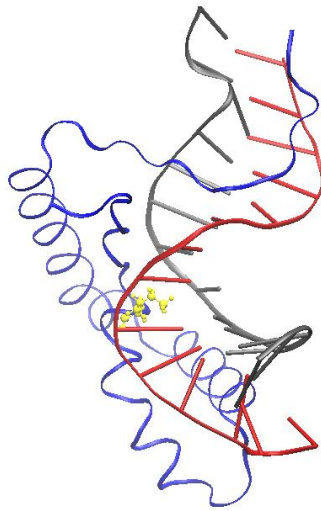




ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΤΕΧΝΙΚΗΣ ΤΩΝ ΥΛΙΚΩΝ

Υπολογιστική Διερεύνηση του Μηχανισμού της
Έκφρασης του Γονιδίου *SOX9*, κατά το Πρώτο Στάδιο
της Μεταγραφής του μέσω του Μεταγραφικού
Παράγοντα SRY



Διπλωματική Εργασία
Αγγελική - Γεωργία Νάσου

Επιβλέπων Καθηγητής
Γεώργιος Κ. Παπαδόπουλος

Αθήνα 2021

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε στο εργαστήριο της ερευνητικής ομάδας ENGIMATER του Τομέα Επιστήμης και Τεχνικής των Υλικών της Σχολής Χημικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου, υπό την επίβλεψη του κ. Γεώργιου Παπαδόπουλου. Επί τη ευκαιρία της ολοκλήρωσής της, θα ήθελα να ευχαριστήσω τους ανθρώπους που συντέλεσαν στην πραγματοποίησή της.

Αρχικά, θα ήθελα να ευχαριστήσω τον Αναπληρωτή Καθηγητή του Ε.Μ.Π. κύριο Γεώργιο Παπαδόπουλο, για την εμπιστοσύνη που μου έδειξε με την ανάθεση του συγκεκριμένου θέματος. Τον ευχαριστώ θερμά για την επίβλεψη και την επιστημονική του υποστήριξη, η οποία ήταν καθοριστική για την πραγματοποίηση αυτού του έργου, και για το ειλικρινές ενδιαφέρον του για την πρόοδο μου και την κατανόηση των όσων κλήθηκα να μάθω στα πλαίσια της εκπόνησης του θέματος.

Οφείλω ένα μεγάλο ευχαριστώ στην Ευαγγελία Παντατοσάκη, μεταδιδακτορική ερευνήτρια στο Εργαστήριο ENGIMATER, για τη στήριξη και την καθοδήγηση της καθ' όλη την διάρκεια της εκπόνησης αυτής της διπλωματικής. Επίσης, για το ενδιαφέρον και την υπομονή με την οποία με εισήγαγε στο αντικείμενο της εργασίας, το σύνολο των γνώσεων που μου μετέδωσε και την επιστημονική εμπειρία που μοιράστηκε μαζί μου. Την ευχαριστώ για τον ενθουσιασμό που μου ενέπνευσε μέσα από το δικό της ζήλο για το αντικείμενο.

Ολοκληρώνοντας τον κύκλο των προπτυχιακών μου σπουδών, θα ήθελα να ευχαριστήσω την οικογένειά μου, για την στήριξη που μου έχει προσφέρει όλα αυτά τα χρόνια, και ιδιαίτερα την αδερφή μου, Ελπίδα, η οποία πιστεύει σε μένα, με συμβουλεύει και με ενθαρρύνει σε κάθε μου βήμα. Τέλος, θα ήθελα να ευχαριστήσω τους φίλους και τους συμφοιτητές μου στους οποίους οφείλονται όλες οι όμορφες στιγμές και αναμνήσεις των φοιτητικών μου χρόνων.

Αγγελική - Γεωργία Νάσου

Μάρτιος 2021

ΠΕΡΙΛΗΨΗ

Στον κυτταρικό πυρήνα, οι διεργασίες που εμπλέκονται στη μεταγραφή των γονιδίων κατά το πρώτο στάδιο της γονιδιακής έκφρασης, περιλαμβάνουν το σχηματισμό μοριακών συμπλεγμάτων, τα οποία απαρτίζονται από πρωτεϊνικά μόρια γνωστά ως μεταγραφικοί ρυθμιστές. Η απορρύθμιση των ανωτέρω διεργασιών μεταγραφικής ρύθμισης έχει σχετιστεί με πλήθος νευροεκφυλιστικών ασθενειών, γενετικών διαταραχών, καθώς και την υπερέκφραση ογκογονιδίων. Ο μηχανισμός που διέπει τα φαινόμενα αυτά δεν έχει ακόμη αποσαφηνιστεί, ωστόσο, η κατανόησή του δύναται να προσφέρει νέα προοπτική στην αντιμετώπιση σοβαρότατων ασθενειών, ανοίγοντας το δρόμο σε νέες θεραπευτικές προσεγγίσεις και τεχνικές της μοριακής ιατρικής. Ο μεταγραφικός παράγοντας SRY διαδραματίζει κρίσιμο ρόλο στην ανάπτυξη του άρρενος φύλου κατά την εμβρυική περίοδο της κύησης. Η λειτουργική δράση του γονιδίου *SOX9*, τη μεταγραφή του οποίου ενεργοποιεί η πρόσδεση του SRY, αποτελεί κύριο μοριακό διακόπτη στον φυλοκαθορισμό των θηλαστικών, ενώ η ανεπαρκής ή ετεροχρονισμένη μεταγραφή του δύναται να προκαλέσει φυλοσύνδετες διαταραχές. Σκοπό της παρούσας Διπλωματικής Εργασίας αποτελεί η διερεύνηση του μηχανισμού της έκφρασης του γονιδίου *SOX9*, κατά το πρώτο στάδιο της μεταγραφής του, μέσω της πρόσδεσης σε αυτό του μεταγραφικού παράγοντα SRY. Πραγματοποιήθηκαν προσομοιώσεις Μοριακής Δυναμικής βασισμένες στη στατιστική μηχανική και θερμοδυναμική, σε δύο συστήματα: (i) το γονίδιο *SOX9* συνδεδεμένο με τον μεταγραφικό παράγοντα SRY, και (ii) το γονίδιο *SOX9* ελεύθερο, εμβαπτισμένα σε ιοντικά υδατικά διαλύματα. Η επεξεργασία των παραγόμενων τροχιών (αλληλουχία μικροκαταστάσεων) για τα δύο συστήματα έδειξε ότι κατά την πρόσδεση του SRY αποβάλλονται μόρια νερού από το κέλυφος ενυδάτωσης του γονιδίου (DNA) και ο SRY προσεγγίζει το DNA μέσω των πλευρικών αλυσίδων των αμινοξέων που κείνται κυρίως στην πρώτη και τη δεύτερη α -έλικα του. Η πρόσδεση του SRY προκαλεί διεύρυνση της μικρής αύλακας του DNA, με ταυτόχρονη ελάττωση του εύρους της μεγάλης αύλακας, που έχει σαν αποτέλεσμα την κάμψη του DNA προς την πλευρά της μεγάλης αύλακας. Επίσης στην περιοχή της αλυσίδας του DNA όπου πραγματοποιείται η πρόσδεση παρατηρείται αύξηση της απόστασης μεταξύ των γειτονικών βάσεων του DNA, με ταυτόχρονη παρεμβολή του αμινοξέος ισολευκίνης, το οποίο κείται στην πρώτη α -έλικα του SRY. Τα αποτελέσματα αυτά βρίσκονται σε συμφωνία με *in vitro* πειραματικές μελέτες από τη βιβλιογραφία. Επιπροσθέτως, τα αποτελέσματα δείχνουν ότι η παρεμβολή της ισολευκίνης γίνεται μεταξύ δύο διαδοχικών βάσεων αδενίνης και χαρακτηρίζεται ως μερική, καθώς μόνο τα άτομα της πλευρικής αλυσίδας της ισολευκίνης εντοπίστηκαν να εισχωρούν στο εσωτερικό της διπλής έλικας. Η αποσαφήνιση του ανωτέρω μηχανισμού πρόσδεσης παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς σε αντίθεση με την πλειονότητα των πρωτεϊνών οι οποίες προσδέονται στη μεγάλη αύλακα του DNA, ο μεταγραφικός παράγοντας SRY εκκινεί τη μεταγραφή του γονιδίου *SOX9* μέσω της πρόσδεσής του στη μικρή αύλακα του DNA.

ABSTRACT

Transcription factors (TF) are proteins that play a prominent role in transcriptional regulation of gene expression and cellular differentiation. Their action has been linked to the formation of multi-molecular assemblies. Misregulation of transcriptional processes has been related to neurodegenerative diseases, genetic disorders and the expression of oncogenes. The molecular mechanism underlying the above phenomena remains unknown, and its understanding may offer new therapeutic routes to the treatment of severe diseases via molecular medicine. The objective of this thesis is the study of the mechanism of binding of the transcription factor SRY to the gene *SOX9*, by means of Molecular Dynamics (MD) full-atom simulations. For this, we study (i) the TF, SRY, bound to the *SOX9* gene, and (ii) the gene *SOX9* free; both systems were modeled in an ionic aqueous solution environment. The geometrical features and water shell of the bound and unbound gene (DNA) were studied, as well as the distances between the atoms of SRY and *SOX9*. The MD trajectory post-processing analysis shows that, during the SRY – *SOX9* binding, water molecules are discarded from the DNA water shell and SRY approaches *SOX9* mostly through the sidechains of the amino acids belonging to the first and second α -helix. Furthermore, the MD simulations revealed that the minor groove of the DNA is widened and the major groove is compressed as a result of the SRY – *SOX9* binding; also, the DNA is bent in the direction of the major groove showing an increase in the rise per base pair at the binding site, as a result of isoleucine intercalation in the DNA double helix. These results are in agreement with the results of in vitro studies found elsewhere. In addition, the MD shows that the isoleucine intercalation occurs between two adjacent adenine bases and is characterized as partial intercalation, since only the sidechain atoms of isoleucine were found to enter in the interior of the double helix. The elucidation of the above binding mechanism is of particular interest, since, in contrast to majority of proteins binding to the major groove, the SRY initiates the transcription of the *SOX9* gene via minor groove binding to DNA.

ΠΕΡΙΕΧΟΜΕΝΑ

Ευχαριστίες	II
Περίληψη.....	III
Abstract.....	IV
I. ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ	1
1. Μεταγραφική Ρύθμιση.....	2
Εισαγωγή	2
1.1. Δομή δεοξυριβονουκλεϊκού οξέος.....	3
1.2. Η διεργασία της μεταγραφής.....	5
1.3. Διεργασίες μεταγραφικής ρύθμισης και μόρια-ρυθμιστές	6
1.4. Μεταγραφική ρύθμιση και νέες θεραπευτικές οδοί	7
1.5. Μεταγραφικοί παράγοντες	8
1.5.1. Δομή πρωτεϊνών.....	8
1.5.2. Δομή μεταγραφικών παραγόντων	11
1.6. Ο μεταγραφικός παράγοντας SRY.....	12
1.7. Ο μεταγραφικός παράγοντας SRY και η διαμόρφωση του φύλου στον άνθρωπο	13
1.8. Μεταγραφικός παράγοντας SRY και φυλοσύνδετες διαταραχές	15
2. Προσομοιώσεις μοριακής δυναμικής.....	16
Εισαγωγή	16
2.1. Βασικές έννοιες στις προσομοιώσεις Μοριακής Δυναμικής	18
2.1.1. Η περιγραφή της τροχιάς στην κλασική μηχανική.....	18
2.1.2. Περιγραφή μοριακών συστημάτων και δυναμικά αλληλεπίδρασης.....	19
2.1.3. Μελέτη μικρών συστημάτων	23
2.2. Κλασική Στατιστική Μηχανική	24
2.2.1. Στατιστικά Σύνολα Gibbs.....	24
2.2.2. Συνήθη στατιστικά σύνολα	28
2.3. Αλγόριθμοι μοριακών προσομοιώσεων	32
2.3.1. Περιγραφή αλγορίθμων Μοριακής Δυναμικής.....	32
2.3.2. Ο ολοκληρωτής Verlet.....	34
2.3.3. Ο ολοκληρωτής leap-frog	35

2.3.4. Ο αλγόριθμος Berendsen	36
2.3.5. Ο αλγόριθμός Steepest Descent	41
II. ΥΠΟΛΟΓΙΣΤΙΚΟ ΜΕΡΟΣ	42
3. Δημιουργία συστημάτων μοριακής προσομοίωσης	43
3.1. Μοντελοποίηση μακρομορίων.....	43
3.2. Αρχική απεικόνιση συμπλέγματος SRY-DNA.....	43
3.2.1. Ψηφιακή ανακατασκευή ελεύθερου DNA.....	44
3.3. Πεδίο δυνάμεων και παράμετροι μοριακής δυναμικής	47
4. Υπολογιστική επίδοση μοριακών προσομοιώσεων (Performance)	52
5. Αποτελέσματα Εξισορρόπησης συστημάτων προσομοίωσης	53
5.1. Εξισορρόπηση συστήματος DNA – SRY σε υδατικό διάλυμα.....	53
5.2. Εξισορρόπηση συστήματος DNA σε υδατικό διάλυμα	57
6. Αποτελέσματα προσομοιώσεων μοριακής δυναμικής.....	62
6.1. Εισαγωγή	62
6.2. Κέλυφος ενυδάτωσης μακρομορίων	67
6.3. Διαμόρφωση DNA.....	69
6.4. Γεωμετρικά χαρακτηριστικά της διπλής έλικας ελεύθερου και συμπλεγμένου DNA.....	71
6.5. Ανάλυση αποστάσεων και επαφών του συμπλέγματος DNA - SRY.....	80
6.6. Νουκλεοτίδια αδενίνης και πρόσδεση της SRY	88
6.7. Μερική παρεμβολή της ισολευκίνης της SRY στη διπλή έλικα του DNA	93
7. Σύνοψη και συμπεράσματα.....	101
8. Βιβλιογραφία	106

I. ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ

1. ΜΕΤΑΓΡΑΦΙΚΗ ΡΥΘΜΙΣΗ

ΕΙΣΑΓΩΓΗ

Η κυτταρική διαφοροποίηση, η διαδικασία με την οποία τα κύτταρα αποκτούν εξειδίκευση ώστε να επιτελούν διαφορετικές λειτουργίες στον οργανισμό, αποτελεί ένα από τα μεγαλύτερα μυστήρια της ζωής. Όλα τα κύτταρα ενός οργανισμού περιέχουν το ίδιο γενετικό υλικό, ωστόσο, σε κάθε ιστό εκφράζονται διαφορετικά γονίδια, τα οποία καθορίζουν τη μορφή και τη λειτουργία τους. Για την επίτευξη αυτής της διαφοροποίησης, το κύτταρο διαθέτει μηχανισμούς ρύθμισης, οι οποίοι καθορίζουν ποιοτικά και ποσοτικά τη γονιδιακή έκφραση.

Η ρύθμιση της γονιδιακής έκφρασης πραγματοποιείται από το κύτταρο με διαφορετικούς μηχανισμούς που λαμβάνουν χώρα, τόσο εντός του κυτταρικού πυρήνα, όσο και εκτός πυρήνα, στο κυτταρόπλασμα. Ένας από τους μηχανισμούς ρύθμισης της γονιδιακής έκφρασης εντός του κυτταρικού πυρήνα είναι η αυτό-συγκρότηση του συμπλέγματος της μεταγραφής το οποίο αποτελείται από πλήθος βιομορίων. Δυσλειτουργίες στο μηχανισμό σχηματισμού αυτού του συμπλέγματος μπορεί να προκαλέσουν σοβαρές ασθένειες, όπως είναι οι νευροεκφυλιστικές διαταραχές και ο καρκίνος.

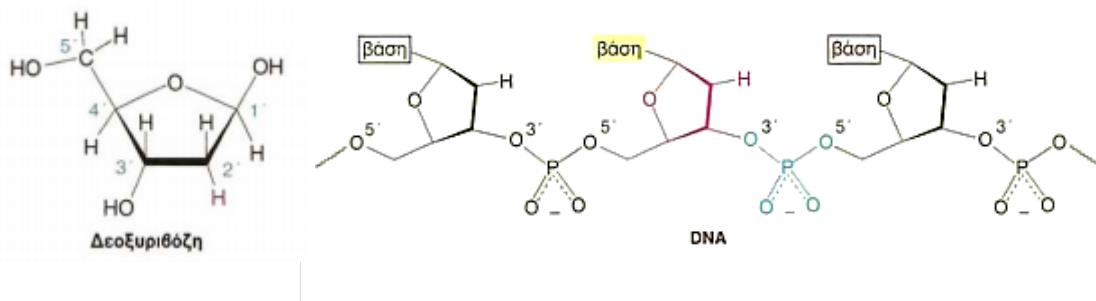
Ο μικροσκοπικός μηχανισμός σχηματισμού του συμπλέγματος έναρξης της μεταγραφής, καθώς και οι παράμετροι που τον επηρεάζουν, παραμένει ένα ανεξερεύνητο πεδίο. Σκοπός της παρούσας διπλωματικής είναι η διερεύνηση του μικροσκοπικού φυσικοχημικού μηχανισμού που οδηγεί στο σχηματισμό του, μέσω μοριακής υπολογιστικής προσομοίωσης, βασισμένης στη στατιστική μηχανική, του συστήματος των βιολογικών μορίων που συμμετέχουν σε αυτό .

Η αποσαφήνιση του μηχανισμού αυτό-συγκρότησης τους συμπλέγματος έναρξης της μεταγραφής και τρόπου με τον οποίο θα μπορούσε να παρεμποδιστεί ο ελαττωματικός σχηματισμός του μπορεί να οδηγήσει σε νέες θεραπευτικές οδούς και το σχεδιασμό αποτελεσματικότερων θεραπειών.

1.1. ΔΟΜΗ ΔΕΟΞΥΡΙΒΟΝΟΥΚΛΕΪΚΟΥ ΟΞΕΟΣ

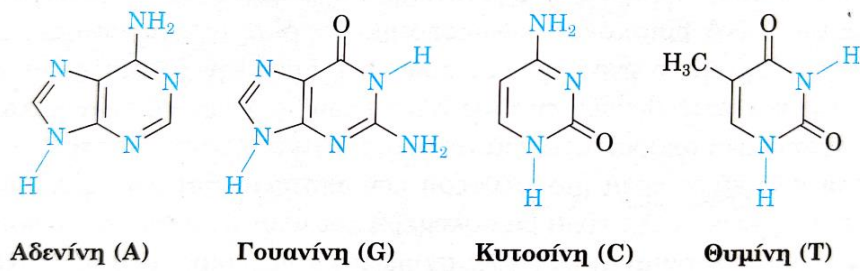
Τα νουκλεϊκά οξέα, δηλαδή το δεοξυριβονουκλεϊκό οξύ (DNA) και το ριβονουκλεϊκό οξύ (RNA), συνιστούν γραμμικά πολυμερή μεγάλου μεγέθους. Αποτελούν μόρια-φορείς της γενετικής πληροφορίας εντός του κυττάρου. Στα μόρια του DNA βρίσκονται κωδικοποιημένες όλες οι πληροφορίες που αφορούν στον προσδιορισμό της φύσης του κυττάρου, τον έλεγχο της ανάπτυξης και της διαίρεσης του, τη βιοσύνθεση των ενζύμων και των πρωτεϊνών που είναι απαραίτητες για τις φυσιολογικές λειτουργίες του κυττάρου.

Τα νουκλεϊκά οξέα είναι βιοπολυμερή τα οποία απαρτίζονται από **νουκλεοτίδια**. Τα νουκλεοτίδια ενώνονται σχηματίζοντας μακρές αλυσίδες. Κάθε νουκλεοτίδιο αποτελείται από ένα σάκχαρο, μια φωσφορική ομάδα και μια ετεροκυκλική βάση. Στο δεοξυριβονουκλεϊκό οξύ το σάκχαρο είναι η 2'-δεοξυριβόζη (Εικόνα 1.1). Τα σάκχαρα συνδέονται με τις φωσφορικές ομάδες με φωσφοδιεστερικούς δεσμούς και αποτελούν τον **κορμό**, ο οποίος είναι κοινός (Εικόνα 1.1) για όλα τα νουκλεοτίδια. Στο ένα άκρο της αλυσίδας του νουκλεϊκού οξέος υπάρχει μια ελεύθερη υδροξυλομάδα στον C3' της δεοξυριβόζης, με αποτέλεσμα το άκρο να καλείται **άκρο 3'**, ενώ στο άλλο άκρο της αλυσίδας υπάρχει μια ελεύθερη φωσφορική ομάδα στον C5' του σακχάρου και καλείται **άκρο 5'**.^{1,2}



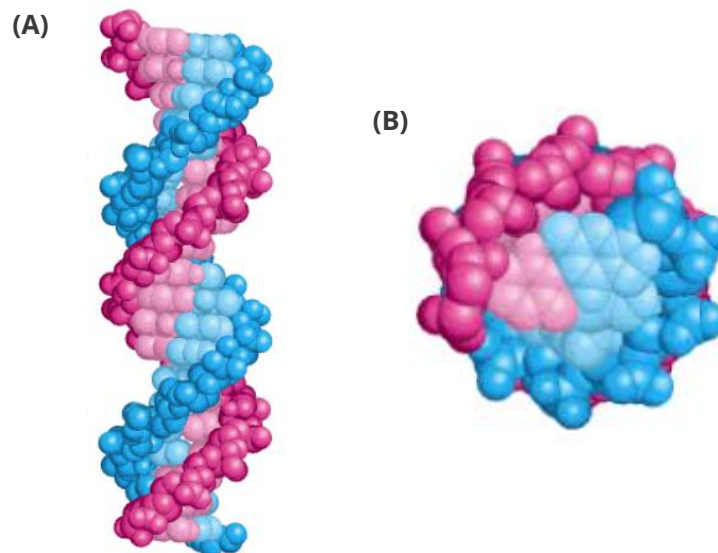
Εικόνα 1.1. (Αριστερά) Δομή δεοξυριβόζης. (Δεξιά) Ο κορμός του DNA. Ο κορμός σχηματίζεται με φωσφοδιεστερικό δεσμό 3→5. Τα σάκχαρα απεικονίζονται με κόκκινο χρώμα και η φωσφορική ομάδα με πράσινο.¹

Οι βάσεις του δεοξυριβονουκλεϊκού οξέος ποικίλουν μεταξύ τεσσάρων τύπων: Αδενίνη (A), Γουανίνη (G), Κυτοσίνη (C) και Θυμίνη (T) και είναι χαρακτηριστικές για το κάθε νουκλεοτίδιο. Η αδενίνη και η γουανίνη ονομάζονται πουρίνες, ενώ η κυτοσίνη και η θυμίνη ονομάζονται πυριμιδίνες. Η αλληλουχία των βάσεων στα νουκλεϊκά οξέα αποτελεί τον κώδικα της αποθήκευσης των γενετικών πληροφοριών. Η δομή των τεσσάρων βάσεων των νουκλεοτιδίων παρουσιάζεται στην Εικόνα 1.2.



Εικόνα 1.2. Δομές βάσεων DNA.²

Οι βάσεις του DNA έχουν την ικανότητα να σχηματίζουν εξειδικευμένα ζεύγη βάσεων τα οποία σταθεροποιούνται με δεσμούς υδρογόνου: η αδενίνη συνδέεται με τη θυμίνη με δύο δεσμούς υδρογόνου, ενώ η γουανίνη συνδέεται με την κυτοσίνη με τρεις δεσμούς υδρογόνου. Τα ζεύγη αυτά δίνουν στο DNA τη μορφή της **διπλής έλικας**, μιας ελικοειδούς δομής αποτελούμενης από δύο αλυσίδες. Σύμφωνα με το μοντέλο Watson-Crick για τη δομή του DNA, δύο ελικοειδείς πολυνουκλεοτιδικές αλυσίδες περιστρέφονται γύρω από ένα κοινό άξονα, έχοντας αντίθετες κατευθύνσεις. Οι φωσφορικές ομάδες και τα σάκχαρα βρίσκονται στο εξωτερικό μέρος ενώ οι βάσεις στο εσωτερικό της έλικας, με τοποθέτηση σχεδόν κάθετη στον άξονα της έλικας. Το μοντέλο της διπλής έλικας του DNA παρουσιάζεται στην Εικόνα 1.3.



Εικόνα 1.3. Το μοντέλο Watson-Crick της διπλής έλικας για το DNA. Η μία πολυνουκλεοτιδική αλυσίδα απεικονίζεται με χρώμα μπλε και η άλλη με κόκκινο. Οι βάσεις έχουν πιο ανοιχτό χρώμα από τον κορμό των φωσφορικών ομάδων και των σακχάρων. (A) Όψη κατά μήκος του άξονα της διπλής έλικας. (B) Όψη εγκάρσια στον άξονα της διπλής έλικας.¹

1.2. Η ΔΙΕΡΓΑΣΙΑ ΤΗΣ ΜΕΤΑΓΡΑΦΗΣ

Όπως προαναφέρθηκε, στο DNA του κυττάρου βρίσκονται αποθηκευμένες όλες οι πληροφορίες που είναι απαραίτητες για τη σύνθεση των πρωτεϊνών οι οποίες ρυθμίζουν τις λειτουργίες και καθορίζουν την ταυτότητα του κυττάρου, κωδικοποιημένες υπό τη μορφή αλληλουχιών βάσεων κατά μήκος της αλυσίδας του. Η αποκωδικοποίηση των πληροφοριών αυτών, ώστε να μπορούν να αξιοποιηθούν από το κύτταρο για τη σύνθεση των πρωτεϊνών πραγματοποιείται σε δύο βασικά στάδια: το μεταγραφικό στάδιο και το στάδιο της μετάφρασης. Κατά μεταγραφικό στάδιο, το οποίο συμβαίνει εντός του κυτταρικού πυρήνα, η πληροφορία που υπάρχει σε ένα γονίδιο στο DNA μεταγράφεται σε ένα αγγελιοφόρο μόριο (mRNA) το οποίο θα τη μεταφέρει εκτός του κυτταρικού πυρήνα. Κατά το στάδιο της μετάφρασης, το οποίο συμβαίνει εκτός του κυτταρικού πυρήνα, το αγγελιοφόρο μόριο με τη βοήθεια κυτταρικών οργανιδίων (ριβόσωμα) θα μεταφράσει την πληροφορία που φέρει προς το σχηματισμό συγκεκριμένης πρωτεΐνης. Μέσω αυτών των δύο βασικών διεργασιών πραγματοποιείται η γονιδιακή έκφραση, η ροή δηλαδή των γενετικών πληροφοριών και ο μετασχηματισμός των πληροφοριών του DNA σε λειτουργικά μόρια.

Αναλυτικότερα, κατά τη διεργασία της **μεταγραφής του DNA**, ένας από τους δύο κλώνους του DNA χρησιμοποιείται ως εκμαγείο με σκοπό την παραγωγή ενός μορίου ριβοζονουκλεϊκού οξέος (RNA) που ονομάζεται αγγελιοφόρο RNA (messenger RNA, **mRNA**). Το mRNA που παράγεται περιέχει την πληροφορία του μεταγραφόμενου γονιδίου κωδικοποιημένη στην αλληλουχία των βάσεων που περιέχει στη δομή του.

Η διεργασία της μεταγραφής πραγματοποιείται σε τρία στάδια. Το πρώτο και ιδιαίτερα σημαντικό στάδιο είναι η έναρξη της μεταγραφής. Η σημασία του συγκεκριμένου σταδίου έγκειται στο γεγονός ότι καθορίζει το είδος και τη συχνότητα των πρωτεϊνών που συνθέτει το κύτταρο. Η έναρξη της μεταγραφής γίνεται όταν ένα ένζυμο που ονομάζεται RNA πολυμεράση II συναντήσει μια συγκεκριμένη αλληλουχία βάσεων πάνω στο DNA η οποία καλείται **υποκινητής** (ή προαγωγέας, promoter) και συνδεθεί σε αυτή. Με τη σύνδεση RNA πολυμεράσης - υποκινητή προκαλείται το άνοιγμα της διπλής έλικας του DNA ώστε να ξεκινήσει η μεταγραφή.

Στους ευκαριωτικούς οργανισμούς η έναρξη της μεταγραφής είναι μια πολύπλοκη διαδικασία: για να μπορέσει η RNA πολυμεράση να ξεκινήσει τη μεταγραφή είναι απαραίτητη η παρουσία πολλών επικουρικών πρωτεϊνικών μορίων που ονομάζονται **γενικοί μεταγραφικοί παράγοντες** (General Transcription Factors, **GTFs**).³ Οι γενικοί μεταγραφικοί παράγοντες ή μεταγραφικοί παράγοντες προσδένονται στους υποκινητές πολλών γονιδίων, κοντά στην αφετηρία της μεταγραφής και βοηθούν στην τοποθέτηση της RNA πολυμεράσης στην κατάλληλη θέση, ώστε να ξεκινήσει η μεταγραφή.

Κατά το δεύτερο στάδιο της μεταγραφής, μια από τις δύο έλικες του ανοιγμένου πλέον DNA χρησιμοποιείται ως εκμαγείο για τη σύνθεση του mRNA. Με τη βοήθεια της RNA

πολυμεράσης τοποθετούνται συμπληρωματικά ριβονουκλεοτίδια στα νουκλεοτίδια της έλικας του DNA. Τα ριβονουκλεοτίδια αυτά σχηματίζουν το mRNA που φέρει την πληροφορία για τη δομή της συγκεκριμένης πρωτεΐνης.

Η μεταγραφή τερματίζει κατά το τρίτο στάδιο, όταν η RNA πολυμεράση συναντήσει συγκεκριμένη αλληλουχία νουκλεοτιδίων πάνω στο DNA, η οποία δίνει το σήμα για τερματισμό της μεταγραφής (αλληλουχία τερματισμού).³

1.3. ΔΙΕΡΓΑΣΙΕΣ ΜΕΤΑΓΡΑΦΙΚΗΣ ΡΥΘΜΙΣΗΣ ΚΑΙ ΜΟΡΙΑ-ΡΥΘΜΙΣΤΕΣ

Η ρύθμιση της έκφρασης των γονιδίων στο στάδιο της μεταγραφής αφορά στην επιλογή των γονιδίων και του χρόνου στον οποίο θα εκφραστούν. Πραγματοποιείται από ειδικά πρωτεϊνικά μόρια τα οποία ονομάζονται **μεταγραφικοί ρυθμιστές (transcriptional regulators)** με διεργασίες οι οποίες στο σύνολό τους απαρτίζουν τη διεργασία της μεταγραφικής ρύθμισης (transcriptional regulation).

Όπως προαναφέρθηκε, η έναρξη της μεταγραφής πραγματοποιείται στους υποκινητές των γονιδίων, καθώς αυτοί περιέχουν μια αλληλουχία νουκλεοτιδίων που σηματοδοτεί τη θέση έναρξης για τη σύνθεση του RNA, με την πρόσδεση σε αυτήν της RNA πολυμεράσης και των μεταγραφικών παραγόντων. Πέρα από τη θέση έναρξης στον υποκινητή, τα περισσότερα γονίδια διαθέτουν επιπρόσθετες ρυθμιστικές αλληλουχίες DNA (DNA regulatory sequences) για τη ρύθμιση της μεταγραφής τους. Οι αλληλουχίες αυτές δεν έχουν αυτοδύναμη δράση, αλλά η δράση τους εξαρτάται από την πρόσδεση σε αυτές ειδικών πρωτεϊνικών μορίων που ονομάζονται μεταγραφικοί ρυθμιστές. Οι προσδεδεμένοι μεταγραφικοί ρυθμιστές μπορεί να δράσουν είτε ενεργοποιητικά, προωθώντας τη μεταγραφή του γονιδίου (ενεργοποιητικές πρωτεΐνες ή ενεργοποιητές – activator proteins or activators) είτε κατασταλτικά (κατασταλτικές πρωτεΐνες ή καταστολείς – repressor proteins or repressors), αποτρέποντάς την.

Οι ρυθμιστικές αλληλουχίες του DNA στις οποίες προσδένονται οι ενεργοποιητικές πρωτεΐνες ονομάζονται ενισχυτικές ακολουθίες ή **ενισχυτές (enhancers)**. Οι ενισχυτές μπορεί να βρίσκονται πριν ή μετά το γονίδιο που μεταγράφεται, ή ακόμα και χιλιάδες ζεύγη νουκλεοτιδίων μακριά από αυτό. Έτσι, για να μπορέσει να ξεκινήσει η διαδικασία της μεταγραφής είναι πολλές φορές απαραίτητη η αναδίπλωση του DNA σε βρόχο, προκειμένου να έρθουν σε επαφή οι ενεργοποιητικές πρωτεΐνες που βρίσκονται προσδεδεμένες στους ενισχυτές με την RNA πολυμεράση ή τους μεταγραφικούς παράγοντες που την πλαισιώνουν. Η σύνδεση αυτή συχνά υποβοηθείται από ομάδες άλλων ειδικών πρωτεϊνών που καλούνται **συνενεργοποιητές (coactivators)**, η σημαντικότερη εκ των οποίων είναι ένα σύμπλεγμα που ονομάζεται Mediator.⁴

Συνεπώς, προκειμένου να ξεκινήσει η μεταγραφή είναι συχνά απαραίτητη η παρουσία και η συνεργατική δράση της RNA πολυμεράσης και των γενικών μεταγραφικών

παραγόντων στον υποκινητή, των ενεργοποιητών στους ενισχυτές και των συνενεργοποιητών. Το σύμπλεγμα μορίων που σχηματίζεται τελικά απαρτίζεται από πλήθος μεταγραφικών παραγόντων, συνενεργοποιητών και άλλων συμπαραγόντων, η δράση των οποίων χαρακτηρίζεται από συνεργασία και συνέργεια.⁵ Τα παραπάνω αποτελούν απαραίτητα στοιχεία για τη σύνθεση του **συμπλέγματος έναρξης της μεταγραφής** (transcription initiation complex).³

1.4. ΜΕΤΑΓΡΑΦΙΚΗ ΡΥΘΜΙΣΗ ΚΑΙ ΝΕΕΣ ΘΕΡΑΠΕΥΤΙΚΕΣ ΟΔΟΙ

Η απορρύθμιση των διεργασιών που εμπλέκονται στη μεταγραφή των γονιδίων έχει σχετιστεί με ποικίλες σοβαρές ασθένειες, όπως οι νευροεκφυλιστικές ασθένειες και ο καρκίνος.^{6,7} Μεταλλάξεις σε κύτταρα οδηγούν στη δημιουργία δομικά παραλλαγμένων πρωτεϊνών (γενικών μεταγραφικών παραγόντων ή μεταγραφικών ρυθμιστών), επηρεάζοντας έτσι την πρόσδεσή τους με το DNA ή/και με άλλες πρωτεΐνες (ενεργοποιητές), με αποτέλεσμα την ελαττωματική ρύθμιση της έκφρασης γονιδίων.^{6,11} Οι μεταλλάξεις στις πρωτεΐνες που οδηγούν σε απορρύθμιση των διεργασιών που εμπλέκονται στη μεταγραφή, συνδέονται με πλήθος νευροεκφυλιστικών ασθενειών όπως η Αμυοτροφική Πλευρική Σκλήρυνση (amyotrophic lateral sclerosis) και η Μετωποκροταφική Άνοια (frontotemporal dementia).^{4,7} Συνδέονται ακόμη και με λοιμώδεις ασθένειες, καθώς υποστηρίζεται ότι κάποιοι ιοί χρησιμοποιούν μεταγραφικά μονοπάτια στα κύτταρα των ξενιστών τους για να διεκπεραιώσουν τους αναπαραγωγικούς τους κύκλους.⁷

Επιπλέον, η απορρύθμιση των διεργασιών που εμπλέκονται στη μεταγραφή, έχει συνδεθεί με την υπερέκφραση ογκογονιδίων, ενώ οι ιδιαίτερες φυσικοχημικές αλληλεπιδράσεις μεταξύ των βιομορίων που εμπλέκονται στη μεταγραφή επηρεάζουν επίσης τις φαρμακοκινητικές ιδιότητες ορισμένων χορηγούμενων αντινεοπλασματικών φαρμάκων, όπως είναι η cis-πλατίνη (cisplatin) και η ταμοξιφαίνη (tamoxifen), τα οποία δρουν στοχεύοντας σε γονίδια – στόχους εντός του κυτταρικού πυρήνα.^{8,9}

Η κατανόηση των μηχανισμών που εμπλέκονται στις διεργασίες της μεταγραφικής ρύθμισης και της δημιουργίας των βιομοριακών συμπλεγμάτων, καθώς και του τρόπου με τον οποίο θα μπορούσε να παρεμποδιστεί ο ελαττωματικός σχηματισμός τους μπορεί να οδηγήσει σε νέες θεραπευτικές οδούς. Επιπλέον, η αποσαφήνιση των μηχανισμών με τους οποίους επιδρούν τα φαρμακευτικά μόρια στα βιομοριακά συμπλέγματα εντός του κυτταρικού πυρήνα θα επιτρέψει τον σχεδιασμό αποτελεσματικότερων θεραπειών.

1.5. ΜΕΤΑΓΡΑΦΙΚΟΙ ΠΑΡΑΓΟΝΤΕΣ

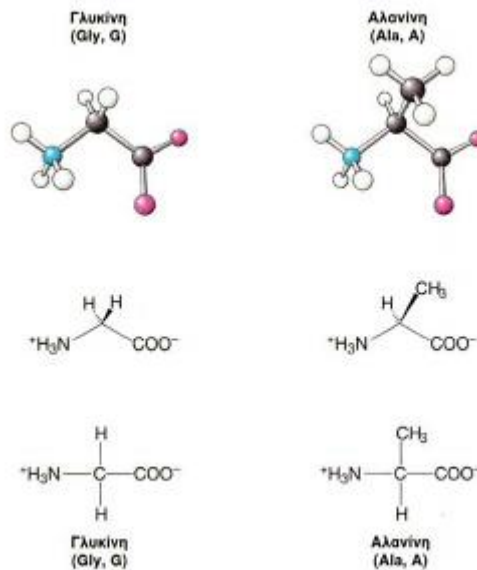
1.5.1. ΔΟΜΗ ΠΡΩΤΕΪΝΩΝ

Οι πρωτεΐνες είναι βιολογικά μακρομόρια που εξυπηρετούν κρίσιμες λειτουργίες στους ζωντανούς οργανισμούς: καταλύουν διεργασίες, μεταφέρουν και αποθηκεύουν άλλα μόρια, παρέχουν μηχανική στήριξη και προστασία του ανοσοποιητικού, μεταδίδουν νευρικές ώσεις και ελέγχουν την ανάπτυξη και τη διαφοροποίηση του κυττάρου. Αποτελούν γραμμικά πολυμερή τα οποία απαρτίζονται από μονομερή που ονομάζονται **αμινοξέα**.¹

Τα αμινοξέα φέρουν δύο λειτουργικές ομάδες: τη βασική αμινομάδα (-NH₂) και την όξινη καρβοξυλομάδα (-COOH). Κάθε αμινοξύ αποτελείται από ένα κεντρικό άτομο άνθρακα το οποίο ονομάζεται **α-άνθρακας** (α carbon, **C_α**), στον οποίο συνδέονται η αμινομάδα και η καρβοξυλομάδα. Επιπλέον, στον α-άνθρακα συνδέεται ένα άτομο Η και μια **πλευρική αλυσίδα** (side chain ή R group), η οποία είναι χαρακτηριστική για το κάθε αμινοξύ.¹ Ο α-άνθρακας είναι τετραεδρικός, καθώς σε αυτόν συνδέονται τέσσερις διαφορετικές ομάδες με αποτέλεσμα τα αμινοξέα να μπορούν να υπάρξουν σε δύο εναντιομερείς μορφές. Ωστόσο, στη φύση χρησιμοποιείται μόνο το ένα εναντιομερές για τη βιοσύνθεση των πρωτεϊνών.^{1,2}

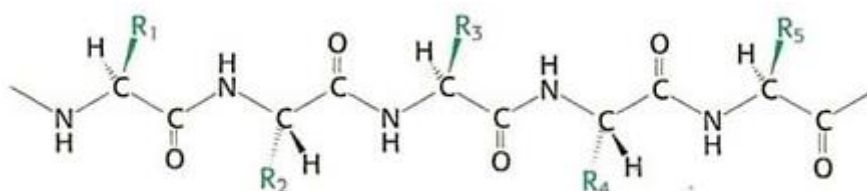
Όλες οι πρωτεΐνες των ζωντανών οργανισμών απαρτίζονται από 20 μόνο αμινοξέα στα οποία αντιστοιχίζονται 20 είδη πλευρικών αλυσίδων που διαφέρουν ως προς το μέγεθος, το φορτίο, την υδροφοβικότητα και την χημική δραστηκότητά τους.

Στην Εικόνα 1.4 παρουσιάζονται δύο πιο απλά αμινοξέα, η Γλυκίνη που έχει ως πλευρική αλυσίδα ένα άτομο Η και η Αλανίνη, που έχει ως πλευρική αλυσίδα μία μεθυλομάδα (-CH₃).



Εικόνα 1.4. Δομές Γλυκίνης και Αλανίνης. (Πάνω) Αναπαράσταση της διάταξης των ατόμων και των δεσμών με σφαίρες και ράβδους. (Μέση) Στερεοχημική απεικόνιση της γεωμετρίας των δεσμών και των ατόμων. (Κάτω) Προβολές κατά Fischer των δομών.¹

Τα αμινοξέα ενώνονται με πεπτιδικούς δεσμούς δημιουργώντας πολυπεπτιδικές αλυσίδες. Κάθε μονάδα αμινοξέος στην πολυπεπτιδική αλυσίδα ονομάζεται κατάλοιπο (residue). Η πολυπεπτιδική αλυσίδα αποτελείται από ένα σταθερά επαναλαμβανόμενο τμήμα που ονομάζεται **κύρια αλυσίδα ή κορμός**, και ένα μεταβλητό τμήμα που περιλαμβάνει τις διαφορετικές **πλευρικές αλυσίδες**, τα οποία παρουσιάζονται στην Εικόνα 1.5.

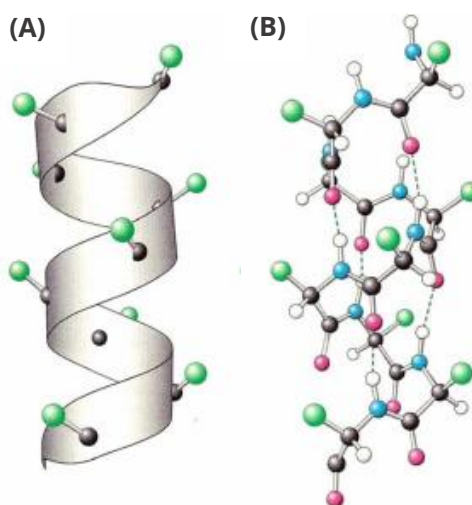


Εικόνα 1.5. Τα τμήματα της πολυπεπτιδικής αλυσίδας. Ο κορμός της αλυσίδας απεικονίζεται με μαύρο χρώμα και οι πλευρικές αλυσίδες με πράσινο χρώμα.¹

Οι φυσικές πολυπεπτιδικές αλυσίδες περιέχουν συνήθως από 50 έως 2000 κατάλοιπα αμινοξέων και ονομάζονται πρωτεΐνες. Το μέσο μοριακό βάρος ενός αμινοξέος είναι περίπου 110 με αποτέλεσμα το μέσο μοριακό βάρος των περισσότερων πρωτεϊνών να κυμαίνεται μεταξύ 5 500 και 220 000. Η μάζα μιας πρωτεΐνης εκφράζεται και σε μονάδες

dalton όπου ένα dalton ισούται με τη μονάδα ατομικής μάζας. Για παράδειγμα μια πρωτεΐνη μοριακού βάρους 50 000 έχει μάζα 50 000 dalton ή 50 kd (kilo-dalton).¹

Η αλληλουχία των αμινοξέων στην πολυπεπτιδική αλυσίδα συνιστά την **πρωτοταγή δομή** της πρωτεΐνης. Οι πολυπεπτιδικές αλυσίδες μπορούν να αναδιπλώνονται σε δομές όπως η α-έλικα, η β-πτυχωτή επιφάνεια, οι στροφές και οι θηλιές, σχηματίζοντας τη **δευτεροταγή δομή** της πρωτεΐνης. Στη δομή **α-έλικας**, η ομάδα CO κάθε αμινοξέος σχηματίζει έναν δεσμό υδρογόνου με την ομάδα NH του αμινοξέος που βρίσκεται τέσσερα κατάλοιπα μπροστά στην αλληλουχία με αποτέλεσμα τη δημιουργία μιας ραβδόμορφης δομής όπου ο κορμός έχει σχήμα σπειρώματος. Η δομή της α-έλικας παρουσιάζεται στην Εικόνα 1.6. Οι πολυπεπτιδικές αλυσίδες μπορούν να αλλάξουν κατεύθυνση δημιουργώντας **στροφές** και θηλιές. Οι στροφές και οι θηλιές βρίσκονται κυρίως στην επιφάνεια των πρωτεϊνών και συμμετέχουν στην αλληλεπίδραση των πρωτεϊνών με άλλα μόρια.¹ Οι υδατοδιαλυτές πρωτεΐνες αναδιπλώνονται στο χώρο σε συμπαγείς δομές με πολικά κέντρα. Η **τριτοταγής δομή** της πρωτεΐνης αναφέρεται στις αναδιπλώσεις ολόκληρης της αλυσίδας στο χώρο. Για πρωτεΐνες που αποτελούνται από δύο ή περισσότερες αλυσίδες, ο τρόπος με τον οποίο τοποθετούνται οι αλυσίδες αυτές στο χώρο συνιστά την **τεταρτοταγή δομή** της πρωτεΐνης.



