

Ταφάρη Μ. Θεοδώρα

Διπλωματική Εργασία

Προσομοιώσεις Μοριακής
Δυναμικής του Ν-τελικού
άκρου της ουβικτίνης

Επιβλέπων Καθηγητής
Νικόλαος Μ. Τζυκός

Δημοκρίτειο Πανεπιστήμιο Θράκης
Ιούλιος 2014

life

is a dynamic equilibrium of creation and destruction. Inside our cells the protein nano-machines, to which we owe our distinction from the inorganic, are perpetually recycled and rebuilt, forever battling the inevitable fate of entropic decay. Once covalently tagged with Ubiquitin, a protein is doomed to destruction by the proteasome, a protein degradation machine found in all of our cells. The component amino acids are then reused to synthesize new proteins. The constant recycling and rebuilding of proteins not only ensures that damaged proteins are removed quickly, but also allows rapid regulation of enzyme levels in the cell. The meandering path of the metal ribbon closely follows the fold and thus the internal structure of Ubiquitin. It features all of the major structural elements of typical proteins, including two alpha helices and a curved beta sheet. Its small size (76 amino acids) makes it one of the most studied proteins for protein folding and dynamics.

"Angel of Death" - Ubiquitin Copper, Steel - 9" x 9" x 16" 2011 - currently on display at

ScienceHouse

Πίνακας περιεχομένων

ΠΕΡΙΛΗΨΗ.....	8
Abstract	11
ΚΕΦΑΛΑΙΟ 1	13
1.1 Το πρόβλημα της αναδίπλωσης των πρωτεϊνών	14
1.2 Ο ρόλος των πεπτιδίων στη μελέτη της αναδίπλωσης	24
1.3 Στόχος εργασίας.....	30
1.4 Η προέλευση του πεπτιδίου 1E0Q	31
1.5 Δομική ανάλυση του πεπτιδίου 1E0Q.....	36
1.6 Άλλες προσομοιώσεις με το πεπτίδιο 1E0Q	40
ΚΕΦΑΛΑΙΟ 2	50
Εισαγωγή στις προσομοιώσεις μοριακής δυναμικής.....	51
2.1 Μοριακή μηχανική και δυναμικά πεδία	52
2.3 Προσομοιώσεις μοριακής δυναμικής	61
ΚΕΦΑΛΑΙΟ 3	63
Το πεπτίδιο από τη σκοπιά των προσομοιώσεων μοριακής δυναμικής.....	64
3.1 Τεχνικά χαρακτηριστικά υπολογιστικών συστημάτων.....	65
3.2 Προαπαιτούμενα αρχεία.....	68
3.3 Προετοιμασία συστήματος.....	74
3.3.1 Ορισμός αρχικών συντεταγμένων	75
3.3.2 Ελαχιστοποίηση και θέρμανση του αρχικού συστήματος.....	78
3.3.3 Παραγωγική φάση.	79
3.3.4 Μείωση του αριθμού μορίων νερού του συστήματος.	81
ΚΕΦΑΛΑΙΟ 4	86
Παρουσίαση και ανάλυση των αποτελεσμάτων.....	87
4.1 RMSD based ανάλυση των τροχιακών	88
4.2 Ανάθεση δευτεροταγούς δομής	120
4.3 Cluster analysis.....	128
4.3.1 Cartesian Principal Component Analysis.....	130

4.3.2 Dihedral angle Principal Component Analysis	149
4.5 Αλληλεπικάλυψη των τροχιακών.....	185
4.6 Σύγκριση αποτελεσμάτων.....	187
ΚΕΦΑΛΑΙΟ 5	193
Συμπεράσματα και Συζήτηση.....	194
Βιβλιογραφικές αναφορές	198

Πίνακας εικόνων

Εικόνα 1.1 (Αναπαράγεται άνευ αδείας από Dill & Chan, Nature Struct Biol., 1997).....	19
Εικόνα 1.2 (αναπαράγεται άνευ αδείας από W.S Horne & S.H. Gellman, Accounts of chemical research, 2008).....	26
Εικόνα 1.3 (αναπαράγεται άνευ αδείας από Sophie E. Jackson , 2006).	32
Εικόνα 1.4 (αναπαράγεται άνευ αδείας από R. Zerella et al., 1999).....	33
Εικόνα 1.5	36
Εικόνα 1.6 (αναπαράγεται άνευ αδείας από R. Zerella et al. 2000).....	39
Εικόνα 1.7 (αναπαράγεται άνευ αδείας από Seokmin Shin και Youngshang Pak, 2002).....	41
Εικόνα 1.8 (αναπαράγεται άνευ αδείας από W. L. Jorgensen et al., 2004)	43
Εικόνα 1.9 (αναπαράγεται άνευ αδείας από S. Jang et al., 2007)	45
Εικόνα 1.10 (αναπαράγεται άνευ αδείας από S. Jang et al., 2008)	46
Εικόνα 1.11 (αναπαράγεται άνευ αδείας από N. Vaidehi et al., 2011)	48
Εικόνα 3.1 Σχηματική αναπαράσταση του υπολογιστικού συστήματος του εργαστηρίου (http://utopia.duth.gr/~glykos/norma.html)	66
Εικόνα 3.2.	77
Εικόνα 3.3	82
Εικόνα 3.4	83
Εικόνα 3.5	84
Εικόνα 3.6	85
Εικόνα 4.1	90
Εικόνα 4.2	91
Εικόνα 4.3	92
Εικόνα 4.4	93
Εικόνα 4.5	94
Εικόνα 4.6	96
Εικόνα 4.7	97

Εικόνα 4.8	99
Εικόνα 4.9	103
Εικόνα 4.10	105
Εικόνα 4.11	109
Εικόνα 4.12	111
Εικόνα 4.13	112
Εικόνα 4.14	116
Εικόνα 4.15	118
Εικόνα 4.16	119
Εικόνα 4.17	122
Εικόνα 4.18	124
Εικόνα 4.19	126
Εικόνα 4.20	127
Εικόνα 4.21	132
Εικόνα 4.22	133
Εικόνα 4.23	136
Εικόνα 4.24	137
Εικόνα 4.25	138
Εικόνα 4.26	140
Εικόνα 4.27	142
Εικόνα 4.28	144
Εικόνα 4.29	145
Εικόνα 4.30	148
Εικόνα 4.31	151
Εικόνα 4.32	155
Εικόνα 4.33	156
Εικόνα 4.34	157
Εικόνα 4.35	160
Εικόνα 4.36	162
Εικόνα 4.37	163

Εικόνα 4.38	164
Εικόνα 4.39	167
Εικόνα 4.40	169
Εικόνα 4.41	170
Εικόνα 4.42	172
Εικόνα 4.43	173
Εικόνα 4.44	175
Εικόνα 4.45	184
Εικόνα 4.46	186
Εικόνα 4.47	190
Εικόνα 4.48	191
Εικόνα 4.49	192

Πίνακας πινάκων

Πίνακας 1.1 (αναπαράγεται άνευ αδείας από N. Vaidehi et al., 2011)	49
Πίνακας 3.1 Χαρακτηριστικά και συνθήκες της εκάστοτε προσομοίωσης.	67
Πίνακας 4.1	108
Πίνακας 4.2	115
Πίνακας 4.3	134
Πίνακας 4.4	135
Πίνακας 4.5	152
Πίνακας 4.6	166
Πίνακας 4.7	180
Πίνακας 4.8	181
Πίνακας 4.9	183
Πίνακας 4.10	186

ΠΕΡΙΛΗΨΗ

Ένα από τα δημοφιλέστερα, άλυστα ζητήματα της επιστημονικής κοινότητας της εποχής μας είναι το πρόβλημα της αναδίπλωσης των πρωτεϊνών, δηλαδή η αποσαφήνιση των διαδικασιών που κατευθύνουν το σχηματισμό της φυσικής δομής μιας πρωτεΐνης μέσα σε λίγα χιλιοστά του δευτερολέπτου. Στα πλαίσια του ζητήματος αυτού επιλέχθηκε ως σύστημα μελέτης το πεπτίδιο 1E0Q, το οποίο είναι το N-τελικό δεκαεπταπεπτίδιο της ουβικιτίνης που φέρει μία σημειακή μετάλλαξη. Η ουβικιτίνη αποτελεί πρωτεΐνη υψίστης σημασίας για την ομαλή λειτουργία των κυττάρων, καθώς αποτελεί τμήμα του πρωτεασώματος και συμβάλει καθοριστικά στη διαδικασία της ουβικιτινίωσης, μία διαδικασία με την οποία επιτυγχάνεται η σωστή αποικοδόμηση των πρωτεϊνών, εντός του κυττάρου. Η μη ομαλή διεξαγωγή αυτής της διαδικασίας μπορεί να οδηγήσει σε απορύθμιση του κυτταρικού κύκλου και ως εκ τούτου να οδηγήσει στην εμφάνιση ασθενειών. Το πεπτίδιο 1E0Q αποτελείται από 17 αμινοξέα και φέρει την μεταλλαγή T9D. Πειράματα NMR και κρυσταλλογραφίας ακτίνων X έχουν δείξει ότι αποκτά δομή β-φουρκέτας και έχει τη δυνατότητα να αναδιπλώνεται ανεξάρτητα από το υπόλοιπο πρωτεϊνικό τμήμα της ουβικιτίνης. Οι προσομοιώσεις μοριακής δυναμικής έχουν χρησιμοποιηθεί εκτενώς για τη μελέτη της πρωτεϊνικής αναδίπλωσης τις τελευταίες δεκαετίες. Το εύρος του μεγέθους των συστημάτων και οι παράμετροι που έχουν χρησιμοποιηθεί ποικίλλουν σε μεγάλο βαθμό.

Στην παρούσα πτυχιακή εργασία πραγματοποιήθηκαν τρεις προσομοιώσεις με το πεπτίδιο 1E0Q, υπό τις ίδιες συνθήκες, αλλά με διαφορετικές παραμέτρους (εμπειρικά δυναμικά πεδία - force fields). Τα πειραματικά δεδομένα συγκρίνονται με τα τροχιακά που προέκυψαν από τρεις προσομοιώσεις. Από τη σύγκριση προέκυψε πως τα δυναμικά πεδία που χρησιμοποιήθηκαν μπορούν να αποδώσουν δομές που πλησιάζουν πολύ στη φυσική δομή. Σύμφωνα με τα

αποτελέσματα των αναλύσεων μας οι δύο από τις τρεις προσομοιώσεις αναδιπλώθηκαν επιτυχώς σε διαμορφώσεις που μοιάζουν με την φυσική. Ωστόσο, διαφέρουν το ενεργειακό τοπίο αλλά και στην κινητική της αναδίπλωσης. Η τρίτη προσομοίωση περιλαμβάνει την πρωτονιομένη κατάσταση του πεπτιδίου. Σε αυτή τα αποτελέσματα δεν βρίσκονται σε συμφωνία με τα πειραματικά δεδομένα και καταλήγουν σε μία εναλλακτική, λιγότερο σταθερή διαμόρφωση.

Η μελέτη του ενεργειακού πεδίου της αναδίπλωσης του πεπτιδίου και των γεγονότων αναδίπλωσης συμπληρώνει τα κινητικά δεδομένα για το πεπτίδιο και θα μπορούσε να αποτελεί συνέχεια της παρούσας πτυχιακής εργασίας.

Abstract

One of the most charming and unresolved issues of the structural and biochemical community is the protein folding problem, which is the identification of processes that direct the formation of the physical structure of a protein in a few milliseconds. In this context, 1E0Q peptide was chosen as the study system, 1E0Q is mutation of the N-terminus of ubiquitin. Ubiquitin, a protein of the utmost importance for proper cell function, is part of the proteasome and plays a major part in the process of ubiquitination. Such procedures play critical roles in the regulation of the cell cycle and are thus important targets for diseases that derive from their deregulation. The 1E0Q peptide comprises 17 amino acids and carries the single point mutation, T9D. According to NMR and X-ray crystallography experiments, the peptide acquires a β -hairpin structure and has the ability to fold independently from the rest of the protein.

Molecular dynamics simulations have been used extensively in the past decade for the study of protein folding. They have been used in a wide range of systems, with many combinations of parameters. In the present thesis, we performed three molecular dynamics simulations using the 1E0Q peptide under the same conditions, but with different force fields. We compared the experimental structures with the structures predicted by our three independent folding simulations. According to our data, two of the used force fields successfully folded to the native structure. However, they differ on the predicted folding landscape and on the kinetics of the folding. The third simulation, concerning the protonated form of the peptide predicted an alternative, less stable, conformation of the peptide.

The study of the energy landscape of the peptide folding and folding events complements the kinetic data for the 1E0Q peptide and can be included in further work for this issue.

REQUAMIO 1

In theory, all one needs to know in order to fold a protein into its biologically active shape is the sequence of its constituent amino acids. Why has nobody been able to put theory into practice?

f.M. Richards

1.1 Το πρόβλημα της αναδίπλωσης των πρωτεϊνών

Το πρόβλημα της αναδίπλωσης των πρωτεϊνών αποτελεί αναμφισβήτητα ένα από τα δημοφιλέστερα, άλυτα προβλήματα της σύγχρονης Μοριακής Βιολογίας. Ο τρόπος με τον οποίο συμβάλει η αμινοξική αλληλουχία μιας πρωτεΐνης στον καθορισμό της τριτοταγούς δομής της αποτελεί το γνωστό “πρόβλημα της αναδίπλωσης των πρωτεϊνών” (Dill et al., 2008). Η επίλυση του προβλήματος της αναδίπλωσης των πρωτεϊνών αποτελεί πρόκληση τις τελευταίες δεκαετίες για την επιστημονική κοινότητα και κατατάσσεται στα μεγαλύτερα άλυτα προβλήματα.

Το πρόβλημα της αναδίπλωσης ανέκυψε μέσα από την ανάλυση των πρώτων πρωτεϊνικών δομών σε ατομικό επίπεδο, κατά τη δεκαετία του '60. Μπορεί να θεωρηθεί “τρισδιάστατο”, καθώς αποτελείται από τρία διαφορετικά προβλήματα -συνιστώσες. Το πρώτο είναι ο κώδικας αναδίπλωσης και υπόκειται στους νόμους της θερμοδυναμικής, δηλαδή τις διαμοριακές δυνάμεις που κατευθύνουν την αναδίπλωση μιας δεδομένης αμινοξικής αλληλουχίας στη δομή της. Το δεύτερο κομμάτι του προβλήματος αφορά τη διαδικασία της αναδίπλωσης και πιο συγκεκριμένα τους νόμους της κινητικής που διέπουν τα μονοπάτια της αναδίπλωσης που ακολουθούν οι πρωτεΐνες, ώστε να αναδιπλωθούν γρήγορα. Τέλος, το τρίτο αποτελεί μια υπολογιστική πρόκληση, και είναι η πρόβλεψη της φυσικής δομής από την αμινοξική αλληλουχία.

Η θερμοδυναμική υπόθεση που διατύπωσε ο Christian Anfinsen μαζί με τα πειράματα που πραγματοποίησε με τους συνεργάτες του στις ριβονουκλεάσες αποτελούν ορόσημο για την επιστήμη μας, καθώς τεκμηρίωσε πως η φυσική δομή μιας πρωτεΐνης αντιστοιχεί στην θερμοδυναμικά σταθερή δομή και εξαρτάται μόνο από την αλληλουχία και τις συνθήκες του διαλύματος και όχι από την οδό της αναδίπλωσης (κινητική). Επιπλέον συμπέρανε πως η

αναδίπλωση είναι ανεξάρτητη της ύπαρξης ριβοσομάτων, πρωτεϊνών συνοδών (chaperones) και του ιδίου του κυττάρου και μπορεί κάλλιστα να πραγματοποιηθεί *in vitro* σε δοκιμαστικό σωλήνα. Τέλος από την εργασία του Anfinsen συνεπάγεται το εξής : παρόλο τη δράση της φυσικής επιλογής σε μια αλληλουχία, η σταθερότητα της αναδίπλωσης και η κινητική της υπόκεινται στους νόμους της φυσικοχημείας (Dill et al. 2008).

Η υπόθεση του Anfinsen έρχεται σε σύγκρουση με την μετέπειτα παραδοχή πως η διαδικασία της αναδίπλωσης ακολουθεί κάποιο κυρίαρχο μονοπάτι ή αποτελείται από παράλληλα γεγονότα (αλληλεπιδράσεις) που περιλαμβάνουν πολλαπλά μονοπάτια, είναι γνωστή ως “παράδοξο του Levinthal”. Ο Levinthal, το 1968, υποστήριξε πως εάν μία πρωτεΐνη επισκεφθεί, κατά την αναδίπλωση της, όλους τους πιθανούς συνδυασμούς (διαμορφώσεις) θα χρειαστεί πολύ μεγαλύτερο χρόνο για να αναδιπλωθεί σε σχέση με τα ελάχιστα δευτερόλεπτα στα οποία συμβαίνει στην πραγματικότητα η αναδίπλωση της. Θεώρησε λοιπόν πως αν γίνουν κατανοητοί οι φυσικοί μηχανισμοί της πρωτεϊνικής αναδίπλωσης, θα είναι δυνατός ο σχεδιασμός αλγορίθμων, οι οποίοι θα μπορούν να προβλέψουν τη φυσική δομή από μια δεδομένη αμινοξική αλληλουχία.

Από την εποχή του Anfinsen όλο και περισσότερα ερωτήματα πλαισιώνουν το κύριο ερώτημα της πρωτεϊνικής αναδίπλωσης. Προβλήματα όπως η μέτρηση γεγονότων αναδίπλωσης σε χρονικά διαστήματα μικρότερα από μερικά χιλιοστά του δευτερολέπτου καθώς και η παρακολούθηση της αναδίπλωσης σε μεμονωμένες αλυσίδες μονομερών ταλάνιζαν εκείνη την εποχή την επιστημονική κοινότητα. Η έλλειψη εργαλείων για τα πειράματα αναδίπλωσης αποτελεί παρελθόν, καθώς υπάρχουν δεκάδες τρόποι για την επίλυση των δύο προαναφερθέντων ζητημάτων. Τεχνικές όπως το Laser temperature-jump (Callender RH et al., 1998), μέθοδοι μεταλλάξεων που αποδίδουν ποσότητες που ονομάζονται τιμές ϕ ή ψ , για την ταυτοποίηση αυτών των αμινοξέων και τον έλεγχο της ταχύτητας της αναδίπλωσης (Fersht AR et al., 2004, Krantz BA et al.,

2004, D.P Goldenberg, 1988, A.R Fersht et al., 1995), η μέθοδος FRET μέσω της οποίας καθίσταται δυνατή η παρακολούθηση σχηματισμού συγκεκριμένων επαφών (B Schuler et al., 2002), η μέθοδος ανταλλαγής υδρογόνου με την οποία φαίνονται γεγονότα δομικής αναδίπλωσης (Maity H et al., 2005, M.M Krishna et al., 2004). Μέχρι σήμερα έχει πραγματοποιηθεί αρκετή πρόοδος στο πεδίο της μελέτης του μηχανισμού αναδίπλωσης, όχι μόνο σε πειραματικό επίπεδο αλλά και σε θεωρητικό, με πολλά μοντέλα να προσπαθούν να εξηγήσουν το φαινόμενο (K.A Dill et al., 2008). Τα κυριότερα εξ αυτών αναφέρονται παρακάτω.

Diffusion-collision model

Σε αυτό το μοντέλο η πρωτεΐνη φαίνεται να αποτελείται από πολλά τμήματα (“elementary microdomains”), καθένα από τα οποία είναι αρκετά μικρό ώστε να υιοθετεί όλες τις πιθανές διαμορφώσεις και να μπορεί να αναδιπλωθεί αρκετά γρήγορα σε σχέση με το χρόνο αναδίπλωσης ολόκληρης της πρωτεΐνης. Η ύπαρξη των “microdomains” στην αναδιπλούμενη πρωτεΐνη παρέχει έναν τρόπο αποφυγής της ανεύρεσης όλων των πιθανών εναλλακτικών διαμορφώσεων. Τα “microdomains” διαχέονται υπό την επίδραση εσωτερικών και τυχαίων εξωτερικών δυνάμεων μέχρι να συγκρουστούν. Έτσι δημιουργούνται μεγαλύτερα συσσωματώματα τα οποία ονομάζονται “subdomains”. Αυτό συμβαίνει όταν τα microdomains έχουν αποκτήσει στοιχεία δευτεροταγούς δομής και η σύγκρουση έχει κατάλληλο προσανατολισμό. Σταδιακά σχηματίζεται ο πρωτεϊνικός σκελετός της φυσικής διαμόρφωσης της πρωτεΐνης. Το τελευταίο στάδιο είναι ο ακριβής σχηματισμός της τριτοταγούς δομής, που περιλαμβάνει το πακετάρισμα των πλευρικών αλυσίδων (Karplus & Weaver, 1979, 1994). Ορισμένες πρωτεΐνες, όπως τα ελικοειδή δεμάτια, ακολουθούν έναν παρόμοιο τρόπο αναδίπλωσης, που ονομάζεται hierarchical diffusion-collision model. Η δευτεροταγής δομή σχηματίζεται με ιεραρχικό τρόπο (J.A McCammon et al., 1977, J.K Myers, 2001).

Nucleation-condensation mechanism

Αυτός ο μηχανισμός προτάθηκε στα μέσα της δεκαετίας του '90 από τον Fersht μετά τα πειράματα που πραγματοποίησε με τον αναστολέα της χυμοθρυψίνης 2 (CI2), μια πρωτεΐνη που αποτελείται από 64 αμινοξέα. Σε αυτό το μοντέλο χρησιμοποιείται η έννοια του πυρήνα για να περιγράψει διάχυτες, εκτεταμένες περιοχές οι οποίες παρατηρούνται σε μεταβατική κατάσταση (transition state) και όχι μια καλοσχηματισμένη δομή γύρω από την οποία διαμορφώνεται η δομή της πρωτεΐνης (A.R Fersht, 1997). Τα πειράματα που πραγματοποιήθηκαν δείχνουν πως η δευτεροταγής και τριτοταγής δομή της πρωτεΐνης σχηματίζονται ταυτοχρόνως και η μια είναι συνέπεια της άλλης, καθώς ο πυρήνας που σχηματίζεται δεν μπορεί να σταθεροποιηθεί χωρίς αλληλεπιδράσεις με απομακρυσμένα κατάλοιπα. Ο μηχανισμός αναδίπλωσης του CI2 φαίνεται να ακολουθεί την κινητική δύο σταδίων, επειδή υπάρχει έλλειψη σταθερότητας στο ενδιάμεσο στάδιο (A.R Fersht et al., 1995). Άλλα παραδείγματα αναδίπλωσης που ακολουθούν την κινητική δύο σταδίων είναι ο IgG-binding domain της πρωτεΐνης G του στρεπτόκοκκου (P. Alexander et al. 1992), ο SH3 domain της σπεκτρίνης (A.R. Viguera et al., 1994), η πρωτεΐνη πρόσδεσης του acyl coenzyme A (B.B Kragelund et al., 1995), ένα τμήμα 80 καταλοίπων του καταστολέα του φάγου λ (G.S. Huang et al., 1995).

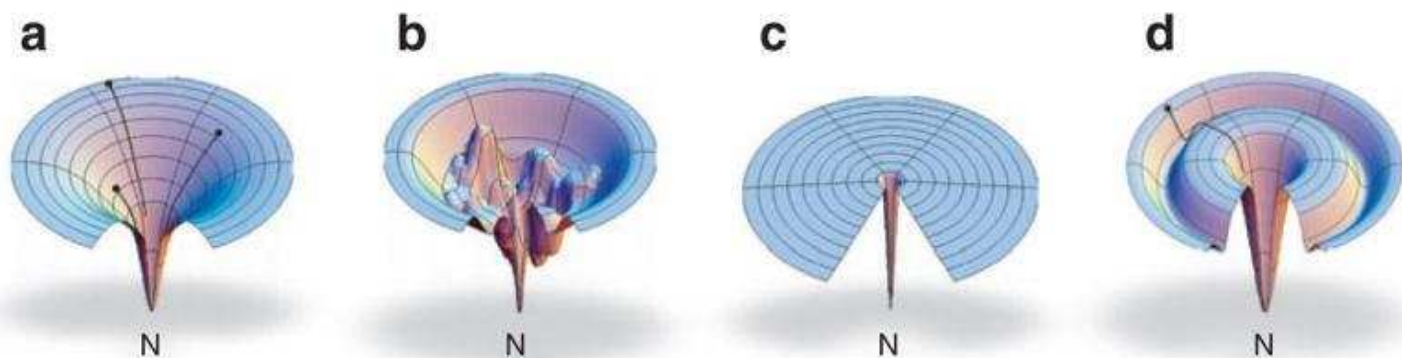
Funnel-Shaped energy landscapes

Η μεγάλη πρόκληση στο πρόβλημα της αναδίπλωσης είναι η περιγραφή των μυριάδων διαμορφώσεων από τις οποίες περνά η αμινοξική αλληλουχία προκειμένου να φτάσει στη φυσική της δομή. Τα τελευταία χρόνια αναπτύχθηκε το μοντέλο του ενεργειακού τοπίου το οποίο καλύπτει το φάσμα όλων των προηγούμενων μοντέλων και των εξαιρέσεων τους. Το ενεργειακό τοπίο μπορεί να απεικονιστεί ως μια χοάνη, στην κορυφή της οποίας αντιστοιχούν οι αρχικές αποδιαταγμένες (denatured structure) δομές μιας πρωτεϊνικής αλληλουχίας. Οι δομές αυτές αντιστοιχούν σε καταστάσεις υψηλής εντροπίας. Εν συνεχεία, βρίσκονται οι δομές ανάμεσα στις αποδιαταγμένες και τη φυσική δομή, που

αντιστοιχούν σε χαμηλότερης εντροπίας παροδικές δομές. Στο τέλος βρίσκεται η φυσική δομή της πρωτεϊνικής αλληλουχίας που αντιστοιχεί στο ενεργειακό ελάχιστο. Το μοντέλο αυτό προτείνει πως συνυπάρχουν παράλληλα μικροσκοπικά γεγονότα αναδίπλωσης που μπορεί να οδηγήσουν σε διαφορετικές οδούς μέχρι να αποκτηθεί η φυσική δομή. Με απλά λόγια, η χοάνη που αντικατοπτρίζει το ενεργειακό πεδίο περιέχει “ενεργειακά πηγάδια” καθένα από τα οποία αντιστοιχεί σε κάποια διαμόρφωση.

Η ιδέα του ενεργειακού πεδίου χρησιμοποιήθηκε προκειμένου να εξηγήσει την αποδιάταξη και την ισορροπία μεταξύ της εντροπίας με τις δυνάμεις που καθοδηγούν την αναδίπλωση. Οι πρωτεΐνες αποδιατάσσονται σε υψηλή θερμοκρασία, λόγω της ύπαρξης πολλών καταστάσεων υψηλής ενέργειας και λιγότερων χαμηλής ενέργειας. Αυτό σημαίνει πως η ενέργεια διοχετεύεται από τη μετουσιωμένη κατάσταση, μέσα από τις μεταβατικές καταστάσεις, προς τη φυσική (Dill et al., 2008). Αυτή είναι η “νέα αντίληψη” για την αναδίπλωση, αντικαθιστά την άποψη ότι η μετάβαση από την αποδιαταγμένη κατάσταση γίνεται μέσω ενός μονοπατιού προς τη φυσική κατάσταση, με την άποψη πως η μετάβαση είναι μια παράλληλη διαδικασία του συνόλου των μορίων μέσω διαφόρων μονοπατιών. Σε ένα ενεργειακό πεδίο, η αναδίπλωση μπορεί να περιγραφεί καλύτερα ως ένας όγκος νερού που ρέει σε μια πλαγιά, με ακανόνιστη μορφή και μέσω διαφόρων μονοπατιών και λιγότερο ως ένας όγκος νερού που ρέει σε ένα αυλάκι (Dill et al., 1997).

Η σχηματική απεικόνιση ενός “funnel” αποτελεί χρήσιμο εργαλείο για την σύνδεση των ιδιοτήτων στατιστικής μηχανικής και της κινητικής της αναδίπλωσης των πρωτεϊνών. Στην εικόνα 1.1 φαίνονται ορισμένες περιπτώσεις απεικόνισης.



Εικόνα 1.1 (Αναπαράγεται άνευ αδείας από Dill & Chan, Nature Struct Biol., 1997)

Ένας τύπος απεικόνισης των ενεργειακών τοπίων χρησιμοποιεί στους άξονες την ελεύθερη ενέργεια σε σχέση με τους βαθμούς ελευθερίας, προκειμένου να αντικατοπτριστεί η ποικιλία και η πυκνότητα των διαμορφώσεων. Ως “N” ορίζεται η φυσική δομή στον πυθμένα της χοάνης και αντικατοπτρίζει την ελάχιστη ελεύθερη ενέργεια. Στο σχήμα παρουσιάζονται: (a) Ένα ομοιόμορφο τοπίο (smooth energy landscape) για ένα ταχέως αναδιπλούμενο μόριο, (b) ένα τραχύ ενεργειακό τοπίο (rugged energy landscape), το οποίο περιέχει ενεργειακές “παγίδες”, δηλαδή πολλούς πυθμένες που αντιστοιχούν σε ενδιάμεσα στάδια, αλλά και στάδια στα οποία χρειάζεται περαιτέρω ενέργεια προκειμένου να υπερπηδηθεί το ενεργειακό εμπόδιο και να αποκτηθεί εν συνεχεία η φυσική δομή, (c) ένα ενεργειακό τοπίο το οποίο ονομάζεται “γήπεδο του γκολφ” (golf course) και η αναδίπλωση σε αυτό κυριαρχείται από αργή, τυχαία αναζήτηση σε όλη την επιφάνειά του, (d) ένα ενεργειακό τοπίο στο οποίο η αναζήτηση της φυσικής δομής πρέπει να περάσει από ένα υποχρεωτικό ενδιάμεσο στάδιο (moat landscape).

Συμπεραίνουμε λοιπόν πως τα ενεργειακά τοπία αποτελούν ένα χρήσιμο εργαλείο και μπορούν να παρέχουν τόσο ποσοτικά, όσο και ποιοτικά αποτελέσματα. Ένα από τα λίγα καθολικά χαρακτηριστικά, που διέπουν την κινητική της πρωτεϊνικής αναδίπλωσης και παρατηρήθηκε για πρώτη φορά από

τους Plaxco, Simons και Baker (PSB), είναι το ότι η ταχύτητα της αναδίπλωσης μιας πρωτεΐνης σχετίζεται με τις τοπολογικές ιδιότητες της φυσικής δομής. Ταχέως αναδιπλούμενα μόρια εμφανίζουν συχνότερα τοπική δομή, όπως οι έλικες, ενώ βραδέως αναδιπλούμενα μόρια εμφανίζουν συχνότερα μη-τοπική δομή, όπως τα β-φύλλα. Ωστόσο, δεν παύουν να υπάρχουν και εξαιρέσεις (K.W Plaxco et al., 1998). Μεταγενέστερα βρέθηκε ότι τα ποσοστά αναδίπλωσης συσχετίζονται και με άλλες τοπολογικές παραμέτρους της φυσικής δομής, όπως το μήκος της πρωτεϊνικής αλυσίδας (protein's effective chain length), δηλαδή το μήκος της αμινοξικής αλληλουχίας αφαιρουμένου των καταλοίπων που βρίσκονται στις έλικες (D.N Ivankov et al., 2004), το μήκος της δευτεροταγούς δομής (J.T Huang et al., 2007), οι επαφές που δημιουργεί κάθε κατάλοιπο με άλλα μακρινά κατάλοιπα (M.M Gromiha et al., 2001), το ποσοστό των επαφών που σχηματίζονται μεταξύ απομακρυσμένων καταλοίπων στην αλληλουχία (L Mirny et al., 2001). Πλέον υπάρχουν ακόμα και μέθοδοι που προβλέπουν την ταχύτητα της αναδίπλωσης ενός πεπτιδίου χρησιμοποιώντας τη σύσταση της αμινοξικής αλληλουχίας, ένα τέτοιο παράδειγμα είναι το πρόγραμμα FOLD-RATE (<http://psfs.cbrc.jp/fold-rate>, M.M Gromiha et al., 2006).

Ο μηχανισμός zipping and assembly (ZA)

Σύμφωνα με το μηχανισμό αυτό, κατά τα αρχικά στάδια της αναδίπλωσης (από picoseconds μέχρι nanoseconds) μικρά πεπτιδικά τμήματα της αλληλουχίας αναζητούν, καθένα ξεχωριστά, μικρές, τοπικές δομές, όπως β-στροφές, βρόγχους και ελικοειδείς στροφές. Ορισμένες από αυτές τις τοπικές δομές είναι ικανές να παραμείνουν σταθερές για μεγαλύτερο χρονικό διάστημα και να επεκταθούν (zipping) σε μεγαλύτερες και σταθερότερες δομές. Σε ακόμη μεγαλύτερα χρονικά διαστήματα ομάδες τέτοιων εκτεταμένων δομών συναρμολογούνται σε μεγαλύτερες δομές που μοιάζουν περισσότερο με τη φυσική δομή και σταθεροποιούνται. Ο μηχανισμός ZA έχει αρκετά κοινά, αλλά και διαφορές με τους μηχανισμούς που προαναφέρθηκαν (Dill et al., 2007).

Το τρίτο κομμάτι του προβλήματος της αναδίπλωσης, δηλαδή το υπολογιστικό κομμάτι που αφορά την πρόβλεψη της φυσικής δομής από μια δεδομένη αμινοξική αλληλουχία, ξεκίνησε να απασχολεί την επιστημονική κοινότητα από πολύ νωρίς. Αρχικά δημιουργήθηκαν πρωτεϊνικές βάσεις δεδομένων που χρησιμοποιούσαν αλγορίθμους πρόβλεψης της δευτεροταγούς δομής (P.Y Chou et al., 1974, 1978). Στη δεκαετία του '80 χρησιμοποιήθηκαν για πρώτη φορά μέθοδοι υπολογιστικής φυσικής, όπως τα ατομικά δυναμικά πεδία, ενώ στις αρχές της δεκαετίας του '90 σημειώθηκε σημαντική πρόοδος στη χρήση βάσεων δεδομένων και αλγορίθμων ανίχνευσης ομολογίας προκειμένου να συναρμολογηθούν δομές από ομόλογες αλληλουχίες και να αναγνωρισθεί η δομή άγνωστων αλληλουχιών από σύγκριση με τρισδιάστατες δομές που περιέχονται σε βάσεις δεδομένων.

Τι σημαίνει, όμως, πρόβλεψη της πρωτεϊνικής δομής με υπολογιστικές μεθόδους; Στο ερώτημα αυτό θα προσπαθήσουμε, αρχικά, να απαντήσουμε με ένα παράδειγμα. Έχοντας μια αμινοξική αλληλουχία (πρωτοταγής δομή) η οποία, σε φυσιολογικές συνθήκες θερμοκρασίας εντός υδατικού διαλύματος, αντιπροσωπεύει μια μονομερή σφαιρική πρωτεΐνη ο στόχος είναι να βρούμε όλα τα ελικοειδή και πτυχωτά τμήματα (έλικες και β-κλώνοι), όλα τα ζεύγη των κλώνων που σχηματίζουν β-φύλλα, όλους τους δισουλφιδικούς δεσμούς μεταξύ των καταλοίπων κυστεΐνης, όλους τους βρόγχους που ενώνουν τα στοιχεία δευτεροταγούς δομής και εν τέλει την τρισδιάστατη δομή της αναδιπλωμένης πρωτεΐνης.

Η πρόοδος των υπολογιστικών μεθόδων πρόβλεψης τριτοταγούς δομής εκτιμάται από τα πειράματα CASP (Critical Assessment of Protein Structure Prediction) που ξεκίνησαν το 1994 με το CASP1 και φτάνουν τις συνολικά 10 εκδοχές το 2012 (CASP10) (<http://predictioncenter.org/>). Στην κοινότητα του CASP οι ερευνητικές ομάδες εφαρμόζουν τις δικές τους μεθόδους πρόβλεψης σε αλληλουχίες για τις οποίες δεν είναι γνωστή η φυσική δομή, αλλά πρόκειται

σύντομα να δημοσιευθεί (J. Moult et al., 1995, 2001, 2003, 2005, J Moult, 1999) . Οι υπολογιστικές μέθοδοι για την πρόβλεψη της δομής των πρωτεϊνών μπορούν να ταξινομηθούν σε τέσσερις ομάδες :

A. Comparative modeling

Η μέθοδος της συγκριτικής μοντελοποίησης βασίζεται στη αρχή πως οι κοντινές εξελικτικά αλληλουχίες έχουν παρόμοια τρισδιάστατη δομή, δηλαδή πως η ομολογία στην αλληλουχία υποδηλώνει και δομική ομολογία (A. Sali et al., 1993). Όσο μεγαλύτερη είναι η ομοιότητα της αλληλουχίας, τόσο μεγαλύτερη μπορεί να είναι και η ομοιότητα της δομής. Σε αλληλουχίες με μεγαλύτερη από 50% ομολογία η πρόβλεψη είναι πολύ καλή. Σε αλληλουχίες με ομολογία από 30% έως 50% έχει δειχθεί πως το 80% των ατόμων Ca αναμένεται να βρίσκονται σε απόσταση 3,5Å από την πραγματική τους θέση. Ενώ σε αλληλουχίες με λιγότερο από 30% ομολογία οι προβλέψεις μπορεί να περιέχουν σημαντικά λάθη.

B. Fold recognition and threading

Η μέθοδος αυτή βασίζεται στην αρχή πως ο αριθμός των διαφόρων αναδιπλωμένων πρωτεϊνικών δομών είναι σημαντικά μικρότερος από τον αριθμό των αλληλουχιών που παράγονται από τα προγράμματα της γονιδιακής αλληλούχισης. Ο αριθμός των διαφόρων αναδιπλωμένων δομών έχει εκτιμηθεί με βάση την ομαδοποίηση των δομών, κατατεθειμένων στην πρωτεϊνική βάση δομών, σε οικογένειες (Berman et al., 2000). Η μέθοδος threading βασίζεται στην προσαρμογή μιας αλληλουχίας-στόχου σε μια γνωστή δομή μέσα σε μια βιβλιοθήκη που περιέχει δομές. Η μέθοδος αυτή εφαρμόστηκε με επιτυχία από το Skolnick και τους συνεργάτες του στα πειράματα CASP5-7 (Skolnick et al., 2003). Σε μια πιο πρόσφατη προσπάθεια του ο Skolnick χρησιμοποίησε μια επαναληπτική προσέγγιση στην οποία αρχικά στοιχίζονται η αλληλουχία-στόχος και οι γνωστές δομές, αγνοώντας το ταίριασμα των αμινοξικών

αλληλεπιδράσεων. Σε μεταγενέστερες στοιχίσεις λαμβάνονται υπόψιν πληροφορίες από προηγούμενες στοιχίσεις προκειμένου να βαθμολογηθούν οι ενέργειες από τις αμινοξικές αλληλεπιδράσεις (Skolnick et al., 2004) .

Γ. First principles methods with database information,

οι οποίες περιλαμβάνουν τις παρακάτω μεθόδους:

1. Fragment-based recombination methods,

Στις μεθόδους ανασυνδυασμού βασιζόμενου σε θραύσματα η κεντρική ιδέα είναι ότι οι τοπικές αλληλεπιδράσεις, που εξαρτώνται από την αλληλουχία, κατευθύνουν την αλυσίδα να δοκιμάσει συγκεκριμένα σύνολα τοπικών διαμορφώσεων, ενώ οι μη-τοπικές αλληλεπιδράσεις προτιμούν χαμηλής ενέργειας διαμορφώσεις, οι οποίες είναι συμβατές με τις προτιμώμενες τοπικές διαμορφώσεις (biased local conformers).

2. Hybrid methods,

Στις υβριδικές μεθόδους αναπτύχθηκαν προσεγγίσεις που συνδυάζουν την πολλαπλή σύγκριση της αλληλουχίας (multiple sequence comparison), το threading, τη βελτιστοποίηση με συναρτήσεις βαθμολόγησης και την ομαδοποίηση (clustering). Η χωροδιάταξη της πρωτεΐνης περιγράφεται από ένα μοντέλο πλέγματος με τρία ή λιγότερα άτομα ανά κατάλοιπο (Y. Zhang et al., 2003). Το threading χρησιμοποιείται για να παρέχει πληροφορίες για τις αλληλεπιδράσεις μεγάλης εμβέλειας, μεταξύ απομακρυσμένων πλευρικών αλυσίδων. Με τη χρήση του SPICKER, ένα εργαλείο για ομαδοποίηση πρωτεϊνικών δομών που παράγονται από μια προσομοίωση, και μετά από επιλογή των πολυπληθέστερων ομάδων, βρέθηκε πως οι διαμορφώσεις στις ομάδες αυτές τείνουν περισσότερο στη φυσική δομή, από ότι οι δομές με τη χαμηλότερη ελεύθερη ενέργεια (Y. Zhang et al., 2003).

3. Μέθοδοι που συνδυάζουν πληροφορίες από τη δευτεροταγή δομή και

επιλεγμένους περιορισμούς της τριτοταγούς με βελτιστοποίηση Monte Carlo (MC optimization) ή ολική αιτιοκρατική βελτιστοποίηση (deterministic global optimization) (Adjiman et al., 1996, 1997, Androulakis et al., 1995. Floudas, 2000, 2005. Floudas et al., 2005) .

Δ. First principles methods without database information

Στις μεθόδους που δεν χρησιμοποιούνται βάσεις δεδομένων, ο σκοπός είναι να ταυτοποιηθεί η δομή μιας πρωτεΐνης με τη χαμηλότερη ελεύθερη ενέργεια στο περιβάλλον της , χρησιμοποιώντας μόνο τους νόμους της φυσικής και την αμινοξική της αλληλουχία. Αυτές οι μέθοδοι μπορούν να χρησιμοποιηθούν για οποιαδήποτε αλληλουχία. Λόγω της μεγάλης ποικιλίας των πρωτεϊνών-στόχων και της αδυναμίας της χρήσης κάποιας βάσης δεδομένων, κατατάσσονται στις δυσκολότερες μεθόδους για την πρόγνωση της δομής (C.A. Floudas, 2007).

Οι υπολογιστικές προσομοιώσεις είναι ένα physics-based εργαλείο για την πρόβλεψη της δομής και τη μελέτη της αναδίπλωσης. Στη μέθοδο αυτή το μέτρο της επιτυχίας είναι η κατανόηση των δυνάμεων και της δυναμικής που διέπουν της ιδιότητες των πρωτεϊνών. Οι επιτυχία των μεθόδων αυτών θα επιτρέψει την υπέρβαση των ορίων της βιοπληροφορικής, στην πρόβλεψη αλλαγών στη διαμόρφωση, που είναι σημαντικές για την ανακάλυψη φαρμάκων με υπολογιστικές μεθόδους, στην κατανόηση των μηχανισμών δράσης των πρωτεϊνών, της κίνησης, της κατάλυσης, της αναδίπλωσης και άλλων καταστάσεων που απαιτούν περισσότερα από μια στατική φυσική δομή, στην κατανόηση της απόκρισης των πρωτεϊνών σε διαλύτες, pH, άλατα, αποδιατακτικούς και άλλους παράγοντες και τέλος στο σχεδιασμό συνθετικών πρωτεϊνών (Dill et al., 2008). Οι μέθοδοι των προσομοιώσεων μοριακής δυναμικής, που είναι και το αντικείμενο της παρούσας πτυχιακής θα συζητηθούν εκτενέστερα στα κεφάλαια 2 και 3.

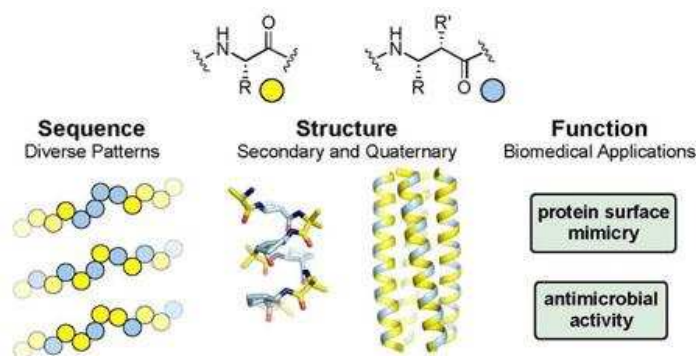
1.2 Ο ρόλος των πεπτιδίων στη μελέτη της αναδίπλωσης

Η ένωση δύο αμινοξέων μέσω πεπτιδικού δεσμού, έχει ως αποτέλεσμα τη δημιουργία ενός πεπτιδίου και συγκεκριμένα ενός διπεπτιδίου. Η συνεχής προσθήκη αμινοξέων στο καρβοξυτελικό άκρο έχει ως αποτέλεσμα την αύξηση του μήκους του πεπτιδίου. Με τον τρόπο αυτό είναι δυνατό να δημιουργηθούν πολυπεπίδια ή πρωτεΐνες αποτελούμενες από εκατοντάδες έως χιλιάδες αμινοξέα. Η πολυπλοκότητα, συνεπώς, μιας μεγάλης αμινοξικής αλληλουχίας (πρωτεΐνη) μπορεί να αντικατασταθεί από μικρότερα μοντέλα, τα οποία είναι ευκολότερο να μελετηθούν. Εφόσον τα πεπτίδια είναι μικρότερα συστήματα από τις πρωτεΐνες, δημιουργούν λιγότερες επαφές και μειώνεται το υπολογιστικό κόστος αλλά και ο υπολογιστικός χρόνος.

Μέσω των προσομοιώσεων μοριακής δυναμικής σε πεπτίδια μπορούν να χαρακτηριστούν η δυναμική και οι μοριακοί μηχανισμοί των αρχικών γεγονότων της πρωτεϊνικής αναδίπλωσης. Η αναδίπλωση των πεπτιδίων πραγματοποιείται στην κλίμακα των nanoseconds. Συνεπώς τα πεπτίδια είναι μια καλή ευκαιρία για τη γεφύρωση θεωρίας και πράξης (S. Gnanakaran et al., 2003). Για παράδειγμα η δουλειά του Snow και των συνεργατών του επιβεβαιώνει απόλυτη ταύτιση μεταξύ των αποτελεσμάτων από τις προσομοιώσεις και τα πειράματα που πραγματοποίησαν στην μικρο-πρωτεΐνη BBA5 που σχεδίασαν, προσφέροντας τη δυνατότητα άμεσης σύγκρισης ανάμεσα στην *in vitro* και την *in silico* πρωτεϊνική αναδίπλωση (C.D Snow et al., 2002).

Ο χειρισμός χαρακτηριστικών όπως ο σχηματισμός ιοντικών δεσμών και δεσμών υδρογόνου, η ενυδάτωση και ο σχηματισμός υδρόφοβου πυρήνα, που καθορίζουν τη σταθερότητα των πρωτεϊνών, αποτελεί πρόκληση για τις προσομοιώσεις μοριακής δυναμικής σε πεπτίδια. Η κατανόηση του μηχανισμού

της αναδίπλωσης σε πεπτίδια και μικρές πρωτεΐνες επιταχύνθηκε από τρεις σημαντικές εξελίξεις. Πρώτον, ο σχεδιασμός μικρών πεπτιδίων που μιμούνται την πολυπλοκότητα των πρωτεϊνών, αλλά είναι ικανοποιητικά μικρά έτσι ώστε να επιτρέπεται η λεπτομερής μελέτη τους (E. de Alba et al., 1999, M Ramirez-Alvarado et al., 2001, MT Pastor et al., 2002, JW Neidigh et al., 2002). Δεύτερον, η ανάπτυξη γρήγορης (nanosecond) time-resolved spectroscopy για την πειραματική μελέτη της δυναμικής της αναδίπλωσης των πεπτιδίων στην ίδια χρονική κλίμακα με τις προσομοιώσεις (WH Woodruff et al., 1996, V Munoz et al., 1997, JC Crane et al., 2000, LJ Lapidus et al., 2000, S Sporlein et al., 2002). Και τρίτον, η ανάπτυξη και η εφαρμογή των αλγορίθμων για τις προσομοιώσεις έχουν βοηθήσει να ξεπεραστούν οι περιορισμοί της ανεπαρκούς δειγματοληψίας (insufficient sampling) (Y Sugita et al., 2000, MR Shirts et al., 2001). Ωστόσο η συνεχής ανάπτυξη και βελτίωση των δυναμικών πεδίων στις προσομοιώσεις είναι σημαντική ώστε να επιτευχθεί μεγαλύτερη συμφωνία μεταξύ των προσομοιώσεων και των πειραμάτων (S. Gnanakaran et al., 2003).



Εικόνα 1.2 (αναπαράγεται άνευ αδείας από W.S Horne & S.H. Gellman, Accounts of chemical research, 2008)

Ετερογενή foldamers που αποτελούνται από α - και β - αμινοξικά κατάλοιπα. Ο στόχος της επιστημονικής κοινότητας είναι από μια αμινοξική αλληλουχία να κατανοήσουμε τη δευτεροταγή και τριτοταγή δομή έτσι ώστε να μπορέσουμε να ερμηνεύσουμε τη λειτουργία και να τη χρησιμοποιήσουμε στις βιοϊατρικές εφαρμογές.

Πέρα από την επαλήθευση των πειραμάτων και την βελτίωση των δυναμικών πεδίων τα πεπτίδια χρησιμοποιούνται και σε θεραπευτικούς σκοπούς. Τα πεπτίδια όμως που παράγονται φυσιολογικά στους οργανισμούς δεν είναι καλοί υποψήφιοι για φάρμακα, λόγω της χαμηλής βιοδιαθεσιμότητας, της δυνατότητας ανοσοποίησης και της χαμηλής μεταβολικής σταθερότητας *in vivo* (P.W Latham, 1999).

Πρόσφατες προσπάθειες επικεντρώθηκαν στη βελτίωση των μειονεκτημάτων των φυσικών πεπτιδίων και στη δημιουργία μη-φυσικών ολιγομερών που μιμούνται τα πεπτίδια. Τα “peptidomimetics” μπορούν να βασιστούν σε οποιοδήποτε ολιγομερές που μιμείται την αρχική δομή μέσω τροποποίησης της. Τα “peptidomimetics” είναι συχνά ανθεκτικά σε πρωτεάσες, έχουν μειωμένη ανοσογονικότητα και βελτιωμένη βιοδιαθεσιμότητα σε σχέση με τα πεπτιδικά ανάλογα. Εκτός από τη μίμηση της αρχικής δομής, τα αποκαλούμενα “foldamers” αποτελούν υποσύνολο των “peptidomimetics” και εμφανίζουν καλά καθορισμένα στοιχεία δευτεροταγούς δομής, όπως έλικες, στροφές, και μικρές δομές που μοιάζουν με φύλλα (εικόνα 1.2) (J.A Patch et al., 2002). Το ερώτημα εδώ είναι αν αυτά τα πολυμερή μπορούν να εξυπηρετήσουν λειτουργίες, όπως η κατάλυση, η αναγνώριση και άλλες ανάλογες με αυτές που υποστηρίζουν τα αναδιπλωμένα βιοπολυμερή. Για το λόγο αυτό μελετάται η δομή τους και η ικανότητα τους να υιοθετούν συγκεκριμένες διαμορφώσεις (Gellman et al.,2009).

Ένα παράδειγμα με “foldamers” που έχουν φαρμακευτική εφαρμογή είναι τα αντιμικροβιακά πεπτίδια (AMPs), τα οποία παρέχουν προστασία εναντίον παθογόνων βακτηρίων και αποτελούν σημαντικό τμήμα του ανοσοποιητικού συστήματος. Τα AMPs που παράγονται από τα ριβοσώματα έχουν μήκος 12-80 κατάλοιπα και υιοθετούν διάφορες διαμορφώσεις, α-έλικες και β-φύλλα κ.α. Ένας τρόπος για να σχεδιαστούν AMPs είναι ο *de novo* σχεδιασμός τους με coarse-grained μοντέλα και η χρήση coarse-grained προσομοιώσεων μοριακής δυναμικής που περιέχουν την αντιμικροβιακή ουσία αλλά και την κυτταρική

μεμβράνη προκειμένου να ελεγχθούν οι αλληλεπιδράσεις μεταξύ τους (GN Tew et al., 2009). Μια δεύτερη μέθοδος σχεδιασμού “foldamers” είναι η sequence-based, στην οποία πραγματοποιείται συστηματική τροποποίηση του πρωτεϊνικού σκελετού στην περιοχή δέσμευσης μιας πρωτεΐνης, με τον τρόπο αυτό τα foldamers μπορούν να δεσμευτούν ισχυρά και επιλεκτικά σε μια πρωτεϊνική επιφάνεια-στόχο (W.S Horne et al., 2008).

Υπάρχουν επίσης βάσεις δεδομένων που σχετίζονται με αντιμικροβιακά πεπτίδια και το σχεδιασμό τους. Η APD2 περιέχει 1228 πεπτίδια εκ των οποίων 65 είναι αντικαρκινικά, 76 αντιϊικά (53 αντί-HIV), 327 αντιμυκητιασικά και 944 αντιβακτηριδιακά. Μπορεί να παρέχει αναζήτηση σε οικογένειες πεπτιδίων, πηγές πεπτιδίων, θέσεις πρόσδεσης καθώς και στατιστική ανάλυση για τη σύσταση ενός πεπτιδίου-φαρμάκου (<http://aps.unmc.edu/AP/main.php>) (G Wang et al., 2009). Η PepBank είναι μια δεύτερη βάση που αφορά πεπτίδια και παρέχει αναζήτηση σε υπάρχοντα πεπτίδια με μήκος μικρότερο των 20 καταλοίπων (<http://pepbank.mgh.harvard.edu/>) (T Shtatland et al., 2007, T Duchrow et al., 2009).

Ένα σημαντικό μειονέκτημα του υπολογιστικού σχεδιασμού των foldamers είναι το ότι υπάρχει έλλειψη δομικών δεδομένων που θα επιτρέψουν την παραμετροποίηση των rotamer βιβλιοθηκών και των ενεργειών. Μια προσπάθεια για τη δημιουργία βιβλιοθηκών rotamers έγινε από τον Scott Shandler και τους συνεργάτες του. Οι τελευταίοι ερευνήσανε την χρήση της μοριακής μηχανικής για την κατασκευή βιβλιοθηκών με rotamers για το σκελετό μη-φυσικών foldamers. Οι παράμετροι που χρησιμοποιήθηκαν για το μήκος δεσμού και τις γωνίες ήταν αυτοί του δυναμικού πεδίου CHARMM27, με τροποποιήσεις στις διεδρες γωνίες του σκελετού. Επιπλέον όλες οι πλευρικές αλυσίδες τοποθετήθηκαν σε β-ομογλυκίνη για κάθε υποψήφιο ικρίωμα. Οι πλευρικές αλυσίδες τοποθετήθηκαν στη θέση C³ του κεντρικού καταλοίπου σε κάθε ικρίωμα για να δημιουργηθεί ένα L-β-αμινοξύ (S.J Shandler et al., 2010).

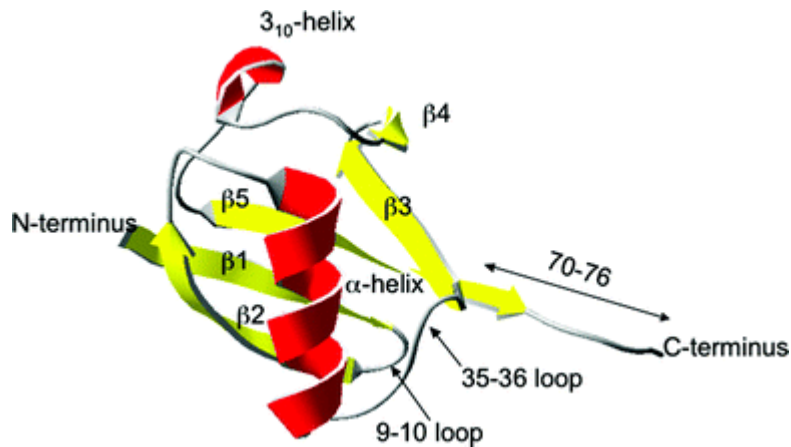
Συνοψίζοντας, είναι εμφανές ότι η χρήση των πεπτιδίων έχει ένα ευρύ φάσμα εφαρμογών στη βιολογία, τη βιοχημεία, τη φυσική, τη βιοπληροφορική, την ιατρική και άλλους επιστημονικούς κλάδους. Οι μελέτες συνεχίζονται, έτσι ώστε να κατανοηθούν απόλυτα οι μηχανισμοί αναδίπλωσης και δράσης των πεπτιδίων και των πρωτεϊνών και στη συνέχεια να χρησιμοποιηθούν σε περισσότερο πρακτικές εφαρμογές.

1.3 Στόχος εργασίας

Έχοντας λάβει υπόψην όλες τις προηγούμενες μελέτες που έχουν πραγματοποιηθεί για το πεπτίδιο 1E0Q και γνωρίζοντας πως δεν έχει πραγματοποιηθεί καμία εκτενής μελέτη αναδίπλωσης του πεπτιδίου βασισμένη σε προσομοιώσεις μοριακής δυναμικής χρησιμοποιώντας τα δυναμικά πεδία AMBER 99SB και AMBER-ILDN, θέτουμε ως στόχο της παρούσας πτυχιακής εργασίας την επίλυση του προβλήματος αναδίπλωσης του πεπτιδίου 1E0Q με τη χρήση physics-based μεθόδων και τη σύγκριση των αποτελεσμάτων με την πειραματικά προσδιορισμένη δομή και τα ευρήματα των προηγούμενων μελετών.

1.4 Η προέλευση του πεπτιδίου 1E0Q

Το υπό μελέτη πεπτίδιο, αποτελεί μια μεταλλαγμένη μορφή του N-τελικού δεκαεπτά-πεπτιδίου της ουβικιτίνης. Η ουβικιτίνη (εικόνα 1.3) είναι μια μικρή πρωτεΐνη που αποτελείται από 76 αμινοξέα και εκφράζεται από το ανθρώπινο γονίδιο UbB, το οποίο εδράζεται στην περιοχή 17p11.1-17p12 (G.C. Webb et al., 1990). Στην πραγματικότητα η οικογένεια των γονιδίων της ουβικιτίνης χωρίζεται σε δύο τύπους, ο πρώτος αποτελείται από επαναλήψεις ουβικιτίνης και ονομάζεται “polyubiquitin gene” και ο δεύτερος αποτελείται από την ουβικιτίνη συνδεδεμένη με μια πρωτεϊνική ουρά και ονομάζεται “ubiquitin fusion gene” (Schlesinger and Bond, 1987). Στον άνθρωπο η οικογένεια των γονιδίων της ουβικιτίνης χωρίζεται σε δύο “ubiquitin fusion gene” υποοικογένειες, UbA52 και UbA80, στις οποίες κωδικοποιείται η ουβικιτίνη συνδεδεμένη με μια ουρά αποτελούμενη από 52 και 80 αμινοξέα αντίστοιχα (Lund et al., 1985. Salvesen et al., 1987. Baker 1988), και δύο “polyubiquitin gene” υποοικογένειες, UbB και UbC, στις οποίες συμπεριλαμβάνονται τρεις και εννέα κωδικές περιοχές αντίστοιχα. Στα αλληλόμορφα UbC έχουν βρεθεί περιστασιακά οκτώ ή επτά κωδικές περιοχές (Wiborg et al., 1985, Baker and Board 1987a, 1989). Επιπλέον σε κάθε υποοικογένεια, εκτός της UbC, έχουν βρεθεί και ψευδογονίδια (Wiborg et al., 1985, Baker and Board 1987a, 1987b, Baker 1988, Cowland et al., 1988).

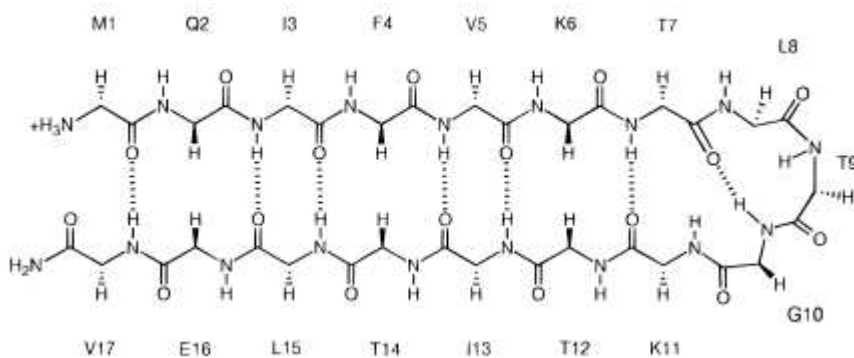


Εικόνα 1.3 (αναπαράγεται άνευ αδείας από Sophie E. Jackson , 2006).

Παρουσίαση της τρισδιάστατης δομής της ουβικιτίνης σε ανάλυση 1.8 Angstroms (1ubq.pdb). Το μόριο της ουβικιτίνης είναι αρκετά συμπαγές καθώς το 87% της πολυπεπτιδικής αλυσίδας συμμετέχει στο σχηματισμό στοιχείων δευτεροταγούς δομής. Με κόκκινο χρώμα φαίνεται η κύρια α-έλικα καθώς και η 3₁₀-έλικα, ενώ με κίτρινο χρώμα παρουσιάζονται οι πέντε β-κλώνοι, οι οποίοι σχηματίζουν ένα β-φύλλο με παράλληλη και αντί-παράλληλη διάταξη. Φαίνονται επίσης οι περιοχές των βρόχων, που έχουν σχεδιαστεί σε διάφορες μελέτες.

Τα αρχικά 17 αμινοξέα, του N-τελικού άκρου, της ουβικιτίνης σχηματίζουν β-φουρκέτα και μελετήθηκαν, ανεξάρτητα από την πρωτεΐνη, για πρώτη φορά από την επιστημονική ομάδα της R. Zerella στα τέλη της δεκαετίας του '90 (R. Zerella et al., 1999). Η R. Zerella και οι συνεργάτες της πραγματοποίησαν πειράματα H^1 NMR (nucleus magnetic resonance) προκειμένου να χαρακτηρίσουν τη δομή του πεπτιδίου, καθώς φαίνεται να παίζει σημαντικό ρόλο στα αρχικά στάδια της αναδίπλωσης της πρωτεΐνης. Η φυσική δομή του δεκαεπτά-πεπτιδίου στην πρωτεΐνη είναι μια β-φουρκέτα, όπως φαίνεται και στην εικόνα 1.4. Από τα πειράματα τους βρέθηκε πως το πεπτίδιο έχει την ικανότητα να αναδιπλώνεται ανεξάρτητα από την υπόλοιπη αμινοξική αλληλουχία και κατά την αναδίπλωση του, σε υδατικό διάλυμα, η δομή του πλησιάζει πολύ τη φυσική δομή. Ωστόσο οι επαφές που σχηματίζουν οι πλευρικές αλυσίδες του απομονωμένου πεπτιδίου διαφέρουν από αυτές της

φυσικής αλυσίδας. Όπως είναι γνωστό οι επαφές που δημιουργούνται τοπικά επηρεάζουν και επηρεάζονται από τις επαφές που δημιουργούνται με απομακρυσμένα κατάλοιπα έτσι ώστε να σταθεροποιηθεί η φυσική δομή. Ωστόσο, η ολική συνέργεια των επαφών δεν αποτελεί απόλυτη προϋπόθεση για την αναδίπλωση, έτσι είναι συχνά ανιχνεύσιμη η αναδίπλωση απομονωμένων στοιχείων δευτεροταγούς δομής. Αυτή η ανεξάρτητη αναδίπλωση μπορεί να παίζει σημαντικό ή υποβοηθητικό ρόλο στην αναδίπλωση άλλων στοιχείων της φυσικής πρωτεΐνης και είναι πιθανό να κατευθύνει πρώιμα γεγονότα της αναδίπλωσης.



Εικόνα 1.4 (αναπαράγεται άνευ αδειάς από R. Zerella et al., 1999).

Σχηματική αναπαράσταση της β-φουρκέτας, που σχηματίζει το N-τελικό άκρο της ουβικιτίνης στη φυσική δομή. Παρουσιάζονται οι δεσμοί υδρογόνου μεταξύ των βραχιόνων της φουρκέτας, καθώς και ο προσανατολισμός των πρωτονίων CaH

Έχουν πραγματοποιηθεί αρκετά πειράματα με πεπτιδικά τμήματα που σχηματίζουν α-έλικές και β-στροφές σε διάλυμα (Dyson et al., 1988, 1992. Scholtz & Baldwin, 1992), τα οποία έχουν αποδειχθεί πολύτιμα για συστηματικές μελέτες του ρόλου των επιμέρους καταλοίπων στη διασφάλιση της σταθεροποίησης τους. Στην περίπτωση των β-φύλλων έχουν πραγματοποιηθεί λιγότερες μελέτες, ωστόσο είναι εμφανές ότι ορισμένες β-φουρκέτες μπορούν να αναδιπλωθούν αυτόνομα, είτε σε υδατικό διάλυμα είτε

παρουσία οργανικού συνδιαλύτη, αποκτώντας δομές παρόμοιες με τη φυσική (Cox et al., 1993. Blanco et al., 1994a, 1994b. Searle et al., 1995, 1996). Σε μερικές περιπτώσεις έχει δειχθεί αναδίπλωση και σε de novo σχεδιασμένες φουρκέτες (Blanco et al., 1993. De Alba et al., 1996. Ramirez-Alvarado et al., 1996. Maynard et al., 1998).

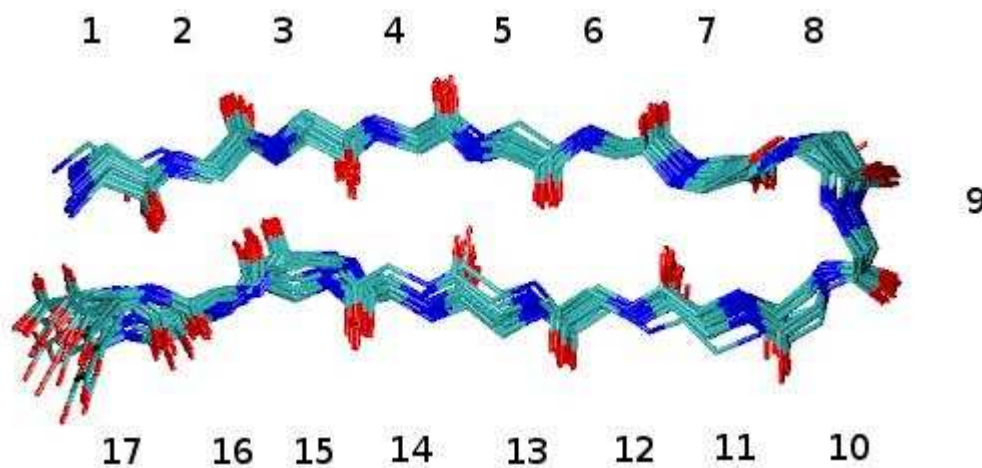
Η μελέτη του δεκαεπτά-πεπτιδίου προέκυψε ύστερα από μελέτη μεγαλύτερων τμημάτων του μορίου της ουβικιτίνης. Πιο συγκεκριμένα, σε ένα διάλυμα ουβικιτίνης, σε pH 2 και αναλογία μεθανόλη-νερό 3:1, βρέθηκε να υπάρχει μία μερικώς αναδιπλούμενη κατάσταση που ονομάστηκε “A-state”, στην κατάσταση αυτή τα περισσότερα από τα κατάλοιπα 1-35 (συμπεριλαμβανομένης της N-τελικής φουρκέτας και της κύριας α-έλικας) εμφανίζουν C_αH μετατοπίσεις που αντιστοιχούν στη φυσική δομή (Harding et al., 1991. Stockman et al., 1993). Ενώ το C-τελικό άκρο στην “A-state” αποτελείται από μη φυσική ελικοειδής δομή (Stockman et al., 1993). Προκειμένου να αξιολογηθεί η έκταση, στην οποία οι φυσικού-τύπου δομές του N-τελικού άκρου μπορεί να είναι αλληλοεξαρτώμενες για τη σταθερότητά τους μελετήθηκαν στις ίδιες συνθήκες δύο πεπτιδικά τμήματα που αντιστοιχούν στα κατάλοιπα 1-21 και 1-35. Και στις δύο περιπτώσεις, οι χημικές μετατοπίσεις των καταλοίπων της φουρκέτας ήταν ταυτόσημες με αυτές του ενσωματωμένου πεπτιδίου στην “A-state” , γεγονός που υποδηλώνει πως πρόκειται για ένα αυτόνομα αναδιπλούμενο στοιχείο σε αυτές τις συνθήκες (Cox et al., 1993). Ωστόσο, υπό αμιγώς υδατικές συνθήκες βρέθηκαν σημαντικά χαμηλότερα επίπεδα αναδίπλωσης, προφανώς επειδή η παρουσία αλκοόλης συμβάλει στη δημιουργία δεσμών υδρογόνου. Παρόλα αυτά, η διατήρηση αυτής της μικρής ποσότητας φυσικού-τύπου δομής στο νερό αποκτά ενδιαφέρον σε σχέση με την κατανόηση της αναδίπλωσης της πρωτεΐνης. Για το λόγο αυτό μελετήθηκε με περισσότερη λεπτομέρεια η συμπεριφορά της αλληλουχίας της φουρκέτας U(1-17) από την ομάδα της Zerella.

Η μελέτη του πεπτιδίου U(1-17) έδειξε πως κυρίαρχο ρόλο στη διαμόρφωση της φουρκέτας έχει η στροφή, καθώς τα άκρα της, παρόλο που φαίνεται να προτιμούν τη β-δομή, δεν εξασφαλίζουν ισχυρούς δεσμούς ζευγαρώματος. Το μόνο ζεύγος αμινοξέων που συνδέεται με δεσμούς υδρογόνου φαίνεται να είναι το Val5-Ile13, πέρα από αυτό δεν φαίνεται να υπάρχει κάποια άλλη ισχυρή ένδειξη, όπως δισουλφιδικοί δεσμοί ή φορτισμένα ζεύγη, που βοηθούν στη σταθεροποίηση της β-δομής (R. Zerella et al., 1999). Στη φυσική δομή της ουβικιτίνης ο τύπος της στροφής αυτής είναι G1 β-bulge. Ο τύπος αυτός κυριαρχεί στις β-φουρκέτες, όπως έχουν δείξει στατιστικές μελέτες και μελέτες σχεδιασμένων πεπτιδίων (Sibanda και Thornton, 1991, De Alba et al., 1996).

