

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ
ΤΜΗΜΑ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ ΚΑΙ ΓΕΝΕΤΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ ΤΟΥ ΠΕΠΤΙΔΙΟΥ
AAGDYY

ΜΑΡΑΔΙΔΟΥ ΣΤΥΛΙΑΝΗ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ : ΝΙΚΟΛΑΟΣ Μ. ΓΛΥΚΟΣ

ΑΛΕΞΑΝΔΡΟΥΠΟΛΗ , ΟΚΤΩΒΡΙΟΣ 2010

Ευχαριστίες

Θα ήθελα να ευχαριστήσω την οικογένεια μου για την αγάπη και την υποστήριξη τους όλα αυτά τα χρόνια.

Ακόμη, θα ήθελα να ευχαριστήσω τον επιβλέπων καθηγητή μου κ. Νικόλαο Γλυκό για την συνεργασία και την κατανόησή του.

Τέλος, θα ήθελα να ευχαριστήσω τους φίλους μου που ήταν δίπλα μου.

ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη	4
Κεφάλαιο 1	
Εισαγωγή	
1.1 Βιολογική σημασία και χρήση των πεπτιδίων	5
1.2 Το πεπτίδιο AAGDYY	7
Κεφάλαιο 2	
Προσομοιώσεις Μοριακής Δυναμικής	
2.1 Προσομοιώσεις Μοριακής Δυναμικής με λίγα λόγια	13
2.2 Στάδια προσομοίωσης	16
2.3 Περιορισμοί	18
Κεφάλαιο 3	
Προσομοίωση Πεπτιδίου	
3.1 Προετοιμασία του συστήματος	23
3.2 Μέθοδοι ανάλυσης τροχιακών	25
Κεφάλαιο 4	
Ανάλυση Προσομοιώσεων	
4.1 RMSD από την αρχική δομή	29
4.2 frame to frame rmsd	30
4.3 Cluster analysis – Ομαδοποίηση των αποτελεσμάτων	32
4.3.1 Cartesian-PCA ανάλυση cluster	36
4.3.2 dihedral-PCA ανάλυση cluster	38
4.3.2.1 Σύγκριση των clusters από την ανάλυση dPCA	51
4.3.3 Σύγκριση των clusters με την πειραματική δομή	52
Επίλογος	56
Βιβλιογραφία	57
Παράρτημα	62

ΠΕΡΙΛΗΨΗ

Στα πλαίσια της κατανόησης του προβλήματος της αναδίπλωσης των πρωτεϊνών μελετάτε η δομή μικρότερων πεπτιδίων που έχουν σταθερή δομή. Στην παρούσα εργασία ως σύστημα μελέτης επιλέχθηκε το εξαπεπτίδιο AAGDYY. Το πεπτίδιο αυτό αποτελεί τμήμα της πρωτεΐνης BLIP και έχει την ικανότητα να αναστέλλει την TEM-1 β-λακταμάση καθώς και άλλες β-λακταμάσες. Έτσι, το πεπτίδιο αυτό μπορεί να χρησιμοποιηθεί για την ανάπτυξη νέων φαρμάκων ενάντια σε μια ποικιλία β-λακταμασών. Από προηγούμενες μελέτες έχει δειχθεί πειραματικά ότι το πεπτίδιο αυτό έχει δομή β-στροφής σε διάλυμα DMSO-d₆. Στόχος μας ήταν να αναπαράγουμε αυτή τη δομή του πεπτιδίου με υπολογιστικές μεθόδους. Χρησιμοποιήσαμε λοιπόν τις προσομοιώσεις μοριακής δυναμικής για να παρακολουθήσουμε την πορεία της αναδίπλωσης του πεπτιδίου σε υδατικό διάλυμα. Ωστόσο από την σύγκριση των αποτελεσμάτων δεν επιβεβαιώθηκε η πειραματικά προσδιορισμένη δομή του πεπτιδίου.

ABSTRACT

In order to understand the problem of protein folding we study the structure of smaller peptides which have a stable structure. In this study we selected the exapeptide AAGDYY. This peptide is part of the protein BLIP and has the ability to inhibit the TEM-1 β-lactamase and other β-lactamases. Thus, this peptide can be used for the development of new drugs against a variety of β-lactamases. Previous studies have shown experimentally that the peptide adopts a β-turn structure in DMSO-d₆ solution. Our goal was to reproduce this structure of the peptide with computational methods. Thus, we used molecular dynamics simulations to watch the folding process of the peptide in aqueous solution. However, by comparing the results, the experimentally determined structure of the peptide was not confirmed.

Κεφάλαιο 1

Εισαγωγή

1.1 Βιολογική σημασία και χρήση των πεπτιδίων

Τα πεπτίδια, όπως και οι πρωτεΐνες, παίρνουν μέρος σε όλες σχεδόν τις λειτουργίες των κυττάρων και ορισμένα έχει δειχθεί πειραματικά ότι εμφανίζουν στοιχεία δευτεροταγούς δομής. Τα πεπτίδια μπορούν να χωριστούν με βάση το βιολογικό τους ρόλο στις εξής κατηγορίες: τις πεπτιδικές ορμόνες, τα νευροπεπτίδια, τα πεπτιδικά αντιβιοτικά και τις πεπτιδικές τοξίνες (Sewald and Jakubke, 2002). Αναπτύχθηκαν πολλές τεχνικές για τη μελέτη των πεπτιδίων όσον αφορά την απομόνωση, τη σύνθεση, τον προσδιορισμό της δομής και τον τρόπο δράσης τους (Harder et al., 2001). Το ενδιαφέρον τους προς τα πεπτίδια στρέφουν οι τομείς της βιοχημείας, της χημείας, της βιολογίας, της φαρμακολογίας και της τεχνολογίας γονιδίων. Σημαντική είναι η συμβολή του τρόπου δράσης των πεπτιδίων στη φαρμακολογία για την ανάπτυξη νέων θεραπειών που αφορούν ασθένειες που σχετίζονται μ' αυτά.

Η επικοινωνία μεταξύ των κυττάρων επηρεάζεται από τα πεπτίδια που αλληλεπιδρούν με τους υποδοχείς και παίρνουν μέρος σε βιοχημικές διεργασίες όπως το μεταβολισμό, την αντίληψη του πόνου, την αναπαραγωγή και την ανοσολογική απόκριση δρώντας ως νευροδιαβιβαστές, νευρορρυθμιστές και ορμόνες (Fricke et al., 2006). Τα νευροπεπτίδια προκύπτουν μετά από πρωτεόλυση των πρόδρομων πρωτεϊνών, αποθηκεύονται στα νευροεκκριτικά κυστίδια, μεταφέρονται στις απολήξεις και έχουν ορμονική δράση σε ορισμένους ιστούς. Κατά τη διαδικασία της πρωτεόλυσης των πρόδρομων πρωτεϊνών εκτός από την παραγωγή των λειτουργικών πεπτιδίων σχηματίζονται και πολλά άλλα πεπτίδια. Αναμένεται ότι αυτά τα πεπτίδια

μπορεί να έχουν λειτουργικούς ρόλους και να μας προσφέρουν καινούρια δεδομένα για τους μηχανισμούς ρύθμισης και σηματοδότησης (Ueki *et al.* 2007).

Μια άλλη μεγάλη κατηγορία πεπτιδίων με βιολογική σημασία είναι τα αντιμικροβιακά πεπτίδια και οι πεπτιδικές τοξίνες που αποτελούν την πρώτη γραμμή άμυνας, ενάντια στις μολύνσεις από ένα μεγάλο φάσμα μικροοργανισμών συμπεριλαμβανομένων των gram- θετικών και αρνητικών βακτηρίων, των μυκήτων και συγκεκριμένων ιών. Αντιμικροβιακά πεπτίδια έχουν ταυτοποιηθεί σε πολλούς οργανισμούς όπως φυτά, έντομα, βακτήρια και σπονδυλωτά (Brown and Hancock 2006). Η πλειοψηφία των αντιμικροβιακών πεπτιδίων αποτελείται από 6 με 50 κατάλοιπα αμινοξέων. Προκειμένου να αντιμετωπιστούν αποτελεσματικότερα περιπτώσεις βακτηριακών μολύνσεων που εμφανίζουν ανθεκτικότητα στα αντιβιοτικά το ενδιαφέρον των επιστημόνων στράφηκε στη χρήση αυτών των πεπτιδίων για την παραγωγή νέων φαρμάκων (Brogden 2005).

Τα πεπτίδια βρίσκουν εφαρμογή στην παραγωγή εμβολίων που ο σχεδιασμός τους βασίζονται είτε σε μια υπομονάδα ενός παθογόνου παράγοντα, είτε σε ανοσολογικά πεπτίδια είτε σε συνθετικά πεπτίδια που αντιστοιχούν σε συντηρημένες περιοχές πρωτεϊνών με παθογόνο δράση (Berzofsky *et al.* 2005). Το σκεπτικό είναι ότι ένα μικρότερο πεπτίδια που η αλληλουχία του ταυτίζεται με αυτή ενός επιτόπου ενάντια στον οποίο παράγονται αντισώματα είναι δυνατόν να έχει την ίδια δράση με μια πρωτεΐνη του παθογόνου παράγοντα που επάγει την ανοσολογική απόκριση του ξενιστή. Ο στόχος αυτής της στρατηγικής είναι ο εμβολιασμός με ένα όσο το δυνατό μικρότερο πεπτίδιο που να αποτελείται από ένα καλά καθορισμένο αντιγόνο και να προκαλεί αποτελεσματικά μια ειδική ανοσολογική απόκριση χωρίς ενδεχόμενο κινδύνου (Rowlands 1992) . Υπάρχουν όμως και αρκετοί περιορισμοί: η αμινοξική αλληλουχία του πεπτιδίου δε θα πρέπει να απαντάται σε πρωτεΐνες του ίδιου οργανισμού ώστε να είναι ειδικό, οι επίτοποι των πρωτεϊνών που μιμείται το πεπτίδιο πρέπει να εκτίθενται στην επιφάνεια για να αντιδρούν με τα αντισώματα, το πεπτίδιο θα πρέπει να είναι σε θέση να παίρνει στερεοδιάταξη παρόμοια με αυτή του φυσικού επιτόπου και τέλος είναι πιθανό ένας μόνο επίτοπος να μην είναι επαρκώς ανοσογόνος. Για να είναι ανοσογόνα τα πεπτίδια-επίτοποι χρειάζεται να συζευχθούν σε μεγάλα μόρια φορείς όπως πρωτεΐνες ή δομές που θυμίζουν το καψίδιο ιών.

Πολλές βιολογικές διεργασίες των κυττάρων ρυθμίζονται από την αλληλεπίδραση μεταξύ των πρωτεϊνών. Ειδικά πεπτίδια στόχοι μπορούν να χρησιμοποιηθούν για την κατανόηση της φύσης των λειτουργικών περιοχών των

πρωτεϊνών, για την ταυτοποίηση και απομόνωση των ορμονικών και νευροδιαβιβαστικών υποδοχέων και το σχεδιασμό φαρμάκων με βάση τη δομή (Hunter and Kuhar 2003). Η χρήση των πεπτιδίων ως αγωνιστές ή ανταγωνιστές για την αναστολή της αλληλεπίδρασης των πρωτεϊνών μπορεί να εφαρμοστεί στην θεραπεία ασθενειών αφού πρώτα αντιμετωπιστούν τα προβλήματα της πρωτεόλυσης και βιοδιαθεσιμότητας.

Η γνώση της δομής και της λειτουργίας των πρωτεϊνών αποτελεί το στόχο της σύγχρονης μοριακής βιολογίας (Dill et al., 2007). Η ικανότητα να προβλέπουμε το μηχανισμό αναδίπλωσης και την τελική δομή μιας πρωτεΐνης βασισμένοι στην αλληλουχία της μπορεί να επηρεάσει δραματικά διάφορους τομείς που εκτείνονται από τη βιοχημεία στη μοριακή ιατρική και τη νανοτεχνολογία. Ο τομέας της αναδίπλωσης των πρωτεϊνών έχει αναπτυχθεί τόσο που σήμερα μπορούμε να εξετάσουμε την αναδίπλωση μικρών πρωτεϊνών, πρωτεϊνών με μικρό χρόνο αναδίπλωσης και μικρών πρωτεϊνικών τμημάτων. Η κατανόηση της διαδικασίας αναδίπλωσης αυτών μας προσφέρει σημαντικές πληροφορίες για το τρόπο που αναδιπλώνονται μεγαλύτερες πρωτεΐνες καθώς συναντιούνται σε αυτές όμοιες δομές (Motta et al. 2005). Ένα σημαντικό εργαλείο με το οποίο μπορούν να αναπαρασταθούν οι πολύπλοκες αλληλεπιδράσεις που λαμβάνουν χώρα κατά την αναδίπλωση πεπτιδίων είναι οι προσομοιώσεις μοριακής δυναμικής (Seibert M. M. et al. 2005).

1.2 Το πεπτίδιο AAGDYY.

Στην παρούσα εργασία ασχοληθήκαμε με το πεπτίδιο AAGDYY το οποίο παρουσιάζει φαρμακολογικό ενδιαφέρον για την παρασκευή αντιβιοτικών καθώς έχει βρεθεί ότι εμπλέκεται στην αναστολή της TEM-1 β-λακταμάσης λόγω της σύνδεσης του στο ενεργό της κέντρο. Στις παραγράφους που ακολουθούν δίνονται βιβλιογραφικά στοιχεία σχετικά μ' αυτό.

Ο σχεδιασμός φαρμάκων με τη χρήση μικρών πεπτιδίων αποτελεί σημαντικό πεδίο έρευνας για την ανάπτυξη νέων θεραπειών. Ένα τέτοιο παράδειγμα είναι και η ανάγκη που δημιουργήθηκε για την εύρεση νέων φαρμάκων που να αναστέλλουν τη δράση των β-λακταμασών. Στη γεωργία, τη βιομηχανία και την ιατρική εφαρμόζονται θεραπείες με αντιβιοτικά β-λακταμών. Μεγάλη χρήση τους παρατηρείται στις χώρες με αναπτυγμένη βιομηχανία με σκοπό τη χρήση τους στη γεωργία και ως

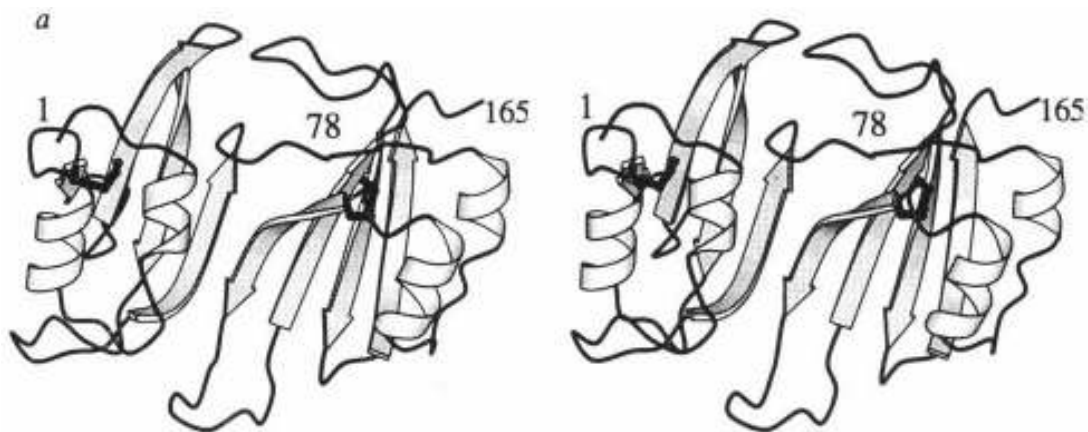
συμπλήρωμα διατροφής στη ζωική παραγωγή. Από τη χρήση αντιβιοτικών β-λακταμασών δεν έχουν παρατηρηθεί σημαντικές παρενέργειες στον άνθρωπο και λόγω της χαμηλής τους τοξικότητας και της μεγάλης δομικής του ποικιλομορφίας θεωρούνται ως η καταλληλότερη θεραπεία για την αντιμετώπιση ασθενειών που προκαλούνται από βακτήρια. Παρ' όλα αυτά όμως εμφανίστηκαν βακτήρια ανθεκτικά σε αυτά και η ανάπτυξη νέων θεραπειών αποτελεί επιτακτική ανάγκη.

Ο μηχανισμός άμυνας των βακτηρίων ενάντια στα αντιβιοτικά β-λακταμών είναι η παραγωγή β-λακταμασών. Οι β-λακταμάσες είναι βακτηριακά ένζυμα που παρέχουν στα βακτήρια ανθεκτικότητα ενάντια στα αντιβιοτικά β-λακταμών, όπως τις πενικιλίνες και τις κεφαλοσπορίνες, που χρησιμοποιούνται ενάντια σε μια μεγάλη ποικιλία βακτηρίων και η επίδραση τους προκαλεί διάσπαση της δομής του κυτταρικού τους τοιχώματος (Wang, J. 2009), καταλύοντας την υδρόλυση τους. Οι β-λακταμάσες κατατάσσονται σε τέσσερις (A-D) κλάσεις. Η TEM-1 β-λακταμάση είναι η πιο διαδεδομένη β-λακταμάση A κλάσης ανάμεσα στα gram-αρνητικά βακτήρια. Η TEM-1 β-λακταμάση παράγεται από το βακτήριο *Escherichia coli*, κωδικοποιείται από το γονίδιο bla_{TEM-1} και ρόλος της είναι να υδρολύει τις πενικιλίνες και τις περισσότερες κεφαλοσπορίνες. Η αλόγιστη χρήση των αντιβιοτικών β-λακταμών έχει προκαλέσει πίεση φυσικής επιλογής στα βακτήρια. Εμφανίστηκαν Εντεροβακτήρια που παράγουν νέες β-λακταμάσες ή παραλλαγές των κλασικών β-λακταμασών που τα καθιστούν ανθεκτικά στα αντιβιοτικά. Τα βακτήρια αυτά αποτελούν μεγάλη απειλή καθώς κάνουν τις τρέχουσες θεραπείες κατά των βακτηρίων αναποτελεσματικές (Therrien, C., 2000) .

Το διήθημα καλλιέργειας του gram-θετικού βακτηρίου *Streptomyces clavuligerus* βρέθηκε ότι εκτός από μια ποικιλία ενώσεων β-λακταμασών περιέχει μια πρωτεΐνη αναστολέα β-λακταμασών (BLIP) (Doran J. L. et al., 1990). Η BLIP είναι μια εξωκυτταρική πρωτεΐνη που δρα ως αναστολέας ενάντια σε μια ποικιλία β-λακταμασών μεταξύ των οποίων και της TEM-1 β-λακταμάσης (Natalie C. J et al., 1994). Προκειμένου να αναπτυχθούν νέες στρατηγικές αντιμετώπισης της ανθεκτικότητας των βακτηρίων ενάντια στα αντιβιοτικά β-λακταμών μελετήθηκε η ανασταλτική δράση της BLIP ενάντια στις β-λακταμάσες καθώς όλα τα μικρά μόρια αναστολείς, όπως clavulanic acid, sulbactam και tazobactam, που χρησιμοποιήθηκαν δεν ήταν πολύ αποτελεσματικά.

Ο αναστολέας β-λακταμασών είναι μια πρωτεΐνη 165 αμινοξέων που αποτελείται από μια διαδοχική επανάληψη μιας επικράτειας 76 αμινοξέων που

συνδέονται από ένα τριπεπτίδιο Ala77-Ser79 (**Εικόνα 1.1**). Πρόκειται για ένα σχετικά επίπεδο μόριο με κάθε επικράτειά του να αποτελείται από το μοτίβο έλικα-στροφή-έλικα τοποθετημένη απέναντι από μια αντιπαράλληλη β πτυχωτή επιφάνεια τεσσάρων κλώνων. Οι δυο επικράτειες τοποθετούνται η μια δίπλα στην άλλη με αποτέλεσμα το σχηματισμό 8-κλώνων αντιπαράλληλης β-πτυχωτής επιφάνειας. Σε κάθε επικράτεια και οι δυο έλικες βρίσκονται παράλληλα στους κλώνους της πτυχωτής επιφάνειας, σχηματίζοντας μεταξύ αυτών ένα μικρό υδροφοβικό πυρήνα. Παρόλο που έχει δυο δομικά όμοιες επικράτειες η στοιχειομετρία της δέσμευσης της BLIP με τα ένζυμα είναι αυστηρά 1:1.



Εικόνα 1.1 Stereo όψη που δείχνει περιοχές της δευτεροταγούς δομής της BLIP. (αναπαράγεται άνευ αδείας από τους Natalie C. J. et al, 1994)

Η β-λακταμάση TEM-1 αναστέλλεται ισχυρά από την BLIP. Ύστερα από σύγκριση της δομής της TEM-1 β-λακταμάσης με β-λακταμάσες άλλων βακτηρίων όπως των *S. Aureus* και *B. Licheniformis* παρατηρήθηκε ότι η διάταξη των βασικών καταλυτικών καταλοίπων είναι όμοια. Γίνονται υποθέσεις λοιπόν για το κατά πόσο είναι πιθανό τα δομικά χαρακτηριστικά της BLIP τα οποία αλληλεπιδρούν ειδικά και αναστέλλουν την περιοχή του ενεργού κέντρου της TEM-1 θα φανούν επίσης χρήσιμα για το σχεδιασμό αναστολέων β-λακταμασών για άλλα είδη (Natalie C. J et al., 1994).

