heating εφαρμόζεται επίσης για συνολικά 146.000 βήματα, δηλαδή 292.000fs (0.292ns). Η παραγωγική φάση (production phase) εφαρμόζεται για 5.000.000 βήματα, δηλαδή 10ns, με τις ατομικές συντεταγμένες να σώζονται κάθε 200 βήματα, δηλαδή κάθε 0.4ps. Ενδεικτικά αναφέρονται στο Παράρτημα τα script που χρησιμοποιήθηκαν για το Namd, όπου αναφέρονται και λεπτομερώς όλες οι παράμετροι και οι τιμές που τους δόθηκαν.



Εικόνα 2.22 Αριστερά Σχηματική αναπαράσταση του προς προσομοίωση συστήματος που περιλαμβάνει την πρωτεΐνη, τα μόρια νερού και τα ιόντα. Η πρωτεΐνη και τα ιόντα αναπαρίστανται με βάση την ακτίνα Van der Waals, ενώ τα νερά απεικονίζονται με wireframe. Με μπλε χρώμα απεικονίζεται το μονομερές που χρησιμοποιήθηκε, με κίτρινο τα ιόντα νατρίου και με γαλάζιο τα ιόντα χλωρίου. Δεξιά Το ίδιο σύστημα μετά από περιστροφή 90° γύρω από έναν άξονα παράλληλο στον επιμήκη άξονα του μορίου.

Κεφάλαιο 4

Ανάλυση

Προσομοιώσεων

Στο κεφάλαιο αυτό περιγράφεται η ανάλυση των προσομοιώσεων των συστημάτων που περιγράφηκαν στο προηγούμενο κεφάλαιο. Σκοπός της ανάλυσης είναι η μελέτη της σταθερότητας και των κινητικών και δυναμικών ιδιοτήτων των μορίων αυτών.

Στα πλαίσια αυτής της ανάλυσης μελετήσαμε κάποιες παραμέτρους όπως το rmsd από την αρχική και μέση δομή, το rmsf και το radius of gyration (R_G). Το rmsd (root-mean-square-deviation) από την αρχική δομή μας δείχνει κατά πόσο το μόριο απέκλινε από την αρχική του δομή κατά τη διάρκεια της προσομοίωσης. Αποτελεί ένα ποσοτικό μέτρο των δομικών αλλαγών που συμβαίνουν σε μεγάλη κλίμακα στο

μόριο σε σχέση με τη δομή από την οποία ξεκίνησε. Βασίζεται στον υπολογισμό των (ευκλείδειων) αποστάσεων των ατόμων από τις αρχικές τους θέσεις και δίδεται από την εξίσωση:

$$\text{RMSD} = \sqrt{(\sum(x_i - x)^2)/N}$$

όπου N είναι ο αριθμός ατόμων, τα x αντιστοιχούν στις αρχικές συντεταγμένες των ατόμων και τα x_i στις συντεταγμένες που αντιστοιχούν στην εκάστοτε δομή κατά τη διάρκεια του trajectory. Στο Παράρτημα παρατίθεται ενδεικτικά ένα παράδειγμα υπολογισμού του rmsd από την αρχική δομή του (με το πρόγραμμα xplor (Brunger, 1992)).

Κατά τη διάρκεια μίας προσομοίωσης, οι ατομικές συντεταγμένες σφάζονται, όπως είδαμε στο Κεφάλαιο 3, κάθε 0.4ps. Αυτό έχει σαν αποτέλεσμα να προκύψει μία οικογένεια από πολλές δομές από τις οποίες περνάει το μόριο κατά τη διάρκεια της προσομοίωσης. Η οικογένεια αυτή δομών (διαδοχικών στο χρόνο) είναι γνωστή ως τροχιά (trajectory). Με βάση αυτές, μπορεί να προκύψει μία μέση δομή (average structure), η οποία δεν είναι πραγματική υπό την έννοια ότι το μόριο κατά τη διάρκεια της προσομοίωσης την υιοθετεί, αλλά μας βοηθά να βγάλουμε χρήσιμα συμπεράσματα. Έτσι υπολογίζοντας το rmsd σε σχέση αυτή τη φορά με τη μέση δομή μπορούμε να δούμε πόσο σταθερό είναι ένα μόριο, αν σταθεροποιείται σε μία δομή ή αν συγκλίνει (converge) σε άλλη/ες δομές. Ο τύπος υπολογισμού του rmsd από τη μέση δομή είναι ο ίδιος μόνο που σε αυτήν την περίπτωση το x δεν αντιπροσωπεύει τις συντεταγμένες των ατόμων στην αρχική δομή αλλά στη μέση.

$$\text{RMSD} = \sqrt{(\sum(x_i - \langle x \rangle)^2)/N}$$

Στο Παράρτημα παρατίθεται ενδεικτικά ένα παράδειγμα υπολογισμού της μέσης δομής ενός μορίου κατά τη διάρκεια του trajectory και του rmsd από αυτή τη μέση δομή (με το πρόγραμμα xplor).

Το rmsf (root-mean-square-fluctuation) είναι ένα μέτρο της κινητικότητας που παρουσιάζει κάθε άτομο σε σχέση με τη θέση του στη μέση δομή κατά τη διάρκεια του trajectory. Υπολογίζεται με την εξίσωση:

$$\text{RMSF} = \sqrt{(\sum(R_i - \langle R_i \rangle)^2)/N}$$

όπου $\langle R_i \rangle$ είναι η μέση θέση του ατόμου (i) κατά τη διάρκεια του trajectory και $(R_i - \langle R_i \rangle)$ είναι η (ευκλείδειος) απόσταση ανάμεσα στη θέση του ατόμου (σε κάποιο στιγμιότυπο) και στη μέση θέση. Με τη μέθοδο υπολογισμού της μέσης δομής που χρησιμοποιήθηκε (με το xplor), οι τιμές αυτές αναγράφονται στο αρχείο της μέσης

δομής (average.pdb) στην στήλη που αντιστοιχεί στους παράγοντες B (B factors). Μάλιστα οι τιμές αυτές μπορούν να συγκριθούν άμεσα με τις πειραματικές τιμές των παραγόντων B από τις κρυσταλλογραφικά προσδιορισμένες δομές (όταν υπάρχουν). Η σχέση που συνδέει τους παράγοντες B με το r_{msf} είναι η ακόλουθη:

$$B_i = (8/3)\pi^2(RMSF)^2$$

Το radius of gyration (R_G) είναι μία παράμετρος (long-range) χαρακτηριστική του μεγέθους ενός σώματος οποιουδήποτε σχήματος. Ισοδυναμεί με την υδροδυναμική ακτίνα μίας σφαίρας με παρόμοιες υδροδυναμικές ιδιότητες και αντανακλά το πόσο συμπαγές είναι ένα μόριο. Στις αναλύσεις χρησιμοποιούμε τον όρο αυτό σαν ένδειξη της σταθερότητας και της διατήρησης της δομής (Zagrovic *et al.*, 2005). Η εξίσωση που περιγράφει το R_G είναι η ακόλουθη:

$$R_G = \sqrt{\langle (r_i - \langle r_i \rangle)^2 \rangle}$$

όπου $\langle r_i \rangle$ είναι η μέση απόσταση από το κέντρο βάρους (mass-weighted) του μορίου κατά τη διάρκεια ολόκληρου του trajectory και r_i είναι η απόσταση από το κέντρο βάρους του μορίου σε κάθε στιγμιότυπο. Με τον τρόπο αυτό υπολογίζουμε πως μεταβάλλεται το R_G του μορίου σε σχέση με μία μέση τιμή. Στο Παράρτημα παρατίθεται ενδεικτικά ένα παράδειγμα του υπολογισμού του R_G του μορίου κατά τη διάρκεια του trajectory (με το πρόγραμμα xplor).

Μία άλλη σειρά αναλύσεων (που πραγματοποιήθηκαν με το πρόγραμμα carma) αφορά στη δημιουργία χαρτών αποστάσεων. Τέτοιου είδους χάρτες μας παρέχουν χρήσιμες πληροφορίες σχετικά με την κινητικότητα της πρωτεΐνης, όπως ποια τμήματα κινούνται σε μεγαλύτερο βαθμό σε σχέση με άλλα κατά τη διάρκεια του trajectory. Υπολογίζοντας τη μέση απόσταση κάθε Ca ατόμου από κάθε άλλο για όλο το trajectory προκύπτει ένας χάρτης αποστάσεων μεταξύ Ca ατόμων (Ca-Ca distance map). Με τον τρόπο αυτό ετοιμάζεται και ένας δεύτερος χάρτης ο οποίος περιέχει τα r_{msd} που αντιστοιχούν σε αυτές τις αποστάσεις μεταξύ των Ca ατόμων. Ενδιαφέρον επίσης παρουσιάζει και ένας χάρτης ο οποίος αντιπροσωπεύει το r_{msd} του κάθε frame από κάθε άλλο (frame-to-frame r_{msd}). Ένας τέτοιος χάρτης συμβάλλει στην εξαγωγή συμπερασμάτων σχετικά με την κινητικότητα του μορίου ενώ καθιστά περισσότερο ξεκάθαρη τη δημιουργία clusters (ομάδων) δομών που αρχίζουν ήδη να διακρίνονται από το r_{msd} από την αρχική δομή.

* Όλοι οι τύποι προετοιμάστηκαν με το πρόγραμμα latex (Lamport, 1985)

Ακολούθως μπορούμε να υπολογίσουμε ένα πίνακα variance-covariance (Mayer *et al.*, 2003) από τον οποίο στη συνέχεια (με κανονικοποίηση των τιμών) προκύπτει ένας πίνακας cross-correlation για τα Ca άτομα (Hünenberger *et al.*, 1995). Ο τελευταίος πίνακας παίρνει τιμές από -1 μέχρι +1. Έχοντας μία γραφική αναπαράσταση ενός πίνακα συσχέτισης των μεταθέσεων του κάθε Ca ατόμου στο καρτεσιανό σύστημα συντεταγμένων (και για τις τρεις συντεταγμένες x, y, z) είμαστε σε θέση να δούμε πως σχετίζονται μεταξύ τους οι διάφορες κινήσεις που πραγματοποιεί κάθε τμήμα του μορίου (Lange *et al.*, 2005) καθώς επίσης τα δευτεροταγή στοιχεία της δομής που κινούνται σαν συλλογικές ομάδες (collective groups) (Ichiye *et al.*, 1991). Έτσι, όταν δύο τμήματα του μορίου κινούνται μαζί θα εμφανίζουν θετική συσχέτιση (correlated, τιμή +1), ενώ όταν κινούνται αντίθετα θα εμφανίζουν αρνητική συσχέτιση (anticorrelated, τιμή -1). Τέλος, οι κινήσεις που πραγματοποιούν μπορεί να μην εμφανίζουν κάποια συσχέτιση (uncorrelated, τιμή 0) (Luo *et al.*, 2002).

Η επόμενη σειρά αναλύσεων ανήκει στο λεγόμενο Principal Component Analysis. Σε αυτό περιλαμβάνεται ο υπολογισμός των eigenvalues (χαρακτηριστική τιμή) και eigenvectors (χαρακτηριστικό άνυσμα) (Amadei *et al.*, 1993). Μία χαρακτηριστική τιμή (eigenvalue) ενός δισδιάστατου πίνακα (square matrix) είναι μία τιμή η οποία αναπαρίσταται με το γράμμα λ . Ένα χαρακτηριστικό άνυσμα είναι ένα μη-μηδενικό άνυσμα, τέτοιο ώστε να ικανοποιείται η εξίσωση:

$$A\mathbf{x} = \lambda\mathbf{x}$$

όπου A είναι ο δισδιάστατος πίνακας (που είναι ο πίνακας variance-covariance που υπολογίστηκε πριν), λ είναι ο eigenvalue και \mathbf{x} είναι ο eigenvector που του αντιστοιχεί. Εάν αναλύσουμε την (σύνθετη) κίνηση που πραγματοποιεί το μόριο σε ένα άθροισμα απλών κινήσεων, αυτές οι παράμετροι (κάθε ζεύγος eigenvalue-eigenvector) θα αντιπροσωπεύουν κάθε μία από τις ανεξάρτητες αυτές απλές κινήσεις.

Ο υπολογισμός της προβολής των ατομικών διακυμάνσεων σε επιλεγμένους eigenvectors (για παράδειγμα των τριών πρώτων που θα αντιπροσωπεύουν τις κύριες κινήσεις που πραγματοποιεί το μόριο), μας επιτρέπει να εξετάσουμε εάν η διάρκεια της προσομοίωσης ήταν αρκετή για να μας δώσει ένα ικανό δείγμα της κίνησης του μορίου (sufficient sampling). Από άποψη κινητικής της πρωτεϊνικής αναδίπλωσης το πρόβλημα έγκειται στο γεγονός ότι οι κινήσεις σε μακρομοριακό επίπεδο είναι πολύ μεγαλύτερης κλίμακας και όχι της τάξης των nanosecond που πραγματοποιείται μία προσομοίωση. Οι μοριακές κινήσεις που εμπíπτουν στο ενδιαφέρον μας τοποθετούνται στην κλίμακα των msec ή και μεγαλύτερη (Larson *et al.*). Ωστόσο, είναι δυνατόν να υπάρχει sufficient sampling για κάποιες κύριες διαμορφώσεις του

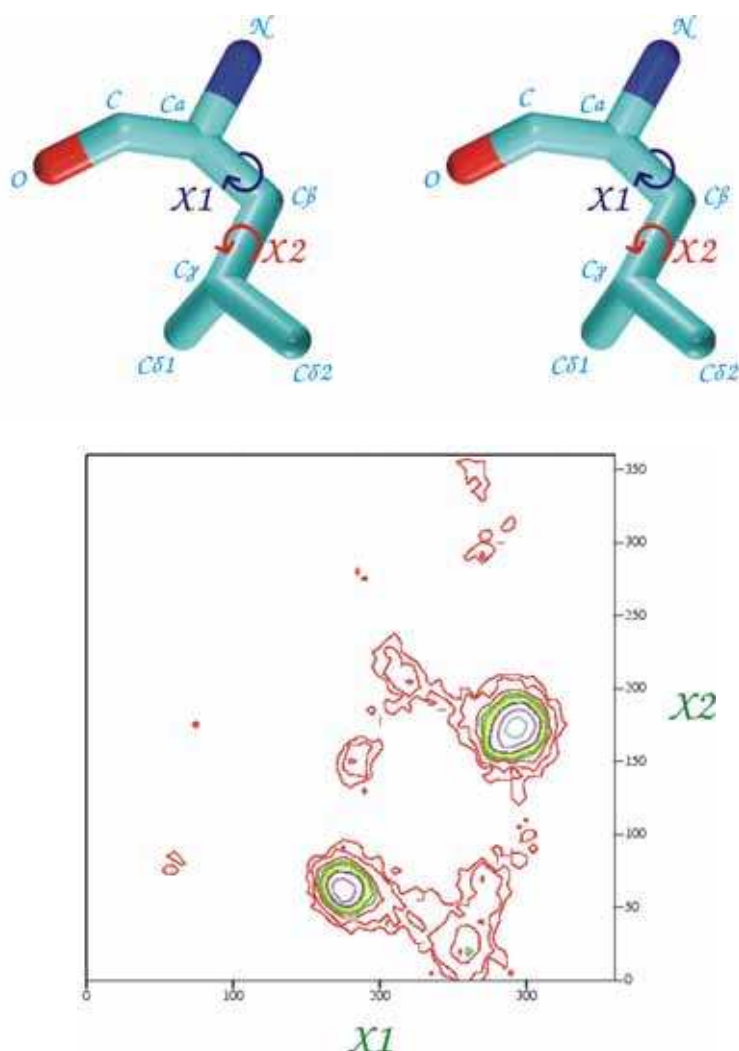
μορίου που βλέπουμε κατά τη διάρκεια της προσομοίωσης. Δεν είμαστε σε θέση ωστόσο να αποκλείσουμε την ύπαρξη επιπλέον διαμορφώσεων τις οποίες δεν προλάβουμε να παρατηρήσουμε.

Τέλος, εφόσον έχουμε sufficient sampling για κάποιο eigenvector, μπορούμε να δημιουργήσουμε ένα τεχνητό trajectory το οποίο θα αναπαριστά την κίνηση που αντιστοιχεί στο συγκεκριμένο eigenvector και με αυτό τον τρόπο να δούμε κάποια χαρακτηριστική κίνηση που πραγματοποιεί το μόριο. Εδώ θα πρέπει να σημειωθεί πως σε αυτό το trajectory δεν υπάρχει η έννοια του χρόνου, απλά αναπαρίσταται η κίνηση. Ολοκληρώνοντας την μελέτη των eigenvectors, εξετάζουμε την διακύμανση των eigenvectors ανά κατάλοιπο. Έτσι μπορούμε να δούμε ποια κατάλοιπα συμμετέχουν σε μεγαλύτερο βαθμό σε κάθε χαρακτηριστική κίνηση. Ακολούθως, υπολογίζοντας την αθροιστική διακύμανση, για παράδειγμα των πέντε πρώτων eigenvector, που όπως θα δούμε στη συνέχεια αντιπροσωπεύουν την πλειοψηφία των κύριων κινήσεων, βλέπουμε ποια κατάλοιπα παρουσιάζουν μεγαλύτερη διακύμανση κατά την κίνησή τους. Τα αποτελέσματα αυτά είναι άμεσα συγκρίσιμα με αυτά που προκύπτουν από το rmsf. Για τον υπολογισμό της αθροιστικής διακύμανσης χρησιμοποιήθηκε ένα απλό πρόγραμμα (SumFluct, Sum of Fluctuations) που δημιούργησε η συγγραφέας σε γλώσσα προγραμματισμού C (ο πηγαίος κώδικας του προγράμματος παρατίθεται στο Παράρτημα).

Μία άλλη ενδιαφέρουσα ανάλυση είναι η ομαδοποίηση των δομών που παρατηρούνται κατά τη διάρκεια του trajectory (cluster analysis). Σε πρώτη προσέγγιση είναι δυνατόν να κατασκευαστεί ένα δενδρόγραμμα ανάλογο με αυτό που αναπαριστά την απόσταση ανάμεσα στα είδη (φυλογενετικό δένδρο) χρησιμοποιώντας σαν πίνακα αποστάσεων, τον πίνακα rmsd. Με αυτόν τον τρόπο όλες οι δομές του trajectory ομαδοποιούνται βάσει του rmsd. Ο αριθμός των clusters που προκύπτουν είναι ενδεικτικός των κύριων δομών που υιοθετεί το μόριο. Μπορεί, στατιστικά, να υπολογιστεί και ένα ποσοστό των δομών που περιέχει το κάθε cluster έτσι ώστε να υπάρχει μία ένδειξη για το πόσο παραμένει το μόριο σε μία διαμόρφωση. Όλοι οι στατιστικοί υπολογισμοί έγιναν με το πρόγραμμα R (R Development Core Team, 2004). Η μέθοδος του cluster analysis ενός trajectory θεωρείται ένας πολύ αποτελεσματικός τρόπος για να διαφανεί κατά πόσο παρατηρείται σύγκλιση (convergence) των δομών (Smith *et al.*, 2002).

Στην Ενότητα 3.2, είδαμε τις υποθέσεις που έχουν διατυπωθεί αναφορικά με τη δομή που υιοθετούν οι διάφορες παραλλαγές της Rop. Για την [Ala₂Leu₂]₈ δεν έχει προσδιορισθεί ακόμη δομή με κρυσταλλογραφία ή φασματοσκοπία NMR. Κατά συνέπεια προετοιμάσαμε δύο συστήματα με διαφορετικές τοπολογίες. Μέσα από την ανάλυση των προσομοιώσεων των συστημάτων αυτών μπορούμε να βγάλουμε συμπεράσματα σχετικά με την σταθερότητα και ίσως ακόμα και την προτίμηση του μορίου προς μία τοπολογία. Στα πλαίσια αυτά ιδιαίτερο ενδιαφέρον παρουσιάζει η

συμπεριφορά των καταλοίπων του υδρόφοβου πυρήνα, που σε αυτήν την περίπτωση συνίσταται αποκλειστικά από κατάλοιπα αλανίνης και λευκίνης. Υπολογίζοντας τις τιμές των διεδρων γωνιών κατά τη διάρκεια του trajectory μπορούμε να προσδιορίσουμε τα rotamers από τα οποία περνάει ένα κατάλοιπο. Με την ομαδοποίηση των παραπάνω αποτελεσμάτων και τον υπολογισμό ποσοστών για την κάθε ομάδα μπορούμε να δούμε ποια είναι τα κυρίαρχα rotamer για το κατάλοιπο που μας ενδιαφέρει και πόσο παραμένει σε αυτά (residence times). Η αλανίνη δεν έχει rotamers (παίρνει μόνο μία διαμόρφωση), καθώς η πλευρική της ομάδα είναι μία μεθυλομάδα. Η πλευρική ομάδα όμως της λευκίνης, της επιτρέπει να έχει αρκετά rotamers. Προκειμένου να προσδιοριστεί ένα rotamer χρειαζόμαστε (στην περίπτωση της λευκίνης) τις διέδρες γωνίες χ_1 και χ_2 . Η γωνία χ_1 ορίζεται με τα άτομα N-Ca-C β και η γωνία χ_2 ορίζεται με τα άτομα Ca-C β -C γ -C δ_1 .



Εικόνα 4.1 Πάνω Τρισδιάστατο μοντέλο (σε stereo αναπαράσταση) της λευκίνης. Κάτω Κατανομή των τιμών των γωνιών χ_1 και χ_2 για τη λευκίνη (αναπαράγεται άνευ αδείας από Kleywegt *et al.*, 1998).

Στην **Εικόνα 4.1** βλέπουμε ένα τρισδιάστατο μοντέλο για τη λευκίνη, όπου διαφαίνονται τα άτομα με βάση τα οποία ορίζονται οι χ_1 και χ_2 γωνίες καθώς και οι άξονες περιστροφής για κάθε γωνία. Στο κάτω τμήμα της εικόνας βλέπουμε ένα δισδιάστατο χάρτη με την κατανομή των τιμών των χ_1 και χ_2 γωνιών, όπως αυτές έχουν βρεθεί σε κατάλοιπα λευκίνης σε κρυσταλλογραφικά προσδιορισμένες δομές πρωτεϊνών (Kleywegt *et al.*, 1998). Στο *Παράρτημα* παρατίθεται ένας τρόπος υπολογισμού των διέδρων γωνιών με τη βοήθεια του προγράμματος `χρlog`, καθώς και ο πηγαίος κώδικας (σε γλώσσα προγραμματισμού C), ενός απλού προγράμματος (DiCon, Dihedral Conversion) που δημιούργησε η συγγραφέας για μετατροπή των τιμών των γωνιών που δίνει το `χρlog`, από τον τριγωνομετρικό κύκλο -180° έως 180° στον τριγωνομετρικό κύκλο 0° έως 360° .

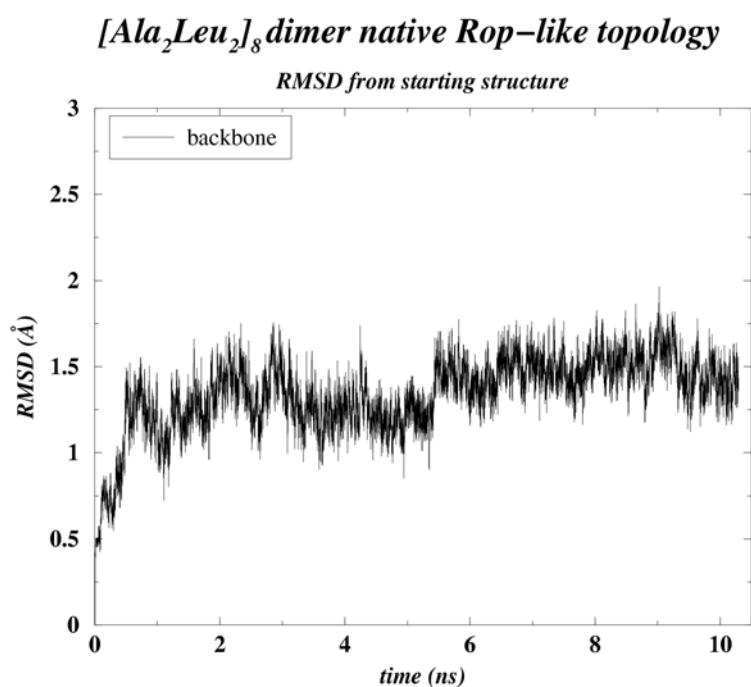
Όπως έχει ήδη αναφερθεί, το μόριο είναι ένα 4- α -ελικοειδές δεμάτιο δύο μονομερών τύπου έλικα-θηλιά-έλικα. Πραγματοποιώντας μία ανάλυση διατήρησης της δευτεροταγούς δομής (secondary structure analysis) μπορούμε να δούμε κατά πόσο διατηρούνται οι έλικες, αν η διαμόρφωσή τους αλλοιώνεται σε κάποια σημεία, για παράδειγμα στα C-τελικά άκρα καθώς και τη συμπεριφορά των στροφών.

Να σημειωθεί πως για όλες τις αναλύσεις που προαναφέρθηκαν θα πρέπει να αφαιρεθούν οι περιστροφές και οι μεταθέσεις του μορίου. Αυτό γίνεται με την υπέρθεση όλων των `frame` στο αρχικό χρησιμοποιώντας τον αλγόριθμο του Kabsch (Kabsch, πρόγραμμα `lsqkab`, 1976). Αυτό το least-square fitting έγινε με το πρόγραμμα `carma`.

4.1 Ανάλυση προσομοίωσης διμερούς $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία

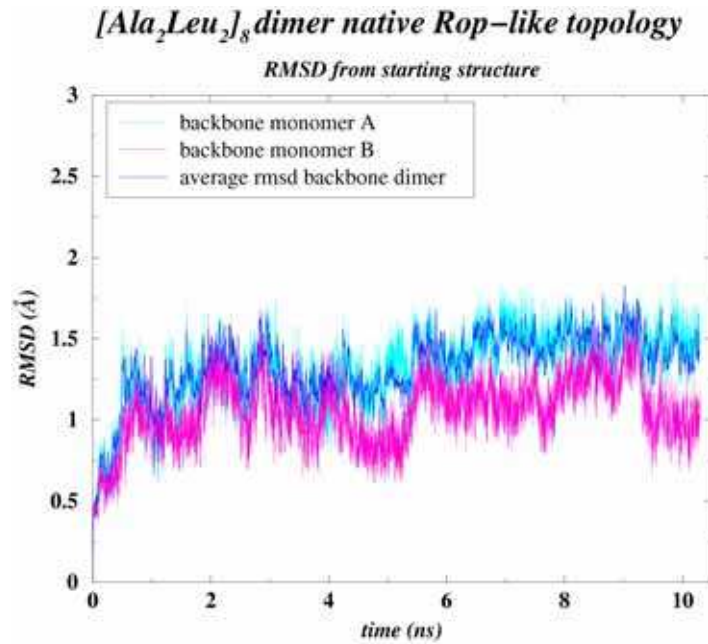
4.1.1 Rmsd από την αρχική δομή

Το πρώτο βήμα σε μία ανάλυση προσομοίωσης είναι ο υπολογισμός του rmsd από την αρχική δομή. Στην **Εικόνα 4.2** παρουσιάζεται το rmsd λαμβάνοντας υπόψη μόνο τα άτομα του σκελετού (backbone) του μορίου. Βλέπουμε πως μετά από ένα σύντομο χρονικό διάστημα (που αντιστοιχεί στο heating) το rmsd σταθεροποιείται γύρω στα 1.3 Å όπου και παραμένει για περίπου τη μισή διάρκεια της προσομοίωσης (μέχρι περίπου τα 5ns). Στη συνέχεια φαίνεται να συμβαίνει μία σχετικά απότομη αλλαγή (φαίνεται σαν σκαλοπάτι στην Εικόνα) και ακολούθως σταθεροποιείται ξανά γύρω στα 1.5 Å μέχρι το τέλος της προσομοίωσης (δηλαδή μέχρι τα 10ns). Κατά συνέπεια το μόριο μάλλον υιοθετεί δύο διακριτές δομές. Η υπόθεση αυτή θα εξεταστεί στη συνέχεια.

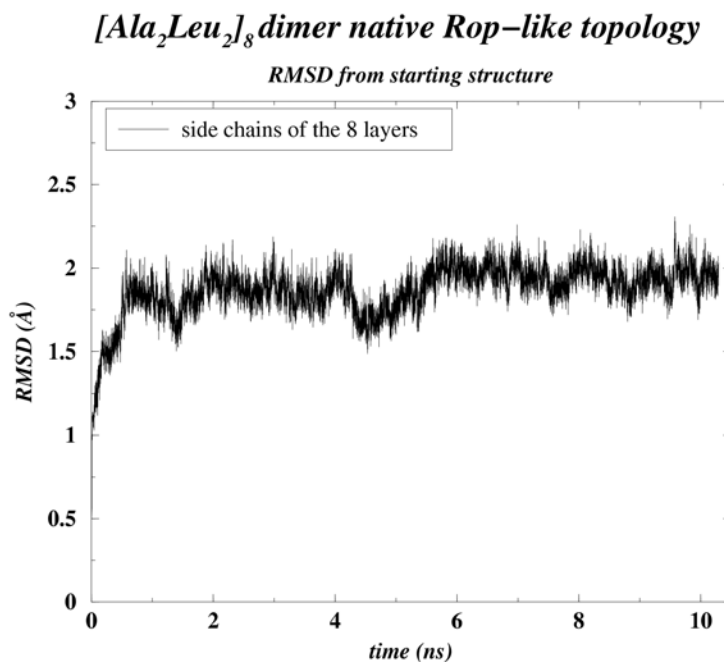


Εικόνα 4.2

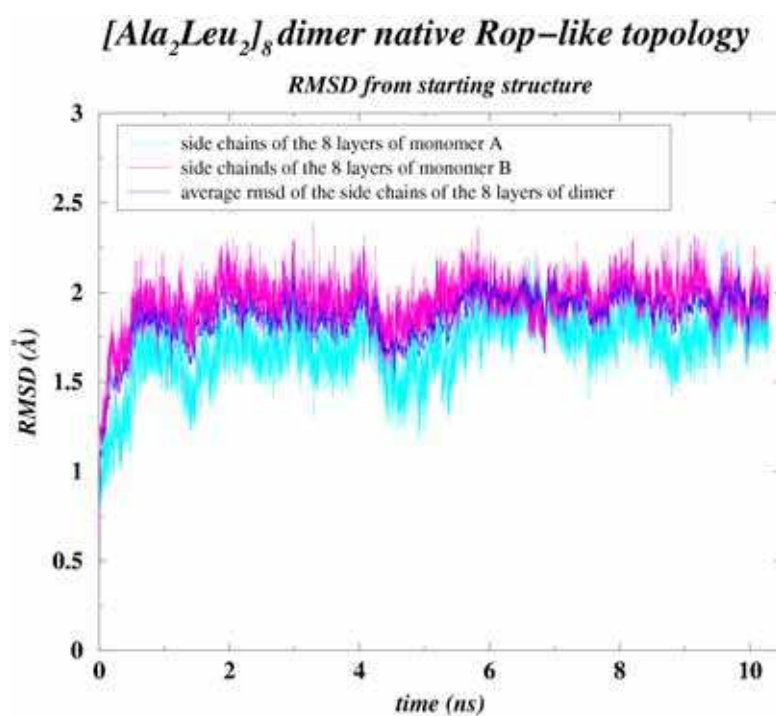
Η αλλαγή αυτή στη δομή φαίνεται πως οφείλεται περισσότερο στο ένα μονομερές, όπως παρατηρούμε και στην **Εικόνα 4.3**. Το μονομερές B παρουσιάζει μικρότερη κινητικότητα κατά τη διάρκεια όλης της προσομοίωσης, σε αντίθεση με το μονομερές A το οποίο φαίνεται να είναι πιο ασταθές, με τιμή rmsd γύρω στα 1.3-1.4 Å. Το μονομερές B έχει γενικά μικρότερες τιμές rmsd. Στο πρώτο μισό της προσομοίωσης κυμαίνεται γύρω στο 1 Å και στη συνέχεια αυξάνεται στα 1.2 Å.



Εκτός από την κινητικότητα των ελίκων σαν συμπαγή μόρια, μας ενδιαφέρει και η κινητικότητα των πλευρικών ομάδων (χωρίς να ληφθούν υπόψη τα άτομα υδρογόνου) των καταλοίπων που συνιστούν τα οκτώ επίπεδα του υδρόφοβου πυρήνα της πρωτεΐνης. Εξετάσαμε λοιπόν το rmsd που παρουσιάζουν κατά τη διάρκεια της προσομοίωσης σε σχέση με τη δομή από την οποία ξεκίνησαν. Με τον τρόπο αυτό έχουμε ένα μέτρο της σταθερότητας του πυρήνα. Όπως βλέπουμε στην **Εικόνα 4.4**, μετά την πάροδο της φάσης του heating, το rmsd σταθεροποιείται γύρω στα 1.8 Å και παρουσιάζει μικρή διακύμανση. Συνεπώς ο πυρήνας φαίνεται να είναι σχετικά σταθερός.



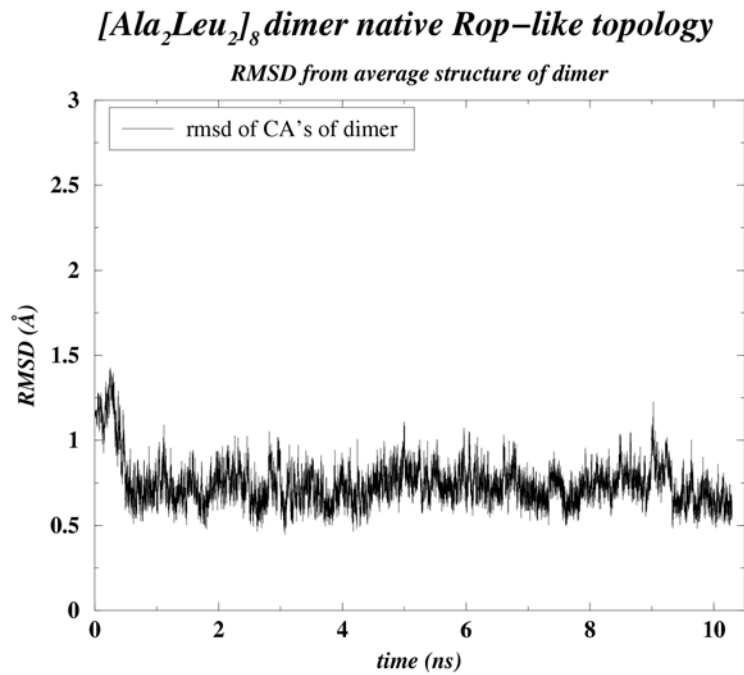
Οι πλευρικές ομάδες του κάθε μονομερούς μεμονωμένα επίσης είναι αρκετά σταθερές κατά τη διάρκεια της προσομοίωσης (**Εικόνα 4.5**). Στην περίπτωση αυτή όμως, βλέπουμε πως οι πλευρικές ομάδες του μονομερούς A έχουν μικρότερο rmsd, γύρω στα 1.6-1.7 Å, ενώ του μονομερούς B κυμαίνονται στα 1.9-2.0 Å, σε αντίθεση με το rmsd που εμφανίζει ο σκελετός του μονομερούς, που είναι μεγαλύτερο για το μονομερές A και μικρότερο για το μονομερές B (βλ. Εικόνα 4.3). Το γεγονός αυτό, ότι οι πλευρικές ομάδες δείχνουν περισσότερη κινητικότητα σε σχέση με τον σκελετό, προξένησε εντύπωση, αφού φαίνεται οι πλευρικές ομάδες να συμπεριφέρονται ανεξάρτητα, παρά το γεγονός πως συνδέονται με το σκελετό.



Εικόνα 4.5

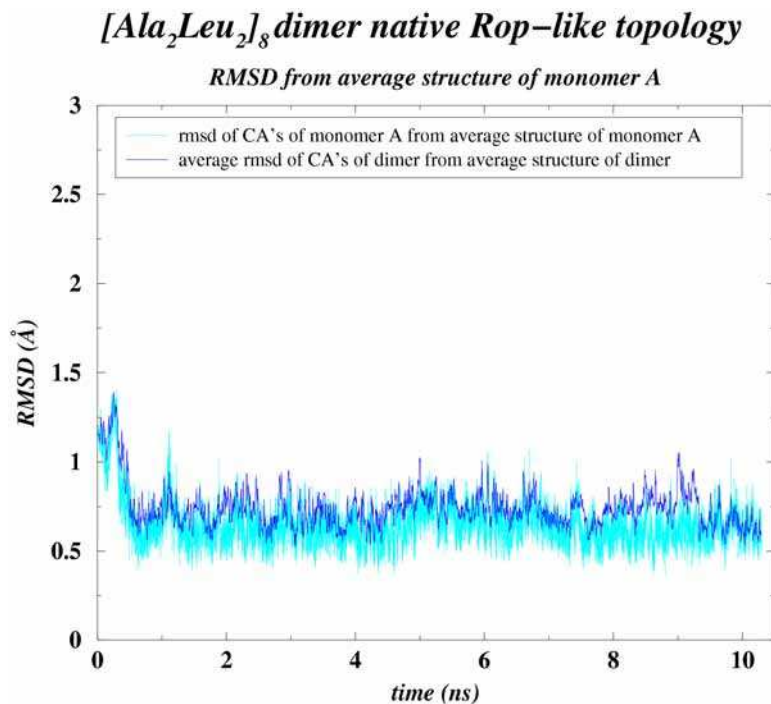
4.1.2 Rmsd από τη μέση δομή

Μετά τον υπολογισμό μίας μέσης δομής, υπολογίζουμε το rmsd που εμφανίζει η δομή σε κάθε στιγμιότυπο (frame) του trajectory από τη μέση δομή. Στην **Εικόνα 4.6** βλέπουμε πως το rmsd σταθεροποιείται γρήγορα (αμέσως μετά τη φάση του heating) σε μικρή τιμή (0.7 Å). Ο υπολογισμός έγινε λαμβάνοντας υπόψη μόνο τα Ca άτομα του διμερούς, σε σχέση με τη μέση δομή του διμερούς. Το σκαλοπάτι που είδαμε στην προηγούμενη ενότητα στην Εικόνα 4.2 γύρω στα 5ns συνεχίζει να υπάρχει και εδώ, μόνο που είναι λιγότερο ευδιάκριτο. Το γεγονός αυτό συνεισφέρει στην άποψη για την ύπαρξη δύο διακριτών δομών.



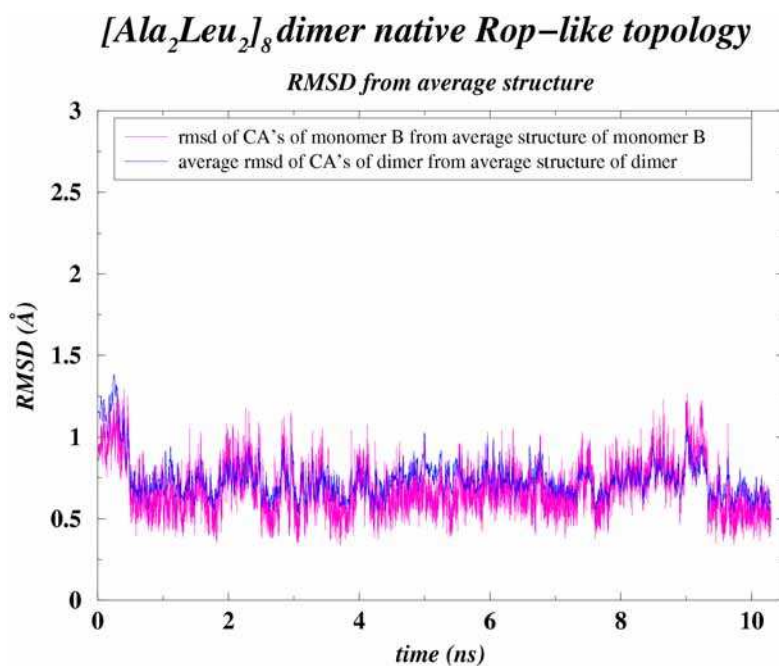
Εικόνα 4.6

Για την περίπτωση του κάθε μονομερούς μεμονωμένα, βλέπουμε στην **Εικόνα 4.7** πως το μονομερές A (λαμβάνοντας υπόψη κατά τον υπολογισμό μόνο τα Ca άτομα) είναι επίσης αρκετά σταθερό, με rmsd από τη μέση δομή (του μονομερούς) λίγο μικρότερη από την αντίστοιχη του διμερούς, στα 0.6 Å. Η μέση δομή του κάθε μονομερούς υπολογίστηκε λαμβάνοντας υπόψη μόνο τα άτομα του σκελετού.



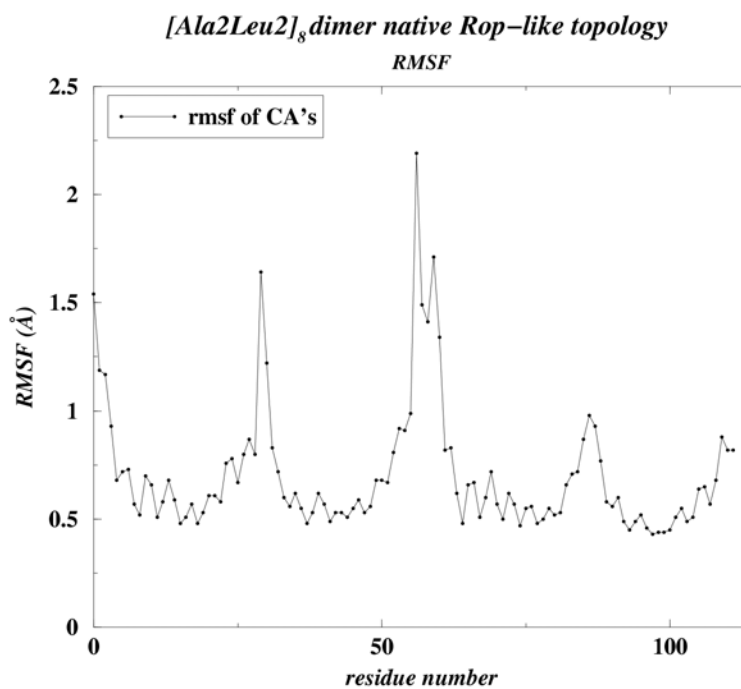
Εικόνα 4.7

Το rmsd του μονομερούς B (λαμβάνοντας υπόψη κατά τον υπολογισμό μόνο τα Ca άτομα) από τη μέση δομή του κυμαίνεται στην ίδια τιμή με αυτήν του μονομερούς A, αλλά εμφανίζει μεγαλύτερη διακύμανση (**Εικόνα 4.8**). Το γεγονός αυτό συμφωνεί με την παρατήρηση που κάναμε στην προηγούμενη ενότητα, ότι το μονομερές B εμφανίζει μεγαλύτερη κινητικότητα σε σχέση με το μονομερές A (βλ. Εικόνα 4.3). Μία μικρή αλλαγή φαίνεται και γύρω στα 9ns, η οποία φαίνεται και στο διμερές (βλ. Εικόνα 4.6).



4.1.3 Rmsf

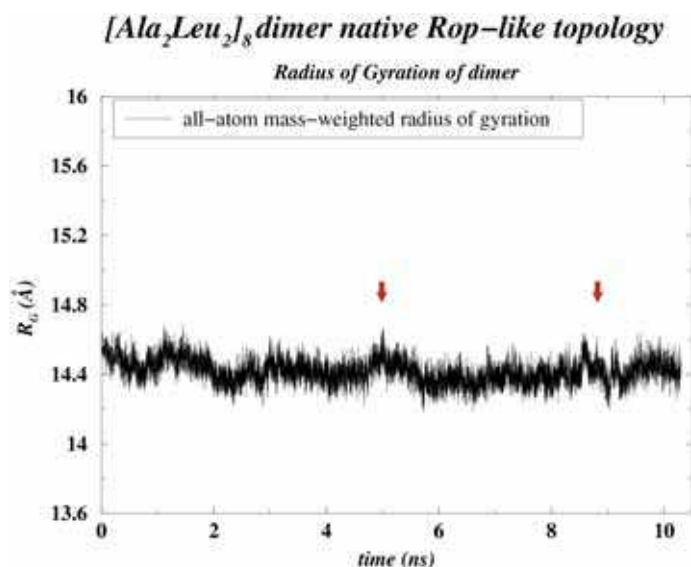
Το rmsf (root-mean-square-fluctuation) είναι μία παράμετρος που δείχνει τη διακύμανση ενός ατόμου σε σχέση με τη θέση που καταλαμβάνει στη μέση δομή. Στην **Εικόνα 4.9** μπορούμε να διακρίνουμε τα κατάλοιπα που παρουσιάζουν μεγαλύτερη διακύμανση, δηλαδή έχουν μεγαλύτερη κινητικότητα. Υψηλές τιμές εμφανίζουν τα κατάλοιπα στην αρχή (1-2), στις στροφές (28-30) και στο τέλος (54-56) κάθε μονομερούς. Τα αποτελέσματα είναι αναμενόμενα, δεδομένου ότι τα κατάλοιπα αυτά δεν συμμετέχουν στο σχηματισμό σταθερής δομής, όπως θα δούμε και στη συνέχεια.



Εικόνα 4.9

4.1.4 Εξέλιξη του Radius of Gyration (R_G)

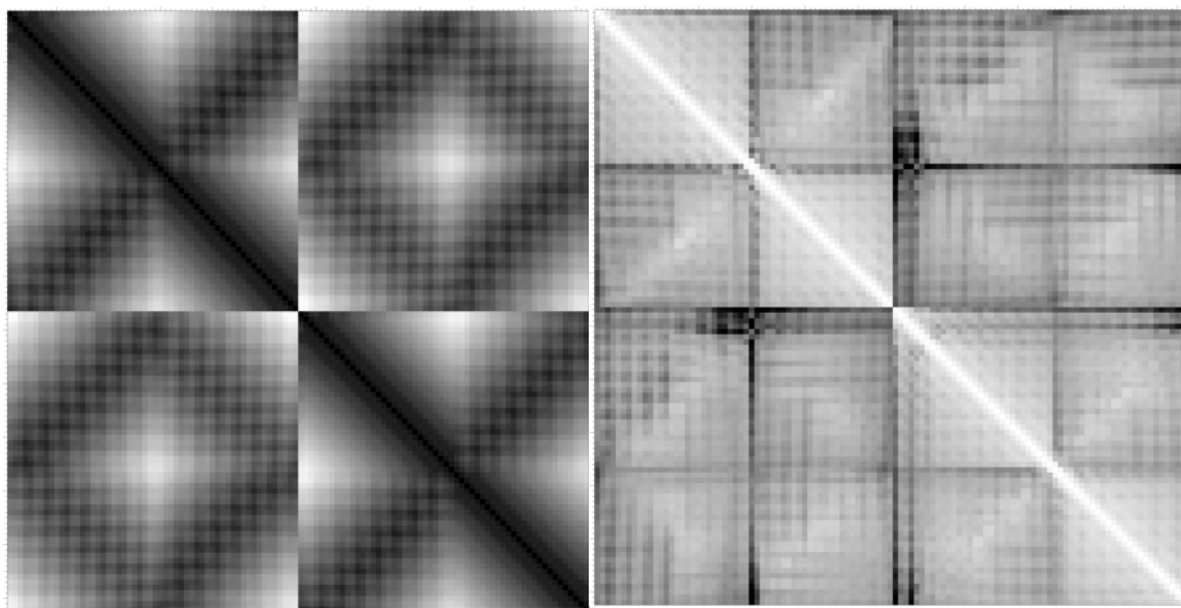
Στην **Εικόνα 4.10** απεικονίζεται η εξέλιξη του R_G του ομοδιμερούς με native (Rop-like) τοπολογία κατά τη διάρκεια του trajectory. Παρατηρούμε ότι κυμαίνεται γύρω από την τιμή των 14.5 Å - 14.6 Å και δεν παρουσιάζει σημαντικές διακυμάνσεις. Το γεγονός αυτό δείχνει πως το μόριο είναι αρκετά σταθερό. Η αρχική τιμή του R_G , που είναι αυτή της κρυσταλλικής δομής, είναι 14.55 Å. Οι διακυμάνσεις που είχαν παρατηρηθεί στις προηγούμενες ενότητες γύρω στα 5 και 9ns συνεχίζουν να φαίνονται και εδώ (υποδεικνύονται με τα κόκκινα βέλη).



Εικόνα 4.10

4.1.5 Χάρτης της μέσης απόστασης μεταξύ Ca ατόμων και αντίστοιχο rmsd

Όπως αναφέρθηκε και στην αρχή του Κεφαλαίου, μπορούμε να φτιάξουμε ένα πίνακα ο οποίος να περιέχει τις μέσες αποστάσεις μεταξύ όλων των Ca ατόμων. Μία αναπαράσταση ενός τέτοιου πίνακα φαίνεται στην **Εικόνα 4.11**. Η πάνω οριζόντια γραμμή και η αριστερή κάθετη στήλη, αντιστοιχούν στα Ca άτομα των καταλοίπων του A και B μονομερούς. (Το πρώτο μεγάλο τετράγωνο αντιστοιχεί στο πρώτο μονομερές και το δεύτερο, στο δεύτερο μονομερές. Τα κατάλοιπα της στροφής εντοπίζονται στο κέντρο του κάθε τετραγώνου.) Κατά συνέπεια η εικόνα είναι συμμετρική, με άξονα συμμετρίας τη διαγώνια γραμμή (υπάρχει και δεύτερος άξονας συμμετρίας, η άλλη διαγώνιος). Η διαβάθμιση του χρώματος από μαύρο σε λευκό ακολουθεί την αύξηση της απόστασης μεταξύ Ca ατόμων, με τις μικρότερες να απεικονίζονται με μαύρο και τις μεγαλύτερες με λευκό. Η μέγιστη απόσταση μεταξύ Ca ατόμων βρέθηκε να είναι 43.7 Å και όπως αναμένεται την παρουσιάζουν τα άτομα της στροφής του ενός μονομερούς σε σχέση με το άλλο μονομερές και τα C-τελικά κατάλοιπα του ενός μονομερούς σε σχέση με το άλλο μονομερές. Μικρές αποστάσεις έχουν βρεθεί μεταξύ είτε ατόμων που ανήκουν στην ίδια έλικα, είτε βρίσκονται στο ίδιο περίπου επίπεδο σε απέναντι έλικες.



Εικόνα 4.11 Ca-Ca distance map

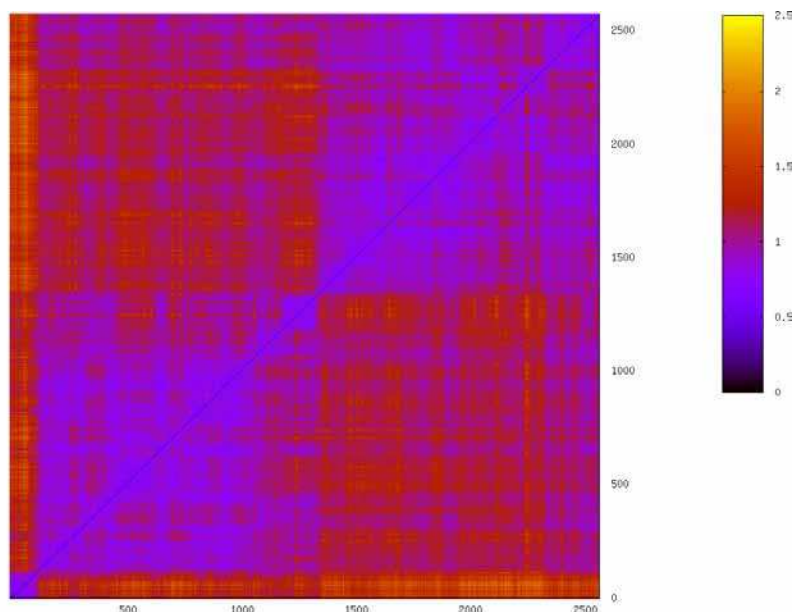
Εικόνα 4.12 rmsd που αντιστοιχεί στον Ca-Ca distance map

Μεγαλύτερο ενδιαφέρον παρουσιάζουν τα αποτελέσματα που απεικονίζονται στην **Εικόνα 4.12**. Στην εικόνα αυτή βλέπουμε μία γραφική αναπαράσταση των rmsd που αντιστοιχούν στην προηγούμενη εικόνα. Ισχύουν όσα προαναφέρθηκαν με τη διαφορά ότι ο χρωματικός κώδικας είναι ανεστραμμένος, δηλαδή οι μικρότερες τιμές rmsd φαίνονται με λευκό και οι μεγαλύτερες με μαύρο. Η μεγαλύτερη τιμή rmsd που

βρέθηκε ήταν 2.07 Å. Στην εικόνα αυτή μπορούμε να διακρίνουμε ποια τμήματα της πρωτεΐνης είναι ιδιαίτερα ευκίνητα. Τα 4 N-τελικά του B μονομερούς κινούνται πολύ σε σχέση με τα κατάλοιπα της στροφής του A μονομερούς. Τα κατάλοιπα αυτά δεν σχηματίζουν σταθερή δομή και είναι εκτεθειμένα προς το διάλυμα. Το πρώτο κατάλοιπο που συμμετέχει στον πυρήνα είναι το 5. Είναι λοιπόν αναμενόμενο να έχουν αυξημένη κινητικότητα. Επίσης μεγάλο rmsd εμφανίζουν τα Ca άτομα των καταλοίπων 1 και 4 του B μονομερούς σε σχέση με όλη τη δεύτερη έλικα του B μονομερούς. Ο λόγος είναι και πάλι πως τα κατάλοιπα αυτά δεν συμμετέχουν στο σχηματισμό σταθερής α-έλικας. Τέλος, φαίνεται να κινούνται πολύ τα κατάλοιπα της στροφής του A μονομερούς σε σχέση με ολόκληρο το B μονομερές.

