



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΙΑΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ & ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

Διπλωματική Εργασία με θέμα

Υπολογισμός του προφίλ σκέδασης ακτίνων-Χ μικρής γωνίας (SAXS) της λυσοζύμης με την χρήση του προγράμματος CRY SOL- Προσομοιώσεις Μοριακής Δυναμικής.

Calculation of Small-Angle X-Ray Scattering profil of lysozyme by using the program CRY SOL- Molecule Dynamics Simulations.

Πατεπιάς Ηλίας
Λάρισα, Φεβρουάριος 2018

Science means constantly walking a tightrope between blind faith and curiosity; between expertise and creativity; between bias and openness; between experience and epiphany; between ambition and passion; and between arrogance and conviction – in short, between an old today and a new tomorrow.

- Heinrich Rohrer -

Co-inventor of Scanning Tunneling Microscope (STM)

Nobel Prize in Physics (1986)

Στοιχεία διπλωματικής εργασίας

Η παρούσα εργασία εκπονήθηκε στην ομάδα Βιοφυσικής, του εργαστηρίου Δομικής και Λειτουργικής Βιοχημείας του τμήματος Βιοχημείας & Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας, κατά το χρονικό διάστημα Μάρτιος 2017 – Φεβρουάριος 2018, υπό την επίβλεψη του Επίκουρου Καθηγητή κ. Γεωργίου Παπαδόπουλου.

Μέλη τριμελούς συμβουλευτικής επιτροπής

- Παπαδόπουλος Γεώργιος, Επίκουρος Καθηγητής (Επιβλέπων)
- Λεωνίδας Δημήτριος, Καθηγητής (Μέλος)
- Σκαμνάκη Βασιλική, Λέκτορας (Μέλος)

Ευχαριστίες

Ξεκινώντας, θα ήθελα να ευχαριστήσω από καρδιάς τον επιβλέποντα καθηγητή μου κ. Παπαδόπουλο για την άριστη συνεργασία που είχαμε, τις πολύτιμες συμβουλές του και την έμπρακτη στήριξή του μέχρι να ολοκληρωθεί η παρούσα εργασία. Κυρίως όμως θα ήθελα να τον ευχαριστήσω για την αστείρευτη υπομονή του, την αισιοδοξία και το χιούμορ με τα οποία αντιμετωπίσαμε μαζί της δυσκολίες που παρουσιάστηκαν, αλλά και την κατανόηση που έδειξε στις δικές μου αδυναμίες. Ένα μεγάλο ευχαριστώ στην οικογένεια μου και τους φίλους μου, που πάντα ήταν δίπλα μου όλα αυτά τα χρόνια, με στήριξαν και με βοηθούσαν, και τέλος ένα ξεχωριστό ευχαριστώ στα παιδιά της ομάδας, Μεχμέτογλου Φερίδέ, Τσομπάνη Θεοδώρα, και την Μάκιου Ανθή, που εκτός του ότι είμαστε καλοί φίλοι αποδειχθήκαμε και εξαιρετικοί συνεργάτες.

ΠΕΡΙΕΧΟΜΕΝΑ

Σκοπός	7
Περίληψη.....	7
Abstract.....	7
ΕΙΣΑΓΩΓΗ	9
<i>A) ΛΙΓΑ ΛΟΓΙΑ ΓΙΑ ΤΙΣ ΠΡΩΤΕΪΝΕΣ</i>	<i>9</i>
<i>1. ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΚΑΙ ΙΔΙΟΤΗΤΕΣ ΤΩΝ ΠΡΩΤΕΪΝΩΝ</i>	<i>9</i>
<i>2. ΠΡΩΤΟΤΑΓΗΣ – ΔΕΥΤΕΡΟΤΑΓΗΣ – ΤΡΙΤΟΤΑΓΗΣ - ΤΕΤΑΡΤΟΤΑΓΗΣ ΔΟΜΗ ΤΩΝ ΠΡΩΤΕΪΝΩΝ</i>	<i>11</i>
<i>3. Η ΛΥΣΟΖΥΜΗ</i>	<i>12</i>
<i>4. Η ΜΕΤΟΥΣΙΩΣΗ ΤΩΝ ΠΡΩΤΕΪΝΩΝ</i>	<i>15</i>
<i>5. ΑΝΩΜΑΛΗ ΑΝΑΔΙΠΛΩΣΗ ΠΡΩΤΕΪΝΩΝ ΚΑΙ ΝΟΣΗΜΑΤΑ</i>	<i>16</i>
<i>B) ΤΕΧΝΙΚΕΣ ΑΝΑΛΥΣΗΣ ΤΗΣ ΔΟΜΗΣ ΤΩΝ ΜΑΚΡΟΜΟΡΙΩΝ</i>	<i>18</i>
<i>1. ΕΙΣΑΓΩΓΗ</i>	<i>18</i>
<i>2. ΚΡΥΣΤΑΛΛΟΓΡΑΦΙΑ ΑΚΤΙΝΩΝ-Χ</i>	<i>19</i>
<i>3. ΚΥΚΛΙΚΟΣ ΔΙΧΡΩΙΣΜΟΣ, CIRCULAR DICHROISM (CD)</i>	<i>19</i>
<i>4. ΠΥΡΗΝΙΚΟΣ ΜΑΓΝΗΤΙΚΟΣ ΣΥΝΤΟΝΙΣΜΟΣ (NMR)</i>	<i>20</i>
<i>5. ΗΛΕΚΤΡΟΝΙΑΚΗ ΜΙΚΡΟΣΚΟΠΙΑ</i>	<i>21</i>
<i>6. ΣΚΕΛΑΣΗ ΑΚΤΙΝΩΝ-Χ ΜΙΚΡΗΣ ΓΩΝΙΑΣ (SAXS)</i>	<i>22</i>
<i>7. ΣΥΓΚΡΙΣΗ ΤΗΣ SAXS ΜΕ ΤΙΣ ΒΑΣΙΚΕΣ ΤΕΧΝΙΚΕΣ ΔΟΜΙΚΗΣ ΑΝΑΛΥΣΗΣ ...</i>	<i>25</i>
ΜΕΘΟΔΟΙ-ΠΡΟΓΡΑΜΜΑΤΑ	26
<i>A) ΜΕΘΟΔΟΙ.....</i>	<i>26</i>
<i>1. ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ</i>	<i>26</i>
<i>B) ΠΡΟΓΡΑΜΜΑΤΑ</i>	<i>29</i>
<i>1. ΠΡΟΓΡΑΜΜΑ ΜΟΡΙΑΚΗΣ ΑΠΕΙΚΟΝΙΣΗΣ (VMD)</i>	<i>29</i>
<i>2. ΠΡΟΓΡΑΜΜΑ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ (NAMD)</i>	<i>30</i>
<i>3. ΤΟ ΠΡΟΓΡΑΜΜΑ CRY SOL</i>	<i>30</i>
<i>4. ΠΡΟΓΡΑΜΜΑ PRIMUS</i>	<i>32</i>

ΥΠΟΛΟΓΙΣΤΙΚΟ ΜΕΡΟΣ	33
1. ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ ΑΡΧΕΤΟΥ PSF & PDB	33
3. ΠΡΟΣΘΗΚΗ ΙΟΝΤΩΝ (NEUTRALIZATION / ADD IONS).....	34
4. ΕΛΑΧΙΣΤΟΠΟΙΗΣΗ ΕΝΕΡΓΕΙΑΣ (ENERGY MINIMIZATION).....	34
5. ΕΞΙΣΟΡΡΟΠΗΣΗ (EQUILIBRATION)	35
6. ΣΤΑΔΙΑ ΥΠΟΛΟΓΙΣΤΙΚΟΥ ΜΕΡΟΥΣ	35
ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ	42
1. ΥΠΟΛΟΓΙΣΜΟΣ ΚΑΙ ΑΝΑΛΥΣΗ RMSD	42
2. ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΓΥΡΟΣΚΟΠΙΚΗΣ ΑΚΤΙΝΑΣ	43
3. ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ D_{max}	45
4. ΣΥΓΚΡΙΣΗ ΚΑΜΠΥΛΩΝ ΤΩΝ ΕΝΤΑΣΕΩΝ ΣΚΕΔΑΣΗΣ	47
ΠΑΡΑΡΤΗΜΑ	49
SCRIPTS.....	49
ΒΙΒΛΙΟΓΡΑΦΙΑ	56

Σκοπός

Σκοπός της παρούσας διπλωματικής εργασίας είναι η προσπάθεια βελτίωσης των προφίλ σκέδασης ακτίνων-X μικρής γωνίας (SAXS) όπως προκύπτουν από το πρόγραμμα CRY SOL με εισαγόμενα δομές της λυσοζύμης από προσομοιώσεις μοριακής δυναμικής με χρήση των force fields CHARMM 27 και CHARMM 36.

Περίληψη

Οι πρωτεΐνες αποτελούν μία ομάδα ιδιαίτερων μακρομορίων, τα οποία παίζουν ποικίλους ρόλους σε διάφορες διεργασίες στους ζώντες οργανισμούς. Οι φυσιολογικές δομές των πρωτεϊνών μπορούν να διαταραχθούν από την έκθεσή τους σε διάφορους χημικούς και φυσικούς παράγοντες οδηγώντας στην εμφάνιση μεγάλου αριθμού νοσημάτων. Έτσι, χάρη στην κεντρικό ρόλο των μακρομορίων αυτών έχουν αναπτυχθεί με την πάροδο του χρόνου διάφορες τεχνικές δομικής ανάλυσης των πρωτεϊνών. Στην παρούσα διπλωματική θα εστιάσουμε στην μέθοδο σκέδασης ακτίνων-X μικρής γωνίας (SAXS), μελετώντας την λυσοζύμη με την βοήθεια του προγράμματος CRY SOL και την χρήση προσομοιώσεων μοριακής δυναμικής. Τέλος, θα συγκρίνουμε τα αποτελέσματα που προκύπτουν από την χρήση δύο διαφορετικών Force Fields, του CHARMM27 και του νέου Force Field CHARMM36.

Abstract

Proteins constitute a group of particular macromolecules, which play a determinant role in an assortment of biological processes in living organisms. The native protein structures can be disturbed by various chemical and physical factors leading to the onset of numerous diseases. Thus, due to their crucial role, various structural analysis techniques have been developed in the past years. The main focus of the present thesis is the small-angle X-ray scattering (SAXS) as it is applied on lysozyme using the program CRY SOL and Molecular Dynamics Simulations. Lastly, the comparison between the Force Fields CHARMM27 and the new Force Field CHARMM36 will be made and discussed.

ΕΙΣΑΓΩΓΗ

Στο πρώτο μέρος της εισαγωγής της διπλωματικής μας εργασίας θα παρατεθούν βασικά στοιχεία για τις πρωτεΐνες και τις ιδιότητες τους, θα αναφερθούμε στον τρόπο με τον οποίο ορισμένες αλλαγές στην δομή των πρωτεϊνών συνδέονται με την εμφάνιση νοσημάτων. Στο δεύτερο μέρος της εισαγωγής θα αναφερθούμε στις βασικές τεχνικές δομικής ανάλυσης των πρωτεϊνών και θα τις συγκρίνουμε με την σκέδαση ακτίνων-Χ μικρής γωνίας, που χρησιμοποιήσαμε στην δική μας μελέτη.

A) ΛΙΓΑ ΛΟΓΙΑ ΓΙΑ ΤΙΣ ΠΡΩΤΕΪΝΕΣ

1. ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΚΑΙ ΙΔΙΟΤΗΤΕΣ ΤΩΝ ΠΡΩΤΕΪΝΩΝ

Οι πρωτεΐνες είναι τα πιο πολυδύναμα μακρομόρια στους ζώντες οργανισμούς και εξυπηρετούν βασικές λειτουργίες σε όλες σχεδόν τις βιολογικές τους διεργασίες. Λειτουργούν ως καταλύτες, ως μεταφορείς και αποθηκευτές άλλων μορίων όπως το οξυγόνο, παρέχουν μηχανική στήριξη και ανοσοπροστασία, δημιουργούν κίνηση, διαβιβάζουν νευρικές ώσεις και ρυθμίζουν την ανάπτυξη και τη διαφοροποίηση. Οι πρωτεΐνες είναι μεγάλα σύνθετα μακρομόρια, με μοριακό βάρος που ποικίλλει από 10 έως και πάνω από 1.000 kDa, αποτελούνται από αμινοξέα, τα οποία ενώνονται μεταξύ τους με πεπτιδικούς δεσμούς σχηματίζοντας μίας γραμμική αλυσίδα, η οποία καλείται αλυσίδα πολυπεπτιδίων. Όλες οι πρωτεΐνες περιέχουν άνθρακα, οξυγόνο και άζωτο και οι περισσότερες εξ αυτών θείο [1]. Διάφορες βασικές ιδιότητες επιτρέπουν στις πρωτεΐνες να συμμετέχουν σε ένα ευρύ φάσμα λειτουργιών.

- *Οι πρωτεΐνες είναι γραμμικά πολυμερή δομούμενα από μονομερή αμινοξέων. Με τον τρόπο αυτό η λειτουργία μιας πρωτεΐνης εξαρτάται άμεσα από την τριδιάστατη δομή της. Το εντυπωσιακό είναι ότι οι πρωτεΐνες αναδιπλώνονται αυθόρμητα σε τριδιάστατες δομές που καθορίζονται από την αλληλουχία των αμινοξέων του πρωτεϊνικού πολυμερούς.*
- *Οι πρωτεΐνες περιέχουν μια μεγάλη σειρά λειτουργικών ομάδων. Οι ομάδες αυτές περιλαμβάνουν αλκοόλες, θειόλες, θειοεστέρες, καρβοξυλικές ομάδες,*

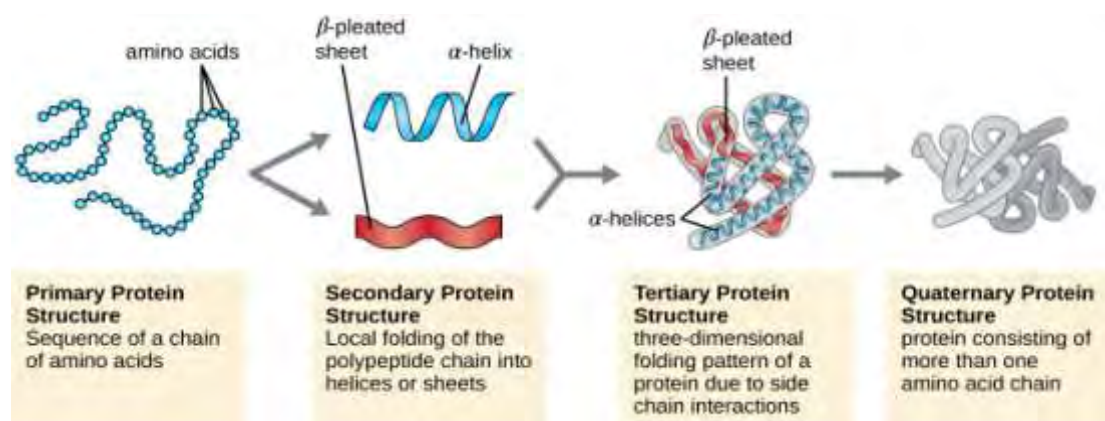
καρβαμίδια και ποικιλία βασικών ομάδων. Όταν συνδυάζονται σε διαφορετικές αλληλουχίες, οι λειτουργικές αυτές ομάδες ερμηνεύουν το φάσμα λειτουργιών των πρωτεϊνών. Παραδείγματος χάριν, η χημική αντιδραστικότητα που οφείλεται στις ομάδες αυτές είναι απαραίτητη για την λειτουργία των ενζύμων, των πρωτεϊνών οι οποίες καταλύουν ειδικές χημικές αντιδράσεις στα βιολογικά συστήματα.