Η μελέτη της κρυσταλλικής δομής της BLIP με την TEM-1 β-λακταμάση έδειξε ότι το κατάλοιπο Asp-49 από το βρόχο της μιας επικράτειας της BLIP δημιουργεί δεσμούς υδρογόνου με τέσσερα συντηρημένα κατάλοιπα του ενεργού κέντρου της TEM-1 (Ser-130, Lys-234, Ser-235 και Arg-244). Αυτά τα τέσσερα

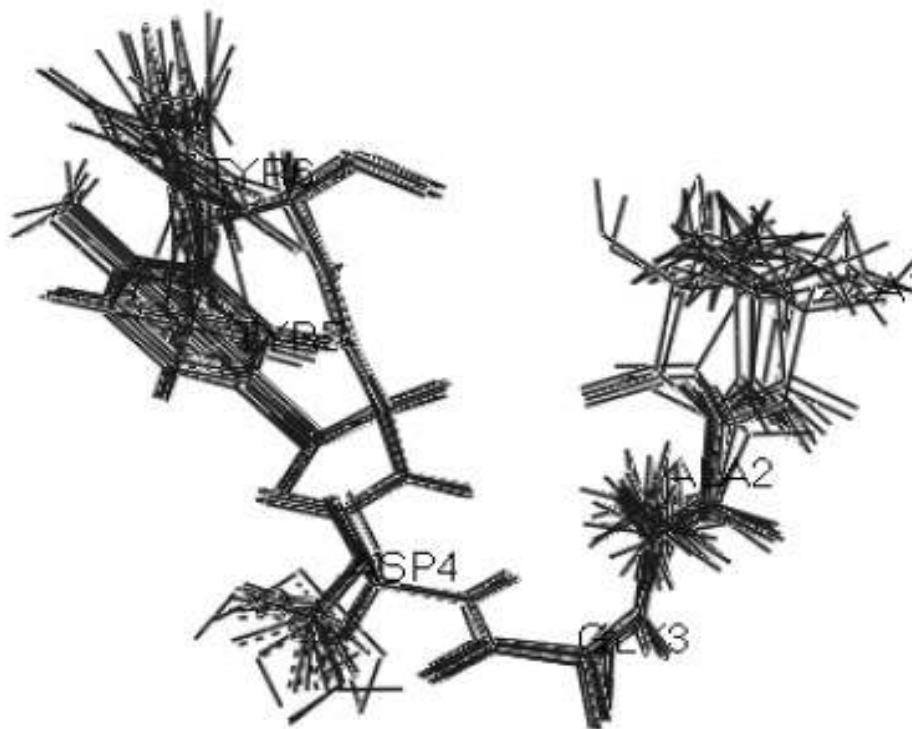
αμινοξέα εμπλέκονται στην δέσμευση και κατάλυση των αντιβιοτικών β-λακταμών και είναι συντηρημένα σε όλες τις β-λακταμάσες κλάσης A. Ένα επίσης σημαντικό κατάλοιπο για την αναστολή της TEM-1 β-λακταμάσης είναι το αμινοξύ Phe-142 από το βρόχο της δεύτερης επικράτειας που έρχεται σε επαφή με τα κατάλοιπα Glu-104, Tyr-105, Asn-170, Ala-237, Gly-238 και Glu-240 της β-λακταμάσης. Τα περισσότερα από αυτά τα κατάλοιπα είναι είτε συντηρημένα στις κλάσης A β-λακταμάσες ή εμπλέκονται στην κατάλυση. Τα κατάλοιπα αυτά μιμούνται τις αλληλεπιδράσεις που δημιουργούνται από την πενικιλίνη όταν δεσμεύεται στο ενεργό κέντρο της β-λακταμάσης (Petrosino *et al.* 1999). Η κρυσταλλική δομή δείχνει ότι οι δυο στροφές της BLIP εισέρχονται στο ενεργό κέντρο της β-λακτομής για να εμποδίσουν την δέσμευση των αντιβιοτικών β-λακταμών.

Το ενδιαφέρον όμως στράφηκε και σε ένα πεπτίδιο έξι αμινοξέων που αποτελείται από τα κατάλοιπα 46 με 51 και έχει δομή στον κρύσταλλο β-στροφής τύπου II' και αλληλεπιδρά με το ενεργό κέντρο της TEM-1β-λακταμάσης. Προκειμένου να μελετηθεί αν το πεπτίδιο διατηρεί την ικανότητα δέσμευσης ώστε να αλληλεπιδρά με το ενεργό κέντρο των β-λακταμασών οι Rudgers *et al.* (2001) δημιούργησαν πεπτίδια που περιέχουν τα έξι αυτά κατάλοιπα 46-AAGDYY-51. Για ένα πεπτίδιο που περιείχε αυτή την αλληλουχία βρέθηκε ότι η ικανότητα δέσμευσης του ήταν επαρκής για την αλληλεπίδραση με την TEM-1. Επιπλέον βρέθηκε ότι δεν αναστέλλει μόνο αυτή αλλά και μια β-λακταμάση κλάσης A και μια κλάσης C. Η ικανότητα αυτή του BLIP πεπτιδίου δείχνει ότι ίσως να έχει ευρύ φάσμα δραστηριότητας και πως ίσως να μπορεί να χρησιμοποιηθεί για την ανάπτυξη νέων μικρών μορφών αναστολέων ενάντια σε μια ποικιλία β-λακταμασών και PBPs (penicillin binding proteins) καθώς η κρυσταλλική τους δομή φανερώνει ότι στην περιοχή του ενεργού τους κέντρου έχουν όμοια τρισδιάστατη δομή.

Την ίδια περίοδο οι Gao *et al.* (2002) μελέτησαν τη δομή του πεπτιδίου χρησιμοποιώντας το ως μοντέλο για την κατανόηση των παραγόντων που συμβάλουν στο σχηματισμό β-στροφών και του προβλήματος της αναδίπλωσης των πρωτεϊνών. Η γνώση της δομής των πεπτιδίων που δρουν ως αναστολείς πρωτεϊνών μπορεί να χρησιμοποιηθεί μετέπειτα στο σχηματισμό πεπτιδίων-αναστολέων. Για να προσδιορίσουν τη διαμόρφωση που υιοθετεί το πεπτίδιο χρησιμοποίησαν τις τεχνικές Far-UV circular dichroic spectroscopy, Infrared spectroscopy, φασματοσκοπία πυρηνικού μαγνητικού συντονισμού και restrained molecular dynamics simulation. Το CD φάσμα καθώς και η φασματοσκοπία υπερύθρου που πραγματοποιήθηκαν σε

υδατικό διάλυμα (50mM phosphate buffer, στους 20° C) υποστηρίζουν την υπόθεση ότι το πεπτιδίο υιοθετεί τη δομή β-στροφής.

Για να λύσουν τη δομή του πεπτιδίου χρησιμοποίησαν φασματοσκοπία NMR σε διάλυμα DMSO-d₆ και υπολογιστικές μεθόδους. Από τα αποτελέσματα που πήραν από το φάσμα NOEs χρησιμοποίησαν τις κορυφές αN(i,i+2) των καταλοίπων Gly3-Tyr5 και αN(i,i+1) μεταξύ των Asp4-Tyr5 και Gly3-Asp5 και κατασκεύασαν μια ομάδα 12 πιθανών δομών. Η υπέρθεση των δομών αυτών επί της κύριας αλυσίδας δείχνει ότι το πεπτιδίο B1 έχει δομή β-στροφής με διεδρες γωνίες [(Φ₃,φ₃)(Φ₄,φ₄)] ίσες με [(-52°, -32°)(-38°, -44°)] που αντιστοιχούν στον τύπο III β-στροφής (Εικόνα 1.2). Επιπλέον στοιχεία που υποστηρίζουν την υπόθεση ότι η δομή του πεπτιδίου είναι αυτή της β-στροφής είναι η ύπαρξη ενός δεσμού υδρογόνου μεταξύ των καταλοίπων Ala2 και Tyr5 καθώς και η μη ένδειξη ύπαρξης αλληλεπίδρασης μεταξύ των καταλοίπων Ala1 και Tyr6 ή Ala2 και Tyr5 καθώς το φάσμα NOEs δεν έδειξε τα κατάλοιπα αυτά να βρίσκονται αρκετά κοντά στο χώρο.



Εικόνα 1.2 Υπέρθεση των 12 δομών του πεπτιδίου B1 σύμφωνα με το φάσμα NOEs (αναπαράγεται άνευ αδειας από Gao et al., 2002)

Τα στοιχεία που έχουμε λοιπόν ως τώρα και περιγράφουν τη δομή της αλληλουχίας των αμινοξέων AAGDYY-NH₂ υποδεικνύουν ότι στο διάλυμα DMSO

το πεπτίδιο υιοθετεί τη μορφή β-στροφής τύπου III σύμφωνα με τα αποτελέσματα της έρευνας που έκαναν οι Gao et al και ότι η ίδια αλληλουχία αμινοξέων στην πρωτεΐνη BLIP σχηματίζει β-στροφή τύπου II' (στο κρυσταλλικό σύμπλοκο TET-1/BLIP). Σημαντικό ρόλο σε αυτή την περίπτωση φαίνεται ότι παίζει το περιβάλλον στο οποίο βρίσκεται το πεπτίδιο.

Στα πλαίσια της κατανόησης του προβλήματος της αναδίπλωσης των πρωτεϊνών μελετάται η δομή μικρότερων και λιγότερο πολύπλοκων, σε σχέση με τις πρωτεΐνες, πεπτιδίων που δημιουργούν σταθερές δομές, όπως οι β-στροφές, ένα δομικό μοτίβο που επικρατεί για παράδειγμα στις σφαιρικές πρωτεΐνες (Marcelino C. et al. 2008). Ωστόσο πολλά ερωτήματα παραμένουν αναπάντητα σχετικά με τους παράγοντες που συμβάλουν στο σχηματισμό και τη σταθερότητα των β-στροφών. Ένα σημαντικό εργαλείο που χρησιμοποιείται για την καλύτερη κατανόηση της δομής και της λειτουργίας των βιολογικών μορίων είναι οι προσομοιώσεις μοριακής δυναμικής.

Στην εργασία αυτή ως σύστημα μελέτης επιλέχθηκε το γραμμικό πεπτίδιο AAGDYY-NH₂ στο νερό και με τη χρήση προσομοιώσεων μοριακής δυναμικής ερευνήθηκε η τάση που έχει να υιοθετεί μια σταθερή διαμόρφωση. Επίσης η σύγκριση των αποτελεσμάτων αυτών με αυτά των Gao et al θα δείξουν κατά πόσο ταιριάζουν τα μοντέλα μεταξύ τους και κατά πόσο το διαφορετικό περιβάλλον επηρεάζει τη δομή του πεπτιδίου.

Κεφάλαιο 2

Προσομοιώσεις Μοριακής Δυναμικής

2.1 Προσομοιώσεις Μοριακής Δυναμικής με λίγα λόγια

Το στατικό μοντέλο των βιομορίων που λαμβάνουμε για παράδειγμα από τη κρυσταλλογραφία ακτίνων Χ παρόλο που μας παρέχει πολύτιμες πληροφορίες σχετικά με τη δομή των βιομορίων είναι ωστόσο ανεπαρκές για να εξαγάγουμε συμπεράσματα για τη μεγάλη ποικιλία της βιολογικής τους δράσης. Η αντιμετώπιση των πρωτεϊνών ως άκαμπτες δομές αντικαταστάθηκε από το δυναμικό μοντέλο όπου τα άτομα που το αποτελούν αλληλεπιδρούν συνεχώς μεταξύ τους αλλά και με το περιβάλλον. Οι εσωτερικές κινήσεις των ατόμων που έχουν ως αποτέλεσμα την αλλαγή της διαμόρφωσης έχουν καίριο ρόλο στη λειτουργία τους. Ένα σημαντικό εργαλείο για τη μελέτη των βιολογικών μακρομορίων σε θεωρητικό επίπεδο είναι οι προσομοιώσεις μοριακής δυναμικής, μια υπολογιστική μέθοδος που μελετά τη χρονικά εξαρτώμενη συμπεριφορά μοριακών συστημάτων.

Με τις προσομοιώσεις μοριακής δυναμικής στοχεύουμε στην μελέτη των μακροσκοπικών ιδιοτήτων ενός συστήματος όπως ο υπολογισμός της ενέργειας και του μηχανισμού αλλαγής της διαμόρφωσης. Η σύνδεση μεταξύ των μικροσκοπικών προσομοιώσεων και των μακροσκοπικών ιδιοτήτων γίνεται μέσω της στατιστικής μηχανικής (Wilde and Singh, 1998). Έτσι μια μικροσκοπική ιδιότητα μετατρέπεται σε μακροσκοπική όπως η πίεση, η ενέργεια και η θερμοότητα. Σύμφωνα με το θεμελιώδες αξίωμα της στατιστικής μηχανικής ο συνολικός μέσος (ensemble average) είναι ίσος με το χρονικό μέσο (time average) που υπολογίζεται από τις προσομοιώσεις μοριακής δυναμικής (2.1). Η βασική ιδέα λοιπόν είναι να αφήσουμε το σύστημα να εξελίσσεται στο χώρο ώστε να περάσει από όλες τις πιθανές θέσεις. Ένας στόχος των

προσομοιώσεων είναι να παράγει αρκετές αντιπροσωπευτικές διαμορφώσεις τέτοιες που να ικανοποιούν αυτή την ισότητα. Σε αυτή την περίπτωση πειραματικά δεδομένα σχετικά με τη δομή, τη δυναμική και τη θερμοδυναμική είναι δυνατόν να υπολογιστούν χρησιμοποιώντας ένα επαρκές σύνολο υπολογιστικών δεδομένων. Η διάρκεια της προσομοίωσης πρέπει να είναι τέτοια που να λαμβάνουμε αυτόν τον επαρκή αριθμό διαμορφώσεων.

$$\langle A \rangle_{\text{ensemble}} = \langle A \rangle_{\text{time}} \quad 2.1$$

Ensemble average = Time average

Εκτός από τη στατιστική μηχανική οι προσομοιώσεις μοριακής δυναμικής βασίζονται και στην κλασική μηχανική. Με την εξίσωση κίνησης ή 2ο Νόμο του Νεύτωνα, γνωρίζοντας τη δύναμη που ασκείται στο κάθε άτομο είναι δυνατό να καθοριστεί η επιτάχυνση του κάθε ατόμου στο σύστημα. Η ολοκλήρωση της εξίσωσης κίνησης μας δίνει τη τροχιά που περιγράφει τη θέση, τη ταχύτητα και την επιτάχυνση των μορίων όπως μεταβάλλονται σε σχέση με το χρόνο. Η μέθοδος είναι ντετερμινιστική αφού με δεδομένες τις ακριβείς θέσεις και ταχύτητες όλων των μορίων σε δεδομένη χρονική στιγμή μπορούμε να υπολογίσουμε τις θέσεις και ταχύτητες όλων των μορίων σε κάθε άλλη χρονική στιγμή στο μέλλον (ή το παρελθόν).

Η εξίσωση κίνησης του Νεύτωνα είναι :

$$F_i = m_i a_i$$

όπου F_i είναι η δύναμη που ασκείται στο άτομο i , m_i είναι η μάζα του και a_i η επιτάχυνση του. Η δύναμη όμως μπορεί να εκφραστεί και ως μεταβολή της δυναμικής ενέργειας:

$$F_i = - dV/dr_i$$

από τις δυο αυτές εξισώσεις έχουμε:

$$-dV/dr_i = m_i d^2r_i/dt^2$$

όπου V είναι η δυναμική ενέργεια του συστήματος. Η εξίσωση κίνησης μπορεί να συσχετίσει την παράγωγο της δυναμικής ενέργειας με την αλλαγή της θέσης ως προς το χρόνο.

Η δυναμική ενέργεια ορίζεται ως το άθροισμα της ενέργειας των εσωτερικών αλληλεπιδράσεων δεσμικών ατόμων (μήκη δεσμών, γωνίες και περιστροφές δεσμών) με την ενέργεια των εξωτερικών αλληλεπιδράσεων μη δεσμικών ατόμων (ηλεκτροστατικές δυνάμεις και δυνάμεις van der Waals).

$$V(\mathbf{R}) = E_{\text{bonded}} + E_{\text{non-bonded}}$$

$$E_{\text{bonded}} = E_{\text{bond-stretch}} + E_{\text{angle-bond}} + E_{\text{rotate-along-bond}}$$

$$E_{\text{non-bonded}} = E_{\text{van der Waals}} + E_{\text{electrostatics}}$$

Οι προσομοιώσεις μοριακής δυναμικής επιτρέπουν τη μελέτη πολύπλοκων δυναμικών διεργασιών που συμβαίνουν στα βιολογικά συστήματα. Αυτές περιλαμβάνουν τη σταθερότητα των πρωτεϊνών, τις αλλαγές στη δομή, την πρωτεϊνική αναδίπλωση, την μοριακή αναγνώριση (πρωτεΐνες, DNA, μεμβράνες, σύμπλοκα) καθώς και τη μεταφορά των ιόντων στα βιολογικά συστήματα. Ακόμα αποτελούν μέσο για την διεξαγωγή ερευνών στο σχεδιασμό φαρμάκων και τον προσδιορισμό της δομής (από ακτίνες X και MNR).

Οι εφαρμογές των μεθόδων προσομοίωσης στον μακρομοριακό τομέα χωρίζονται σε τρεις τύπους (Karplus et al., 2002). Ο πρώτος αναφέρεται στην απλή μελέτη της διαμόρφωσης στο χώρο (sampling configuration space) για τον προσδιορισμό ή τη βελτιστοποίηση των δομών με δεδομένα που παίρνουμε από τα πειράματα. Ο δεύτερος χρησιμοποιείται για να περιγράψουμε ένα σύστημα σε κατάσταση ισορροπίας όπως και τις δομικές και κινητικές του ιδιότητες καθώς και των θερμοδυναμικών παραμέτρων. Στον τρίτο, οι προσομοιώσεις χρησιμοποιούνται για να εξετάσουμε τη δυναμική του συστήματος στο χρόνο.

Η ανάπτυξη των υπολογιστικών μεθόδων έφερε αλλαγές στη σχέση μεταξύ θεωρίας και πειράματος (Daura X. et al., 1999). Από την μια πλευρά, οι προσομοιώσεις αύξησαν τις απαιτήσεις για την ακρίβεια των μοντέλων. Από την άλλη πλευρά, οι προσομοιώσεις έρχονται συχνά πολύ κοντά στις πειραματικές συνθήκες σε τέτοιο βαθμό που τα υπολογιστικά αποτελέσματα μπορούν να συγκριθούν απευθείας με τα πειραματικά αποτελέσματα. Σ' αυτή την περίπτωση, η προσομοίωση γίνεται ένα εξαιρετικά ισχυρό εργαλείο όχι μόνο για την κατανόηση

και εξήγηση των πειραμάτων σε μικροσκοπικό επίπεδο, αλλά για να μελετήσουμε τομείς που δεν είναι προσιτοί πειραματικά ή που απαιτούν πολύ ακραίες πειραματικές συνθήκες. Οι προσομοιώσεις δίνουν τη δυνατότητα να πραγματοποιηθούν θεωρητικά πειράματα που τα αποτελέσματα μας παρέχουν χρήσιμες πληροφορίες για την κατανόηση των φαινομένων. Τα πειράματα παίζουν σημαντικό ρόλο στην εγκυρότητα της μεθοδολογίας των προσομοιώσεων. Με σύγκριση των δεδομένων προσομοίωσης και των πειραματικών δεδομένων μπορούμε να εξετάσουμε την ακρίβεια των υπολογιστικών αποτελεσμάτων. Μια τέτοια σύγκριση είναι ιδιαίτερα σημαντική μια και τα συστηματικά λάθη που εισάγονται με τη χρήση των εμπειρικών δυναμικών (empirical potentials) είναι δύσκολο να υπολογιστούν με θεωρητικές εκτιμήσεις.

2.2 Στάδια προσομοίωσης

Τα βήματα που ακολουθούμε για την πραγματοποίηση των προσομοιώσεων μοριακής δυναμικής είναι η προετοιμασία του συστήματος, η ελαχιστοποίηση της ενέργειας, η εξισορρόπηση και το τελικό στάδιο που είναι η παραγωγική φάση.

Η αρχικοποίηση (initialization) του συστήματος απαιτεί τον καθορισμό των αρχικών συντεταγμένων και ταχυτήτων του μακρομορίου καθώς και του διαλύματος στο οποίο βρίσκεται και των ιόντων που πιθανών να προστεθούν. Οι αρχικές συντεταγμένες μπορεί να προέρχονται από την κρυσταλλογραφία ακτίνων X ή NMR και στην περίπτωση που δεν είναι διαθέσιμα τέτοια πειραματικά δεδομένα χρησιμοποιούνται άλλες τεχνικές όπως ομόλογη μοντελοποίηση (homology modeling) για την κατασκευή μιας θεωρητικής δομής. Ακόμα ως αρχικές συντεταγμένες και ταχύτητες μπορούν να χρησιμοποιηθούν οι τελικές τιμές που προκύπτουν από προηγούμενη προσομοίωση. Ιδιαίτερη προσοχή χρειάζεται στην ενυδάτωση του μορίου και την προσθήκη ιόντων προκειμένου να πετύχουμε συνθήκες όσο γίνεται πιο κοντά σ' αυτές που επικρατούν σε ένα φυσιολογικό υδατικό περιβάλλον. Για τον προσδιορισμό των ορίων του συστήματος εφαρμόζονται περιοδικές οριακές συνθήκες και επιτυγχάνεται αποτελεσματικά η ελάττωση της επιφανειακής τάσης των μορίων νερού. Έτσι μια προσομοίωση εκτελείται χρησιμοποιώντας ένα σχετικά μικρό αριθμό μορίων με τέτοιο τρόπο ώστε τα μόρια να υφίσταται δυνάμεις σαν να βρίσκεται σε όγκο διαλύματος. Στις περιοδικές οριακές συνθήκες το σύστημα τοποθετείται σε ένα κεντρικό κουτί και θεωρείται ότι έχει

απειρώς πολλά είδωλα και στις τρεις διαστάσεις. Το κεντρικό κουτί μπορεί να έχει σχήμα κυβικό, εξαγωνικού πρίσματος, ρομβικού δωδεκαέδρου, επιμηκυμένου δωδεκαέδρου ή κολοβωμένου οκταέδρου. Η τοποθέτηση των ιόντων γίνεται με ειδικούς αλγορίθμους (ion-placement algorithms) που είναι διαθέσιμοι στα προγράμματα μοριακής δυναμικής ώστε να πετύχουμε την επιθυμητή συγκέντρωση ιόντων. Τέλος, η αρχική τιμή των ταχυτήτων επιλέγεται τυχαία από μια κατανομή Maxwell ώστε η αρχική ορμή να είναι μηδέν.

Ακολουθεί η ελαχιστοποίηση της δυναμικής ενέργειας του συστήματος σύμφωνα με τα δυναμικά πεδία (force fields) για τις αρχικές συντεταγμένες. Ο στόχος αυτής της διαδικασίας είναι να οδηγηθούμε σε μια σταθερή προσομοίωση.

Από την στιγμή που έχουν οριστεί οι αρχικές τιμές των συντεταγμένων και ταχυτήτων το σύστημα θερμαίνεται (heating) σταδιακά ώστε να φτάσει στην επιθυμητή θερμοκρασία όπου θα πραγματοποιηθεί η προσομοίωση. Το επόμενο βήμα είναι η εξισορρόπηση (equilibration). Ο σκοπός για τον οποίο εκτελείται αυτό το βήμα είναι οι ιδιότητες του συστήματος όπως η πίεση, η θερμοκρασία και η ενέργεια να σταθεροποιηθούν στο χρόνο.

Το τελικό βήμα στη προσομοίωση είναι να αφήσουμε το σύστημα να τρέξει στην παραγωγική φάση (production phase) για το επιθυμητό χρονικό διάστημα που μπορεί να κυμαίνεται από αρκετές εκατοντάδες ps μέχρι ns ή περισσότερο. Στη διάρκεια αυτής της φάσης υπολογίζονται οι θερμοδυναμικές παράμετροι του συστήματος.

