ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ

Π.Μ.Σ. ΜΟΡΙΑΚΗ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΗ ΒΙΟΛΟΓΙΑ ΦΥΤΩΝ - ΠΡΑΣΙΝΗ ΒΙΟΤΕΧΝΟΛΟΓΙΑ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΕΙΔΙΚΕΥΣΗΣ

# Αλληλεπίδραση των πρωτεϊνών LHCX1 και FCP του διατόμου *Phaeodactylum tricornutum*

ΧΡΥΣΑΦΟΥΔΗ ΑΝΘΗ

Ηράκλειο, Οκτώβριος 2020

UNIVERSITY OF CRETE BIOLOGY DEPARTMENT

POSTGRADUATE STUDIES PROGRAMMES

PLANT MOLECULAR AND APPLIED BIOLOGY-GREEN BIOTECHNOLOGY

POSTGRADUATE SPECIALIZATION THESIS

# Interaction of LHCX1 and FCP proteins of the diatom *Phaeodactylum tricornutum*

CHRYSAFOUDI ANTHI

Heraklio, October 2020

**Επιβλέπων Καθηγητής:**

Δασκαλάκης Βαγγέλης, Επίκουρος καθηγητής, Τμήμα Επιστήμης και Τεχνολογίας Περιβάλλοντος, Τεχνολογικό Πανεπιστήμιο Κύπρου

**Τριμελής εξεταστική επιτροπή:**

Δασκαλάκης Βαγγέλης, Επίκουρος καθηγητής, Τμήμα Επιστήμης και Τεχνολογίας Περιβάλλοντος, Τεχνολογικό Πανεπιστήμιο Κύπρου

Κοτζαμπάσης Κυριάκος, Καθηγητής, Τμήμα Βιολογίας, Πανεπιστήμιο Κρήτης

Κοκκινίδης Μιχαήλ, Ομότιμος καθηγητής, Τμήμα Βιολογίας, Πανεπιστήμιο Κρήτης

**Ευχαριστίες**

Αρχικά, ευχαριστώ τον καθηγητή, κύριο Κοτζαμπάση Κυριάκο και τον ομότιμο καθηγητή, κύριο Κοκκινίδη Μιχαήλ, για την πολύτιμη συμβολή τους στην ολοκλήρωση αυτής της εργασίας, ως μέλη της τριμελούς επιτροπής.

Επίσης, νιώθω την ανάγκη να ευχαριστήσω ιδιαίτερα τον επιβλέποντα καθηγητή της εργασίας μου, κύριο Βαγγέλη Δασκαλάκη, Επίκουρο καθηγητή του Τεχνολογικού Πανεπιστημίου Κύπρου, για την υπομονή, την αμέριστη υποστήριξη και ενθάρρυνση που μου παρείχε όλο αυτό το διάστημα.

Φυσικά, η εργασία αυτή δεν θα είχε πραγματοιηθεί χωρίς την στήριξη της οικογένειας μου, που μου έμαθαν πως όλα είναι δυνατό να συμβούν και των φίλων μου που με στήριζαν και με συμβούλευαν καθ' όλη τη διάρκεια των μεταπτυχιακών μου σπουδών.

Τέλος, θα ήθελα να εκφράσω τις ευχαριστίες μου στον καθηγητή Alexander V. Ruban, του πανεπιστημίου Queen Mary στο Λονδίνο, καθώς και τον μεταδιδακτορικό ερευνητή δρ. Vasco Giovagnetti, για τη καθοδήγηση και τις πολύτιμες συμβουλές τους.

# ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ ..... 6
ABSTRACT ..... 7
1. ΕΙΣΑΓΩΓΗ ..... 8
1.1 Τι είναι τα διάτομα ..... 8
1.2 Το διάτομο Phaeodactylum tricornutum ..... 8
1.3 Φωτοσυνθετική ικανότητα διατόμων ..... 9
1.4 Οι πρωτεΐνες FCP ..... 10
1.5 Οι πρωτεΐνες LHCX1 ..... 11
1.6 Αλληλεπιδράσεις μεταξύ πρωτεϊνών με τη χρήση Μοραικής Δυναμικής ..... 12
2. ΣΚΟΠΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ ..... 15
3. ΜΕΘΟΔΟΙ-ΥΛΙΚΑ ..... 16
3.1 Δομή και δυναμική της πρωτεΐνης LHCXI ..... 16
3.1.1 Υπολογιστική επίλυση δομής LHCX1 ..... 16
3.1.2 Τοποθέτηση της LHCX1 στη θυλακοειδή μεμβράνη ..... 19
3.1.3 Δυναμική της πρωτεΐνης LHCX1 ..... 20
3.2 Το πρωτεϊνικό σύμπλοκο LHCX1-FCP ..... 24
3.3 Πρωτονίωση του πρωτεϊνικού συμπλόκου LHCX1-FCP ..... 26
3.4 Τοποθέτηση συμπλόκου στην θυλακοειδή μεμβράνη ..... 27
3.5 Δυναμική του συμπλόκου πρωτεϊνών LHCX1 και FCP ..... 29
3.5.1 Ελαχιστοποίηση ενέργειας και εξισορρόπηση συστήματος ..... 29
3.5.2 Παράγωγη Μοριακής Δυναμικής ..... 34
4.ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ ..... 37
4.1 Ανάλυση αποτελεσμάτων της δυναμικής του συμπλόκου πρωτεϊνών FCP και LHCX1 ..... 37
4.2 Συζήτηση ..... 45
4.3 Επίλογος-Μελλοντική ερευνητική εργασία ..... 46
ΒΙΒΛΙΟΓΡΑΦΙΑ ..... 47

# ΠΕΡΙΛΗΨΗ

Η φωτοσύνθεση βασίζεται σε σύμπλοκα κεραίων τα οποία συλλέγουν το ηλιακό φως και το μετατρέπουν σε χημική ενέργεια στα ενεργά κέντρα της φωτοσυνθετικής μηχανής. Τα διάτομα είναι ευκαριωτικά φύκη που βρίσκονται σε γλυκά νερά και ωκεανούς. Εκφράζουν πρωτεΐνες-κεραίες που περιέχουν φουκοξανθίνη και χλωροφύλλες a/ c (FCP), για τη συλλογή φωτός. Οι FCP έχουν εξαιρετικές ικανότητες συλλογής φωτός στη μπλε-πράσινη περιοχή του φάσματος, επίσης έχουν ισχυρό φωτοπροστατευτικό μηχανισμό που τους επιτρέπει να προσαρμόζονται υπό το έντονα κυμαινόμενο φως στην επιφάνεια του ωκεανού, σε συνδυασμό με τις πρωτεΐνες-φωτοπροστασίας (LHCX). Η LHCX1, μπορεί να προκαλέσει αλλαγές στη διαμόρφωση των FCP πρωτεϊνών μέσω δημιουργίας συμπλόκου FCP-LHCX1. Στη βιβλιογραφία λείπουν σημαντικά δομικά και δυναμικά δεδομένα για την οικογένεια των πρωτεϊνών LHCX των διατόμων. Χρησιμοποιώντας συγκριτική και δομική μοντελοποίηση, έχουμε προσδιορίσει τη δομή της LHCX1. Χρησιμοποιώντας μεθόδους όπως η Μοριακή Δυναμική, διερευνήθηκε (α) η απόκριση της LHCX1 στο pH, η οποία είναι ένας βασικός παράγοντας που σχετίζεται με τη φωτοπροστασία, καθώς επίσης (β) η δυναμική του συμπλόκου FCP-LHCX1. Τα διάτομα είναι ικανά να αφομοιώσουν μεγάλες ποσότητες διοξειδίου του άνθρακα και να συμβάλλουν στο 20% της παγκόσμιας πρωτογενούς παραγωγής. Ρυθμίζοντας το μηχανισμό φωτοπροστασίας τους, το ποσοστό αυτό θα μπορούσε να αυξηθεί. Τα αποτελέσματα συνεπώς, μπορούν να συνδυαστούν άμεσα με εφαρμογές για το μεταβαλλόμενο κλίμα και μπορούν επίσης λόγω ομοιότητας να συμβάλουν στην αποσαφήνιση του φωτοπροστατευτικού μηχανισμού στα ανώτερα φυτά.

**Λέξεις κλειδιά:** φωτοπροστατευτικός μηχανισμός, Μοριακή Δυναμική, *Phaeodactylum tricornutum*, LHCX1, FCP

## ABSTRACT

The photosynthetic process relies on light-harvesting antenna complexes (LHCs) that absorb sunlight and convert it into chemical energy within reaction centers. The LHCs of the diatoms (algae) are Fucoxanthin and Chlorophyll-a/c binding Proteins (FCP). FCPs have exceptional light-harvesting capabilities at the blue-green region, and they have also developed a robust photoprotective mechanism that enables diatoms to adapt under fluctuating (or high) light at the ocean surface. A special family of photoprotective proteins (LHCX) is essential for the latter mechanism and the adaptation of the diatoms in the changing environment. LHCX1 binds to FCP proteins under photoprotective conditions. Structural and dynamic data for the LHCX family of proteins are lacking in the literature. Employing comparative and structural modeling, we have determined the structure of LHCX1. Using methods such as Molecular Dynamics Simulations, we have monitored (a) the response of LHCX1 to pH, which is a key factor related to photoprotection, and (b) the dynamics of the FCP-LHCX1 complex. Diatoms are able to fix large amounts of carbon dioxide, and contribute to 20% of the global primary production. By adjusting their photoprotection mechanism, this percentage could be increased. Results can thus be directly associated with applications for the changing climate and can help in elucidating also, in a timely manner, the photoprotective mechanism in vascular plants.



**Key words:** photoprotection mechanism, Molecular Dynamics, *Phaeodactylum tricornutum*, LHCX1, FCP

# 1. ΕΙΣΑΓΩΓΗ

## 1.1 Τι είναι τα διάτομα

Τα διάτομα είναι ευκαριωτικοί οργανισμοί της οικογένειας Bacillariophyceae, συναντώνται σε ωκεανούς, λίμνες, ποτάμια αλλά και γενικά υγρότοπους . Είναι μονοκύτταροι οργανισμοί οι οποίοι εμφανίζονται μεμονωμένοι ή σε αποικίες, έχουν ποικίλα σχήματα, όπως αστέρια, βεντάλιες, νήματα, τεθλασμένες γραμμές και το μήκος τους μπορεί να φτάσει από 2 ως 200 μικρόμετρα (Hasle, 1996). Αναπαράγονται αγενώς, με κυτταρική διαίρεση και σε πολύ ευνοϊκές συνθήκες η αναπαραγωγή μπορεί να γίνεται κάθε 24 ώρες, ενώ έχουν περίπου 6 μέρες ζωής. Κινούνται κυρίως μέσω της ροής πεδίου των νερών – ρεμάτων. Ένα μοναδικό χαρακτηριστικό των διατόμων, όσον αφορά στην ανατομία τους, είναι πως περιβάλλονται από ένα κυτταρικό τοίχωμα από διοξείδιο του πυριτίου, το οποίο ονομάζεται κρούστα (Horner, 2002). Είναι κυρίως φωτοσυνθετικοί οργανισμοί, αν και υπάρχουν και ετερότροφοι για να είναι δυνατή η επιβίωση τους χωρίς την παρουσία ηλικιακού φωτός (Armstrong, 2000).

Η αξία των διατόμων είναι ιδιαίτερα μεγάλη, καθώς μέσω της φωτοσύνθεσης, μετατρέπουν την ηλιακή ενέργεια σε χημική και με αυτόν τον τρόπο, παράγουν πολύ μεγάλο ποσοστό οξυγόνου, ενώ απομακρύνουν σημαντικό ποσοστό διοξειδίου του άνθρακα από την ατμόσφαιρα (Bailleul, 2015). Επίσης, λόγω της ευαισθησίας ορισμένων διατόμων σε περιβαλλοντικές αλλαγές, χρησιμοποιούνται ως βιολογικοί δείκτες ποιότητας νερού και με τη βοήθεια απολιθωμάτων μπορούμε να έχουμε συμπεράσματα για την εξελικτική πορεία των οικοσυστημάτων όπου απαντώνται.

## 1.2 Το διάτομο Phaeodactylum tricornutum

Για την παρούσα εργασία, θα εξετάσουμε το διάτομο *Phaeodactylum tricornutum* (Bohlin. 1898), το οποίο θεωρείται διάτομο μοντέλο καθώς είναι διμερώς συμμετρικό και ένα από τα πρώτα διάτομα, στα οποία διερευνήθηκε πλήρως η αλληλουχία του γονιδιώματος του.

Όσον αφορά την ταξινόμηση του

- Επικράτεια: Ευκαριωτικά
- Βασίλειο: Πρώτιστα
- Φύλο: Bacilloriophyta
- Κλάση: Bacilloriophycaceae
- Τάξη: Bacilloriales
- Οικογένεια: Phaeodactylaceae
- Γένος: Phaeodactylum
- Είδος: Tricornutum

Η ιδιαιτερότητα του διατόμου αυτού βασίζεται στο γεγονός πως μπορεί να υπάρξει σε τρεις μορφότυπους, χωρίς να υπάρχει πάγια εξήγηση για το φαινόμενο αυτό. Πιθανώς,

οι περιβαλλοντικές συνθήκες επάγουν μορφογενετικούς μηχανισμούς που ελέγχονται επιγενετικά (De Martino, 2007).

Η επιλογή του συγκεκριμένου διατόμου δεν είναι τυχαία, καθώς έχει ορισμένα μοναδικά χαρακτηριστικά, εξαιτίας των οποίων, υπερτερεί σε σχέση με τα υπόλοιπα διάτομα για τις πειραματικές μελέτες στη βιβλιογραφία. Μπορεί να επιβιώσει σε ποικίλες συνθήκες, γι' αυτό και συναντάται σε όλο τον πλανήτη, συνήθως σε παράκτιες περιοχές. Είναι ιδιαίτερα ανθεκτικό σε στρεσογόνες συνθήκες, όπου οι τιμές του pH μπορεί να φτάσουν εώς και 10 (Goldman, 1982), ή σε συνθήκες έντονης αλατότητας με τιμές έως και 65 ppt (Abdullahi, 2006). Τέλος, αν και είναι το μόνο διάτομο, όπου η παρουσία πυριτίου δεν είναι απαραίτητη για την ανάπτυξη του, έχει ιδιαίτερα γρήγορο ρυθμό ανάπτυξης και βιοσυνθετική ικανότητα (Hemple, 2011).

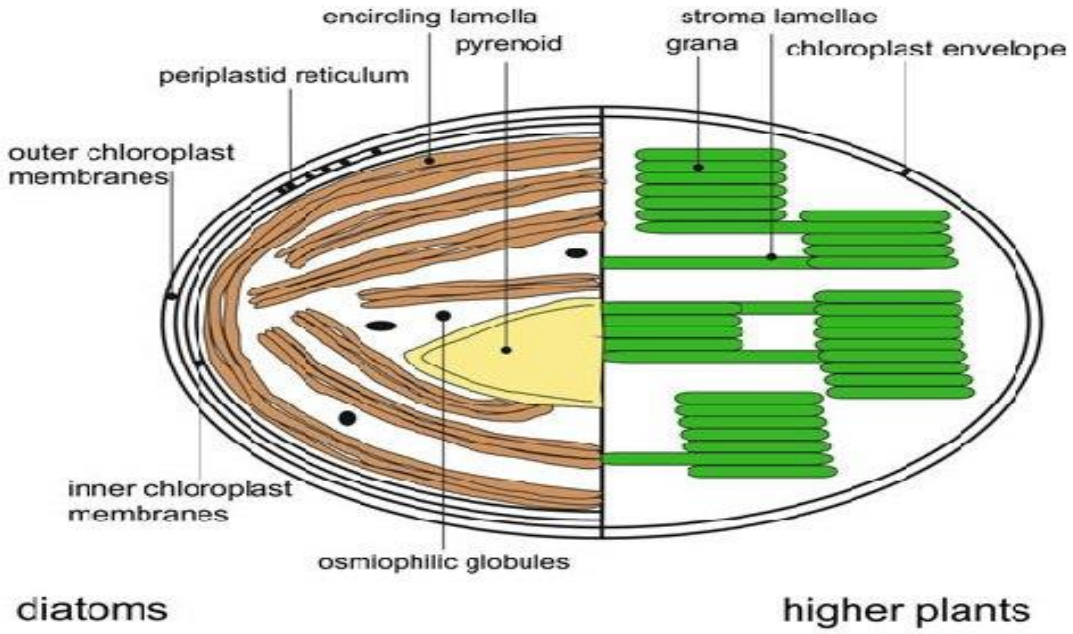
Εξαιτίας όλων των παραπάνω, έχει χρησιμοποιηθεί έντονα στη βιοτεχνολογία, σε ποικίλους τομείς, όπως η παραγωγή βιοκαυσίμου, ως ξενιστής για την έκφραση ανασυνδυασμένων πρωτεϊνών. Εκχύλισμά του έχει χρησιμοποιηθεί για τη μείωση κατεστραμμένων πρωτεϊνών από οξείδωση, σε κύτταρα που εκτέθηκαν σε UVA και UVB ακτινοβολία (Angerhofer, 2009), ενώ φαίνεται πως μελλοντικά θα είναι η κύρια πηγή της χρωστικής φουκοξανθίνης, η οποία χρησιμοποιείται σε υδατοκαλλιέργειες ή ακόμη και ως αντιοξειδωτικό.

### 1.3 Φωτοσυνθετική ικανότητα διατόμων

Όπως έχουμε ήδη αναφερθεί, τα διάτομα κινούνται παθητικά, επομένως βρίσκονται σε περιοχές με έντονη ροή – ανάμιξη υδάτων, επομένως και η ένταση της ακτινοβολίας που προσπίπτει ή διαχέεται σε αυτά δεν έχει σταθερή ένταση. Ως αποτέλεσμα, φαίνεται να επηρεάζεται και η φωτοσυνθετική τους ικανότητα (Lichtman, 2001).

Είναι οργανισμοί οι οποίοι μπορούν να επιβιώσουν σε όλο το φάσμα του ηλιακού φωτός και μπορούμε να τους συναντήσουμε εώς την επιπελαγική ζώνη (0-200m) (Fogg, 1999). Η ικανότητα τους να επιβιώνουν σε ένα περιβάλλον με μεγάλες διακυμάνσεις φωτός, δείχνει φωτοσυνθετική ευελιξία και επομένως μια ικανή προστασία από στρες. Έχουν μηχανισμούς οι οποίοι ρυθμίζουν άμεσα τη μεταφορά και χρήση της ηλιακής ενέργειας όπου θα χρησιμοποιήσουν σε φωτοχημικές και μη διεργασίες.

Λόγω της εξελικτικής διαδικασίας, ο φωτοσυνθετικός μηχανισμός των διατόμων διαφέρει από αυτόν των ανώτερων φυτών. Αν και κατά κύριο λόγο δεν υπάρχει στοίβαξη grana, έχει παρατηρηθεί ότι κάποιοι οργανισμοί, και ιδιαίτερα τα φαιοφύκη, συγκεντρώνουν το PSI (φωτοσύστημα I) στη πλευρά του στρώματος. Τα θυλακοειδή των διατόμων βρίσκονται κατά μήκος του πλασμιδίου, σε ζώνες των τριών (Pyszniac, 1992) και οι κεραίες συλλογής φωτός (LHC) κατά μήκος των θυλακοειδών μεμβρανών.




*Εικόνα 1.* Σχηματική απεικόνιση του χλωροπλάστη και της θυλακοειδούς δομής σε διάτομα (αριστερά) και ανώτερα φυτά (δεξιά). Τα πλαστίδια των διατόμων κατακλύζονται από τέσσερις μεμβράνες, τα θυλακοειδή τους σχηματίζουν τριπλά ελάσματα. Ένα από αυτά περικυκλώνει το στρώμα του χλωροπλάστη, το πυρενοειδές είναι μια δομή εμπλουτισμένη με RubisCO σε χλωροπλάστες φυκών. (Grouneva, 2013)

Ακόμη μια διαφορά μεταξύ ανώτερων φυτών και διατόμων, σχετικά με τον φωτοσυνθετικό μηχανισμό, είναι η περιεκτικότητα σε χρωστικές. Στα ανώτερα φυτά, υπάρχει μεγαλύτερο ποσοστό χλωροφυλλών (α και β) που τους προσδίδει το χαρακτηριστικό πράσινο χρώμα, ενώ οι ξανθοφύλλες έχουν ρόλο τόσο στη συλλογή φωτός, όσο και την προστασία της φωτοσυνθετικής μηχανής. Αντίθετα, τα διάτομα περιέχουν χλωροφύλλες (α και γ), ενώ έχουν μεγαλύτερο ποσοστό ξανθοφυλλών, δίνοντας τους μια καφέ απόχρωση (Wihlelm, 1990), ενώ λαμβάνουν περισσότερο ενεργό μέρος στη ρύθμιση της απορρόφησης και διάχυσης της ηλιακής ενέργειας. Οι ξανθοφύλλες των διατόμων με την μεγαλύτερη περιεκτικότητα είναι η φουκοξανθίνη (FX), η οποία είναι υπεύθυνη για τη σταθερότητα της δομής του συμπλόκου συλλογής φωτός (LHC) εντός της θυλακοειδούς μεμβράνης (Kuhlbrandt, 1994), ακολουθεί η διαδινοξανθίνη (DD), η οποία φαίνεται πως συμμετέχει στη μεταφορά ενέργειας προς την χλωροφύλλη α (Chl-a), και τέλος είναι η διατοξανθίνη (DT), η οποία έχει ρόλο στον κύκλο των ξανθοφυλλών ο οποίος ρυθμίζεται από την ένταση της ηλιακής ακτινοβολίας.

## 1.4 Οι πρωτεΐνες FCP

Οι φωτοσυνθετικοί οργανισμοί έχουν πρωτεΐνες-πολυπεπτίδια συλλογής φωτός, στις οποίες οι χρωστικές αποτελούν τις προσθετικές ομάδες έτσι ώστε να μπορούν να απορροφούν και να διοχετεύουν ενέργεια στον πυρήνα του φωτοσυστήματος, προς τα ενεργά κέντρα. Οι πρωτεΐνες συλλογής φωτός των διατόμων ονομάζονται πρωτεΐνες

δέσμευσης φουκοξανθίνης (Fx)- χλωροφύλλης (chl-a/ chl-c) και ανήκουν στην οικογένεια των διαμεμβρανικών πρωτεϊνών συλλογής φωτός (LHC). Η κεραία συλλογής φωτός των διατόμων αποτελείται από FCPs, οι οποίες λόγω των χρωστικών τους, έχουν εξαιρετικές δυνατότητες συλλογής φωτός αλλά και φωτοπροστασίας, με αποτέλεσμα να συμβάλλουν στην επιβίωση των διατόμων σε περιβάλλοντα όπου η ηλιακή ακτινοβολία δεν είναι σταθερή (Kuszynska, 2015). Κάθε μονομερές της FCP δεσμεύει χλωροφύλλη a/c, φουκοξανθίνη(Fx), διαδινοξανθίνη(Dd), διατοξανθίνη(Dt) και β-καροτένιο (Lepetit, 2010). Όπως και στα φυτά, οι FCPs περιβάλλουν τον πυρήνα του φωτοσυστήματος με αποτέλεσμα να σχηματίζουν ένα εξελιγμένο δίκτυο χρωστικών- πρωτεϊνών για την αποδοτικότερη συλλογή φωτός.