Παρόλο που τα δεδομένα των προαναφερθέντων πειραμάτων έδειξαν πως υπάρχει σημαντικός βαθμός αναδίπλωσης και η δομή που σχηματίζεται είναι φυσικού-τύπου, το πεπτίδιο U(1-17) δεν απέδωσε επαρκή δεδομένα NMR για να πραγματοποιηθεί πλήρης δομική ανάλυση. Τα αποτελέσματα της μελέτης του πεπτιδίου U(1-17) στάθηκαν αφορμή για περαιτέρω πειράματα, προκειμένου να εξακριβωθεί η αιτία της σταθερότητας της δομής αυτής. Σε προηγούμενες εργασίες, με άλλα μεταλλαγμένα πεπτίδια, προτάθηκε πως μεταλλάξεις στην αλληλουχία της στροφής μπορούν να ρυθμίσουν τη σταθερότητα της αναδιπλωμένης δομής του πεπτιδίου (Searle et al., 1995). Σε μια προσπάθεια να ερευνηθούν εκτενέστερα οι επιπτώσεις μιας μετάλλαξης στην στροφή μιας β-φουρκέτας, δημιουργήθηκε το πεπτίδιο U(1-17)T9D, το οποίο αποτελεί κοντινό ανάλογο του πεπτιδίου U(1-17). Η μεταλλαγή T9D αφορά την αντικατάσταση του ενάτου αμινοξέως από ασπαρτικό (Asp), αντί της θρεονίνης (Thr) που βρίσκεται στη φυσική αλυσίδα. Η δομή του μεταλλαγμένου πεπτιδίου, U(1-17)T9D, έχει καταχωρηθεί στην Protein Data Bank με τον κωδικό 1E0Q.

Το πεπτίδιο παρασκευάστηκε συνθετικά σε στερεά φάση συνεχούς ροής, χρησιμοποιώντας τη μέθοδο Fmoc-polyamide. Επιπλέον χρησιμοποιήθηκε NovaSyn KR ρητίνη σε παρασκευαστή NovaSyn Crystal. Το ακατέργαστο

πεπτιδίο διαλύθηκε σε 5% οξικό οξύ και καθαρίστηκε σε στήλη Waters Deltapak reverse-phase C18, για την έκλουση χρησιμοποιήθηκε 0.1% TFA/H₂O και 0.1% TFA/acetonitrile. Τα πειράματα NMR πραγματοποιήθηκαν μετά από διάλυση του πεπτιδίου σε διάλυμα αναλογίας 9H₂O:1D₂O. Οι συγκεντρώσεις των πεπτιδικών δειγμάτων είχαν εύρος 1.0-2.0 mM και το pH του διαλύματος είχε τιμή 3.8. Τα αποτελέσματα συγκεντρώθηκαν και οι δομές υπολογίστηκαν με τη χρήση του προγράμματος X-PLOR 3.8 (Brünger, 1992). Στην πρωτεϊνική βάση δεδομένων (PDB) καταχωρήθηκαν οι συντεταγμένες 27 δομών, όπως φαίνεται και στην εικόνα 1.5, για το πεπτιδίο U(1-17)T9D (κωδικός αναφοράς 1E0Q).



Εικόνα 1.5

Υπέρθωση του σκελετού (C, O, N άτομα) των 27 συγκλίνων δομών του πεπτιδίου U(1-17)T9D, που έχουν κατατεθεί στην πρωτεϊνική βάση δεδομένων. Ο χρωματισμός έγινε κατ' άτομο, με κόκκινο χρώμα εμφανίζονται τα άτομα του οξυγόνου, με μπλε σκούρο του αζώτου και με σιέλ του άνθρακος (τα άτομα του υδρογόνου δεν εμφανίζονται στην εικόνα). Κάθε αμινοξύ αριθμείται.

1.5 Δομική ανάλυση του πεπτιδίου 1E0Q

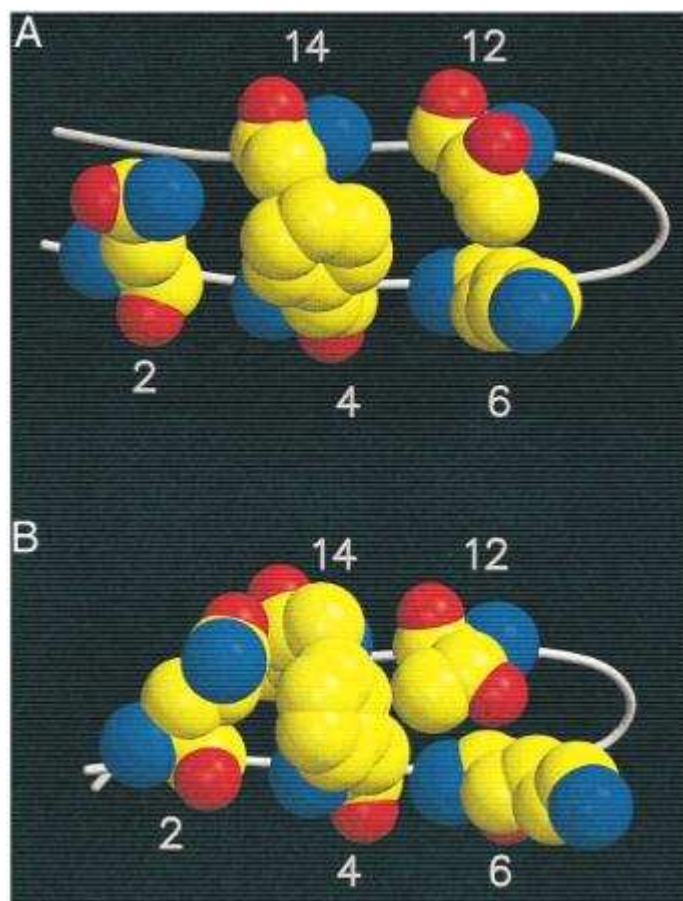
Τα αποτελέσματα από τα πειράματα H^1 NMR για το U(1-17)T9D έδωσαν φάσμα με μεγαλύτερη διασπορά σε σχέση με το U(1-17), καθώς το U(1-17)T9D είναι ευδιάλυτο στο νερό σε τιμές pH μικρότερες του 4.0. Επιπλέον υπάρχουν σαφείς ενδείξεις για ένα μεγάλο πληθυσμό β-φουρκέτας στο φάσμα του U(1-17)T9D.

Το πεπτίδιο U(1-17)T9D σχηματίζει δομή β-φουρκέτας αγρίου τύπου στο διάλυμα. Η δομή είναι καλά καθορισμένη, καθώς ο πεπτιδικός σκελετός (άτομα: N, Ca, C, O) έχει root mean square deviation (RMSD) 0.59 Å. Στα απομονωμένα άκρα (κατάλοιπα 1-7 και 11-17) το RMSD του πεπτιδικού σκελετού είναι 0.4 Å, ενώ το RMSD του πεπτιδικού σκελετού των καταλοίπων της στροφής (7-11) είναι 0.25 Å. Οι σχετικά μικρές τιμές RMSD δεν έρχονται σε αντίθεση με μια περισσότερο δυναμική δομή, σε σχέση με τη δομή της β-φουρκέτας σε μια πρωτεΐνη. Επίσης οι μεγαλύτερης έκτασης δεσμοί υδρογόνου που παρατηρήθηκαν στο πεπτίδιο U(1-17)T9D σε σχέση με το αντίστοιχο τμήμα της ουβικιτίνης, συνάδουν με μια περισσότερο δυναμική συμπεριφορά του U(1-17)T9D. Με τη δομή της β-φουρκέτας συμφωνούν και οι τιμές των φ/ψ γωνιών, οι οποίες δεν βρίσκονται σε απαγορευμένες περιοχές του Ramachandran plot. Οι αποστάσεις και η γεωμετρία των ατόμων του σκελετού συνάδουν με την ύπαρξη πέντε δεσμών υδρογόνου στο πεπτίδιο. Συνεπώς είναι πιθανόν το μοτίβο των δεσμών υδρογόνου του πεπτιδίου U(1-17)T9D να συμπίπτει με εκείνο του ενσωματωμένου πεπτιδίου στην ουβικιτίνη.

Στις 22 από τις 27 συνολικά δομές οι φ/ψ γωνίες των καταλοίπων, στις θέσεις $i+1$ και $i+2$ της στροφής (Leu8 and Asp9), ταιριάζουν περισσότερο με εκείνες που αναμένονται για την φυσικού-τύπου type I στροφή (Hutchinson et Thornton, 1994), παρά με εκείνες που χαρακτηρίζουν κάθε άλλο είδος β-στροφής. Ωστόσο η ψ ($i+2$), που αντιστοιχεί στο μεταλλαγμένο κατάλοιπο Asp9 αποκλίνει από την

ιδανική τιμή των 0° και βρίσκεται διαρκώς γύρω στις -38° . Στις υπόλοιπες 5 δομές οι φ/ψ γωνίες δεν βρίσκονται 100% στην πιο ευνοημένη περιοχή του Ramachandran plot.

Πέρα από τη διαμόρφωση των ατόμων του σκελετού του πεπτιδίου, έχει χαρακτηριστεί αρκετά καλά και η διαμόρφωση των πλευρικών αλυσίδων, κυρίως των καταλοίπων Ile3, Phe4, Val5, Lys6, Thr12, και Thr14. Άλλες πλευρικές ομάδες είναι λιγότερο καλά καθορισμένες, όπως των Thr7, Leu8, Ile13, και Leu15. Η πλευρική ομάδα του μεταλλαγμένου καταλοίπου, Asp9, φαίνεται να προσδίδει ασυνήθιστη σταθερότητα στο πεπτίδιο, πιθανόν από την αλληλεπίδραση με την Lys11. Δεν υπάρχουν όμως ισχυρές αποδείξεις για ιοντικές αλληλεπιδράσεις μεταξύ των πλευρικών αλυσίδων των δύο αυτών καταλοίπων. Οι πλευρικές αλυσίδες των καλά καθορισμένων αμινοξέων αλληλεπιδρούν τοπικά, έτσι ώστε να πακετάρονται εκατέρωθεν του σκελετού του πεπτιδίου. Τα κατάλοιπα Ile3 και Val5 φαίνεται να πακετάρονται μαζί, ταυτοχρόνως τα κατάλοιπα Ile13 και Leu15 αλληλεπιδρούν με κάποιο τρίτο αλλά και με τα κατάλοιπα Ile3 και Val5 μέσω δυνάμεων Wan der Waals. Τα δεδομένα δεν είναι επαρκή, έτσι ώστε να καθοριστούν απόλυτα οι αλληλεπιδράσεις των πλευρικών ομάδων. Από την άλλη επιφάνεια της φουρκέτας η μεθυλομάδα της Thr14 πακετάρεται απέναντι από τον δακτύλιο της Phe4 και η μεθυλομάδα της Thr12 δημιουργεί επαφές με τις ομάδες μεθυλενίου της πλευρικής αλυσίδας της Lys6 (Εικόνα 1.6).



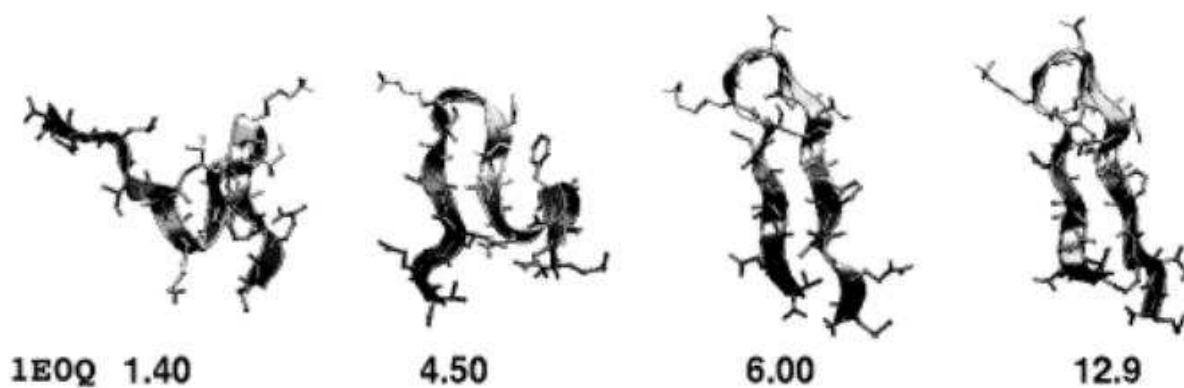
Εικόνα 1.6 (αναπαράγεται άνευ αδείας από R. Zerella et al. 2000)

Σχηματική απεικόνιση των αλληλεπιδράσεων που σχηματίζουν την υδρόφιλη επιφάνεια της β-φουρκέτας. A. Παρουσιάζεται η NMR δομή του πεπτιδίου U(1-17)T9D. Φαίνεται η αλληλεπίδραση του δακτυλίου της Phe4 με την πλευρική ομάδα της Thr14, καθώς και η αλληλεπίδραση των πλευρικών αλυσίδων των καταλοίπων Lys6 και Thr12. Σημειώνεται επίσης η απουσία αλληλεπίδρασης μεταξύ των καταλοίπων Phe4 και Gln2.

B. Παρουσιάζεται η κρυσταλλική δομή του ενσωματωμένου πεπτιδίου στην ουβικιτίνη, όπου οι αλληλεπιδράσεις των πλευρικών ομάδων είναι διαφορετικές. Στις αλληλεπιδράσεις συμμετάσχουν τα κατάλοιπα Phe4 με Thr12, Thr14, και Gln2.

1.6 Άλλες προσομοιώσεις με το πεπτίδιο 1E0Q

Οι προσομοιώσεις μοριακής δυναμικής, όπως έχει προαναφερθεί, αποτελούν ένα ισχυρό εργαλείο για την πρόβλεψη της δομής και τη μελέτη της αναδίπλωσης. Μια πρώτη προσπάθεια για την πρόβλεψη της δομής του πεπτιδίου 1E0Q έγινε από την επιστημονική ομάδα των Seokmin Shin και Youngshang Pak το 2003. Στην προαναφερθείσα μελέτη χρησιμοποιήθηκε το δυναμικό πεδίο CHARMM19 (all-atom based δυναμικό πεδίο) σε συνδυασμό με το μοντέλο ενυδάτωσης GB (Generalized Born implicit solvent model). Στις προσομοιώσεις τους τα N και C άκρα του πεπτιδίου επιδιορθώθηκαν σύμφωνα με την τυπική μορφή NH_3^+ και CO_2^- και η αναδίπλωση ξεκίνησε από πλήρως εκτεταμένο πεπτίδιο. Για τον έλεγχο της θερμοκρασίας χρησιμοποιήθηκε ο θερμοστάτης Berendsen και για τη ρύθμιση των αποστάσεων των δεσμών που σχηματίζονται από βαρέα άτομα και υδρογόνα χρησιμοποιήθηκε ο αλγόριθμος SHAKE. Ενώ δεν χρησιμοποιήθηκαν non-bond energy cutoffs για τον υπολογισμό της πλήρους ενέργειας ενυδάτωσης GB. Συνολικά ελήφθησαν έξι ανεξάρτητα τροχιακά σε θερμοκρασία 360K με χρονικό βήμα των 0.15 fs για 15.0 ns. Στα πέντε από τα έξι τροχιακά η δομή της β-φουρκέτας σχηματίστηκε εντός των χρονικών ορίων και μεταξύ 0.4 - 7.0 ns. Τα αποτελέσματα έδειξαν πως η προβλεπόμενη δομή με την ελάχιστη ενέργεια συμφωνεί απόλυτα με την πειραματικά προσδιορισμένη δομή από το NMR, καθώς το RMSD των δύο αυτών δομών είναι 1,36 Å. Στην εικόνα 1.7 παρουσιάζονται ορισμένα στιγμιότυπα από τα τροχιακά της προαναφερθείσας μελέτης (S. Shin και Y. Pak et. al, 2002).



Εικόνα 1.7 (αναπαράγεται άνευ αδείας από Seokmin Shin και Youngshang Pak, 2002)

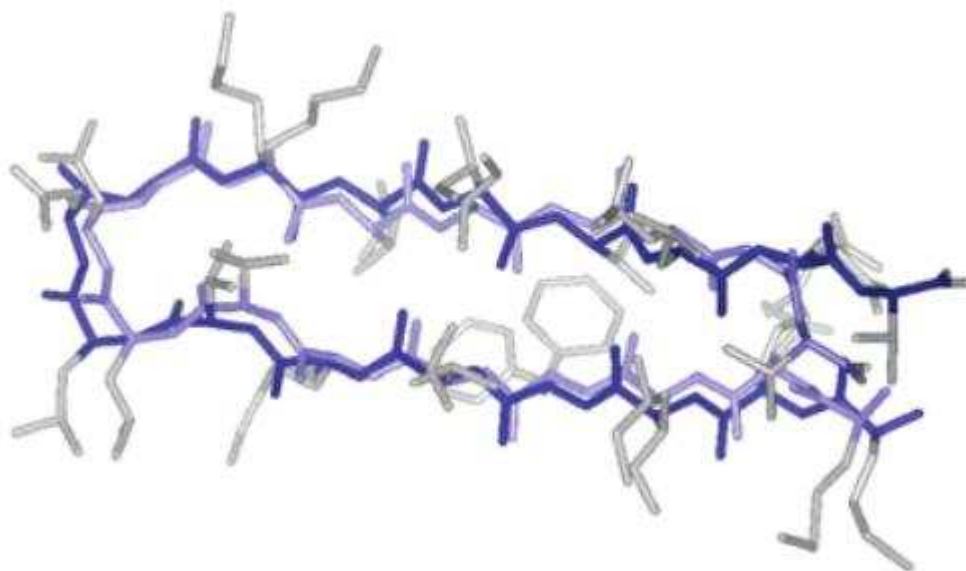
Παρουσιάζεται αντιπροσωπευτικό τροχιακό του πεπτιδίου 1E0Q. Οι τιμές αντιπροσωπεύουν το χρόνο της προσομοίωσης σε ns. Το C-τελικό άκρο βρίσκεται από την αριστερή πλευρά.

Ένα χρόνο αργότερα η ομάδα του W. L. Jorgensen πραγματοποίησε μια σειρά από οκτώ προσομοιώσεις Monte Carlo ξεκινώντας από πλήρως εκτεταμένη δομή του πεπτιδίου U(1-17)T9D, σε θερμοκρασία 303K. Το πρόγραμμα που χρησιμοποιήθηκε ήταν το MCPRO (W. L. Jorgensen et al., 2004, 2005). Υιοθετήθηκαν κανονικά επίπεδα πρωτονίωσης για pH 7.0 (αποπρωτονιωμένα καρβοξυλικά οξέα, πρωτονιωμένες αμίνες και γουανιδίνες) και χρησιμοποιήθηκε το δυναμικό πεδίο OPLS-AA (W. L. Jorgensen et al., 1996) για την αξιολόγηση της δυναμικής ενέργειας. Η ολική δυναμική ενέργεια υπολογίσθηκε χωρίς cutoffs για τις μη δεσμικές αλληλεπιδράσεις και με τιμή 1 για την διηλεκτρική σταθερά στις αλληλεπιδράσεις Coulomb. Όλες οι προσομοιώσεις οδήγησαν σε σταθερή δομή β-φουρκέτας. Επιπλέον σε όλες τις προσομοιώσεις παρατηρήθηκε μια ταχεία χαλάρωση της εκτεταμένης δομής σε μια πιο συμπαγή κατάσταση που χαρακτηρίζεται από ένα συνεστραμμένο πεπτιδικό σκελετό. Στην αναδιπλωμένη κατάσταση το RMSD του αναδιπλούμενου πεπτιδίου κυμαίνεται μεταξύ 2.5 – 5 Å σε σχέση με την πειραματικά προσδιορισμένη δομή.

Η δομή που έχει την χαμηλότερη ενέργεια είναι επίσης η πιο κοντινή, δομικά, στην πειραματικά προσδιορισμένη, καθώς έχει τους σωστούς δεσμούς υδρογόνου μεταξύ των ατόμων του σκελετού (Met1(O)-Val17(N), Lys11(O)-Thr7(N) και τα ζεύγη δεσμών μεταξύ Ile3-Leu15 και Val5-Ile13) και τον σωστό τύπο στροφής (Type I β -turn). Σε όλες τις προσομοιώσεις παρατηρήθηκε η δημιουργία στροφής μόνο στο κέντρο του πεπτιδίου, σημείο κλειδί για τον σχηματισμό της φουρκέτας, και επιπλέον δεν παρατηρήθηκε καμία ελικοειδής δομή. Σημειώνεται επίσης ότι φέρει ιδιαίτερο ενδιαφέρον ο ρόλος του καταλοίπου Gly10, του οποίου η ευελιξία ευθύνεται για τον σχηματισμό της στροφής στην περιοχή αυτή. Στην πρώτη προσομοίωση η Gly10 βρίσκεται στην τέταρτη θέση στην αλληλουχία της στροφής (η σειρά των καταλοίπων της στροφής είναι Thr7-Leu8-Asp9-Gly10), αλλά θα μπορούσε κάλλιστα το συγκεκριμένο κατάλοιπο να βρίσκεται στην τρίτη θέση λόγω της προτίμησης που έχει για αυτή τη θέση σε μία β -στροφή. Σε αυτή την περίπτωση οι δύο αντιπαρατιθέμενες αλυσίδες θα είχαν μετατοπιστεί κατά ένα κατάλοιπο, πράγμα που προτάθηκε από την ομάδα της R. Zerella, αλλά δεν παρατηρήθηκε στα πειράματα NMR. Στις επόμενες επτά προσομοιώσεις η χαμηλότερη ενέργεια των δομών είναι μεγαλύτερη κατά 15-20 Kcal σε σχέση με την χαμηλότερη ενέργεια που παρατηρήθηκε στην πρώτη προσομοίωση και η Gly10 βρίσκεται στην τρίτη θέση της στροφής (τα κατάλοιπα της στροφής σε αυτές τις δομές είναι Leu8-Asp9-Gly10-Lys11), πράγμα που σημαίνει πως πρόκειται για TypeII β -turn. Στην περίπτωση αυτή οι δεσμοί υδρογόνου μεταξύ των β -κλώνων δεν συμπίπτουν με τη φυσική δομή.

Στην πρώτη προσομοίωση που πραγματοποιήθηκε παρατηρήθηκαν επίσης και οι σωστές αλληλεπιδράσεις μεταξύ των πλευρικών αλυσίδων των καλά προσδιορισμένων καταλοίπων, σύμφωνα με τα πειράματα NMR, οι οποίες αναπαριστούν τη συμπαγή δομή που σχηματίζεται εκατέρωθεν του σκελετού του πεπτιδίου. Στις υπόλοιπες προσομοιώσεις, εξαιτίας της αναστροφής που παρουσιάζει ο ένας κλώνος της φουρκέτας, οι δομές των πλευρικών αλυσίδων

είναι εντελώς διαφορετικές και επιπλέον εμφανίζονται διάφορες μη-φυσικές επαφές μεταξύ των καταλοίπων. Στην εικόνα 1.8 παρουσιάζεται η δομή του πρώτου κύκλου MC στην οποία αντιστοιχεί η χαμηλότερη ενέργεια σε σχέση με την πειραματικά προσδιορισμένη δομή (W. L. Jorgensen et al., 2004).



Εικόνα 1.8 (αναπαράγεται άνευ αδείας από W. L. Jorgensen et al., 2004)

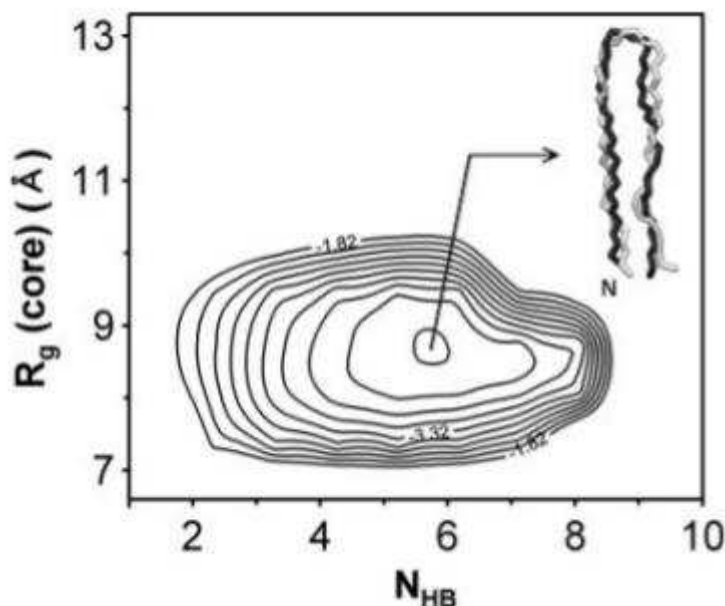
Υπέρθωση της NMR δομής και της δομής με τη χαμηλότερη ενέργεια από την πρώτη προσομοίωση MC. Υπάρχει αξιοσημείωτη ομοιότητα μεταξύ των δύο δομών. Ο τύπος στροφής είναι και στις δύο περιπτώσεις Type I β -turn. Ο αριθμός και η θέσεις των υδρογονοδεσμών είναι ταυτόσημες στις δύο δομές (δεν έχουν σημειωθεί οι δεσμοί υδρογόνου). Το αμινοτελικό άκρο βρίσκεται στον κάτω κλώνο.

Τέσσερα χρόνια αργότερα η ομάδα του S. Jang χρησιμοποιώντας Replica Exchange Molecular Dynamics (REMD) ως μέθοδο προσομοίωσης, ένα τροποποιημένο ενεργειακό μοντέλο, βασιζόμενο στο param99 δυναμικό πεδίο και το GB μοντέλο ενυδάτωσης πραγματοποίησαν εκτεταμένες ab initio μελέτες αναδίπλωσης σε καλά χαρακτηρισμένα πεπτίδια με α ή β κλώνους, μεταξύ των οποίων μελετήθηκε και το πεπτίδιο 1E0Q. Χρησιμοποιώντας το δυναμικό πεδίο

param99MOD3/GBSA (S. Jang et al., 2007) και ξεκινώντας από πλήρως εκτεταμένη δομή του πεπτιδίου πραγματοποιήθηκαν προσομοιώσεις REMD με 16-18 αντίγραφα (σε θερμοκρασία 250-650 K) για τα πεπτίδια C-peptide, EK-peptide, 1E0Q και GB1. Χρησιμοποιήθηκε βήμα των 2.0 fs με τον αλγόριθμο RATTLE (BJ Palmer, 1993, H.C Andersen, 1983) για τη ρύθμιση του μήκους δεσμού που περιέχει άτομο υδρογόνου. Η αποθήκευση κάθε τροχιακού έγινε κάθε 500 fs. Επιπλέον χρησιμοποιήθηκε cutoff των 24 Å για τις μη δεσμικές αλληλεπιδράσεις και για την GB ενυδάτωση. Όλες οι προσομοιώσεις εκτελέστηκαν με το πακέτο TINKER (J. W. Ponder, 2003).

Στην προαναφερθείσα μελέτη χρησιμοποιήθηκαν συνολικά 18 αντίγραφα του πεπτιδίου 1E0Q σε θερμοκρασίες 260-171 K και ο χρόνος προσομοίωσης (REMD simulation) του εκάστοτε αντιγράφου ήταν 10 ns. Στις προηγούμενες προσομοιώσεις που αναφέρθηκαν ελέγχθηκε η ικανότητα αναδίπλωσης του πεπτιδίου σε συγκεκριμένες συνθήκες, ενώ στην περίπτωση αυτή εξετάζεται με λεπτομέρεια η επιφάνεια της ελεύθερης ενέργειας του πεπτιδίου σε σχέση με το πειραματικά προσδιορισμένο πεπτίδιο από το NMR. Από τον χρόνο των 10 ns τα τελευταία 7 ns χρησιμοποιήθηκαν για τον υπολογισμό της ελεύθερης ενέργειας. Τα κατάλοιπα του υδρόφοβου πυρήνα του πεπτιδίου είναι τα MET1, ILE3, PHE4, VAL5, LEU8, ILE13, LEU15 και VAL17. Δημιουργήθηκε ένας χάρτης ελεύθερης ενέργειας (εικόνα 1.9) στην ίδια θερμοκρασία που καθορίστηκε η NMR δομή (275 K). Στους άξονες του σχεδιαγράμματος αναπαρίσταται η γυρεοσκοπική ακτίνα [R_g (core)] σε σχέση με τον αριθμό των φυσικών υδρογονοδεσμών (N_{HB}). Στο χάρτη παρουσιάζεται ένα εξαιρετικά απλό χαρακτηριστικό, μόνο μία κορυφή με κέντρο την φυσικού-τύπου δομή που έχει RMSD σκελετού 2.5-3.0 Å. Συνεπώς η επιφάνεια της ελεύθερης ενέργειας υποδηλώνει πως η αναδίπλωση του πεπτιδίου 1E0Q γίνεται σχεδόν χωρίς ενεργειακά φράγματα. Επιπλέον, η λεπτομερής ανάλυση του REMD τροχιακού πρότεινε την πιθανότητα η συνολική αναδίπλωση του πεπτιδίου να προχωρά με σχεδόν ταυτόχρονο σχηματισμό του υδρόφοβου πυρήνα και των

υδρογονοδεσμών μεταξύ των δύο β-κλώνων (S. Jang et al., 2007).



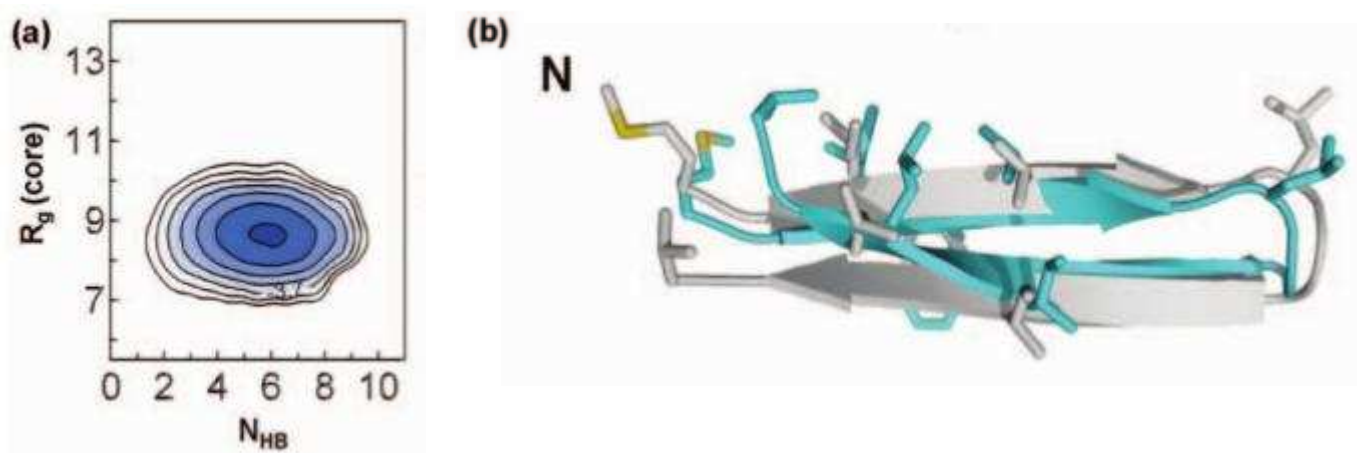
Εικόνα 1.9 (αναπαράγεται άνευ αδείας από S. Jang et al., 2007)

Ο χάρτης ελεύθερης ενέργειας του πεπτιδίου 1E0Q συναρτήσει της ακτίνας περιστροφής του υδρόφοβου πυρήνα [$R_g(\text{core})$] και του αριθμού των φυσικών υδρογονοδεσμών (N_{HB}) σε θερμοκρασία 275K. Οι ισοϋψείς σχεδιάζονται με βήμα 0.3 kcal/mol. Η δομή με την χαμηλότερη ελεύθερη ενέργεια είναι τοποθετημένη στο σημείο $N_{HB}, R_g(\text{core}) = (5.8, 8.7 \text{ \AA})$. επάνω δεξιά φαίνεται η υπέρθεση μεταξύ της αντιπροσωπευτικής δομής της πιο σταθερής κατάστασης (μαύρο χρώμα) με την NMR δομή (γκρι χρώμα).

Η ίδια επιστημονική ομάδα σε μια προσπάθεια βελτιστοποίησης του δυναμικού πεδίου param99MOD3 πραγματοποίησε πειράματα REMD προσομοιώσεων χρησιμοποιώντας και πάλι πεπτίδια μεταξύ των οποίων είναι και το 1E0Q και κατέληξε στην δημιουργία του δυναμικού πεδίου param99MOD5/GBSA.

Ο συνολικός χρόνος προσομοίωσης για το πεπτίδιο 1E0Q αυτή τη φορά ήταν 80ns, εκ των οποίων τα πρώτα 10 ns αφαιρέθηκαν από τους υπολογισμούς. Τα αποτελέσματα της δεύτερης μελέτης συμφωνούν με τα αποτελέσματα της πρώτης, καθώς ο χάρτης που περιγράφει την ελεύθερη ενέργεια [συναρτήσει

των N_{HB} και R_g (core)] είναι ίδιος (εικόνα 1.10). Ωστόσο σε σχέση με την NMR δομή το φυσικού-τύπου πεπτιδίο που προσδιορίστηκε στη συγκεκριμένη μελέτη είναι αρκετά παραμορφωμένο λόγω της τάσης για πιο συμπαγή πακετάρισμα των πλευρικών αλυσίδων των καταλοίπων Ile3, Val5, Ile13 και Leu15 (πιθανώς οφείλεται σε τεχνητό λάθος του μοντέλου ενυδάτωσης GB). Επιπροσθέτως, παρόλο που το κατάλοιπο Gly10 βρίσκεται στην τέταρτη θέση της αλληλουχίας της στροφής (TypeI β -turn), οι ϕ/ψ τιμές για το συγκεκριμένο κατάλοιπο δεν ήταν οι αναμενόμενες, συγκριτικά με την NMR δομή.



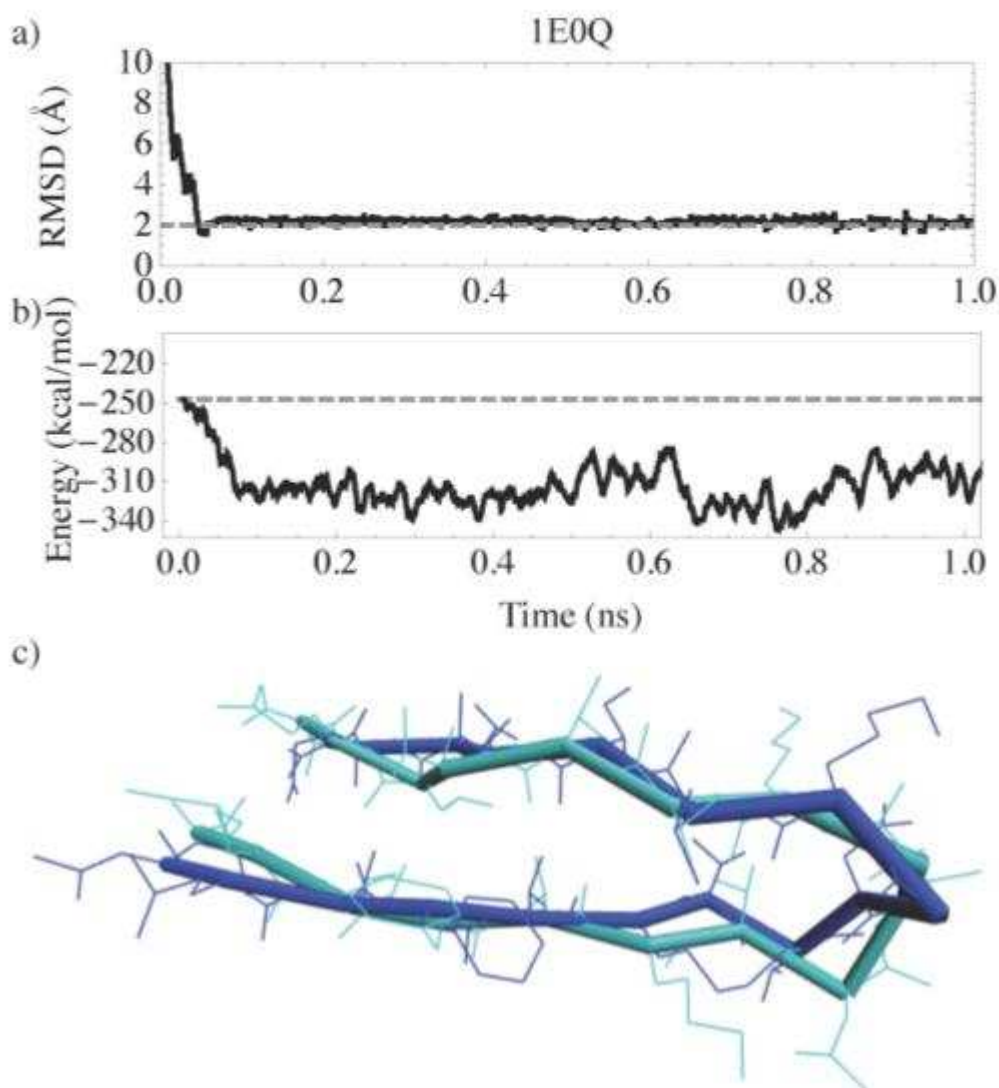
Εικόνα 1.10 (αναπαράγεται άνευ αδείας από S. Jang et al., 2008)

(a) Ο χάρτης ελεύθερης ενέργειας του πεπτιδίου 1E0Q συναρτήσει της ακτίνας περιστροφής του υδρόφοβου πυρήνα [R_g (core)] και του αριθμού των φυσικών υδρογονοδεσμών (N_{HB}) σε θερμοκρασία 275K. Η δομή με την χαμηλότερη ελεύθερη ενέργεια είναι τοποθετημένη στο σημείο N_{HB}, R_g (core) = (6, 8.5 Å). (b) Η πειραματικά προσδιορισμένη NMR δομή (γκρι χρώμα) και η δομή με την ελάχιστη ελεύθερη ενέργεια (κυανό χρώμα) αλληλεπικαλύπτονται. Το RMSD του σκελετού είναι 2.0 Å. Τα κατάλοιπα του υδρόφοβου πυρήνα υποδεικνύονται με ραβδοειδή αναπαράσταση.

Πιο πρόσφατα μια μελέτη που πραγματοποιήθηκε από την ομάδα της N. Vaidehi είχε ως επίκεντρο τη σύγκριση των μεθόδων constrained molecular dynamics και all-atom MD, για το κατά πόσο η αναζήτηση διαμορφώσεων χρησιμοποιώντας την πρώτη μέθοδο που είναι πιο βελτιωμένη και

εμπλουτισμένη με φυσικού-τύπου δομές σε σχέση με την δεύτερη (N. Vaidehi et al., 2011). Στα πειράματά τους συμπεριλήφθηκε και η μελέτη της αναδίπλωσης του πεπτιδίου 1E0Q.

Με το πεπτίδιο 1E0Q διενεργήθηκαν torsional REMD προσομοιώσεις με οκτώ αντίγραφα και χρόνο προσομοίωσης 20 ns ανά αντίγραφο, ξεκινώντας από μία εκτεταμένη δομή. Συγκριτικά με την NMR δομή το RMSD του σκελετού της καλύτερα αναδιπλούμενης δομής είναι 1.5 Å, ενώ το RMSD όλων των ατόμων είναι 2.8 Å. Τα αποτελέσματα αυτά είναι συμβατά με τα αποτελέσματα των προαναφερθέντων μελετών. Η ομοιότητα μεταξύ μιας αντιπροσωπευτικής δομής από τις προσομοιώσεις και της NMR δομής αναπαρίσταται και στην εικόνα 1.11c. Επιπλέον, αναφέρεται πως το μέσο RMSD του σκελετού του πεπτιδίου κατά τη διάρκεια της φάσης αναδίπλωσης της προσομοίωσης είναι 2.2 Å. Σε αυτό το σημείο αξίζει να σημειωθεί πως η περιοχή των β-κλώνων είχε χαμηλότερες διακυμάνσεις συγκριτικά με την περιοχή της β-στροφής, με μέσο RMSD σκελετού 2.0 Å για την περιοχή των κλώνων έναντι των 2.6 Å για την στροφή. Όπως φαίνεται και από τη σύγκριση των εικόνων 1.11a και 1.11b η δυναμική ενέργεια μειώνεται καθώς η δομή πλησιάζει στην στη φυσική. Αυτό σημαίνει πως η συνάρτηση της ενέργειας είναι επαρκής για την αναδίπλωση απλών πεπτιδίων με constrained MD (N. Vaidehi et al., 2011).



Εικόνα 1.11 (αναπαράγεται άνευ αδείας από N. Vaidehi et al., 2011)

(a) Γραφική παράσταση του RMSD σε σχέση με το χρόνο, για το αντίγραφο των REMD προσομοιώσεων που έδωσε τη φυσικού-τύπου δομή. (b) Γραφική παράσταση της δυναμικής ενέργειας σε σχέση με το χρόνο για το ίδιο πεπτίδιο. Η διακεκομμένη γραμμή αναπαριστά την αρχική δυναμική ενέργεια. (c) Σύγκριση της NMR δομής (κυανό χρώμα) με μια αντιπροσωπευτική δομή από την προσομοίωση.

Πέρα από το RMSD ταυτοποιήθηκε και ο σωστός τύπος στροφής, σύμφωνα με την NMR δομή, με την Gly10 στην τέταρτη θέση της στροφής (type I β -turn). Η μόνη απόκλιση από την NMR δομή ήταν ο δεσμός υδρογόνου μεταξύ Leu15(O)-

Ile3(N) , ενώ οι υπόλοιποι δεσμοί υδρογόνου συμφωνούν με τα προηγούμενα πειράματα. Το μήκος δεσμού της προσομοίωσης αφορά το μέσο μήκος κατά τα τελευταία 19ns. Στον πίνακα 1.1 γίνεται σύγκριση των δεσμών υδρογόνου του σκελετού μεταξύ της NMR δομής και των δομών από τις προσομοιώσεις. Η μεγάλη απόκλιση του μήκους δεσμού του ζεύγους Leu15(O)-Ile3(N) οφείλεται στη συστροφή των καταλοίπων που βρίσκονται στα άκρα των β-κλώνων.

H-bond	NMR-data [Å]	Simulation [Å]
Met1(O)-Val17(N)	4.0 ± 0.4	4.4 ± 1.2
Leu15(O)-Ile3(N)	3.4 ± 0.2	6.8 ± 0.4
Ile3(O)-Leu15(N)	3.13 ± 0.08	2.7 ± 0.2
Ile13(O)-Val5(N)	3.7 ± 0.4	3.0 ± 0.1
Val5(O)-Ile13(N)	3.8 ± 0.3	3.0 ± 0.2
Lys11(O)-Thr7(N)	3.6 ± 0.2	3.4 ± 0.3

Πίνακας 1.1 (αναπαράγεται άνευ αδείας από N. Vaidehi et al., 2011)

Σύγκριση των αποστάσεων των υδρογονοδεσμών του σκελετού της NMR δομής και με τις προσομοιώσεις για το πεπτιδίο 1E0Q.

Τέλος στην μελέτη της ομάδας της N. Vaidehi υποστηρίζεται πώς η αυξημένη σταθερότητα του πεπτιδίου 1E0Q συγκριτικά με την σταθερότητα του φυσικού πεπτιδίου U(1-17) οφείλεται στο ιοντικό ζεύγος (salt bridge) που δημιουργείται μεταξύ της NH_3^+ ομάδας της Lys11 και του COO^- του μεταλλαγμένου καταλοίπου Asp9. Η ανάλυση του torsional MD τροχιακού έδειξε πως η μέση N – O απόσταση μεταξύ των πλευρικών αλυσίδων των καταλοίπων Lys11 – Asp9 ήταν περίπου 2.8 Å (N. Vaidehi et al., 2011), πράγμα που υποστηρίζει τη δημιουργία ιοντικού ζεύγους.

REQUAMIO 2

“One of the strongest motives that lead men to art and science is escape from everyday life with its painful crudity and hopeless dreariness, from the fetters of one's own ever-shifting desires. A finely tempered nature longs to escape from the personal life into the world of objective perception and thought”

-Albert Einstein-

Εισαγωγή στις προσομοιώσεις μοριακής δυναμικής

Η μοριακή μοντελοποίηση (molecular modeling) συνδυάζει την επιστήμη και την τέχνη με σκοπό την απόδοση *in silico* των ατομικών και μοριακών αλληλεπιδράσεων που διέπουν την μικροσκοπική και μακροσκοπική συμπεριφορά των φυσικών συστημάτων. Η σύνδεση του μακροσκοπικού κόσμου με το μικροσκοπικό, που είναι και η ουσία του μοριακού μοντελισμού, επιτυγχάνεται μέσω της θεωρίας της στατιστικής μηχανικής (Molecular Dynamics Simulation - Michel Cuendet - EMBL 2008).

Τα μοντέλα που προκύπτουν χρησιμοποιούνται για τη μελέτη της μοριακής δομής και λειτουργίας. Προκειμένου να διεκπεραιωθεί η μοριακή μοντελοποίηση χρησιμοποιούνται υπολογιστικά εργαλεία και προγράμματα απεικόνισης. Στα υπολογιστικά εργαλεία συμπεριλαμβάνονται μέθοδοι όπως η εξ' αρχής (*ab initio*) και η ήμι-εμπειρική (*semi-empirical*) κβαντική μηχανική, η εμπειρική (μοριακή) μηχανική (*empirical molecular mechanics*), η μοριακή δυναμική (*molecular dynamics*), οι μέθοδοι ελεύθερης ενέργειας και ενυδάτωσης (*free energy and solvation methods*) και πολλές άλλες καθιερωμένες τεχνικές (T. Schlick, *Molecular modeling and simulation*, 2010).

2.1 Μοριακή μηχανική και δυναμικά πεδία

Παρόλο που πραγματοποιήθηκαν πολλές διαφορετικές προσεγγίσεις της κβαντικής μηχανικής, αρκετά προβλήματα εξακολουθούν να υφίστανται και δεν μπορούν να παρακαμφθούν ή να λυθούν. Συνεπώς η επιστημονική κοινότητα στράφηκε προς μια εναλλακτική λύση, τις μεθόδους της μοριακής μηχανικής.

Η μοριακή μηχανική συχνά αναφέρεται και ως μέθοδος δυναμικών πεδίων (force fields) ή δυναμικής ενέργειας (potential energy) και παραπέμπει στη χρήση απλών συναρτήσεων δυναμικού και ενέργειας με στόχο την αναπαράσταση μοριακών συστημάτων. Η μοριακή μηχανική χρησιμοποιείται ευρέως σε προσομοιώσεις μοριακής δυναμικής, αλλά έχει και άλλες εφαρμογές, όπως στις προσομοιώσεις Monte Carlo και στη βελτίωση της μοριακής δομής (molecular structure refinement) (J.A McCammon and S.A Adcock, 2006). Στην μοριακή μηχανική το μόριο αποτελείται από άτομα, που συνδέονται μεταξύ τους μέσω δεσμών, οι οποίοι λειτουργούν ως ελατήρια. Ως αποτέλεσμα των διαμοριακών και ενδομοριακών δυνάμεων το μόριο μπορεί να διαστέλλεται, να κάμπτεται και να περιστρέφεται σε σχέση με τους δεσμούς. Οι εσωτερικές δυνάμεις περιγράφονται, συνήθως, χρησιμοποιώντας απλές συναρτήσεις, για παράδειγμα οι δεσμικές αλληλεπιδράσεις περιγράφονται από το νόμο του Hooke, ενώ οι μη δεσμικές αλληλεπιδράσεις μπορεί να περιγράφονται από το δυναμικό Lennard-Jones. Χρησιμοποιώντας αυτά τα μοντέλα και τις εξισώσεις κίνησης του Νεύτωνα, οι προσομοιώσεις μοριακής δυναμικής μπορούν να παράγουν τη μεταβολή μιας δομής σε συνάρτηση με το χρόνο, δηλαδή τη θέση, την ταχύτητα και την επιτάχυνση όλων των ατόμων του μορίου κατά τη διάρκεια της προσομοίωσης (τροχιακό).

Για ένα σύστημα το οποίο διέπεται από τους νόμους του Νεύτωνα κάθε άτομο (i) αντιπροσωπεύεται από ένα σημείο, μάζας m , για το οποίο ισχύει η εξίσωση 2.4

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i \quad i = 1, 2, \dots, N \quad (2.4)$$

Στην εξίσωση 2.4 το πρώτο μέλος αντιπροσωπεύει το γινόμενο της μάζας επί το διάνυσμα της επιτάχυνσης και το δεύτερο μέλος αντιστοιχεί στο διάνυσμα της δύναμης που δρα στο άτομο i . Εάν στην εξίσωση αυτή το δεύτερο μέλος παραμένει σταθερό τότε η εξίσωση περιγράφει ένα απομονωμένο σύστημα, που αντιστοιχεί σε μικροκανονικά σύνολα (NVE ensembles) με σταθερό αριθμό ατόμων, όγκου και ενέργειας. Επειδή όμως η πρωτεϊνική αναδίπλωση συμβαίνει σε περιβάλλον που εμπλέκεται και η θερμοκρασία χρειαζόμαστε κανονικά σύνολα (NVT ensembles) με σταθερό αριθμό ατόμων όγκου και θερμοκρασίας ή ισόθερμα-ισοβαρή σύνολα (NPT ensembles) με σταθερό αριθμό ατόμων, πίεσης και θερμοκρασίας. Αυτό μπορεί να επιτευχθεί μεταβάλλοντας τις ταχύτητες έτσι ώστε να προσαρμοσθεί η κινητική ενέργεια του συστήματος στην επιθυμητή θερμοκρασία (H.A. Scheraga et al, 2008).

Τρεις είναι οι άξονες γύρω από τους οποίους περιστρέφεται η θεωρία της μοριακής μηχανικής. Ο πρώτος αφορά τη θερμοδυναμική, ο δεύτερος την προσθετικότητα και ο τρίτος τη μεταβίβαση.

Η υπόθεση της θερμοδυναμικής (Thermodynamic Hypothesis), δηλώνει πως, σε αντίθεση με τις περισσότερες βιολογικές διεργασίες, η αναδίπλωση των βιομορίων συμβαίνει αυτόματα, χωρίς τη βοήθεια κάποιου ενζύμου το οποίο βοηθά στην μείωση του ενεργειακού φραγμού έτσι ώστε να αποκτηθεί η τελική, φυσική, δομή. Το γεγονός αυτό υποδηλώνει πως μια ισχυρή θερμοδυναμική ενέργεια οδηγεί τα βιομόρια να αναδιπλωθούν από μία τυχαία (υψηλής ενέργειας) διαμόρφωση προς τη φυσική τους δομή (ελάχιστη ενέργεια). Μελέτες έχουν δείξει πως η πιθανότητα απόκτησης θερμοδυναμικά σταθερής δομής μιας τυχαίας διαμόρφωσης είναι αντιστρόφως ανάλογη της θερμοκρασίας (A.M. Gutin και E.I. Shakhnovich, 1990).

Η αρχή της προσθετικότητας υποδηλώνει ότι το άθροισμα των “δεσμικών” (bonded) και “μη δεσμικών” (non-bonded) ενεργειών αποτελεί την ολική δυναμική ενέργεια. Στις δεσμικές αλληλεπιδράσεις συμπεριλαμβάνονται οι ομοιοπολικοί δεσμοί (μεταξύ δύο ατόμων), οι γωνίες (μεταξύ τριών ατόμων), οι διέδρες γωνίες (μεταξύ τεσσάρων ατόμων) και οι μη κανονικές διέδρες γωνίες (impropers), ενώ στις μη δεσμικές αλληλεπιδράσεις συμπεριλαμβάνονται οι Van der Waals (Lennard-Jones potential) και οι ηλεκτροστατικές (Coulomb law). Η ολική δυναμική ενέργεια περιγράφεται από την εξίσωση 2.5, οι όροι της εξίσωσης αναφέρονται αντίστοιχα στα μήκη των δεσμών, στις γωνίες, στις διέδρες γωνίες, στις μη κανονικές διέδρες, στις αλληλεπιδράσεις Van der Waals και στις ηλεκτροστατικές αλληλεπιδράσεις .

$$\begin{aligned}
 U(\vec{R}) = & \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 \\
 & + \sum_{\text{dihedral}} \times K_\chi \left(1 + \cos(n\chi - \delta) \right) + \sum_{\text{impropers}} K_\varphi(\varphi - \varphi_0)^2 \\
 & + \sum_{\text{nonbond}} \left(\epsilon_{ij} \left[\left(\frac{R \min_{ij}}{r_{ij}} \right)^{12} - \left(\frac{R \min_{ij}}{r_{ij}} \right)^6 \right] \right) + \frac{q_i q_j}{\epsilon r_{ij}}
 \end{aligned} \tag{2.5}$$

Όπου: b είναι το μήκος δεσμού, θ είναι η γωνία μεταξύ δύο ομοιοπολικών δεσμών (που συνδέουν τρία άτομα), θ είναι η διέδρη γωνία ή γωνία περιστροφής, φ είναι η μη κανονική διέδρη γωνία και r_{ij} είναι η απόσταση μεταξύ των ατόμων i και j . Παράμετροι και όροι που διέπουν το δυναμικό πεδίο είναι η σταθερά δύναμης δεσμού (K_b) και η απόσταση εξισορρόπησης b_0 , η σταθερά δύναμης της ομοιοπολικής γωνίας (K_θ) και η γωνία εξισορρόπησης (θ_0), η σταθερά δύναμης της διέδρης γωνίας (K_χ), το πλήθος (n) και η γωνία φάσης (δ), η σταθερά δύναμης για τις μη κανονικές γωνίες (K_φ) και η μη κανονική γωνία εξισορρόπησης φ_0 . Αυτές οι παράμετροι αποτελούν τις ενδομοριακές παραμέτρους ενός δυναμικού πεδίου. Όσο αφορά τις εξωτερικές παραμέτρους για τις μη δεσμικές αλληλεπιδράσεις, αυτές είναι τα φορτία των ατόμων i και j (q_i και q_j αντίστοιχα), η σταθερά LJ well-depth ϵ_{ij} , και η ελάχιστη απόσταση

αλληλεπίδρασης των ατόμων $i, j \in R_{\min, ij}$, η οποία χρησιμοποιείται για τις αλληλεπιδράσεις Van der Waals. Η διηλεκτρική σταθερά ϵ συνήθως αντιστοιχεί στη μονάδα (κενό) στις προσομοιώσεις που συμπεριλαμβάνουν αλληλεπιδρόν διάλυμα (implicit solvent) (A.D Mackerell, 2004). Την αθροιστική αυτή ιδιότητα εκμεταλλεύονται τα διάφορα πρωτόκολλα προσομοιώσεων μοριακής δυναμικής, προκειμένου να υπολογισθούν με διαφορετική “βαρύτητα” οι δεσμικές και μη δεσμικές αλληλεπιδράσεις σε συνάρτηση με το χρόνο. Η προθετικότητα αξιοποιείται, επίσης, στο σχεδιασμό αλγορίθμων ελαχιστοποίησης.

Τέλος, η αρχή της μεταβίβασης δηλώνει ότι, εφόσον τα μόρια αποτελούνται από τις ίδιες χημικές υπομονάδες και το σύστημα βρίσκεται σε κανονικές συνθήκες (χωρίς την επιρροή ειδικών δυνάμεων), τα αποτελέσματα για τη δυναμική που απορρέουν από πειράματα σε αντιπροσωπευτικές δομές μπορούν να χρησιμοποιηθούν (μεταβιβασθούν) σε μεγάλα βιομόρια. Πρακτικά αυτό μπορεί να χρησιμοποιηθεί είτε μεταφέροντας τη γεωμετρία των δεσμών μεταξύ ατόμων, μικρών μορίων, χρησιμοποιώντας συναρτήσεις και λαμβάνοντας υπόψιν την εξάρτηση από το περιβάλλον (N.L Allinger et al., 1987), είτε χρησιμοποιώντας μεταβλητές για κάθε τύπο ατόμου (ο κάθε τύπος ατόμου δηλώνει το περιβάλλον του). Οι ατομικοί τύποι χρησιμοποιούνται στα δυναμικά πεδία και είναι αποθηκευμένοι σε αρχεία παραμετροποίησης και τοπολογίας (M. Karplus et al., 1998, T. Schlick, 2002).

Τα δυναμικά πεδία που συμπεριλαμβάνουν όλα τα άτομα ενός συστήματος στους υπολογισμούς ονομάζονται “all-atom force fields”. Εναλλακτικά τα υδρογόνα μπορούν να παραλειφθούν στα “united atom force fields”, εκτός από τα πολικά υδρογόνα τα οποία σχηματίζουν υδρογονικούς δεσμούς. Στις μέρες μας χρησιμοποιούνται και οι δύο τύποι δυναμικών πεδίων.

Ένα ακόμη σημαντικό στοιχείο στη μοριακή δυναμική είναι η ο ακριβής χειρισμός του διαλύτη (υδατικό περιβάλλον), ο χειρισμός αυτός

πραγματοποιείται με τη χρήση μεθόδων explicit ή implicit solvent. Η πρώτη μέθοδος έρχεται πιο κοντά με τις πραγματικές συνθήκες, αλλά είναι χρονοβόρα σε σχέση με τη δεύτερη μέθοδο. Τα μοντέλα νερού που χρησιμοποιούνται στις explicit μεθόδους προσομοιώσεων είναι τα TIP3P, TIP4P, SPC, extended SPC/E, και F3C98. Τα TIP3P μόρια νερού είναι τα ευρέως διαδεδομένα. Χρησιμοποιήθηκαν για την παραμετροποίηση των περισσοτέρων δυναμικών πεδίων, αλλά και για τις προσομοιώσεις των συστημάτων της παρούσας πτυχιακής εργασίας. Τα μοντέλα διαφέρουν στη γεωμετρία και σε κάποιες περιπτώσεις υπάρχουν διορθώσεις στην εσωτερική ενέργεια που αποδίδουν καλύτερα τη δομή και τις ιδιότητες διάχυσης (SPC/E). Ωστόσο τα TIP3P αναπαριστούν καλύτερα την ενέργεια και την πυκνότητα του νερού και έχουν συμβατότητα με το δυναμικό πεδίο Amber.

Οι μη δεσμικές αλληλεπιδράσεις είναι οι σημαντικότερες για τη συμπεριφορά ενός βιομορίου και ταυτόχρονα οι πιο απαιτητικές για το δυναμικό πεδίο. Για το λόγο αυτό έχουν αναπτυχθεί προσεγγίσεις που περιορίζουν τις αποστάσεις για τις ηλεκτροστατικές και vdW αλληλεπιδράσεις. Ωστόσο μεγαλύτερη ακρίβεια παρέχεται χρησιμοποιώντας όσο το δυνατόν περισσότερες μη δεσμικές αλληλεπιδράσεις. Όσο αφορά τις vdW αλληλεπιδράσεις, για να ληφθεί υπόψη η επίδραση τους, θεωρείται ως ομοιογενές στιδίοτε πέρα από την απόσταση αποκοπής. Για τις ηλεκτροστατικές αλληλεπιδράσεις έχει αναπτυχθεί η μέθοδος particle mesh Ewald (PME). Για την εφαρμογή αυτής της μεθόδου απαιτούνται περιοδικές οριακές συνθήκες επαρκούς μεγέθους, για να αποφευχθούν φαινόμενα αλληλεπίδρασης μεταξύ της πρωτεΐνης και του ειδώλου της. Σε περιπτώσεις μακρομορίων που αυτό είναι ανέφικτο χρησιμοποιούνται στοχαστικές οριακές συνθέσεις και όχι περιοδικές. Σε αυτή την προσέγγιση το μόριο περιβάλλεται από μια σφαίρα νερού και είναι περιτριγυρισμένο από ένα δυναμικό το οποίο διατηρεί την πυκνότητα του συστήματος. Σε ορισμένες περιπτώσεις χρησιμοποιούνται Langevin dynamics στα μόρια του νερού για τη συμπλήρωση του πεδίου αντίδρασης (A.D. Mackerell, 2004).

Με τον όρο δυναμικά πεδία αναφερόμαστε σε όλες αυτές τις συναρτήσεις και τις αντίστοιχες παραμέτρους, που χρησιμοποιούνται για την περιγραφή της κίνησης των ατόμων ενός μορίου. Τα τελευταία χρόνια υπάρχει μια ταχύτατη ανάπτυξη των υπολογιστικών δυνατοτήτων. Αυτό έχει συμβάλει καθοριστικά στην ανάπτυξη νέων δυναμικών πεδίων, τα οποία αποδίδουν αποτελεσματικότερα τα μονοπάτια που ακολουθούν τα βιομόρια κατά τη διαδικασία της αναδίπλωσης τους και συμφωνούν με τα πειραματικά αποτελέσματα. Παρόλο που τα περισσότερα νέα δυναμικά πεδία αποτελούν προσαυξήσεις προηγούμενων, η ταχύτατη αυτή ανάπτυξη αποτελεί ένα ισχυρό εργαλείο, που χρησιμοποιείται από τις προσομοιώσεις μοριακής δυναμικής, έτσι ώστε να παραχθεί λεπτομερώς η ατομική περιγραφή των μονοπατιών της αναδίπλωσης (D.E Shaw et al, 2011).

Σε αυτό το σημείο ακολουθεί μια μικρή ιστορική αναδρομή σχετικά με την εξέλιξη των δυναμικών πεδίων. Τα δυναμικά πεδία εμφανίστηκαν ως εναλλακτική λύση της κβαντικής μηχανικής αρκετό καιρό πριν δημιουργηθούν οι προσομοιώσεις μοριακής δυναμικής. Στη δεκαετία του '70 έγιναν οι πρώτες προσπάθειες ανάπτυξης συναρτήσεων δυναμικής ενέργειας για τον κλάδο της οργανικής χημείας που αποτέλεσαν βασικά σημεία στην ανάπτυξη των δυναμικών πεδίων. Τα δυναμικά ECEPP (Moman et al., 1975) καθώς και το CFF (Lifson και Warshel, 1968) ήταν από τα πρώτα που αναπτύχθηκαν. Επειδή δεν μπορούμε να αναφέρουμε όλα τα δυναμικά πεδία που έχουν αναπτυχθεί μέχρι σήμερα, θα εστιάσουμε την προσοχή στο AMBER (Assisted Model Building with Energy Refinement) (Weiner και Kollman, 1981) το οποίο χρησιμοποιείται για τη μελέτη της πρωτεϊνικής αναδίπλωσης.

Η πρώτη προσπάθεια έγινε στις αρχές της δεκαετίας του '80 από τον Kollman και τους συνεργάτες του. Τα φορτία που χρησιμοποιήθηκαν προήλθαν από υπολογισμούς κβαντικής χημείας (Hartree- Fock STO-3G), οι όροι για τις Van der Waals αλληλεπιδράσεις προήλθαν από προηγούμενα πειράματα προσομοιώσεων υγρής φάσης (Jorgensen, 1981) και πειράματα σε κρυστάλλους (Hagler et al., 1974), οι σταθερές για τις δυνάμεις, τους δεσμούς και τις γωνίες

προήλθαν, επίσης, από προηγούμενα πειράματα σε κρυστάλλους και τέλος οι σταθερές για τις περιστροφικές κινήσεις προήλθαν από υπολογισμούς κβαντικής χημείας. Το πρώτο δυναμικό πεδίο είναι μια “polar hydrogen only” προσέγγιση γιατί δεν περιλαμβάνει όλα τα άτομα στους υπολογισμούς. Το 1986 δημοσιεύθηκε μια επέκταση του πρώτου δυναμικού πεδίου που συμπεριλάμβανε όλα τα άτομα (Weiner et al., 1986). Στις αρχές της δεκαετίας του '90 ο Kollman και οι συνεργάτες του ανέπτυξαν ένα νέο δυναμικό πεδίο, γνωστό ως ff94 (Cornell et al., 1995). Σε αυτή την έκδοση πραγματοποιήθηκαν βελτιώσεις στις παραμέτρους αλλά και στον αλγόριθμο από τον οποίο προέρχονται οι παράμετροι αυτοί, με απώτερο σκοπό την προσομοίωση και άλλων μορίων εκτός των πρωτεϊνών. Ο σκοπός αυτός επετεύχθη με τη δημιουργία του προγράμματος antechamber, περίπου μια δεκαετία αργότερα. Επειδή όλες οι προηγούμενες προσομοιώσεις γινόταν σε αέρια φάση, ο επόμενος στόχος ήταν η δημιουργία προσομοιώσεων σε υγρή φάση, έτσι ώστε να υπάρχουν και αλληλεπιδράσεις με το διάλυμα. Προκειμένου να επιτευχθεί αυτός ο στόχος έπρεπε να προσαρμοστούν οι παράμετροι για τα φορτία και για της αλληλεπιδράσεις Van der Waals έτσι ώστε τα αποτελέσματα να πλησιάζουν περισσότερο στα πειραματικά δεδομένα. Ένα ακόμη σημείο στο οποίο δόθηκε έμφαση, για τη βελτίωση του ff94 δυναμικού πεδίου, ήταν οι παράμετροι για τις ϕ και ψ γωνίες του σκελετού, οι οποίες παίζουν καθοριστικό ρόλο στη διαμόρφωση κάθε αμινοξέος και επηρεάζουν την ενέργεια για τις έλικες, τα φύλλα και τις στροφές στις πρωτεΐνες. Οι επόμενες εκδόσεις ήταν το ff96 (Kollman et al., 1997) και το ff99 (Wang et al., 2000), στις οποίες η δυναμική των ϕ και ψ γωνιών ταίριαζε με τα αποτελέσματα ενεργειών από πειράματα κβαντικής μηχανικής σε τετραπεπίδια και διπεπίδια. Από το 2000 υπάρχει η δυνατότητα δοκιμής, των δυναμικών πεδίων σε μικρά πεπίδια, και σύγκρισης με τα πειραματικά δεδομένα από τεχνικές όπως το NMR (Nuclear Magnetic Resonance) και ο κυκλικός διχρωισμός (J.W. Ponder και D.A. Case, 2003). Οι προσεγγίσεις που χρησιμοποιήθηκαν, για τις ϕ και ψ γωνίες, στο δυναμικό πεδίο ff99 οδήγησαν σε λάθος δομικές προτιμήσεις στα κατάλοιπα γλυκίνης, επιπλέον παρατηρήθηκε

υπερ-σταθεροποίηση των α -ελίκων. Από την άλλη, οι αλλαγές που παρουσιάστηκαν στο δυναμικό πεδίο ff96 οδήγησαν σε προτίμηση των β -φύλλων. Συνεπώς, ήταν αναγκαία η δημιουργία ενός δυναμικού πεδίου το οποίο δεν θα φέρει προτίμηση σε μια διαμόρφωση. Προκειμένου να ξεπεραστούν τα προβλήματα αυτά χρησιμοποιήθηκαν παράμετροι για τις επιπλέον δίεδρες γωνίες (C^β) που περιέχονται στα αμινοξικά κατάλοιπα που φέρουν πλευρικές ομάδες (πλην της γλυκίνης) και επηρεάζουν την περιστροφή των ϕ/ψ γωνιών του C^α . Οι δίεδρες αυτές γωνίες ονομάστηκαν ϕ' και ψ' ($\phi' = C-N-C^\alpha-C^\beta$ και $\psi' = C^\beta-C^\alpha-C-N$) και η νέα εκδοχή του ff99 ονομάστηκε ff99SB (C. Simmerling et al, 2006). Μία διαφορετική τροποποίηση των ff94/ff99, η οποία περιελάμβανε μια τελείως διαφορετική αντίληψη για την προέλευση των επιμέρους ατομικών φορτίων αποτέλεσε νέο δυναμικό πεδίο, το οποίο ονομάστηκε ff03 (Duan et al, 2003). Μία τελευταία εκδοχή του Amber είναι το ff99SB-ILDN, μέσω της οποίας βελτιώνεται ακόμη περισσότερο η ακρίβεια των αποτελεσμάτων. Σε αυτή την εκδοχή έγινε βελτιστοποίηση της δυναμικής της χ_1 γωνίας συστροφής (χ_1 torsion angle) των καταλοίπων Ile, Leu, Asp, Asn. Πραγματοποιήθηκαν συγκριτικές μελέτες ανάμεσα στα αποτελέσματα που έδινε το δυναμικό πεδίο ff99SB για τις χ_1 torsions και τα πειραματικά δεδομένα της PDB, και παρατηρήθηκε πως στα αμινοξέα ισολευκίνη, λευκίνη, ασπαρτικό και ασπαραγίνη υπάρχει σημαντική απόκλιση, αυτό έδειξε πως το ff99SB δεν χειρίζεται σωστά τις πλευρικές ομάδες αυτών των αμινοξέων. Παρατηρήθηκαν νέα δυναμικά για τις χ_1 torsions μέσω της προσαρμογής των παραμέτρων του δυναμικού πεδίου σε κβαντικό επίπεδο (ab initio quantum with DF-LMP2 calculations). Τέλος επιβεβαιώθηκε το τροποποιημένο δυναμικό πεδίο χρησιμοποιώντας δεδομένα NMR τα οποία επιβεβαίωσαν τις διαμορφώσεις των πλευρικών αλυσίδων (D.E. Shaw et al, 2010).

Εκτός από το AMBER, δυναμικά πεδία τα οποία χρησιμοποιούνται ευρέως για τη μελέτη της πρωτεϊνικής αναδίπλωσης είναι το OPLS/AA και το CHARMM22. Το OPLS και το CHARMM χρησιμοποιούν δεδομένα από το HF/ 6-31G* για τα ατομικά φορτία . Το CHARMM και το AMBER είναι συμβατά με το TIP3 μοντέλο νερού, ενώ το OPLS είναι συμβατό με τα TIP3P, TIP4P και SPC μοντέλα. Μεταξύ των τριών αυτών δυναμικών πεδίων υπάρχουν διαφορές στην κατανομή των φορτίων, πράγμα που οδηγεί σε διαφορές στην ισορροπία των τοπικών αλληλεπιδράσεων. Επιπλέον διαφορές βρίσκονται και στον υπολογισμό των ϕ και ψ γωνιών, το CHARMM χρησιμοποιεί τη μέθοδο CMAP για να διορθωθούν τα προβλήματα με το σκελετό των πρωτεϊνών, ενώ το AMBER χρησιμοποιεί τις προσεγγίσεις που αναφέρθηκαν προηγουμένως. Μελέτες έχουν δείξει πως το OPLS συμφωνεί περισσότερο με τα πειραματικά δεδομένα, ενώ τα αποτελέσματα από το AMBER και το CHARMM είναι παρόμοια (A.D. Mackerell, 2004).