Με το τέλος της προσομοίωσης παίρνουμε ένα πλήθος δεδομένων που τα χρησιμοποιούμε για να εξάγουμε συμπεράσματα κατά την ανάλυση. Οι χρονικά εξαρτώμενες ιδιότητες μπορούν να προσδιοριστούν γραφικά με έναν από τους άξονες να αντιστοιχεί στο χρόνο και τον άλλο στον παράγοντα που μας ενδιαφέρει όπως η ενέργεια, το rmsd κ.τ.λ. Οι μέσες δομές μπορούν να υπολογιστούν και να συγκριθούν με τις πειραματικές δομές. Οι προσομοιώσεις μοριακής δυναμικής βοηθούν στην απεικόνιση και κατανόηση των αλλαγών στη διαμόρφωση σε ατομικό επίπεδο όταν συνδυάζονται με προγράμματα γραφικών που μπορούν να παρουσιάσουν τις δομικές παραμέτρους σε σχέση με το χρόνο.

2.3 Περιορισμοί

Έχουν περάσει πάνω από τριάντα χρόνια από τη δημοσίευση της πρώτης προσομοίωσης μακρομοριακού βιολογικού συστήματος που αφορούσε μια μικρή πρωτεΐνη την BPTI (McCammon *et al.*, 1977). Από τότε γίνονται συνεχείς προσπάθειες για να προσεγγιστεί μέσω των προσομοιώσεων μια πιο ρεαλιστική αναπαράσταση των βιολογικών συστημάτων αυτό όμως έχει ως αποτέλεσμα την εισαγωγή πολλών περιορισμών στους υπολογισμούς. Ο ρεαλισμός των δυναμικών τροχιακών εξαρτάται από τις εμπειρικές συναρτήσεις της δυναμικής ενέργειας $V(R)$, τον υπολογισμό των δυνάμεων μεγάλου εύρους (ηλεκτροστατικές αλληλεπιδράσεις και Lennard-Jones), τη διάρκεια του χρονικού βήματος Δt , κλπ.

Οι θεωρητικές μελέτες των βιολογικών μορίων επιτρέπουν τη μελέτη της σχέσης μεταξύ δομής, λειτουργίας και δυναμικής σε ατομικό επίπεδο. Τα βιολογικά συστήματα αποτελούνται από μεγάλο αριθμό ατόμων γεγονός που κάνει ανέφικτους αυτούς τους υπολογισμούς χρησιμοποιώντας τη κβαντομηχανική. Τη λύση στο πρόβλημα δίνουν οι εμπειρικές συναρτήσεις δυναμικής ενέργειας ή δυναμικά πεδία που έχουν λιγότερο υπολογιστικό κόστος (MacKerell *et al.*, 1998). Στην ανάπτυξη των δυναμικών πεδίων μεγάλη ήταν η συμβολή των Bruce και Karplus που η προσπάθεια τους, με δεδομένη την αμινοξική αλληλουχία και ένα σύνολο συντεταγμένων (όπως αυτές που παίρνουμε από δομές ακτίνων-X), να υπολογίσουν την ενέργεια του συστήματος και άλλων παραγώγων της ως συνάρτηση των ατομικών θέσεων είχε ως αποτέλεσμα τη δημιουργία του προγράμματος που αργότερα ονομάστηκε Pre-CHARMM (Karplus 2003). Οι πιο ευρέως χρησιμοποιούμενες συναρτήσεις δυναμικής ενέργειας είναι οι AMBER, CHARMM, GROMOS και OPLS/AMBER (Ponder and Case, 2003), (Guvench and MacKerell, 2008). Τα δυναμικά αυτά πεδία εξασφαλίζουν μια ικανοποιητική λύση μεταξύ της ακρίβειας των αποτελεσμάτων και της υπολογιστικής επάρκειας. Η ικανότητα τους να αναπαράγουν φυσικές ιδιότητες γίνεται με σύγκριση των αποτελεσμάτων με πειραματικά δεδομένα.

Η δυναμική ενέργεια με τα δυναμικά πεδία υπολογίζεται ως το άθροισμα των ενεργειών λόγω των αλληλεπιδράσεων δεσμικών ατόμων και μη δεσμικών ατόμων. Παρ' όλες τις προσπάθειες για βελτιστοποίηση των δυναμικών πεδίων λόγω των προσεγγίσεων που γίνονται υπάρχουν ορισμένοι περιορισμοί.

Ένας από τους πιο σημαντικούς περιορισμούς είναι ότι καμία δραστική αλλαγή όπως η δημιουργία ή το σπάσιμο δεσμών δεν επιτρέπονται στην

ηλεκτρονιακή δομή. Για να δοθεί λύση στο πρόβλημα γίνονται μελέτες για την ανάπτυξη συνδυασμού δυναμικών πεδίων κβαντομηχανικής και μοριακής μηχανικής. Κατά τον προσδιορισμό των παραμέτρων στα δυναμικά πεδία τα άτομα αντιπροσωπεύονται από ατομικούς τύπους μιας συγκεκριμένης δεσμικής κατάστασης (π.χ. ένα άτομο άνθρακα σε έναν δεσμό sp^3 έχει διαφορετικές ιδιότητες από ένα άτομο άνθρακα που βρίσκεται στο δακτύλιο ιστιδίνης). Έχοντας ως σκοπό τη μείωση του αριθμού των ατομικών τύπων προχωρούν σε ομαδοποίηση τους και έτσι κάθε άτομο στο μόριο δεν παρουσιάζεται ως μοναδικό με μοναδική σειρά παραμέτρων με αποτέλεσμα τον ανακριβή υπολογισμό της δυναμικής ενέργειας, Άτομα όπως οι αλειφατικοί άνθρακες ή το υδρογόνο που οι ιδιότητες τους δεν είναι τόσο ευαίσθητες σε σχέση με το περιβάλλον τους μια απλή σειρά παραμέτρων μπορεί να δουλέψει αρκετά καλά. Αντίθετα άτομα όπως το οξυγόνο και το άζωτο που επηρεάζονται περισσότερο από τα γειτονικά τους άτομα απαιτούν περισσότερους τύπους και παραμέτρους για τα διαφορετικά δεσμικά περιβάλλοντα.

Μια προσέγγιση που έγινε για την μείωση του υπολογιστικού κόστους είναι η κατά ζεύγη προσθετική προσέγγιση (pair-wise additive approximation). Σ' αυτή την περίπτωση η ενέργεια αλληλεπίδρασης μεταξύ ενός ατόμου και του υπόλοιπου συστήματος υπολογίζεται ως άθροισμα των κατά-ζεύγη αλληλεπιδράσεων (άτομο με άτομο), σαν τα ζεύγη ατόμων να είναι απομονωμένα από τα άλλα άτομα του συστήματος. Η ταυτόχρονη αλληλεπίδραση μεταξύ τριών ή περισσότερων ατόμων δεν υπολογίζεται με αποτέλεσμα να μην συμπεριλαμβάνονται ορισμένα φαινόμενα πόλωσης. Αυτό μπορεί να οδηγήσει σε λεπτές διαφορές μεταξύ υπολογισθέντων και πειραματικών αποτελεσμάτων π.χ. διαφορά στη τιμή pK .

Μια σημαντική παράληψη στη συνάρτηση της δυναμικής ενέργειας είναι η εντροπία. Μια ελάχιστη τιμή E που υπολογίζεται ως άθροισμα των δυναμικών συναρτήσεων δεν αντιστοιχεί απαραίτητα στην ισορροπία ή την πιο πιθανή κατάσταση του συστήματος, αυτό αντιστοιχεί στην ελάχιστη τιμή της ελεύθερης ενέργειας. Εξαιτίας του γεγονότος ότι τα πειράματα διεξάγονται γενικώς υπό ισόθερμες-ισοβαρείς συνθήκες (σταθερή πίεση, θερμοκρασία και μέγεθος του συστήματος) η θέση ισορροπίας αντιστοιχεί στο ελάχιστο της ελεύθερης ενέργειας του Gibbs, G . Ενώ, δηλαδή, οι ενεργειακοί υπολογισμοί αγνοούν τα εντροπικά φαινόμενα του συστήματος, αυτά περιλαμβάνονται στις προσομοιώσεις μοριακής δυναμικής.

Μια τυπική προσομοίωση εκτελείται σε συστήματα που περιέχουν μεγάλο αριθμό ατόμων. Το υπολογιστικό κόστος σε κάθε βήμα είναι μεγάλο για τον υπολογισμό της ενέργειας και των παραγώγων της και περιορίζεται ο συνολικός χρόνος της προσομοίωσης. Παρ' όλα αυτά η ανάγκη για να μελετηθούν πιο περίπλοκα και ρεαλιστικά μοντέλα είναι μεγαλύτερη από την ανάγκη για αύξηση της χρονικής διάρκειας της προσομοίωσης. Για την ολοκλήρωση της εξίσωσης κίνησης των προσομοιώσεων μοριακής δυναμικής χρησιμοποιούνται οι αλγόριθμοι ολοκλήρωσης. Όλοι οι αλγόριθμοι ολοκλήρωσης για τον προσδιορισμό των θέσεων, ταχυτήτων και επιταχύνσεων βασίζονται στην σειρά εξισώσεων Taylor (Chang et al., 2006):

$$\begin{aligned}r(t+\delta t) &= r(t) + v(t)\delta t + 1/2a(t)\delta t^2 + \dots \\v(t+\delta t) &= v(t) + a(t)\delta t + 1/2b(t)\delta t^2 + \dots \\a(t+\delta t) &= a(t) + b(t)\delta t + \dots\end{aligned}$$

όπου r είναι η θέση, v η ταχύτητα (η πρώτη παράγωγος του χρόνου), a η επιτάχυνση (η δεύτερη παράγωγος του χρόνου) κ.τ.λ.

Αυτές οι εξισώσεις είναι προσεγγιστικές και υπεισέρχονται λάθη που σχετίζονται με την περικοπή της επέκτασης (truncation errors) της εξίσωσης Taylor και στα λάθη ολοκλήρωσης (round-off errors). Τα πρώτα δεν εξαρτώνται από την εφαρμογή του αλλά εντοπίζονται στον αλγόριθμο, ενώ τα λάθη ολοκλήρωσης σχετίζονται με μια συγκεκριμένη εφαρμογή του αλγόριθμου (π.χ. στον πεπερασμένο αριθμό ψηφίων που χρησιμοποιούνται στους υπολογισμούς). Και τα δύο μπορούν να περιοριστούν με μείωση του χρονικού διαστήματος Δt (Ercolessi, 1997).

Για τη μείωση του υπολογιστικού κόστους των προσομοιώσεων μπορούμε να μειώσουμε τους υπολογισμούς ανά χρονικό βήμα ή να αυξήσουμε το χρονικό βήμα ολοκλήρωσης Δt (Schlick 2001).

Η μείωση των υπολογισμών ανά βήμα μπορεί να επιτευχθεί με πολλούς τρόπους όπως η χρήση cutoff ranges για τις μη-δεσμικές αλληλεπιδράσεις, η σπάνια ανανέωση της λίστας μη-δεσμικών ζευγών και η μείωση των βαθμών ελευθερίας με περιορισμό των συγκεκριμένων περιοχών του μορίου.

Η μέγιστη τιμή για το χρονικό βήμα που μπορεί να χρησιμοποιηθεί επιτυχώς περιορίζεται από την ακρίβεια και σταθερότητα. Η αστάθεια ενός τροχιακού ανιχνεύεται εύκολα όταν οι συντεταγμένες και οι ενέργειες αυξάνονται ανεξέλεγκτα. Αυτό συχνά σημαίνει ότι το χρονικό βήμα είναι πολύ μεγάλο. Ωστόσο, αν δεν είναι

πολύ μεγάλο το βήμα, αυτή η αστάθεια ίσως είναι φαινόμενη ενός τροχιακού διάρκειας εκατοντάδων ps ή μεγαλύτερη.

Η ακρίβεια περιορίζει τα χρονικά βήματα με τη λογική ότι η λύση που παράγεται πρέπει να είναι κατάλληλη για τη διαδικασία που ακολουθήθηκε. Έτσι μια προσομοίωση ίσως είναι σταθερή αλλά το χρονικό βήμα πολύ μεγάλο για να συλλάβει τις διακυμάνσεις υψηλής συχνότητας με ακρίβεια.

Ο αλγόριθμος που χρησιμοποιείται συχνότερα είναι ο Verlet που υπολογίζει τη θέση χρησιμοποιώντας τις θέσεις από τα δυο προηγούμενα βήματα και τις επιταχύνσεις από το προηγούμενο βήμα σύμφωνα με την εξίσωση:

$$r(t+\delta t)=2r(t)-r(t-\delta t)+a(t)\delta t^2$$

Ένας τέτοιος αλγόριθμος μπορεί εύκολα να τροποποιηθεί για να εισάγει περιορισμούς στις εσωτερικές συντεταγμένες, όπως μήκη δεσμών και γωνίες.

Μιας και οι εξισώσεις κίνησης λύνονται προσεγγιστικά σε διακριτά χρονικά βήματα οι περιορισμοί που εισάγονται δεν θα τηρούνται όσο προχωρά η προσομοίωση. Τη λύση στο πρόβλημα αυτό έδωσε ο προσδιορισμό των δυνάμεων περιορισμού με τέτοιο τρόπο ώστε οι περιορισμοί να ικανοποιούνται ακριβώς στο τέλος κάθε χρονικού βήματος. Ένας τέτοιος αλγόριθμος είναι ο SHAKE.

Όπως έχει είδη αναφερθεί, οι προσομοιώσεις μοριακής δυναμικής γίνονται σε περιοδικές οριακές συνθήκες (Periodic Boundary Conditions). Σ' αυτές τις συνθήκες το σύστημα προς μελέτη τοποθετείται σε ένα κουτί και θεωρείται ότι περιβάλλεται από πανομοιότυπα κουτιά και στις τρεις διαστάσεις. Στην περίπτωση που ένα άτομο ξεπεράσει το όριο του κουτιού ένα αντίγραφο του εισάγεται για να το αντικαταστήσει και έτσι ο αριθμός των μορίων παραμένει σταθερός. Στις προσομοιώσεις των βιομορίων το κατώφλι (cutoff) ορίζεται έτσι ώστε τα άτομα στο κεντρικό κουτί να μην βλέπουν την εικόνα τους στα γειτονικά κουτιά. Οι δυνάμεις που ασκούνται στα κύρια μόρια υπολογίζονται από τα μόρια του ίδιου κουτιού όπως και από τα μόρια των γειτονικών.

Μια από τις πιο χρονοβόρες διαδικασίες στις προσομοιώσεις μοριακής δυναμικής είναι ο υπολογισμός των μεγάλου εύρους ηλεκτροστατικών αλληλεπιδράσεων. Με το να αυξήσουμε το κατώφλι της απόστασης μπορεί να έχουμε δραματική αύξηση του υπολογιστικού κόστους. Αν και πολύ συχνά οι

ηλεκτροστατικές αλληλεπιδράσεις μεγάλου εύρους αγνοούνται, υπάρχουν περιπτώσεις που παίζουν σημαντικό ρόλο (π.χ. οι αλληλεπιδράσεις van der Waals μεταξύ ατόμων είναι σημαντικές κάτω από συγκεκριμένες γεωμετρικές συνθήκες).

Έχουν αναπτυχθεί πολλά μοντέλα που επιτρέπουν τον υπολογισμό των ηλεκτροστατικών επιδράσεων μεγάλου εύρους όπως το άθροισμα Ewald στην προσομοίωση πρωτεϊνών και ενζύμων σε κρυσταλλική φάση. Η μέθοδος Particle Mesh Ewald (PME) (Essmann et al., 1995), μια παραλλαγή του αθροίσματος Ewald, χρησιμοποιείται για τον υπολογισμό της ηλεκτροστατικής ενέργειας σε περιοδικά συστήματα.

Θεωρώντας ότι έχουμε να κάνουμε με φορτισμένα ιόντα, υποθέτουμε ότι κάθε ιόν περιβάλλεται από μια σφαιρικά συμμετρική κατανομή με αντίθετο φορτίο που εξουδετερώνει το ιόν. Αυτή η κατανομή είναι περιορισμένη τοπικά δηλαδή περιέχεται σε ένα κατώφλι απόστασης (cutoff distance) που προστατεύει αποτελεσματικά τις αλληλεπιδράσεις μεταξύ ιόντων.

Η μέθοδος Ewald έχει τις βάσεις της στην κρυσταλλογραφία (Deserno and Holm, 1998). Για τον υπολογισμό της ηλεκτροστατικής ενέργειας εισάγονται δυο όροι για τα ζεύγη ατόμων στο άμεσο χώρο (πλέγμα) και για τα ζεύγη ατόμων που αντιστοιχούν στις αλληλεπιδράσεις με τα είδωλα των ατόμων του κεντρικού κουτιού. Ο πρώτος υπολογίζεται με άμεσο υπολογισμό και ο δεύτερος με μετασχηματισμό Fourier. Οι Fast Fourier μετασχηματισμοί χρησιμοποιούνται για τον υπολογισμό της ηλεκτροστατικής ενέργειας που αντιστοιχεί στο αντίστροφο πλέγμα.

Κεφάλαιο 3

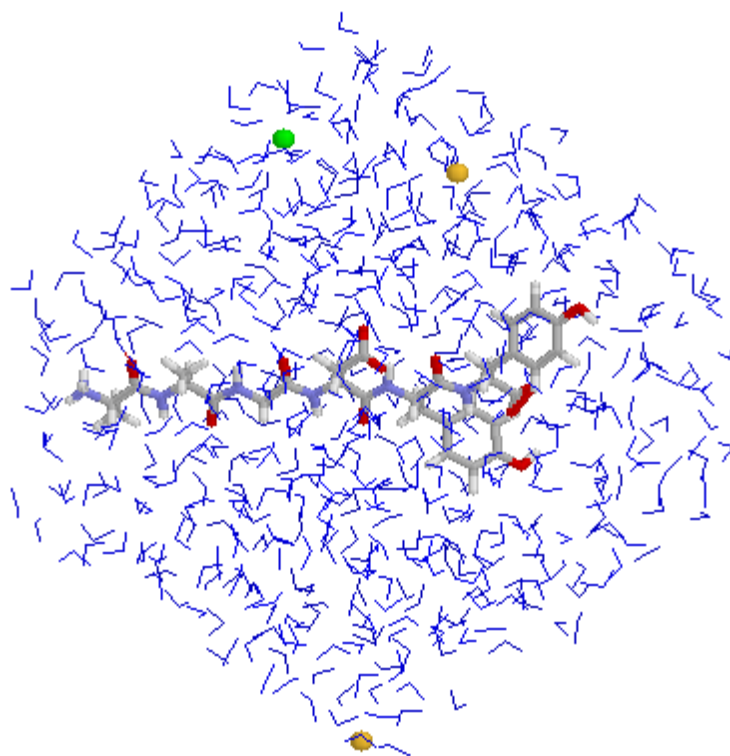
Προσομοίωση πεπτιδίου

3.1 Προετοιμασία του συστήματος

Σ' αυτήν την εργασία πραγματοποιήθηκαν τέσσερις ανεξάρτητες προσομοιώσεις με τις ίδιες παραμέτρους. Το πεπτίδιο ξεκίνησε την προσομοίωση από την εκτεταμένη διαμόρφωση (extended conformation). Το αρχείο για το πεπτίδιο δημιουργήθηκε χρησιμοποιώντας το πρόγραμμα ribosome. Στη συνέχεια με το πρόγραμμα moleman (Kleywegt, 1995) θέσαμε το κέντρο βάρους του στην αρχή των αξόνων και ευθυγραμμίστηκαν οι αδρανειακοί άξονες (axes of inertia) σύμφωνα με το ορθοκανονικό σύστημα αξόνων. Για να τελειοποιήσουμε τη δομή του πεπτιδίου (συμπλήρωση ατόμων) χρησιμοποιήθηκε το πρόγραμμα psfgen που είναι τμήμα του προγράμματος NAMD και έτσι δημιουργήθηκε ένα νέο αρχείο pdb με το αρχικό πεπτίδιο και ένα αρχείο psf με τις πληροφορίες της δομής του που χρειάζονται για το NAMD (Phillips J. C., 2005).

Στην συνέχεια έγινε η ενυδάτωση του πεπτιδίου. Οι προσομοιώσεις γίνονται υπό περιοδικές οριακές συνθήκες. Έτσι το πεπτίδιο τοποθετείται σε ένα κολοβωμένο οκτάεδρο (truncated octahedron), ώστε να έχουμε όσο γίνεται πιο συμπυκνωμένο πλέγμα, με μόρια νερού τύπου TIP3. Παρόλο που τα μόρια νερού τύπου TIP3 δεν συμφωνούν με τις πειραματικά προσδιορισμένες δομικές ιδιότητες του νερού, όπως για παράδειγμα οι τύποι SPCE και TIP5 (Pasciuk 2004), αναπαριστούν καλύτερα την ενέργεια και την πυκνότητα του υγρού νερού και αναμένεται να είναι πιο κοντά στην πραγματικότητα. Ένας ακόμα λόγος για την επιλογή του τύπου TIP3 είναι η συμβατότητα του με το πρόγραμμα CHARMM (Patel *et al.*, 2004).

Στο σύστημα προστίθενται ο απαραίτητος αριθμός ιόντων, δυο ιόντα Na^+ και ένα ιόν Cl^- . Σκοπός είναι να εξουδετερωθεί το φορτίο του συστήματος και να προκύψει ουδέτερο σύστημα. Αυτή είναι μια προϋπόθεση που απαιτείται για την εφαρμογή της μεθόδου particle-mesh Ewald (PME) (Darden et al., 1993) που χρησιμοποιείται στα περιοδικά συστήματα (*Šhirts et al., 2006*) για τον υπολογισμό της ενέργειας λόγω ηλεκτροστατικών αλληλεπιδράσεων. Η τοποθέτηση των ιόντων μπορεί να γίνει πολύ εύκολα με τη βοήθεια του προγράμματος VMD. Το σύστημα που προκύπτει αποτελείται από 1800 άτομα εκ των οποίων τα 84 είναι άτομα του πεπτιδίου, 1713 άτομα νερού, 2 άτομα Na και ένα άτομο Cl (**Εικόνα 3.1**).



Εικόνα 3.1 Η στοιχειώδης κυψελίδα που αναπαριστά το σύστημα προσομοίωσης του πεπτιδίου AAGDYY. Στο κέντρο με μορφή wireframe απεικονίζεται το πεπτίδιο. Το κίτρινο χρώμα αντιστοιχεί στα άτομα Na και το πράσινο το άτομο Cl. Με μπλε χρώμα απεικονίζεται το νερό. Η εικόνα δημιουργήθηκε με το πρόγραμμα rasmol (Bernstein, 1999).