Η κρυσταλλική δομή του διμερούς της FCP έχει προσδιοριστεί, με τα τμήματά της στο στρώμα και το εσωτερικού της θυλακοειδούς μεμβράνης (lumen) να έχουν βασιστεί σε δομική σύγκριση με τις πρωτεΐνες – κεραίες LHC των ανωτέρων φυτών (Wang, 2019). Αν και η κρυσταλλική δομή υπάρχει στη βιβλιογραφία, δεν υπάρχει πληροφορία για την απόκριση της δομής σε διαφορετικές τιμές pH στο lumen. Τα κατάλοιπα ασπαρτικού (Asp) και γλουταμικού (Glu) που βρίσκονται στο lumen, δεν αλληλεπιδρούν με χρωστικές ή τη μεμβράνη, αλλά είναι εκτεθειμένα στην υδάτινη φάση. Τα αμινοξέα αυτά, υπό όξινες συνθήκες (pH < ~5.5), μπορούν να πρωτονιωθούν με αποτέλεσμα να προκαλέσουν αλλαγές στη δευτεροταγή διαμόρφωση της πρωτεΐνης (α- έλικες, β-πτυχωτές και βρόγχους). Επομένως, καθώς η φωτοπροστασία ενεργοποιείται σε όξινες συνθήκες στο lumen, είναι σημαντικό να μελετήσουμε τις πιθανές αυτές αλλαγές. Δομικές αλλαγές που εξαρτώνται από το pH, οδηγούν και αλλαγές στη δυναμική της πρωτεΐνης, πιθανώς σχετικές με το μηχανισμό φωτοπροστασίας (Nagao, 2020).

### 1.5 Οι πρωτεΐνες LHCX1

Οι κεραίες συλλογής φωτός των διατόμων έχουν την δυνατότητα να συλλέγουν μπλε-πράσινο φως και να αποκρίνονται γρήγορα στις διακυμάνσεις του φωτός, με την ενεργοποίηση του φωτοπροστατευτικού μηχανισμού. Όπως έχει ήδη αναφερθεί, με αυτό το μηχανισμό, τα διάτομα προσαρμόζονται κάτω από συνθήκες μεταβαλλόμενου φωτός. Για την ενεργοποίηση αυτού του μηχανισμού, ενεργό ρόλο αναλαμβάνουν οι πρωτεΐνες LHCX (Buck,2019).

Η οικογένεια των πρωτεϊνών LHCX ελέγχει την κύρια συνιστώσα του μηχανισμού της μη φωτοχημικής απόσβεσης, που εξαρτάται από την ενέργεια (qE). Ιδιαίτερο ρόλο παρουσιάζει η ισομορφή LHCX1. Κάτω από συνθήκες χαμηλού φωτισμού συμβάλει σχεδόν εξ ολοκλήρου στην απόσβεση της περίσσειας ενέργειας ,υπό μορφή θερμότητας, μέσω μη φωτοχημικής απόσβεσης, ενώ και οι υπόλοιπες ισομορφές έχουν παρόμοια λειτουργία σε συνθήκες φωτοπροστασίας (Buck, 2019).

Για την απόσβεση της περίσσειας ενέργειας, η παρουσία της LHCX1 οδηγεί σε αποτελεσματικότερη απόκριση στο φως, ελέγχει ποσοτικά το επίπεδο του

φωτοπροστατευτικού μηχανισμού μη φωτοχημικής απόσβεσης (NPQ), χωρίς ο τελευταίος μηχανισμός να είναι πλήρως κατανοητός (Bailleul, 2010). Έχει προταθεί πως η LHCX1 μπορεί να αισθανθεί αλλαγές στις τιμές του pH στο εσωτερικό της θυλακοειδούς μεμβράνης (lumen) και φαίνεται πως έχει επιπλέον δομικό ρόλο στο φωτοσύστημα II (Zhu, 2010).

Στη βιβλιογραφία είναι αποδεκτό πως η πρωτεΐνη LHCX1 είναι η αντίστοιχη της PsbS (Photosystem II Subunit S) που βρίσκεται στο φωτοσύστημα II των ανώτερων φυτών (Bailleul, 2010). Προτείνεται πως η LHCX1 μπορεί να προκαλέσει αλλαγές στη διαμόρφωση αλλά και την ευελιξία των μεμβρανικών FCP στα διάτομα. Η ικανότητα της LHCX1 να αλληλεπιδρά με τις FCP έχει ως αποτέλεσμα την απόρριψη της περίσσειας ενέργειας μέσω θερμότητας (Buck, 2019), επομένως μειώνεται και η υπερδιέγερση/ αναγωγική πίεση στο φωτοσύστημα II.

Γενικά οι πρωτεΐνες LHCX έχουν κύριο ρόλο στη ρύθμιση της ισορροπίας μεταξύ συλλογής ενέργειας και φωτοπροστασίας. Ιδιαίτερα η LHCX1 κατέχει τον κυρίαρχο ρόλο στην απόσβεση ενέργειας με μη φωτοχημικές διαδικασίες (NPQ). Ωστόσο, δεν υπάρχουν πολλές καταγραφές για τα δομικά και λειτουργικά χαρακτηριστικά των LHCX στη βιβλιογραφία.

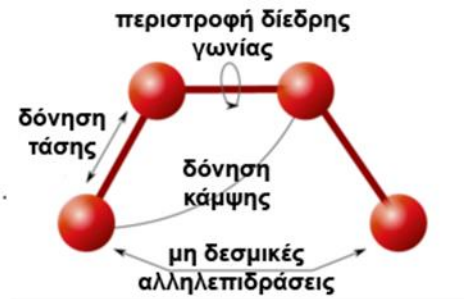
### 1.6 Αλληλεπιδράσεις μεταξύ πρωτεϊνών με τη χρήση Μοραικής Δυναμικής

Η μέθοδος της Μοριακής Δυναμικής χρησιμοποιείται ευρέως στην προσομοίωση της συμπεριφοράς των υλικών σε ατομικό επίπεδο, η οποία δεν μπορεί να επιτευχθεί με άλλες θεωρητικές μεθόδους λόγω μεγέθους του υπό μελέτη συστήματος ή πειραματικές διατάξεις λόγω μειωμένης διακριτικής ικανότητας. Εφαρμόζεται σε επιστήμες όπως η φυσικοχημεία, η επιστήμη υλικών και η βιοφυσική (Freddolino, 2010). Η πρώτη αναφορά υπολογισμού Μοριακής Δυναμικής έγινε στη μελέτη αλληλεπιδράσεων σε ένα σύστημα σκληρών σφαιρών (Alder and Wainwright, 1959), ενώ η πρώτη προσομοίωση πρωτεΐνης με βάση τη Μοριακή Δυναμική έγινε το 1977 από τον McCammon.

Η Μοριακή Δυναμική (ΜΔ) συνδυάζει τη Μοριακή Μηχανική για την περιγραφή των αλληλεπιδράσεων μεταξύ ατόμων. Για παράδειγμα χρησιμοποιεί τον νόμο του ελατηρίου του Hooke για δεσμικές αλληλεπιδράσεις, το νόμο του Coulomb και το Δυναμικό Lennard-Jones για μη δεσμικές αλληλεπιδράσεις, υπολογίζοντας την ενέργεια του συστήματος E ως σύνολο επί μέρους ενεργειών (Σχήμα 1):

$$E=E(\text{δεσμικές})+E\ (\text{μη δεσμικές})$$

$$E(r) = \sum_{\text{Δεσμοί}} k_b(b-b_0)^2 + \sum_{\text{Γωνίες}} k_\theta(\theta-\theta_0)^2 + \sum_{\text{Δίεδρες}} k_\phi\left(1+\cos(n\phi-\phi_0)\right) + \sum_{\text{Δίεδρες}} k_\psi(\psi-\psi_0)^2 + \sum_{\substack{\text{Lenard-Jones}\\ \text{(μη Δεσμικά)}}} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{\substack{\text{Coulomb}\\ \text{(μη Δεσμικά)}}} \frac{q_i q_j}{\varepsilon_D r_{ij}}.$$




*Σχήμα 1.* Δεσμικές και μη Δεσμικές αλληλεπιδράσεις μεταξύ των ατόμων σε ένα σύστημα.

Η ενέργεια εξαρτάται από ένα σύνολο παραμέτρων και μεταβλητών (Σχήμα 1). Οι παράμετροι kb, kθ αναφέρονται στις σταθερές του ελατηρίου κατά Hooke, που προσομοιώνει την τάση δεσμών και κάμψη γωνιών. Τα kφ και kψ αναφέρονται σε παραμέτρους που ελέγχουν την περιστροφή δίεδρων γωνιών, τα εij, σij αφορούν στην ενέργεια και ακτίνα van der Waals (υδρόφοβες αλληλεπιδράσεις), ενώ τα εD, qi και qj αναφέρονται στη διηλεκτρική σταθερά και τα φορτία των ατόμων i-j. Τα b0, θ0, φ0 και ψ0 αναφέρονται στο μήκος δεσμού ισορροπίας, στη γωνία ισορροπίας, και στις δίεδρες γωνίες ισορροπίας (T=0K). Για κάθε χρονική στιγμή, οι παράμετροι ισορροπίας παίρνουν τις μεταβλητές τιμές b, θ, φ και ψ αντίστοιχα. Το σύνολο των παραμέτρων ονομάζεται πεδίο δυνάμεων ή Force Field και περιγράφει πλήρως ένα μοριακό συστήμα.

Η Μοριακή Δυναμική ακολουθεί την Κλασσική Μηχανική με τις εξισώσεις κίνησης του Νεύτωνα, για σώματα, να δίνουν την κίνηση στα άτομα, με βάση τις δυνάμεις που ασκούνται σε αυτά:

F=-∇E

F=m×a, όπου F η δύναμη, m η μάζα και a η επιτάχυνση).

Οι προσομοιώσεις Μοριακής Δυναμικής μπορούν να παράγουν τη μεταβολή μιας πρωτεϊνικής δομής σε συνάρτηση με τον χρόνο κατά τη διάρκεια της προσομοίωσης (Haile, 1997), καθώς επιλύεται η παρακάτω διαφορική εξίσωση, για κάθε άτομο i:

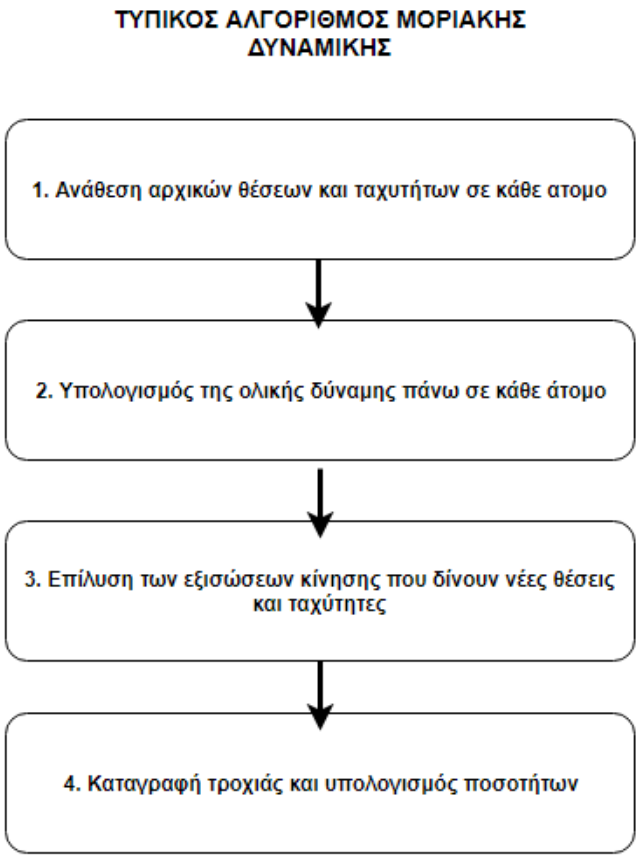
-∇E=m×a ⇒-dE(i)/dr(i) =m×(d^2 r(i))/⟦dt⟧^2 ⇒r(i)(t)=⋯, όπου r(i) οι συντεταγμένες του κάθε ατόμου i και t ο χρόνος.

Με την παραπάνω διαδικασία, οδηγούμαστε σε μια χρονοεξαρτώμενη θέση του κάθε ατόμου i στο χώρο ή την 'τροχιά' των ατόμων της πρωτεΐνης. Η τροχιά αυτή περιλαμβάνει όλες τις διαμορφώσεις που θα μπορούσαν να βρεθούν πειραματικά σε ένα διάλυμα της πρωτεΐνης, μια δεδομένη χρονική στιγμή, δεδομένων συνθηκών πίεσης, και θερμοκρασίας.

Όπως και σε μια πειραματική διαδικασία, έτσι και με τη Μοριακή Δυναμική μπορούμε να παρακολουθήσουμε τη χρονική εξέλιξη ενός συστήματος, έτσι ώστε το σύστημα να περάσει από όλες τις πιθανές καταστάσεις. Ωστόσο, με την προσομοίωση μπορούμε να εξετάσουμε πως τα συστήματα εξελίσσονται και περνάνε από διαφορετικές

διαμορφώσεις, με ελεγχόμενες μεταβλητές (θερμοκρασία, πίεση, pH), σε ατομικό επίπεδο, το οποίο δεν είναι πάντοτε εφικτό πειραματικά (π.χ. αδυναμία κρυστάλλωσης μιας πρωτεΐνης στην ενεργή ή ανενεργή διαμόρφωση, ή στη διαμόρφωση σε διάλυμα/ μεμβράνη).

Για να γίνει η προσομοίωση πρέπει να ακολουθήσουμε συγκεκριμένα βήματα. Όπως και στο πείραμα, θα πρέπει να γίνει προετοιμασία του συστήματος, θα πρέπει να δηλώσουμε τις αρχικές θέσεις και κατανομές ταχύτητες των σωματιδίων (ώστε να προσομοιώσουμε συγκεκριμένη θερμοκρασία - πίεση) καθώς έπειτα να βρεθεί μια αρχική δομή ελάχιστης ενέργειας που απαντάται στις πειραματικές διατάξεις. Ακολουθεί η παραγωγική φάση της προσομοίωσης κατά την οποία υπολογίζονται οι ενέργειες, οι δυνάμεις στα άτομα, οι νέες ταχύτητες και θέσεις τους την επόμενη χρονική στιγμή. Μετά το τέλος της προσομοίωσης συλλέγονται και αναλύονται τα δεδομένα (εικόνα 2).




*Εικόνα 2*. Διάγραμμα τυπικού αλγόριθμου της Μοριακής Δυναμικής. Τα βήματα 2, 3, 4 επαναλαμβάνονται για κάθε χρονικό βήμα.

# 2. ΣΚΟΠΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Σκοπός της εργασίας είναι να διερευνηθεί η αλληλεπίδραση μεταξύ των πρωτεϊνών FCP και LHCX1, στο διάτομο Phaeodactylum tricornutum, εντός της θυλακοειδούς μεμβράνης και να παρέχουμε πληροφορίες που λείπουν από τη βιβλιογραφία και είναι σημαντικές στη διερεύνηση του φωτοπροστατευικού μηχανισμού των διατόμων, σε μια μελέτη βασικής έρευνας.

Για την υλοποίηση του σκοπού αυτού, θα πρέπει πρώτα να δημιουργήσουμε και έπειτα να βελτιστοποιήσουμε τη δομή της LHCX1 μέσω μοντελοποίησης κατ' ομολογία και εκτεταμένων προσομοιώσεων μοριακής δυναμικής, υπό την επίδραση διαφορετικών τιμών pH. Η δομή της LHCX1 δεν έχει λυθεί πειραματικά, οπότε δεν έχουμε στη διάθεσή μας μια αρχική θέση ατόμων της LHCX1 στο χώρο.

Εφόσον καταλήξουμε σε μια σταθερή δομή της πρωτεΐνης LHCX1, θα πρέπει να δημιουργήσουμε μοντέλα σύνδεσης μεταξύ των LHCX1 και FCP. Τα πιο σταθερά σύμπλοκα θα βελτιστοποιηθούν στη δομή μέσα από προσομοιώσεις μοριακής δυναμικής. Η βελτιστοποίηση της δομής έγκειται στην επίλυση των εξισώσεων κίνησης του Νεύτωνα, ώστε οι δυνάμεις F στα άτομα να είναι οι ελάχιστες. Με τη διαδικασία αυτή θα πάρουμε σταθερά σύμπλοκα LHCX1-FCP που αν και παίζουν ενεργό ρόλο στη φωτοπροστασία στα διάτομα, δεν έχουν προσδιοριστεί πειραματικά.

Τέλος, στην επιφάνεια διεπαφής μεταξύ των πρωτεϊνών, θα γίνει ταυτοποίηση των καταλοίπων των αμινοξέων που έχουν σημαντικό ρόλο στη σταθερότητα του συμπλόκου αλλά και πώς αυτή ρυθμίζεται από τις αλληλεπιδράσεις των αμινοξέων που ρυθμίζονται από τις αλλαγές του pH.

Τα αποτελέσματα αυτών των διαδικασιών θα συμπληρώσουν το βιβλιογραφικό κενό τόσο για την αλληλεπίδραση μεταξύ των δυο πρωτεϊνών, όσο και για τον ρόλο της LHCX1 των διατόμων στις αλλαγές της διαμόρφωσης της FCPs, που σχετίζονται με τη φωτοπροστασία, σε ατομικό επίπεδο.

# 3. ΜΕΘΟΔΟΙ-ΥΛΙΚΑ

Ο ρόλος της πρωτεΐνης LHCX1 είναι σημαντικός στη φωτοπροστασία ή την προσαρμογή του διατόμου *P.tricornutum* σε περιβάλλοντα με διακύμανση του φωτός (Buck, 2019). Ωστόσο, δεν υπάρχουν στοιχεία σε ατομική κλίμακα για τη δομή και τη δυναμική της LHCX1 στη βιβλιογραφία. Αρχικά, λοιπόν, θα πρέπει να προσδιορίσουμε τη δομή της LHCX1. Για να μπορέσουμε να μελετήσουμε το ρόλο της LHCX1 στις παραπάνω διεργασίες φωτοπροστασίας, στοχεύουμε, στη συνέχεια, στις αλληλεπιδράσεις μεταξύ των πρωτεϊνών της LHCX1 και της Fucoxanthin (Fx) και Chlorophyll (Chl) a/c-binding Protein (FCP) (Giovagnetti, 2018), με την τελευταία να έχει κεντρικό ρόλο στην απορρόφηση φωτός και να έχει λυθεί δομικά (Wang, 2019).

## 3.1 Δομή και δυναμική της πρωτεΐνης LHCXI

### 3.1.1 Υπολογιστική επίλυση δομής LHCX1

Για την υπολογιστική επίλυση της δομής της LHCX1 χρησιμοποιήσαμε το Jalview, το οποίο είναι ένα εργαλείο βιοπληροφορικής, και χρησιμοποιείται για την παραγωγή, προβολή και επεξεργασία αλληλουχιών καταλοίπων (πρωτοταγή δομή πρωτεϊνών). Πραγματοποιήσαμε ευθυγράμμιση αλληλουχιών (sequence aligment), για να βρούμε το ποσοστό ομολογίας της ζητούμενης πρωτεΐνης LHCX1 με άλλες, γνωστής δομής και παρόμοιας λειτουργίας, ώστε να χρησιμοποιηθούν ως βάση, στην πρόβλεψη μίας αρχικής τρισδιάστατης δομής της LHCX1. Στη συνέχεια, τα αποτελέσματα της ευθυγράμμισης δίνονται στο εργαλείο Phyre2 (Kelley, 2015), το οποίο είναι διαδικτυακή υπηρεσία που προβλέπει δομές πρωτεϊνών με βάση γνωστές δομές παρόμοιας αλληλουχίας.

Κατά την παραπάνω διαδικασία διαπιστώσαμε πως οι εξ ορισμού παράμετροι πρόβλεψης της δομής της LHCX1 στο Phyre2, με βάση την πρωτοταγή της δομή (αλληλουχία), παράγει διαμορφώσεις μη συμβατές με το δομικό μοτίβο των LHC από τα ανώτερα φυτά. Έγινε σύγκριση της αλληλουχίας/ δομής της LHCX1 με αυτές των πρωτεϊνών: LHCII (σπανάκι) (Wei, 2016), PsbS (σπανάκι) (Fan, 2015), Fucoxanthine/Chlorophyll a-b binding protein (*Chaetoceros gracilis*) (Pi, 2019) και CP29 (arabidopsis) (Yakushevska,2003). Με σκοπό την καλύτερη πρόβλεψη της δομής της LHCX1, τροποποιήσαμε την αλληλουχία της, προεκτείνοντας τα N-/ C- τελικά της άκρα, με αμινοξέα από τις αλληλουχίες των LHC στα ανώτερα φυτά, που φαίνεται να απουσιάζουν στην LHCX1, όταν πραγματοποιείται «sequence alignment» με το JalView. Διαπιστώσαμε πως η προσθήκη αμινοξέων στο C- τελικό άκρο στην αλληλουχία της LHCX1 οδηγεί σε πολύ καλύτερη πρόβλεψη δομής με το Phyre2, καθώς δημιουργεί δομικά μοτίβα σε αυτές τις περιοχές συμβατά με τα δομικά χαρακτηριστικά των LHC από ανώτερα φυτά. Συγκεκριμένα, από την επεξεργασία των αλληλουχιών με το Jalview, προκύπτουν τέσσερις νέες αλληλουχίες από τις οποίες, καθεμιά περιλαμβάνει την

αλληλουχία της πρωτεΐνης LHCX1 και κάποια αμινοξέα τα οποία ανήκουν στο C-τελικό άκρο των υπόλοιπων LHC πρωτεϊνών.