Εικόνα 1.6. Η δομή μιας α-έλικας. (A) Απεικόνιση κορδέλας στην οποία ξεχωρίζουν τα άτομα α-άνθρακα (μαύρο) και οι πλευρικές αλυσίδες (πράσινο). (B) Μια πλάγια όψη του μοντέλου με σφαίρες και ράβδους όπου διακρίνονται οι δεσμοί υδρογόνου (διακεκομμένες γραμμές) μεταξύ των ομάδων NH και CO.¹

1.5.2. ΔΟΜΗ ΜΕΤΑΓΡΑΦΙΚΩΝ ΠΑΡΑΓΟΝΤΩΝ

Οι μεταγραφικοί παράγοντες είναι πρωτεϊνικά μόρια που έχουν διπλή λειτουργία: αφενός συνδέονται με το DNA του προς μεταγραφή γονιδίου, και αφετέρου συνδέονται με τους ενεργοποιητές, συνενεργοποιητές και με άλλους μεταγραφικούς παράγοντες για τη δημιουργία του συμπλέγματος έναρξης της μεταγραφής. Ένας μεταγραφικός παράγοντας αποτελείται από δύο ειδών δομικές περιοχές: τις **δομικές περιοχές σύνδεσης με το DNA** (DNA-binding domains, **DBDs**) και τις **περιοχές ενεργοποίησης** (activation domains, **ADs**).¹⁰

Οι **δομικές περιοχές σύνδεσης με το DNA** αποτελούν περιοχές της πρωτεΐνης με καθορισμένη δομή στον τρισδιάστατο χώρο όπως είναι τα δομικά μοτίβα: Έλικά-στροφή-έλικα, Φτερωτή έλικά-στροφή-έλικα, Δάκτυλος ψευδαργύρου, Αυτόκλειστο λευκίνης, β-πτυχωτή επιφάνεια, β-φουρκέτα.^{5, 11} Οι περιοχές αυτές συνδέονται σε σχετικά μικρές αλληλουχίες νουκλεοτιδίων του DNA, περίπου 6-12 ζευγάρια βάσεων.

Οι **δομικές περιοχές ενεργοποίησης** των μεταγραφικών παραγόντων συνδέονται με τους συνενεργοποιητές και - σε αντίθεση με τις περιοχές σύνδεσης με το DNA - δεν έχουν καθορισμένη τρισδιάστατη δομή και καλούνται **ενδογενώς αδόμητες περιοχές** (intristically disordered regions).¹⁰ Οι ενδογενώς αδόμητες περιοχές αποτελούν κοινό χαρακτηριστικό των συνενεργοποιητών και των μεταγραφικών παραγόντων και η σύνδεση που πραγματοποιείται μεταξύ τους οφείλεται στις αλληλεπιδράσεις μεταξύ των αδόμητων τμημάτων IDR της πολυπεπτιδικής τους αλυσίδας.^{6, 10}

Ο προσδιορισμός του μικροσκοπικού μηχανισμού με τον οποίο οι μεταγραφικοί παράγοντες προσδένονται με το DNA και ελέγχουν τη μεταγραφή είναι μια πρόκληση για την κατανόηση του ρόλου τους, την αποσαφήνιση συγκεκριμένων λειτουργιών του γονιδιώματος και τη χαρτογράφηση των μονοπατιών έκφρασης σημαντικών γονιδίων των σύνθετων οργανισμών.⁵

1.6. Ο ΜΕΤΑΓΡΑΦΙΚΟΣ ΠΑΡΑΓΟΝΤΑΣ SRY

Για τη μελέτη, μέσω μοριακών προσομοιώσεων, του μηχανισμού πρόσδεσης των μεταγραφικών παραγόντων στο DNA, η οποία συντελείται κατά το πρώτο στάδιο της μεταγραφής, επιλέχθηκε το σύστημα ενός μεταγραφικού παράγοντα και του αντίστοιχου τμήματος του γονιδίου στο οποίο αυτός προσδένεται για να ενεργοποιήσει τη μεταγραφή του. Ο μεταγραφικός παράγοντας που επιλέχθηκε είναι η πρωτεΐνη Περιοχής Φυλοκαθορισμού του Χρωμοσώματος Y (Sex-determining region Y protein or SRY protein), γνωστή και ως Παράγοντας Καθορισμού Όρχεων (Testis-determining factor), η οποία διαδραματίζει κρίσιμο ρόλο στην ανάπτυξη του αρρενοσ φύλου κατά την εμβρυική περίοδο της κύησης. Το γονίδιο στο οποίο προσδένεται ο μεταγραφικός παράγοντας SRY^a, προς ενεργοποίηση της μεταγραφής του, είναι το γονίδιο SOX9. Η λειτουργική δράση του συγκεκριμένου συστήματος αποτελεί κύριο μοριακό διακόπτη στον φυλοκαθορισμό των θηλαστικών, ενώ η ανεπαρκής ή ετεροχρονισμένη μεταγραφή του γονιδίου SOX9 δύναται να προκαλέσει φυλοσύνδετες διαταραχές. Το σύστημα SRY - SOX9 παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς ο μεταγραφικός παράγοντας SRY συνδέεται στη μικρή αύλακα του DNA, σε αντίθεση με την πλειοψηφία των μεταγραφικών παραγόντων, οι οποίοι συνδέονται στη μεγάλη αύλακα του DNA.

Το γονίδιο που κωδικοποιεί την SRY εδράζεται στο βραχύ σκέλος του χρωμοσώματος Y και ανήκει στην SOX οικογένεια γονιδίων, τα οποία κωδικοποιούν SOX μεταγραφικούς παράγοντες. Οι μεταγραφικοί αυτοί παράγοντες διαθέτουν μια ειδική περιοχή που ονομάζεται **κυτίο ομάδας υψηλής κινητικότητας** (High Mobility Group Box, **HMG-Box**) χάρη στην οποία επιτυγχάνεται η σύνδεσή τους με το DNA. Η περιοχή HMG-Box χαρακτηρίζεται από τρεις άλφα-έλικες που σχηματίζουν μια δομή σχήματος «L», η οποία επιτρέπει την πρόσδεση στην μικρή αύλακα του DNA.

Ο μεταγραφικός παράγοντας SRY στον άνθρωπο αποτελείται από 204 αμινοξέα και έχει μοριακό βάρος 23,88 kDa.¹² Περιλαμβάνει μια N-τελική περιοχή (N-terminal domain, NTD), μια περιοχή HMG-Box 79 αμινοξέων και μια C-τελική περιοχή (C-terminal domain, CTD). Η HMG-Box είναι εξαιρετικά συντηρημένη εξελικτικά. Αντίθετα, τα άκρα NTD και CTD δεν είναι εξελικτικά συντηρημένα και εμφανίζουν διαφοροποιήσεις στα διάφορα είδη θηλαστικών.¹³

^a Για τη συντομογραφία των πρωτεϊνών και των γονιδίων ακολουθείται η σημειογραφία σύμφωνα με την American Society for Clinical Investigation (ASCI, <https://www.jci.org/kiosk/publish/genestyle>): οι πρωτεΐνες σημειώνονται με κεφαλαία, ενώ τα γονίδια με κεφαλαία πλάγια γράμματα

1.7. Ο ΜΕΤΑΓΡΑΦΙΚΟΣ ΠΑΡΑΓΟΝΤΑΣ SRY ΚΑΙ Η ΔΙΑΜΟΡΦΩΣΗ ΤΟΥ ΦΥΛΟΥ ΣΤΟΝ ΑΝΘΡΩΠΟ

Ο καθορισμός του φύλου, τόσο στον άνθρωπο όσο και στα περισσότερα θηλαστικά, γίνεται βάσει του συστήματος φυλετικών χρωμοσωμάτων XX – XY. Κατά αυτό το σύστημα, το ωκύτταρο φέρει ένα X χρωμόσωμα ενώ το σπερματοζώαριο δύναται να φέρει είτε X είτε Y χρωμόσωμα. Κατά τη γονιμοποίηση μπορεί συνεπώς να προκύψει είτε XX καρυότυπος ο οποίος αντιστοιχεί σε φαινότυπο θήλεος, είτε XY καρυότυπος ο οποίος αντιστοιχεί σε φαινότυπο άρρενος. Με αυτόν τον τρόπο τη στιγμή της γονιμοποίησης καθορίζεται το **γενετικό φύλο**.

Παρότι το γενετικό φύλο καθορίζεται τη στιγμή της γονιμοποίησης, στην ανθρώπινη διάπλαση οι γεννητικοί αδένες αναπτύσσονται όμοια και στα δύο φύλα μέχρι την 7^η εβδομάδα της κύησης και ονομάζονται **αδιαφοροποίητοι γεννητικοί αδένες ή πρωταρχικές γονάδες**.^{14, 15} Κατά την κύηση, τα κύτταρα των πρωταρχικών γονάδων (αναπαραγωγικά κύτταρα) είναι σε διδύναμη κατάσταση, που σημαίνει ότι έχουν τη δυνατότητα να γίνουν είτε ανδρικά κύτταρα (κύτταρα Sertoli και Leydig) ή γυναικεία κύτταρα (κύτταρα ωοθυλακίου και θήκης). Η μετέπειτα διαφοροποίηση των των διδύναμων γονάδων εξαρτάται από την ύπαρξη ή απουσία του φυλοκαθοριστικού γονιδίου *SRY*, που βρίσκεται πάνω στο Y χρωμόσωμα, και είναι υπεύθυνο για την έναρξη του καθορισμού του ανδρικού φύλου στα θηλαστικά.¹⁶

Το γονίδιο *SRY* κωδικοποιεί την πρωτεΐνη που καλείται πρωτεΐνη Περιοχής Φυλοκαθορισμού του Χρωμοσώματος Y (Sex-determining region Y protein or SRY protein), ή πρωτεΐνη SRY, γνωστή και ως Παράγοντας Καθορισμού Όρχεων (Testis-determining factor), η οποία λειτουργεί ως κύριος μοριακός διακόπτης στον φυλοκαθορισμό των θηλαστικών.¹⁷ Η πρωτεΐνη SRY, ούσα η ίδια ένας μεταγραφικός παράγοντας, πυροδοτεί ένα κύμα γονιδιακών δραστηριοτήτων που στο σύνολό τους ωθούν την ανάπτυξη του άρρενος φύλου του εμβρύου έναντι του θήλεος.¹³ Ενεργοποιεί άλλους μεταγραφικούς παράγοντες που εμπλέκονται στον καθορισμό του ανδρικού φύλου, όπως ο μεταγραφικός παράγοντας *SOX9*, οι οποίοι μπορούν να αναστείλουν τη δημιουργία του θηλυκού αναπαραγωγικού συστήματος κατά την εμβρυική ανάπτυξη (6-8 εβδομάδες από τη δημιουργία του εμβρύου), καθώς και να αλληλεπιδράσουν με άλλα γονίδια ώστε να προάγουν την ανάπτυξη των ανδρικών γεννητικών οργάνων. Κατά τη διαδικασία αυτή επιτρέπουν τη διαφοροποίηση των διδύναμων γονάδων σε κύτταρα Sertoli, και ακολούθως τον πολλαπλασιασμό και την ανάπτυξή τους, ώστε να δημιουργήσουν τον πρωταρχικό ανδρικό γεννητικό αδέν. Τα κύτταρα Sertoli παράγουν την αντιμυλλέρεια ορμόνη, AMH, η οποία προκαλεί την υποστροφή των πόρων του Müller, που σχηματίζουν τη μήτρα και τους ωαγωγούς, και εκκρίνεται στο ανθρώπινο σώμα σε μεγάλο βαθμό μέχρι και την εφηβεία.¹⁵

Η ενεργοποίηση της μεταγραφής του γονιδίου *SOX9* για τη σύνθεση του μεταγραφικού παράγοντα *SOX9* συμβαίνει όταν η πρωτεΐνη *SRY* συμπλέκεται με τον στεροειδογενετικό παράγοντα 1 (steroidogenic factor 1, *SF1*, γνωστός επίσης και ως *NR5A1*), έναν έτερο μεταγραφικό ρυθμιστή, και αλληλεπιδρούν με την περιοχή της ενισχυτικής αλληλουχίας *TESCO* (testis-specific enhancer of *SOX9* core), δηλαδή της ενισχυτικής αλληλουχίας που καθορίζει την ανάπτυξη των ανδρικών γεννητικών αδένων, που βρίσκεται πάνω στο γονίδιο *SOX9*. Ειδικότερα, η ένωση με το DNA επιτυγχάνεται μέσω της «δομικής περιοχής πρόσδεσης στο DNA» (DNA-binding domain) της πρωτεΐνης *SRY*, η οποία προσδένεται στη μικρή αύλακα του DNA στοχευμένα στη συγκεκριμένη αλληλουχία *TESCO*, προκαλώντας ξεδίπλωμα του DNA. Η επίτευξη της πρόσδεσης διευκολύνει τη μεταγραφή του γονιδίου *SOX9*,¹⁸ έχοντας ως αποτέλεσμα τη θετική ρύθμιση του μεταγραφικού παράγοντα *SOX9*. Αφού εκφραστεί η πρωτεΐνη *SOX9*, αποτελεί και η ίδια παράγοντα αυτορρύθμισης του γονιδίου που την κωδικοποιεί, ανατροφοδοτώντας θετικά την παραγωγή της, με αλληλεπίδραση με την αλληλουχία *TESCO*.^{17, 19}

Οι πρωτεΐνες *SRY* και *SOX9* αποτελούν μεταγραφικούς παράγοντες κάποιων κοινών γονιδίων και μάλιστα σε πολλά από αυτά συνδέονται στην ίδια περιοχή των υποκινητών.¹⁹ Ωστόσο, υπάρχουν ενδείξεις ότι η *SOX9* μπορεί να αποτελέσει όχι μόνο αναγκαίο αλλά και ικανό παράγοντα για τη διαφοροποίηση του άρρενος φύλου.¹⁴ Σε πειραματικές μελέτες σε ποντίκια, η εκτοπική έκφραση της *SOX9* οδηγεί στην ανάπτυξη των ανδρικών γεννητικών αδένων ακόμα και χωρίς την παρουσία λειτουργικής *SRY*.¹³ Αν και τα γονιδιακά μονοπάτια του καθορισμού του φύλου δεν έχουν ακόμη χαρτογραφηθεί πλήρως, μια τέτοια πληροφορία υποδεικνύει ότι ο λειτουργικός ρόλος της *SRY* είναι ουσιαστικά η ενεργοποίηση της μεταγραφής της *SOX9*.¹⁹ Συνεπώς εγείρεται το κρίσιμο ζήτημα του τρόπου με τον οποίο η *SRY* ενεργοποιεί την μεταγραφή της *SOX9* μέσω της πρόσδεσής της στο γονίδιο *SOX9*. Η διερεύνηση του τρόπου πρόσδεσης της *SRY* στο DNA αποτελεί αντικείμενο της παρούσας εργασίας.

1.8. ΜΕΤΑΓΡΑΦΙΚΟΣ ΠΑΡΑΓΟΝΤΑΣ SRY ΚΑΙ ΦΥΛΟΣΥΝΔΕΤΕΣ

ΔΙΑΤΑΡΑΧΕΣ

Διαταραχές της ανάπτυξης του φύλου (Disorders of Sex Development, DSDs) είναι παθήσεις οι οποίες επηρεάζουν την ανάπτυξη των γονάδων ή των εξωτερικών γεννητικών οργάνων.¹⁴ Για την ανάπτυξη των όρχεων σε XY καρυότυπο είναι αναγκαία έκφραση του γονιδίου *SOX9* έγκαιρα και σε ικανοποιητικό βαθμό. Η καθυστερημένη ή ανεπαρκής έκφραση του γονιδίου αυτού οδηγεί σε διαταραχές στην ανάπτυξη των όρχεων και σε φαινόμενα αντιστροφής φύλου¹³

Το γονίδιο *SRY*, το οποίο κωδικοποιεί την πρωτεΐνη SRY, περιέχει μια ισχυρά συντηρημένη αλληλουχία μεταξύ τρωκτικών και άλλων θηλαστικών, η οποία ονομάζεται «κυτίο HMG» (κουτί Ομάδας Υψηλής Ενεργότητας, High Mobility Group (HMG) box) και κωδικοποιεί τη «δομική περιοχή πρόσδεσης στο DNA» της πρωτεΐνης SRY. Μεταλλάξεις στο κυτίο HMG προκαλούν αναστροφή ανάπτυξης του φύλου.²⁰ Γενικότερα, οι μεταλλάξεις στο γονίδιο *SRY* (και συνεπώς πιθανές αλλαγές στη δομή της πρωτεΐνης και τη λειτουργία της) οδηγούν σε ένα ευρύ φάσμα φυλοσύνδετων διαταραχών με ποικίλες επιδράσεις στο φαινότυπο και το γονότυπο του ατόμου.

Πιο συγκεκριμένα, μεταλλάξεις στην *SRY* ή στο ίδιο το γονίδιο *SOX9* δύνανται να οδηγήσουν σε αναστροφή φύλου κατά την οποία το άτομο, αν και φέρει XY καρυότυπο, αναπτύσσει εξωτερικά γεννητικά χαρακτηριστικά θήλεος (*46, XY αναστροφή φύλου 1*). Τέτοιου είδους μεταλλάξεις οδηγούν σε εκφύλιση των γονάδων, κατά την οποία ο αναπαραγωγικός ιστός αντικαθίσταται από συνδετικό ιστό με αποτέλεσμα την αδυναμία ανάπτυξης δευτερογενών χαρακτηριστικών του φύλου κατά την εφηβεία και την μη ικανότητα αναπαραγωγής.¹²

Μεταλλάξεις κατά τις οποίες το *SRY* γονίδιο έχει μετατοπιστεί στο X χρωμόσωμα καρυότυπου δύνανται να προκαλέσουν την ανάπτυξη ανδρικών γονάδων (*46, XX αναστροφή φύλου 1*). Ο φαινότυπος που προκύπτει σε αυτήν την περίπτωση ποικίλει και μπορεί να αφορά στη συνύπαρξη ωοθηκών και όρχεων ή μόνο στην ανάπτυξη όρχεων.^{12,}

21

2. ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ

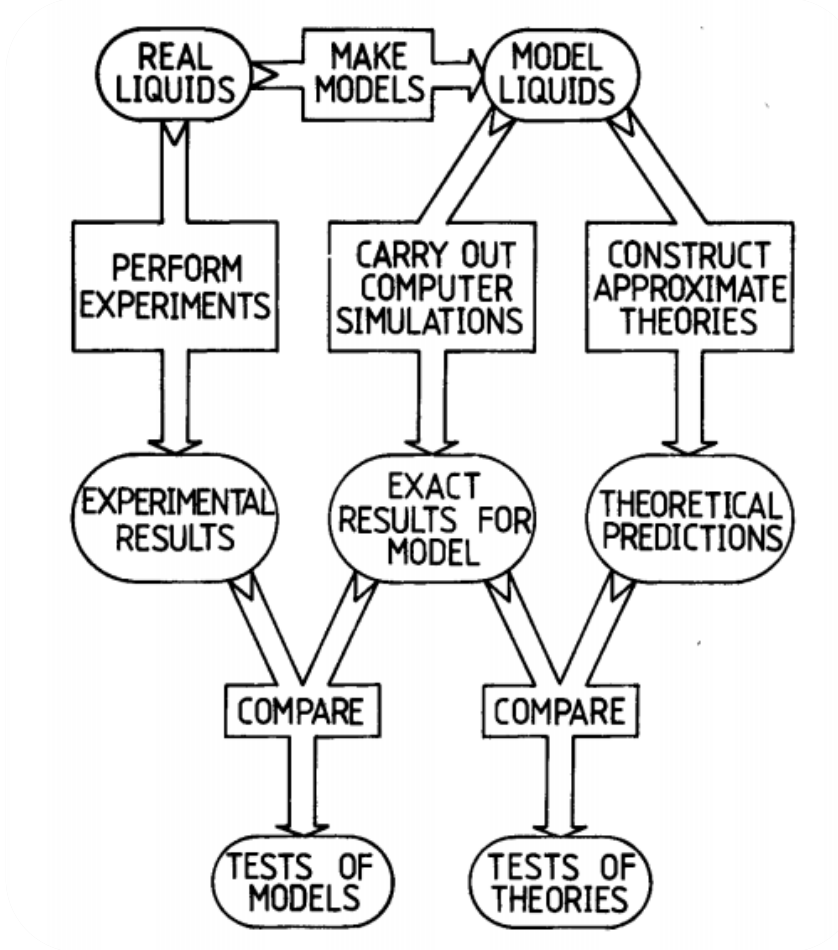
ΕΙΣΑΓΩΓΗ

Η κατανόηση της δομής και της δυναμικής των σωματιδίων που χαρακτηρίζει την υγρή φάση της ύλης αποτελεί ένα πρόβλημα που έχει απασχολήσει εδώ και πολλά χρόνια τόσο πειραματικούς όσο και θεωρητικούς επιστήμονες. Οι πειραματικοί παρατηρούν τη συμπεριφορά των υγρών μέσω πειραμάτων, ενώ οι θεωρητικοί επιστήμονες κατασκευάζουν μοντέλα για να εξηγήσουν τη συμπεριφορά τους. Τα μοντέλα που κατασκευάζονται, αναλύονται με τη βοήθεια υπολογιστικών προσομοιώσεων.

Για την ανάλυση μοριακών μοντέλων υγρών με υπολογιστική προσομοίωση χρησιμοποιούνται δύο τεχνικές: η Monte Carlo προσομοίωση και η προσομοίωση Μοριακής Δυναμικής. Και οι δύο αυτές τεχνικές αποσκοπούν στη δημιουργία της τροχιάς των σωματιδίων στο χωροχρόνο, ωστόσο διαφέρουν ως προς τον τρόπο με τον οποίο την εξάγουν. Η Μοριακή δυναμική, η οποία χρησιμοποιείται στην παρούσα εργασία, αφορά στην επίλυση των κλασικών εξισώσεων κίνησης του Νεύτωνα για ένα σύνολο μορίων. Η τεχνική αυτή υλοποιήθηκε για πρώτη φορά για ένα σύστημα σκληρών σφαιρών από τους Alder και Wainwright το 1959. Αργότερα, το 1964, επιτεύχθηκε η επίλυση των εξισώσεων κίνησης για σωματίδια Lennard-Jones.²² Έκτοτε, η Μοριακή δυναμική έχει γνωρίσει ραγδαία ανάπτυξη και πλέον επιλύονται εξισώσεις για σύνθετα πεδία δυνάμεων.

Η εύρεση της τροχιάς των σωματιδίων οδηγεί, μέσω της θεωρίας της στατιστικής μηχανικής, στον προσδιορισμό των θερμοδυναμικών ιδιοτήτων του συστήματος. Με αυτόν τον τρόπο, οι υπολογιστικές προσομοιώσεις αποτελούν έναν άμεσο σύνδεσμο των μικροσκοπικών ιδιοτήτων του συστήματος (μάζες ατόμων και αλληλεπιδράσεις, μοριακή γεωμετρία κλπ) και των μακροσκοπικών ιδιοτήτων πειραματικού ενδιαφέροντος. Παρέχουν ακριβή αποτελέσματα σε προβλήματα στατικής μηχανικής, τα οποία μπορεί να μην ήταν επιλύσιμα με άλλες μεθόδους και αποτελούν έναν τρόπο ελέγχου των αντίστοιχων θεωριών. Ταυτόχρονα, τα αποτελέσματά τους συγκρίνονται με αυτά των πρακτικών πειραμάτων, προσφέροντας μια διαφορετική οπτική σε αυτά και συνεισφέροντας στην ερμηνεία νέων αποτελεσμάτων.

Συνεπώς, οι υπολογιστικές προσομοιώσεις αποτελούν μια γέφυρα, τόσο μεταξύ των μοντέλων και των θεωρητικών προβλέψεων, όσο και μεταξύ των μοντέλων και των πειραματικών αποτελεσμάτων. Λόγω αυτού του συνδυαστικού ρόλου και του τρόπου με τον οποίο οι προσομοιώσεις διεξάγονται και αναλύονται, συχνά αποκαλούνται «υπολογιστικά πειράματα». Ο συνδυαστικός ρόλος των υπολογιστικών προσομοιώσεων παρουσιάζεται στην Εικόνα 2.1.



Εικόνα 2.1. Απεικόνιση της σύνδεσης μεταξύ πειράματος, θεωρίας και υπολογιστικής προσομοίωσης.

2.1. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΣΤΙΣ ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ

2.1.1. Η ΠΕΡΙΓΡΑΦΗ ΤΗΣ ΤΡΟΧΙΑΣ ΣΤΗΝ ΚΛΑΣΙΚΗ ΜΗΧΑΝΙΚΗ

Η περιγραφή της κίνησης των σωμάτων στον χώρο με την πάροδο του χρόνου πραγματοποιείται μέσω της κλασικής μηχανικής. Η βάση της κλασικής μηχανικής είναι ο νόμος του Νεύτωνα,

$$\mathbf{F} = m \cdot \mathbf{a} \quad (2.1.1)$$

όπου \mathbf{F} είναι η δύναμη που ασκείται σε ένα σώμα μάζας m και \mathbf{a} η επιτάχυνση που έχει το σώμα τη χρονική στιγμή που του ασκείται η δύναμη. Η θεωρία αυτή έχει εφαρμογή σε μεγάλο εύρος μεγέθους σωμάτων: από τα μόρια μέχρι τους πλανήτες. Για τα συστήματα πολλών μορίων, που ενδιαφέρουν την παρούσα εργασία, ορίζεται ένα πρότυπο κατά το οποίο οι θέσεις N σωματιδίων μπορούν να περιγραφούν μέσω ανυσμάτων:

$$\{\mathbf{r}\} = \{x_1, y_1, z_1, \dots, x_N, y_N, z_N\} \quad (2.1.2)$$

με βάση συγκεκριμένο σημείο αναφοράς (O). Οι ανεξάρτητες συνιστώσες των $\{\mathbf{r}\}$ στο σύνολό τους αποτελούν τους **βαθμούς ελευθερίας** του συστήματος. Το πρότυπο καθορίζεται από ένα σύνολο ιδιοτήτων συμπεριλαμβανομένου του χώρου στον οποίο κινούνται τα σωματίδια, τις μάζες του και τις δυνάμεις που ασκούνται σε αυτά, οι οποίες αποτελούν συναρτήσεις της θέσης και του χρόνου, $\mathbf{F}(\mathbf{r}, t)$.

Η κλίση της συνάρτησης του δυναμικού δίνει το αντίστοιχο μέγεθος της δύναμης που ασκείται στο σώμα, σύμφωνα με την εξίσωση (1.3), η οποία με τη σειρά της καθορίζει την αρχική επιτάχυνση του σώματος.

$$\mathbf{F} = -\nabla V \quad (2.1.3)$$

Οι εξισώσεις **Lagrange** χρησιμοποιούνται για την κλασική μηχανική περιγραφή του συστήματος. Για τις γενικευμένες μεταβλητές θέσης των σωματιδίων, \mathbf{q}_k , η Lagrangian εξίσωση της κίνησης είναι η:

$$\frac{d}{dt} \frac{\partial}{\partial \dot{\mathbf{q}}_k} L - \frac{\partial L}{\partial \mathbf{q}_k} = 0, \quad (2.1.4)$$

όπου η Lagrangian συνάρτηση $L(\mathbf{q}, \dot{\mathbf{q}})$ ορίζεται από την διαφορά της δυναμικής από την κινητική ενέργεια:

$$L = K - V \quad (2.1.5)$$

και αποτελεί συνάρτηση των γενικευμένων συντεταγμένων, \mathbf{q}_k , και των χρονικών τους παραγώγων $\dot{\mathbf{q}}_k$. Η γενικευμένη ορμή \mathbf{p}_k ορίζεται ως:

$$\mathbf{p}_k = \frac{\partial L}{\partial \dot{\mathbf{q}}_k} \quad (2.1.6)$$

Με μετασχηματισμό Legendre προκύπτει η Χαμιλτονιανή του συστήματος, $\mathcal{H}(\mathbf{p}, \mathbf{q})$, η οποία ορίζεται από την εξίσωση:

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = \dot{\mathbf{q}} \cdot \mathbf{p} - L(\mathbf{q}, \dot{\mathbf{q}}) \quad (2.1.7)$$

Η Χαμιλτονιανή όπως περιγράφεται στην εξ. (2.1.7), με εξάρτηση από το χρόνο μόνον μέσω των γενικευμένων ορμών, \mathbf{p} , και γενικευμένων συντεταγμένων, \mathbf{q} , όταν περιγράφει συντηρητικό σύστημα (συντηρητικό πεδίο δυνάμεων), ισούται με την ενέργεια του συστήματος, και αποτελεί σταθερά της κίνησης του συστήματος. Οι εξισώσεις κίνησης προκύπτουν από τις σχέσεις:

$$\dot{\mathbf{q}}_k = \frac{\partial \mathcal{H}}{\partial p_k} \quad \text{και} \quad \dot{\mathbf{p}}_k = -\frac{\partial \mathcal{H}}{\partial q_k} \quad (2.1.8)$$

όπου $k = 1, 2, \dots, 3N$

2.1.2. ΠΕΡΙΓΡΑΦΗ ΜΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΚΑΙ ΔΥΝΑΜΙΚΑ ΑΛΛΗΛΕΠΙΔΡΑΣΗΣ

Η μικροσκοπική κατάσταση ενός συστήματος προσδιορίζεται από τις θέσεις και τις ορμές των σωματιδίων του, δηλαδή των ατόμων και των μορίων. Η προσέγγιση Born-Oppenheimer επιτρέπει την έκφραση της Χαμιλτονιανής ενός συστήματος N σωματιδίων ως άθροισμα των συναρτήσεων της κινητικής, K , και της δυναμικής, V , ενέργειας του συστήματος. Έτσι, για ένα σύστημα στο οποίο κάθε σωματίδιο i έχει συντεταγμένες q_i και ορμές p_i , ισχύουν:

$$\mathbf{q} = (q_1, q_2, \dots, q_{3N}) \quad (2.1.9)$$

$$\mathbf{p} = (p_1, p_2, \dots, p_{3N}) \quad (2.1.10)$$

$$\mathcal{H}(\mathbf{p}, \mathbf{q}) = K(\mathbf{p}) + V(\mathbf{q}) \quad (2.1.11)$$

Η **κινητική ενέργεια** K , σε αρκετά φυσικά συστήματα λαμβάνει τη μορφή:

$$K = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i} \quad (2.1.12)$$

όπου m_i είναι η μάζα του μορίου.

Η κίνηση των μορίων εξαρτάται από τη δομή τους, τις αλληλεπιδράσεις μεταξύ τους καθώς επίσης και από εξωτερικές δυνάμεις που μπορεί να ενεργούν στο σύστημα. Η μοριακή δομή καθορίζει τους **βαθμούς ελευθερίας** του συστήματος, οι οποίοι συνίστανται σε βαθμούς μεταφοράς, περιστροφής και δόνησης.

Για δεδομένη μοριακή δομή γίνεται καθορισμός των αλληλεπιδράσεων των μορίων μέσω των δυναμικών αλληλεπίδρασης. Τα δυναμικά αυτά αντιστοιχούν στην δυναμική ενέργεια, V , που έχουν τα μόρια όταν βρεθούν κοντά μεταξύ τους. Συνεπώς, η δυναμική ενέργεια εμπεριέχει χρήσιμες πληροφορίες για τις διαμοριακές αλληλεπιδράσεις. Γνωρίζοντάς την, μπορούμε, μέσω της Χαμιλτονιανής, να κατασκευάσουμε της εξισώσεις κίνησης οι οποίες καθορίζουν τη χρονική εξέλιξη του συστήματος και τις μηχανικές του ιδιότητες. Η Χαμιλτονιανή ή η δυναμική ενέργεια αποτελούν τις βασικές εισόδους σε υπολογιστικές προσομοιώσεις. Η **δυναμική ενέργεια** V αναλύεται σε όρους οι οποίοι εξαρτώνται από τις συντεταγμένες των ατόμων, ζευγών ατόμων ή τριάδων ατόμων:

$$V = \sum_i u_1(\mathbf{r}_i) + \sum_i \sum_{j>i} u_2(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \sum_{j>1} \sum_{k>j>i} u_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \quad (2.1.13)$$

όπου τα αθροίσματα αφορούν σε όλα τα διακριτά ζεύγη (ή τριάδες κλπ), χωρίς να λαμβάνεται κάποιο ζεύγος δεύτερη φορά (πχ ij και ji). Ο πρώτος όρος της εξίσωσης (2.1.13), $u_1(\mathbf{r}_i)$, αναπαριστά την επίδραση εξωτερικού πεδίου στο σύστημα ενώ όλοι οι υπόλοιποι τις αλληλεπιδράσεις των σωματιδίων.

Ο δεύτερος όρος, u_2 , ονομάζεται δυναμικό ζεύγους και εκφράζει την αλληλεπίδραση μεταξύ δύο σωματιδίων. Τα σωματίδια εκλαμβάνονται ως σχεδόν σφαιρικά «κέντρα αλληλεπίδρασης» και το δυναμικό αλληλεπίδρασης εξαρτάται μόνο από το διάνυσμα μεταξύ των κέντρων αλληλεπίδρασης, $\mathbf{r}_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$.

Η μαθηματική περιγραφή του δυναμικού αλληλεπίδρασης συναρτήσει των συντεταγμένων του συστήματος σωματιδίων, μαζί με τις παραμέτρους για τον υπολογισμό του, συνιστούν το **πεδίο δυνάμεων**. Αυτό μπορεί να εκφραστεί σε όρους δεσμικών και μη δεσμικών αλληλεπιδράσεων:

$$V = V_{bonds} + V_{angles} + V_{dihedrals} + V_{improper\ dihedrals} + V_{vdW} + V_{ES} \quad (2.1.14)$$

Οι πρώτοι τέσσερις όροι της εξ. (2.1.14) αναφέρονται σε δεσμικές (ενδομοριακές) αλληλεπιδράσεις και αφορούν στα μήκη δεσμών, τις γωνίες δεσμών τις δίεδρες γωνίες και τις μη κανονικές δίεδρες αντίστοιχα.

Οι δύο τελευταίοι όροι της εξ (2.1.14) αναφέρονται σε μη δεσμικές (διαμοριακές) αλληλεπιδράσεις και συγκεκριμένα στις αλληλεπιδράσεις van der Waals και στις ηλεκτροστατικές αλληλεπιδράσεις.

Αν και οι διαμοριακές δυνάμεις είναι βασικά ηλεκτροστατικής φύσης, διακρίνονται τέσσερις τύποι διαμοριακών δυνάμεων:²³

Ηλεκτροστατικές δυνάμεις

Ασκούνται μεταξύ μοριακών ειδών που χαρακτηρίζονται από μόνιμη κατανομή φορτίων. Οι δυνάμεις αυτές δεν απαιτούν παραμόρφωση των ηλεκτρονιακών νεφών για να εκφραστούν (π.χ. ιόντα, δίπολα, τετράπολα και ούτω καθεξής).

Δυνάμεις εξ επαγωγής

Ασκούνται μεταξύ ενός μη πολικού μορίου και ενός μορίου της προηγούμενης κατηγορίας. Οι δυνάμεις αυτές προκαλούν διαταραχή της κατανομής του ηλεκτρονιακού φορτίου του μη πολικού μορίου και προκύπτει παραμόρφωση του ηλεκτρονιακού νέφους.

Δυνάμεις διασποράς

Είναι ασθενείς δυνάμεις που ασκούνται μεταξύ μη πολικών μορίων εξαιτίας των διακυμάνσεων του ηλεκτρονιακού νέφους που έχουν ως αποτέλεσμα τη δημιουργία στιγμιαίων διπόλων με αποτέλεσμα ελκτική αλληλεπίδραση.

Ειδικές ή Χημικές δυνάμεις

Οι δυνάμεις αυτές περιλαμβάνουν τη μεταφορά ηλεκτρονιακού φορτίου από το ένα κέντρο στο άλλο και ενέχουν το χαρακτήρα χημικού δεσμού. Είναι ισχυρές και έχουν σαφείς προτιμήσεις προσανατολισμού.

Ειδικότερα οι δυνάμεις που προέρχονται από αλληλεπιδράσεις του τύπου:

Μονίμου διπόλου – Μονίμου διπόλου, Μονίμου διπόλου – Στιγμιαίου διπόλου, και Στιγμιαίου διπόλου – Στιγμιαίου διπόλου, αναφέρονται ως δυνάμεις **van der Waals**.

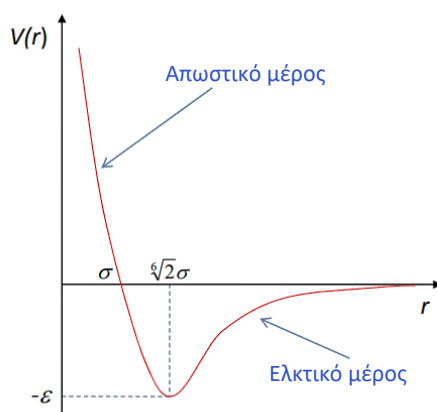
Κοινό χαρακτηριστικό των ανωτέρω τύπων δυναμικών αλληλεπίδρασης είναι η συνάρτησή τους με την απόσταση: r^{-6} . Η τιμή δε των δυναμικών αυτών είναι πάντα αρνητική διότι περιγράφουν ελκτικές αλληλεπιδράσεις μόνο. Ένα ευρέως χρησιμοποιούμενο στις υπολογιστικές προσομοιώσεις δυναμικό αλληλεπίδρασης το οποίο περιγράφει αλληλεπιδράσεις που οφείλονται σε δυνάμεις van der Waals είναι το δυναμικό Lennard-Jones 12-6. Το δυναμικό αυτό περιγράφει αλληλεπιδράσεις που

οφείλονται σε δυνάμεις van der Waals και αποτελείται από ένα ελκτικό μέρος κβαντομηχανικής προέλευσης με τη μορφή:

$$V(r) = 4\varepsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right), \quad (2.1.15)$$

όπου r η απόσταση μεταξύ των δυο σωματιδίων, ε το βάθος του πηγαδιού του δυναμικού και σ η απόσταση στην οποία η δυναμική ενέργεια μεταξύ των σωματιδίων είναι μηδενική.

Το δυναμικό Lennard-Jones παρουσιάζεται στην Εικόνα 2.2.



Εικόνα 2.2. Διαγραμματική απεικόνιση του δυναμικού Lennard-Jones συναρτήσεως του διανύσματος των κέντρων αλληλεπίδρασης. Το δυναμικό λαμβάνει την ελάχιστη τιμή του

σε απόσταση $r = r_m = 2^{1/6} \sigma$, $\left(\frac{dV}{dr} \Big|_{r=r_m} = 0 \text{ και } V(r_m) = -\varepsilon \right)$.

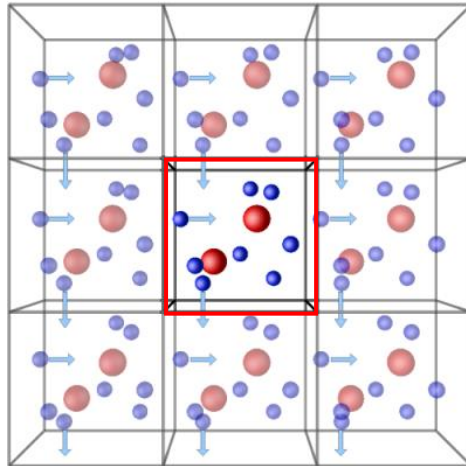
2.1.3. ΜΕΛΕΤΗ ΜΙΚΡΩΝ ΣΥΣΤΗΜΑΤΩΝ

Οι υπολογιστικές προσομοιώσεις αφορούν συνήθως σε έναν μικρό αριθμό μορίων, της τάξεως των δεκάδων χιλιάδων έως και εκατομμυρίων, ο οποίος εξαρτάται από την διαθέσιμη μνήμη του υπολογιστή και την ταχύτητα εκτέλεσης των υπολογισμών. Αν τα μόρια αυτά ανήκουν σε ένα σύστημα όπως είναι μια μικρή σταγόνα υγρού ή ένας μικροκρύσταλλος, οι συνεκτικές δυνάμεις μεταξύ των μορίων να επαρκούν για να τα συγκρατήσουν στο σύστημα. Στην περίπτωση όμως των υδάτινων όγκων, για να συγκρατηθούν τα μόρια θα ήταν απαραίτητη η εφαρμογή ενός δυναμικού που αναπαριστά το δοχείο το οποίο τα αποτρέπει από την μεταξύ τους απομάκρυνση. Αυτό, ωστόσο, θα σήμαινε ότι ένας μεγάλος αριθμός μορίων θα βρισκόταν στην επιφάνεια και θα δεχόταν διαφορετικές δυνάμεις από τον κύριο όγκο του υγρού.

Για την αντιμετώπιση αυτού του προβλήματος στις υπολογιστικές προσομοιώσεις εφαρμόζονται **περιοδικές οριακές συνθήκες** [Bor and von Karma 1912]. Έστω ότι τα μόρια του συστήματος είναι διατεταγμένα σε ένα χωρίο ή αλλιώς εμπεριέχονται σε ένα κυβικό χωρίο. Το χωρίο αυτό αναπαράγεται στο χώρο δημιουργώντας ένα άπειρο πλέγμα. Κατά την προσομοίωση, καθώς τα μόρια κινούνται στο πρωτεύον χωρίο, οι περιοδικές τους εικόνες σε καθένα από τα γειτονικά χωρία κινούνται ακριβώς με τον ίδιο τρόπο. Κατά συνέπεια, αν ένα μόριο φύγει από το πρωτεύον χωρίο, μια από τις περιοδικές του εικόνες θα εισέλθει σε αυτό από την απέναντι έδρα. Με αυτόν τον τρόπο δεν υπάρχει κάποιο τείχος στα όρια του πρωτεύοντος χωρίου και συνεπώς κανένα μόριο δεν βρίσκεται στην επιφάνεια αυτού. Μια αναπαράσταση κυβικού χωρίου με τις περιοδικές οριακές του εικόνες απεικονίζεται στην Εικόνα 2.3.

Το κυβικό χωρίο αποτελεί το πιο απλό γεωμετρικά χωρίο. Ωστόσο, και άλλα χωροπληρωτικά πολύεδρα, όπως το ρομβικό δωδεκάεδρο ή το κόλουρο οκτάεδρο χρησιμοποιούνται σε υπολογιστικές προσομοιώσεις καθώς, αν και πιο πολύπλοκα, έχουν μικρότερο όγκο σε σχέση με κυβικά χωρία ίδιας διάστασης.

Με την εφαρμογή των περιοδικών οριακών συνθηκών εγείρεται το θέμα του υπολογισμού των ιδιοτήτων του συστήματος, βάσει των αλληλεπιδράσεων των μορίων. Όπως προαναφέρθηκε, βασικό σημείο στη Μοριακή Δυναμική αποτελεί ο υπολογισμός της δυναμικής ενέργειας, η οποία λαμβάνει υπόψιν τις αλληλεπιδράσεις των μορίων με όλα τα υπόλοιπα μόρια. Στην περίπτωση των περιοδικών οριακών συνθηκών κάτι τέτοιο θα σήμαινε ότι για τον υπολογισμό της δυναμικής ενέργειας για ένα μόριο i του αρχικού κουτιού, θα λαμβάνονταν υπόψιν και οι αλληλεπιδράσεις του με τα μόρια όλων των περιοδικών εικόνων του κουτιού. Για την αποφυγή αυτού εφαρμόζεται το **όριο αποκοπής**, το οποίο ορίζεται από την απόσταση από τον κέντρο του μορίου πέρα από την οποία η αλληλεπίδρασή του με όλα τα υπόλοιπα μόρια θεωρείται ότι έχει μηδενική συνεισφορά στον υπολογισμό της δυναμικής ενέργειας, ισχύει δηλαδή: $V(r) = 0$ για $r > r_c$, όπου το r_c όριο αποκοπής.



Εικόνα 2.3. 3-διάστατη αναπαράσταση περιοδικού συστήματος πρωτεύοντος χωρίου προσομοίωσης (μέσον). Τα μόρια εισέρχονται και φεύγουν από τα κυτία προς όλες τις κατευθύνσεις.

2.2. ΚΛΑΣΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΜΗΧΑΝΙΚΗ

Η Στατιστική μηχανική είναι ο κλάδος της φυσικής που μελετά μακροσκοπικά συστήματα επί τη βάσει μικροσκοπικής ή μοριακής θεώρησης. Στόχος της Στατιστικής Μηχανικής είναι η κατανόηση και πρόβλεψη μακροσκοπικών φαινομένων, καθώς και ο υπολογισμός μακροσκοπικών ιδιοτήτων από τις ιδιότητες των μορίων ως οντότητες οι οποίες απαρτίζουν το μακροσκοπικό σύστημα.²⁴ Η θεωρία αυτή αναπτύχθηκε στο δεύτερο μισό του 19^{ου} αιώνα πρωτίστως από τον Josiah Willard Gibbs, ο οποίος πρότεινε την ιδιοφυή θεώρηση του στατιστικού συνόλου (Gibbs Ensemble), και τους Ludwig Eduard Boltzmann και James Clerk Maxwell.

2.2.1. ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΟΛΑ GIBBS

Το στατιστικό σύνολο εισήχθη ως έννοια από τον J. W. Gibbs στο σύγγραμμά του: Josiah Willard Gibbs (1902) *Elementary Principles in Statistical Mechanics*, New York και αποτέλεσε τη βάση της σύγχρονης στατιστικής θερμοδυναμικής. Σύμφωνα με τη θεώρηση Gibbs ένα μακροσκοπικό σύστημα θεωρείται ως μια ιδεατή συλλογή ενός συνόλου πολύ μεγάλου αριθμού, k , συστημάτων, καθένα από τα οποία αποτελεί ξεχωριστή μικροσκοπική κατάσταση (μικροκατάσταση, i , με ιδιοτιμή ενέργειας E_i , $i = 1, 2, \dots, k$), αλλά όλα υπόκεινται στους ίδιους μακροσκοπικούς περιορισμούς με αυτούς της μακροσκοπικής κατάστασης.

Μαθηματική περιγραφή του στατιστικού συνόλου.

Έστω σύστημα N αλληλεπιδρώντων μορίων.