4.1.6 Rmsd μεταξύ των δομών του τροχιακού

Με το rmsd από την αρχική δομή έχουμε μία άποψη της κινητικότητας του μορίου και των δομών που υιοθετεί. Ωστόσο, υπολογίζοντας το rmsd ανάμεσα σε κάθε στιγμιότυπο (frame) του trajectory μπορούμε να έχουμε μία πιο ξεκάθαρη εικόνα. Μία γραφική απεικόνιση ενός τέτοιου υπολογισμού φαίνεται στην **Εικόνα 4.13**. Στον οριζόντιο και κάθετο άξονα αντιστοιχούν τα frame (συνολικά $2.573 \times 2.573 = 6.620.329$, μαζί με τη φάση του heating, καθώς ο υπολογισμός έγινε με βήμα 10). Όπως φαίνεται στο ένθετο δεξιά, η διαβάθμιση του χρώματος από μαύρο σε κίτρινο ακολουθεί την αύξηση στην τιμή του rmsd. Η μέγιστη τιμή βρέθηκε να είναι 2.5 Å. Από τα χρώματα της εικόνας συνεπάγεται ότι η μεγάλη πλειοψηφία των τιμών rmsd βρέθηκαν περίπου από 1 έως 2 Å. Οι τιμές αυτές σε γενικές γραμμές συμφωνούν με τα αποτελέσματα που αναπαρίστανται στην Εικόνα 4.2.

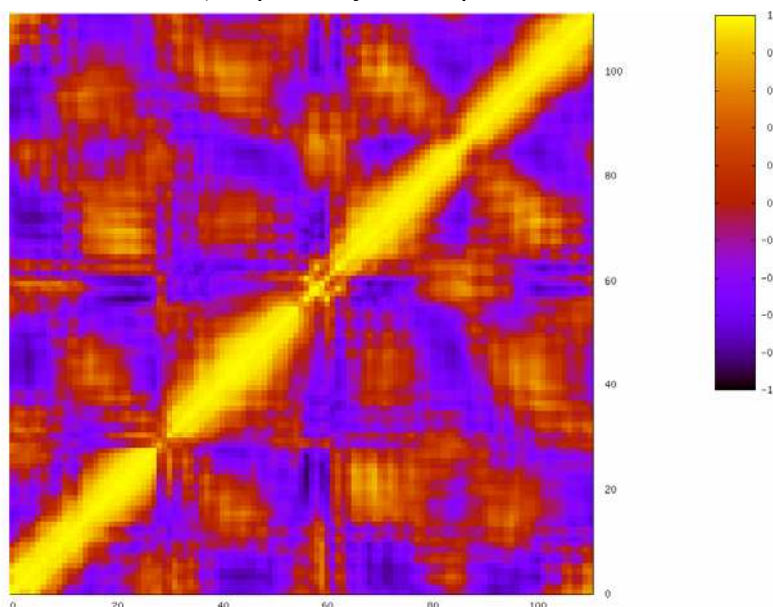


Εικόνα 4.13 frame-to-frame rmsd. Οι μονάδες στους άξονες είναι σε frame, με τις τιμές να είναι μία τάξη μεγέθους μικρότερες (επειδή ο υπολογισμός έγινε με βήμα 10).

Μία πιο προσεκτική παρατήρηση μας επιτρέπει να διακρίνουμε τη δημιουργία ομάδων δομών (clusters). Τα clusters φαίνονται σαν τετράγωνα διαφορετικού χρώματος που το μέγεθος τους αντικατοπτρίζει τον αριθμό των frame στα οποία παρατηρείται η συγκεκριμένη δομή. Οι δομές που συναντάμε σε κάθε frame δεν είναι πανομοιότυπες, αλλά το rmsd που παρουσιάζει η δομή του ενός frame σε σχέση με τη δομή του άλλου είναι τόσο μικρό, ώστε να μπορούμε να κατατάξουμε τις δομές των δύο frame στην ίδια ομάδα δομών (βλ. Ενότητα 4.1.9). Στα πρώτα 1000 (100 x 10) περίπου frame (δηλαδή περίπου 0.4ns) το μόριο φαίνεται να συμπεριφέρεται διαφορετικά και παρουσιάζει σχετικά υψηλή τιμή rmsd (γύρω στα 2 Å) σε σχέση με τα υπόλοιπα. Δύο μεγάλα cluster είναι εμφανή. Το ένα αφορά κατά προσέγγιση τα frame 1000 (100 x 10) – 14000 (1400 x 10) (δηλαδή περίπου από 0.4 – 5.6 ns) και το δεύτερο τα frame 14000 (1400 x 10) – 25730 (2573 x 10) (5.6 – 10.2 ns). Επίσης, διαφαίνεται μία έντονη γραμμή γύρω στο frame 22000 (8.8ns) που οδηγεί στη δημιουργία μίας ακόμη διακριτής ομάδας με μικρή όμως απόκλιση στην τιμή του rmsd. Βλέπουμε λοιπόν πως τα δεδομένα συνεχίζουν να υποστηρίζουν την ύπαρξη δύο διακριτών δομών, ενώ επαληθεύονται όλα τα προηγούμενα αποτελέσματα. Σε αυτήν την περίπτωση όμως είμαστε σε θέση να ξέρουμε με ακρίβεια πότε συμβαίνουν αλλαγές στο μόριο.

4.1.7 Variance-Covariance και cross-correlation

Όπως είδαμε στην εισαγωγική ενότητα του κεφαλαίου, ένας πίνακας variance-covariance and cross-correlation θα μας δίνει πληροφορίες σχετικά με το ποια τμήματα του μορίου κινούνται μαζί ή αντίθετα. Μία γραφική απεικόνιση αυτού του πίνακα (του cross-correlation) παρουσιάζεται στην **Εικόνα 4.14**.



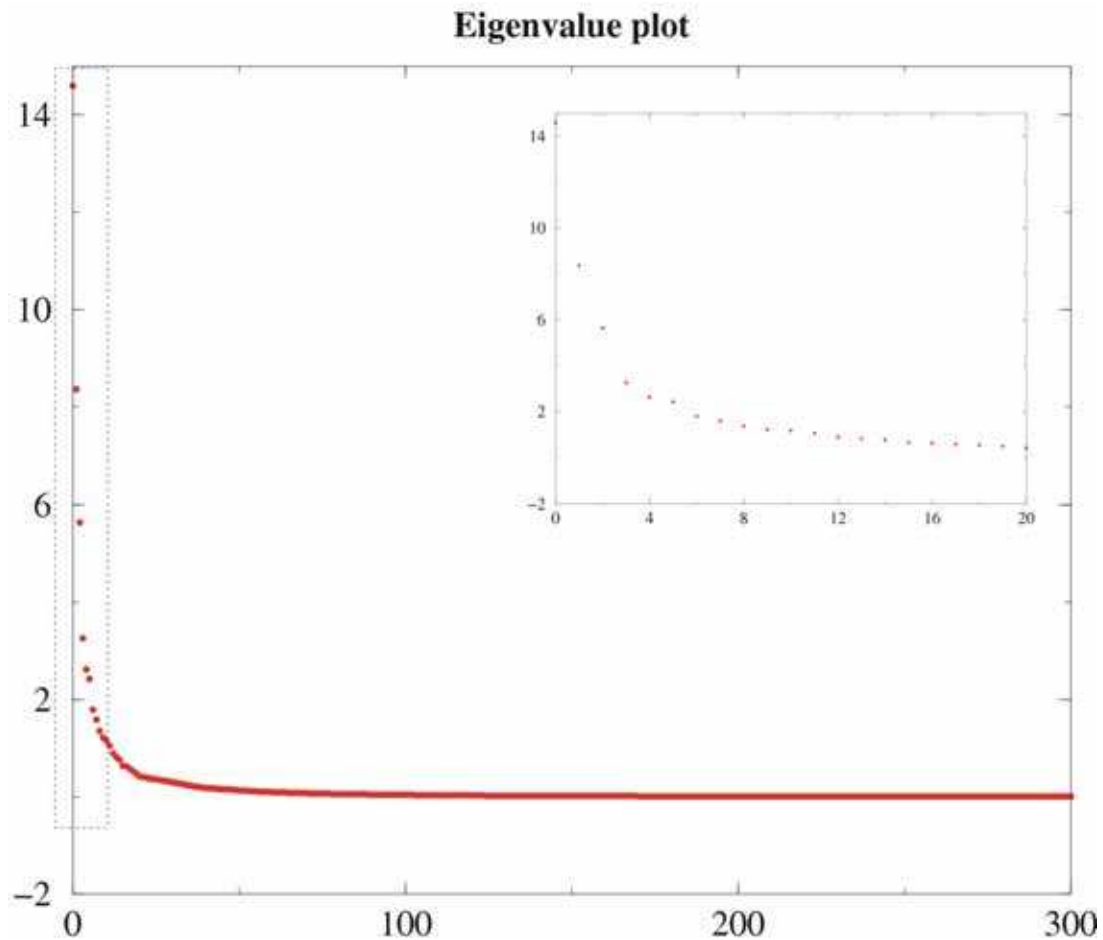
Εικόνα 4.14 variance-covariance and cross-correlation

Στον οριζόντιο και κάθετο άξονα αντιστοιχούν τα κατάλοιπα αρχικά του A μονομερούς (1-56) και στην συνέχεια του B (57-112). Στο ένθετο δεξιά αναφέρεται η διαβάθμιση του χρώματος από μαύρο σε κίτρινο. Η συσχέτιση παίρνει τιμές από -1 (αρνητική συσχέτιση) έως +1 (θετική συσχέτιση). Με κόκκινο σημειώνονται οι κινήσεις που δεν εμφανίζουν συσχέτιση (τιμή 0).

Η διαγώνιος γραμμή αντιστοιχεί στο ίδιο κατάλοιπο (ως εκ τούτου θα έχει πλήρη συσχέτιση με τον εαυτό του και για αυτό το λόγο φαίνεται με έντονο κίτρινο χρώμα). Το πάχος της γραμμής όμως περιλαμβάνει περισσότερα από ένα κατάλοιπα. Αυτό υποδεικνύει ότι και τα γειτονικά κατάλοιπα κινούνται μαζί, κάτι που είναι αναμενόμενο αφού είναι ομοιοπολικά συνδεδεμένα. Γύρω από τις θέσεις περίπου 30, 60, 85 και 110 φαίνονται ομάδες σημείων με πορτοκαλί χρώμα. Οι θέσεις αυτές αντιστοιχούν στα κατάλοιπα των στροφών και στα N-τελικά άκρα, των οποίων η κίνηση φαίνεται να σχετίζεται (άλλωστε συμμετέχουν στο σχηματισμό του πρώτου και όγδοου επιπέδου του υδρόφοβου πυρήνα). Το ίδιο φαινόμενο παρατηρείται και για τα κατάλοιπα της κάθε έλικας, όπως υποδεικνύεται από τις πορτοκαλί διαγώνιες γραμμές που είναι κάθετες στη κεντρική κίτρινη διαγώνιο. Ακόμη, είναι αναμενόμενο τα κατάλοιπα που συμμετέχουν στο σχηματισμό ενός επιπέδου ή γειτονικών επιπέδων να κινούνται μαζί, αφού με τις υδρόφοβες αλληλεπιδράσεις που αναπτύσσονται ανάμεσά τους συμβάλλουν στην σταθεροποίηση του πυρήνα. Κινήσεις που δεν σχετίζονται μεταξύ τους εμφανίζονται ανάμεσα σε κατάλοιπα που ανήκουν σε διαφορετικές έλικες και δεν βρίσκονται στο ίδιο επίπεδο (φαίνονται με μωβ διαγώνιες γραμμές που είναι κάθετες στη κεντρική κίτρινη διαγώνιο και παρεμβάλλονται στις πορτοκαλί διαγώνιους που προαναφέρθηκαν). Τέλος, τα κατάλοιπα 1 και 4 του B μονομερούς φαίνεται να κινούνται αντίθετα σε σχέση με όλη την υπόλοιπη έλικα στην οποία ανήκουν (έντονες σκούρες γραμμές), κάτι το οποίο είδαμε και στην Εικόνα 4.12.

4.1.8 Eigenvalues και Eigenvectors

Προκειμένου να δούμε τις χαρακτηριστικές κινήσεις που πραγματοποιεί το μόριό μας, θα πρέπει, όπως αναφέρθηκε στην αρχή του Κεφαλαίου να υπολογίσουμε τους eigenvalues και eigenvectors. Πριν προβούμε όμως σε οποιοδήποτε άλλου είδους ανάλυση θα πρέπει να εξετάσουμε εάν υπάρχει sufficient sampling (για λεπτομέρειες βλ. Ενότητα 4.1). Για το σκοπό αυτό υπολογίζουμε τις διακυμάνσεις (fluctuations) των τριών πρώτων eigenvectors, οι οποίοι, όπως φαίνεται στην **Εικόνα 4.15** θα αντιπροσωπεύουν την πλειοψηφία των κύριων κινήσεων του μορίου.



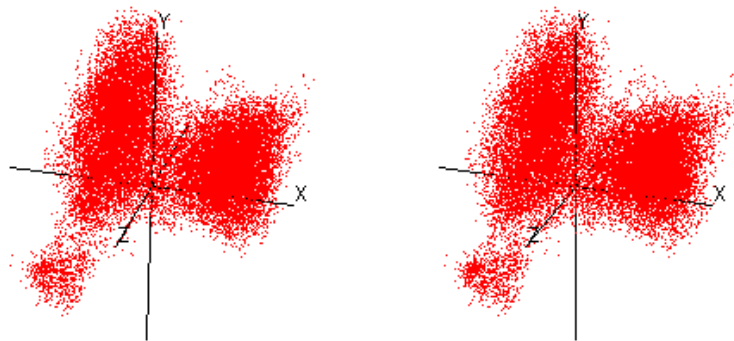
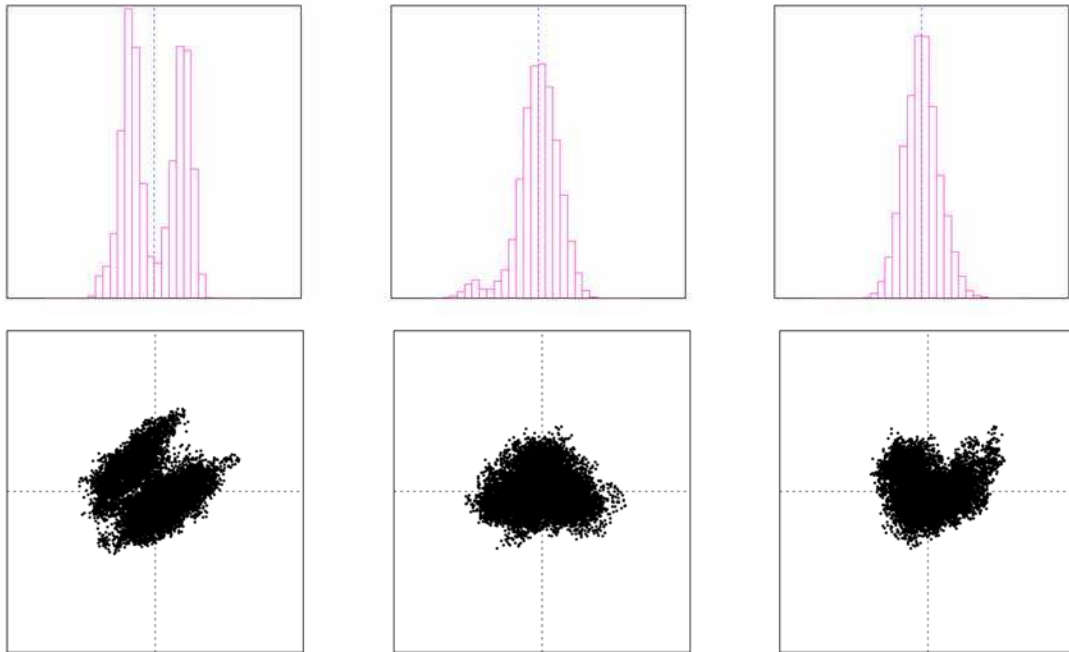
Εικόνα 4.15 Διαγραμματική απεικόνιση των eigenvalues. Στο ένθετο δεξιά φαίνεται σε μεγέθυνση το τμήμα του γραφήματος που περικλείεται με το διακεκομμένο κουτί. Οι πρώτοι 10 eigenvectors μπορούν να εξηγήσουν σχεδόν πλήρως τη κίνηση του μορίου.

Στην **Εικόνα 4.16** παρατίθεται μία γραφική απεικόνιση των τριών πρώτων eigenvectors. Στην περίπτωση που ο χρόνος της προσομοίωσης ήταν αρκετός για να έχουμε sufficient sampling κάποιων κινήσεων, θα δούμε στο ιστόγραμμα (πρώτη σειρά) μία κανονική κατανομή επικεντρωμένη γύρω από το μηδέν.

Στην αναπαράσταση των διακυμάνσεων στο επίπεδο που σχηματίζουν οι δύο πρώτοι eigenvectors θα πρέπει να έχουμε μία κανονική κατανομή επικεντρωμένη γύρω από τη θέση 0,0, (η αναπαράσταση γίνεται στις δύο διαστάσεις) και τέλος στην αναπαράσταση των διακυμάνσεων στο επίπεδο που σχηματίζουν οι τρεις πρώτοι eigenvectors θα έχουμε μία κανονική κατανομή επικεντρωμένη γύρω από την θέση 0,0,0 (εφόσον η αναπαράσταση γίνεται στις τρεις διαστάσεις).

Βλέπουμε λοιπόν, πως sufficient sampling έχει επιτευχθεί για τον πρώτο eigenvector, αλλά είναι ευδιάκριτη η δημιουργία δύο cluster. Αυτό αποτελεί ένδειξη ότι ίσως δε θα πρέπει να μελετήσουμε τους eigenvectors κατά τη διάρκεια ολόκληρου του trajectory. Εφόσον από τη μέχρι τώρα ανάλυση έχει προκύψει πως στο μόριο, στη μέση περίπου του χρόνου της προσομοίωσης, συμβαίνει μία ευδιάκριτη αλλαγή,

ίσως να έχει περισσότερο νόημα να μελετήσουμε τις κινήσεις που πραγματοποιεί το μόριο, ξεχωριστά για το πρώτο και δεύτερο μισό του trajectory (βλ. αρχή κεφαλαίου, Ενότητα 4.1.1 και Ενότητα 4.1.4 και Ενότητα 4.1.6). Αντιθέτως, για τους επόμενους δύο eigenvectors, φαίνεται πως ο χρόνος της προσομοίωσης ήταν αρκετός για να τους παρακολουθήσουμε.

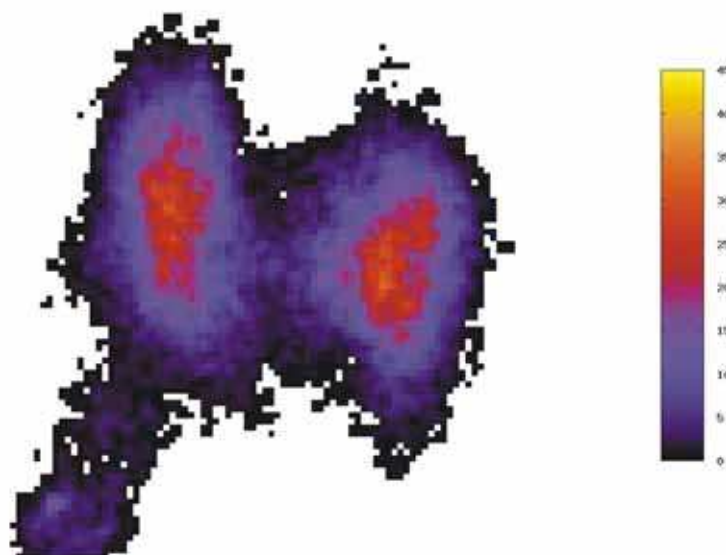


Εικόνα 4.16 Πάνω σειρά Προβολή των διακυμάνσεων των τριών πρώτων eigenvectors στη μία διάσταση.

Μεσαία σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων δύο eigenvectors (δύο διαστάσεις).

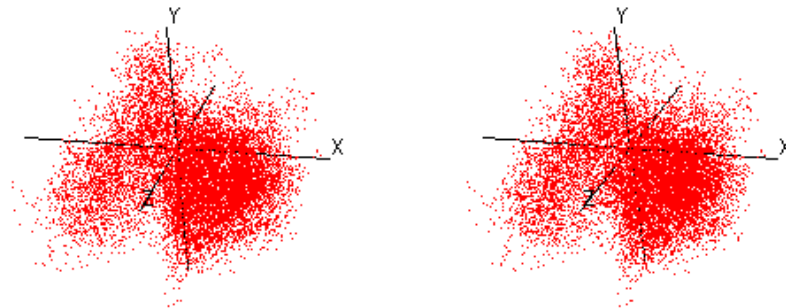
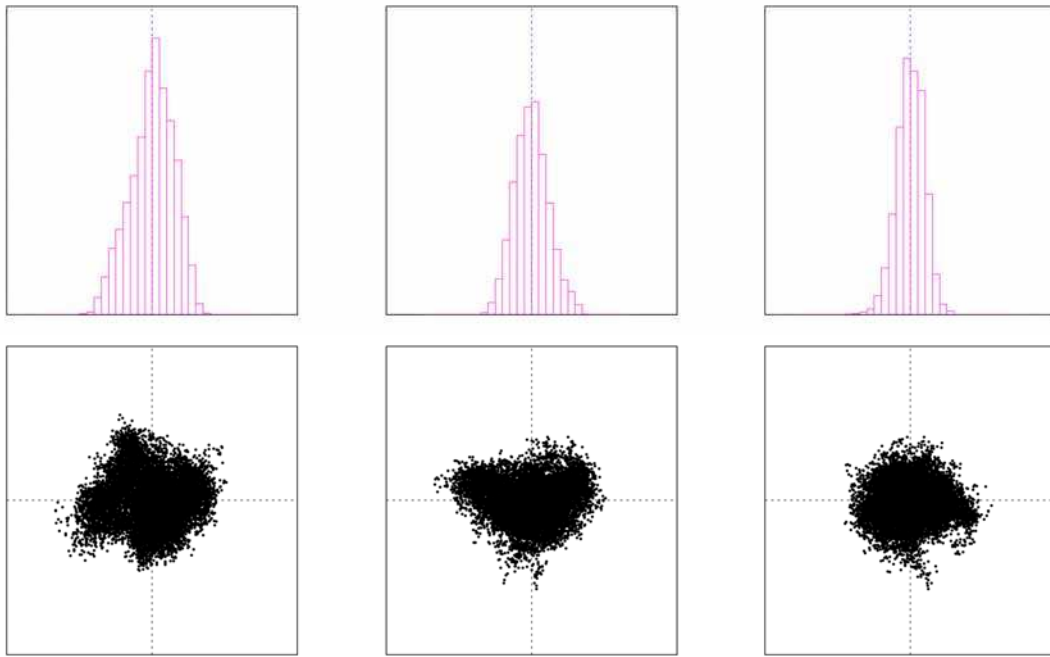
Κάτω σειρά Stereo προβολή των διακυμάνσεων στο επίπεδο των τριών πρώτων eigenvectors (τρεις διαστάσεις).

Στο κάτω μέρος της Εικόνας 4.16, που αναπαρίστανται και οι τρεις eigenvectors σε μία τρισδιάστατη απεικόνιση, θα μπορούσε να ισχυριστεί κάποιος ότι φαίνονται τρία διαφορετικά clusters. Όμως σε μία τέτοιου είδους απεικόνιση δεν είναι δυνατό να γνωρίζουμε τη συνεισφορά που υπάρχει σε κάθε cluster. Με άλλα λόγια, μπορεί κάποιο να φαίνεται σαν ξεχωριστή ομάδα, όμως ο αριθμός των σημείων από τα οποία αποτελείται να μην είναι αρκετά μεγάλος ώστε να θεωρηθεί σαν σημαντικό cluster που θα πρέπει να ληφθεί υπόψη. Σε τέτοιου είδους περιπτώσεις είναι χρήσιμη μία απεικόνιση στην οποία θα υπάρχει πληροφορία αναφορικά με την πυκνότητα (τον αριθμό) των σημείων (distribution of density of points), κάτι ανάλογο δηλαδή με τις ισοϋψείς που χρησιμοποιούνται στα δισδιάστατα γραφήματα. Στην **Εικόνα 4.17** βλέπουμε μία τέτοια προσπάθεια. Η διαβάθμιση του χρώματος από μαύρο (0) σε κίτρινο (45) ακολουθεί την αύξηση στην πυκνότητα των σημείων. Κατά συνέπεια σημαντικά είναι μόνο τα δύο cluster, ενώ στο τρίτο δεν υπάρχει σημαντικός αριθμός σημείων.



Εικόνα 4.17 Προβολή των διακυμάνσεων ανά ζεύγη eigenvectors
βάση πυκνότητας σημείων.

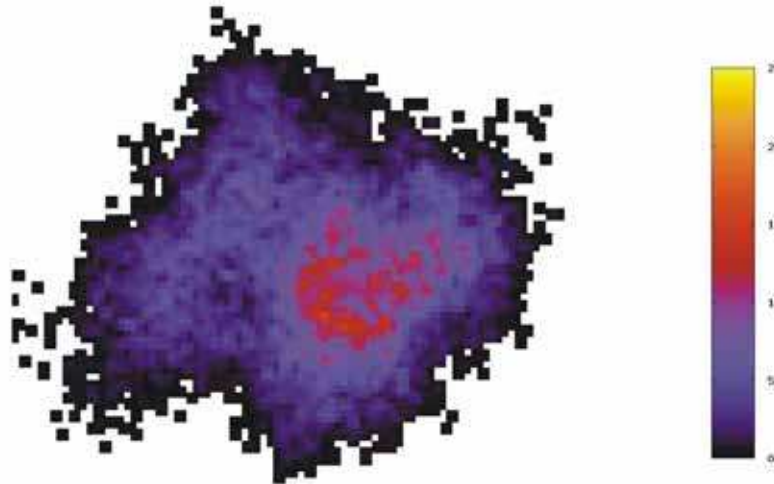
Αν επαναλάβουμε την διαδικασία που περιγράφηκε παραπάνω λαμβάνοντας υπόψη ένα τμήμα μόνο του trajectory, τα αποτελέσματα, αναφορικά με το αν έχουμε sufficient sampling, διαφοροποιούνται. Στις **Εικόνες 4.18** και **4.19** βλέπουμε την αντίστοιχη εικόνα που προκύπτει υπολογίζοντας καινούργια eigenvectors και eigenvalues για το πρώτο μισό της προσομοίωσης. Πιο συγκεκριμένα ο υπολογισμός έγινε για τα frame 1200-13500 (δηλαδή από 0.48ns μέχρι 5.4ns). Η επιλογή αυτή έγινε με βάση τα αποτελέσματα που προέκυψαν υπολογίζοντας το frame-to-frame rmsd (Ενότητα 4.1.6). Με αυτόν τον τρόπο φαίνεται πως υπάρχει sufficient sampling και για τους τρεις eigenvectors σε αυτό το διάστημα της προσομοίωσης.



Εικόνα 4.18 Πάνω σειρά Προβολή των διακυμάνσεων των τριών πρώτων eigenvectors στη μία διάσταση για το πρώτο μισό της προσομοίωσης.
 Μεσαία σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων δύο eigenvectors (δύο διαστάσεις) για το πρώτο μισό της προσομοίωσης.
 Κάτω σειρά Stereo προβολή των διακυμάνσεων στο επίπεδο των πρώτων τριών eigenvectors (τρεις διαστάσεις) για το πρώτο μισό της προσομοίωσης.

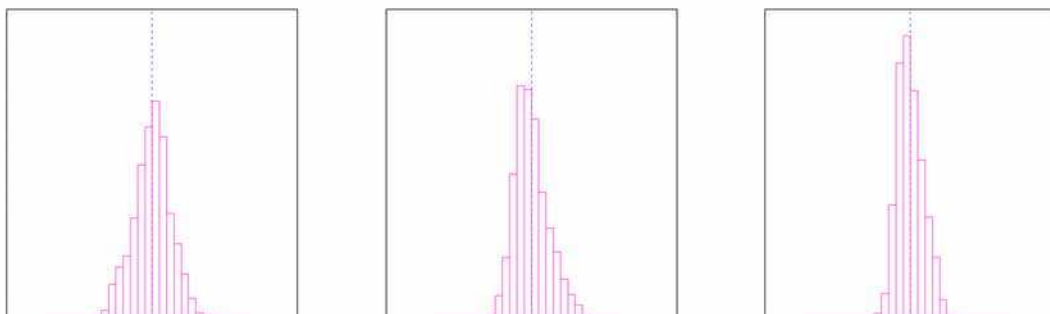
Ακολούθως υπολογίσαμε eigenvalues και eigenvectors για τα frame 13600-25700 (δηλαδή από 5.44ns μέχρι 10.28ns). Η αναπαράσταση των διακυμάνσεών τους φανερώνει πως χρειαζόταν περισσότερος χρόνος για να μελετηθεί ολοκληρωμένα η κίνηση του μορίου για το δεύτερο μισό της προσομοίωσης (**Εικόνα 4.20**). Ενώ δεν φαίνεται να υπάρχει πρόβλημα όσον αφορά το sufficient sampling, για τον πρώτο

eigenvector είναι και πάλι εμφανή τα δύο cluster που δημιουργούνται. Όπως φαίνεται στην **Εικόνα 4.21**, το ένα από τα δύο συνίσταται από σχεδόν τριπλάσιο αριθμό σημείων.

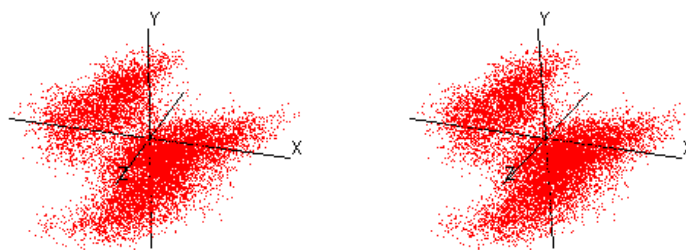
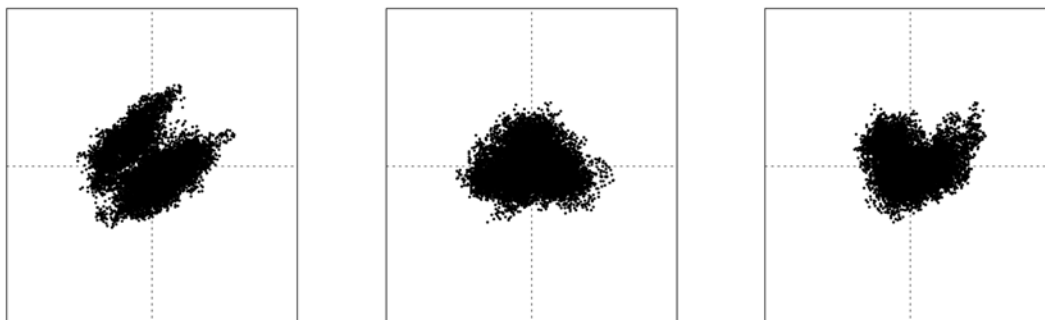


Εικόνα 4.19 Προβολή των διακυμάνσεων ανά ζεύγη eigenvectors βάση πυκνότητας σημείων για το πρώτο μισό της προσομοίωσης.

Η διαδικασία αυτή θα μπορούσε να επαναληφθεί και για επιμέρους τμήματα του δεύτερου μισού της προσομοίωσης, ώστε να φτάσουμε να έχουμε μία σφαιρική κατανομή για τον πρώτο eigenvector. Ωστόσο, δεν θα προβούμε σε κάτι τέτοιο, αφενός γιατί το δεύτερο cluster δεν περιέχει ικανοποιητικό αριθμό σημείων, αφετέρου γιατί η μελέτη των cluster γίνεται σε καλύτερο βαθμό στην Ενότητα 4.1.9 που ακολουθεί.



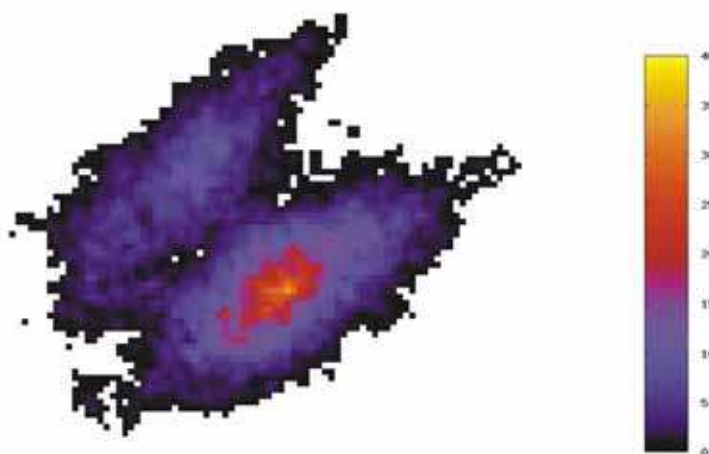
Συνέχεια της εικόνας στην επόμενη σελίδα



Εικόνα 4.20 Πάνω σειρά Προβολή των διακυμάνσεων των τριών πρώτων eigenvectors στη μία διάσταση για το δεύτερο μισό της προσομοίωσης .

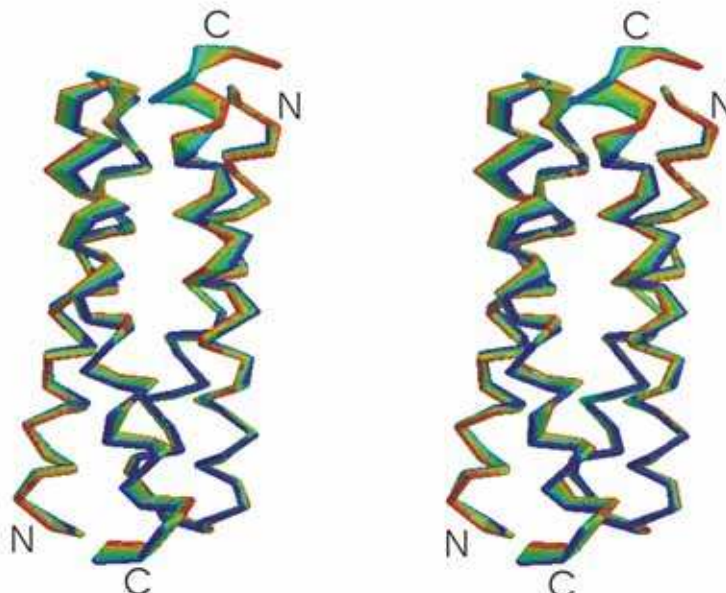
Μεσαία σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων δύο eigenvectors (δύο διαστάσεις) για το δεύτερο μισό της προσομοίωσης.

Κάτω σειρά Stereo προβολή των διακυμάνσεων στο επίπεδο των πρώτων τριών eigenvectors (τρεις διαστάσεις) για το δεύτερο μισό της προσομοίωσης.



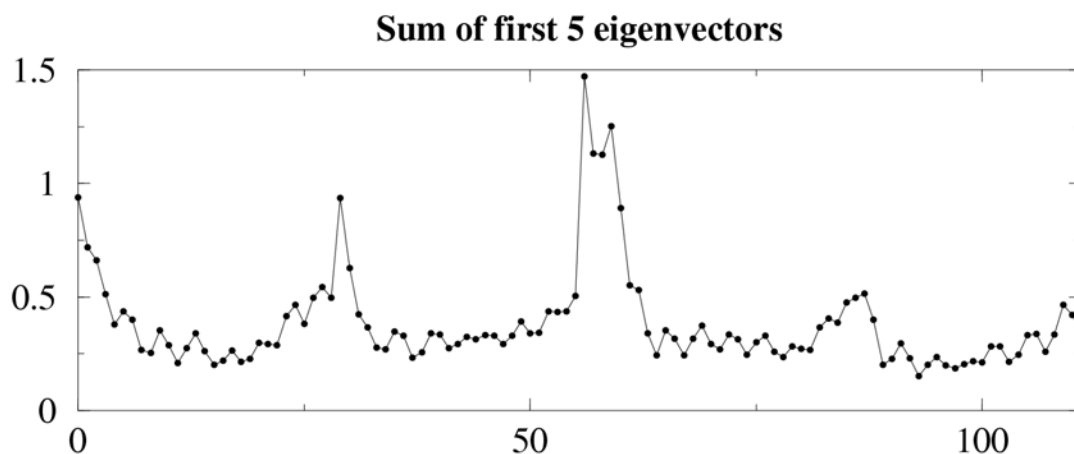
Εικόνα 4.21 Προβολή των διακυμάνσεων ανά ζεύγη eigenvectors βάση πυκνότητας σημείων για το δεύτερο μισό της προσομοίωσης.

Στην **Εικόνα 4.22** βλέπουμε μία γραφική απεικόνιση της κίνησης που πραγματοποιεί το μόριο βάση του πρώτου eigenvector. Ο υπολογισμός έγινε για ολόκληρο το trajectory. Η διαβάθμιση του χρώματος από μπλε σε κόκκινο φανερώνει το εύρος της κίνησης.

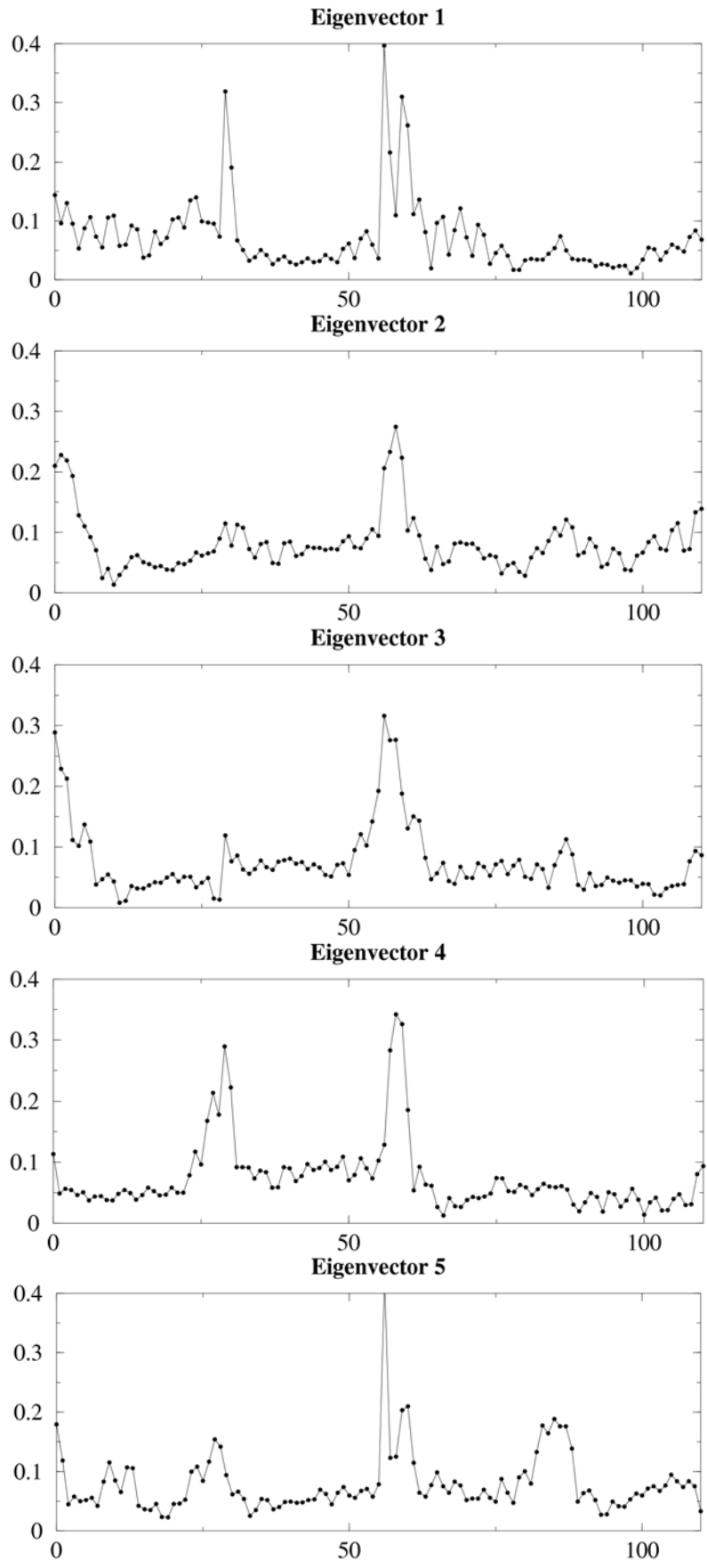


Εικόνα 4.22 Stereo αναπαράσταση της κίνησης του μορίου που περιγράφεται από το πρώτο ζεύγος eigenvalue – eigenvector.

Ολοκληρώνοντας την μελέτη των eigenvectors, μπορούμε να δούμε την κατανομή των διακυμάνσεων του κάθε eigenvector ανά κατάλοιπο. Με τον τρόπο αυτό έχουμε μία άποψη της ποσότητας των σχετιζόμενων διακυμάνσεων του eigenvector μεταξύ των καταλοίπων. Στην **Εικόνα 4.23** βλέπουμε αθροιστικά την διακύμανση των πέντε πρώτων eigenvectors για κάθε κατάλοιπο. Αυτή η αθροιστική διακύμανση συμφωνεί με τα αποτελέσματα από το rmsf (βλ. Ενότητα 4.1.3, Εικόνα 4.9). Στην **Εικόνα 4.24** έχουμε αναλυτικά την αναπαράσταση για κάθε έναν eigenvector ξεχωριστά.



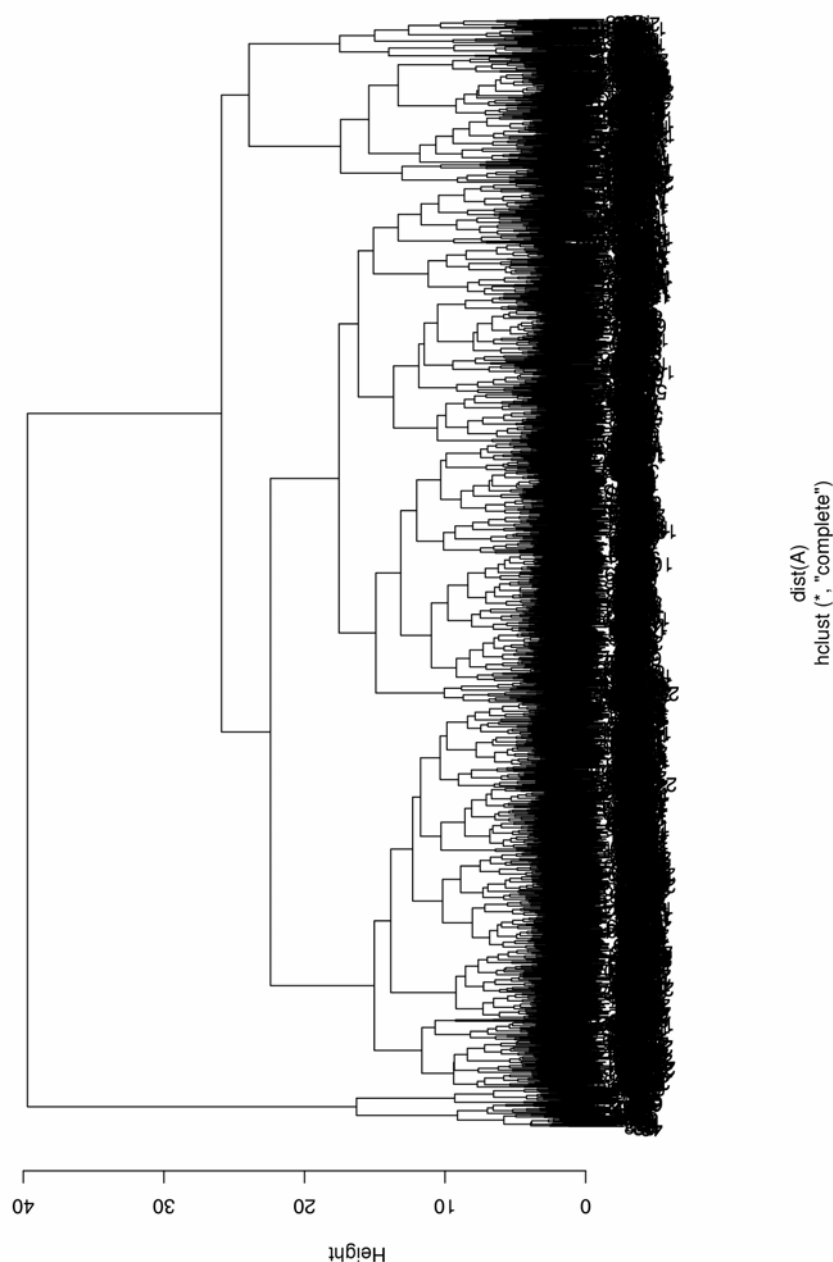
Εικόνα 4.23 Αναπαράσταση της αθροιστικής διακύμανσης των πρώτων 5 eigenvector ανά κατάλοιπο.



Εικόνα 4.24 Αναπαράσταση των διακυμάνσεων των πρώτων 5 eigenvector ανά κατάλοιπο.

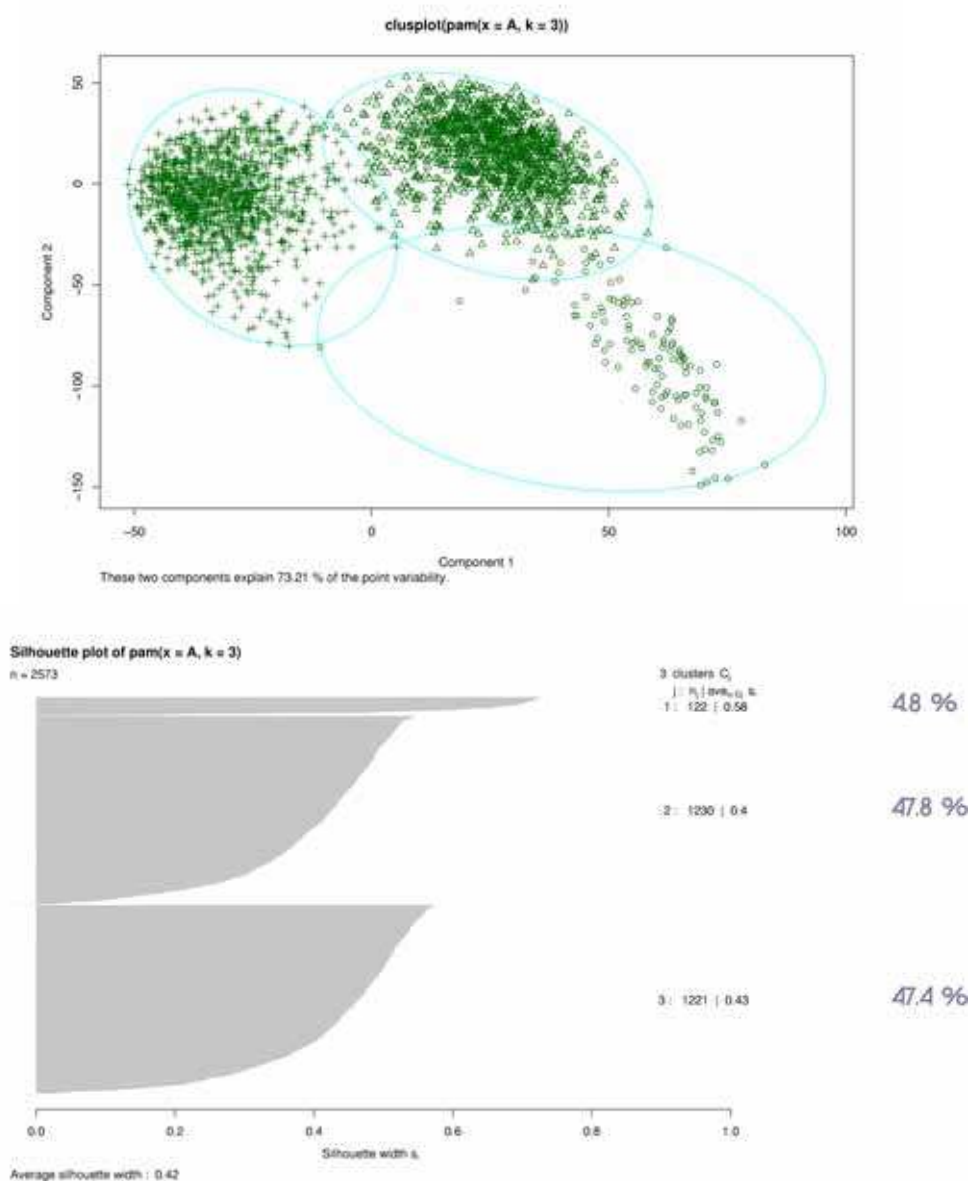
4.1.9 Ομαδοποίηση δομών (cluster analysis) του trajectory

Το πρώτο βήμα στην ομαδοποίηση των δομών που παρατηρούνται κατά τη διάρκεια ενός trajectory είναι η δημιουργία ενός δένδρου. Σε ένα τέτοιο δένδρο, όπως προαναφέρθηκε, λαμβάνεται υπόψη το rmsd ανάμεσα στις δομές (Εικόνα 4.25). Με βάση αυτό το δένδρο μπορούμε να συνάγουμε τον αριθμό των κυρίαρχων cluster δομών. Ένας πιο αξιόπιστος τρόπος για να γίνει αυτό είναι βάση ενός στατιστικού προγράμματος. Με τον τρόπο αυτό προκύπτουν τα στατιστικά σημαντικά clusters.



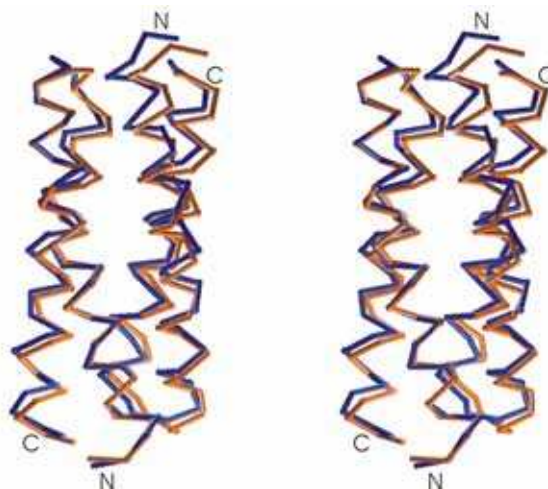
Εικόνα 4.25 Δενδρόγραμμα των δομών που παρατηρούνται κατά τη διάρκεια της προσομοίωσης του διμερούς $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία, βάση του rmsd.

Όπως φαίνεται στην **Εικόνα 4.26** υπάρχουν τρεις σημαντικές ομάδες δομών. Ο αλγόριθμος που χρησιμοποιήθηκε στους παρακάτω υπολογισμούς ήταν ο pam. Ένα σημαντικό πλεονέκτημα της χρήσης της στατιστικής στην ανάλυση, είναι η δυνατότητα υπολογισμού ποσοστών για το κάθε cluster. Με τη μέθοδο αυτή μπορούμε να γνωρίζουμε με ακρίβεια σε ποιο cluster δομών ανήκει το κάθε frame. Έτσι προκύπτει πως το πρώτο cluster, που αποτελεί το 4.8% των δομών περιλαμβάνει τα frame 1-1060 (δηλαδή έως τα 0.42ns). Το δεύτερο cluster αποτελεί το 47.8% των δομών και περιλαμβάνει τα frame 1070-13540 (δηλαδή έως τα 5.42ns) και το τρίτο cluster, που αποτελεί το 47.4% των δομών, περιλαμβάνει τα frame 13550-25730 (δηλαδή έως τα 10.29ns). Τα αποτελέσματα αυτά φαίνονται και από το frame-to-frame rmsd, Εικόνα 4.13.



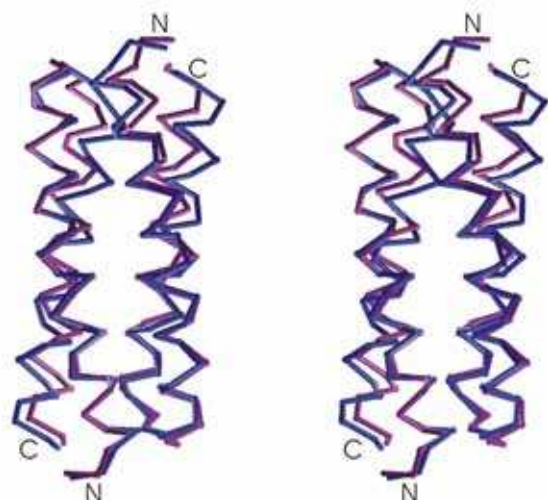
Εικόνα 4.26 Πάνω Προσδιορισμός των clusters. Κάτω Κατανομή των δομών στα τρία clusters και υπολογισμός ποσοστών.