Το τελευταίο βήμα που πρέπει να γίνει πριν ξεκινήσει η προσομοίωση είναι ο καθορισμός των ορίων του συστήματος. Μια γρήγορη μέθοδος για να καθοριστούν

είναι το πρόγραμμα pdbset. Για το σύστημα που μελετάμε τα όρια του πλαισίου είναι:

Orthogonal Coordinate limits in output file:

	Minimum	Maximum	Centre	Range
On X :	-16.69	16.78	0.05	33.47
On Y :	-16.86	16.84	-0.01	33.69
On Z :	-16.70	17.00	0.15	33.69

Οι προσομοιώσεις μοριακής δυναμικής πραγματοποιήθηκαν με το πρόγραμμα NAMD χρησιμοποιώντας το δυναμικό πεδίο CHARMM27. Το σύστημα αρχικά υπέστη ελαχιστοποίηση της ενέργειας για 2000 βήματα για τα άτομα των πλευρικών αλυσίδων και για άλλα 2000 βήματα για όλα τα άτομα. Έπειτα το σύστημα θερμάνθηκε σταδιακά μέχρι η τελική θερμοκρασία να φτάσει τους 320K με βήμα $\Delta T=20K$. Κατά τη σταδιακή θέρμανση τα άτομα Ca παραμένουν σταθερά. Μετά ακολουθεί εξισορρόπηση του συστήματος υπό ισοβαρείς-ισόθερμες συνθήκες NpT (σταθερή πίεση, θερμοκρασία και σταθερός αριθμός ατόμων) χωρίς περιορισμούς. Η πίεση και η θερμοκρασία ρυθμίζονται με το δυναμικό Nose-Hoover Langevin και έτσι η πίεση διατηρείται στη 1atm στους 320K. Για την παραγωγική φάση χρησιμοποιήθηκε ο αλγόριθμος Verlet-I και για τον προσδιορισμό των περιορισμών ο SHAKE (Ryckaert et al., 1977). Το χρονικό βήμα ήταν 2fs, οι μικρού εύρους μη δεσμικές αλληλεπιδράσεις υπολογίστηκαν κάθε δυο χρονικά βήματα και οι μεγάλου εύρους ηλεκτροστατικές αλληλεπιδράσεις κάθε τέσσερα χρονικά βήματα χρησιμοποιώντας τη μέθοδο particle-mesh Ewald. Τα τροχιακά που λαμβάνονται αποθηκεύοντας τις ατομικές συντεταγμένες όλου του συστήματος κάθε 0.4ps. Η συνολική διάρκεια των τεσσάρων προσομοιώσεων που πραγματοποιήθηκαν είναι 241ns.

3.2 Μέθοδοι ανάλυσης τροχιακών

Από τη στιγμή που έχουμε συγκεντρώσει τα δεδομένα που πήραμε από τις προσομοιώσεις μοριακής δυναμικής ακολουθεί το στάδιο της ανάλυσης των αποτελεσμάτων. Προκειμένου να εξάγουμε συμπεράσματα σχετικά με τη σταθερότητα και τις κινητικές και δυναμικές ιδιότητες του πεπτιδίου μελετήθηκε μια σειρά παραμέτρων όπως περιγράφονται αναλυτικά παρακάτω.

Μια χρήσιμη παράμετρος ενδεικτική της σταθερότητας του μορίου είναι η ρίζα της μέσης τετραγωνικής απόκλισης (root mean square deviation, rmsd). Με τον υπολογισμό του rmsd βλέπουμε την απόκλιση της διαμόρφωσης του μορίου από μια δομή αναφοράς π.χ. την αρχική ή τη μέση δομή κατά τη διάρκεια της προσομοίωσης. Το rmsd από την αρχική δομή υπολογίζεται σύμφωνα με τον τύπο:

$$\text{RMSD} = (1/N_i \sum (\mathbf{x}_i - \mathbf{x})^2)^{1/2}$$

όπου N είναι ο αριθμός των ατόμων, το x αντιστοιχεί στις συντεταγμένες των ατόμων στην αρχική δομή και το x_i αντιστοιχεί στις συντεταγμένες της εκάστοτε δομής κατά τη διάρκεια του trajectory.

Μια ακόμα παράμετρος που μας δίνει πληροφορίες για τη σταθερότητα και τη διατήρηση της δομής του μορίου είναι το radius of gyration (R_G) (Kuszewski et al., 1999). Το R_G μιας ομάδας ατόμων ορίζεται ως η απόσταση rms για κάθε άτομο του μορίου από το κέντρο βάρους του μορίου και δίνεται από την εξίσωση:

$$\text{RadiusGyration} = (1/N_i \sum (\mathbf{r}_i - \mathbf{r}_{\text{com}})^2)^{1/2}$$

όπου N είναι ο συνολικός αριθμός ατόμων, r_i είναι η απόσταση από το κέντρο βάρους του μορίου σε κάθε στιγμιότυπο και r_{com} είναι η μέση απόσταση από το κέντρο βάρους του μορίου κατά τη διάρκεια ολόκληρου του trajectory.

Προκειμένου τώρα να μελετήσουμε τις κινητικές και δυναμικές ιδιότητες του μορίου υπολογίζονται το rmsf, χάρτες αποστάσεων, πίνακες variance-covariance και οι eigenvectors-eigenvalues.

Η κινητικότητα των ατόμων σε σχέση με τη θέση τους στη μέση δομή κατά τη διάρκεια του τροχιακού δίνεται από τον υπολογισμό της ρίζας της μέσης τετραγωνικής διακύμανσης (root mean square fluctuation, rmsf). Μέσω της μελέτης του rmsf μπορούμε να προσδιορίσουμε τα περισσότερο κινητικά κατάλοιπα. Το rmsf υπολογίζεται από τον τύπο

$$\text{RMS}_i^{\text{fluct}} = (1/N_f \sum (\mathbf{r}_i^f - \mathbf{r}_i^{\text{ave}})^2)^{1/2}$$

όπου N ο συνολικός αριθμός των ατόμων, r_i^{mc} είναι η μέση θέση του κάθε ατόμου i και r_i^f η θέση του ατόμου i στο στιγμιότυπο f .

Η κινητικότητα του μορίου μπορεί να μελετηθεί από μια σειρά χαρτών αποστάσεων. Ο χάρτης αποστάσεων μεταξύ των Ca ατόμων μας δίνει πληροφορίες σχετικά με τη μέση απόσταση κάθε Ca ατόμου από κάθε άλλο για όλο το τροχιακό. Με τον υπολογισμό ενός τέτοιου χάρτη αποστάσεων υπολογίζεται και ένας χάρτης που περιέχει τα αντίστοιχα rmsd. Ακόμα ο χάρτης που περιλαμβάνει το rmsd του κάθε frame από κάθε άλλο (frame to frame) εκτός από τα συμπεράσματα για τη κινητικότητα του μορίου κάνει φανερή και μια πρώτη ομαδοποίηση των δομών.

Επιπλέον μπορούμε να μελετήσουμε τη συσχέτιση των κινήσεων μεταξύ των ατόμων στο μόριο. Για το λόγο αυτό αρχικά υπολογίζεται ένας πίνακας variance-covariance και ακολούθως με κανονικοποίηση των τιμών παίρνουμε έναν πίνακα cross-correlation για τα άτομα που επιλέξαμε (Ca, heavy atoms κλπ). Στον πίνακα αυτό όταν δυο άτομα κινούνται μαζί, δηλαδή έχουν θετική συσχέτιση, η τιμή είναι +1, όταν κινούνται αντίθετα, δηλαδή έχουν αρνητική συσχέτιση, η τιμή είναι -1 και όταν η κίνηση των ατόμων μεταξύ τους δεν έχουν καμία συσχέτιση η τιμή είναι 0. Ο πίνακας λοιπόν παίρνει τιμές από -1 έως +1 και γίνονται εμφανή δευτεροταγή στοιχεία της δομής.

Τέλος, μια άλλη σειρά αναλύσεων που μας δίνουν χρήσιμες πληροφορίες είναι η ανάλυση των κύριων συνιστωσών (Principal Component Analysis, PCA). Η κεντρική ιδέα είναι να μειώσουμε τις διαστάσεις των δεδομένων, που αποτελούνται από ένα μεγάλο αριθμό συσχετιζόμενων μεταβλητών, διατηρώντας όσο το δυνατόν τη διακύμανση που παρουσιάζει το σύνολο των δεδομένων (Jolliffe 2002). Αυτό επιτυγχάνεται μετατρέποντας τις σε ένα νέο σύνολο μεταβλητών, τις κύριες συνιστώσες, που είναι γραμμικοί συνδυασμοί των αρχικών μεταβλητών και δεν είναι συσχετιζόμενες. Οι κύριες συνιστώσες διατάσσονται έτσι ώστε στις πρώτες να οφείλονται οι περισσότερες διακυμάνσεις των αρχικών μεταβλητών. Έτσι δημιουργείται ένας δισδιάστατος πίνακας συσχέτισης μεταξύ των μεταβλητών από τον οποίο υπολογίζονται οι χαρακτηριστικές τιμές (eigenvalues) και τα αντίστοιχα χαρακτηριστικά ανύσματα (eigenvectors) ώστε να ικανοποιείται η εξίσωση:

$$\sum \alpha_k = \lambda \alpha_k$$

όπου Σ ο δισδιάστατος πίνακας μιας μεταβλητής, λ είναι η χαρακτηριστική τιμή και α_k το χαρακτηριστικό άνυσμα που του αντιστοιχεί και είναι μη μηδενικό. Από κάθε ζεύγος eigenvectors – eigenvalues μπορούν να χαρακτηριστούν οι επιμέρους κινήσεις της σύνθετης κίνησης του μορίου. Ένα από τα σημαντικότερα συμπεράσματα που μπορούμε να εξάγουμε από τη μελέτη των ζευγών eigenvectors – eigenvalues είναι εάν η διάρκεια της προσομοίωσης ήταν επαρκής ώστε να μας δώσει ικανοποιητικό δείγμα της κίνησης του μορίου. Αυτό μπορεί να γίνει μέσα από τα γραφήματα της ελεύθερης ενέργειας ΔG για όλους τους συνδυασμούς των επιλεγμένων eigenvectors. Η ανάλυση των κύριων συνιστωσών μπορεί να γίνει είτε με βάση το καρτεσιανό σύστημα, Cartesian-PCA, είτε χρησιμοποιώντας τις διέδρες γωνίες, dihedral-PCA (Altis A., 2007). Μια ακόμα ανάλυση που πραγματοποιείται με βάση τον αριθμό των κύριων συνιστωσών που έχουν επιλεγθεί είναι η ομαδοποίηση των τροχιακών (Yeung and Ruzzo, 2001).

Κεφάλαιο 4

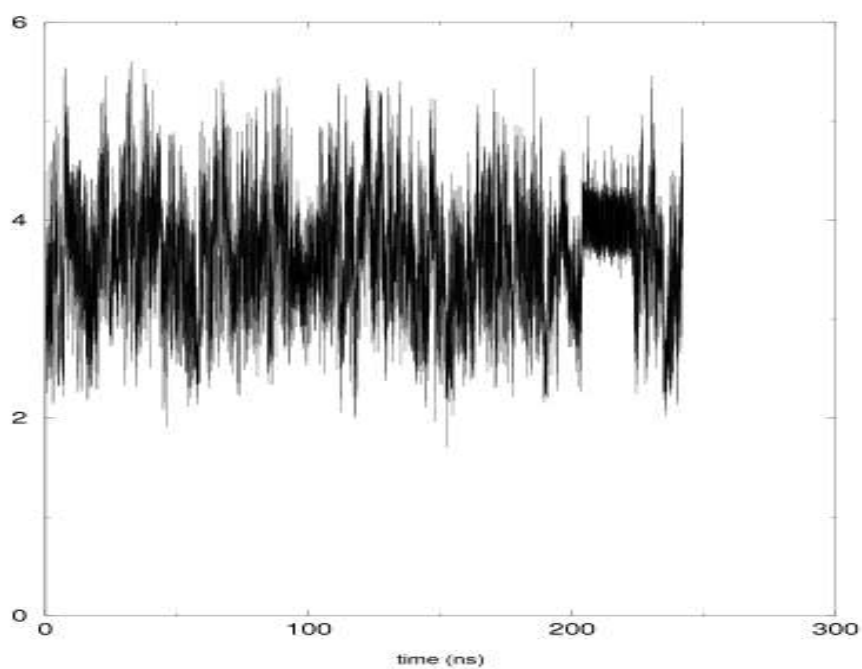
Ανάλυση Προσομοιώσεων

Ο στόχος αυτής της εργασίας ήταν να επιβεβαιωθεί μέσω προσομοιώσεων μοριακής δυναμικής η πειραματικά προσδιορισμένη δομή, από τους Gao *et al.* (2002), του εξαπεπτιδίου AAGDYY. Μετά όμως από μια σειρά αναλύσεων των αποτελεσμάτων που πήραμε από τις προσομοιώσεις, το πεπτίδιο φαίνεται να περνά από πολλές διαμορφώσεις που δεν συμπίπτουν όμως με αυτή που βρέθηκε πειραματικά. Οι αναλύσεις που ακολουθούν έγιναν με το πρόγραμμα ανάλυσης μοριακής δυναμικής carma (Glykos 2006).

4.1 RMSD από την αρχική δομή

Προκειμένου να διαπιστώσουμε αν το πεπτίδιο υιοθετεί την πειραματικά προσδιορισμένη δομή κατά την διάρκεια των προσομοιώσεων υπολογίστηκε το rmsd από την αρχική δομή. Για το λόγο αυτό τα τέσσερα trajectories των προσομοιώσεων ενώθηκαν σε ένα και σ' αυτά προστέθηκε ως αρχική δομή το πρώτο από τα 12 μοντέλα της πειραματικά προσδιορισμένης δομής.

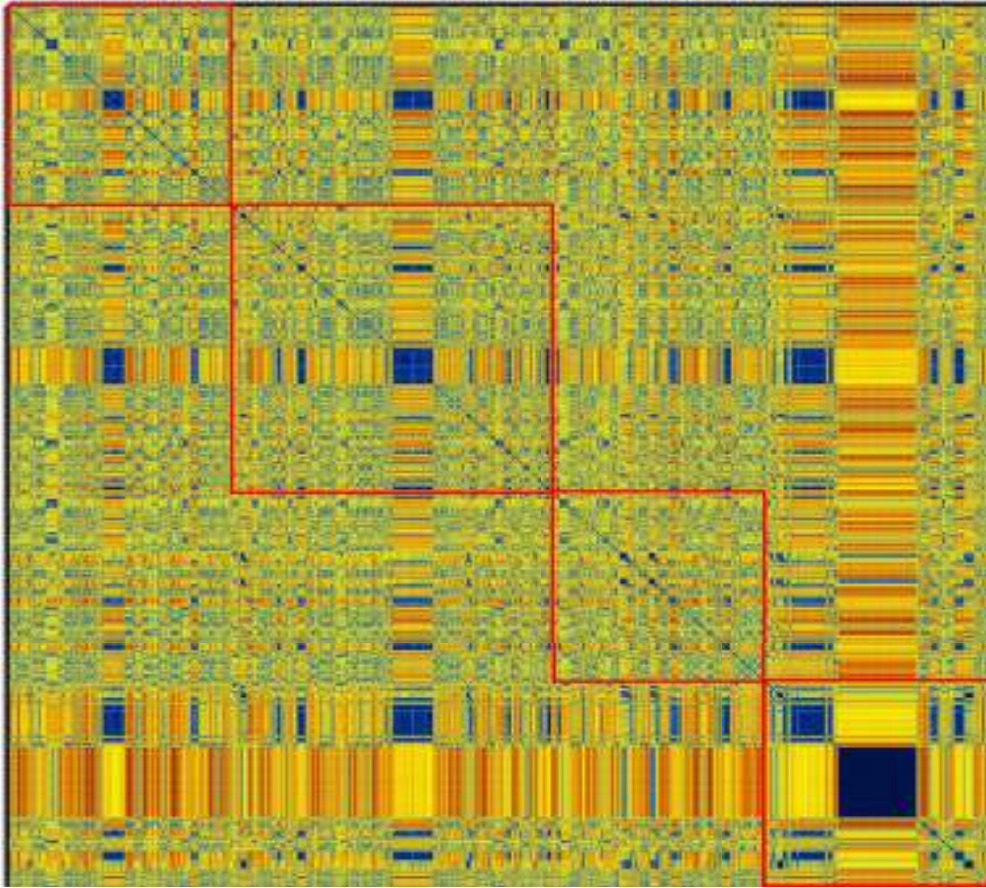
Στην **Εικόνα 4.1** παρουσιάζεται το rmsd από την αρχική δομή λαμβάνοντας υπόψη όλα τα άτομα του μορίου εκτός από τα πρωτόνια (heavy atoms). Παρατηρούμε πως το rmsd για το μεγαλύτερο χρονικό διάστημα έχει τιμή γύρω στο 3.8 angstroms με πολύ μεγάλη διακύμανση κοντά στα 2angstroms, κάποια στιγμή όμως σταθεροποιείται κοντά στα 4angstroms με μικρότερη διακύμανση στα 0.5angstroms. Κατά συνέπεια το πεπτίδιο φαίνεται να αποκλίνει αρκετά από την αρχική δομή και πως μάλλον υιοθετεί μια άλλη διακριτή δομή.



Εικόνα 4.1 RMSD από την αρχική δομή

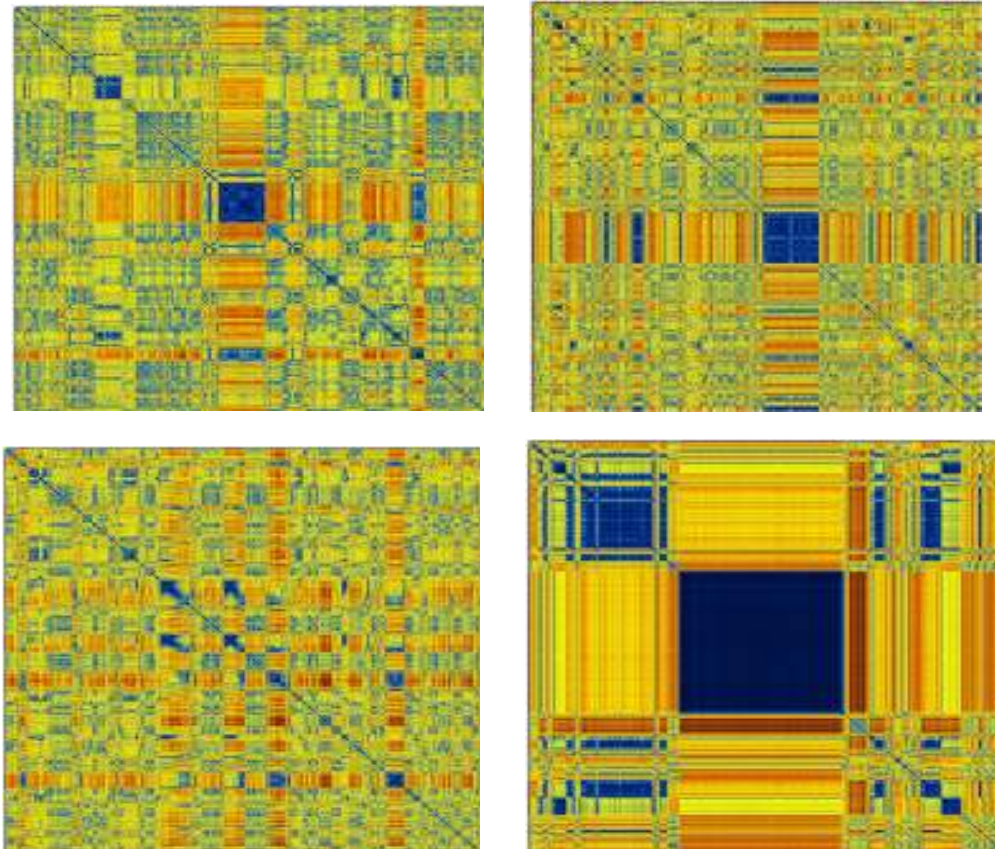
4.2 frame to frame rmsd

Ένας τρόπος για να μελετήσουμε την κινητικότητα του μορίου και τις δομές που υιοθετεί κατά τη διάρκεια της προσομοίωσης είναι ο υπολογισμός του rmsd του κάθε frame από το κάθε άλλο. Η δημιουργία ενός τέτοιου δισδιάστατου χάρτη που περιέχει όλες τις δομές από τις οποίες περνά το πεπτίδιο κατά τη διάρκεια των προσομοιώσεων με βήμα 100 στιγμιότυπα μας δίνει και μια πρώτη ομαδοποίηση των αποτελεσμάτων. Έτσι τα frame που το rmsd της δομής τους δεν έχουν μεγάλη απόκλιση από το rmsd της δομής άλλων frame θεωρείται ότι ανήκουν στο ίδιο cluster (ομάδα).



Εικόνα 4.2 frame to frame rmsd

Στον παραπάνω χάρτη **Εικόνα 4.2** απεικονίζονται στον οριζόντιο και κάθετο άξονα τα frames (604741 x 604741) ανά 100 και η διαβάθμιση του χρώματος από μπλε σε κίτρινο ακολουθεί την αύξηση του rmsd λαμβάνοντας υπόψη όλα τα άτομα του μορίου εκτός των πρωτονίων. Όπως παρατηρούμε το πεπτίδιο δεν είναι σταθερό και αλλάζει πολλές διαμορφώσεις. Παρατηρούμε όμως και τη δημιουργία κάποιων cluster που απεικονίζονται ως μικρά μπλε τετράγωνα. Ο χάρτης αυτός επαληθεύει τα αποτελέσματα που παίρνουμε από το διάγραμμα rmsd από την πειραματικά προσδιορισμένη δομή. Στην συνέχεια θα υποστηρίξουμε την ύπαρξη μιας περισσότερο διακριτής ομάδας που προκύπτει από τις προσομοιώσεις που όμως η αντιπροσωπευτική της δομή διαφέρει από αυτή των Gao *et al.*



Εικόνα 4.3 frame to frame rmsd των 4 προσομοιώσεων

Στην **Εικόνα 4.3** παρατίθενται οι χάρτες frame to frame rmsd για κάθε μια προσομοίωση ξεχωριστά. Τα τετράγωνα με κόκκινο περίγραμμα που φαίνονται στην **Εικόνα 4.2** αντιστοιχούν με τη σειρά στους παραπάνω χάρτες. Από τη σύγκριση των δυο εικόνων παρατηρούμε ότι στο ενωμένο trajectory υπάρχει ένα κύριο cluster που αντιστοιχεί στο κυρίαρχο cluster του trajectory της τέταρτης προσομοίωσης. Και έχει ακόμα ένα cluster που επαναλαμβάνεται τρεις φορές και αντιστοιχεί στο δεύτερο cluster της τέταρτης προσομοίωσης καθώς και στο κυρίαρχο cluster της δεύτερης και στο μεγαλύτερο της τρίτης προσομοίωσης.