- *Οι πρωτεΐνες μπορούν να αλληλεπιδράσουν μεταξύ τους και με άλλα βιολογικά μακρομόρια, για να δημιουργήσουν πολύπλοκα σύμπλοκα. Μέσα στα σύμπλοκα αυτά, οι πρωτεΐνες μπορούν να δράσουν συνεργειακά και να εμφανίσουν ιδιότητες που δεν προϋπήρχαν στα ανεξάρτητα μακρομόρια. Στα σύμπλοκα περιλαμβάνονται μακρομοριακές μηχανές που επιτυγχάνουν την ακριβή αντιγραφή του DNA, τη μεταγωγή σήματος μέσα στο κύτταρο, αλλά και πολλές άλλες απαραίτητες διεργασίες.*
- *Μερικές πρωτεΐνες είναι σχεδόν άκαμπτες, ενώ υπάρχουν άλλες που εμφανίζουν μία σχετική ευκαμψία. Οι άκαμπτες μονάδες μπορούν να λειτουργήσουν ως δομικά στοιχεία του κυτταρικού σκελετού ή του συνδετικού ιστού. Τα τμήματα των πρωτεϊνών που παρουσιάζουν περιορισμένη ευκαμψία μπορούν να χρησιμοποιηθούν ως αρθρωτά τμήματα, ελατήρια ή μοχλοί απαραίτητα για την πρωτεϊνική λειτουργία, για την σύνδεση πρωτεϊνών μεταξύ τους και με άλλα μόρια σε πολύπλοκα σύμπλοκα και για την μεταβίβαση πληροφοριών μεταξύ κυττάρων αλλά και σε επικοινωνία εντός του κυττάρου. Έχει βρεθεί ένας σημαντικός αριθμός νοσημάτων και ασθενειών που συνδέονται με μεταβολές στην δομή, και κατ' επέκταση με την λειτουργία των πρωτεϊνών [2].*

2. ΠΡΩΤΟΤΑΓΗΣ – ΔΕΥΤΕΡΟΤΑΓΗΣ – ΤΡΙΤΟΤΑΓΗΣ - ΤΕΤΑΡΤΟΤΑΓΗΣ ΔΟΜΗ ΤΩΝ ΠΡΩΤΕΪΝΩΝ

Οι πρωτεΐνες χαρακτηρίζονται από την ύπαρξη πρωτοταγούς, δευτεροταγούς, τριτοταγούς και τεταρτοταγούς δομής [2]. Οι δομές αυτές προσδίδουν στην κάθε πρωτεΐνη την προσωπικότητά της και καθορίζουν την λειτουργία και τον ρόλο της στους ζώντες οργανισμούς.

- Η *πρωτοταγής δομή* καθορίζεται από την αλληλουχία των αμινοξέων στην πολυπεπτιδική αλυσίδα.
- Η *δευτεροταγής δομή* καθορίζει ποιες διαμορφώσεις των πεπτιδίων επιτρέπονται στερεοχημικά και ποιες αξιοποιούν στο έπακρο την δυνατότητα για σχηματισμό δεσμών υδρογόνου των ομάδων -NH και -CO της κύριας αλυσίδας.
- Η *τριτοταγής δομή* αναφέρεται στην αναδίπλωση της πολυπεπτιδικής αλυσίδας πάνω στον εαυτό της και στο τελικό σχήμα που θα πάρει στον χώρο. Σ' αυτή την διαμόρφωση σημαντικό ρόλο παίζουν οι πλευρικές αλυσίδες (R) των αμινοξέων. Ανάμεσα σε αυτές τις ομάδες μπορούν να αναπτυχθούν τόσο ομοιοπολικοί όσο και δευτερεύοντες δεσμοί (πολικοί και μη πολικοί). Η αναδίπλωση γίνεται έτσι ώστε οι υδρόφοβες πλευρικές ομάδες να βρίσκονται στο εσωτερικό. Οι ομοιοπολικοί δεσμοί είναι κυρίως δισουλφιδικοί δεσμοί (γέφυρες θείου). Σπανιότερα αναπτύσσονται δεσμοί τύπου αλδιμίνης (κολλαγόνο, ελαστίνη), πεπτιδικοί δεσμοί ανάμεσα στην γ-καρβοξυλομάδα του γλουταμινικού οξέος και της ε-αμινομάδας της λυσίνης (ινώδες), θειεστερικοί δεσμοί ανάμεσα στην γ-καρβοξυλομάδα του γλουταμινικού οξέος και της σουλφιδρυλομάδας της κυστεΐνης και πιο σπάνια ο αιθερικός δεσμός ανάμεσα σε δύο τυροσίνες.
- Η *τεταρτοταγής δομή* αφορά στις πρωτεΐνες που αποτελούνται από περισσότερες από μία πολυπεπτιδικές αλυσίδες (υπομονάδες). Αναφέρεται στις σχέσεις στον χώρο των διαφόρων πολυπεπτιδικών αλυσίδων ενός πρωτεϊνικού μορίου. Η διαμόρφωση αυτή σταθεροποιείται μόνο από δευτερεύοντες δεσμούς (υδρόφοβοι, δ.Υ. και γέφυρες άλατος). Οι περιοχές επαφής μεταξύ των υπομονάδων έχουν πολύ συχνά λειτουργικό ρόλο.



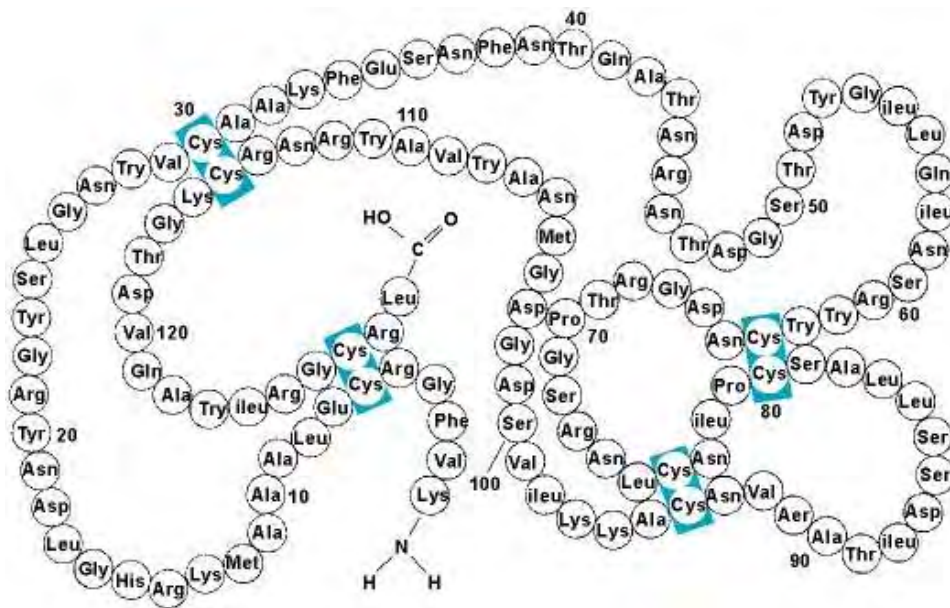
Εικόνα 1: Εικόνα όπου φαίνεται η πρωτοταγής, δευτεροταγής, τριτοταγής και τεταρτοταγής διαμόρφωση μίας πρωτεΐνης, με σειρά από τα αριστερά προς τα δεξιά.

3. Η ΛΥΣΟΖΥΜΗ

Στην διπλωματική μας εργασία ασχοληθήκαμε με μια από τις πιο πολύ μελετημένες εργαστηριακά πρωτεΐνες, την λυσοζύμη. Η λυσοζύμη, είναι μια υδρολάση, γνωστή επίσης και ως μουραμιδάση ή N-ακέτυλομουραμιδική γλυκανδρουλάση. Είναι ένα ένζυμο με αντιμικροβιακή δράση και αποτελεί μέρος του έμφυτου ανοσοποιητικού συστήματος. Στον άνθρωπο το ένζυμο της λυσοζύμης κωδικοποιείται από το γονίδιο LYZ, το προϊόν του οποίου αποτελείται από 129 αμινοξέα και έχει μοριακό βάρος 14,6 kDa [3, 4]. Κατά μήκος του μορίου υπάρχουν 4 δισουλφιδικοί δεσμοί, οι οποίοι βοηθούν το μόριο να διατηρήσει την σταθερότητά του. Για την μελέτη μας χρησιμοποιήσαμε την λυσοζύμη του οργανισμού Gallus gallus (όρνια), η πλήρης ακολουθία της οποίας είναι η εξής:

```
>1DPX:A|PDBID|CHAIN|SEQUENCE
KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQINS
RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNC AKKIVSDGNMNAWVAWRNRCKGTDV
QAWIRGCR L
```

(πηγή: <http://www.uniprot.org/uniprot/O6LEL2.fasta>)



Εικόνα 2: Πρωτοταγής δομή - Πλήρης αλληλουχία των αμινοξέων της λυσοζύμης. MrBorden's Biology Rattler Site Room 664 – WordPress.com

Ο Alexander Fleming, ο οποίος ανακάλυψε την πενικιλίνη, ήταν ο πρώτος που παρατήρησε για πρώτη φορά την αντιβακτηριακή δράση της λυσοζύμης, όταν έλαβε καλλιέργειες με ρινική βλέννα από έναν ασθενή που έπασχε από κρυολόγημα [5]. Η λυσοζύμη είναι άφθονη σε εκκρίσεις όπως τα δάκρυα, το σάλιο, η βλέννα, το μητρικό γάλα, ενώ η μεγάλη ποσότητα λυσοζύμης εντοπίζεται επίσης και στο ασπράδι του αυγού. Το ένζυμο αυτό προκαλεί ρήξη των γλυκοσιδικών δεσμών της πεπτιδογλυκάνης, η οποία απαντάται στα κυτταρικά τοιχώματα των βακτηρίων, και ιδιαίτερα των θετικών κατά Gram βακτηρίων. Προσβάλλει την πεπτιδογλυκάνη που αποτελεί το φυσικό υπόστρωμα μεταξύ του N-ακέτυλομουραμικού οξέος (NAM), και του τέταρτου ατόμου άνθρακα της N-ακέτυλογλυκοζαμίνης (NAG). Η λυσοζύμη μπορεί επίσης να διασπάσει τους γλυκοσιδικούς δεσμούς στην χιτίνη, με μικρότερη βέβαια αποτελεσματικότητα από τις χιτινάσες [6].

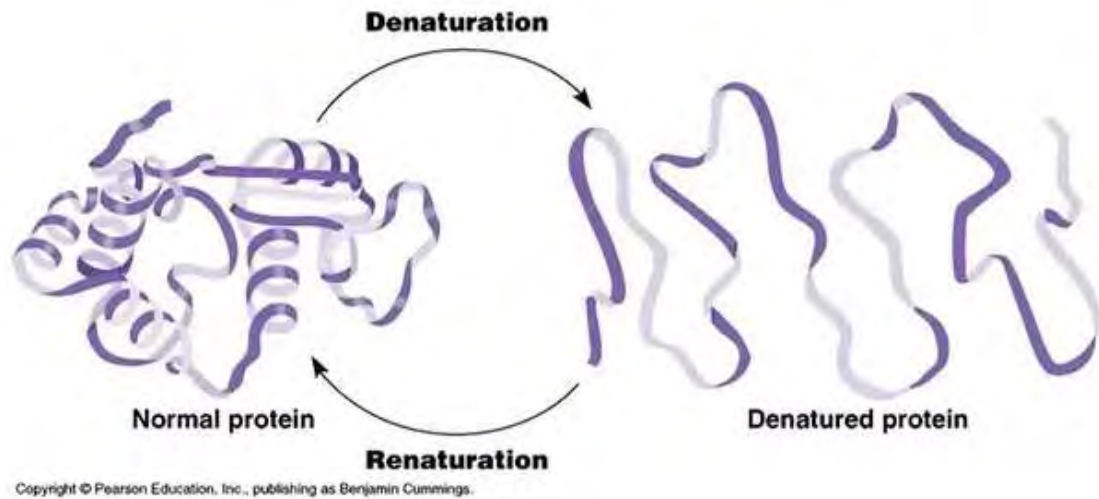


Εικόνα 3: Η κρυσταλλική δομή της λυσοζύμης (PDB code: 2LZT)

Όπως θα δούμε στην συνέχεια, η δομή και κατ' επέκταση η λειτουργία των πρωτεϊνών επηρεάζεται από ορισμένους φυσικούς και χημικούς παράγοντες, ενώ από την άλλη πλευρά, λόγω της υψηλής τους ευαισθησίας, έχουμε την εμφάνιση ενός μεγάλου αριθμού νοσημάτων και ασθενειών.

4. Η ΜΕΤΟΥΣΙΩΣΗ ΤΩΝ ΠΡΩΤΕΪΝΩΝ

Οποιαδήποτε αλλαγή στην διαμόρφωση του μορίου της πρωτεΐνης χαρακτηρίζεται ως μετουσίωση. Η μετουσίωση αφορά αλλαγές στη δευτεροταγή, τριτοταγή ή τεταρτοταγή δομή της πρωτεΐνης, ενώ δεν αναφέρεται σε αλλαγές της πρωτοταγούς δομής. Κατά την μετουσίωση οι δεσμοί που έχουν αναπτυχθεί μεταξύ των πλευρικών ομάδων των αμινοξέων διασπώνται, με συνέπεια την καταστροφή της τριδιάστατης δομής της πρωτεΐνης και την απώλεια έτσι της λειτουργικότητας της. Η μετουσίωση μπορεί να προκύψει από την επίδραση διαφόρων φυσικών και χημικών παραγόντων και συνδυασμών τους, όπως θερμότητα, επίδραση οξέων ή αλκάλων, επίδραση πυκνών διαλυμάτων άλμης, ακτινοβολία, μηχανική επίδραση, υψηλή πίεση, υπέρηχοι και λοιπά. Κατά το φαινόμενο αυτό προκύπτουν νέες διαμορφώσεις προσωρινές ή μόνιμες και συνήθως ενδιάμεσες και μικρής διάρκειας ζωής. Η μετουσίωση μπορεί να είναι αντιστρεπτή ή μη-αντιστρεπτή. Στην πρώτη περίπτωση, όταν σταματήσει η επίδραση του παράγοντα μετουσίωσης τότε η διαμόρφωση επανέρχεται στην αρχική της μορφή, κάτι που δεν συμβαίνει στην περίπτωση της μη-αντιστρεπτής μετουσίωσης. Χαρακτηριστικό παράδειγμα μη-αντιστρέψιμης μετουσίωσης αποτελεί η θέρμανση του λευκώματος του αυγού που δίνει συμπαγές ασπράδι (μετουσίωση αλβουμίνης). Συνήθως όταν υπάρξει μόνο εκτύλιξη της πολυπεπτιδικής αλυσίδας, τότε η μετουσίωση είναι αντιστρεπτή. Εάν διασπαστούν οι δισουλφιδικοί δεσμοί που μπορεί να συμμετέχουν στην διαμόρφωση της δομής της, τότε συνήθως η μετουσίωση είναι μη-αντιστρεπτή. Τέλος, η ευαισθησία των πρωτεϊνών στους διάφορους παράγοντες μετουσίωσης διαφέρει και εξαρτάται από την ικανότητα του παράγοντα να διασπάσει τους δεσμούς που σταθεροποιούν την δευτεροταγή, τριτοταγή ή τεταρτοταγή δομή της πρωτεΐνης.



Εικόνα 4: Εικόνα όπου φαίνεται η επίδραση της μετουσίωσης στην διαμόρφωση μιας πρωτεΐνης. Πρόκειται για αντιστρεπτή μετουσίωση, αφού η πρωτεΐνης μπορεί να επανέλθει στην αρχική της μορφή μετά την πάροδο της επίδρασης του παράγοντα μετουσίωσης.

5. ΑΝΩΜΑΛΗ ΑΝΑΔΙΠΛΩΣΗ ΠΡΩΤΕΪΝΩΝ ΚΑΙ ΝΟΣΗΜΑΤΑ

Έχει διαγνωσθεί ένας μεγάλος αριθμός ασθενειών που ευθύνονται σε "ανώμαλη" αναδίπλωση των πρωτεϊνών (δυσδιαμόρφωση) [7]. Αυτού του τύπου οι ασθένειες μπορούν να διαχωριστούν σε τρεις κύριες κατηγορίες:

1. *Ασθένειες απώλειας λειτουργίας:* Οι περισσότερες ασθένειες απώλειας λειτουργίας προκαλούνται από την ανώμαλη αναδίπλωση πρωτεϊνών, λόγω κληρονομούμενων μεταλλάξεων. Χαρακτηριστικά παραδείγματα ασθενειών της κατηγορίας αυτής αποτελεί το σύνδρομο ανεπάρκειας ανθρακικής ανυδράσης II (CADS), το οποίο εκδηλώνεται εξαιτίας σημειακών μεταλλάξεων στο γονίδιο της ανθρώπινης ανθρακικής ανυδράσης II (HCA II), αλλά και το Parkinson, στην περίπτωση του οποίου παρατηρείται απώλεια λειτουργίας της α -συνουκλείνης, μίας πρωτεΐνης που εμφανίζεται σε αφθονία στον ανθρώπινο εγκέφαλο και βρίσκεται κυρίως στις άκρες των νευρώνων.

2. *Ασθένειες αύξησης λειτουργίας-τοξικότητας*: Στην κατηγορία αυτή έχουμε μεταβολές του προϊόντος ενός γονιδίου, κάτι που μπορεί να οδηγήσει σε υπερέκφραση πρωτεϊνών με ανώμαλη λειτουργία. Χαρακτηριστικά παραδείγματα αποτελούν η νόσος Huntington (HD) και η νόσος Charcot-Marie-Tooth (υπερέκφραση του γονιδίου PMP22).

3. Οι μόνες μολυσματικές ασθένειες οφειλόμενες σε δυσδιαμόρφωση είναι εκείνες των πρωτεϊνών Prions. Χαρακτηριστικά παραδείγματα της κατηγορίας αυτής είναι η νόσος Kuru, και η ασθένεια Gerstmann-Straussler Scheinker (GSS).

Πίνακας 1. Παραδείγματα ασθενειών που οφείλονται στην ανώμαλη αναδίπλωση πρωτεϊνών, καθώς και οι πρόδρομες πρωτεΐνες που εμπλέκονται στις ασθένειες αυτές.