LHCXI
MKFAATILALIGSAAAFAPAQTSRASTSLQYAKEDLVGAIPPVGFFDPLGFADKADSPTLKRYREAELTHGR
VAMLAVVGFLVGEAVEGSSFLFDASISGPAITHLSQVPAPFWVLLTIAIGASEQTRAVIGWVDPADAPVD
KPGLLRDDYVPGDLGFDPLGLKPSDPEELITLQTKELQNGRLAMLAAAGFMAQELVNGKGILENLQG

LHCXI + LHCII
MKFAATILALIGSAAAFAPAQTSRASTSLQYAKEDLVGAIPPVGFFDPLGFADKADSPTLKRYREAELTHGR
VAMLAVVGFLVGEAVEGSSFLFDASISGPAITHLSQVPAPFWVLLTIAIGASEQTRAVIGWVDPADAPVD
KPGLLRDDYVPGDLGFDPLGLKPSDPEELITLQTKELQNGRLAMLAAAGFMAQELVNGKGILENLQGHLA
DPVNNNAWNFATNFVPGK

LHCXI + PSBS
MKFAATILALIGSAAAFAPAQTSRASTSLQYAKEDLVGAIPPVGFFDPLGFADKADSPTLKRYREAELTHGR
VAMLAVVGFLVGEAVEGSSFLFDASISGPAITHLSQVPAPFWVLLTIAIGASEQTRAVIGWVDPADAPVD
KPGLLRDDYVPGDLGFDPLGLKPSDPEELITLQTKELQNGRLAMLAAAGFMAQELVNGKGILENLQGETG
VPINEIEPLVLLNVVFFFIAAINPGTGKFITDDEEED

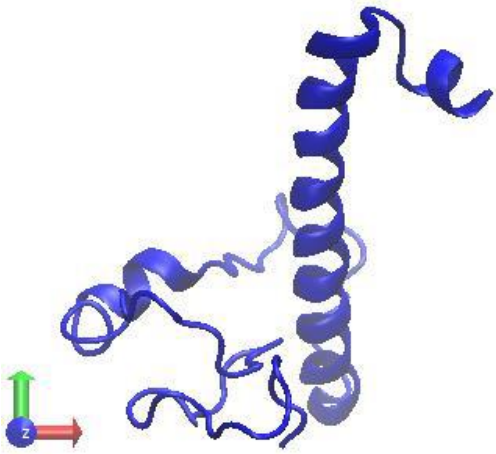
LHCXI + CP29
MKFAATILALIGSAAAFAPAQTSRASTSLQYAKEDLVGAIPPVGFFDPLGFADKADSPTLKRYRAELTHGR
VAMLAVVGFLVGEAVEGSSFLFDASISGPAITHLSQVPAPFWVLLTIAIGASEQTRAVIGWVDPADAPVD
KPGLLRDDYVPGDLGFDPLGLKPSDPEELITLQTKELQNGRLAMLAAAGFMAQELVNGKGILENLQGQF

LHCXI + chlorophyll a-b binding protein
MKFAATILALIGSAAAFAPAQTSRASTSLQYAKEDLVGAIPPVGFFDPLGFADKADSPTLKRYREAELTHGR
VAMLAVVGFLVGEAVEGSSFLFDASISGPAITHLSQVPAPFWVLLTIAIGASEQTRAVIGWVDPADAPVD
KPGLLRDDYVPGDLGFDPLGLKPSDPEELITLQTKELQNGRLAMLAAAGFMAQELVNGKGILENLQGHLS
DPLHTTIIDTFSSS

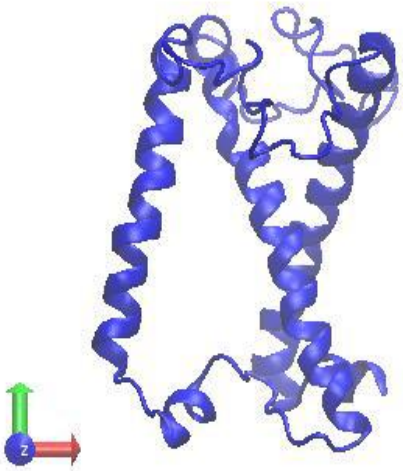
*Εικόνα 3.* Καταγραφή νέων αλληλουχιών, με την προσθήκη αμινοξέων στο καρβοξυτελικό άκρο της LHCX1.

Με τη βοήθεια του Phyre2, προκύπτουν πέντε τρισδιάστατες δομές της LHCX1 πρωτεΐνης για όλες τις περιπτώσεις στην εικόνα 3. Από τις τελευταίες καλούμαστε να επιλέξουμε την καλύτερη δυνατή, ώστε να προχωρήσουμε περαιτέρω. Σε κάθε περίπτωση, μετά την πρόβλεψη από το Phyre2, αφαιρείται το τμήμα της δομής που δεν ανήκει στην LHCX1, από το τελικό C-άκρο.

Ακολουθεί οπτική αξιολόγηση των δομών και σύγκριση μέσω υπέρθεσης με γνωστές δομές LHC από τα ανώτερα φυτά. Συγκεκριμένα, για το σκοπό αυτό, χρησιμοποιούμε το VMD, ένα πρόγραμμα μοριακής απεικόνισης. Όπως αναφέραμε παραπάνω, με την πρωτοταγή αλληλουχία της LHCX1 δεν μπορούμε να παράγουμε δομή συμβατή με το δομικό μοτίβο των LHC των ανώτερων φυτών. Βλέπουμε πως δημιουργείται μόλις μια α-έλικα (εικόνα 4), επομένως δεν μπορούμε να τη χρησιμοποιήσουμε εφόσον η δομή της δεν είναι ολοκληρωμένη. Σε αντίθεση, οι προβλέψεις με τα επαυξημένα C-άκρα, εμφανίζονται ολοκληρωμένες, με περισσότερες διαμεμβρανικές έλικες (εικόνα 5).



*Εικόνα 4.* Δομή της αλληλουχίας LHCX1, χωρίς την προσθήκη αμινοξέων στο καρβοξυτελικό άκρο.



*Εικόνα 5.* Δομή της αλληλουχίας LHCX1, με την προσθήκη αμινοξέων από την CP29 στο καρβοξυτελικό άκρο.

Συγκρίνοντας τις τέσσερις δομές που προκύπτουν με το Phyre2, η δομή της LHCX1 με προέκταση του C-τελικού άκρου με βάση την CP29 παρουσιάζει βέλτιστη αναπαράσταση για τις α-έλικες καθώς και τις θηλιές (εικόνα 4), ενώ βρίσκεται σε πλήρη συμβατότητα με αυτές της LHC οικογένειας, όπως παρατηρείται μετά από τη δομική σύγκριση. Επομένως, μπορούμε να την χρησιμοποιήσουμε περαιτέρω. Σημειώνεται πως η επιλογή της καταλληλότερης δομής έγινε οπτικά, με δομική ευθυγράμμιση (structural alignment), και με βάση την εμπειρίας μας σε δομές συμπλόκων LHCII από ανώτερα φυτά (Daskalakis, 2020).

### 3.1.2 Τοποθέτηση της LHCX1 στη θυλακοειδή μεμβράνη

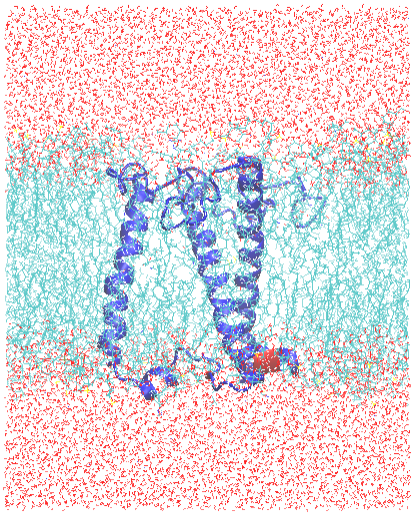
Επόμενο βήμα, αφότου προσδιορίσαμε μια στατική τρισδιάστατη δομή της LHCX1, είναι η τοποθέτηση της μέσα σε μοντέλο πλήρως ενυδατωμένης θυλακοειδούς μεμβράνης από τα διάτομα (van Eerden, 2015). Η μεμβράνη αποτελείται από λιπίδια (monogalactosyldiacylglycerol - MGDG, digalactosyldiacylglycerol - DGDG, sulfoquinovosyldiacylglycerol - SQDG, phosphatidylglycerol - PG) (Lepetit, 2012). Τα λιπίδια απεικονίζονται με γαλάζιο χρώμα, νερό (SOL) και ιόντα Νατρίου (Na) απεικονίζονται με κόκκινο-άσπρο χρώμα (εικόνα 4). Η βασική αρχή της τοποθέτησης μιας πρωτεΐνης σε μεμβράνη, είναι οι α-έλικες να βρίσκονται στην περιοχή των λιπιδίων, ενώ υδρόφιλα τμήματα να βρίσκονται σε επαφή με τις υδάτινες φάσεις, στις επιφάνειες της μεμβράνης.

Η διαδικασία της τοποθέτησης γίνεται έχοντας ως σημείο αναφοράς το κέντρο μάζας της μεμβράνης, στο οποίο θα αντιστοιχίσουμε και το κέντρο μάζας της πρωτεΐνης. Ακολουθεί η αφαίρεση λιπιδίων που αλληλεπικαλύπτονται με κατάλοιπα της πρωτεΐνης σε απόσταση μικρότερη του 1.5 Άνγκστρομ (Å). Η τιμή του 1.5A προέκυψε μετά από αρκετές προσπάθειες δοκιμής-λάθους. Αποτέλεσμα όλης αυτής της διεργασίας, είναι να δημιουργηθεί ένα κενό στη θυλακοειδή μεμβράνη, στο οποίο ταιριάζει να ενσωματωθεί η δομή της LHCX1.

Ωστόσο, για να μπορέσουμε να τοποθετήσομε την πρωτεΐνη στη θυλακοειδή μεμβράνη, πρέπει να αναπαρίστανται σωστά οι καταστάσεις πρωτονίωσης των κατάλοιπων για συγκεκριμένες τιμές pH. Μέσω του PDB2PQR Server (μέθοδος propka, (Dolinsky, 2004)), υπολογίζουμε το ισοϊοντικό σημείο αμινοξέων (pKa) όπως γλουταμικού (GLU), ασπαρτικού (ASP) και ιστιδίνης (HIS), για κάθε κατάλοιπο της LHCX1. Οι πρωτονιώσεις προβλέπονται για τιμές pH ~7.0 και ~5.5 που θεωρούνται σχετικές για τη μετάβαση των LHCX1/ FCP από την κατάσταση απορρόφησης φωτός, στην κατάσταση φωτοπροστασίας αντίστοιχα (Giovegnetti, 2018) και με βάση τις προβλέψεις προστίθενται πρωτόνια όπου κρίνεται απαραίτητο.

Αποτέλεσμα όλων των παραπάνω βημάτων είναι η δημιουργία μιας στατικής δομής της πρωτεΐνης LHCX1 (εικόνα 6), με συγκεκριμένες πρωτονιώσεις ανά τιμή pH, που περιγράφονται αναλυτικά και έχουν επιβεβαιωθεί από την πειραματική βιβλιογραφία να σχετίζονται με τη μετάβαση συλλογής φωτός – φωτοπροστασίας. Τα Asp (D) και Glu (E) είναι αποπρωτονιωμένα, εκτός των E-65 και E-85 τα οποία είναι πρωτονιωμένα σε όλες τις περιπτώσεις. Συγκεκριμένα, σύμφωνα με τον Ballottari (2016), υπάρχουν αμινοξέα στις φωτοπροστατευτικές πρωτεΐνες, τα οποία επηρεάζονται από την αλλαγή του pH και έχουν σημαντικό ρόλο στη φωτοπροστασία. Ωστόσο, ο Ballottari ασχολήθηκε με τη πρωτεΐνη LHCSR3, οπότε πρέπει να εντοπίσουμε τα αντίστοιχα αμινοξέα στην LHCX1. Αξίζει να ασχοληθούμε με τα D-95 και E-205, καθώς φαίνεται να είναι τα αντιστοιχία που παρατήρησε ο Ballottari και φαίνεται πως έχουν κάποια σημασία στην επαγωγή του φωτοπροστατευτικού μηχανισμού. Επομένως, με βάση τα παραπάνω λαμβάνουμε υπόψη τις περιπτώσεις: α) για τιμή pH ~5.5 τα D-95/ E-205 είναι

πρωτονιωμένα, β) για τιμή pH 7.0 και τα δύο είναι αποπρωτονιωμένα.



*Εικόνα 6.* Δομή της πρωτεΐνης LHCX1 μέσα στη θυλακοειδή μεμβράνη, σε pH 7, μετά τον καθαρισμό λιπιδίων, νερού και ιόντων νατρίου, και τη προσθήκη πρωτονίων, σε 1,5 Å.

### 3.1.3 Δυναμική της πρωτεΐνης LHCX1

Μετά τη δημιουργία της στατικής δομής της LHCX1, ακολουθεί η διερεύνηση της δυναμικής της μέσω της μεθόδου της Μοριακή Δυναμικής (λογισμικό Gromacs, http://www.gromacs.org/). Στη διαδικασία αυτή, όπως αναφέρθηκε, προσομοιώνεται η κίνηση των ατόμων στο χώρο και τον χρόνο, δίνοντας διαφορετικές διαμορφώσεις της πρωτεΐνης στο χρόνο.

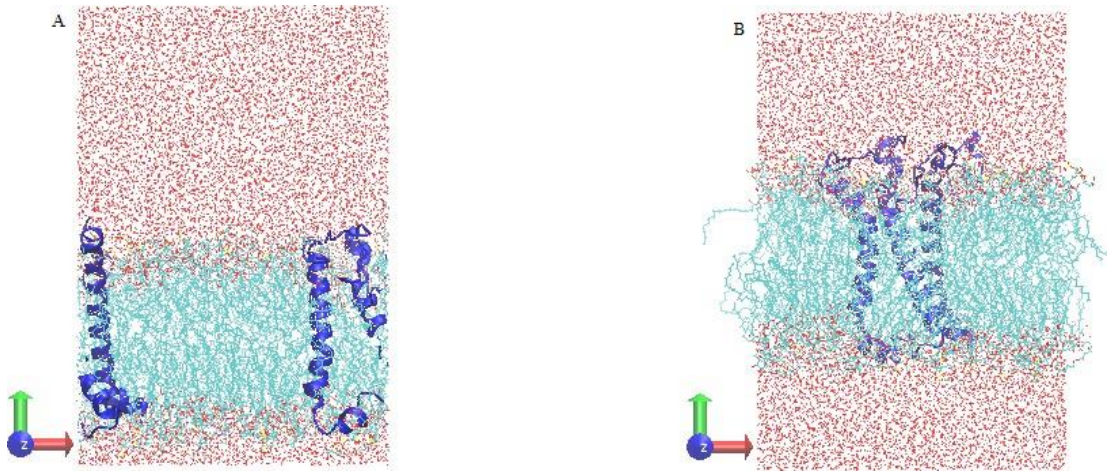
Έχοντας δημιουργήσει το σύστημα πρωτεΐνης- μεμβράνης, πρέπει να διασφαλίσουμε ότι το σύστημα δεν έχει ακατάλληλη γεωμετρία ή δεν θα έχει συγκρούσεις λόγω της χωρικής διάταξης των μορίων από την αρχική τοποθέτηση στη μεμβράνη. Θα ακολουθήσουμε μια διαδικασία όπου θα δώσουμε στο σύστημα μας την κατάλληλη γεωμετρία/ προσανατολισμό στα μόρια, και κατανομή ταχυτήτων στα άτομα με βάση τη θερμοκρασία (310K), και πίεση (1 atm).

Η διαδικασία αυτή περιλαμβάνει τα εξής στάδια: ελαχιστοποίηση ενέργειας (energy minimization EM), εξισορρόπηση συστήματος (equilibration). Αρχικά, πρέπει να βρούμε τη δομή ελάχιστης ενέργειας για το σύστημα. Για παράδειγμα, αν δύο άτομα έχουν έρθει σε πολύ κοντινή απόσταση στη στατική δομή, λόγω τυχαίας τοποθέτησής τους στο σύστημα, θα αποκτήσουν υψηλές ταχύτητες σε αντίθετες διευθύνσεις, ώστε θα απομακρυνθούν εκρηκτικά, και η θερμοκρασία του συστήματος θα αυξηθεί σημαντικά στα επόμενα βήματα. Για το σκοπό αυτό, παραγωγίζουμε την εξίσωση της ενέργειας ως προς τις συντεταγμένες των ατόμων και επιλύουμε ως προς τις συντεταγμένες αυτές

θέτοντας την παράγωγο ως μηδέν (βλ. εξισώσεις κίνησης). Μετά την εύρεση της δομής ελάχιστης ενέργειας, θα πρέπει να σταθεροποιήσουμε τη θερμοκρασία και την πίεση του συστήματος μέσα από διαδοχικά στάδια στα οποία κάθε φορά χαλαρώνουμε τους περιορισμούς κίνησης της πρωτεΐνης ώστε να προσαρμόσουμε τη θυλακοειδή μεμβράνη στη γεωμετρία της πρωτεΐνης. Όταν το σύστημα καταστεί ισορροπημένο σε δυνάμεις που ασκούνται στα άτομα, στην επιθυμητή θερμοκρασία και πίεση, εξάγουμε πέντε δομές που απέχουν 1 ns στο χρονικό πλαίσιο από τα τελευταία 5ns της διαδικασίας εξισορρόπησής. Στο επόμενο στάδιο, παράγουμε τροχιές ατόμων με βάση τη μοριακή δυναμική για τη συλλογή δεδομένων (5x 0.5μs).

Το αποτέλεσμα της Μοριακής Δυναμικής είναι αρχεία τα οποία περιέχουν τις συντεταγμένες των ατόμων της πρωτεΐνης σε συνάρτηση με τον χρόνο, και ουσιαστικά περιέχουν δομές οι οποίες στο σύνολο τους δηλώνουν τη δυναμική της πρωτεΐνης.

Έπειτα από την παραπάνω διαδικασία, η πρωτεΐνη μπορεί να φαίνεται «σπασμένη» ή να είναι σε οποιοδήποτε σημείο της μεμβράνης (εικόνα 7Α). Αυτό είναι απόλυτα φυσιολογικό, καθώς έχουμε εφαρμόσει περιοδικές συνθήκες στο σύστημά μας. Οπότε, το πρώτο βήμα πριν προχωρήσουμε σε οποιαδήποτε διαδικασία ανάλυσης είναι να τοποθετήσουμε την πρωτεΐνη στο κέντρο της μεμβράνης (εικόνα 7Β) καθ' όλη τη διάρκεια των τροχιών.

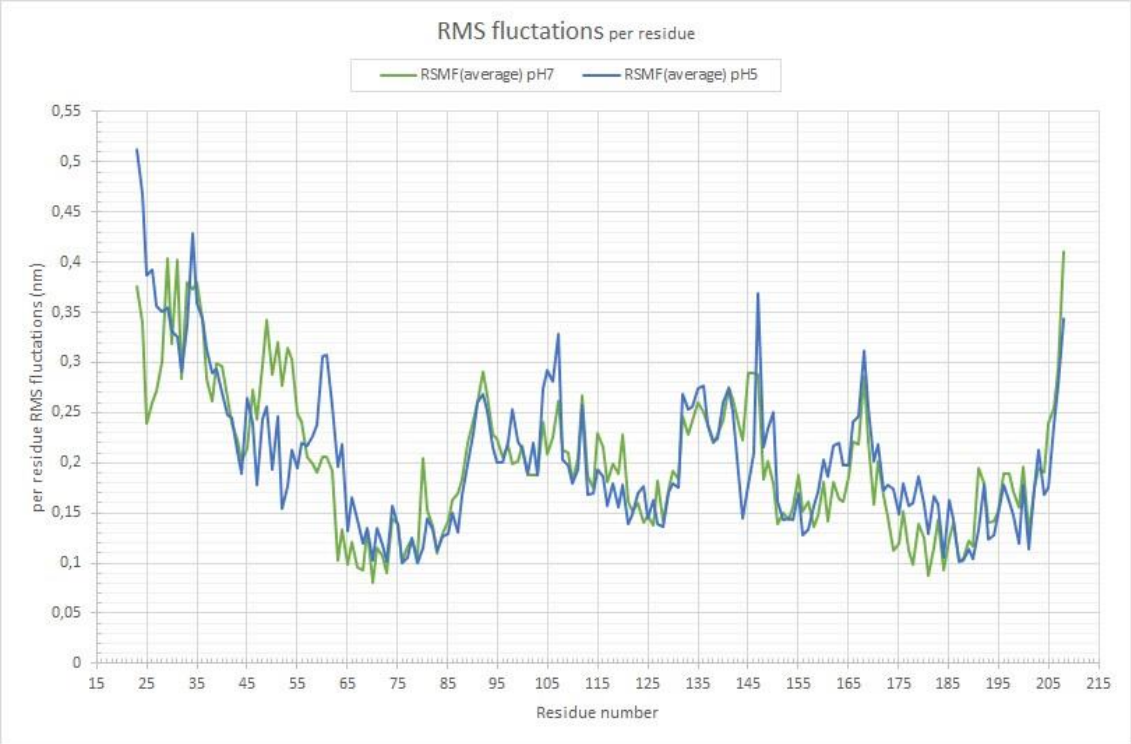



Εικόνα 7. Δομή και θέση της πρωτεΐνης LHCX1 στη θυλακοειδή μεμβράνη Α) αμέσως μετά τη διαδικασία production MD Β) έπειτα από επανατοποθέτηση στο κέντρο της μεμβράνης.

Ακολουθεί η εξέταση της κινητικότητας των αμινοξέων της πρωτεΐνης μέσω της παραμέτρου Root Mean Square Fluctuations (RMSF) ανά κατάλοιπο, για τις δυο τιμές pH (~5.5 και 7.0) όπου πραγματοποιήθηκαν οι υπολογισμοί. Σκοπός της ανάλυσης είναι να ελέγξουμε εάν η κινητικότητα των αμινοξέων επηρεάζεται από την αλλαγή του pH, εάν δηλαδή η κινητικότητα ορισμένων περιοχών της πρωτεΐνης αυξάνεται όταν ενεργοποιείται ο φωτοπροστατευτικός μηχανισμός (pH ~5.5) ή ακόμα και αν έχουμε

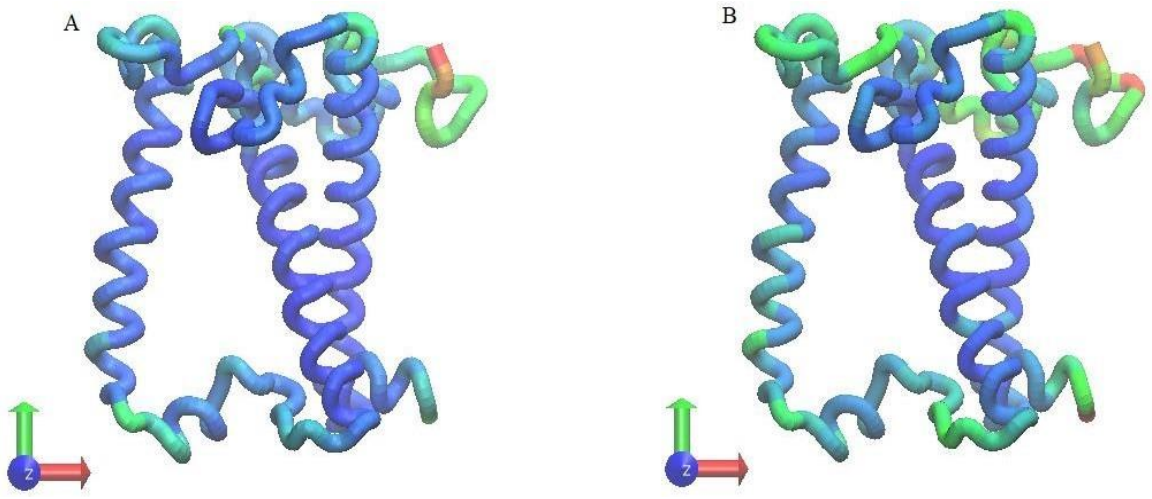
επιβράδυνση της κίνησης, από τις αλλαγές του pH.