Για κάθε μικροκατάσταση του συνόλου Gibbs, ορίζεται ένα διάνυσμα, (\mathbf{q}, \mathbf{p}) ,

όπου $\mathbf{q} \equiv (q_1, q_2, \dots, q_{3N})$ οι γενικευμένες συντεταγμένες (**χώρος απεικονίσεων**),

και $\mathbf{p} \equiv (p_1, p_2, \dots, p_{3N})$ οι γενικευμένες ορμές (**χώρος ορμών**).

Οι μικροκαταστάσεις του συστήματος μεταβάλλονται με τη πάροδο του χρόνου λόγω των μοριακών αλληλεπιδράσεων συνιστώντας έναν $6N$ -διάστατο διανυσματικό χώρο Γ , (**χώρος φάσεων**) στον οποίο κάθε μικροκατάσταση ορίζει ένα υπερδιάνυσμα: (\mathbf{q}, \mathbf{p}) .

Κάθε μικροκατάσταση εμφανίζεται στο Γ σύμφωνα με μια **συνάρτηση πυκνότητας πιθανότητας**, $\rho(\mathbf{q}, \mathbf{p}, t)$, η οποία σε σύστημα εν ισορροπία αναμένεται να εξαρτάται από το χρόνο, μόνο μέσω των \mathbf{q} και \mathbf{p} , δεδομένου ότι στην ισορροπία όλες οι μετρούμενες πειραματικά ιδιότητες ενός συστήματος παραμένουν χρονικά αμετάβλητες. Συνεπώς θα ισχύει $\rho_{ισορ.} = \rho(\mathbf{q}, \mathbf{p})$ και

$$\int_{\Gamma} \rho(\mathbf{q}, \mathbf{p}) d^{3N} q d^{3N} p = 1 \quad (2.2.1)$$

Επομένως, η τιμή μιας μηχανικής ιδιότητας του συστήματος στη ισορροπία (πχ πίεση) στη μικροκατάσταση i , (\mathbf{q}, \mathbf{p}) , αναμένεται η συνάρτηση $A(\mathbf{q}, \mathbf{p})$. Καθώς το σύστημα εξελίσσεται στο χώρο Γ σύμφωνα με θεμελιώδες αίτημα της στατιστικής μηχανικής κατα Gibbs: η πειραματικά μετρούμενη μακροσκοπική ιδιότητα $A_{πειρ.}$, δίδεται ως μέση τιμή όλων των $A(\mathbf{q}, \mathbf{p})$, ανάλογα με την συχνότητα εμφάνισης κάθε μικροκατάστασης στο **χώρο φάσεων**. Η υπόθεση αυτή εκφράζεται στην εξ. (2.2.2).

$$A_{πειρ.} \equiv \langle A \rangle_{\Gamma} = \int_{\Gamma} A(\mathbf{q}, \mathbf{p}) \rho_{ισορ.}(\mathbf{q}, \mathbf{p}) d^{3N} q d^{3N} p \quad (2.2.2)$$

Σε σύστημα εν ισορροπία σύμφωνα με το Εργοδικό Θεώρημα²⁵ η μέση τιμή στατιστικού συνόλου μιας ιδιότητας πρέπει να ισούται με τη μέση τιμή χρόνου αυτής, δηλαδή

$$\langle A \rangle_{\Gamma} = \langle A \rangle_t = \lim_{t \rightarrow \infty} \frac{1}{t} \int_{t_0}^{t+t_0} A(\mathbf{q}(t), \mathbf{p}(t)) dt \quad (2.2.3)$$

Η χρονική εξέλιξη του συστήματος καθορίζεται, σύμφωνα με την κλασική μηχανική, από τις εξισώσεις κίνησης του Νεύτωνα. Οι εξισώσεις αυτές αποτελούν ένα πεπλεγμένο σύστημα διαφορικών εξισώσεων δευτέρας τάξης και η αναλυτική τους επίλυση δεν είναι δυνατή λόγω του προβλήματος του μεγάλου αριθμού σωματίων που προαναφέρθηκε. Η

επίλυση τους πραγματοποιείται προσεγγιστικά μέσω αριθμητικών υπολογιστικών τεχνικών με τη χρήση ειδικών προγραμμάτων εφαρμογών σε ισχυρούς ηλεκτρονικούς υπολογιστές και συνήθως αφορά σε χιλιάδες σωματίων, έναντι των σωματίων της τάξης του 10^{23} που στην πραγματικότητα αντιπροσωπεύουν τη μακροσκοπική κατάσταση. Επιπλέον, η λήψη μέσων τιμών δεν είναι δυνατόν να πραγματοποιηθεί για άπειρο χρόνο, και έτσι επιλέγεται ένα μεγάλο χρονικό διάστημα, t . Στις προσομοιώσεις μοριακής δυναμικής οι εξισώσεις κίνησης επιλύονται βήμα προς βήμα για έναν αριθμό, τ , χρονικών βημάτων, με βήμα ολοκλήρωσης $\delta t = t / \tau$.

Κατά τη χρονική εξέλιξη του χώρου φάσεων κανένα σύστημα δεν καταστρέφεται ή δημιουργείται και, σύμφωνα με το θεώρημα του Liouville για την πυκνότητα πιθανότητας, προκύπτει $d\rho/dt = 0$, όπου d/dt η ολική παράγωγος ως προς το χρόνο. Για ένα σύστημα N ατόμων ισχύει:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \sum_i \dot{\mathbf{r}}_i \cdot \nabla_{\mathbf{r}_i} + \sum_i \dot{\mathbf{p}}_i \cdot \nabla_{\mathbf{p}_i} = \frac{\partial}{\partial t} + \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}_i} + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}_i} \quad (2.2.4)$$

όπου οι $\nabla_{\mathbf{r}_i}, \nabla_{\mathbf{p}_i}$ είναι παράγωγοι ως προς την θέση και την ορμή των ατόμων και $\dot{\mathbf{r}}_i, \dot{\mathbf{p}}_i$ οι χρονικές παράγωγοι της θέσης και της ορμής. Αν οριστεί ο συντελεστής Liouville, L , ως

$$iL = \left(\sum_i \dot{\mathbf{r}}_i \cdot \nabla_{\mathbf{r}_i} + \sum_i \dot{\mathbf{p}}_i \cdot \nabla_{\mathbf{p}_i} \right) = \dot{\mathbf{r}} \cdot \nabla_{\mathbf{r}_i} + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}_i} \quad (2.2.5)$$

Έτσι ώστε $d/dt = \partial/\partial t + iL$, τότε με χρήση του θεωρήματος του Liouville, η χρονική μεταβολή της πυκνότητας πιθανότητας των μικροκαταστάσεων δίνεται από τη σχέση:

$$\frac{\partial \rho_{ens}(\mathbf{\Gamma}, t)}{\partial t} = -iL \rho_{ens}(\mathbf{\Gamma}, t) \quad (2.2.6)$$

Η παραπάνω εξίσωση δηλώνει ότι ο ρυθμός μεταβολής της ρ_{ens} σε ένα συγκεκριμένο σημείο του χώρου φάσεων σχετίζεται με τις ροές μέσα και έξω από το σημείο αυτό. Η εξίσωση (2.6) έχει επίλυση:

$$\rho_{ens}(\mathbf{\Gamma}, t) = \exp(-iLt) \rho_{ens}(\mathbf{\Gamma}, 0) \quad (2.2.7)$$

στην οποία μπορεί να χρησιμοποιηθεί η σειρά: $\exp(-iLt) = 1 - iLt - \frac{1}{2} L^2 t^2 + \dots$

Και συνεπώς η εξίσωση της κίνησης μιας συνάρτησης όπως η $A(\mathbf{\Gamma})$ παίρνει τη μορφή:

$$A(\mathbf{\Gamma}(t)) = iLA(\mathbf{\Gamma}(t)) \quad (2.2.8)$$

ή

$$A(\Gamma(t)) = \exp(iLt)A(\Gamma(0)) \quad (2.2.9)$$

Αν η $\rho_{ens}(\Gamma)$ αντιπροσωπεύει στατιστικό σύνολο σε ισορροπία τότε η χρονική εξάρτηση εξαφανίζεται, $\partial\rho_{ens}/\partial t = 0$. Στην περίπτωση αυτή η πειραματικά μετρούμενη τιμή της ιδιότητας A περιγράφεται από την παρακάτω σχέση:

$$A_{\text{πειρ}} = \langle A \rangle_{ens} = \sum_{\Gamma} A(\Gamma) \rho_{ens}(\Gamma) \quad (2.2.10)$$

Η οποία αντιστοιχεί στην αντικατάσταση της μέσης χρονικής τιμής της εξίσωσης (2.2.3) από την μέση τιμή που λαμβάνεται από όλα τα σημεία του στατιστικού συνόλου «παγωμένα» σε μία χρονική στιγμή.

Η $\rho_{ens}(\Gamma)$ δύναται να γραφεί ως γινόμενο της συνάρτησης βάρους (weight function), $w_{ens}(\Gamma)$ η οποία ικανοποιεί τις παρακάτω εξισώσεις:

$$\rho_{ens}(\Gamma) = Q_{ens}^{-1} w_{ens}(\Gamma) \quad (2.2.11)$$

$$Q_{ens} = \sum_{\Gamma} w_{ens}(\Gamma) \quad (2.2.12)$$

$$\langle A \rangle_{ens} = \sum_{\Gamma} w_{ens}(\Gamma) A(\Gamma) / \sum_{\Gamma} w_{ens}(\Gamma) \quad (2.2.13)$$

Η συνάρτηση βάρους $w_{ens}(\Gamma)$ αποτελεί μη κανονικοποιημένη μορφή της $\rho_{ens}(\Gamma)$ και η συνάρτηση επιμερισμού (partition function, ή κανονικό άθροισμα καταστάσεων, sum over the states), Q_{ens} δρα ως παράγοντας κανονικοποίησης. Η συνάρτηση επιμερισμού παρέχει ένα μέτρο για το πλήθος των ενεργειακών σταθμών που μπορεί να καταλάβει το σύστημα για συγκεκριμένες θερμοδυναμικές μεταβλητές. Η συνάρτηση αυτή είναι βασική για τη στατιστική μηχανική καθώς μπορεί να παράξει κάθε μακροσκοπική ποσότητα του συστήματος. Η σύνδεσή της με την κλασική θερμοδυναμική γίνεται μέσω της σχέσης:

$$\Psi_{ens} = -\ln Q_{ens} \quad (2.2.14)$$

όπου Ψ_{ens} το θερμοδυναμικό δυναμικό το οποίο λαμβάνει την ελάχιστη τιμή του στη θερμοδυναμική ισορροπία.

Μια προσέγγιση για τον υπολογισμό των θερμοδυναμικών ποσοτήτων θα ήταν η άμεση εκτίμηση του Q_{ens} για ένα στατιστικό σύνολο μέσω της εξ. (2.2.12). Το άθροισμα, ωστόσο, όλων των πιθανών καταστάσεων δεν είναι εφικτό για συστήματα πολλών μορίων καθώς

υπάρχουν πάρα πολλές καταστάσεις, οι περισσότερες από τις οποίες αντιστοιχούν σε πολύ χαμηλή τιμή της συνάρτησης βάρους. Οι χαμηλές τιμές της συνάρτησης βάρους οφείλονται σε μη φυσιολογικές επικαλύψεις πυρήνων μορίων. Είναι επιθυμητή η εξαίρεση των καταστάσεων αυτών με σκοπό στο άθροισμα καταστάσεων να λαμβάνονται μόνο αυτές με υψηλή πιθανότητα. Η ιδέα της παραγωγής μιας ομάδας καταστάσεων του χώρου φάσεων τα οποία λαμβάνονται από ολόκληρο το σύνολο σε συμφωνία με την πυκνότητα κατανομής $\rho_{ens}(\Gamma)$ αποτελεί τη βάση στις προσομοιώσεις Monte Carlo.

Στη μοριακή δυναμική η μέση τιμή του στατιστικού συνόλου της εξ. (2.2.13) αντικαθίσταται από τη μέση τιμή της τροχιάς. Οι εξισώσεις κίνησης του Νεύτωνα παράγουν μια σειρά καταστάσεων οι οποίες είναι σύμφωνες με τη συνάρτηση κατανομής ρ_{NVE} για το σταθερό - NVE ή αλλιώς μικροκανονικό σύνολο. Η μετάβαση από το μικροκανονικό σύνολο σε άλλο στατιστικό σύνολο πραγματοποιείται υπό συγκεκριμένες προϋποθέσεις. Σε αυτές τις περιπτώσεις, η μέση τιμή του στατιστικού συνόλου είναι ίση με ένα είδος μέσης χρονικής τιμής που δίνεται από τη σχέση:

$$A_{\text{πειρ}} = \langle A \rangle_{\text{time}} = \frac{1}{\tau_{\text{πειρ}}} \sum_{\tau=1}^{\tau_{\text{πειρ}}} A(\Gamma(\tau)) \quad (2.2.15)$$

2.2.2. ΣΥΝΗΘΗ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΟΛΑ

Θεωρούνται τέσσερα συνήθη στατιστικά σύνολα: το μικροκανονικό ή αλλιώς σταθερό - NVE, το κανονικό ή αλλιώς σταθερό - NVT, το ισόθερμο - ισοβαρές ή αλλιώς σταθερό - NPT και το μεγάλο κανονικό ή αλλιώς σταθερό - μVT κανονικό σύνολο. Για κάθε σύστημα οι θερμοδυναμικές μεταβλητές που αναφέρονται είναι σταθερές. Οι υπόλοιπες θερμοδυναμικές ποσότητες ορίζονται από τις μέσες τιμές του συνόλου και για κάθε συγκεκριμένη κατάσταση οι στιγμιαίες τιμές της της κατάλληλης συνάρτησης εξάγονται από την αντίστοιχη μέση τιμή.

Μικροκανονικό Στατιστικό Σύνολο (Microcanonical, NVE)

Το μικροκανονικό στατιστικό σύνολο, που προαναφέρθηκε, αποτελείται από συστήματα τα οποία είναι απομονωμένα από το περιβάλλον τους και συνεπώς έχουν τελείως καθορισμένη ολική ενέργεια, E , αφού δεν δύνανται να ανταλλάξουν ενέργεια με το περιβάλλον τους, και καθορισμένο πλήθος σωματιδίων, N , αφού είναι απομονωμένα. Επιπλέον, τα συστήματα του συνόλου έχουν τον ίδιο όγκο, V .

Με βάση το μικροκανονικό σύνολο διατυπώνεται το βασικό αίτημα της Στατιστικής μηχανικής που ονομάζεται **Αίτημα των a priori ίσων πιθανοτήτων: Στα συστήματα**

που αποτελούν ένα μικροκανονικό σύνολο σε κατάσταση θερμοδυναμικής ισορροπίας θα αντιπροσωπεύονται όλες οι δυνατές προσιτές καταστάσεις του με την ίδια πιθανότητα.²⁶

$$\rho_{NVE}(\Gamma) = \begin{cases} \text{σταθ.}, & \text{αν } E - \delta E < \mathcal{H}(\Gamma) < E \\ 0 & , \text{σε κάθε άλλη περίπτωση} \end{cases}$$

Καθώς η ενέργεια του μικροκανονικού συστήματος σε ισορροπία έχει σταθερή τιμή μέσα σε μια περιοχή $E \pm \delta E$.

Η πυκνότητα κατανομής για το μικροκανονικό σύνολο είναι ανάλογη του:

$$\delta(\mathcal{H}(\Gamma) - E)$$

όπου η Γ αναπαριστά τις θέσεις και τις ορμές των σωματιδίων και η $\mathcal{H}(\Gamma)$ είναι η Χαμιλτονιανή. Η συνάρτηση δ κάνει επιλογή των σωματιδίων που έχουν την επιθυμητή τιμή E , από τα N σωματίδια που περιέχονται στον όγκο V . Όταν οι καταστάσεις είναι διακριτές τότε η δ αντιστοιχεί στο δέλτα του Κρόνεκερ και λαμβάνει τιμές 0 και 1, ενώ όταν είναι συνεχείς αντιστοιχεί στη συνάρτηση Dirac. Η συνάρτηση επιμερισμού για μικροκανονικό σύστημα γράφεται ως:

$$Q_{NVE} = \sum_{\Gamma} \delta(\mathcal{H}(\Gamma) - E) \quad (2.2.16)$$

Η ιονοει-κλασική έκφραση της παραπάνω συνάρτησης για ένα ατομικό σύστημα είναι η:

$$Q_{NVE} = \frac{1}{N!} \frac{1}{h^{3N}} \int d\mathbf{q} d\mathbf{p} \delta(\mathcal{H}(\mathbf{q}, \mathbf{p}) - E) \quad (2.2.17)$$

Όπου h , η σταθερά του Plank.

Το κατάλληλο θερμοδυναμικό δυναμικό είναι το αρνητικό της εντροπίας S ,

$$-S / k_B = -\ln Q_{NVE} \quad (2.2.18)$$

όπου k_B η σταθερά του Boltzmann.

Κανονικό Στατιστικό Σύνολο (Canonical, NVT)

Το κανονικό σύνολο αντιπροσωπεύει κλειστά συστήματα με διαθερμικά τοιχώματα, τα οποία έχουν σταθερό αριθμό σωματιδίων N , όγκο V και θερμοκρασία T . Στο σύνολο αυτό τα συστήματα δεν έχουν την ίδια ενέργεια καθώς πραγματοποιείται ανταλλαγή ενέργειας μεταξύ του συστήματος και ενός πολύ μεγαλύτερου συστήματος σταθερής

θερμοκρασίας T με το οποίο έρχεται σε επαφή. Έτσι η ενέργεια διακυμαίνεται γύρω από τη μέση ενέργεια, $\langle E \rangle$, του συστήματος. Η συσχέτιση του μικροκανονικού συνόλου με το κανονικό επιτρέπει το μετασχηματισμό των αποτελεσμάτων του πρώτου σε αυτά του δεύτερου.

Η πυκνότητα κατανομής για το κανονικό σύνολο είναι ανάλογη του:

$$\exp(-\mathcal{H}(\Gamma)/k_B T)$$

Η συνάρτηση επιμερισμού είναι η:

$$Q_{NVT} = \sum_{\Gamma} \exp(-\mathcal{H}(\Gamma)/k_B T) \quad (2.2.19)$$

και η οιονεί-κλασική μορφή της είναι η:

$$Q_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} \int d\mathbf{q} d\mathbf{p} \exp(-\mathcal{H}(\mathbf{q}, \mathbf{p})/k_B T) \quad (2.2.20)$$

Η κατάλληλη θερμοδυναμική συνάρτηση είναι η ελεύθερη ενέργεια Helmholtz A :

$$A/k_B T = -\ln Q_{NVT} \quad (2.2.21)$$

Ισόθερμο-Ισόβαρες Στατιστικό Σύνολο (Isothermal-Isobaric, NPT)

Το σύνολο αυτό αποτελείται από συστήματα τα οποία έχουν καθορισμένο πλήθος σωματιδίων N , πίεση P και θερμοκρασία T . Χρησιμοποιείται σε υπολογιστικές μελέτες ιδιαίτερα όταν αντικείμενο της μελέτης είναι οι ιδιότητες υπό σταθερή πίεση. Η συνάρτηση πυκνότητας πιθανότητας του Ισόθερμου - Ισοβαρούς στατιστικού συνόλου είναι ανάλογη του:

$$\exp(-(\mathcal{H} + PV)/k_B T)$$

Σημειώνεται ότι η μέση τιμή της ποσότητας μέσα στην εκθετική συνάρτηση δίνει την θερμοδυναμική ενθαλπία: $H = \langle \mathcal{H} \rangle + P \langle V \rangle$. Η συνάρτηση επιμερισμού είναι η:

$$Q_{NPT} = \sum_{\Gamma} \sum_V \exp(-(\mathcal{H} + PV)/k_B T) = \sum_V \exp(-PV/k_B T) \quad (2.2.22)$$

και η οιονεί-κλασική μορφή της είναι η:

$$Q_{NPT} = \frac{1}{N!} \frac{1}{h^{3N}} \frac{1}{V_0} \int dV \int d\mathbf{q} d\mathbf{p} \exp(-(\mathcal{H} + PV)/k_B T) \quad (2.2.23)$$

όπου V_0 μια βασική μονάδα όγκου η οποία επιλέγεται έτσι ώστε να κάνει το Q_{NPT} αδιάστατο.

Η θερμοδυναμική συνάρτηση που αντιστοιχεί στο ισόθερμο-ισοβαρές σύνολο είναι η ελεύθερη ενέργεια Gibbs, G :

$$G/k_B T = -\ln Q_{NPT} \quad (2.2.24)$$

Μεγάλο Κανονικό Στατιστικό Σύνολο (Grand Canonical, μVT)

Το σύνολο αυτό αποτελείται από ανοιχτά θερμοδυναμικά συστήματα τα οποία ανταλλάσσουν ενέργεια και ύλη με το περιβάλλον τους. Τα συστήματα αυτά έχουν το ίδιο χημικό μηχανικό μ , τον ίδιο όγκο V , και την ίδια θερμοκρασία T . Η συνάρτηση πυκνότητας για το μεγάλο κανονικό στατιστικό σύνολο είναι ανάλογη του:

$$\exp(-(\mathcal{H} + \mu N)/k_B T)$$

Η αντίστοιχη συνάρτηση επιμερισμού είναι η:

$$Q_{\mu VT} = \sum_{\Gamma} \sum_N \exp(-(\mathcal{H} + \mu N)/k_B T) = \sum_N \exp(\mu N/k_B T) \quad (2.2.25)$$

και η οιονεί-κλασική μορφή της είναι η:

$$Q_{\mu VT} = \sum_N \frac{1}{N!} \frac{1}{h^{3N}} \exp(\mu N/k_B T) \int d\mathbf{r} d\mathbf{p} \exp(-\mathcal{H}/k_B T) \quad (2.2.26)$$

Η κατάλληλη θερμοδυναμική συνάρτηση είναι η:

$$-PV/k_B T = -\ln Q_{NPT} \quad (2.2.27)$$

2.3. ΑΛΓΟΡΙΘΜΟΙ ΜΟΡΙΑΚΩΝ ΠΡΟΣΟΜΟΙΩΣΕΩΝ

Ο υπολογισμός της τροχιάς των σωματιδίων ενός συστήματος υλοποιείται, όπως προαναφέρθηκε, με τη χρήση αλγορίθμων σε ισχυρούς υπολογιστές. Για το σκοπό αυτό έχει κατασκευαστεί πλήθος αλγορίθμων, οι οποίοι διαφοροποιούνται ανάλογα με τις ιδιαιτερότητες του κάθε συστήματος. Στις ενότητες που ακολουθούν παρουσιάζεται η βάση των αλγορίθμων της μοριακής δυναμικής και στη συνέχεια αναλύονται οι αλγόριθμοι που χρησιμοποιούνται στην προσομοίωση της παρούσας εργασίας. Παρουσιάζεται, επίσης, ο αλγόριθμος για την ενεργειακή ελαχιστοποίηση του συστήματος, ο οποίος, αν και δεν αποτελεί αλγόριθμο μοριακής δυναμικής, είναι απαραίτητος για την προσομοίωση του συστήματος.

2.3.1. ΠΕΡΙΓΡΑΦΗ ΑΛΓΟΡΙΘΜΩΝ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ

Στην Μοριακή Δυναμική επιχειρείται η επίλυση των διαφορικών εξισώσεων κίνησης του Νεύτωνα στο χρόνο. Αυτή πραγματοποιείται μέσω μεθόδων πεπερασμένων διαφορών. Έχοντας γνωστές τις θέσεις και τις ταχύτητες των σωματιδίων σε μια χρονική στιγμή t , επιχειρείται ο προσδιορισμός των θέσεων και των ταχυτήτων σε μια επόμενη χρονική στιγμή $(t + \delta t)$ με επαρκή ακρίβεια. Οι εξισώσεις επιλύονται βήμα προς βήμα και επιλογή του χρονικού διαστήματος, δt , εξαρτάται από τη μέθοδο που χρησιμοποιείται και οφείλει να είναι αρκετά μικρότερο του χρόνου που χρειάζεται ένα σωματίδιο για να διανύσει το μήκος του, περίπου $10^{-14} s$.²²

Όλοι οι αλγόριθμοι ολοκλήρωσης των διαφορικών εξισώσεων προσεγγίζουν τις θέσεις, τις ταχύτητες και τις επιταχύνσεις των σωματιδίων στον χρόνο $(t + \delta t)$ μέσω των **σειρών Taylor**:

$$\begin{aligned}\mathbf{r}^p(t + \delta t) &= \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \frac{1}{6} \delta t^3 \mathbf{b}(t) + \dots \\ \mathbf{v}^p(t + \delta t) &= \mathbf{v}(t) + \delta t \mathbf{a}(t) + \frac{1}{2} \delta t^2 \mathbf{b}(t) + \dots \\ \mathbf{a}^p(t + \delta t) &= \mathbf{a}(t) + \delta t \mathbf{b}(t) + \dots \\ \mathbf{b}^p(t + \delta t) &= \mathbf{b}(t) + \dots\end{aligned}\tag{2.3.1}$$

Οι τιμές που προκύπτουν αποτελούν τις «προβλεπόμενες» τιμές, οι οποίες στη συνέχεια διορθώνονται. Οι \mathbf{r} αντιπροσωπεύουν τις θέσεις, οι \mathbf{v} τις ταχύτητες, οι \mathbf{a} τις επιταχύνσεις και \mathbf{b} όλες τις παραγώγους τρίτης τάξης του \mathbf{r} . Οι εξισώσεις (3.1) δεν αρκούν για να παραχθούν οι σωστές τροχιές με την πάροδο του χρόνου, καθώς δεν περιέχουν τις εξισώσεις κίνησης. Οι τελευταίες εισάγονται στον αλγόριθμο της Μοριακής Δυναμικής μέσω ενός **βήματος διόρθωσης**: Από τις νέες θέσεις \mathbf{r}^p υπολογίζονται οι

δυνάμεις στο χρόνο $(t + \delta t)$ και από αυτές οι σωστές επιταχύνσεις $\mathbf{a}^c(t + \delta t)$. Οι \mathbf{a}^c συγκρίνονται στη συνέχεια με τις επιταχύνσεις που προβλέπονται από τις εξ. (3.1) και εκτιμάται το σφάλμα του βήματος πρόβλεψης:

$$\Delta \mathbf{a}(t + \delta t) = \mathbf{a}^c(t + \delta t) - \mathbf{a}^p(t + \delta t) \quad (2.3.2)$$

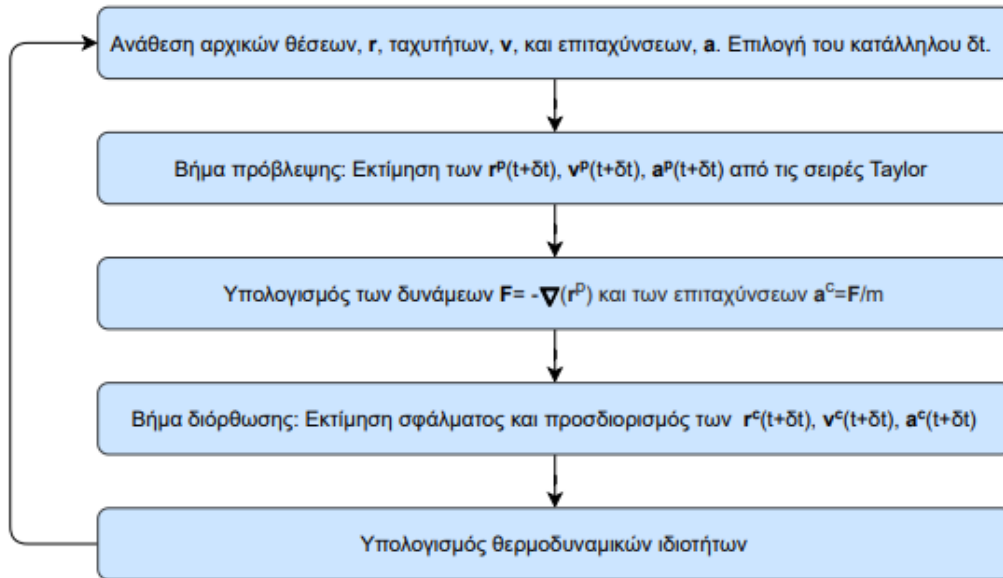
Το σφάλμα και τα αποτελέσματα του βήματος πρόβλεψης εισάγονται στο βήμα διόρθωσης:

$$\begin{aligned} \mathbf{r}^c(t + \delta t) &= \mathbf{r}^p(t + \delta t) + c_0 \Delta \mathbf{a}(t + \delta t) \\ \mathbf{v}^c(t + \delta t) &= \mathbf{v}^p(t + \delta t) + c_1 \Delta \mathbf{a}(t + \delta t) \\ \mathbf{a}^c(t + \delta t) &= \mathbf{a}^p(t + \delta t) + c_2 \Delta \mathbf{a}(t + \delta t) \\ \mathbf{b}^c(t + \delta t) &= \mathbf{b}^p(t + \delta t) + c_3 \Delta \mathbf{a}(t + \delta t) \end{aligned} \quad (2.3.3)$$

όπου $\mathbf{r}^c(t + \delta t)$ κλπ. αποτελούν καλύτερες εκτιμήσεις των πραγματικών θέσεων, ταχυτήτων κλπ. Οι συντελεστές $c_0, c_1, c_2, c_3 \dots$ δίνονται από τον αλγόριθμο Gear. Η δομή ενός αλγορίθμου μοριακής δυναμικής περιλαμβάνει, συνεπώς, τα εξής βήματα:

1. Πρόβλεψη των θέσεων και ταχυτήτων κλπ. για χρόνο $(t + \delta t)$ μέσω των σειρών Taylor
2. Εκτίμηση των δυνάμεων και από αυτές των επιταχύνσεων, από τις νέες θέσεις
3. Διόρθωση των νέων θέσεων που προβλέφθηκαν μέσω των διορθωμένων επιταχύνσεων
4. Υπολογισμός των επιθυμητών θερμοδυναμικών ιδιοτήτων, που θα χρησιμοποιηθούν για τον υπολογισμό των χρονικών μέσων όρων.

Η δομή αυτή αποτελεί τη βάση των αλγορίθμων μοριακής δυναμικής και παρουσιάζεται στην Εικόνα 2.4. Οι αλγόριθμοι αυτοί διαφοροποιούνται στα διάφορα μέρη τους ανάλογα με τις ανάγκες της κάθε προσομοίωσης. Στις παραγράφους που ακολουθούν παρουσιάζονται κάποιοι ολοκληρωτές, που χρησιμοποιούνται για την ολοκλήρωση των εξισώσεων κίνησης στους αλγορίθμους μοριακής δυναμικής, και ο αλγόριθμος Berendsen ο οποίος χρησιμοποιήθηκε στην παρούσα εργασία.



Εικόνα 2.4. Διάγραμμα ροής Αλγορίθμου Μοριακής Δυναμικής.

2.3.2. Ο ΟΛΟΚΛΗΡΩΤΗΣ VERLET

Ο ολοκληρωτής αυτός αποτελεί την πιο διαδεδομένη μέθοδο ολοκλήρωσης των εξισώσεων κίνησης. Εκτελεί άμεση επίλυση των διαφορικών εξισώσεων δευτέρου βαθμού: $m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i$ και βασίζεται στις θέσεις $\mathbf{r}(t)$, τις επιταχύνσεις $\mathbf{a}(t)$ και τις θέσεις $\mathbf{r}(t - \delta t)$ του προηγούμενου βήματος:²²

$$\mathbf{r}(t + \delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \delta t) + \delta t^2 \mathbf{a}(t) \quad (2.3.4)$$

Στην παραπάνω εξίσωση δεν εμφανίζονται οι ταχύτητες, οι οποίες έχουν εξαλειφθεί από τα αθροίσματα που προκύπτουν από τις σειρές Taylor, καθώς:

$$\begin{aligned} \mathbf{r}(t + \delta t) &= \mathbf{r}(t) + \delta t \mathbf{v}(t) + (1/2 \delta t^2) \mathbf{a}(t) + \dots \\ \mathbf{r}(t - \delta t) &= \mathbf{r}(t) - \delta t \mathbf{v}(t) + (1/2 \delta t^2) \mathbf{a}(t) - \dots \end{aligned} \quad (2.3.5)$$

Οι ταχύτητες, οι οποίες δεν είναι αναγκαίες για τον υπολογισμό των ταχυτήτων, αλλά απαιτούνται για την εκτίμηση της κινητικής ενέργειας, υπολογίζονται ως εξής:

$$\mathbf{v}(t) = \frac{\mathbf{r}(t + \delta t) - \mathbf{r}(t - \delta t)}{2\delta t} \quad (2.3.6)$$

Σημειώνεται ότι ο αλγόριθμος με αυτήν την μορφή είναι κεντραρισμένος (λόγω της συμμετρίας των $\mathbf{r}(t + \delta t)$, $\mathbf{r}(t - \delta t)$) και έτσι είναι χρονικά αντιστρέψιμος. Επιπλέον, η προέλαση των θέσεων γίνεται σε ένα βήμα, σε αντίθεση με τα δύο βήματα πρόβλεψης-διόρθωσης του αλγορίθμου που προαναφέρθηκε, με αποτέλεσμα και ο κώδικας του Verlet να φέρει τις αντίστοιχες διαφοροποιήσεις.

2.3.3. Ο ΟΛΟΚΛΗΡΩΤΗΣ LEAP-FROG

Ο ολοκληρωτής leap-frog, επονομαζόμενος «ολοκληρωτής μισού-βήματος», αποτελεί τροποποίηση του Verlet και γράφεται ως:²²

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) - \delta t \mathbf{v}\left(t + \frac{1}{2} \delta t\right) \quad (2.3.7)$$

$$\mathbf{v}\left(t + \frac{1}{2} \delta t\right) = \mathbf{v}\left(t - \frac{1}{2} \delta t\right) + \delta t \mathbf{a}(t) \quad (2.3.8)$$

Οι ποσότητες που αποθηκεύονται είναι οι τρέχουσες θέσεις, $\mathbf{r}(t)$ και επιταχύνσεις $\mathbf{a}(t)$ μαζί με τις ταχύτητες του ενδιάμεσου βήματος $\mathbf{v}(t - 1/2\delta t)$. Αρχικά εφαρμόζεται η εξίσωση των ταχυτήτων (3.8) και μέσω των συντεταγμένων υπολογίζονται οι ταχύτητες για το επόμενο ενδιάμεσο βήμα, $\mathbf{v}(t + 1/2\delta t)$. Σε αυτό το βήμα, οι τρέχουσες ταχύτητες υπολογίζονται ως:

$$\mathbf{v}(t) = \frac{1}{2} \left(\left(\mathbf{v}(t) + \frac{1}{2} \delta t \right) + \mathbf{v}\left(t - \frac{1}{2} \delta t\right) \right) \quad (2.3.9)$$

Η τελευταία εξίσωση χρησιμοποιείται για τον υπολογισμό της ενέργειας και οποιωνδήποτε άλλων ποσοτήτων την χρονική στιγμή t . Στη συνέχεια, χρησιμοποιείται η εξίσωση (3.7) για τον υπολογισμό των θέσεων, που θα βρεθούν πάλι μπροστά από τις ταχύτητες, και υπολογίζονται οι νέες επιταχύνσεις.

2.3.4. Ο ΑΛΓΟΡΙΘΜΟΣ BERENDSEN

Στον αλγόριθμο αυτό, πραγματοποιείται σταθεροποίηση της θερμοκρασίας ή της πίεσης του συστήματος, όχι με τροποποίηση της Χαμιλτονιανής, αλλά με ασθενή σύζευξη με εξωτερικό λουτρό, το οποίο βρίσκεται στην επιθυμητή θερμοκρασία ή πίεση αντίστοιχα. Χρησιμοποιείται η αρχή της ελάχιστης τοπικής διαταραχής σε συμφωνία με την απαιτούμενη συνολική σύζευξη. Αποτελεί μια προσέγγιση η οποία πλησιάζει τα ιδανικά φυσικά συστήματα που δεν είναι σε ισορροπία.²⁷

Αρχικά, θεωρείται σύζευξη με λουτρό θερμότητας το οποίο βρίσκεται σε σταθερή θερμοκρασία, T_0 . Για την επίτευξη της σύζευξης, εισάγεται στις εξισώσεις κίνησης ο στοχαστικός όρος $R(t)$ και ο όρος της τριβής $m_i \gamma_i u_i$, έτσι ώστε να προκύπτει μια εξίσωση Langevin:

$$m_i \dot{u}_i = F_i - m_i \gamma_i u_i + R(t) \quad (2.3.10)$$

όπου F_i η δύναμη που ασκείται στο σωματίδιο i από τα υπόλοιπα σωματίδια, γ_i η σταθερά τριβής που ασκείται στο σωματίδιο i λόγω της επαφής του με τα υπόλοιπα, και $R(t)$ η Gaussian στοχαστική μεταβλητή, η οποία έχει μηδενική μέση τιμή και ισχύει:

$$\langle R_i(t) R_j(t + \tau) \rangle = 2m_i \gamma_i k T_0 \delta(\tau) \delta_{ij} \quad (2.3.11)$$

Η παραπάνω εξίσωση αντιπροσωπεύει τις συχνές συγκρούσεις ελαφριών σωματιδίων τα οποία απαρτίζουν ιδανικό αέριο σε θερμοκρασία T_0 . Θεωρείται ίδια σταθερά τριβής για όλα τα σωματίδια $\gamma_i = \gamma$. Ο προσδιορισμός της χρονική εξάρτησης της θερμοκρασίας T γίνεται μέσω της ολική κινητικής ενέργειας, E_k :

$$\frac{dE_k}{dt} = \lim_{\delta t \rightarrow 0} \left[\left\{ \sum_{i=1}^{3N} \frac{1}{2} m_i u_i^2(t + \delta t) - \sum_{i=1}^{3N} \frac{1}{2} m_i u_i^2(t) \right\} / \delta t \right] \quad (2.3.12)$$

όπου N ο αριθμός των σωματιδίων. Η παραπάνω εξίσωση, με δεδομένο ότι η $R_i(t')$ είναι μη συσχετισμένη με την $u_i(t)$ και την $R_i(t)$ αντιστοιχεί στην:

$$\frac{dE_k}{dt} = \sum_{i=1}^{3N} u_i F_i + 2\gamma \left(\frac{3N}{2} k T_0 - E_k \right) \quad (2.3.13)$$

Ο πρώτος όρος στην παραπάνω εξίσωση ισούνται με το αρνητικό της χρονικής παραγώγου της δυναμικής ενέργειας και ο δεύτερος περιγράφει τη συνολική σύζευξη με

το λουτρό θερμότητας. Σε όρους θερμοκρασίας ο δεύτερος όρος της παραπάνω εξίσωσης εκφράζεται ως:

$$\left(\frac{dT}{dt}\right)_{\text{λουτρού}} = 2\gamma(T_0 - T) \quad (2.3.14)$$

Σημειώνεται ότι η χρονική σταθερά τ_T της παρούσας σύζευξης ισούται με $(2\gamma)^{-1}$. Συνεπώς η εξίσωση κίνησης για συνολική σύζευξη θερμοκρασίας γίνεται:

$$m_i \dot{u}_i = F_i + m_i \gamma_i \left(\frac{T_0}{T} - 1\right) u_i \quad (2.3.15)$$

Η εξίσωση αυτή αντιπροσωπεύει μία αναλογική στάθμιση των ταχυτήτων ανά χρονικό βήμα του αλγορίθμου από u σε λu με:

$$\lambda = 1 + \frac{\delta t}{2\tau_T} \left(\frac{T_0}{T} - 1\right) \quad (2.3.16)$$

Η μεταβολή της θερμοκρασίας ανά βήμα μπορεί να γίνει ακριβώς ίση με $(T_0 - T) \delta\tau / \tau_T$, δίνοντας:

$$\lambda = \left[1 + \frac{\delta t}{\tau_T} \left(\frac{T_0}{T} - 1\right) \right]^{1/2} \quad (2.3.17)$$

Με την ίδια λογική πραγματοποιείται και η σύζευξη με λουτρό σταθερής πίεσης: ένας επιπλέον όρος θα προστεθεί στην εξίσωση κίνησης ώστε να επηρεάζει την μεταβολή της πίεσης:

$$\left(\frac{dP}{dt}\right)_{\text{λουτρού}} = \frac{P_0 - P}{\tau_P} \quad (2.3.18)$$

όπου η πίεση δίνεται από το θεώρημα virial:

$$P = \frac{2}{3V} (E_K - \mathcal{E}) \quad (2.3.19)$$

και \mathcal{E} δίνεται από την εξίσωση:

$$\mathcal{E} = -\frac{1}{2} \sum_{i < j} \mathbf{r}_{ij} \mathbf{F}_{ij} \quad (2.3.20)$$

$$\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j \quad (2.3.21)$$

Στην οποία \mathbf{F}_{ij} είναι η δύναμη που ασκείται στο σωματίδιο i από σωματίδιο j . Οι ενδομοριακές αλληλεπιδράσεις δεν συνεισφέρουν στις εξ. (2.3.19), (2.3.20), με αποτέλεσμα να μπορούν να προσδιοριστούν με χρήση μόνο των κέντρων μάζας των συντεταγμένων, ταχυτήτων και δυνάμεων. Για την ελαχιστοποίηση τοπικών διαταραχών χρησιμοποιείται μια απλή αναλογική στάθμιση συντεταγμένων, συνακόλουθη της στάθμισης του όγκου. Έτσι εισάγεται ένας νέος όρος, στην εξίσωση: $\dot{x} = u$:

$$\dot{x} = u + ax \quad (2.3.22)$$

Ο όγκος μεταβάλλεται αναλόγως:

$$\dot{V} = 3aV \quad (2.3.23)$$

Η μεταβολή της πίεσης σχετίζεται με την ισόθερμη συμπιεστότητα β :

$$\frac{dP}{dt} = -\frac{1}{\beta V} \frac{dV}{dt} = -\frac{3a}{\beta} \quad (2.3.24)$$

Και με βάση την εξ. (3.18) καθορίζεται το α :

$$\alpha = -\beta(P_0 - P) / 3\tau_p \quad (2.3.25)$$

Συνεπώς, η τροποποιημένη εξίσωση της κίνησης που προκύπτει είναι η:

$$\dot{x} = u - \frac{\beta(P_0 - P)}{3\tau_p} x \quad (2.3.26)$$

Η εξίσωση αυτή αντιπροσωπεύει μία αναλογική στάθμιση των συντεταγμένων και του μήκους του κουτιού l ανά χρονικό βήμα του αλγορίθμου από x σε μx και από l σε μl :

$$\mu = 1 - \frac{\beta \delta t (P_0 - P)}{3\tau_p} \quad (2.3.27)$$

ή ισοδύναμα,

$$\mu = \left[1 - \frac{\delta t (P_0 - P)}{\tau_p} \right]^{1/3} \quad (2.3.28)$$

Οι όροι στάθμισης των ταχυτήτων και των συντεταγμένων επιλύονται σε αλγόριθμο που χρησιμοποιεί ολοκληρωτή leap-frog για την επίλυση των εξισώσεων κίνησης. Επιβάλλονται περιορισμοί σύμφωνα με τους οποίους η χωρίς περιορισμούς διάταξη x μετατρέπεται σε περιορισμένη διάταξη y με μετατόπιση των διανυσμάτων σε μια

κατεύθυνση η οποία δίνεται από την διάταξη αναφοράς x_{ref} . Αυτό συμβολίζεται ως $[x \rightarrow y; x_{ref}]$.

Δεδομένων των θέσεων $\mathbf{r}(t)$ και των ταχυτήτων $\mathbf{v}(t-1/2\delta t)$ όλων των ατόμων, του μήκους και κουτιού $l(t)$ και όγκου $V(t)$, για ισοτροπικά συστήματα, παρουσιάζονται τα βήματα του αλγορίθμου:

Βήμα 1. Προσδιορισμός των δυνάμεων $\mathbf{F}(t)$ σε όλα τα άτομα: $\mathbf{a}_i(t) = \mathbf{F}_i(t) / \mathbf{m}_i$.

Βήμα 2. Σύζευξη Πίεσης.

Προσδιορίζεται η πίεση, μέσω της virial Ξ και η κινητική ενέργεια, μέσω των κέντρων μάζας:

$$\Xi(t) = \frac{1}{2} \sum_{\alpha < \beta} \mathbf{R}_{\alpha\beta}(t) \mathbf{F}_{\alpha\beta}(t) \quad (2.3.29)$$

όπου $\mathbf{R}_{\alpha\beta} = \mathbf{R}_\alpha - \mathbf{R}_\beta$ είναι η σχετική θέση του διανύσματος των κέντρων μάζας των μορίων α και β και $\mathbf{F}_{\alpha\beta} = \sum_{i\alpha j\beta} F_{i\alpha j\beta}^{pair}$ είναι το άθροισμα όλων των αλληλεπιδράσεων ζεύγους των ατόμων $i\alpha$ του μορίου α με τα άτομα $j\beta$ του μορίου β . Οι αλληλεπιδράσεις που δεν αντιστοιχούν σε ζεύγη (γωνίες, δίεδρα) παραλείπονται.

$$E_{κιν}^{κμ} \left(t - \frac{1}{2} \delta t \right) = \sum_{\alpha} \frac{1}{2} M_{\alpha} V_{\alpha}^2 \left(t - \frac{1}{2} \delta t \right), \quad (2.3.30)$$

όπου M_{α} η συνολική μάζα του μορίου α , $V_{\alpha} = M_{\alpha}^{-1} \sum_i m_{i\alpha} v_{i\alpha}$ το κέντρο της μάζας ταχυτήτων του μορίου α .

$$P(t) = \frac{1}{3V(t)} \left[E_{κιν}^{κμ} \left(t - \frac{1}{2} \delta t \right) - \Xi(t) \right] \quad (2.3.31)$$

ο παράγοντας στάθμισης της πίεσης γίνεται:

$$\mu = \left\{ 1 + \frac{\delta t}{\tau_p} \beta [P(t) - P_0] \right\}^{1/3} \quad (2.3.32)$$

Βήμα 3. Σύζευξη Θερμοκρασίας.