	Disease	Precursor protein
Loss-of-function	Cancer	P53 tumor suppressor
	Cystic Fibrosis	CFTR
	Carbonic Anhydrase II Deficiency Syndrome	HCA II
Gain-of-toxic function	Alzheimers disease	Aβ-protein
	Type II diabetes	IAPP
	ALS	Superoxide dismutase
	Parkinson's disease	α-synuclein
Infectious-misfolding	Creutzfeldt-Jakob disease	Prion protein
	BSE	Prion protein
	Kuru	Prion protein
	Scrapie	Prion protein

B) ΤΕΧΝΙΚΕΣ ΑΝΑΛΥΣΗΣ ΤΗΣ ΔΟΜΗΣ ΤΩΝ ΜΑΚΡΟΜΟΡΙΩΝ

Σε αυτό το τμήμα της εισαγωγής, θα αναφερθούμε στις βασικές τεχνικές που έχουμε στην διάθεση μας, για να λάβουμε μία εικόνα της δομής των διαφόρων μακρομορίων, και θα συγκρίνουμε τις μεθόδους αυτές με την μέθοδο SAXS που χρησιμοποιήσαμε στις μελέτες που πραγματοποιήσαμε στην παρούσα διπλωματική εργασία.

1. ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια έγιναν δύο επιστημονικές επαναστάσεις που επηρέασαν άμεσα τις βιολογικές επιστήμες. Η επανάσταση στη γενομική (genomic revolution) και η λεγόμενη δομική επανάσταση (structural revolution) ή αλλιώς επανάσταση της δομικής βιολογίας. Η πρώτη περιλαμβάνει την αλληλούχιση των γονιδίων ολόκληρων οργανισμών από τα βακτήρια, μέχρι και τον άνθρωπο. Ο συσχετισμός της γενετικής πληροφορίας με τον λειτουργικό ρόλο των προϊόντων κωδικοποίησης αποτελεί πρόκληση για την μετα-γενομική βιολογία [8]. Η δομική επανάσταση περιλαμβάνει την εξέλιξη και ανάπτυξη τεχνικών για την επίλυση δομών των βιολογικών μακρομορίων. Η λεπτομερής περιγραφή των δομών αποτέλεσε την εκλογικευμένη βάση της διερεύνησης του ρόλου των βιολογικών μακρομορίων και βοήθησε στην ανάπτυξη σημαντικών τομέων της βιοχημείας, όπως η δομική ενζυμολογία, η οποία διαθέτει διάφορες σημαντικές εφαρμογές όπως π.χ στον κατευθυνόμενο από τη δομή σχεδιασμό φαρμάκων.

Έτσι, δεν είναι υπερβολή να πούμε ότι σήμερα θα μπορούσε να αναδιατυπωθεί το κεντρικό δόγμα της βιολογίας ως: *Η αλληλουχία καθορίζει την δομή, η οποία με την σειρά της καθορίζει την λειτουργία* [8]. Γενικά, τα βιολογικά μακρομόρια είναι μόρια μεγάλου μοριακού βάρους όπως είναι οι πρωτεΐνες (η λυσοζύμη με την οποία θα ασχοληθούμε), τα νουκλεϊκά οξέα, οι πολυσακχαρίτες, αλλά και σύμπλοκα αυτών που απαντώνται στα βιολογικά συστήματα [9]. Στο κομμάτι που ακολουθεί θα προσπαθήσουμε να δώσουμε μία εικόνα των βασικών τεχνικών που παρέχουν δομικές πληροφορίες για τα διάφορα μακρομόρια, όπως είναι η κρυσταλλογραφία ακτίνων-X, ο πυρηνικός μαγνητικός συντονισμός (NMR), η ηλεκτρονιακή μικροσκοπία και ο κυκλικός διχρωϊσμός (CD), και έπειτα θα συγκρίνουμε τις μεθόδους αυτές και θα παραθέσουμε πληροφορίες για την τεχνική που χρησιμοποιήθηκε στη δική μας μελέτη, την σκέδαση ακτίνων-X μικρής γωνίας (SAXS).

2. ΚΡΥΣΤΑΛΛΟΓΡΑΦΙΑ ΑΚΤΙΝΩΝ-X

Η κρυσταλλογραφία ακτίνων-X αποτελεί την πιο συνηθισμένη τεχνική επίλυσης μακρομοριακής δομής σε ατομικό επίπεδο, και θα λέγαμε ότι βρίσκεται στην βάση της πυραμίδας των δομικών τεχνικών, καθώς είναι η παλαιότερη ιστορικά και έχει χρησιμοποιηθεί για την ανίχνευση πάνω από 132055 πρωτεϊνικών δομών αλλά και δομών νουκλεϊκών οξέων [10]. Η θέση του ατόμου ενός μακρομορίου στον τριδιάστατο χώρο, σε σχέση με όλα τα άλλα άτομα του μορίου, ονομάζεται ατομική συντεταγμένη (atomic coordinate). Με την κρυσταλλογραφία ακτίνων-X είναι δυνατός ο προσδιορισμός των ατομικών συντεταγμένων των ατόμων ενός μακρομορίου. Η τεχνική αυτή στηρίζεται στην επεξεργασία των δεδομένων περίθλασης ακτίνων-X από κρυστάλλους μακρομορίων. Υπό συγκεκριμένες συνθήκες τα βιολογικά μακρομόρια, όπως οι πρωτεΐνες, μπορούν να περάσουν από την διαλυτή κατάσταση στην κρυσταλλική. Η κρυσταλλογραφία ακτίνων-X δηλαδή, προϋποθέτει την κρυστάλλωση των μακρομορίων, η οποία δεν είναι πάντοτε εφικτή, αλλά μπορεί να επιλύσει μακρομοριακές δομές σε υψηλή διακριτική ικανότητα (ευκρίνεια 0.5 – 3 Å). Οι κρύσταλλοι αποτελούνται από κανονικές τριδιάστατες συστοιχίες μορίων, που συγκρατούνται μεταξύ τους με μη-ομοιοπολικές αλληλεπιδράσεις. Τα στάδια της επίλυσης της δομής ενός μακρομορίου με κρυσταλλογραφία ακτίνων-X περιλαμβάνουν την κρυστάλλωσή του, τη συλλογή των κρυσταλλογραφικών δεδομένων σε συμβατικές ή συγχροτρονικές πηγές ακτίνων-X, και την δημιουργία ενός ατομικού δομικού μοντέλου (επίλυση φάσεων και κατασκευή χαρτών ηλεκτρονιακής πυκνότητας που δείχνουν την κατανομή των ηλεκτρονίων γύρω από τα άτομα σε ένα μόριο) και τη βελτιστοποίηση του σε συμφωνία με τα κρυσταλλογραφικά πειραματικά δεδομένα.

3. ΚΥΚΛΙΚΟΣ ΔΙΧΡΩΙΣΜΟΣ, CIRCULAR DICHROISM (CD)

Μία οπτικά ενεργή ουσία απορροφά διαφορετικά την δεξιά κυκλικά πολωμένη ακτινοβολία από την αριστερά κυκλικά πολωμένη ακτινοβολία. Αυτή η άνιση απορρόφηση των δύο συνιστούντων διανυσμάτων, έχει ως αποτέλεσμα το συνιστάμενο διάνυσμα να ιχνογραφεί μια έλλειψη, να είναι δηλαδή ελλειπτικά

πολωμένο. Το φαινόμενο αυτό καλείται κυκλικός διχρωϊσμός (CD). Τα φάσματα κυκλικού διχρωϊσμού καταγράφουν τη διαφορά απορρόφησης της δεξιά και αριστερά πολωμένης ακτινοβολίας ως συνάρτηση του μήκους κύματος. Ο κυκλικός διχρωϊσμός αποτελεί ένα σημαντικό μέσο για την διερεύνηση των ποσοστών των στοιχείων της δευτεροταγούς δομής μιας πρωτεΐνης, μπορεί να ανιχνεύσει αλλαγές στην διαμόρφωση πρωτεϊνών που οφείλονται σε μεταβολές του pH, της θερμοκρασίας, μεταλλάξεων, πρόσδεσης υποκαταστατών, ενώ παρέχει πληροφορίες και για την πρωτεϊνική αναδίπλωση [11].

4. ΠΥΡΗΝΙΚΟΣ ΜΑΓΝΗΤΙΚΟΣ ΣΥΝΤΟΝΙΣΜΟΣ (NMR)

Η τεχνική της φασματοσκοπίας NMR ανακαλύφθηκε το 1946 από τους Bloch και Purcell [12] και μπορεί να χρησιμοποιηθεί για τον υπολογισμό των αποστάσεων των ατομικών πυρήνων σε ένα μόριο, οδηγώντας έτσι σε ένα τριδιάστατο μοντέλο του μακρομορίου και παρέχοντας πληροφορίες για την δυναμική του. Στηρίζεται στο φαινόμενο της διέγερσης ατομικών πυρήνων με περιττό ατομικό αριθμό, που διαθέτουν στροφορμή (spin) υπό την επίδραση ηλεκτρομαγνητικής ακτινοβολίας όταν βρίσκονται σε μαγνητικό πεδίο. Για παράδειγμα, όταν τα μόρια μιας πρωτεΐνης τοποθετούνται σε ένα ισχυρό μαγνητικό πεδίο τα spin μέρους των ατόμων του υδρογόνου ευθυγραμμίζονται με το πεδίο (κατάσταση ισορροπίας). Παλμοί ραδιοσυχνότητας (RF) διεγείρουν τους πυρήνες, οι οποίοι όταν επανέρχονται σε κατάσταση ισορροπίας εκπέμπουν ακτινοβολία RF. Η συχνότητα αυτής της ακτινοβολίας εξαρτάται από το μοριακό περιβάλλον του πυρήνα [13]. Τα σήματα της ακτινοβολίας καταγράφονται ως φάσματα διαφόρων τύπων, στην συνέχεια επεξεργάζονται και ταυτοποιούνται. Το τελικό αποτέλεσμα είναι μία λίστα με περιορισμούς στην απόσταση μεταξύ συγκεκριμένων ζευγών ατόμων υδρογόνου, από όπου μπορούν να αναγνωριστούν π.χ δομικά στοιχεία της δευτεροταγούς και της τριτοταγούς δομής μιας πρωτεΐνης, και να κατασκευαστεί ένα τριδιάστατο δομικό μοντέλο. Σημαντικό πλεονέκτημα της NMR είναι ότι με αυτήν μπορεί να μελετηθεί η δυναμική συμπεριφορά των μακρομορίων (χημική ισορροπία, κινητικότητα, υπολογισμός τιμών K_d προσδετών, αποδιάταξη). Επίσης, σε σύγκριση με την κρυσταλλογραφία ακτίνων-X, δεν απαιτείται κρυστάλλωση και έτσι μπορεί να χρησιμοποιηθεί σε εύρος πειραματικών συνθηκών (θερμοκρασία, pH, ιοντική ισχύ).

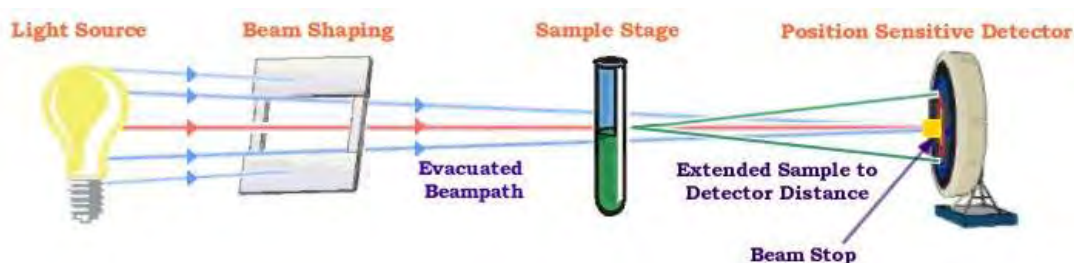
Στα βασικά μειονεκτήματα της μεθόδου συγκαταλέγονται ο περιορισμός στο μέγεθος των μακρομορίων, που πρέπει να είναι μικρότερα από 50 kD, καθώς και η αδυναμία χαρακτηρισμού μιας μοναδικής δομής, αφού η NMR δίνει ως αποτέλεσμα ένα σετ πολλαπλών δομών του υπό εξέταση μακρομορίου.

5. ΗΛΕΚΤΡΟΝΙΑΚΗ ΜΙΚΡΟΣΚΟΠΙΑ

Η τεχνική της ηλεκτρονιακής μικροσκοπίας μας επιτρέπει να περιγράψουμε την δομή κυτταρικών οργανιδίων, ιών, μακρομορίων και συμπλόκων τους σε ατομική ευκρίνεια. Υπάρχουν διάφορα είδη ηλεκτρονιακής μικροσκοπίας, όπως είναι η ηλεκτρονική μικροσκοπία διέλευσης (TEM), ή μικροσκοπία σάρωσης (SEM) ή EELS, με την πιο συνηθισμένη να είναι η TEM. Η αρχή της λειτουργίας του ηλεκτρονικού μικροσκοπίου είναι ανάλογη του οπτικού μικροσκοπίου, εάν αντικαταστήσουμε το φως με ακτίνες ηλεκτρονίων που παράγεται από θερμαινόμενο μεταλλικό νημάτιο σε υψηλό κενό, και τον γυάλινο φακό με ηλεκτρομαγνητικό πεδίο ώστε να επιτυγχάνεται η εστίαση των ηλεκτρονίων. Το είδωλο του αντικειμένου ανιχνεύεται με ειδική κάμερα (CCD detector). Η πιο σημαντική διαφορά ανάμεσα στο οπτικό και το ηλεκτρονικό μικροσκόπιο είναι η ευκρίνεια της εικόνας, η οποία καθορίζεται από το μήκος κύματος της πηγής της ακτινοβολίας (όσο πιο μικρό το μήκος κύματος, τόσο μεγαλύτερη είναι η ευκρίνεια). Ένα μεγάλο μειονέκτημα της τεχνικής αυτής, είναι οι βλάβες (ρήξη χημικών δεσμών και δημιουργία ελεύθερων ριζών) που προκαλούνται στα βιολογικά δείγματα λόγω της ηλεκτρονιακής ακτινοβολίας. Το πρόβλημα αυτό αντιμετωπίζεται με την μέθοδο της αρνητικής χρώσης. Με την τεχνική αυτή λοιπόν, είναι δυνατή η μελέτη της δομής κυττάρων, ιών και πρωτεϊνών σε ευκρίνεια μέχρι και 20 Å (χαμηλή ευκρίνεια). Ένα επίσης αρνητικό της μεθόδου αυτής, είναι ότι δεν μπορεί να χρησιμοποιηθεί για την ανάλυση μακρομορίων και πρωτεϊνών που περιέχονται σε διάλυμα. Η ηλεκτρονιακή μικροσκοπία χρησιμοποιείται για να επιλυθούν σε χαμηλή ευκρίνεια δομές μεγάλων μακρομοριακών συμπλόκων. Ωστόσο, η εξέλιξη της τεχνικής της κρυο-ηλεκτρονιακής μικροσκοπίας υπόσχεται βελτίωση της ευκρίνειας μέχρι το ατομικό επίπεδο.

6. ΣΚΕΔΑΣΗ ΑΚΤΙΝΩΝ-Χ ΜΙΚΡΗΣ ΓΩΝΙΑΣ (SAXS)

Η βασική ιδέα πίσω από την σκέδαση ακτίνων-Χ μικρής γωνίας είναι απλή: Με την βοήθεια του X-ray Generator έχουμε απελευθέρωση μίας δέσμης ακτίνων-Χ και την πρόσπτωσή της στο πρωτεϊνικό διάλυμα, και στην συνέχεια μετράμε πόσο οι ακτίνες αυτές αποκλίνουν. Οι περισσότερες ακτίνες-Χ ταξιδεύουν απευθείας μέσα από την ύλη του πρωτεϊνικού διαλύματος χωρίς να αλλάζουν κατεύθυνση, αλλά περίπου το 1/1.000.000 εκτρέπεται κατά μία μικρή γωνία. Το μοτίβο σκέδασης μικρής γωνίας σε έναν ανιχνευτή ακτίνων-Χ μοιάζει πολύ με ένα ηλιοβασίλεμα, με την δέσμη που δεν εμφανίζει απόκλιση να μοιάζει με τον ήλιο: υπάρχει μία φωτεινή λάμψη πολύ κοντά την δέσμη που εξασθενεί προς το σκοτάδι (σε υψηλότερες γωνίες). Για να προστατέψουμε τον ευαίσθητο ανιχνευτή μας, αποκλείουμε την άμεση δέσμη με ένα μικρό μεταλλικό ορθογώνιο που ονομάζεται beamstop. Έτσι, με την χρήση των σύγχρονων αλγορίθμων και των δεδομένων που λαμβάνονται από τον ανιχνευτή ακτίνων-Χ, μπορούμε να λάβουμε πολύτιμες δομικές πληροφορίες για τα βιομόρια που μας ενδιαφέρουν. Το SAXS με τον τρόπο αυτό γίνεται ένα απαραίτητο εργαλείο της επιστήμης, και δίνει την δυνατότητα στους ερευνητές να προσδιορίσουν την μάζα και το μέγεθος ενός βιομορίου στο διάλυμα, να εξακριβώσουν εάν αυτό είναι άκαμπτο ή εύκαμπτο, αλλά και να καταλάβουν πώς πολλαπλά μόρια μπορούν να ταιριάξουν μαζί ώστε να σχηματίσουν περίπλοκα μοριακά σύμπλοκα. Το SAXS λοιπόν μας βοηθά να κατανοήσουμε τον τρόπο λειτουργίας των βιομορίων, αλλά και τις αλληλεπιδράσεις τους μέσα σε ένα ζωντανό κύτταρο. Στην **εικόνα 5** έχουμε την σχηματική απεικόνιση της μεθόδου SAXS:



Εικόνα 5: Πειραματική διάταξη SAXS

Η πειραματική διαδικασία περιλαμβάνει επαναλήψεις ζευγών μετρήσεων του προφίλ σκέδασης του πρωτεϊνικού διαλύματος και του αντίστοιχου διαλύτη (buffer) για διαφορετικές συγκεντρώσεις της πρωτεΐνης.