Ακολούθως προσδιορίστηκαν οι μέσες δομές που προκύπτουν από το δεύτερο και τρίτο cluster που αποτελούν τις στατιστικά σημαντικές ομάδες δομών και καταλαμβάνουν μεγάλο τμήμα του trajectory. Στις **Εικόνες 4.27, 4.28 και 4.29** που ακολουθούν γίνεται σύγκριση των δομών αυτών τόσο μεταξύ τους όσο και με την δομή από την οποία ξεκίνησαν. Σε κάθε περίπτωση η στοίχιση έγινε βάσει της πρώτης έλικας του μονομερούς A (αριστερά). Δεν φαίνεται να υπάρχει σημαντική διαφορά ανάμεσα στις δύο μέσες δομές. Η μεγαλύτερη απόκλιση μεταξύ τους φαίνεται να υπάρχει στην περιοχή της στροφής του μονομερούς A.



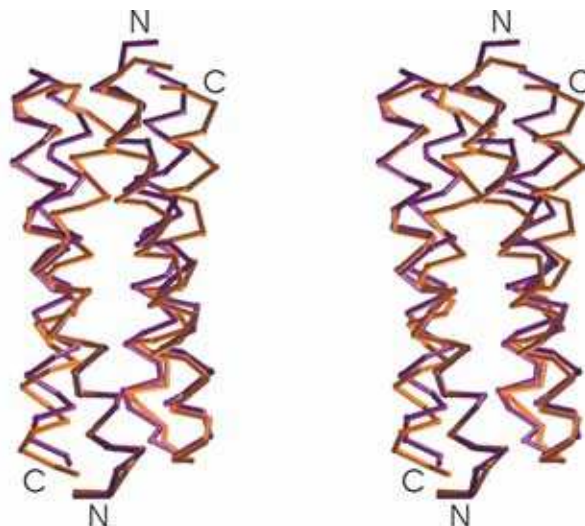
Εικόνα 4.27 Stereo σύγκριση της μέσης δομής του πρώτου cluster (απεικονίζεται με μπλε χρώμα) με τη μέση δομή του δεύτερου cluster (απεικονίζεται με πορτοκαλί χρώμα).

Η μέση δομή του πρώτου cluster φαίνεται να είναι αρκετά κοντά στη δομή από την οποία ξεκίνησε, κάτι που είναι αναμενόμενο βάσει των μέχρι τώρα αποτελεσμάτων. Μικρές αποκλίσεις συναντάμε στα καρβοξυτελικά άκρα και των δύο μονομερών.



Εικόνα 4.28 Stereo σύγκριση της μέσης δομής του πρώτου cluster (απεικονίζεται με μπλε χρώμα) με την αρχική δομή του trajectory (απεικονίζεται με μωβ χρώμα).

Οι πιο εμφανείς διαφορές εμφανίζονται ανάμεσα στη μέση δομή του δεύτερου cluster και την αρχική δομή του trajectory. Το γεγονός αυτό εξηγεί γιατί βλέπουμε το σκαλοπάτι στην τιμή του rmsd και ακολούθως μικρή αύξησή του. Είναι εμφανής τόσο η αλλαγή στη διαμόρφωση της στροφής του μονομερούς A, όσο και η απόκλιση της δεύτερης έλικας του μονομερούς B, ιδιαίτερα προς το καρβοξυτελικό άκρο.

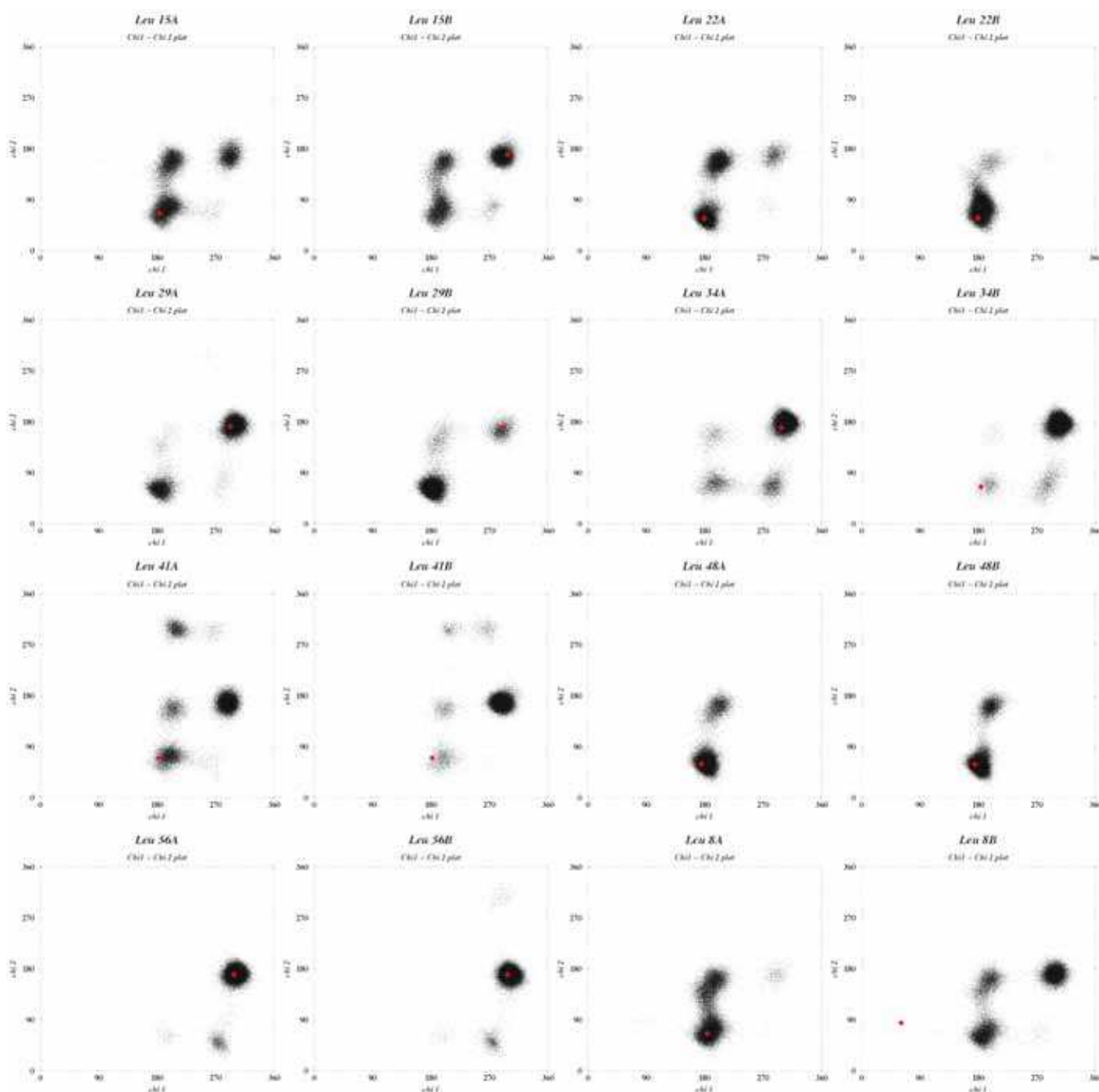


Εικόνα 4.29 Stereo σύγκριση της μέσης δομής του δεύτερου cluster (απεικονίζεται με πορτοκαλί χρώμα) με την αρχική δομή του trajectory (απεικονίζεται με μωβ χρώμα).

Οι εικόνες αυτές εξηγούν σε ικανοποιητικό βαθμό τις δομικές αλλαγές που συμβαίνουν στο μόριο και στις οποίες οφείλονται οι διακυμάνσεις των τιμών των παραμέτρων που μελετήθηκαν μέχρι τώρα. Για την αποσαφήνιση της ακριβούς φύσης των αλλαγών αυτών θα πρέπει να γίνει σε βάθος δομική μελέτη κατά την οποία, η προσοχή μας θα επικεντρωθεί σε συγκεκριμένα κατάλοιπα, τις αλληλεπιδράσεις που αναπτύσσονται και πως αυτές μεταβάλλονται.

4.1.10 Ανάλυση διεδρων γωνιών (χ -value)

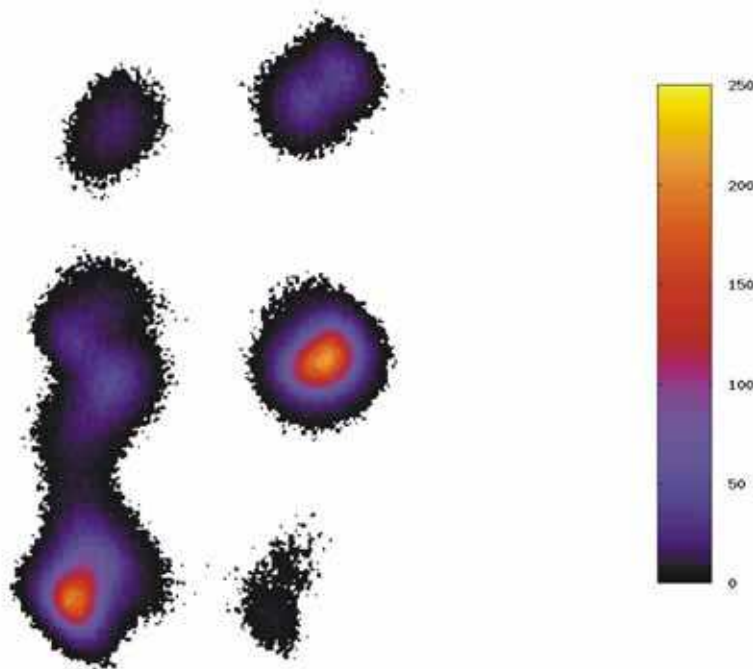
Το πρώτο βήμα σε μία τέτοια ανάλυση, όπως αναφέρθηκε στην αρχή του Κεφαλαίου είναι ο υπολογισμός κάθε διεδρης γωνίας των πλευρικών ομάδων για ολόκληρο το trajectory. Ένας τρόπος απεικόνισης των αποτελεσμάτων είναι σε ένα διδιάστατο γράφημα, όπου ο οριζόντιος άξονας θα αντιστοιχεί στις τιμές της χ_1 γωνίας και ο κάθετος στις τιμές της χ_2 γωνίας.



Εικόνα 4.30 Κατανομή των χ_1 και χ_2 διεδρων γωνιών των καταλοίπων λευκίνης του υδρόφοβου πυρήνα για ολόκληρο το trajectory. Με κόκκινο ρόμβο σημειώνεται το rotamer από το οποίο ξεκίνησε το κατάλοιπο.

Στην **Εικόνα 4.30** παρουσιάζεται η κατανομή των χ_1 και χ_2 γωνιών για όλα τα κατάλοιπα λευκίνης που συνιστούν τον υδρόφοβο πυρήνα της πρωτεΐνης. Με τον τρόπο αυτόν δημιουργούνται clusters (ομάδες) από συνδυασμούς τιμών για τις γωνίες χ_1 και χ_2 που αντιστοιχούν στα κυρίαρχα rotamers.

Καταρχήν παρατηρούμε πως τα κυρίαρχα rotamers που προκύπτουν από τις προσομοιώσεις μοριακής δυναμικής συμφωνούν με τα rotamers που έχουν προσδιορισθεί κρυσταλλογραφικά (Εικόνα 4.1). Το γεγονός αυτό συνδράμει σε σημαντικό βαθμό στην αξιοπιστία των αποτελεσμάτων που προκύπτουν από τις προσομοιώσεις. Για λόγους σύγκρισης, στην **Εικόνα 4.31** παρουσιάζεται η κατανομή των τιμών των δίεδρων γωνιών για όλα τα κατάλοιπα λευκίνης αθροιστικά (Σύγκριση με Εικόνα 4.1).



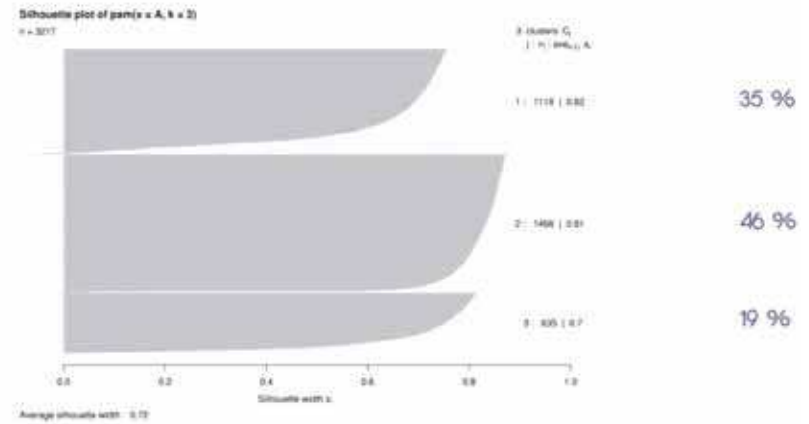
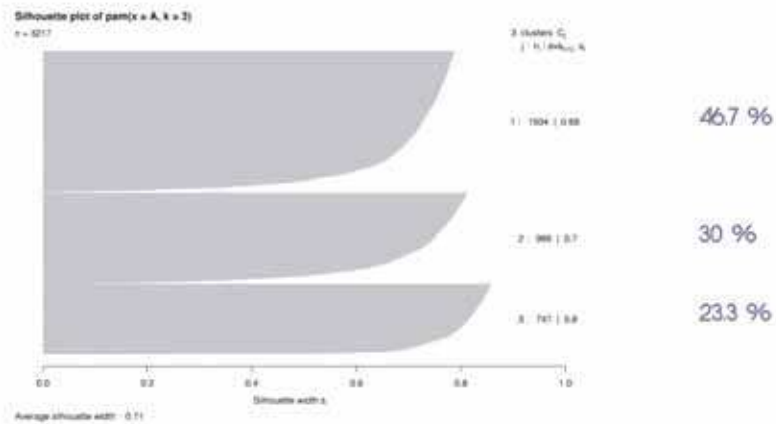
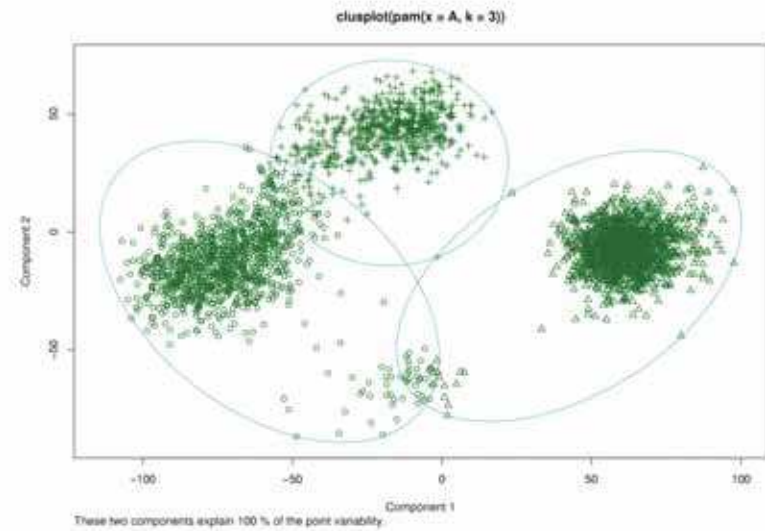
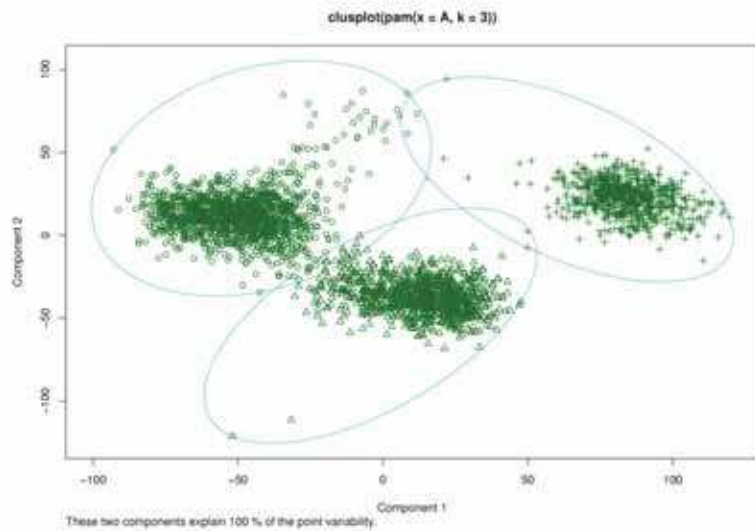
Εικόνα 4.31 Αθροιστική κατανομή των τιμών των δίεδρων χ_1 και χ_2 γωνιών για τα κατάλοιπα λευκίνης του υδρόφοβου πυρήνα όπως προέκυψαν από τις προσομοιώσεις.

Είναι αξιοσημείωτο το γεγονός ότι τα κατάλοιπα δεν «παγιδεύτηκαν» στο rotamer από το οποίο ξεκίνησαν (κόκκινος ρόμβος). Έτσι υπάρχουν περιπτώσεις που ενώ ένα κατάλοιπο ξεκίνησε από το πιο συχνό rotamer, πέρασε και σε ένα λιγότερο συχνό (όπως τα κατάλοιπα 15A, 15B, 22A, 22B, 29A, 34A, 48A, 48B, 56A, 56B, 8A). Παρατηρείται όμως και το αντίστροφο φαινόμενο, δηλαδή ενώ ξεκίνησε από λιγότερο συχνό rotamer, αυτό δεν το εμπόδισε να περάσει και στο πιο συχνό (όπως τα κατάλοιπα 29B, 34B, 41A, 41B). Τέλος, συναντήσαμε και την ακραία περίπτωση, στο κατάλοιπο 8B, όπου το αρχικό rotamer ήταν ένα από τα πιο σπάνια που έχουν

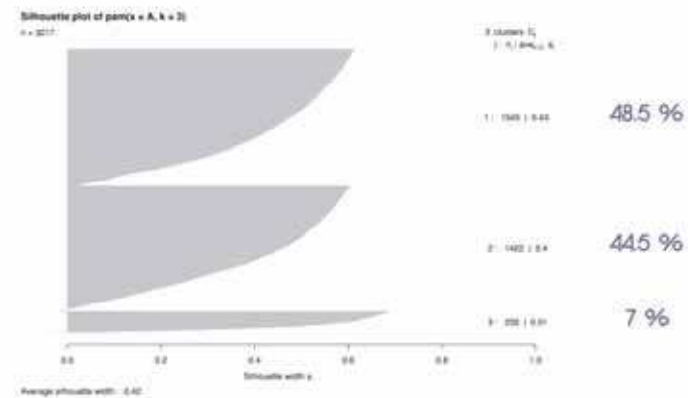
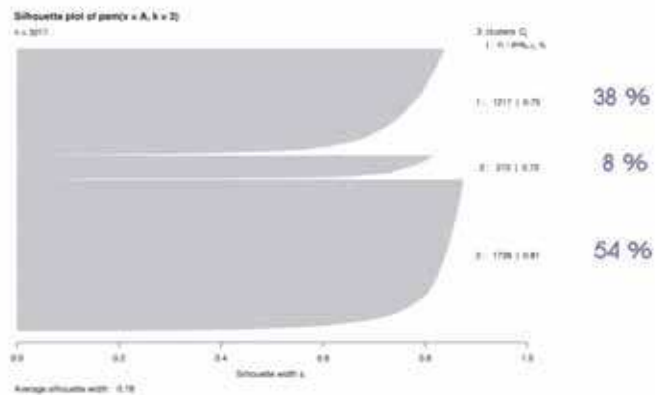
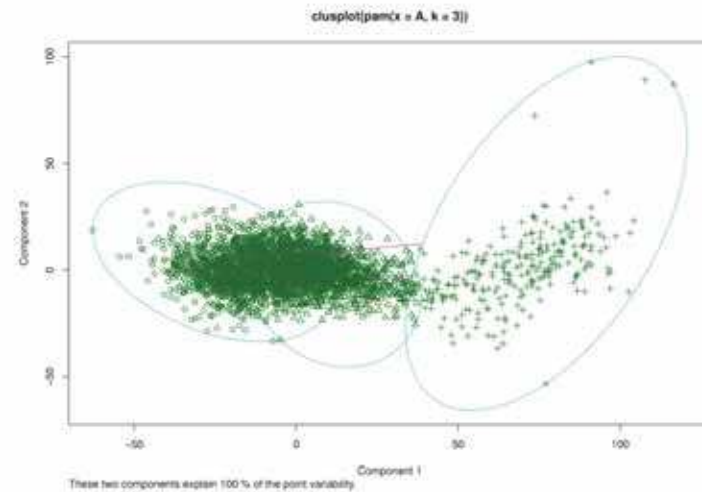
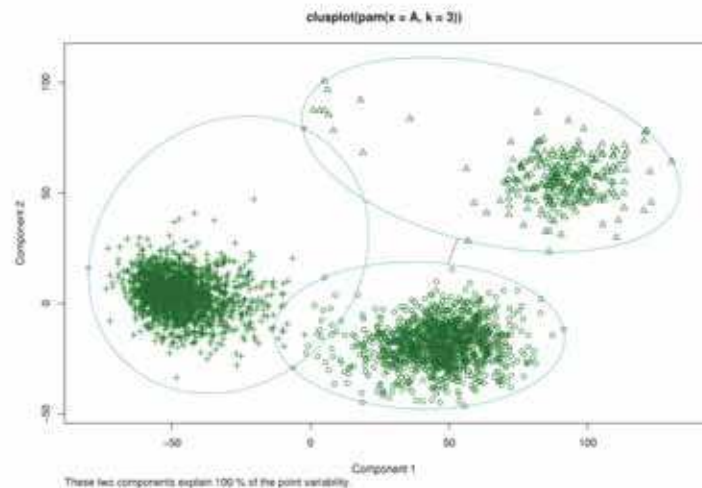
παρατηρηθεί. Παρόλα αυτά, στη συνέχεια κατά τη διάρκεια του trajectory παρατηρήθηκαν μόνο τα συνήθη rotamers.

Αξιοσημείωτο είναι το γεγονός πως παρατηρείται πανομοιότυπο πρότυπο στην κατανομή των τιμών για τα ίδια κατάλοιπα στο κάθε μονομερές, ενδεχομένως λόγω συμμετρίας, αφού πρόκειται για ομοδιμερές. Δηλαδή βλέπουμε πως, για παράδειγμα, για το κατάλοιπο 15 στο μονομερές A παρατηρούνται τα ίδια rotamers με το κατάλοιπο 15 στο μονομερές B. Στο σημείο αυτό να υπενθυμίσουμε πως στη native (Rop-like) τοπολογία τα κατάλοιπα αυτά δεν συμμετέχουν στο σχηματισμό του ίδιου επιπέδου (layer). Η λευκίνη 15A συμμετέχει στο επίπεδο L5, ενώ η λευκίνη 15B συμμετέχει στο επίπεδο L4 (Εικόνα 3.7). Τα επίπεδα αυτά είναι γειτονικά και συνεπώς υπάρχει αλληλεπίδραση ανάμεσά τους. Όμως το φαινόμενο αυτό είναι γενικευμένο και αφορά όλα τα κατάλοιπα λευκίνης του πυρήνα. Έτσι παρατηρείται το ίδιο πρότυπο ακόμα για τα κατάλοιπα 56A - 56B, και 8A - 8B, παρόλο που το κατάλοιπο 56A συμμετέχει στο επίπεδο L8, το κατάλοιπο 56B συμμετέχει στο επίπεδο L1, το κατάλοιπο 8A συμμετέχει στο επίπεδο L7 και το κατάλοιπο 8B συμμετέχει στο επίπεδο L2, που είναι ιδιαίτερα απομακρυσμένα. Μάλιστα, το φαινόμενο είναι τόσο ισχυρό που ακόμα και αν το αρχικό rotamer του καταλοίπου στο ένα και το άλλο μονομερές ήταν διαφορετικό, αυτό δεν είχε καμία επίδραση και εμφανίστηκε το ίδιο πρότυπο (για παράδειγμα τα κατάλοιπα 15A - 15B, 34A - 34B, 8A - 8B). Τίθεται λοιπόν το ερώτημα γιατί τα κατάλοιπα αυτά συμπεριφέρονται με τον ίδιο τρόπο; Άραγε το γεγονός αυτό έχει συνέπειες στην αναδίπλωση του μορίου; Διαδραματίζει κάποιο ρόλο στην δραματική αύξηση του χρόνου αναδίπλωσης και αποδιάταξης που εμφανίζει η $[Ala_2Leu_2]_8$ σε σχέση με τη φυσική Rop;

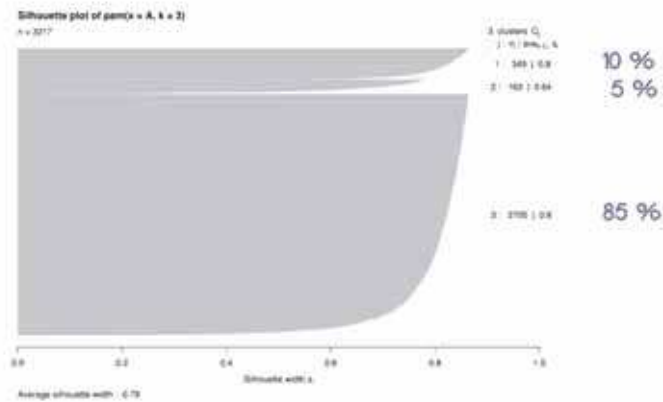
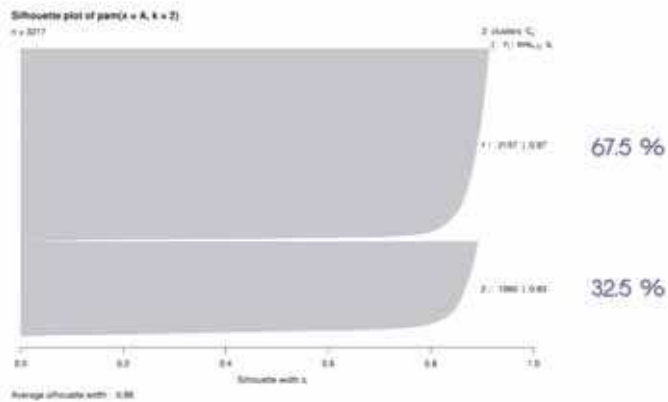
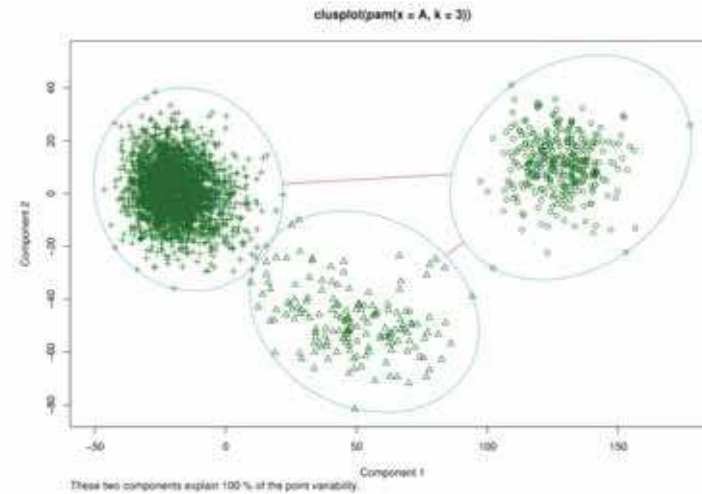
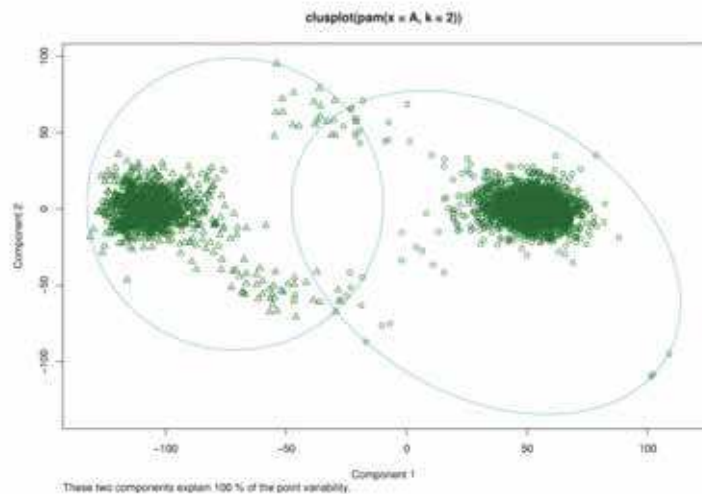
Στις **Εικόνες 4.32 – 4.39** που ακολουθούν περιλαμβάνεται μία στατιστική ανάλυση των τιμών των διέδρων χ_1 και χ_2 γωνιών του κάθε καταλοίπου, ανάλογη με αυτή που παρουσιάστηκε ήδη στην Ενότητα 4.2.8 κατά την ανάλυση των cluster δομών που συναντούνται σε ένα trajectory. Στην κάθε εικόνα παρατίθενται τα αποτελέσματα για το ίδιο κατάλοιπο και για τα δύο μονομερή (αριστερά για το μονομερές A και δεξιά για το μονομερές B), για λόγους σύγκρισης με όσα προαναφέρθηκαν. Επίσης αναφέρονται τα ποσοστά που προκύπτουν για το κάθε cluster. Έτσι έχουμε μία άποψη για τον χρόνο τον οποίο πέρασε το κάθε κατάλοιπο σε κάθε rotamer, το οποίο αντιπροσωπεύεται από ένα cluster (residence times). Τα ποσοστά αυτά θα μπορούσαν να συμβάλλουν σε μία πρόβλεψη για το rotamer το οποίο αναμένεται να παρατηρηθεί κρυσταλλογραφικά ή με φασματοσκοπία πυρηνικού μαγνητικού συντονισμού (NMR, Nuclear Magnetic Resonance spectroscopy).



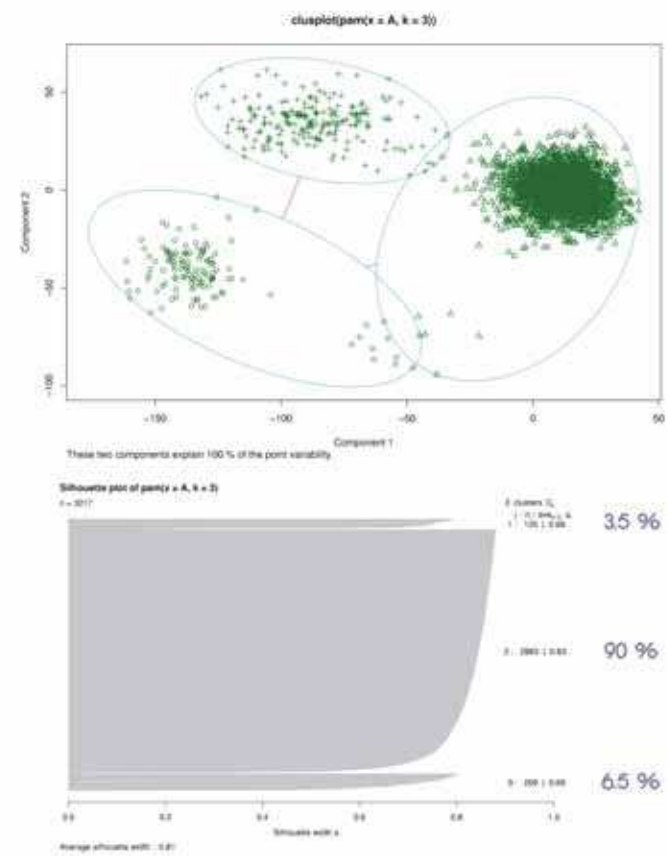
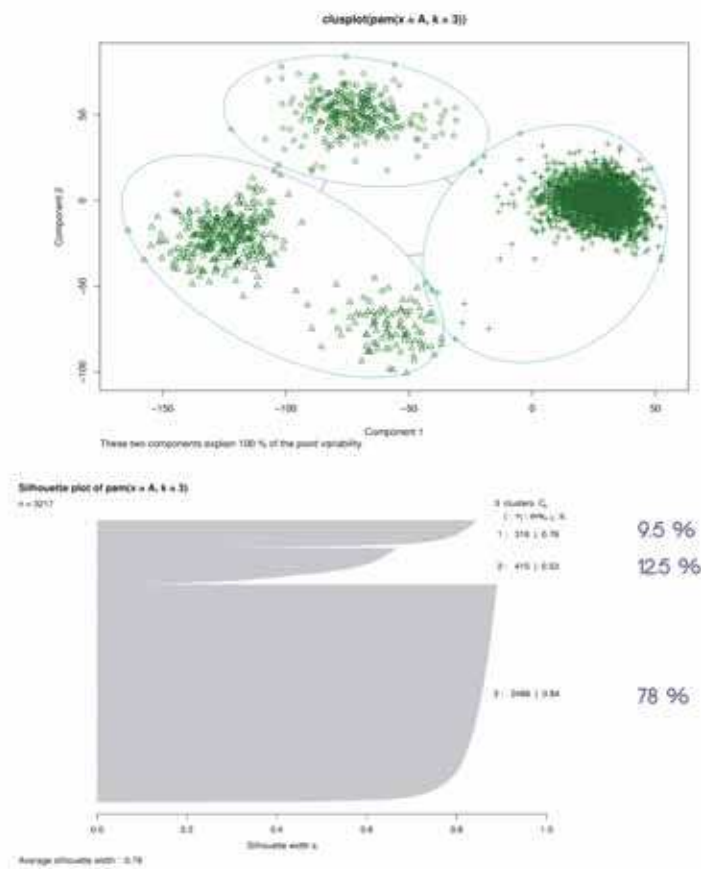
Εικόνα 4.32 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 15A και δεξιά για το κατάλοιπο 15B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 15A και δεξιά για το κατάλοιπο 15B.



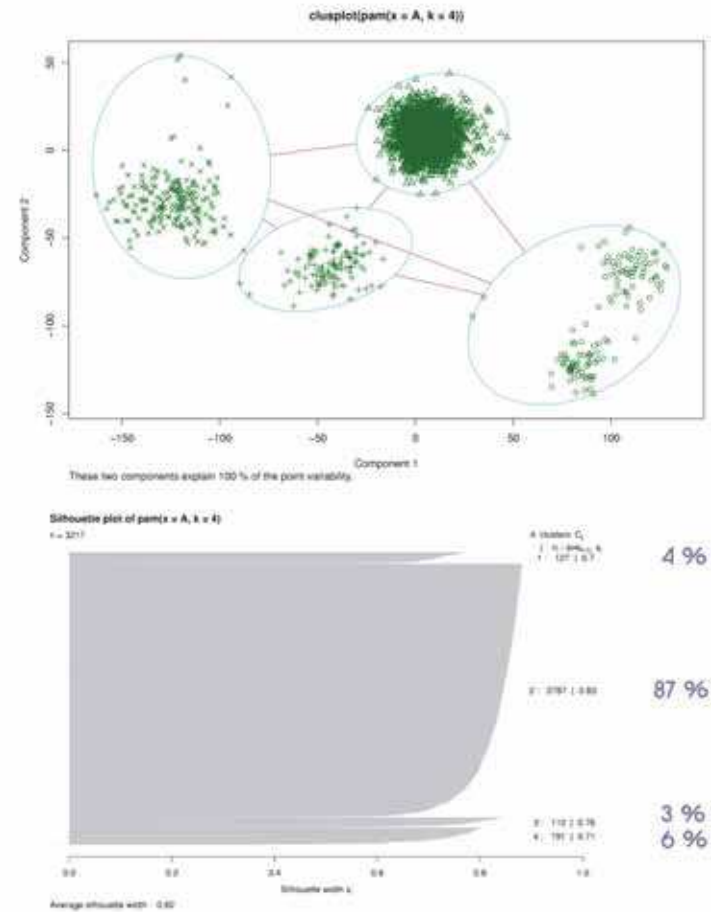
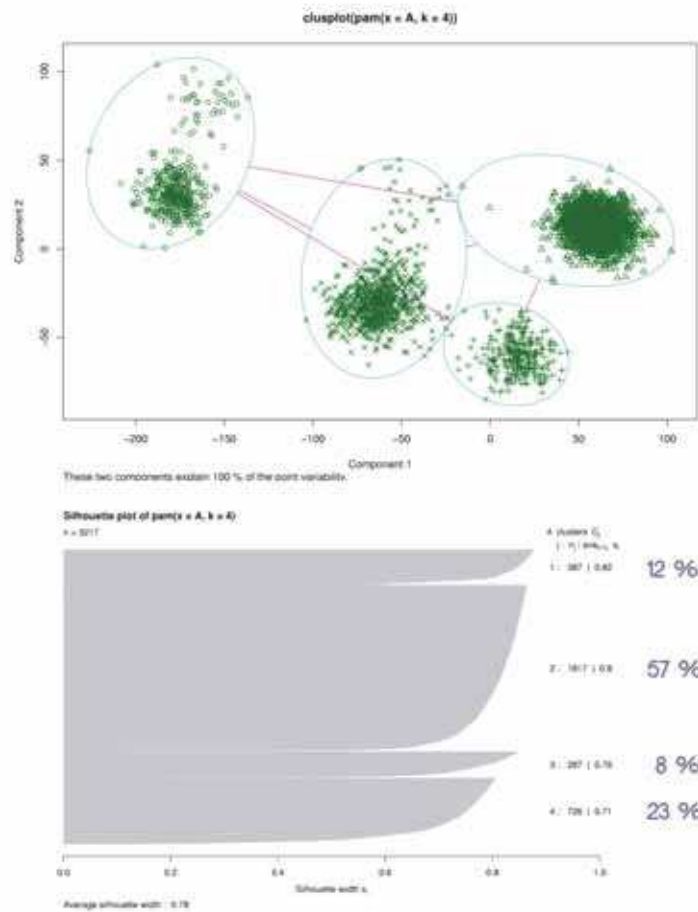
Εικόνα 4.33 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, *αριστερά* για το κατάλοιπο 22A και *δεξιά* για το κατάλοιπο 22B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, *αριστερά* για το κατάλοιπο 22A και *δεξιά* για το κατάλοιπο 22B.



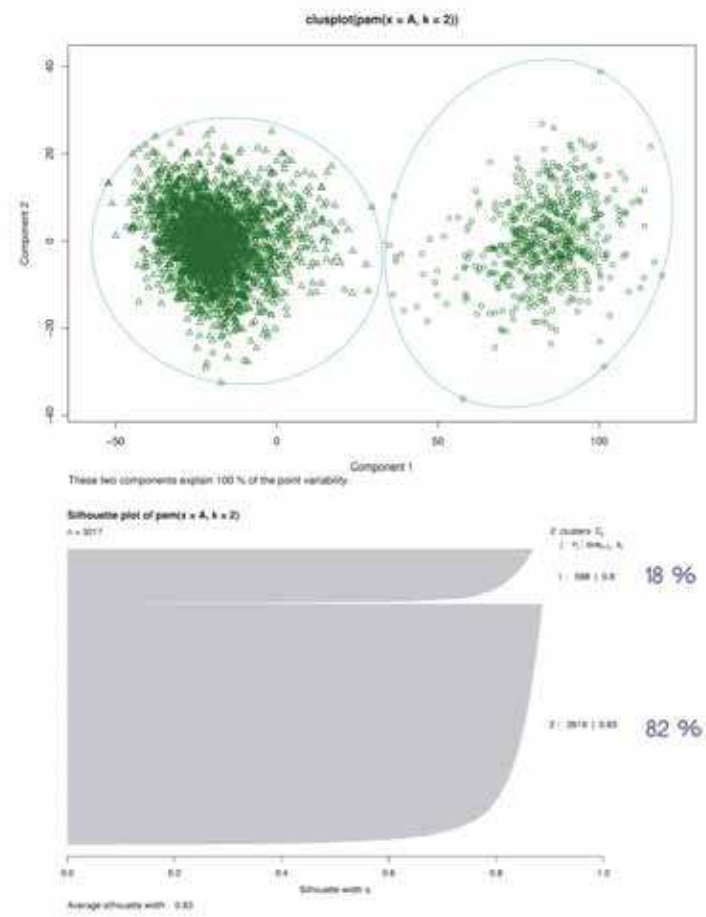
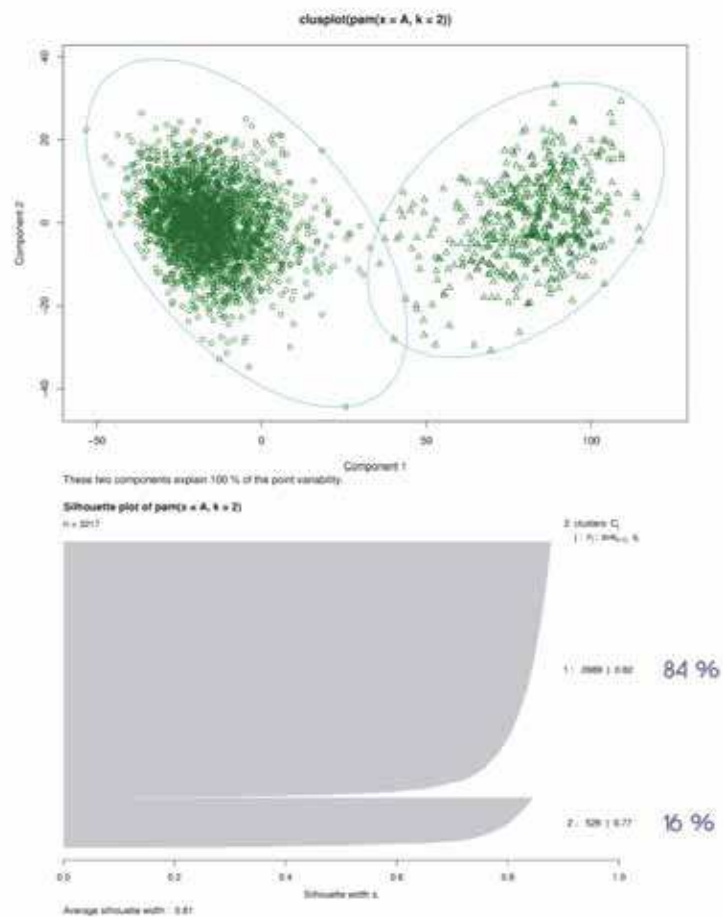
Εικόνα 4.34 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 29A και δεξιά για το κατάλοιπο 29B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 29A και δεξιά για το κατάλοιπο 29B.



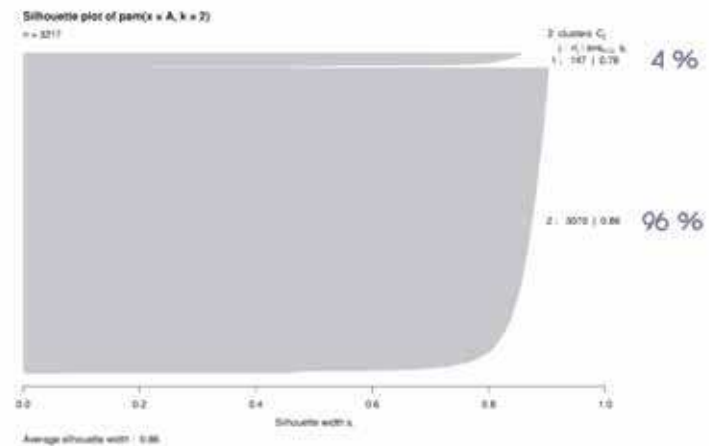
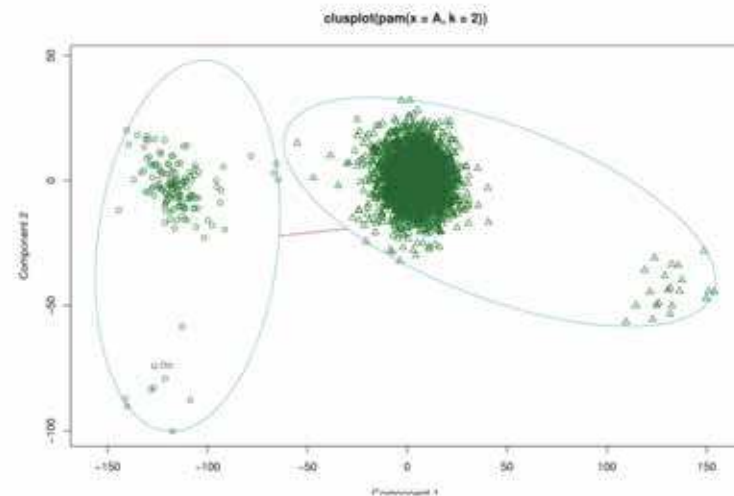
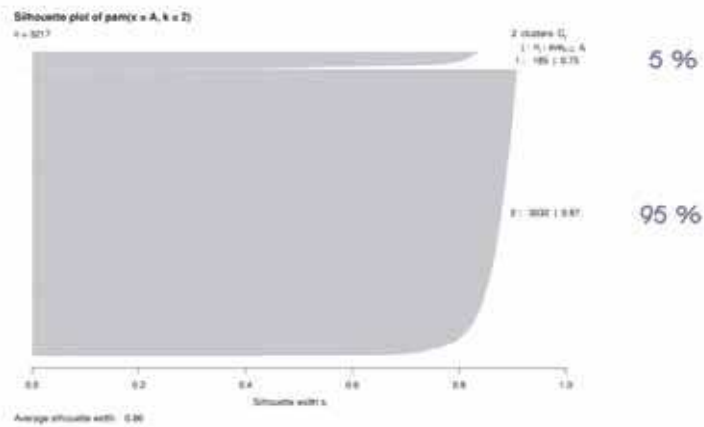
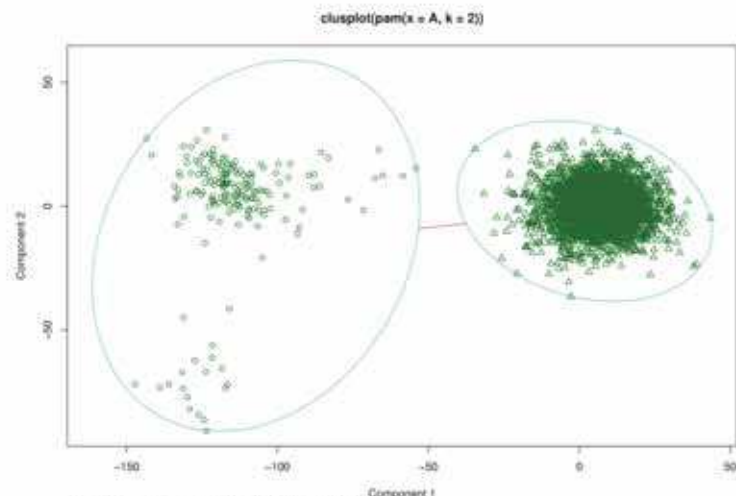
Εικόνα 4.35 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 34A και δεξιά για το κατάλοιπο 34B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 34A και δεξιά για το κατάλοιπο 34B.



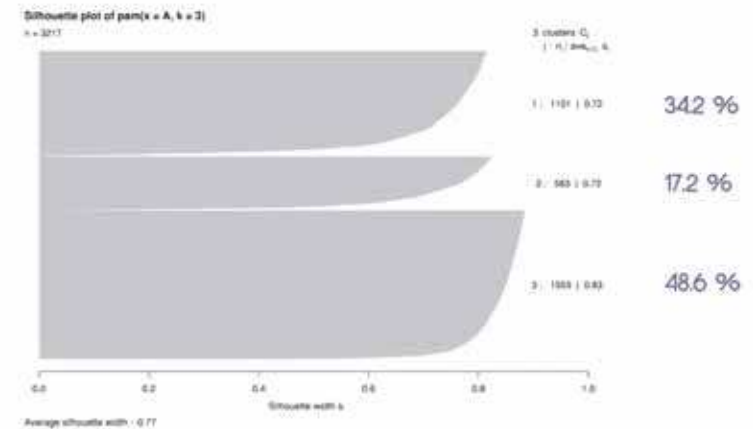
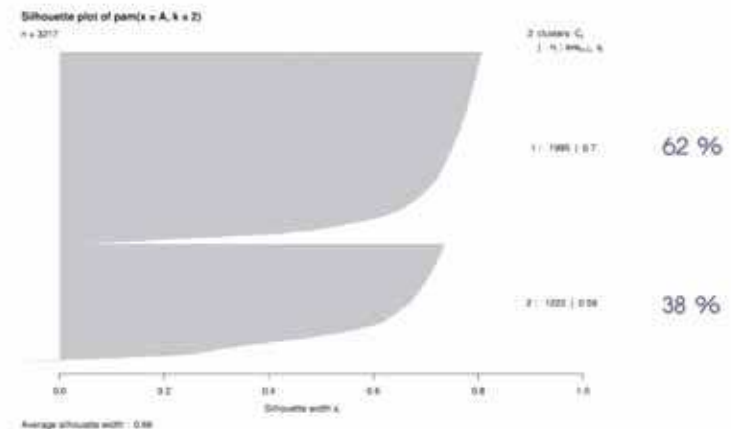
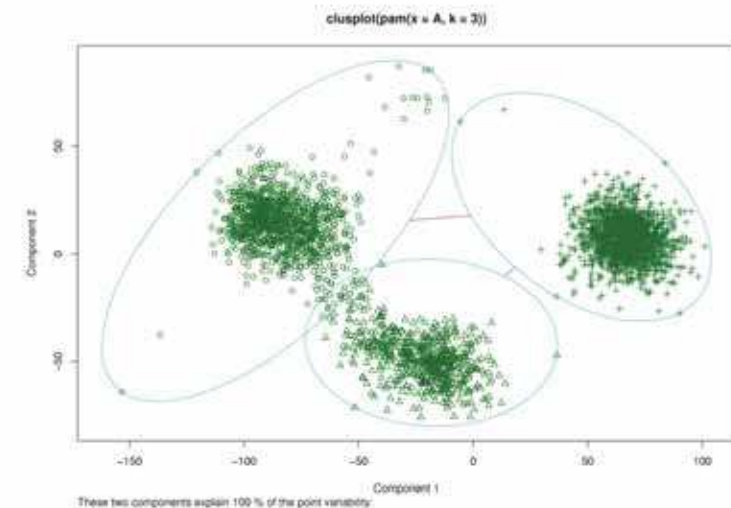
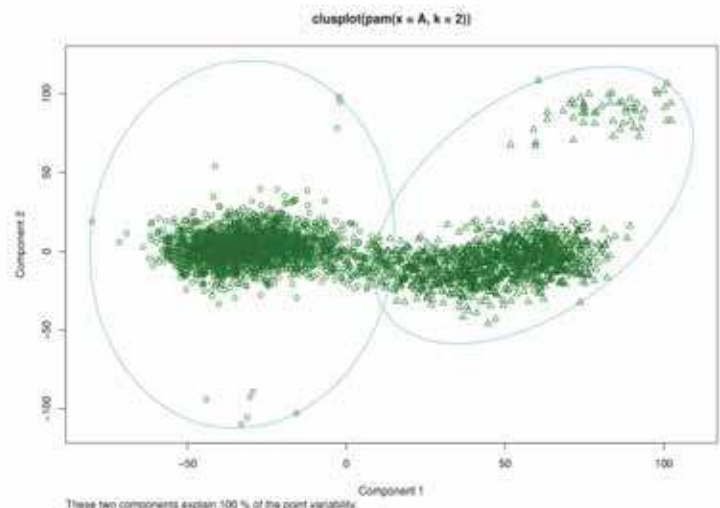
Εικόνα 4.36 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 41A και δεξιά για το κατάλοιπο 41B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 41A και δεξιά για το κατάλοιπο 41B.



Εικόνα 4.37 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, *αριστερά* για το κατάλογο 48A και *δεξιά* για το κατάλογο 48B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, *αριστερά* για το κατάλογο 48A και *δεξιά* για το κατάλογο 48B.



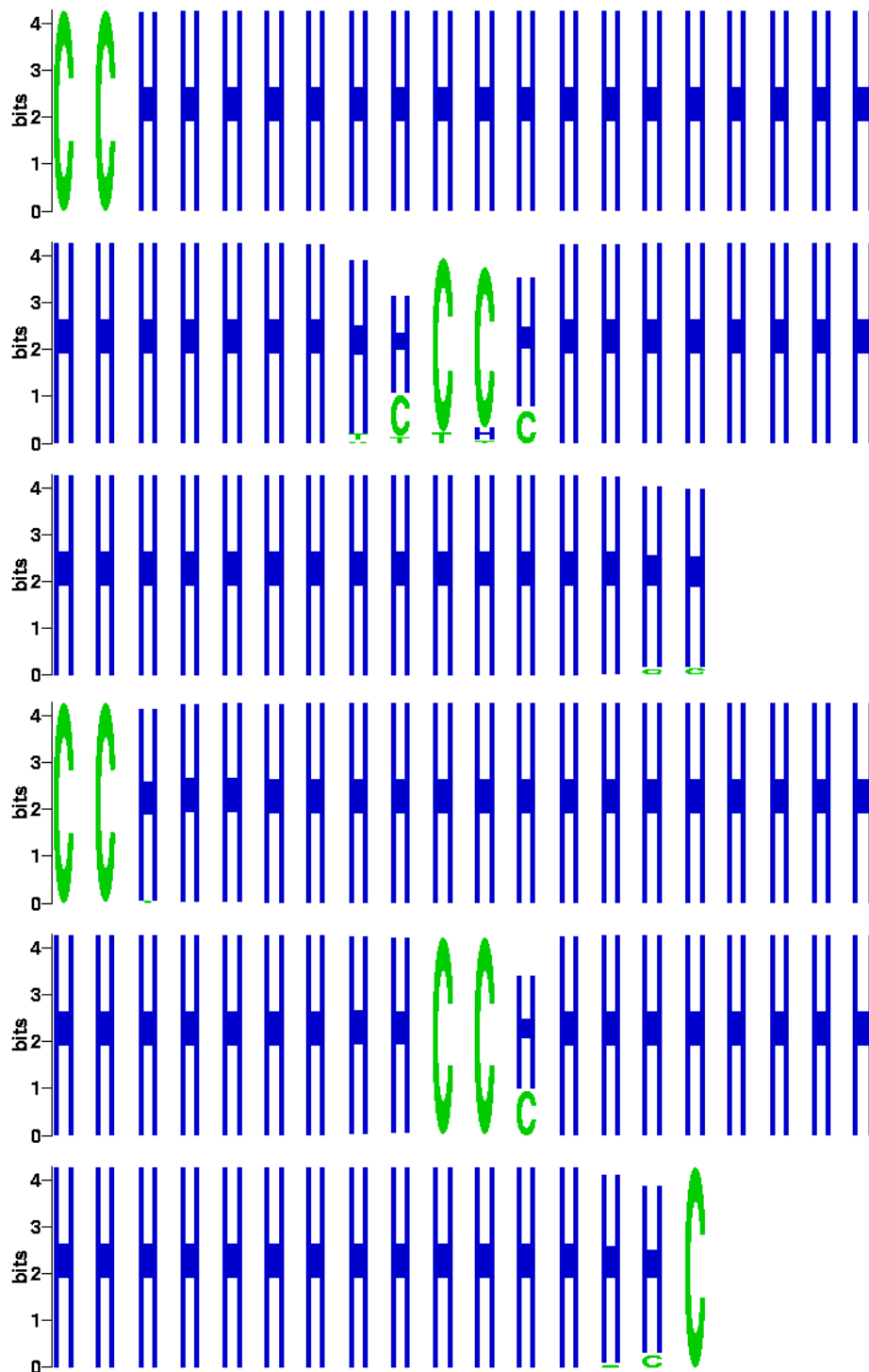
Εικόνα 4.38 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 56A και δεξιά για το κατάλοιπο 56B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 56A και δεξιά για το κατάλοιπο 56B.