Προσδιορίζεται η θερμοκρασία:

$$T\left(t - \frac{1}{2}\delta t\right) = \frac{2k}{3N - M - 3} E_{\kappa iv}\left(t - \frac{1}{2}\delta t\right), \quad (2.3.33)$$

όπου N ο αριθμός των ατόμων, M ο αριθμός των περιορισμών, και

$$E_{\kappa iv}\left(t - \frac{1}{2}\delta t\right) = \sum \frac{1}{2} m_i u_i^2\left(t - \frac{1}{2}\delta t\right) \quad (2.3.34)$$

και ο παράγοντας στάθμισης της θερμοκρασίας γίνεται:

$$\lambda = \left[1 + \frac{\delta t}{\tau_T} \left\{ \frac{T_0}{T\left(t - \frac{1}{2}\delta t\right)} - 1 \right\} \right]^{1/2} \quad (2.3.35)$$

Βήμα 4. Υπολογισμός ταχυτήτων από την εξ. (2.3.8):

$$\mathbf{v}\left(t + \frac{1}{2}\delta t\right) = \mathbf{v}\left(t - \frac{1}{2}\delta t\right) + \delta t \mathbf{a}(t)$$

Βήμα 5. Σύζευξη θερμοκρασίας. Στάθμιση $\mathbf{v} : \mathbf{v}\left(t + \frac{1}{2}\delta t\right) \leftarrow \lambda \mathbf{v}\left(t + \frac{1}{2}\delta t\right)$.

Βήμα 6. Υπολογισμός των νέων, μη περιορισμένων θέσεων από την εξ. (2.3.7):

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) - \delta t \mathbf{v}\left(t + \frac{1}{2}\delta t\right)$$

Βήμα 7. Εφαρμογή περιορισμών στις συντεταγμένες: $[\mathbf{r}(t + \delta t) \rightarrow \mathbf{r}(t + \delta t); \mathbf{r}(t)]$

Βήμα 8. Υπολογισμός περιορισμένων ταχυτήτων:

$$\mathbf{v}\left(t + \frac{1}{2}\delta t\right) = [\mathbf{r}(t + \delta t) - \mathbf{r}(t)] / \delta t$$

Βήμα 9. Σύζευξη Πίεσης. Στάθμιση πίεσης στις συντεταγμένες και το μέγεθος του κουτιού:

$$\mathbf{r}(t + \delta t) \leftarrow \mu \mathbf{r}(t + \delta t)$$

$$l(t + \delta t) = \mu l(t)$$

$$V(t + \delta t) = \mu^3 V(t)$$

2.3.5. Ο ΑΛΓΟΡΙΘΜΟΣ STEEPEST DESCENT

Στις μοριακές προσομοιώσεις είναι πολλές φορές απαραίτητη η ενεργειακή ελαχιστοποίηση του συστήματος προτού υποβληθεί σε αλγόριθμο μοριακής δυναμικής. Αυτή έχει ως σκοπό την ελαχιστοποίηση της δυναμικής ενέργειας του συστήματος, έτσι ώστε να μπορεί στην συνέχεια να εφαρμοστεί αποτελεσματικά ο αλγόριθμος της μοριακής δυναμικής. Επιχειρείται, συνεπώς, η τοποθέτηση των μορίων του συστήματος σε ευνοϊκές θέσεις, οι οποίες αντιστοιχούν σε δυνάμεις μικρότερες από μια συγκεκριμένη τιμή, και οι οποίες στη συνέχεια θα αποτελέσουν τις αρχικές θέσεις του αλγορίθμου μοριακής δυναμικής. Στην παρούσα εργασία για το σκοπό αυτό χρησιμοποιείται ο αλγόριθμος Steepest Descent.²⁸

Στον Steepest Descent ορίζεται το διάνυσμα \mathbf{r} ως το διάνυσμα των $3N$ συντεταγμένων και δίνεται μια αρχική μέγιστη μετατόπιση h_0 . Υπολογίζονται οι δυνάμεις, \mathbf{F} και η δυναμική ενέργεια. Οι θέσεις κάθε νέο βήμα, $n+1$, υπολογίζονται από τη σχέση:

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \frac{\mathbf{F}_n}{\max(|\mathbf{F}_n|)} h_n \quad (2.3.36)$$

όπου h_n η μέγιστη μετατόπιση, \mathbf{F}_n η δύναμη ή η αρνητική κλίση της δυναμικής ενέργειας V και $\max(|\mathbf{F}_n|)$ αναφέρεται στη μέγιστη από τις απόλυτες τιμές των συστατικών της δύναμης.

Οι δυνάμεις και η ενέργεια υπολογίζονται ξανά για τις νέες θέσεις. Αν $(V_{n+1} < V_n)$, οι νέες θέσεις είναι αποδεκτές και $h_{n+1} = 1.2h_n$, ενώ αν $(V_{n+1} \geq V_n)$, οι νέες θέσεις απορρίπτονται και $h_n = 0.2h_n$. Ο αλγόριθμος σταματάει είτε όταν έχει συμπληρωθεί ένας μεγάλος αριθμός βημάτων ο οποίος ορίζεται από το χρήστη, είτε όταν οι μέγιστη των απολύτων τιμών των συστατικών της δύναμης είναι μικρότερη από την τιμή που ορίστηκε.

II. ΥΠΟΛΟΓΙΣΤΙΚΟ ΜΕΡΟΣ

3. ΔΗΜΙΟΥΡΓΙΑ ΣΥΣΤΗΜΑΤΩΝ ΜΟΡΙΑΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ

3.1. ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΜΑΚΡΟΜΟΡΙΩΝ

3.2. ΑΡΧΙΚΗ ΑΠΕΙΚΟΝΙΣΗ ΣΥΜΠΛΕΓΜΑΤΟΣ SRY-DNA

Για τη μελέτη, μέσω μοριακών προσομοιώσεων, του μηχανισμού πρόσδεσης των μεταγραφικών παραγόντων στο DNA, χρησιμοποιείται το σύστημα του μεταγραφικού παράγοντα SRY και του αντίστοιχου τμήματος του γονιδίου *SOX9* στο οποίο αυτός προσδένεται για να ενεργοποιήσει τη μεταγραφή του εν λόγω γονιδίου. Για την αρχική δομή του συστήματος χρησιμοποιήθηκε το αρχείο 1J46 της Protein Data Bank³², το οποίο περιέχει το σύστημα SRY-DNA όπως αυτό απομονώθηκε σε πειράματα Φασματοσκοπίας Πυρηνικού Μαγνητικού Συντονισμού (NMR).³⁰

ΔΟΜΗ SRY

Το τμήμα του SRY που προσομοιώνεται αποτελείται από τα 84 αμινοξέα που βρίσκονται στις θέσεις 57 – 140 της αλληλουχίας των 204 αμινοξέων και περιέχουν το HMG-κυτίο της SRY. Επιπλέον, μία μεθειονίνη βρίσκεται στην αρχή του τμήματος του μακρομορίου, η οποία αποτελεί τεχνητή προσθήκη και είναι αποτέλεσμα της πειραματικής διαδικασίας απομόνωσης της δομής. Η δευτεροταγής δομή του εν λόγω τμήματος του SRY περιλαμβάνει τρεις α-έλικες που σχηματίζουν μια δομή σχήματος «L», δύο στροφές και τρεις βρόγχους. Η πρωτοταγής και δευτεροταγής δομή του τμήματος του μεταγραφικού παράγοντα SRY που χρησιμοποιείται στην προσομοίωση παρουσιάζεται στην Εικόνα 3.1. Η επεξήγηση των συντομογραφιών για τα πρωτεϊνικά αμινοξέα παρουσιάζεται στον Πίνακα 3.1.

Πίνακας 3.1. Επεξήγηση συντομογραφιών πρωτεϊνικών αμινοξέων.

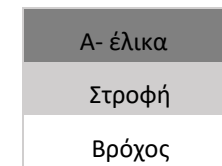
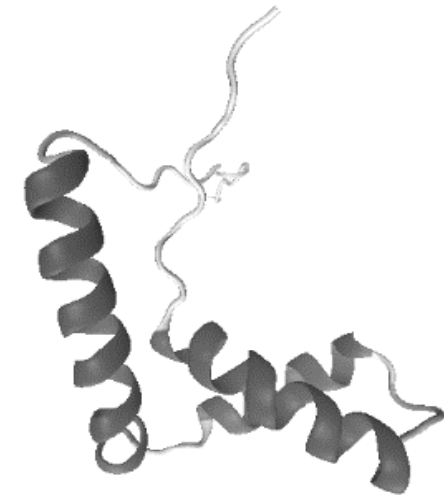
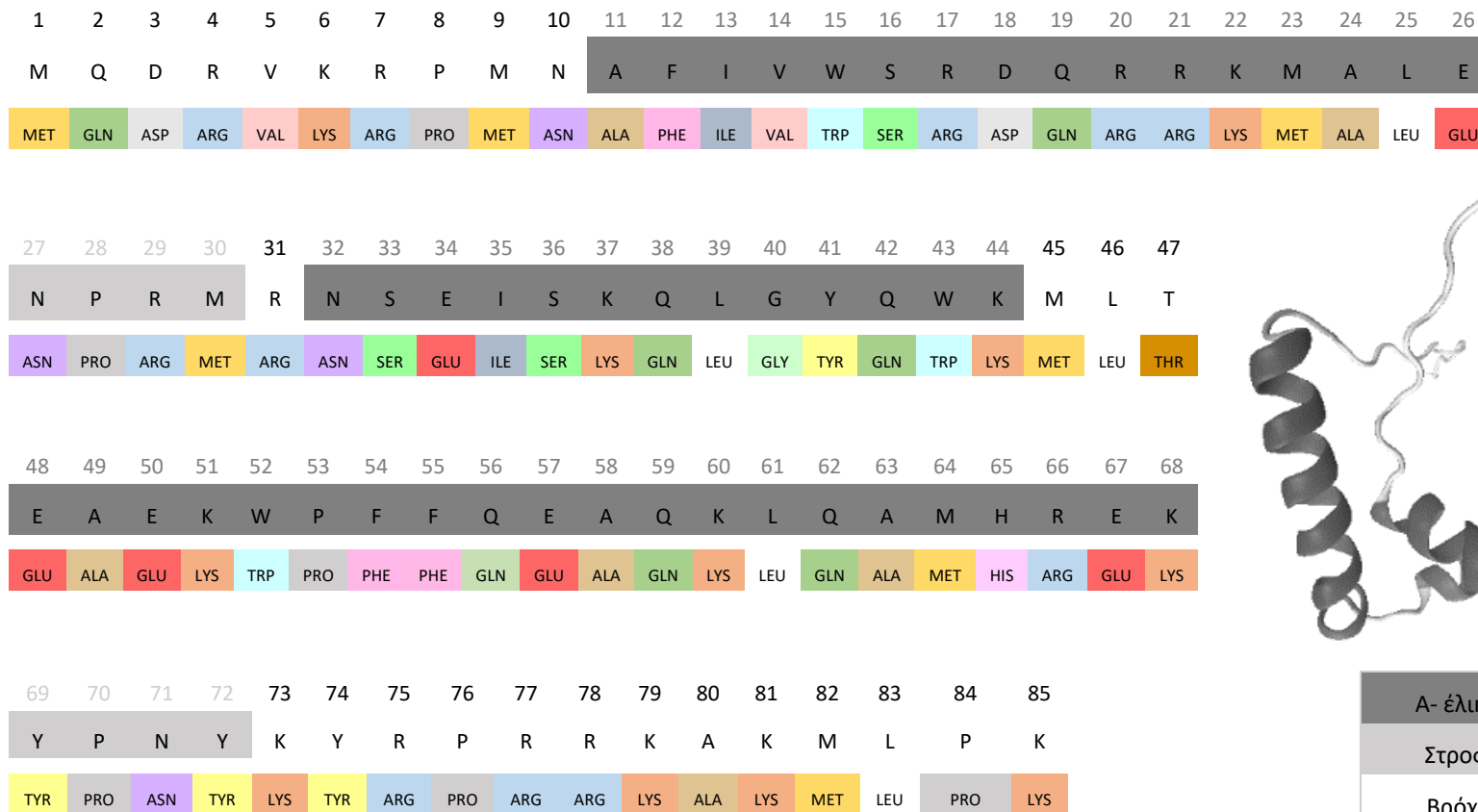
A	Ala	Αλανίνη	M	Met	Μεθειονίνη
C	Cys	Κυστεΐνη	N	Asn	Ασπαραγίνη
D	Asp	Ασπαρτικό οξύ	P	Pro	Προλίνη
E	Glu	Γλουταμικό οξύ	Q	Gln	Γλουταμίνη
F	Phe	Φαινυλαλανίνη	R	Arg	Αργινίνη
G	Gly	Γλυκίνη	S	Ser	Σερίνη
H	His	Ιστιδίνη	T	Thr	Θρεονίνη
I	Ile	Ισολευκίνη	V	Val	Βαλίνη
K	Lys	Λυσίνη	W	Trp	Θρυπτοφάνη
L	Leu	Λευκίνη	Y	Tyr	Τυροσίνη

ΔΟΜΗ ΓΟΝΙΔΙΟΥ *SOX9*

Το γονίδιο *SOX9* εδράζεται στο αυτοσωμικό χρωμόσωμα 17.³¹ Η SRY προσδένεται κατά προτίμηση την αλληλουχία (A/T)ACAA(T/A) από την πλευρά της μικρής αύλακας του γονιδίου *SOX9*.³² Στην παρούσα προσομοίωση χρησιμοποιείται αλληλουχία 14 ζευγών νουκλεοτιδίων, που περιλαμβάνουν την αλληλουχία αυτή. Η δομή του εν λόγω τμήματος του γονιδίου παρουσιάζεται στην Εικόνα 3.2.

3.2.1. ΨΗΦΙΑΚΗ ΑΝΑΚΑΤΑΣΚΕΥΗ ΕΛΕΥΘΕΡΟΥ DNA

Ανακατασκευάστηκε μια αρχική απεικόνιση για το τμήμα του γονιδίου *SOX9* μήκους 14 νουκλεοτιδίων, όταν αυτό είναι ελεύθερο, δηλαδή χωρίς τον προσδεδεμένο μεταγραφικό παράγοντα SRY με σκοπό την προσομοίωση του ελεύθερου DNA για λόγους σύγκρισης με το προσδεδεμένο DNA. Χρησιμοποιήθηκε η αλληλουχία των βάσεων που εμφανίζεται στην Εικόνα 3.2. Η ανακατασκευή σε ατομιστική λεπτομέρεια πραγματοποιήθηκε με χρήση του Web 3DNA 2.0 server.³³

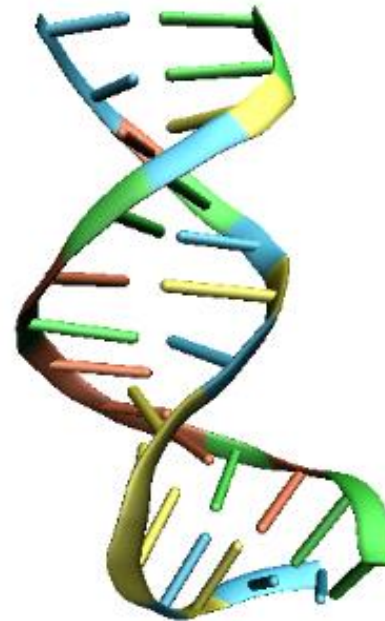


α)

β)

Εικόνα 3.1. Πρωτοταγής (α) και δευτεροταγής (β) δομή του τμήματος του μεταγραφικού παράγοντα SRV. α) Αλληλουχία των 85 αμινοξέων του τμήματος του SRV: στην πρώτη σειρά παρουσιάζεται ο αύξων αριθμός των αμινοξέων, στη δεύτερη η συντομογραφία των αμινοξέων, με χρωματική διαβάθμιση του γκρι σε αντιστοιχία με την απεικόνιση της δευτεροταγούς δομής και στην τρίτη η συντομογραφία με χρωματισμό ανάλογο του είδους του αμινοξέος. β) Απεικόνιση δευτεροταγούς δομής του τμήματος της SRV: αποτελείται από τρεις α-έλικες, δύο στροφές και τρεις βρόχους.

	Chain B	Chain C	
	5'-	3'-	
1	C	G	28
2	C	G	27
3	T	A	26
4	G	C	25
5	C	G	24
6	A	T	23
7	C	G	22
8	A	T	21
9	A	T	20
10	A	T	19
11	C	G	18
12	A	T	17
13	C	G	16
14	C	G	15
	-3'	-5'	



	Chain B	Chain C	
	5'-	3'-	
1	C	G	28
2	C	G	27
3	T	A	26
4	G	C	25
5	C	G	24
6	A	T	23
7	C	G	22
8	A	T	21
9	A	T	20
10	A	T	19
11	C	G	18
12	A	T	17
13	C	G	16
14	C	G	15
	-3'	-5'	

Εικόνα 3.2. Αλληλουχία βάσεων του DNA που αποτελεί τμήμα του γονιδίου *SOX9* στο οποίο συνδέεται ο μεταγραφικός παράγοντας *SRY*. Με A συμβολίζεται η αδενίνη, C η κυτοσίνη, G η γουανίνη και T η θυμίνη. Με Chain B και Chain C συμβολίζονται οι δύο έλικες του DNA.

3.3. ΠΕΔΙΟ ΔΥΝΑΜΕΩΝ ΚΑΙ ΠΑΡΑΜΕΤΡΟΙ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ

Πραγματοποιήθηκαν προσομοιώσεις μοριακής δυναμικής για την καταγραφή της χρονικής εξέλιξης της διαμόρφωσης του DNA και του μεταγραφικού παράγοντα. Οι προσομοιώσεις πραγματοποιήθηκαν στο ισόθερμο-ισοβαρές στατιστικό σύνολο (N, P, T) όπου N ο συνολικός αριθμός των ατόμων του συστήματος, P η πίεση και T η θερμοκρασία. Προσομοιώθηκαν τα εξής συστήματα:

- I. Σύστημα DNA – SRY, σε νερό, παρουσία ιόντων Na^+
- II. Σύστημα DNA σε νερό, παρουσία ιόντων Na^+

Η θερμοκρασία και η πίεση των συστημάτων ρυθμίστηκαν ώστε να πλησιάζουν τις συνθήκες στον ανθρώπινο οργανισμό, $T = 310 \text{ K}$ και $P = 1 \text{ bar}$. Οι τροχιές της μοριακής δυναμικής είχαν χρονική διάρκεια 100 ns.

ΠΕΔΙΟ ΔΥΝΑΜΕΩΝ

Η αριθμητική ολοκλήρωση των εξισώσεων κίνησης³⁴ για όλους τους βαθμούς ελευθερίας των συστημάτων έγινε μέσω του λογισμικού πακέτου GROMACS 2016.1. με τη χρήση των κατάλληλων πεδίων δυνάμεων για το DNA, την πρωτεΐνη και το νερό. Για το DNA χρησιμοποιήθηκε το πεδίο δυνάμεων AMBER ff94³⁵ και για την πρωτεΐνη το πεδίο δυνάμεων AMBER99sb-ildn³⁶, το οποίο έχει χρησιμοποιηθεί με επιτυχία σε προσομοιώσεις πρωτεϊνών με νουκλεϊκά οξέα.^{37, 38} Για το νερό χρησιμοποιήθηκε το πεδίο δυνάμεων TIP3P³⁹, το οποίο έχει χρησιμοποιηθεί επιτυχώς σε προσομοιώσεις ενυδάτωσης νουκλεϊκών οξέων μαζί με το πεδίο δυνάμεων AMBER⁴⁰.

Στα πεδία δυνάμεων AMBER, οι δεσμοί και οι γωνίες περιγράφονται ως αρμονικές ταλαντώσεις, οι δίεδρες προσδιορίζονται βάσει των κεντρικών ατόμων, μέσω ενός συνόλου παραμέτρων, οι αλληλεπιδράσεις van der Waals περιγράφονται από το δυναμικό Lennard-Jones 12-6, και οι ηλεκτροστατικές αλληλεπιδράσεις περιγράφονται ως αλληλεπιδράσεις Coulomb μεταξύ σημειακών φορτίων που βρίσκονται στα κέντρα των ατόμων. Το πεδίο δυνάμεων παρουσιάζεται στην εξίσωση (3.2.1).

$$V_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\varphi - \gamma)] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \quad (3.2.1)$$

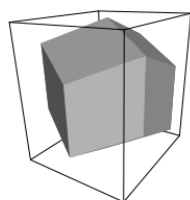
Στην παραπάνω εξίσωση οι πρώτοι τρεις όροι αναφέρονται σε δεσμικές αλληλεπιδράσεις. Ο πρώτος όρος αποτελεί εξίσωση αρμονικού δυναμικού και περιγράφει τη δόνηση των δεσμών γύρω από τη θέση ισορροπίας. Ο δεύτερος όρος περιγράφει την κάμψη των γωνιών και ο τρίτος περιγράφει την ενέργεια από τις αλληλεπιδράσεις που προκύπτουν από τις μεταβολές των διεδρων γωνιών.

Το τελευταίο άθροισμα περιγράφει τις μη δεσμικές αλληλεπιδράσεις ατόμων. Οι πρώτοι δύο όροι εκφράζουν το δυναμικό Lennard-Jones 12-6 για την περιγραφή των δυνάμεων διασποράς. Ο τελευταίος όρος αποτελεί έκφραση του νόμου Coulomb για τις ηλεκτροστατικές δυνάμεις που αναπτύσσονται λόγω των μερικών φορτίων των ιόντων.

Συγκεκριμένα στην εξίσωση συμβολίζονται με r το μήκος δεσμού, θ η γωνία του δεσμού και ο δείκτης eq αναφέρεται στις τιμές ισορροπίας. Στο δυναμικό διεδρων, φ η διεδρη γωνία, η παράμετρος n αναφέρεται στην πολλαπλότητα της διεδρης γωνίας και η γ στην αλλαγή φάσης κατά τη διάρκεια μιας πλήρους περιστροφής. Στο δυναμικό μη δεσμικών αλληλεπιδράσεων οι παράμετροι A_{ij} και B_{ij} αντιστοιχούν στους παράγοντες $4\epsilon\sigma^{12}$ και $4\epsilon\sigma^6$ του δυναμικού Lennard-Jones 12-6 που περιγράφεται στην εξ. (2.1.15) και οι όροι q_i και q_j τα μερικά φορτία των σωματιδίων i και j τα οποία έχουν απόσταση R_{ij} . Το ϵ αποτελεί έκφραση της διηλεκτρικής σταθεράς και οι όροι K_r, K_θ, V_n αποτελούν σταθερές που εξαρτώνται από τα άτομα που συμμετέχουν στο δεσμό.

ΠΡΟΣΟΜΟΙΟΥΜΕΝΟ ΧΩΡΙΟ

Για την προσομοίωση επιλέχτηκε χωρίο σχήματος ρομβικού δωδεκάεδρου με χαρακτηριστικά που παρουσιάζονται στην Εικόνα 3.3. Το ρομβικό δωδεκάεδρο κελί αποτελεί το μικρότερο χωροπληρωτικό κελί, με όγκο 71% του κυβικού κελιού ιδίων διαστάσεων, με αποτέλεσμα να εξοικονομεί υπολογιστικό χρόνο σε σχέση με το κυβικό κελί. Ενδείκνυται για διατάξεις μακρομορίων που προσεγγίζουν σφαιρικό σχήμα, διαλυμένα σε διαλύτη, καθώς απαιτούνται λιγότερα μόρια διαλύτη για να πληρούται το χωρίο της προσομοίωσης,²⁸ και εξοικονομείται υπολογιστικός χρόνος.



box type	image distance	box volume	box vectors			box vector angles		
			a	b	c	∠bc	∠ac	∠ab
cubic	d	d^3	d 0 0	0 d 0	0 0 d	90°	90°	90°
rhombic dodecahedron (xy-square)	d	$\frac{1}{2}\sqrt{2}d^3$ $0.707d^3$	d 0 0	0 d 0	$\frac{1}{2}d$ $\frac{1}{2}d$ $\frac{1}{2}\sqrt{2}d$	60°	60°	90°

Εικόνα 3.3. Χαρακτηριστικά και απεικόνιση ρομβικού δωδεκάεδρου κελιού.²⁸

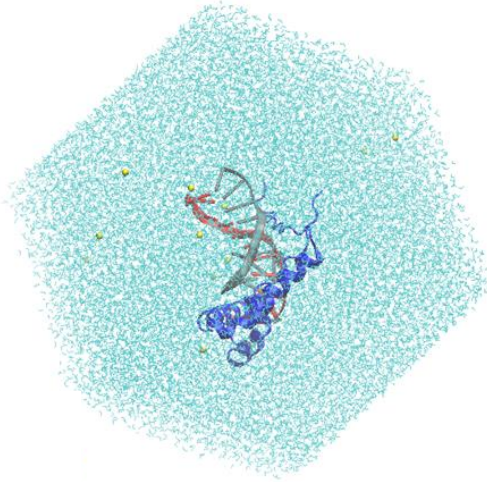
Οι διαστάσεις του χωρίου καθορίστηκαν σε κάθε προσομοίωση έτσι ώστε τα μακρομόρια να καλύπτονται από στρώμα νερού τουλάχιστον 2.0 nm προς κάθε κατεύθυνση. Η τιμή αυτή είναι αρκετά μεγαλύτερη από την προτεινόμενη, 0.8 – 1.2 nm στα πρωτόκολλα για τις προσομοιώσεις νουκλεϊκών οξέων⁴¹ και εξασφαλίζει την παρουσία του κελύφους ενυδάτωσης του δίκλωνου μορίου νουκλεϊκού οξέος και την απαραίτητη ποσότητα περιβάλλοντος νερού.

Στο χωρίο της προσομοίωσης προστέθηκε νερό ως διαλύτης. Στο σύστημα η πρωτεΐνη είναι θετικά φορτισμένη, το DNA αρνητικά φορτισμένο, ενώ το συνολικό φορτίο του συστήματος δεν είναι ουδέτερο. Για την εξουδετέρωση του φορτίου ήταν απαραίτητη η προσθήκη κατιόντων Na. Επιλέχθηκαν με τυχαίο τρόπο, μέσω γεννήτριας τυχαίων αριθμών, μόρια νερού από το σύστημα, τα οποία αντικαταστάθηκαν από κατάλληλο αριθμό ατόμων Na⁺ ώστε το φορτίο να είναι ουδέτερο.

Στον Πίνακα 3.2 παρουσιάζονται τα δεδομένα για το μέγεθος των δύο συστημάτων και στην Εικόνα 3.4 η αρχική απεικόνιση της διάταξης για το σύστημα DNA – SRY σε νερό παρουσία ιόντων.

Πίνακας 3.2. Αριθμός ατόμων και μέγεθος χωρίου προσομοίωσης των συστημάτων

	Σύστημα 1: DNA - SRY	Σύστημα 2: Ελεύθερο DNA
Άτομα DNA	886	886
Άτομα Πρωτεΐνης	1513	-
Άτομα νερού	65109 (21703 μόρια H ₂ O)	51678 (17226 μόρια H ₂ O)
Άτομα Na ⁺	12	26
Συνολικός αριθμός ατόμων:	67520	52590
Όγκος κουτιού (nm³)	684.35	542.36



Εικόνα 3.4. Αρχική απεικόνιση του χωρίου προσομοίωσης DNA – SRY σε νερό παρουσία ιόντων. Με γαλάζιες γραμμές απεικονίζονται τα μόρια νερού. Με μπλε χρώμα απεικονίζεται η πρωτεΐνη, και με κόκκινο και γκρι οι δύο κλώνοι του DNA. Τα μόρια Na^+ απεικονίζονται ως κίτρινες σφαίρες.

ΕΞΙΣΟΡΡΟΠΗΣΗ ΣΥΣΤΗΜΑΤΩΝ

Πριν την εκτέλεση των μοριακών προσομοιώσεων για την ανάλυση των παραγόμενων τροχιών, προηγήθηκε εξισορρόπηση των συστημάτων, η οποία ακολούθησε συγκεκριμένο πρωτόκολλο⁴², το οποίο περιλαμβάνει τρία διακριτά βήματα. Καταρχάς, πραγματοποιήθηκε ενεργειακή ελαχιστοποίηση των συστημάτων με χρήση του αλγορίθμου steepest-descent προκειμένου να εξαλειφθούν τυχόν επικαλύψεις ατόμων στην αρχική απεικόνιση. Η μέγιστη δύναμη ορίστηκε στην τιμή 10^3 J/mol/nm. Κατά τη διάρκεια της ελαχιστοποίησης τα άτομα των μακρομορίων κρατήθηκαν ακίνητα στις αρχικές τους διαμορφώσεις με επιβολή αρμονικού δυναμικού σε όλα τα άτομα της δομής τους εκτός των υδρογόνων, με σταθερά δύναμης 10^6 Jmol⁻¹nm⁻².

Στο δεύτερο και τρίτο βήμα, διεξήχθησαν τρεξίματα μοριακής δυναμικής (i) στο κανονικό στατιστικό σύνολο, (N, V, T) , χρονικής διάρκειας 1 ns, για την εξισορρόπηση της θερμοκρασίας του συστήματος, και (ii) στο ισόθερμο – ισοβαρές στατιστικό σύνολο (N, P, T) χρονικής διάρκειας έως 6 ns για την εξισορρόπηση της πίεσης του συστήματος και της πυκνότητας του διαλύτη. Στα ανωτέρω τρεξίματα τα άτομα των μακρομορίων κρατήθηκαν επίσης ακίνητα στις αρχικές τους διαμορφώσεις όπως περιγράφηκε παραπάνω.

ΕΚΤΕΛΕΣΗ ΠΡΟΣΟΜΟΙΩΣΕΩΝ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ

Έπειτα από την εξισορρόπηση των συστημάτων, αφαιρέθηκαν οι περιορισμοί στις κινήσεις των ατόμων των μακρομορίων και ακολούθησαν προσομοιώσεις μοριακής δυναμικής χρονικής διάρκειας 100 ns. Χρησιμοποιήθηκε ο τροποποιημένος Berendsen θερμοστάτης V-rescale⁴³ για την σύζευξη της θερμοκρασίας και ο βαροστάτης Berendsen²⁷ για τη σύζευξη της πίεσης. Εφαρμόστηκαν περιορισμοί στα μήκη των δεσμών στους οποίους εμπλέκονται άτομα υδρογόνου⁴⁴, όπως συνηθίζεται στα βιομοριακά συστήματα, για την εξοικονόμηση υπολογιστικού χρόνου χωρίς να έχει μεγάλη επίπτωση στην ακρίβεια των υπολογισμών. Στον αλγόριθμο Berendsen χρησιμοποιήθηκε χρονική σταθερά 2.0 ps και η συμπίεστικότητα του νερού ορίστηκε στην τιμή $4.5 \cdot 10^{-5} \text{ bar}^{-1}$. Η σύζευξη της θερμοκρασίας εφαρμόστηκε σε δύο ξεχωριστές ομάδες ατόμων: DNA – SRY (και αντίστοιχα μόνο το DNA στο σύστημα του ελεύθερου DNA) και νερό – ιόντα, με χρονική σταθερά 0.1 ps. Για τις αλληλεπιδράσεις διασποράς εφαρμόστηκε κατώφλι δυναμικού 1.0 nm. Για τον υπολογισμό των ηλεκτροστατικών αλληλεπιδράσεων χρησιμοποιήθηκε η μέθοδος particle-mesh Ewald⁴⁵. Η ολοκλήρωση των εξισώσεων κίνησης έγινε με τον αλγόριθμο leap-frog και οι παραγόμενες τροχιές καταγράφονταν ανά 10 ps με βήμα 2 fs.

Η οπτικοποίηση όλων των τροχιών και η δημιουργία εικόνων έγινε με χρήση του προγράμματος VMD⁴⁶. Τα γεωμετρικά χαρακτηριστικά του DNA εξήχθησαν από το πρόγραμμα CURVES⁴⁷.

4. ΥΠΟΛΟΓΙΣΤΙΚΗ ΕΠΙΔΟΣΗ ΜΟΡΙΑΚΩΝ ΠΡΟΣΟΜΟΙΩΣΕΩΝ (PERFORMANCE)

Τα τρεξίματα των μοριακών προσομοιώσεων διεξήχθησαν στη συστοιχία υπολογιστών (Beowulf Linux cluster) "Pericles" της ερευνητικής ομάδας ENGIMATER του Τομέα Επιστήμης και Τεχνικής των Υλικών, με επικεφαλής τον Καθηγητή Γ. Κ. Παπαδόπουλο. Χρησιμοποιήθηκαν εξυπηρετητές 1950 Servers QUAD Core - Intel XEON E5430 @ 2.66 GHz, 64 Gb Memory. Στους πίνακες που ακολουθούν παρουσιάζονται τα δεδομένα για την υπολογιστική επίδοση των προσομοιώσεων: ο χρόνος προσομοίωσης, το μέγεθος του κάθε συστήματος (αριθμός ατόμων), το χρονικό διάστημα της ολοκλήρωσης (dt), ο αριθμός επαναληπτικών βημάτων της ολοκλήρωσης (nsteps), οι κόμβοι, οι πυρήνες CPU ανά κόμβο και ο χρόνος εκτέλεσης (wall time). Σημειώνεται ότι στα τρεξίματα τα αποτελέσματα καταγράφονται στα αρχεία εξόδου ανά 5000 βήματα.

Πίνακας 4.1. Δεδομένα υπολογιστικής επίδοσης για την εκτέλεση της προσομοίωσης του συστήματος DNA - SRY

ΣΥΣΤΗΜΑ DNA - SRY			
	NVT εξισορρόπηση	NPT εξισορρόπηση	Προσομοίωση ΜΔ
Χρόνος προσομοίωσης	1 ns	2ns	100ns
Μέγεθος συστήματος	67520 άτομα	67520 άτομα	67520 άτομα
dt	0.002 ps	0.002 ps	0.002 ps
nsteps	500000	1000000	50000000
Κόμβοι (nodes)	2	2	4
Πυρήνες ανά κόμβο	8	8	8
Σύνολο πυρήνων	16	16	32
Wall time	2 h 59 min 06 s	5 h 59 min 59 s	7 d 03 h 03 min 19 s

Πίνακας 4.2. Δεδομένα υπολογιστικής επίδοσης για την εκτέλεση της προσομοίωσης του συστήματος DNA.

ΣΥΣΤΗΜΑ DNA			
	NVT εξισορρόπηση	NPT εξισορρόπηση	Προσομοίωση ΜΔ
Χρόνος προσομοίωσης	1 ns	6ns	100ns
Μέγεθος συστήματος	52590 άτομα	52590 άτομα	52590 άτομα
dt	0.002 ps	0.002 ps	0.002 ps
nsteps	500000	3000000	50000000
Κόμβοι (nodes)	7	7	7
Πυρήνες ανά κόμβο	8	8	8
Σύνολο πυρήνων	56	56	56
Wall time	1 h 30 min 34 s	9h 06 min 36 s	6 d 7 h 49 min 58 s

5. ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΞΙΣΟΡΡΟΠΗΣΗΣ ΣΥΣΤΗΜΑΤΩΝ ΠΡΟΣΟΜΟΙΩΣΗΣ

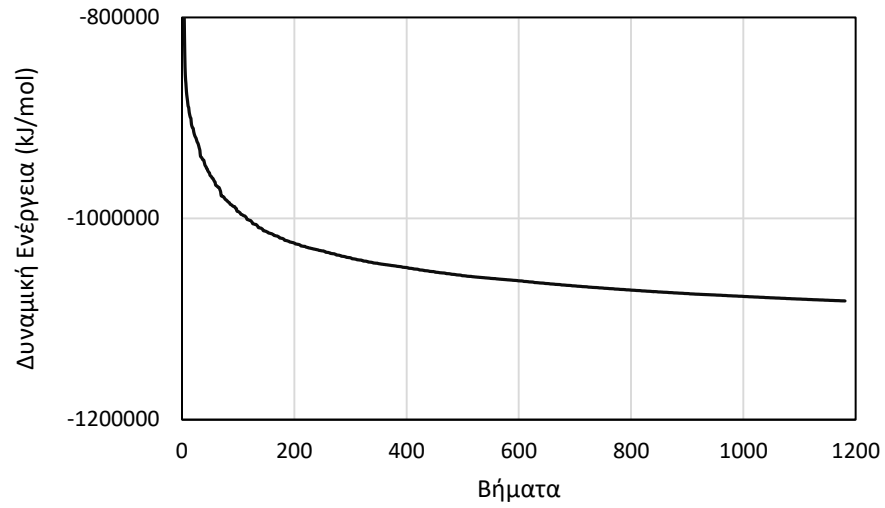
5.1. ΕΞΙΣΟΡΡΟΠΗΣΗ ΣΥΣΤΗΜΑΤΟΣ DNA – SRY ΣΕ ΥΔΑΤΙΚΟ ΔΙΑΛΥΜΑ

Η ενεργειακή ελαχιστοποίηση του συστήματος DNA – SRY στο νερό παρουσία ιόντων ολοκληρώθηκε σε 1181 βήματα εφόσον η μέγιστη δύναμη ήταν μικρότερη των 1000.0 kJ/mol/nm. Το Διάγραμμα 5.1, που ακολουθεί, παρουσιάζει τη μεταβολή της δυναμικής ενέργειας κατά την εξέλιξη της ενεργειακής ελαχιστοποίησης. Η ελάχιστη τιμή της δυναμικής ενέργειας είναι -1082013 kJ/mol.

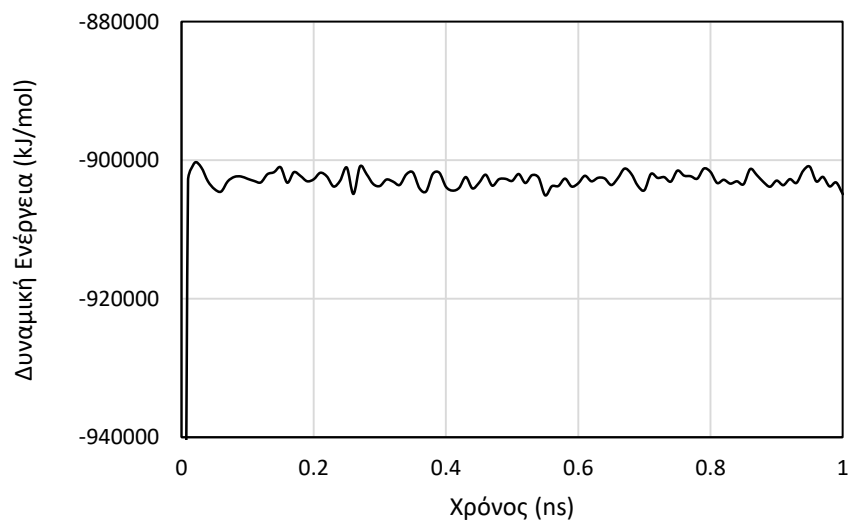
Η εξισορρόπηση έγινε σε NVT είχε διάρκεια 1ns. Η χρονική εξέλιξη της δυναμικής ενέργειας παρουσιάζεται στο Διάγραμμα 5.2 και η χρονική εξέλιξη της θερμοκρασίας κατά τη διάρκεια της προσομοίωσης παρουσιάζεται στο Διάγραμμα 5.3. Τα δεδομένα έχουν καταγραφεί ανά 10ps της τροχιάς και ο μέσος όρος της θερμοκρασίας είναι 309.7 K.

Η εξισορρόπηση NPT, για πίεση $P=1$ bar και θερμοκρασία $T= 310$ K είχε χρονική διάρκεια 2 ns. Η χρονική εξέλιξη της δυναμικής ενέργειας παρουσιάζεται στο Διάγραμμα 5.4 και ο μέσος όρος της είναι -902672 kJ/mol η χρονική εξέλιξη της θερμοκρασίας στο Διάγραμμα 5.5 με μέσο όρο 310.0 K της πίεσης στο Διάγραμμα 5.6 με μέσο όρο 1.19 bar και της πυκνότητας στο Διάγραμμα 5.7, με μέσο όρο 993.7 kg/m³. Τα δεδομένα έχουν καταγραφεί ανά 10ps και από αυτά έχουν εξαχθεί οι μέσοι όροι. Στα Διαγράμματα 5.5 – 5.7 παρουσιάζονται επίσης οι κυλιόμενοι μέσοι όροι. Η τιμή του κυλιόμενου μέσου όρου σε μία χρονική στιγμή προκύπτει από τον μέσο όρο των ιδιοτήτων των 10 τελευταίων χρονικών στιγμών.

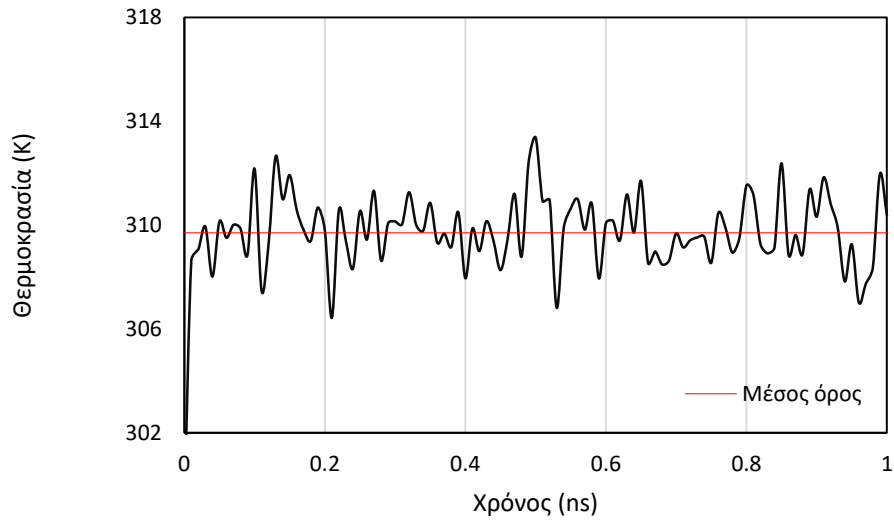
Σημειώνεται ότι η πυκνότητα του συστήματος, προσεγγίζει την πυκνότητα του νερού στους 310 K, η οποία είναι 993.4 kg/m³.⁴⁸



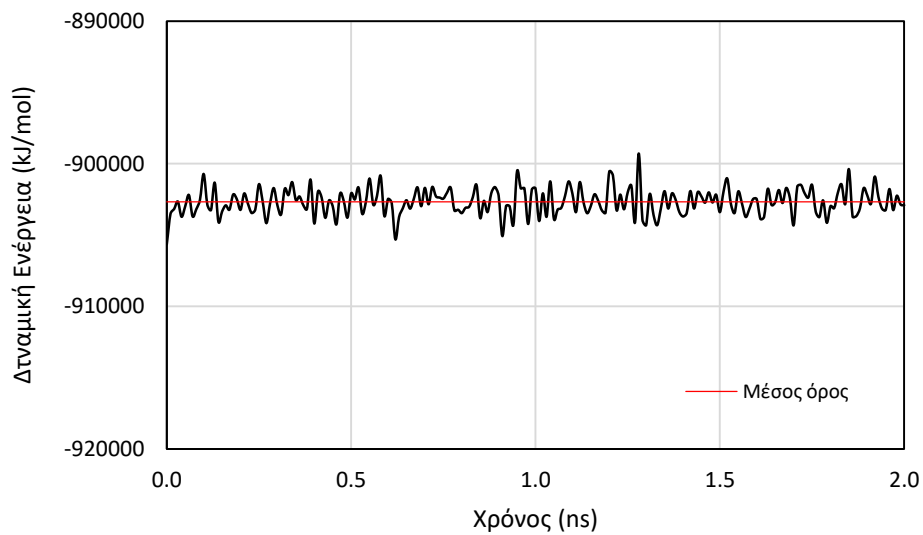
Διάγραμμα 5.1. Μεταβολή της δυναμικής ενέργειας κατά την ενεργειακή ελαχιστοποίηση συστήματος DNA – SRY.



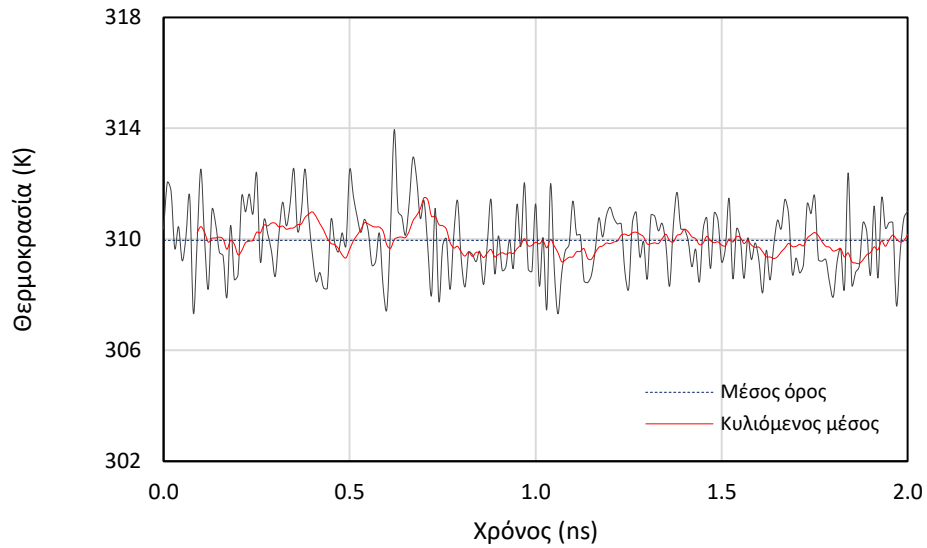
Διάγραμμα 5.2. Χρονική εξέλιξη της δυναμικής ενέργειας κατά την NVT εξισορρόπηση του συστήματος DNA – SRY.