Σημειώνεται πως η κίνηση των καταλοίπων της πρωτεΐνης συνδέεται άμεσα με τη λειτουργία της και τη μετάβαση μεταξύ διαμορφώσεων. Αποτέλεσμα της ανάλυσης RMSF, είναι το παρακάτω διάγραμμα (εικόνα 8) στο οποίο μπορούμε να εντοπίσουμε πως διαφοροποιείται η κινητικότητα της πρωτεΐνης LHCX1 εάν αλλάξουμε την τιμή του pH σε ορισμένες περιοχές.




*Εικόνα 8.* Μέση τετραγωνική ρίζα διακύμανσης (root mean square fluctuation RMSF), της LHCX1, ανά κατάλοιπο για pH 7 (πράσινο χρώμα) και pH 5 (μπλε χρώμα).

Μπορούμε να υποθέσουμε πως ως γενική εικόνα η πρωτεΐνη είναι αρκετά ευλύγιστη, λόγω των υψηλών τιμών RMSF, και ακολουθεί το ίδιο μοτίβο κίνησης και στις δυο περιπτώσεις pH. Δεν μπορούμε να βγάλουμε ένα γενικό συμπέρασμα για την κινητικότητα σε pH ~5.5 σε σχέση με την κινητικότητα σε pH 7.0. Θα πρέπει να ερευνηθεί αυτή η κινητικότητα ανά περιοχή. Εάν οπτικοποιήσουμε το παραπάνω διάγραμμα στη δομή της LHCX1, θα πάρουμε το αποτέλεσμα που φαίνεται στην εικόνα 9, όπου χρησιμοποιείται ένας χρωματικός κώδικας με μεγάλη κλίμακα κινητικότητας (κόκκινες περιοχές), μεσαία (κίτρινες-πράσινες περιοχές) και μικρή (μπλε περιοχές).




*Εικόνα 9.* Χρωματισμός καταλοίπων της πρωτεΐνης LHCX1 συμφώνα με την κινητικότητα τους Α) σε pH 5 και Β) σε pH 7. Τα κατάλοιπα που έχουν μεγαλύτερη κινητικότητα απεικονίζονται με κόκκινο χρώμα ενώ αυτά με τη μικρότερη με μπλε χρώμα σε μια παλέτα από το μπλε στο πράσινο, κίτρινο, πορτοκαλί και τέλος στο κόκκινο.

Η γενική εικόνα δείχνει πως οι α-έλικες δεν επηρεάζονται ιδιαίτερα από την αλλαγή στο pH, εκτός από περιπτώσεις που τυχαίνει τα κατάλοιπα να βρίσκονται κοντά σε πτυχώσεις (coils) ή στροφές (turns). Στα σημεία εκείνα φαίνεται πως η επιρροή της αλλαγής του pH παίζει ρόλο. Εξετάζοντας συνδυαστικά το διάγραμμα και τη δομή μπορούμε να εντοπίσουμε τις περιοχές με την περισσότερη κινητικότητα. Σημαντική διαφορά στην κινητικότητα εμφανίζεται στο C-τελικό άκρο της πρωτεΐνης κατά την αλλαγή pH.

Για pH ~5.5, βάζοντας τα κατάλοιπα σε φθίνουσα σειρά, όσον αφορά τη κινητικότητα (εικόνα 9Α), έχουμε με κόκκινο χρώμα το 23, με πορτοκαλί το 24, ακολουθούν τα 25-37 (με το 34 να έχει λίγο μεγαλύτερη κινητικότητα), τα 147-148, το 208, τα 106-107, τα 168-169, τα 60-63. Τα κατάλοιπα που δεν έχουν αναφερθεί έχουν πολύ μικρές τιμές RMSF. Αντίστοιχα για pH ~7.0 (εικόνα 9Β), με κόκκινο χρώμα απεικονίζονται το 208, τα 28-29, τα 30-31, ακολουθούν με πορτοκαλί τα 23, 33-36 και έπειτα τα 47-54, τα 91-93, τα 144-148 και τέλος τα 167-168.

Συγκρίνοντας τη δυναμική της LHCX1 σε διαφορετικές τιμές pH, μπορούμε να βγάλουμε το συμπέρασμα πως η LHCX1 γίνεται περισσότερο σταθερή, ειδικότερα όσο πλησιάζει σε συνθήκες φωτοπροστασίας (pH ~5.5). Στις συνθήκες αυτές βλέπουμε πως είναι σχετικά σταθερή και δεν έχει πολλά τμήματα τα οποία κινούνται ταχύτερα σε σχέση με άλλα. Θα μπορούσαμε να συμπεράνουμε πως η LHCX1 'κλειδώνει' σε μια συγκεκριμένη διαμόρφωση, σε pH ~5.5. Αντίθετα, όταν βρίσκεται σε ουδέτερες συνθήκες (pH ~7.0), όλα τα κατάλοιπα που ανήκουν σε πτυχώσεις και στροφές κινούνται πολύ γρηγορότερα από αυτά που βρίσκονται στις α-έλικες. Ενώ και ορισμένες α-έλικες παρουσιάζουν αυξημένη κινητικότητα.

Αφού έχουμε χαρακτηρίσει πλήρως δομικά και δυναμικά την LHCX1 σε διαφορετικές τιμές pH, μπορούμε να προχωρήσουμε στην πρόβλεψη της αλληλεπίδρασης LHCX1-FCP στο διάτομο P. tricornutum με βάση τη λυμένη δομή της FCP στη βιβλιογραφία (pdb: 6A2W, (Wang, 2019)).

### 3.2 Το πρωτεϊνικό σύμπλοκο LHCX1-FCP

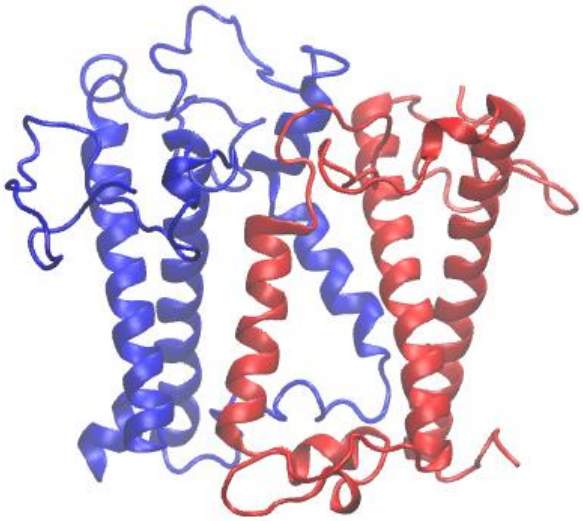
Για να μπορέσουμε να εξετάσουμε τις αλληλεπιδράσεις μεταξύ των δυο πρωτεϊνών πρέπει να τις τοποθετήσουμε μέσα στη θυλακοειδή μεμβράνη στην κατάλληλη διάταξη. Για τον λόγο αυτό, πρέπει να πραγματοποιηθεί η διαδικασία συμπλοκοποίησης- δέσμευσης (*docking*), με την οποία υπολογίζουμε και προβλέπουμε τον καταλληλότερο προσανατολισμό ενός μορίου σε σχέση με ένα άλλο, όταν σχηματίζουν ένα σταθερό σύμπλοκο.

Η δομή της FCP έχει λυθεί πειραματικά με ακτίνες X και η θέση των ατόμων της στο χώρο είναι αποτέλεσμα μέσων τιμών από ένα μεγάλο σύνολο δομών στον κρύσταλλο (*ensemble average*). Προτού ξεκινήσουμε τη διαδικασία του docking, θα πρέπει να προσδιορίσουμε μια αντίστοιχη δομή της LHCX1 με τις θέσεις των ατόμων της να αντιστοιχούν σε ένα μεγάλο σύνολο δομών (*time-average* από τη Μοριακή Δυναμική). Ενώνοντας τις πέντε τροχιές οι οποίες προέκυψαν από τη διαδικασία παραγωγής της Μοριακής Δυναμικής στην LHCX1, έχουμε ένα ικανό δείγμα δομών (2.5μs). Στη συνέχεια ομαδοποιούμε τις διαφορετικές δομές σε αυτό το σύνολο σε *clusters*, ώστε να επιλέξουμε τις πιο αντιπροσωπευτικές δομές για να συνεχίσουμε στο *docking*. Από την κατηγοριοποίηση των δομών, προέκυψαν τρεις αντιπροσωπευτικές δομές για pH ~5.5 και πέντε για pH ~7.0.

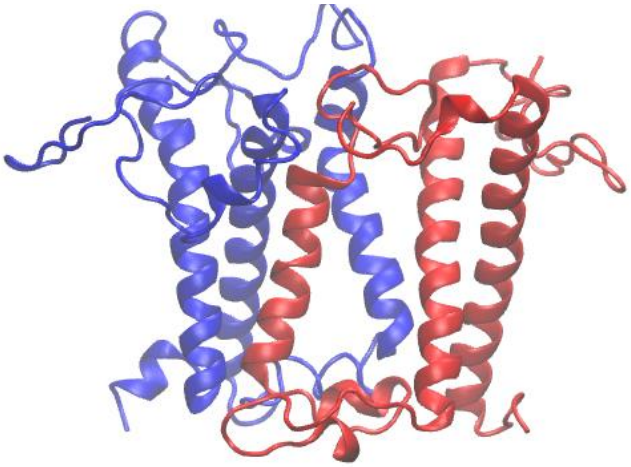
Για τη διαδικασία του *docking* χρησιμοποιήσαμε τον αλγόριθμο HADDOCK (van Zundert, 2016), στο οποίο πρέπει να ορίσουμε ενεργές και ανενεργές περιοχές για αλληλεπίδραση στις δύο πρωτεΐνες χωριστά. Οι ενεργές περιοχές αποτελούνται από αμινοξέα που συμμετέχουν στην αλληλεπίδραση των δυο πρωτεϊνών, βρίσκονται στην επιφάνεια διεπαφής τους και περιφερειακά, ενώ οι ανενεργές αποτελούνται από κατάλοιπα τα οποία δεν συμμετέχουν στην αλληλεπίδραση και βρίσκονται σε επαφή με τις υδάτινες φάσεις της θυλακοειδούς μεμβράνης ή τα λιπίδια.

Συνολικά δημιουργήσαμε οχτώ ζεύγη δεδομένων εισαγωγής, για σύμπλοκα FCP-LHCX1 χρησιμοποιώντας τις οχτώ αντιπροσωπευτικές δομές από LHCX1, και την κρυσταλλική δομή της FCP. Δημιουργήθηκαν πάνω από 80 σύμπλοκα-προβλέψεις (για κάθε συνδυασμό), ωστόσο σε pH ~7.0 βρέθηκε μόλις μια δομή η οποία είχε κατάλληλο προσανατολισμό και θέση (εικόνα 10), ενώ σε pH ~5.5 βρέθηκαν τρεις δομές, από τις οποίες επιλέχθηκε η αρτιότερη (εικόνα 11). Σε αυτό το σημείο θα πρέπει να επισημάνουμε πως ψάχνουμε δομές με σωστό προσανατολισμό (στρώμα-lumen) ταυτόχρονα και για τις δύο πρωτεΐνες όταν το σύμπλοκο τοποθετηθεί στη θυλακοειδή μεμβράνη. Γνώμονα έχουμε την αλληλεπίδραση των υδρόφοβων αμινοξέων με τα

θυλακοειδή λιπίδια, ενώ τα υδρόφιλα θα πρέπει να βρίσκονται στις υδάτινες φάσεις στις δύο πλευρές της θυλακοειδούς μεμβράνης. Πλήρως συμβατή με το κριτήριο αυτό είναι μόνο μία δομή που παρουσιάζεται στην εικόνα 11.



*Εικόνα 10.* Αποτέλεσμα της διαδικασίας της συμπλοκοποίησης (docking) των πρωτεϊνών LHCX1 (μπλε χρώμα) και FCP (κόκκινο χρώμα) σε τιμή pH 7.



*Εικόνα 11*. Αποτέλεσμα της διαδικασίας της συμπλοκοποίησης (docking) των πρωτεϊνών LHCX1 (μπλε χρώμα) και FCP (κόκκινο χρώμα) σε τιμή pH 5.

Εφόσον έχουμε πλέον καταλήξει στο σύμπλοκο FCP-LHCX1 (εικόνα 11) που μπορούμε να χρησιμοποιήσουμε, θα πρέπει να προσθέσουμε πρωτόνια στο σύμπλοκο FCP-LHCX1 για τις διαφορετικές τιμές pH (~5.5, ~7.0) που εξετάζουμε στην εργασία αυτή.

### 3.3 Πρωτονίωση του πρωτεϊνικού συμπλόκου LHCX1-FCP

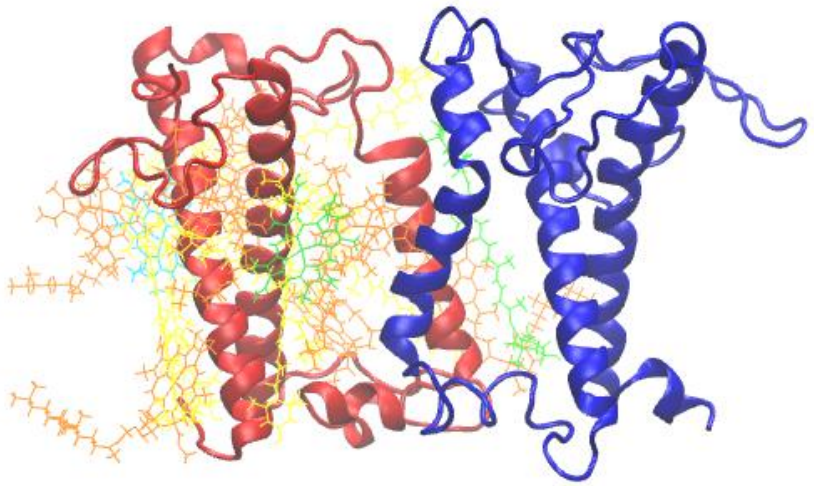
Για να μπορέσουμε να προσθέσουμε πρωτόνια όπου αυτό κρίνεται απαραίτητο, χρησιμοποιήσαμε και πάλι το εργαλείο PDB2PQR για να βρούμε τα σημεία κορεσμού οξέος (pKa) του κάθε γλουταμικού (GLU), ασπαρτικού (ASP) και ιστιδίνης (HIS) στο σύμπλοκο.

Όσον αφορά την LHCX1, τα Asp (D) και Glu (E) είναι αποπρωτονιωμένα, εκτός των E-65 και E-85 τα οποία είναι πρωτονιωμένα σε όλες τις περιπτώσεις. Ωστόσο, όπως αναφέρθηκε, σύμφωνα με τον Ballottari (2016), υπάρχουν αμινοξέα τα οποία επηρεάζονται από την αλλαγή του pH και έχουν σημαντικό ρόλο στη φωτοπροστασία. Στην περίπτωση του συμπλόκου FCP-LHCX1 λαμβάνουμε υπόψη μας περισσότερες περιπτώσεις πρωτονίωσης, καθώς έχει αναφερθεί στη βιβλιογραφία πως η δημιουργία συμπλόκου FCP-LHCX1 οδηγεί σε αλλαγές στις τιμές pKa των καταλοίπων των δύο πρωτεϊνών (Giovagnetti, 2018). Λαμβάνουμε υπόψη τις περιπτώσεις: α) για τιμή pH ~5.5 τα D-95/ E-205 είναι πρωτονιωμένα, β) για τιμή pH ~5.5 το D-95 είναι πρωτονιωμένο ενώ το E-205 αποπρωτονιωμένο, γ) για τιμή pH ~5.5 το D-95 είναι αποπρωτονιωμένο ενώ το E-205 είναι πρωτονιωμένο και τέλος, δ) για τιμή pH 7.0 όπου και τα δύο είναι αποπρωτονιωμένα.

Αντίστοιχη είναι η διαδικασία και για την FCP, στην οποία όμως η διαδικασία είναι πιο ξεκάθαρη. Έχουμε μόλις δυο περιπτώσεις α) έκτος από τα D64, E-54, E-72, E-82, E-158, τα οποία φαίνεται να έχουν ιδιαίτερη σημασία στην φωτοπροστασία (Nagao 2020, Wang, 2019), όλα τα υπόλοιπα Asp και Glu είναι αποπρωτονιωμένα για pH ~5.5. και β) όλα τα Asp και Glu είναι αποπρωτονιωμένα για pH 7.0. Οπότε σύμφωνα με τις πρωτονιώσεις των αμινοξέων, έχουμε ένα σύμπλοκο σε pH 7 και τρία σε pH ~5.5.

Ωστόσο η κρυσταλλική δομή της FCP περιλαμβάνει και χρωστικές οι οποίες με τη σειρά τους θα πρέπει να πρωτονιωθούν, να τις τοποθετήσουμε στην πρωτεΐνη και έχοντας πλέον το σύστημα χρωστικες-FCP-LHCX1 να το τοποθετήσουμε στη μεμβράνη. Οι χρωστικές που υπάρχουν στην FCP και μας ενδιαφέρουν για την αλληλεπίδραση των δυο πρωτεϊνών είναι: χλωροφύλλη a (Chl-a), χλωροφύλλη C1 και 2 (KC1, KC2), διαδινοξανθίνη (DD6) και φουκοξανθίνη (A86). Με τη βοήθεια του Maestro (Schrödinger Release 2020-3: Maestro, Schrödinger, LLC, New York, NY, 2020), το οποίο είναι ένα λογισμικό μοριακής αναπαράστασης, διορθώνουμε τους δεσμούς μεταξύ των ατόμων όπου αυτό κρίνεται απαραίτητο, προσθέτουμε τα πρωτόνια στις χρωστικές και εφόσον οι χρωστικές είναι πλέον ολοκληρωμένες, τις προσθέτουμε στη δομή της

FCP και το σύστημα μας πλέον είναι ολοκληρωμένο (εικόνα 12). Θα πρέπει να σημειώσουμε πως πολλές φορές η ποιότητα του κρυστάλλου για την επίλυση μιας πρωτεϊνικής δομής οδηγεί σε μη ιδανικές θέσεις ή απουσία κάποιων ατόμων στις χρωστικές (π.χ. υδρογόνα), με αποτέλεσμα να πρέπει να ακολουθηθεί η διαδικασία που μόλις περιγράψαμε.

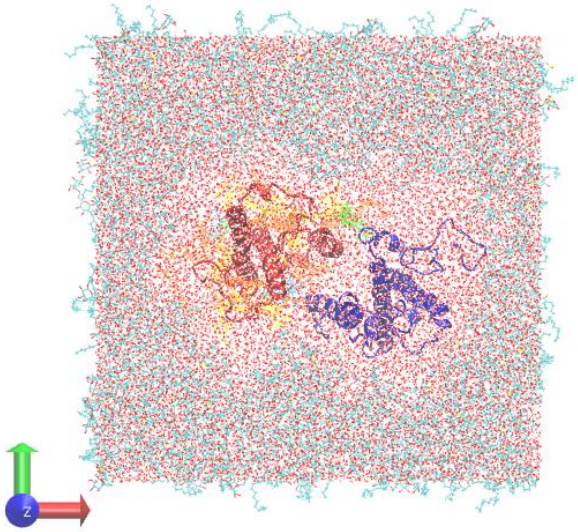


*Εικόνα 12.* Δομή των πρωτεϊνών LHCX1 (μπλε), FCP (κόκκινο) και των χρωστικών CLA (πορτοκάλι), KC1 και KC2 (γαλάζιο), DD6 (πράσινο), A86 (κίτρινο), μετά την προσθήκη πρωτονίων, σε pH 5.

### 3.4 Τοποθέτηση συμπλόκου στην θυλακοειδή μεμβράνη

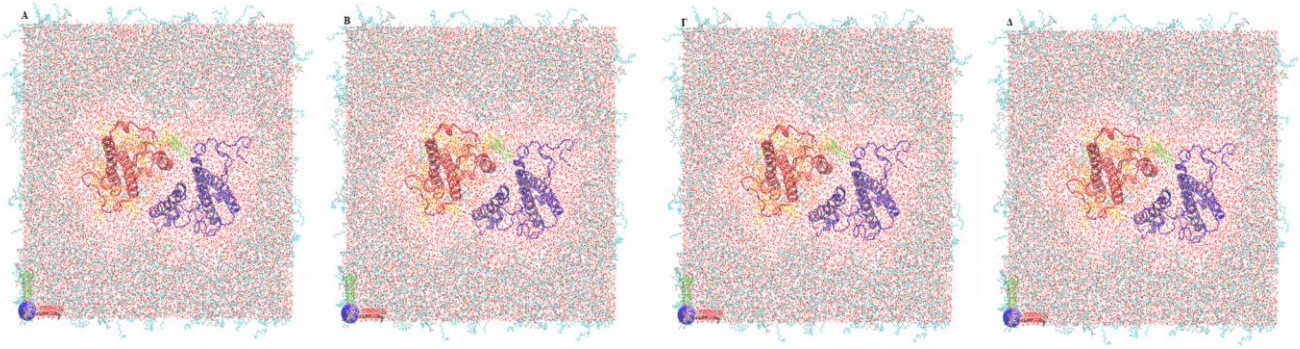
Εφόσον οι δομές των πρωτεϊνών είναι πλέον έτοιμες προς χρήση, το μόνο που απομένει είναι να «συναρμολογήσουμε» το σύμπλοκο εντός της θυλακοειδούς μεμβράνης. Από τα αποτελέσματα του docking, έχουμε την σχετική θέση των πρωτεϊνών, οπότε μπορούμε να τις τοποθετήσουμε στη μεμβράνη, έχοντας σαν σημεία αναφοράς τα κέντρα μάζας της μεμβράνης και του συμπλόκου πρωτεϊνών, τα οποία θα αντιστοιχίσουμε μεταξύ τους.

Αφού τοποθετήσουμε το σύστημα στη μεμβράνη, θα πρέπει να αφαιρέσουμε τα λιπίδια τα οποία αλληλοκαλύπτονται με κατάλοιπα αμινοξέων, σε απόσταση μέχρι 1.6 Άνγκστρομ (Å). Το αποτέλεσμα όλης της παραπάνω διαδικασίας φαίνεται στην εικόνα 13, με μπλε χρώμα απεικονίζεται η LHCX1, με κόκκινο η FCP, και οι χρωστικές, με πράσινο, ενώ έχει δημιουργηθεί ένας ξεκάθαρος κενός χώρος ώστε το σύστημα να είναι τοποθετημένο σωστά, χωρίς να υπάρχουν αλληλεπικαλύψεις μεταξύ ατόμων, που θα οδηγούσαν σε ανάπτυξη ισχυρότατων δυνάμεων μεταξύ των ατόμων αυτών και αστάθεια της προσομοίωσης.