Εικόνα 4.39 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 8A και δεξιά για το κατάλοιπο 8B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 8A και δεξιά για το κατάλοιπο 8B.

4.1.11 Ανάλυση δευτεροταγούς δομής

Εκτός από τις διεδρες γωνίες χ_1 και χ_2 , ενδιαφέρον παρουσιάζει και η μελέτη των ϕ και ψ γωνιών (Ramachandran plots). Μία ϕ γωνία ορίζεται με τα άτομα C_{n-1} -N-C α -C και μία ψ με τα άτομα N-C α -C- N_{n+1} . Επομένως (όπως προκύπτει και από τα άτομα από τα οποία ορίζονται) οι γωνίες αυτές καθορίζουν το πώς συνδέονται τα κατάλοιπα.



Εικόνα 4.40 Μεταβολή της δευτεροταγούς δομής για κάθε θέση της αλληλουχίας.

Η διακύμανση των τιμών τους είναι ενδεικτική της διατήρησης της δευτεροταγούς δομής. Για να αποφύγουμε μία ανάλυση σαν αυτή που περιγράφηκε για τις χ_1 και χ_2 γωνίες στην Ενότητα 4.2.9, που θα χρειαζόταν να υπολογίσουμε ρ_{hi} και ρ_{si} γωνίες για κάθε κατάλοιπο, μπορούμε εναλλακτικά να δούμε κατά πόσο διατηρούνται οι έλικες κατά τη διάρκεια του trajectory.

Μία τέτοια αναπαράσταση με χρήση logo (Shaner *et al.*, 2003) παρουσιάζεται στην **Εικόνα 4.40**, όπου με H (Helix) αναπαρίστανται οι α -έλικες και με C (Coils) οι ελεύθερες θηλιές. Κάθε σειρά περιλαμβάνει 20 θέσεις της αλληλουχίας. Βλέπουμε πως οι δομές έλικας διατηρούνται κατά τη διάρκεια ολόκληρου του trajectory και επομένως είναι αρκετά σταθερές. Τα κατάλοιπα 1, 2 και 56 του μονομερούς A και του μονομερούς B δεν σχηματίζουν σταθερή δομή. Η περιοχή της στροφής (κατάλοιπα 28-32) του μονομερούς A υιοθετεί δομή θηλιάς όπως και η αντίστοιχη περιοχή του μονομερούς B (κατάλοιπα 30-32). Επομένως η περιοχή της στροφής του μονομερούς A είναι περισσότερο αποδιαταγμένη (πέντε κατάλοιπα έχουν τη δομή θηλιάς) και αυτό εξηγεί το μεγαλύτερο rmsd του μονομερούς A σε σχέση με το μονομερές B (Εικόνα 4.3). Το γεγονός ότι η στροφή του ενός μονομερούς συμπεριφέρεται διαφορετικά από την στροφή του άλλου μονομερούς προξενεί εντύπωση δεδομένου ότι υπάρχει συμμετρία (το μόριο είναι ομοδιμερές).

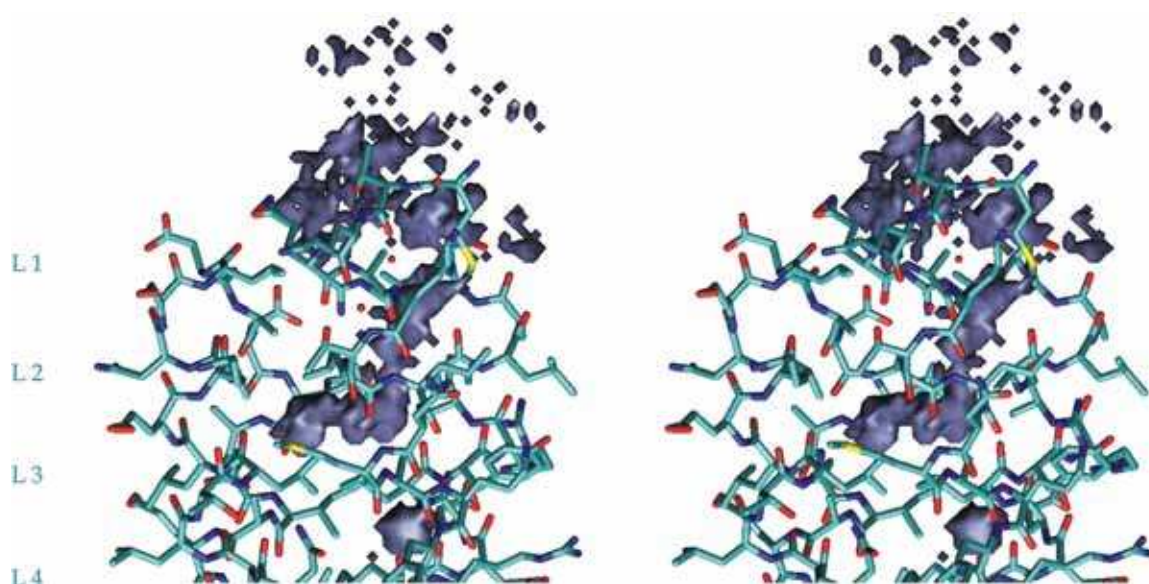
4.1.12 Υπολογισμός της κατανομής των μορίων νερού (calculation of water distribution)

Κατά την προετοιμασία του συστήματος της $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία υπάρχει ένα στάδιο στο οποίο η πρωτεΐνη ενυδατώνεται (solvate) με μόρια νερού και ιόντα (Ενότητα 3.3). Η προσθήκη των μορίων νερού γίνεται με τυχαίο τρόπο ώστε να δημιουργηθεί ένα κουτί συγκεκριμένων διαστάσεων. Λόγω της ύπαρξης μίας κοιλότητας (cavity) στη φυσική πρωτεΐνη και άρα στην $[Ala_2Leu_2]_8$ βρέθηκε ότι τοποθετήθηκαν τρία μόρια νερού σε αυτήν την κοιλότητα. Να σημειωθεί στο σημείο αυτό ότι δεδομένου της συμμετρίας του μορίου είναι αναμενόμενο η κοιλότητα αυτή να παρατηρείται και στα επίπεδα L5-L7. Ωστόσο δεν βρέθηκαν μόρια νερού εκεί.

Έχουν βρεθεί τέτοιες κοιλότητες σε αρκετές σφαιρικές πρωτεΐνες (globular proteins), εκ των οποίων κάποιες είναι κενές ενώ άλλες περιέχουν κρυσταλλογραφικά προσδιορισμένα μόρια νερού. Η πιθανότητα να περιέχουν νερά αυξάνεται με το μέγεθος της κοιλότητας (Rashin *et al.*, 1986). Η παρουσία των νερών σε έναν πρωτεϊνικό πυρήνα προωθεί τη σταθερότητα ενυδατώνοντας τα πολικά άτομα. Ειδικότερα, τα καλά διατηρημένα νερά βρίσκονται συνήθως σε περιοχές χωρίς

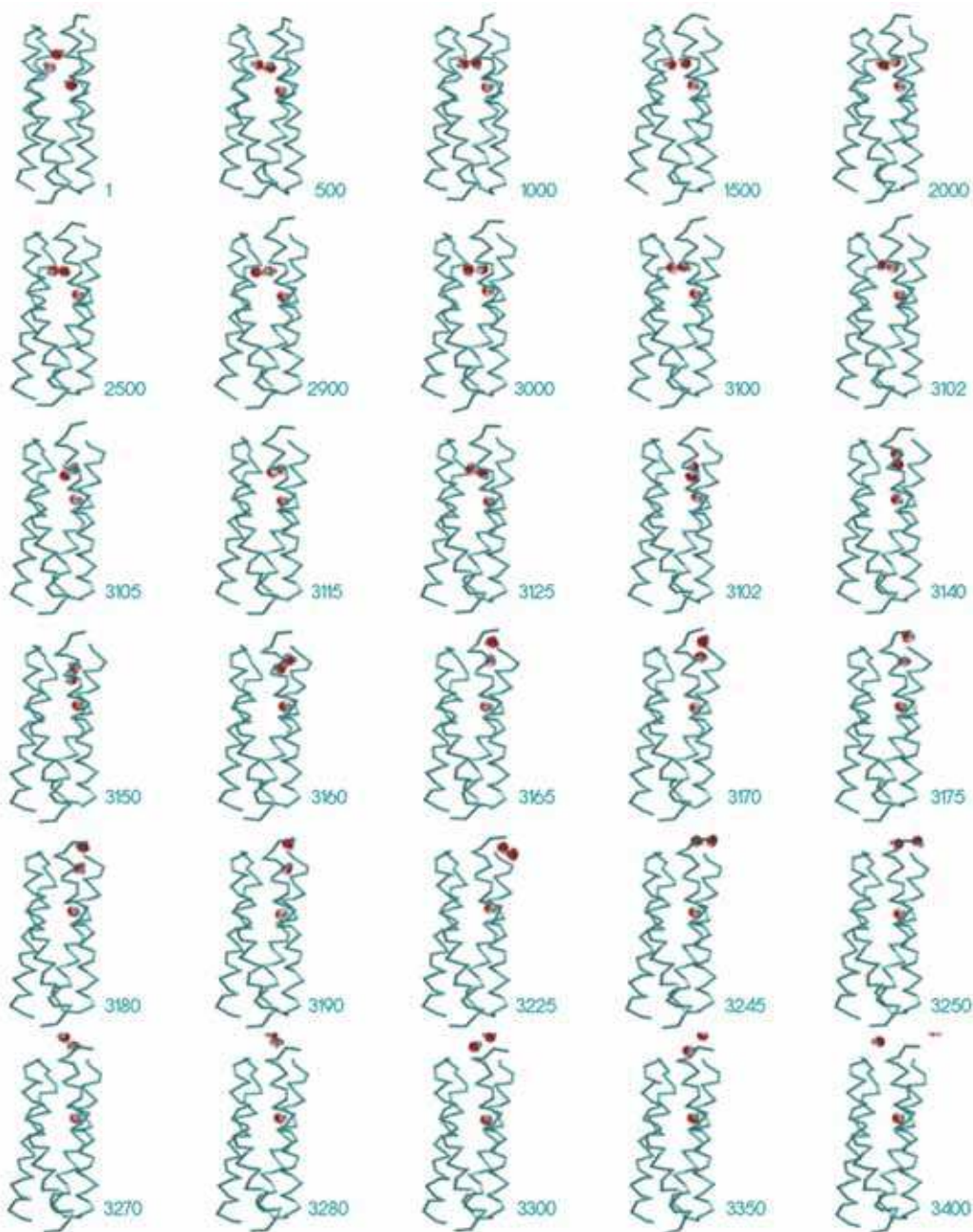
δευτεροταγή δομή και αναπτύσσουν δεσμούς υδρογόνου με άτομα της κύριας αλυσίδας που διαφορετικά δεν δημιουργούν δεσμούς υδρογόνου, ενώ η πλειοψηφία των πολικών πλευρικών αλυσίδων που δημιουργούν δεσμούς υδρογόνου με αυτά τα θαμμένα νερά (buried waters) σχηματίζουν και ενδομοριακούς δεσμούς υδρογόνου. Παρόλα αυτά δεν αποκλείεται η πιθανότητα εύρεσης μορίων νερού σε ένα μη πολικό περιβάλλον, όπου μπορεί να διατηρούνται χωρίς να αναπτύσσουν σταθερούς δεσμούς υδρογόνου. Πράγματι έχουν βρεθεί τέτοια ελαττώματα στο πακετάρισμα (“packing defects”) στο εσωτερικό κάποιων πρωτεϊνών, αρκετά μεγάλα ώστε να μπορούν να φιλοξενήσουν ένα ή περισσότερα μόρια νερού (Park *et al.*, 2005). Κάτι τέτοιο φαίνεται να συμβαίνει και σε αυτήν την περίπτωση.

Θελήσαμε λοιπόν να μελετήσουμε την συμπεριφορά αυτών των μορίων νερού κατά τη διάρκεια του trajectory. Ένας τρόπος για να γίνει αυτό είναι δημιουργώντας κάτι ανάλογο με ένα χάρτη ηλεκτρονικής πυκνότητας (Reddy *et al.*, 2003). Ένας τέτοιος χάρτης αντί για πυκνότητα ηλεκτρονίων θα δείχνει κατανομή μορίων νερού (hydration pattern). Στην **Εικόνα 4.41** βλέπουμε ένα τέτοιο χάρτη. Το πλεονέκτημά του είναι ότι δείχνει τόσο την πορεία που ακολούθησαν τα μόρια νερού, όσο και το χρόνο που παρέμειναν στη κάθε θέση. Η αρχική θέση των μορίων νερού ήταν, ένα μεταξύ των επιπέδων L4 και L3 και δύο μεταξύ των επιπέδων L3 και L2. Τα δύο αυτά μόρια νερού φεύγουν σχετικά γρήγορα, ενώ το πρώτο παραμένει μέχρι το τέλος του χρόνου της προσομοίωσης. Κατά την έξοδό τους παραμένουν και ανάμεσα στα επίπεδα L2 και L1 και για λίγο περισσότερο χρόνο στο επίπεδο L1.



Εικόνα 4.41 Χάρτης της κατανομής των τριών μορίων νερού της κοιλότητας κατά τη διάρκεια του trajectory.

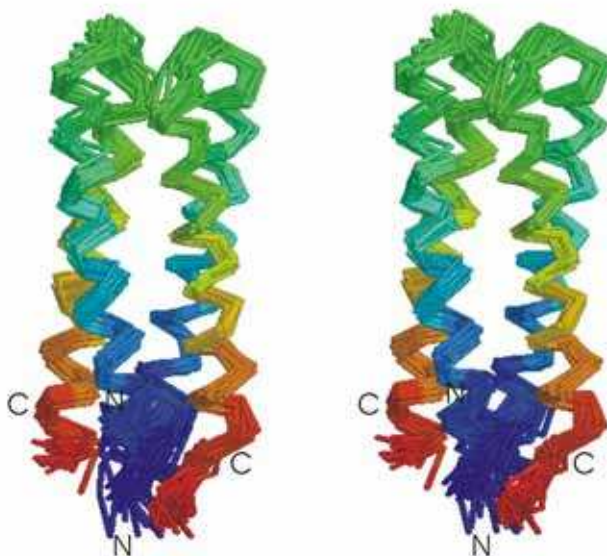
Στην **Εικόνα 4.42** παρατίθενται στιγμιότυπα του trajectory που δείχνουν την κίνηση των μορίων νερού και περιλαμβάνονται τα frame 1-3400 (δηλαδή μέχρι 1.36ns). Τα στιγμιότυπα επιλέχθηκαν κατά τρόπο ώστε να αναπαρασταθεί όσο το δυνατόν καλύτερα η κίνηση των μορίων νερού. Δίπλα στο κάθε στιγμιότυπο αναφέρεται ο ακριβής χρόνος του (σε frame).



Εικόνα 4.42 Στιγμιότυπα (ο χρόνος αναφέρεται σε frame) του trajectory για την κίνηση των μορίων νερού του cavity.

4.2 Ανάλυση προσομοίωσης διμερούς [Ala₂Leu₂]₈ με *syn* τοπολογία

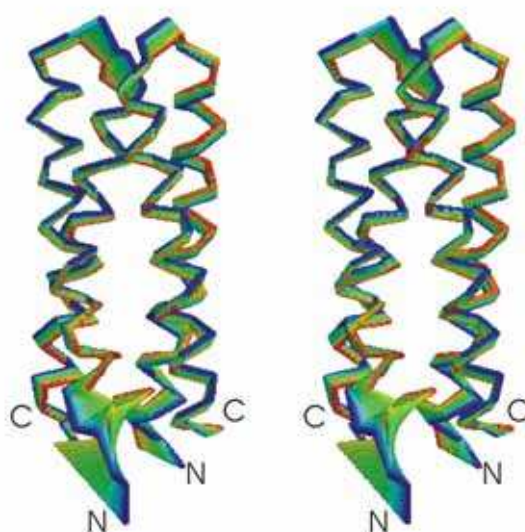
Όπως είδαμε στην Ενότητα 3.4 για τη δημιουργία (υπολογιστικά) του διμερούς [Ala₂Leu₂]₈ με *syn* τοπολογία χρησιμοποιήθηκε σαν αρχική δομή, η δομή μίας παραλλαγής της Rop, της [Ala₂-Ile₂]₆ (PDB entry 1F4N και 1F4M). Για το μόριο αυτό έχουν προσδιορισθεί οι ατομικές συντεταγμένες για τα κατάλοιπα 1-60 για το A μονομερές και για τα κατάλοιπα 2-57 για το B μονομερές (βλ. Πίνακα 3.2). Αυτό σημαίνει πως υπάρχουν επιπλέον κατάλοιπα στα αμινοτελικά και καρβοξυτελικά άκρα του μορίου τα οποία δεν συμμετέχουν στο σχηματισμό του υδρόφοβου πυρήνα και ως εκ τούτου σε μία σταθερή δομή αλλά βρίσκονται εκτεθειμένα προς το διάλυμα με τη μορφή θηλιάς (loop). Το γεγονός αυτό μπορεί να επηρεάσει σημαντικά την αξιοπιστία των αποτελεσμάτων που προκύπτουν κατά την ανάλυση της προσομοίωσης του συστήματος αυτού. Ο λόγος είναι πως σε περίπτωση που το μόριο διαθέτει ιδιαίτερα ευκίνητες ουρές, θα αλλοιώνονται τα αποτελέσματα εάν ληφθούν και αυτές υπόψη κατά τους υπολογισμούς. Η κίνηση όμως των ουρών δεν περιέχει κάποια πληροφορία που να μας ενδιαφέρει. Όταν για παράδειγμα υπολογίζουμε μία μέση δομή για το μόριο ο συνυπολογισμός των καταλοίπων στις ουρές θα οδηγήσει σε διαφορετικό αποτέλεσμα. Αυτό όμως που επιθυμούμε εμείς είναι να προσεγγίσουμε μία μέση δομή του μορίου βασιζόμενοι στο τμήμα που σχηματίζει το σκελετό της δομής του μορίου και όχι στο τμήμα που πραγματοποιεί τυχαία κίνηση στο διάλυμα.



Εικόνα 4.43 Stereo άποψη της υπέρθεσης των στιγμιότυπων του trajectory του μορίου κάθε 500 frame. Αναπαρίστανται μόνο ο σκελετός του μορίου. Η διαβάθμιση του χρώματος από μπλε σε κόκκινο ακολουθεί από το N-τελικό προς το C-τελικό άκρο.

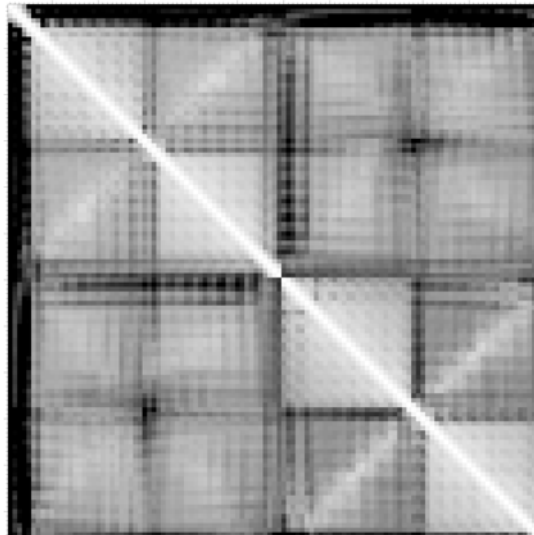
Προκειμένου να εξεταστεί εάν τα κατάλοιπα στα άκρα πραγματοποιούν μία τυχαία και άτακτη κίνηση εξετάσαμε το trajectory που ακολούθησε το μόριο. Στην **Εικόνα 4.43** βλέπουμε σε υπέρθεση στιγμιότυπα του trajectory. Είναι εμφανές ότι τα κατάλοιπα στα αμινοτελικά άκρα τόσο του ενός όσο και του άλλου μονομερούς κινούνται σε μεγάλο βαθμό σε σχέση με το υπόλοιπο μόριο αλλά και με τα κατάλοιπα στα καρβοξυτελικά άκρα.

Η κίνηση αυτή είναι πιο εμφανής εάν παρακολουθήσουμε την κίνηση που πραγματοποιεί το μόριο βάση του πρώτου eigenvector (**Εικόνα 4.44**). Ο υπολογισμός έγινε για ολόκληρο το trajectory. Η διαβάθμιση του χρώματος από μπλε σε κόκκινο φανερώνει το εύρος της κίνησης. Εδώ φαίνεται καθαρά η διαφορά στο εύρος της κίνησης των καταλοίπων στα N-τελικά άκρα.

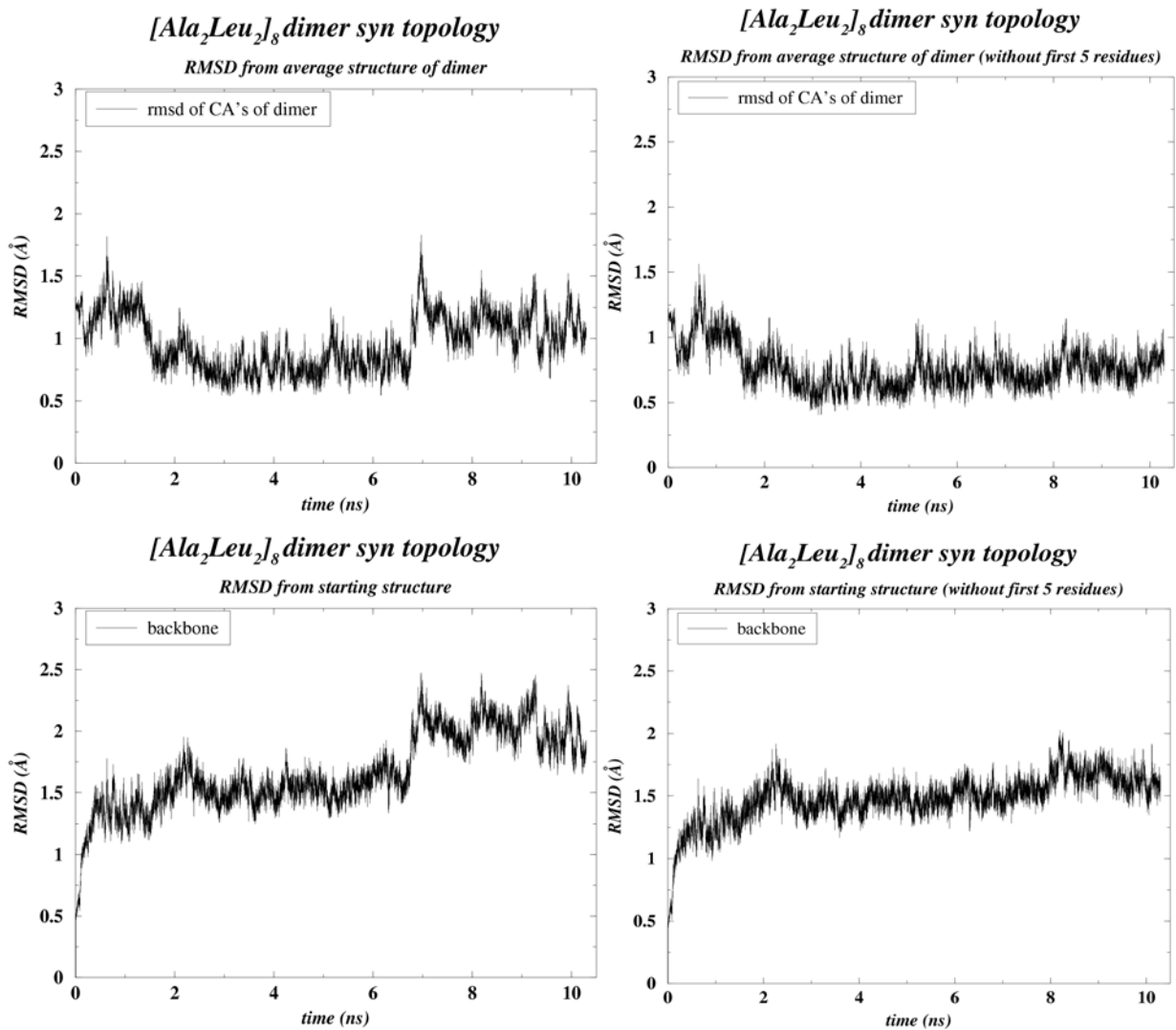


Εικόνα 4.44 Stereo αναπαράσταση της κίνησης του μορίου που περιγράφεται από το πρώτο ζεύγος eigenvalue – eigenvector λαμβάνοντας υπόψη όλα τα κατάλοιπα.

Είναι εύλογο να υποθέσει κανείς ότι τα κατάλοιπα τα οποία θα πρέπει να μη ληφθούν υπόψη στους υπολογισμούς είναι τα κατάλοιπα στα αμινοτελικά άκρα που προηγούνται του πρώτου επιπέδου που συμμετέχει στο σχηματισμό του υδρόφοβου πυρήνα. Να υπενθυμίσουμε στο σημείο αυτό, ότι επειδή η διεύθυνση των δύο μονομερών σε αυτή τη δομή είναι διαφορετική, το επίπεδο εδώ σχηματίζεται από τα κατάλοιπα 5A, 5B, 56A, 56B (βλ. Εικόνα 3.11). Ένας πιο αξιόπιστος τρόπος για να δούμε ποια κατάλοιπα κινούνται πολύ θα ήταν βάσει του rmsd που παρουσιάζουν σε σχέση με τα άλλα. Στην **Εικόνα 4.45** βλέπουμε μία γραφική αναπαράσταση των rmsd που αντιστοιχούν στις αποστάσεις των Ca (Ca-Ca distance map), όπου όπως προαναφέρθηκε, οι μικρότερες τιμές rmsd φαίνονται με λευκό και οι μεγαλύτερες με μαύρο. Η μεγαλύτερη τιμή rmsd που βρέθηκε ήταν 3.29 Å. Κατά συνέπεια, τα ακραία κατάλοιπα που είναι ιδιαίτερα ευκίνητα είναι τα πρώτα πέντε στο αμινοτελικό άκρο κάθε μονομερούς.



Εικόνα 4.45 rmsd που αντιστοιχεί στον Cα-Cα distance map για την $[Ala_2Leu_2]_8$ με syn τοπολογία, λαμβάνοντας υπόψη όλα τα κατάλοιπα.

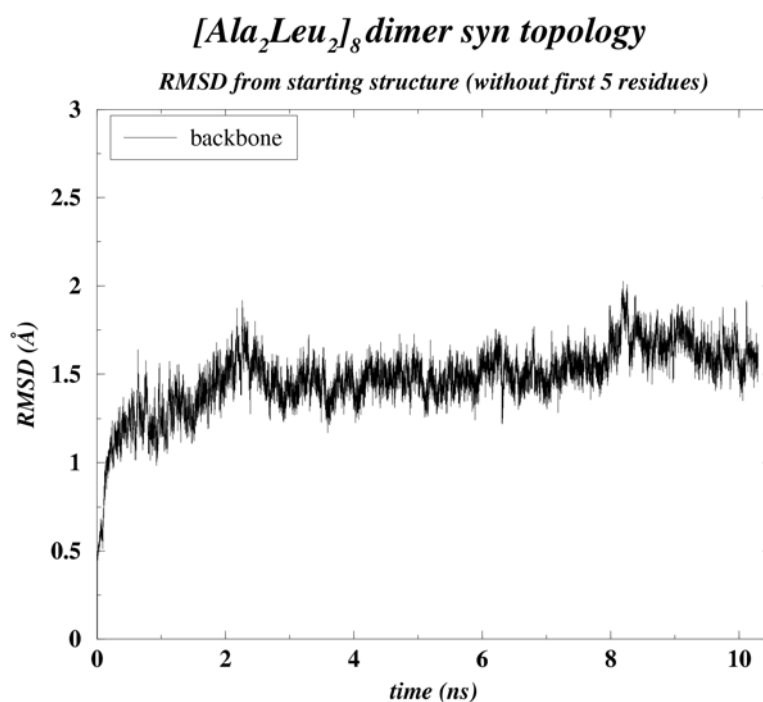


Εικόνα 4.46

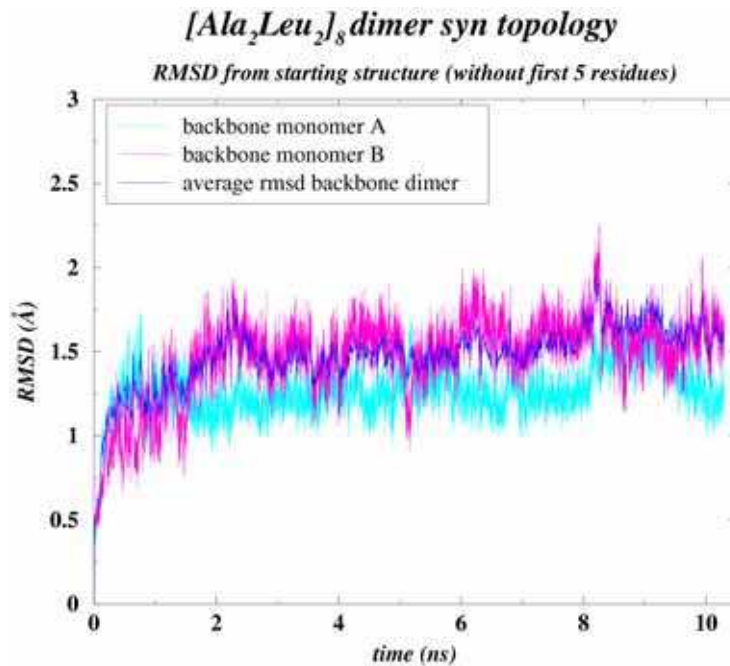
Στην **Εικόνα 4.46** παρατίθενται ενδεικτικά το rmsd από την αρχική και μέση δομή του μορίου στο συγκεκριμένο σύστημα. Αριστερά ο υπολογισμός έγινε για όλα τα κατάλοιπα του συστήματος ενώ δεξιά δεν λήφθηκαν υπόψη τα πέντε πρώτα κατάλοιπα κάθε μονομερούς. Βλέπουμε χαρακτηριστικά την επίδραση που έχει ο συνυπολογισμός των ευκίνητων καταλοίπων. Προκειμένου λοιπόν να διασφαλίσουμε κατά το δυνατόν, την αξιοπιστία των αναλύσεων, τα κατάλοιπα αυτά δεν συμπεριλήφθηκαν στους υπολογισμούς που περιγράφονται στη συνέχεια (τα κατάλοιπα δεν αφαιρέθηκαν από το μόριο, απλά δεν λαμβάνονται υπόψη).

4.2.1 Rmsd από την αρχική δομή

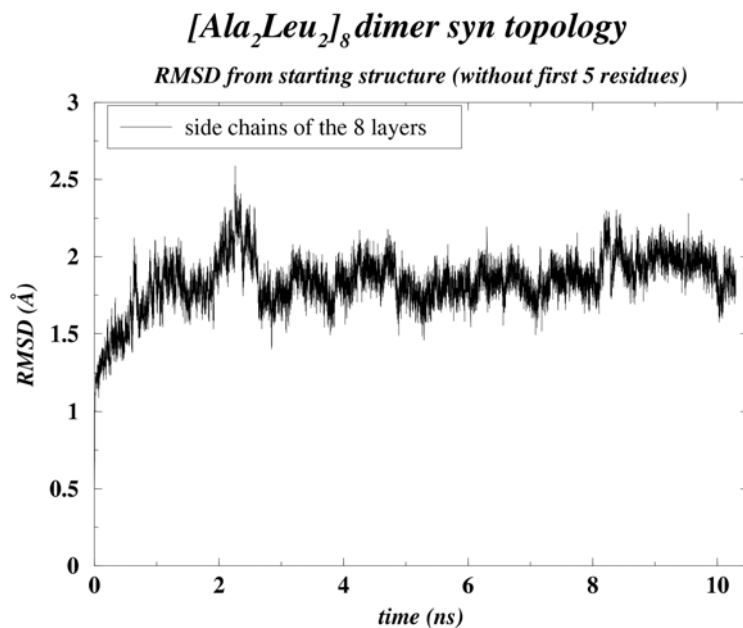
Στην **Εικόνα 4.47** παρουσιάζεται το rmsd λαμβάνοντας υπόψη μόνο τα άτομα του σκελετού (backbone) του μορίου. Εδώ το σύστημα αργεί λίγο περισσότερο να σταθεροποιηθεί και μετά την πάροδο της φάσης του heating εμφανίζεται ένα σκαλοπάτι. Ακολούθως παραμένει γύρω στα 1.5 Å για το μεγαλύτερο μέρος του χρόνου της προσομοίωσης (περίπου από τα 2.5ns μέχρι και τα 8ns). Στη συνέχεια εμφανίζεται και ένα δεύτερο σκαλοπάτι μετά το οποίο το rmsd κυμαίνεται μεταξύ 1.6 Å και 1.7 Å. Είναι πιθανό και σε αυτήν την περίπτωση το μόριο να υιοθετεί περισσότερες από μία κύριες δομές, κάτι το οποίο θα εξεταστεί στη συνέχεια. Ωστόσο, εδώ οι αλλαγές που συμβαίνουν είναι περισσότερες αλλά λιγότερο ευδιάκριτες, κάτι που μπορεί να σημαίνει μία πιο σταδιακή μετάβαση.



Εικόνα 4.47

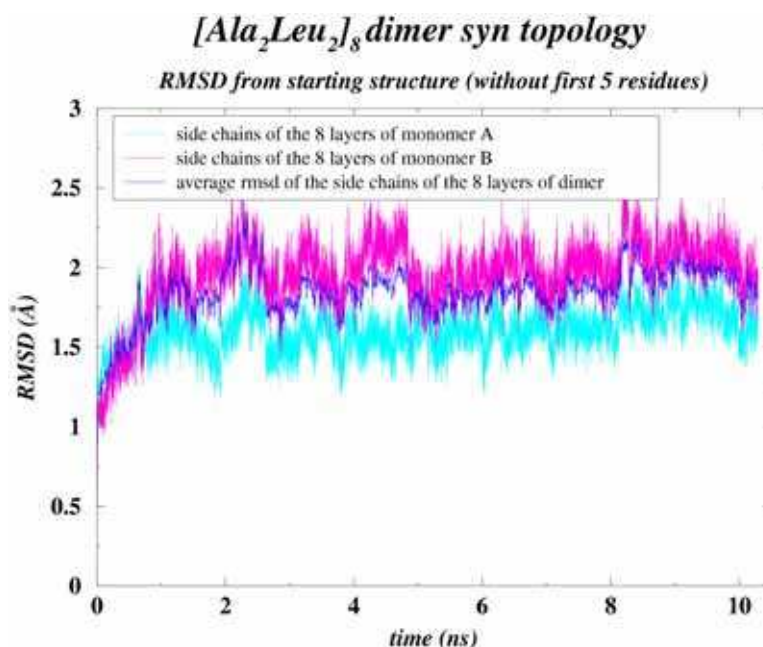


Οι αλλαγές στη δομή εδώ οφείλονται σε μεγαλύτερο βαθμό σε κινητικότητα του μονομερούς B, όπως προκύπτει από την **Εικόνα 4.48**. Το μονομερές A παρουσιάζει μικρότερη κινητικότητα κατά τη διάρκεια όλης της προσομοίωσης, σε αντίθεση με το μονομερές B το οποίο φαίνεται να είναι πιο ασταθές, με μεγαλύτερες διακυμάνσεις, η τιμή rmsd του οποίου είναι γύρω στα 1.6-1.7 Å. Το μονομερές B έχει γενικά μικρότερες τιμές rmsd. Στο μεγαλύτερο μέρος της προσομοίωσης κυμαίνεται γύρω στα 1.2-1.3 Å και στη συνέχεια (γύρω στα 8ns – 9ns) αυξάνεται στα 1.4 Å.



Όσον αφορά το rmsd των πλευρικών ομάδων (χωρίς να ληφθούν υπόψη τα άτομα υδρογόνου) των καταλοίπων που συνιστούν τα οκτώ επίπεδα του υδρόφοβου πυρήνα της πρωτεΐνης, βλέπουμε στην **Εικόνα 4.49** ότι παρουσιάζει σημαντικές διακυμάνσεις, γεγονός που δεν συνδράμει με ιδιαίτερη σταθερότητα του πυρήνα. Στην αρχή φτάνει μέχρι και τα 2.5 Å, αλλά στη συνέχεια (μετά το ίδιο σκαλοπάτι στα 2ns περίπου) κατεβαίνει στα 1.7-1.8 Å, ενώ διατηρείται και το σκαλοπάτι στα 8ns – 9ns.

Οι πλευρικές ομάδες του κάθε μονομερούς μεμονωμένα επίσης εμφανίζουν διακυμάνσεις κατά τη διάρκεια της προσομοίωσης (**Εικόνα 4.50**). Το μονομερές B εξακολουθεί να είναι περισσότερο ασταθές και όσον αφορά τη συμπεριφορά των πλευρικών ομάδων των καταλοίπων του υδρόφοβου πυρήνα σε σχέση με το μονομερές A. Ενώ οι πλευρικές ομάδες του μονομερούς B εμφανίζουν τιμές rmsd γύρω στα 1.6 Å, οι πλευρικές ομάδες του μονομερούς A εμφανίζουν τιμές rmsd γύρω στα 2 Å.

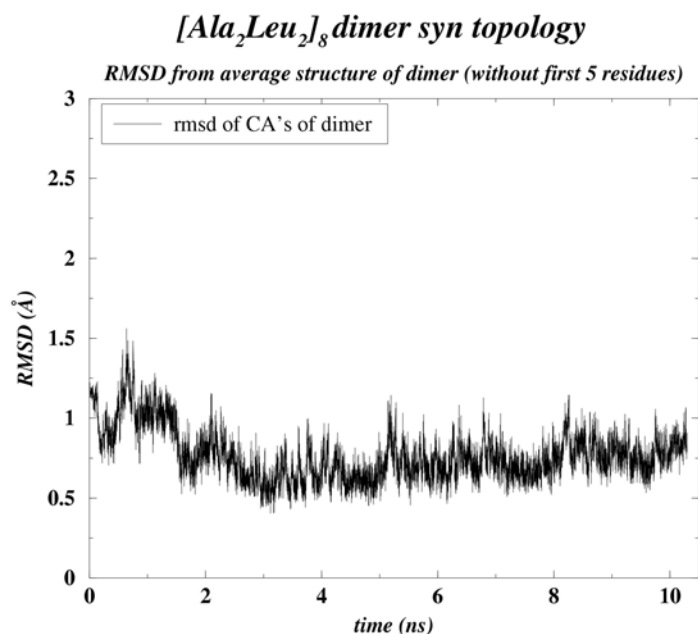


Εικόνα 4.50

4.2.2 Rmsd από τη μέση δομή

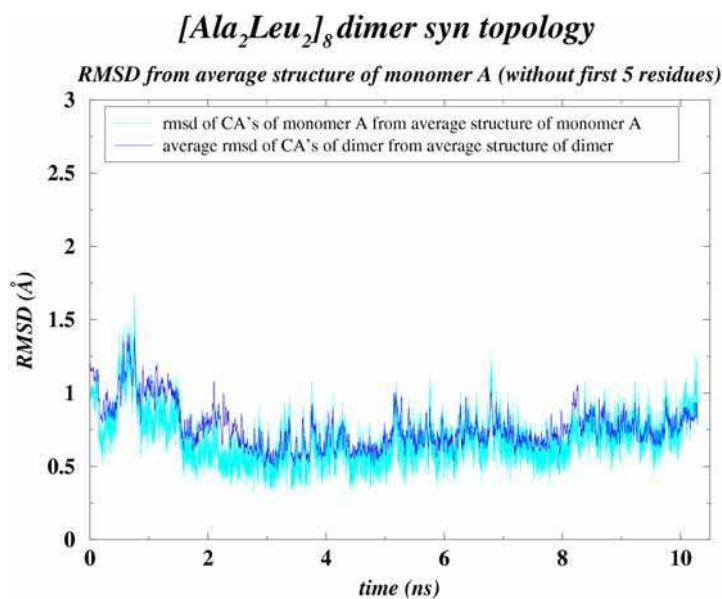
Μετά τον υπολογισμό μίας μέσης δομής, υπολογίζουμε το rmsd που εμφανίζει η δομή σε κάθε στιγμιότυπο (frame) του trajectory από τη μέση δομή. Στην **Εικόνα 4.51** βλέπουμε πως και σε αυτήν την περίπτωση το rmsd εμφανίζει σημαντικές διακυμάνσεις. Ωστόσο διατηρείται σε αρκετά χαμηλές τιμές (0.6Å -0.7Å). Ο

υπολογισμός έγινε λαμβάνοντας υπόψη μόνο τα Ca άτομα του διμερούς, σε σχέση με τη μέση δομή του διμερούς. Οι διακυμάνσεις στα 2ns και 8ns – 9ns συνεχίζουν να διαφαίνονται.



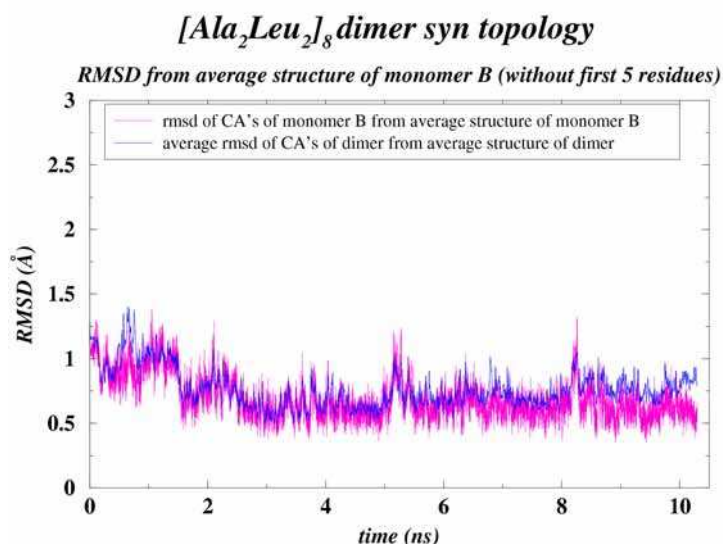
Εικόνα 4.51

Για την περίπτωση του κάθε μονομερούς μεμονωμένα, βλέπουμε στην **Εικόνα 4.52** πως το μονομερές A (λαμβάνοντας υπόψη κατά τον υπολογισμό μόνο τα Ca άτομα) είναι σχετικά σταθερό, με rmsd από τη μέση δομή (του μονομερούς) λίγο μικρότερη από την αντίστοιχη του διμερούς, στα 0.6 Å. Η μέση δομή του κάθε μονομερούς υπολογίστηκε λαμβάνοντας υπόψη μόνο τα άτομα του σκελετού.



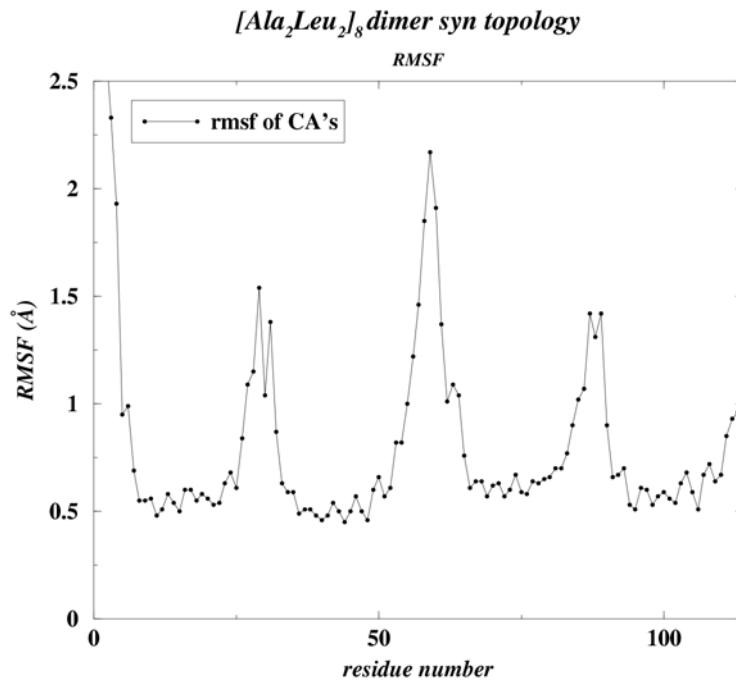
Εικόνα 4.52

Το rmsd του μονομερούς B (λαμβάνοντας υπόψη κατά τον υπολογισμό μόνο τα Ca άτομα) από τη μέση δομή του κυμαίνεται στην ίδια τιμή με αυτήν του μονομερούς A (Εικόνα 4.53). Έτσι, σε γενικές γραμμές το μόριο παρουσιάζει κάποια σταθερότητα, όπως προκύπτει από τις χαμηλές τιμές rmsd από τη μέση δομή. Εμφανίζει όμως μικρές διακυμάνσεις σε όλη τη διάρκεια του χρόνου της προσομοίωσης.



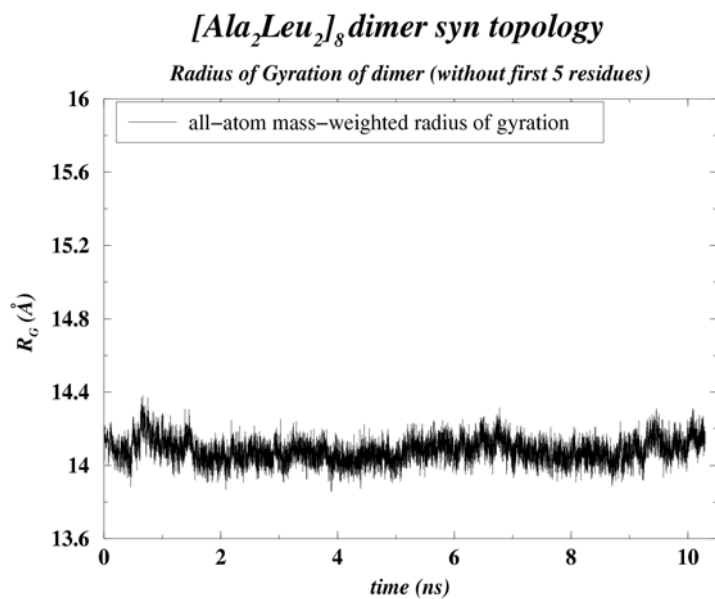
4.2.3 Rmsf

Στην Εικόνα 4.54 αναπαρίσταται το rmsf που εμφανίζει κάθε κατάλοιπο. Μεγαλύτερη διακύμανση παρουσιάζουν στο μονομερές A, τα κατάλοιπα στο αμινοτελικό άκρο [(1-4, (για λόγους σύγκρισης με το διμερές με την άλλη τοπολογία τα πρώτα τρία κατάλοιπα δεν φαίνονται στην εικόνα γιατί η τιμή του rmsf ξεπερνά τα 3 Å)], στη στροφή (27-31) και στο καρβοξυτελικό άκρο (56-60). Για το μονομερές B, υψηλότερες τιμές εμφανίζουν τα κατάλοιπα στις ίδιες περιοχές αλλά με λίγο μικρότερο εύρος. Δηλαδή τα κατάλοιπα 2-3 στο αμινοτελικό άκρο, 28-30 στη στροφή και 56-57 στο καρβοξυτελικό άκρο. Όπως έχει προαναφερθεί τα κατάλοιπα αυτά είναι περισσότερο ευκίνητα και πραγματοποιούν τυχαία κίνηση, κατά συνέπεια αναμένεται να εμφανίζουν μεγαλύτερη διακύμανση από τη μέση θέση τους (τη θέση τους στη μέση δομή του μορίου).



4.2.4 Εξέλιξη του Radius of Gyration (R_G)

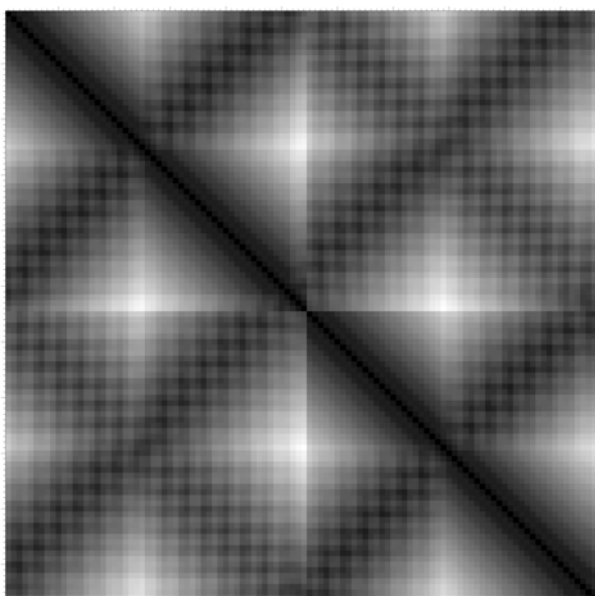
Στην **Εικόνα 4.55** απεικονίζεται η εξέλιξη του R_G της $[Ala_2Leu_2]_8$ με *syn* τοπολογία κατά τη διάρκεια του trajectory. Παρατηρούμε ότι κυμαίνεται γύρω από την τιμή των 14.1 Å -14.2 Å και δεν παρουσιάζει σημαντικές διακυμάνσεις. Το γεγονός αυτό δείχνει πως το μόριο είναι αρκετά συμπαγές. Η αρχική τιμή του R_G , που είναι αυτή της κρυσταλλικής δομής, είναι 14.55 Å.



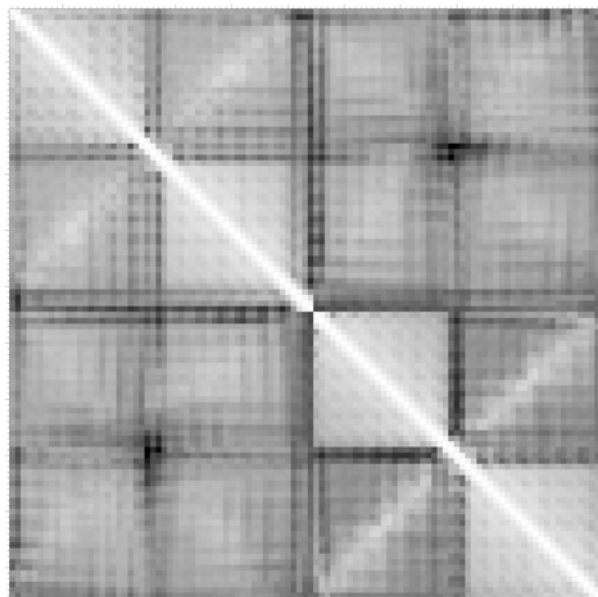
Οι απότομες διακυμάνσεις που είχαν παρατηρηθεί στα 2ns και 8ns – 9ns δεν είναι τόσο ευδιάκριτες εδώ. Αυτό μπορεί να σημαίνει ότι η αλλαγή στο μόριο ίσως να μην οδηγεί σε σημαντική μεταβολή της δομής του. Μελέτη των κύριων διαμορφώσεων που υιοθετεί το μόριο κατά τη διάρκεια του trajectory (cluster analysis), θα αποσαφηνίσει τη φύση των αλλαγών που λαμβάνουν χώρα.