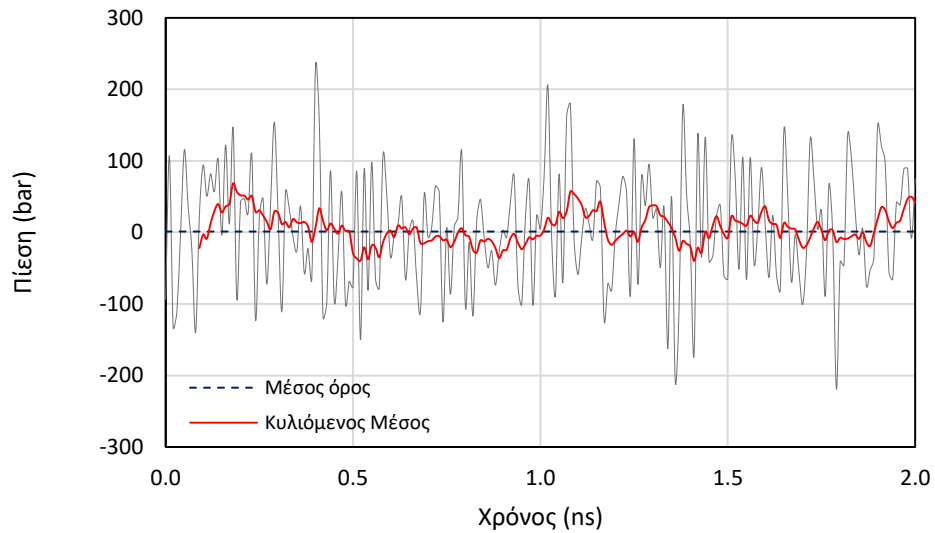
Διάγραμμα 5.3. Χρονική εξέλιξη της θερμοκρασίας κατά την εξισορρόπηση NVT του συστήματος DNA – SRY. Με κόκκινη γραμμή παρουσιάζεται ο μέσος όρος των τιμών της θερμοκρασίας.



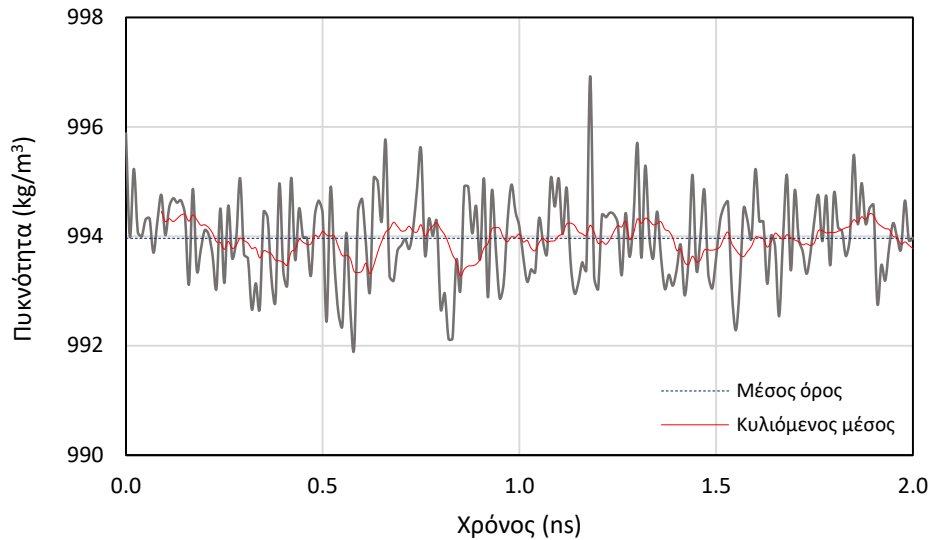
Διάγραμμα 5.4. Χρονική εξέλιξη της δυναμικής ενέργειας κατά την εξισορρόπηση NPT του συστήματος DNA – SRY . Με κόκκινη γραμμή παρουσιάζεται ο μέσος όρος των τιμών της δυναμικής ενέργειας.



Διάγραμμα 5.5. Χρονική εξέλιξη της θερμοκρασίας κατά την εξισορρόπηση NPT του συστήματος DNA – SRY. Με κόκκινη γραμμή παρουσιάζεται ο κυλιόμενος μέσος και με μπλε διακεκομμένη ο μέσος όρος.



Διάγραμμα 5.6. Χρονική εξέλιξη της πίεσης κατά την εξισορρόπηση NPT του συστήματος SRY – DNA. Με κόκκινη γραμμή παρουσιάζεται ο κυλιόμενος μέσος και με μπλε διακεκομμένη ο μέσος όρος.



Διάγραμμα 5.7. Χρονική εξέλιξη της πυκνότητας κατά την εξισορρόπηση NPT του συστήματος SRV – DNA. Με κόκκινη γραμμή παρουσιάζεται ο κυλιόμενος μέσος και με μπλε διακεκομμένη ο μέσος όρος.

5.2. ΕΞΙΣΟΡΡΟΠΗΣΗ ΣΥΣΤΗΜΑΤΟΣ DNA ΣΕ ΥΔΑΤΙΚΟ ΔΙΑΛΥΜΑ

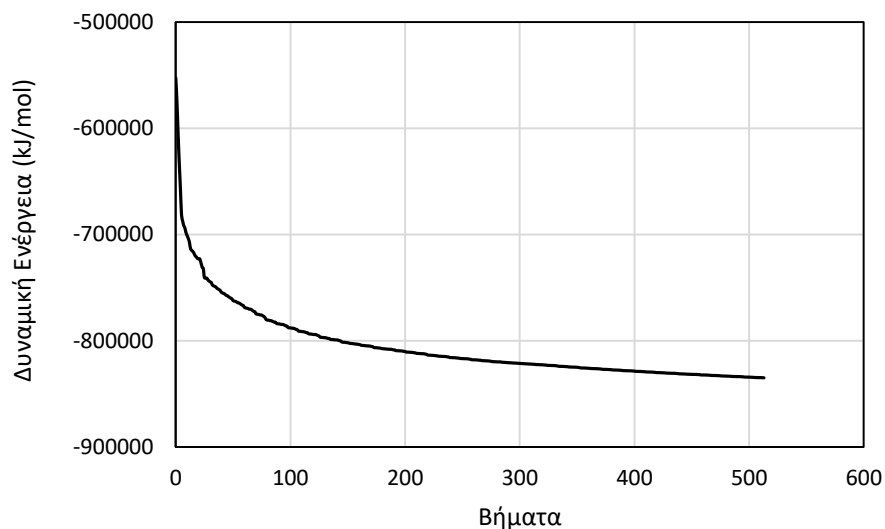
Η ενεργειακή ελαχιστοποίηση του συστήματος του ελεύθερου DNA στο νερό παρουσία ιόντων ολοκληρώθηκε σε 514 βήματα εφόσον η μέγιστη δύναμη ήταν μικρότερη των 1000.0 kJ/mol/nm. Το Διάγραμμα 5.8, που ακολουθεί, παρουσιάζει τη μεταβολή της δυναμικής ενέργειας κατά την εξέλιξη της ενεργειακής ελαχιστοποίησης. Η ελάχιστη τιμή της δυναμικής ενέργειας είναι -834799 kJ/mol.

Η εξισορρόπηση έγινε σε NVT είχε διάρκεια 1ns. Η χρονική εξέλιξη της δυναμικής ενέργειας παρουσιάζεται στο Διάγραμμα 5.9 και η χρονική εξέλιξη της θερμοκρασίας κατά τη διάρκεια της προσομοίωσης παρουσιάζεται στο Διάγραμμα 5.10. Τα δεδομένα έχουν καταγραφεί ανά 10ps της τροχιάς και η μέση τιμή των δεδομένων της θερμοκρασίας είναι 309.9 K.

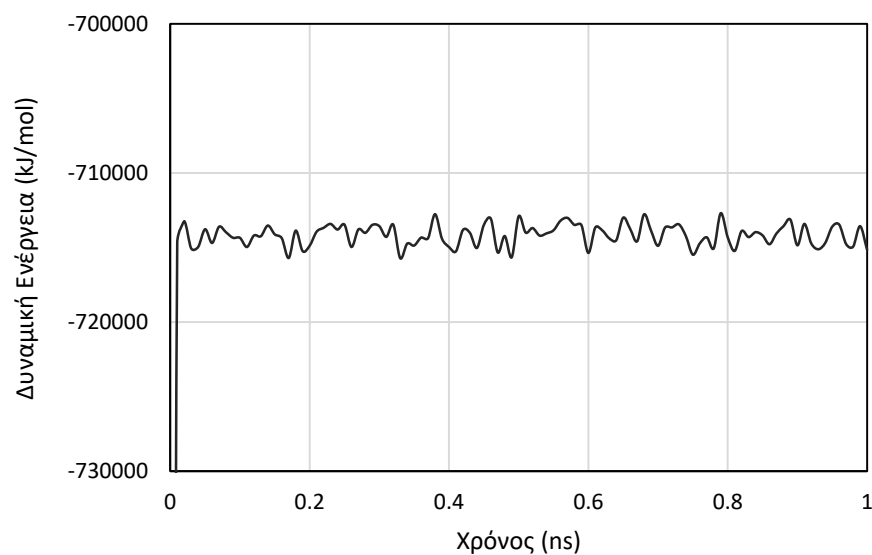
Η εξισορρόπηση NPT, για πίεση $P = 1$ bar και θερμοκρασία $T = 310$ K είχε χρονική διάρκεια 6 ns. Η χρονική εξέλιξη της δυναμικής ενέργειας παρουσιάζεται στο Διάγραμμα 5.11 και ο μέσος όρος της είναι -716698 kJ/mol, η χρονική εξέλιξη της θερμοκρασίας στο Διάγραμμα 5.12 με μέσο όρο 310.0 K, της πίεσης στο Διάγραμμα 5.13 με μέσο όρο 0.96 bar και της

πυκνότητας στο Διάγραμμα 5.14 , με μέσο όρο 991.4 kg/m^3 . Τα δεδομένα έχουν καταγραφεί ανά 10ps και από αυτά εξαχθεί οι μέσοι όροι και οι κυλιόμενοι μέσοι όροι.

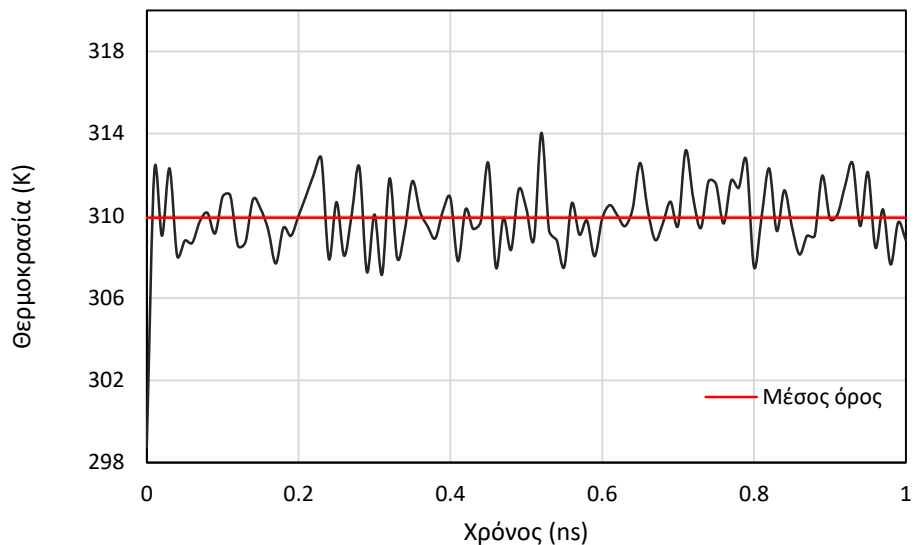
Σημειώνεται ότι η πυκνότητα του συστήματος, προσεγγίζει την πυκνότητα του νερού στους 310 K, η οποία είναι 993.4 kg/m^3 .⁵¹



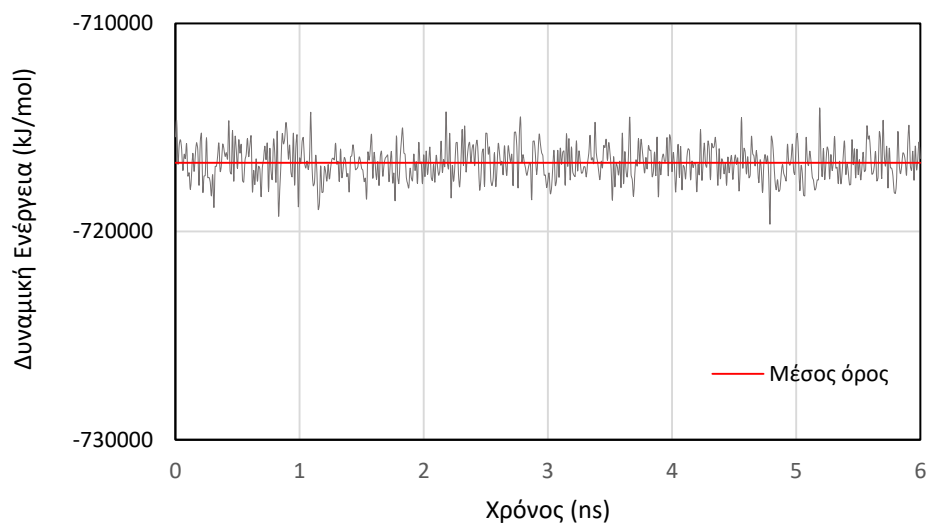
Διάγραμμα 5.8. Μεταβολή της δυναμικής ενέργειας κατά την ενεργειακή ελαχιστοποίηση συστήματος ελεύθερου DNA



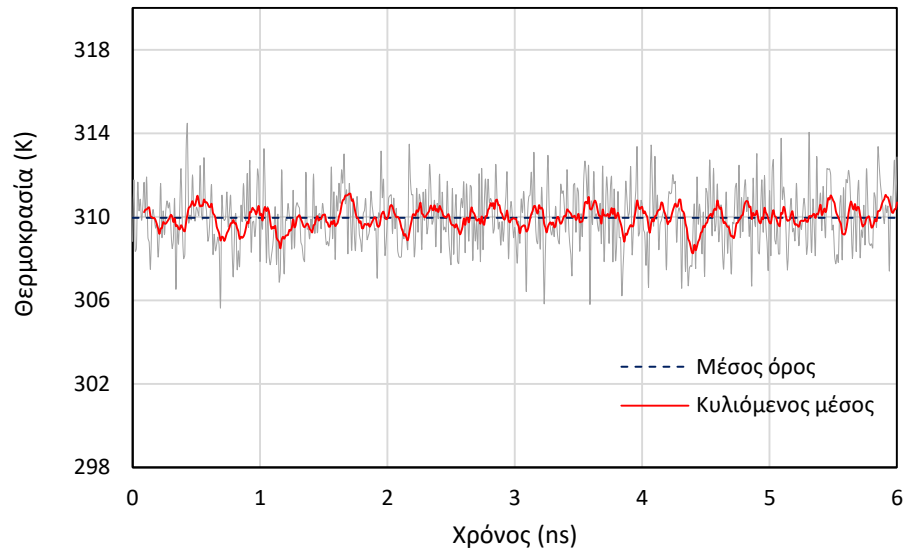
Διάγραμμα 5.9. Χρονική εξέλιξη της δυναμικής ενέργειας κατά την εξισορρόπηση NVT για το ελεύθερο DNA.



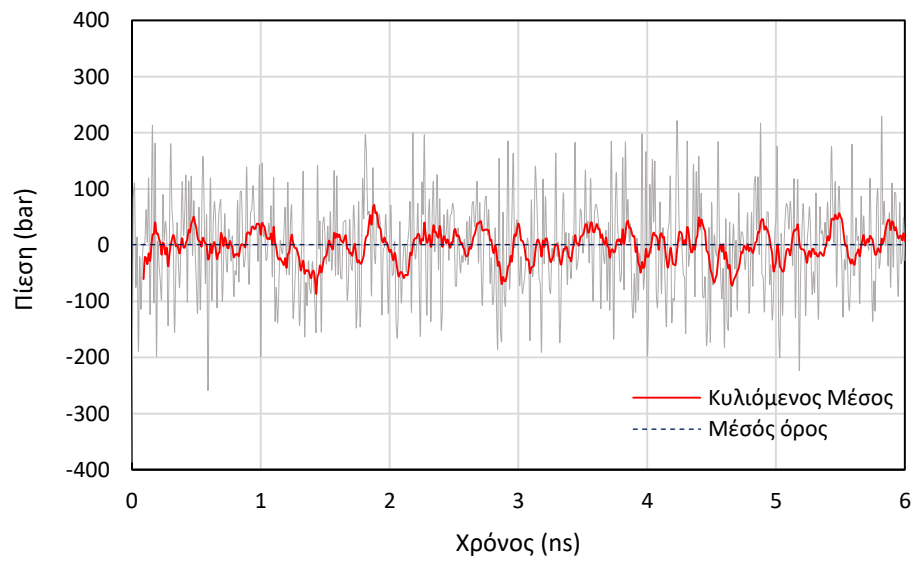
Διάγραμμα 5.10. Χρονική εξέλιξη της θερμοκρασίας κατά την εξισορρόπηση NVT για το ελεύθερο DNA. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών θερμοκρασίας.



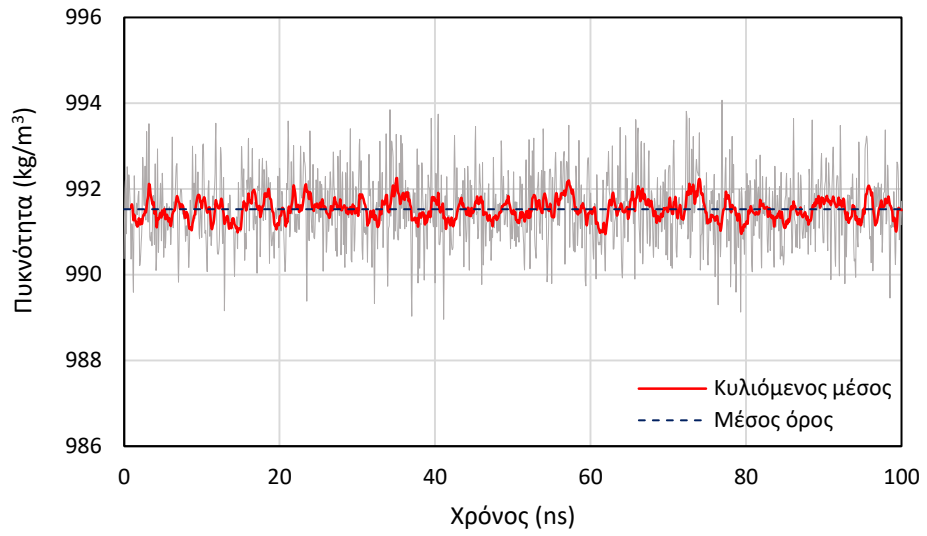
Διάγραμμα 5.11. Χρονική εξέλιξη της δυναμικής ενέργειας κατά την εξισορρόπηση NPT. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών της δυναμικής ενέργειας.



Διάγραμμα 5.12. Χρονική εξέλιξη της θερμοκρασίας κατά την NPT εξισορρόπηση. Με κόκκινη γραμμή απεικονίζεται ο κυλιόμενος μέσος όρος και με μπλε διακεκομμένη ο μέσος όρος των τιμών.



Διάγραμμα 5.13. Χρονική εξέλιξη της πίεσης κατά την NPT εξισορρόπηση. Με κόκκινη γραμμή απεικονίζεται ο κυλιόμενος μέσος όρος και με μπλε διακεκομμένη ο μέσος όρος των τιμών.



Διάγραμμα 5.14. Χρονική εξέλιξη της πυκνότητας κατά την NPT εξισορρόπηση. Με κόκκινη γραμμή απεικονίζεται ο κυλιόμενος μέσος όρος και με μπλε διακεκομμένη ο μέσος όρος των τιμών.

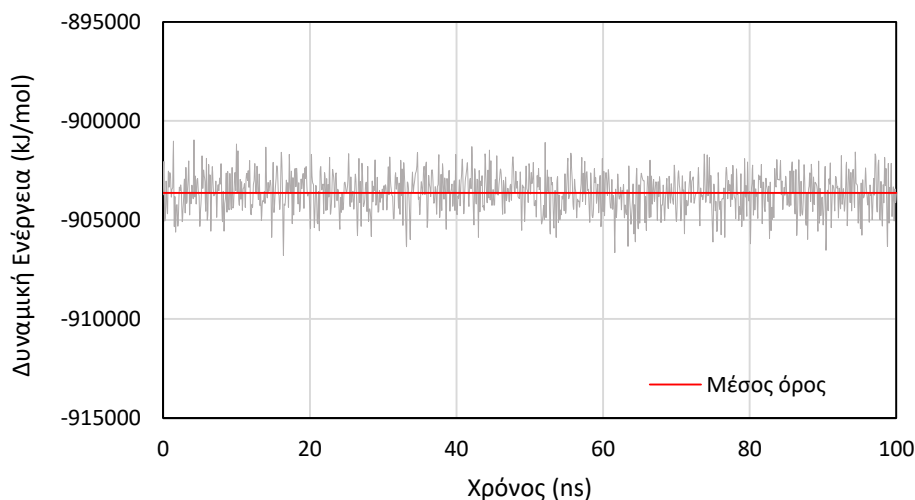
6. ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΣΟΜΟΙΩΣΕΩΝ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα της μοριακής δυναμικής, τα οποία προκύπτουν από τις τροχιές που παρήχθησαν από τις δύο προσομοιώσεις: του συστήματος DNA – SRY και του συστήματος ελεύθερου DNA. Οι τροχιές αυτές έχουν χρονική διάρκεια 100 ns και τα μεγέθη που προκύπτουν έχουν εξαχθεί από τιμές ανά 100 ps της τροχιάς, τα οποία αντιστοιχούν σε 1001 χρονικές στιγμές της τροχιάς.

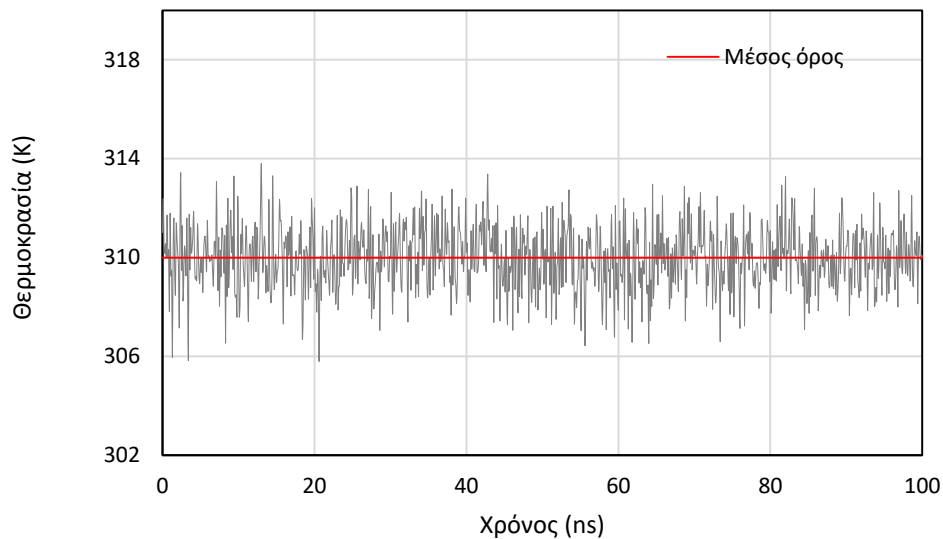
6.1. ΕΙΣΑΓΩΓΗ

ΣΥΣΤΗΜΑ DNA - SRY

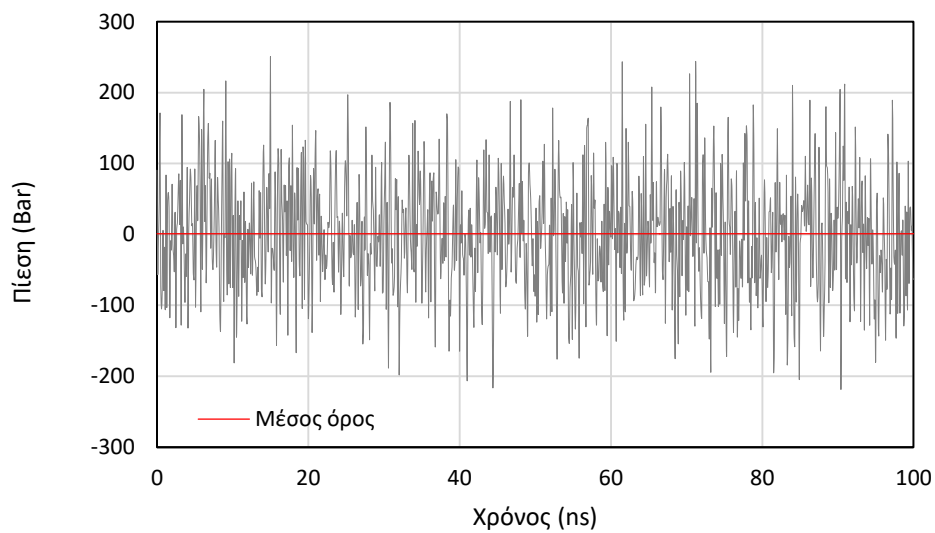
Στα Διαγράμματα 6.1 – 6.4 παρουσιάζονται τα διαγράμματα δυναμικής ενέργειας, θερμοκρασίας, πίεσης και πυκνότητα κατά τη διάρκεια της προσομοίωσης μοριακής δυναμικής για το σύστημα DNA – SRY. Ο μέσος όρος για την δυναμική ενέργεια είναι -903636 kJ/mol, για την θερμοκρασία 310.0 K, για την πίεση 0.95 bar και για την πυκνότητα 994.4 kg/m³.



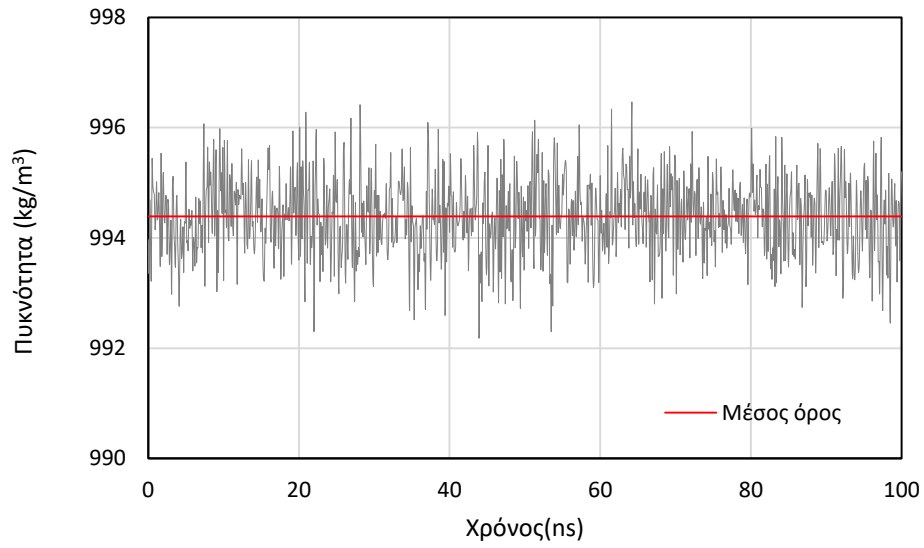
Διάγραμμα 6.1. Χρονική εξέλιξη της δυναμικής ενέργειας κατά την προσομοίωση μοριακής δυναμικής του συστήματος DNA – SRY. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών.



Διάγραμμα 6.2. Χρονική εξέλιξη της θερμοκρασίας κατά την προσομοίωση μοριακής δυναμικής του συστήματος DNA – SRY. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών.



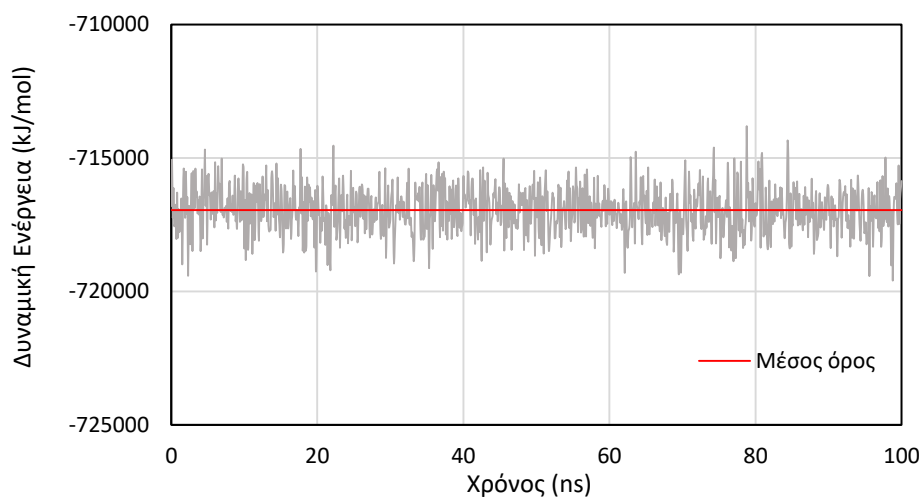
Διάγραμμα 6.3. Χρονική εξέλιξη της πίεσης κατά την προσομοίωση μοριακής δυναμικής του συστήματος DNA – SRY. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών.



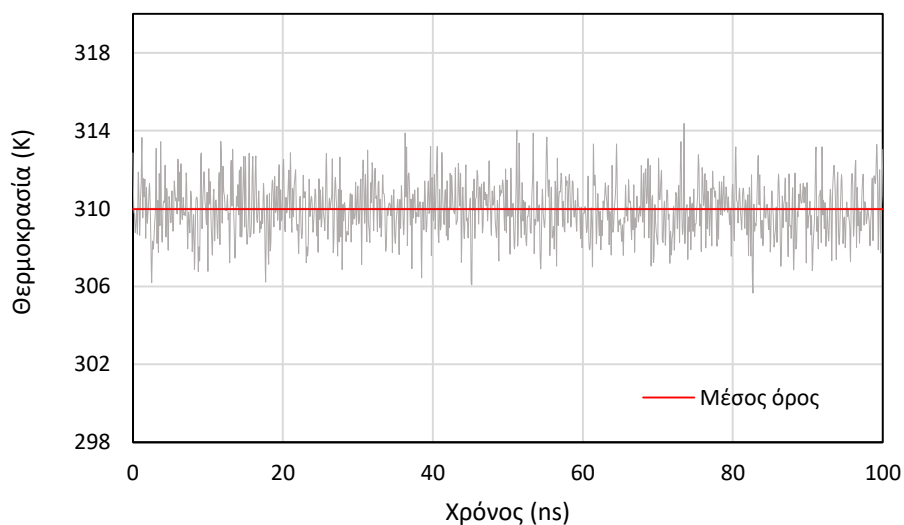
Διάγραμμα 6.4. Χρονική εξέλιξη της πυκνότητας κατά την προσομοίωση μοριακής δυναμικής του συστήματος DNA – SRY. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών.

ΣΥΣΤΗΜΑ ΕΛΕΥΘΕΡΟΥ DNA

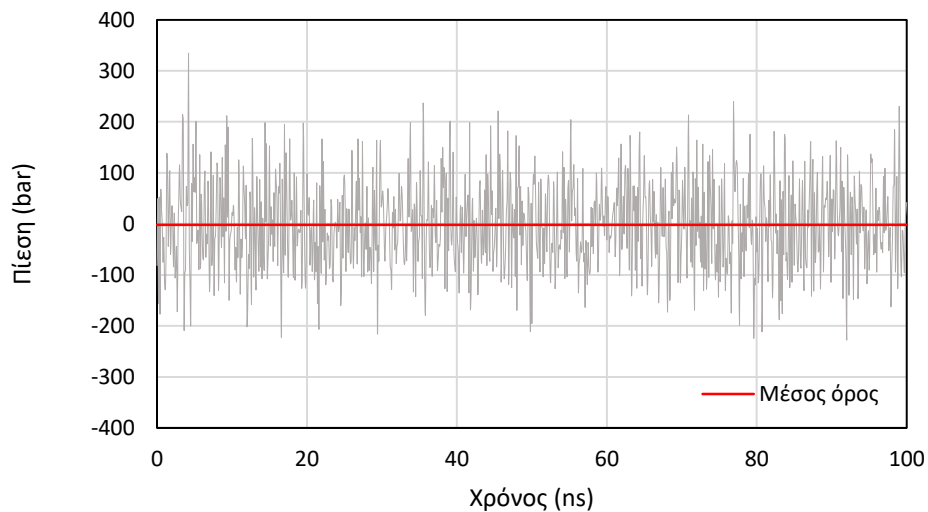
Στα Διαγράμματα 6.5 - 6.8 παρουσιάζονται τα διαγράμματα δυναμικής ενέργειας, θερμοκρασίας, πίεσης και πυκνότητας κατά τη διάρκεια της προσομοίωσης μοριακής δυναμικής για το ελεύθερο DNA. Ο μέσος όρος για τη δυναμική ενέργεια είναι -7167954 kJ/mol, για τη θερμοκρασία 310.0 K, για την πίεση 1.03 bar και για την πυκνότητα 991.5 kg/m³.



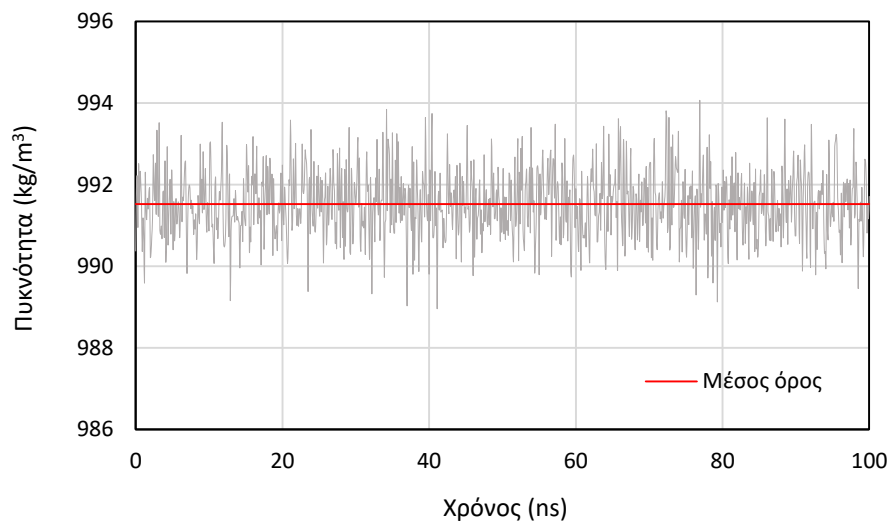
Διάγραμμα 6.5. Χρονική εξέλιξη της δυναμικής ενέργειας κατά την προσομοίωση μοριακής δυναμικής του συστήματος ελεύθερου DNA. Με κόκκινη γραμμή ο μέσος όρος των τιμών.



Διάγραμμα 6.6. Χρονική εξέλιξη της θερμοκρασίας κατά την προσομοίωση μοριακής δυναμικής του συστήματος ελεύθερου DNA. Με κόκκινη γραμμή ο μέσος όρος των τιμών.



Διάγραμμα 6.7. Χρονική εξέλιξη της πίεσης κατά την προσομοίωση μοριακής δυναμικής του συστήματος ελεύθερου DNA. Με κόκκινη γραμμή ο μέσος όρος των τιμών.



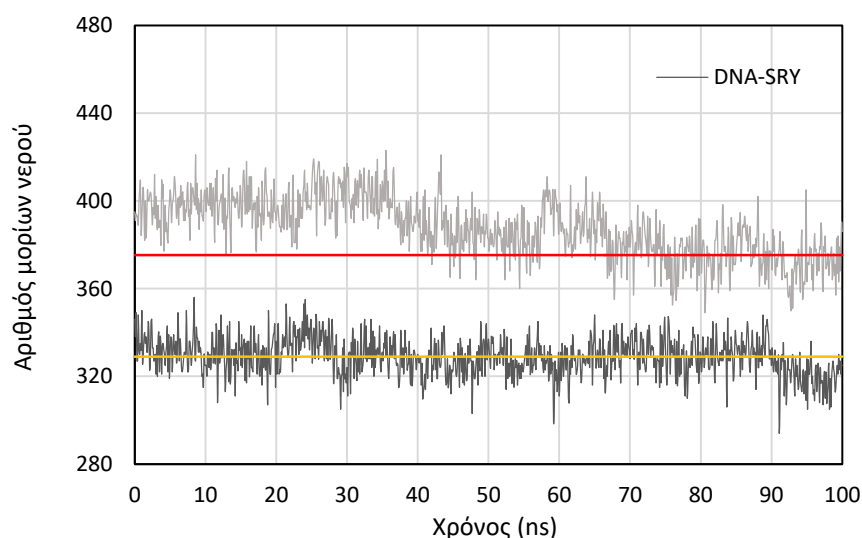
Διάγραμμα 6.8. Χρονική εξέλιξη της πυκνότητας κατά την προσομοίωση μοριακής δυναμικής του συστήματος ελεύθερου DNA. Με κόκκινη γραμμή ο μέσος όρος των τιμών.

6.2. ΚΕΛΥΦΟΣ ΕΝΥΔΑΤΩΣΗΣ ΜΑΚΡΟΜΟΡΙΩΝ

Η ενυδάτωση του DNA είναι ιδιαίτερα σημαντική για τη διαμόρφωσή και τη λειτουργία του.¹ Τα μόρια του νερού που περιβάλλουν το DNA συμβάλλουν στη σύνδεσή του με μικρά μόρια ή πρωτεΐνες. Τα μόρια νερού που περιβάλλουν τις πρωτεΐνες συμβάλλουν στη διαμόρφωση της τριτοταγούς δομής τους. Υπολογίστηκε το κέλυφος ενυδάτωσης από τον αριθμό των μορίων νερού που βρίσκονται σε ακτίνα 0.35 nm από τα άτομα των μακρομορίων. Η απόσταση αυτή επιλέχτηκε έτσι ώστε να εξασφαλίζεται ότι συμπεριλαμβάνεται το πρώτο μονομοριακό στρώμα νερού που είναι ροφημένο στα μακρομόρια. Σημειώνεται ότι κατά IUPAC η απόσταση μεταξύ των ηλεκτρωνητικών ατόμων ενός δεσμού υδρογόνου είναι μικρότερη από 0.35 nm.⁴⁹

ΚΕΛΥΦΟΣ ΕΝΥΔΑΤΩΣΗΣ DNA

Υπολογίστηκε το κέλυφος ενυδάτωσης του DNA για το σύστημα DNA – SRY, και του ελεύθερου DNA, τα οποία παρουσιάζονται στα διαγράμματα που ακολουθούν. Ο μέσος όρος για το DNA στο σύστημα DNA – SRY είναι 329 μόρια νερού. Παρατηρείται ότι το κέλυφος ενυδάτωσης του ελεύθερου DNA μειώνεται με το χρόνο και εμφανίζει ένα πλάτο στα τελευταία 30 ns. Αυτή η μεταβολή υποδηλώνει ότι η διαμόρφωση του ελεύθερου DNA, το οποίο έχει δημιουργηθεί με ψηφιακή ανακατασκευή, χρειάζεται κάποιο χρόνο για να εξισορροπηθεί, σε αντίθεση με τη διαμόρφωση του DNA στο σύστημα DNA – SRY, η οποία έχει ληφθεί από πειραματικά δεδομένα. Έτσι, ο μέσος όρος των μορίων νερού στο κέλυφος ενυδάτωσης του ελεύθερου DNA λαμβάνεται από τα τελευταία 30 ns της εξισορροπημένης τροχιάς και έχει τιμή 375 μόρια.

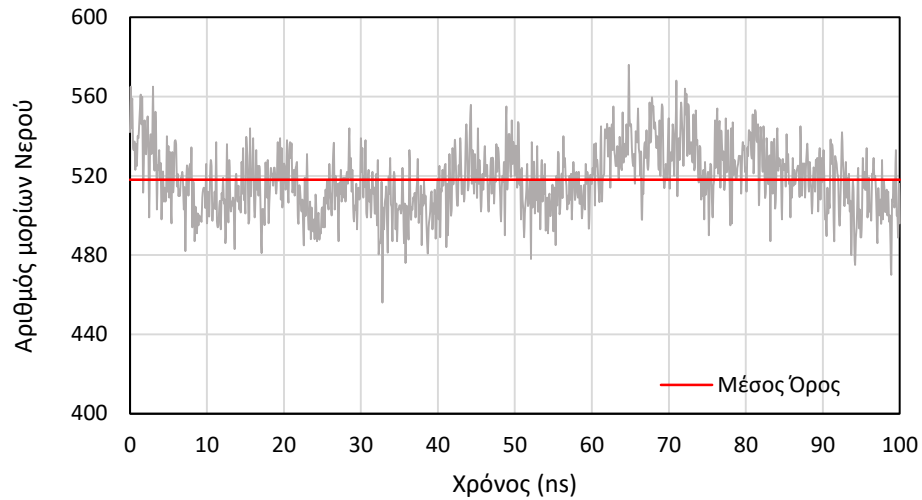


Διάγραμμα 6.9. Χρονική εξέλιξη του κελύφους ενυδάτωσης του DNA για το σύστημα DNA – SRY (σκούρο γκρι χρώμα και μέσος με κίτρινη γραμμή) και το ελεύθερο DNA (ανοιχτό γκρι χρώμα και μέσος με κόκκινη γραμμή).

Παρατηρείται ότι το κέλυφος ενυδάτωσης του DNA μειώνεται κατά 12% (διαφορά 46 μορίων νερού) κατά την πρόσδεσή του με την SRY.

ΚΕΛΥΦΟΣ ΕΝΥΔΑΤΩΣΗΣ SRY

Το κέλυφος ενυδάτωσης της πρωτεΐνης παρουσιάζεται στο διάγραμμα που ακολουθεί. Ο μέσος όρος αριθμών μορίων νερού στο κέλυφος ενυδάτωσης της είναι 518.



Διάγραμμα 6.10. Χρονική εξέλιξη του κελύφους ενυδάτωσης της SRY σε ακτίνα 0.35 nm για το σύστημα DNA – SRY. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των ατόμων νερού στο κέλυφος ενυδάτωσης

6.3. ΔΙΑΜΟΡΦΩΣΗ DNA

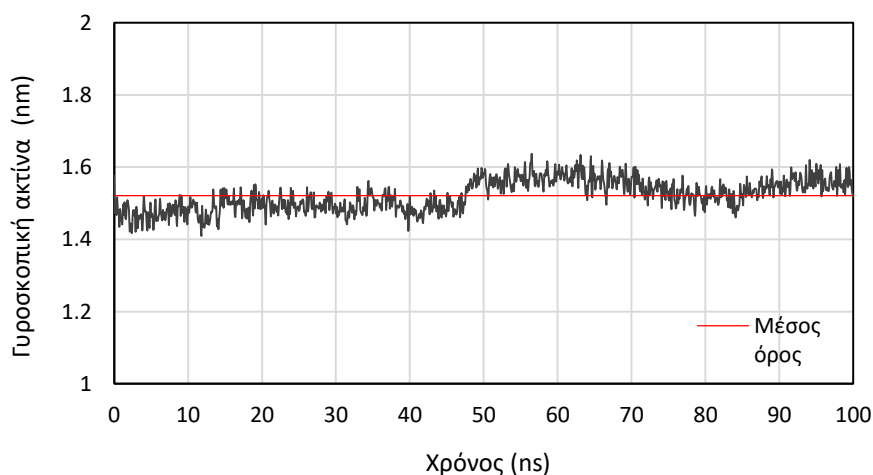
Οι διαστάσεις των μακρομοριακών αλυσίδων μπορούν να χαρακτηριστούν μέσω των αποστάσεων όλων των μονομερών από το κέντρο μάζας. Ένα μέγεθος για τον χαρακτηρισμό τους είναι η γυροσκοπική ακτίνα, R_g , η οποία, για ένα μακρομόριο αποτελούμενο από N άτομα, δίνεται από τη σχέση:⁵⁰

$$R_g = \sqrt{\frac{\sum_i^N m_i |\vec{r}_i|^2}{\sum_i^N m_i}}$$

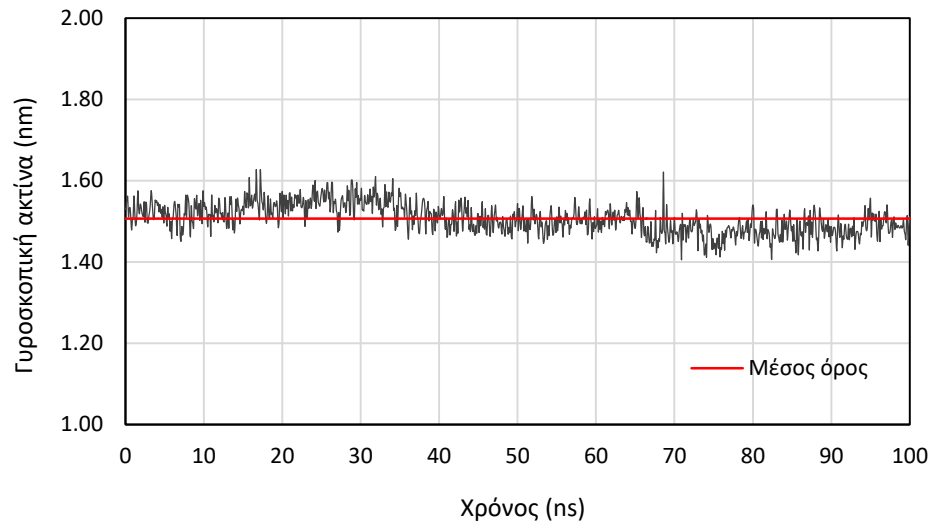
όπου m_i , η μάζα του ατόμου i και \vec{r}_i το διάνυσμα θέσεως αυτού από το κέντρο μάζας του μακρομορίου σε μια συγκεκριμένη απεικόνιση. Η γυροσκοπική ακτίνα παρέχει ένα μέτρο του πόσο εκτεταμένο ή συρρικνωμένο είναι το μακρομόριο.