4.3 Cluster analysis – Ομαδοποίηση των αποτελεσμάτων

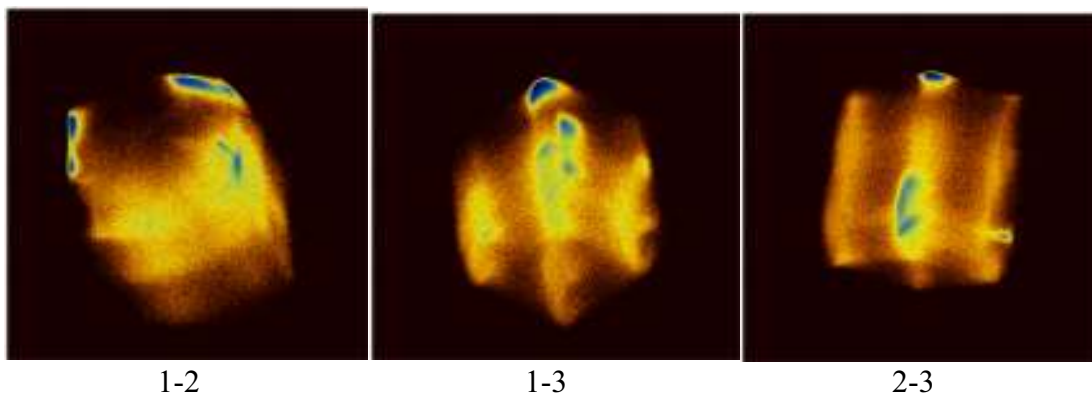
Η ομαδοποίηση των αποτελεσμάτων μπορεί να γίνει κάνοντας ανάλυση των κύριων συνιστωσών (Principal Component Analysis, PCA) των trajectories. Για να βρούμε τις ομάδες (clusters) που δημιουργούνται στο trajectory που αντιπροσωπεύει το σύνολο των προσομοιώσεων πραγματοποιήθηκε ανάλυση των κύριων συνιστωσών

χρησιμοποιώντας τις διεδρες γωνίες (dihedral PCA). Για το διαχωρισμό των clusters από το πρόγραμμα χρησιμοποιήθηκαν οι τρεις πρώτες κύριες συνιστώσες. Ο αριθμός των cluster που υπολογίστηκαν και τα frame που αντιστοιχούν στο καθένα είναι:

Σύνολο	604.741 frames
1 ^ο cluster	15.640 frames
2 ^ο cluster	23.077 frames
3 ^ο cluster	4.634 frames
4 ^ο cluster	3.367 frames
5 ^ο cluster	1.611 frames
6 ^ο cluster	475 frames

Από τα έξι αυτά cluster κυρίαρχο θεωρείται αυτό που αντιστοιχεί στα περισσότερα frame δηλαδή το δεύτερο cluster. Παρατηρούμε ότι το πεπτιδίο βρίσκεται σε κάποια διαμόρφωση καταλαμβάνοντας μόλις το 8% του συνόλου των frames.

Το πρόγραμμα ταυτόχρονα με την ομαδοποίηση των frame σε cluster υπολογίζει και την ελεύθερη ενέργεια και για τον υπολογισμό της βασίζεται στις τρεις πρώτες κυρίαρχες συνιστώσες. Από τα γραφήματα της ελεύθερης ενέργειας για τους ανά ζεύγη τρεις πρώτους eigenvectors μπορούμε να βγάλουμε συμπεράσματα σχετικά με την επάρκεια του χρόνου των προσομοιώσεων που θα μας δώσει ικανοποιητικό δείγμα της κίνησης του πεπτιδίου.



Εικόνα 4.4 Γράφημα της ελεύθερης ενέργειας με βάση τους τρεις πρώτους eigenvectors

Όπως παρατηρούμε στην **Εικόνα 4.4**, τα γραφήματα της ελεύθερης ενέργειας για τους τρεις πρώτους eigenvectors σε ζεύγη (1-2,1-3,2-3) είναι εμφανή τα cluster που δημιουργήθηκαν από το πρόγραμμα. Διαπιστώνουμε ότι υπάρχουν ελάχιστες τιμές της ελεύθερης ενέργειας που αντιστοιχούν σε συγκεκριμένες δομές. Μπορούμε

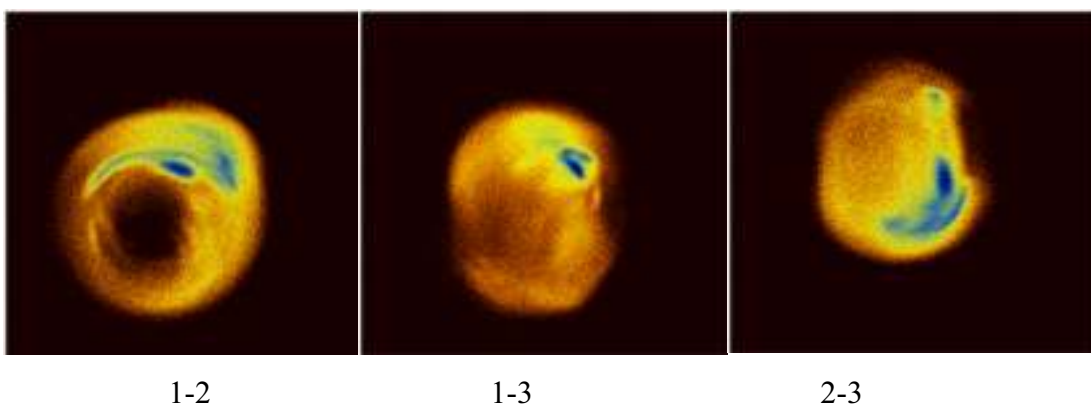
λοιπόν να ισχυριστούμε ότι το δείγμα που έχουμε από τις προσομοιώσεις είναι ικανοποιητικό ώστε να βγάλουμε ασφαλή συμπεράσματα (Mu Y. *et al.*, 2005).

Η ομαδοποίηση των αποτελεσμάτων μπορεί να γίνει και με ανάλυση των κύριων συνιστωσών με βάση το καρτεσιανό σύστημα (Cartesian-PCA, cPCA). Και σε αυτήν την περίπτωση για το διαχωρισμό των clusters από το πρόγραμμα χρησιμοποιήθηκαν οι τρεις πρώτες κύριες συνιστώσες. Ο αριθμός των cluster που υπολογίστηκαν και τα frame που αντιστοιχούν στο καθένα είναι:

Σύνολο	604.741 frames
1° cluster	32.908 frames
2° cluster	263 frames
3° cluster	140 frames
4° cluster	56 frames
5° cluster	23 frames
6° cluster	5 frames
7° cluster	5 frames

Τα cluster που προκύπτουν από την ανάλυση cPCA είναι επτά με μόνο ένα όμως από αυτά να αντιστοιχεί σε μεγάλο αριθμό frames που καταλαμβάνουν το 5,4% του συνόλου του τροχιακού. Παρατηρούμε ότι οι δυο αναλύσεις δεν δίνουν τα ίδια αποτελέσματα μιας και υπολογίστηκαν λαμβάνοντας υπόψη διαφορετικά στοιχεία.

Τα γραφήματα της ελεύθερης ενέργειας που υπολογίζονται από την ανάλυση cPCA για τους ανά ζεύγη τρεις πρώτους eigenvectors είναι τα ακόλουθα.

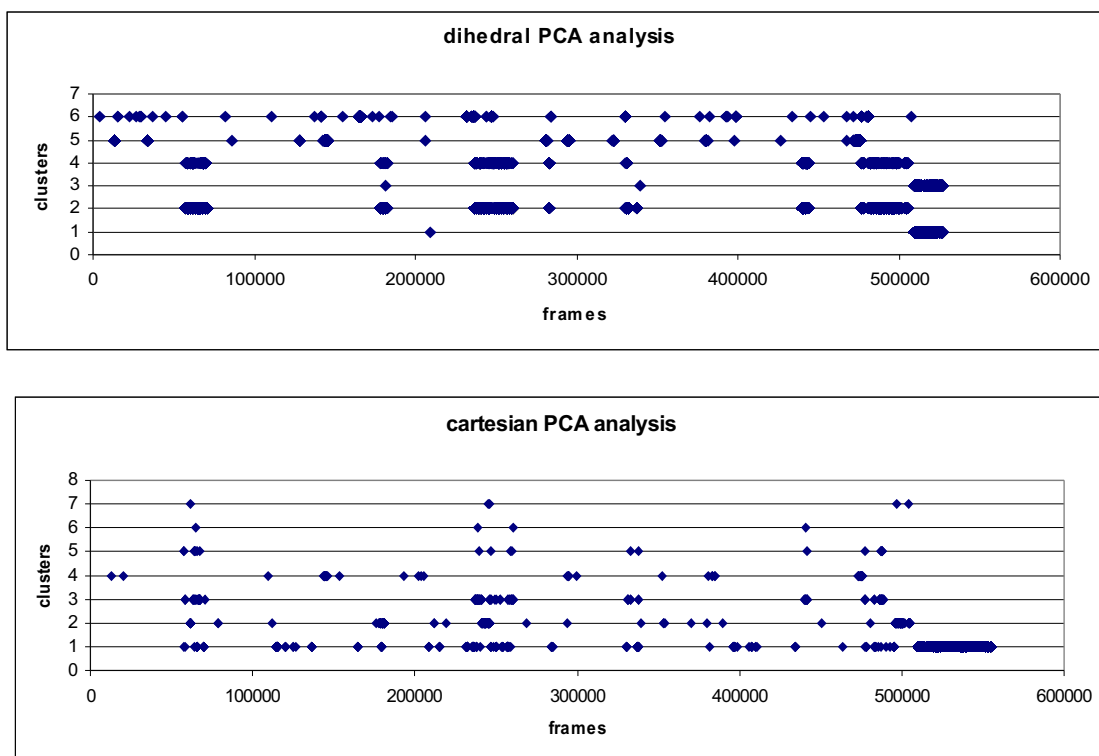


Εικόνα 4.5 Γράφημα της ελεύθερης ενέργειας με βάση τους τρεις πρώτους eigenvectors

Από τις **Εικόνες 4.4** και **4.5** παρατηρούμε πως τα γραφήματα της ελεύθερης ενέργειας για τις δυο αναλύσεις διαφέρουν. Ενώ στην περίπτωση του dPCA

εμφανίζονται αρκετά ελάχιστα στις ελεύθερης ενέργειας, στο γράφημα της ελεύθερης ενέργειας που προκύπτει από το cPCA έχουμε απεικόνιση ενός ελάχιστου.

Ένα διάγραμμα κατανομής των frames σε κάθε cluster που προκύπτει από τις δυο αναλύσεις μας δίνει μια καλύτερη εικόνα του κάθε cluster και μπορούμε να δούμε και τις διαφορές που εμφανίζουν οι δυο ομαδοποιήσεις.



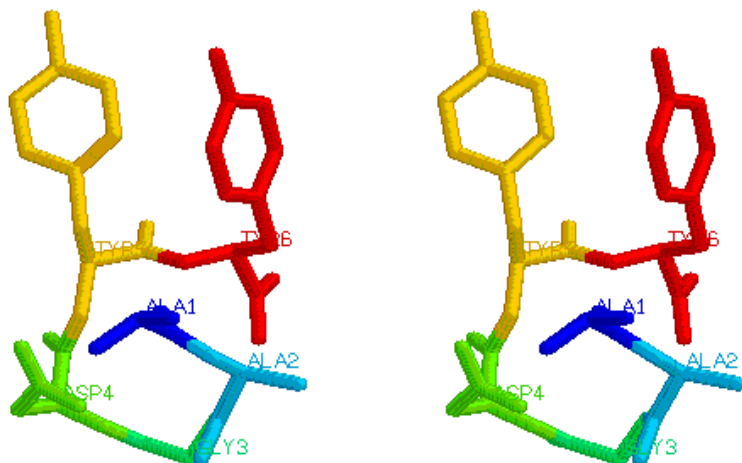
Εικόνα 4.6 Κατανομή των frames σε κάθε cluster

Από την **Εικόνα 4.6** παρατηρούμε ότι τα frames του κυρίαρχου cluster από το cPCA συμπεριλαμβάνουν και τα frames του κυρίαρχου cluster καθώς και ορισμένα από τα frames του δεύτερου cluster από το dPCA. Από την σύγκριση τώρα της κατανομής των frames σε κάθε cluster του dPCA και της **Εικόνα 4.2** διαπιστώνουμε ότι τα frames των clusters επαληθεύουν την ομαδοποίηση των frames όπως προκύπτει από το χάρτη frame to frame rmsd.

Γνωρίζοντας τα frame του τροχιακού (trajectory) που αντιστοιχούν σε κάθε cluster μπορούμε να μελετήσουμε το καθένα με περισσότερες λεπτομέρειες. Μελετάτε λοιπόν η σταθερότητα και η κινητικότητα κάθε cluster ξεχωριστά.

4.3.1 Cartesian-PCA ανάλυση cluster

Όπως είδαμε και στην προηγούμενη ενότητα από την ανάλυση των κύριων συνιστωσών με βάση το καρτεσιανό σύστημα παίρνουμε ένα κυρίαρχο cluster η αντιπροσωπευτική δομή (frame 26875) του οποίου είναι:



Εικόνα 4.7 stereo όψη του κυρίαρχου cluster από το ενωμένο trajectory του cPCA

Κινητικότητα του πεπτιδίου στο cluster

Μέσω του υπολογισμού της παραμέτρου rmsf μπορούμε μετέπειτα να απεικονίσουμε την κινητικότητα του πεπτιδίου. Δημιουργείται ένα αρχείο pdb που περιέχει στιγμιότυπα του cluster κάθε 1500 frame και η στήλη που αντιστοιχεί στον παράγοντα B αντικαθίσταται από τις τιμές του rmsf.

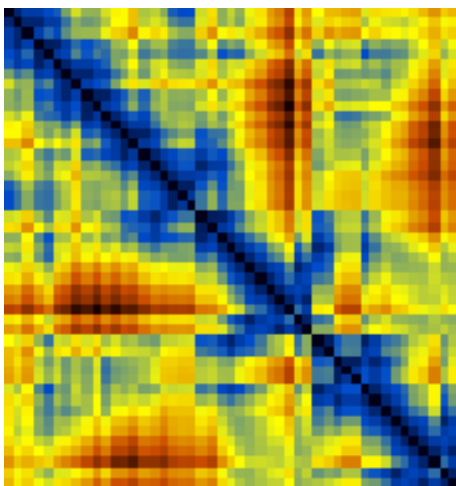


Εικόνα 4.8 Κινητικότητα πεπτιδίου με βάση το rmsf

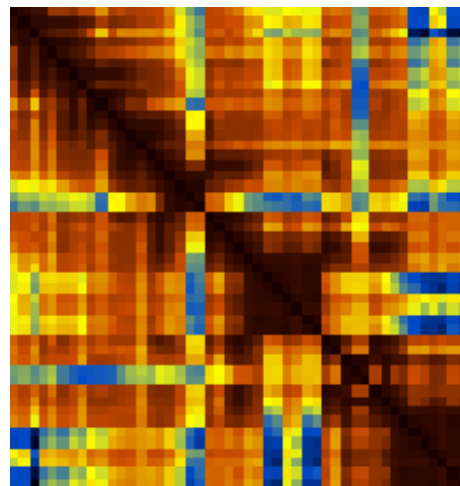
Στην **Εικόνα 4.8** φαίνεται η κινητικότητα των καταλοίπων του πεπτιδίου όπως απεικονίζεται μέσω του rasmol επιλέγοντας το χρωματισμό με βάση τη θερμοκρασία (colors > temperature). Με μπλε χρώμα απεικονίζονται τα λιγότερο κινητικά άτομα ενώ με κόκκινο χρώμα αυτά που εμφανίζουν μεγάλη διακύμανση. Παρατηρούμε ότι ο σκελετός του πεπτιδίου δεν εμφανίζει κινητικότητα σε αντίθεση με τις πλευρικές ομάδες που εμφανίζουν μεγαλύτερη κινητικότητα.

Χάρτης της μέσης απόστασης μεταξύ όλων των ατόμων εκτός των υδρογόνων (heavy atoms) και το αντίστοιχο rmsd:

Στις **Εικόνες 4.9** και **4.10** παρατηρούμε το χάρτη της μέσης απόστασης μεταξύ όλων των ατόμων εκτός των υδρογόνων και το rmsd που αντιστοιχεί σε αυτή για το κυρίαρχο cluster από το cPCA για το ενωμένο trajectory. Στην **Εικόνα 4.9** με μπλε χρώμα απεικονίζονται τα άτομα που εμφανίζουν χαμηλή κινητικότητα, με κόκκινο αυτά με μεγάλη κινητικότητα και με κίτρινο αυτά που βρίσκονται σε ενδιάμεση κατάσταση. Η διαβάθμιση των χρωμάτων στην **Εικόνα 4.10** είναι ακριβώς η αντίθετη. Από το χάρτη αποστάσεων συμπεραίνουμε ότι οι πλευρικές ομάδες των δυο τελευταίων καταλοίπων τυροσίνης εμφανίζουν μεγαλύτερες τιμές. Και από το χάρτη για τα αντίστοιχα rmsd παρατηρούμε ότι το πεπτίδιο είναι σταθερό στο μεγαλύτερο τμήμα του εκτός από τα δυο τελευταία κατάλοιπα που είναι πιο ευκίνητα.



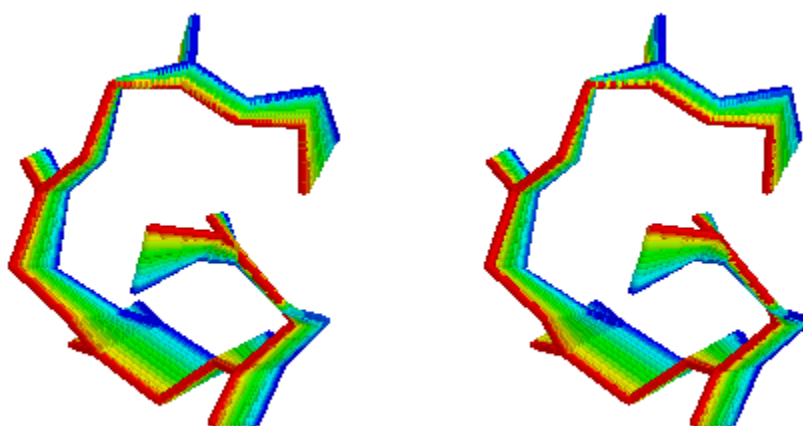
Εικόνα 4.9 non-hydrogen χάρτης αποστάσεων



Εικόνα 4.10 rmsd που αντιστοιχεί στο non-hydrogen χάρτη αποστάσεων

Κίνηση του πεπτιδίου με βάση τους eigenvectors-eigenvalues

Κάνοντας cPCA για το κυρίαρχο cluster μπορούμε να μελετήσουμε την κίνηση του πεπτιδίου στο cluster με βάση τους eigenvectors. Στην **Εικόνα 4.11** έχουμε την αναπαράσταση της κίνησης του σκελετού του πεπτιδίου με βάση τον πρώτο eigenvector. Οι πλευρικές ομάδες λείπουν καθώς οι πλευρικές ομάδες των καταλοίπων τυροσίνης εμφανίζουν μεγάλη κινητικότητα που δεν μπορεί να αναπαρασταθεί. Το εύρος της κίνησης φαίνεται από τη διαβάθμιση του χρώματος από μπλε σε κόκκινο.

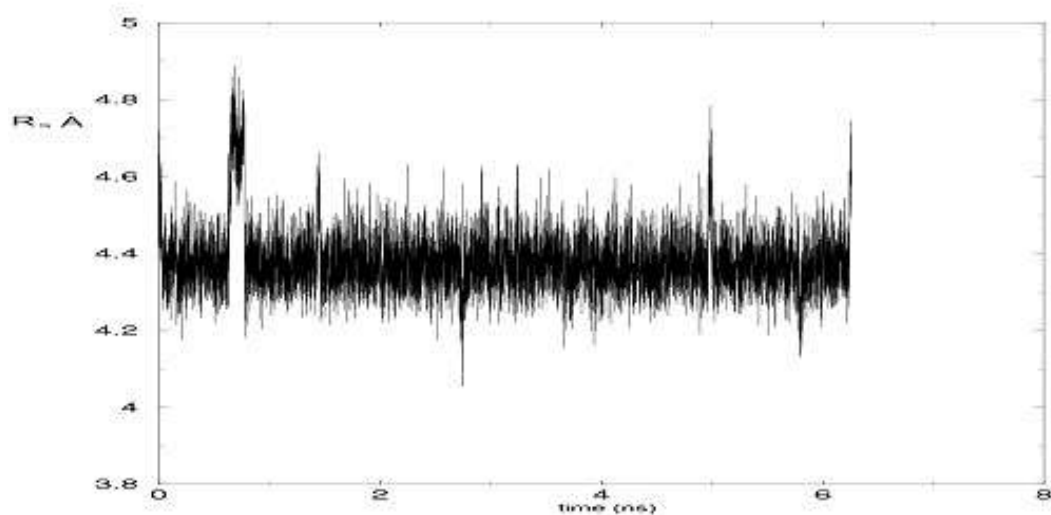


Εικόνα 4.11 Κίνηση του πεπτιδίου με βάση τον 1^ο eigenvector σε stereo όψη

4.3.2 dihedral-PCA ανάλυση cluster

1^ο Cluster:

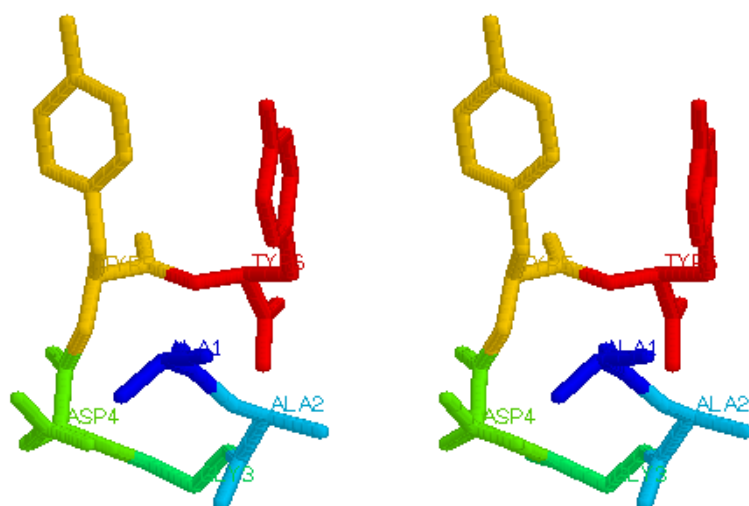
Μια παράμετρος που μας δείχνει πόσο συμπαγές είναι το μόριο είναι η εξέλιξη του radius of Gyration (R_G). Από την εξέλιξη του R_G μπορούμε να βγάλουμε συμπεράσματα σχετικά με τη σταθερότητα και τη διατήρηση της δομής.



Εικόνα 4.12 radius of Gyration για το πρώτο cluster

Στην **Εικόνα 4.12** δίνεται η εξέλιξη του R_G για το πρώτο cluster του ενωμένου trajectory. Σύμφωνα με το διάγραμμα το cluster φαίνεται να είναι αρκετά σταθερό καθώς η τιμή του R_G παρατηρείται να κυμαίνεται γύρω στα 4.4angstroms με διακύμανση 0.2angstroms και να αυξάνεται στο 4.7angstroms για ένα ελάχιστο χρονικό διάστημα 0.2ns.