4.2.5 Χάρτης της μέσης απόστασης μεταξύ Ca ατόμων και αντίστοιχο rmsd

Στην αρχή του Κεφαλαίου είδαμε πως μπορούμε να φτιάξουμε ένα πίνακα ο οποίος να περιέχει τις μέσες αποστάσεις μεταξύ όλων των Ca ατόμων. Μία αναπαράσταση ενός τέτοιου πίνακα φαίνεται στην **Εικόνα 4.56**. Όπως είδαμε στην Ενότητα 4.1.5, η πάνω οριζόντια γραμμή και η αριστερή κάθετη στήλη, αντιστοιχούν στα Ca άτομα των καταλοίπων του A και B μονομερούς και η διαβάθμιση του χρώματος από μαύρο σε λευκό ακολουθεί την αύξηση της απόστασης μεταξύ Ca ατόμων, με τις μικρότερες να απεικονίζονται με μαύρο και τις μεγαλύτερες με λευκό. Η μέγιστη απόσταση μεταξύ Ca ατόμων βρέθηκε να είναι 45.8 Å και όπως αναμένεται την παρουσιάζουν τα άτομα της στροφής του ενός μονομερούς σε σχέση με το άλλο μονομερές και τα C-τελικά κατάλοιπα του ενός μονομερούς σε σχέση με το άλλο μονομερές. Μικρές αποστάσεις έχουν βρεθεί μεταξύ είτε ατόμων που ανήκουν στην ίδια έλικα, είτε βρίσκονται στο ίδιο περίπου επίπεδο σε απέναντι έλικες.



Εικόνα 4.56 Ca-Ca distance map



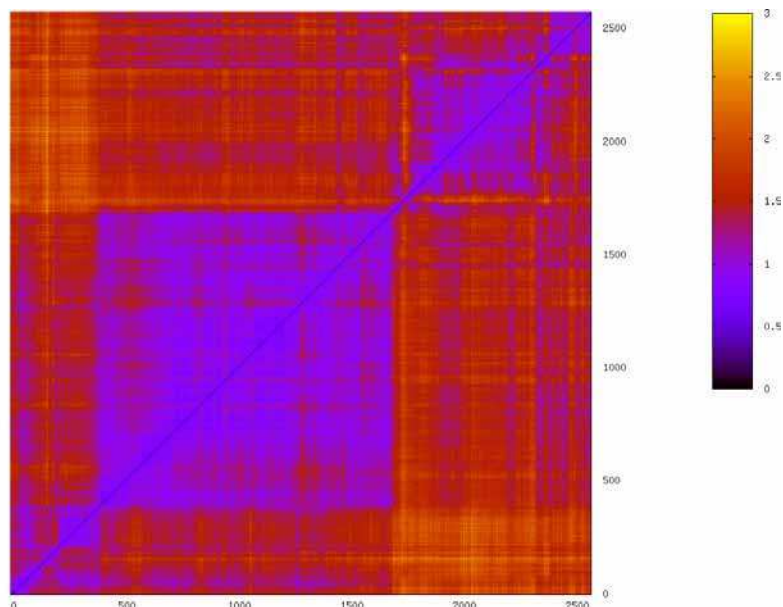
Εικόνα 4.57 rmsd που αντιστοιχεί στον Ca-Ca distance map

Μεγαλύτερο ενδιαφέρον παρουσιάζουν τα αποτελέσματα που απεικονίζονται στην **Εικόνα 4.57**. Στην εικόνα αυτή βλέπουμε μία γραφική αναπαράσταση των rmsd που αντιστοιχούν στην προηγούμενη εικόνα. Ισχύουν όσα προαναφέρθηκαν με τη διαφορά ότι ο χρωματικός κώδικας είναι ανεστραμμένος, δηλαδή οι μικρότερες τιμές rmsd φαίνονται με λευκό και οι μεγαλύτερες με μαύρο. Η μεγαλύτερη τιμή rmsd που βρέθηκε ήταν 2.27 Å. Στην εικόνα αυτή μπορούμε να διακρίνουμε ποια τμήματα της πρωτεΐνης είναι ιδιαίτερα ευκίνητα. (Να σημειωθεί ότι η αρίθμηση ξεκινά από το κατάλοιπο 6, καθώς τα πέντε πρώτα κατάλοιπα δεν λήφθηκαν υπόψη, για λόγους που αναλύθηκαν στην αρχή της ενότητας). Τα κατάλοιπα 57-60 στο καρβοξυτελικό άκρο Α μονομερούς κινούνται πολύ σε σχέση με όλο το υπόλοιπο μονομερές. Τα κατάλοιπα αυτά δεν συμμετέχουν στο σχηματισμό του υδροφοβου πυρήνα, αλλά σχηματίζουν μία δομή θηλιάς και είναι εκτεθειμένα προς το διάλυμα. Είναι λοιπόν αναμενόμενο να έχουν αυξημένη κινητικότητα. Επίσης μεγάλο rmsd εμφανίζουν τα Ca άτομα των καταλοίπων της στροφής του Β μονομερούς σε σχέση με τα αντίστοιχα της στροφής του Α μονομερούς. Να υπενθυμίσουμε στο σημείο αυτό, ότι στην *syn* τοπολογία τα κατάλοιπα στις στροφές των δύο μονομερών βρίσκονται απέναντι και συμμετέχουν στο σχηματισμό του πρώτου επιπέδου (βλ. Εικόνα 3.10 και Εικόνα 3.11). Μεγάλη κινητικότητα παρουσιάζουν και τα κατάλοιπα 6-7 του Α μονομερούς σε σχέση με όλο το Β μονομερές και τα κατάλοιπα 6-7 του Β μονομερούς σε σχέση με όλο το Α μονομερές. Ο λόγος είναι ότι το όγδοο επίπεδο, που συνίσταται από τα κατάλοιπα 5A, 5B, 56A, 56B, δεν είναι σταθερό και συνεπώς τα γειτονικά κατάλοιπα θα δείχνουν επίσης αυξημένη κινητικότητα (αφού είναι ομοιοπολικά συνδεδεμένα). Βλέπουμε λοιπόν πως σε αυτήν την τοπολογία, το πρώτο και όγδοο επίπεδο που συνίσταται από τα κατάλοιπα της στροφής και των ουρών αντίστοιχα, δεν είναι ιδιαίτερα σταθερά. Τέλος, μεγάλο rmsd παρουσιάζει το κατάλοιπο 57 του Β μονομερούς σε σχέση με το Α μονομερές, το οποίο υποδεικνύει ότι το κατάλοιπο αυτό δεν λαμβάνει μέρος σε κάποια σταθερή αλληλεπίδραση.

4.2.6 Rmsd μεταξύ των δομών του τροχιακού

Με το rmsd από την αρχική δομή έχουμε μία άποψη της κινητικότητας του μορίου και των δομών που υιοθετεί. Ωστόσο, υπολογίζοντας το rmsd ανάμεσα σε κάθε στιγμιότυπο (frame) του trajectory μπορούμε να έχουμε μία πιο ξεκάθαρη εικόνα. Μία γραφική απεικόνιση ενός τέτοιου υπολογισμού φαίνεται στην **Εικόνα 4.58**. Στον οριζόντιο και κάθετο άξονα αντιστοιχούν τα frame (συνολικά $2.573 \times 2.573 = 6.620.329$, μαζί με τη φάση του heating, καθώς ο υπολογισμός έγινε με βήμα 10). Όπως φαίνεται στο ένθετο δεξιά, η διαβάθμιση του χρώματος από μαύρο σε κίτρινο

ακολουθεί την αύξηση στην τιμή του rmsd. Η μέγιστη τιμή βρέθηκε να είναι 3 Å. Από τα χρώματα της εικόνας συνεπάγεται ότι η μεγάλη πλειοψηφία των τιμών rmsd βρέθηκαν μεταξύ 1 και 2 Å. Οι τιμές αυτές σε γενικές γραμμές συμφωνούν με τα αποτελέσματα που αναπαρίστανται στην Εικόνα 4.47.

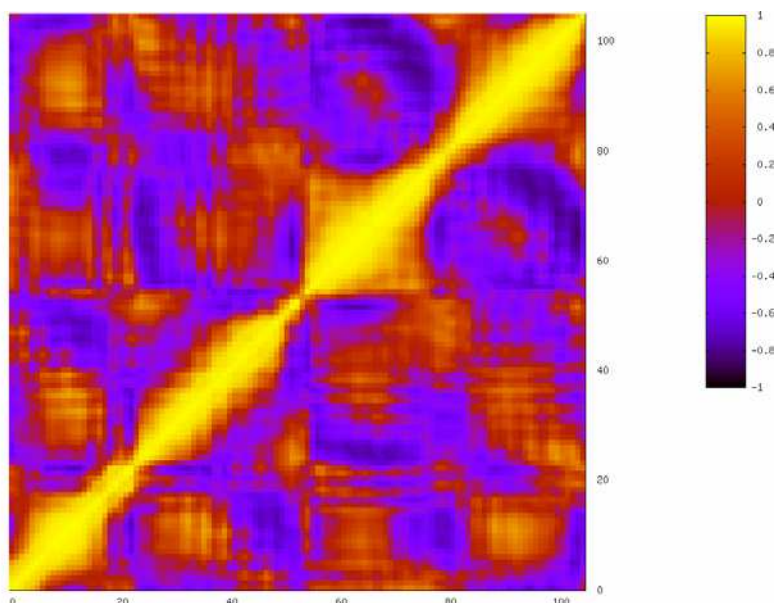


Εικόνα 4.58 frame-to-frame rmsd. Οι μονάδες στους άξονες είναι σε frame, με τις τιμές να είναι μία τάξη μεγέθους μικρότερες (επειδή ο υπολογισμός έγινε με βήμα 10).

Μία πιο προσεκτική παρατήρηση μας επιτρέπει να διακρίνουμε τη δημιουργία ομάδων δομών (clusters). Τα clusters φαίνονται σαν τετράγωνα διαφορετικού χρώματος που το μέγεθος τους αντικατοπτρίζει τον αριθμό των frame στα οποία παρατηρείται η συγκεκριμένη δομή. Στα πρώτα 4000 (400x10) περίπου frame (δηλαδή περίπου μέχρι τα 1.6ns) το μόριο φαίνεται να συμπεριφέρεται διαφορετικά και παρουσιάζει σχετικά χαμηλή τιμή rmsd (γύρω στα 1 - 1.5 Å) σε σχέση με τα υπόλοιπα. Είναι εμφανές ένα μεγάλο cluster που περιλαμβάνει τα frame 4000 (400x10) – 17000 (1700x10) (δηλαδή περίπου 1.6ns – 6.8ns) το οποίο παρουσιάζει σχετικά χαμηλές τιμές rmsd (1 – 1.5 Å). Ακολούθως, οι τιμές rmsd αυξάνουν και κυμαίνονται γύρω στα 2 – 2.5 Å, μέχρι και το τέλος του trajectory. Ένα ακόμα cluster, μικρότερο αυτή τη φορά, περιλαμβάνει κατά προσέγγιση τα frame 17000 (1700x10) – 23000 (2300x10) (δηλαδή περίπου 6.8ns – 9.2ns). Τέλος, τα frame 24000 (2400x10) -25730 (2573x10) (δηλαδή περίπου 9.6ns – 10.3ns), φαίνεται επίσης να ανήκουν σε μία ομάδα με χαμηλότερη τιμή rmsd. Βλέπουμε λοιπόν πως τα δεδομένα συνεχίζουν να υποστηρίζουν την ύπαρξη ευδιάκριτων αλλαγών, οι οποίες στην Εικόνα 4.47 φαίνονται σαν απότομες διακυμάνσεις (σκαλοπάτια). Σε αυτήν την περίπτωση όμως είμαστε σε θέση να ξέρουμε με ακρίβεια πότε συμβαίνουν αλλαγές στο μόριο.

4.2.7 Variance-Covariance και cross-correlation

Η γραφική απεικόνιση του cross-correlation πίνακα που προκύπτει από τον variance-covariance, παρουσιάζεται στην **Εικόνα 4.59**. Στον οριζόντιο και κάθετο άξονα αντιστοιχούν τα κατάλοιπα αρχικά του A μονομερούς (6-60) και στην συνέχεια του B (6-57). Στο ένθετο δεξιά αναφέρεται η διαβάθμιση του χρώματος από μαύρο σε κίτρινο. Η συσχέτιση παίρνει τιμές από -1 (αρνητική συσχέτιση) έως +1 (θετική συσχέτιση). Με κόκκινο σημειώνονται οι κινήσεις που δεν εμφανίζουν συσχέτιση (τιμή 0).



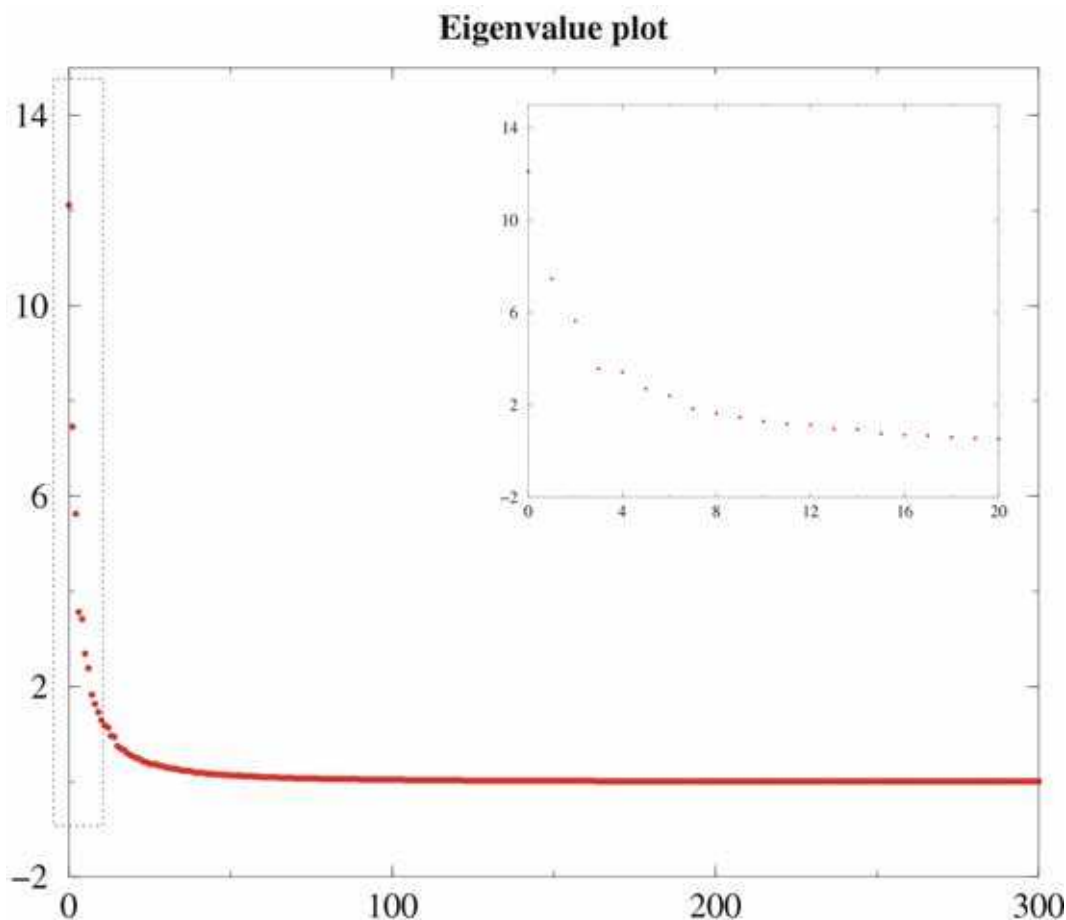
Εικόνα 4.59 variance-covariance and cross-correlation

Η διαγώνιος γραμμή αντιστοιχεί στο ίδιο κατάλοιπο (ως εκ τούτου θα έχει πλήρη συσχέτιση με τον εαυτό του και για αυτό το λόγο φαίνεται με έντονο κίτρινο χρώμα). Το πάχος της γραμμής όμως περιλαμβάνει περισσότερα από ένα κατάλοιπα. Αυτό υποδεικνύει ότι και τα γειτονικά κατάλοιπα κινούνται μαζί, κάτι που είναι αναμενόμενο αφού είναι ομοιοπολικά συνδεδεμένα. Μάλιστα σε ορισμένες περιπτώσεις όπως, για τα κατάλοιπα της πρώτης έλικας του A μονομερούς και τα κατάλοιπα της πρώτης και δεύτερης έλικας του B μονομερούς, το φαινόμενο είναι πολύ ισχυρό. Το γεγονός αυτό σημαίνει ότι οι έλικες αυτές μάλλον κινούνται σαν συμπαγή σώματα, δηλαδή τα κατάλοιπα που ανήκουν σε αυτές κινούνται με τον ίδιο τρόπο. Ο ισχυρισμός αυτός επιβεβαιώνεται και από τις πορτοκαλί διαγώνιες γραμμές που είναι κάθετες στη κεντρική κίτρινη διαγώνιο. Αυτές οι διαγώνιες γραμμές περιλαμβάνουν κατάλοιπα που συμμετέχουν στο σχηματισμό ενός επιπέδου (για την *syn* τοπολογία και τα κατάλοιπα που σχηματίζουν τα επίπεδα σε αυτήν βλ. Εικόνες 3.10 και 3.11) και επομένως είναι αναμενόμενο εξαιτίας των υδρόφοβων

αλληλεπιδράσεων που αναπτύσσονται ανάμεσά τους και συμβάλλουν στην σταθεροποίηση του πυρήνα. Κινήσεις που δεν σχετίζονται μεταξύ τους εμφανίζονται ανάμεσα σε κατάλοιπα που ανήκουν σε διαφορετικές έλικες και δεν βρίσκονται στο ίδιο επίπεδο (φαίνονται με μωβ διαγώνιες γραμμές που είναι κάθετες στη κεντρική κίτρινη διαγώνιο και παρεμβάλλονται στις πορτοκαλί διαγώνιους που προαναφέρθηκαν). Τέλος, κάτι που προξενεί ιδιαίτερη εντύπωση είναι η έντονα αρνητική συσχέτιση που εμφανίζουν στη κίνησή τους η πρώτη με τη δεύτερη έλικα του B μονομερούς (έντονα σκουρόχρωμη διαγώνιος πάνω δεξιά).

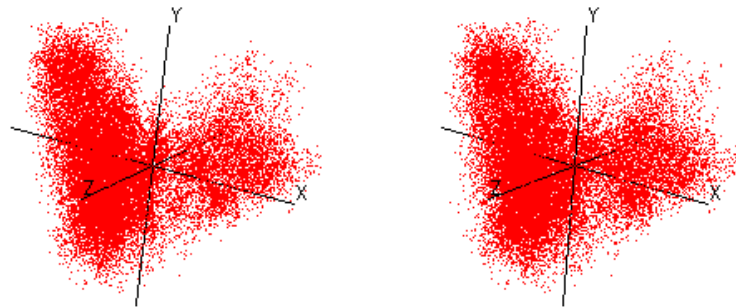
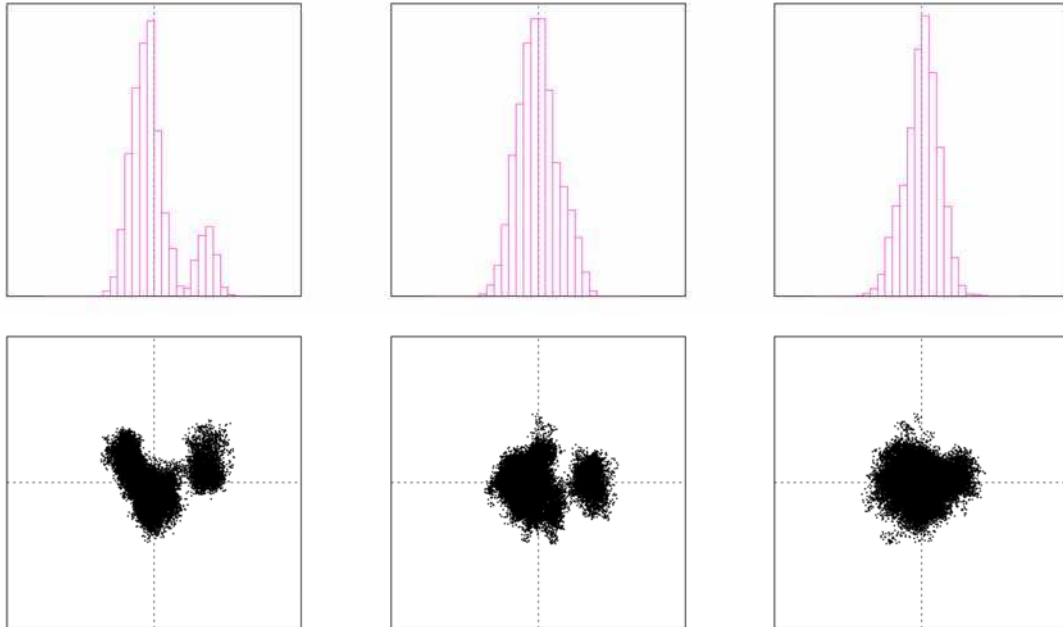
4.2.8 Eigenvalues και Eigenvectors

Όπως είδαμε και στην Ενότητα 4.1.8, προκειμένου να δούμε τις χαρακτηριστικές κινήσεις που πραγματοποιεί το μόριό μας, θα πρέπει, αρχικά να εξετάσουμε εάν υπάρχει sufficient sampling υπολογίζοντας τις διακυμάνσεις (fluctuations) των τριών πρώτων eigenvectors, οι οποίοι, όπως φαίνεται στην **Εικόνα 4.60** θα αντιπροσωπεύουν την πλειοψηφία των κύριων κινήσεων του μορίου.



Εικόνα 4.60 Διαγραμματική απεικόνιση των eigenvalues. Στο ένθετο δεξιά φαίνεται σε μεγέθυνση το τμήμα του γραφήματος που περικλείεται με το διακεκομμένο κουτί. Οι πρώτοι 10 eigenvectors μπορούν να εξηγήσουν σχεδόν πλήρως τη κίνηση του μορίου.

Στην **Εικόνα 4.61** παρατίθεται μία γραφική απεικόνιση της προβολής των διακυμάνσεων των τριών πρώτων eigenvectors. Όπως είδαμε στην προηγούμενη ενότητα για να υπάρχει sufficient sampling κάποιων κινήσεων, θα πρέπει σε κάθε αναπαράσταση να βλέπουμε μία κανονική κατανομή.



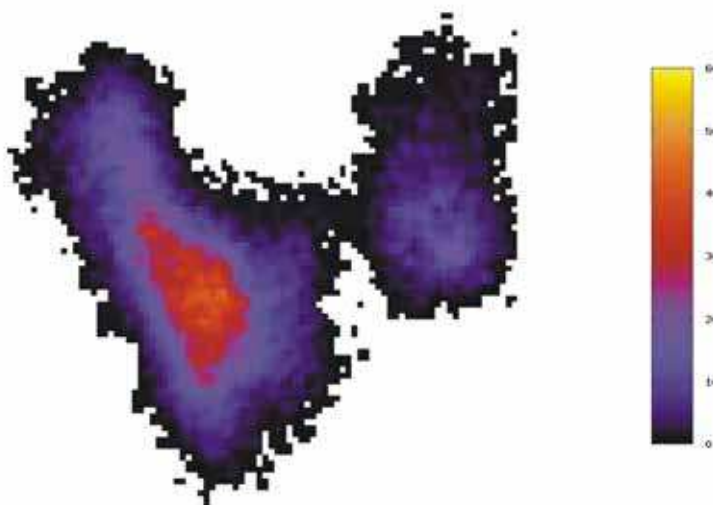
Εικόνα 4.61 Πάνω σειρά Προβολή των διακυμάνσεων των τριών πρώτων eigenvectors στη μία διάσταση.

Μεσαία σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων δύο eigenvectors (δύο διαστάσεις).

Κάτω σειρά Προβολή των διακυμάνσεων στο επίπεδο των τριών πρώτων eigenvectors (τρεις διαστάσεις).

Βλέπουμε λοιπόν, πως sufficient sampling δεν έχει επιτευχθεί για τον πρώτο eigenvector, και είναι ευδιάκριτη η δημιουργία δύο cluster, εκ των οποίων το ένα φαίνεται να αποτελείται από πολύ μεγαλύτερο αριθμό σημείων. Αυτό αποτελεί ένδειξη ότι ίσως δε θα πρέπει να μελετήσουμε τους eigenvectors κατά τη διάρκεια ολόκληρου του trajectory. Εφόσον από τη μέχρι τώρα ανάλυση έχει προκύψει πως στο μόριο συμβαίνουν κάποιες ευδιάκριτες αλλαγές σε καθορισμένα σημεία του trajectory, ίσως να έχει περισσότερο νόημα να μελετήσουμε τις κινήσεις που πραγματοποιεί το μόριο, σε επιμέρους τμήματα του (βλ. Ενότητα 4.3.1, Ενότητα 4.3.3 και Ενότητα 4.3.5). Οι επόμενοι δύο eigenvectors, φαίνεται να εμφανίζουν μία κατανομή που πλησιάζει αρκετά την κανονική.

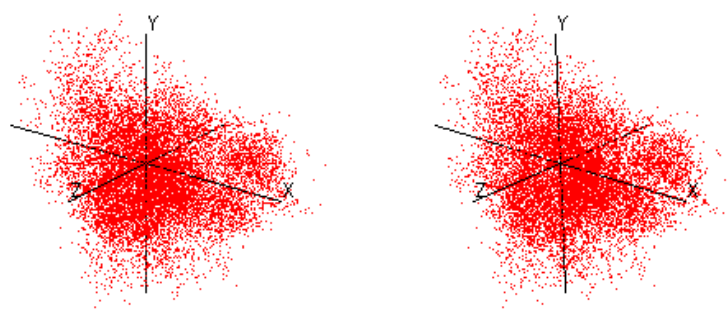
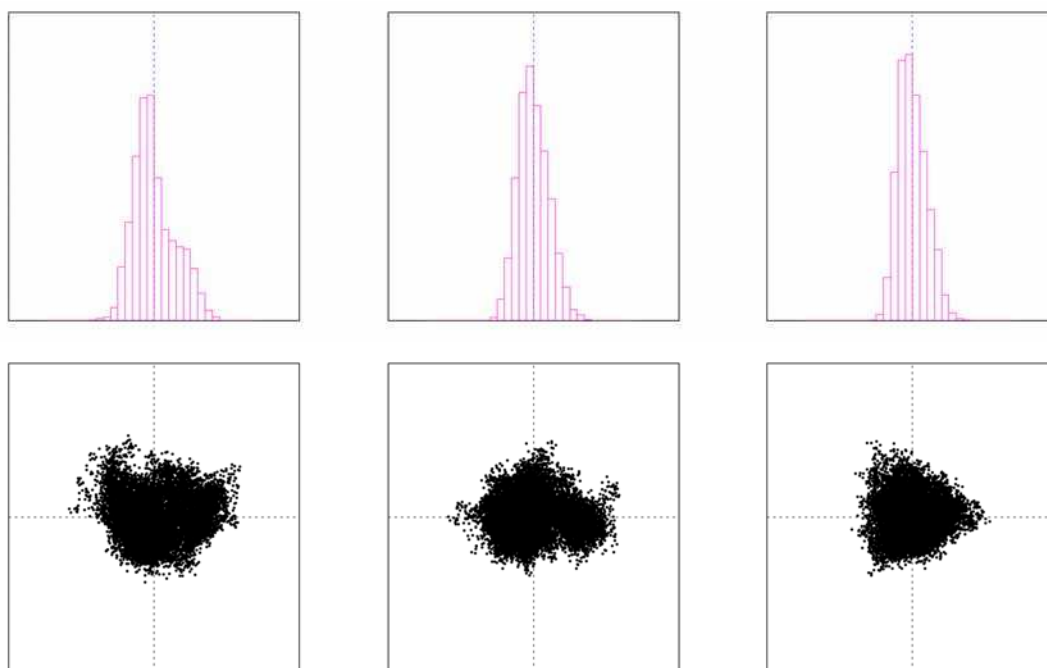
Στην **Εικόνα 4.62** βλέπουμε μία απεικόνιση των διακυμάνσεων των προβολών των δύο πρώτων eigenvector με βάση την πυκνότητα (αριθμό) των σημείων (distribution of density of points). Η διαβάθμιση του χρώματος από μαύρο (0) σε κίτρινο (60) ακολουθεί την αύξηση στην πυκνότητα των σημείων. Από την εικόνα αυτή βλέπουμε πως το δεύτερο cluster περιέχει σημαντικά μικρότερο αριθμό σημείων. Στην επόμενη ενότητα (ομαδοποίηση δομών του trajectory), θα μελετηθούν τόσο οι ομάδες δομών που εμφανίζονται, όσο και ο χρόνος παραμονής του μορίου σε κάθε μία από αυτές τις διαμορφώσεις.



Εικόνα 4.62 Προβολή των διακυμάνσεων ανά ζεύγη eigenvectors
βάση πυκνότητας σημείων.

Αν επαναλάβουμε την διαδικασία που περιγράφηκε παραπάνω λαμβάνοντας υπόψη ένα τμήμα μόνο του trajectory, τα αποτελέσματα, αναφορικά με το αν έχουμε sufficient sampling, διαφοροποιούνται. Στις **Εικόνες 4.63** και **4.64** βλέπουμε την αντίστοιχη εικόνα που προκύπτει υπολογίζοντας καινούργια eigenvectors και eigenvalues για τα frame 6000 - 20000 (δηλαδή από 2.4ns μέχρι 8ns). Τα frame αυτά

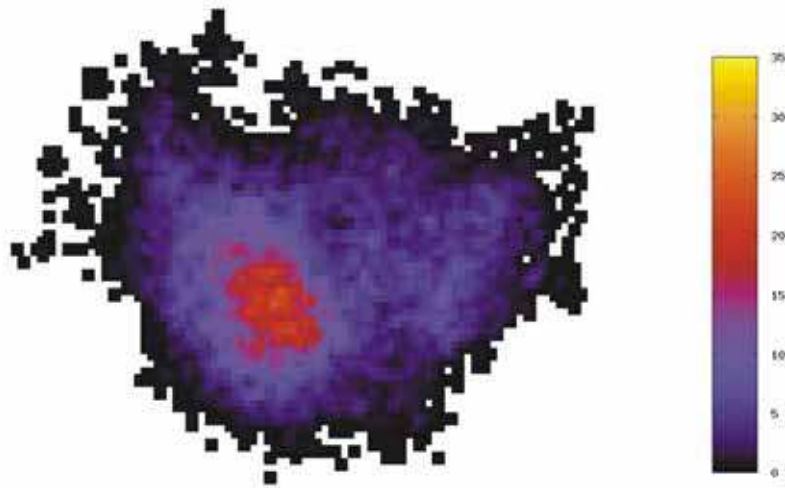
αποτελούν το μεγαλύτερο τμήμα του χρόνου της προσομοίωσης, στα όρια των οποίων παρατηρήθηκαν οι έντονες διακυμάνσεις που προαναφέρθηκαν. Η επιλογή έγινε με βάση τα αποτελέσματα που προέκυψαν υπολογίζοντας το frame-to-frame rmsd (Ενότητα 4.2.6) και επιβεβαιώνονται από το cluster analysis που ακολουθεί. Με αυτόν τον τρόπο φαίνεται πως υπάρχει sufficient sampling και για τους τρεις eigenvectors σε αυτό το διάστημα της προσομοίωσης καθώς βλέπουμε κανονικές κατανομές και στις τρεις διαστάσεις.



Εικόνα 4.63 Πάνω σειρά Προβολή των διακυμάνσεων των τριών πρώτων eigenvectors στη μία διάσταση που υπολογίστηκαν από 2.4ns μέχρι 8ns.

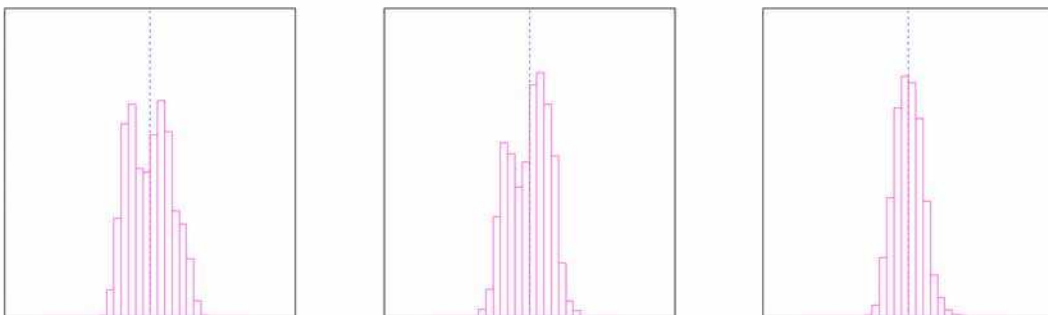
Μεσαία σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων δύο eigenvectors (δύο διαστάσεις) που υπολογίστηκαν από 2.4ns μέχρι 8ns.

Κάτω σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων τριών eigenvectors (τρεις διαστάσεις) που υπολογίστηκαν από 2.4ns μέχρι 8ns.

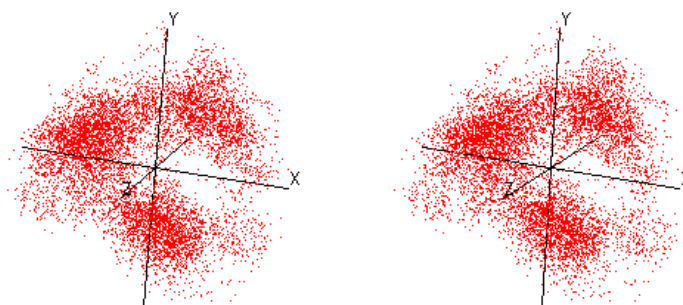
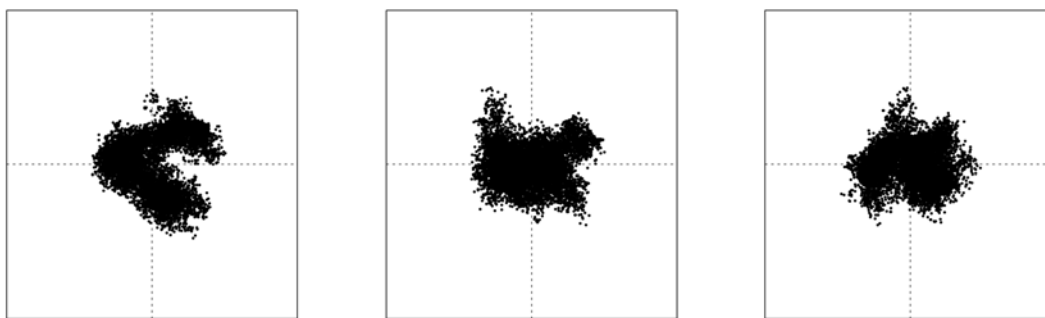


Εικόνα 4.64 Προβολή των διακυμάνσεων ανά ζεύγη eigenvectors βάση πυκνότητας σημείων που υπολογίστηκαν από 2.4ns μέχρι 8ns.

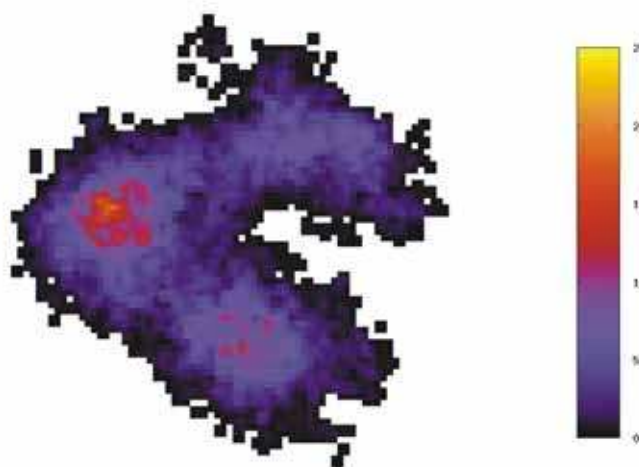
Ακολούθως υπολογίσαμε eigenvalues και eigenvectors για τα frame 21000-25000 (δηλαδή από 8.4ns μέχρι 10ns). Η αναπαράσταση των διακυμάνσεών τους φανερώνει πως χρειαζόταν περισσότερος χρόνος για να μελετηθεί ολοκληρωμένα η κίνηση του μορίου προς το τέλος της προσομοίωσης (**Εικόνα 4.65**). Ενώ δεν φαίνεται να υπάρχει πρόβλημα όσον αφορά το sufficient sampling, για τον δεύτερο και τρίτο eigenvector, για τον πρώτο είναι και εμφανή τρία cluster. Όπως φαίνεται στην **Εικόνα 4.66**, και τα τρία cluster περιέχουν σημαντικό αριθμό σημείων. Αυτό σημαίνει ότι ο χρόνος της προσομοίωσης δεν ήταν επαρκής για να μελετηθεί ολοκληρωμένα το μόριο αυτό.



Συνέχεια της εικόνας στην επόμενη σελίδα

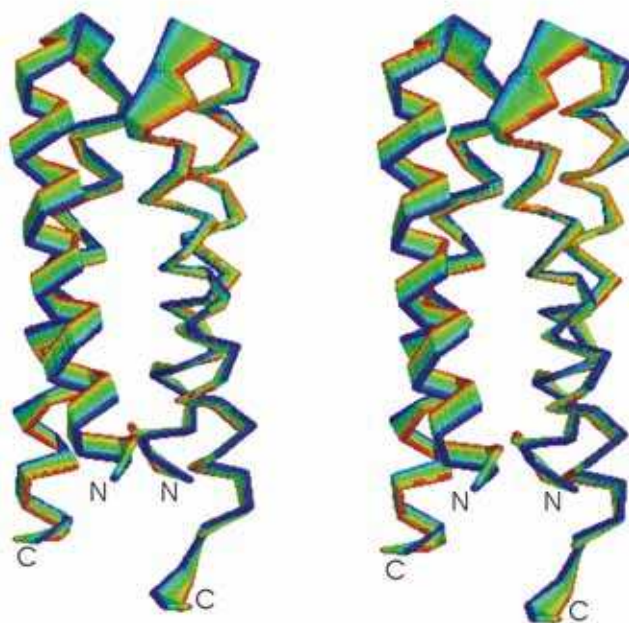


Εικόνα 4.65 Πάνω σειρά Προβολή των διακυμάνσεων των τριών πρώτων eigenvectors στη μία διάσταση που υπολογίστηκαν από 8.4ns μέχρι 10ns.
 Μεσαία σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων δύο eigenvectors (δύο διαστάσεις) που υπολογίστηκαν από 8.4ns μέχρι 10ns.
 Κάτω σειρά Προβολή των διακυμάνσεων στο επίπεδο των πρώτων τριών eigenvectors (τρεις διαστάσεις) που υπολογίστηκαν από 8.4ns μέχρι 10ns.



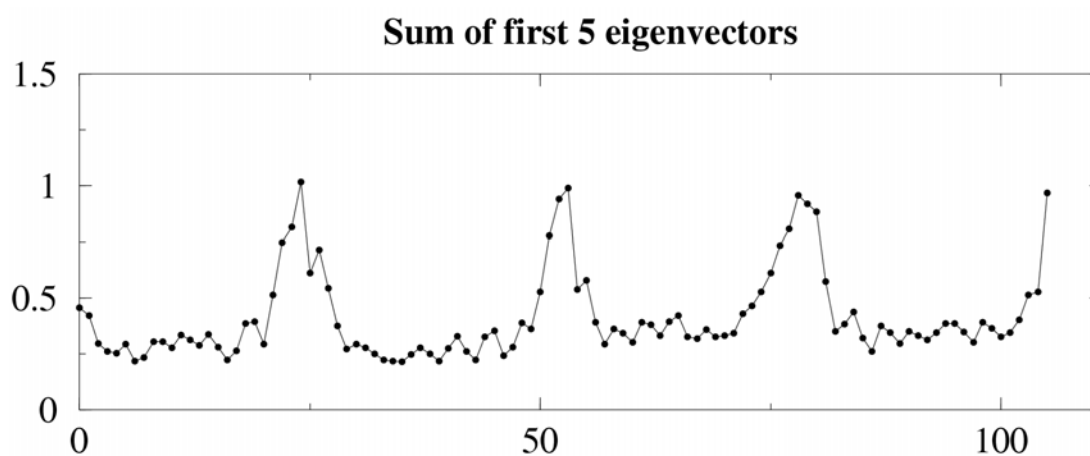
Εικόνα 4.66 Προβολή των διακυμάνσεων ανά ζεύγη eigenvectors βάση πυκνότητας σημείων που υπολογίστηκαν από 8.4ns μέχρι 10ns.

Στην **Εικόνα 4.67** βλέπουμε μία γραφική απεικόνιση της κίνησης που πραγματοποιεί το μόριο βάση του πρώτου eigenvector. Ο υπολογισμός έγινε για ολόκληρο το trajectory. Η διαβάθμιση του χρώματος από μπλε σε κόκκινο φανερώνει το εύρος της κίνησης. Βλέπουμε πως οι στροφές των δύο μονομερών κινούνται σε σημαντικά μεγαλύτερο βαθμό σε σχέση με το υπόλοιπο μόριο. Είναι αξιοσημείωτο το γεγονός ότι και οι δύο έλικες του A μονομερούς εμφανίζουν κινητικότητα του ίδιου εύρους, σε αντίθεση με την πρώτη έλικα του B μονομερούς, που παρουσιάζει πολύ μικρότερο εύρος. Αυτό έρχεται σε συμφωνία με τα αποτελέσματα που προέκυψαν από τον πίνακα variance-covariance (βλ. Ενότητα 4.2.7). Στην Εικόνα 4.44 είδαμε την απεικόνιση του πρώτου ζεύγους eigenvalue – eigenvector λαμβάνοντας υπόψη κατά τον υπολογισμό όλα τα άτομα. Στην εικόνα αυτή φαίνεται κυρίως η κίνηση των αμινοτελικών καταλοίπων και των καταλοίπων των στροφών, ενώ για το υπόλοιπο μόριο δεν φαίνεται αν κάποιο τμήμα εμφανίζει μεγαλύτερο εύρος στη κίνησή του. Παραλείποντας τα κατάλοιπα που γνωρίζουμε εκ των προτέρων ότι κινούνται σε μεγάλο βαθμό επειδή δε συμμετέχουν σε κάποια σταθερή δομή (βλ. αρχή Ενότητας 4.2) μπορούμε να εστιάσουμε καλύτερα στη κίνηση του υπόλοιπου μορίου. Έτσι αναπαρίσταται καλύτερα τόσο η κίνηση των στροφών όσο και των ελίκων.

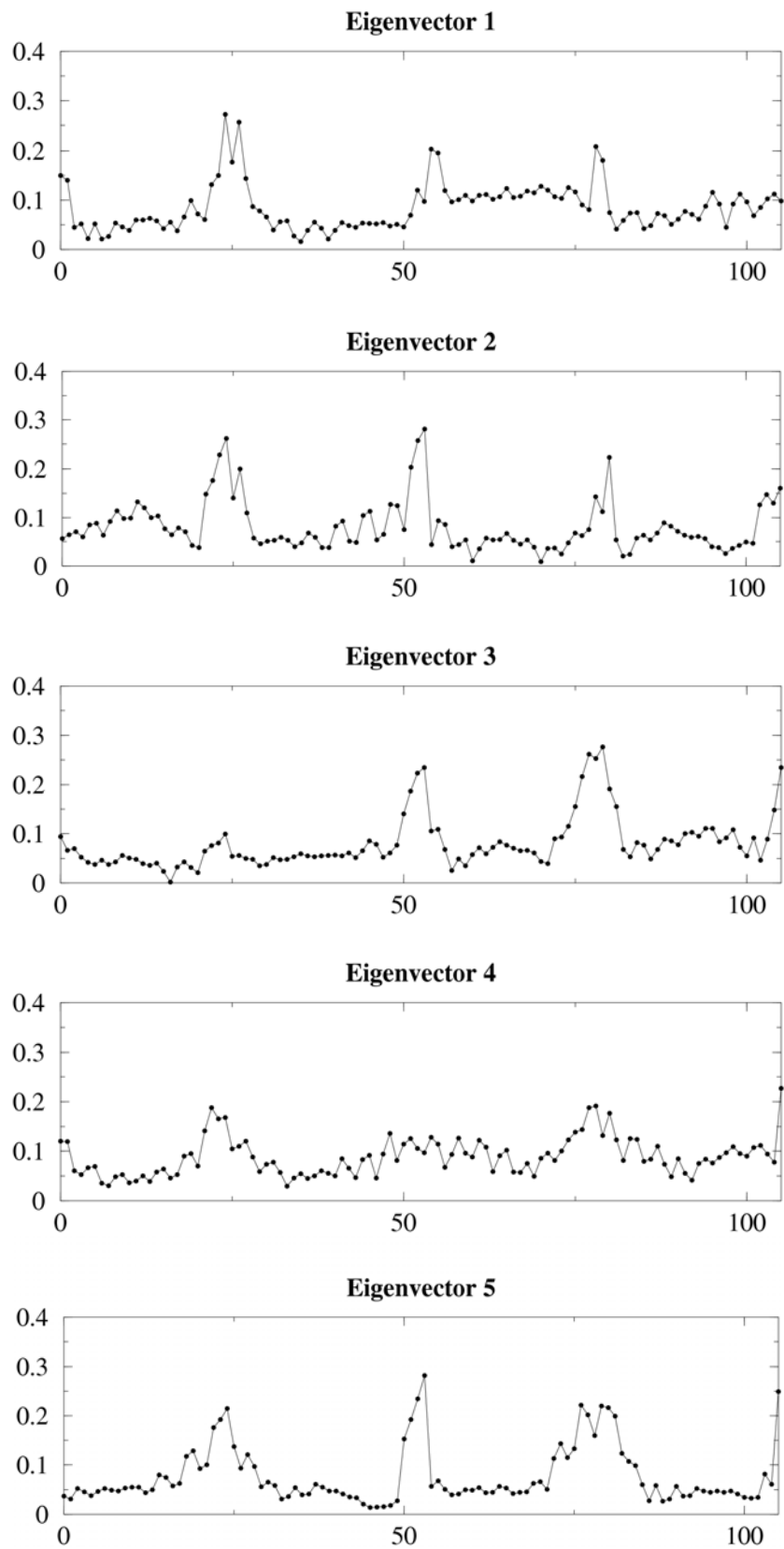


Εικόνα 4.67 Stereo αναπαράσταση της κίνησης του μορίου που περιγράφεται από το πρώτο ζεύγος eigenvalue – eigenvector. Τα πρώτα πέντε αμινοτελικά κατάλοιπα κάθε μονομερούς έχουν αφαιρεθεί. Η ίδια αναπαράσταση για το πλήρους μήκους μόριο φαίνεται στην Εικόνα 4.44.

Ολοκληρώνοντας την μελέτη των eigenvectors, μπορούμε να δούμε την κατανομή των διακυμάνσεων του κάθε eigenvector ανά κατάλοιπο. Με τον τρόπο αυτό έχουμε μία άποψη της ποσότητας των σχετιζόμενων διακυμάνσεων του eigenvector μεταξύ των καταλοίπων. Στην **Εικόνα 4.68** βλέπουμε αθροιστικά την διακύμανση των πέντε πρώτων eigenvectors για κάθε κατάλοιπο. Αυτή η αθροιστική διακύμανση συμφωνεί με τα αποτελέσματα από το rmsf (βλ. Ενότητα 4.2.3, Εικόνα 4.54). Μεγάλη κινητικότητα εμφανίζουν τα κατάλοιπα στις περιοχές των στροφών καθώς και στα αμινοτελικά και καρβοξυτελικά άκρα. Στην **Εικόνα 4.69** έχουμε αναλυτικά την αναπαράσταση για κάθε έναν eigenvector ξεχωριστά.



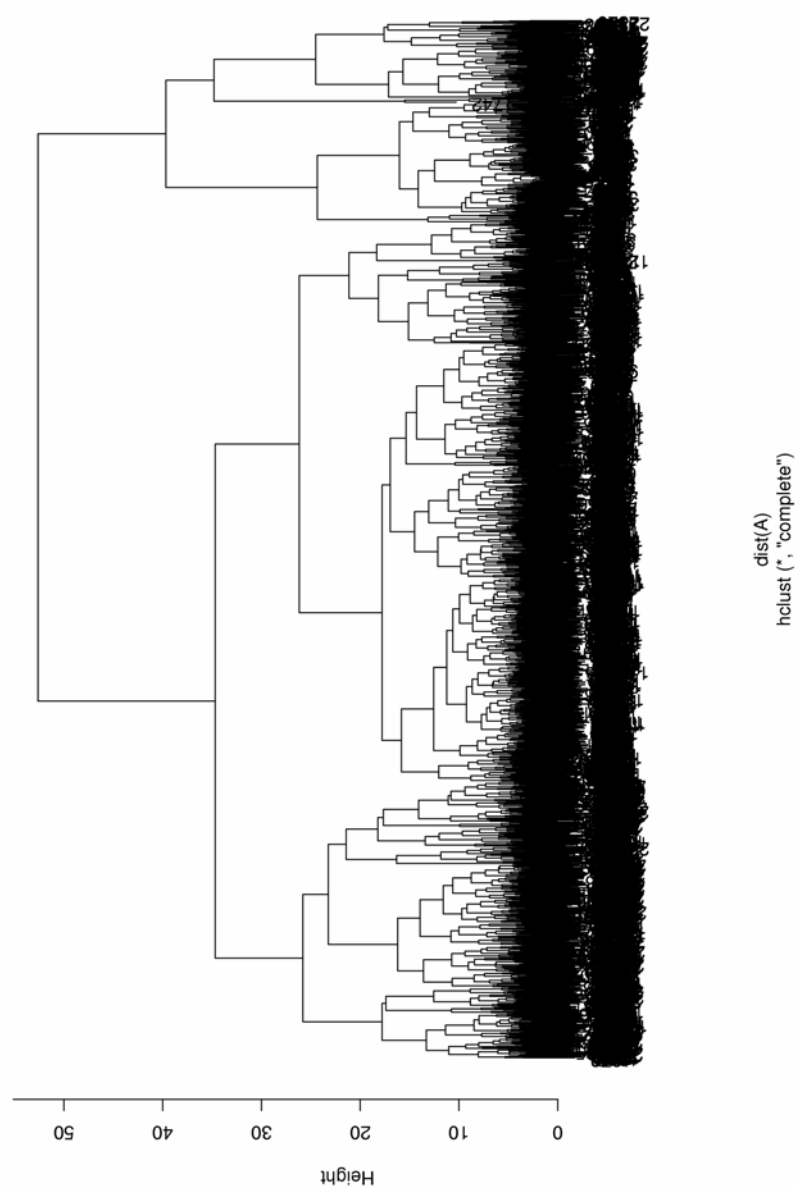
Εικόνα 4.68 Αναπαράσταση της αθροιστικής διακύμανσης των πρώτων 5 eigenvector ανά κατάλοιπο.



Εικόνα 4.69 Αναπαράσταση των διακυμάνσεων των πρώτων 5 eigenvector ανά κατάλοιπο.

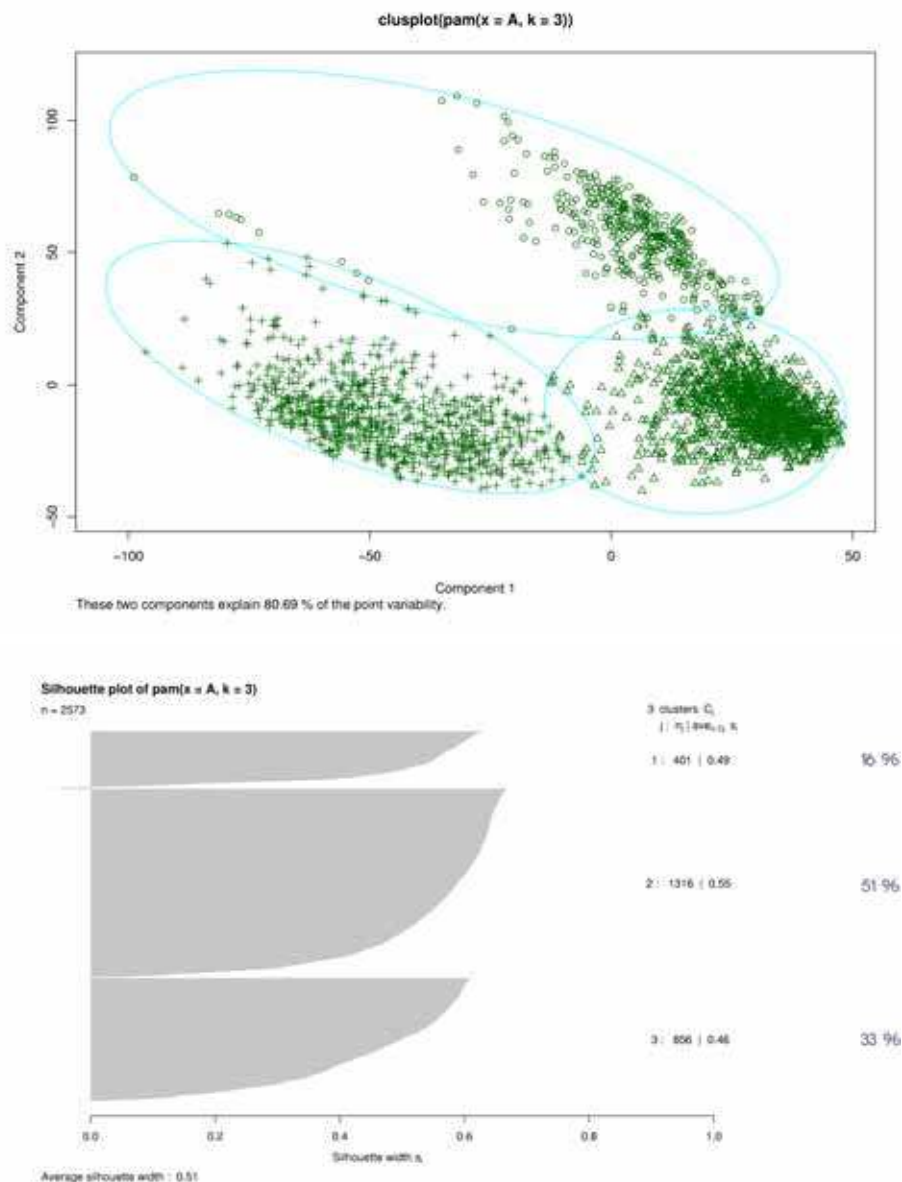
4.2.9 Cluster analysis του trajectory

Όπως είδαμε και στην Ενότητα 4.1.9, όταν πραγματοποιούμε cluster analysis του trajectory, καταρχήν δημιουργούμε ένα δένδρο το οποίο αναπαριστά τις δομές που παρατηρούνται κατά τη διάρκεια της προσομοίωσης, βάση του rmsd που έχουν (Εικόνα 4.70). Με βάση αυτό το δένδρο μπορούμε να συνάγουμε τον αριθμό των κυρίαρχων cluster δομών. Ένας πιο αξιόπιστος τρόπος για να γίνει αυτό είναι βάση ενός στατιστικού προγράμματος. Με τον τρόπο αυτό προκύπτουν τα στατιστικά σημαντικά clusters.