*Εικόνα 13.* Δομή των πρωτεϊνών LHCX1 (μπλε), FCP(κόκκινο) και των χρωστικών CLA (πορτοκάλι), KC1 και KC2 (γαλάζιο), DD6 (πράσινο), A86 (κίτρινο), μέσα στη θυλακοειδή μεμβράνη, σε pH 5, μετά τον καθαρισμό λιπιδίων, νερού και ιόντων νατρίου, και τη προσθήκη πρωτονίων, σε 1,6 Å.

Φυσικά όλη η διαδικασία έγινε τέσσερις φορές, όσοι και οι συνδυασμοί στις τιμές pH που εξετάζουμε, επομένως τα αποτελέσματα ήταν τέσσερα συστήματα παρόμοιας μορφής (εικόνα 14).




*Εικόνα 14.* Δομή των πρωτεϊνών LHCX1 (μπλε), FCP(κόκκινο) και των χρωστικών CLA(πορτοκάλι), KC1 και KC2(γαλάζιο), DD6(πράσινο), A86 (κίτρινο), μέσα στη θυλακοειδή μεμβράνη, μετά τον καθαρισμό λιπιδίων, νερού και ιόντων νατρίου, και τη προσθήκη πρωτονίων, σε 1,6 Å, για Α) pH ~5.5 όπου D-95 και E-205 είναι πρωτονιωμένα, Β) pH ~5.5 όπου το D-95 είναι πρωτονιωμένο ενώ το E-205 αποπρωτονιωμένο, Γ) pH ~5.5 όπου το D-95 είναι αποπρωτονιωμένο ενώ το E-205 είναι πρωτονιωμένο και τέλος, Δ) pH 7 όπου D-95 και E-205 είναι αποπρωτονιωμένα.

### 3.5 Δυναμική του συμπλόκου πρωτεϊνών LHCX1 και FCP

Αντίστοιχα με την LHCX1, χρησιμοποιήσαμε το GROMACS για τη δυναμική προσομοίωση του συστήματος μας και να εξετάσουμε τυχόν σφάλματα στη διαδικασία ανάπτυξης του μοντέλου μας. Ακολουθούμε και πάλι την ίδια διαδικασία, που περιλαμβάνει τα στάδια ελαχιστοποίησης ενέργειας, εξισορρόπησης συστήματος, παραγωγή τροχιών μοριακής δυναμικής.

#### 3.5.1 Ελαχιστοποίηση ενέργειας και εξισορρόπηση συστήματος

Η ελαχιστοποίηση ενέργειας πραγματοποιείται διότι η δομή με την χαμηλότερη δυναμική ενεργεία, είναι μια σταθερή δομή της πρωτεΐνης στη φύση, σε συνθήκες απόλυτου μηδενός (Κούτσελος, 2015). Σε περίπτωση που ξεκινούσαμε από οποιαδήποτε άλλη δομή, θεωρητικά δεν θα μπορούσαμε να εξασφαλίσουμε πως η δυναμικής της πρωτεΐνης που υπολογίζεται δίνει ένα σύνολο διαμορφώσεων της πρωτεΐνης που απαντώνται στις πειραματικές διαδικασίες.

Ξεκινώντας από την ελαχιστοποίηση ενέργειας στο σύστημα (εικόνα 15), ακολουθεί η εξισορρόπηση των διατομικών δυνάμεων/ θερμοκρασίας και πίεσης μέσα από διαδοχικά βήματα.

```
; LINES STARTING WITH ';' ARE COMMENTS
; Parameters describing what to do, when to stop and what to save

integrator      = steep         ; Algorithm options
define          = -DPOSRES
                                ; steep = steepest descent minimization
                                ; MD = Leap Frog algorith for integrating Newton (equations of motion)
emtol           = 5.0           ; Stop minimization when the energy changes by less than emtol kJ/mol.
emstep          = 0.01          ; initial step-size
nsteps          = 1000          ; Maximum number of (minimization) steps to perform
nstenergy       = 10            ; Write energies to disk every nstenergy steps
nstxout-compressed = 0          ; Which coordinate group(s) to write to disk
compressed-x-grps  = System     ; Which energy group(s) to write to disk

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions
cutoff-scheme   = Verlet
nstlist         = 10            ; Frequency to update the neighbor list and long range forces
ns_type         = grid          ; Method to determine neighbor list (simple, grid)
rlist           = 1.2           ; Cut-off for making neighbor list (short range forces)
coulombtype     = PME           ; Treatment of long range electrostatic interactions
fourierspacing  = 0.15          ; Grid Spacing for PME
coulomb-modifier = Potential-shift-Verlet
; Dispersion correction
DispCorr        = Ener          ; account for cut-off vdW scheme
rcoulomb        = 1.2           ; long range electrostatic cut-off
vdwtype         = Cut-off       ; Cut-off Vdw
vdw-modifier    = Potential-Switch  ; potential tapering...
rvdw-switch     = 0.8           ; potential tapering start at...
rvdw            = 1.2
constraints     = h-bonds       ; Bond types to replace by constraints
lincs-order     = 8             ; lincs order (>8 for cg)
constraint-algorithm = Lincs    ; Shake or Lincs
pbc             = xyz           ; Periodic Boundary Conditions (yes/no)
```

*Εικόνα 15.* Παράμετροι για την εκκίνηση ελαχιστοποίησης ενέργειας μέσω GROMACS.

Η εξισορρόπηση του συστήματος ξεκινά ανεβάζοντας τη θερμοκρασία σε 100 K και αφαιρώντας τη σύζευξη με την πίεση (εικόνα 16), στο επόμενο στάδιο κρατάω σταθερή τη θερμοκρασία στους 100 K και αυξάνω το χρονικό βήμα αριθμητικής επίλυσης διαφορικών εξισώσεων κίνησης για τη δημιουργία τροχιών (εικόνα 17), έπειτα κρατώντας την θερμοκρασία στους 100 K, αυξάνω την πίεση σε 1 atm (εικόνα 18). Αφού έχουμε θέσει μια αρχική τιμή και για τις δυο παραμέτρους που εξετάζουμε, πλέον ανεβάζουμε τη θερμοκρασία σε 250 K και αφαιρούμε τη σύζευξη με την πίεση (εικόνα 19) και τέλος δίνουμε στη θερμοκρασία την τελική της τιμή η οποία είναι 310 K και διατηρούμε την πίεση σε 1 atm (εικόνα 20). Αφού δώσουμε στο σύστημα τις κατάλληλες τιμές πίεσης και θερμοκρασίας, αφαιρούμε σταδιακά τις δυνάμεις από τα Ca άτομα της πρωτεΐνης, εώς ότου η κίνηση τους χαλαρώσει μέχρι το επιθυμητό στάδιο (εικόνα 21-24).

```
; This is where you chose dynamics e.g. md-vv, bd, cg
integrator               = bd
define                   = -DPOSRES
; bd friction
bd-fric                  = 0
;
; starting time for the run
tinit                    = 0
; time step, number of steps
dt                       = 0.001
nsteps                   = 50000
;
; output x, v to coordinate file
nstxout                  = 0
nstvout                  = 0
; output log - energy
nstlog                   = 25000
nstenergy                = 25000
nstcalcenergy            = 25000
; Write more info - compressed
nstxout-compressed       = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme            = Verlet
ns-type                  = grid
pbc                      = xyz
;
; Electrostatics
coulombtype              = PME            ; Treatment of long range electrostatic interactions
coulomb-modifier         = Potential-shift-Verlet
fourierspacing           = 0.15          ; Grid Spacing for PME
rcoulomb                 = 1.2
rlist                    = 1.2
ewald-rtol               = 1e-5
;
; Vdw
```

```
; Vdw
vdwtype                  = Cut-off       ; Cut-off Vdw
vdw-modifier             = Potential-Switch
rvdw-switch              = 0.8           ; potential tapering start at...
rvdw                     = 1.2
ewald-rtol-lj            = 1e-3
; Dispersion correction
DispCorr        = Ener  ; account for cut-off vdw scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tc-grps                  = system
tcoupl                   = berendsen
tau-t                    = 0.001
ref-t                    = 100
; P coupling no, berendsen, Parrinello-Rahman, MTTK
Pcoupl                   = no
pcoupltype               = isotropic
tau-p                    = 0.001
ref-p                    = 1.0
compressibility          = 4.5e-5
;
; Generate Velocities
gen-vel                  = yes
gen-temp                 = 100
gen-seed                 = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINCS, SHAKE
constraints              = h-bonds       ; Bond types to replace by constraints
constraint-algorithm     = LINCS
morse                    = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
nstcomm         = 1
comm-mode       = Linear
comm-grps       = System
```

*Εικόνα 16.* Πρώτο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS, ανεβάζουμε τη θερμοκρασία σε 100K, χωρίς σύζευξη με πίεση

```
; This is where you chose dynamics e.g. md-vv, bd, cg
define                   = -DPOSRES
integrator               = md-vv
;
annealing                  = no
;
; starting time for the run
tinit                    = 0
; time step, number of steps
dt                       = 0.001
nsteps                   = 100000
; remove center of mass rotation and translation
; comm-mode                  = Angular
;
; output x, v to coordinate file
nstxout                  = 0
nstvout                  = 0
; output log - energy
nstlog                   = 50000
nstenergy                = 50000
nstcalcenergy            = 50000
; Write more info - compressed
compressed-x-grps        = Protein
nstxout-compressed       = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme            = Verlet
ns-type                  = grid
pbc                      = xyz
;
; Electrostatics
coulombtype              = PME           ; Treatment of long range electrostatic interactions
coulomb-modifier         = Potential-shift-Verlet
rcoulomb                 = 1.2
rlist                    = 1.2
ewald-rtol               = 1e-5
fourierspacing           = 0.15          ; Grid Spacing for PME
```

```
; Vdw
vdwtype                  = Cut-off       ; Cut-off Vdw
vdw-modifier             = Potential-Switch
rvdw-switch              = 0.8           ; potential tapering start at...
rvdw                     = 1.2
ewald-rtol-lj            = 1e-3
; Dispersion correction
DispCorr        = Ener          ; account for cut-off vdw scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tc-grps                  = system
tcoupl                   = berendsen
tau-t                    = 0.1
ref-t                    = 100
; P coupling no, berendsen, Parrinello-Rahman, MTTK
Pcoupl                   = no
pcoupltype               = isotropic
tau-p                    = 50.0
ref-p                    = 1.0
compressibility          = 4.5e-5
;
; Generate Velocities
gen-vel                  = yes
gen-temp                 = 100
gen-seed                 = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINCS, SHAKE
constraints              = h-bonds       ; Bond types to replace by constraints
constraint-algorithm     = LINCS
morse                    = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
nstcomm         = 1
comm-mode       = Linear
comm-grps       = Protein lipids SOL_K
```

*Εικόνα 17.* Δεύτερο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS, κρατάμε τις ίδιες συνθήκες πίεσης και θερμοκρασίας, αυξάνουμε το βήμα.

```
; This is where you chose dynamics e.g. md-vv, bd, cg
define                  = -DPOSRES
integrator              = md-vv
;
;Annealing no, single
annealing                 = no
;
; starting time for the run
tinit                   = 0
; time step, number of steps
dt                      = 0.001
nsteps                  = 100000
; remove center of mass rotation and translation
; comm-mode               = Angular
;
; output x, v to coordinate file
nstxout                 = 0
nstvout                 = 0
; output log - energy
nstlog                  = 50000
nstenergy               = 50000
nstcalcenergy           = 50000
; Write more info - compressed
compressed-x-grps       = Protein
nstxout-compressed      = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme           = Verlet
ns-type                 = grid
pbc                     = xyz
;
; Electrostatics
coulombtype             = PME            ; Treatment of long range electrostatic interactions
coulomb-modifier        = Potential-shift-Verlet
rcoulomb                = 1.2
rlist                   = 1.2
ewald-rtol              = 1e-5
fourierspacing          = 0.15          ; Grid Spacing for PME
```

```
; VdW
vdwtype                 = Cut-off       ; Cut-off Vdw
vdw-modifier            = Potential-Switch
rvdw-switch             = 0.8           ; potential tapering start at...
rvdw                    = 1.2
ewald-rtol-lj           = 1e-3
; Dispersion correction
DispCorr        = EnerPres      ; account for cut-off vdW scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                  = berendsen
tau-t                   = 0.1
ref-t                   = 100
; P coupling no, berendsen, Parrinello-Rahman, MTTK
tc-grps                 = system
Pcoupl                  = berendsen
pcoupltype              = semiisotropic
tau-p                   = 50.0
ref-p                   = 1.0 1.0
compressibility         = 4.6e-5 4.5e-5
refcoord-scaling        = all
;
; Generate Velocities
gen-vel                 = yes
gen-temp                = 100
gen-seed                = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINS, SHAKE
constraints             = h-bonds       ; Bond types to replace by constraints
constraint-algorithm    = LINCS
morse                   = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
nstcomm         = 1
comm-mode       = Linear
comm-grps       = Protein lipids SOL_K
```

*Εικόνα 18.* Τρίτο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS, κρατάμε τη θερμοκρασία σε 100 K και διατηρούμε την πίεση σε 1 atm.

```
; This is where you chose dynamics e.g. md-vv, bd, cg
define                  = -DPOSRES
integrator              = md-vv
;
;Annealing no, single
annealing                 = single
annealing-npoints       = 9
annealing-time          = 0 250 750 1000 1500 1750 2250 2500 3000
annealing-temp          = 100 150 150 200 200 230 230 250 250
;
; starting time for the run
tinit                   = 0
; time step, number of steps
dt                      = 0.001
nsteps                  = 3000000
; remove center of mass rotation and translation
; comm-mode               = Angular
;
; output x, v to coordinate file
nstxout                 = 0
nstvout                 = 0
; output log - energy
nstlog                  = 500000
nstenergy               = 500000
nstcalcenergy           = 500000
; Write more info - compressed
compressed-x-grps       = Protein
nstxout-compressed      = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme           = Verlet
ns-type                 = grid
pbc                     = xyz
;
; Electrostatics
coulombtype             = PME            ; Treatment of long range electrostatic interactions
coulomb-modifier        = Potential-shift-Verlet
rcoulomb                = 1.2
rlist                   = 1.2
ewald-rtol              = 1e-5
fourierspacing          = 0.15          ; Grid Spacing for PME
```

```
;
; VdW
vdwtype                 = Cut-off       ; Cut-off Vdw
vdw-modifier            = Potential-Switch
rvdw-switch             = 0.8           ; potential tapering start at...
rvdw                    = 1.2
ewald-rtol-lj           = 1e-3
; Dispersion correction
;DispCorr       = EnerPres      ; account for cut-off vdW scheme
DispCorr                = Ener          ; apply long range dispersion corrections for
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                  = v-rescale
tau-t                   = 0.1
ref-t                   = 250
; P coupling no, berendsen, Parrinello-Rahman, MTTK
tc-grps                 = system
Pcoupl                  = no
pcoupltype              = isotropic
tau-p                   = 2.0
ref-p                   = 1.0
compressibility         = 4.5e-5
;
; Generate Velocities
gen-vel                 = yes
gen-temp                = 100
gen-seed                = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINS, SHAKE
constraints             = all-bonds     ; Bond types to replace by constraints
constraint-algorithm    = LINCS
morse                   = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
nstcomm         = 1
comm-mode       = Linear
comm-grps       = Protein lipids SOL_K
```

*Εικόνα 19.* Τέταρτο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS, αφαιρούμε τη σύζευξη με την πίεση, ανεβάζουμε τη θερμοκρασία σε 250K .

```
; This is where you chose dynamics e.g. md-vv, bd, cg
define                  = -DCa_POSRES
integrator              = md-vv
;
;Annealing no, single
annealing               = single
annealing-npoints       = 6
annealing-time          = 0 250 375 650 750 2000
annealing-temp          = 250 250 275 275 310 310
;
; starting time for the run
tinit                   = 0
; time step, number of steps
dt                      = 0.002
nsteps                  = 1000000
; remove center of mass rotation and translation
; comm-mode               = Angular
;
; output x, v to coordinate file
nstxout                 = 0
nstvout                 = 0
; output log - energy
nstlog                  = 200000
nstenergy               = 200000
nstcalcenergy           = 200000
; Write more info - compressed
compressed-x-grps       = Protein
nstxout-compressed      = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme           = Verlet
ns-type                 = grid
pbc                     = xyz
;
; Electrostatics
coulombtype             = PME           ; Treatment of long range electrostatic interactions
coulomb-modifier        = Potential-shift-Verlet
rcoulomb                = 1.2
rlist                   = 1.2
ewald-rtol              = 1e-5
fourierspacing          = 0.15          ; Grid Spacing for PME
;
; VdW
vdwtype                 = Cut-off       ; Cut-off Vdw
vdw-modifier            = Potential-Switch
rvdw-switch             = 0.8           ; potential tapering start at...
rvdw                    = 1.2
ewald-rtol-lj           = 1e-3
; Dispersion correction
DispCorr        = EnerPres      ; account for cut-off vdW scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                  = v-rescale
tau-t                   = 0.1
ref-t                   = 310
; P coupling no, berendsen, Parrinello-Rahman, MTTK
tc-grps                 = system
Pcoupl                  = berendsen
pcoupltype              = semiisotropic
tau-p                   = 2.0
ref-p                   = 1.0 1.0
compressibility         = 4.6e-5 4.5e-5
refcoord-scaling        = all
;
; Generate Velocities
gen-vel                 = no
gen-temp                = 250
gen-seed                = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINS, SHAKE
constraints             = all-bonds       ; Bond types to replace by constraints
constraint-algorithm    = LINCS
morse                   = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
nstcomm         = 1
comm-mode       = Linear
comm-grps       = Protein lipids SOL_K
```

*Εικόνα 20*. Πέμπτο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS, ανεβάζουμε τη θερμοκρασία σε 310K, διατηρούμε την πίεση σε 1 atm.

```
; This is where you chose dynamics e.g. md-vv, bd, cg
integrator              = md
define                  = -DCa_POSRES
;
;Annealing no, single
annealing               = no
;
; starting time for the run
tinit                   = 0
; time step, number of steps
dt                      = 0.002
nsteps                  = 2500000
; remove center of mass rotation and translation
; comm-mode               = Angular
;
; output x, v to coordinate file
nstxout                 = 0
nstvout                 = 0
; output log - energy
nstlog                  = 500000
nstenergy               = 500000
nstcalcenergy           = 500000
; Write more info - compressed
compressed-x-grps       = Protein_CHL
nstxout-compressed      = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme           = Verlet
ns-type                 = grid
pbc                     = xyz
;
; Electrostatics
coulombtype             = PME           ; Treatment of long range electrostatic interactions
coulomb-modifier        = Potential-shift-Verlet
rcoulomb                = 1.2
rlist                   = 1.2
fourierspacing          = 0.15          ; Grid Spacing for PME
ewald-rtol              = 1e-5
;
;
; VdW
vdwtype                 = Cut-off       ; Cut-off Vdw
vdw-modifier            = Potential-Switch
rvdw-switch             = 0.8           ; potential tapering start at...
rvdw                    = 1.2
ewald-rtol-lj           = 1e-3
; Dispersion correction
DispCorr        = EnerPres      ; account for cut-off vdW scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                  = v-rescale
tc-grps         = Protein_CHL LIPIDS    SOL_K   ; three coupling groups - more accurate
tau_t           = 0.1   0.1     0.1             ; time constant, in ps
ref_t           = 310   310     310             ; reference temperature, one for each group, in K
; P coupling no, berendsen, Parrinello-Rahman, MTTK
Pcoupl                  = berendsen
pcoupltype              = semiisotropic
tau-p                   = 2.0
ref-p                   = 1.0 1.0
compressibility         = 4.6e-5 4.5e-5
;
; Generate Velocities
gen-vel                 = no
gen-temp                = 250
gen-seed                = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINS, SHAKE
constraints             = h-bonds       ; Bond types to replace by constraints
constraint-algorithm    = LINCS
morse                   = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
nstcomm         = 1
comm-mode       = Linear
comm-grps       = Protein_CHL LIPIDS SOL_K
refcoord-scaling = all
```

*Εικόνα 21*. Έκτο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS.

```
; This is where you chose dynamics e.g. md-vv, bd, cg
integrator               = md
define                   = -DCa_POSRES
;
;Annealing no, single
annealing                = no
;
; starting time for the run
tinit                     = 0
; time step, number of steps
dt                       = 0.002
nsteps                   = 2500000
;
; output x, v to coordinate file
nstxout                  = 0
nstvout                  = 0
; output log - energy
nstlog                   = 500000
nstenergy                = 500000
nstcalcenergy            = 500000
; Write more info - compressed
nstxout-compressed       = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme            = Verlet
ns-type                  = grid
pbc                      = xyz
;
; Electrostatics
coulombtype              = PME           ; Treatment of long range electrostatic interactions
coulomb-modifier         = Potential-shift-Verlet
rcoulomb                 = 1.2
rlist                    = 1.2
ewald-rtol               = 1e-5
fourierspacing           = 0.15          ; Grid Spacing for PME
;
```

```
;
; VdW
vdwtype                  = Cut-off       ; Cut-off Vdw
vdw-modifier             = Potential-Switch
rvdw-switch              = 0.8           ; potential tapering start at...
rvdw                     = 1.2
ewald-rtol-lj            = 1e-3
; Dispersion correction
DispCorr        = EnerPres      ; account for cut-off vdW scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                   = v-rescale
tc-grps         = Protein_CHL LIPIDS SOL_K  ; three coupling groups - more accurate
tau_t           = 0.5 0.5 0.5             ; time constant, in ps
ref_t           = 310 310 310             ; reference temperature, one for each group, in K
; P coupling no, berendsen, Parrinello-Rahman, MTTK
Pcoupl                   = Parrinello-Rahman
pcoupltype               = semiisotropic ; isotropic (xyz) or semiisotropic (xy z) if you add 2 values for tau-p and compress
tau-p                    = 2.0
ref-p                    = 1.0 1.0
compressibility          = 4.6e-5 4.5e-5
; Generate Velocities
gen-vel                  = no
gen-temp                 = 310
gen-seed                 = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINCS, SHAKE
constraints              = h-bonds       ; Bond types to replace by constraints
constraint-algorithm     = Lincs
morse                    = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
;Linear, Angular, None
nstcomm         = 1
comm-mode       = Linear
comm-grps       = Protein_CHL LIPIDS SOL_K
refcoord-scaling = all
```