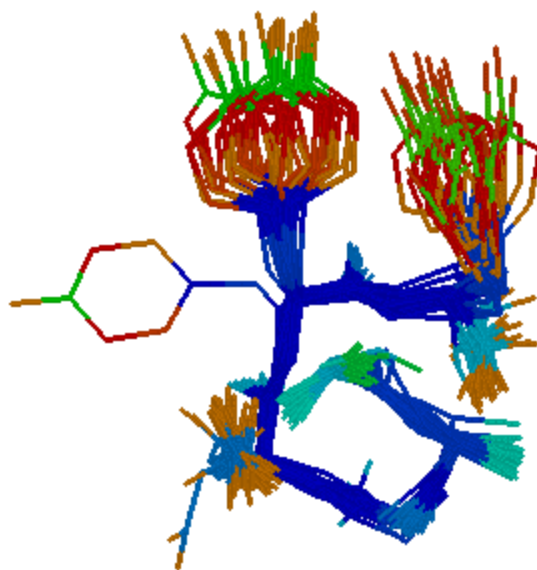
Η αντιπροσωπευτική δομή του cluster που αντιστοιχεί στην μικρότερη τιμή του rmsd από τη μέση δομή (frame 15555) είναι:



Εικόνα 4.13 stereo όψη του πρώτου cluster από το ενωμένο trajectory

Κινητικότητα του πεπτιδίου στο cluster

Για να μελετήσουμε την κινητικότητα των καταλοίπων του πεπτιδίου υπολογίστηκε η παράμετρος rmsf (root mean square fluctuation). Έτσι μπορούμε να δούμε τη διακύμανση ενός ατόμου σε σχέση με τη μέση δομή. Για την αναπαράσταση του πεπτιδίου σύμφωνα με το rmsf στο αρχείο pdb που περιλαμβάνει στιγμιότυπα του cluster με βήμα 500 frame αντικαθιστούμε την στήλη που αντιστοιχεί τις τιμές του B παράγοντα (B-factor) με αυτές που υπολογίστηκαν.

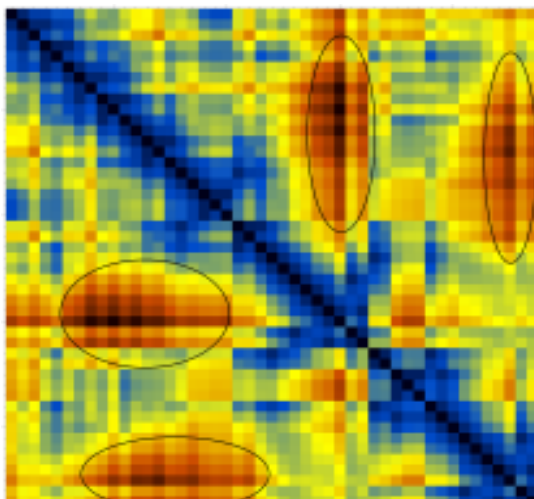


Εικόνα 4.14 Κινητικότητα πεπτιδίου με βάση το rmsf

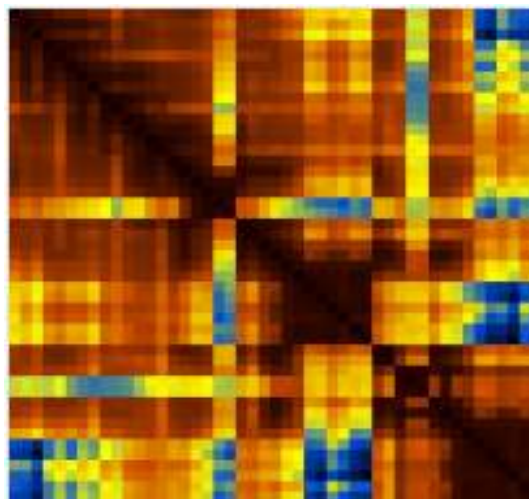
Στην **Εικόνα 4.14** φαίνεται η κινητικότητα των καταλοίπων του πεπτιδίου όπως απεικονίζεται μέσω του rasmol επιλέγοντας το χρωματισμό με βάση τη θερμοκρασία (colors > temperature). Με μπλε χρώμα απεικονίζονται τα λιγότερο κινητικά άτομα ενώ με κόκκινο χρώμα αυτά που εμφανίζουν μεγάλη διακύμανση. Παρατηρούμε ότι οι πλευρικές ομάδες των καταλοίπων τυροσίνης (Y) (πιο συγκεκριμένα ο αρωματικός δακτύλιος) εμφανίζουν μεγαλύτερη κινητικότητα, ενώ ο σκελετός παραμένει σταθερός.

Χάρτης της μέσης απόστασης μεταξύ όλων των ατόμων εκτός των υδρογόνων (heavy atoms) και το αντίστοιχο rmsd:

Στις **Εικόνες 4.15** και **4.16** παρατηρούμε το χάρτη της μέσης απόστασης μεταξύ όλων των ατόμων εκτός των υδρογόνων και το rmsd που αντιστοιχεί σε αυτή για το πρώτο cluster που βρέθηκε με την ανάλυση dPCA του ενωμένου trajectory. Στην **Εικόνα 4.15** με μπλε χρώμα απεικονίζονται τα άτομα που εμφανίζουν χαμηλή κινητικότητα, με κόκκινο αυτά με μεγάλη κινητικότητα και με κίτρινο αυτά που βρίσκονται σε ενδιάμεση κατάσταση. Η διαβάθμιση των χρωμάτων στην **Εικόνα 4.16** είναι ακριβώς η αντίθετη. Η περιοχές σε κύκλο που εμφανίζουν μεγαλύτερη κινητικότητα αντιστοιχούν στις τυροσίνες. Το πεπτίδιο εκτός από τις πλευρικές ομάδες των τυροσινών δεν εμφανίζει μεγάλη κινητικότητα.



Εικόνα 4.15 non-hydrogen χάρτης αποστάσεων



Εικόνα 4.16 rmsd που αντιστοιχεί στο non-hydrogen χάρτη αποστάσεων

Πίνακας της μέσης απόστασης μεταξύ των Ca ατόμων και το αντίστοιχο rmsd:

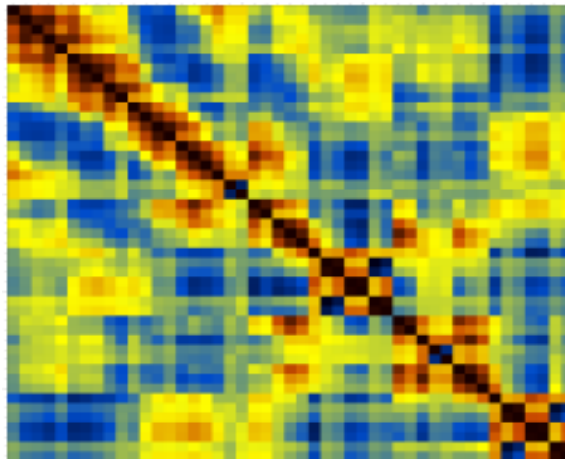
0.0	3.83456	6.37815	5.79442	5.31389	6.10599
0.05863	0.0	3.83355	5.64782	5.92269	4.70315
0.16678	0.05998	0.0	3.80387	5.95105	5.55791
0.20204	0.16053	0.05900	0.0	3.84553	6.01255
0.25601	0.23379	0.20413	0.05903	0.0	3.83582
0.28717	0.28656	0.25988	0.19779	0.05877	0.0

Η διαγώνιος του πίνακα αντιστοιχεί στην απόσταση που έχουν τα Ca άτομα με το εαυτό τους, πάνω από τη διαγώνιο έχουμε τη μέση απόσταση μεταξύ των Ca

ατόμων και κάτω το αντίστοιχο rmsd. Από ένα τέτοιο πίνακα μπορούμε να βγάλουμε συμπεράσματα για τη δομή του πεπτιδίου σύμφωνα με τη απόσταση των Ca ατόμων. Παρατηρούμε ότι η μεγαλύτερη τιμή της μέσης απόστασης είναι μεταξύ του πρώτου και τρίτου Ca ατόμου γεγονός που σημαίνει ότι τα άτομα αυτά είναι απομακρυσμένα στη δομή, ενώ η μικρότερη τιμή αντιστοιχεί μεταξύ του τέταρτου και τρίτου Ca.

Χάρτης Normalised cross-correlation

Από τον πίνακα συσχέτισης (Εικόνα 4.17) μπορούμε να βγάλουμε συμπεράσματα για το ποια άτομα του μορίου κινούνται μαζί ή αντίθετα. Ο χάρτης συσχέτισης παίρνει τιμές από -1, αρνητική συσχέτιση, (κίτρινο χρώμα) έως +1, θετική συσχέτιση, (μαύρο χρώμα) και 0 (μπλε χρώμα) όταν δεν υπάρχει συσχέτιση.

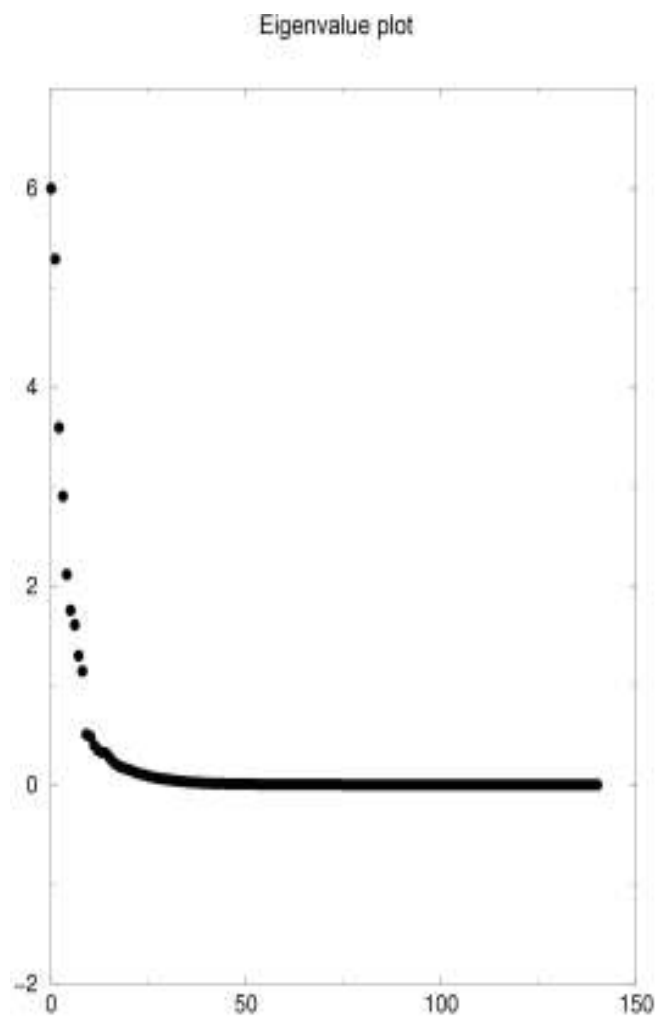


Εικόνα 4.17 normalised cross-correlation πίνακας για heavy atoms

Στην Εικόνα 4.17 η διαγώνιος αντιστοιχεί στην συσχέτιση που έχουν τα άτομα με τον εαυτό τους, πλήρης συσχέτιση, και γι αυτό το λόγο έχει μαύρο χρώμα. Ακόμα βλέπουμε με πορτοκαλί χρώμα τμήματα του πεπτιδίου που κινούνται μαζί. Τέτοια τμήματα είναι η πλευρική ομάδα του τελευταίου καταλοίπου τυροσίνης με τη γλυκίνη.

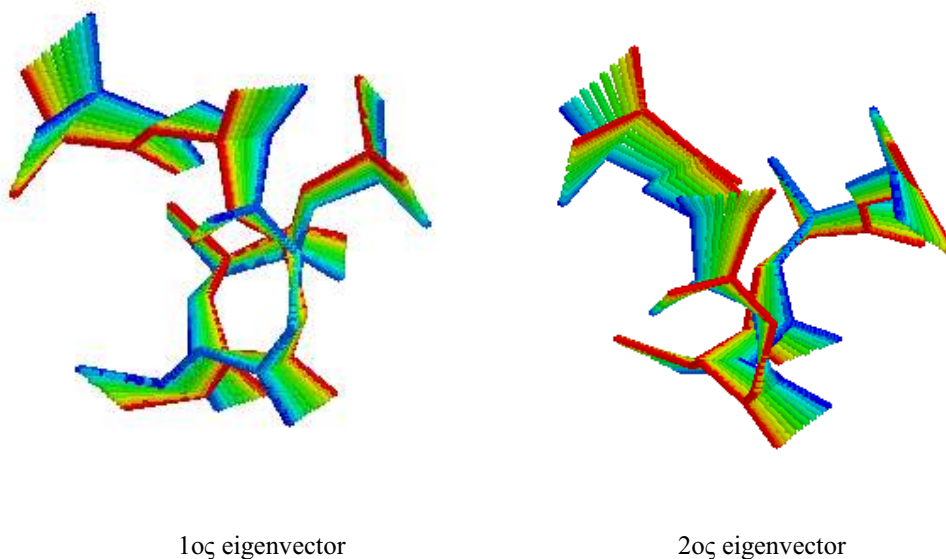
Κίνηση πεπτιδίου για τους δυο πρώτους eigenvectors από το cPCA

Προκειμένου να δούμε τις χαρακτηριστικές κινήσεις που πραγματοποιεί το πεπτίδιο πρέπει να υπολογίσουμε τους eigenvectors και eigenvalues, γι αυτό το λόγο πρώτα έγινε ανάλυση cPCA για το cluster.



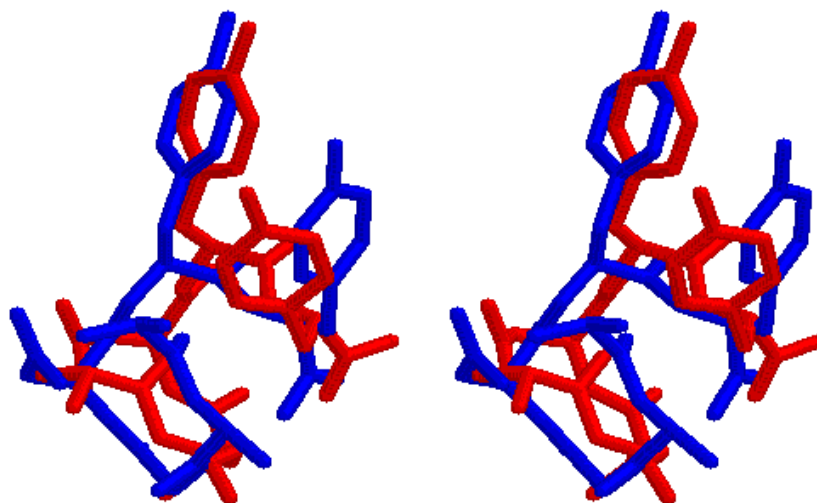
Εικόνα 4.18 Διάγραμμα απεικόνισης των eigenvalues

Στην **Εικόνα 4.18** βλέπουμε την γραφική απεικόνιση των eigenvalues. Οι πρώτοι 10 eigenvectors αντιπροσωπεύουν την πλειοψηφία των κύριων κινήσεων του πεπτιδίου. Για να παρακολουθήσουμε την κίνηση του πεπτιδίου επιλέχθηκαν οι δυο πρώτοι eigenvectors. Στην **Εικόνα4.19** που ακολουθεί φαίνεται η κίνηση του πεπτιδίου με βάση των 1^ο και 2^ο eigenvector με τη διαβάθμιση του χρώματος από μπλε σε κόκκινο να δείχνει το εύρος της κίνησης.



Εικόνα 4.19 Αναπαράσταση της κίνησης του πεπτιδίου με βάση τον 1^ο και 2^ο eigenvector

Σύγκριση των πρώτων clusters από τις αναλύσεις cPCA και dPCA

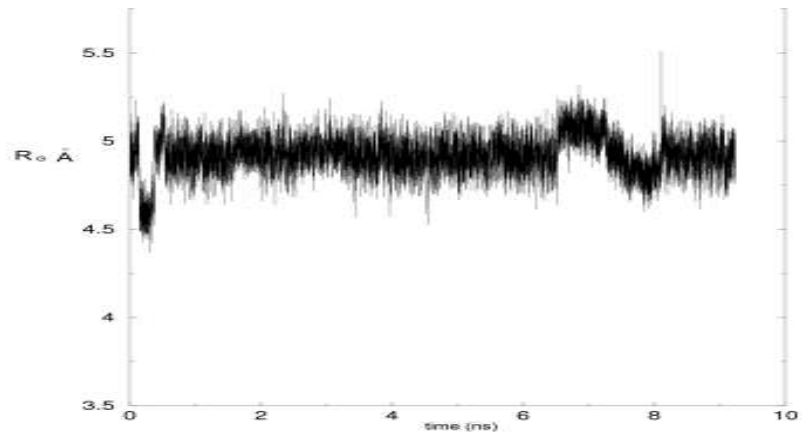


Εικόνα 4.20 stereo όψη της υπέρθεσης των δυο πρώτων cluster από cPCA (μπλε χρώμα) και dPCA (κόκκινο χρώμα)

Για να διαπιστώσουμε κατά πόσο μοιάζουν τα δυο πρώτα cluster από τις ομαδοποιήσεις των frames σύμφωνα με τις αναλύσεις cPCA και dPCA κάναμε υπέρθεση των δυο αντιπροσωπευτικών δομών του κάθε cluster. Αυτό που παρατηρούμε είναι ότι οι δυο δομές μοιάζουν αρκετά. Αυτό δικαιολογείται μιας και τα δυο cluster έχουν σε μεγάλο ποσοστό κοινά frames.

2o cluster

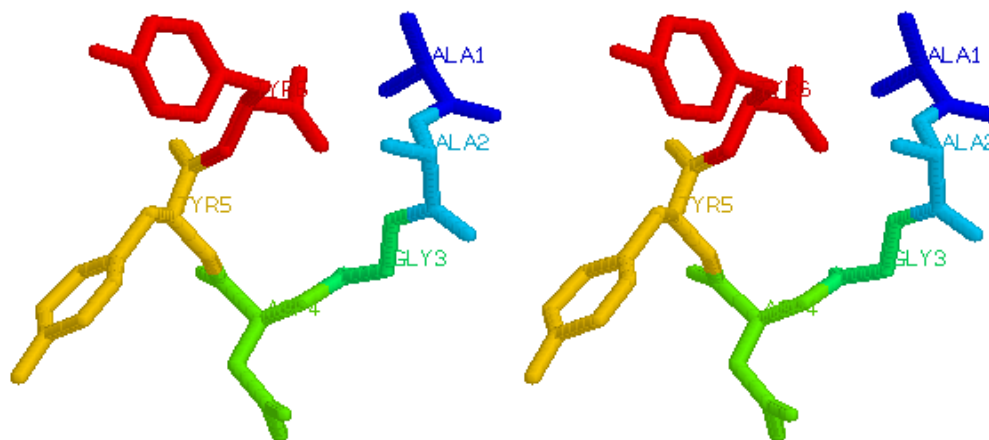
Η εξέλιξη του radius of Gyration για το δεύτερο και μεγαλύτερο cluster από την ανάλυση dPCA θα μας δείξει πόσο συμπαγές είναι το πεπτίδιο. Το δεύτερο είναι και το κυρίαρχο cluster, αποτελείται από 23077 frames και καταλαμβάνει το 3,82% του συνόλου του trajectory.



Εικόνα 4.21 radius of Gyration για το δεύτερο cluster από dPCA

Η **Εικόνα 4.21** δείχνει την εξέλιξη του radius of Gyration για το δεύτερο cluster από dPCA. Όπως παρατηρούμε το R_G εμφανίζει κάποιες διακυμάνσεις για μικρά διαστήματα παραμένοντας όμως για το μεγαλύτερο συνεχές χρονικό διάστημα γύρω από τη τιμή 5.0angstroms με διακύμανση 0.3angstroms.

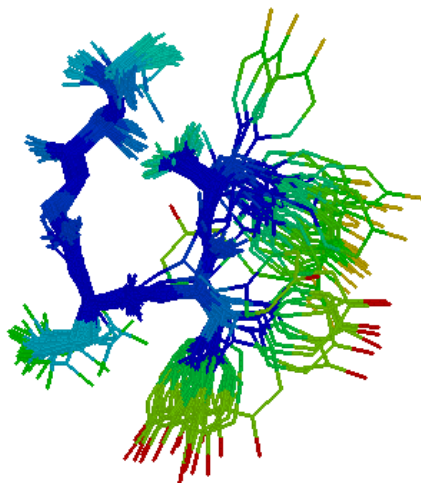
Η αντιπροσωπευτική δομή του δεύτερου cluster που αντιστοιχεί στην μικρότερη τιμή του rmsd από τη μέση δομή (frame 5584):



Εικόνα 4.22 stereo όψη της αντιπροσωπευτικής δομής του δεύτερου cluster από dPCA του ενωμένου trajectory

Κινητικότητα του πεπτιδίου στο cluster

Για να μελετήσουμε την κινητικότητα των καταλοίπων του πεπτιδίου υπολογίστηκε η παράμετρος rmsf (root mean square fluctuation). Έτσι μπορούμε να δούμε τη διακύμανση ενός ατόμου σε σχέση με τη μέση δομή. Για την αναπαράσταση του πεπτιδίου σύμφωνα με το rmsf στο αρχείο pdb αντικαθιστούμε την στήλη που αντιστοιχεί στις τιμές του B παράγοντα (B-factor) με αυτές που υπολογίστηκαν.



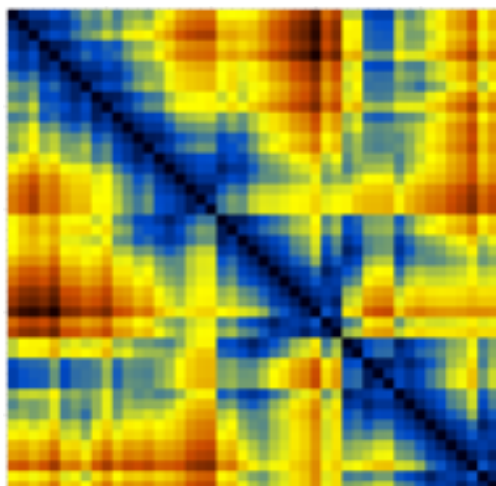
Εικόνα 4.23 Κινητικότητα πεπτιδίου με βάση το rmsf

Από την **Εικόνα 4.23**, όπου απεικονίζεται τη κινητικότητα του πεπτιδίου μέσω του προγράμματος rasmol επιλέγοντας το χρωματισμό με βάση τη θερμοκρασία, βγάζουμε το συμπέρασμα ότι το cluster δεν εμφανίζει μεγάλη κινητικότητα. Η διακύμανση του χρώματος από μπλε σε κόκκινο εκφράζει την αύξηση της κινητικότητας των ατόμων. Ο σκελετός του πεπτιδίου είναι σταθερός και η πλευρικές ομάδες εμφανίζουν μια κινητικότητα.