Εικόνα 4.70 Δενδρόγραμμα των δομών που παρατηρούνται κατά τη διάρκεια της προσομοίωσης του διμερούς $[Ala_2Leu_2]_8$ με *syn* τοπολογία, βάση του rmsd.

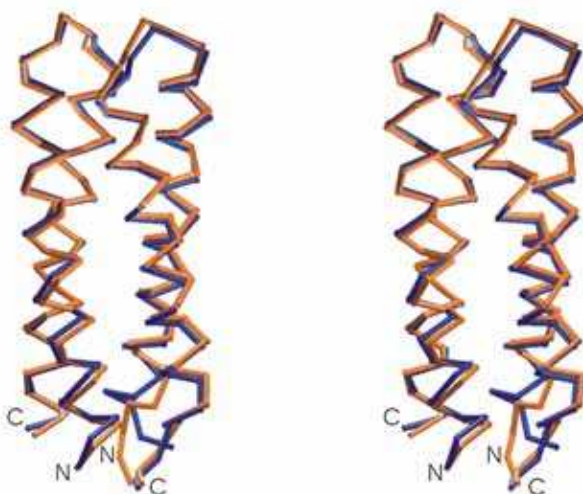
Όπως φαίνεται στην **Εικόνα 4.71** υπάρχουν τρεις σημαντικές ομάδες δομών. Ο αλγόριθμος που χρησιμοποιήθηκε στους παρακάτω υπολογισμούς ήταν ο pam. Ένα σημαντικό πλεονέκτημα της χρήσης της στατιστικής στην ανάλυση, είναι η δυνατότητα υπολογισμού ποσοστών για το κάθε cluster. Με τη μέθοδο αυτή μπορούμε να γνωρίζουμε με ακρίβεια σε ποιο cluster δομών ανήκει το κάθε frame. Έτσι προκύπτει πως το πρώτο cluster που αποτελεί το 16% των δομών περιλαμβάνει τα frame 1-3870 (δηλαδή έως τα 1.5ns). Το δεύτερο cluster αποτελεί το 51% των δομών περιλαμβάνει τα frame 3880-16900 (δηλαδή έως τα 6.8ns) και το τρίτο cluster που αποτελεί το 33% των δομών, περιλαμβάνει τα frame 16910-25730 (δηλαδή έως τα 10.29ns). Τα αποτελέσματα αυτά φαίνονται και από το frame-to-frame rmsd, Εικόνα 4.58.



Εικόνα 4.71 Πάνω Προσδιορισμός των clusters.

Κάτω Κατανομή των δομών στα τρία clusters και υπολογισμός ποσοστών.

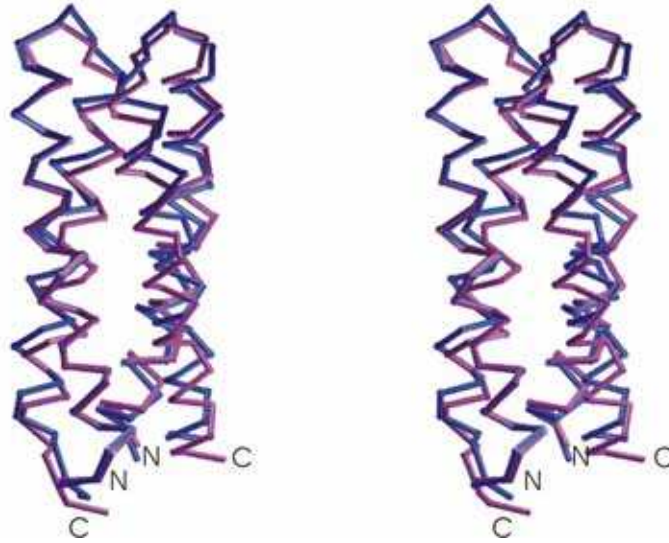
Ακολούθως προσδιορίστηκαν οι μέσες δομές που προκύπτουν από το δεύτερο και τρίτο cluster που αποτελούν τις στατιστικά σημαντικές ομάδες δομών και καταλαμβάνουν το μεγαλύτερο τμήμα του trajectory. Στις **Εικόνες 4.72, 4.73 και 4.74** που ακολουθούν γίνεται σύγκριση των δομών αυτών τόσο μεταξύ τους όσο και με την δομή από την οποία ξεκίνησαν. Σε κάθε περίπτωση η στοίχιση έγινε βάσει της πρώτης έλικας του μονομερούς A (αριστερά). Οι μέσες δομές που υπολογίστηκαν για τα δύο cluster φαίνεται να ταιριάζουν αρκετά καλά μεταξύ τους. Η μόνη σημαντική απόκλιση που παρατηρείται, είναι στο αμινοτελικό άκρο του μονομερούς B, όπου και πρέπει να οφείλεται η αύξηση στη τιμή του rmsd (σκαλοπάτι) που βλέπουμε γύρω στα 8ns (βλ. Εικόνα 4.47) αλλά και το μεγαλύτερη τιμή rmsd του μονομερούς B σε σχέση με το μονομερές A (βλ. Εικόνα 4.48).



Εικόνα 4.72 Stereo σύγκριση της μέσης δομής του δεύτερου cluster (απεικονίζεται με μπλε χρώμα) με τη μέση δομή του τρίτου cluster (απεικονίζεται με πορτοκαλί χρώμα).

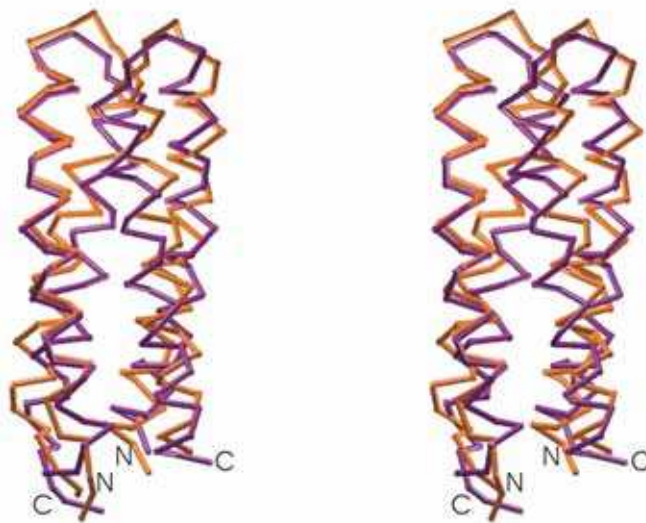
Η μέση δομή του δεύτερου cluster φαίνεται να είναι αρκετά κοντά στη δομή από την οποία ξεκίνησε, κάτι που είναι αναμενόμενο βάσει των μέχρι τώρα αποτελεσμάτων. Μικρές αποκλίσεις συναντάμε στα αμινοτελικά και καρβοξυτελικά άκρα και των δύο μονομερών καθώς και στις περιοχές των στροφών.

Οι πιο εμφανείς διαφορές εμφανίζονται ανάμεσα στη μέση δομή του τρίτου cluster και την αρχική δομή του trajectory. Υπάρχει απόκλιση στη περιοχή της στροφής του μονομερούς A αλλά και του μονομερούς B, όπως επίσης και στα N-τελικά και C-τελικά κατάλοιπα του μονομερούς A. Ωστόσο, βλέπουμε πως κατά τη διάρκεια της προσομοίωσης παρατηρείται μεγάλη απόκλιση στις έλικες, σαν συμπαγή σώματα, ιδιαίτερα στο μονομερές B (έτσι δικαιολογείται και η υψηλότερη τιμή rmsd).



Εικόνα 4.73 Stereo σύγκριση της μέσης δομής του δεύτερου cluster (απεικονίζεται με μπλε χρώμα) με την αρχική δομή του trajectory (απεικονίζεται με μωβ χρώμα).

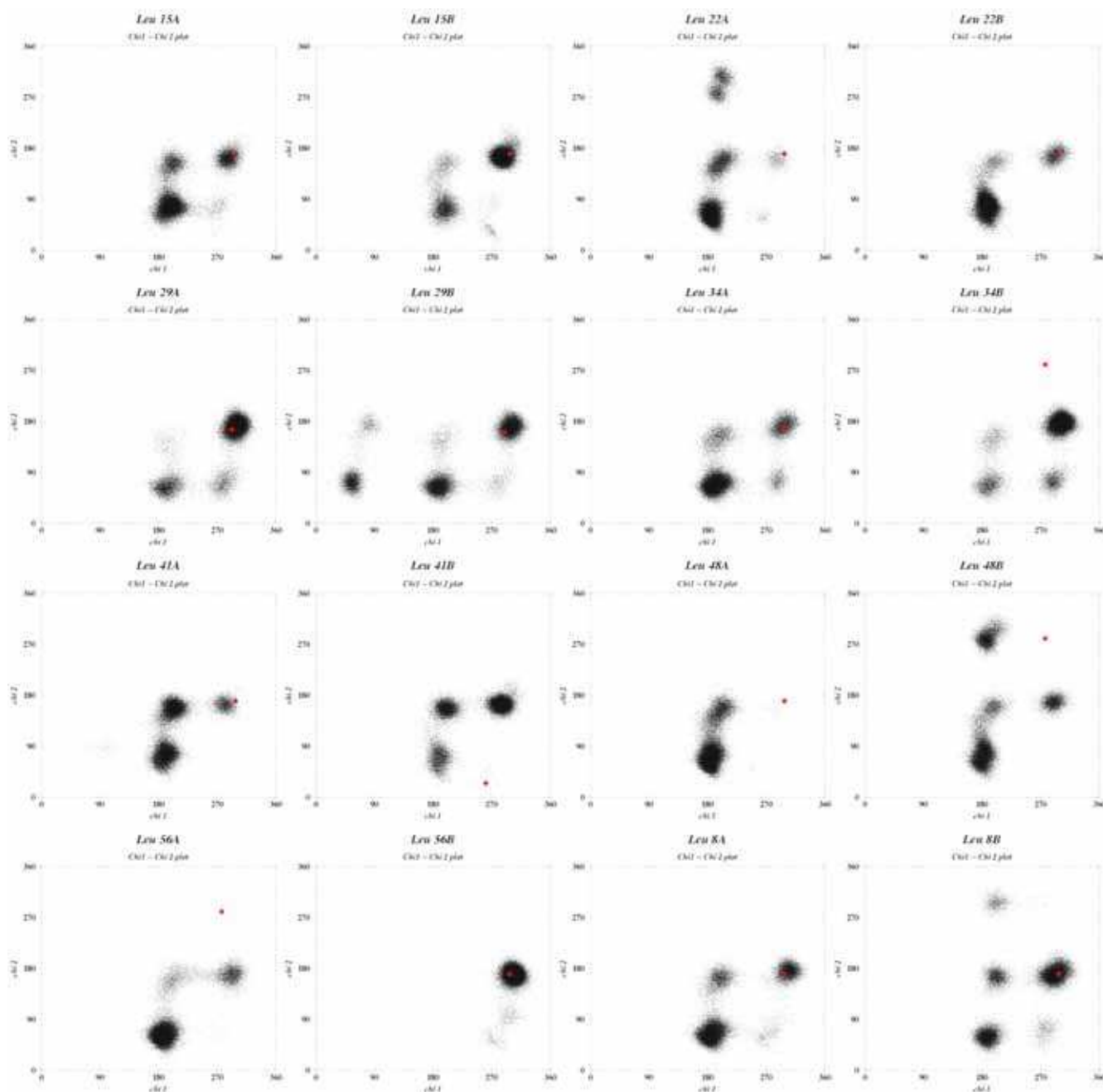
Οι εικόνες αυτές εξηγούν σε ικανοποιητικό βαθμό τις δομικές αλλαγές που συμβαίνουν στο μόριο και στις οποίες οφείλονται οι διακυμάνσεις των τιμών των παραμέτρων που μελετήθηκαν μέχρι τώρα. Για την αποσαφήνιση της ακριβούς φύσης των αλλαγών αυτών θα πρέπει να γίνει σε βάθος δομική μελέτη κατά την οποία, η προσοχή μας θα επικεντρωθεί σε συγκεκριμένα κατάλοιπα, τις αλληλεπιδράσεις που αναπτύσσονται και πως αυτές μεταβάλλονται. Πάντως σε γενικές γραμμές, από τη μελέτη των μέχρι τώρα αποτελεσμάτων, φαίνεται πως ο υπολογιστικός χρόνος της προσομοίωσης του συστήματος αυτού δεν ήταν επαρκής για να γίνει ολοκληρωμένη μελέτη του μορίου με την *syn* τοπολογία.



Εικόνα 4.74 Stereo σύγκριση της μέσης δομής του τρίτου cluster (απεικονίζεται με πορτοκαλί χρώμα) με την αρχική δομή του trajectory (απεικονίζεται με μωβ χρώμα).

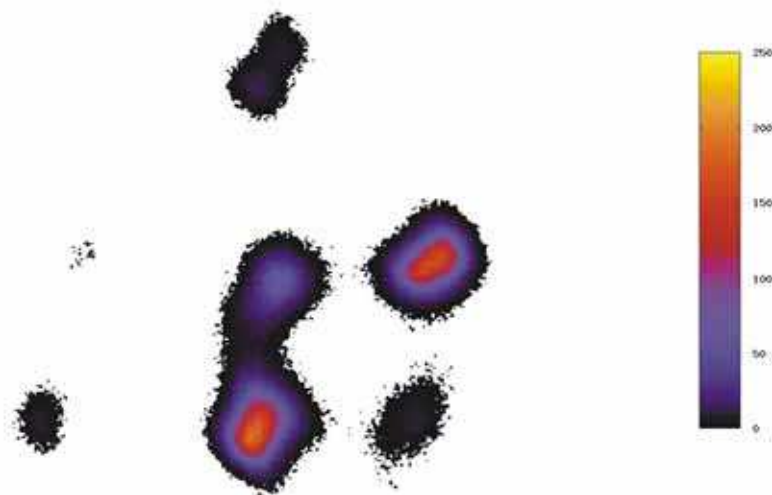
4.2.10 Ανάλυση διεδρων γωνιών (χ -value)

Σε αυτήν την ανάλυση, υπολογίζουμε, όπως αναφέρθηκε στην αρχή του Κεφαλαίου, διέδρες γωνίες πλευρικών ομάδων όλων των καταλοίπων λευκίνης που συμμετέχουν στο σχηματισμό του υδρόφοβου πυρήνα της πρωτεΐνης. Στην **Εικόνα 4.75** παρουσιάζεται ένα δυσδιάστατο γράφημα των χ_1 και χ_2 γωνιών. Με τον τρόπο αυτόν δημιουργούνται clusters (ομάδες) από συνδυασμούς τιμών για τις γωνίες χ_1 και χ_2 που αντιστοιχούν στα κυρίαρχα rotamers.



Εικόνα 4.75 Κατανομή των χ_1 και χ_2 διεδρων γωνιών των καταλοίπων λευκίνης του υδρόφοβου πυρήνα για ολόκληρο το trajectory. Με κόκκινο ρόμβο σημειώνεται το rotamer από το οποίο ξεκίνησε το κατάλοιπο.

Καταρχήν και για αυτό το μόριο, τα rotamer τα οποία παρατηρούνται κατά την διάρκεια της προσομοίωσης έρχονται σε συμφωνία με τα rotamer της λευκίνης που έχουν προσδιορισθεί κρυσταλλογραφικά στις διάφορες δομές. (Εικόνα 4.1). Το γεγονός αυτό συμβάλλει στην αξιοπιστία των αποτελεσμάτων που προκύπτουν από τις προσομοιώσεις. Για λόγους σύγκρισης, στην **Εικόνα 4.76** παρουσιάζεται, αθροιστικά, η κατανομή των τιμών των διεδρων γωνιών για όλα τα κατάλοιπα λευκίνης της $[Ala_2Leu_2]_8$ με *syn* τοπολογία.



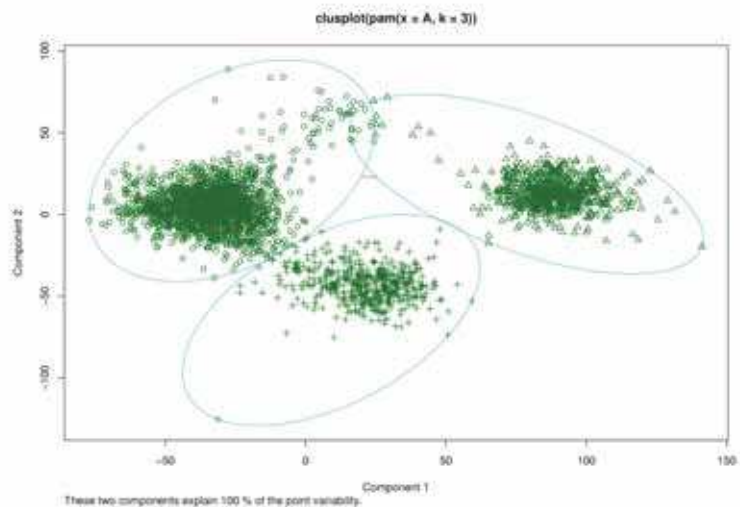
Εικόνα 4.76 Αθροιστική κατανομή των τιμών των διεδρων χ_1 και χ_2 γωνιών για τα κατάλοιπα λευκίνης του υδρόφοβου πυρήνα όπως προέκυψαν από τις προσομοιώσεις.

Και σε αυτό το μόριο, τα κατάλοιπα δεν «παγιδεύτηκαν» στο rotamer από το οποίο ξεκίνησαν (κόκκινος ρόμβος). Βλέπουμε και εδώ περιπτώσεις που ενώ ένα κατάλοιπο ξεκίνησε από το πιο συχνό rotamer, πέρασε και σε ένα λιγότερο συχνό (όπως τα κατάλοιπα 15B, 29A, 29B, 8B), καθώς και το αντίστροφο φαινόμενο, δηλαδή ενώ ξεκίνησε από λιγότερο συχνό rotamer, αυτό δεν το εμπόδισε να περάσει και στο πιο συχνό (όπως τα κατάλοιπα 15A, 22A, 22B, 34A, 41A, 8A). Εδώ βλέπουμε αρκετές περιπτώσεις όπου το αρχικό rotamer ήταν ιδιαίτερα σπάνιο, ωστόσο στη πορεία της προσομοίωσης, τα κατάλοιπα κατέλαβαν τις συνήθεις διαμορφώσεις τους και παρατηρήθηκαν συγκεκριμένα cluster συνδυασμών διεδρων γωνιών (όπως τα κατάλοιπα 34B, 41B, 48A, 48B, 56A). Τέλος, το κατάλοιπο 56B, φαίνεται να συμπεριφέρεται με ένα μοναδικό τρόπο, σε σχέση με όλα τα άλλα κατάλοιπα λευκίνης του υδρόφοβου πυρήνα και των δύο μορίων με τις διαφορετικές τοπολογίες, δημιουργώντας ένα μόνο cluster. Το γεγονός ότι παρατηρούνται ελάχιστες σποραδικές περιπτώσεις που αποκλίνει από το cluster αυτό συνιστά ένδειξη ότι δεν είναι παγιδευμένο σε αυτήν μόνο τη διαμόρφωση. Προκειμένου να διευκρινιστεί ο λόγος αυτής της προτίμησης, δηλαδή εάν το κατάλοιπο δημιουργεί κάποια αλληλεπίδραση με κάποιο άλλο κατάλοιπο ή εάν απλά είναι ενεργειακά

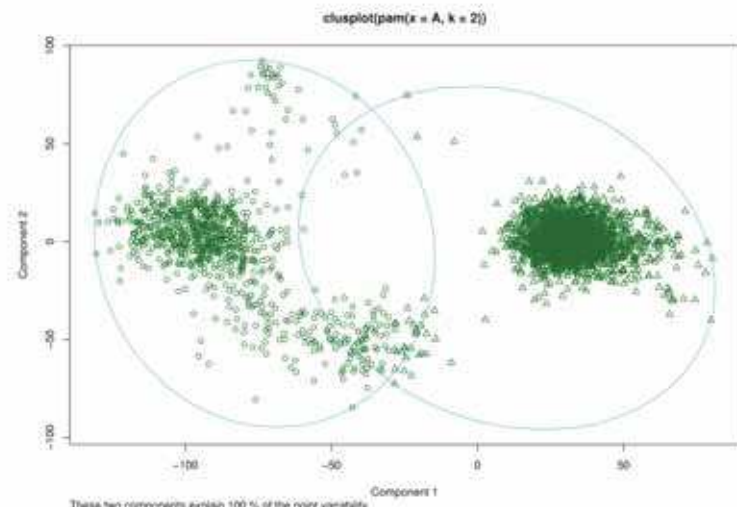
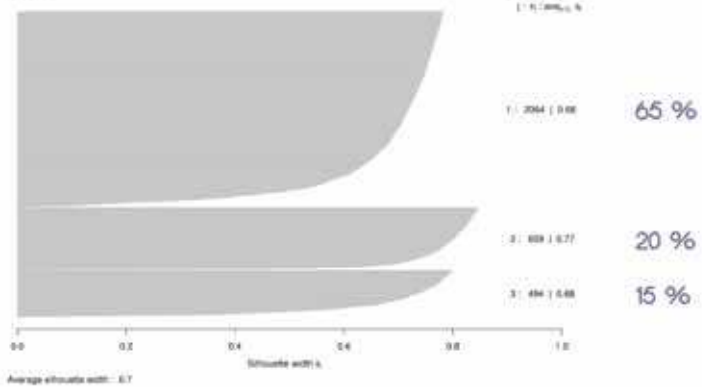
περισσότερο ευνοϊκή η συγκεκριμένη διαμόρφωση, χρειάζεται περισσότερο ενδελεχή μελέτη της συμπεριφοράς του καταλοίπου αυτού κατά τη διάρκεια του trajectory.

Στην Ενότητα 4.1.10, όπου έγινε μελέτη των διέδρων γωνιών της $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία, παρατηρήσαμε πανομοιότυπο πρότυπο στην κατανομή των τιμών για τα ίδια κατάλοιπα στο κάθε μονομερές. Μάλιστα τονίστηκε ότι το φαινόμενο ήταν αρκετά ισχυρό, ώστε να ισχύει για όλα τα κατάλοιπα, παρά το γεγονός ότι σε αυτήν την τοπολογία τα ίδια κατάλοιπα σε διαφορετικά μονομερή είναι απομακρυσμένα στη δομή. Το ίδιο φαινόμενο παρατηρείται και στην $[Ala_2Leu_2]_8$ με *syn* τοπολογία, αλλά είναι λιγότερο ισχυρό και λιγότερο γενικευμένο. Έτσι υπάρχουν περιπτώσεις που παρατηρείται πανομοιότυπο πρότυπο, όπως για παράδειγμα, για τα κατάλοιπα 15A-15B, 34A-34B, 41A-41B, 8A-8B, όπου το πρότυπο διατηρείται ανεξάρτητα από το αρχικό rotamer. Ωστόσο, υπάρχουν και κάποιες διαφορές στο χρόνο παραμονής στο κάθε rotamer (residence times, βλ. παρακάτω). Σε αυτό το μόριο, όμως, μερικά κατάλοιπα εμφανίζουν μεγαλύτερο αριθμό από cluster σε σχέση με το αντίστοιχο στο άλλο μονομερές, όπως για παράδειγμα τα κατάλοιπα 22A-22B, 29A-29B, 48A-48B, 56A-56B. Να υπενθυμίσουμε ότι τα κατάλοιπα αυτά συμμετέχουν στο σχηματισμό του ίδιου επιπέδου σε αυτήν την τοπολογία (βλ. Εικόνα 3.11). Εάν δεχθούμε την υπόθεση των Levy *et al.* ότι στην περίπτωση της $[Ala_2Leu_2]_8$ υπάρχει μίγμα των δύο τοπολογιών τότε αυτό που παρατηρούμε πειραματικά είναι οι μέσες θερμοδυναμικές και κινητικές ιδιότητες. Στην περίπτωση αυτή καθίσταται πιο δύσκολο να διαπιστώσουμε κατά πόσο η συμπεριφορά των καταλοίπων του υδρόφοβου πυρήνα έχει επίδραση στην κινητική ή τη σταθερότητα του μορίου.

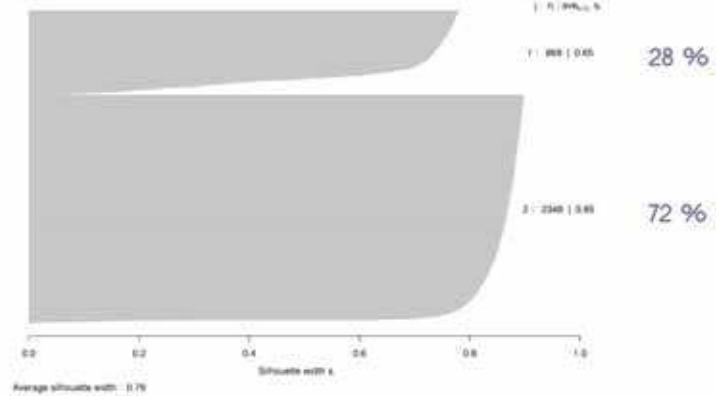
Στις **Εικόνες 4.77 – 4.84** που ακολουθούν περιλαμβάνεται μία στατιστική ανάλυση των τιμών των διέδρων χ_1 και χ_2 γωνιών του κάθε καταλοίπου, ανάλογη με αυτή που παρουσιάστηκε ήδη στην Ενότητα 4.1.10 κατά την ανάλυση των τιμών των διέδρων γωνιών, χ_1 και χ_2 , των καταλοίπων λευκίνης του υδρόφοβου πυρήνα της $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία. Στην κάθε εικόνα παρατίθενται τα αποτελέσματα για το ίδιο κατάλοιπο και για τα δύο μονομερή (αριστερά για το μονομερές A και δεξιά για το μονομερές B), για λόγους σύγκρισης με όσα προαναφέρθηκαν. Επίσης αναφέρονται τα ποσοστά που προκύπτουν για το κάθε cluster. Έτσι έχουμε μία άποψη για τον χρόνο τον οποίο πέρασε το κάθε κατάλοιπο σε κάθε rotamer, το οποίο αντιπροσωπεύεται από ένα cluster (residence times). Τα ποσοστά αυτά θα μπορούσαν να συμβάλλουν σε μία πρόβλεψη για το rotamer το οποίο αναμένεται να παρατηρηθεί κρυσταλλογραφικά ή με φασματοσκοπία πυρηνικού μαγνητικού συντονισμού (NMR, Nuclear Magnetic Resonance spectroscopy). Στην επόμενη ενότητα που γίνεται συγκριτική μελέτη των δύο μορίων με τις διαφορετικές τοπολογίες, περιλαμβάνεται και ένας συγκεντρωτικός πίνακας με τα rotamer που παρατηρούνται για κάθε κατάλοιπο και τους χρόνους παραμονής.



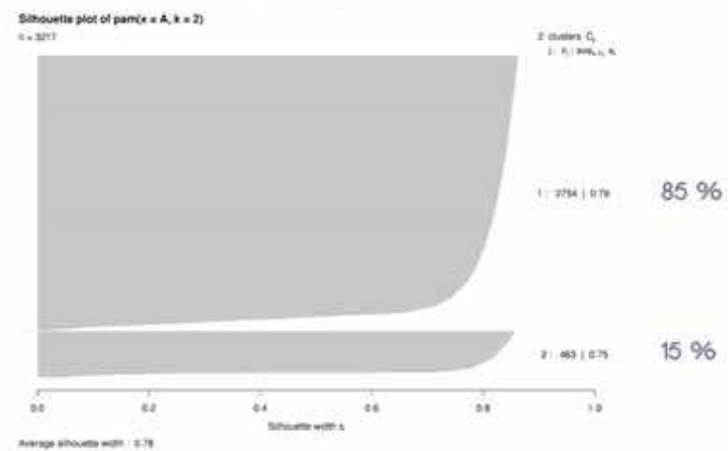
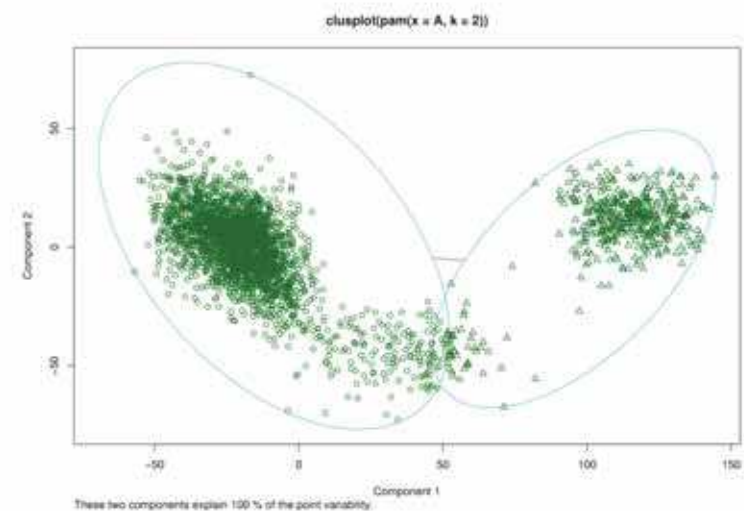
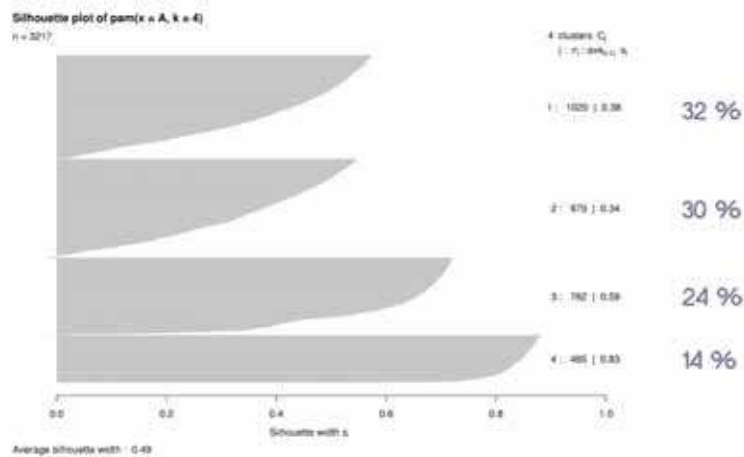
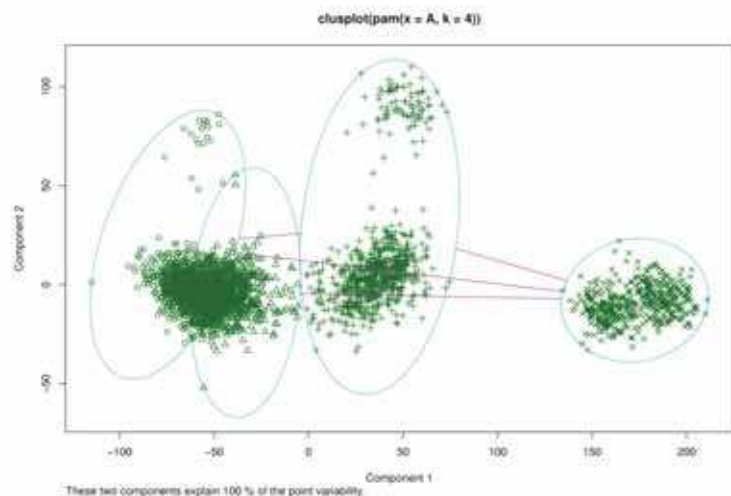
Silhouette plot of pam(x = A, k = 3)
n = 3117



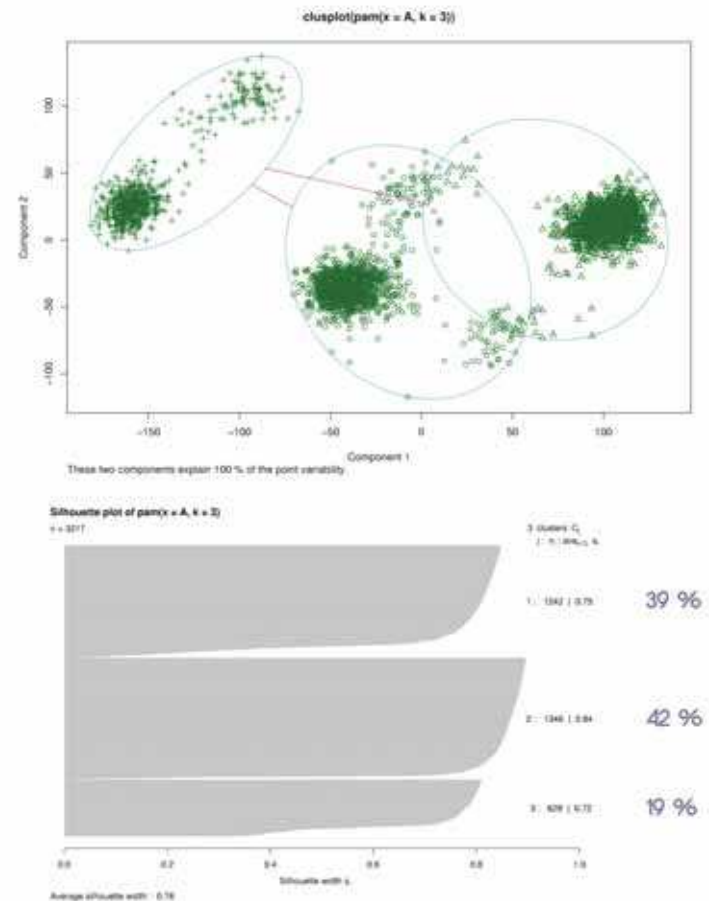
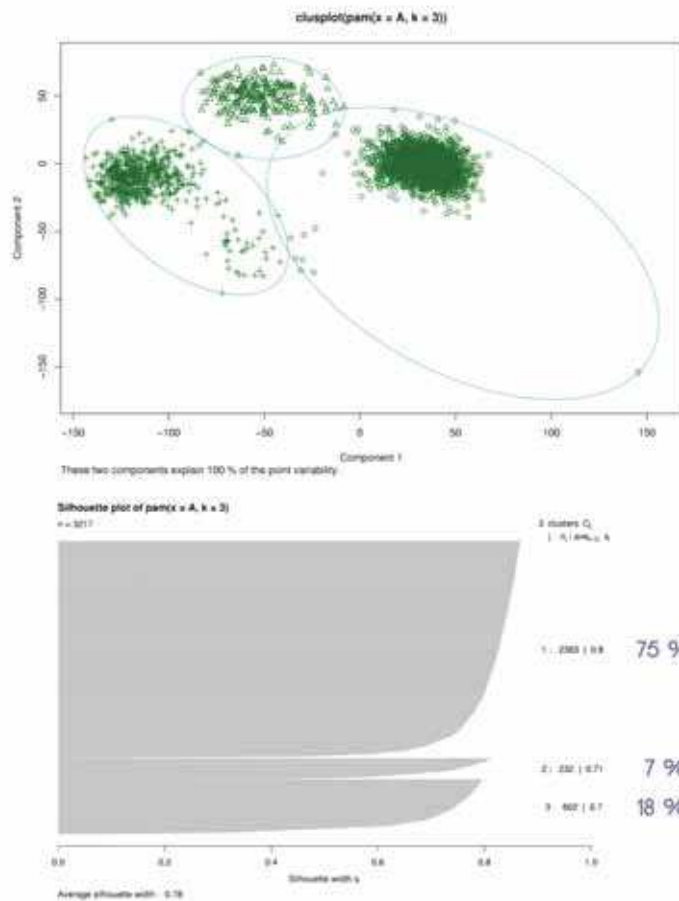
Silhouette plot of pam(x = A, k = 2)
n = 3217



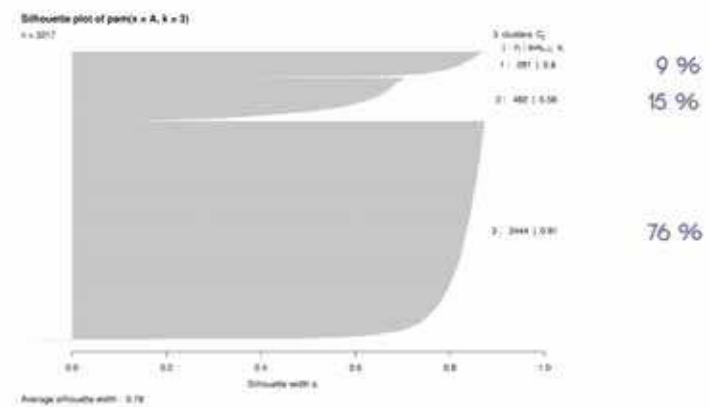
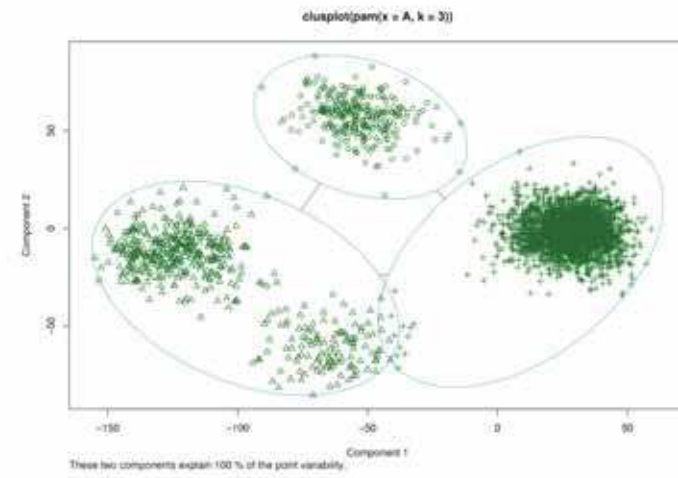
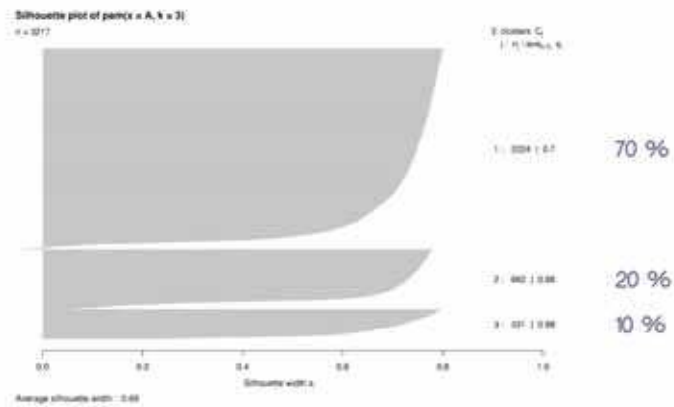
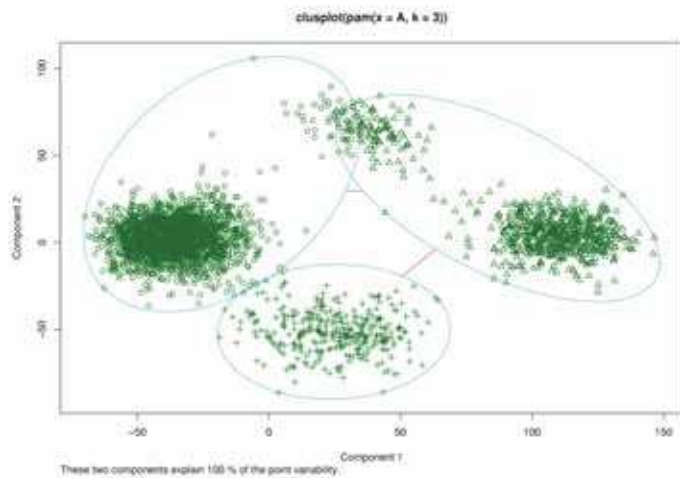
Εικόνα 4.77 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 15A και δεξιά για το κατάλοιπο 15B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 15A και δεξιά για το κατάλοιπο 15B.



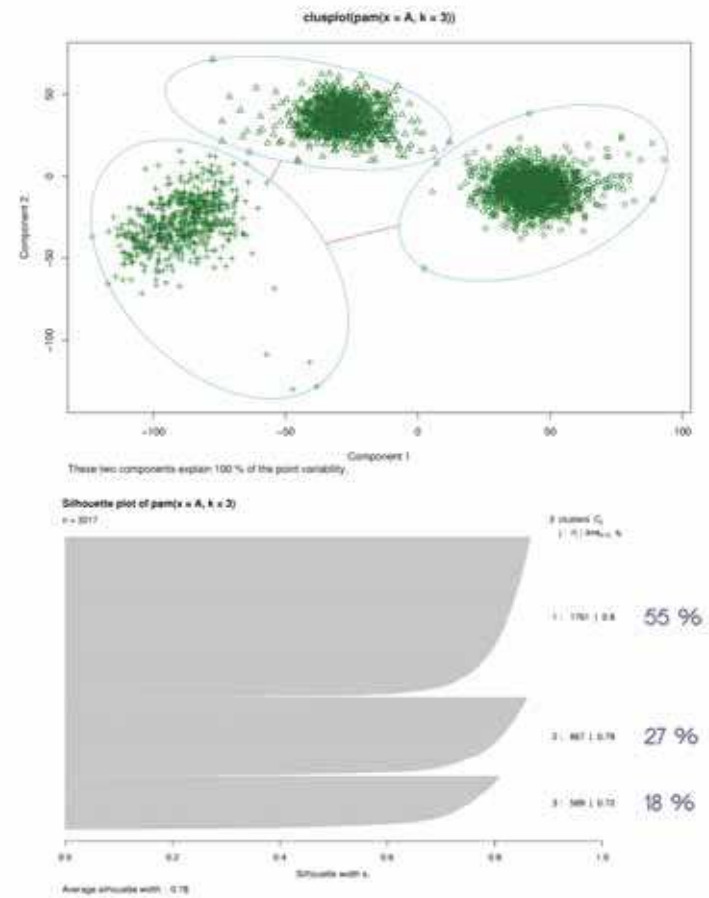
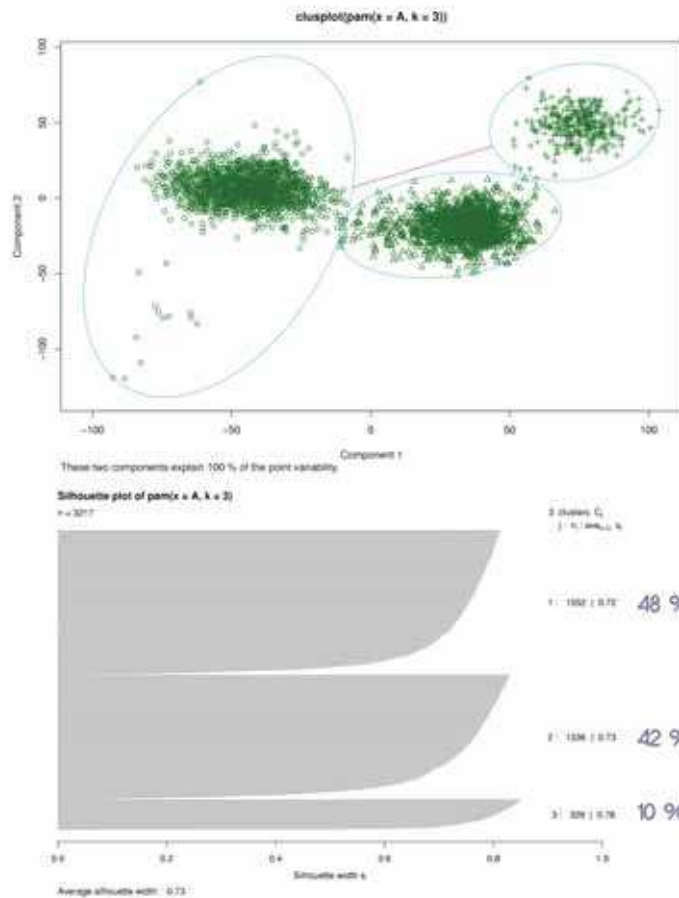
Εικόνα 4.78 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλογο 22A και δεξιά για το κατάλογο 22B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλογο 22A και δεξιά για το κατάλογο 22B.



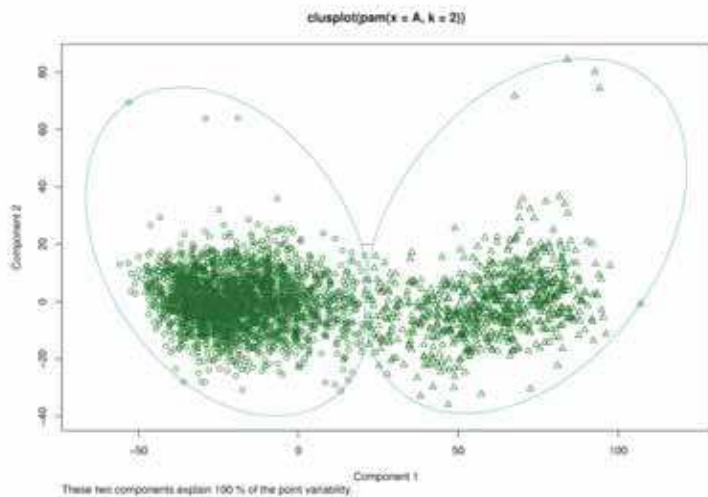
Εικόνα 4.79 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 29A και δεξιά για το κατάλοιπο 29B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 29A και δεξιά για το κατάλοιπο 29B.



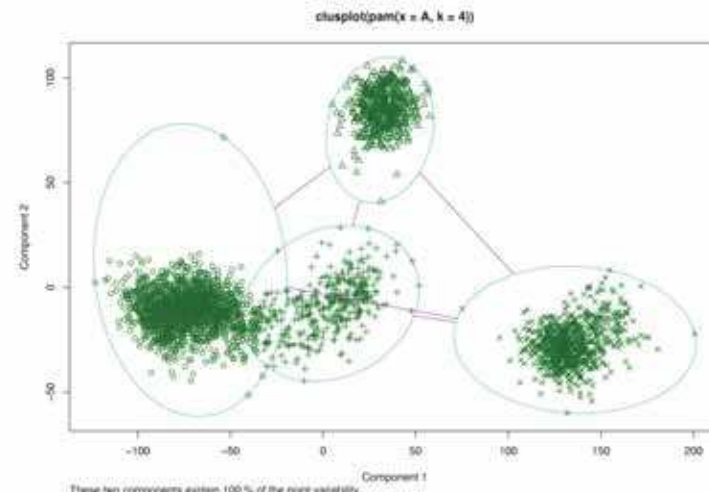
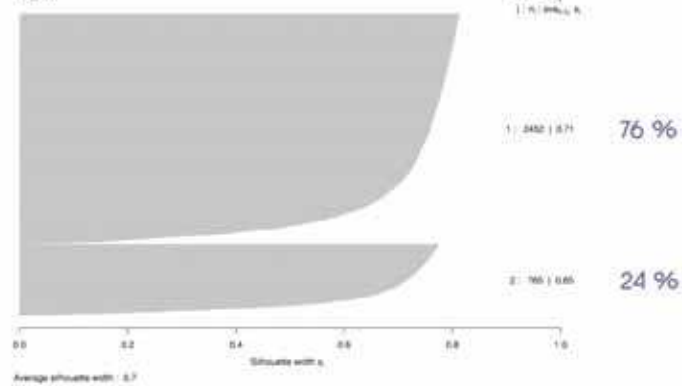
Εικόνα 4.80 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 34A και δεξιά για το κατάλοιπο 34B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 34A και δεξιά για το κατάλοιπο 34B.



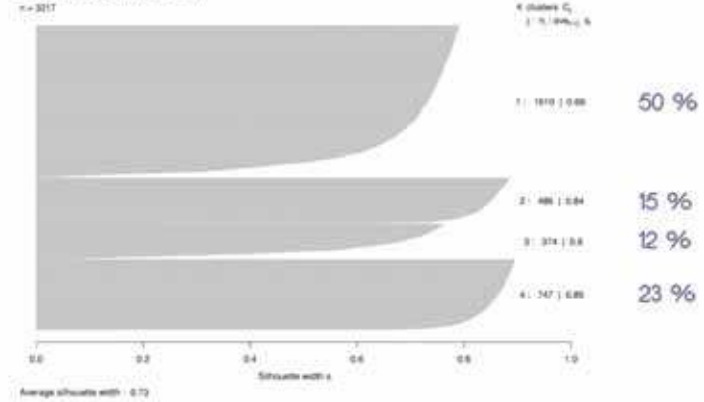
Εικόνα 4.81 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 41A και δεξιά για το κατάλοιπο 41B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 41A και δεξιά για το κατάλοιπο 41B.



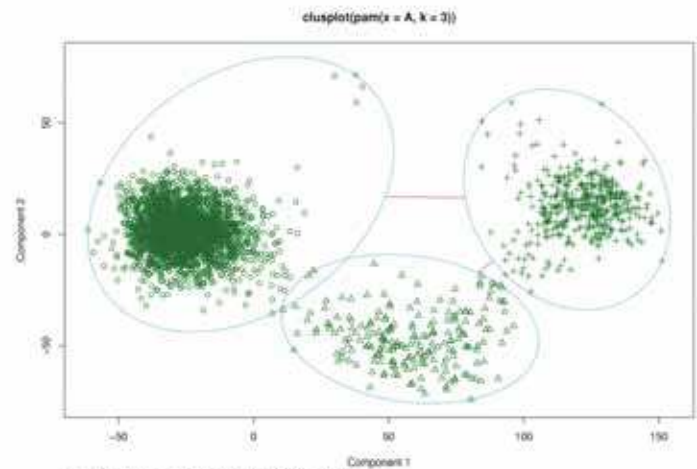
Silhouette plot of pam(x = A, k = 2)



Silhouette plot of pam(x = A, k = 4)

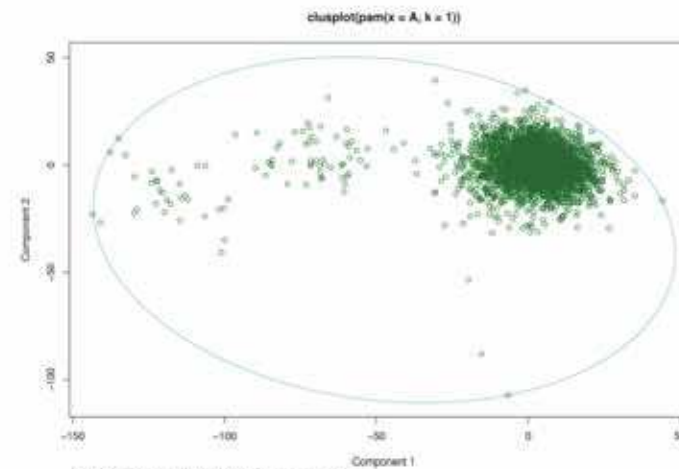
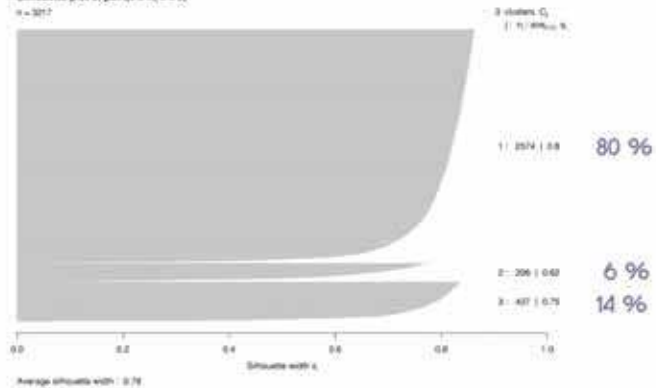


Εικόνα 4.82 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 48A και δεξιά για το κατάλοιπο 48B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 48A και δεξιά για το κατάλοιπο 48B.



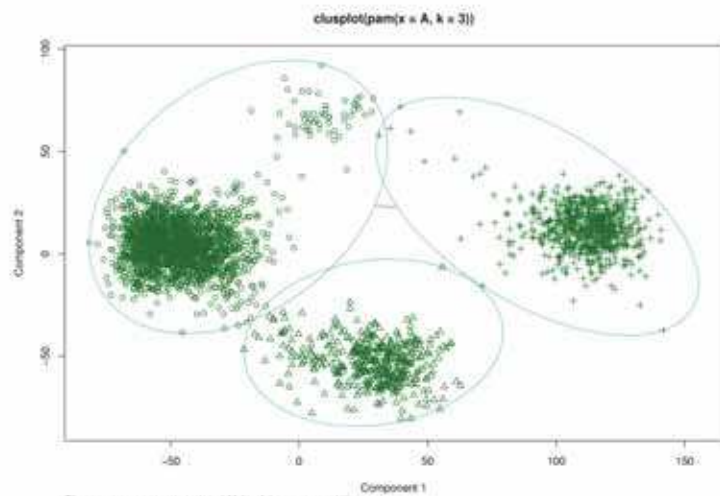
These two components explain 100 % of the point variability.