2.2 Προσομοιώσεις μοριακής δυναμικής

Οι προσομοιώσεις μοριακής δυναμικής είναι μέθοδοι που χρησιμοποιούνται ευρέως και μας επιτρέπουν να λαμβάνουμε πληροφορίες σχετικά με την χρονική εξέλιξη των διαμορφώσεων των βιολογικών μορίων καθώς και πληροφορίες σχετικά με την κινητική και τη θερμοδυναμική. Παρέχουν μεγάλη ακρίβεια σχετικά με την εξέλιξη της θέσης κάθε σωματιδίου σε συνάρτηση με το χρόνο και μπορούν να χρησιμοποιηθούν για την ποσοτικοποίηση των ιδιοτήτων ενός συστήματος σε χρονική κλίμακα που δεν μπορεί να προσεγγιστεί με καμία άλλη μέθοδο.

Το 1957 έγινε η πρώτη προσομοίωση μοριακής δυναμικής (MD) χρησιμοποιώντας το αποκαλούμενο *hard-sphere model*, στο οποίο τα άτομα αλληλεπιδρούσαν μέσω συγκρούσεων (B.J Alder και T.E Wainwright, 1957). Στη δεκαετία του '70 πραγματοποιήθηκε η πρώτη MD προσομοίωση πρωτεΐνης που αφορούσε την ΒΡΤΙ, η προσομοίωση περιλάμβανε περίπου 500 άτομα και είχε διάρκεια μόλις 9,2 ns (JA McCammon et al., 1977). Η πρώτη προσομοίωση που είχε διάρκεια ενός μ s καταγράφηκε το 1998 (P.A Kollman et Y. Duan, 1998). Σήμερα πραγματοποιούνται προσομοιώσεις πολύ μεγαλύτερης διάρκειας που συμπεριλαμβάνουν 10^4 - 10^6 άτομα, ενώ οι δημοσιεύσεις που συμπεριλαμβάνουν προσομοιώσεις διαδέχονται η μια την άλλη με ταχύτατους ρυθμούς (J.A. McCammon et al, 2006, R.C Walker et al., 2013, D.E Shaw et al., 2013).

Τα δυναμικά πεδία Amber, CHARMM και GROMOS αναπτύχθηκαν αρχικά στα πλαίσια των λογισμικών μοριακής μηχανικής και δυναμικής Amber, CHARMM και GROMOS, ενώ το OPLS-AA αναπτύχθηκε στα πλαίσια των λογισμικών BOSS και MCPPro. Ωστόσο τα δυναμικά πεδία είναι συμβατά με ποικίλα πακέτα λογισμικών. Για ακαδημαϊκούς σκοπούς υπάρχουν τα παρακάτω λογισμικά που χρησιμοποιούνται για προσομοιώσεις μοριακής δυναμικής:

Amber (<http://amber.scripps.edu>),
CHARMM (<http://www.charmm.org>),
GROMACS (<http://www.gromacs.org>),
NAMD (<http://www.ks.uiuc.edu/Research/namd>),
Tinker (<http://dasher.wustl.edu/tinker>) και
Desmond (http://www.deshawresearch.com/resources_desmond.html).

MEMO 3

"A learning experience is one of those things that says, 'You know that thing you just did?'"

Don't do that."

-Douglas Adams-



Το πεπτίδιο από τη σκοπιά των προσομοιώσεων μοριακής δυναμικής

Σκοπός αυτού του κεφαλαίου είναι να γίνει η παρουσίαση των υπολογιστικών συστημάτων που χρησιμοποιήθηκαν (τεχνικά χαρακτηριστικά), καθώς και η βήμα προς βήμα διαδικασία που ακολουθήθηκε για την πραγματοποίηση των προσομοιώσεων. Πραγματοποιήθηκαν τρεις προσομοιώσεις μοριακής δυναμικής, στις οποίες το υπό μελέτη πεπτίδιο έχει την αλληλουχία :

MQIFVKTLDGKTITLEV

Η πειραματικά προσδιορισμένη δομή της παραπάνω αλληλουχίας είναι καταχωρημένη στην Protein Data Bank (PDB) με τον κωδικό “1E0Q” (R. Zerella et al., 2000). Η διαφορά των δύο πρώτων προσομοιώσεων έγκειται στο δυναμικό πεδίο (AMBER 99SB και AMBER 99SB ILDN), ενώ στην τρίτη προσομοίωση το πεπτίδιο είναι πρωτονιομένο και χρησιμοποιείται το δυναμικό πεδίο AMBER 99SB ILDN. Εν συνεχεία οι προσομοιώσεις θα αναφέρονται με την ακόλουθη ονομασία :

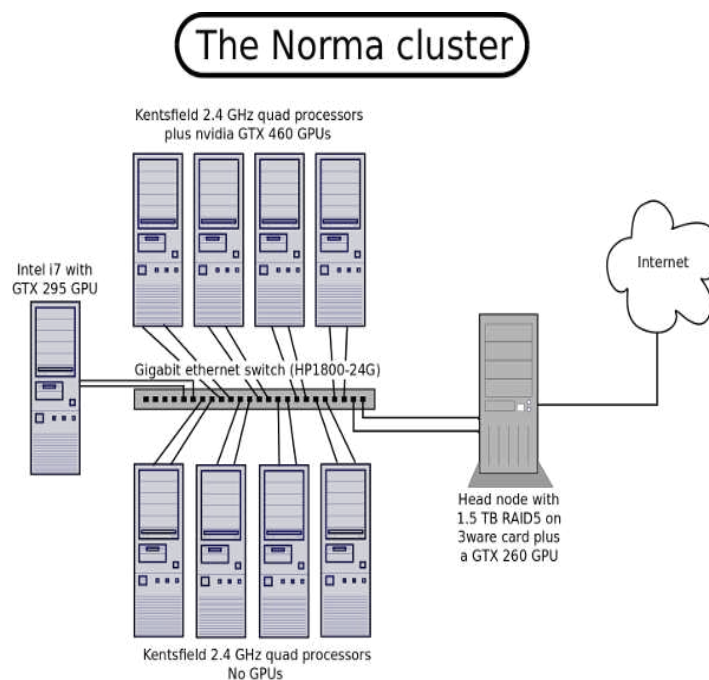
“99SB” , για την πρώτη προσομοίωση στην οποία χρησιμοποιήθηκε το AMBER 99SB δυναμικό πεδίο,

“ILDN” , για την δεύτερη προσομοίωση στην οποία χρησιμοποιήθηκε το AMBER 99SB ILDN δυναμικό πεδίο και

“protonated” , για την τρίτη προσομοίωση στην οποία χρησιμοποιήθηκε το ίδιο δυναμικό πεδίο με τη δεύτερη και το πεπτίδιο βρίσκεται σε πρωτονιομένη μορφή.

3.1 Τεχνικά χαρακτηριστικά υπολογιστικών συστημάτων

Οι προσομοιώσεις πραγματοποιήθηκαν σε μία συστοιχία υπολογιστών τύπου Beowulf (The Norma computing cluster). Η Norma περιλαμβάνει συνολικά 40 CPU (central processing unit) πυρήνες, συνολική φυσική μνήμη 46 Gbytes και 6 GPDPUs (General-purpose computing on graphics processing units) κατανεμημένα σε 10 κόμβους (εικόνα 3.1). Οι οκτώ κόμβοι (πλην του σκληρού δίσκου) προσφέρουν, έκαστος, τέσσερις πυρήνες, τέσσερα Gbytes φυσικής μνήμης και δύο (gigabit) διασυνδέσεις δικτύου. Ο ένατος κόμβος (Intel i7 965 extreme) προσφέρει έξι Gbytes φυσικής μνήμης και μία κάρτα γραφικών με δυνατότητα αύξησης της επεξεργαστικής ισχύος μέσω CUDA (Compute Unified Device Architecture) GXT-295. Οι κόμβοι συνδέονται μέσω ενός ειδικού ελεγκτή δικτύου HP ProCurve 1800-24G Gigabit Ethernet switch. Ο κύριος κόμβος παρέχει τέσσερις πυρήνες οκτώ Gbytes φυσικής μνήμης, 1,5 Tbytes αποθηκευτικής μνήμης, τρεις (gigabit) διασυνδέσεις δικτύου, και μια NVIDIAGTX-260 GPU (graphics processing unit).



Εικόνα 3.1 Σχηματική αναπαράσταση του υπολογιστικού συστήματος του εργαστηρίου (<http://utopia.duth.gr/~glykos/norma.html>)

Το πρόγραμμα που χρησιμοποιήθηκε προκειμένου να εκτελεστούν οι προσομοιώσεις στη Norma είναι το NAMD (J.C. Phillips et al, 2005), το οποίο χρησιμοποιείται σε συνδυασμό με το πρόγραμμα VMD (Visual Molecular Dynamics) (W. Humphrey et al, 1996) για την εκκίνηση των προσομοιώσεων καθώς και για την ανάλυση των παραγόμενων τροχιακών και είναι συμβατό με τις παραμέτρους και τις μορφές αρχείων των δυναμικών πεδίων AMBER και CHARMM. Επιπλέον για την οπτικοποίηση των τροχιακών που παράχθηκαν χρησιμοποιήθηκε το πρόγραμμα CARMA (N.M Glykos, 2006). Ο τρόπος με τον οποίο χρησιμοποιήθηκαν τα προαναφερθέντα προγράμματα για την προετοιμασία των συστημάτων αναλύεται στις παρακάτω παραγράφους.

Ο φυσικός χρόνος που απαιτήθηκε για να ολοκληρωθούν οι προσομοιώσεις διαφέρει σε κάθε περίπτωση. Η προσομοίωση “99SB” έχει διάρκεια 4,02808 μs, παράχθηκαν 5.035.100 διαμορφώσεις (frames) και πραγματοποιήθηκε σε 154 ημέρες (φυσικός χρόνος). Η προσομοίωση “ILDN” έχει διάρκεια 3,00682 μs,

παράχθηκαν 3.758.525 διαμορφώσεις (frames) και πραγματοποιήθηκε σε 125 ημέρες (φυσικός χρόνος). Τέλος, η προσομοίωση “protonated” έχει διάρκεια 2,97184μs, παράχθηκαν 3.714.800 διαμορφώσεις και ολοκληρώθηκε σε 116 ημέρες (φυσικός χρόνος). Ο παρακάτω συγκεντρωτικός πίνακας (πίνακας 3.1) παρουσιάζει τα πειράματα που διεξήχθησαν και τα χαρακτηριστικά του εκάστοτε συστήματος.

	“99SB”	“ILDN”	“PROTONATED”
Διάρκεια προσομοίωσης (μs)	4.03	3.01	2.97
Φυσικός χρόνος (days)	154	125	116
ns/day	26.16	24.05	25.62
Συνολικός αριθμός ατόμων	5616	5616	5596
Αριθμός ατόμων πεπτιδίου	285	285	287
Αριθμός ατόμων νερού	5331 (1777 μόρια H ₂ O)	5331 (1777 μόρια H ₂ O)	5307 (1769 μόρια H ₂ O)
Ιόντα	-	-	2* Cl ⁻
Θερμοκρασία (K)	320	320	320
Διαστάσεις Box (A)	40,161 x 40,161 x 40,161	40,161 x 40,161 x 40,161	40,161 x 40,161 x 40,161

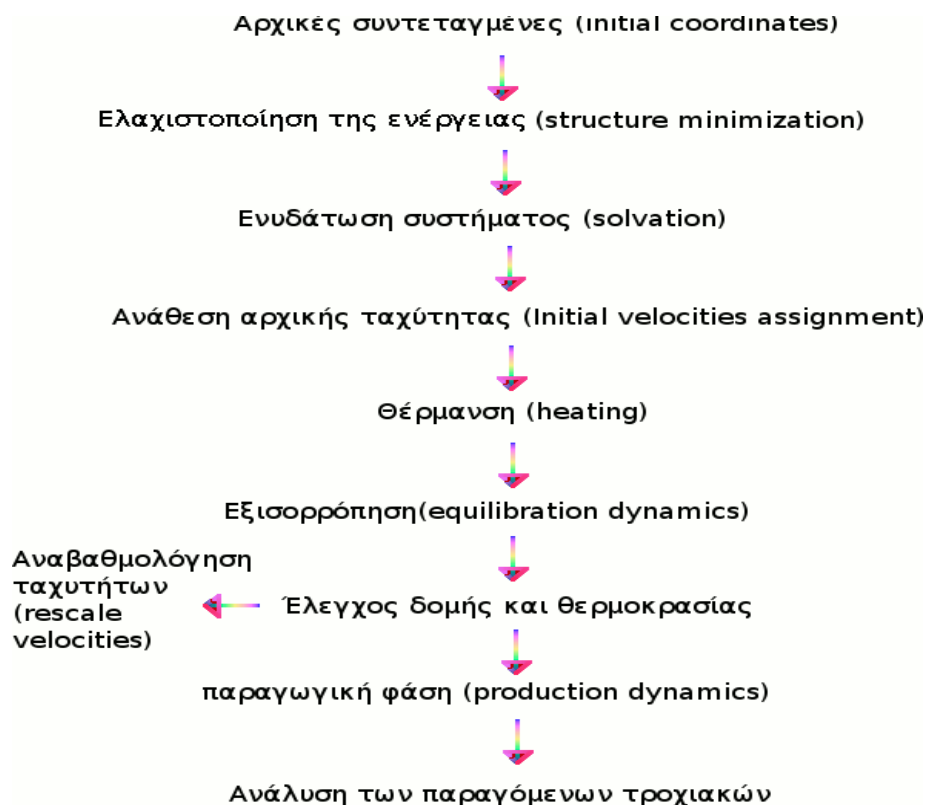
Πίνακας 3.1 Χαρακτηριστικά και συνθήκες της εκάστοτε προσομοίωσης.



3.2 Προαπαιτούμενα αρχεία

3.2 Προαπαιτούμενα αρχεία

Κατά κανόνα όλα (σχεδόν) τα πρωτόκολλα προσομοιώσεων μοριακής δυναμικής απαιτούν την ύπαρξη (ή δημιουργία) συγκεκριμένων αρχείων προκειμένου να γίνει η προετοιμασία των συστημάτων που πρόκειται να μελετηθούν, επίσης σε όλα (σχεδόν) τα πρωτόκολλα η διαδικασία της προετοιμασίας των συστημάτων είναι ίδια -πλην ορισμένων παραλλαγών- με αυτή που αναλύεται παρακάτω.



Σχεδιάγραμμα 1. Διάγραμμα ροής προσομοιώσεων μοριακής δυναμικής

Όπως προαναφέρθηκε οι προσομοιώσεις πραγματοποιήθηκαν με το πρόγραμμα **NAMD** χρησιμοποιώντας το δυναμικό πεδίο “**AMBER 99SB**” και μια παραλλαγή του, το “**AMBER 99SB ILDN**”. Το **NAMD** χρειάζεται τέσσερις τύπους αρχείων πριν αρχίσει η διαδικασία της προσομοίωσης. Η ελάχιστη είσοδος, λοιπόν, για μια προσομοίωση είναι :

Ένα αρχείο που περιέχει τις ατομικές συντεταγμένες ή/και τις ταχύτητες του συστήματος. Τέτοια αρχεία είναι τα **.pdb**. Αυτά μπορούν να δημιουργηθούν εξολοκλήρου με το χέρι, αλλά επειδή συνήθως υπάρχουν πολλές στερικές συγκρούσεις μεταξύ των μορίων του συστήματος, θα χρειαστεί να γίνει μια σύντομη ελαχιστοποίηση της ενέργειας (**structure minimization**) πριν συνεχιστεί η διαδικασία. Εναλλακτικά, τα αρχεία αυτά είναι κατατεθειμένα στην **Protein Data Bank**, για αρκετές χιλιάδες δομές που έχουν χαρακτηριστεί μέσω τεχνικών όπως κρυσταλλογραφία **NMR** και **X-ray**.

Ένα αρχείο που έχει αποθηκευμένες παραμέτρους για το δυναμικό πεδίο και την τοπολογία (.prmtop file), δηλαδή τον τρόπο με τον οποίο είναι ομοιοπολικά συνδεδεμένα τα μόρια του συστήματος. Περιέχει πληροφορίες όπως ο τύπος των ατόμων, ο αριθμός των ατόμων ανά μόριο, το φορτίο, οι δεσμοί που υπάρχουν μεταξύ των ατόμων του συστήματος (αριθμούνται οι δεσμοί που περιλαμβάνουν υδρογόνα και αυτοί που δεν περιλαμβάνουν υδρογόνα), οι γωνίες μεταξύ των ατόμων. Το συγκεκριμένο αρχείο θα μπορούσε να δημιουργηθεί μέσω του AMBER για το NAMD, αλλά για οικονομικούς λόγους επιλέχθηκε το AMBERTOOLS, το οποίο είναι ελεύθερο προς χρήση και συμβατό με το NAMD, για τη δημιουργία αυτού του αρχείου και συγκεκριμένα χρησιμοποιείται το εργαλείο Xleap. Το script που χρησιμοποιεί το Xleap βρίσκεται στο leap.script.

Επιπλέον, είναι απαραίτητο να υπάρχει ένα αρχείο στο οποίο ο χρήστης καθορίζει όλες τις επιλογές τις οποίες το NAMD θα υιοθετήσει κατά τη διάρκεια της προσομοίωσης. Το αρχείο αυτό είναι το αρχείο διαμόρφωσης ή configuration file (με κατάληξη .conf ή συνηθέστερα .namd) και συμπεριλαμβάνει τα εξής εδάφια :

Το αρχικό κομμάτι αφορά τα αρχεία που χρησιμοποιούνται ως είσοδο και περιλαμβάνουν τις παραμέτρους και την τοπολογία (.prmtop file). Επίσης αναφέρονται τα αρχεία στα οποία καταγράφονται συνεχώς νέα δεδομένα από την εξέλιξη της προσομοίωσης, όπως η ταχύτητα και οι συντεταγμένες, τα αρχεία στα οποία αποθηκεύονται οι πληροφορίες αυτές είναι αρχεία τύπου .coor (coordinates), .vel (velocity) και .xsc (eXtended System Configuration) και χρησιμοποιούνται για την περίπτωση που θα χρειαστεί να γίνει επανεκκίνηση της προσομοίωσης από το σημείο στο οποίο σταματήσαμε. Επίσης στην αρχή διευκρινίζεται το force field που θα χρησιμοποιηθεί (στη δική μας περίπτωση, το AMBER).

Εν συνεχεία, αναφέρεται το όνομα των αρχείων εξόδου, η συχνότητα καταγραφής των .dcd αρχείων (αρχεία γραμμένα στο δυαδικό σύστημα, που

περιέχουν τις συντεταγμένες όλων των ατόμων) καθώς και των αρχείων επανεκκίνησης (.restart.coor, .restart.vel, restart.xsc).

Το επόμενο εδάφιο αφορά τις συχνότητες καταγραφής των log και .xst (eXtended System Trajectory) αρχείων (outputEnergies, outputTiming, xstFreq). Κατά τη διάρκεια της προσομοίωσης δημιουργούνται log αρχεία στα οποία καταγράφεται η πρόοδος της προσομοίωσης και ό,τι εκτελείται την ανάλογη χρονική στιγμή. Τα αρχεία αυτά αποθηκεύουν πληροφορίες με βήμα, που καθορίζεται από τον χρήστη. Το .xst αρχείο περιέχει μια καταγραφή του περιοδικού κουτιού και των μεταβλητών του εκτεταμένου συστήματος κατά τη διάρκεια της προσομοίωσης.

Αμέσως μετά ορίζεται το χρονικό βήμα για την προσομοίωση, που συνήθως περιορίζεται στο 1 fs, είναι δυνατό να επεκταθεί στα 2 fs καθορίζοντας το μήκος όλων των δεσμών στο μόριο που περιέχουν άτομο υδρογόνου. Αυτό γίνεται μέσω της επιλογής rigidBonds --> all (η οποία βρίσκεται σε επόμενο εδάφιο), και έχει ως επιπλέον συνέπεια να γίνουν τα μόρια νερού τελείως άκαμπτα. Στην δική μας περίπτωση το timestep έχει επεκταθεί στα 2 fs. Επίσης ορίζεται το βήμα για τις ηλεκτροστατικές αλληλεπιδράσεις (nonBondedFreq, fullElectFrequency) σύμφωνα με την μέθοδο particle mesh Ewald (PME) και τα βήματα ανά κύκλο (stepsPerCycle) που συνδέονται με την επιλογή pairlistdist (αναφέρεται στο επόμενο εδάφιο) και δημιουργούνται ως ζεύγος στην αρχή κάθε κύκλου. Η επιλογή stepsPerCycle ρυθμίζει πόσα βήματα πρέπει να ολοκληρωθούν για να δημιουργηθεί ένα νέο ζεύγος, δηλαδή σηματοδοτεί την έναρξη ενός νέου κύκλου. Έτσι το μήκος της προσομοίωσης αντιστοιχεί σε έναν ακέραιο αριθμό κύκλων.

Στα επόμενα εδάφια συμπεριλαμβάνονται παράμετροι και επιλογές που αφορούν το δυναμικό πεδίο, το πλέγμα στο οποίο θα υπολογιστούν οι ηλεκτροστατικές αλληλεπιδράσεις (μέσω PME για full-system periodic electrostatics), τις περιοδικές συνθήκες, τη δυναμική Langevin (έλεγχος της πίεσης και των ιδιοτήτων του βαρόμετρου)

Τέλος ορίζεται ποιο θα είναι το πρώτο timestep και πόσα ns θα τρέξει η προσομοίωση.

Για την ανάλυση των προσομοιώσεων που θα πραγματοποιηθούν με το πρόγραμμα CARMA είναι απαραίτητη η δημιουργία ενός αρχείου .psf, που έχει αποθηκευμένες όλες τις απαραίτητες πληροφορίες για την πρωτεϊνική τη δομή ενός μορίου (protein structure file, .psf). Το .psf αρχείο στην αρχή εμφανίζει τον τίτλο του αρχείου, καθώς και μια περιγραφή (σχόλιο) για τον τρόπο δημιουργίας του. Στη συνέχεια ακολουθούν πέντε κύρια εδάφια με τους εξής τίτλους :

ATOMS:

Στην εικόνα φαίνεται ο τύπος του αρχείου (PSF) και ο τίτλος (x-plor format file). Το εδάφιο !NATOMS περιέχει εννέα στήλες στις οποίες αναγράφονται ο η ατομική ταυτότητα (atom ID), ο τομέας (συνήθως αλυσίδα, εφόσον μιλάμε για πρωτεϊνικές αλληλουχίες), η ταυτότητα του καταλοίπου (residue ID) στο οποίο ανήκει το εκάστοτε άτομο, το όνομα του καταλοίπου (residue name), το όνομα του ατόμου (atom name), ο τύπος του ατόμου (atom type), το φορτίο, η μάζα και ένα 0 στην τελευταία στήλη (δεν χρησιμοποιείται).

```
PSF
1 !NTITLE
REMARKS VMD generated structure x-plor psf file

46026 !NATOM
1 A 1 MET N N3 0.159200 14.0100 0
2 A 1 MET H1 H 0.198400 1.0080 0
3 A 1 MET H2 H 0.198400 1.0080 0
```

BONDS:

Αναφέρεται στους ομοιοπολικούς δεσμούς και στο αρχείο φαίνονται τέσσερα ζεύγη ατόμων ανά σειρά, όπως φαίνεται και στην εικόνα.

```
46026 !NBOND: bonds
1 2 1 3 1 4 1 5
5 6 5 7 5 18 7 8
7 9 7 10 10 11 10 12
```


ANGLES:

Εμφανίζονται τρεις τριάδες ατόμων στην κάθε σειρά.

```
46264 !NTHETA: angles
  2      1      3      2      1      4      2      1      5
  3      1      4      3      1      5      4      1      5
  1      5      6      1      5      7      1      5      18
  6      5      7      6      5      18      7      5      18
```

DIHEDRALS και IMROPERS:

Εμφανίζονται δύο τετράδες ατόμων ανά σειρά.

```
46493 !NPHI: dihedrals
  2      1      5      6      2      1      5      7
  2      1      5      18      3      1      5      6
  3      1      5      7      3      1      5      18
  4      1      5      6      4      1      5      7
  4      1      5      18      1      5      7      8
```

Τα .psf αρχεία μπορεί να είναι σε CHARMM ή X-PLOR μορφή (format) και θα πρέπει η σειρά των ατόμων που βρίσκονται στα αρχεία .pdb και .dcd να ταιριάζει ακριβώς με τη σειρά τους στο .psf αρχείο. Το .psf δημιουργείται μέσω του VMD, χρησιμοποιώντας τα αρχεία .pdb και .prmtop.

3.3 Προετοιμασία συστήματος

Σε αυτό το τμήμα του κεφαλαίου θα αναλυθεί καθένα από τα βήματα του σχεδιαγράμματος ροής των προσομοιώσεων μοριακής δυναμικής (σχεδιάγραμμα 3.1), έτσι ώστε να παρουσιασθεί η πειραματική μέθοδος που ακολουθήθηκε από την αρχή μέχρι και το τέλος των προσομοιώσεων. Η εκτενής ανάλυση των δεδομένων παρουσιάζεται στο Κεφάλαιο 4.

Οι προσομοιώσεις μοριακής δυναμικής, όπως περιγράφηκε στο κεφάλαιο 2, είναι μια μέθοδος υπολογισμού της θέσης, της ταχύτητας και της ενέργειας κάθε ατόμου σε ένα σύστημα στη ροή του χρόνου. Μέσω των προσομοιώσεων μοριακής δυναμικής έχουμε τη δυνατότητα να παρατηρήσουμε τη χρονική εξέλιξη της αναδίπλωσης ενός πεπτιδίου από την εκτεταμένη, υψηλής ενέργειας, διαμόρφωση έως την τελική, ελάχιστης ενέργειας διαμόρφωση του πεπτιδίου. Κάθε βήμα στη διαδικασία αυτή (χρονικά ισούται με 2fs) αντιστοιχεί στην καταγραφή των ιδιοτήτων του συστήματος και το σύνολο των βημάτων αποτελεί ένα τροχιακό (trajectory), με απλά λόγια ένα τροχιακό είναι μια ταινία που δείχνει τις διαμορφώσεις που αποκτά το πεπτίδιο μέχρι το τέλος της προσομοίωσης. Σε αυτό το σημείο πρέπει να επισημανθεί πως στις τρεις προσομοιώσεις που πραγματοποιήθηκαν ακολουθήθηκε η ίδια διαδικασία προετοιμασίας των συστημάτων, με κάποιες μικρές αποκλίσεις που θα επισημανθούν στη συνέχεια, για το λόγο αυτό το πρωτόκολλο προετοιμασίας θα παρουσιαστεί μια φορά για όλες τις προσομοιώσεις.

Πριν ξεκινήσει η προετοιμασία του συστήματος μας θα πρέπει να έχουμε πάρει κάποιες αποφάσεις για το πώς θα γίνουν οι προσομοιώσεις. Θα πρέπει να απαντήσουμε σε ερωτήσεις, όπως : “Με ποιο πρόγραμμα θα γίνουν οι προσομοιώσεις;”, “Ποιο είναι το καταλληλότερο δυναμικό πεδίο για να έχω αποτελέσματα που έρχονται πολύ κοντά στην πραγματικότητα, ανάλογα με το

πεπτίδιο που θέλω να μελετήσω;”, “Σε τι συνθήκες θα γίνει η προσομοίωση (πίεση, θερμοκρασία, όγκος συστήματος);”, “Θα χρησιμοποιήσω νερά στο σύστημά μου;”, “Θα χρησιμοποιήσω περιοδικές οριακές συνθήκες ή στοχαστικές μεθόδους;”. Τα προαναφερθέντα είναι μερικά από τα ερωτήματα που απορρέουν από το κύριο ερώτημα “Μπορώ να προβλέψω την τριτοταγή δομή ενός μορίου χρησιμοποιώντας τους νόμους της φυσικής, έχοντας μόνο την αλληλουχία του (πρωτοταγής δομή);”

3.3.1 Ορισμός αρχικών συντεταγμένων

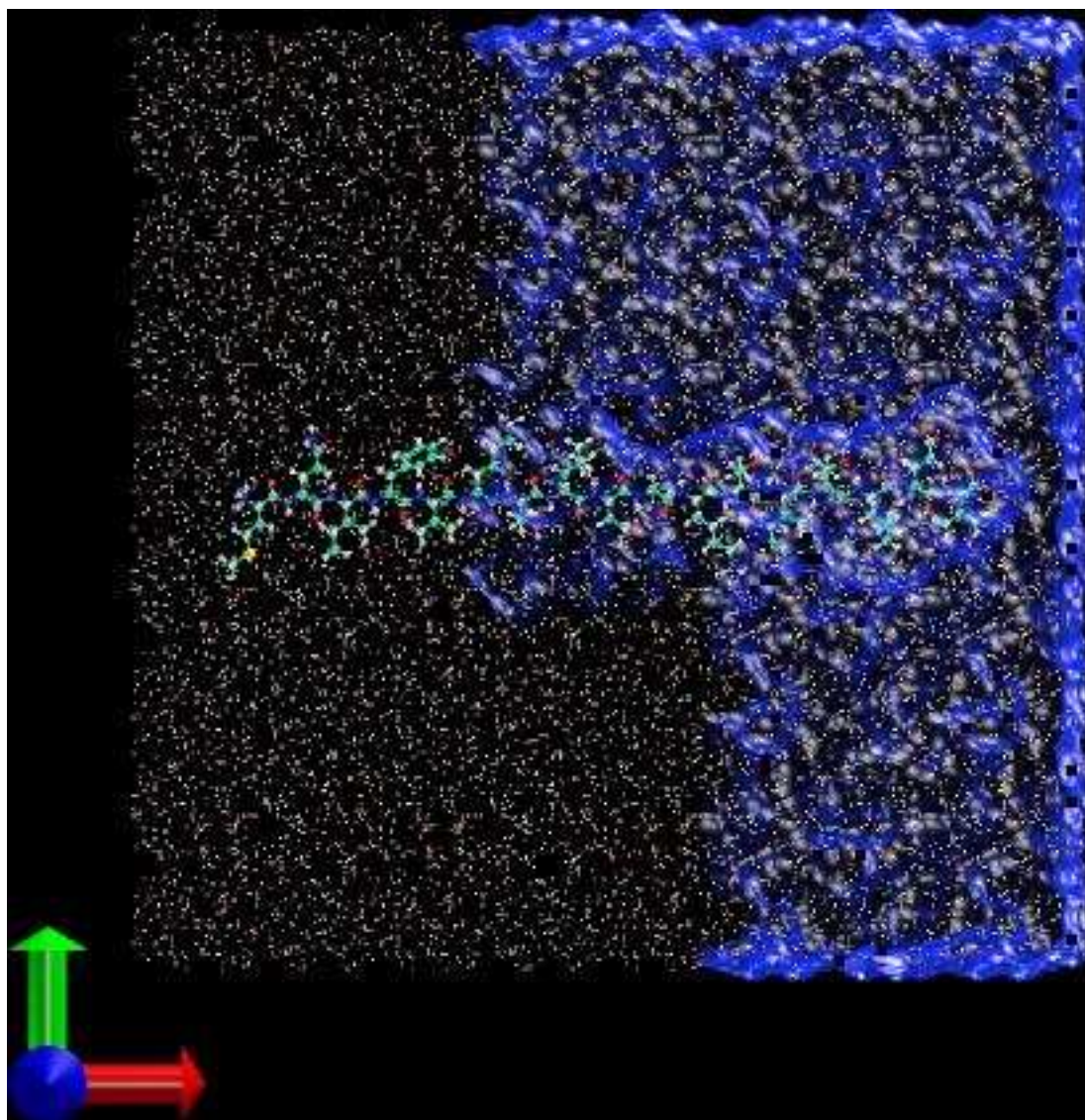
Ένα τροχιακό αποτελείται από τρεις φάσεις : 1) Αρχικοποίηση, 2) Θέρμανση - Εξισορρόπηση και 3) Παραγωγική φάση. Το πρώτο στάδιο στις προσομοιώσεις μοριακής δυναμικής είναι ο ορισμός των αρχικών συντεταγμένων όλων των ατόμων. Το αρχείο που περιέχει τις αρχικές συντεταγμένες, όπως έχουμε προαναφέρει, είναι το .pdb το οποίο βρίσκεται στην Protein Data Bank. Το αρχικό PDB αρχείο είναι το ίδιο για όλες τις προσομοιώσεις και περιέχει την πλήρως εκτεταμένη μορφή του πεπτιδίου. Το αρχείο αυτό περιλαμβάνει τις συντεταγμένες των ατόμων του μεταλλαγμένου δεκαεπταπεπτιδίου του αμινοτελικού άκρου της ουβικιτίνης [U(1-17)T9D].

Επειδή το πεπτίδιο σε διαφορετική τιμή pH εμφανίζει κάποια κατάλοιπα πρωτονιομένα, πραγματοποιήθηκαν δύο είδους προσομοιώσεις, η μία αφορά το πεπτίδιο χωρίς πρωτονιομένα κατάλοιπα και πραγματοποιήθηκε με δύο force fields, το AMBER99SB και το AMBER99SB-ILDN, και η δεύτερη αφορά το πεπτίδιο πρωτονιομένο στα κατάλοιπα Asp και Glu, στην οποία προστέθηκαν δύο ιόντα χλωρίου για την ουδετεροποίηση. Η δεύτερη πραγματοποιήθηκε σε όξινο pH και συγκεκριμένα στο pKa του Asp (pH = 3,8) χρησιμοποιώντας το AMBER99SB-ILDN force field.

Το επόμενο βήμα μετά την απόκτηση του .pdb αρχείου είναι η ευθυγράμμιση των αδρανειακών αξόνων του πεπτιδίου με το ορθοκανονικό σύστημα αξόνων και η μετατόπιση του κέντρου βάρους στην αρχή των αξόνων. Έπειτα η διαδικασία προετοιμασίας συμπεριλαμβάνει τη δημιουργία του αρχείου .prmtop το οποίο όπως έχουμε ήδη αναφέρει γίνεται μέσω του Xleap (ambertools). Σύμφωνα με το .pdb αρχείο που χρησιμοποιείται ως αρχείο εισόδου και το δυναμικό πεδίο με το οποίο θα γίνει η προσομοίωση, το Xleap φορτώνει αρχικά τις βιβλιοθήκες ονοματολογίας για οποιοδήποτε τύπο ονόματος υπάρχει περίπτωση να χρησιμοποιηθεί κατά την προσομοίωση (άτομα, αμινοξέα, νερά, ιόντα κ.ο.κ), ελέγχει τη συμβατότητα με τα ονόματα που υπάρχουν στο .pdb αρχείο και προσθέτει άτομα όπου είναι απαραίτητο, συνήθως συμπληρώνει τα υδρογόνα που λείπουν, για παράδειγμα κατά την προετοιμασία του αρχείου .prmtop της αρχικής εκτέλεσης προστέθηκαν συνολικά 129 άτομα εκ των οποίων τα 125 ήταν υδρογόνα. Εν συνεχεία το Xleap δημιουργεί το κουτί με τα νερά το οποίο περιβάλλει το πεπτίδιο. Σε όλα τα συστήματα χρησιμοποιήθηκαν TIP3P μοντέλα νερού. Το αρχικό κουτί είχε διαστάσεις 78.825 x 78.825 x 78.825, έτσι ώστε να υπάρχει ένα στρώμα νερού, τουλάχιστον, 6Å γύρω από το πεπτίδιο (Εικόνα 3.2). Αυτό είναι επιθυμητό σε προσομοιώσεις με περιοδικές οριακές συνθήκες για να είμαστε σίγουροι πως το πεπτίδιο δεν θα αλληλεπιδρά με κάποιο από τα είδωλα που το περιβάλλουν. Το αρχικό σύστημα αποτελείται από 46026 άτομα και συμπεριλαμβάνει 15247 μόρια νερού. Στη συνέχεια γίνεται η προσθήκη ιόντων εάν είναι απαραίτητο. Τα ιόντα τοποθετούνται αυτόματα σύμφωνα με την κατανομή των φορτίων του πεπτιδίου. Υπάρχουν αλγόριθμοι οι οποίοι τοποθετούν θετικά ιόντα σε ηλεκτροαρνητικές περιοχές και αρνητικά ιόντα σε ηλεκτροθετικές περιοχές, έτσι ώστε επιτευχθεί μια επιθυμητή ολική ιοντική συγκέντρωση (T. Schlick, 2002). Στα συστήματα “99SB” και “ILDN” δεν προστέθηκαν ιόντα. Στο σύστημα “protonated” προστέθηκαν δύο ιόντα χλωρίου. Εφόσον έχουν ολοκληρωθεί επιτυχώς όλα τα προαναφερθέντα στάδια, το Xleap δημιουργεί τα .prmtop και .crd όπου αποθηκεύονται όλες οι παράμετροι για την τοπολογία και τον τρόπο σύνδεσης των ατόμων. Τέλος

δημιουργείται το νέο .pdb αρχείο το οποίο περιέχει τις συντεταγμένες όλων των ατόμων.

Χρησιμοποιώντας τα αρχεία που δημιουργήθηκαν με το Xleap (.prmtop, .pdb) ετοιμάζουμε το .psf αρχείο (protein structure file) με τη βοήθεια του VMD.



Εικόνα 3.2.

Το αρχικό σύστημα συμπεριλαμβάνει το ευθυγραμμισμένο πεπτίδιο (U(1-17)T9D) στο κέντρο του συστήματος και τα νερά που το περιβάλλουν. Η εικόνα ετοιμάσθηκε με το VMD. Τα νερά αντιπροσωπεύονται από κουκίδες σε όλη την έκταση του κουτιού, ενώ στο δεξί τμήμα της εικόνας αντιπροσωπεύονται και από

τη μοριακή επιφάνεια, η οποία παράγεται μέσω του SURF αλγόριθμου που χρησιμοποιείται από το VMD. Το πεπτίδιο αναπαρίσταται με CPK.

3.3.2 Ελαχιστοποίηση και θέρμανση του αρχικού συστήματος.

Εφόσον έχει ολοκληρωθεί προσεκτικά η πρώτη φάση που περιλαμβάνει τον καθορισμό των αρχικών συντεταγμένων, μπορούμε να περάσουμε στο στάδιο της ελαχιστοποίησης της ενέργειας του συστήματος και της σταδιακής θέρμανσης, μέχρι την απόκτηση της επιθυμητής θερμοκρασίας.

Σε αυτό το σημείο χρειαζόμαστε το αρχείο διαμόρφωσης `heat.namd`, το οποίο ορίζει πως ακριβώς θα γίνει η διαδικασία της ελαχιστοποίησης και της θέρμανσης του συστήματος. Τα αρχεία που χρησιμοποιούνται ως είσοδος για τη δημιουργία του `heat.namd` είναι το `.prmtop` και το `.pdb` που δημιουργήθηκαν στην προηγούμενη φάση της προετοιμασίας. Η ελαχιστοποίηση πραγματοποιείται αρχικά για τα άτομα που δεν ανήκουν στον πρωτεϊνικό σκελετό για 2000 βήματα (ένα βήμα αντιστοιχεί σε 2 fs). Στη συνέχεια η θερμοκρασία του συστήματος ορίζεται στους 20K και αυξάνεται σταδιακά με βήμα 20 μέχρι να φτάσει στους 320K. Καθ' όλη τη διάρκεια της θέρμανσης διατηρούνται σταθερά τα άτομα Ca. Έπειτα ακολουθεί για 12000 βήματα η εξισορρόπηση του όγκου με τα άτομα Ca ακινητοποιημένα, χρησιμοποιώντας το βαρόμετρο Langevin. Η συχνότητα καταγραφής των `.dcd` αρχείων είναι 200, ενώ η συχνότητα καταγραφής των αρχείων επανεκκίνησης είναι 10000. μέχρι το τέλος της φάσης αυτής έχουν τρέξει 30000 βήματα.

Στο `heat.namd` καθορίζεται η μέθοδος υπολογισμού των ηλεκτροστατικών αλληλεπιδράσεων η οποία είναι η `particle-mesh Ewald (PME)`. Στις περιοδικές συνθέσεις το σύστημα τοποθετείται στη στοιχειώδη κυψελίδα και θεωρείται πως έχει απείρως πολλές εικόνες στο χώρο. Αυτό σημαίνει ένα άπειρο πλέγμα

τριών διαστάσεων. Η μέθοδος PME χρησιμοποιείται για να περιγραφούν οι ηλεκτροστατικές αλληλεπιδράσεις σε ένα τρισδιάστατο πλέγμα.

Οι προσομοιώσεις έγιναν σε ισοβαρείς – ισόθερμες συνθήκες (NPT ensemble). Προκειμένου να παραμείνει σταθερή η πίεση άλλαξε ο όγκος του συστήματος. Για τον έλεγχο του βαρόμετρου χρησιμοποιήθηκε η μέθοδος Langevin dynamics. Η δυναμική Langevin είναι μια εναλλακτική της νευτώνειας δυναμικής. Όταν ένα σύστημα (συμπεριλαμβανομένου του διαλύτη) αντιμετωπίζεται σε πλήρως ατομικό επίπεδο οι δυνάμεις είναι μόνο δυναμικές, ειδάλλως οι δυνάμεις σύγκρουσης και τριβής θα πρέπει να εισαχθούν και να μιμηθούν τις συγκρούσεις του διαλυμένου μορίου με το περιβάλλον του. Αυτή η αντιμετώπιση αναφέρεται ως Langevin ή Brownian dynamics.

3.3.3 Παραγωγική φάση.

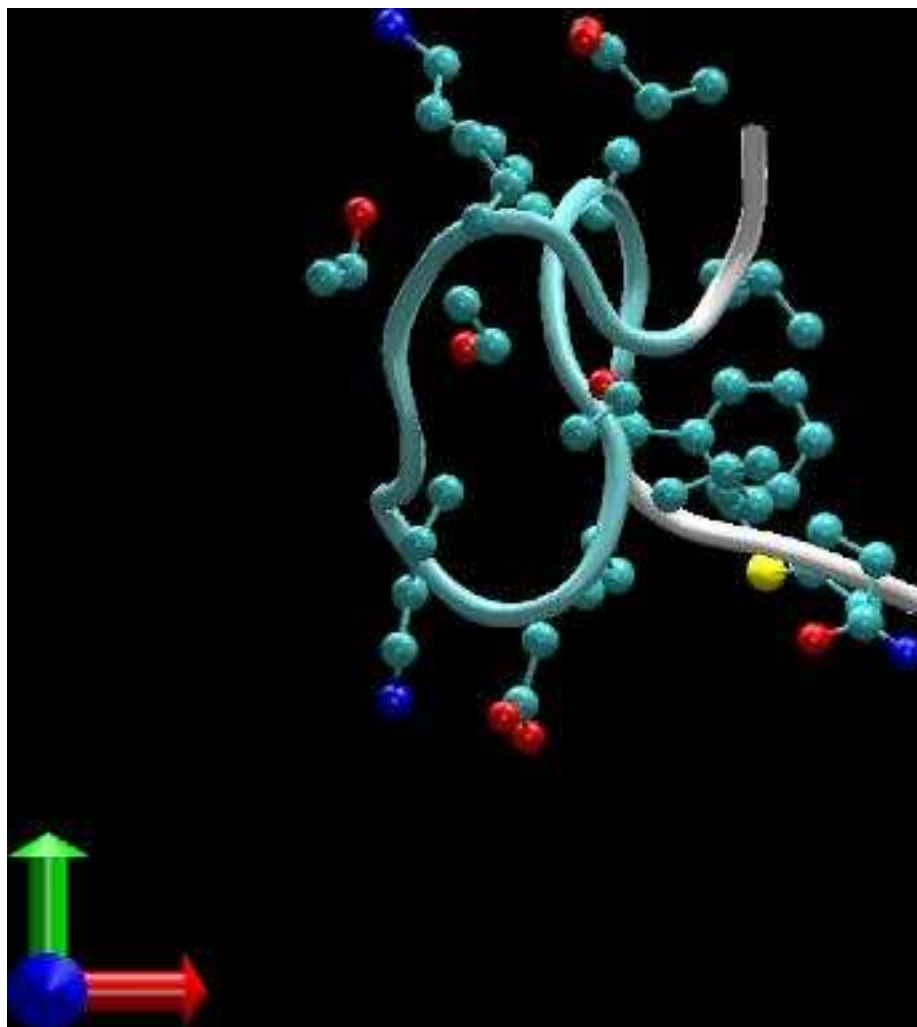
Σε αυτή τη φάση το σύστημα έχει αποκτήσει την επιθυμητή θερμοκρασία και οι συνθήκες πίεσης και θερμοκρασίας παραμένουν σταθερές μέχρι το τέλος της προσομοίωσης. Για να ξεκινήσει αυτή η φάση είναι απαραίτητο ένα αρχείο διαμόρφωσης της παραγωγικής φάσης, το `equi.namd` file, στο οποίο καθορίζεται από τον χρήστη ο τρόπος που θα γίνει η προσομοίωση. Οι συντεταγμένες, οι ταχύτητες και η ενέργεια του συστήματος που καταγράφηκαν στο τελευταίο βήμα της προηγούμενης φάσης χρησιμοποιούνται ως είσοδο για την έναρξη της παραγωγικής φάσης. Στο σύστημά μας η αρχική παραγωγική φάση ξεκινάει από το βήμα 30000 (150 frames) και ορίζουμε να συνεχίσει μέχρι τα 500.000.000 βήματα. Το βήμα καταγραφής των `.dcd` αρχείων, καθώς και της ενέργειας αυξάνεται σε 400. Στην κυριολεξία αυτό που γίνεται μετά από αρκετό αριθμό βημάτων είναι η διακοπή της προσομοίωσης και η επανέναρξή της μετά την μείωση των μορίων νερού του συστήματος. Αυτό γίνεται για να είναι πιο αποδοτική η προσομοίωση, εφόσον μικρότερος αριθμός ατόμων συνεπάγεται και

μικρότερο υπολογιστικό κόστος. Η αρχική παραγωγική φάση σταματά στο timestep 14064000, και μέχρι τώρα έχουν παραχθεί συνολικά 35235 frames στο .dcd αρχείο.

3.3.4 Μείωση του αριθμού μορίων νερού του συστήματος.

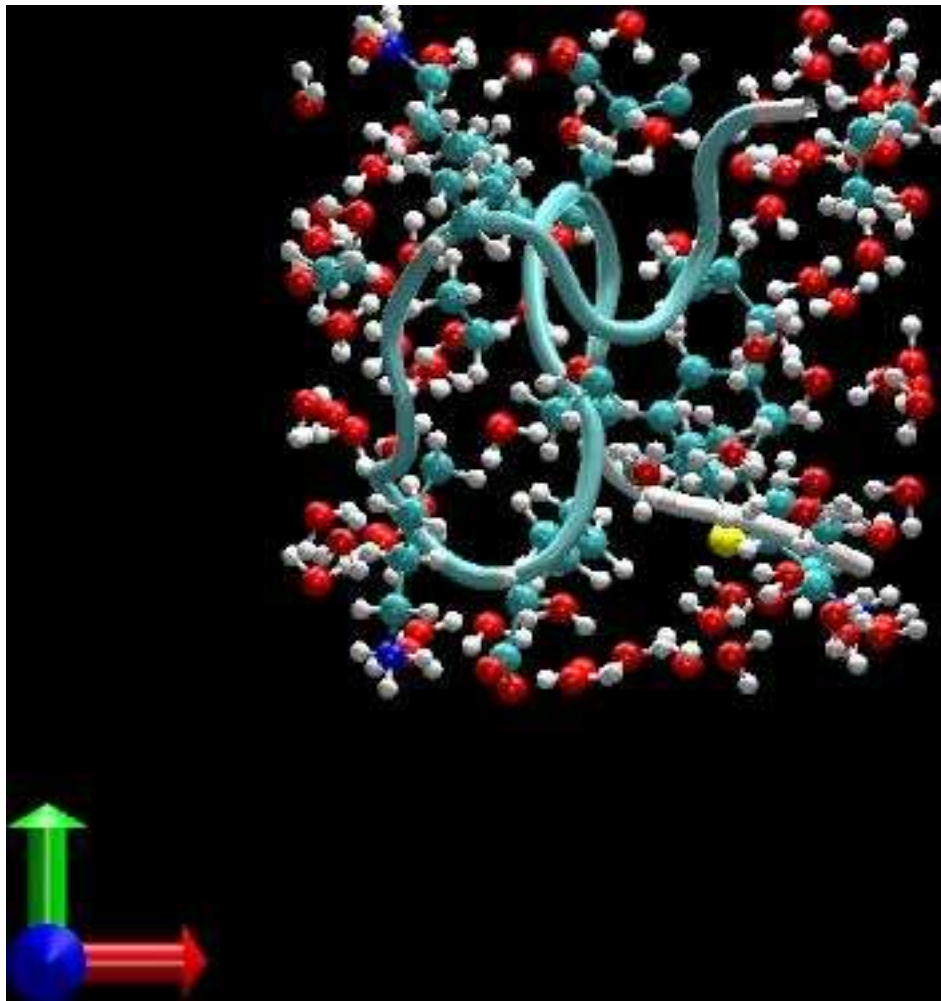
Το αρχικό σύστημα περιελάμβανε συνολικά 46026 άτομα εκ των οποίων τα 45741 αντιστοιχούσαν σε άτομα των μορίων νερού. Λαμβάνοντας το τελευταίο frame της αρχικής προσομοίωσης δημιουργήθηκε ένα .pdb αρχείο με το όνομα starting.pdb (εικόνα 3.3). Το starting.pdb περιελάμβανε μόνο 131 άτομα και χρησιμοποιήθηκε μέσω του Xleap για να δημιουργηθούν τα ubi.prmtop, ubi.crd και ubi.pdb (εικόνα 3.5) αρχεία που συμπεριλαμβάνουν συνολικά 5616 άτομα, εκ των οποίων τα 5331 ανήκουν σε μόρια νερού. Το νέο μέγεθος του κουτιού ορίστηκε στα 40,161 x 40,161 x 40,161 Å , έτσι ώστε να υπάρχει πλέον ένα στρώμα νερού, πάχους, τουλάχιστον 5Å, γύρω από το πεπτίδιο.

Στη συνέχεια πραγματοποιήθηκε ελαχιστοποίηση και θέρμανση του συστήματος με τον τρόπο παρόμοιο με αυτόν που περιγράφηκε στο εδάφιο 3.3.2 για πολύ μικρό χρονικό διάστημα. Αρχικά έγινε η ελαχιστοποίηση και μετά τέθηκε η θερμοκρασία στους 20K, με αυτή την θερμοκρασία τρέξαμε 1000 βήματα και στη συνέχεια η θερμοκρασία αυξήθηκε σταδιακά με βήμα 20 μέχρι τους 320K. Συνολικά έγιναν 20000 βήματα τα οποία αντιστοιχούν σε 95 frames του .dcd αρχείου. Συνολικά η προσομοίωση “99SB” περιέχει 5035100 frames που αντιστοιχούν σε 4,03 μs. Με τον ίδιο τρόπο ετοιμάστηκε η προσομοίωση “ILDN”, η μόνη διαφορά έγκειται σε κάποιες παραμέτρους του δυναμικού πεδίου. Το συνολικό τροχιακό της “ILDN” αποτελείται από 3758525 frames. Το συνολικό τροχιακό της “protonated” συμπεριλαμβάνει 3714800 frames. Το πρώτο frame της “protonated” είναι αυτό που χρησιμοποιήθηκε για να δημιουργηθούν τα μικρότερα συστήματα των δύο άλλων προσομοιώσεων. Το κουτί συμπεριλαμβάνει 5596 άτομα εκ των οποίων τα 5307 αντιστοιχούν σε μόρια νερού, τα 287 αντιστοιχούν στο πεπτίδιο και υπάρχουν και δύο ιόντα χλωρίου.



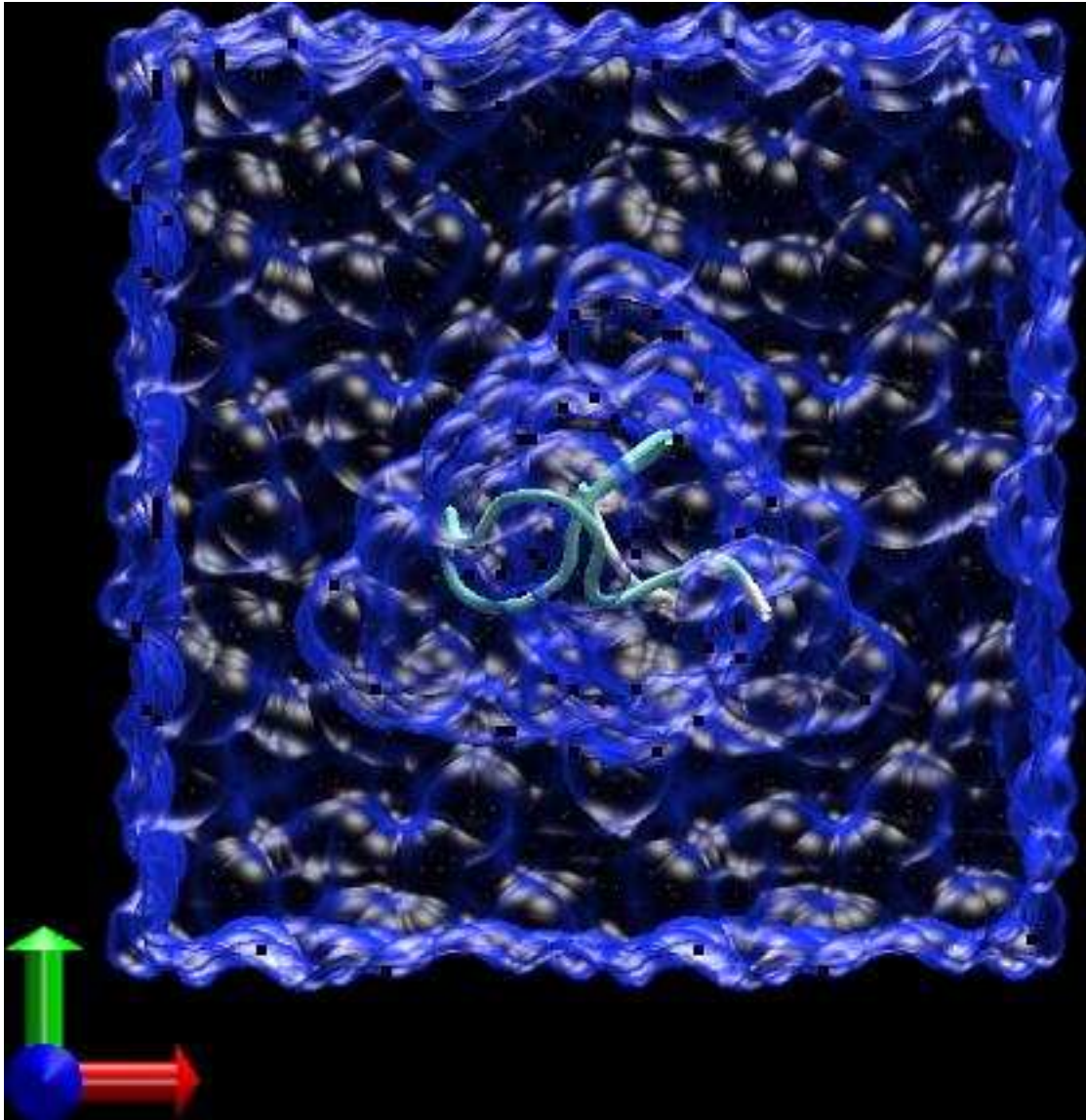
Εικόνα 3.3

Παρουσιάζεται η starting δομή η οποία προέκυψε από την τελική καταγραφή της αρχικής προσομοίωσης. Ο σκελετός της πρωτεΐνης αντιπροσωπεύεται με new cartoon, ενώ οι πλευρικές ομάδες με CPK. Το σύστημα συμπεριλαμβάνει μόλις 131 άτομα.



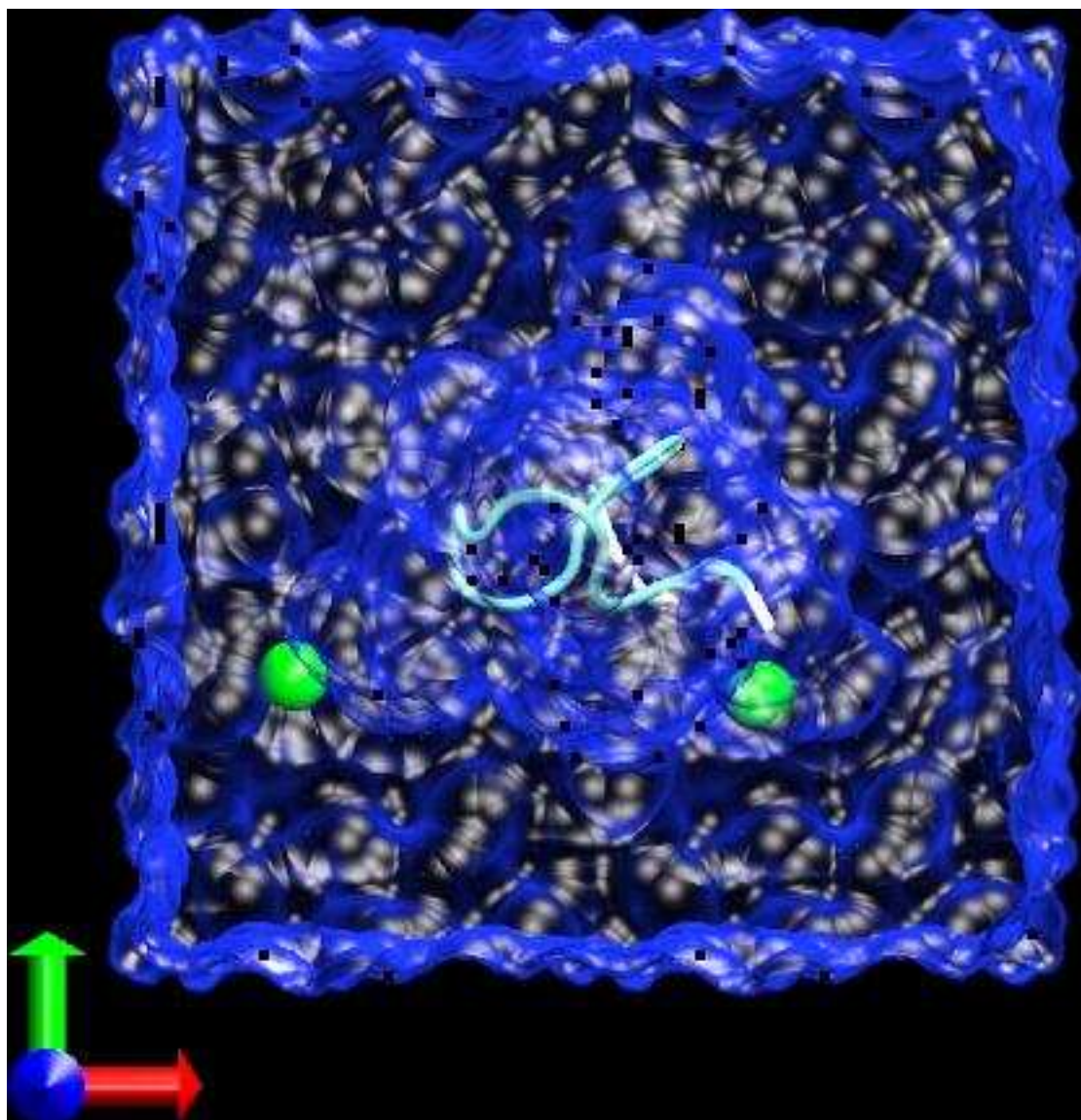
Εικόνα 3.4

Παρουσιάζεται ενυδατωμένο το πεπτίδιο, υπάρχουν 79 μόρια νερού, ενώ συνολικά το σύστημα περιλαμβάνει 522 άτομα.



Εικόνα 3.5

Αναπαριστά το πεπτίδιο έναρξης της δεύτερης παραγωγικής φάσης (κυρίως φάση της προσομοίωσης. Το πεπτίδιο είναι πλήρως ενυδατωμένο ενώ το μέγεθος του κουτιού έχει γίνει πιο μικρό (40x40x40). Στο κέντρο του κουτιού βρίσκεται το πεπτίδιο, όπου αντιπροσωπεύεται με new cartoon. Τα νερά αντιπροσωπεύονται με surf (molecular surface) χρώματος μπλέ. Το σύνολο των ατόμων είναι 5616, από τα οποία τα 5331 ανήκουν σε μόρια νερού.



Εικόνα 3.6

Αναπαριστά την αρχική δομή του πρωτονιαμένου πεπτιδίου. Στο κέντρο βρίσκεται το πεπτίδιο με μορφή new cartoon, οι πράσινες σφαίρες αντιπροσωπεύουν τα ιόντα χλωρίου (vdW αναπαράσταση) και το νερό εμφανίζεται με μπλε χρώμα (molecular surface). Ο συνολικός αριθμός ατόμων ανέρχεται σε 5596, εκ των οποίων τα 5307 ανήκουν σε μόρια νερού.

Εφόσον η περιγραφή του πρωτοκόλλου ολοκληρώθηκε είμαστε σε θέση να περάσουμε στο τελευταίο και κυριότερο κομμάτι των προσομοιώσεων μοριακής δυναμικής, που δεν είναι άλλο από την ανάλυση των τροχιακών που παράχθηκαν.

QUESTION 4

“What do you think science is? There's nothing magical about science. It is simply a systematic way for carefully and thoroughly observing nature and using consistent logic to evaluate results. Which part of that exactly do you disagree with? Do you disagree with being thorough? Using careful observation? Being systematic? Or using consistent logic?”

-Steven Novella-

Παρουσίαση και ανάλυση των αποτελεσμάτων

Μετά από την αναλυτική περιγραφή του προβλήματος της αναδίπλωσης των πρωτεϊνών και των προσομοιώσεων μοριακής δυναμικής, ως physics based μέθοδο επίλυσής του, περνάμε στο τελευταίο τμήμα της παρούσας πτυχιακής εργασίας, που αφορά την παρουσίαση των προγραμμάτων που χρησιμοποιήθηκαν για την ανάλυση των τροχιακών που παρήχθησαν κατά τη διάρκεια των προσομοιώσεων μοριακής δυναμικής του δεκαεπτά-πεπτιδίου U(1-17)T9D ή 1E0Q και την παρουσίαση των αποτελεσμάτων και της σύγκρισης τους με προηγούμενες μελέτες.

Πραγματοποιήθηκαν τρεις διαφορετικές προσομοιώσεις μοριακής δυναμικής του πεπτιδίου 1E0Q με αλληλουχία MQIFVKTLDGKTITLEV, που θα αναφέρονται στη συνέχεια ως “99SB”, “ILDN” και “protonated”. Τα τροχιακά των προσομοιώσεων καταγράφηκαν και χρησιμοποιήθηκαν για την ανάλυση της δομής και της συμπεριφοράς του πεπτιδίου σε συγκεκριμένες συνθήκες, όπως έχουμε αναφέρει σε προηγούμενο κεφάλαιο. Στις αναλύσεις συμπεριλαμβάνονται οι εξής: 1) Ανάλυση κυρίων συνιστωσών (Principal Component Analysis), 2) ανάλυση χαρτών και γραφημάτων βασισμένων στη μέση απόκλιση (RMSD based matrices), 3) ανάλυση δευτεροταγούς δομής (STRIDE και DSSP), και 4) ανάλυση βασισμένη σε ομαδοποίηση (cluster analysis). Στη συνέχεια του κεφαλαίου αναφέρονται οι βασικές αρχές κάθε ανάλυσης καθώς και των προγραμμάτων που χρησιμοποιήθηκαν. Στο τέλος θα συγκρίνουμε τις δομές που προέκυψαν από τα πειράματά μας με την πειραματικά προσδιορισμένη NMR δομή και με δομές από άλλες προσομοιώσεις, καθώς και από τα διαφορετικά δυναμικά πεδία που χρησιμοποιήθηκαν στις προσομοιώσεις.

4.1 RMSD based ανάλυση των τροχιακών

Μια ανάλυση που χρησιμοποιείται κατά κόρον στις αναλύσεις προσομοιώσεων είναι η RMSD (Root-Mean-Square Deviation). Το RMSD είναι ένα ποσοτικό μέτρο της δομικής αλλαγής του μορίου κατά τη διάρκεια της προσομοίωσης, μας δείχνει δηλαδή κατά πόσο αποκλίνει η αρχική δομή από τις υπόλοιπες κατά τη διαδικασία της αναδίπλωσης. Το RMSD ουσιαστικά είναι η μέτρηση μεταξύ των θέσεων των Ca ατόμων (και όχι μόνο) και μπορεί να γίνει για οποιοδήποτε ζεύγος Ca και όχι απαραίτητα μόνο μεταξύ της αρχικής δομής και των υπολοίπων. Όσο μικρότερη είναι η τιμή RMSD τόσο πιο χωροταξικά ισοδύναμες είναι οι δύο δομές. Για να θεωρηθούν δύο δομές ταυτόσημες θα πρέπει το RMSD να έχει τιμή 0.0 Å. Κατά σύμβαση δομές που έχουν RMSD μικρότερο του 2.0 Å μπορούν να θεωρηθούν όμοιες. Το RMSD υπολογίζεται με την εξίσωση (1):

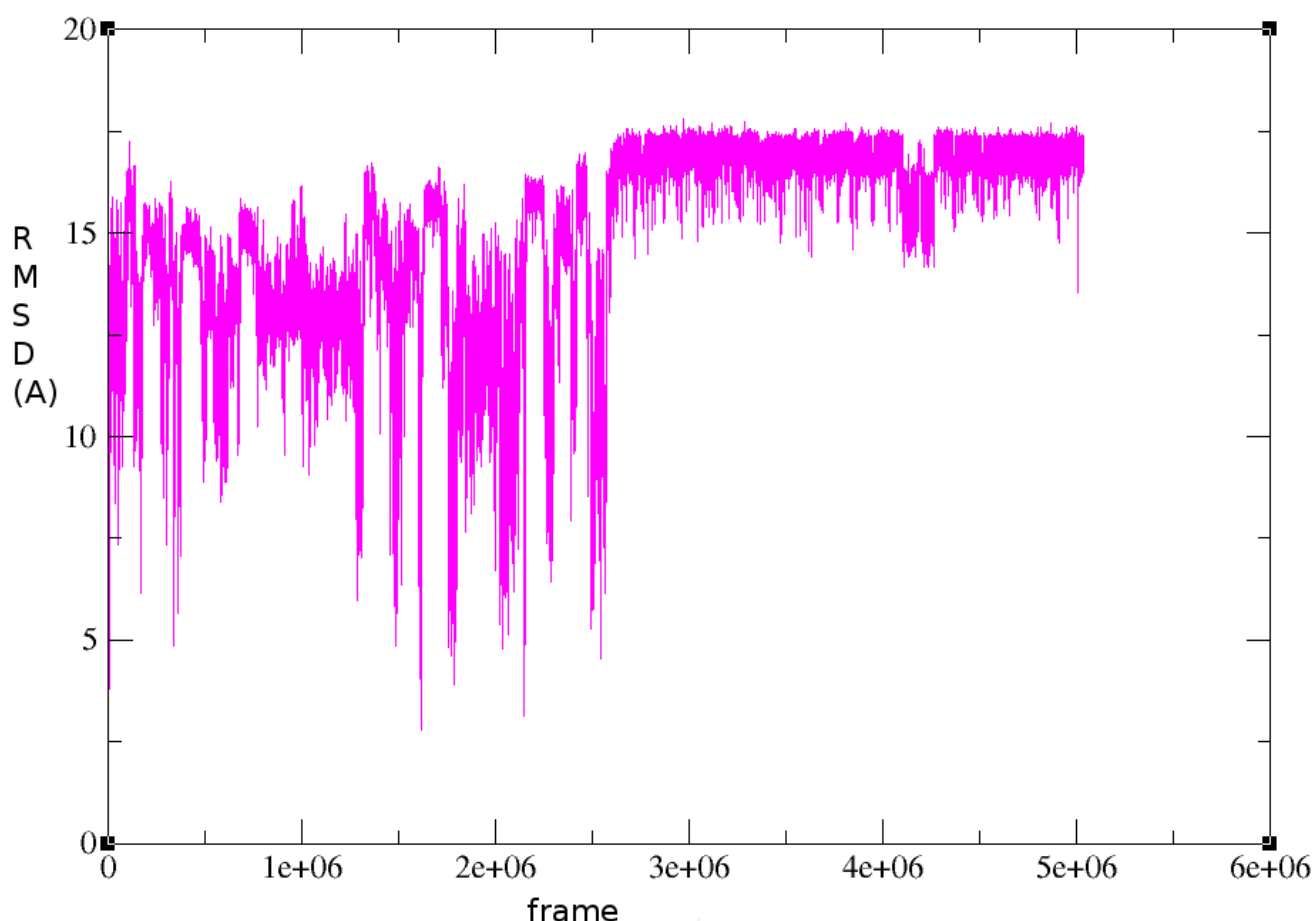
$$\text{RMSD} = \sqrt{[\sum(\chi_i - \chi_{\text{ref}})^2 / N]} \quad (1),$$

όπου, χ_i είναι οι συντεταγμένες της εκάστοτε δομής κατά τη διάρκεια της προσομοίωσης, χ_{ref} είναι οι συντεταγμένες της δομής αναφοράς και N είναι το πλήθος των ατόμων. Μια παραλλαγή (2) του παραπάνω τύπου αποδίδει την απόκλιση ή σύγκλιση ολόκληρου του τροχιακού ως προς τη μέση δομή, η οποία δεν αντιστοιχεί σε πραγματική δομή αλλά αποτελεί το μέσο για το σύνολο των συντεταγμένων ενός τροχιακού. Η διακύμανση από τη μέση δομή αντιστοιχεί στο RMSF και αποτελεί μέτρο κινητικότητας. Μας βοηθά να κατανοήσουμε τη σταθερότητα της δομής σε σχέση με το χρόνο, καθώς αυτή τη φορά η μέση δομή είναι η δομή αναφοράς με την οποία συγκρίνεται ολόκληρο το τροχιακό.

$$\text{RMSF} = \sqrt{[\sum(\chi - \langle \chi_i \rangle)^2 / N]} \quad (2),$$

όπου το $\langle \chi_i \rangle$ αντιστοιχεί στις συντεταγμένες του ατόμου i της μέσης δομής. Για να φέρουμε την παραπάνω μέτρηση πιο κοντά στην πραγματικότητα μπορούμε να βρούμε ποια δομή από το τροχιακό έχει το χαμηλότερο RMSD σε σχέση με τη μέση δομή και να χρησιμοποιήσουμε αυτή ως αντιπροσωπευτική δομή του τροχιακού. Έτσι οι υπολογισμοί θα αντιστοιχούν σε πραγματικές συντεταγμένες που υιοθετήθηκαν κατά τη διάρκεια της προσομοίωσης και θα μπορούμε να ελέγξουμε κατά πόσο, και για πόσο χρονικό διάστημα, αποκλίνουν ή συγκλίνουν οι δομές που παρήχθησαν με την αντιπροσωπευτική δομή.