ΓΥΡΟΣΚΟΠΙΚΗ ΑΚΤΙΝΑ DNA

Υπολογίστηκε η γυροσκοπική ακτίνα του DNA στο σύστημα DNA – SRY και στο σύστημα του ελεύθερου DNA, οι οποίες παρουσιάζονται στα διαγράμματα 6.11 και 6.12. Ο μέσος όρος της γυροσκοπικής ακτίνας για το DNA στο σύστημα DNA – SRY είναι 1.52 nm ενώ στο ελεύθερο DNA είναι 1.53 nm. Οι τιμές της γυροσκοπικής ακτίνας είναι πολύ κοντινές για τα δύο συστήματα, γεγονός που υποδεικνύει ότι η ένωση με την SRY δεν προκαλεί μεταβολή στη γυροσκοπική ακτίνα του DNA.



Διάγραμμα 6.11. Χρονική εξέλιξη της γυροσκοπικής ακτίνας DNA στο σύστημα DNA – SRY. Με κόκκινο χρώμα παρουσιάζεται ο μέσος όρος των τιμών.



Διάγραμμα 6.12. Χρονική εξέλιξη γυροσκοπικής ακτίνας ελεύθερου DNA. Με κόκκινο χρώμα παρουσιάζεται ο μέσος όρος των τιμών.

6.4. ΓΕΩΜΕΤΡΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΗΣ ΔΙΠΛΗΣ ΕΛΙΚΑΣ ΕΛΕΥΘΕΡΟΥ ΚΑΙ ΣΥΜΠΛΕΓΜΕΝΟΥ DNA

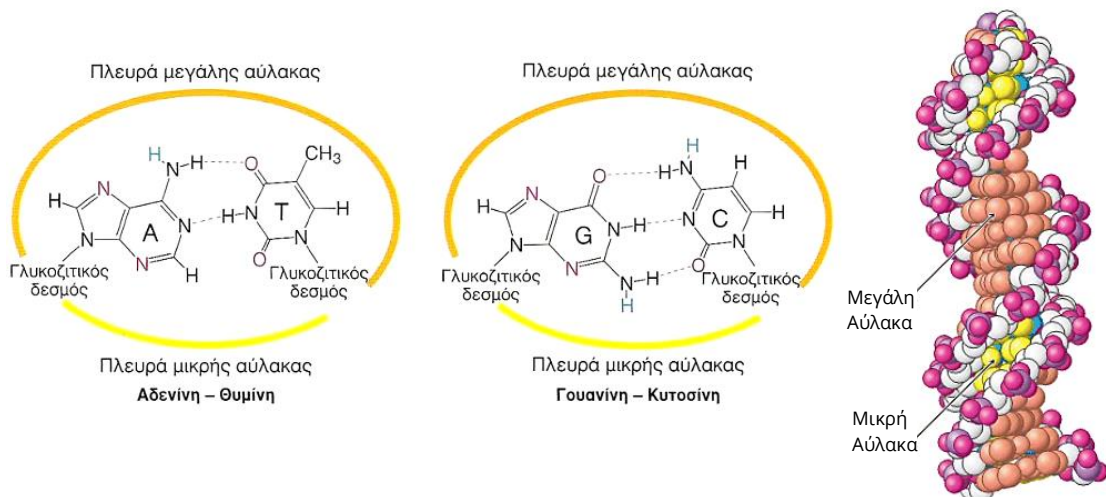
Σε αυτό το στάδιο μελετήθηκε το DNA ως προς την γεωμετρία του, τόσο στην ελεύθερή του μορφή όσο και στην μορφή όπου είναι δεσμευμένο από την SRY. Πραγματοποιήθηκε υπολογισμός των εξής μεγεθών:

- Άξονας διπλής έλικας

Η διπλή έλικα του DNA περιστρέφεται γύρω από έναν άξονα στον οποίο οι βάσεις του DNA είναι σχεδόν κάθετες. Ο άξονας αυτός παρουσιάζεται στην Εικόνα 6.2.

- Εύρος μικρής και μεγάλης αύλακας

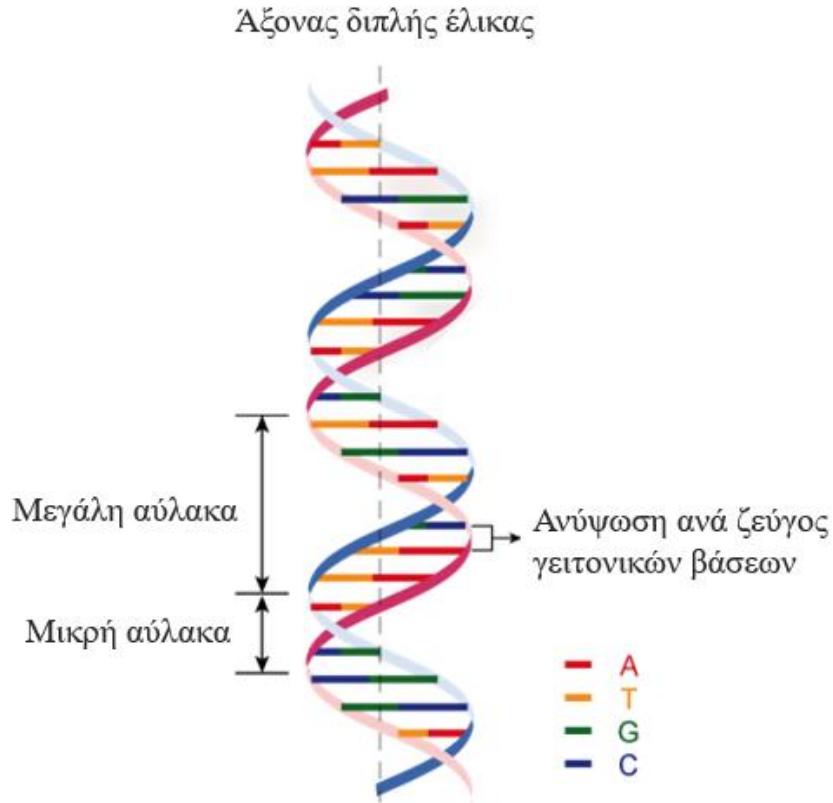
Η διπλή έλικα του DNA αποτελείται από δύο αύλακες: τη μικρή και τη μεγάλη αύλακα. Αυτές δημιουργούνται καθώς οι γλυκοζιτικοί δεσμοί που συνδέουν τις βάσεις ενός ζεύγους με το σκελετό σακχάρου - φωσφόρου δεν είναι αντιδιαμετρικοί. Το O-2 της πυριμιδίνης και το N-3 της πουρίνης του ζεύγους βάσεων βρίσκονται στη μικρή αύλακα, ενώ στη μεγάλη αύλακα περιέχεται η απέναντι πλευρά του ζεύγους και - στην περίπτωση της θυμίνης - η μεθυλική ομάδα. Στο B-DNA, το οποίο αποτελεί την πιο διαδεδομένη μορφή δεοξυριβονουκλεϊκού οξέος, η μικρή αύλακα έχει εύρος 6 Å και βάθος 7.5 Å, ενώ η μεγάλη έχει εύρος 12 Å και βάθος 8.5 Å. Στην Εικόνα 6.1 παρουσιάζονται οι πλευρές της κάθε αύλακας, όπως διαμορφώνονται από τους δεσμούς μεταξύ των ζευγών βάσεων, και στην Εικόνα 6.1 απεικονίζεται η μεγάλη και η μικρή αύλακα του B-DNA.



Εικόνα 6.1. (Αριστερά) Οι πλευρές τις μικρής και της μεγάλης αύλακας για τα ζεύγη βάσεων. (Δεξιά) Η μεγάλη και η μικρή αύλακα στο μόριο του B-DNA. Με κίτρινο χρώμα απεικονίζεται η μικρή αύλακα και με πορτοκαλί η μεγάλη. Με λευκό χρώμα απεικονίζονται τα άτομα άνθρακα του κορμού.¹

- Ανύψωση ανά ζεύγος γειτονικών βάσεων

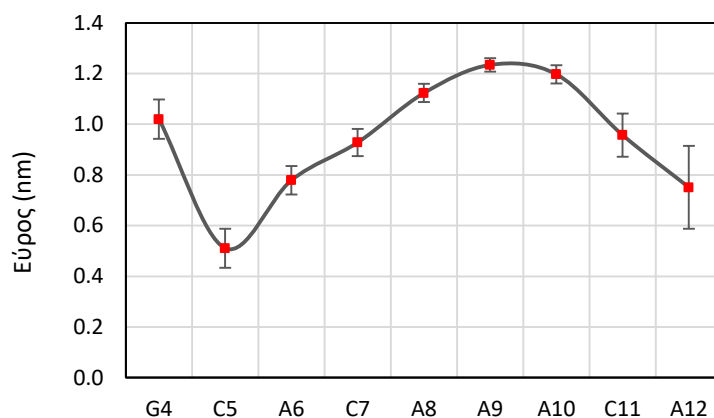
Η ανύψωση ανά ζεύγος γειτονικών βάσεων αναφέρεται στην απόσταση μεταξύ δύο γειτονικών βάσεων που ανήκουν στον ίδιο κλώνο του του δίκλωνου DNA. Αυτές, σε ελεύθερο B-DNA, απέχουν 3.4 Å. Ανά 10.5 βάσεις (34 Å) προκύπτει στροφή της έλικας και ανά βάση υπάρχει στροφή 36° οπότε ανά 34 Å προκύπτει πλήρης στροφή της έλικας, 360°. Η διάμετρος της έλικας είναι 20 Å.¹ Τα χαρακτηριστικά αυτά απεικονίζονται στην Εικόνα 6.2.



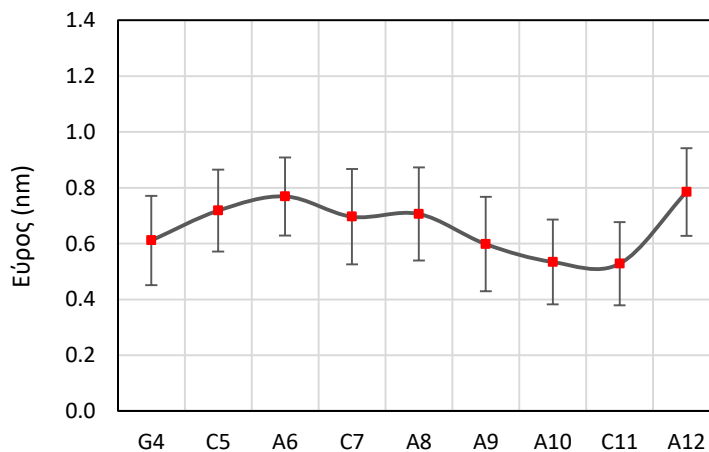
Εικόνα 6.2. Σχεδιαστική αναπαράσταση της δομής του DNA στην οποία παρουσιάζονται ο άξονας της έλικας, το μήκος της στροφής και η διάμετρος της έλικας, η ανύψωση των γειτονικών βάσεων και οι αύλακες του B-DNA.

ΕΥΡΟΣ ΜΙΚΡΗΣ ΑΥΛΑΚΑΣ

Υπολογίστηκε το εύρος της μικρής αύλακας για το DNA στο σύστημα DNA – SRY και στο σύστημα ελεύθερου DNA. Οι μέσοι όροι των τιμών του εύρους της μικρής αύλακας για τα 100 ns της τροχιάς παρουσιάζονται στα διαγράμματα 6.13 και 6.14. Κάθε ζεύγος νουκλεοτιδίων για το οποίο μετρήθηκε το εύρος της μικρής αύλακας αναπαρίσταται στον οριζόντιο άξονα των διαγραμμάτων, από το σύμβολο της βάσης που φέρει το νουκλεοτίδιο της 5'-3' αλυσίδας, συνοδευόμενο από τον αύξοντα αριθμό του, όπως παρουσιάζεται στην Εικόνα 3.2 της ενότητας 3.1. *Μοντελοποίηση μακρομορίων*. Στα διαγράμματα απεικονίζεται επίσης η τυπική απόκλιση για κάθε τιμή του εύρους της μικρής αύλακας.



Διάγραμμα 6.13. Εύρος της μικρής αύλακας του DNA για το σύστημα DNA – SRY. Με κόκκινα τετράγωνα απεικονίζονται οι τιμές που αντιστοιχούν στους μέσους όρους για τα 100 ns.



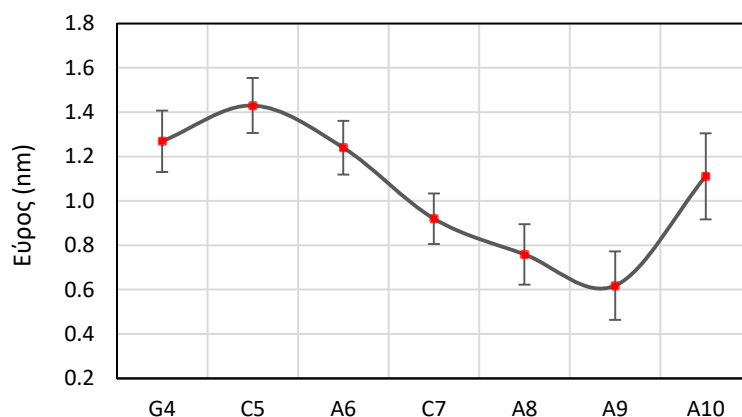
Διάγραμμα 6.14. Εύρος της μικρής αύλακας του DNA για το σύστημα του ελεύθερου DNA. Με κόκκινα τετράγωνα απεικονίζονται οι τιμές που αντιστοιχούν στους μέσους όρους για τα 100 ns.

Παρατηρείται διεύρυνση της μικρής αύλακας στην αλληλουχία CAAAC για το σύστημα DNA - SRY σε σύγκριση με το σύστημα του ελεύθερου DNA. Μάλιστα, το εύρος της μεγάλης αύλακας εμφανίζεται πάνω από διπλάσιο για τα νουκλεοτίδια A9 και A10 και σχεδόν διπλάσιο για τα νουκλεοτίδια A8 στην περίπτωση του συμπλεγμένου DNA.

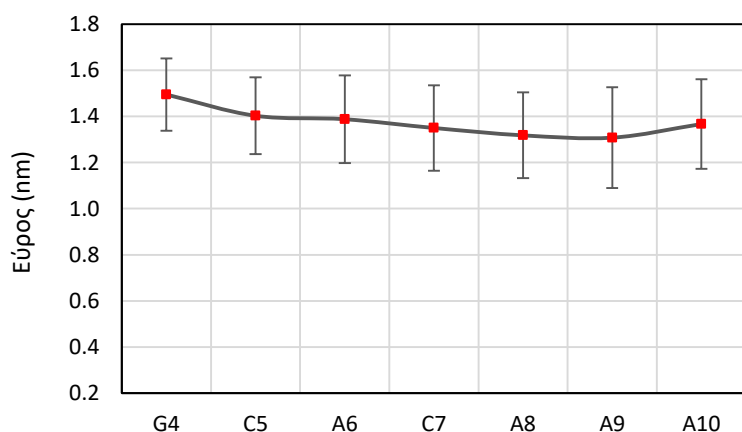
ΕΥΡΟΣ ΜΕΓΑΛΗΣ ΑΥΛΑΚΑΣ

Υπολογίστηκε το εύρος της μεγάλης αύλακας για το DNA στο σύστημα DNA – SRY και στο σύστημα ελεύθερου DNA. Οι μέσοι όροι των τιμών του εύρους της μικρής αύλακας για τα 100 ns της τροχιάς παρουσιάζονται στα Διαγράμματα 6.15 και 6.16. Στα διαγράμματα απεικονίζεται επίσης η τυπική απόκλιση για κάθε τιμή του εύρους μικρής αύλακας.

Παρατηρείται μείωση του εύρους της μεγάλης αύλακας του συμπλεγμένου DNA σε σχέση με αυτό του ελεύθερου DNA στην αλληλουχία CAAA. Συγκεκριμένα, στο νουκλεοτίδια A8 και A9 το εύρος της μεγάλης αύλακας παρουσιάζεται σχεδόν υποδιπλάσιο στην περίπτωση του συμπλεγμένου DNA.



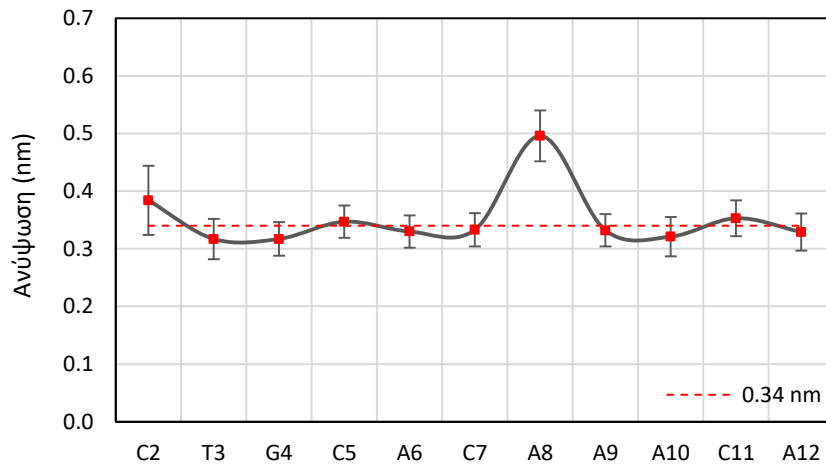
Διάγραμμα 6.15. Εύρος της μεγάλης αύλακας του DNA για το σύστημα DNA – SRY. Με κόκκινα τετράγωνα απεικονίζονται οι τιμές που αντιστοιχούν στους μέσους όρους για τα 100 ns, ενώ εμφανίζονται και οι αντίστοιχες τυπικές αποκλίσεις.



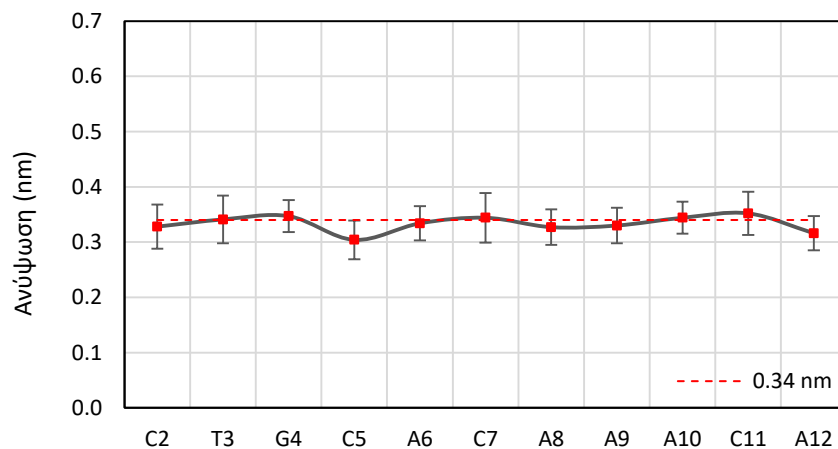
Διάγραμμα 6.16. Εύρος της μεγάλης αύλακας του DNA για το σύστημα του ελεύθερου DNA. Με κόκκινα τετράγωνα απεικονίζονται οι τιμές που αντιστοιχούν στους μέσους όρους για τα 100 ns, ενώ εμφανίζονται και οι αντίστοιχες τυπικές αποκλίσεις.

ΑΝΥΨΩΣΗ ΓΕΙΤΟΝΙΚΩΝ ΒΑΣΕΩΝ

Μετρήθηκε η απόσταση των γειτονικών βάσεων για τα νουκλεοτίδια της 5'-3' έλικας του DNA στο σύστημα DNA – SRY και στο σύστημα του ελεύθερου DNA. Οι μέσοι όροι, από τα 100 ns της τροχιάς, των αποστάσεων των γειτονικών βάσεων παρουσιάζονται στα Διαγράμματα 6.17 και 6.18 μαζί με τις αντίστοιχες τυπικές αποκλίσεις.



Διάγραμμα 6.17. Ανύψωση των γειτονικών βάσεων για το σύστημα DNA – SRY. Με κόκκινα τετράγωνα απεικονίζονται οι τιμές που αντιστοιχούν στους μέσους όρους για τα 100 ns. Με κόκκινη διακεκομμένη γραμμή απεικονίζεται η βιβλιογραφική τιμή¹ για την ανύψωση γειτονικών βάσεων για το DNA.

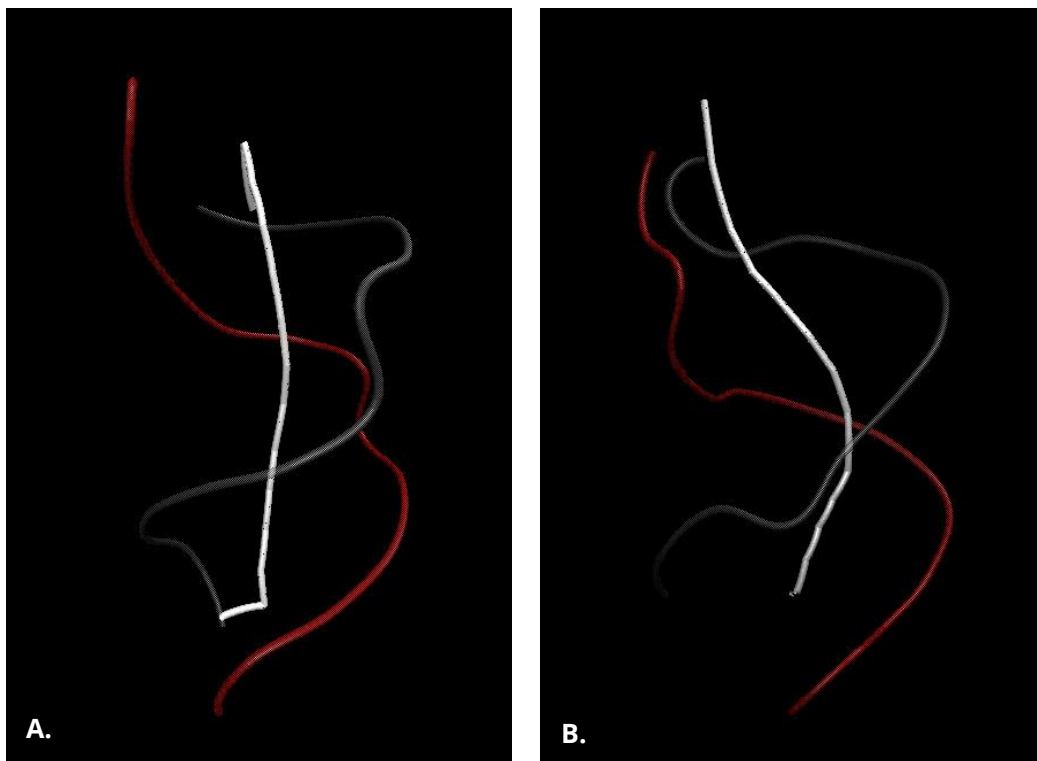


Διάγραμμα 6.18. Ανύψωση γειτονικών βάσεων για το σύστημα DNA – SRY. Με κόκκινες κουκίδες απεικονίζονται οι τιμές που αντιστοιχούν στους μέσους όρους για τα 100 ns. Με κόκκινη διακεκομμένη γραμμή απεικονίζεται η βιβλιογραφική τιμή¹ για την ανύψωση γειτονικών βάσεων για το DNA.

Παρατηρείται αύξηση κατά 50% της ανύψωσης των γειτονικών βάσεων στο νουκλεοτίδιο A8 για το σύστημα DNA – SRY σε σχέση με το σύστημα DNA. Αυτή η αύξηση στην ανύψωση των γειτονικών βάσεων υποδεικνύει την παρεμβολή της SRY στο σημείο αυτό, η οποία εξετάζεται στα κεφάλαια 6.6 και 6.7.

ΑΞΟΝΑΣ ΕΛΙΚΑΣ

Υπολογίστηκε ο άξονας της διπλής έλικας για την τελική διαμόρφωση του DNA στο σύστημα DNA – SRY και στο σύστημα του ελεύθερου DNA. Παρουσιάστηκε εμφανής κάμψη του άξονα στο σύστημα DNA – SRY σε σχέση με το ελεύθερο DNA. Οι άξονες για τα δύο συστήματα παρουσιάζονται στην Εικόνα 6.3.



Εικόνα 6.3. Απεικόνιση του άξονα της διπλής έλικας για το DNA: A. Στο σύστημα ελεύθερου DNA και B. Στο σύστημα DNA – SRY. Με άσπρο χρώμα απεικονίζεται ο άξονας, με κόκκινο ο κορμός της chain B και με γκρι ο κορμός της chain C του DNA.

Εξετάζεται η κάμψη που παρουσιάστηκε στην περίπτωση του δεσμευμένου DNA, στο ύψος της αλληλουχίας CAAAC, στην οποία παρατηρήθηκε η διεύρυνση της μικρής αύλακας και η στένωση της μεγάλης. Υπολογίστηκε η γωνία η οποία περιγράφει την κάμψη του άξονα της έλικας του DNA, όπως αυτή φαίνεται στην Εικόνα 6.4. Για το σκοπό αυτό, επιλέχθηκαν τέσσερα σημεία στον άξονα: AXI4, AXI6, AXI8, AXI11, τα οποία ορίζουν τα διανύσματα:

$$\mathbf{a} = (x_{AXI6} - x_{AXI4}, y_{AXI6} - y_{AXI4}, z_{AXI6} - z_{AXI4}) = (a_1, a_2, a_3)$$

$$\mathbf{b} = (x_{AXI11} - x_{AXI8}, y_{AXI11} - y_{AXI8}, z_{AXI11} - z_{AXI8}) = (b_1, b_2, b_3)$$

Το εσωτερικό γινόμενο των παραπάνω διανυσμάτων δίνεται από την εξίσωση:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

και το μέτρο του κάθε διανύσματος υπολογίζεται από τις σχέσεις:

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2} \quad \text{και} \quad |\mathbf{b}| = \sqrt{b_1^2 + b_2^2 + b_3^2}$$

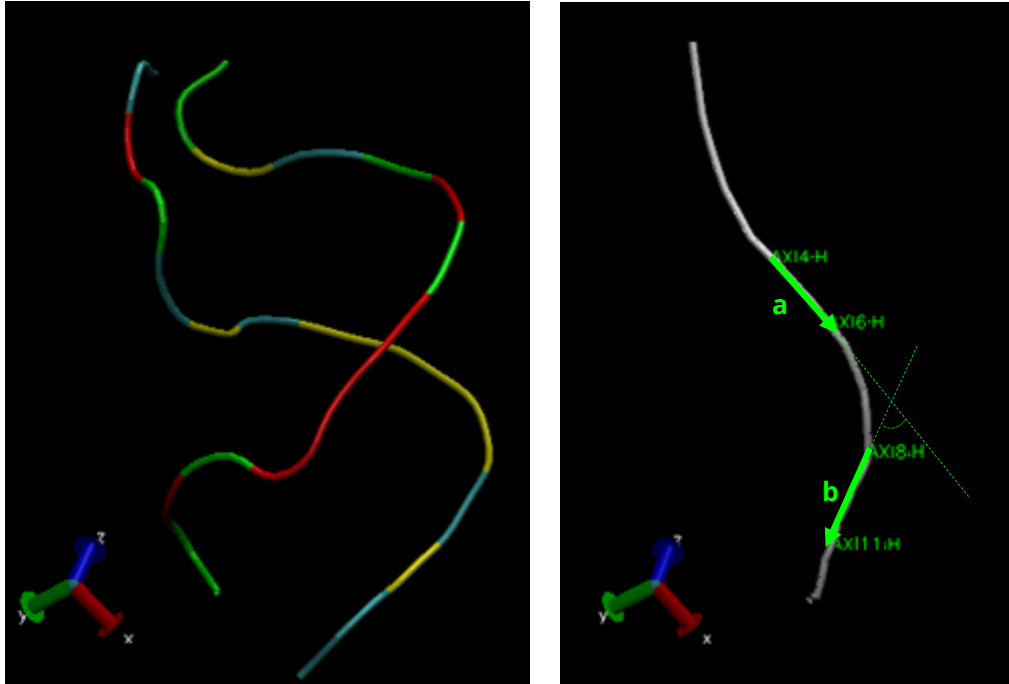
Το συνημίτονο της γωνίας που σχηματίζουν τα διανύσματα \mathbf{a} και \mathbf{b} υπολογίζεται από την εξίσωση:

$$\cos(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|}$$

Από τους παραπάνω υπολογισμούς προέκυψε ότι $(\mathbf{a}, \mathbf{b}) = 63.4^\circ$.

Η τιμή αυτή είναι σε καλή συμφωνία με την πειραματική τιμή της κάμψης του DNA κατά την πρόσδεσή του με τον SRY, η οποία προσδιορίστηκε στις 60° .^{18,51} Η κάμψη αυτή είναι σημαντική καθώς, σύμφωνα με πειραματικές μελέτες, μεταλλάξεις της SRY που περιορίζουν την κάμψη οδηγούν σε σημαντική ελάττωση της μεταγραφής του γονιδίου SOX9 και κατά συνέπεια στην εμφάνιση διαταραχών του φύλου.¹⁸

Στην Εικόνα 6.4, που ακολουθεί, παρουσιάζονται τα σημεία που επιλέχθηκαν για τον ορισμό των διανυσμάτων \mathbf{a} και \mathbf{b} καθώς επίσης και ο κορμός του DNA, σε όμοιο προσανατολισμό, χρωματισμένος ανάλογα με τα νουκλεοτίδια του.



Εικόνα 6.4. (Αριστερά) Απεικόνιση του κορμού του DNA: με κίτρινο χρώμα απεικονίζεται το τμήμα του κορμού που αντιστοιχεί σε νουκλεοτίδια αδενίνης, με κόκκινο θυμίνης, με γαλάζιο κυτοσίνης και με πράσινο γουανίνης. (Δεξιά) Απεικόνιση του άξονα της διπλής έλικας: Με πράσινες ετικέτες παρουσιάζονται τα σημεία που επιλέχθηκαν και ως πράσινα βέλη εμφανίζονται τα διανύσματα τα οποία ορίζονται από τα σημεία.

6.5. ΑΝΑΛΥΣΗ ΑΠΟΣΤΑΣΕΩΝ ΚΑΙ ΕΠΑΦΩΝ ΤΟΥ ΣΥΜΠΛΕΓΜΑΤΟΣ DNA - SRY

Στο στάδιο αυτό εξετάστηκαν οι θέσεις των ατόμων της πρωτεΐνης σε σχέση με τα άτομα του DNA προκειμένου να εξαχθούν συμπεράσματα ως προς τον τρόπο σύνδεσής τους. Υπολογίστηκαν τα εξής μεγέθη:

- **Ελάχιστες αποστάσεις**

Το μέγεθος αυτό αφορά στην ελάχιστη απόσταση που εμφανίζεται την κάθε χρονική στιγμή της τροχιάς μεταξύ δύο ομάδων ατόμων. Για παράδειγμα, αν επιλεχθούν οι ομάδες DNA και SRY τότε για κάθε χρονική στιγμή θα υπολογιστεί μία τιμή, η οποία αντιστοιχεί στη μικρότερη απόσταση που εντοπίζεται μεταξύ των ατόμων DNA και SRY και αφορά σε ένα άτομο του DNA με ένα άτομο της SRY.

- **Αριθμός επαφών**

Το μέγεθος αυτό αφορά στον αριθμό των επαφών μεταξύ των ατόμων τα οποία βρίσκονται σε απόσταση μικρότερη ή ίση με μια ορισμένη τιμή και ανήκουν σε διαφορετικές ομάδες. Για παράδειγμα, αν επιλεχθεί απόσταση 0.3 nm για τις ομάδες DNA και SRY και την χρονική στιγμή $t = 0$ ps υπολογιστούν "5 επαφές", αυτό σημαίνει ότι 5 ζεύγη ατόμων DNA - SRY βρίσκονται σε απόσταση μικρότερη των 0.3 nm. Σημειώνεται ότι ένα άτομο DNA μπορεί να εμφανίζεται σε περισσότερες από μια επαφές με άτομα της SRY (πχ άτομο 1_{DNA} - άτομο 1_{SRY} , άτομο 1_{DNA} - άτομο 2_{SRY} , άτομο 1_{DNA} - άτομο 3_{SRY} μας δίνει 3 επαφές για το άτομο 1_{DNA} σε μια χρονική στιγμή).

Για την ανάλυση των αποστάσεων, τα άτομα της SRY ομαδοποιήθηκαν ως εξής: η πρωτεΐνη αφορά σε όλα τα άτομα της SRY, η κύρια αλυσίδα αφορά στα άτομα της κύριας αλυσίδας (κορμού) της πρωτεΐνης, όπως αυτά περιγράφονται στο κεφάλαιο 1.5, οι πλευρικές αλυσίδες, όπως ορίζονται στο κεφάλαιο 1.5 και η ραχοκοκαλιά της πρωτεΐνης, η οποία αποτελείται από τα άτομα της κύριας αλυσίδας εξαιρουμένων των O και των H. Η ομαδοποίηση αυτή παρουσιάζεται στον Πίνακα 6.1.

Τα άτομα των ατόμων του DNA ομαδοποιήθηκαν ανάλογα με τη βάση που φέρει το νουκλεοτίδιο στο οποίο ανήκουν. Η ομαδοποίηση αυτή παρουσιάζεται στον Πίνακα 6.2.

Πίνακας 6.1. Ομάδες ατόμων SRY

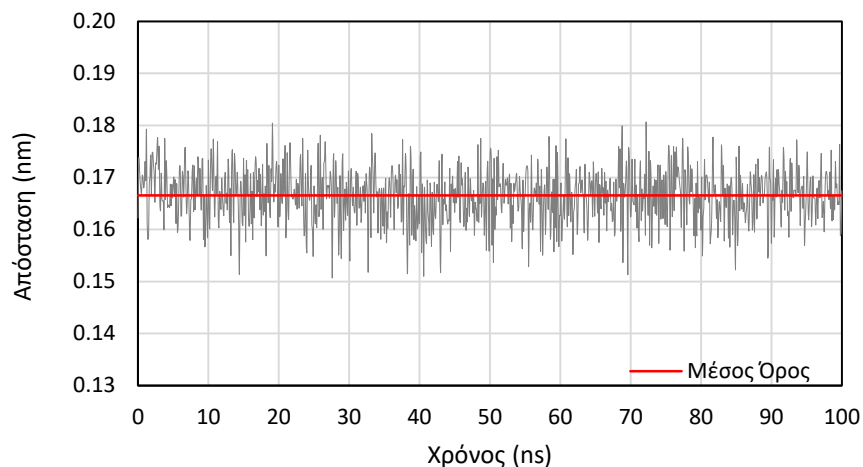
Ομάδα	Είδος Ατόμων	Αριθμός Ατόμων
πρωτεΐνη	Όλα τα άτομα της πρωτεΐνης	1513
πλευρικές αλυσίδες	Όλα τα άτομα των πλευρικών αλυσίδων	1091
κύρια αλυσίδα	N, C _α , C, O (συμπεριλαμβανομένων τον O του C-άκρου της πρωτεΐνης), H των αμιδίων και του N-άκρου της πρωτεΐνης	422
ραχοκοκαλιά	N, C _α , C της κύριας αλυσίδας	255

Πίνακας 6.2. Ομάδες ατόμων DNA

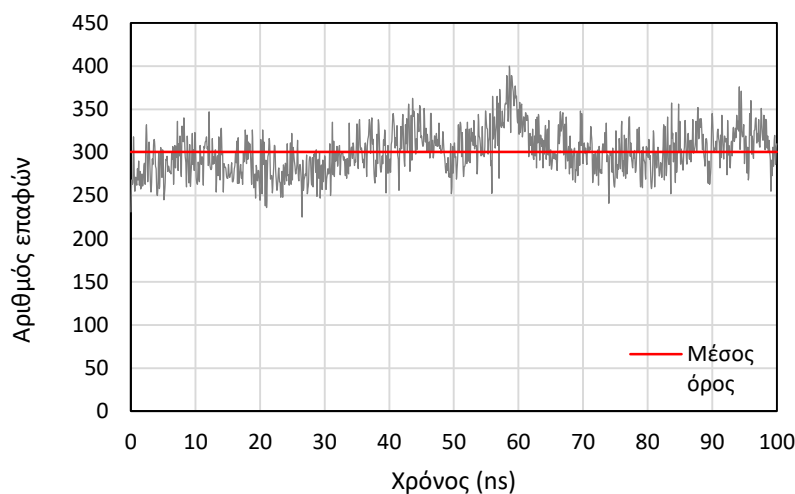
Ομάδα	Είδος Ατόμων	Αριθμός Ατόμων
DNA	Όλα τα άτομα του DNA	886
DG	Άτομα νουκλεοτιδίων με βάση Γουανίνης	263
DA	Άτομα νουκλεοτιδίων με βάση Αδενίνης	192
DC	Άτομα νουκλεοτιδίων με βάση Κυτοσίνης	239
DT	Άτομα νουκλεοτιδίων με βάση Θυμίνης	192

ΑΝΑΛΥΣΗ ΑΠΟΣΤΑΣΕΩΝ

Αρχικά, υπολογίστηκαν οι ελάχιστες αποστάσεις της SRY από το DNA, οι οποίες κυμαίνονται μεταξύ 0.15 nm – 0.18 nm και έχουν μέσο όρο 0.167 nm. Υπολογίστηκε, επίσης ο αριθμός επαφών κάτω των 0.3 nm για το DNA και την SRY, ο οποίος κυμαίνεται μεταξύ 250 – 400 επαφές με μέσο όρο 300.6 επαφές. Αυτά παρουσιάζονται στα Διαγράμματα 6.19 και 6.20.

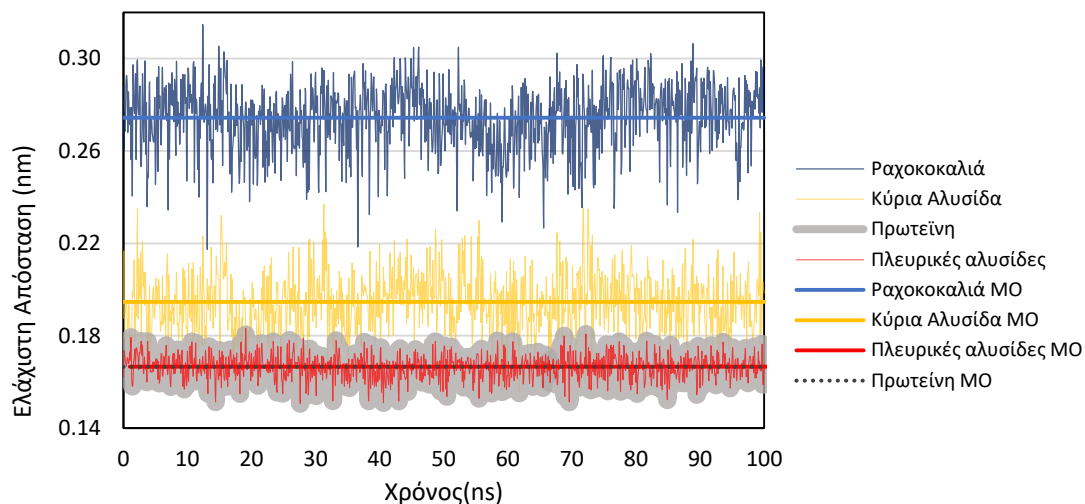


Διάγραμμα 6.19. Χρονική εξέλιξη των ελαχίστων αποστάσεων μεταξύ του DNA και του SRY. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών.

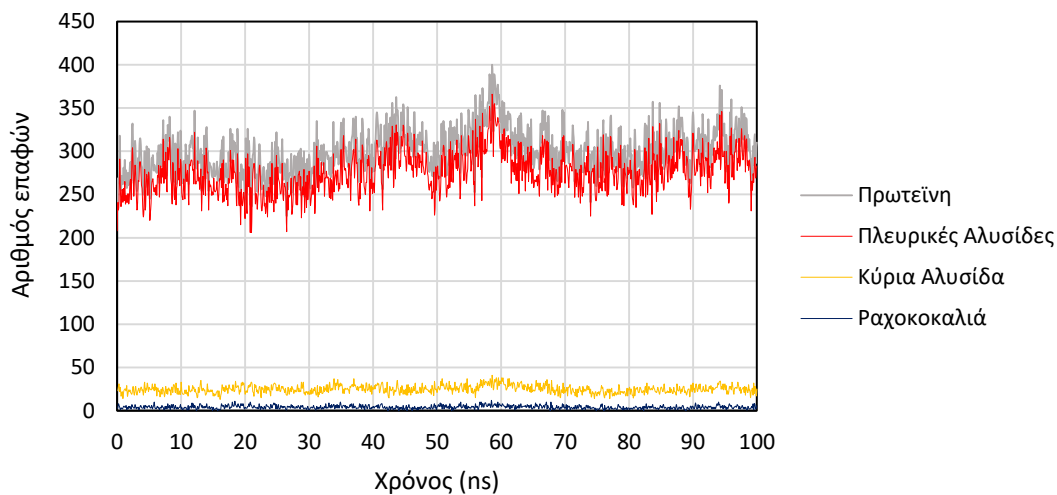


Διάγραμμα 6.20. Χρονική εξέλιξη του αριθμού επαφών μεταξύ του DNA και του SRY. Με κόκκινη γραμμή απεικονίζεται ο μέσος όρος των τιμών.

Στη συνέχεια, υπολογίστηκαν οι ελάχιστες αποστάσεις ομάδων ατόμων της SRY από το DNA και οι αντίστοιχοι αριθμοί επαφών. Ο μέσος όρος ελαχίστων αποστάσεων πρωτεΐνης - DNA είναι 0.167 nm, καθώς αυτή ταυτίζεται με την SRY, ο μέσος όρος για τις πλευρικές αλυσίδες - DNA είναι 0.167 nm, για την κύρια αλυσίδα - DNA είναι 0.195 nm και για τη ραχοκοκαλιά - DNA 0.274 nm. Ο μέσος όρος αριθμών επαφών για την πρωτεΐνη είναι 300.6 καθώς αυτή ταυτίζεται με την SRY, για τις πλευρικές αλυσίδες είναι 275.7, για την κύρια αλυσίδα είναι 24.9 και για τη ραχοκοκαλιά 4.1. Τα μεγέθη αυτά παρουσιάζονται στα Διαγράμματα 6.21 και 6.22.



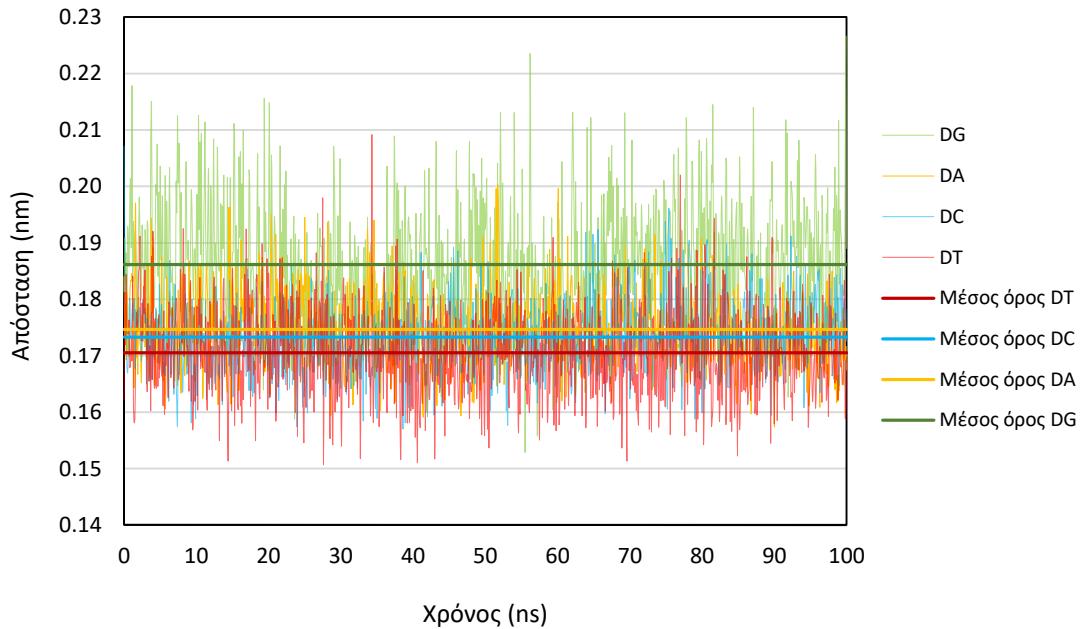
Διάγραμμα 6.21. Χρονική εξέλιξη των ελαχίστων αποστάσεων των ομάδων του SRY από το DNA. Με γκρι χρώμα παρουσιάζονται οι ελάχιστες αποστάσεις της πρωτεΐνης – DNA, με κόκκινο των πλευρικών αλυσίδων – DNA, με κίτρινο της κύριας αλυσίδας – DNA και με μπλε της ραχοκοκαλιάς – DNA. Με γραμμές ίδιου χρώματος παρουσιάζονται οι αντίστοιχοι μέσοι όροι.



Διάγραμμα 6.22. Χρονική εξέλιξη του αριθμού επαφών μεταξύ των ομάδων ατόμων του SRY και του DNA. Με γκρι χρώμα παρουσιάζονται οι επαφές πρωτεΐνης – DNA, με κόκκινο χρώμα οι επαφές πλευρικών αλυσίδων – DNA, με κίτρινο της κύριας αλυσίδας – DNA και με μπλε της ραχοκοκαλιάς – DNA.

Παρατηρείται ότι η ελάχιστη απόσταση πλευρικών αλυσίδων - DNA ταυτίζεται με την ελάχιστη απόσταση πρωτεΐνης - DNA. Επιπλέον τα άτομα των πλευρικών αλυσίδων παρουσιάζουν πολύ μεγαλύτερους αριθμούς επαφών σε σχέση με τα άτομα των άλλων ομάδων της SRY. Η κύρια αλυσίδα παρουσιάζει μεγαλύτερη ελάχιστη απόσταση και μικρότερο αριθμό επαφών από τις πλευρικές αλυσίδες, ενώ στο τέλος έρχεται η ραχοκοκαλιά της SRY.