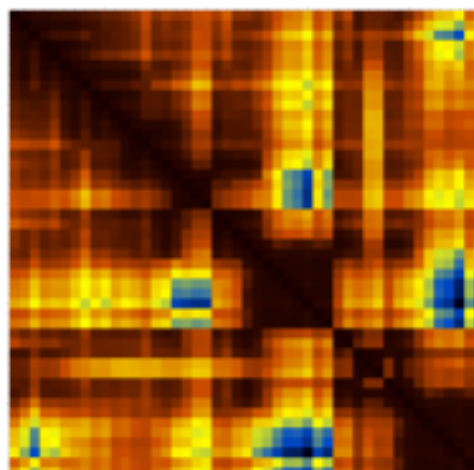
Χάρτης της μέσης απόστασης μεταξύ όλων των ατόμων εκτός των υδρογόνων (heavy atoms) και το αντίστοιχο rmsd:

Στις **Εικόνες 4.23** και **4.24** παρατηρούμε το χάρτη της μέσης απόστασης μεταξύ όλων των ατόμων εκτός των υδρογόνων και το rmsd που αντιστοιχεί σε αυτή για το δεύτερο cluster που βρέθηκε με την ανάλυση dPCA στο ενωμένο trajectory. Στην **Εικόνα 4.23** με μπλε χρώμα απεικονίζονται τα άτομα που εμφανίζουν χαμηλή κινητικότητα, με κόκκινο αυτά με μεγάλη κινητικότητα και με κίτρινο αυτά που

βρίσκονται σε ενδιάμεση κατάσταση. Η διαβάθμιση των χρωμάτων στην **Εικόνα 4.24** είναι ακριβώς η αντίθετη. Όπως παρατηρούμε από το χάρτη των rmsd το πεπτίδιο είναι στο σύνολο του είναι σταθερό με μόνο τις πλευρικές ομάδες των καταλοίπων τυροσίνης να εμφανίζουν μεγάλη κινητικότητα.



Εικόνα 4.24 non-hydrogen χάρτης αποστάσεων



Εικόνα 4.25 rmsd που αντιστοιχεί στο non-hydrogen χάρτη αποστάσεων

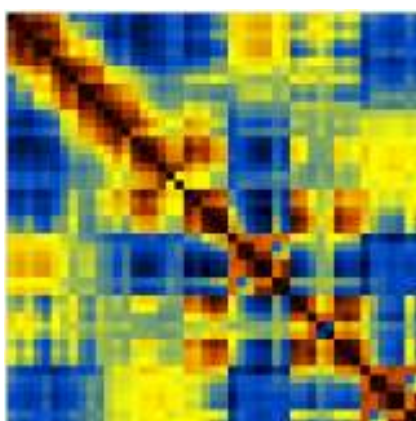
Πίνακας της μέσης απόστασης μεταξύ των Ca ατόμων και το αντίστοιχο rmsd:

0.0	3.83361	5.85507	8.30959	8.32752	5.41691
0.05830	0.0	3.81892	6.91681	7.49672	5.28332
0.25975	0.05991	0.0	3.80034	6.06183	5.83971
0.38397	0.24172	0.05986	0.0	3.83105	5.99179
0.27917	0.33123	0.20210	0.05841	0.0	3.81992
0.22452	0.26208	0.31899	0.21258	0.05957	0.0

Η διαγώνιος του πίνακα αντιστοιχεί στην απόσταση που έχουν τα Ca άτομα με το εαυτό τους, πάνω από τη διαγώνιο έχουμε τη μέση απόσταση μεταξύ των Ca ατόμων και κάτω το αντίστοιχο rmsd. Μας δίνεται έτσι η δυνατότητα να βγάλουμε συμπεράσματα για τη δομή του πεπτιδίου με βάση την απόσταση των Ca ατόμων. Παρατηρούμε ότι η μεγαλύτερη τιμή απόστασης είναι μεταξύ του πρώτου και πέμπτου Ca ατόμου που σημαίνει ότι είναι απομακρυσμένα, ενώ η μικρότερη τιμή αντιστοιχεί στα άτομα του τέταρτου και τρίτου Ca που σημαίνει ότι είναι κοντά στη δομή.

Χάρτης Normalised cross-correlation

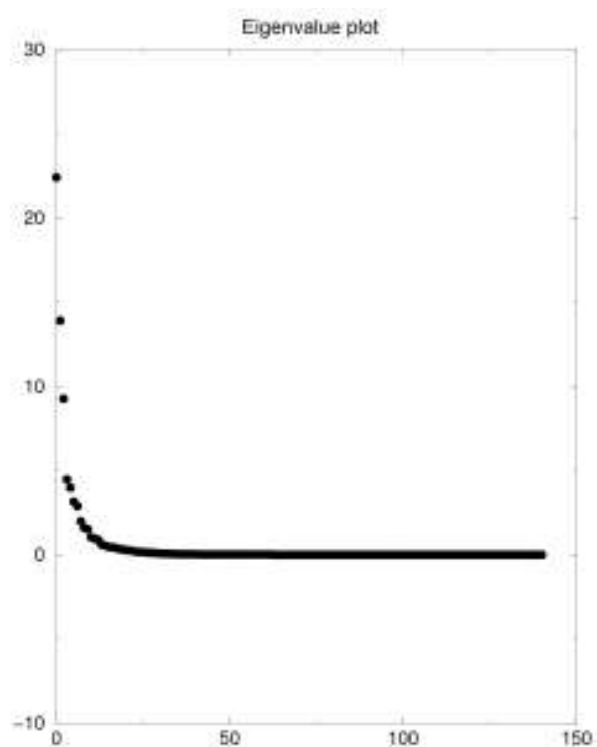
Για να μελετήσουμε τώρα ποιες ομάδες ατόμων κινούνται μαζί, ποιες αντίθετα και για ποια άτομα δεν υπάρχει συσχέτιση στην κίνηση τους δημιουργήσαμε το χάρτη συσχέτισης (**Εικόνα 4.26**). Ο χάρτης συσχέτισης παίρνει τιμές από -1 αρνητική συσχέτιση (κίτρινο χρώμα) έως +1 θετική συσχέτιση (μαύρο χρώμα) και 0 (μπλε χρώμα) όταν δεν υπάρχει συσχέτιση. Παρατηρούμε ότι η πλευρική ομάδα του πέμπτου καταλοίπου με το πρώτο κατάλοιπο έχουν αρνητική συσχέτιση, όπως και του έκτου καταλοίπου με το τέταρτο κατάλοιπο.



Εικόνα 4.25 normalised cross-correlation πίνακας για heavy atoms

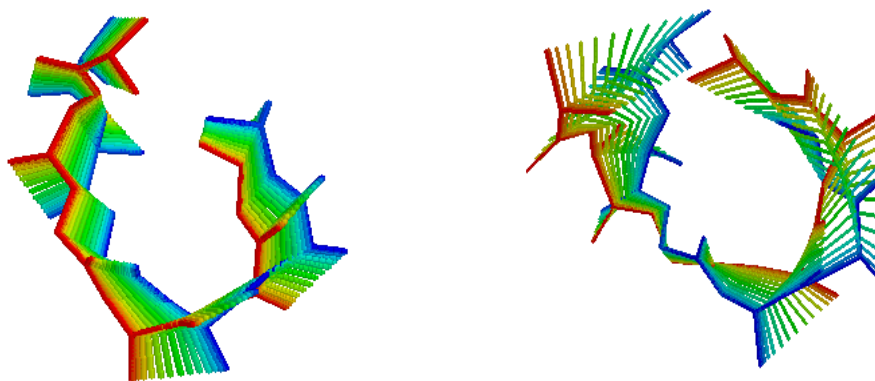
Κίνηση πεπτιδίου για τους δυο πρώτους eigenvectors από το cPCA

Για να παρακολουθήσουμε την κίνηση του πεπτιδίου σύμφωνα με τους eigenvectors κάναμε cPCA ανάλυση για το δεύτερο cluster. Στην **Εικόνα 4.26** βλέπουμε το διάγραμμα των eigenvalues. Παρατηρούμε ότι οι πρώτοι 5 eigenvectors αντιπροσωπεύουν την πλειοψηφία των κύριων κινήσεων του πεπτιδίου.



Εικόνα 4.26

Στην **Εικόνα 4.27** που ακολουθεί φαίνεται η κίνηση του πεπτιδίου με βάση των 1^ο και 2^ο eigenvector με τη διαβάθμιση του χρώματος από μπλε σε κόκκινο να δείχνει το εύρος της κίνησης του.



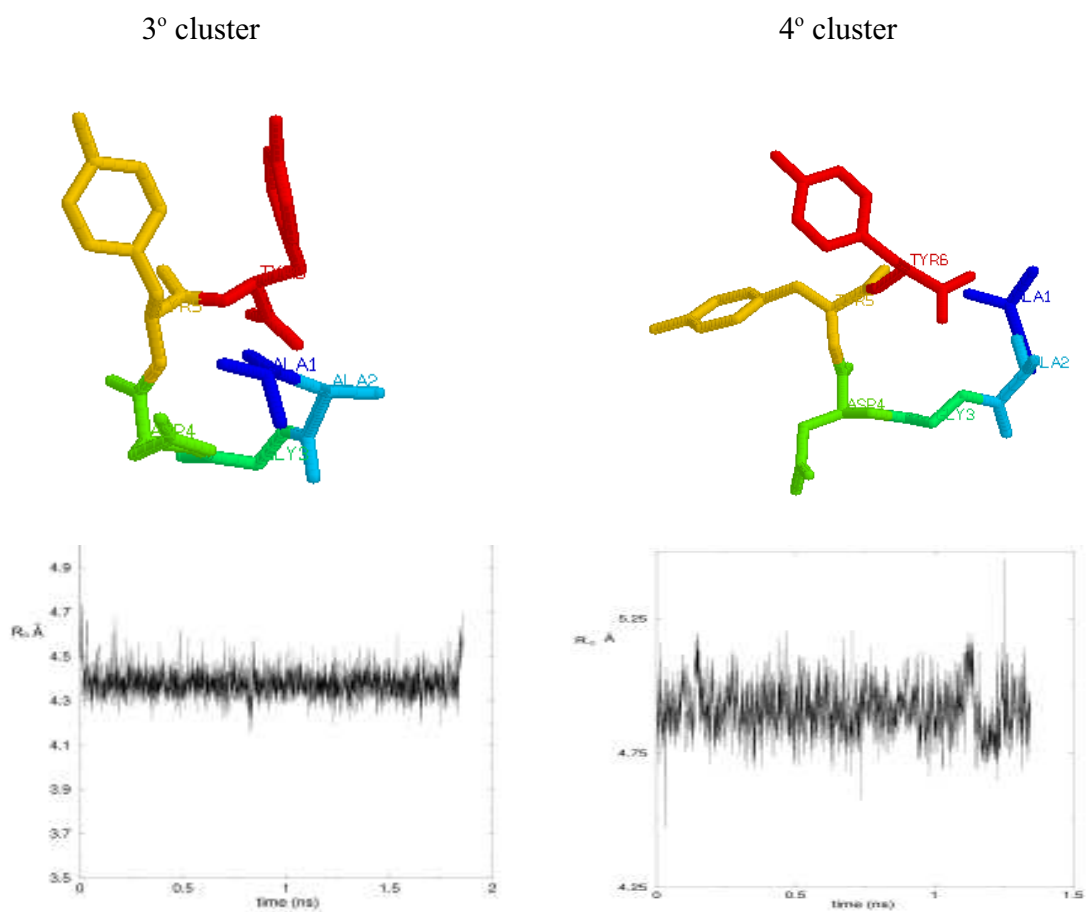
1st eigenvector

2nd eigenvector

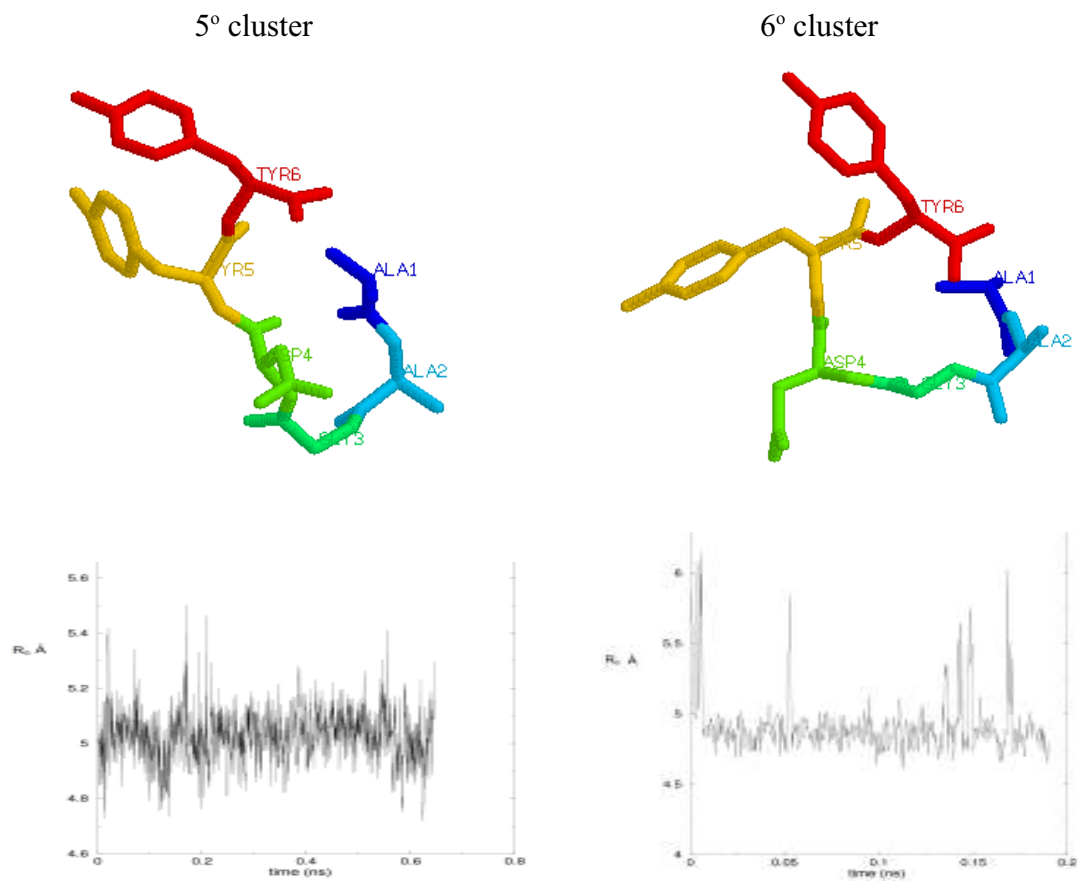
Εικόνα 4.27

Άλλα clusters

Έως τώρα ασχοληθήκαμε εκτενέστερα με την ανάλυση των κυριότερων cluster που σχηματίστηκα από την ομαδοποίηση των trajectories. Από την ανάλυση dPCA όμως προέκυψαν και άλλα cluster πέρα από τα δυο κυρίαρχα που συνολικά αποτελούνται από 10087 frames (1,66% του trajectory). Στην **Εικόνα 4.28** και στην **Εικόνα 4.29** φαίνονται οι αντιπροσωπευτικές δομές του κάθε cluster καθώς και το αντίστοιχο διάγραμμα της εξέλιξης του radius of gyration που δείχνει το κατά πόσο είναι σταθερά αυτά τα cluster.



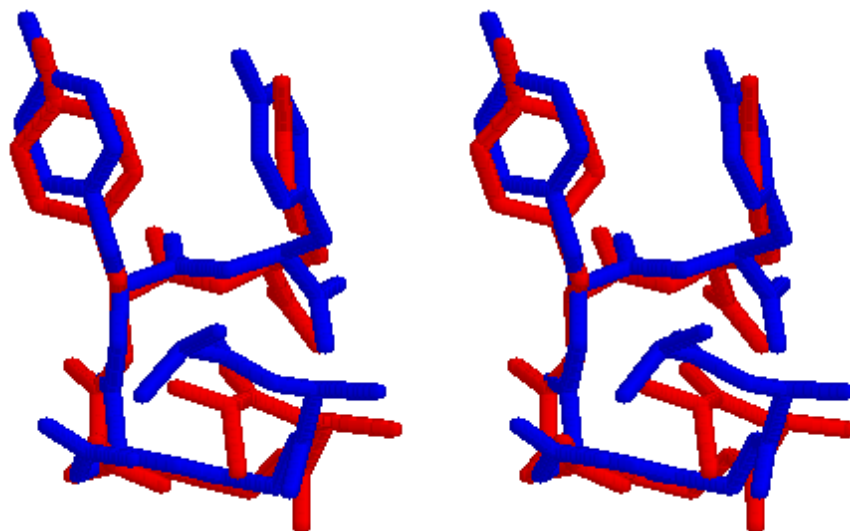
Εικόνα 4.28



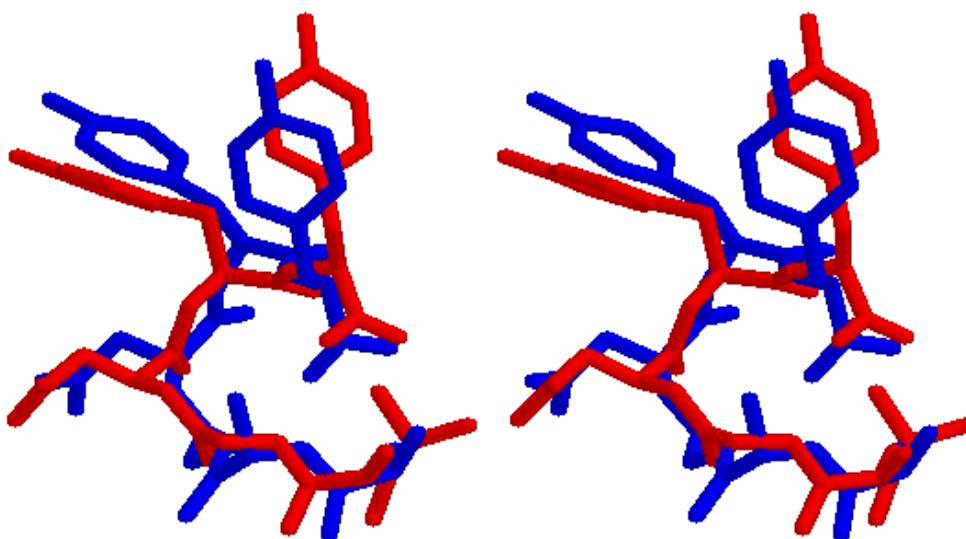
Εικόνα 4.29

4.3.2.1 Σύγκριση των clusters από την ανάλυση dPCA

Ρίχνοντας μια πιο προσεκτική ματιά στα cluster που προέκυψαν από την ανάλυση dPCA παρατηρούμε ότι οι δομές ορισμένων μοιάζουν μεταξύ τους. Τέτοιες περιπτώσεις είναι το 1^ο και 3^ο cluster καθώς και το 2^ο και 4^ο cluster. Στην **Εικόνα 4.30** και στην **Εικόνα 4.31** δίνεται η υπέρθεση των ζευγών αυτών όπου φαίνεται η ομοιότητα τους. Οι αντιπροσωπευτικές δομές του 1^ο και 3^ο cluster εμφανίζουν τα ίδια δομικά χαρακτηριστικά. Τα πέντε πρώτα κατάλοιπα συμμετέχουν στο σχηματισμό στροφής και δημιουργούνται δεσμοί υδρογόνου μεταξύ της NH ομάδας των τυροσινών (Tyr-5, Tyr-6) και της C=O ομάδας της Ala-1.



Εικόνα 4.30 Υπέρθεση 1^{ου} (μπλε χρώμα) και 3^{ου} (κόκκινο χρώμα) cluster

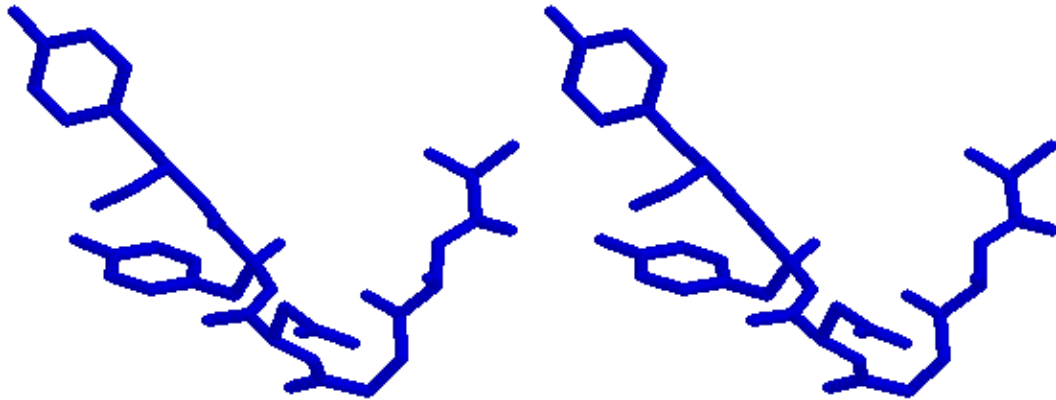


Εικόνα 4.31 Υπέρθεση 2^{ου} (μπλε χρώμα) και 4^{ου} (κόκκινο χρώμα) cluster

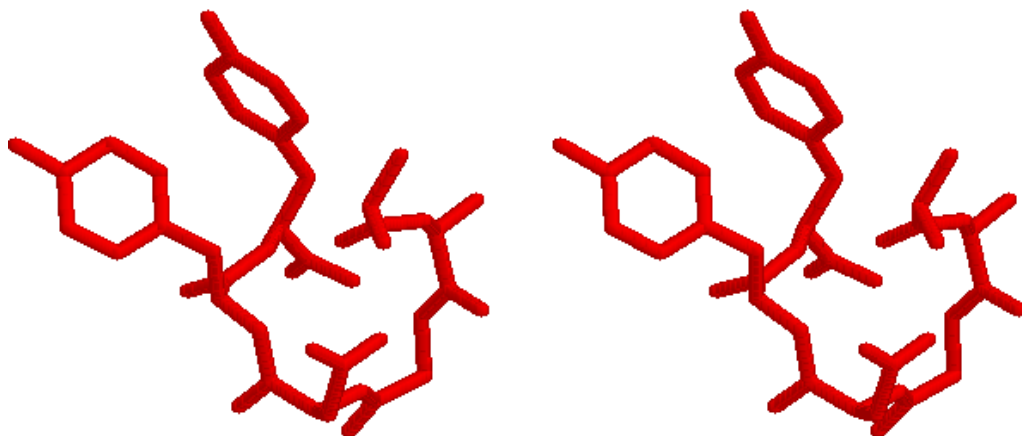
4.3.3 Σύγκριση των clusters

Στη προηγούμενη ενότητα ασχοληθήκαμε με την ανάλυση των clusters που προέκυψαν από την dPCA ανάλυση των trajectories με σκοπό να μελετήσουμε τις δομές τις οποίες υιοθετεί το πεπτιδίο κατά την διάρκεια της προσομοιώσεως. Θα συνεχίσουμε με τη σύγκριση των clusters με την πειραματικά προσδιορισμένη δομή για να διαπιστώσουμε αν οι προσομοιώσεις που πραγματοποιήθηκαν με τις παραμέτρους που επιλέχθηκαν μπορούν να αναπαράγουν αυτή τη δομή. Στην **Εικόνα 4.32** που ακολουθεί παρατίθενται σε stereo όψη η πειραματικά προσδιορισμένη δομή

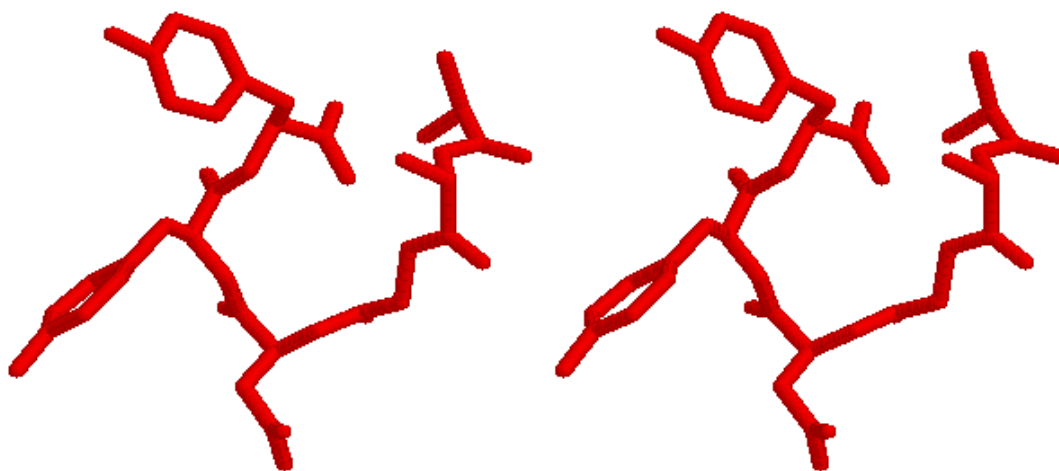
του πεπτιδίου και ακολούθως η αντιπροσωπευτική δομή του κάθε cluster. Οι δομές του 3^{ου} και 4^{ου} cluster λείπουν καθώς όπως δείξαμε στην προηγούμενη ενότητα έχουν ομοιότητες με τις δομές του 1^{ου} και 2^{ου} cluster αντίστοιχα.