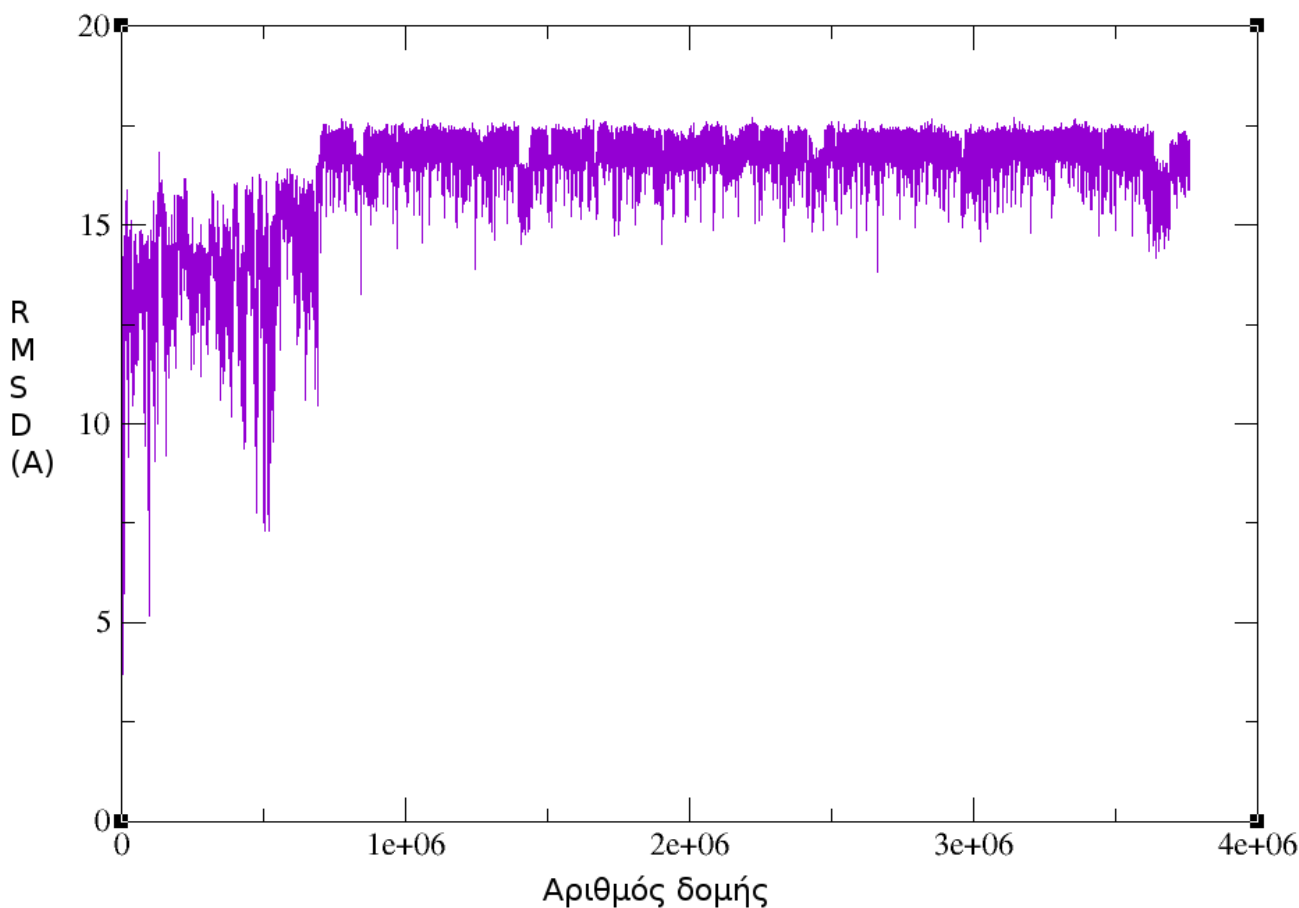
Τα αποτελέσματα από μία ανάλυση RMSD είναι μία γραφική παράσταση στην οποία ο κάθετος άξονας αντιστοιχεί στο RMSD (Å) και ο οριζόντιος στο χρόνο. Όπως φαίνεται και από την εικόνα 4.1 το τροχιακό της προσομοίωσης 99SB αποκτά σταθερή δομή στη μέση περίπου της προσομοίωσης, καθώς το RMSD από την αρχική δομή (πλήρως εκτεταμένο πεπτίδιο) αρχικά ταλαντεύεται μεταξύ 2.5 – 17.0 Å, στη μέση περίπου της προσομοίωσης σταθεροποιείται γύρω στα 17.0 Å για αρκετό χρονικό διάστημα, εν συνεχεία φαίνεται να αλλάζει η δομή καθώς το RMSD μικραίνει για μικρό χρονικό διάστημα, ενώ και πάλι προς το τέλος αυξάνεται και σταθεροποιείται. Από το RMSD αυτό φαίνεται να υπάρχει μία κύρια δομή στο τροχιακό της 99SB.



Εικόνα 4.1

Γραφική παράσταση του RMSD για όλα τα άτομα Ca της αρχικής δομής σε σχέση με τα άτομα Ca των δομών ολόκληρου του τροχιακού της προσομοίωσης 99SB.

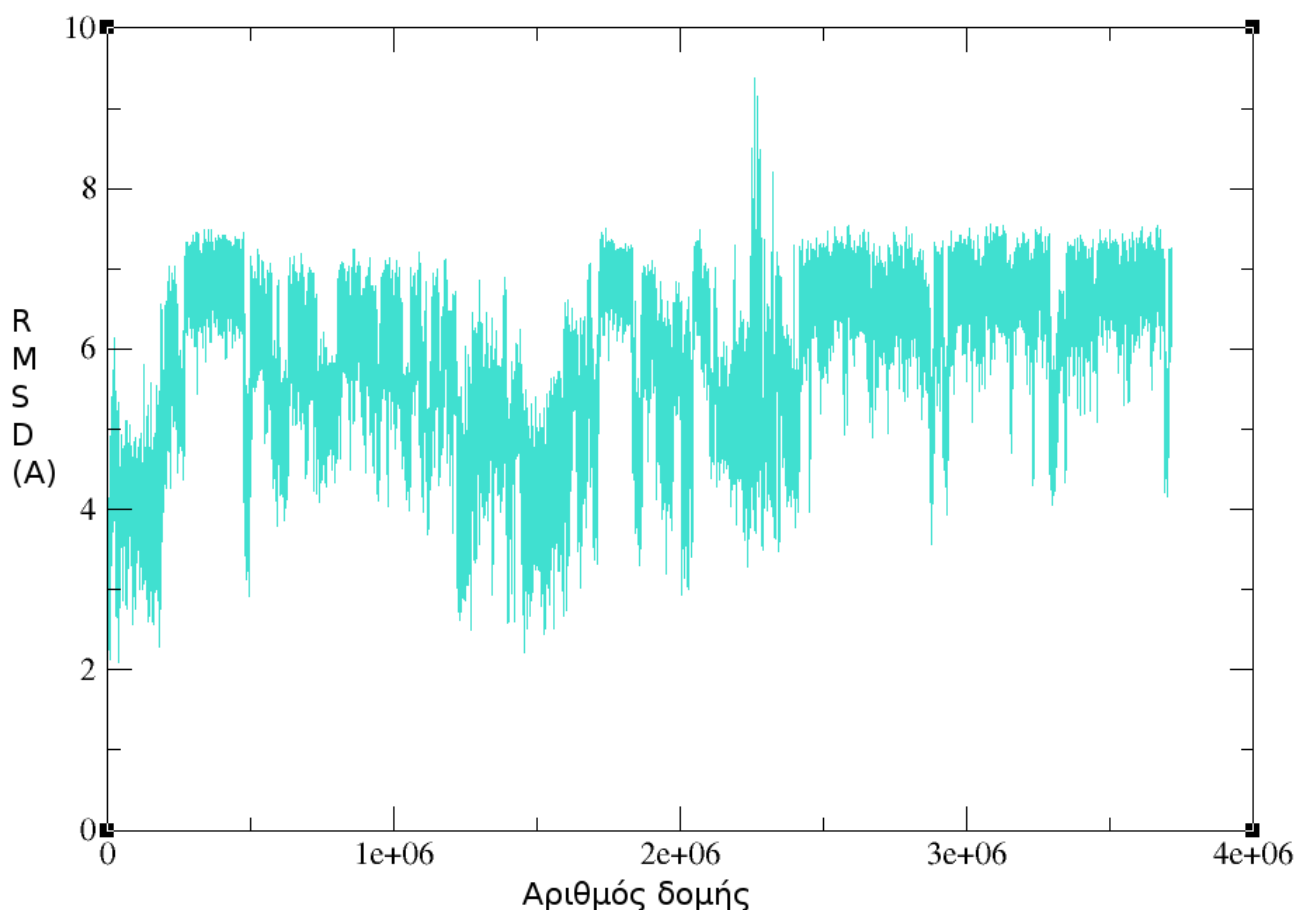
Στο τροχιακό της προσομοίωσης ILDN η σταθεροποίηση της δομής γίνεται πιο γρήγορα σε σχέση με το τροχιακό της 99SB. Όπως φαίνεται και στην εικόνα 4.2 στην αρχή το RMSD έχει χαμηλές τιμές, ενώ περίπου στο 1/3 της προσομοίωσης (περίπου μετά από $1 \cdot 10^6$ στιγμιότυπα) η δομή σταθεροποιείται και παραμένει σταθερή, με μικρές διακυμάνσεις του RMSD μέχρι το τέλος της προσομοίωσης.



Εικόνα 4.2

Γραφική παράσταση του RMSD για όλα τα άτομα Ca της αρχικής δομής σε σχέση με τα άτομα Ca των δομών ολόκληρου του τροχιακού της προσομοίωσης ILDN.

Τέλος στην προσομοίωση protonated η διακύμανση του RMSD είναι χαμηλότερη συγκριτικά με τις δύο προηγούμενες προσομοιώσεις, καθώς το εύρος των τιμών RMSD είναι 2 – 9 Å και παρατηρείται σταθεροποίηση για μικρά χρονικά διαστήματα κοντά στην τιμή 7 Å. Το γεγονός ότι η δομή δεν παραμένει σταθερή για μεγάλο χρονικό διάστημα μέχρι το τέλος της προσομοίωσης πιθανόν να σημαίνει πως ο χρόνος της προσομοίωσης δεν ήταν επαρκής για τη σταθεροποίηση του προτονιωμένου πεπτιδίου. Επιπλέον οι τιμές RMSD των δομών του τροχιακού protonated σε σχέση με την αρχική δομή είναι αρκετά χαμηλότερες από τις τιμές των δυο προηγούμενων τροχιακών. Είναι πιθανόν να υπάρχουν 6 δομές που σταθεροποιούνται για μικρό χρονικό διάστημα. Στην εικόνα 4.3 φαίνεται το RMSD του τροχιακού protonated.

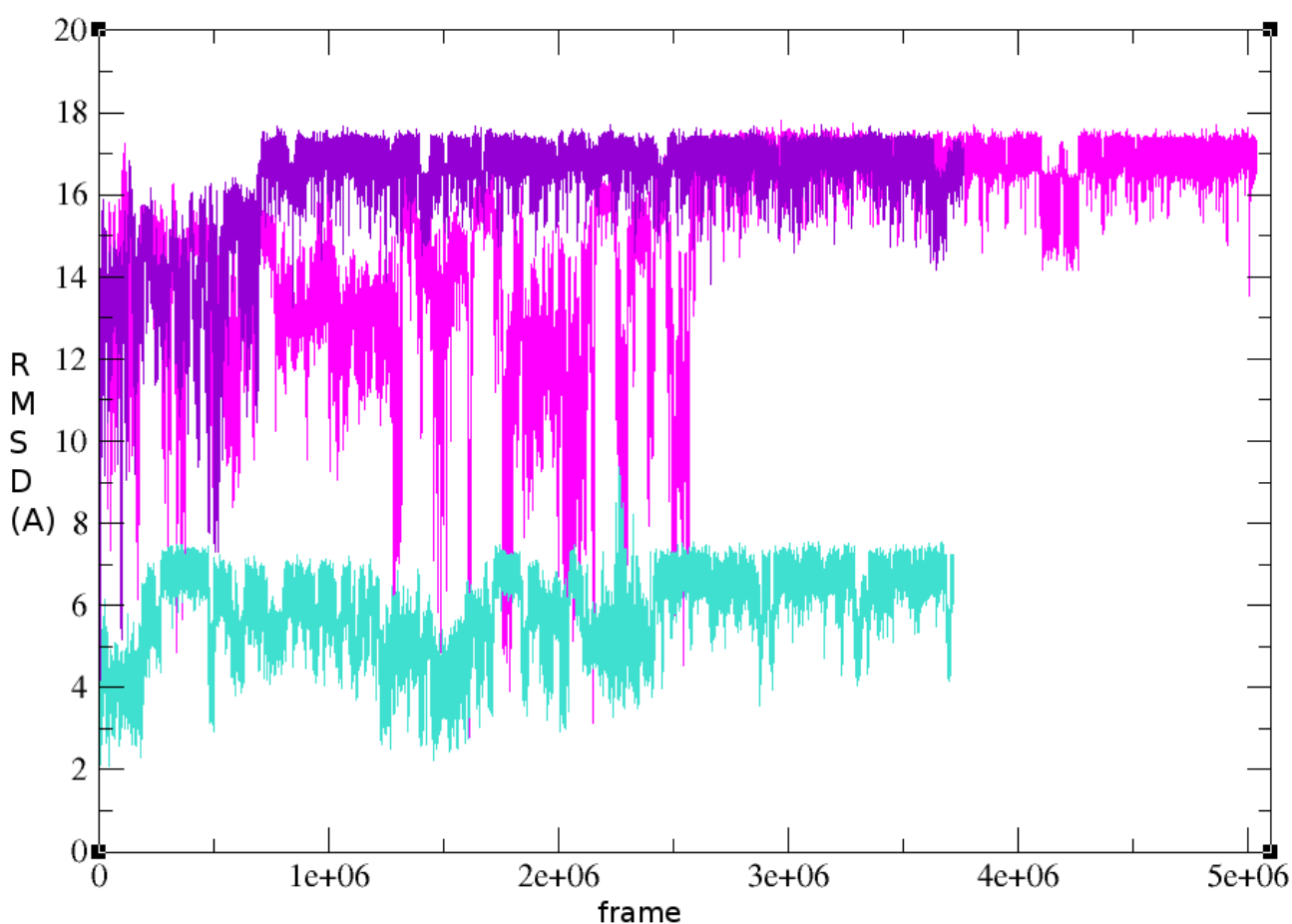


Εικόνα 4.3

Γραφική παράσταση του RMSD για όλα τα άτομα Ca της αρχικής δομής σε σχέση με τα άτομα Ca των δομών ολόκληρου του τροχιακού της προσομοίωσης protonated.

Με μια πρώτη ματιά κοιτάζοντας μόνο τα αποτελέσματα από το RMSD θα μπορούσαμε να πούμε πως υπάρχουν γεγονότα αναδίπλωσης σε όλα τα τροχιακά, καθώς υπάρχει σταθερότητα στις τιμές του RMSD έστω και για μικρό χρονικό διάστημα και στα τρία τροχιακά. Συγκριτικά μπορούμε να πούμε πως το τροχιακό protonated έχει αυξημένη σταθερότητα σε σχέση με τα άλλα δύο τροχιακά καθώς και τις χαμηλότερες τιμές RMSD σε σχέση με την αρχική δομή. Οι τιμές RMSD του protonated κυμαίνονται μεταξύ 2 και 9 Å, ενώ οι τιμές RMSD για το 99SB κυμαίνονται μεταξύ 2.5 και 17 Å και του ILDN μεταξύ 5 και 17 Å. Το τροχιακό ILDN έχει αυξημένη μεταβολή του RMSD στην αρχή της

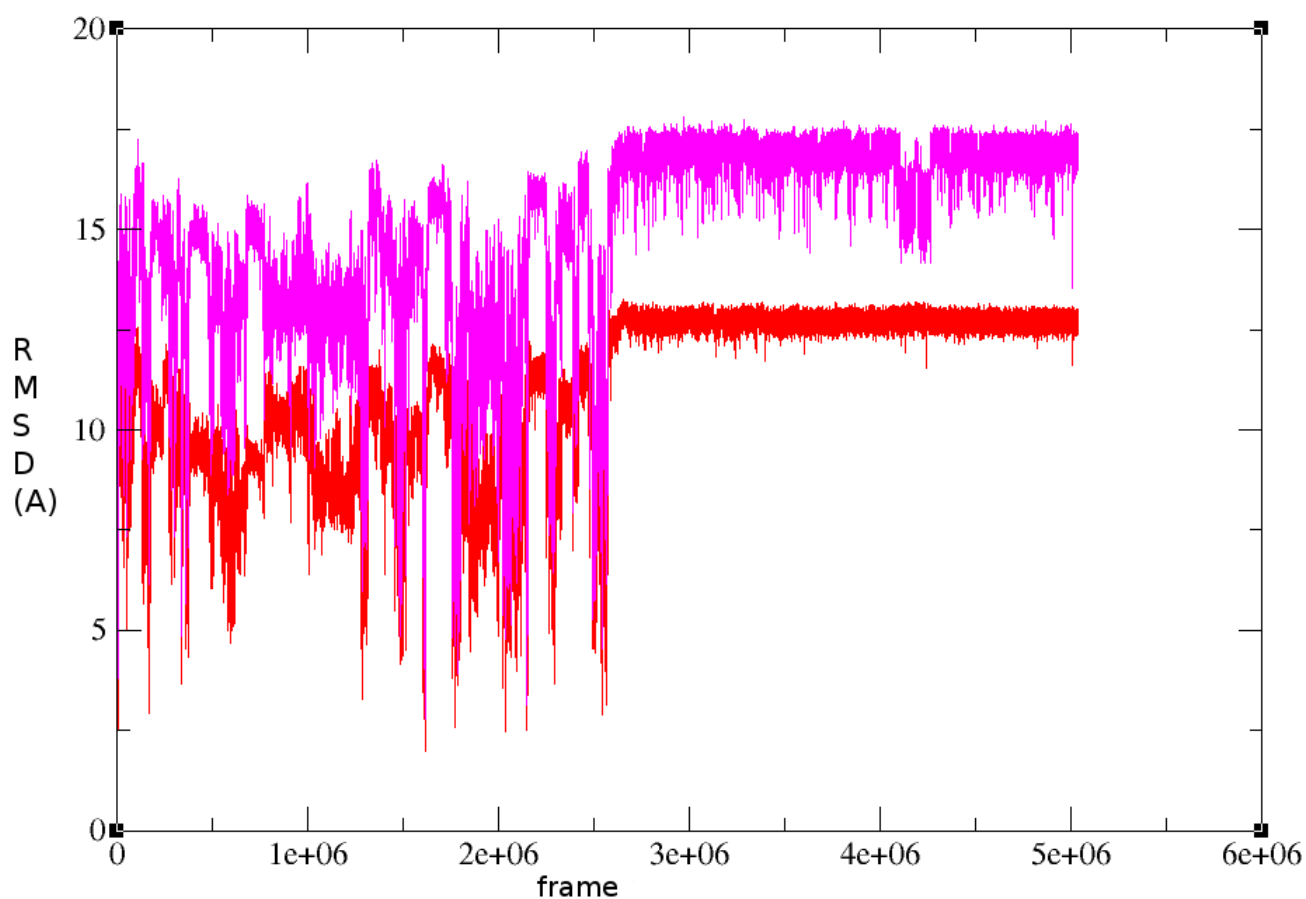
προσομοίωσης, πιθανόν λόγω του heating κατά την έναρξη, αλλά στη συνέχεια εμφανίζονται μεγάλα διαστήματα στα οποία η δομή είναι σταθερή. Θα μπορούσαμε να πούμε πως το ILDN τροχιακό παρουσιάζει τα μεγαλύτερα “τμήματα σταθερότητας” σε σχέση με τα άλλα δύο τροχιακά. Συγκριτικά το RMSD από τα τρία τροχιακά φαίνεται στην εικόνα 4.4. Οι προσομοιώσεις δεν έχουν το ίδιο μέγεθος και οι δομές προφανώς δεν είναι οι ίδιες, αλλά αυτό που θέλουμε να εξετάσουμε σε αυτό το σημείο από τη σύγκρισή του RMSD των τροχιακών είναι συγκριτική η κινητικότητα του κάθε τροχιακού και όχι οι δομές αυτές καθ' αυτές.



Εικόνα 4.4

Γραφική παράσταση που απεικονίζει το RMSD των ατόμων Ca των τροχιακών που παρήχθησαν. Με ροζ απεικονίζεται το τροχιακό της 99SB προσομοίωσης, με βιολετί το τροχιακό της ILDN προσομοίωσης και με κίτρινο το τροχιακό της protonated.

Όπως έχει παρατηρηθεί και από προηγούμενες προσομοιώσεις τα άκρα της β-φουρκέτας που σχηματίζει το πεπτίδιο παρουσιάζουν μεγάλη κινητικότητα (N. Vaidehi et al., 2011), για το λόγο αυτό μετρήθηκαν οι τιμές RMSD που παρουσιάζουν τα άτομα Ca των καταλοίπων 3 έως 15 (αφαιρέθηκαν από τους υπολογισμούς τα δύο πρώτα και τα δύο τελευταία κατάλοιπα). Το γεγονός αυτό φαίνεται και από την εικόνα 4.5, όπου στο τροχιακό της 99SB προσομοίωσης το εύρος των τιμών RMSD για τα άτομα Ca των καταλοίπων 3-15 είναι 2.5 – 13 Å. Επιπλέον παρατηρείται αυξημένη σταθερότητα στο δεύτερο μισό του τροχιακού σε σχέση με το RMSD όλων των ατόμων Ca το ίδιου τροχιακού.

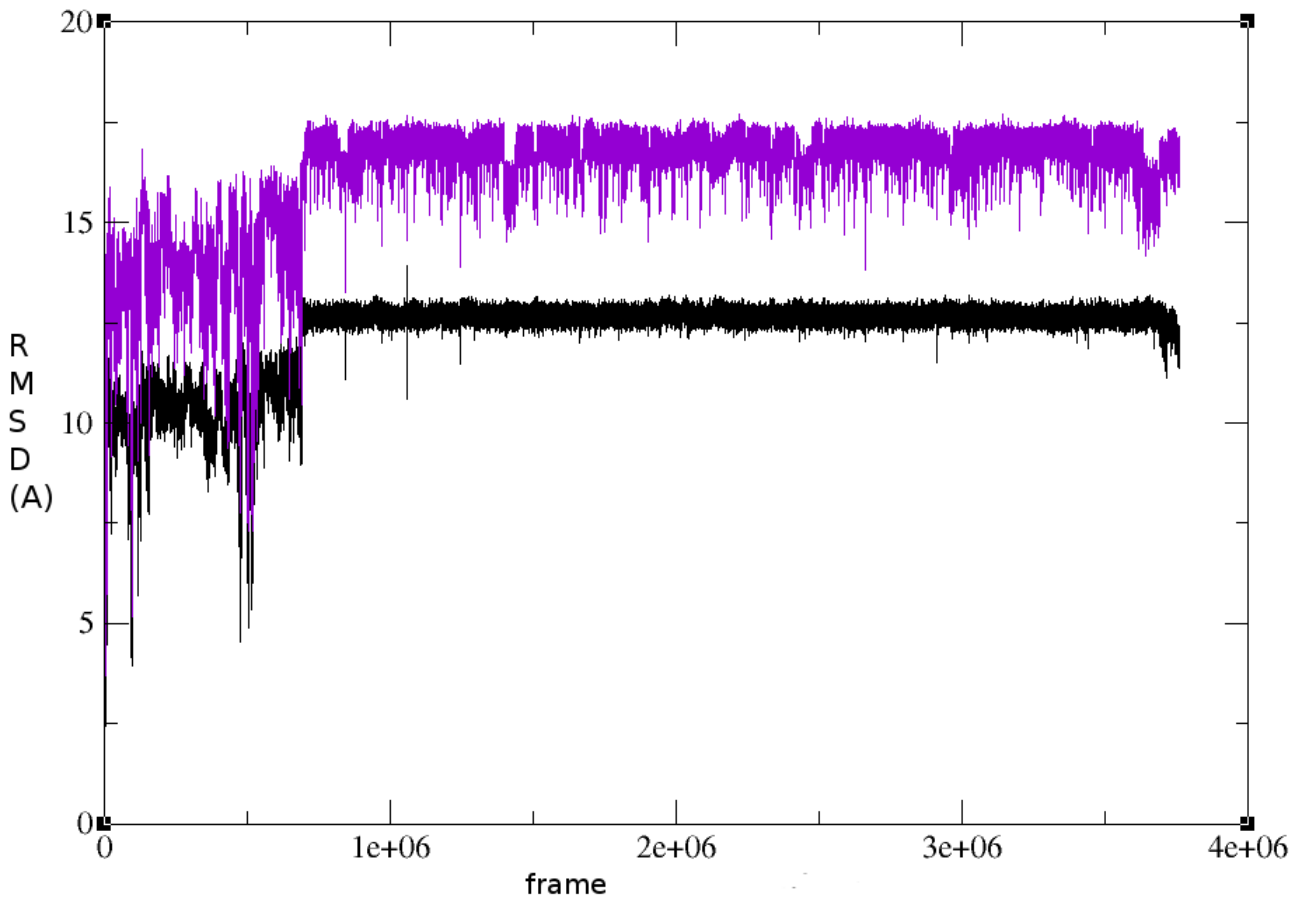


Εικόνα 4.5

Γραφική παράσταση του τροχιακού 99SB, όπου γίνεται σύγκριση του RMSD των ατόμων Ca για τα κατάλοιπα 1-17 (με ροζ χρώμα) και του RMSD των ατόμων Ca για τα κατάλοιπα 3-15 (με κόκκινο χρώμα). Είναι εμφανής η σταθερότητα που

παρουσιάζει το δεύτερο μισό του τροχιακού στα κατάλοιπα 3-15, καθώς και η μικρότερη διακύμανση του RMSD του οποίου το εύρος τιμών είναι μικρότερο σε σχέση με τη διακύμανση που παρουσιάζει ολόκληρο το πεπτίδιο.

Παρόμοιο αποτέλεσμα με το τροχιακό 99SB έχουμε και στο τροχιακό της προσομοίωσης ILDN. Η διακύμανση του RMSD των ατόμων Ca των καταλοίπων 3-15 είναι επίσης χαμηλότερη σε σχέση με τη διακύμανση του RMSD όλων των ατόμων Ca του πεπτιδίου. Επίσης η μέγιστη τιμή RMSD από την αρχική δομή είναι χαμηλότερη από τη μέγιστη τιμή RMSD ολόκληρου του πεπτιδίου σε σχέση με την αρχική δομή. Το εύρος των τιμών RMSD για τα άτομα Ca των καταλοίπων 3-15 κυμαίνεται μεταξύ 2.5 και 13 Å περίπου. Βλέποντας την εικόνα 4.6 μπορούμε να πούμε πως τα κατάλοιπα 3-15 του τροχιακού ILDN είναι πιο σταθερά κατά τη διάρκεια της προσομοίωσης σε σχέση με ολόκληρο το πεπτίδιο.

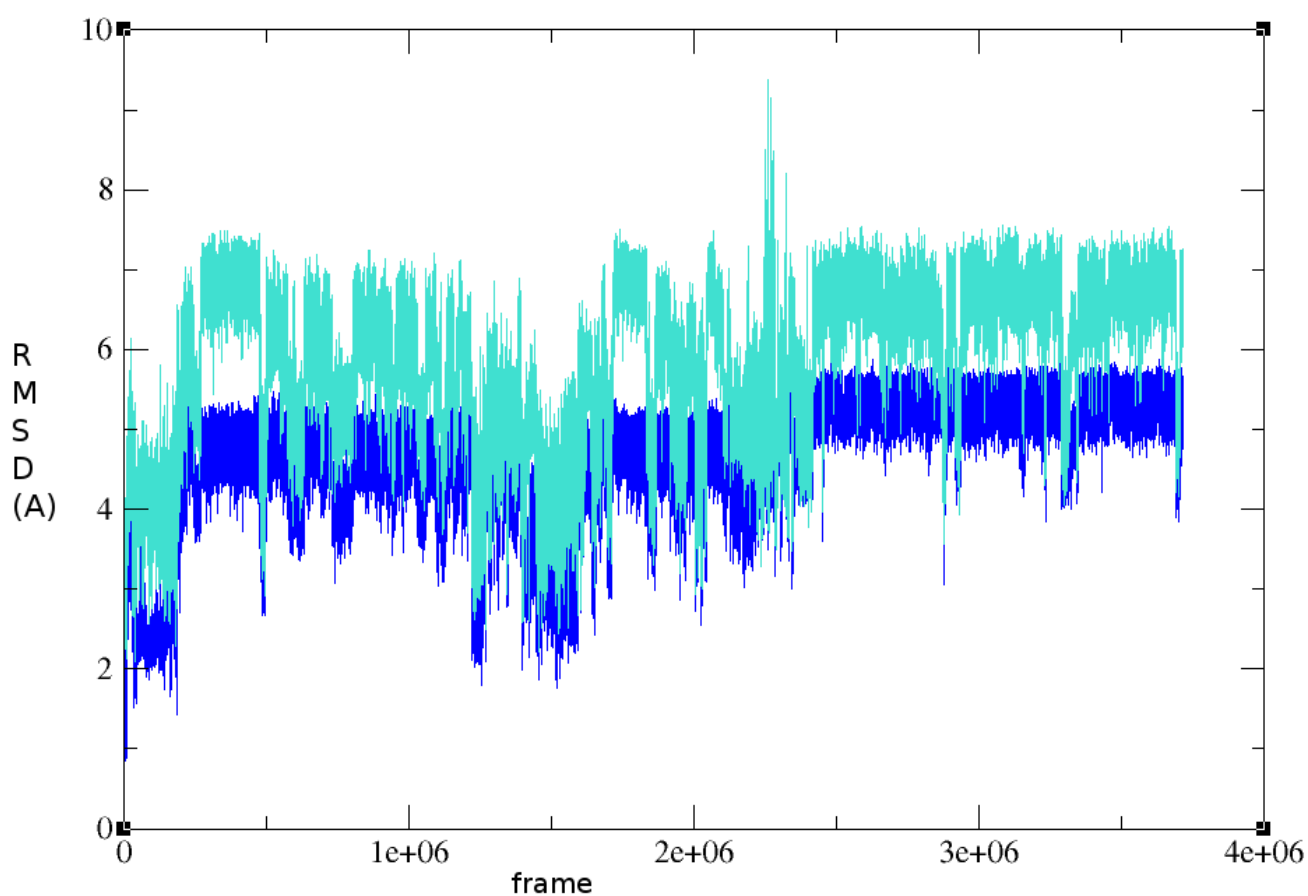


Εικόνα 4.6

Γραφική παράσταση του τροχιακού ILDN, όπου γίνεται σύγκριση του RMSD των ατόμων Ca για τα κατάλοιπα 1-17 (με βιολετί χρώμα) και του RMSD των ατόμων Ca για τα κατάλοιπα 3-15 (με μαύρο χρώμα). Είναι εμφανής η αυξημένη σταθερότητα που παρουσιάζουν τα κατάλοιπα 3-15 μετά από μικρό χρονικό διάστημα από την έναρξη της προσομοίωσης, καθώς και η μικρότερη διακύμανση του RMSD του οποίου το εύρος τιμών είναι μικρότερο σε σχέση με όλα τα κατάλοιπα του τροχιακού.

Την ίδια συμπεριφορά με τα δύο προηγούμενα τροχιακά παρουσιάζει και το τροχιακό της προσομοίωσης protonated. Η σύγκριση των RMSDs των ατόμων Ca για τα κατάλοιπα 1-17 και 3-15 μας δείχνει πως το εύρος των τιμών που παίρνει το RMSD μικραίνει για τα κατάλοιπα 3-15, σε σχέση με το εύρος των τιμών RMSD ολόκληρου του πεπτιδίου. Επιπλέον και σε αυτό το τροχιακό

μειώνεται το μέγιστο RMSD από την αρχική δομή, όπως και στις δύο προηγούμενες προσομοιώσεις. Συνεπώς τα κατάλοιπα 3-15 παρουσιάζουν μεγαλύτερη σταθερότητα συγκριτικά με ολόκληρο το πεπτίδιο στην προσομοίωση protonated (εικόνα 4.7). Και από τις τρεις προσομοιώσεις είναι εμφανές πως το τμήμα 3-15 του πεπτιδίου είναι πιο σταθερό από ολόκληρο το πεπτίδιο στις συνθήκες που εξετάστηκε. Είναι πιθανόν οι αναμενόμενοι δεσμοί υδρογόνου της περιοχής των άκρων της φουρκέτας να δημιουργούνται και να διασπώνται συχνά, λόγω της μεγάλης κινητικότητας των άκρων.



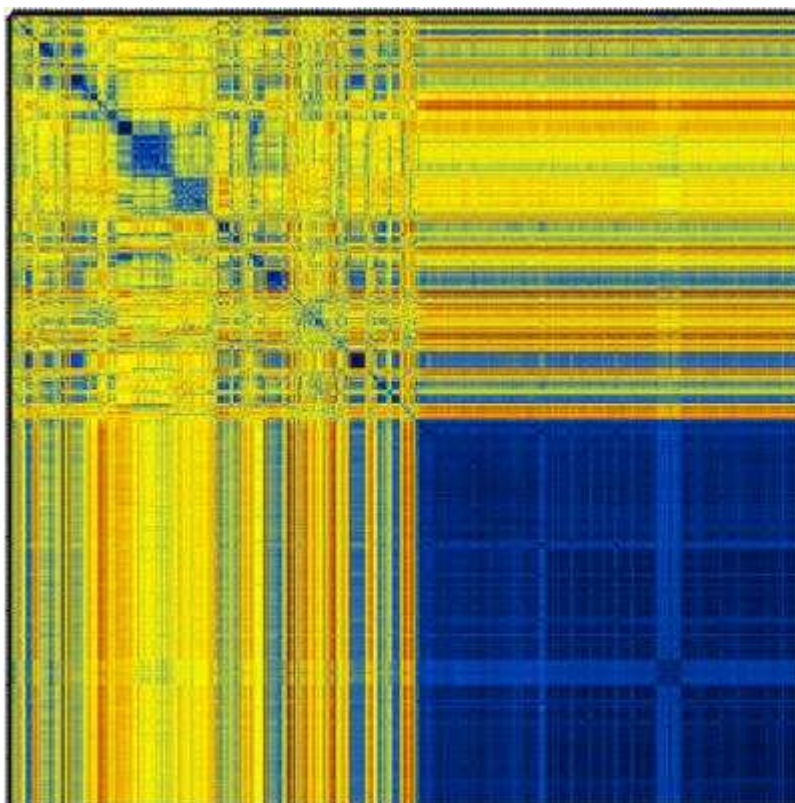
Εικόνα 4.7

Γραφική παράσταση του τροχιακού protonated, όπου γίνεται σύγκριση του RMSD των ατόμων Ca για τα κατάλοιπα 1-17 (με κυανό χρώμα) και του RMSD των ατόμων Ca για τα κατάλοιπα 3-15 (με μπλε χρώμα). Είναι εμφανής η μικρότερη διακύμανση του RMSD του οποίου το εύρος τιμών είναι μικρότερο σε σχέση με

όλα τα κατάλοιπα του τροχιακού. Τα κατάλοιπα 3-15 παρουσιάζουν μεγαλύτερη σταθερότητα προς το τέλος της προσομοίωσης protonated, με τιμή RMSD κοντά στο 5 σε σχέση με την αρχική δομή.

Ένας δεύτερος εύχρηστος τρόπος παρουσίασης των αποστάσεων επιλεγμένων ατόμων μεταξύ διαφορετικών δομών του ίδιου τροχιακού είναι οι πίνακες RMSD. Με τους πίνακες RMSD έχουμε τη δυνατότητα να συγκρίνουμε την απόσταση οποιουδήποτε ατόμου (ατόμων Ca, Backbone, Not hydrogen κ.λπ) μεταξύ όλων των δομών που υιοθετούνται κατά τη διάρκεια της προσομοίωσης και όχι μόνο μίας δομής με όλες τις υπόλοιπες, όπως έγινε με τα διαγράμματα RMSD. Έτσι μπορούμε να έχουμε μια πιο ξεκάθαρη εικόνα της κίνησης κατά τη διάρκεια της προσομοίωσης. Ο πίνακας RMSD είναι ένα τετράγωνο ανάλογο του μήκους της εκάστοτε προσομοίωσης. Στους δύο άξονες τοποθετούνται οι δομές που θέλουμε να συγκρίνουμε με αύξουσα σειρά, ξεκινώντας από την πρώτη δομή και τελειώνοντας με την τελευταία δομή του τροχιακού. Σε περίπτωση που το τροχιακό συμπεριλαμβάνει πολύ μεγάλο αριθμό δομών υπάρχει η δυνατότητα χρήσης βήματος στη δημιουργία του πίνακα, προκειμένου να μικρύνει το μέγεθός του. Η βαθμολόγηση του RMSD μεταξύ των δομών γίνεται σε χρωματική κλίμακα. Η διαγώνιος του πίνακα, η οποία αντιστοιχεί στην απόσταση εαυτών ατόμων έχει πάντα το χρώμα που αντιστοιχεί στο ελάχιστο RMSD. Καθώς αυξάνει το RMSD των ατόμων αλλάζει και η απόχρωση στον πίνακα. Οι πίνακες που θα παρουσιασθούν στη συνέχεια κατασκευάσθηκαν χρησιμοποιώντας το πρόγραμμα CARMA και έχουν κλίμακα από σκούρο μπλε (ελάχιστο RMSD) έως κόκκινο (μέγιστο RMSD). Αυτό που μας ενδιαφέρει σε έναν πίνακα RMSD είναι η παρατήρηση των χαμηλότερων τιμών κατά τη διάρκεια της προσομοίωσης, έτσι ώστε να δούμε τη σταθερότητα των δομών στο χρόνο. Πρακτικά αυτό φαίνεται από το σχηματισμό μπλε τετραγώνων στη διαγώνιο γραμμή, όσο μεγαλύτερα είναι τα τετράγωνα τόσο περισσότερο παραμένει σταθερή μια δομή. Επιπλέον με τον πίνακα RMSD μπορούμε να ελέγξουμε αν οι σταθερές δομές που σχηματίζονται κατά τη διάρκεια της

προσομοίωσης είναι όμοιες, αυτό γίνεται παρατηρώντας τις προβολές των δομών στα αντίστοιχα τετράγωνα. Επειδή όμως μια εικόνα αντιστοιχεί σε χίλιες λέξεις και οι περιγραφές δεν αρκούν άνευ σχήματος ας δούμε τον πίνακα RMSD της εικόνας 4.8, που περιγράφει την απόσταση των ατόμων Ca μεταξύ των διαμορφώσεων του τροχιακού 99SB.



Εικόνα 4.8

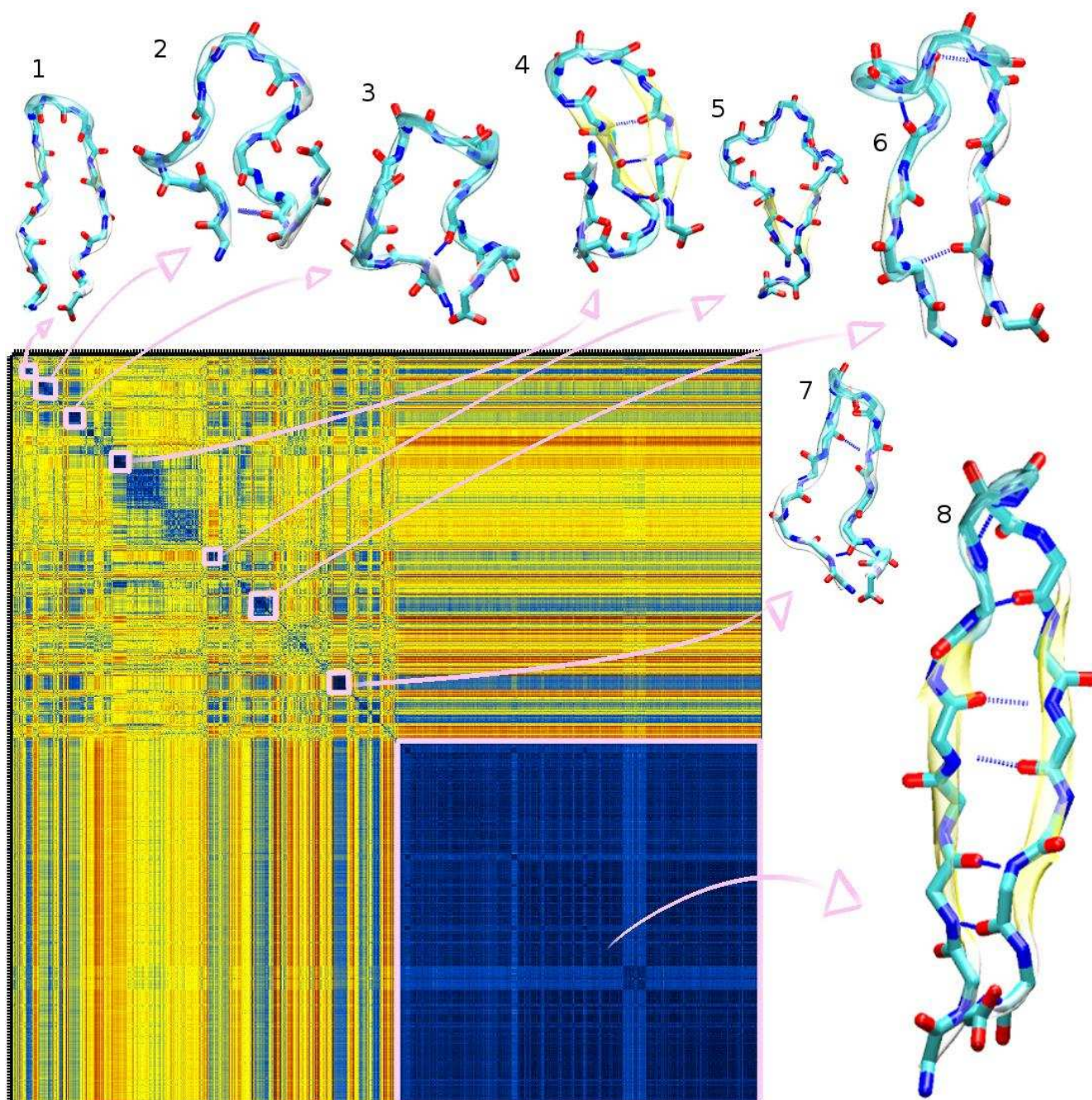
Πίνακας RMSD του τροχιακού 99SB. Είναι τοποθετημένες με αύξουσα σειρά οι διαμορφώσεις που σχηματίστηκαν κατά την προσομοίωση στον κάθετο και οριζόντιο άξονα, ξεκινώντας με την πρώτη δομή από επάνω και αριστερά, αντίστοιχα. Το συνολικό μήκος της προσομοίωσης αντιστοιχεί σε 5035100 διαμορφώσεις και χρόνο 4.03 μ s. Ο πίνακας δημιουργήθηκε με βήμα 1000 προκειμένου να μικρύνουμε το μέγεθος του. Το νέο μέγεθος του πίνακα είναι 5036 διαμορφώσεις. Η βαθμολόγηση της διαβάθμισης έγινε από 0Å (σκούρο μπλε) έως 13Å (κόκκινο). Κατά μήκος της διαγωνίου παρατηρούνται τα μπλε κουτιά που αντιστοιχούν σε δομές των οποίων τα Ca άτομα έχουν χαμηλές τιμές RMSD.

Φαίνεται να σχηματίζονται αρκετές δομές οι οποίες παραμένουν σταθερές για ορισμένο χρονικό διάστημα. Όπως και στην εικόνα 4.1 και εδώ στο δεύτερο μισό της προσομοίωσης η δομή φαίνεται να σταθεροποιείται μέχρι το τέλος.

Στην προσομοίωση 99SB η πιο κοντινή στη μέση δομή έχει τιμή RMSD 3.97 Å. Προκειμένου να δούμε τις δομές που παραμένουν σταθερές για κάποιο χρονικό διάστημα επιλέξαμε από κάθε ομάδα δομών (μπλε τετράγωνο) τη μέση δομή και συγκρίναμε με αυτή τις δομές που συμπεριλαμβάνονται στην ομάδα, έτσι ώστε να καταλήξουμε στην πιο κοντινή δομή με τη μέση δομή του κάθε τετραγώνου. Στην εικόνα 4.9 παρουσιάζονται οι αντιπροσωπευτικές δομές για την κάθε ομάδα δομών. Συνολικά επιλέχθηκαν οκτώ δομές. Το RMSD των δομών αυτών είναι αρκετά χαμηλό σε σχέση με το RMSD της μέσης δομής από την περιοχή που προέρχονται. Η πρώτη δομή έχει τιμή RMSD σε σχέση με τη μέση 0.46Å και υιοθετήθηκε στα 0.091μs. Παρόλο που έχει σχηματιστεί η στροφή και το σχήμα της μοιάζει με φουρκέτα, δε φέρει στοιχεία δευτεροταγούς δομής καθώς δεν έχει σχηματισμένους δεσμούς υδρογόνου. Η δεύτερη δομή έχει τιμή RMSD σε σχέση με τη μέση 1.14 Å και υιοθετήθηκε στα 0.18 μs. Στη δομή αυτή δεν υφίσταται κανένα στοιχείο δευτεροταγούς δομής, αλλά έχει σχηματιστεί ένα δεσμός υδρογόνου μεταξύ Gln2(N) - Thr14(O). Η τρίτη δομή έχει τιμή RMSD 0.56 Å και υιοθετείται στα 0.35 μs. Και αυτή η διαμόρφωση δεν φέρει κανένα στοιχείο δευτεροταγούς δομής, παρόλο που έχουν σχηματιστεί δυο δεσμοί υδρογόνου. Ο πρώτος δεσμός σχηματίστηκε μεταξύ Met1(N) – Val17(O) και ο δεύτερος μεταξύ Gln2(N) – Ile13(O), οι οποίοι δεν ανήκουν στους δεσμούς που έχει η φυσική δομή. Η τέταρτη δομή έχει τιμή RMSD 0.50 Å από τη μέση δομή της περιοχής της και υιοθετήθηκε στα 0.60 μs. Σε αυτή τη διαμόρφωση έχουν αρχίσει να σχηματίζονται β-δομές και έχουν σχηματιστεί και δύο δεσμοί υδρογόνου στο σκελετό του πεπτιδίου. Ο πρώτος σχηματίστηκε μεταξύ των καταλοίπων Thr7(O) – Glu16(N) και ο δεύτερος μεταξύ των καταλοίπων Asp9(N) – Thr14(O), επιπλέον έχουν σχηματιστεί και τρεις δεσμοί υδρογόνου μεταξύ των πλευρικών αλυσίδων αυτοί οι δεσμοί αφορούν τα κατάλοιπα (a) Gln2(HE2) –

Asp9(OD2), (b) Lys6(HZ3) – Val17(OXT) και (c) Asp9(OD1) – Thr12(HG1). Στη διαμόρφωση αυτή το μεταλλαγμένο κατάλοιπο Asp9 σχηματίζει συνολικά τρεις δεσμούς υδρογόνου. Η πέμπτη δομή έχει τιμή RMSD 0.73 Å από τη μέση δομή της περιοχής που παραμένει σταθερή και υιοθετείται στα 1.07 μs. Στη διαμόρφωση αυτή παρατηρείται μια μικρή περιοχή του πεπτιδίου να αποκτά β-δομή που συμπεριλαμβάνει τα κατάλοιπα Gln2, Ile3, Ile13, Thr14 και είναι εμφανής ο έντονα στρεβλωμένος πεπτιδικός σκελετός. Επιπλέον σχηματίζονται δύο δεσμοί υδρογόνου στον σκελετό του πεπτιδίου μεταξύ των καταλοίπων Gln2(O) – Thr14(N) και Asp9(O) – Thr12(N) και ένας δεσμός υδρογόνου μεταξύ των πλευρικών ομάδων των καταλοίπων Thr14(HG1) - Glu16(OE2). Η έκτη διαμόρφωση έχει τιμή RMSD 0.50 Å από τη μέση δομής της περιοχής στην οποία ανήκει και υιοθετήθηκε στα 1.36 μs. Στην έκτη διαμόρφωση δεν παρατηρείται κανένα στοιχείο δευτεροταγούς δομής αν και έχουν σχηματισθεί τρεις δεσμοί υδρογόνου στον πεπτιδικό σκελετό, ο οποίος είναι αρκετά παραμορφωμένος. Οι δεσμοί υδρογόνου που σχηματίστηκαν στον πεπτιδικό σκελετό αφορούν τα κατάλοιπα (a) Ile3(N) – Leu15(O), (b) Val5(O) – Asp9(N) και (c) Asp9(O) – Thr12(N), ενώ έχει δημιουργηθεί και ένας δεσμός μεταξύ της πλευρικής αλυσίδας του καταλοίπου Thr14(OG1) και του καταλοίπου Lys6(H). Η έβδομη διαμόρφωση έχει τιμή RMSD 0.34 Å από τη μέση δομή της περιοχής από την οποία προέρχεται και υιοθετήθηκε στα 1.78 μs. Ούτε σε αυτή τη δομή βλέπουμε να υφίστανται στοιχεία δευτεροταγούς δομής, καθώς αποτελείται εξ' ολοκλήρου από turn και coil διαμορφώσεις. Παρόλα αυτά έχουν σχηματισθεί δύο δεσμοί υδρογόνου στον πεπτιδικό σκελετό. Οι δεσμοί αυτοί είναι (a) Gln2(N) – Thr14(O) και (b) Lys6(O) – Thr12(N). Επιπροσθέτως δημιουργήθηκαν και δύο δεσμοί υδρογόνου μεταξύ πλευρικών ομάδων και ατόμων του σκελετού. Οι δεσμοί αυτοί είναι (a) Lys6(H) – Thr12(OG1) και (b) Thr7(O) – Lys11(HZ2). Η τελευταία δομή, από τις περιοχές που επιλέχθηκαν ως οι σταθερότερες του τροχιακού 99B, έχει τιμή RMSD 0.46 Å από τη μέση δομή της περιοχής από την οποία προέρχεται και υιοθετήθηκε στα 3.8 μs. Η δομή αυτή φαίνεται να υιοθετεί την διαμόρφωση της β-φουρκέτας, καθώς έχει

σχηματισθεί τόσο η β-στροφή όσο και οι β-κλώνοι της φουρκέτας. Η στροφή αποτελείται από τα κατάλοιπα Thr7 – Leu8 – Asp9 – Gly10. Η Gly είναι στην τέταρτη θέση στην αλληλουχία της στροφής πράγμα που σημαίνει πως πρόκειται για τον τύπο I β-στροφής, που ταυτοποιήθηκε στην NMR δομή και στις προηγούμενες μελέτες αναδίπλωσης του πεπτιδίου (βλ. Κεφάλαιο 1). Πέρα από τον σωστό τύπο στροφής η όγδοη διαμόρφωση, που είναι η κοντινότερη στη μέση δομή του μεγάλου μπλε τετραγώνου, έχει τις σωστές θέσεις δεσμών υδρογόνου μεταξύ των άκρων της φουρκέτας, σύμφωνα με την NMR δομή. Οι δεσμοί είναι οι εξής: (1) Ile3(N) – Leu15(O), (2) Ile3(O) – Leu15(N), (3) Val5(N) – Ile13(O), (4) Val5(O) – Ile13(N), (5) Thr7(N) – Lys11(O), (6) Thr7(O) – Gly10(N). Ο μόνος δεσμός υδρογόνου που δεν σχηματίζεται στη δομή αυτή είναι ο Met1(O) – Val17(N), ο οποίος ανήκει στα κατάλοιπα των άκρων της φουρκέτας. Τα κατάλοιπα των άκρων, όπως έχει ήδη αναφερθεί παρουσιάζουν έντονη κινητικότητα, πιθανώς για το λόγο αυτό να μην σχηματίζεται ο έβδομος δεσμός υδρογόνου όπως θα έπρεπε σύμφωνα με την NMR δομή. Σε προηγούμενες μελέτες έχουν βρεθεί οι έξι δεσμοί υδρογόνου μερικοί εκ των οποίων είχαν αρκετή απόκλιση από την NMR δομή (N. Vaidehi et al., 2011), αλλά σε καμία μελέτη του πεπτιδίου έως τώρα δεν έχει αναφερθεί ο έβδομος δεσμός υδρογόνου, μεταξύ Thr7(O) – Gly10(N), ο οποίος βρίσκεται στα κατάλοιπα της β-στροφής. Τέλος στη δομή αυτή παρατηρείται και η αλληλεπίδραση μεταξύ των καταλοίπων Asp9(OD1) και Thr7(HG1). Οι επαφές που δημιουργεί το μεταλλαγμένο κατάλοιπο πιθανολογείται πως είναι υπεύθυνες για την αυξημένη σταθερότητα του πεπτιδίου 1E0Q σε σχέση με το φυσικό πεπτίδιο. Σε προηγούμενες προσομοιώσεις, καθώς και στο NMR πείραμα οι επαφές που δημιουργεί το ένατο κατάλοιπο με την πλευρική αλυσίδα της Lys11 θεωρούνται υπεύθυνες για τη σταθερότητα της β-φουρκέτας, χωρίς όμως να υπάρχουν ισχυρές ενδείξεις (R. Zerella et al., 2000, N. Vaidehi et al., 2011).

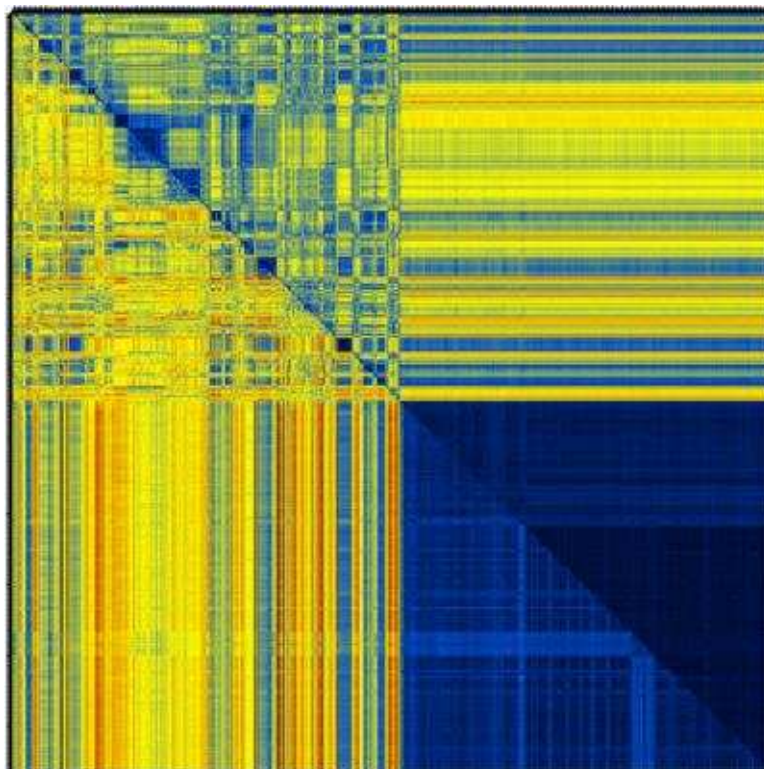


Εικόνα 4.9

Πίνακας RMSD του τροχιακού 99SB. Παρουσιάζονται οι δομές του τροχιακού που φαίνεται να παραμένουν σταθερές για χρονικό διάστημα που είναι μεγαλύτερο του 5% του συνολικού χρόνου της προσομοίωσης . Κάθε δομή ανήκει σε μια διαφορετική μπλε περιοχή και είναι συνδεδεμένη με αυτή με βέλος. Ξεκινώντας από επάνω αριστερά και κατευθυνόμενοι με τη φορά που ρολογιού βλέπουμε την εξέλιξη του τροχιακού 99SB και χρονικά. Η όγδοη είναι η τελική δομή του πεπτιδίου και η αντιπροσωπευτική του τροχιακού 99SB, καθώς παραμένει

σταθερή για μεγάλο χρονικό διάστημα μέχρι και το τέλος της προσομοίωσης. Χαμηλό RMSD σε σχέση με την όγδοη δομή έχουν οι δομές 1, 6 και 7. Οι δομές σχεδιάστηκαν με το πρόγραμμα VMD. Παρουσιάζεται μόνο ο σκελετός του πεπτιδίου χωρίς τα υδρογόνα (C – κυανό, N – μπλε, O – κόκκινο) και οι δεσμοί υδρογόνου μεταξύ των ατόμων του σκελετού (μπλε διακεκομμένη γραμμή). Επιπλέον σε όσες δομές έχει σχηματιστεί β-δομή αναπαρίσταται με κίτρινο βέλος. Το μέγεθος των δομών αναλογεί στο χρόνο παραμονής της δομής στην αντίστοιχη διαμόρφωση.

Για να επαληθευτεί η σταθερότητα των δομών που σχηματίστηκαν στο τροχιακό 99SB δημιουργήσαμε και έναν πίνακα RMSD στον οποίο το κάτω τμήμα της διαγωνίου αφορά τα άτομα Ca όλων των καταλοίπων και το επάνω τμήμα της διαγωνίου τα άτομα Ca των καταλοίπων 3-15 (εικόνα 4.10). Με τη σύγκριση αυτή βλέπουμε πως στο κομμάτι του πεπτιδίου που δεν συμπεριλαμβάνονται τα άκρα τα άτομα Ca είναι πιο σταθερά κατά τη διάρκεια της προσομοίωσης και οι δομές που σχηματίζονται έχουν χαμηλότερες τιμές RMSD και συνεπώς μοιάζουν περισσότερο. Πιθανώς έτσι ενισχύεται η άποψη ότι ο δεσμός υδρογόνου μεταξύ των ατόμων Met1(O) – Val17(N) δεν σχηματίζεται λόγω της αυξημένης κινητικότητας των άκρων.



Εικόνα 4.10

Πίνακας RMSD του τροχιακού 99SB. Σύγκριση των τιμών RMSD για τα άτομα Cas των καταλοίπων 3-15 (επάνω) και για τα άτομα Cas όλων των καταλοίπων (κάτω). Το μπλε χρώμα είναι πιο έντονο στη δεξιά πλευρά (χαμηλότερο RMSD) και οι διαμορφώσεις φαίνεται να συμφωνούν περισσότερο μεταξύ τους, σε σχέση με την αριστερή πλευρά. Η χρωματική κλίμακα διατηρείται σταθερή μεταξύ 0 και 12Å σε όλους του χάρτες RMSD.

Συμπερασματικά, μπορούμε να πούμε πως στην προσομοίωση 99SB η φυσικού – τύπου δομή σχηματίστηκε μετά τα πρώτα 1.78 μs, συνεπώς φαίνεται πως το δυναμικό πεδίο AMBER99SB χρειάζεται αρκετό υπολογιστικό (και φυσικό χρόνο) μέχρι να υιοθετηθεί η φυσικού – τύπου δομή της β-φουρκέτας.

Ο πίνακας RMSD των ατόμων Ca του τροχιακού ILDN παρουσιάζεται στην εικόνα 4.11. Στον πίνακα αυτό τα πράγματα είναι πιο ξεκάθαρα για την κινητικότητα των ατόμων Ca κατά τη διάρκεια της προσομοίωσης ILDN,

καθώς η τελική δομή αποκτάται πιο γρήγορα σε σχέση με την προσομοίωση 99SB και παραμένει σταθερή για πολύ μεγάλο χρονικό διάστημα. Με εξαίρεση τα κατάλοιπα των άκρων η υπόλοιπη δομή δεν αλλάζει σημαντικά για μεγάλο χρονικό διάστημα. Ας πάρουμε όμως τις διαμορφώσεις, που παρατηρήσαμε πως παραμένουν σταθερές για ορισμένο σεβαστό χρονικό διάστημα, από την αρχή της προσομοίωσης. Η προσομοίωση ξεκίνησε με το πεπτίδιο σε πλήρως εκτεταμένη διαμόρφωση ακολούθησαν τα βήματα του heating και της εξισορρόπησης και στην παραγωγική φάση άρχισαν να σχηματίζονται οι πρώτες διαμορφώσεις που θυμίζουν δομή φουρκέτας. Το πρώτο τμήμα όπου φαίνεται να σταθεροποιείται για μικρό χρονικό διάστημα η δομή βρίσκεται πολύ νωρίς στην προσομοίωση και η δομή που έχει το χαμηλότερο RMSD συγκριτικά με τη μέση δομή του τμήματος αυτού υιοθετήθηκε μόλις στα 0.13 μ s με τιμή RMSD 0.6 Å. Στη δομή αυτή σχηματίζεται β -διαμόρφωση σε μία μικρή περιοχή που συμπεριλαμβάνει τα κατάλοιπα Lys6, Thr7, Thr14 και Leu15. Επιπλέον φέρει δύο δεσμούς υδρογόνου μεταξύ των καταλοίπων (a) Lys6(O) – Leu15(N) και (b) Leu8(N) – Ile13(O), οι οποίοι δεν υφίστανται στη φυσικού-τύπου δομή. Η δεύτερη δομή που αντιστοιχεί στο δεύτερο σταθερό διάστημα έχει τιμή RMSD 1.3 Å από τη μέση δομής της περιοχής που απομονώθηκε και υιοθετήθηκε στα 0.19 μ s. Η δεύτερη δομή μοιάζει αρκετά με την πρώτη καθώς φέρει β -διαμόρφωση σε μια μικρή περιοχή του πεπτιδίου και φέρει και τον ίδιο δεσμό υδρογόνου μεταξύ των καταλοίπων Lys6(O) – Leu15(N). Εκτός από το δεσμό αυτό σχηματίζεται και ένας δεύτερος μεταξύ των καταλοίπων Lys6(N) – Leu15(O). Στη συνέχεια η β -διαμόρφωση χάνεται και το πεπτίδιο σχηματίζει διαμορφώσεις coil και loop. Η τρίτη δομή φέρει μόνο έναν δεσμό υδρογόνου μεταξύ των καταλοίπων Leu8(N) – Ile13(O), η τιμή RMSD από τη μέση δομή της περιοχής από την οποία προέρχεται είναι 1.5 Å και υιοθετήθηκε στα 0.28 μ s. Στην τέταρτη δομή τα άκρα της φουρκέτας απομακρύνονται και δεν υπάρχει κανένας δεσμός υδρογόνου μεταξύ των ατόμων του πεπτιδικού σκελετού. Η δομή αυτή έχει τιμή RMSD 1.8 Å από τη μέση δομή της περιοχής από την οποία προέρχεται. Υιοθετήθηκε στα 0.38 μ s και αποτελείται εξ' ολοκλήρου από coil

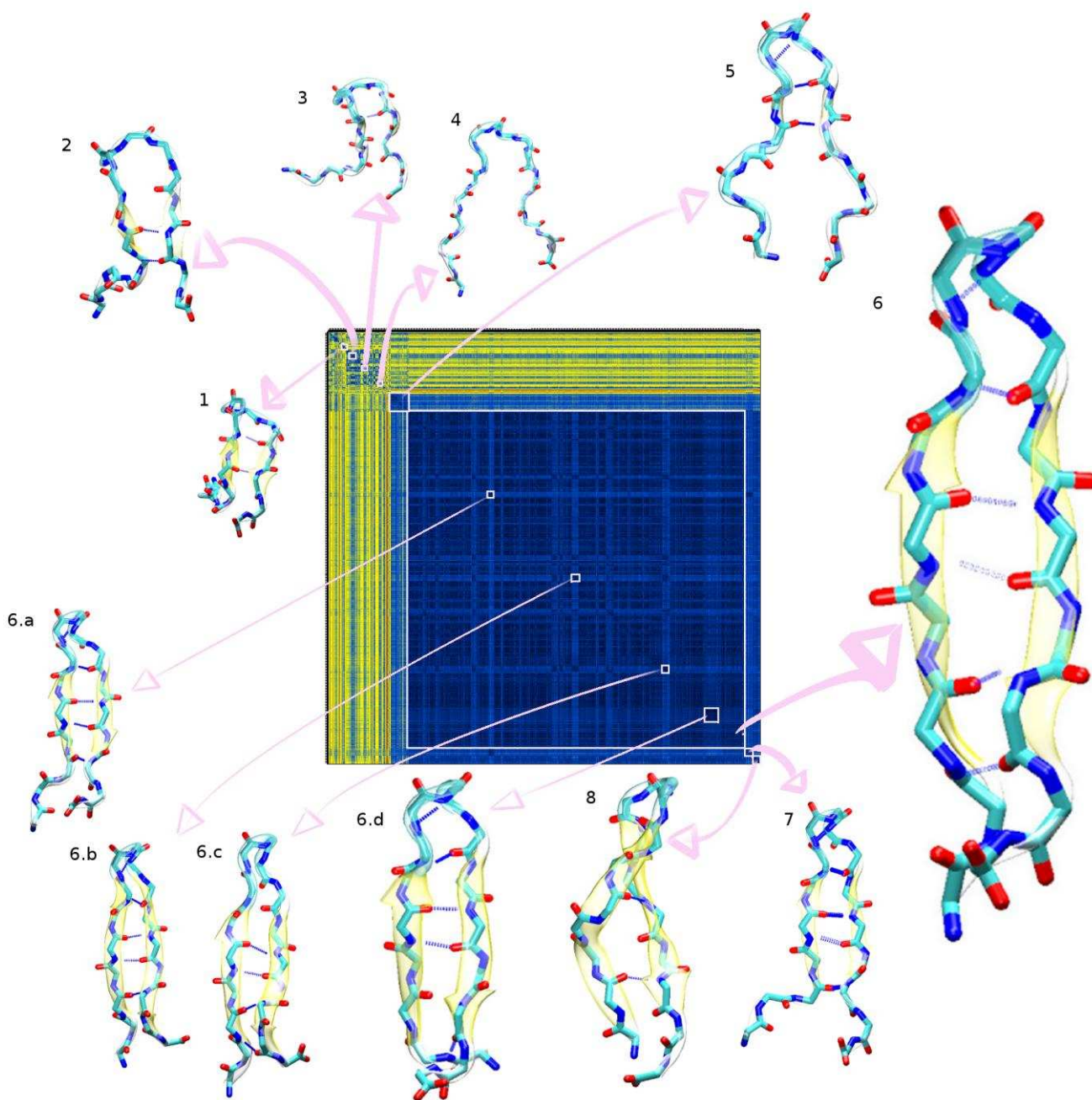
διαμόρφωση. Στην πέμπτη δομή εμφανίζονται και πάλι οι δεσμοί υδρογόνου στο σκελετό και σχηματίζεται η τύπου I β-στροφή με τη Gly10 στην τέταρτη θέση της στροφής. Η δομή αυτή έχει τιμή RMSD 1.0 Å από τη μέση δομή της περιοχής από την οποία επιλέχθηκε και υιοθετήθηκε στα 0.47 μs. Οι δεσμοί υδρογόνου που σχηματίζονται στα άτομα του σκελετού είναι οι εξής : (a) Val5(O) – Ile13 (N), (b) Thr7(N) – Lys11(O) και (c) Thr7(O) – Gly10(N). οι δεσμοί αυτοί έχουν ταυτοποιηθεί στην τελική δομή του τροχιακού 99SB αλλά και στην NMR δομή. Συμπεραίνουμε λοιπόν πως κατά την προσομοίωση με το ILDN δυναμικό πεδίο στα πρώτα 0.47 μs έχει σχηματιστεί ο σωστός τύπος β-στροφής και τρεις σωστοί δεσμοί υδρογόνου (σύμφωνα με την NMR δομή). Η έκτη διαμόρφωση, είναι η κυρίαρχη του τροχιακού ILDN, καθώς φαίνεται να παραμένει για μεγάλο χρονικό διάστημα. Το RMSD της έκτης δομής σε σχέση με τη μέση δομή της περιοχής από την οποία προέρχεται είναι 0.4 Å και υιοθετήθηκε στα 1.26 μs. Έχει διαμορφωθεί η β-φουρκέτα, η οποία έχει τον σωστό τύπο I β-στροφής και τους σωστούς δεσμούς υδρογόνου στο σκελετό. Οι δεσμοί που δημιουργήθηκαν είναι : (1) Ile3(N) – Leu15(O), (2) Ile3(O) – Leu15(N), (3) Val5(N) – Ile13(O), (4) Val5(O) – Ile13(N), (5) Thr7(N) – Lys11(O), (6) Thr7(O) – Gly10(N). Οι ίδιοι δεσμοί δημιουργήθηκαν και στην τελική δομή του τροχιακού 99SB. Και σε αυτή τη δομή λείπει ο δεσμός υδρογόνου, στα άκρα της φουρκέτας, μεταξύ των καταλοίπων Met1(O) – Val17(N). Από την περιοχή που προέρχεται η έκτη δομή παρουσιάζονται άλλες τέσσερις παρόμοιες διαμορφώσεις στην εικόνα 4.11 (6a, 6b, 6c, 6d). Οι δομές αυτές έχουν χαμηλό RMSD σε σχέση με τη μέση δομή της περιοχής και μοιάζουν αρκετά με την έκτη δομή. Και οι τέσσερις δομές σχηματίζουν τον σωστό τύπο β-στροφής και έχουν εκτεταμένη β-δομή. Οι δεσμοί υδρογόνου όλων των δομών που απομονώθηκαν από το τροχιακό ILDN φαίνονται στον πίνακα 4.1. Παρόλο που η έκτη διαμόρφωση φαίνεται να είναι η κυρίαρχη του τροχιακού απομονώθηκαν άλλες δύο δομές στη συνέχεια, διότι παρατηρήθηκε αλλαγή στο RMSD. Η έβδομη δομή υιοθετήθηκε στα 2.94 μs και έχει τιμή RMSD από τη μέση δομή της περιοχής από την οποία προέρχεται 0.5 Å. Στη δομή αυτή τα κατάλοιπα των άκρων έχουν απομακρυνθεί και πάλι με

συνέπεια τη διάσπαση του δεσμού Ile3(N) – Leu15(O) που υπήρχε στην προηγούμενη δομή. Η στροφή και οι υπόλοιποι δεσμοί παραμένουν ως έχει. Τέλος στην όγδοη διαμόρφωση το πεπτίδιο αποκτά μία έντονη συστροφή, με αποτέλεσμα την απώλεια δεσμών υδρογόνου και τη δημιουργία ενός “λανθασμένου” δεσμού, παρόλο που η περιοχή των β-κλώνων παραμένει η ίδια. Οι δεσμοί υδρογόνου στην όγδοη διαμόρφωση είναι :(a) Gln2(O) – Leu15(N), (b) Val5(O) – Ile13(N) και (c) Thr7(N) – Lys11(O). Ο δεσμός (a) δεν υφίσταται στην NMR δομή. Αυτό πιθανόν συμβαίνει λόγω της έντονης κάμψης που παρουσιάζει ο β-κλώνος του N-τελικού άκρου του πεπτιδίου.