Η ένταση $I(s)$ των ακτίνων-X στον ανιχνευτή (μετά την αφαίρεση της σκέδασης του buffer) δίνεται από την εξίσωση [24]:

$$I(s) = 4\pi \int_0^{D_{\max}} r^2 \gamma(r) \frac{\sin sr}{sr} dr,$$

όπου $s = 4\pi \sin(\theta)/\lambda$, 2θ είναι η γωνία σκέδασης, r είναι το διάνυσμα θέσης μέσα στο δείγμα και γ είναι η σφαιρική μέση τιμή της συνάρτησης αυτοσυσχέτισης της περιόσμιας της πυκνότητας σκέδασης.

Σε συστήματα μονοδιασποράς (monodisperse) και για πολύ μικρές τιμές του s ισχύει

$$I(s) = I(0)e^{-R_g^2 s^2/3},$$

όπου R_g είναι η γυροσκοπική ακτίνα του μορίου της διαλυμένης ουσίας, που αποτελεί μέτρο των διαστάσεων και της κατανομής της μάζας ενός πολυμερούς. Μια πρωτεΐνη θα έχει μικρότερη R_g στην σφαιρική, συμπαγή της διαμόρφωση από ό,τι στην αποδιαταγμένη. Η R_g υπολογίζεται από την κλίση του διαγράμματος Guinier $\left[\ln I(s) = \ln I(0) - \frac{R_g^2 s^2}{3} \right]$.

Πιο συγκεντρωτικά λοιπόν, με την σκέδαση ακτίνων-X μικρής γωνίας, καθορίζονται παράμετροι όπως: α) Η γυροσκοπική ακτίνα R_g (τυπική ακρίβεια 2%), β) Το μοριακό βάρος (τυπική ακρίβεια 10%), γ) Ο βαθμός αναδίπλωσης, ή η ύπαρξη διαταραχής ή μετουσίωσης στην υπό εξέταση πρωτεΐνη, ενώ τα δεδομένα που προκύπτουν από την SAXS μπορούν να χρησιμοποιηθούν από τους ερευνητές για (i) τον προσδιορισμό της φυσιολογικής ολιγομερούς κατάστασης, (ii) την μελέτη των αλληλεπιδράσεων πρωτεΐνης-πρωτεΐνης υπό διαφορετικές συνθήκες διαλύματος, (iii) την οπτικοποίηση των διαταραγμένων περιοχών (disordered regions) που δεν παρατηρούνται στην κρυσταλλογραφία ακτίνων-X, και τέλος, (iv) η ανάλυση του προφίλ SAXS παρέχει πληροφορίες σχετικά με το μέγεθος και το σχήμα των σωματιδίων του διαλύματος. Η τεχνική SAXS σε συνδυασμό με NMR και υπολογιστικές τεχνικές προσφέρει τη δυνατότητα χαρακτηρισμού των ενδογενώς διαταραγμένων πρωτεϊνών (IDPs). Έχει βρεθεί ότι περισσότερο από το 35% των ανθρώπινων πρωτεϊνών παρουσιάζουν

περιοχές με σημαντικές διαταραχές, και ότι περίπου το 25% πιθανόν να είναι εντελώς διαταραγμένες, δηλαδή δεν έχουν σταθερή τριτοταγή δομή στο διάλυμα .

Από την άλλη πλευρά τώρα, η εφαρμογή της μεθόδου SAXS ορισμένες φορές παρεμποδίζεται και δεν καθίσταται εφικτή. Κάτι τέτοιο μπορεί να συμβαίνει όταν το μόριο που θέλουμε να μελετήσουμε είναι πολύ μεγάλο και όταν το δείγμα μας είναι πολύ αραιό ή είναι ετερογενές (σύμπλοκα πρωτεΐνης-DNA-λιπιδίου). Η τεχνική SAXS είναι μια μέθοδος, τα αποτελέσματα της οποίας μπορούν να μεταβληθούν από διαφορές της θερμοκρασίας, του pH, λόγω αναντιστοιχίας buffer, είτε λόγω εμφάνισης συσσωματωμάτων [19].

Έχοντας πλέον μια εικόνα των βασικών τεχνικών δομικής ανάλυσης των μακρομορίων, μπορούμε να συνεχίσουμε στην σύγκριση της τεχνικής που χρησιμοποιήθηκε στην μελέτη μας, την σκέδασης ακτίνων-X μικρής γωνίας (SAXS), με τις προαναφερθείσες τεχνικές.

7. ΣΥΓΚΡΙΣΗ ΤΗΣ SAXS ΜΕ ΤΙΣ ΒΑΣΙΚΕΣ ΤΕΧΝΙΚΕΣ ΔΟΜΙΚΗΣ ΑΝΑΛΥΣΗΣ

Η σκέδαση ακτίνων-X μικρής γωνίας συμπληρώνει την μέθοδο της κρυσταλλογραφίας ακτίνων-X και της φασματοσκοπίας NMR. Προκειμένου να αξιολογηθεί πλήρως το δυναμικό της σκέδασης ακτίνων-X μικρής γωνίας, είναι απαραίτητο όχι μόνο να βελτιωθούν οι πειραματικές τεχνικές, αλλά και να αναπτυχθούν ακριβή και αποτελεσματικά υπολογιστικά προγράμματα για την αναφορά ατομικών μοντέλων στις μετρήσεις [14]. Με την σκέδαση ακτίνων-X μικρής γωνίας μπορούμε να ανιχνεύσουμε αδρά δομικά χαρακτηριστικά των μακρομορίων που περιέχονται σε ένα διάλυμα, χωρίς να απαιτείται η κρυστάλλωση της πρωτεΐνης, κάτι που είναι σίγουρα χρονοβόρο και σε ορισμένες περιπτώσεις δεν καθίσταται εφικτό. Στην κρυσταλλογραφία, όπου η κρυστάλλωση της πρωτεΐνης που θέλουμε να μελετήσουμε είναι απαραίτητη, οι πρωτεϊνικές κινήσεις περιορίζονται, ενώ αντίθετα οι πρωτεΐνες σε ένα διάλυμα υπόκεινται σε φυσική δυναμική και οι αλληλεπιδράσεις μεταξύ τους είναι ελεύθερες. Έτσι, μέσα σε ένα διάλυμα, λόγω της θερμικής κίνησης και της επαφής με τον διαλύτη, τα μόρια της πρωτεΐνης διέρχονται από διαφορετικές διαμορφώσεις, με συχνότερη αυτήν της κρυσταλλικής. Συχνά τα άκρα της πρωτεΐνης καθώς και οι αναστροφές εμφανίζουν μεγάλες αποκλίσεις. Αντίθετα, μία κρυσταλλωμένη πρωτεΐνη συναντάται σε μία μόνο διαμόρφωση. Μερικές φορές μάλιστα, προκειμένου να καταστεί εφικτή η κρυστάλλωση αποκόπτονται ευέλικτα μέρη της πρωτεΐνης.

Προκειμένου να γίνει εφικτή η σύγκριση προτεινόμενων μοριακών μοντέλων με πειραματικά προφίλ SAXS, έχουν αναπτυχθεί προγράμματα (Crysol κ.α) για τον υπολογισμό αυτών των προφίλ από γνωστές κρυσταλλικές δομές ή μοντέλα. Δυστυχώς τα προγράμματα αυτά εμφανίζουν σχετική μόνο επιτυχία και δεν μπορούν να αναπαραγάγουν αρκετά ικανοποιητικά τα πειραματικά προφίλ σκέδασης. Οι πιθανοί λόγοι για αυτό εντοπίζονται τόσο στα προαναφερθέντα, όσο και στο ότι χρησιμοποιούν μοντέλο συνεχούς διαλύτη (νερού). Στο πλαίσιο αυτής της εργασίας επιχειρείται να παρακαμφθούν οι λόγοι που οδηγούν στην μερική ασυμφωνία μεταξύ των υπολογιζόμενων από τα πειραματικά προφίλ SAXS, χρησιμοποιώντας αφενός μεν δομές από προσομοιώσεις μοριακής δυναμικής, αφετέρου δε μοντέλο άμεσου διαλύτη.

ΜΕΘΟΔΟΙ-ΠΡΟΓΡΑΜΜΑΤΑ

Στο μέρος αυτό, θα αναφερθούν τα χαρακτηριστικά και ο τρόπος λειτουργίας του λογισμικού και των μεθόδων που χρησιμοποιήθηκαν στην ανά χείρας διπλωματική εργασία.

A) ΜΕΘΟΔΟΙ

1. ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ

Μία από τις θεωρητικές μεθόδους με τις οποίες μπορούμε να μελετήσουμε τις πρωτεΐνες, και γενικότερα τα μακρομόρια, είναι οι προσομοιώσεις Μοριακής Δυναμικής (Molecular Dynamics Simulations). Όπως υποδηλώνει και το όνομα, ουσιαστικά πραγματοποιείται μία προσομοίωση των μεταβολών των ιδιοτήτων (θέσεις ατόμων, ταχύτητες, κτλ.) ενός μοριακού συστήματος, ως συνάρτηση του χρόνου. Οι προσομοιώσεις Μοριακής Δυναμικής έγιναν δυνατές χάρη στην ανάπτυξη των ηλεκτρονικών υπολογιστών. Ουσιαστικά, όλη η εργασία γίνεται σε έναν ηλεκτρονικό υπολογιστή με την βοήθεια ειδικών προγραμμάτων. Οι προσομοιώσεις που εκτελούνται στον υπολογιστή, λειτουργούν ως γέφυρα μεταξύ της μικροσκοπικής θεώρησης και του μακρομοριακού κόσμου του εργαστηρίου. Επίσης, λειτουργούν ως γέφυρα μεταξύ θεωρίας και πειράματος [15].

Η Μοριακή Δυναμική σχετίζεται με διαδικασίες που εξαρτώνται από τον χρόνο και που λαμβάνουν χώρα σε μοριακά συστήματα. Κάθε δυναμική διεργασία (π.χ κίνηση) έχει μία χρονική κλίμακα, ένα πλάτος καθώς και ένα ενεργειακό εύρος. Τα μακρομόρια, και συγκεκριμένα οι πρωτεΐνες, εμφανίζουν ένα μεγάλο εύρος χαρακτηριστικών κινήσεων, από τις ατομικές ταλαντώσεις που είναι πολύ γρήγορες (fs) και πολύ εντοπισμένες, έως τις αργές κινήσεις που γίνονται στην κλίμακα

ολόκληρου του μορίου, όπως μία αναδίπλωση (μs-ms). Αρκετές από τις κινήσεις, παίζουν σημαντικό ρόλο στη βιοχημική λειτουργία μίας πρωτεΐνης.

Υπάρχουν δύο κύριες κατηγορίες αναλύσεων Μοριακής Δυναμικής, ανάλογα με το μοντέλο (και κατ' επέκταση τον μαθηματικό φορμαλισμό) που επιλέγεται για να αναπαραστήσει ένα φυσικό σύστημα: Οι αναλύσεις που χρησιμοποιούν τους νόμους της κλασικής μηχανικής και οι αναλύσεις που λαμβάνουν υπ' όψιν τους την κβαντομηχανική φύση του χημικού δεσμού. Προς το παρόν, μόνο η <<κλασική>> Μοριακή Δυναμική, αυτή δηλαδή που χρησιμοποιεί τους νόμους της κλασικής μηχανικής, είναι πρακτικά εφαρμόσιμη σε προσομοιώσεις βιομοριακών συστημάτων [16]. Μια προσομοίωση Μοριακής Δυναμικής αποτελείται από την αριθμητική, βήμα προς βήμα, επίλυση των κλασικών εξισώσεων της κίνησης [15] :

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i \quad \mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} \mathcal{U}$$

Η πρώτη από τις οποίες είναι η γνωστή εξίσωση του 2^{ου} Νόμου του Νεύτωνα, που μας λέει ότι η δύναμη \mathbf{f}_i , ισούται με την μάζα m_i επί την επιτάχυνση (δεύτερη παράγωγος του \mathbf{r}_i ως προς τον χρόνο). Η δεύτερη εξίσωση μας λέει ότι η δύναμη είναι η κλίση της δυναμικής ενέργειας \mathcal{U} συναρτήσει της θέσης \mathbf{r}_i .

Όπως γίνεται αντιληπτό, τα άτομα ενός βιομορίου δέχονται ένα σύνολο δυνάμεων λόγω της θέσης τους, για τον υπολογισμό των οποίων χρησιμοποιούνται συγκεκριμένες μαθηματικές εκφράσεις που περιγράφουν το δυναμικό συναρτήσει της θέσης (συναρτήσεις δυναμικού ή force fields), όπως είναι το AMBER και το CHARMM. Στις προσομοιώσεις της παρούσας εργασίας χρησιμοποιήθηκε η συνάρτηση δυναμικού CHARMM, όπως θα δούμε πιο αναλυτικά παρακάτω. Μια μαθηματική συνάρτηση δυναμικού, περιλαμβάνει εξισώσεις που περιγράφουν την μεταβολή της δυναμικής ενέργειας που οφείλεται α) σε αλληλεπιδράσεις μεταξύ ατόμων που δεν συνδέονται με ομοιοπολικό δεσμό και β) σε αλληλεπιδράσεις ατόμων που συνδέονται με ομοιοπολικό δεσμό. Παρακάτω παρουσιάζονται οι εξισώσεις αυτές: Η εξίσωση:

$$v^{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$$

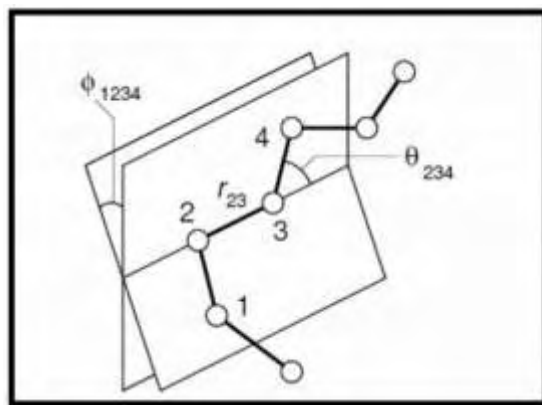
Αφορά στην δυναμική ενέργεια λόγω αλληλεπιδράσεων Van der Waals [17]. Το δυναμικό αυτό καλείται δυναμικό Lennard-Jones. Το σ , είναι η απόσταση κατά την οποία η δυναμική ενέργεια μεταξύ δύο ατόμων είναι μηδέν. Το ϵ έχει διαστάσεις ενέργειας. Το r είναι η απόσταση μεταξύ δύο ατόμων. Η εξίσωση:

$$v^{\text{Coulomb}}(r) = \frac{Q_1 Q_2}{4\pi\epsilon_0 r}$$

αφορά στην δυναμική ενέργεια που οφείλεται στις ηλεκτροστατικές αλληλεπιδράσεις, και οι οποίες διέπονται από τον νόμο του Coulomb. Τα Q_1 και Q_2 είναι φορτία δύο ατόμων, το r είναι η απόσταση μεταξύ των ατόμων, ενώ το ϵ_0 , είναι η διηλεκτρική σταθερά του κενού και το π είναι η γνωστή σταθερά που ισούται με 3,1416. Τέλος, η εξίσωση:

$$\begin{aligned} U_{\text{intramolecular}} = & \frac{1}{2} \sum_{\text{bonds}} k_{ij}^r (r_{ij} - r_{\text{eq}})^2 \\ & + \frac{1}{2} \sum_{\text{bend angles}} k_{ijk}^\theta (\theta_{ijk} - \theta_{\text{eq}})^2 \\ & + \frac{1}{2} \sum_{\text{torsion angles}} \sum_m k_{ijkl}^{\phi, m} (1 + \cos(m\phi_{ijkl} - \gamma_m)) \end{aligned}$$

Αφορά στην δυναμική ενέργεια που οφείλεται στις αλληλεπιδράσεις ατόμων που συνδέονται με ομοιοπολικό δεσμό. Το δυναμικό αυτό αποτελείται από τρεις όρους: Ο πρώτος όρος περιγράφει την δυναμική ενέργεια λόγω απομάκρυνσης των ατόμων από την θέση ισορροπίας σε ομοιοπολικό δεσμό. Ο δεύτερος όρος αναφέρεται στην δυναμική ενέργεια λόγω κάμψεων και ο τρίτος στην δυναμική ενέργεια λόγω στρέψεων. Όλα τα παραπάνω γίνονται κατανοητά με την βοήθεια της **εικόνας 6**:



Εικόνα 6

Το r_{23} συμβολίζει την απόσταση μεταξύ των ατόμων 2 και 3, τα οποία είναι συνδεδεμένα με ομοιοπολικό δεσμό. Η θ_{234} συμβολίζει την γωνία μεταξύ των φορέων 2 διαδοχικών ομοιοπολικών δεσμών. Τέλος, η ϕ_{1234} συμβολίζει την γωνία στρέψης [15]. Συνοψίζοντας λοιπόν, ένα πρόγραμμα προσομοιώσεων Μοριακής Δυναμικής, χρησιμοποιεί μια συνάρτηση δυναμικού και υπολογίζει από τις τρέχουσες συντεταγμένες των ατόμων σε κάθε βήμα της προσομοίωσης και για κάθε άτομο, την συνισταμένη δύναμη που ασκείται πάνω του. Από τις δυνάμεις υπολογίζονται οι επιταχύνσεις και ακολούθως οι νέες θέσεις και ταχύτητες για χρονικό βήμα Δt (συνήθως 2 fs). Η διαδικασία αυτή συνεχίζεται για χιλιάδες ή εκατομμύρια βήματα.