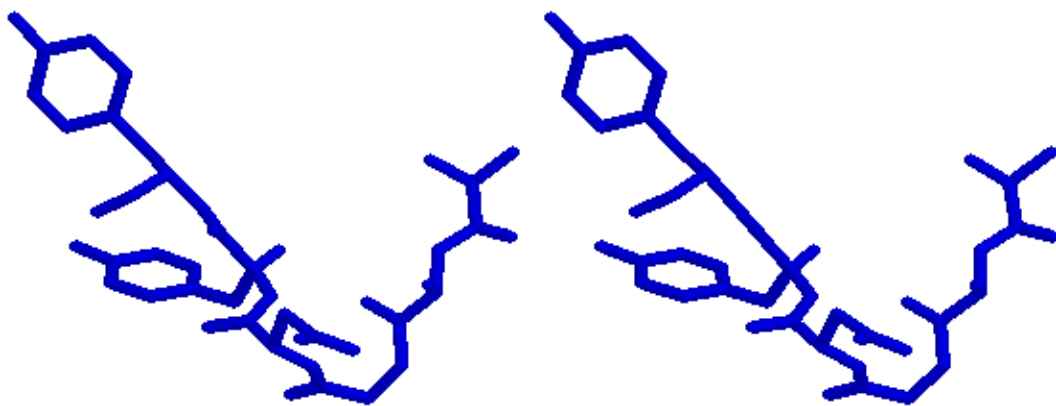
Πειραματική δομή



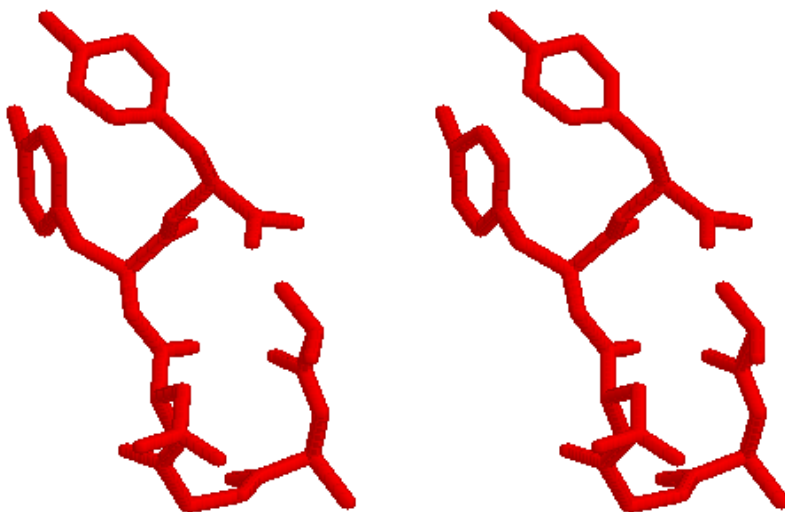
1^ο cluster



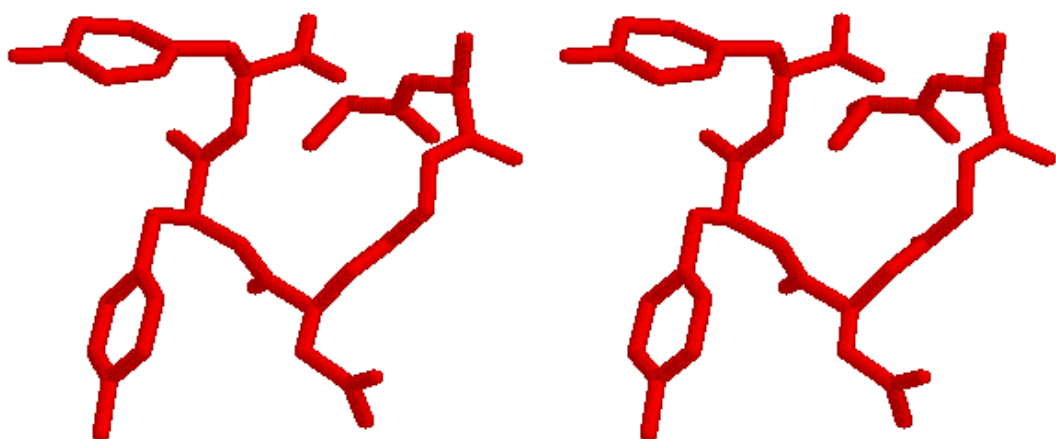
2^ο cluster



Πειραματική δομή



5° cluster



6° cluster

Εικόνα 4.32 stereo όψη των δομών του πεπτιδίου. Με μπλε απεικονίζεται η πειραματικά προσδιορισμένη δομή, ενώ με κόκκινο οι δομές των cluster που προέκυψαν από τις προσομοιώσεις. Οι δομές έγιναν με το πρόγραμμα rasmol.

Από την παράθεση των δομών στην **Εικόνα 4.32** παρατηρούμε ότι οι υπολογιστικά προσδιορισμένες δομές του πεπτιδίου διαφέρουν από την πειραματική δομή. Πιο συγκεκριμένα η δομή του πεπτιδίου που προσδιορίστηκε πειραματικά, έχει έναν δεσμό υδρογόνου που σχηματίζεται μεταξύ της C=O ομάδας της Ala-2 και της NH ομάδας της Tyr-5, που είναι χαρακτηριστικός στη δομής β-στροφής. Ενώ, όπως αναφέρθηκε στην προηγούμενη ενότητα, στο 1^ο cluster τα πέντε πρώτα κατάλοιπα συμμετέχουν στο σχηματισμό στροφής και δημιουργούνται δεσμοί υδρογόνου μεταξύ της NH ομάδας των τυροσινών (Tyr-5, Tyr-6) και της C=O ομάδας της Ala-1. Προκειμένου να διαπιστώσουμε τη δομή των υπόλοιπων clusters κάναμε ανάλυση της δομής ανά κατάλοιπο μέσω του προγράμματος VMD (Humphrey *et al.*, 1996). Διαπιστώσαμε ότι σ' αυτά το πεπτίδιο έχει τη δομή σπειράματος. Καμία λοιπόν δομή δεν συμπίπτει με αυτή που δόθηκε πειραματικά στο πεπτίδιο.

Επίλογος

Σκοπός αυτής της εργασίας ήταν να διαπιστώσουμε αν οι προσομοιώσεις μοριακής δυναμικής χρησιμοποιώντας ως force field το CHARMM27 και η επιλογή των συνθηκών του συστήματος μπορούν να αναπαράγουν την πειραματικά προσδιορισμένη δομή του πεπτιδίου AAGDYY. Τα αποτελέσματα όμως που πήραμε από τις προσομοιώσεις δεν είναι συμβατά με τα πειραματικά δεδομένα. Το πεπτίδιο καθώς εξελίσσεται η προσομοίωση είναι ασταθές καθώς περνά από διαφορετικές δομές χωρίς όμως να σταθεροποιείται σε κάποια από αυτές. Συνολικά το πεπτίδιο παραμένει σε κάποια δευτεροταγή δομή σε ποσοστό 8% του συνόλου του trajectory.

Στη σύγκριση των αποτελεσμάτων πρέπει να λάβουμε υπόψη μας ότι τα πειραματικά δεδομένα που έχουμε στη διάθεση μας προέκυψαν από πειράματα που έγιναν κάτω από διαφορετικές συνθήκες σε σχέση με τις προσομοιώσεις. Το πεπτίδιο μελετήθηκε πειραματικά σε διάλυμα DMSO-d₆ σε θερμοκρασία 20°C (293K) ενώ οι προσομοιώσεις έγιναν σε υδατικό διάλυμα στους 320K. Αυτός μπορεί να είναι ένας λόγος για τον οποίο οι προσομοιώσεις δεν μπόρεσαν να προβλέψουν τα πειραματικά δεδομένα. Έτσι μια προσομοίωση υπό συνθήκες όμοιες με αυτές του πειράματος μπορεί να δώσει αποτελέσματα συγκρίσιμα μ' αυτά του πειράματος.

Παρ' όλα αυτά, η συνολική διάρκεια της προσομοιώσεις που είναι 0.241μs μας επιτρέπει να προσδιορίσουμε τη δομή του πεπτιδίου με βάση τις παραμέτρους που επιλέχθηκαν. Το πρόβλημα της επάρκειας του δείγματος, που παίρνουμε από τις προσομοιώσεις, ώστε να εξάγουμε ασφαλή συμπεράσματα στην προκειμένη περίπτωση δεν υφίσταται. Μια προσομοίωση διάρκειας της τάξης των 10⁻⁷s είναι δυνατό να προβλέψει στοιχεία δευτεροταγούς δομής για ένα πεπτίδιο. Μπορούμε λοιπόν να πούμε ότι η δομή του πεπτιδίου που προέκυψε είναι αυτή που αντιστοιχεί στο δυναμικό πεδίο και στις συνθήκες κάτω από τις οποίες έγιναν οι προσομοιώσεις.

BIBΛΙΟΓΡΑΦΙΑ

1. Altis A., Nguyen P. H., Hegger R. and Stock G. (2007). Dihedral angle principal component analysis of molecular dynamics simulations. *J. Chem. Phys.* **126**, 244111(10).
2. Bernstein H. J. (1999). Rasmol 2.7.1 Molecular Graphics Visualisation Tool.
3. Berzofsky J., Oh S. and Terabe M. (2005). Peptide Vaccines Against Cancer. *Cancer Treatment and Research* **123**, 115-136.
4. Brogden K. A. (2005). Antimicrobial peptides: pore formers or metaboloc inhibitors in bacteria? *Nature Reviews Microbiology* **3**, 238-250.
5. Brown K. L. and Hancock R. E. (2006). Cationic host defense (antimicrobial) peptides. *Current Opinion in Immunology* **18**, 24-30.
6. Chang X., Wang Y. and Hu L. (2006). An implicit Taylor series numerical calculation method for power system transient simulation. *The Guide to Computing Literature*, 82-85.
7. Darden T., York D. and Pedersen L. (1993). An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092.
8. Daura X., Gademann K, Jaun B, Seebach D, van Gunsteren W.F and Mark A.E (1999). Peptide folding: when simulation meets experiment. *Angew. Chem.* **111**, 249-253.
9. Deserno M. and Holm C. (1998). How to mesh up Ewald sums (I): A theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.* **109**, 7678-7693.
10. Dill A K, Ozkan S B, Weikl T R, Chodera J D and Voelz V A (2007). The protein folding problem: when will it be solved? *Current Opinion in Structural Biology* **17**, 342–346.
11. Doran J. L., Leskiw B. K., Aippersbach S. and Jensen S. E. (1990). Isolation and characterization of a beta-lactamase-inhibitory protein from *Streptomyces clavuligerus* and cloning and analysis of the corresponding gene. *J Bacteriol.* **172**, 4909-4918
12. Ercolessi , F. (1997) A molecular dynamics primer. International School for Advanced Studies (Sissa-isas). Italy.

13. Essmann U., Perera L. and Berkowitz M. L. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577-8593.
14. Fricker, L. D., Lim, J., Pan, H. and Che, F.-Y. (2006), Peptidomics: Identification and quantification of endogenous peptides in neuroendocrine tissues. *Mass Spectrometry Reviews* **25**, 327-344.
15. Gao F., Wang Y., Qiu Y., Li Y., Sha Y., Lai L., and Wu H. (2002). β -turn formation by a six-residue linear peptide in solution. *Peptide* **60**, 75-80.
16. Garemyr, R. & Elofsson A. (1999). Study of the Electrostatics Treatment in molecular dynamics simulations. *PROTEINS: Structure, function and genetics* **37**, 417-428.
17. Glykos, N. M. (2006) Carma: A molecular dynamics analysis program, *J. Comput. Chem.* (in press).
18. Guvench O., MacKerell Jr. A. D. (2008). Comparison of Protein Force Fields for Molecular Dynamics Simulations. *Methods in Molecular Biology* **443**, 63-88.
19. Harder J., Bartels J., Christophers E. and Schroder J. M. (2001). Isolation and Characterization of Human b-Defensin-3, a Novel Human Inducible Peptide Antibiotic. *The Journal of Biological Chemistry* **276**, 5707–5713.
20. Humphrey W., Dalke A. and Schulten K. (1996). VMD Visual Molecular Dynamics. *J. Mol. Graphics* 14.1, 33-38.
21. Hunter R.G. and Kuhar M.J. (2003). CART Peptides as Targets for CNS Drug Development. *Current Drug Targets-CNS & Neurological Disorders* **2**, 201-205.
22. Jolliffe T. (2002). Principal component analysis (Springer, New York)
23. Karplus M. and McCammon, J. A. (2002). Molecular dynamics simulation of biomolecules. *Nat. Struct. Biol.* **9**, 646-652.
24. Karplus M. (2003). Molecular Dynamics of biological macromolecules: a brief history and perspective. *Biopolymers* **68**, 350-358.
25. Kleywegt G. J. (1995). [O/X-PLOR DICTIONARIES] Dictionaries for Heteros, ESF/CCP4. *Newsletter* **31**, 45-50.
26. Kuszewski J, Gronenborn AM and Clore GM (1999). Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. *J Am Chem Soc* 121, 2337–2338.

27. MacKerell, Jr., A.D., Bashford, D., Bellott, M., Dunbrack Jr., R., Evanseck, J., Field, M., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F., Mattos, C., Michnick, S., Ngo, T., Nguyen, D., Prodhom, B., Reiher, III, W., Roux, B. and Schlenkrich, M., Smith, J., Stote, R. and Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., & Karplus, M. (1998). All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **102**, 3586-3616.
28. Marcelino C. Anna Maria and Lila M. Gierasch (2008). Roles of β -Turns in Protein Folding: From Peptide Models to Protein Engineering. *Biopolymers* **89**, 380-391.
29. McCammon, J. A., Gelin, B. R. & Karplus, M. (1977). Dynamics of folded proteins. *Nature*, **267**, 585-590.
30. Motta A., Reches M., Pappalardo L., Andreotti G. and Gazit E. (2005). The Preferred Conformation of the Tripeptide Ala-Phe-Ala in Water is an Inverse γ -Turn: Implications for Protein Folding and Drug Design. *Biochemistry* **44**, 14170-14178.
31. Mu Y., Nguyen P. H. and Stock G. (2005). Energy Landscape of Small Peptide Revealed by Dihedral Angle Principal Component Analysis. *PROTEINS: Structure, Function, and Bioinformatics* **58**, 45-52.
32. Natalie C. J. Strynadka, Susan E. Jensen, Kathy Johns, Helen Blanchard, Malcolm Page, Andre Matagne, Jean-Marie Frere & Michael N. G. James (1994). Structural and kinetic characterization of a β -lactamase-inhibitor protein. *Nature* **368**, 657-660.
33. Paschek D. (2004). Temperature Dependence of the Hydrophobic Hydration and Interaction of Simple Solutes: An Examination of Five Popular Water Models. *J. Chem. Phys.* **120**, 6674-6690.
34. Patel, S.; MacKerell, Jr. AD; Brooks III, Charles L (2004). "CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model". *J Comput Chem* **25** (12): 1504–1514.
35. Petrosino J., Rudgers G., Hiram Gilbert H., and Palzkill T (1999). Contributions of Aspartate 49 and Phenylalanine 142 Residues of a Tight Binding Inhibitory Protein of β -Lactamases. *The journal of biological chemistry* **274**, 2394-2400.

36. Phillips J. C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C., Skeel R. D., Kale L., and Schulten K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **26**,1781-1802.
37. Ponder J. W. and Case D. A. (2003). Force Fields for Protein Simulations. *Advances Protein Chemistry* **66**, 27-86.
38. Rowlands D.J. (1992). How can peptide vaccines work? *FEMS Microbiology Letters* **100**, 479-481.
39. Rudgers G. W., Huang W., and Palzkill T. (2001). Binding Properties of a Peptide Derived from β -Lactamase Inhibitory Protein. *Antimicrobial Agents and Chemotherapy* **45**, 3279–3286.
40. Ryckaert, J-P, Ciccotti G, Berendsen HJC (1977). "Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes". *Journal of Computational Physics* **23**, 327–341.
41. Schlick T. (2001). Time-trimming tricks for dynamic simulations: Splitting force updates to reduce computational work. *Structure* **9**, 45-53.
42. Seibert M. M., Patriksoo A., Hess B. and David van der Spoel, (2005). Reproducible Polypeptide Folding and Structure Prediction using Molecular Dynamics Simulations. *J. Mol. Biol.* **354**, 173-183.
43. Sewald N. and Hans-Dieter Jakubke (2002). Peptides: Chemistry and Biology. Wiley, Second Edition.
44. Sewald N., Hans-Dieter Jakubke (2002). Peptides: Chemistry and Biology. *Wiley-VCH Verlag GmbH & Co. KgaA*
45. Shirts RB, Burt SR, Johnson AM. (2006). Periodic boundary condition induced breakdown of the equipartition principle and other kinetic effects of finite sample size in classical hard-sphere molecular dynamics simulation. *J Chem Phys* **125**, 164102.
46. Therrien C., Levesque R.C. (2000). Molecular basis of antibiotic resistance and β -lactamase inhibition by mechanism-based inactivators: perspectives and future directions. *FEMS Microbiology Reviews* **24**, 251-262.
47. Ueki N., Someya K., Matsuo Y., Wakamatsu K., and Mukai H. (2007). Cryptides: Functional Cryptic Peptides Hidden in Protein Structures. *Peptide Science* **88**, 190-198.

48. Wang J., Palzkill T., Dar-Chone Chow (2009). Structural Insight into the Kinetics and C_p of Interactions between TEM-1 β -Lactamase and β -Lactamase Inhibitory Protein (BLIP). *The journal of biological chemistry* **284**, 595–609.
49. Wilde R. E. and Singh S. (1998). *Statistical Mechanics, Fundamentals and Modern Applications*. John Wiley & Sons, Inc, New York.
50. Yeung, K. Y. and Ruzzo, W. L. (2001). An empirical study of Principal Component Analysis for clustering gene expression data. *Bioinformatics*.
51. [EMBnet molecular dynamics simulation Charmm](http://www.ch.embnet.org/MD_tutorial/)
http://www.ch.embnet.org/MD_tutorial/

Παράρτημα

Minimization and heating up script (heat.namd)

```
#
# 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
DCDunitcell      on

#
# 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       11
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        36          # <===== CHANGE ME
PmeGridsizeY        36          # <===== CHANGE ME
PmeGridsizeZ        36          # <===== CHANGE ME

#
# Periodic boundary things
#
wrapWater           on
wrapNearest         on

cellBasisVector1    36.00 00.00 00.00    # <===== CHANGE ME
cellBasisVector2    00.00 36.00 00.00    # <===== CHANGE ME
cellBasisVector3    18.00 18.00 18.00    # <===== CHANGE ME
cellOrigin          00.00 00.00 00.00    # <===== CHANGE ME

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

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

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

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

useGroupPressure    yes

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

```

```

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

#
# turn off pressure control until later
#
langevinPiston          off

#
# minimize nonbackbone atoms
#
minimize                400                ;# <===== CHANGE ME
output                  output/min_fix

#
# min all atoms
#
fixedAtoms              off
minimize                400                ;# <===== CHANGE ME
output                  output/min_all

#
# heat with CAs restrained
#
set temp 20;
while { $temp < 321 } {                ;# <===== Check me
  langevinTemp          $temp
  run                   400                ;# <===== CHANGE ME
  output                output/heat_ca
  set temp [expr $temp + 20]
}

#
# equilibrate volume with CAs restrained
#
langevinPiston          on
run                   400                ;# <===== CHANGE ME
output                  output/equil_ca

#
# equilibrate volume without restraints
#
constraintScaling       0
run                   2000                ;# <===== CHANGE ME

```


Equilibration script (equi.namd)

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

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

#
# 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              11
cutoff                  12
pairlistdist            13.5

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

```

#
# Particle Mesh Ewald parameters.
#
Pme                on
PmeGridsizeX      36          # <===== CHANGE ME
PmeGridsizeY      36          # <===== CHANGE ME
PmeGridsizeZ      36          # <===== CHANGE ME

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

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

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

useGroupPressure   yes

firsttimestep      9600        # <===== CHANGE ME
run                100000000   ;# <===== CHANGE ME

```