*Εικόνα 22*. Έβδομο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS.


```
; This is where you chose dynamics e.g. md-vv, bd, cg
integrator               = md
define                   = -DCa_POSRES
;
;Annealing no, single
annealing                = no
;
; starting time for the run
tinit                     = 0
; time step, number of steps
dt                       = 0.002
nsteps                   = 2500000
;
; output x, v to coordinate file
nstxout                  = 0
nstvout                  = 0
; output log - energy
nstlog                   = 500000
nstenergy                = 500000
nstcalcenergy            = 500000
; Write more info - compressed
nstxout-compressed       = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme            = Verlet
ns-type                  = grid
pbc                      = xyz
;
; Electrostatics
coulombtype              = PME           ; Treatment of long range electrostatic interactions
coulomb-modifier         = Potential-shift-Verlet
rcoulomb                 = 1.2
rlist                    = 1.2
ewald-rtol               = 1e-5
fourierspacing           = 0.15          ; Grid Spacing for PME
;
```

```
;
; VdW
vdwtype                  = Cut-off       ; Cut-off Vdw
vdw-modifier             = Potential-Switch
rvdw-switch              = 0.8           ; potential tapering start at...
rvdw                     = 1.2
ewald-rtol-lj            = 1e-3
; Dispersion correction
DispCorr        = EnerPres      ; account for cut-off vdW scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                   = v-rescale
tc-grps         = Protein_CHL LIPIDS SOL_K  ; three coupling groups - more accurate
tau_t           = 0.5 0.5 0.5             ; time constant, in ps
ref_t           = 310 310 310             ; reference temperature, one for each group, in K
; P coupling no, berendsen, Parrinello-Rahman, MTTK
Pcoupl                   = Parrinello-Rahman
pcoupltype               = semiisotropic ; isotropic (xyz) or semiisotropic (xy z) if you add 2 values for tau-p and compress
tau-p                    = 2.0
ref-p                    = 1.0 1.0
compressibility          = 4.6e-5 4.5e-5
; Generate Velocities
gen-vel                  = no
gen-temp                 = 310
gen-seed                 = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINCS, SHAKE
constraints              = h-bonds       ; Bond types to replace by constraints
constraint-algorithm     = Lincs
morse                    = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
;Linear, Angular, None
nstcomm         = 1
comm-mode       = Linear
comm-grps       = Protein_CHL LIPIDS SOL_K
refcoord-scaling = all
```


*Εικόνα 23*. Όγδοο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS.

```
; This is where you chose dynamics e.g. md-vv, bd, cg
integrator                = md
;
;Annealing no, single
annealing                 = no
;
; starting time for the run
tinit                     = 0
; time step, number of steps
dt                        = 0.002
nsteps                    = 5000000
;
; output x, v to coordinate file
nstxout                   = 500000
nstvout                   = 500000
; output log - energy
nstlog                    = 500000
nstenergy                 = 500000
nstcalcenergy             = 500000
; Write more info - compressed
compressed-x-grps         = System
nstxout-compressed        = 0
;
; Periodic Boundary Conditions xyz, no, xy
cutoff-scheme             = Verlet
ns-type                   = grid
pbc                       = xyz
;
; Electrostatics
coulombtype               = PME            ; Treatment of long range electrostatic interactions
coulomb-modifier          = Potential-shift-Verlet
rcoulomb                  = 1.2
rlist                     = 1.2
ewald-rtol                = 1e-5
fourierspacing            = 0.15           : Grid Spacing for PME
```

```
; VdW
vdwtype                   = Cut-off        ; Cut-off Vdw
vdw-modifier              = Potential-Switch
rvdw-switch               = 0.8            ; potential tapering start at...
rvdw                      = 1.2
ewald-rtol-lj             = 1e-3
; Dispersion correction
DispCorr        = EnerPres      ; account for cut-off vdw scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                    = v-rescale
tc-grps         = Protein_CHL LIPIDS SOL_K  ; three coupling groups - more accurate
tau_t           = 0.5 0.5 0.5             ; time constant, in ps
ref_t           = 310 310 310             ; reference temperature, one for each group, in K
; P coupling no, berendsen, Parrinello-Rahman, MTTK
Pcoupl                    = Parrinello-Rahman
pcoupltype                = semiisotropic ; isotropic (xyz) or semiisotropic (xy z) if you add 2 values for tau-p and compress
tau-p                     = 2.0
ref-p                     = 1.0 1.0
compressibility           = 4.6e-5 4.5e-5
; Generate Velocities
gen-vel                   = no
gen-temp                  = 310
gen-seed                  = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINCS, SHAKE
constraints               = h-bonds        ; Bond types to replace by constraints
constraint-algorithm      = Lincs
morse                     = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
;Linear, Angular, None
nstcomm         = 100
comm-mode       = Linear
```

*Εικόνα 24.* Ένατο βήμα ελαχιστοποίησης ενέργειας μέσω GROMACS.

Έχοντας ισορροπήσει τις διατομικές δυνάμεις το σύστημα μας και αφού έχουμε σταθεροποιήσει την πίεση και τη θερμοκρασία, μπορούμε να προχωρήσουμε στη διαδικασία παραγωγής Μοριακής Δυναμικής, για να συλλέξουμε τα δεδομένα που θα χρησιμοποιήσουμε στην ανάλυση.

### 3.5.2 Παράγωγη Μοριακής Δυναμικής

Με την παραπάνω διαδικασία της εξισορρόπησης προκύπτει μια τροχιά των ατόμων του συστήματος για 10ns, απ' όπου εξάγουμε δύο δομές που απέχουν χρονικά μεταξύ τους 5 ns, ξεκινώντας  από την εξαγωγή της πρώτης στα 5 ns. Από αυτές τις δομές, αφού πρώτα ευθυγραμμιστούν, θα ξεκινήσει η παραγωγή μοριακής δυναμικής (production MD) για 0.5μs ανά περίπτωση pH/ πρωτονίωσης.

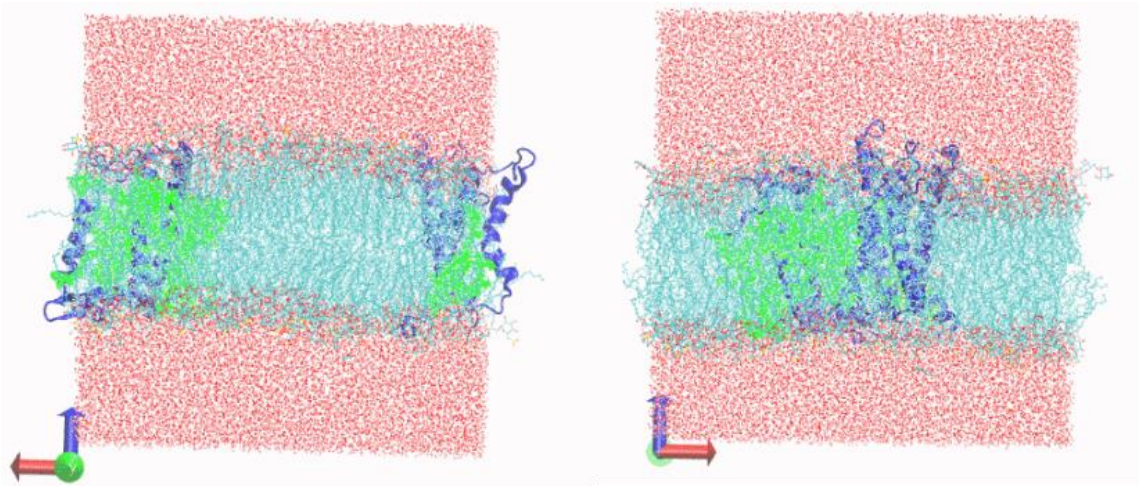
Το αρχείο παραμέτρων για το στάδιο της μοριακής δυναμικής είναι της ίδιας μορφής με το στάδιο της εξισορρόπησης, κρατώντας την πίεση σε 1atm αλλά αλλάζουμε τη θερμοκρασία σε 310 K, επίσης σώζουμε διαμορφώσεις του συμπλόκου κάθε 1ns με χρονικό όριο τα 0.5 μs (εικόνα 25).

```
 This is where you chose dynamics e.g. md-vv, bd, cg
ntegrator               = md

nnealing no, single
nnealing                = no

 starting time for the run
nit                     = 0
 time step, number of steps
t                       = 0.002
steps                   = -1

 output x, v to coordinate file
stxout                  = 50000000
stvout                  = 50000000
 output log - energy
stlog                   = 500000
stenergy                = 500000
stcalcenergy            = 500000
 Write more info - compressed
ompressed-x-grps        = System
stxout-compressed       = 500000

 Periodic Boundary Conditions xyz, no, xy
utoff-scheme            = Verlet
s-type                  = grid
bc                      = xyz

 Electrostatics
oulombtype              = PME           ; Treatment of long range electrostatic interactions
oulomb-modifier         = Potential-shift-Verlet
oulomb                  = 1.2
list                    = 1.2
wald-rtol               = 1e-5
ourierspacing           = 0.15          ; Grid Spacing for PME


j
```

```
;
; VdW
vdwtype                 = Cut-off       ; Cut-off Vdw
vdw-modifier            = Potential-Switch
rvdw-switch             = 0.8           ; potential tapering start at...
rvdw                    = 1.2
ewald-rtol-lj           = 1e-3
; Dispersion correction
DispCorr        = EnerPres      ; account for cut-off vdW scheme
;
; T coupling berendsen, nose-hoover, v-rescale
tcoupl                  = v-rescale
tc-grps         = Protein_CHL LIPIDS SOL_K  ; three coupling groups - more accurate
tau_t           = 0.5 0.5 0.5             ; time constant, in ps
ref_t           = 298 298 298             ; reference temperature, one for each group, in K
; P coupling no, berendsen, Parrinello-Rahman, MTTK
Pcoupl                  = Parrinello-Rahman
pcoupltype              = semiisotropic ; isotropic (xyz) or semiisotropic (xy z) if you add 2 values for tau-p and compress
tau-p                   = 2.0
ref-p                   = 1.0 1.0
compressibility         = 4.6e-5 4.5e-5
; Generate Velocities
gen-vel                 = no
gen-temp                = 310
gen-seed                = 173729
;
; Constraints none (explicitly in top), h-bonds, all-bonds
; Algorithms LINCS, SHAKE
constraints             = h-bonds       ; Bond types to replace by constraints
constraint-algorithm    = Lincs
morse                   = no
; COM motion removal
; These options remove motion of the protein/bilayer relative to the solvent/ions
;Linear, Angular, None
nstcomm         = 100
comm-mode       = Linear
comm-grps       = Protein_CHL LIPIDS SOL_K
```

*Εικόνα 25.* Παράμετροι για τη διαδικασία παραγωγής Μοριακής Δυναμικής, μέσω GROMACS.

Τα αποτελέσματα της συνολικής αυτής διαδικασίας που προηγήθηκε είναι ένα σύνολο από δομές του συμπλόκου των πρωτεϊνών που δηλώνουν τη δυναμική του. Ουσιαστικά μας παρουσιάζουν τις συντεταγμένες κάθε ατόμου σε μια συγκεκριμένη χρονική στιγμή. Έχουμε τρέξει όλη τη διαδικασία για δύο τιμές pH με τέσσερις διαφορετικές καταστάσεις πρωτονίωσης καταλοίπων, για κάθε τιμή παράγουμε δύο τροχιές, οπότε συνολικά εξετάζουμε 8* 0.5 μs= 4 μs της δυναμικής του LHCX1-FCP συμπλόκου.



*Εικόνα 26.* Δομή και θέση του συστήματος πρωτεϊνών LHCX1-FCP στη θυλακοειδή μεμβράνη A) αμέσως μετά τη διαδικασία production MD B) έπειτα από επανατοποθέτηση στο κέντρο της μεμβράνης, για pH ~5.5. Οι πρωτεΐνες απεικονίζονται με μπλε χρώμα, ενώ οι χρωστικές με πράσινο.

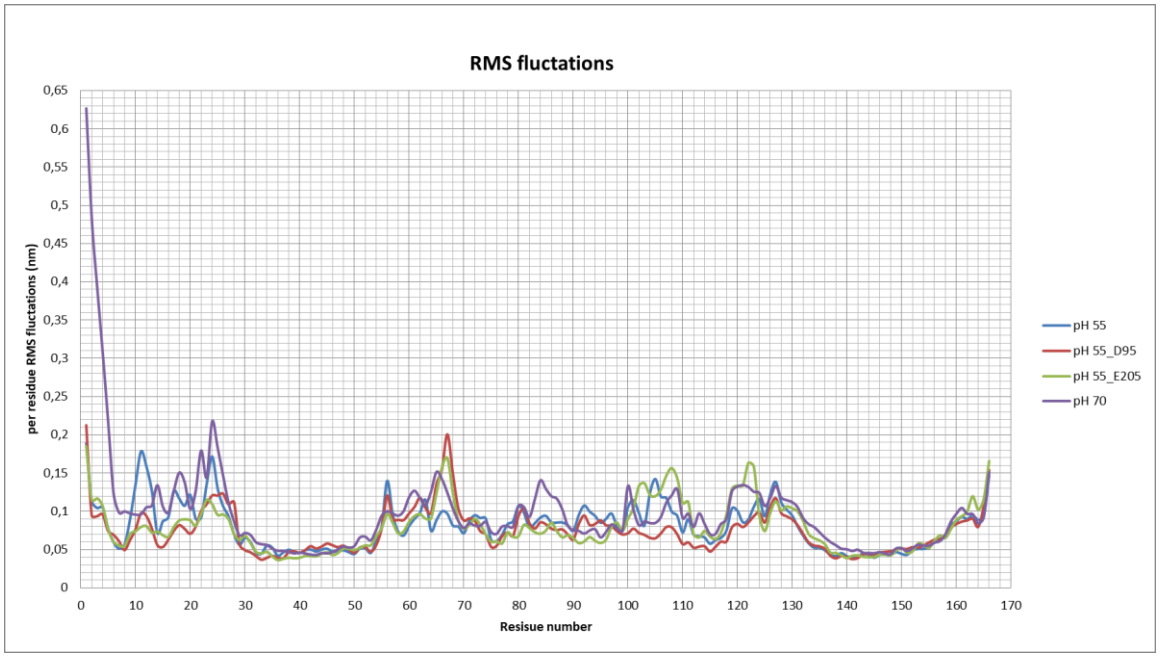
Πριν προχωρήσουμε σε οποιαδήποτε ενέργεια, θα πρέπει να τοποθετήσουμε και πάλι τις πρωτεΐνες στο κέντρο της μεμβράνης. Λόγω των περιοδικών συνθηκών που εφαρμόσαμε στο σύστημα, ενδέχεται οι πρωτεΐνες να έχουν απομακρυνθεί ή να φαίνονται σπασμένες (εικόνα 26). Με το στάδιο αυτό ολοκληρώνεται η διαδικασία προσομοίωσης και μπορούμε να προχωρήσουμε στην ανάλυση δεδομένων.

# 4.ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ

## 4.1 Ανάλυση αποτελεσμάτων της δυναμικής του συμπλόκου πρωτεϊνών FCP και LHCX1

Στην ανάλυση των αποτελεσμάτων που προκύπτουν από τη διαδικασία της Μοριακής Δυναμικής, θα επικεντρωθούμε κυρίως (α) στην κινητικότητα των αμινοξέων στις διαφορετικές συνθήκες pH που έχουμε μελετήσει, (β) τυχόν αλλαγές στη διαμόρφωση των πρωτεϊνών σε αυτές τις συνθήκες, όπως και τα σημαντικότερα κατάλοιπα με ισχυρή αλληλεπίδραση μεταξύ των FCP και LHCX1.

Μέσω της παραμέτρου RMSF ανά κατάλοιπο, θα συγκρίνουμε συλλογικά τις διαφορές στην κινητικότητα των αμινοξέων και κατά πόσο αυτά επηρεάζονται από τις αλλαγές στο μικροπεριβάλλον τους.

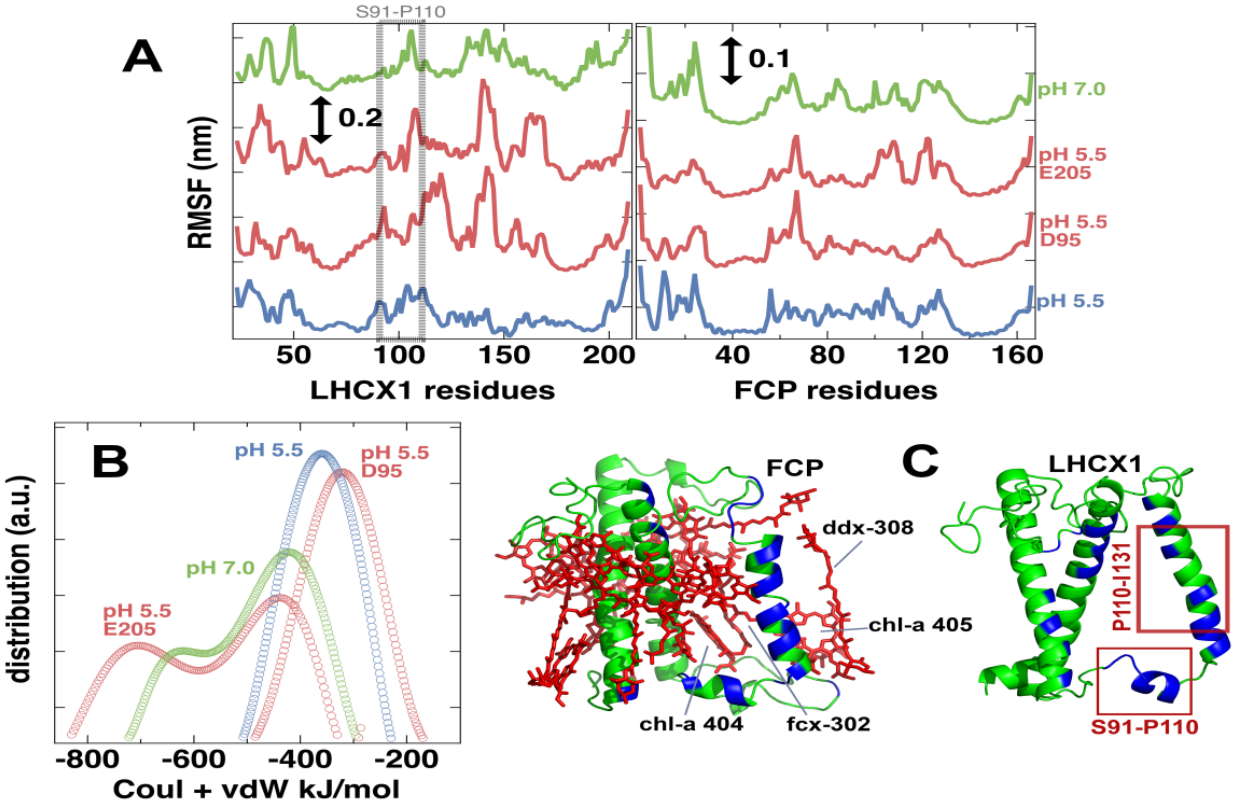



*Εικόνα 27.* Μέση τετραγωνική ρίζα διακύμανσης (root mean square fluctuation RMSF), της FCP,ανά κατάλοιπο για pH ~5.5 (μπλε χρώμα), για pH ~5.5 με πρωτονιωμένο το D-95 και αποπρωτονιωμένο το E-205 (κόκκινο χρώμα), για pH ~5.5 με αποπρωτονιωμένο το D-95 και προτονισμένο το E-205 (πράσινο χρώμα), για pH 7 (μωβ χρώμα).

Στην εικόνα 27 μπορούμε να δούμε τα αμινοξέα τα οποία επηρεάζονται από τις αλλαγές και αυτά που είναι ιδιαίτερα σταθερά. Ξεκινώντας από αριστερά, βλέπουμε αμινοξέα τα οποία και στις τέσσερις περιπτώσεις παρουσιάζουν κάποια κινητικότητα το οποίο είναι αναμενόμενο καθώς τα αμινοτελικά και καρβοξυτελικά άκρα παρουσιάζουν πάντα κάποια κινητικότητα. Καθώς προχωράμε προς τα δεξιά βλέπουμε στην περιοχή των αμινοξέων 35-50 πως είναι πολύ σταθερά και δεν παρουσιάζουν αλλαγές, όπως συμβαίνει και στα αμινοξέα 130-160, τα οποία απαρτίζουν τις δύο από τις τρεις α-έλικες

της πρωτεΐνης. Αυτό που αξίζει να σημειωθεί είναι πως εκτός από τα αμινοξέα 57-75 τα οποία ανήκουν σε στροφή και είναι αναμενόμενο να έχουν διακυμάνσεις στην κινητικότητα τους, βλέπουμε τα αμινοξέα 76-99 πως παρουσιάζουν έντονες αλλαγές αν και βρίσκονται σε α-έλικα και οι αλλαγές γίνονται εντονότερες στα 100-124 τα οποία βρίσκονται σε στροφές.

Οπότε ένα πρώτο συμπέρασμα είναι πως υπάρχουν περιοχές της πρωτεΐνης οι οποίες όντως παρουσιάζουν αλλαγές καθώς αλλάζει η τιμή του pH, και μάλιστα σχετίζονται με αλλαγές στη διαμόρφωση της πρωτεΐνης.

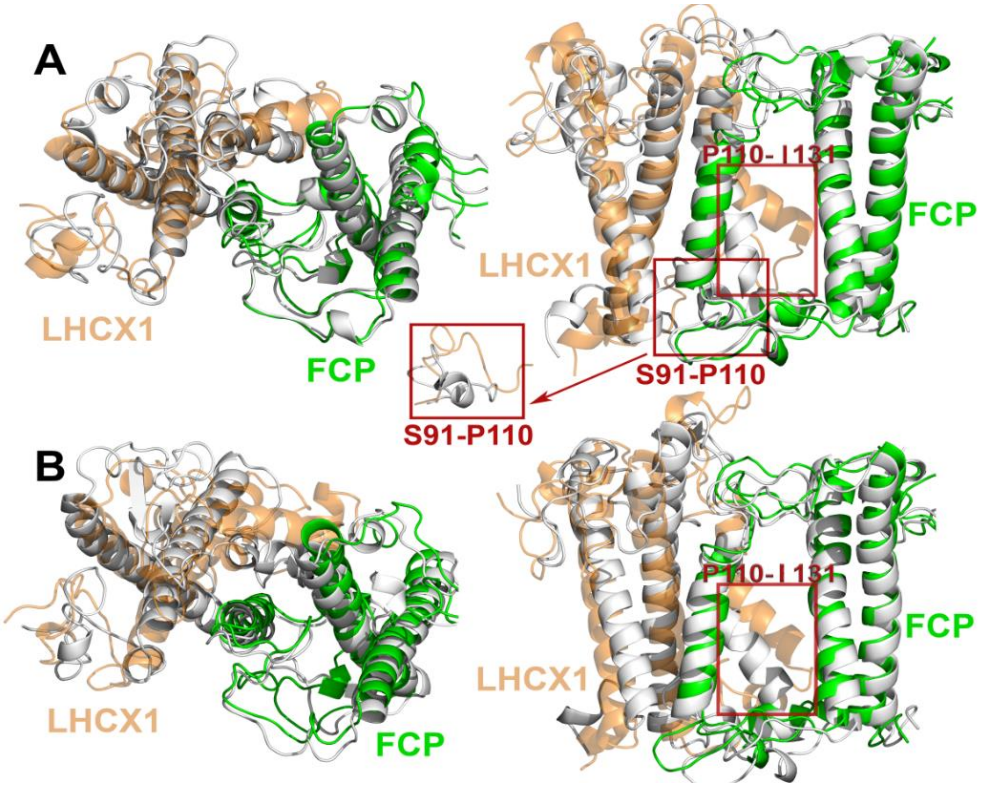



*Εικόνα 28.* Α) Μέση τετραγωνική ρίζα διακύμανσης (root mean square fluctuation RMSF) ανά κατάλοιπο για τις πρωτεΐνες LHCX1 και FCP στις διάφορες περιπτώσεις πρωτονίωσης. Β) Αλληλεπιδράσεις Coulomb και δυνάμεις vdW μεταξύ των πρωτεϊνών LHCX1 και FCP για τις τέσσερις περιπτώσεις πρωτονίωσης. Γ) Δομή των πρωτεϊνών LHCX1 και FCP.