	1	2	3	4	5	6	6a	6b	6c	6d	7	8
Met1(O) – Val17(N)	-	-	-	-	-	-	-	-	-	-	-	-
Ile3(N) – Leu15(O)	-	-	-	-	-	+	-	-	-	-	-	-
Ile3(O) – Leu15(N)	-	-	-	-	-	+	+	+	+	+	+	-
Val5(N) – Ile13(O)	-	-	-	-	-	+	+	+	+	+	+	-
Val5(O) – Ile13(N)	-	-	-	-	+	+	+	+	+	+	+	+
Thr7(N) – Lys11(O)	-	-	-	-	+	+	+	+	+	+	+	+
Thr7(O) – Gly10(N)	-	-	-	-	+	+	+	+	-	+	+	-

Πίνακας 4.1

Σημειώνονται ποιοι δεσμοί υδρογόνου της NMR δομής σχηματίζονται στις αντιπροσωπευτικές δομές που παρατηρήθηκαν στο τροχιακό της ILDN προσομοίωσης με χρόνο παραμονής μεγαλύτερο του 5% του συνολικού χρόνου προσομοίωσης. Η δομή έξι είναι η κυρίαρχη δομή του τροχιακού και φέρει τους περισσότερους δεσμούς υδρογόνου της NMR δομής.

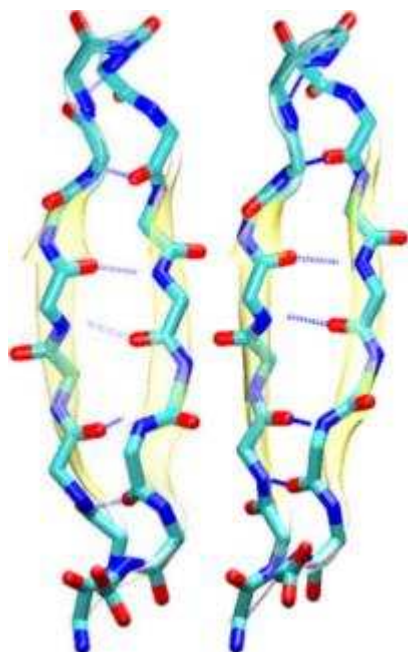


Εικόνα 4.11

Πίνακας RMSD του τροχιακού ILDN. Με αριθμητική σειρά παρουσιάζονται οι κυριότερες διαμορφώσεις που φαίνεται να παραμένουν σταθερές για ορισμένο χρονικό διάστημα. Το μέγεθος του πίνακα αντιστοιχεί σε 3,758,525 διαμορφώσεις. Είναι εμφανές πως η δομή του πεπτιδίου σταθεροποιείται σχετικά γρήγορα και δεν αλλάζει σημαντικά η διαμόρφωσή του μέχρι το τέλος της προσομοίωσης. Η έκτη δομή είναι η κυρίαρχη δομή του τροχιακού ILDN. Με βέλος σημειώνεται η περιοχή του πίνακα στην οποία ανήκει κάθε δομή. Ο χρωματισμός των δομών

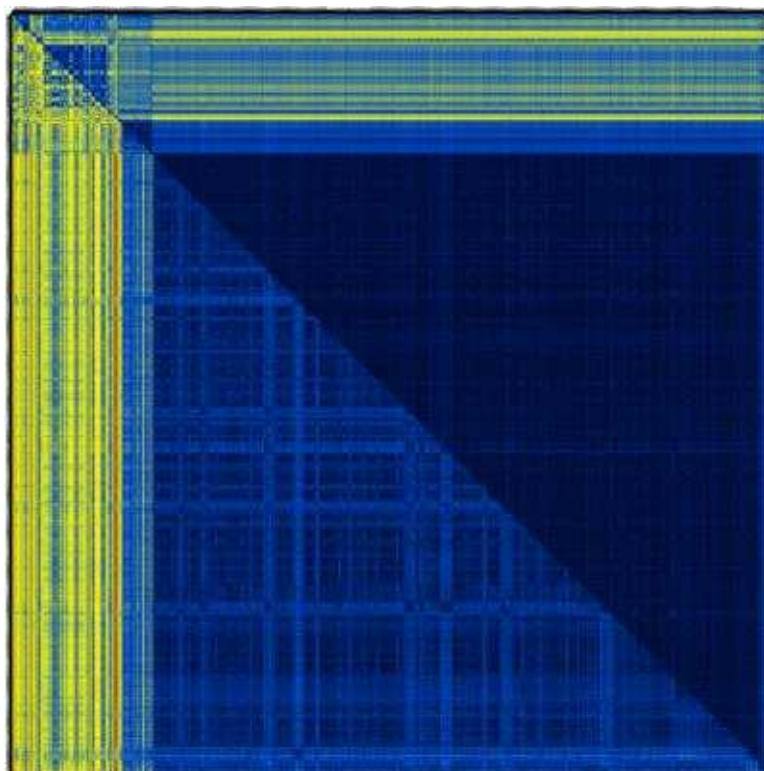
είναι ίδιος με την εικόνα 4.10. Τα αποτελέσματα του πίνακα φαίνεται να είναι σε συμφωνία τα αποτελέσματα της εικόνας 4.2.

Συμπερασματικά μπορούμε να πούμε πως στην προσομοίωση ILDN το πεπτίδιο αναδιπλώθηκε πιο γρήγορα συγκριτικά με την προσομοίωση 99SB, αν αναλογιστούμε πως στο τροχιακό 99SB μέχρι τα 1.78 μs δεν πραγματοποιήθηκε πλήρης αναδίπλωση και η αντιπροσωπευτική δομή υιοθετήθηκε στα 3.8 μs, ενώ στο τροχιακό ILDN δεν υπήρξε υιοθέτηση της φυσικής δομής μέχρι τα 0.47 μs και η αντιπροσωπευτική δομή υιοθετήθηκε στα 1.26 μs. Η αντιπροσωπευτική δομή του τροχιακού ILDN παρουσιάζει πιο έντονη συστροφή των β-κλώνων σε σχέση με την αντιπροσωπευτική δομή του τροχιακού 99SB όπως φαίνεται και στην εικόνα 4.12. Επιπροσθέτως, στον πίνακα RMSD των ατόμων Ca των καταλοίπων 3-15 παρουσιάζεται αυξημένη σταθερότητα συγκριτικά με το RMSD όλων των ατόμων Ca (εικόνα 4.13). Το φαινόμενο αυτό ίσως δικαιολογεί την απουσία του πρώτου δεσμού.



Εικόνα 4.12

Οι κυρίαρχες δομές των τροχιακών 99SB (δεξιά) και ILDN (αριστερά). Και οι δύο δομές έχουν τη διαμόρφωση β-φουρκέτας. Ο τύπος της στροφής είναι type I b-turn και οι δεσμοί υδρογόνου είναι οι ίδιοι με την NMR δομή, με εξαίρεση τον πρώτο δεσμό. Παρατηρήθηκε μια μικρή συστροφή και στις δύο διαμορφώσεις, αλλά στη δομή του τροχιακού ILDN είναι εντονότερη.



Εικόνα 4.13

Πίνακας RMSD του τροχιακού ILDN. Σύγκριση των τιμών RMSD για τα άτομα Ca των καταλοίπων 3-15 (επάνω από τη διαγώνιο) και για τα άτομα Ca όλων των καταλοίπων (κάτω από τη διαγώνιο). Οι τιμές RMSD είναι χαμηλότερες για τα κατάλοιπα 3-15, δηλαδή το κομμάτι αυτό του πεπτιδίου είναι πιο σταθερό κατά τη διάρκεια της προσομοίωσης.

Η ανάλυση των τροχιακών βασιζόμενη στο RMSD θα κλείσει με την παρουσίαση των αντίστοιχων αποτελεσμάτων για το τροχιακό της προσομοίωσης protonated. Στην προσομοίωση protonated παρήχθησαν 3714800 διαμορφώσεις και η συνολική διάρκεια της προσομοίωσης ήταν 2.97 μ s. Το δυναμικό πεδίο που χρησιμοποιήθηκε ήταν το AMBER99SB-ILDN, αλλά η διαφορά με την ILDN προσομοίωση έγκειται στην πρωτονίωση του πεπτιδίου.

Στην εικόνα 4.14 παρουσιάζεται ο πίνακας RMSD για τα άτομα Ca του πεπτιδίου κατά τη διάρκεια της προσομοίωσης και οι αντίστοιχες δομές με τις πιο σταθερές περιοχές του τροχιακού. Επιλέχθηκαν συνολικά έξι περιοχές και

αντίστοιχα έξι δομές, στις οποίες αλλάζουν εντελώς τα ζεύγη δεσμών υδρογόνου συγκριτικά με τις δομές των δύο προηγούμενων τροχιακών. Η πρώτη δομή έχει τιμή RMSD 1.2 Å από τη μέση δομή της περιοχής στην οποία ανήκει και υιοθετήθηκε στα 0.10 μs. Έχουν σχηματιστεί τέσσερις δεσμοί υδρογόνου στον πεπτιδικό σκελετό και μία μικρή περιοχή φέρει β-διαμόρφωση. Οι δεσμοί υδρογόνου είναι οι εξής: (a) Phe4(O) – Lys6(N), (b) Lys6(O) – Thr14(N), (c) Leu8(N) – Thr12(O) και (d) Leu8(O) – Lys11(N). Κανένας από τους δεσμούς της δομής αυτής δεν αντιστοιχεί σε δεσμό που υφίσταται στην NMR δομή και στις δομές των τροχιακών 99SB και ILDN. Επιπλέον στην πρώτη δομή υπάρχουν δύο στροφές. Η δομή αυτή είναι αρκετά διαφορετική από τις επόμενες του τροχιακού, καθώς εμφανίζονται πολύ υψηλές τιμές RMSD ανάμεσα σε αυτή την περιοχή και τον υπόλοιπο πίνακα, πρόκειται δηλαδή για μία παροδική διαμόρφωση η οποία δεν εμφανίζεται ξανά. Η δεύτερη δομή έχει τιμή RMSD 0.54 Å από τη μέση δομή της περιοχής από την οποία προέρχεται και υιοθετήθηκε στα 0.36 μs. Στη δομή αυτή έχουν σχηματιστεί οι παρακάτω επτά δεσμοί υδρογόνου στον πεπτιδικό σκελετό: (1) Met1(O) – Ile3(N), (2) Phe4(N) – Glu16(O), (3) Phe4(O) – Glu16(N), (4) Lys6(N) – Thr14(O), (5) Lys6(O) – Thr14(N), (6) Leu8(N) – Thr12(O), (7) Leu8(O) – Lys11(N). Κανένας από τους δεσμούς αυτούς δεν δημιουργείται στην NMR δομή και στις προηγούμενες δομές των τροχιακών 99SB και ILDN. Στο πεπτίδιο διαμορφώνεται εκτεταμένη β-δομή στα κατάλοιπα 4-8 και 12-16, ενώ φαίνεται να σχηματίζεται στροφή στα κατάλοιπα 9-11 και 1-3. Το RMSD της δομής αυτής είναι υψηλό σε σχέση με τις επόμενες αντιπροσωπευτικές δομές. Η τρίτη δομή έχει τιμή RMSD 0.47 Å από τη μέση δομή της περιοχής από την οποία προέρχεται και υιοθετήθηκε στα 1.39μs. Φέρει επτά δεσμούς υδρογόνου στον πεπτιδικό σκελετό : (1) Met1(N) – Val17(O), (2) Phe4(N) – Glu16(O), (3) Phe4(O) – Glu16(N), (4) Lys6(N) – Thr14(O), (5) Lys6(O) – Thr14(N), (6) Leu8(N) – Thr12(O), (7) Leu8(O) – Thr12(N). Με εξαίρεση τον πρώτο δεσμό όλοι οι υπόλοιποι είναι ίδιοι με τη δεύτερη δομή. Η εκτεταμένη β-διαμόρφωση συμπεριλαμβάνει, επίσης τα ίδια κατάλοιπα με τη δεύτερη δομή και η στροφή φαίνεται να σχηματίζεται στα

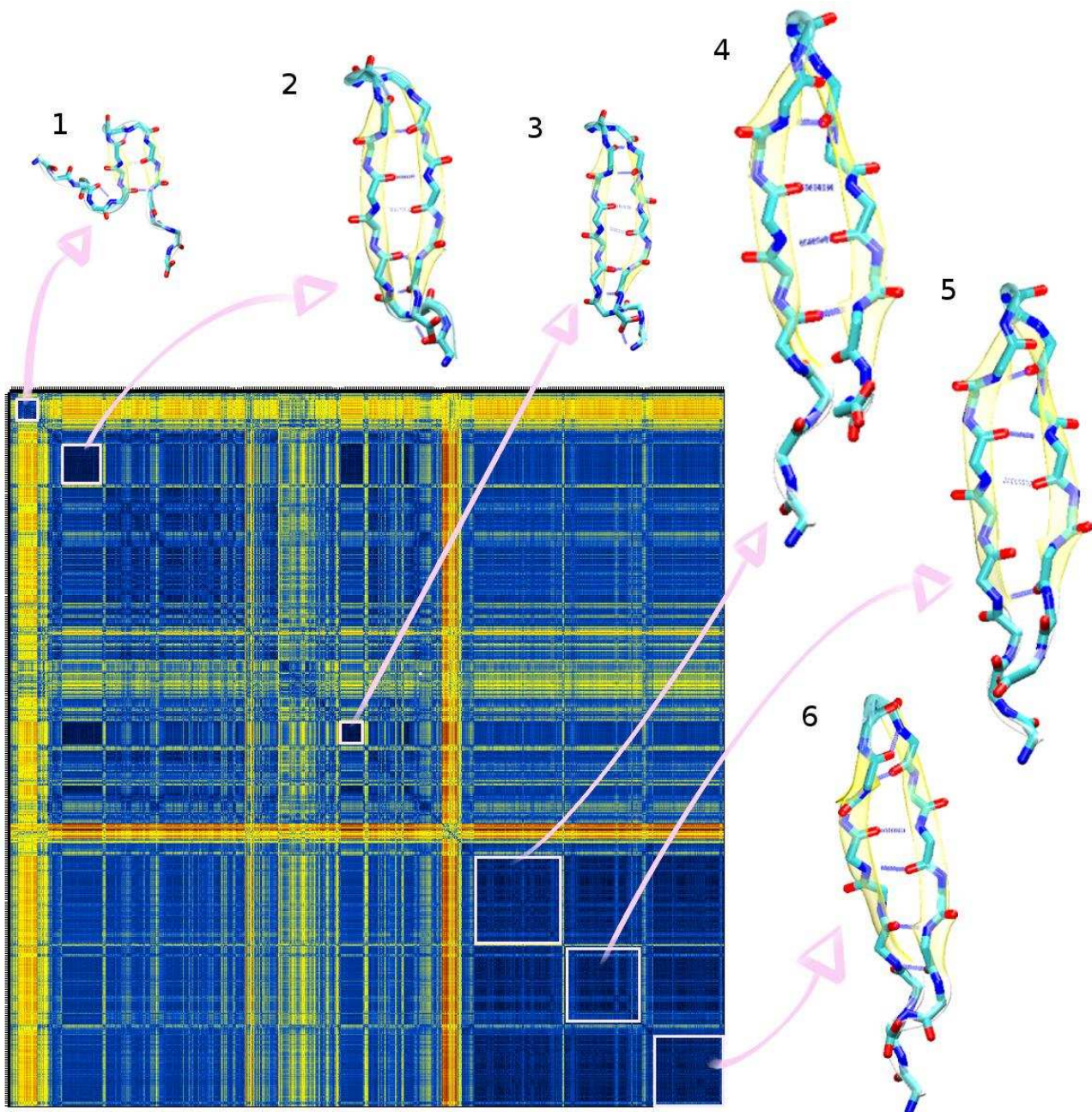
κατάλοιπα 9-11. Στην τέταρτη δομή η τιμή RMSD από τη μέση δομή της περιοχής είναι 0.63 Å και υιοθετείται στα 1.97 μs. Στη δομή αυτή σχηματίστηκαν τέσσερις δεσμοί υδρογόνου μεταξύ των καταλοίπων : (1) Phe4(O) – Leu15(N), (2) Lys6(N) – Ile13(O), (3) Lys6(O) – Ile13(N) και (4) Leu8(N) – Lys11(O). Οι δεσμοί αυτοί παρόλο που δεν αντιστοιχούν στην πειραματικά προσδιορισμένη NMR δομή, συμφωνούν με την πιθανή εναλλακτική δομή που παρουσίασε η ομάδα της R. Zerella. Στην εναλλακτική δομή έχει δημιουργηθεί μετατόπιση των κλώνων με αποτέλεσμα τη δημιουργία διαφορετικών δεσμών στο σκελετό του πεπτιδίου. Στην τέταρτη δομή τα άκρα της φουρκέτας είναι αρκετά απομακρυσμένα και δε δημιουργούνται δύο δεσμοί υδρογόνου, όπως επίσης δεν έχει δημιουργηθεί και ο δεσμός μεταξύ των καταλοίπων Leu8(O) – Lys11(N), που προβλέπει η εναλλακτική δομή. Η περιοχή της β-διαμόρφωσης και της στροφής είναι ίδια με την προηγούμενη δομή. Στην πέμπτη δομή η τιμή RMSD από τη μέση δομή της περιοχής από την οποία προέρχεται είναι 0.71 Å και υιοθετήθηκε στα 2.46 μs. Στη δομή αυτή σχηματίστηκαν οι παρακάτω πέντε δεσμοί υδρογόνου στον πεπτιδικό σκελετό: (1) Phe4(N) – Leu15(O), (2) Phe4(O) – Leu15(N), (3) Lys6(N) – Ile13(O), (4) Lys6(O) – Ile13(N), (5) Leu8(N) – Lys11(O). Οι δεσμοί υδρογόνου αντιστοιχούν στους δεσμούς της εναλλακτικής δομής, αλλά δεν έχουν δημιουργηθεί οι δεσμοί Leu8(O) – Lys11(N), Gln2(N) – Val17(O) και Gln2(O) – Val17(N). Οι δύο τελευταίοι βρίσκονται στα άκρα της φουρκέτας τα οποία έχουν μεγάλη απόσταση μεταξύ τους καθώς παρουσιάζουν αυξημένη κινητικότητα σε σχέση με το υπόλοιπο πεπτιδικό τμήμα. Τα κατάλοιπα που συμμετέχουν στη στροφή είναι ίδια με τις προηγούμενες δομές, όπως και τα κατάλοιπα της β-διαμόρφωσης. Η έκτη και τελευταία δομή του τροχιακού protonated έχει τιμή RMSD 0.47 Å από τη μέση δομή της περιοχής από την οποία προέρχεται και υιοθετήθηκε στα 2.69 μs. Στην έκτη δομή έχουν σχηματιστεί έξι δεσμοί υδρογόνου στον πεπτιδικό σκελετό, οι οποίοι είναι οι εξής: (1) Phe4(N) – Leu15(O), (2) Phe4(O) – Leu15(N), (3) Lys6(N) – Ile13(O), (4) Lys6(O) – Ile13(N), (5) Leu8(N) – Lys11(O) και (6) Leu8(O) – Lys11(N). Οι δεσμοί συμφωνούν με το πρότυπο δεσμών της

εναλλακτικής δομής, αλλά απουσιάζει το ζεύγος δεσμών των καταλοίπων Gln2 – Val17, τα οποία έχουν διαμόρφωση coil στη συγκεκριμένη δομή. Τα κατάλοιπα της στροφής και της β-διαμόρφωσης παραμένουν ίδια με τις προηγούμενες δομές. Οι δεσμοί υδρογόνου που δημιουργήθηκαν και αντιστοιχούν στην εναλλακτική δομή φαίνονται συνολικά στον πίνακα 4.2

	1	2	3	4	5	6
Gln2(N) – Val17(O)	-	-	-	-	-	-
Gln2(O) – Val17(N)	-	-	-	-	-	-
Phe4(N) – Leu15(O)	-	-	-	-	+	+
Phe4(O) – Leu15(N)	-	-	-	+	+	+
Lys6(N) – Ile13(O)	-	-	-	+	+	+
Lys6(O) – Ile13(N)	-	-	-	+	+	+
Leu8(N) – Lys11(O)	-	-	-	+	+	+
Leu8(O) – Lys11(N)	+	+	-	-	-	+

Πίνακας 4.2

Σημειώνονται ποιοι δεσμοί υδρογόνου της εναλλακτικής δομής σχηματίζονται στις δομές που παρουσιάζουν κατοχή στο τροχιακό protonated προσομοίωσης μεγαλύτερη του 5% του συνόλου. Η δομή έξι φαίνεται να είναι η αντιπροσωπευτική δομή του τροχιακού και φέρει τους περισσότερους δεσμούς υδρογόνου της εναλλακτικής δομής.

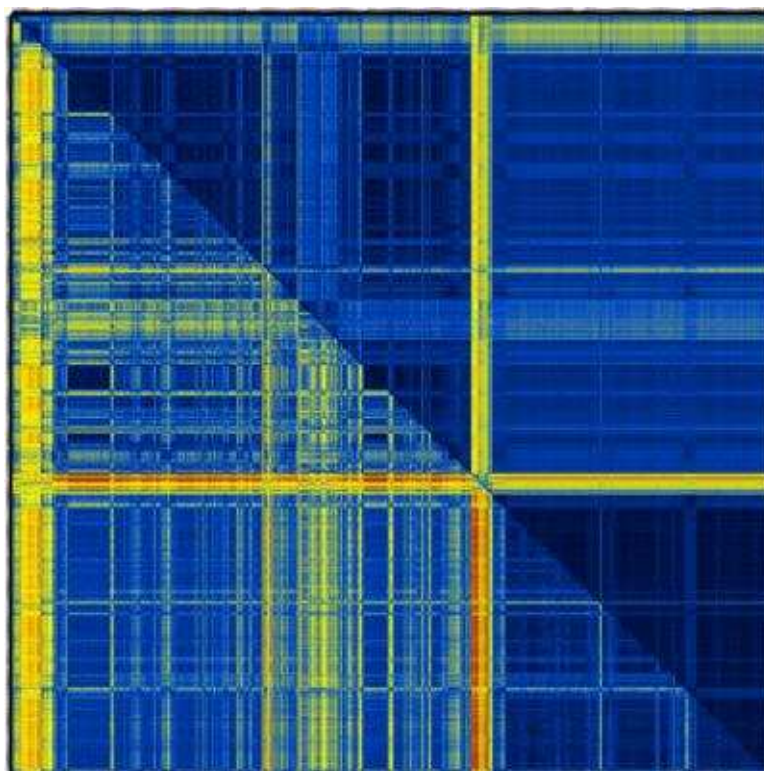


Εικόνα 4.14

Πίνακας RMSD του τροχιακού protonated. Με αριθμητική σειρά παρουσιάζονται οι κυριότερες διαμορφώσεις που παρουσιάζουν κατοχή (occipancy) στο τροχιακό μεγαλύτερη του 5% του συνόλου.. Το μέγεθος του πίνακα αντιστοιχεί σε 3,714,800 διαμορφώσεις. Ο σχηματισμός της φουρκέτας εμφανίζεται μόλις από την δεύτερη δομή (0.36 μ s) και με ορισμένες μεταβολές στους δεσμούς υδρογόνου επανεμφανίζεται στην τρίτη δομή. Οι δομές 4, 5, και 6 μοιάζουν αρκετά μεταξύ τους, συνεπώς μπορούμε να πούμε ότι μετά την τέταρτη δομή το πεπτιδίο

σταθεροποιείται (0.97 μ s). Η έκτη δομή φαίνεται να είναι η αντιπροσωπευτική δομή του τροχιακού protonated. Με βέλος σημειώνεται η περιοχή του πίνακα στην οποία ανήκει κάθε δομή. Ο χρωματισμός των δομών είναι ίδιος με την εικόνα 4.10. Ο πίνακας φαίνεται να συμφωνεί με τη γραφική παράσταση της εικόνας 4.3.

Τα κατάλοιπα των άκρων όπως και στις δύο προηγούμενες προσομοιώσεις δε φαίνεται να σχηματίζουν δεσμούς υδρογόνου, καθώς βρίσκονται σε μεγάλη απόσταση μεταξύ τους. Για να ελεγχθεί η σταθερότητα του “πυρήνα” του πεπτιδίου παρουσιάζεται στην εικόνα 4.15 ο πίνακας RMSD για τα άτομα Ca των καταλοίπων 3-15 και τα άτομα Ca όλων των καταλοίπων. Όπως και στους δύο προηγούμενους συγκριτικούς πίνακες και εδώ φαίνεται πως η απόσταση των ατόμων Ca για τα κατάλοιπα 3-15 είναι μικρότερη σε σχέση με την απόσταση των ατόμων Ca για όλα τα κατάλοιπα. Συνεπώς το τμήμα 3-15 του πεπτιδίου είναι σταθερότερο κατά τη διάρκεια της προσομοίωσης protonated. Η απώλεια συνεπώς των δεσμών υδρογόνου των άκρων προσδίδεται στην αυξημένη κινητικότητα των άκρων.

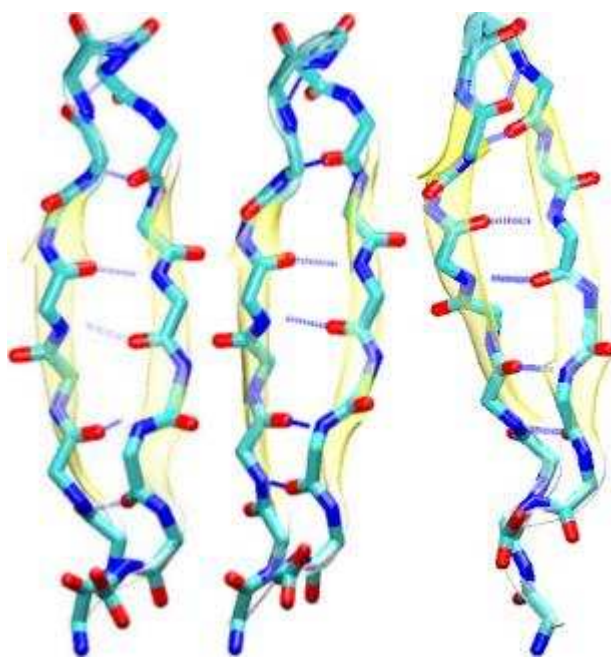


Εικόνα 4.15

Πίνακας RMSD του τροχιακού protonated. Σύγκριση των τιμών RMSD για τα άτομα Ca των καταλοίπων 3-15 (επάνω) και για τα άτομα Ca όλων των καταλοίπων (κάτω από τη διαγώνιο). Οι τιμές RMSD είναι χαμηλότερες για τα κατάλοιπα 3-15, με συνέπεια το κομμάτι αυτό του πεπτιδίου να είναι πιο σταθερό κατά τη διάρκεια της προσομοίωσης.

Συνοψίζοντας όλα τα παραπάνω συμπεραίνουμε πως στην προσομοίωση protonated το πεπτίδιο σχημάτισε την εναλλακτική δομή και όχι την NMR δομή, όπως οι προηγούμενες δύο προσομοιώσεις. Ο χρόνος αναδίπλωσης του πεπτιδίου είναι μικρός συγκριτικά με την προσομοίωση 99SB, καθώς φαίνεται να σχηματίζεται η στροφή και οι β-κλώνοι μόλις στα 0.36μs, ενώ το πεπτίδιο αρχίζει να σταθεροποιείται γύρω στα 0.97 μs και η τελική δομή αποκτάται 2.69 μs. Στην προσομοίωση 99SB η δημιουργία της στροφής και των β-κλώνων έγινε ταυτόχρονα με την απόκτηση της σταθερής δομής, περίπου στα 3.8 μs. Η διαφορά στο χρόνο αναδίπλωσης στις δύο προσομοιώσεις φαίνεται να είναι της τάξης του 1 μs – σημαντικός υπολογιστικός και φυσικός χρόνος σε

προσομοιώσεις μοριακής δυναμικής. Στην περίπτωση της προσομοίωσης ILDN το πεπτίδιο αναδιπλώθηκε αρκετά γρήγορα αλλά και σε αυτή την προσομοίωση ο σχηματισμός της στροφής και των β-κλώνων πραγματοποιήθηκε ταυτόχρονα με την υιοθέτηση της κυρίαρχης δομής (1.26 μ s) και όχι πιο νωρίς, όπως στο τροχιακό της protonated. Στην εικόνα 4.16 παρουσιάζονται οι αντιπροσωπευτικές δομές των τριών τροχιακών.



Εικόνα 4.16

Οι τρεις αντιπροσωπευτικές δομές των προσομοιώσεων που πραγματοποιήθηκαν. Από αριστερά είναι (1) ILDN, (2) 99SB, (3) protonated. Στη δομή protonated είναι εμφανής η μετάθεση των κλώνων κατά ένα κατάλοιπο, καθώς ο β-κλώνος του N-τελικού άκρου (αριστερά) είναι μεγαλύτερος κατά ένα κατάλοιπο σε σχέση με τον β-κλώνο του C-τελικού άκρου. Οι δύο πρώτες δομές συμφωνούν μεταξύ τους και με την NMR δομή, ενώ η τρίτη δομή έχει διαφορετική στροφή και διαφορετικούς δεσμούς υδρογόνου από τις άλλες δύο και από την NMR δομή, αλλά συμφωνεί με την εναλλακτική δομή που προτείνεται από την ομάδα της R. Zerella στο NMR πείραμα.

4.2 Ανάθεση δευτεροταγούς δομής

Στην προηγούμενη ενότητα είδαμε την κινητικότητα του πεπτιδίου κατά τη διάρκεια των τριών προσομοιώσεων και ορισμένες δομές από στιγμιότυπα των τροχιακών. Οι αναλύσεις που θα ακολουθήσουν σε αυτή την ενότητα αφορούν την δευτεροταγή δομή που υιοθετούν τα κατάλοιπα του πεπτιδίου κατά τη διάρκεια της προσομοίωσης. Τα προγράμματα που χρησιμοποιήθηκαν για το σκοπό αυτό είναι το STRIDE (D. Frishman et al., 1995, M. Heinig et al., 2004) και το DSSP (W. Kabsch et al., 1983).

Η αντιστοίχιση στοιχείων δευτεροταγούς δομής είναι ένα σημαντικό βήμα στο χαρακτηρισμό της τρισδιάστατης δομής μιας πρωτεΐνης και επίσης είναι η αφετηρία για πολλές θεωρητικές μελέτες που αφορούν την πρόβλεψη δευτεροταγούς δομής, την μοντελοποίηση με βάση την ομολογία, την ανάστροφη πρωτεϊνική αναδίπλωση, την περιγραφή των αναδιπλούμενων μοτίβων και άλλες παρόμοιες μελέτες. Παρόλο που η αναγνώριση διαμορφώσεων α-ελίκων και β-φύλλων φαίνεται να είναι εύκολη, στην πραγματικότητα μια αλγοριθμική λύση περιπλέκεται από την ασαφή φύση των στοιχείων αυτών. Έχουν αναπτυχθεί αρκετοί αλγόριθμοι για την επίλυση της ανάθεσης δευτεροταγούς δομής και δεν προκαλεί έκπληξη το γεγονός ότι κάθε μέθοδος προσεγγίζει το πρόβλημα από διαφορετική σκοπιά με αποτέλεσμα να υπάρχει διαφωνία στα αποτελέσματα της ανάθεσης.

Το STRIDE (STRuctural IDentification) είναι ένας αυτόματος αλγόριθμος για την ανάθεση δευτεροταγούς δομής από τις ατομικές συντεταγμένες και βασίζεται στη συνδυασμένη χρήση της ενέργειας δεσμών υδρογόνου και σε στατιστικές πληροφορίες που αφορούν τη γωνία στρέψης των ατόμων του σκελετού (D. Frishman et al., 1995) . Πριν τη δημιουργία του STRIDE, ο πιο κοινός αλγόριθμος ήταν αυτός του προγράμματος DSSP (Define Secondary

Structure of Proteins), ο οποίος βασίζεται στην αναγνώριση μοτίβων δεσμών υδρογόνου. Τα βασικά μοτίβα που παρατηρούνται στον αλγόριθμο DSSP είναι τα εξής : “n-turns” : με ένα δεσμό υδρογόνου μεταξύ της ομάδας CO του καταλοίπου i και της ομάδας NH του καταλοίπου i+n, (n=2, 3, 4) και “bridges” : με δεσμούς υδρογόνου μεταξύ μη-σειριακών καταλοίπων της αλληλουχίας. Η επανάληψη μοτίβων “4-turns” αντιστοιχεί σε α-έλικες, ενώ η επανάληψη του μοτίβου “bridges” αντιστοιχεί σε β-δομή. Όλες οι άλλες μορφές των βασικών μοτίβων αντιστοιχούν σε 3_{10} -έλικες, π-έλικες, μονές στροφές και μονές β-bridges (W. Kabsch et al., 1983) .

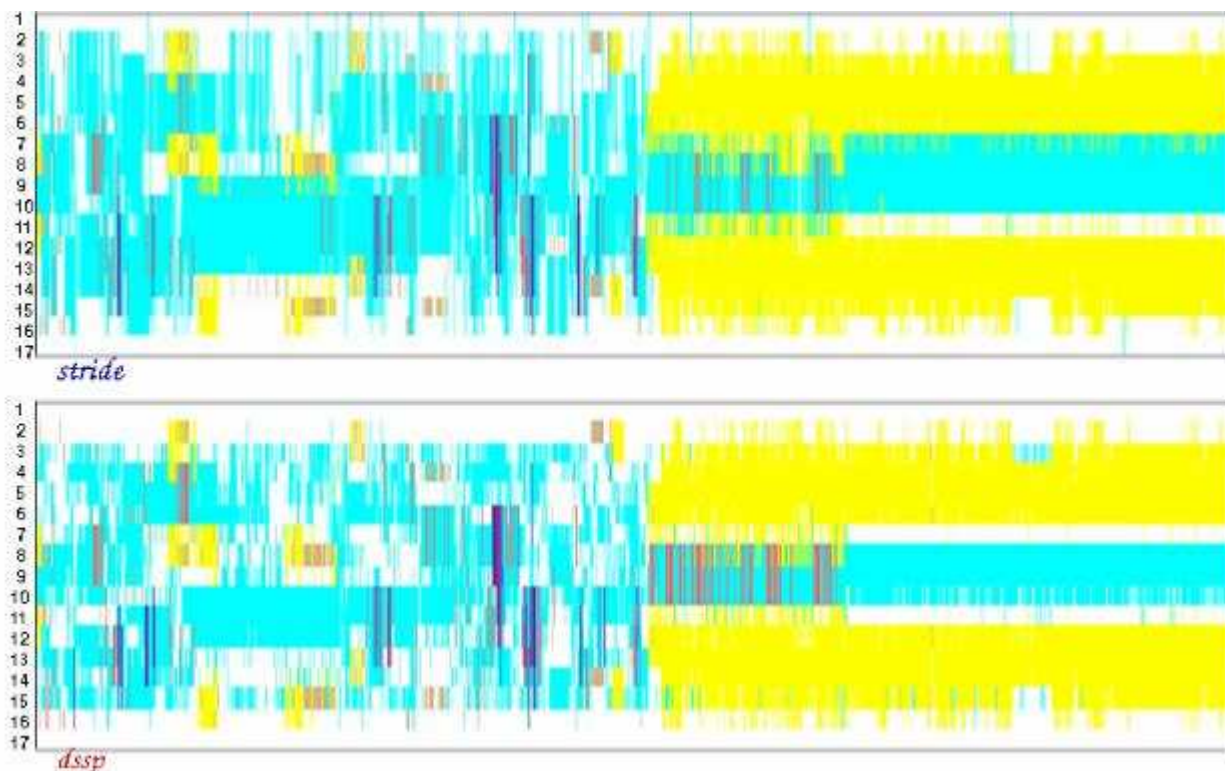
Το αποτέλεσμα από τον αλγόριθμο STRIDE είναι ένα γράμμα που αντιστοιχεί στη διαμόρφωση κάθε καταλοίπου της αλληλουχίας καθ' όλη την έκταση του τροχιακού. Η αντιστοίχιση γράμματος και δευτεροταγούς δομής είναι η παρακάτω :

- H α-έλικα
- G 3_{10} έλικα
- I π-έλικα
- E Εκτεταμένη β-διαμόρφωση
- B β-μεμονωμένη γέφυρα
- T Στροφή
- C Σπείραμα (καμία από τις παραπάνω διαμορφώσεις)

Η αντίστοιχη έξοδος του αλγόριθμου DSSP είναι:

- H α-έλικα
- G 3_{10} έλικα
- I π-έλικα
- E Εκτεταμένη β-διαμόρφωση
- B β-μεμονωμένη γέφυρα
- T Στροφή με δεσμό υδρογόνου
- S Στροφή

Στην εικόνα 4.17 εμφανίζονται τα αποτελέσματα για το τροχιακό 99SB από την ανάθεση δομής σύμφωνα με τους αλγόριθμους DSSP και STRIDE. Προκειμένου να γίνει πιο ευανάγνωστο το αποτέλεσμα, τα γράμματα αντικαταστάθηκαν με χρώμα.



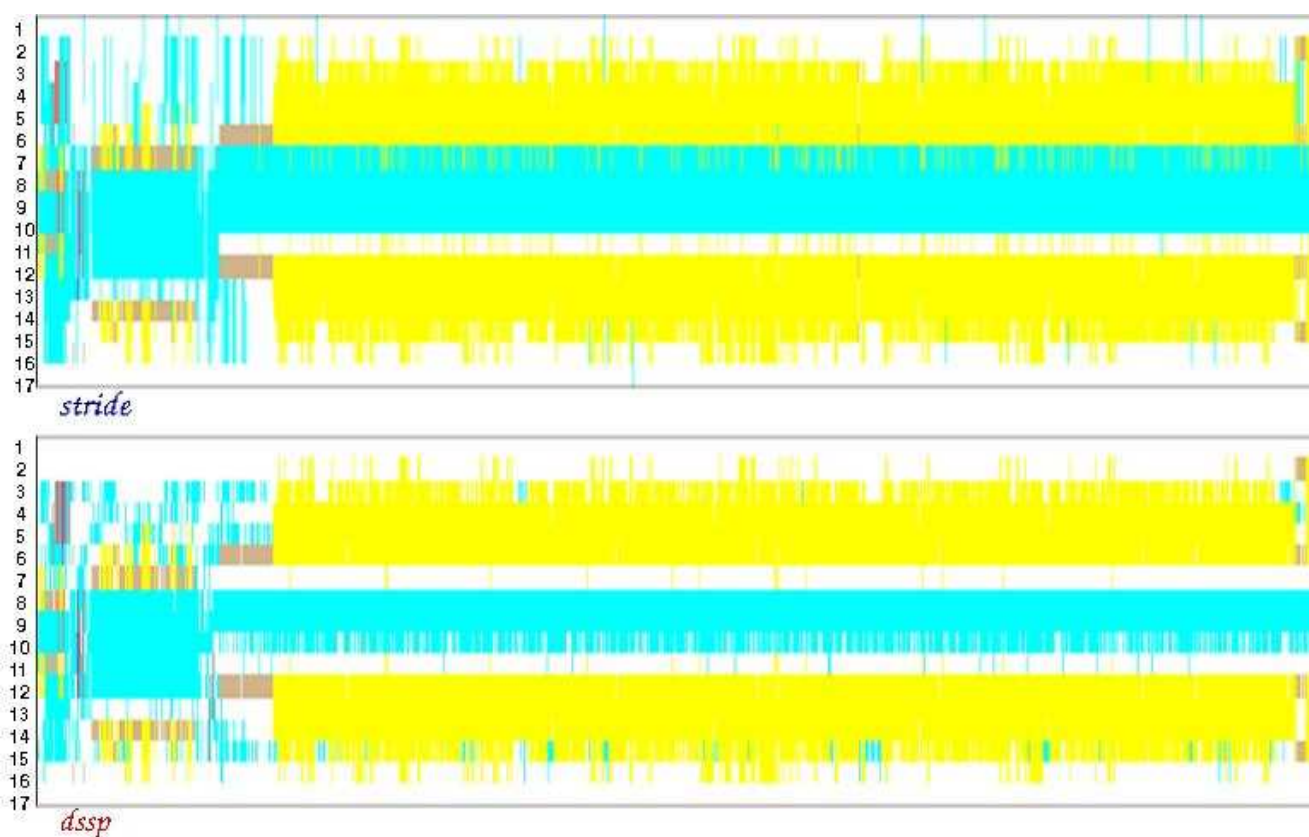
Εικόνα 4.17

Γραφική παράσταση ανάθεσης δομής από τους αλγόριθμους STRIDE (επάνω) και DSSP (κάτω) του τροχιακού 99SB. Στον οριζόντιο άξονα είναι τοποθετημένες οι διαμορφώσεις που υιοθετήθηκαν κατά τη διάρκεια της προσομοίωσης 99SB με βήμα 1000, ενώ στον κάθετο άξονα αριθμούνται τα κατάλοιπα του πεπτιδίου. Βλέπουμε την εξέλιξη κάθε καταλοίπου κατά μήκος του τροχιακού. Το κίτρινο χρώμα αντιστοιχεί σε εκτεταμένη β-διαμόρφωση, το κυανό σε στροφή, το λευκό σε σπείραμα, το κόκκινο σε α-έλικα, το μοβ σε 3_{10} -έλικα, το μπλε σε π-έλικα και το γκρι σε μεμονωμένη γέφυρα.

Όπως ήταν αναμενόμενο τα αποτελέσματα των αλγορίθμων συμφωνούν στην πλειοψηφία των διαμορφώσεων καθώς λαμβάνονται υπόψιν και στις δύο περιπτώσεις παρόμοιες συνισταμένες για την απόδοση δευτεροταγούς δομής. Στην ανάθεση που προκύπτει από τον αλγόριθμο STRIDE τα δύο πρώτα και τα δύο τελευταία κατάλοιπα δεν φαίνεται να υιοθετούν β-διαμόρφωση, πράγμα που συμφωνεί και με το αποτέλεσμα των πινάκων RMSD περί αυξημένης κινητικότητας των άκρων της φουρκέτας. Επιπλέον μετά τη μέση περίπου του τροχιακού φαίνεται μια πιο ξεκάθαρη διαμόρφωση του πεπτιδίου, με τα κατάλοιπα 3 – 6 και 12 – 15 να υιοθετούν εκτεταμένη β-διαμόρφωση και τα κατάλοιπα 7 – 10 να βρίσκονται στη στροφή της φουρκέτας σχηματίζοντας τον τύπο I β-στροφής. Στην ανάθεση που προκύπτει από τον αλγόριθμο DSSP τα κατάλοιπα 1 – 2 και 16 – 17, επίσης, δεν φαίνεται να υιοθετούν β-διαμόρφωση. Μετά τη μέση περίπου του τροχιακού φαίνεται τα κατάλοιπα της εκτεταμένης β-διαμόρφωσης να είναι τα ίδια με την ανάθεση του αλγορίθμου STRIDE, ενώ ως κατάλοιπα της στροφής αποδίδονται τα 8 – 10, γεγονός που δεν συμφωνεί με την NMR δομή, ενώ δεν φαίνεται να αποδίδεται καμία διαμόρφωση στο έβδομο κατάλοιπο. Γενικότερα ο αλγόριθμος DSSP αποδίδει σε περισσότερα κατάλοιπα διαμόρφωση σπειράματος σε σχέση με τον STRIDE. Και οι δύο αλγόριθμοι αποδίδουν σε ορισμένα κατάλοιπα διαμόρφωση έλικας, αλλά φαίνεται να είναι περιστασιακή.

Στο τροχιακό της ILDN προσομοίωσης τα αποτελέσματα από την απόδοση δευτεροταγούς δομής στα κατάλοιπα του πεπτιδίου (εικόνα 4.18) είναι παρόμοια με το τροχιακό της 99SB προσομοίωσης. Στα κατάλοιπα 1 – 2 και 16 – 17 δεν φαίνεται να αποδίδεται καμία διαμόρφωση καθ' όλη την έκταση του τροχιακού, ενώ στα κατάλοιπα 1 – 6 και 12 – 15 αποδίδεται εκτεταμένη β-διαμόρφωση, στη μεγαλύτερη έκταση του τροχιακού και με τους δύο αλγορίθμους. Ο αλγόριθμος STRIDE αποδίδει, στην μεγαλύτερη έκταση του τροχιακού, διαμόρφωση στροφής στα κατάλοιπα 7 – 10. Τα αποτελέσματα αυτά είναι σύμφωνα με την

NMR δομή. Ενώ ο αλγόριθμος DSSP αποδίδει διαμόρφωση στροφής στα κατάλοιπα 8 – 10, για την ίδια έκταση του τροχιακού. Στο τροχιακό της ILDN προσομοίωσης δεν αποδίδεται διαμόρφωση α-έλικας, όπως στο προηγούμενο τροχιακό. Όσο αφορά τη σύγκριση των δύο αλγορίθμων σε σχέση με την απόδοση διαφορετικής δευτεροταγούς διαμόρφωσης στα ίδια κατάλοιπα είναι πιθανόν να οφείλεται σε ψευδώς θετικά ή αληθώς αρνητικά αποτελέσματα. Επιπλέον ο αλγόριθμος STRIDE είναι σε θέση να αντιστοιχεί κατάλοιπα που συμμετέχουν σε GI β-bulge στην εκτεταμένη διαμόρφωση, ενώ το DSSP όχι (D. Frishman et al., 1995).

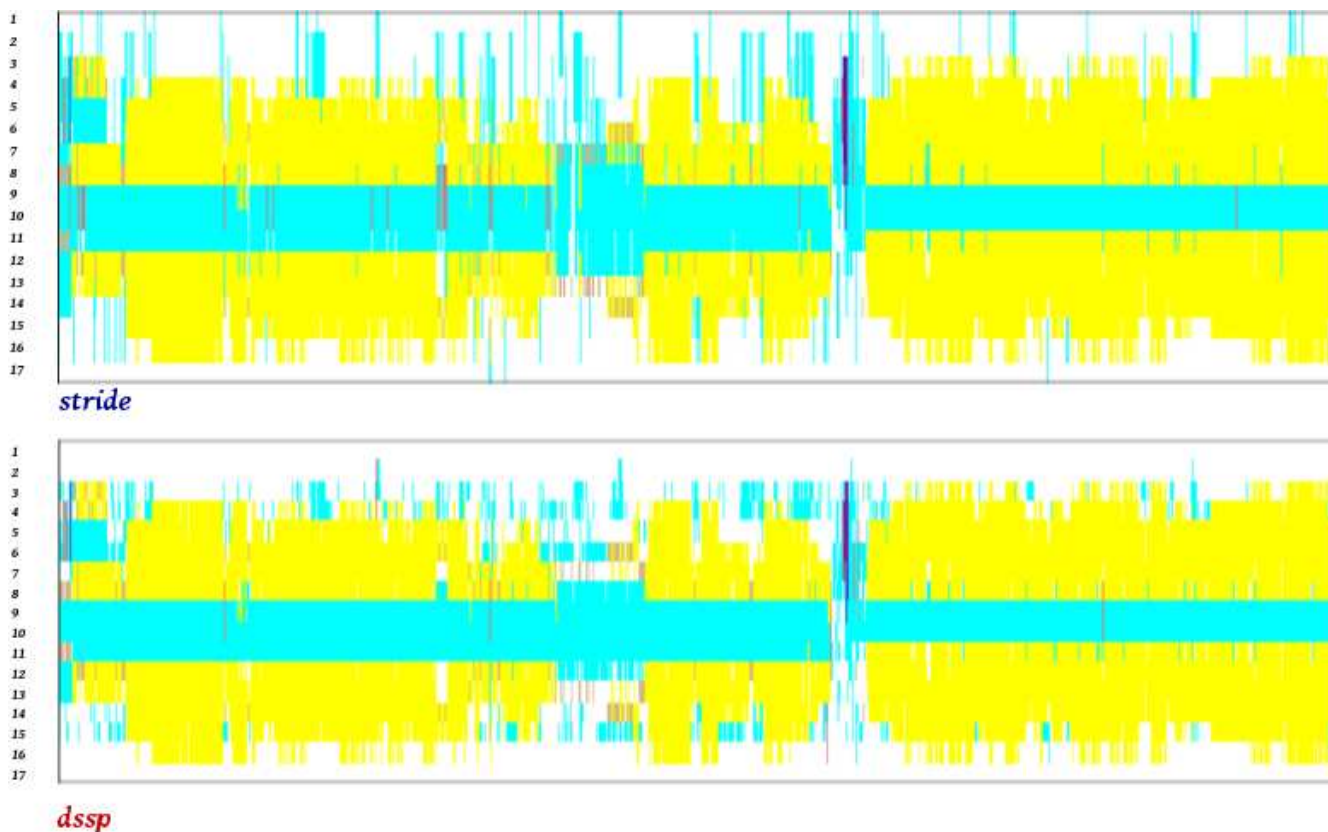


Εικόνα 4.18

Γραφική παράσταση ανάθεσης δομής από τους αλγορίθμους STRIDE (επάνω) και DSSP (κάτω) του τροχιακού ILDN. Στον οριζόντιο άξονα είναι τοποθετημένες οι διαμορφώσεις που υιοθετήθηκαν κατά τη διάρκεια της προσομοίωσης ILDN με βήμα 1000, ενώ στον κάθετο άξονα αριθμούνται τα κατάλοιπα του πεπτιδίου.

Βλέπουμε την εξέλιξη της δευτεροταγούς δομής κάθε καταλοίπου κατά μήκος του τροχιακού. Ο χρωματισμός της δευτεροταγούς διαμόρφωσης είναι ίδιος με την εικόνα 4.17.

Στο τροχιακό της προσομοίωσης protonated η ανάθεση δομής από τους αλγορίθμους STRIDE και DSSP φαίνεται στην εικόνα 4.19. Τα αποτελέσματα από τους δύο αλγορίθμους στο τροχιακό protonated μοιάζουν περισσότερο μεταξύ τους σε σχέση με τις δύο προηγούμενες προσομοιώσεις. Και σε αυτή την περίπτωση τα κατάλοιπα των άκρων της φουρκέτας δε φαίνεται να υιοθετούν σταθερή δευτεροταγή δομή για μεγάλο χρονικό διάστημα. Σε αντίθεση με τα δύο προηγούμενα τροχιακά εδώ φαίνεται να αποδίδεται εξ' αρχής η δομή της στροφής στα κατάλοιπα 9 – 11, αλλά στη συνέχεια περιορίζεται μόνο στα κατάλοιπα 9 – 10. Η εκτεταμένη β-διαμόρφωση ανατίθεται στα κατάλοιπα 4 – 8 και 12 – 15 στη μεγαλύτερη έκταση του τροχιακού. Η διαμόρφωση του ενδέκατου καταλοίπου φαίνεται να μεταβάλλεται από στροφή σε εκτεταμένη β-διαμόρφωση λίγο μετά τη μέση του τροχιακού.

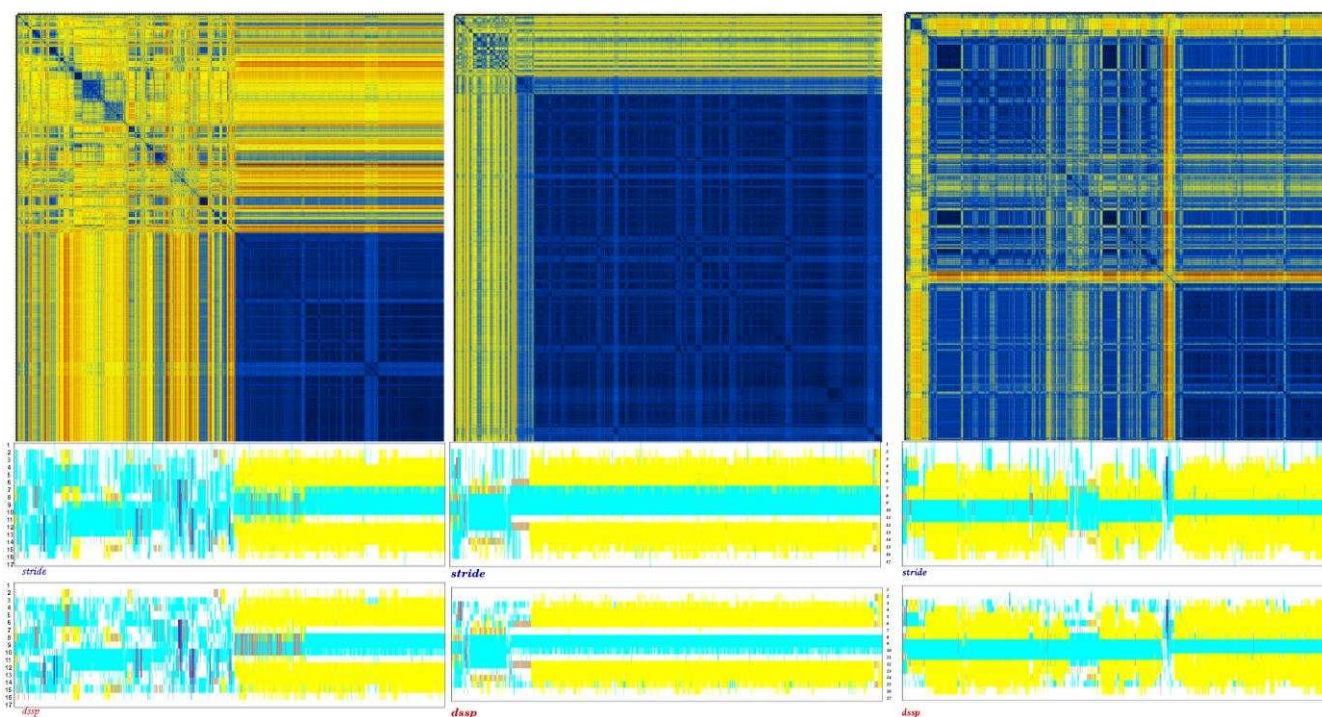


Εικόνα 4.19

Γραφική παράσταση ανάθεσης δομής από τους αλγόριθμους STRIDE (επάνω) και DSSP (κάτω) του τροχιακού protonated. Στον οριζόντιο άξονα είναι τοποθετημένες οι διαμορφώσεις που υιοθετήθηκαν κατά τη διάρκεια της προσομοίωσης protonated με βήμα 1000, ενώ στον κάθετο άξονα αριθμούνται τα κατάλοιπα του πεπτιδίου. Βλέπουμε την εξέλιξη της δευτεροταγούς δομής κάθε καταλοίπου κατά μήκος του τροχιακού. Ο χρωματισμός της δευτεροταγούς διαμόρφωσης είναι ίδιος με την εικόνα 4.17.

Στην εικόνα 4.20 φαίνεται η σύγκριση των πινάκων RMSD και της εξέλιξης στο χρόνο της δευτεροταγούς δομής ανά κατάλοιπο του εκάστοτε τροχιακού. Τα αποτελέσματα από τους πίνακες RMSD και τους αλγόριθμους ανάθεσης δευτεροταγούς δομής φαίνεται να συμφωνούν μεταξύ τους. Οι αλλαγές στο RMSD που παρατηρούνται στους πίνακες αντικατοπτρίζονται και σε αλλαγές της δευτεροταγούς δομής από τους αλγόριθμους. Το πρώτο συμπέρασμα που διατυπώθηκε από την μελέτη των πινάκων RMSD, για την κινητικότητα των

καταλοίπων 1 – 2 και 16 – 17 και την αδυναμία σχηματισμού δεσμών υδρογόνου στα κατάλοιπα αυτά, επιβεβαιώνεται και από την πρόβλεψη δευτεροταγούς δομής από τους αλγόριθμους STRIDE και DSSP, καθώς δεν προβλέπεται η απόκτηση δευτεροταγούς δομής από τα κατάλοιπα αυτά. Εφόσον δεν σχηματίζονται δεσμοί υδρογόνου δεν μπορεί να προβλεφθεί και η δευτεροταγής δομή τους. Επιπλέον και στα υπόλοιπα κατάλοιπα τα αποτελέσματα των τριών μεθόδων είναι σύμφωνα μεταξύ τους, καθώς ο σχηματισμός της στροφής και της εκτεταμένης β-διαμόρφωσης φαίνεται να είναι σταθερά κατά τη διάρκεια των τριών τροχιακών.



Εικόνα 4.20

Πίνακες RMSD και ανάθεση δομής για τα τροχιακά 99SB, ILDN και protonated. Από αριστερά προς τα δεξιά παρουσιάζονται με τη σειρά τα τροχιακά 99SB, ILDN και protonated. Όλοι οι πίνακες δημιουργήθηκαν με βήμα 1000. Τα αποτελέσματα από τις τρεις μεθόδους φαίνεται να συμφωνούν για κάθε προσομοίωση.

Οι αλγόριθμοι πρόβλεψης δευτεροταγούς δομής υποστηρίζουν την ύπαρξη β-φουρκέτας και στα τρία τροχιακά και μπορούμε να παρατηρήσουμε την μετάθεση β-κλώνου που δημιουργήθηκε στο πεπτίδιο στο τροχιακό της protonated προσομοίωσης.

4.3 Cluster analysis

Οι προσομοιώσεις μοριακής δυναμικής έχουν γίνει ισχυρό εργαλείο για την μελέτη της αναδίπλωσης βιολογικών μορίων τα τελευταία χρόνια. Καθώς η διεκπεραίωση τους σε χρονική κλίμακα μερικών μs αποτελεί πλέον διαδικασία ρουτίνας, ένας μεγάλος όγκος πληροφοριών συλλέγεται προκειμένου να αναλυθεί. Αυτή η ακολουθία δεδομένων καθορίζει πλήρως την ιστορία των ατομικών κινήσεων από πλευράς ενός διαδοχικού, χρονικά-εξαρτώμενου συνόλου μοριακών διαμορφώσεων από την προσομοίωση και ενός μεγαλύτερου συνόλου ιδιοτήτων προερχόμενο από το τροχιακό (όπως ενέργειες, μήκη δεσμών και κατανομές γωνιών).

Παρόλο που ορισμένες από τις ιδιότητες που προέρχονται από το τροχιακό είναι σχετικά εύκολο να προσδιοριστούν (όπως το RMSD από την αρχική δομή ή άλλες επιλεγόμενες διαμορφώσεις), άλλες ιδιότητες είναι δυσκολότερο να προσδιοριστούν και είναι πιο χρονοβόρες. Επιπλέον οι σχέσεις μεταξύ των μοριακών διαμορφώσεων είναι κρυμμένες στην πολυπλοκότητα των δεδομένων. Ένας πολύ χρήσιμος τρόπος για να διαφανούν οι σχέσεις μεταξύ των διαμορφώσεων είναι η ομαδοποίηση των μοριακών διαμορφώσεων σε υποσύνολα με κριτήριο την ομοιότητα των διαμορφώσεων του τροχιακού. Η ομαδοποίηση είναι μία γενική τεχνική άντλησης δεδομένων που μπορεί να εφαρμοστεί σε οποιαδήποτε συλλογή πληροφοριακών στοιχείων (σημεία), για την οποία είναι διαθέσιμη μία συνάρτηση μέτρησης της απόστασης ανάμεσα σε ζεύγη σημείων. Ένας αλγόριθμος ομαδοποίησης διαιρεί τα σημεία σε μια συλλογή από ασυνεχή σύνολα τα οποία ονομάζονται συστάδες (clusters). Τα σημεία ενός cluster είναι, ιδανικά, περισσότερο όμοια μεταξύ τους, σε σχέση με τα σημεία ενός άλλου cluster.

Στην παρούσα πτυχιακή εργασία για την ομαδοποίηση διαμορφώσεων των τροχιακών 99SB, ILDN και protonated χρησιμοποιήθηκε η ανάλυση των κυρίαρχων συνιστωσών στο καρτεσιανό σύστημα (Cartesian principal component analysis) και η ανάλυση κυρίαρχων συνιστωσών των δίδρων γωνιών ϕ/ψ (dihedral principal component analysis). Ένας τρόπος ομαδοποίησης διαμορφώσεων μπορεί να γίνει και μέσω της ανάλυσης RMSD (Gordon et al., 1992), που προηγήθηκε στην ενότητα 4.1.

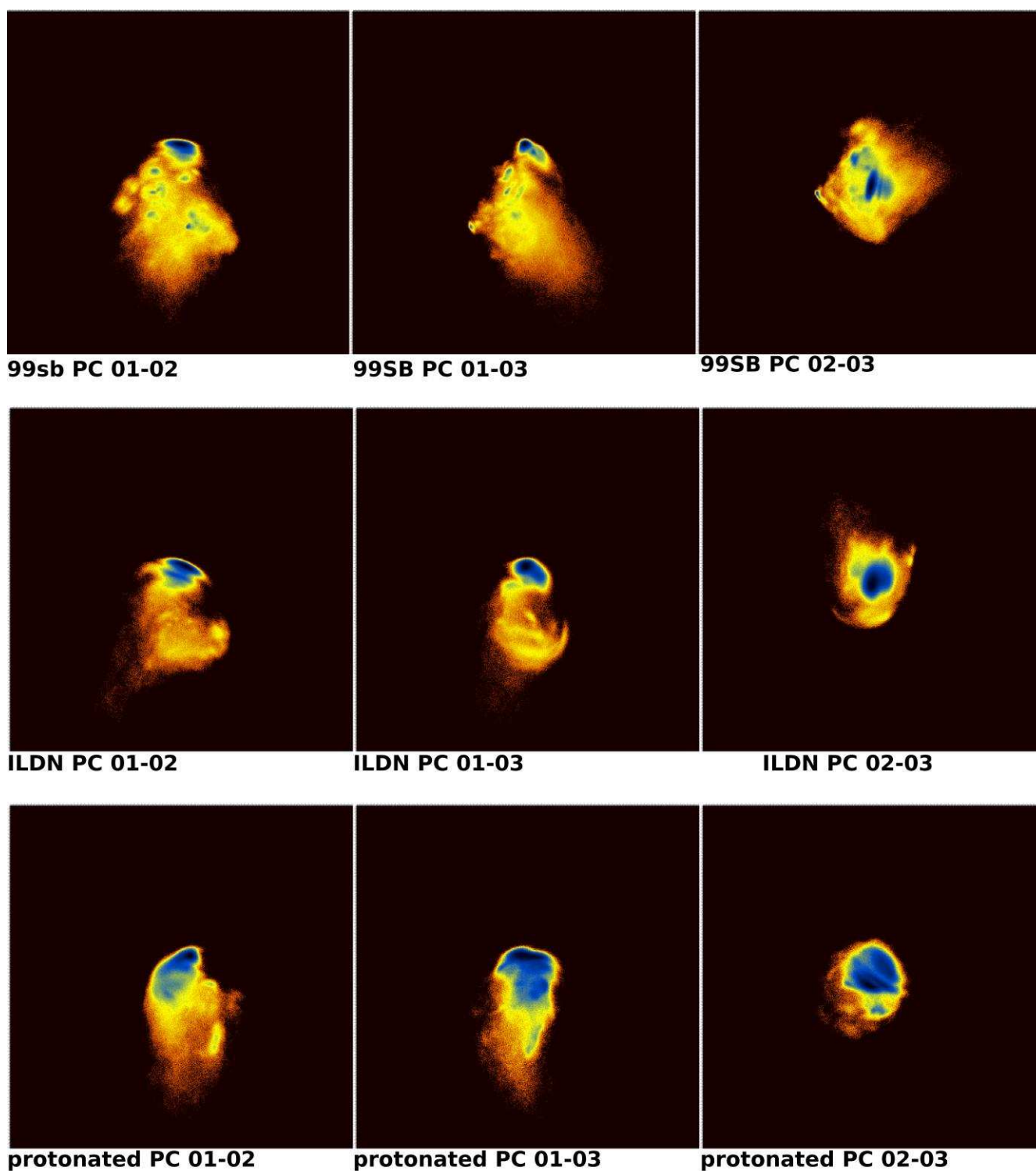
4.3.1 Cartesian Principal Component Analysis

Η κεντρική ιδέα της ανάλυσης κυρίων συνιστωσών (PCA) είναι η μείωση των διαστάσεων από ένα σύνολο δεδομένων που αποτελείται από μεγάλο αριθμό αλληλένδετων μεταβλητών, διατηρώντας ταυτόχρονα όσο το δυνατόν περισσότερο τη διασπορά που παρουσιάζουν τα δεδομένα. Αυτό επιτυγχάνεται με τη μετατροπή σε ένα νέο σύνολο μεταβλητών, τις κύριες συνιστώσες (principal components, PCs), οι οποίες είναι μη συσχετιζόμενες (η συνδιακύμανση δυο οποιονδήποτε συνιστωσών είναι μηδενική) και ταξινομούνται έτσι ώστε ορισμένες εκ των πρώτων να διατηρούν το μεγαλύτερο μέρος της διακύμανσης που παρουσιάζεται σε όλες τις αρχικές μεταβλητές (I.T. Jolliffe, 2002).

Η προσέγγιση αυτή βασίζεται σε έναν πίνακα συνδιακύμανσης (covariance matrix), ο οποίος παρέχει πληροφορίες σχετικά με τις συσχετίσεις (correlations) δύο σημείων του συστήματος. Η PCA αντιπροσωπεύει ένα γραμμικό μετασχηματισμό του πίνακα συνδιακύμανσης σε διαγώνιο πίνακα και επομένως αφαιρεί τις στιγμιαίες γραμμικές συσχετίσεις μεταξύ των μεταβλητών (Alexandros Altis et al., 2007). Κατατάσσοντας τις ιδιοτιμές (eigenvalues) του μετασχηματισμού μειούμενα, έχει δεχθεί ότι ένα μεγάλο τμήμα των διακυμάνσεων του συστήματος μπορεί να περιγραφεί από μερικές μόνο κύριες

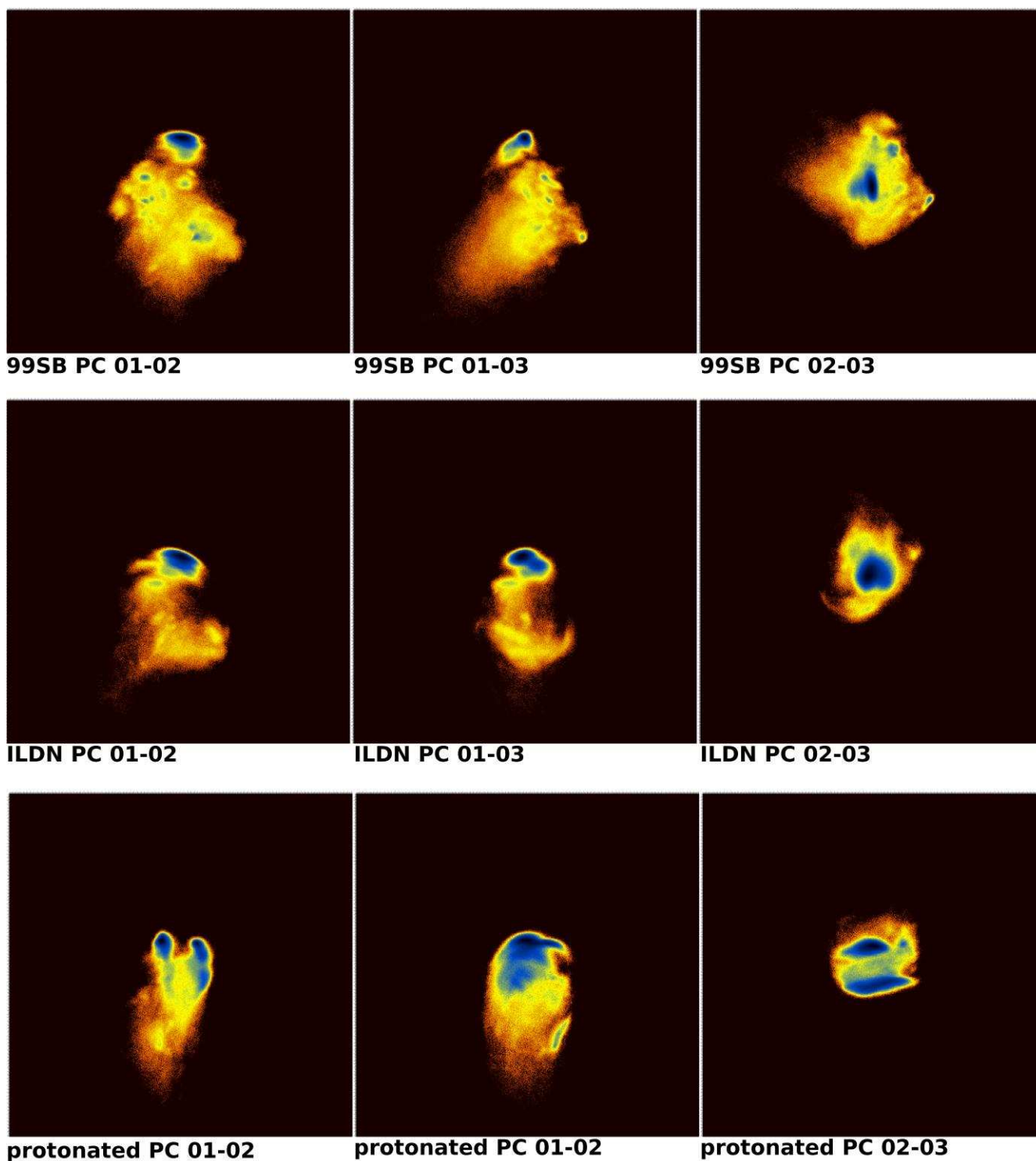
συνιστώσες (eigenvectors) (M. Karplus et al., 1991, A.E Garcia, 1992, A. Amadei, et al., 1993, S. Hayward et al., 1993, O.M Becker, 1997, O.F Lange et al., 2006, F. Noe et al., 2006). Ο τύπος PCA που χρησιμοποιεί καρτεσιανές συντεταγμένες για τις αναλύσεις ονομάζεται Cartesian Principal Component Analysis (cPCA).