B) ΠΡΟΓΡΑΜΜΑΤΑ

1. ΠΡΟΓΡΑΜΜΑ ΜΟΡΙΑΚΗΣ ΑΠΕΙΚΟΝΙΣΗΣ (VMD)

Το VMD [20] είναι ένα πρόγραμμα γραφικών, σχεδιασμένο για την διαδραστική οπτικοποίηση και ανάλυση των βιοπολυμερών, όπως πρωτεΐνες, νουκλεϊκά οξέα, λιπίδια, αλλά και μεμβράνες. Το VMD τρέχει σε όλα τα είδη Unix (όπως το linux), σε Apple MacOS X, και σε Microsoft Windows. Το πρόγραμμα αυτό μπορεί να φορτώσει τροχιές συντεταγμένων ατόμων, από προγράμματα μοριακών προσομοιώσεων όπως τα AMBER, Charmm, DLPOLY, Gromacs, MMTK, NAMD, X-PLOR, και πολλά άλλα. Το VMD ενσωματώνει διερμηνευτές (interpreters) για τις γλώσσες scripting Python και Tcl, με την βοήθεια των οποίων μπορούμε να επεξεργασθούμε τα δεδομένα (θέσεις ατόμων και ταχύτητες) που προκύπτουν από τις προσομοιώσεις, ως συνάρτηση του χρόνου.

2. ΠΡΟΓΡΑΜΜΑ ΜΟΡΙΑΚΗΣ ΔΥΝΑΜΙΚΗΣ (NAMD)

Οι Προσομοιώσεις Μοριακής Δυναμικής (ΠΜΔ), υπολογίζουν τις ατομικές τροχιές με την επίλυση των εξισώσεων κίνησης χρησιμοποιώντας εμπειρικές συναρτήσεις δυναμικής ενέργειας (Force Fields, FF), όπως το πεδίο δυνάμεων CHARMM, που προσεγγίζει τα πραγματικά ατομικά δυναμικά στα συστήματα βιοπολυμερών. Για την διεξαγωγή ΠΜΔ έχουν αναπτυχθεί διάφορα προγράμματα υπολογιστών. Αυτά τα προγράμματα αναπτύχθηκαν αρχικά για σειριακές μηχανές. Η προσομοίωση μεγάλων μορίων, όμως, απαιτεί τεράστια υπολογιστική ισχύ. Ένας τρόπος για να επιτευχθεί η προσομοίωση είναι η αξιοποίηση παράλληλων υπολογιστών. Για τις ανάγκες της μελέτης μας, χρησιμοποιήθηκε το πρόγραμμα NAMD που είναι σχεδιασμένο να λειτουργεί αποτελεσματικά σε παράλληλους επεξεργαστές, για την προσομοίωση μεγάλων μορίων. Το Force Field που χρησιμοποιείται από το NAMD είναι ίδιο με εκείνο που χρησιμοποιείται από τα προγράμματα CHARMM και X-PLOR. Το NAMD ενσωματώνει τον αλγόριθμο Particle Mesh Ewald (PME), ο οποίος λαμβάνει υπόψη όλες τις ηλεκτροστατικές αλληλεπιδράσεις. Αυτός ο αλγόριθμος μειώνει την υπολογιστική πολυπλοκότητα από $O(N^2)$ σε $O(N \log n)$. Ένα ακόμη πλεονέκτημα του NAMD είναι ότι λειτουργεί σε συνεργασία με το VMD. Δηλαδή, ένα σύστημα που υποβάλλεται σε προσομοίωση με το NAMD μπορεί να προβληθεί και να τροποποιηθεί με το VMD [21].

3. ΤΟ ΠΡΟΓΡΑΜΜΑ CRY SOL

Το CRY SOL [22] είναι ένα πρόγραμμα υπολογισμού του προφίλ SAX μακρομορίων σε διάλυμα, όταν οι ατομικές δομές των μακρομορίων αυτών είναι γνωστές. Το πρόγραμμα CRY SOL χρησιμοποιεί το πολυπολικό ανάπτυγμα των πλατών σκέδασης για τον υπολογισμό του σφαιρικού μέσου όρου σκέδασης, λαμβάνοντας υπόψη το κέλυφος ενυδάτωσης.

Ως δεδομένο εισόδου (input) μπορεί να χρησιμοποιηθεί ένα αρχείο PDB με δομή ακτίνων-X ή NMR μίας πρωτεΐνης ή ενός συμπλόκου πρωτεΐνης-DNA (RNA).

Η ένταση της σκέδασης ενός σωματιδίου σε ένα διάλυμα δίνεται από τον τύπο:

$$I(s) = \left\langle \left| A_a(s) - \rho_s A_s(s) + \delta\rho_b A_b(s) \right|^2 \right\rangle_{\Omega}$$

$A_a(s)$: Η ατομική σκέδαση στο κενό.



$A_s(s)$: Η σκέδαση από το excluded volume.



$A_b(s)$: Η σκέδαση του κελύφους ενυδάτωσης.



ρ_s : Η ηλεκτρονιακή πυκνότητα του διαλύτη ($e/\text{\AA}^3$).

$\delta\rho_b$: Η αντίθεση του κελύφους ενυδάτωσης ($e/\text{\AA}^3$).

Με δεδομένες τις ατομικές συντεταγμένες μπορούμε λοιπόν να προβλέψουμε:

- Την καμπύλη σκέδασης του διαλύματος.
- Μας βοηθά να χαρακτηρίσουμε την πειραματική καμπύλη σκέδασης, χρησιμοποιώντας μόνο δύο ελεύθερες παραμέτρους :
 1. Τον μέσο όρο του εκτοπισμένου διαλύτη ανά ατομική ομάδα.
 2. Την αντίθεση του κελύφους ενυδάτωσης.

4. ΠΡΟΓΡΑΜΜΑ PRIMUS

Το πρόγραμμα PRIMUS [23], μας επιτρέπει να πραγματοποιήσουμε χειρισμούς στα πειραματικά δεδομένων της σκέδασης ακτίνων μικρής γωνίας, όπως: εντοπισμός μέσου όρου μεταξύ δύο σετ, αφαίρεση, συγχώνευση, επέκταση σε μηδενική συγκέντρωση, προσαρμογή καμπύλης καθώς και υπολογισμό από τα Guinier και Porod plots, παραμέτρων όπως: ακτίνα περιστροφής (για επίπεδα, ραβδοειδή και σφαιροειδή σωματίδια), όγκο Porod, ένταση για $s = 0$, και μοριακό βάρος. Το PRIMUS έχει γραφτεί σε Intel Fortran v.11.0 με βιβλιοθήκη γραφικών QuickWin.

Στο κεφάλαιο που ακολουθεί θα αναλύσουμε την διαδικασία με την οποία μπορούμε ξεκινώντας από ένα αρχείο τροχιάς dcd να καταλήξουμε στον υπολογισμό του προφίλ σκέδασης. Αρχικά λοιπόν, παίρνουμε dcd από τις προσομοιώσεις που γίνανε με το NAMD, οι οποίες είναι τρεις: (i) Προσομοίωση πρωτεϊνικού διαλύματος λυσοζύμης με την βοήθεια του Force Field CHARMM 27, (ii) Προσομοίωση πρωτεϊνικού διαλύματος λυσοζύμης με την βοήθεια του Force Field CHARMM 36, (iii) Μια προσομοίωση μόνο με νερό. Στην συνέχεια, από τις προσομοιώσεις αυτές παίρνουμε κάθε φορά 400 στιγμιότυπα (frames), και σώζουμε ξεχωριστά την πρωτεΐνη με το κέλυφος ενυδάτωσης καθώς και τον αποκλειόμενο όγκο νερού. Τέλος, χρησιμοποιώντας το πρόγραμμα CRY SOL υπολογίζουμε το profil σκέδασης.

ΥΠΟΛΟΓΙΣΤΙΚΟ ΜΕΡΟΣ

Στο πλαίσιο αυτής της εργασίας υπολογίσθηκαν τα προφίλ SAXS της λυσοζύμης από δύο σειρές προσομοιώσεων, μια με το FF CHARMM27 (C27) και μια με το νεότερο CHARMM36 (C36). Χρειάστηκε επίσης και μια τρίτη προσομοίωση με καθαρό νερό. Στα ακόλουθα εδάφια περιγράφεται ο τρόπος δημιουργίας των απαραίτητων αρχείων ενδεικτικά για το FF CHARMM27.

1. ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ ΑΡΧΕΙΟΥ PSF & PDB

Ένα αρχείο pdb, δεν περιέχει καμία πληροφορία για τον τρόπο σύνδεσης των ατόμων, των οποίων τις συντεταγμένες περιέχει. Για να πραγματοποιηθεί όμως μία προσομοίωση, η συγκεκριμένη πληροφορία είναι απαραίτητη. Έτσι, πρέπει να δημιουργηθεί ένα αρχείο psf, το οποίο περιέχει πληροφορίες για τους τύπους των ατόμων που υπάρχουν στο αρχείο παραμέτρων, για την συνδεσμολογία μεταξύ των ατόμων αυτών, καθώς και πληροφορίες για τα φορτία τους. Το αρχείο αυτό μπορεί να δημιουργηθεί από το πρόγραμμα psfgen, με βάση το αρχικό αρχείο pdb και ένα αρχείο τοπολογίας (topology file), το οποίο στην δική μας ανάλυση είναι το top_all27_prot_na.inp. Το αρχείο παραμέτρων που χρησιμοποιήθηκε είναι ένα αρχείο που περιέχει τις σταθερές της συνάρτησης δυναμικού για κάθε τύπο ατόμου και είναι το par_all27_prot_na.inp για πρωτεΐνες και νουκλεϊκά οξέα. Ακολουθώντας τις οδηγίες του tutorial του προγράμματος NAMD, συντάξαμε ένα tcl script, δηλαδή ένα σύνολο εντολών βάσει των οποίων το psfgen δημιουργεί ένα αρχείο psf, καθώς και ένα αρχείο pdb (ΠΑΡΑΡΤΗΜΑ: Ρυθμιστικό script A)

2. ΕΝΥΔΑΤΩΣΗ

Πριν από την ενεργειακή ελαχιστοποίηση (energy minimization), πρέπει να διαμορφωθεί το κατάλληλο περιβάλλον γύρω από το μόριο της πρωτεΐνης ώστε να προσεγγίζει όσο γίνεται πιστότερα την πραγματικότητα. Επειδή το άμεσο περιβάλλον

της πρωτεΐνης που μελετάμε είναι υδατικό, αυτό που πρέπει να κάνουμε είναι να τροποποιήσουμε το μοντέλο του συστήματος προσθέτοντας νερό. Το μοντέλο νερού που χρησιμοποιήσαμε είναι το TIP3. Συνήθως, δημιουργούμε έναν κύβο ή μια σφαίρα στο εσωτερικό των οποίων βρίσκεται η προς μελέτη πρωτεΐνη. Προσθέσαμε 16210 μόρια νερού γύρω από την πρωτεΐνη. Το μέγεθος του κύβου εξαρτάται κυρίως από το μέγεθος της πρωτεΐνης, τον αριθμό των μορίων νερού, αλλά και από το είδος της μελέτης που θέλουμε να διεξαγάγουμε.

3. ΠΡΟΣΘΗΚΗ ΙΟΝΤΩΝ (NEUTRALIZATION / ADD IONS)

Τα αρχεία psf και pdb που λάβαμε από το solvation περιέχουν το σύστημα πρωτεΐνης-νερού. Στις προσομοιώσεις μας χρησιμοποιήθηκε για τον υπολογισμό των ηλεκτροστατικών αλληλεπιδράσεων η μέθοδος PME, η οποία απαιτεί το σύστημα να είναι ηλεκτρικά ουδέτερο. Για τον σκοπό αυτό, προσθέσαμε οκτώ ιόντα Cl⁻ έτσι ώστε να εξασφαλισθεί η ηλεκτρική ουδετερότητα στο διάλυμα. Αυτό έγινε ως εξής: *Extensions Modeling* → *Add Ions* → *Only neutralize* → *Autoionize* ώστε να πάρουμε την λυσοζύμη με νερά και ιόντα.

4. ΕΛΑΧΙΣΤΟΠΟΙΗΣΗ ΕΝΕΡΓΕΙΑΣ (ENERGY MINIMIZATION)

Πριν από την καθαυτό προσομοίωση (εξισορρόπηση) της πρωτεΐνης πρέπει να ελαχιστοποιήσουμε την δυναμική ενέργεια του συστήματος πρωτεΐνης-νερού-ιόντων. Για να τρέξουμε την ελαχιστοποίηση χρειάζεται να συντάξουμε ένα configuration script. Σε αυτό θα καθορίζονται οι συνθήκες στις οποίες θα γίνει η διαδικασία, τα αρχεία που θα διαβάσει το NAMD, καθώς επίσης και ένα αρχείο με τις παραμέτρους (par_all27_prot_na.inp). Πρέπει να προσέξουμε όλα αυτά να περιέχονται μέσα στον φάκελο εργασίας που βρίσκεται και το script. Στο πεδίο coordinates, βάζουμε το αρχείο pdb που προέκυψε από την προσθήκη ιόντων, ενώ στο πεδίο structure, το αντίστοιχο αρχείο psf. Ορίζουμε το όνομα που επιθυμούμε για το αρχείο που θα προκύψει στο πεδίο output name.

Τρέξαμε την ελαχιστοποίηση ενέργειας για 4000 steps. Κατά την διάρκεια της ελαχιστοποίησης ενέργειας το πρόγραμμα namd ενημερώνει περιοδικά ένα αρχείο κειμένου (logfile) που περιέχει τις τρέχουσες ενέργειες καθώς και το "gradient tolerance". Θα έχουμε μία ικανοποιητική ελαχιστοποίηση της ενέργειας του συστήματος όταν αυτός ο δείκτης είναι μονοψήφιος (< 9) και όσο το δυνατόν μικρότερος, κάτι που μπορεί να επιτευχθεί αυξάνοντας τα βήματα.

Όταν τελειώσει η διαδικασία αυτή, θα δημιουργηθούν στο φάκελο τρία νέα αρχεία τύπου: .coor, .vel, xsc, με όνομα αυτό που ορίσαμε σαν output name στο script του minimization. Το αρχείο .coor θα χρησιμοποιηθεί αργότερα για την εξισορρόπηση.

5. ΕΞΙΣΟΡΡΟΠΗΣΗ (EQUILIBRATION)

Μετά την ολοκλήρωση της ελαχιστοποίησης της ενέργειας, προχωρήσαμε στην εξισορρόπηση σε τρία στάδια, ένα σύντομο υπό σταθερό πλήθος ατόμων, σταθερή πίεση και θερμοκρασία (NPT) και δύο διαδοχικά διάρκειας 10ns το καθένα υπό σταθερό πλήθος ατόμων, σταθερό όγκο και θερμοκρασία (NVT). Οι παράμετροι του FF δίνονταν στο αρχείο par_all27_prot_na.inp. Το πρώτο στάδιο χρησίμευσε στην προσαρμογή του όγκου της κυψελίδας προσομοίωσης στις συνθήκες $P = 1 \text{ Atm}$ και $T = 300 \text{ K}$. Το δεύτερο και τρίτο στάδιο αποτέλεσαν την κυρίως προσομοίωση (ΠΑΡΑΡΤΗΜΑ: Ρυθμιστικό script B). Μετά το πέρας της εξισορρόπησης, δημιουργούνται στον φάκελο τέσσερα νέα αρχεία τύπου: .dcd, .coor, .vel, .xsc, και έξι νέα αρχεία τύπου .restart.xsc.old, restart.xsc, .restart.vel.old, restart.vel, .restart.coor.old, .restart.coor, με όνομα αυτό που ορίσαμε σαν output name στο script.