Συνεχίζοντας την σύγκριση ανάμεσα στις τέσσερις περιπτώσεις πρωτονίωσης, μέσω του RMSF, μπορούμε να εντοπίσουμε την περιοχή καταλοίπων S91-P110 της LHCX1. Η συγκεκριμένη περιοχή θα μπορούσε να έχει κύριο ρόλο στη φωτοπροστασία καθώς αλλάζει ιδιαίτερα η κινητικότητα της περιοχής και η διαμόρφωσή της σε διαφορετικές καταστάσεις πρωτονίωσης (εικόνα 28A).

Εκτός από την κινητικότητα των αμινοξέων, μπορούμε να δούμε πως οι ηλεκτροστατικές αλληλεπιδράσεις (Coulomb) και οι δυνάμεις Van Der Waals (vdW) αλλάζουν στις τέσσερις περιπτώσεις πρωτονίωσης. Η ισχυρότερη αλληλεπίδραση γίνεται σε pH 7.0 και pH ~5.5 με πρωτονιωμένο το E-205 και αποπρωτονιωμένο το D-95, ενώ όταν προσθέσουμε πρωτόνιο στο D-95 η αλληλεπίδραση αυτή εξασθενεί (εικόνα 28B). Αυτό έρχεται σε αντίθεση με τα αποτελέσματα που αναμέναμε, καθώς θα έπρεπε όσο μειώνεται το pH και ενεργοποιείται ο φωτοπροστατευτικός μηχανισμός, η αλληλεπίδραση μεταξύ των δυο πρωτεϊνών να γίνεται πιο ισχυρή. Εκτός από τις αλληλεπιδράσεις Coulomb και τις δυνάμεις vdW, μπορεί να υπάρχουν και άλλα φαινόμενα που συμβάλλουν στην επαγωγή της φωτοπροστατευτικής λειτουργίας, όπως μεταφορά ενέργειας μεταξύ χρωστικών, ή αλλαγή στη διαμόρφωση χρωστικών στην περιοχή της διεπιφάνειας που έχει ως άμεσο αποτέλεσμα τη μετάβαση από μια κατάσταση συλλογής φωτός σε μια κατάσταση φωτοπροστασίας. Συγκεκριμένα, στην εικόνα 28C με μπλε εμφανίζονται οι περιοχές αλληλεπίδρασης στις FCP/ LHCX1 στο σύμπλοκο που δημιουργείται. Υπάρχουν χαρακτηριστικές χρωστικές στην περιοχή (chl-a 405, ddx-308, fcx-302 και Chl-a 404) που θα μπορούσαν να αλλάξουν τον προσανατολισμό τους με την αλληλεπίδραση FCP-LHCX1.

Η αλληλεπίδραση με βάση τις δυνάμεις Coulomb και vdW φαίνεται πως επηρεάζει σε κάποιο βαθμό τη δευτεροταγή δομή της LHCX1 (εικόνες 29A, 29B). Η διαμόρφωση της LHCX1 στις περιοχές S91-P110, P110-I131 εξαρτάται από την τιμή του pH και την προσθήκη πρωτονίων στην πλευρά του lumen. Συγκεκριμένα, εάν συγκρίνουμε τις δομές του συμπλόκου για pH ~5.5, όπου η FCP απεικονίζεται με πράσινο χρώμα και η LHCX1 με πορτοκαλί, και pH 7, όπου και οι δύο πρωτεΐνες απεικονίζονται με λευκό, (εικόνα 29A), βλέπουμε πως στις περιοχές S91-P110 και P110-I131 υπάρχει αλλαγή μεταξύ α-έλικας και απουσίας δομικών στοιχείων, επομένως είναι ξεκάθαρο πως όταν το pH αλλάζει, και ενεργοποιείται ο φωτοπροστατευτικός μηχανισμός, επηρεάζεται και η δομή των πρωτεϊνών. Αντίστοιχα, στην εικόνα 29B βλέπουμε τις δομές σε pH ~5.5 με πρωτονιωμένο το D-95 και αποπρωτονιωμένο το E-205, όπου η FCP απεικονίζεται με πράσινο χρώμα και η LHCX1 με πορτοκαλί, και pH ~5.5 με αποπρωτονιωμένο το D-95 και πρωτονιωμένο το E-205, όπου και οι δύο πρωτεΐνες απεικονίζονται με λευκό.




*Εικόνα 29.* Δομή του συμπλόκου πρωτεϊνών A) σε pH ~5.5, όπου η FCP αναπαρίσταται με πράσινο και η LHCX1 με πορτοκαλί και σε pH 7, όπου FCP και LHCX1 αναπαρίστανται με λευκό B) σε pH ~5.5 με πρωτονιωμένο το D-95, όπου η FCP αναπαρίσταται με πράσινο και η LHCX1 με πορτοκαλί και σε pH ~5.5 με πρωτονιωμένο το E-205, όπου FCP και LHCX1 αναπαρίστανται με λευκό.

Παρατηρούμε, λοιπόν, πως τα μοντέλα σε pH ~5.5 και pH ~5.5 με πρωτονιωμένο D-95, έχουν παρόμοια συμπεριφορά όσον αφορά τις περιοχές της LHCX1, οι οποίες φαίνονται να έχουν κάποιο ρόλο (91-110, 110-131) στη φωτοπροστασία. Επίσης τα μοντέλα στα οποία το D-95 δεν είναι πρωτονιωμένο, pH ~5.5 και 7, παρουσιάζουν παρόμοιες διαμορφώσεις στις περιοχές 91-110 και 110-131 της LHCX1. Αυτό δηλώνει πως το κύριο κατάλοιπο που είναι κατά πάσα πιθανότητα υπεύθυνο για την ενεργοποίηση της LHCX1 στη φωτοπροστατευτική διαμόρφωση είναι το D-95. Το τελευταίο πρωτονιώνεται σε χαμηλό pH (~5.5), ενώ αποπρωτονιώνεται σε υψηλότερες τιμές pH και αποτελεί τον διακόπτη μεταξύ διαμόρφωσης συλλογής φωτός και φωτοπροστασίας.

Στην εικόνα 30 μπορούμε να δούμε στην αλληλουχία των FCP/ LHCX1, ποια κατάλοιπα επηρεάζονται από την αλλαγή του pH και πως αλλάζει η διαμόρφωση τους.

*Εικόνα 30.* Πρωτοταγής και δευτεροταγής δομή των πρωτεϊνών FCP και LHCX1 και πως αυτές διαφοροποιούνται ανάλογα με τις πρωτονιώσεις.

Πράγματι, στην LHCX1, φαίνεται πως οι περιοχές S91-P110, P110-I131 επηρεάζονται από τις αλλαγές του pH. Σε pH ~5.5, η περιοχή S91-P110 φαίνεται πως είναι σε στροφή, όταν προσθέσουμε πρωτόνιο στο D-95 έχουμε τη δημιουργία α-έλικας (κόκκινη επισήμανση), σε ~5.5 με πρωτονιωμένο E-205 και αποπρωτονιωμένο D-95 δημιουργείται π-έλικα (μπλε επισήμανση) και σε pH 7 έχουμε και πάλι δημιουργία π-έλικας (μπλε επισήμανση). Η περιοχή P110-I131, σε pH ~5.5 φαίνεται πως είναι δυο κομμάτια α-έλικας με μια στροφή ενδιάμεσα, σε pH ~5.5 με πρωτονιωμένο D-95 έχουμε μια α-έλικα μικρού μήκους, σε pH ~5.5 με πρωτονιωμένο E-205 έχουμε μια ακόμα μικρότερη α-έλικα αλλά δημιουργείται μια π-έλικα και σε pH 7 δημιουργείται μια α-έλικα μεγάλου μήκους.

Στην FCP, οι αλλαγές δεν είναι τόσο έντονες, ωστόσο μπορούμε να παρατηρήσουμε πως υπάρχουν διαφοροποιήσεις στην περιοχή L20-D23, όπου σε pH ~5.5 έχουμε στροφή ενώ στις υπόλοιπες περιπτώσεις δημιουργείται α-έλικα, στην F70-E72, όπου έχουμε π-έλικα εκτός από το pH ~5.5 με πρωτονιωμένο D-95 και στην W125-T127, όπου σε pH ~5.5 και pH ~5.5 με πρωτονιωμένο E-205 έχουμε π-έλικα, ενώ σε pH ~5.5 με πρωτονιωμένο D-95 και pH7 έχουμε στροφή.

Αφού έχουμε μελετήσει την επίδραση της αλλαγής του pH στην δευτεροταγή δομή των υπό μελέτη πρωτεϊνών στο σύμπλοκο, θα πρέπει να εξετάσουμε και πως αυτές οι αλλαγές στη δευτεροταγή δομή επηρεάζουν τις αλληλεπιδράσεις μεταξύ FCP-LHCX1 στο σύμπλοκο. Για το σκοπό αυτό, θα χρησιμοποιήσουμε διαγράμματα Ramachandran (Ramachandran, 1963), τα οποία είναι μια οπτικοποίηση των διερδικών γωνιών των καταλοίπων που βρίσκονται στον σκελετό (*backbone*) μιας πρωτεΐνης. Με βάση αυτά τα διαγράμματα, μπορούμε να εντοπίσουμε τις δίεδρες γωνίες που σχηματίζει κάθε κατάλοιπο με τα δυο γειτονικά του. Το εύρος των γωνιών αυτών αφορά σε δύο κατηγορίες: γωνίες που ανήκουν στις ‘επιτρεπτές περιοχές’, και γωνίες με τιμές στις μη επιτρεπτές περιοχές του διαγράμματος Ramachandran, λόγω της χωρικής διάταξης των καταλοίπων. Σύμφωνα με τα παραπάνω, με τη βοήθεια ενός διαγράμματος Ramachandran, μπορούμε να αποκτήσουμε εικόνα για τη δομή των πεπτιδίων,

μελετώντας μόνο τη διαμόρφωση των γωνιών μεταξύ των καταλοίπων τους, αλλά να εντοπίσουμε και τμήματα της πρωτεΐνης όπου οι δίεδρες γωνίες πλησιάζουν ή βρίσκονται σε μη επιτρεπτές περιοχές. Τέτοιες περιοχές φαίνεται να είναι ιδιαίτερα σημαντικές για τη λειτουργία των πρωτεϊνών, καθώς έχουν 'αποθηκευμένο' μεγάλο περιεχόμενο σε δυναμική ενέργεια που μπορεί να διοχετευτεί σε απαραίτητες αλλαγές στη διαμόρφωσή κατά τη λειτουργία τους (Penner, 2020).

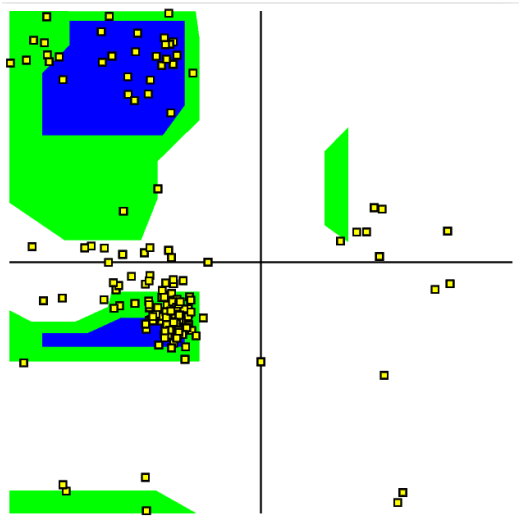
Στις εικόνες 31-34 βλέπουμε τα διαγράμματα αυτά ,για την πρωτεΐνη FCP, στις διαφορετικές περιπτώσεις πρωτονίωσης του συμπλόκου FCP-LHCX1. Σύμφωνα με τις γωνίες που σχηματίζουν τα κατάλοιπα μεταξύ τους, οι πλέον επιτρεπόμενες περιοχές είναι χρωματισμένες με μπλε, ενώ αυτές που είναι μερικώς επιτρεπόμενες είναι χρωματισμένες με πράσινο. Σε ένα διάγραμμα Ramachandran, στις χρωματισμένες περιοχές βρίσκονται, θεωρητικά, όλες οι πιθανές δευτεροταγείς δομές. Επομένως, είναι αναμενόμενο να βλέπουμε πως σε αυτές τις περιοχές συγκεντρώνονται τα περισσότερα κατάλοιπα, εφόσον πρόκειται για α-έλικες ή για β-πτυχωτές. Αντίθετα, στις λευκές περιοχές είναι δύσκολο να υπάρξουν οι αντίστοιχες διατάξεις των γωνιών, λόγω της στερικής παρεμπόδισης μεταξύ των ατόμων . Ωστόσο, αν και μια τέτοια διάταξη είναι πολύ δύσκολο να συμβεί, μπορούμε να υποθέσουμε πως αφού η FCP ειναι ικανή να τις υιοθετήσει στη δομή της, πρόκειται για μια αρκετά ευέλικτη πρωτεΐνη ή οι περιοχές αυτές της FCP σχετίζονται άμεσα με αλλαγές στη διαμόρφωσή της για τη λειτουργία της απορρόφησης φωτός – φωτοπροστασίας. Τα κατάλοιπα που βρίσκονται στις λευκές περιοχές χρειάζονται μεγάλη ενέργεια για να μπορέσουν να σχηματίσουν τη διαμόρφωση αυτή και γι' αυτό τον λόγο ενδέχεται να την εκλύουν σε συνθήκες φωτοπροστασίας με έντονες αλλαγές στη διαμόρφωση.

Εκτός από τα κατάλοιπα της γλυκίνης (Gly), των οποίων το διάγραμμα διαφοροποιείται από τον γενικό κανόνα των διαγραμμάτων Ramachandran καθώς είναι δυνατόν να καταγράψουν γωνίες σε οποιαδήποτε περιοχή (Ho, Brasseur, 2005), τα κατάλοιπα που βρίσκονται εκτός των επιτρεπόμενων περιοχών για την FCP ειναι: α) για pH ~5.5, **LEU-18, VAL-21, ASP-23, ASN-118**, β) για pH~5.5 (D95 πρωτονιωμένο), **THR-107, PHE-116**, γ) για pH~5.5 (E-205 πρωτονιωμένο), **ARG-104, ASN-118**, δ) για pH 7, **GLU-99, ARG-104, LEU-106**. Σε αυτό το σημείο, θα πρέπει να σημειώσουμε πως τα **Asn-118**, **Thr-107, Leu-106** αλληλεπιδρούν με τη φουκοξανθίνη-304, η **Leu-18** με την φουκοξανθίνη-306, η **Phe-116** με τη φουκοξανθίνη 302, ενώ το **Glu-99** είναι αξονικός υποκαταστάτης της Chl-a 406. Η αποθηκευμένη ενέργεια στις δίεδρες αυτές γωνίες, που ανήκουν σε μη επιτρεπτές περιοχές του διαγράμματος Ramachandran, μπορεί να οδηγήσει σε αλλαγές στους προσανατολισμούς των παραπάνω χρωστικών, με άμεση συσχέτιση με το μηχανισμό φωτοπροστασίας στην FCP.

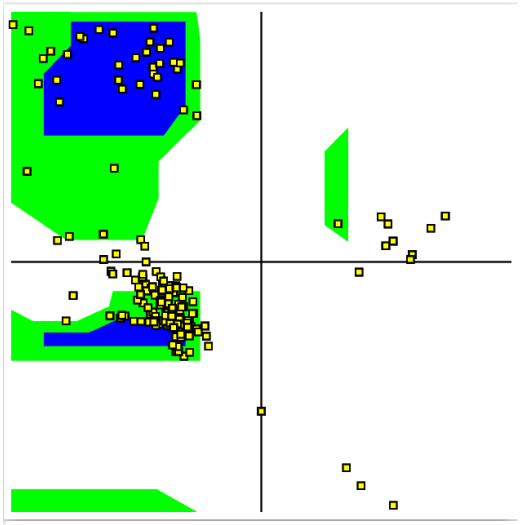
Επίσης, στον πίνακα της εικόνας 35 παρουσιάζονται τα αποτελέσματα από χάρτες επαφής FCP-LHCX1 των τεσσάρων περιπτώσεων πρωτονίωσης, καθώς και τα κατάλοιπα που αλληλεπιδρούν μεταξύ τους, με ένα όριο τα 4 Å .

Από τις παραπάνω αναλύσεις, παρατηρούμε πως όταν το D-95 δεν είναι πρωτονιωμένο, τα κατάλοιπα απομακρύνονται έντονα από τις δευτεροταγείς δομές (εικόνα 33, 34), ενώ όταν είναι πρωτονιωμένο, αν και πάλι, υπάρχουν αμινοξέα τα οποία βρίσκονται σε λευκές περιοχές, είναι πιο κοντά στην τυπική διάταξη της πρωτεΐνης (εικόνα 31, 32). Επίσης, όταν το D-95 δεν είναι πρωτονιωμένο, υπάρχουν περισσότερα κατάλοιπα μεταξύ των πρωτεϊνών τα οποία αλληλεπιδρούν. Οι περιοχές αλληλεπίδρασης αυτών των καταλοίπων παρουσιάζουν μια σημαντική διασπορά. Σε κάθε άλλη περίπτωση, που το D-95 πρωτονιώνεται, υπάρχουν τρεις κύριες περιοχές που φαίνονται σχετικά εντοπισμένες, όπου τα κατάλοιπα των δύο πρωτεϊνών βρίσκονται αρκετά κοντά μεταξύ τους.

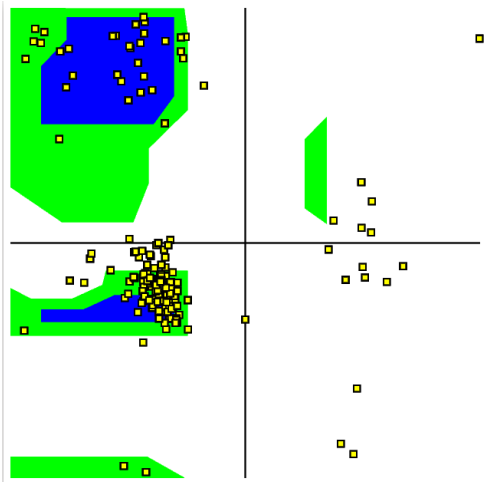
Οι αλληλεπιδράσεις των καταλοίπων των πρωτεϊνών FCP και LHCX1 είναι τριών τύπων, α) αλληλεπιδράσεις vdW, στις οποίες δίνεται βαρύτητα στις υδρόφοβες αλληλεπιδράσεις μεταξύ μορίων, με αποστάσεις που δεν πρέπει να υπερβαίνουν τα 0.5 Å, β) δεσμοί υδρογόνου, οι οποίοι εξαρτώνται από την απόσταση των δύο μορίων η οποία δεν ξεπερνά τα 3.5 Å άλλα και από τη γωνία μεταξύ του δότη/ δέκτη που δεν πρέπει να ξεπερνά τις 63º, γ) αλληλεπιδράσεις π-π *stacking* οι οποίες αφορούν αρωματικά κατάλοιπα και η απόσταση των αρωματικών τους δακτυλίων δεν μπορεί να ξεπερνά τα 6.5 Å, και να εμφανίζονται παράλληλα για μέγιστη αλληλοεπικάλυψη ηλεκτρονικών νεφών.



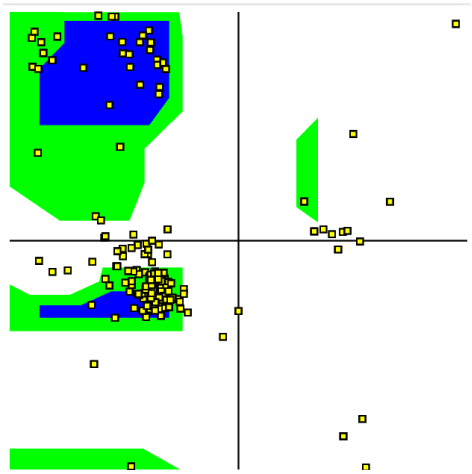
*Εικόνα 31.* Διάγραμμα Ramachandran της πρωτεΐνης FCP για τιμή pH ~5.5.



*Εικόνα 32.* Διάγραμμα Ramachandran της πρωτεΐνης FCP για τιμή pH ~5.5.με πρωτονιωμένο το D-95.



*Εικόνα 33.* Διάγραμμα Ramachandran της πρωτεΐνης FCP για τιμή pH ~5.5 με πρωτονιωμένο το E-205.