Στις αναλύσεις cPCA που πραγματοποιήθηκαν με το πρόγραμμα CARMA (N.M. Glykos, 2006) ελήφθησαν υπόψιν αρχικά μόνο τα άτομα του πεπτιδικού σκελετού και εν συνεχεία όλα τα βαριά άτομα, έτσι ώστε να εξετάσουμε την κίνηση του πεπτιδίου συμπεριλαμβανομένων και των πλευρικών αλυσίδων πέραν του πεπτιδικού σκελετού. Στην εικόνα 4.21 παρουσιάζεται το ενεργειακό τοπίο που προκύπτει από την προβολή των τροχιακών στο επίπεδο που ορίζεται από τους τρεις πρώτους eigenvectors με τα μεγαλύτερα eigenvalues, ανά ζεύγη. Έχει εφαρμοστεί fitting (απομάκρυνση περιστροφών και μεταθέσεων) στα Ca άτομα του πεπτιδίου και η ανάλυση cPCA αφορά τα άτομα του σκελετού. Τα πολυπληθέστερα cluster εμφανίζονται με μπλε χρώμα. Ιδανικά θα θέλαμε να υπάρχει ένα μόνο συμπαγές cluster που θα αντιπροσωπεύει μία κυρίαρχη διαμόρφωση. Στην εικόνα 4.22 παρουσιάζεται η ανάλυση cPCA λαμβάνοντας υπόψιν όλα τα βαριά άτομα των τροχιακών.



Εικόνα 4.21

Παρουσίαση των ενεργειακών τοπίων (ΔG , energy landscapes) στο επίπεδο των τριών κυρίαρχων συνιστωσών (Pcs) ανά ζεύγη (οριζοντίως) για τα τροχιακά 99SB, ILDN και protonated (καθέτως). Στην ανάλυση χρησιμοποιήθηκαν τα 68 άτομα (Ca, C, N, O) του πεπτιδικού σκελετού.



Εικόνα 4.22

Παρουσίαση των ενεργειακών τοπίων (ΔG , energy landscapes) στο επίπεδο των τριών κυρίαρχων συνιστωσών (Pcs) ανά ζεύγη (οριζοντίως) για τα τροχιακά 99SB, ILDN και protonated (καθέτως). Στην ανάλυση χρησιμοποιήθηκαν 135 άτομα, που αντιστοιχούν στα βαρέα άτομα.

Πριν την ανάλυση cPCA εφαρμόστηκε fitting και στις δύο περιπτώσεις στα άτομα Ca, γεγονός που σημαίνει πως έχουν απομακρυνθεί μεταθέσεις και περιστροφές και έχουν αποκλειστεί τα άτομα υδρογόνου και τα μόρια νερού. Επιπλέον έχουν εξαιρεθεί από τους υπολογισμούς το πρώτο και το τελευταίο αμινοξύ του πεπτιδίου (Met1, Val17) λόγω της αυξημένης κινητικότητας που παρουσιάζουν. Και ως δομή αναφοράς για την υπέρθεση όλων των δομών του εκάστοτε τροχιακού έχει οριστεί η πρώτη δομή του τροχιακού.

Φαίνεται ξεκάθαρα ο σχηματισμός ενός μόνο cluster στο τροχιακό ILDN. Στο τροχιακό 99SB φαίνεται να σχηματίζεται ένα αρκετά συμπαγές cluster, αλλά ταυτοχρόνως σχηματίζονται και ορισμένα πολύ μικρότερα. Τέλος στο τροχιακό protonated στην πρώτη ανάλυση cPCA φαίνεται να σχηματίζεται ένα αρκετά μεγάλο τροχιακό και ένα πολύ μικρότερο, ενώ στην δεύτερη ανάλυση που συμπεριλαμβάνονται οι πλευρικές ομάδες φαίνεται να σχηματίζονται δύο διακριτά clusters τα οποία βρίσκονται σε κοντινές αποστάσεις στον χάρτη της ελεύθερης ενέργειας. Τα clusters που προέκυψαν από κάθε ανάλυση μαζί με το επί τοις εκατό ποσοστό του καθενός φαίνονται στους πίνακες 4.3 και 4.4.

	Cluster 1	Cluster 2	Cluster 3	Variance explained
99SB	37.9%	0.6%	0.5%	96%
ILDN	71.4%	-	-	98%
protonated	59.9%	0.3%	-	92%

Πίνακας 4.3

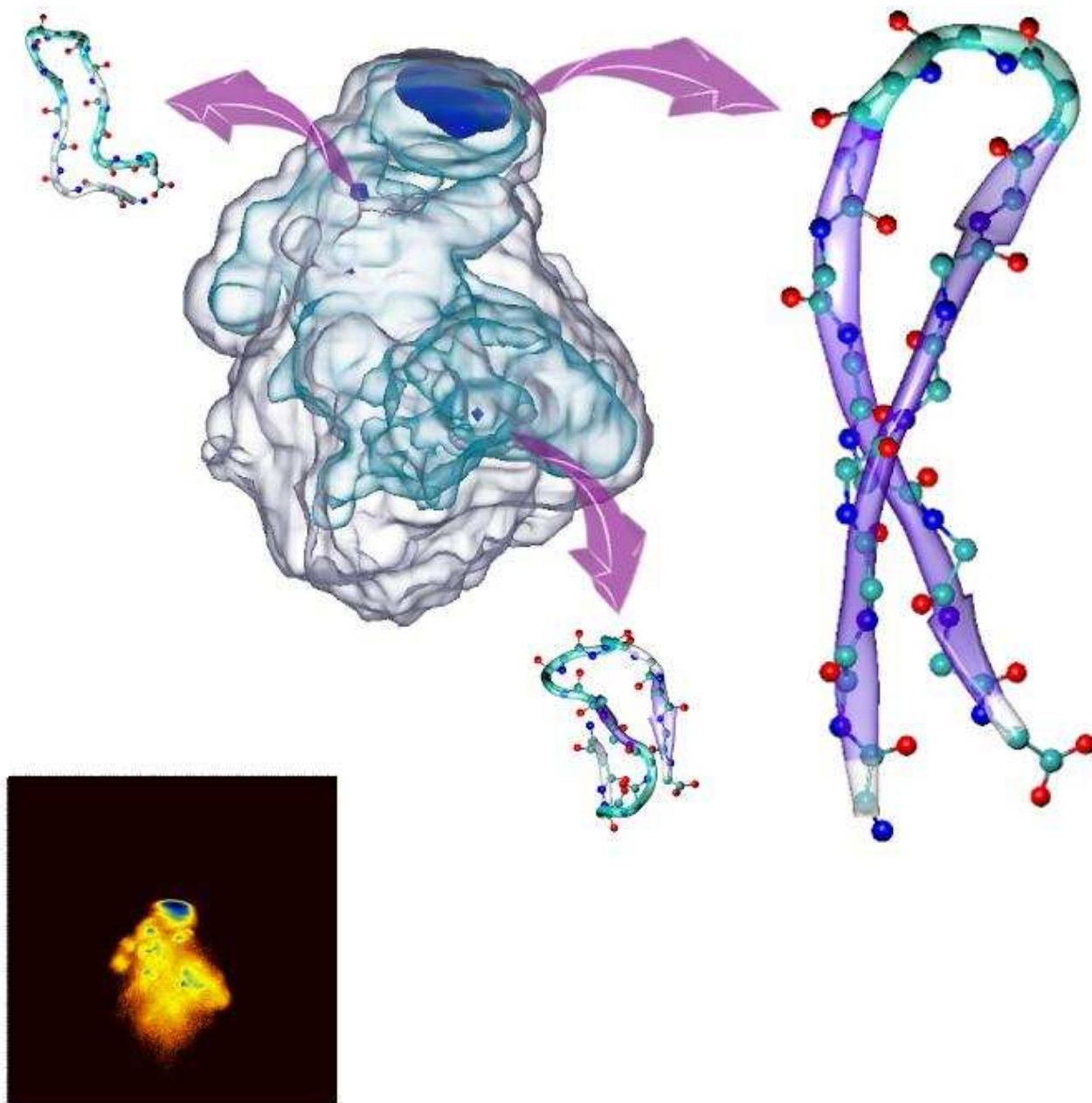
Συγκριτική παρουσίαση των clusters που σχηματίσθηκαν από την ανάλυση cPCA για τα άτομα της κύριας αλυσίδας των τροχιακών 99SB, ILDN και protonated. Στις στήλες 2 – 4 αναγράφεται το ποσοστό που καταλαμβάνει το κάθε cluster σε σχέση με το συνολικό μέγεθος του τροχιακού, ενώ στην πέμπτη στήλη αναγράφεται το ποσοστό της διακύμανσης που καλύπτεται από τα clusters.

	Cluster 1	Cluster 2	Cluster 3	Variance explained
99SB	37.5%	0.5%	0.3%	96%
ILDN	69.6%	-	-	97%
protonated	20.3%	10.5%	-	75%

Πίνακας 4.4

Συγκριτική παρουσίαση των clusters που σχηματίστηκαν από την ανάλυση cPCA για τα βαριά άτομα των τροχιακών 99SB, ILDN και protonated. Στις στήλες 2 – 4 αναγράφεται το ποσοστό που καταλαμβάνει το κάθε cluster σε σχέση με το συνολικό μέγεθος του τροχιακού, ενώ στην πέμπτη στήλη αναγράφεται το ποσοστό της διακύμανσης που καλύπτεται από τα clusters.

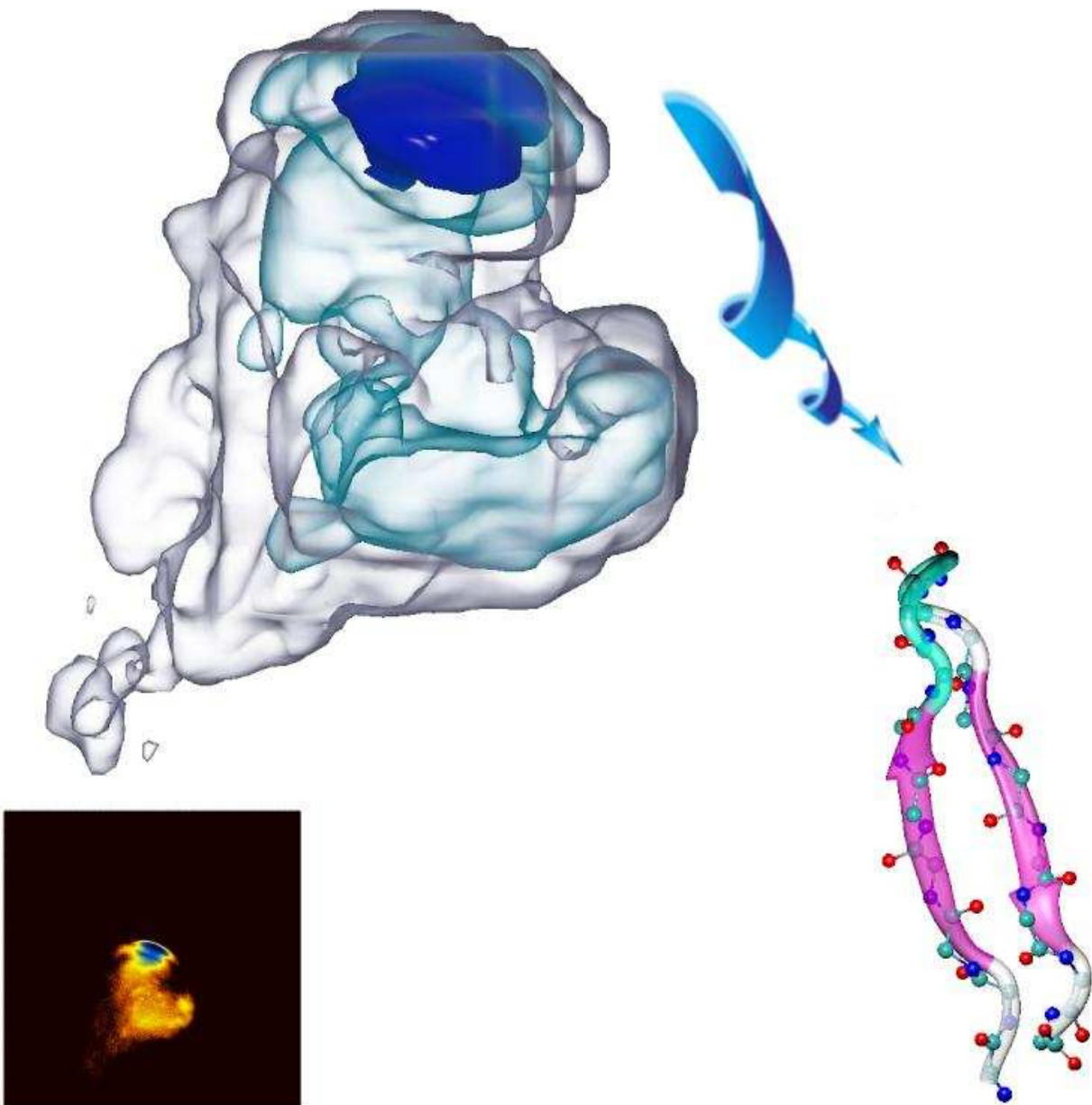
Εν συνεχεία, από κάθε cluster που εντοπίστηκε, βρέθηκε η μέση δομή και στη συνέχεια η κοντινότερη δομή του cluster προς τη μέση δομή με βάση το RMSD, η οποία αποτελεί την αντιπροσωπευτική δομή του cluster. Στις εικόνες 4.23, 4.24 και 4.25 παρουσιάζεται το τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης των τροχιακών 99SB, ILDN και protonated με τις αντίστοιχες αντιπροσωπευτικές δομές για το κάθε cluster. Το ενεργειακό τοπίο προέκυψε από την ανάλυση cPCA χρησιμοποιώντας μόνο τα 68 άτομα της κύριας αλυσίδας του πεπτιδίου.



Εικόνα 4.23

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού 99SB. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το Cartesian PCA για τα 68 άτομα της κύριας αλυσίδας του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού 99SB στους τρεις principal components της ανάλυσης Cartesian PCA. Παρουσιάζονται τρία επίπεδα ισο-επιφάνειας και οι μπλε κορυφές αντιστοιχούν στα cluster που σχηματίσθηκαν.

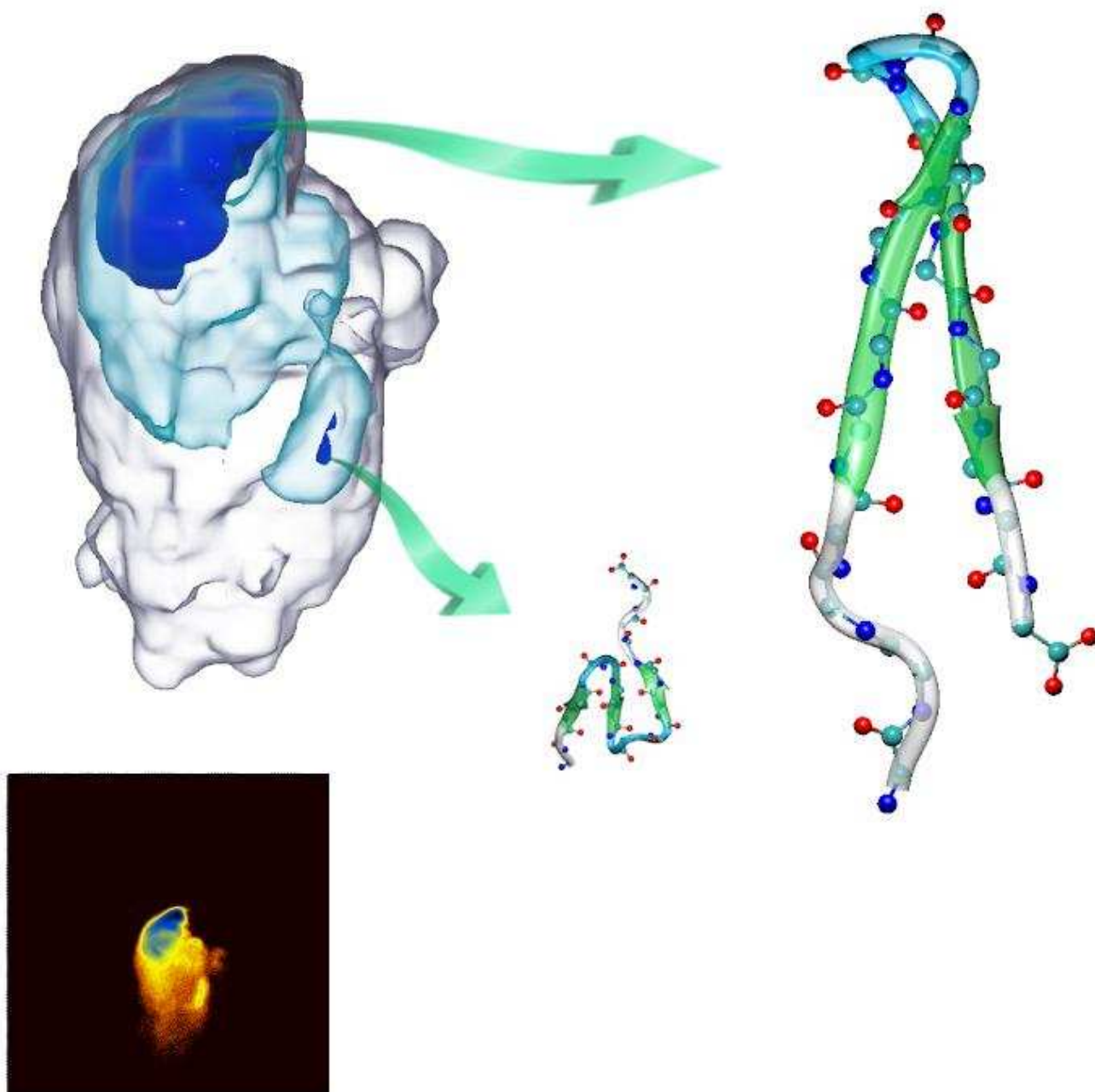
Κάθε cluster συνδέεται με βέλος με την αντιπροσωπευτική δομή του (κοντινότερη στη μέση δομή του cluster). Το μέγεθος των δομών αντιπροσωπεύει το μέγεθος του cluster σε σχέση με το συνολικό μέγεθος της προσομοίωσης.



Εικόνα 4.24

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού ILDN. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το

*Cartesian PCA για τα 68 άτομα της κύριας αλυσίδας του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού 99SB στους τρεις *principal components* της ανάλυσης *Cartesian PCA*. Παρουσιάζονται τρία επίπεδα ισο-επιφάνειας και οι μπλε κορυφές αντιστοιχούν στα *cluster* που σχηματίστηκαν. Κάθε *cluster* συνδέεται με βέλος με την αντιπροσωπευτική δομή του (κοντινότερη στη μέση δομή του *cluster*). Το μέγεθος των δομών αντιπροσωπεύει το μέγεθος του *cluster* σε σχέση με το συνολικό μέγεθος της προσομοίωσης.*

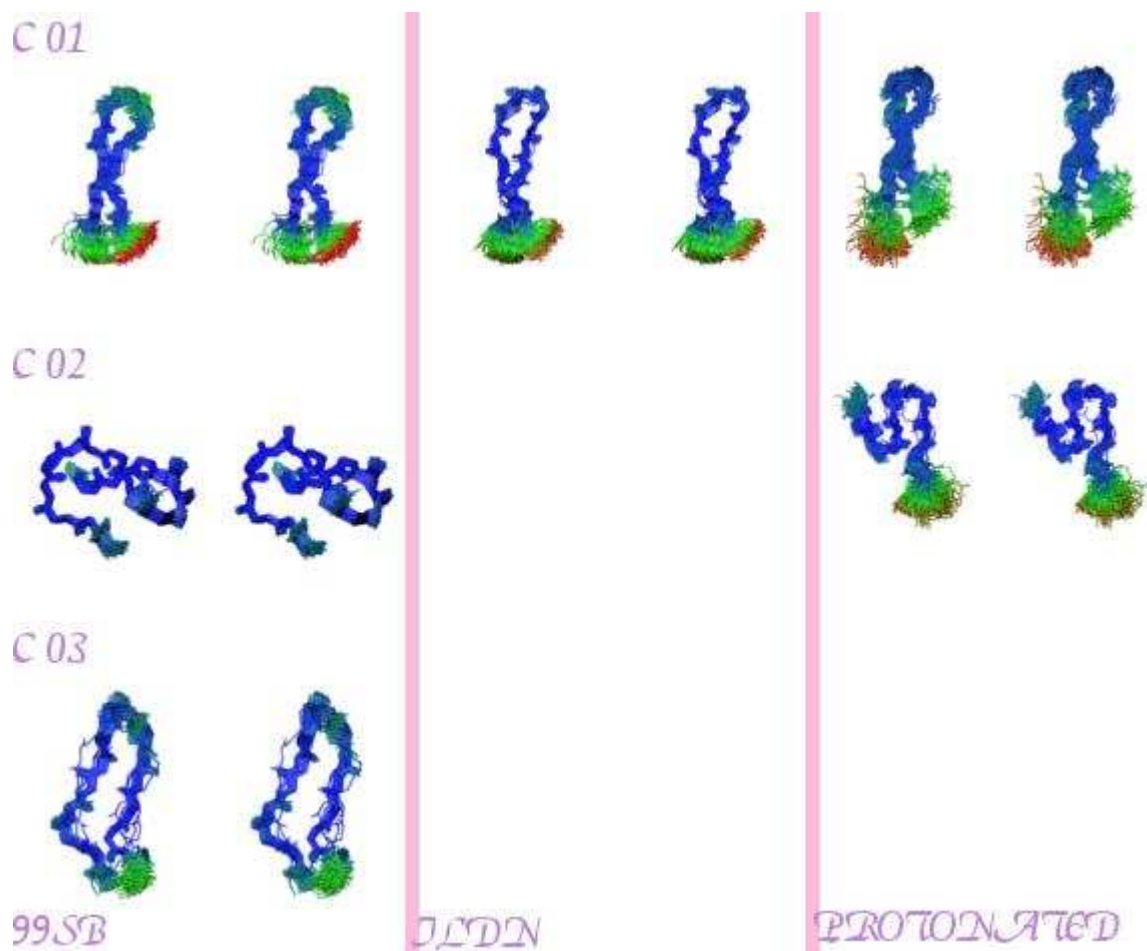


Εικόνα 4.25

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού protonated. Κάτω

αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το *Cartesian PCA* για τα 68 άτομα της κύριας αλυσίδας του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού 99SB στους τρεις *principal components* της ανάλυσης *Cartesian PCA*. Παρουσιάζονται τρία επίπεδα ισο-επιφάνειας και οι μπλε κορυφές αντιστοιχούν στα *cluster* που σχηματίστηκαν. Κάθε *cluster* συνδέεται με βέλος με την αντιπροσωπευτική δομή του (κοντινότερη στη μέση δομή του *cluster*). Το μέγεθος των δομών αντιπροσωπεύει το μέγεθος του *cluster* σε σχέση με το συνολικό μέγεθος της προσομοίωσης.

Από τις αναλύσεις *Cartesian PCA* μπορούμε να συμπεράνουμε τον σχηματισμό β – φουρκέτας σε όλα τα κυρίαρχα *clusters*. Μάλιστα οι δομές που παρουσιάζονται είναι αρκετά όμοιες με τις δομές που εξήχθησαν από τις αναλύσεις RMSD (Εικόνα 4.19). Κρατώντας σταθερά τα άτομα Ca σε όλη την ανάλυση και εκτελώντας την *Cartesian PCA* μόνο για τα άτομα του σκελετού η προσοχή επικεντρώνεται στην κίνηση του και συνεπώς στην δημιουργία της β – φουρκέτας που είναι και η φυσική διαμόρφωση. Στην εικόνα που ακολουθεί παρουσιάζεται η υπέρθεση των δομών από το κάθε *cluster* που προέκυψε από την ανάλυση *Cartesian PCA* για τα άτομα του πεπτιδικού σκελετού (Ca, C, N, O).



Εικόνα 4.26

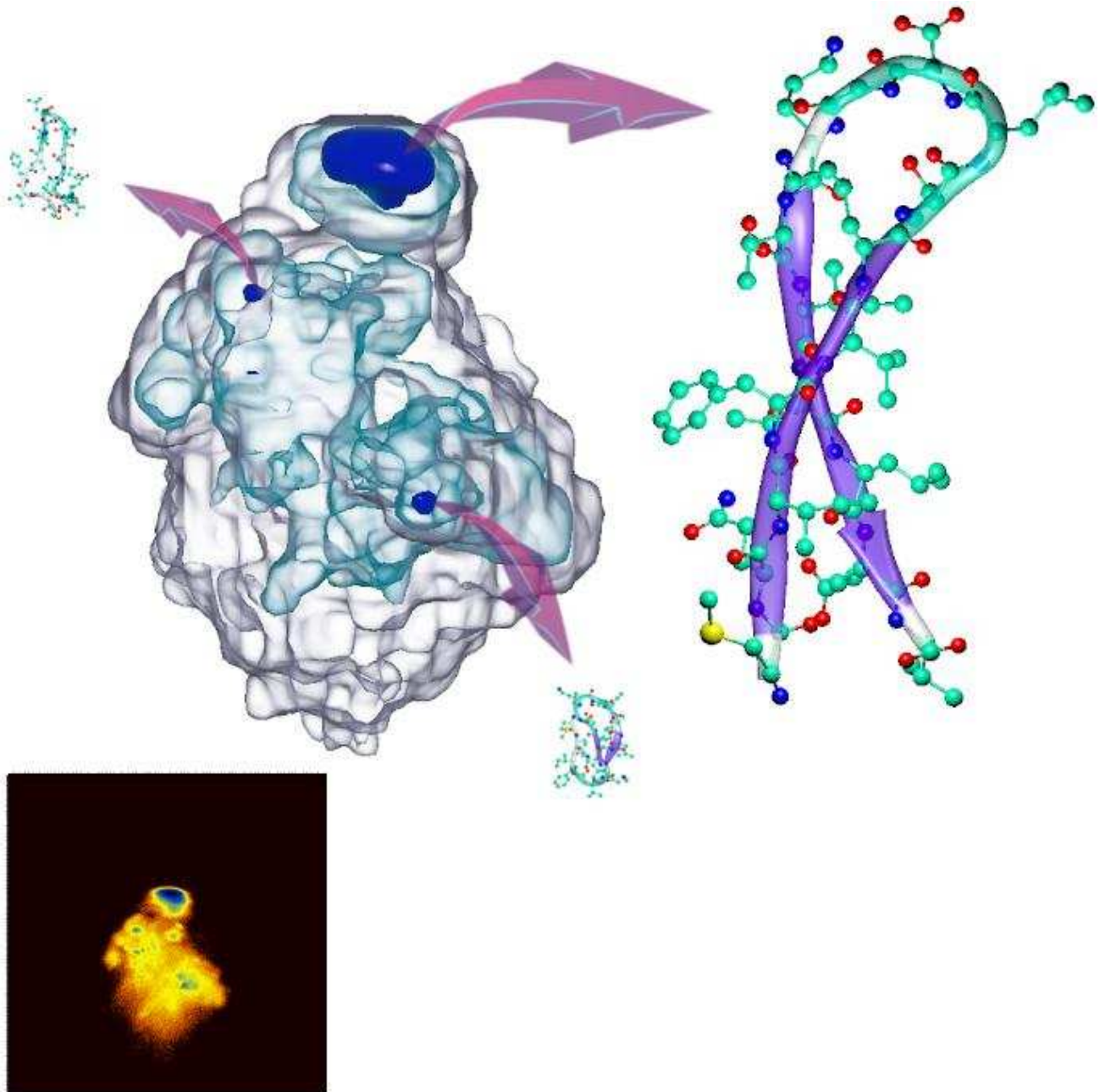
Παρουσίαση υπέρθεσης των αντιπροσωπευτικών δομών, σε stereo view, των clusters, όπως προέκυψε από την ανάλυση Cartesian PCA για τα τρία τροχιακά. Για την υπέρθεση χρησιμοποιήθηκαν μόνο τα άτομα του πεπτιδικού σκελετού (Ca, C, N, O). Ο χρωματισμός έγινε με βάση τις ατομικές διακυμάνσεις, η κλίμακα διατηρείται σταθερή για τα clusters που ανήκουν στο ίδιο τροχιακό.

Στη συνέχεια πραγματοποιήθηκαν οι αναλύσεις Cartesian PCA για όλα τα βαρέα άτομα (εξαιρουμένων των ατόμων υδρογόνου) και χρησιμοποιώντας σταθερά τα άτομα του πεπτιδικού σκελετού. Οι τρισδιάστατες απεικονίσεις του ενεργειακού τοπίου των τριών τροχιακών αναπαράγονται στις εικόνες 4.27, 4.28 και 4.29. Σε κάθε cluster αντιστοιχεί μία κυρίαρχη δομή, η οποία είναι η κοντινότερη στη μέση δομή του cluster. Η ανάλυση αυτή πραγματοποιήθηκε προκειμένου να εξετάσουμε την κινητικότητα των πλευρικών ομάδων πέραν του πεπτιδικού

σκελετού. Εξετάζοντας την κινητικότητα των πλευρικών ομάδων μπορούμε να εντοπίσουμε το σχηματισμό υδρόφοβου κλωβού εκατέρωθεν της κύριας αλυσίδας, καθώς και την κινητικότητα του μεταλλαγμένου καταλοίπου (Asp9), το οποίο φαίνεται να ευθύνεται για την αυξημένη σταθερότητα του πεπτιδίου, καθώς σχηματίζει επαφές με τις πλευρικές ομάδες γειτονικών καταλοίπων.

Στην κυρίαρχη δομή του τροχιακού 99SB έχει σχηματιστεί η δομή της β-φουρκέτας. Η πλευρική ομάδα του καταλοίπου Asp9 φαίνεται να σχηματίζει αλληλεπιδράσεις με την πλευρική ομάδα του καταλοίπου Lys11. Η πλευρική ομάδα του καταλοίπου Thr14 βρίσκεται πολύ κοντά με τις πλευρικές ομάδες των καταλοίπων Phe4 και Gln2 και τα κατάλοιπα Lys6 και Thr12 σχηματίζουν επαφές μεταξύ τους. Με αυτή την διάταξη των πλευρικών ομάδων επιβεβαιώνεται η NRM δομή, καθώς οι επαφές που σχηματίζονται αντιστοιχούν στον υδρόφοβο κλωβό της μιας επιφάνειας της φουρκέτας που προσδιορίστηκε πειραματικά από την R. Zerella. Στην άλλη επιφάνεια της φουρκέτας τα κατάλοιπα Ile3, Val5, Ile13 και Leu15 φαίνεται να σχηματίζουν τον δεύτερο κλωβό του πεπτιδίου, όπως προέβλεψε και η μελέτη της NMR δομής, παρόλο που τα δεδομένα δεν ήταν επαρκή για την εξακρίβωση των επαφών που δημιουργούνται. Παρόμοια αποτελέσματα για τη θέση των πλευρικών αλυσίδων βρέθηκαν και για το κυρίαρχο cluster της προσομοίωσης ILDN. Οι πλευρικές αλυσίδες πακετάρονται έτσι ώστε να σχηματιστούν οι δύο επιφάνειες εκατέρωθεν της φουρκέτας, αλλά η πλευρική ομάδα του καταλοίπου Asp9 φαίνεται να έχει μεγαλύτερη απόσταση από την πλευρική ομάδα του καταλοίπου Lys11. Τέλος, στο κυρίαρχο cluster του τροχιακού protonated φαίνεται να αλλάζει το πρότυπο των επαφών μεταξύ των πλευρικών αλυσίδων που σχηματίζουν τις δύο επιφάνειες της φουρκέτας. Τα κατάλοιπα Gln2, Phe4 και Lys6 φαίνεται να βρίσκονται στην ίδια πλευρά της β-φουρκέτας, όπως και στις προηγούμενες δομές, ενώ τα κατάλοιπα Thr12 και Thr14 φαίνεται να βρίσκονται στην άλλη πλευρά της φουρκέτας και τη θέση τους φαίνεται να πήραν τα κατάλοιπα Ile13, Leu15. Επιπλέον στην ίδια πλευρά βρίσκεται και το κατάλοιπο

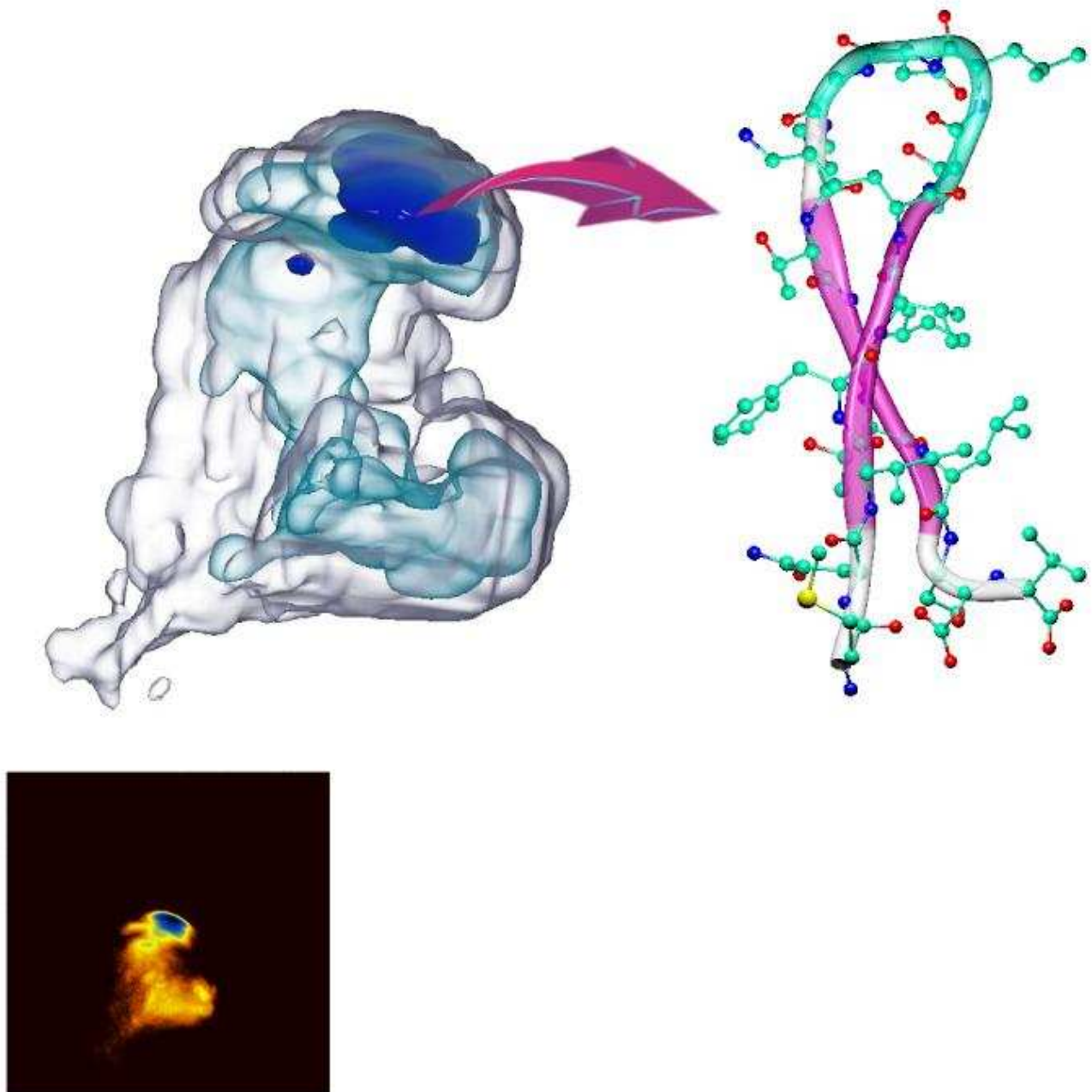
Val17. Συνολικά έχουμε στη μία πλευρά τα κατάλοιπα Gln2, Phe4, Lys6, Ile13, Leu15 και Val17, ενώ στην αντίθετη πλευρά έχουμε τα κατάλοιπα Met1, Ile3, Val5, Thr12, Thr14 και Glh16 (προτονιομένο κατάλοιπο Glu).



Εικόνα 4.27

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού 99SB. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το

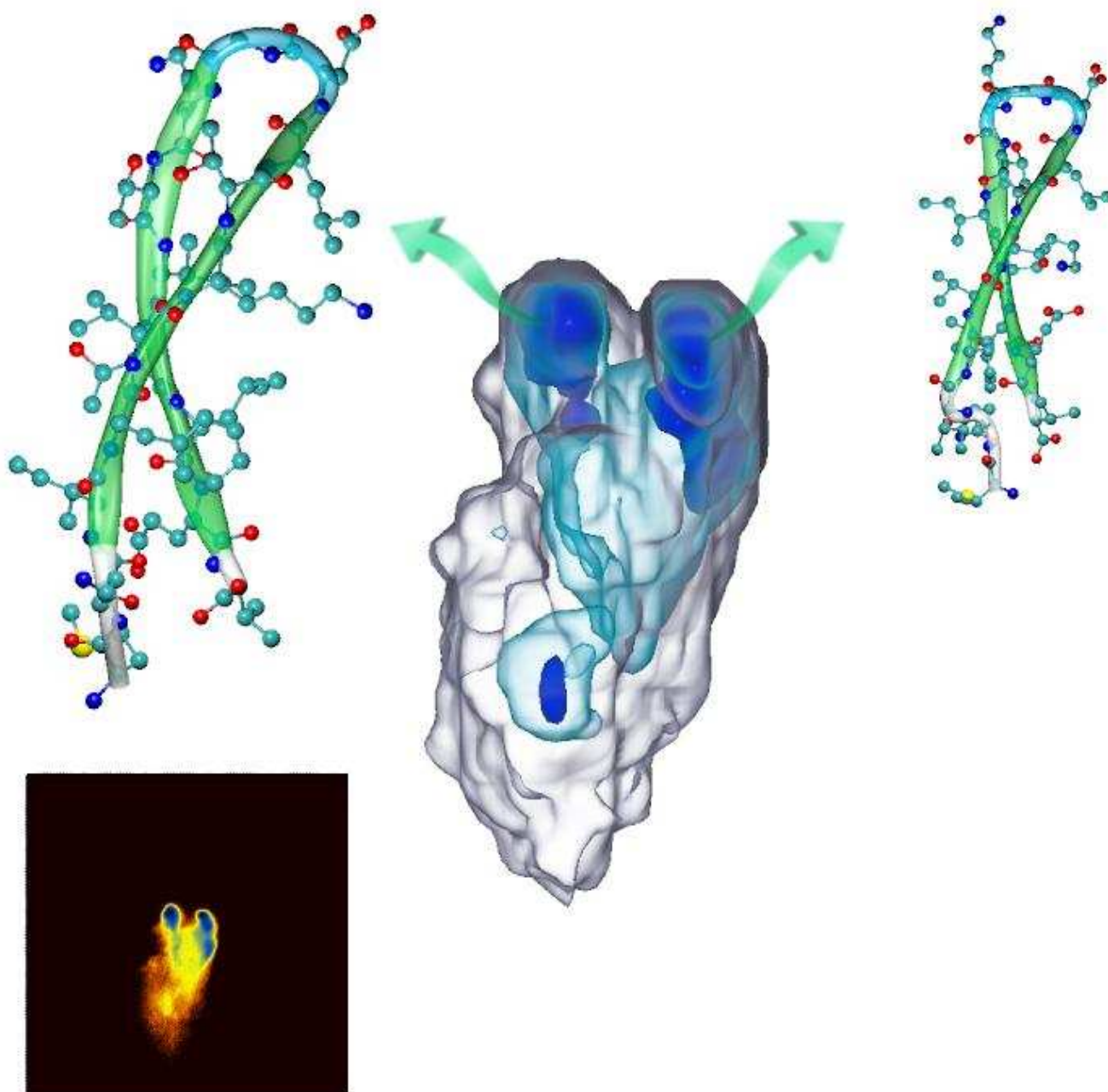
Cartesian PCA για τα 135 βαρέα άτομα του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού 99SB στους τρεις principal components της ανάλυσης Cartesian PCA. Παρουσιάζονται τρία επίπεδα ισο-επιφάνειας και οι μπλε κορυφές αντιστοιχούν στα cluster που σχηματίστηκαν. Κάθε cluster συνδέεται με βέλος με την αντιπροσωπευτική δομή του (κοντινότερη στη μέση δομή του cluster). Το μέγεθος των δομών αντιπροσωπεύει το μέγεθος του cluster σε σχέση με το συνολικό μέγεθος της προσομοίωσης.



Εικόνα 4.28

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού ILDN. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το Cartesian PCA για τα 135 βάρια άτομα του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού 99SB στους τρεις principal components της ανάλυσης Cartesian PCA. Παρουσιάζονται τρία επίπεδα ισο-επιφάνειας και οι μπλε κορυφές αντιστοιχούν στα cluster που σχηματίστηκαν. Κάθε cluster

συνδέεται με βέλος με την αντιπροσωπευτική δομή του (κοντινότερη στη μέση δομή του cluster). Το μέγεθος των δομών αντιπροσωπεύει το μέγεθος του cluster σε σχέση με το συνολικό μέγεθος της προσομοίωσης.



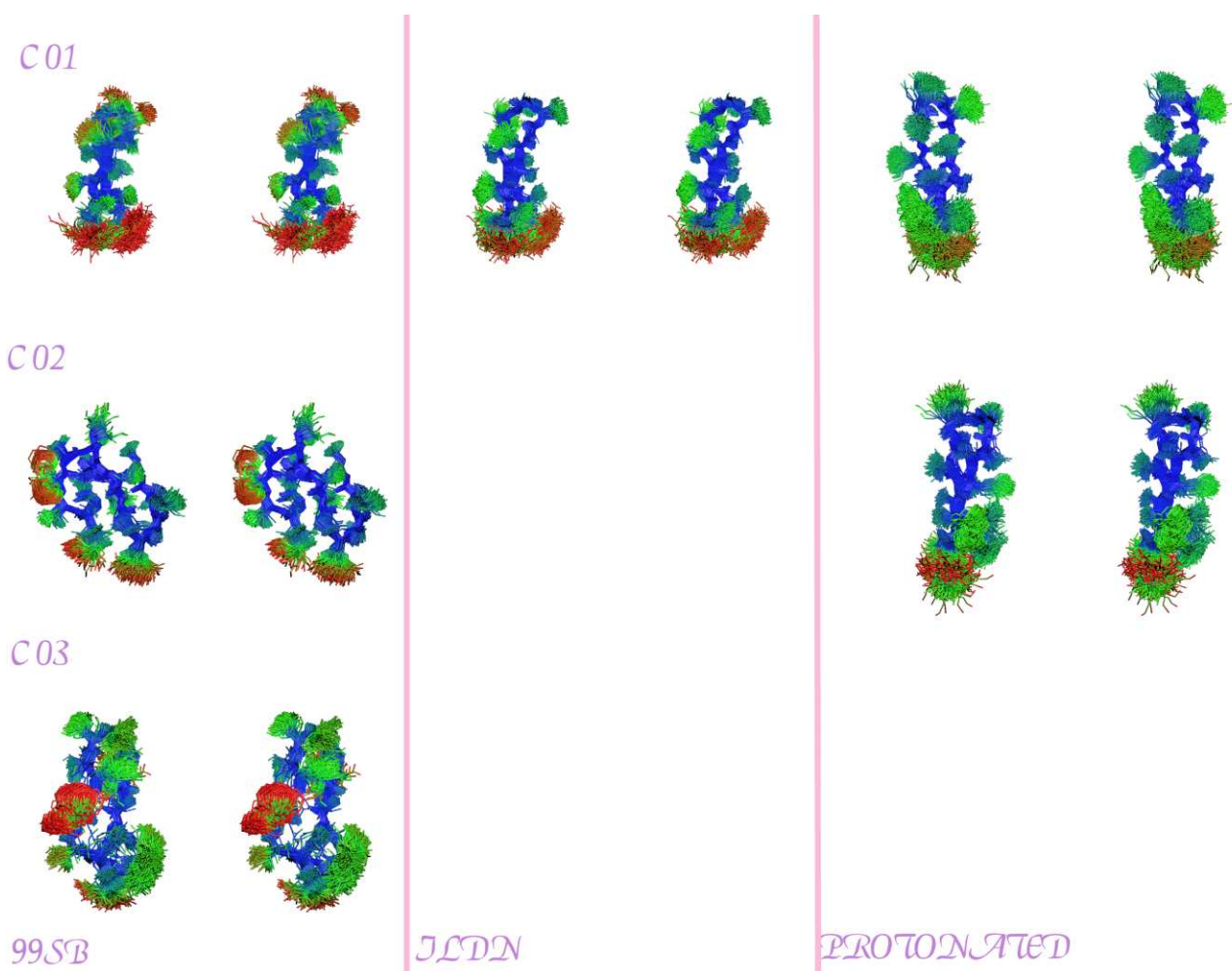
Εικόνα 4.29

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού protonated. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το

Cartesian PCA για τα 135 βαρέα άτομα του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού protonated στους τρεις principal components της ανάλυσης Cartesian PCA. Παρουσιάζονται τρία επίπεδα ισο-επιφάνειας και οι μπλε κορυφές αντιστοιχούν στα cluster που σχηματίστηκαν. Κάθε cluster συνδέεται με βέλος με την αντιπροσωπευτική δομή του (κοντινότερη στη μέση δομή του cluster). Το μέγεθος των δομών αντιπροσωπεύει το μέγεθος του cluster σε σχέση με το συνολικό μέγεθος της προσομοίωσης.

Από τα clusters που σχηματίστηκαν από την ανάλυση Cartesian PCA στα 135 βαρέα άτομα του πεπτιδίου πραγματοποιήθηκε υπέρθεση των δομών που σχηματίζουν το εκάστοτε cluster, προκειμένου να γίνει περισσότερο εμφανής η κινητικότητα των πλευρικών ομάδων. Στην εικόνα 4.30 παρουσιάζεται η υπέρθεση των δομών των clusters που παρήχθησαν από την ανάλυση cPCA. Στο πρώτο cluster του κάθε τροχιακού έχει σχηματισθεί η β – φουρκέτα, καθώς και οι υδρόφοβες επιφάνειες εκατέρωθεν του πεπτιδικού σκελετού. Οι πλευρικές ομάδες των ακραίων καταλοίπων παρουσιάζουν αυξημένη κινητικότητα σε σχέση με τις υπόλοιπες πλευρικές ομάδες, γεγονός που συμφωνεί με τις προηγούμενες αναλύσεις. Επιπλέον στο πρώτο cluster του τροχιακού 99SB παρουσιάζεται αυξημένη κινητικότητα και στις πλευρικές ομάδες των καταλοίπων Asp9, Leu8 και Lys11. Η πλευρική ομάδα του καταλοίπου Asp9 φαίνεται να ταλαντεύεται μεταξύ των πλευρικών ομάδων των καταλοίπων Leu8 και Lys11 σχηματίζοντας επαφές προκειμένου να σταθεροποιηθεί η δομή της β-φουρκέτας. Στο Cluster που προέκυψε για το τροχιακό της προσομοίωσης ILDN παρατηρείται επίσης αυξημένη κινητικότητα στα κατάλοιπα των άκρων του πεπτιδίου, σε σχέση με τα υπόλοιπα κατάλοιπα, αλλά η πλευρική ομάδα του μεταλλαγμένου καταλοίπου φαίνεται να βρίσκεται διαρκώς κοντά στην πλευρική ομάδα της Lys11, με την οποία σχηματίζει επαφές. Τέλος στο πρώτο cluster του τροχιακού protonated τα κατάλοιπα των άκρων συντηρούν την κινητικότητα που παρατηρήθηκε στα προηγούμενα τροχιακά και το μεταλλαγμένο κατάλοιπο βρίσκεται μεταξύ των καταλοίπων Leu8 και Lys11.

Είναι πιθανό να σχηματίζει, περιοδικά, επαφές και με τα δύο αυτά κατάλοιπα. Οι υπόλοιπες πλευρικές ομάδες φαίνεται να σχηματίζουν δύο σταθερές επιφάνειες εκατέρωθεν του πεπτιδικού σκελετού και στα τρία αντιπροσωπευτικά clusters. Παρόλο που το πρότυπο των επαφών μεταξύ πλευρικών ομάδων στο κυρίαρχο cluster του τροχιακού protonated αλλάζει, καθώς οι δύο επιφάνειες της β-φουρκέτας αποτελούνται από διαφορετικά κατάλοιπα σε σχέση με τις διαμορφώσεις των τροχιακών 99SB και ILDN, σχηματίζονται δύο σταθερές επιφάνειες εκατέρωθεν του πεπτιδικού σκελετού. Τα υπόλοιπα clusters αντιπροσωπεύουν διαφορετικές δομές σε σχέση με τα κυρίαρχα, αλλά το ποσοστό εμφάνισης των δομών αυτών στα τροχιακά είναι αρκετά χαμηλό (<10%) και θεωρούνται αμελητέα προς περαιτέρω ανάλυση.



Εικόνα 4.30

Παρουσίαση υπέρθεσης των clusters, σε stereo view, όπως προέκυψε από την ανάλυση Cartesian PCA για τα τρία τροχιακά. Για την υπέρθεση χρησιμοποιήθηκαν όλα τα βαρέα άτομα. Ο χρωματισμός έγινε με βάση τις ατομικές διακυμάνσεις, η κλίμακα διατηρείται σταθερή για τα clusters που ανήκουν στο ίδιο τροχιακό.

Σε όλες τις προσομοιώσεις που αναλύθηκαν τα ενεργειακά τοπία φαίνεται να έχουν από ένα έως τρία clusters και στα κυρίαρχα cluster να υπάρχει το ενεργειακό ελάχιστο του κάθε τροχιακού. Γεγονός που σημαίνει πως τα ενεργειακά τοπία που παράγονται χρησιμοποιώντας Cartesian PCA είναι σχετικά ομαλά. Προκειμένου να εξετασθεί εκτενέστερα το ενεργειακό τοπίο της

αναδίπλωσης των τριών τροχιακών ακολουθεί στην επόμενη παράγραφο η ανάλυση κυρίων συνιστωσών χρησιμοποιώντας τις διέδρες γωνίες.

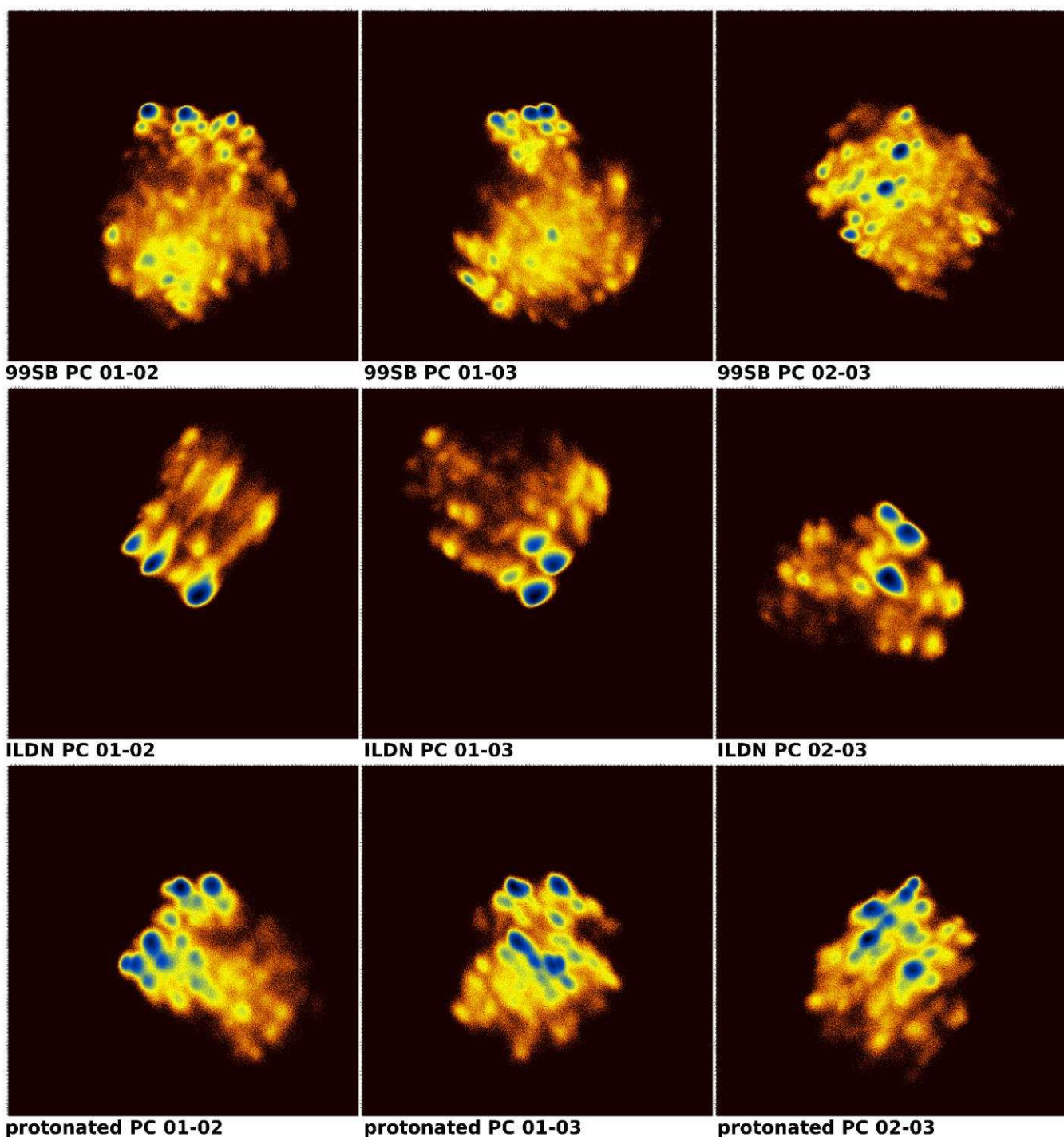
4.3.2 Dihedral angle Principal Component Analysis

Οι αναλύσεις κυρίων συνιστωσών χρησιμοποιώντας τις διέδρες γωνίες παρουσιάστηκαν για πρώτη φορά από την επιστημονική ομάδα του Yuguang Mu το 2005 οι οποίοι πραγματοποίησαν προσομοιώσεις αναστρέψιμης αναδίπλωσης και αποδιάταξης στην πέντα-αλανίνη. Το εύρημα τους από τις αναλύσεις dihedral angle PCA, όσο αφορά το ενεργειακό τοπίο των δύο διαδικασιών, ήταν διαφορετικό συγκριτικά με τις κλασικές αναλύσεις Cartesian PCA. Προκειμένου να μελετηθεί η διαμορφωτική δυναμική ενός μορίου είναι αναγκαίο να γίνει διάκριση μεταξύ της εσωτερικής κίνησης και της συνολικής κίνησης, η οποία είναι η μετάθεση και η συνολική περιστροφή. Η μετάθεση μπορεί εύκολα να διαχωριστεί θέτοντας σταθερό το κέντρο μάζας του μορίου. Ωστόσο, η εξάλειψη της συνολικής περιστροφής είναι απλή μόνο για κάποιο άκαμπτο σώμα ή στην περίπτωση των μικρών διακυμάνσεων γύρω από μία δομή αναφοράς. Το πρόβλημα της ολικής περιστροφής είναι δυνατόν να προσπελαστεί χρησιμοποιώντας τις αναλύσεις κυρίων συνιστωσών για τις διέδρες γωνίες (Mu et al., 2005).

Το ενεργειακό τοπίο που προκύπτει από τις αναλύσεις dPCA (dihedral PCA) εμφανίζει πολλά ενεργειακά ελάχιστα τα οποία αντιστοιχούν σε καλά καθορισμένες δομές. Συνεπώς είναι πιο τραχύ από το (λείο) ενεργειακό τοπίο που παράγουν οι αναλύσεις cPCA. Αυτό οφείλεται στο γεγονός ότι με την ανάλυση cPCA δεν μπορεί να διαχωριστεί η εσωτερική από την ολική κίνηση.

Στην παρούσα πτυχιακή εργασία θα πραγματοποιηθούν δύο αναλύσεις dPCA, η πρώτη αφορά τις διέδρες γωνίες φ και ψ , ενώ η δεύτερη αφορά τις διέδρες

γωνίες συμπεριλαμβανομένης και της χ_1 γωνίας. Οι αναλύσεις dPCA περιορίζονται στην κίνηση του πεπτιδικού σκελετού και αγνοούν τις πλευρικές ομάδες, συνεπώς δεν είναι απαραίτητη η απομάκρυνση ολικών περιστροφών και μεταθέσεων, όπως έγινε στην Cartesian PCA. Στην εικόνα 4.31 παρουσιάζονται τα ενεργειακά τοπία που προκύπτουν από την προβολή των τροχιακών στο επίπεδο που ορίζεται από τους τρεις πρώτους eigenvectors με τα μεγαλύτερα eigenvalues, ανά ζεύγη. Τα ενεργειακά τοπία προέκυψαν από την ανάλυση dPCA που αφορά τις γωνίες φ και ψ . Παρόλο που δε συμπεριλαμβάνονται οι πλευρικές ομάδες στην ανάλυση dPCA, τα ενεργειακά τοπία που προέκυψαν φαίνεται να έχουν περισσότερα clusters σε σχέση με τα ενεργειακά τοπία των αναλύσεων Cartesian PCA, όπως αναμένεται από αυτή την ανάλυση. Επιβεβαιώνεται συνεπώς ο κανόνας των τραχέων ενεργειακών πεδίων που προκύπτουν από την dPCA. Με την ανάλυση dPCA μπορούμε να δούμε περισσότερες ομαδοποιημένες δομές που αντιστοιχούν σε χαμηλά ενεργειακά επίπεδα και οι οποίες δε διαφαίνονται στην ανάλυση cPCA.



Εικόνα 4.31

Παρουσίαση των ενεργειακών τοπίων (ΔG , energy landscapes) στο επίπεδο των τριών κυρίαρχων συνιστωσών (Pcs) ανά ζεύγη (από δεξιά προς τα αριστερά) για τα τροχιακά 99SB, ILDN και protonated (από επάνω προς τα κάτω). Τα ενεργειακά τοπία είναι αποτέλεσμα των αναλύσεων dPCA για τις 32 ϕ και ψ γωνίες του πεπτιδίου.

Τα clusters που δημιουργήθηκαν στην ανάλυση dPCA για τους τρεις πρώτους eigenvectors είναι περισσότερα σε αριθμό σε σχέση με τα clusters της cPCA. Στον πίνακα 4.5 παρουσιάζονται τα clusters που σχηματίστηκαν για το κάθε τροχιακό, καθώς και το ποσοστό που καταλαμβάνει το κάθε cluster σε σχέση με το συνολικό τροχιακό. Στο τροχιακό 99SB δημιουργήθηκαν 13 clusters, εκ των οποίων τα δύο πρώτα φαίνεται να έχουν σημαντικό ποσοστό συμμετοχής στο τροχιακό καθώς τα υπόλοιπα καταλαμβάνουν λιγότερο του 10% του τροχιακού. Στο τροχιακό ILDN σχηματίστηκαν τρία clusters τα οποία έχουν σημαντικό ποσοστό συμμετοχής στο τροχιακό. Τέλος στο τροχιακό protonated σχηματίστηκαν 7 clusters, εκ των οποίων τα δύο πρώτα φαίνεται να είναι τα σημαντικότερα σε μέγεθος.

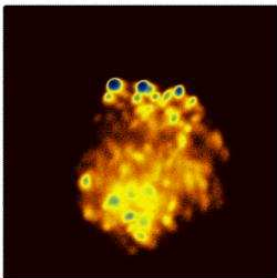
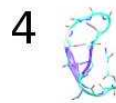
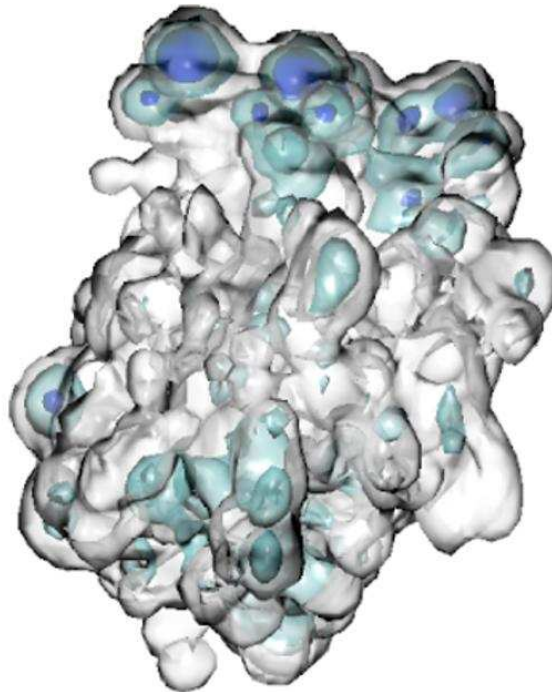
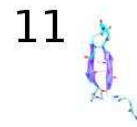
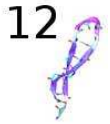
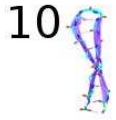
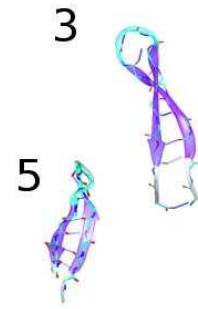
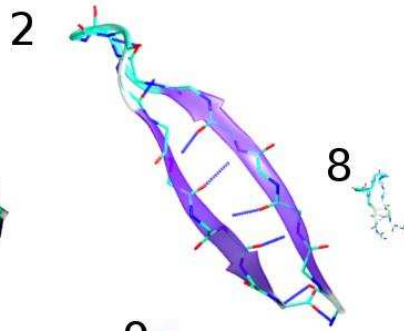
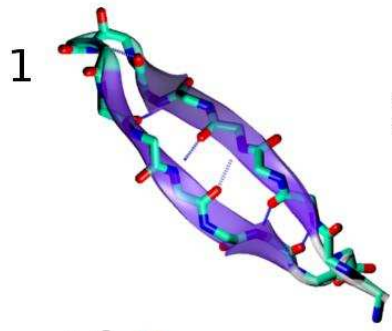
No. Cluster	1	2	3	4	5	6	7	8	9	10	11	12	13	Variance Explained
99SB	14.9 %	10.5 %	3.0%	0.8%	0.5%	0.6%	0.4%	0.5%	0.4%	0.4%	0.3%	0.2%	0.1%	96%
ILDN	38.2 %	20.1 %	6.4%	-	-	-	-	-	-	-	-	-	-	97%
Protonated	11.8 %	11.8 %	8.2%	7.6%	0.6%	0.4%	0.07%	-	-	-	-	-	-	90%

Πίνακας 4.5

Συγκριτική παρουσίαση των clusters που σχηματίστηκαν από την ανάλυση dPCA για τις γωνίες ϕ και ψ των τροχιακών 99SB, ILDN και protonated. Στις στήλες 2 – 12 αναγράφεται το ποσοστό που καταλαμβάνει το κάθε cluster σε σχέση με το συνολικό μέγεθος του τροχιακού, ενώ στην τελευταία στήλη αναγράφεται το ποσοστό της διακύμανσης που καλύπτεται από τα clusters.

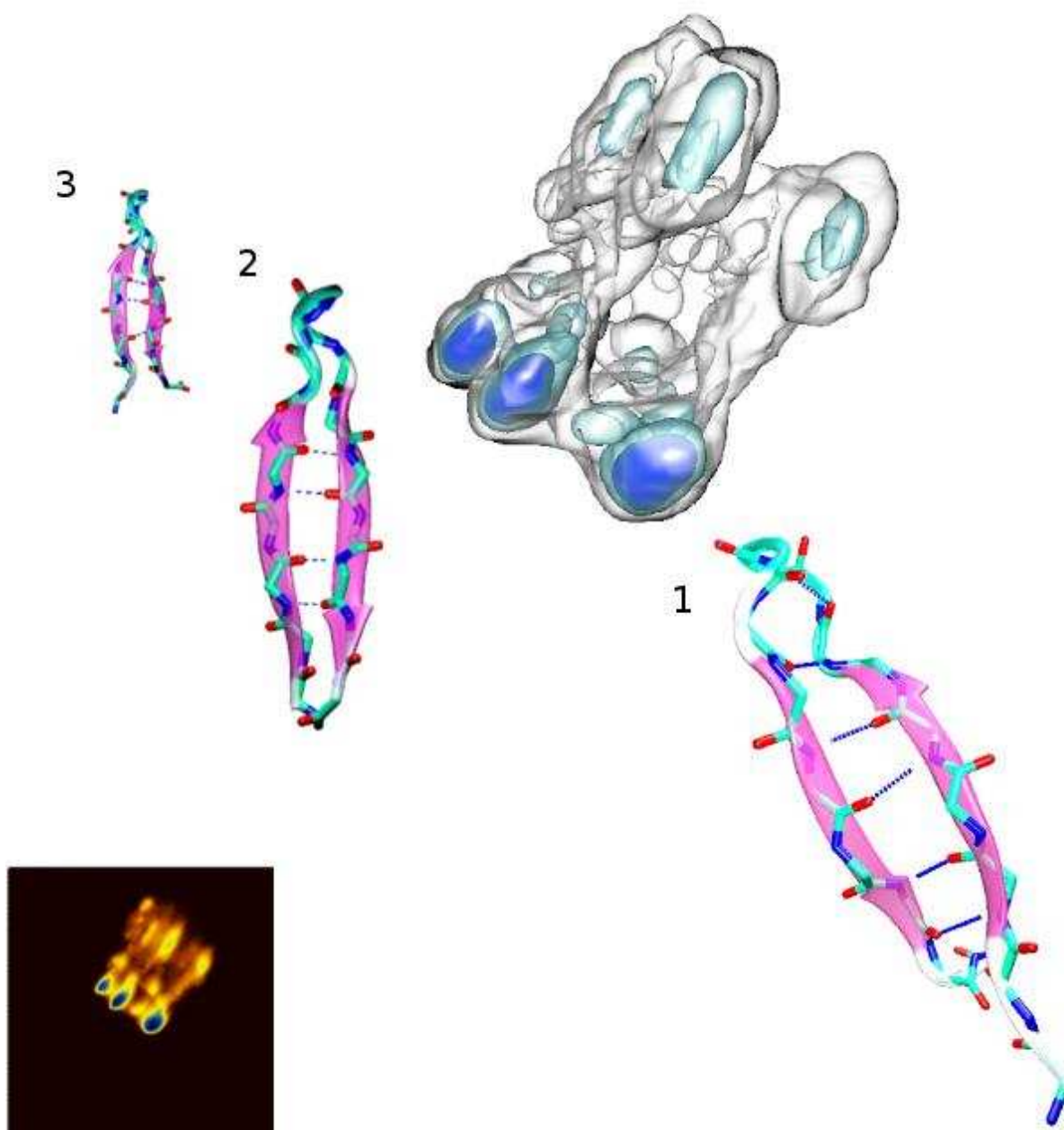
Στις εικόνες 4.32, 4.33, 4.34 παρουσιάζονται τα τρισδιάστατα ενεργειακά τοπία όπως προέκυψαν από την ανάλυση dihedral PCA για τις 32 δίεδρες γωνίες του πεπτιδίου. Το τρισδιάστατο γράφημα του ενεργειακού πεδίου παρουσιάζεται,

όπως και το δισδιάστατο ενεργειακό γράφημα που σχηματίζεται από τους *eigenvectors* 01 – 02, με τρία διαφορετικά επίπεδα ισοεπιφάνειας τα οποία προσδιορίζουν τα σημεία στα οποία σχηματίστηκαν τα πολυπληθέστερα clusters. Σε κάθε ενεργειακό τοπίο συμπεριλαμβάνονται και οι αντιπροσωπευτικές δομές από την κάθε “κορυφή” του γραφήματος (οι κορυφές αντιστοιχούν στα διαφορετικά clusters). Κάθε αντιπροσωπευτική δομή αντιστοιχεί στη δομή που έχει το χαμηλότερο RMSD από τη μέση δομή του κάθε cluster. Επιπλέον το μέγεθος των δομών στο γράφημα είναι ενδεικτικό για το μέγεθος του cluster. Η θέση των δομών επάνω στο γράφημα είναι τέτοια, έτσι ώστε να αποτελεί την προβολή των clusters στην επιφάνεια της εικόνας. Επειδή, όμως, ορισμένα ενεργειακά πεδία εμφανίζουν πολλές κορυφές και για να μην υπάρχει σύγχυση με την αντιστοιχία δομών και clusters κάθε δομή αριθμείται ανάλογα με τον αριθμό του cluster στο οποίο ανήκει.