6. ΣΤΑΔΙΑ ΥΠΟΛΟΓΙΣΤΙΚΟΥ ΜΕΡΟΥΣ

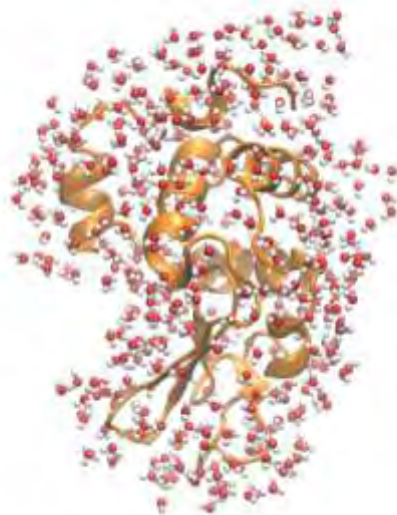
- *1^ο στάδιο:* Αποθήκευση αρχείων pdb από το dcd της προσομοίωσης με CHARMM36

Στο πρώτο στάδιο της διαδικασίας, αποθηκεύσαμε αρχεία από το dcd, σε μορφή pdb, που είναι η μορφή που δέχεται ως εισαγόμενο το πρόγραμμα Crysol. Για να φτάσουμε όμως εκεί, (i) αρχικά ανοίγουμε ένα terminal στο linux, μεταφερόμαστε

στον φάκελο που υπάρχουν τα αρχεία μας (psf, int, dcd), και από εκεί πληκτρολογώντας VMD+Enter ανοίγουμε την κονσόλα του προγράμματος VMD, (ii) File → New Molecule → Determine File Type → Επιλογή psf → Load → Browse → Επιλέγω το dcd eq_min_lyz_NVT_c36_5_400fr_cent.dcd → Load → Extensions → Tk Console → Play save_pdb_from_dcd.tcl (ΠΑΡΑΡΤΗΜΑ: Script Γ) → Παραγωγή αρχείων lyso36_hs3A.pdb, τα οποία περιέχουν την λυσοζύμη με το κέλυφος ενυδάτωσης. Η διαδικασία που περιγράψαμε μόλις, είναι κοινή για την εφαρμογή όλων των tcl script, και διεκπεραιώνεται μέσω του προγράμματος VMD. Αντίθετα, όλα τα bash scripts μπορούν να αξιοποιηθούν με την βοήθεια του Terminal.

- *2^ο στάδιο*

Έπειτα, στα αρχεία που δημιουργήσαμε (lyso36_hs3A.pdb), χρησιμοποιήσαμε ένα 2^ο bash script, με όνομα HSE2HIS_END2TER.sh (ΠΑΡΑΡΤΗΜΑ: Script Δ), το οποίο μετονομάζει σε κάθε ένα από τα αρχεία pdb τα HSE σε HIS, καθώς και τα END σε TER, δημιουργώντας 400 νέα αρχεία με όνομα lyso36-hs3A_HSE2HIS-END2TER_\$.pdb. Οι μετονομασίες αυτές είναι απαραίτητες για την σωστή αναγνώριση των αμινοξέων από το πρόγραμμα CRYSOLE.



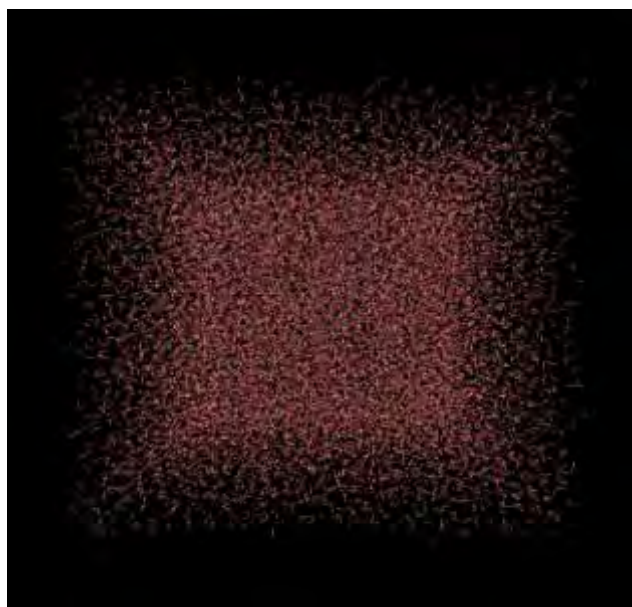
Εικόνα 7: Η πρωτεΐνη με το κέλυφος ενυδάτωσης έκτασης 3 Å.

- **3^ο στάδιο**

Στην συνέχεια ασχοληθήκαμε με την προσομοίωση καθαρού νερού. Η προσομοίωση αυτή χρησιμοποιείται για την παραγωγή αρχείων pdb με το εκτοπιζόμενο από την πρωτεΐνη νερό. Σε πρώτη φάση παρήχθησαν 400 αρχεία με όνομα lyso36-solv-nero_\$i.pdb με τη βοήθεια του script save_solve-pdb_from_dcd.tcl, τα οποία περιέχουν τον κύβο του νερού.

- **4^ο στάδιο**

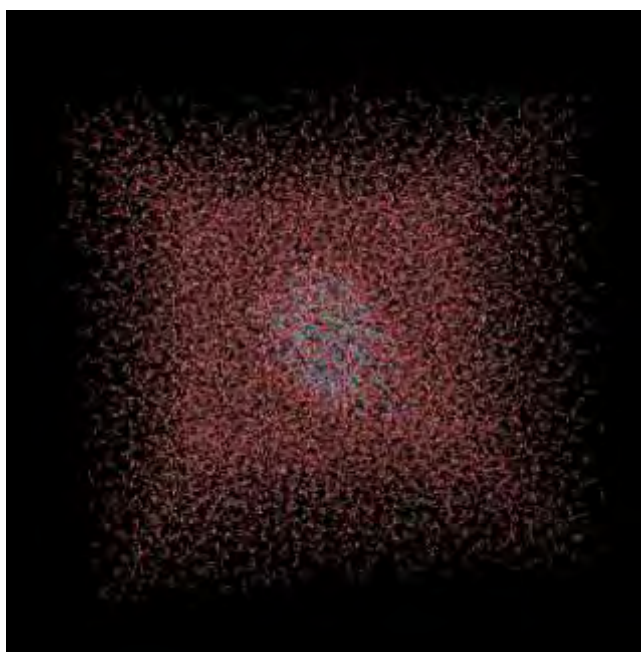
Χρήση bash script με όνομα skip_1st_line-TIP3toNERO.sh (ΠΑΡΑΡΤΗΜΑ: Script ΣΤ), με στόχο την απομάκρυνση της πρώτης σειράς από τα αρχεία solv-lyso.pdb, καθώς και της μετατροπής του TIP3 σε NERO στα αρχεία αυτά. Τα αρχεία που δημιουργήθηκαν στο 4^ο στάδιο έχουν όνομα solv-lyso36_TIP3toNERO_\$i.pdb. Η απομάκρυνση της πρώτης σειράς-γραμμής είναι απαραίτητη για την δημιουργία λειτουργικών αρχείων pdb, κατά την διαδικασία της συνένωσης που ακολουθεί στο επόμενο στάδιο, ενώ η μετονομασία του TIP3 σε NERO μας βοηθά να ξεχωρίσουμε τα μόρια του νερού του κελύφους ενυδάτωσης από τα μόρια του νερού του κύβου.



Εικόνα 8: Ο κύβος του νερού των αρχείων solv-lyso36_\$i.pdb.

- *5^ο στάδιο*

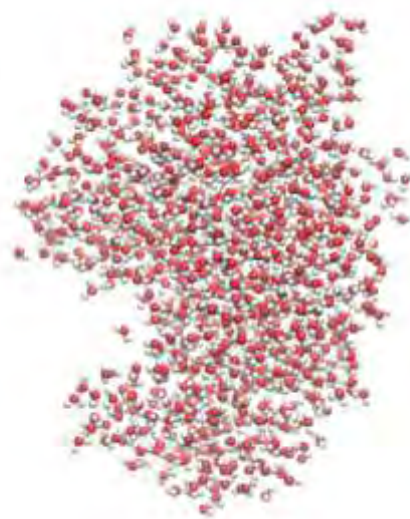
Χρήση 5^{ου} script, και πιο συγκεκριμένη ενός bash script με όνομα concatenate.sh, το οποίο συνενώνει τα αρχεία pdb που δημιουργήθηκαν στο 2^ο και στο 4^ο στάδιο, δηλαδή των αρχείων pdb lyso36-hs3A_HSE2HIS-END2TER_\$.pdb και solv-lyso36_TIP3toNERO_\$.pdb, για την παραγωγή 400 αρχείων pdb με όνομα lyso-sol-nero_\$.pdb. Τα αρχεία αυτά, περιέχουν τόσο την πρωτεΐνη με το κέλυφος ενυδάτωσης, όσο και τον κύβο νερού.



Εικόνα 9: Η λυσοζύμη με το κέλυφος ενυδάτωσης στον κύβο του νερού.

- *6^ο στάδιο*

Χρήση 6^{ου} script με όνομα save_excl_solv.tcl (ΠΑΡΑΡΤΗΜΑ: Script Z), με την βοήθεια του οποίου έχουμε παραγωγή 400 αρχείων με όνομα lyso36-excl_3A_\$.pdb, που περιέχουν μόνο το excluded volume, δηλαδή των όγκο του νερού που αποκλείεται από την λυσοζύμη.



Εικόνα 10: Ο όγκος του νερού που αποκλείεται από την λυσοζύμη.

- *7^ο στάδιο*

Στο στάδιο αυτό λοιπόν, έχοντας υπολογίσει πλέον τόσο τον όγκο του νερού που αποκλείεται από την λυσοζύμη, όσο και την πρωτεΐνη με το κέλυφος ενυδάτωσης, κάνουμε εφαρμογή του προγράμματος CRY SOL (ΠΑΡΑΡΤΗΜΑ: DOS script H) στα αρχεία αυτά, έτσι ώστε να δημιουργηθούν 400 αρχεία int και 400 αρχεία alm, δηλαδή να υπολογίσουμε την ένταση του σκέτους νερού, αλλά και τις εντάσεις στην πρωτεΐνη. Οι παράμετροι που χρησιμοποιήθηκαν για την εφαρμογή του CRY SOL φαίνονται στο log file που ακολουθεί:

```

*** ----- ***
***   C R Y S O L   Wintel/UNIX/Linux version 2.8.3   ***
*** Please reference: D.Svergun, C.Barberato           ***
***   & M.H.J.Koch (1995) J. Appl.Cryst., 28, 768-773 ***
***   Version (LMAX=50) for small and wide angles     ***
***           Last revised --- 12/08/15 14:00         ***
***   Copyright (c) ATSAS Team                        ***
***   EMBL, Hamburg Outstation, 1995 - 2015          ***
*** ----- ***

```

```

-----
                Program options :
0 - evaluate scattering amplitudes and envelope
1 - evaluate only envelope and Flms
2 - read CRY SOL information from a .sav file

```


- *8^ο στάδιο*

Με την βοήθεια του προγράμματος `almavr.exe`, μπορούμε τώρα να υπολογίσουμε την μέση τιμή των εντάσεων, τόσο της λυσοζύμης με το κέλυφος ενυδάτωσης, όσο και του διαλύτη-buffer ξεχωριστά. Για να υλοποιηθεί αυτό, αρχικά πήραμε τα 400 αρχεία με όνομα `lyso36-excl_3A_$i.pdb` και με την βοήθεια του `run.almavr.bat` υπολογίστηκε η μέση τιμή των συγκεκριμένων εντάσεων (`out-av-excl.alm`). Αντίστοιχα πραγματοποιήθηκε η ίδια διαδικασία στα αρχεία `lyso36-hs3A_$i.pdb`, παίρνοντας το `out-av-lyso36.alm`

- *9^ο στάδιο*

Τέλος, χρησιμοποιήσαμε το `subalm.exe`, έτσι ώστε να γίνει η αφαίρεση μεταξύ της μέσης τιμής των εντάσεων του διαλύτη από την μέση τιμή των εντάσεων της λυσοζύμης με το κέλυφος ενυδάτωσης, ώστε να υπολογισθεί η αντίστοιχη ένταση. Με άλλα λόγια, στο στάδιο αυτό χρησιμοποιήσαμε τα `out-av-excl.alm` και `lyso36-hs3A.alm` που βρήκαμε στο προηγούμενο στάδιο και τα αφαιρέσαμε.

Όλες οι διαδικασίες που αναπτύχθηκαν στην ενότητα αυτήν πραγματοποιήθηκαν μία φορά με το Force Field CHARMM 36 και μια με το Force Field CHARMM 27. Η ανάλυση των αποτελεσμάτων ακολουθεί στην ενότητα που ακολουθεί.

ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ

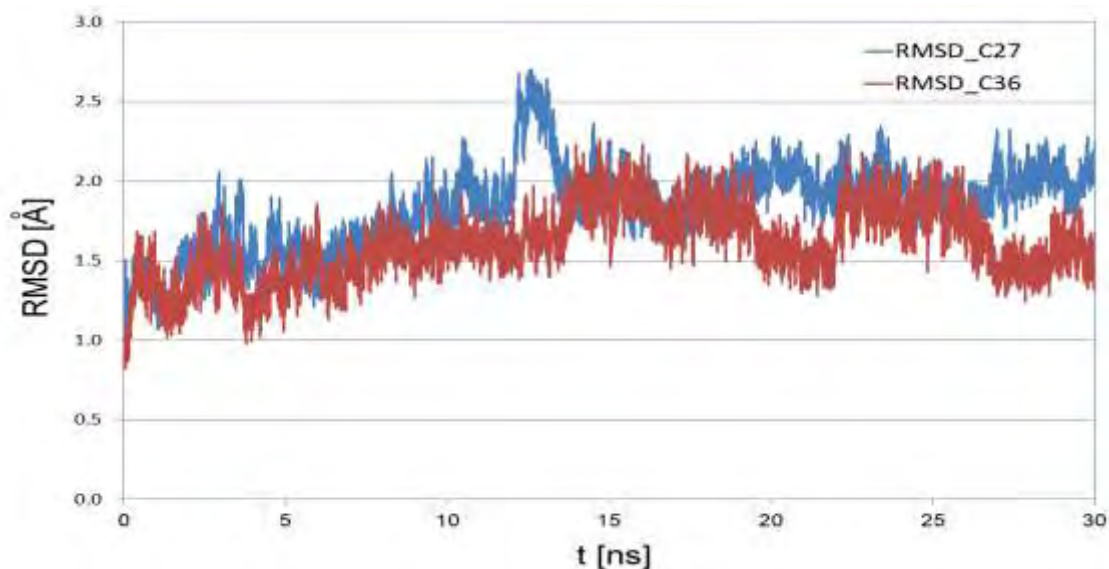
1. ΥΠΟΛΟΓΙΣΜΟΣ ΚΑΙ ΑΝΑΛΥΣΗ RMSD

Μια από τις παραμέτρους που χρησιμοποιήθηκε ως μέτρο της μεταβολής της δομής σε σχέση με την αρχική δομή της λυσοζύμης συναρτήσει του χρόνου, είναι το RMSD (root mean square deviation). Μπορούμε σε ένα διάγραμμα, να αναπαραστήσουμε την μεταβολή του RMSD για κάθε στιγμιότυπο της προσομοίωσης.

$$RMS = \left\langle \left(r_i^\alpha - r_i^\beta \right)^2 \right\rangle^{\frac{1}{2}} = \sqrt{\frac{1}{N_i} \sum_i \left(r_i^\alpha - r_i^\beta \right)^2}$$

Εξισ.1: Εδώ r_i^α είναι η θέση του ατόμου i στην δομή α , ενώ το r_i^β είναι η θέση του ίδιου ατόμου στην δομή β και N είναι το πλήθος των ατόμων.