*Εικόνα 34.* Διάγραμμα Ramachandran της πρωτεΐνης FCP για τιμή pH ~7.

| pH ~5.5 | | | pH ~5.5_D95 | | | pH ~5.5_E205 | | | pH 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A:65: :TYR | VDW:SC_SC | S:197: :LEU | A:77: :PHE | VDW:SC_SC | S:101: :PRO | A:77: :PHE | PIPISTACK:SC_SC | S:113: :PHE | A:65: :TYR | HBOND:SC_MC | S:197: :LEU |
| A:80: :PHE | PIPISTACK:SC_SC | S:105: :HIS | A:106: :LEU | VDW:MC_SC | S:133: :TRP | A:77: :PHE | VDW:SC_SC | S:113: :PHE | A:65: :TYR | VDW:SC_SC | S:197: :LEU |
| A:80: :PHE | VDW:SC_SC | S:105: :HIS | A:107: :THR | VDW:MC_SC | S:133: :TRP | A:80: :PHE | VDW:SC_SC | S:104: :THR | A:80: :PHE | VDW:SC_SC | S:112: :PRO |
| A:93: :PHE | VDW:SC_SC | S:186: :MET | A:106: :LEU | VDW:SC_SC | S:62: :ARG | A:80: :PHE | VDW:SC_SC | S:112: :PRO | A:84: :PRO | VDW:SC_SC | S:197: :LEU |
| A:105: :ASP | HBOND:SC_MC | S:62: :ARG | A:100: :SER | VDW:SC_SC | S:69: :THR | A:80: :PHE | VDW:SC_SC | S:116: :LEU | A:88: :LEU | VDW:SC_SC | S:114: :TRP |
| A:106: :LEU | HBOND:SC_MC | S:62: :ARG | A:101: :SER | HBOND:SC_SC | S:70: :HIS | A:89: :VAL | VDW:SC_SC | S:193: :MET | A:89: :VAL | VDW:SC_SC | S:190: :ALA |
| A:106: :LEU | VDW:SC_SC | S:62: :ARG | A:92: :LEU | VDW:SC_SC | S:98: :ILE | A:90: :GLN | VDW:SC_SC | S:197: :LEU | A:92: :LEU | VDW:SC_SC | S:114: :TRP |
| A:100: :SER | VDW:SC_SC | S:66: :ALA | A:89: :VAL | VDW:SC_SC | S:99: :SER | A:92: :LEU | VDW:SC_SC | S:119: :ILE | A:92: :LEU | VDW:SC_SC | S:115: :VAL |
| A:93: :PHE | PIPISTACK:SC_SC | S:70: :HIS | A:80: :PHE | VDW:SC_MC | S:101: :PRO | A:93: :PHE | VDW:SC_SC | S:193: :MET | A:100: :SER | VDW:SC_SC | S:126: :GLN |
| A:93: :PHE | VDW:SC_SC | S:70: :HIS | A:81: :LYS | VDW:SC_SC | S:101: :PRO | A:104: :ARG | VDW:SC_SC | S:126: :GLN | A:104: :ARG | VDW:SC_SC | S:130: :VAL |
| A:96: :GLY | VDW:SC_MC | S:70: :HIS | | | | A:106: :LEU | VDW:SC_SC | S:49: :LEU | A:104: :ARG | HBOND:SC_MC | S:133: :TRP |
| A:97: :VAL | VDW:SC_SC | S:70: :HIS | | | | A:106: :LEU | VDW:SC_SC | S:51: :PHE | A:104: :ARG | PICATION:SC_SC | S:133: :TRP |
| A:92: :LEU | VDW:MC_SC | S:100: :GLY | | | | A:107: :THR | VDW:SC_SC | S:51: :PHE | A:104: :ARG | VDW:SC_SC | S:133: :TRP |
| A:85: :ALA | VDW:SC_SC | S:101: :PRO | | | | A:107: :THR | VDW:SC_MC | S:62: :ARG | A:101: :SER | HBOND:SC_MC | S:62: :ARG |
| A:92: :LEU | VDW:SC_SC | S:103: :ILE | | | | A:100: :SER | VDW:SC_SC | S:70: :HIS | A:106: :LEU | VDW:SC_SC | S:62: :ARG |
| | | | | | | A:85: :ALA | VDW:MC_SC | S:101: :PRO | A:97: :VAL | VDW:SC_SC | S:63: :TYR |
| | | | | | | A:89: :VAL | VDW:SC_SC | S:101: :PRO | A:93: :PHE | PIPISTACK:SC_SC | S:70: :HIS |
| | | | | | | A:85: :ALA | VDW:SC_SC | S:102: :ALA | A:93: :PHE | VDW:SC_SC | S:70: :HIS |
| | | | | | | A:88: :LEU | VDW:SC_SC | S:104: :THR | A:96: :GLY | HBOND:SC_MC | S:70: :HIS |
| | | | | | | A:159: :GLN | VDW:SC_SC | S:113: :PHE | A:96: :GLY | VDW:SC_MC | S:70: :HIS |
| | | | | | | | | | A:85: :ALA | VDW:SC_SC | S:101: :PRO |

*Εικόνα 35.* Τύποι κυριότερων αλληλεπιδράσεων μεταξύ των καταλοίπων των πρωτεϊνών FCP(chain A) και LHCX1 (chain S) για τις τέσσερις περιπτώσεις πρωτονίωσης.

Στην εικόνα 35, βλέπουμε τα κατάλοιπα με την ισχυρότερη αλληλεπίδραση, καθώς και το είδος αυτής. Σε τιμή pH 7 και τιμή pH ~5.5 με πρωτονιωμένο το E-205, έχουμε περισσότερες αλληλεπιδράσεις, κυρίως vdW, σε σύγκριση με pH ~5.5 και pH ~5.5 με πρωτονιωμένο το D-95. Συμπεραίνουμε πως οι αλληλεπιδράσεις vdW και Coulomb εξασθενούν όταν πλησιάζουμε στην επαγωγή του φωτοπροστατευτικού μηχανισμού για την αλληλεπίδραση LHCX1 - FCP. Όπως στην εικόνα 28B παρατηρούμε πως οι καμπύλες σε pH ~5.5 και pH ~5.5 με D-95 πρωτονιωμένο είναι παρόμοιες και οι καμπύλες σε pH ~5.5 με E-205 πρωτονιωμένο και pH 7.0 είναι και αυτές παρόμοιες μεταξύ τους. Στην εικόνα 35 βλέπουμε πως τα κατάλοιπα των δύο πρωτεϊνών που αλληλεπιδρούν, ομαδοποιούνται ανά ζεύγη pH ~5.5, pH ~5.5 με D-95 πρωτονιωμένο και pH ~5.5 με E-205 πρωτονιωμένο και pH 7.

## 4.2 Συζήτηση

Σύμφωνα με όλα τα παραπάνω, μπορούμε να βγάλουμε τα εξής συμπεράσματα.
(Α) Η φωτοπροστατευτική πρωτεΐνη LHCX1 είναι αρκετά ευέλικτη και η αλλαγή του pH φαίνεται πως επηρεάζει σημαντικά τη δευτεροταγή της δομή στις περιοχές S91-P110 και P110-I131. Τα κατάλοιπα της LHCX1 που επηρεάζονται από την αλλαγή αυτή, πρέπει να έχουν κάποια σημασία στον φωτοπροστατευτικό μηχανισμό, ιδιαίτερα το D-95.
(Β) Όταν εξετάσουμε το πρωτεϊνικό σύμπλοκο FCP - LHCX1 παρατηρούμε πως το αμινοξυ-τελικό άκρο της FCP διαφέρει έντονα σε διαμόρφωση μεταξύ pH 7.0 και pH ~5.5, καθώς φαίνεται πως η κινητικότητα του μειώνεται. Επίσης, καθώς αλλάζει η δευτεροταγής δομή της LHCX1 ή οι πρωτονιώσεις στη lumen πλευρά της LHCX1, αλλάζει σημαντικά η αλληλεπίδραση μεταξύ LHCX1 - FCP.

Οι περιπτώσεις πρωτονιώσεις της LHCX1 που ελέγχθηκαν ήταν:

1. pH~5.5, τα Asp (D) και Glu (E) είναι αποπρωτονιωμένα, εκτός των D-95, E-205, E-65 και E-85 τα οποία είναι πρωτονιωμένα
2. pH~5.5_D95, τα Asp (D) και Glu (E) είναι αποπρωτονιωμένα, εκτός των D-95, E-65 και E-85 τα οποία είναι πρωτονιωμένα
3. pH~5.5_E-205, τα Asp (D) και Glu (E) είναι αποπρωτονιωμένα, εκτός των E-205, E-65 και E-85 τα οποία είναι πρωτονιωμένα
4. pH 7, τα Asp (D) και Glu (E) είναι αποπρωτονιωμένα, εκτός των E-65 και E-85 τα οποία είναι πρωτονιωμένα

(Γ) Η διαμόρφωση της FCP, με τη συμπλοκοποίηση FCP-LHCX1, φαίνεται να επηρεάζεται με παρόμοιο τρόπο στις παραπάνω περιπτώσεις 1-2 και 3-4. Η πρωτονίωση του D-95 παίζει το ρόλο του ¨διακόπτη¨ που αντιλαμβάνεται - ανιχνεύει τις τιμές του pH στο lumen, το βαθμό οξύνισής του, και προτείνουμε με βάση τα μοντέλα που παρουσιάστηκαν σε αυτή τη διατριβή, πως με αυτόν τον τρόπο ενεργοποιείται η φωτοπροστασία στο σύμπλοκο FCP-LHCX1. Όπως είδαμε προηγουμένως, εκτός από τις στροφές οι οποίες είναι περιοχές με έντονη κινητικότητα και είναι αναμενόμενο να μεταβάλλονται καθώς αλλάζει το pH, αλλάζει και η διαμόρφωση μιας α-έλικας, αν και θεωρητικά είναι πιο συντηρημένη περιοχή. Επομένως, μπορούμε με βεβαιότητα να πούμε πως οι αλλαγές στο pH, επάγουν και αλλαγές στη διαμόρφωση των πρωτεϊνών και αυτές με τη σειρά τους μεταβάλλουν τη λειτουργία τους.
(Δ) Ένα ακόμη αξιοσημείωτο αποτέλεσμα είναι πως όταν το pH μειώνεται, οι αλληλεπιδράσεις vdW και Coulomb εξασθενούν αν και αναμέναμε το αντίθετο, μιας και σε pH ~5.5 θα έπρεπε να ενισχύεται η αλληλεπίδραση μεταξύ FCP και LHCX1. Ενδεχομένως, καθώς το pH μειώνεται, μπορεί να υπάρχουν επιμέρους αλληλεπιδράσεις οι οποίες να γίνονται ισχυρότερες ή η επίδρασή τους στην FCP να είναι σημαντική, και με αυτόν τον τρόπο να επάγεται ο φωτοπροστατευτικός μηχανισμός. Μια άλλη εξήγηση είναι πως το σύμπλοκο, αντιλαμβάνεται τη μείωση του pH ως μια στρεσογόνα συνθήκη κατά την οποία μπαίνει σε μια διαδικασία συντήρησης, δαπανώντας την ελάχιστη ενέργεια ώστε να μην αποδιαταχθεί, ενώ διοχετεύει όλη την υπόλοιπη ενέργεια, ώστε να ενεργοποιηθεί η διαδικασία της φωτοπροστασίας και να συνεχίσει η λειτουργία της φωτοσύνθεσης στον οργανισμό.

### 4.3 Επίλογος-Μελλοντική ερευνητική εργασία

Με την παρούσα εργασία, στόχος μας ήταν να προσδιορίσουμε μια τρισδιάσταση δομή του συμπλόκου FCP-LHCX1, και να διαπιστώσουμε τις αλλαγές στις διαμορφώσεις των πρωτεϊνών που συμμετέχουν στο σύμπλοκο, σε συνάρτηση του pH στη lumen περιοχή. Ο στόχος αυτός επετεύχθη και μάλιστα ταυτοποιήσαμε ένα σημαντικό κατάλοιπο στην LHCX1 (D-95) που φαίνεται να παίζει το ρόλο του διακόπτη για την ενεργοποίηση της πρωτεΐνης, τις αλληλεπιδράσεις με την FCP και τη μετάβαση του συμπλόκου FCP-LHCX1 μεταξύ της κατάστασης συλλογής φωτός και φωτοπροστασίας στα διάτομα. Η παρούσα εργασία πραγματοποιείται σε συνεργασία με το εργαστήριο του καθ. Alexander Ruban, στο Queen Mary University of London. Παράλληλα, στο εργαστήριο αυτό, πραγματοποιούνται πειράματα σε μεταλλάγματα της LHCX1 στα κατάλοιπα D-95 και E-205 και αναμένουμε να δώσουν σημαντικές πληροφορίες για την αξιολόγηση των μοντέλων μας, αλλά και σε συνδυασμό με τη θεωρητική δουλειά να διασαφηνίσουν το μηχανισμό φωτοπρστασίας στα διάτομα.

Ένα σημαντικό ερώτημα που παραμένει αναπάντητο είναι η αλληλεπίδραση μεταξύ χρωστικών στο σύμπλοκο LHCX1-FCP σε διαφορετικές τιμές pH, σε επίπεδο μεταφοράς ενέργειας εξιτονίου (exciton energy transfer, EET) ή μεταφοράς φορτίου (charge transfer, CT) που έχει άμεση σχέση με την απόσβεση περίσσειας ενέργειας στο μηχανισμό φωτοπροστασίας. Για την απάντηση αυτού του ερωτήματος, θα πρέπει να ανατρέξουμε σε πιο ακριβείς και πολύπλοκες, αλλά κοστοβόρες χρονικά, μεθόδους όπως είναι οι *ab initio* (Daskalakis, 2020), σε μελλοντικό στάδιο της έρευνας για τα διάτομα.

# ΒΙΒΛΙΟΓΡΑΦΙΑ


Abdullahi A.S., Underwood G.J.C., Gretz M.R., (2006). Extracellularmatrix assembly in diatoms (Bacillariophyceae). V. Environmental effects on polysaccharide synthesis in the model diatom, Phaeodactylum tricornutum, J. Phycol, vol. 42, pp. 363–378. doi.org/10.1111/j.1529-8817.2006.00193.x

Alder B.J., Wainwright T.E., (1957). Phase Transition for a Hard Sphere System. The Journal of Chemical Physics, 27, 1208-1209. https://doi.org/10.1063/1.1743957

Angerhofer C.K., Maes D., Giacomoni P.U., (2009). The Use of Natural Compounds and Botanicals in the Development of Anti-Aging Skin Care Products, Book: Skin Aging Handbook, Elsevier. DOI: 10.1016/B978-0-8155-1584-5.50014-4

Armstrong E., Rogerson,A. Leftley J.W., (2000). Utilisation of seaweed carbon by three surface-associated heterotrophic protists, Stereomyxa ramosa, Nitzschia alba and Labyrinthula sp, Aquatic Microbial Ecology, vol. 21, pp. 49–57.

Bailleul B., Berne N., Murik O., Petroutsos D, Prihoda J, Tanaka A., Villanova V., Bligny R., Flori S., Falconet D., Krieger-Liszkay A., Santabarbara S., Rappaport F., Joliot P., Tirichine L., Falkowski P.G., Cardol P., Bowler C., Finazzi G., (2015). Energetic coupling between plastids and mitochondria drives CO2 assimilation in diatoms. Nature, 524(7565):366-9. doi: 10.1038/nature14599.

Ballottari M., Truong T.B., De Re E., Erickson E., Stella G.R., Fleming G.R., Bassi R., Niyogi K.K., (2016). Identification of pH-sensing Sites in the Light Harvesting Complex Stress-related 3 Protein Essential for Triggering Non-photochemical Quenching in Chlamydomonas reinhardtii. J. Biol. Chem., 1;291(14):7334-46. doi: 10.1074/jbc.M115.704601

Barker H.A., (1935). Photosynthesis in diatoms. Archiv. Mikrobiol. 6, 141–156. doi.org/10.1007/BF00407284

Bowler C., Allen A., Badger J. et al., (2008). The Phaeodactylum genome reveals the evolutionary history of diatom genomes. Nature 456, 239–244. doi.org/10.1038/nature07410

Büchel C., (2019). How diatoms harvest light. Science, 365(6452):447-448. doi:10.1126/science.aay3036

Buck, J.M., Sherman, J., Bártulos, C.R. et al., (2019). Lhcx proteins provide photoprotection via thermal dissipation of absorbed light in the diatom Phaeodactylum tricornutum. Nat Commun 10, 4167. https://doi.org/10.1038/s41467-019-12043-6

Daskalakis V., Papadatos S., Stergiannakos T., (2020). The conformational phase space of the photoprotective switch in the major light harvesting complex II. Chem Commun (Camb)., 10.1039/d0cc04486e. doi:10.1039/d0cc04486e

Dolinsky T.J, Nielsen J.E., McCammon A.J., Baker N.A., (2004). PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations, Nucleic Acids Research, Volume 32, Issue suppl_2, Pages W665–W667. https://doi.org/10.1093/nar/gkh381

Fogg G.E., (1998). The biology of polar habitats- Biology of habitats. Oxford University Press, Oxford University Press

Freddolino P.L., Harrison C.B., Liu Y., Schulten K., (2010). Challenges in protein folding simulations: Timescale, representation, and analysis. Nat Phys., 6(10):751-758. doi:10.1038/nphys1713

Geiderlo R.J., Platt T., Raven J.A., (1986). Size dependence of growth and photosynthesis in diatoms: a synthesis, MARINE ECOLOGY - PROGRESS SERIES, Vol. 30: 93-104

Giovagnetti V., Ruban A.V., (2018). The evolution of the photoprotective antenna proteins in oxygenic photosynthetic eukaryotes. Biochem Soc Trans., 46(5):1263-1277. doi:10.1042/BST20170304

Goldman J.C., Riley C.B., Dennett M.R., (1982). The effect of pH in intensive microalgal cultures. II. Species competition, Journal of Experimental Marine Biology and Ecology, Volume 57, Issue 1, pp. 15-24. doi.org/10.1016/0022 -0981(82)90141-1

Grouneva I., Gollan P.J., Kangasjärvi S. *et al.*, (2013). Phylogenetic viewpoints on regulation of light harvesting and electron transport in eukaryotic photosynthetic organisms. *Planta* 237, 399–412. doi.org/10.1007/s00425-012-1744-5

Haile J.M., (1997). Molecular Dynamics Simulation: Elementary Methods. Wiley-Interscience

Halse G.R., Syvertsen E.E., (1996). Marine Diatoms, Identifying Marine Diatoms and Dinoflagellates , Chapter 2, Elsevier, doi.org/10.1016/B978-012693015-3/50005-X

Hempel F., Lau J., Klingl A., Maier U.G., (2011). Algae as Protein Factories: Expression of a Human Antibody and the Respective Antigen in the Diatom Phaeodactylum tricornutum. PLoS ONE, 6(12): e28424. doi.org/10.1371/journal.pone.0028424

Ho B.K., Brasseur R., (2005). The Ramachandran plots of glycine and pre-proline. BMC Struct Biol **5,** 14.doi.org/10.1186/1472-6807-5-14

Holm L., Sander C., (1996). Mapping the protein universe. Science, 273 (5275): 595–603. doi:10.1126/science.273.5275.595

Horner R.A., (2002). A taxonomic guide to some common marine phytoplankton. Biopress, pp. 25–30. ISBN 9780948737-657

Kelley L., Mezulis S., Yates C. et al., (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10, 845–858. https://doi.org/10.1038/nprot.2015.053

Kuczynska P., Jemiola-Rzeminska M., Strzalka K., (2015). Photosynthetic Pigments in Diatoms, Mar Drugs, 13(9):5847-5881. doi: 10.3390/md13095847

Kühlbrandt W., Wang D., Fujiyoshi Y., (1994). Atomic model of plant light-harvesting complex by electron crystallography. Nature 367, 614–621. https://doi.org/10.1038/367614a0

Lavaud J., (2007). Fast Regulation of Photosynthesis in Diatoms: Mechanisms, Evolution and Ecophysiology, Functional Plant Science and Biotechonology, 1, pp.267-287. ffhal-01094678

Liguori N., Periole X., Marrink S. et al., (2015). From light-harvesting to photoprotection: structural basis of the dynamic switch of the major antenna complex of plants (LHCII). Sci Rep 5, 15661. doi.org/10.1038/srep15661

Litchman E., Steiner D., Bossard P., (2003). Photosynthetic and growth responses of three freshwater algae to phosphorus limitation and daylength. Freshwater Biology, 48: 2141-2148. doi:10.1046/j.1365-2427.2003.01157.x

Martin-Jézéquel, V., Tesson, B., (2012). Phaeodactylum tricornutum polymorphism: An overview.Adv. Algal CellBiol. Walter Gruyter, pp. 43–80

McCammon J., Gelin B., Karplus M., (1977). Dynamics of folded proteins, Nature 267, 585–590. doi://doi.org/10.1038/267585a0

Nagao R., Yokono M., Ueno Y., Shen J.R., Akimoto S., (2020). Acidic pH-Induced Modification of Energy Transfer in Diatom Fucoxanthin Chlorophyll a/c-Binding Proteins, The Journal of Physical Chemistry B, 124 (24), 4919-4923. DOI: 10.1021/acs.jpcb.0c04231

Papagiannakis E., H.M. van Stokkum I., Fey H., Büchel C., van Grondelle R., (2005). Spectroscopic characterization of the excitation energy transfer in the fucoxanthin-chlorophyll protein of diatoms. Photosynth Res., 86(1-2):241-250. doi:10.1007/s11120-005-1003-8

Penner R., (2020). Backbone Free Energy Estimator Applied to Viral Glycoproteins. Journal of Computational Biology. doi: 10.1089/cmb.2020.0120.

Pi X, Zhao S, Wang W, et al., (2019). The pigment-protein network of a diatom photosystem II-light-harvesting antenna supercomplex. Science, 365(6452):eaax4406. doi:10.1126/science.aax4406

Pudney A., Gandini C., Economou C.K. et al., (2019). Multifunctionalizing the marine diatom Phaeodactylum tricornutum for sustainable co-production of omega-3 long chain polyunsaturated fatty acids and recombinant phytase. Sci Rep 9, 11444. https://doi.org/10.1038/s41598-019-47875-1

Pyszniac A.M., Gibbs S.P., (1992). Immunocytochemical localization of photosystem I and the fucoxanthin-chlorophyll a/c light-harvesting complex in the diatom Phaeodactylum tricornutum. Protoplasma 166, 208-217. doi.org/10.1007/BF01322783

Qin X., Pi X., Wang W. et al., (2019). Structure of a green algal photosystem I in complex with a large number of light-harvesting complex I subunits. Nat. Plants 5, 263–272.doi.org/10.1038/s41477-019-0379-y

Ramachandran G.N., Ramakrishnan C., Sasisekharan V., (1963). Stereochemistry of polypeptide chain configurations. Journal of Molecular Biology. 7: 95–9. doi:10.1016/S0022-2836(63)80023-6

Roberta Croce R., van Amerongen H., (2020). Light harvesting in oxygenic photosynthesis: Structural biology meets spectroscopy, Science, Vol. 369, Issue 6506, eaay2058. DOI: 10.1126/science.aay2058

Schaller-Laudel S., Latowski D., Jemioła-Rzemińska M., Strzałka K., Daum S., Bacia K., Wilhelm C., Goss, R., (2017). Influence of thylakoid membrane lipids on the structure of aggregated light-harvesting complexes of the diatom Thalassiosira pseudonana and the green alga Mantoniella squamata. Physiol Plantarum, 160: 339-358. doi:10.1111/ppl.12565

Su X., Ma J., Pan X. et al., (2019). Antenna arrangement and energy transfer pathways of a green algal photosystem-I–LHCI supercomplex. Nat. Plants 5, 273–281. doi.org/10.1038/s41477-019-0380-5

Taddei L, Stella GR, Rogato A, et al., (2016). Multisignal control of expression of the LHCX protein family in the marine diatom Phaeodactylum tricornutum. J Exp Bot., 67(13):3939-3951. doi:10.1093/jxb/erw198

Van Eerden F.J., de Jong D.H., de Vries A.H., Wassenaar T.A., Marrink S.J., (2015). Characterization of thylakoid lipid membranes from cyanobacteria and higher plants by

molecular dynamics simulations, Biochim. Biophys. Acta - Biomembranes, 1848, pp. 1319-1330

Van Zundert GCP, Rodrigues JPGLM, Trellet M, et al. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J Mol Biol. 2016;428(4):720-725. doi:10.1016/j.jmb.2015.09.014

Wang W, Yu LJ, Xu C, et al., (2019). Structural basis for blue-green light harvesting and energy dissipation in diatoms. Science, vol. 363, issue 6427, eaav0365. doi:10.1126/science.aav0365

Wilhelm C., Manns L., (1991). Changes in pigmentation of phytoplankton species during growth and stationary phase-consequences for reliability of pigment-based methods of biomass determination. J Appl Phycol 3, 305–310. doi.org/10.1007/BF00026092

Zhu S.H., Green B.R., (2010). Photoprotection in the diatom Thalassiosira pseudonana: Role of LI818-like proteins in response to high light stress. Biochim Biophys Acta 1797:1449–1457. doi.org/10.1016/j.bbabio.2010.04.003