Silhouette plot of pam(x = A, k = 3)



These two components explain 100 % of the point variability.

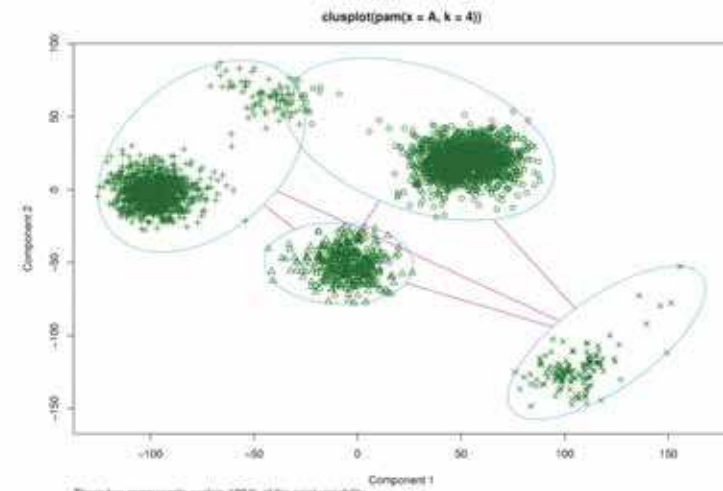
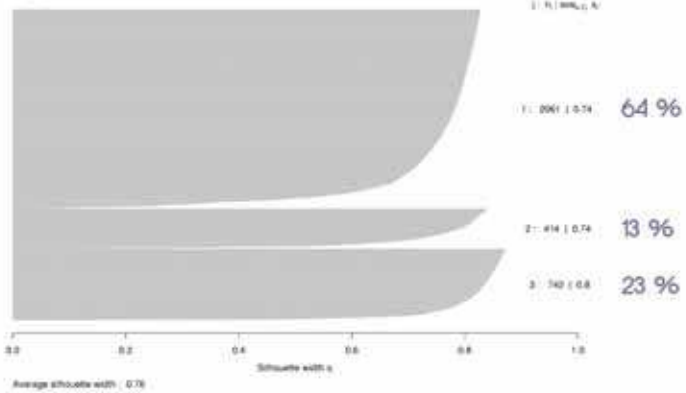
Εικόνα 4.83 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 56A και δεξιά για το κατάλοιπο 56B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 56A και δεξιά για το κατάλοιπο 56B.



These two components explain 100 % of the point variability.

Silhouette plot of pam(x = A, k = 3)

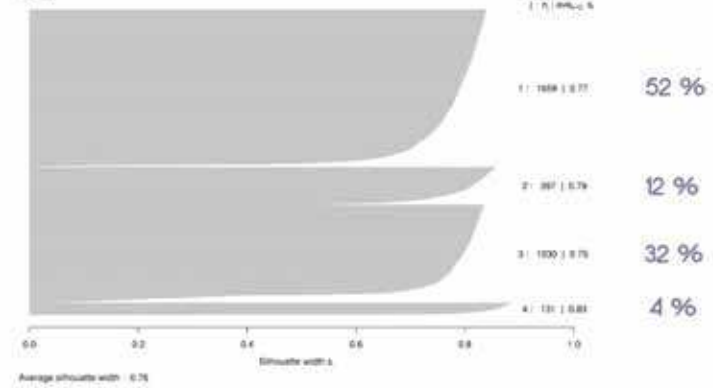
n = 3017



These two components explain 100 % of the point variability.

Silhouette plot of pam(x = A, k = 4)

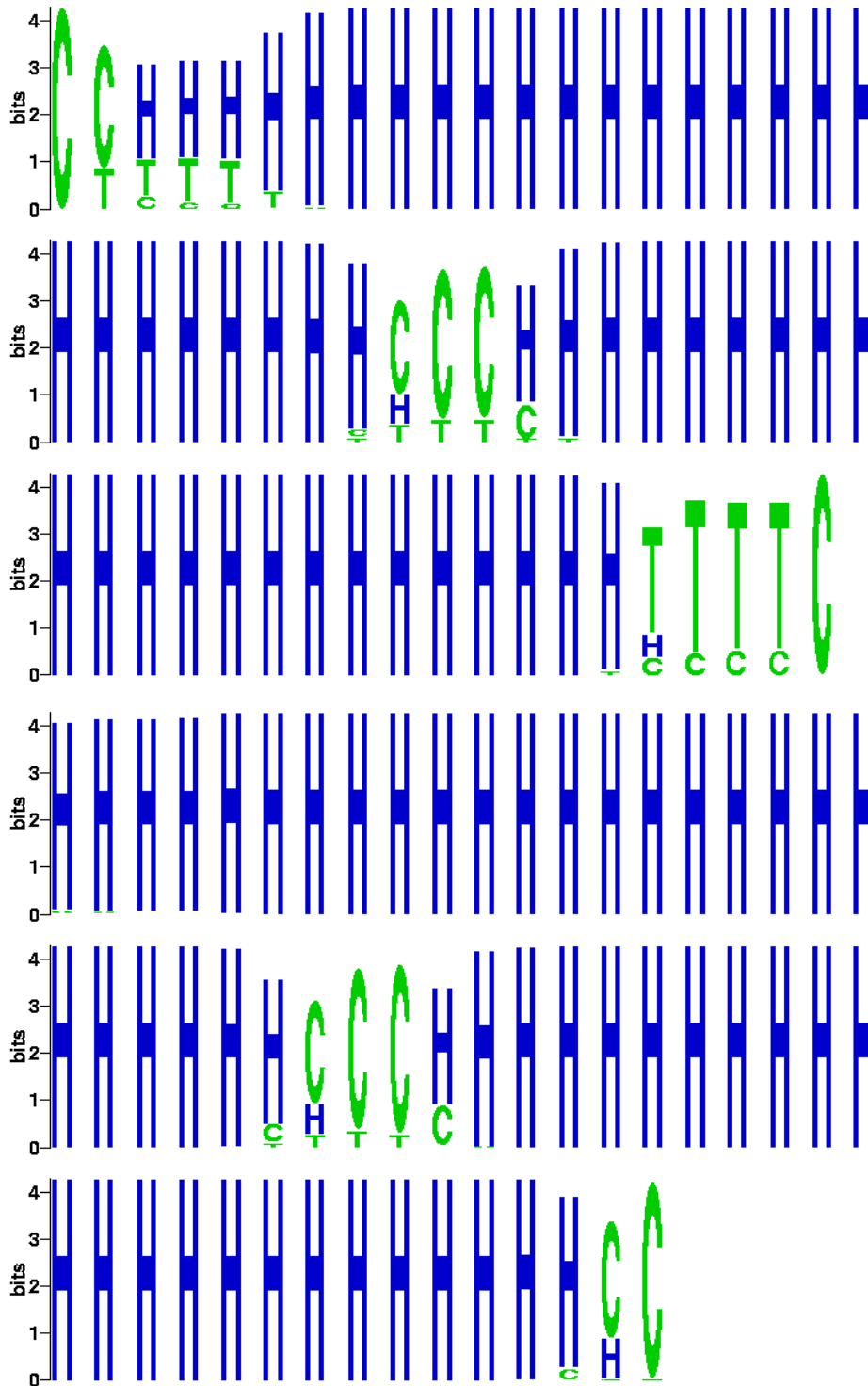
n = 3017



Εικόνα 4.84 Πάνω Προσδιορισμός των clusters τιμών χ_1 και χ_2 γωνιών, αριστερά για το κατάλοιπο 8A και δεξιά για το κατάλοιπο 8B. Κάτω Κατανομή των rotamers και υπολογισμός ποσοστών, αριστερά για το κατάλοιπο 8A και δεξιά για το κατάλοιπο 8B.

4.2.11 Ανάλυση δευτεροταγούς δομής

Είδαμε στην Ενότητα 4.1.11, ότι ένας εναλλακτικός τρόπος για να μελετήσουμε τις ϕ και ψ γωνίες, είναι παρακολουθώντας τη διατήρηση των στοιχείων δευτεροταγούς δομής του μορίου, εφόσον, οι γωνίες αυτές καθορίζουν το πώς συνδέονται τα κατάλοιπα.



Εικόνα 4.85 Μεταβολή της δευτεροταγούς δομής για κάθε θέση της αλληλουχίας.

Μία τέτοια αναπαράσταση με χρήση logo παρουσιάζεται στην **Εικόνα 4.85**, όπου με H (Helix) αναπαρίστανται οι α-έλικες, με C (Coils) οι ελεύθερες θηλιές και με T (Turn) οι στροφές. Κάθε σειρά περιλαμβάνει 20 θέσεις της αλληλουχίας. Βλέπουμε πως οι δομές έλικας διατηρούνται κατά τη διάρκεια ολόκληρου του trajectory και επομένως είναι αρκετά σταθερές. Όπως είναι αναμενόμενα τα αμινοτελικά, τα καρβοξυτελικά κατάλοιπα και τα κατάλοιπα της στροφής και των δύο μονομερών δεν συμμετέχουν στη δημιουργία σταθερών στοιχείων δευτεροταγούς δομής, αλλά παίρνουν πιο ελεύθερες δομές. Πιο συγκεκριμένα, τα 6 πρώτα αμινοτελικά κατάλοιπα του μονομερούς A, αμφιταλαντεύονται ανάμεσα στις τρεις διαμορφώσεις κατά τη διάρκεια του trajectory. Η περιοχή της στροφής (κατάλοιπα 28-32) του μονομερούς A υιοθετεί σε μεγαλύτερο ποσοστό δομή θηλιάς, αλλά σχηματίζει και στροφή (T) όπως και η αντίστοιχη περιοχή του μονομερούς B (κατάλοιπα 27-31). Τα τελικά κατάλοιπα του μονομερούς A (56-60) είναι περισσότερο αποδιαταγμένα σε σχέση με τα τελικά κατάλοιπα του μονομερούς B (56-57), κάτι το οποίο είναι αναμενόμενο δεδομένου ότι για το μονομερές A έχει προσδιορισθεί μεγαλύτερος αριθμός καταλοίπων, τα οποία δεν συμμετέχουν στο σχηματισμό σταθερής δομής.

4.3 Συγκριτική μελέτη ανάμεσα στην $[Ala_2Leu_2]_8$ με native (Rop-like) και με *syn* τοπολογία

Στις προηγούμενες δύο ενότητες περιγράφηκε μία συνολική μελέτη για το κάθε διμερές. Σκοπός της μελέτης δύο ομοδιμερών με διαφορετικές τοπολογίες ήταν, όπως είδαμε στην Ενότητα 3.2, να εξετασθεί η ορθότητα της υπόθεσης που διατυπώθηκε από τους Levy *et al.* αναφορικά με τις τοπολογίες που υιοθετούν οι διάφορες παραλλαγές (variants) της ROP. Στην ενότητα αυτή θα επιχειρηθεί μία σύγκριση των αποτελεσμάτων που προέκυψαν από τις προσομοιώσεις αυτών των συστημάτων. Η σύγκριση δεν θα επεκταθεί σε όλη την ανάλυση που πραγματοποιήθηκε, αλλά θα περιοριστεί σε κάποιες σημαντικές παραμέτρους, ενδεικτικές της σταθερότητας του μορίου.

Η πρώτη ανάλυση που πραγματοποιήθηκε όπως είδαμε, ήταν η μελέτη του rmsd από την αρχική δομή λαμβάνοντας υπόψη τα άτομα του σκελετού του μορίου (Εικόνες 4.2 και 4.47). Για την $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία είδαμε τιμές rmsd από 1.3Å μέχρι 1.5Å, ενώ υπήρχαν δύο κυρίαρχα cluster δομών (47.8% και 47.4%), που το καθένα καταλάμβανε περίπου το μισό του χρόνου της προσομοίωσης (βλ. Εικόνα 4.26). Για την $[Ala_2Leu_2]_8$ με *syn* τοπολογία η τιμή του rmsd κυμαίνεται για το μεγαλύτερο μέρος του χρόνου της προσομοίωσης στα 1.5Å και στη συνέχεια αυξάνεται στα 1.6Å – 1.7Å. Σε γενικές γραμμές εμφανίζει ελαφρώς υψηλότερες τιμές rmsd. Εδώ υπάρχουν τρία σημαντικά cluster δομών (16%, 51% και 33%), με το δεύτερο να είναι το κυρίαρχο (βλ. Εικόνα 4.71).

Όσων αφορά τη σταθερότητα του πυρήνα, μπορούμε να δούμε πως μεταβάλλεται κατά τη διάρκεια του trajectory το rmsd από την αρχική δομή, λαμβάνοντας υπόψη τις πλευρικές ομάδες των καταλοίπων που συνιστούν τα οκτώ επίπεδα του υδρόφοβου πυρήνα της πρωτεΐνης (Εικόνες 4.4 και 4.49). Για την $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία το rmsd κυμαίνεται στα 1.8Å, παρουσιάζοντας μικρές διακυμάνσεις. Για την $[Ala_2Leu_2]_8$ με *syn* τοπολογία αρχικά οι τιμές εμφανίζονται αρκετά υψηλές, γύρω στα 2.5Å και αργούν περισσότερο να σταθεροποιηθούν στα 1.7Å με 1.8Å, ενώ οι διακυμάνσεις είναι περισσότερο έντονες. Και στις δύο περιπτώσεις, ωστόσο, το μονομερές B παρουσιάζει μεγαλύτερη κινητικότητα σε σχέση με το μονομερές A (βλ. Εικόνες 4.5 και 4.50).

Μία ακόμα σημαντική παράμετρος είναι η μεταβολή που παρουσιάζει το rmsd από τη μέση δομή κατά τη διάρκεια του trajectory (Εικόνες 4.6 και 4.51). Όταν η τιμή αυτή είναι σταθερή σημαίνει ότι το μόριο δεν αποκλίνει σημαντικά από τη μέση δομή του και το γεγονός αυτό συνιστά ένδειξη σταθερότητας. Για την $[Ala_2Leu_2]_8$ με native (Rop-like) τοπολογία το rmsd από τη μέση δομή σταθεροποιείται γύρω από την τιμή 0.7Å. Για την $[Ala_2Leu_2]_8$ με *syn* τοπολογία οι

τιμές rmsd είναι επίσης αρκετά χαμηλές, στα 0.6Å – 0.7Å, αλλά οι διακυμάνσεις είναι αρκετά πιο έντονες.

Όσον αφορά το rmsf, φαίνεται ότι είναι κοινές οι περιοχές που εμφανίζουν μεγαλύτερη διακύμανση και στα δύο μόρια, δηλαδή τα κατάλοιπα στα αμινοτελικά και καρβοξυτελικά άκρα καθώς και στις περιοχές των στροφών (Εικόνες 4.9 και 4.54). Βέβαια οι τιμές για την [Ala₂Leu₂]₈ με *syn* τοπολογία είναι σημαντικά υψηλότερες.

Την μεγαλύτερη διαφορά τη συναντάμε στην τιμή του R_G (radius of gyration) (Εικόνες 4.10 και 4.55). Για την [Ala₂Leu₂]₈ με native (Rop-like) τοπολογία η τιμή αυτή κυμαίνεται στα 14.5Å - 14.6Å, ενώ για την [Ala₂Leu₂]₈ με *syn* τοπολογία στα 14.1Å - 14.2Å. Η τιμή αυτή για την *syn* τοπολογία προέκυψε χωρίς να ληφθούν υπόψη τα κατάλοιπα στο αμινοτελικό άκρο. Τα αποτελέσματα διαφοροποιούνται σημαντικά αν ο υπολογισμός γίνει για ολόκληρο το μόριο, δίνοντας τιμές γύρω στα 14.6Å - 14.7Å, δηλαδή και πάλι ελαφρώς υψηλότερες από τις αντίστοιχες της native (Rop-like) τοπολογίας. Ωστόσο, οι αποκλίσεις στις τιμές είναι αρκετά μικρές και δεν υπάρχει κάποια σημαντική ένδειξη ότι μία από τις δύο τοπολογίες δεν μπορεί να οδηγήσει σε σταθεροποιημένο μόριο.

Αξίζει να σημειωθεί ότι βλέποντας αθροιστικά τη συνεισφορά των πρώτων πέντε eigenvector ανά κατάλοιπο και στα δύο μόρια με τις εναλλακτικές τοπολογίες (Εικόνες 4.23 και 4.68) βλέπουμε ότι στις στροφές και στα άκρα των μορίων η συνεισφορά είναι πιο σημαντική. Όμως στην *syn* τοπολογία ο αριθμός των καταλοίπων που συμμετέχει στην κίνηση είναι μεγαλύτερος. Οι μικρότερες τιμές που παρατηρούνται οφείλονται στο γεγονός ότι έχουν παραληφθεί τα ευκίνητα αμινοτελικά κατάλοιπα.

Ένας ακόμη τρόπος για να δούμε ποια από τις δύο τοπολογίες είναι ενδεχομένως πιο σταθερή, είναι υπολογίζοντας την έκταση της επιφάνειας του μορίου που είναι προσβάσιμη (accessible surface area, ASA) στα μονομερή και στα διμερή. Η λογική είναι πως η διαφορά ανάμεσα στο άθροισμα της ASA των δύο μονομερών και της ASA του διμερούς, είναι η επιφάνεια η οποία «θάβεται» (buried surface) κατά τη δημιουργία του συμπλόκου. Όσο περισσότερη επιφάνεια θάβεται, τόσο περισσότερο σταθερό είναι το ομοδιμερές (αφού αυξάνεται η επιφάνεια επαφής των μονομερών). Η παραπάνω διαδικασία πραγματοποιήθηκε για τις μέσες δομές και των δύο ομοδιμερών με τις εναλλακτικές τοπολογίες (**Πίνακας 4.1**). Όλοι οι υπολογισμοί έγιναν με το πρόγραμμα areaimol, από το πακέτο προγραμμάτων CCP4 (Kabsch *et al.*, 1994). Για την [Ala₂Leu₂]₈ με native (Rop-like) τοπολογία υπολογίστηκε μία συνολική επιφάνεια 6452 Å². Η τιμή της ASA για το A μονομερές είναι 4511 Å² και για το B μονομερές 4556 Å². Επομένως η επιφάνεια που θάβεται από την [Ala₂Leu₂]₈ με native (Rop-like) τοπολογία κατά τη δημιουργία του συμπλόκου είναι (4511 + 4556) – 6452 = 2615 Å². Για την [Ala₂Leu₂]₈ με *syn*

τοπολογία υπολογίστηκε μία συνολική επιφάνεια 6597 Å². Η τιμή της ASA για το A μονομερές είναι 4719 Å² και για το B μονομερές 4455 Å². Επομένως η επιφάνεια που θάβεται από την [Ala₂Leu₂]₈ με *syn* τοπολογία κατά τη δημιουργία του συμπλόκου είναι (4719 + 4455) – 6597 = 2577 Å². Βλέπουμε λοιπόν πως και στα δύο διμερή θάβεται σχεδόν η ίδια επιφάνεια, με ελαφρώς υψηλότερη τιμή για την [Ala₂Leu₂]₈ με native (Rop-like) τοπολογία. Να σημειωθεί στο σημείο αυτό ότι η αντίστοιχη επιφάνεια που θάβεται από τη φυσική πρωτεΐνη είναι μόλις 1344.5 Å². Δηλαδή η [Ala₂Leu₂]₈ φαίνεται να είναι αρκετά πιο σταθερή, κάτι το οποίο είναι αναμενόμενο αν σκεφτεί κανείς ότι στο πακετάρισμα του πυρήνα συμμετέχουν περισσότερο υδρόφοβα κατάλοιπα σε σχέση με τη φυσική πρωτεΐνη.

| | ASA A μονομερούς | ASA B μονομερούς | ASA ομοδιμερούς (σύμπλοκο) | Buried Surface |
|---|---------------------|---------------------|----------------------------------|---------------------|
| [Ala ₂ Leu ₂] ₈ με native (Rop-like) τοπολογία | 4511 Å ² | 4556 Å ² | 6452 Å ² | 2615 Å ² |
| [Ala ₂ Leu ₂] ₈ με <i>syn</i> τοπολογία | 4719 Å ² | 4455 Å ² | 6597 Å ² | 2577 Å ² |

Πίνακας 4.1

Πριν ολοκληρώσουμε την σύγκριση των δύο εναλλακτικών τοπολογιών, αξίζει να εξετάσουμε κατά πόσο τα κατάλοιπα του υδρόφοβου πυρήνα που κατέχουν την ίδια θέση, δηλαδή βρίσκονται στο ίδιο επίπεδο, στα δύο μόρια, συμπεριφέρονται με παρόμοιο τρόπο. Στον **Πίνακα 4.2** βλέπουμε την κατανομή των ποσοστών των rotamer για τα κατάλοιπα της λευκίνης του υδρόφοβου πυρήνα, όπως υπολογίστηκαν στην προηγούμενη ενότητα. Η πλειοψηφία των καταλοίπων που καταλαμβάνουν αντίστοιχη θέση σε κάθε επίπεδο στις δύο τοπολογίες δεν φαίνεται να συμπεριφέρεται με παρόμοιο τρόπο. Έτσι στις περισσότερες περιπτώσεις τα αντίστοιχα κατάλοιπα εμφανίζουν τόσο διαφορετικό αριθμό από πιθανά rotamer όσο και διαφορετικούς χρόνους παραμονής. Εξαιρέσεις αποτελούν τα κατάλοιπα 34A – 34A, 15B – 41B και 48A – 48A στις δύο τοπολογίες αντίστοιχα.

| | | [Ala ₂ Leu ₂] ₈ με native (Rop-like) τοπολογία | | [Ala ₂ Leu ₂] ₈ με <i>syn</i> τοπολογία |
|------|------|--|------|---|
| L1 | 29 A | 67.5 % | 29 A | 75 % |
| | | 32.5 % | | 7 % |
| | | | | 18 % |
| | 56 B | 4 % | 29 B | 39 % |
| | | 96 % | | 42 % |
| | | | | 19 % |
| L2 | 34 A | 9.5 % | 34 A | 70 % |
| | | 12.5 % | | 20 % |
| | | 78 % | | 10 % |
| | 8 B | 34.2 % | 34 B | 9 % |
| | | 17.2 % | | 15 % |
| | | 48.6 % | | 76 % |
| L3 | 22 A | 38 % | 22 A | 32 % |
| | | 8 % | | 30 % |
| | | 54 % | | 24 % |
| | 48 B | 18 % | 22 B | 14 % |
| | | 82 % | | 85 % |
| | | | | 15 % |
| L4 | 41 A | 57 % | 41 A | 55 % |
| | | 12 % | | 27 % |
| | | 8 % | | 18 % |
| | | 23 % | | |
| | 15 B | 35 % | 41 B | 55 % |
| | | 46 % | | 27 % |
| 19 % | | 18 % | | |
| L5 | 15 A | 46.7 % | 15 A | 65 % |
| | | 30 % | | 20 % |
| | | 23.3 % | | 15 % |
| | 41 B | 4 % | 15 B | 28 % |
| | | 87 % | | |
| | | 3 % | | |
| | | 6 % | | |
| | | | | 72 % |

| | | | | |
|----|------|--------|------|-------|
| L6 | 48 A | 84 % | 48 A | 76 % |
| | | 16 % | | 24 % |
| | 22 B | 48.5 % | 48 B | 50 % |
| | | 44.5 % | | 15 % |
| | | 7 % | | 12 % |
| | | | | 23 % |
| L7 | 8 A | 62 % | 8 A | 64 % |
| | | 38 % | | 13 % |
| | | | | 23 % |
| | 34 B | 3.5 % | 8 B | 52 % |
| | | 90 % | | 12 % |
| | | 6.5 % | | 32 % |
| L8 | 56 A | 5 % | 56 A | 80 % |
| | | 95 % | | 6 % |
| | | | | 14 % |
| | 29 B | 10 % | 56 B | 100 % |
| | | 5 % | | |
| | | 85 % | | |

Πίνακας 4.2 Ποσοστιαία κατανομή των rotamer των καταλοίπων λευκίνης του υδρόφοβου πυρήνα για τα δύο μόρια με τις εναλλακτικές τοπολογίες. Η σειρά των καταλοίπων ακολουθεί τη σειρά που κατέχουν στη δομή (δηλαδή το επίπεδο στο οποίο ανήκουν).

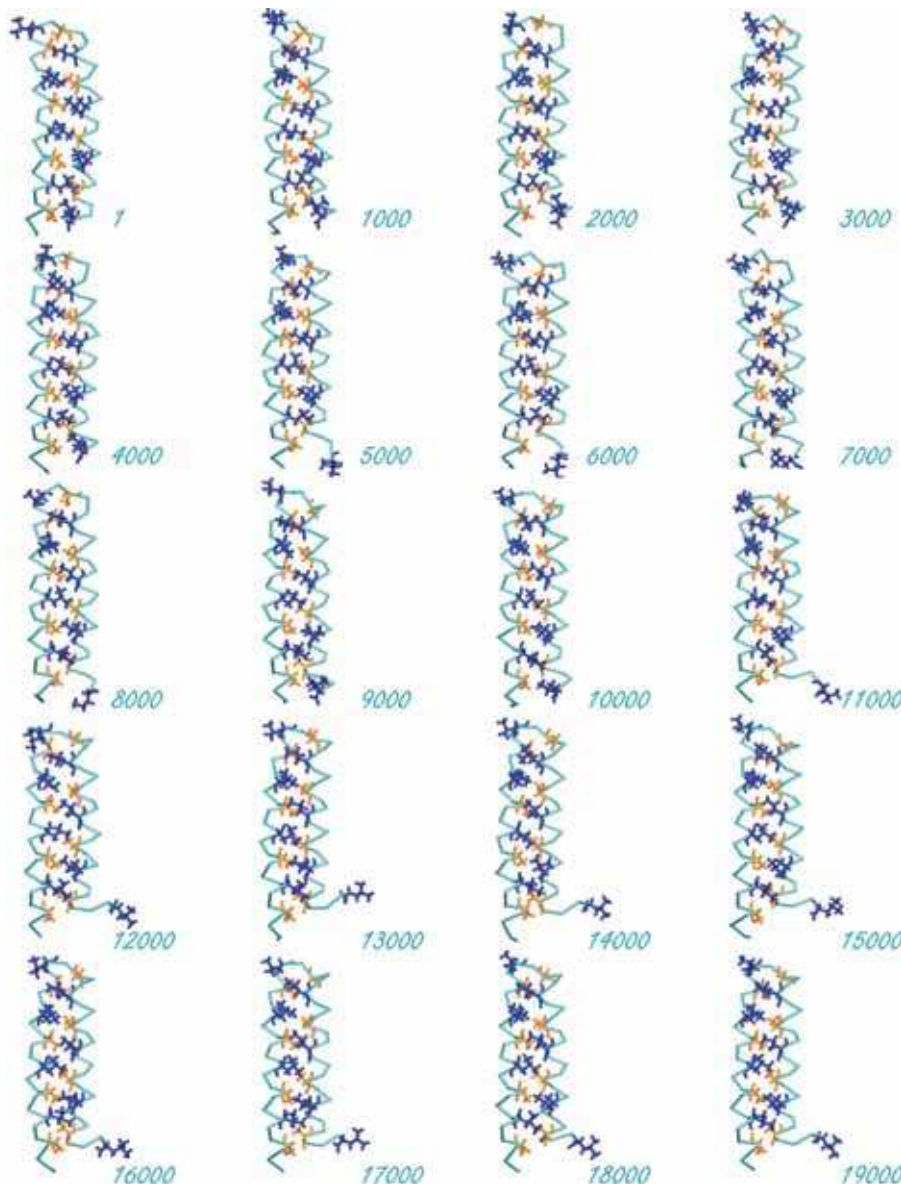
Επομένως από την μέχρι τώρα ανάλυση, μπορούμε να ισχυριστούμε ότι, η υπόθεση των Levy *et al.*, ότι οι παραλλαγές της ROP που έχουν περιγραφεί κατά καιρούς μπορούν να λάβουν οποιαδήποτε από τις δύο εναλλακτικές διαμορφώσεις, μπορεί να ισχύει, τουλάχιστον για την περίπτωση της $[Ala_2Leu_2]_8$, καθώς δεν υπάρχει κάποια ισχυρή ένδειξη που να υποδεικνύει κάποια «προτίμηση» προς τη μία τοπολογία, βασιζόμενοι στις προσομοιώσεις μοριακής δυναμικής.

4.5 Ανάλυση προσομοίωσης *native* μονομερούς

Εξαιτίας της δυσαναλογίας ανάμεσα στο διαθέσιμο χρόνο και τον όγκο των υπολογιστικών δεδομένων, λεπτομερής ανάλυση πραγματοποιήθηκε μόνο για τα δύο

συστήματα των διμερών. Η ανάλυση για τις προσομοιώσεις των μονομερών θα περιοριστεί σε απλή παρουσίαση αντιπροσωπευτικών στιγμιότυπων από τα trajectory στις ενότητες που ακολουθούν.

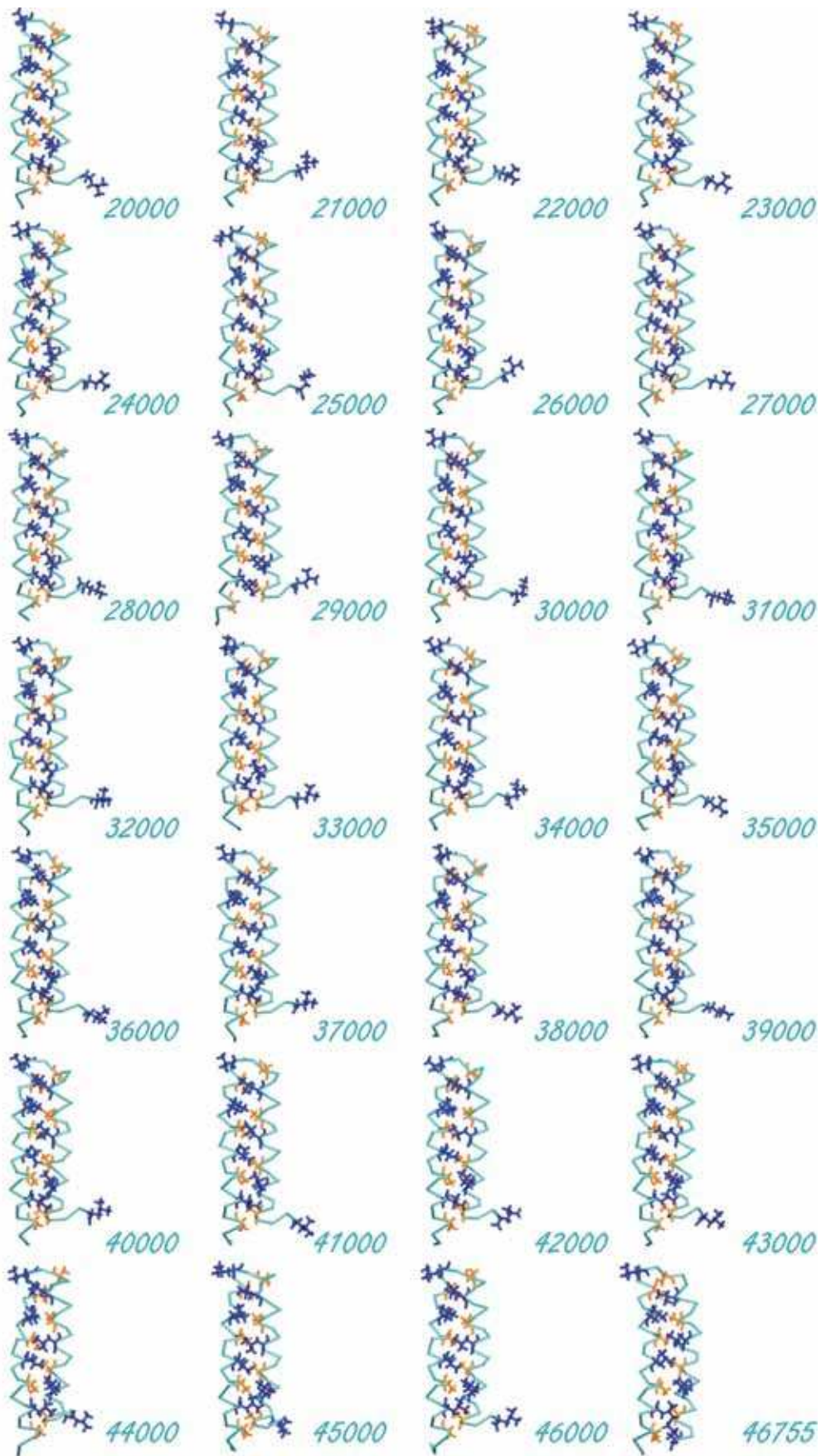
Το *native* μονομερές δεν επιδεικνύει ιδιαίτερη απόκλιση από τη δομή από την οποία ξεκίνησε (**Εικόνα 4.76**). Οι πλευρικές ομάδες των καταλοίπων που συνιστούν τα οκτώ επίπεδα του υδρόφοβου πυρήνα της πρωτεΐνης είναι αρκετά σταθερές, με εξαίρεση τη λευκίνη 29, στη στροφή του μονομερούς αλλά κυρίως τη λευκίνη 56, στο καρβοξυτελικό άκρο, η οποία είναι ιδιαίτερα ευκίνητη. Ωστόσο, μέχρι το τέλος του χρόνου της προσομοίωσης (20ns), οι πλευρικές ομάδες παρέμειναν στη μορφή που τις συναντάμε στο ομοδιμερές, δηλαδή στραμμένες προς το άλλο μονομερές. Επομένως, το μονομερές, στη *native* μορφή του, δηλαδή όπως το συναντάμε στο λειτουργική μονάδα του διμερούς, δεν μπορεί να σχηματίσει από μόνο του ένα ελικοειδές σπείραμα (coiled coil).



Εικόνα 4.76

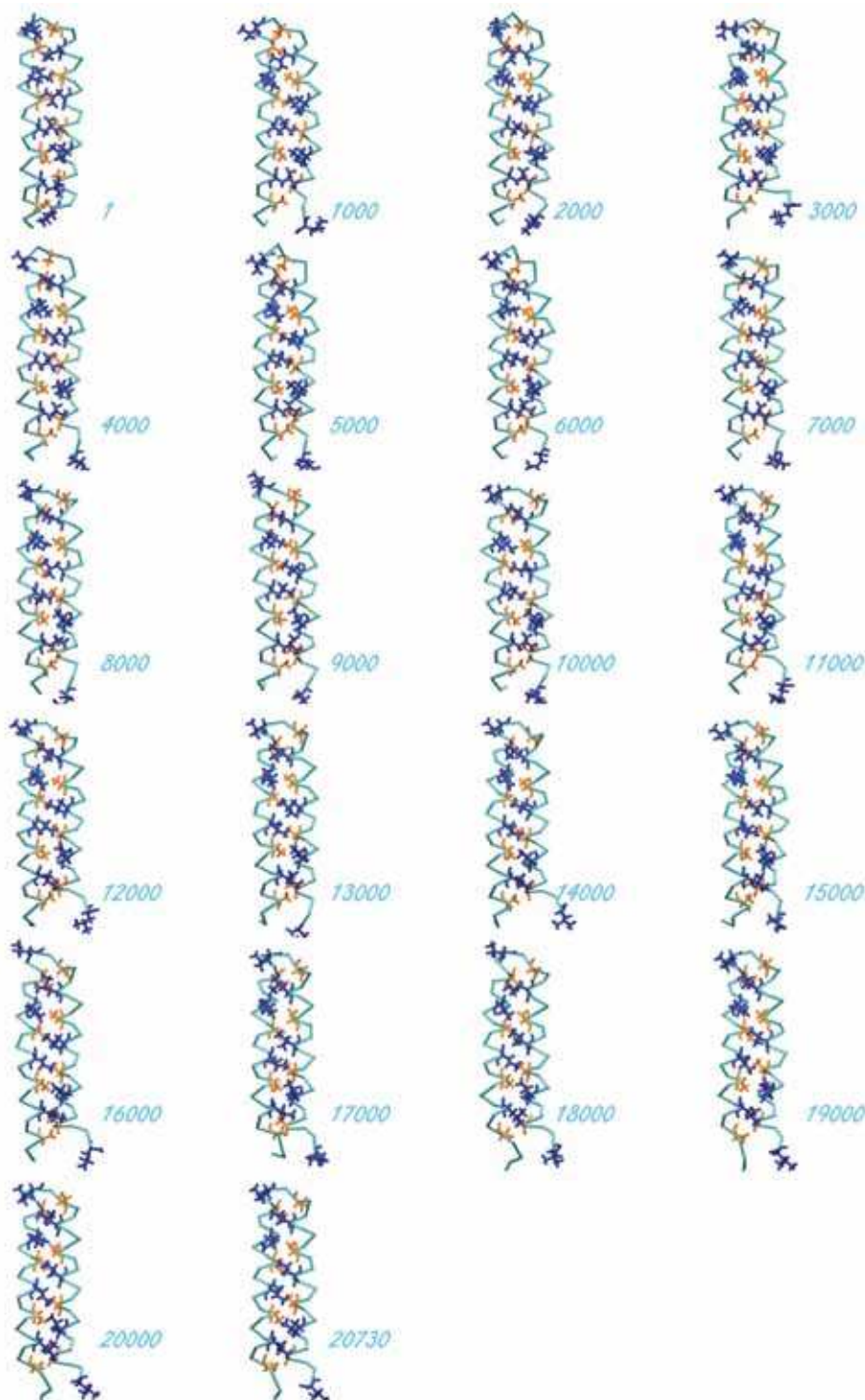
Αντιπροσωπευτικά στιγμιότυπα της προσομοίωσης του *native* μονομερούς. Κάτω δεξιά σε κάθε στιγμιότυπο αναφέρεται ο χρόνος (σε frame). Ο σκελετός του μορίου απεικονίζεται με wireframe. Οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα απεικονίζονται με sticks. Με πορτοκαλί φαίνονται τα κατάλοιπα *Ala* και με μπλε τα κατάλοιπα *Leu*.

Συνέχεια στην επόμενη σελίδα



4.6 Ανάλυση προσομοίωσης *artificial* μονομερούς

Και στην περίπτωση του *artificial* μονομερούς δεν παρατηρείται κάποια ιδιαίτερη απόκλιση σε σχέση με την αρχική δομή (Εικόνα 4.77). Η κινητικότητα που επιδεικνύει οφείλεται κυρίως στα κατάλοιπα που συμμετέχουν στο πρώτο και όγδοο επίπεδο. Και πάλι οι πλευρικές ομάδες παρέμειναν σε σημαντικό βαθμό κοντά στην αρχική τους διαμόρφωση.



Εικόνα 4.77

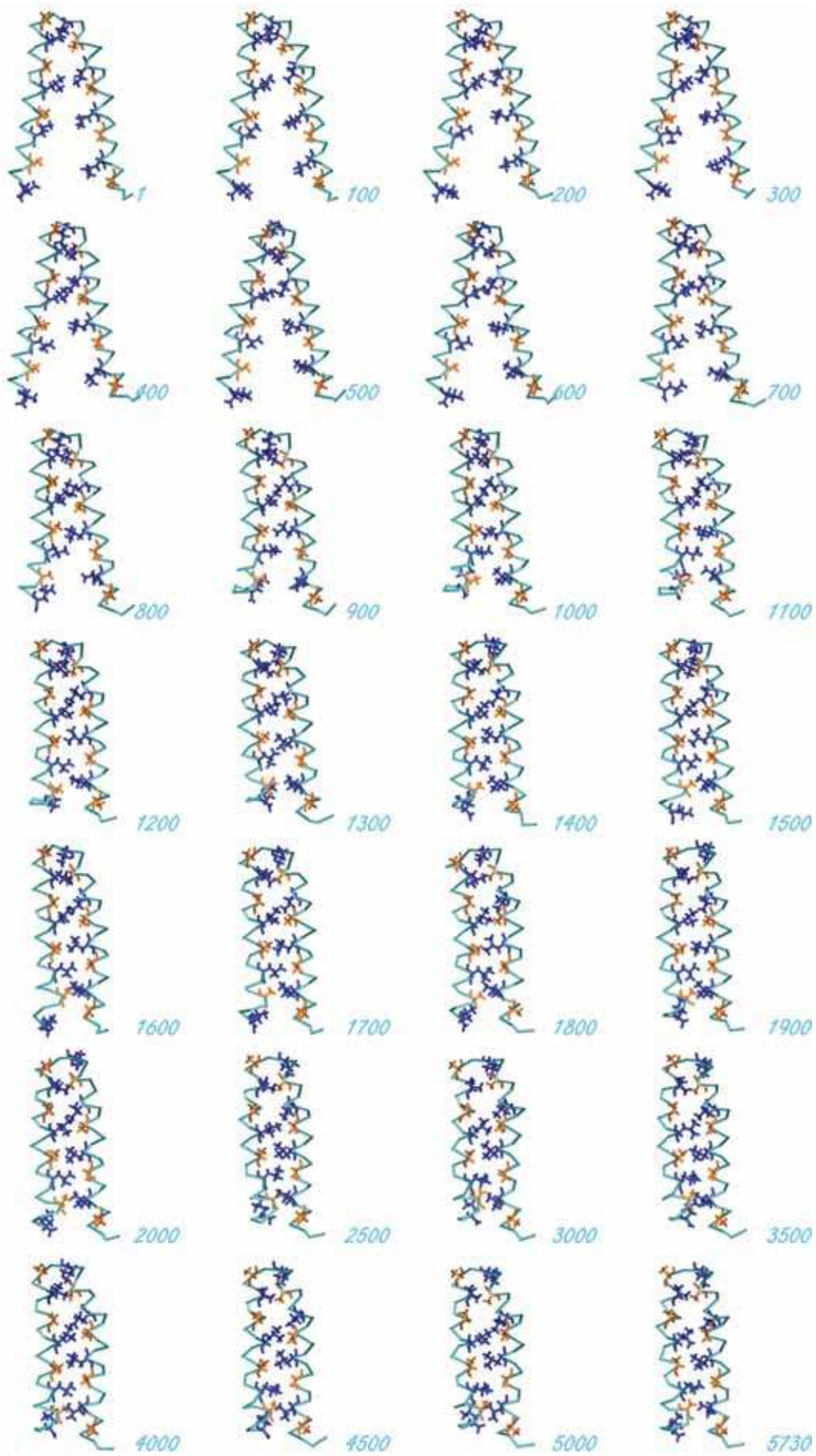
Αντιπροσωπευτικά στιγμιότυπα της προσομοίωσης του *artificial* μονομερούς. Κάτω δεξιά σε κάθε στιγμιότυπο αναφέρεται ο χρόνος (σε frame). Ο σκελετός του μορίου απεικονίζεται με wireframe. Οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα απεικονίζονται με sticks. Με πορτοκαλί φαίνονται τα κατάλοιπα *Ala* και με μπλε τα κατάλοιπα *Leu*.

4.7 Ανάλυση προσομοίωσης *opened* μονομερούς

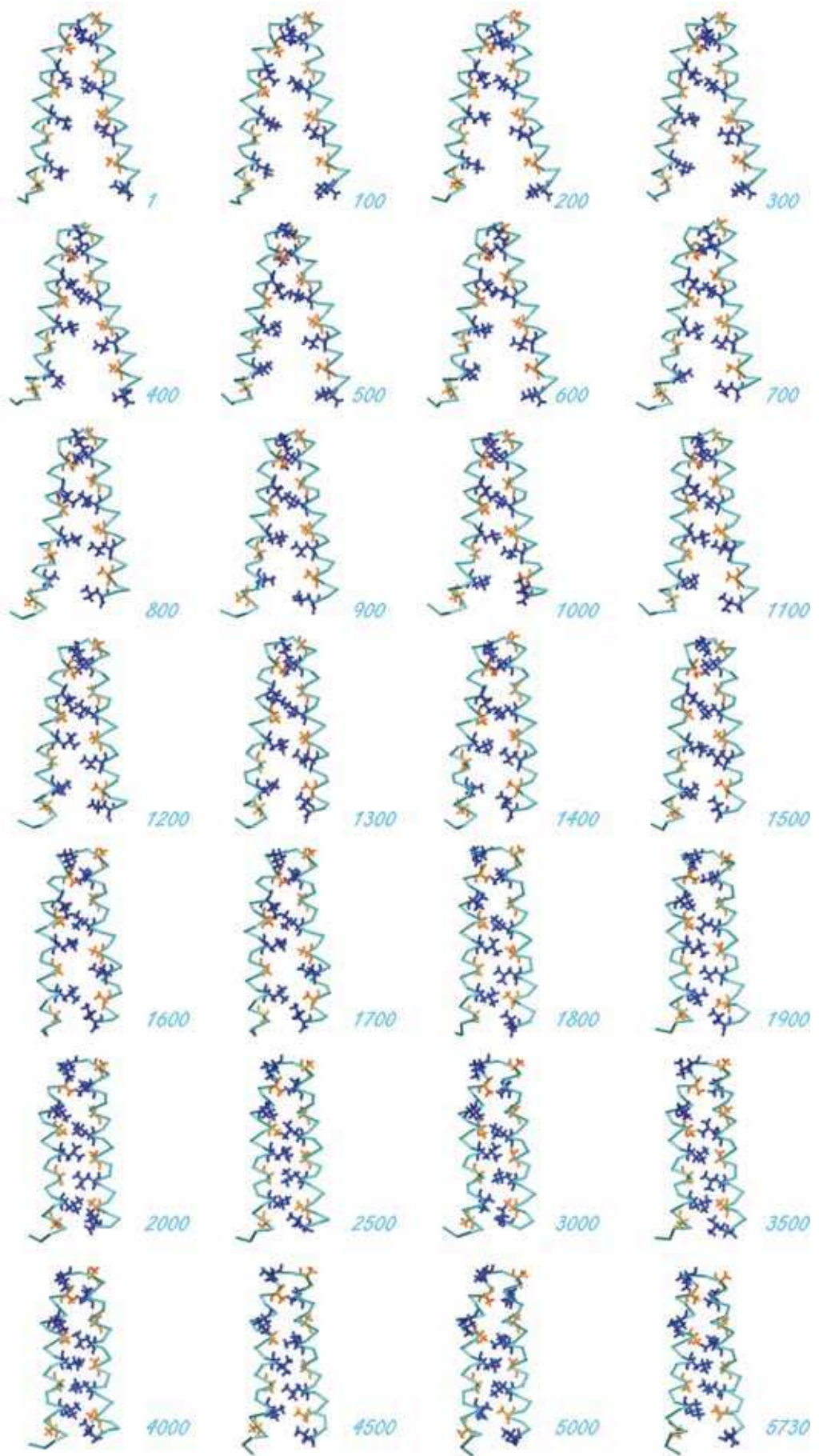
Όπως είδαμε στην Ενότητα 3.7, η A31P τοπολογία υιοθετεί μία αρκετά ανοικτή διαμόρφωση, αφήνοντας με αυτόν τον τρόπο αρκετό χώρο στις πλευρικές ομάδες για να κινηθούν ελεύθερα. Οι πρώτες έξι προσομοιώσεις που πραγματοποιήθηκαν ήταν σύντομης διάρκειας (2ns). Ο σκοπός ήταν να μελετήσουμε κατά πόσο το αποτέλεσμα που παρατηρούμε είναι επαναλήψιμο (**Εικόνες 4.78 – 4.83**).

Το πιο αξιοσημείωτο χαρακτηριστικό όλων αυτών των προσομοιώσεων, ήταν ότι οι πλευρικές ομάδες των καταλοίπων που συμμετέχουν στο σχηματισμό των επιπέδων του υδρόφοβου πυρήνα στη native ROP, κινούνται πολύ γρήγορα. Έτσι οι δύο έλικες πλησιάζουν γρήγορα μεταξύ τους και οι πλευρικές ομάδες αλληλεπιδρούν (όπως και στα επίπεδα της native πρωτεΐνης), σταθεροποιώντας το μονομερές σε αυτή τη διαμόρφωση. Η κίνηση αυτή, αν λάβουμε υπόψη την κατανομή του φορτίου στην κάθε έλικα με βάση τα κατάλοιπα από τα οποία συνίσταται, φαίνεται να είναι το αποτέλεσμα ηλεκτροστατικών αλληλεπιδράσεων. Ήδη από τη φάση του heating (τα πρώτα 730 frame), βλέπουμε πως τα πέντε πρώτα «επίπεδα» έχουν σχηματιστεί. Στην πρώτη προσομοίωση, στα 1000 frame φαίνεται να έχουν «κλείσει» και τα οκτώ επίπεδα, ενώ στις επόμενες αυτό συμβαίνει λίγο αργότερα γύρω στα 1300 – 1500 frame. Ακολουθώντας και μέχρι το τέλος του χρόνου της προσομοίωσης, το μονομερές παραμένει σταθεροποιημένο σε αυτή τη διαμόρφωση. Σε κάποια στιγμιότυπα ορισμένων προσομοιώσεων, οι έλικες εμφανίζουν μία αλλοίωση στη δομή τους, όπως για παράδειγμα στην Εικόνα 4.78, frame 3000 – 4000 και στην Εικόνα 4.81, frame 3000 – 4000. Στην Εικόνα 4.80, στο frame 3500 βλέπουμε ένα άνοιγμα (κοιλότητα), λόγω αδυναμίας δημιουργίας του τρίτου «επιπέδου», το οποίο διατηρείται μέχρι το τέλος. Στην πέμπτη προσομοίωση (Εικόνα 4.82), τα δύο τελευταία επίπεδα δε φαίνεται να διατηρούνται μέχρι το τέλος. Στην έκτη από αυτήν τη σειρά των προσομοιώσεων, που διαρκεί λίγο περισσότερο (2.7ns), βλέπουμε μία αλλοίωση στη δομή των ελίκων που οδηγεί σε έντονη κυρτότητα του μορίου, από το frame 2000 και μέχρι το τέλος.

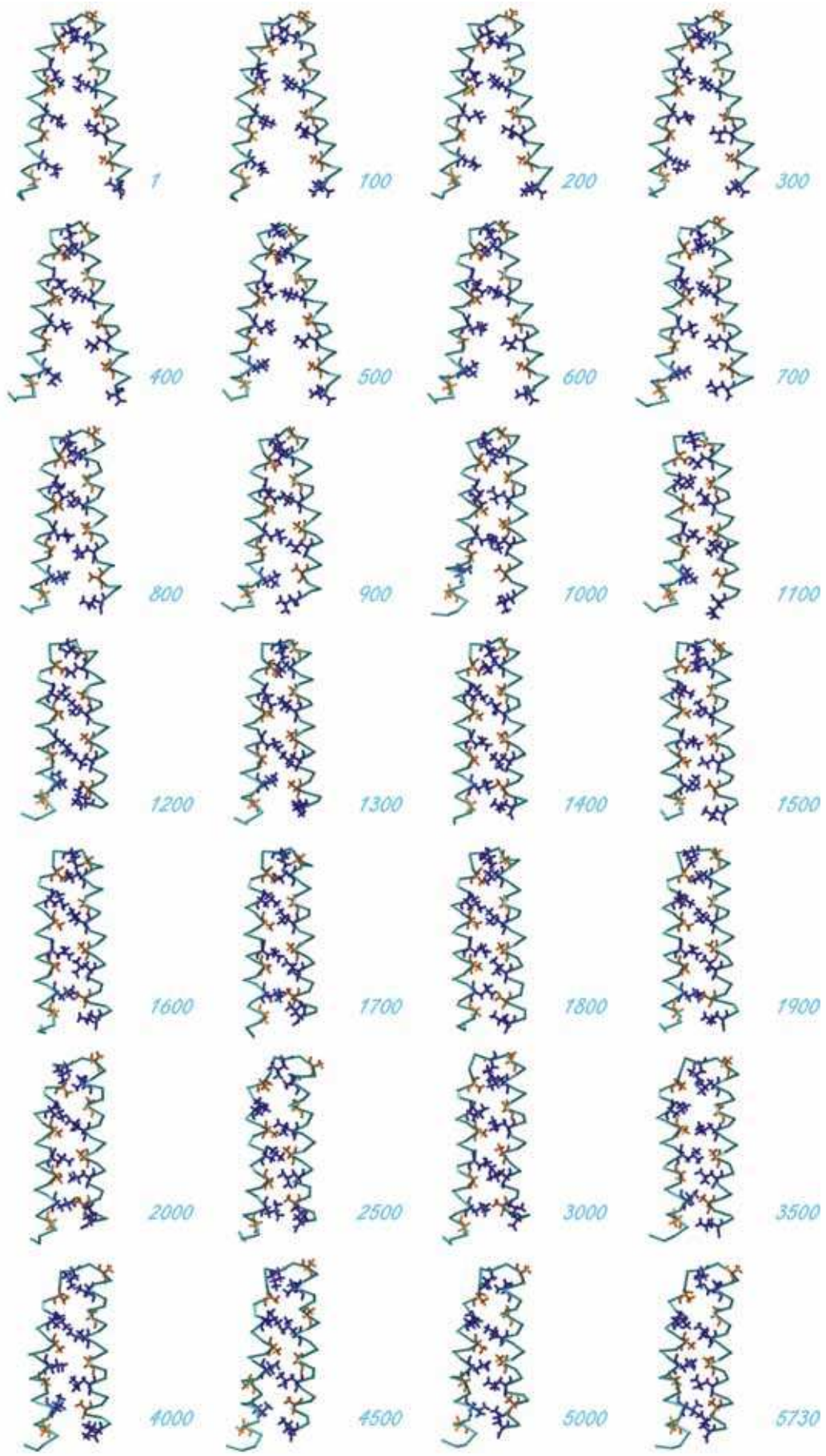
Αυτήν την αλλοίωση της δομής των ελίκων και την έντονη κυρτότητα μπορούμε να την παρακολουθήσουμε καλύτερα στην τελευταία και μεγάλης διάρκειας (10ns) προσομοίωση του συστήματος αυτού (**Εικόνα 4.84**). Στα πρώτα 1000 frame έχουν ήδη σχηματιστεί τα περισσότερα επίπεδα, με εξαίρεση το όγδοο, λόγω της μεγάλης κινητικότητας της λευκίνης 56. Από το frame 2000 και μετά αρχίζει να είναι εμφανής η αλλοίωση της δομής των ελίκων που οδηγεί σε κυρτότητα του μορίου, ενώ απουσιάζει και εδώ το τρίτο επίπεδο. Στην πορεία της προσομοίωσης βλέπουμε πως χάνεται και το έκτο επίπεδο (frame 3500 – 5000). Οι πλευρικές ομάδες των καταλοίπων που συνιστούν τα επίπεδα είναι αρκετά ευκίνητες και τα επίπεδα δεν μπορούν να σχηματιστούν. Αυτήν την συμπεριφορά το μόριο την υιοθετεί μέχρι το τέλος της προσομοίωσης, με αποτέλεσμα η δομή του να μην διατηρείται σταθερή.