Υπολογίσαμε την χρονική εξέλιξη του (Μέση τετραγωνική απόκλιση των ατομικών θέσεων) με την βοήθεια του "RMSD Trajectory Tool" του VMD, τόσο με το FF CHARMM27, όσο και με το CHARMM36. Τα αποτελέσματα φαίνονται **στην εικόνα 11:**



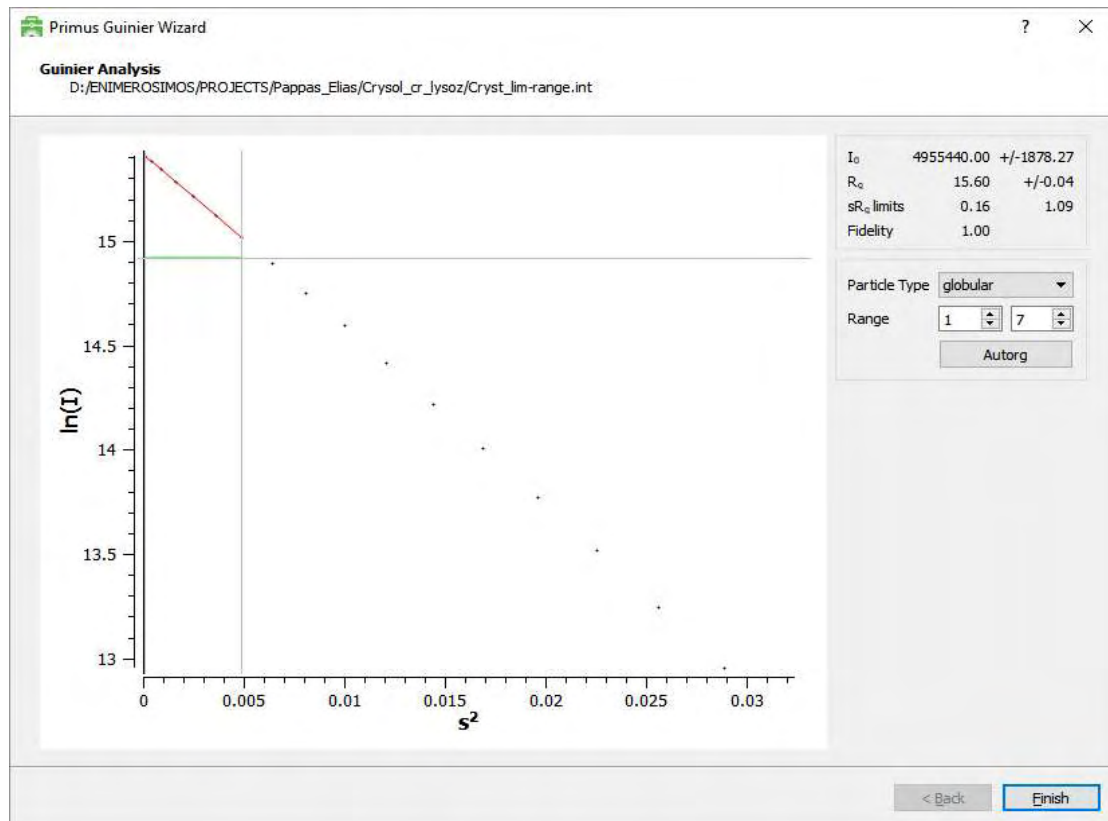
Εικόνα 11: Σύγκριση RMSD μεταξύ του C27 (μπλέ καμπύλη) και του C36 (κόκκινη καμπύλη).

Όπως φαίνεται στην εικόνα 11, υπάρχει μια διαφορά μεταξύ των δύο Force Fields, χωρίς όμως να αποκλείεται η σύγκλιση εάν συνεχίζαμε την διαδικασία για περισσότερα από 30ns.

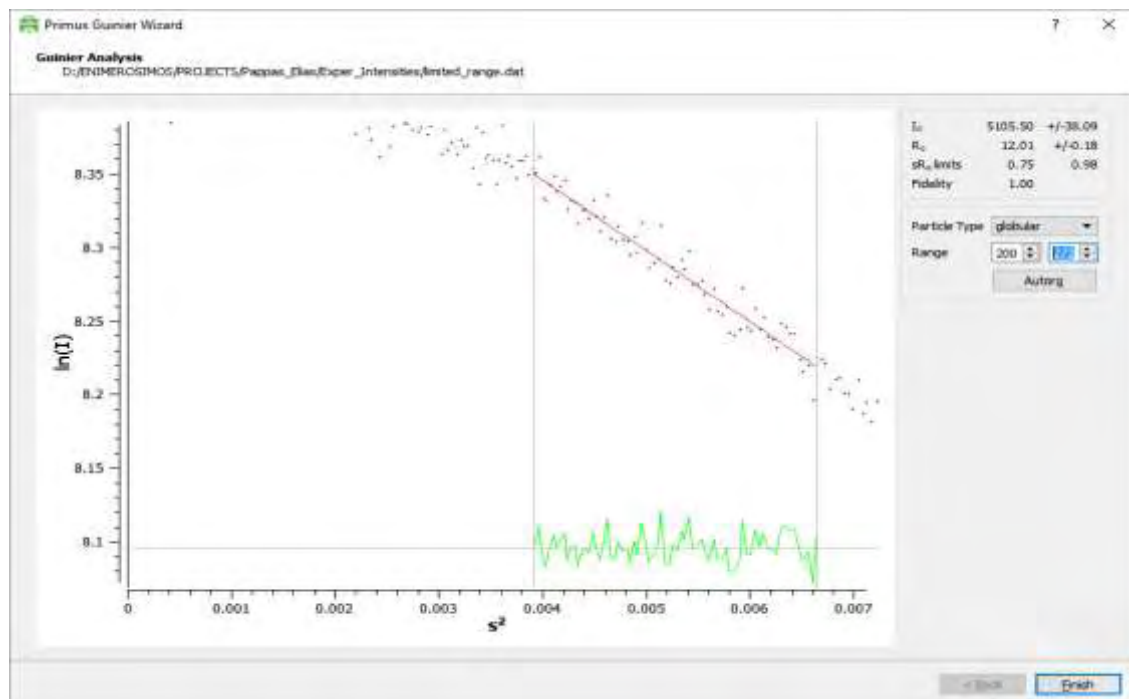
2. ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΓΥΡΟΣΚΟΠΙΚΗΣ ΑΚΤΙΝΑΣ

Η γυροσκοπική ακτίνα (R_g) ορίζεται ως η μέση απόσταση των στοιχειωδών μονάδων μιας πολυμερικής αλυσίδας από το κέντρο βάρους της. Στην παρούσα διπλωματική εργασία προχωρήσαμε στον υπολογισμό της R_g με τη χρήση ειδικών προγραμμάτων (Primus, VMD), ώστε να γίνει σύγκριση μεταξύ πειραματικής - CRYSOLE κρυσταλλικής - C36 - C27 τιμής.

Ξεκινήσαμε με τον υπολογισμό της γυροσκοπικής ακτίνας για το Force Field C27 και το C36 με το πρόγραμμα Primus. Με την βοήθεια του Primus, οι τιμές του R_g της λυσοζύμης για C27 και C36 υπολογίστηκαν ότι είναι 15,2 Å και 14,95 Å αντίστοιχα. Στην συνέχεια κάναμε χρήση του προγράμματος VMD με τις τιμές που προέκυψαν να είναι 14,45 Å για το C27 και 14,46 Å για το C36. Συμπερασματικά λοιπόν, τα δύο αυτά Force Field δεν παρουσιάζουν διαφορά ως προς το R_g . Ακολούθησε ο υπολογισμός του R_g από πειραματική και κρυσταλλική δομή με την βοήθεια του προγράμματος Primus. Η τιμές του R_g της πειραματικής ήταν 12,0 Å ενώ αυτή της κρυσταλλικής δομής 15,6 Å. Στις εικόνες που ακολουθούν μπορούμε να πάρουμε μια γεύση από το περιβάλλον του προγράμματος Primus, αλλά και να δούμε οπτικά το τμήμα της καμπύλης από το οποίο υπολογίζεται η γυροσκοπική ακτίνα (κόκκινο τμήμα της καμπύλης).



Εικόνα 12: Υπολογισμός της γυροσκοπικής ακτίνας της κρυσταλλικής δομής με την χρήση του προγράμματος Primus και την βοήθεια της ανάλυσης Guinier.

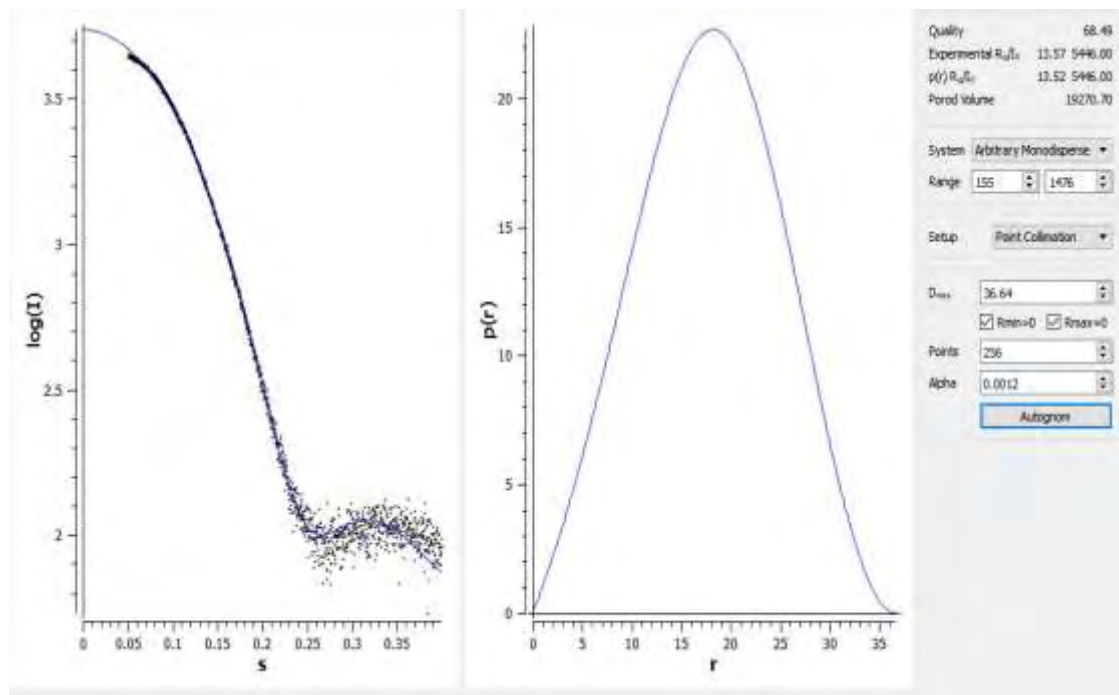


Εικόνα 13: Υπολογισμός της γυροσκοπικής ακτίνας από την πειραματική με το πρόγραμμα Primus και την βοήθεια της ανάλυσης Guinier.

3. ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ D_{max}

Ως D_{max} ορίζουμε την μέγιστη απόσταση μεταξύ δύο ατόμων στο μόριο της πρωτεΐνης. Στην παρούσα μελέτη υπολογίσαμε την D_{max} της πειραματικής - CRY SOL κρυσταλλικής - C36 - C27 τιμής, με τρόπο παρόμοιο με αυτόν που αναλύσαμε πριν. Οι τιμές που πήραμε για την πειραματική - CRY SOL κρυσταλλική - C36 - C27 ήταν 36,64 Å, 47,1 Å, 44,75 Å και 48,32 Å αντίστοιχα. Από τα αποτελέσματα που λάβαμε παρατηρούμε ότι η πειραματική τιμή του D_{max} αποκλίνει από την τιμή των προσομοιώσεων και από αυτήν που προκύπτει από την κρυσταλλική δομή. Υπάρχει μια μαθηματική έκφραση που σχετίζει την συνάρτηση κατανομής των ενδομοριακών αποστάσεων $P(r)$, από όπου προκύπτει το D_{max} , με τον αντίστροφο μετασχηματισμό Fourier της καμπύλης της έντασης. Η μεγάλη λοιπόν απόκλιση της πειραματικής D_{max} από τις υπόλοιπες που λάβαμε μπορεί να είναι αποτέλεσμα ενός εκ των παρακάτω λόγων:

1. Το fit που πραγματοποιεί το Primus δεν είναι ακριβές, αφού έχουμε σφάλματα στις πειραματικές εντάσεις.
2. Η ποιότητα του διαλύματος λυσοζύμης που χρησιμοποιήθηκε για τις μετρήσεις της μελέτης μας δεν ήταν ικανοποιητική-
3. Ο αριθμός μετρήσεων του διαλύματος της λυσοζύμης ήταν περιορισμένος.
4. Η έκταση του κελύφους ενυδάτωσης μπορεί να πρέπει να θεωρηθεί μεγαλύτερη (> 3 Å)

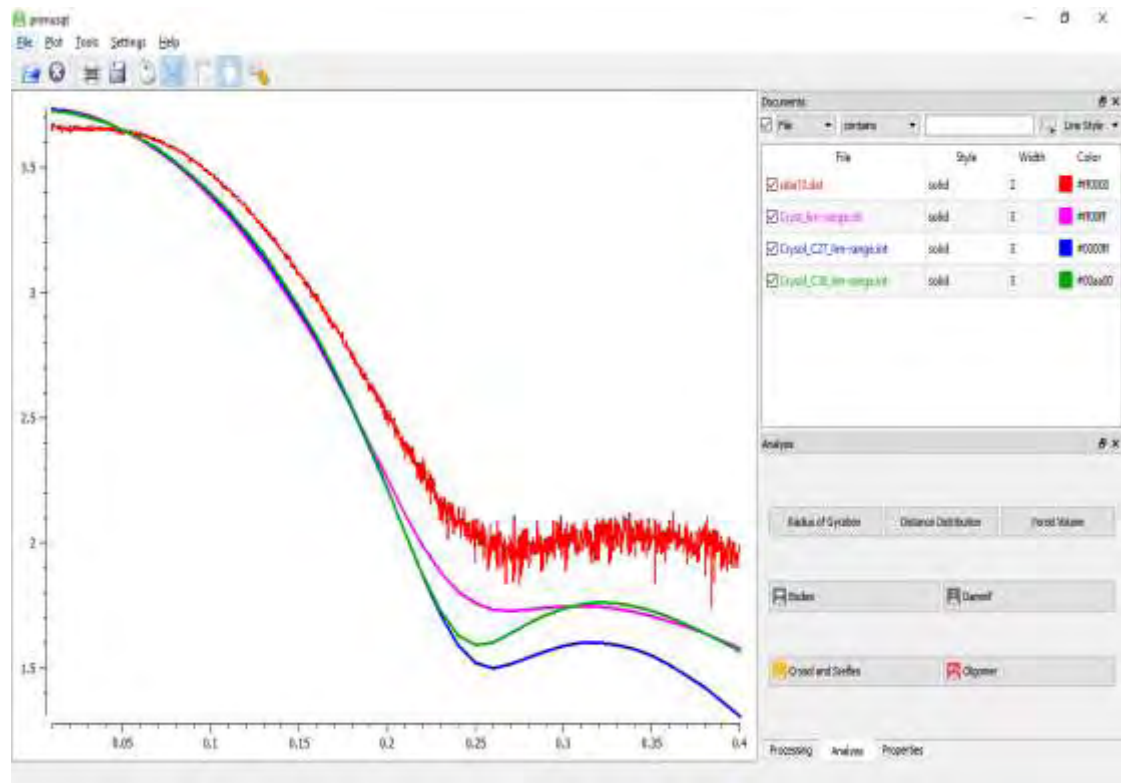


Εικόνα 14: Υπολογισμός της πειραματικής τιμής του D_{max} . Ο υπολογισμός της τιμής αυτής έγινε με την βοήθεια του προγράμματος Primus, το οποίο χρησιμοποιεί για την ανάλυση το AUTOGNOM.

Πίνακας 2: Οι τιμές για το R_g και το D_{max} της λυσοζύμης όπως προσδιορίστηκαν με το πρόγραμμα Primus από πειραματικό προφίλ SAXS καθώς και εκείνες που προέκυψαν από ΠΜΔ με το πρόγραμμα VMD και τα force fields C27 και το C36.

	R_g [Å]	D_{max} [Å]
Πειραματική	12	36,6
Κρυσταλλική δομή	15,6	47,1
Προσομοίωση με C27	15,2	48,3
Προσομοίωση με C36	15,0	44,8

4. ΣΥΓΚΡΙΣΗ ΚΑΜΠΥΛΩΝ ΤΩΝ ΕΝΤΑΣΕΩΝ ΣΚΕΛΑΣΗΣ



Εικόνα 15: Προφίλ SAXS στο πρόγραμμα Primus. (Κόκκινο-πειραματική, μωβ- CRYSOLO με κρυσταλλική, μπλε-CRYSOLO με CHARMM27, πράσινο-CRYSOLO με CHARMM36).

Στην εικόνα 15 παρουσιάζεται το προφίλ SAXS από το πείραμα (κόκκινο), από την κρυσταλλική δομή υπολογισμένο με το Crysol, όπου το νερό θεωρείται συνεχές μέσο (μωβ), από 400 δομές των προσομοιώσεων με FF C27 (μπλε) και από 400 δομές των προσομοιώσεων με FF C36 (πράσινο). Είναι φανερό ότι η προσπάθεια να βελτιωθεί η προσέγγιση του υπολογιζόμενου από το Crysol προφίλ λαμβάνοντας υπόψη την ποικιλία των δομών σε διάλυμα καθώς και την άμεση παρουσία των μορίων νερού με την βοήθεια των προσομοιώσεων μοριακής δυναμικής δεν απέδωσε καρπούς. Το FF C36 προσεγγίζει λίγο καλύτερα το πειραματικό προφίλ από ό,τι το FF C27, αλλά παραμένει χειρότερο από αυτό της κρυσταλλικής δομής. Εστιάζοντας στο μέρος του προφίλ κοντά στο $s = 0$, παρατηρούμε σύγκλιση όλων των υπολογισμένων προφίλ τα οποία όμως αποκλίνουν από το πειραματικό. Το τελευταίο δείχνει πολύ μικρή αρνητική κλίση στην περιοχή αυτή, κάτι το οποίο δεν είναι αναμενόμενο. Αυτό

μάλλον δηλώνει ότι η ποιότητα του διαλύματος λυσοζύμης στο οποίο έγινε η μέτρηση SAXS δεν ήταν πολύ καλή.