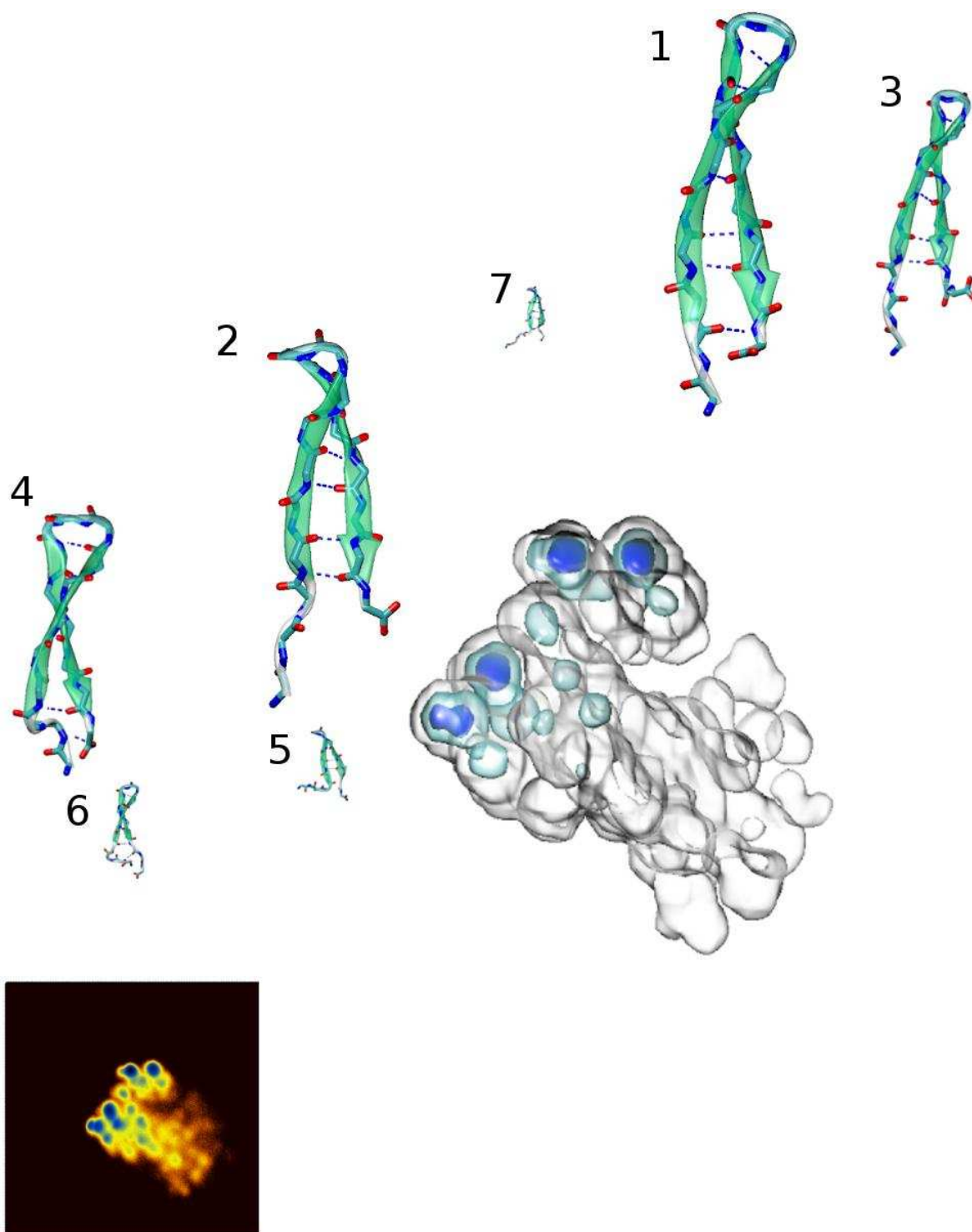
Εικόνα 4.32

*Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού 99SB. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το *dihedral PCA* για τις 32 διέδρες γωνίες του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού 99SB στους τρεις *principal components* της ανάλυσης *dihedral PCA*. Οι δομές είναι τοποθετημένες έτσι ώστε να φαίνεται η προβολή των τροχιακών στο επίπεδο ολόκληρης της εικόνας. Κάθε *cluster* συνδέεται, επιπλέον, αριθμητικά με την αντιπροσωπευτική δομή του (κοντινότερη στη μέση δομή του *cluster*), για να μην υπάρχει σύγχυση στην αντιστοιχία. Το μέγεθος των δομών είναι ενδεικτικό του μέγεθους του *cluster*.*



Εικόνα 4.33

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού ILDN. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το dihedral PCA για τις 32 διέδρες γωνίες του πεπτιδίου. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού ILDN στους τρεις principal components της ανάλυσης dihedral PCA.



Εικόνα 4.34

Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού protonated. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το dihedral PCA για τις 32 διέδρες γωνίες του πεπτιδίου. Στο τρισδιάστατο γράφημα

φαίνεται η προβολή του τροχιακού *protonated* στους τρεις *principal components* της ανάλυσης *dihedral PCA*. Οι δομές είναι τοποθετημένες έτσι ώστε να φαίνεται η προβολή των τροχιακών στο επίπεδο ολόκληρης της εικόνας.

Όπως έχει ήδη προαναφερθεί ο αριθμός των clusters για τα τρία τροχιακά είναι μεγαλύτερος σε σχέση με τις αναλύσεις Cartesian PCA. Αναλυτικότερα, στο τροχιακό 99SB βρέθηκαν 3 clusters με την ανάλυση cPCA για τα άτομα του πεπτιδικού σκελετού, ενώ με την ανάλυση dPCA βρέθηκαν 13, στο τροχιακό ILDN σχηματίστηκε ένα μόνο cluster με την ανάλυση cPCA για τα άτομα του πεπτιδικού σκελετού, ενώ με την ανάλυση dPCA σχηματίστηκαν τρία, τέλος στο τροχιακό *protonated* με την cPCA σχηματίστηκαν δύο clusters, ενώ με την dPCA επτά. Παρόλο που τα ποσοστά που αντιστοιχούν σε μερικά clusters είναι αμελητέα προς μελέτη, στις εικόνες των τρισδιάστατων ενεργειακών πεδίων συμπεριλαμβάνονται οι δομές από όλα τα clusters που βρέθηκαν, διότι η ενέργεια που αντιστοιχεί στις δομές αυτές είναι αρκετά χαμηλή έτσι ώστε να συμπεριλαμβάνονται στην μελέτη.

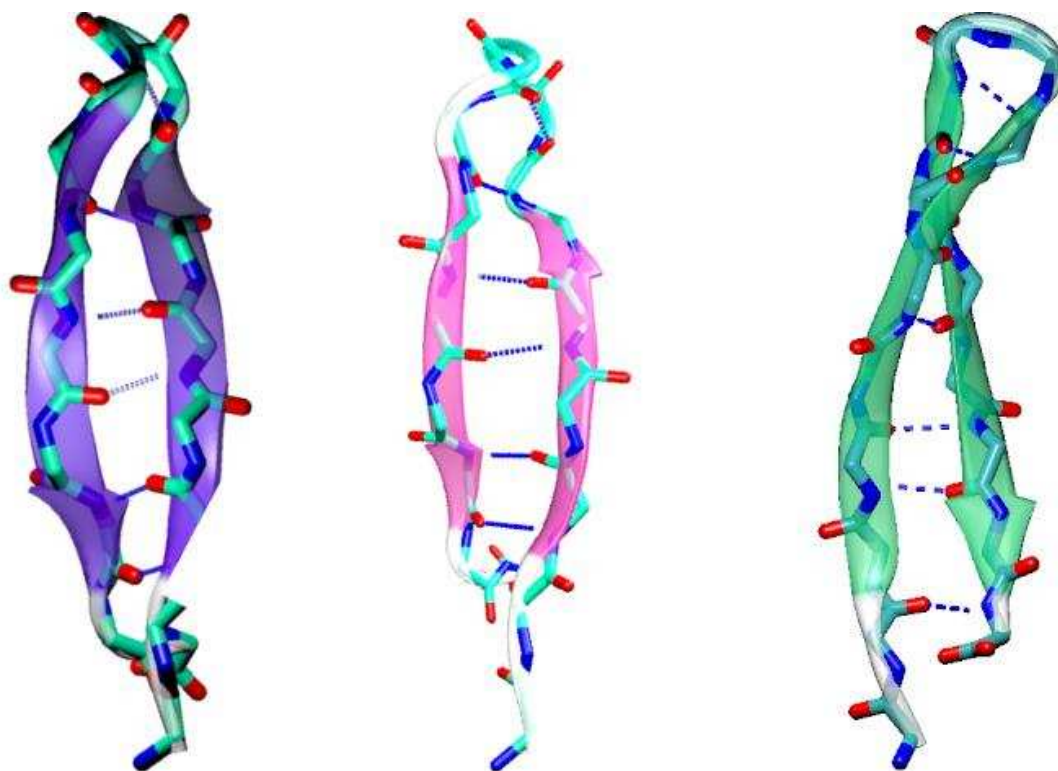
Από την ανάλυση dPCA για τις 32 δίεδρες γωνίες (ϕ και ψ) του πεπτιδίου προέκυψε μία κυρίαρχη δομή για κάθε τροχιακό. Ως κυρίαρχη δομή θεωρούμε τη δομή που έχει το χαμηλότερο RMSD σε σχέση με τη μέση δομή του πρώτου cluster. Σε όλα τα τροχιακά το πρώτο cluster αντιστοιχεί στο πολυπληθέστερο και ενεργειακά ελάχιστο cluster σε σχέση με όλα τα clusters που προέκυψαν. Στην εικόνα 4.35 παρουσιάζονται οι κυρίαρχες δομές από το κάθε τροχιακό.

Στην κυρίαρχη δομή του τροχιακού 99SB φαίνεται να έχει σχηματιστεί η β -φουρκέτα των καταλοίπων Ile3 – Lys6 και Lys11 – Leu15 και η στροφή μεταξύ των καταλοίπων Thr7 – Gly10. Τα δύο πρώτα και τα δύο τελευταία κατάλοιπα δεν φαίνεται να υιοθετούν κάποια διαμόρφωση, όπως έγινε και στις δομές που παρουσιάστηκαν από τις προηγούμενες αναλύσεις. Η στροφή φαίνεται να είναι Type I β -turn, καθώς η γλυκίνη βρίσκεται στην τρίτη θέση των καταλοίπων της

στροφής. Οι δεσμοί υδρογόνου σχηματίζονται μεταξύ των καταλοίπων Ile3(N) – Leu15(O), Ile3(O) – Leu15(N), Val5(N) – Ile13(O), Val5(O) – Ile13(N), Thr7(N) – Lys11(O) και Thr7(O) – Gly10(N). Οι παραπάνω δεσμοί υδρογόνου ταυτίζονται με τους δεσμούς υδρογόνου της πειραματικά προσδιορισμένης δομής, πλην του δεσμού υδρογόνου μεταξύ των καταλοίπων Met1(O) – Val17(N), καθώς και με την κυρίαρχη δομή του τροχιακού 99SB που εξήχθη από τις αναλύσεις RMSD και Cartesian PCA.

Στην κυρίαρχη δομή του τροχιακού ILDN η επιφάνεια της εκτεταμένης β-δομής εκτείνεται μεταξύ των καταλοίπων Ile3 – Lys6 και Thr12 – Leu15, ενώ τα κατάλοιπα της στροφής φαίνεται να είναι τα Thr7, Leu8, Asp9 και Gly10, με την Gly στη θέση i+3 της στροφής. Παρατηρώντας το δεσμό υδρογόνου που σχηματίζεται μεταξύ των καταλοίπων i(CO) – i+3(NH) και τον προσανατολισμό των καταλοίπων i+1 και i+2 συμπεραίνουμε πως πρόκειται για τον τύπο I β-στροφής και σε αυτή τη δομή. Οι δεσμοί υδρογόνου που σχηματίστηκαν κατά μήκος της β-φουρκέτας είναι οι ίδιοι με τους δεσμούς της κυρίαρχης δομής του τροχιακού 99SB. Επιπλέον και στις δύο περιπτώσεις οι επιφάνειες που σχηματίζονται εκατέρωθεν του πεπτιδικού σκελετού αποτελούνται από τις πλευρικές ομάδες των καταλοίπων Gln2, Phe4, Lys6, Thr12, Thr14 (για την μία επιφάνεια) και Ile3, Val5, Ile13 και Leu15 (για την άλλη επιφάνεια). Η διάταξη αυτή ταιριάζει με τις κυρίαρχες δομές των προηγούμενων αναλύσεων αλλά και με την πειραματικά προσδιορισμένη δομή.

Τέλος στο τροχιακό protonated η διαμόρφωση του εκτεταμένου β-φύλλου περιλαμβάνει τα κατάλοιπα Ile3 – Leu8 και Lys11 – Glu16, ενώ η στροφή αποτελείται από τα κατάλοιπα Leu8 – Lys11 με την γλυκίνη να βρίσκεται στη θέση i+2. Η τρίτη δομή φαίνεται να πλησιάζει περισσότερο στην δεύτερη εναλλακτική δομή που προτείνει η ομάδα της R. Zerella, η οποία παρουσιάζει μια μετάθεση των κλώνων κατά ένα κατάλοιπο.



Εικόνα 4.35

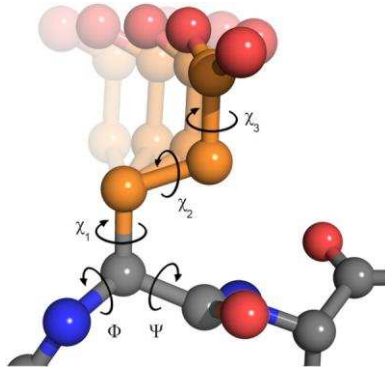
Συνοπτική παρουσίαση των κυρίαρχων δομών του κάθε τροχιακού, όπως προέκυψαν από την ανάλυση dPCA. Από αριστερά προς τα δεξιά έχουμε τις δομές που αντιστοιχούν στα τροχιακά 99SB, ILDN και protonated. Ο χρωματικός συνδυασμός είναι σύμφωνος με τις εικόνες των προηγούμενων αναλύσεων. Στις τρεις δομές φαίνεται ξεκάθαρα ο σχηματισμός της β-φουρκέτας. Η τρίτη δομή διαφέρει από τις δύο πρώτες ως προς το σχηματισμό των δεσμών υδρογόνου και τον τύπο της στροφής.

Προκειμένου να παρατηρήσουμε πόσο μεγάλη είναι η απόκλιση των δομών που σχηματίζουν τα παραγόμενα clusters πραγματοποιήθηκε τρισδιάστατη υπέρθεση δομών, όπως έγινε και μετά την ανάλυση cPCA, που απαρτίζουν το κάθε cluster. Η απόκλιση του πεπτιδικού άξονα (Ca, C, N, O) των δομών δεν ήταν μεγάλη. Οι περιοχές που εμφανίζουν μεγαλύτερη απόκλιση είναι τα άκρα του πεπτιδίου, γεγονός που μας παραπέμπει στην αυξημένη κινητικότητα των άκρων που παρατηρήθηκε από τις πρώτες κιάλας αναλύσεις.

Στην εικόνα 4.37 παρουσιάζονται οι τρισδιάστατες υπερθέσεις των δομών που απαρτίζουν τα πολυπληθέστερα clusters. Επειδή τα clusters που απορρέουν από την ανάλυση dPCA είναι αρκετά σε αριθμό κρίθηκε σκόπιμο να παρουσιασθούν μόνο αυτά τα οποία καταλαμβάνουν ένα σεβαστό ποσοστό του συνολικού τροχιακού. Συνεπώς παρουσιάζονται μόνο τα clusters που καταλαμβάνουν περισσότερο του 5% του συνολικού όγκου του τροχιακού. Αναλυτικότερα θα παρουσιασθούν τα δύο πρώτα από τα δεκατρία clusters που προέκυψαν για το τροχιακό 99SB, όλα τα clusters του τροχιακού ILDN και τα τέσσερα πρώτα clusters που προέκυψαν για το τροχιακό protonated.

Στον συγκεκριμένο τύπο υπέρθεσης αξιολογούμε την κινητικότητα του πεπτιδικού σκελετού σε μια ομάδα δομών. Οι δομές αυτές προέκυψαν από την dPCA ανάλυση, η οποία όπως έχει ήδη αναφερθεί λαμβάνει υπόψιν μόνο τις γωνίες ϕ και ψ .

Προκειμένου να προχωρήσουμε ένα βήμα παραπέρα, ενσωματώνουμε στους υπολογισμούς των δίεδρων γωνιών την γωνία χ_1 (CHI1 angle-dihedral PCA). Όπως είναι προφανές, με τον τρόπο αυτό μπορούμε να σχηματίσουμε ομάδες δομών (clusters) που να αποτελούνται από δομές όμοιες με βάση τα νέα κριτήρια που θέσαμε, δηλαδή τις δίεδρες γωνίες ϕ και ψ και την χ_1 . Η χ_1 γωνία υπόκειται σε ορισμένους περιορισμούς που προκύπτουν από την στερεοχημική παρεμπόδιση των ατόμων της πλευρικής αλυσίδας και της κύριας αλυσίδας (εικόνα 4.36). Οι διαφορετικές διαμορφώσεις της πλευρικής αλυσίδας ως συνάρτηση της χ_1 γωνίας αναφέρονται ως gauche(+), trans και gauche(-).

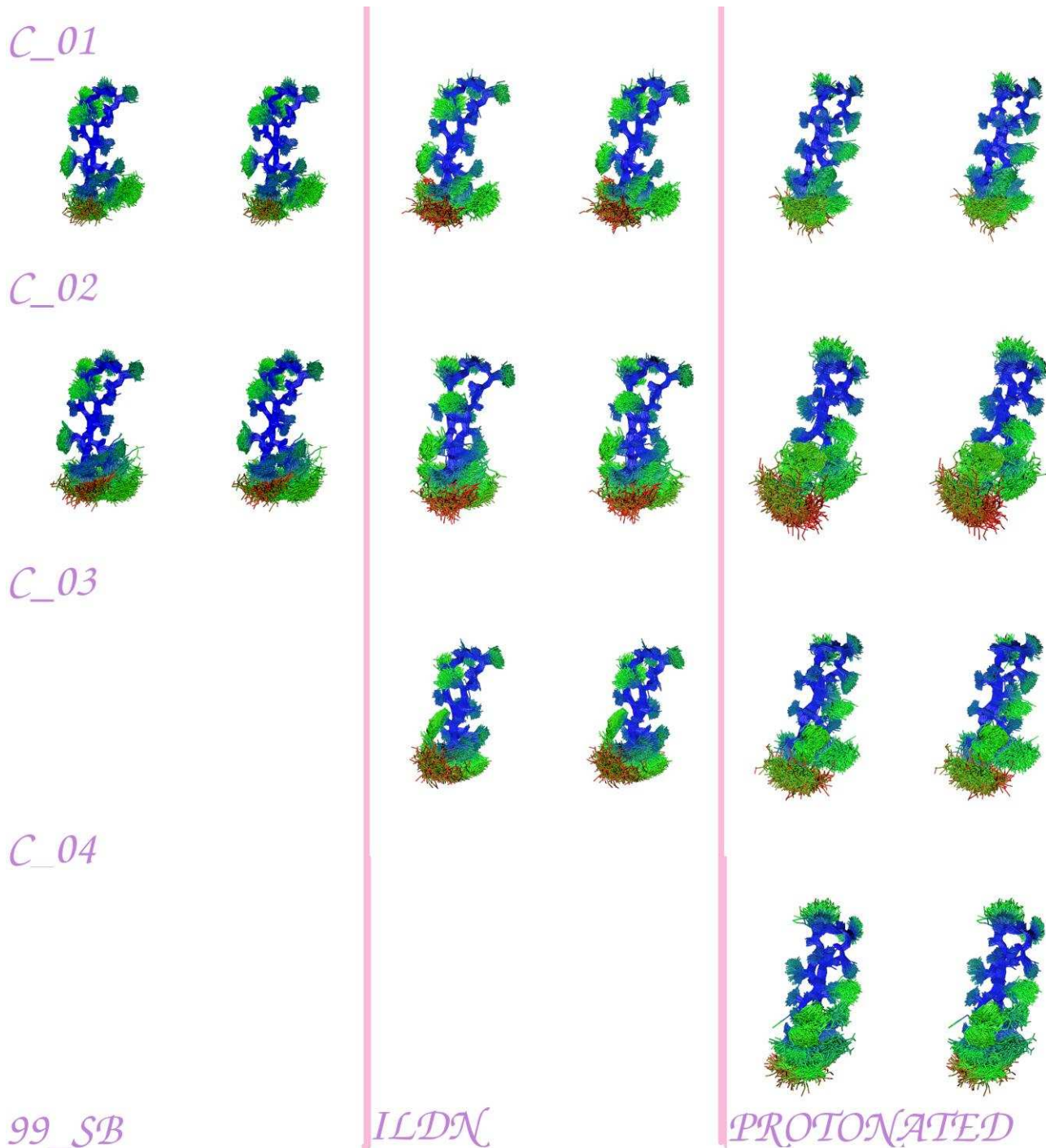


Εικόνα 4.36

Σχηματική αναπαράσταση (*ball and stick representation*) του γλουταμικού οξέος, το οποίο διαθέτει τρεις χ γωνίες στην πλευρική του ομάδα (πορτοκαλί). Πέραν των χ γωνιών, παρουσιάζονται οι ϕ και ψ γωνίες του σκελετού του αμινοξέως (γκρι). Οι χ γωνίες των αμινοξέων μπορεί να κυμαίνονται από καμία έως τέσσερις, στα 20 βασικά αμινοξέα. Οι δίεδρες γωνίες ενός πεπτιδίου είναι οι μόνοι βαθμοί ελευθερίας για τον πεπτιδικό σκελετό (ϕ και ψ) και τις πλευρικές ομάδες (χ_n). (Αναπαράγεται άνευ αδείας από Harder et al., 2010)

Αρχικά παρουσιάζονται τα αποτελέσματα από την ανάλυση dPCA με ενσωματωμένη την χ_1 γωνία, που αφορούν το ενεργειακό τοπίο. Στην εικόνα 4.38 παρουσιάζονται τα ενεργειακά τοπία όλων των τροχιακών, όπως παρήχθησαν από την ανάλυση dPCA συμπεριλαμβανομένης της γωνίας χ_1 .

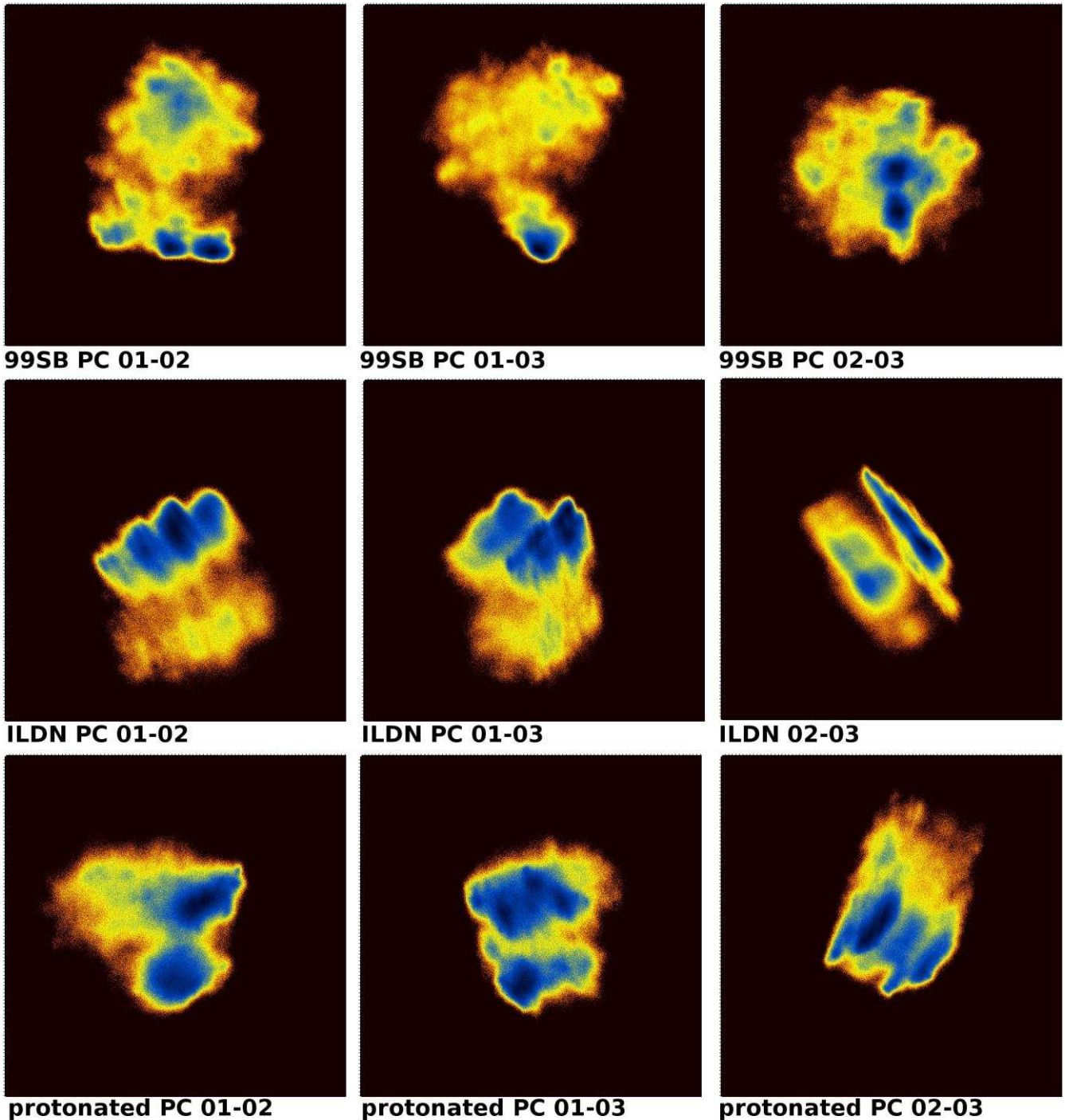
Συγκριτικά με την προηγούμενη ανάλυση dPCA, τα clusters έχουν μειωθεί σε αριθμό σε όλα τα τροχιακά. Μπορούμε να πούμε ότι αυτό δικαιολογείται από την άποψη ότι εφόσον έχουμε προσθέσει ένα επιπλέον κριτήριο στην ανάλυση (την γωνία χ_1) είναι δυσκολότερο να υπάρξουν πολλά clusters που να έχουν ικανοποιητικό πλήθος δομών έτσι ώστε να καλύψουν μεγάλο εύρος της διακύμανσης των τροχιακών. Επιπλέον, δεν πρέπει να ξεχνάμε ότι με την προσθήκη της γωνίας χ_1 αυτό που μελετάμε στην ουσία είναι η συμπεριφορά των πλευρικών ομάδων, οι οποίες παρουσιάζουν πολύ μεγαλύτερη διακύμανση-κινητικότητα από αυτή του πεπτιδικού σκελετού (Ca, C, N, O), η παρατήρηση αυτή θα γίνει και οπτικά αντιληπτή στην υπέρθεση δομών αργότερα.



Εικόνα 4.37

Παρουσίαση των αντιπροσωπευτικών clusters, σε stereo view, όπως προέκυψε από την ανάλυση dihedral PCA για τα τρία τροχιακά. Η υπέρθεση των δομών βασίζεται στον πεπτιδικό σκελετό. Η διάταξη του πεπτιδικού σκελετού δεν φαίνεται να παρουσιάζει μεγάλη διακύμανση. Αυξημένη κινητικότητα, σε σχέση με το υπόλοιπο πεπτιδικό τμήμα, παρουσιάζεται στα πεπτιδικά άκρα. Δεν παρουσιάζονται όλες οι ομάδες που παρήχθησαν από την ανάλυση. Ο χρωματισμός έγινε με βάση τις ατομικές διακυμάνσεις, η κλίμακα διατηρείται

σταθερή για τα clusters που ανήκουν στο ίδιο τροχιακό. Οι πλευρικές ομάδες συμπληρώθηκαν στην εικόνα προκειμένου να παρουσιαστεί περισσότερο ολοκληρωμένη η πεπτιδική δομή.



Εικόνα 4.38

Παρουσίαση των ενεργειακών τοπίων (ΔG , energy landscapes) στο επίπεδο των τριών κυρίαρχων συνιστωσών (Pcs) ανά ζεύγη (οριζοντίως) για τα τροχιακά 99SB,

ILDN και protonated (καθέτως). Τα ενεργειακά τοπία είναι αποτέλεσμα των αναλύσεων dPCA για τις 32 φ και ψ δίεδρες γωνίες του πεπτιδικού σκελετού και τις 16 χ1 γωνίες των πλευρικών ομάδων (συνολικά 48 γωνίες, που σχηματίζονται από 83 άτομα, συμπεριελήφθησαν στους υπολογισμούς της dPCA για κάθε προσομοίωση). Η κλίμακα βαθμολόγησης αντιστοιχεί στο χρωματικό φάσμα μεταξύ μπλε (ελάχιστη βαθμολόγηση) και ερυθρού (μέγιστη βαθμολόγηση).

Συγκρίνοντας τα ενεργειακά τοπία που παρήχθησαν από την dPCA που συμπεριλαμβάνει την χ1 γωνία με αυτά από την cPCA στην οποία μελετήθηκαν όλα τα βαρέα άτομα, παρατηρούμε ότι το ενεργειακό τοπίο σε όλα τα τροχιακά είναι πιο τραχύ από αυτό της cPCA και καλύπτεται μικρότερο ποσοστό του χάρτη διακύμανσης.

Στο τροχιακό της 99SB προσομοίωσης παρήχθησαν συνολικά 4 clusters που καλύπτουν το 85% της διακύμανσης. Το πρώτο cluster είναι μεγάλο, σε αντίθεση με τα υπόλοιπα τα οποία κατέχουν λιγότερο από το 5% του συνόλου του τροχιακού. Στο τροχιακό της ILDN προσομοίωσης παρήχθησαν συνολικά 2 clusters τα οποία καλύπτουν το 83% της διακύμανσης. Ενώ τέλος στο τροχιακό της προσομοίωσης protonated παρήχθησαν συνολικά 3 τροχιακά τα οποία καλύπτουν μόνο το 54% της διακύμανσης. Το μέγεθος των clusters είναι επίσης αρκετά μικρό σε σχέση με το μέγεθος του τροχιακού. Το ποσοστό που κατέχει κάθε cluster σε κάθε τροχιακό που παράχθηκε από την ανάλυση, αλλά και το ποσοστό της διακύμανσης που καλύπτεται από τα clusters κάθε τροχιακού παρουσιάζονται αναλυτικά στον πίνακα 4.6.

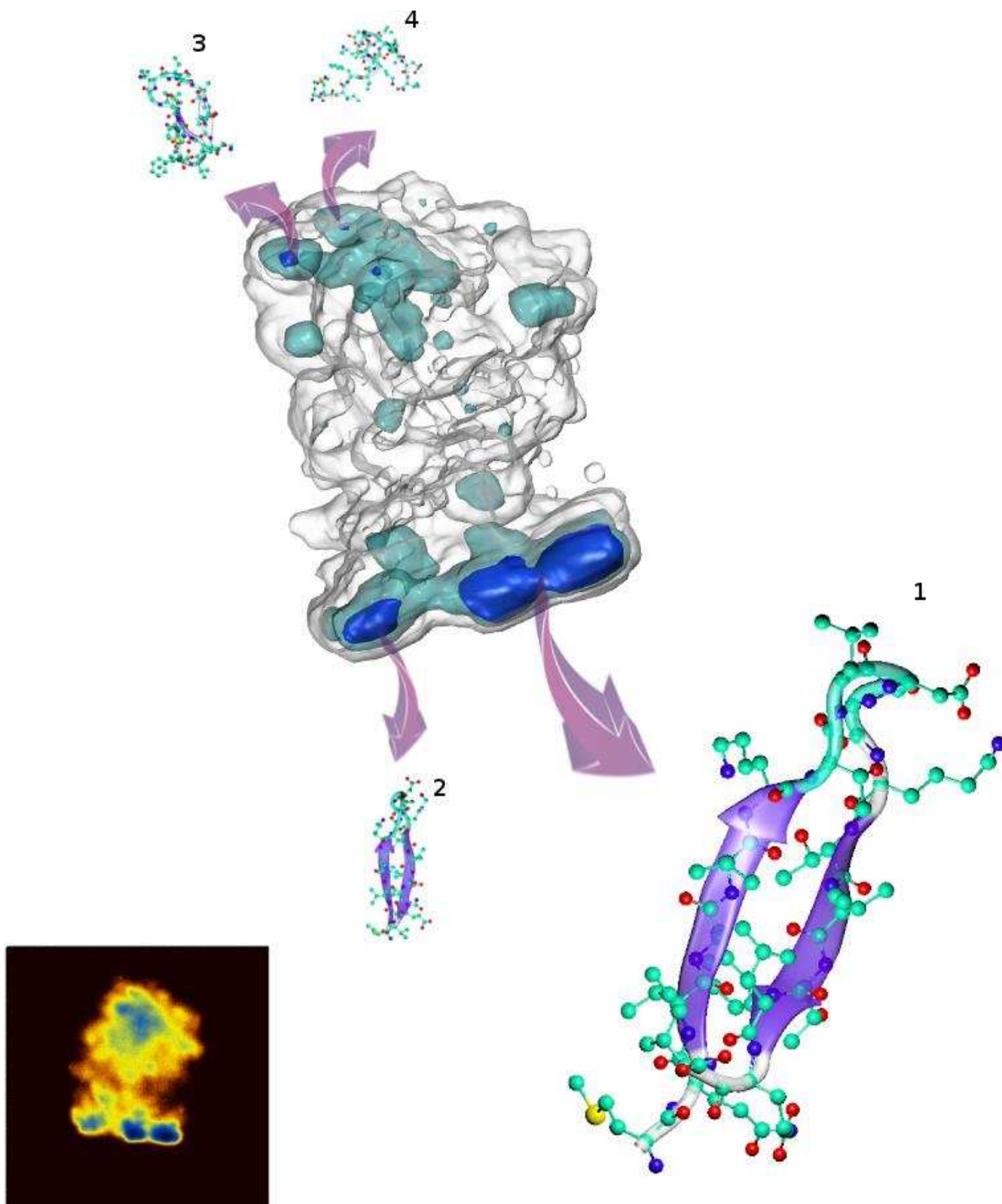
No. Cluster /1	2	3	4	Variance Explained	
trajectory					
99SB	23.6%	0.1%	0.02%	0.02%	85%
ILDN	35%	7.0%	-	-	83%
Protonated	8.9%	4.5%	4.4%	-	54%

Πίνακας 4.6

Συγκριτική παρουσίαση των clusters που σχηματίσθηκαν από την ανάλυση dPCA για τις γωνίες ϕ , ψ και χ_1 των τροχιακών 99SB, ILDN και protonated. Στις στήλες 2 – 5 αναγράφεται το ποσοστό που καταλαμβάνει το κάθε cluster σε σχέση με το συνολικό μέγεθος του τροχιακού, ενώ στην τελευταία στήλη αναγράφεται το ποσοστό της διακύμανσης που καλύπτεται από τα clusters.

Η διακύμανση που καλύπτεται από τα clusters που παρήχθησαν για τα τροχιακά 99SB και ILDN, είναι ικανοποιητική και αντιστοιχεί 85% και 83% για την κάθε μία. Δεν συμβαίνει όμως το ίδιο για την διακύμανση που καλύπτουν τα clusters που παρήχθησαν για το τροχιακό protonated. Στην τελευταία περίπτωση η τα clusters καλύπτουν μόνο το 54% της διακύμανσης. Εάν προσπαθήσουμε να συμπεριλάβουμε περισσότερα clusters στην ανάλυση θα διαπιστώσουμε ότι, ενώ μπορούμε να φτάσουμε μέχρι και το 80% περίπου της διακύμανσης τα clusters που χρησιμοποιούμε σε αυτή την περίπτωση έχουν πολύ μικρό αριθμό δομών και δεν είναι στατιστικά σημαντικά προς μελέτη.

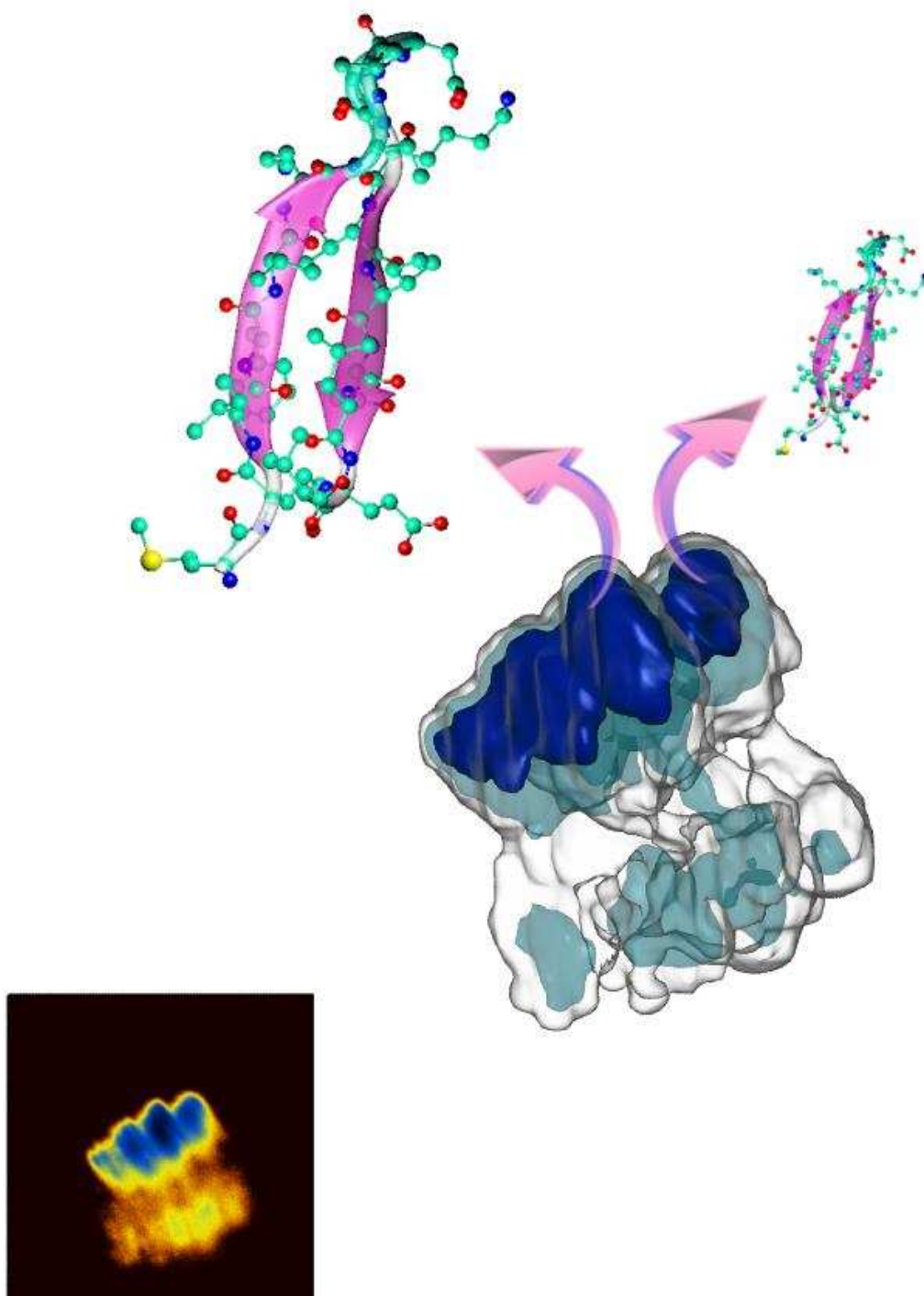
Εν συνεχεία θα παρουσιασθούν τα τρισδιάστατα ενεργειακά τοπία με ενσωματωμένες τις αντιπροσωπευτικές δομές από κάθε cluster (εικόνες 3.39, 3.40 και 3.41) και τέλος η υπέρθεση δομών των clusters που παρήχθησαν από την ανάλυση (εικόνα 3.42). Όπως και στην προηγούμενη ανάλυση dPCA, οι τρεις κύριες δομές από κάθε προσομοίωση θα παρουσιασθούν μαζί προκειμένου να είναι ευκολότερη η σύγκριση μεταξύ τους (εικόνα 3.43).



Εικόνα 4.39

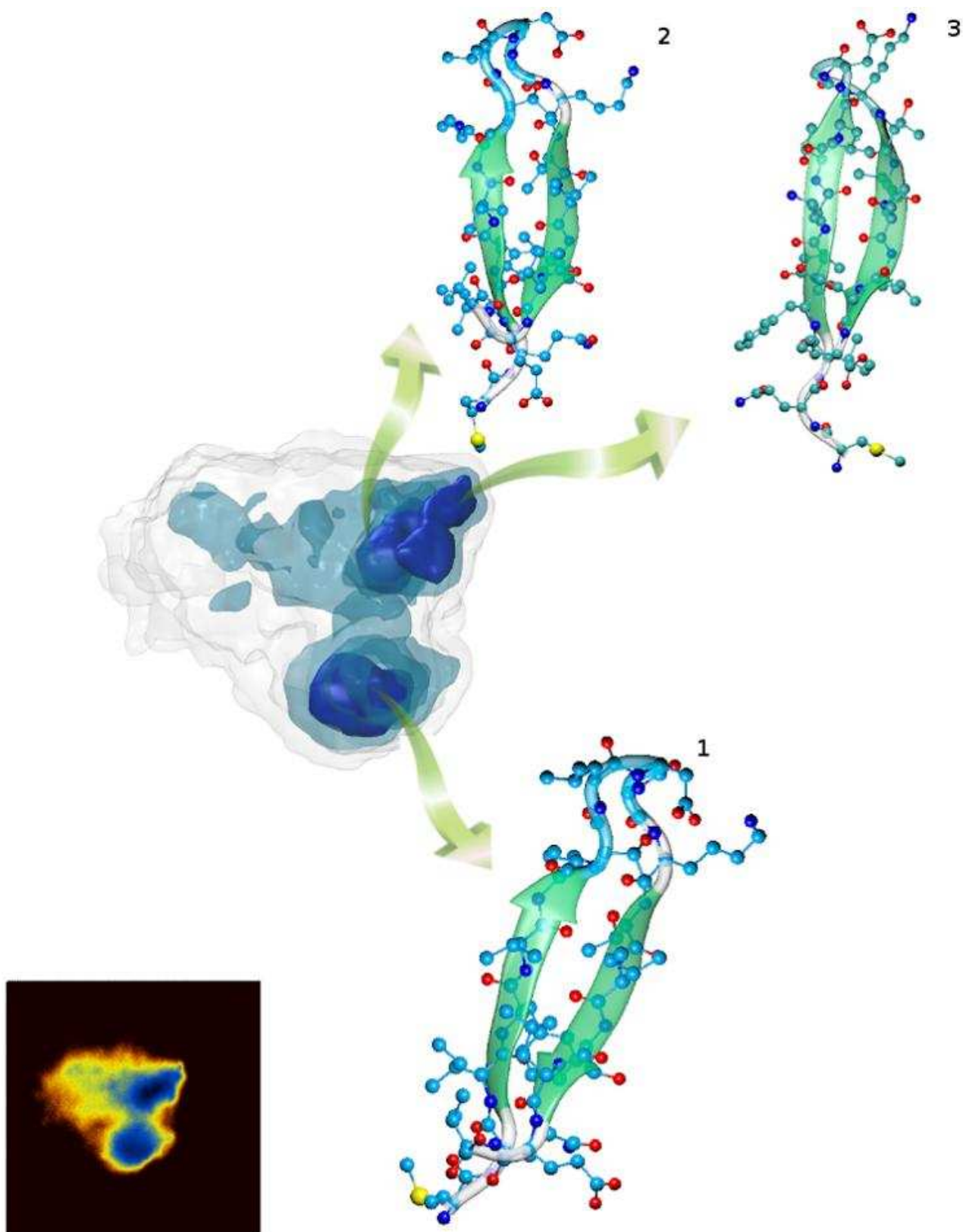
Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού 99SB. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το

dihedral PCA για τις 32 ϕ και ψ διεδρες γωνίες του πεπτιδίου και τις 16 χ_1 γωνίες. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού *ILDN* στους τρεις principal components της ανάλυσης *dihedral PCA*.



Εικόνα 4.40

*Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού ILDN. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το *dihedral PCA* για τις 32 ϕ και ψ δίεδρες γωνίες του πεπτιδίου και τις 16 χ_1 γωνίες. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού ILDN στους τρεις *principal components* της ανάλυσης *dihedral PCA*.*

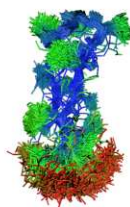
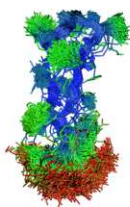


Εικόνα 4.41

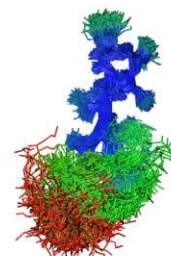
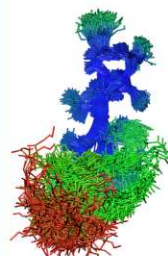
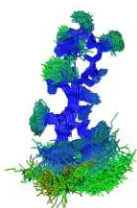
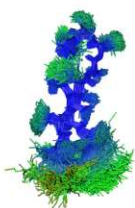
Τρισδιάστατο ενεργειακό τοπίο της αναδίπλωσης του τροχιακού protonated. Κάτω αριστερά φαίνεται η δισδιάστατη γραφική παράσταση της ελεύθερης ενέργειας στο επίπεδο των δύο πρώτων κύριων συνιστωσών όπως υπολογίσθηκε από το

dihedral PCA για τις 32 φ και ψ διέδρες γωνίες του πεπτιδίου και τις 16 χ1 γωνίες. Στο τρισδιάστατο γράφημα φαίνεται η προβολή του τροχιακού ILDN στους τρεις principal components της ανάλυσης dihedral PCA.

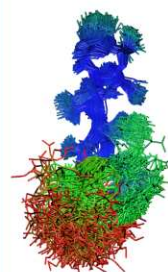
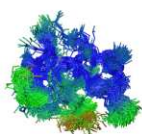
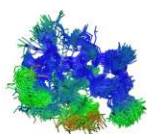
C_01



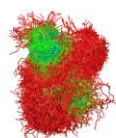
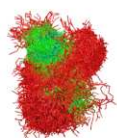
C_02



C_03



C_04



99SB

ILDN

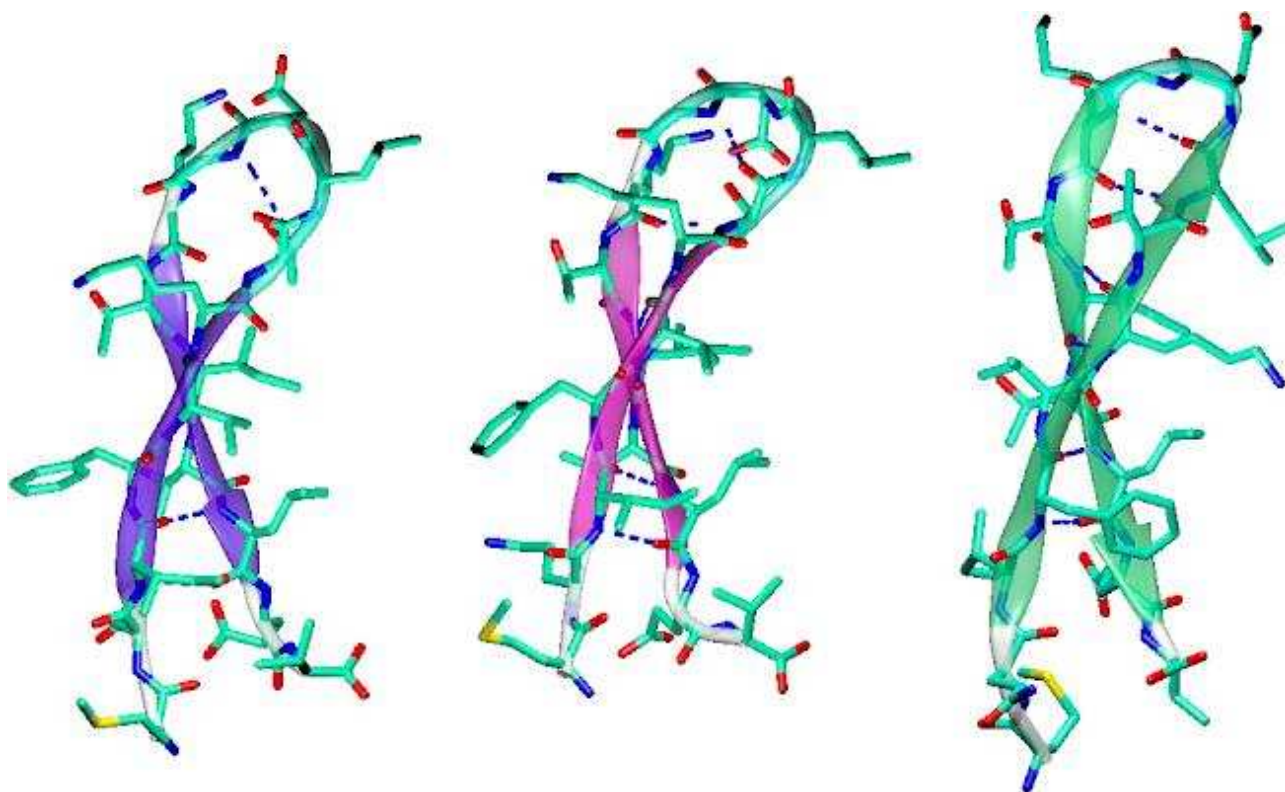
PROTONATED

Εικόνα 4.42

Υπέρθωση δομών των clusters, σε stereo view, όπως προέκυψε από την ανάλυση dihedral PCA, η οποία συμπεριλαμβάνει την γωνία χ_1 , για τα τρία τροχιακά. Αυξημένη κινητικότητα, σε σχέση με το υπόλοιπο πεπτιδικό τμήμα, παρουσιάζεται στα άκρα. Ο χρωματισμός έγινε με βάση τις ατομικές διακυμάνσεις, η κλίμακα διατηρείται σταθερή για τα clusters που ανήκουν στο ίδιο τροχιακό.

Στην παραπάνω εικόνα (εικ. 4.42) παρουσιάζονται όλα τα clusters των τροχιακών 99SB, ILDN και protonated, όπως παρήχθησαν από την dPCA που περιλαμβάνει την γωνία χ_1 . Το κυρίαρχο cluster σε κάθε περίπτωση είναι το πρώτο. Από το κυρίαρχο cluster προέρχεται και η κυρίαρχη δομή του εκάστοτε τροχιακού. Για το τροχιακό 99SB μπορούμε να πούμε πως δεν παρουσιάζει μεγάλη διακύμανση στην υπέρθεση, καθώς το μεγαλύτερο τμήμα είναι χρωματισμένο μπλε και πράσινο. Όπως είναι αναμενόμενο το πράσινο χρώμα βρίσκεται στις πλευρικές ομάδες, πράγμα που σημαίνει ότι παρουσιάζουν κάποια μικρή κινητικότητα εντός του cluster. Στο τροχιακό ILDN και τα δύο cluster που παρήχθησαν καταλαμβάνουν ποσοστό μεγαλύτερο του 5% του τροχιακού, όπως έχει ήδη αναφερθεί. Ωστόσο το πρώτο cluster είναι που μεγαλύτερο από το δεύτερο και η αντιπροσωπευτική δομή του τροχιακού ILDN για την ανάλυση αυτή προέρχεται από αυτό το cluster. Στο cluster αυτό παρουσιάζεται μικρή κινητικότητα στις πλευρικές ομάδες του πεπτιδίου και μεγάλη κινητικότητα στις πλευρικές ομάδες των καταλοίπων των άκρων. Ωστόσο, στο δεύτερο cluster του τροχιακού ILDN η κινητικότητα των πλευρικών ομάδων φαίνεται να είναι μικρότερη. Τέλος στο τροχιακό protonated φαίνεται να έχουμε το ίδιο πρότυπο κινητικότητας, με τα άκρα του πεπτιδίου να έχουν την πιο έντονη κίνηση, ενώ το υπόλοιπο τμήμα του πεπτιδίου παρουσιάζει μικρότερη κίνηση. Συνολικά τα κατάλοιπα της στροφής αλλά και τα κατάλοιπα των β -κλώνων που σχηματίζουν τους υδρόφοβους πυρήνες εκατέρωθεν του πεπτιδικού σκελετού δεν παρουσιάζουν αυξημένη κινητικότητα. Η παρατήρηση αυτή είναι σημαντική γιατί υποστηρίζει τη συμβολή των επαφών που δημιουργούν οι πλευρικές

ομάδες στη σταθερότητα της δομής του πεπτιδίου. Εφόσον οι επαφές που δημιουργούν οι πλευρικές ομάδες μεταξύ τους είναι ισχυρές συμβάλουν στη σταθεροποίηση της δομής, με αποτέλεσμα να αναμένουμε η κινητικότητα αυτών των πλευρικών ομάδων να είναι χαμηλή.

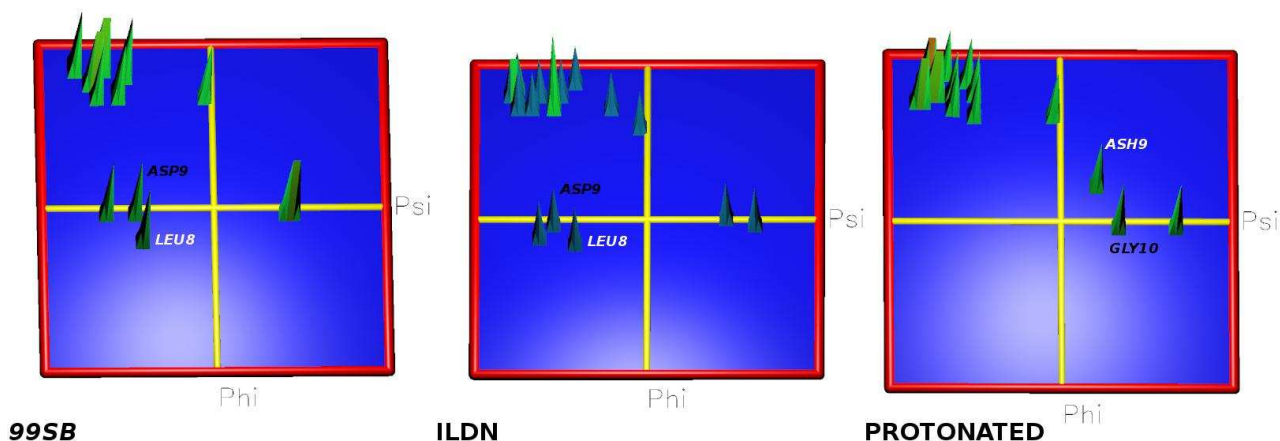


Εικόνα 4.43

Συνοπτική παρουσίαση των κυρίαρχων δομών του κάθε τροχιακού, όπως προέκυψαν από την ανάλυση dPCA που συμπεριλαμβάνει την γωνία χ_1 . Από αριστερά προς τα δεξιά έχουμε τις αντιπροσωπευτικές δομές των τροχιακών 99SB, ILDN και protonated. Στις τρεις δομές σχηματίζεται β -φουρκέτα. Οι δύο πρώτες δομές είναι σύμφωνες με την πειραματικά προσδιορισμένη δομή, ενώ η τρίτη δομή διαφέρει ως προς το μοτίβο των δεσμών υδρογόνου, τον τύπο της στροφής και την διεύθυνση των πλευρικών ομάδων. Ωστόσο η τρίτη δομή αντιστοιχεί στην εναλλακτική δομή που προτείνουν οι R. Zerella et al.. Ο χρωματικός συνδυασμός είναι σε συμφωνία με τις εικόνες των προηγούμενων αναλύσεων.

Οι αντιπροσωπευτικές δομές των τροχιακών 99SB, ILDN και protonated, που προέκυψαν από την ανάλυση dPCA (με την χ_1) και dPCA (χωρίς την χ_1) συμφωνούν μεταξύ τους. Η δομή της β -φουρκέτας σχηματίσθηκε και στις τρεις περιπτώσεις. Στις δύο πρώτες δομές (99SB και ILDN) σχηματίσθηκαν όλοι οι δεσμοί υδρογόνου που περιέχει η πειραματικά προσδιορισμένη δομή, εκτός από το δεσμό, στα άκρα του πεπτιδίου, μεταξύ MET1(O) – VAL17(N). Στην πρώτη δομή δεν φαίνεται να σχηματίζεται και ο δεσμός μεταξύ των καταλοίπων THR7(N) – LYS11(O). στην τρίτη δομή το πρότυπο των δεσμών υδρογόνου του πεπτιδικού σκελετού φαίνεται να ακολουθεί αυτό της εναλλακτικής δομής, που προτείνουν οι R. Zerella et al., αλλά και σε αυτή την περίπτωση δεν σχηματίζεται ο δεσμός υδρογόνου στα άκρα του πεπτιδίου.

Η στροφή που σχηματίζεται στις δύο πρώτες δομές, επίσης, συμφωνεί με την στροφή της πειραματικά προσδιορισμένης δομής και είναι τύπου I β -στροφή. Η στροφή σχηματίζεται μεταξύ των καταλοίπων 7-11. Ο δεσμός υδρογόνου σχηματίζεται μεταξύ των καταλοίπων THR7(O) – GLY10(N) και η γλυκίνη βρίσκεται στην θέση $i+3$, επιπλέον οι ϕ και ψ γωνίες των καταλοίπων $i+1$ και $i+2$ καταδεικνύουν ότι πρόκειται για τον τύπο I β -στροφής. Στην τρίτη δομή ο δεσμός υδρογόνου σχηματίζεται μεταξύ των καταλοίπων LEU8(O) – LYS11(N), οπότε η γλυκίνη βρίσκεται στην θέση $i+2$, ενώ στην θέση $i+1$ βρίσκεται το ASH9. Οι γωνίες ϕ και ψ των καταλοίπων $i+1$ και $i+2$ καταδεικνύουν ότι πρόκειται για τον τύπο I' β -στροφής. Στην εικόνα 4.44 παρουσιάζονται τα διαγράμματα Ramachandran για τις γωνίες ϕ και ψ της κάθε δομής.



Εικόνα 4.44

Διάγραμμα Ramachandran για τις φ και ψ γωνίες των καταλοίπων κάθε αντιπροσωπευτικής δομής. Σημειώνονται τα κατάλοιπα $i+1$ (λευκό χρώμα) και $i+2$ (μαύρο χρώμα) σε κάθε περίπτωση.

Τέλος παρατηρούμε ότι αλλάζει και το πρότυπο του υδρόφοβου πυρήνα γύρω από τον πεπτιδικό σκελετό στην τρίτη δομή. Στις δύο πρώτες δομές το πακετάρισμα των πλευρικών αλυσίδων γίνεται όπως και στην πειραματικά προσδιορισμένη δομή, δηλαδή στην μία πλευρά σχηματίζουν επαφές τα κατάλοιπα Thr14, Phe4, Gln2 και Lys6, Thr12, ενώ στην άλλη πλευρά βρίσκονται σε κοντινή απόσταση τα κατάλοιπα Ile3, Val5, Ile13 και Leu15. Επιπλέον στις δύο πρώτες δομές το κατάλοιπο Asp9 φαίνεται να βρίσκεται πού κοντά στο κατάλοιπο Lys11. Ο σχηματισμός δεσμών υδρογόνου μεταξύ αυτών των πλευρικών ομάδων πιθανών να αυξάνει τη σταθερότητα της αναδίπλωσης, η επαφή αυτή δεν φαίνεται να υπάρχει στην τρίτη δομή, καθώς τα κατάλοιπα αυτά είναι απομακρυσμένα.

Συνοψίζοντας, στην παράγραφο 4.3 είδαμε τις αναλύσεις που βασίζονται στην ομαδοποίηση των δεδομένων και συγκεκριμένα τις αναλύσεις cPCA και dPCA. Οι διαφορετικοί τρόποι με τους οποίους ομαδοποιήσαμε τα τροχιακά κατέληξαν

σε όμοια και συγκρίσιμα αποτελέσματα. Οι δομές οι οποίες εξήχθησαν από τις αναλύσεις αυτές είναι όμοιες για κάθε τροχιακό και μπορούν να συγκριθούν δομικά και με τις δομές που παρουσιάζονται στις αναλύσεις που βασίζονται στο RMSD (εικόνα 4.16).

Καταλήγοντας σε κάποιες δομές, είναι σκόπιμο να συγκρίνουμε την πειραματικά προσδιορισμένη δομή με τις δομές που περιέχονται στα τροχιακά 99SB, ILDN και protonated, προκειμένου να εξετάσουμε ομοιότητες και διαφορές. Στην επόμενη παράγραφο θα παρουσιασθεί η σύγκριση της πειραματικά προσδιορισμένης δομής με τα τρία τροχιακά μέσω της ανάλυσης RMSD.

4.4 Σύγκριση RMSD πειραματικά προσδιορισμένης δομής

Στην παρούσα ενότητα του τετάρτου κεφαλαίου θα παρουσιαστούν οι αναλυτικοί πίνακες με το RMSD των 27 κατατεθειμένων δομών στην pdb (protein data bank) σε σχέση με τις δομές που εμπεριέχονται σε κάθε τροχιακό, σε διαφορετικές περιπτώσεις ατόμων και καταλοίπων. Με τον τρόπο αυτό θα μπορέσουμε να ελέγξουμε ποια (ή ποιες) από τις 27 κατατεθειμένες δομές πλησιάζει περισσότερο με τις δομές που παρήχθησαν κατά τις προσομοιώσεις μοριακής δυναμικής, στις διαφορετικές αυτές περιπτώσεις που θα παρουσιασθούν.

Δημιουργήθηκαν δύο πίνακες για τον συγκεκριμένο σκοπό, επειδή όπως έχουμε προαναφέρει αρκετές φορές τα πεπτιδικά άκρα παρουσιάζουν μεγάλη κινητικότητα κατά τις προσομοιώσεις και επηρεάζουν τις τιμές διαφόρων παραμέτρων, όπως το RMSD. Για το λόγο αυτό, ο ένας πίνακας (πίνακας 4.7) παρουσιάζει την χαμηλότερη τιμή RMSD για τα κατάλοιπα 1-17 κάθε μίας από τις 27 κατατεθειμένες δομές σε σχέση με τις δομές των τροχιακών, υπό διάφορες παραμέτρους. Ο δεύτερος πίνακας (πίνακας 4.8) παρουσιάζει την χαμηλότερη τιμή RMSD για τα κατάλοιπα 3-15 κάθε μίας από τις 27 κατατεθειμένες δομές σε σχέση με τις δομές των τροχιακών, υπό τις ίδιες παραμέτρους.

Οι διαφορετικές περιπτώσεις που εξετάστηκαν στους δύο πίνακες αφορούν την επιλογή των ατόμων στα οποία βασίστηκε η μέτρηση RMSD. Στην πρώτη περίπτωση εξετάζεται το RMSD χρησιμοποιώντας τα άτομα Ca για τις καταχωρημένες δομές και για τα τρία τροχιακά, στην δεύτερη περίπτωση εξετάζονται όλα τα άτομα του πεπτιδικού σκελετού και στην τρίτη περίπτωση εξετάζονται όλα τα βαρέα άτομα του πεπτιδίου (αφαιρούνται, δηλαδή, τα άτομα υδρογόνου).

Με μία πρώτη ματιά στον πίνακα 4.7 παρατηρούμε ότι υπάρχουν δύο

κατηγορίες τιμών RMSD (με εξαίρεση μία περίπτωση στο τροχιακό protonated). Θεωρητικά μπορούμε να ονομάσουμε δύο δομές όμοιες όταν έχουν RMSD χαμηλότερο των 2Å. Με τον τρόπο αυτό μπορούμε να κατατάξουμε δύο κατηγορίες τιμών στον πίνακα, η πρώτη είναι αυτή στην οποία οι τιμές RMSD είναι χαμηλότερες των 2Å (πράσινο χρώμα στον πίνακα), και συνεπώς αναφέρονται σε όμοιες δομές, και η δεύτερη είναι αυτή στην οποία οι τιμές RMSD είναι μεγαλύτερες των 2Å (κόκκινο χρώμα στον πίνακα), και συνεπώς αναφέρονται σε λιγότερο όμοιες δομές. Κοιτάζοντας αναλυτικότερα τον πίνακα μπορούμε να διαπιστώσουμε ότι δεν υπάρχει μία δομή η οποία να έχει σε όλες τις περιπτώσεις τη χαμηλότερη τιμή RMSD, αλλά υπάρχει ένα μοτίβο χαμηλών τιμών, ειδικά για τα τροχιακά 99SB και ILDN τα οποία περιέχουν ακριβών τον ίδιο αριθμό ατόμων. Τα τροχιακά αυτά έχουν χαμηλότερες τιμές RMSD σε σχέση με το τροχιακό protonated.

Στον πίνακα 4.8 παρουσιάζονται δύο κατηγορίες τιμών, επίσης, και δεν φαίνεται να υπάρχει μία δομή που να έχει σε κάθε περίπτωση το ελάχιστο RMSD. Οι τιμές είναι χαμηλότερες σε σχέση με τις τιμές του πρώτου πίνακα, γεγονός που υποστηρίζει την την αυξημένη κινητικότητα των άκρων. Επιπλέον, όπως και στον προηγούμενο πίνακα έτσι και στον πίνακα 4.8 οι μεγαλύτερες τιμές παρουσιάζονται στο τροχιακό protonated. Η διαφορά στην τιμή είναι της τάξης του ενός Å περίπου.

Amino acids : 1-17

	99SB			ILDN			PROTONATED-ILDN		
	Ca	backbone	Not hydrogen	Ca	backbone	Not hydrogen	Ca	backbone	Not hydrogen
NMR model	RMSD (A)								
1	0.78	0.98	3.33	0.64	0.81	3.30	1.75	1.93	4.15
2	1.01	1.19	3.40	0.85	1.03	3.30	1.69	1.79	4.10
3	1.09	1.34	3.40	0.84	1.11	3.28	1.76	1.91	4.18
4	0.63	0.84	3.27	0.60	0.67	3.24	1.68	1.83	4.11
5	1.04	1.30	3.38	0.69	0.99	3.26	1.74	1.95	4.25
6	0.72	0.93	3.27	0.60	0.77	3.25	1.92	2.11	4.18
7	1.02	1.25	3.36	0.85	1.09	3.27	1.74	1.84	4.13
8	0.75	0.93	3.32	0.58	0.73	3.30	1.94	2.12	4.20
9	0.73	0.97	3.29	0.63	0.78	3.26	1.97	2.16	4.23
10	0.62	0.93	3.29	0.56	0.71	3.24	1.89	2.02	4.27
11	1.09	1.30	3.44	0.61	1.06	3.33	1.91	2.16	4.36
12	0.82	1.02	3.29	0.59	0.82	3.26	1.80	1.96	4.31
13	0.91	1.09	3.35	0.63	0.94	3.30	1.99	2.18	4.35
14	0.68	0.84	3.30	0.63	0.71	3.27	1.66	1.81	4.11
15	0.80	0.97	3.31	0.62	0.79	3.29	1.95	2.14	4.23
16	0.67	0.95	3.31	0.58	0.75	3.28	2.01	2.19	4.32
17	0.76	1.01	3.30	0.65	0.88	3.30	1.66	1.83	4.07
18	0.66	0.84	3.30	0.63	0.71	3.27	1.66	1.81	4.11
19	0.71	1.12	3.29	0.60	0.83	3.21	1.81	2.02	4.27

20	0.64	0.97	3.24	0.58	0.77	3.22	1.71	1.91	4.12
21	1.09	1.16	3.36	0.93	0.98	3.27	1.70	1.79	4.07
22	0.76	0.97	3.29	0.57	0.73	3.26	1.91	2.09	4.18
23	0.65	0.87	3.27	0.63	0.74	3.26	1.72	1.89	4.09
24	1.03	1.22	3.39	0.85	1.08	3.29	1.89	1.99	4.21
25	0.64	0.87	3.29	0.58	0.69	3.26	1.62	1.79	4.11
26	0.74	0.92	3.33	0.62	0.72	3.28	1.99	2.11	4.26
27	0.64	0.93	3.30	0.55	0.73	3.29	1.74	1.91	4.16

Πίνακας 4.7

Amino acids : 3-15									
	99SB			ILDN			PROTONATED-ILDN		
	Ca	backbone	Not hydrogen	Ca	backbone	Not hydrogen	Ca	backbone	Not hydrogen
NMR model	RMSD (A)								
1	0.54	0.62	3.24	0.50	0.58	3.16	1.55	1.67	3.94
2	0.77	0.97	3.27	0.75	0.94	3.25	1.53	1.68	3.89
3	0.80	0.98	3.20	0.69	0.87	3.17	1.59	1.74	3.93
4	0.50	0.60	3.16	0.44	0.57	3.10	1.46	1.58	3.87
5	0.52	0.64	3.12	0.40	0.53	3.15	1.60	1.84	4.02
6	0.50	0.58	3.15	0.41	0.49	3.14	1.75	1.86	3.97
7	0.76	0.96	3.28	0.75	0.93	3.25	1.55	1.70	3.90
8	0.47	0.56	3.22	0.40	0.49	3.19	1.78	1.89	4.01

9	0.48	0.59	3.20	0.41	0.49	3.19	1.81	1.93	4.02
10	0.41	0.53	3.20	0.36	0.44	3.15	1.72	1.83	3.99
11	0.46	0.56	3.17	0.37	0.49	3.20	1.88	2.06	4.11
12	0.46	0.60	3.12	0.37	0.48	3.11	1.68	1.85	4.00
13	0.50	0.57	3.18	0.39	0.50	3.18	1.91	2.06	4.11
14	0.51	0.61	3.18	0.45	0.58	3.12	1.43	1.54	3.87
15	0.47	0.56	3.21	0.39	0.49	3.19	1.79	1.90	4.02
16	0.47	0.58	3.18	0.43	0.47	3.18	1.83	1.93	4.12
17	0.50	0.60	3.20	0.46	0.59	3.16	1.45	1.57	3.82
18	0.51	0.61	3.18	0.45	0.58	3.12	1.43	1.54	3.87
19	0.42	0.55	3.17	0.35	0.44	3.13	1.70	1.81	3.95
20	0.49	0.59	3.15	0.45	0.55	3.10	1.51	1.61	3.89
21	0.81	0.97	3.21	0.85	0.97	3.19	1.53	1.62	3.87
22	0.48	0.60	3.19	0.43	0.49	3.17	1.72	1.82	4.00
23	0.49	0.56	3.18	0.46	0.54	3.11	1.53	1.61	3.86
24	0.76	0.91	3.19	0.74	0.90	3.19	1.72	1.80	4.01
25	0.49	0.60	3.21	0.42	0.56	3.14	1.44	1.58	3.83
26	0.46	0.52	3.20	0.36	0.49	3.18	1.84	1.94	4.05
27	0.49	0.57	3.23	0.41	0.51	3.15	1.58	1.68	3.92

Πίνακας 4.8

Και στους δύο πίνακες οι χαμηλότερες τιμές RMSD, εντός των τροχιακών, παρατηρούνται για τα άτομα του πεπτιδικού σκελετού και για το άτομο Ca, όπως αναμένεται. Αντιθέτως η σύγκριση ολόκληρου του πεπτιδίου (πλην των ατόμων υδρογόνου) σε όλες τις περιπτώσεις παρουσιάζει τιμή RMSD

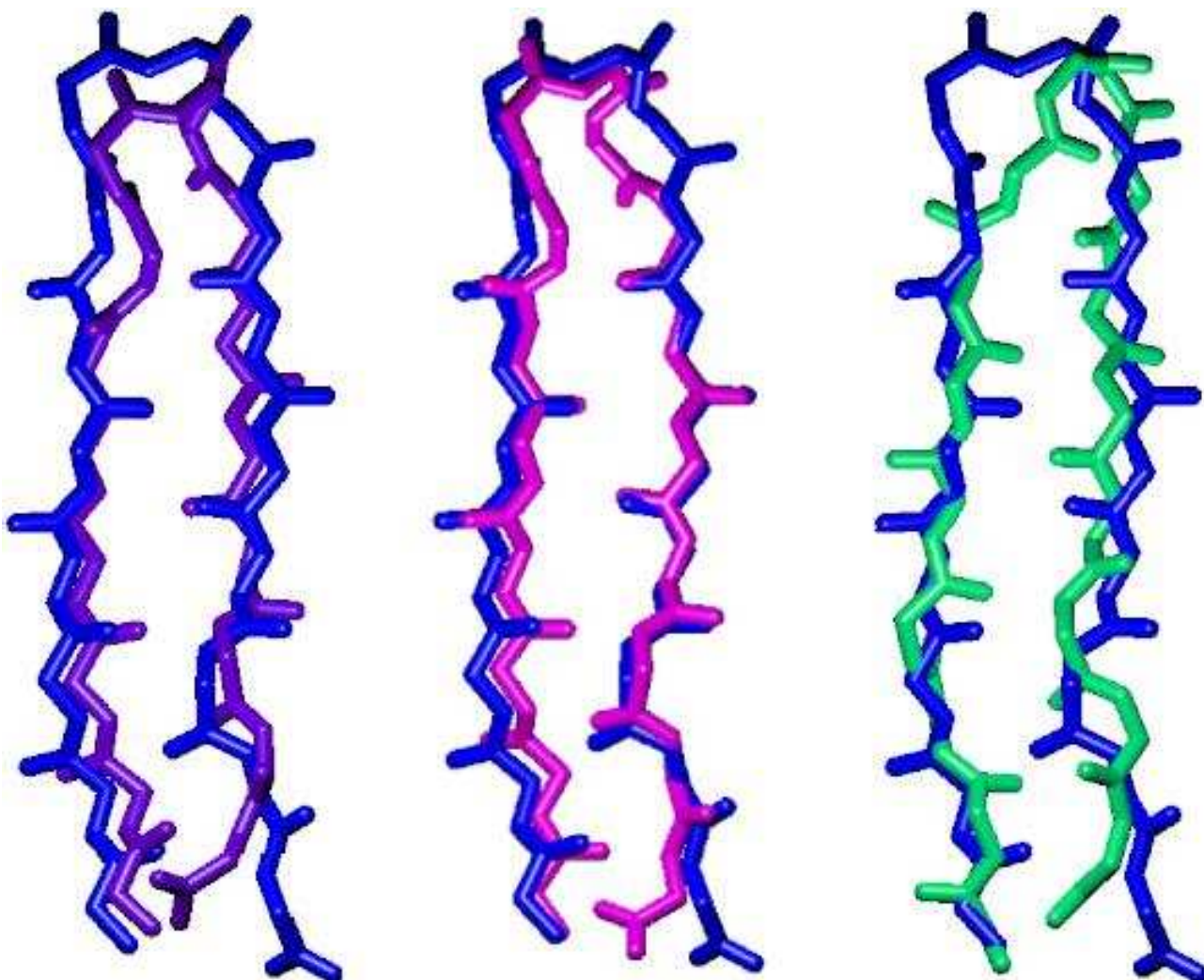
μεγαλύτερη των 3Å. Ως τελευταία παρατήρηση για την σύγκριση των RMSD των πειραματικά προσδιορισμένων δομών και των δομών που εμπεριέχονται στα τροχιακά που παρήχθησαν από τις προσομοιώσεις, αναφέρεται πως την χαμηλότερη τιμή RMSD κατέχει η κατατεθειμένη δομή 10, ενώ την υψηλότερη τιμή κατέχει η δομή 11. Στους πίνακες υπερτονίζονται με κίτρινο χρώμα οι χαμηλότερες τιμές RMSD σε κάθε περίπτωση και με γκρι χρώμα οι υψηλότερες. Στον πίνακα 4.9 παρουσιάζεται ο συνολικός αριθμός μεγίστων και ελαχίστων τιμών που παρουσιάζει η κάθε NMR δομή. Με τον τρόπο αυτό μπορούμε να βρούμε την NMR δομή που εμφανίζει την μεγαλύτερη ομοιότητα (περισσότερα ελάχιστα RMSD) με τα τροχιακά των προσομοιώσεων.