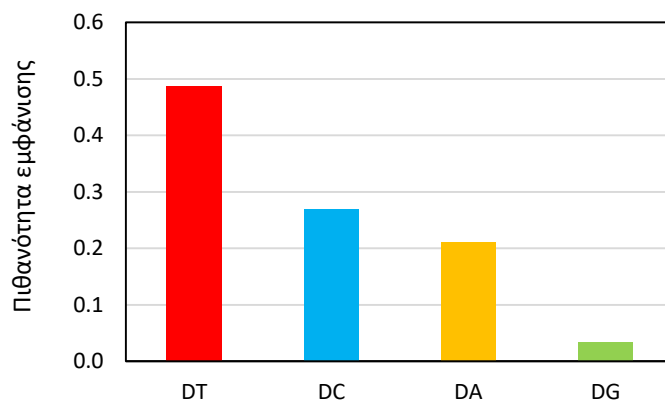
Εν συνεχεία, λόγω της εγγύτητας των πλευρικών αλυσίδων με το DNA που παρατηρήθηκε, μελετώνται οι ελάχιστες αποστάσεις των πλευρικών αλυσίδων από τις ομάδες ατόμων του DNA. Αυτές παρουσιάζονται στο Διάγραμμα 6.3. Οι ελάχιστες αποστάσεις πλευρικών αλυσίδων - DT έχουν μέσο όρο 0.171 nm, οι ελάχιστες αποστάσεις πλευρικών αλυσίδων - DC έχουν μέσο όρο 0.173 nm, οι αντίστοιχες για τις πλευρικές αλυσίδες - DA έχουν μέσο όρο 0.175 nm και αυτές για τις πλευρικές αλυσίδες - DG έχουν μέσο όρο 0.186 nm.



Διάγραμμα 6.23. Χρονική εξέλιξη των ελαχίστων αποστάσεων των πλευρικών αλυσίδων από τα νουκλεοτίδια του DNA: νουκλεοτίδια γουανίνης (DG, πράσινο), νουκλεοτίδια αδενίνης (DA, κίτρινο), νουκλεοτίδια κυτοσίνης (DC, γαλάζιο), νουκλεοτίδια θυμίνης (DT, κόκκινο). Με έντονες ευθείες γραμμές απεικονίζονται οι αντίστοιχοι μέσοι όροι.

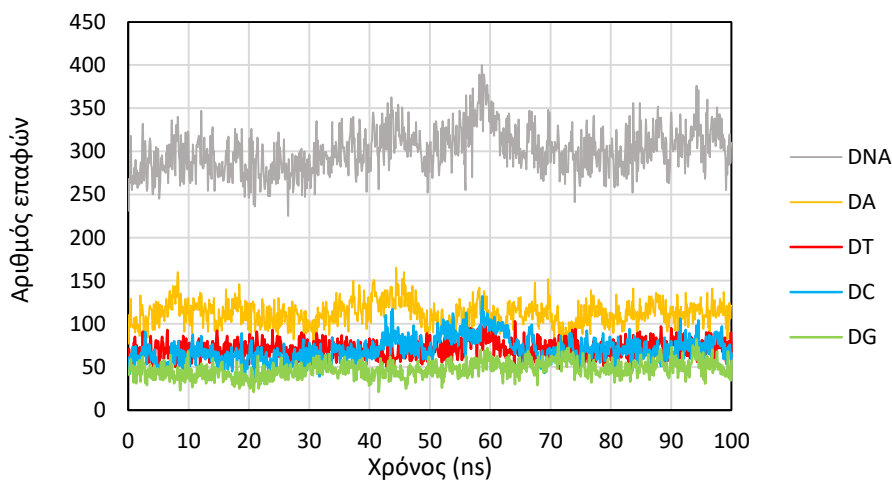
Τα νουκλεοτίδια που φέρουν βάσεις θυμίνης παρουσιάζουν μικρότερες ελάχιστες αποστάσεις από τις πλευρικές αλυσίδες σε σχέση με τα υπόλοιπα νουκλεοτίδια. Ακολουθούν τα νουκλεοτίδια που φέρουν βάσεις κυτοσίνης και στη συνέχεια αυτά που φέρουν βάσεις θυμίνης, ενώ σε μεγαλύτερες αποστάσεις βρίσκονται τα νουκλεοτίδια που φέρουν βάσεις γουανίνης.

Υπολογίστηκε, επίσης, η πιθανότητα του κάθε είδους νουκλεοτιδίων να βρεθεί σε ελάχιστη απόσταση από άτομα της SRY. Κατά τη διάρκεια της τροχιάς τα νουκλεοτίδια που φέρουν βάσεις θυμίνης εμφανίστηκαν με πιθανότητα 0.49, κυτοσίνης με πιθανότητα 0.27, αδενίνης με πιθανότητα 0.21 και γουανίνης με πιθανότητα 0.03. Οι τιμές αυτές παρουσιάζονται στο Διάγραμμα 6.24.



Διάγραμμα 6.24. Πιθανότητα εμφάνισης των νουκλεοτιδίων του DNA σε ελάχιστη απόσταση με την SRY.

Για την περαιτέρω διερεύνηση των αποστάσεων από την SRY ανά νουκλεοτίδιο υπολογίστηκε ο αριθμός επαφών εντός 0.3 nm των νουκλεοτιδίων του DNA από την SRY. Αυτές παρουσιάζονται στο διάγραμμα 6.25. Ο μέσος όρος των αριθμών επαφών μεταξύ του DNA και της SRY είναι, όπως προαναφέρθηκε, 300.6. Ο μέσος όρος των αριθμών επαφών για τα νουκλεοτίδια που φέρουν βάση αδενίνης είναι 112.0, των νουκλεοτιδίων με βάση θυμίνης είναι 71.1, των νουκλεοτιδίων με βάση κυτοσίνης είναι 70.1 και των νουκλεοτιδίων με βάση γουανίνης είναι 47.4.

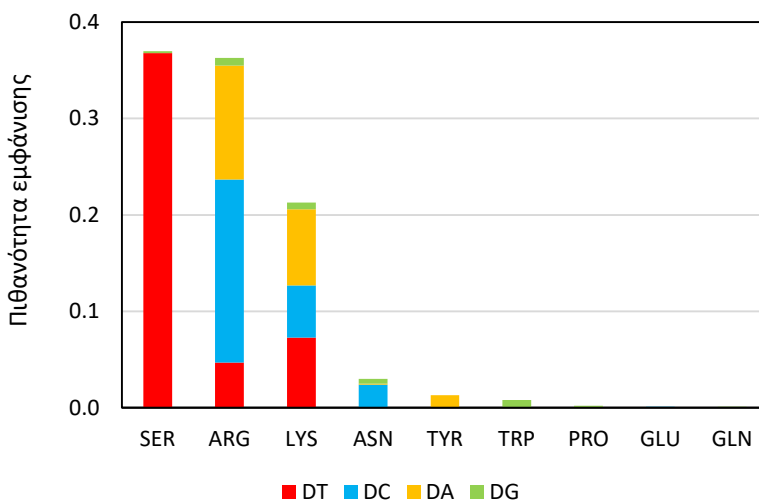


Διάγραμμα 6.25. Χρονική εξέλιξη του αριθμού επαφών του SRY με το DNA (γκρι), τα νουκλεοτίδια του DNA που φέρουν βάση αδενίνης (κίτρινο), θυμίνης (κόκκινο), κυτοσίνης (γαλάζιο), γουανίνης (πράσινο).

Παρατηρείται ότι τα νουκλεοτίδια που φέρουν βάσεις αδενίνης παρουσιάζουν περισσότερες επαφές με την πρωτεΐνη σε σχέση με τα υπόλοιπα. Ακολουθούν τα νουκλεοτίδια με βάσεις θυμίνης, μετά με κυτοσίνης και μετά με γουανίνης.

Ακολούθησε υπολογισμός της πιθανότητας εμφάνισης σε ελάχιστη απόσταση με το DNA του κάθε αμινοξέος, και η πιθανότητα του κάθε νουκλεοτιδίου να βρίσκεται σε ελάχιστη απόσταση με το αντίστοιχο αμινοξύ. Αυτές παρουσιάζονται στο Διάγραμμα 6.26. Αμινοξέα σερίνης εμφανίζονται σε ελάχιστη απόσταση με πιθανότητα 0.370, με ελάχιστες αποστάσεις σερίνης – DT να έχουν πιθανότητα 0.368 και σερίνης – DG να εμφανίζονται με πιθανότητα 0.002. Η πιθανότητα εμφάνισης αργινίνης είναι 0.363, λυσίνης 0.213, ασπαραγίνης 0.003, προλίνης 0.002, γλουταμινικού οξέος 0.001 και γλουταμίνης 0.001.

Συγκεκριμένα, 37% των ελαχίστων αποστάσεων DNA – SRY αντιστοιχούν σε αποστάσεις του αμινοξέος 36SER με το νουκλεοτίδιο T19. Το νουκλεοτίδιο T19 είναι το ζεύγος του νουκλεοτιδίου A10, στο οποίο παρατηρείται μεγάλη διεύρυνση της μικρής αύλακας, όπως αναφέρθηκε στο κεφάλαιο 6.4.

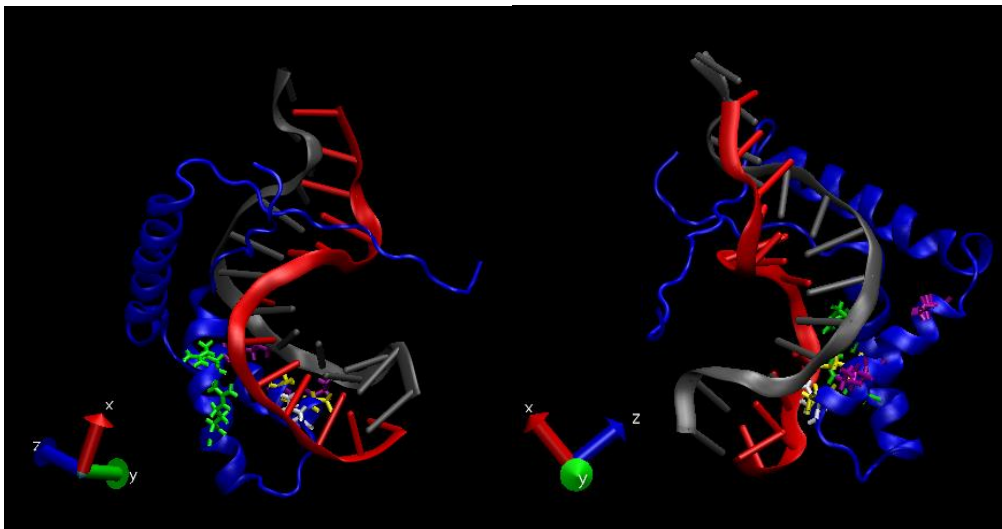


Διάγραμμα 6.26. Πιθανότητα εμφάνισης των αμινοξέων της SRY σε ελάχιστη απόσταση με τα νουκλεοτίδια του DNA.

Εξετάστηκαν τα αμινοξέα που εμφανίζονται σε ελάχιστη απόσταση με πιθανότητα μεγαλύτερη του 1%. Από αυτά, στη δεύτερη α-έλικα της SRY ανήκουν οι σερίνες (36SER, 33SER), μία λυσίνη (37LYS), και η ασπαραγίνη 32ASN και η λυσίνη 44LYS, οι οποίες βρίσκονται στην αρχή και το τέλος αντίστοιχα της δεύτερης α-έλικας. Στην πρώτη α-έλικα της SRY ανήκουν οι αργινίνες η 17ARG και 21ARG. Τα υπόλοιπα αμινοξέα (78ARG, 75ARG, 73LYS, 79LYS, 81LYS, 85LYS, 6LYS) ανήκουν στους βρόχους στα άκρα της πρωτεΐνης.

Συνεπώς, τα αμινοξέα της δεύτερης α-έλικας της SRY είναι αυτά που παρουσιάζονται πιο συχνά σε ελάχιστη απόσταση, υποδεικνύοντας ότι η έλικα αυτή προσεγγίζει πολύ το DNA. Ακολουθούν τα αμινοξέα της πρώτης α-έλικας. Τα αμινοξέα που βρίσκονται σε ελάχιστη απόσταση από το DNA με πιθανότητα μεγαλύτερη του 1% και ανήκουν στην πρώτη και δεύτερη α-έλικα παρουσιάζονται στην Εικόνα 6.5.

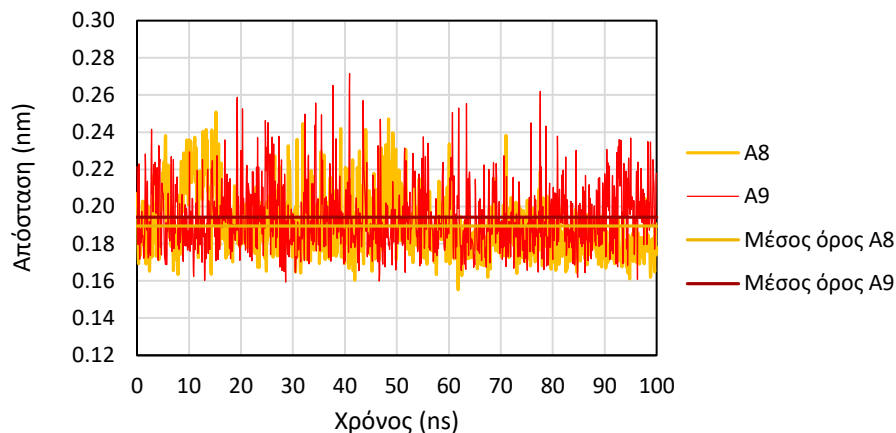
Τα αμινοξέα στα άκρα της πρωτεΐνης που επίσης εμφανίζονται συχνά σε ελάχιστη απόσταση, δεν μπορούν να είναι ενδεικτικά της τοποθέτησης του SRY ως προς το DNA, καθώς στην πραγματικότητα η πρωτεΐνη δεν διακόπτεται στα άκρα αυτά.



Εικόνα 6.5. Αμινοξέα από εμφανίζονται σε ελάχιστη απόσταση με το DNA με πιθανότητα > 1%. Με κόκκινο χρώμα απεικονίζεται η chain B και με σκούρο γκρι η chain C του DNA, ενώ με μπλε η πρωτεΐνη. Επισημαίνονται με πράσινο χρώμα οι αργινίνες, με κίτρινο οι σερίνες, με μωβ οι λυσίνες, και με άσπρο η ασπαραγίνη. (Αριστερά) Όψη στην οποία βρίσκεται μπροστά η πρώτη α-έλικα. (Δεξιά) Όψη στην οποία βρίσκεται μπροστά η δεύτερη α-έλικα.

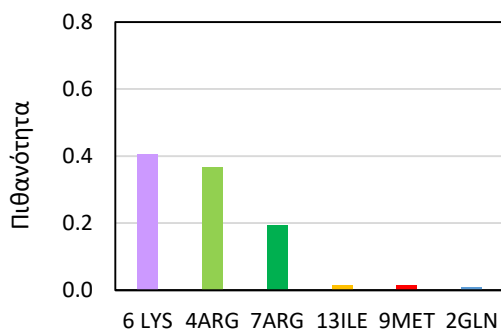
6.6. ΝΟΥΚΛΕΟΤΙΔΙΑ ΑΔΕΝΙΝΗΣ ΚΑΙ ΠΡΟΣΔΕΣΗ ΤΗΣ SRY

Λόγω της ανύψωσης μεταξύ των βάσεων των νουκλεοτιδίων A8 και A9 που παρατηρήθηκε, υπολογίστηκαν οι ελάχιστες αποστάσεις αυτών από την SRY. Αυτές κυμαίνονται μεταξύ 0.16-0.26 nm για την A8 και 0.16-0.27 nm για την A9 με μέσου όρους 0.190 nm και 0.194 nm αντίστοιχα, και παρουσιάζονται στο Διάγραμμα 6.27.

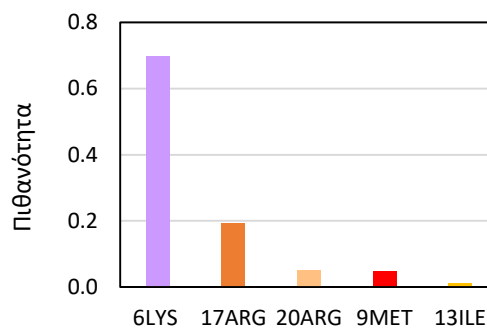


Διάγραμμα 6.27. Χρονική εξέλιξη των ελαχίστων αποστάσεων των νουκλεοτιδίων A8 και A9 από την SRY. Με κίτρινο απεικονίζονται οι ελάχιστες αποστάσεις μεταξύ της SRY και του νουκλεοτιδίου A8 και ο αντίστοιχος μέσος όρος και με κόκκινο οι ελάχιστες αποστάσεις μεταξύ της SRY και του νουκλεοτιδίου A9 και ο μέσος όρος τους.

Στα Διαγράμματα 6.28 και 6.29 παρουσιάζεται η πιθανότητα του κάθε αμινοξέος της SRY να βρεθεί σε ελάχιστη απόσταση από το νουκλεοτίδιο A8 και A9 αντίστοιχα.



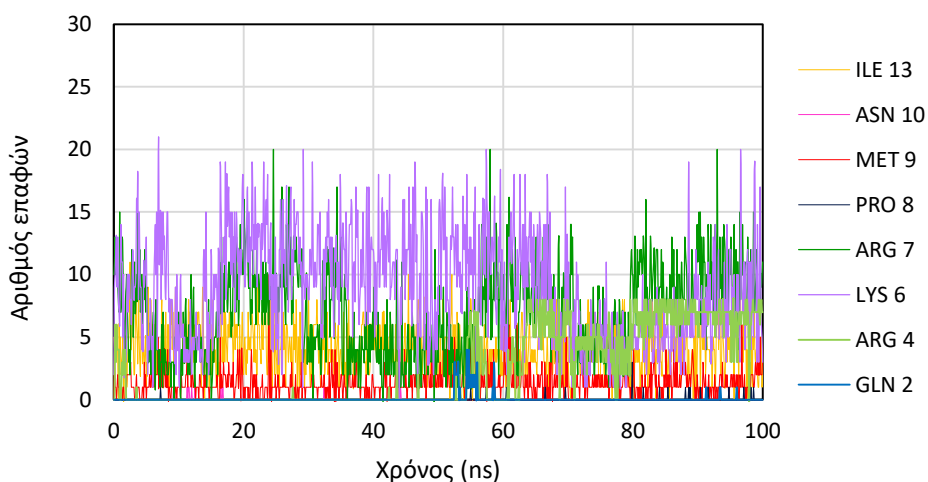
Διάγραμμα 6.28. Πιθανότητα εμφάνισης νουκλεοτιδίων σε ελάχιστη απόσταση με A8.



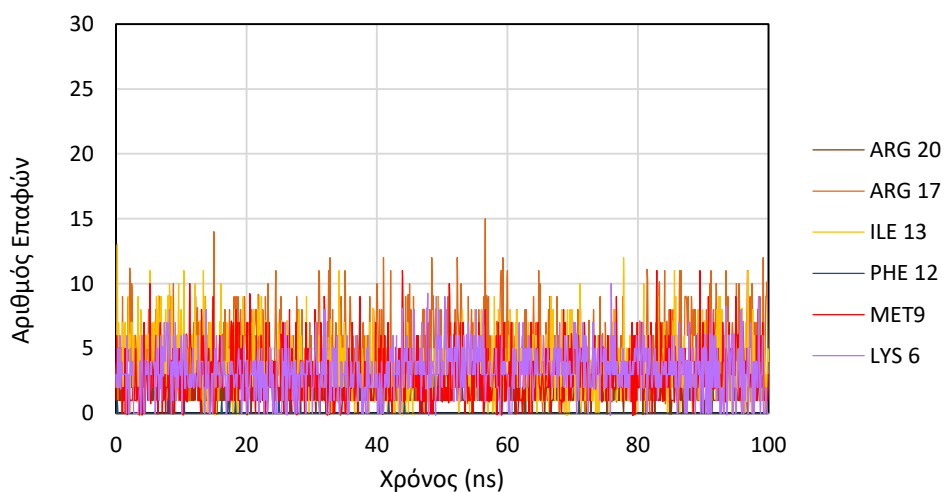
Διάγραμμα 6.29. Πιθανότητα εμφάνισης νουκλεοτιδίων σε ελάχιστη απόσταση με A9.

Παρατηρείται ότι στις ελάχιστες αποστάσεις SRY – A8 εμφανίζεται το αμινοξύ 6LYS με πιθανότητα 0.406, 4ARG με πιθανότητα 0.368, 7ARG με πιθανότητα 0.193, 13ILE με πιθανότητα 0.014, 9MET με πιθανότητα 0.013 και 2GLN με πιθανότητα 0.007. Στις ελάχιστες αποστάσεις SRY – A9 εμφανίζεται το αμινοξύ 6LYS με πιθανότητα 0.696, 17ARG με πιθανότητα 0.192, 20ARG με πιθανότητα 0.051, 9MET με πιθανότητα 0.049 και 13ILE με πιθανότητα 0.012.

Υπολογίστηκαν οι αριθμοί επαφών εντός 0.3 nm του κάθε αμινοξέος της SRY με τα νουκλεοτίδια A8 και A9 με την SRY, η χρονική εξέλιξη των οποίων παρουσιάζεται στα Διαγράμματα 6.30 και 6.31:

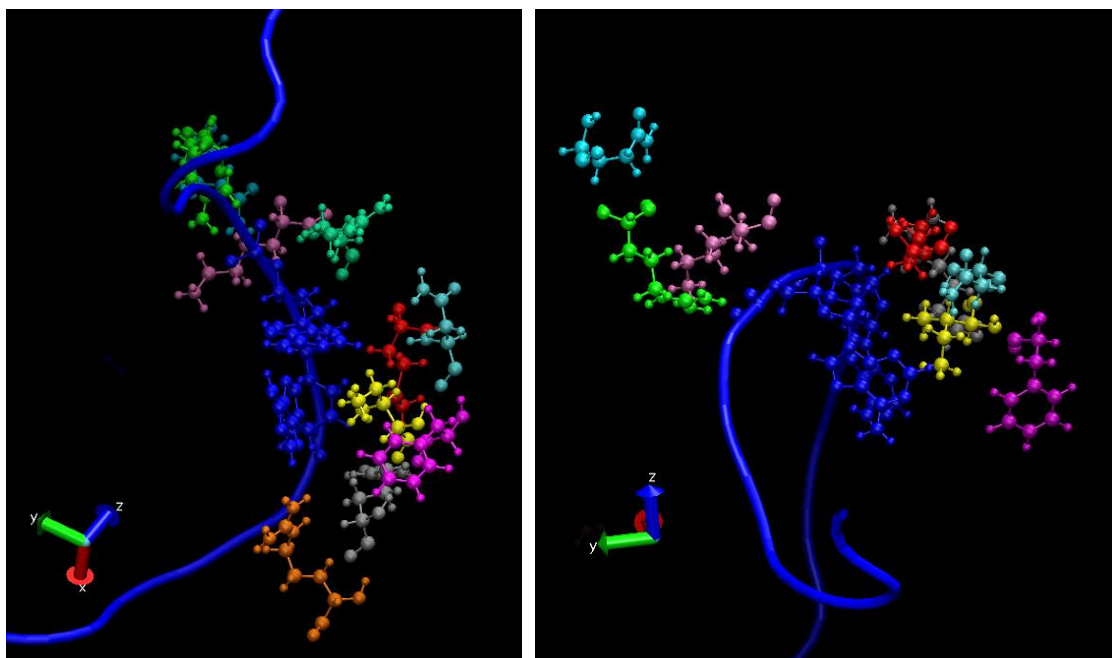


Διάγραμμα 6.30. Χρονική εξέλιξη του αριθμού των επαφών μεταξύ διαφορετικών αμινοξέων της SRY και του νουκλεοτιδίου A8.



Διάγραμμα 6.31. Χρονική εξέλιξη του αριθμού των επαφών μεταξύ διαφορετικών αμινοξέων της SRY και του νουκλεοτιδίου A9.

Τα αμινοξέα τα οποία εμφανίζονται στα Διαγράμματα 6.28-6.31 απεικονίστηκαν προκειμένου να εξεταστεί η θέση τους ως προς τα νουκλεοτίδια A8 και A9. Η απεικόνισή τους παρουσιάζεται στην Εικόνα 6.6.



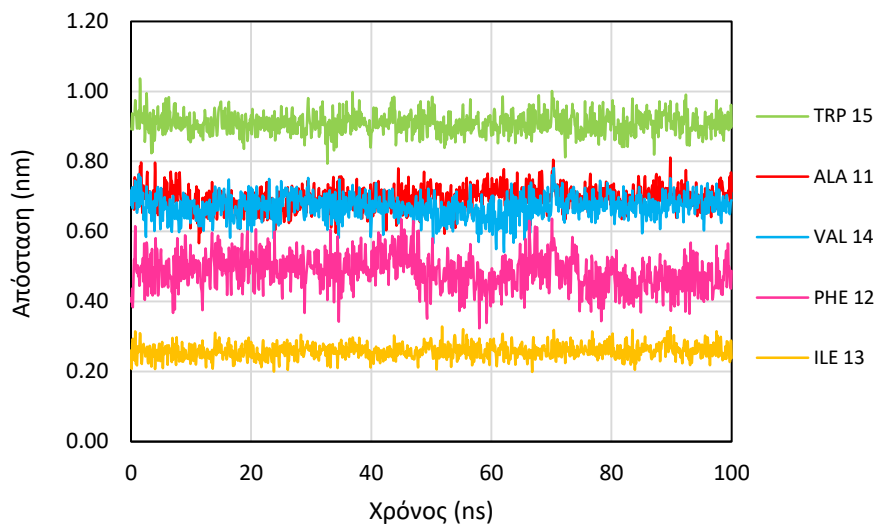
A)

B)

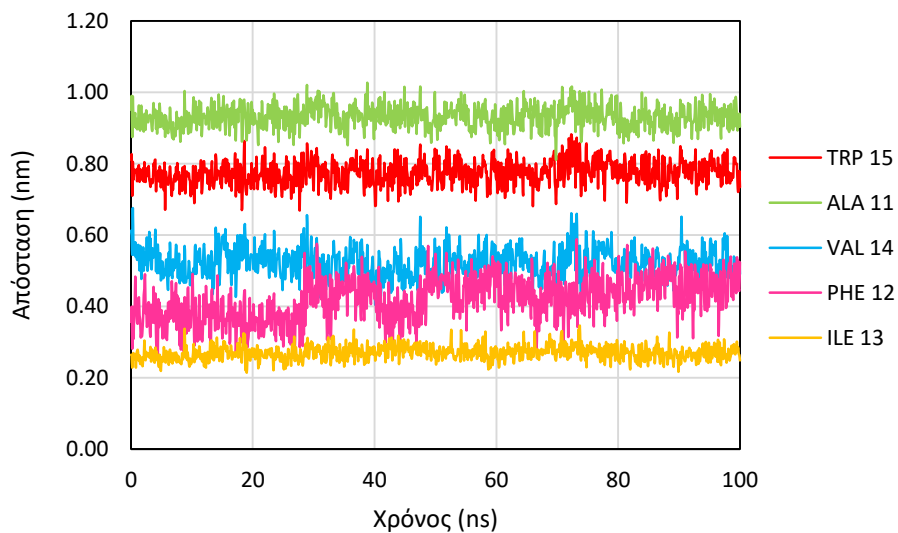
Εικόνα 6.6. Απεικόνιση των αμινοξέων που βρέθηκαν σε ελάχιστες αποστάσεις και σε επαφές εντός των 0.3 nm με τα A8 και A9. (A) Όψη κατά μήκος του άξονα της διπλής έλικας. (B) Όψη σχεδόν εγκάρσια στον άξονα της διπλής έλικας. Ως μπλε σωλήνας απεικονίζεται η νοητή γραμμή που συνδέει τους φωσφόρους των νουκλεοτιδίων του ένα κλώνου της διπλής έλικας. Με μπλε χρώμα απεικονίζονται τα νουκλεοτίδια A8 και A9. Με κίτρινο χρώμα απεικονίζεται η 13ILE, με σκούρο ροζ η 6LYS, με φωτεινό πράσινο η 4ARG, σκούρο πράσινο η 7ARG, με πορτοκαλί η 20ARG, με κόκκινο η 9MET, με ανοιχτό γαλάζιο η 2GLN, με σκούρο γαλάζιο η 10ASN, με γκρι η 17ARG και με μωβ η 12PHE. Στην εικόνα

Παρατηρείται ότι η 13ILE βρίσκεται πιο κοντά στο εσωτερικό της διπλής έλικας καθώς φαίνεται να έχει τη μικρότερη απόσταση από τις βάσεις των νουκλεοτιδίων, μεταξύ των οποίων παρατηρείται αυξημένη ανύψωση.

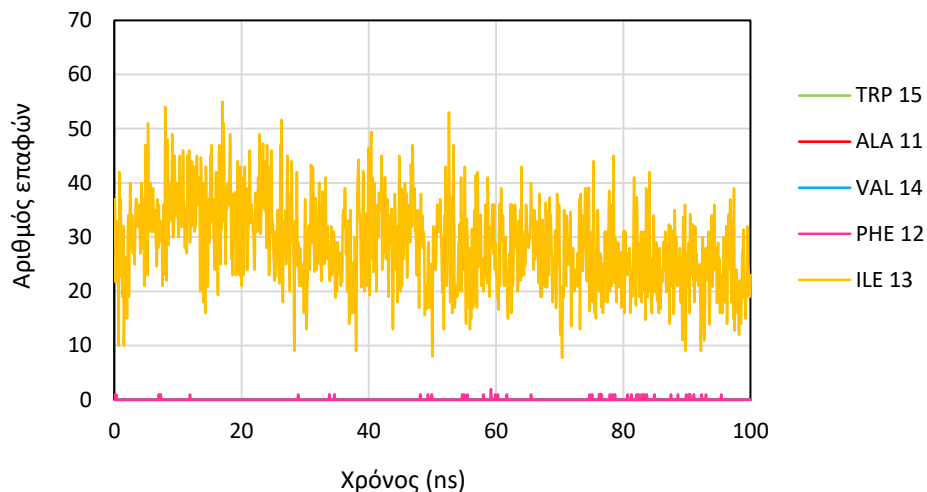
Υπολογίστηκαν, επίσης, οι ελάχιστες αποστάσεις και ο αριθμός επαφών εντός των 0.4 nm με τα νουκλεοτίδια A8 και A9 για τα γειτονικά αμινοξέα της 13ILE : 11ALA, 12PHE, 14VAL, 15TRP. Αυτά παρουσιάζονται στα Διαγράμματα 6.32 – 6.35.



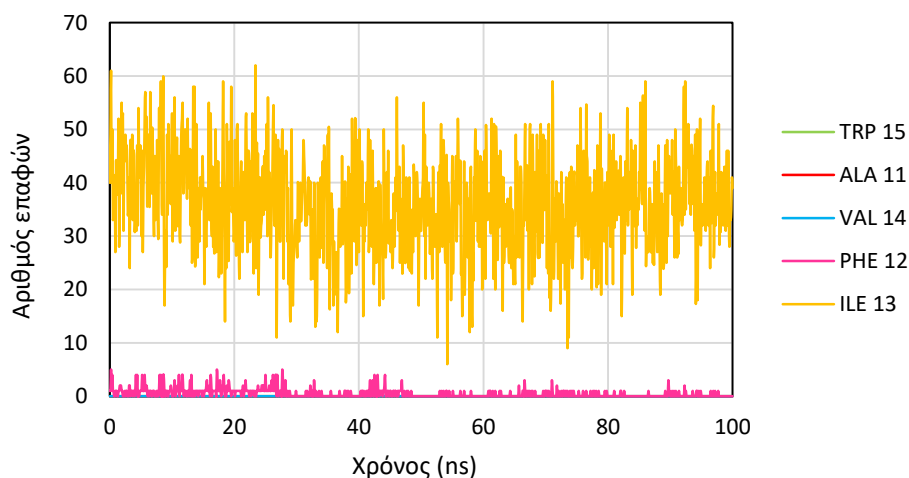
Διάγραμμα 6.32. Χρονική εξέλιξη των ελαχίστων αποστάσεων μεταξύ του νουκλεοτιδίου A8 και των γειτονικών αμινοξέων της 13ILE.



Διάγραμμα 6.33. Χρονική εξέλιξη των ελαχίστων αποστάσεων μεταξύ του νουκλεοτιδίου A9 και των γειτονικών αμινοξέων της 13ILE.



Διάγραμμα 6.34. Χρονική εξέλιξη του αριθμού επαφών εντός των 0.4 nm μεταξύ του νουκλεοτιδίου A8 και των γειτονικών αμινοξέων της 13ILE.



Διάγραμμα 6.35. Χρονική εξέλιξη του αριθμού επαφών εντός των 0.4 nm μεταξύ του νουκλεοτιδίου A9 και των γειτονικών αμινοξέων της 13ILE.

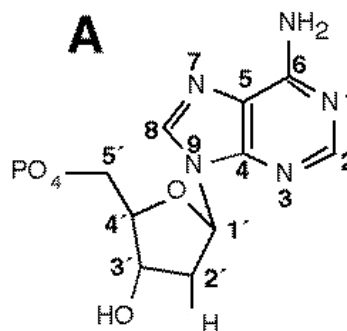
Παρατηρείται ότι τα γειτονικά αμινοξέα της 13ILE απέχουν αρκετά από τα νουκλεοτίδια A8 και A9. Επιπλέον, τα αμινοξέα αυτά δεν εμφανίζονται καθόλου ή εμφανίζονται ελάχιστα σε ελάχιστες αποστάσεις με τα νουκλεοτίδια που μας ενδιαφέρουν. Ωστόσο, ακόμα και αν τα αμινοξέα αυτά δεν προσεγγίζουν σε μικρές αποστάσεις τα νουκλεοτίδια A8, και A9 οι θέσεις τους και οι διαμορφώσεις τους επηρεάζουν τη θέση της 13 ILE.

6.7. ΜΕΡΙΚΗ ΠΑΡΕΜΒΟΛΗ ΤΗΣ ΙΣΟΛΕΥΚΙΝΗΣ ΤΗΣ SRY ΣΤΗ ΔΙΠΛΗ ΕΛΙΚΑ ΤΟΥ DNA

Στην ενότητα αυτή παρουσιάζεται η ανάλυση των ελαχίστων αποστάσεων και του αριθμού επαφών του αμινοξέος 131LE της SRY από το DNA και από τα νουκλεοτίδια A8 και A9. Για την ανάλυση αυτή είναι σκόπιμο να προηγηθεί η περιγραφή της δομής της ισολευκίνης και των νουκλεοτιδίων αδενίνης:

- **Δομή νουκλεοτιδίων αδενίνης**

Η δομή της δεοξυριβόζης και της βάσης αδενίνης περιγράφηκε στο κεφάλαιο 1.1. Ολόκληρο το νουκλεοτίδιο αδενίνης παρουσιάζεται στην Εικόνα 6.7. Τα άτομα ή οι ομάδες ατόμων που την απαρτίζουν συμβολίζονται με το χημικό τους σύμβολο, συνοδευόμενο από έναν αριθμό: τα άτομα της βάσης αριθμούνται από το 1 μέχρι το 9, ενώ τα άτομα της ζάχαρης από το 1' μέχρι το 5'. Τα άτομα που ενώνονται με άτομα της βάσης ή της ζάχαρης έχουν τον ίδιο αριθμό με το άτομο με το οποίο ενώνονται.

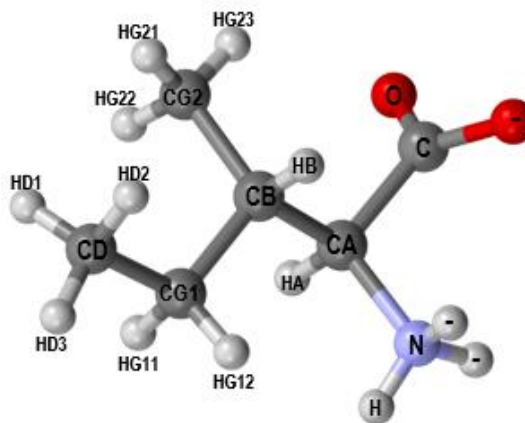


Εικόνα 6.7. Αρίθμηση ατόμων στο νουκλεοτίδιο της αδενίνης

- **Δομή ισολευκίνης**

Η δομή της ισολευκίνης παρουσιάζεται στην Εικόνα 6.8.

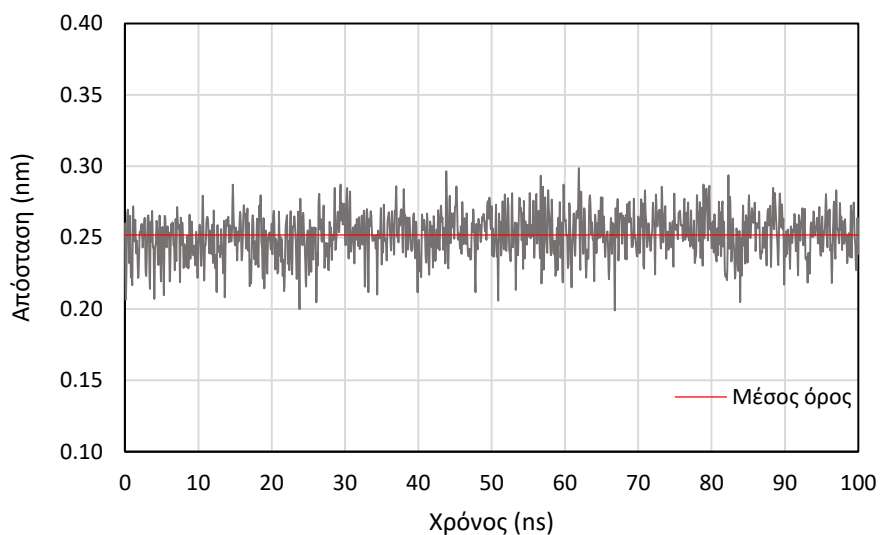
Με γκρι χρώμα απεικονίζονται τα άτομα άνθρακα, με κόκκινο τα άτομα οξυγόνου, με μπλε του αζώτου και με άσπρο τα υδρογόνα. Στην εικόνα παρουσιάζονται οι συμβολισμοί των ατόμων που χρησιμοποιούνται στην παρούσα εργασία. Με " - " έχουν επισημανθεί τα άτομα της ισολευκίνης που δεν είναι παρόντα στην SRY, λόγω της ένωσης του αμινοξέος με τα γειτονικά του κατά μήκος της πολυπεπτιδικής αλυσίδας.



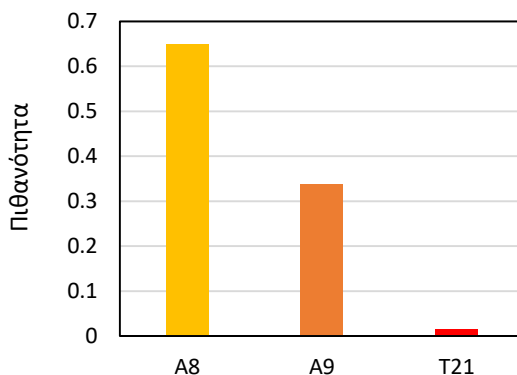
Εικόνα 6.8. Απεικόνιση του αμινοξέος της ισολευκίνης.

ΑΝΑΛΥΣΗ ΠΑΡΕΜΒΟΛΗΣ

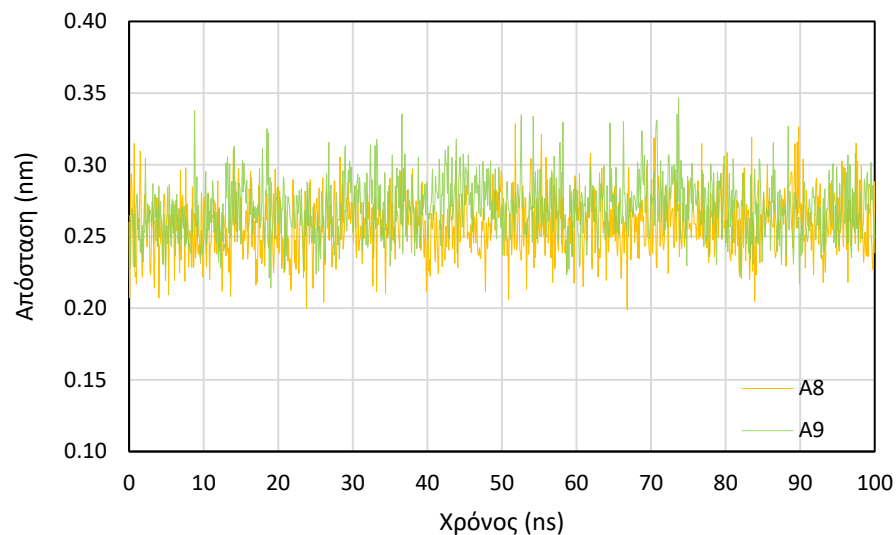
Υπολογίστηκαν οι ελάχιστες αποστάσεις της ισολευκίνης 13 από το DNA. Αυτές κυμαίνονται μεταξύ 0.20 - 0.30 nm με μέσο όρο 0.25 nm. Στις αποστάσεις αυτές εμφανίζεται η A8 με πιθανότητα 0.649, η A9 με πιθανότητα 0.337 και η T21 με πιθανότητα 0.014. Οι ελάχιστες αποστάσεις παρουσιάζονται στο Διάγραμμα 6.36 και η πιθανότητα εμφάνισης των αμινοξέων αδενίνης και θυμίνης σε αυτές παρουσιάζεται στο Διάγραμμα 6.37. Σημειώνεται ότι η T21 είναι ζεύγος της A8.



Διάγραμμα 6.36. Χρονική εξέλιξη των ελαχίστων αποστάσεων της 13ILE από το DNA.

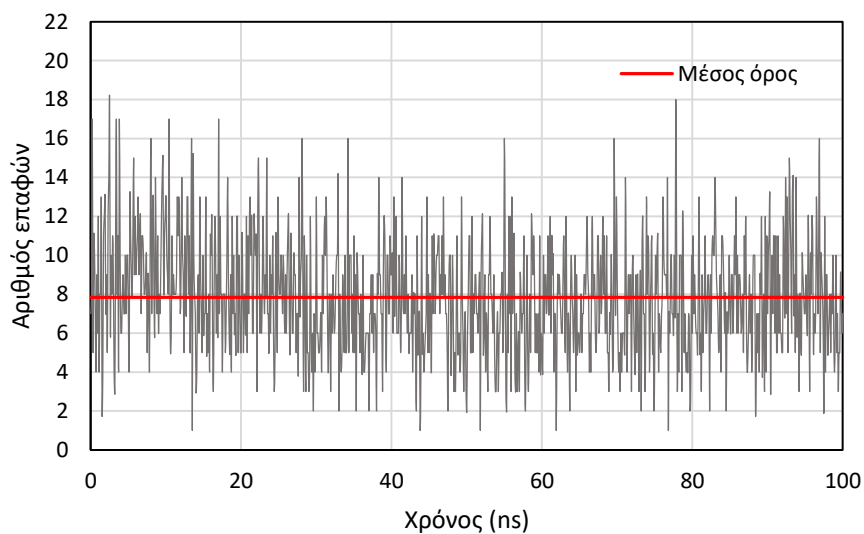


Διάγραμμα 6.37. Πιθανότητα εμφάνισης των νουκλεοτιδίων A8, A9 και T21 σε ελάχιστη απόσταση με την 13ILE.

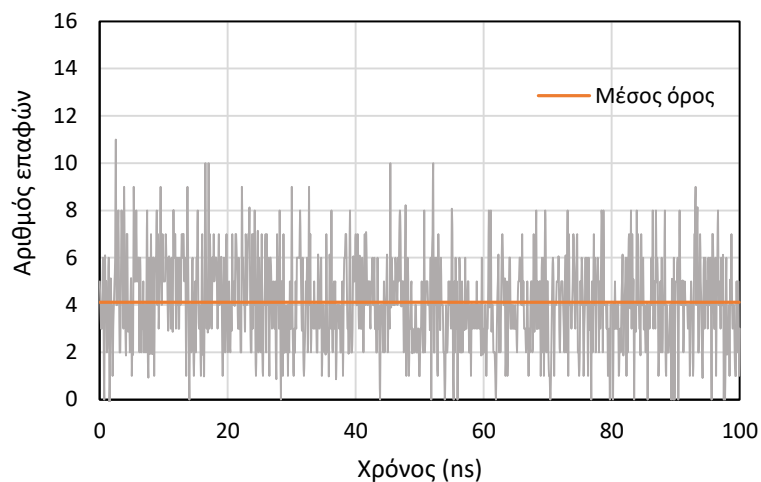


Διάγραμμα 6.38. Χρονική εξέλιξη των ελαχίστων αποστάσεων της 13ILE από τα νουκλεοτίδια A8 (κίτρινο χρώμα) και A9 (ανοιχτό πράσινο χρώμα).