Εν κατακλείδι, μια μελλοντική εκτενέστερη μελέτη υπολογισμού των προφίλ SAXS από προσομοιώσεις ΜΔ θα πρέπει να χρησιμοποιήσει καλύτερης ποιότητας πειραματικό προφίλ SAXS αναφοράς που θα προκύπτει ως μέση τιμή από περισσότερες μετρήσεις. Επίσης, θα πρέπει να δοκιμασθούν κελύφη ενυδάτωσης μεγαλύτερης έκτασης ($> 3 \text{ \AA}$).

ΠΑΡΑΡΤΗΜΑ

SCRIPTS

Στο μέρος αυτό, αναφέρονται ενδεικτικά ορισμένα από τα bash και tcl script που χρησιμοποιήθηκαν για να αυτοματοποιηθούν οι εργασίες που πραγματοποιήθηκαν στην παρούσα διπλωματική εργασία.

A) Ρυθμιστικό script για το πρόγραμμα psfgen

```
#package require psfgen
topology top_all127_prot_na.inp
#topology sp_GOL_4RN5.rtf

alias residue HIS HSE

alias residue HOH TIP3
alias atom TIP3 O OH2

segment A {
pdb packmol_2LZT.pdb
}
patch DISU A:6 A:127
patch DISU A:30 A:115
patch DISU A:64 A:80
patch DISU A:76 A:94
coordpdb packmol_2LZT.pdb A
guesscoord

segment W {
pdb packmol_TIP3W.pdb
}
coordpdb packmol_TIP3W.pdb W

#guesscoord

segment Y {
pdb packmol_TIP3Y.pdb
}
coordpdb packmol_TIP3Y.pdb Y

#guesscoord
```

```

segment C {
pdb packmol_CLA.pdb
}
coordpdb packmol_CLA.pdb C

#guesscoord

writepdb lyz-wat.pdb
writepsf lyz-wat.psf
resetpsf

```

B) Ρυθμιστικό script για την εξισορρόπηση

```

structure          lyz-wat.psf
coordinates        min_lyz-wat.coor
bincoordinates    eq_lyz-wat_NPT_4.coor
set temperature    300
set outputname    eq_lyz-wat_NVT_0
firsttimestep     0

# Input
paraTypeCharmm    on
parameters        par_all27_prot_na.inp
#parameters       3SP3_SUC.par
temperature       $temperature

# Force-Field Parameters
exclude           scaled1-4
1-4scaling        1.0
cutoff            14.0
switching         on
switchdist        12.0
pairlistdist      16.5

# Integrator Parameters
timestep          2.0 ;# 2fs/step
rigidBonds        all ;# needed for 2fs steps
nonbondedFreq     1
fullElectFrequency 2
stepspercycle     10

# Constant Temperature Control
langevin          on ;# do langevin dynamics
langevinDamping   5 ;# damping coefficient (gamma) of
5/ps
langevinTemp      $temperature
langevinHydrogen  off ;# don't couple langevin bath to
hydrogens

# Periodic Boundary Conditions
cellBasisVector1  80.0  0.  0.

```

```

cellBasisVector2      0.    80.0   0.
cellBasisVector3      0.     0.   80.0
cellOrigin             -0.106395 -0.0359 -0.0083
wrapWater              on
wrapAll                on

# PME (for full-system periodic electrostatics)
PME                    yes
PMEGridSizeX          84
PMEGridSizeY          84
PMEGridSizeZ          84

# Output
outputName             $outputname

restartfreq            200      ;# 500steps = every 1ps
dcdfreq               500
xstFreq               500
outputEnergies        100
outputPressure        100

run      5000000      ;# 10 ns

```

Γ) Tcl script για αποθήκευση αρχείων από το dcd, σε μορφή pdb (αρχεία πρωτεΐνης και κελύφους ενυδάτωσης).

```

set nf [molinfo top get numframes]
for {set i 0} {$i < $nf} {incr i} {
  [atomselect top "protein or same residue as (water or resname
  SOD) and within 3 of protein" frame $i] writepdb lyso36-
  hs3A_$i.pdb
}

```

Δ) Bash script για μετονομασία του HSE σε HIS και του END σε TER σε αρχεία pdb.

```

for i in {0..399}
do
  sed 's/HSE/HIS/g' lyso36-hs3A_$i.pdb > lyso36-hs3A_$i_HIS.pdb
  sed 's/END/TER/g' lyso36-hs3A_$i_HIS.pdb > lyso36-hs3A_HSE2HIS-
  END2TER_$i.pdb
  rm lyso36-hs3A_$i_HIS.pdb
done

```

E) Script σε Tcl και bash για την αποθήκευση αρχείων από το dcd, σε μορφή pdb (αρχεία με νερό)

```

set nf [molinfo top get numframes]
for {set i 0} {$i < $nf} {incr i} {
[atomselect top "water" frame $i] writepdb solv-lyso36_$i.pdb
}

for i in {0..399}
do
sed '1d' solv-lyso36_$i.pdb > nsolv-lyso36_$i.pdb
sed 's/TIP3/NERO/g' nsolv-lyso36_$i.pdb > solv-lyso36_TIP3toNERO_$i.pdb
rm nsolv-lyso36_$i.pdb
done

```

ΣΤ) Bash script για απομάκρυνση της πρώτης σειράς από τα αρχεία pdb και μετατροπή TIP3 σε NERO.

```

#!/bin/bash
pathlyso='/home/hlias/Desktop/lyso_c36'
pathsolv='/home/hlias/Desktop/lyso_c36'

for i in {0..399}
do
# wat_TIP3toNERO_$i.pdb after deleting first line from
wat_$i.pdb and replacing TIP3 with NERO
cat $pathlyso/lyso36-hs3A_HSE2HIS-END2TER_$i.pdb
$pathlyso/solv-lyso36_TIP3toNERO_$i.pdb > $pathlyso/lyso36-
solv-nero_$i.pdb
done

```

Z) Tcl script για την παραγωγή αρχείων pdb του εκτοπιζόμενου από την λυσοζύμη νερού

```

for {set i 0} {$i < 400} {incr i} {
mol load pdb lyso36-solv-nero_$i.pdb
set excl [atomselect top "same residue as rename NERO and
within 3 of protein"]
$excl writepdb lyso36-excl_3A_$i.pdb
mol delete top
}

```

H) DOS script για την εφαρμογή του προγράμματος CRY SOL

```

For %%f in (* .pdb) do (
Crysol0.exe -ns 201 -sm 2.0 -dns 0 -dro 0 -eh -lm 50
%%~nf.pdb > log_%%~nf.txt
)

```

Θ) Tcl script για τον υπολογισμό της γυροσκοπικής ακτίνας με το πρόγραμμα VMD.

```
# Calculates the average Rg over a dcd for protein,
protein_1hs, protein_1hs2hs, 1hs and 1hs2hs

# Start of procedure gyr_radius definition =====
proc gyr_radius {sel} {
  # make sure this is a proper selection and has atoms
  if {[ $\$sel$  num] <= 0} {
    error "gyr_radius: must have at least one atom in
selection"
  }
  # gyration is sqrt( sum((r(i) - r(center_of_mass))^2) / N)
  set com [measure center  $\$sel$  weight mass]
  set sum 0
  foreach coord [ $\$sel$  get {x y z}] {
    set sum [vecadd  $\$sum$  [veclength2 [vecsub  $\$coord$   $\$com$ ]]]
  }
  return [expr sqrt( $\$sum$  / ( $[\$sel$  num] + 0.0))]
}

# Start of procedure lavg definition =====
proc lavg L {expr ([join  $\$L$  +])/[llength  $\$L$ ].}

# Start of procedure mean2 definition =====
proc mean2 L {
  set sum 0
  foreach i  $\$L$  {set sum [expr { $\$sum$ + $\$i$ * $\$i$ }] }
  expr {double( $\$sum$ )/[llength  $\$L$ ]}
}

# Start of procedure stddev definition =====
proc stddev L {
  set m [lavg  $\$L$ ]
  expr {sqrt([mean2  $\$L$ ]- $\$m$ * $\$m$ )}
}

# End of procedure definition =====

mol load psf run.psf dcd no_field_last_5000fr.dcd
# Read cell dimensions for wrap.
package require pbctools
pbc readxst res1.xst -molid top
set outfile1 [open list_Rg_no_field_last_1ns.txt w]
set outfile2 [open aver_Rg_no_field_last_1ns.txt w]
set nf [molinfo top get numframes]

set rgpList {}
set rgp1hsList {}
set rgp1hs2hsList {}
set rglhsList {}
#set rglhs2hsList {}
```

```

# Rg calculation loop
for {set i 0 } { $i < $nf } { incr i } {
# Wrap all at the center of mass
  pbc wrap -centersel "protein" -center com -compound
  residue -all
# Select atoms
  set selp [atomselect top "protein" frame $i]
  set sel $selp
  set pRg [gyr_radius $sel]
  lappend rgpList $pRg
  unset sel
  unset selp
  set selplhs [atomselect top "protein or same residue as
water and within 3 of protein" frame $i]
  set sel $selplhs
  set plhsRg [gyr_radius $sel]
  lappend rgplhsList $plhsRg
  unset selplhs
  unset sel
  set selplhs2hs [atomselect top "protein or same residue as
water and within 6 of protein" frame $i]
  set sel $selplhs2hs
  set plhs2hsRg [gyr_radius $sel]
  lappend rgplhs2hsList $plhs2hsRg
  unset sel
  unset selplhs2hs
  set sellhs [atomselect top "same residue as water and within
3 of protein" frame $i]
  set sel $sellhs
  set lhsRg [gyr_radius $sel]
  lappend rglhsList $lhsRg
  unset sel
  unset sellhs
  set sellhs2hs [atomselect top "same residue as water and
within 6 of protein" frame $i]
  set sel $sellhs2hs
  set lhs2hsRg [gyr_radius $sel]
  lappend rglh2hsList $lhs2hsRg
  unset sel
  unset sellhs2hs
puts $outfile1 "$i $pRg $plhsRg $plhs2hsRg $lhsRg $lhs2hsRg"
}

#set avpRg [expr ([join $rgpList +])/[llength $rgpList].]
set L $rgpList
set avpRg [lavg $L]
set devpRg [stddev $L]
puts [llength $L]
puts $avpRg
set txt1 "<Prot Rg> ="
set txt2 "StDev ="
puts $outfile2 "$txt1 $avpRg $txt2 $devpRg"
set rgpList {}
set L {}

```

```

set L $rgplhsList
set avplhsRg [lavg $L]
set devplhsRg [stddev $L]
puts [llength $L]
puts $avplhsRg
set txt1 "<Prot+S1 Rg> ="
set txt2 "StDev ="
puts $outfile2 "$txt1 $avplhsRg $txt2 $devplhsRg"
set rgplhsList {}
set L {}

set L $rgplhs2hsList
set avplhs2hsRg [lavg $L]
set devplhs2hsRg [stddev $L]
puts [llength $L]
puts $avplhs2hsRg
set txt1 "<Prot+S1+S2 Rg> ="
set txt2 "StDev ="
puts $outfile2 "$txt1 $avplhs2hsRg $txt2 $devplhs2hsRg"
set rgplhs2hsList {}
set L {}

set L $rglhsList
set avlhsRg [lavg $L]
set devlhsRg [stddev $L]
puts [llength $L]
puts $avlhsRg
set txt1 "<S1 Rg> ="
set txt2 "StDev ="
puts $outfile2 "$txt1 $avlhsRg $txt2 $devlhsRg"
set rglhsList {}
set L {}

set L $rglh2hsList
set avlhs2hsRg [lavg $L]
set devlhs2hsRg [stddev $L]
puts [llength $L]
puts $avlhs2hsRg
set txt1 "<S1+S2 Rg> ="
set txt2 "StDev ="
puts $outfile2 "$txt1 $avlhs2hsRg $txt2 $devlhs2hsRg"
set rglhs2hsList {}
set L {}

close $outfile1
close $outfile2

mol delete top

exit

```

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Η αναδίπλωση των πρωτεϊνών: Ένα φαινόμενο με πληθώρα εφαρμογών, Πατρινέλη Α., *Περισκόπιο της Επιστήμης*, τεύχος 216 (Απρίλιος 1998), 70-76
- [2] **Βιοχημεία**. Stryer L., Tymoczko J., Berg J., (2006), 45-83
- [3] **Human lysozyme: sequencing of a cDNA, and expression and secretion by *saccharomyces cerevisiae***. Yoshimura K, Toibana A, Nakahama K, *Biochemical and Biophysical Research Communications* (January 1988). 150(2): 794-801.
- [4] **The human lysozyme gene. Sequence organization and chromosomal localization**. Peters CW, Kruse U, Pollwein R, Grzeschik KH, Sippel AE, *European Journal of Biochemistry* (July 1989). 182(3): 507-516.
- [5] **On a remarkable bacteriolytic element found in tissues and secretions**. Fleming A, *Proceedings of the Royal Society B* (May 1922), 93(653): 306-317
- [6] **Adsorption and reactions of chitinase and lysozyme on chitin**. Skujiņš J, Puķīte A, McLaren AD, *Molecular and Cellular Biochemistry*. (December 1973). 2 (2): 221-228.
- [7] **Protein Misfolding in Human Diseases**, Karin Almstedt, Department of Physics, Chemistry and Biology Linköping University(2009), SE-581 83 Linköping, Sweden Linköping.
- [8] **Primers in Biology**. Petsko, G.A., Ringe, D., *Protein Structure and Function*. New Science Ltd, 2004
- [9] **The New Penguin Dictionary of Biology**, Eighth Edition
- [10] **Crystallography :Made Crystal Clear**. Rodes G., Academic Press, Second Edition
- [11] **Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases** . Lee Whitmore, B. A. Wallace. Department of Crystallography, Birkbeck College, University of London.

- [12] **Nuclear Induction.** Bloch, E., Physical Rev. (1946) 70,460-474
- [13] **Εισαγωγή στη Δομή των πρωτεϊνών.** Branden, C., Tooze, J., (2006) Ακαδημαϊκές εκδόσεις . Δεύτερη έκδοση.
- [14] **Simulated x-ray scattering of protein solutions using explicit-solvent models,** Sanghyun Park, Jaydeep P. Bardhan, Benoit Roux, Lee Makowski, The Journal of Chemical Physics (2009), 130, 134114
- [15] **Computational Soft Matter. From Synthetic polymers to Proteins, Lecture Notes,** Michael P. Allen, Norbert Attig, Kurt Binder, Helmut Grubmuller, Kurt Kremer (Eds.), John von Neumann. Introduction to Molecular Dynamics Simulation. Institute for Computing, Juelich, NIC Series. 2004;23:1-28.
- [16] <http://staff.science.nus.edu.sg/~phywjs/lecture-notes/cpnote3.pdf>
- [17] **Exosite interactions contribute to tension-induced cleavage of von Willebrand factor by the antithrombotic ADAMTS13 metalloprotease, PNAS.** Weiqiang Gao, P. J. Anderson, E. M. Majerus, E. A. Tuley and J. E. Sandler, December 12, 2006; 103:19099-19140.
- [18] **A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins.** Alexey G. Kikhney, Dmitri I. Svergun (14 September 2015) FEBS Letters, Issue 19, PartA, pages 2431-2660.
- [19] **Small-angle scattering for structural biology – Expanding the frontier while avoiding the pitfalls,** David A. Jacques, Jill Trehwella, 2010 Apr; 19(4): 642-657.
- [20] **Visual Molecular Dynamics.** Humphrey W, Dalke A and Schultem K, VMD – J. Molec. Graphics, (1996), vol. 14, pp.33-38
- [21] **Scalable molecular dynamics with NAMD.** Phillips J.C., Braun R., Wang W., Gumbart J., Tajkhorshid E., Villa E., Chipot C, Skeel R.D., Kale L. and Schulten K. Journal of Computational Chemistry (2005), Volume26, Issue 16, pages1781-1802. DOI:10.1002/jcc.20289.
- [22] **CRY SOL-a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates,** D. Svergun, C.Barberato and M. H. J. Koch. European Molecular Biology Laboratory, Hamburg Outstation, Notkestrasse 85, D-22603 Hamburg, J. Appl. Cryst. (1995), 28, 768-773.

[23] **PRIMUS: a Windows PC-based system for small-angle scattering data analysis**, P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch and D. I. Svergun, *J. Appl. Cryst.* (2003). 36, 1277-1282.

[24] **Small-angle scattering studies of biological macromolecules in solution**, D. I. Svergun and M. H. J. Koch, *Rep. Prog. Phys.* 66 (2003) 1735–1782 PII: S0034-4885(03)12688-7