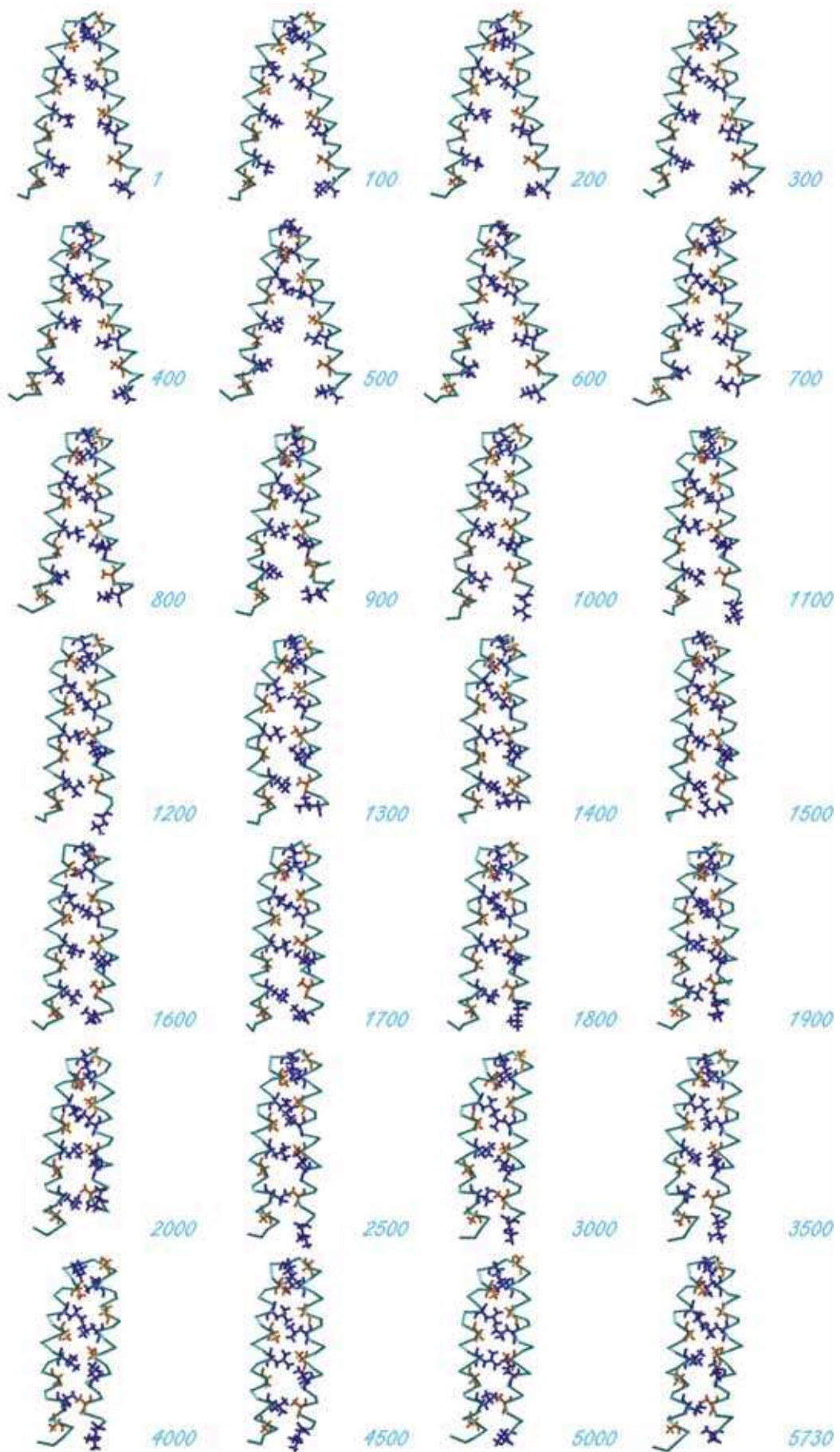
Εικόνα 4.78



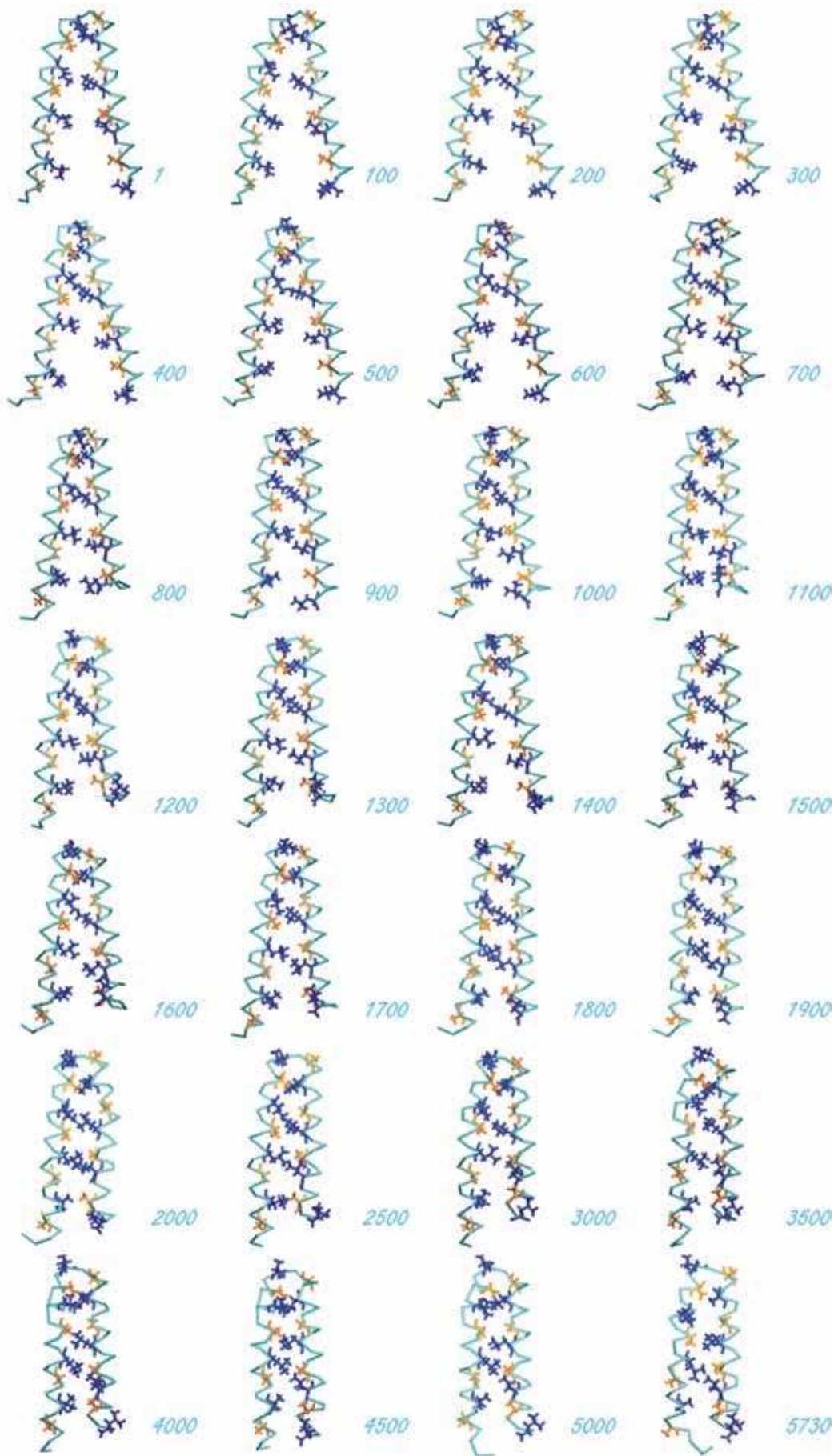
Εικόνα 4.79



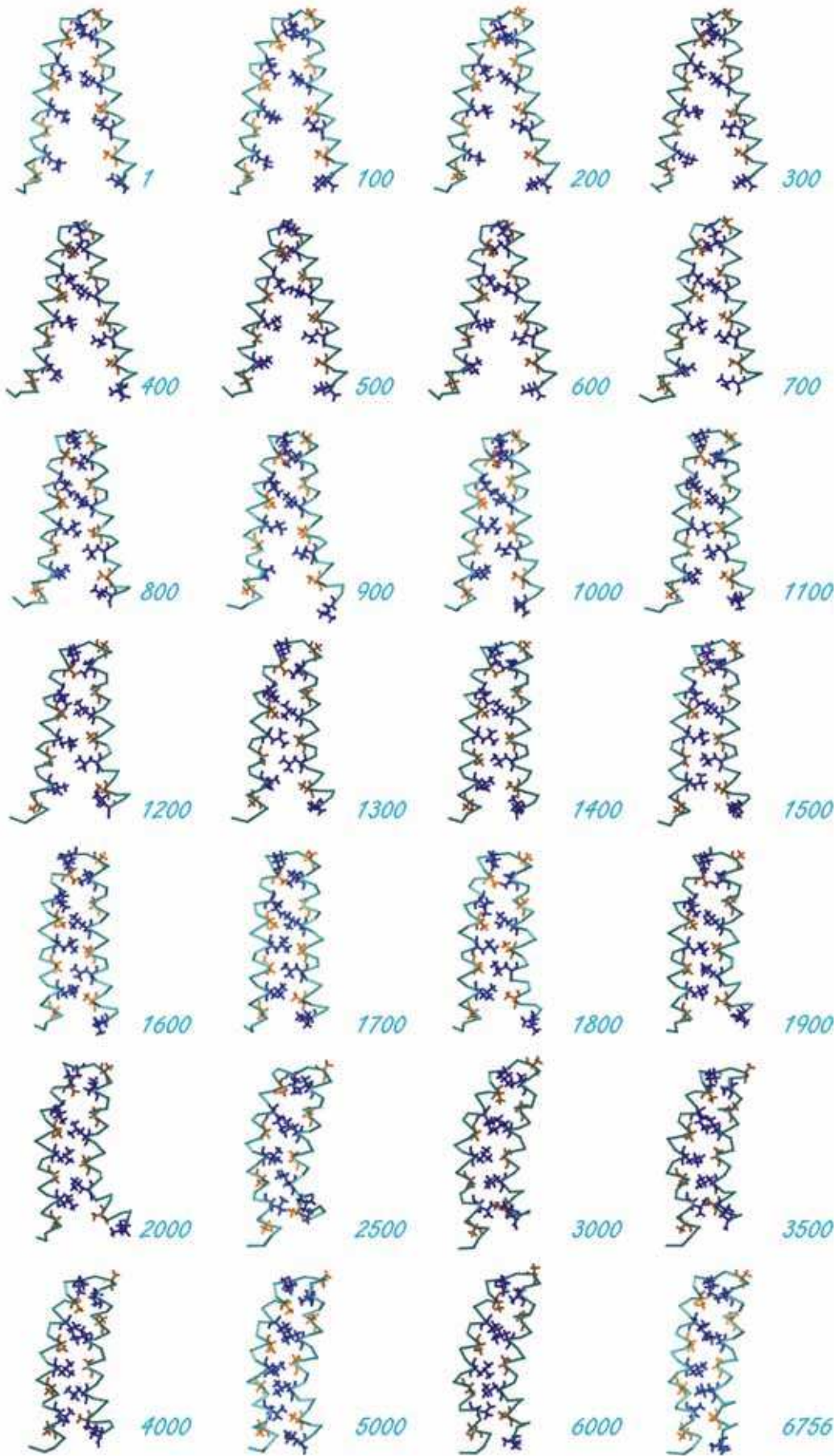
Εικόνα 4.80



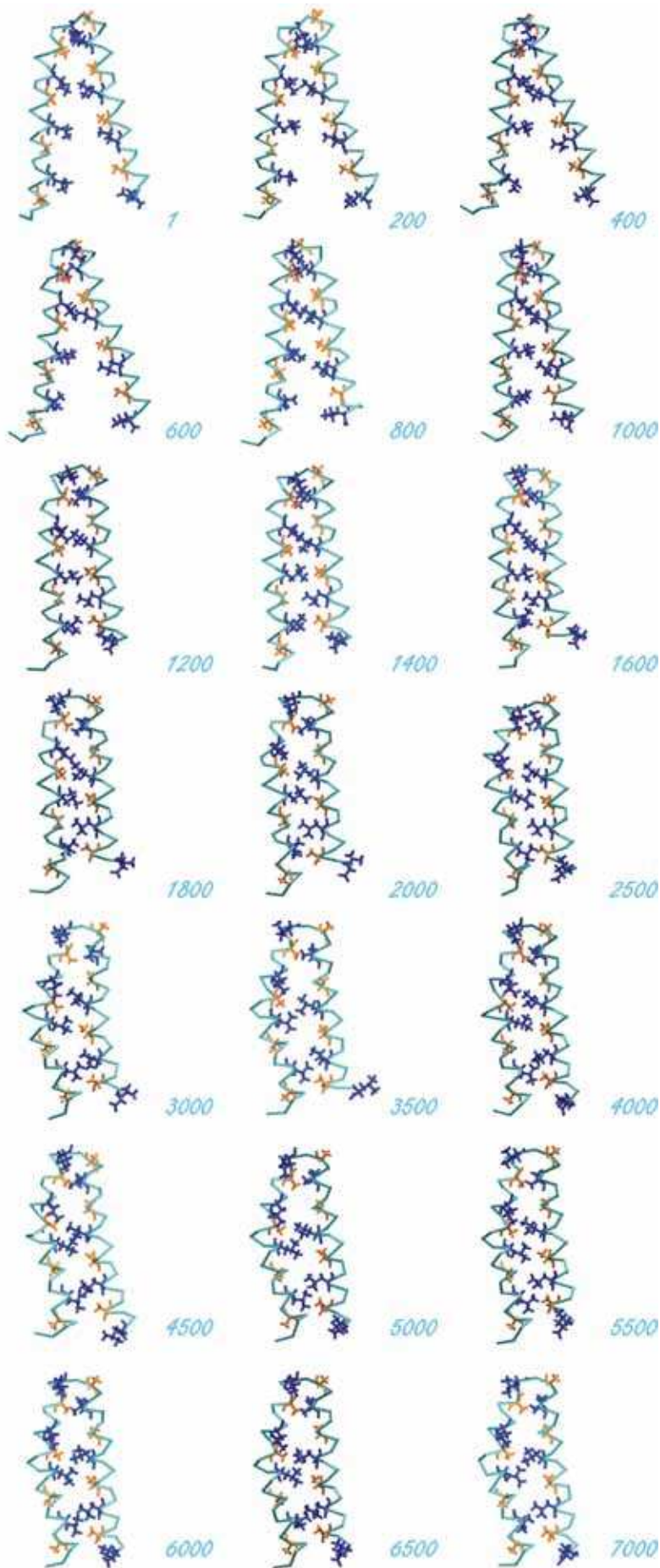
Εικόνα 4.81



Εικόνα 4.82



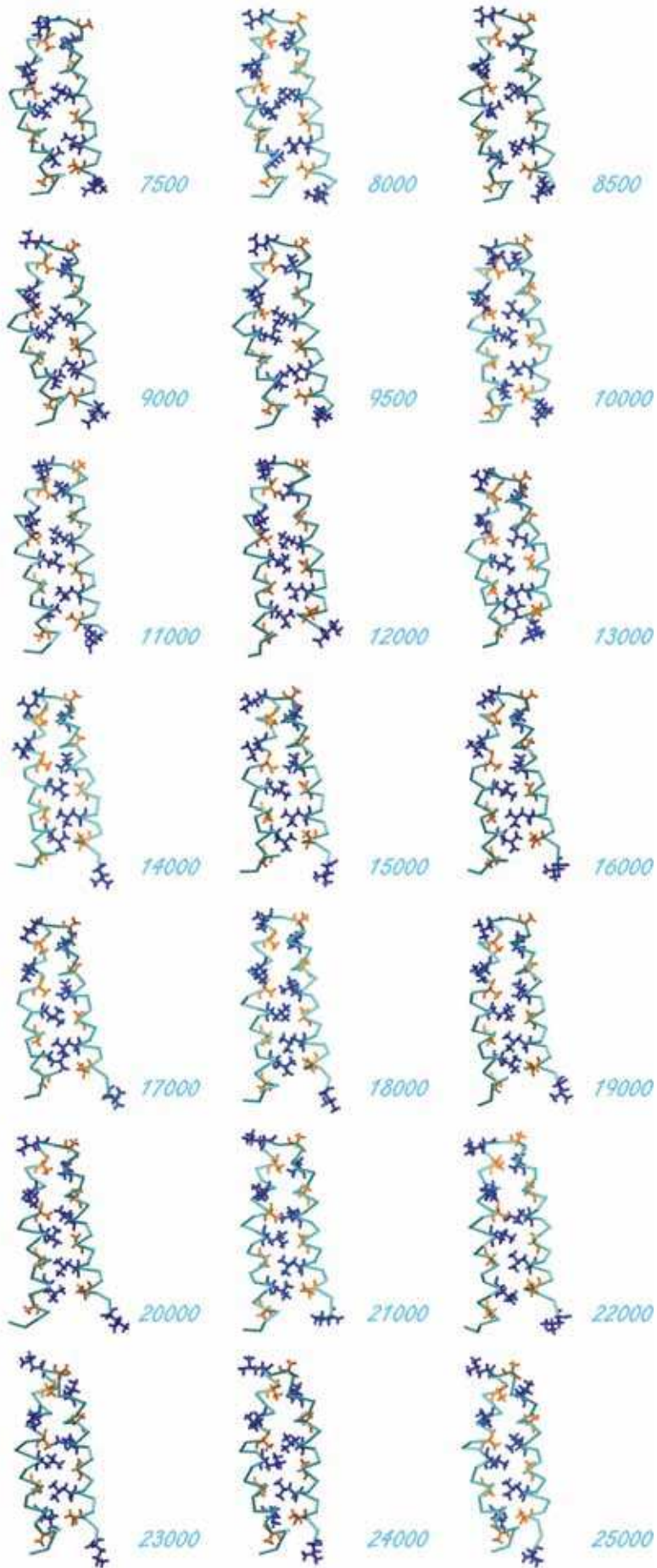
Εικόνα 4.83



Εικόνα 4.84

Αντιπροσωπευτικά στιγμιότυπα της τελευταίας προσομοίωσης (10ns) του *opened* μονομερούς. Κάτω δεξιά σε κάθε στιγμιότυπο αναφέρεται ο χρόνος (σε frame). Ο σκελετός του μορίου απεικονίζεται με wireframe. Οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα απεικονίζονται με sticks. Με πορτοκαλί φαίνονται τα κατάλοιπα *Ala* και με μπλε τα κατάλοιπα *Leu*.

Συνέχεια στη
επόμενη σελίδα

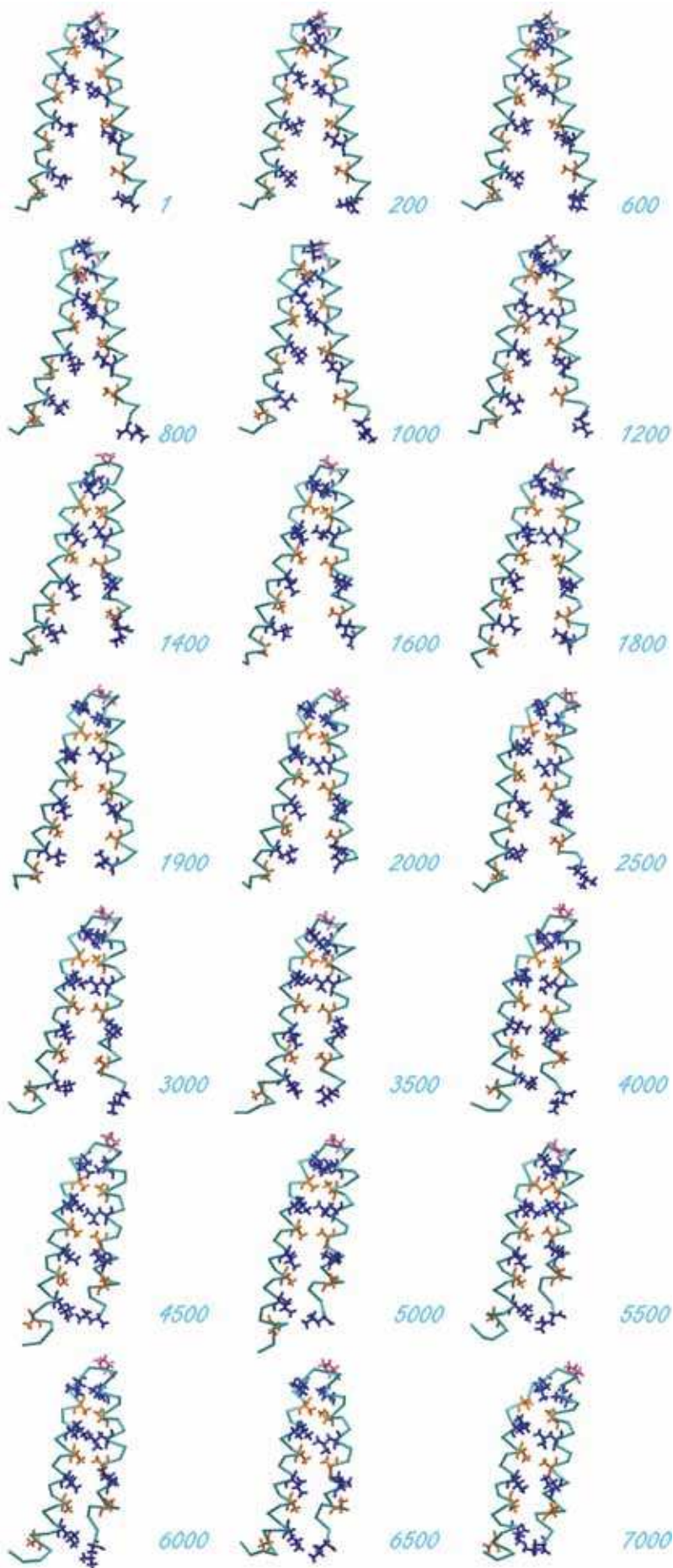


4.8 Ανάλυση προσομοίωσης *opened P31* μονομερούς

Είναι γενικά παραδεκτό ότι η παρουσία ενός καταλοίπου προλίνης οδηγεί σε κάμψη της πρωτεΐνης σε κρίσιμα σημεία της δομής της (MacArthur *et al.*, 1991). Στην Ενότητα 3.8 αναφέρθηκε ένα ακραίο παράδειγμα όπου η παρουσία μίας προλίνης σε μία στροφή ενός 4-α-ελικοειδούς δεματίου οδήγησε σε δραματική αλλαγή στη δομή, με αποτέλεσμα μία πρωτεΐνη με τελείως διαφορετική τοπολογία. Τα μονομερή προσφέρονται ιδιαίτερα για τη διερεύνηση της επίδρασης της προλίνης. Σε πρώτη φάση μπορεί να μελετηθεί η συμπεριφορά και σταθερότητα του *opened P31* μονομερούς μέσα από τις προσομοιώσεις μοριακής δυναμικής. Ακολούθως μπορεί να γίνει συγκριτική ανάλυση με τις προσομοιώσεις του *opened* μονομερούς. Θεωρητικά και δεδομένου ότι όλες οι προσομοιώσεις πραγματοποιούνται κάτω από τις ίδιες συνθήκες, οποιαδήποτε μεγάλη αλλαγή ανάμεσα στα δύο αυτά μονομερή θα πρέπει να οφείλεται στην παρουσία της προλίνης. Παρά το ενδιαφέρον που παρουσιάζει η συγκεκριμένη ανάλυση, στην παρούσα εργασία (λόγω έλλειψης χρόνου για την επεξεργασία τόσο μεγάλου όγκου δεδομένων) θα αρκεστούμε σε απλή παρουσίαση αντιπροσωπευτικών στιγμιότυπων από το *trajectory*.

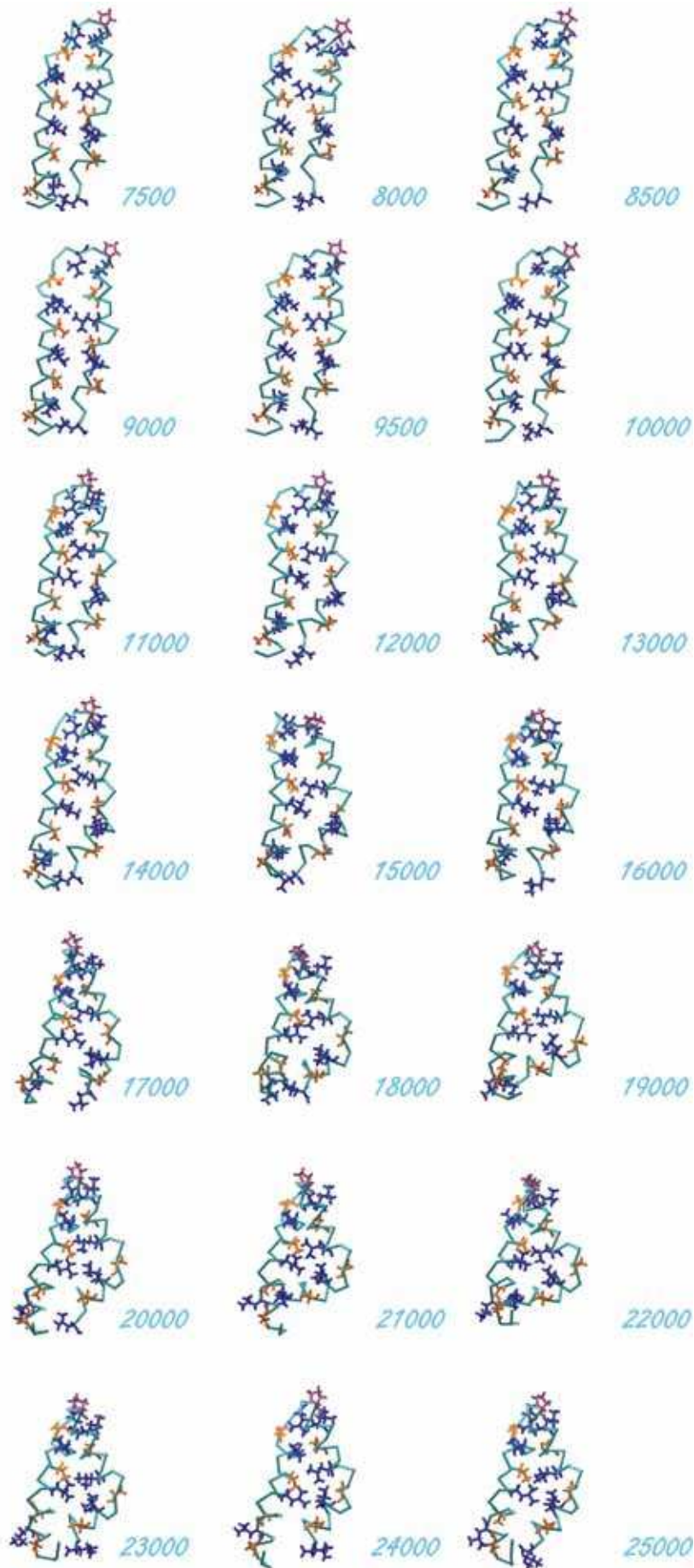
Είναι εντυπωσιακό το γεγονός ότι και σε αυτό το μόριο παρατηρείται ο σχηματισμός των επιπέδων. Από την τοπολογία της A31P, θα ανέμενε κανείς ότι το μονομερές θα πρέπει να παραμένει στην ανοιχτή διαμόρφωση. Σίγουρα σημαντικό ρόλο διαδραματίζει η διαφορετική αλληλουχία με κατάλοιπα αλανίνης και λευκίνης, και θεωρητικά πιο ισχυρά πακεταρισμένο «πυρήνα». Στα 1000 frame έχουν σχηματιστεί περίπου τα μισά επίπεδα, ενώ οι δύο έλικες δεν πλησιάζουν αρκετά μέχρι και τα πρώτα 2000 frame. Ωστόσο δεν φαίνεται να σχηματίζονται όλα τα επίπεδα, όπως για παράδειγμα το έκτο και το έβδομο μέχρι και το 5000 frame. Ακολούθως και πιθανότατα λόγω απουσίας των επιπέδων που σταθεροποιούν τη δομή, παρατηρείται μία αλλοίωση στις έλικες. Εδώ η αλλοίωση είναι αρκετά σημαντική ώστε να μην μιλάμε πλέον για κυρτότητα αλλά για διάλυση της ακεραιότητας των ελίκων και κατά συνέπεια του μορίου (frame 15000). Οι αλληλεπιδράσεις ανάμεσα στις πλευρικές ομάδες δεν είναι πλέον καλά καθορισμένες. Προς το τέλος της προσομοίωσης το μόριο παίρνει ακαθόριστη, σχεδόν διαλυμένη δομή.

Επαναληπτικές προσομοιώσεις θα δείξουν κατά πόσο μπορεί να υφίσταται μία τέτοια τοπολογία για ένα μόριο με αυτήν την αλληλουχία ή εάν η δομή διαλύεται κάθε φορά.



Εικόνα 4.85
 Αντιπροσωπευτικά στιγμιότυπα της προσομοίωσης του *opened P31* μονομερούς. Κάτω δεξιά σε κάθε στιγμιότυπο αναφέρεται ο χρόνος (σε frame). Ο σκελετός του μορίου απεικονίζεται με wireframe. Οι πλευρικές ομάδες των καταλοίπων του υδρόφοβου πυρήνα απεικονίζονται με sticks. Με πορτοκαλί φαίνονται τα κατάλοιπα *Ala*, με μπλε τα κατάλοιπα *Leu*, ενώ με μωβ σημειώνεται η *Pro*.

Συνέχεια στη
 επόμενη σελίδα



ΕΠΙΜΟΙΩΣΗ

Η Rop είναι μία μικρή πρωτεΐνη 63 καταλοίπων που υιοθετεί ένα απλό μοτίβο 4-ελικοειδούς δεματίου. Παρά την απλότητά της, ο μηχανισμός αναδίπλωσης αυτής και των διάφορων παραλλαγών της αποτελεί αντικείμενο μελέτης εδώ και αρκετά χρόνια. Μέχρι πρόσφατα υπήρχε η πεποίθηση ότι οι διάφορες παραλλαγές της θα υιοθετούν τη δομή της φυσικής πρωτεΐνης. Ωστόσο δύο μεταλλαγμένα μόρια, η A31P και $[Ala_2Ile_2]_6$ δίνουν δύο εναλλακτικές τοπολογίες. Μάλιστα η τοπολογία της $[Ala_2Ile_2]_6$, που αναφέρεται ως *syn*, φαίνεται να είναι αρκετά σταθερή. Έτσι προτάθηκε, όπως έχει προαναφερθεί, ότι οι διάφορες παραλλαγές μπορεί να απαντούν σε μίγμα των δύο τοπολογιών και μάλιστα ότι τα μόρια που επιδεικνύουν γρήγορη κινητική αναδίπλωσης, όπως η $[Ala_2Leu_2]_8$ θα τείνουν να πάρουν την *syn* τοπολογία.

Από την ανάλυση των προσομοιώσεων μοριακής δυναμικής σε δύο συστήματα ομοδιμερών, ένα με τη native (Rop-like) και ένα με την *syn* τοπολογία, προκύπτει ότι και τα δύο συστήματα είναι εξίσου σταθερά. Συνεπώς δε μπορούμε να αποκλίσουμε μία από τις δύο τοπολογίες, βάση σταθερότητας. Ωστόσο, σε όλες τις αναλύσεις οι τιμές για τη native (Rop-like) τοπολογία είναι ελαφρώς χαμηλότερες. Με βάση το γεγονός αυτό, θα είναι αναμενόμενο να υπάρχει κάποια προτίμηση όχι προς τη *syn* τοπολογία, αλλά προς την τοπολογία της φυσικής πρωτεΐνης. Οι αποκλίσεις στις τιμές όμως είναι πολύ μικρές για να ισχυριστούμε με βεβαιότητα ότι το μόριο θα έχει μόνο μία από τις εναλλακτικές τοπολογίες και ποια.

Απάντηση στο ερώτημα αυτό θα μπορούσε να δοθεί μέσα από τα συστήματα των μονομερών. Σε πρώτη φάση πραγματοποιήθηκαν προσομοιώσεις απομονωμένων μονομερών ξεκινώντας από διάφορες διαμορφώσεις (*native*, *artificial*, *opened*, *opened P31*). Η ανάλυση των προσομοιώσεων των συστημάτων αυτών (δεν περιλαμβάνεται στα πλαίσια της εργασίας αυτής, λόγω δυσαναλογίας όγκου δεδομένων και διαθέσιμου χρόνου) θα δείξει ποια είναι πιο σταθερά και θα αναδείξει υποψήφιους για την πραγματοποίηση, στη συνέχεια, υπολογιστικών συγκρούσεων. Μέσα από τις συγκρούσεις μονομερών περιμένουμε να παρατηρήσουμε πραγματικά γεγονότα αναδίπλωσης και να δούμε εάν προκύπτει μόνο η μία τοπολογία, ή στην περίπτωση που προκύπτουν και οι δύο, ποια είναι πιο σταθερή. Να υπενθυμίσουμε, ότι για τα συστήματα των ομοδιμερών ξεκινήσαμε από τις κρυσταλλογραφικά προσδιορισμένες δομές και πραγματοποιήσαμε υπολογιστικά τις μεταλλάξεις. Έτσι διαθέτουμε και ένα μέτρο σύγκρισης για τα διμερή που θα προκύψουν από τις συγκρούσεις των μονομερών.

BIBLIOGRAFIA

1. Allen, M.P. (2004). Introduction to molecular dynamics simulation. *N/C 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. Anfinsen, 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. Banner, D. W., Kokkinidis, M. & Tsernoglou, D. (1987). Structure of the ColE1 Rep protein at 1.7 Å resolution. *J. Mol. Biol.* **196**, 657-675.
6. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. (2000). The Protein Data Bank. *Nucleic Acids Res* **28**, 235-242.
7. Bernstein, H. J. (1999). Rasmol 2.7.1. Molecular Graphics Visualisation Tool.
8. Brunger, A. T. (1992). X-PLOR Version 3.1. A system for X-ray crystallography and NMR. Connecticut: Yale University Press
9. Caflisch, A. (2004). Protein folding: simple models for a complex process. *Structure* **12**, 1750-1752.
10. Castagnoli, L., Scarpa, M., Kokkinidis, M., Banner, D. W., Tsernoglou, D. & Cesareni, G. (1989). Genetic and structural analysis of the ColE1 Rep (Rom) protein. *EMBO J.* **8**, 621-629.
11. 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 Rep. *J. Mol. Biol.* **237**, 378-387.

12. Ceruso, M. A., Grottesi, A. & Di Nola, A. (1999). Effects of core-packing on the structure, function, and mechanics of a four-helix-bundle protein ROP. *Proteins* **6**, 436-446.
13. Ceruso, M. A., Pericle, 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.
14. Cheatham III, T. E & Brooks, B. R. (1998). Recent advances in molecular dynamics simulation towards the realistic representation of biomolecules in solution. *Theor. Chem. Acc.* **99**, 279-288.
15. Cheatham III, T.E & Kollman, P.A. (1996). Observation of the *A*-DNA to *B*-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.* **259**, 434-444.
16. 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.
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. (1990). α -helical coiled coils and bundles: how to design an α -helical protein. *Proteins* **7**, 1-15.
19. 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.
20. Daidone, I., Amadei, A. & Di Nola, A. (2005). Thermodynamic and kinetic characterization of a β -hairpin peptide in solution: an extended phase space sampling by molecular dynamics simulations in explicit water. *Proteins* **59**, 510-518.
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. DeMarco, M. L., Alonso, D. O. V. & Daggett, V. (2004). Diffusing and colliding: the atomic level folding/unfolding pathway of a small helical protein. *J. Mol. Biol.* **341**, 1109-1124.
24. Dill, K.A. & Chan, H.S. (1997). From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10-19.
25. Doruker, P., Atilgan, A. R. & Bahar, I. (2000). Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to α -amylase inhibitor. *Proteins* **40**, 512-524.
26. 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.
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. Duan, Y., Wang, L. & Kollman, P. A. (1998). The early stage of folding of villin headpiece subdomain observed in a 200-nanosecond fully solvated molecular dynamics simulation. *PNAS* **95**, 9897-9902.
29. Eberle, W., Pastore, A., Sander, C. & Rosch, P. (1991). The structure of the ColE1 rop in solution. *J. Biomol. NMR* **1**, 71-82.
30. Fadoulglou, V., Glykos, N. M. & Kokkinidis, M. (2001). Side-chain conformations in 4- α -helical bundles. *Protein Eng.* **14**, 321-328.
31. Frisman, D. & Argos, P. (1995). Knowledge-based secondary structure assignment. *Proteins* **23**, 566-579.
32. Gerini, M. F., Roccatano, D., Baciocchi, E. & Nola, A. (2003). Molecular dynamics simulations of Lignin Peroxidase in solution. *Biophys. J.* **84**, 3883-3893.
33. 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.
34. 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.

35. 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.
36. Grayson, P., Tajkhorshid, E. & Schulten, K. (2003). Mechanisms of selectivity in channels and enzymes studied with interactive molecular dynamics. *Biophys J.* **85**, 36-48.
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. Hess, B. (2002). Convergence of sampling in protein simulations. *Phys. Rev.* **E65**, 031910, 107-128.
40. 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.
41. Humphrey, W., Dalke, A. & Schulten, K. (1996). "VMD —Visual Molecular Dynamics". *J. Mol. Graphics* **14**.1, 33-38.
42. Hunenberger, 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.
43. 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.
44. Isralewitz, B., Gao, M. & Schulten, K. (2001). Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.* **11**, 224-230.
45. Izaguirre, J. A., Reich, S. & Skeel, R. D. (1999). Longer time steps for molecular dynamics. *J. Chem. Phys.* **110**, 9853-9864.
46. Jemth, P., Gianni, S., Day, R., Li, B., Johnson, C. M., Daggett, V. & Fersht, A. R. (2004). Demonstration of a low-energy on-pathway intermediate in a fast-folding protein by kinetics, protein engineering, and simulation. *PNAS* **101**, 6450-6455.

47. Jorgensen, M. L., Chandrasekhar, J., Madura, J.D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935.
48. Kabsch, W. (1994). LSQKAB, version 4.2. Collaborative Computational Project, Number 4. The CCP4 suite: Programs for Protein Crystallography, *Acta Cryst.* **D50**, 760-763.
49. Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Cryst.* **A32**, 922-923.
50. 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.
51. Kamtekar, S. & Hecht, M. H. (1995). The four-helix bundle: what determines a fold? *FASEB* **9**, 1013-1022.
52. Karplus, M. (2003). Molecular Dynamics of biological macromolecules: a brief history and perspective. *Biopolymers* **68**, 350-358.
53. Karplus, M. & McCammon, J. A. (2002). Molecular dynamics simulation of biomolecules. *Nat. Struct. Biol.* **9**, 646-652.
54. Kleywegt, G.J. (1995). [O/X-FLOR DICTIONARIES] Dictionaries for Heteros, ESF/CCP4 Newsletter 31, 45-50.
55. Kleywegt, G.J. & Jones, T.A. (1998). Databases in protein crystallography. *Acta Cryst.* **D54**, 1119-1132.
56. Koretke, K. K., Luthey-Schulten, Z. & Wolynes, P. G. (1998). Self-consistently optimized energy functions for protein structure prediction by molecular dynamics. *Biophysics* **95**, 2932-2937.
57. Kresse, H. P., Czubayko, M., Nyakatura, G., Vriend, G., Sander, C. & Bloecker, H. (2001). Four-helix bundle topology re-engineered: monomeric Pop protein variants with different loop arrangements. *Protein Eng.* **14**, 897-901.
58. Lahr, S. J., Engel, D. E., Stayrook, S. E., Maglio, O., North, B., Geremia, S., Lombardi, A. & DeGrado, W. F. (2005). Analysis and design of turns in α -helical hairpins. *J. Mol. Biol.* **346**, 1441-1454.

59. Lampion, L. (1985). *Latex A document preparation system*, Addison Wesley, version Web2c 7.5.4.
60. Lange, O. F., Grubmüller, 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.
61. Larson, S. M., Snow, C. D., Shirts, M. & Pande V. S. *Folding@Home and Genome@Home: Using distributed computing to tackle previously intractable problems in computational biology*. Computational Genomics, 2003.
62. 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.
63. Levinthal, C. (1968). Are there pathways for protein folding? *J. Chim Phys* **65**, 44-45.
64. 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 dimer-folding mystery. *PNAS* **102**, 2373-2378.
65. 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.
66. 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.
67. MacArthur, M. W. & Thornton, J.M. (1991). Influence of proline residues on protein conformation. *J. Mol. Biol.* **218**, 397-412.
68. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evansck, 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., Wiorkiewicz-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.
69. 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.

70. Magliery, T. J. & Regan, L. (2004). Combinatorial approaches to protein stability and structure. *Eur. J. Biochem.* **271**, 1595-1608.
71. Mayer, K. L., Earley, M. R., Gupta, S., Fichumani, K., Regan, L. & Stone, M. J. (2003). Covariation of backbone motion throughout a small protein domain. *Nat. Struct. Biol.* **10**, 962-965.
72. 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.
73. McCammon, J.A., Gelin, B.R. & Karplus, M. (1977). Dynamics of folded proteins. *Nature* **267**, 585-590.
74. McRee, D. E. (1992). A visual protein crystallographic software system for X11/Xview. *J. Mol. Graph.* **10**, 44-46.
75. Merritt, E. A. & Bacon, D. J. (1997). Raster3D —photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.
76. Munson, M., Anderson, K. S. & Regan, L. (1997). Speeding up protein folding: mutations that increase the rate at which Pop folds and unfolds by over four orders of magnitude. *Fold. Des* **2**, 77-87.
77. 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.
78. Munson, M., O'Brien, R., Sturtevant, J. M. & Regan, L. (1994). Redesigning the hydrophobic core of a four-helix-bundle protein. *Prot. Sci.* **3**, 2015-2022.
79. 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.
80. 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.
81. 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.
82. Pantano, S., Tyagi, M., Giacca, M. & Carloni, P. (2004). Molecular dynamics simulation on HIV-1 Tat. *Eur. Biophys J.* **33**, 344-351.

83. Park, S. & Saven, J. G. (2005). Statistical and molecular dynamics studies of buried waters in globular proteins. *Proteins* **60**, 450-463.
84. Peters, K., Hinz, H.-J. & Cesareni, G. (1997). Introduction of a proline residue into position 31 of the loop of the dimeric 4- α -helical protein ROP causes a drastic destabilization. *Biol. Chem* **378**, 1141-1152.
85. Flaxco, 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.
86. 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.
87. Predki, P. F., Nayak, L. M., Gottlieb, M. B. C & Regan, L. (1995). Dissecting RNA-protein interactions: RNA-RNA recognition by Rop. *Cell* **80**, 41-50.
88. Predki, P. F. & Regan, L. (1995). Redesigning the topology of a four-helix-bundle protein: monomeric Rop. *Biochemistry* **34**, 9834-9839.
89. Presnell, S. R. & Cohen, F. E. (1989). Topological distribution of four- α -helix bundles. *PNAS* **86**, 6592-6596.
90. Price, D. J. & Brooks III, C. L. (2002). Modern protein force fields behave comparably in molecular dynamics simulations. *J. Comput. Chem* **23**, 1045-1057.
91. Polisky, B. (1988). ColE1 replication control circuitry: sense from antisense. *Cell* **55**, 929-932.
92. 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>.
93. Rapaport, D.C. (1997). Interactive molecular dynamics. *Physica A.* **240**, 246-254.
94. Rashin, A.A., Iofin, M. & Honig, B. (1986). Internal cavities and buried waters in globular proteins. *Biochemistry* **25**, 3619-3625.
95. Reddy, S. Y., Lederer, F. & Karplus, M. (2003). DNA polymorphism: A comparison of force fields for nucleic acids. *Biophysical J.* **84**, 1421-1449.
96. Rist, M. J. & Marino, J. P. (2001). Association of an RNA kissing complex analysed using 2-aminopurine fluorescence. *Nucleic Acids Res* **29**, 2401-2408.

97. Rose, G.D. & Creamer, T.P. (1994). Protein folding: Predicting predicting. *Proteins* **19**, 502-513.
98. Ryckaert, J. P., Cicotti, G. & Berendsen, H. J. C. (1977). Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys* **23**, 327-341.
99. Sander, C. (1994). Design of protein structures: helix bundles and beyond. *Trends Biotechnol.* **12**, 163-167.
100. Shaner, M.C, Blair, I.M & Schneider, T.D. (Mudge, T.N., Milutinovic, V. & Hunter, L. eds) (2003) Sequence logos : A powerful, yet simple tool, Proceedings of the Twenty-Sixth Annual Hawaii International Conference on System Sciences, Volume 1: Architecture and Biotechnology Computing", 813-821", IEEE Computer Society Press", Los Alamitos, CA", 1993 From the "Alternative Approaches to Sequence Representation" minitrack of the Biotechnology Computing Track, Hawaii International Conference on System Sciences - 26, Kauai, Hawaii - January 5-8, 1993
101. 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.
102. Smith, L. J., Daura, X. & van Gunsteren, W. F. (2002). Assessing equilibration and convergence in biomolecular simulations. *Proteins* **48**, 487-496.
103. 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.
104. 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.
105. Straasma, T. P. & Berendsen, H. J. C. (1988). Free energy of ionic hydration: analysis of a thermodynamic integration technique to evaluate free energy differences by molecular dynamics simulations. *J. Chem. Phys* **89**, 5876-5886.
106. Suárez, D. & Field, M. J. (2005). Molecular dynamics simulations of human butyrylcholinesterase. *Proteins* **59**, 104-117.

107. Tajkhorshid, E, Nollert, P., Jensen, M. Miercke, L. J. W., O'Connell, J., Stroud, R. M. & Schulten, K. (2002). Control of the selectivity of the Aquaporin water channel family by global orientational tuning. *Science* **296**, 525-530.
108. 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.
109. 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.
110. 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.
111. Wright, Lindorff —Larsen, K., Randles, L.G. & Clarke, J. (2003). Parallel protein-folding pathways revealed and mapped. *Nat. Struct. Biol.* **10**, 658-662.
112. 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.
113. 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.
114. 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.
115. Zielkiewicz, J. (2005). Structural properties of water: comparison of the SPC, SPCE, TIP4P and TIP5P models of water. *J. Chem. Phys* **123**, 104501.
116. Zoetewey, D. L., Tripet, B. P. Kutateladze, T. G., Overduin, M. J., Wood, J. M. & Hodges, R. S. (2003). Solution structure of the C-terminal antiparallel coiled-coil domain from *Escherichia coli* osmosensor ProP. *J. Mol. Biol.* **334**, 1063-1076.
117. <http://origin.imbb.forth.gr/~glykos/> *carma*, locally written software.

PARASTHENA

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
PmeGridsizeY      64
PmeGridsizeZ      60

#
# Periodic boundary things
#
wrapWater          on
wrapNearest        on

cellBasisVector1   78.59 00.00 00.00
cellBasisVector2   00.00 60.14 00.00
cellBasisVector3   00.00 00.00 56.83
cellOrigin          0.02 -0.01 2.42

#
# Fixed atoms for initial heating-up steps
#
fixedAtoms         on
fixedAtomsForces   on
fixedAtomsFile     fixe_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
langevinHydrogen   on

langevinPiston     on
langevinPistonTarget 1.01325
langevinPistonPeriod 200
langevinPistonDecay 100
langevinPistonTemp 320

useGroupPressure   yes

#####
# The actual minimisation and heating-up
# protocol follows.
#####

```

```

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

#
# turn off pressure control until later
#
langevinPiston  off

#
# minimize nonbackbone atoms
#
minimize      2000
output       output/min_fix

#
# min all atoms
#
fixedAtoms    off
minimize      2000
output       output/min_all

#
# heat with CAs restrained
#
set temp 20;
while (Stemp < 321) {
  langevinTemp  Stemp
  run          2000
  output       output/heat_ca
  set temp [expr Stemp + 20]
}

#
# equilibrate volume with CAs restrained
#
langevinPiston  on
run            10000
output         output/equil_ca

#
# equilibrate volume without restraints
#
constraintScaling  0
run               100000

```

Equilibration script (production phase)

```
#
# Input files
#
structure          ionized.psf
coordinates        heat_out.coor
velocities         heat_out.vel
extendedSystem    heat_out.xsc
parameters        par_all127_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      135

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

```
#
# Particle Mesh Ewald parameters
#
Pme                on
PmeGridsizeX      80
PmeGridsizeY      64
PmeGridsizeZ      60

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

#
# Langevin dynamics parameters
#
langevin           on
langevinDamping    1
langevinTemp       320
langevinHydrogen   on

langevinPiston     on
langevinPistonTarget 1.01325
langevinPistonPeriod 200
langevinPistonDecay 500
langevinPistonTemp 320

useGroupPressure   yes

firsttimestep      146000
run                5000000
```


Calculation of rmsd from starting structure (xplor script)

```
structure @protein.psf end
coordinates disposition-comp @psfgen.pdb

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

evaluate ($current - $first)

read trajectory

ascii=false
input=protein.dcd
begin=$first
skip=$step
stop=2000000000
end

set display=(backbone.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) end

  coor rms selection-((name CA or name C or name O or name HN or name HA or name N)
    and not hydrogen) end

  evaluate ($rms = $RESULT)

  display $current $rms

  evaluate ($current = $current + $step)

  read trajectory next end

end loop traj

stop
```

Calculation of rmsd from average structure (xplor script)

```
structure @protein.psf end
coordinates disposition-comp @aver_and_rms.pdb

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

evaluate ($current - $first)

read trajectory

ascii-false
input-protein.dcd
begin-$first
skip-$step
stop-20000000
end

set display-aver vs MD.rmsd end

while ($status # "COMPLETE") loop traj

  coor fit selection-(name ca) end
  coor rms selection-(name ca) end
  evaluate ($rms_CA_ALL - $RESULT)

  display $current $rms_CA_ALL

  evaluate ($current - $current + $step)

  read trajectory next end

end loop traj

stop
```

Calculation of radius of gyration (xplor script)

```
structure @protein.psf end

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

evaluate ($current - $first)

read trajectory

ascii=false
input=protein.dcd
begin=$first
step=$step
stop=20000000
end

set display-gyration.list end

while ($status # 'COMPLETE') loop traj

coor gyration mass=true end

display $current $RG

evaluate ($current - $current + $step)

read trajectory next end

end loop traj

stop
```

Calculation of χ_1 dihedral (xplor script)

```
structure @protein.psf end

dynamics analysis
  pick dihedral
  ( residue 15 and name n and segid="A" )
  ( residue 15 and name ca and segid="A" )
  ( residue 15 and name cb and segid="A" )
  ( residue 15 and name cg and segid="A" ) geometry

asci=false
input=protein.dcd
begin=0 skip=1 stop=2000000

output=chi1_15A.list

end

stop
```

Calculation of χ_2 dihedral (xplor script)

```
structure @protein.psf end

dynamics analysis
  pick dihedral
  ( residue 15 and name ca and segid="A" )
  ( residue 15 and name cb and segid="A" )
  ( residue 15 and name cg and segid="A" )
  ( residue 15 and name cd1 and segid="A" ) geometry

asci=false
input=protein.dcd
begin=0 skip=1 stop=2000000

output=chi2_15A.list

end

stop
```

SumFluct (Sum of Fluctuations)

```
#include <stdio.h>
#include <math.h>

main()
{
  float x, y, z, w, q;
  int c;

  while ( c = scanf("%f %f %f %f %f", &x, &y, &z, &w, &q) == 5 )
    printf("%10.8f\n", ( x + y + z + w + q ));
  if ( c > 0 )
    fprintf(stderr, "ERROR\n");
}
```

DiCon (Dihedral Conversion)

```
#include <stdio.h>
#include <math.h>

main ()
{
  float val1;
  float val2;

  while ( scanf ("%f %f", &val1, &val2) == 2 )
  {
    if ( val2 > 0.0 )
      printf( "%15.6f %15.6f\n", val1, val2);
    else
    {
      val2 += 360.0;
      printf( "%15.6f %15.6f\n", val1, val2);
    }
  }
}
```