	1-17		3-15		TOTAL	
	MIN	MAX	MIN	MAX	MIN	MAX
1	-	-	-	-	-	-
2	1	-	-	1	1	1
3	-	2	-	2	-	4
4	2	-	1	-	3	-
5	-	-	1	-	1	-
6	-	-	-	-	-	-
7	-	1	-	2	-	3
8	-	-	-	-	-	-
9	-	-	-	-	-	-
10	2	-	2	-	4	-
11	-	4	-	1	-	5
12	-	-	1	-	1	-
13	-	-	-	2	-	2

14	1	-	2	-	3	-
15	-	-	-	-	-	-
16	-	2	-	1	-	3
17	1	-	1	-	2	-
18	1	-	2	-	3	-
19	1	-	2	-	3	-
20	1	-	1	-	2	-
21	2	2	-	3	2	5
22	-	-	-	-	-	-
23	-	-	-	-	-	-
24	-	-	-	-	-	-
25	2	-	-	-	2	-
26	-	-	1	-	1	-
27	-	-	-	-	-	-

Πίνακας 4.9

Από την δέκατη δομή θα παρουσιαστούν οι πλησιέστερες προς αυτή δομές από το κάθε τροχιακό. Στην εικόνα 4.45 παρουσιάζεται η υπέρθεση της πειραματικά προσδιορισμένης δομής την κοντινότερή της σε κάθε τροχιακό.



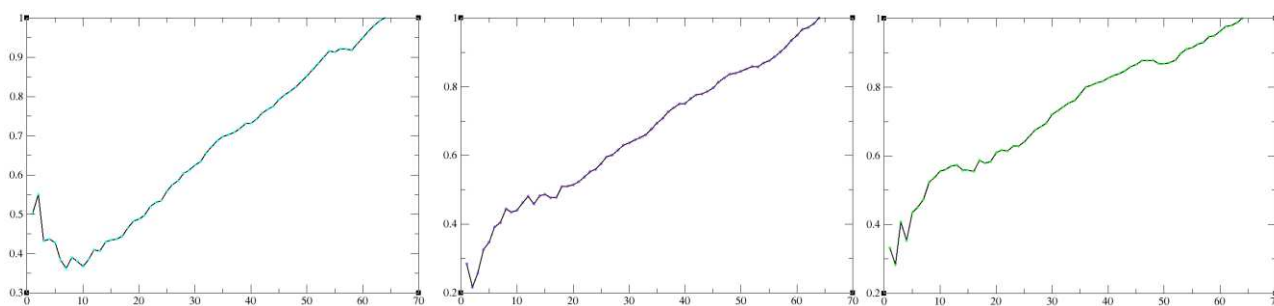
Εικόνα 4.45

Υπέρθωση του πεπτιδικού σκελετού της αντιπροσωπευτικής δομής από κάθε τροχιακό με την πειραματικά προσδιορισμένη δομή. Η NMR δομή έχει μπλε χρώμα, η 99SB έχει μοβ χρώμα, η ILDN έχει ροζ χρώμα και η protonated έχει πράσινο χρώμα.

4.5 Αλληλεπικάλυψη των τροχιακών

Προκειμένου να έχουν νόημα οι συγκρίσεις των τροχιακών που παρήχθησαν θα πρέπει να γνωρίζουμε εάν τα τροχιακά συγκλίνουν και κατά πόσο. Ένας τρόπος σύγκρισης της σύγκλισης είναι μέσω της επικάλυψης των eigenvectors. Χρησιμοποιώντας τα ζεύγη eigenvectors – eigenvalues μπορούμε να συγκρίνουμε διαφορετικά τροχιακά που προέρχονται από διαφορετικά δυναμικά πεδία και για διαφορετικούς συνδυασμούς ατόμων, σε επίπεδο ομοιότητας της κίνησης. Ο σκοπός του ελέγχου της αλληλεπικάλυψης των eigenvectors είναι ο έλεγχος για την επάρκεια του χρόνου προσομοίωσης, έτσι ώστε να έχουμε επαρκή δειγματοληψία.

Στην εικόνα 4.46 παρουσιάζονται οι γραφικές παραστάσεις αλληλεπικάλυψης (sufficient sampling) των πρώτων 64 eigenvectors, με τα υψηλότερα eigenvalues, από την ανάλυση dihedral-PCA για όλους τους συνδυασμούς των τροχιακών. Η μεγαλύτερη αλληλεπικάλυψη φαίνεται να υπάρχει μεταξύ των τροχιακών 99SB – ILDN. Το πρότυπο αλληλεπικάλυψης είναι παρόμοιο και στις τρεις περιπτώσεις, αλλά οι τιμές είναι χαμηλές για τους πρώτους eigenvectors. Συνεπώς τα τροχιακά δεν συγκλίνουν σε μεγάλο βαθμό και η κίνηση τους διαφέρει. Το γεγονός ότι μετά τους 20 πρώτους περίπου eigenvectors η αλληλεπικάλυψη ξεπερνά το 0.7, σημαίνει ότι τα τροχιακά περιγράφουν διαφορετικές κινήσεις. Στον πίνακα 4.10 παρουσιάζονται οι τιμές βαθμολόγησης της σύγκλισης των πρώτων τριών eigenvectors για τα ζεύγη 99SB-ILDN, 99SB-protonated και ILDN-protonated.



Εικόνα 4.46

Eigenspace overlap μεταξύ των τροχιακών ως συνάρτηση των 64 eigenvectors με τα υψηλότερα eigenvalues από την ανάλυση Dihedral-PCA. Αριστερά, με κυανό, φαίνεται η σύγκριση μεταξύ των τροχιακών 99SB-ILDN, στο κέντρο, με μοβ, φαίνεται η σύγκριση μεταξύ των τροχιακών 99SB-protonated και δεξιά, με πράσινο, φαίνεται η σύγκριση μεταξύ των τροχιακών ILDN- protonated.

overlap	99SB-ILDN	99SB-protonated	ILDN- protonated
Eigenvector 1	0.5022	0.2842	0.3330
Eigenvector 2	0.5509	0.2164	0.2841
Eigenvector3	0.4324	0.2569	0.4085

Πίνακας 4.10

4.6 Σύγκριση αποτελεσμάτων

Στο τέταρτο κεφάλαιο παρουσιάστηκαν τα αποτελέσματα της παρούσας πτυχιακής εργασίας μέσα από διάφορες αναλύσεις, οι οποίες συμπεριλαμβάνουν την μέτρηση της τιμής RMSD μεταξύ διαφόρων συνδυασμών τροχιακών, δομών, καταλοίπων και ατόμων, την ανάλυση πινάκων RMSD, τις αναλύσεις DSSP και STRIDE, την ανάλυση Cartesian-PCA, την ανάλυση Dihedral-PCA, με χ_1 γωνία και χωρίς χ_1 γωνία και τέλος την ανάλυση σύγκλισης τροχιακών.

Ο σκοπός όλων των παραπάνω αναλύσεων είναι ο έλεγχος του περιεχομένου των τροχιακών που παρήχθησαν από τις προσομοιώσεις και η σύγκρισή του με την πειραματικά προσδιορισμένη δομή, η οποία είναι κατατεθειμένη στην PDB. Επιπλέον, χρησιμοποιώντας τα αποτελέσματα θα μπορέσουμε να κάνουμε και μια σύγκριση των παραμέτρων που χρησιμοποιήθηκαν για να πραγματοποιηθούν οι προσομοιώσεις. Δηλαδή, θα συγκρίνουμε τα δυναμικά πεδία μεταξύ τους, για να δούμε ποιο ανταποκρίθηκε καλύτερα στους σκοπούς της παρούσας μελέτης.

Από τις RMSD αναλύσεις προέκυψε ένας χάρτης RMSD για κάθε τροχιακό, που παρουσιάζει αποστάσεις ατόμων για κάθε δομή του τροχιακού. Στους χάρτες αναλύθηκαν οι περιοχές οι οποίες εμφανίζουν χαμηλό RMSD για κάποιο χρονικό διάστημα. Με τον τρόπο αυτό εξήχθησαν από τα τροχιακά συνολικά 26 δομές, από τις οποίες 8 βρίσκονται στο τροχιακό 99SB, 12 στο τροχιακό ILDN και 6 στο τροχιακό protonated. Οι δομές αυτές είναι οι πλησιέστερες στη μέση δομή κάθε περιοχής με χαμηλό RMSD. Με τον τρόπο αυτό μπορούμε να δούμε και τα στάδια από τα οποία πέρασε το πεπτίδιο κατά τη διάρκεια της προσομοίωσης. Στις εικόνες 4. 47, 4.87, 4.49 παρουσιάζεται η εναλλαγή των σταθερότερων δομών των τροχιακών που αναλύθηκαν. Από τις 26 συνολικά δομές καταλήξαμε σε 3 αντιπροσωπευτικές δομές, μία για κάθε τροχιακό (εικόνα 4.16). Μία

σημαντική πρώτη παρατήρηση από την ανάλυση RMSD είναι ο χρόνος που χρειάστηκε για να αναδιπλωθούν τα μόρια σε κάθε προσομοίωση. Για το τροχιακό 99SB χρειάστηκαν 3.8 μs προκειμένου να παρατηρηθεί η δομή της β -φουρκέτας, ενώ για το τροχιακό η παρατήρηση αυτή έγινε μόλις στα 1.26 μs . Τέλος στο τροχιακό protonated χρειάστηκαν 2.69 μs για να αποκτηθεί η δομή της β -φουρκέτας. Ως πρώτη παρατήρηση λοιπόν μπορούμε να πούμε ότι η προσομοίωση ILDN ήταν η πιο γρήγορη, από τις τρεις προσομοιώσεις που πραγματοποιήθηκαν, στο σχηματισμό της “φυσικής δομής”.

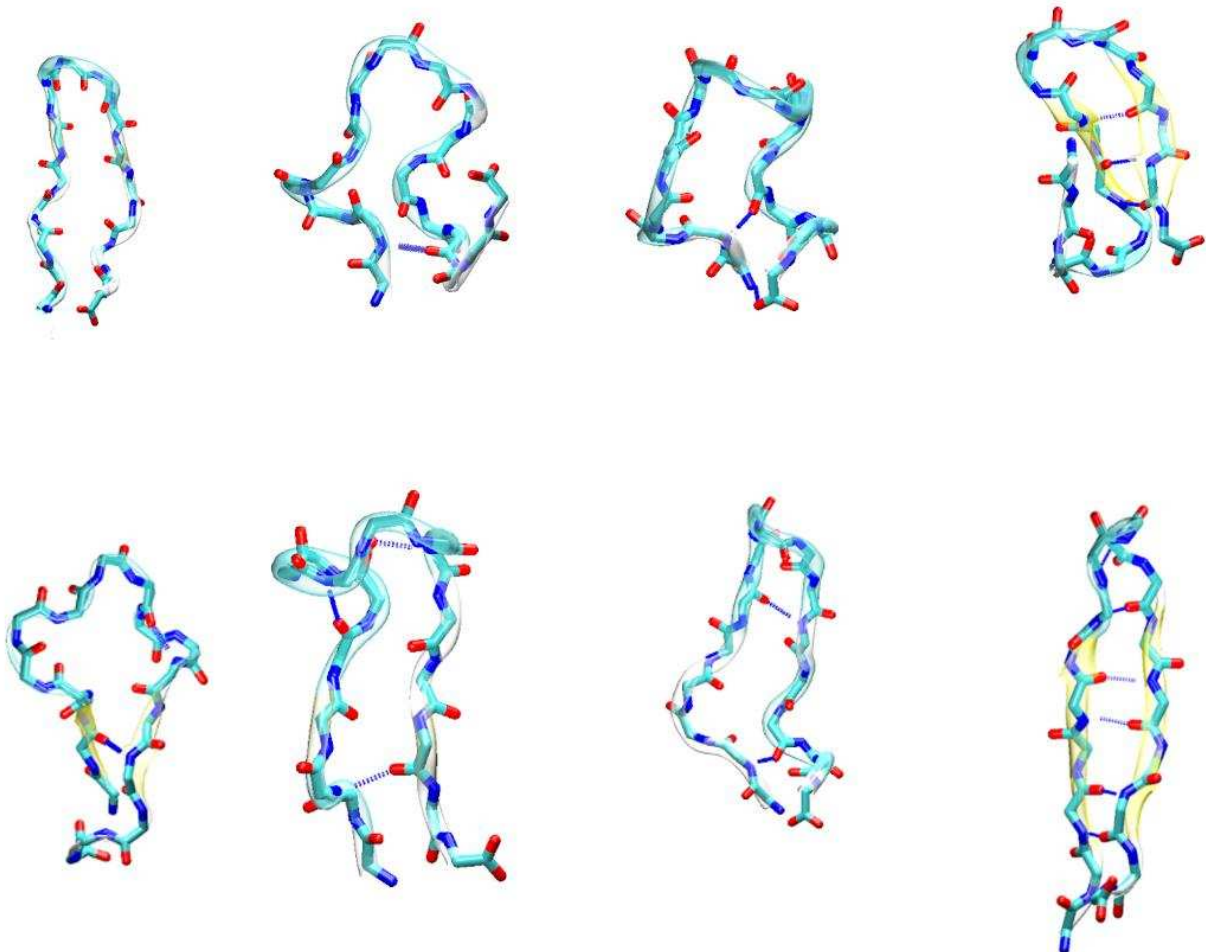
Στη συνέχεια πραγματοποιήθηκαν οι αναλύσεις δευτεροταγούς δομής με τους αλγόριθμους DSSP και STRIDE. Οι αναλύσεις δευτεροταγούς δομής μας δείχνουν εμφανώς την απόκτηση β -δομής και στα τρία τροχιακά με κάποιες διαφορές στα κατάλοιπα. Από την θέση των καταλοίπων της στροφής μπορούμε να συμπεράνουμε αν πρόκειται για ανάστροφη στροφή ή για στροφή δύο καταλοίπων. Στις περιπτώσεις 99SB και ILDN η στροφή είναι τύπου I, ενώ στην περίπτωση Protonated η στροφή είναι τύπου I'. Η θέση των καταλοίπων στη στροφή, βέβαια είναι απαραίτητη αλλά όχι επαρκής για τον προσδιορισμό του τύπου. Κύριο ρόλο παίζουν οι ϕ και ψ διέδρες γωνίες των καταλοίπων που βρίσκονται στη στροφή. Από την ανάλυση δευτεροταγούς δομής παρατηρούμε, εκτός από την απόκτηση της β -δομής, την διατήρησή της σε όλη την υπόλοιπη έκταση του τροχιακού. Τέλος στο τροχιακό protonated φαίνεται και η μετάθεση κατά ένα κατάλοιπο στην β -δομή, γεγονός που υποστηρίζει ότι η αντιπροσωπευτική δομή του τροχιακού διαφέρει από τις άλλες δύο και μοιάζει περισσότερο με την δεύτερη προτεινόμενη δομή των R. Zerella et al.. Στην εικόνα 4.20 φαίνεται η πλήρης συμφωνία μεταξύ των πινάκων RMSD και των αλγορίθμων DSSP και STRIDE.

Η επόμενη κατηγορία αναλύσεων αφορά την ομαδοποίηση δομών, που εμπεριέχονται στα τροχιακά, με βάση ορισμένα κοινά χαρακτηριστικά. Το πρώτο κομμάτι αφορά την Cartesian Principal Components Analysis, ενώ το δεύτερο

κομμάτι την Dihedral Principal Components Analysis. Η γενική ιδέα είναι η μελέτη ενεργειακών πεδίων και η εύρεση ενεργειακών ελαχίστων, υπό την προϋπόθεση ότι η φυσική δομή κατέχει την ελάχιστη ενέργεια και συνεπώς στις ομάδες με την χαμηλότερη ενέργεια πρέπει να βρίσκονται οι αντιπροσωπευτικές δομές των τροχιακών. Στην Cartesian-PCA εφαρμόστηκαν δύο τεχνικές ομαδοποίησης, στην πρώτη εφαρμόζεται η ανάλυση για τα άτομα του πεπτιδικού σκελετού και στην δεύτερη για όλα τα βαριά άτομα. Τα ενεργειακά πεδία περιέχουν λίγα ενεργειακά ελάχιστα και φαίνονται πιο ομαλά σε σχέση με αυτά που παρήχθησαν από την Dihedral-PCA. Τα Clusters που προέκυψαν και από τις δύο τεχνικές είναι αρκετά μεγάλα και καλύπτουν ικανοποιητικό ποσοστό του χάρτη διακύμανσης (πίνακες 4.1 και 4.2). Το πρώτο cluster της κάθε ανάλυσης έχει τη χαμηλότερη ελεύθερη ενέργεια και από αυτό προέρχεται η αντιπροσωπευτική δομή κάθε τροχιακού. Από την υπέρθεση δομών παρατηρείται ελάχιστη απόκλιση μεταξύ των δομών των κυρίαρχων ομάδων. Στις αντιπροσωπευτικές δομές που απομονώθηκαν δημιουργήθηκε η δομή της β-φουρκέτας και στα τρία τροχιακά, με την διαφορά ότι στο τροχιακό protonated υπάρχει μετάθεση κλώνου, διαφορετικό μοτίβο υδρογονοδεσμών και διαφορετικός τύπος στροφής. Στην Dihedral-PCA πραγματοποιήθηκαν επίσης δύο διαφορετικές τεχνικές, η πρώτη ανάλυση έγινε με βάση τις διέδρες γωνίες ϕ και ψ , και η δεύτερη συμπεριελάμβανε εκτός αυτών και την χ_1 γωνία. Η πρώτη περίπτωση εστιάζει στον πεπτιδικό σκελετό, ενώ η δεύτερη στις πλευρικές ομάδες. Τα ενεργειακά πεδία από την dPCA περιλαμβάνουν περισσότερες ομάδες δομών και χαρακτηρίζονται περισσότερο τραχεία (εικόνα 4.31, 4.38). Επιπλέον, το μέγεθος των ομάδων είναι μικρότερο σε σχέση με αυτά που παρήχθησαν στην cPCA. Στην πρώτη περίπτωση η κάλυψη του χάρτη διακύμανσης από τα clusters είναι αρκετά υψηλή, όπως και με την cPCA, αλλά στην δεύτερη περίπτωση μειώνεται (ειδικά στην περίπτωση του τροχιακού protonated). Ωστόσο τα κυρίαρχα τροχιακά περιλαμβάνουν δομές που σχηματίζουν τη δομή της β-φουρκέτας, αν και στην περίπτωση του τροχιακού protonated εμφανίζεται η εναλλακτική δομή (εικόνα 4.43). Η υπέρθεση των κυρίαρχων τροχιακών δεν

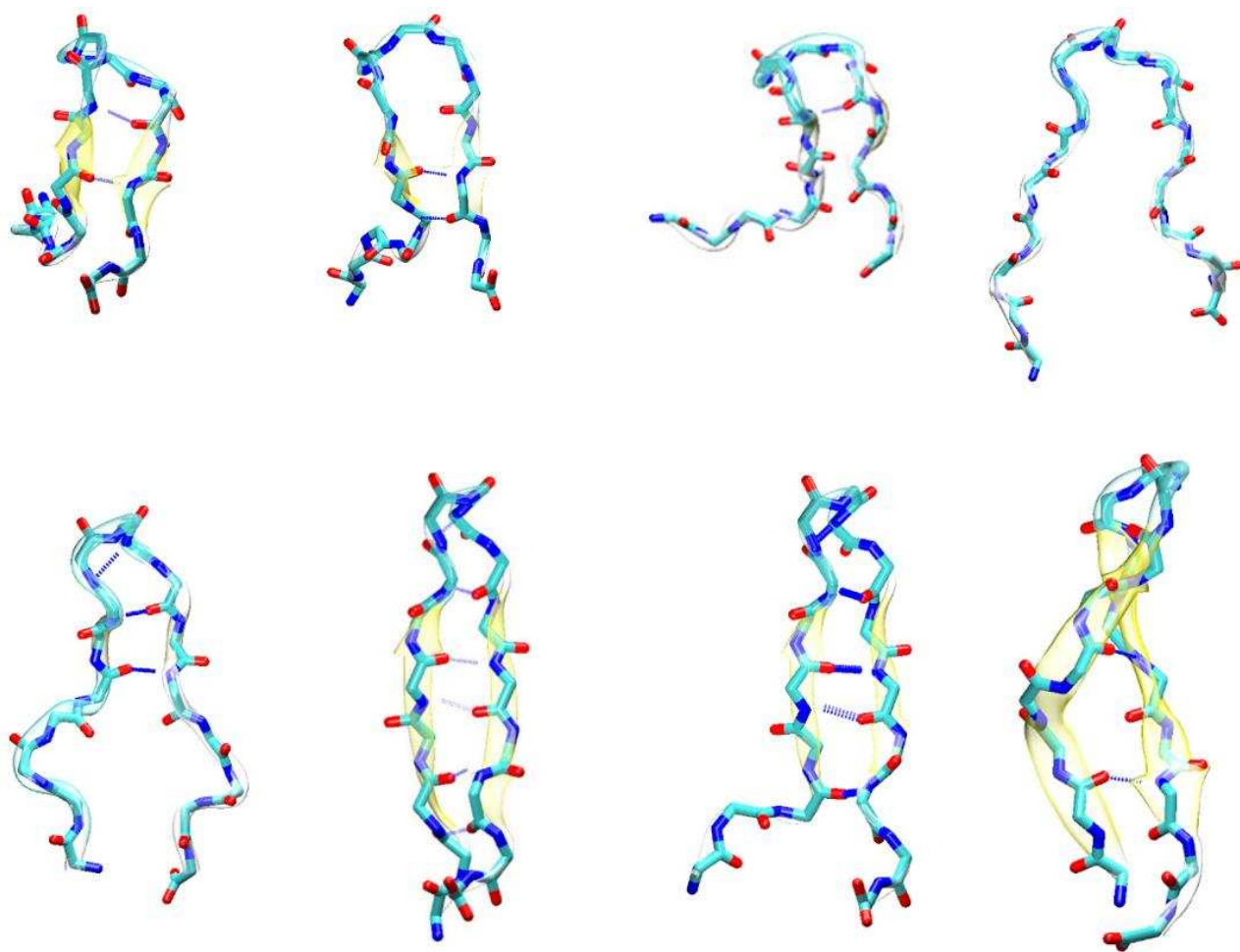
παρουσιάζει μεγάλη απόκλιση των δομών (εικόνες 4.37 και 4.42).

Η σύγκριση του RMSD των δομών έχουν κατατεθεί την πρωτεϊνική βάση δεδομένων παρουσιάζει ένα εύρος τιμών από 0.56 Å έως 4.36 Å. Οι μεγαλύτερες τιμές παρατηρούνται στην σύγκριση όλων των βαρέων ατόμων του πεπτιδίου, ενώ οι χαμηλότερες στην σύγκριση των ατόμων Ca. Από τις κατατεθειμένες δομές μεγαλύτερη ομοιότητα φαίνεται να παρουσιάζει η δέκατη δομή, καθώς έχει τα περισσότερα ελάχιστα τους πίνακες RMSD, ενώ η ενδέκατη δομή φαίνεται να παρουσιάζει μεγαλύτερη διαφορά από τις δομές των τροχιακών, καθώς παρουσιάζει τα περισσότερα μέγιστα στους πίνακες RMSD (πίνακας 4.9).



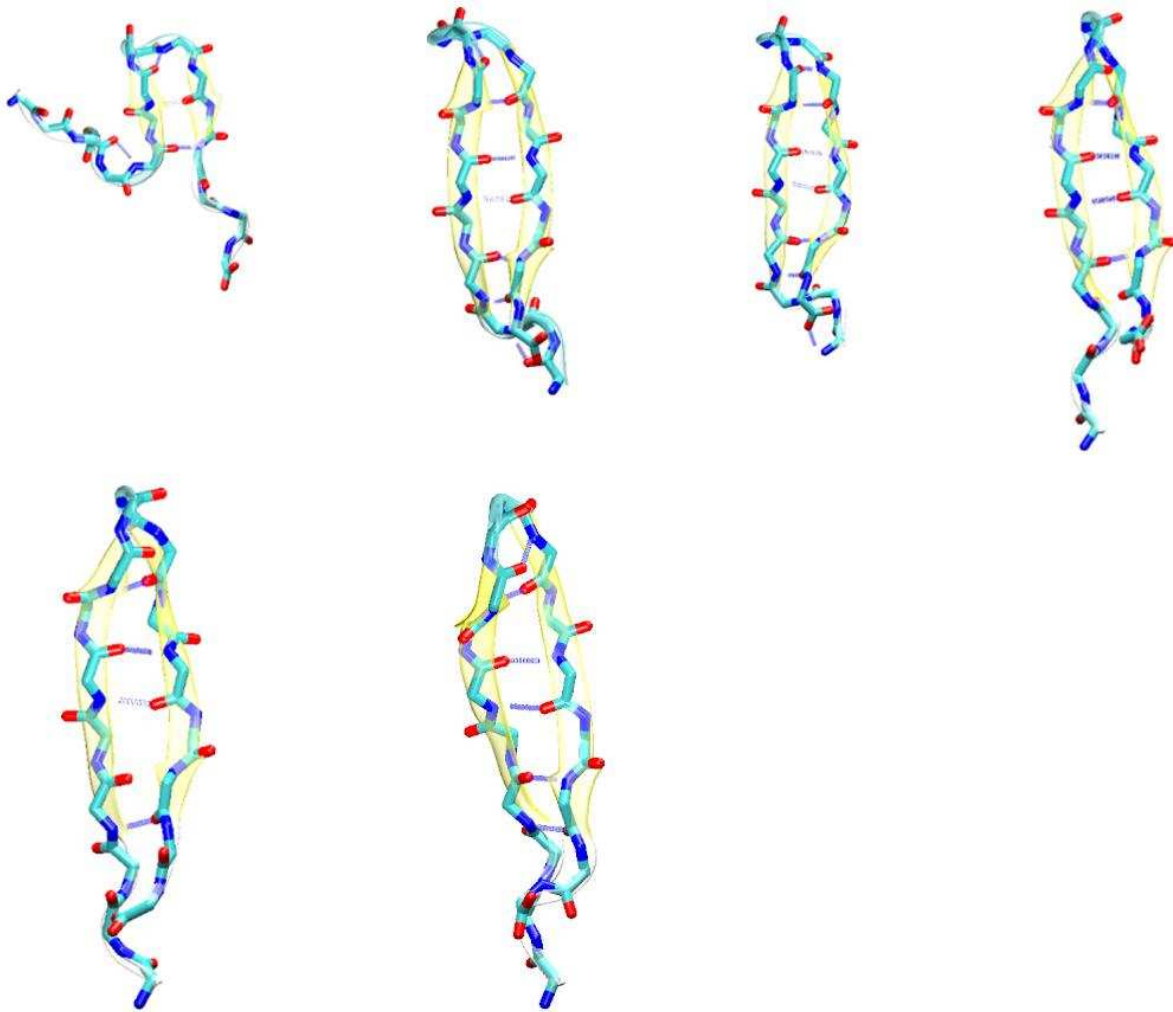
Εικόνα 4.47

Στιγμιότυπα από το τροχιακό 99SB, όπως απομονώθηκαν από τους πίνακες RMSD.



Εικόνα 4.48

Στιγμιότυπα από το τροχιακό ILDN, όπως απομονώθηκαν από τους πίνακες RMSD.



Εικόνα 4.49

Στιγμιότυπα από το τροχιακό protonated, όπως απομονώθηκαν από τους πίνακες RMSD

REFLECTIONS

"Faith does not protect you. Medicine and airbags... Those are the things that protect you. God does not protect you. Intelligence protects you. Enlightenment. Put your faith in something with tangible results. How long has it been since someone walked on water? Modern miracles belong to science.. Computers, vaccines, space stations... Even the divine miracle of creation. Matter from nothing... In a lab. Who needs God? No! Science is God!"

Dan Brown

Συμπεράσματα και Συζήτηση

Η ουβικιτίνη είναι μία από τις σημαντικότερες πρωτεΐνες που συμμετέχουν στη διαδικασία της ουβικιτινίωσης διαφόρων υποστρωμάτων εντός του κυττάρου. Είναι σχετικά μικρή πρωτεΐνη αποτελούμενη από 76 αμινοξέα. Το N-τελικό δεκαεπταπεπτίδιο της ουβικιτίνης αποτελεί αντικείμενο μελέτης ως προς την δομή και τη λειτουργία του από τη δεκαετία του '90. Μία παραλλαγή του φυσικού πεπτιδίου αποτελεί και τον στόχο τη παρούσας πτυχιακής εργασίας. Το πεπτίδιο 1E0Q φέρει στην αλληλουχία του την μετάλλαξη T9D, η οποία, όπως υποστηρίζεται από προηγούμενες μελέτες, προσδίδει σταθερότητα στη δομή του πεπτιδίου. Πειράματα NMR και κρυσταλλογραφίας ακτίνων X απέδειξαν πως η δομή του 1E0Q είναι μία β-φουρκέτα με τύπου I β-στροφή.

Πραγματοποιήθηκαν τρεις προσομοιώσεις μοριακής δυναμικής συνολικής διάρκειας 10 μs χρησιμοποιώντας το πεπτίδιο 1E0Q με διαφορετικές παραμέτρους σε κάθε προσομοίωση. Στην πρώτη προσομοίωση χρησιμοποιήθηκε το δυναμικό πεδίο AMBER99SB, ενώ στις άλλες δύο το AMBER99SB-ILDN. Από τα αποτελέσματά μας είμαστε σε θέση να κάνουμε μία αρχική σύγκριση των παραγόμενων δομών με τις πειραματικά προσδιορισμένες δομές και να αποδώσουμε κάποια συγκριτικά συμπεράσματα για την απόδοση των δυναμικών πεδίων που χρησιμοποιήθηκαν. Όπως προαναφέρθηκε σε προηγούμενο κεφάλαιο, ο στόχος μας είναι η μελέτη της αναδίπλωσης του πεπτιδίου 1E0Q, η οποία θα συμβάλει στη μελέτη της αναδίπλωσης των πρωτεϊνών από τη σκοπιά των προσομοιώσεων μοριακής δυναμικής, δηλαδή με “physics based” μοντέλα.

Τα αποτελέσματα των προσομοιώσεων συμφωνούν ως επί το πλείστον με τα πειραματικά δεδομένα και φαίνεται ότι η φυσική δομή έχει προσεγγισθεί σε μεγάλο βαθμό. Στις δύο από τις τρεις προσομοιώσεις οι κυρίαρχες δομές

συγκλίνουν αρκετά με τις πειραματικές, ενώ στην τρίτη η κυρίαρχη δομή προσεγγίζει μια εναλλακτική διευθέτηση. Ωστόσο, στο τροχιακό ILDN το πεπτίδιο απέκτησε πιο γρήγορα την native-like διαμόρφωση και την διατήρησε μέχρι το τέλος, σε αντίθεση με την προσομοίωση 99SB, στην οποία το πεπτίδιο, παρόλο που πέρασε από αρκετά ενδιάμεσα στάδια προκειμένου να καταλήξει στην native-like διαμόρφωση, δεν την διατήρησε σταθερή μέχρι το τέλος του τροχιακού.. Το αποτέλεσμα αυτό σημαίνει πως το δυναμικό πεδίο AMBER99sb-ILDN είναι πιο αποτελεσματικό από το AMBER99sb ή χάνεται η πληροφορία των ενδιάμεσων σταδίων χρησιμοποιώντας το; Προκειμένου να απαντηθεί η παραπάνω ερώτηση θα πρέπει να πραγματοποιηθούν νέες προσομοιώσεις μοριακής δυναμικής, με νέα δυναμικά πεδία και να συγκριθεί η συμπεριφορά του πεπτιδίου κατά την αναδίπλωση.

Το τροχιακό το οποίο δεν συμφωνεί με τα πειραματικά δεδομένα είναι το protonated, το οποίο κατέληξε μεν, σε διαμόρφωση β-φουρκέτας, αλλά δεν είναι η σωστή διαμόρφωση σύμφωνα με τα πειράματα που προηγήθηκαν στο πεπτίδιο. Η εναλλακτική διαμόρφωση που απέκτησε το πρωτονιομένο πεπτίδιο χαρακτηρίζεται από μία μετατόπιση των κλώνων της φουρκέτας κατά ένα κατάλοιπο και από πιο στενή στροφή. Η διαμόρφωση αυτή ταιριάζει στην εναλλακτική διαμόρφωση που προτείνεται από την ομάδα της R. Zerella, αλλά δεν υπάρχει κατατεθειμένη δομή προκειμένου να προβούμε σε συγκρίσεις. Το αποτέλεσμα αυτό είναι συνέπεια των συνθηκών της αναδίπλωσης (pH); Αν ναι, τι συνέπειες μπορεί να έχει η αλλαγή της διαμόρφωσης του πεπτιδίου σε ολόκληρη της πρωτεΐνη, εντός του κυττάρου; Η αλλαγή αυτή παρατηρείται και στο φυσικό πεπτίδιο ή είναι ιδιότητα του μεταλλαγμένου πεπτιδίου μόνο;

Τέλος το ενεργειακό πεδίο φαίνεται να είναι ομαλό και να έχει τη μορφή χοάνης στην προσομοίωση με το δυναμικό πεδίο AMBER99SB-ILDN, καθώς από τα αποτελέσματα παρατηρείται μόνο ένα γεγονός αναδίπλωσης το οποίο παραμένει μέχρι το τέλος της προσομοίωσης. Στις άλλες δύο προσομοιώσεις το ενεργειακό

τοπίο φαίνεται να είναι περισσότερο τραχύ και να συμπεριλαμβάνει περισσότερα γεγονότα αναδίπλωσης και ενδιάμεσα στάδια. Θα μπορούσαν όλες οι παραπάνω παρατηρήσεις να βοηθήσουν στο γενικότερο πρόβλημα της αναδίπλωσης των πρωτεϊνών και αν ναι με ποιο τρόπο;

Όλα τα ερωτήματα που δημιουργούνται εξετάζοντας τα αποτελέσματα της παρούσας πτυχιακής εργασίας τονίζουν την πολυπλοκότητα του κύριου προβλήματος και αποτελούν απόδειξη της εξαιρετικής δυσκολίας που αντιμετωπίζει η επιστημονική κοινότητα στην λύση του προβλήματος της αναδίπλωσης των πρωτεϊνών, καθώς θέτοντας ένα ερώτημα απορρέουν δεκάδες επόμενα που απαιτούν διευκρίνηση.

Τα αποτελέσματα μιας μελέτης δεν είναι ποτέ επαρκή, για να απαντήσουν τέτοιου είδους ερωτήματα, αν δεν είναι συγκρίσιμα με αποτελέσματα συμπληρωματικών μελετών. Είναι ενθαρρυντικό το γεγονός ότι έχουν ήδη αρχίσει να πραγματοποιούνται μελέτες μοριακής δυναμικής σε μεγάλα πρωτεϊνικά συστήματα, συμπεριλαμβανομένης και της ουβικιτίνης, οι οποίες έχουν ως στόχο την παρουσίαση της δομής αλλά και τη βαθμολόγηση και τη σύγκριση των δυναμικών πεδίων (D.E Shaw et al., 2012, A. Ganoth et al., 2013). Αξιοποιώντας την ποικιλία των πειραματικών δεδομένων και την αυξημένη υπολογιστική ισχύ μπορούμε να συνεχίσουμε της μελέτες στο πεδίο αυτό πραγματοποιώντας προσομοιώσεις με το μεταλλαγμένο αλλά και το φυσικό πεπτιδίο, ελεύθερο αλλά και ενσωματωμένο στην ουβικιτίνη, με διαφορετικά δυναμικά πεδία και συνθήκες, έτσι ώστε να καλύψουμε ένα μεγάλο τμήμα των συνθηκών του κυττάρου.

Προχωρώντας ένα βήμα παραπέρα θα μπορούσαμε να σχηματίσουμε *in silico* μορφές του πεπτιδίου, αλλά και της πρωτεΐνης ολόκληρης, που εμπλέκονται σε διάφορες ασθένειες και να ελέγξουμε τη λειτουργικότητα τους, την επαγωγή αλλά και την καταστολή τους. Με τον τρόπο αυτό μπορούμε να συνδέσουμε τις

προσομοιώσεις μοριακής δυναμικής με το σχεδιασμό και την ανάπτυξη
φαρμακευτικών στόχων.

Βιβλιογραφικές αναφορές

- 1. Ken a. Dill, s. Banu ozkan, m. Scott shell, thomas r. Weikl, the protein folding problem, annu rev biophys. 2008; 9; 37: 289–316.**
- 2. Callender rh, dyer rb, gilmanshin r, woodruff wh. Fast events in protein folding: the time evolution of primary processes. Annu rev phys chem. 1998; 49:173–202.**
- 3. Fersht ar, sato s. Φ -value analysis and the nature of protein-folding transition states. Proc natl acad sci usa. 2004; 101:7976–81**
- 4. Sosnick tr, dothager rs, krantz ba. Differences in the folding transition state of ubiquitin indicated by Φ and ψ analyses. Proc natl acad sci usa. 2004; 101:17377–82**
- 5. Goldenberg dp. Genetic studies of protein stability and mechanisms of folding. Annu rev biophys biomol struct. 1988; 17:481–507**
- 6. Itzhaki ls, otzen de, fersht ar. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. J mol biol. 1995; 254:260–88.**
- 7. Schuler b, lipman ea, eaton wa. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. Nature. 2002; 419:743–47.**
- 8. Maity h, maity m, krishna mm, mayne l, englander sw. Protein folding: the stepwise assembly of foldon units. Proc natl acad sci usa. 2005; 102:4741–46.**
- 9. Krishna mm, hoang l, lin y, englander sw. Hydrogen exchange methods to study protein folding. Methods. 2004;34:51–64.**
- 10. Karplus m, weaver dl. Diffusion-collision model for protein folding. Biopolymers. 1979; 18:1421–37.**
- 11. Karplus m, weaver dl. Protein folding dynamics: the diffusion-collision model and experimental data. Protein sci. 1994; 3:650–68.**

12. Mccammon ja, gelin br, karplus m. Dynamics of folded proteins. *Nature*. 1977; 267:585–90.
13. Myers jk, oas tg. Preorganized secondary structure as an important determinant of fast protein folding. *Nat struct biol*. 2001; 8:552–58.
14. Fersht ar. Nucleation mechanisms in protein folding. *Curr opin struct biol*. 1997; 7:3–9.
15. Itzhaki ls, otzen de, fersht ar. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J mol biol*. 1995; 254:260–88.
16. Alexander, patrick, stephen fahnestock, timothy lee, john orban, and philip bryan. "thermodynamic analysis of the folding of the streptococcal protein g igg-binding domains b1 and b2: why small proteins tend to have high denaturation temperatures." *Biochemistry* 31, no. 14, 1992; 3597-3603.
17. Viguera ar, serrano l, wilmanns m. Different folding transition states may result in the same native structure. *Nat struct biol*. 1996; 3:874–80.
18. Kragelund, birthe b., Carol v. Robinson, jens knudsen, christopher m. Dobson, and flemming m. Poulsen. "folding of a four-helix bundle: studies of acyl-coenzyme a binding protein." *Biochemistry* 34, no. 21 1995; 7217-7224.
19. Huang, guewha s., and terrence g. Oas. "submillisecond folding of monomeric lambda repressor." *Proceedings of the national academy of sciences* 92, no. 15 1995; 6878-6882.
20. Dill ka, chan hs. From levinthal to pathways to funnels. *Nat struct biol*. 1997; 4:10–19.
21. Plaxco kw, simons kt, baker d. Contact order, transition state placement and the refolding rates of single domain proteins. *J mol biol*. 1998; 277:985–94.
22. Ivankov dn, finkelstein av. Prediction of protein folding rates from the

- amino acid sequence-predicted secondary structure. *Proc natl acad sci usa*. 2004; 101:8942–44.
23. Huang jt, cheng jp, chen h. Secondary structure length as a determinant of folding rate of proteins with two- and three-state kinetics. *Proteins*. 2007; 67:12–17.
24. Gromiha mm, selvaraj s. Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J mol biol*. 2001; 310:27–32.
25. Mirny l, shakhnovich e. Protein folding theory: from lattice to all-atom models. *Annu rev biophys biomol struct*. 2001; 30:361–96.
26. Gromiha mm, thangakani am, selvaraj s. Fold-rate: prediction of protein folding rates from amino acid sequence. *Nucleic acids res*. 2006; 34:w70–74.
27. Ozkan sb, wu gha, chodera jd, dill ka. Protein folding by zipping and assembly. *Proc natl acad sci usa*. 2007; 104:11987–92.
28. Chou py, fasman gd. Prediction of protein conformation. *Biochemistry*. 1974; 13:222–45.
29. Chou py, fasman gd. Empirical predictions of protein conformation. *Annu rev biochem*. 1978; 47:251–76.
30. J, pedersen jt, judson r, fidelis k. A large-scale experiment to assess protein structure prediction methods. *Proteins*. 1995; 23:ii–iv.
31. Moulton j, hubbard t, bryant sh, fidelis k, pedersen jt. Critical assessment of methods of protein structure prediction (casp)— round ii. *Proteins s* 1997; 1:2–6.
32. Moulton j, john, krzysztof fidelis, adam zemla, and tim hubbard. "critical assessment of methods of protein structure prediction (casp): round iv." *Proteins: structure, function, and bioinformatics* 45, no. S5 2001; 2-7.
33. Venclovas c, zemla a, fidelis k, moulton j. Assessment of progress over the casp experiments. *Proteins*. 2003; 53:585–95.

34. Moulton J. A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction. *Curr Opin Struct Biol.* 2005; 15:285–89.
35. Moulton J. "Predicting protein three-dimensional structure." *Current Opinion in Biotechnology* 10, no. 6 1999; 583-588.
36. Moulton J, Krzysztow Fidelis, Adam Zemla, and Tim Hubbard. "Critical assessment of methods of protein structure prediction (CASP)-round V." *Proteins: structure, function, and bioinformatics* 53, no. S6 2003; 334-339.
37. Moulton J, Krzysztow Fidelis, Burkhard Rost, Tim Hubbard, and Anna Tramontano. "Critical assessment of methods of protein structure prediction (CASP)—round 6." *Proteins: structure, function, and bioinformatics* 61, no. S7 2005; 3-7.
38. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* 1993; 234:779–815.
39. Berman H, Westbrook J, Feng Z, Gilliland T, N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. "The protein data bank." *Nucleic acids research* 28, no. 1 2000; 235-242.
40. Skolnick J, Yang Zhang, Adrian K. Arakaki, Andrzej Kolinski, Michal Boniecki, András Szilágyi, and Daisuke Kihara. "Touchstone: a unified approach to protein structure prediction." *Proteins: structure, function, and bioinformatics* 53, no. S6 2003; 469-479.
41. Skolnick J, Daisuke Kihara, and Yang Zhang. "Development and large scale benchmark testing of the PROSPECTOR_3 threading algorithm." *Proteins: structure, function, and bioinformatics* 56, no. 3 2004; 502-518.
42. Zhang Y, Andrzej Kolinski, and Jeffrey Skolnick. "Touchstone II: a new approach to ab initio protein structure prediction." *Biophysical Journal* 85, no. 2 2003; 1145-1164.
43. Zhang Y, and Jeffrey Skolnick. "Scoring function for automated assessment of protein structure template quality." *Proteins: structure, function, and bioinformatics* 57, no. 4 2004; 702-710.
44. Zhang Y, and Jeffrey Skolnick. "SPICKER: a clustering approach to

- identify near-native protein folds." *Journal of computational chemistry* 25, no. 6 2004; 865-871.
45. Adjiman cs, androulakis ip, maranas cd, floudas ca.. A global optimization method, abb, for process design. *Comp chem eng* 1996 ; 20:s419–s424.
46. Adjiman cs, androulakis ip, floudas ca.. Global optimization of minlp problems in process synthesis and design. *Comp. Chem. Eng* . 1997; 21:s445–s450.
47. Androulakis ip, maranas cd, floudas ca.. abb: a global optimisation method for general constrained nonconvex problems. *J global optim* 1995; 7:337–363.
48. Floudas ca. *Deterministic global optimization: theory, methods and applications. Nonconvex optimization and its applications. Dordrecht, the netherlands: kluwer academic publishers, 2000.*
49. Floudas ca.. Research challenges, opportunities, and synergism in systems engineering and computational biology. *Aiche j* 2005; 51:1872–1884.
50. Floudas ca, akrotirianakis ig, caratzoulas s, meyer ca, kallrath j. Global optimization in the 21st century: advances and challenges. *Comp chem eng* 2005; 29:1185–1202.
51. Floudas, c. A. "computational methods in protein structure prediction." *Biotechnology and bioengineering* 97, no. 2 2007; 207-213.
52. Gnanakaran, s., Hugh nymeyer, john portman, kevin y. Sanbonmatsu, and angel e. Garcia. "peptide folding simulations." *Current opinion in structural biology* 13, no. 2 2003; 168-174.
53. Snow, christopher d., Houbi nguyen, vijay s. Pande, and martin gruebele. "absolute comparison of simulated and experimental protein-folding dynamics." *Nature* 420, no. 6911 2002; 102-106.
54. De alba e, antoro j, rico ma, jimenez ma, de novo design of a monomeric three-stranded anti-parallel beta-sheet. *Protein sci* 1999; 8:854-865.
55. Ramirez-alvarado m, blanco fj, serrano l, elongation of the bh8 beta-hairpin peptide: electrostatic interactions in beta-hairpin formation and

- stability. *Protein sci* 2001 10:1381-1392.
56. Pastor mt, de la paz ml, lacroix e, serrano l, perez-paya e, combinatorial approaches: a new tool to search for highly structured beta-hairpin peptides. *Proc natl acad sci usa* 2002;99:614-619.
57. Neidigh jw, fesinmeyer rm, andersen nh, designing a 20-residue protein. *Nat struct biol* 2002; 9:425-430.
58. Williams s, causgrove tp, gilmanshin r, fang ks, callender rh, Woodruff wh, dyer rb, fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry* 1996; 35:691-697.
59. Munoz v, thompson pa, hofrichter ja, eaton wa, folding dynamics and mechanism of beta-hairpin formation. *Nature* 1997; 390:196-199.
60. Crane jc, koepf ek, kelly jw, gruebele m, mapping the transition state of the ww domain beta-sheet. *J mol biol* 2000; 298:283-292.
61. Lapidus lj, eaton wa, hofrichter j: measuring the rate of intramolecular contact formation in polypeptides. *Proc natl acad sci usa* 2000; 97:7220-7225.
62. Sporlein s, carstens h, satzger h, renner c, behrendt r, moroder l, tavan p, zinth w, wachtveitl j: ultrafast spectroscopy reveals subnanosecond peptide conformation dynamics and validates molecular dynamics simulation. *Proc natl acad sci usa* 2002; 99:7998-8002.
63. Sugita y, okamoto y, replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape. *Chem phys lett* 2000; 329:261-270.
64. Shirts mr, pande vs: mathematical analysis of coupled parallel simulations. *Phys rev lett* 2001; 86:4983-4987.
65. Latham, peter w. "therapeutic peptides revisited." *Nature biotechnology* 17, 1999; 755-758.
66. Patch, james a., and annelise e. Barron. "mimicry of bioactive peptides via non-natural, sequence-specific peptidomimetic oligomers." *Current opinion in chemical biology* 6, no. 6, 2002; 872-877.

- 67.Horne, w. Seth, and samuel h. Gellman. "foldamers with heterogeneous backbones." *Accounts of chemical research* 41, no. 10, 2008; 1399-1408.
- 68.Tew, gregory n., Richard w. Scott, michael l. Klein, and william f. Degrado. "de novo design of antimicrobial polymers, foldamers, and small molecules: from discovery to practical applications." *Accounts of chemical research* 43, no. 1, 2009; 30-39.
- 69.Horne, w. Seth, melissa d. Boersma, matthew a. Windsor, and samuel h. Gellman. "sequence-based design of α/β -peptide foldamers that mimic bh3 domains." *Angewandte chemie international edition* 47, no. 15, 2008; 2853-2856.
- 70.Wang, guangshun, xia li, and zhe wang. "apd2: the updated antimicrobial peptide database and its application in peptide design." *Nucleic acids research* 37, no. suppl 1, 2009; d933-d937.
- 71.Shtatland, timur, daniel guettler, misha kossodo, misha pivovarov, and ralph weissleder. "pepbank-a database of peptides based on sequence text mining and public peptide data sources." *Bmc bioinformatics* 8, no. 1, 2007; 280.
- 72.Duchrow, timo, timur shtatland, daniel guettler, misha pivovarov, stefan kramer, and ralph weissleder. "enhancing navigation in biomedical databases by community voting and database-driven text classification." *Bmc bioinformatics* 10, no. 1 2009; 317.
- 73.Shandler, scott j., Maxim v. Shapovalov, roland l. Dunbrack, jr, and william f. Degrado. "development of a rotamer library for use in β -peptide foldamer computational design." *Journal of the american chemical society* 132, no. 21, 2010; 7312-7320.
- 74.Webb, g. C., R. T. Baker, k. Fagan, and p. G. Board. "localization of the human ubb polyubiquitin gene to chromosome band 17p11. 1-17p12." *American journal of human genetics* 46, no. 2, 1990; 308.
- 75.Schlesinger mj, bond mj ubiquitin genes. *Oxf surv . Eukaryotic genes*, 1987; 4:77-91 .

76. Lund pk, moats-staats bm, simmons jg, hoyt e, d'ercole Aj, martin f, van wykjj, nucleotide sequence analysis of a cDNA encoding human ubiquitin reveals that ubiquitin is synthesized as a precursor. *J Biol Chem*, 1985; 260: 7609-7613 .
77. Salvesen g, loyd c, farley d, cDNA encoding a human homolog of yeast ubiquitin 1. *Nucleic acids res* 1987; 15:5485 .
78. Baker rt molecular analysis of human ubiquitin genes. Ph.D. thesis, the Australian National University, Canberra, 1988 .
79. Baker rt, board pg the human ubiquitin gene family: structure of a gene and pseudogenes from the ubb subfamily. *Nucleic acids res* 1987a; 15:443-463 ,
80. Wiborg O, pedersen ms, wind a, berglund le, marcker ke, vuustj the human ubiquitin multigene family: some genes contain multiple directly repeated ubiquitin coding sequences. *Embo j* 1985; 4:755-759 .
81. Baker rt, board pg , unequal crossover generates variation in ubiquitin coding unit number at the human ubcpolyubiquitin locus. *Am j hum genet* 1989; 44:534-542 11.
82. Baker rt, board pg , nucleotide sequence of a human ubiquitin ubb processed pseudogene. *Nucleic acids res* 1987b; 15:4352 .
83. Cowland jb, wiborg O, vuust j, human ubiquitin genes: one member of the ubb gene subfamily is a tetrameric non processed pseudogene. *Febs lett* 1988; 231:187-191 .
84. Zerella, rosa, philip a. Evans, j. O. H. N. Ionides, len c. Packman, b. Trotter, joel p. Mackay, and dudley h. Williams. "autonomous folding of a peptide corresponding to the n-terminal β -hairpin from ubiquitin." *Protein science* 8, no. 6, 1999; 1320-1331.
85. Dyson hj, rance m, houghten ra, lerner ra, wright pe. Folding of immunogenic peptide fragments of proteins in water solution. 1. Sequence requirements for the formation of a reverse turn. *J mol biol* 1988; 201:161-200.

86. Dyson hj, merutka g, waltho jp, lerner ra, wright pe. Folding of peptide-fragments comprising the complete sequence of proteins—models for initiation of protein folding. 1. Myohemerythrin. *J mol biol* 1992; 226:795–817.
87. Scholtz jm, baldwin rl. The mechanism of α -helix formation by peptides. *Ann rev biophys biomol struct* 1992; 21:95–118.
88. Cox jpl, evans pa, packman lc, williams dh, woolfson dn. Dissecting the structure of a partially folded protein—circular dichroism and nuclear magnetic resonance studies of peptides from ubiquitin. *J mol biol* 1993; 234:483–492.
89. Blanco fj, jimenez ma, pineda a, rico m, santoro j, nieto jl. Nmr solution structure of the isolated n-terminal fragment of protein-g b-1 domain—evidence of trifluoroethanol induced native-like b-hairpin formation. *Biochemistry* 1994a; 33:6004–6014.
90. Blanco fj, rivas g, serrano l. A short linear peptide that folds into a native stable b-hairpin in aqueous solution. *Nature struct biol* 1994b; 1:584–590.
91. Searle ms, williams dh, packman lc. A short linear peptide derived from the n-terminal sequence of ubiquitin folds into a water-stable non-native b-hairpin. *Nature struct biol* 1995; 2:999–1006.
92. Searle ms, zerella r, williams dh, packman lc. Native-like b-hairpin structure in an isolated fragment from ferredoxin—nmr and cd studies of solvent effects on the n-terminal-20 residues. *Protein eng* 1996; 9:559–565.
93. Blanco fj, jimenez ma, herranz j, rico m, santoro j, nieto jl. Nmr evidence of a short linear peptide that folds into a b-hairpin in aqueous solution. *J am chem soc* 1993; 115:5887–5888.
94. De alba e, jimenez ma, rico m, nieto jl. Conformational investigation of designed short linear peptides able to fold into b-hairpin structures in aqueous solution. *Folding design* 1996; 1:133–144 .
95. Ramirez-alvarado m, blanco fj, serrano l. De novo design and structural analysis of a model b-hairpin peptide system. *Nature struct biol* 1996; 206

3:604– 612.

96. Maynard aj, sharman gj, searle ms. Origin of beta-hairpin stability in solution: structural and thermodynamic analysis of the folding of a model peptide supports hydrophobic stabilization in water. *J am chem soc* 1998; 120:1996–2007.
97. Harding mm, williams dh, woolfson dn. Characterization of a partially denatured state of a protein by 2-dimensional nmr—reduction of the hydrophobic interaction in ubiquitin. *Biochemistry* 1991; 30:3120–3128.
98. Stockman bj, euvrard a, scahill ta. Heteronuclear 3-dimensional nmr spectroscopy of a partially denatured protein—the a-state of human ubiquitin. *J biomol nmr*, 1993; 3:285–296.
99. Sibanda bl, thornton jm. Conformation of b-hairpins in protein structures—classification and diversity in homologous structures. *Methods enzymol* 1991; 202:59–82.
100. Hutchinson eg, thornton jm. A revised set of potentials for b-turn formation in proteins. *Protein sci* 1994; 3:2207–2216.
101. Jang, soonmin, seokmin shin, and youngshang pak. "molecular dynamics study of peptides in implicit water: ab initio folding of β -hairpin, β -sheet, and $\beta\beta\alpha$ -motif." *Journal of the american chemical society* 124, no. 18 2002; 4976-4977.
102. Ulmschneider, jakob p., and william l. Jorgensen. "polypeptide folding using monte carlo sampling, concerted rotation, and continuum solvation." *Journal of the american chemical society* 126, no. 6 2004; 1849-1857.
103. Jorgensen, william l., and julian tirado–rives. "molecular modeling of organic and biomolecular systems using boss and mcpro." *Journal of computational chemistry* 26, no. 16 2005; 1689-1700.
104. Jorgensen, william l., David s. Maxwell, and julian tirado-rives. "development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids." *Journal of*
207

- the american chemical society 118, no. 45 1996; 11225-11236.
105. Ulmschneider, jakob p., and william l. Jorgensen. "polypeptide folding using monte carlo sampling, concerted rotation, and continuum solvation." *Journal of the american chemical society* 126, no. 6 2004; 1849-1857.
106. Jang, soonmin, eunae kim, and youngshang pak. "direct folding simulation of α -helices and β -hairpins based on a single all-atom force field with an implicit solvation model." *Proteins: structure, function, and bioinformatics* 66, no. 1 2007; 53-60.
107. Andersen, hans c. "rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations." *Journal of computational physics* 52, no. 1 1983; 24-34.
108. Palmer bj. Direct application of shake to the velocity verlet algorithm. *J comput phys* 1993; 104:470–472.
109. Ponder jw. *Tinker 4.1: software tools for molecular design*. St. Louis, mo: washington university; 2003.
110. Balaraman, gouthaman s., In-hee park, abhinandan jain, and nagarajan vaidehi. "folding of small proteins using constrained molecular dynamics." *The journal of physical chemistry b* 115, no. 23 2011; 7588-7596.
111. S. A. Adcock and j. A. Mccammon, "molecular dynamics: survey of methods for simulating the activity of proteins", *chem. Rev.* 2006; 106: 1589–1615 .
- A. Liwo, c. Czaplewski, s. Oldziej, and h. A. Scheraga. Computational techniques for efficient conformational sampling of proteins. *Curr. Opin. Struct. Biol.*, 2008.; 18:134–139.
112. E.i. Shakhnovich and a.m. Gutin. Implications of thermodynamics on protein folding for evolution of primary sequences. *Nature*, 1990; 346:773–775.
- A. D. Mackerell, jr. Empirical force fields for biological

- macromolecules: overview and issues. *J. Comput. Chem.*, 2004; 25:1584–1604.
113. J. T. Sprague, j. C. Tai, y. Yuh, and n. L. Allinger. The mmp2 calculational method. *J. Comput. Chem.*, 1987; 8:581–603.
114. L. S. D. Caves, j. D. Evanseck, and m. Karplus. Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin. *Prot. Sci.*, 1998; 7:649–666.
- A. D. Mackerell, jr., M. Feig, and c. L. Brooks, iii. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.*, 2004; 25:1400–1415.
115. F. A. Momany, r. F. Mcguire, a. W. Burgess, and h. A. Scheraga. Energy parameters in polypeptides. Vii. Geometric parameters partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.*, 1975; 79:2361–2381.
116. S. Lifson and a. Warshel. Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and n-alkane molecules. *J. Chem. Phys.*, 1968; 49:5116–5129.
117. Weiner, paul k.; Kollman, peter a. Amber: assisted model building with energy refinement. A general program for modeling molecules and their interactions. *Journal of computational chemistry*, 1981; 2.3: 287-303.
118. Jorgensen, william I. Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *Journal of the american chemical society*, 1981; 103.2: 335-340.
119. Hagler, a. T.; Huler, e.; Lifson, s. Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals. *Journal of the american chemical society*, 1974;

- 96.17: 5319-5327.
120. W. D. Cornell, p. Cieplak, c. I. Bayly, i. R. Gould, k. M. Merz, jr., D. M. Ferguson, d. C. Spellmeyer, t. Fox, j. W. Caldwell, and p. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Amer. Chem. Soc.*, 1995; 117:5179–5197.
 121. Kollman, peter, et al. The development/application of a ‘minimalist’ organic/biochemical molecular mechanic force field using a combination of ab initio calculations and experimental data. In: *computer simulation of biomolecular systems*. Springer netherlands, 1997; 83-96.
 122. Wang, junmei, piotr cieplak, and peter a. Kollman. "how well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules?." *Journal of computational chemistry* 21.12 2000; 1049-1074.
 123. Ponder, jay w., and david a. Case. "force fields for protein simulations." *Advances in protein chemistry* 66 2003; 27-86.
 124. Hornak, viktor, et al. "comparison of multiple amber force fields and development of improved protein backbone parameters." *Proteins: structure, function, and bioinformatics* 65.3 2006; 712-725.
 125. Duan, yong, et al. "a point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations." *Journal of computational chemistry* 24.16 2003; 1999-2012.
 126. Shaw, david e., et al. "atomic-level characterization of the structural dynamics of proteins." *Science* 330.6002 2010; 341-346.
 127. Lindorff-larsen, kresten, et al. "improved side-chain torsion potentials for the amber ff99sb protein force field." *Proteins: structure, function, and bioinformatics* 78.8 2010; 1950-1958.
 128. B. J. Alder and t. E. Wainwright. *Studies in molecular dynamics. I. General method*. *J. Chem. Phys.*, 1959; 31:459–466.

129. J. A. Mccammon, b. R. Gelin, and m karplus. Dynamics of folded proteins. *Nature*, 1977; 267:585–590.
130. Y. Duan, p. A. Kollman, and s. C. Harvey. Protein folding and beyond. In e. Keinan and i. Schechter, editors, *chemistry for the 21st century*. Wiley-vch, weinheim, germany, 2000.
131. Y. Duan, l. Wang, and p. A. Kollman. The early stage of folding of villin headpiece subdomain observed in a 200-nanosecond fully solvate molecular dynamics simulation. *Proc. Natl. Acad. Sci. Usa*, 1998; 95:9897–9902.
132. Salomon-ferrer, romelia, et al. "routine microsecond molecular dynamics simulations with amber on gpus. 2. Explicit solvent particle mesh ewald." *Journal of chemical theory and computation* 9.9 2013; 3878-3888.
133. Piana, stefano, kresten lindorff-larsen, and david e. Shaw. "atomic-level description of ubiquitin folding." *Proceedings of the national academy of sciences* 110.15 2013; 5915-5920.
134. Zerella, rosa, et al. "structural characterization of a mutant peptide derived from ubiquitin: implications for protein folding." *Protein science* 9.11 2000; 2142-2150.
135. Phillips, james c., et al. "scalable molecular dynamics with namd." *Journal of computational chemistry* 26.16 2005; 1781-1802.
136. Humphrey, william, andrew dalke, and klaus schulten. "vmd: visual molecular dynamics." *Journal of molecular graphics* 14.1 1996; 33-38.
137. Glykos, nicholas m. "software news and updates carma: a molecular dynamics analysis program." *Journal of computational chemistry* 27.14 2006; 1765-1768.
138. Frishman, dmitrij, and patrick argos. "knowledge-based protein secondary structure assignment." *Proteins: structure, function, and bioinformatics* 23.4 1995; 566-579.
139. Heinig, matthias, and dmitrij frishman. "stride: a web server for

- secondary structure assignment from known atomic coordinates of proteins." *Nucleic acids research* 32.suppl 2 2004; w500-w502.
140. Kabsch, wolfgang, and christian sander. "dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features." *Biopolymers* 22.12 1983; 2577-2637.
 141. Gordon, heather l., and rajmund l. Somorjai. "fuzzy cluster analysis of molecular dynamics trajectories." *Proteins: structure, function, and bioinformatics* 14.2 1992; 249-264.
 142. Jolliffe it. *Principal component analysis*. Springer; new york: 2002.
 143. Altis, alexandros, et al. "dihedral angle principal component analysis of molecular dynamics simulations." *The journal of chemical physics* 126.24 2007; 244111.
 144. Ichiye, toshiko, and martin karplus. "collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations." *Proteins: structure, function, and bioinformatics* 11.3 1991; 205-217.
 145. García, Angel E. "Large-amplitude nonlinear motions in proteins." *Physical review letters* 68.17 1992; 2696.
 146. Amadei, Andrea, Antonius Linssen, and Herman JC Berendsen. "Essential dynamics of proteins." *Proteins: Structure, Function, and Bioinformatics* 17.4 1993; 412-425.
 147. Hayward, S., et al. "Effect of solvent on collective motions in globular protein." *Journal of molecular biology* 234.4 1993; 1207-1217.
 148. Becker, Oren M. "Geometric versus topological clustering: an insight into conformation mapping." *Proteins: Structure, Function, and Bioinformatics* 27.2 1997; 213-226.
 149. Lange, Oliver F., and Helmut Grubmüller. "Can principal components yield a dimension reduced description of protein dynamics on long time scales?." *The Journal of Physical Chemistry B* 110.45 2006; 22842-22852.

150. Noé, Frank, et al. "Transition networks for the comprehensive characterization of complex conformational change in proteins." *Journal of Chemical Theory and Computation* 2.3 2006; 840-857.
151. Mu, Yuguang, Phuong H. Nguyen, and Gerhard Stock. "Energy landscape of a small peptide revealed by dihedral angle principal component analysis." *Proteins: Structure, Function, and Bioinformatics* 58.1 2005; 45-52.
152. Harder, Tim, et al. "Beyond rotamers: a generative, probabilistic model of side chains in proteins." *BMC bioinformatics* 11.1 2010; 306.
153. Schlick, Tamar. *Molecular Modeling and Simulation: An Interdisciplinary Guide: An Interdisciplinary Guide*. Vol. 21. Springer, 2010.
154. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, et al. Systematic Validation of Protein Force Fields against Experimental Data. *PloS ONE* 7,2 2012; e32131.
155. Ganoth, Assaf, Yossi Tsfadia, and Reuven Wiener. "Ubiquitin: Molecular modeling and simulations." *Journal of Molecular Graphics and Modelling* 46 2013; 29-40.
156. <http://psfs.cbrc.jp/fold-rate>
157. <http://predictioncenter.org>
158. <http://aps.unmc.edu/AP/main.php>
159. <http://pepbank.mgh.harvard.edu>
160. <http://amber.scripps.edu>
161. <http://www.charmm.org>
162. <http://www.gromacs.org>
163. <http://www.ks.uiuc.edu/Research/namd>
164. <http://dasher.wustl.edu/tinker>
165. http://www.deshawresearch.com/resources_desmond.html

It

is a brave new world, moving from our mechanical world, born in the industrial revolution, to the biotechnological future. The gears of life are complex, encoded polymers, nano machines that create order out of disorder and harness the free energy of the sun to perpetuate the information they carry in to the unknown. Now, man begins to alter that very microscopic machinery that constitutes his existence, that gives rise to his consciousness out of inorganic matter, ever driven forward by his curiosity and desire to manipulate his surroundings and himself.

-Mike Tika-

"Angel of Death" - Ubiquitin