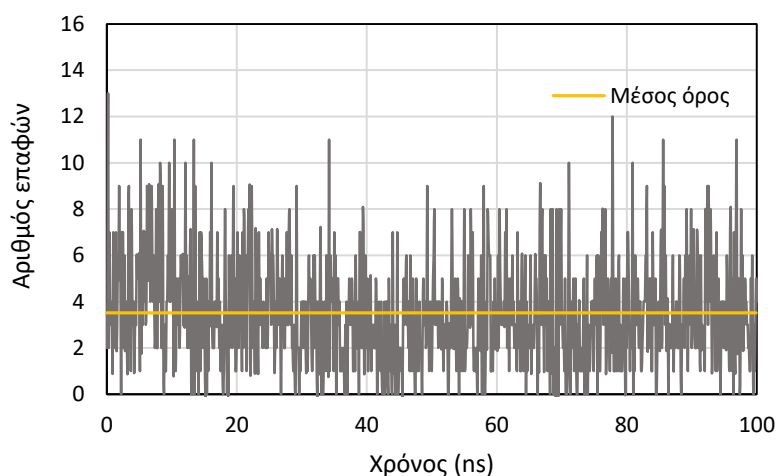
Υπολογίστηκαν, επίσης, οι αριθμοί επαφών 13ILE - DNA, οι οποίοι κυμαίνονται μεταξύ 1 - 18 με μέσο όρο 7.8, οι αριθμοί επαφών 13ILE - A8, οι οποίοι κυμαίνονται μεταξύ 0 - 11 με μέσο όρο 4.1 και οι αριθμοί επαφών 13ILE - A9, οι οποίοι κυμαίνονται μεταξύ 0 - 13 με μέσο όρο 3.5. Αυτοί παρουσιάζονται στα Διαγράμματα 6.39 - 6.41.



Διάγραμμα 6.39. Χρονική εξέλιξη του αριθμού επαφών μεταξύ της 13ILE και του DNA.

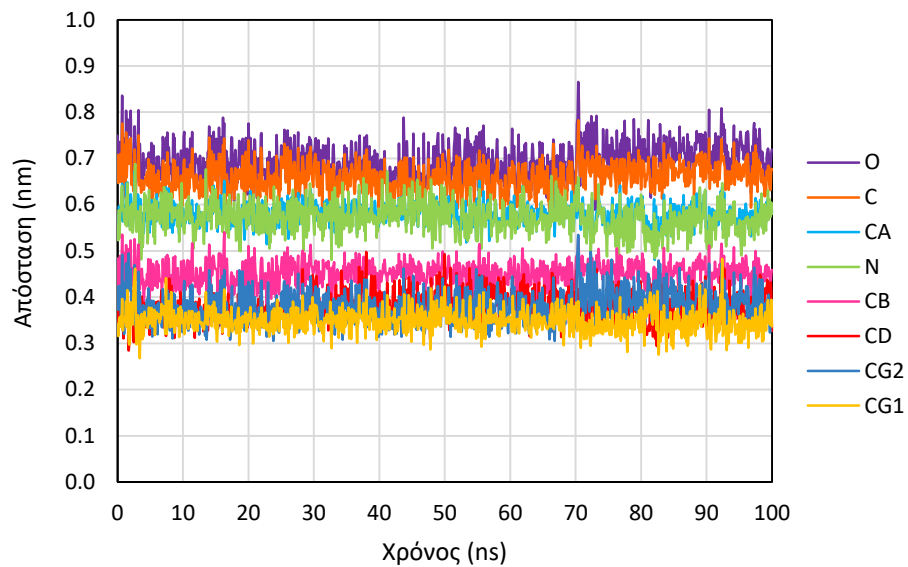


Διάγραμμα 6.40. Χρονική εξέλιξη του αριθμού επαφών μεταξύ της 13ILE και του νουκλεοτιδίου A8.

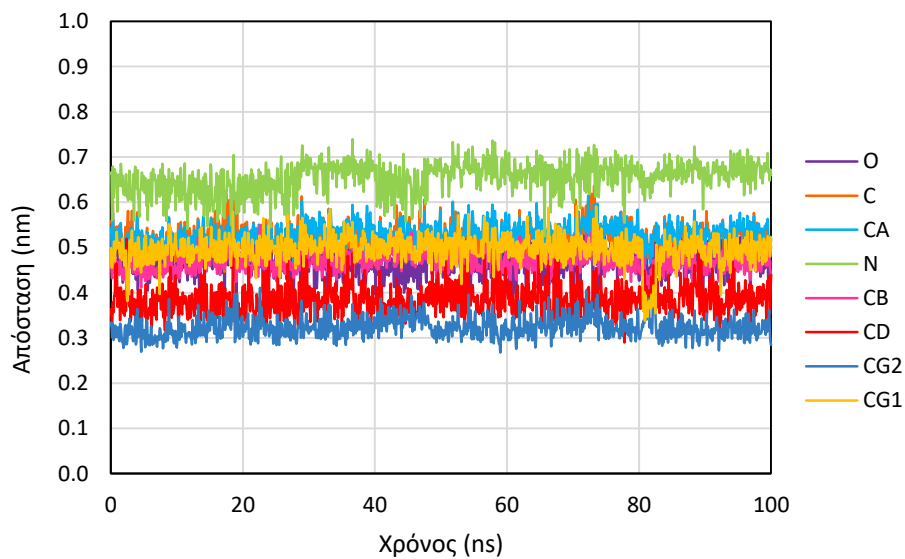


Διάγραμμα 6.41. Χρονική εξέλιξη του αριθμού επαφών μεταξύ της 13ILE και του νουκλεοτιδίου A9.

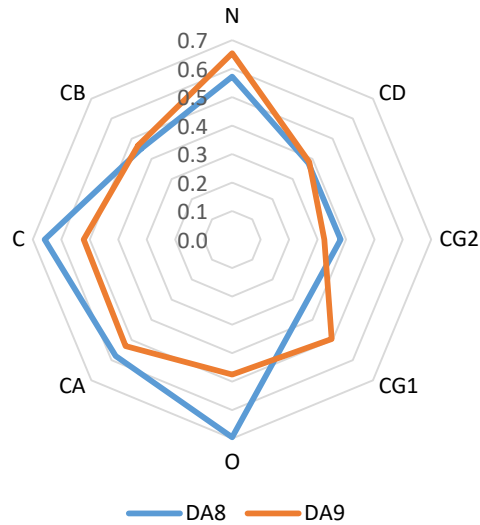
Υπολογίστηκαν, στη συνέχεια, οι ελάχιστες αποστάσεις της A8 και A9 από τα άτομα της 13ILE, εξαιρουμένων των ατόμων H. Αυτές παρουσιάζονται στα Διαγράμματα 6.42 και 6.43. Οι μέσοι όροι των ελαχίστων αποστάσεων για κάθε άτομο παρουσιάζονται στο αραχνοειδές γράφημα 6.40. Παρατηρείται ότι οι άνθρακες CG1, CG2 και CD βρίσκονται σε πιο κοντινές αποστάσεις με το νουκλεοτίδιο A8, ενώ το νουκλεοτίδιο A9 προσεγγίζει σε κοντινότερη απόσταση οι άνθρακες CG2 και CD.



Διάγραμμα 6.42. Χρονική εξέλιξη των ελαχίστων αποστάσεων των ατόμων της 13ILE από το νουκλεοτίδιο A8.



Διάγραμμα 6.43. Χρονική εξέλιξη των ελαχίστων αποστάσεων των ατόμων της 13ILE από το νουκλεοτίδιο A9.



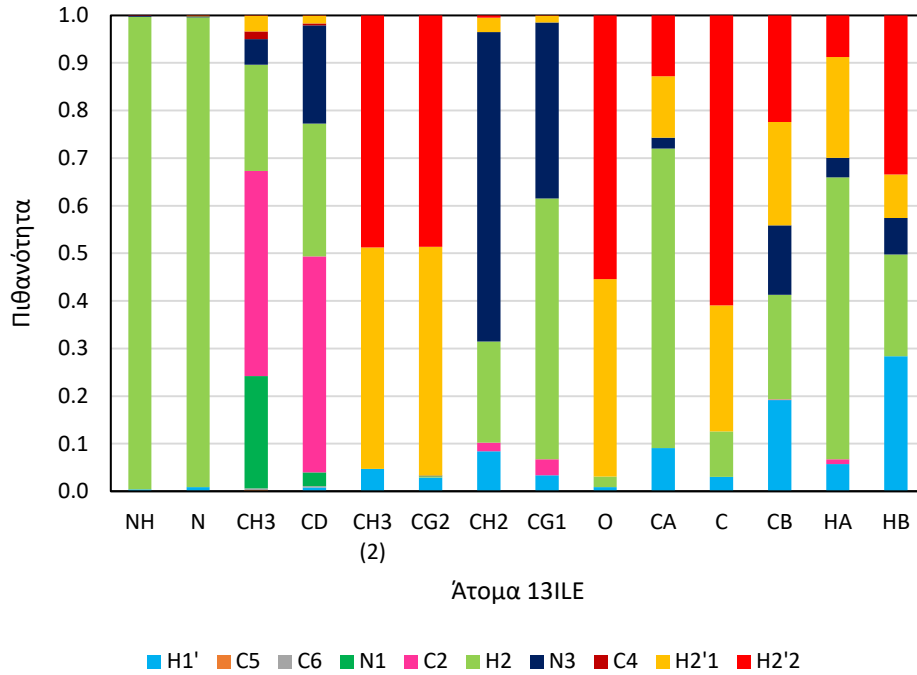
Διάγραμμα 6.44. Αραχνοειδές γράφημα στο οποίο παρουσιάζονται οι μέσοι όροι των ελαχίστων αποστάσεων των ατόμων 13ILE από τα νουκλεοτίδια A8 και A9.

Ακολούθως, εξετάστηκε η πιθανότητα εμφάνισης σε ελάχιστη απόσταση του κάθε ατόμου των νουκλεοτιδίων για κάθε άτομο της ισολευκίνης, η οποία παρουσιάζεται στα Διαγράμματα 6.45 και 6.46.

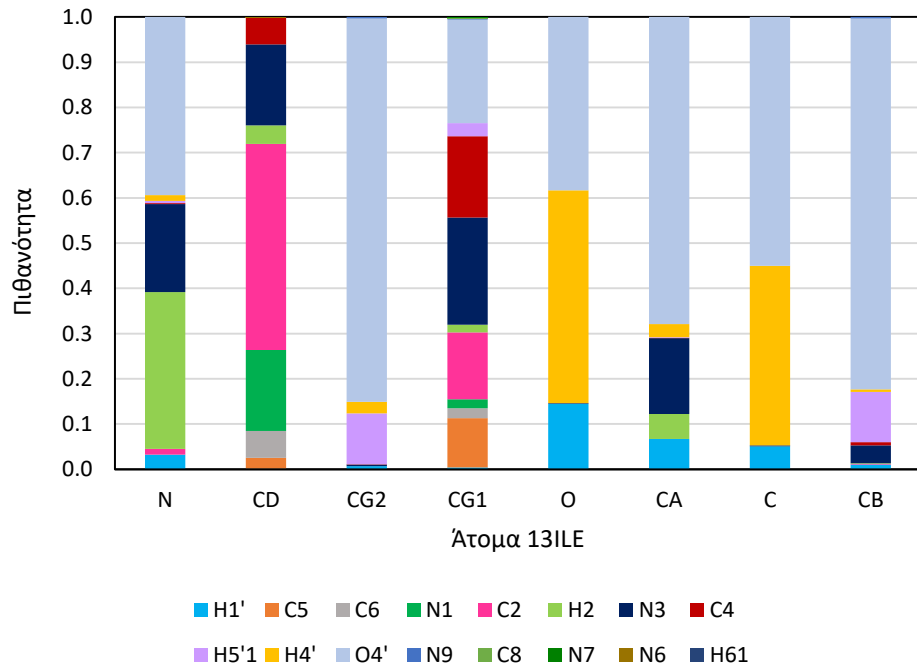
Στην περίπτωση του A8 παρατηρείται ότι ο CG1 βρίσκεται σε ελάχιστη απόσταση από το H2 της βάσης με πιθανότητα 0.55 και σε ελάχιστη απόσταση από το N3 της βάσης με πιθανότητα 0.37. Ο CG2, αν και βρίσκεται σε πολύ μικρή απόσταση από το A8, η απόσταση αυτή με τα άτομα H2'1 και H2'2 του σακχάρου του νουκλεοτιδίου και όχι της βάσης. Ο CD βρίσκεται σε ελάχιστη απόσταση από τον C2 της βάσης με πιθανότητα 0.45, από το H2 της βάσης με πιθανότητα 0.28 και από το N3 της βάσης με πιθανότητα 0.2.

Ο CG2 βρίσκεται στην μικρότερη ελάχιστη απόσταση από το A9, ωστόσο αυτό γίνεται κατά 85% με το O4' της ζάχαρης. Αντίθετα, οι ελάχιστες αποστάσεις του CD από την A9 αφορούν τον άνθρακα C2 της βάσης με πιθανότητα 0.46 και τα N3 και N1 της βάσης με ίση πιθανότητα, 0.18.

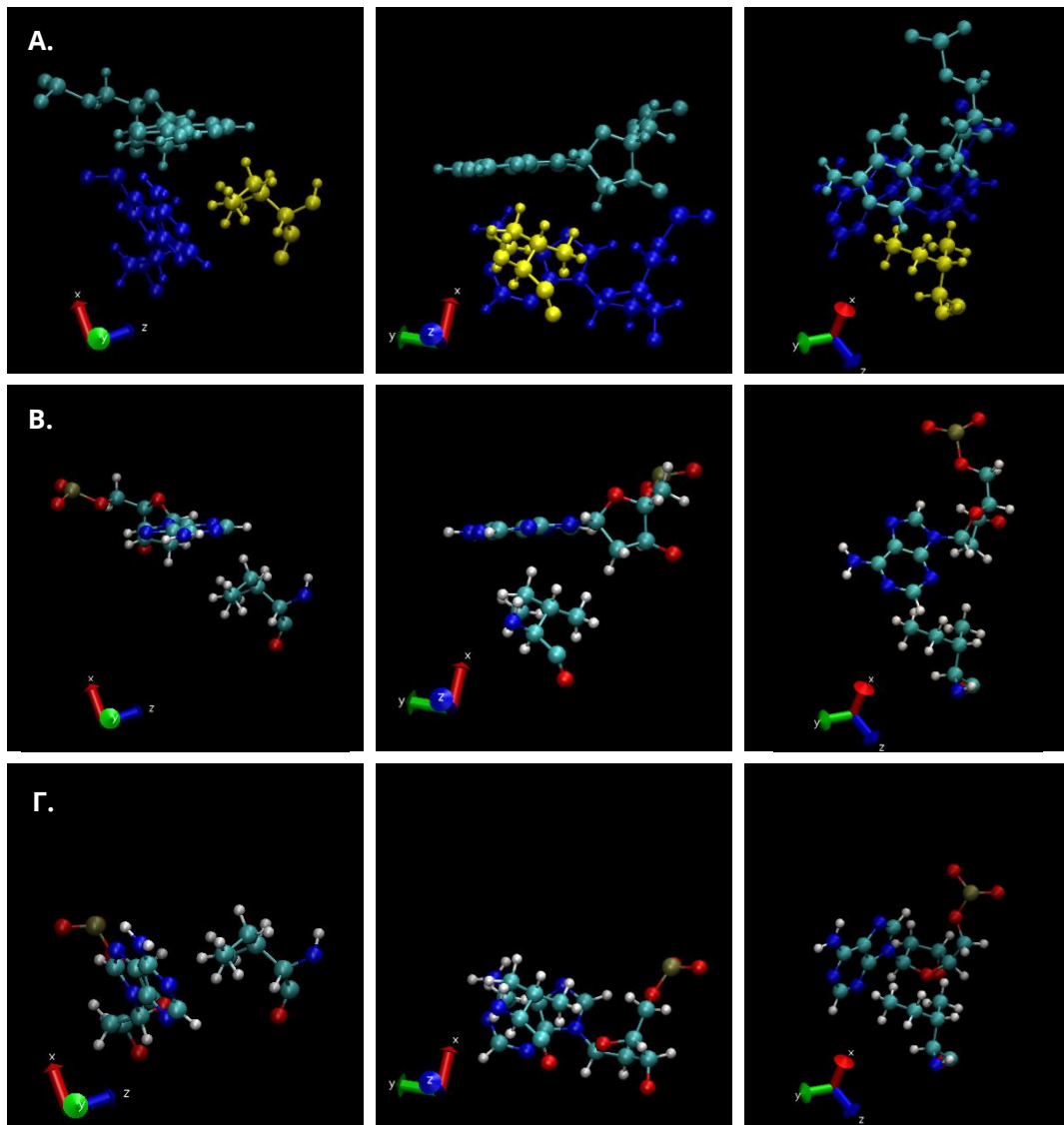
Με βάση τα ανωτέρω συνάγεται ότι η παρεμβολή της ισολευκίνης μεταξύ των βάσεων των δύο νουκλεοτιδίων προς το εσωτερικό της διπλής έλικας δεν πραγματοποιείται από ολόκληρο το αμινοξύ, αλλά μόνο από τα μόρια CD και CG1 της πλευρικής της αλυσίδας του. Επομένως, η παρεμβολή της ισολευκίνης στο DNA χαρακτηρίζεται ως μερική.



Διάγραμμα 6.45. Πιθανότητα εμφάνισης των ατόμων του νουκλεοτιδίου A8 σε ελάχιστη απόσταση με τα άτομα της 13ILE.



Διάγραμμα 6.46. Πιθανότητα εμφάνισης των ατόμων του νουκλεοτιδίου A9 σε ελάχιστη απόσταση με τα άτομα της 13ILE.



Εικόνα 6.9. Ατομιστική απεικόνιση των A8, A9 και 13ILE σε διαφορετικές όψεις: A. A8 (ανοιχτό μπλε), A9 (σκούρο μπλε), 13ILE (κίτρινο). B. 13ILE και A8. Γ. 13ILE και η A9. Τα άτομα στις Β., Γ. σειρές απεικονίζονται ως: C γαλάζιο, H άσπρο, O κόκκινο, N μπλε, P καφέ.

Η ατομιστική απεικόνιση της μερικής παρεμβολής της 13ILE μεταξύ των A8, A9 παρουσιάζεται στην Εικόνα 6.9. Διακρίνεται ότι, εξαιτίας της μερικής παρεμβολής της ισολευκίνης, τα επίπεδα των βάσεων που ανήκουν στα νουκλεοτιδία A8 και A9 παύουν να είναι παράλληλα μεταξύ τους, όπως είναι στην ελεύθερη διαμόρφωση του DNA. Η απόσταση μεταξύ των βάσεων αυξάνεται προς την πλευρά της ισολευκίνης (τα επίπεδα των βάσεων αποκλίνουν) και μειώνεται προς την αντίθετη πλευρά (τα επίπεδα των βάσεων συγκλίνουν). Οι τοπικές αυτές διαμορφώσεις αποτελούν την αιτία για την τοπική κάμψη του DNA που εμφανίζεται στην περιοχή αυτή, όπως μελετήθηκε στο Κεφάλαιο 6.4.

7. ΣΥΝΟΨΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα εργασία πραγματοποιήθηκαν προσομοιώσεις μοριακής δυναμικής δύο συστημάτων: (i) του μεταγραφικού παράγοντα SRY (πρωτεΐνη) προσδεσμένο στο γονίδιο SOX9 (DNA) και (ii) του ελεύθερου γονιδίου SOX9, εμβαπτισμένα στο νερό παρουσία ιόντων. Η προσομοίωση των δύο συστημάτων πραγματοποιήθηκε σε συνθήκες σταθερής πίεσης, $P = 1 \text{ bar}$ και θερμοκρασίας, $T = 310 \text{ K} = 37^\circ\text{C}$, οι οποίες επικρατούν στον ανθρώπινο οργανισμό. Οι τροχιές που προέκυψαν από τις δύο προσομοιώσεις αναλύθηκαν και εξήχθησαν πληροφορίες σχετικά με τη διαμόρφωση του DNA και του SRY κατά τη σύνδεσή τους.

Σε πρώτο στάδιο εξετάστηκε το κέλυφος ενυδάτωσης σε ακτίνα 0.35 nm για το DNA στην συμπλεγμένη και την ελεύθερή του μορφή, καθώς αυτό πολλές φορές επιδρά στη σύνδεσή του με άλλα μακρομόρια. Παρατηρήθηκε μείωση του κελύφους ενυδάτωσης κατά 12% (διαφορά 46 μορίων νερού) στην περίπτωση του δεσμευμένου DNA. Συνεπώς, η σύνδεση του DNA με τον SRY προκαλεί την αποβολή μορίων νερού από το κέλυφος ενυδάτωσης του DNA.

Στη συνέχεια, εξετάστηκαν τα γεωμετρικά χαρακτηριστικά της διπλής έλικας του συμπλεγμένου και του ελεύθερου DNA. Αρχικά, υπολογίστηκε το εύρος της μικρής και της μεγάλης αύλακας του δεσμευμένου και του ελεύθερου DNA. Παρουσιάστηκε διεύρυνση της μικρής αύλακας στο δεσμευμένο DNA στο ύψος της αλληλουχίας CAAAC, η οποία διπλασιάζεται στο ύψος των δύο τελευταίων A της αλληλουχίας αυτής. Ταυτόχρονα παρουσιάστηκε μείωση του εύρους της μεγάλης αύλακας στο ύψος της CAAA, έως και υποδιπλασιασμός αυτού στο ύψος των δύο πρώτων A της αλληλουχίας αυτής. Επιπλέον εξετάστηκε η κάμψη του άξονα της διπλής έλικας του DNA στην περιοχή CAAAC, και υπολογίστηκε η γωνία κάμψης ίση με 63.4° , τιμή η οποία είναι σε καλή συμφωνία με πειραματικές μελέτες.^{21, 53}

Συμπεραίνεται, ότι η σύνδεση της SRY στο DNA πραγματοποιείται στην αλληλουχία CAAA του DNA, στην πλευρά της μικρής αύλακας, σε συμφωνία με πειραματικές μελέτες³⁵ με διεύρυνση της μικρής και στένωση της μεγάλης αύλακας του μακρομορίου, προκαλώντας την κάμψη του DNA προς την πλευρά της μεγάλης αύλακας.

Σε επόμενο στάδιο, εξετάστηκαν οι αποστάσεις και οι επαφές των ατόμων του δεσμευμένου DNA από τα άτομα του SRY. Οι ελάχιστες αποστάσεις αυτών κυμαίνονται μεταξύ $0.15 \text{ nm} - 0.18 \text{ nm}$, και αντιστοιχούν σε άτομα του DNA με άτομα των πλευρικών αλυσίδων του SRY. Η κύρια αλυσίδα της SRY πλησιάζει το DNA με ελάχιστη απόσταση της τάξεως των 0.20 nm και η ραχοκοκαλιά είναι πιο απομακρυσμένη, με ελάχιστη απόσταση 0.27 nm . Επιπλέον, η μεγάλη πλειοψηφία των επαφών SRY – DNA εντός των 0.3 nm αφορούν σε αποστάσεις των

πλευρικών αποστάσεων, και λίγες επαφές της κύριας αλυσίδας (και ελάχιστες του σκελετού) με το DNA εμφανίζονται εντός των 0.3 nm.

Συμπεραίνεται, συνεπώς ότι η πρωτεΐνη SRY προσεγγίζει το DNA μέσω των πλευρικών της αλυσίδων. Το αποτέλεσμα αυτό είναι αναμενόμενο, καθώς είναι γνωστό ότι η σύνδεση των πρωτεϊνών πραγματοποιείται μέσω των πλευρικών αλυσίδων τους.

Υπολογίστηκαν, έπειτα, οι αποστάσεις και οι επαφές των πλευρικών αλυσίδων της SRY από τα νουκλεοτίδια του DNA. Οι πλευρικές αλυσίδες της SRY παρουσιάζουν μικρότερες ελάχιστες αποστάσεις με τα νουκλεοτίδια θυμίνης, οι οποίες όμως βρίσκονται πολύ κοντά με αυτές των νουκλεοτιδίων αδενίνης και κυτοσίνης, με μέσους όρους από 0.171 – 0.175 nm. Συγκρίσιμη διαφορά στις ελάχιστες αποστάσεις από τις πλευρικές αλυσίδες παρουσιάζουν μόνο τα νουκλεοτίδια γουανίνης (με μέσο όρο 0.186 nm). Τα νουκλεοτίδια θυμίνης εμφανίζονται με μεγαλύτερη συχνότητα σε ελάχιστες αποστάσεις από τις πλευρικές αλυσίδες, ακολουθούμενο από αυτά της γουανίνης και της αδενίνης. Τα νουκλεοτίδια αδενίνης εμφανίζουν αισθητά περισσότερες επαφές με το DNA σε σχέση με τα νουκλεοτίδια με βάσεις θυμίνης και κυτοσίνης. Τα νουκλεοτίδια με βάσεις γουανίνης εμφανίζουν και λιγότερες επαφές και μεγαλύτερες ελάχιστες αποστάσεις.

Συμπεραίνεται από τα παραπάνω ότι τα νουκλεοτίδια θυμίνης βρίσκονται σε πιο μικρές αποστάσεις από την SRY, αλλά τα νουκλεοτίδια αδενίνης έχουν περισσότερες κοντινές επαφές με αυτήν.

Στη συνέχεια υπολογίστηκε η πιθανότητα εύρεσης σε ελάχιστη απόσταση των αμινοξέων της SRY από το DNA. Οι ελάχιστες αποστάσεις μεταξύ SRY και DNA διαμορφώνονται κυρίως από τα αμινοξέα που βρίσκονται στη δεύτερη α-έλικα της SRY και δευτερευόντως από τα αμινοξέα που βρίσκονται στην πρώτη α-έλικα της SRY. Μεγάλο ποσοστό αυτών (37%) αντιστοιχεί στο αμινοξύ σερίνη (36SER), που βρίσκεται προς το κέντρο της δεύτερης α-έλικας, και το νουκλεοτίδιο θυμίνης T19, που είναι ζεύγος του νουκλεοτιδίου αδενίνης A10, στο ύψος των οποίων η μικρή αύλακα του DNA είναι ιδιαίτερα διευρυμένη.

Συμπεραίνεται ότι κατά την πρόσδεση του SRY, η δεύτερη α-έλικα του προσεγγίζει το DNA πολύ στενά και το αμινοξύ σερίνη (36SER) που κείται στο κέντρο αυτής πλησιάζει τον ένα κλώνο του DNA στην περιοχή που παρατηρείται η διεύρυνση της μικρής αύλακας του. Επιπλέον, η πρώτη α-έλικα πλησιάζει επίσης πολύ κοντά στο DNA.

Από τα γεωμετρικά χαρακτηριστικά του DNA υπολογίστηκε, επίσης, η ανύψωση των βάσεων κατά μήκος της διπλής έλικας. Ο υπολογισμός αυτός αποκάλυψε, για την περίπτωση του συμπλεγμένου DNA, αύξηση κατά 50% της ανύψωσης μεταξύ των νουκλεοτιδίων A8 και A9, τα οποία βρίσκονται στο κέντρο της αλληλουχίας CAAA, σε σχέση

με την ανύψωση που υπάρχει στο ελεύθερο DNA. Η αύξηση αυτή αποδίδεται στην παρεμβολή τμήματος της SRY μεταξύ των νουκλεοτιδίων A8 και A9.

Ακολούθως, πραγματοποιήθηκε μελέτη της περιοχής του DNA μεταξύ των νουκλεοτιδίων A8 και A9, όπου παρατηρήθηκε αυξημένη ανύψωση βάσεων. Υπολογίστηκαν, αρχικά, οι ελάχιστες αποστάσεις της SRY από τα A8, A9, οι οποίες έχουν μέσο όρο 0.190 nm και 0.194 nm αντίστοιχα και είναι αυξημένες κατά 0.02 nm από τις ελάχιστες αποστάσεις SRY – DNA. Υπολογίστηκε ο αριθμός επαφών εντός των 0.3 nm για A8-SRY και A9-SRY. Εντοπίστηκαν τα αμινοξέα τα οποία συμμετέχουν σε αυτές τις ελάχιστες αποστάσεις και εμφανίζονται σε ελάχιστες επαφές. Τα αμινοξέα αυτά εξετάστηκαν ως προς την θέση τους σε σχέση με τα νουκλεοτίδια A8 και A9 και εκτιμήθηκε ότι η ισολευκίνη 131LE είχε εισχωρήσει αρκετά μέσα στην έλικα, πλησιάζοντας τις βάσεις των παραπάνω εξεταζόμενων νουκλεοτιδίων.

Συμπεραίνεται ότι το αμινοξύ της SRY που παρεμβάλλεται μεταξύ βάσεων των νουκλεοτιδίων A8 και A9 είναι η ισολευκίνη 131LE, η οποία κείται στην πρώτη α-έλικα του SRY.

Μελετήθηκε η θέση των γειτονικών αμινοξέων της ισολευκίνης 131LE σε σχέση με τα νουκλεοτίδια A8 και A9. Υπολογίστηκαν οι ελάχιστες αποστάσεις και ο αριθμός επαφών εντός των 0.4 nm για τα δύο αμινοξέα που προηγούνται και τα δύο αμινοξέα που έπονται της 131LE στην αλληλουχία της πολυπεπτιδικής αλυσίδας της SRY. Οι ελάχιστες αποστάσεις αυτών από τα A8 και A9 προέκυψαν πολύ μεγαλύτερες από αυτές της 131LE και οι επαφές τους με τα A8 και A9 είναι ελάχιστες έως μηδενικές.

Συμπεραίνεται ότι τα γειτονικά αμινοξέα της 131LE δεν προσεγγίζουν σε μικρές αποστάσεις τα νουκλεοτίδια A8 και A9. Ωστόσο, καθώς τα αμινοξέα αυτά ανήκουν στην ίδια πολυπεπτιδική αλυσίδα του SRY, τόσο οι διαμορφώσεις όσο και οι θέσεις τους επηρεάζουν τη θέση και τη διαμόρφωση της 131LE σε σχέση με το DNA.

Στο στάδιο που ακολούθησε, μελετήθηκε η παρεμβολή της 131LE στα νουκλεοτίδια A8 και A9. Εξετάστηκε η τοποθέτηση της 131LE στο χώρο, σε σχέση με το DNA και τα νουκλεοτίδια A8 και A9. Οι ελάχιστες αποστάσεις της 131LE – DNA έχουν διακύμανση 0.20 – 0.30 nm και μέσο όρο 0.25, αριθμοί μεγαλύτεροι από τους αντίστοιχους για την SRY – DNA. Αυτές οι αποστάσεις αντιστοιχούν αποκλειστικά στα ζεύγη νουκλεοτιδίων A8 – T21 και A9 – T20, και αφορούν κατά κύριο λόγο την A8 με πιθανότητα εμφάνισης 0.65. Επιπλέον, ο μέσος όρος ελαχίστων αποστάσεων 131LE – A8 εμφανίζεται λίγο μικρότερος από τον μέσο όρο ελαχίστων αποστάσεων 131LE – A9 (0.26 nm και 0.27 nm αντίστοιχα). Ο αριθμός επαφών της 131LE – A8 προκύπτει επίσης ελαφρώς μεγαλύτερος από τον αριθμό επαφών της 131LE – A9.

Συμπεραίνεται ότι η μικρότερη απόσταση της 13ILE εντοπίζεται από το νουκλεοτίδιο A8, και το νουκλεοτίδιο αυτό προσεγγίζει λίγο πιο στενά την 13ILE σε σχέση με το νουκλεοτίδιο A9.

Για την πιο λεπτομερή μελέτη της παρεμβολής της 13ILE έγινε διερεύνηση των ελαχίστων αποστάσεων των ατόμων της 13ILE (εξαιρουμένων των ατόμων H) από τα νουκλεοτίδια A8 και A9. Προέκυψε ότι οι μικρότερες ελάχιστες αποστάσεις της 13ILE από το νουκλεοτίδιο A8 αντιστοιχούν στα άτομα CG1, CG2 και CD, της πλευρικής αλυσίδας της 13ILE. Οι μικρότερες ελάχιστες αποστάσεις της 13ILE από το νουκλεοτίδιο A8 αντιστοιχούν στα άτομα CG2 και CD της 13ILE. Ακολούθως, εξετάστηκε με ποια άτομα των νουκλεοτιδίων A8 και A9 έρχονται τα άτομα της 13ILE σε ελάχιστη απόσταση. Προέκυψε ότι ο CG1 βρίσκεται σε ελάχιστη απόσταση από τα άτομα H2 και N3 της βάσης του A8. Ο CD βρίσκεται σε ελάχιστη απόσταση από τα άτομα H2 και N3 της βάσης του A8 και από τα άτομα C2, N3 και N1 της βάσης A9. Ο CG2 βρίσκεται σε μικρή απόσταση από τα A8 και A9, ωστόσο η απόσταση αυτή είναι από τα άτομα της ζάχαρης των νουκλεοτιδίων και όχι των βάσεων.

Συμπεραίνεται ότι η παρεμβολή της 13ILE μεταξύ των βάσεων των νουκλεοτιδίων A8 και A9 πραγματοποιείται μέσω των ατόμων CD και CG1 της πλευρικής αλυσίδας της. Η παρεμβολή αυτή χαρακτηρίζεται ως μερική, καθώς δεν παρεμβάλλεται ολόκληρο το αμινοξύ της ισολευκίνης, αλλά μόνο ο ένας κλάδος της πλευρικής του αλυσίδας.

Επιπλέον, παρατηρήθηκε ότι κατά τη μερική παρεμβολή της ισολευκίνης, τα επίπεδα των βάσεων των νουκλεοτιδίων αποκλίνουν προς το μέρος της ισολευκίνης και συγκλίνουν προς την αντίθετη πλευρά. Η διαμόρφωση αυτή αποτελεί την αιτία για την κάμψη του μορίου του DNA.

Μελλοντικοί στόχοι

Σύμφωνα με πειραματικές μελέτες, η σύνδεση του μεταγραφικού παράγοντα SRY στο γονίδιο *SOX9* επηρεάζεται από μεταλλάξεις που μπορούν να συμβούν είτε στην πρωτεΐνη είτε στο DNA. Δεδομένου αυτού, θα ήταν χρήσιμο να μελετηθεί η πρόσδεση SRY-*SOX9*, κατασκευάζοντας είτε μια μεταλλαγμένη μορφή του μεταγραφικού παράγοντα SRY στην οποία να απουσιάζουν τα αμινοξέα ισολευκίνη 131LE, ή/και σερίνη 36SER, ή μια μεταλλαγμένη μορφή του γονιδίου *SOX9* το οποίο να περιέχει μια διαφορετική αλληλουχία βάσεων, αντί για την κρίσιμη αλληλουχία CAAAC στην οποία πραγματοποιείται η πρόσδεση του SRY.

Επιπλέον, η σύνδεση στο γονίδιο, μέσω της *δομικής περιοχής σύνδεσης με το DNA* του μεταγραφικού παράγοντα, η οποία μελετήθηκε στην παρούσα διπλωματική εργασία, αποτελεί το πρώτο βήμα για την εκκίνηση της μεταγραφής του γονιδίου. Σε επόμενο στάδιο ο μεταγραφικός παράγοντας συνδέεται με τους συνενεργοποιητές, μέσω της *δομικής περιοχής ενεργοποίησης*, και, με τη βοήθεια και άλλων πρωτεϊνών, συγκροτείται το σύμπλεγμα της έναρξης της μεταγραφής.

Ως μελλοντικός στόχος τίθεται η μελέτη της σύνδεσης της δομικής περιοχής ενεργοποίησης του μεταγραφικού παράγοντα, με την αντίστοιχη περιοχή ενός άλλου μεταγραφικού παράγοντα ή συνενεργοποιητή. Είναι γνωστό ότι η SRY συνδέεται με τον στεροειδογενετικό παράγοντα 1, SF1, για την ενεργοποίηση της μεταγραφής του *SOX9*, και, επομένως, η ατομιστική μελέτη αυτής της σύνδεσης, θα προσφέρει πολύτιμη πληροφορία για τον μηχανισμό δράσης των μεταγραφικών παραγόντων.

Σε επόμενο στάδιο, και εφόσον η συγκρότηση του συμπλέγματος έναρξης της μεταγραφής προϋποθέτει την συγκέντρωση πλήθους μεταγραφικών ρυθμιστών και συνενεργοποιητών, προτείνεται η προσομοίωση του συστήματος των ανωτέρω βιομορίων. Η προσομοίωση αυτή, θα είναι εφικτή με αδροποίηση του συστήματος, λόγω του μεγάλου αριθμού των εμπλεκόμενων μορίων.

8. ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, Fifth Edition, W H Freeman, 2002.
2. McMurry, J. *Οργανική Χημεία*. (Μετάφραση στα ελληνικά). Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο 2015.
3. Alberts, B.; Bray, D.; Hopkin, K.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Essential Cell Biology*, 3rd Edition. Garland Science, 2010.
4. Sabari, B. R.; Dall'Agnese, A.; Boija, A.; Klein, I. A.; Coffey, E. L.; Shrinivas, K.; Abraham, B. J.; Hannett, N. M.; Zamudio, A. V.; Manteiga, J. C.; Li, C. H.; Guo, Y. E.; Day, D. S.; Schuijers, J.; Vasile, E.; Malik, S.; Hnisz, D.; Lee, T. I.; Cisse, I. I.; Roeder, R. G.; Sharp, P. A.; Chakraborty, A. K.; Young, R. A. Coactivator condensation at super-enhancers links phase separation and gene control. *Science (New York, N.Y.)*. **2018**, 361(6400), eaar3958.
5. Lambert, S. A.; Jolma, A.; Campitelli, L. F.; Das, P. K.; Yin, Y.; Albu, M.; Chen, X.; Taipale, J.; Hughes, T. R.; Weirauch, M. T. The Human Transcription Factors. *Cell*. **2018**, 172(4), 650–665.
6. Shrinivas, K.; Sabari, B. R.; Coffey, E. L.; Klein, I. A.; Boija, A.; Zamudio, A. V.; Schuijers, J.; Hannett, N. M.; Sharp, P. A.; Young, R. A.; Chakraborty, A. K. Enhancer Features that Drive Formation of Transcriptional Condensates. *Mol Cell*. **2019**, 75(3), 549–561.e7
7. Alberti, S.; Dormann, D. Liquid-Liquid Phase Separation in Disease. *Annu. Rev. Genet.* **2019**, 53, 171–194.
8. Viny, A. D., Levine, R. L. Drug modulation by nuclear condensates. *Science*. **2020**, 368, 6497.
9. Klein, I. A.; Boija, A.; Afeyan, L. K.; Hawken, S. W.; Fan, M., Dall'Agnese, A., Oksuz, O.; Henninger, J. E.; Shrinivas, K.; Sabari, B. R.; Sagi, I.; Clark, V. E.; Platt, J. M.; Kar, M.; McCall, P. M.; Zamudio, A. V.; Manteiga, J. C.; Coffey, E. L.; Li, C. H.; Hannett, N. M.; Guo, Y. E.; Decker, T. M.; Lee, T. I.; Zhang, T.; Weng, J. K.; Taatjes, D. J.; Chakraborty, A.; Sharp, P. A.; Chang, Y. T.; Hyman, A. A.; Gray, N. S.; Young, R. A. Partitioning of cancer therapeutics in nuclear condensates. *Science*. **2020**, 368, 6497, 1386–1392.
10. Boija, A.; Klein, I. A.; Sabari, B. R.; Dall'Agnese, A.; Coffey, EL.; Zamudio, A. V.; Li, C. H.; Shrinivas, K.; Manteiga, J. C.; Hannett, N. M.; Abraham, B. J.; Afeyan, L. K.; Guo, Y. E.; Rimel, J. K., Fant , C. B.; Schuijers, J.; Lee, T. I.; Taatjes, D. J.; Young, R. A. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*. **2018**, 175(7), 1842-1855.e16.
11. Fulton, D. L.; Sundararajan, S.; Badis, G.; Hughes, T. R.; Wasserman, W. W.; Roach, J. C.; Sladek, R. TFCat: The curated catalog of mouse and human transcription factors. *Genome Biol.* **2009**, 10, R29.
12. The UniProt Consortium. UniProt: the universal protein knowledge base in 2021. *Nucleic Acids Res.* 49:D1, 2021.
13. She, Z. Y.; Yang, W. X.; Sry and SoxE genes: How they participate in mammalian sex determination and gonadal development? *Semi Cell Dev. Biol.* **2017**, 63, 13-22.

14. Croft, B.; Ohnesorg, T.; Hewitt, J. *et al.* Human sex reversal is caused by duplication or deletion of core enhancers upstream of SOX9. *Nat. Commun.* **2018**, *9*, 5319.
15. Keith, L.; Moore, Persaud, T. V. N. *The Developing Human / Clinically Oriented Embryology*, 8th Edition. *Saunders Elsevier*. **2008**
16. Berta, P.; Hawkins, J. R.; Sinclair, A. H.; Taylor, A.; Griffiths, B. L.; Goodfellow, P. N.; Fellous, M. Genetic evidence equating SRY and the testis-determining factor. *Nat.* **1990**, *348*, 448-450.
17. Kashimada, K.; Koopman, P. Sry: the master switch in mammalian sex determination. *Development*, **2010**, *137*, 3921-3930.
18. Harley, V. R.; Clarkson, M. J.; Argentaro, A. The molecular action and regulation of the testis-determining factors, SRY (sex-determining region on the Y chromosome) and SOX9 [SRY-related high-mobility group (HMG) box 9]. *Endocr. Rev.* **2003**, *24*, 466-487
19. Li, Y.; Zheng, M.; Lau, Y. F. The sex-determining factors SRY and SOX9 regulate similar target genes and promote testis cord formation during testicular differentiation. *Cell Rep.* **2014**, *8*, 723-33.
20. Schafer, A. J.; Goodfellow, P. N. Sex determination in humans. *BioEssays*. **1996**, *18*, 955-963.
21. MalaCards – the human disease database. www.malacards.org. Rappaport, N.; Twik, M.; Plaschkes, I.; Nudel, R.; Iny, S. T.; Levitt, J.; Gershoni, M.; Morrey, CP.; Safran, M.; Lancet D. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search, *Nucleic Acids Research*(2016). https://www.malacards.org/card/46xx_sex_reversal_1_2
22. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*. Clarendon Press; Oxford, 1991.
23. Hiemenz, P.C. *Principles of Colloid and Surface Chemistry*, New York, 1977.
24. McQuarrie, D.A. *Statistical Mechanics*, California, 2000.
25. Birkhoff, G. D. Proof of the Ergodic Theorem. *PNAS.* **1931**, *17*, 656
26. Huang, K. *Statistical Mechanics*. Wiley, New York, 1963.
27. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *Chem. Phys.* **1984**, *81*, 3684.
28. GROMACS Reference Manual Version 2016.1. The GROMACS development teams at the Royal Institute of Technology and Uppsala University, Sweden.
29. RCSB Protein Data Bank. 1J46. <https://www.rcsb.org/structure/1J46>
30. Murphy, E. C.; Zhurkin, V. B.; Louis, J. M.; Cornilescu, G.; Clore G. M. Structural Basis for SRY-dependent 46-X,Y Sex Reversal: Modulation of DNA Bending by a Naturally Occurring Point Mutation, *J. Mol. Biol.* **2001**, *312*, 3, 481-499.
31. Sadler, T. W. *Langman's Βασική Ιατρική Εμβρυολογία, (Μετάφραση)*, Ιατρικές Εκδόσεις Π. Χ. Πασχαλίδης, 2006
32. Harley, V.R.; Lovell-Bagde, R.; Goodfellow, P.N. Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res.* **1994**, *22(8)*,1500-1501. doi: 10.1093/nar/22.8.1500

33. Li, S.; Olson, W. K.; Lu, X.-J. Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Res.* 2019, *47*, W26–W34
34. (a) Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19-25.
(b) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H.J.C. GROMACS: Fast, Flexible and Free. *J. Comp. Chem.* **2005**, *26*, 1701-1718.
(c) Berendsen, H.J.C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **1995**, *91*, 43-56.
35. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
36. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dorr, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the AMBER ff99SB protein force field. *PROTEINS: Struct. Funct. Gen.* **2010**, *78*, 1950-1958.
37. Balaceanu, A.; Pérez, A.; Dans, P. D.; Orozco, M. Allostery and signal transfer in DNA. *Nucleic Acids Res.* **2018**, *46*, 7554-7565.
38. Marklund, E. G.; Mahmutovic, A.; Berg, O. G.; Hammar, P.; van der Spoel, D.; Fange, D.; Elf, J. Transcription-factor binding and sliding on DNA studied using micro- and macroscopic models. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19796-19801.
39. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935
40. Kührová, P.; Otyepka, M.; Šponer, J.; Banáš, P. Are Waters around RNA More than Just a Solvent? – An Insight from Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 401-411.
41. Galindo-Murillo, R.; Bergonzo, C.; Cheatham III, T. E. Molecular Modeling of Nucleic Acid Structure: Setup and Analysis. *Curr. Protoc. Nucleic Acid Chem.* **2014**, *56*, 7.10.1-7.10.21
42. Pantatosaki, E.; Papadopoulos, G. K. Binding Dynamics of siRNA with selected Lipopeptides: A Computer-Aided Study of the Effect of Lipopeptides' Functional Groups and Stereoisomerism. *J. Chem. Theory Comput.* **2020**, *16*, 3842-3855.
43. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
44. (a) Hess, B., Bekker, H., Berendsen, H. J. C., Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **1997**, *18*, 1463-1472.
(b) Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116-122.
45. (a) Darden, T., York, D., Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092

- (b) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., Pedersen, L. G. A smooth particle mesh ewald potential. *J. Chem. Phys.* **1995**, *103*, 8577-8592.
46. Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Molec. Graphics* **1996**, *14*, 33-38;
47. Lavery, R.; Moakher M, Maddocks JH, Petkeviciute D, Zakrzewska K. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* **2014**, *37*, 5917-29.
48. Eric, W. L.; Mark, O.; McLinden, Daniel G. F.; «Thermophysical Properties of Fluid Systems" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899.
49. Arunan, E.; Desiraju, G. R.; Klein, R. A.; Sadlej, S.; Scheiner, S.; Alkorta, I.; Clary, D. C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P.; Kjaergaard, H. G.; Legon, A. C.; Mennucci, B.; Nesbitt, D. J. Definition of the hydrogen bond (IUPAC Recommendations 2011). *Pure Appl Chem.* **2011**, *83* (8), 1637-1641
50. Flory, P. J.; Jackson-C, J. C.; Wood, J. *Statistical Mechanics of chain molecules*. John Wiley & Sons, INC., 1969.
51. Werner, M. H.; Huth, J. R.; Gronenborn, A. M.; Clore, G. M. Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell.* **1995**, *81*(5), 705-714.